Non-Exercise Activity Thermogenesis Monitoring Using Wearable Technology

A Thesis Submitted

in Partial Fulfilment of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

by

Ankita Dewan

(2017csz0012)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY ROPAR

JANUARY, 2025

Dedicated to my family !!

Declaration of Originality

I hereby declare that the work which is being presented in the thesis entitled **Non-Exercise Activity Thermogenesis Monitoring Using Wearable Technology** has been solely authored by me. It presents the result of my own independent investigation/research conducted during the time period from January 2018 to September 2023 under the supervision of Dr. Venkata M. V. Gunturi at IIT Ropar. To the best of my knowledge, it is an original work, both in terms of research content and narrative, and has not been submitted or accepted elsewhere, in part or in full, for the award of any degree, diploma, fellowship, associateship, or similar title of any university or institution. Further, due credit has been attributed to the relevant state-of-the-art and collaborations (if any) with appropriate citations and acknowledgments, in line with established ethical norms and practices. I also declare that any idea/data/fact/source stated in my thesis has not been fabricated/falsified/misrepresented. All the principles of academic honesty and integrity have been followed. I fully understand that if the thesis is found to be unoriginal, fabricated, or plagiarized, the Institute reserves the right to withdraw the thesis from its archive and revoke the associated Degree conferred. Additionally, the Institute also reserves the right to appraise all concerned sections of society of the matter for their information and necessary action (if any). If accepted, I hereby consent for my thesis to be available online in the Institute's Open Access repository, inter-library loan, and the title & abstract to be made available to outside organizations.

Ankita.

Signature

Ankita Dewan
2017csz0012
PhD
Department of CSE
Indian Institute of Technology Ropar
Rupnagar, Punjab 140001

Date: 22 January, 2025

Acknowledgement

First of all, I begin by expressing my deepest gratitude to almighty God, whose unending guidance and grace have sustained me throughout this entire journey of PhD. I also want to thank IIT Ropar for providing me with the platform and all the facilities required for completing my PhD.

I would like to thank my supervisor, Dr. Venkata M. V. Gunturi, for the patient guidance, encouragement, and advice he has provided throughout my PhD study. I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my questions and queries so promptly. I owe him way more than my words could ever express.

I must extend my special gratitude to Dr. Vinayak Naik for being an extraordinary mentor and collaborator. I have learned a lot of things from him throughout my journey of PhD. His willingness to share expert insights have significantly contributed to my both professional and personal growth as a researcher.

I am immensely grateful to all of my doctoral committee members: Dr. Abhinav Dhall, Dr. Mukesh Saini, Dr. Himanshu Paliwal, Dr. Suman Kumar, Dr. Basant Subba and Dr. Apurva Mudgal for evaluating my research work periodically and their support. I am indebted to IIT Ropar for providing me with all the necessary facilities for this research.

Very special gratitude goes out to Kousik Kumar Dutta, Shivam Gupta, Anil, Priyankar, Vivek, Surbhi, Akanksha, Aroof, Jaspinder, Neeru, Amit, Shubhada, Upkar, Anamika Chabhra, Taran, Gopika, Versha, Birender, Vansha, Neha and Ankit. I also place on record, my sense of gratitude to one and all, who directly or indirectly, have assisted me.

Last but not least, I must express my very profound gratitude to my husband and family for providing me with unfailing support and continuous encouragement throughout my years of study.

This achievement would not have been possible without all of you. Thank you for being a part of my PhD journey.

– Ankita Dewan

Certificate

This is to certify that the thesis entitled **Non-Exercise Activity Thermogenesis Monitoring Using Wearable Technology**, submitted by **Ankita Dewan (2017csz0012)** for the award of the degree of **Doctor of Philosophy** of Indian Institute of Technology Ropar, is a record of bonafide research work carried out under our guidance and supervision. To the best of my knowledge and belief, the work presented in this thesis is original and has not been submitted, either in part or full, for the award of any other degree, diploma, fellowship, associateship or similar title of any university or institution. In our opinion, the thesis has reached the standard fulfilling the requirements of the regulations relating to the Degree.

GViswanath

Signature of the Supervisor
Dr. Venkata M. V. Gunturi
Department of CSE
Indian Institute of Technology Ropar
Rupnagar, Punjab 140001
Date: 22 January 2025

Lay Summary

Our daily movements, even small ones like cooking, cleaning, or simply pacing around the house, play a significant role in our health. These small actions, known as NEAT (Non-Exercise Activity Thermogenesis), are important for maintaining a healthy lifestyle. This research focuses on using smartwatches to recognize and differentiate these activities in a simple, practical, and affordable way. Most systems that track activities rely on complex setups or high-frequency data collection, which can be inconvenient or drain a device's battery quickly. Our work addresses this by developing smart models that work efficiently with less data, making them more suitable for everyday use.

In the first part of this study, we designed a model that could recognize common household activities, like sweeping or resting, using only basic data from a smartwatch. This was a big step towards creating a system that is easy to use and reliable. We then expanded this system to include more activities, covering both active tasks and quiet moments. The smartwatch gathers information from the wearer and processes it to identify what they are doing. To handle situations where activities don't fit into neat categories, we created a flexible option to classify these moments as "others" making the system adaptable to real life. Finally, we made the system even smarter and more independent. By using advanced technology, we created a solution that doesn't rely on Wi-Fi or other devices to work. This makes it convenient for homes where internet connectivity might be limited. We also ensured that the technology works smoothly on a smartwatch without draining the battery too quickly.

This research shows that smartwatches can do much more than count steps—they can help us understand our everyday movements in detail. By recognizing and tracking these activities, this technology could help people stay healthier, improve their fitness, and gain valuable insights into their daily habits.

Abstract

Our thesis explores the complex task of distinguishing NEAT (Non-Exercise Activity Thermogenesis) activities from non-NEAT activities in a home environment using data from wearable smartwatch sensors. It presents a multifaceted problem where low-frequency sensor data is used to differentiate thirteen distinct household activities. Existing research often prioritizes high-frequency data or multiple sensors, overlooking fundamental home activities. To address these limitations, this research introduces innovative AI models, capable of achieving superior accuracy while working on low frequency. This dissertation unfolds in three key steps, each contributing to the overall understanding and advancement of NEAT activity recognition.

In our first work (Chapter Three), we introduced the Hierarchical Model, tailored to differentiate seven home activities using low-frequency (1Hz) sensor data. Our experiments revealed the model's remarkable accuracy, outperforming traditional flat models like XGB. Even when the sampling frequency was reduced to 1Hz, the Hierarchical Model maintained substantial accuracy. These results provided a foundation for recognizing NEAT activities using wearable technology, highlighting the potential of our approach.

In our second work (Chapter Four) we delved deeper into the system's development. The system, worn on a smartwatch, identifies and classifies 13 distinct activities, including both physical and sedentary actions. The user initiates data collection, which is transmitted to a central server for interpretation. Key parameters, such as battery depletion rate and data sampling rate, were evaluated, with the goal of future on-device deployment. Notably, our system accommodates unclassified activities through the introduction of the "OTHERS" class, a versatile approach that can be adjusted based on desired strictness or leniency.

In our last work (Chapter Five) we tackled the challenge of distinguishing household activities using low-frequency data without relying on external connectivity. Our innovative Hybrid Model achieved superior accuracy compared to other models, making it more user-friendly in environments without readily available Wi-Fi or paired devices. The choice of a 10-second window length was found to balance accuracy and battery efficiency. Additionally, the deployment of neural networks, including 1d-CNN, LSTM, and Bidirectional LSTM, allowed us to propose a Hybrid Model, integrating TensorFlow Lite (TFLite) for improved performance and efficiency on the smartwatch.

This dissertation provides a comprehensive understanding of NEAT activity recognition using wearable technology in a home environment. From the introduction of the Hierarchical Model to the development of the system and the implementation of neural networks, our findings pave the way for future advancements in the field. Distinguishing NEAT and non-NEAT activities has the potential to revolutionize how we monitor and understand daily activities, impacting areas like healthcare, fitness, and beyond.

Keywords: NEAT (Non-Exercise Activity Thermogenesis), Wearable Technology, Low-Frequency Sensor Data, Activity Recognition

List of Publications

Journal

1. **Dewan, A.**, Gunturi, V. M., & Naik, V. (2023, January), "NEAT Activity Detection using Smartwatch," in *International Journal of Ad Hoc and Ubiquitous Computing. Inderscience Publishers.*

Conferences

1. **Dewan, A.**, Gunturi, V. M., Naik, (2023, November) "Deep Learning Models for NEAT Activity Detection on Smartwatch" in *In proceedings of The 20th IEEE International Conference on Ubiquitous Intelligence and Computing*
2. **Dewan, A.**, Gunturi, V. M., Naik, V., & Dutta, K. K. (2021, October). "NEAT Activity Detection using Smartwatch at Low Sampling Frequency," in *2021 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Internet of People and Smart City Innovation*, (SmartWorld/SCALCOM/UIC/ATC/IOP/SCI) (pp. 25-32). IEEE.
3. **Dewan, A.**, Gunturi, V. M., Naik, V., Vishwakarma, K., & Bohra, S. (2019, December). "A hierarchical classifier for detecting metro-journey activities in data sampled at low frequency." In *Proceedings of the 17th International Conference on Advances in Mobile Computing & Multimedia* (pp. 46-55).

Contents

Declaration **iv**

Acknowledgement **v**

Certificate **vi**

Lay Summary **vii**

Abstract **viii**

List of Publications **ix**

List of Figures **xv**

List of Tables **xvii**

1 Introduction **1**

1.1 Neat / non-NEAT 2

1.1.1 Activities 4

1.2 Research Gaps 13

1.3 Research Questions 14

1.4 Objective Of Thesis 14

1.5 Contribution Of Thesis 15

1.6 Significance of NEAT Monitoring Devices 16

1.7 Organisation of thesis 17

2 Data Collection and Deployment for NEAT Monitoring **21**

2.1 Data Collection 21

2.1.1 Important Elements of data collection 21

2.1.2 Data collection using smartwatch 23

2.2 Factors considered in building NEAT monitoring device 24

2.2.1 Device Configuration 24

2.2.2 Choice of sensors 24

2.2.3 Sensor frequency selection 26

2.2.4 Deployment in smartwatch 27

3 Related Work **31**

3.1 Activity Detection 31

3.2 Usage of Multiple Sensors 32

3.3 Usage of high energy consuming sensors like Heart Rate and GPS 35

3.4 Usage of High Frequency data . . . 38
3.5 Latest research trends . . . 42
3.6 Key limitation and challenges of Related Work . . . 43

4 NEAT Activity Detection using Smartwatch at Low Sampling Frequency 45
4.1 Problem Scope . . . 45
4.1.1 Challenges . . . 46
4.2 Scope of chapter . . . 46
4.3 Contributions . . . 47
4.4 Our Proposed Approach . . . 47
4.4.1 Inspiration . . . 47
4.4.2 Preprocessing and Features . . . 48
4.4.3 Proposed Hierarchical model . . . 49
4.4.4 Model Selection Rationale . . . 50
4.5 Experimental Analysis . . . 51
4.5.1 Candidate Algorithms . . . 52
4.5.2 Training & evaluation matrix . . . 52
4.5.3 Consequences of varying window length on final accuracy . . . 52
4.5.4 F1-scores of individual classes . . . 54
4.6 Conclusion and Limitations . . . 56

5 NEAT Activity Detection using Smartwatch 59
5.1 Problem Scope . . . 59
5.2 Scope of chapter . . . 60
5.3 Contributions . . . 60
5.4 Proposed architecture . . . 61
5.4.1 Detection of others classes . . . 62
5.5 Experimentation . . . 63
5.5.1 Candidate Algorithms: . . . 63
5.5.2 Experimental Goals . . . 63
5.5.3 Comparison of Feature selection algorithm - TsFresh vs ECDF vs Static Features . . . 68
5.5.4 Real Time Testing . . . 69
5.5.5 Key Results . . . 71
5.6 Conclusion and Limitations . . . 74

6 Neural Networks for NEAT Activity Detection on Smartwatch 77
6.1 Problem Scope . . . 77
6.2 Scope of chapter . . . 77
6.3 Contributions . . . 77
6.4 Candidate neural network classifiers . . . 78

6.4.1 **1D-CNN** . . . 78
6.4.2 **ANN** . . . 79
6.4.3 **LSTM** . . . 80
6.4.4 **BiDirectional-LSTM** . . . 81
6.4.5 **Our Hybrid Model** . . . 82
6.5 Experimentation . . . 84
6.5.1 Effect of varying window length on total Accuracy & battery consumption . . . 85
6.6 Conclusion and Limitations . . . 87

7 Conclusions and Future Work **89**
7.1 Research Conclusion . . . 89
7.2 Contributions to Existing Knowledge and Distinctiveness of This Research . 91
7.2.1 Contribution to Existing Knowledge . . . 92
7.2.2 Distinctiveness of This Research . . . 92
7.3 Limitations of the research . . . 93
7.4 Future work . . . 95
7.4.1 Further battery optimisation . . . 95
7.4.2 Integrating Other Types of Sensors . . . 97
7.4.3 Applying Models in Different Environments . . . 97
7.4.4 Domain Adaptation . . . 97
7.4.5 Use of Functional API: . . . 98

References **101**

List of Figures

1.1 Total Daily Energy Expenditure . . . 1
1.2 Total Daily Energy Expenditure (TDEE) . . . 3
1.3 Cooking . . . 5
1.4 Sweeping . . . 6
1.5 Mopping . . . 7
1.6 Eating . . . 9
1.7 Driving . . . 10

2.1 Fossil Sport Smartwatch 43mm - FTW4019 , Image credits : Amazon.com . 24

4.1 Proposed hierarchical model for distinguishing the NEAT and non-NEAT activities. . . . 49
4.2 Effect of window length on overall accuracy when all features are used. . . . 53
4.3 Effect of window length on overall accuracy when only accelerometer or gyroscope features are used. . . . 54

5.1 Architecture for Testing on Live Data. . . . 62
5.2 Architecture for Data Sampling . . . 63
5.3 Rate of Battery Depletion while using ECDF features as compared to the Static features. The X-axis represents the time where 1 unit represents 10 seconds, and the Y-axis represents the Battery left (mAh) in the smartwatch. 64
5.4 Rate of Battery Depletion while taking different window lengths, such as 2 sec, 4 sec, and 6 seconds. The X-axis represents the time where 1 unit represents 10 seconds, and the Y-axis represents the Battery left (mAh) in the smartwatch. . . . 65
5.5 Rate of Battery Depletion while using 1Hz sampling rate as compared to 10Hz sampling rate. The X-axis represents the time where 1 unit represents 10 seconds, and the Y-axis represents the Battery left (mAh) in the smartwatch. . . . 66
5.6 Rate of Battery Depletion while varying the file upload rate to the server. The X-axis represents the time where 1 unit represents 10 seconds, and the Y-axis represents the Battery left (mAh) in the smartwatch . . . 66
5.7 Effect of different overlapping windows percentage on overall accuracy for 5Hz frequency and ECDF Features. . . . 67
5.8 Effect of different overlapping windows percentage on overall accuracy for 5Hz frequency and Static Features. . . . 67
5.9 Effect of different overlapping windows percentage on overall accuracy for 10Hz frequency and ECDF Features. . . . 67

5.10 Effect of different overlapping windows percentage on overall accuracy for 10Hz frequency and Static Features. 68
5.11 Comparison of Tsfresh vs Static vs ECDF Features selection algorithms w.r.t. Accuracies and Computation time 70
5.12 Live Data Accuracy for all activities 71

6.1 1D-CNN 79
6.2 ANN 80
6.3 LSTM 81
6.4 Bidirectional LSTM 82
6.5 Our Proposed Hybrid Model 83
6.6 Accuracy and Battery Depletion Rate of neural networks 85
6.7 Comparison of Model Accuracy at Different Data Sampling Frequencies (1 Hz vs 10 Hz) 87

7.1 Sequential API 98
7.2 Functional API 99

List of Tables

2.1 Advantages and Disadvantages of High vs. Low Sampling Rate Sensors . . 27
2.2 **Server Deployment and Proposed Solution** 30

4.1 Total Battery consumed in 5 hours 45
4.2 Accuracy of different classifiers on 1Hz, 5Hz, and 10Hz sampled data (2 sec windows 50% overlap.) 46
4.3 Experimenting on various known classifiers by selecting multiple combinations for classes. Class 1 - Cooking, Class 2 - Sweeping, Class 3 - Mopping, Class 4 - Walking, Class 5 - Climb up, Class 6 - Climb Down, Class 7 - non-NEAT 51
4.4 Total Training Data 51
4.5 F1-scores of Our Hierarchical Model and Flat Classifier XGBoost (XGB) for $\theta = 0.50$ and frequency = 1Hz and window length = 2seconds 53
4.6 F1-scores of Our Hierarchical Model and Flat Classifier XGBoost (XGB) for $\theta = 0.50$, frequency = 1Hz, window = 4 seconds 54
4.7 F1-scores of Our Hierarchical Model and Flat Classifier XGBoost (XGB) for $\theta = 0$, frequency = 1Hz, window = 2 seconds 55
4.8 F1-scores of Our Hierarchical Model and Flat Classifier XGBoost (XGB) for $\theta = 0$, frequency = 1Hz, window = 4 seconds 55
4.9 F1-scores of Our Hierarchical Model and Flat Classifier XGBoost (XGB) for $\theta = 0.50$, frequency = 5Hz, window = 2 seconds 55
4.10 F1-scores of Our Hierarchical Model and Flat Classifier XGBoost (XGB) for $\theta = 0.50$, frequency = 5Hz, window = 4 seconds 56

5.1 Accuracy of different classifiers (along with the standard deviation after 5 runs) on 1Hz, 5Hz, and 10Hz sampled data for 2-sec window length and 0% overlap.) 60
5.2 Key Results 61
5.3 F1-scores and overall accuracy from the Classifiers for ECDF vs Static for overlapping $\theta = 0$ and window length = 2 seconds. 69
5.4 F1-scores and overall accuracy from the Classifiers for ECDF vs Static for overlapping $\theta = 0.5$ and window length = 2 seconds 70
5.5 Key Results 72

6.1 Total Training Data 84
6.2 F1-Scores of Different Classifiers for Each Activity for 3 seconds window length 86
6.3 F1-Scores of Hybrid Model for Different Window Lengths for Each Activity 86

7.1 NEAT Activity Detection System Objectives and Fulfillment 94

Chapter 1

Introduction

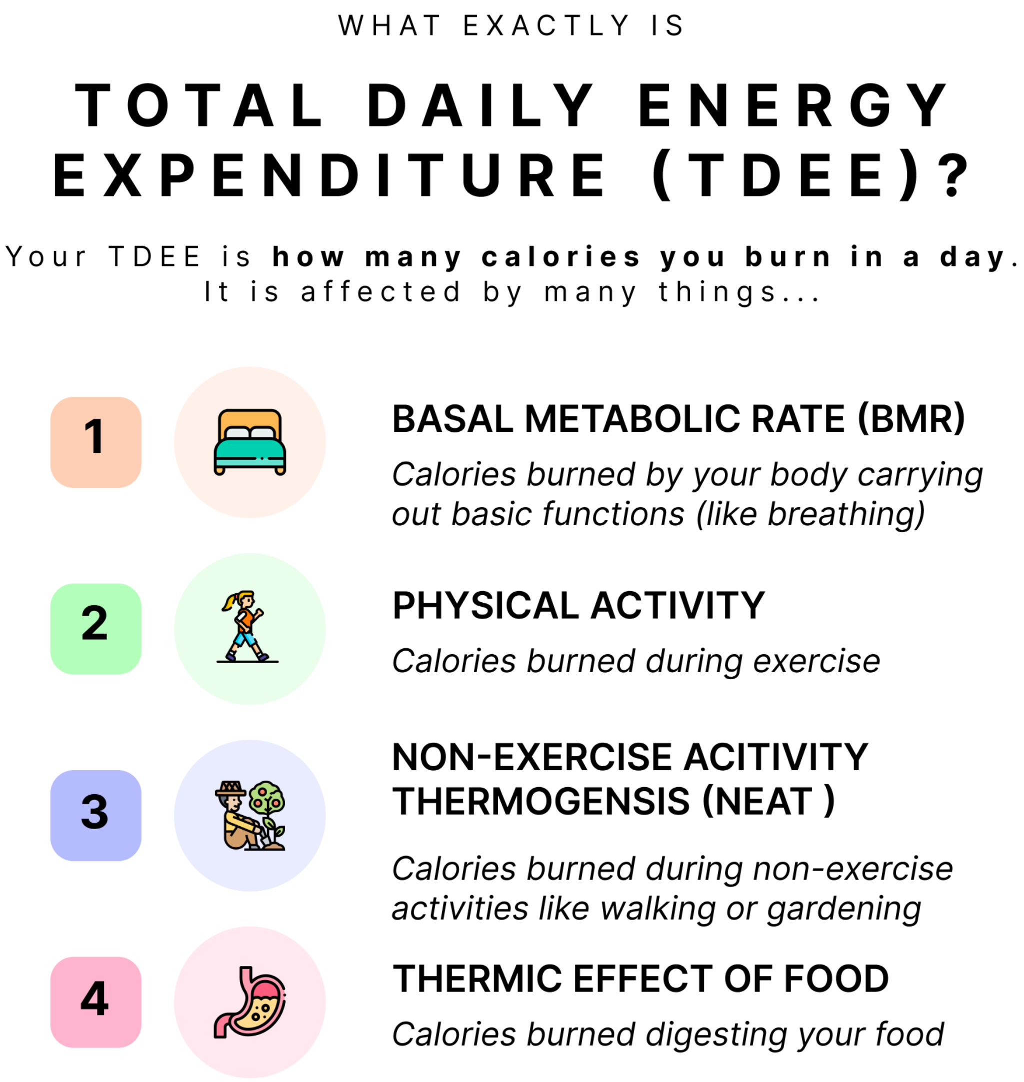


Figure 1.1: Total Daily Energy Expenditure

In our fast-paced world, understanding human activity has become increasingly important for promoting overall health and well-being depicted in Figure 1.1. Human activity recognition has emerged as a pivotal area of study, shedding light on how we go about our daily lives. One critical aspect of this research is the distinction among Non-Exercise Activity Thermogenesis (NEAT) and non-NEAT activities. NEAT encompasses the energy we expend in our day-to-day tasks, excluding exercise, while non-NEAT activities encompass everything else we do.

The advent of wearable technology, particularly smartwatches, has opened up new avenues for monitoring these activities. Smartwatches, often equipped with multiple sensors, have the potential to provide valuable insights into the NEAT and non-NEAT activities of individuals. These sensors, though typically low-end compared to specialized equipment, can offer a cost-effective and convenient way to capture and analyze human activity patterns.

This thesis explores human activity recognition, with the use of cutting-edge technologies like Mobiquitous computing. The Mobiquitous computing paradigm leverages the power of mobile and ubiquitous devices to integrate data collection and analysis into our daily lives seamlessly. This approach paves the way for more efficient and accurate activity recognition methods.

The primary objective of this thesis is to demonstrate the effectiveness of smartwatch sensors in activity recognition, despite challenges such as limited battery life and restricted hardware capabilities including storage capacity and computation power. We aim to showcase that, from the data collected through smartwatch sensors, our proposed hierarchical model outperforms conventional flat models such as MLP, SVM, RF, and XGB. Later in our thesis, we strive to balance accuracy with battery life in smartwatch devices. Recognizing that users have different priorities—some favoring longer battery life and others prioritizing higher accuracy in activity tracking—we provide comparisons to allow users to choose options that best suit their needs. Additionally, our research focuses on delivering real-time feedback to users using neural networks such as 1d-CNN, LSTM, Bi-directional LSTM and our proposed hybrid model all of which are embedded in their smartwatches. These models have been used to enhance security and perform processing entirely on the watch. Through these innovations, we aim to contribute meaningfully to the evolving field of wearable technology, with a specific focus on enhancing human health and well-being.

1.1 Neat / non-NEAT

NEAT (Non-Exercise Activity Thermogenesis):

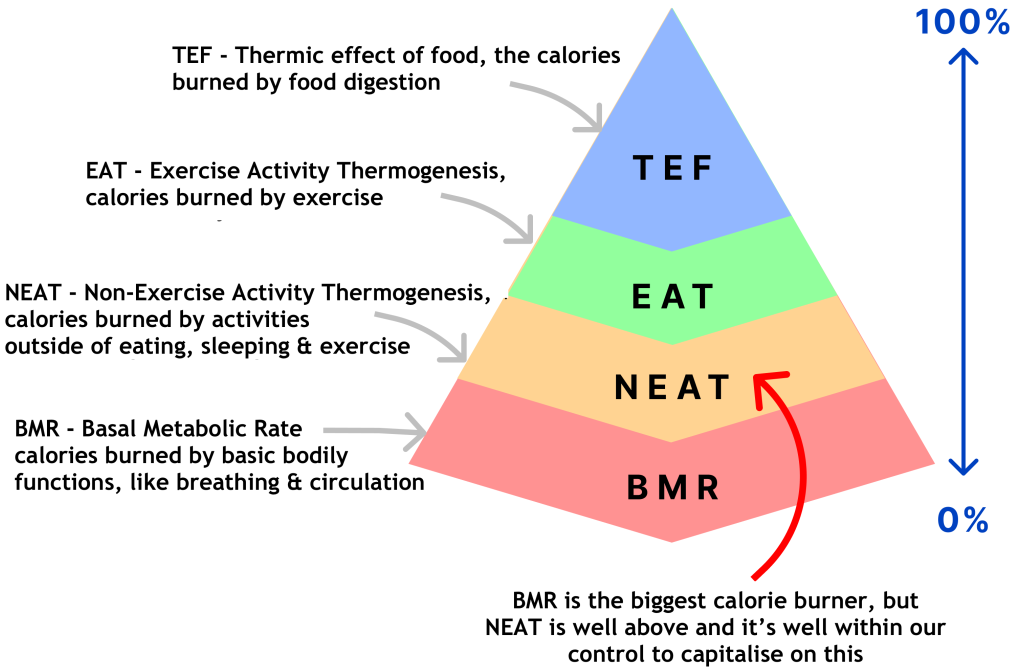


Figure 1.2: Total Daily Energy Expenditure (TDEE)

NEAT encompasses a wide spectrum of activities that contribute significantly to the energy expenditure of our bodies depicted in Figure 1.2. These activities occur spontaneously throughout the day and are not part of planned exercise routines. NEAT includes actions such as walking, climbing stairs, fidgeting, standing, and even tasks like gardening or housework. These movements may seem inconsequential on their own, but collectively, they account for a substantial portion of our daily calorie burn. NEAT is an integral component of our overall daily energy expenditure and plays a vital role in regulating body weight and metabolic health.

non-NEAT:

non-NEAT activities, in contrast, refer to physical exercises that are structured, planned, and intentional. These activities are purposefully performed to enhance fitness, build strength, or achieve specific health objectives. Examples of non-NEAT activities encompass activities such as running, weightlifting, cycling, swimming, and participation in organized sports. Unlike NEAT, which occurs naturally as part of daily life, non-NEAT activities are typically scheduled and carried out with the explicit goal of improving physical health and fitness.

In summary, NEAT and non-NEAT represent two distinct categories of physical activity. NEAT captures the myriad of everyday movements we make without conscious effort, while non-NEAT activities are planned exercises designed to target specific health and fitness goals. Both aspects of physical activity are essential, and wearable technology plays a crucial role in monitoring and quantifying these activities. Understanding and analyzing

NEAT and non-NEAT activities through wearable technology have profound implications for health and well-being, as they offer insights into how our daily choices and routines impact our overall health, metabolism, and fitness levels.

1.1.1 Activities

Categorization of Activities

We choose multiple activities that cater to a wide range of human behaviours. Studying these activities can help garner a deeper understanding of human daily routines and, in turn, help understand NEAT better. All these chosen activities cover most of the activities that we could perform during intermittent curfew due to the COVID-19 pandemic. Below is the categorisation of activities we choose to incorporate into our research.

1. **Daily Movements (walking, climbing up and down, eating):** Such activities are part of everyone's daily routine. Though they are common occurring activities still can be differentiated as the intensity and frequency with which they are performed can be varied. Hence, by monitoring daily movements we can set a benchmark for baseline energy consumption throughout the day.

2. **Occupational Activities (cooking, sweeping, mopping, working on a laptop):** These activities are heavily dependent on a person's job and household activities. Activities like cooking and cleaning require different levels of physical exertion, whereas working on a laptop is more of a sedentary activity. Both forms of physical activity impact NEAT, particularly highlighting the energy expenditure.

3. **Leisure Activities (watching TV, browsing on phone):** These activities though being low on energy expenditure are still critical for monitoring and understanding NEAT. These activities account to a major portion of person's time spent during a day, specially in this modern technology-driven lifestyle. Monitoring these help in understanding how much time is spent in low-energy vs more active leisure activities.

4. **Transportation Activities (cycling, driving, sitting in a car):** These activities have a vast range of energy expenditure, ranging from a high-energy activity like cycling to a low-energy activity like sitting in a car. Analysing these activities provides a comprehensive understanding of how transportation activities contribute to NEAT monitoring.

These activities were chosen, citing their frequency of occurrence and the similar hand or body movement in which the activity is conducted. All these activities are commonly performed in a person's daily routine and are susceptible to behaviour modifications. These can be easily utilised in public health interventions, making them ideal for studying NEAT. This can further help enhance and build effective public health strategies.

Formal Definition of Activities

1. **Cooking:**

Figure 1.3: Cooking

Cooking is a common daily activity that involves various movements and gestures in the kitchen, such as chopping ingredients, stirring, flipping, and moving pots and pans. Monitoring this activity can provide valuable insights into an individual's daily energy expenditure and overall physical activity levels. Smartwatches equipped with sensors like accelerometers and gyroscopes can play a pivotal role in detecting the activity of cooking.

When you are cooking, you perform various arm and hand movements, including lifting, shaking, and stirring. These movements generate accelerations in different directions. By analyzing the patterns of acceleration, the smartwatch can recognize the specific movements associated with cooking. For example, the accelerometer can detect the repetitive up-and-down motion of chopping vegetables or the back-and-forth motion of stirring a pot.

When cooking, you often make precise and deliberate movements with your hands and arms, such as flipping a chapati or turning a spatula. The gyroscope can capture the rotational dynamics of these movements. By analyzing the gyroscope data, the smartwatch can identify the characteristic rotations and orientations associated with cooking activities.

Data from both the accelerometer and gyroscope allows the smartwatch to create a more accurate and comprehensive picture of the cooking activity. Using these sensors and machine learning algorithms, a smartwatch on the dominant hand can effectively monitor cooking activities and provide insights into Non-Exercise activity thermogenesis. This information can be valuable for individuals interested in tracking their daily physical activity and calorie expenditure, as well as for researchers studying the impact of cooking on energy balance and overall health.

2. **Sweeping:**

Figure 1.4: Sweeping

Sweeping is a common daily activity. Specifically in India, that involves using a broom or similar tool to clean and remove dust, dirt, or debris from floors, surfaces, or outdoor areas. It requires a combination of repetitive arm and hand movements along with body motion as the individual sweeps back and forth.

When a person is sweeping, they engage in repetitive back-and-forth arm movements. The accelerometer in the smartwatch can detect these rhythmic motions by tracking changes in acceleration along the wrist's x and y axes. The sensor will pick up on the continuous oscillations, indicative of the sweeping activity. Similarly while sweeping, there are subtle but consistent rotations and changes in wrist orientation as the person moves their arm and hand in different directions. The gyroscope can detect these rotational movements and changes in orientation.

When accelerometer and gyroscope data are combined, a smartwatch can recognize the unique motion patterns associated with sweeping. Machine learning algorithms can be trained on this sensor data to distinguish sweeping from other activities. For example, the algorithm can be trained to recognize the specific frequency, duration, and direction of the wrist movements characteristic of sweeping.

3. **Wet mopping:**

Figure 1.5: Mopping

In India, Wet mopping is a physical activity that involves cleaning a surface, typically a floor, using a damp mop or cloth to remove dirt, stains, and debris. It has a series of complex activities, preparation of cleaning solutions, saturating the mop stick in the solution, the actual physical scrubbing during wet mopping, and the removal of debris. It requires a combination of repetitive arm and hand movements, as well as body movements such as bending, swaying the mop stick and walking around the cleaning area. The effectiveness of wet mopping largely depends on the intensity and duration of these movements, making it an interesting activity to monitor for various purposes, including fitness tracking and household management.

Accelerometers in smartwatches measure changes in acceleration, which can be used to track the rhythmic back-and-forth or side-to-side movements typically associated with mopping. These sensors can detect the oscillatory patterns of arm and hand movements, helping to identify when someone is engaged in wet mopping.

Gyroscopes, on the other hand, measure angular velocity and orientation. When combined with accelerometers, they can provide data on the orientation and angle of the smartwatch relative to the Earth's gravity. This information can be used to distinguish between mopping activities and other daily activities like walking or sitting. Wet mopping often involves movements that are distinct from typical wrist movements during other activities.

4. **Walking:** Walking is a fundamental human activity characterized by the rhythmic movement of one's legs to propel the body forward. It is a low-impact aerobic exercise that is not only a common mode of transportation but also a popular form of physical activity for maintaining health and fitness. Monitoring walking can provide valuable insights into an individual's daily physical activity levels and overall well-being.

   Smartwatches can register changes in acceleration when you walk, your body experiences rhythmic up-and-down movements, which are detected by the accelerometer. As you take a step, the accelerometer senses the acceleration when your foot strikes the ground and the deceleration when it leaves the ground. These repetitive patterns of acceleration and deceleration are indicative of walking.

   Gyroscope measures angular velocity or orientation changes. While walking, your wrist and arm experience a specific pattern of rotation as you swing them back and forth. This rotational movement is detected by the gyroscope and can be used to further confirm the walking activity. Gyroscope data helps differentiate between walking and other activities with similar accelerations, like cycling or running.

   By combining data from both the accelerometer and gyroscope, a smartwatch can accurately detect walking and provide valuable information on step count, calorie expenditure, cadence and pace, activity tracking, and health monitoring. This information is valuable for promoting physical activity, tracking fitness goals, and assessing overall health and well-being. It is a crucial part of research and applications related to Non-Exercise activity thermogenesis monitoring using wearable technology.

5. **Climb up/Down:** Climbing up and down, whether it's on stairs, hills, or any elevated surface, is a common physical activity with a notable impact on an individual's energy expenditure. Wearable technology, like smartwatches equipped with accelerometers and gyroscopes, offers insights into daily physical activity and energy expenditure.

   Vertical movement during climbing involves ascending or descending steps, stairs, or inclines, engaging muscle groups such as the legs, glutes, and core. It typically demands more effort and energy than walking on flat ground.

   Smartwatch accelerometers measure linear motion changes, detecting upward and downward movements related to climbing. As an individual climbs or descents, the accelerometer registers each step's acceleration impact. Gyroscopes measure angular velocity and rotation, distinguishing various movements and orientation changes. For instance, during ascent, the gyroscope can detect wrist tilting indicative of upward hand motion, while descent reveals the opposite tilt as the hand moves downward.

By combining data from both the accelerometer and gyroscope, smartwatches accurately distinguish between climbing up and climbing down activities. These sensors quantify duration, intensity, and frequency of these movements, enabling users and researchers to track physical activity levels and estimate energy expenditure. This information is invaluable for fitness tracking, health monitoring, and research on physical activity patterns.

6. **Eating:**

Figure 1.6: Eating

Eating is a fundamental human activity involving the consumption of food and beverages. It encompasses hand-to-mouth actions, as in India that is the most prevalent way of eating. Monitoring eating habits is crucial for various health and lifestyle applications, including dietary assessments, weight management, and understanding eating disorders. Accelerometers in smartwatches can detect subtle hand movements associated with bringing food or hand / utensils to the mouth. These sensors are sensitive to changes in motion and can recognize the repeated pattern of raising one's hand toward the mouth during eating. Gyroscopes can be used to detect the rotational movements of the wrist and forearm during the chewing and swallowing phases of eating. By analyzing the data from both accelerometers and gyroscopes, smartwatches can track the duration of the Eating activity by continuously monitoring the hand-to-mouth gestures and wrist movements. Moreover, they can provide real-time feedback to users about their eating habits, such as suggesting slower eating to aid digestion or promoting healthier eating patterns. Additionally, long-term data analysis can help individuals and healthcare professionals identify trends and make informed decisions about diet and nutrition.

7. **Driving:**

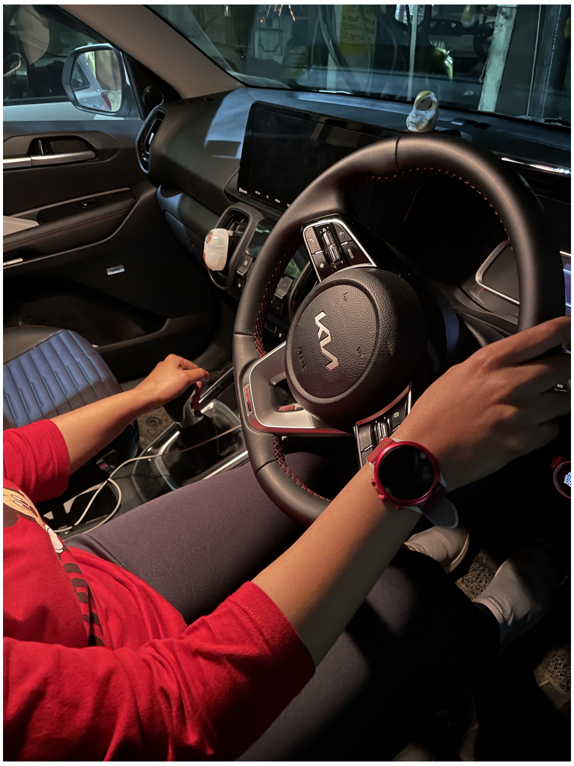


Figure 1.7: Driving

Driving in India, where right-hand driving is the norm, presents a unique context for monitoring this non-exercise activity using wearable sensor technology. With the increasing prevalence of automatic vehicles and the dominant use of the dominant hand for vehicle controls, wearable sensors can be strategically employed to gather valuable data. Driving is a common non-exercise activity that involves operating a vehicle to move from one location to another. It is a fundamental part of modern life for many individuals in India, whether for daily commuting, running errands, or leisure travel. Monitoring driving as a Non-Exercise activity can provide valuable insights into an individual's daily routine and overall health. Smartwatch sensors can play a pivotal role in detecting this activity by utilizing various technologies and data inputs tailored to the Indian driving scenario. These sensors are equipped to capture and interpret specific aspects of driving. Smartwatches are equipped with accelerometers that measure changes in motion and acceleration. When a person is driving, the smartwatch can detect the vibrations and movements associated with the vehicle's motion, considering the right-hand control of the vehicle in India. The accelerometer data, in conjunction with the gyroscope, can provide a comprehensive picture of the wearer's wrist orientation, angular velocity, and hand movements during driving. By training the smartwatch to recognize specific driving patterns associated with right-hand control, such as paddle shifts, steering wheel movements, and hand gestures, the accuracy of activity detection can be significantly improved. Overall, smartwatches and wearable technology have the potential to accurately detect Non-Exercise driving activity in India's right-hand driving environment. These technologies, when tailored to the local driving context, can provide valuable insights into daily routines, stress levels, and promote safer driving habits among individuals engaged in the act of driving.

8. **Working on a laptop:** Working on a laptop is a sedentary activity that involves individuals using a laptop or computer for various purposes such as office work, studying, web browsing, or entertainment. It typically requires a person to sit at a desk or in a stationary position for an extended period, often involving tasks that require typing, using a mouse, or engaging with the laptop's screen. While this activity is essential in today's digital age, it is associated with prolonged sitting and reduced physical activity, which can have adverse health effects if not monitored and managed. Smartwatches are equipped with various sensors that can play a pivotal role in detecting the "Working on Laptop" activity. The accelerometer in a smartwatch can detect motion and measure changes in velocity. When a person is working on a laptop, their arm movements (e.g., typing, moving the mouse) generate subtle but distinct patterns of motion. By analyzing these motion patterns, the smartwatch can identify periods of laptop use. The gyroscope sensor can provide information about the orientation and angular velocity of the wrist. When someone is actively typing or using a laptop touchpad, their wrist orientation may exhibit specific patterns. The gyroscope can help differentiate between laptop use and other activities. By combining data from these sensors and applying machine learning techniques, a smartwatch can effectively monitor and detect the "Working on Laptop" activity. This information can be valuable for users to track their sedentary time, promote breaks, and maintain a healthier balance between work and physical activity, ultimately contributing to their overall well-being.

9. **Browsing on the phone:** In today's interconnected world, "Browsing on the Phone" has become a ubiquitous and integral part of daily life. This common sedentary activity involves individuals using their mobile phones for a variety of purposes, from social media scrolling to web browsing and texting. One defining feature of Browsing on the Phone is how individuals hold their devices. Typically, the phone is cradled in the user's dominant hand, with the screen positioned in front of their face or slightly lower. This natural and ergonomic gesture facilitates ease of use, enabling users to navigate the touchscreen, swipe, tap, and type effortlessly. It sheds light on our hand movements during extended periods of inactivity, reflecting our evolving relationship with technology.

10. **Cycling:** Cycling, the act of riding a bicycle with hands firmly gripping the

handlebars, is a multifaceted activity that combines physical exertion, recreation, and transportation. While the cyclist's hands maintain a consistent position on the handlebars, the smartwatch's sensors play a pivotal role in capturing the nuanced motions associated with this activity. Cycling is not merely a mode of movement; it's a lifestyle. It encompasses various forms, from road cycling for speed enthusiasts to urban commuting and the thrill of navigating rugged terrains on mountain bikes. The heart of cycling lies in the rhythmic pedaling, balance, and the connection between the cyclist and the bicycle. Modern smartwatches are equipped with an array of sophisticated sensors, including accelerometers and gyroscopes. These sensors are designed to detect a range of movements and changes in orientation. In the context of cycling, they emerge as invaluable tools, capable of capturing intricate aspects of this activity. While the cyclist's hands maintain a relatively stable grip on the handlebars, the smartwatch's sensors have the remarkable capability to detect subtle, continuous jerks, and vibrations arising from the motion of cycling. These sensors work like a silent companion on your cycling journey, turning the seemingly simple act of gripping your handlebars into a vibrant well of information.

11. **Sitting in Car:** Sitting in Car is an activity that represents a state of idleness while traveling in a vehicle. While it may seem like a passive activity, it holds significant relevance in understanding human behavior and mobility during transportation. This activity typically involves an individual occupying a seat within a motorized vehicle, such as a car, and experiencing motion as the vehicle moves from one location to another. One notable aspect of "Sitting in Car" is the potential for the smartwatch to detect slight motion. This motion is primarily attributed to the vehicle's movement. Even when seated, passengers can experience subtle shifts and vibrations as the car accelerates, decelerates, turns, or encounters bumps in the road. These subtle movements can be recorded by the smartwatch's sensors, providing insights into the level of motion experienced during the journey.

12. **Watching TV:** Watching TV is a sedentary behavior during which individuals remain seated or reclined in front of a television or screen for prolonged periods while consuming various content such as TV shows, movies, or streaming services. It involves minimal physical activity, often accompanied by low energy expenditure. This activity is pervasive in modern lifestyles and can contribute to extended periods of inactivity, potentially impacting one's health and overall well-being.

    During the act of watching TV, physical movement is typically minimal. A smartwatch equipped with accelerometers can detect and record this lack of motion by observing the wearer's wrist, which remains relatively stable. The wrist's limited or virtually absent movements serve as a key indicator that aids the smartwatch in identifying and categorizing this sedentary activity. Similarly, a gyroscope can register the stable orientation of the wrist, further reinforcing the identification of this inactive behavior. By combining data from both the accelerometer and gyroscope, the smartwatch can attain a more precise understanding of the user's activity level. For instance, it can recognize when the wrist remains stationary and there are no significant changes in wrist orientation, both of which are characteristic of watching TV.

1.2 Research Gaps

Exploring the extensive body of research in the field of activity detection, we have identified several critical research gaps that, if addressed, could significantly advance the domain. One primary gap is in determining the optimal types of sensors that balance accuracy, battery life, and practical usability in real-world scenarios. Additionally, there is a need to establish the ideal frequency of data collection that maximizes the accuracy of activity detection while minimizing energy consumption and data processing demands.

Research has shown that the use of multiple sensors can increase complexity, raise power consumption, and introduce potential issues with sensor data fusion. This multi-sensor approach often leads to practical inefficiencies. There is a pressing need to develop methods to efficiently manage and process data from multiple sensors without overwhelming the device's computational and battery capabilities.

Privacy and security concerns are particularly acute with vision-based systems that involve cameras, which could inadvertently capture sensitive personal data. It is crucial to ensure data security and user privacy in sensor-based systems, addressing vulnerabilities that could potentially expose user data.

In terms of real-time analysis and user feedback, there is a gap in implementing effective real-time data processing techniques to provide immediate feedback to users about their activity levels. Overcoming latency issues in real-time data analysis and feedback mechanisms is essential to enhance user experience and engagement.

Another significant research gap is the GPS and heart rate monitors are critical for accurate activity monitoring but are also major contributors to battery drain in wearable devices. This has a direct impact of high energy consumption on user experience, particularly in terms of device usability, frequency of charging, and the practicality of continuous monitoring. There is a need for using low-power sensors, energy-efficient data transmission methods, and algorithms that require fewer resources. All without compromising the accuracy and reliability of activity detection.

Practicality and user adoption also pose major challenges. Ensuring the wearability and comfort of devices, particularly when multiple sensors are required for accurate activity detection, is essential. Enhancing the user adoption rate by focusing on the practicality and convenience of wearable devices in everyday use is crucial.

Finally, there is a need for exploring advanced machine learning and neural networks to improve the accuracy of activity detection, especially in complex scenarios involving multiple activities. These research gaps reflect the current challenges and opportunities in the field of activity detection using wearable technology. Addressing these issues can lead to significant advancements in the development of more efficient, accurate, and user-friendly wearable devices.

1.3 Research Questions

In this thesis, we address several pivotal research questions designed to advance the functionality and user experience of smartwatches with a focus on Non-Exercise Activity Thermogenesis (NEAT) and non-NEAT activity detection:

Low Sampling Sensor Data and Battery Efficiency: How can utilizing low sampling sensor data enhance battery efficiency while maintaining sufficient accuracy for activity detection in smartwatches?

Optimal Sensor Selection for Activity Detection: What sensors available in smartwatches best balance the requirements for generalizability and battery optimization in detecting our chosen set of 13 activities?

Optimal Data Collection Frequency: What is the ideal frequency for data collection that maximizes accuracy in detecting NEAT activities while also conserving battery life?

Advanced Models for NEAT Detection: Can more complex models, such as hierarchical models, significantly improve the accuracy of NEAT activity detection over traditional flat models?

Deployable Models for Smartwatches: Are there deployable machine learning or neural networks that can efficiently overcome latency and battery limitations on smartwatches?

Balancing Accuracy and Battery Efficiency: What are the key parameters like features used for classification, the rate at which file needs to uploaded to server, data sampling frequency, and choice of window length for balancing accuracy in activity detection with battery efficiency, and how can these parameters be optimized in smartwatch technology?

These questions aim to deepen our understanding of activity detection using wearable devices, particularly in enhancing battery life, improving the accuracy of activity detection, and introducing real-time feedback within the constraints of wearable technology.

1.4 Objective Of Thesis

The primary objective of this thesis is to conceptualize, design, and construct an innovative Non-Exercise Activity Thermogenesis (NEAT) monitoring device that leverages wearable technology. This form of thermogenesis has gained significance due to its potential impact on overall health, energy balance, and weight management.

The envisioned NEAT monitoring device aims to address several key aspects:

1. **Efficiency**: The application on the smartwatch is engineered to efficiently and accurately monitor NEAT-related activities. This efficiency extends to both data

collection and analysis, ensuring that the smartwatch is practical for daily use and capable of producing meaningful insights.

2. **Wearable Technology**: Leveraging the capabilities of modern wearable technology, the application is designed to be unobtrusive and user-friendly. By integrating it into a smartwatch, our objective is to facilitate continuous and non-disruptive data collection during daily activities. We have worked on providing users with regular feedback directly on the smartwatch.

3. **Data sanity**: An essential aspect of the application is the precision and reliability of data collection. It utilizes meaningful sensors, such as the accelerometer and gyroscope sensors present in a smartwatch, and classification algorithms that can distinguish and quantify NEAT activities, enabling accurate assessment and measurement.

4. **Battery Efficiency**: The application running on the smartwatch is designed to optimize battery efficiency, thereby extending the smartwatch's battery life while maintaining a low sampling frequency rate. This balance between energy consumption and data quality ensures the device remains operational throughout the day without frequent recharging, enhancing its practicality.

1.5 Contribution Of Thesis

In this thesis, the NEAT Activity Detection System contributes to multiple areas, including robust activity detection, energy efficiency, privacy protection, comprehensive activity tracking, and real-time feedback. By leveraging a hierarchical model and low-frequency data from wearable sensors, the system accurately classifies a range of NEAT activities using advanced feature extraction techniques like statistical methods and ECDF, demonstrating high accuracy and robustness through real-world data. Its energy-efficient design, with low sampling frequencies and on-device processing, significantly reduces power consumption and extended battery life, making it practical for long-term use. The system successfully classifies thirteen diverse NEAT activities, representing everyday movements at home and outdoors, providing comprehensive insights into physical activity patterns. It also has a provision to scale to "Others" activity if not classified in any of the stated activities giving the model more scalability. It also addressed privacy concerns using non-intrusive wearable sensors like accelerometer and gyroscope, avoiding sensitive visual or audio data and enhancing user trust. Additionally, the system offers real-time feedback to users, encouraging increased physical activity and healthier behaviours through immediate insights, with on-device neural networks enabling instant updates and aiding consistent user engagement.

1.6 Significance of NEAT Monitoring Devices

The field of wearable technology has witnessed remarkable advancements in recent years, enabling the monitoring and analysis of various physiological parameters in real-time. One prominent area of research within this domain is the application of wearable technology in monitoring Non-Exercise Activity Thermogenesis (NEAT) and non-NEAT activities. This section of the thesis focuses on illuminating the application scope of utilizing wearable technology to monitor NEAT. The integration of wearable devices, such as accelerometers, gyroscopes, provides an unprecedented opportunity to capture and analyze NEAT data in a non-invasive and continuous manner. The potential applications of NEAT monitoring using wearable technology are multifaceted and extend across several domains. Monitoring Non-Exercise activity thermogenesis (NEAT) using wearable technology holds significant potential in the **healthcare** field. Tracking NEAT can provide valuable insights into an individual's daily physical activity, which is crucial for overall health and wellness. NEAT monitoring can help individuals and healthcare professionals better understand the impact of daily activities on energy expenditure. This information can be used to design personalized weight management plans *"Interindividual variation in posture allocation: possible role in human obesity"* published in *"American Association for the Advancement of Science"* [1] - Levine et al. concludes that obesity is a result of an imbalance between energy intake and energy expenditure. They specifically focus on non-exercise activity thermogenesis (NEAT) and their study found that mildly obese individuals spent, on average, about 2 hours more each day in a seated position compared to lean individuals, and this difference was biologically determined rather than a result of body weight changes. The author suggests that if obese individuals adopted the NEAT-enhanced behaviors of lean individuals, they could potentially burn an additional 350 calories per day, which could aid in weight management as per stated by [1].

Wearable devices can track prolonged periods of sedentary behavior, which has been associated with various health risks. Monitoring NEAT can prompt individuals to break up sedentary time with short, active breaks. *"Sedentary time and cardio-metabolic biomarkers in US adults: NHANES 2003–06"* published in *"European heart journal"* [2] - Healy et al. conclude that these findings, which are the first to represent the general population, demonstrate the harmful links between extended periods of sedentary behavior and adverse cardio-metabolic and inflammatory bio-markers. This suggests that healthcare practitioners and public health messages should emphasize the importance of reducing and interrupting sedentary time as a preventive measure to mitigate cardiovascular disease risk.

Smartwatches equipped with NEAT monitoring capabilities can provide users with a comprehensive view of their daily activity levels. This data can be used to create personalized fitness plans that are tailored to an individual's lifestyle and activity patterns. For example, someone with a sedentary job may need a different workout plan than a person with an active occupation. *"Could the Blue Zone's exercise secrets hold the key to* *vital ageing?"* [3] - The author clearly portrays that Boosting NEAT is a straightforward and accessible approach to incorporate more movement in daily life. The beauty of increasing NEAT lies in its simplicity; it doesn't necessitate a gym membership or intense high-intensity interval training (HIIT) sessions. *"Objective quantification of physical activity in bariatric surgery candidates and normal-weight controls"* [4] - The study's by Bond et al. findings reveal a concerning trend among bariatric surgery candidates, indicating that they tend to have notably low levels of physical activity. Moreover, it appears that they infrequently partake in physical activities that meet the necessary criteria for duration and intensity required to maintain and enhance their overall health. This underscores the significance of further research aimed at identifying effective strategies for boosting physical activity levels in this particular demographic of individuals who are considering bariatric surgery.

NEAT plays a significant role in weight management, as it contributes to daily calorie expenditure. By understanding how their non-exercise activities impact their energy balance, individuals can make more effective decisions regarding calorie intake and exercise frequency. *"Accuracy of armband monitors for measuring daily energy expenditure in healthy adults"* published in *"Medicine and science in sports and exercise"* [5] - The study aimed to assess the accuracy of energy expenditure estimations obtained from two portable armband devices: the SenseWear Pro3 Armband (SWA) monitor and the SenseWear Mini Armband (Mini) monitor, while volunteers engaged in their daily activities in a real-world, free-living setting. Monitoring NEAT can also help individuals avoid over training. If excessive NEAT is combined with strenuous exercise, it can lead to burnout and increased risk of injury. Smartwatches can provide insights into the balance between activity and rest. Sedentary behavior is associated with various health risks, including cardiovascular disease and obesity. Smartwatches can serve as a tool to promote a more active and healthier lifestyle. *"Too little exercise and too much sitting: inactivity physiology and the need for new recommendations on sedentary behavior"* published in *"Current cardiovascular risk reports"* [6] - Hamilton et al. in this study emphasises that while we know the benefits of physical activity, we should also recognise the independent and harmful effects of prolonged sitting on cardiovascular and metabolic health. This warrants a reconsideration of how we approach and address sedentary behavior in public health guidelines and recommendations.

1.7 Organisation of thesis

In Chapter four [4], we present a systematic approach to developing a hierarchical classification model for distinguishing Non-Exercise Activity Thermogenesis (NEAT) activities in a home setting. Our objective was to create a model that effectively differentiates between NEAT and non-NEAT activities, using seven representative tasks: cooking, sweeping, mopping, walking, climbing stairs, and sedentary activities like watching TV or desk work. Data collection was performed using smartwatches with accelerometer and gyroscope sensors at a low sampling rate of 1Hz. The raw data was processed and structured into a .csv format, followed by feature extraction. Initial models, including KNN, MLP, SVM, Logistic, Random Forest, Naive Bayes, and XGBoost, struggled with low-frequency data. This led us to develop a hierarchical model using binary classifiers, simplifying the seven-class problem and yielding better performance in accuracy and efficiency. Our model was validated with real-world data from 10 volunteers. This number was chosen due to COVID-19 restrictions, which limited the feasibility of gathering larger groups.

In the fifth chapter [5], we built on our previous chapter by adding six new activities to our study, expanding the total to thirteen: Cooking, Sweeping, Mopping, Walking, Climbing Up, Climbing Down, Eating, Driving, Working on a Laptop, Browsing on a Phone, Cycling, Sitting in a Car, and Watching TV. This broader scope offered richer insights into daily behaviors. We focused on key parameters like feature selection, data upload frequency, sampling rate, and window length, deploying models on a server for real-time user feedback. We challenged the assumption that high-frequency data is always needed, exploring low frequencies (10 Hz and 1 Hz) to see their effects on battery longevity and classification accuracy.

Our hypothesis emphasized balancing energy efficiency and classification accuracy for NEAT activities, using smartwatches. We achieved this through reduced sampling rates (1Hz, 5Hz, 10Hz) and feature processing. Experiments with ten volunteers revealed that lower sampling rates extended battery life but made distinguishing activities more complex, highlighting the trade-off between accuracy and frequency.

We also added classification for 'other' activities, enhancing system flexibility. By deploying models on server, we enabled real-time feedback via Wi-Fi smartwatches. However, the system's reliance on Wi-Fi and server-based processing posed limitations, especially in connectivity-poor areas, affecting usability and response time. Additionally, data transmission raised security concerns, stressing the balance between real-time feedback and user privacy.

Chapter six [6] of our thesis represents a significant shift in our approach, addressing the challenges identified in chapter five concerning real-time data processing and feedback provision in a server-based environment. We outline our move towards deploying neural networks directly on the smartwatch, transitioning from reliance on server processing to a more autonomous, device-centric model. This chapter elucidates the limitations we faced and the innovative solutions we developed to enhance NEAT activity monitoring using wearable technology. We begin by exploring the initial constraints of our previous model, including the dependency on a continuous Wi-Fi connection for data transfer, latency issues affecting real-time feedback, and data security concerns with network-layer transfers. This discussion sets the stage for our strategic pivot to on-smartwatch deployment. We transitioned to neural networks using TensorFlow Lite to handle on device processing efficiently. Our research focused on continuous data collection from smartwatch sensors, ensuring precise activity monitoring. We carefully analyzed battery consumption, balancing model accuracy with energy efficiency. By shifting to raw data instead of extensive preprocessing, we streamlined our process, making the system more responsive. Additionally, we added real-time feedback to enhance user experience and encourage a healthier lifestyle. Finally, we evaluated model performance, assessing accuracy in detecting NEAT activities and the impact on battery life.

Chapter 2

Data Collection and Deployment for NEAT Monitoring

2.1 Data Collection

2.1.1 Important Elements of data collection

In our research on "Non-Exercise Activity Thermogenesis Monitoring Using Wearable Technology" we came across an interesting factor of research problem which revolved around collecting high quality data to investigate the thermogenic response associated with Non-Exercise activities. This builds the basis of our research, as its essential to understand how wearable technologies can effectively monitor these activities.

Data Collection Method

Our primary focus was on developing a robust data collection method. To address this, we utilized a smartwatch, equipped with accelerometer and gyroscope sensors, which were instrumental in tracking volunteers' movements (Device configuration: 2.2.1). This smartwatch was comfortably worn on the dominant hand of each volunteer, serving as a subtle and convenient means to record various physiological parameters and motion data.

These sensors continuously monitored and recorded data, actively pinging and detecting changes once a volunteer initiates an activity. This real-time activity detection ensured that the precise moments of activity were captured, providing us with a detailed and granular dataset for our research.

Wearable Application Composition:

To streamline the data collection process and ensure user-friendliness, we developed a customized wearable application. This application featured a comprehensive list of activities that the volunteers could engage in, accompanied by user-friendly controls, including start, pause, and stop buttons. Upon starting the tracking process, volunteers were prompted to select an activity of their choice. Subsequently, data tracking commenced. When the activity concluded, volunteers could press the stop button, which would save the recorded data in .csv format on the smartwatch.

Taking into consideration that collecting data for extended periods can be physically taxing, we also provided a pause button, which allowed users to temporarily halt the data recording if they felt the need to rest between activities. It's important to note that we emphasized to volunteers that the usage of this pause button should be minimized to ensure the integrity and continuity of data collection.

Number of Volunteers

A total of 10 volunteers were carefully chosen to participate in our data collection process. These volunteers were selected from the age group of 30-60 years. This selection examined how NEAT activity might vary across different age groups, offering valuable insights into the data collected. Choosing 10 volunteers helped keep the data management and analysis practical, citing the COVID restrictions levied during our research. We acknowledge while the sample size allows for detailed data collection and analysis, it may limit the generalizability of our findings. Further studies with larger and more varied volunteer pools can validate and expand on our findings, ensuring broader reliability and applicability.

Environment

In this research, data was collected in a partly controlled setting. Volunteers could go about their daily activities naturally, but if they choose to perform an activity out of the selected 13 activities, they could use the Start feature on the smartwatch and record their activity. This setup was chosen to ensure the data was realistic while keeping the environment focused. The smartwatch app had a pause feature to handle unexpected interruptions and keep the data accurate. This lets volunteers temporarily stop tracking if needed.

The data collection lasted for two weeks for each volunteer. This period was chosen to get enough data without making it too difficult for the volunteers. During these two weeks, volunteers used the smartwatch to choose and track their activities, allowing us to capture detailed accounts of their interactions and daily movements.

Data Cleaning

To enhance the accuracy and reliability of our collected data, we implemented a data cleaning process. This involved the removal of initial and final data points during each activity. These specific data points were removed to mitigate the impact of unwanted motions that often occurred at the onset and conclusion of an activity, which could potentially distort the data quality.

Data Sanity

Given that the sensors were an out-of-the-box implementation of the smartwatch, we adopted a rigorous approach to maintain the integrity and sanity of the data. To achieve this, we decided to use raw sensor values directly, without any manipulation. By utilizing these raw sensor values, we ensured that the data we collected was in its most unaltered state, guaranteeing that the values we extracted were unquestionably genuine and untainted.

This approach allowed us to maintain a high level of confidence in the authenticity of the data, making it a reliable foundation for our research and analysis. By preserving the raw sensor values, we minimized the potential for unintended biases or inaccuracies, resulting in a robust dataset for our work.

In summary, our research problem revolved around effective data collection that could be instrumental in studying Non-Exercise activity thermogenesis. The choice of wearable technology, the development of a user-friendly application, the selection of an age-specific sample, and the meticulous attention to data integrity in a semi-controlled environment collectively addressed this critical challenge, forming the foundation for our research.

2.1.2 Data collection using smartwatch

Our research chapter commenced with a rigorous assessment of prominent existing datasets, including the Human Activity Recognition UCI HAR Dataset, UCI WISDM Smartphone and Smartwatch Activity and Biometrics Dataset, and ActiGraph datasets. These datasets, while valuable, presented certain limitations in terms of the sensor technologies employed and the data collection methods employed, thereby posing significant challenges aligned with our research objectives.

Firstly, these datasets featured different sets of activities, making it difficult to achieve a standardized and comprehensive view of Non-Exercise activity thermogenesis (NEAT). Moreover, the activities were conducted in controlled environments, which did not accurately represent the diverse and dynamic real-world scenarios we aimed to investigate. Our core challenge arose from the need for a more naturalistic and adaptable source of data collection. We recognized the significance of obtaining sensor data from everyday devices, particularly smartphones, which have become integral to modern life. However, even with smartphones, it proved challenging to obtain a seamless stream of data for various activities due to their physical nature and inherent limitations in sensing capabilities.

To address these limitations and to pave the way for our research, we identified wearable technology as a promising solution. We saw the potential to harness the sensors embedded in smartwatches to collect sensor data while individuals engaged in various activities. This shift to wearable technology not only offered a novel approach to data collection but also presented a feasible means to acquire our raw data. This allowed us to gather authentic and real-world sensor data, providing a foundation for conducting in-depth research into Non-Exercise activity thermogenesis.

Our research commenced with these challenges and culminated in the innovative use of wearable technology to collect comprehensive and naturalistic sensor data. This transformation not only overcame our initial research hurdles but also enabled us to explore Non-Exercise activity thermogenesis in a manner that was both practical and pioneering.

Challenges faced during data collection

We faced multiple challenges during research, particularly in ensuring volunteers were compliant and managing environmental factors. The major challenge was ensuring that the volunteer wore the smartwatch consistently on the dominant hand throughout the data collection period and was performing the activity consistently, as any inconsistency in the data collected could lead to misclassification and reduced accuracy of activity detection. To mitigate this issue, volunteers were given clear instructions and a trial run, which included familiarising them with the environment in which data was to be collected, the usage of the app and the data collection process. Additionally, physical obstructions also added a challenge to the data collection process. To overcome this challenge, we collected data in various typical home environments, ensuring the model can generalise across different conditions. We also gave them detailed guidelines and training sessions on using the application and conducted regular follow-ups to ensure adherence to guidelines to maintain data quality.

2.2 Factors considered in building NEAT monitoring device

2.2.1 Device Configuration

Figure 2.1: Fossil Sport Smartwatch 43mm - FTW4019 , Image credits : Amazon.com

We used Fossil Sport Smartwatch 43mm - FTW4019 (Figure: 2.1), equipped with a Snapdragon Wear 3100 processor, 512MB RAM, and 4GB internal storage, as our data collection tool. This works on Wear OS by Google, running the latest version of Android, making it easier for other Wear OS powered devices to download our app. It also has higher power and storage compared to other fitbits and heart rate monitoring devices. This is a commonly used smartwatch in budget that has all the sensors required for activity detection.

2.2.2 Choice of sensors

NEAT encompasses various daily activities that contribute to energy expenditure, such as walking, standing, fidgeting, and other subtle movements that are not deliberate exercises. To address this problem effectively, we needed to choose the most suitable sensors from a wide array of options available.

The available sensors that we considered for capturing NEAT activities included accelerometers, gyroscopes, barometers, heart rate sensors, step count sensors, and magnetometers. While previous researches like [7] [8] [9] has demonstrated the potential of combining these sensors to create models for activity detection, we identified specific challenges and limitations associated with some of these sensors that influenced our decision-making process.

Accelerometer: An accelerometer is a sensor that measures acceleration forces experienced by the device it is integrated into. These forces can be due to both static gravity (e.g., when the device is at rest) and dynamic motion (e.g., during physical activities). Accelerometers provide information about changes in velocity and direction, making them essential for tracking movements and physical activities. Accelerometer data typically consists of three values, corresponding to the three spatial axes: X, Y, and Z. These values represent the acceleration along each axis at a given point in time.

Gyroscope: A gyroscope is a sensor that measures angular velocity or rotational motion. It helps determine the orientation and changes in orientation of the device. Gyroscopes are crucial for recognizing movements such as tilting, rotation, and angular changes, which can be valuable for understanding various physical activities.

Barometer: A barometer is a sensor that measures atmospheric pressure. It is used to estimate changes in altitude or elevation. In some cases, it can be employed to gauge subtle changes in height, like climbing stairs or ascending slopes. However, one of the challenges we encountered was the limited availability of barometers in wearable devices. Barometers are considered premium sensors and are not commonly found in everyday smartwatches or fitness trackers. This limited their practicality for widespread NEAT activity monitoring.

Heart Rate Sensor: A heart rate sensor, often based on photoplethysmography (PPG) technology, measures the user's heart rate by detecting changes in blood volume through the skin. While valuable for assessing cardiovascular activity and exercise intensity, heart rate sensors tend to consume significant amounts of power, making them less suitable for continuous long-term monitoring. Although heart rate sensors are commonly integrated into wearable devices, they pose a significant drawback in terms of battery consumption. Monitoring NEAT activities over an extended period requires a sustainable power source, making heart rate sensors less suitable for continuous tracking.

Magnetometer: A magnetometer is a sensor that measures the strength and direction of magnetic fields. It is primarily used for tasks like compass orientation or detecting magnetic anomalies. Magnetometers have limited relevance in NEAT activity monitoring. They did not provide any significant additional value for our research, as NEAT activities are predominantly characterized by physical movements rather than magnetic field variations.

Step Count Sensors: Step count sensors, often integrated with accelerometers or as standalone components, track the number of steps taken by the user. They are commonly used to estimate the distance walked or run and are a key feature in fitness trackers. However, their application is somewhat specific to activities that involve walking or running. Duplication of data collection by including both step count sensors and accelerometers may lead to unnecessary complexity without substantial benefits.

Considering these factors, our research found that accelerometer and gyroscope sensors emerged as the most sensible choices for monitoring NEAT activities. These sensors are relatively common in wearable devices, provide essential motion data, and are power-efficient, making them well-suited for continuous and sustainable NEAT activity tracking.

By addressing these challenges and limitations associated with sensor selection, our research aims to provide a practical and effective solution for monitoring NEAT activities using wearable technology, thus contributing to the advancement of health monitoring and physical activity research.

2.2.3 Sensor frequency selection

To monitor NEAT accurately and continuously, researchers face the challenge of selecting an appropriate sensor frequency, specifically choosing between high sampling rate and low sampling rate sensors.

High Sampling Rate Sensors:

High sampling rate sensors tend to capture data very quickly, recording several measurements per second. This can prove to be a great choice where we need real-time information and high temporal resolution. However, because the work at such a high frequency, they tend to consume more power, requires higher data storage and can be more prone to motion artifacts (*Motion artifacts are unwanted signals in sensor data caused by user movement, leading to inaccurate measurements in wearable and medical devices. For example, if a heart rate monitor gets bumped around, it might show an incorrect heart rate. Similarly, an accelerometer meant to measure specific movements might also pick up unrelated body movements. Addressing these artifacts is crucial for ensuring the accuracy and reliability of sensor data*).

Low Sampling Rate Sensors:

Low sampling rate sensors collect data at a slower pace, with fewer data points over a given time interval. They are appropriate for applications where high temporal resolution is not critical, and energy efficiency is a priority. Low sampling rate sensors consume less power, generate smaller data volumes, and are less prone to motion artifacts.

Choice of Low Sampling Frequency (≤10Hz): In the context of monitoring NEAT using wearable technology, it's essential to balance the need for data precision with the sustainability of the solution. Recognizing the importance of prolonged monitoring periods with minimal battery drainage and the practicality of using smartwatches, we choose a low sampling frequency of less than or equal to 10Hz.

This choice of a low sampling frequency aligns with our goal of achieving a sustainable, long-term monitoring solution. It reduces power consumption, extending the battery life of the wearable device, and ensures that the data collected is sufficient for capturing trends in NEAT while maintaining a reasonable level of accuracy. This decision enabled us to conduct comprehensive NEAT monitoring without the constraints of frequent battery recharges or excessive data management burdens.

Table 2.1: Advantages and Disadvantages of High vs. Low Sampling Rate Sensors

Metric	High Sampling Rate Sensors	Low Sampling Rate Sensors
Data Precision	High precision due to frequent measurements.	Lower precision as data is less granular.
Power Consumption	Tend to consume more power, reducing battery life.	Lower power consumption, longer battery life.
Data Volume	Generates large data volumes, requiring more storage and bandwidth.	Smaller data volumes, easier to manage.
Motion Artifacts	More susceptible to motion artifacts and noise.	Reduced impact of motion artifacts.
Wearable Comfort	May require bulkier and less comfortable devices.	Allows for sleeker and more comfortable wearables.
Application Scope	Suitable for applications demanding high temporal resolution.	Better for applications with less time-sensitive requirements.

Table 2.1 compares the merits and constraints of high and low sampling rate sensors in wearable devices. High sampling rate sensors capture more precise data, capturing small changes over time, crucial for detailed real-time analysis. However, they use more power, reduce battery life, and produce a lot of data, needing more storage and processing power. They are also more susceptible to errors from unwanted movement, requiring complex filtering to get accurate readings. This can make devices bulkier and less comfortable for user compliance.

Low sampling rate sensors consume less power, extending battery life and allowing for smaller, more comfortable devices. While they provide less detailed data, they are less affected by movement errors, making them suitable for applications that want prolonged data capturing.

Choosing between high and low sampling rate sensors depends on the application's specific needs. High sampling rate sensors are necessary for tasks that need detailed time-specific data, while low sampling rate sensors are better for tasks where comfort and longer battery life are more important.

2.2.4 Deployment in smartwatch

As we moved forward in the "Non-Exercise Activity Thermogenesis Monitoring Using Wearable Technology," we encountered a complex research problem requiring us to think outside the box. Initially, like many previous studies, we used mobile technology to collect data from wearable devices, and then we processed that data on remote servers *(we used python as a scripting language and used heroku to setup our server)*. However, this approach came with several tough challenges that needed creative solutions:

High Battery Drain: Using mobile technology for data collection, which required sending data to a server at specific intervals, placed a considerable strain on the wearable device's battery. This strain was primarily due to the need for an active Wi-Fi connection during these data transmission periods. As a result, this energy-intensive process rendered the wearable impractical for long-term monitoring, posing a significant obstacle to achieve continuous and dependable data collection.

Expensive Server Setup: Setting up and maintaining server infrastructure for data processing was costly, making it less accessible for research projects with limited budgets. This also limited the scalability of wearable technology solutions, particularly when considering the potential for increased dataset size in the future. The infrastructure's capacity to handle larger datasets became a significant concern, potentially bottle-necking the system's ability to accommodate the growing demands of Non-Exercise activity thermogenesis monitoring.

Security Concerns: Transmitting sensitive user data from mobile devices to remote servers introduced security risks. Data breaches and privacy concerns became serious issues, requiring strong security measures and data protection.

Latency Problems: Data collected from wearable devices often experienced delays in processing on remote servers, leading to slow feedback to users about their activities. Real-time insights were compromised, affecting the effectiveness of the monitoring system.

Processing Cost for Flat Models: In server-based deployment, preprocessing of data is necessary before it can be fed into flat models. This preprocessing step consumes computational resources on the smartwatch, incurring a processing cost that can affect overall performance and responsiveness.

In response to these challenges, we realized the need for a more efficient and user-friendly approach. We shifted our research strategy toward deploying machine learning models directly on smartwatches, bypassing the server-based processing phase. However, this transition came with its own unique set of challenges.

Flat models, known for their straightforward architectures and minimal computational demands, had long been a staple in various applications. However, their compatibility with wearable devices, such as smartwatches, presented substantial hurdles. One notable limitation was their inability to be deployed directly on smartwatches, primarily due to the constraints associated with the pickle format file. This challenge prompted us to rethink our approach and seek alternative solutions.

The transition from traditional flat models to more computationally efficient neural networks was a critical step in our research. As mobile technology has evolved, we've witnessed the emergence of tools like TensorFlow Lite API, which allows for the integration of lightweight neural networks directly into Android apps. These TensorFlow Lite (TFLite) models are specifically designed to be highly efficient and resource-friendly, making them ideal for resource-constrained devices like smartwatches. With TFLite, we deployed neural networks directly within the app, eliminating the need for sending data to remote servers for processing. This not only streamlined the process but also enhanced data privacy and

security by keeping sensitive information on the device.
However, it's important to note that while TFLite excels in running lightweight neural networks, it may not be suitable for flat models, which typically have different architectural requirements. Therefore, our shift toward more computationally efficient neural networks was not only in response to the limitations of traditional flat models but also in alignment with the technological advancements that allow for direct, on-device model deployment, enhancing the overall efficiency and effectiveness of our monitoring system.
Our innovative solution involved developing and deploying hybrid neural networks directly on smartwatches. This approach offered several significant advantages:
Improved Security: Eliminating data transmission to external servers eliminated the security risks associated with data exposure during transit. This direct processing approach enhanced user privacy and data security.
Real-Time Feedback: Our novel deployment method enabled immediate, real-time feedback to users. They could now receive timely and accurate insights into their ongoing activities, greatly enhancing the overall effectiveness of our monitoring system.
No Preprocessing for neural networks: In contrast, when deploying neural networks, preprocessing becomes unnecessary as these models have the capability to directly ingest raw sensor data. This streamlined approach not only eliminates the need for preprocessing but also enhances the efficiency of the processing pipeline on the smartwatch, leading to improved overall performance.
Assumptions: During data collection using smartwatch we operated under two assumptions. We assumed that the way our 10 volunteers perform the activity will generalise for a larger group of volunteers as our sample includes individuals from diverse age groups. We also assumed that the accuracy of sensors in smartwatch is reliable for performing data collection and to train our models.
In summary, our research addressed the initial challenges of high energy consumption, costly server setups, security vulnerabilities, and latency issues associated with mobile technology and server-based data processing. We overcame these obstacles by developing and deploying hybrid neural networks directly on smartwatches, resulting in no requirement of data preprocessing, enhanced data security, and real-time user feedback. This innovative approach represents a significant advancement in the field of Non-Exercise activity thermogenesis monitoring, making wearable technology solutions more accessible and efficient for everyone.

Table 2.2: **Server Deployment and Proposed Solution**

Parameter	**Challenge**	**Proposed Solution**
Data Collection Method	Mobile technology with data transmission to servers, high battery drain, security risks.	Direct deployment of machine learning models on smartwatches.
Server Infrastructure	Expensive server setup, scalability issues with growing datasets.	On-device model deployment with TensorFlow Lite API.
Data Security	Security risks with sensitive data transmission to remote servers.	Enhanced user privacy through local processing.
Data Processing Latency	Delays in data processing on remote servers, impacting real-time feedback.	Real-time insights through on-device model deployment.
Processing Cost for Flat Models	Preprocessing data on a smartwatch with processing cost.	Elimination of preprocessing with neural networks.
Compatibility with Smartwatches	Inability to deploy flat models directly on smartwatches.	Development and deployment of hybrid neural networks.
TensorFlow Lite (TFLite)	Integration of lightweight neural networks into Android apps.	Efficient and secure on-device model deployment.

Chapter 3

Related Work

This chapter presents a condensed review of the state-of-the-art contributions with respect to the following areas within the domain of activity detection. The underlying principle and limitations, along with the improvements made to these seminal contributions, have also been highlighted.

3.1 Activity Detection

Activity detection lies at the heart of NEAT monitoring, and in order to correctly detect activity being performed, various approaches have been explored in the past. Out of these approaches, sensor-based and vision-based approaches have been the forerunners. Many studies explore these approaches and showcase how they can be utilised in activity detection. In wearable devices, sensors like accelerometers, gyroscopes, and others help detect activity by capturing movement data. Researchers like [10] [11] [12] have extensively worked on building algorithms that can effectively recognise activities like walking, running, cycling, and even more specific tasks like stair climbing. [13] [14] worked on technology that allowed continuous tracking of daily activities and provided valuable insights into the physical well-being of an individual. On the other hand, the vision-based approach uses a camera to capture information and analyse human activity, as represented in this study [15]. With a vision-based approach, activity recognition comes with its own challenges. One of the biggest concerns lies in privacy and data security; using a camera that captures visual information may inadvertently capture PII(Personally Identifiable Information) information , which raises security concerns. Moreover, there is a hardware challenge due to increasing power and storage demand; devices are becoming bulkier and less convenient for users to use in day-to-day life [16] [17]. Lighting conditions become another major factor as vision-based detection does not work under low-lighting conditions or environmental factors; thus, recording all kinds of activity without a proper lighting setup is impossible. Data processing intricacies involved with visual analysis can take time to process and lead to delay and synchronisation challenges. Vision-based can bring in a lot of contextual details but lacks precision in quantifying smaller and repetitive activity metrics. To bring accuracy to the system, a vision-based approach needs a proper camera calibration and lighting setup to give optimal results, which might be counter intrusive in daily activity detection. Considering both approaches and judging the situation of COVID-19 protocols, we decided to go ahead with a sensor-based approach using smartwatch sensors. The rationale behind the sensor-based approach was that we wanted to capture granular physical motions, keeping in mind the large-scale deployment for individual monitoring. We wanted to build high-precision activity detection with the thought of implementing real-time health monitoring and eradicating privacy-sensitive concerns. Lastly, we wanted to pick a convenient mode of data collection, which was fulfilled by the use of wearable smartwatch sensors.

Activity detection in the healthcare industry has created opportunities to enhance the quality of life for individuals, especially the elderly [16] [18], by preventing illness, accidents, and disease. Wearable technology has ushered in an era of preventive healthcare [19]. Continuous monitoring of vital health data, such as heart rate, sleep patterns, and activity levels, allows for early detection of potential health issues. This proactive approach can help individuals and healthcare professionals take timely actions to prevent illness and improve overall health [20].

In addition to that, wearable devices equipped with fall detection and emergency response features are invaluable for the elderly [21]. These devices can automatically alert caregivers or emergency services in the event of a fall or medical emergency, providing peace of mind to both the individuals and their families [22].

3.2 Usage of Multiple Sensors

In the domain of activity recognition research, various studies have sought to achieve high accuracy in differentiating between activities through the use of sensor-based approaches. However, a common challenge faced by these studies is the heavy reliance on multiple sensors for data collection, which may hinder their practicality in real-world scenarios.

The field of activity recognition has seen a plethora of research approaches, each shedding light on unique aspects and challenges. In the realm of general activity recognition, studies such as those conducted by Zheng et al. (2017) [23] and Nandy et al. (2019) [24] have harnessed the power of multiple sensors to achieve remarkable levels of accuracy. Zheng and colleagues involved 20 volunteers in their study, demonstrating an average recognition accuracy of approximately 96.0%. However, their dependence on a wide array of sensors, operating within the frequency range of 5Hz to 50Hz, has raised concerns regarding the practicality of these systems in real-world scenarios. Similarly, Nandy and the research team embarked on a comprehensive exploration of activity recognition, encompassing various activities like sitting, walking, and running. They harnessed the capabilities of wearable and smartphone-embedded sensors, culminating in an impressive 94% recognition accuracy. Despite their achievements, the extensive reliance on multiple sensors for data collection has posed practical challenges.

In the pursuit of advanced activity recognition, innovative approaches have surfaced, exemplified by the work of Nandy (2020) [25] and Roychowdhury et al. (2018) [26]. Nandy's work represents a significant leap, capable of identifying both static and dynamic intense activities, including walking while carrying weight. This advanced system combines data from a smartphone's accelerometer and a wearable heart rate sensor. Nandy's work achieved a 96% accuracy using an ensemble model and novel feature extraction techniques. The approach combined traditional classifiers such as LDA, DT, K-NN, and SVM, with an ensemble model using stack generalization and a neural network meta-learner. This method effectively handled variability in smartphone positioning and user behaviors. However, the practicality of such systems is a point of contention, given the considerable reliance on multiple sensors. Roychowdhury and the team chose to employ feature extraction and learning methods to distinguish various detailed activities, such as slow and brisk walking. Their introduction of a novel feature based on "jerk" for activity detection resulted in a remarkable 95% accuracy rate. While the efficacy of these methods is evident, concerns regarding the practicality of multiple sensor utilization in real-world settings persist.

The application of Deep Convolutional Neural Networks (CNNs) has emerged as a promising avenue in activity recognition, a trend observed in studies beyond those explicitly mentioned here. Deep learning models, particularly when tasked with combining data from various sensors, exhibit notable accuracy. Noori et al. (2020) [27], in their work, have likely delved into this challenge, shedding light on the intricacies of achieving precision through the amalgamation of sensor data using CNNs. However, a common challenge associated with these systems is their reliance on multiple sensors. This raises a pertinent question: is the pursuit of high accuracy in activity recognition inherently at odds with the practicality of sensor usage in real-world applications?

Stepping into the domain of specific activity recognition, Kumari et al. (2021) [28] have adopted a unique focus on body sensor networks for detecting simple activities like sitting, standing, walking, and resting. While their research provides granularity in the recognition of activities, the practicality of employing numerous body-worn sensors in everyday life remains a legitimate concern. Similarly, Akiduki et al. (2022) [29] have proposed a sophisticated system for detecting inattentive driving, leveraging not only a heart rate sensor but also body-worn inertial sensors. The combination of three detection models and the deployment of a smart anomaly detection algorithm has resulted in pinpoint accuracy in identifying driver behavior. However, the incorporation of multiple sensors in data collection may give rise to practical challenges when it comes to real-world implementation.

Within the realm of sensor placement and fusion strategies for optimizing activity recognition, Tang (2020) [30] has explored the impact of sensor locations. Their findings suggest that using sensors on areas other than wrists yields superior results, with a recommended approach of using at least two sensors on non-wrist locations. This underscores the significance of sensor placement in the pursuit of recognition accuracy.

Maylor (2023) [31] delves into the relationship between physical activity volume and intensity and their impact on cardiometabolic health. However, their study, which involves accelerometers worn on both the thigh and wrist, raises concerns about the practicality of deploying multiple sensors on various body parts in real-world scenarios.

Sensor fusion strategies, as showcased by Awais et al. (2016) [32], Attal et al. (2015) [33] and Ullah et al. (2021) [34], offer an avenue for enhanced recognition accuracy.

Attal's study effectively utilizes three wearable accelerometers situated at the subject's chest, right thigh, and left ankle for the classification of typical daily human activities. Ullah et al. introduce a framework consisting of a 3-axis accelerometer, 3-axis gyroscope, and temperature sensor worn on the body for classifying physical activities, achieving commendable recognition rates. Nevertheless, these findings underscore the potential limitations of employing multiple sensors in real-world scenarios.

In the era of deep learning, Saeed et al. (2022) [35] demonstrate that deep learning methods outperform traditional machine learning techniques for activity recognition. The remarkable accuracy and efficiency of deep learning methods are exciting but come with practicality concerns, especially when multiple sensors are required.

Anand et al. (2021) [36] take an innovative approach, utilizing sensors placed on individuals' bodies to teach computers how to learn, akin to training a dog to perform tricks step by step. While their approach introduces a novel perspective on activity recognition, it also underscores the practical challenges of employing multiple sensors in real-world applications. In a similar vein, Eska et al. (2023) [37] offers an interesting perspective on the practicality aspect. They developed a system called REPLAY, aimed at providing real-time feedback during exercise using heart rate monitors and motion trackers. This system employs sensors, including heart rate monitors and motion trackers, to show users their physiological data while engaging in high-intensity interval training (HIIT). However, a drawback to their approach is the need for multiple sensors to collect this valuable information.

IoT systems and wearable sensors, as explored by Uday et al. (2018) [38], offer real-time monitoring and data transmission to cloud-based platforms. These systems, while effective in monitoring stress levels and physiological data, confront issues associated with the use of numerous sensors. These diverse research avenues provide promising insights into the capabilities of multiple sensors and challenges related to their practical implementation in real-world scenarios.

In a study with a focus on monitoring the health and activities of older adults, Bourke et al. (2017) [39] conducted research that involved two distinct settings: a controlled laboratory environment and the participants' daily routines. The participants were equipped with up to 12 sensors to capture their movements, and high-speed video recordings were used to track their motions. While this methodology yielded valuable insights into the health and activities of older individuals, it also brought to the fore concerns regarding privacy, particularly due to the use of video recording. Furthermore, the feasibility of relying on a multitude of sensors for real-life applications in the context of older adults' health monitoring deserves consideration.

The deployment of multiple sensors on the body for activity recognition presents a complex trade-off. While Aly et al. (2015) achieved accurate real-time activity monitoring using off-the-shelf 3D accelerometers, the practicality of using multiple body-worn sensors may not always be feasible [40]. Davoudi et al. (2021) delved into the placement of accelerometer devices on various body parts to estimate energy expenditure but acknowledged the limitations and impracticality of employing multiple sensors in everyday situations [41]. Wolff et al. (2018) explored head-mounted sensors for activity recognition, highlighting the challenges of practical, real-life applications, particularly for capturing subtle movements [42]. Similarly, Nam et al. (2013) argued against the practicality of simultaneously employing a camera and multiple sensors for activity detection in real-world scenarios [43]. Hoelzemann et al. (2019) synchronized signals from body-worn sensors but noted that the utilization of multiple sensors might not always be realistic [44]. Rokni et al. (2017) discussed the addition of new sensors at various body locations but also acknowledged the limitations of employing multiple body sensors [45]. Staudenmayer et al. (2009) employed a multi-sensor network, recognizing that the use of multiple accelerometers for activity detection may impose constraints on real-world practicality [46]. These diverse studies collectively underscore the inherent tension between the pursuit of high accuracy in activity recognition and the practical challenges associated with deploying multiple sensors on the body in real-world scenarios.

3.3 Usage of high energy consuming sensors like Heart Rate and GPS

In the field of physical activity monitoring research, the utilization of high-energy intensive sensors has presented a persistent challenge when it comes to monitoring physical activity over extended periods. Researchers from various categories have extensively explored this issue, shedding light on the limitations and practical challenges associated with these high-energy sensors. Here's a more detailed expansion of the insights and findings from the mentioned studies:

Precise Energy Expenditure Estimation: Two notable studies, Costa et al. (2015) [47] and Suh et al. (2017) [48], have tackled the challenge of estimating energy expenditure during physical activities. Costa et al. employed a combination of GPS and heart rate monitors, aiming to provide highly accurate estimates. While their approach did yield precise results, it wasn't without its challenges. Notably, the energy demands of heart rate monitors posed practical issues, potentially limiting the viability of this method for long-term activity tracking. Suh et al., on the other hand, explored various physical activities using accelerometers alongside high-energy sensors such as electrocardiograms. Their findings emphasized the limitations of continuous and extended activity tracking, where energy consumption emerged as a significant constraint.

Pedometers and Heart Rate Sensors: Freedson et al. (2000) [49] and Eckard et al. (2019) [50] have contributed to the discourse on precise physical activity tracking, primarily utilizing pedometers and heart rate monitors. Freedson's work involved the use of a pedometer and a heart rate sensor, but it became evident that continuous monitoring with energy-intensive sensors like heart rate monitors posed practical challenges. Eckard turned to heart rate sensors as a means to measure physical activity during exercises, yet similar to Freedson's findings, the use of high-energy sensors for prolonged activity detection was found to be unfeasible.

Wearable Devices for Monitoring Older Adults: Schrack et al. (2018) [51] focused on monitoring physical activity in older adults using wearable gadgets such as physical activity trackers and heart rate monitors. While wearable devices have shown promise in tracking activity, Schrack et al.'s research revealed that the energy requirements of these sensors could present hurdles for long-term monitoring, particularly in the context of older adults who may require extended tracking due to health considerations.

Fatigue Assessment: Allik et al. (2022) [52] explored the assessment of fatigue with heart rate sensors. Their work brought to light the limitations associated with using high-energy sensors like heart rate monitors for extended activity monitoring. These limitations included not only practical issues but also the potential discomfort for individuals undergoing extended monitoring.

Combining Heart Rate and GPS: In a study conducted by Michanikou et al. (2023) [53], individuals were equipped with smartwatches containing both heart rate monitors and GPS for physical activity tracking. While this combination seemed promising for precision and versatility, Michanikou questioned the practicality of using energy-consuming sensors like GPS and heart rate monitors for extended monitoring. The power demands of these sensors might pose challenges in terms of battery life and, therefore, continuous monitoring.

Sports Monitoring: In a research paper by Grigoroiu et al. (2021) [54], a study was conducted with a cohort of six female athletes, each equipped with personalized monitoring tools. These tools included wristwatches and chest belts designed for heart rate measurement. Additionally, athletes maintained dedicated Polar Flow accounts, allowing the seamless transmission of data captured by their wristwatches. The study utilized a Polar M400 Heart Rate Monitor to measure heart rate in beats per minute accurately. The accompanying software played a pivotal role in enabling athletes to amass a wide array of data during various activities, including physical training sessions, tennis practices, and routine daily activities. Notably, the data was represented as percentages, providing a comprehensive overview of their daily exertions and progress. However, it's worth noting that the use of energy-intensive sensors, like heart rate sensors, for extended activity monitoring may present practical challenges.

In addition to the above categories, several other studies have explored physical activity monitoring using high-energy sensors. For instance, Eskandari (2022) [55] evaluated heart rate responses using wearable electrocardiogram recorders, achieving notable classification results but encountering challenges due to sensor energy demands. Charvatova (2017) [56] correlated heart rate with changes in altitude using GPS and heart rate sensors, highlighting the impracticality of energy-draining sensors like GPS for extended monitoring.

Allahbakhshi (2020) [57] combined GPS and accelerometer data for physical activity classification, recognizing the challenges posed by energy-consuming sensors like GPS for prolonged monitoring. Nakanishi (2015) [58] used triaxial accelerometers and heart rate monitors to classify physical activities, acknowledging the difficulties of long-term monitoring with high-energy sensors like heart rate monitors.

Muggeridge (2021) [59] discussed identifying exercise types using heart rate monitoring devices but noted the limitations of relying on energy-draining sensors like heart rate monitors for extended periods. In the scholarly work detailed in the research paper by Garcia (2020) [60], the dataset under examination is an amalgamation of sensor-generated measurements harnessed from a smartphone. The encompassed sensors include the accelerometer, gyroscope, magnetometer, and GPS. These measurements are purposefully associated with four specific activity categories: inactivity, general physical activity, walking, and driving. Nevertheless, the endeavor to sustain prolonged activity tracking encounters significant hurdles when contingent upon the utilization of high-energy sensors such as GPS. Koffman et al. (2023) [61] have embarked upon a mission to quantify physical activity using easily accessible wearable devices, with a particular focus on monitoring heart rate and step count. The central objective of their study revolves around the identification of clinically significant subgroups among individuals who have undergone the experience of a stroke, but continuous monitoring with energy-consuming sensors like heart rate monitors presented challenges. Yan et al. (2014) [62] introduce the Dynamo lifestyle intervention, a comprehensive program strategically crafted to stimulate physical activity and diminish sedentary behavior among children manifesting cardio-metabolic risk factors. The study conducted by the authors encompassed the real-life tracking of children's movement, activity levels, and heart rates over a span of one week. The dependence on energy-intensive sensors, with heart rate monitors as the prominent example, engenders substantial complexities in the context of continuous, long-term physical activity monitoring. An in-depth examination was conducted by Li (2022) [63] on ten distinct dance teaching methodologies, with a particular emphasis on integrating heart rate monitoring in tandem with acceleration motion sensors. This integrated approach allowed the researchers to quantify the energy expenditure associated with sports dance instruction, offering a precise and comprehensive measurement framework, revealing limitations in relying solely on high-energy sensors like heart rate monitors for extended activity tracking. An innovative approach by Floris et al. (2020) [64] to physiological monitoring was investigated. The study involved the observation of thirty healthy participants in various resting positions, including sitting, standing, and supine. During these observations, participants were equipped with accelerometers, gyroscopes, and a virtual reality (VR) headset for a duration of 30 seconds. Mean heart rate (HR) estimation was derived from a 1-lead electrocardiogram (ECG). The primary limitation of this study is the reliance on multiple devices, which could introduce discomfort and mobility issues for participants, as well as the method's inherent focus on heart rate estimation, making it less versatile for the analysis of diverse physical activities, necessitating caution when interpreting findings from a data science standpoint. Doddabasappla (2021) [65] classified cough signals using a tri-axial accelerometer sensor but highlighted concerns about battery consumption due to the higher data collection frequency.

In summary, these studies collectively underscore the challenge of utilizing high-energy intensive sensors for prolonged and continuous activity monitoring. While these sensors have demonstrated the precision of data collected, their practicality in real-world scenarios is limited, emphasizing the need for more energy-efficient sensor technologies to overcome these challenges. Developing sensors that strike a balance between accuracy and energy efficiency is crucial to advancing the field of physical activity monitoring, making it more practical for real-world, long-term applications.

3.4 Usage of High Frequency data

Numerous groundbreaking studies have made remarkable strides in classifying various human modes of locomotion and activities within the domain of activity recognition utilizing high-frequency sensors. These studies harness an array of sensor technologies, providing us with valuable insights into the intricacies of high-frequency data capture and analysis.

High-Frequency Accelerometer-Based Recognition: Junker et al. (2004) [66] and Figueira et al. (2016) [67] emerged as trailblazers in the realm of accurate classification of human locomotion using body-worn acceleration sensors. Junker et al. prioritize the classification of different human modes of locomotion by leveraging body-worn acceleration sensors. Their results demonstrate that precise classification can be achieved with a relatively high data sampling frequency of 20 Hz and a modest 2-bit resolution. While Figueira et al. they achieved a remarkable accuracy of 94.5% in human activity detection. This was accomplished through the utilization of two sensors: the accelerometer, which was sampled at a rate of 30 Hz, and the barometer, which was sampled at 5 Hz. Nonetheless, it's important to highlight that the increased data sampling rate presents a significant challenge in the form of heightened power consumption, which could potentially have implications for battery life in real-world applications.

Integration of Multiple Sensors: Wan et al. (2020) [68] embarked on a novel path by integrating a medley of sensors into a real-time data transfer platform, enabling data retrieval at frequencies reaching up to 100 Hz. While their primary focus revolved around recognizing motion sequences, this ambitious study illuminated challenges related to processing demands, battery drainage, and the practicality of employing multiple sensors. Bruno et al. (2014) [69] and Reiss et al. (2012) [70] ventured into the high-frequency sensor data arena by collecting information from multiple body-worn Inertial Measurement Units (IMUs) at an impressive 100 Hz. In the research conducted by Bruno et al., a dataset derived from nine subjects participating in 18 distinct activities was employed. These subjects were equipped with three Inertial Measurement Units (IMUs) sampled at a rate of 100 Hz, in addition to heart rate monitors operating at 9 Hz. Similarly, Reiss et al. curated a dataset encompassing 18 distinct activities executed by 9 subjects. Each subject wore three Inertial Measurement Units (IMUs) situated on the wrist, chest, and ankle, with data recorded at a sampling rate of 100 Hz. This dataset has been made publicly accessible and was subsequently employed in four classification tasks employing standard processing techniques and five different classifiers. However, it's important to understand that the high data collection frequency can genuinely raise concerns about potential data disturbances, greater demand on battery life, and the energy requirements of heart rate sensors. These are vital considerations for data scientists working with sensor-based datasets.

Simultaneously, Stisen et al. (2015) [71], Banos et al. (2014, 2015) [72] [73], and Chavarriaga et al. (2013) [7] expanded the horizons of data capture by introducing smartwatches, smartphones, and a variety of body-worn and environmental sensors. In the study by Stisen et al., a distinctive data collection method involving four smartwatches and eight smartphones situated at the waist or in pouches was employed with a notably high sampling frequency. Banos and colleagues, in their work, presented mHealthDroid, an open-source Android implementation for biomedical app development. They utilized Shimmer wearable devices on ten subjects at a high 50 Hz frequency. In another paper by Banos et al., a precise activity recognition model was crafted and successfully validated under both offline and online conditions. However, it's worth noting that they adopted a notably high data sampling frequency of 50 Hz and incorporated multiple Shimmer wearable sensors positioned on the right wrist, left ankle, and chest. The challenge here lies in managing the higher frequency and the integration of multiple body sensors. Chavarriaga et al. employed an array of sensors, including body-worn sensors at various positions, such as the upper body, hip, leg, and shoes, as well as environmental sensors. While this comprehensive approach provides rich data sources, it's important to note that using multiple sensors in real-world activity recognition scenarios poses certain challenges and limitations.

High-Frequency Smartphone-Based Recognition: Hnoohom et al. (2017) [74], Reyes et al. (2014) [8], Khan et al. (2018) [75], and Zhu et al. (2019) [76] ventured into the realm of high-frequency data with recognition systems primarily reliant on smartphones. Across these studies, a common thread unravels as authors navigate the intricacies of high-frequency data sampling and the integration of multiple sensors or devices. Hnoohom's data collection involved the placement of a smartphone in both the left and right trouser pockets, and data was sampled at a rate of 50 Hz. The primary objective was to identify and classify six physical activities: standing, sitting, lying down, walking, walking upstairs, and walking downstairs. The experimental findings shed light on the synergistic potential between accelerometer and gyroscope data, which led to significant improvements in the accuracy of activity recognition. Moreover, the integration of ensemble learning techniques proved instrumental in enhancing accuracy, culminating in an impressive 91.16% accuracy rate. In a similar vein, Reyes and their research team introduced an online Human Activity Recognition (HAR) system with an impressive 50 Hz sampling rate. This frequency, although valuable for capturing nuanced details of human activity and postural transitions, came with a distinct trade-off. The higher data collection frequency invariably meant more intensive usage of smartphone batteries. This challenge was significant, as it could potentially limit the duration such systems could operate in the real world, where extended data collection periods are often necessary. The work by Khan et al. introduced a transductive transfer learning model that sought to minimize the reliance on extensive labeled training data. Data collection involved two smartphones and a smartwatch, all synchronized at a 50 Hz sampling rate. Here, too, the emphasis on high data frequency and the incorporation of multiple sensors presented its unique set of challenges. The data processing requirements escalated, and the energy consumption surged. These challenges could pose substantial hurdles in practical, real-world applications. Lastly, Zhu and colleagues proposed a human activity recognition framework employing convolutional neural networks (CNNs) in tandem with smartphone sensor data, encompassing the accelerometer, gyroscope, and magnetometer. Data was diligently collected at a 50 Hz frequency, with smartphones positioned in different configurations, such as handheld, trouser pocket, and backpack. Once again, the dual challenges of high data frequency and the involvement of multiple sensors emerged as pivotal limitations. While the richer dataset enabled sophisticated analysis, the increased data load and battery consumption could make deployment in real-world settings a complex endeavor. These limitations underscore the critical importance of meticulously balancing the pursuit of high-quality data with the practical constraints of battery life and data processing capabilities in sensor-based research and applications.

Deep Learning Approaches: Dua (2021) [77], Chen et al. (2015) [78], and Wang et al. (2021) [79] dived deep into high-frequency datasets sampled at 100 Hz to deploy deep learning methodologies for activity classification via smartphones. They opened up new possibilities by circumventing the need for manual feature extraction. In the study conducted by Dua, a high-frequency dataset operating at 100 Hz was employed to construct a Deep Neural Network model. This model cleverly integrated components of Convolutional Neural Networks and Gated Recurrent Units, allowing for automatic feature extraction and accurate activity classification. An intriguing aspect of this approach is the utilization of raw sensor data, which eliminates the need for manual feature extraction. However, it's crucial to remain mindful of potential data noise and the increased battery consumption associated with the elevated sampling frequency used in this study. Similarly, in the research paper authored by Chen et al., the authors proposed a human activity recognition method leveraging acceleration data and Convolutional Neural Networks (CNN). Data was collected from smartphones situated at both waist and trouser pockets, sampled at a high frequency of 100 Hz. The emphasis on a higher data frequency and the use of multiple body sensors introduced distinct challenges in data processing and energy consumption. In the work by Wang et al., a recurrent attention network (RAN) was introduced to address the complex tasks of sequential, weakly labeled multi-activity recognition and location estimation. Data collection involved iPhones sampled at a rate of 50 Hz. However, it's vital to acknowledge that the increased data frequency, although beneficial for the tasks at hand, may potentially impact battery life, a crucial consideration in the practical deployment of such systems. As data scientists, balancing data quality and the constraints of real-world implementation remains a constant challenge in sensor-based research.

Wearable Device-Based Recognition: Sztyler et al. (2017) [80], Kruger et al. (2017) [81], and Bhat et al. (2020) [82] embark on the exploration of various aspects related to wearable devices and data collection. Sztyler et al. investigate the recognition of wearable device positions through a series of comprehensive experiments. These experiments involve the placement of devices on different parts of the body, with data sampled at a frequency of 50 Hz. Similarly, in the study by Kruger and colleagues [81], the authors curate a dataset using wearable devices positioned at various locations on the body. Notably, the data collection in this context occurs at a higher sampling rate of 110 Hz. In a separate study by Bhat et al. [82], the authors introduce the w-HAR dataset, a resource containing labeled data from 22 users, for the purpose of classifying seven distinct activities. They leverage design space exploration to optimize a neural network architecture for activity classification to accomplish this task. Furthermore, two online learning algorithms are applied to adapt the classifier for users not initially included during the design phase. Data collection in this study is noteworthy for its exceptionally high frequency, recorded at 250 Hz, and originates from a wearable device placed on the right ankle. However, it is of utmost importance to acknowledge the challenges stemming from this intensified data collection frequency and the incorporation of multiple body sensors. These challenges encompass effective data processing and mindful energy consumption, both essential considerations in the realm of practical applications.

In the paper authored by Laput et al. [83], the development of a custom smartwatch kernel is documented, which allowed for a remarkable increase in the accelerometer's sampling rate, reaching an impressive 4000 Hz. This enhancement was instrumental in enabling the recognition of a wide array of hand gestures, thus enhancing expressive input capabilities. However, it's crucial to be cognizant of the challenges introduced by this high sampling frequency, notably the potential for data noise and increased battery usage, aspects that warrant careful consideration in practical applications.

Similarly, in the research paper by Anguita and his colleagues [84], an Activity Recognition database was meticulously curated, involving 30 subjects engaging in various daily activities while equipped with smartphones containing embedded inertial sensors positioned at the waist. The study also integrated acceleration sensors at 12 different positions, capturing data at a relatively high frequency of 30 Hz. While the results discussed the effectiveness of a multiclass Support Vector Machine (SVM), it's imperative to acknowledge the challenges that emerge from the higher data frequency and the incorporation of multiple body sensors, both of which can pose substantial complexities in data processing and analysis.

In the article authored by Chatzaki et al. [85], a comprehensive assessment focused on the utilization of a smartphone's acceleration sensor for human activity and fall detection. This assessment spanned 12 distinct activities of daily living (ADLs) and four types of falls. The data collected was sampled at an approximate frequency of 100 Hz, which introduced noteworthy challenges linked to higher battery depletion due to the elevated data collection frequency. It's crucial to address these considerations, particularly in scenarios where prolonged data collection is essential.

In the research paper authored by Weiss [86], data collection employed both smartphones and smartwatches, featuring a sampling rate of 20 Hz. Smartphones were placed in the right pants pocket, while smartwatches adorned the subjects' dominant hand. Nevertheless, it's essential to recognize the challenges associated with the increased data frequency and the integration of multiple body sensors, which can introduce complexities that demand careful management and analysis.

In summary, these comprehensive studies have significantly advanced the field of activity recognition by harnessing the capabilities of high-frequency sensor data. They have concurrently underscored the persistent challenges linked to the management of increased processing demands, the mitigation of data noise, and the preservation of battery life—crucial considerations for practical applications in real-world scenarios.

3.5 Latest research trends

Researchers have made significant advancements in Activity Recognition. This field has gained popularity in health care, smart homes, sports, game controls and elderly care. [87] focused on distinguishing basic actions and transition actions (e.g., standing and sitting) in HAR. Their primary contribution is the introduction of hybrid deep learning models with a combination of CNN (convolutional neural networks) and LSTM (Long Short-Term Memory) networks. This model uses CNN for local feature extraction and LSTM for time-dependent relationships. The research later presents superiority over other models using similar datasets. Recent research by [88] introduces a HAR system using smartphone and smartwatch data to identify 18 different physical activities. It combines a CNN and BiGRU (bidirectional gated recurrent unit) for feature extraction. It uses SLFN (single-hidden-layer feed-forward neural network) with RELM (regularised extreme learning machine) for activity recognition. This system significantly enhances activity detection, providing a reliable method to monitor mental and physical health and improving elderly care. Another research [89] compares various deep learning techniques for HAR using smartwatch data. They evaluate three methods for recognising daily activities. Later, they showcase how DeepConvLSTM architecture is the most effective. Results showcase the potential to understand user routines, benefitting applications in monitoring daily tasks and physical tasks. Another research [90] in the domain of health monitoring explores the impact of physical activity (PA) on myopia. They developed a smartwatch-based model to differentiate indoor and outdoor PA activities accurately. They merged light intensity sensing with accelerometer data for precise PA segmentation. The study highlights SVM's effectiveness in handling high-dimensional data and its performance in PA classification compared to other machine learning models.

The progress in activity recognition technology promises enhanced quality of life by enabling more efficient health monitoring. As researchers continue to explore and refine these technologies, activity recognition is projected to become an integral part of various fields, driving innovation and improving everyday life.

3.6 Key limitation and challenges of Related Work

Related work in NEAT activity detection showcases several limitations. Researchers working with wearable sensor technology raise a significant challenge with sensor accuracy and calibration specially those who are working with multiple devices at multiple positions. The major challenge is the inconsistency of sensor calibration across those multiple devices and brands, which leads to discrepancies in data accuracy. Our objective is to use a single sensor device source for data collection to minimise the challenge with data collection such as calibration. Multiple devices and multiple positions pose practical challenges too which are to be removed by use of a single sensor.

Multiple studies highlight the challenge of coping with the battery life of wearable devices. Data collection done on a higher frequency majorly reduces the battery life, which in turn limits the continuous monitoring of a device. Frequent recharging of the device defeats the device's practicality and negatively affects user compliance, as the user is less likely to wear the device for a longer period, reducing the reliability of the data collected. Use of low-frequency data limits the diversity and richness of features that can be extracted and in turn impacts the performance of the activity recognition models. Our objective is to work with as low as possible frequency without compromising accuracy of the data.

Differences in user behaviour, device usage and environment challenges introduce biases and reduce the generalizability of findings, posing another challenge. Our objective is to work in a semi controlled environment so that the training data for the model is good.

Continuous monitoring with vision-based systems raises significant privacy concerns due to the storage and processing of video data containing personal information. To address these issues, we need secure, efficient, and user-friendly solutions that protect privacy and build user trust. Our objective is not only to preserve individuals' privacy but also to use lightweight data, such as text-based or structured data like csv files, which is easier to process compared to audio or video data.

Chapter 4

NEAT Activity Detection using Smartwatch at Low Sampling Frequency

4.1 Problem Scope

Our work aims to build a classification model that can differentiate between typical Non-Exercise Activity Thermogenesis (NEAT) activities within a home environment. Focused activities include cooking, sweeping, mopping, walking, climbing up, climbing down, and non-NEAT activities (e.g., watching television and desk work). Due to COVID-19 restrictions and curfews, we choose to perform activities that could be easily performed in-home setup and are part of daily routine. The challenge lies in building a model capable of working with low-frequency data (1Hz), as NEAT activities are not easily separable at this granularity. The primary objective is to develop a model discerning typical home activities using data from smartwatch sensors, specifically the accelerometer and gyroscope [91]. Targeting seven activities, the classifier aims to operate with data sampled at a low frequency of 1Hz. The model's value extends to health monitoring applications, especially during pandemic scenarios where individuals are confined to home environments [92]. Such applications could track NEAT activities, crucial for maintaining physical activity during a lockdown. The ideal NEAT activity recognition solution should utilize readily available hardware (smartphone or smartwatch) and be energy-efficient, considering the likelihood of NEAT activities spreading throughout the day. This work presents a smartwatch-based solution for NEAT activity recognition, emphasizing energy efficiency achieved by reducing sensor sampling rates ([93]). Table 4.1 shows battery consumption at different sampling rates, based on our experiments conducted at various frequencies. We performed the data collection process for 5 hours for each frequency. The findings discussed here are further elaborated in [94].

Frequency	**Battery**
10 Hz	93%
5 Hz	86%
1 Hz	78%

Table 4.1: Total Battery consumed in 5 hours

While low sampling frequency ensures good battery life, it is essential to note that *activities focused here cannot be easily discerned in low-frequency data.* Table 5.1 illustrates the total accuracy (for our activities of interest) obtained by different classification techniques on data sampled at 10Hz, 5Hz, and 1Hz. As the table shows, the performance of the classifiers drops down significantly as the sampling frequency is reduced to 1Hz.

Frequency	**1Hz**	**5Hz**	**10Hz**
KNN	74	85	89
Multi-layer Perceptron	71	82	89
SVM	78	87	92
Logistic	61	71	81
Random Forest	47	53	61
Naive Bayes	60	67	73
XGBoost	81	89	94
***Our Model* 4.1**	87	94	96

Table 4.2: Accuracy of different classifiers on 1Hz, 5Hz, and 10Hz sampled data (2 sec windows 50% overlap.)

4.1.1 Challenges

One challenge in our research was dealing with the limited feature set caused by fewer data points at low sampling frequencies. This makes it harder to extract a wide range of detailed features for recognizing activities. High-frequency data provide more opportunities for capturing diverse and detailed features. To tackle this, we focused on extracting strong features using ECDF and statistical features (mode, max, median, lower quartile and standard deviation) over varied window lengths (e.g., 2, 4 and 6 seconds) to ensure we had enough information for accurate classification.

Another challenge was distinguishing between similar activities, which can be tricky with low-frequency data. For example, slow walking and standing might look similar in accelerometer readings at 1Hz. To address this, we used a hierarchical model that classifies activities step by step, gradually narrowing down the options and improving accuracy. By using multiple classifiers in a hierarchical structure, the system could better handle similar activities and improve overall recognition performance.

4.2 Scope of chapter

This chapters objective is to explore the potential of using low-frequency (1Hz) smartwatch data, specifically from accelerometer and gyroscope sensors, to accurately detect Non-Exercise Activity Thermogenesis (NEAT) in a home environment. Given the inherent challenges of working with low-frequency data, the research is guided by the following objectives:

1. To determine whether low-frequency (1Hz) data from smartwatch sensors,

specifically accelerometer and gyroscope, can accurately distinguish between various NEAT and non-NEAT activities.

2. To improve the classification accuracy of a hierarchical model compared to traditional flat classification models in identifying activities from low-frequency smartwatch data.

3. To identify which features derived from accelerometer and gyroscope data, such as mode, max, median, lower quartile, and standard deviation, are most influential in distinguishing NEAT activities at low sampling rates.

4. To assess how the overlap in data window length impacts the accuracy of activity classification at low sampling rates.

4.3 Contributions

1. **Hierarchical model to identify 7 activities:** Our hierarchical model (Figure 4.1) is novel since existing work on these activities does not address a hierarchical approach, and even if a hierarchical model exists in prior research like [95] and [96], their models do not include the specific set of activities (our chosen seven) that we have considered. Our major contribution is the development of a hierarchical model that can accurately identify the following seven different types of activities typically performed in a home setting: (a) cooking, (b) sweeping, (c) mopping, (d) walking, (e) climbing up, (f) climbing down, and (g) non-NEAT activities.

2. **Low Frequency:** Our proposed model (Figure 4.1) can work with data sampled at low frequency (1Hz). Although there has been existing work done using high-frequency sampling rates, we are focusing on achieving the best accuracy despite working with low sampling frequencies. Our focus is to maximize accuracy while minimizing the sampling frequency, thus making it an energy-efficient solution. Additionally, we try to use battery-efficient sensors such as accelerometers and gyroscopes, instead of battery-draining sensors like GPS and heart rate monitors.

3. **Real-Data:** We experimentally evaluated (trained and tested) our proposed model (Figure 4.1) on real data collected using a smartwatch (Due to COVID restrictions and in-home setting we could choose limited number of volunteers). Additionally, we compared our model's accuracy with many other models, and our model's accuracy proved to be the best among them.

4.4 Our Proposed Approach

4.4.1 Inspiration

This chapter is influenced by the work I presented in [97], where a hierarchical model was created for classifying activities during a metro journey. Initially employing flat models for efficient activity classification, we extend this approach to detect NEAT (Non-Exercise Activity Thermogenesis) and non-NEAT activities using smartwatch sensors. This chapter follows a similar methodology to adapt the hierarchical model's success in metro journey activity classification to the nuanced context of distinguishing NEAT and non-NEAT activities through smartwatch sensor data.

4.4.2 Preprocessing and Features

The raw sensor values are in the form of time series data $\mathcal{T}$, which are divided into a set of overlapping windows ($\mathcal{W}$) for the purpose of training. Two temporally adjacent windows w_i and w_{i+1} in $\mathcal{W}$ can have a degree of overlap defined by the parameter θ. We set the value of θ to be 0 and 0.5. When $\theta = 0$, there is no overlap. When $\theta = 0.5$, there is 50% overlap among the consecutive windows. This overlapping of 50% essentially creates another data point between two otherwise non-overlapping (but temporally adjacent) windows. This, in turn, helps in learning a more robust model by reducing the effect of outlier data. In fact, our experiments also show that all models obtain higher accuracy in the case of overlapping windows (details in Section 5.5). Each window is of length ω. We considered ω to be 2, 4, and 6 seconds.

Accelerometer Individual Axis Features: Accelerometer gives the output in all three dimensions, i.e., x-axis, y-axis, and z-axis. Therefore, for a specific window $w_i \in \mathcal{W}$, there will be a three-time series of acceleration values consisting of a_x, a_y, and a_z. We apply five statistics features on each three-time series in a window – (1) mode, (2) max, (3) median, (4) lower quartile, and (5) standard deviation. Hence, we have 15 combinations possible, i.e., each axis (3) paired with each of the statistical features (5). Please note that we have considered the raw sensor values, and no filter was applied before sending it for calculating the statistical features.

Accelerometer Magnitude Features: This feature is found using the formula $a_{mag} = \sqrt{a_x^2 + a_y^2 + a_z^2}$. So given the individual axis of the accelerometer, we can find out its magnitude for a given time instant. For this feature as well, we find out the same five statistical features (mode, max, median, lower quartile, and standard deviation) for each window $w_i \in \mathcal{W}$. There is no filter used before finding the magnitude here as well.

Gyroscope Individual Axis Features: Gyroscope sensor gives the angular motion speed of the device worn or carried by a person. Just like an accelerometer, a gyroscope also gives a three-dimensional output for the x-axis, y-axis, and z-axis. We denote the resulting time series as g_x, g_y, and g_z. Along with accelerometer values, the orientation of the device given by the gyroscope sensor is also useful in detecting the user's motion. We find the previously mentioned five statistical features for each of the axes given in a time series window having a total of 15 statistical features.

Gyroscope Magnitude Features: Just like accelerometer magnitude, we also find out the overall magnitude of the gyroscope sensor along the individual axis as $g_{mag} = \sqrt{g_x^2 + g_y^2 + g_z^2}$. After computing the magnitude, we preprocess it by finding out the five statistical features (mode, max, median, lower quartile, and standard deviation) for each

g_{mag} over the window length. Henceforth, we get five features from this category as well.

4.4.3 Proposed Hierarchical model

We demonstrate our proposed hierarchical learning model for distinguishing the seven NEAT activities, viz., Cooking, Sweeping, Mopping, Walking, Climbing up, Climbing down, and non-NEAT in Figure 4.1.

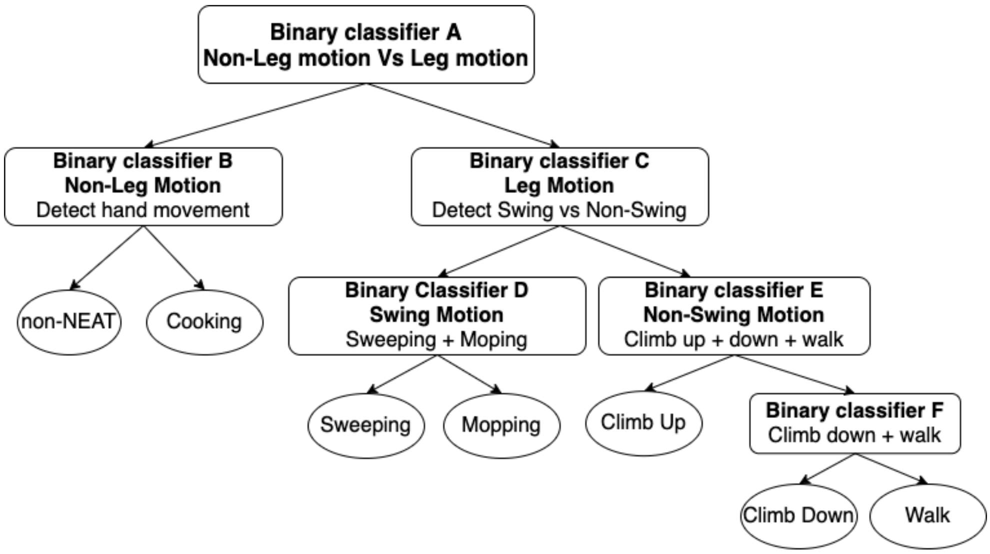


Figure 4.1: Proposed hierarchical model for distinguishing the NEAT and non-NEAT activities.

As shown in Figure 4.1, our proposed hierarchical model is a combination of various binary classifiers (A, B, C, D, E, and F), i.e., at each level, we are doing the classification only between two classes. The two classes were formed so that the most alike ones were clubbed together; they were separated from the rest of the ones. We first check if the action involves a change in the location, i.e., leg motion. If it does not, then we separate it at the root node itself. That is why our topmost level classifier "A" distinguishes "Cooking" and "non-NEAT" from the rest of the classes. We merge the Cooking and non-NEAT classes into one class and the rest into another. On the second level on the left-hand side, we separated the non-leg motion data and made a classifier "B", which distinguishes further into non-NEAT and Cooking. We know that there is hand movement in cooking with minimal leg motion, unlike non-NEAT with minimal or no hand movement involved. Hence, those are distinguishable among themselves.

In the second level on the right-hand side (Classifier "C"), we separate the hand swing motion from the non-swing motion. If there is a detection of swing motion, we move further to the left subtree and categorize it into Sweeping + Mopping. If there is a detection of non-swing hand motion, we move further to the right subtree and categorize it into the climb up + down + walk. Later, Classifier "D" classifies the data into sweeping and mopping. In the last level, we have distinguished climb up from Climb down + walk using classifier "E" because of the similarity in hand motion between the climb down and walk classes. Later, classifier “F” classifies into climb down and walk.

Implementation of the Learning Model: We have seven different class labels – 1 being the cooking class, 2 being the sweeping class, 3 being the mopping class, 4 being the walking class, 5 and 6 are climb up and climb down respectively, and 7 is the non-NEAT class. We take the training data, perform the masking, and send it through all the available classifiers. The classifier, that gives the best accuracy among all, is chosen as the final classifier for that level. At the root level, we first mask our training data into “17” vs “23456”. The data labels of “17” correctly classified are directly given to the classifier “B” for further classification between 1 and 7. Once the “23456” data is passed through the top-level Classifier “A” and correctly classified, it is masked again into “23” vs. “456”. Here, the wrongly classified ones are thrown away. The data labels of “23” correctly classified are directly given to a classifier “D” for further classification between 2 and 3. Once the “456” data is passed through the second level classifier “C” and correctly classified, it is again masked into “5” vs “46”, and the wrongly classified ones are thrown away. At the last levels, “4” and “6” are given to a classifier “F” for further classification. The best models with the best accuracy on training data are stored for the testing phase at each level.

After the data points reach a leaf node, they are no longer trained further. In the test phase, the data points in the form of windows are passed through each level classifier, and the final labels are matched with the actual labels. Hereafter, the confusion matrix is formed. Note that we discard those window points that do not pass through the correct labels during the training phase. For example, in the training phase, if the data points belonging to Cooking or non-NEAT are classified as “Leg motion”, then those data points are discarded before making a new masked training data of “Swing motion” vs. “Non-Swing motion.” We did this for all the lower-level classification since it ensures the quality of the final model without running into an overfitting problem. Moreover, those discarded data points were anyway not crucial in the training phase.

We used the following classical Learning algorithms at each level - (1) K-Nearest Neighbour with k=5), (2) Multi-Layer Perceptron (3-hidden layers), (3) Support Vector Machine (with RBF Kernel), (4) Random Forest, and (5) XGB. XGB has the maximum accuracy.

4.4.4 Model Selection Rationale

Different combinations of activity classes are grouped together in various splits in table 4.3. For instance, “123 456 7” means activities 1, 2, and 3 are in one group, 4, 5, and 6 are in another group, and 7 is in its own group. Each cell in the table shows the performance (likely accuracy) of the corresponding model for a specific class split. The performances are highlighted with colors to indicate their quality:

Dark Green: Highest performance. Light Green: High performance. Yellow: Moderate performance. Red: Low performance.

We observed that:

1. The XGB model stands out with consistently high performance across various class splits.
2. The KNN, MLP, and SVM models also show strong performance; therefore, we proceeded with these top four classifiers.
3. Logistic Regression, Random Forest, and Naive Bayes do not perform well for almost all splits.
4. The performance varies significantly depending on the class split, indicating that the choice of class grouping can substantially impact the effectiveness of the models.
5. Lastly and most importantly, the class split "17 23456" yields the highest accuracies across all models, making it the most effective grouping for this task. This insight led to the decision to proceed with this particular class split.

Table 4.3: Experimenting on various known classifiers by selecting multiple combinations for classes. Class 1 - Cooking, Class 2 - Sweeping, Class 3 - Mopping, Class 4 - Walking, Class 5 - Climb up, Class 6 - Climb Down, Class 7 - non-NEAT

Class splits	KNN	MLP	SVM	Logistic Regression	Random Forest	Naive Bayes	XGB
123 456 7	88%	87%	90%	77%	77%	73%	92%
12 3456 7	85%	86%	88%	77%	71%	76%	90%
123 4567	88%	88%	90%	80%	81%	70%	91%
1 23456 7	92%	92%	92%	88%	84%	83%	95%
1 23 456 7	85%	83%	86%	72%	68%	69%	88%
15 2346 7	86%	85%	87%	79%	82%	77%	89%
17 23456	94%	93%	85%	85%	86%	86%	95%

4.5 Experimental Analysis

Table 4.4: Total Training Data

Window Length	1Hz (50% overlap)	1Hz (0% overlap)
2 seconds	11335	5670
4 seconds	5670	2835
6 seconds	3780	1890

Training Data - Table 4.4 shows the number of instances belonging to each window length and overlapping percentage parameter. These numbers decrease as we increase the window length. This decrease occurs because we are expanding the time frame for which we want to determine the input statistical features and their corresponding output labels. We observed an improvement from two seconds to four seconds, but there was no significant improvement from four seconds to six seconds. Moreover, increasing the window length reduces the number of sampling data points; hence, we decided not to proceed with fewer sampling points. Additionally, if we deploy this model on a server in the near future, it will delay the output of class labeling.

4.5.1 Candidate Algorithms

We have compared our Hierarchical Model with the Flat Model. When we say Flat model, we refer to the in-built classifiers, which are readily available in common machine learning libraries (e.g., sklearn in Python). The flat models attempt to learn a single decision boundary amongst all our classes of interest. For the flat models, we chose the following classifiers - a) KNN (K-nearest neighbor with k=5) b) MLP (multi-layer perceptron with 3 hidden layers having 13 neurons each) c) SVM (support vector machine with kernel='rbf') d) Logistic Regression e) Random Forest f) Gaussian Naive Bayes g) XGB (Extreme Gradient Boosting).

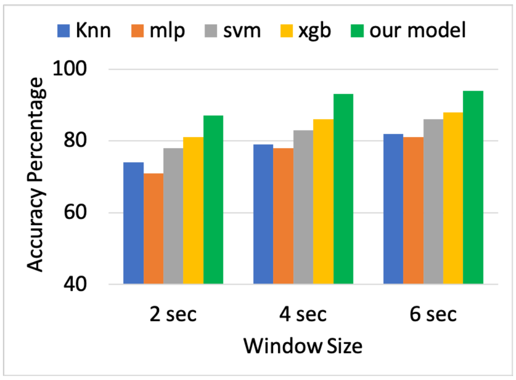
In the case of our hierarchical model, we tried different classifiers for each binary classifier mentioned in the tree (Figure 4.1) and chose the classifier that gives the best accuracy. In our implementation, we made this decision on the basis of the training accuracy as the test data is considered to be "hidden" by definition. In our experiments, we found out that in most of the cases, XGB was chosen at each level. We use python language for the implementation of our models.

4.5.2 Training & evaluation matrix

We have used window lengths of 2sec, 4sec, and 6sec in our experiments. The window overlap parameter θ was varied across 0.5 and 0 (i.e., no overlap amongst temporally consecutive windows). For a given set of windows $\mathcal{W}$ with its corresponding θ values, we divide it into a train and a test data set in the ratio of $4 : 1$, i.e., 80% of the data is allocated for training, and the rest 20% is allocated for testing. To get reliable results, we divided our given dataset into training and test portions 10 times (randomly). Following this, we trained (and tested the learned model) on each of the previously mentioned 10 splits. Finally, we report the average of F1-scores and Accuracies obtained across those 10 test datasets.

4.5.3 Consequences of varying window length on final accuracy

Figure 4.2 and Figure 4.3 illustrate the results of this experiment. The test accuracy of our model is shown in the form of bar graphs along with the rest of the flat models. In this experiment, we tried window lengths of 2sec, 4sec, and 6secs. Overlap parameter θ was taken as 0 and 0.50. Figure 4.2 shows the results corresponding to the case where all features were used in training. Whereas, Figure 4.3 displays the results corresponding to cases where only the accelerometer was used (Figure 4.3a and Figure 4.3b), or only the gyroscope was used (Figure 4.3c and Figure 4.3d). It is important to note that, the performance of all the models decreased when only gyroscope features were used. Overall, we observed that our proposed model outperformed the alternative approaches consistently. Moreover, our experimental results also indicate that all models perform better when windows overlap (i.e., $\theta = 0.50$). A similar increase in performance with an increase in overlap has also been reported in other works [97]. This is possibly due



(a) All features were used. $\theta = 0.5$, frequency = 1Hz

(b) All features were used. $\theta = 0.5$, frequency = 5Hz

(c) All features were used. $\theta = 0$, frequency = 1Hz

(d) All features were used. $\theta = 0$, frequency = 5Hz

Figure 4.2: Effect of window length on overall accuracy when all features are used.

to the fact that in the case of overlapping windows, the effect of outliers is reduced as "good data" sort of "spawns" more "good data" (inadvertently) through the process of overlapping the windows.

Table 4.5: F1-scores of Our Hierarchical Model and Flat Classifier XGBoost (XGB) for $\theta = 0.50$ and frequency = 1Hz and window length = 2seconds

	F1-Scores of Our Model							
	Cook	Sweep	Mop	Walk	CU	CD	non-NEAT	Overall Accuracy
All Features	0.8	0.9	0.86	0.83	0.81	0.9	0.99	87%
Only Accelo	0.76	0.84	0.82	0.75	0.77	0.84	0.98	82%
Only Gyro	0.34	0.78	0.8	0.57	0.73	0.71	0.96	71%
	F1-Scores of XGB Model							
	Cook	Sweep	Mop	Walk	CU	CD	non-NEAT	Overall Accuracy
All Features	0.81	0.83	0.86	0.78	0.68	0.68	0.96	80%
Only Accelo	0.75	0.68	0.75	0.69	0.53	0.58	0.97	71%
Only Gyro	0.37	0.57	0.65	0.55	0.46	0.44	0.93	57%

(a) Only Accelerometer, $\theta = 0.5$, frequency = 1Hz

(b) Only Accelerometer, $\theta = 0.5$, frequency = 5Hz

(c) Only Gyroscope, $\theta = 0.5$, frequency = 1Hz

(d) Only Gyroscope, $\theta = 0.5$, frequency = 5Hz

Figure 4.3: Effect of window length on overall accuracy when only accelerometer or gyroscope features are used.

Table 4.6: F1-scores of Our Hierarchical Model and Flat Classifier XGBoost (XGB) for $\theta = 0.50$, frequency = 1Hz, window = 4 seconds

	F1-Scores of Our Model							
	Cook	Sweep	Mop	Walk	CU	CD	non-NEAT	Overall Accuracy
All Features	0.9	0.96	0.93	0.89	0.91	0.93	0.99	93%
Only Accelo	0.82	0.88	0.84	0.82	0.82	0.89	0.98	87%
Only Gyro	0.45	0.89	0.85	0.59	0.74	0.79	0.99	76%
	F1-Scores of XGB Model							
	Cook	Sweep	Mop	Walk	CU	CD	non-NEAT	Overall Accuracy
All Features	0.84	0.88	0.86	0.85	0.75	0.79	0.98	85%
Only Accelo	0.81	0.83	0.82	0.81	0.65	0.7	0.96	80%
Only Gyro	0.46	0.65	0.71	0.58	0.53	0.51	0.95	63%

4.5.4 F1-scores of individual classes

The illustration of f1-scores for each of the classes can be seen in Table 4.5, Table 4.6, Table 4.7, Table 4.8, Table 4.9 and Table 4.10. The upper section in table represents our model, and the lower section represents the flat model. We have shown only XGBoost

Table 4.7: F1-scores of Our Hierarchical Model and Flat Classifier XGBoost (XGB) for $\theta = 0$, frequency = 1Hz, window = 2 seconds

	F1-Scores of Our Model							
	Cook	Sweep	Mop	Walk	CU	CD	non-NEAT	Overall Accuracy
All Features	0.75	0.84	0.83	0.71	0.73	0.88	0.92	83%
Only Accelo	0.71	0.63	0.71	0.68	0.7	0.83	0.94	75%
Only Gyro	0.41	0.68	0.78	0.51	0.68	0.76	0.93	69%
	F1-Scores of XGB Model							
	Cook	Sweep	Mop	Walk	CU	CD	non-NEAT	Overall Accuracy
All Features	0.78	0.79	0.82	0.65	0.68	0.63	0.91	76%
Only Accelo	0.72	0.58	0.74	0.73	0.48	0.55	0.89	71%
Only Gyro	0.34	0.52	0.61	0.57	0.38	0.39	0.87	53%

Table 4.8: F1-scores of Our Hierarchical Model and Flat Classifier XGBoost (XGB) for $\theta = 0$, frequency = 1Hz, window = 4 seconds

	F1-Scores of Our Model							
	Cook	Sweep	Mop	Walk	CU	CD	non-NEAT	Overall Accuracy
All Features	0.8	0.89	0.89	0.76	0.78	0.92	0.99	86%
Only Accelo	0.74	0.69	0.74	0.71	0.74	0.87	0.98	78%
Only Gyro	0.44	0.73	0.82	0.55	0.72	0.79	0.98	72%
	F1-Scores of XGB Model							
	Cook	Sweep	Mop	Walk	CU	CD	non-NEAT	Overall Accuracy
All Features	0.81	0.82	0.87	0.7	0.7	0.68	0.96	80%
Only Accelo	0.76	0.61	0.78	0.75	0.52	0.58	0.92	73%
Only Gyro	0.36	0.56	0.64	0.6	0.4	0.4	0.9	56%

Table 4.9: F1-scores of Our Hierarchical Model and Flat Classifier XGBoost (XGB) for $\theta = 0.50$, frequency = 5Hz, window = 2 seconds

	F1-Scores of Our Model							
	Cook	Sweep	Mop	Walk	CU	CD	non-NEAT	Overall Accuracy
All Features	0.89	0.96	0.93	0.91	0.91	1	0.99	94%
Only Accelo	0.89	0.93	0.92	0.89	0.94	1	0.98	92%
Only Gyro	0.58	0.87	0.89	0.71	0.82	0.91	0.99	82%
	F1-Scores of XGB Model							
	Cook	Sweep	Mop	Walk	CU	CD	non-NEAT	Overall Accuracy
All Features	0.92	0.95	0.94	0.93	0.86	0.89	0.98	93%
Only Accelo	0.88	0.88	0.89	0.9	0.77	0.83	0.98	88%
Only Gyro	0.57	0.77	0.79	0.74	0.69	0.64	0.95	74%

(XGB) since it was the best among all flat models. We have demonstrated f1-scores for

Table 4.10: F1-scores of Our Hierarchical Model and Flat Classifier XGBoost (XGB) for $\theta = 0.50$, frequency = 5Hz, window = 4 seconds

	F1-Scores of Our Model							
	Cook	Sweep	Mop	Walk	CU	CD	non-NEAT	Overall Accuracy
All Features	0.96	0.98	0.97	0.95	0.95	0.97	0.99	97%
Only Accelo	0.88	0.83	0.92	0.9	0.96	0.98	0.98	90%
Only Gyro	0.61	0.92	0.89	0.79	0.86	0.96	0.99	85%
	F1-Scores of XGB Model							
	Cook	Sweep	Mop	Walk	CU	CD	non-NEAT	Overall Accuracy
All Features	0.93	0.94	0.95	0.94	0.9	0.91	0.98	93%
Only Accelo	0.87	0.82	0.81	0.91	0.82	0.86	0.97	86%
Only Gyro	0.68	0.83	0.81	0.77	0.77	0.78	0.96	80%

three different types of feature sets in all the tables (Table 4.5, Table 4.6, Table 4.7, Table 4.8, Table 4.9 and Table 4.10). In the first set, we have used all statistical features (i.e., accelerometer x, y, z-axis, the magnitude of accelerometer, gyroscope x, y, z-axis, and the magnitude of a gyroscope). In the second set, we have used only accelerometer features (i.e., individual axis and its magnitude), and in the third set, we have considered only gyroscope features (same as an accelerometer). In all the tables (Table 4.5, Table 4.6, Table 4.7, Table 4.8, Table 4.9 and Table 4.10) our experiments indicate the following: (a) Our proposed hierarchical model outperforms XGBoost (and other flat models) for all the parameter values of θ, window lengths, and data sampling frequency explored in the experiments. (b) Best accuracy (and F1 scores) is obtained when we use both accelerometer and gyroscope features together. (c) As expected, both our model and XGBoost perform better when high-frequency data (5Hz) is used. (d) Both models perform better with $\theta = 0.50$ and a window length 4secs as per showcased in Table 4.6 and Table 4.10

4.6 Conclusion and Limitations

Distinguishing the typical home activities in a home environment from smartwatch sensor data (low-frequency sampling data) is not a trivial problem. Presently, the work closely related to our application deals with high-frequency or multiple sensors. Moreover, they do not include basic home activities like sweeping, mopping, or cooking on a smartwatch. In contrast, our proposed model can distinguish these seven activities from each other using data sampled at low frequency **(1Hz)**. Choosing a 1Hz sampling frequency for smartwatch data in NEAT activity monitoring balances data accuracy and battery efficiency. While there are limitations in capturing high-intensity activities, this is not a significant concern for NEAT monitoring. The extended battery life and reduced device strain make the approach practical and user-friendly, supporting long-term and continuous monitoring.

The hierarchical model and feature extraction methods employed in this research ensure that data collected at this frequency remains accurate and reliable for distinguishing between different NEAT activities. Our experiments show that the proposed approach gives better overall accuracy compared to all flat models especially when both sensors (Gyroscope + Accelerometer) are used together. It was also observed that when both feature extraction methods (ECDF and Statistical) are combined, they produce the best results. For example, for 5Hz frequency and 50% overlapping among windows, the accuracy of our proposed Hierarchical model is 97% whereas for the same configuration the accuracy of the best flat model i.e. XGB is 93%. Similarly for 1Hz frequency and 50% overlapping among windows, the accuracy of our proposed Hierarchical model is 93% whereas for the same configuration, the accuracy of XGB is 85%. This addresses all the research questions outlined in the initial section of the chapter.

However, there are some limitations to our work that should be acknowledged. First, the activities considered in our research were limited. We began our work during the COVID-19 pandemic, which restricted us to a few basic home activities that could be performed during the curfew in India. With leaving home restricted, we confined our activities to basic household tasks. Second, due to the curfew restrictions, the number of volunteers was limited to my family members only. Despite these constraints, our findings demonstrate the potential of using low-frequency smartwatch data for activity recognition, paving the way for future research.

Chapter 5

NEAT Activity Detection using Smartwatch

5.1 Problem Scope

Extending the work done in Chapter 3, this chapter aims to create a robust system for distinguishing Non-Exercise Activity Thermogenesis (NEAT) and non-NEAT activities within a home setting. The goal is to perform a detailed study of various parameters like **features** used for classification, the rate at which file needs to **uploaded to server**, data sampling **frequency**, and choice of **window length** on the battery depletion rate and classification accuracy. The current state of the art in the area of Activity Recognition has not focused on NEAT activities e.g. cooking, sweeping, mopping, etc. Moreover, many works assume high frequency, whereas, we work with data sampled at lower frequencies (10 Hz and 1 Hz).

In our opinion, a perfect solution for NEAT activity recognition must possess the following two characteristics. To begin with, it must use a single and readily available hardware device (for example, a smartwatch or a smartphone). Secondly, as NEAT activities are likely to occur throughout the day, our recognition system should be energy efficient. Any user would probably like the ability to record meaningful data without needing to charge their device numerous times during the day to get significant results. A smartwatch-based solution is developed in this work to recognize NEAT and non-NEAT activities. If energy efficiency is the goal, reducing the sampling rate of the sensors is the most efficient way of achieving it; we know that battery consumption is directly proportional to the data sampling rate [93]. Other than the sampling rate, we can also vary the amount of preprocessing of the features, timestamp length of the window, etc. We will discuss this in detail in the section 5.5. The work outlined here is also referenced in [98].

Although good battery life is guaranteed by a low sampling frequency, it should be noted that the thirteen activities focused on cannot be easily distinguished in low-frequency data. Therefore, the user must make an intelligent choice in determining the frequency that suits them to achieve efficient battery life for their smartwatch. The total accuracy (for our thirteen activities of interest) obtained by different classification algorithms on data sampled at 1Hz, 5Hz, and 10Hz is demonstrated in Table 5.1. As the frequency of sampling data increases, the accuracy is expected to increase."

Frequency	1Hz	5Hz	10Hz
KNN	77 (0.2)	87 (0.5)	88 (0.1)
Multi-layer Perceptron	74 (1.1)	85(0.6)	87 (0.4)
SVM	78 (0.4)	88 (0.3)	88 (0.3)
Logistic	64 (0.7)	77 (0.4)	78 (0.6)
Random Forest	48 (2.7)	55 (1.5)	56 (1.9)
Naive Bayes	62 (0.8)	71 (0.8)	71 (0.8)
XGBoost	80 (0.4)	91 (0.4)	92 (0.3)

Table 5.1: Accuracy of different classifiers (along with the standard deviation after 5 runs) on 1Hz, 5Hz, and 10Hz sampled data for 2-sec window length and 0% overlap.)

5.2 Scope of chapter

This chapter is structured around the following objectives:

1. To evaluate the impact of different feature extraction methods (e.g. ECDF, statistical features) on the classification accuracy and battery consumption of a smartwatch-based NEAT activity recognition system.

2. To assess the effect of varying data sampling frequencies (e.g. 1Hz vs 10Hz) on the classification accuracy and battery life of the smartwatch.

3. To determine the impact of file upload rates on the server on battery consumption and overall performance of the NEAT activity recognition system.

4. To identify the optimal window length that provides the best balance between classification accuracy and battery consumption in a smartwatch-based NEAT activity recognition system.

5. To compare the performance of different classification algorithms (e.g. MLP, SVM, Random Forest, XGBoost) in terms of accuracy and energy efficiency for NEAT activity detection.

5.3 Contributions

1. **Increase number of activities:** In previous studies [99], [100], [101], there has been a limited set of activities in home and outdoor settings, thus the models built do not encompass a broader range of activity detection. This chapter builds on previous chapter 4, where we concentrated limited set of activities. In this chapter, we are extending our analysis to include thirteen activities: (1) Cooking, (2) Sweeping, (3) Mopping, (4) Walking, (5) Climbing Up, (6) Climbing Down, (7) Eating, (8) Driving, (9) Working on a Laptop, (10) Browsing on a Phone, (11) Cycling, (12) Sitting in a Car, and (13) Watching TV.

2. **Deployment of classification models:** Reviewing previous research in the activity detection domain reveals that researchers has focused on developing models, but deployment and real-time feedback to the user were not addressed. In this chapter, we discuss how we deployed our classification models on a server and used them to achieve the desired accuracy. We deployed four robust classification models—XGB, MLP, SVM, and Random Forest—on the Heroku server using the Flask API. This API is accessed via a smartwatch. The prerequisites for using our smartwatch include establishing a Bluetooth connection between the device and a compatible smartphone. Furthermore, to obtain accurate output signals from the server API, both the smartwatch and smartphone must be connected to Wi-Fi. The classification model used by the smartwatch operates on data derived from the accelerometer and gyroscope sensors present in a typical smartwatch.

3. **Worked on others class:** Lastly, we worked on the "Others" class which means if the activity detected does not belong to any of the above-mentioned thirteen classes then the prediction is made as "Others". This works on the basis of threshold which is configurable depending on the strictness a user desires.

4. The brief results can be seen in the Table 5.2.

Parameters	**Best Battery Efficiency**	**Best Accuracy**
Features	Statistical	ECDF
Data sampling Frequency	1 Hz	10 Hz
File Upload Rate	As High as possible	No Effect on Accuracy on changing this parameter
Window Length	$\geq$ 6 seconds	$\geq$ 6 seconds
Classifier	No Effect on Battery	XGB (n_estimators = 100)

Table 5.2: Key Results

5.4 Proposed architecture

After preparing the back-end models, we deploy them on a server. Figure 5.1 gives an overview of how the live data flows in the system and how we get the desired output in a smartwatch by using Wi-Fi connectivity. The process starts with a person wearing a smartwatch with an active Wi-Fi connection; it automatically synchronizes with the smartphone's Wi-Fi. With the press of the start button, we begin capturing the raw inputs from the sensors. We then pre-process this raw data using static or ECDF features and save it into a .csv file in the smartwatch itself. The processed .csv file gets sent to the Heroku server regularly, hitting an API that expects two parameters - the .csv file and the classifier's name. After the model processes the .csv, the server sends back the result to the smartwatch. We repeat this process until the user presses the stop button

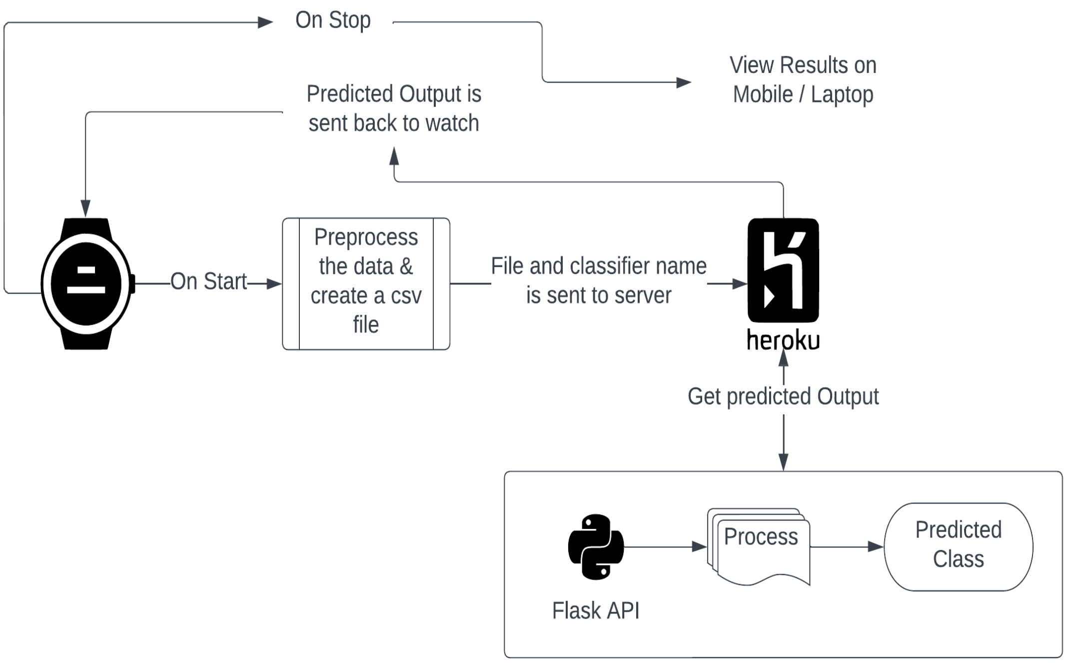


Figure 5.1: Architecture for Testing on Live Data.

and transfers the entire prediction and actual activity class file to a smartphone. We use this method to test the accuracy of all activities (we will show this in the section 5.5.4).

5.4.1 Detection of others classes

Along with the instantaneous detection of the 13 NEAT and non-NEAT activities, a user performs multiple other activities too. So to generalize our approach towards deployment, we detect those non-listed unknown activities named the "Others" class. We use various clustering algorithms and outlier detection methods to detect this activity. We will discuss in brief each of them.

1. **Gaussian Mixture Model:** Its library has a predict_proba function, which gives us a probability of a test point belonging to one of the classes. Even if the points lie far from all the clusters, it gives a high probability of belonging to one of the clusters.

2. **OPTICS Algorithm:** We find different numbers of clusters for our 13 classes and calculate the mean (the average value of all data points belonging to one class) and variance (the spread of all the data points of a class from its class mean) of these clusters. Then, we compared the distance between an outside class data point and the mean of each cluster and compared it with 3 times the variance since for a normal distribution 99% of the data lies within the 3 times the standard deviation.

In order to identify "Others" class we apply threshold concept. We claim that increasing the threshold value will make the classification stricter. Let's say we have 6 seconds of data for prediction and we can get prediction at every second. If we set the threshold at 70%, then at least 4 seconds of instances should belong to one particular class. If that is not the case, we will classify the data as "Others" by default. Hence, the classification as "Others" depends on the strictness of the threshold we set. But, here we have experimented with 60% threshold.

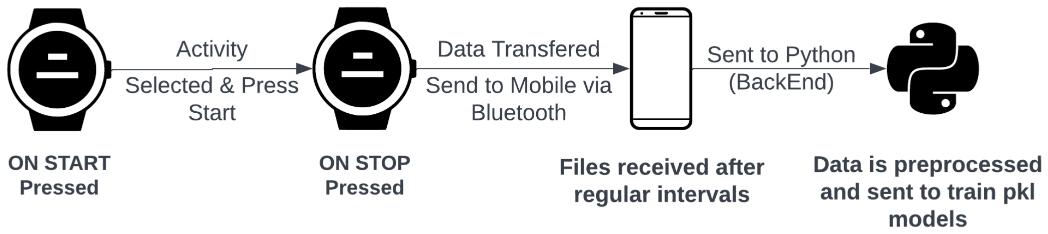


Figure 5.2: Architecture for Data Sampling

5.5 Experimentation

5.5.1 Candidate Algorithms:

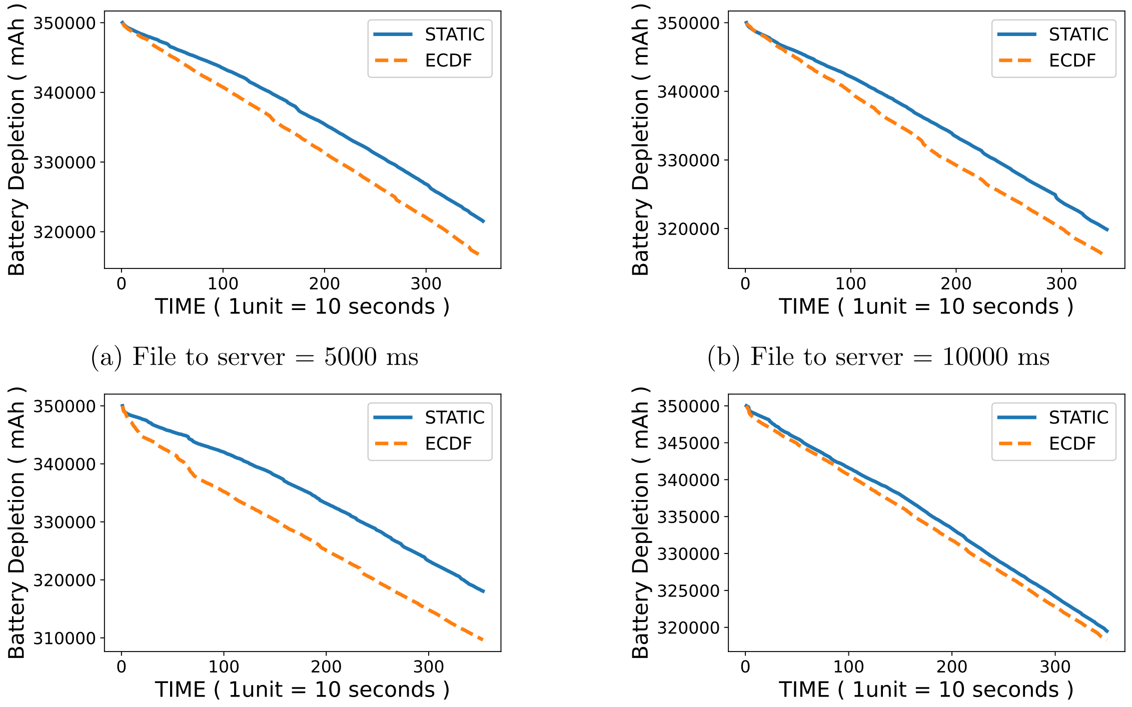
Here, we are using the most common and most effective classification models. These are in-built classifiers found in most machine learning libraries (for example, sklearn in Python). For the classification purpose, we chose the following classifiers - a) MLP (multi-layer perceptron with three hidden layers having 13 neurons each), b) SVM (support vector machine with kernel='rbf'), c) Random Forest, d) XGB (Extreme Gradient Boosting), and e) AutoML (Automated Machine Learning with parameters as time_left_for_this_task: 5*60 (5 minutes) per_run_time_limit: 50 seconds initial_configurations_via_metalearning: 0) [102].

5.5.2 Experimental Goals

Effect of parameters on Battery

We have some parameters on which we performed the battery experiments, i.e., we calculate the battery depletion rate. The four basic parameters are -

1. **Features: -** We factor in two kinds of features, Statical and ECDF. We considered five statistical features, (1) mode, (2) max, (3) median, (4) lower quartile, and (5) standard deviation. Similarly, the Empirical Cumulative Distribution Function is a step function for n data points that jumps up by 1/n each time. For ECDF, we find out 5 data points at equal intervals between the minimum and maximum values. We witness the battery depletion for both ECDF and statistical features in Figure 5.3.



(a) File to server = 5000 ms

(b) File to server = 10000 ms

(c) File to server = 20000 ms

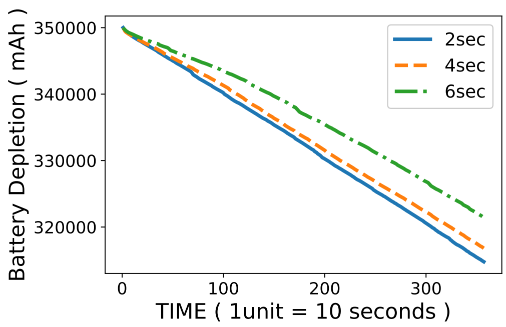
(d) File to server = 40000 ms

Figure 5.3: Rate of Battery Depletion while using ECDF features as compared to the Static features. The X-axis represents the time where 1 unit represents 10 seconds, and the Y-axis represents the Battery left (mAh) in the smartwatch.

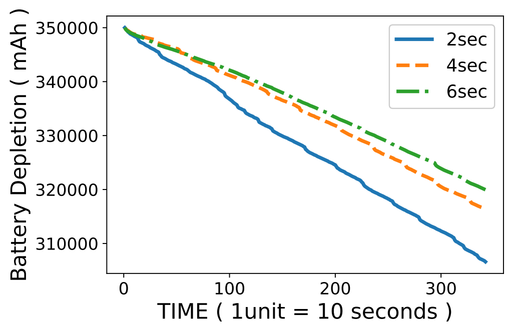
The computation of ECDF is more complicated than simple statistical features that we are calculating for preprocessing and, subsequently, the rate of battery depletion.

2. **Window Length : -** We denote the window's length by ω. In our study, ω equates to 2, 4, and 6 seconds as shown in Figure 5.4. The longer the window length, the lesser the battery consumption. Like "File upload rate," a smaller window length will mean multiple windows; subsequently, it will pass more instances through PKLs on the server.

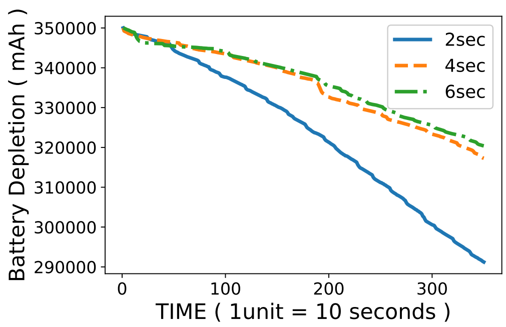
3. **Sensor sampling rate: -** Although there are four types of modes available in our Android, i.e., Normal, UI, Game, and Fastest, we are using the slowest two, i.e., Normal and UI (considering the low frequency $\rightarrow$ low battery depletion phenomena). These are the delays that the sensors provide. We get approximately 3-4 pings/second using normal mode, and UI mode delivers approximately 13-15 pings/second. As we insert the values of the accelerometer and gyroscope, we consider the latest value that the sensor has dispatched after every second and every 1/10th of a second for Normal and UI mode, respectively. Our application's *onSensorChanged()* callback method receives sensor events based on the data delay (or sampling rate). We can witness the battery depletion in Figure 5.5, and, the higher frequency data sampling rate, i.e., 10 Hz (UI mode), depletes the battery at a faster rate as compared to the 1Hz (Normal mode).



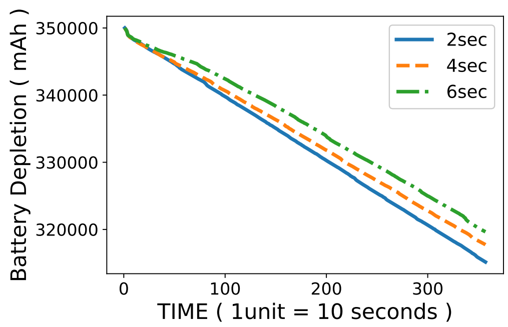
(a) File to server = 5000 ms



(b) File to server = 10000 ms



(c) File to server = 20000 ms

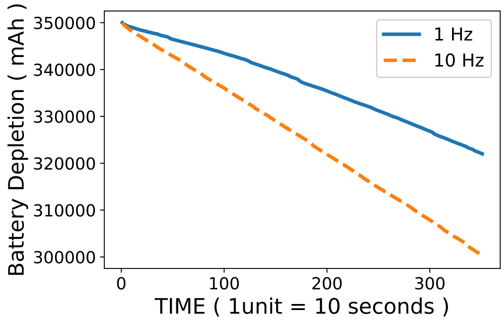


(d) File to server = 40000 ms

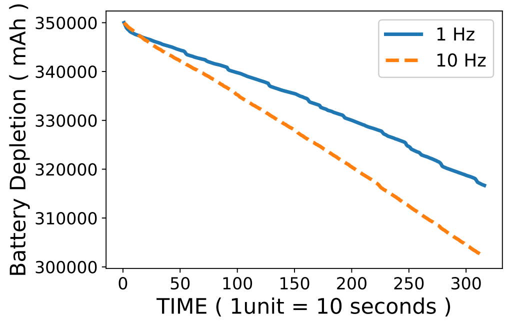
Figure 5.4: Rate of Battery Depletion while taking different window lengths, such as 2 sec, 4 sec, and 6 seconds. The X-axis represents the time where 1 unit represents 10 seconds, and the Y-axis represents the Battery left (mAh) in the smartwatch.

4. **File upload rate: -** This is the time we fix for sending our file to the server for processing through PKLs models (prepared classification models). We observe the battery depletion graph in Figure 5.6. We vary the time intervals from 10 seconds (10k ms) to 70 seconds (70k ms). The smaller the gap for sending files, the higher would be the battery depletion since the server is getting pinged frequently. The time for Sending files to the server can be every few seconds or after full storage capacity it's a personal choice but the optimal one would be "as high as possible" since this is hardware-dependent. We can increase this time duration to some minutes or a few hours till the time our application doesn't freeze.

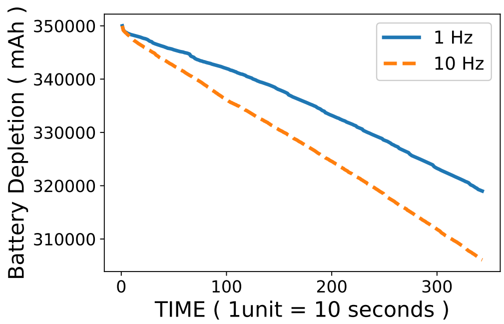
We performed the above four experiments for one hour each and noted the battery level at intervals of 10 seconds. The entire set of experiments was done twice to ensure that the graph behaves in a similar manner, and it did. As we are calculating the battery depletion rate per second, which involves determining the amount of battery left in the smartwatch at a given point in time, we could conclude that the graph behaved in a similar way in both runs. We started with the battery fully charged and noted the depletion rate concerning the last value reported. As the absolute battery value may vary at full charge (100%), we had to normalize the values to start from one absolute value. We did this to get a better visualization of the graphs.



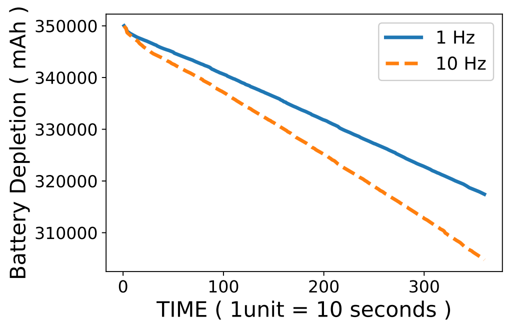
(a) File to server = 5000 ms



(b) File to server = 10000 ms



(c) File to server = 20000 ms



(d) File to server = 40000 ms

Figure 5.5: Rate of Battery Depletion while using 1Hz sampling rate as compared to 10Hz sampling rate. The X-axis represents the time where 1 unit represents 10 seconds, and the Y-axis represents the Battery left (mAh) in the smartwatch.

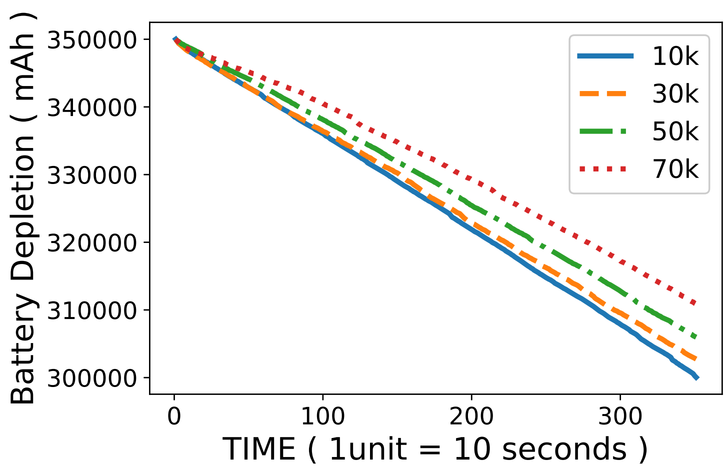
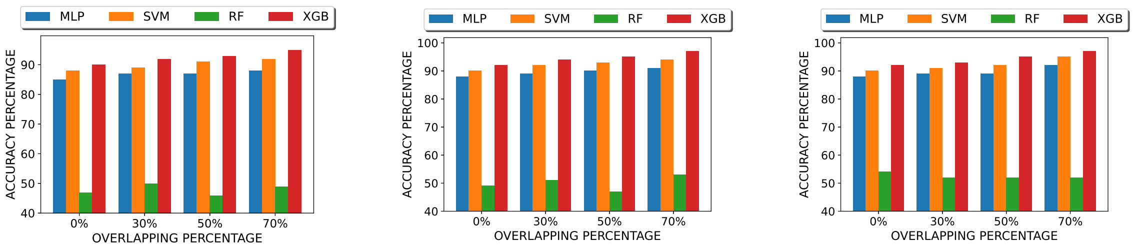


Figure 5.6: Rate of Battery Depletion while varying the file upload rate to the server. The X-axis represents the time where 1 unit represents 10 seconds, and the Y-axis represents the Battery left (mAh) in the smartwatch

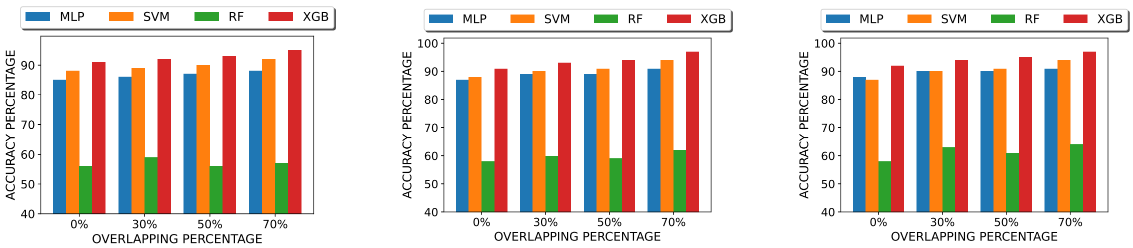
Effect of parameters on Accuracy

Training and Evaluation Metrics We use 2-second, 4-second, and 6-second window lengths in our experiments. We vary a window overlap parameter θ between 0 (i.e., no overlapping between consecutive time windows) and 0.7 (i.e., seventy percent overlap between consecutive time windows). Taking a set of $\mathcal{W}$ values with their corresponding θ values, we divide it into a 4:1 ratio, that is, assigning 80% of the data for training, and the rest 20% is set for testing. Our dataset is randomly divided into training and test sections (10 times) to obtain reliable results. We trained on each of the ten sections (and tested the learned model). Lastly, we present the average of the F1-scores and Accuracy values obtained across these 10 test datasets.



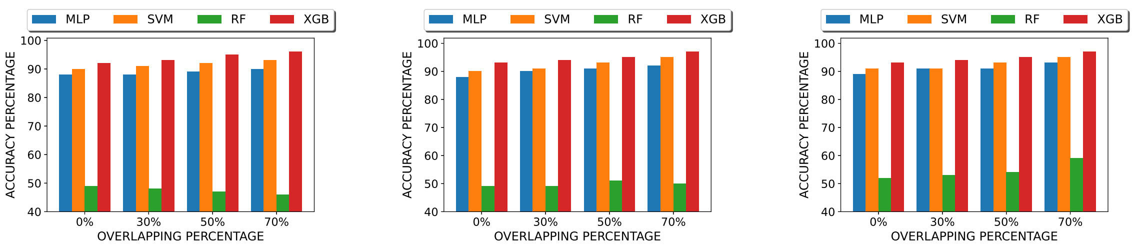
(a) 2 seconds window (b) 4 seconds window (c) 6 seconds window

Figure 5.7: Effect of different overlapping windows percentage on overall accuracy for 5Hz frequency and ECDF Features.



(a) 2 seconds window (b) 4 seconds window (c) 6 seconds window

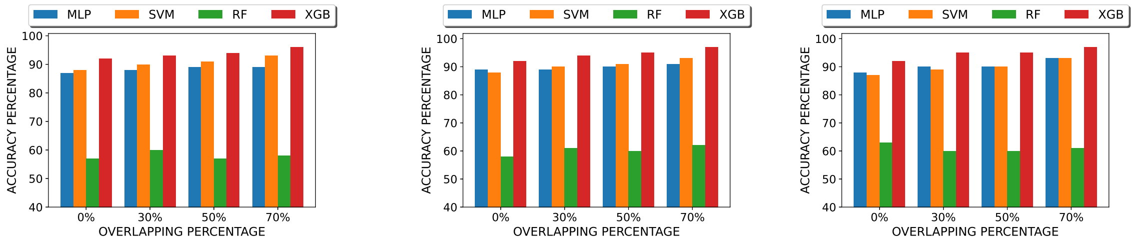
Figure 5.8: Effect of different overlapping windows percentage on overall accuracy for 5Hz frequency and Static Features.



(a) 2 seconds window (b) 4 seconds window (c) 6 seconds window

Figure 5.9: Effect of different overlapping windows percentage on overall accuracy for 10Hz frequency and ECDF Features.

1. **Effect of varying window length and overlapping percentage on the final accuracy** The average of 10 results is shown in the graphs of Figure 5.7, 5.8, 5.9, and 5.10. In all these Figures, a) represents 2 seconds window, b) represents 4 seconds window, and c) represents 6 seconds window. There is a refinement of at least 2 to 3% accuracy for each classifier to increase the window length. The



(a) 2 seconds window (b) 4 seconds window (c) 6 seconds window

Figure 5.10: Effect of different overlapping windows percentage on overall accuracy for 10Hz frequency and Static Features.

higher window length denotes a lesser amount of battery consumption (Figure 5.4) and more amount of data to recognize the activity. Similarly, when we increase the overlapping percentage of windows, there is an improvement of at least 2 to 3% in accuracy for each classifier.

2. **Effect of varying Frequencies and Preprocessing Features on the F1-Scores** As we increase the frequencies of pings from the sensors, not only does the battery consumption increase (Figure 5.5) but the F1-scores boost too. Henceforth, the accuracies also increase. The higher the frequencies of sensors, the better the quality of information to predict. Likewise, we have chosen two types of preprocessing features - ECDF and Statics. ECDF consumes additional battery as compared to Statics (as shown in Figure 5.3) but gives more promising F1-scores and henceforth the accuracies for our thirteen activities.

We see in Table 5.3, for 0% overlapping among windows and a 2 second window length, ECDF features give better F1-scores and accuracies as compared to Statistical features. The same observation can be seen in Table 5.4, i.e., for 50% overlapping among windows and a 2 second window length. In all of the eight tables, we see that the F1-scores and accuracies are better for 10Hz than 5Hz.
Please note - we alter the frequencies of sensor data by adjusting the modes and designating the timer at which we desire our pings. To work on 10Hz data, we use the UI mode, but for 5Hz, we down-sampled the 10Hz data.

5.5.3 Comparison of Feature selection algorithm - TsFresh vs ECDF vs Static Features

We chose three different feature selection algorithms: Tsfresh, ECDF, and statistical methods. We compared the results of all three in terms of accuracy and computation time. Tsfresh [103] provides a comprehensive set of feature extraction methods, ranging from simple statistical moments to more advanced features such as entropy, fractal dimension, and autocorrelation. A large number of features(789 for each of the six axes) provided by tsfresh makes it a useful tool for time series analysis, as it provides a rich set of features that can be used to characterize the structure and behavior of time series data. Figure

Table 5.3: F1-scores and overall accuracy from the Classifiers for ECDF vs Static for overlapping $\theta = 0$ and window length = 2 seconds.

(a) Overlapping $\theta = 0$, frequency = 5Hz, features = ECDF

	Cook	Sweep	Mop	Walk	Climb up	Climb down	Eat	Drive	Laptop	On Phone	Cycling	In Car	Watching TV	Overall Accuracy
MLP	0.75	0.87	0.83	0.87	0.76	0.82	0.85	0.88	0.95	0.98	0.98	0.93	0.93	88%
SVM	0.88	0.95	0.92	0.93	0.90	0.91	0.89	0.91	0.95	0.99	0.99	0.94	0.93	93%
RF	0.56	0.42	0.06	0.45	0.02	0.09	0.41	0.12	0.71	0.94	0.89	0.0	0.0	45%
XGB	0.95	0.97	0.94	0.97	0.96	0.97	0.96	0.96	0.97	1.0	0.99	0.99	0.98	97%
AUTOML	0.88	0.92	0.89	0.93	0.89	0.92	0.92	0.95	0.98	0.99	0.99	0.98	0.97	94%

(b) Overlapping $\theta = 0$, frequency = 5Hz, features = Static

	Cook	Sweep	Mop	Walk	Climb up	Climb down	Eat	Drive	Laptop	On Phone	Cycling	In Car	Watching TV	Overall Accuracy
MLP	0.72	0.84	0.81	0.83	0.71	0.80	0.85	0.88	0.93	0.98	0.99	0.92	0.89	86%
SVM	0.76	0.90	0.86	0.87	0.79	0.86	0.85	0.88	0.94	0.99	0.98	0.89	0.90	88%
RF	0.54	0.56	0.08	0.43	0.05	0.35	0.69	0.80	0.96	0.93	0.70	0.63	57%	
XGB	0.81	0.90	0.88	0.88	0.82	0.86	0.90	0.90	0.94	0.98	0.99	0.94	0.95	90%
AUTOML	0.78	0.90	0.89	0.88	0.75	0.83	0.87	0.88	0.96	0.99	0.98	0.94	0.94	89%

(c) Overlapping $\theta = 0$, frequency = 10Hz, features = ECDF

	Cook	Sweep	Mop	Walk	Climb up	Climb down	Eat	Drive	Laptop	On Phone	Cycling	In Car	Watching TV	Overall Accuracy
MLP	0.79	0.90	0.83	0.87	0.78	0.85	0.85	0.89	0.94	0.98	0.98	0.93	0.92	89%
SVM	0.90	0.96	0.93	0.95	0.93	0.95	0.89	0.92	0.96	0.99	0.99	0.95	0.93	94%
RF	0.56	0.62	0.05	0.5	0.02	0.06	0.39	0.0	0.76	0.94	0.90	0.0	0.0	47%
XGB	0.94	0.98	0.96	0.98	0.98	0.98	0.96	0.96	0.99	1.0	1.0	0.98	0.98	98%
AUTOML	0.92	0.96	0.93	0.95	0.91	0.94	0.94	0.95	0.98	0.99	1.0	0.97	0.96	95%

(d) Overlapping $\theta = 0$, frequency = 10Hz, features = Static

	Cook	Sweep	Mop	Walk	Climb up	Climb down	Eat	Drive	Laptop	On Phone	Cycling	In Car	Watching TV	Overall Accuracy
MLP	0.77	0.88	0.81	0.84	0.75	0.83	0.85	0.87	0.93	0.98	0.98	0.94	0.90	87%
SVM	0.75	0.88	0.85	0.88	0.80	0.87	0.85	0.87	0.94	0.99	0.98	0.90	0.90	88%
RF	0.54	0.68	0.90	0.46	0.02	0.04	0.35	0.48	0.80	0.96	0.91	0.67	0.62	57%
XGB	0.83	0.91	0.88	0.88	0.85	0.88	0.90	0.91	0.95	0.98	0.99	0.95	0.93	91%
AUTOML	0.84	0.91	0.87	0.89	0.82	0.90	0.87	0.89	0.94	0.99	0.99	0.95	0.93	91%

5.11 (a) and (b) shows the accuracy comparison for 5Hz and 10Hz sampling data and we can observe that for 4 second and 6 second windows, the accuracy of Tsfresh is better as compared to static and ECDF. But in Figure 5.11 (c) and (d) we can observe the computation time of Tsfresh as compared to static and ECDF is much much higher for both 5Hz and 10Hz sampling data. There are several other feature extraction tools that are similar to tsfresh such as - Featuretools, tsfeat, rfeat, sktime, and tsaug which we will be exploring in depth in our future work.

5.5.4 Real Time Testing

To confirm the accuracy of our model, we conducted a brief real-time user test by selecting a new volunteer to perform all activities in this live testing phase. We chose to use only one volunteer for this quick assessment of our model. In the real-time testing, we used the XGB classifier, as it delivered the best accuracy during training and testing. We chose a data sampling rate of 10Hz, which was an optimal choice among the two modes (normal and UI). The file upload rate was set to once every 70,000 ms, as this proved to be more battery efficient, and a window length of 6 seconds was chosen for the best balance of accuracy and battery efficiency. The volunteer performed each activity for varied amounts

Table 5.4: F1-scores and overall accuracy from the Classifiers for ECDF vs Static for overlapping $\theta = 0.5$ and window length = 2 seconds

(a) Overlapping $\theta = 0.5$, frequency = 5Hz, features = ECDF

	Cook	Sweep	Mop	Walk	Climb up	Climb down	Eat	Drive	Laptop	On Phone	Cycling	In Car	Watching TV	Overall Accuracy
MLP	0.78	0.88	0.83	0.87	0.77	0.83	0.84	0.90	0.94	0.97	0.98	0.92	0.89	87%
SVM	0.90	0.96	0.95	0.94	0.89	0.92	0.92	0.94	0.96	0.99	0.99	0.95	0.94	94%
RF	0.56	0.41	0.06	0.44	0.03	0.01	0.41	0.16	0.75	0.95	0.89	0.00	0.00	45%
XGB	0.96	0.98	0.98	0.97	0.95	0.96	0.98	0.97	0.99	1.00	1.00	0.99	0.99	98%
AUTOML	0.86	0.91	0.89	0.86	0.80	0.87	0.84	0.88	0.94	0.98	0.98	0.91	0.91	89%

(b) Overlapping $\theta = 0.5$, frequency = 5Hz, features = Static

	Cook	Sweep	Mop	Walk	Climb up	Climb down	Eat	Drive	Laptop	On Phone	Cycling	In Car	Watching TV	Overall Accuracy
MLP	0.77	0.87	0.83	0.84	0.74	0.82	0.85	0.89	0.92	0.98	0.97	0.93	0.91	87%
SVM	0.83	0.92	0.89	0.88	0.80	0.87	0.89	0.90	0.94	0.99	0.98	0.93	0.91	90%
RF	0.56	0.62	0.08	0.46	0.04	0.00	0.36	0.27	0.82	0.96	0.89	0.60	0.65	55%
XGB	0.87	0.91	0.92	0.91	0.84	0.89	0.93	0.94	0.96	0.99	0.99	0.97	0.95	93%
AUTOML	0.85	0.91	0.91	0.87	0.81	0.88	0.88	0.90	0.96	0.99	0.99	0.95	0.95	91%

(c) Overlapping $\theta = 0.5$, frequency = 10Hz, features = ECDF

	Cook	Sweep	Mop	Walk	Climb up	Climb down	Eat	Drive	Laptop	On Phone	Cycling In Car	Watching TV	Overall Accuracy
MLP	0.82	0.91	0.86	0.88	0.82	0.88	0.85	0.93	0.97	0.99	0.93	0.92	90%
SVM	0.93	0.97	0.96	0.95	0.93	0.95	0.96	0.99	0.99	0.99	0.96	0.94	95%
RF	0.55	0.61	0.03	0.52	0.02	0.01	0.04	0.72	0.93	0.92	0.50	0.47	47%
XGB	0.98	0.98	0.98	0.97	0.98	0.98	0.99	1.0	0.99	1.0	0.99	0.99	98%
AUTOML	0.87	0.90	0.89	0.89	0.83	0.86	0.89	0.95	0.98	0.98	0.93	0.91	91%

(d) Overlapping $\theta = 0.5$, frequency = 10Hz, features = Static

	Cook	Sweep	Mop	Walk	Climb up	Climb down	Eat	Drive	Laptop	On Phone	Cycling In Car	Watching TV	Overall Accuracy
MLP	0.80	0.90	0.86	0.87	0.78	0.85	0.88	0.91	0.95	0.98	0.99	0.93	89%
SVM	0.84	0.93	0.90	0.91	0.83	0.88	0.91	0.99	0.98	0.99	0.93	0.90	91%
RF	0.56	0.70	0.48	0.48	0.03	0.15	0.82	0.96	0.92	0.61	0.0	0.55	55%
XGB	0.89	0.94	0.91	0.92	0.88	0.93	0.95	0.96	0.99	0.99	0.97	0.95	94%
AUTOML	0.85	0.93	0.91	0.92	0.87	0.91	0.88	0.99	0.98	0.99	0.96	0.95	92%

5 Hz	2 seconds	4 seconds	6 seconds
Tsfresh	90%	93%	92%
Static	91%	91%	92%
ECDF	97%	92%	92%

(a) Accuracies for 5Hz sampling data

10 Hz	2 seconds	4 seconds	6 seconds
Tsfresh	93%	94%	94%
Static	92%	92%	92%
ECDF	97%	93%	93%

(b) Accuracies for 10Hz sampling data

5 Hz	2 seconds	4 seconds	6 seconds
Tsfresh	3931.671	2283.186	1636.572
Static	18.701	8.942	6.355
ECDF	22.159	11.342	8.442

(c) Computation Time for 5Hz sampling data

10 Hz	2 seconds	4 seconds	6 seconds
Tsfresh	4324.854	2593.414	1933.929
Static	19.299	9.414	6.864
ECDF	23.561	11.907	8.327

(d) Computation Time for 10Hz sampling data

Figure 5.11: Comparison of Tsfresh vs Static vs ECDF Features selection algorithms w.r.t. Accuracies and Computation time

of time; therefore, we considered the ratio of the time duration of correctly classified instances to the total time duration of the activity. The results can be seen in Figure 5.12.

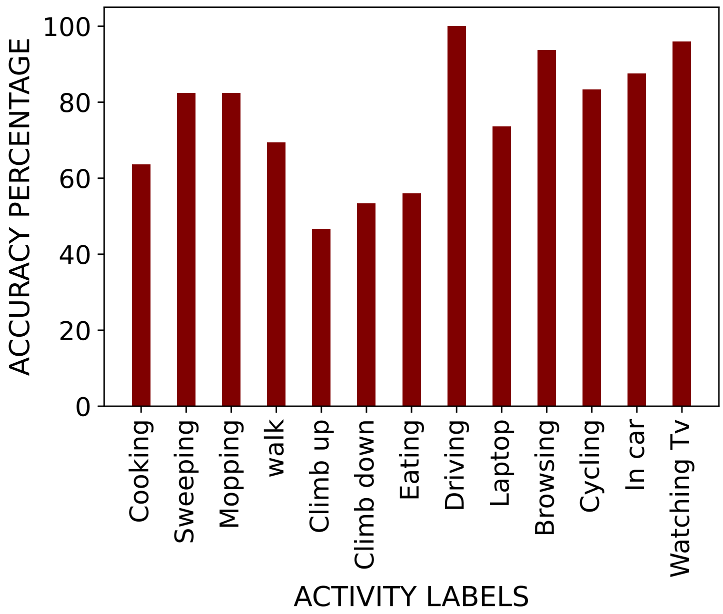


Figure 5.12: Live Data Accuracy for all activities

5.5.5 Key Results

We deployed our models at lower frequencies, specifically 10 Hz and 1 Hz. The four parameters used to assess the battery depletion rate—Features, File upload rate, Data sampling rate, and Window lengths—varied as shown in Table 5.5, which presents our key results.

We evaluated two distinct types of features: **Statistical** and **ECDF (Empirical Cumulative Distribution Function)**. Each of these feature sets has its own impact on both battery consumption and accuracy. **Statistical features**, which are generally less computationally intensive, exhibit lower energy consumption, depleting approximately 38,000 milliampere-hours of battery within one hour. This efficiency makes them an attractive option for scenarios where conserving battery life is a priority. On the other hand, **ECDF features** are more computationally demanding due to their complex nature. This complexity leads to higher battery consumption, with a recorded depletion of around 43,000 milliampere-hours in the same timeframe. However, this increased energy use comes with a significant trade-off: the accuracy of models using ECDF features is notably superior. Specifically, **ECDF** achieves an accuracy of 97%, whereas **Statistical features** reach only 91%. To summarize, while **Statistical features** are more energy-efficient and ideal for applications where battery preservation is critical, **ECDF features** are the preferred choice when the highest accuracy is required, despite their higher battery consumption. **Thus, Statistical features are recommended for energy-sensitive applications, while ECDF features are favored in accuracy-critical scenarios**.

File upload rate is the interval after which data files are sent from the device to the server for result prediction. This parameter has a direct impact on battery consumption, as more frequent uploads lead to higher energy use. From our experiments, we observed

Table 5.5: Key Results

Parameters	Summary on Battery Consumption	Summary on Accuracy	Constants
Features	**Statistical** (Battery Depletion - "38,000" milliamperes in one hour) **ECDF** (Battery Depletion - "43,000" milliamperes in one hour)	**Statistical** (91%) **ECDF** (97%)	Fixed Data sampling frequency as 10 Hz Window length as 6 seconds File Upload Rate as 50k Classifier as XGB
Data sampling Frequency	**1 Hz** (Battery Depletion - "36,000" milliamperes in one hour) **10 Hz** (Battery Depletion - "43,000" milliamperes in one hour)	**1 Hz** (92%) **10 Hz** (97%)	Fixed Features as ECDF Window length as 6 seconds File Upload Rate as 50k Classifier as XGB
File Upload Rate	**Once every 10 seconds** - 50,000 milliamperes in one hour **Once every 30 seconds** - 47,000 milliamperes in one hour **Once every 50 seconds** - 43,000 milliamperes in one hour **Once every 70 seconds** - 39,000 milliamperes in one hour	No Effect on Accuracy on changing this parameter	Fixed Features as ECDF Data sampling frequency as 10 Hz Window length as 6 seconds Classifier as XGB
Window Length	**2 sec** - 43,000 milliamperes in one hour **4 sec** - 33,600 milliamperes in one hour **6 sec** - 30,000 milliamperes in one hour	**2 sec** - 91% **4 sec** - 93% **6 sec** - 97%	Fixed Features as ECDF Data sampling frequency as 10 Hz File Upload Rate as 50k Classifier as XGB
Classifier	No Effect on Accuracy on changing this parameter	**MLP** - 91% **SVM** - 93% **RF** - 56% **XGB** - 97%	Fixed Features as ECDF Data sampling frequency as 10 Hz Window length as 6 seconds File Upload Rate as 50k
	Best Battery efficiency is Approx "33,000" milliamperes in one hour for an accuracy of 87%. Parameters are - Statistical features, Normal Mode, File Upload Rate once every 70 seconds, and window length as 6 seconds	Best Accuracy of 97% depletes Approx "37,000" milliamperes in one hour. Parameters are - ECDF features, UI Mode, File Upload Rate once every 70 seconds and window length as 6 seconds	

that reducing the frequency of file uploads (i.e., increasing the upload interval) significantly reduces battery depletion. Specifically: Uploading once every 10 seconds results in the highest battery depletion, approximately 50,000 milliamperes in one hour. Uploading once every 30 seconds reduces battery depletion to about 47,000 milliamperes in one hour. Uploading once every 50 seconds further lowers battery consumption to 43,000 milliamperes in one hour. Uploading once every 70 seconds results in the lowest observed battery depletion, around 39,000 milliamperes in one hour.

Interestingly, the file upload rate does not affect the accuracy of predictions, making it an independent parameter in terms of accuracy. Therefore, the file upload rate should be as high as possible to maximize energy efficiency. However, practical constraints with devices like smartwatches showed that longer intervals (such as an hour) could cause the device to become unresponsive. Thus, we limited our experiments to a maximum of 70 seconds, where we found an optimal balance between battery efficiency and device performance. **Hence, we report the File upload rate to be "as high as possible" to enhance energy efficiency while maintaining device stability.**

Data Sampling Rate - Android provides developers with four distinct data sampling frequencies, each designed to cater to different performance needs: **Normal**, **UI**, **Game**, and **Fastest**. These modes vary significantly in terms of how they balance battery consumption against the responsiveness of data collection. The **Normal** mode operates at a 1 Hz frequency, making it the most energy-efficient option available. This mode is suitable for applications where battery life is a top priority, as it only depletes around 36,000 milliampere-hours in one hour. On the other hand, the **UI** mode, which runs at a 10 Hz frequency, provides a faster data collection rate. This higher frequency allows for more data points to be captured within the same time frame, significantly improving the accuracy of decisions made by the application. However, this comes at the cost of higher battery consumption, depleting around 43,000 milliampere-hours in one hour. The **Game** and **Fastest** modes, while not considered in our battery efficiency goal, follow a similar pattern: the faster the data collection, the greater the battery drain. The order of battery depletion follows this sequence: **Fastest >Game >UI >Normal**. Given that our focus is on optimizing battery efficiency without sacrificing too much accuracy, we primarily consider **UI mode** (10 Hz sampling rate) and **Normal mode** (1 Hz sampling rate). The higher the frequency, the more information is available to make precise and definitive decisions, which enhances the overall accuracy of the system. **Thus, we report that 1 Hz (Normal mode) is energy-efficient, while 10 Hz (UI mode) is accuracy-driven**.

Lastly, we consider **Window Length** which refers to the time duration over which the classifier collects data before making a prediction about the activity. This parameter is crucial as it directly influences both the accuracy of the predictions and the battery consumption of the device. When the window length is increased, the classifier has more data to analyze, which generally leads to better accuracy. This is because the longer the time period, the more information is available for the classifier to make a well-informed prediction. However, this improvement in accuracy is only valid up to a certain point. Beyond this limit, increasing the window length further does not yield any significant gains in accuracy. Our experiments showed that extending the window length beyond 6 seconds did not result in any additional accuracy improvements. We considered three different window lengths in our study: 2, 4, and 6 seconds. Here's what we found: 2-second window length: This resulted in the highest battery consumption, approximately 43,000 milliamperes in one hour. The frequent predictions required more energy, but the accuracy was relatively lower at 91%. 4-second window length: Reducing the prediction frequency by increasing the window length to 4 seconds decreased battery consumption to around 33,600 milliamperes in one hour, with an improved accuracy of 93%. 6-second window length: Further increasing the window length to 6 seconds resulted in the lowest battery consumption, approximately 30,000 milliamperes in one hour, with the highest accuracy of 97%. From these results, it's clear that a shorter window length leads to more frequent predictions, which increases battery consumption. On the other hand, a 6-second window length strikes the best balance between energy efficiency and prediction accuracy. **Hence, we recommend a window length of 6 seconds as it is both energy-efficient and**

accuracy-driven. Beyond 6 seconds, additional accuracy gains are minimal, making this duration optimal for applications that require a balance between energy consumption and performance.

So the best set of parameters for good battery efficiency would be → Statistical features, File upload rate is chosen as once every 70,000ms, 1hz data sampling rate, 6 seconds or above window length and lastly XGB as classifier (no effect though since classifier is present on the server). The best parameters for good accuracy would be → ECDF features, File upload rate as once every 70,000ms, 10hz data sampling rate, 6 seconds or above window length, and lastly XGB as classifier (as shown in Figure 5.5).

5.6 Conclusion and Limitations

We developed a system that tells the user what kind of activity they are doing by taking data from the smartwatch that the user has worn on his wrist. We have a total of 13 activities - (1) Cooking, (2) Sweeping, (3) Moping, (4) Walking, (5) Climbing up, (6) Climbing down, (7) Eating, (8) Driving, (9) Working on Laptop, (10) Browsing on the phone, (11) Cycling, (12) Sitting in a car, and (13) Watching TV. The process to get this output is collecting the raw input data from the smartwatch sensors by pressing a start button and then dispatching the data to the server to get interpreted and passed through classification models deployed there. The communication between the smartwatch and the server happens through Wi-Fi connectivity. We evaluate the final results in terms of battery depletion rate for each parameter - preprocessing features (Static and ECDF), File upload rate, Data sampling rate as 1hz or 10hz, and timestamp length of the window. The battery depletion rate would enable a user to make the best choice for determining an optimal set of parameters. Since these classification models are not deployable on a smartwatch, we had to take support from the intermediate server, but in our upcoming research, we tried to deploy everything on the smartwatch to mitigate the quotient of the internet availability.

We implement that if the user does not perform an activity from the above chosen 13 activities, then our model will point to another class termed the Fourteenth class, i.e., “others“ We term this when there is no majority voting to any of the above-stated classes. We used different clustering techniques to learn about the “others“ class, but it was impossible to separate the outliers from the actual plethora of activities in high-dimensional data. Henceforth, we applied the concept of dominating class (60% and above instances, if they belong to the same class for a particular window, then we say yes to that class; otherwise, the label is “others“). We can adjust this threshold to 40% or 80%, depending on how lenient or strict we want our system to be.

Even after creating a deployable model for real-time activity recognition using smartwatch data, there are some limitations to our work. Firstly, despite expanding the range of activities, we faced challenges in increasing the number of volunteers due to the intermittent nature of COVID restrictions. Consequently, we could only enlist a total of 10 volunteers from our neighborhood. Despite this limitation, we endeavored to diversify our volunteer pool by including individuals aged 15 to 65. Secondly, there was a constant WiFi and Bluetooth connectivity usage, as the ML models were not deployable on the smartwatch. We opted to deploy the classification models on the server, which required a constant connection to the server via WiFi to receive live output predictions.

Chapter 6

Neural Networks for NEAT Activity Detection on Smartwatch

6.1 Problem Scope

This chapter focuses on developing neural network models for Non-Exercise Activity Thermogenesis (NEAT) recognition, extending beyond typical activities. Our aim is to create a smartwatch-compatible model, offering real-time updates on NEAT activities. We scrutinized key parameters, notably window length, considering their impact on battery consumption and accuracy. Our ideal solution integrates seamlessly with common devices and prioritizes energy efficiency for continuous data collection. Utilizing neural networks on smartwatches, we prioritize user privacy with on-device predictions. Departing from conventional practices, our lower frequency data sampling rate of 10 Hz, notably with bidirectional LSTM, yields optimal accuracy and minimal battery depletion. Our hybrid model, combining 1D-CNN and bidirectional LSTM, excels in accuracy, suggesting a 10-second window for practical real-world application. The findings discussed here are further elaborated in [104].

6.2 Scope of chapter

1. **Exploring neural networks:** To research on neural networks. This involved exploring various neural network architectures, evaluating their accuracy, and feasibility of deployment in smartwatch.

2. **Server-less tracking of NEAT activities:** To deploy classification models directly in smartwatch for a server-less tracking of NEAT activities building upon the foundation laid in previous chapter 5.

3. **Optimising energy efficiency:** To identify the optimal window length, which achieves a balance between maintaining high accuracy and minimizing battery consumption, making the solution practical for continuous use in real-world settings.

6.3 Contributions

This chapter makes several important contributions to the field of NEAT activity detection on smartwatches through the development and experimentation with neural networks.

These contributions include:

1. **Server-less Real-Time Predictions:** By focusing on on-device predictions without reliance on external servers, we prioritized user privacy and developed a solution compatible with smartwatch hardware, paving the way for scalable deployment.
2. **Performance Superiority of Hybrid Model:** Through comparative analysis, the hybrid model demonstrated superior performance, achieving the highest accuracy with lower battery consumption compared to other models like 1D-CNN and Bi-LSTM, particularly at higher data sampling rates of 10 Hz.

These contributions not only enhance NEAT activity recognition accuracy but also advance the development of energy-efficient, privacy-preserving solutions for wearable devices.

6.4 Candidate neural network classifiers

6.4.1 1D-CNN

The model architecture presented in Figure 6.1 includes eight layers. The first layer is a Conv1D layer with 64 filters, a kernel size of 3, and a ReLU activation function. The layer takes input in the shape of (# window length, # no. of features), where the window length refers to the number of sensor instances considered for prediction and training, and the number of features is six, as we have three axes for the accelerometer and three axes for the gyroscope.

The second layer is another Conv1D layer with 128 filters, a kernel size of 3, and a ReLU activation function. A Dropout layer with a rate of 0.5 is the third layer, which randomly drops out some of the neurons during training, to prevent overfitting.

The fourth layer is a MaxPooling1D layer with a pool size of 2. It reduces the spatial size of the output from the previous layer by taking the maximum value in each of the 2 neighboring positions, which helps prevent overfitting and reduce the number of parameters.

The fifth layer is a Flatten layer that flattens the output from the previous layer into a 1D array. The sixth layer is a Dense layer with 100 neurons and the ReLU activation function.

The seventh layer is another Dense layer with n_outputs neurons, where n_outputs is the number of classes in the output. It uses the softmax activation function, which outputs probabilities for each class and is commonly used as the last layer in a classification problem.

Finally, the model is compiled with a categorical cross-entropy loss function, an optimizer, and accuracy as the evaluation metric. The categorical cross-entropy is used as a loss function for multi-class classification problems, while the optimizer updates the model parameters during training.

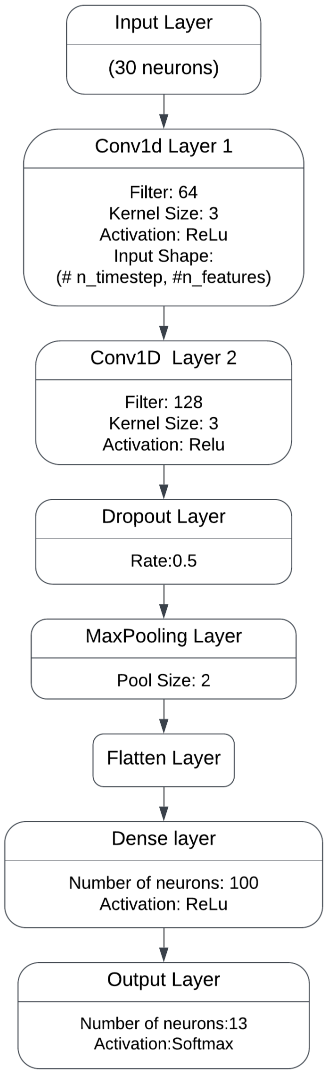


Figure 6.1: 1D-CNN

6.4.2 ANN

To create an Artificial Neural Network (ANN) using Keras' Sequential model, we begin by adding the input layer which consists of 30 neurons and uses ReLU activation shown in Figure 6.2. Next, we add four hidden layers with 40 neurons each using ReLU activation, along with a Dropout layer with a rate of 0.2 for each hidden layer. Dropout randomly removes neurons during training to prevent overfitting. Finally, we add an output layer with 13 neurons and a softmax activation function. The softmax function is commonly used for multi-class classification problems as it outputs a probability distribution over the possible classes. In summary, our classifier defines an ANN with four hidden layers, each with 40 neurons, a dropout rate of 0.2, and an output layer with 13 neurons using the softmax activation function.

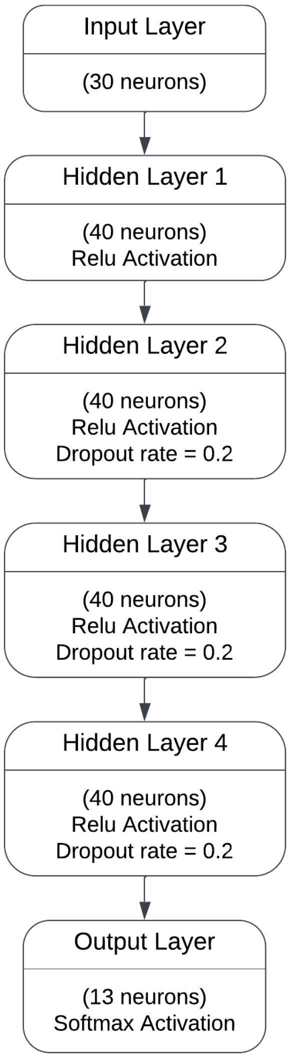


Figure 6.2: ANN

6.4.3 LSTM

To start, we initialize the RNN model depicted in Figure 6.3. Following this, we add the first layer of LSTM with 120 units, where the return_sequences parameter is set to True to return sequences from this layer to the next. Additionally, the input_shape parameter is set to the shape of the input data X_train, which is a 3D array with shape (number of samples, number of time steps, number of features). We then apply Dropout regularization to the first LSTM layer with a parameter of 0.05, representing the fraction of input units to be dropped.

Next, we add the second layer of LSTM with 150 units, again with return_sequences set to True to return sequences to the next layer, followed by Dropout regularization. We then add the third layer of LSTM with 100 units, again with return_sequences set to True, followed by Dropout regularization. Finally, we add the fourth layer of LSTM with 50

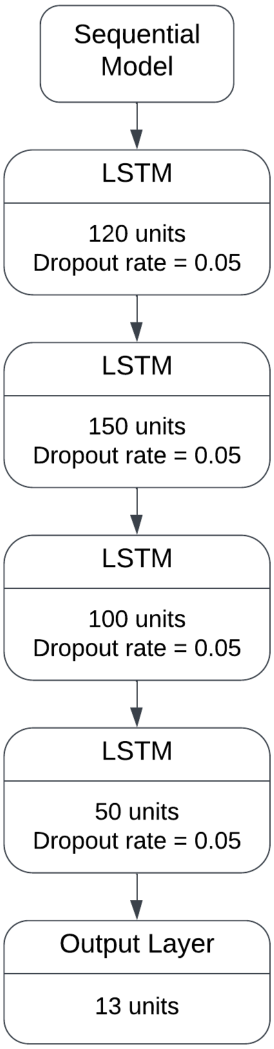


Figure 6.3: LSTM

units, with return_sequences not set, as this is the last LSTM layer in the network, followed by Dropout regularization.
Lastly, we add the output layer with 13 units, using the Dense function to create a fully connected layer in which each input unit is connected to each output unit. The output of this layer represents the predicted value.

6.4.4 BiDirectional-LSTM

Our model architecture shown in Figure 6.4 includes a Bidirectional LSTM layer with 128 units, a type of RNN that captures temporal dependencies in sequential data. The layer's input shape is based on the training data's shape, i.e., # window length and # no. of features, where the window length refers to the number of sensor instances considered for prediction and training, and the number of features is 6 as discussed above.
To prevent overfitting, we add a dropout layer with a rate of 0.5, randomly dropping out

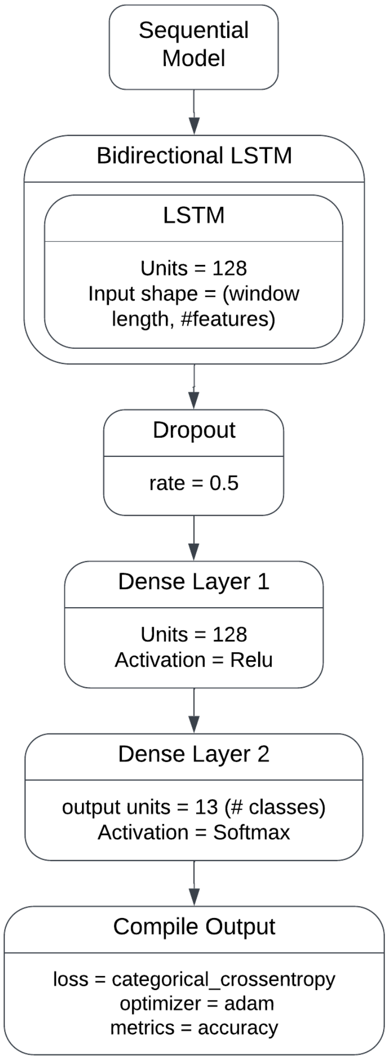


Figure 6.4: Bidirectional LSTM

some of the neurons during training.
We add a dense layer with 128 units and ReLU activation, followed by an output layer with a number of units equal to the number of predicted classes. The output layer uses the softmax activation function, which generates a probability distribution over the possible classes.
The model is compiled using the categorical cross-entropy loss function, which is commonly used in multiclass classification problems, and the Adam optimizer.

6.4.5 Our Hybrid Model

The presented diagram in Figure 6.5 delineates the architecture of a neural network designed to process sequential data. The model is structured sequentially in this arrangement, with each component serving a distinct purpose.
The initial component is a Conv1D layer aimed at detecting patterns in the input data.

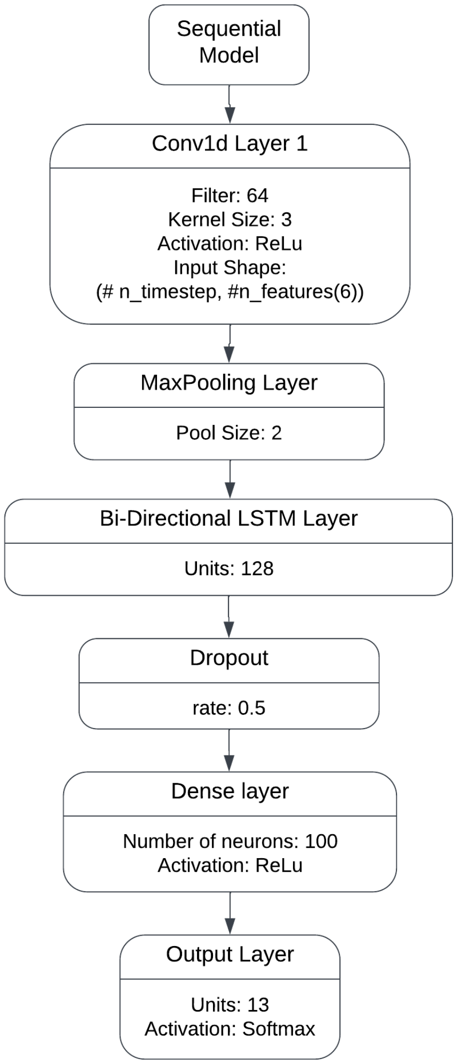


Figure 6.5: Our Proposed Hybrid Model

It comprises 64 filters, each scanning through a window of size 3. The 'ReLU' activation function is employed to introduce non-linearity. This layer is designed to process data shaped according to a specified window size and possesses six features.

Subsequently, a MaxPooling1D layer follows, tasked with reducing data dimensions through a process known as pooling. In this case, the pooling operation selects the maximum value from every two adjacent values, thereby downsizing the data representation.

The architecture progresses to incorporate a Bidirectional Long Short-Term Memory (LSTM) layer, renowned for its prowess in capturing sequential dependencies. This specific LSTM configuration comprises 128 units and operates in both forward and backward directions, allowing the network to glean insights from patterns in either direction.

A Dropout layer is introduced to tackle the risk of overfitting – a phenomenon where a model becomes overly adapted to training data. This layer systematically deactivates

a fraction of neurons during training, effectively preventing the model from relying too heavily on particular neurons.
A Dense layer with 128 units is added, enhancing the model's capacity to identify intricate patterns. The 'ReLU' activation function applied to this layer introduces non-linearity, enabling the model to accommodate complex relationships.
Concluding the architecture is an output Dense layer meticulously tailored for classification tasks. The number of units within this layer aligns with the total count of classes, denoted as 'num_classes.' The 'softmax' activation function is utilized here, generating a probability distribution that indicates the likelihood of each class being the correct prediction.
In summary, this neural network configuration comprises a sequence of layers, each contributing to the model's ability to process sequential data and produce accurate class predictions.

6.5 Experimentation

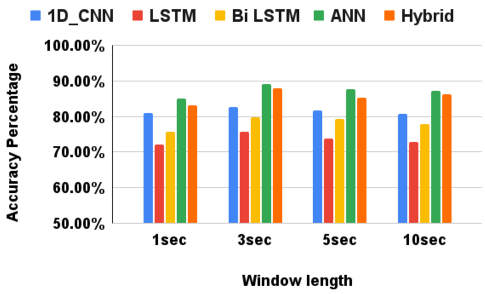
Table 6.1: Total Training Data

Window Length	**10 Hz**
1 second	113,400
3 seconds	37,800
5 seconds	22,680
10 seconds	11,340

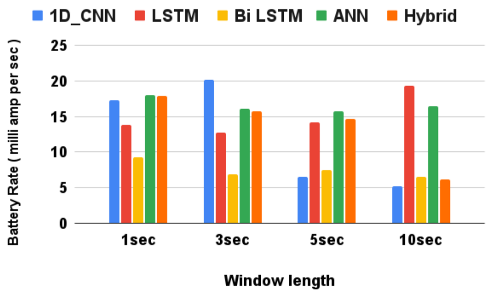
Table 6.1 presents the distribution of occurrences for different window lengths. With an increase in the window length, the frequency of occurrences declines. This trend arises because we are expanding the timeframe for analyzing input statistical features and their respective output labels. We deliberately decided not to extend the window length further, as it did not yield substantial enhancements.
Training and evaluation metrics: In our experimental setup, we utilized window lengths of 1 second, 3 seconds, 5 seconds, and 10 seconds. To ensure the robustness of outcomes, we partitioned the dataset into training and testing subsets with an 80:20 ratio, allocating 80% for training and 20% for testing. Among this 80% training data we had a validation split of 10% and 90% i.e. our training data is divided into further training and validation sets. A validation split of 0.1 (or 10%) is a commonly used practice in machine learning and neural networks for the purpose of model evaluation. The primary goal of using a validation split is to assess how well our neural network is generalizing to data that it has not seen during training. Here's why it's done and what it accomplishes:
Model Evaluation: During the training process, the model learns from the training data. However, the model's ultimate goal is to perform well on unseen or new data. The validation split allows us to set aside a portion of our data (in this case, 10%) for the

purpose of model evaluation.
Preventing Overfitting: When a model is trained for too many epochs or is too complex, it may start to "overfit" the training data. Overfitting occurs when the model performs very well on the training data but poorly on new, unseen data. The validation data helps to detect overfitting. If the validation loss starts to increase while the training loss continues to decrease, it's a sign that our model might be overfitting.
Early Stopping: Validation data is crucial for implementing early stopping, which is a technique used to prevent training the model for too long. When the validation loss starts to increase consistently, early stopping allows us to stop training, saving time and potentially preventing overfitting.

6.5.1 Effect of varying window length on total Accuracy & battery consumption



(a) Accuracy



(b) Rate of Power Consumption

Figure 6.6: Accuracy and Battery Depletion Rate of neural networks

We compared our approach against [104] 1D_CNN, LSTM, Bidirectional LSTM, and ANN models for NEAT and non-NEAT activity detection. Figure 6.6a illustrates the performance of our Hybrid model in comparison to others. In this experiment, we took the following window lengths: 1 second, 3 seconds, 5 seconds, and 10 seconds. Following two conclusions were drawn from our experiment. First, our Hybrid Model and ANN obtained higher accuracy than all the other models (1D_CNN, Bi-LSTM, and LSTM), and this trend was consistent across all the window lengths. Secondly, the performance of all the models remained relatively consistent across the window lengths, i.e., no major trends were seen across window lengths.
Figure 6.6b presents the rate of battery depletion, measured per second, for our five designated neural networks: 1D_CNN, LSTM, Bidirectional LSTM, ANN, and our proposed Hybrid Model. We used the function *BatteryManager.BATTERY_PROPERTY_CHARGE_COUNTER* in Android to extract the cumulative charge stored in the smartwatch battery. Figure 6.6b shows that a window length of 10 secs leads to the lowest battery consumption (around 6 milli-amperes per second) for 1D_CNN, Bi-LSTM, and our Hybrid Model. Further, in our experiments, we observed that when we set the window length to 10 seconds (in our Hybrid Model) the app consumed around 10% of

battery in one hour. However, when we set it to 3 seconds or 5 seconds, the app consumed around 25% of battery in just one hour.
To summarize, we make the following conclusions from our experiments: (1) Accuracy of all models remains consistent across different window lengths. (2) Our proposed Hybrid Model categorically outperforms 1D_CNN, Bi-LSTM, and LSTM in terms of accuracy (across all window lengths). (3) While the accuracy of our Hybrid Model is close to that of ANN (across all window lengths), our model obtained a significantly lower battery consumption rate than ANN (while maintaining comparable accuracy) at a window length of 10 seconds. In other words, a window length of 10 seconds (for our model) balances the trade-off between accuracy and battery efficiency.

Classifiers	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9	Class 10	Class 11	Class 12	Class 13
1D-CNN	0.789	0.9	0.595	0.833	0.845	0.948	0.724	0.877	0.92	0.978	0.975	0.887	0.79
LSTM	0.738	0.833	0.36	0.767	0.722	0.879	0.635	0.795	0.874	0.981	0.977	0.815	0.656
Bi-LSTM	0.815	0.936	0.607	0.923	0.774	0.96	0.627	0.863	0.764	0.979	0.96	0.48	0.63
ANN	0.81	0.90	0.80	0.81	0.83	0.91	0.85	0.91	0.97	0.96	0.99	0.89	0.80
Hybrid	0.858	0.95	0.537	0.817	0.891	0.984	0.806	0.896	0.925	0.984	0.987	0.909	0.822

Table 6.2: F1-Scores of Different Classifiers for Each Activity for 3 seconds window length

Window	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9	Class 10	Class 11	Class 12	Class 13
1 second	0.738	0.928	0.5	0.858	0.841	0.884	0.59	0.88	0.92	0.98	0.98	0.872	0.755
3 seconds	0.858	0.95	0.537	0.817	0.891	0.984	0.806	0.896	0.925	0.984	0.987	0.909	0.822
5 seconds	0.861	0.968	0.542	0.801	0.886	0.974	0.62	0.89	0.903	0.989	0.984	0.872	0.74
10 seconds	0.868	0.937	0.51	0.784	0.863	0.979	0.723	0.91	0.923	0.979	1	0.905	0.77

Table 6.3: F1-Scores of Hybrid Model for Different Window Lengths for Each Activity

We present the F1-scores for various classes achieved by different neural networks, including our proposed hybrid model, in Table 6.2. Our hybrid model either matches or surpasses the performance of the other models across most of the classes (with class 3 and class 4 being the exception). In Table 6.3, we illustrate the F1-scores across different window lengths obtained by our hybrid model.
Additional experimentation was done for two different frequencies (1hz and 10hz) to understand the trend for different neural networks on both frequencies in figure 6.7. The bar chart compares the accuracy of various machine learning models : 1D-CNN, Bi-LSTM, Hybrid, LSTM, and ANN when trained and evaluated using two different data sampling frequencies: 1 Hz (blue bars) and 10 Hz (red bars).
1D-CNN and ANN models show relatively similar performance across both frequencies, with a slight improvement at 10 Hz. Bi-LSTM and LSTM models demonstrate a noticeable increase in accuracy when the sampling frequency is increased from 1 Hz to 10 Hz, indicating that these models benefit more from higher-frequency data. The Hybrid model consistently performs the best at both sampling rates, showing the highest accuracy among all the models tested.
This chart highlights the importance of data sampling frequency in model performance, with most models benefiting from the higher 10 Hz frequency. However, the degree of improvement varies across different model architectures.

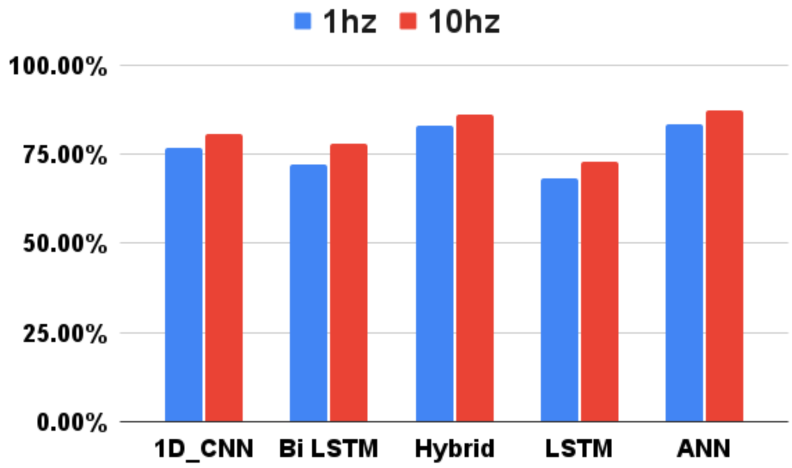


Figure 6.7: Comparison of Model Accuracy at Different Data Sampling Frequencies (1 Hz vs 10 Hz)

6.6 Conclusion and Limitations

Identifying everyday household activities using data from a smartwatch is quite challenging. Most past studies have focused on data transfer from sensor devices, high-frequency data, or multi-sensor approaches, often neglecting basic domestic tasks. In contrast, our innovative model can accurately distinguish thirteen different activities using low-frequency data directly on the smartwatch. Our results show that this Hybrid Model outperforms others in accuracy. We also found that a 10-second window length is the optimal choice, balancing the trade-off between accuracy and battery life efficiency.

Though our research has developed an innovative model that accurately distinguishes thirteen different household activities using low-frequency data directly on a smartwatch, it still has some limitations. Firstly, we persisted with the same volunteers as before, thus encountering the ongoing limitation of having a limited number of volunteers. Secondly, our objective was to achieve real-time prediction output on the smartwatch without relying on any external device or server. To achieve this, we deployed neural networks on the smartwatch. However, this approach required compromising on accuracy.

Chapter 7

Conclusions and Future Work

7.1 Research Conclusion

In chapter four [4] we address the challenging task of distinguishing 7 home activities within a home environment using data from wearable smartwatch sensors. We encountered a multifaceted problem where distinguishing such activities from low-frequency sensor data presented a non-trivial challenge. We noted that existing research in this area predominantly focused on high-frequency data or utilized multiple sensors, often overlooking basic home activities like sweeping, mopping, or cooking when working with smartwatches.

In response to these limitations, we introduced our innovative model, designed specifically to differentiate between seven distinct home activities using sensor data sampled at a low frequency, specifically 1Hz. Through a series of experiments, we demonstrated the effectiveness of our proposed approach in achieving superior accuracy when compared to flat models, particularly when incorporating both gyroscope and accelerometer data. For instance, with a 5Hz sampling frequency and 50% window overlapping, our Hierarchical model achieved an impressive accuracy of 97%, outperforming the best flat model, XGB, which achieved an accuracy of 93%. Similarly, when we reduced the sampling frequency to 1Hz with the same window overlapping, our proposed Hierarchical model maintained a substantial accuracy of 93%, surpassing XGB's accuracy of 85%.

This sets the stage for the exploration of non-energy activity thermogenesis monitoring using wearable technology, highlighting the potential of our approach in accurately recognizing and differentiating household activities through the analysis of low-frequency smartwatch sensor data. The promising results obtained from our experiments in the chapter underscore the relevance and significance of our research, providing a solid foundation for the subsequent work.

Building upon this foundation, in the fifth chapter [5], we delve deeper into the development of our innovative system. This system is designed to furnish users with real-time information about the specific activities they are engaged in, all through the use of a smartwatch worn on their wrist. Our system successfully identifies and classifies a total of 13 distinct activities, encompassing a wide range of daily tasks, such as cooking, sweeping, mopping, walking, and more. These activities include not only physical activities but also sedentary ones, like working on a laptop or watching TV. The process begins with the user initiating data collection by pressing a start button on the smartwatch. This raw sensor data is then transmitted to a central server for interpretation, where classification models are deployed to accurately recognize and categorize the user's ongoing activity. Communication between the smartwatch and the server is facilitated through Wi-Fi connectivity.

To evaluate the system's performance, we considered several crucial parameters, including battery depletion rate, preprocessing features (Static and ECDF), file upload rate, data sampling rate (1Hz or 10Hz), and the length of the timestamp window. Understanding the battery depletion rate is of paramount importance, as it empowers users to make informed decisions about the selection of optimal parameters. While the classification models currently reside on the server due to computational limitations on the smartwatch, our future goal was to optimize the system for on-device deployment, thereby reducing dependence on internet connectivity.

One noteworthy aspect of our system was its ability to handle situations where the user's activity does not fall within the predefined 13 categories. In such cases, our model employs a Fourteenth class, aptly named "OTHERS." This class serves as a catch-all for unclassified activities, ensuring that users receive meaningful feedback even when their actions do not align with the predefined list.

Throughout our investigation, we explored various clustering techniques in an attempt to understand and classify activities within the "OTHERS" class. However, the complexity of high-dimensional data made it challenging to differentiate genuine outliers from legitimate activities. Consequently, we introduced the concept of the dominating class, where the system assigns that label if a particular class constitutes 60% or more of instances within a given window. In instances where no single class achieves this threshold, the label defaults to "OTHERS." This approach provides flexibility, as the threshold can be adjusted to 40% or 80%, depending on the desired level of leniency or strictness in our system.

This chapter illustrates the development and functionality of our activity recognition system, highlighting its capability, robustness, and potential for future enhancements, particularly with the objective of deploying classification models directly on the smartwatch to further reduce reliance on external servers and internet connectivity.

Thus in chapter six [6], we took on the formidable challenge of distinguishing everyday household activities from smartwatch sensor data without relying on Wi-Fi or Bluetooth connectivity. In contrast to the prevailing research trends, which often prioritize high-frequency data or multi-sensor approaches, our innovative model set out to differentiate thirteen activities using low-frequency data sampling. By doing so, we eliminated the need for external connectivity, not only enhancing user privacy but also making real-world implementation more feasible and user-friendly.

Our empirical evaluations have yielded compelling results, firmly establishing the superior performance of our proposed Hybrid Model in terms of accuracy when compared to other models. This is a significant achievement, as it means that our model can reliably classify activities even when users are operating their smartwatches in environments where Wi-Fi or paired devices are not readily available. This extends the utility of our system to a broader range of practical scenarios, adding to its real-world applicability.

One of the key takeaways from our research pertains to the choice of window length for data analysis. Our findings highlight that a window length of 10 seconds strikes a harmonious balance between accuracy and battery efficiency. This practical decision allows users to monitor their activities without the undue burden of excessive power consumption, thus ensuring the continued user-friendliness of our system.

Moreover, it is worth noting that we have deployed neural networks, including 1d-CNN, LSTM, and Bidirectional LSTM. Through extensive experimentation, we observed that the highest accuracy was obtained with the Bidirectional LSTM model, while the 1d-CNN model exhibited the least battery consumption. These insights led us to propose a Hybrid Model that leverages both 1d-CNN and Bidirectional LSTM, achieving an overall improvement in accuracy compared to the standalone models. The combination of these models has been seamlessly integrated into the smartwatch using TensorFlow Lite (TFLite), eliminating the need to transfer data to servers or systems to predict accuracy.

Chapter six of our thesis stands as a significant milestone in our research journey. It demonstrates our ability to successfully address the challenge of recognizing household activities using a low-frequency, standalone smartwatch system, while simultaneously ensuring resource and operational efficiency. The insights gained from this chapter will guide further development and refinement of our system, advancing the field of activity recognition using wearable technology.

Our research journey has allowed us to create models that recognize activities using wearable technology at home. These models show great promise, potentially benefiting areas like healthcare and smart home automation. By seamlessly integrating neural networks into standalone smartwatches, we have made a significant step towards more user-friendly and efficient activity recognition systems.

As we wrap up this thesis, it's clear that our work goes beyond the individual chapters. Each innovative solution we have presented comes together to create a well-rounded approach to activity recognition. The success of our models and the insights from our experiments all contribute to advancing the field. Looking ahead, future research could focus on optimizing on-device deployment, improving classification algorithms, and broadening the range of activities recognized.

In essence, our thesis address the immediate challenges of activity recognition while laying the groundwork for future advancements in wearable technology. By taking an interdisciplinary approach, we have bridged gaps in existing research and paved the way for more refined, efficient, and user-friendly solutions in the ever-evolving field of home activity recognition.

7.2 Contributions to Existing Knowledge and Distinctiveness of This Research

This section highlights how this research sets apart from the existing work done in the field of NEAT activity detection. This also showcases how this research contributes to existing research and scales it.

7.2.1 Contribution to Existing Knowledge

1. **Comprehensive NEAT Activity Monitoring:** This research extends the understanding of NEAT by including a detailed and thorough analysis of 13 activities which covers a wide range of daily activities using wearable technology. Other research generally limited their approach to working with high-intensity activities or specific isolated activities. With this research, we aim to broaden the scope of activity detection by performing daily activities and generating a holistic view of NEAT behavior.

2. **Low-Frequency Data Utilisation:** In past research, there has been a prevailing notion that high-frequency data is necessary for accurate activity detection, and in order to curb that, we successfully built a hierarchical classification system that operates at a low sampling frequency of 1Hz. This is a significant contribution as it showcases that practical and long-term monitoring can be achieved without compromising on accuracy. This, in turn, helps to build user compliance and broaden the scope for a scalable solution.

3. **Hierarchical Classification Model:** The hierarchical classification model presented in this research represents a novel approach to activity recognition. By breaking complex multi-class classification tasks into simpler binary decisions, this model improves accuracy and robustness in classifying between similar activities. The creation of this model paves the path for future studies to improve the accuracy of activity detection.

4. **Robust Feature Extraction Techniques:** This research sets a new standard for preparing data in NEAT activity detection by using strong statistical methods to extract features from low-frequency data. By effectively using features like max, mean, mode, variance and lower quartile it ensures enough information is captured for accurate classification, even when data is sparse.

5. **Practical Implications for Health Monitoring:** This research helps improve health monitoring and personal fitness by providing practical and easy-to-use solutions. The system gives real-time feedback and uses little power, allowing continuous use. This supports long-term health changes and better lifestyle habits.

7.2.2 Distinctiveness of This Research

1. What sets this research apart is using low frequency (1Hz-10Hz) while maintaining a high classification accuracy. This addresses a significant gap in the existing research where high-frequency data collection often leads to impractical solutions for everyday use.

2. Including the "Others" class for unclassified activities highlights the system's capability to classify the class if it is not part of any of the 13 activities we choose. This approach ensures that our research remains relevant and informative across a wide range of user activities and thus improves upon activity recognition.

3. The model's performance was evaluated using real-world data collected from volunteers in a typical home environment. Previous research relies on data collected from controlled lab environments, which may not reflect accurately upon everyday conditions. This approach ensures that our findings can be directly applied in real-life scenarios, enhancing the external validity of the research.

4. Our research's primary focus of the research was to build an energy-efficient solution, such as on-device processing and low-frequency data usage. The major benefit of using energy-efficient solutions is better practical applicability. This sets our research apart from other research that may achieve high accuracy at the cost of high energy consumption and limited battery life.

5. As we demonstrated that accurate NEAT activity detection can be achieved with low-frequency data, this research paves the way for more efficient and user-friendly wearable devices. The advancement can influence the design and functionality of future wearable technologies, promoting broader adoption and sustainable usage.

Conclusion
In summary, this research makes significant contributions to the field of NEAT activity detection by demonstrating the feasibility of low-frequency data utilisation, introducing a hierarchical classification model, and emphasising practical, user-friendly solutions. The study sets itself apart by building energy-efficient solutions with real-world applicability, ensuring the findings are innovative and directly relevant to improving health monitoring technologies. Table 7.1 showcases our objective and fulfilment's of building a NEAT Activity detection system.

7.3 Limitations of the research

Our research has a few limitations that need to be acknowledged.

1. The range of activities we performed was limited because we started our research during the COVID-19 pandemic. This meant we could only perform household activities during curfews in India. This restriction narrowed the activities we could analyse. Still overcoming the challenge, we tested out 13 different activities that we could perform indoors and outdoors.

2. There was a limited number of volunteers. At first, only family members could participate due to lockdown restrictions. Later, we included ten volunteers from our

Objective	**Fulfillment**	**Evidence**
Develop a Robust NEAT Activity Detection System	Implementation of a hierarchical model leveraging low-frequency data from wearable sensors, ensuring accurate classification of various NEAT activities. Advanced feature extraction techniques (Statistical and ECDF) enhance robustness and reliability.	High accuracy and robustness demonstrated through real-world data collected from volunteers. Effective handling of complex multi-class classification tasks, accurately detecting activities like walking, standing, and cooking.
Ensure Energy Efficiency and Practicality	Designed for energy efficiency and practicality with a low sampling frequency, balancing data accuracy and battery life for long-term use without frequent recharging. Enhances user compliance and practical application feasibility.	Significant reduction in power consumption compared to high-frequency methods. Extended battery life and minimized energy usage through on-device processing, making the system practical and efficient for everyday use.
Incorporate a Comprehensive Range of NEAT Activities	Successfully included a wide range of NEAT activities representing daily movements, capturing a holistic view of NEAT behaviors and providing valuable insights into physical activity patterns.	Accurately classified thirteen different activities, chosen for their prevalence and representativeness in typical home and outdoor environments. Demonstrates comprehensive coverage and real-world applicability.
Address Privacy Concerns	Avoids intrusive vision-based systems, using non-intrusive wearable sensors to monitor movements without capturing sensitive visual data, enhancing user trust and acceptance.	Prioritizes privacy by utilizing accelerometer and gyroscope data, avoiding visual or audio information. Sets the research apart from existing solutions that raise privacy issues with continuous video recording.
Provide Real-Time Feedback	Provides real-time feedback to users based on NEAT activity levels, motivating increased physical activity and healthier lifestyles through immediate insights.	Demonstrated capability for real-time data processing and instant feedback using on-device neural networks. Users receive immediate updates on activity levels, promoting consistent engagement and positive behavior changes.

Table 7.1: NEAT Activity Detection System Objectives and Fulfillment

neighbourhood for data collection. However, the ongoing COVID restrictions made getting a larger, more diverse group difficult. This small group of volunteers may affect how broadly our findings can be applied.

3. In Chapter 5 of our research, we mentioned using the server for real-time feedback to the user. We needed a constant connection to WiFi and Bluetooth to execute this because they couldn't run directly on the smartwatch. This constant connectivity might be problematic in real-world situations where WiFi isn't always available. However, later in research, we mitigated this in Chapter 6 and deployed models directly in the smartwatches.

4. To make real-time predictions on the smartwatch without using external devices or servers, we ran neural networks directly on the smartwatch. However, we had to give up better accuracy generated through flat models compared to neural networks to use this method.

5. We used only two smartwatches of the same brand to collect data and ensure we did not have a disparity in the data set. The application was built in Android (Wear OS), so any device that uses the same OS can easily use it. We must use multiple devices to scale and generalise the findings in future research.

Despite these challenges, our research shows that using low-frequency smartwatch data for recognising activities is promising and sets the stage for future studies.

7.4 Future work

In recent years, NEAT (Non-Exercise Activity Thermogenesis) activity detection using smartwatches has emerged as a promising technology, facilitating a deeper understanding of individual physical activity and promoting healthier lifestyles. As technology continues to evolve, several future scopes are poised to enhance the efficacy and practicality of NEAT activity detection. In this section, we will explore critical areas of development that promise to shape the future of this field.

7.4.1 Further battery optimisation

When it comes to deploying neural networks for the monitoring of Non-Exercise Activity Thermogenesis (NEAT) activities on smartwatches, the main challenge is the constrained battery life of these compact wearables. Smartwatches, despite their versatility, come equipped with miniature power sources that inherently face limitations in terms of energy capacity. This limitation results from the appropriate balance between power and form factor, as smartwatches aim to accommodate various functions while remaining comfortable to wear. The continuous and resource-intensive monitoring required for NEAT activities can rapidly deplete the smartwatch's battery, bringing forth several critical issues. Firstly, this battery limitation significantly affects the user experience. Frequent recharging or a short battery life can disrupt the seamless integration of the smartwatch into daily routines. The inconvenience of frequent charging and the constant concern about running out of battery can deter users from fully utilizing the smartwatch's potential. Secondly, the limited battery life directly influences the scope of NEAT activity monitoring. A shorter battery life may necessitate fewer monitoring intervals, limiting the ability to capture a comprehensive picture of a user's daily activities.

Innovative Solutions to Optimize Energy Efficiency:

Researchers are actively engaged in developing innovative solutions to address these challenges and optimize the energy efficiency of neural networks and sensor operations within smartwatches. Several key strategies are being employed:

Model Compression: One of the primary strategies involves model compression. Researchers are working diligently to reduce the complexity of neural networks without compromising accuracy. This reduction in model size results in more streamlined computations during inference, leading to significantly lower energy demands. The development of compressed or lightweight models ensures that activity recognition remains highly accurate while substantially conserving battery life.

Sensor Management and Fusion for Efficient Data Collection: Smartwatches are equipped with an array of sensors, including accelerometers and gyroscopes. To conserve power during periods of inactivity, researchers are developing intelligent sensor management algorithms that selectively activate sensors only when necessary, allowing the smartwatch to gather data efficiently without constantly draining the battery. By turning sensors on and off strategically, energy consumption is minimized, and accurate monitoring is maintained. Integrating advanced, energy-efficient sensors capable of capturing precise data enables intermittent or event-triggered data acquisition, reducing the need for continuous monitoring. Smartwatches equipped with these sensors can intelligently collect data only when relevant, lowering the energy overhead associated with non-essential monitoring. Ongoing research in sensor fusion techniques and data collection methodologies further enhances energy efficiency, ensuring smartwatches perform their tasks without overtaxing their batteries.

Low-Power Model Architectures: Mitigating the battery life challenge involves a crucial focus on the development and deployment of low-power neural networks architectures. These models are tailored to the specific constraints of smartwatches and edge devices. Notable examples include MobileNet and TinyML, designed to minimize energy consumption while preserving high-performance capabilities. Such models enable smartwatches to operate with extended battery life, ensuring that the wearers can enjoy uninterrupted activity monitoring without compromising the accuracy of detection. This field of research is continually evolving, with researchers pushing the boundaries of energy-efficient model design.

7.4.2 Integrating Other Types of Sensors

Future research can make NEAT activity data more accurate and detailed by adding more types of sensors. Using heart rate monitors, barometers, and GPS modules can give extra information, helping us understand user activities better. For instance, heart rate monitors can tell us how intense an activity is, giving us more precise details about NEAT activities. Barometric pressure sensors can notice changes in height, like when someone climbs stairs, and GPS modules can show the location of outdoor activities, making it easier to track activities in different places.
However, adding more sensors comes with challenges. More sensors can use up more power, so we need to find ways to save energy to keep the devices working for a long time. Also, combining and syncing data from different sensors can be tricky and might need advanced computing power.

7.4.3 Applying Models in Different Environments

To make sure NEAT activity detection models are strong and work well everywhere, future studies should test these models in diverse environments. This means trying them out in cities, countryside, indoors, outdoors, and in various weather conditions. Testing in these different places will help the models recognize NEAT activities accurately no matter where they are used, making them more reliable.
Testing with a diverse groups of people and in different settings will also make the models more useful and trustworthy for everyone. However, different places have different types of noise and disruptions that the models need to handle, which is a big challenge. Also, collecting enough varied data from many places takes a lot of planning and resources, making it a tough task.

7.4.4 Domain Adaptation

In the world of NEAT (Non-Exercise Activity Thermogenesis) activity detection using smartwatches, research is on the brink of some exciting breakthroughs. The future isn't just about making these systems work better; it's about making them smarter and more adaptable. Imagine smartwatches that can seamlessly adjust to various situations, cater to each person's unique habits, and stay reliable even when conditions change unexpectedly, all while maintaining top performance. Here's a look at the future research paths that could make this vision a reality:
Personalized Models: The path forward in NEAT activity detection research leads to the development of models that are not just intelligent but profoundly personalized. This envisages a future where research efforts empower AI to understand individuals at a personal level. It involves crafting models that can discern the nuances in an individual's movements and behaviors, akin to the way a close friend would. The research paper "Personalized Human Activity Recognition Using Convolutional Neural Networks" by Rokni et al. in 2018 [105] serves as a guiding light for this direction. Future research

will likely focus on models that adapt to be as unique as each individual, accurately recognizing activities, regardless of how they are performed.

Transfer Learning: In the ever-evolving world of AI research, the ability to learn quickly and adapt is becoming increasingly crucial. Transfer learning is at the heart of this capability. Imagine AI models that can leverage the experience of previous models and rapidly fine-tune themselves with specific new data, much like a seasoned professional mastering a new job. A groundbreaking study by Chen et al., "Fedhealth: A federated transfer learning framework for wearable healthcare" (2020) [106], showcased the immense potential of this technique in healthcare. Extending this approach to NEAT activity detection holds exciting possibilities. The future will likely see AI systems that are not only more reliable but also more versatile, ready to meet the challenges of our dynamic world.

In conclusion, the future of NEAT activity detection research holds the promise of personalized, adaptable, and resilient systems, much like a close friend who understands, learns quickly, and effortlessly adjusts to the world around them. This future envisions NEAT activity detection that is not just accurate and reliable but deeply relatable, ushering in a new era of human-AI collaboration.

7.4.5 Use of Functional API:

The conventional sequential API (Figure 7.1) for neural networks on smartwatches can be limiting. Future developments will explore the use of the Functional API (Figure 7.2) to build more complex and versatile models that can better capture and interpret NEAT activities.

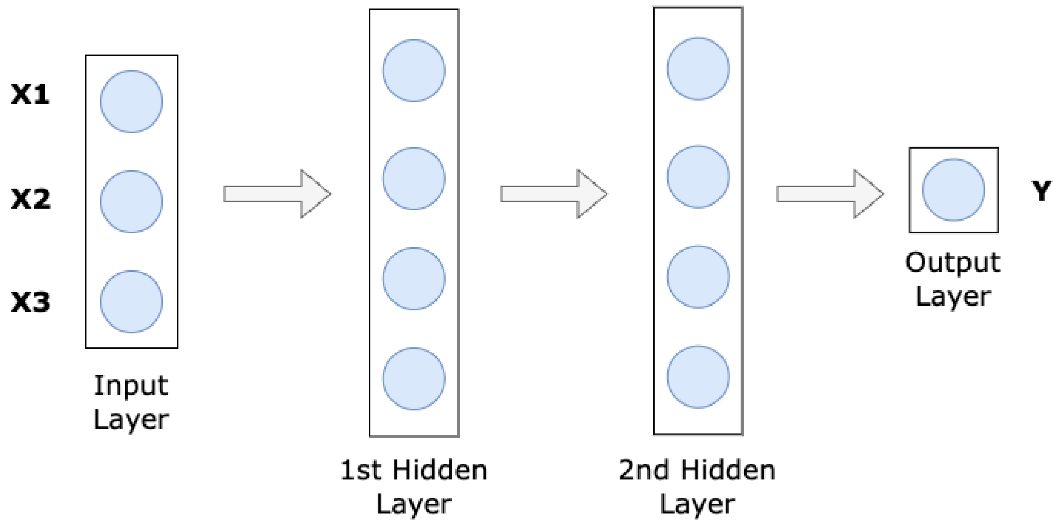


Figure 7.1: Sequential API

Flexible Model Architectures: The Functional API allows for the creation of flexible, non-linear model architectures. This flexibility is particularly useful when dealing with complex NEAT activity patterns that may not follow a strictly sequential structure. While

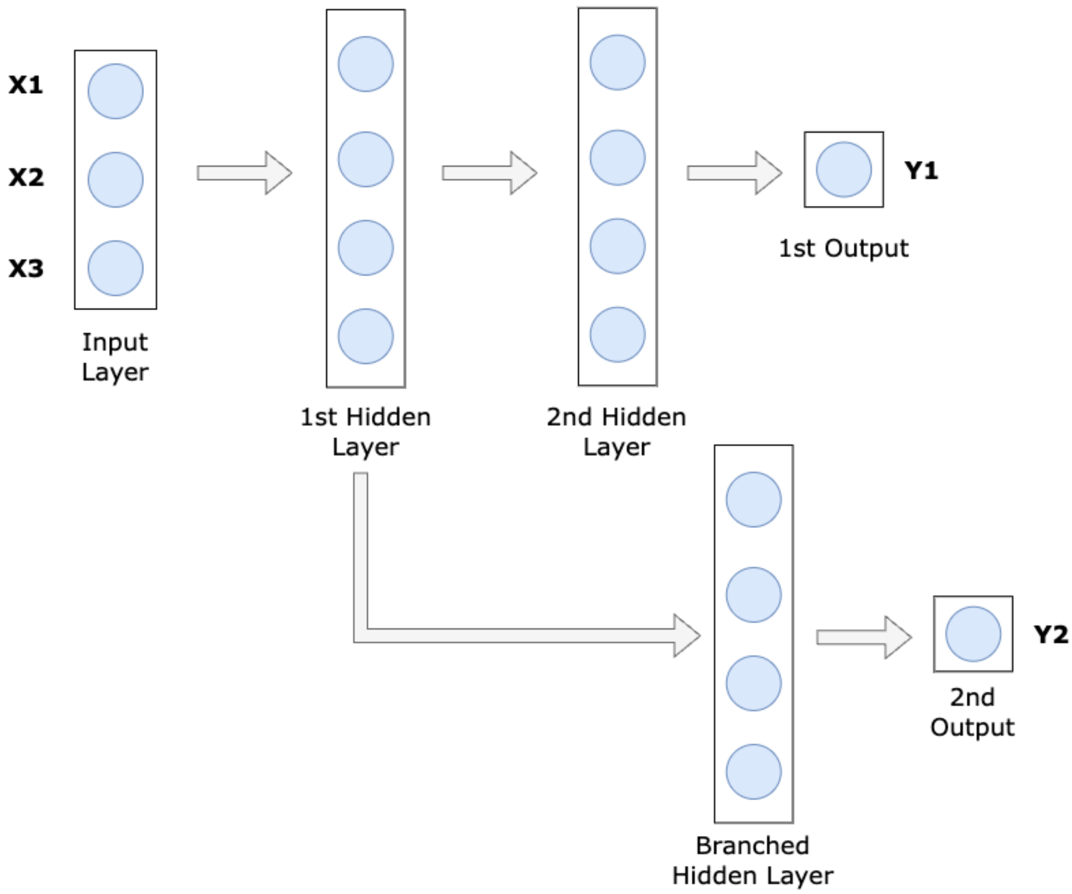


Figure 7.2: Functional API

specific papers on the use of the Functional API in smartwatch-based activity detection may be limited, research on the Functional API's applications in neural networks can be adapted to this context.

Feature Fusion: Leveraging the Functional API, researchers can develop models that efficiently fuse data from multiple sensors and sources. Feature fusion is essential for better data representation and enhanced activity detection. Papers like "Deep Feature Fusion for COVID-19 Detection Using Chest X-rays" by Siva Sundara Rajan and Sudharsan (2021) illustrate the concept of feature fusion in deep learning.

In conclusion, the future of NEAT activity detection using smartwatches and neural networks holds great promise. Future developments in Battery Optimization, Integrating Other Types of Sensors, Applying Models in Different Environments, Exploring Real-Time Feedback Mechanisms, Domain Adaptation, and the use of the Functional API will significantly enhance the effectiveness and practicality of these devices, ultimately contributing to improved health and well-being for users. Researchers and developers in this field should draw inspiration from existing research and adapt these insights to the unique challenges and opportunities presented by smartwatch-based NEAT activity detection.

References

[1] James A Levine, Lorraine M Lanningham-Foster, Shelly K McCrady, Alisa C Krizan, Leslie R Olson, Paul H Kane, Michael D Jensen, and Matthew M Clark. Interindividual variation in posture allocation: possible role in human obesity. *Science*, 307(5709):584–586, 2005.

[2] Genevieve N Healy, Charles E Matthews, David W Dunstan, Elisabeth AH Winkler, and Neville Owen. Sedentary time and cardio-metabolic biomarkers in us adults: Nhanes 2003–06. *European heart journal*, 32(5):590–597, 2011.

[3] Gideon Remfry. Could the blue zone's exercise secrets hold the key to vital ageing?

[4] Dale S Bond, John M Jakicic, Sivamainthan Vithiananthan, J Graham Thomas, Tricia M Leahey, Harry C Sax, Dieter Pohl, GD Roye, Beth A Ryder, and Rena R Wing. Objective quantification of physical activity in bariatric surgery candidates and normal-weight controls. *Surgery for Obesity and Related Diseases*, 6(1):72–78, 2010.

[5] Accuracy of armband monitors for measuring daily energy expenditure in healthy adults. *Medicine and science in sports and exercise*, 42(11):2134–2140, 2010.

[6] Marc T Hamilton, Genevieve N Healy, David W Dunstan, Theodore W Zderic, and Neville Owen. Too little exercise and too much sitting: inactivity physiology and the need for new recommendations on sedentary behavior. *Current cardiovascular risk reports*, 2(4):292–298, 2008.

[7] Ricardo Chavarriaga, Hesam Sagha, Alberto Calatroni, Sundara Tejaswi Digumarti, Gerhard Tröster, José del R Millán, and Daniel Roggen. The opportunity challenge: A benchmark database for on-body sensor-based activity recognition. *Pattern Recognition Letters*, 34(15):2033–2042, 2013.

[8] Jorge-Luis Reyes-Ortiz, Luca Oneto, Alessandro Ghio, Albert Samá, Davide Anguita, and Xavier Parra. Human activity recognition on smartphones with awareness of basic activities and postural transitions. In *Artificial Neural Networks and Machine Learning–ICANN 2014: 24th International Conference on Artificial Neural Networks, Hamburg, Germany, September 15-19, 2014. Proceedings 24*, pages 177–184. Springer, 2014.

[9] Heesu Park, Suh-Yeon Dong, Miran Lee, and Inchan Youn. The role of heart-rate variability parameters in activity recognition and energy-expenditure estimation using wearable sensors. *Sensors*, 17(7):1698, 2017.

[10] Weijun Tao, Tao Liu, Rencheng Zheng, and Hutian Feng. Gait analysis using wearable sensors. *Sensors*, 12(2):2255–2283, 2012.

[11] Mi Zhang and Alexander A Sawchuk. Motion primitive-based human activity recognition using a bag-of-features approach. In *Proceedings of the 2nd ACM SIGHIT international health informatics symposium*, pages 631–640, 2012.

[12] Tahmina Zebin, Patricia J Scully, and Krikor B Ozanyan. Evaluation of supervised classification algorithms for human activity recognition with inertial sensors. In *2017 IEEE SENSORS*, pages 1–3. IEEE, 2017.

[13] Hongnian Yu, Shuang Cang, and Yan Wang. A review of sensor selection, sensor devices and sensor deployment for wearable sensor-based human activity recognition systems. In *2016 10th international conference on software, knowledge, information management & applications (skima)*, pages 250–257. IEEE, 2016.

[14] Ahmad Jalal, Naeha Sarif, Jeong Tai Kim, and Tae-Seong Kim. Human activity recognition via recognized body parts of human depth silhouettes for residents monitoring services at smart home. *Indoor and built environment*, 22(1):271–279, 2013.

[15] Tomasz Hachaj. Improving human motion classification by applying bagging and symmetry to pca-based features. *Symmetry*, 11(10):1264, 2019.

[16] Ahmad Jalal, Shaharyar Kamal, and Daijin Kim. A depth video sensor-based life-logging human activity recognition system for elderly care in smart indoor environments. *Sensors*, 14(7):11735–11759, 2014.

[17] Seba Susan, Prachi Agrawal, Minni Mittal, and Srishti Bansal. New shape descriptor in the context of edge continuity. *CAAI Transactions on Intelligence Technology*, 4 (2):101–109, 2019.

[18] Yu Tingting, Wang Junqian, Wu Lintai, and Xu Yong. Three-stage network for age estimation. *CAAI Transactions on Intelligence Technology*, 4(2):122–126, 2019.

[19] Subhas Chandra Mukhopadhyay. Wearable sensors for human activity monitoring: A review. *IEEE sensors journal*, 15(3):1321–1330, 2014.

[20] Maryam Naseer Malik, Muhammad Awais Azam, Muhammad Ehatisham-Ul-Haq, Waleed Ejaz, and Asra Khalid. Adlauth: Passive authentication based on activity of daily living using heterogeneous sensing in smart cities. *Sensors*, 19(11):2466, 2019.

[21] Eduardo Casilari, Moisés Álvarez-Marco, and Francisco García-Lagos. A study of the use of gyroscope measurements in wearable fall detection systems. *Symmetry*, 12(4):649, 2020.

[22] Jie Yin, Qiang Yang, and Jeffrey Junfeng Pan. Sensor-based abnormal human-activity detection. *IEEE transactions on knowledge and data engineering*, 20(8):1082–1090, 2008.

[23] Lingxiang Zheng, Dihong Wu, Xiaoyang Ruan, Shaolin Weng, Ao Peng, Biyu Tang, Hai Lu, Haibin Shi, and Huiru Zheng. A novel energy-efficient approach for human activity recognition. *Sensors*, 17(9):2064, 2017.

[24] Asmita Nandy, Jayita Saha, Chandreyee Chowdhury, and Kundan PD Singh. Detailed human activity recognition using wearable sensor and smartphones. In *2019 International Conference on Opto-Electronics and Applied Optics (Optronix)*, pages 1–6. IEEE, 2019.

[25] Asmita Nandy, Jayita Saha, and Chandreyee Chowdhury. Novel features for intensive human activity recognition based on wearable and smartphone sensors. *Microsystem Technologies*, 26:1889–1903, 2020.

[26] Ishan RoyChowdhury, Jayita Saha, and Chandreyee Chowdhury. Detailed activity recognition with smartphones. In *2018 Fifth International Conference on Emerging Applications of Information Technology (EAIT)*, pages 1–4. IEEE, 2018.

[27] Farzan Majeed Noori, Michael Riegler, Md Zia Uddin, and Jim Torresen. Human activity recognition from multiple sensors data using multi-fusion representations and cnns. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 16(2):1–19, 2020.

[28] Kriti Kumari, Pankaj H Chandankhede, and Abhijit S Titarmare. Design of human activity recognition system using body sensor networks. In *2021 6th international conference on communication and electronics systems (ICCES)*, pages 1011–1016. IEEE, 2021.

[29] Takuma Akiduki, Jun Nagasawa, Zhong Zhang, Yuto Omae, Toshiya Arakawa, and Hirotaka Takahashi. Inattentive driving detection using body-worn sensors: feasibility study. *Sensors*, 22(1):352, 2022.

[30] QU Tang, Dinesh John, Binod Thapa-Chhetry, Diego Jose Arguello, and Stephen Intille. Posture and physical activity detection: impact of number of sensors and feature type. *Medicine and science in sports and exercise*, 52(8):1834, 2020.

[31] Benjamin D Maylor, Charlotte L Edwardson, Alexandra M Clarke-Cornwell, Melanie J Davies, Nathan P Dawkins, David W Dunstan, Kamlesh Khunti, Tom Yates, and Alex V Rowlands. Physical activity assessed by wrist and thigh worn accelerometry and associations with cardiometabolic health. *Sensors*, 23(17):7353, 2023.

[32] Muhammad Awais, Luca Palmerini, and Lorenzo Chiari. Physical activity classification using body-worn inertial sensors in a multi-sensor setup. In *2016 IEEE*

2nd International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI), pages 1–4. IEEE, 2016.

[33] Ferhat Attal, Samer Mohammed, Mariam Dedabrishvili, Faicel Chamroukhi, Latifa Oukhellou, and Yacine Amirat. Physical human activity recognition using wearable sensors. *Sensors*, 15(12):31314–31338, 2015.

[34] Farman Ullah, Asif Iqbal, Sumbul Iqbal, Daehan Kwak, Hafeez Anwar, Ajmal Khan, Rehmat Ullah, Huma Siddique, and Kyung-Sup Kwak. A framework for maternal physical activities and health monitoring using wearable sensors. *Sensors*, 21(15): 4949, 2021.

[35] Umair Saeed, Kamlesh Kumar, Asif Ali Laghari, Mansoor Ahmed Khuhro, Noman Islam, Ghulam Muhammad Shaikh, Fahad Hussain, and Aftab Ahmed Shaikh. A comparative analysis of classification techniques for human activity recognition using wearable sensors and smart-phones. *EAI Endorsed Transactions on Pervasive Health and Technology*, 8(30):e1–e1, 2022.

[36] Sujatha Jamuna Anand, S Magesh, and I Arockiamary. Contemporary human activity recognition based predictions by sensors using random forest classifier. *Journal of Computational and Theoretical Nanoscience*, 18(4):1243–1250, 2021.

[37] Bettina Eska, Marco Philip, Pawel W Wozniak, Albrecht Schmidt, and Jakob Karolus. [don't] let the bodies hiit the floor: Fostering body awareness in fast-paced physical activity using body-worn sensors. *Proceedings of the ACM on Human-Computer Interaction*, 7(MHCI):1–27, 2023.

[38] Sreedevi Uday, C Jyotsna, and J Amudha. Detection of stress using wearable sensors in iot platform. In *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, pages 492–498. IEEE, 2018.

[39] Alan Kevin Bourke, Espen Alexander F Ihlen, Ronny Bergquist, Per Bendik Wik, Beatrix Vereijken, and Jorunn L Helbostad. A physical activity reference data-set recorded from older adults using body-worn inertial sensors and video technology—the adapt study data-set. *Sensors*, 17(3):559, 2017.

[40] Heba Aly and Mohamed A Ismail. ubimonitor: Intelligent fusion of body-worn sensors for real-time human activity recognition. In *Proceedings of the 30th Annual ACM Symposium on Applied Computing*, pages 563–568, 2015.

[41] Anis Davoudi, Mamoun T Mardini, David Nelson, Fahd Albinali, Sanjay Ranka, Parisa Rashidi, Todd M Manini, et al. The effect of sensor placement and number on physical activity recognition and energy expenditure estimation in older adults: Validation study. *JMIR mHealth and uHealth*, 9(5):e23681, 2021.

[42] Johann P Wolff, Florian Grützmacher, Arne Wellnitz, and Christian Haubelt. Activity recognition using head worn inertial sensors. In *Proceedings of the 5th international Workshop on Sensor-based Activity Recognition and Interaction*, pages 1–7, 2018.

[43] Yunyoung Nam, Seungmin Rho, and Chulung Lee. Physical activity recognition using multiple sensors embedded in a wearable device. *ACM Transactions on Embedded Computing Systems (TECS)*, 12(2):1–14, 2013.

[44] Alexander Hoelzemann, Henry Odoemelem, and Kristof Van Laerhoven. Using an in-ear wearable to annotate activity data across multiple inertial sensors. In *Proceedings of the 1st International Workshop on Earable Computing*, pages 14–19, 2019.

[45] Seyed Ali Rokni and Hassan Ghasemzadeh. Synchronous dynamic view learning: a framework for autonomous training of activity recognition models using wearable sensors. In *Proceedings of the 16th ACM/IEEE International Conference on Information Processing in Sensor Networks*, pages 79–90, 2017.

[46] John Staudenmayer, David Pober, Scott Crouter, David Bassett, and Patty Freedson. An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. *Journal of applied physiology*, 107(4):1300–1307, 2009.

[47] Silvia Costa, David Ogilvie, Alice Dalton, Kate Westgate, Søren Brage, and Jenna Panter. Quantifying the physical activity energy expenditure of commuters using a combination of global positioning system and combined heart rate and movement sensors. *Preventive medicine*, 81:339–344, 2015.

[48] P Dong Suh-Yeon, L Miran, and Y Inchan. The role of heart-rate variability parameters in activity recognition and energy-expenditure estimation using wearable sensors. *Sensors (Basel)*, (7):1698, 2017.

[49] Patty S Freedson and Kelly Miller. Objective monitoring of physical activity using motion sensors and heart rate. *Research quarterly for exercise and sport*, 71(sup2): 21–29, 2000.

[50] Matthew L Eckard, Hana C Kuwabara, and Carole M Van Camp. Using heart rate as a physical activity metric. *Journal of Applied Behavior Analysis*, 52(3):718–732, 2019.

[51] Jennifer A Schrack, Andrew Leroux, Jerome L Fleg, Vadim Zipunnikov, Eleanor M Simonsick, Stephanie A Studenski, Ciprian Crainiceanu, and Luigi Ferrucci. Using heart rate and accelerometry to define quantity and intensity of physical activity in older adults. *The Journals of Gerontology: Series A*, 73(5):668–675, 2018.

[52] Ardo Allik, Kristjan Pilt, Moonika Viigimäe, Ivo Fridolin, and Gert Jervan. A novel physical fatigue assessment method utilizing heart rate variability and pulse arrival time towards personalized feedback with wearable sensors. *Sensors*, 22(4): 1680, 2022.

[53] Antonis Michanikou, Panayiotis Kouis, Kostas Karanicolas, and Panayiotis K Yiallouros. Setup of consumer wearable devices for exposure and health monitoring in population studies. *JoVE (Journal of Visualized Experiments)*, (192):e63275, 2023.

[54] Carmen GRIGOROIU, Teodora WESSELLY, Adrian PRICOP, Mihaela NETOLITZCHI, and Camelia BRANEȚ. Individualisation of the training process of tennis players using the heart rate monitor. *Discobolul-Physical Education, Sport & Kinetotherapy Journal*, 60(4), 2021.

[55] Maryam Eskandari, Saman Parvaneh, Hossein Ehsani, Mindy Fain, and Nima Toosizadeh. Frailty identification using heart rate dynamics: A deep learning approach. *IEEE Journal of Biomedical and Health Informatics*, 26(7):3409–3417, 2022.

[56] Hana Charvátová, Aleš Procházka, Saeed Vaseghi, Oldřich Vyšata, and Martin Vališ. Gps-based analysis of physical activities using positioning and heart rate cycling data. *Signal, Image and Video Processing*, 11:251–258, 2017.

[57] Hoda Allahbakhshi, Lindsey Conrow, Babak Naimi, and Robert Weibel. Using accelerometer and gps data for real-life physical activity type detection. *Sensors*, 20 (3):588, 2020.

[58] Motofumi Nakanishi, Shintaro Izumi, Sho Nagayoshi, Hironori Sato, Hiroshi Kawaguchi, Masahiko Yoshimoto, Takafumi Ando, Satoshi Nakae, Chiyoko Usui, Tomoko Aoyama, et al. Physical activity group classification algorithm using triaxial acceleration and heart rate. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 510–513. IEEE, 2015.

[59] David Joseph Muggeridge, Kirsty Hickson, Aimie Victoria Davies, Oonagh M Giggins, Ian L Megson, Trish Gorely, and Daniel R Crabtree. Measurement of heart rate using the polar oh1 and fitbit charge 3 wearable devices in healthy adults during light, moderate, vigorous, and sprint-based exercise: validation study. *JMIR mHealth and uHealth*, 9(3):e25313, 2021.

[60] Daniel Garcia-Gonzalez, Daniel Rivero, Enrique Fernandez-Blanco, and Miguel R Luaces. A public domain dataset for real-life human activity recognition using smartphone sensors. *Sensors*, 20(8):2200, 2020.

[61] Lily J Koffman, Ciprian M Crainiceanu, Ryan T Roemmich, and Margaret A French. Identifying unique subgroups of individuals with stroke using heart rate and steps to characterize physical activity. *Journal of the American Heart Association*, page e030577, 2023.

[62] Kestens Yan, Barnett Tracie, Mathieu Marie-Ève, Henderson Mélanie, Bigras Jean-Luc, Thierry Benoit, Maxime St-Onge, Lambert Marie, et al. Innovation through wearable sensors to collect real-life data among pediatric patients with cardiometabolic risk factors. *International journal of pediatrics*, 2014.

[63] Lin Li, Yuanyuan Liu, Yue Gu, Zhe Zhu, et al. Application of heart rate combined with acceleration motion sensor in sports dance teaching. *Journal of Sensors*, 2022.

[64] Claudia Floris, Sarah Solbiati, Federica Landreani, Gianfranco Damato, Bruno Lenzi, Valentino Megale, and Enrico Gianluca Caiani. Feasibility of heart rate and respiratory rate estimation by inertial sensors embedded in a virtual reality headset. *Sensors*, 20(24):7168, 2020.

[65] Kruthi Doddabasappla and Rushi Vyas. Analysis of frequency spectral features of coughing activity from tri-axial accelerometer sensor at multiple body points. In *2021 IEEE Industrial Electronics and Applications Conference (IEACon)*, pages 181–184. IEEE, 2021.

[66] Sampling frequency, signal resolution and the accuracy of wearable context recognition systems. In *Eighth International Symposium on Wearable Computers*, volume 1, pages 176–177. IEEE, 2004.

[67] Carina Figueira, Ricardo Matias, and Hugo Gamboa. Body location independent activity monitoring. 2016.

[68] Shaohua Wan, Lianyong Qi, Xiaolong Xu, Chao Tong, and Zonghua Gu. Deep learning models for real-time human activity recognition with smartphones. *Mobile Networks and Applications*, 25:743–755, 2020.

[69] Barbara Bruno, Fulvio Mastrogiovanni, and Antonio Sgorbissa. A public domain dataset for adl recognition using wrist-placed accelerometers. In *the 23rd IEEE International Symposium on Robot and Human Interactive Communication*, pages 738–743. IEEE, 2014.

[70] Attila Reiss and Didier Stricker. Introducing a new benchmarked dataset for activity monitoring. In *2012 16th international symposium on wearable computers*, pages 108–109. IEEE, 2012.

[71] Allan Stisen, Henrik Blunck, Sourav Bhattacharya, Thor Siiger Prentow, Mikkel Baun Kjærgaard, Anind Dey, Tobias Sonne, and Mads Møller Jensen. Smart devices are different: Assessing and mitigating mobile sensing heterogeneities for

activity recognition. In *Proceedings of the 13th ACM conference on embedded networked sensor systems*, pages 127–140, 2015.

[72] Oresti Banos, Rafael Garcia, Juan A Holgado-Terriza, Miguel Damas, Hector Pomares, Ignacio Rojas, Alejandro Saez, and Claudia Villalonga. mhealthdroid: a novel framework for agile development of mobile health applications. In *Ambient Assisted Living and Daily Activities: 6th International Work-Conference, IWAAL 2014, Belfast, UK, December 2-5, 2014. Proceedings 6*, pages 91–98. Springer, 2014.

[73] Oresti Banos, Claudia Villalonga, Rafael Garcia, Alejandro Saez, Miguel Damas, Juan A Holgado-Terriza, Sungyong Lee, Hector Pomares, and Ignacio Rojas. Design, implementation and validation of a novel open framework for agile development of mobile health applications. *Biomedical engineering online*, 14(2):1–20, 2015.

[74] Narit Hnoohom, Sakorn Mekruksavanich, and Anuchit Jitpattanakul. Human activity recognition using triaxial acceleration data from smartphone and ensemble learning. In *2017 13th international conference on signal-image technology & internet-based systems (SITIS)*, pages 408–412. IEEE, 2017.

[75] Md Abdullah Al Hafiz Khan, Nirmalya Roy, and Archan Misra. Scaling human activity recognition via deep learning-based domain adaptation. In *2018 IEEE international conference on pervasive computing and communications (PerCom)*, pages 1–9. IEEE, 2018.

[76] Ran Zhu, Zhuoling Xiao, Ying Li, Mingkun Yang, Yawen Tan, Liang Zhou, Shuisheng Lin, and Hongkai Wen. Efficient human activity recognition solving the confusing activities via deep ensemble learning. *Ieee Access*, 7:75490–75499, 2019.

[77] Nidhi Dua, Shiva Nand Singh, and Vijay Bhaskar Semwal. Multi-input cnn-gru based human activity recognition using wearable sensors. *Computing*, 103:1461–1478, 2021.

[78] Yuqing Chen and Yang Xue. A deep learning approach to human activity recognition based on single accelerometer. In *2015 IEEE international conference on systems, man, and cybernetics*, pages 1488–1492. IEEE, 2015.

[79] Kun Wang, Jun He, and Lei Zhang. Sequential weakly labeled multiactivity localization and recognition on wearable sensors using recurrent attention networks. *IEEE Transactions on Human-Machine Systems*, 51(4):355–364, 2021.

[80] Timo Sztyler, Heiner Stuckenschmidt, and Wolfgang Petrich. Position-aware activity recognition with wearable devices. *Pervasive and mobile computing*, 38:281–295, 2017.

[81] Frank Krüger, Albert Hein, Kristina Yordanova, and Thomas Kirste. Recognising user actions during cooking task (cooking task dataset) imu data. *University Library, University of Rostock*, 2017.

[82] Ganapati Bhat, Nicholas Tran, Holly Shill, and Umit Y Ogras. w-har: An activity recognition dataset and framework using low-power wearable devices. *Sensors*, 20 (18):5356, 2020.

[83] Gierad Laput, Robert Xiao, and Chris Harrison. Viband: High-fidelity bio-acoustic sensing using commodity smartwatch accelerometers. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pages 321–333, 2016.

[84] Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra, et al. A public domain dataset for human activity recognition using smartphones.

[85] Charikleia Chatzaki, Matthew Pediaditis, George Vavoulas, and Manolis Tsiknakis. Human daily activity and fall recognition using a smartphone's acceleration sensor. In *Information and Communication Technologies for Ageing Well and e-Health: Second International Conference, ICT4AWE 2016, Rome, Italy, April 21-22, 2016, Revised Selected Papers 2*, pages 100–118. Springer, 2017.

[86] Gary M Weiss, Kenichi Yoneda, and Thaier Hayajneh. Smartphone and smartwatch-based biometrics using activities of daily living. *IEEE Access*, 7: 133190–133202, 2019.

[87] Sophie Huhn, Miriam Axt, Hanns-Christian Gunga, Martina Anna Maggioni, Stephen Munga, David Obor, Ali Sié, Valentin Boudo, Aditi Bunker, Rainer Sauerborn, et al. The impact of wearable technologies in health research: scoping review. *JMIR mHealth and uHealth*, 10(1):e34384, 2022.

[88] Tan-Hsu Tan, Jyun-Yu Shih, Shing-Hong Liu, Mohammad Alkhaleefah, Yang-Lang Chang, and Munkhjargal Gochoo. Using a hybrid neural network and a regularized extreme learning machine for human activity recognition with smartphone and smartwatch. *Sensors*, 23(6):3354, 2023.

[89] Ariany F Cavalcante, Victor H de L Kunst, Thiago de M Chaves, Júlia DT de Souza, Isabela M Ribeiro, Jonysberg P Quintino, Fabio QB da Silva, André LM Santos, Veronica Teichrieb, and Alana Elza F da Gama. Deep learning in the recognition of activities of daily living using smartwatch data. *Sensors*, 23(17):7493, 2023.

[90] Daixi Chen, Yuchen Du, Yuan Liu, Jun Hong, Xiaojian Yin, Zhuoting Zhu, Jingjing Wang, Junyao Zhang, Jun Chen, Bo Zhang, et al. Development and validation of a smartwatch algorithm for differentiating physical activity intensity in health monitoring. *Scientific Reports*, 14(1):9530, 2024.

[91] Greg Milette and Adam Stroud. *Professional Android Sensor Programming.* Wrox Press Ltd., Birmingham, UK, UK, 1st edition, 2012. ISBN 1118183487.

[92] James A. Levine. Non-exercise activity thermogenesis (neat). *Best Practice and Research Clinical Endocrinology and Metabolism*, 16(4):679–702, 2002.

[93] Lingxiang Zheng and et al. A novel energy-efficient approach for human activity recognition. *Sensors*, 17(9), 2017.

[94] Ankita Dewan, Venkata M. V. Gunturi, Vinayak Naik, and Kousik Kumar Dutta. Neat activity detection using smartwatch at low sampling frequency. In *2021 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/IOP/SCI)*, pages 25–32, 2021. doi: 10.1109/SWC50871.2021.00014.

[95] Accelerometer's position free human activity recognition using a hierarchical recognition model. In *The 12th IEEE international conference on e-health networking, applications and services*, pages 296–301. IEEE, 2010.

[96] Yuhuang Zheng. Human activity recognition based on the hierarchical feature selection and classification framework. *Journal of Electrical and Computer Engineering*, 2015:34–34, 2015.

[97] Ankita Dewan, Venkata MV Gunturi, Vinayak Naik, Kartik Vishwakarma, and Shrehal Bohra. A hierarchical classifier for detecting metro-journey activities in data sampled at low frequency. In *Proceedings of the 17th International Conference on Advances in Mobile Computing & Multimedia*, pages 46–55, 2019.

[98] Ankita Dewan, Venkata M. V. Gunturi, and Vinayak Naik. Neat activity detection using smartwatch. *International Journal of Ad Hoc and Ubiquitous Computing*, 2023. doi: 10.1504/IJAHUC.2023.10057574.

[99] Amari Vaughn, Paul Biocco, Yang Liu, and Mohd Anwar. Activity detection and analysis using smartphone sensors. In *2018 IEEE International Conference on Information Reuse and Integration (IRI)*, pages 102–107. IEEE, 2018.

[100] Alvina Anjum and Muhammad U Ilyas. Activity recognition using smartphone sensors. In *2013 ieee 10th consumer communications and networking conference (ccnc)*, pages 914–919. IEEE, 2013.

[101] Xizhe Yin, Weiming Shen, Jagath Samarabandu, and Xianbin Wang. Human activity detection based on multiple smart phone sensors and machine learning algorithms. In *2015 IEEE 19th international conference on computer supported cooperative work in design (CSCWD)*, pages 582–587. IEEE, 2015.

[102] Dhaval Patel, Shrey Shrivastava, Wesley Gifford, Stuart Siegel, Jayant Kalagnanam, and Chandra Reddy. Smart-ml: A system for machine learning model exploration using pipeline graph. In *2020 IEEE International Conference on Big Data (Big Data)*, pages 1604–1613, 2020. doi: 10.1109/BigData50022.2020.9378082.

[103] Dhaval Patel, Syed Yousaf Shah, Nianjun Zhou, Shrey Shrivastava, Arun Iyengar, Anuradha Bhamidipaty, and Jayant Kalagnanam. Flops: On learning important

time series features for real-valued prediction. In *2020 IEEE International Conference on Big Data (Big Data)*, pages 1624–1633, 2020. doi: 10.1109/BigData50022.2020.9378499.

[104] Ankita Dewan, Venkata M. V. Gunturi, and Vinayak Naik. Deep learning models for neat activity detection on smartwatch. In *2023 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/IOP/SCI)*, 2023.

[105] Seyed Ali Rokni, Marjan Nourollahi, and Hassan Ghasemzadeh. Personalized human activity recognition using convolutional neural networks. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.

[106] Yiqiang Chen, Xin Qin, Jindong Wang, Chaohui Yu, and Wen Gao. Fedhealth: A federated transfer learning framework for wearable healthcare. *IEEE Intelligent Systems*, 35(4):83–93, 2020.