Particle Swarm Optimisation based Convolutional Neural Network for Human Activity Prediction

Preethi Gunishetty Devarakonda

A dissertation submitted in partial fulfilment of the requirements of
Dublin Institute of Technology for the degree of
M.Sc. in Computing (Data Analytics)

Sep 2020

Declaration

I certify that this dissertation which I now submit for examination for the award of MSc in Computing (Data Analytics), is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

This dissertation was prepared according to the regulations for postgraduate study of the Dublin Institute of Technology and has not been submitted in whole or part for an award in any other Institute or University.

The work reported on in this dissertation conforms to the principles and requirements of the Institute's guidelines for ethics in research.

Signed: Preethi Gunishetty Devarakonda

Date: 01 September 2020

Abstract

The increased usage of smartphones for daily activities has created a huge demand and opportunities in the field of ubiquitous computing to provide personalized services and support to the user. In this aspect, Sensor-Based Human Activity Recognition (HAR) has seen an immense growth in the last decade playing a major role in the field of pervasive computing by detecting the activity performed by the user. Thus, accurate prediction of user activity can be valuable input to several applications like health monitoring systems, wellness and fit tracking, emergency communication systems etc.,

Thus, the current research performs Human Activity Recognition using a Particle Swarm Optimization (PSO) based Convolutional Neural Network which converges faster and searches the best CNN architecture. Using PSO for the training process intends to optimize the results of the solution vectors on CNN which in turn improve the classification accuracy to reach the quality performance compared to the state-of-the-art designs. The study investigates the performances PSO-CNN algorithm and compared with that of classical machine leaning algorithms and deep learning algorithms. The experiment results showed that the PSO-CNN algorithm was able to achieve the performance almost equal to the state-of-the-art designs with a accuracy of 93.64%. Among machine learning algorithms, Support Vector machine found to be best classifier with accuracy of 95.05% and a Deep CNN model achieved 92.64% accuracy score.

Keywords: *Human Activity Recognition, Particle Swarm Optimisation, Convolutional Neural Network, Time Series Classification, Deep Learning, Sensors*

Acknowledgments

I would first like to express my sincere thanks to my supervisor Prof. Bojan Bozic for his continuous support, guidance and advice throughout the project. It's been a good experience and honor to work under his supervision.

I would like to thank Dr. Luca Longo, my thesis coordinator for helping me to formulate project proposal and this helped me in completing my thesis through various phases.

Not to forget, I would like to thank all my friends and family for all their encouragement, support and motivation during my studies. Special gratitude to my parents Sheshadri G.D. and Nirmala G.D. , and my sister Ankitha for their love, support and encouragement throughout my studies. This accomplishment would not have been possible without them.

Contents

List of Figures

List of Tables

List of Acronyms

HAR	Human Activity Recognition
ADL	Activities of Daily Living
CRISP-DM	Cross Industry Standard Process for Data Mining
DT	Decison Tree
ML	Machine Learning
LR	Multinomial Logistic Regression
SVM	Support Vector Machine
HMM	Hidden Markov Model
LSTM	Long Short Term Memory
CNN	Convolutional Neural Network
PSO	Particle Swarm Optimisation
PCA	Principal Component Analysis
FN	False Negavtive
FP	False Positive
TN	True Negative
TP	True Positive

Chapter 1

Introduction

1.1 Background

Activity Recognition aims at identifying the activity of users based on series of observations collected during the activity in a definite context environment. Applications that are enabled with activity recognition are gaining huge attention, as users get personalized services and support based on their contextual behaviour. The proliferation of wearable devices and smartphones has provided real-time monitoring of human activities through sensors that are embedded in smart devices such as proximity sensors, cameras, microphone, magnetometers accelerometers, gyroscopes, GPS etc., Thus, understanding human activities in inferring the gesture or position has created a competitive challenge in building personal health care systems, examining wellness and fit characteristics, and most pre-dominantly in elderly care, abnormal activity detection, diabetes or epilepsy disorders etc.,

Initially, Human Activity Recognition (HAR) experiment was carried out by attaching one or more dedicated on-body sensors to specific parts of human body to collect time series data (He, Zhang, Ren, & Sun, 2015; Ioffe & Szegedy, 2015) As, the usage of smart phones for daily activities has increased extensively, HAR research has employed to collect data from built-in sensors embedded in smart phones (Guo, Liu, & Chen, 2016; Subasi, Fllatah, Alzobidi, Brahimi, & Sarirete, 2019; bin Abdullah, Negara, Sayeed, Choi, & Muthu, 2012; Ehatisham-Ul-Haq, Azam, Amin, & Naeem, 2020). The raw data from the sensors are analysed using several machine learning and deep learning algorithms to classify the activity with appropriate evaluation metric. The activity recognition performance has significantly made strides since the of research, but the experiment set up can be varied, for example, the types of exercises performed by human subjects, the sorts of sensors utilized, the rate at which signal is sampled , the segment length of time series data. Apart from choosing classifier learning algorithms, the approaches are varied in terms of applying various feature processing techniques namely feature selection, extraction and transformation. These choices made comparative evaluation of different Human Activity Recognition (HAR) approaches complex. Thus, Human Activity Recognition (HAR) plays a significant part in enhancing people's lifestyle, as it should be competent enough in learning high level quality information from raw sensor data. Effective HAR applications are incorporated for contextual behaviour analysis (Aurangzeb et al., 2019), video surveillance analysis (Prati, Shan, & Wang, 2019; Hwang, Park, & Har, 2019), gait investigation (to determine any abnormalities in walking or running), gesture and position recognition (Kang, Zhang, & Dong, 2019).

1.2 Research Problem

Human Activity Recognition (HAR) is evolving to be a challenging time series classification task which involves predicting the human activity based on sensor data where the data points are recorded at regular intervals. Though HAR seems to be the straightforward approach of performing HAR, there are numerous issues and challenges that are encountered in selecting the appropriate feature processing technique and thus choosing the correct modelling algorithm for the time series data is crucial. Apart from this, there are also few resource constraints like limited battery power (due to continuous sampling on mobile devices), memory and storage capabilities.

Thus, the conventional approaches have made extraordinary progress on Human Activity Recognition (HAR) by incorporating machine learning algorithms such as Naïve Bayes, Decision Tree, Support Vector Machine, Logistic Regression as there are only few labelled data. It requires domain knowledge to manually process the feature extraction. On the other hand, deep learning algorithms has seen high performance in areas like Natural language Processing, Object Recognition etc., In spite of these advancements, another line of research has emerged in applying nature-inspired meta heuristic optimization techniques like Particle Swarm Optimization, Genetic Algorithms on Neural Networks. The research question that is aimed to be addressed in the current study can be concisely stated as below –

"To what extent can the Particle Swarm Optimized Convolutional Neural Network significantly enhance the recognition of human activity from raw inertial sensor data when compared with supervised machine learning algorithms and Deep Learning Algorithms"

Algorithms: Naive Bayes, Support Vector Machine (SVM),Random Forest (RF) Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN), PSO Optimized Convolutional Neural Network

1.3 Research Objectives

The most feasible solution in overcoming these challenges could be by looking into existing works and analysing the experimental set up. Thus, picking the right sensor and right gestures with demonstrated capabilities can significantly eliminate the chances of inaccurate sensor data. The traditional machine learning algorithms require large amount of labelled static data and manually performing the feature selection tasks. But in real applications most of the activity data are unlabelled and entire data needs to be analysed. Since Deep Learning methods can perform training on the entire data, and analysing the complex features, this study focuses on investigating the performance various deep learning models in classifying the time series data.

Convolutional Neural Network requires large number of parameters to tune and it is time consuming. The study explores the optimization using metaheuristic algorithm Particle Swarm Optimization for Convolutional Neural Network. Thus, the study has a deep investigation towards major approaches followed in HAR namely, machine learning using hand-crafted features and deep learning using raw inertial signals. The process in which the research carried to achieve the results is mentioned below.

1. Exploring the previous works on Human Activity Recognition, identifying gaps in the research with a detailed analysis.
2. Data Preparation and Data Pre-Processing is conducted by considering two versions of the dataset.
3. Designing a solution to perform Human Activity Recognition by using Particle Swarm Optimized Convolutional Neural Network.
4. Implementing the solution fortified in the proposed design and tune the models to to obtain the expected accuracy.
5. Evaluate the performance of the various models.
6. Comparing the results obtained for different models and place the findings in the study.

1.4 Research Methodologies

The study aims at enhancing the prediction accuracy of human activity by applying Particle Swarm Optimisation for training Convolutional Neural Network on the existing dataset, and hence secondary research is employed. Considering the existing research on human activity recognition, a detailed literature review is conducted on machine learning algorithms, deep learning algorithms and Optimization techniques to get the deep understanding of the project undertaken.

Additionally, the research follows the quantitative methodology and it is empirical in nature. As the experiment involves identifying the best classification algorithm, where the performance is measured through best accuracy and F1 score of various models built. The research form is empirical as it involves finding the solutions to the research problem using systemic modelling experiments.

The study is performed based on reviewing the existing works, where a hypothesis is designed that can be accepted/ rejected through the experiment. Based on the results obtained from the hypothesis tests, the solution can be generalized to address a specific context. Thus, the current study is inductive in nature.

1.5 Scope and Limitations

The current study explores two versions of the same dataset that is with hand-crafted features and with raw inertial signals. However, several research in HAR studies in HAR are made by considering multiple datasets in order to evaluate the robustness of the classification model (Mannini, Rosenberger, Sabatini, & Intille, 2017).

From the reviews of the existing works, Data Dimensionality Reduction is performed on the hand-crafted features to get the optimal subset of features to enhance the accuracy of machine learning model. In the current experiment, Data Dimensionality Reduction is not performed to before modelling the data with classical machine modelling algorithms. Performing this would have contributed to the performance of the models.

Parameter tuning in the models and can sometimes enhance the accuracy of the state-of-the-art models. to Since the focus of the experiment is to investigate the performance of the multiple models , the study had to limit the number of parameters to manipulate for algorithms. Due to computing power and time constraints , limited parameter tuning is performed.

1.6 Document Outline

- Chapter 2 provides an overview on Human Activity Recognition and its applications. Various approaches for HAR task are discussed. Particularly, Sensor based HAR is detailed with different sensor modalities. This chapter also gives an overview of the modelling approaches for HAR . Each modelling approach is discussed with its theory and its applicability in HAR.

- Chapter 3 explains the design of the current study. The structure of the chapter is aligned with the CRISP-DM methodology, which outlines data understanding, data preparation and modelling algorithms chosen. The chapter concludes with the evaluation metrics and model optimization.

- Chapter 4 explains the practical implementation of the experiment. This explains the parameters settings and topology of models implemented. It represents the results obtained in the form of chosen evaluation metric

- Chapter 5 performs the evaluation of the experimental results obtained. Performance comparison of various models implemented are presented. Bases on result obtained the research hypothesis are evaluated. The chapter concludes by discussing the strength and limitations of the experiment performed.

- Chapter 6 summarizes the overall workflow of the process and findings of the research project undertaken. Additionally, possible areas of future work is highlighted.

Chapter 2

Review of existing literature

This section provides an overview on Human Activity Recognition and its applications. Various approaches for HAR task are discussed. Particularly, Sensor based HAR is detailed with different sensor modalities. This chapter also gives an overview of the modelling approaches for HAR . Each modelling approach is discussed with its theory and its applicability in HAR.

2.1 Human Activity Recognition (HAR)

Due to the advancement in ubiquitous computing, Activity Recognition has been one of the major research areas in mobile technology that has seen the rapid demand over the past few years (Al-Obeidat, Belacel, & Spencer, 2019). This covers major areas like smart homes (Q. Li, Ning, Zhu, Cui, & Chen, 2019), wellness and fitness monitoring (Nawaratne, Alahakoon, De Silva, Kumara, & Yu, 2019), video surveillance analysis for security purposes (Voicu, Dobre, Bajenaru, & Ciobanu, 2019), behavioural analysis (Bhat, Deb, & Ogras, 2019; Patel & Shah, 2019; Longstaff, Reddy, & Estrin, 2010), emergency services (Yin, Yang, & Pan, 2008) etc., Considering the security and privacy of user data, various steps have been enforced while collecting data from the smartphones for user authentication. Additionally, advances in the field of healthcare and medicine have significantly improved the quality of life which in turn has seen significant rise in life expectancy. Increasing cost for health treatment, especially for elderly have created demand in implementing cost-reducing actions from various health-care institutes. Apart from improving infrastructure facilities, technical improvements have contributed for efficient heath care systems.

Accordingly, Human Activity Recognition (HAR) has evolved as a trending technology with the ability to benefit elderly people and disabled (Kwapisz, Weiss, & Moore, 2011). The basic principle of Human Activity Recognition is to classify the body gesture, position, motion and then predict the different states of action or behaviour (Chen & Xue, 2015; Abowd, Dey, Orr, & Brotherton, 1998).

2.2 Video Based Human Activity Recognition

Video-based Human Activity Recognition collects data from Camera and analyses videos or images containing human movement. The task of Video-Based Human activity recognition can be achieved by following approaches like; space-time approaches in computer vision and sequential approaches using time series analysis (Cook, Feuz, & Krishnan, 2013).

In vision-based approach, human activity is recognized using similarity and difference among the volume of images or video sequences. (Derpanis, Sizintsev, Cannons, & Wildes, 2012) proposed a method where the series of images are observed for changes in shape, that corresponds to a motion of user. Similar approach was followed by (Bobick & Davis, 2001) by comparing the shape of images in patches of fixes size. By extracting important features from the image sequence. citerodriguez2010spatio has seen best results in recognizing the activity. With the advancement in the use of deep learning for real time activity recognition (L. Wang, Qiao, & Tang, 2015; Donahue et al., 2015; Venugopalan et al., 2014) applied neural networks based on learnings from a semantic trajectory data collected from raw video.

In sequential approaches which followed a time series analysis initially used, statistical techniques for HAR tasks. (Yacoob & Black, 1999) used Principle Component Analysis (PCA) with Singular Value Decompositions (SVD) to process one signal at one time which assumes all the variability are the linear combinations with weighted statistical features. In this aspect, Hidden Markov models (HMM) is employed for HAR and achieved satisfactory results. (Oliver, Rosario, & Pentland, 2000) in his work proposed new Coupled-Hidden Markov models which can recognize complex human activities.

2.3 Sensor Based Human Activity Recognition

Due to the immense growth of sensor technology and ubiquitous computing, sensor-based Human Activity Recognition is gaining attention which is widely used with enhanced protection and privacy. According to (Chavarriaga et al., 2013), the HAR task can be achieved by placing the sensors at different locations to recognize human activity for specific context Table 2.1 . based on sensor placements at different locations. HAR with different sensor modalities are listed below –

Modality	**Description**	**Sensor Types**
Wearable	Usually Worn by the user to capture the body movements	Smartphone, watches gyroscope ,accelerometer
Object Sensors	Mounted on objects to capture objects movements	RFID, accelerometer on objects
Ambient Sensors	Mounted in environment surroundings to record user interaction	Bluetooth,Sound,WiFi,
Hybrid Sensors	Combination of multiple sensors	Multiple types, often deployed in smart environments

Table 2.1: Sensor modalities for Human Activity Recognition tasks
(J. Wang, Chen, Hao, Peng, & Hu, 2019)

2.3.1 Object Sensors

Object sensors are generally placed on various objects or items to detect the movement of a specific object (Okeyo, Chen, & Wang, 2014). These work differently from wearable sensors which capture human motion details. The object sensors are focuses to detect the the change in movement of specific objects so that to decide human activities. For example, in-order to identify drinking activity of a human the accelerometer is fixed to a glass.

The adoption of Activity Recognition in smart homes to provide has seen a significant improvement in making intelligent decisions and provide better user experience (Hong & Ohtsuki, 2011; Sarkar, Lee, Lee, et al., 2011; Yang, Lee, & Choi, 2011; Tolstikov et al., 2011; Van Kasteren, Englebienne, & Kröse, 2010).These systems are capable of recognising complex activities like washing dishes, eating, taking a shower etc., as the decisions depend upon the data collected from object sensors placed in other locations at home.

As in case of smart home features a Radio Frequency Identifier (RFID) which is treated as object sensors. (Fang & Hu, 2014; Vepakomma, De, Das, & Bhansali, 2015; Yang, Nguyen, San, Li, & Krishnaswamy, 2015). and medical monitoring systems (A. Wang, Chen, Shang, Zhang, & Liu, 2016; X. Li, Zhang, Marsic, Sarcevic, & Burd, 2016).The RFID can provide more detailed information for recognizing complex activities. However, these are less in use when compared to wearable sensors as they are difficult to deploy.

2.3.2 Ambient Sensors

The interactions between the human movements and other surroundings are being recorded by Ambient sensors. Some of the Ambient sensors are sound sensors (to record changes in noise signals), pressure sensors(to record the humidity of the surrounding), and temperature sensors are used to record the changes in various weather conditions. Many studies say that ambient sensors can be used to detect human activities like hand gestures and limb movements (Kim & Toomajian, 2016; Lane, Georgiev,

& Qendro, 2015). In addition, ambient sensors can be damaged easily by the environment, human activities only of certain for certain application can be inferred correctly (Y. Wang, Cang, & Yu, 2019).

Furthermore, Cameras has also served the purpose of HAR where the activities are classified based on the movements, gestures and positions from video sequences. This has seen a remarkable benefits in intrusion detection applications (Turaga, Chellappa, Subrahmanian, & Udrea, 2008; Candamo, Shreve, Goldgof, Sapper, & Kasturi, 2009; Ahad, Tan, Kim, & Ishikawa, 2008).

2.3.3 Hybrid Sensors

Different Combination of sensor data are used collectively to recognize human activity. As explained in (Hayashi, Nishida, Kitaoka, & Takeda, 2015) the combination acceleration sensor with acoustic information data could significantly improve the accuracy of HAR task which is a combination of ambient sensor and object sensor. As a result of which both environment changes and object movements can be captured.

(Hammerla, Halloran, & Plötz, 2016; Hannink et al., 2016; A. Wang et al., 2016) built a smart home environment named A-Wristocracy , which utilized data from combination of multiple sensors . Large number of simple and complex activities were able to detect using ambient, object and wearable sensors.

2.3.4 Body-worn sensors / wearable sensors

Wearable sensors are one of the widely used sensor modalities in HAR. These sensors are often worn or attached to the users, namely an accelerometer, gyroscope, and magnetometer. As the human body moves, the acceleration and angular velocity are varied, this data is further analysed to predict the activity. These sensors can be embedded in smart phones, smart watches, fit bands, headbands etc., The figure (2.1) shows the different wearable sensors that can be used by humans (Piwek, Ellis, Andrews, & Joinson, 2016) (Cleland et al., 2013) studies the significance of sensor and its appropriate position to be placed on the body of the user. Many research are conducted to investigate the variability in accuracies by placing the sensors on different pasts of human body. One such study is performed by.

(Parkka et al., 2006) by placing sensors on chest and wrist for duration of two hours which gave 83% classification accuracy.

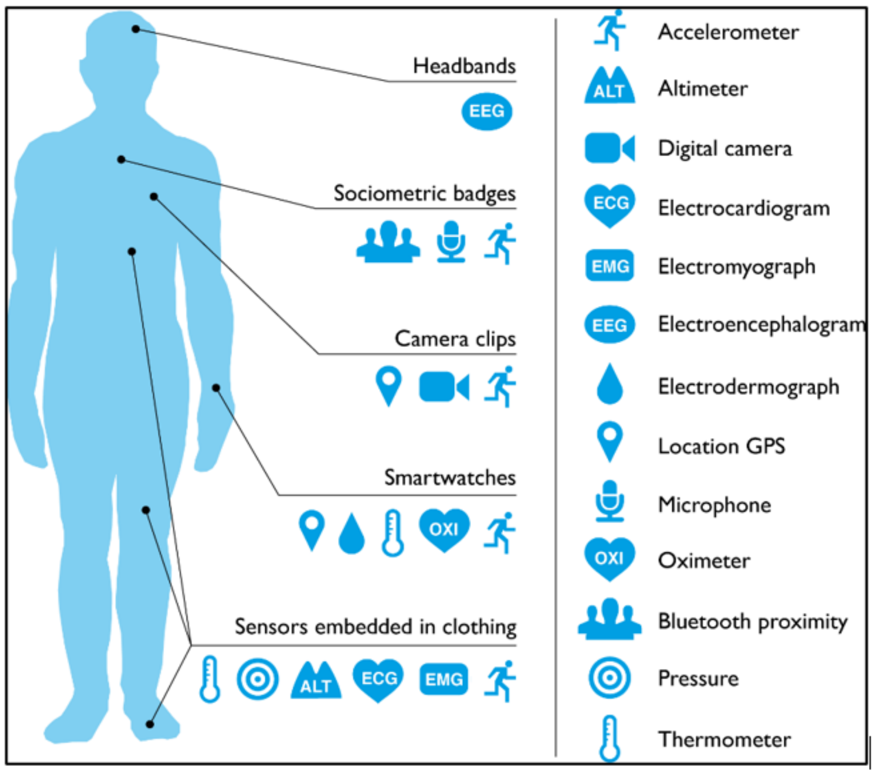

Figure 2.1: Body Worn sensors
(Piwek et al., 2016)

Thus, wearable sensors were widely used for HAR (Jiang & Yin, 2015; Chen & Xue, 2015; Plötz, Hammerla, & Olivier, 2011a; Yang et al., 2015; Zeng et al., 2014) in various health monitoring systems. In recent days, inertial sensing, that uses movement-based Sensors which can be attached on user's body has been studied widely (Yu, Cang, & Wang, 2016). Among those work, the accelerometer is mostly used sensor for collecting position details. Gyroscope and magnetometer are also used in combination with accelerometer.

2.4 Modeling Approaches for Human Activity Recognition

In any data mining project, the choice of the appropriate modeling algorithm does not depends only on the type of problem to solve, but also on the type of input data. Due to the natural ordering of the temporal feature data, the Human Activity Recognition is considered as a typical pattern-recognition system where it involves classifying the human activity based on the series of data. The main difference between Machine Learning Algorithms and Deep Learning Algorithms in recognising human activity is the way the input features are extracted. In this aspect, the below sections explain the methodology chosen for the task of Human Activity Recognition.

2.5 Methodology based on Machine Learning Algorithms

2.5.1 Theory

Machine Learning is a is one of the application of Artificial Intelligence branch, which enhances the capability of computers to learn from past experiences in the data without any human intervention. The algorithms are evolved from fields like statistical analysis, computing science etc., There are several Machine algorithms designed to not only to address the problem but also efficient in terms of computing power and storage efficiency. There are mainly four categories of Machine Leaning Algorithms namely:

- Supervised Learning: This requires a labelled input data to learn and train from the data. Supervised learning is used when there are set of input variables (x) and an output variable (Y) and modeling algorithm is employed to learn the relationship between input and output. The task is to find the approximate this mapping function, such that the model can predict the output for a new input.
- UnSupervised Learning: Unsupervised learning is used when there is (X) and no

corresponding output variable. The aim for unsupervised learning is to find the underlying pattern or distribution in the input data.

- Semi-Supervised Learning: This is used when there is large amount of input data (X) and among them, only few of the data is labelled (Y).

- Reinforcement Learning: method focus at using data collected from the interaction with environment and then actions are taken that would minimize or maximize the error. The learning continues until the algorithm explores the full range of possible values.

2.5.2 Machine Learning Algorithms for Human Activity Recognition

Considering HAR as one of the pattern recognition problem, the conventional pattern recognition methods have seen extraordinary results by utilizing machine learning algorithms like hidden Markov models, decision tree support vector machine, naive Bayes (Lara & Labrador, 2012). The Figure 2.2 illustrates the process of Human Activity Recognition using hand-crafted features modelled with machine learning algorithms. The raw inertial activity signals received from the sensors are subjected to feature -extraction process by domain knowledge experts (Bengio, 2013). The features that are usually extracted are based on two main domain features namely; time domain and frequency domain (Figo, Diniz, Ferreira, & Cardoso, 2010).

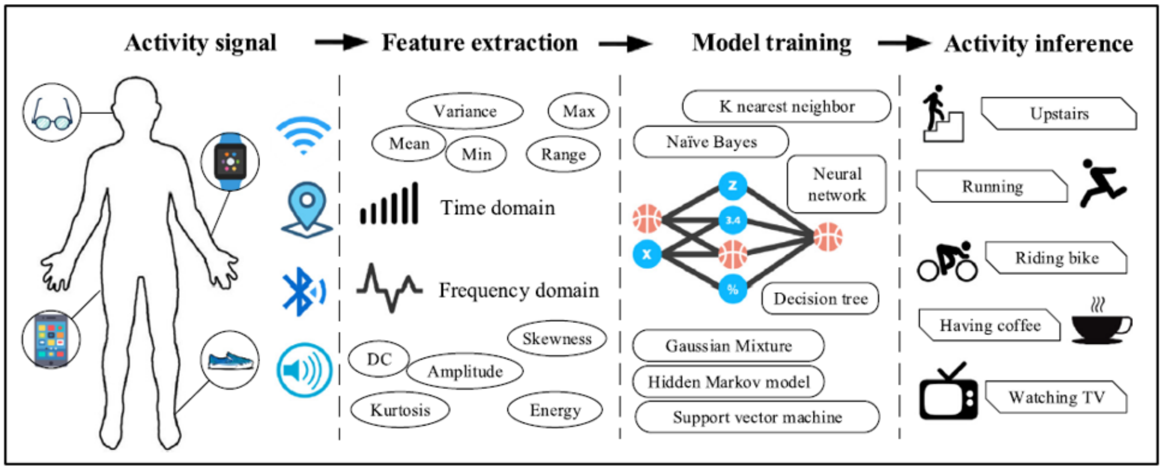

Figure 2.2: Process of Human Activity Recognition using hand-crafted features modelled with machine learning algorithms

The time domain features are computed based on mathematical functions to extract statistical details from the signals. The frequency domain features possess mathematical functions that record recursive patterns of signals. Thus, in machine leaning approach for HAR, the input data are always extracted from human engineered hand-crafted features. These features may be further pre-proceed using Data Dimensionality Reduction techniques to select the significant features.

(Khusainov, Azzi, Achumba, & Bersch, 2013) explains selecting important features is more significant than choosing a classification algorithm. This is because poor quality features may hinder the performance of the classifier. (Hassan, Uddin, Mohamed, & Almogren, 2018) in his recent work, employed Kernel Principal Component Analysis (KPCA) which works based on statistical analysis before applying modeling. Furthermore, (Khan, Siddiqi, & Lee, 2013) employed Stepwise Linear Discriminant Analysis (SWLDA) which is a non-linear method, selects the subset of features by using regression combined with F-test. The model showed enhanced performance after applying Data Dimensionality Technique.

Different modelling algorithms have been employed to predict the human activity recognition. (N. Ravi, Dandekar, Mysore, & Littman, 2005) in his work used Naïve Bayes classifier with few parameter settings to classify 8 different activities, which outperformed other classification algorithms. Several research employed Naive Byes as the primary classifier for human activity recognition (Yang, 2009; Kose, Incel, & Ersoy, 2012; Lu, Pan, Lane, Choudhury, & Campbell, 2009)

In recent times, learning algorithms which are based on error computation namely; Artificial Neural Networks (Kwapisz et al., 2011; Tang, Teng, Zhang, Min, & He, 2020; Irvine, Nugent, Zhang, Wang, & Ng, 2020) Support Vector Machine (Shaafi, Salem, & Mehaoua, 2020; Aslan, Durdu, & Sabanci, 2020; ?, ?) new are used for predicting HAR without any Data pre-processing technique applied.

The most used modelling algorithms that showed efficient results as per the study are Naive Bayes, Multinormal Logistic Regression, K - Nearest Neighbour Hidden Markov Models, Support Vector Machine and Artificial Neural Network.

2.6 Methodology based on Deep Learning Algorithms

2.6.1 Theory

Deep Learning uses multiple layers of neural network to extract higher level information or features from the raw input data. Deep Learning algorithms have made a tremendous progress in areas like image analysis, speech recognition, text analysis, self-driving cards, time series forecasting fraud detection etc., (Mannini et al., 2017). They are extensively used in applications where the input data are huge and requires high computational resources. (W. Liu et al., 2017).

(Dong & Wang, 2016) explains the process of machine learning, which begins with selecting relevant features from the set of raw inputs. These subset of features are then used for modeling. On the other hand, relevant features are extracted automatically from the input data in Deep Learning. This helps in eliminating the need of doamin expertise and human intervention for perfroming feature extraction. Different types Deep Learning models are listed below -

- <u>Autoencoder:</u> An autoencoder is type of artificial neural network that is capable of learning representations using various coding procedures. Multilayer perceptron is one of the simple form of encoder which contains input layer, one or more number of hidden layers and an output layer. The number of nodes at the output layer makes the difference between multilayer perceptron autoencoder. In autoencoder, the number of nodes at input and output layers are same. This follows a unsupervised learning methodology.

- Restricted Boltzmann Machines: RBM are stochastic Deep Learning models which consists of two types of nodes namely hidden and visible nodes . Every two consecutive layers of RBM is treated as Deep Belief Network (DBN). DBN or RBM are mainly connected by fully-connected layers.

- Recurrent Neural Network (RNN) : Recurrent Neural Network (RNN) is commonly used in speech recognition and text processing which works by making use of the temporal relationship between the neurons. With the help of RNN's internal memory state, they possess capability to process variable length of series of data.

- Convolutional Neural Network (CNN) : CNNs are capable enough to extract features from raw input data(1-Dimensional, 2-Dimensional, 3-Dimensional) and it showed promising results in image classification, sentiment analysis and speech recognition. The two advantages of CNNs are , they employ Parameter sharing and local connectivity (Zhang et al., 2019). Parameter sharing is sharing of weights by all of the neurons in a specific feature map. Local connectivity is the concept where each neuron is connected only to a nearby subset of the input which are correlated. This helps greatly to lessen the number of parameters in the entire architecture system , thus makes the computation faster and efficient.

2.6.2 Deep Learning Approaches for Human Activity Recognition

Though, conventional Pattern Recognitions (PR) approaches gained satisfactory results in HAR, these methods heavily rely on hand crafted feature generation usually done by domain expertise (Arel, Rose, & Karnowski, 2010). This sometimes leads to error in collection data and missing some significant data points. On the other hand, Deep Neural Networks are capable of automatic feature extraction without human intervention. In fact, the model becomes more robust when data is large (Najafabadi et al., 2016).

The Figure 2.3 illustrates the process of HAR followed by Deep Learning Algorithms. Initially, the raw sensor signals collected from inertial sensors (accelerometer, gyroscope etc,.) are it is directly subjecting to modelling, where no feature extraction step is performed. Additionally, deep learning follows a unsupervised, incremental learning which makes it more feasible to implement HAR tasks. (Plötz, Hammerla, & Olivier, 2011b)

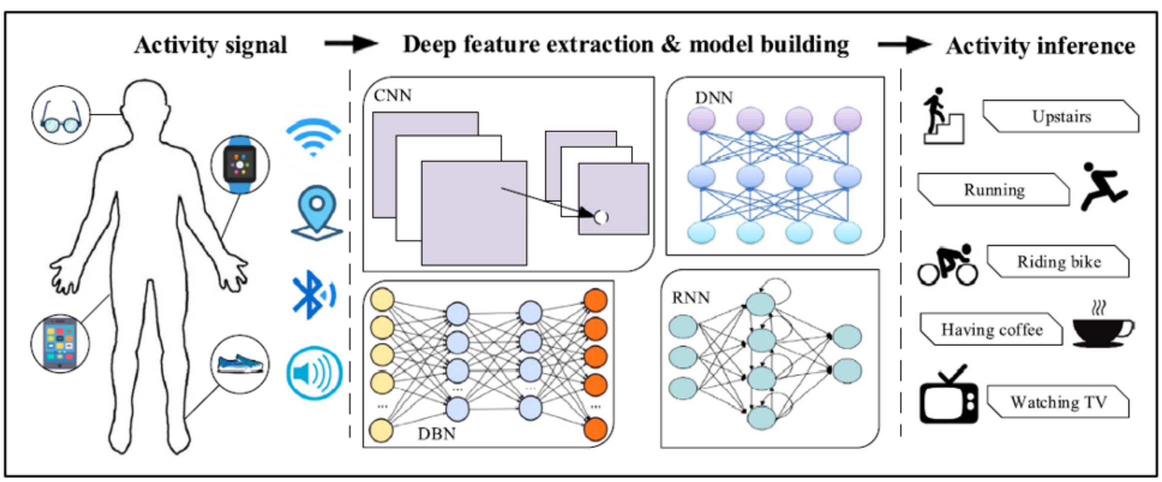

Figure 2.3: Process of Human Activity Recognition using raw inertial signals modelled with deep learning algorithms

Several Deep learning mode were employed to perform Activity Recognition in various contexts. (Z. Liu, Wu, Zhu, & Zhang, 2016) investigated the performance of

Restricted Boltzmann Machines for Activity Recognition from data collected through smart watches. The method outperformed other models and gained high accuracy results with less computation time.

Additionally, Long short-term memory(LSTM) models has been utilized to predict the activity performed for unbalanced real world data where the model performance wss evaluated using f1 score due to imbalance nature of data. (Guan & Plötz, 2017). (Vepakomma et al., 2015) the hand-crafted features are obtained from inertial sensors, and these features are added into DNN algorithm. In this aspect, (Walse, Dharaskar, & Thakare, 2016) used PCA as a Dimensionality Reduction Technique before modelling to Deep Neural Network (DNN). However, since domain knowledge is used for feature extraction, the model cannot be generalized.

Some works used Recurrent Neural Network (RNN) for the HAR (Edel & Köppe, 2016; Guan & Plötz, 2017; Hammerla et al., 2016; Inoue, Inoue, & Nishida, 2018), where the learning rate and computational power are the main constraints . More time is invested in finding the optimal set of hyper parameters that provides the best results. (Inoue et al., 2018) identified various model parameters and recommended a model that would achieve high accuracy of HAR by turning the hyperparameters. (Edel & Köppe, 2016) investigated the performance of binarized–BLSTM–RNN model in which weight, input and output parameter values of hidden layers are treated as binary. The main constraint of RNN based Human Activity Recognition models is to deal with the time, power constraint environment, while still thriving to achieve good performance results.

Furthermore, CNN's are used more extensively for HAR tasks with varied experimental settings. In general, CNNs are mostly used for image classification using 2-Dimensional Convolution since it accepts the data with shape n * n. Several works resized the single dimension input data to a 2D image so as to make use of 2D convolution. (Ha, Yun, & Choi, 2015) used similar approach in reshaping the input data to a 2D image. While (Jiang & Yin, 2015) designed a complex design of CNN algorithm by transforming time series data into an image. Other works include (Singh, Pondenkandath, Zhou, Lukowicz, & Liwickit, 2017; X. Li et al., 2016; D. Ravi, Wong, Lo,

& Yang, 2016) performed data transformation to achieve CNN model driven results.

2.7 Why Particle Swarm Optimization ?

As seen from the above section, deep learning networks have gained better results with less efforts in parameter settings. In particular, Deep Convolotional Neural Networks are used extensively due to its flexibility in both data driven approach (Using 1D Convolution for signal data) and model driven approach (data transformation of signal data to a 2D image). In order to gain higher performance of the model , several layers has to be used and parameter initialization has to be done carefully. This needs a detailed knowledge on CNN architecture and also on the dataset.

Thus, to find the optimal CNN architecture automatically without human intervention, a meta heuristic algorithm Particle Swarm Optimisation is utilized which is easy to implement with lower computational cost.

Theory

Particle Swarm Optimization (PSO) is a nature inspired, meta-heuristic algorithm often used for discrete, continuous and sometimes for x'combination optimization problems. The PSO was first introduced by Kennedy and Eberhart in 2001 (Kennedy, 2006) which is inspired by the pattern followed by a flock of words during flying. PSO works by making only few or no assumptions regarding the problem being optimized and possess the ability to search large spaces of candidate solutions in a efficient manner.

In PSO, a particle is called a single solution and the total of all such solutions is termed as swarm. The main ideology behind PSO is that each particle is well known of its velocity and the best configuration achieved in the past (pBest), and the particle which is the current global best configuration in the swarm of particles(gBest). Hence, at every current iteration, each particle updates its velocity in such a way that its new position will be close enough to global gBest and its own pBest at the same time. The velocity and particle vector are adjusted based to the following equations 2.1 and 2.2 respectively:

$$v_{id}(t+1) = w * v_{id}(t) + c_1 * r_1 * (P_{id} - x_{id}(t)) + c_2 * r_2 * (P_{gd} - x_{id}(t) \tag{2.1}$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \tag{2.2}$$

where $\mathrm{v_{id}}$ indicates the velocity of $\mathrm{i_{th}}$ particle in the $\mathrm{d_{th}}$ dimension, $\mathrm{z_{id}}$ indicates the position of $\mathrm{i_{th}}$ particle in the $\mathrm{d_{th}}$ dimension, $\mathrm{P_{id}}$ and $\mathrm{P_{gd}}$ represents the local best and the global best in the $\mathrm{d_{th}}$ dimension, $\mathrm{r_1}$ and $\mathrm{r_2}$ are the random numbers between the range 0 and 1, $\mathrm{c_1}$, $\mathrm{c_2}$ and w, are acceleration coefficient for exploitation ,acceleration coefficient for exploration and inertia weight respectively. Since the encoded vector in the proposed method is fixed-length and consists of decimal values, and PSO is effective to search for the optimal solution in a fixed-length search space of decimal values, the proposed method will use PSO as the search algorithm. One of the advantages of PSO is that they converge at a faster rate than Genetic Algorithms(GA) (Sahu, Panigrahi, & Pattnaik, 2012; Hu & Yen, 2013)

2.7.1 Gaps in the Research

The literature review outlines the existing works on Human Activity Recognition in terms of the modelling approaches chosen. However, certain gaps are found in both the approaches.

Some research works employed Machine Learning Approach to perform HAR with hand-crafted features faced low performance as only shallow features are explored and learned by the classifiers (Yang et al., 2015; Rashidi, Cook, Holder, & Schmitter-Edgecombe, 2010).Before deep learning was uses extensively, shallow neural network classifiers, that is Multi-Layer Perceptron (MLP), was considered to be a promising algorithm for HAR. In this aspect, (Godino-Llorente & Gomez-Vilda, 2004) performed HAR with algorithms like logistic regression, decision tree and MLP and MLP outperformed the other two models.

As deep convolutional neural networks (CNNs) has been used to obtain the excellent results in most of the image classification benchmarks datasets, they have overcome the need of human experts for classification (Chen & Atwood, 2007; Shin et al., 2010) But still, it remains a challenging task to find the meaningful CNN architecture that would apply for all type of domains. As a result of some of the successful CNN's architecture like ResNet

(Mihanpour, Rashti, & Alavi, 2020), VGG16 (?, ?), DenseNet (Iandola et al., 2014) and (?, ?) were introduced recently considering domain knowledge. The results from this outperformed the state-of-the-art baseline CNN model. However, the CNN's architecture are designed by doing lot of trial and error methods and are suitable to handle problems only in specific context.

PSO algorithm was employed to train an Artificial Neural Network by (Gudise & Venayagamoorthy, 2003). The results show that ANN's training time was reduced with PSO greatly. In this aspect, (Carvalho & Ludermir, 2006, 2007; Ojha, Abraham, & Snášel, 2017) designed PSO algorithms for two tasks that is to train ANN and to find better architecture. This resulted in achieving competitive result than other models.

Thus, most of the works in PSO was used to find optimal architectures in full connected networks (Dehuri, Roy, Cho, & Ghosh, 2012; Cao, Zhao, & Zaïane, 2013), but these cannot be used for tasks like image classification , activity recognition which indeed used a complex deep layers. Human Activity Recognition. Recently, (Junior & Yen, 2019) came up with PSO trained for CNN architecture that is suitable for only image classification with 2-Dimensional Convolutions. The experiment was performed with 10 benchmark datasets, and the results were outstanding.

However, there is not much work done on using 1-Dimensional Convolution for finding optimal architectures in CNN. Considering the gaps in the literature review , the current research aims to address few issues and find solutions that generalizes the models for Human Activity Recognition tasks.

2.7.2 Research Question

Thus , considering the gaps in the above mentioned literature review, this research aims to address the below research question.

"To what extent can the Particle Swarm Optimized Convolutional Neural Network significantly enhance the recognition of human activity from raw inertial sensor data when compared with supervised machine learning algorithms and Deep Learning Algorithms"

Chapter 3

Experiment design and methodology

3.1 Design Flow

This chapter gives the plan and the research methodology used for performing the research. To ensure quality outcome in a data mining project, a standard methodology must be followed. There are many data mining methodologies evolved and revised to rectify a robust and well-organized methodology for performing the data mining project. Among them, the Cross Industry Standard Process for Data Mining (CRISP–DM), a well proven methodology with a structured approach (Piatetsky, 2014) is employed to conduct the current study. This methodology possess flexibility, practicality and is idealised with a sequence of events. The current experiment addresses each of the phases in a separate section.

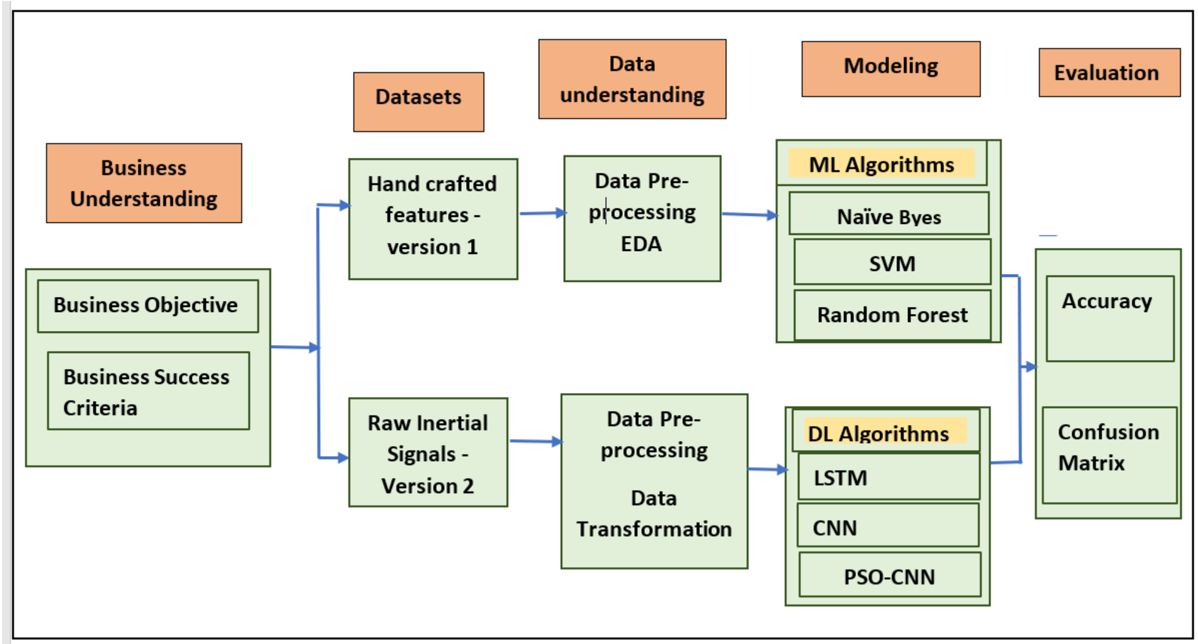

Figure 3.1: Design Flow

The figure shows the design flow to be followed for the current research. The experiment begins with the Business Understanding phase, which indicates what is to be accomplished from a business perspective. The expected outputs of this phase form the main objectives of the project. Here the insights and goals of the project are defined. In order to answer the research question, the experiment is conducted with two versions of the datasets which is explained in Data Understanding Phase. Additionally, data description report is prepared to understand each filed description. This is done separately for both the datasets.

The third phase is the Data preparation stage. Here the data is checked for duplicates, null records and appropriate action is taken to address them. Further, new derived fields can be formed based on the domain knowledge. Data from multiple databases are integrated to form the final dataset for modelling. The fourth step is the modeling stage.

Based on initial analysis done from the literature review, suitable modelling technique is chosen and applied on the two versions of the dataset. Next phase is Evaluation phase. Based on the evaluation criteria, models are evaluated to see if it meets the business objective.

3.2 Business Understanding

Human activity recognition has a great business value which is discussed in the previous chapter. The primary aim of the project is to identify a best classification algorithm which identifies the different human activities in motion accurately. Thus, the predicted activity can be applied to multiple applications like health monitoring and controlling systems, wellness and fit tracking, alarming to emergency situations etc., This provides great help to track the activities of elderly people, infants, patients with disabilities who require immediate medical response in case of unexpected accidents. It is achieved either by wearable sensor data or data collected from the smart phones. Hence accurate classification of activities is crucial, as it is related to providing emergency support. Accordingly, the study investigates the performance of multiple algorithms in three directions namely supervised machine learning, deep learning and optimisation algorithms as discussed in Chapter 2.

The laptop used for the execution process is is 64-bit i7-7500 Intel® Core™ with CPU speed @ 2.70 GHz processor and RAM of 8GB. The programming is done using Python Language of version 3.6 in Anaconda tool. The experiment could encounter the inability of executing certain algorithms that require robust computer with high computing power.

- Business Objective: As the experiment deals with providing quick and timely assistance to users, the algorithms must be robust enough to classify activities with high accuracy. Any undetected activity or error in number of false positive and false negatives leads to erroneous caution and panicking alerts leading to loss of time, manpower and other resources. Analysing the significance and in order to find the robust algorithm with high classification accuracy, the objective of the study is to investigate the performance of multiple algorithms in three directions namely supervised machine learning, deep learning and optimisation algorithms. And pick the best algorithm that predict the Human Activity .

- Business Success Criteria: The solution for the research problem must not only find the models which performs in classifying target data, but also ensure to show the confirmation that the results obtained are significant and is consistent when tried to repeat the solution. By considering the above business objectives, evaluation criteria and constraints, below hypothesis is formed to answer the research question –

Hypothesis 1

- *Null Hypothesis(H0):*If Particle Swarm Optimised Convolutional Neural Network is used instead of supervised machine learning then there is no significant improvement in classification of Human Activity in terms of accuracy and the F1 score.
- *Alternate Hypothesis(HA):*If Particle Swarm Optimised Convolutional Neural Network is used instead of supervised machine learning , then there is significant improvement in classification of Human Activity in terms of accuracy and the F1 score.

Hypothesis 2

- *Null Hypothesis(H0):*If Particle Swarm Optimised Convolutional Neural Network is used instead of deep learning algorithms, then there is no significant improvement in classification of Human Activity in terms of accuracy and the F1 score.
- *Alternate Hypothesis(HA):*If Particle Swarm Optimised Convolutional Neural Network is used instead of deep learning algorithms, then there is significant improvement in classification of Human Activity in terms of accuracy and the F1 score.

3.3 Data Understanding

3.3.1 Data Description

The dataset used in this study is downloaded from UCI Machine Learning Repository created at SmartLab, one of the Research Laboratories at DIBRIS at University of Genova in Italy whose major research area follows investigating techniques and algorithms for Computational Intelligence and Data Analytics. (Anguita, 2006) experimented on a group of 30 volunteers within a range of age between 19-48 years who were performing daily activities. Each subject are volunteer performed daily activities which are show in Table 3.3 while carrying a smartphone (Samsung Galaxy S II) that is waist-mounted. The smartphone was embedded with inertial sensors.

With the help of this embedded accelerometer and gyroscope, 3-axial linear acceleration and 3-axial angular velocity were captured at a constant rate of 50Hz. To label the data manually, the experiments are captured in the form of video. Additionally, the sensor signals accelerometer and gyroscope were processed by including noise filters and sampled with fixed width sliding windows of 2.56 seconds and 50% overlap that indicates 128 readings per sliding window. The resulting processed signals is a combination of gravity acceleration and body acceleration components and were subjected into a low pass filter to get the separated components with the gravitational force components cut off at the lower end of the filter. As such, vectors are formed from each window to obtain time and frequency domain variables.

Thus, each record in the dataset includes below features:

- total acceleration and approximate body acceleration which is obtained from Triaxial Accelerometer.
- Triaxial Angular velocity is extracted from gyroscope.
- A 561-feature vector with time and frequency domain variables.
- activity as target.
- An identifier which indicates the participant who performed the experiment.

Raw Signal	Defnition
tBodyAcc-XYZ	Body acceleration in time
tGravityAcc-XYZ	Gravity acceleration in time
tBodyAccJerk-XYZ	Jerk in body acceleration in time
tBodyGyro-XYZ	Body gyroscope measure in time
tBodyGyroJerk-XYZ	Jerk in body gyroscope measure in time
tBodyAccMag	Magnitude of body acceleration in time
tGravityAccMag	Magnitude of gravity acceleration in time
tBodyAccJerkMag	Magnitude of jerk in body acceleration in time
tBodyGyroMag	Magnitude of body gyroscope measure in time
tBodyGyroJerkMag	Magnitude of jerk in body gyroscope measure in time
fBodyAcc-XYZ	Body acceleration in frequency
fBodyAccJerk-XYZ	Jerk in body acceleration in frequency
fBodyGyro-XYZ	Body gyroscope measure in frequency
fBodyAccMag	Magnitude of body acceleration in frequency
fBodyAccJerkMag	Magnitude of jerk in body acceleration in frequency
fBodyGyroMag	Magnitude of body gyroscope measure in frequency
fBodyGyroJerkMag	Magnitude of jerk in body gyroscope measure in frequency

Table 3.1: Description of raw signals from HAR experiment

The features collected from the accelerometer and gyroscope sensors are tabulated in 3.1. The initially obtained raw signals, tAcc-XYZ and tGyroXYZ were the tri-axial signals obtained from the accelerometer and the gyroscope, the total acceleration was further split into tBodyAcc-XYZ and tGravityAcc-XYZ.

To obtain The Jerk signals; tBodyAccJerk-XYZ and tBodyGyroJerk-XYZ, the body ,the angular velocity and acceleration data are recorded with same time difference. Euclidean norm is computed to calculate the magnitude of these 3-dimensional signals. Which resulted in components like tBodyAccMag, tGravityAccMag, tBodyAccJerkMag, tBodyGyroMag, tBodyGyroJerkMag.

A Fast Fourier Transform is included to some of the signals which resulted in the features as fBodyAcc-XYZ, fBodyAccJerk-XYZ, fBodyGyroXYZ, fBodyAccJerkMag, fBody-

GyroMag, fBodyGyroJerkMag.

The 'XYZ' in the tables signifies the three axis directions X, Y, Z for each of the tri-axial signals; where 't' indicates time domain variables and 'f' indicates frequency domain variables. The Table 3.2 shows the complete list of the set of descriptive variables computed for each of the above raw signal.

Descriptive	Defnition
mean()	Average value
std()	Standard deviation
mad()	Median absolute deviation
max()	Maximum value
min()	Minimum value
sma()	Signal magnitude area
energy()	Energy value
iqr()	Interquartile range
entropy()	Signal entropy value
arCoeff()	Autoregression coefficient
correlation()	Correlation coefficient
maxInds()	Index of the largest magnitude frequency component
meanFreq()	Weighted average of the frequency component
skewness()	Skewness of the frequency domain signal
kurtosis()	Kurtosis of the frequency domain signal
bandsEnergy()	Energy of the frequency within the FFT of each window
angle()	Angle between the vectors

Table 3.2: Description of derived signals from HAR experiment

In addition to this, an identifier feature, 'Subject', indicating the user who carried out the activity is included in the dataset. This is considered as 30 variable labels and each label is considered as a each subject indicating an experiment and physical activity of each of this subject in this study is treated as a target feature.

Finally, the target feature indicating the physical activity performed by the subjects during which the sensor data are collected and is stored as 'Activity'. It is a 6 factor categorical variable. The activities carried out and recorded are illustrated in Table 3.3.

Activty ID
Walking
Walking_Upstairs
Walking_Downstairs
Sitting
Standing
Laying

Table 3.3: Target Features - Activities performed by the subjects

3.3.2 Data Exploration

The initial phase of exploring the data set is to examine with the data source to make sure that the data possess correct number of dimensions. Once the data integrity is validated, each of these features has to be understood along with its description with respect to domain.

Analysing the statistical description is one of the ways to get an overview about the shape of the features. The statistical descriptions are different based on the data type of the feature. Mode is usually calculated for for categorical features and for numerical variables median, mean, inter-quartile range is significant to understand the shape of the data.

In order to understand the uniformity and distribution of the data, visualising with plots is significant. This makes easy to identify outliers and the relationship between multiple attributes can be understood in a better way. Additionally, aggregations help to understanding the data at domain level.

To conclude, the current research problem is determined as a Classification task. Thus from the previous data understanding phase, the HAR dataset is explored and consists of total of 563 features, where 561 are numerical independent features, and one categorical feature representing the target, Activity label and other describing the identifier, indicating the subject. Statistical summary along with data visualizations and other explorations would be a part of the current study , listed in chapter 4 .

3.4 Data Preparation

Data preparation is one of the most significant phases in a data mining project. In order to get the desired and meaningful insights in the evaluation phase quality of data plays a major role. Data Preparation involves manipulating the data, so that it is free from noise and other inconsistency.

3.4.1 Data Cleaning

To begin with, the entire dataset should be examined for the missing data Any data point is said to have missing value if the corresponding value shave N/A, 999 or blank field. The second task under data cleaning is outlier detection and analysis. In most cases, an outlier may be caused due to human error in coding or while collecting data. A data point is considered to be an outlier if it is significantly far from the normal curve, so to visualize them box plots and histograms can help greatly to understand. Another significant step in data cleaning phase is to eliminate data redundancy, where the dataset is checked for dulpicate records.

After performing the above data preparation tasks, it is ensured that the data is consistent and accurate, with no noise.Then the final data is ready for modelling stage.

3.5 Modelling

As shown in the figure 4.1, the modelling phase is performed with two version of datasets using algorithms in different categories.. The below section describes the modelling algorithms and its theory. In order to address the research question, the data which is ready after performing data cleaning, data integration is subjected to modelling. Each modelling algorithm is tuned with various best configuration parameters to have the best working model that is intact with the data. The literature review has described multiple classification algorithms typically used for HAR problems. Thus, 6 modeling algorithms are applied on HAR dataset to identify the best model that yield high classification accuracy.

3.5.1 Machine Learning Algorithms

The machine algorithms chosen for Human Activity Recognition are mentioned below.

- Naive Byes: Naïve Byes classifier implements Bayes Theorem providing probabilistic classification [30]. This is suitable for fast computation especially in huge data. They are used for various application like sentiment analysis, text classification, spam filtering etc.,

- Support Vector Machine :Support Vector Machine is one of the baseline models which gained highest accuracy in human activity recognition when compared with other classical machine learning algorithms. SVM is implemented for both classification and regression tasks and this works by building builds the hyperplane margin between classes.

- Random Forest : Random Forest Classifier is nothing but a combination of multiple decision Trees which is an an ensemble learning method for classification, regression and other machine learning tasks

3.5.2 Deep Learning Algorithms

The deep learning algorithms chosen for Human Activity Recognition are mentioned below.

- Long Short Term Memory(LSTM): Long Short-Term Memory networks – are a special kind of Recurrent Neural Network (RNN), capable of learning long-term dependencies. LSTMs are designed to avoid the long-term dependency problem. The LSTM learns to map each sliding window of sensor data to an activity, where the data points or samples in the input sequence are read one at a time, and each time step may consist of one or more variables.

- Convolutional Neural Network (CNN): CNN has been used for time series classification problems especially in classifying real time activity data. The signals in HAR has high similarity between them and is referred as Local dependency, while scale invariant means that the scale remains same for different time and domain frequencies.

3.5.3 Particle Swarm Optimization Based CNN

Though CNN's have showed good results in HAR, there a re multiple parameters to take care to find the optimal CNN architecture. The main focus of any neural network is to minimize the error between training targets and predicted outputs. It is cross-entropy in case of CNN's, which is carried out by backpropagation and gradient descent. Even a simple CNN's have many parameters to tune them. Thus, it is significant to find algorithms which finds and evaluates CNN architecture with less time. Thus, motivated from this, a new PSO-CNN is utilized for Human Activity Recognition. The below Figure 3.2 shows the working of the model

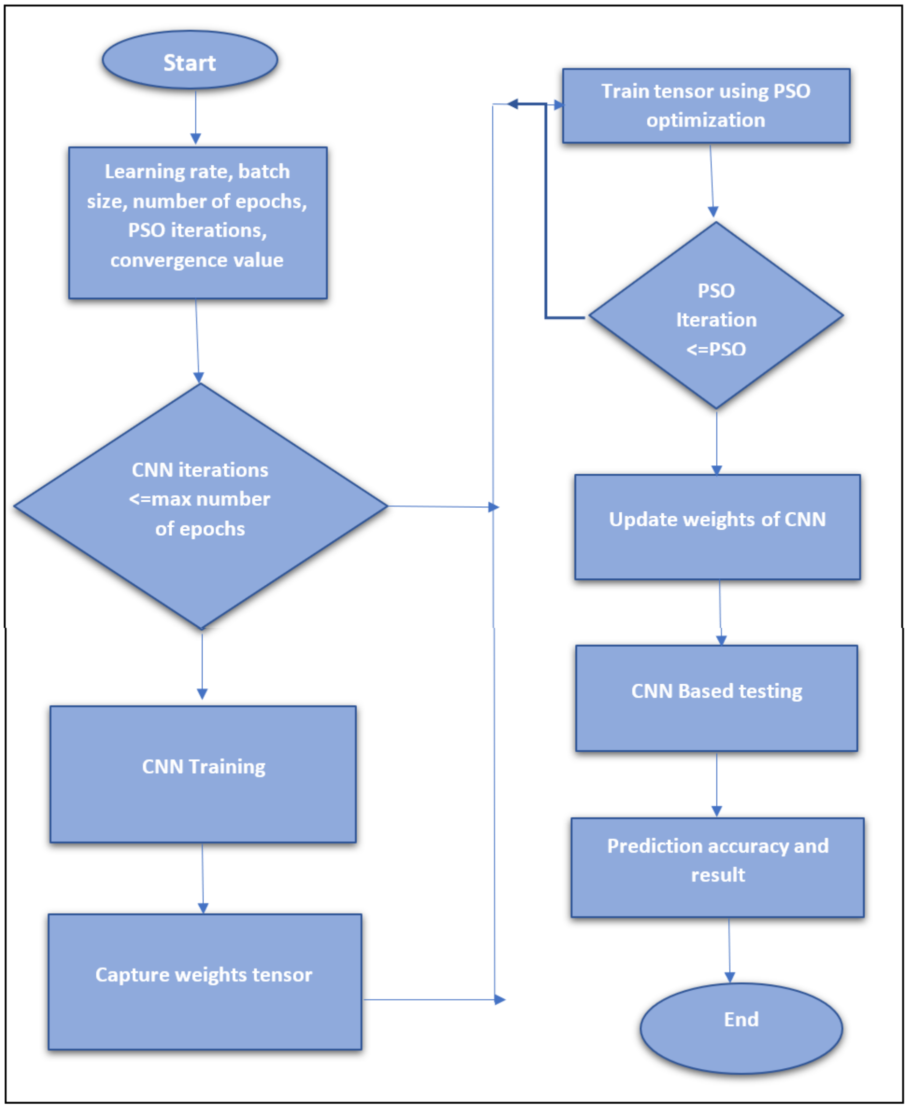

Figure 3.2: Particle Swarm Optimization Training for Convolutional Neural Network

The working of PSO-CNN can be divided into five stages as below -

- CNN Training – The CNN is trained with some pre-denied weights initialixed. . It uses a CNN with 1D convolutional layer, since the HAR dataset consist of signals in shape [samples, time_steps, no of features]. The output is one hot vector encoded which is 6 (target activity to be predicted).

- Pre-PSO Training – Here weights are captured from CNN training and it is converted to particle.
- Particle Swarm Optimization Training - After initializing the values of convergence, cognitive value, social value, number of particles, stopping condition and number of epochs, PSO algorithms searches the hyperplane for optimized vector using the CNN loss function .
- Update CNN Architecture - Using the values of weight in previous phase, the final results are computed. A new CNN architecture is created is created based on these weights rather than basis of the output.
- Computation of Prediction Accuracy and Results - The output of the CNN is formed and the final accuracy , loss valued are evaluated.

3.6 Evaluation Metrics

In order to evaluate the performance of the modeling algorithms, appropriate metric is chosen. Based on the research question that the study is going to address, appropriate performance criteria and its measure has to be chosen. Should also consider the ability and feasibility of the work and the study.

Confusion Matrix Confusion Matrix - Confusion matrix is also known as contingency table, provides a overall performance of the classification model. The Figure 3.3 shows the format of a confusion matrix.

		Predicted	
		Positive	Negative
Actual	Positive	TP	FN
	Negative	FP	TN

Figure 3.3: Confusion Matrix

This confusion matrix provides important elements of the modelling results, they are described below

- True Positives (TP): The number of observations in the positive target class which were correctly classified by the model.
- False Negatives (FN): The number of observations in the positive target class which were incorrectly classified as in the negative target class by the model.
- False Positives (FP): The number of observations in the negative target class which were incorrectly classified as in the positive target class by the model.
- True Negatives (TN): The number of observations in the negative target class which were correctly classified by the model.

Each of these measures have significance that depends on the research question and type of data. Multiple metrics can be concluded from the confusion matrix –
Accuracy – Represents the total number of correct predictions made by the model.

$$Formula : \frac{TP + TN}{TP + TN + FP + FN}$$

Recall – It is the proportion of relevant positive classes identified correctly by the classifier

$$Formula : \frac{TP}{TP + FN}$$

Precision – It is the proportion of positive classes correctly identified considering all the positive classes predicted.

$$Formula : \frac{TP}{TP + FP}$$

F1 Score – The F1 score is the harmonic mean of the precision and recall metrics. This metric allows for a balanced expression of the model's recall and precision.

$$Formula : 2 * \frac{Precision * Recall}{Precision + Recall}$$

Chapter 4

Implementation

This chapter outlines the execution process The research was carried out using the Cross Industry Standard Process for Data Mining (CRISP-DM) methodology which helps in providing a uniform framework and guidelines for data mining projects. This methodology consists of six phases – Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation and Deployment. The Business Understanding phase was covered in chapter three – Design and Methodology section.

4.1 Data Understanding

4.1.1 Data Gathering

Two versions of the data was made available for modelling purposes. These are mentioned as follows -

- Hand-crafted features of activity windows- Version 1: Each recorded window possess a 561 column vector with time and frequency domain variables rectified separately, a activity label ID indicating the activity performed by the subject and an identifier or the ID of the subject who carried out the experiment.
- Raw Inertial sensor data - Version 2 : Raw signals which are tri-axial from the gyroscope and accelerometer sensor are collected by placing wearable sensors on the volunteer called subjects.

4.2 Data Exploration

4.2.1 Distribution of Data points over each activity

In order to get the distribution of each activity, bar graph is plotted which shows that the data points are almost equally distributed which is illustrated in Figure 4.1

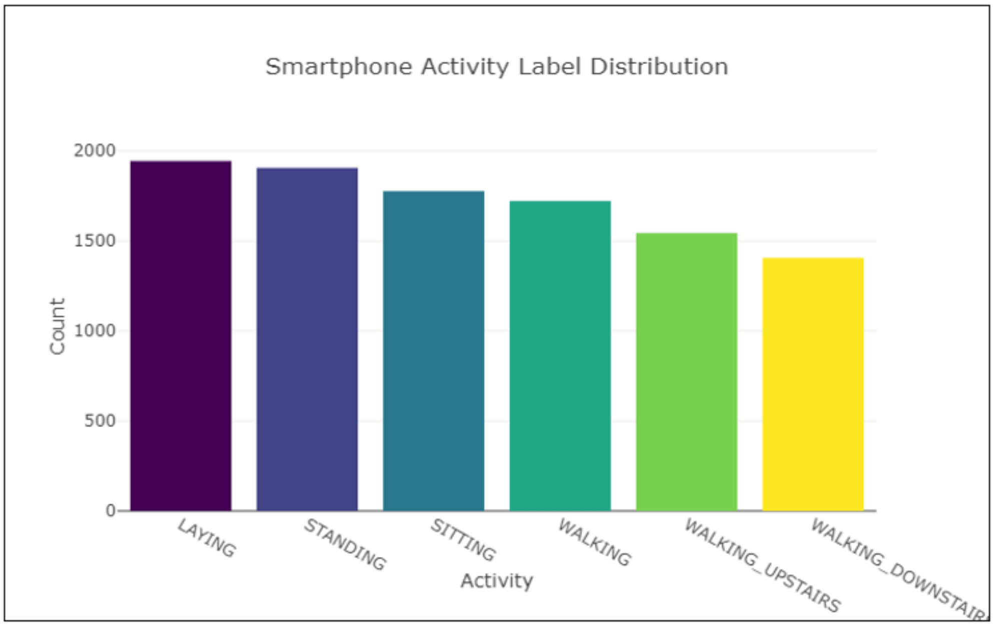

Figure 4.1: Distribution of Data points over each activity

4.2.2 Activity Data provided by each user

Using simple aggregation, each subject with their activity level data are plotted. It is observed from the Figure 4.2 we can say that the distribution is uniform, with respect each of the subjects and also with target class.

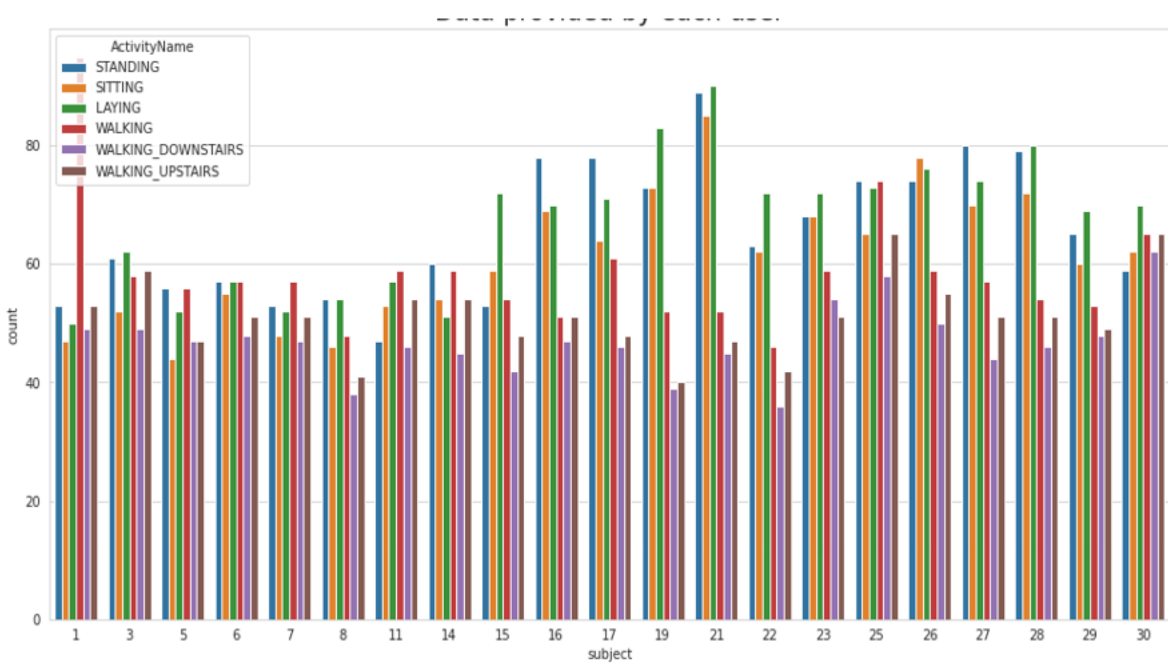

Figure 4.2: Activity Data provided by each user

4.2.3 Analyze the variation in data over Static and Dynamic activities

In static activities like Sitting, Standing and Laying down motion information is not significant, whereas dynamic activities like Walking, Walking Upstairs and Walking Downstairs motion details will be significant. In order to variability of the motion data, the below graph is plotted

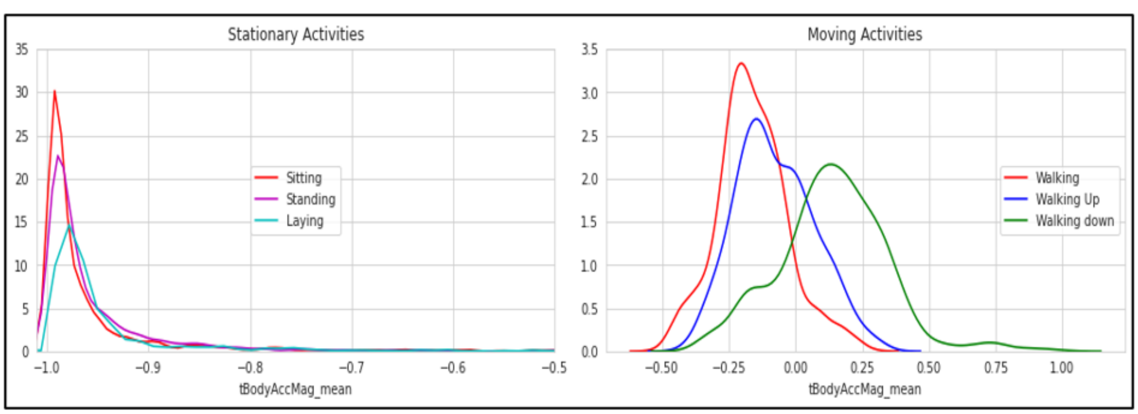

Figure 4.3: Distribution of data points over Static and Dynamic activities

4.2.4 Separation of activities by Magnitude of Acceleration

To find the acceleration varying over each activity, box plot is plotted Figure 4.4 where interesting observations are made which can separate static and dynamic activities. Some of them are listed below -

- If tAccMean is < -0.8 then the activities performed are either Sitting, Laying or Standing.
- If tAccMean is > -0.6 then the activities are either Walking or Walking Downstairs or Walking Upstairs.
- If tAccMean > 0.0 then the activity performed is Walking Downstairs.
- More than 75% of the labels can be classified with some errors.

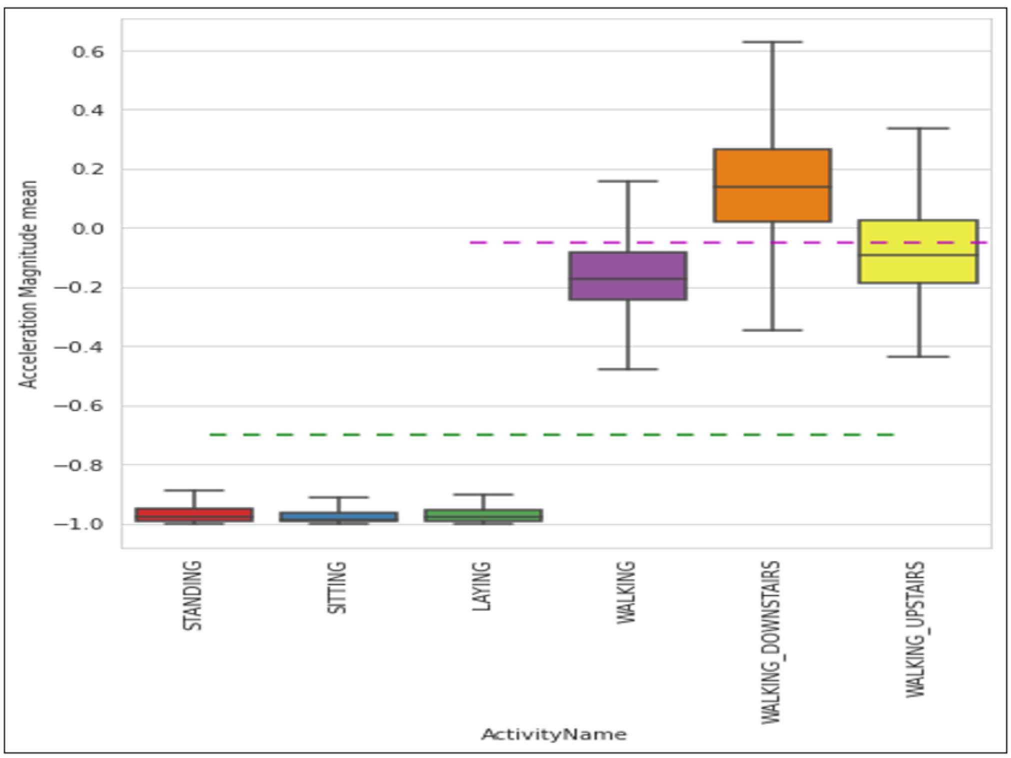

Figure 4.4: Separation of activities by Magnitude of Acceleration

There was no noisy data found while performing the data cleaning process. The range of time and frequency domain variables were seen between +1 and -1, indicating normalization were performed while creating the dataset. The data is checked for missing values and duplicate rows. No such scenarios were found over the entire data.

4.3 Data Preparation

Data Preparation steps are performed separately for both the versions of the data. The below sections justify the same.

4.3.1 Hand-crafted features - version 1

Data Integration is performed to merge different files. The datasets are present a zip folder which consisted of actual data files, feature files and their description separately. Some of the files used for integration are listed below.

- 'activity_labels': a text file that contains the six different activities mentioned.
- 'features': text file that contains the entire names of 561 features.
- features_info': text file that contains detailed descriptions of all of these features.
- 'train' folder 'test' folder: folders containing training and test datasets. Each file is read using pd.read_csv() function into a data frame.

4.3.2 Raw Inertial signals - version 2

The raw data signals are present in folder 'Inertial Signals' for train and test separately. It contains 9 files where 3 files correspond to each x, y and z axis. The files are training data listed below. Similar structure is found for test dataset.

- 'total_acc_x_train': a text file indicating the acceleration signal from the smartphone's accelerometer. Each row represents a 128-element vector. The same description holds same for the 'total_acc_y_train.txt' and 'train.txt' files for the Y and Z axis, respectively.
- 'body_acc_x_train': a text file indicating the body acceleration signal computed by subtracting the gravity from the total acceleration. The same description holds same

for the body_acc_y_train'.txt' and body_acc_z_train'.txt' files for the Y and Z axis, respectively.

- 'body_gyro_x_train': a text file indicating the angular velocity vector which is measured by the gyroscope for each window sample. The same description holds same for the body_gyro_y_train.txt' and body_gyro_z_train.txt' files for the Y and Z axis, respectively.

For Time Series Classification the data must be represented in an appropriate manner to fit the model. This allows the model to associate the signal data with activity class. Each window can correspond to a specific activity. A given window of data samples might contain multiple variables, namely x, y, and z axes of a sensor.

4.3.2.1 Data Transformation

Below steps are performed to transform the data to required shape.

- Using reshape () function in pandas, the data is split into window of 128- time steps with a 50% overlap. Thus, the training array is with shape
  train = (7352, 128, 9) 7352 samples, with 128 timesteps, represented with 9 features where each feature data are contained in a separate file

- **test = (2947, 128, 9)** 2947 samples, with 128 timesteps, represented with 9 features where each feature data are contained in a separate file.

- **train_Y = (7352, 6)** indicating 7352 samples and 6 indicating activity type performed.

- **test_Y = (2947, 6)** indicating 2947 samples 6 indicating activity type performed.

- **get_dummies()** function is used to convert the categorical target activity variable into dummy/indicator variables.

- The training and test data are scaled for normalizing the signal data using Scale.transform() function.

4.4 Modelling

The below Figure shows the distribution of target activities. Modelling algorithms are sued to predict the target class.

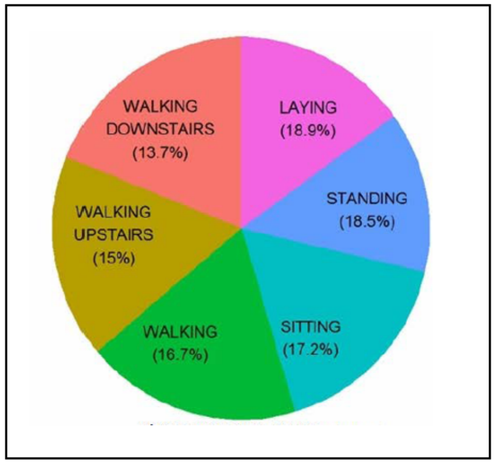

Figure 4.5: Distribution of Target Class

4.4.1 Machine Learning Algorithms – with Hand-crafted features – Version 1

As discussed in chapter 3, Version -1 dataset is modelled with classical machine learning algorithms. Each modeling algorithms along with the parameter settings are discussed below.

4.4.1.1 Naïve Byes

Naïve Byes classifier implements Bayes Theorem providing probabilistic classification [30]. This is suitable for fast computation especially in huge data. They are used for various application like sentiment analysis, text classification, spam filtering etc., They mainly assume that presence or absence of each attribute value is independent from the presence or absence of other values. Here the modelling with Naïve Byes is done with the default setting.

4.4.1.2 Random Forest

Random Forest Classifier is nothing but a combination of multiple decision Trees which is an an ensemble learning method for classification, regression and other machine learning tasks. This performs training by building multiple decision trees and designating the output of the class which is mode in case of classification of the individual trees. In most cases, the default hyperparameters of Random Forest provides good result and it is good at avoiding overfitting (Pretorious, Bierman Steel, 2016). GridSearchCV is run on Random Forest to find all the best combinatiosn of parameters, but there was no much improvement found from the GridSearchCV.

4.4.1.3 Support Vector Machine

Support Vector Machine is one of the baseline models which gained highest accuracy in human activity recognition when compared with other classical machine learning algorithms. SVM is implemented for both classification and regression tasks and this works by building builds the hyperplane margin between classes. To classify the data points into different classes, there could be many possible hyperplanes available. The algorithms intend to find the hyperplane which has the, maximum place which is nothing but maximum distance between the classes. It provides the options to choose the kernels like linear, Radial Basis Function and polynomial. For the modelling, linear kernel is chosen.

4.4.2 Deep Learning Algorithms – With Raw Inertial Signals – Version 2

Sensor based activity recognition requires domain-level knowledge about human activities to analyse even the minute details of sensor data. Though traditional machine algorithms have shown some extra-ordinary performance in classifying human activities, it requires domain knowledge and few labelled data [34,35]. In contrast, Deep Learning exhibits the capability of training real time activity data that are coming in stream or sequence. Considering Human Activity Recognition as a Time Series Classifications problem which aims at classifying sequences of sensor data, two well-known algorithms LSTM and CNN are modelled on the raw inertial signal data - Version 2. The below sections explains the model parameter settings.

4.4.2.1 Long Short-Term Memory (LSTM)

Long Short-Term Memory networks – are a special kind of Recurrent Neural Network (RNN), capable of learning long-term dependencies. LSTMs are designed to avoid the long-term dependency problem. The LSTM learns to map each sliding window of sensor data to an activity, where the data points or samples in the input sequence are read one at a time, and each time step may consist of one or more variables.

Layer Topology

Single Layer LSTM is used with the batch size is 16, and the activation function used in the LSTM layers is the basic Sigmoid function.

The LSTM requires a three-dimensional input with [samples, time_steps, features].So, the data is loaded, where one sample is equal to one window of the time series data, where each window consists 128 time steps, and a time step possess 9 features. Further, each series of data points is partitioned into overlapping windows of 2.56 seconds of data, or 128-time steps with 9 features. Thus the input_shape =[timesteps, features] corresponds to [128, 9]. Here the output for the model will be a six-element vector indicating the probability of a given window corresponds to each of the 6 activity types. Below Figure 4.6 represents the summary of the LSTM model implemented.

```
************Multi Layer LSTM**************
Model: "sequential_11"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lstm_16 (LSTM)               (None, 128, 32)           5376
_________________________________________________________________
dropout_16 (Dropout)         (None, 128, 32)           0
_________________________________________________________________
lstm_17 (LSTM)               (None, 28)                6832
_________________________________________________________________
dropout_17 (Dropout)         (None, 28)                0
_________________________________________________________________
dense_11 (Dense)             (None, 6)                 174
=================================================================
Total params: 12,382
Trainable params: 12,382
Non-trainable params: 0
_________________________________________________________________
```

Figure 4.6: LSTM Model Description for HAR Data

4.4.2.2 Convolutional Neural Network (CNN)

CNN has achieved good results in image classification, sentiment analysis and speech recognition task by extracting features from signals. CNN has been used for time series classification problems especially in classifying real time activity data because of scaling invariable and local dependencies. Local dependency means the nearby signals in Human Activity Recognition (HAR) are likely to be correlated to each other, while scale invariant means that the scale remains same for different time and domain frequencies. Thus, CNN has a better understanding of learning features that are present in recursive patterns.

Layer Topology

Keras library provides easy methods for creating Convolutional Neural Networks (CNNs) of 1, 2, or 3 dimensions; Conv1D, Conv2D and Conv3D. Generally, CNNs are used for image classification, in which model accepts a 2-diemnsiomal input indicating the image's height , width and colour channels for feature learning. The same ideology is implemented for 1-dimensional series of data. The model learns the features from series of data and maps the internal features of series. Hence 1D CNN is efficient for learning features from a fixed length window of the entire dataset.

The architecture of our network is summarized in Figure 4.7 It contains fuly connected layers and activation layers, two 1D convolutional, one max-pooling and fully connected layers. Each layer of the CNN consisted of a convolutional layer, and it is followed by a spatial pooling layer. 70% of the data is used for training and the 30% were considered as the test data and evaluated at the time of prediction. The first layer, the 1D convolutional layer, is convolved with a filter and applied activation function to get the output feature map. Each output feature map further merge convolutions with multiple input features. Every convolution layer is followed by a max-pooling layer

The max-pooling layer serves the purpose of reducing the size of the input map. In many cases, it is done simply for the size reduction. The max-pooling ensures splitting up the matrix form of each convolution layer's outputs into small non-overlapping grids, and proceeds by picking up the the maximum value in each non-overlapping grid as the value in the reduced size matrix. Thus, spatial abstractness of features can be increased by using max-pooling layer.

The next layer is a fully connected with a linear function the output dimension depends on input dimensions. This layer picks all neurons in the previous layer (either fully connected, or pooling, or convolutional) and in turn connects them to every single neuron .In the end, to translate the input signals to output signals , activation layers are used.

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv1d_1 (Conv1D)            (None, 126, 32)           896
_________________________________________________________________
conv1d_2 (Conv1D)            (None, 124, 32)           3104
_________________________________________________________________
dropout_21 (Dropout)         (None, 124, 32)           0
_________________________________________________________________
max_pooling1d_1 (MaxPooling1 (None, 62, 32)            0
_________________________________________________________________
flatten_1 (Flatten)          (None, 1984)              0
_________________________________________________________________
dense_13 (Dense)             (None, 50)                99250
_________________________________________________________________
dense_14 (Dense)             (None, 6)                 306
=================================================================
Total params: 103,556
Trainable params: 103,556
Non-trainable params: 0
_________________________________________________________________
```

Figure 4.7: CNN Model Description for HAR Data

4.4.3 Particle Swarm Optimized Convolutional Neural Network (PSO-CNN)

CNN's have proved to obtain best results in most of the image classification benchmark datasets, overcoming the need of hand-crafted features to learn. [1,2] Nevertheless, it is a challenging task to design a meaningful CNN architecture. Hence, PSO is utilized to search for an optimal architecture in deep CNNs using varied length of the particles. Particles can increment in size without any upper limit. PSO-CNN architecture is explained in chapter 3, and parameter settings are can be classified into three categories which are mentioned as below.

4.4.3.1 Parameter Settings for Particle Swarm Optimization

The parameters used in this category control the behaviour of the Particle Swarm Optimisation algorithm. It consists three parameters namely, the number of iterations, the size of the swarm, (Cg) represents the probability of selecting a layer from global best while computing each particle's velocity. The number of iterations specify the actual number of iterations that the optimal search algorithm will run before optimization is completed. The best CNN architecture that is with best accuracy is saved at after the optimization of the last particle. The swarm size indicates number of particles in the PSO algorithm. Here, each individual particle is a one complete CNN architecture whose performance to be tested by the algorithm. Table 4.1

Description	**Value**
Number of iterations	10
Swarm Size	20
Cg	0.5

Table 4.1: Parameter Initialization for Particle Swarm Optimization

4.4.3.2 Parameter Settings for initializing CNN architecture

The parameter settings used in the second category control the initial movement of the particles. It involves eight parameters listed in the table below Table 4.2. In this step, an initial population of swarm which is of CNN architectures. This initial population consists of individuals withCNN architectures picked randomly as defined by these parameters. To limit the number of feature maps from a output of a convolution layer, minimum and maximum number of outputs must be defined. The size of the kernel is always chosen will between the range of the minimum and maximum size of a convolutional kernel. Only the initial architecture is controlled by these parameters, after first initialization the architecture is updated based on design specified.

Description	Value
Minimum number of outputs from a conv layer	3
Maximum number of outputs from a conv layer	256
Minimum number of neurons in a FC layer	1
Minimum number of layers	3
Maximum number of layers	20
Minimum size of a Conv kernel	3x3
Maximum size of a Conv kernel	7x7

Table 4.2: Parameter Settings for initializing CNN architecture

4.4.3.3 Parameter Settings for training Convolutional Neural Network

The parameters here specify the training process of each particle. It includes four parameters to set which is listed in Table 4.3 These parameters control the weight updating process during the training of each particle. The number of epochs specifies the total number of times the particle is trained using entire dataset before its accuracy is evaluated. The dropout parameter is used in the particle to avoid overfitting.

Description	Value
epochs for particle evaluation	100
epochs for global best	256
Dropout rate	0.5
Batch normalizer layer outputs	yes

Table 4.3: Parameter Settings for training Convolutional Neural Network

Furthermore, the model includes batch normalization between the layers to avoid overfitting during training process.

4.5 Evaluation Method

The performance of the modelling algorithms is done using various measures. HAR is a Multilablel classification problem. The main challenge in classification task is to correctly classify the target variables. Only accuracy score cannot give us the overall performance of the model. Hence, confusion matrix which gives the actual number of correct and incorrect predictions made for each target class is considered. Additionally, precison, recall and f1 score is computed. But for comparisions accuracy and f1 score are considered. Details about the metrics are discussed in Chapter 3.

Chapter 5

Results, Evaluation and Discussion

5.1 Model Results

5.1.1 Machine Learning Algorithms Results - Using hand crafted features - Version 1

Since, this a classification task, the evaluation metric chosen is Accuracy and F1 score. Along with this, confusion matrix is also computed.

Naive Byes: Test Accuracy = 77.02%

Confusion Matrix

Figure 5.1: Confusion Matrix for Naive Byes

Classification Report

```
-------------------------
| Classifiction Report |
-------------------------
                    precision    recall  f1-score   support

            LAYING       0.96      0.60      0.74       537
           SITTING       0.58      0.75      0.65       491
          STANDING       0.80      0.86      0.83       532
           WALKING       0.82      0.84      0.83       496
WALKING_DOWNSTAIRS       0.83      0.61      0.70       420
  WALKING_UPSTAIRS       0.76      0.96      0.84       471

          accuracy                           0.77      2947
         macro avg       0.79      0.77      0.77      2947
      weighted avg       0.79      0.77      0.77      2947
```

Figure 5.2: Classification Report for Naive Byes

Analysis:

The Naive bayes model is basic generative model that performs based on the probability function. The Table 5.1 represents the confusion matrix of the Naive bayes algorithm modelled with hand-crafted fevature data achieved the accuracy of 77.02%. The accuracy is not satisfactory when compared with the other algorithms. Most of the cases in LAYING are misclassified as SITTING, and SITTING is wrongly misclassified as STANDING. From the classification report, f1 score is 74% which shows poor performance in classifying activities.

Support Vector Machine: Test Accuracy = 95.04%

Confusion Matrix

Figure 5.3: Confusion Matrix for Support Vector Machine

Classification Report

```
------------------------
| Classifiction Report |
------------------------
                    precision    recall  f1-score   support

            LAYING       1.00      1.00      1.00       537
           SITTING       0.94      0.89      0.91       491
          STANDING       0.91      0.95      0.93       532
           WALKING       0.94      0.98      0.96       496
WALKING_DOWNSTAIRS       0.99      0.91      0.95       420
  WALKING_UPSTAIRS       0.93      0.96      0.94       471

          accuracy                           0.95      2947
         macro avg       0.95      0.95      0.95      2947
      weighted avg       0.95      0.95      0.95      2947
```

Figure 5.4: Classification Report for Support Vector Machine

Analysis:

Support Vector Machine showed good results with accuracy of 95.04%. The Table 5.3 shows the confusion matrix for SVM. From the table it is evident that , all the activity types were classified almost except for STANDING and SITTING. This might be due to the more similiarities in the data points because of which a hyperplane of separation was not able to form.

Additionally, SVM achived great results with Version -1 Dataset without any parameter tuning. Thus, SVM can be considered as of the best base classifier to achive high accuracy without changing the state-of-the art of the model.

Random Forest Test Accuracy = 92.36%

Confusion Matrix

Figure 5.5: Confusion Matrix for Random Forest

<u>Classification Report</u>

```
-------------------------
| Classifiction Report |
-------------------------
                    precision    recall  f1-score   support

            LAYING       1.00      1.00      1.00       537
           SITTING       0.91      0.88      0.89       491
          STANDING       0.90      0.92      0.91       532
           WALKING       0.89      0.98      0.93       496
WALKING_DOWNSTAIRS       0.96      0.84      0.90       420
  WALKING_UPSTAIRS       0.90      0.90      0.90       471

          accuracy                           0.92      2947
         macro avg       0.92      0.92      0.92      2947
      weighted avg       0.92      0.92      0.92      2947
```

Figure 5.6: Classification Report for Random Forest

Analysis:

The table 5.5 exhibits the results for the random forest with achieved accuracy of 92.36%.

Though Random Forest has achieved satisfactory results in classifying activities, it failed to classify the activities between SITTING and STANDING . Some misinterpretations are also seen for activities like WALKING_UPSTAIRS and WALKING_DOWNSTARIRS. One of the reason could be the Random Forest performs well in case of categorical features.This is because, algorithms can easily create rules and partition the data if the variables are categorical.

5.1.2 Deep Learning Algorithms Results

Long Short Term Memory (LSTM) : Test Accuracy = 84.71

Confusion Matrix

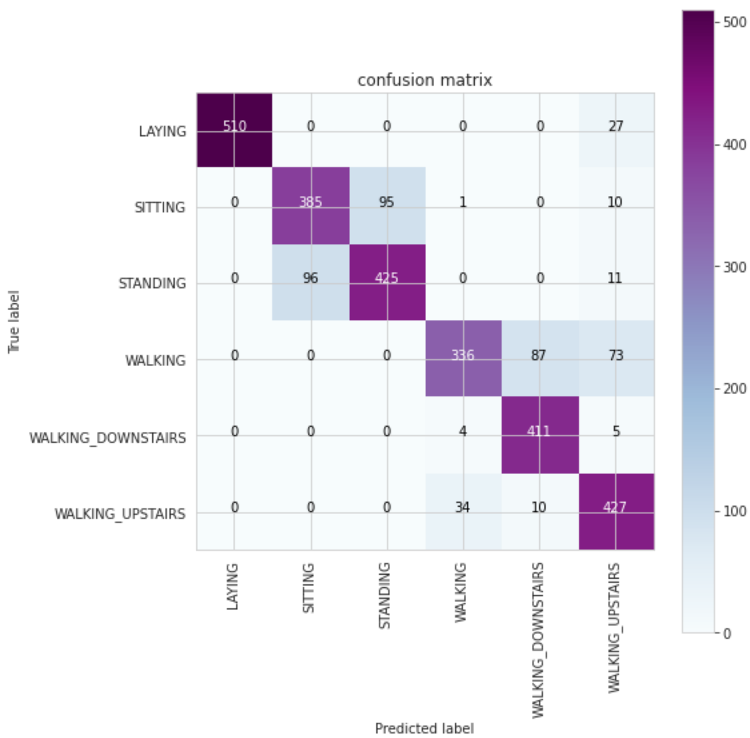

Figure 5.7: Confusion Matrix for LSTM

Classification Report

	precision	recall	f1-score	support
Downstairs	0.90	0.68	0.77	496
Jogging	0.77	0.91	0.83	471
Sitting	0.81	0.98	0.89	420
Standing	0.80	0.78	0.79	491
Upstairs	0.82	0.80	0.81	532
Walking	1.00	0.95	0.97	537
accuracy			0.85	2947
macro avg	0.85	0.85	0.84	2947
weighted avg	0.85	0.85	0.84	2947

Figure 5.8: Classification Report for LSTM

Analysis:

LSTM network models is one of the type of recurrent neural network has a feature of learning and remembering long sequences of continuous input data. They are mainly used for data with long sequences upto 500 time steps. Figure 5.7 shows the confusion matrix for LSTM obtained. The data is continous signal input. Howver, LSTM did not achive great score in accuracy , this might be due to the window size chosen. The accuracy score is 84.71% and f1 score is 84.42% From the classification matrix, it is observed that some of the activities are classified incorrectly. The classifier misinterpreted nearly 95 cases from STANDING to SITTING. Additionally, WALKING and WALKING_DOWNSTAIRS are misclassified.

Convolutional Neural Network (CNN): Test Accuracy = 92.64%

Confusion Matrix

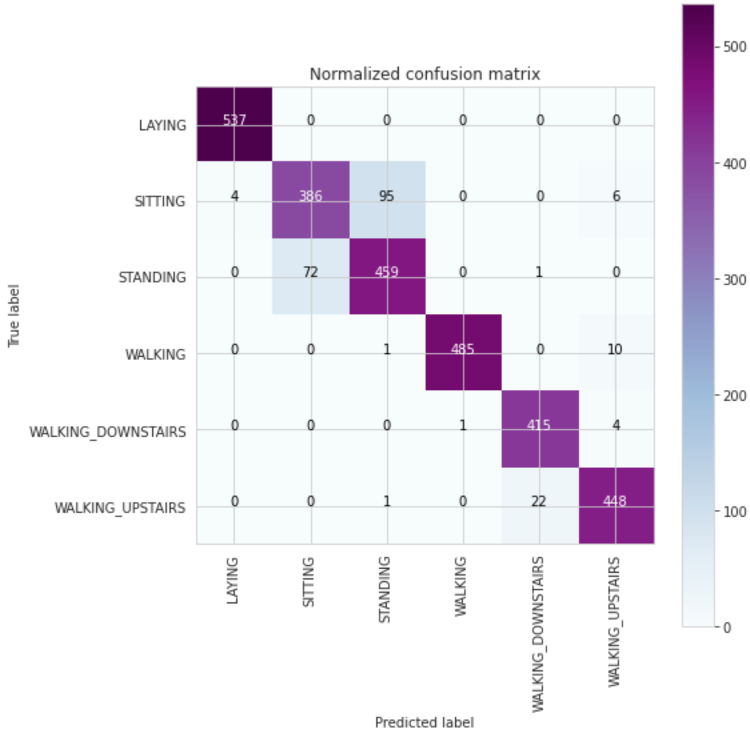

Figure 5.9: Confusion Matrix for CNN

Classification Report

	precision	recall	f1-score	support
Downstairs	1.00	0.98	0.99	496
Jogging	0.96	0.95	0.95	471
Sitting	0.95	0.99	0.97	420
Standing	0.84	0.79	0.81	491
Upstairs	0.83	0.86	0.84	532
Walking	0.99	1.00	1.00	537
accuracy			0.93	2947
macro avg	0.93	0.93	0.93	2947
weighted avg	0.93	0.93	0.93	2947

Figure 5.10: Classification Report for CNN

Analysis:

The Figure 5.9 shows the confusion matrix for CNN model obtained. The model showed accuracy of 92.64% which is a good result with input as raw signal data. The data is not hand-crafted and with the signals directly from the sensors can be directly applied to model without any much data pre-procesing effort. Additionally, the f1 score achieved is 92.7% .From the table 4.15, it is evident that the model was able to classify all the activity with least misinterpretations except for STANDING and SITTING. This is observed in all models, sometimes caused when the users does not show much variability in movement while SITTING and STANDNING. Neverthless, the CNN can be considered as one of the best classifier in terms of activity classification.The model was trained with only 30 epochs, and with less parameter tuning. Thus, CNN is capable of achieving promising results with raw sensor data.

5.1.3 PSO-CNN Results

Confusion Matrix

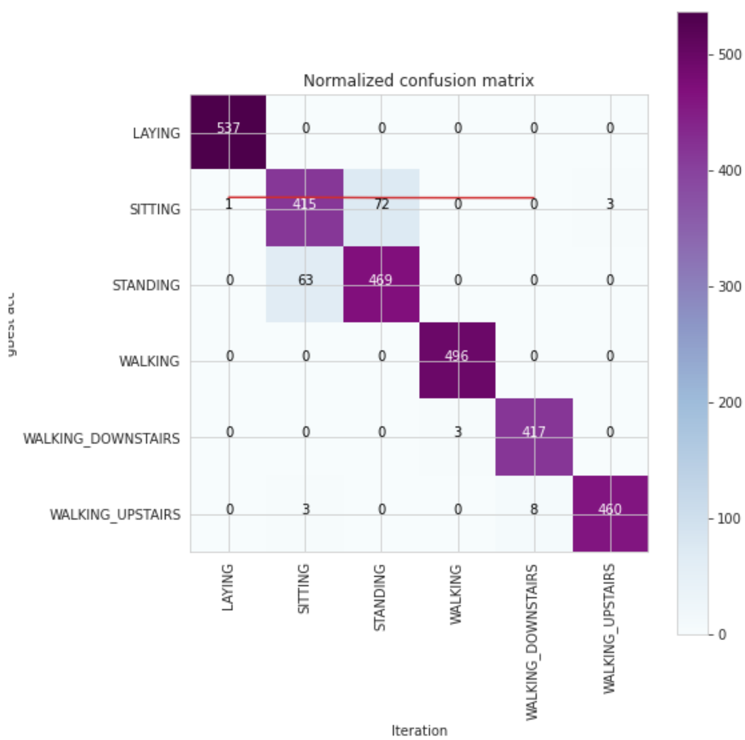

Figure 5.11: Confusion Matrix for PSO-CNN

Classification Report

	precision	recall	f1-score	support
Downstairs	1.00	0.99	1.00	496
Jogging	0.99	0.96	0.98	471
Sitting	0.95	1.00	0.98	420
Standing	0.85	0.81	0.83	491
Upstairs	0.84	0.87	0.85	532
Walking	1.00	1.00	1.00	537
accuracy			0.94	2947
macro avg	0.94	0.94	0.94	2947
weighted avg	0.94	0.94	0.94	2947

Figure 5.12: Classification Report for PSO-CNN

Analysis:

The Figure 5.12 shows the confusion matrix for PSO-CNN model obtained. The model showed accuracy of 93.64% which is a good result with input as raw signal data. The data is not hand-crafted and with the signals directly from the sensors can be directly applied to model without any much data pre-procesing effort. The model was able to classify activities correctly when compared to other models. Additionally, the f1 score achieved is 93.57% .From the Figure 5.12, it is evident that the model was able to classify all the activity with least misinterpretations except for STANDING and SITTING. PSO-CNN model was built to find the best CNN architecture with minimum effort. This also overcomes the local minima problem of the backpropogation training algorithms. The experiment was conducted with 20 epochs , which is less than base CNN model.

5.2 Evaluation Of results

The performance of PSO-CNN is evaluated against Machine Learning and Deep Learning Algorithms

5.2.1 Comparing PSO-CNN with machine learning algorithms

The analysis of results is performed by comparing PSO-CNN with Machine Learning Algorithms. The below Figure 5.13 shows the results.

Dataset	Algorithms	Accuracy	F1 score
Hand-crafted features -Version 1	Naïve Byes	77.02%	71.59%
	Support Vector Machine	95.04%	95.1%
	Random Forest	92.36%	91.78%
Raw Inertial Signal – Version 2	PSO-CNN	93.64%	93.62%

Figure 5.13: Comparsion of PSO-CNN with Machine Learning ALgorithms

From the table, it is evident that Support Vector Machine achieved accuracy of 95.04% and F1 score 95.1%. The machine learning models were built using Hand-crafted features -Version 1 Dataset. The model achieved satisfactory results without performing any Data Dimensionality reduction techniques. On the other hand, PSO-CNN also achieved considerable results with raw sensor data with accuracy of 93.64%. However, the hand-crafted feature extraction process requires human effort to manually design the features.

5.2.2 Comparing PSO-CNN with Deep learning algorithms

The analysis of results is performed by comparing PSO-CNN with Deep Learning Algorithms. The below Figure 5.14 shows the results.

Dataset	Algorithms	Accuracy	F1 score
Raw Inertial Signal – Version 2	LSTM	84.71%	84.42%
	CNN	92.64%	92.71%
	PSO-CNN	93.64%	93.62%

Figure 5.14: Comparsion of PSO-CNN with Deep Learning ALgorithms

From the table, it is clear that PSO-CNN was able to achieve high performance of accuracy when compared with LSTM and CNN models. LSTM performance was low with accuracy 84.71% and F1 score with 84.42%. This, PSO-CNN gained better results than the state-of-the art CNN model. For a classification problem, the capability of the modeling algorithm to classify each target class correctly also plays a major role. Each The algorithm's ability to classify each activity like walking, sitting, laying are discussed in section 5.1 . From the classification report of PSO-CNN Figure 5.11,it is evident that PSO-CNN was able to classify most number of activities correctly.

5.3 Hypothesis Evaluation

Two hypothesis are formed for the evaluation of the experiment

Hypothesis 1

- *Null Hypothesis(H0):*If Particle Swarm Optimised Convolutional Neural Network is used instead of supervised machine learning then there is no significant improvement in classification of Human Activity in terms of accuracy and the F1 score.
- *Alternate Hypothesis(HA):*If Particle Swarm Optimised Convolutional Neural Network is used instead of supervised machine learning , then there is significant improvement in classification of Human Activity in terms of accuracy and the F1 score.

From the Figure 5.13 which compares the results of PSO-CNN with machine learning algorithms, it is evident that PSO-CNN did not achieve better results than SVM. Hence, there is no significant evidence to reject the null hypothesis.

Hypothesis 2

- *Null Hypothesis(H0):*If Particle Swarm Optimised Convolutional Neural Network is used instead of deep learning algorithms, then there is no significant improvement in classification of Human Activity in terms of accuracy and the F1 score.
- *Alternate Hypothesis(HA):*If Particle Swarm Optimised Convolutional Neural Network is used instead of deep learning algorithms, then there is significant improvement in classification of Human Activity in terms of accuracy and the F1 score.

From the Figure 5.14, it is evident that PSO-CNN achieved higher accuracy than deep learning algorithms CNN and LSTM.Hence we have significant evidence to reject the null hypothesis.

5.4 Strength and Limitations

This section outlines the strengths and limitations of the design flow. To current research employs various multiple modelling algorithms , so that the dataset in study is completely explored. Additionally, two versions of the datasets are being tested. This provides a strong evidence to prove the results. The proposed approach not only compares the performance of the algorithms, but also provide suggestions based on experimental results to choose the appropriate methodology (machine learning or deep learning) to handle Human Activity Recognition tasks.

However, the experiments were conducted without much hyper tuning the parameters. The results may vary to some extent more time is invested in setting the parameters.

Chapter 6

Conclusion

This section provides a overall review of the current research. Gives a summary of the research overview, problem definition along with key findings in Experiment Design. Suggestions for future work are highlighted.

6.1 Research Overview

The main aim of the research was to predict the human activity using the data collected from Inertial sensors. Detailed analysis on the existing research is made which includes the type of data used and approaches used for Human Activity Recognition task. Broadly, HAR task was achieved using two approaches namely machine learning and deep learning by which used hand – crafted features data and raw inertial sensor data, respectively.

There was not much exploration done on the meta-heuristic optimisation algorithms like Particle Swarm Optimization. With this motivation and considering the gaps in research, the current study was aimed to perform the experiments with two versions of the datasets using various algorithms. As a result, the data gathering, and preparation was performed separately for both the datasets. To compare the performance of PSO based CNN, various modelling algorithms were chosen, and was evaluated using the performance metrics.

The research conducted with the objective to find the classifier with high predictive accuracies compared with two different family of modelling algorithms.

6.2 Problem Definition

Identifying the gaps in the existing literature, and analysing the best approaches for HAR, the performance of PSO-CNN was compared and evaluated against the two family of algorithms namely classical machine learning algorithms and Deep Learning algorithms.

"To what extent can the Particle Swarm Optimized Convolutional Neural Network can significantly enhance the recognition of human activity from raw inertial sensor data when compared with supervised machine learning algorithms and Deep Learning Algorithms "

6.3 Design/Experimentation, Evaluation & Results

CRISP-DM approach was followed through out the project to get the best outcomes at each step. Accordingly, the implementation began with performing Data Gathering, Data Understanding and Data Preparation for both the data sets separately.

The design involved performing experiments with two versions of the data sets. For hand-crafted features, Machine Learning Algorithms were used for modelling. And since, deep neural networks have capability to take the raw input without any domain-knowledge applied, raw inertial signals data was used. Furthermore, the performance of PSO-CNN is evaluated with suitable metric.

The results were tabulated and detailed analysis was given and it is proved that PSO-CNN showed good results than Deep Learning algorithms , but failed to achieve satisfactory results compared to machine learning algorithms.

6.4 Contributions and impact

Detailed literature review was performed emphasizing on the applications of Human Activity Recognition in various fields. In particluar, Sensor Based HAR is highlighted for the readers. This also detailed about the current state of the art techniques in HAR.

A systematic investigation is done for importing two versions of the sensor datasets. This can be used as reference for future works.

Illustrated that PSO based CNN proved to be the best classifier for data where human-engineered feature knowledge is not needed.

Additionally, the work tries to enhance the performance of state-of-the-art design of the CNN model by using Optimisation. This adds up to the generalization of using PSO-CNN model for other Activity Recognition tasks.

In the current research PSO algorithm is used to find the optimal architectures in deep convolutional neural network. Furthermore it make use of the benefits of global and local exploration capabilities of the particle swarm optimization technique PSO and the gradient descent back-propogation thereby to form a efficient searching algorithm this is because the performance of of deep convolution network extremely depends on their network structure used and hyper-parameter selections.

In order to find the best hyper parameters lot of training time is employed which requires the deep understanding of CNN architecture and also the domain knowledge. Hence PSO-CNN is employed to optimize these parameter configurations and through which efficient parameters are evolved that would increase the performance with less training time.

6.4.1 Future Work and recommendations

The current research can be explored and improved in many ways so as to improve the human activity recognition tasks. The proposed approach also provides a flexible methodology where one can change the initial parameter settings of both PSO and CNN. In this way a trade-off between the model generalization capabilities and complexity of the model can be justified. From the experimental results it is illustrated that PSO has been shown to converge faster and find the best configuration with less training time. This exceed he performance of state-of-the-art results obtained in the domain of HAR.

To some extent, the algorithm failed to recognize the similar he activities like WALKING_UPSTAIRS and WALKING, LAYING and SITTING . This may be due to the insufficient data. The solution can be further explored with large time series data.

The experiment is conducted to explore the capability of deep learning algorithms in HAR tasks. In order to generalize the model capability, this can be applied to other Activity Recognition tasks which includes Time Series data.

References

Abowd, D., Dey, A. K., Orr, R., & Brotherton, J. (1998). Context-awareness in wearable and ubiquitous computing. *Virtual Reality*, *3*(3), 200–211.

Ahad, M. A. R., Tan, J., Kim, H., & Ishikawa, S. (2008). Human activity recognition: Various paradigms. In *2008 international conference on control, automation and systems* (pp. 1896–1901).

Al-Obeidat, F., Belacel, N., & Spencer, B. (2019). Combining machine learning and metaheuristics algorithms for classification method proaftn. In *Enhanced living environments* (pp. 53–79). Springer.

Arel, I., Rose, D. C., & Karnowski, T. P. (2010). Deep machine learning-a new frontier in artificial intelligence research [research frontier]. *IEEE computational intelligence magazine*, *5*(4), 13–18.

Aslan, M. F., Durdu, A., & Sabanci, K. (2020). Human action recognition with bag of visual words using different machine learning methods and hyperparameter optimization. *Neural Computing and Applications*, *32*(12), 8585–8597.

Aurangzeb, K., Haider, I., Khan, M. A., Saba, T., Javed, K., Iqbal, T., ... Sarfraz, M. S. (2019). Human behavior analysis based on multi-types features fusion and von nauman entropy based features reduction. *Journal of Medical Imaging and Health Informatics*, *9*(4), 662–669.

Bengio, Y. (2013). International conference on statistical language and speech processing.

Bhat, G., Deb, R., & Ogras, U. Y. (2019). Openhealth: Open-source platform for wearable health monitoring. *IEEE Design & Test*, *36*(5), 27–34.

bin Abdullah, M. F. A., Negara, A. F. P., Sayeed, M. S., Choi, D.-J., & Muthu, K. S. (2012). Classification algorithms in human activity recognition using smartphones. *International Journal of Computer and Information Engineering*, *6*(77-84), 106.

Bobick, A. F., & Davis, J. W. (2001). The recognition of human movement using temporal templates. *IEEE Transactions on pattern analysis and machine intelligence*, *23*(3), 257–267.

Candamo, J., Shreve, M., Goldgof, D. B., Sapper, D. B., & Kasturi, R. (2009). Understanding transit scenes: A survey on human behavior-recognition algorithms. *IEEE transactions on intelligent transportation systems*, *11*(1), 206–224.

Cao, P., Zhao, D., & Zaïane, O. R. (2013). A pso-based cost-sensitive neural network for imbalanced data classification. In *Pacific-asia conference on knowledge discovery and data mining* (pp. 452–463).

Carvalho, M., & Ludermir, T. B. (2006). Particle swarm optimization of feed-forward neural networks with weight decay. In *2006 sixth international conference on hybrid intelligent systems (his'06)* (pp. 5–5).

Carvalho, M., & Ludermir, T. B. (2007). Particle swarm optimization of neural network architectures andweights. In *7th international conference on hybrid intelligent systems (his 2007)* (pp. 336–339).

Chavarriaga, R., Sagha, H., Calatroni, A., Digumarti, S. T., Tröster, G., Millán, J. d. R., & Roggen, D. (2013). The opportunity challenge: A benchmark database for on-body sensor-based activity recognition. *Pattern Recognition Letters*, *34*(15), 2033–2042.

Chen, Y., & Atwood, M. E. (2007). Context-centered design: bridging the gap between understanding and designing. In *International conference on human-computer interaction* (pp. 40–48).

Chen, Y., & Xue, Y. (2015). A deep learning approach to human activity recognition based on single accelerometer. In *2015 ieee international conference on systems, man, and cybernetics* (pp. 1488–1492).

Cleland, I., Kikhia, B., Nugent, C., Boytsov, A., Hallberg, J., Synnes, K., ... Finlay, D. (2013). Optimal placement of accelerometers for the detection of everyday activities. *Sensors*, *13*(7), 9183–9200.

Cook, D., Feuz, K. D., & Krishnan, N. C. (2013). Transfer learning for activity recognition: A survey. *Knowledge and information systems*, *36*(3), 537–556.

Dehuri, S., Roy, R., Cho, S.-B., & Ghosh, A. (2012). An improved swarm optimized functional link artificial neural network (iso-flann) for classification. *Journal of Systems and Software*, *85*(6), 1333–1345.

Derpanis, K. G., Sizintsev, M., Cannons, K. J., & Wildes, R. P. (2012). Action spotting and recognition based on a spatiotemporal orientation analysis. *IEEE transactions on pattern analysis and machine intelligence*, *35*(3), 527–540.

Donahue, J., Anne Hendricks, L., Guadarrama, S., Rohrbach, M., Venugopalan, S., Saenko, K., & Darrell, T. (2015). Long-term recurrent convolutional networks for visual recognition and description. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 2625–2634).

Dong, B., & Wang, X. (2016). Comparison deep learning method to traditional methods using for network intrusion detection. In *2016 8th ieee international conference on communication software and networks (iccsn)* (pp. 581–585).

Edel, M., & Köppe, E. (2016). Binarized-blstm-rnn based human activity recognition. In *2016 international conference on indoor positioning and indoor navigation (ipin)* (pp. 1–7).

Ehatisham-Ul-Haq, M., Azam, M. A., Amin, Y., & Naeem, U. (2020). C2fhar: Coarse-to-fine human activity recognition with behavioral context modeling using smart inertial sensors. *IEEE Access*, *8*, 7731–7747.

Fang, H., & Hu, C. (2014). Recognizing human activity in smart home using deep learning algorithm. In *Proceedings of the 33rd chinese control conference* (pp. 4716–4720).

Figo, D., Diniz, P. C., Ferreira, D. R., & Cardoso, J. M. (2010). Preprocessing techniques for context recognition from accelerometer data. *Personal and Ubiquitous Computing*, *14*(7), 645–662.

Godino-Llorente, J. I., & Gomez-Vilda, P. (2004). Automatic detection of voice impairments by means of short-term cepstral parameters and neural network based detectors. *IEEE Transactions on Biomedical Engineering*, *51*(2), 380–384.

Guan, Y., & Plötz, T. (2017). Ensembles of deep lstm learners for activity recognition using wearables. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *1*(2), 1–28.

Gudise, V. G., & Venayagamoorthy, G. K. (2003). Comparison of particle swarm optimization and backpropagation as training algorithms for neural networks. In *Proceedings of the 2003 ieee swarm intelligence symposium. sis'03 (cat. no.03ex706)* (p. 110-117).

Guo, Q., Liu, B., & Chen, C. W. (2016). A two-layer and multi-strategy framework for human activity recognition using smartphone. In *2016 ieee international conference on communications (icc)* (pp. 1–6).

Ha, S., Yun, J.-M., & Choi, S. (2015). Multi-modal convolutional neural networks for activity recognition. In *2015 ieee international conference on systems, man, and cybernetics* (pp. 3017–3022).

Hammerla, N. Y., Halloran, S., & Plötz, T. (2016). Deep, convolutional, and recurrent models for human activity recognition using wearables. *arXiv preprint arXiv:1604.08880*.

Hannink, J., Kautz, T., Pasluosta, C. F., Gaßmann, K.-G., Klucken, J., & Eskofier, B. M. (2016). Sensor-based gait parameter extraction with deep convolutional neural networks. *IEEE journal of biomedical and health informatics*, *21*(1), 85–93.

Hassan, M. M., Uddin, M. Z., Mohamed, A., & Almogren, A. (2018). A robust human activity recognition system using smartphone sensors and deep learning. *Future Generation Computer Systems*, *81*, 307–313.

Hayashi, T., Nishida, M., Kitaoka, N., & Takeda, K. (2015). Daily activity recognition based on dnn using environmental sound and acceleration signals. In *2015 23rd european signal processing conference (eusipco)* (pp. 2306–2310).

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the ieee international conference on computer vision* (pp. 1026–1034).

Hong, J., & Ohtsuki, T. (2011). A state classification method based on space-time signal processing using svm for wireless monitoring systems. In *2011 ieee 22nd international symposium on personal, indoor and mobile radio communications* (pp. 2229–2233).

Hu, W., & Yen, G. G. (2013). Adaptive multiobjective particle swarm optimization based on parallel cell coordinate system. *IEEE Transactions on Evolutionary Computation*, *19*(1), 1–18.

Hwang, M. G., Park, H. J., & Har, D. H. (2019). A method of smart phone original video identification by using unique compression ratio pattern. *Forensic science international*, *304*, 109889.

Iandola, F., Moskewicz, M., Karayev, S., Girshick, R., Darrell, T., & Keutzer, K. (2014). Densenet: Implementing efficient convnet descriptor pyramids. *arXiv preprint arXiv:1404.1869*.

Inoue, M., Inoue, S., & Nishida, T. (2018). Deep recurrent neural network for mobile human activity recognition with high throughput. *Artificial Life and Robotics*, *23*(2), 173–185.

Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*.

Irvine, N., Nugent, C., Zhang, S., Wang, H., & Ng, W. W. (2020). Neural network ensembles for sensor-based human activity recognition within smart environments. *Sensors*, *20*(1), 216.

Jiang, W., & Yin, Z. (2015). Human activity recognition using wearable sensors by deep convolutional neural networks. In *Proceedings of the 23rd acm international conference on multimedia* (pp. 1307–1310).

Junior, F. E. F., & Yen, G. G. (2019). Particle swarm optimization of deep neural networks architectures for image classification. *Swarm and Evolutionary Computation*, *49*, 62–74.

Kang, W., Zhang, Y., & Dong, X. (2019). Body gesture recognition based on polarimetric micro-doppler signature and using deep convolutional neural network. *Progress In Electromagnetics Research*, *79*, 71–80.

Kennedy, J. (2006). Swarm intelligence. In *Handbook of nature-inspired and innovative computing* (pp. 187–219). Springer.

Khan, A. M., Siddiqi, M. H., & Lee, S.-W. (2013). Exploratory data analysis of acceleration signals to select light-weight and accurate features for real-time activity recognition on smartphones. *Sensors*, *13*(10), 13099–13122.

Khusainov, R., Azzi, D., Achumba, I. E., & Bersch, S. D. (2013). Real-time human ambulation, activity, and physiological monitoring: Taxonomy of issues, techniques, applications, challenges and limitations. *Sensors*, *13*(10), 12852–12902.

Kim, Y., & Toomajian, B. (2016). Hand gesture recognition using micro-doppler signatures with convolutional neural network. *IEEE Access*, *4*, 7125–7130.

Kose, M., Incel, O. D., & Ersoy, C. (2012). Online human activity recognition on smart phones. In *Workshop on mobile sensing: From smartphones and wearables to big data* (Vol. 16, pp. 11–15).

Kwapisz, J. R., Weiss, G. M., & Moore, S. A. (2011). Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter*, *12*(2), 74–82.

Lane, N. D., Georgiev, P., & Qendro, L. (2015). Deepear: robust smartphone audio sensing in unconstrained acoustic environments using deep learning. In *Proceedings of the 2015 acm international joint conference on pervasive and ubiquitous computing* (pp. 283–294).

Lara, O. D., & Labrador, M. A. (2012). A survey on human activity recognition using wearable sensors. *IEEE communications surveys & tutorials*, *15*(3), 1192–1209.

Li, Q., Ning, H., Zhu, T., Cui, S., & Chen, L. (2019). A hybrid approach to inferring the internet of things for complex activity recognition. *EURASIP Journal on Wireless Communications and Networking*, *2019*(1), 1–11.

Li, X., Zhang, Y., Marsic, I., Sarcevic, A., & Burd, R. S. (2016). Deep learning for rfid-based activity recognition. In *Proceedings of the 14th acm conference on embedded network sensor systems cd-rom* (pp. 164–175).

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing*, *234*, 11–26.

Liu, Z., Wu, M., Zhu, K., & Zhang, L. (2016). Sensafe: A smartphone-based traffic safety framework by sensing vehicle and pedestrian behaviors. *Mobile Information Systems*, *2016*.

Longstaff, B., Reddy, S., & Estrin, D. (2010). Improving activity classification for health applications on mobile devices using active and semi-supervised learning. In *2010 4th international conference on pervasive computing technologies for healthcare* (pp. 1–7).

Lu, H., Pan, W., Lane, N. D., Choudhury, T., & Campbell, A. T. (2009). Soundsense: scalable sound sensing for people-centric applications on mobile phones. In *Proceedings of the 7th international conference on mobile systems, applications, and services* (pp. 165–178).

Mannini, A., Rosenberger, M., Sabatini, A., & Intille, S. (2017, 04). Activity recognition in youth using single accelerometer placed at wrist or ankle. *Medicine and science in sports and exercise*, *49*, 801-812. doi: 10.1249/MSS.0000000000001144

Mihanpour, A., Rashti, M. J., & Alavi, S. E. (2020). Human action recognition in video using db-lstm and resnet. In *2020 6th international conference on web research (icwr)* (pp. 133–138).

Najafabadi, M., Villanustre, F., Khoshgoftaar, T., Seliya, N., Wald, R., & Muharemagic, E. (2016, 09). Deep learning techniques in big data analytics. In (p. 133-156). doi: 10.1007/978-3-319-44550-2_5

Nawaratne, R., Alahakoon, D., De Silva, D., Kumara, H., & Yu, X. (2019). Hierarchical two-stream growing self-organizing maps with transience for human activity recognition. *IEEE Transactions on Industrial Informatics*.

Ojha, V. K., Abraham, A., & Snášel, V. (2017). Metaheuristic design of feedforward neural networks: A review of two decades of research. *Engineering Applications of Artificial Intelligence*, *60*, 97–116.

Okeyo, G., Chen, L., & Wang, H. (2014). Combining ontological and temporal formalisms for composite activity modelling and recognition in smart homes. *Future Generation Computer Systems*, *39*, 29–43.

Oliver, N. M., Rosario, B., & Pentland, A. P. (2000). A bayesian computer vision system for modeling human interactions. *IEEE transactions on pattern analysis and machine intelligence*, *22*(8), 831–843.

Parkka, J., Ermes, M., Korpipaa, P., Mantyjarvi, J., Peltola, J., & Korhonen, I. (2006). Activity classification using realistic data from wearable sensors. *IEEE Transactions on information technology in biomedicine*, *10*(1), 119–128.

Patel, A., & Shah, J. (2019). Sensor-based activity recognition in the context of ambient assisted living systems: A review. *Journal of Ambient Intelligence and Smart Environments*, *11*(4), 301–322.

Piwek, L., Ellis, D. A., Andrews, S., & Joinson, A. (2016). The rise of consumer health wearables: promises and barriers. *PLoS medicine*, *13*(2), e1001953.

Plötz, T., Hammerla, N., & Olivier, P. (2011b, December). Feature learning for activity recognition in ubiquitous computing..

Plötz, T., Hammerla, N. Y., & Olivier, P. L. (2011a). Feature learning for activity recognition in ubiquitous computing. In *Twenty-second international joint conference on artificial intelligence.*

Prati, A., Shan, C., & Wang, K. I.-K. (2019). Sensors, vision and networks: From video surveillance to activity recognition and health monitoring. *Journal of Ambient Intelligence and Smart Environments*, *11*(1), 5–22.

Rashidi, P., Cook, D. J., Holder, L. B., & Schmitter-Edgecombe, M. (2010). Discovering activities to recognize and track in a smart environment. *IEEE transactions on knowledge and data engineering*, *23*(4), 527–539.

Ravi, D., Wong, C., Lo, B., & Yang, G.-Z. (2016). Deep learning for human activity recognition: A resource efficient implementation on low-power devices. In *2016 ieee 13th international conference on wearable and implantable body sensor networks (bsn)* (pp. 71–76).

Ravi, N., Dandekar, N., Mysore, P., & Littman, M. L. (2005). Activity recognition from accelerometer data. In *Aaai* (Vol. 5, pp. 1541–1546).

Sahu, A., Panigrahi, S. K., & Pattnaik, S. (2012). Fast convergence particle swarm optimization for functions optimization. *Procedia Technology*, *4*, 319–324.

Sarkar, J., Lee, Y.-K., Lee, S., et al. (2011). Gpars: A general-purpose activity recognition system. *Applied Intelligence*, *35*(2), 242–259.

Shaafi, A., Salem, O., & Mehaoua, A. (2020). Improving human activity recognition algorithms using wireless body sensors and svm. In *2020 international wireless communications and mobile computing (iwcmc)* (pp. 607–612).

Shin, C., Kim, H., Kang, C., Jang, Y., Choi, A., & Woo, W. (2010). Unified context-aware augmented reality application framework for user-driven tour guides. In *2010 international symposium on ubiquitous virtual reality* (pp. 52–55).

Singh, M. S., Pondenkandath, V., Zhou, B., Lukowicz, P., & Liwickit, M. (2017). Transforming sensor data to the image domain for deep learning—an application to footstep detection. In *2017 international joint conference on neural networks (ijcnn)* (pp. 2665–2672).

Subasi, A., Fllatah, A., Alzobidi, K., Brahimi, T., & Sarirete, A. (2019). Smartphone-based human activity recognition using bagging and boosting. *Procedia Computer Science*, *163*, 54–61.

Tang, Y., Teng, Q., Zhang, L., Min, F., & He, J. (2020). Efficient convolutional neural networks with smaller filters for human activity recognition using wearable sensors. *arXiv preprint arXiv:2005.03948*.

Tolstikov, A., Hong, X., Biswas, J., Nugent, C., Chen, L., & Parente, G. (2011). Comparison of fusion methods based on dst and dbn in human activity recognition. *Journal of Control Theory and Applications*, *9*(1), 18–27.

Turaga, P., Chellappa, R., Subrahmanian, V. S., & Udrea, O. (2008). Machine recognition of human activities: A survey. *IEEE Transactions on Circuits and Systems for Video technology*, *18*(11), 1473–1488.

Van Kasteren, T., Englebienne, G., & Kröse, B. J. (2010). An activity monitoring system for elderly care using generative and discriminative models. *Personal and ubiquitous computing*, *14*(6), 489–498.

Venugopalan, S., Xu, H., Donahue, J., Rohrbach, M., Mooney, R., & Saenko, K. (2014). Translating videos to natural language using deep recurrent neural networks. *arXiv preprint arXiv:1412.4729*.

Vepakomma, P., De, D., Das, S. K., & Bhansali, S. (2015). A-wristocracy: Deep learning on wrist-worn sensing for recognition of user complex activities. In *2015 ieee 12th international conference on wearable and implantable body sensor networks (bsn)* (pp. 1–6).

Voicu, R.-A., Dobre, C., Bajenaru, L., & Ciobanu, R.-I. (2019). Human physical activity recognition using smartphone sensors. *Sensors*, *19*(3), 458.

Walse, K. H., Dharaskar, R. V., & Thakare, V. M. (2016). Pca based optimal ann classifiers for human activity recognition using mobile sensors data. In *Proceedings of first international conference on information and communication technology for intelligent systems: Volume 1* (pp. 429–436).

Wang, A., Chen, G., Shang, C., Zhang, M., & Liu, L. (2016). Human activity recognition in a smart home environment with stacked denoising autoencoders. In *International conference on web-age information management* (pp. 29–40).

Wang, J., Chen, Y., Hao, S., Peng, X., & Hu, L. (2019). Deep learning for sensor-based activity recognition: A survey. *Pattern Recognition Letters*, *119*, 3–11.

Wang, L., Qiao, Y., & Tang, X. (2015). Action recognition with trajectory-pooled deep-convolutional descriptors. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 4305–4314).

Wang, Y., Cang, S., & Yu, H. (2019). A survey on wearable sensor modality centred human activity recognition in health care. *Expert Systems with Applications*, *137*, 167–190.

Yacoob, Y., & Black, M. J. (1999). Parameterized modeling and recognition of activities. *Computer Vision and Image Understanding*, *73*(2), 232–247.

Yang, J. (2009). Toward physical activity diary: motion recognition using simple acceleration features with mobile phones. In *Proceedings of the 1st international workshop on interactive multimedia for consumer electronics* (pp. 1–10).

Yang, J., Lee, J., & Choi, J. (2011). Activity recognition based on rfid object usage for smart mobile devices. *Journal of Computer Science and Technology*, *26*(2), 239–246.

Yang, J., Nguyen, M. N., San, P. P., Li, X., & Krishnaswamy, S. (2015). Deep convolutional neural networks on multichannel time series for human activity recognition. In *Ijcai* (Vol. 15, pp. 3995–4001).

Yin, J., Yang, Q., & Pan, J. J. (2008). Sensor-based abnormal human-activity detection. *IEEE Transactions on Knowledge and Data Engineering*, *20*(8), 1082–1090.

Yu, H., Cang, S., & Wang, Y. (2016). 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)* (pp. 250–257).

Zeng, M., Nguyen, L. T., Yu, B., Mengshoel, O. J., Zhu, J., Wu, P., & Zhang, J. (2014). Convolutional neural networks for human activity recognition using mobile sensors. In *6th international conference on mobile computing, applications and services* (pp. 197–205).

Zhang, Y., Xu, G., Wang, Y., Liang, X., Wang, L., & Huang, T. (2019). Empower event detection with bi-directional neural language model. *Knowledge-Based Systems*, *167*, 87–97.