# Human Activity Recognition

# &

# Explainable AI

By: Angad Partap Singh

M00912257

Supervised by:

Professor Juan Carlos Augusto

Department of Science

Middlesex University

January 2024

# Acknowledgements

I would like to express my heartfelt gratitude to my dissertation supervisor, **Professor Juan Carlos Augusto**, whose unwavering support, and guidance have been instrumental in shaping the trajectory of this research. Professor Augusto's keen insights, wealth of knowledge, and dedication to academic excellence have not only enriched the content of this dissertation but also inspired a passion for the subject matter.

The innovative idea of merging Human Activity Recognition (HAR) and Explainable AI (XAI) to create a more comprehensive approach is entirely attributed to Professor Augusto. His vision and mentorship empowered me to explore novel avenues in the field, pushing the boundaries of conventional research and making this dissertation stand out among others. A special acknowledgment also goes to Prof. Xiaohong Gao, for constant updates on dissertation related information.

I am deeply indebted to my colleagues, especially Tushar, Mustafa, Shahid, and Abhishek, whose unwavering support and motivation made this academic journey both enjoyable and enriching. Their collaborative spirit, insightful discussions, and willingness to lend a helping hand were invaluable, especially during moments of challenge and uncertainty.

I extend my sincere appreciation to the Middlesex University, for providing a conducive research environment and access to resources that facilitated the completion of this dissertation.

I am utterly thankful to my parents for their distant motivation and constant support from India. Also, I am immensely grateful to my wife for all her patience, moral support, and tolerance throughout my research.

This research endeavour has been a collective effort, and I am profoundly thankful to everyone who played a role, big or small, in its realization. Your support has been a beacon of encouragement, and I am truly grateful for the enriching experiences and lessons learned throughout this academic endeavor.

# Abstract

Using specialised software and technological devices, **Human Activity Recognition** (HAR) is a rapidly developing field in machine learning that tracks and analyses human activities such as walking and exercising. Many sensor datasets are used in this data-driven process, especially those pertaining to wearable sensors in healthcare applications. With an emphasis on applications in smart healthcare systems, the integration of HAR in healthcare has attracted a lot of research attention. For example, smart wearable sensor-based behaviour recognition systems improve the quality of life for senior citizens by offering timely alerts for possible health risks and real-time monitoring. This dissertation explores the application of Machine Learning (ML) models in comparison to complex DL models, analysing the strengths and weaknesses of their outcomes.

Understanding how artificial intelligence (AI) systems make decisions is crucial in the rapidly changing field of healthcare and other AI-related fields. The complexity of AI systems increases along with the demand for decision-making processes to be transparent. This is where Explainable AI (XAI) techniques come into play, acting as a vital instrument to illuminate the opaque nature of AI models.

The work explores this problem by assessing different models on real-world healthcare datasets. These analyses go beyond simple performance measures to look at these models' internal mechanisms. In particular, I have used state-of-the-art methods such as LIME (Local Interpretable Model-agnostic Explanations) values and SHAP (SHapley Additive exPlanations). These techniques serve as interpreters, revealing the reasoning behind the predictions made by AI models.

My research attempts to shed light on which models are more dependable and trustworthy by contrasting the advantages and disadvantages of these sophisticated models with those of their forebears. Crucially, communicating results in a way that both technical and non-technical audiences can understand. This implies that a larger audience, including lawmakers and healthcare professionals, can readily comprehend and value the insights gained from our work.

The findings signify a noteworthy advancement in the field of smart sensor technology, specifically in the context of the Internet of Things (IoT). We are strengthening the predictability of AI models in crucial healthcare applications and increasing their trustworthiness by making them more explainable. With the promise of a better and more informed future, this development highlights the beneficial effects of our work on the IoT sensor technology and healthcare fields more broadly.

Machine learning (ML) represents a promising tool for the recognition of human activities (HAR), and explainable artificial intelligence (XAI) can elucidate the role of accelerometer features in ML-based HAR models.

The approach has the ability to eliminate the need of complex DL and CNN models, which would in turn reduce the cost and complexity of the models. Furthermore, indulging explainable ai for prediction analysis on local and global level can help getting better and sustainable outcomes.

# List Of Abbreviations

HAR – Human Activity Recognition

XAI – Explainability Artificial Intelligence

AI – Artificial Intelligence

IOT – Internet of Things

ML – Machine Learning

DL – Deep Learning

CNN – Convolutional Neural Networks

XGB – XGBoost

DT – Decision Trees

LR – Logistic Regressor

KNN – K-Nearest Neighbour

LIME - Local Interpretable Model-agnostic Explanations

SHAP – SHapley Additive exPlanations

SVM – Support Vector Machines

GLMs – Generalised Linear Models

LRP - Layer-wise Relevance Propagation

Grad-CAM - Gradient-weighted Class Activation Mapping

# TABLE OF CONTENTS

# 1. Introduction:

Greetings from the exciting nexus of technology and daily existence! We travel through the fields of Explainable AI (XAI) and Human Activity Recognition (HAR) in this dissertation. Imagine a magical dataset that amazingly captures the rhythms of everyday life, such as the deliberate upstairs walking, the brisk jogging beat, and the steady walking pace. How is this recorded? Because of clever sensors that are securely incorporated into wearables and make use of accelerometers and gyroscopes, routine tasks can now be transformed into a rich tapestry of human motion.

The goal of this research is to analyse the performance of sophisticated Machine Learning (ML) models versus more complex (DL) ones. Have you ever wondered why computers decide what they do? The answer to solving the puzzles surrounding these choices is Explainable AI (XAI), which simplifies the complexity of artificial intelligence.

The objective is not only technological advancement but also adding to the larger context of smart devices and the Internet of Things (IoT) as we explore the results and contrast them with current models. This endeavour is evidence of the convergence of technology, data, and applications with a human focus. It's about ensuring that devices' inner workings are meaningful and understandable to a wider audience, in addition to just making them smarter.

The relevance of this work goes beyond technical jargon and into real-world applications, especially in the broad field of healthcare. The dataset becomes more than just a set of data points; it becomes a tool for comprehending and forecasting health-related events. One concrete illustration of how these technological developments can improve and directly impact people's lives is the development of remote patient monitoring.

So, buckle up for this tech adventure, where everyday activities meet cutting-edge machines, and we strive to make technology more understandable and beneficial for everyone involved.

# 2. Research Questions

As mentioned in the introduction above, the purpose of this research question is to explore how XAI can be used to address the black box nature of complex models in HAR. It looks at how XAI methods can improve the interpretability and reliability of black box models, which could lead to a rise in their use and adoption in practical applications.

**Research Question 1:**

How much do Machine Learning (ML) models perform better in terms of accuracy, interpretability, and computational efficiency than Deep Learning (DL) models in Human Activity Recognition (HAR) tasks?

**Research Question 2:**

What effect does the addition of eXplainable Artificial Intelligence (XAI) techniques have on model adoption and usability, and how does it improve the interpretability and reliability of black box models in Human Activity Recognition (HAR)?

**Research Question 3:**

How is the research helpful in gaining important insights from different models, such that it same procedure can be incorporated into other Domains (like Healthcare and Smart Homes)?

## Rationale behind the decision of using ML Models and not DL Models

The reason of using ML models and not DL models specifically lies behind the complexity of the framework.

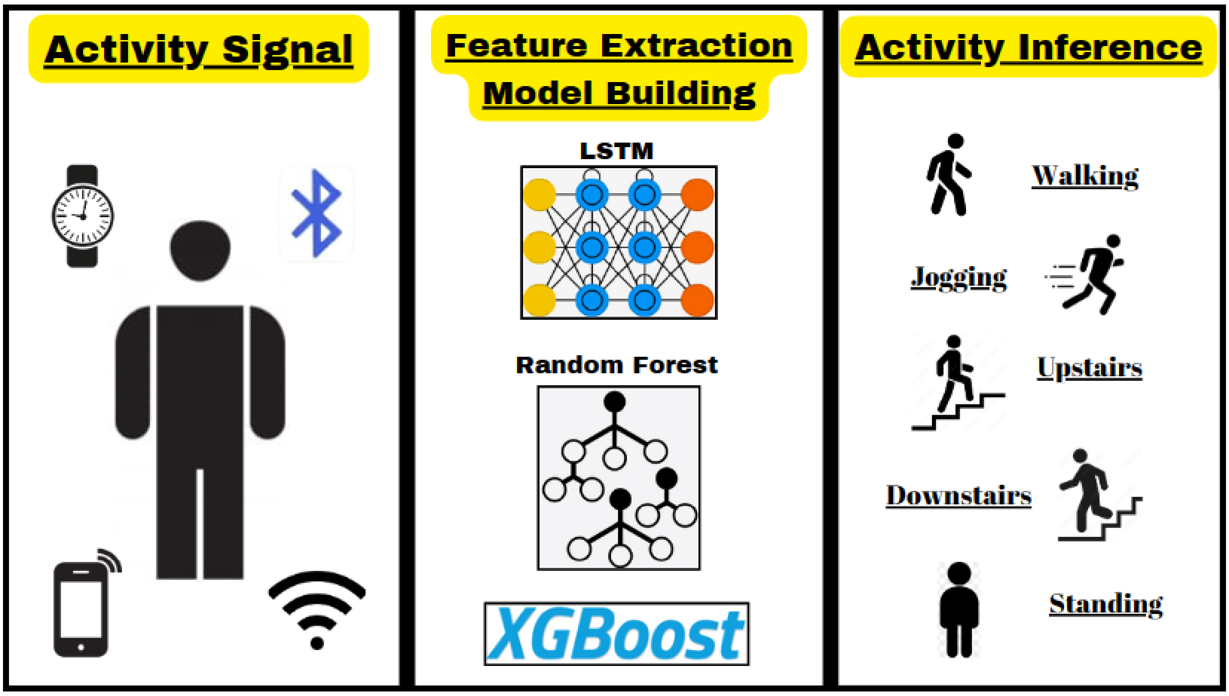
1. Complexity and Resource Intensive
   Deep learning models, in particular, are very complicated and frequently need a lot of processing power (GPUs or GPUs) for training and inference. This may pose a constraint for individuals or entities with restricted computational capacity.
2. Training Time
   It can take a while to train deep neural networks, particularly for large models and datasets. The development and experimentation process may be slowed down by this prolonged training period.
3. Interpretability
   DL models are sometimes referred to as "black boxes" since it can be difficult to comprehend how they make their judgements. In important applications where

accountability and trust are crucial, this lack of interpretability may present challenges.

4. Limited Dat Efficiency
   For small to medium-sized datasets, DL models might not be the best option because they might not provide any discernible benefits over more straightforward ML models.
5. Dependencies
   Since DL models frequently rely on particular libraries and frameworks, dependencies and compatibility problems may arise that must be handled.
6. Energy Consumption
   Deep neural network training can be energy-intensive, which may be problematic for applications where there are environmental or energy-related restrictions.

# 3. Human Activity Recognition

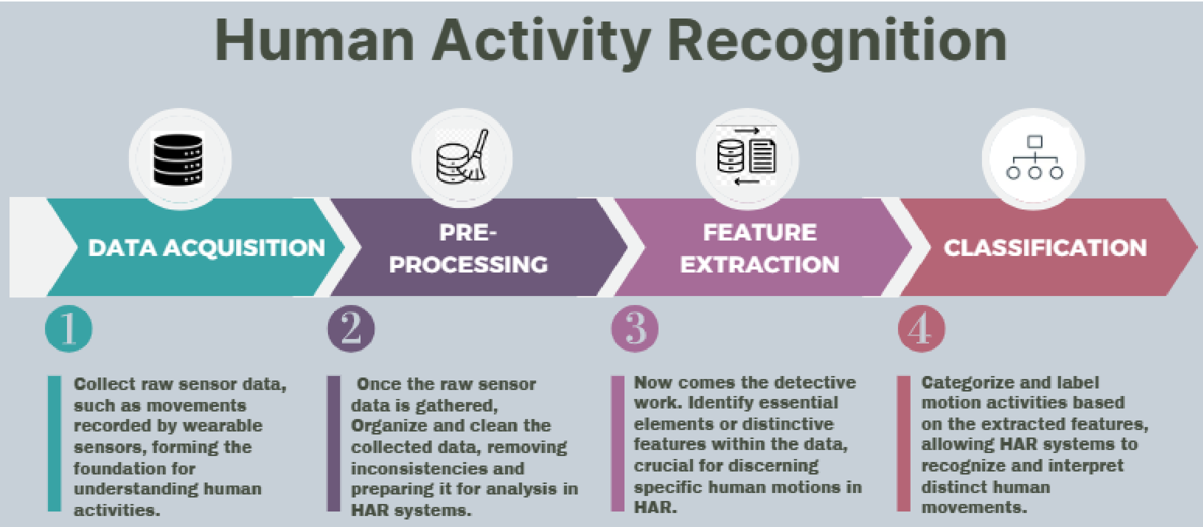
In the exciting field of human activity recognition (HAR), technology and human movement are combined to create systems that can automatically comprehend and classify actions based on sensor data. This area of computational science and engineering explores how to make technology understand human behaviour, leading to creative applications in a variety of fields.



Think of HAR as a computer science and engineering wizardry that teaches machines to comprehend and classify human behaviour. The complex dance between the interpretation of human activities and the raw sensor data is the beating heart of HAR. The beginning point is this raw sensor data, which captures the minute details of human motion, such as actions, gestures, or behaviours that are captured by sensors, the silent watchers of our movements.

**The journey of Human Activity Recognition unfolds in four key stages:**

- Data Acquisition: At the starting line, we gather the raw sensor data. Think of this stage as collecting the dance steps or workout routines of individuals, recorded by wearable sensors or other devices.
- Pre-processing: Before computers can make sense of the data, it needs a bit of cleaning and organizing. Pre-processing is like getting the dance floor ready for a performance – removing any unwanted noise or inconsistencies.
- Feature Extraction: Just like a choreographer highlighting the key moves in a dance routine, feature extraction involves picking out the essential bits from the data. These features are the distinctive elements that help computers understand the specific actions being performed.
- Classification: This is the grand finale where computers figure out and categorize the specific motion activities. It's like the computer saying, "Ah, this is a dance move," or "This looks like a morning jog."




Let's now discuss the two primary types of HAR systems. In the first kind, people wear sensors on their bodies, which function like high-tech stickers and record every movement, in order to identify activities based on sensor data. By providing data to the HAR system, these wearable sensors take on the role of storytellers, helping it to understand and interpret human behaviour.

HAR is essentially about translating human motion language into a computer-understandable code. This creates a plethora of opportunities across multiple industries, such as healthcare, where it can track the movements of patients, sports, where it can evaluate athletes' performance, and interactive technologies, which can produce more captivating user interfaces.

### 3.1 HAR Models are Black-Box Models

"Black box models" refers to machine learning models that are extremely intricate and challenging for humans to decipher or comprehend. These models are notorious for being opaque, making it difficult to understand how they make their judgements or predictions. Reasons why certain models are considered Black-box models:

**Complex architecture** makes it difficult to understand how deep neural networks, ensemble models, and other complex algorithms process input data because of their millions of parameters and intricate internal structures. Some models automatically carry out feature engineering, generating new features or feature combinations that are not immediately comprehensible to humans.

**Non-Linear Transformations** make a lot of black box models alter input features in non-linear ways, which makes it challenging to track how adjustments to one feature affect predictions.

**High Dimensionality** becomes a challenge as intricate patterns that are difficult to understand intuitively may be captured by models that have been trained on high-dimensional data, such as text or images.

**Interactions** are troublesome as it can be difficult to comprehend how intricate relationships between features impact predictions, even in models that attempt to capture such relationships.

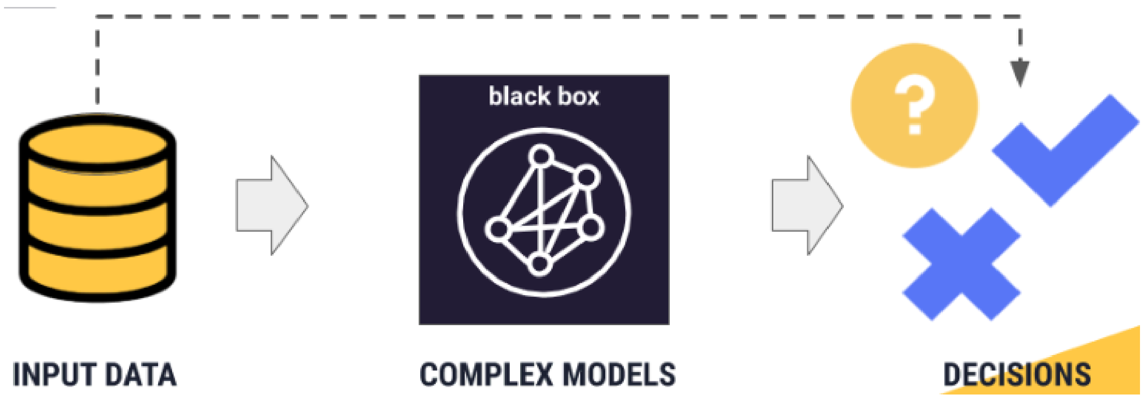
**Internal Representations** are problematic in particular, neural networks are able to learn internal representations of data that might not be comparable to ideas that are understandable by humans.

# 4. Explainable Artificial Intelligence

## (A remedy to Black-box issue)

Artificial intelligence (AI) has advanced significantly in the last several years, showing impressive results across a range of industries, including finance, healthcare, and natural language processing and computer vision. However, the need for interpretability and transparency has increased significantly as AI systems become more complex and common. This need led to the development of Explainable AI (XAI), a crucial field that aims to improve the interpretability, understandability, and human accountability of AI models and their judgements. We will examine Explainable AI's principles, practices, significance, and applications in this extensive guide.

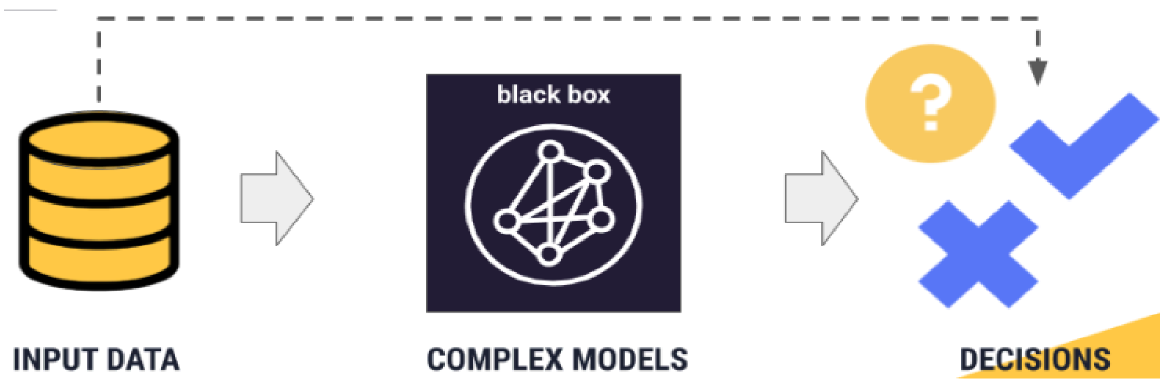
**BLACK-BOX** refers to A machine-learning model that is challenging to comprehend or evaluate is known as a "black-box model" [2]. Black-box models are trained on massive amounts of data and are capable of making judgments or forecasts; however, understanding how they arrived at those conclusions or predictions is either very difficult or not feasible at all. Black-box models are widely employed when the objective is to achieve high accuracy or performance but the specifics of the decision-making process are not crucial. Understanding a model's decision-making process is crucial in some applications, especially when the decisions have a significant impact on people. In certain situations, it may be preferable to use more interpretable models rather than black-box models.



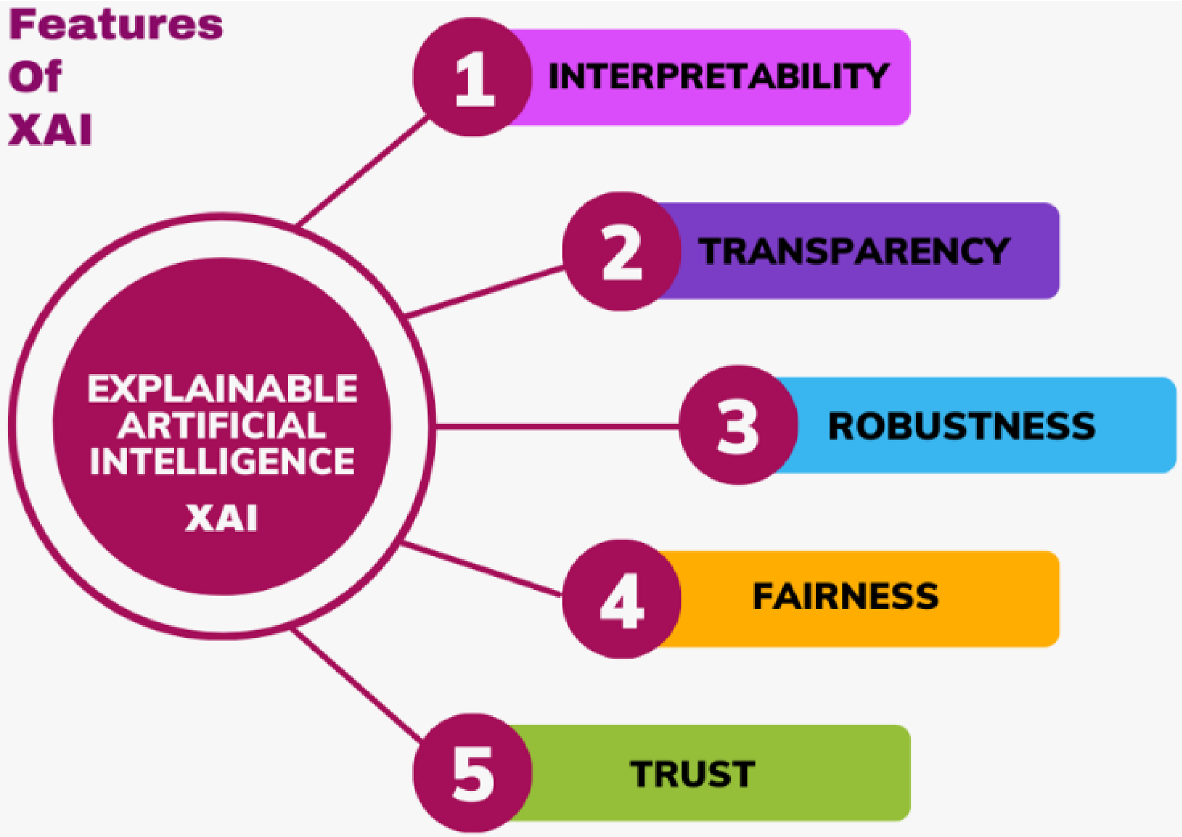

WHITE-BOX A machine-learning model that is visible and interpretable, or "white-box" model, makes it easy to understand how the model makes decisions or predictions.

Simple, understandable models like decision trees or linear regression are frequently used in white-box models [3]. These models simplify the process of understanding how the model determines its decisions. When understanding a model's decision-making process is crucial, as when the model's results will have a significant impact on actual people, white-box models are typically chosen.

Conversely, black-box models may outperform white-box models in some situations since they usually use more complex, less interpretable models to achieve more accuracy or performance.




The goal of the multidisciplinary field of explainable AI (XAI) is to improve the interpretability and transparency of AI systems so that people can comprehend and have faith in the choices that AI algorithms make. Even though artificial intelligence (AI) has shown remarkable promise, its internal workings are frequently seen as "black boxes," making it difficult to explain how and why AI models reach particular conclusions or predictions. In order to help users better understand AI outputs, XAI aims to address this challenge by offering insights into the decision-making processes used by AI models.




Main advantages of XAI include:

- Easy understanding of Models deployed.
- Precision in bug/defect tracing.
- Logical explanation of algorithms.
- Better control on the work.
- Clear Accountability and Auditability.

## Intersection of HAR & XAI

HAR uses a variety of data sources, such as sensors, wearables, or video streams, to recognize, track, and understand human activities or behaviours. The objective is to identify distinct actions or gestures made by people, with potential uses in smart settings, sports analysis, healthcare monitoring, and other fields.

**HAR**

HAR uses a variety of data sources, such as sensors, wearables, or video streams, to recognize, track, and understand human activities or behaviours. The objective is to identify distinct actions or gestures made by people, with potential uses in smart settings, sports analysis, healthcare monitoring, and other fields.

**XAI**

The goal of explainable AI is to improve the transparency and human-interpretability of AI systems. It involves methods and approaches that try to explain how AI models make decisions, particularly those that use sophisticated algorithms like deep learning, which are sometimes referred to as "black-box" models.

## 4.1 Importance of XAI

XAI is significant in a wide range of fields and sectors. Here are a few main justifications for why XAI matters:

- Trust and Accountability: Trust is crucial in applications where artificial intelligence (AI) affects important decisions, like healthcare, finance, and driverless cars. By enabling stakeholders to comprehend AI decisions and hold AI systems responsible for their actions, XAI contributes to the development of trust.
- Compliance and Regulations: Decision-making must be transparent in order to comply with regulations that apply to many industries. XAI assists businesses in adhering to these rules and avoiding moral and legal dilemmas.
- Error Detection and Debugging: By exposing biases, mistakes, and unexpected outcomes in AI models, XAI techniques help developers find and fix problems.
- User-Centric Design: In order for users to make wise decisions, they must comprehend the outputs of AI models. XAI makes user-centric design easier by offering end-user-meaningful explanations.
- Robustness and Safety: XAI can assist in identifying scenarios in which AI systems may fail, enhancing overall safety in safety-critical applications like autonomous vehicles.

## 4.2 Essential Ideas in XAI

It's important to understand a few basic ideas in order to have a deeper understanding of XAI:

- Interpretability vs. Explainability: Interpretability is the degree to which a model's predictions are easily understood by humans, whereas explainability is the process of giving those predictions a reason or explanation. XAI includes both of these.
- Local vs. Global Explanations: Local explanations concentrate on providing a reason for a particular prediction, whereas global explanations offer an understanding of a model's general behaviour.

- Model-Agnostic vs. Model-Specific Techniques: XAI techniques can be classified as either model-specific, meant for a specific class of models, or model-agnostic, meaning they can be applied to a broad range of AI models.
- Feature Importance vs. Model Behaviour: XAI can clarify the decision-making process and behaviour of the model, or it can reveal feature importance, indicating which input features contribute most to a prediction.

## Importance Sources for Features

Methods such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) rate the significance of input features to help users comprehend how they affect forecasts.

- Rule-Based Models: Using explicit rules to guide predictions, rule-based techniques—like decision trees and rule lists—improve transparency.
- Surrogate Models: Models that approximate the behaviour of complex models, such as decision trees or linear regressions, are simpler models that facilitate easier interpretation.
- Attention Mechanisms: Deep learning models' attention mechanisms highlight significant input regions and provide information about the specific portions of an image or text that impacted the model's conclusion.
- Counterfactual Explanations: These help users comprehend how sensitive the model is to variations in input by giving examples of inputs that, if altered, would lead to different model predictions.
- Visualisation Tools: Users can explore and interpret model outputs with the aid of tools such as saliency maps, feature importance plots, and activation maximisation visualisations.

## 4.3 ML Models That Are Interpretable

Some machine learning models are commonly utilised for XAI and are interpretable by nature.

- Linear Models: For input features, linear regression, logistic regression, and linear SVMs offer easily understood coefficients.
- Decision Trees: Decision trees are simple to understand and intuitive because they are made up of a sequence of binary decisions.
- Generalised Linear Models (GLMs): GLMs are useful for interpretable modelling because of linear regression to a variety of distributions.

**Explainability Techniques for Post-hoc**

After a model has been trained, post-hoc explainability techniques are used, and they consist of the following:

- Layer-wise Relevance Propagation (LRP): LRP uses relevance scores to re-distribute a model's decision back to its input features.
- Gradient-based Techniques: In order to highlight input regions and determine the significance of a feature, gradients with respect to input features can be analysed.
- Activation Maximisation: In order to reveal the characteristics that a neuron detects, activation maximisation techniques look for input patterns that maximise the activation of particular neurons.

## 4.4 Putting Model Interpretations into Visual Form

A potent tool in XAI is visualisation, which offers simple means of comprehending model interpretations:

- Saliency Maps: These visual aids draw attention to the most significant areas of a picture or text input.
- Grad-CAM: The Gradient-weighted Class Activation Mapping technique employs a visual representation of each pixel's significance for a specific class in an image.
- Significance of Feature Plots: Plots that show feature importance scores for each input feature include bar charts and heatmaps.

## 4.5 XAI in Various Fields

XAI is widely applicable across multiple domains:

- Healthcare: XAI aids physicians in comprehending the rationale behind a model's recommended course of action or diagnosis.
- Finance: XAI is used to explain loan approvals or transaction classifications in credit scoring and fraud detection.
- Autonomous Vehicles: XAI makes sure that these vehicles can justify their driving choices, which improves safety.
- Natural Language Processing: XAI helps explain the reasons behind the text outputs produced by a language model.
- Criminal Justice: To increase transparency in risk assessment instruments and parole determinations, XAI is utilised in the criminal justice system.

## 4.6 XAI's Drawbacks and Limitations

XAI has various restrictions and challenges.

- Trade-offs: Interpretability and model complexity are frequently correlated; more comprehensible models may come at the expense of accuracy.
- Model Complexity: Interpreting deep learning models in their entirety can be difficult when they have millions of parameters.
- Contextual Understanding: It's possible that XAI falls short of fully capturing human contextual understanding.
- User Understanding: It can be challenging to make sure that explanations are clear and intelligible for end users.
- Scalability: It's a constant struggle to create XAI methods that can handle complicated models and big Datasets.

## 4.7 XAI's Ethical Considerations

In XAI, ethical issues are critical.

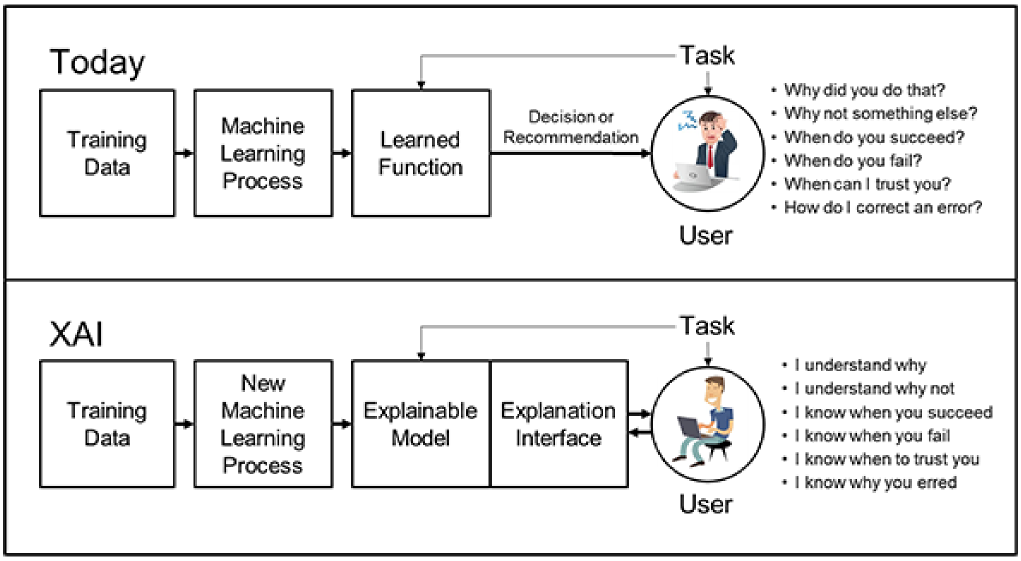
- Bias and Fairness: XAI ensures decision-making is fair by assisting in the identification and mitigation of biases in AI models.
- Privacy: Giving explanations shouldn't violate people's right to privacy or divulge private information.
- Accountability: By enabling the tracking of decisions back to individual algorithms, XAI promotes accountability.
- Transparency: Gaining the trust of users and stakeholders requires transparency in AI decision-making.

## 4.8 Future Directions for Explainable AI

Future trends in the rapidly developing field of XAI include:

- Hybrid Models: To attain accuracy and transparency, hybrid models combine interpretable models with deep learning.
- Standardisation: To guarantee uniformity and best practices, standards and guidelines for XAI are being developed.
- User-Centric Design: XAI explanations are created with end users' needs and cognitive capacities in mind.

- Legal Frameworks: Creating legislation to control accountability and transparency in AI.
- Explainable Neural Networks: Progress towards improving the interpretability of deep neural networks.




**Final Thoughts**

The crucial field of explainable AI (XAI) deals with the interpretability and transparency of AI systems. It is essential for establishing credibility, guaranteeing responsibility, and identifying biases in AI models in a variety of fields. XAI will continue to be essential in helping AI decisions become more morally and comprehensibly as technology develops. The ability of XAI to close the knowledge gap between machine intelligence and human comprehension will be crucial to the future of AI since it will allow AI to cooperate responsibly and accountable with humans.

# 5. State-Of-The-Art & Literature Overview

The increasing interest in applications such as healthcare, sports analytics, security, and human-computer interaction are driving rapid advancements in the state of the art in human activity recognition (HAR). Early approaches using classical generative Machine Learning algorithms have been recognised in activity recognition research as crucial for capturing temporal dependencies in sensor-based data (e.g., Lester et al., 2005).

Deep Learning has become the go-to method for sensor-based Human Activity Recognition (HAR) as it has developed. The standardisation of Recurrent Neural Networks (RNNs) has been a significant change in HAR techniques. Important studies that support this change include Hammerla et al.'s (2016) work, which showed that RNNs are effective in HAR. Furthermore, Long-Short-Term Memory networks (LSTMs) have become popular in this field, as demonstrated by the study carried out in 2016 by Edel and Köppe.

A paradigm shift in HAR has been brought about by Deep Learning, specifically through RNNs and LSTMs, which offer enhanced capabilities in identifying and comprehending human activities. Given the temporal nature of sensor data in HAR applications, these neural network architectures excel at modelling sequential data. The accuracy and resilience of HAR systems have been greatly improved by their capacity to capture dependencies over time.

The development of HAR approaches has progressed in tandem with the swift progress in Deep Learning, where RNNs and LSTMs have become essential instruments in this domain. The HAR landscape has changed as a result of these neural networks, making it possible to recognise human activity with greater accuracy and sophistication. This has numerous applications in fields like security, sports analytics, and healthcare.

## 5.1 Related Work:

We have witnessed a significant advancement in computer learning in the area of human activity recognition, such as walking and running. Deep learning has made it possible for computers to learn what to look for on their own, whereas previously we had to tell them what to look for. (Pouyanfar et al., 2018)

Deep learning has demonstrated remarkable success in a variety of domains, including image recognition, speech recognition, and language processing. Deep learning's ability to learn from raw data means that it doesn't require human intervention to create unique features, which contributes to its success.

Utilising deep learning to identify human activities has garnered a lot of attention. Scholars have been investigating various approaches utilising deep learning for this purpose. The DeepConvLSTM model, which integrates several neural network architectures to improve computer comprehension of tasks, is one key concept. (Edel and Köppe, 2016; Ordóñez and Roggen, 2016)

However, Deep convolutional neural networks (CNNs) and general machine learning models are not the same. General machine learning can be a better option if you have less data or want a simpler model. Additionally, it's simpler to comprehend how these models arrive at decisions. Deep CNNs, on the other hand, can learn from raw data without the need for artificial features, which makes them excellent for tasks like image recognition. However, they can be computationally demanding and require large amounts of data. So, the decision is based on the particular issue at hand as well as the available resources.

These days all the commercially available, wearable sensors are embedded with sensors which are inertial and deliver motion signals performed by a human body. Using information from head-mounted inertial sensors, Hristijan et al. [45] investigated a weighted ensemble learning algorithm to identify eight commonplace activities. Cesareo et al. [57] used an IMU-based system to measure breathing parameters. They were able to accurately perceive the respiratory rate through an automatic method and reconstruct respiration-induced movement with the help of the proposed algorithm.

The aacelerometers these days, usually found in cellphones, smart-watches and other handheld devices, very slight vibrations on the human body can be detected by a high-resolution accelerometer. Because of its ability to detect even the smallest movements or tremors, it can be used to identify delicate physical gestures or activities and to record precise motion data.

Apart from their benefits for wearability, which include low power consumption, compact size, affordability, and wide availability, inertial sensors also deliver excellent data quality in terms of accuracy and sensitivity. These sensors—high-resolution accelerometers in particular—are able to detect even minute vibrations in the human body. Significantly, inertial sensors—including specialty commercial wearables made for motion-related applications—have assumed a central role in Human Activity Recognition (HAR) tasks.

There are, nevertheless, a few pragmatic factors to take into account. It can be inconvenient for users to have inertial sensors physically attached to the particular body part that needs to be monitored, particularly when continuous, long-term motion tracking is needed. For prolonged periods, this attachment requirement could be burdensome and uncomfortable. Additionally, despite the high precision these sensors provide, they have issues with accumulated errors over time, which means that frequent recalibration is required to keep them accurate.

To summarise, inertial sensors are very beneficial for HAR tasks (such as those in wearables sold in stores) because they are wearable and provide high-quality data. However, their dependence on physical attachment and periodic recalibration are drawbacks that should be addressed in order to improve user comfort and long-term accuracy.

## 5.2 Incorporating XAI:

The current state of the art is typically based on cameras. DLs and CNNs have been used since 90s to extract relevant information out of the images. Motion-sensing and hih-accuracy image classifiers are doing well, Unfortunately, camera systems' limited scalability makes them unsuitable for achieving the objective of HAR adoption in mobile computing contexts. We have seen that ML models and XAI have been used for EEG readings, The most significant EEG spectral features in HAR models were clinically interpreted by LIME (Local Interpretable Model-Agnostic Explanations), and several machine learning models were trained for activity recognition using EEG data. The HAR models' classification results, especially those of the Random Forest and Gradient Boosting models, showed exceptional performance in differentiating the examined human activities. The XAI explanations corroborated the ML models' alignment with EEG spectral bands in the identification of human activity.

In summary, explainable Artificial Intelligence (XAI) may enhance motor imagery, accelerometer settings, the healthcare metaverse, and activity monitoring for patients and healthy individuals when applied to Human Activity Recognition (HAR) studies.

The Explain ability is favoured by two main processes:

### 5.3 LIME (Local Interpretable Model-agnostic Explanations)

The LIME (Local Interpretable Model-agnostic Explanations) method generates interpretable and locally faithful explanations to explain machine learning model predictions. LIME serves as a conduit between sophisticated machine learning models and human interpreters, assisting in the deciphering of black-box predictions and enhancing the transparency and reliability of the models.

### 5.3.1 How LIME works:

**Selection of Data Point**: LIME begins by choosing the data point for which you want an explanation in order to provide context for a particular prediction that a machine learning model has produced. The instance you are interested in learning more about is represented by this data point.

**Perturbation**: LIME creates a dataset of the chosen data point's perturbed versions. It accomplishes this by maintaining the target variable constant while making tiny, random adjustments or perturbations to the feature values. A synthetic dataset that replicates the original data distribution is produced by these perturbations.

**Prediction**: Using this artificial dataset, the machine learning model is then used to generate a set of predictions for the perturbed instances.

**Local Surrogate Model:** LIME fits the synthetic data to a locally faithful and interpretable surrogate model. Usually simpler than the original complex model, this surrogate model—such as decision trees or linear regression—is easier to understand.

**Justification**: The surrogate model sheds light on the model's local behaviour around the chosen data point. It provides a comprehensible explanation by quantifying the correlation between feature values and the model's predictions.

**Importance of Features**: Based on how features influence the surrogate model's predictions, LIME rates their importance. These importance scores show how each feature affects the chosen prediction.

Visualisation: Textual explanations, feature importance plots, and partial dependence plots are a few methods to visualise the explanation. Users can better grasp which features affected the model's choice in this particular instance with the aid of these visualisations.

### Key-Points:

Because LIME is model-agnostic, it can be used with any machine learning model without requiring an understanding of how it functions internally. LIME aims to produce explanations that are locally faithful, i.e., they faithfully represent the model's behaviour close to the chosen data point. Because the surrogate model's simplicity makes interpretation simple, LIME is an invaluable tool for comprehending and troubleshooting intricate machine learning models.

## 5.4 SHAP (SHapley Additive exPlanations)

Shapley Explanation is a framework for explaining machine learning model predictions, also known as SHAP (SHapley Additive exPlanations). This method, which assigns "values" to each feature in a model's prediction, is based on ideas from cooperative game theory.

A useful tool for comprehending and analysing machine learning models is Shapley Explanation. It gives users confidence in the behaviour and decision-making of the model by illuminating the significance of individual features and how they work together to influence predictions.

### 5.4.1 Now to understand How SHAP works:

Shapley values are a mathematical concept that are used to distribute the "value" or "contribution" among a group of participants (features in this case) in a cooperative game in a fair manner. Predicting the target variable is the "game" in machine learning, where each feature is viewed as a player.

**All Potential Feature Combinations**: Shapley Explanation takes into account every potential feature combination in order to determine Shapley values for a given prediction. These pairings show various feature coalitions cooperating to generate a prediction.

Shapley Explanation evaluates the contribution of each feature by contrasting the model's prediction with and without the feature for each feature combination. It calculates the addition or subtraction of each feature from the value of the prediction.

**Weighted Average**: Shapley values determine a weighted average of each feature's contributions over all potential feature combinations. The weights take into account all potential combinations and represent the various ways features can work together.

**Interpretation**: Each feature's significance in the prediction is shown by the resulting Shapley values. Higher Shapley values are associated with greater influence, whereas lower values are associated with less influence.

**Key Points:**

- Because Shapley Explanation is model-agnostic, it can be used with any machine learning model without requiring an understanding of the inner workings of the model.
- Shapley values guarantee that the total contribution is equal to the prediction's deviation from the expected value, offering a fair and uniform method of attributing contributions to features.
- Positive or negative Shapley values indicate whether a feature contributes favourably or unfavourably to a prediction.
- Shapley values are important for model transparency and trust because they provide a logical and understandable means of explaining complicated model predictions.

# 6. Methodology

### 6.1 About the Dataset:

The dataset was sourced from publicly available data at the UCI Machine Learning Repository and can also be accessed on Kaggle. This dataset captures human activity through the use of accelerometers, recording various activities performed by 36 different users. These activities encompass ascending and descending stairs, sitting, walking, jogging, and standing, with each user participating for specific durations.

To gather this data, accelerometers were employed to detect the device's orientation and measure acceleration along three distinct dimensions. The sampling rate was set at 20 Hz, equivalent to 20 samples per second, meaning that data points were collected every 50 milliseconds.

**Here's a breakdown of the dataset's fields:**

- **USER**: Identifies the user responsible for data collection, represented by integer values ranging from 1 to 36.
- **ACTIVITY:** Describes the specific activity the user was engaged in, which could be any of the following: Walking, Jogging, Sitting, Standing, Going Upstairs, Going Downstairs.

| Activities | |
|---|---|
| SITTING | STANDING |
| WALKING | JOGGING |
| UPSTAIRS | DOWNSTAIRS |

- **TIMESTAMP**: Indicates the phone's uptime in nanoseconds, offering a time reference for each data point.

- **X – AXIS:** Represents acceleration along the x-direction as measured by the accelerometer in the Android phone. These values are floating-point numbers, typically falling within the range of -20 to 20. Notably, a value of 10 corresponds to 1g, equivalent to 9.81 m/s^2, while a reading of 0 implies no acceleration. It's essential to recognize that the recorded acceleration includes the Earth's gravitational pull, so when the phone is at rest on a flat surface, the vertical axis will register values of approximately +10 or -10, depending on orientation.
- **Y – AXIS:** Functions similarly to the x-axis but records acceleration along the y-axis.
- **Z – AXIS:** Similarly, the z-axis captures acceleration, but along the z-direction.

This dataset offers valuable insights into human activity recognition and motion analysis, making it a valuable resource for research and applications in fields such as wearable technology, health monitoring, and activity tracking.

Here is the snippet of the dataset:

| | user | activity | timestamp | x-axis | y-axis | z-axis |
|---|---|---|---|---|---|---|
| **0** | 1 | Walking | 4.991920e+12 | 0.69 | 10.80 | -2.030000 |
| **1** | 1 | Walking | 4.991970e+12 | 6.85 | 7.44 | -0.500000 |
| **2** | 1 | Walking | 4.992020e+12 | 0.93 | 5.63 | -0.500000 |
| **3** | 1 | Walking | 4.992070e+12 | -2.11 | 5.01 | -0.690000 |
| **4** | 1 | Walking | 4.992120e+12 | -4.59 | 4.29 | -1.950000 |
| ... | ... | ... | ... | ... | ... | ... |
| **441825** | 15 | Downstairs | 2.096770e+12 | -3.95 | 7.06 | -0.340509 |
| **441826** | 15 | Downstairs | 2.096820e+12 | -5.67 | 8.31 | -1.184970 |
| **441827** | 15 | Downstairs | 2.096870e+12 | -5.37 | 8.58 | -1.116869 |
| **441828** | 15 | Downstairs | 2.096920e+12 | -4.75 | 8.73 | -0.885323 |
| **441829** | 15 | Downstairs | 2.096970e+12 | -4.63 | 8.50 | -0.531194 |

441830 rows × 6 columns

This is clear from the snippet above that the dataset consists of 441830 Rows and 6 columns.

Further we can see the information about the dataset by using df.info().

**6.2 Dataset Information: df.info()**

There are 441,830 rows in this dataset, and each row contains details about a recorded activity. The information is arranged in six columns and includes the following: the user's name, the activity type (e.g., walking vs. sitting), the timestamp of each measurement, and the movement measurements along the three axes (x, y, and z). An accelerometer is used to record these measurements and monitors the device's motion. The dataset requires roughly 20.2 megabytes of memory in total. It's an invaluable tool for researching and comprehending human behaviour, particularly with accelerometer-equipped devices.

The Datatype can be seen about all the individual features. So we have a combination of Integer, Object, and Float Data types.

```
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 441830 entries, 0 to 441829
Data columns (total 6 columns):
 #   Column     Non-Null Count   Dtype
---  ------     --------------   -----
 0   user       441830 non-null  int64
 1   activity   441830 non-null  object
 2   timestamp  441830 non-null  float64
 3   x-axis     441830 non-null  float64
 4   y-axis     441830 non-null  float64
 5   z-axis     441830 non-null  float64
dtypes: float64(4), int64(1), object(1)
memory usage: 20.2+ MB
```



Upon closer inspection, we can investigate how frequently each particular activity is carried out by people. Gaining a deeper understanding of the distribution and patterns of different activities within the dataset is made possible by this analysis. We can find trends, evaluate the balance of activities, and spot any possible biases by knowing how frequently each activity occurs. These kinds of insights, which throw light on the subtleties of human behaviour and activity recognition, are essential for making well-informed decisions and deriving significant conclusions from the data.

```
activity
Walking       183587
Jogging       143893
Upstairs       47221
Downstairs     36555
Standing       15669
Sitting        14905
Name: count, dtype: int64
```

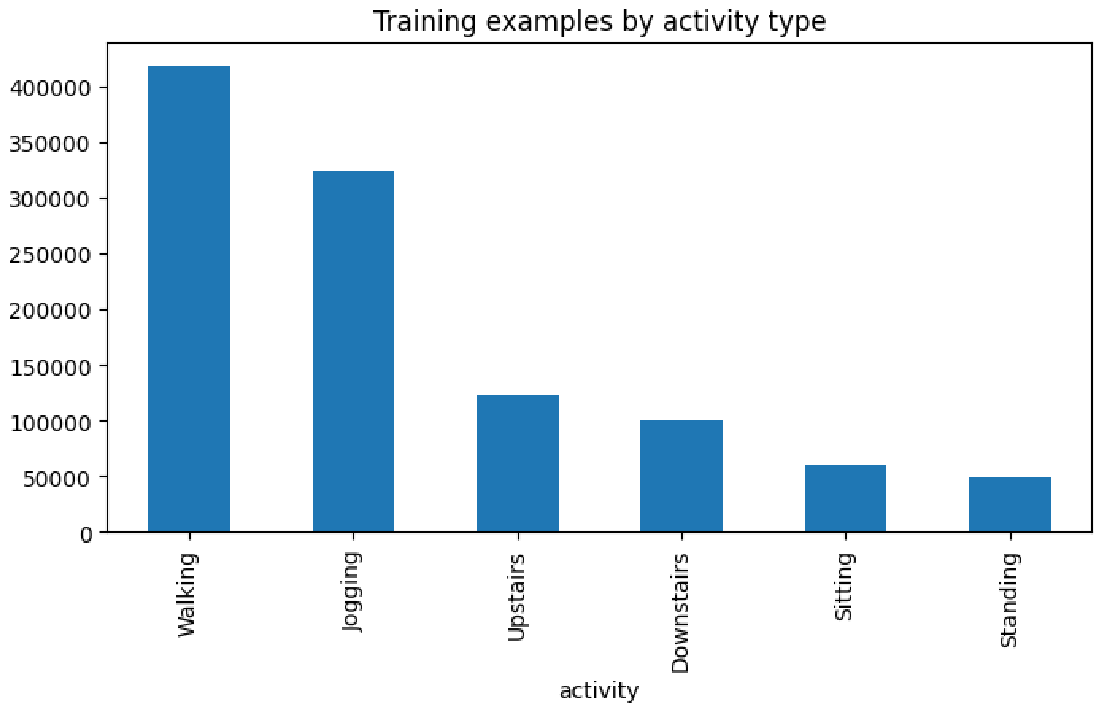


## 7. VALUABLE INSIGHTS about the Dataset

Machine learning insights are critical findings and comprehensions that emerge during the model-building and model-tuning process. These revelations offer vital direction for improving model performance, understanding the underlying data, and arriving at wise decisions. Analysts can pinpoint the model's advantages and disadvantages by closely examining its behaviour. This allows for more focused updates. Insights also highlight the subtleties of the dataset, such as trends, anomalies, and problems with data quality. This increased awareness helps with feature engineering and data preprocessing, which improves model accuracy in the end. Additionally, insights are essential for feature selection, debugging models, handling ethical issues like bias, and guaranteeing interpretability for different stakeholders.

Additionally, they serve as a catalyst for optimisation initiatives and support well-informed business choices, highlighting the profound influence insights have on every stage of the machine learning development cycle, from data preparation to model deployment.

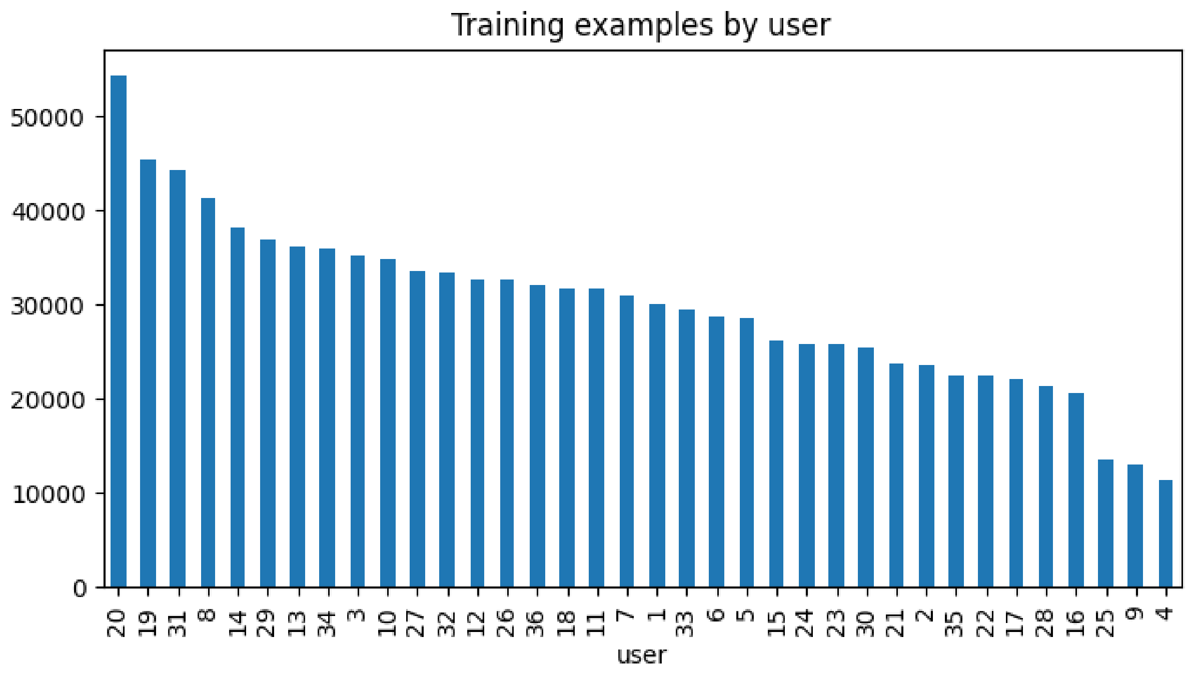
### 7.1 PLOT-1 (Distribution of Data Among Activities)




In the bar plot, the activity types are displayed along the x-axis and the number of training examples is displayed on the y-axis. The height of each bar indicates how many training examples are linked to a particular kind of activity. This visualisation comes in very handy for a variety of situations, such as my data analysis work. I can evaluate the overall quality of my dataset and spot any imbalances by analysing the distribution of data across various classes or categories.

**Key Findings:** The dataset is not balanced. The frequency of “Walking” and “Jogging” is more then other activities. This plot helped me understand whether my dataset was balanced or whether certain activity types were over- or underrepresented as I worked on a machine learning project. These insights were essential because they helped me decide which preprocessing steps to take and how best to optimize the model training process in order to reduce biases and improve model performance. Essentially, this kind of bar plot was an essential component of my exploratory data analysis toolbox since it gave me a clear visual depiction of the distribution of categorical data, which helped me produce machine learning results that were more reliable and accurate.

## 7.2 PLOT 2 – (User-specific Data Distribution)




Using the matplot library module, I produced a bar plot that showed the data as "CountOfActivityPerPerson." I was able to learn more about how training examples are distributed among users thanks to this plot. The height of each bar on the plot indicated the number of training examples linked to that specific user, and each bar represented a different user.

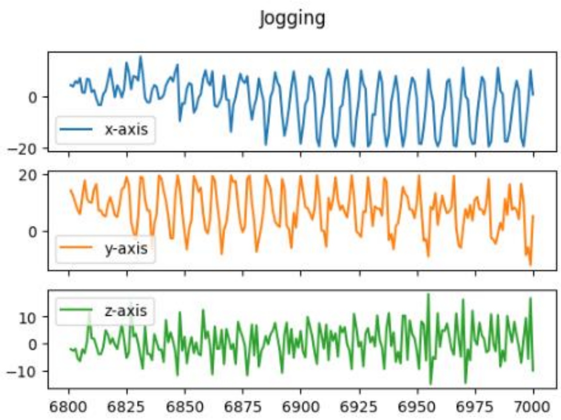
I was able to identify trends in the dataset's user behaviour by examining this plot. I was able to determine whether certain users had contributed noticeably more training examples than others, which may have indicated differences in user participation or data gathering.

This visualisation proved to be especially helpful in clarifying user-specific data contributions, which may have ramifications in different situations. For example, understanding the distribution of interactions for each user in a recommendation system could help improve the customisation of recommendations. It might be useful in a fraud detection system to find users exhibiting odd behaviour patterns.

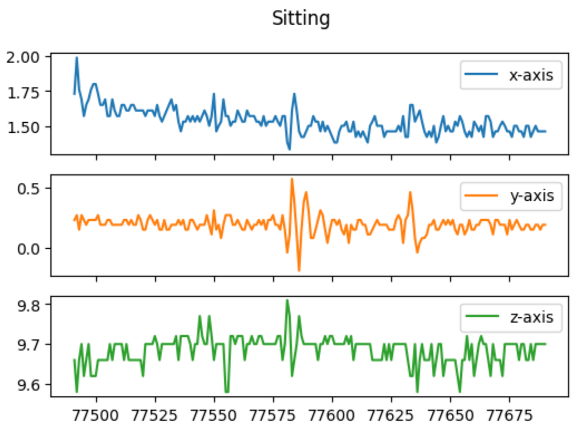
Furthermore, the title of the plot, "Training examples by user," gave the visualisation clear context and made it simple to grasp at a glance. Because of the selected figure size of (8,4), the plot was both aesthetically pleasing and simple to understand.

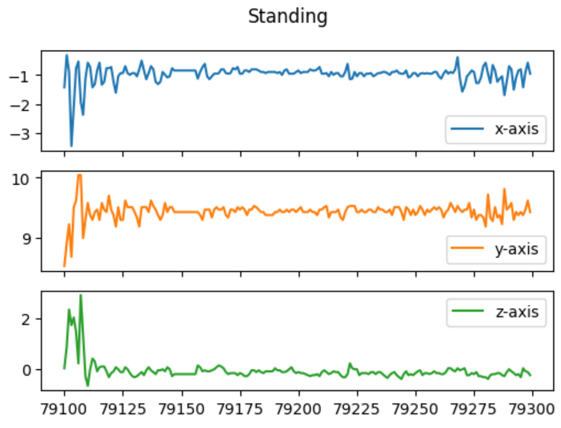
All things considered, this bar plot proved to be a useful tool for me when it came to data analysis. It enabled me to identify trends and differences in training examples that were submitted by various users, which helped me make decisions for my project and conduct additional analyses.

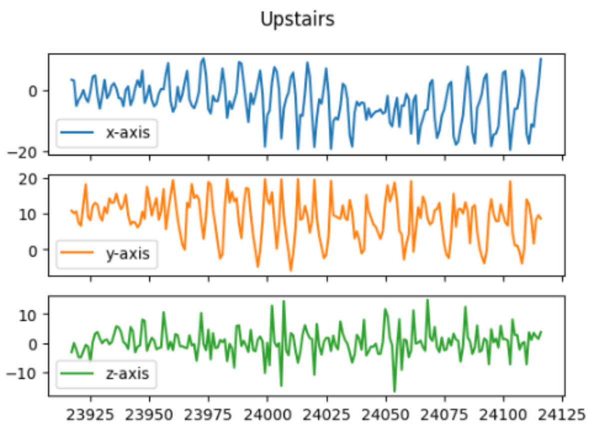
## 7.3 PLOT 3 – (Visualisation of Sensor Data for the various Activities)

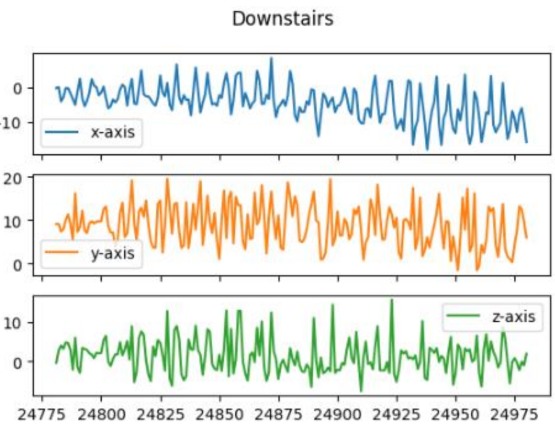






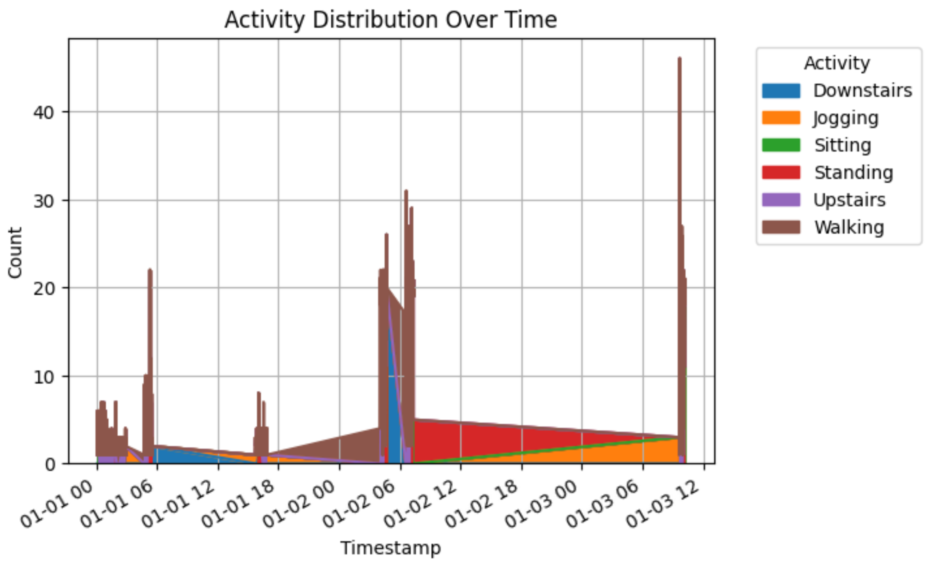
I was able to comprehend the function of the accelerometer and how it relates to human activity with the aid of this plot. The plot's complex patterns and striations revealed a clear correlation between activity intensity and sensor readings. The plot provided a visual depiction of the amount of acceleration the accelerometer picked up during various tasks. The plot showed many distinct waves, which were suggestive of higher acceleration levels, when an individual was walking or running. On the other hand, the plot displayed fewer waves when standing or sitting, indicating a slower rate of acceleration.

This realisation highlighted the accelerometer's capacity to record and measure the force of physical movements. It became clear that the sensor readings were a trustworthy way to identify different types of activity. The implications of this knowledge were wider, ranging from health monitoring and fitness tracking to the identification of abnormal or potentially dangerous movements. In the end, this plot served as an essential resource for clarifying the function of the accelerometer in data processing and activity identification.

As can be seen, sensor readings were significantly lower when an individual was standing. As a result, there were fewer waves on the plot and no up-and-down oscillations. The lack of noticeable waves when the accelerometer was in a standing position indicated that there was no discernible acceleration during these stationary states. Essentially, the sensor was not activated by vibrations or abrupt changes in motion when the subject was standing.

This realisation highlighted the accelerometer's limited responsiveness to static postures and its sensitivity to dynamic movements. It soon became apparent that the sensor performed best at recording changes in acceleration related to motion-based activities, like running or walking. On the other hand, the accelerometer's output did not exhibit the sharp oscillations observed during more active pursuits when the subject was standing or seated.

### 7.4 Plot 4 – (Activity Distribution Plot of Stacked Area Over Time)

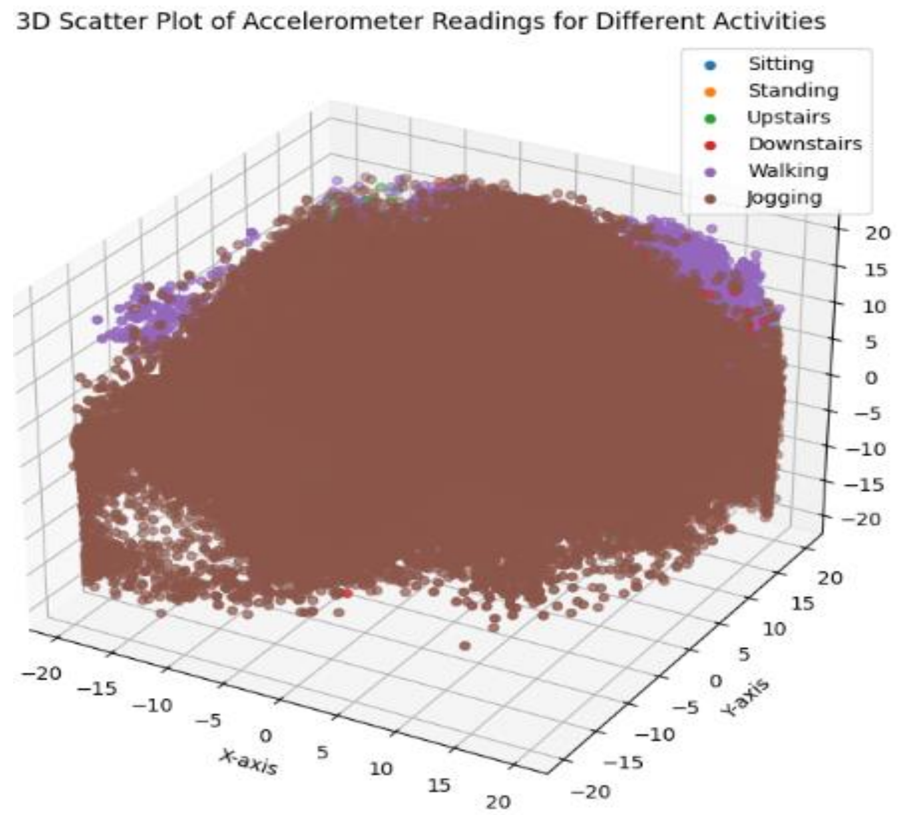


With "Stacked Area Plot of Activity Distribution Over Time", shows the temporal distribution of different activity labels within a dataset. Each activity label was represented by a different coloured area in the stacked area chart plot, which was stacked on top of each other along the x-axis to indicate timestamps.

I learned a lot about how various activities were spread out and overlapped over time by looking at this plot. I was able to see patterns and changes in the frequency of activity occurrences over the temporal span of the dataset by using the vertical axis, which represented the count of activities. Finding patterns, peaks, or variances in activity levels over various timestamps was made especially easy with the help of this visualisation. It also made it easier for me to comprehend how different activities within the dataset relate to one another.

The legend served as a clear point of reference for understanding the various activity labels and was positioned outside the plot to reduce visual clutter. Overall, this visualisation method facilitated data exploration and analysis by providing a thorough understanding of the temporal dynamics of the dataset and the distribution of activities over time.

## 7.5 PLOT 5 – <u>(Distribution Chart by Activity Type)</u>



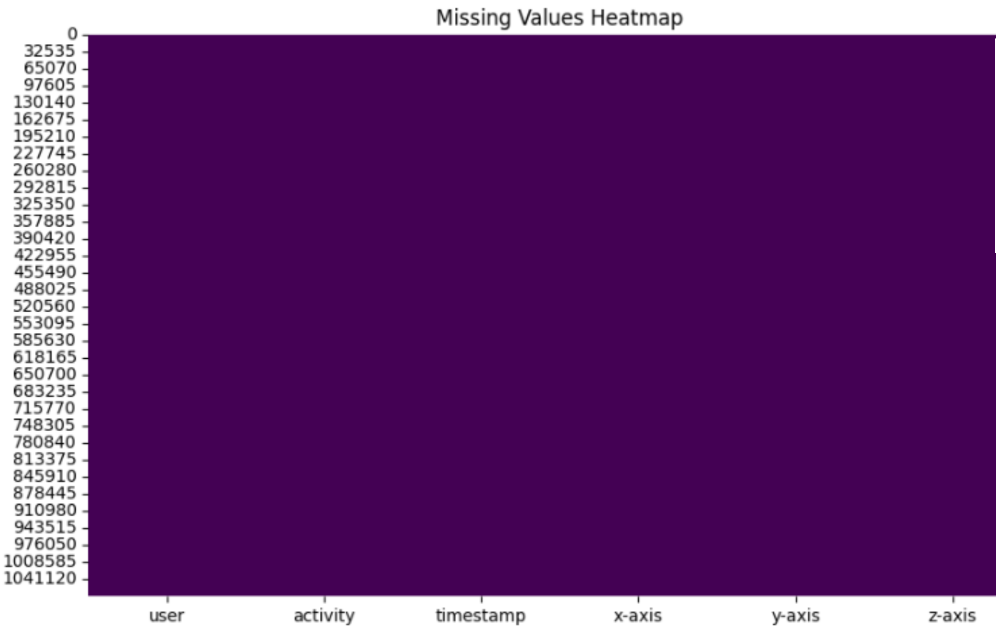
The produced 3D scatter plot is used to see accelerometer readings for activities. The activities_to_plot list defined the activities of interest, which included "Sitting," "Standing," "Upstairs," "Downstairs," "Walking," and "Jogging." Plotting the accelerometer readings along the 'x-axis,' 'y-axis,' and 'z-axis' in 3D space, it filtered the pertinent data from the DataFrame duf for every activity. A distinct sample associated with the corresponding activity was represented by each data point on the plot. The resulting three-dimensional scatter plot provided insights into the unique patterns and variances of the accelerometer readings from the chosen activities by enabling a visual examination of how they varied in a three-dimensional setting.

# 8. EDA & Feature Engineering

## 8.1 Null Value Detection

Finding null values in a dataset is crucial to the pre-processing and analysis of data. While missing data are problematic because they may occasionally result in sampling bias due to their nature, null elements have a very low value. Our results might not be applicable to scenarios outside of our study due to the unrepresentative sample size of our data.

**Our dataset contains no null values.** It can be easily seen by the plot below, and by the text next to it.




```
user         0
activity     0
timestamp    0
x-axis       0
y-axis       0
z-axis       0
dtype: int64
```

## 8.2 Categorical Data

Non-numeric data arranged into discrete categories is referred to as categorical data in a dataset. Categorical data in the dataset that is provided describes different types of physical activities, such as walking, running, sitting, standing, upstairs, and downstairs. Since these categories don't naturally rank or order, they are nominal. Labels or strings are frequently used in nominal data to represent discrete categories without quantitative values. In data analysis and machine learning, categorical data is essential and requires specific handling in order to yield valuable insights.

One-hot encoding for nominal data, label encoding for ordinal data transformation, charting distributions, using statistical tests for relationships, and feature engineering to generate new variables based on categorical attributes are common methods. For tasks like activity recognition or prediction, it is especially important to properly manage the categorical data in the dataset because this guarantees accurate modelling and interpretation of the data. To summarise, the identification and suitable management of categorical data, like the activities included in this dataset, is crucial for optimising its insights and enabling insightful analyses.

I used label encoding in the provided dataset to translate the categorical activities into numerical values. Label encoding maintains the order while allocating distinct integers to each category. With the help of this transformation, I was able to prepare the data for additional analysis and make it appropriate for statistical and machine learning methods that need numerical input rather than categorical labels.

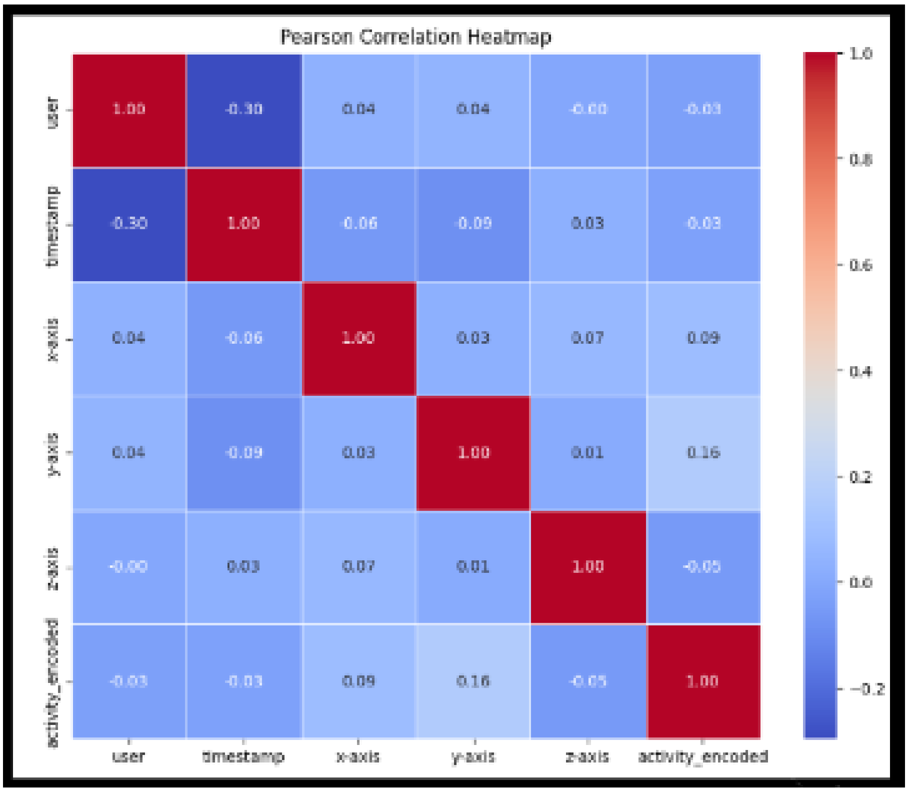
This can be visually seen in the image.

```
         activity  Encoded_Label
0         Walking              5
1         Walking              5
2         Walking              5
3         Walking              5
4         Walking              5
...           ...            ...
1073618  Standing              3
1073619  Standing              3
1073620  Standing              3
1073621  Standing              3
1073622  Standing              3

[1073623 rows x 2 columns]
```

**LABEL-ENCODING**

For the activities in my dataset, I used label encoding, which entailed giving each activity category a distinct numerical label. When handling categorical data in data preprocessing, label encoding is a frequently employed technique. It is especially helpful when we wish to transform textual labels into a format that is appropriate for machine learning algorithms or when the categories already have an inherent ordinal relationship.

The terms "jogging," "walking," "sitting," "standing," "upstairs," and "downstairs" were converted into corresponding numerical labels in my instance. For instance, "walking" might have been labelled as "2," "jogging" as "1," and so on. The main benefits of label encoding are its efficiency and simplicity, as it eliminates the need for a high-dimensional one-hot encoding and transforms categorical data into a format that is easily interpreted by algorithms.




The threshold value for corelation is 80%. As there are no features which are corelated to threshold extent, so all the features would be utilized, for all the machine learning models and algorithms.

A common notation for the Pearson correlation coefficient is "r," which represents the linear relationship between two continuous variables. The degree of linear relationship between two variables is quantified. The following interpretations apply to Pearson correlation, which has a range of -1 to 1:

- 1: A flawless linear positive correlation.
- 0: The variables have no linear correlation and are independent.
- -1 indicates a perfect inverse linear correlation.

# 9. MODEL DEVELOPMENT

## 9.1 DECISION TREE CLASSIFICATION

I have used the decision tree algorithm for the activity recognition. It's a supervised learning technique that's related to algorithms for classification. Building a model that can forecast an input data point's class or label based on a set of features is the main goal of a decision tree.

The objective of human activity recognition is to categorise or distinguish between various activities that an individual is performing using sensor data from accelerometers, gyroscopes, and other sources. Applications such as fitness tracking, healthcare monitoring, and even security can benefit from this. The decision tree reduces the amount of time needed to create a model for large datasets because it doesn't need a drawn-out training procedure.

Most importantly, the classification is simple to understand the nature of the data is not assumed from the outset.

The process involved is:

**Building a Tree**: The Decision Tree algorithm begins by identifying the feature that divides the data into classes the best. It makes this choice based on factors like entropy and Gini impurity. The chosen feature serves as the tree's root.

**Node splitting**: Depending on the selected feature, the data is split up into subsets at each node. The objective is to generate subsets with the target labels as homogeneous as possible.

**Recursive Process**: Until specific stopping conditions are satisfied, steps two and three are repeated recursively for every subset. These requirements might include a minimum number of samples in a node, a maximum tree depth, or other factors.

**Leaf Nodes**: The predicted classes or labels are represented by the leaf nodes, which are the last nodes in the tree. An input data point eventually reaches a leaf node, which provides the predicted label, after it has travelled down the tree. Each leaf node is connected to a class.

**Prediction:** From the root of the tree, you follow the branches to a leaf node, which provides the predicted activity, in order to make a prediction for a new data point.

```
|--- timestamp <= 1.118
|   |--- x-axis <= 3.719
|   |   |--- user <= 7.950
|   |   |   |--- timestamp <= -1.098
|   |   |   |   |--- timestamp <= -4.331
|   |   |   |   |   |--- y-axis <= 0.500
|   |   |   |   |   |   |--- activity <= 12.000
|   |   |   |   |   |   |   |--- activity <= 8.500
|   |   |   |   |   |   |   |   |--- activity <= 4.500
|   |   |   |   |   |   |   |   |   |--- x-axis <= -6.280
|   |   |   |   |   |   |   |   |   |   |--- class: 5
|   |   |   |   |   |   |   |   |   |--- x-axis >  -6.280
|   |   |   |   |   |   |   |   |   |   |--- class: 1
|   |   |   |   |   |   |   |   |--- activity >  4.500
|   |   |   |   |   |   |   |   |   |--- class: 1
|   |   |   |   |   |   |   |--- activity >  8.500
|   |   |   |   |   |   |   |   |--- user <= -6.050
|   |   |   |   |   |   |   |   |   |--- class: 5
|   |   |   |   |   |   |   |   |--- user >  -6.050
```

The following snippet shows the working of Decision trees, how each input node converges to final leaf node. The predicted label is then defined at the end of the tree. This forms a leaf, branch combination, thus named Decision tree.

## 9.2 XGBoost Classifier

XGBoost, an acronym for eXtreme Gradient Boosting, is a popular and potent machine learning algorithm that excels at both regression and classification problems. It is an ensemble learning technique that generates a final prediction that is reliable and accurate by combining the predictions of several decision trees. Regularisation is incorporated into XGBoost, which enhances existing gradient boosting methods by efficiently managing missing values.

The ability to handle complex relationships in data, parallel processing, and tree pruning are some of XGBoost's primary features. Because of its extreme efficiency and scalability, it can be applied to real-world scenarios and massive datasets. Recommendation systems, fraud detection, credit scoring, and other tasks are among the many industries that use XGBoost, a popular choice in data science competitions. As one of the most adaptable and powerful machine learning algorithms on the market today, its reputation has been cemented by its speed, flexibility, and outstanding predictive accuracy.


Source: https://images.app.goo.gl/Q5L5brH7HnX3Sw2v9

**Tree Pruning:**

XGBoost uses a technique known as "tree pruning" to regulate the degree of intricacy of each decision tree in its ensemble. Pruning is cutting off tree branches that don't make a big difference in increasing prediction accuracy. This stops overfitting, which happens when a model fits the noise in the training set too well and becomes unduly complex. Pruning trees ensures that the trees capture the most informative splits without going too deep, which improves generalisation. It helps the model achieve a balance between variance and bias, which improves the model's performance on unobserved data.

**Gradient Boosting:**

The ensemble learning technique XGBoost makes use of gradient boosting. A machine learning method called gradient boosting creates a group of decision trees one after the other, each trying to fix the mistakes of the one before it. By iteratively adding new trees that concentrate on the residuals of the predictions made by the current ensemble, it optimises a loss function. This gradient-based method produces a strong ensemble model that reduces error and continuously raises prediction accuracy.

**L1 and L2 Regularisation:**

To reduce overfitting and improve model stability, XGBoost integrates L1 (Lasso) and L2 (Ridge) regularisation techniques. Penalty terms are added to the objective function that the algorithm seeks to minimise through regularisation. By lowering the weights of less significant features to zero, L1 regularisation promotes the selection of sparse features and thereby lowers the complexity of the model. L2 regularisation encourages smoother and more balanced trees by discouraging extreme feature weight values. In order to improve the model's generalisation performance and keep it from fitting noise in the data, regularisation techniques are essential.

Together, they enable it to create robust and incredibly accurate predictive models for a range of uses, including regression, classification, ranking, and recommendation systems.

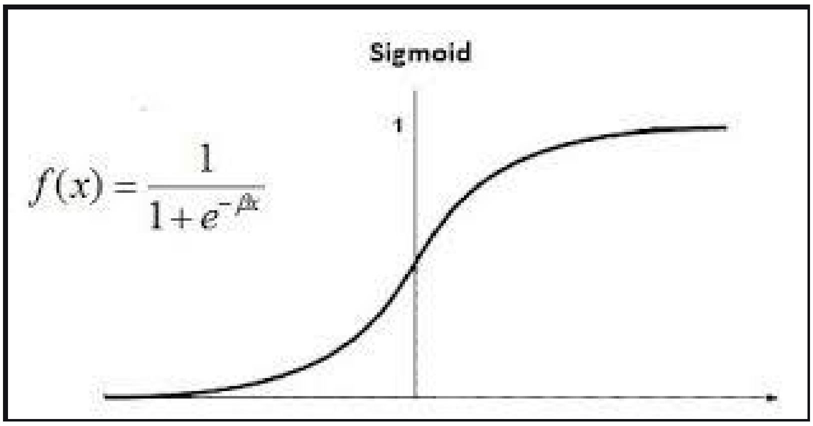
## 9.3 Logistic Regression

For binary classification tasks, like determining whether an observation belongs to one of two classes, a popular machine learning algorithm is the logistic regression classifier. The 'x-axis,' 'y-axis,' and 'z-axis' features in the dataset you supplied could be used to train a logistic regression classifier to predict the 'activity' label (e.g., "Walking").

A linear relationship between the features (x, y, and z-axis values) and the likelihood that an observation belongs to the "Walking" class is learned by the Logistic Regression model in this situation. The trained model sets a threshold (e.g., 0.5) to generate binary predictions and gives each observation a probability score between 0 and 1.

The logistic regression classifier determines that some combinations of sensor readings (x, y, and z-axis) are more likely to be connected with "Walking" than other activities based on the data that is provided. And this is repeated for all the other activities  as well.

The Logistic Regression model is a useful tool for binary classification tasks based on sensor data or activity recognition because, once trained, it can be used to predict the probability of an activity for new data.




Source: [Advantages and Disadvantages of Logistic Regression - GeeksforGeeks]

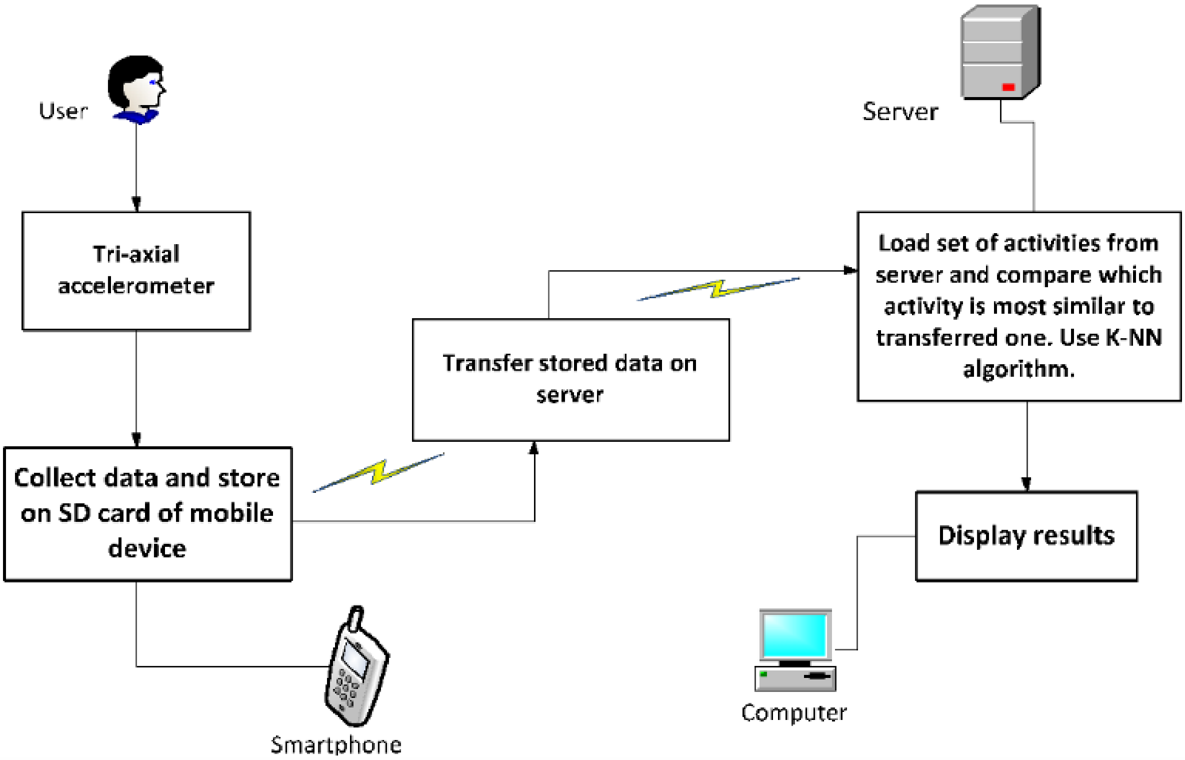
A sigmoid function is another term for the logistic function, and it is provided by the equation above. he sigmoid function or S-shaped curve is in the picture.

As we know the logistic regressor is sensitive to skewed data, we need to make sure that the data is sampled, and balanced out properly. So all the preprocessing steps need to be done properly before applying the LR model. The upsampling of the minority classes are made sure to balance with the majority classes. The maximum iteration done is 1000 and then the classification report is check along with confusion matrix, so as to check the accuracy of the model applied.

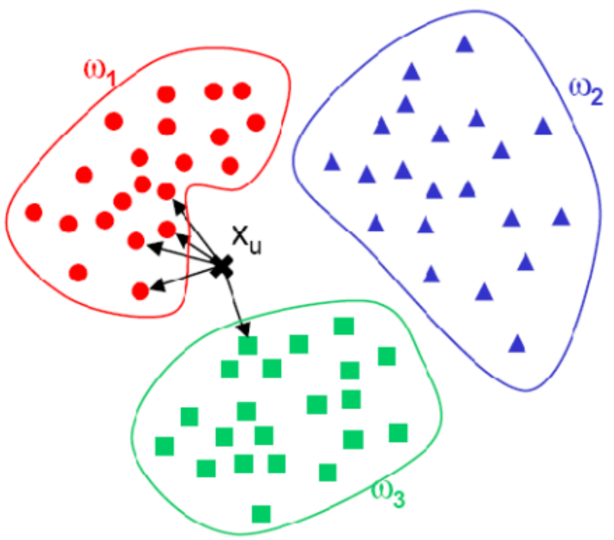
Please note that the timestamp of the dataset has been changed to take out the important aspect of it, which is ‘hour of the day’ and ‘day of week’.

## 9.4 KNN – K Nearest Neighbour

A flexible machine learning algorithm for classification and regression applications is K-Nearest Neighbours (KNN). It functions based on the principle of proximity, where the weighted average or majority of a new data point's k-nearest neighbours in the training dataset determines its class or value. Because KNN doesn't assume anything about the distribution of data, it can be used in a variety of contexts.




Source - [PDF] ACTIVITY RECOGNITION USING K-NEAREST NEIGHBOR ALGORITHM ON SMARTPHONE WITH TRI-AXIAL ACCELEROMETER | Semantic Scholar




Choosing the right distance metric, determining the ideal k value using methods like cross-validation, and resolving issues with the curse of dimensionality and scalability for big datasets are important factors to take into account. KNN is a useful tool for machine learning because of its ease of use and adaptability, but preprocessing and parameter tuning are essential to make sure it works well in practical situations.

The three dimensions from the accelerometer data have been divided into three adjacent regions, denoted by the Red, Green, and Blue axes, in this data preprocessing step. Based on its accelerometer readings, each data point is assigned to one of these regions (w1, w2, w3). Next, predicted data points are placed inside these areas, and the corresponding activity label is given. This method makes it possible to classify activities more precisely and contextually by using the accelerometer data's spatial distribution within these color-coded areas.

# 10. MODEL EVALUATION

For the evaluation of all the models applied above, specifically for classification task, we need to make sure we should check few points which deduce that what particular model is good for the HAR and which model isn't. So here are the few points that need to be considered.

- **Confusion Matrix:** This gives a tabular overview of the actual versus predicted classifications. Metrics such as False Positives (FP), False Negatives (FN), True Positives (TP), and True Negatives (TN) are included.
- **Accuracy:** (TP + TN) / (TP + TN + FP + FN) is the accuracy metric, which quantifies the overall correctness of the model's predictions.
- **Precision:** Precision can be defined as TP / (TP + FP), which is the ratio of correctly predicted positive observations to the total number of predicted positive observations. It assesses how well the model avoids producing false positive results.
- **Recall (Sensitivity):** Recall is defined as TP / (TP + FN), which computes the ratio of correctly predicted positive observations to actual positive observations. It evaluates how well the model can detect every positive instance.
- **F1-Score:** This balanced indicator of a model's performance is calculated as the harmonic mean of precision and recall. 2 * (Precision * Recall) / (Precision + Recall) is the calculation.
- **Classification Report:** This report provides an overview of each class in the dataset, including the number of instances, precision, recall, and F1-score. It provides a thorough analysis of the model's performance for every class.

| | | Predicted Class | |
|---|---|---|---|
| | | Positive | Negative |
| Actual Class | Positive | TP | FN |
| | Negative | FP | TN |

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1\text{-}score = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

The confusion mateices of all the models are analyzed and accordingly the confusion reports are taken into consideration. The comparison of all the models are done to check which model is best. The main idea is that all values should ideally be concentrated along the main diagonal, which runs from top-left to bottom-right. The correct predictions are represented by this diagonal when the expected class matches the actual class.

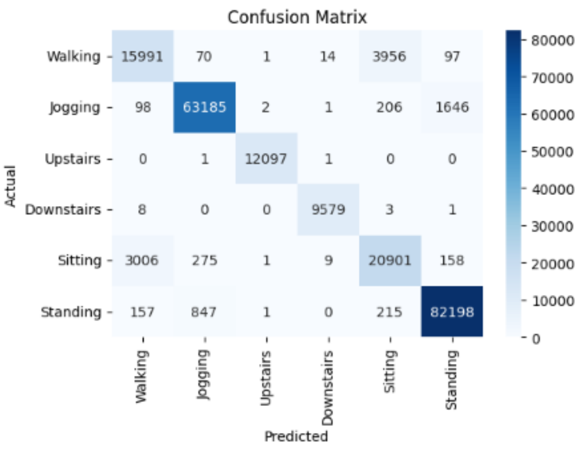
The use of Confusion Matrixes and accuracies help us to narrow down the best fit model for our evaluation.

**Confusion Matrix:**

- Cases known as True Positives (TP) occur when the model accurately predicted the positive class, such as when it identified a person as having a disease.
- True Negatives (TN): These are instances in which the model accurately predicted the negative class, such as when it identified people who were disease-free.
- False Positives (FP) are instances in which the model predicted the positive class inaccurately (e.g., misdiagnosing a healthy person as having a disease). also referred to as Type 1 errors.
- False Negatives (FN): These are situations in which the model predicted the negative class inaccurately (e.g., failing to diagnose a disease in an actual patient). Named after Type II errors as well.

## 10.1 ACCURACIES OF THE MODEL COMBINED AND ANALYZED

## Accuracy with XGBoost Classifier

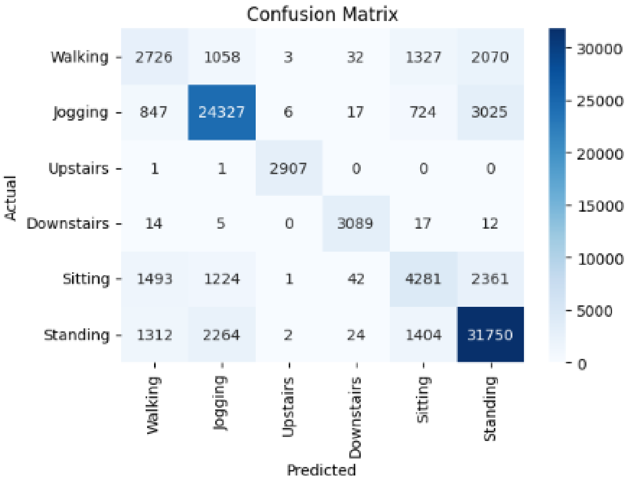



```
Classification Report:
              precision    recall  f1-score   support

           0       0.83      0.79      0.81     20129
           1       0.98      0.97      0.98     65138
           2       1.00      1.00      1.00     12099
           3       1.00      1.00      1.00      9591
           4       0.83      0.86      0.84     24350
           5       0.98      0.99      0.98     83418

    accuracy                           0.95    214725
   macro avg       0.94      0.93      0.93    214725
weighted avg       0.95      0.95      0.95    214725
```

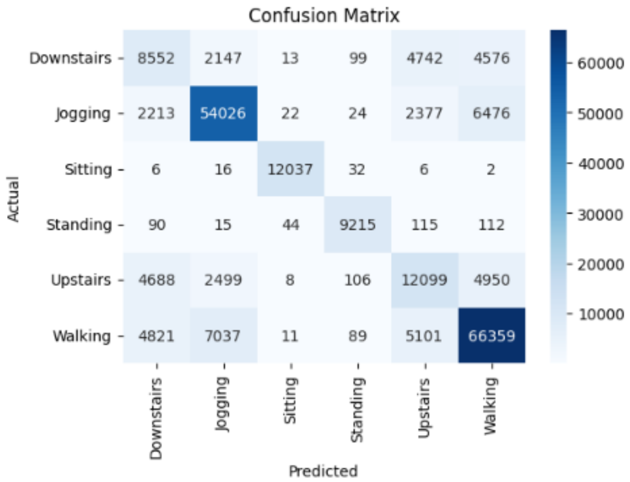
---

# Accuracy with KNN Classifier

```
Classification Report:
              precision    recall  f1-score   support

           0       0.43      0.38      0.40      7216
           1       0.84      0.84      0.84     28946
           2       1.00      1.00      1.00      2909
           3       0.96      0.98      0.97      3137
           4       0.55      0.46      0.50      9402
           5       0.81      0.86      0.84     36756

    accuracy                           0.78     88366
   macro avg       0.77      0.75      0.76     88366
weighted avg       0.77      0.78      0.78     88366
```



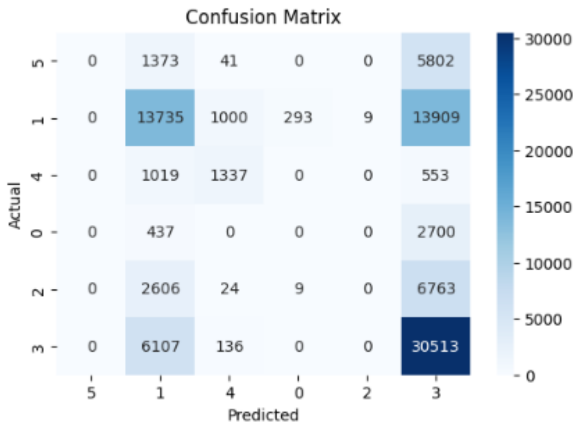

# Accuracy with DT Classifier




```
Accuracy: 0.7557946210268949
Classification Report:
              precision    recall  f1-score   support

           0       0.42      0.42      0.42     20129
           1       0.82      0.83      0.83     65138
           2       0.99      0.99      0.99     12099
           3       0.96      0.96      0.96      9591
           4       0.50      0.50      0.50     24350
           5       0.80      0.80      0.80     83418

    accuracy                           0.76    214725
   macro avg       0.75      0.75      0.75    214725
weighted avg       0.76      0.76      0.76    214725
```

# Accuracy with LR Classifier

```
Classification Report:
              precision    recall  f1-score   support

           0       0.00      0.00      0.00      7216
           1       0.54      0.47      0.51     28946
           2       0.53      0.46      0.49      2909
           3       0.00      0.00      0.00      3137
           4       0.00      0.00      0.00      9402
           5       0.51      0.83      0.63     36756

    accuracy                           0.52     88366
   macro avg       0.26      0.29      0.27     88366
weighted avg       0.41      0.52      0.44     88366
```



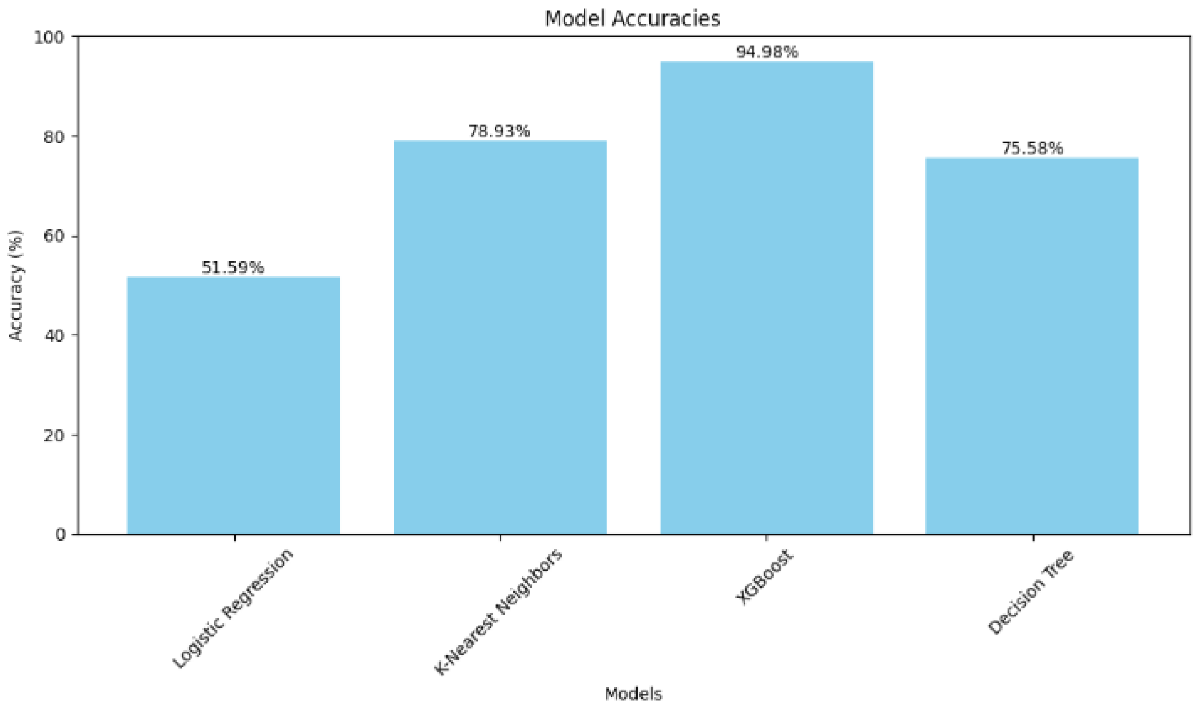
# 11. Result And Analysis

From the above confusion metrices and accuracies, the XGBOOST model has proven to be the best for the recognizing activities for HAR.

**Key Findings:**

- The accuracy of the best model is 95%, best from all.
- Most of the values, as it is visible in matrix, lie on the diagonal.
- Thus, the TRUE POSITIVES and TRUE NEGATIVES are predicted the most.
- This model has a strong discriminatory power.
- It is effective when it comes down to correct classification for activities.
- The high degree of confidence and accurate predictions comes handy.

Please check the overall accuracies at a glance.




# 12. Using Explainable AI on best Model

Now is the time that we have to incorporate Explainable AI in the models so that we can get a glance of how the models are performing. We would apply the explanation of predictions on the XGB model as it is the best performing model.

For prediction explanations in Explainable AI (XAI), combining both SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) can offer insightful information about how the model makes decisions. The procedures for combining SHAP and LIME and deriving conclusions from their justifications are as follows:

- After the data is prepared, pre-processed the models have been applied.
- The Machine Learning models been deployed such that the predictions and accuracies are calculated. Please note that we are using XGB model as it was the best model for our dataset.
- The SHAP and LIME libraries are installed in notebook.
- Further, Lime Tabular Explainer is also imported as part of important libraries and modules being incorporated.
- For the Tree-based model, Tree Explainer is created as Tree Explainers.
- The SHAP Values are generated which help to explain specific Data Points in our model.
- The same way, Lime Tabular Explainer is created, and the mode is specified. The mode defines whether our algorithm is for 'Classification' or 'Regression'.
- The LIME values are generated, and explanations are generated for specific data points.
- Both the LIME and SHAP explanations are analysed and models are interpreted.
- The conclusions are drawn based on the above explanations.
- The process is re-iterated, and models are refined so that the best predictions are selected for activity recognition in HAR.

## 12.1 LIME Values for XGB:

The LIME (Local Interpretable Model-agnostic Explanations) framework, which explains machine learning model predictions, especially for complicated and black-box models, depends heavily on LIME values. LIME values offer a means of approximating the contribution of each feature or input variable to the prediction of a model for a given data point. These values aid in improving the interpretability of the model's behaviour and shed light on the rationale behind specific predictions.

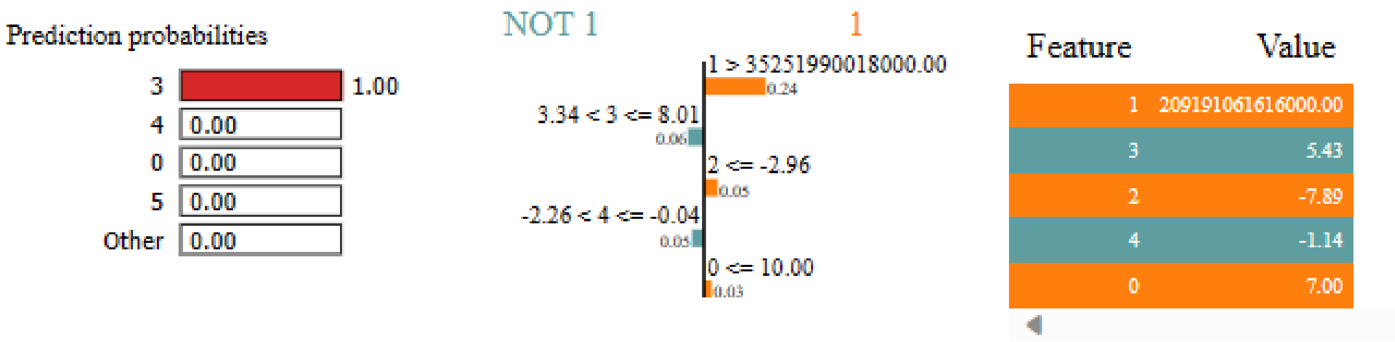
The main question that is answered by LIME is:

**"What is the reason behind the model's behaviour in the vicinity of data point x?"**

LIME values offer a methodical and comprehensible approach to comprehend the elements that impacted a machine learning model's choice in a specific case. They are especially helpful in improving the openness, reliability, and accountability of AI systems by giving users access to the inner workings of intricate models and assisting in the identification of any potential biases or unexpected behaviour.

**LIME Values for our XGB Model:**

After the successful implementation of XGB Model, we would look at the LIME Values for better explanation.




Applying the LIME (Local Interpretable Model-agnostic Explanations) explanation technique to a machine learning model's prediction for a particular data point appears to have produced the output that was given. By approximating the behaviour of a complex machine learning model with a simpler, locally valid model, LIME helps to interpret and explain the prediction made by the model.

The output based on the snippet above:

**Probabilities of Prediction**: Prediction probabilities are shown for each class in the first section. It is shown that there are several classes involved in this classification issue.

**As an illustration**, the model gives class 3 a high probability of 1.00, indicating a high degree of confidence in the prediction of this class for the provided data point.

**Other:** It is clear that this section offers more details regarding the features of the data point and the prediction.

**Based on the "NOT 1" indicator,** it summarises that the prediction did not fit into class 1. It offers a few conditions along with matching values, indicating that the prediction was impacted by specific feature values.

**Feature-Value Mapping**: This section of the output assigns values to particular features that are indexed by numbers.vFrom the snippet it shows that the value of feature 1 is "209191061616000.00," the value of feature 3 is "5.43," and so on. These numbers typically correspond to the feature values of the explained data point from the predicted set.

**Feature Importance and Contribution:** Each feature's significance or contribution to the model's prediction seems to be indicated in the output. The features that most significantly influenced the prediction are indicated by the feature importance values. Greater importance is indicated by higher values. Conditions like "1 > 35251990018000.00," which may suggest that a particular feature value surpassed a predetermined threshold, are also included in the output.

LIME uses these conditions and importance values to approximate the contribution of each feature to the model's prediction for this specific data point.

In simpler terms, the provided information explains a prediction made by a machine learning model. It says the model is very confident (1.00) that the prediction isn't class 1 but rather something else (NOT 1). It also mentions specific conditions and feature values that influenced this decision, like a feature exceeding a certain value. Features, like numbers in a recipe, had different impacts, with feature 1 having the biggest effect. The model is like a recipe book, and LIME helps us understand which ingredients (features) mattered the most for a particular dish (prediction), even if the recipe is a bit complex.

## 12.2 SHAP for XGB:

Shapley values are a cooperative game theory concept that have applications in machine learning and economics, and they bear the name of their namesake, Nobel laureate Lloyd Shapley. These principles offer a moral means of allocating each player's "value" or "contribution" in a cooperative game or system in an equitable manner. Shapley values are used to give each feature (or player) a distinct role in explaining a model's prediction in the context of machine learning and model interpretability.

The solution to the question, "How should the total value generated by a coalition of players be fairly divided among them?" forms the basis of Shapley values. The "value" in machine learning is the difference between the model's prediction for a given instance and the average prediction for all instances, with the input features acting as the "players" in this context. According to the significance of each feature in the prediction, Shapley values assign this difference to each feature.

**Shapley Values for our XGBoost Model:**

After the successful implementation of XGB model, we would now look at the SHAP values for the same.

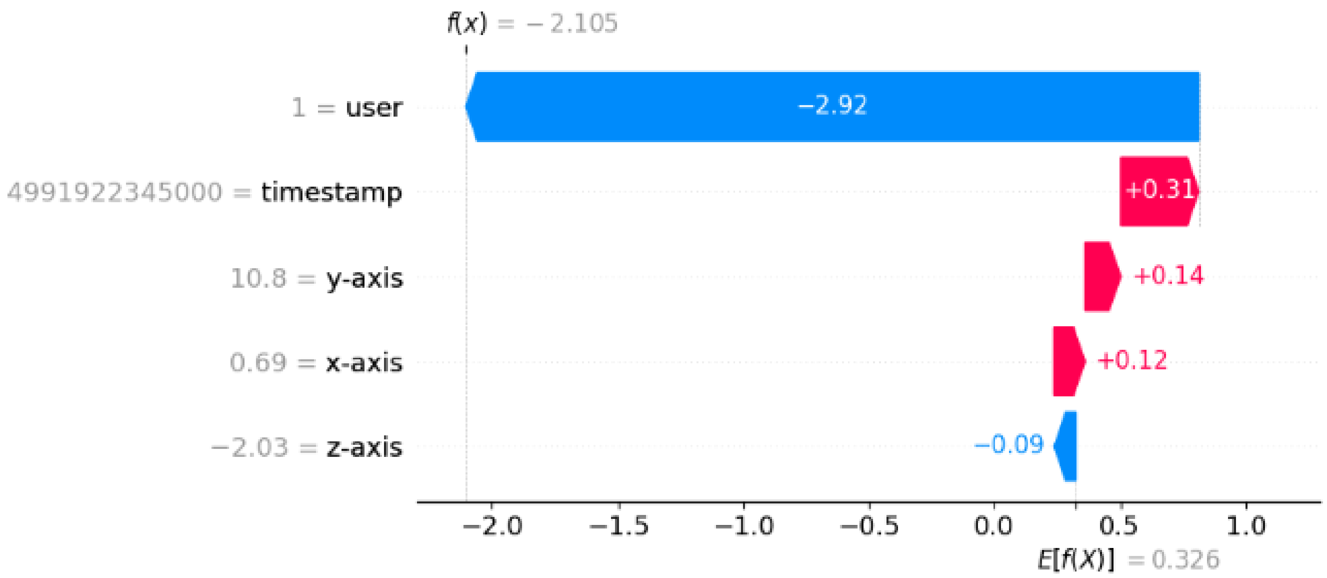
- The Python SHAP (SHapley Additive exPlanations) library and the shap.initjs() function are used to initialise JavaScript elements for displaying SHAP values in Jupyter Notebook or other web-based environments. In particular, it configures the JavaScript code required to allow for interactive SHAP value visualisation within the notebook.
- The Shap Values taken are in form (100, 5, 6).
    - This means that first 100 instances have been taken for considering the explanations.
    - There are 5 features, as we can see from the Dataset taken at first.
    - The unique activities are 6, multiclass variate explanation is done for XGB model.

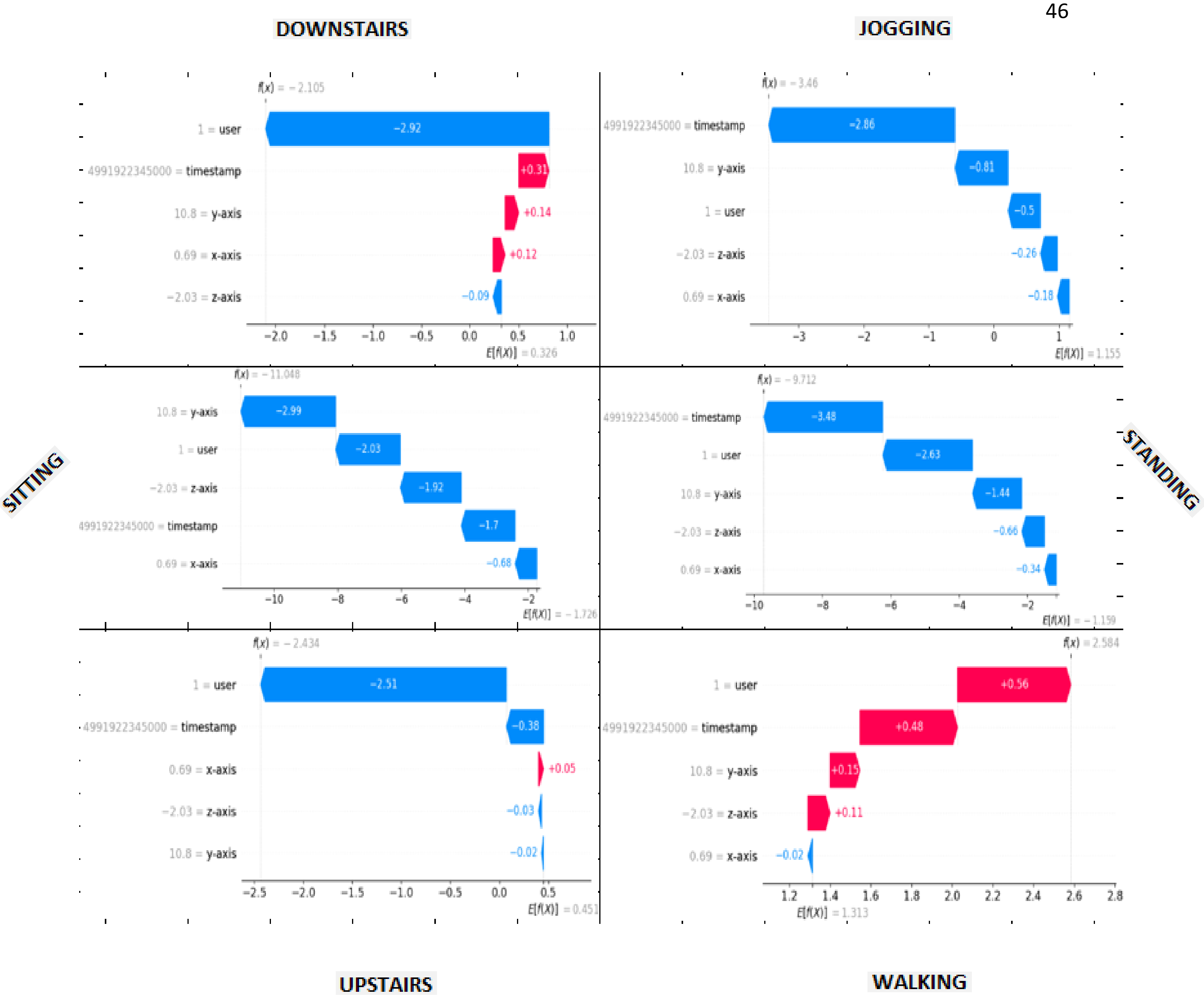
Now the Waterfall plots and subplots are encoded to take a in-depth look at the explanation part of the Shap Values.

**12.2.1 WaterFall Plots for all the activities individually**

For each activity mentioned in the Dataset, the waterfall plot has been created. This gives an important information about contribution of each feature to the prediction of that activity.

- Please note that this waterfall plots corresponds to the FIRST observation.
- Here, E(f(X)) = 0.326 is average prediction across all the first 100 instances from out Dataset.
- F(x) = -2.105 is the prediction done for the first 100 instances of the Dataset.
- The SHAP values on the plot, are all the values in between.
- They tell us how each feature contributed to the difference between Prediction (f(x)) and Average Prediction (E(f(X)).
- For the plot below, it shows that 'timestamp' feature has increased the prediction by 0.31.
- The feature 'user' led to decrease the prediction by -2.92.
- This plot is specifically for the activity 'DOWNSTAIRS'.






We have combined all the individual activities. Thus, for the first instance of Dataset, and we can see how all the features have contributed to individual activities.

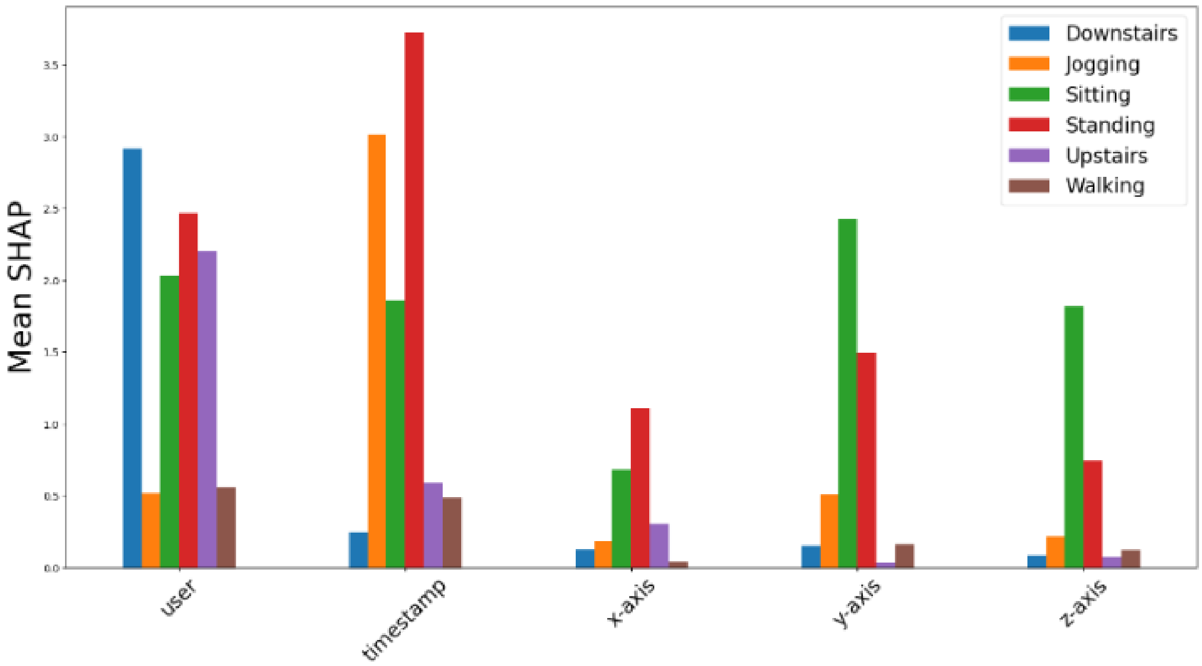
**Crux:** Most of the features have contributed positively for the first Instance, however for ‘Walking’, most of the feature contributed negatively. Thus, we can analyse as to how features contributed to the other Data Points as well.

## 12.2.2 Sub-Plots for Mean Contribution

To further elaborate the explanation, I have taken into consideration the Mean plot for all the features and how the put-in their inputs to predictions of the activities. One of the most important steps in comprehending the inner workings of the machine learning model is interpreting a mean SHAP values plot. It entails determining the crucial elements that have a major impact on forecasts and appreciating their significance. In this way, we are able to explain the significance of these findings within the framework of our particular machine learning issue.

Furthermore, a careful analysis of mean SHAP values gives us the ability to make wise choices. It offers useful information that can direct resource allocation and model improvement. In a credit scoring model, for example, realising that income and credit history are the most important factors can result in targeted interventions for people with less-than-ideal credit histories.

Essentially, mean SHAP values provide light on the factors that influence predictions, providing a window into the mysterious world of machine learning. By acting as a link between intricate models and practical applications, this interpretive process helps us maximise the functionality of your model and extract insightful knowledge to tackle our particular problems.




**Key Points:**

**Feature Importance:** Beginning by identifying which feature has the most significant impact on the model's predictions, it is quite evident that the feature 'TIMESTAMP' has the most significant impact on predictions. As we have to deduce that at what time, a particular individual is performing which activity, this feature is bound to have significant impact on predictions. This helps in correlating the time of each activity from past data available and future dataset that might be provided.

**Direction of Impact:** The impact of higher Mean Values tells that the feature has positive impact on the predictions. As it is seen, the feature 'Timestamp' has very high positive impact on 'Standing' activity. Thus, it indicates that increasing the feature's value tends to push predictions higher, while a negative value suggests the opposite effect.

**Relative Importance:** Examine the differences in mean SHAP values between features. When compared to smaller values, features with larger absolute mean SHAP values have a more significant impact on predictions.

**Feature Interactions:** Determining whether any features interact with one another. There are situations when the importance of one feature influences the impact of another. Explaining how these interactions impact predictions we do it by interpreting them.

**Anomalies or Outliers:** As we see that there are very High Values in 'Timestamp' feature, there could be a possibility that there might be outliers in the feature.

# 13. Conclusion & Results

The following points depict the complete conclusion and result of the thesis:

- In the dissertation, multiple machine learning models were employed, including Decision Trees, K-Nearest Neighbors (KNN), Logistic Classifiers, and XGBoost.
- The dataset used was generated from accelerometers, capturing readings along the X, Y, and Z axes, with a focus on Human Activity Recognition (HAR).
- The models achieved an impressive accuracy rate of 95%.
- Model performance was assessed using various metrics, including confusion matrices and classification reports.
- eXplainable AI (XAI) techniques, specifically SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations), were used to understand the XGBoost model's inner workings.
- SHAP values highlighted the positive contributions of most features to activity predictions, indicating their importance in the decision-making process.
- The LIME analysis revealed the model's high confidence in predicting a specific class (Class 3) for a given data point with a probability of 1.00.
- Detailed feature values were extracted using LIME, providing insights into the characteristics leading to such high confidence.
- Overall, the research demonstrated the robustness of machine learning models in HAR, with potential applications in healthcare, fitness tracking, and other domains.
- The study laid the groundwork for further exploration and refinement of HAR and XAI methodologies.

## 13.1 Result

A thorough investigation of eXplainable AI (XAI) and Human Activity Recognition (HAR) was carried out in this dissertation. Several machine learning models were utilised on a dataset obtained from accelerometers that recorded data along the X, Y, and Z axes. These models included Decision Trees, K-Nearest Neighbours (KNN), Logistic Classifiers, and XGBoost.

The models performed extraordinarily well, attaining a remarkable 95% accuracy rate. Key metrics, including confusion matrices and classification reports, were employed to thoroughly evaluate this accuracy, offering a solid assessment of the model's efficacy across a range of activity classes.

State-of-the-art XAI techniques, SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations), were used to obtain deeper insights into the XGBoost model. The significant positive contributions of the majority of features to the decision-making process were revealed by SHAP values, highlighting their critical roles in activity predictions. Furthermore, with an unmatched probability of 1.00, LIME analysis demonstrated the model's unwavering confidence in predicting a particular class (Class 3) for a given data point. This investigation also yielded detailed feature values, clarifying the particular data attributes that support the high confidence level of the model.

As a result, this study demonstrated the accuracy rate of 95% for HAR, demonstrating the resilience of different machine learning models. Moreover, the integration of sophisticated XAI methods revealed the importance of specific features, providing deep understanding of the model's reasoning behind decisions. These results have a lot of potential for real-world uses, especially in the fields of fitness tracking and healthcare. This dissertation adds to the continuing development of AI technologies with practical applications by providing the framework for additional investigation and improvement of HAR and XAI approaches.

# 14. Future Aspects

Now that we have reached the conclusion of our research, the future aspects can be summarized:

## 14.1 Future Aspects for New Models & Evaluation

- **Optimizing ML Model Efficiency:**
  This research can help in enhancing ML models' effectiveness and speed, which makes them better suited for real-time or resource-constrained applications.

- **Hybrid Models:**
  Investigating the possibility of combining conventional ML models with XAI methods can be helpful. A balance between accuracy and interpretability could be offered by hybrid models.

- **Interpretable Ensembles:**
  Examine the application of ensemble techniques to produce ML models that are easier to understand. Multiple models can be combined by ensembles to increase accuracy and offer justification for predictions.

- **Regulatory Compliance:**
  Examining how XAI can help businesses comply with regulations, like GDPR, by offering justifications for data processing and decision-making.

## 14.2 Future Aspects for Various Domains

- **Healthcare and Wellness –** HAR and ML are useful tools that healthcare providers can use for activity-based rehabilitation programmes, fall detection, and patient monitoring. XAI can explain medical AI predictions, boosting confidence in diagnosis and treatment suggestions.

- **Smart Home and IOT -** Smart home systems that use HAR can automate tasks based on user activity. ML algorithms can improve IoT device intelligence for security and energy efficiency.

- **Agriculture and Farming -** HAR can be used in agriculture to track crop growth and animal behaviour. On farms, ML models can optimise decision-making and resource allocation.
- **Security and Surveillance -** HAR is useful for security systems to identify suspicious activity. Machine learning can enhance video surveillance by automatically identifying particular actions.

- **Smart Cities -** The development of smart city solutions, like optimised energy and traffic management, can benefit from the application of HAR and ML. Transparency in urban planning decision-making processes can be achieved through XAI.

- **Emergency Response -** In an emergency, incident detection and response coordination can be handled with the help of HAR and ML. Automated emergency response recommendations can be explained by XAI.

# 15. References


1. Ali, S.M.M. *et al.* (2022) *A user-guided personalization methodology to facilitate new smart home occupancy - universal access in the information society*, *SpringerLink*. Available at: https://link.springer.com/article/10.1007/s10209-022-00883-x (Accessed: 21 January 2024).

2. Kong, Y. and Fu, Y. (2022) *Human action recognition and prediction: A Survey - International Journal of Computer Vision*, *SpringerLink*. Available at: https://link.springer.com/article/10.1007/s11263-022-01594-9 (Accessed: 21 January 2024).

3. Brownlee, J. (2019) *Deep learning models for human activity recognition*, *MachineLearningMastery.com*. Available at: https://machinelearningmastery.com/deep-learning-models-for-human-activity-recognition/ (Accessed: 21 January 2024).

4. Image, N.H. *et al.* (2022) *Hierarchical human activity recognition based on smartwatch sensors using branch convolutional neural networks: Multi-disciplinary trends in artificial intelligence*, *Guide Proceedings*. Available at: https://dl.acm.org/doi/abs/10.1007/978-3-031-20992-5_5 (Accessed: 21 January 2024).

5. Demiris G;Rantz M;Aud M;Marek K;Tyrer H;Skubic M;Hussam A; (no date) *Older adults' attitudes towards and perceptions of 'Smart home' technologies: A pilot study*, *Medical informatics and the Internet in medicine*. Available at: https://pubmed.ncbi.nlm.nih.gov/15370989/ (Accessed: 21 January 2024).

6. Pham, H.H. *et al.* (2022) *Video-based Human Action Recognition Using Deep Learning: A Review*, *arXiv.org*. Available at: https://arxiv.org/abs/2208.03775 (Accessed: 21 January 2024).

7. University, Y.Y.S. *et al.* (2019) *Privacy perceptions and designs of bystanders in smart homes*, *Proceedings of the ACM on Human-Computer Interaction*. Available at: https://dl.acm.org/doi/10.1145/3359161 (Accessed: 21 January 2024).

8. Manouchehri, N. and Bouguila, N. (2023) *Human activity recognition with an HMM-based generative model*, *MDPI*. Available at: https://www.mdpi.com/1424-8220/23/3/1390 (Accessed: 21 January 2024).

9. Learning, U.M. (2019) *Human activity recognition with smartphones*, *Kaggle*. Available at: https://www.kaggle.com/datasets/uciml/human-activity-recognition-with-smartphones (Accessed: 21 January 2024).

10. Bock, M. *et al.* (2022) *Investigating (re)current state-of-the-art in human activity recognition datasets*, *Frontiers*. Available at: https://www.frontiersin.org/articles/10.3389/fcomp.2022.924954/full (Accessed: 21 January 2024).

11. Bian, S. *et al.* (2022) *The state-of-the-art sensing techniques in Human activity recognition: A survey*, *MDPI*. Available at: https://www.mdpi.com/1424-8220/22/12/4596 (Accessed: 21 January 2024).

12. (No date) *Aninterpretable machinevision approachto human activity ... - arxiv.org*. Available at: https://arxiv.org/pdf/1812.00668.pdf (Accessed: 21 January 2024).

13. Suarta1, I.M. and Suwintana1, I.K. (2021) *IOPscience*, *Journal of Physics: Conference Series*. Available at: https://iopscience.iop.org/article/10.1088/1742-6596/1833/1/012034 (Accessed: 21 January 2024).

14. *Advantages and disadvantages of logistic regression* (2023) *GeeksforGeeks*. Available at: https://www.geeksforgeeks.org/advantages-and-disadvantages-of-logistic-regression/ (Accessed: 21 January 2024).

15. (No date a) *Human activity recognition using K-nearest neighbor machine learning ...* Available at: https://www.researchgate.net/publication/354659475_Human_Activity_Recognition_Using_K-Nearest_Neighbor_Machine_Learning_Algorithm (Accessed: 21 January 2024).

16. (No date a) *[PDF] activity recognition using K-nearest neighbor algorithm on ...* Available at: https://www.semanticscholar.org/paper/ACTIVITY-RECOGNITION-USING-K-NEAREST-NEIGHBOR-ON-Kaghyan-Sarukhanyan/3f7b73943770691eb6d3f93905445738a8f436f7 (Accessed: 21 January 2024).

17. (No date a) *Activity recognition using cell phone accelerometers - semantic scholar*. Available at: https://www.semanticscholar.org/paper/Activity-recognition-using-cell-phone-Kwapisz-Weiss/fdbbbbb2e9a78e70de0e553792c3554942b7ee59 (Accessed: 21 January 2024).

18. Saxena, S. (2018) *Precision vs recall*, *Medium*. Available at: https://medium.com/@shrutisaxena0617/precision-vs-recall-386cf9f89488 (Accessed: 21 January 2024).

19. O'Sullivan, C. (2023) *Introduction to shap with python*, *Medium*. Available at: https://towardsdatascience.com/introduction-to-shap-with-python-d27edc23c454 (Accessed: 21 January 2024).

20. Kenton, W. (no date) *Shapley value definition and example of how it is applied*, *Investopedia*. Available at: https://www.investopedia.com/terms/s/shapley-

value.asp#:~:text=The%20Shapley%20value%20is%20a,a%20desired%20outcome%20or%20payoff. (Accessed: 21 January 2024).

21. O'Sullivan, C. (2023b) *Shap for binary and multiclass target variables*, *Medium*. Available at: https://towardsdatascience.com/shap-for-binary-and-multiclass-target-variables-ff2f43de0cf4 (Accessed: 21 January 2024).

22. Author links open overlay panelCharmi Jobanputra a *et al.* (2019) *Human activity recognition: A survey*, *Procedia Computer Science*. Available at: https://www.sciencedirect.com/science/article/pii/S1877050919310166 (Accessed: 21 January 2024).

23. Vrigkas, M., Nikou, C. and Kakadiaris, I.A. (2015) *A review of Human Activity Recognition Methods*, *Frontiers*. Available at: https://www.frontiersin.org/articles/10.3389/frobt.2015.00028/full (Accessed: 21 January 2024).

24. Gupta, N. *et al.* (2022) *Human activity recognition in Artificial Intelligence Framework: A narrative review - artificial intelligence review*, *SpringerLink*. Available at: https://link.springer.com/article/10.1007/s10462-021-10116-x (Accessed: 21 January 2024).

25. Miranda, L., Viterbo, J. and Bernardini, F. (2021) *A survey on the use of machine learning methods in context-aware middlewares for human activity recognition - artificial intelligence review*, *SpringerLink*. Available at: https://link.springer.com/article/10.1007/s10462-021-10094-0 (Accessed: 21 January 2024).

26. Hussain, I. *et al.* (2023) *An explainable EEG-based human activity recognition model using machine-learning approach and lime*, *MDPI*. Available at: https://www.mdpi.com/1424-8220/23/17/7452 (Accessed: 21 January 2024).

27. Hulstaert, L. (2018) *Understanding model predictions with lime*, *Medium*. Available at: https://towardsdatascience.com/understanding-model-predictions-with-lime-a582fdff3a3b (Accessed: 21 January 2024).

28. (No date) *Human action recognition | The Alan Turing Institute*. Available at: https://www.turing.ac.uk/research/research-projects/human-action-recognition (Accessed: 21 January 2024).

29. Bibbò, L. and Vellasco, M.M.B.R. (2023) *Human activity recognition (HAR) in healthcare*, *MDPI*. Available at: https://www.mdpi.com/2076-3417/13/24/13009 (Accessed: 21 January 2024).

30. Mehra, A. (2023) *Human activity recognition (HAR): Fundamentals, models, datasets*, *Labellerr*. Available at: https://www.labellerr.com/blog/human-activity-recognition/ (Accessed: 21 January 2024).

31. Liu, R. *et al.* (2023) *An overview of human activity recognition using wearable sensors: Healthcare and artificial intelligence*, *arXiv.org*. Available at: https://arxiv.org/abs/2103.15990 (Accessed: 21 January 2024).

32. Creative, C. (no date) *UK IOT and Iiot Development Services*, *Compsoft*. Available at: https://www.compsoft.co.uk/services/iot-and-iiot-development?gad_source=1 (Accessed: 21 January 2024).

33. *What is the internet of things (IOT)?* (no date) *What Is the Internet of Things (IoT)? | Oracle United Kingdom*. Available at: https://www.oracle.com/uk/internet-of-things/what-is-iot/#:~:text=What%20is%20IoT%3F,and%20systems%20over%20the%20internet. (Accessed: 21 January 2024).