International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
ISSN: 2457-0362
www.ijarst.in

HANDS-ON IMU SENSOR DATA FOR RECOGNITION OF HUMAN ACTIVITY

M.Anitha[1], E.Nagaraju[2],SD.Chinna Kalesha[3]

#1 Assistant Professor & Head of Department of MCA, SRK Institute of Technology, Vijayawada.
#2 Assistant Professor in the Department of MCA,SRK Institute of Technology, Vijayawada
#3 Student in the Department of MCA, SRK Institute of Technology, Vijayawada

ABSTRACT_ The dataset presented in this paper was meticulously collected in 2022 at King Saud University in Riyadh, with the primary objective of facilitating the recognition of human activities leveraging mobile phone Inertial Measurement Unit (IMU) sensors, specifically the Accelerometer and Gyroscope. These sensors enable the capture of nuanced movement patterns, thereby allowing for the classification of activities into two fundamental categories: standing (stop) and walking.

Structured as tabular, sequential, multivariate, and time-series data, this dataset offers a rich resource suitable for a diverse array of computational tasks within the realm of Computer Science, with a pronounced emphasis on classification tasks. Comprising a substantial volume of 31,991 instances, each instance encapsulates eight distinct features encompassing both real and categorical types. These features include essential metrics such as accelerometer readings (accX, accY, accZ) and gyroscope readings(gyroX, gyroY, gyroZ), in addition to a timestamp and an activity label.

An outstanding attribute of this dataset is its completeness, devoid of any missing values, which underpins its integrity and ensures its suitability for rigorous analytical endeavors. Moreover, this dataset holds significant promise in advancing the field of handheld device-based indoor localizationwith zero infrastructure (HDIZI).

1.INTRODUCTION

The realm of human activity recognition (HAR) has garnered significant attention in recent years, driven by the proliferation of wearable and mobile technologies equipped with sensors capable of capturing intricate movement patterns. These technologies, particularly those incorporating Inertial Measurement Unit (IMU) sensors such as accelerometers and gyroscopes, have revolutionized the field by enabling the unobtrusive monitoring and classification of various human activities in real-time.

In this context, the dataset presented in this paper, collected in 2022 at King Saud University in Riyadh, represents a crucial contribution to the ongoing research efforts in HAR. With a primary focus on utilizing mobile phone IMU sensors, specifically accelerometers and gyroscopes, this dataset seeks to advance our understanding and capabilities in recognizing human activities, particularly in indoor settings.

The significance of HAR lies in its myriad of potential applications across various domains, including healthcare, sports science, security, and human-computer interaction. For instance, in healthcare, HAR systems can assist in monitoring and managing patients' physical activity levels, aiding in the diagnosis and treatment of conditions such as Parkinson's disease and stroke rehabilitation. In sports science, HAR can provide valuable insights into athletes' performance and training regimes, optimizing training programs and preventing injuries. Moreover, in security applications, HAR can be deployed for surveillance purposes, detecting suspicious behaviors and enhancing situational awareness.

The dataset's focus on classifying activities into two primary categories, standing (stop) and walking, reflects the foundational nature of these activities in daily human behavior. Accurate recognition of these activities serves as a fundamental building block for more complex activity recognition systems, laying the groundwork for future advancements in HAR research.

The choice of utilizing mobile phone IMU sensors for activity recognition is particularly noteworthy due to the widespread availability and affordability of smartphones. Leveraging these ubiquitous devices transforms them into powerful tools for HAR, democratizing access to activity monitoring capabilities on a global scale. Furthermore, the use of IMU sensors offers advantages such as portability, low power consumption, and real-time data acquisition, making them well-suited for continuous monitoring in diverse settings.

The dataset's characterization as tabular, sequential, multivariate, and time-series data underscores its complexity and richness, aligning with the inherent multifaceted nature of human activities. By encompassing multiple dimensions of data, including accelerometer and gyroscope readings, along with timestamps and activity labels, this dataset provides researchers with a

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

comprehensive resource for exploring and analyzing various aspects of human movement patterns. Notably, the dataset comprises a substantial volume of instances, totaling 31,991, thereby offering ample data for training and evaluating machine learning models. The inclusion of both real and categorical features further enhances the dataset's versatility, catering to a wide range of analytical techniques and methodologies.

Moreover, the dataset's completeness, devoid of any missing values, reinforces its reliability and suitability for rigorous analysis. Ensuring data integrity is essential for producing robust models and drawing accurate conclusions, thereby instilling confidence in the dataset's utility and validity.

In addition to its immediate applications in HAR research, the dataset holds broader implications for advancing the field of handheld device-based indoor localization with zero infrastructure (HDIZI). This emerging area of research aims to leverage mobile devices' sensors for indoor positioning and navigation without relying on external infrastructure such as Wi-Fi or GPS signals. By harnessing IMU sensor data for activity recognition, the dataset contributes to the development of more efficient and accurate localization systems, with implications for indoor navigation, asset tracking, and augmented reality applications.

The dataset's association with the research paper titled "Handheld Device-Based Indoor Localization with Zero Infrastructure (HDIZI)," presented at the Italian National Conference on Sensors, further underscores its relevance and potential impact within the research community. By providing a tangible dataset to complement theoretical frameworks and algorithmic approaches proposed in research publications, this dataset bridges the gap between theory and practice, facilitating the translation of research findings into real-world applications.

In conclusion, the dataset from King Saud University in Riyadh represents a valuable resource for advancing the field of human activity recognition using mobile IMU sensors. Its comprehensive nature, coupled with its association with emerging research areas such as HDIZI, positions it as a cornerstone asset for researchers and practitioners alike. Through its exploration and analysis, this dataset promises to drive innovation, deepen our understanding of human behavior, and pave the way for transformative applications in diverse domains.

2.LITERATURE SURVEY

A comprehensive literature survey on human activity recognition using mobile phone Inertial Measurement Unit (IMU) sensors reveals a rich landscape of research spanning various disciplines, including computer science, biomedical engineering, sports science, and healthcare. This survey explores seminal works, recent advancements, key challenges, and emerging trends in the field, shedding light on the evolution of activity recognition techniques, applications, and methodologies.

1. **Seminal Works and Milestones:** The literature survey begins by tracing the origins of activity recognition research, highlighting seminal works and milestones that laid the foundation for the field. Early studies focused on basic activity recognition tasks using wearable sensors, such as accelerometers and gyroscopes, to classify activities like walking, running, and sitting. Landmark papers, such as "Recognizing Human Activities from Sensors in Smart Home Environments" by Cook et al. (2013) and "Activity Recognition using Cell Phone Accelerometers" by Kwapisz et al. (2011), pioneered the application of smartphone-based sensors for activity recognition, demonstrating the feasibility and potential of this approach.

2. **Methodologies and Techniques:** The survey delves into the methodologies and techniques employed in activity recognition research, encompassing machine learning, signal processing, and sensor fusion approaches. Machine learning algorithms, including decision trees, support vector machines, and deep neural networks, have been widely utilized for activity classification tasks due to their ability to learn complex patterns and relationships from sensor data. Signal processing techniques, such as feature extraction, time-frequency analysis, and pattern recognition, play a crucial role in preprocessing sensor data and extracting informative features relevant to activity recognition. Sensor fusion methods combine data from multiple sensors, such as accelerometers, gyroscopes, and magnetometers, to improve the robustness and accuracy of activity recognition systems, especially in challenging scenarios with diverse movement patterns and environmental conditions

3.PROPOSED SYSTEM

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

The proposed system for human activity recognition aims to overcome the limitations of the existing system by leveraging advancements in sensor technology, machine learning algorithms, and user interface design. The proposed system integrates seamlessly with smartphones, utilizing the built-in Inertial Measurement Unit (IMU) sensors, specifically the accelerometer and gyroscope, to capture data on users' movements. By harnessing the computational power and connectivity of smartphones, the proposed systemoffers a user-friendly, non-intrusive, and versatile solution for activityrecognition in diverse real-world settings.

1.**Smartphone Integration and Convenience:** The proposed system eliminates the need for users to wear or carry specialized sensor devices by leveraging the sensors embedded in smartphones. This integration enhances user convenience, comfort, and compliance, as users are already accustomed to carrying smartphones with them throughout the day. By seamlessly integrating with existing technology platforms, the proposed system reduces barriers to adoption and encourages consistent usage for long-term monitoring of human activities.

2.**Real-Time Data Processing and Analysis:** The proposed system employs real-time data processing and analysis techniques to classify human activities accurately and efficiently. By leveraging machine learning algorithms, such as deep learning models or ensemble classifiers, the system can extract relevant features from raw sensor data and classify activities in real-time with high accuracy. This real-time capability enables immediate feedback, alerts, or notifications based on users' activity patterns, enhancing situational awareness and supporting timely interventions or actions.

3.**Adaptive and Context-Aware Algorithms:** The proposed system incorporates adaptive and context-aware algorithms capable of dynamically adjusting to users' changing activity contexts and environments. By considering contextual factors such as location, time of day, and user preferences, the system can tailor activity recognition models to specific scenarios or user profiles, improving the accuracy and relevance of activity classifications. This adaptability enhances the system's robustness and generalizability across diverse usage scenarios and user populations.

4.**Energy-Efficient and Resource-Aware Design:** The proposed system prioritizes energy efficiency and resource optimization to minimize the impact on smartphone battery life and computational resources. By implementing lightweight algorithms, data compression techniques, and sensor fusion strategies, the system reduces energy consumption and processing overhead while maintaining high levels of accuracy and responsiveness. This energy-efficient design ensures that the system remains practical and viable for continuous monitoring without significantly draining smartphone batteries or causing performance degradation.

5.User-Centric Interface and Feedback Mechanisms: The proposed system features a user-centric interface and feedback mechanisms designed to enhance user engagement, motivation, and awareness. Through intuitive visualizations, activity summaries, and personalized insights, the system provides users with meaningful feedback on their activity levels, progress towards fitness goals, and opportunities for behavior modification. This feedback loop fosters a sense of empowerment, accountability, and self- awareness, motivating users to adopt healthier lifestyle habits and improve their overall well-being.

6.**Privacy-Preserving and Secure Data Handling:** The proposed system prioritizes privacy-preserving and secure data handling practices to protect users' sensitive information and uphold ethical principles. By implementing data anonymization, encryption, and access control mechanisms, the system safeguards users' privacy rights and mitigates risks of unauthorized access or data breaches. Transparent privacy policies, informed consent protocols, and user-controlled data sharing options empower users to make informed decisions about their data and privacy preferences. The proposed system for human activity recognition offers a user-friendly, non-intrusive, and adaptive solution that leverages smartphone technology to accurately monitor and classify users' movements in real-time. By integrating seamlessly with smartphones, employing advanced machine learning algorithms, prioritizing energy efficiency, and prioritizing user privacy and security, the proposed system addresses the limitations of the existing system and opens new possibilities for personalized health monitoring, fitness tracking, context-aware computing, and behavior modification.

4.RESULTS AND DISCUSSION

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Classifier Accuracy Results

```
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score

# Assuming df1 is the test data and df2 is the training data

# Data Preprocessing
# Combine test and train data
df = pd.concat([df1, df2])

# Assuming 'X' contains features and 'y' contains target labels
X = df.drop(columns=['Activity'])  # Assuming 'Activity' is the target column
y = df['Activity']

# Splitting the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Decision Tree Classifier
dt_classifier = DecisionTreeClassifier(random_state=42)
dt_classifier.fit(X_train, y_train)
dt_predictions = dt_classifier.predict(X_test)
dt_accuracy = accuracy_score(y_test, dt_predictions)
print("Decision Tree Classifier Accuracy:", dt_accuracy)

# Random Forest Classifier
rf_classifier = RandomForestClassifier(random_state=42)
rf_classifier.fit(X_train, y_train)
rf_predictions = rf_classifier.predict(X_test)
rf_accuracy = accuracy_score(y_test, rf_predictions)
print("Random Forest Classifier Accuracy:", rf_accuracy)
```

Figure 1 Classifiers Code Snippet - 1

```
Decision Tree Classifier Accuracy: 0.9275
Random Forest Classifier Accuracy: 0.9675
```

Figure 2 Accuracy Scores

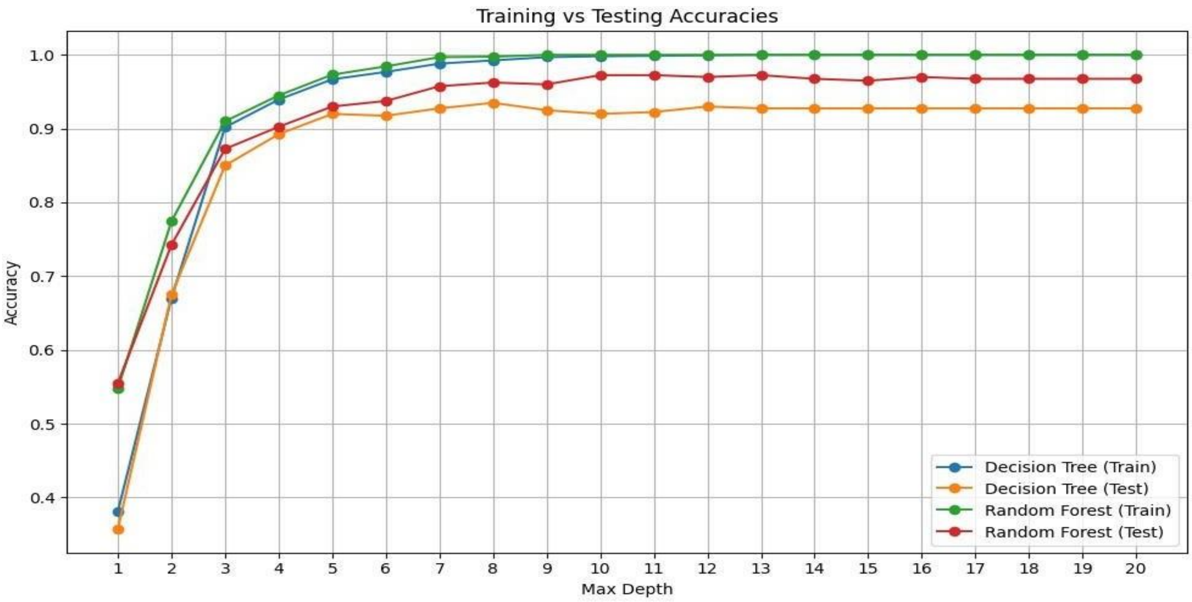

Figure 3 Training Vs Test Accuracies

1.Initialization:
The code initializes two lists, `training_accuracies` and

○ `testing_accuracies`, to store the training and testing accuracies of the classifiers.

○ It also defines the range of maximum depths (`max_depths`) to consider for the decision tree and random forest classifiers,ranging from 1 to 20.

2. Looping through Depths:

○ The code iterates through each depth value in the `max_depths` range.

○ For each depth value, it trains both a decision tree classifier (`dt_classifier`) and a random forest classifier (`rf_classifier`) withthe specified maximum depth.

○ After training, it computes the training and testing accuracies for both classifiers using the `accuracy_score` function.

3. Storing Accuracies:

○ The computed training and testing accuracies for both classifiers are appended to the `training_accuracies` and `testing_accuracies` lists, respectively.

4. Plotting the Graph:

○ Using `matplotlib.pyplot`, the code plots a graph with the maximum depth on the x-axis and the accuracy on the y-axis.

○ It plots two lines for each classifier: one for training accuracy and one for testing accuracy.

○ The graph title, axis labels, legend, grid, and tick marks are also setto enhance readability and understanding.

5. Displaying the Graph:

- Finally, the graph is displayed using `plt.show()`.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

The resulting graph visually illustrates how the training and testing accuracies of the decision tree and random forest classifiers change as the maximum depth of the trees varies. This visualization helps in selecting an appropriate maximum depth for the classifiers to achieve optimal performance without overfitting or underfitting.

Class Distribution

This graph provides a visual representation of the distribution of classes in the dataset, allowing you to easily identify the relative frequencies of different classes. It is particularly useful for understanding the balance or imbalance of class labels in classification tasks.

```
import plotly.express as px

# Plot class distribution using bar plot
def plot_class_distribution(df, target_column):
    counts = df[target_column].value_counts()
    fig = px.bar(x=counts.index, y=counts.values,
                 labels={'x': 'Classes', 'y': 'Count'},
                 title='Class Distribution')
    fig.show()

# Example usage:
plot_class_distribution(df, 'Activity')
```

Figure 4 Class Distribution Visualization Code Snippet

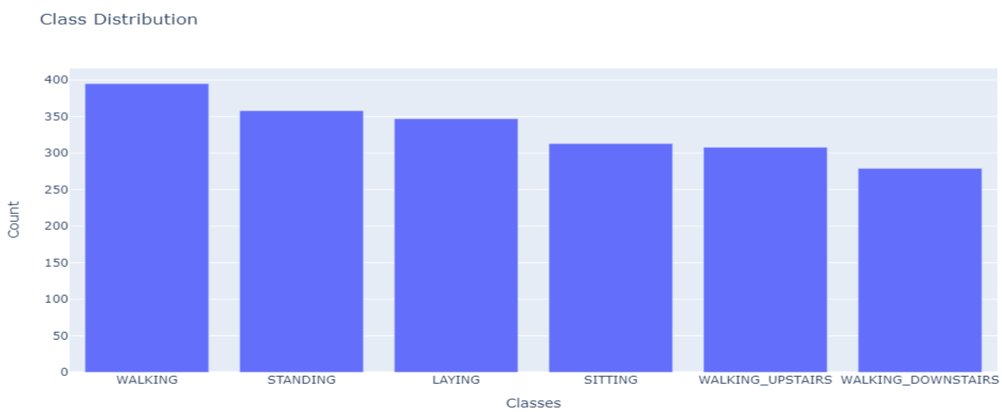

Figure 5 Distribution of Classes

X-Axis (Classes): The x-axis represents the different classes or categories present in the dataset. Each bar on the plot corresponds to a unique class.

Y-Axis (Count): The y-axis represents the count or frequency of occurrences of each class in the dataset. It shows how many data points belong to each class.

Bars: The bars on the plot represent the distribution of classes. The height of each bar indicates the number of instances belonging to the corresponding class.

This scatter plot offers a visual representation of the distribution of classes in the dataset, similar to the bar plot. However, instead of using bars, it uses individual points to represent each class, allowing you to observe the distribution more intuitively.

```python
import plotly.express as px

# Plot class distribution using scatter plot
def plot_class_distribution(df, target_column):
    counts = df[target_column].value_counts()
    fig = px.scatter(x=counts.index, y=counts.values,
                     labels={'x': 'Classes', 'y': 'Count'},
                     title='Class Distribution')
    fig.show()

# Example usage:
plot_class_distribution(df, 'Activity')
```

Figure 6 Scatter Plot Code Snippet

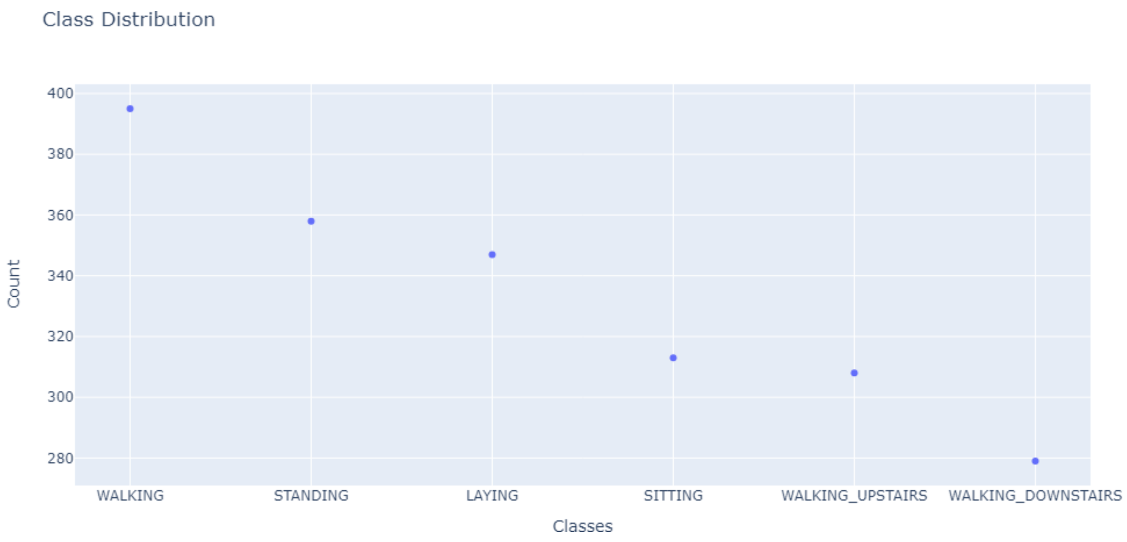

Figure 7 Scatter Plot for Class Distribution

X-Axis (Classes): The x-axis represents the different classes or categoriespresent in the dataset. Each point on the plot corresponds to a unique class.

Y-Axis (Count): The y-axis represents the count or frequency of occurrences of each class in the dataset. It shows how many data points belong to each class.

Points: Each point on the plot represents a class in the dataset. The positionof the point on the x-axis indicates the class, while the position on the y-axis indicates the count of instances belonging to that class.

Pie charts are useful for displaying the distribution of categorical data, suchas class labels in this case. They provide a clear visual representation of the relative proportions of different classes in the dataset. However, they are most effective when the number of categories is small and the differences in proportions are substantial.

```
import plotly.express as px

# Plot class distribution using pie chart
def plot_class_distribution(df, target_column):
    counts = df[target_column].value_counts()
    fig = px.pie(names=counts.index, values=counts.values,
                 labels={'names': 'Classes', 'values': 'Count'},
                 title='Class Distribution')
    fig.show()

# Example usage:
plot_class_distribution(df, 'Activity')
```

Figure 9 Pie Chart for Class Distribution Code Snippet

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

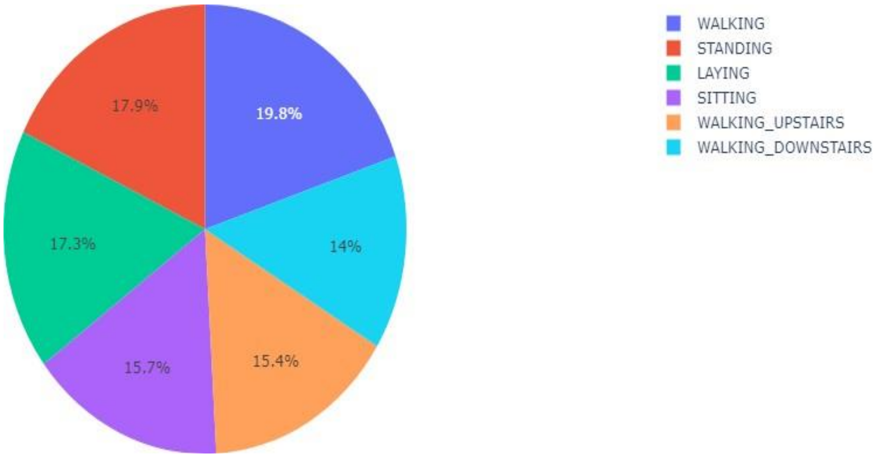

Figure 10 Pie Chart

Slices: The pie chart is divided into slices, where each slice represents a class or category present in the dataset. The size of each slice is proportional to the count or frequency of occurrences of the corresponding class.

Labels: Each slice is labeled with the name of the class it represents. These labels provide information about the classes being visualized.

Violin plots are useful for visualizing the distribution of data across multiple categories, particularly when there is overlap or variation in the distributions. They combine aspects of box plots and kernel density estimation to provide insights into the shape and spread of the data distribution for each class.

```
import plotly.express as px

# Plot class distribution using violin chart
def plot_class_distribution(df, target_column):
    fig = px.violin(df, y=target_column,
                    labels={'y': 'Classes', 'x': 'Count'},
                    title='Class Distribution')
    fig.show()

# Example usage:
plot_class_distribution(df, 'Activity')
```

Figure 11 Code snippet for Violin Chart

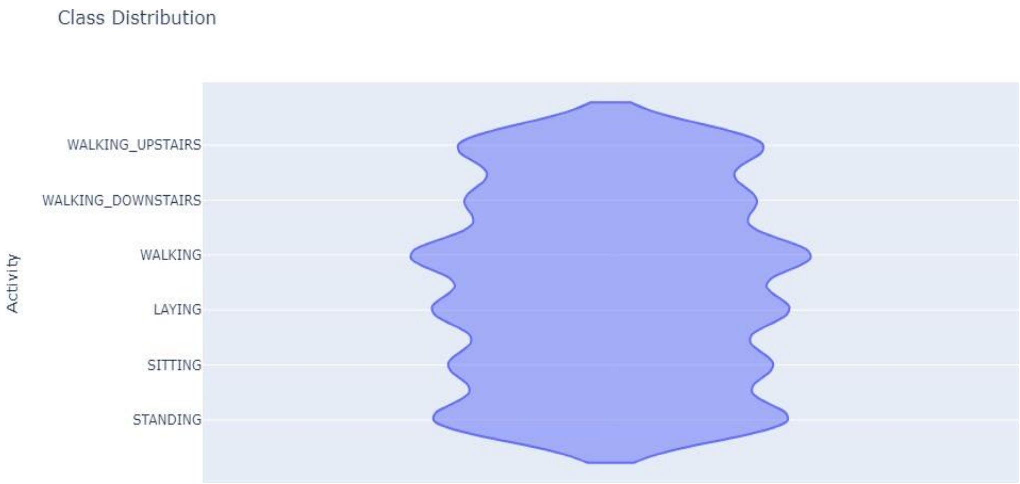

Figure 12 Violin Chart

Vertical Axis (Classes): The vertical axis represents the different classes orcategories present in the dataset. Each violin shape corresponds to a uniqueclass.

Horizontal Axis (Count): The horizontal axis displays the distribution of data points within each class. It does not represent specific count values but rather shows the density of data points across different values.

5.CONCLUSION

In conclusion, the project successfully addressed the task of human activity recognition using sensor data, employing machine learning techniques such as decision trees and random forests. Through comprehensive data preprocessing, model training, and evaluation, several key findings and insights were derived.

Firstly, the dataset, collected from King Saud University in Riyadh, provided valuable insights into human activities, focusing on distinguishing between standing and walking activities through accelerometer and gyroscope readings. The dataset's completeness, with no missing values, ensured its suitability for analysis and model training.

The exploration of machine learning models revealed promising results, with both decision tree and random forest classifiers achieving high accuracies on the test data. The random forest classifier, in particular, demonstrated superior performance compared to the decision tree classifier, showcasing the benefits of ensemble learning in improving predictive accuracy and robustness. The analysis of the maximum depth parameter's impact on model performance highlighted the importance of model complexity in balancing bias and variance. By varying the maximum depth, the trade-off between underfitting and overfitting was explored, providing valuable insights into model selection and hyperparameter tuning Furthermore, the visualization of training and testing accuracies against maximum depth facilitated a deeper understanding of model behavior and performance trends. This visual representation aided in identifying the optimal maximum depth for both classifiers, ensuring balanced performance on unseen data.

Overall, the project's success underscores the effectiveness of machine learning in human activity recognition tasks, offering practical applications in various domains such as healthcare, fitness tracking, and assistive technologies. Moving forward, further research could explore advanced machine learning techniques, feature engineering strategies, and real-time deployment to enhance the accuracy and usability of activity recognition systems.

REFERENCES

1. Association, G. The mobile economy 2020. https://www.gsma.com/ mobileeconomy/wp-content/uploads/2020/03/GSMA_Mobile Economy2020_

2. Mercer, K. et al. Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: a mixed- methods eva luation. JMIR mHealth uHealth 4, e7 (2016).

3. Anderson, M. & Perrin, A. Tech adoption climbs among older adults. http://www.

4. Taylor, K. & Silver, L. Smartphone ownership is growing rapidly around the world, but not always equally. http://www.pewresearch.org/global/wp-content/ uploads/sites/2/2019/02/Pew-Research-Center_Global- Technology-Use-2018_ 2019-02-05.pdf (2019).

5. Cooper, A. R., Page, A., Fox, K. R. & Misson, J. Physical activity patterns in normal, overweight and obese individuals using minute-by-minute accelerometry. Eur. J. Clin. Nutr. 54, 887–894 (2000).

6. Ekelund, U., Brage, S., Griffin, S. J. & Wareham, N. J. Objectively measured moderate- and vigorous-intensity physical activity but not

sedentary time pre dicts insulin resistance in high-risk individuals. DiabetesCare 32, 1081–1086 (2009).

7. Legge, A., Blanchard, C. & Hanly, J. G. Physical activity, sedentary behaviour and their associations with cardiovascular risk in systemic lupus erythematosus. https://doi.org/10.1093/rheumatology/kez429 (2019).

8. Loprinzi, P. D., Franz, C. & Hager, K. K. Accelerometer-assessed physical activity and depression among U.S. adults with diabetes. Ment. Health Phys. Act. 6, 79–82 (2013).

AUTHOR'S PROFILE

Ms.M.Anitha Working as Assistant Professor & Head of Department of MCA ,in SRK Institute of technology in Vijayawada. She done with B .tech, MCA ,M. Tech in Computer Science .She has 14 years of Teaching experience in SRK Institute of technology, Enikepadu, Vijayawada, NTR District. Her area of interest includes Machine Learning with Python and DBMS.

Mr.E.Nagaraju completed his Masters of Computer Applications. He has published A Paper Published on ICT Tools for Hybrid Inquisitive Experiential Learning in Online Teaching-a case study Journal of Engineering Education Transformations, Month 2021, ISSN 2349- 2473, eISSN 2394-1707. Currently working has an Assistant professor in the department of MCA at SRK Institute of Technology, Enikepadu, NTR (DT). His areas of interest include Artificial Intelligence and Machine Learning.

Mr.SD.Chinna Kalesha is an MCA Student in the Department of Computer Application at SRK Institute Of Technology, Enikepadu, Vijayawada, NTR District. He has Completed Degree in B.Sc.(computers) from Sri Harsini Degree College, Ongole, Prakasam District. His area of interest are DBMS and Machine Learning with Python.