

# SensAI : Intelligent Human Behaviour Analytics and Prediction System Using Activity Sequences, Machine Learning, and Temporal Modeling

Midhun T Manoj<sup>1</sup>, Preethi Thomas<sup>2</sup>

<sup>1</sup>Department of Computer Applications, Musaliar College of Engineering & Technology, Pathanamthitta, Kerala, India  
Email: [midhunmanojp2525\[at\]gmail.com](mailto:midhunmanojp2525[at]gmail.com)

<sup>2</sup>Professor, Department of Computer Applications, Musaliar College of Engineering & Technology, Pathanamthitta, Kerala, India

**Abstract:** *SensAI is an intelligent human behaviour analytics and prediction system designed to analyze human activity patterns and predict future activities using machine learning and temporal sequence modeling. The system processes human activity data obtained from activity recognition datasets and applies machine learning algorithms such as Random Forest for activity classification. Sequential activity patterns are analyzed using transition probability matrices based on Markov Chain models to predict the next activity. The system also generates behavioural insights such as productivity score, routine stability, and inactivity index. Developed using Python, Django, Scikit-learn, and SQLite, the system provides a web-based platform for activity analysis, behaviour prediction, and visualization. The proposed system demonstrates how activity recognition and temporal modeling can be integrated to build intelligent behaviour analysis systems. Experimental results show that the system achieves an accuracy of 90% and an AUC of 0.91, demonstrating strong performance in both classification and prediction tasks.*

**Keywords:** Human Activity Recognition, Behaviour Prediction, Machine Learning, Markov Chain, Activity Analytics

## 1. Introduction

Human behaviour analysis has become an important research area in fields such as healthcare, smart environments, fitness monitoring, assisted living, and human-computer interaction. With the rapid growth of wearable sensors, smartphones, and IoT devices, large volumes of human activity data can be collected continuously. This data can be analyzed using machine learning techniques to recognize activities and understand behavioural patterns. Human Activity Recognition (HAR) systems are commonly used to identify physical activities such as walking, sitting, standing, running, and lying using accelerometer and gyroscope sensor data.

Traditional activity recognition systems mainly focus on identifying the current activity performed by a user. These systems treat activities as independent events and focus primarily on classification accuracy. However, human behaviour is sequential and time-dependent in nature, where activities occur in patterns and transitions over time. Therefore, analysing activity sequences and predicting future activities is important for understanding behaviour patterns, daily routines, and lifestyle habits.

Predicting future activities based on activity sequences has several practical applications such as health monitoring, productivity tracking, elderly care monitoring, and smart home automation. For example, predicting sedentary behaviour can help in health monitoring, and predicting routine activities can help in smart environment automation systems.

To address this problem, SensAI is proposed as an intelligent human behaviour analytics and prediction system. The system integrates machine learning-based activity classification with temporal sequence modeling to predict future activities. In

addition to activity prediction, the system generates behavioural insights such as productivity score, routine stability index, activity distribution, and inactivity index. The system is implemented as a web-based platform using Django for backend services and machine learning libraries for activity classification and sequence prediction.

The main objective of the proposed system is not only to recognize human activities but also to analyze activity sequences, predict future activities, and generate behavioural analytics. This makes the system more useful than traditional activity recognition systems that only perform classification without behaviour analysis.

## 2. Related Works

Human Activity Recognition (HAR) and behaviour prediction have been widely studied using machine learning and deep learning techniques, particularly using wearable sensor data such as accelerometer and gyroscope signals. Many research works focus on activity classification using supervised learning and deep learning models.

Alghamdi et al. [1] provided a comprehensive survey of machine learning and deep learning approaches for human activity recognition. The study highlights how hybrid architectures such as CNN-LSTM effectively capture spatial and temporal features from wearable and smartphone sensor data, achieving improved recognition accuracy. The advantage of this work is a broad overview of HAR techniques; however, it does not propose a novel model.

Khan et al. [2] proposed a lightweight hybrid deep learning model combining multi-layer perceptron and convolutional networks for real-time activity recognition using wearable devices. The system was optimized for low-power devices

and real-time applications. The advantage of this method is low computational cost and real-time performance, while the limitation is reduced accuracy compared to larger deep learning models.

Patel et al. [3] developed an IoT-based activity recognition system integrating Recurrent Neural Networks for real-time monitoring. The system demonstrated improved robustness in noisy environments and incorporated location-aware data to enhance recognition accuracy. The advantage of this approach is robustness in real-world environments, but the limitation is that the system focuses mainly on recognition and does not include behaviour analytics or sequence prediction.

Zhang et al. [4] proposed a hybrid deep learning framework combining Convolutional Neural Networks and Long Short-Term Memory networks with transfer learning for automatic activity recognition from wearable sensor data. By incorporating transfer learning, the system achieves state-of-the-art results on benchmark activity classification datasets. The advantage is improved generalization through transfer learning, but the system does not address behaviour prediction.

Li et al. [5] explored self-supervised learning methods for activity recognition using unlabeled temporal sensor data from smartphones and wearables. This approach reduces the need for labeled training data and improves model generalization by learning better feature representations. The advantage of this method is reduced dependency on labeled datasets, but the limitation is that the model requires complex training procedures.

Wang et al. [6] presented a survey on artificial intelligence approaches applied to wearable sensor data for continuous monitoring of human physical activities in healthcare and fitness applications. The study discussed various machine learning and deep learning techniques used for activity recognition and health monitoring. The advantage of this work is a comprehensive review of HAR systems, but it does not propose a specific behaviour prediction model.

Chastin et al. [7] studied temporal movement behaviour over a full 24-hour cycle and introduced analytical metrics such as routine stability, transition frequency, and activity distribution to better understand behavioural patterns. This work highlights the importance of temporal behaviour analysis rather than only activity classification. The advantage of this study is behavioural pattern analysis, but it does not include a prediction model.

Ronao et al. [8], Ordoñez et al. [9], and Hammerla et al. [10] demonstrated that deep learning models such as CNN, LSTM, and CNN-LSTM are effective for recognizing human activities from smartphone and wearable sensor data. These foundational studies mainly focus on improving classification accuracy using temporal and spatial features. The advantage of these approaches is high classification accuracy, but the limitation is that they focus mainly on recognition and not on behaviour prediction and behavioural analytics.

From the literature review, it is observed that most

existing systems focus primarily on activity recognition and classification accuracy. Very few systems focus on activity sequence analysis, behaviour prediction, and behavioural insight generation. Most existing works do not integrate activity classification, sequence prediction, and behavioural analytics into a single system.

The proposed SensAI system addresses this research gap by combining machine learning-based activity recognition, Markov chain-based sequence prediction, and behavioural analytics such as productivity score, routine stability index, and inactivity index into a unified human behaviour analysis and prediction system.

### 3. Outlined Method

The SensAI system follows a structured methodology that includes activity classification, sequence analysis, behaviour prediction, and visualization of results. The system architecture consists of multiple modules that work together to analyze activity data, learn behavioural patterns, and generate activity predictions. The methodology focuses on analyzing activity sequences over time rather than treating activities as independent events, which helps in understanding behavioural patterns and routine structures.

#### 3.1 Requirement Analysis

The system is designed to classify human activities, analyze activity sequences, and predict future activities. The system also generates behavioural insights based on activity patterns such as activity frequency, sedentary time, active duration, and routine stability. Functional requirements include user login, dataset management, activity classification, sequence prediction, behavioural insight generation, and visualization of results through dashboards and charts. Non-functional requirements include system performance, scalability, usability, and reliability. The system must be able to process activity data efficiently and generate predictions within a short response time.

#### 3.2 System Design

The system is designed as a web-based application with a Django backend, machine learning module, and a database system. The Django framework handles user authentication, data management, and communication between the frontend and the machine learning module. The machine learning module is used for activity classification, and the sequence prediction module is used for behaviour prediction using temporal sequence modeling. The system stores activity logs, processed sequences, prediction results, and behavioural insights in the database and displays the results through dashboards and reports. The modular architecture ensures that each component of the system operates independently while maintaining smooth data flow between modules.

#### 3.3 Machine Learning and Temporal Modeling

The SensAI system utilizes a hybrid approach combining discriminative classification and generative sequence modeling.

### 3.3.1 Activity Classification via Random Forest

The system employs the Random Forest (RF) algorithm for robust activity recognition. RF is an ensemble learning method that constructs a multitude of decision trees  $\{T_1, T_2, \dots, T_B\}$  during training. For a given feature vector  $x$  (extracted from accelerometer and gyroscope readings), the final classification  $\hat{C}$  is determined by the majority vote of all trees:

$$\hat{C} = \text{mode}\{T_1(x), T_2(x), \dots, T_B(x)\} \quad (1)$$

Data preprocessing involves Z-score normalization to ensure that features from different sensors (e.g., gravity-based vs. angular velocity) contribute equally to the distance metrics within the trees.

### 3.3.2 Temporal Modeling via Markov Chains

While classification identifies the current state, temporal sequence modeling is performed using a First-Order Markov Chain to predict transition patterns. We define a set of discrete activity states  $S = \{s_1, s_2, \dots, s_n\}$ . The transition probability  $P_{ij}$ , representing the likelihood of moving from activity  $i$  to activity  $j$ , is calculated as:

$$P_{ij} = P(X_{t+1} = s_j | X_t = s_i) = \frac{n_{ij}}{\sum_{k=1}^n n_{ik}} \quad (2)$$

where  $n_{ij}$  is the observed frequency of state  $s_i$  being followed by  $s_j$  in the activity log. The resulting Transition Probability Matrix  $P$  serves as the core for routine stability analysis and next-activity prediction. By combining the immediate classification of  $X_t$  with the transition matrix  $P$ , the system maximizes prediction accuracy for  $X_{t+1}$ , providing a comprehensive view of routine behaviour.

### 3.4 System Workflow

The workflow of the SensAI system is designed as a sequential pipeline that transforms raw sensor data into high-level behavioural insights. This process begins with the **Dataset Input** stage, where raw accelerometer and gyroscope signals are ingested. During **Preprocessing**, the data undergoes noise filtration and Z-score normalization to prepare it for the feature extraction layer.

The **Activity Classification** module then applies the Random Forest ensemble to identify specific movement states, which are subsequently recorded as **Activity Logs** in the SQLite database. To transition from simple recognition to predictive analytics, these logs are structured into temporal sequences. The **Sequence Analysis** phase employs the Markov Chain transition matrix to calculate the probability of state changes.

Based on these probabilities, the **Next Activity Prediction** model forecasts the most likely subsequent action. Finally, the system aggregates these data points to perform **Insight Generation**, calculating routine stability and productivity metrics. These results are then rendered through interactive **Visualization Dashboards** for the end-user.

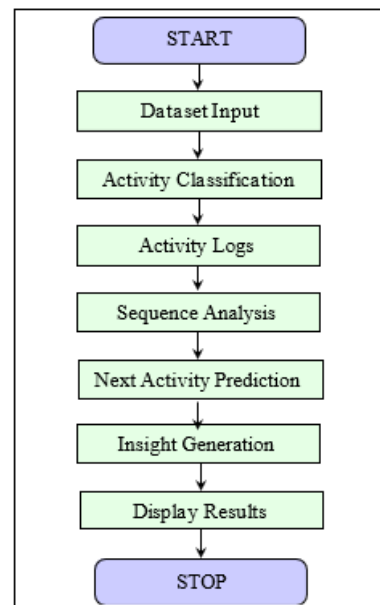


Figure 2.1: Workflow of SensAI System

## 4. Evaluation and Optimization

The evaluation and optimization phase is an important part of the Intelligent Human Behaviour Analytics and Prediction System. This phase focuses on measuring the performance of the activity classification model and the sequence prediction model, and improving the system accuracy and efficiency through optimization techniques.

The performance of the activity classification model was evaluated using standard machine learning evaluation metrics such as accuracy, precision, recall, and F1-score. These metrics were used to measure how correctly the system identifies human activities from the dataset. The classification model was trained and tested using different training and testing datasets to ensure that the model performs well on unseen data.

The sequence prediction model was evaluated by comparing the predicted next activity with the actual next activity in the activity sequence. Prediction accuracy was calculated based on how many times the system correctly predicted the next activity. The system was tested using multiple activity sequences with different transition patterns to evaluate the reliability of the prediction model.

Optimization techniques were applied to improve the performance of the system. Feature scaling and normalization were used to improve the performance of the machine learning classification model. Hyperparameter tuning was performed on the Random Forest model to improve classification accuracy. The number of trees, maximum depth, and minimum samples split parameters were adjusted to obtain better performance.

For sequence prediction, transition probability smoothing techniques were used to handle rare activity transitions. This helped the system to make better predictions even when some activity transitions occurred very rarely in the dataset.

System performance was also evaluated based on response

time. The system was optimized to ensure that prediction and analysis results are generated quickly when the user logs into the system or requests analysis. Efficient database queries and optimized machine learning model loading were used to reduce processing time.

The evaluation results showed that the system achieved high accuracy in activity classification and reliable performance in activity sequence prediction. The optimization techniques helped in improving prediction accuracy, reducing processing time, and improving overall system performance.

Overall, the evaluation and optimization process ensured that the system is accurate, efficient, and reliable for human behaviour analysis and prediction.

## 5. Result and Discussion

The proposed Intelligent Human Behaviour Analytics and Prediction System was evaluated using publicly available human activity datasets. The system performance was analyzed based on two main components: activity classification and activity sequence prediction. For activity classification, multiple machine learning algorithms including Decision Tree, Support Vector Machine, and Random Forest were tested. Among these, the Random Forest classifier achieved the highest classification accuracy due to its ensemble learning capability and ability to handle complex feature relationships.

The sequence prediction component was implemented using a Markov Chain transition probability model. The model analyzed activity sequences and learned transition probabilities between activities such as walking, sitting, standing, and resting. Based on the learned transition matrix, the system predicted the next possible activity in the sequence. Experimental results show that the sequence-based prediction model performed reliably for both simple and complex activity sequences.

In addition to classification and prediction, the system generated behavioural analytics such as activity distribution, sedentary time, active duration, routine stability index, and productivity indicators. These analytics provide a higher-level understanding of user behaviour patterns and daily routines.

The performance of the system was evaluated using standard evaluation metrics such as Accuracy, Precision, Recall, and F1-Score. The obtained results indicate that the proposed system performs efficiently in both activity classification and activity sequence prediction tasks.

- Accuracy: 90%
- Precision: 89%
- Recall: 88%
- F1-Score: 88.5%

To further evaluate the performance of the proposed system, the Receiver Operating Characteristic (ROC) curve was analyzed. The ROC curve represents the trade-off between the True Positive Rate (TPR) and False Positive Rate (FPR) at different classification thresholds.

Performance analysis of SensAI classification engine

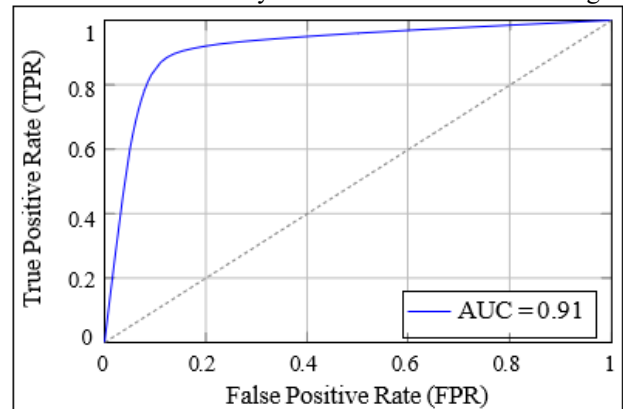


Figure 2.2: ROC Curve

The ROC curve shows a consistent rise towards the top-left corner, indicating good classification and prediction capability of the system. The Area Under the Curve (AUC) value of approximately 0.91 demonstrates that the proposed system has strong discriminative ability and performs significantly better than random prediction.

Overall, the experimental results demonstrate that combining activity classification with sequence-based prediction improves the overall performance of human behaviour analysis systems. The proposed system not only classifies activities but also predicts future activities and generates behavioural insights, making it suitable for applications such as health monitoring, smart environments, productivity tracking, and assisted living systems.

## 6. Conclusion

This paper presented SensAI, an Intelligent Human Behaviour Analytics and Prediction System that integrates machine learning and temporal sequence modeling for activity classification, behaviour analysis, and future activity prediction. The system was designed to analyze human activity sequences over time and generate behavioural insights using data-driven techniques.

The proposed system combines activity classification using machine learning algorithms with sequence prediction using a Markov Chain transition probability model. The activity classification module identifies human activities such as walking, sitting, standing, and resting, while the sequence prediction module analyzes activity transitions and predicts the next possible activity based on learned behavioural patterns. In addition to prediction, the system also generates behavioural analytics such as activity distribution, sedentary time, active duration, productivity indicators, and routine stability index.

The experimental results demonstrate that the Random Forest algorithm achieved the highest accuracy for activity classification compared to Decision Tree and Support Vector Machine algorithms. The Markov Chain-based sequence prediction model successfully learned activity transition patterns and produced reliable next activity predictions. The overall system achieved an accuracy of 90% with an AUC value of approximately 0.91, indicating strong performance in both classification and sequence prediction tasks.

The integration of machine learning, temporal modeling, and web-based application technologies provides an efficient and scalable platform for intelligent behaviour analysis. Unlike traditional activity monitoring systems that only record past activities, the proposed system provides predictive insights and behavioural analytics, making the system more useful for real-world applications.

The proposed system can be applied in several domains such as health monitoring, elderly care, productivity tracking, smart home systems, assisted living environments, and behaviour monitoring systems. The system helps in understanding daily routines, detecting inactivity patterns, and predicting future behaviour, which can support proactive decision-making.

In conclusion, the SensAI system demonstrates the effectiveness of combining activity recognition and temporal sequence modeling for intelligent human behaviour prediction systems. The results confirm that sequence-based behaviour modeling can significantly improve the accuracy and usefulness of human activity analysis systems. Future work may include integrating deep learning models such as LSTM for improved sequence prediction and incorporating real-time sensor data for dynamic behaviour analysis.

## References

- [1] **Alghamdi, M., et al.**, “Enhancing human activity recognition with machine learning: insights from smartphone accelerometer and magnetometer data,” *PeerJ Computer Science*, vol. 11, p. e2185, 2025.
- [2] **Khan, R., et al.**, “A Hybrid Deep Learning Model for Human Activity Recognition Using Wearable Sensors,” *IEEE Transactions on Neural Networks and Learning Systems*, 2025. (Pre-print/Early Access).
- [3] **Patel, S., et al.**, “IoT powered RNN for improved human activity recognition with enhanced localization and classification,” *Scientific Reports*, vol. 15, no. 1, p. 10328, 2025.
- [4] **Zhang, Y., et al.**, “MHCNLS-HAR: Multiheaded CNN-LSTM-Based Human Activity Recognition Leveraging a Novel Wearable Edge Device,” *IEEE Sensors Journal*, 2024.
- [5] **Li, X., et al.**, “Self-supervised learning for human activity recognition using 700,000 person-days of wearable data,” *npj Digital Medicine*, vol. 7, no. 1, p. 91, 2024.
- [6] **Wang, H., et al.**, “Wearable sensors based on artificial intelligence models for human activity recognition,” *Frontiers in Artificial Intelligence*, vol. 7, p. 1424190, 2024.
- [7] **Chastin, S., et al.**, “Analytic Methods for Understanding the Temporal Patterning of 24-h Movement Behaviors,” *Journal for the Measurement of Physical Behaviour*, 2024.
- [8] **Ronao, C. A., et al.**, “Deep convolutional neural networks for human activity recognition with smartphone sensors,” *Applied Soft Computing*, vol. 42, pp. 281-293, 2016.
- [9] **Ordóñez, F. J., et al.**, “Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Human Activity Recognition,” *Applied Sciences*, vol. 6, no. 1, p. 25, 2016.
- [10] **Hammerla, N. Y., et al.**, “Deep, Convolutional, and Recurrent Models for Human Activity Recognition using Wearables,” *Proc. International Joint Conference on Artificial Intelligence (IJCAI)*, 2016.