International Research Journal of Multidisciplinary Scope (IRJMS), 2024; 5(1):113-123

Original Article | ISSN (O): 2582-631X DOI: 10.47857/irjms.2024.v05i01.0165

Predicting Human Activity and Behavior Patterns from Cell Phone Data Using Machine Learning

Arivoli A*, Soumya Brata Paul

School of Computer Science and Engineering, VIT University, Vellore, Tamil Nadu, India. *Corresponding Author's Email: arivoli.a@vit.ac.in

Abstract

Across the globe, Smartphones enjoy a high ownership rate. A person's day-to-day activities are incomplete without the use of smartphones. These sensor-rich devices are often repurposed to collect various forms of data from the user in the domains of 1) app usage, 2) overall phone activity, 3) location, 4) day and night-time activity, 5) communication, 6) social media behavior and so on. This information collected from sensor and log data taken from smartphones can be used to create predictive measures extending to an individual's Big Five personality traits and examine them. Human activity recognition (HAR) as a field has always tried to improve its performance in the aspects of feature extraction and improve the accuracy of recognition and prediction by performing various machine learning techniques on the data received from accelerometer and gyroscope sensors and creating convolutional neural networks to recognize the multiple activities such as walking, jogging, walking up or down the stairs, etc. Although HAR is a recognized research field, there have been significant challenges such as a) different movements for different individuals, b) limited training data, and c) complexity of performed activities. The idea of this project is to create a HAR model with the help of Machine Learning and Deep Learning, which will be more efficient in terms of performance and accuracy than the conventional models and try to overcome the challenges. That will be applied to the Smartphone data to predict a smartphone user's activity at one point of the day. The idea is to also get an analysis of the behavioral traits of the user in terms of online expenditure and predict the date of the next purchase. That will be done to send personalized information such as offers and announcements to the user catered according to their needs to enhance their personal experience.

Keywords: Human activity recognition, Convolutional neural network, Smartphones, Sensor, Machine learning, Deep learning.

Introduction

In this digital age, a person can only imagine performing daily activities and doing essential tasks with a smartphone. They have a portable operating system with computing capabilities, interconnectivity, and application programming interfaces (APIs) for running third-party tools and applications. Smartphones also have features such as image capture, storage, internet browsing, and embedded sensors like gyroscopes, accelerometers, and magnetometers(1), allowing applications to be developed based on a user's location, movement, and context. To create helpful smartphone applications, it's crucial to use context recognition and situational awareness of the device's user. One platform for this is Activity Recognition, which can be handled by the built-in sensors and is used in various fields such as business, healthcare, security, transportation, etc. Other types of sensors include wearable sensors that can detect movement and enable the transfer of information between devices using Bluetooth (2-3). Detecting and recording is possible with these portable sensors, which aid in recognizing subjects continuously. Activities can also be captured when monitored in their favored environment. Human Movement Recognition is a substantial and challenging field of research with many applications in intelligent workspaces, healthcare, and national security (4-5). PC vision-based methods have generally been used for human movement monitoring (6). A more efficient approach is to process data from inertial measurement unit sensors worn on the user's body (7-8), or built into a user's smartphone to track their movements. A model is developed to recognize various activities under real-world conditions using data collected by a single triaxial accelerometer built into a phone (9-10).

(Received 20th October 2023; Accepted 09th January 2024; Published 30th January 2024)

Considering the limited memory available on smart devices, the overall performance of classifiers with controlled training facts is considered. Gathering training records and using them directly for classification steps reduces the burden on users.
A method using machine-learning techniques was utilized to uncover patterns in various areas of behavior, including communication and social behavior, listening to music, using apps, mobility, general smartphone activity, and whole-day activity. The models used in this approach included an inner loop for preprocessing and adjusting parameters and an outer loop for accurately evaluating the models. Both loops utilized cross-validation techniques, with the inner circle using 5-fold cross-validation and the outer loop using 10X10-fold cross-validation (11).
A different approach suggested using a multi-layer perceptron with an activation function applied to the output and utilizing the DL4J Java library for building the neural network. However, limitations in memory constrained the network's size and depth, resulting in poor performance when the Android app was run on less powerful devices (12).
A different study introduces a new combination model utilizing CNN to enhance the identification precision. The CNN-7 model recognizes seven distinct activities, while the CNN-2 model is specifically created to differentiate between two activities that produce very similar signal patterns: ascending stairs and walking (13). Additionally, a study employs an LSTM-based method for extracting features to identify human activities using data from tri-axial accelerometers (14). A different study suggests a design incorporating a shallow CNN for removing local elements without supervision, along with features representing global characteristics of the time series. However, the limited number of activities in the study makes it difficult to compare its results to existing solutions (15). Another research (16) illustrates the connection between individual variations and actual actions and how this can aid in comprehending the behavioral foundation of character.

Methodology

UCI HAR Dataset: This dataset was used for the Human Activity Recognition part. The data in this dataset was captured from 30 people who carried out various actions while attaching a smartphone (Samsung Galaxy S II) to their waist. The UCI HARdata was collected using sensors in smartphones worn by individuals while performing various activities. It includes data collected from the accelerometers and gyroscope sensors of smartphones worn by volunteers while performing multiple activities. The most important purpose of this dataset is to develop and evaluate machine learning algorithms for activity recognition. The main features include time-domain and frequency-domain measurements from the accelerometer and gyroscope. This dual-source data enables the progress of more superior and accurate models by leveraging multiple types of information. Each data point is mapped with a specific activity, making it suitable for supervised machine-learning tasks. The smartphone contained sensors, such as an accelerometer and gyroscope, which captured the data during the experiment. Additionally, the experiment was also filmed for manual labeling of the data. In the study, the team in charge created new features for the time series by using a window 2.56 seconds in width and moving it along the string. Due to the 50% overlap of the windows, the resulting data points were evenly spaced at 1.28 seconds. The process was recorded on video for later manual labeling of the data. The data was categorized into six labels, namely "Walking," "Walking Downstairs," "WalkingUpstairs," "Sitting," "Standing," and "Lying," by Supervised learning problem as shown in Figure 1. Almost all participants' data captured was more for walking upstairs. So, participants took more time walking upstairs than downstairs, considering an equal number of steps.

Online Retail Dataset: This data set encompasses all transactions from a UK-based, non-store online retail company specializing in unique, all-purpose gifts. The Online Retail Dataset was also collected based on activities from the UCI machine learning repository but not the APP-based dataset. That dataset is based on a UK-based, non-store, online retail company where cell phones were popular in 2010. Each record in the dataset usually represents a single transaction or line item in an invoice. Multiple documents

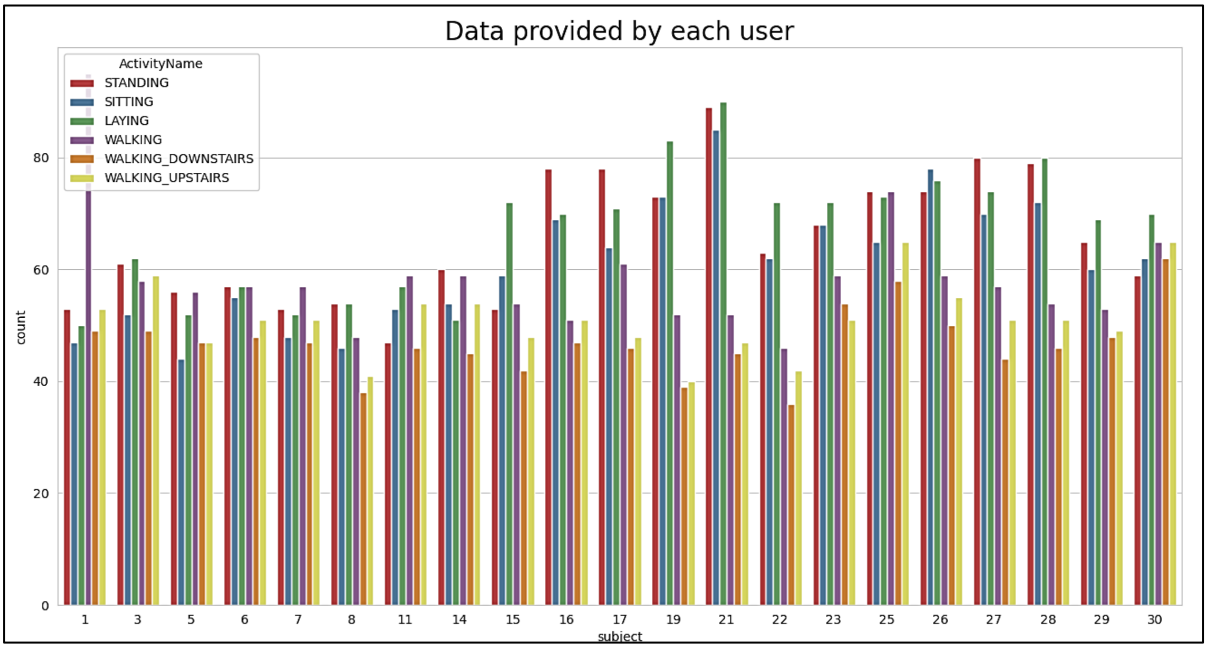

Figure 1: Data provided by each user perform the six types of activity

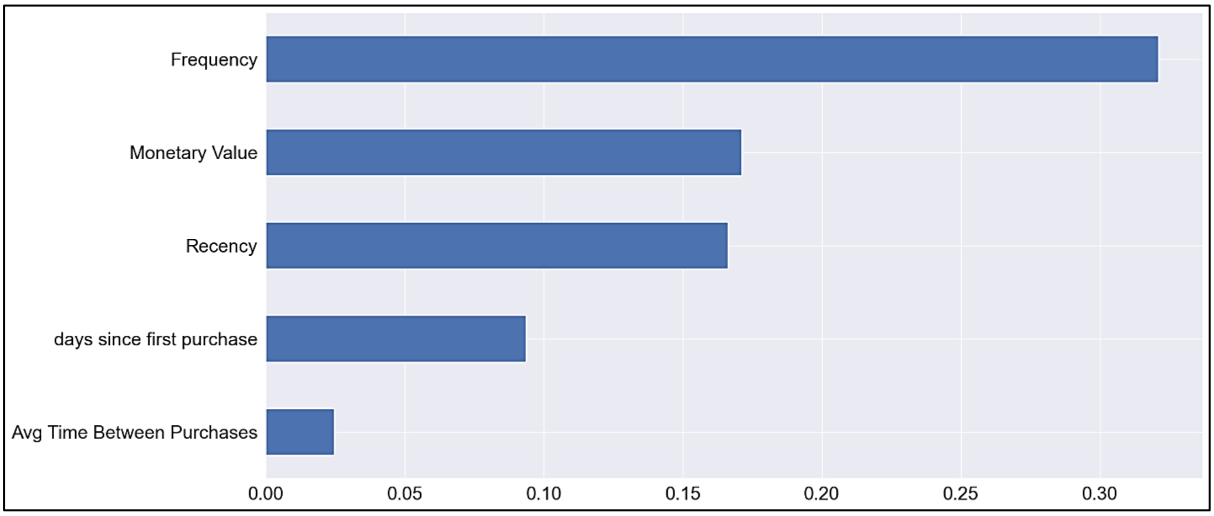

Figure 2: Various features generated from Preprocessed data of Online Retail dataset

may belong to the same transaction if various products are in one purchase. The transactions were recorded between December 1, 2010, and December 9, 2011. The company has a large number of wholesale clients. The preprocessed data is shown in Figure 2.

The HAR Dataset was obtained from the UCI dataset repository and then divided into two parts: the RAW dataset and the pre-engineered by domain or signal expert engineers. Therefore, we first utilize the pre-engineered dataset with traditional Machine Learning techniques to train and predict human activity. Afterward, we can employ the RAW dataset with Deep Learning models to learn and predict human movement.

The dataset has obtained '3-axial linear acceleration' (*tAcc-XYZ*) and '3-axial angular velocity' (*tGyro-XYZ*) from a smartphone's accelerometer and gyroscope, respectively, using various configurations. In these metrics, the prefix 't' represents the aspect of time, and the suffix 'XYZ' indicates signals in the X, Y, and Z directions that measure three different axes.

The expert took raw sensor data and transformed it into valuable features by filtering out noise and dividing it into windows of 2.56 seconds with 50% overlap. They calculated variables in both the time and frequency domain from each window. The acceleration signal was divided into the Body acceleration signal and the Gravity acceleration signal (tBodyAcc-XYZ and tGravityAcc-XYZ), using a low pass filter with a corner frequency of 0.3Hz. Afterward, they calculated the body's linear acceleration and angular velocity in real-time to obtain jerk signals (tBodyGyroJerk-XYZ and tBodyAccJerk-XYZ). The magnitude of these 3-dimensional signals was used as features with names such as tBodyAccMag_, tGravityAccMag, tBodyAccJerkMag, _tBodyGyroMag, and tBodyGyroJerkMag.They also separated the acceleration signal into body and gravity acceleration and derived linear acceleration and angular velocity to get jerk signals. They then calculated the magnitude of these signals, labeled them with the prefix 'f', and named them accordingly. These signals are categorized as fBodyGyroMag,fBodyAcc-XYZ, etc.

The dataset assigns numerical values from 1 to 6 as identifiers for the Y_labels, with one representing WALKING, 2 representing WALKING_UPSTAIRS, 3 representing WALKING_DOWNSTAIRS, 4 representing SITTING, 5 representing STANDING, and 6 representing LAYING.

The readings from 30% of the volunteers (9) were taken as test data, and the remaining 70% of the subjects' recordings (21) were taken for training data. If the participants walked the same number of stairs going up and down, and the smartphones recorded data at the same rate, there should be an equal number of data points for both walking up and down the stairs. If there is no inaccurate data, the participants appeared to walk down the stairs about 10% faster than they walked up to them.

Motion info will be significant in the case of Moving activities and not so much in the case of Stationary activities, as shown in Figure 3.

The data must be separated into different classes to differentiate the activity being performed. The magnitude of acceleration is one of the factors used to separate it. Suppose the average acceleration (tAccMean) is less than -0.8. In that case, the activities are likely to be standing, sitting, or laying, and if it is >-0.6, they are likely to be walking downstairs or upstairs. Also, if the angle between the X-axis and Gravity_mean is > 0, the activity is Lying. In the sitting and Standing position, the magnitude of acceleration is still close to 0. Suppose you are standing and sitting without moving (tAccMean is equal to 0). In this sitting behavior, the person is characteristically stationary and sitting with their torso and legs forming a right angle or a similar seated position. In this standing behavior, the person is static and standing on their feet. Like other activities, the smartphone's sensors record acceleration and angular velocity data while the person is standing and sitting. The dataset's "SITTING" and "STANDING" behaviors are necessary components for activity recognition models, and accurately categorizing these behaviors can provide insights into an individual's daily activities and behavior patterns. As shown from the t-SNE cluster in Figure 4, all activities are separated nicely.

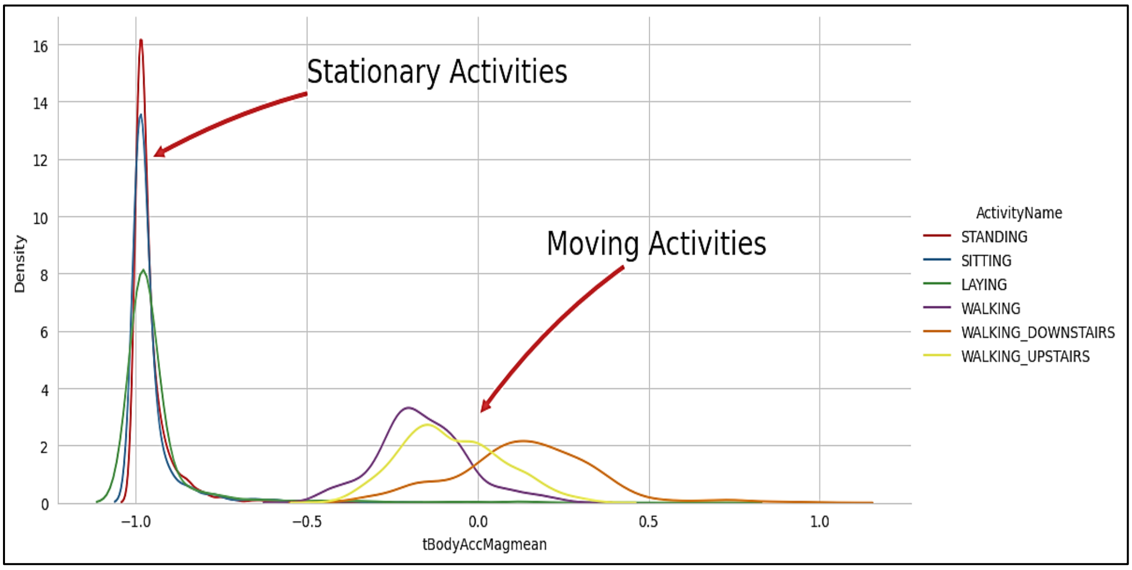

Figure 3: Differentiating Stationary and Moving activities

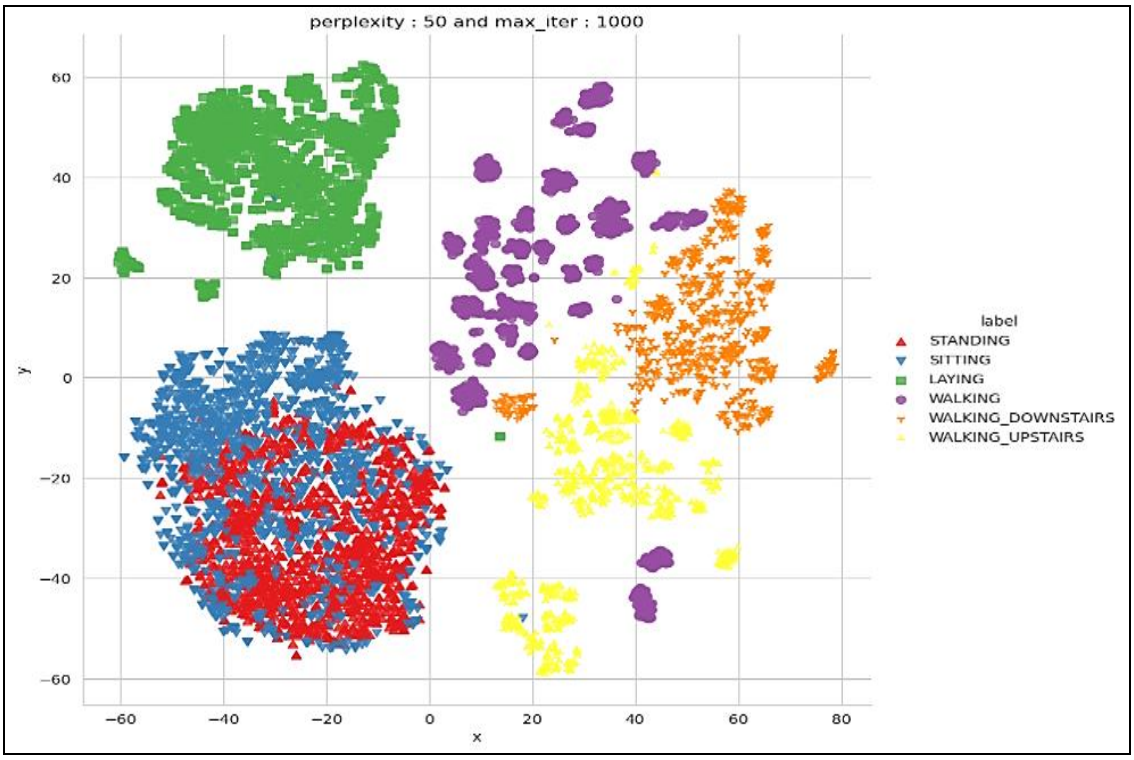

Figure 4: t-SNE clusters of the data generating various activities

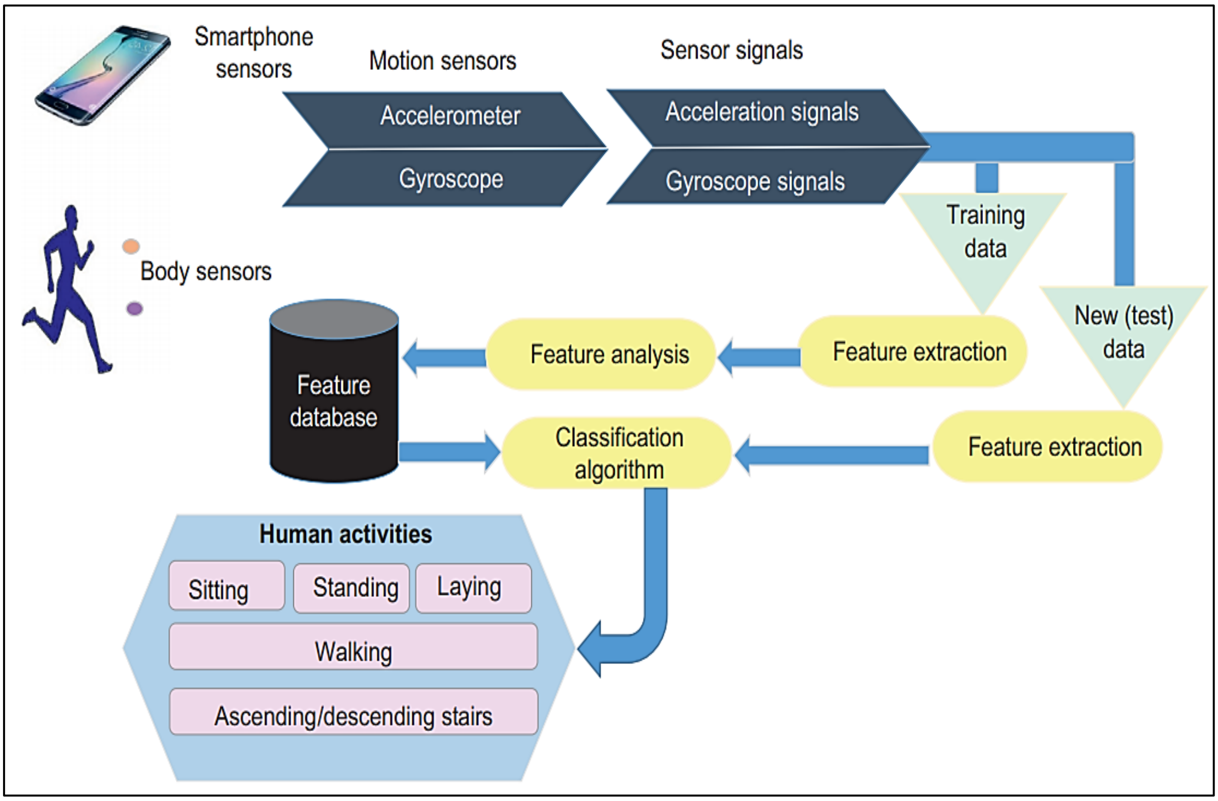

Figure 5: Predicting human activity using the traditional ML models

Experimental Procedure

a. Creating ML models: Initially, the dataset had 564 features, and we will remove the three that tell the activity is being performed directly. We will apply the traditional machine learning models to the remaining 561 features and predict the movement, as shown in Figure 5. The Machine Learning models used are: 1. Logistic Regression, 2. Kernel SVM, 3. Random Forest Classifier, 4. Linear SVC, and 5. Decision Tree.

b. Creating LSTM Model: We developed a deep learning model using LSTM (long short-term memory) on raw time series data in this part. The goal was to classify human activities such as standing, climbing stairs, and going downstairs.LSTM, a Recurrent Neural Network (RNN), was the most effective method for recognizing these human activities. Initially, a 1-Layer LSTM Model was created, showing excellent results without hyperparameter tuning. After that, the 2-layer of the LSTM Model was defined with more hyperparameter tuning, and it was fitted on the raw time series data.

Predict Purchase Date: This section aims to employ machine learning to forecast the likelihood of a customer purchasing within the next week. Feature engineering was performed to clean the data to get accurate predictions and generate features based on recency, frequency, and monetary value, setting seven days as the target window for prediction. It is also assumed that determining an average number of days between purchases helps predict future purchases, as shown in Figure 7. Subsequently, we utilized various machine learning techniques, including logistic regression, XGBoost, Random Forest, AdaBoost, and Gaussian Naive Bayes, to create classification models. According to the AUC score, the random forest model is the most precise and was chosen to generate propensity scores to identify customers with the highest probability of purchasing within the next seven days, as shown in Table 1. Random forest features, such as customer democracies, purchase history, and website interaction, are good when predicting the next purchase date. Train a Random Forest regression model on the training data. Predicting numerical values (product purchases) is essential in the context of regression rather than classification. The model acquires patterns and relationships between the features and product purchases in the training data. Random Forest is a multipurpose algorithm for time series predicting and can capture complex connections between elements and product purchases. It's important to fine-tune the model, consider feature importance, and incorporate domain knowledge to make informed decisions based on the predictions. Combining the forecast with business strategies can help optimize product availability and marketing efforts. In this way, companies can send notifications to such users at specific points of time during the day to increase the likelihood of customer's purchase activity.

Table 1: Sample of Customers with the highest purchase probability within the next seven days

CustomerID	Recency	Frequency	Monetary Value	days since first purchase	Avg Time Between Purchases	Purchase Probability
15498.0	10	26	12451.46	359	13.423077	0.942407
16525.0	9	25	12705.73	359	14.000000	0.941355
13113.0	7	23	11906.76	359	15.304348	0.940367
13319.0	10	25	10536.45	358	13.920000	0.938644
17428.0	15	27	16841.25	361	12.814815	0.936842

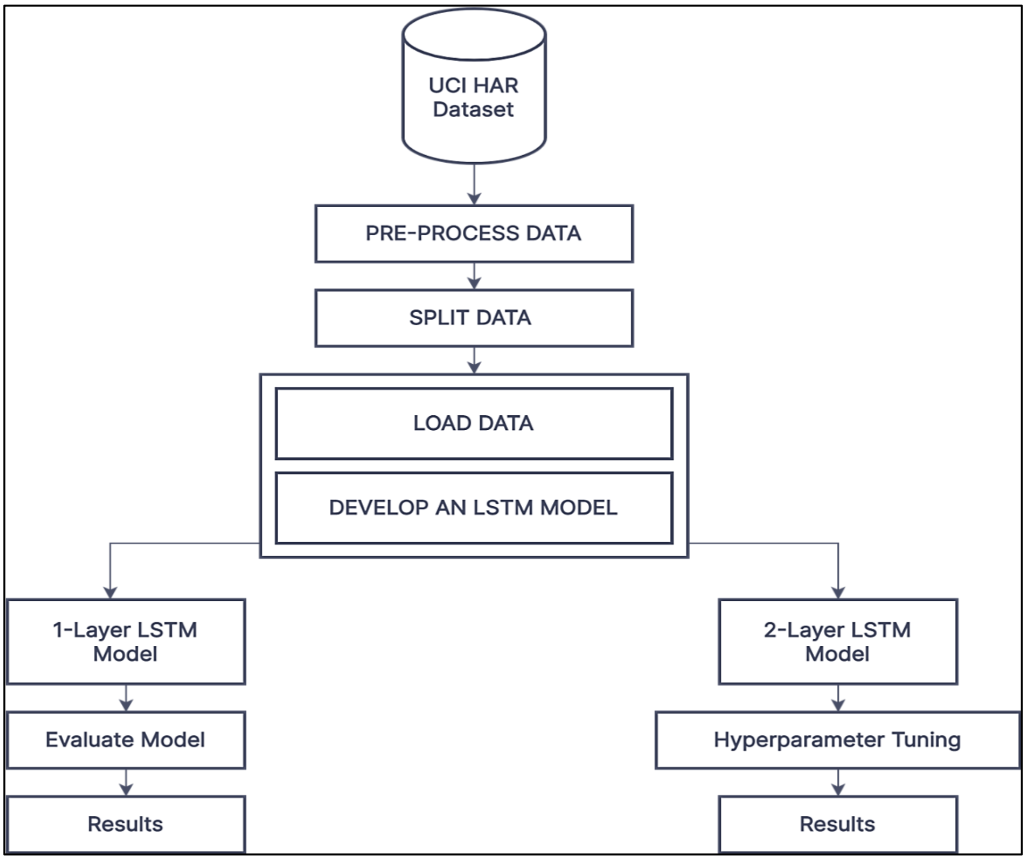

Figure 6: Apply LSTM Model on Raw time series data

Figure 7: Average Number of Days Between Purchases for all Customers

Results and Discussions

We applied traditional Machine Learning algorithms on the dataset containing expert-generated features and evaluated the accuracy as shown in Table 2. As we can observe, Linear SVC has the highest accuracy of 96.78%, closely followed by RBF SVM. So, any one of these can be used.

The confusion matrix was also generated, as shown in Figure 8(a), and we can observe that the ML model can predict the human activities accurately except sometimes confusing “Sitting” with“Standing,” as seen in the third column.

Regarding the Deep Learning model, we can achieve impressive results using only raw data with LSTM without needing additional feature engineering, as shown in Table 3.

The 1-Layer LSTM model alone achieved 90.09 % with a loss value 0.477. When we increase the layers and do hyperparameter tuning, the accuracy increases, and the Cross-entropy value decreases. The confusion matrix can be observed in Figure 8 (b). The model can predict most human activities correctly with raw time series data, which is impressive. Hyperparameter tuning is the iterative process of finding the optimal solution to the problem. Choosing the correct hyperparameters achieves better accuracy, lower error rates, and improved model generalization. The tuning hyperparameters model familiarizes the characteristics of the sensor data, such as noise levels and sampling rates.HAR models are likely overfitting, mainly when the dataset is small or imbalanced. Hyperparameter tuning can mitigate overfitting issues and improve the model's robustness. LSTM-based models are powerful tools for modeling human activity behavior using sensor data. The model has been compared with LSTM with 1_layer, LSTM with 2_layer, and LSTM with 2_Layer with neurons 32, 48, and 64, along with the hyperparameter tuning.

Table 2: Performances of applied traditional ML Models

Model	**Accuracy**	**Error**
Logistic Regression	95.79%	4.208%
Linear SVC	96.78%	3.224%
rbf SVM classifier	96.27%	3.733%
DecisionTree	87.92%	12.08%
Random Forest	92.26%	7.737%

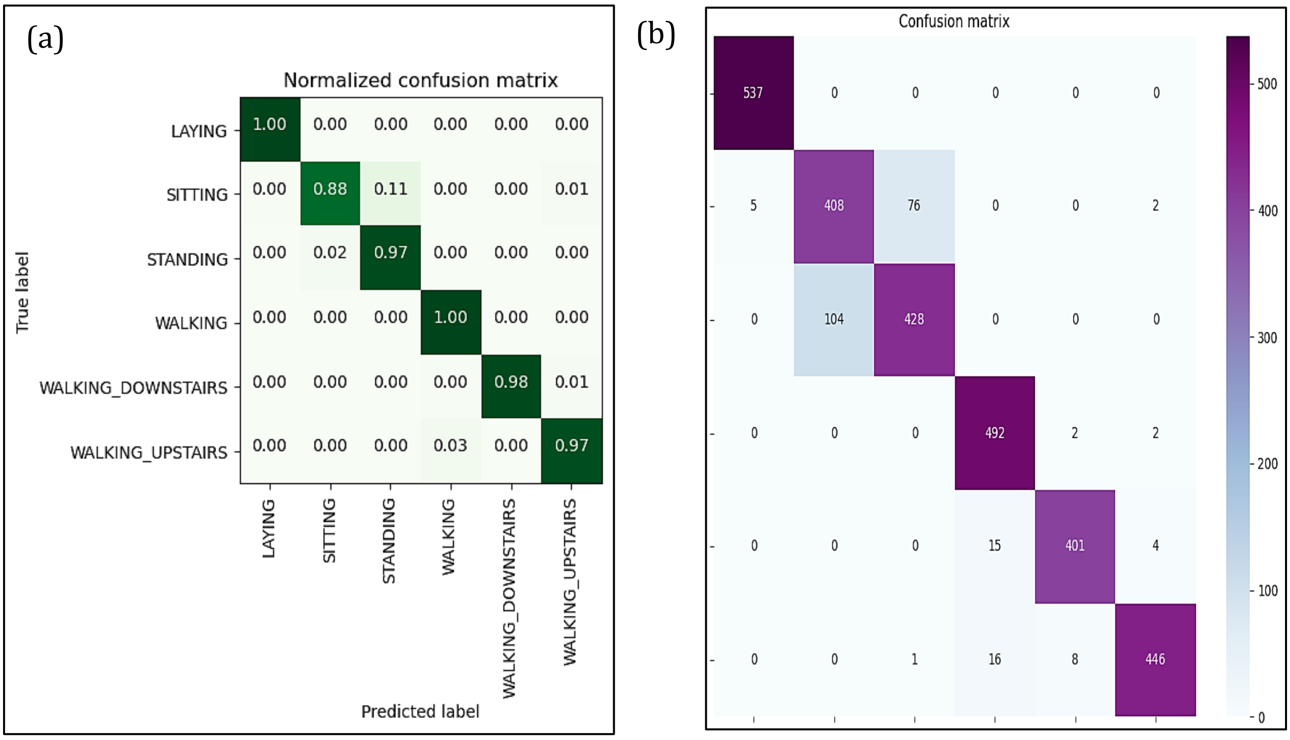

Figure 8: Confusion matrix for (a) Linear SVC and (b) 2-Layer LSTM Model

Table 3: Performances of created LSTM Models

Model Name	Cross entropy	Accuracy
LSTM With 1_Layer (neurons:32)	0.477	90.09%
LSTM With 2_Layer (neurons:64, neurons:48)	0.343	91.17%
LSTM With 2_Layer (neurons: 128, neurons: 64)	0.238	92.37%

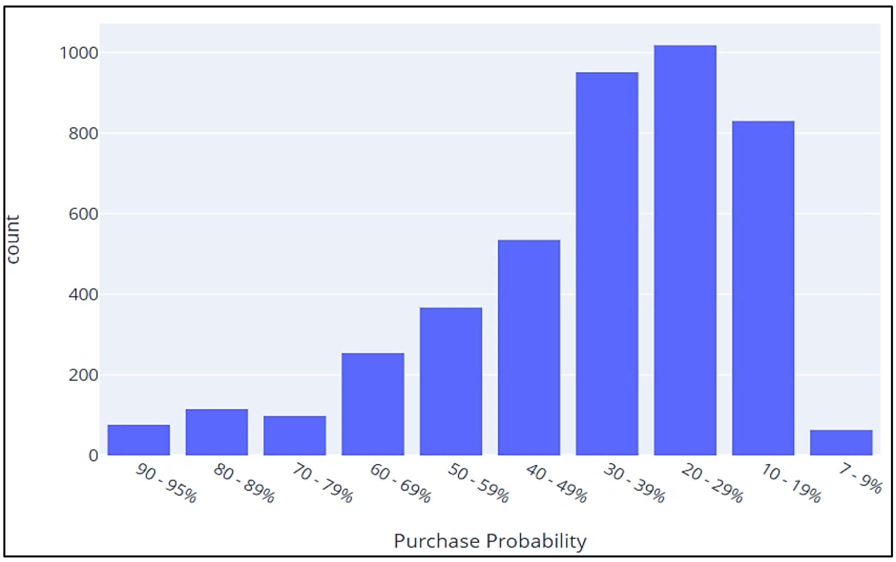

Figure 9: Graph showing Probability of customers making a purchase

After predicting the activity with the Deep Learning model, we used a Random Forest classifier to make the model for predicting the next date of purchase as it had the best AUC score among the other models. Finally, the purchase probability of customers can be observed in Figure 9.

If financial and monetary constraints are not an obstacle, we can provide marketing incentives to customers in the 70-79%, 60-69%, and 50-59% segments to boost their likelihood of purchasing. Customers in the 90-95% and 80-89% segments already have a high chance of purchasing, so incentives may not be necessary. However, incentives should be offered if they fail to purchase within the expected timeframe. When resources are limited, incentives should only be given to customers in the 70-95% segments who do not purchase within seven days.

Cross Validation

The best model for each algorithm is formulated using the derived best parameters. Then, the models are evaluated on the test data. Test data is divided into five random folds (stratified), each containing the whole dataset. We got the average accuracy for the five different models using the data. We also tested with a five-by-two-fold cross-validation t-paired test to confirm the winner algorithm. The whole dataset (training and test data) is divided into two random folds and run five times in this test. Then, the results are collected, and the t statistic is calculated.

Conclusion

When applying the ML Classical Model, we can select Linear SVC, RBF SVM classifier, or Logistic Regression as our preferred model. However, when using the LSTM Model, we use raw data rather than engineered data created by experts. Layers are added in LSTM in Layer 2 to increase the accuracy, but the accuracy of Layer 1 was only 90.09%. Despite this, LSTM still performs exceptionally well and achieves a high accuracy of 91% with 2-layer LSTM and hyperparameter tuning. Additionally, as we increase the number of LSTM layers and fine-tune the hyperparameters, the cross-entropy value decreases and accuracy increases. This LSTM model with more hyperparameter tuning provides more accuracy than the existing ML algorithms in the comparative work. The Random Forest model was used to predict the probability of purchase for the recurring customers, and we discussed how the companies can strategize their marketing to increase sales.

Abbreviation

Nil

Acknowledgement

I would like to thank Dr. Arivoli A of VIT University for her guidance and suggestions which helped to make this project better. I would also like to thank VIT for providing me the knowledge, opportunity and resources to implement this project. I would always be grateful for the opportunity provided to me.

Authors contribution

All authors contributed to the study conception and design.

Conflicts of interest

The authors declare that they have no competing interests.

Ethical approval

The work is not being considered for publication or released as a preprint elsewhere.

Funding

No funding received by any government or private concern.

References

1. Lane ND, Miluzzo E, Lu H, Peebles D, Choudhury T, and Campbell AT. A survey of mobile phone sensing. IEEE Communications Magazine. 2010; 48(9): 124-156.
2. Saponas S, Lester J, Froehlich JE, Fogarty J, & Landay JA. iLearn on the iPhone: Real-Time Human Activity Classification on Commodity Mobile Phones.2008.
3. Brezmes T, Gorricho JL, and Cotrina J. Activity recognition from accelerometer data on a mobile phone. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2009;5518.
4. Khan AM, Lee YK, Lee SY, and Kim TS. A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer. IEEE Transactions on Information

Technology in Biomedicine. 2010; 14(5):1765-1780.
5. Westerterp KR. Assessment of physical activity: A critical appraisal. European Journal of Applied Physiology. 2009; 105(6).
6. Poppe R. A survey on vision-based human action recognition. Image Vis Comput. 2009;28(6).
7. Casale P, Pujol O, and Radeva P. Human activity recognition from accelerometer data using a wearable device. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2011;6669.
8. Krishnan NC, Colbry D, Juillard C, and Panchanathan S. Real time human activity recognition using tri-axial accelerometers. Sensors Signals and Information Processing Workshop. 2008.
9. Kwapisz JR, Weiss GM, and Moore SA. Cell phone-based biometric identification. IEEE 4th International Conference on Biometrics: Theory, Applications and Systems. 2010.
10. Kwapisz JR, Weiss GM, and Moore SA. Activity recognition using cell phone accelerometers. ACM SIGKDD Explorations Newsletter. 2011; 12(2).
11. Stachl C. Predicting personality from patterns of behavior collected with smartphones. Proc Natl Acad Sci. 2020; 117(30).
12. Voicu RA, Dobre C, Bajenaru L, and Ciobanu RI. Human physical activity recognition using smartphone sensors. Sensors (Switzerland). 2019; 19(3).
13. Zhu R. Efficient Human Activity Recognition Solving the Confusing Activities Via Deep Ensemble Learning. IEEE Access. 2019; 7.
14. Chen Y, Zhong K, Zhang J, Sun Q, and Zhao X. LSTM Networks for Mobile Human Activity Recognition. 2016.
15. Ignatov A. Real-time human activity recognition from accelerometer data using Convolutional Neural Networks. Applied Soft Computing Journal. 2018;62.
16. Stachl C. Personality Traits Predict Smartphone Usage. Eur J Pers. 2017; 31(6).