WalkSmart – A New Machine Learning-Based Real-Time Gait Analysis Technique using Smartphones to Mitigate Ambulatory Fall Incidents

Shruti Badrish
Redmond High School
Redmond, WA, USA
shrutibadrish@gmail.com

***Abstract*—Ambulatory fall incidents – which denote unanticipated falls during normal walking – are well known to be a leading cause of injury and injury-related deaths, especially in the elderly and in those with neurological disorders. Every year, one in four older adults experiences a fall, and every 11 seconds, someone is treated in an emergency department for an ambulatory fall-related injury. In this work, I aim to mitigate ambulatory falls by using smartphones to detect and alert people when their risk of falling increases. I propose WalkSmart – a new machine learning-based technique for ambulatory fall prevention based on analyzing *gait variance* using sensor data collected using smartphones. WalkSmart runs as a smartphone app and interacts with my cloud-based machine learning model to detect and alert people of impending falls, thereby potentially saving lives.**

WalkSmart uses a new machine learning model for classifying gait patterns as steady or unsteady, using a long short-term memory (LSTM) based recurrent neural network as its basis. WalkSmart was trained using hours of real-world walking data collected using smartphones. Experiments show that it can effectively infer unsteady gait with a high accuracy of 96.6%, using a small window of accelerometer and gyroscope sensor data captured by a smartphone.

***Index Terms*—fall prevention, machine learning, gait, smartphone, sensor, data processing**

I. Introduction

The Centers for Disease Control and Prevention (CDC) reports that ambulatory falls are the leading cause of injury and injury-related deaths among adults aged 65 years and older in the United States [2]. Every year, one in four older adults experiences such a fall, and every 11 seconds, someone is treated in an emergency department for a fall-related injury. Ambulatory falls can lead to serious injuries such as hip fractures, head trauma, and lacerations, which can result in hospitalization, loss of independence, and even death. Each year at least 300,000 older people are hospitalized for hip fractures.

A large body of research in this space focuses on *fall detection*, which focuses on automatically detecting that someone has fallen, so that post-fall assistive measures such as calling emergency personnel can be automatically performed. However, given the seriousness of this problem and its severe consequences on the subsequent quality of life for a fragile demographic, it is of vital importance that we design practical yet proactive techniques to prevent the occurrence of ambulatory falls in the first place.

As the first step towards a solution, we make the observation that smartphones have become a ubiquitous part of adult life today. According to Pew Research Center, the share of Americans that own a smartphone is now 85%, up from just 35% in 2011 [1]. Further, a vast majority of smartphones carry sensors such as accelerometers and gyroscopes. Accelerometers detect the orientation of the smartphone, whereas gyroscopes track rotation or twisting movements of the device. All sensor readings can be captured and processed in real time by phone apps and downloaded for processing on computers.

Second, we observe the existence of a correlation between the *gait* of an individual and their risk of suffering an ambulatory fall. Gait is defined as the walking pattern of an individual. It is long known that irregularities in one's gait can be effectively used to predict ambulatory fall risk [9], [15]. A number of factors can cause such irregularities, such as the development of neurological disorders, injuries, impairment, a sudden onset of dizziness, or substance abuse. These irregularities are in turn highly likely to lead to such falls.

In this paper, I propose a new method for ambulatory fall prevention based on the detection of gait irregularities through smartphone accelerometer and gyroscope sensor data – WalkSmart, a smartphone app that can effectively predict such irregularities and alert the user. The overall approach is as follows: we study a corpus of normal and abnormal gait patterns and use machine learning to create a time-sequence-based deep learning model called LSTM (for long short-term memory) [6] that can detect abnormal gait. We can then use this learned model to perform inference over smartphone sensor readings collected and processed in real time using the smartphone. Sensor readings are then used by an app on the phone to detect gait abnormalities and provide timely (within few seconds in the ideal case) warnings to the user. WalkSmart lowers the likelihood of a subsequent fall, and thus is a vital tool for preventing falls before they occur.

The rest of this paper is organized as follows. I describe my data preparation in Section II, covering data collection and pre-processing. Next, Section III describes my new neural network, including training and the integration of inference using my smartphone app. I report results in Section IV showing the accuracy of my solution, its application in an app, and its ability to perform inference on unseen test data

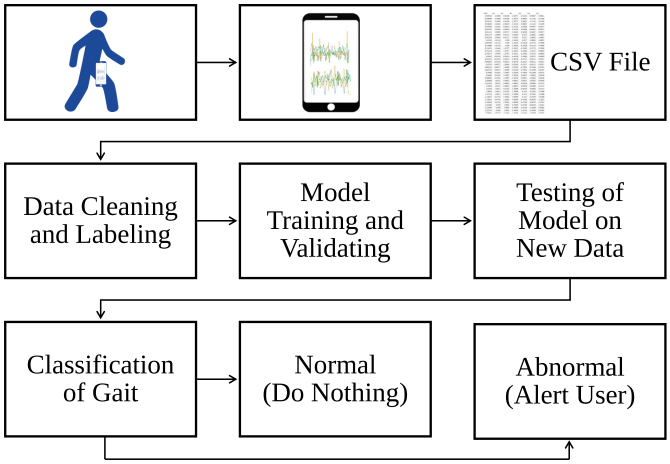

Fig. 1. Flowchart of the Gait Classification process

in real time. I survey related work in Section V and conclude in Section VI.

II. Data Preparation

The objective of this study is to categorize the walking patterns of individuals who walk with a steady walking pattern that is characteristic of a normal gait and those who have an unsteady walking pattern which is indicative of an abnormal gait putting them at a higher risk for ambulatory falls. This classification was done based on sensor data collected from the smartphone they carried. Specifically, this study collected smartphone accelerometer and gyroscope readings from participants while they walked with varying types of gait, including normal and simulated abnormal gaits. The data was used to train and test a new machine learning model called WalkSmart, that I describe in the next section. My model is able to distinguish abnormal gait patterns from normal gait patterns. Fig. 1 provides a flow chart of the process.

A. Data

The mobile application 'Physics Toolbox Sensor Suite' was used to collect smartphone sensor (accelerometer and gyroscope) data. The data was exported into a comma-separated value (CSV) file format. Accelerometers and gyroscopes are two of many sensors present in most smartphones, and measure the dynamics of the phone along the x, y, and z axes. Specifically, the linear accelerometer measures linear acceleration in m/s^2, and the gyroscope measures angular rotational velocity in rad/s. Together, the sensors reflect the smartphone user's motion, and can be used to determine their gait pattern. The data was collected at a 50 Hz sample rate in order to optimize the amount of data that needs to be processed by the model and thus reduce computation time.

B. Data Collection

I performed a study where raw accelerometer and gyroscope data was collected through smartphone sensors while participants walked for fixed time periods. Participants in the study were of a range of ages, from adolescents to adults. All participants were healthy and were capable of walking for the entire period. The walking interval was set to 3 minutes and participants walked multiple times resulting in over fifty hours of training data. All data collected and activities performed in the study were done with prior consent from participants. During data collection, each participant carried a smartphone in their right front pocket. The participants were instructed to walk in 3-minute intervals, during which they performed several variations of walking. Specifically, they walked three times in their normal walking pattern for 3 minutes each time at different speeds. Additionally, they walked for 3 minutes with their left shoe on and right shoe off, with the left shoe raised by 4cm and repeated the same this time with the right shoe on and raised by 4cm and left shoe off. Finally, they walked for another 3 minutes with the left knee unbent during their walk. These walking patterns were chosen to simulate the antalgic gaits [3] and left leg pegged patterns. This process was repeated multiple times with each participant. The walking patterns are summarized in Table I. The data collected included x, y, and z angular velocities and linear accelerations, which were measured 50 times per second and exported for analysis. This data is saved in CSV files. The file names include the word "normal" or "abnormal" based on the walking it corresponds to, to help with data preprocessing.

C. Data Preprocessing

Each CSV file corresponds to a single participant and contains data collected during various walking conditions. For each file, the data is preprocessed by removing null values, converting the accelerometer and gyroscope values to floating point values, and assigning a 'normal' or 'abnormal' label. This cleaned data is appended into a pandas DataFrame. The resulting DataFrame contains all the cleaned and labeled data from all the participants, which I use for training my model.

D. Data Visualization

Data visualization is carried out by plotting the accelerometer and gyroscope data against time. A clear difference is visible in the normal and abnormal walking gaits. Figures 2 and 3 show samples of normal and abnormal gait data respectively.

III. WalkSmart Model and App

WalkSmart consists of a neural network model that is based on a Long Short-Term Memory (LSTM) network. The neural network was trained on my collected data in order to classify the gait into one of the two categories. It was implemented using the Python Keras Sequential API. This machine learning model forms the basis of the WalkSmart smartphone app, which will be capable of inferring unsteady gait based on short intervals of smartphone accelerometer and gyroscope data.

A. The Model

An LSTM is a type of recurrent neural network (RNN) [11] used for sequence prediction. LSTMs are designed to handle

TABLE I
WALKING GAITS

Gait style	Description
Normal brisk	Participants walk in their normal walking pattern at their usual pace.
Normal fast	Participants walk in their normal walking pattern at a fast pace.
Normal slow	Participants walk in their normal walking pattern at a slow pace.
Left antalgic	Participants walk with their left shoe on for a height difference of 4cm between their legs.
Right antalgic	Participants walk with their right shoe on for a height difference of 4cm between their legs.
Left pegged (straight) leg	Participants walk without bending their left knee.

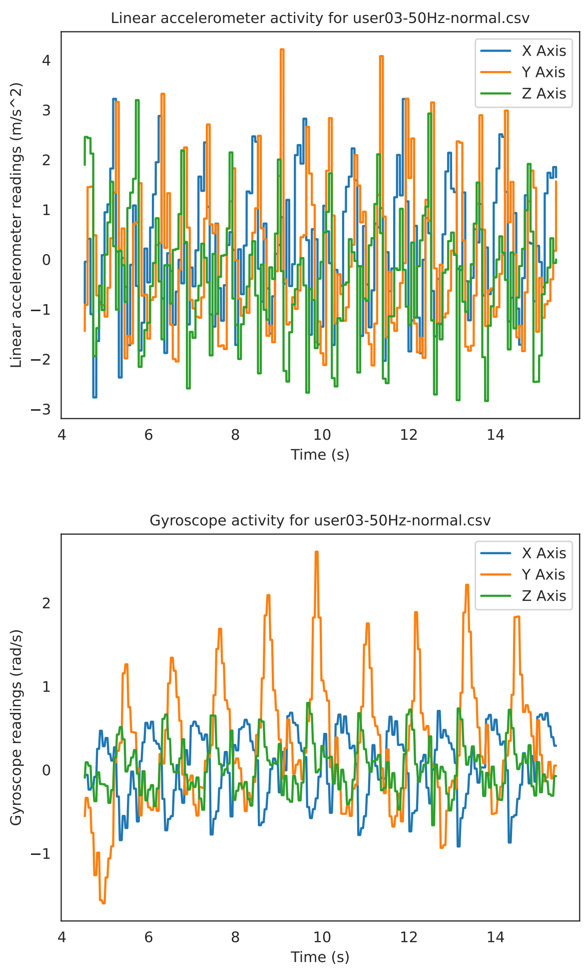

Fig. 2. Example graphs of linear accelerometer and gyroscope readings for normal gait.

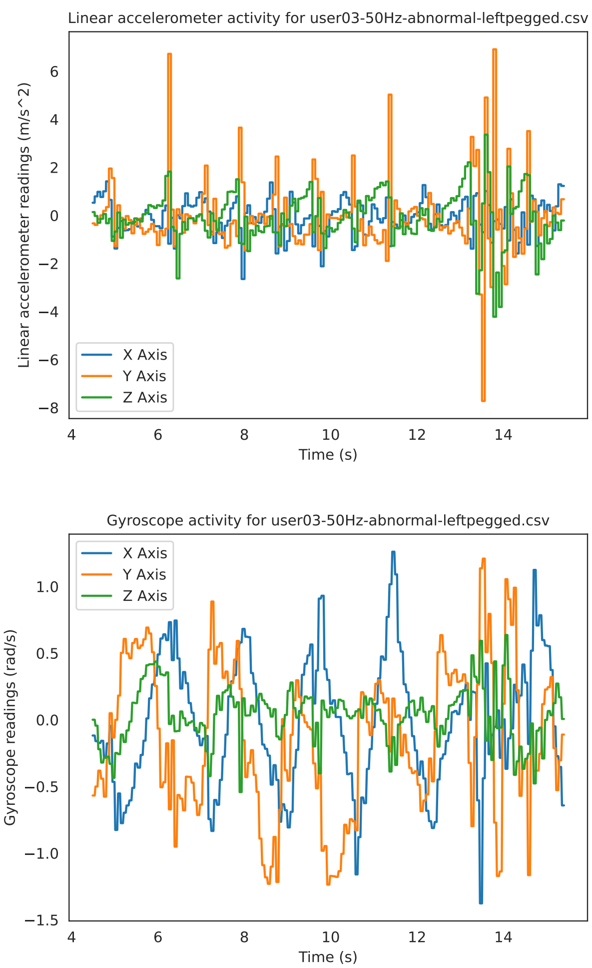

Fig. 3. Example graphs of linear accelerometer and gyroscope readings for abnormal gait.

sequential data, such as (as used in this project) time-series data. They do so by processing the data one element at a time, while maintaining a memory of previous elements. The model I implemented has four layers.

The first layer in the WalkSmart model is an LSTM layer with 128 units. This layer is a type of RNN that allows the model to learn and remember long-term dependencies in the input data, making it well-suited for sequential data such as time series.

The second layer is a dropout layer [13] with a rate of 0.5. This layer randomly drops out half of the units in the previous layer during each training iteration, which helps to prevent overfitting and improve generalization performance.

The third layer is a dense layer with 64 units and a rectified linear unit (ReLU) activation. The ReLU function returns 0 for negative values and is linear for positive values. It can be written as $f(x) = \max(0, x)$.

The fourth and final layer is a softmax layer with a number of units equal to the number of classes in the output (i.e., 2). This layer computes the probability of each class given the

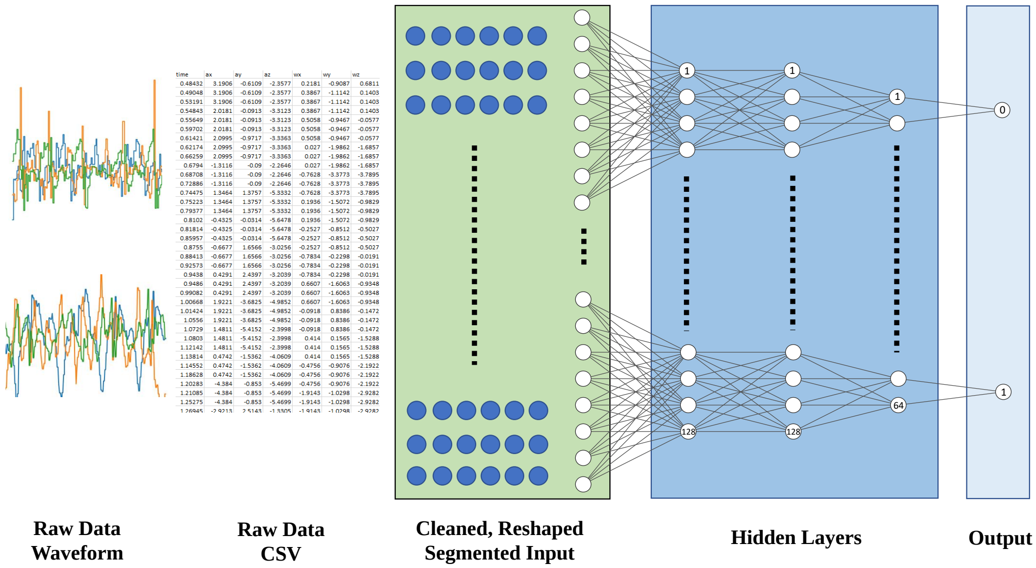

Fig. 4. Visualization of the LSTM RNN model

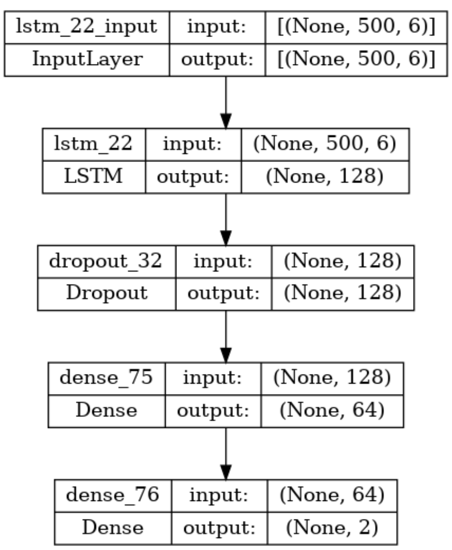

Fig. 5. Visualization of the LSTM RNN model

input data and ensures that the probabilities sum to one.

Ultimately, this model classifies sequential data into one of two classes using an RNN with an LSTM layer, dropout for regularization, and dense and softmax layers for classification.

Figures 4 and 5 show a visualization of my model.

Data is classified and trained over 10 second intervals with a 1.0 second sliding window. I chose 10 second intervals of data to make sure that a slight change in gait in otherwise normal walking does not get flagged as abnormal. The number of training epochs was set to 200. The number of epochs in LSTM training is the number of times the algorithm goes through the entire training dataset during the training process. I found that going beyond 200 epochs provided diminishing accuracy returns as compared to the time taken to train the model. Additionally, overfitting can occur if the number of epochs is too high, which can lead to poor generalizability for new data.

B. *Smartphone Integration*

The WalkSmart neural network model is integrated into a new smartphone app that I built to help mitigate ambulatory falls. Users of the app are expected to keep the phone (with the app running) on their self. The WalkSmart app is implemented with Android Studio, and consists of several components, described next.

- *Data Collection*: The app allows users to select a sensor data source, for example, a trace of collected smartphone accelerator and gyroscope readings. The data collector uses data collected by a separate app, the Physics Sensor Toolbox Suite.
- *Cloud Inference*: I deploy the WalkSmart machine learning model on a cloud server, and expose it as a web service. The server is implemented with Python's Flask web application framework. It accepts a small window of sensor data through a HTTP POST request, through a REST API. It then uses the Keras and Tensorflow library to perform inference on the data, and returns the classification and fall likelihood back to the smartphone app in a JSON formatted result.
- *Result Display*: The WalkSmart app displays the result of inference, and alerts the user with an audible sound and vibration when the likelihood of a fall exceeds a set threshold. Figure 6 shows the output as displayed in the WalkSmart app.

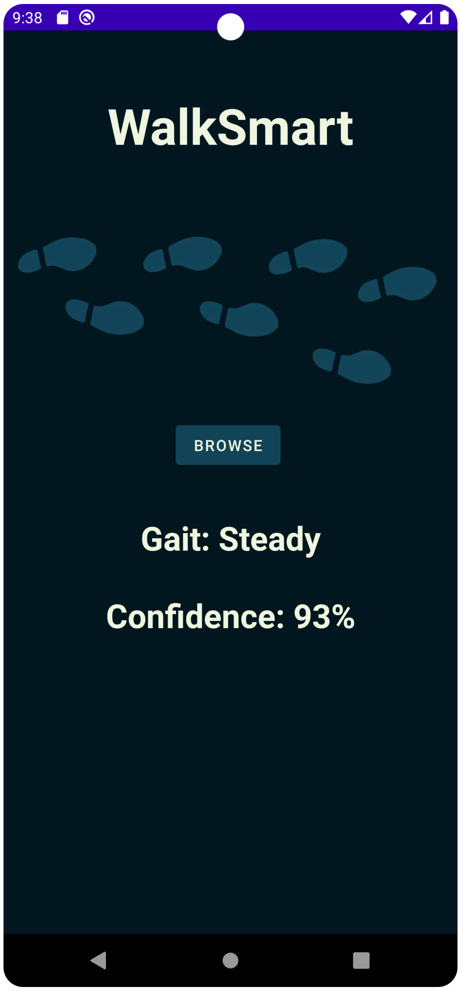

Fig. 6. Result display of the WalkSmart App

IV. RESULTS

This system for monitoring one's gait was used to detect irregularities and, thus, the likelihood of a future fall. The model was trained over the pre-determined training data, for 200 epochs (Figure 7).

The accuracy of the trained model (i.e., the percentage of data that is classified correctly) is 96.6%. The loss of the trained model is computed using a cross-entropy loss function, defined as $L(y, \hat{y}) = -\left[y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})\right]$, where y is the true label (either 0 or 1) and $\hat{y}$ is the predicted probability of the positive class. In my case, the loss is calculated to be 0.127.

Figure 8 depicts the confusion matrix for the model, in which positive is defined as abnormal gait, which provides further information on the model's performance. Each square in a confusion matrix depicts one of four possible outcomes: true positive (TP, top left), false positive (FP, bottom left), false negative (FN, top right), or true negative (TN, bottom right). These values can be used to calculate the model's precision, recall, and F-score.

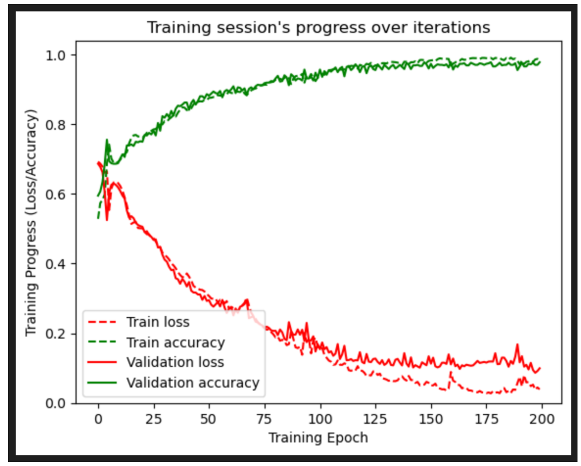

Fig. 7. Model training accuracy and loss model over epochs.

The precision of a model is defined as how many predicted positive cases were actually positive, or

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

The precision value for this model is 0.963. The recall of a model is defined as how many positive cases were predicted as positive, or

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

The recall value for this model is 0.979. The F-score of a model is defined as the harmonic mean of recall and precision, or

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

The F-score value for this model is 0.971.

From the confusion matrix, it can be seen that the model behind WalkSmart predicts the type of gait with great accuracy.

V. RELATED WORK

The paper [5] provides a good survey of prior work.

Sensors such as inertial measurement units [4] or hall-effect sensors [3] have been previously used for gait analysis. However special sensors can be cumbersome to carry around for the sole purpose of fall detection. This work utilizes smartphone sensors to solve a similar problem, thus increasing the accessibility of the gait analysis method.

In [17], the authors propose smartphone-based gait recognition as a less obtrusive way to capture biometric information, compared to fingerprints, iris, face, and voice recognition. In [16], a person's drunken gait is determined using a deep neural network algorithm, achieving a 79% test accuracy. In [10], accelerometer data is used to identify gait changes in Parkinson's patients by using a random forest model which achieves a 92.5% test accuracy.

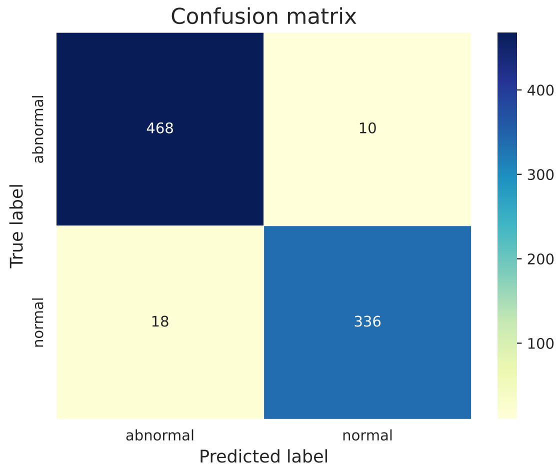

Fig. 8. Confusion matrix for trained model performance on testing data.

Fall detection systems [14] solve a related problem: detecting falls in order to help the elderly and their caregivers, and subsequently triggering alarms and notifying emergency services. However, these systems focus on detecting a fall after it has occurred. My research focuses on preventing falls before they occur.

Many existing fall prevention systems assess users' medical and behavioral histories and surveys to predict the likelihood of a fall [12]. My work does not currently use user information.

There are two related works that focus on using smartphones in predicting falls [7], [8]. Their work takes a feature extraction approach: they combine the horizontal accelerometer signals with pitch and roll data to create 4 tilt-invariant signals, and extract 3 quantitative features from these signals to create a feature vector. The system uses a decision tree classifier to sort the data into classes based on the features of the training data. Their results show that this approach works best only if the user's data was a part of the training process. For new data, the system accuracy is poor. In contrast, my system uses an LSTM-based approach and is able to classify data from a user whose data was not used while training with a high accuracy.

Furthermore, this research implements the created model into a smartphone app, making the idea practical and accessible. WalkSmart synthesizes past efforts to capitalize on gait irregularities and builds upon them to be a highly accurate, accessible, and useful application that has many implications for ambulatory fall prevention.

VI. Conclusions and Future Work

I presented a new neural network model called WalkSmart for detecting and mitigating ambulatory falls. My results show that smartphone accelerometer and gyroscope sensor data can be used as an accurate and effective predictor of gait variance. The LSTM model is able to predict normal, abnormal, and irregular gait with a test accuracy of 96.6%, a precision of 0.963 and a recall of 0.979. These values indicate the strong possibility of using smartphone data to detect falls, as irregular gait is a common predictor of falls. I have also integrated my model into a smartphone app that interacts with a cloud server to perform machine learning inferencing on the user's gait data, and provides alerts of potential falls.

These results also have many future implications. By increasing the number of epochs the model trains for as well as the amount of participants in the study, the model can predict steady and unsteady gait patterns with a higher accuracy. Additionally, adaptive learning can be used to personalize the model to the user, thereby making it more accurate. As future work, I plan to allow users to report falls and upload their smartphone sensor data through the WalkSmart app, so that the model accuracy can be further improved.

Acknowledgments

I would like to thank my parents for providing educational resources to investigate the problem, and my participants for providing test data for use in the project.

References

[1] Demographics of mobile device ownership and adoption in the united states, pew research center. https://www.pewresearch.org/internet/fact-sheet/mobile/.

[2] Keep on your feet - preventing older adult falls. https://www.cdc.gov/injury/features/older-adult-falls/index.html.

[3] C. Chheng and D. Wilson. Abnormal Gait Detection Using Wearable Hall-Effect Sensors. *Sensors (Basel)*, 21(4), Feb 2021.

[4] Trupti Gujarathi and Kalyani Bhole. Gait analysis using imu sensor. In *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pages 1–5, 2019.

[5] Mohammad Ashfak Habib, Mas S. Mohktar, Shahrul Bahyah Kamaruzzaman, Kheng Seang Lim, Tan Maw Pin, and Fatimah Ibrahim. Smartphone-based solutions for fall detection and prevention: Challenges and open issues. *Sensors*, 14(4):7181–7208, 2014.

[6] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Comput.*, 9(8):1735–1780, nov 1997.

[7] Akm Jahangir Alam Majumder, Farzana Rahman, Ishmat Zerin, William Ebel, and Sheikh Iqbal Ahamed. iPrevention: Towards a novel real-time smartphone-based fall prevention system. In *Proceedings of the 28th Annual ACM Symposium on Applied Computing*, SAC '13, page 513–518, New York, NY, USA, 2013. Association for Computing Machinery.

[8] Akm Jahangir Alam Majumder, Ishmat Zerin, Miftah Uddin, Sheikh I. Ahamed, and Roger O. Smith. SmartPrediction: A real-time smartphone-based fall risk prediction and prevention system. In *Proceedings of the 2013 Research in Adaptive and Convergent Systems*, RACS '13, page 434–439, New York, NY, USA, 2013. Association for Computing Machinery.

[9] Muyinat Y. Osoba, Ashwini K. Rao, Sunil K. Agrawal, and Anil K. Lalwani. Balance and gait in the elderly: A contemporary review. *Laryngoscope Investigative Otolaryngology*, 4(1):143–153, 2019.

[10] Albert Pierce, Niklas König Ignasiak, Wilford K. Eiteman-Pang, Cyril Rakovski, and Vincent Berardi. Mobile phone sensors can discern medication-related gait quality changes in Parkinson's patients in the home environment. *Computer Methods and Programs in Biomedicine Update*, 1:100028, 2021.

[11] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. *Learning Internal Representations by Error Propagation*, page 318–362. MIT Press, Cambridge, MA, USA, 1986.

[12] S. Sato, T. Kakamu, T. Hayakawa, T. Kumagai, T. Hidaka, Y. Masuishi, S. Endo, and T. Fukushima. Predicting falls from behavioral and psychological symptoms of dementia in older people residing in facilities. *Geriatr Gerontol Int*, 18(11):1573–1577, Nov 2018.

[13] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.*, 15(1):1929–1958, jan 2014.

[14] Carlo Tacconi, Sabato Mellone, and Lorenzo Chiari. Smartphone-based applications for investigating falls and mobility. In *2011 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops*, pages 258–261, 2011.

[15] Joe Verghese, Anne Felicia Ambrose, Richard B. Lipton, and Cuiling Wang. Neurological gait abnormalities and risk of falls in older adults. *Journal of Neurology*, 257:392–398, 2010.

[16] Tirra Hanin Mohd Zaki, Musab Sahrim, Juliza Jamaludin, Sharma Rao Balakrishnan, Lily Hanefarezan Asbulah, and Filzah Syairah Hussin. The study of drunken abnormal human gait recognition using accelerometer and gyroscope sensors in mobile application. In *2020 16th IEEE International Colloquium on Signal Processing Its Applications (CSPA)*, pages 151–156, 2020.

[17] Qin Zou, Yanling Wang, Qian Wang, Yi Zhao, and Qingquan Li. Deep learning-based gait recognition using smartphones in the wild. *IEEE Transactions on Information Forensics and Security*, 15:3197–3212, 2020.