Predictive Modelling of Biomechanical Joint Angles via Ensemble Learning

Name: Bingi Srinath

Roll no:x23335483

I. INTRODUCTION

Human gait has taken its position as a pillar in the research area of biomechanics, clinical diagnostics and rehabilitation science. Gait data is the form of obtaining and processing the movement of the body joints, being protected at different conditions, when walking. The data included in the project is taken on the UCI Machine Learning Repository and includes more than 180,000 observations that have been gathered on the basis of several subjects. Other characteristics which include leg (left/right), type of joint, and replication assist in representing the multivariate state of the human walking patterns.

The aim behind using this data is to model and comprehend the biomechanical alteration on the joint angles depending upon how the gait was captured. This allows a greater understanding of the response of the various joints to movement as a consequence of time, this in turn can be useful in application to physiotherapists, design of prosthetics and study of motion pattern. A quantitative analysis of such gait patterns allows researchers and clinicians to discover possible abnormalities as well as optimise rehabilitation efforts and improve performance monitoring both in medical and in sporting situations.

II. GOALS / BUSINESS VALUES

The main idea of this project is to come up with a predictive system which may work out well to estimate the angles of the joints by analyzing multivariate data of gait. The project will examine important variables like time during the gait cycle, walking status, joint type and leg side and how this affects the pattern of joints movement. Such predictive functionality is not only useful in observing a normal human locomotion but also in forming a basis of identifying deviations, which can signal physical impairment conditions. Based on such analysis, the project is data-driven and aims at precise and reliable prediction and analysis of joint angles with high accuracy that is critical in biomechanical assessment.

The potential business and clinical value in this project is to support healthcare professionals, sports scientists and assistive technology developers through health and illness management. Having proper gait modeling can also be useful in physiotherapy where treatment programs can be made specific through prediction of joint behavior. It is also capable of playing a major role in identifying the presence of gait abnormalities at an early stage, eliminating the use of costly and invasive tests to obtain a diagnosis. Moreover, such predictive models, in their turn, can be used in the design of the prosthetics and orthotic devices, performance monitoring system in athletes, and the creation of the smart rehabilitation and wearable health technology.

III. LITERATURE SURVEY

Gait analysis of humans plays a crucial part aspect in present-day biomechanics, physiotherapy, and human-computer interaction. Gait-related angles of joints and motion of the limbs can now highly accurately be predicted with the emergence of wearable sensors and machine learning models. Such predictive models are extensively used in clinical tests, sports analysis, and rehabilitation treatment. Modeling of human gait data, especially joint angles, leads to extraction of patterns by learning scientists that can lead to prediction of the future motion or detect anomalies.

Mannini and Sabatini (2010)[2] are the first who used machine learning to classify physical activity based on body-worn accelerometers. They emphasized that it is possible to achieve correct activity recognition using the combination of statistical characteristics and classic classifiers. This paper opened the way to the predictive model of the raw motion data acquired by the help of the sensors.

The study by ***Hannink et al. (2017)[3]*** addresses the applicability of deep convolutional neural networks (CNNs) in deriving gait parameters out of data captured by wearable instruments. They improve the accuracy of stride length and walking speed estimation without feature selection. The method is especially useful in clinical practice, when it comes to evaluating the mobility-impaired patients. The experiment shows the possibility of the use of wearable technology and deep learning to monitor gait in real-time automatically. It presents the contribution of AI to the enhancement of the diagnostic and treatment measures within healthcare.

The ***Alaqtash et al. (2011)[4]*** work estimates the gait phases and joint angles on IMU-based wearable sensors in real-time by using fuzzy logic. Their method was very promising in regard to translating motion signals into biomechanical angles but restricted by problems of generalization. As it is stated in their work, there is a necessity in scalable data-oriented models.

Based on these studies, one will quickly recognize that gait prediction does not occur anymore through heuristic and fuzzy schemes but rather through well-developed learning systems. But problems with outliers, encoding (categorical) or generalizability has beset earlier models. We complement these in our project by introducing the robust kind of preprocessing (e.g. one hot encoding, robust scaling) and the use of XGBoost Regression, which is a high performance model and is robust when faced with skewed and noisy data. This assists in better prediction of angles and reducing the limitations observed in the past studies.

IV. ETHICAL CONCERNS

Utilization of human gait data in any project becomes important; this is due to the usage of the data and privacy and consent. Due to the curious fact that gait patterns can be identified by people, the anonymity of the participants is of highest concern in terms of their security. The publicly available hyper-local dataset is going to be used; however, it will have to be confirmed in practice that nobody would tamper with it by trying to re-identify the subjects or tamper with data and infringe the rights of individuals. Moreover, the developers and the analysts should not make any unprofessional health diagnosis or declare claims unsubstantiated clinically and having no domain expertise.

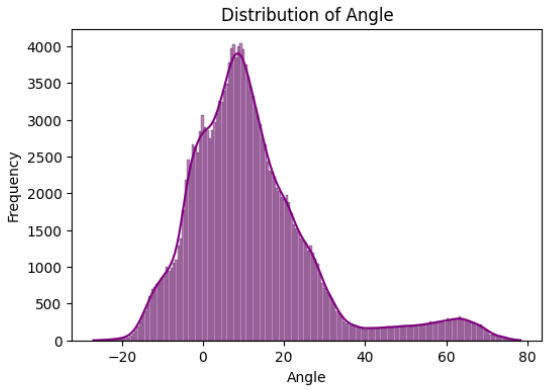
V. MODEL TRAINING

Implementation process commenced by carrying out extensive exploratory data analysis (EDA) in order to interpret structure and distribution of multi variate gait data. The outlier detection was made on the basis of the Interquartile Range (IQR) method that has led to the discovery of 10,762 outliers in the column of angle. No OUTLIERS were found in the other columns. This assisted in concentrating the attention on the preprocessing. Results of the imbalance check (Perfect balance across all categorical features) implied that the dataset was not imbalanced and thus there was neither a need to resample the dataset.

As a skewed value was observed in the target variable angle (skewness 1.38), it was necessary to think of transformation or robust ways of modeling reliably. Several visualizations were developed to aid data comprehension which includes bar plot, boxplot, pair plot, and distribution plot. These images gave information about feature relationship and variance in angle conditionally and across joints and time.

The categorical variables that include: condition, leg, and joint have been converted into appropriate machine learning format by using One-Hot encoder to encode them into a number structure. The skew and the outliers in the target numerical column were used to apply the RobustScaler method to only the angle column specifically and stabilized the data to eliminate highly deviated values. The dataset was also divided into training and testing sets. The XGBoost Regressor model was trained and prediction made on the basis of its high accuracy, regularization and the fact that it can easily work with structured tabular data.

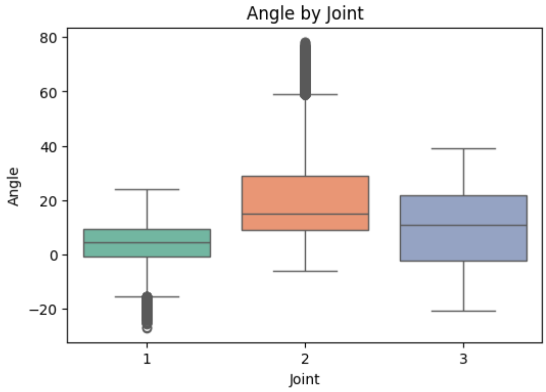
Figure 1: Angle Distribution Plot



Insights:

- The preparation is skewed to the right which proves that the target variable is not normally distributed.

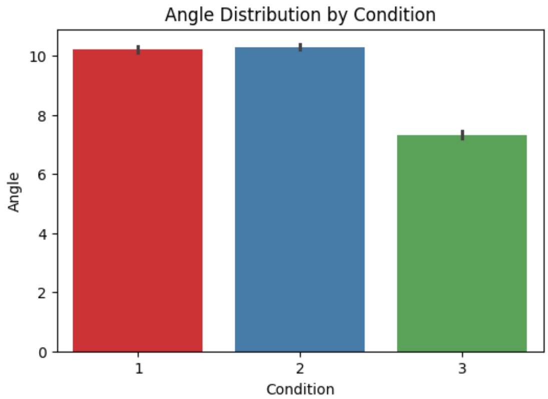
Figure 2:Joint Type Boxplot of Angle



Insights:

- The hip joint has the highest number of values of angles and this indicates a more dynamic movement.

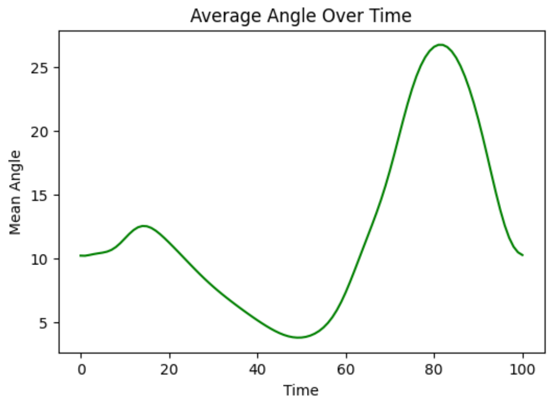
Figure 3:barplot Mean Angle condition



Insights:

- The average joint angle that is recorded by the walking condition of Slow is slightly lower than the normal walking and Fast.

Figure 4:Time Dependent Mean Joint Angle



Insights:

- Stable Movement Pattern: The tendency of the joint angle over time is smooth and sinusoidal, which is typical of a rhythmic motion of joints when walking.

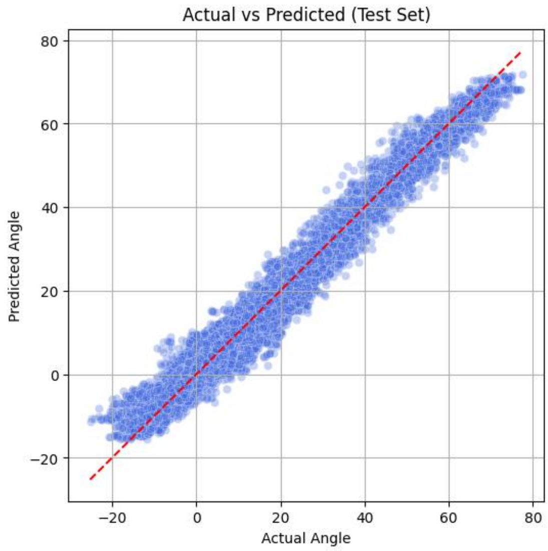
VI. RESULTS & BUSINESS VALUE INTERPRETATION

XGBoost Regressor model has been applied to the multivariate gait dataset, where the outcomes were very accurate. The model was evaluated and produced the following results after data preprocessing actions (outlier detection, consideration of skewness with robust scaling, encoding of categorical data, and subset separation), respectively:

- RMSE (Root Mean Squared Error) training: 2.643
- Scores of Trainings R 2: 0.973
- RMSE Testing result: 2.693
- R 2 Score of Testing: 0.972

These measures indicate that the model has good generalizability on unexplained data with minimal error, and that it is able to explain more than 97 percent of the variance in the target (joint angle) values thereby pointing to a resounding and sound forecasting system.

Figure 5:Predicted Vs Actual Angle Plots



Insights:

- The points are southeast with the diagonal line meaning that there is a vehement symmetry between the forecast values and true values.

Business Value and Interpretation Qualitative

Biomechanically, this predictive model is very useful in business and clinical terms concerning rehabilitation monitoring and analysis. The possibility of predicting joint angles that can be obtained on time series and condition of movement data with considerable accuracy can:

- Assist the physiotherapists to recognize early abnormal movement of the joint and individualize the exercise programs.

- ➢ Get incorporated in the wearable devices to monitor gait, wherein the elderly patients or patients post-surgery can get real-time feedback.
- ➢ Be used as a data-driven base of automated gait analysis systems and unmanned the number of efforts needed to provide manual analysis and raise the diagnostic accuracy.

These results have shown that the possibility of high-performance measures and significant real-world results can be accomplished by employing advanced regressive methods like XGBoost along with a good preprocessing and assessment when it comes to healthcare and bio-mechanics.

VII. REFERENCES

[1] Dataset: Multivariate Gait Data - UCI Machine Learning Repository

[2] A.Mannini and A. M. Sabatini, "Machine Learning Methods for Classifying Human Physical Activity from On-Body Accelerometers," Sensors, vol. 10, no. 2, pp. 1154–1175, 2010. [Online]. Available: https://doi.org/10.3390/s100201154

[3] J. Hannink, M. Kautz, M. Pasluosta, J. Klucken, B. M. Eskofier, "Sensor-based gait parameter extraction with deep convolutional neural networks," IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 1, pp. 85–93, Jan. 2017. [Online]. Available: https://doi.org/10.1109/JBHI.2016.2636456

[4] M. Alaqtash et al., "Application of wearable sensors for human gait analysis using fuzzy computational algorithm," Engineering Applications of Artificial Intelligence, vol. 24, no. 6, pp. 1018–1025, 2011. [Online]. Available: https://doi.org/10.1016/j.engappai.2011.03.008