National College of Ireland

Genetic Algorithm Optimized Deep Learning Model for Parkinson Disease Severity Detection

MSc Research Project
Data Analytics

Shveta Srivastava
Student ID: 18194851

School of Computing
National College of Ireland

Supervisor: Dr. Rashmi Gupta

National College of Ireland
Project Submission Sheet
School of Computing

National College of Ireland

Student Name:	Shveta Srivastava
Student ID:	18194851
Programme:	Data Analytics
Year:	2021
Module:	MSc Research Project
Supervisor:	Dr. Rashmi Gupta
Submission Due Date:	16/08/2021
Project Title:	Genetic Algorithm Optimized Deep Learning Model for Parkinson Disease Severity Detection
Word Count:	6923
Page Count:	24

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	15th August 2021

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Genetic Algorithm Optimized Deep Learning Model for Parkinson Disease Severity Detection

Shveta Srivastava
18194851

Abstract

Parkinson's disease (PD) is reflected by several motor function disabilities such as tremors, loss of balance, speech impairment, etc. due to loss of dopamine neurotransmitter. While researchers have been building models for diagnosing and classifying PD patients based on gait data, speech data or handwriting data accounting for visible symptoms only, PD patients show non-motor symptoms such as sleep disorder, neuropsychological symptoms, cognitive impairment, olfactory loss, much before the actual diagnosis. This study involves coupling both motor as well as non-motor symptoms of PD patients from up to 10 years longitudinal records in PPMI database and building an optimized deep learning model for PD severity classification based on the Hoehn & Yahr index. This longitudinal complex dataset brings along challenges of dealing with high volume of missing and inconsistent data in various assessments at different time points. To deal with such complexity, the proposed model for this study is a type of Recurrent Neural Network, Long Short-Term Memory (LSTM) model which learns well from data with long term dependencies. This multi-time step model gives high accuracy of 88% for multi-class severity prediction. The LSTM model is also coupled with heuristic evolutionary search algorithm, Genetic Algorithm (GA) considering the vast dimensionality of the longitudinal heterogenous records and to find the optimal window size and number of LSTM units to minimize the loss function, MSE. The results have been compared with baseline model, state-of-the-art approach, MLP (Multi-layer Perceptron) which is a feed forward network. The novel GA-LSTM model used in this project shows reduced RMSE score of 0.33 as compared to 0.72 in MLP. Multiple Machine learning algorithms have also been implemented where XGBoost shows the highest accuracy of 89%.

1 Introduction

Parkinson Disease (PD) is second most common progressive neurodegenerative disorder in the world and is associated with loss of Dopamine producing cells i.e., Substantia Nigra in the mid-brain section. Since this chemical Dopamine plays an important role in controlling the movement of the body, the patient with PD experiences loss in control of the movement of their limbs and experiences tremors, bradykinesia (slower movement), impairment in speech and gait, etc.[1]. Apart from motor symptoms, Parkinson's disease is also accompanied by certain non-motor symptoms like sleeplessness at night, fatigue,

[1] `https://www.parkinson.org/understanding-parkinsons`

anxiety, depression, cognitive impairment, or urinary problem. Mostly diagnosis of PD has been carried out based on tremors or postural instability such as from data related to gait, speech or handwriting which suggests distinctive patterns to distinguish a PD patient from normal controls. Whereas there has been evidence that non-motor symptoms exist in PD much earlier than advancement to motor symptoms and continues throughout at all stages. It has been recently found that almost all the PD patients experiences some sort of non-motor symptoms along with the usual motor symptoms which depends on the level of severity of disease (Rodriguez-Blazquez et al.; 2021). Figure 1 shows the motor and non-motor symptoms of Parkinson's Disease represented by an iceberg where visible motor symptoms are just like the 'tip of an iceberg' which multiple underlying invisible non-motor symptoms. Hence, evaluation of non-motor symptoms is an important element and should be considered for effective diagnosis of PD and its severity (Armañanzas et al.; 2013; Zhang et al.; 2019).

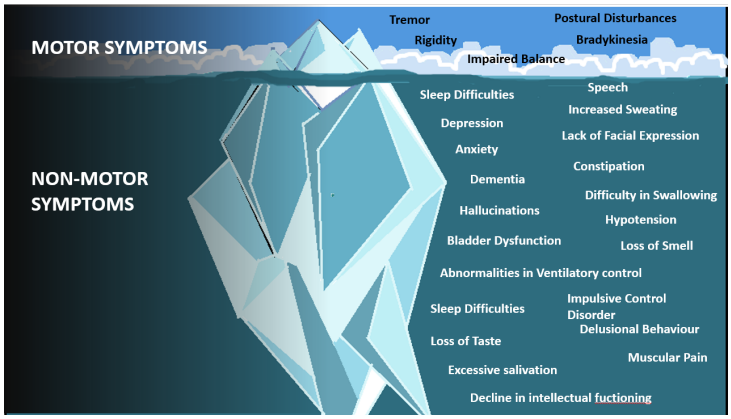

Figure 1: Motor and Non-Motor Symptoms of Parkinson's Disease represented by an 'Iceberg' where motor symptoms are just the tip of the iceberg

Diagnosis of Parkinson's disease severity in patients is proved to be beneficial for patient's counselling based on disease prognosis and considering plausible options for treatment (Samantha, K. et. al, 2017). This paper attempts to build a deep learning algorithm which assesses PD patient's clinical data over a period of 5-10 years, indicating their motor and non-motor symptoms for accurate detection of severity of disease. This model will help the medical practitioners identify patients in need of special attention and medication based on their severity. Several clinimetric scales have been employed by researchers and neurologists to detect the severity of Parkinson's disease patients. The most widely used scale is Hoehn and Yahr (HY) which quantifies the disease progression as five stages of disability, starting from least severe 'stage I' to most severe bed-ridden patients in 'stage V' (Hoehn and Yahr; 1967). Another very common scale is Movement Disorder Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS) (Opara et al.; 2017). This paper classifies patients on the basis of their HY scores.

Longitudinal records of patients consists of multivariate time series records which can be very precious in terms of building models but comes with the complexities of missing records, missing variables, irregular sampling, multi-dimensionality issues (Lipton et al.; 2015). These complexities accompanied by the interdependence of vast number of predictor variables cannot be captured by traditional machine learning models (Su et al.; 2021). Hence, a Recurrent Neural Network (RNN) based model, Long Short-Term

Memory (LSTM) which learns well with the long-term dependencies of the patients data has been proposed in this study (Lipton et al.; 2015; Razavian et al.; 2016; Djerioui et al.; 2020). LSTM has been successfully applied for sequential data in other areas such as stock market prediction (Chung and Shin; 2018), load forecasting (Al Mamun et al.; 2019) and natural language processing (Gorgolis et al.; 2019), while its application in medical field is still lacking and could prove to be beneficial for diagnosis and prediction of illnesses.

Another approach to deal with complex medical data gaining popularity now-a-days is evolutionary algorithm called Genetic Algorithm (GA)(Huang and Wang; 2006). GA uses optimisation approach based on the biological theory by Charles Darwin called 'Survival of the Fittest' to select the best feature window for building model with minimal error and maximum accuracy (Salmanpour et al.; 2020; Soumaya et al.; 2021; Kara; 2021). This novel approach of GA coupled with deep learning model LSTM has been employed in this research as an attempt to give an optimised model and has shown improved results over previous work for classification of PD severity of a patient.

1.1 Research Question

This paper attempts to answer the following research question, "How well can Genetic algorithm optimised LSTM model on longitudinal clinical assessments data, help in determining Parkinson's disease severity?".

Training deep learning model with longitudinal data could be a complicated task considering the multiple time steps of each patient owing to the visits and complex multi-variate data. Hence, coupling the model to find optimal inputs and minimized error could help deal with complexities.

1.2 Research Objective and Contribution

The paper derives its motivation from the importance of inclusion of non-motor symptoms in the diagnosis of severity of PD along with the motor symptoms. The main objective of the paper is to build a deep learning model that learns from the longitudinal dependencies of complex clinical data. Applying the novel approach of optimised hybrid model using GA-LSTM for prediction of PD severity on the longitudinal data gives improved results when compared against baseline model, multi-layer perceptron (MLP).

The challenges of diagnosis of various diseases from multiple parameters or symptoms associated with patients leads to the compelling need of building automated tools which is an ongoing research area. There is still a need for more research for building more accurate and improved models as compared to previous ones. Parkinson's disease (PD) needs particular attention in relation to building automated diagnostic models because of the increased number of cases each year and the complexity associated with this neurological disorder. This automated deep learning model can be helpful in accurate diagnosis of Parkinson's disease severity of the patient for further treatment options and counselling by the practitioners or neurologists.

1.3 Plan of Paper

The structure of the paper is as follows. Section 2 entails the previous research work related to Parkinson's disease and the machine learning techniques used in this project.

Section 3 explains the methodology used in this project. Section 4 shows the design specification. Section 5 explains step by step approach of implementation of the proposed model. Section 6 shows evaluation of results. Finally, Section 7 consists of conclusion and future work.

2 Related Work

Substantial amount of research has been conducted to build machine learning models for diagnosis of PD or its classification, mostly based on visible motor symptoms of the patients such as gait or speech variability.

2.1 Biomarkers for PD diagnosis

Because of its associated complications, Parkinson's disease has been diagnosed using multiple biomarkers that mark the onset of this disease by different researchers. A recent study used Vertical Ground Reaction Force (VGRF) from gait data to build Artificial Neural Network (ANN) to diagnose PD with 97% accuracy and determine patient's severity using HY scale with 87% accuracy (Veeraragavan et al.; 2020). Similar work using ANN for prediction of patient's disease severity using HY scale on gait data was conducted by (Varrecchia et al.; 2021). The study included about equal number of PD patients and healthy controls. The PD patients were on constant medication for at least last two weeks and with no FOG (Freezing of Gait) episode. They used PCA for feature selection and found no significant difference between all the PCA features and the proposed subset after selection. Apart from gait, speech variability is another common symptom of advanced PD and has been extensively used to distinguish PD from normal controls (Soumaya et al.; 2021). Many researchers have used multiple modalities to apply regression and classification techniques. For example, Lei et al. (2018) used clinical study variables like sleep, depression, olfactory, cognitive scores, etc. along with patient's MRI scans, CSF biomarkers, to diagnose PD. The non-motor symptoms originate in PD much earlier in time and shows variability as the disease progresses. This shows that including non-motor symptoms is very important along with motor symptoms for PD diagnosis and disease severity prediction models.

The first study combining non-motor symptoms with motor symptoms for PD classification was done in 2013 by (Armañanzas et al.; 2013). They combined the two assessment scales, HY and CISI-PD scales where HY stage scale refers to motor symptoms only, whereas CISI-PD comprises other cognitive disabilities and non-motor complications. They used Estimation of Distributed Algorithms (EDA) as feature selection for building machine learning algorithms i.e., Artificial neural networks (ANN), C4.5 Decision trees (DT), Linear Discriminant analysis (LDA), Naïve Bayes (NB) and K-nearest neighbours (KNN). They showed the relative importance of using non-motor symptoms of PD patients in their diagnosis and stage classification. In another work combining the two scales UPDRS and HY scales to determine PD severity explored a combination of different methods for feature selection and machine learning (Tsiouris et al.; 2017). Wrapper-based feature selection method when combined with an evolutionary decision tree model called RIPPER (Repeated Incremental Pruning to Produce Error Reduction) algorithm showed the highest accuracy in this study. Prashanth and Roy (2018) concluded from their study involving combined UPDRS and HY scales that the features such as tremor, handwriting, bradykinesia and facial expression contribute the most for

building an accurate model as determined by Random Forest. The model they built to determine PD severity prediction comprised support vector machine (SVM), ordinal logistic regression (OLR), AdaBoost and RUSBoost with 97% accuracy. Severity prediction deep learning model (DNN) based on UPDRS scores showed improved accuracy as compared to SVR as discussed by Grover et al. (2018).

In another novel approach to identify PD patient's non-motor subtypes based on motor subtypes, (Ren et al.; 2020) found that Tremor Dominant (TD)/ Indeterminate /postural instability and gait disturbance (PIGD) based classification is the most suitable motor subtype for identification of patient's non-motor symptoms. It has been quite evident from all the research works that there is a strong connection between patient's motor and non-motor symptoms. The motor symptoms examined by UPDRS scale is hence combined with non-motor assessment exams from PPMI database in this study, to form an effective tool for building machine learning model for disease stage classification.

2.2 Challenges accompanying medical data

Medical data constitutes longitudinal records of patients over the period of several weeks to years of follow-up visits. One of the biggest challenges in working with longitudinal data is dealing with inconsistent patient records. Since many patients drop out in between or do not turn up for the follow-up assessments, this inconsistency in records leads to bias in data (Fitzmaurice et al.; 2008). Another challenge is missing data of patient for a number of assessment tests by the medical practitioner that are not necessary at that moment of time.

Lipton et al. (2015) used resampling of all the missing values with forward and backward filling within the window of one hour for each visit. When the entire variable record is missing, the authors imputed them with the value that is clinically normal as denoted by experts. Su et al. (2021) in a recent study, used K-nearest neighbour (knn) imputation method to fill the missing records in an approach for prediction of depression in elderly people living in China. According to the authors this approach will obtain a value from all the related cases in the complete data frame based on Euclidean distance. Here, KNN is a non-parametric approach for imputation and is independent of features and label. Bang et al. (2020) in a comprehensive study on missing values in longitudinal data for electronic health records (EHR) of patients, applied phased-LSTM model to address the issue of irregular sampling of medical data.

However, all the approaches used for imputation of missing values in medical data discards some of the useful information about the patient's observation and its time. Also, missing values should not be imputed if they are meant to be left blank by the medical practitioner owing to some assumptions about normalcy of record or fulfillment of some underlying diagnostic criteria. Hence, even if researchers benefit from studying temporal aspect of longitudinal data, its complicated analysis due to missing variables and values, irregular visits by patients, irregular time-varying effects and complex correlation among variables, is a very challenging problem that needs to be addressed (Garcia and Marder; 2017).

2.3 Learning with Longitudinal Data

Various researchers have built machine learning algorithms to understand the disease progression from this complex time-series data. Holden et al. (2018) performed mixed

model regression to obtain the MDS-UPDRS scores from PD cohort over 5 years data. The results indicated a substantial increase in scores by 4.7 points every year indicating the disease progression in PD patients. On the other hand, Evers et al. (2019) used Gaussian Linear state Space Model to estimate the disease progression. The MDS-UPDRS scores were used to model the longitudinal data by calculating the differences in variance introduced by PD progression. They also considered the short-term effects in patients scores as well as the measurement errors. The study indicated the presence of high error variance in the scores because of heterogeneity of the patient's data and highlighted the need to deal with this heterogeneity of PD progression in longitudinal data with a more reliable tool.

Dealing with such issue, deep learning models have seen improved results and have been successfully implemented in longitudinal records of patients with other illnesses too. Zhao et al. (2019) used clinical electronic health record (EHR) of 10-years to build machine learning models for cardiovascular disease prediction. Their base models comprised Random Forest (RF), Logistic Regression (LR), Gradient Boosting Tree (GBT) for longitudinal features. Whereas after incorporating genetic data to the previous clinical records, the results improved using deep learning models, CNN and LSTM with AUROC of 0.79. In another work related with longitudinal records of Alzheimer's disease patients, Lei et al. (2020) predicted clinical scores using Support Vector Regression (SVR). They applied a novel feature encoding method using Deep Polynomial Network (DPN) and discovered a relationship between patient's clinical scores and MRI scans over time.

One of the popular approaches for handling complex long-term time-series data is Long Short-Term Memory (LSTM) model (Lipton et al.; 2015; Razavian et al.; 2016). In a study related to Parkinson's disease, Zhao et al. (2018) built a hybrid CNN-LSTM model where Convolutional Neural Network (CNN) learns well from the spatial features from gait data whereas LSTM learns well with the patient's temporal features in order to predict PD patient's severity. This spatio-temporal model achieved better improved results as compared to base models for classification. Zhang et al. (2019) worked on identification of PD patient's subtypes, which referred to disease severity from mild baseline with moderate progression of motor symptoms to moderate baseline with mild progression and finally to most severe baseline having rapid disease progression. They used data for motor and non-motor symptoms combined with their neuroimaging scans such as SPECT and DaTScan, and their biospecimen such as Cerebrospinal fluid. In their novel approach, they learnt patient's representations through sequence of records using LSTM and then used Dynamic Time Warping (DTW) to calculate the distance between the patients for k-means clustering for patient subtyping. Djerioui et al. (2020) compared two deep learning algorithms, Multi-layer Perceptron (MLP) and LSTM models for prediction of heart diseases and found LSTM to be better performing model for this longitudinal data analysis. Hence, Long-short term memory model works well with longitudinal data, but training the model with manual selection of hyperparameters could be complicated owing to multi-dimentionality.

2.4 Evolutionary Methods for Optimization of Model

Large number of features from complex longitudinal data lead to generalisation of machine learning models and may lead to redundancy. This issue is commonly termed as a curb of dimensionality and usually results in overfitting of the models, hampering their accuracy. Therefore, one of the most essential steps in machine learning and data

mining is feature selection or dimensionality reduction for categorisation of features into meaningful representation of components.

Salmanpour et al. (2020) recently in their study with PPMI clinical study concluded that evolutionary feature selection methods, genetic algorithm and ant-colony optimisation are the best methods for better prediction of PD. They used features from several assessments from PPMI database such as patient's demographics, baseline MDS-UPDRS scores as well as DAT SPECT images. They combined different feature selection methods such as Simulated Annealing (SA), Particle Swarm Optimization Algorithm (PSO), Differential Evolution Algorithm (DEA), Genetic Algorithm (GA), Nondominated sorting genetic algorithm (NSGAII) and Ant Colony Optimisation (ACA) with 11 different machine learning algorithms. Local Linear Model Trees (LOLIMOT) combined with GA and ACA minimised the error significantly resulting in the best model. Soumaya et al. (2021) in their study for detecting PD using speech signals, coupled genetic algorithm with SVM classifier and achieved high accuracy of 91%. GA facilitated in minimising the dimensions of the feature space while maximising the accuracy of the model. Genetic Algorithm has also been applied on functional MRI scans to find the regions of interest impacting patient's condition. Wutzl et al. (2019) coupled GA with SVM to classify severe chronic disorder of consciousness from patient's MRI scans.

Many studies have showed the application of GA to optimize the hyperparameters of deep learning models. Kilicarslan et al. (2021) classified anaemia patients into three categories by estimating Anaemic patient's blood count test results using optimized Convolutional Neural Network (CNN) and Stacked Auto-encoders (SAE) with GA. They proved their model showed improved accuracy of 98% compared to base models.

2.5 Hybrid Learning Algorithm

Since longitudinal PD records comes with the long-term time-series format, LSTM model has shown to work with higher accuracy as compared to other machine learning algorithms such as SVM, ANN, etc. Hence, related work with the combination of two techniques GA and LSTM has been explored and found that application of GA coupled LSTM model is still lacking in the medical sector related to classification or prediction of diseases.

Apart from its application as feature selection method, GA has also been used for hyperparameter optimisation of machine learning algorithms. Chung and Shin (2018) in their study related to stock market prediction built GA optimised LSTM model and determined the topology and size of the time window using this approach. Al Mamun et al. (2019) built a similar model to predict electrical load forecasting using GA-LSTM. Another work in Natural Language Processing (NLP) domain by Gorgolis et al. (2019) successfully implemented hyperparameter optimisation of LSTM model using Genetic Algorithm.

Besides these domain, there have been studies related to prediction of illnesses using GA-LSTM models. Nejedly et al. (2019) applied Genetic Algorithm to find optimal hyperparameters to train Long Short-Term Memory (LSTM) Neural Network model for prediction of sepsis for ICU (Intensive Care Unit) patients in the hospital. Study by Rashid et al. (2019) aimed at building a machine learning tool in the health sector for disease classification and prediction in conjunction with two optimization techniques, Biogeography-based Optimisation (BBO) and GA. They worked for prediction of two common illnesses, Diabetes and Cancer. The authors concluded that LSTM coupled with GA and BBO gave significantly better accuracy than previously commonly used al-

gorithms on health data such as Multi-layer Perceptron and SVM which has the drawback of learning with time-series data.

In Table 1 the list of some of the related work is provided.

Table 1: Summary details of Related work for Parkinson's Disease Detection

Authors	**Method**	**Data**	**Result**
Armañanzas et al. (2013)	NB, KNN, LDA, DT, ANN	PPMI clinical scores	Five different classifiers gave accuracy between 72%-92%
Bang et al. (2020)	Phased-LSTM	Longitudinal EHR data	AUC of 74%
Lei et al. (2020)	Deep Polynomial Network	clinical data for Alzheimer's disease prediction +MRI	MAE was compared with RNN and LSTM at different time points, DPN outperformed
Ashour et al. (2020)	LSTM	PD Gait data for FOG Prediction	83% average accuracy was achieved with LSTM model
Soumaya et al. (2021)	GA+SVM	Speech data for PD detection	Accuracy of 91% with model
Varrecchia et al. (2021)	ANN	Gait data for PD prediction	AUC of 0.73-0.88
Grover et al. (2018)	DNN	PD clinical data	DNN with 63% accuracy as compared to SVR
Djerioui et al. (2020)	MLP +LSTM	Heart disease prediction	LSTM outperformed with accuracy of 96.5%
Zhao et al. (2019)	NB, KNN, LDA, DT, ANN	Clinical data for EHR + genetic data	incorporating genetic data into clinical data showed improved AUROC values for all the models
Lipton et al. (2015)	LSTM	clinical data for ICU	Model outperformed MLP with AUC of 0.86

3 Methodology

The research methodology used for this project is Knowledge Discovery in Databases (KDD). The main requirement of KDD is proper understanding and knowledge of data and project goals and its domain for application. The design specification for the model used in this research project is illustrated in Figure 2.

The different stages of methodology involve starting with data collection and merging the datasets together, then pre-processing and standardization, followed by transform-

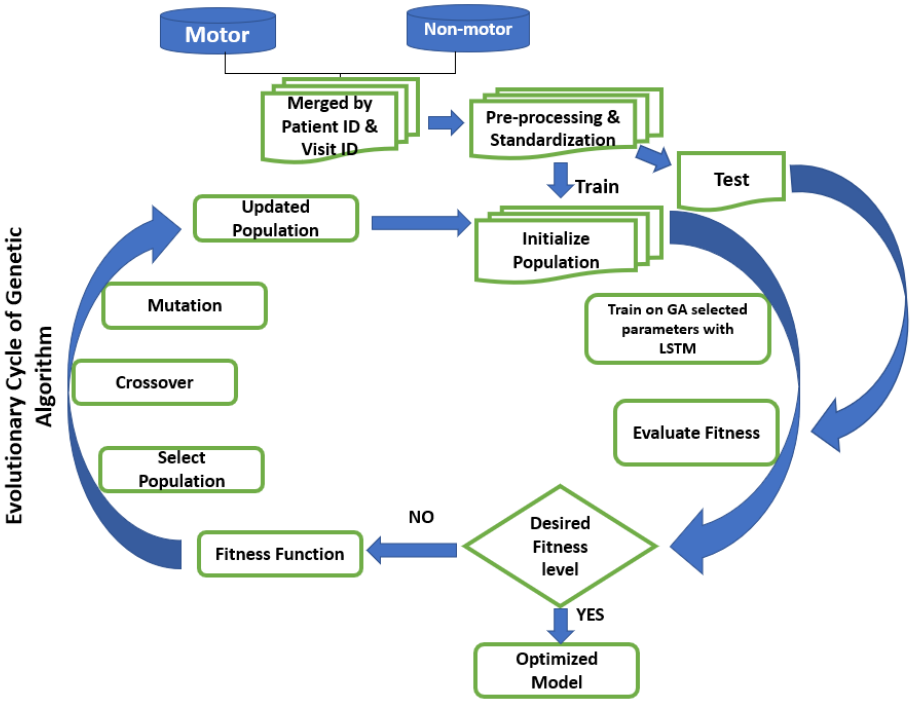

Figure 2: Design Specification for GA-LSTM model

ation using GA, modelling using LSTM and evaluation. Pre-processed feature space of train data is used to initialize the population to go through the genetic algorithm cycle for producing the best fit LSTM model. The following subsections explains all the steps.

3.1 Data Collection and Understanding

The Parkinson's disease study datasets were collected from Parkinson's Progressive Markers Initiative (PPMI) founded by Michael J. Fox Foundation [2] where over 2000 PD patients along with healthy controls participate with their consent for collection of their data in this central repository following standard ethical protocols for data acquisition. Patients are monitored for their disease severity and progression at various assessment scales at regular intervals for a period of 5-10 years. The complete evaluation of PD involves different cognitive tests for motor and non-motor function impairments, brain MRI scans, genetic data, etc. These comprehensive datasets are available for analysis and research on request.

The proposed research involves combination of motor and non-motor symptoms assessments. Previous work on longitudinal data has shown variance with regards to patient's symptoms during follow-ups over years and hence will be appropriate to apply in the prediction model of disease severity (Armananzas et. al, 2013; Simuni et. al, 2018). Combining other modalities has been avoided in this project as non-linearity of complex longitudinal data coupled with different modalities would add complexity which would interfere with model building process (Prashanth and Roy, 2018).

Understanding PPMI data required comprehensive decoding of the assessment codes from the datasets and careful selection of data for building model. Motor assessments

[2]Data Source: `http://www.ppmi-info.org/data`

Table 2: Few of the clinical assessments related to MDS-UPDRS

MDS-UPDRS I	MDS-UPDRS II	MDS-UPDRS III
Cognitive Assessments	Speech	Speech Assessment
Hallucinations	Saliva and Drooling	Rigidity
Psychosis	Chewing and Swallowing	Facial Expression
Depressed Moods	Doing activities, hobbies	Finger Tapping
Anxiety	Handwriting	Toe Tapping,Leg Agility
Dopamine-dysregulation	Freezing	Tremors(kinetic,rest)
Apathy	Getting out of bed/car	Postural stability
	Walking and balance	Gait & Freezing of Gait

Table 3: Non-Motor assessments related to Parkinson's Disease

ASSESSMENT CODE	NON-MOTOR ASSESSMENTS
MOCA	Montreal Cognitive Assessment
STAI	State-Trait Anxiety Inventory
GDS	Geriatric Depression Scale
COGNITIVE	Cognitive Categorisation
EPWORTH	Epworth Sleepiness Scale
HVLT	Hopkins Verbal Learning Test
LNS	Letter Number Sequencing Test
SCOPA	Scale for Outcomes in PD – Autonomic Dysfunction
QUIP	Questionnaire for Impulsive-compulsive Disorder
BENTON	Benton Judgement of Line Orientation
REM	Rapid Eye Movement Sleep Behaviour Disorder
O-UPSIT	Olfactory test for smell
SDM	Symbol Digit Modalities
SFT	Semantic Fluency Test
LFT	Lexical Fluency Test
CLOCK	Clock Drawing Test
BOSTON	Boston Naming Test

were evaluated using MDS-UPDRS scores which comprised four stages of patient's motor impairment assessments pertaining to daily-lifestyle experiences. Few examples of this scale assessments has been shown in table 2. Stage IV evaluates the side-effects of medication leading to motor complications which is irrelevant in case of disease progression (Dinov et. al, 2016) and was not included for study. Furthermore, only off-medication data is taken into account for this project as medication might interfere with assessment scores and give false reading for disease severity.

The non-motor assessments includes tests for neuro-behavioural, Neuro-psychological and cognitive impairments as shown in table 3. The data sets for assessments 'BOSTON', 'CLOCK', 'LFT' and 'TRAIL' have less than 500 records, hence these have been omitted from analysis.

This project aims to build a model to mark the progression of Parkinson's disease patients who need medical attention and counselling for further treatment to slow down the progression . Figure 3 shows the five stages of PD on HY scale.

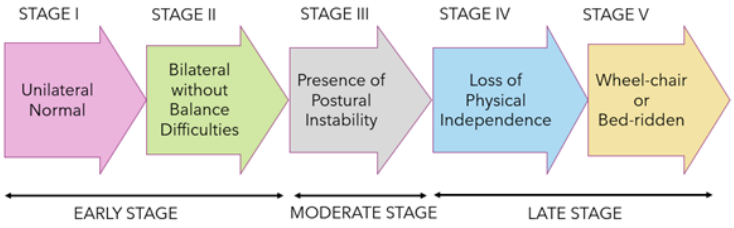

Figure 3: HY scale stages for Parkinson's Disease severity

Prashanth and Roy (2018) combined the 5 stages of PD as shown in figure 3 to early, moderate and late stages and further discarding 'late stage' due to class-imbalance and avoid motor complexities. This project builds deep learning models based on 3 stages as well as define them into 2 classes, 'severe' or 'non-severe' where stage II or more are classified as 'Severe'.

3.2 Data Preparation

PPMI database consists of vast repository of 113 data sets, out of which 18 data sets have been accessed which constituted assessments for motor and non-motor symptoms. Each data comprised records by unique patient ID and their visits for a duration of 5-10 years, where each visit has a dedicated Event_ID. These data sets have been merged together by unique 'PATNO' and 'Event_ID'. The final merged data comprises 9472 observations of 175 variables at different time steps. Data contains assessment records from 1473 participants.

One of the biggest challenges to work with longitudinal data is dealing with missing values. Longitudinal data have lots of missing records, as not all the patients get assessed for all the measures at the same time and very few patients participate till the last visit for the follow up visits over such a long period of time. The data sets has records from baseline visits to follow-up visits till sixteenth visit which is specified by 'EVENT ID' as can be seen in figure 4. There are random missing visits from all the data sets. Hence, after careful selection of visits with most records, these have been assigned an integer value from 1 to 9. Visits with very few records have been discarded. The plot for the number of patients for total selected visits can be seen in figure 5.

BL	PW	ST	U01	V01	V02	V03	V04	V05	V06	V08	V10	V12	V13	V14	V15	V16
1558	10	201	2	12	1029	2	1535	2	1399	1135	985	818	489	524	262	42

Figure 4: Number of records for different visits

Missing Value imputation is a common technique used in analysing data with either forward fill or backward fill to retain the amount of data (Lipton et al.; 2015; Bang et al.; 2020; Su et al.; 2021). Another method to deal with missing values is to replace them with mean of that particular column. But this is not possible as each column comprises range of records from different time periods of different participants. In this project, the missing values have been filled by using 'interpolate' function for each of the patient's records after grouping the records according to the individual patient's records and their visits (Bang et al.; 2020).

Next, plot of the target variable NHY as seen in figure 6 shows a big class imbalance at later stages of the patients. Last two stages of severe to very severe patients have negligible

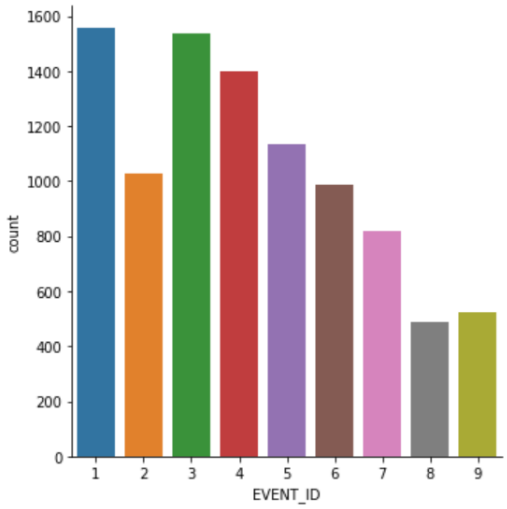

Figure 5: Participation of patients from baseline to last visit

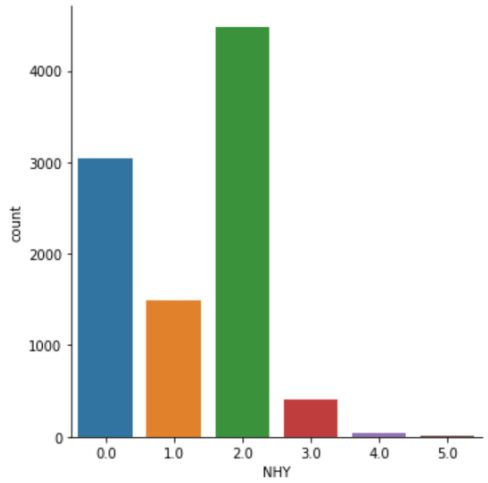

Figure 6: Severity of patients from normal to most severe

records, whereas stage 3 is moderate PD patients still needing medical attention. Hence, the last three stages were merged to get severe patients. The final data comprises four classes for further analysis (0-normal, 1-mild, 2-moderate, 3-severe) (Prashanth and Roy; 2018).

3.3 Data Standardization

Data is normalized to address the potential problem of over-fitting while training the model (Kara; 2021). Min-max standardization has been used here in this project to have all the features in a normal range from -1 to +1. Normalized value 'X' can be represented by the following equation.

$$X = (x - x_{\min})/(x_{\max} - x_{\min})(1)$$

3.4 Multi-variate Long Short-term Memory (LSTM)

LSTM is a type of Recurrent Neural Network (RNN) which deals well with RNN's vanishing gradient and overcome's the problem of exploding gradients. LSTM has found to be excellent for time-series prediction as they can learn well with long term dependencies of the data. Also, during the classification process, the interdependence of the variables in longitudinal data can be captured by LSTM units, which is almost impossible with traditional machine learning (ML) models (Su et.al, 2020).

The basic architecture of an LSTM model as illustrated in figure 7 (Kara; 2021), includes a memory cell that is capable of maintaining all the information record at each time step,and a forget gate which decides if information is irrelevant and needs to be forgotten from the previous state. Besides memory and forget, there are input and output gates.

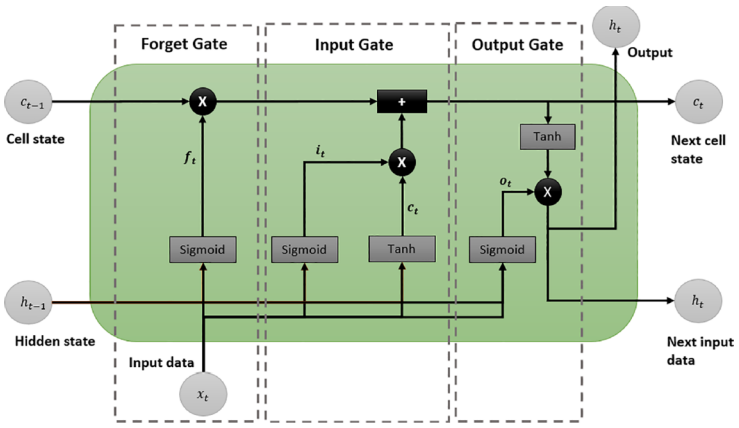

Figure 7: Architecture of an LSTM model

Computationally, the four gates can be represented for time step t as

$$i_t = \sigma_g(W_i x_t + R_i h_{t-1} + b_i) \tag{2}$$

$$f_t = \sigma_g(W_f x_t + R_f h_{t-1} + b_f) \tag{3}$$

$$o_t = \sigma_g(W_o x_t + R_o h_{t-1} + b_o) \tag{4}$$

$$m_t = \sigma_c(W_m x_t + R_m h_{t-1} + b_m) \tag{5}$$

where i, f,o and m represent the four gates of LSTM model, input, forget, output and memory cell respectively. 'W' represents the input weight and 'R' represent the recurrent weight, whereas 'b' represents the bias (Lipton et al.; 2015; Kara; 2021).

The layers of a typical LSTM model consists of an input layer where at each time step t, it has an input size of number of input nodes. Next is the LSTM layer with a number of hidden layers and the last number of nodes corresponding to output mode. The layer consists of an activation function for regular update of the cell. Next layer consists of a fully connected output layer corresponding to the number of classes with softmax as activation function and cross-entropy for calculating model loss (Ashour et al.; 2020).

3.5 Genetic Algorithm

Genetic Algorithm (GA) is an evolutionary search engine which helps to search the best fit individuals from the entire population. It is based on the evolutionary theory of

'Survival of the Fittest'. GA has been recently utilised to find the best hyper-parameters of a machine learning model by optimisation of the fitness function.The flowchart in figure 8 summarizes the basic Genetic Algorithm

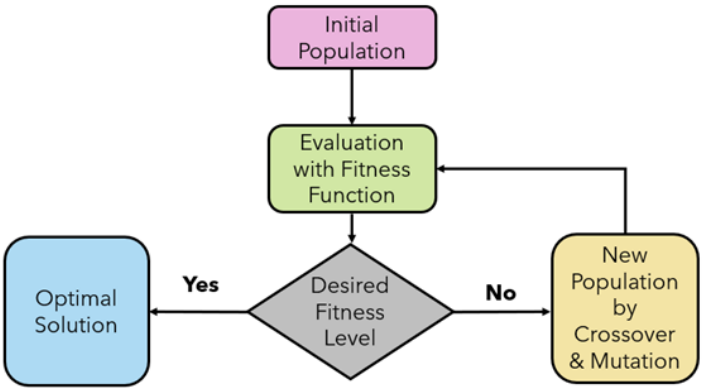

Figure 8: Basic Architecture of Genetic Algorithm

The process of Genetic Algorithm starts with random search through initial population consisting multiple chromosomes or solutions expressed in terms of 0's and 1's. These binary strings further accommodates simulating genes that is a set of variables (Chun and Shin, 2018) (Figure 9).The process starts with uniform spreading of the search space with solutions. GA then carefully selects the variables that are best fit for the applied model with cross-over and mutation (Huang and Wang, 2006). Gene associating the optimal feature is represented by '1' whereas non-contributing feature is represented by '0' for each instance (Fayyazifar and Samadiani, 2017).

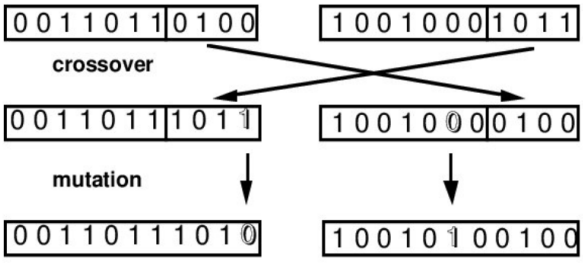

Figure 9: Crossover and Mutation in Genetic Algorithm

The cycle of Genetic Algorithm has the following phases:

1. Initializing the population with some random set of variables to start the gradient search. Higher population might lead to reach the desired solution faster, but to run this might take much longer time and can be heavy on system memory.

2. Define the Fitness Function and evaluate to get the best fitness for the model. The fitness function in this case is model 'loss' which should be minimized for best results.

3. If the desired fitness criteria is not met, the algorithm randomly selects the next set of population. Fitness function that is complex with higher population leads to higher computational costs.

4. It goes through crossover and mutation to get the new population. This process makes sure that the same solution is not repeated in every generation.

5. The algorithm checks the fitness function again and the search stops if the criteria is met with the optimal solution, else the genetic algorithm search continues for specified generations.

Genetic Algorithm has been applied in this study to get the best fit LSTM model. This hybrid model will help to get the best fitness score. The fitness function used in this project is 'loss' function. This means the GA runs through several generations to get the fittest LSTM model with best 'window size' and best 'number of units' with minimal loss.

3.6 Baseline Approaches

Apart from the novel approach, several baseline models have been implemented in this project. The models included a basic deep learning model, Multi-layer Perceptron (MLP) along with machine learning models commonly used to solve classification problem in Parkinson's disease such as Logistic Regression (LR), Support Vector Machine (SVM), Naive Bayes (NB), Gradient Boosting Classifier (XGB), Random Forest (RF) and Decision Tree (DT). The models have been used to 'serious' parkinson's disease patient with 'HY' severity score of 2 and more. These patients need special medical attention in order to slow down the disease progression.

For multi-layer Perceptron model, the training algorithm obtains the required bias vectors as well as weight matrices based on the traditional backpropagation error. However, objective function commonly used for training model now-a-days is based on Mean Square Error (MSE) score (Djerioui et al.; 2020). Hence, MSE has also been used for evaluation of model as a metric for comparison against the novel hybrid GA-LSTM model.

4 Implementation

This section explains end to end implementation of the proposed project. Data sets understanding, merging and transformation has been carried out using R programming, whereas rest of the modelling has been carried out using Python programming in an open source online Google server called Google Collaboratory (Colab). Colab is an excellent free online platform to run complicated and data heavy machine learning algorithms on GPUs and TPUs provided by Google.

4.1 Long Short-Term Memory Model (LSTM)

The normalized data was split into training data and testing data using 80/20 split. The model input for LSTM network should be in three dimensions, hence the train and test data were reshaped into samples, time-steps and features. LSTM model was then implemented on the pre-processed Parkinson's data. The model was constructed using random selection of hyperparameters of LSTM model with hidden dense layers. The LSTM model can be summarized as follows-

The LSTM network was constructed with an LSTM layer with 44 input units and two hidden dense layers of 20 units and 10 units, with 'ReLu' activation function. The output layer consisted of 4 layers with 'softmax' activation function for this multi-class classification model. Further, 10 percent 'dropout' was used to randomly drop 10% of units from the dense layer to avoid over-fitting. The model used 'adam' optimizer and 'sparse-categorical accuracy' as metrics for model evaluation. Model loss was calculated

```
Model: "sequential_3"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lstm_3 (LSTM)                (None, 44)                13904
_________________________________________________________________
dense_9 (Dense)              (None, 20)                900
_________________________________________________________________
activation_9 (Activation)    (None, 20)                0
_________________________________________________________________
dropout_3 (Dropout)          (None, 20)                0
_________________________________________________________________
dense_10 (Dense)             (None, 10)                210
_________________________________________________________________
activation_10 (Activation)   (None, 10)                0
_________________________________________________________________
dense_11 (Dense)             (None, 4)                 44
_________________________________________________________________
activation_11 (Activation)   (None, 4)                 0
=================================================================
Total params: 15,058
Trainable params: 15,058
Non-trainable params: 0
_________________________________________________________________
```

Figure 10: Summary of the Long Short-Term Memory model

using both 'sparse categorical cross-entropy' as well as 'mean-squared error' where the former gave better results with loss function and model performance.

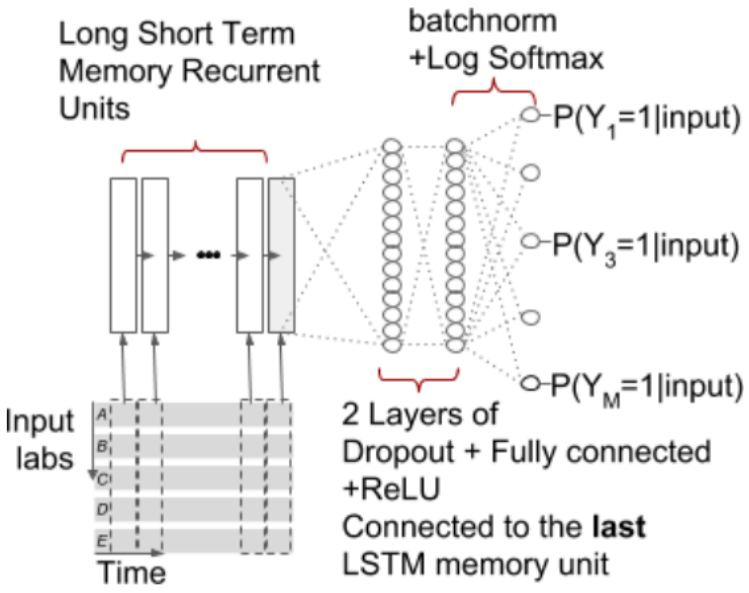

Figure 11: Long Short-Term Memory model for Parkinson's disease severity prediction

Complete graphical representation of LSTM network has been illustrated in figure 11 with an LSTM recurrent unit that represents input of the multiple time steps, multilabel data with two hidden dense layers and fully connected outer layer. The layers are connected to the LSTM memory unit to calculate the information that has to be saved by the model (Razavian et al.; 2016).

Since, selection of model parameters is such a tedious task, LSTM model was also implemented coupled with the heuristic search engine 'Genetic Algorithm' to find the

best parameters to train the model and get optimized output.

4.2 Hybrid Approach: GA optimized LSTM model

Due to vast dimensionality of longitudinal records, medical practitioners and researchers face many challenges dealing with the complexity that accompany the patient's data for diagnosis of illnesses. From previous studies, Genetic Algorithm (GA) has been found to be a recent favourite for optimisation of machine learning algorithms by reducing the feature space and errors associated with modelling (Dinov et. al, 2016; Alvarez et. al, 2019; Varrechhia et. al, 2021).

DEAP (Distributed Evolutionary Algorithm in Python) framework has been used to implement GA for optimization of LSTM model. This GA-LSTM model helps in finding the optimal window size and number of hidden units to achieve the best accuracy of the prediction model. The steps involved in the model are:

1. Reshaped the data into a single array for input.

2. Defined a function for preparing the data set. The dataset has been divided into portions of population with the pair of X and Y where X represents the past values and Y represents the future solutions at each time step.

3. Decoding the solution by Genetic algorithm for window size and number of LSTM units.The data has been split into train and test with 80/20 split and used to build the LSTM model architecture.

4. Defined and calculated the fitness function which is reduced model 'loss' in this case.

5. Ran the GA-LSTM model by defining the population size of 10, number of generations and gene length of 4. These has been randomly selected after multiple iterations and gives the best results.

6. The algorithm's search engine found the top solutions for optimal window size and number of units for LSTM model.

7. Using this best configuration, the model has been trained and compiled. Model has been fitted with a batch size of 32 for 50 epochs using 'Adam' optimizer and 'mean-squared error' as model loss.

8. The model has been evaluated and finally used to make predictions.

4.3 Multi-Layer Perceptron

Multi-layer perceptron (MLP) is the basic artificial neural network model consisting of input layer, multiple dense layers in between and an output layer. This feed-forward neural network model has been constructed using keras package based on previous research. The model consisted of 5-layers with random manual selection and tuning of its hyperparameters, such as number of units for input, hidden layer units or number of layers to achieve the best model.

10 percent Dropouts of nodes were used after each layer to prevent overfitting, hence improving model's performance. This was followed by batch normalisation for input means or variances. The Input layer and Dense layers were built with 'ReLu' activation function, whereas 'Sigmoid' activation function was used for the output layer as the output has two classes-'Severe' and 'Non-Severe'patients. The model is finally compiled using 'Adam' optimizer which is based on stochastic gradient descent and 'Binary Crossentropy' as model loss which defines probability based on two classes. Metrics used here

is 'Accuracy'. 'Early stopping' and 'Model checkpoint' have been used.

Apart from these multiple machine learning models have been implemented on the dataset for classification and evaluated the results for comparison with LSTM and MLP models.

5 Evaluation

5.1 LSTM

The LSTM model is implemented using 100 epochs and a batch size of 32. The shuffle mode has been turned off due to time series sequence of data. The model has been created using callbacks, 'Early Stopping' to monitor minimum validation loss with a patience level of 10 and 'Model Checkpoint' to monitor maximum validation sparse categorical accuracy.

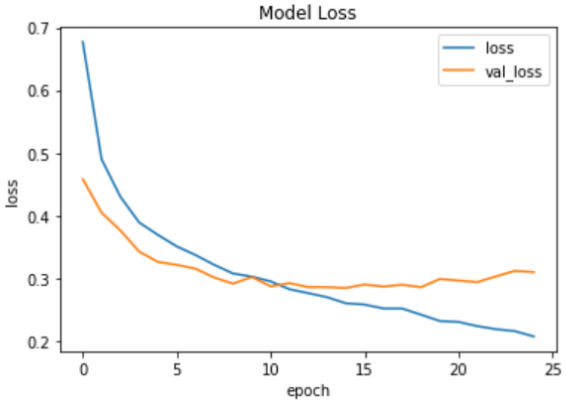

Figure 12: Model Loss Plot

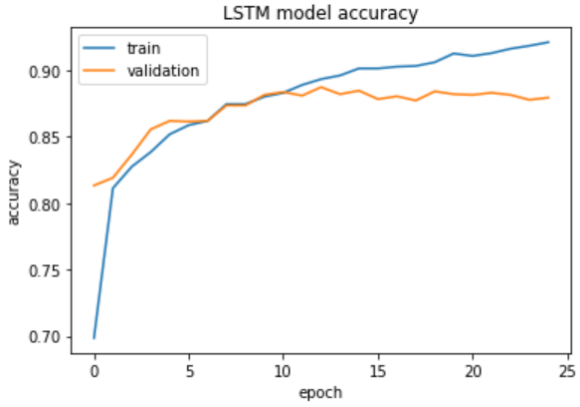

Figure 13: Model Accuracy Plot

The 'best accuracy' options have been turned on for the model. Further the model has been implemented after several iterations of model parameters to give the desired results.The best accuracy of the model is 88%. Further, the LSTM model results can be visualized by the Model loss plot and Model accuracy plot for the training data and validation data as shown in figure 12 and 13 respectively. The plots show that the data is not over-fitting as there is not much divergence between the train and validation curves in the plot . Finally, predictions were made using test data.

5.2 GA-LSTM

The Genetic Algorithm optimized Long Short-term Memory was implemented for population size of 4 for four generations. The solution with length of gene is used as 10. These parameters have been used after multiple iterations. The top solution after the evolutionary search are Window Size of 12 and Number of Units 8.

```
Top Solution for Window Size:  12 , Num of Units:  8
```

Figure 14: Top solution for Genetic Search

LSTM model has been integrated with this evolutionery method and has been recompiled with the best solutions for the parameters defined by GA and implemented. The loss function of the model 'Mean squared Errorr' has been minimized and test 'RMSE' is equal to 0.33 for the hybrid model.

5.3 Baseline Approach: Multi-Layer Perceptron

Feed forward neural network model showed only 49% accuracy with the defined architecture. The loss of the model defined as 'binary cross-entropy' has been shown in figure 15.

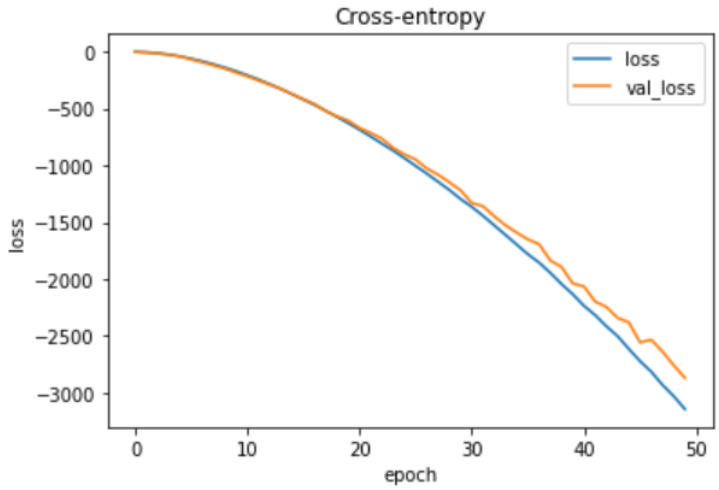

Figure 15: Model Loss for Multi-layer perceptron Network

Validation RMSE score has been calculated for comparison with GA-LSTM model and is found to be 0.72 which is significantly higher than the GA-LSTM model.

5.4 Multiple Baseline Machine Learning Algorithms

Based on literature review, multiple baseline machine learning algorithms have been built for classification of Parkinson's disease and their performance have been evaluated for comparison with the deep learning models, Multi-layer Perceptron and Long Short Term Memory Neural networks. This can be seen in figure 16

The machine learning model selected were based on the previous studies and included K-Nearest neighbours, Support Vector Machine (Linear, Polynomial, RBF), Decision

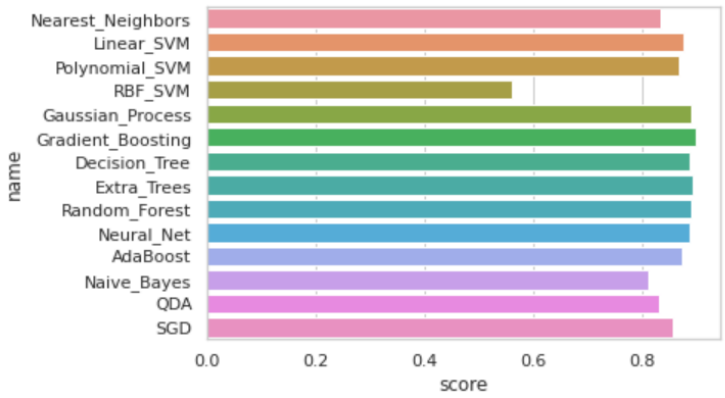

Figure 16: Plot for Comparison of Multiple Baseline Machine Learning models for PD prediction

Tree classifier, Random Forest, Gradient Boosting, Extra trees, AdaBoost, Naive Bayes, SGD classifier, Quadratic Discriminent Analysis. All of these are compared against each other for model accuracy score. Even though most of them seem to be performing well, it should be noted that these models did not consider the time step aspect of longitudinal data. Hence, the proposed approach of LSTM model that learns from the long-term memory of data dependencies should be considered for future studies.

5.5 Discussion

This research considers multiple factors while building a prediction model for Parkinson's disease. Firstly, from previous research work the importance of underlying non-motor symptoms is evident. Hence, this research combines the non-motor assessments with motor assessments to build a deep-learning model. Secondly, the longitudinal temporal aspect of clinical data for Parkinson's disease patients has been ignored by many researchers. Hence, this project addresses the issue of dealing with complexity and multi-dimensionality of a longitudinal clinical data of Parkinson's disease patients using LSTM model. Lastly, Genetic Algorithm has been integrated with LSTM model to optimize its hyperparameters and minimize the model loss 'MSE' and thereby determining the optimal neural network architecture. The proposed GA-LSTM model inculcates the non-linear propoerties represented by the superior features of Parkinson's disease temporal data. The model is successful in reducing the loss function as compared to the fully-connected multi-layer perceptron model. Several state-of-the-art methods using machine learning models have also been analyzed and compared with neural network where Gradient Boosting, Random Forest, Extra trees have all performed well along with Neural Network with about 88-89% accuracy. Whereas, they do not consider the longitudinal aspect of the data.

Even though Genetic Algorithm has optimized the LSTM model, the contraints of designing the model architecture with fewer number of generations and populations might have had an impact on model building process. Furthermore, the input for GA-LSTM model has to be a single array of population which hinders the model evaluation process at later stages. Furthermore,the model worked well for multi-class classification, with accuracy of 88%, it didn't give as good results for binary classification.

6 Conclusion and Future Work

Deep learning models, Multi-layer Perceptron along with proposed model LSTM, were built in this project for prediction of severity of Parkinson's disease. LSTM model was further optimised using Genetic Algorithm to find the best solutions for hyperparameters such as window size and number of input units to be used in LSTM network. Genetic Algorithm is implemented for minimizing the model 'loss' function giving much lower 'RMSE' than the baseline model Multi layer Perceptron. Application of the proposed model can significantly enhance the performance of disease prediction models by its optimized design and be used in healthcare industries.

Future work will intend to deal with all model complexities and work with more functionalities of Genetic Algorithm such as involving other parameters for optimising, dealing with the issue of multi-dimensionality for multi-step longitudinal data. Model can also incorporate other clinical assessments such as scans and genetic data for a deeper analysis.

7 Acknowledgements

This research work could not have been completed without constant support of my family and colleagues. My sincere gratitude to the contributors of PPMI for granting access to the database. Finally, I acknowledge the support and advice of my supervisor, Dr. Rashmi Gupta and National College of Ireland.

References

Al Mamun, A., Hoq, M., Hossain, E. and Bayindir, R. (2019). A hybrid deep learning model with evolutionary algorithm for short-term load forecasting, *2019 8th International Conference on Renewable Energy Research and Applications (ICRERA)*, IEEE, pp. 886–891.

Armañanzas, R., Bielza, C., Chaudhuri, K. R., Martinez-Martin, P. and Larrañaga, P. (2013). Unveiling relevant non-motor parkinson's disease severity symptoms using a machine learning approach, *Artificial intelligence in medicine* **58**(3): 195–202.

Ashour, A. S., El-Attar, A., Dey, N., Abd El-Kader, H. and Abd El-Naby, M. M. (2020). Long short term memory based patient-dependent model for fog detection in parkinson's disease, *Pattern recognition letters* **131**: 23–29.

Bang, S.-J., Wang, Y. and Yang, Y. (2020). Phased-lstm based predictive model for longitudinal ehr data with missing values.

Chung, H. and Shin, K.-s. (2018). Genetic algorithm-optimized long short-term memory network for stock market prediction, *Sustainability* **10**(10): 3765.

Djerioui, M., Brik, Y., Ladjal, M. and Attallah, B. (2020). Heart disease prediction using mlp and lstm models, *2020 International Conference on Electrical Engineering (ICEE)*, IEEE, pp. 1–5.

Evers, L. J., Krijthe, J. H., Meinders, M. J., Bloem, B. R. and Heskes, T. M. (2019). Measuring parkinson's disease over time: the real-world within-subject reliability of the mds-updrs, *Movement Disorders* **34**(10): 1480–1487.

Fitzmaurice, G., Davidian, M., Verbeke, G. and Molenberghs, G. (2008). *Longitudinal data analysis*, CRC press.

Garcia, T. P. and Marder, K. (2017). Statistical approaches to longitudinal data analysis in neurodegenerative diseases: Huntington's disease as a model, *Current neurology and neuroscience reports* **17**(2): 14.

Gorgolis, N., Hatzilygeroudis, I., Istenes, Z. and Gyenne, L.-G. (2019). Hyperparameter optimization of lstm network models through genetic algorithm, *2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA)*, IEEE, pp. 1–4.

Grover, S., Bhartia, S., Yadav, A., Seeja, K. et al. (2018). Predicting severity of parkinson's disease using deep learning, *Procedia computer science* **132**: 1788–1794.

Hoehn, M. and Yahr, M. (1967). Parkinsonism: onset, progression, and mortality., v. 17, n.

Holden, S. K., Finseth, T., Sillau, S. H. and Berman, B. D. (2018). Progression of mds-updrs scores over five years in de novo parkinson disease from the parkinson's progression markers initiative cohort, *Movement disorders clinical practice* **5**(1): 47–53.

Huang, C.-L. and Wang, C.-J. (2006). A ga-based feature selection and parameters optimizationfor support vector machines, *Expert Systems with applications* **31**(2): 231–240.

Kara, A. (2021). Multi-step influenza outbreak forecasting using deep lstm network and genetic algorithm, *Expert Systems with Applications* **180**: 115153.

Kilicarslan, S., Celik, M. and SAHIN, Ş. (2021). Hybrid models based on genetic algorithm and deep learning algorithms for nutritional anemia disease classification, *Biomedical Signal Processing and Control* **63**: 102231.

Lei, B., Yang, M., Yang, P., Zhou, F., Hou, W., Zou, W., Li, X., Wang, T., Xiao, X. and Wang, S. (2020). Deep and joint learning of longitudinal data for alzheimer's disease prediction, *Pattern Recognition* **102**: 107247.

Lei, H., Huang, Z., Han, T., Luo, Q., Cai, Y., Liu, G. and Lei, B. (2018). Joint regression and classification via relational regularization for parkinson's disease diagnosis, *Technology and Health Care* **26**(S1): 19–30.

Lipton, Z. C., Kale, D. C., Elkan, C. and Wetzel, R. (2015). Learning to diagnose with lstm recurrent neural networks, *arXiv preprint arXiv:1511.03677* .

Nejedly, P., Plesinger, F., Viscor, I., Halamek, J. and Jurak, P. (2019). Prediction of sepsis using lstm neural network with hyperparameter optimization with a genetic algorithm, *2019 Computing in Cardiology (CinC)*, IEEE, pp. Page–1.

Opara, J., Małecki, A., Małecka, E. and Socha, T. (2017). Motor assessment in parkinsons disease, *Ann Agric Environ Med* **24**(3): 411–415.

Prashanth, R. and Roy, S. D. (2018). Novel and improved stage estimation in parkinson's disease using clinical scales and machine learning, *Neurocomputing* **305**: 78–103.

Rashid, T. A., Hassan, M. K., Mohammadi, M. and Fraser, K. (2019). Improvement of variant adaptable lstm trained with metaheuristic algorithms for healthcare analysis, *Advanced Classification Techniques for Healthcare Analysis*, IGI Global, pp. 111–131.

Razavian, N., Marcus, J. and Sontag, D. (2016). Multi-task prediction of disease onsets from longitudinal laboratory tests, *Machine learning for healthcare conference*, PMLR, pp. 73–100.

Ren, J., Hua, P., Li, Y., Pan, C., Yan, L., Yu, C., Zhang, L., Xu, P., Zhang, M. and Liu, W. (2020). Comparison of three motor subtype classifications in de novo parkinson's disease patients, *Frontiers in Neurology* **11**: 1803.

Rodriguez-Blazquez, C., Schrag, A., Rizos, A., Chaudhuri, K. R., Martinez-Martin, P. and Weintraub, D. (2021). Prevalence of non-motor symptoms and non-motor fluctuations in parkinson's disease using the mds-nms, *Movement Disorders Clinical Practice* **8**(2): 231–239.

Salmanpour, M. R., Shamsaei, M., Saberi, A., Klyuzhin, I. S., Tang, J., Sossi, V. and Rahmim, A. (2020). Machine learning methods for optimal prediction of motor outcome in parkinson's disease, *Physica Medica* **69**: 233–240.

Soumaya, Z., Taoufiq, B. D., Benayad, N., Yunus, K. and Abdelkrim, A. (2021). The detection of parkinson disease using the genetic algorithm and svm classifier, *Applied Acoustics* **171**: 107528.

Su, D., Zhang, X., He, K. and Chen, Y. (2021). Use of machine learning approach to predict depression in the elderly in china: A longitudinal study, *Journal of Affective Disorders* **282**: 289–298.

Tsiouris, K. M., Rigas, G., Antonini, A., Gatsios, D., Konitsiotis, S., Koutsouris, D. D. and Fotiadis, D. I. (2017). Mining motor symptoms updrs data of parkinson's disease patients for the development of hoehn and yahr estimation decision support system, *2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, IEEE, pp. 445–448.

Varrecchia, T., Castiglia, S. F., Ranavolo, A., Conte, C., Tatarelli, A., Coppola, G., Di Lorenzo, C., Draicchio, F., Pierelli, F. and Serrao, M. (2021). An artificial neural network approach to detect presence and severity of parkinson's disease via gait parameters, *Plos one* **16**(2): e0244396.

Veeraragavan, S., Gouwanda, D., Ahmad, S. A. et al. (2020). Parkinson's disease diagnosis and severity assessment using ground reaction forces and neural networks, *Frontiers in physiology* **11**: 1409.

Wutzl, B., Leibnitz, K., Rattay, F., Kronbichler, M., Murata, M. and Golaszewski, S. M. (2019). Genetic algorithms for feature selection when classifying severe chronic disorders of consciousness, *PloS one* **14**(7): e0219683.

Zhang, X., Chou, J., Liang, J., Xiao, C., Zhao, Y., Sarva, H., Henchcliffe, C. and Wang, F. (2019). Data-driven subtyping of parkinson's disease using longitudinal clinical records: a cohort study, *Scientific reports* **9**(1): 1–12.

Zhao, A., Qi, L., Li, J., Dong, J. and Yu, H. (2018). A hybrid spatio-temporal model for detection and severity rating of parkinson's disease from gait data, *Neurocomputing* **315**: 1–8.

Zhao, J., Feng, Q., Wu, P., Lupu, R. A., Wilke, R. A., Wells, Q. S., Denny, J. C. and Wei, W.-Q. (2019). Learning from longitudinal data in electronic health record and genetic data to improve cardiovascular event prediction, *Scientific reports* **9**(1): 1–10.