EAI Endorsed Transactions on Pervasive Health and Technology Research Article **EAI.EU**

An LSTM based DNN Model for Neurological Disease Prediction Using Voice Characteristics

Anila M[1,*], G Kiran Kumar[2], D Malathi Rani[3], M V V Prasad Kantipudi[4] and D Jayaram[5]

[1,2]Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Gandipet, Hyderabad, Telangana State, India

[3]Department of Electronics and Communications Engineering, Marri Laxman Reddy Institute of Technology and Management, Dundigal, Hyderabad, Telangana State, India

[4]Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune 412115, Maharashtra, India

[5]Department of Information Technology, Chaitanya Bharathi Institute of Technology (A), Gandipet, Hyderabad, Telangana State, India

Abstract

INTRODUCTION: A neurological condition known as Parkinson's disease (PD); it affected millions of individuals worldwide. An early diagnosis can help enhance the quality of life for those who are affected with this disease. This paper presents a novel Deep neural network model based on Long Short-Term Memory (LSTM) design for the identification of PD using voice features.
OBJECTIVES: This research work aims to Identify the presence of PD using voice features of individuals. To achieve this, a Deep neural Network with LSTM is to be designed. Objective of the work is to analyse the voice data and implement the model with good accuracy.
METHODS: The proposed model is a Deep Neural Network with LSTM.
RESULTS: The proposed method uses the features gleaned from voice signals for training phase of LSTM model which achieved an accuracy of 89.23%, precision value as 0.898, F1-score of 0.965, and recall value as 0.931and is observed as best when compared to existing models.
CONCLUSION: Deep Neural Networks are more powerful than ANNs ahd when associated with LSTM , the model outperformed the job of identifying PD using voice data.

Keywords: Voice features, Deep Neural Network, LSTM, Parkinson's Disease, Machine Learning (ML)

Received on 06 December 2023, accepted on 08 March 2024, published on 14 March 2024

doi: 10.4108/eetpht.10.5424

*Corresponding author. Email: anilarao.m@gmail.com

1. Introduction

Worldwide, many people are afflicted by the neurological condition, called Parkinson's disease [1]. It is observed with biomarkers like with tremors, stiffness, and reduced motor skills that largely affects the central nervous system. Early detection of PD can enhance the patient's quality of life by enabling early intervention and treatment [2]. However, traditional methods for detecting Parkinson's disease are subjective and depend on the experience of the clinician. Therefore, there is a need for automated and objective methods that can accurately detect PD.

Many Machine Learning and Deep Learning models developed by various researchers are observed as automated methods that help detecting Parkinson's disease [3]. LSTM neural networks is observed with great potential for processing time-series data and have been used in different applications.

The proposed technique in this paper utilizes the features extracted from speech signals to train an LSTM model that accurately classifies the samples. The approach is of three

main phases: pre-processing of data, feature extraction, and training an LSTM neural network, and testing the trained model on a dataset containing speech signals recorded from patients affected with the disease and healthy controls.
Overall, the proposed approach presented in this paper has the potential to be used as a non-invasive and automated tool for the detection of PD using speech characteristics. The rest of the paper is organised as: Related research on detection of Parkinson's disease is included in Section II. The pre-processing methods, feature extraction, LSTM neural network architecture, and evaluation measures are all covered in detail in Section III of the proposal. Section IV reviews the assessment of the proposed technique and offers various findings. The paper is concluded with Section V, which explores the possible uses of the proposed strategy in the early identification of Parkinson's disease.

2. Literature Survey

People with neurological diseases like PD also suffer from Cognitive Impairment (CI). The cognitive impairment causes speech disfluencies and errors. The author of this study [4] examines how CI affects the changes in read patterns brought on by errors and disfluencies. They begin by examining mistakes and disfluencies in hand transcripts and annotating them with information about their location and type. Then, they use Automatic Speech Recognition (ASR) to create transcripts while simultaneously detecting the same faults and disfluencies in manual transcripts without comments. They discovered that using a regression framework, these traits can be utilised to forecast MoCA Scores. The mistake and disfluency patterns may vary, but they predict that these aspects will be vulnerable to cognitive impairment in adults without PD.
Ghosh, H. et al. (2023) Demonstrates [5] the use of convolutional neural networks for potato leaf disease prediction, contributing to agricultural technology and crop health. They draw the conclusion that even though much research employs auditory characteristics like jitter and shimmer, phonatory techniques have not yet demonstrated their value in the diagnosis of neurological diseases. Therefore, it is impossible to tell individuals with neurodegenerative illnesses apart from controls using these traits. Phonatory features, such noise, can be used to distinguish between the two, according to several recent studies; nonetheless, they are well suited for automated detection and assessment methods than for usage as standalone biomarkers [6]. With regard to articulatory methods, it is crucial to keep in mind that every single patient is making an effort to speak English.
Two classifier designs are used in this work [7]: the traditional pipeline technique and an end-to-end approach. In the former approach, they use glottal structures to extract the articulatory and phonatory features. They employ the quasi-closed phase with iterative adaptive inverse filtering glottal inverse filtering techniques in this instance. The end-to-end method makes use of DL models that are trained on speech waveforms and voice source waveforms [8-9]. Since voice problems influence the vocal folds, glottal source signals are used as a raw waveform in addition to using the speech waveform. The accuracy of the SVM-based pipeline method was improved from 65% to 67% when the baseline features were merged. Additionally, they concluded that their model may be expanded to also predict the neurological condition of PD patients.
The performance of methods for detecting neurological diseases of the system degrades for any mismatch found between the signals, which is mostly brought about by signal deterioration in test signals. They try to control the quality of the voice signals (recordings) using two different methods with identification of short and long-term degradations and any violations in protocol. Using the mPower PD data set and the evaluation of their hypotheses is done. To optimise vocal features and increase the efficiency of these detecting systems, they primarily seek to enhance or regulate the voice signal quality.
A complete automated classification algorithm is used with total of 5 Convolution Neural Networks are used for training phase and recorded with accuracy of 85%. The EEG artifact rejection is observed to be not improved but resulted in faster training [10]. An incremental SVM algorithm is used for the prediction of target labels of dataset. As part of pre-processing, self-organising maps are used to get the clustering task done [11]. Error values of MAE=0.4656 and MAE=0.4967 are recorded for total and motor UPDRS.Another study [12] used a centre of pressure to study the posture stability thereby by observe the balance over performing the dual task and cognitive tasks. These findings imply that patients with PD prioritise postural control above other concurrent tasks, at least in the early stages of the disease, just like healthy participants do.

3. Proposed Methodology

3.1 Proposed algorithm for detection of PD using voice features

1. Perform Pre-processing on the voice dataset in two parts.
    a. Selection of relevant attributes using principal component analysis.
    b. Reshape the data to fit into the LSTM model as it requires a 3D tensor input.
2. Split the dataset into train set and test set with a ratio of 80:20.
3. Train the neural network with ReLU at the hidden layer and sigmoid function at the output layer.
4. Evaluate the model using test set.
5. Generate the confusion matrix and evaluate the other metrics such as precision, recall, and F1 score and accuracy.
6. Model can be used to predict the output for any new sample.

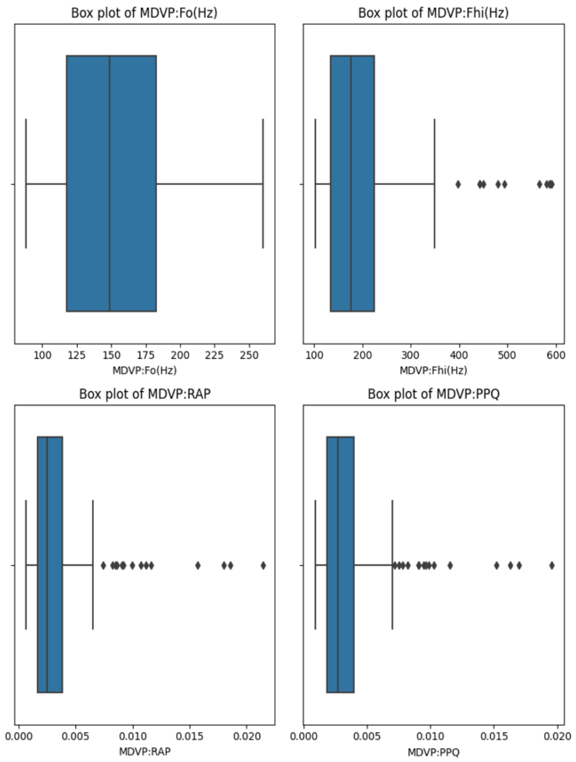

Figure 2: Sample Boxplots

Since the model is sequential, its layers are stacked in a linear fashion. A Rectified Linear Unit activation function with an LSTM layer with 256 units make up the top layer. In this situation, there are 22 features because that is the number of features that were extracted, and the input shape is (1, X_train.shape[2]), which indicates that the input is a 3D tensor with one time step.

A dropout layer with a rate of 0.2 makes up the second layer. By asynchronously eliminating 20% of the units in the LSTM layer during training, this helps avoid overfitting.

The following three layers, which each have 128, 64, and 48 units, are dense lay-ers that are fully connected. Each of these layers adds non-linearity to the model using a ReLU activation function. The binary classification is carried out at out-put layer that has single unit and a Sigmoid function.

Figure 1 represent the flow of steps involved in the proposed methodology.

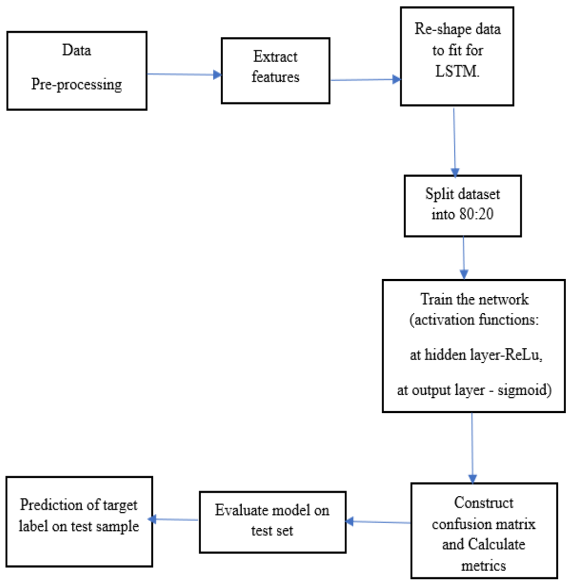

Figure 1: Proposed flow diagram

3.2 Dataset Description and Visualization:

3.2.1 Description of Dataset

Data is collected from the UCI Parkinson's Dataset, which included biomedical voice characteristics from both PD affected patients and healthy controls. The recordings included sustained vowel phonation with corresponding phonation features from 195 individuals.

Recordings were taken using a specialized software called "Vocal Loading Test" and the phonation features were extracted using the PRAAT software. The dataset consists of 24 features, including measurements of frequency, shimmer, jitter, and noise ratio. Two of the features, patient name, and status, were later ignored in the analysis.

Each record in the dataset has a label of either '1' indicating "Parkinson's" or '0' indicating "healthy control." This labelling helped identifying PD affected and not affected based on voice features. The UCI Parkinson's Dataset is an essential resource for researchers and clinicians working towards improving the analysis and behaviour of those affected.

3.2.2 Dataset Visualization:

Data visualization is an essential tool for communicating complex information and insights in a clear and concise manner. It allows us to represent data visually through graphs, scatter plots, and other types of visual aids, thereby understand patterns and relationships in the data. So, data visualization is a critical tool for anyone working with data.

It is important to have the knowledge of our data, and one of the majorly used visualization is a scatter plot, from which we can observe the patterns of our da-ta. Box plots do present the statistics about the data like mean, median and quar-tiles. Few box plots are shown in the figure2 which are plotted taking the features of dataset.

3.3 Activation Functions used in the Methodology

3.3.1. ReLU (Rectified Linear Unit) Activation Function:

ReLU activation function is a non-linear function that returns the input value if it is positive and returns 0 if it is negative. This means that the output of the function is always non-negative, which can help prevent the vanishing gradient problem that can occur in deep neural networks. The vanishing gradient problem happens when the gradients of the cost function become very small, making the learning process very slow.

The ReLU function is also computationally efficient, as it is simple to calculate and has a fast convergence rate. Additionally, it has been shown to perform well in various types of neural network architectures and in different types of problems.

In this specific code, ReLU activation function is used in the first LSTM layer and in the three dense layers. The first LSTM layer has 256 units, while the dense layers have 128, 64, and 48 units. The purpose of these layers is to extract useful features from the input data and to transform the extracted features into meaningful output. Using ReLU activation function in these layers helps in creating non-linearity in the model, which can lead to better performance in complex datasets.

3.3.2. Sigmoid Activation Function:

Sigmoid activation function is a non-linear function that maps any input value to a value between 0 and 1. It is commonly used in binary classification problems where the output should be a probability between 0 and 1. In this case, the output of the model should be a probability of a person having Parkinson's disease.

The Sigmoid function has the property that its output is always between 0 and 1, which makes it useful for converting any input value into a probability. This probability value can then be used to make predictions about whether the input belongs to a particular class or not. Here, Sigmoid activation function is used in the last dense layer, which has only one unit. The output of the model is passed through this layer to obtain a probability value between 0 and 1, which is then compared to a threshold value (0.5) to make the final prediction about whether the input belongs to a particular class or not.

3.4 Confusion Matrix

To measure the performance of any binary classifier, initially true positive, true negative, false positive and false negative count will be generated. Using the above four values, a matrix of the below format can be represented as in Table 1.

Table 1: Confusion Matrix

	Predicted Positive	**Predicted Negative**
Actual Positive	TP count	FN count
Actual Negative	FP count	TN count

Here,

- TP count represents the count of correctly classified positive samples.
- FP count represents count of incorrectly classified positive samples.
- TN count represents the count of correctly classified negative samples.
- FN count represents the count of incorrectly predicted negative samples.

The purpose of confusion matrix is to find the performance of the model by calculating measures like accuracy, precision, recall, and F1-score as presented their formulas below. For example, if the model is predicting false positives, it may be overly sensitive to certain features in the data, while if it is predicting false negatives, it may be missing important features that are indicative of the positive class.
The following metrics used to evaluate the model and are given as below from (1) to (4):

$$Accuracy = \frac{\textit{True positives} + \textit{True negatives}}{\textit{All samples}} \quad (1)$$

$$Precision = \frac{\textit{True positives}}{\textit{True Positives} + \textit{False negatives}} \quad (2)$$

$$Recall = \frac{\textit{True positives}}{\textit{True positives} + \textit{False negatives}} \quad (3)$$

$$F1 - score = \frac{Precision\ .Recall}{Precision + Recall} \quad (4)$$

Confusion matrix helps to provide insights and limitations regarding the model’s performance. Any machine learning model can be evaluated using the confusion matrix, providing insight into the model's strengths and weaknesses, and guiding improvements in the model's architecture and training process.

4. Results and Discussions

The outcomes show the performance metrics used to assess how well the LSTM model identified Parkinson's illness. The summary of the model can be visualized via figure 3.A confusion matrix is generated with the true positive, true negative, false positive and false negative rates. In addition to these, the model is evaluated with metrics like precision, F1 score, recall and accuracy. The proposed model received scores of 89.23% for maximum accuracy, 0.898 for Precision, 0.965 for Recall, and 0.931 for F1 Score. The LSTM-DNN model is the compared to

```
Layer (type)                 Output Shape              Param #
=================================================================
lstm_5 (LSTM)                (None, 256)               285696

dropout_5 (Dropout)          (None, 256)               0

dense_20 (Dense)             (None, 128)               32896

dense_21 (Dense)             (None, 64)                8256

dense_22 (Dense)             (None, 48)                3120

dense_23 (Dense)             (None, 1)                 49

=================================================================
Total params: 330,017
Trainable params: 330,017
Non-trainable params: 0
```

Figure 3: Model summary

the other classification models and yield the results shown by figures 4 to figure 7. The LSTM-DNN model outperformed all the other classification models with the highest accuracy of them all.

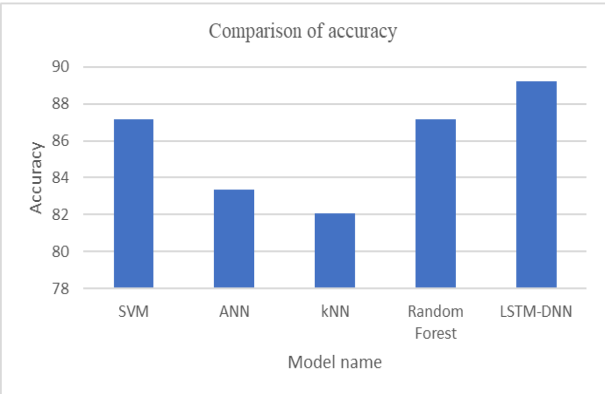

Figure 4: Comparison of accuracy value of proposed model with other models

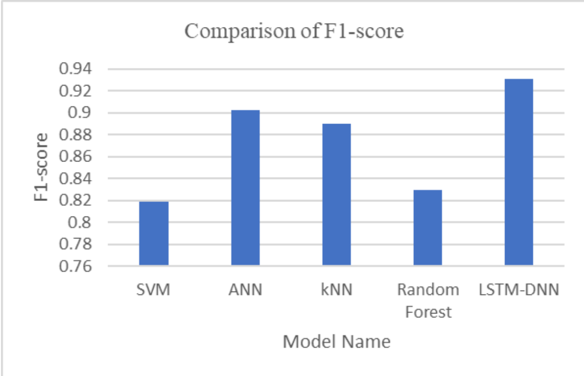

Figure 5: Comparison of F1-score value of proposed model with other models

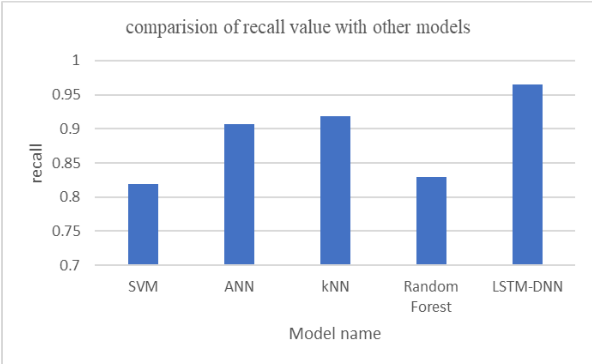

Figure 6: Comparison of recall value of proposed model with other models

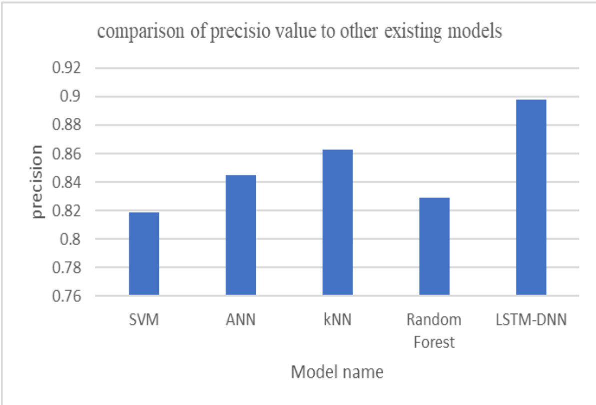

Figure 7: Comparison of precision of proposed model with other models

Table 2: Result Analysis

Method Name	Accuracy (%)	Precision	Recall	F1-Score
SVM	87.17	0.819	0.819	0.819
ANN	83.33	0.845	0.907	0.902
kNN	82.05	0.863	0.919	0.890
Random Forest	87.17	0.829	0.829	0.829
LSTM-DNN	89.23	0.898	0.965	0.931

From the above Table 2, it can be observed that the proposed model achieved better accuracy when compared to existing models.

5. Conclusion and Future Scope

A promising method for the early and precise identification of Parkinson's disease (PD) utilising speech features is the Parkinson's disease detection system employing LSTM. With great accuracy and precision, the LSTM model can identify data in two categories, either affected with PD or healthy control by learning complicated patterns in the voice signal. The use of the UCI Parkinson's dataset provides a standardized and reliable dataset for training and evaluating the LSTM model. The dataset contains many voice data entries from individuals with and without Parkinson's disease, allowing the model to learn well. In order to assess the performance of the model, various metrics like accuracy, precision, recall, and F1 score are used which could assist pinpoint areas that need development. The model's accuracy rating of 89.23%, which is extremely desirable, shows that the model may be employed for accurate and timely PD detection. The optimization of the model architecture and hyperparameters, as well as the use of regularization techniques and early stopping, helps prevent overfitting. Overall, the Parkinson's disease detection system based on LSTM is an extremely effective method for detecting PD early and precisely utilising speech recordings. Through facilitating early detection and prompt intervention, it has the potential to enhance the lives of people with Parkinson's disease (PD) through improving disease management and quality of life. The proposed method is LSTM-based and observed with divergence, it can be enhanced by inclusion of more hidden layers and hyperparameter tuning with other values.

In addition to voice recordings, other modalities such as gait analysis and hand-writing analysis can also provide valuable information for PD detection. Integrating multiple modalities could enhance the accuracy and consistency of the PD detection system. The LSTM model can be integrated into wearable devices such as smartwatches or earbuds for real-time monitoring of PD symptoms. This would enable early detection of symptom onset and could improve the management of PD. Large-scale studies involving diverse populations could further validate the effectiveness of the LSTM model for PD detection. Such studies would also enable the identification of additional features or modalities that also im-prove the accuracy of PD detection system. Also, LSTM model could be fine-tuned using data from related diseases such as Alzheimer's or Huntington's dis-ease. This would enable the development of a more generalized disease detection system that could be used for multiple neurodegenerative diseases.

References

[1] Laureano Moro-Velazquez, Jorge A. Gomez-Garcia, Julian D. Arias-Londoño, Najim Dehak, Juan I. Godino-Llorente, "Advances in Parkinson's Disease detection and assessment using voice and speech: A review of the articulatory and phonatory aspects," 2021, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2021.102418

[2] Amir Hossein Poorjam, Mathew Shaji Kavalekalam, Liming Shi, Yordan P. Raykov, Jesper Rindom Jensen, Max A. Little, Mads Græsbøll Christensen, "Automatic quality control and enhancement for voice-based remote Parkinson's disease detection," 2019, arxiv, https://doi.org/10.48550/arXiv.1905.11785

[3] Amrit Romana, John Bandon, Matthew Perez, Stephanie Gutierrez, Richard Richter, Angela Roberts, Emily Mower Provost, "Automatically Detecting Errors and Disfluencies in Read Speech to Predict Cognitive Impairment in People with Parkinson's Disease," 2021, ISCA, https://doi.org/10.21437/Interspeech.2021-1694

[4] O. Karaman, H. Çakın, A. Alhudhaif, K. Polat, "Robust automated Parkinson disease detection based on voice signals with transfer learning," Expert Syst. Appl. 178 (2021), 115013, https://doi.org/10.1016/j.eswa.2021.115013

[5] Pei-Fang Guo, Prabir Bhattacharya and Nawwaf Kharma "Advances in Detecting Parkinson's Disease" 2010, Springer.

[6] Ali H. Al-Fatlawi, Mohammed H. Jabardi, Sai Ho Ling, "Efficient Diagnosis System for Parkinson's Disease Using Deep Belief Network" 2016, Elsevier https://doi.org/10.1109/CEC.2016.7743941

[7] Mehrbakhsh Nilashi, Othman Ibrahim, Hossein Ahmadi, Leila Shahmoradi, Mohammadreza Farahmand "A hybrid intelligent system for the prediction of Parkinson's Disease progression using machine learning techniques" 2017, Elsevier.

[8] Salim Lahmiri, Debra Ann Dawson, Amir Shmuel "Performance of machine learning methods in diagnosing Parkinson's disease based on dysphonia measures" 2017, Springer.

[9] Postuma, R. & Montplaisir, J. Predicting Parkinson's disease-why, when, and how? Parkinsonism & related disorders 15, S105–S109 (2009).

[10] Ishihara, L., and Brayne, C., A systematic review of depression and mental illness preceding Parkinson's disease. Acta Neurol. Scand. 113(4)211–220, 2006. doi:10.1111/j.1600-0404.2006.00579.x.

[11] B. E. Sakar, M. E. Isenkul, C. O. Sakar, A. Sertbas, F.Gurgen, S. Delil, H. Apaydin, O. Kursun, "Collection and Analysis of a Parkinson Speech Dataset with Multiple Types of Sound Recordings", IEEE Journal of Biomedical and Health Informatics, vol. 17, no. 4, pp. 828-834, July 2013.https://doi.org/10.1109/JBHI.2013.2245674

[12] Benba A., Jilbab A. Hammouch A., "Hybridization of best acoustic cues for detecting persons with Parkinson's disease," 2014 Second World Conference on Complex Systems (WCCS), Agadir, 2014, pp. 622-625.