http://dx.doi.org/10.35596/1729-7648-2023-21-3-102-110

Original paper

UDC 615.47:616-072.7

IT DIAGNOSTICS OF PARKINSON'S DISEASE BASED ON THE ANALYSIS OF VOICE MARKERS AND MACHINE LEARNING

ULADZIMIR A. VISHNIAKOU, YIWEI XIA

Belarusian State University of Informatics and Radioelectronics (Minsk, Republic of Belarus)

Submitted 26.01.2023

Abstract. The results of studying the parameters of the spectra of speech signals by machine learning with the use of neural networks are presented. This study was carried out in order to confirm experimentally the possibility of performing an assessment of these parameters for the detection of Parkinson's disease in the early stages (IT diagnostics). During the study, the public database was used, which systematized the spectra of vowel sounds uttered by patients with Parkinson's disease. The applied method is binary data classification. In the course of the study, the speech data spectrum was first preprocessed, which consisted of filtering it in order to remove its noise components and eliminate bursts and gaps in it. Then the parameters of the processed spectrum of speech data were determined: average value, maximum and minimum, peak, wavelet coefficients, MFCC and TQWT. After that, the object was selected using the PCA algorithm. The model was trained using the Knn and Random Forest algorithms, as well as the Bayesian neural network. The Bayesian optimization algorithm and the GridSearch method were used to find the best model hyperparameters. It has been established that when using Knn, Random Forest and Bayesian neural network, it is possible to increase the accuracy of recognition of Parkinson's disease by 94.7; 88.16 and 74.74 %, respectively. A similar study by other scientists showed that the recognition accuracy of data sets was only 86 %.

Keywords: machine learning, neural networks, binary classification, hyperparameter optimization, decision tree, sensitivity, precision.

Conflict of interests. The authors declare no conflict of interests.

For citation. Vishniakou U. A., YiWei Xia (2023) IT Diagnostics of Parkinson's Disease Based on the Analysis of Voice Markers and Machine Learning. *Doklady BGUIR*. 21 (3), 102–110. http://dx.doi.org/10.35596/1729-7648-2023-21-3-102-110.

IT-ДИАГНОСТИКА БОЛЕЗНИ ПАРКИНСОНА НА ОСНОВЕ АНАЛИЗА ГОЛОСОВЫХ МАРКЕРОВ И МАШИННОГО ОБУЧЕНИЯ

В. А. ВИШНЯКОВ, ИВЭЙ СЯ

Белорусский государственный университет информатики и радиоэлектроники (г. Минск, Республика Беларусь)

Поступила в редакцию 26.01.2023

Аннотация. Представлены результаты исследования параметров спектров речевых сигналов с помощью машинного обучения с применением нейронных сетей, проведенного в целях экспериментального подтверждения возможности выполнения оценки этих параметров для выявления болезни Паркинсона на ранних стадиях (IT-диагностика). В ходе исследования использовали общедоступную базу данных, в которой систематизированы спектры гласных звуков, произнесенных пациентами с болезнью Паркинсона. Примененный метод – бинарная классификация данных. Сначала выполняли предварительную обработку спектра речевых данных, состоявшую в его фильтрации, для удаления из него шумов и устранения присутствующих в нем всплесков и пробелов. Затем определяли параметры обработанного спектра речевых данных: среднее значение, максимум, минимум, пик, вейвлет-коэффициенты, MFCC и TQWT. После этого выбирали объект с помощью алгоритма PCA. Для обучения модели использовали алгоритмы Knn и Random Forest и нейронной сети Байеса. Для нахождения наилучших гиперпараметров модели применяли алгоритм оптимизации Байеса и метод GridSearch. Установлено, что при использовании Knn, Random Forest и нейронной сети Байеса можно обеспечить увеличение точности распознавания болезни Паркинсона на 94,7; 88,16 и 74,74 % соответственно. Аналогичное исследование, проведенное другими учеными, показало, что точность распознавания наборов данных составила всего 86 %.

Ключевые слова: машинное обучение, нейронные сети, бинарная классификация, гиперпараметрическая оптимизация, дерево решений, чувствительность, точность.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Для цитирования. Вишняков, В. А. IT-диагностика болезни Паркинсона на основе анализа голосовых маркеров и машинного обучения / В. А. Вишняков, Ивэй Ся // Доклады БГУИР. 2023. Т. 21, № 3. С. 102–110. http://dx.doi.org/10.35596/1729-7648-2023-21-3-102-110.

Introduction

Parkinson's disease (PD) is a common neurodegenerative disease of the elderly. Medical research [1] suggests that the pathology of Parkinson's disease has two main aspects: depigmentation of the substantia nigra due to degeneration of neurons containing melanin and dopamine; and the formation of Lewy bodies in the substantia nigra and other brain areas, such as the nucleus acumens and parts of the cortex.

Lewy bodies are considered as biological marker of Parkinson's disease and thought to be responsible for the loss of dopaminergic neurons. Studies have shown that Lewy bodies spread through nervous system in a specific pattern of dissemination, H. Braak et al. [2] divided this pattern into six stages, with olfactory and speech disturbances occurring in the first two stages. According to the study, 89 % of Parkinson's disease patients have varying degrees of speech impairment. Studying the symptoms of speech disorders in Parkinson's disease patients is an excellent direction to achieve IT diagnosis disease in its early stages.

Review

The investigation is done to classify Parkinson diseased and healthy people by extracting fourteen phonological features and twelve cepstral features of speech using Upadhya methodology [3]. Fourteen phonological features including five jitter variants, six shimmer variants, two harmonic features and the mean autocorrelation of the fundamental frequency feature. The final classification can be based on machine learning technology using a neural network approach. Parkinson disease detection using pattern recognition method has been presented in literature by Putri [4]. It extracted 22 speech features and 12 EMG features and used artificial neural networks as a classification method.

In this article we used machine learning and neural network to identify Parkinson's disease with K-nearest neighbor (Knn), Random Forest and Bayesian algorithms. Based on machine learning, neural network algorithms can be divided into the following four categories: supervised, unsupervised, semi-supervised and reinforcement. The general process of machine learning algorithms can be divided into 5 steps, as shown in Fig. 1.

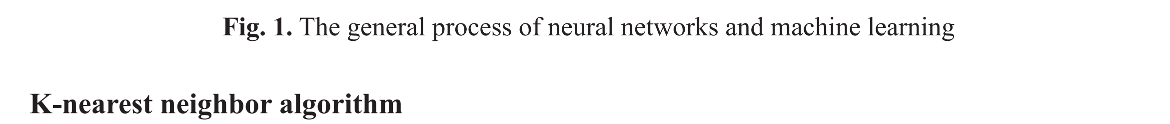


Fig. 1. The general process of neural networks and machine learning

K-nearest neighbor algorithm

The K-nearest neighbor [5] algorithm is a simple machine learning algorithm. It has two basic parameters:

– distance calculation methods. Common methods for calculation of distance between two points are Manhattan distance, Euclidean distance and Minkowski distance;

– the k value. Find the k nearest neighbors.

The implementation process of the Knn algorithm contains the following steps:

1) inputting training datasets and testing datasets;

2) determining the parameter of k;

3) calculating the distance of the current point against the given training data (distance between all training data points and the current point, current point is from testing datasets). Euclidean distance can be used here;

4) ranking the calculated distances in ascending order;

5) selecting k points with the smallest distance from the current point (k points are from the training datasets);

6) The label with the highest frequency among the k points is returned as the predicted label for the current point.

The advantage of the Knn algorithm is that it is computationally simple, but the disadvantage is that it is computationally intensive and has a large memory overhead. A suitable value of k needs to be chosen.

Random Forest algorithm

Random Forest algorithm [6] is a decision tree model based on the bagging framework, it consists of following steps:

1) if N is the number of samples in the training datasets and a constant $n < N$ is specified, then n samples are randomly selected from the training set with a put-back, selected n samples are used as the training datasets, samples not selected are used as the testing sets to evaluate the error;

2) if the number of features in the training set is M, then m features ($m << M$) are randomly selected from the selected n samples to build the decision tree;

3) the above steps are repeated k times, resulting in k decision trees;

4) after generating k decision trees, for each testing data, the category with the highest number of classification results among the k decision trees is used as the classification result of the whole Random Forest.

Fig. 2 shows the flow of the Random Forest classification algorithm.

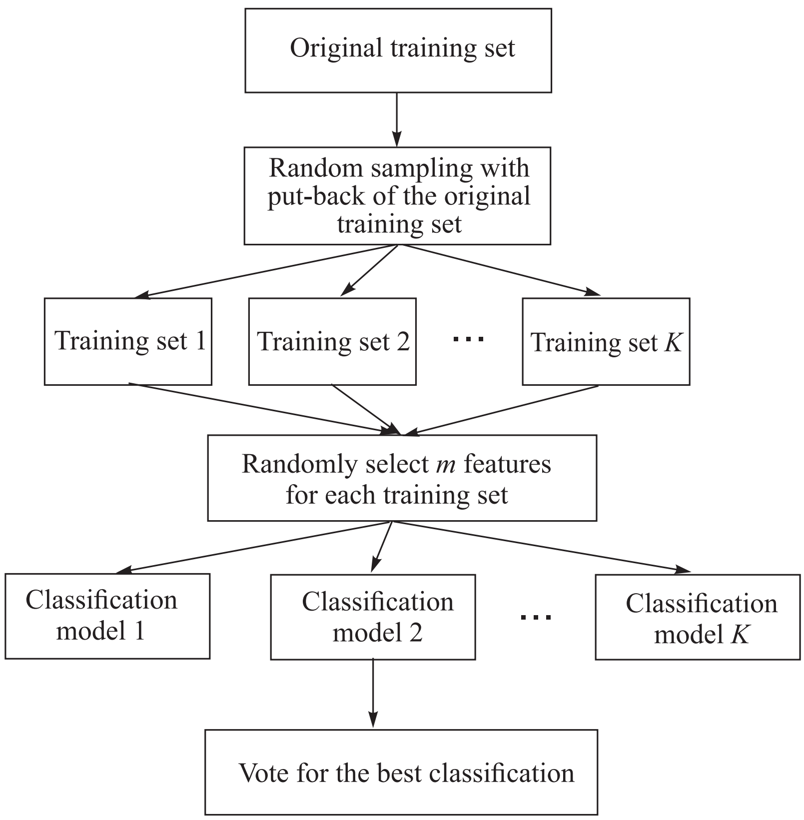


Fig. 2. The flow of the Random Forest classification algorithm

The hyperparameters of the Random Forest classification algorithm are as follows:
– the number of decision trees;
– the number of features in each decision tree;
– whether to adopt a random sampling with put-back;
– the maximum depth of the tree, beyond which branches will be pruned.

The advantage of Random Forest algorithm is that it can be trained in parallel and can handle high-dimensional data and unbalanced data very well, but the disadvantage is that it is not good at handling small data or low-dimensional data.

Bayesian neural network algorithm

Bayesian neural networks (BNN) [7] have the same network structure as fully connected neural networks. But BNN combined probabilistic modelling with neural networks. Bayesian neural networks differ from normal neural networks in that the weight parameters are random variables rather than definite values. The Fig. 3 shows the comparison of parameters between Deep neural networks (DNN) and BNN.

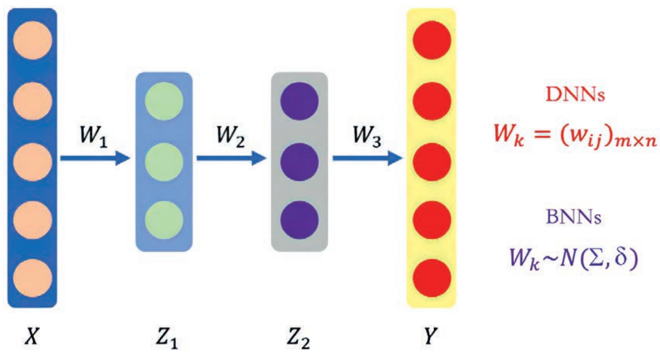


Fig. 3. The comparison of parameters between Deep neural networks and Bayesian neural networks

The hyperparameters of the BNN are as follows:
– the number of hidden layers;
– regularization parameters. Regularization coefficients affect the generalization ability of the model;
– the number of neurons per layer;
– the number of epoch. An epoch means that all the data is fed into the network, completing a forward computation plus backward propagation process;
– learning rate. Learning rate schedule for weight updates;
– the activation function of the hidden layer. Activation functions such as logistic regression function, tanh function, relu function, etc.

The steps of BNN algorithm are as follows:
1) initializing the training model parameters;
2) inputting training datasets into the model and perform a forward propagation to obtain the prediction results;
3) calculating the error between the prediction result and the real result;
4) the error is back-propagated once and the model parameters are optimised using an optimisation algorithm;
5) repeating the training for *k* epochs;
6) outputing the model parameters.

Methodology

The public datasets [8] used in this article were gathered from 188 patients with PD (107 men and 81 women) with ages ranging from 33 to 87 at the Department of Neurology in Cerrahpa Faculty of Medicine, Istanbul University. The control group consisted of 64 healthy individuals (23 males and 41 females), aged between 41 and 82 years old. During data collection, the microphone was set to 44.1 KHz and then repeated three times to collect sustained pronunciation of the vowel “a” from each participant.

After data cleaning and pretreatment of the datasets, for each speech record, we extracted 21 baseline features (5 jitter variants features [9], 6 shimmer variants features [9], 5 fundamental frequency parameters features [10], 2 harmonic parameters [11], 1 recurrence period density entropy feature [11], 1 detrended fluctuation analysis feature [11] and 1 pitch period entropy feature [11]), 11 temporal frequency features (3 intensity parameters features, 4 formant frequencies features and 4 bandwidth features), 84 Mel frequency cepstral coefficients features, 182 Wavelet transform based features [12], 22 vocal fold features [13] (3 glottis quotient features, 6 glottal to noise excitation features, 7 vocal fold excitation ratio and 6 empirical mode decomposition features), 432 tunable Q-factor wavelet transform features [14]. The data features set was then normalized, divided into training datasets and testing datasets in the ratio of 9:1. The training datasets were trained and tested using 5-fold cross validation with 5 times' repetition. Testing datasets were used to test the final results. Since the data were labeled data, the problems were supervised learning of classification problems. In this article, we used the Knn algorithm, Random Forest and the Bayesian neural network algorithm. The Fig. 4 shows the flow chart of training datasets and testing datasets.

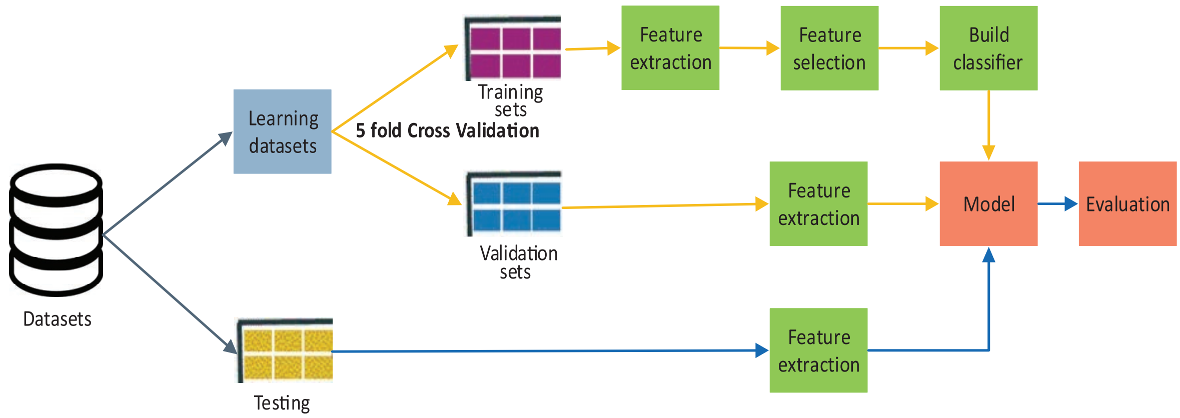


Fig. 4. The flow chart of training datasets and testing datasets

Experiments and results

Python is an interpreted, object-oriented, dynamically data-typed high-level programming language. Because it is simple and has a large number of useful libraries. So, we used Python to implement machine learning algorithms and neural network algorithms. Here were the Python libraries used in our experiments:

– Numpy library was used to load and store audio data;
– Matplotlib library was used to draw confusion matrices and ROC plots;
– Sklearn library was used to implement machine learning algorithms;
– Keras library was used to implement BNN models;
– Pandas library was used to read csv files.

In this experiment, we used the confusion matrix [15] to evaluate the model. Here were the experiments we conducted. First of all, we applied the Bayesian optimizer to optimize the hyperparameters of Knn. Its acquisition function was an expected improvement per second. Then the number of calculation iterations was 30 times. The minimal classification error plot on speech features using Knn algorithm was shown in the Fig. 5.

From Fig. 5, we can see that error reached a minimum value at the 29th iteration, so we took the KNN parameter value of the 29th iteration as the KNN parameter value for this experiment. We used the PCA (Principal components analysis) method to select 10 features [16]. The number of neighbours k was 2. The distance measurement was Cosine function, distance weighting was the inverse distance weighting.

The recognition results of the Parkinson's disease speech datasets based on the Knn algorithm were as follows. The confusion matrix and Receiver Operating Characteristic (ROC) [17] plots of the experimental results for the training datasets were shown in Fig. 6. The confusion matrix and ROC curve plots of the experiment results for the testing datasets were shown in Fig. 7. The experimental results of Parkinson's recognition based on Knn algorithm on the testing datasets were shown in Tab. 1.

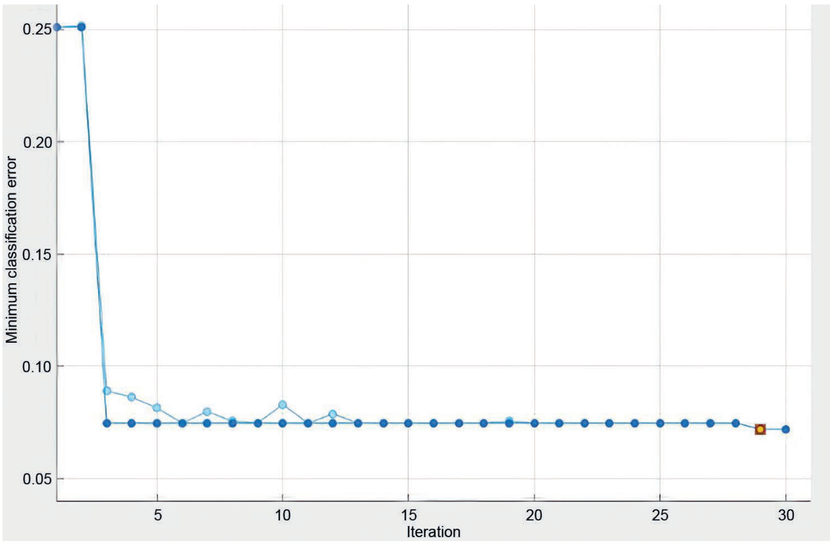
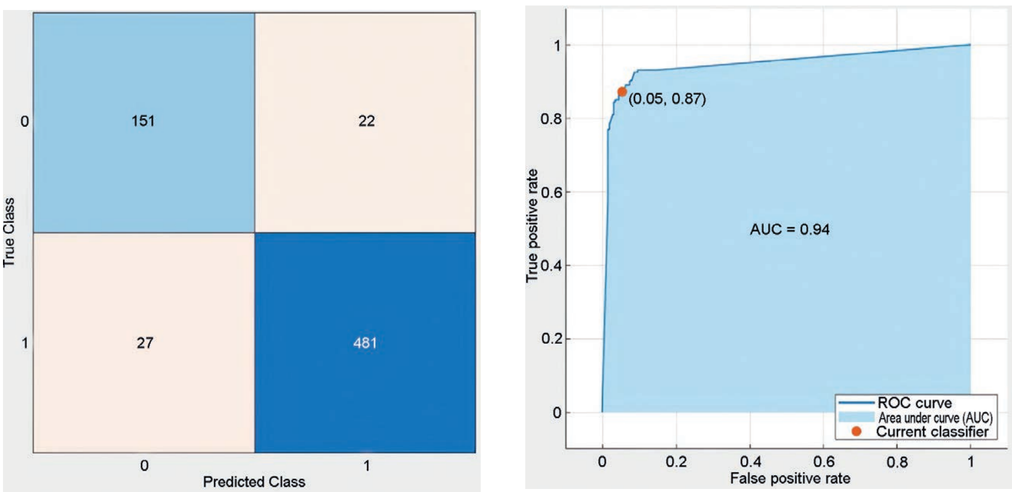
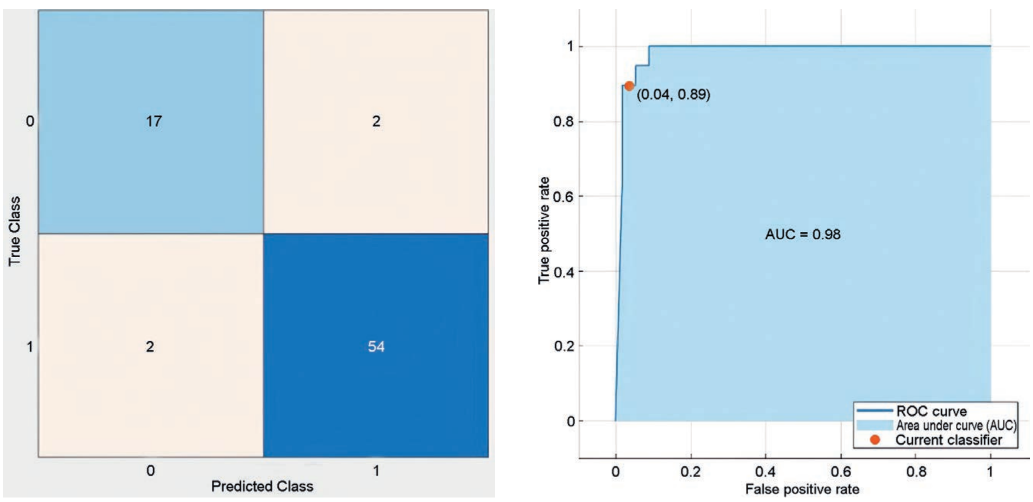


Fig. 5. The minimal classification error plot on speech features using Knn algorithm



a *b*

Fig. 6. Results of training datasets:
a – confusion matrix of training datasets; *b* – Receiver Operating Characteristic plots of training datasets



a *b*

Fig. 7. Results of testing datasets:
a – confusion matrix of testing datasets; *b* – Receiver Operating Characteristic plots of testing datasets

Table 1. Experimental results of Parkinson's recognition based on Knn algorithm

Dataset	Average value, %			Test accuracy, %
	Precision	Sensitivity	F1 score	
Pd_speech	92.95	92.95	92.95	94.7

In summary, the accuracy of the training datasets was 92.8 % and the accuracy of the testing datasets was 94.7 %. The accuracy of the testing sets was 1.9 % higher than that of the training sets. This means that the data sets were very small. When the datasets were divided into training and testing dataset, the data from different categories are unevenly distributed, resulting in the accuracy of the testing datasets being greater than the accuracy of the training dataset.

In the Random Forest algorithm, the GridSearchCV method was used to select the appropriate hyperparameters. The number of decision trees in this experiment was searched between 1 and 100 and the best result was 81; the number of features on each decision tree was searched between 1 and 100 and the best result was 23. The Gini coefficient was chosen as a measure of impurity, the maximum depth of the tree was not restricted. Random sampling with put-back was used. The experimental results of Parkinson's recognition based on Random Forest algorithm on the testing datasets were shown in Tab. 2.

Table 2. Experimental results of Parkinson's recognition based on Random Forest algorithm

Dataset	Average value, %			Test accuracy, %
	Precision	Sensitivity	F1 score	
Pd_speech	76.91	89.31	82.65	88.16

From this experiment it could be seen that, the test accuracy of the recognition results of the speech datasets of Parkinson's disease based on Random Forest algorithm was only 88.16 %. They were lower than the model based on the Knn algorithm. Although the Random Forest algorithm outperformed the Knn algorithm in most cases. However, when the amount of data was particularly small, the Knn algorithm still outperformed the Random Forest algorithm in classification. We used a BNN model, which utilized 2 hidden layers with first layer of size 10 and second layer of size 10, and a ReLU function as the activation function. The experimental results of Parkinson's recognition based on BNN algorithm on the testing datasets were shown in Tab. 3.

Table 3. Experimental results of Parkinson's recognition based on Bayesian neural network algorithm

Dataset	Average value, %			Test accuracy, %
	Precision	Sensitivity	F1 score	
Pd_speech	64.15	57.14	60.44	74.74

In summary, experiments indicated that the BNN reached 74.74 % recognition accuracy on the testing datasets. The other index parameters were very low and the recognition effect was not good, which means that the BNN algorithm was not suitable for the model of small data. The neural network algorithm required larger data for better recognition. We compared the results of IT diagnosis of Parkinson's disease with the known results of other researchers. The Tab. 4 compares the test accuracy of this study and existing studies on the same dataset of Pd.

Table 4. Compare with others' research

Dataset	Researcher	Research method	Test accuracy, %
Pd_speech	C. O. Sakar [8]	SVM (RBF)	86
		SVM (Linear)	83
		Multilayer perceptron	84
	Authors	Knn	94
		Random Forest	88
		Bayesian neural network	74

The Tab. 4 showed that the Knn algorithm has the highest test accuracy with 94 % on the same dataset. This indicates that Knn algorithm has better performance on small data sets.

Conclusion

1.The article presented speech recognition systems, which are most relevant in Parkinson's disease, discussed their results. Using public datasets, the authors extracted basic characteristics, time frequency characteristics, characteristics of low-frequency cepstral coefficients, characteristics based on the wavelet transform, vocal fold characteristics and configurable characteristics of the Q-factor wavelet transform. Datasets of patients were used in a machine learning model, the model was then trained using the Knn algorithm, Random Forest algorithm, Bayesian neural network algorithm.

2. Based on the speech data of patients, IT diagnostics of Parkinson's disease results were: the recognition accuracy for the Knn algorithm reached 94.7 %, the Random Forest algorithm – 88.16 %, the Bayesian neural network algorithm – 74.74 %. The authors results of IT diagnostics of Parkinson's disease were compared with the known results of other researchers, the best of which are 86 %.

3. The results of the experiment showed that the early detection of Parkinson's disease based on the speech data was effective. But it was also possible to see from the experiment that the data set had a very important impact on the results of the experiment. A more balanced dataset with more data could provide better recognition results. On small data sets, the Knn algorithm surpassed the Random Forest and the Bayesian neural network algorithm.

References

1. Davie C. A. (2008) A Review of Parkinson's Disease. *British Medical Bull*etin. 86 (1), 109–127.
2. Braak H., Ghebremedhin E., Rüb U., Bratzke H., Del Tredici K. (2004) Stages in the Development of Parkinson's Disease-Related Pathology. *Cell and Tissue Research*. 318 (1), 121–134.
3. Upadhya S. S., Cheeran A. N. (2018) Discriminating Parkinson and Healthy People Using Phonation and Cepstral Features of Speech. *Computer Science*. 143, 197–202.
4. Putri F., Caesarendra W., Pamanasari E. D., Ariyanto M., Setiawan J. D. (2018) Parkinson Disease Detection Based on Voice and EMG Pattern Classification Method for Indonesian Case Study. *Journal of Energy, Mechanical, Material, and Manufacturing Engineering*. 3 (2), 87–98.
5. Zhang M. L., Zhou Z. H. (2007) ML-KNN: a Lazy Learning Approach to Multi-Label Learning. *Pattern Recognition*. 40 (7), 2038–2048.
6. Palimkar P., Shaw R. N., Ghosh A. (2022) Machine Learning Technique to Prognosis Diabetes Disease: Random Forest Classifier Approach. *Advanced Computing and Intelligent Technologies.* Springer, Singapore. 219–244.
7. Kareem S. W, Alyousuf F. Q., Ahmad K., Hawezi R., Awla H. (2022) Structure Learning of Bayesian Network: a Review. *Qalaai Zanist Journal*. 7 (1), 956–975.
8. Sakar C. O., Serbes G., Gunduz A., Tunc H. C., Nizam H. (2019) A Comparative Analysis of Speech Signal Processing Algorithms for Parkinson's Disease Classification and the Use of the Tunable Q-Factor Wavelet Transform. *Applied Soft Computing*. 74 (4), 255–263.
9. Upadhya S. S., Cheeran A. N., Nirmal J. H. (2017) Statistical Comparison of Jitter and Shimmer Voice Features for Healthy and Parkinson Affected Persons. *Electrical, Computer and Communication Technologies. IEE*E. 1–6.
10. Holmes J. R., Oates M. J., Phyland J. D., Hughes J. A. (2000) Voice Characteristics in the Progression of Parkinson's Disease. *International Journal of Language & Communication Disorders*. 35 (3), 407–418.
11. Abdurrahman G., Sintawati M. (2020) Implementation of Xgboost for Classification of Parkinson's Disease. *Journal of Physics*. 1538 (1), 1–12.
12. Tsanas A., Little M., McSharry P. E., Ramig L. O. (2010) New Nonlinear Markers and Insights into Speech Signal Degradation for Effective Tracking of Parkinson's Disease Symptom Severity. *NOLTA*, *Poland*. 457–460.
13. Erdogdu S. B., Serbes G., Sakar C. O. (2017) Analyzing the Effectiveness of Vocal Features in Early Telediagnosis of Parkinson's Disease. *PloS One*. 12 (8), 1–18.
14. Selesnick I. W. (2011) Wavelet Transform with Tunable Q-Factor. *IEEE Transactions on Signal Processing.* 59 (8), 3560–3575.
15. Visa S., Ramsay B., Ralescu A. L., Esther van der Knaap (2011) Confusion Matrix-Based Feature Selection. *Artificial Intelligence and Cognitive Science*. 710 (1), 120–127.

16. Maćkiewicz A., Ratajczak W. (1993) Principal Components Analysis (PCA). *Computers & Geosciences*. 19 (3), 303–342.
17. Fawcett T. (2006) An Introduction to ROC Analysis. *Pattern Recognition Letters*. 27 (8), 861–874.

Authors' contribution

The authors contributed equally to the writing of the article.

Сведения об авторах

Вишняков В. А., д. т. н., профессор кафедры инфокоммуникационных технологий Белорусского государственного университета информатики и радиоэлектроники

Ивэй Ся, аспирант кафедры инфокоммуникационных технологий Белорусского государственного университета информатики и радиоэлектроники

Адрес для корреспонденции

220013, Республика Беларусь,
г. Минск, ул. П. Бровки, 6
Белорусский государственный университет
информатики и радиоэлектроники
Тел.: +375 44 486-71-82
E-mail: vish@bsuir.by
Вишняков Владимир Анатольевич

Information about the authors

Vishniakou U. A., Dr. of Sci. (Tech.), Professor at the Department of Infocommunication Technologies of the Belarusian State University of Informatics and Radioelectronics

YiWei Xia, Postgraduate at the Department of Infocommunication Technologies of the Belarusian State University of Informatics and Radioelectronics

Address for correspondence

220013, Republic of Belarus,
Minsk, P. Brovki St., 6
Belarusian State University
of Informatics and Radioelectronics
Tel.: +375 44 486-71-82
E-mail: vish@bsuir.by
Vishniakou Uladzimir Anatolievich