International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Uncertainty Based Chaotic Pigeon Inspired Optimized Feature Selection for Effective Dyslexia Prediction using Density Peak Clustering

Vinodh. M. R[1], P. J. Arul Leena Rose[2]

Submitted: 12/09/2023 **Revised:** 28/10/2023 **Accepted:** 15/11/2023

Abstract: The role of cognitive computing in the domain of medical monitoring substantially aids the specialist in detecting concealed disorders at an initial phase. Dyslexia is not a disorder but a long-term problem affecting children's learning and reading abilities. The indicators used for dyslexic detection are vague and uncertain due to their poor quality. Hence, it is essential to examine the importance of features based on the amount of information each attribute can contribute. While all the attributes are involved in prediction of dyslexia, it affects the performance of the prediction model due to high dependency among attributes known as redundancy. Most of the existing feature selection models does not focus on handing uncertainty among attributes. This paper focuses to handle uncertainty by accomplishing clustering of instances and discovering the most relevant features in each cluster class for effective prediction of dyslexic among children. Based on the density of the nearest neighboring instances, the centroids are selected and clustered by developing a density peak clustering algorithm. The features which are highly relevant to predict the dyslexia presence or absence is determined by inducing a nature-inspired algorithm known as Pigeon Inspired Optimization (PIO). Unlike conventional PIO, which selects the population in a random manner, this work utilizes chaotic mapping-based PIO is designed to optimize the feature selection more efficiently for dyslexia prediction. The simulation results validated using support vector machine proved that the proposed Density Peak Clustering based pigeon inspired optimization (DPC-PIO) produced a higher accuracy rate compared to other states of arts algorithms for feature selection in dyslexia prediction.

Keywords: Dyslexia, Uncertainty, Density Peek Clustering, Chaotic Mapping, Pigeon Inspired Optimization.

1. Introduction

An impairment in reading and spelling skills, when compared to normal children's intellectual potential in schooling, is a significant cognitive deficit that affects the normal way of life of school children [1]. The presence of dyslexia is determined by requiring children suspected of having dyslexia to complete a non-writing graphical test. The specialist assigns a test rating based on the child's performance from the ensuing score, dyslexia or other abnormalities is diagnosed [2]. However, when specialists assess scores, they may encounter confusion on certain criteria, which is the biggest issue in the field of cognitive computing.

Prevalence of dyslexia in India is estimated to be 15% and it is also most commonly agreed as one of the hereditary components [3]. Despite the fact that majority of children in such families do not develop dyslexia, but child with dyslexic are poor in reading and speaking. Families in India place a high value on academic performance and the importance of education. The necessity of focusing on dyslexia diagnosis in children at its initial stage is very essential. When dyslexia are overlooked, it can lead to a stressful life for children. Without understanding the impact of dyslexia, parents and teachers may believe that the child is not putting forth sufficient effort to succeed in their studies, which can negatively impact the child's mental health [4]. The educational environment has indeed been greatly influenced by empowerment in the area of software applications and digital technology [5]. In recent years, researchers have begun to focus on machine learning technologies that could help dyslexic students improve their academic performance [6].

The voluminous and the low quality of the dataset are the primary and major issues in accurate detection of dyslexia. This research work focuses on developing an uncertainty-based feature selection algorithm which influence the performance of the dyslexia prediction model to improve the accuracy rate and reduce the misclassification rate.

The section 2 discusses about the existing literatures on feature selection models in dyslexia, the proposed methodology of uncertainty-based feature selection in dyslexia dataset is explained in section 3. The results and discussions are elaborately presented in the section 4 and the final section is the conclusion about this research work.

[1]Research Scholar,Department of Computer Science, SRM Institute of Science & Technology, Kattankulathur, Chennai, Tamil Nadu, India
[*]Corresponding Author Email: vinodhmr116@gmail.com

[2]Professor, Department of Computer Science, SRM Institute of Science & Technology, Kattankulathur, Chennai, Tamilnadu, India

2. Related Work

Gilles Richard et al [7] in their work assed the presence of dyslexic/dysgraphic using machine learning and artificial intelligence. By examining the children audio recording and text pictures discrimination of normal and dyslexic child are analyzed.

The classification model works on basic features such as audio and picture files.

Jothi et al [8] developed a statistical model by analysing eye movement to discover the dyslexic and non-dyslexic individuals. Eye tracker is used to examine eye movements with attributes like saccades, fixations and transient. To predict the dyslexia, particle swarm optimization-based hybrid kernel is used and the features are extracted using principal component analysis.

Shamsuddin et al [9] designed a correlation-based feature selection to detect dyslexia at its early stage. The correlation among the features in the dataset is computed and sorted accordingly. The classification model to validate the feature subset is done by decision table, simple logistics and bayes net.

Appadurai et al [10] detected the dyslexia severity by screening individuals brain image. The machine learning models is used for predicting the dyslexia in distributed network. The support vector machine is used as the classification of dyslexic and non-dyslexic individuals.

Opeyemi et al [11] developed a residue number system with deep learning approach is used to detect the secure dyslexia biomarkers. The pixel bitstream encoder is used to encrypt the MRI brain image by constructing cascaded DNN in dyslexia prediction.

Li Mun et al [12] presented a fuzzy inference system model is used to screen the presence of dyslexia in a rapid way. The two risk factors considered are low and high fuzzy system with the rule-based classifier. The weka tool is used to predict the performance of FIS using naïve bayes, random forest and decision table.

Tamboer et al [13] Presumed five behavioural aspects, including dyslexia assessment by structural model factors for investigating dyslexia reality with the use of linear regression Spelling, pronunciation, temporary memory entire reading ability, and visual recognition disturbance were among the five criteria considered. Z-score normalization was used to bring the values of the components involved in dyslexia prediction within that category.

2.1 Problem Definition

The raw dyslexia data is generally comprised of redundant and inappropriate attributes. The main challenge in analysing the pattern of dyslexic a non-dyslexia child depends on the attributes value and if they are irrelevant and redundant it is very challenging and impacts the performance of the prediction model. Thus, it is essential to develop an optimized feature selection model whose main principle is to eradicate the redundant and inappropriate attributes to sustain only the significant attributes. The feature selection improves the accuracy rate of classification model in dyslexia detection by involving only the potential sub feature set instead of using whole feature set. Though, the feature selection is attracted by many researchers and several literatures are in existence, presence of uncertainty in selection of significant features is not well treated by them. Thus, this research work aims to develop an uncertainty based optimized feature selection model for enhancing the dyslexia detection.

2.2 Importance of Feature Selection

The task of selecting the smallest subset of features from a large dataset is known as features selection. It is critical to compute the feature score depending on the quantity of information that it can provide clear perspective to many challenges in constructing optimized prediction models. The presence of noise in input or class labels results in uncertainty. When all the features are used in prediction process, then it will affect the accuracy of the model. Thus, it is important to select the features that are related with the dyslexia detection and avoid features which affect the results of the prediction model due to high dependency among features and irrelevancy with class labels. Though there are many feature selection algorithms are in existence, the optimized feature selection is the main objective of this research work to handle the uncertainty in selecting best subset of features in dyslexia dataset.

In recent research works, the metaheuristic models are deployed as optimizer to improve the process of feature selection. Hence this research work adapted a pigeon inspired metaheuristic algorithm which has the ability to reach the best solutions in feature selection, which improves the efficiency of dyslexia prediction.

3. Methodology: Uncertainty Quantified Density Peak Clustering based Pigeon Inspired Optimization Feature Selection Model for Effective Dyslexia Prediction

In this research work dyslexia dataset is collected from Keel Repository [17] with 12 medical features as indicators related to the skills of writing, reading and calculus understanding rate. The test was conducted on 1065 school children between the age of 6 and 8.

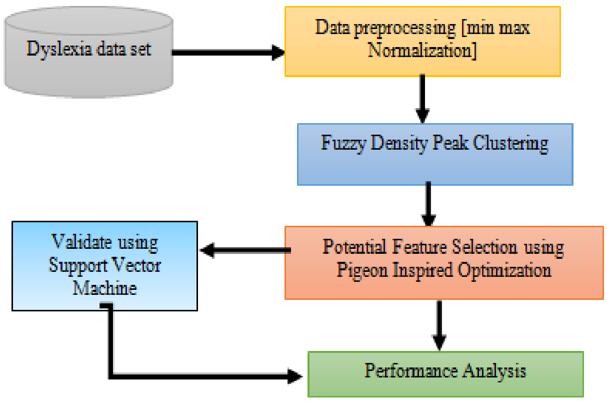

Fig. 1. Overall Flow of the Uncertainty Quantified Density Peak Clustering based Pigeon Inspired Optimization Feature Selection Model for Effective Dyslexia Prediction

The dyslexia comprised of vague input and output values. Initially, the collected dataset is of different range of attribute values, it cannot be used as such, so normalization is applied for converting them to a common range. The instances are clustered using density peek clustering model and the significant attributes which influence the prediction model is identified using Pigeon Inspired Optimization to overcome the problem of impreciseness and uncertainty. The Support vector machine is used for classification of presence or absence of dyslexia as a validation process. The figure 1 depicts the overall framework of the DPC-PIO based uncertainty handling significant feature selection in Dyslexia dataset.

3.1 Dataset Preprocessing

The dyslexia dataset comprised of vague information and each attribute vary in their range of values. While performing feature selection, the values of attributes should not influence the process of prediction, Hence, it is important to improve the quality of the dataset by applying normalization on each attribute to convert them to common range of values i.e 0 to 1. To achieve this min-max normalization is followed. It is mathematically denoted as

$$NM(f) = \frac{f - Min(f_{1..n})}{Max(f_{1..n}) - Min(f_{1..n})}$$

Where f refers to an attribute value, min and max are the minimum and maximum value of the attribute set of values.

4. Density Peek Clustering

Following the normalization phase, the similarity between the dyslexia dataset instances is determined using a novel clustering approach termed as Density Peek Clustering (DPC). The DPC is built on the concept that cluster centroids exhibit higher densities than surrounding regions and also that the distance between cluster centers is relatively large [14]. The three main factors of the DPC are

Determines the local density βj of an instance j

Discovers minimal distance among instance j and other instances with a higher density λj

Calculate the product of two parameters υj = βj * λj

The parameters βj and λj belongs to two different hypothesis of the algorithm DPC. The first assumption is based on the instance treated as cluster centers which have the higher density than the surrounding areas. The second assumption is that, the cluster centers has greater distance form the instance belongs to other clusters than the instances within its cluster.

The detailed description of the Density peak algorithm is mathematically represented in details

Consider the dyslexia dataset is denoted as Z = {z1, z2, z3,..zn} with n number of records. Each instance zi has p attributes. Hence, zij is the jth attribute of instance zi, the distance among the instances are computed using Ruzicka similarity measure (RSM) instead of Euclidean measure. The distance among two instances zl = {z11,z12, z13… z1p } and zm ={z11,z12, z13… zmp }

$$\text{RSM}\ (z_l, z_m) = \frac{\sum_{i=} \min\ (z_{li}, z_{mi})}{\sum_{i=} \max\ (z_{li}, z_{mi})}$$

Where z_{li}, z_{mi} refers to ith attribute of two different instances zl and zm respectively.

The βj which is the local density of an instance zi is formulated as

$$\beta\text{j} = \sum_{j \neq l} \mu\big(RSM(z_l - z_j) - r_c\big)$$

$$\mu(Z) = \begin{cases} 1, & z < 0 \\ 0, & z > 0 \end{cases}$$

Where rc refers to cutoff distance of the cluster,βj refers to nearest number of instances to the zl. The minimum distance among zl and any other instances Zj' with a higher density βj′ it is formulated as

$$\lambda_j = \begin{cases} \min\limits_{k:\beta'_j > \beta_j} \big(RSM_{jk}\big), if\ \exists j^j \beta'_j > \beta_j \\ \max\limits_{j'}\big(RSM_{jj'}\big), otherwise \end{cases}$$

βj and λj are computed for each instance Zjused for finding cluster centers

Algorithm

Input: Dyslexia Dataset Z = {z1, z2, z3,..zn}

Output: Clustered instances

Procedure

Begin

1. Normalize dataset Z using Min-Max using equation (1) to convert the range of values [0,1]
2. Compute the distance between instances as matrix using the Ruzicka similarity measure (RSM) formula

$$\text{RSM}\ (z_l, z_m) = \frac{\sum_{i=} \min\ (z_{li}, z_{mi})}{\sum_{i=} \max\ (z_{li}, z_{mi})}$$

3. Calculate the local density βj of each instances using the equation
4. Compute distance from the nearest larger density point λ_j
5. Estimate the decision value υ according to equation and arrange it in ascending order υ′.
6. Find the smallest distance among instance 'j' that is clustered and instance 'i' not yet clustered
7. Add the instance i to the cluster of instance j
8. Repeat the process 6 and 7 until all the instances are clustered

End

5. Feature selection using Chaos Pigeon Inspired Optimization Algorithm

After clustering all the instances of dyslexia instances using density peak clustering, the feature selection is done using the newly developed Chaos Pigeon Inspired Optimization Algorithm (CPIO). The objective of CPIO is to discover the contribution of each features in the dataset and rank them based on their fitness value. Those features with highest fitness value are ranked higher and those subsets are used for classification of dyslexia instead of using all the features [15]. The detailed process of CPIO based optimized feature selection is explained in the following subsection.

A novel pigeon inspired metaheuristic algorithm which is developed spurred on pigeon's behavior of returning to its home. The factors involved in PIO is sun, magnetic field and landmarks. The map to return home is shaped in pigeon's mind using the magnetic field. According to the height and angle of the sun, pigeon's flight direction is updated. At the same time, overall population's directions will be updated [16]. The landmarks nearby its home, assist pigeons to fly close to their terminus. In PIO method two important operators are used to explain the behavior of the pigeons they are map and compass as shown in the figure 2. The sun and the magnetic field influence the performance of the artificial pigeons and landmark plays its vital role while returning to its home.

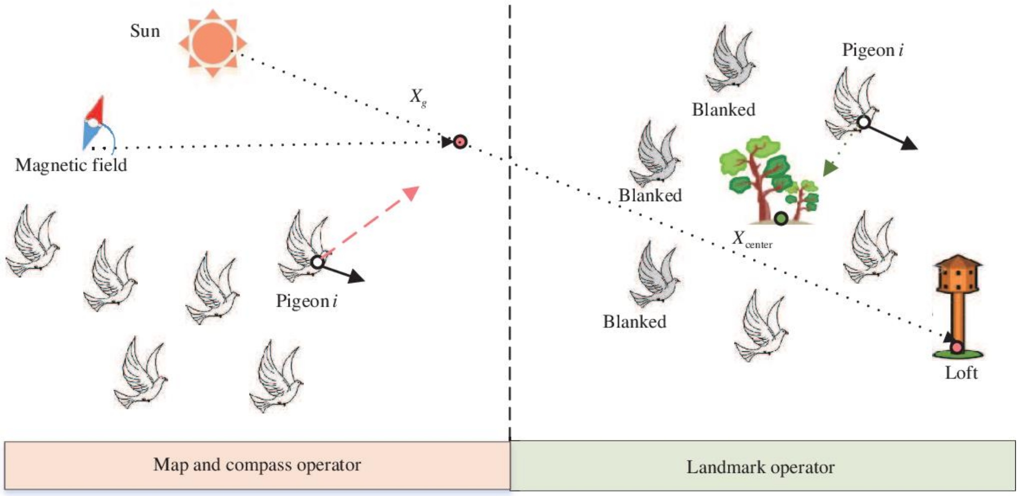

Fig. 2. Operators for PIO Map, compass and Landmark

5.1. Process of Velocity and Position Operator

Initialize the whole pigeon flock, number of pigeons are denoted as Ng, the dimension space of pigeons is denoted as Dm, R is the factor of compass and map. The position and velocity of the pigeon is mathematically formulated as

Pgi= [Pg1, Pg2, …Pgd]

Vli = [Vl1, Vl2,…Vld]

In PIO algorithm, the entire pigeon flock delivers a strong global searching strategy using the map and compass operators i.e position and velocity. They velocity and position are updated during each iteration using the formula

$$pg_i^{(t+1)} = pg_i^{(t)} + vl_i^{(t+1)}$$

$$vl_i^{(t+1)} = e^{-R*t} * vl_i^{(t)} + \tau_1(pg_{gb} - pg_i)$$

As shown in the equation, pg_{gb} denote the pigeon position with the highest fitness value of their whole flock during each update, t refers to number of iterations and τ_1 is a constant variable limited to 0 and 1. Unlike PSO algorithm which considers the individual bird optimal position, the PIO has the ability of strong global search to eradicate the issue of getting trapped in a local optimum.

6. Process of Landmark Operator of Pigeon

When the pigeon floc is reaching its destination, the operators maps and compass have less impact on returning behaviour of the entire population. Till the pigeons come home, the target nearest to the location would then serve as the new navigation guide. The pigeon inspired method replicates this physical spectacle by employing the pigeon with the highest fitness value in the whole population as center and it has most reasonable manner to drive the entirety to continuously update. Initially, entire pigeon flock must be omitted from the pigeons which cannot discover it to avert these pigeons from disturbing the repetition track of whole population formulated as

$$N_g^t = \frac{N_g^{t-1}}{2}$$

Selecting the pigeon that has maximum headship in whole flock computer based on the fitness value of each pigeon $\mathcal{F}(\mathrm{Pg}_i^t)$ is represented in the following equation

$$\mathrm{Pg}_{center}^t = \frac{\sum_{i=1}^{N_g^t} \mathrm{Pg}_i^t * \mathcal{F}(\mathrm{Pg}_i^t)}{N_g^t * \sum_{i=1}^{N_g^t} \mathcal{F}(\mathrm{Pg}_i^t)}$$

The pigeon which leads the whole flock with a constant value $\mathcal{E}$ which avoid the fitness value to fall zero and it is iterated as follows

$$\mathrm{Pg}_i^{t+1} = \mathrm{Pg}_i^t + \mathbb{Z}_4(\mathrm{Pg}_{center}^t - \mathrm{Pg}_i^t)$$

The maximum and minimum fitness value in each iteration is depicted as

$$\mathcal{F}(\mathrm{Pg}_i^t) = \begin{cases} \mathcal{F}(\mathrm{Pg}_i^t) formaximumvalue \\ \dfrac{1}{\mathcal{F}(\mathrm{Pg}_i^t) + \mathcal{E}} forminimumvalue \end{cases}$$

6.1 Chaos theory-based pigeon inspired algorithm

In this research work the chaos theory is induced to select the initial population, instead of using randomness in standard pigeon inspired optimization algorithm. The selection of random pigeons to discover the most influential attributes in dyslexia dataset will not offer guaranteed uniform understanding of features. This randomness results in local optima while performing clustering process-based feature selection. This work aims to improve the process of potential feature selection by adapting the chaotic mapping which avoids early convergence to results. It is mathematically formulated as

$$\mathrm{hi} + 1 = \begin{cases} 4\beta\rho_i(0.5 - \rho_i) & ,0 \le \rho_i < 0.5 \\ 1 - 4\beta\rho_i(0.5 - \rho_i)(1 - \rho_i), & 0.5 \le \rho_i, \le 1 \end{cases}$$

where $3.6 \le \beta \le 4$, $\beta = 4$, $\rho\ 0 =$ rnd $\in (0,1)$

6.2 Algorithm for Uncertainty Based Feature Selection for Effective Dyslexia Prediction

Input: Clustered Dyslexia Dataset

Output: Potential Feature selection

Procedure

1. Initialize the velocity and position of the parameters
2. Initialize the flock of pigeons using chaotic theory
3. Compute the fitness value of each pigeon in the flock
4. Choose the best position of the pigeon flock.
5. Update each pigeon's velocity and position
6. Calculate each pigeon to find the fitness values and remove the poor pigeon based on the historical global values of features.

7. Compute the centre position of flock using landmark operator using equation
8. Update global optimum and fitness value
9. If maximum iteration reached stop the process, else go to 7
Stop

6.3 Support Vector Machine

In this work to validate the subset of features the support vector machine is used to predict the accuracy of Dyslexia. The ultimate goal of Support Vector Machine (SVM) [ALA17] approach is to discover a hyperplane in N- number of attributes that noticeably classifies the data instances or records of dyslexia dataset. SVM is a simple machine learning algorithm as it generates more accuracy with less computation complexity. The SVM is commonly used in classification and regression tasks. But most of the time it is used as a classifier.

In order to split the two classes (Absence or Presence) of data instances, more possible hyperplanes are available among which one could be chosen. But the objective of SVM is to discover the maximum margin plane which has maximum distance among data instances of two classes. While choosing the plane with maximum margin it will reinforce the more confidence in classifying future incoming dyslexia dataset. The instances which lie closer to the hyperplane involves in influencing the orientation and position of the hyperplane.

7. Results and Discussions

This section discusses in detail about the performance of Density Peak Clustering based Pigeon Inspired Optimization (DPC-PIO) for identifying the significant feature subset which contributes more information to build an effective classification model for dyslexia prediction. The DPC-PIO is implemented using python software. The dataset is collected from keel repository with 12 medical features of 1065 children. The test conducted to detect the dyslexia are the 12 features namely Vocabulary, Verbal orders, Color, Visual memory, Visual-motor coordination, Perception of shapes, Spatial relations, Auditive perception, rhythm, pronunciation, Spatial orientation and Analysis of reading and writing. The proposed DPC-PIO algorithm selects feature subset which comprised of 6 potential attributes. The existing algorithms used for comparison are Mutual Information, Entropy feature selection and correlation-based feature selection.

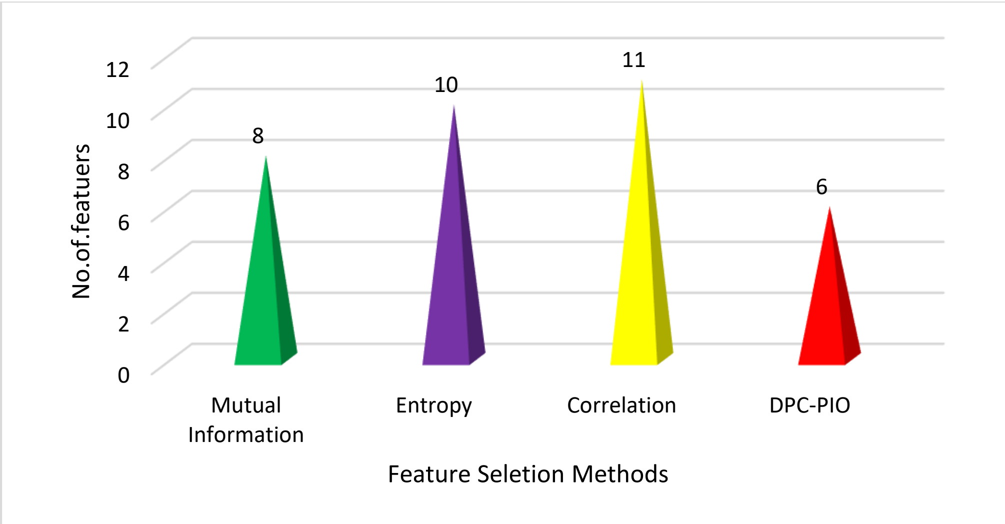

Fig. 3. Comparison of features selected

The figure 3 shows the number of features selected as significant features for dyslexia detection using four different models. The Mutual information which works based on the amount of information generated by the attributes depending on the class label selected 8 attributes in its feature subset. The entropy-based feature selection selects 10 attributes and based on the relationship among the attributes 11 attributes are selected as feature subset by correlation algorithm. The proposed model DPC-PIO selects feature subset with 6 potential attributes with the objective of maximal relevancy and minimal redundancy even in presence of uncertainty in selection of features from dyslexia dataset. The features selected by the entropy-based feature selection are Vocabulary, Verbal orders, Color, perception of shapes, Visual-motor coordination and Analysis of reading and writing.

Table 1 shows the performance comparison of three four different feature selection models namely Mutual Information (MI), Correlation based feature selection, Entropy algorithm and proposed DPC-PIO algorithm for dyslexia prediction based on precision, recall and accuracy.

Table 1. Performance comparison of feature selection using SVM

	Precision	Recall	Accuracy
DPC-PIO	0.87	0.75	0.92
MI	0.80	0.69	0.87
Correlation	0.71	0.64	0.82
Entropy	0.67	0.57	0.77

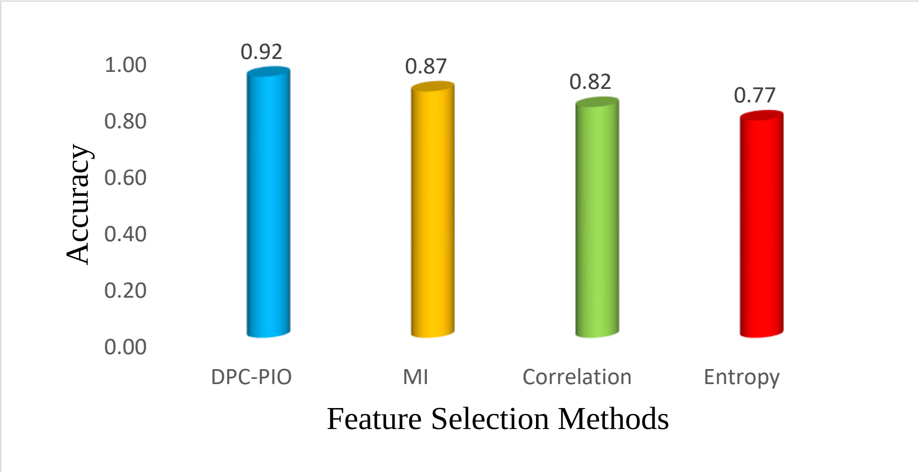

Fig. 4. Performance of accuracy in Dyslexia prediction

$$\text{Accuracy} = \frac{\text{No. of Correctly preidcted presence and absence of Dyslexia}}{\text{Total No. of. Instances}}$$

The result shows the accuracy performance of four different feature subset generated by four different feature selection models to classify the dyslexic dataset. While involving features for prediction, it is important to determine the contribution of each features and based on their score the best suited must be selected. The sub feature set selected by the proposed density peak clustering-based Pigeon Inspired Optimization produced highest accuracy rate compared to the state of art algorithms such as mutual information, Entropy and correlation feature selection models. DPC-PIO handles uncertainty in dataset and generates the score for each attribute based on the contribution in prediction of dyslexia by offering more information.

Fig. 5. Performance of precision rate in Dyslexia prediction

$$\text{Precision} = \frac{\text{Correctely predicited Dyslexic Patients}}{\text{Total No. of. Patients detected as Dyselxic}}$$

The figure 5 explores the precision rate of dyslexia prediction model which uses four different feature subset algorithms. The proposed model uses the Pigeon Inspired Optimization to take better decision by focusing on the uncertainty and risk of the dyslexia dataset. The similar pattern of instances is clustered using density peak clustering which obtain the clusters in a single step regardless of the shape and dimensionality of the space. With the clustered instances the features which highly offers information in clustering and improving the process of prediction is ranked using Pigeon Inspired Optimization and the best suited attributes are involved in dyslexia prediction. Thus, the proposed model DPC-PIO produced higher precision rate compared to other models namely entropy, correlation and mutual information-based feature selection algorithms.

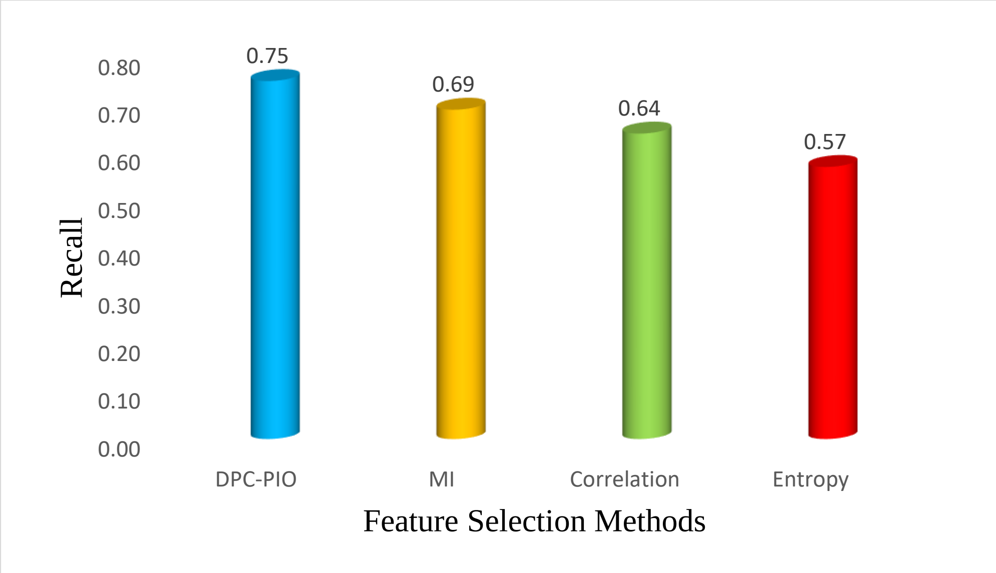

Fig. 6. Performance of recall in Dyslexia prediction

$$\text{Recall} = \frac{\text{No. of. Correctly predicited Dyslexic cases}}{\text{Total No. of. actual Dysexliex Cases in the dataset}}$$

The proposed density peak clustering-based Pigeon Inspired Optimization algorithm accomplishes highest recall rate compared to other feature selection models as depicted in the figure 6. The quality of dyslexia dataset is very low and this results in irrelevant feature selection by existing algorithms. The performance of the prediction model has close relationship with the nature of dataset used, so the proposed model handles the uncertainty by determining the similarity of instances in dyslexia dataset using density peek clustering which has the higher density of instances near the centroids than its surrounding regions. The imprecise measurement in selection of significant features is handled prominently by quantifying uncertainty using Pigeon Inspired Optimization algorithm.

8. Conclusion

The feature selection process is very challenging when there is uncertainty due to impreciseness, vagueness in dyslexic dataset. The accuracy of a prediction or classification model is greatly influenced by the features involved. When without considering the feature selection, if entire features are used for prediction, due to redundancy and irrelevancy it affects accuracy rate in detection of dyslexia. This paper discusses about the proposed unsupervised clustering with metaheuristic model for optimized feature selection to handle the issue of uncertainty in selection of potential feature subset in dyslexia dataset. The similarity among the instances are identified using density peak clustering algorithm and potential feature which provides highest information to distinguish the instances of different cluster is discovered by adapting chaotic pigeon inspired optimization. From the results obtained it is explored that the conventional mutual information, correlation and entropy models doesn’t have the ability to handle uncertainty in significant feature subset selection for dyslexia detection. The proposed model DPC-PIO well handled the uncertainties in potential feature subset selection to improve dyslexia prediction accuracy rate by accomplishing global optimization.

Acknowledgement

Funding

Not Applicable

Compliance with ethical standards

Conflict of interest

The authors declare that they have no conflict of interest.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Authors’ contributions

Dr. P. J. Arul Leena Rose supervised every step of the work and provided critical review and valuable input. All authors read and approved the final manuscript.

References

[1] International Dyslexia Association, Definition of Dyslexia Retrieved, 2017, pp. 15-06-2017.

[2] Andre Esteva, Brett Kuprel, Roberto, A.,Novoa, Justin Ko et al., “Dermatologist-level classification of skin cancer with deep neural networks”, Nature, pp. 542:115 – 118, Jan 2017.

[3] Poole, Alexandra, FarhanaZulkernine, and Catherine Aylward. "Lexa: A tool for detecting dyslexia through auditory pro-cessing", Computational Intelligence (SSCI), IEEE Sympo-sium Series on., 2017.

[4] American Psychiatric Association, Diagnostic and statistical manual of mental disorders: DSM-5, Autor, Washington, DC, 5th ed. edition, 2013.

[5] Andreadakis, Ioannis M.,Aslanides, and Maria Papadopouli., “Dyslexml: Screening toolfor dyslexia using machine learning”, CoRR, abs/1903.06274, 2019.

[6] Alex Frid and Larry M.,“Manevitz. Features and machine learning for correlating and classifying between brain areas and dyslexia”,CoRR, abs/1812.10622, 2018.

[7] Gilles Richard, Mathieu Serrurier, Dyslexia and Dysgraphia,“Prediction: A new machine learning approach”, Psychology, Computer Science, Mathematics, 2020.

[8] Jothi Prabha&Bhargavi, R., “Prediction of Dyslexia from Eye Movements Using Machine Learning”, Proceedings of the third

International Conference on Microelectronics Computing and Communication Systems, pp. 23-24, January 2019.
[9] Shamsuddin, S. N. W., Mat, N. S. F. N., Makhtar, M., "Journal of Fundamental and Applied Sciences", 9(6S), pp. 886-899, 2017.
[10] Appadurai, Jothi &Bhargavi, R., Ragala and Ramesh,"Prediction of dyslexia using support vector machine in distributed environment", International Journal of Engineering and Technology (UAE), 7, pp. 2795-2799, 2018.
[11] Opeyemi Lateef Usman, Ravie ChandrenMuniyandi, "CryptoDL: Predicting Dyslexia Biomarkers from Encrypted Neuroimaging Dataset Using Energy-Efficient Residue Number System and Deep Convolutional Neural Network", Symmetry,12(836), pp. 1-24, 2020.
[12] Ng Li Mun, Nur AnidaJumadi, "A Comparative Classification Models Study for Development of Early Dyslexia Screening System", Universal Journal of Educational Research, 8(3B), pp. 1-15, 2020.
[13] Tamboer, P., Vorst, H. C. M., Ghebreab, S., Scholte, H. S., "Machine learning and dyslexia: Classification of individual structural neuro-imaging scans of students with and without dyslexia", NeuroImage: Clinical, 11, pp. 508–514, 2016.
[14] Du, M., Ding, S., and Jia, H., "Study on density peaks clustering based on k-nearest neighbors and principal component analysis", Knowledge-Based Systems, 99(1), pp. 135–145, 2016.
[15] Duan, H., Qiao, P.,"Pigeon-inspired optimization: a new swarm intelligence optimizer for air robot path planning", Int. J. Intell. Comput. Cybern., 7, pp. 24–37, 2014.
[16] Bin, F., Wang, Z., Sun, J., "Niche quantum-behaved particle swarm optimization with chaotic mutation operator", Comput. Appl. Software, 26(1),pp. 50–52, 2019.https://sci2s.ugr.es/keel/datasets.php