https://doi.org/ 10.33472/AFJBS.6.9.2024.4547-4561

African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

Integrated ANN-GA method for classification of uncertain protein sequence with estimation of objective function

***T.SudhaRani**, Associate Professor, CSE, Aditya University, Surampalem, INDIA, Research Scholar, Department of CSE, Jawaharlal Nehru Technological University, Kakinada ,
sudharani.tirukoti@aec.edu.in

Dr.A.Yesu Babu[2]
Professor, Department of CSE, Sir C R Reddy College of Engineering

Dr.D.Haritha[3]
Professor, Department of CSE, Eluru, Jawaharlal Nehru Technological University, Kakinada

Article History: Volume 6,Issue 9, 2024

Received: 30 Apr 2024

Accepted : 17 May 2024

doi: 10.33472/AFJBS.6.9.2024.4547-4561

Abstract

The amino acids consist of vast range of uncertain protein sequences which are key elements for structural analysis of the cell function. This paper has proposed an ANN (Artificial Neural Network) incorporated through GA (Genetic Algorithm) for classification of protein sequences. The proposed ANN incorporated through GA(ANN-GA) utilized flower pollination algorithm (FPA) for identification of optimal points in the code of a protein data pattern. Components Essential Amino Acids composite consist of structural and solvent factors. Through formulation of objective function structural composites of protein sequences for classification. The developed ANN-GA estimates the sequence of protein fold prediction within the collected dataset. The dimensional analysis of data expressed that set of protein classification incorporates distinct structural folds and information. The ANN-GA performance is evaluated by means of accuracy of classification for recognized structural characteristics of protein sequences. The simulation analysis expressed that developed ANN-GA exhibits improved performance compared to the conventional techniques.

Keywords: Protein classification, Neural Network, Artificial Neural Network – Genetic Algorithm, ANN-GA, Bioinformatics, Flower Pollination Algorithm (FPA).

1. Introduction

Proteins, which are encoded in an organism's DNA, catalyse the majority of biochemical events inside cells and are also responsible for the transport of nutrients and the transmission of signals from one cell to another [1]. The study of how genetic information is decoded to produce a functioning protein is thus an important part of bioinformatics [2]. The fundamental concepts behind this decoding are well-established. The genetic code dictates that in which a gene's nucleotide sequence is converted into an amino acid sequence. To be functional, a protein's linear chain of amino acids has to take on a 3D form, known as its "native" shape [3].

Homologous proteins may be difficult to classify together if just sequence information is provided in today's genomics [4]. Comparisons based on sequence similarity are not a reliable indicator of evolutionary relationships, making this challenge problematic. New sequences' potential functions may be predicted by comparing their sequence–structure and sequence–sequence relationships. Proteins that have a tight evolutionary link may be identified via pairwise sequence alignment. However, Physically related proteins without considerable sequence similarity, it is inefficient [5]. Despite this, simple algorithms have yielded respectable outcomes in the real world. Methods that have been presented so far rely heavily on thresholding some measure of the distance between the sequences in question. Setting the threshold at a number that provides effective groups is obviously critical [6]. For example, two sequences with a high degree of sequence similarity will only be allocated to the same cluster when the threshold is set to conservatively. That's because it's previously been demonstrated that proteins that have a high degree of sequence similarity are possible to evolutionarily linked; a traditional criterion will likewise lead numerous singular groups [7]. If the threshold is loosened, on the other hand, numerous unrelated proteins will be grouped together.

In this work, an integrated approach for accurate classification of protein sequences was developed. The proposed model integrates Genetic Algorithm (AG) and ANN. To improve the accuracy Flower Pollination Algorithm (FPA) is implemented as a Genetic Algorithm with the formulation of the objective function. The developed method involved in identification of optimal points in the protein sequences and increases the overall accuracy of protein sequence classification. Simulation analysis expressed that proposed ANN incorporated through GA (ANN-GA) exhibits significant performance compared to the existing models.

This study is organizing as follows: The segment II facilitates the literature related to existing techniques utilized for protein sequence classification. In Section III materials and methods considered for classification of protein sequences is presented. In section IV, ANN incorporated through GA (ANN-GA) with optimization algorithm FPA is presented. The section V provides simulation results and followed by conclusion in section VI.

2. Related Works

This section presents existing studies focused on amino acid analysis. The amino acids are representation of numbers those are plotted in geometric representation [8]. The analysis of protein sequences is based on those estimated sequences. Similarly, sequence of protein is estimated based on the Dayhoff's amino acid estimation with inclusion of substitution matrix [9]. The feature of

amino acids is estimated using Bayesnet Naïve of classification of features. The analysis is performed based on the consideration of 1D, 2D and 3D features of amino acids. The analysis expressed that for Bayesnet Naïve 2D is effective.

Similarly, a Logistic function-based machine learning algorithm for classification of protein sequences was developed for the study of classification of protein sequences [10]. The proposed model utilizes dimension reduction technique for identification of amino acid representation. The analysis of results stated that for logistics function, 2D sequences are more effective. Naive Bayes classifier is developed for visualization of protein sequences [11]. The analysis shows that vector representation improves the order of amino acid with vector representation of proteins (VRP). Similarly, amino acid is represented in geometrical form using classifier Support vector machine[12]. The incorporation of Support Vector machine (SVM)significantly illustrates the geometric feature of protein sequences.

To approximate sequence as well as calculation of correlations Partial Decision Tree are used [13]. The developed classifier involved in detection of protein sequences with evolution of protein and prediction of cleavage site proteins in the human body. To predict those leverage proteins in the human body, aDecision tree J4.8 estimates the overall sequence representation of protein with calculation of physicochemical characteristics was developed [14]. The analysis is based on the composition of protein sequences for information evaluation. Further, protein sequences are subjected to high dimensional factors for cleavage of the protein sites or domains for classification [15].

The analysis of protein sequence classification expressed that existing literature does not provide significant information about classification characteristics. The analysis is based on geometric features alone. Also, information about the protein sequence datasets is not presented. To overcome those limitation this research developed ANN incorporated through GA (ANN-GA) for classification of uncertain protein sequences. The proposed approach utilizes Flower Pollination Algorithm (FPA) for improving performance of the classification [16].

3. Materials and Methods

The protein sequence dataset is collected from Rost-Sander data set with diverse domains for classification. The collected dataset incorporates 2930 sequences with CATH v2.4 domains [17]. This paper focused on classification of uncertain folded protein sequences with consideration of different factors such as CATH, Architecture, topology and class of protein. CAT classification as fold is defined by set of structures called Set structure [17]. Use of set of structures with diversity of sequences means which information is duplicate -free and that issue of identifying constructional similarities among all protein pairs considered is non-trivial [18]. Structures for 2930 sequences were chosen as follows:

(i) By their SPACI score, sort all 34,287 CATH v2.4 domains

(ii) Begin through highest domain SPACI score and delete all domains that share major sequence similarity from list.

(iii) Redo stage (ii) for all domains in set which have not been deleted. From this process, 2930 domains are produced which is referred as CATH2930. In which it contains 769

folds and single element (482). In order to facilitate comparative study, five of CATH2930's most populous folds were selected as a more extensive test range that includes at least one-fold from each CATH class: 3.20.20, (TIM-like, 67 representatives), and 3.60.40 (immunoglobulin-like, 169 representatives) are three alternating alpha/ folds in CATH fold 1.10.10 (arc repressor) (Rossmann fold, 215 representatives). There are a total of 605 CATH2930 proteins in these five folds (Set CATH605).

(iv) Examples of protein structures for 5folds are represented in Figure-1. New publication of CATH, CATH3.5, which was available in June 2020, is 2ndset of proteins included in this study.

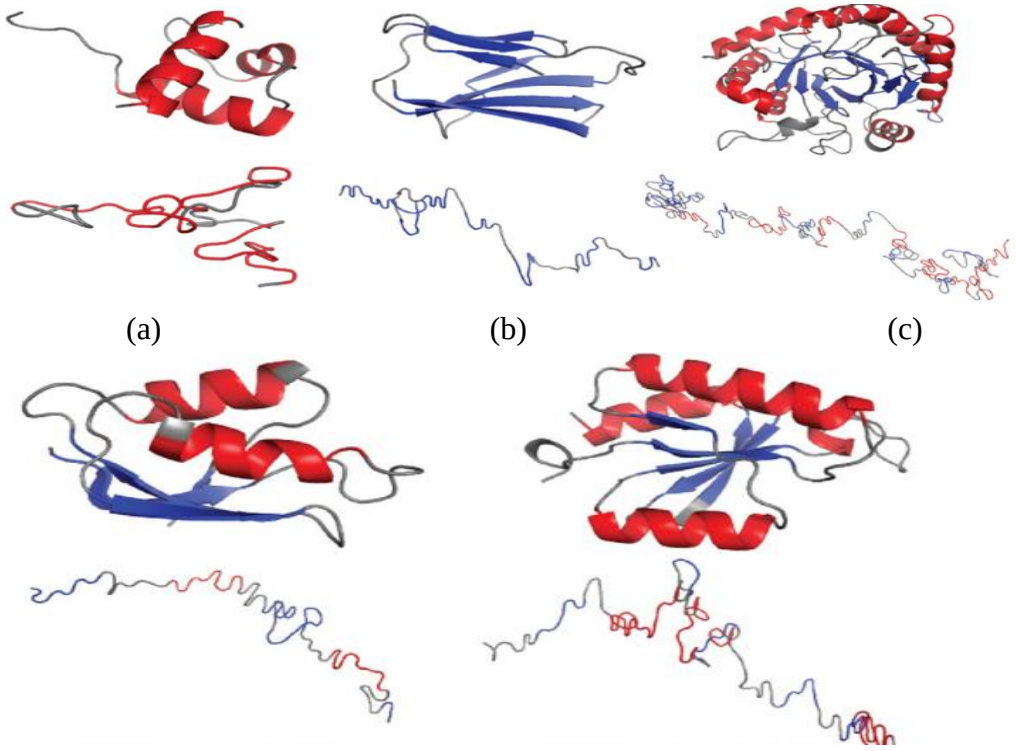

Figure 1: Structure of Protein a) CATH topology 1.10.10 b) CATH topology 2.60.40 c) CATH topology 3.20.20 d) CATH topology 3.30.70 e) CATH topology 3.40.50

To produce CATH2930, similar to this method is used in which low similarities sequence was designed.

(i) Randomize CATH3.5's in all 173,536 domains;

(ii) In randomized list, start with first domain and delete all domains that share major series similarities with it (FASTA E-value b 10−4).

(iii) (iii) With all domains in list which is not deleted, repeat step (ii). CATH35 E4 is referred to as range of 8862 domains arising from this process [19].

Integrated ANN-GA for Protein Sequence Classification

In this paper, protein sequence data were classified for identification of features within the dataset [20]. Generally, protein sequence is related to the multi-label optimization issue. Function of Fitness is represented in equ (1).

$$\chi^* = \min P(\chi) \quad (1)$$

$$P(\chi) = I(\chi) + a_1 S(\chi) \quad (2)$$

Where,

(i) Labeling function variable χ allows every point in protein sequence domain Ψ to label L .protein data intensity is expressed in equ (3).

$$\chi: q\epsilon\psi \rightarrow \chi(q)\epsilon\Re \quad (3)$$

$\mathcal{R} \subset B$ is represented by possible output intensities.

(i) Data term I is represented as

$$I(\chi) = \sum_{q\epsilon\Psi} I_q(\chi(q)) \quad (4)$$

$$I(\chi) = \sum_{L\epsilon B} \sum_{q\epsilon s_1} [w_1(L - v_1(q))^2 + [w_2(L - v_2(q))^2] \quad (5)$$

From equ (5), protein sequence data are denoted as$v_1: \Psi \rightarrow S, v_1: \Psi \rightarrow S$. L-label region $S_1$is $\{q \in \Psi | \chi(q) = L|\}$and weight is $w_1$andw_2.

$$r_1 = |\nabla v_1| * K \; and \; r_2 = |\nabla v_2| * K \quad (6)$$

$$w_1 = \frac{r_1}{r_1 + r_2} \text{ and } w_2 = \frac{r_2}{r_1 + r_2} \quad (7)$$

(i) The data points of protein have equivalent data values encouraged by smooth solution are represented as S .

$$S(\chi) = \sum_{\{q,p\}\epsilon M} s(\chi(q), \chi(p)) \quad (8)$$

Set M consists of everydata class pair as q and p . Truncated absolute value is utilized to defines local neighborhood of q and$s(\chi(q), \chi(p))$, which is given in equ (9).

$$s(\chi(q), \chi(p)) = \min(a_2, |L_q - L_p|) \quad (9)$$

Where positive constant is represented as a_2 .Protein sequence groups and marks require one-to-one correspondence [21]. Using number of labels, number of all possible data values are represented which is equal to output image.In dynamic ranges in images, larger computational load occurs [22]. Through reformulating data terms into accountability labelling, number of labels reduces [23]. Performance of image fusion is role of data from protein sequence as$v_1$andv_2.

$$v_\beta = \beta_{v_1} + (1 - \beta)v_2 \quad (10)$$

From the above equation, $\beta$is the transparency of data for protein sequence identification asi.e.,$\beta(q)\epsilon[0,1]\forall q\epsilon\Psi$.

Equ (2) represents data term,

$$I(\chi) = \sum_{L\epsilon B} \sum_{q\epsilon S_1} [w_1 \left(v_\beta(q,L) - v_1(q)\right)^2 + w_2 \left(v_\beta(q,L) - v_2(q)\right)^2] \quad (11)$$

Hence,

$$v_\beta(q,L) = \frac{L}{M_1} v_1(q) + (1 - \frac{L}{M_1}) v_2(q) L\epsilon \Re_\beta \quad (12)$$

M_l is user-specified number of labels.$\Re_\beta$ i.e. {1,2….M_l is new set of non-negative integer label.

Protein Sequence Analysis with Flower Pollination Algorithm (FPA) In ANN - GA:

Algorithm 1: ANN - GA with FPA

Initialize

Begin FPA specifications like no. Of flowers (m) , switching probability (SP)

Begin m flower population (P)

$[P = \overrightarrow{y^1}, \overrightarrow{y^2}, \ldots \overrightarrow{y^m}]$

Determine optimal solution $y^* \in P$

$i \leftarrow 1$

Dofor each $t \in (1,m)$

if $V(0,1) < SP$

for each $a \in (1, DV)$ do *Global pollination

if $V(0,1) < 0.5$

$y_{i+1}^{t_a} = y_i^{t_a} + \gamma L(y^* - y_i^{t_a})$

else

$y_{i+1}^{t_a} = \left(\max\left(y_i^{j_a}, y_i^{k_b}\right) - \min\left(y_i^{j_a}, y_i^{k_b}\right) \cdot s_2 + \min\left(y_i^{j_a}, y_i^{k_b}\right)\right)$

Where $j,k \in (1,m), j \neq k$

end if

end for

else

$y_{i+1}^t = y_i^t + V(0,1) \times (y_i^l - y_i^n)$ * Local pollination

Where $l,n \in (1,m), l \neq n$

Endif

if $(Fy_{i+1}^t) < F(y_i^t)$ then

update y_i^t with y_{i+1}^t then

end if

if $(Fy_{i+1}^t) < F(y^*)$ then

update y^* with y_{i+1}^t then
end if
end for
$i \leftarrow i+1$
met $i > i_{\max}$
End

In Flower Pollination (FPA) optimization algorithm, flower pollination is a natural process that solves huge optimization issues. Two basic modes of pollination for reproductive phase are biotic & crosspollination as well as abiotic or self-polllination [24]. During cross-pollination, pollen was borne over long distances by beetles, bees, bats and butterflies [25]. Water diffusion or wind moves pollen to shorter lengths during self-pollination [26] FPA principal flower or solution population is formulated in Eq. (13).

$$P = \begin{bmatrix} \overrightarrow{Y^1} = \begin{bmatrix} y_1^1 & y_2^1 & \cdots & y_{DV}^1 \end{bmatrix} \\ \overrightarrow{Y^2} = \begin{bmatrix} y_1^2 & y_2^2 & \cdots & y_{DV}^2 \end{bmatrix} \\ \vdots \quad \vdots \quad \vdots \quad \ddots \quad \vdots \\ \overrightarrow{Y^m} = \begin{bmatrix} y_1^m & y_2^m & \cdots & y_{DV}^m \end{bmatrix} \end{bmatrix} = \begin{bmatrix} F\left(\overrightarrow{Y^1}\right) \\ F\left(\overrightarrow{Y^2}\right) \\ \vdots \\ F\left(\overrightarrow{Y^m}\right) \end{bmatrix} \tag{13}$$

Where m denotes number of solution or $(\overrightarrow{y^1}, \overrightarrow{y^2}, \ldots \overrightarrow{y^m})$ denotes flower population and DV is denoted as number of decision variable [27]. $F(\overrightarrow{Y^k})$ is solution fitness for K= {1, 2…, m}.Biotic or cross-pollination is to reflect Levy flights assets with the global search. Using Levy flight's function, solution updation is represented in Eq. (14).

$$y_{i+1}^t = y_i^t + \gamma L(y^* - y_i^t) \tag{14}$$

Depends on Levy distribution as well as scaling factor, pollination power is represented as L andγ.

$$L \sim \frac{\chi \Gamma(\chi) \sin(\Pi\chi/2)}{\Pi} \frac{1}{R^{1+\chi}} \tag{15}$$

Standard gamma function is distributed from $R > 0$ equ (15) to steps.Abiotic or self-pollination represents local process on basis of two randomly chosen solutions. Equ (16) denotes local pollination of FPA.

$$y_{i+1}^t = y_i^t + s_1 \times (y_i^j - y_i^k) \tag{16}$$

Therefore, in related species, 2 random solutions picked from various flowers are y_i^j andy_i^k. At interval 0 and 1, s_1 value is uniformly distributed.i.e. [0, 1]. Improving search efficacy of an algorithm is successful use of manipulation and discovery search techniques [28]. Furthermore,

we adjust investigation of FPA and then current solution conveniently uses new solution. Exploration search is a technique of increasing convergence speed.FPA discusses problem of search space generated by option of simple global pollination as well as heuristic search space [29]. Two randomly selected parent data will reduce heuristic bounded search space to a few search space regions, which is represented in equ (17).

$$y_{i+1}^{t_a} = \left(\max\left(y_i^{j_a}, y_i^{k_b}\right) - \min\left(y_i^{j_a}, y_i^{k_b}\right)\right) \cdot s_2 + \min\left(y_i^{j_a}, y_i^{k_b}\right)\Big) \quad (17)$$

At i iteration, t^{th} solution of a^{th} variable is $y_i^{t_a}$. The 2 random solutions y_i^a & y_i^b by s_1 and s_2 uniformly allocated to intervals 0 and 1. Heuristic bounded search space method concentrated on area based on current population experience [30]. To avoid algorithm from stuck in local minima by exploring entire search field, simple global pollination is necessary [31]. ANN-GA with FPA measurements is explained in Algorithm 1.

4. Results and Discussion

This paper focused on classification of protein sequence through proposed ANN-GA with FPA [32]. The approach designed for classification of protein sequences for classification is based on formulation of objective function [33]. The simulation analysis consists of collected from Rost-Sander data set, which consists of subset of that dataset [34]. The collected protein dataset is evaluated in MATLAB 2019b with development of ANN for protein sequences classification [35]. Performance analysis of ANN-GA is calculated based on Cross-Entropy with gradient conjugate scaling [36].

For evaluation of prediction system and given by equation (18), overall accuracy, which can be considered as most significant measure, is used [37].

$$Accuracy = \sum_{i-1}^{i-k} \frac{TP_i}{N} \quad (18)$$

Where N is the total no. of sequences and TPi is true positive in i families.Using Elman RNN, relationship between no. of d & average precision of all MIVs, it is accuracy of varies considerably from 68.58% (d1⁄412) to 89.21% on dataset concerning d. For class Ci, sensitivity is correctly predicted fraction of proteins from class Ci; for class Ci, specificity is described as correctly predicted portion of proteins not from class Ci.Formula is given below for each calculation in equations (19) and (20):

$$Se = \left(TP \times 100\right) / TP + FN \quad (19)$$

$$Sp = \left(TN \times 100\right) / (TP + FN) \quad (20)$$

To determine feasibility of our process, we also create another separate dataset contains 1100 proteins, obtained from all families [38]. In figure 2 best validation performance measured for various epoch are presented [39].

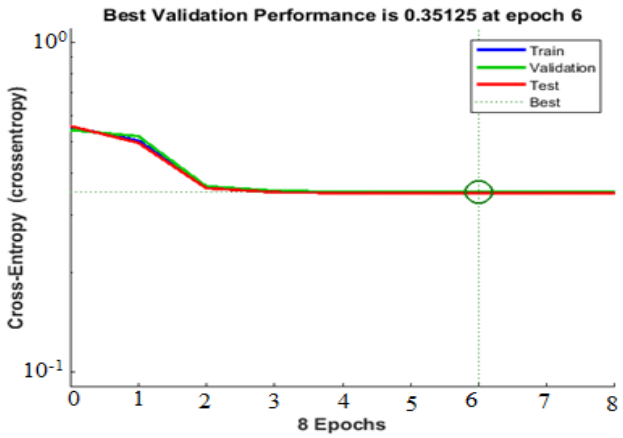

Figure 2: Estimation of best validation

Best validation shall be evaluated concerning research, processing, and validation of dataset [40]. Analysis is concerning consideration of epochs and optimal point is identified at epochs 6. In figure 3 gradient measured for 8 epochs is presented for proposed ANN-GA.

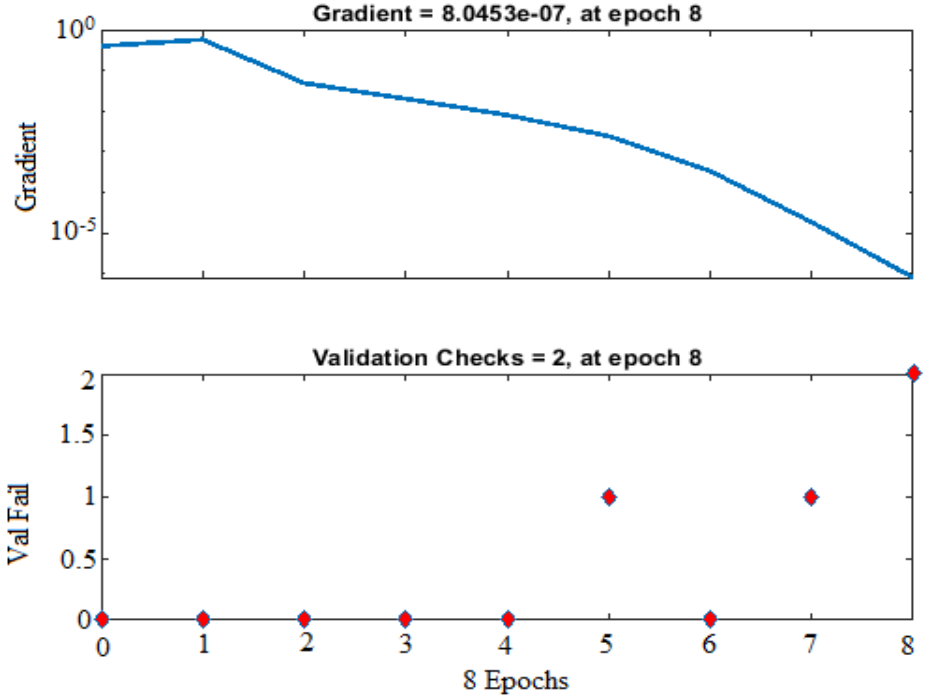

Figure 3: Gradient Measured for ANN-GA

As stated in figure 2 optimal point is observed at epochs 6 this leads to gradient measured at 6 are falls and fluctuating. Analysis expressed that gradient values are significantly higher with minimum number of epochs. In figure 4 confusion matrix designed for proposed ANN-GA is presented.

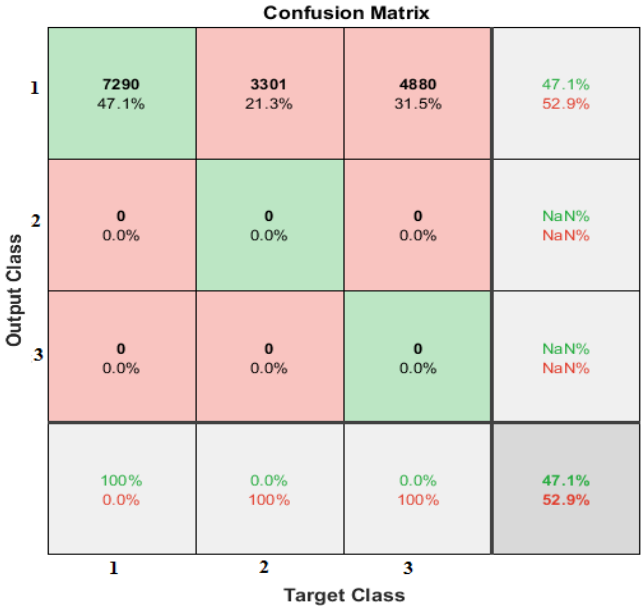

Figure 4: Confusion Matrix for ANN-GA

The confusion matric estimated for ANN-GA exhibits effective performance for calculated TP, TN, FN and FP values. Analysis is based on consideration of FPA optimal values for protein sequence classification. In figure 5, proposed ANN-GA ROC is presented for different classes of protein sequences.

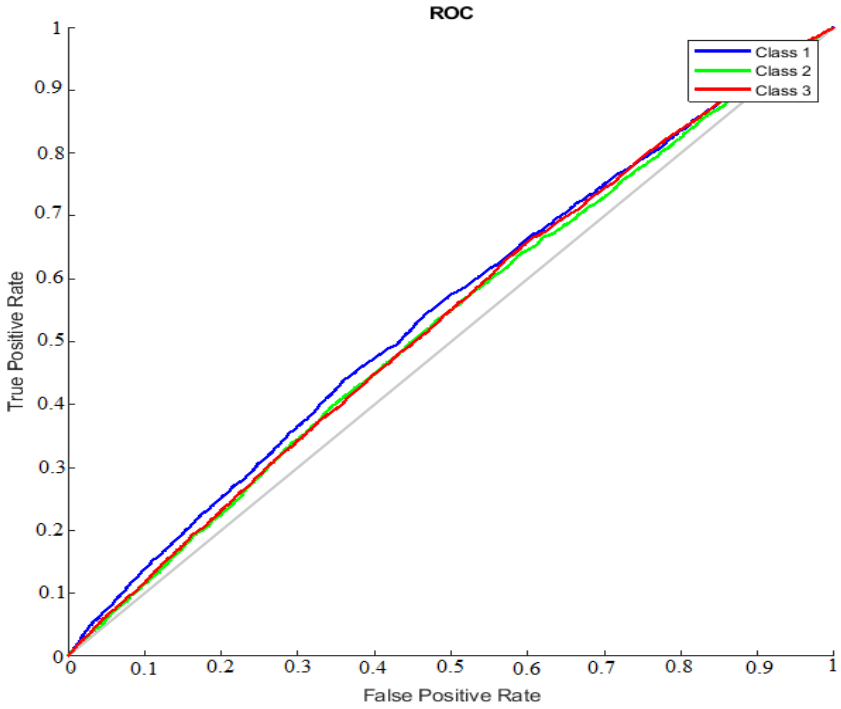

Figure 5: Analysis ROC

Chosen data set consists of 3 data groups for classification of protein chain, ROC is plotted on basis of TPR and FPR computed for proposed method. Analysis is based on consideration of data point characteristics. The smooth ROC curve expressed a significant curve which increases the

overall classification accuracy of protein sequences. Sensitivity and specificity of each classification are summarized in table. 1. The parameter considered for analysis hidden layer, input layer, & output later those are indicated as ξ_x, ξ_y and ξ_z. The performance of proposed ANN-GA with FPA is comparatively examined with existing technique for measured accuracy %. In table 1 overall accuracy measure within neural network input layer.

Table 1: Comparison of Accuracy

Machine Learning	ξ_x	ξ_y	ξ_z
SVM	79.83	88.75	95.85
Partial DT	79.25	84.70	91.40
Naive Bayes	79.43	83.6	91
Logistic function	84.56	82	94
ANN-GA with FPA	88.45	91.23	96.23
DT J4.8	74.10	89.20	92.10
BayesnetNaı¨ve	85.22	79.23	87.2

The comparative analysis of results expressed that proposed ANN-GA with FPA exhibits higher accuracy value for every layer compared with existing techniques. When compared with other methods, proposed method gives more accuracy.

5. Conclusion

This paper presented ANN incorporated through GA (ANN-GA) model for classification of protein sequences using Flower Pollination Algorithm (FPA). The incorporated FPA is based on evaluation of objective function of data labels and estimate protein sequences. The proposed model incorporates various classification techniques, which provides discrete model for resolving protein sequences of different length. Further, the proposed model incorporates considerable range of sequence order that perform effectively on hidden or complicated sequences of protein. The utilization of FPA algorithm is involved in effective identification of optimal data points for classification and improves the overall accuracy rather than existing technique. In future, the proposed model will be implemented with web-server method for user-friendly applications.

References

1. Li, J., & Koehl, P. (2014). 3D representations of amino acids—applications to protein sequence comparison and classification. Computational and structural biotechnology journal, 11(18), 47-58.
2. Bharill, N., Tiwari, A., & Rawat, A. (2015). A novel technique of feature extraction with dual similarity measures for protein sequence classification. Procedia Computer Science, 48, 795-801.
3. Kundu, S. (2017). Mathematical basis of improved protein subfamily classification by a HMM-based sequence filter. Mathematical Biosciences, 293, 75-80.

4. Kumar, V. V., Raghunath, K. M. K., Rajesh, N., Venkatesan, M., Joseph, R. B., & Thillaiarasu, N. (2021). Paddy Plant Disease Recognition, Risk Analysis, and Classification Using Deep Convolution Neuro-Fuzzy Network. Journal of Mobile Multimedia. doi:10.13052/jmm1550-4646.1829.
5. Udagedara, S. R., Wijekoon, C. J., Xiao, Z., Wedd, A. G., & Maher, M. J. (2019). The crystal structure of the CopC protein from Pseudomonas fluorescens reveals amended classifications for the CopC protein family. Journal of Inorganic Biochemistry, 195, 194-200.
6. You, Z. H., Li, X., & Chan, K. C. (2017). An improved sequence-based prediction protocol for protein-protein interactions using amino acids substitution matrix and rotation forest ensemble classifiers. Neurocomputing, 228, 277-282.
7. Chen, C., Zhang, Q., Yu, B., Yu, Z., Lawrence, P. J., Ma, Q., & Zhang, Y. (2020). Improving protein-protein interactions prediction accuracy using XGBoost feature selection and stacked ensemble classifier. Computers in Biology and Medicine, 123, 103899.
8. Molkenthin, N., Mühle, S., Mey, A. S., & Timme, M. (2020). Self-organized emergence of folded protein-like network structures from geometric constraints. Plos one, 15(2), e0229230.
9. Mirceva, G., Ivanoska, I., Naumoski, A., & Kulakov, A. (2019, May). Feature Selection for Improved Classification of Protein Structures. In 2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO) (pp. 1013-1018). IEEE.
10. Lambert, S. A., Yang, A. W., Sasse, A., Cowley, G., Albu, M., Caddick, M. X., ... & Hughes, T. R. (2019). Similarity regression predicts evolution of transcription factor sequence specificity. Nature genetics, 51(6), 981-989.
11. Meng, C., Jin, S., Wang, L., Guo, F., & Zou, Q. (2019). AOPs-SVM: a sequence-based classifier of antioxidant proteins using a support vector machine. Frontiers in Bioengineering and Biotechnology, 7, 224.
12. Swain, S. N., Makunin, A., Dora, A. S., & Barik, T. K. (2019). SNP barcoding based on decision tree algorithm: A new tool for identification of mosquito species with special reference to Anopheles. Acta tropica, 199, 105152.
13. Li, J., Zhu, P., & Zou, Q. (2019, May). Prediction of thermophilic proteins using voting algorithm. In International Work-Conference on Bioinformatics and Biomedical Engineering (pp. 195-203). Springer, Cham.
14. Bowman, M. A., Riback, J. A., Rodriguez, A., Guo, H., Li, J., Sosnick, T. R., & Clark, P. L. (2020). Properties of protein unfolded states suggest broad selection for expanded conformational ensembles. Proceedings of the National Academy of Sciences, 117(38), 23356-23364.
15. Praveen Sundar, P.V., Ranjith, D., Vinoth Kumar, V. et al. Low power area efficient adaptive FIR filter for hearing aids using distributed arithmetic architecture. Int J Speech Technol 23, 287–296 (2020). https://doi.org/10.1007/s10772-020-09686-y.
16. Reddy, D. A., Shambharkar, S., Jyothsna, K., Kumar, V. M., Bhoyar, C. N., & Somkunwar, R. K. (2022, September). Automatic Vehicle Damage Detection Classification framework using Fast and Mask Deep learning. In 2022 Second International Conference on Computer Science, Engineering and Applications (ICCSEA) (pp. 1-6). IEEE.
17. Kumar, A., Janakirani, M., Anand, M., Sharma, S., Vivekanand, C. V., & Chakravarti, A. (2022). Comparative Performance Study of Difference Differential Amplifier Using 7 nm and 14 nm FinFET Technologies and Carbon Nanotube FET. Journal of Nanomaterials, 2022.

18. Vijetha, T., Mallick, P. S., Karthik, R., & Rajan, K. (2022). Effect of Scattering Angle in Electron Transport of AlGaN and InGaN. Advances in Materials Science and Engineering, 2022.
19. Kumar, H., Prasad, R., Kumar, P., & Hailu, S. A. (2022). Friction and Wear Response of Friction Stir Processed Cu/ZrO 2 Surface Nano-Composite. Journal of Nanomaterials, 2022.
20. Bhanuprakash, Lokasani, Soney Varghese, and Sitesh Kumar Singh. "Glass Fibre Reinforced Epoxy Composites Modified with Graphene Nanofillers: Electrical Characterization." Journal of Nanomaterials (2022).
21. Ramaiah, V. S., Singh, B., Raju, A. R., Reddy, G. N., Saikumar, K., & Ratnayake, D. (2021, March). Teaching and Learning based 5G cognitive radio application for future application. In 2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE) (pp. 31-36). IEEE.
22. Mohammad, M. N., Kumari, C. U., Murthy, A. S. D., Jagan, B. O. L., & Saikumar, K. (2021). Implementation of online and offline product selection system using FCNN deep learning: Product analysis. Materials Today: Proceedings, 45, 2171-2178.
23. Padmini, G. R., Rajesh, O., Raghu, K., Sree, N. M., & Apurva, C. (2021, March). Design and Analysis of 8-bit ripple Carry Adder using nine Transistor Full Adder. In 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 1982-1987). IEEE.
24. Dr. k. Raju, A. Sampath Dakshina Murthy, Dr. B. Chinna Rao, Sindhura Bhargavi, G. Jagga Rao, K. Madhu, K. Saikumar. A Robust And Accurate Video Watermarking System Based On SVD Hybridation For Performance Assessment International Journal of Engineering Trends and Technology, 68(7),19-24.
25. Saba, S. S., Sreelakshmi, D., Kumar, P. S., Kumar, K. S., & Saba, S. R. (2020). Logistic regression machine learning algorithm on MRI brain image for fast and accurate diagnosis. International Journal of Scientific and Technology Research, 9(3), 7076-7081.
26. Saikumar, K. (2020). RajeshV. Coronary blockage of artery for Heart diagnosis with DT Artificial Intelligence Algorithm. Int J Res Pharma Sci, 11(1), 471-479.
27. Saikumar, K., Rajesh, V. (2020). A novel implementation heart diagnosis system based on random forest machine learning technique International Journal of Pharmaceutical Research 12, pp. 3904-3916.
28. Raju K., Chinna Rao B., Saikumar K., Lakshman Pratap N. (2022) An Optimal Hybrid Solution to Local and Global Facial Recognition Through Machine Learning. In: Kumar P., Obaid A.J., Cengiz K., Khanna A., Balas V.E. (eds) A Fusion of Artificial Intelligence and Internet of Things for Emerging Cyber Systems. Intelligent Systems Reference Library, vol 210. Springer, Cham. https://doi.org/10.1007/978-3-030-76653-5_11
29. Sankara Babu B., Nalajala S., Sarada K., Muniraju Naidu V., Yamsani N., Saikumar K. (2022) Machine Learning Based Online Handwritten Telugu Letters Recognition for Different Domains. In: Kumar P., Obaid A.J., Cengiz K., Khanna A., Balas V.E. (eds) A Fusion of Artificial Intelligence and Internet of Things for Emerging Cyber Systems. Intelligent Systems Reference Library, vol 210. Springer, Cham. https://doi.org/10.1007/978-3-030-76653-5_12
30. Kiran Kumar M., Kranthi Kumar S., Kalpana E., Srikanth D., Saikumar K. (2022) A Novel Implementation of Linux Based Android Platform for Client and Server. In: Kumar P., Obaid A.J., Cengiz K., Khanna A., Balas V.E. (eds) A Fusion of Artificial Intelligence and Internet of Things

for Emerging Cyber Systems. Intelligent Systems Reference Library, vol 210. Springer, Cham. https://doi.org/10.1007/978-3-030-76653-5_8

31. Shravani, C., Krishna, G. R., Bollam, H. L., Vatambeti, R., & Saikumar, K. (2022, January). A Novel Approach for Implementing Conventional LBIST by High Execution Microprocessors. In 2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 804-809). IEEE.
32. Nagendram, S., Nag, M. S. R. K., Ahammad, S. H., Satish, K., & Saikumar, K. (2022, January). Analysis For The System Recommended Books That Are Fetched From The Available Dataset. In 2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 1801-1804). IEEE.
33. Jothsna, V., Patel, I., Raghu, K., Jahnavi, P., Reddy, K. N., & Saikumar, K. (2021, March). A Fuzzy Expert System for The Drowsiness Detection from Blink Characteristics. In 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 1976-1981). IEEE.
34. Appalaraju, V., Rajesh, V., Saikumar, K., Sabitha, P., & Kiran, K. R. (2021, December). Design and Development of Intelligent Voice Personal Assistant using Python. In 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N) (pp. 1650-1654). IEEE.
35. Naidu, T. P., Gopal, K. A., Ahmed, S. R., Revathi, R., Ahammad, S. H., Rajesh, V., ... & Saikumar, K. (2021, December). A Hybridized Model for the Prediction of Heart Disease using ML Algorithms. In 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N) (pp. 256-261). IEEE.
36. Teju, V., Sowmya, K. V., Yuvanika, C., Saikumar, K., & Krishna, T. B. D. S. (2021, December). Detection of Diabetes Melittus, Kidney Disease with ML. In 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N) (pp. 217-222). IEEE.
37. Mannepalli, K., Raju, K. B., Sirisha, J., Saikumar, K., & Reddy, K. S. (2021, December). LOW complex OFDM channel design using underwater-acoustic-communication using machine learning techniques. In 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 1505-1513). IEEE.
38. Kumar, K. S., Vatambeti, R., Narender, M., & Saikumar, K. (2021, December). A real time fog computing applications their privacy issues and solutions. In 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 740-747). IEEE.
39. Ajay, T., Reddy, K. N., Reddy, D. A., Kumar, P. S., & Saikumar, K. (2021, December). Analysis on SAR Signal Processing for High-Performance Flexible System Design using Signal Processing. In 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 30-34). IEEE.
40. Raju, K. B., Lakineni, P. K., Indrani, K. S., Latha, G. M. S., & Saikumar, K. (2021, October). Optimized building of machine learning technique for thyroid monitoring and analysis. In 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC) (pp. 1-6). IEEE.

Authors Profile

Mrs.T.Sudha Rani, working as Assoc. Professor, Department of Computer Science and Engineering, Aditya Engineering College. Currently Pursuing Ph.D at JNTUK, Kakinada. Her Area of Research are Bioinformatics and Data Mining.
E-mail: sudha.mahi84@gmail.com

Dr. A.Yesu Babu, Currently working as a Professor, Department of Computer Science and Engineering, Sir C R Reddy College of Engineering, Eluru. He is having 31 years of Academic, Research &Academic Administration experience. Published 43 Research Papers in International journals and 6 chapters. Reviewer of Research publications for premier publishing groups like Springer, Elsevier, Inderscience and a number of SCOPUS and SCI indexed journals.

Dr. D. Haritha, She is working as Associate Professor in Computer science and Engineering Department at Jawaharlal Nehru Technological University Kakinada. She has 17+ years of experience. She guided 50 M.Tech students and 15 MCA students for their project. Her research interest is on Image Processing, Data Structures, Software Engineering and Networking. She published 12 research papers in international journals. She published 11 research papers in international conferences.