Functional Expansions Based Multilayer Perceptron Neural Network for Classification Task

Umer Iqbal[1,2], Rozaida Ghazali[1], Muhammad Faheem Mushtaq[1], Afshan Kanwal[3]

[1] University Tun Hussein Onn Malaysia,
Faculty of Computer Science and Information Technology, Johor,
Malaysia

[2] University, Faisalabad Campus,
Pakistan Riphah College of Computing Riphah International,
Malaysia

[3] University Tun Hussein Onn Malaysia, Department of Mathematics and Statistics, Johor,
Malaysia

umeriqbal15@gmail.com, rozaida@uthm.edu.my

Abstract. Artificial neural network has been proved among the best tools in data mining for classification tasks. Where, Multilayer Perceptron (MLP) is known as benchmarked technique for classification tasks due to common use and easy implementation. Meanwhile, it is fail to make high combination of inputs from lower feature space to higher feature space. In this paper, Shifted Genocchi polynomials and Chebyshev Wavelets functional expansions based Multilayer Perceptron techniques with Levenberg Marquardt back propagation learning are proposed to deal with high dimension problems in classification tasks. Five datasets from UCI repository and KEEL datasets were collected to evaluate the performance in terms of five evaluation measures. T-test was applied to check the significance of the proposed techniques. The comparison results show that the proposed models outperform in terms of accuracy, sensitivity, specificity, precision and f-measure.

Keywords. Data mining, classification, shifted Genocchi polynomials, Chebyshev wavelets, multilayer perceptron.

1 Introduction

Among the different data mining tasks, classification is appeared as revolutionary task. Classification problems occur when we have to allocate an object in specific group or class on the base of their features or attributes. Classification task depends on two phases.

First phase is to construct the model, which consists of precogitated classes group. Second phase is to classify the unknown objects. The impact of classification task can be viewed in real life phenomena's such as stock exchange [1], marketing [2], leukemia classification [3], EMG classification [4] gait type classification on inertial sensors data [5] and health care data classification [6].

From last few decades, a variety of models have been developed for data mining. Statistical and artificial neural network models are prominent models. With the passage of time, artificial neural networks (ANNs) have gained much popularity as useful alternative of statistical techniques and due to having the variety of applications in real life [7]. MLP is type of ANNs, which consist of input, hidden and output nodes. In MLP each node is connected with other node in next layer to make connection between them. To train the MLP, different learning algorithms have been used with back propagation such as adaptive momentum to improve gradient descent accuracy [8] and Levenberg Marquardt for classification task. MLP was used to find missing values in data [9], fault detection in gearbox [10], pathological brain detection [11] and power quality disturbance [12].

MLP has large number of applications, meanwhile, MLP has also some drawbacks such

ISSN 1405-5546
doi: 10.13053/CyS-22-4-2602

as; firstly, multilayer structure causes increase in computational work by stuck in local minima. Secondly, it cannot be used for unsupervised learning. Thirdly, MLP was unable to make high combination of inputs to tackle nonlinear high dimensional problems. To overcome the nonlinear higher order dimensional problem, Chebyshev Multilayer Perceptron (CMLP) was proposed [13]. In this neural network, Chebyshev polynomials based functional expansion layer was introduced to confront high dimensional nonlinear problems. Chebyshev polynomials were used to make standard MLP an efficient tool to perform different types of data mining tasks. This network has been used for classification task.

In this paper, two functional expansions were introduced with MLP such as Shifted Genocchi polynomials and Chebyshev Wavelets. The reason behind using these expansions is well explained in section 2. Simulation results were compared with CMLP based on five evaluation measures.

The contributions made by this study are as follow:

- We proposed Shifted Genocchi Polynomials and Chebyshev Wavelets based Multilayer Perceptron for classification.
- The properties of Shifted Genocchi polynomials such as; firstly, less number of terms as compared to the shifted Chebyshev, Legendre and Chebyshev polynomials, which means that with increasing degree of polynomials, the number of terms also increases. Secondly, the coefficients of individual terms in shifted Genocchi polynomials are smaller than the coefficients of individual terms in the classical orthogonal polynomials. These properties encourage us to implement these polynomials as functional expansion.
- The properties of chebyshev wavelets such as; orthonormality, compact support and function approximation with different resolutions made this proposed model novel, where these wavelets are used as functional expansion.
- A comparative analysis of proposed models with CMLP was completed using five datasets. The performance of all models was verified in terms of five evaluation measures.

The remainder of this paper is organized as follows. Section 2 describes a brief about proposed Models. The experimental design used in this work is discussed in section 3. In section 4 we present the results and discussion. Finally, the conclusion is given in section 5.

2 Functional Expansions Based Multilayer Perceptron Neural Network: Proposed Models

This section describes about functional expansions based proposed models with their network structure.

2.1 Shifted Genocchi Multilayer Perceptron

Functional expansion plays a vital role to make high combination of inputs. These expansions are based on basis functions and selection of basis function is very important task, because basis functions are selected according to the nature of the problem. According to approximation theory, usually orthogonal polynomials are considered as good approximates such as Chebyshev orthogonal polynomial. Researchers have used different types of basis functions as functional expansion such as Chebyshev polynomials, Laguerre polynomials and Legendre polynomials [14].

We introduce shifted Genocchi (non-orthogonal) polynomial, which is better approximation property as compared to orthogonal polynomials due to certain characteristics. Firstly, Shifted Genocchi polynomials have less number of terms than the shifted Chebyshev, Legendre and Chebyshev polynomials, which means that with increasing degree of polynomials, the number of terms also increases. For example, on third order degree Shifted Genocchi have three numbers of terms and shifted Chebyshev have four numbers of terms leading to less computational work. Secondly, the coefficients of individual terms in shifted Genocchi polynomials are smaller than the coefficients of individual terms in the classical orthogonal polynomials.

Since the computational errors in the polynomial are related to the coefficient of individual terms, the computational errors are less by using shifted Genocchi polynomials. Motivated

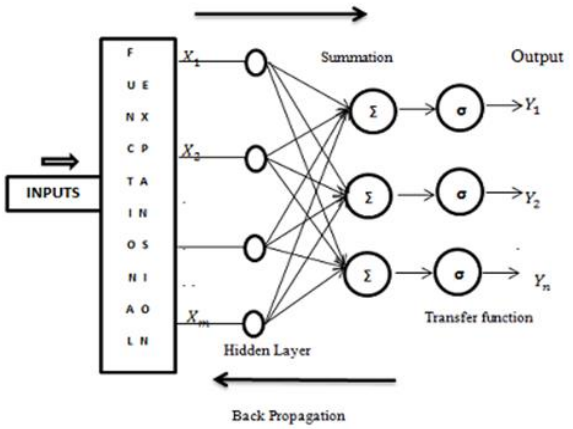

Fig. 1. Shifted Genocchi Multilayer Perceptron Neural Network

from above properties, Shifted Genocchi polynomials are proposed. The equation to derive the polynomials is given in Equation 1 as follows:

$$\begin{aligned} SG_0(t) &= 1, \\ SG_1(t) &= 2t - 3, \\ SG_r(t) &= \sum_{k=0}^{i} \sum_{r=0}^{k} \frac{(-1)^{i-r}\, i!\, G_{i-k}}{(i-k)!\,(k-r)!\,r!} t^r, \end{aligned} \tag{1}$$

where $SG_0(t)$, $SG_1(t)$ are the first and second polynomials respectively. $SG_i(t)$ is the analytical form of n^{th} shifted Genocchi polynomials and 'G' is the Genocchi number. The network structure is presented in Figure 1, where, $X_1, X_2 \ldots X_m$ are the inputs and $Y_1, Y_2, \ldots Y_n$ indicates the output of the neural network.

2.2 Chebyshev Wavelets Multilayer Perceptron

As discussed earlier that, basis functions have very important role as functional expansions to tackle the high dimensions nonlinear problems. Here, Chebyshev wavelets which are derived from Chebyshev polynomials are also used as functional expansion. These wavelets were derived from the dilation and translation of single function. The properties of orthonormality, compact support and due to functions approximation with different resolutions in Chebyshev wavelets made them better as compared to Chebyshev polynomials and Shifted Genocchi polynomials [15]. To understand the reason behind using wavelets, we have discussed three cases regarding to these properties.

Case 1. Orthonormality

Two vectors in an inner product space are orthonormal if they are orthogonal and unit vectors. In more simple way, "A set of vectors form an orthonormal set if all vectors in the set are mutually orthogonal and all of unit length". An orthonormal set which forms a basis is called an orthonormal basis. In case of functional expansion, the constant of expansion for the wavelets is more accurate due to orthonomality property as compared to polynomials, where constant of expansion is not more accurate due to orthogonality.

Case 2. Compact Support

In mathematics, compact support can be defined as, "A function has compact support if it is zero outside of a compact set". For example, set 'A' has a compact support means that it has support which is closed and bounded. On the other hand, the function $f: x \to x^2$ in its entire domain (i.e., $f: R \to R^+$) does not have compact support. In case of wavelets as functional expansion, wavelets with compact support have more advantage over that without compact support, because function approximation within the interval will be more accurate as compared to out of the interval.

Case 3. Function approximation with different resolution

In case of functional expansion, for wavelets we can maintain the degree of governing polynomial and increase the resolution by increasing values of k(integer). Therefore, we have the advantage of seeing the values of expansion at different intervels. For polynomials, just with the maintenance of degree, we cannot see the advantage of increase in accuracy.

Therefore, with the advantages of above three properties, we have proposed Chebyshev Wavelets based MLP (CWMLP) in this study. In this proposed method, these wavelets were used as functional expansion for more accurate classification. Chebyshev wavelets can be derived as following.

Chebyshev wavelets for piecewise polynomial spaces can be construct on interval [0, 1].

Some notations are needs to be introduce for this work, such as:

$N_0 \coloneqq \{0,1,\dots\}$, N denotes the natural numbers and $Z_\delta \coloneqq \{0,1,2,\dots,\delta-1\}$ for a positive integerδ.

For an integer $\delta > 1$, contractive mappings is considered on $I \coloneqq [0,1]$:

$$\psi_\upsilon(t) = \frac{t+\upsilon}{\delta}, t \in [0,1], \upsilon \in Z_\delta.$$

Mappings $\{\psi_\upsilon\}$ clearly fulfil the following properties:

$$\psi_\upsilon(I) \subset I, \ \forall \upsilon \in Z_\delta,$$

$$\bigcup_{\upsilon \in Z_\delta} \psi_\upsilon(I) = I.$$

Let G_0 is finite dimensional linear space on [0, 1], which is spanned by Chebyshev polynomials:

$$T_0\,(2x-1), T_1(2x-1), \dots, T_{M-1}(2x-1),$$

where $M \in N$ and $T_m\,(x)$ is the polynomial of m-order, namely:

$$G_0 \coloneqq span\{T_m(2x-1),/\ x \in [0,1]\ m \in\ Z_M\}.$$

It is known that Chebyshev polynomials $T_m\,(x)$ are orthogonal w.r.t weight function $w(x) = (1-x^2)^{\frac{-1}{2}}$ on [-1, 1]. Some Chebyshev polynomials are:

$$T_0(x) = 1,$$

$$T_{\ 1}(x) = x,$$

$$T_{m+1}(x) = 2x\,T_m(x) - T_{m-1}(x), \quad m = 1,2,3.$$

For construction of orthonormal base of $L^2[0,1]$, we define for each $\upsilon \in Z_\delta$ an isometry R_υ on $L^2[0,1]$:

$$R_\upsilon f)(x) \coloneqq \begin{cases} \sqrt{\delta f\left(\psi_\upsilon^{-1}(x)\right)}, & x \in \psi_\upsilon(I), \\ 0, & x \notin \psi_\upsilon(I). \end{cases}$$

Starting from G_0 space, we define the sequence of spaces $\{G_k\ k \in N_0\}$ via the recurrence relation:

$$G_{k+1} \coloneqq \bigoplus_{\upsilon \in Z_\delta} R_\upsilon\, G_k \qquad k \in N_0\,,$$

where

$$A \oplus B \coloneqq \{f+g : f \in A, g \in B\}.$$

Generally denotes the direct sum of A and B spaces:

$$G_0 \subset G_1 \subset \cdots \subset G_k \subset G_{k+1} \subset \cdots,$$

and

$$dim G_k = M\delta^k, \qquad k \in N_0,$$

$$\overline{\bigcup_{k=0}^{\infty} G_k} = L^2[0,1].$$

Now by constructing the orthonormal base for each space G_k:

$$H_0 \coloneqq \{\eta_m T_m(2x-1), \quad x\epsilon[0,1] | m \in Z_m\},$$

where:

$$\eta_m = \begin{cases} \sqrt{\frac{2}{\pi}} & m = 0, \\ \frac{2}{\pi} & m > 0\,, \end{cases} \tag{2}$$

is an orthonormal base of space G_0 and for all $f(x) \in L^2[0,1]$.

$$supp\{R_\upsilon f\} \cap supp\{R_{\upsilon'} f\} = \phi, \upsilon \neq \upsilon'\,,$$

$$\psi^{(k)}{}_{n,m}(x) = \begin{cases} \eta_m\, \delta^{\frac{k}{2}}\, T_m(2\delta^k x - 2n + 1), & \frac{n-1}{\delta^k} \le x < \frac{n}{\delta^k} \\ 0 & otherwise\,, \end{cases}$$

Then the set of $\{\psi^{(k)}{}_{n,m}(x) \,|n = 1,2,\dots,\delta^k, m \in Z_M\}$ forms an orthonormal base for space G_k w.r.t weight function $w_n(x)$ where:

$$w_n(x) = \begin{cases} w(2\delta^k x - 2n + 1), & \frac{n-1}{\delta^k} \le x < \frac{n}{\delta^k}, \\ 0 & otherwise. \end{cases}$$

For $\delta = 2, k = 1$ and $M = 3$, Chebyshev wavelets on interval [0, 1) are as follows:

$$\left.\begin{aligned}\psi_{1,0}^{(1)}(x) &= \tfrac{2}{\sqrt{\pi}},\\ \psi_{1,1}^{(1)}(x) &= 2\sqrt{\tfrac{2}{\pi}}(4x-1),\\ \psi_{1,2}^{(1)}(x) &= 2\sqrt{\tfrac{2}{\pi}}(1-16x+32x^2),\end{aligned}\right\} \quad 0 \le x < \tfrac{1}{2}$$

$$\left.\begin{aligned}\psi_{2,0}^{(1)}(x) &= \tfrac{2}{\sqrt{\pi}},\\ \psi_{2,1}^{(1)}(x) &= 2\sqrt{\tfrac{2}{\pi}}(4x-3),\\ \psi_{2,2}^{(1)}(x) &= 2\sqrt{\frac{2}{\pi}}(17-48x+32x^2).\end{aligned}\right\} \quad \tfrac{1}{2} \le x < 1$$

The network structure of Chebyshev Wavelets Multilayer Perceptron (CWMLP) is depicted in Figure 2, where, $X_1, X_2 \ldots X_m$ are the inputs and $Y_1, Y_2, \ldots Y_n$ indicates the output of the neural network.

The construction of orthonormal basis box was used to construct the wavelets because without constructing orthonormal basis wavelets cannot be derived. Functional expansion layer is the representation of high combination of inputs.

2.3 Correct / Wrong Classification Examples

To check that, whether the both proposed techniques such as SGMLP and CWMLP can classify the data correctly or not firstly we have considered the XOR binary dataset. The summary of the dataset is given in Table 1.

This dataset is too small to train and test the performance of proposed techniques therefore we have taken twenty times of this dataset for the experiments. The simulation results are shown in Table 2.

The simulation results of training, testing sets and all samples training results using SGMLP and CWMLP are summarized in Table 2. The results have shown that, SGMLP has performed 92.38 % and 94.16% accuracy on training and testing dataset respectively.

On the other hand, the performance of CWMLP was found 98.73% on testing set, which is more accurate on XOR binary dataset which means that, both models have ability to perform better on multi-class dataset due to addition of functional expansional layer for data Classification.

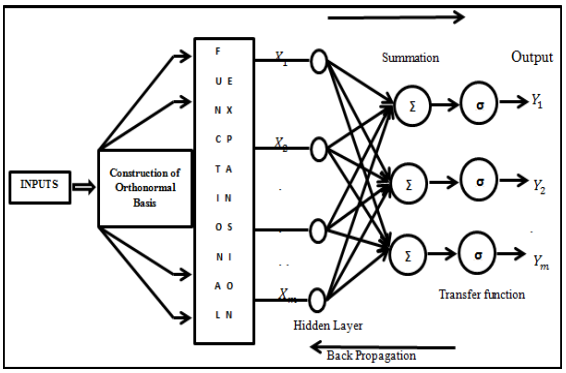

Fig. 2. Chebyshev Wavelets Multilayer Perceptron Neural Network

Table 1. XOR dataset

X_1	X_2	Output
0	0	0
0	1	1
1	0	1
1	1	0

Table 2. XOR dataset simulation results

XOR (Dataset)	SGMLP	CWMLP
Training set	92.38 %	95.97%
Testing set	94.16%	98.73%
When $X_1 = 0, X_2 = 0,$	100%	100%
When $X_1 = 0, X_2 = 1$	96.02%	98.89%
When $X_1 = 1, X_2 = 0$	100%	100%
When $X_1 = 1, X_2 = 1$	97.25%	98.46%
Average	93.27%	97.35%

Moreover, these simulation results have also shown that, both techniques have correctly classified the binary dataset.

2.4 Basic Definitions

This section comprises of some basic definitions that are related to this research work.

Table 3. Description of datasets

Datasets	No. of Samples	No. of Features	No. of Classes
Banana	5300	2	2
Titanic	2210	3	2
Ringnorm	7400	20	2
Breast Cancer	699	10	2
Bank Note Authentication	1372	5	2

Table 4. Description of Evaluation Measure

Evaluation Matrices	Mathematical Equations
Accuracy	$\frac{Tp+Tn}{Tp+Tn+Fp+Fn}$
Sensitivity	$\frac{Tp}{Tp+Fn} \times 100$
Specificity	$\frac{Tn}{Tn+Fp} \times 100$
Precision	$\frac{Tp}{Tp+Fp} \times 100$
F-Measure	$2 * \frac{precision.Sensitivity}{precision+Senstivity}$
Epochs	1000
Area Under the Curve	$\int_{\infty}^{-\infty} A_0(t)\, a_1(t)dt,$ AUC= $\int_{\infty}^{-\infty} A_0(A_1'(x))dx,$ here, $a_1(t) = \frac{dA_1(t)}{dt}$
Mean squared error	$\frac{1}{n}\sum_{i=0}^{n}(\hat{y}_i - y_i)^2$

Imbalanced Datasets: Imbalanced datasets are a special case for classification problem where the class distribution is not uniform among the classes. Typically, they are composed by two classes:

The majority (negative) class and the minority (positive) class.

Balanced Datasets: A balanced dataset is a set that contains all elements observed in all time frames.

Supervised Learning: All data is labeled and the algorithms **learn** to predict the output from the input data.

Unsupervised Learning: All data is unlabeled and the algorithms **learn** to deduce structure from the input data.

3 Experimental Design

This section explains step by step about experimental design of all techniques. The datasets, evaluation measures, data pre-processing and network topology of comparison and proposed techniques were discussed.

3.1 Data Collection

The data sets for classification analysis, which is the requisite input to the models, are obtained from UCI Repository[1] and KEEL datasets[2] [16, 17].

Here the UCI Machine Learning Repository is a collection of databases, domain theories, and data generators that are used by the machine learning community for the empirical analysis of machine learning algorithms and KEEL (Knowledge Extraction based on Evolutionary Learning) is an open source (GPLv3) Java software tool that can be used for a large number of different knowledge data discovery tasks.

We have collected five datasets namely, Ringnorm, Banana, Titanic, Breast Cancer and Bank Note Authentication datasets. Each dataset was divided into two parts; training set and testing set. The data ratio of 70% and 30% was set for training and testing respectively. The details of the used data sets are described in Table 3.

3.2 Evaluation Measures

The performance of comparison and proposed models is evaluated on the base of five evaluation measures. The formulae of evaluation measures is given in Table 4, where Tp, Tn, Fp and Fn are the true positive, true negative, false positive and false negative values respectively.

[1] http://archive.ics.uci.edu/ml datasets.html

[2] http://sci2s.ugr. es/keel/datasets.php

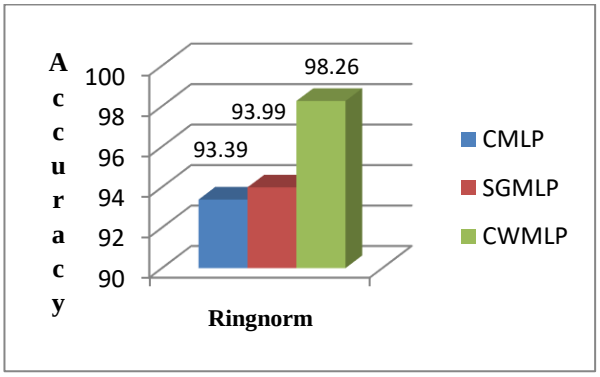

Fig. 3. Comparison in terms of accuracy in Ringnorm

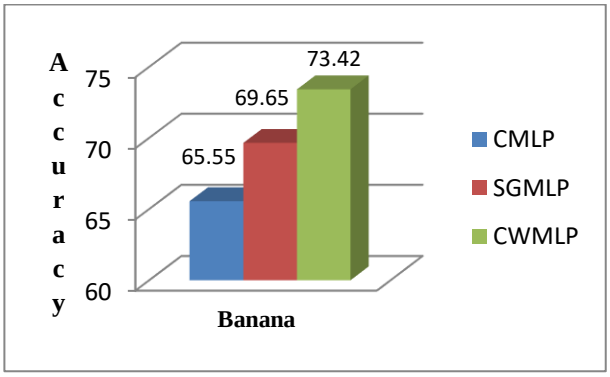

Fig. 4. Comparison in terms of accuracy in Banana

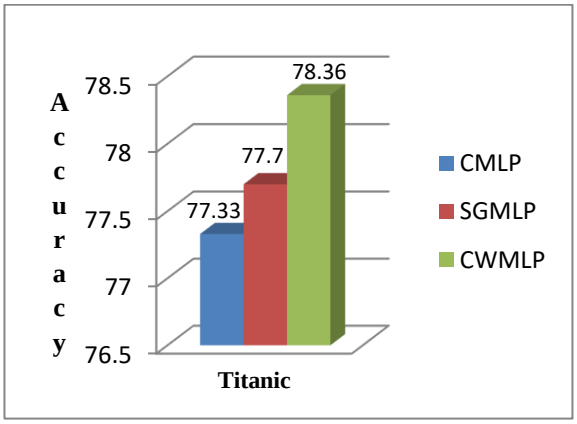

Fig. 5. Comparison in terms of accuracy in Titanic

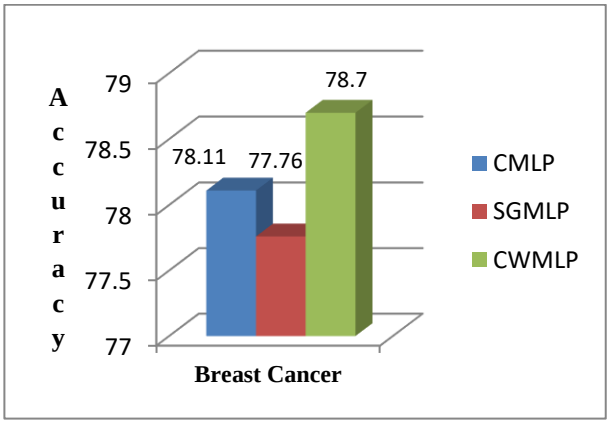

Fig. 6. Comparison in terms of accuracy in Breast Cancer

3.3 Data Pre-Processing

Data transformation is the process to normalize each data set into useful data. Data was normalized to the range [0.2, 0.8] and minimum and maximum normalization method was applied as given in Equation 2:

$$\hat{x} = (max_b - min_b) * \left(\frac{x - min_a}{max_a - min_a}\right) + min_b, \qquad (3)$$

where $\hat{x}$ is the normalize value of x. min_b and max_b are the minimum and maximum values of new range, and min_a and max_a are the minimum and maximum values of all observations.

3.4 Training and Network Topology

The proposed models topology of SGMLP and CWMLP is shown in Table 5. Settings were selected empirically.

Levenberg Marquardt (LM) back propagation was used as learning algorithm with all techniques.

4 Results and Discussion

This section consists of simulation results of comparison and proposed models. The experiments were performed 10 times and then average was taken to obtain verified results.

Table 5. Network Topology

Setting	Value
Activation Function	Sigmoid function
Genocchi Polynomial degree	3
Stopping Criteria	Maximum no. of epochs=1000
Learning rate	[0.1-0.3]
Momentum	[0.3-0.9]
Learning Algorithm	LM back propagation

Table 6. Comparison in terms of sensitivity, specificity and precision

Datasets	Techniques	Sensitivity	Specificity	Precision
Ringnorm	CMLP	93.41	93.41	93.41
	SGMLP	**94.79**	**94.79**	**94.79**
	CWMLP	**98.26**	**98.26**	**98.26**
Banana	CMLP	71.97	71.97	80.09
	SGMLP	**72.54**	**72.54**	**80.71**
	CWMLP	**77.08**	**77.08**	78.95
Titanic	CMLP	80.09	80.09	80.09
	SGMLP	**80.65**	**80.71**	**80.49**
	CWMLP	**80.77**	**80.79**	**80.77**
Breast Cancer	CMLP	74.88	72.13	70.93
	SGMLP	**76.25**	**76.25**	**76.25**
	CWMLP	**77.32**	**77.22**	**77.02**
Bank Note Authentication	CMLP	90	90	90
	SGMLP	**100**	**100**	**100**
	CWMLP	**100**	**100**	**100**

1000 number of iterations was taken in all experiments. Levenberg Marquardt back propagation was used as learning algorithm in all models. Classification accuracy of all models in all datasets was shown in bar graph from Figure 3 to Figure 7.

The accuracy is used to check the overall effectiveness of the techniques. It can be seen that, CWMLP performance was better on Ringnorm and Banana datasets, whereas, CWMLP perform slightly better on Titanic dataset as compared to SGMLP and CMLP.

The reason behind these significant results is that, the proposed solution helps CWMLP to raise and find more appropriate settings during the training which helps to enhance the classification performance for network. This better performance is also due the reason that, Chebyshev wavelets can generate more numbers of basis functions using same degree as compared to Shifted Genocchi polynomials based MLP technique that help to generate more enhanced values with small coefficients and reduced the computational task. Meanwhile, it seems that SGMLP was not fully

supported by parameter settings that cause less improvement with CWMLP.

In case of Brest Cancer dataset, SGMLP performance was lower than the CMLP, but CWMLP perform better than both of the techniques with 78.70% classification accuracy. This is due to the reason that, sometime due to the imbalanced dataset, techniques are not able to classify the data correctly. In Bank Note Authentication dataset, CWMLP and SGMLP performance was 100% to classify the data. The same parameters setting and number of epochs are the reason behind the same results on the other hand CMLP performance was found to be 90%. Over all, it can be seen that CWMLP performance was much better than the SGMLP and CMLP on all datasets in terms of classification accuracy.

The best performance with CMLP, SGMLP and CWMLP using the Banana, Titanic, Ringnorm, Breast Cancer and Bank Note Authentication datasets in terms of Sensitivity, Specificity and Precision is shown in Table 6. As it can be noticed from Table 6, CWMLP identify the proportion of positive values (Sensitivity), proportion of negative values (specificity) and selected items, which are relevant (Precision) correctly with higher performance on Bank note Authentication and Ringnorm datasets.

In Titanic dataset, there is slightly difference between all comparison techniques. In Breast Cancer dataset, SGMLP perform well as compared to CMLP, because Shifted Genocchi can generate more better enhanced inputs due to small coefficient value and less degree, similarly, CWMLP perform slightly high than the SGMLP in terms of Sensitivity, Specificity, and Precision. Over all, it can be seen that CWMLP outperform on all datasets in comparison of SGMLP and CMLP. Bold font was used to prominent the proposed techniques results.

Test's accuracy was measured by F-Measure. Precision and Sensitivity were used to compute the f1-score. The higher value shows its best, lower value shows its worst. In the Figure 8 bar graph represents the F-Measure results of all the comparison techniques. In Bank Note Authentication, SGMLP and CWMLP gives their best with 100% f1-score as compared to CMLP. After that, in Ringnorm dataset CWMLP perform

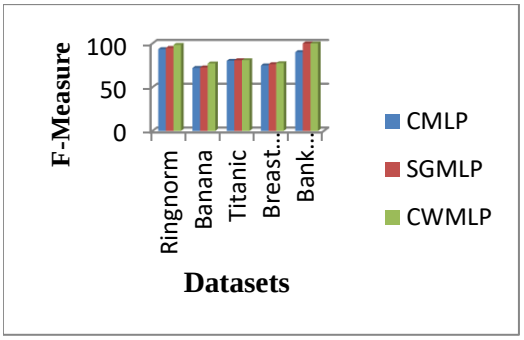

Fig. 8. Comparison in terms of f-measure on all datasets

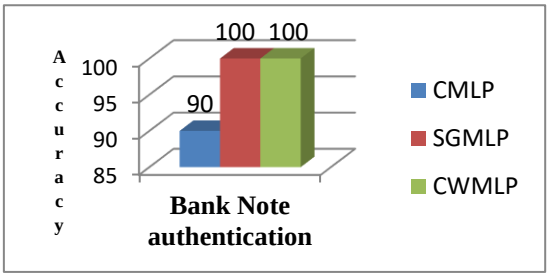

Fig. 7. Comparison in terms of accuracy in BankNote Authentication

well with 98.26 % as compare to SGMLP and CMLP.

In Titanic dataset all the comparison techniques score was slightly different from each other. From results we can conclude that, CWMLP performance was much better than the other techniques in terms of F- Measure. On the whole, it can viewed that the performance of CWMLP is much better due to their property of generating more number of small value coefficient based wavelets as compared to other basis functions. Whereas, the enhanced inputs generated by Chebyshev polynomials are not able to reduce the computational complexity of the network.

4.1 Significance Using T-test

To check that how much significant are our proposed techniques, we have applied paired two samples for means t-test in relation of accuracy. We found that, in all datasets SGMLP has shown significance except in Ringnorm dataset when

compared with CMLP. Insignificance has caused due to approximately similar high accuracy of both proposed and comparison techniques. In Breast Cancer dataset, SGMLP again remain insignificant due to approximately similar accuracy when compared with CWMLP. In rest of datasets, CMLP remained insignificant in all datasets. The hypothetic value was taken 0.05. If we achieved 'value < 0.05' then called as significant and 'value >0.05' then called as insignificant.

5 Conclusion

In this paper, two functional expansions based multilayer perceptron models were proposed to tackle higher dimension nonlinear problems. The concept of adding new functional expansion layer in MLP using Shifted Genocchi polynomials and Chebyshev wavelets were proved much better. It is due to reason that, the number of enhanced inputs generated by these polynomials and wavelets are small in values and less computational as compared to CMLP. Moreover, Chebyshev wavelets have ability to produce more number of basis functions on the same degree as compared to Shifted Genocchi and Chebyshev polynomials. Which is helpful to increase the accuracy of the classifier.

CWMLP and SGMLP were experimentally trained and tested on five benchmarked data sets, which were taken from UCI repository and KEEL datasets. The performance of the proposed models indicates their validity for classification task. The evaluation measures performance show that the SGMLP and CWMLP has better performance in terms of accuracy, sensitivity, specificity, precision and F-Measure over CMLP. Overall, CWMLP performance was outstanding as compared to rest of the techniques.

Similarly, t-test clearly validates the significance of proposed techniques over the benchmarked approaches.

References

1. **Bebarta, D. K., Rout, A. K., Biswal, B., & Dash, P. K. (2012).** Forecasting and classification of Indian stocks using different polynomial functional link artificial neural networks. *Annual IEEE India Conference (INDICON´12)*, pp. 178–182. DOI: 10.1109/INDCON.2012.6420611.
2. **Decker, R. & Kroll, F. (2007).** Classification in marketing research by means of LEM2-generated rules. *Advances in Data Analysis,* pp. 425–432 DOI: 10.1007/978-3-540-70981-7_48.
3. **Lim, H.N., Francis, E.U., Mashor, M.Y., & Hassan, R. (2016).** Classification of bone marrow acute leukemia cells using multilayer perceptron network. *International Conference Electronic Design (ICED´16),* pp. 486–490. DOI: 10.11 09/ICED.2016.7804693.
4. **Artameeyanant, P., Sultornsanee, S., & Chamnongthai, K. (2015).** Classification of electromyogram using weight visibility algorithm with multilayer perceptron neural network. *7th International Conference on Knowledge and Smart Technology (KST´15),* pp. 190–194. DOI:10.1109/ KST.2015.7051485.
5. **Szczepański, D. (2016).** Multilayer perceptron for gait type classification based on inertial sensors data. *Federated Conference on Computer Science and Information Systems (FedCSIS),* pp. 947–950. DOI: 10.15439/2016F265.
6. **Naraei, P., Abhari, A., & Sadeghian, A. (2016).** Application of multilayer perceptron neural networks and support vector machines in classification of healthcare data. *Future Technologies Conference (FTC)*, pp. 848–852. DOI: 10.1109/FTC.2016. 7821702.
7. **Paliwal, M. & Kumar, U.A. (2009).** Neural networks and statistical techniques: A review of applications. *Expert systems with applications*, Vol. 36, No. 1, pp. 2–17. DOI: 10.1016/j.eswa.2007. 10.005.
8. **Nawi, N.M., Ransing, R.S., Salleh, M.N.M., Ghazali, R., & Hamid, N.A. (2010).** An improved back propagation neural network algorithm on classification problems. *Database Theory and Application, Bio-Science and Bio-Technology*, pp. 177–188. DOI: 10.1007/978-3-642-17622-7_18.
9. **Silva-Ramírez, E.L., Pino-Mejías, R., & López-Coello, M. (2015).** Single imputation with multilayer perceptron and multiple imputation combining multilayer perceptron and k-nearest neighbours for monotone patterns. *Applied Soft Computing*, Vol. 29, pp. 65–74. DOI: 10.1016/j.asoc.2014.09.052.
10. **Jedliński, Ł. & Jonak, J. (2015).** Early fault detection in gearboxes based on support vector machines and multilayer perceptron with a continuous wavelet transform. *Applied Soft Computing*, Vol. 30, pp. 636–641. DOI: 10.1016/j.asoc.2015.02.015.

11. **Zhang, Y., Sun, Y., Phillips, P., Liu, G., Zhou, X., & Wang, S. (2016).** A multilayer perceptron based smart pathological brain detection system by fractional Fourier entropy. *Journal of medical systems*, Vol. 40, No. 7, pp. 1–11. DOI: 10.1007/s10916-016-0525-2.
12. **Borges, F.A., Fernandes, R.A., Silva, I.N., & Silva, C.B. (2016).** Feature extraction and power quality disturbances classification using smart meters signals. *IEEE Transactions on Industrial Informatics*, Vol. 12, No. 2, pp. 824–833. DOI:10.1109/ TII.2015.2486379.
13. **Iqbal, U., & Ghazali, R. (2016).** Chebyshev Multilayer Perceptron Neural Network with Levenberg Marquardt-Back Propagation Learning for Classification Tasks. *International Conference on Soft Computing and Data Mining,* pp. 162–170. DOI: 10.1007/978-3-319-51281-5_17.
14. **Mall, S. & Chakraverty, S. (2015).** Multi-layer versus functional link single layer neural network for solving nonlinear singular initial value problems. *Proceedings of the third international symposium on women in computing and informatics* pp. 678–683. (ACM). DOI: 10.1145/2791405.2791542.
15. **Aminikhah, H. & Hosseini, S. (2015).** Numerical solution of nonlinear system of integro-differential equations using Chebyshev wavelets. *Journal of Applied Mathematics, Statistics and Informatics*, Vol. 11, No. 2, pp. 15–34. DOI: 10.1515/jamsi-2015-0009
16. **Blake, C. & Merz, C.J. (1998).** Repository of machine learning databases.
17. **Alcalá, J., Fernández, A., Luengo, J., Derrac, J., García, S., Sánchez, L., & Herrera, F. (2010).** Keel data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework. *Journal of Multiple-Valued Logic and Soft Computing*, Vol. 17, No. 2-3, pp. 255–287.

Article received on 02/04/2017; accepted on 28/02/2018.
Corresponding author is Umer Iqbal.