An Optimal Deep Neural Network Model for Lymph Disease Identification And Classification

J. Junia Deborah, Dr. Latha Parthiban

Abstract— At present days, deep learning (DL) models find useful in various domains especially healthcare sector. This paper introduces a new DL based lymph disease identification and classification model. An optimal deep neural network (DNN) model is applied to classify the lymph data utilizing the stacked autoencoders (SA) which is generally used to extract the features from the dataset which is further classified by the utilization of SoftMax layer. In addition, to make the DNN model more efficient, DL based Adam optimizer (AO) is used which is an adaptive learning rate optimization algorithm which has been developed particularly to train DNN. The presented DNN with AO model called DNN-AO model is tested against benchmark Lymphography dataset and the results are assessed under diverse measures. The experimental outcome verified that effective classification performance is showcased by the DNN-AO model with the maximum sensitivity, specificity, accuracy and kappa value of 95.59, 95.95, 95.95 and 91.69 respectively.

Index Terms— Lymph disease, Classification, Deep Learning, Adam Optimizer

1 INTRODUCTION

In last decades, Computer Aided Diagnosis (CAD) techniques have been applied widely in medical domain. The Medicinal analysis is referred as subjective as it is based on doctor’s practical knowledge regarding the available information. For doctor’s diagnosing purpose [1, 2], it urges that system conversion would grasp the portions of solutions. In order to create CAD systems, a novel method named as Machine Learning (ML) has been proposed. The ML models are most significant and applied in several domains since it is highly capable of retrieving complex association of data in biomedical region [3, 4]. The data analysis could be performed by applying maximum amount of medicinal information by integrating with various applicable techniques in classification process. Therefore, it assumed to be a major challenging issue in accuracy where classification model has been used in examining the anomalies of individual. Here, diverse type of information from medical field is composed with maximum dimensionality [5, 6]. Generally, high definition data requires the mining process with descriptive features which must be selected whereas dataset directional should be declined [7]. For avoiding the irrelevant attributes from dataset [8, 9], reducing dimension might be considered as major action while analyzing the system. By removing the number of unwanted attributes tends to execute a satisfied technique where the screen is sampled in a rapid manner with lower expense. Additionally, there is a point of enhancing the accuracy from diagnosing process, the current research work focus on determining best feature subset for lymphography dataset.

In general, lymphatic system reinforce immune system through the elimination as well as by degrading the cancer cells, rejecting unwanted sources, pathogens, dead blood cells, and debris. It helps to assimilate the fat-soluble as well as fat vitamins from digestive system and provide to body tissues. Additionally, the interstitial spaces among cells remove the unnecessary substance and extra fluids from body. Here, lymphatic system is fully composed with lymph nodes, 2 collecting ducts [10] and thin-walled lymphatic vessels. Along with Lymph vessels, circulatory system vessels are connected. Larger lymph vessels would be similar as veins. From the whole human body, the Lymph capillaries could be distributed. Moreover, the skeletal muscle contraction simulates the movement of valves in lymph fluid. Also, Lymph nodes differ in dimension that resembles the structure of a kidney. From all over the human body, it is scattered from abdomen, neck, armpit, groin as well as pelvis.

The Lymph nodes would have post of T, B and alternate immune tissues. Furthermore, the partial amount of lymphatic cells as well as lymph nodes might be present nearby gastro intestinal tract. Before getting back to circulatory system, the node has to filter the lymph. However, it is reduced; if the node gets maximized then, all nodes would be damaged and leads to failure of simulating fluids. With the application of lymphography medical imaging techniques [11], the present condition of lymphatic system could be predicted. In non-invasive analysis of lymph node, magnetic resonance lymphography has been employed. Whenever there is a diagnosis of abnormal lymph nodes [12], this model applies tiny super paramagnetic components of iron oxide which has extreme sensitivity and specificity. Thus, the current state of lymph nodes along with retrieved information from lymphography model is capable of computing the accuracy in classification process [13]. Therefore, swelling of lymph node is a key for complex scenario and expand the lifetime of which is in threat or risk [14]. Also, the state of lymph node helps to predict the cancer disease [9]. From the current study it has been revealed that, Lymph disease prediction might be a predominant operation. In order to resolve the above complexities, a novel hybrid classifying model has been presented by Polat and Gunes [15] which depends upon C4.5 Decision Tree (DT) classification as well as one-against-all approach for classifying multiple problems that is involved in segmenting images where lymphography and dermatology dataset have been acquired from UCI ML database. The C4.5 DT basically runs across the overall dataset to gain a maximum accuracy for dividing images, dermatology and lymph dataset applies 10-fold cross validation(CV). In such cases, the proposed technique provides most optimal accuracy in classification process. Also, the multiclass classification involves in lymph disease set, Iannello et al. [16] deployed a decomposing model such as Error-Correcting Output Codes (ECOC), Pair Wise Coupling(PWC) as well as One-per-Class (OpC). By the application of different techniques, a general classification method like kernel machine which is said to be a
The common AE is depicted in Fig. 2. Generally, input layer comprises of D neurons, hidden layer has C neurons and D is assumed to be the estimation of input vector and C is evaluation of code vector. Thus, AE consist of two phases like encoding and decoding as illustrated in Fig. 2. The precedent of AE is provided as input for encoding and the final result has been offered to hidden layer of AE. While approaching the potentials of input vector, it has been transferred as code with the application of encoding technique. Also, input of decoding would be the simulation outcome of hidden layer and the final outcome attained from decoder can be declared as output of AE. Hence, decoding operation is provided to reform new input vector from the parameters of encoding method. Therefore, mapping of incoming result from encoder would be offered along with

\[ c = f(Wz p + b) \]  

where \( W = [w_1, w_2, \ldots, w_C] \) denotes the weight matrix, \( Wl = [w_{l1}, w_{l2}, \ldots, w_{lc}] \) implies weight of attached neurons \( l \) to \( C \) neurons, from hidden layer, \( p = [p_1, p_2, p_3]^T \) represents the input vector in \( p_l \) interms of all variables, \( b = [b_1, b_2, b_3]^T \) indicates the biasing vector as \( bl \) is corresponding to bias linked neurons, \( c = [c_1, c_2, c_3]^T \) specifies the code in \( c_l \) that states code formed by neuron \( l \) is hidden layer and \( f \) present in activating process of neurons. Subsequently, mapping of input from output which is obtained from decoder could be offered with

\[ p' = f(W'z c + b') \]  

where \( W = [w_1', w_2', w_c'] \) signifies matrix weight which computes the weight of linked hidden layer of networks, \( p' = [p_1', p_2', p_d]^T \) indicates reformed input vector from \( p' \) which operates all attributes of input vector, \( b = [b_1, b_2, b_c]^T \) represents the biasing vector, such that \( b_1 \) denotes bias association from all neurons whereas \( c = [c_1, c_2, c_C]^T \) implies the feature built using encoder as well as \( f \) resembles the activation process of neurons. In general, mapping process of encoding can be represented by

\[ c = f_e(W, b; p) \]  

where \( f_e \) specifies encoding process and mapping function of decoding might be projected as,

\[ p' = f_d(W', b'; c) \]
where $f_D$ denotes the decoding process in AE. It is formed with the help of layers of encoding and decoding from the linear model as depicted in Fig. 2.

![Fig. 2. General AE structure.](image)

### 2.2 DNN based classification

For promoting the significant codes of deep networks, the DNN based architecture which applies AE that has been projected for applying the CS which helps to improve the evaluation value of classification issue. Hence, DNN classifying process from CS techniques are built with the application of SA as well as softmax layer that is assumed from primary stage. Furthermore, some of the datasets comprises of 8 parameters and class variable which is declared as linear segmentation. Thus, 8 features are induced as instance of input layer. Additionally, DNN forms a stack of two layers in AE where networks are composed with 2 hidden layers from all neurons. Hence, softmax layers has been linked to consequent hidden layer of classification process. Therefore, final resultant layer would offer the feasibility of credit approval (CA) that is also known as non-credit approval (NCA) respectively.

![Fig. 3. SA with L layers.](image)

### Training process of layers

Let $N$ be the input vectors that is filled for training AE as $\{p_{(1)}, p_{(2)}, \ldots, p_{(N)}\}$. The recreation of input would be fully fledged along with training AE is provided as

$$p' = f_o(W', b'; f_e(W, b; p))$$

(5)

This could be exposed as,

$$p' = f_{AE}(W, b, W', b'; p)$$

(6)

where $f_{AE}$ is the performance which is mapped with input to output of AE. Thus, AE is trained with adaptive function which is offered with fault function as,

$$E_{Tot} = E_{MSE} + E_{Reg} + E_{spars}$$

(7)

where $E_{MSE}$, $E_{spars}$, $E_{Reg}$ denotes mean square error, sparsity factor and regularizing value where mean square error $E_{MSE}$ could be determined by using

$$E_{MSE} = \frac{1}{N} \sum_{i=1}^{N} e_i^2$$

(8)

where $e_i$ indicates an error which is attained from certain output, $p_{(i)}$ as well as the predicted result is $p'_{(i)}$. Hence, fault $e_i$ is operated with the help of,

$$e_i = \|p_{(i)} - p'_{(i)}\|$$

(9)

A DNN is capable of learning every point from samples database that tends in over fitting technique. It is a major challenge in DNN, since the result are based on the primary sampled data. In order to resolve this problem, regularization factor, $E_{Reg}$ is considered from objective process which can be determined by the consumption of,

$$E_{Reg} = \lambda \left( \sum_{l=1}^{L} \|w_l\| + \sum_{l=1}^{L} \|w'_l\| \right)$$

(10)

where $\lambda$ represent the period of regularizing process. The sparsity allows a method to learn appeal codes which is obtained from current data. Also, sparsity factor $E_{spars}$ is processed with the help of,

$$E_{spars} = \beta \sum_{m=1}^{c} KL(p||p_m)$$

(11)

where $\beta$ implies the sparsity weight and $KL(p||p_m)$ signifies Kullback-Leibler distinction that is given by,

$$KL(p||p_m) = p \log \frac{p}{p_m} + (1-p) \left( \frac{1-p}{1-p_m} \right)$$

(12)

where sparsity is given by $p$ and $p_m$ implies the standard activation metric of $m$th neurons which is calculated by the application of

$$p_m = \frac{1}{Z} \sum_{i=1}^{Z} f^m(p_{(i)})$$

(13)

where $f^m(p_{(i)})$ denotes activate process of $m$th neurons forming hidden layer in AE.

### Layers in stacked auto-encoder

A DNN which applies AE has been formed along with cascade of encoding layers that is shown in Fig. 3. By understanding the mapping process of AE in Eq. (6), it would be written as

$$f_{AE} = f^1_{AE} \circ f^2_{AE} \circ \ldots \circ f^N_{AE}$$

(14)

where the linear AE process can be resembled as $f_{SAE}$. For each layer of stacked AE, encoding performance has been implemented. It is significant to state that decoding process could not be compiled for all layers [17].
Fig. 4. (a) Network of auto-encoder 1 (b) Network of auto-encoder 2 (c) Softmax classifier (d) DNN [17]
Softmax classification layer

The Softmax classification is considered as multi-class segmenting that applies Logistic Regression (LR) which is helpful in dividing the data. A Softmax layer uses controlled learning method which applies in LR in multiclass classifying process. Therefore, LR depends upon softmax classification operation. The multi class division issue is linked with softmax classification that measures the feasibility of all classes. Here, softmax process is useful to normalize the operation and exponential model helps to define the probable classes. Thus, softmax layer is associated with SA by using the notation of \( f_{sc} \) correspondingly.

2.3 Optimal Tuning Process

Here, Adam algorithm is described with corresponding features. If \( f(\theta) \) signifies a noisy function then, a stochastic scalar function could be applied as: 

\[
\Delta_{s} = a \cdot \frac{\sqrt{v_{t}}}{\sqrt{m_{t}}}.
\]

Then it captures the collection of 2 upper limits: \(|\Delta_{s}| \leq \alpha \cdot \sqrt{1-\beta_{1}^{2}}\), \(|\Delta_{s}| > \sqrt{1-\beta_{2}^{2}}\), as well as \(|\Delta_{s}| \leq \alpha\). Hence, primary case has been conducted in several sparsity cases, when a gradient is assumed to be zero for all execution process by leaving current state process. The lower sparse case consist of effective phase which is lower in size. If \(1-\beta_{1} = \sqrt{1-\beta_{2}}\), then, it might be mentioned as 

\[
\frac{m_{t}}{\sqrt{v_{t}}} < 1, \text{ so } |\Delta_{s}| < \alpha.
\]

An effective magnitude of above steps are filled with parameter space where it is limited with sample values \(a.i.e., |\Delta_{s}| \leq \alpha\). Then, Adam utilizes the of initialization bias correcting terms.

2.4 Training Process of DNN Classifier

The DNN classification is provided along with SA under the application of softmax classifying method. The SA has 2 more AES. Fig. 4(d) depicts DNN classification with the help of SA that is comprised with 2 AE. If input vector is induced as input for DNN \( \{x_{(1)}, x_{(2)}, ..., x_{(D)}\} \) and similar results of variables might be \( \{y_{(1)}, y_{(2)}, ..., y_{(N)}\} \). By using trained input vectors, it mainly aims in training the method which is helpful in tuning parameters of DNN in learning input vectors and division of identical result with maximum accuracy.

Consequently, it is then applied to train DNN classification as follows:

- Initially, primary AE layers undergo training with the help of new input vector \( \{x_{(1)}, x_{(2)}, ..., x_{(D)}\} \) which is identical to destination vector. It mainly aims to reform the input by eliminating the codes \( \{c_{(1,1)}, c_{(1,2)}, ..., c_{(1,C)}\} \) through deployment of AE as illustrated in Fig. 4. (a).
- Alternate AE is trained using the final vector of basic AE, \( \{c_{(1,1)}, c_{(1,2)}, ..., c_{(1,C)}\} \) since input vector and forms a resultant vector \( \{c_{(2,1)}, c_{(2,2)}, ..., c_{(2,R)}\} \).
- Second AE tries to recreate the input of \( c_{(1,i)}, l = 1, 2, ..., R \) that is expressed in Fig.4. (b).
- The stack AE is reduced with the help of softmax classification. Thus, it is trained by holding the next AE’s result, \( c_{(2,i)}, l = 1, 2, ..., R \) since input vector and new class variables \( \{y_{(1)}, y_{(2)}, ..., y_{(N)}\} \) is declared as destination vectors which is obtained from learned data. Therefore, softmax classification is depicted in Fig.4. (c).
- Finally, the improvement of classification from DNN, ADAM is applied for fine-tuning that is considered to be a generalized technique along with networks which is by trained data.

3 PERFORMANCE VALIDATION

To validate the effectual performance of the presented DNN-AO model, a series of simulation takes place on benchmark dataset and the results are compared with several measures.

3.1. Dataset

For experimentation, lymphography database [18] from the University Medical Centre, Oncology Institute is employed. The dataset includes a collection of 148 samples without any missing value. A set of 4 assigned class labels exist namely malign lymph, fibrosis, metastases and normal with a total of 18 numeric attributes. The information related to the dataset is shown in Table 1. The frequency distribution of all the existing 18 attributes is shown in Fig. 5.
### TABLE1 Dataset Description

<table>
<thead>
<tr>
<th>Attribute Number</th>
<th>Attribute description</th>
<th>Possible values of attributes</th>
<th>Assigned values</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lymphatic</td>
<td>Normal = 1, arched = 2, deformed = 3, displaced = 4</td>
<td>1-4</td>
<td>2.74</td>
<td>0.82</td>
</tr>
<tr>
<td>2</td>
<td>Block of afferent</td>
<td>No, Yes</td>
<td>1-2</td>
<td>1.55</td>
<td>0.50</td>
</tr>
<tr>
<td>3</td>
<td>Block of lymph c (superior and inferior flaps)</td>
<td>No, Yes</td>
<td>1-2</td>
<td>1.17</td>
<td>0.38</td>
</tr>
<tr>
<td>4</td>
<td>Block of lymph s (lazy incision)</td>
<td>No, Yes</td>
<td>1-2</td>
<td>1.04</td>
<td>0.21</td>
</tr>
<tr>
<td>5</td>
<td>By pass</td>
<td>No, Yes</td>
<td>1-2</td>
<td>1.24</td>
<td>0.43</td>
</tr>
<tr>
<td>6</td>
<td>Extravasates (force out of lymph)</td>
<td>No, Yes</td>
<td>1-2</td>
<td>1.51</td>
<td>0.50</td>
</tr>
<tr>
<td>7</td>
<td>Regeneration</td>
<td>No, Yes</td>
<td>1-2</td>
<td>1.07</td>
<td>0.25</td>
</tr>
<tr>
<td>8</td>
<td>Early uptake</td>
<td>No, Yes</td>
<td>1-2</td>
<td>1.70</td>
<td>0.46</td>
</tr>
<tr>
<td>9</td>
<td>Lymph nodes diminish</td>
<td>0-3</td>
<td>1-2</td>
<td>1.06</td>
<td>0.31</td>
</tr>
<tr>
<td>10</td>
<td>Lymph nodes enlarge</td>
<td>1-4</td>
<td>1-4</td>
<td>2.47</td>
<td>0.84</td>
</tr>
<tr>
<td>11</td>
<td>Changes in lymph</td>
<td>Bean = 1, oval = 2, round =3</td>
<td>1-3</td>
<td>2.40</td>
<td>0.57</td>
</tr>
<tr>
<td>12</td>
<td>Defect in node</td>
<td>No = 1, lacunar = 2, lacunar marginal = 3, lacunar central = 4</td>
<td>1-4</td>
<td>2.97</td>
<td>0.87</td>
</tr>
<tr>
<td>13</td>
<td>Changes in node</td>
<td>No, lacunar, lacunar marginal, lacunar central</td>
<td>1-4</td>
<td>2.80</td>
<td>0.76</td>
</tr>
<tr>
<td>14</td>
<td>Changes in structure</td>
<td>No, grainy, drop-like, coarse, diluted, reticular, stripped, faint</td>
<td>1-8</td>
<td>5.22</td>
<td>2.17</td>
</tr>
<tr>
<td>15</td>
<td>Special forms</td>
<td>No, Chalices, vesicles</td>
<td>1-3</td>
<td>2.33</td>
<td>0.77</td>
</tr>
<tr>
<td>16</td>
<td>Dislocation</td>
<td>No, Yes</td>
<td>1-2</td>
<td>1.67</td>
<td>0.48</td>
</tr>
<tr>
<td>17</td>
<td>Exclusion of node</td>
<td>No, Yes</td>
<td>1-2</td>
<td>1.80</td>
<td>0.41</td>
</tr>
<tr>
<td>18</td>
<td>Number of nodes</td>
<td>0-80</td>
<td>1-8</td>
<td>2.60</td>
<td>1.91</td>
</tr>
<tr>
<td>19</td>
<td>Target Class</td>
<td>Normal = 0, Metastases = 1, Malign Lymph = 2, Fibrosis = 3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 5. Sample Frequency distribution
3.2. Results analysis

Table 2 displays the results attained by the presented DNN-AO model during the classification of lymphography dataset. While classifying the data under the class normal, i.e. class 0, the DNN-AO model achieves effective classifier outcome with the maximum accuracy of 100, sensitivity of 100, specificity of 100, precision of 100, MCC of 100, F-measure of 100, AUC of 100 and kappa value of 100 respectively. Similarly, while classifying the data under the class metastases, i.e. class 1, the DNN-AO model achieves effective classifier outcome with the maximum accuracy of 91.89, sensitivity of 93.82, specificity of 89.55, precision of 91.56, MCC of 83, F-measure of 92.68, AUC of 91.69, and kappa value of 83.59 respectively. While classifying the data under the class Malign Lymph, i.e. class 2, the DNN-AO model achieves effective classifier outcome with the maximum accuracy of 91.89, sensitivity of 88.52, specificity of 90.55, precision of 91.56, MCC of 83, F-measure of 92.68, AUC of 91.69 and kappa value of 83.59 respectively. Similarly, while classifying the data under the class Fibrosis, i.e. class 3, the DNN-AO model achieves effective classifier outcome with the maximum accuracy of 100, sensitivity of 100, specificity of 100, MCC of 100, F-measure of 100, AUC of 100 and kappa value of 100 respectively.

Table 3 shows the average classifier results of the DNN-AO model on the applied dataset. The table values pointed out that the DNN-AO model classifies the data with the accuracy of 95.95, sensitivity of 95.95, specificity of 95.95, precision of 95.77, MCC of 91.50, F-measure of 95.67 and kappa value of 91.69 respectively.

An extensive comparative analysis of the DNN-AO with the existing models takes place under diverse measures. The attained comparison results are tabulated in Table 4 and Figs. 6-7. On measuring the results in terms of accuracy, it is noticed that ineffective classifier outcome is exhibited by the Fisher RF model with the least accuracy of 75.50. Next to that, slightly effective classifier results are attained by the SFFS-RF model with slightly higher accuracy of 77.10. Then, the No-FS reaches to a nearly equivalent and moderate accuracy of 80.40 and 81.80 respectively. At the same time, the PCA-RF and Relief-F-RF models outperforms the previous methods and obtained near identical accuracy values of 83.90 and 83.40 respectively. Simultaneously, the Relief-F-RF model shows manageable results with the accuracy value of 84.20 whereas the GA-RF model offers good classifier outcome with the high accuracy of 92.20. It is also noted that the PSO-RF model showcased outperforming results over the compared methods except DNN-AO model with the accuracy of 94.00. At last, the DNN-AO model reaches to a maximum classification performance with the utmost accuracy of 95.95.

On measuring the results in terms of sensitivity, it is noticed that ineffective classifier outcome is exhibited by the Fisher RF model with the least accuracy of 74.30. Next to that, slightly effective classifier results are attained by the SFFS-RF model with slightly higher accuracy of 76.40. Then, the No-FS and SFFS-RF reaches to a nearly equivalent and moderate accuracy of 80.40 and 81.80 respectively. At the same time, the PCA-RF and Relief-F-RF models outperforms the previous methods and obtained near identical accuracy values of 83.10. Simultaneously, GA-RF model shows manageable results with the accuracy value of 89.50 whereas the PSO-RF showcased outperforming results over the compared methods except DNN-AO model with the accuracy of 91.00. At last, the DNN-AO model reaches to a maximum classification performance with the utmost accuracy of 95.95.

On measuring the results in terms of specificity, it is noticed that ineffective classifier outcome is exhibited by the Fisher RF model with the least accuracy of 84.80 whereas the GA-RF model offers good classifier outcome with the high accuracy of 88.90. It is also noted that the PSO-RF model outperforms the previous methods except DNN-AO model with the accuracy of 90.00. At last, the DNN-AO model reaches to a maximum classification performance with the utmost accuracy of 95.95.

On measuring the results in terms of precision, it is noticed that ineffective classifier outcome is exhibited by the Fisher RF model with the least accuracy of 84.70 respectively. Similarly, while classifying the data under the class normal, i.e. class 0, the DNN-AO model achieves effective classifier outcome with the maximum accuracy of 95.77, sensitivity of 95.77, specificity of 95.77, precision of 95.77, MCC of 95.77, F-measure of 95.77, AUC of 95.77 and kappa value of 95.77 respectively. While classifying the data under the class metastases, i.e. class 1, the DNN-AO model achieves effective classifier outcome with the maximum accuracy of 92.68, AUC of 91.69 and kappa value of 83.59 respectively. Similarly, while classifying the data under the class Malign Lymph, i.e. class 2, the DNN-AO model achieves effective classifier outcome with the maximum accuracy of 92.68, AUC of 91.69 and kappa value of 83.59 respectively. Similarly, while classifying the data under the class Fibrosis, i.e. class 3, the DNN-AO model achieves effective classifier outcome with the maximum accuracy of 100, sensitivity of 100, specificity of 100, MCC of 100, F-measure of 100, AUC of 100 and kappa value of 100 respectively.

Table 2.

RESULTS ANALYSIS OF PROPOSED METHOD

<table>
<thead>
<tr>
<th>Performance index</th>
<th>Normal (0)</th>
<th>Metastases (1)</th>
<th>Malign Lymph (2)</th>
<th>Fibrosis (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>100</td>
<td>91.89</td>
<td>91.89</td>
<td>100</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>100</td>
<td>93.82</td>
<td>88.52</td>
<td>100</td>
</tr>
<tr>
<td>Specificity</td>
<td>100</td>
<td>89.55</td>
<td>94.25</td>
<td>100</td>
</tr>
<tr>
<td>Precision</td>
<td>100</td>
<td>91.56</td>
<td>91.53</td>
<td>100</td>
</tr>
<tr>
<td>MCC</td>
<td>100</td>
<td>83.00</td>
<td>83.00</td>
<td>100</td>
</tr>
<tr>
<td>F-Measure</td>
<td>100</td>
<td>92.68</td>
<td>90.00</td>
<td>100</td>
</tr>
<tr>
<td>AUC</td>
<td>100</td>
<td>91.69</td>
<td>91.38</td>
<td>100</td>
</tr>
<tr>
<td>Kappa</td>
<td>100</td>
<td>83.59</td>
<td>83.18</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 3.

AVERAGE RESULTS ANALYSIS OF PROPOSED METHOD

<table>
<thead>
<tr>
<th>Performance index</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>95.95</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>95.59</td>
</tr>
<tr>
<td>Specificity</td>
<td>95.95</td>
</tr>
<tr>
<td>Precision</td>
<td>95.77</td>
</tr>
<tr>
<td>MCC</td>
<td>91.50</td>
</tr>
<tr>
<td>F-Measure</td>
<td>95.67</td>
</tr>
<tr>
<td>AUC</td>
<td>95.77</td>
</tr>
<tr>
<td>Kappa</td>
<td>91.69</td>
</tr>
</tbody>
</table>

Fisher RF model with the least accuracy of 75.50. Next to that, slightly effective classifier results are attained by the SFFS-RF model with slightly higher accuracy of 77.10. Then, the No-FS reaches to a moderate accuracy of 81.20. At the same time, the PCA-RF and SFFS-RF models outperforms the previous methods and obtained near identical accuracy values of 83.90 and 83.40 respectively. Simultaneously, the Relief-F-RF model shows manageable results with the accuracy value of 84.20 whereas the GA-RF model offers good classifier outcome with the high accuracy of 92.20. It is also noted that the PSO-RF model showcased outperforming results over the compared methods except DNN-AO model with the accuracy of 94.00. At last, the DNN-AO model reaches to a maximum classification performance with the utmost accuracy of 95.95.
At the same time, the No-FS model outperforms the previous methods and obtained Precision values of 80.70. Simultaneously, the PCA-RF and Relief-F-RF models show manageable and nearly identical results with the value of 82.70 and 83.10 whereas the GA-RF model offers good classifier outcome with the high accuracy of 87.40. It is also noted that the PSO-RF model showcased outperforming results over the compared methods except DNN-AO method with the accuracy of 89.00. At last, the DNN-AO model reaches to a maximum classification performance with the utmost accuracy of 91.50.

On measuring the results in terms of MCC, it is noticed that ineffective classifier outcome is exhibited by the Fisher RF model with the least accuracy of 50.80. Next to that, slightly effective classifier results are attained by the SFFS-RF model with slightly higher accuracy of 54.60. Then, the No-FS reaches to a moderate accuracy of 62.70. At the same time, the SFBS-RF and PCA-RF models outperform the previous methods and obtained near identical accuracy values of 66.00 and 87.80 respectively. Simultaneously, the Relief-F-RF model shows manageable results with the accuracy value of 68.80 whereas the GA-RF model offers good classifier outcome with the high accuracy of 87.70. It is also noted that the PSO-RF model showcased outperforming results over the compared methods except DNN-AO method with the accuracy of 89.00. At last, the DNN-AO model reaches to a maximum classification performance with the utmost accuracy of 95.77.

On measuring the results in terms of F-Measure, it is noticed that ineffective classifier outcome is exhibited by the Fisher RF model with the least accuracy of 74.30. Next to that, slightly effective classifier results are attained by the SFFS-RF model with slightly higher accuracy of 75.20. Then, the No-FS reaches to a moderate accuracy of 82.90. At the same time, SFBS-RF outperforms the previous methods and obtained accuracy values of 80.80. Simultaneously, the PCA-RF and Relief-F-RF models shows manageable results with the accuracy value of 82.80 and 82.70 whereas the GA-RF model offers good classifier outcome with the high accuracy of 87.90. It is also noted that the PSO-RF model showcased outperforming results over the compared methods except DNN-AO method with the accuracy of 89.00. At last, the DNN-AO model reaches to a maximum classification performance with the utmost accuracy of 95.77.

On measuring the results in terms of AUC, it is noticed that ineffective classifier outcome is exhibited by the Fisher RF model with the least accuracy of 51.29. Next to that, slightly effective classifier results are attained by the SFFS-RF model with slightly higher accuracy of 54.60. Then, the No-FS reaches to a moderate accuracy of 62.79. At the same time, the SFBS-RF and PCA-RF models outperform the previous methods and obtained near identical accuracy values of 87.70 and 87.90 respectively. Simultaneously, the Relief-F-RF model shows manageable results with the accuracy value of 85.10 whereas the GA-RF model offers good classifier outcome with the high accuracy of 89.00. It is also noted that the PSO-RF model showcased outperforming results over the compared methods except DNN-AO method with the accuracy of 89.00. At last, the DNN-AO model reaches to a maximum classification performance with the utmost accuracy of 95.59.

On measuring the results in terms of Kappa, it is noticed that ineffective classifier outcome is exhibited by the Fisher RF model with the least accuracy of 41.69. Next to that, slightly effective classifier results are attained by the SFFS-RF model with slightly higher accuracy of 45.44. Then, the No-FS reaches to a moderate accuracy of 62.54. At the same time, the SFBS-RF and PCA-RF models outperform the previous methods and obtained near identical accuracy values of 87.72 and 87.90 respectively. Simultaneously, the Relief-F-RF model shows manageable results with the accuracy value of 85.10 whereas the GA-RF model offers good classifier outcome with the high accuracy of 89.00. It is also noted that the PSO-RF model showcased outperforming results over the compared methods except DNN-AO method with the accuracy of 89.00. At last, the DNN-AO model reaches to a maximum classification performance with the utmost accuracy of 95.67.
of 87.70. At the same time, the PCA-RF and No-FS models outperforms the previous methods and obtained near identical accuracy values of 89.90 and 90.90 respectively. Simultaneously, the Relief-F-RF model shows manageable results with the accuracy value of 91.30 whereas the GA-RF model offers good classifier outcome with the high accuracy of 95.40. It is also noted that the DNN-AO model showcased outperforming results over the compared methods except PSO-RF method with the accuracy of 95.77. At last, the PSO-RF model reaches to a maximum classification performance with the utmost accuracy of 97.00. On measuring the results in terms of kappa, it is noticed that ineffective classifier outcome is exhibited by the Fisher RF model with the least accuracy of 51.29. Next to that, slightly effective classifier results are attained by the SFFS-RF model with slightly higher accuracy of 53.81. Then, the No-PC and PCA-RF reaches to moderate and identical values with the accuracy of 62.77. At the same time, the Relief-F-RF model outperforms the previous methods and obtained accuracy values of 62.79 respectively. Simultaneously, the SFBS-RF model shows manageable results with the accuracy value of 65.17 whereas the GA-RF model offers good classifier outcome with the high accuracy of 87.90. It is also noted that the PSO-RF model showcased outperforming results over the compared methods except DNN-AO method with the accuracy of 89.00. At last, the DNN-AO model reaches to a maximum classification performance with the utmost accuracy of 91.69. The above tables and figures clearly portrayed the superior characteristics of the presented model.

4 Conclusion

This paper has introduced a new DL based lymph disease identification and classification model based on DNN-AO model. Here, optimal DNN model is applied to classify the lymph data utilizing SA as a feature extractor and SoftMax layer as a classifier. In addition, to make the DNN model more efficient, DL based AO is used which is an adaptive learning rate optimization algorithm which has been developed particularly to train DNN. The presented DNN with AD model called DNN-AO model is tested against benchmark Lymphography dataset and the results are assessed under diverse measures. The experimental outcome verified that effective classification performance is showcased by the DNN-AO model with the maximum sensitivity, specificity, accuracy and kappa value of 95.59, 95.95, 95.95 and 91.69 respectively.

REFERENCES