International Journal of Future Generation Communication and Networking
Vol. 13, No. 3, (2020), pp. 73 – 87

OP-DNN: Lung Cancer Survivability Prediction using Novel Optimized-Deep Neural Network Classification Method

Pradeep K.R.[1*], Naveen N.C.[2]
[1*]*Dept. of CSE, K.S. School of Engineering and management, Bengaluru Karnataka*
[2]*Dept. of CSE, JSS Academy of Technical Education, Bengaluru, Karnataka*
[1,2]*Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India.*
[1*]*pradeep2kr2@gmail.com*, [2]*ncnaveen@gmail.com*

Abstract

Lung cancer disease is the most widely recognized deadly disease in the world for the loss of life. Throughout this research, Electronic Health Records (EHRs) textual data are investigated, and survivability rates for lung cancer affected patients are predicted. If the Lunga cancer patients are survivable for more than one year, chemotherapy treatment can be started for those patients. This research paper examines an effective Batch Size-Optimizer based Deep Neural Network (Op-DNN) classifier framework model, which is developed to predict the patient's survivability. Here the textual data set is classified and processed in batches for each iteration. The errors generated from the original classification of the initial batch size is fed back to the Op-DNN classification algorithm for further iterations with the reduced error loss that is free from underfitting and overfitting. The proposed method is compared with various parameters for Machine learning classifier algorithms demonstrating that the Op-DNN model has achieved a better accuracy rate of 91.353 % for prediction of Survivability.

Keywords: *Lung Cancer, Diabetes, Artificial Neural network, DNN, Optimizer, ADAM, ReLU, Epoch, Batchsize*

1. Introduction

Lung cancer is liable for the major number of deaths (1.8 million deaths, 18.4% of the total), since of the poor prediction for this cancer globally [1].

However, the improvement in medical Artificial Intelligence (AI) has been related to the development of AI programs, planned to support in the design of a prediction model that can be used to make decisions. In this research work, an effective Optimized based Deep Neural Network (Op-DNN) Classification technique, an extended version of DNN is proposed, focusing upon the lung cancer survivability prediction using the large textual data. This helps in the early detection of lung cancer, so proper treatment can bet taken at the initial stage itself by providing the result for treatment and curing the lung cancer affected patients. The survival period study is to be measured clinically, significant in order to access patient prediction to undergo chemotherapy. In this study, the North Central Cancer Treatment Group (NCCTG) lung cancer data set is utilized along with the real time of new patient data to get survivability prediction.

A DNN is an Artificial Neural Network (ANN) with multiple layers between the input and output layers. This model, with the combination of batch size, epoch, and optimizer, provides survivability prediction by addressing the problem of overfitting and underfitting capabilities as compared to using a single methodology. The classification based on DNN is primarily a Back Propagation (BP) algorithm, whose prototype usually implements BP

ISSN: 2233-7857 IJFGCN

Neural Network (BPNN) [2]. The classification approach centered on the neural networks has a high learning capability. However, the degree of convergence and the generalization capability is not strong, and it is easy for running into a local minimum point, which increases data grouping accuracy, classification period efficiency by computing the gradient of a multi-variable function. This exercise also qualifies for predictive rate approaches, applied by the preferred features to examine the lung cancer patient's data. The dataset used here emphases on extent accessible at the period of diagnosis, representing a more active set of survival predictor class.

Applying Machine learning techniques (MLT) such as DNN, classification trees (C4.5), Naive Bayes (NBs), Support Vector Machine (SVM)-Linear are compared with the Op-DNN using the lung cancer data for precise survivability prediction. This is valuable for both the doctors and patients, helping out in making decisions by the finest track of treatment for lung cancer affected patients.

The manuscript is systematized as follows: Background section reviews current research work on the MLT and DNN for the affected lung cancer patients using various classification techniques, ensue by a proposed section which comprises the methodology defining novel Op-DNN system architecture, in contrast of different MLT with outcomes and application. Performance evaluation comprises trials and results, which is also described in this section, and the last part includes the conclusion section and future work.

2. Background

The MLT in the field of biomedical, along with several other applications, prove to lead the way in health management [3]. Supervised and unsupervised learning are the primary types of ML. The categorized group of data trained includes input, along with the desired result in supervised learning. The use of ML techniques, Electronic Health Records (EHRs) finds an application that aids the management of patients to care and improvisation of the performance in hospital care management [4].

Researchers have performed work to improve clinical outcome predictions by using deep learning concepts that can fit in imaging scans at various time points. Done by means of time series scans for significantly improving predictive of survival and lung cancer-specific outcomes [5].

Medical researchers proposed a Hybrid 3D-Deep Convolutional Neural Networks (CNN) architecture by classification in order to diagnose lung cancer on CT scan images dataset, to ease the feature mining progression by natural methods with the concept of extracting extra valuable low-level to high-level features having the substantial outcomes [6].

Researchers demonstrated a deep learning application using CNN in medical imaging authorization for the automated quantification of radiographic features and hypothetically improving patient stratification for Non-small-cell lung cancer (NSCLC) patients by varying clinical courses and outcomes, uniform in the equal tumor stage [7].

Medical researchers presented an ANN clinical tool for identification of lung cancer risk prediction using the concept of multi-parameterized by using personal health information accessible in EMR systems, having high specificity and modest sensitivity [8].

Researchers have combined ANN weights after optimization with pruning along with Genetic Algorithm (GA) showing higher convergence, greater success rate, and lesser execution time during the test stage. This method minimizes the resources required for computation and also proves to be an alternative for the neuro-genetic design of neural classifiers [9].

ISSN: 2233-7857 IJFGCN

Artificial Intelligence (AI) finds its applications in several healthcare applications, and for structured data, ML can be used. For the unstructured data, the Neural Network (NN), classical SVM, Natural Language Processing (NLP), and deep learning may be applied. The AI is applied in identifying the diseases related to neurology, cancer, and cardiology [10].

In the research work carried out for prediction of the Parkinson diseases are based on directed systems, basically end up to DNN architecture, which is trained on extracting rules after networks, tested and then applied as complete systems, to get the expected outputs [11].

Researchers have proposed a model that extracts deep features from computed tomography images using pre-trained CNN model in contrast to non-small cell adenocarcinoma lung cancer, using trained classifiers to predict short- term and long-term survivors [12].

The different MLTs like fuzzy logic, NBs, clustering, genetic, SVM, NN, random tree and decision tree, etc. finds extensive applications in the detection, diagnosis, classification, and risk assessment of cancer [13].

A Deep convolution neural network performs better in the classification of the images, different visual tasks, and object detection. The different neurons are activated through the weighted networks over the earlier active neurons [14].

Lung cancer digital image analysis is done to predict the survival rate of a patient through the computational procedure that sort images into groups by histogram equalization then fed into the NN classifier to verify patient status if it is in a normal or abnormal state [15].

Researchers have projected the use of ensemble voting with five decision tree-based classifiers and meta-classifiers to find lung cancer survival prediction using SEER data. The qualified outcome is predicted by estimating the risk of mortality, and performance is based on accuracy and area under the ROC curve [16].

Researchers have proposed a predictive method by using Genetic Algorithm NN (GANN) by means of Bayes' proposition related with logistic regression, which was related by investigating the target outcome of patients living or dead status based on the accuracy of classifying, at 6 months' time interval, later surgery is inspected [17].

The literature review is carried out in the machine learning algorithms applied to detect lung cancer at the early stage to predict the probability of survivability.

3. Op-DNN Methodology

The methodology adopted in the research work attempts to predict the survivability of lung cancer patients with Deep learning technique considers, gender, diabetes, smoking along with other features is achieved, helping the doctors to recommend for chemotherapy based on the number of months of survivability. The proposed methodology is built on the DNN for the classification, where the classifier is forecasting the status of lung cancer patient survivability status is as shown below.

3.1 System Architecture

The proposed system architecture of Op-DNN is shown in Figure 1, consisting of four phases. 1. Lung cancer dataset acquisition.2. Exploratory data analysis includes missing value action and normalization, along with the process of splitting data for training and testing.3. OP-DNN classifier approach. 4.Comparison of algorithms with visualization.

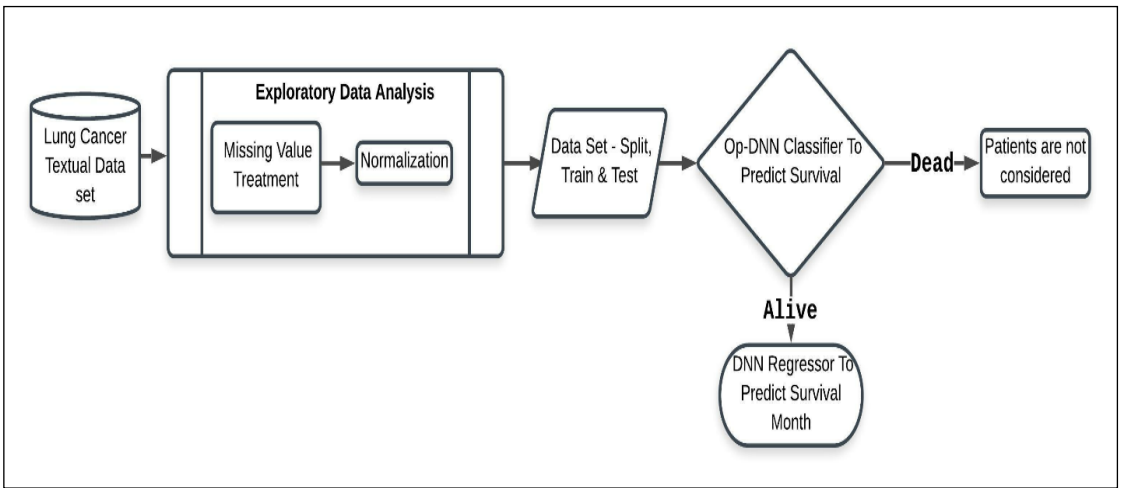

Figure 1. OP-DNN system architecture

3.2 Phase One-Lung cancer dataset acquisition

The data collection is done from the patients after admitting to the hospital. The various data gathered are age, gender, tumor location, stage timing, n-stage, t-stage, diabetes, smoker. The survey is carried out along with the doctors as well as pathologists at different hospitals to select the features. The NCCTG lung cancer data set [18], along with a new patient, data is used and data features are extracted for evaluating the performance. This forms the required features which are needed for the prediction with a probability of more information and non-redundant [19]. This feature set mined is added to OP-DNN, and prediction is performed. The results of the proposed Op-DNN model are compared with other MLT classifier algorithms like DNN, SVM- Linear, NBs, C4.5, J48, where the proposed model provides the best-predicted results**.**

3.2 Phase Two- Exploratory data analysis

In this phase, the total number of textual data present in a dataset is 3065, and the feature set available count is 18 that is facilitated for training plus testing of data essential for measuring decent accuracy. In a total count of 3065 lung cancer data set 2137 are male, and 928 are female lung cancer patient's data in which 1726 lung cancer patients are Type 1 Diabetic (T1D), 916 are Type 2 Diabetic (T2D), and non-diabetic are of 424, as of to consider smoking feature set, 1762 are smoker, and 1303 are non-smoker lung cancer patients. In the data set, lung cancer patient's survival status includes the count of 2628 are alive, and 437 are dead.

a. Missing value treatment and Normalization: Table 1(a) and Table 1(b) summarizes the information of 18 features from the dataset which count of samples, mean, standard deviation, min, 25%, 59%, 75%, max values, and Table 2 and Table 3 shows the normalization, done to get rid of the missing values.
b. Split data for training and testing: A large set of labeled data are trained by using deep learning models. In this regard, out of 3065 lung cancer data sets, train data and test data are considered in the ratio of 80% and 20%, so the train data count is 2452, and the test data count is 613.

Table 1(a). Count, mean, standard deviation, min, 25%, 59%, 75%, max for lung cancer data set sample

	Gender	Age	smoker	Tumor location	t_stage	n_stage	Stage	Timing	Diabetes
count	3065								
mean	1.30	66.41	1.43	4.31	2.71	2.95	1.63	2.15	1.16
std	0.46	9.59	0.49	2.25	1.13	1.00	0.48	0.64	0.64
min	1.00	36.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00
25%	1.00	60.00	1.00	4.00	2.00	3.00	1.00	2.00	1.00
50%	1.00	67.00	1.00	4.00	2.00	3.00	2.00	2.00	1.00
75%	2.00	74.00	2.00	6.00	4.00	4.00	2.00	3.00	2.00
max	2.00	87.00	2.00	18.00	4.00	4.00	2.00	3.00	2.00

Table 1(b). Count, mean, standard deviation, min, 25%, 59%, 75%, max for lung cancer data set sample

	Status	meal_cal	wt_loss	ph_ecog	ph_karno	pat_karno	Survived month	Survived Year	Survived Status
count	3065								
mean	1.68	882.40	9.53	0.96	82.48	79.56	32.16	26.79	0.86
std	0.47	432.82	12.54	0.83	11.84	14.43	54.42	45.34	0.35
min	1.00	96.00	24.00	0.00	50.00	30.00	1.28	0.10	0.00
25%	1.00	568.00	0.00	0.00	80.00	70.00	7.81	0.67	1.00
50%	2.00	910.00	7.00	1.00	80.00	80.00	16.00	1.38	1.00
75%	2.00	1150.00	15.00	1.00	90.00	90.00	32.78	2.69	1.00
max	2.00	5733.00	68.00	8.00	100.00	100.0	95.30	79.40	1.00

Table 2. Normalization for missing value treatment for Gender, age, smoker tumor location, t_satge, n-_stage, and Stage for lung cancer dataset.

study_id	Gender	Age	smoker	Tumor location	t_stage	n_stage	Stage
1	-0.302773	-0.066769	-0.425122	0.099453	0.430343	0.348668	0.369331
2	0.697227	-0.007945	-0.425122	-0.194665	0.430343	0.015334	0.369331
3	-0.302773	-0.164808	0.574878	-0.018194	0.430343	-0.651332	0.369331
4	0.697227	-0.066769	0.574878	0.099453	-0.569657	0.348668	0.369331
5	-0.302773	-0.145200	0.574878	0.099453	-0.236324	0.015334	-0.630669

Table 3. Normalization for missing value treatment for timing, diabetes, status, meal_cal, wt_loss, ph_echog, ph_karno and pat_karno for lung cancer dataset

study_id	Timing	Diabetes	Status	meal_cal	wt_loss	ph_ecog	ph_karno	pat_karno
1	0.4247	0.4195	0.3207	0.0519	0.2442	0.0053	0.1504	0.2920
2	0.4247	0.4195	0.3207	0.0607	0.0594	-0.1196	0.1504	0.1491
3	0.4247	0.4195	-0.6790	-0.081	0.0594	-0.1196	0.1504	0.1491
4	0.4247	0.4195	0.3207	0.0474	0.0159	0.0053	0.1504	-0.2797
5	-0.6306	0.4247	0.4195	0.3207	-0.0405	-0.1035	-0.1196	-0.6306

Class labeling is done to know the survivability of the patients. The label more and less shows that the patient survivability is greater than one year, and survivability is lesser than one year, respectively. The classifier algorithms are used to do the analysis of dataset training for the creation of a new Op-DNN. For the test dataset, prediction evaluation is carried out using several other algorithms.

3.3 Phase Three-OP-DNN Classifier approach

The network-based on deep learning is separated from the single hidden layer with respect to the layers of nodes by which the data passes in recognition of data. Any network with greater than three layers is called as deep learning. With respect to the unlabeled data training, every node layer within the deep network acquires the features automatically. The OP-DNN architecture with the optimizer, as shown in Figure 2.

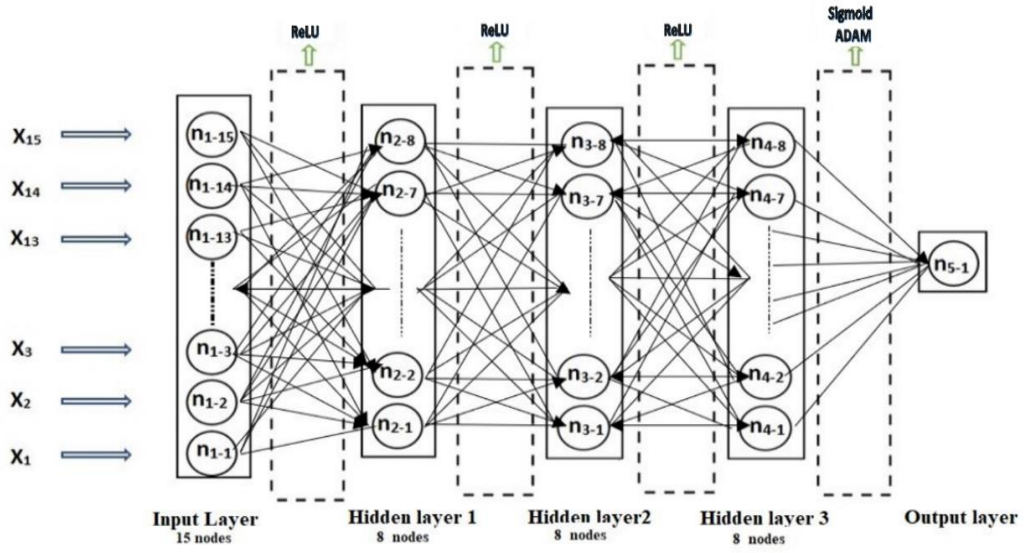

Figure 2. OP-DNN architecture with the optimizer

OP-DNN classifier follows five steps to predict the survivability of lung cancer patients; the flow chart of the OP- DNN classifier is shown in Figure 3.

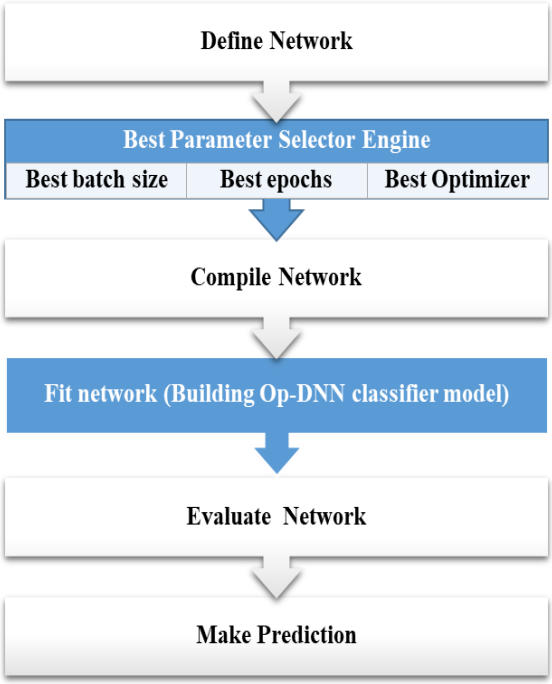

Figure 3. Op-DNN classifier flow

3.3.1 Define network: This network includes initializing the DNN classifier, here Keras Sequential model is applied for initializing the DNN classifier that involves a sequential model by passing a list of layer instances specifying the input shape for the situation which would anticipate. As a part of deep learning model, the embedding layer is set with random weights and will learn along with the model itself, for all of the data in the training dataset. The following are the sequence of actions that take place in defining the DNN for one 15 input feature set units of Lung Cancer patient's data, three hidden layers each consisting of eight units and one output layer having a total of 41 nodes and 264 edges.

a. Adding the input layer and the first hidden layer includes to input shape of a sequential classifier to the first layer consists of a tuple of integers where the batch dimension is not considered, for the specification of the input shape 2D layer dense approach has been applied for the input dimension of 16 feature set of data with eight units in the first hidden layer, having neurons that work in the message of weight, bias and their particular activation function here the activation function used is RELU.
b. In each hidden layer, back-propagation of the NN is done by Keras Dense and Keras Kernel Initializer that propagates the updating of weights and biases of the neurons on the source of error at the output followed by a non-linear activation function.
c. A non-linear activation function will help to learn as per the modification w.r.t error. Activation functions create the back-propagation likely as the gradients are provided through the error, to impart the biases and weights.

$$Y = activation\ function \sum(weights * inputs) + bias \quad (1)$$

$$Z(X,1) = W(1)X + b(1)$$

Here, Z (1) is the vectorized output of hidden layer 1 specifying class label Dead or alive, W (1) be the vectorized weights assigned to 8 units of neurons in hidden layer1, i.e., w_1, w_2, w_3.,X be the vectorized 15 input feature set of, i.e., i_1, i_2 till i_{15},b is the vectorized bias assigned to neurons in the hidden layer, i.e., b_1 and b2, a (1) is the vectorized form of any linear function.

d. ReLU a Rectified Linear Unit used as a non-linear activation function in normal DNN

which are applied to the hidden layers of NN, that back-propagates the errors and activate the neurons multiple layers This provides an output x if x is positive and 0 otherwise with Value Range of [0,inf]

$$A(x) = max(0, x) \quad (2)$$

In this section, analysis is done that provides insight towards the loss, as an estimate or a higher bound to per sample gradient norm. So, let x_i, y_i be the i^{th} input-output pair from the training set, $\Psi(\cdot;\ \theta)$ be a deep learning model parameterized by the vector θ, and $L(.;.)$ be the loss function to be minimized during training. To achieve it, let L (ψ, y): D → R be either the negative log-likelihood through a sigmoid or the squared error loss function defined respectively as

$$L1(\psi, y) = -\log\left(\frac{exp(y\psi)}{1 + exp(y\psi)}\right) \quad y \in \{-1, 1\}\ \psi \in R \quad (3)$$

$$L2(\psi, y) = \|y - \psi\|_2^2 \quad y \in Rd \quad \psi \in R_d \quad (4)$$

Given our upper bound to the gradient norm, from equations 3 and 4, defines equation 5.

$$\|\nabla\theta t\ L(\Psi(x_i;\ \theta_t),\ y_i)\|\ 2\ \leq\ L\rho\ \|\nabla\Psi L(\Psi(x_i;\ \theta_t),\ y_i)\|2 \quad (5)$$

e. With the estimation of the above, due facts, selection proportionally to the loss lessens the variance only when the majority of the samples have losses close to 0. Our hypothesis is confirmed from trials, where the loss scuffles to attain a speedup in the initial phases of training. Adding the second, third hidden layer, each of 8 nodes develops the normal DNN and weights, bias regulation continues for BP adjusting the weights with the ReLU activation for each layer of nodes as stated in the step1 until it reaches the final output layer.

f. Adding the output layer includes Keras Dense with one output node, Keras Kernel Initializer, and Sigmoid activation used while building neural networks. The multilayer perceptron uses the sigmoid as the transfer function. Sigmoid activation considers a two-class problem, i.e., binary classification for lung cancer patient's data, with classes identified as staying alive for more than one year as S1 and Less representing that the patient might stay alive less than one year as S2 indicated to S1 to 1 and S2 as 0.

g. The sigmoid activation function produces a constant value in the range 0 to 1, shown in equation 6.

$$output_{i=}\frac{1}{1+e^{-x}} \qquad \text{Where} x = \text{gain.activation}_i \quad (6)$$

A variant of the sigmoid transfer function is the hyperbolic tangent function shown in equation7.

$$output_{i=}\frac{e^{x} - e^{-x}}{e^{x}+e^{-x}} \quad (7)$$

3.3.2 Best Parameter Selector Engine: This stage involves identifying the best batch size, best epoch, and besting optimizer from the Op-DNN classifier with the input of 5 multiple value combinations chosen from the factors of batch size, epochs, optimizer, scoring, and cross-validation. The selector engine, as shown in Figure 4, helps in finding the right batch size resolving to utilize less memory space to update the weights in each pass. Op-DNN classifier is optimized by cross-validation, done by using the Gridsearch CV algorithm, implementing a "fit" and a "score" method to attain the best batch size required for the data set, epoch size and optimized data set of output. The Steps are shown in Figure 4(a),4(b), along with the details.

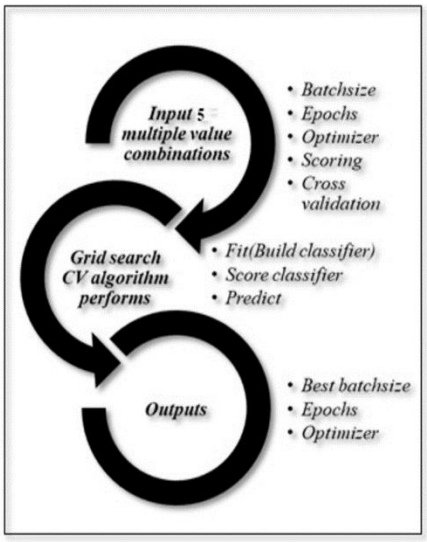

Figure 4 (a). Gridsearch CV algorithm diagram

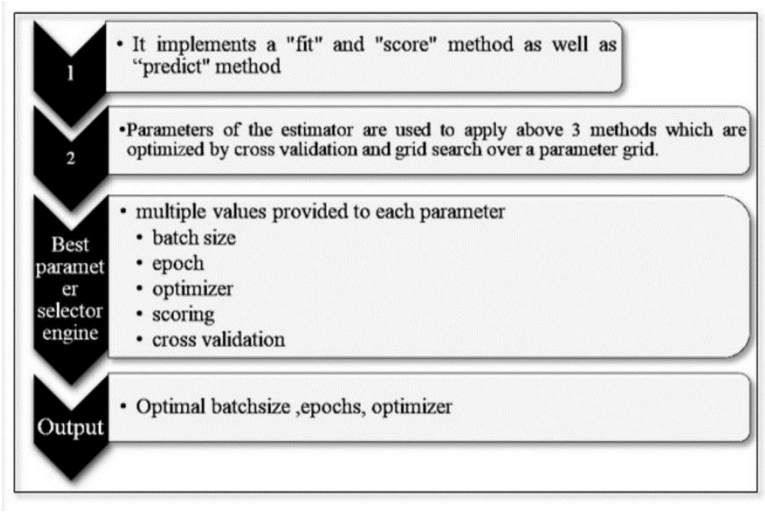

Figure 4(b). Gridsearch CV algorithm flow

a. Fit the values for each parameter chosen from: 'batch size': [10,25,32,50], 'epochs': [50,100, 500], and 'optimizer': ['Adam', 'RMSprop']} worked for all combinations and finding out the best combination. Here values of parameters are fixed to reduce memory consumption and time. Apply 10-fold cross-validation on each parameter's combination, calculated with the estimator, then outputs 10 kinds of scores on each parameter combination.

b. Calculate the mean score from Step 2 to one of the parameter combinations having the lowest score as the most optimized parameter combinations on CV, and each loop leads to the different combinations as the optimized one. Therefore, repeat this process many times, and the best parameter is chosen based as the most optimized state.

c. Repeat step 3 several times to have one set of parameter combinations, providing the best results. From step iv. Gridsearch CV produces one set of parameter combinations; as a result, Op-DNN classifier crops to best parameter combination having batch size as 50, Epochs as 50, and optimizer as Adam.

3.3.3 Compiling the Network: Involves training OP-DNN classifier used in the best

parameter combination of optimization method along with gradient descent [20]. This technique computes the gradient of a loss function through all the weights in the network.

The gradient is nourished to the optimization method, which in turn uses it to update the weights in an effort to minimize the loss function. ADAM optimizer [21] is used along with the solution to minimize the loss; Binary Cross Entropy is used. Here the main objective of the function is having less value of mean square error function (loss/cost function) and finding optimize values weights of normal DNN. Along with this, calculation of metrics that includes accuracy, f2 score, precision, recall is done based on Wrapper classes for turning Tensor Flow metrics into Keras metrics. The equations below explain the working feature involved. The Adam optimization algorithm is a blend of gradient descent with momentum and RMSprop algorithms, where it works well even with a slight fine-tuning of hyperparameters, which are shown below.

a. The first step, calculate an exponentially weighted average of previous gradients, store it in variables G_w and G_b eearlier to bias correction and $G_{w^{corrected}}$ and $G_{b^{corrected}}$ done with bias correction.
b. The second step, at this instant, calculate an exponentially weighted average of the squares of the past gradients and store it in variables SG_w and SG_b earlier to bias correction along with $SG_{w^{corrected}}$ and $SG_{b^{corrected}}$, on iteration i, calculate the results of weights: w & bias: b, using current mini-batch.
c. Lastly, update parameters on joining information from the first step and second step as shown in equations 8 to 16.

i. Initialize G_w , SG_w, G_b and SG_b to zero, update G_w and G_b like momentum

$$G_w = \beta_1 * G_w + (1-\beta_1) * \mathrm{w} \quad (8)$$

$$G_b = \beta_1 * G_b + (1-\beta_1) * \mathrm{b} \quad (9)$$

ii. Update SG_w and SG_b like Rmsprop

$$SG_w = \beta_2 * SG_w + (1-\beta_2) * \mathrm{w}^2 \quad (10)$$

$$SG_b = \beta_2 * SG_b + (1-\beta_2) * b^2 \quad (11)$$

iii. In Adam optimization technique, implement bias correction

$$G_{w^{corrected}} = \frac{G_w}{\left(1-\beta_{1}{}^{i}\right)} \text{ and } G_{b^{corrected}} = \frac{G_b}{\left(1-\beta_{1}{}^{i}\right)} \quad (12)$$

$$SG_{w^{corrected}} = \frac{S_w}{\left(1-\beta_{2}{}^{i}\right)} \text{ and } SG_{b^{corrected}} = \frac{S_b}{\left(1-\beta_{2}{}^{i}\right)} \quad (13)$$

iv. Update parameters w and b

$$w = w - LR * \left(G_{w^{corrected}} \Big/ \sqrt{(SG_{w^{corrected}} + \varepsilon)} \right) \quad (14)$$

$$b = b - LR * \left(G_{b^{corrected}} \Big/ \sqrt{(SG_{b^{corrected}} + \varepsilon)} \right) \quad (15)$$

Where epsilon ε is a very small number to avoid dividing by zero, β_1 and β_2 are hyperparameters that regulate two exponentially weighted averages, here the default values for $\beta_1 = 0.9$ and $\beta_2 = 0.999$.

$$-\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log\left(\hat{y}_i\right) + (1-y_i)\log\left(1-\hat{y}_i\right)\right] \quad (16)$$

3.3.4 Fit the network: Building the Op-DNN classifier model with the preparation of training and tuning the lung cancer dataset. For training the data, fit method implied again in step 2 involving the best batch size and Epoch that fits the Novel Op-DNN classifier model. Epoch specifies one forward pass as well as one backward pass of all the training samples, done through the total number of iterations, specifying the number of passes, done using batch size.

The best value of batch size and epoch is 10 and 50, as acknowledge by the Best Parameter Selector Engine from steps 2, 3 and step 4 passes on the next step 5 for the

evaluation and prediction of the test data set. In this stage, evaluation metrics are used to know Op-DNN model finest, free from underfitting and overfitting for the trained data. The Table 4 and graph shown below in Figure 5, depicts the novel Op-DNN classifier model for the training data set the size of 2452 out of 3065 lung cancer data set, performs well with reduced error loss and maximum accuracy when compared with existing DNN model graph Figure 6, in each iteration of epochs vs. evaluation Metrics.

Table 4. Op-DNN Classifier and DNN Classifier results for trained data

	Op-DNN Classifier	**DNN Classifier**
"ADAM"	**batch_size = 50 epochs = 50**	
Train data =2452		
units/step	60us/step	116us/step
loss	0.2569	0.3389
accuracy	0.91353	0.87645
f2_score	0.8503	0.8552
precision	0.9702	0.92653
recall	0.93442	0.919032

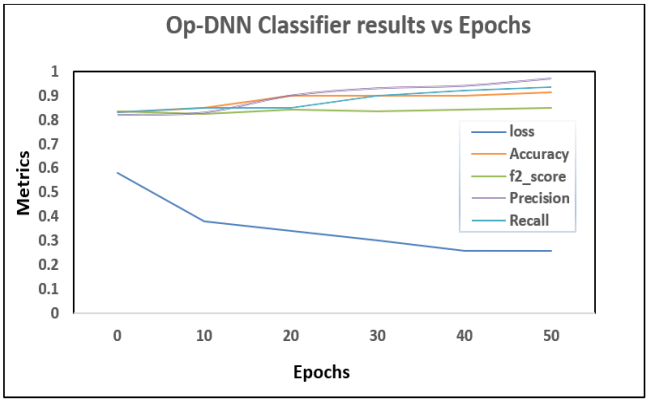

Figure 5. Graph of Op-DNN Classifier results vs. epochs for each iteration

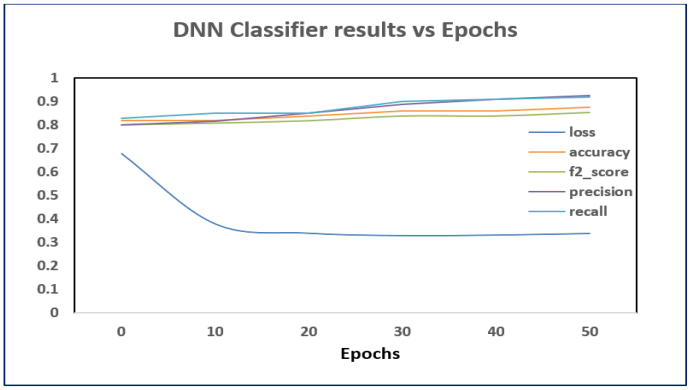

Figure 6. Graph of DNN Classifier results vs. epochs for each iteration

Stating to the conclusion for the trained data, the novel Op-DNN model is best when compared to the regular DNN model with the reduced error loss and accuracy. Evaluation and prediction of the results need to be done for the test data set, which is of size 613 in step 5.

3.3.5 Evaluation and prediction: Novel Op-DNN model is set to be identified by the performance level calculated by different types of evaluation metrics, as shown below.

Confusion Matrix: Provides a matrix as output and defines the thorough performance on the testing model on 613 datasets showing the predicted class that the lung cancer patient will survive more than one year or not versus the actual class is shown in Table 6.

Table 6. Confusion matrix of the test data obtained by Novel Op-DNN classifier model

Test data =613	**Predicted class= dead or alive**		
Actual class dead or alive		**Class = Yes**	**Class = No**
	Class = yes	TP=454	FN=83
	Class = No	FP=36	TN=40

Where FP=False Positives, TP =True positive, FN=False Negatives, TN=True Negatives.

3.4 Phase Four-Results and Discussion

The novel Op-DNN classifier model is built using Jupyter notebook with Skit learn, Keras, and Tensorflow. Performance evaluation is measured with output as "Less," i.e., the patient will survive for 1 month to the 1-year range or "More," i.e., the patient will survive for more than 1 year. The plan for treatment action for the lung cancer patient will be prepared, consequently by the referring pathologist. Op-DNN classifier survivability prediction outputs are shown in Table 7 for the following conditions such as No diabetes with lung cancer, T1D with lung cancer, and T2D with lung cancer.

Table 7. Op-DNN Classifier Survivability Prediction Output

SMOKER								
Lung Cancer Stage		**2nd stage**		**3rd stage**			**Total=Any stage**	
Gender		**Male**	**Female**	**Male**	**Female**		**Male**	**Female**
NO DIABETES	Survived Count	98	34	154	40		252	74
	Dead Count	0	0	0	0		0	0
T1D	Survived Count	257	65	427	205		684	270
	Dead Count	0	10	10	0		10	10
T2D	Survived Count	79	61	130	52		209	113
	Dead Count	22	11	80	27		102	38
NON-SMOKER								

Lung Cancer Stage		2nd stage		3rd stage			Total=Any stage	
Gender		Male	Female	Male	Female		Male	Female
NO DIABETES	Survived Count	8	15	50	24		58	39
	Dead Count	0	0	0	0		0	0
T1D	Survived Count	257	57	278	141		535	198
	Dead Count	0	11	0	8		0	19
T2D	Survived Count	90	31	141	88		231	119
	Dead Count	10	16	46	32		56	48

The survivability prediction and recommendation of the treatment to the patients is the main objective of this research work. To accomplish this objective, the task is to know the efficiency of predicting lung cancer survivability with the different algorithm's performance like OP-DNN, DNN, SVM-Linear, NBs, and C4.5 are compared to know the precision, accuracy, and AUC. Table 8 describes that the accuracy and precision value is high (91.35% and 97.02 %) for the proposed novel Op-DNN classifier model.

Table 8. Lists the performance of algorithms

	Accuracy	Precision	Recall	specificity	FPR	fmeasure
OP-DNN	0.91353	0.970204	0.934424	0.709096	0.290913	0.951504
DNN	0.87645	0.926532	0.919032	0.697434	0.302521	0.922816
C4.5	0.72727	0.827273	0.689394	0.784091	0.215909	0.752063
NBs	0.56181	0.661818	0.551515	0.526971	0.422727	0.601612
SVM-linear	0.42830	0.549091	0.457576	0.250825	0.620001	0.499175

The performance evaluation graphs for the proposed and existing methods are shown in Figure 7.

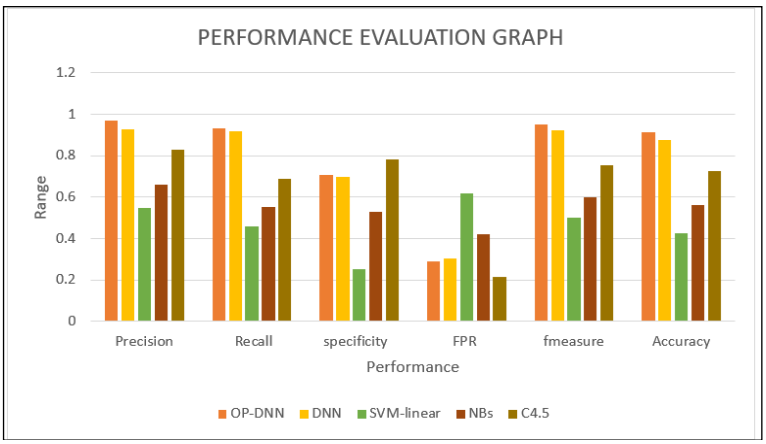

Figure 7. Performance evaluation graph

4 Conclusion

In this work, an efficient method for the classification of lung cancer survivability using an Optimized based deep learning approach is proposed. The performance of the Optimized based deep learning method, i.e., Op-DNN, is compared with the existing Machine learning classifiers. In addition, the metrics like accuracy, precision, recall, specificity, FPR, and fmeasure values are computed for the proposed and existing classification techniques to analyze the classification performance. The prediction accuracy obtained by combining the optimized approach based on batch size and Gridsearch CV technique with the DNN is 91.35% for the proposed method, and it is greater compared to other existing approaches. Moreover, from the outcomes, it could be concluded that the proposed method is capable of predicting the less or more survivability of patients suffering from lung cancer disease. In the future, the proposed work will be extended that can be applied for lung cancer disease prediction for Wide range of real-time data and at the same time, maintaining the prediction accuracy. Op-DNN could also be employed with Regression techniques in the future for better results.

References

[1] WHO Press Release N° 263, "Latest global cancer data", Retrieved from https://www.who.int/cancer/PRGlobocanFinal.pdf, **(2018).**

[2] Saravanan, K., & Sasithra, S., "Review on Classification Based on Artificial Neural Networks. The International Journal of Ambient Systems and Applications, vol. 2, no. 4, **(2014),** pp. 11-18. doi:10.5121/ijasa.2014.2402.

[3] A. Appari, M. Eric Johnson, and D. L. Anthony, "Meaningful Use of Electronic Health Record Systems and Process Quality of Care: Evidence from a Panel Data Analysis of U.S. Acute-Care Hospitals," Health Services Research, vol. 48, no. 2pt1, Jul. **(2012)**, pp. 354–375, doi: 10.1111/j.1475-6773.2012.01448.x.

[4] FitzHenry, F., Murff, H., Matheny, M., Gentry, N., Fielstein, E., & Brown, S. et al., "Exploring the Frontier of Electronic Health Record Surveillance," Medical Care, vol. 51, no. 6, Jun. **(2013)**, pp. 509–516, doi: 10.1097/mlr.0b013e31828d1210.

[5] Xu, Y., Hosny, A., Zeleznik, R., Parmar, C., Coroller, T., Franco, I.,Aerts, H. J. W. L, "Deep Learning Predicts Lung Cancer Treatment Response from Serial Medical Imaging," Clinical Cancer Research, vol. 25, no. 11, Apr. **(2019)**, pp. 3266–3275, doi: 10.1158/1078-0432.ccr-18-2495.

[6] Polat, H., & Danaei Mehr, H. "Classification of Pulmonary CT Images by Using Hybrid 3D-Deep Convolutional Neural Network Architecture," Applied Sciences, vol. 9, no. 5, Mar. **(2019)** , p. 940, doi:10.3390/app9050940.

[7] Hosny, A., Parmar, C., Coroller, T. P., Grossmann, P., Zeleznik, R., Kumar, A., … Aerts, H. J. W. L.,"Deep learning for lung cancer prognostication: A retrospective multi-cohort radiomics study," PLOS Medicine, vol. 15, no. 11, Nov. **(2018)**, p. e1002711, doi:10.1371/journal.pmed.1002711.

[8] Hart, G. R., Roffman, D. A., Decker, R., & Deng, J, "A multi-parameterized artificial neural network for lung cancer risk prediction," PLOS ONE, vol. 13, no. 10, Oct. **(2018)**, p. e0205264, doi:10.1371/journal.pone.0205264.

[9] S. Sakshi and R. Kumar, "A Neuro-Genetic Technique for Pruning and Optimization of ANN Weights," Applied Artificial Intelligence, vol. 33, no. 1, Oct. **(2018)**, pp. 1–26, doi: 10.1080/08839514.2018.1525524

[10] Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., & Shen, H, "Artificial intelligence in healthcare: past, present and future," Stroke and Vascular Neurology, vol. 2, no. 4, Jun. **(2017)**, pp. 230–243, doi: 10.1136/svn-2017-000101.

[11] D. Kollias, A. Tagaris, A. Stafylopatis, S. Kollias, and G. Tagaris, "Deep neural architectures for prediction in healthcare," Complex & Intelligent Systems, vol. 4, no. 2, Nov. **(2017)**, pp. 119–131, doi: 10.1007/s40747-017-0064-6..

[12] Paul, R., Hawkins, S. H., Hall, L. 0., Goldgof, D. B., & Gillies, R. J,"Deep Feature Transfer Learning in Combination with Traditional Features Predicts Survival Among Patients with Lung Adenocarcinoma," Tomography, vol. 2, no. 4, Dec. **(2016),** pp. 388–395, doi: 10.18383/j.tom.2016.00211.

[13] S. Agrawal and J. Agrawal, "Neural Network Techniques for Cancer Prediction: A Survey," Procedia Computer Science, vol. 60**, (2015),** pp. 769–774, doi: 10.1016/j.procs.2015.08.234.

[14] J. Schmidhuber, "Deep learning in neural networks: An overview," Neural Networks, vol. 61, Jan. **(2015)**, pp. 85–117, https://doi.org/10.1016/j.neunet.2014.09.003.

[15] Ada, R. K.," Early detection and prediction of lung cancer survival using neural network classifier", International Journal of Application or Innovation in Engineering & Management,vol. 2, no. 6, (2013),pp. 375-383.

[16] A. Agrawal, S. Misra, R. Narayanan, L. Polepeddi, and A. Choudhary, "Lung Cancer Survival Prediction using Ensemble Data Mining on Seer Data," Scientific Programming, vol. 20, no. 1, (2012), pp. 29–42, doi: 10.1155/2012/920245.

[17] M. F. Jefferson, N. Pendleton, S. B. Lucas, and M. A. Horan, "Comparison of a genetic algorithm neural network with logistic regression for predicting outcome after surgery for patients with nonsmall cell lung carcinoma," Cancer, vol. 79, no. 7, Apr. **(1997)**, pp. 1338–1342.
[18] Loprinzi, C., Laurie, J., Wieand, H., Krook, J., Novotny, P., & Kugler, J., "Prospective evaluation of prognostic variables from patient-completed questionnaires. North Central Cancer Treatment Group.," Journal of Clinical Oncology, vol. 12, no. 3, , Mar. **(1994)** , pp. 601–607,doi: http://dx.doi.org/10.1200/jco.1994.12.3.601.
[19] Pradeep, K. R., & Naveen, N, C., "Lung Cancer Survivability Prediction based on Performance Using Classification Techniques of Support Vector Machines, C4.5 and Naive Bayes Algorithms for Healthcare Analytics," Procedia Computer Science, vol. 132, **(2018)**, pp. 412–420,doi: http://dx.doi.org/10.1016/j.procs.2018.05.162
[20] Bahar, P., Alkhouli, T., Peter, J., Brix, C. J., & Ney, H. ,"Empirical Investigation of Optimization Algorithms in Neural Machine Translation", The Prague Bulletin of Mathematical Linguistics,vol. 108,no. 1, **(2017)**,pp. 13-25. doi:10.1515/pralin-2017-0005.
[21] Kingma, D., & Ba, J." Adam: a method for stochastic optimization (2014)", **(2015)** CoRR, abs/1412.6980.

Authors Profile

Pradeep K R received his B.E and M. Tech. degree from Visvesvaraya Technological University, Belagavi, Karnataka, India. He is currently working as an Assistant Professor in Dept. of CSE, KSSEM and pursuing Ph.D. in computer science, VTU.

Dr. Naveen N C is a Professor and HOD, Department of CSE, JSS Academy of Technical Education, Bengaluru, India. He has completed Ph.D. from SRM University,Chennai. He has published more than 35 research papers in reputed international journals and conferences. His main research work focuses on Big data Analytics, Data mining and artificial intelligence.