Article

# Software Defect Detection and Prevention in Agile based process using Artificial learning methods

**Manikandan N [1,†], Ruby D [1,†,‡], Gayathri A [2,†,‡] and Chebrolu sai prasanna Abhinav [1,†]**

[1] School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India; mkyadhav@yahoo.com

[2] School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, India;

* Correspondence: mkyadhav@yahoo.com;

Version October 17, 2022 submitted to Journal Not Specified

**Abstract:** The currently available predicting approaches for establishing the reliability of software have become obsolete and static requiring additional manual maintenance for adjusting to the ever-evolving data sets pertaining to the Software organizations. Even though Artificial Learning (AL) is competent to address the issues of manually maintaining, certain changes are necessary as to the functioning of companies with defect- predictions. Agile methodology helps in getting the defect data on fly. this helps the prediction process more effective.This paper attempts to identify the diverse advantages of Artificial intellience against the existing approaches and the barriers for their adoption in practice. we have taken data from two existing projects and as a result, we ponder on the estimation of the impact of factors such as competence, and costs, in addition to accuracy, to assist the companies to arrive at informed decisions for the adoption of the techniques of ML for the prediction of the defect in the software.

**Keywords:** Software defect prediction. Deep learning, Software quality, Software reliability, machine learning.

## 1. Introduction

A ubiquitous and enhanced interest in the processes of automation has been on the rise in software industries with the advent and evolution of sophisticated automated tools. Eventually, this automation results in the generation of an enormous quantity of data which often remains unused. Early defect-detection can be advantageous in several areas such as time, schedules and cost by considering the points of extensive floods of data for applying to Artificial Learning (AL) and identifying the defects before their actual occurrence as per the agile process. Software defect-prediction becomes crucial in the analysis of diverse components and assessment of parts prone to defect. Several studies have been in progress for enhancing the prediction quality either for an individual project or for a cluster of projects. However, there is a lack of clarity in the required rules for prediction of errors, particularly while considering inter-project data assisted by heuristic data.

Method Initially, three categories of predictors were constructed on the basis of the six typical classifiers in three diverse scenarios, employing the size of the specific software metric set. Then, predictor acceptable performance is validated on the basis of Top-k metrics considering the statistical methods. Finally, the subset data of Top-k metric is minimized through the process of removal of redundant metrics. The minimum metric subset is tested for its stability on one-way ANOVA tests.

Expected Results 34 releases have been considered taken from 10 open-source projects at the PROMISE repository. As per the output findings, it is evident that the minimum metric subset or Top-k metrics can return an acceptable outcome in comparison with the benchmark predictors. Table 12 presents the procedures for the selection of an appropriate metric set in diverse situations.

The following are the indications from the experimental output results: 1. The specific accuracy requirement should be the metric for the selection of training data pertaining to defect-prediction. 2. In the event of limited resources, the predictor developed out of simplified metrics could be useful and function well. 3. The performance of even simple classifiers (e.g., Naive Bayes) could be considerable when simplified metric sets are employed for prediction of the defect. 4. In several situations, the minimum metric subsets could facilitate the processes of general prediction of the defect with permitted levels of prediction precision.

This paper organizes as follows: Firstly, Section 4.2 presents the literature review related to our work, Next, Section 3 deals with the description of the algorithm and other segmentation algorithms used for comparison. Later, Section 4 demonstrates the results. Finally, Section**??** concludes the paper.

## 2. Motivation

The prime purpose of this study is to authenticate the predictor's feasibility developed with a simplified set of metrics for defect-prediction of software in diverse situations and to study the practical and relevant principles for the selection of metric subset, training data, and classifier of a specific project.

### *2.1. Literature survey*

Different conventional classification Methodologies are depicted in this section.

In the literature, researchers suggested denoising methods to suppress the noise of an image without the degradation of the attributes of the original image. To reduce the noise from MR images, numerous denoising filters like bilateral, PCA, non-local means and bilateral can be utilised. Denoising filter analysis is carried out using various denoising techniques and revealed that the Spatially Adaptive Non-Local Means filter gives finest results than existing ones [1]. In general, clustering algorithms are classified into two types: hard and fuzzy [2]. The robust fuzzy c-means algorithm [3,4] is proposed with the modification in objective function of the conventional FCM, including the local spatial term allowing the computation of the smooth membership-grade. It improves the segmentation process and it's in-sensitive to noise only till a certain level. By considering an additional term in the objective function, the fuzzy clustering with spatial constraints [5] is proposed to allow smoothing of pixels by its neighbouring pixels to overcome the difficulty of the intensity in-homogeneity. This process is not noise sensitive, however it consumes more run time as it encompasses the computation of neighbouring pixels in complete iterations. The FCM-S1 and FCM-S2 in [6,7] are suggested to reduce the execution time of the FCM-S algorithm. FCM-S1 and FCM-S2 methods are used to compute mean and median filtered images to replace neighbourhood pixels of the FCM-S algorithm. An improved fuzzy c-means (Im-FCM) proposed [8], in acquaintance with the neighbourhood desirability term with its distance-measure, depending on two factors: (i) pixel intensities, (ii) spatial location of the adjacent pixels. The Im-FCM uses two parameters, whose finest values are acquired using ANN to amend the degree of two factors which needs more run time to acquire the parameter. The fast-generalized fuzzy-c-means (FGFCM) [9] combines the grey level as well as local spatial information using a similarity measure factor. The fuzzy local information c-means (FLICM) algorithm [10] stated that by adding a new fuzzy local neighbourhood factor in the objective function the intensity-level and neighbourhood relationship in the spatial domain can be found and parameter setting can be avoided. However, images are treated as fuzzy because of ambiguity resulting in classification, with regard to areas, boundaries and imperfect grey-levels. The fuzzy clustering is most often considered in-case of partial membership clusters of an element. In image segmentation, clustering considers image voxel as the data object, especially every voxel is assigned to a cluster depending on its similarity of prominent features [11]. While discussing the ambiguity, indecision arises in the image while defining the membership function in the hybrid algorithms [12] [13]. Ever since the membership grades are inaccurate and varies on individual's options, there are several kinds of uncertainties to an extent,

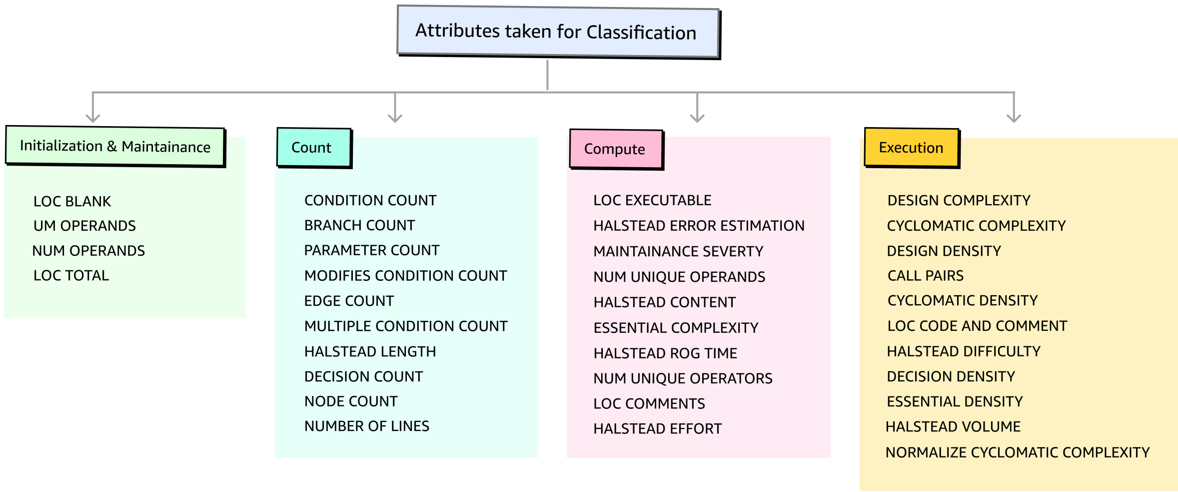


**Figure 1.** Attributes for defect prediction

arising due to absence of well-defined information in outlining the membership function. This leads to an definition of higher order fuzzy sets, referred as intuitionistic fuzzy set theory (IFS) proposed by Atanassov in 1983 [14], which considers the membership and the non-membership grades also. In an intuitionistic fuzzy set, because of the hesitation degree [15], the non-complement of the membership grade is greater than the membership grade. Many methodologies were introduced to mitigate the disadvantages of FCM. The basics of optimal sets is presented in [16] where the number of sets are not defined, by using the Shannon's entropy a standard function is hosted to capitalise the good points in the class. The type 2 fuzzy clustering is proposed [17] , with the ambiguity in a fuzzy set Type 2 membership by giving triangular membership grades for Type 1 fuzzy. The new grades are attained and cluster centres are improved using a standard FCM by taking the Type 2 fuzzy membership. The clustering technique [18–20] is defined in which an intuitionistic fuzzy similarity matrix is transmuted to interval valued fuzzy equivalence matrix, depending on the -cutting matrix of the intuitionistic fuzzy equivalence matrix. In [21], intuitionistic fuzzy sets are proposed based on the theory of association matrix as well as equivalent association matrix. Clustering is performed using the -cutting matrix of the related to association equivalence matrix. Hence, there is a need to develop an algorithm which solves this issue.

## 3. Proposed System model

### 3.1. Data Modeling

Main focus on data collection is towards getting the past records about the occurance of defect in various circumstances. It requires more amount of data for getting correct result. So data set is gathered from NASA MDP dataset with the following attributes. Details about the defect data set are presented in Figure. 1. With the data gathered, Rapid miner is used to populate the data distribution in disfferent classes. Following graph shows the Sample attributes like decision count, cyclomatic complexity, design complexity and LOC count are taken to populate the input data distribution.Same data is used in first step of prediction using various algorithms.

There are 27 attributes taken for the deep learning based enhancement. Fe w of the familiar algorithms are applied for predicting the defect.

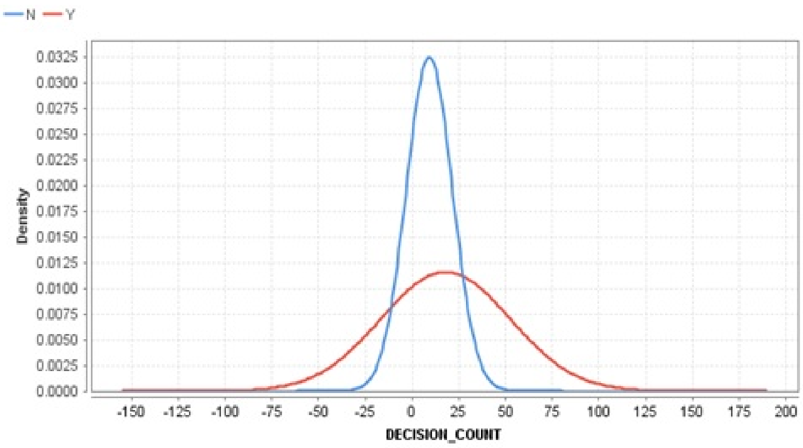


**Figure 2.** Data distribution for complexity and LOC

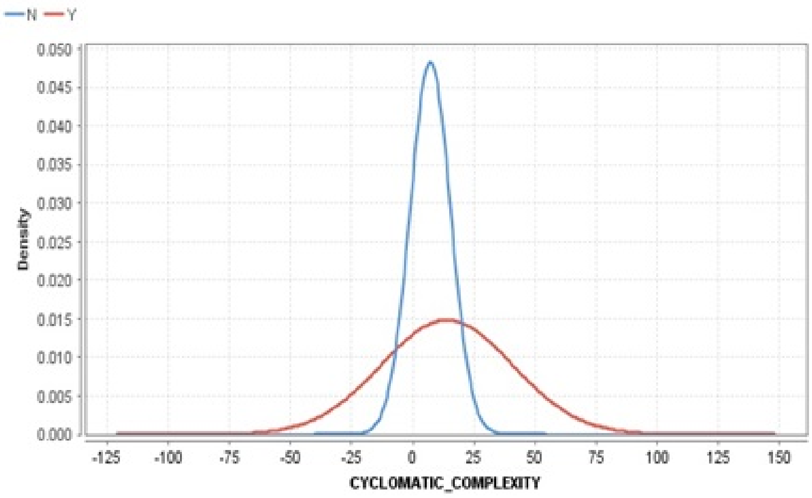


**Figure 3.** Data distribution for complexity and LOC

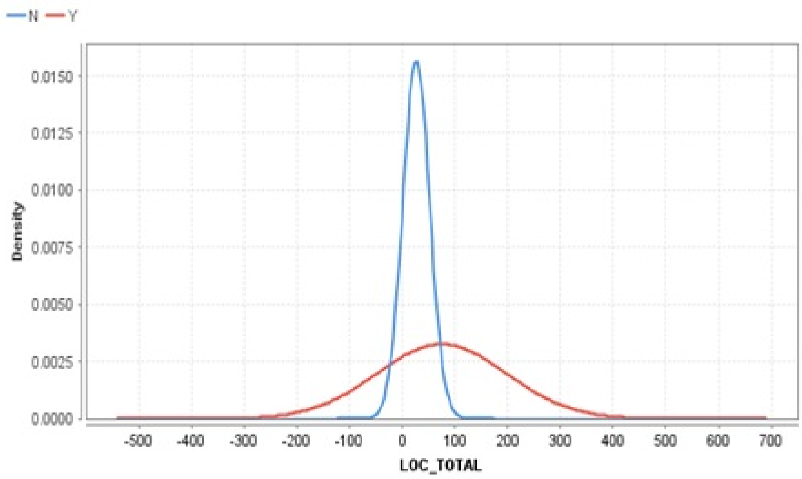


**Figure 4.** Data distribution for complexity and LOC

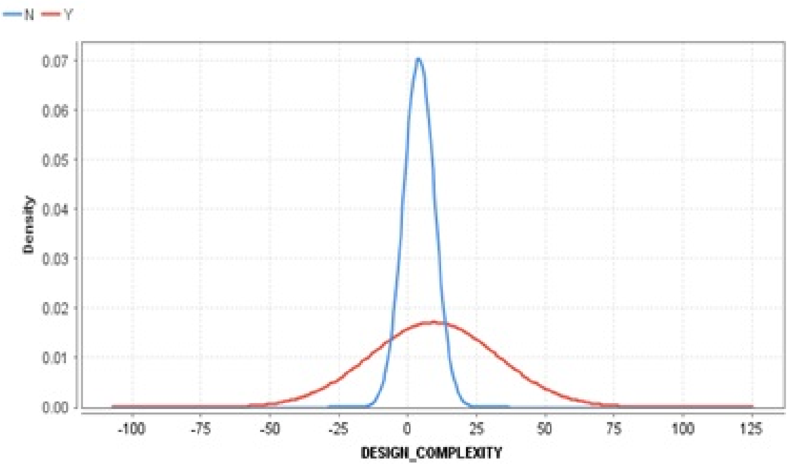


**Figure 5.** Data distribution for complexity and LOC

### *3.2. Algorithm Modeling*

#### 3.2.1. Naive Bayes Classifier

It is basically a classification technique based on Bayes' Theorem that assumes result based on independence among predictors. In easier terms, Naive Bayes classifier adopts that the incidence of a particular feature in a class is not related to the occurrence of any other feature. Though these features depend on each other. all these properties autonomously present to the expected probability. With respect to implementations, Naive Bayes model is easy to construct and useful for large data sets like defect data analysis. In addition to simplicity, Naive Bayes is familiar for its performance. Bayes theorem gives a solution to calculate posterior probability P(c∣x) from P(x) ,P(c) and P(x∣c). consider the following equation,Eq. 3.2.1

$$P(c \mid x) = \frac{p(x \,\|)P(c)}{p(x)}$$

(1)

Where, P(x)= Probability of prediction P(c)=Probability Postereior P(x∣c)=class probability posterior P(c∣x)= likelihood

#### 3.2.2. The Generalized Linear Model (GLZ)

This method is a variant after generalization of the general linear and In its simplest form, a linear model stipulates the linear relationship between a dependent and called as response variable O, and a group of predictor variables, X's, so that Eq. 3.2.2

$$O = a0 + a1x1 + a2x2 + .... + akxk$$

In above Eq. 3.2.2 a0 is the regression coefficient for intercept and the ai values are the regression coefficients computed from the data (for i=1 to k).

#### 3.2.3. logistic regression

The name Logistic regression came bcause of the logistic function used at the primary function of the method.Logistic function is also called the sigmoid function exclusively developed by statisticians to designate properties of population growth in ecosystem, it is an S-shaped curve that can accept

**Table 1.** Defect decision table

| Defective? | |
|---|---|
| Yes | No |
| 9 | 5 |

any real-time number and map it into a range between 0 and 1, but not ever precisely at those range. Consider the following Eq. 3.2.3

$$\frac{1}{1+e^{-value}}$$

Here, e- Base of natural log value, familiarly called as Euler's number. In the software defect prediction, Logistic regression used to model probability of the defective class. For eg. if the modeling defective as yes or know from their complexity, then the first class could be yes and the logistic regression model could be engraved as the probability of yes given for complexity of the code. It can be more formally written Eq. 3.2.3

$$P(defective = yes|complexityofcode)$$

It can be considered in another way, we are displaying the likelihood that an input (A) fits to the default class (Y=1). Again, It can be more formally written as Eq. 3.2.3

$$P(X) = P(Y = 1|X)$$

### 3.2.4. Decision tree

To construct a decision tree, we need to compute two types of entropy values using frequency tables. First, entropy calculation with one attribute Eq. 3.2.4

$$E(s) = \sum_{i}^{e} -P\log_2 Pi$$

the resultant value will be checked with following sequence Entrophy(defective)= Entrophy(5,9) Table 1

(0.36 * log2 0.36)-(0.64 log 2 0.64) =0.94

Second, Entropy calculation using two attributes.

E(defective, LOC)=P(LOC$_B LANK * E(3,2)) + P(LOC_C OMMENTS * E(0,4)) + P(LOC_T OTAL) * E(2,3))$

=(5/14)*0.971 +(4/14)*0.0 +(5/14)*0.971 =0.693

In this paper two attribute method is used in the above manner for prediction.

### 3.2.5. Gradient Boosted Trees

Gradient boosting is a familiar machine learning methodology for regression as well as classification problems. it produces a prediction model as an group of weak prediction models

**Table 2.** Defect decision table

typically like decision trees. Objective of selecting any supervised learning algorithm is to describe a loss function and minimize it. Following equaton is used to define

3.2.6. Random Forest

Random forest algorithm is on of the supervised classification algorithm.This algorithm generates the forest with a number of trees.The more trees in the forest the more robust the forest looks like is generated. The random forest classifier with the higher the number of trees in the forest provides the high accuracy results. • Random forest classifier handles the missing values. • The random forest algorithm / classifier can use for both the regression task and classification. • If more trees in the forest, then random forest classifier won't become overfit . • The random forest classifier for categorical values. Random Forest pseudocode

Step 1:Randomly select "k" features from total "m" features.

Step 2: Where k « m

Step 3:Among the "k" features, calculate the node "d" using the best split point.

Step 4:Split the node into daughter nodes using the best split. Repeat 1 to 4 steps until "l" number of nodes has been reached. Build forest by repeating steps 1 to 4 for "n" number times to create "n" number of trees.

3.2.7. Support Vector Machine

SVM is a supervised machine learning algorithm which can be used for both classification or regression challenges. However, it is mostly used in classification problems. In this algorithm, we plot each data item as a point in n-dimensional space (where n is the number of features you have) with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiate the two classes very well.

1. TRAININGSET xi, yi, i=1..l
2. WHEIGHTS qi, i=1..l
3. BIAS b
4. TRAININGSET PARTITION INTO SUPPOTSET(S) , ERRORSET(E) AND REMAININGSET(R)
5. PARAMS: e , C , KERNELTYPE AND KERNELPARAMs
6. R MATRIX
7. NEW SAMPLE C = (xc, yc)

Prediction results from various algorithms are plotted in table.2

## 4. implementation analysis and Results

### *4.1. Neural network construction*

To implement the algorithms with defect data set in neural network with appropriate input and hidden layer should be formed. As per the results shown in tabe 2, results of deep learning compared with other techniques with all 27 attributes listed. When forming the deep learning, few attributes that are not directly involved in the classified results can be removed. As the firs of this process, greedy based attribute greedy attribute selection to select the appropriate attributes to reduce the error rate of deep learning. Modified approach is presented in Fig. 6.

| Models | Accuracy | Classification error | Sensitivity | Specifity |
|---|---|---|---|---|
| Naive Bayes | 89.8 | 10.2 | 42.7 | 93.4 |
| Generalized Linear Model | 92.2 | 7.8 | 15.0 | 99.0 |
| Logistic Regression | 91.7 | 8.3 | 22.0 | 98.0 |
| Deep Learning | 88.5 | 11.5 | 48.0 | 92.4 |
| Decision Tree | 91.7 | 8.3 | 0 | 99.5 |
| Random Forest | 91.7 | 8.3 | 0 | 99.5 |
| Gradient Boosted Trees | 92.2 | 7.8 | 25 | 98.0 |
| Support Vector Machine | 91.7 | 8.3 | 0 | 99.5 |

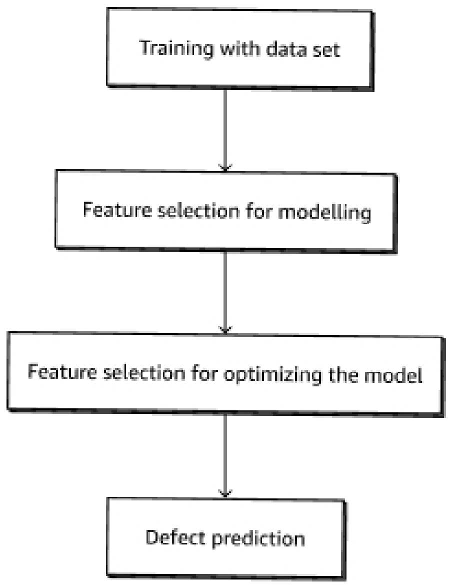


**Figure 6.** Data distribution for complexity and LOC

**Table 3.** Greedy attributes setting

### *4.2. Deep learning with feature Reduction*

In this paper, we have taken few classification algorithms against the performance improvement of deep learning to classify the possibilities of defect. Following figure 2 represents the model for deep learning. Training data set with the attributes listed in table 1 is taken as input.

**Algoriithm 1**

1. Load the data set with attribute X0 to Xn Problem domain D.
2. Define a attribute reduction Function F a sub set of domain D and it is called as Submodular.
3. For every S and T belongs to D, F(S)+F(T)F(SUT) + F(ST)
4. The greedy algorithm constructed with a set by incrementally accumulating the element that increases the highest at each step. Output of this process is a set that is at least (1-1/e)max(F(M)) where MD and (1-1-e)0.64
5. Reduced attributes are stored in input layer Ii Where i= 1 to number of attributes

feature selection plays vital role in improving the performance of classification algorithm. The algorithm fails to scale up size of the sample over time. To understanding the domain in easy and better way also cheaper to collect reduce set of predictors, the classification algorithm is used.

1. The candidate set, from which a solution is produced.
2. The selection function, which selects the best candidate to be further to the solution
3. The viability function, that is used to find out, if a candidate can be used to give to a solution
4. The objective function, which gives a value to a solution. The solution function, which will designate when we have revealed a complete solution.

### *4.3. Greedy Stepwise Algorithm*

Greedy Stepwise Algorithm with forward or backward search over and done with the space of attribute subsets. Possibly will start with no or all attributes from an arbitrary point in the space. Stops when the deletion or addition of any attributes remaining will be results in a decrease in evaluation. The ranked list of attributes by traversing one side to the other side of the space. Recording the order that attributes are selected and results are produced.

The table. 3 describes the selections obtainable for Greedy Stepwise. Over all outcome shows that direct implementation of nenral network algorithms produces better accuracy than deep learning algorithms. So it requires special process to to reduce the less related attributes. Consider the following technique By using all the above attributes against the 3 different prediction paramenters we obtained the following result. The set of predictors with V features, the target variable T, to find out min set F with max classification performance over T is given. The

A sample output sequence for th execution result of Greedy stepwise forward Feature selection is given below

Evaluator: attributeSelection.CfsSubsetEval
-1.7976931348623157E308 -N -1
Relation: PC1
Instances: 759
Attributes: 38
Evaluation mode:evaluate on all training data
=== Attribute Selection on all input data ===
Search Method:
Greedy Stepwise (forwards).

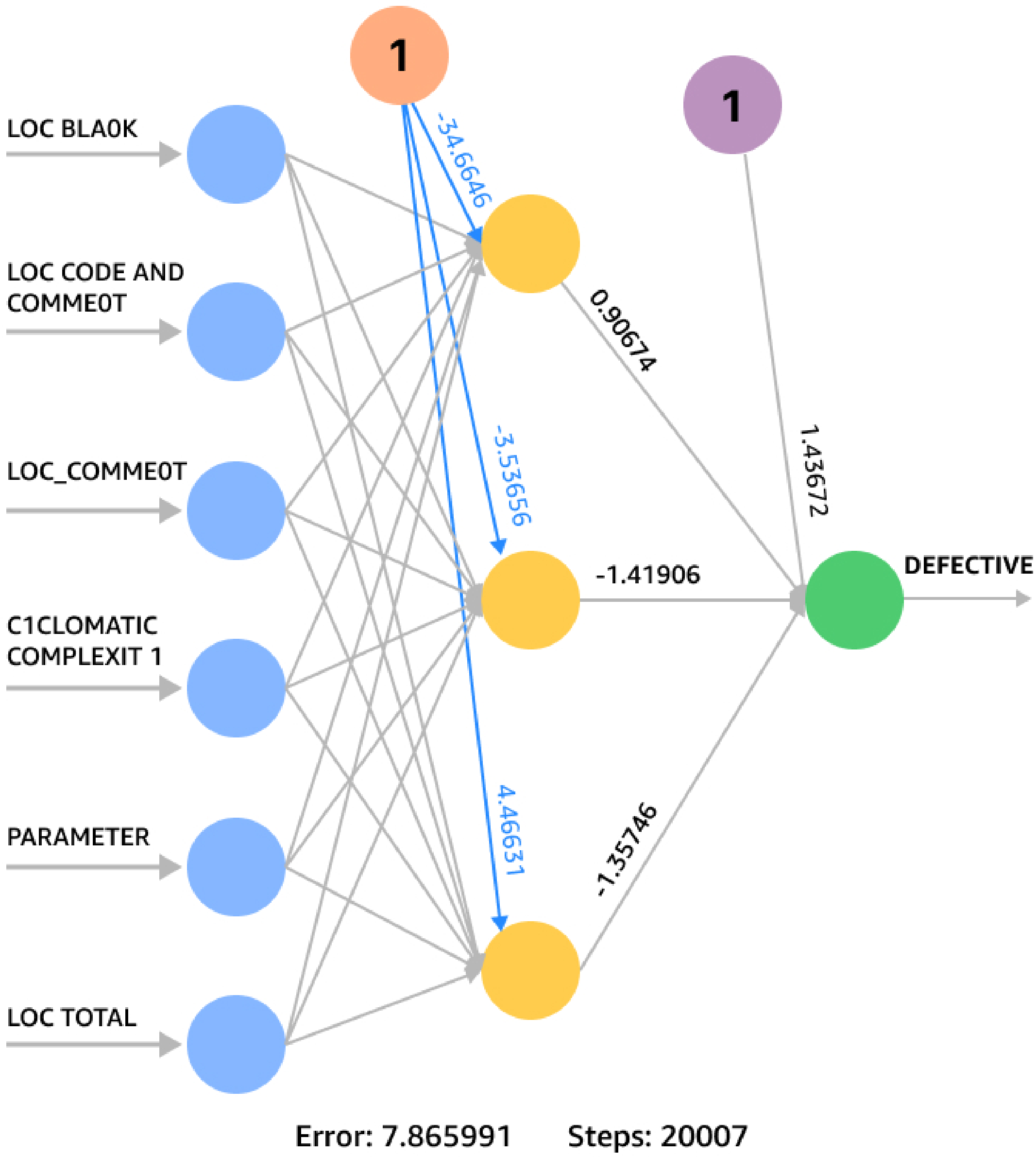


**Figure 7.** Modified Neural network after attribute selection

| Selections | Description |
|---|---|
| ConservativeForwardSelection | If true selection with forward search is choosen,the attributes will remain to be added to the best subset during the whole time of merit does not degrade. |
| Generate Ranking | Set the value to true if a ranked list is essential. |
| Num to select | Specify the number of attributes to retain. If the default value is set to be -1, It indicates that total attributes are to be retained. |
| Search backwards | Search backwards rather than forwards. |
| Start set | The start point is specified as a comma seperated list off attribute indexes beginning at 1. |
| Threshold | Set threshold when attributes need to be discarded. The default value will results in no attributes being rejected. |

Start set: no attributes
Merit of best subset found: 0.15
Attribute Subset Evaluator (supervised, Class (nominal): 38
Defective):
CFS Subset Evaluator
Including locally predictive attributes
Selected attributes: 1,4,5,17,18,30,33,37 : 8
LOC$_{B}LANKLOC_{C}ODE_{A}ND_{C}OMMENT \quad LOC_{C}OMMENTS$
$PARAMETER_{C}OUNT \quad HALSTEAD_{C}ONTENT$
$NORMALIZED_{C}YLOMATIC_{C}OMPLEXITY \quad NUM_{U}NIQUE_{O}PERANDS \quad LOC_{T}OTAL$

sample result of the process carried out is,

Description: Metrics reported on full training frame
model id: rm-h2o-model-deep$_{l}earning - 416142$
$frameid : rm - h2o - frame - deep_{l}earning - 448037$
$MSE$ : 0.05728833
$R^2$ : 0.23161286
$AUC$ : 0.8822808
$logloss$ : 0.19972481

CM: Confusion Matrix (vertical: actual; across: predicted):
N Y Error Rate
N 409 10 0.0239 = 10 / 419
Y 24 13 0.6486 = 24 / 37
Totals 433 23 0.0746 = 34 / 456
Gains/Lift Table (Avg response rate: 8.11
The results produced so far is the simplified CNN with different NN algorithms support. One clear point in all convolution approaches are it requires more time and attributes. Accuracy and error rate are not considerably good when predicting with minimum and essential attributes. So Greedy feature reduction is applied and fter feature reduction the following results were obtained.

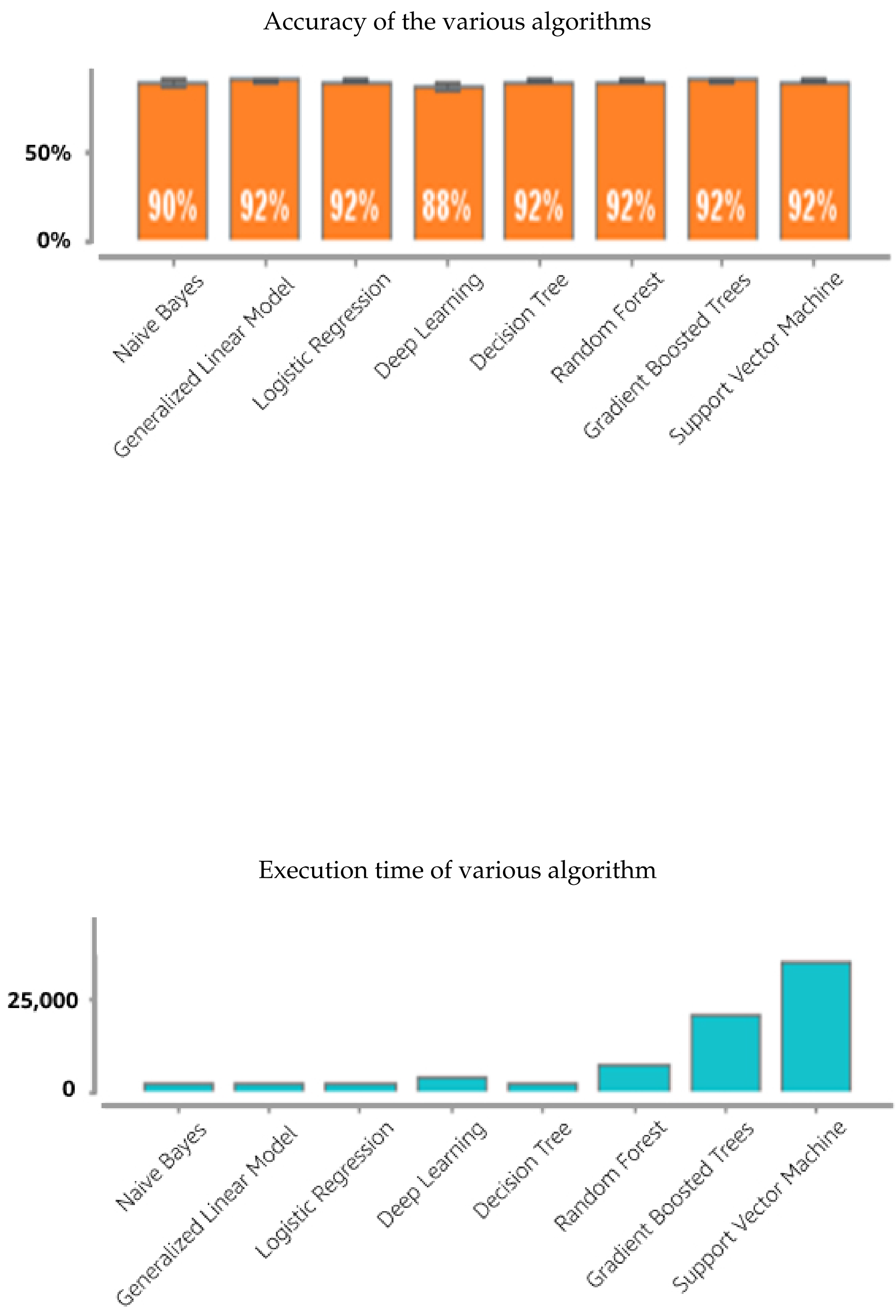


**Figure 8.** Resuls of traditional neural network

## 5. Conclusion

To estimate the effectiveness of the segmentation of the MRI brain tissue and tumours, a novel deep learning method Gaussian kernelised improved intuitionistic fuzzy c-means algorithm (GKEIFCM) is proposed. To decrease the spatial distance between the pixels, Gaussian kernelized distance metric is used. This modification proved that GKEIFCM deep learning algorithm is efficient in segmenting the tissues and tumours with higher accuracy, DSI and JI metrics. Our hybrid technique has provided accurate and reliable segmentation results than FCM, FCM with spatial functions S1 and S2, FLICM, IFCM, IIFCM techniques. With a factor 98.54% of Accuracy, 98.37% of Dice-Similarity-Index, and 96.68% of Jaccard-Index for 3% of noise is attained with the proposed algorithm. From performance analysis, it is evident that the proposed hybrid GKEIFCM segmentation algorithm is robust to noise artefacts and also can take care of shielding effects and correctly locating the tumour in affected regions.

## References

1. Saladi, S.; Amutha Prabha, N. Analysis of denoising filters on MRI brain images. *International Journal of Imaging Systems and Technology* **2017**, *27*, 201–208.
2. Yepuganti, K.; Saladi, S.; Narasimhulu, C.V. Segmentation of tumor using PCA based modified fuzzy C means algorithms on MR brain images. *International Journal of Imaging Systems and Technology* **2020**, *30*, 1337–1345.
3. Monalisa, A.; Swathi, D.; Karuna, Y.; Saladi, S. Robust intuitionistic fuzzy c-means clustering algorithm for brain image segmentation. 2018 International Conference on Communication and Signal Processing (ICCSP). IEEE, 2018, pp. 0781–0785.
4. Pham, D.L. Spatial models for fuzzy clustering. *Computer vision and image understanding* **2001**, *84*, 285–297.
5. Ahmed, M.N.; Yamany, S.M.; Mohamed, N.; Farag, A.A.; Moriarty, T. A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data. *IEEE transactions on medical imaging* **2002**, *21*, 193–199.
6. Kumar, N.P.; Sriram, A.; Karuna, Y.; Saladi, S. An improved type 2 fuzzy C means clustering for MR brain image segmentation based on possibilistic approach and rough set theory. 2018 International Conference on Communication and Signal Processing (ICCSP). IEEE, 2018, pp. 0786–0790.
7. Chen, S.; Zhang, D. Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **2004**, *34*, 1907–1916.
8. Shen, S.; Sandham, W.; Granat, M.; Sterr, A. MRI fuzzy segmentation of brain tissue using neighborhood attraction with neural-network optimization. *IEEE transactions on information technology in biomedicine* **2005**, *9*, 459–467.
9. Cai, W.; Chen, S.; Zhang, D. Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation. *Pattern recognition* **2007**, *40*, 825–838.
10. Krinidis, S.; Chatzis, V. A robust fuzzy local information C-means clustering algorithm. *IEEE transactions on image processing* **2010**, *19*, 1328–1337.
11. Ouchicha, C.; Ammor, O.; Meknassi, M. A new approach based on exponential entropy with modified kernel fuzzy c-means clustering for MRI brain segmentation. *Evolutionary Intelligence* **2022**, pp. 1–15.
12. Arulselvarani, S.; Manimekalai, S. Brain Tumor Segmentation & Detection of Mr Images Using Intuitionistic Fuzzy Clustering Mean (Ifcm). *Annals of the Romanian Society for Cell Biology* **2021**, *25*, 7669–7680.
13. Saladi, S.; Amutha Prabha, N. MRI brain segmentation in combination of clustering methods with Markov random field. *International Journal of Imaging Systems and Technology* **2018**, *28*, 207–216.
14. Atanassov, K. Intuitionistic fuzzy sets. *International journal bioautomation* **2016**, *20*, 1.
15. Saladi, S.; Prabha-Nagarajan, A. A novel fuzzy factor for MRI brain image segmentation using intuitionistic fuzzy kernel clustering approach. *Journal of Advanced Research in Dynamical and Control Systems* **2018**, *10*.
16. Ferahta, N.; others. New fuzzy clustering algorithm applied to RMN segmentation. *IEEE Transactions on Engineering, Computing, and Technology* **2006**, *12*, 9–13.
17. Rhee, F.C.H.; Hwang, C. A type-2 fuzzy C-means clustering algorithm. Proceedings joint 9th IFSA world congress and 20th NAFIPS international conference (Cat. No. 01TH8569). IEEE, 2001, Vol. 4, pp. 1926–1929.

18. Zhang, H.m.; Xu, Z.s.; Chen, Q. On clustering approach to intuitionistic fuzzy sets. *Control and Decision* **2007**, *22*, 882.
19. Kala, R.; Deepa, P. Spatial rough intuitionistic fuzzy C-means clustering for MRI segmentation. *Neural Processing Letters* **2021**, *53*, 1305–1353.
20. Dahiya, S.; Gosain, A. A novel type-II intuitionistic fuzzy clustering algorithm for mammograms segmentation. *Journal of Ambient Intelligence and Humanized Computing* **2022**, pp. 1–16.
21. Xu, Z.; Chen, J.; Wu, J. Clustering algorithm for intuitionistic fuzzy sets. *Information sciences* **2008**, *178*, 3775–3790.