VOL. 18, NO. 5, MARCH 2023 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

A COMPREHENSIVE MACHINE LEARNING TECHNIQUE FOR OVARIAN CANCER PREDICTION

V. Mnssvkr Gupta, K. Vssr Murthy, R. Shiva Shankar and J. Varshini
Department of Computer Science and Engineering, Sagi Rama Krishnam Raju Engineering College, Bhimavaram, Andhra Pradesh, India
E-Mail: guptavkrao@gmail.com

ABSTRACT

On the list of most prevalent cancers in women, ovarian cancer comes in eighth place. Because of standard screening and surveillance limitations, it is clinically impractical to analyse tumour molecular markers to predict therapy response. A relevant dataset with the necessary features was chosen for the prediction task of whether the patient has ovarian cancer or not. This approach uses Multilayer perceptron, ELM with AdaBoost, XGboost, LSTM, and a new CNN with Random Forest algorithms to predict Ovarian Cancer. The assessment parameters for each model were calculated for accuracy, precision, recall, F1-Score, Jaccard Index, and Error rate. The algorithms are then compared based on these metrics to determine the best algorithm. The CNN with Random Forest was the best method, with 95 to 100% accuracy. The CNN with Random Forest algorithm outperformed the individual existing ensemble techniques studied.

Keywords: ovarian cancer (OC), machine learning (ML), convolutional neural network (CNN), multi-layer perceptron (MLP), random forest (RF), ensemble algorithms (EA).

1. INTRODUCTİON

Cancer is a disease in which somebody's cells develop uncontrollably and spread throughout the body. These cells have the potential to form a tumour or malignant mass. The most common cancers in women are breast, ovarian, and cervical. OC is less well-known than other types of cancer in women and when cells in the ovaries become malignant. There are usually no symptoms in the early stages of OC. Symptoms are associated with the later stages. However, they can be vague, such as back pain and weight loss [1]. OC is uncommon, yet it kills more women than other female reproductive malignancies. The earlier OC is detected and treated, the better your chances of a successful recovery. Despite these difficulties, detecting OC at an early stage is possible [2]. After the condition has advanced, women with OC may have no symptoms or very moderate ones. As a result, it becomes more challenging to treat, and other symptoms may also be, such as a heavy sensation in the pelvic and lower abdominal discomfort. Having vaginal bleeding, Gaining or losing weight, Exceptional periods, Back discomfort that doesn't seem to go away, Gas, nausea, vomiting, or an inability to eat [3].

Women diagnosed with OC will have their condition evaluated by doctors to see whether and how far the disease has spread. According to the cancer stage, the body's cancerous cells are distributed throughout. It helps determine the malignancy's degree and the patient's best therapy choices. When discussing cancer survival statistics, doctors also consider the disease's stage. OC is classified into various stages, I (1) through IV (4). In general, less cancer has spread when the number is low. Higher numbers, such as IV, indicate an advanced stage of cancer. Although every person's experience with cancer is different, tumours at comparable stages often respond similarly to treatment and have similar outcomes [4]. Various OC types are shown in Figure-1.

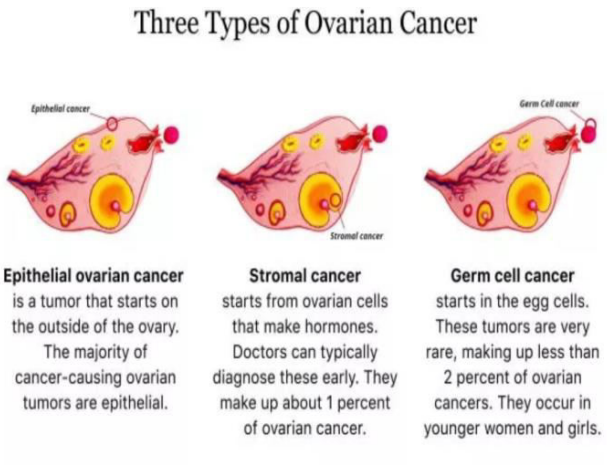

Figure-1. The types of ovarian cancer.

Ovaries, or female reproductive glands, are only found in females (women). Eggs are produced in the ovaries for reproductive purposes (ova). The fertilised eggs migrate from the ovaries to the uterus through the fallopian tubes before settling and developing into a baby. The ovaries are primarily responsible for producing the feminine hormones progesterone and oestrogen. Ovaries are located on each side of the uterus. All three kinds of cells that make up the ovaries are found there. Depending on the cell type, a tumour may take on various forms. Ovarian epithelial malignancies develop from the cells lining the ovaries' outer surface. It's the most common ovarian tumour, accounting for around 80% of all cases. Ovarian cells [5] give rise to cancerous germ cell growth. Stromal tumours are created by the structural tissue cells that produce the female hormones oestrogen and progesterone and keep the ovary together. Some of these tumours are malignant but never move beyond the ovary.

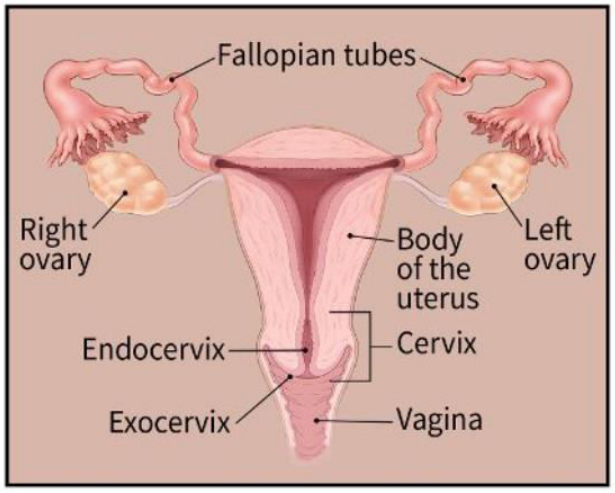

Figure-2. Female ovaries.

Cancerous tumours can spread to other body parts and cause death, whether malignant (cancerous) or borderline (low malignant potential) ovarian tumours. During a pelvic exam, the ovaries and uterus are examined for size, shape, and consistency. The Female Ovaries are shown in Figure-2. In some instances, a pelvic exam may detect tumours in women early; however, most early ovarian tumours are difficult or impossible to feel with a pelvic exam. Various malignancies and illnesses specific to women may be detected during a pelvic exam. Women should talk to their doctors about whether or not they need to get these tests [6].

2. RELATED WORK

According to Lu M *et al.* [7], ML models may be utilised to identify complicated disorders like OC. Based on two biomarkers (Human Epididymis Protein 4 [HE4] and Carcino Embryonic Antigen [CEA]), researchers discovered that a simple decision tree model could correctly identify both BOT and OC. According to the data, 349 Chinese patients were studied using 49 factors, including demographics, blood routine tests, general chemistry, and tumour indicators, which were all tested. For patients with epithelial ovarian cancer (EOC), Pail ES et al. [8] want to use gradient boosting (GB) to create a novel prognostic classification and compare the accuracy of this new model to that of the standard statistical technique. The effectiveness of the final model was externally tested using patient data from Asian Medical Center (validation cohort, n=229) (validation cohort, n=229). The goal of Akazawa M and Hashimoto K [9] was to utilise AI to identify OC pathological diagnoses using patient data from preoperative exams. Numerous reports in OC combine the evaluation of medical images with multi-omics analysis of clinical and genomic data powered by AI.

Kawakami E *et al*. [10] created a cancer of the ovary–specific prediction framework that considers clinical stage, histotype, the amount of remaining tumour burden, and prognosis. It was expected that ML systems would provide critical diagnostic and prognostic information before initiating intervention in the EOC patient population. According to Lee Z. J [11], microamy includes an infinite quantity of data. Genetic indicators may help explain some of the consequences of OC. Gene selection and classification have been combined in OC microamy; this algorithm is called OC. Researchers Elhoseny M *et al*. [12] have discovered a kind of cancer that exclusively affects women. It's challenging to detect cancer at an early stage. One of the essential factors in differentiating OC SOM will be ORNN. AHSO will play a significant role in the search for the OC. Negi S *et al.* [13] claim that HBL OLED offers a 47 %increase in energy efficiency compared to multi-layered OLED architecture. A high electron injection rate and efficient hole blocking were two additional variables that contributed to the 74% improvement in luminous power efficiency. Because of this, the Triple OLED is used for OC diagnostics. Mostavi M *et al.* [14] report that OC is the most often diagnosed malignancy in women nowadays. Using the SIFT method, we can locate and monitor the spread of this cancerous condition.

According to Dong J *et al*. [15], the quantitative evaluation of immunohistochemical images requires a significant contribution from the Division of Immunohisto chemistry picture. They developed an automated segmentation approach based on the characteristics of colour immunohistochemical pictures. The results of the experiments reveal that the strategy effectively divides OC immunohistochemical images into smaller segments. Using protein mass spectra to identify biomarker patterns has piqued the attention of several researchers, including Assareh A and Moradi MH [16]. The difficulties raised by these studies must be addressed before mass spectra-based proteomic patterns may be used. The selection of biomarkers from the high-dimensional input data is unquestionably the most critical aspect of any design. According to a study by Elhoseny M *et al*. [17], OC has a high death rate since it affects women's ovaries and is challenging to identify in the early stages. It is possible to detect OC using the OC data supplied by the Internet of Medical Things. Feature selection using the SOM method was found to be more accurate, as was the separation of medical data into useful metrics. On top of all that, we use an optimum classifier-optimal recurrent neural network (ORNN).

According to Rahman MA *et al.* [18], OC is a deadly condition for older women. Many studies using Artificial Neural Networks have been conducted on OC (ANN). For clinicians to make an informed judgement categorisation accuracy is critical. It can save lives and lower death rates by obtaining an accurate diagnosis as early as possible. The suggested model has a classification accuracy of 98.7% for OC, which is more promising and higher than existing classification methods.

S.S Reddy *et al* [19] worked on the problems of women who are suffering from various problems like Gestational Diabetes [20], and Breast cancer [21-23] by using various machine learning Algorithms. Diabetes mellitus was predicted [24] by using a voting strategy [25], with correlated ailments [26] and multiple ailments

[27]. Chung Shan *et al*. [28] hope their study combines sophisticated ML approaches with ensemble learning, typically considered the most successful way of creating an objective for an inferential issue of recurrent OC. Five ML algorithms, including SVM, C5.0, ELM, MARS (multivariate adaptive regression splines), and RF, were utilised to identify the essential risk variables and predict OC recurrence study. According to Mansi Mathur [29] and others, the ovaries are the primary energy source for the female reproductive system. A large part of the development of female sex hormones is due to the relevance of these tiny glands. When the ovarian ligaments connect two microscopic glandular bodies, they become one unit. There are many possible causes of obsessive-compulsive disorder, but a simple blood test can identify it. A distinction between ordinal labels and nominal labels with no natural ordering was introduced by Misganaw B and Vidyasagar M [30] for multiclass ML issues. Here are a few instances of ordinal labels that are frequent in biology. Using order information when it is present in the issue emphasises the relevance of this comment. According to Park J S et al., OC is the sixth leading cause of mortality for women in the United States [31]. Borderline severe ovarian tumours (SBOTs) are considered less malignant or early serous ovarian carcinomas (SOCs). It is common for SBOTs to proceed to later stages despite the lack of any symptoms. Biomarker genes for OC development and progression have been identified utilising multicategory ML algorithms applied to DNA microarray data, including identifying the SNTN and AOX1 markers.

Yaar *et al*. [32] suggested that a new deep neural network architecture for predicting chemo-sensitivity in OC patients be discovered in this study. Using MIL and a new type of Learning with Privileged Information, the proposed method (LUPI). Input space features that can be used for training and inference may benefit from LUPI's ability to transfer information from highly informative privileged features that are only accessible during training. According to PouryaNaderi and Taghi Mostafavi [33], OC is the leading cause of mortality among women in the United States. Diagnostic molecular biomarkers for OC have been presented as a crucial way to decrease death rates. Research into the tissue expression-to-serum indicators for HOXC6 and PAX8 gene expression may be conducted using techniques similar to those employed in previous studies. Female reproductive cancers like OC are particularly deadly since there are no detectable signs or early-stage screening techniques, according to a study by Xiaoyan Yang *et al.* [34]. Detection occurs when cancer has reached an advanced stage and has a poor chance of recovery. Mingyan Zhou *et al*. [35] stated that patients might suffer catastrophic injury if ovarian endometriomas (OEs) burst spontaneously. They created the PSO-RF model, a PSO-enhanced RF model, to assist in the preoperative diagnosis of ovarian endometriomas spontaneously ruptured.

3. METHODOLOGY

3.1 Objectives

- A relevant dataset with the necessary features was chosen for the prediction task of whether the patient has OC.
- The patient's cancer stage can be identified through the dataset's features.

3.2 Dataset Description

The dataset has 51 features and 350 records with a binary classification dataset. The target variables have two classes, OC detected, and Normal detected.

Table 1: Description of the Features

Feature	Description	Feature	Description
Age	Age of the patient in years	BASO#	Is a blood test a normal range of 5% to 1%
AFP	Aipha Fetoprotein value is given in ng/ml, and the normal range is 10 t0 20 ng/ml	ALT	Alanine transferase value is given in IU/L, and the normal range is between 0 to 45 IU/L
ALB	Ablumin, value is given in IU/L, and the normal range is between 34 to 54 IU/L	AST	Aspartate aminotransferase value is given in the normal range as between 0 to 35 IU/L
BASO%	Is a blood test a standard range of 5% to 1%	CA	Cancer and normal carcinoma range from 8.6 to 10.3
BUN	Blood urea nitrogen normal range 10 to 1	CA125	Cancer antigen 125 normal range 0 to 3.5
CEA	Whether the patient is tested positive or negative for Carcinoembryonic antigen	CA72-4	Whether the patient has a tumour marker or not
C02CP	C02CP normal range 23 to 29 mmol/L	CL	Chloride test normal range 96 to 106 mEQ/L
EO#	Ethylene oxide normal range 30 to 350	DBIL	Direct bilirubin normal range 0.3mg/dL
GGT	Gamma-glutamyl transferase in the blood range from 5 to 40 U/L	EO%	Ethylene oxide normal range 30 to 350

A detailed description of the dataset is given in Table-1. The blood test components are OC contains many attributes like AFP, ALP, ALT, AST, BASO#, BASO%, and BUN. Age is one factor for OC due to the changes in the menstrual cycle. We can predict OC in Women using the dataset's attributes.

3.3 Existing Algorithms Used

A. MLP Algorithm

A feed-forward neural network enhancement known as an MLP is a multi-layer perceptron (MLP). It comprises three parts: an input layer, an output layer, and a concealed-layer component. After being processed, the output signal is sent to and received by this layer. Genotype data are included in the list of OC patients in the input. Prediction and categorisation are within the purview of this layer. The MLP's real computational engine is a slew of hidden layers sandwiched between the input and output, as shown in Figure 3. When using an MLP, data is sent forward from input to output, much like a feed-forward network. The MLP's neurons are trained using the backpropagation learning technique.

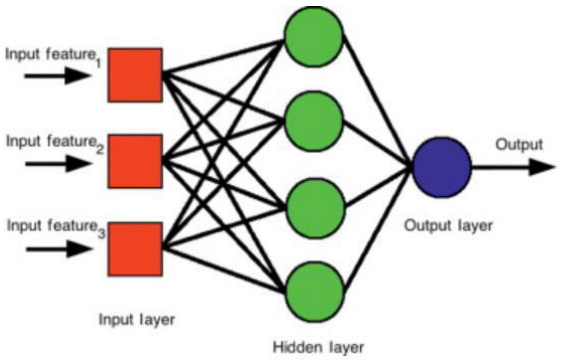

Figure-3. MLP Algorithms architecture.

Three layers are used in the MLP algorithms: an input layer, a hidden layer, and an output layer. A simple distribution of the dataset's characteristics occurs in the input layer, which is then used by the hidden layer. The hidden layer may then predict the values to be produced.

B. ELM with AdaBoost and XgBoost Algorithm

As an Ensemble technique, the AdaBoost algorithm, known as AdaBoost for short, which stands for Adaptive Boosting, was shown in Figure-4.

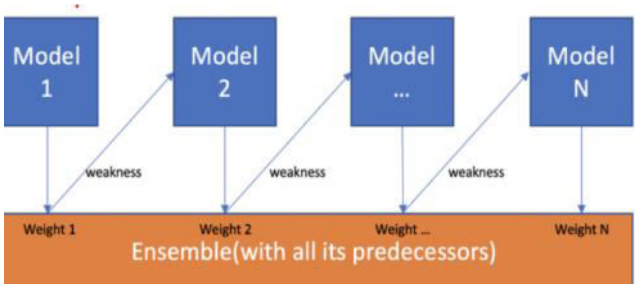

Figure-4. Ada boost algorithms architecture.

The weights are reallocated to each instance, with greater weights assigned to incorrectly recognised cases. In supervised learning, boosting is used to decrease bias and variance. Based on the idea that one learns by doing, it's a progression. XgBoost is a scalable GBDT ML toolbox that distributes gradient-boosted decision trees across nodes. Regression, classification, and ranking are all made easier with this ML package, which also contains parallel tree boosting. Xg boost architecture was shown in Figure-5.

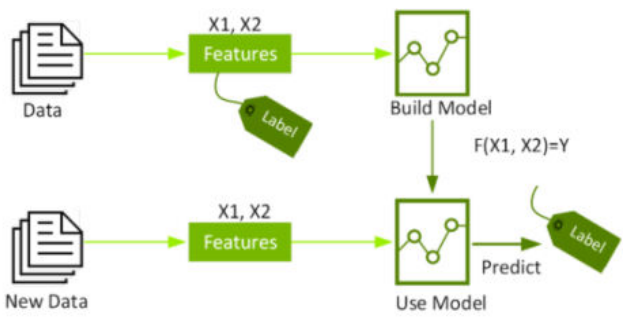

Figure-5. XgBoost algorithms architecture.

OC may be predicted using this model, which uses Ada and Xg boost techniques. Adaboost re-assigns weights to each instance, adding heavier weights to incorrectly recognised instances. Xg Boost relies on an ensemble machine learning framework to forecast unstructured data that uses gradient boosting to build decision trees and predict.

C. LSTM Algorithm

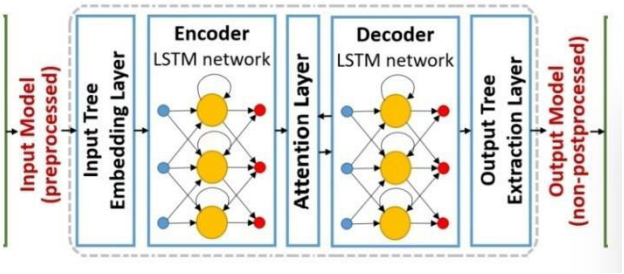

Figure-6. LSTM Algorithms architecture.

The Long Short-Term Memory Network (LSTMN) is an improved version of the regular RNN. RNN's difficulty with vanishing gradients is no longer an issue. At a high level, LSTM is identical to an RNN cell in terms of functionality. The LSTM has three unique portions, each serving a specific function. The previous timestamp's information should be remembered or destroyed in the first component of the rule. In the second phase, the cell tries to pick up additional information from its input.

The LSTM algorithms are divided into three components, as shown in Figure-6. The first portion determines whether or not the initial timestamp's information is to be remembered or discarded. In the second section, the cell attempts to learn new data from the input that this cell receives. Finally, in the third section, the cell transmits the current timestamp's modified information to the next timestamp.

D. Proposed (CNN with Random Forest) Algorithm

A CNN is a DLNN designed to analyse structured data arrays like representations. CNN's are multi-layered FFNNs constructed by stacking multiple layers on top of each other in a predetermined sequence.

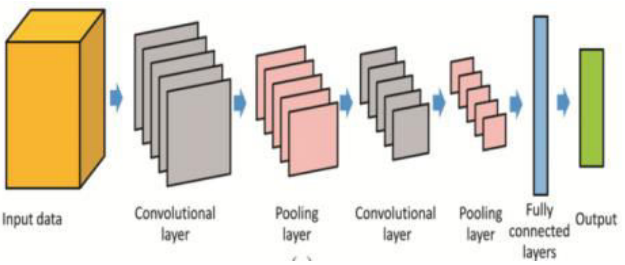

Figure-7. CNN Algorithms architecture.

Because of its sequential structure, CNN is capable of learning hierarchical characteristics. Grouping and hidden layers are commonly followed by convolutional layers in CNNs, whereas convolutional layers often follow activation layers.

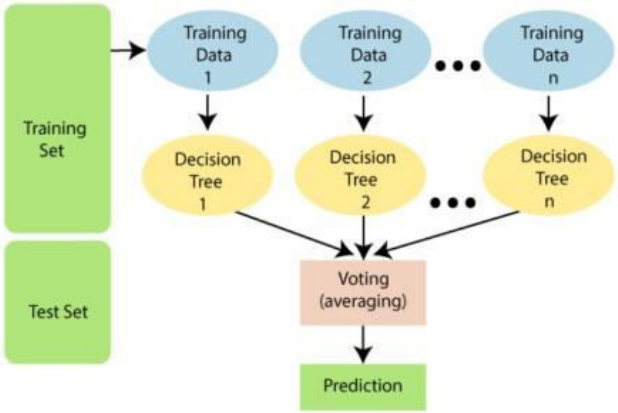

Figure-8. Random forest algorithms architecture.

Random forest

It is healthy knowledge that the well-known Random Forest ML algorithm uses supervised learning methods. It may be used for both classification and regression in ML. Ensemble learning is a technique for combining many classifiers.

Using CNN and Random Forest Algorithms, we extract features from CNN and then retrain these retrieved features using the Random Forest algorithm and the experiment results with CNN and Random Forest by OC patient information for prediction. Figure-7 and Figure-8 are the Architectures for CNN and Random Forest.

Proposed Algorithm 1:	
INPUT:	Ovarian Cancer Dataset
OUTPUT:	Predictions made for the Input Data
Step1:	Import Libraries and Pre-Processing techniques
Step2:	Splitting dataset of Ovarian Cancer into two components training and testing
Step3:	Now apply CNN algorithm on the dataset. The dataset goes to three layers of CNN
Step4:	If we give (-3), the third layer will be extracted
Step5:	The third layers will do the prediction and the feature will return
Step6:	Using this feature, we will claim the random forest
Step 7:	While training the CNN, we will get the optimized feature then the random forest will train on that optimized feature
Step 8:	In random forest, we will use a scaling feature on the dataset
Step 9:	Here classifier.fit() function is used to train the model
Step 10:	Now apply on the dataset, Due to this predict on the random forest, the accuracy will be high
Step 11:	Based on evaluation metrics, the target type with mix value will be estimated output from CNN with Random Forest

Explanation

In the Proposed model, we are using a combination of two algorithms CNN with Random Forest to predict Ovarian Cancer in a patient. In step 1, we are taking details of a patient's medical reports with some features which are related to Ovarian Cancer. Step 2 Import libraries and pre-processing techniques in Python programming that splits the dataset into training and testing in step 3. We apply CNN algorithm to the dataset. The dataset goes to three layers of CNN. A convolution layer, a pooling layer, and a fully connected layer. If we give (-3), the third layer will be extracted from the CNN. When we extract the third layer then, the third layer will do the prediction. Once the prediction is done, the feature will return.

Using this feature, we will claim the random forest. While training the CNN, we will get the optimized feature then the random forest will train on that optimized feature. In random forest, we will use a scaling feature by using this feature the data divide into x trains and y trains to a small range. Now, we will use the classifier.fit() function to train the model. Now apply to the dataset, Due to this, the prediction on the random forest accuracy will be high. We are going to predict Ovarian Cancer in affected patients based on the evaluation metrics, the target type with mix value will be estimated output from CNN with Random Forest.

4. RESULTS

4.1 Performance Metrics

Table-2. Evaluation of various metrics.

S. No	Metrics	Equation/Formulae
1	Accuracy	$\text{Ay} = \frac{(\text{TP}+\text{TN})}{(\text{TP}+\text{TN}+\text{FP}+\text{FN})}$
2	Precision	$\text{Pn} = \frac{(\text{TP})}{(\text{TP}+\text{FP})}$
3	Recall	$\text{Rl} = \frac{(\text{TP})}{(\text{TP}+\text{FN})}$
4	F-Score	$\text{Fe} = \frac{(\text{TP})}{(\text{TP}+\frac{1}{2(\text{FP}+\text{FN})})}$
5	Jaccard Index	$\text{J}(\text{A},\text{B}) = \frac{\lvert A \cap B \rvert}{\lvert A \cup B \rvert}$
6	Error Rate	$\text{Ert} = \frac{(\text{FP}+\text{FN})}{(\text{P}+\text{N})}$
7	AUC-ROC Curve	$\text{AR crv} = \frac{(1+\text{TPR}-\text{FPR})}{2}$

From Table-2, various metrics were evaluated by using the confusion matrix of multiple models.

4.2 Results Obtained

The acquired results for all methods are listed in Table 3. Based on the five metrics examined, the proposed ensemble approach CNN with Random Forest has higher values. It was discovered that when MLP and ELM were compared to the CNN with Random Forest, an ensemble of two basic classifiers was shown to be effective. Performed better than expected. Three ensemble approaches were also used. Compared to CNN with Random Forest, as well as LSTM, even though the three groups

CNN with Random Forest approaches outperformed individual CNN. All of them were outperformed by CNN with Random Forest. For the CNN with Random Forest, the accuracy, precision, recall, F1-Score, Error Rate, Jaccard Index, and AUC curve values were calculated, and the values are 98.57, 98.75, 98.38, 98.54, 1.42, 97.17, 98.38.

Table-3. Results of all algorithms.

Algorithm	Accuracy (%)	Precision	Recall	Fscore	Error Rate	Jaccard Index	AUC Curve
MLP	84.28	84.04	84.24	84.12	15.71	72.90	84.24
ELM	88.57	89.02	89.41	88.56	11.42	79.53	89.41
LSTM	55.71	27.85	50.0	35.77	44.28	31.04	50.0
CNN with RF	98.57	98.75	98.38	98.54	1.42	97.17	98.38

From Table-3, the comparison of all the techniques based on accuracy is shown in Figure 9, and each metric is illustrated in Figure-10.

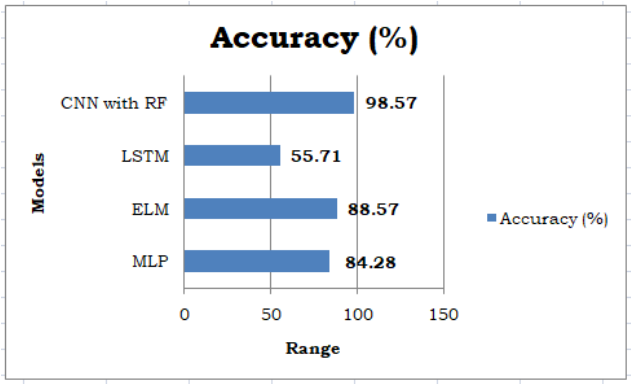

Figure-9. Comparison graph for accuracy with various models.

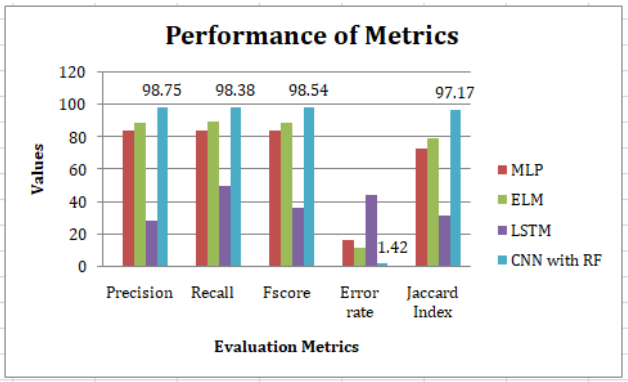

Figure-10. Comparison graph for various metrics with various models.

5. CONCLUSIONS

OC is becoming a more common disease in Women due to different lifestyles. Here we can predict Ovarian Cancer by using Run MLP, ELM, CNN, LSTM, and CNN with Random Forest and also we compare all five algorithms. In that, CNN with Random Forest performs better than other algorithms. The CNN with Random Forest gave 98% to 100% accuracy, which is more significant than different existing algorithms. CNN with Random Forest was identified as a better performing algorithm with accuracy, precision, recall, f1-score, error rate, Jaccard index, and AUC curve of 98.57, 98.75,98.38,98.54,1.42,97.17,98.38 respective. After a reasonable and fair examination of all outcomes, the CNN with Random Forest was determined to be superior to other ensembles in predicting Ovarian Cancer. In the future, it is expected to see a lot more work in the medical field using the best combination of ML algorithms.

REFERENCES

[1] Smith Mayoclinic. [Available online:] https://www.mayoclinic.org/diseases-conditions/ovarian-cancer/symptoms-causes/syc-20375941

[2] The Drugwatch. [Available online:] https://www.drugwatch.com/health/ cancer/ovari an-cancer/

[3] The Medlineplus. [Available online:] https://medlineplus.gov/ovariancl

[4] TheCancerorg.Availableonline:https://www.cancer.org/cancer/ovarian-cancer/detection-diagnosis-staging/staging.html

[5] The Cancer center. [Available online:] https://www.cancercenter.com/cancertypes/ ovarian-cancer/diagnosis-and-detection

[6] Cancer.org. [Available online:] https://www.cancer.org/cancer/ovarian-cancer/ about/what-is-ovarian-cancer.html

[7] Lu M., Fan Z., Xu B., Chen L., Zheng X., Li J., Znati T., Mi Q., Jiang J. 2020. Using machine learning to predict ovarian cancer. International Journal of Medical Informatics. 141: 104195.

[8] Paik E. S., Lee J. W., Park J. Y., Kim J. H., Kim M., Kim T. J., Choi C. H., Kim B. G., Bae D. S., Seo S. W. 2019. Prediction of survival outcomes in patients with epithelial ovarian cancer using machine learning methods. Journal of gynecologic oncology. 30(4).

[9] Akazawa M., Hashimoto K. 2020. Artificial intelligence in ovarian cancer diagnosis. Anticancer Research. 40(8): 4795-800.

[10] Kawakami E., Tabata J., Yanaihara N., Ishikawa T., Koseki, K., Iida, Y., Saito, M., Komazaki, H., Shapiro J.S., Goto C. and Akiyama Y. 2019. Application of artificial intelligence for preoperative diagnostic and prognostic prediction in epithelial ovarian cancer based on blood biomarkers. Clinical Cancer Research. 25(10): 3006-3015.

[11] Elhoseny M., Bian G. B., Lakshmanaprabu S. K., Shankar K., Singh A. K., Wu W. 2019. Effective features toclassify ovarian cancer data in internet of medical things. Computer Networks. 159: 147-56.

[12] Lee Z. J. 2008. An integrated algorithm for gene selection and classification applied to microarray data of ovarian cancer. Artificial Intelligence in Medicine. 42(1): 81-93.

[13] Negi S., Mittal P., Kumar B. 2020. In-Depth Analysis of Structures, Materials, Models, Parameters, and Applications of Organic Light-Emitting Diodes. Journal of Electronic Materials. 49: 4610-36.

[14] Mostavi M., Chiu Y. C., Huang Y., Chen Y. 2020. Convolutional neural network models for cancer type prediction based on gene expression. BMC medical genomics. 13(5): 1-3.

[15] Dong J., Li J., Fu A., Lv H. 2010. Automatic segmentation for ovarian cancer immuno histo chemical image based on yuv color space. In2010 International Conference on Biomedical Engineering and Computer Science, Apr 23 (pp. 1-4). IEEE.

[16] Assareh A., Moradi M. H. 2007. Extracting efficient fuzzy if-then rules from mass spectra of blood samples to early diagnosis of ovarian cancer. In2007 IEEE Symposium on Computational Intelligence and Bioinformatics and Computational Biology, Apr 1 (pp. 502-506). IEEE.

[17] Elhoseny M., Bian G. B., Lakshmanaprabu S. K., Shankar K., Singh A. K., Wu W. 2019. Effective features to classify ovarian cancer data in internet of medical things. Computer Networks. 159: 147-56.

[18] Rahman M. A., Muniyandi R. C., Islam K. T., Rahman MM. 2019. Ovarian Cancer Classification Accuracy Analysis Using 15-Neuron Artificial Neural Networks Model. In 2019 IEEE Student Conference on Research and Development (SCOReD), 15 (pp. 33-38). IEEE.

[19] Reddy S. S., Sethi N., Rajender R. 2020. Evaluation of deep belief network to predict hospital readmission of diabetic patients. In2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), (pp. 5-9). IEEE.

[20] Shankar R. S., Raju V. S., Murthy K. V., Ravibabu D. 2021. Optimized Model for Predicting Gestational Diabetes using ML Techniques. In2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA), (pp. 1623-1629). IEEE.

[21] Narasingarao M. R., Rajender R., Venkata Raju Kallipalli, V. Ramakrishna, Satya Aruna Varanasi, A. S. Lalitha & R. Shiva Shankar. 2022. Prediction of the breast cancer disease using machine learning techniques-a comparison. Harbin Gongye Daxue Xuebao/Journal of Harbin Institute of Technology. 54(1): 126-33.

[22] Shiva Shankar R., Murthy K. V., Someswara Rao C. 2020. Breast cancer disease prediction with recurrent neural networks (RNN). International Journal of Industrial Engineering & Production Research. 31(3): 379-86.

[23] Shankar R. S., Gupta V. M., Murthy K. V., Rao C. S. 2019. Breast Cancer Data Classification Using Machine Learning Mechanisms. Indian Journal of Public Health Research & Development. 10(5).

[24] Reddy S. S., Sethi N., Rajender R. 2020. A Comprehensive Analysis of Machine Learning Techniques for Incessant Prediction of Diabetes Mellitus. International Journal of Grid and Distributed Computing. 13(1): 1-22.

[25] Reddy S. S., Rajender R., Sethi N. 2019. A data mining scheme for detection and classification of diabetes mellitus using voting expert strategy. International Journal of Knowledge-Based and Intelligent Engineering Systems. 23(2): 103-8.

[26] Reddy S. S., Sethi N., Rajender R. 2019. A review of data mining schemes for prediction of diabetes mellitus and correlated ailments. In2019 5th International Conference On Computing, Communication, Control And Automation (ICCUBEA), (pp. 1-5). IEEE.

[27] Reddy S. S., Sethi N., Rajender R. 2021. Mining of multiple ailments correlated to diabetes mellitus. Evolutionary Intelligence. 14(2): 733-40.

[28] Lu Y. C., Lu C. J., Chang C. C. and Lin Y. W. 2017, November. A hybrid of data mining and ensemble learning forecasting for recurrent ovarian cancer. In 2017 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS) (pp. 216-216). IEEE.

[29] Mathur M., Jindal V. and Wadhwa G. 2020, November. Detecting Malignancy of Ovarian Tumour using Convolutional Neural Network: A Review. In 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC) (pp. 351-356). IEEE.

[30] Lotfi M., Misganaw B. and Vidyasagar M. 2017, May. Prediction of time to tumor recurrence in ovarian cancer: comparison of three sparse regression methods. In International Symposium on Bioinformatics Research and Applications (pp. 1-11). Springer, Cham.

[31] Park J. S., Choi S. B., Chung J. W., Kim S. W., Kim D. W. 2014. Classification of serous ovarian tumors based on microarray data using multicategory support vector machines. In 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, (pp. 3430-3433). IEEE.

[32] Yaar A., Asif A., Raza S. E., Rajpoot N., Minhas F. 2020. Cross-Domain Knowledge Transfer for Prediction of Chemo sensitivity in Ovarian Cancer Patients. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. pp. 928-929.

[33] Eganeh P. N., Mostafavi M. T. 2018. Use of machine learning for diagnosis of cancer in ovarian tissues with a selected mRNA panel. In 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), (pp. 2429-2434). IEEE.

[34] Yang X., Khushi M., Shaukat K. 2020. Biomarker CA125 Feature Engineering and Class Imbalance Learning Improves Ovarian Cancer Prediction. In2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), (pp. 1-6). IEEE.

[35] Zhou M., Lin F., Hu Q., Tang Z., Jin C. 2020. Ai-enabled diagnosis of spontaneous rupture of ovarian endometriomas: A PSO enhanced random forest approach. IEEE Access. 8: 132253-64.