Deep learning algorithm vs XGBoost using Wisconsin breast cancer diagnosis

Wan Aezwani Wan Abu Bakar*[a], Muhamad Amierusyahmi Zuhairi[a], Mustafa Man[b], Julaily Aida Jusoh[a], Nur Laila Najwa Josdi[a]
[a]Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Besut Campus, 22200, Besut, Terengganu, Malaysia;
[b]Faculty of Ocean Engineering Technology and Informatics, Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu, Malaysia

ABSTRACT

Breast cancer is causing a significant increase in the number of deaths every year. It is the most common kind of cancer and the leading cause of death in women all over the world. Any improvement in cancer illness prediction and detection is critical for a healthy life. As a result, high accuracy in cancer prediction is critical for keeping patients' treatment and survival standards up to date. Machine learning and deep learning approaches, which have been shown to have a significant impact on the process of breast cancer prediction and early diagnosis, have become a research hotspot and have been proven to be a powerful technique. We went through one machine learning method, XGBoost and deep neural network in this research on the Wisconsin Breast Cancer Diagnostic (WBCD) dataset. After getting the outcomes, these distinct classifiers' performance is evaluated and compared. The major goal of this research is to use machine-learning and deep learning algorithms to predict and diagnose breast cancer, and to determine which algorithms are the most efficient in terms of confusion matrix, precision, and accuracy.

Keywords: Breast cancer, prediction, diagnostic, machine learning, deep learning, xgboost, accuracy, precision, neural network

1. INTRODUCTION

A malignant growth in or around the breast tissue is known as breast cancer. A lump or calcium deposit generally forms as a result of aberrant cell development. The majority of breast lumps are harmless, however some are precancerous or cancerous. Breast cancer can be confined (occurring just in the breast) or metastatic (appearing elsewhere in the body) (spread to another part of the body). Doctor will likely perform a physical exam to evaluate a breast lump. To determine whether that lump is benign, your doctor may order mammography, breast ultrasound, breast MRI, PET/CT or Scintimammography[1]. If the lump is indeed benign, no further action may be needed. However, your doctor may want to monitor it to see if it changes, grows, or disappears over time. If the tests are inconclusive, your doctor may perform a biopsy using ultrasound, x-ray, or magnetic resonance imaging guidance. Breast cancer treatment depends on the tumor's size, extent of disease spread, type, receptor status, tumor growth rate and the patient's general health[1]. Treatments include surgery, radiation therapy, chemotherapy, hormone therapy or a combination thereof. The second largest cause of cancer mortality among women is breast cancer. Every year, only lung cancer claims the lives of more women. Breast cancer kills roughly 1 in every 39 women (about 2.6 percent). Breast cancer death rates in women under 50 have been stable since 2007, although they have continued to decline in older women. The mortality rate decreased by 1% year from 2013 to 2018[2]. These reductions are thought to be due to earlier detection of breast cancer through screening and more awareness, as well as improved therapies. This issue is presented as a two-class (B-benign, M-malignant) classification problem in this article, similar to Stahl[3] and Geekette[4].

Machine learning is crucial in the categorization of breast cancer. There are several diagnostic processes that have been explained above, each of which gives visuals. Machine learning is utilised to classify these sorts of diagnostic images. Machine learning is an AI subfield. Machine learning is used by many developers to retrain current models and improve their performance. For linear data, machine learning is employed. Machine learning produces superior outcomes when the data is tiny, but not when the data is vast. Machine learning is used to train the model in three different ways. With the

* wanaezwani@unisza.edu.my

Third International Conference on Computer Science and Communication Technology (ICCSCT 2022)
edited by Yingfa Lu, Changbo Cheng, Proc. of SPIE Vol. 12506, 125066Y
 · 0277-786X · doi: 10.1117/12.2663128

assistance of the supervisor, supervised machine learning works on known data. Machine learning that is unsupervised is performed without the presence of a supervisor. Machine learning based on reinforcement is becoming less popular. These algorithms use the best data from previous knowledge to make precise decisions.

Machine learning has a sub-field called deep learning. There are decision trees and decision tree-based (traditional machine learning methods) mostly used before in publications such as methods[5]. Deep learning is an unsupervised learning technique that learns from data. Unstructured or unlabelled data is possible. When a deep neural network has more than two hidden layers, it is referred to be a deep network. The input layer is the first one, while the output layer is the second. In comparison to a neural network, the intermediate layer is known as the hidden layer, and it contains more layers. Neurons are the nodes that contain the layers. Machine learning and deep learning vary in that deep learning is closer to its purpose than machine learning. For instance, the supervised deep learning approach has recently gained popularity. Stahl[3] and Geekette[4] used this approach to diagnose breast cancer using feature values obtained from a digitised picture of a Fine Needle Aspirate (FNA) of a breast lump in the WBCD dataset[6]. Refer to[3], He tested with three types of deep neuron networks as 1, 2, and 3 hidden layers of 30 neurons without any data pre-processing using the WBCD dataset with derived characteristics such as mean, standard error, etc. As for Geekette[4] did not give specific details on the data pre-processing but using the original identified features like centre and scale. Furthermore, this article will provide the comparison result of XGBoost[7-8] and deep neural network.

2. RELATED WORK

Machine learning and deep learning are two artificial intelligence subcategories that have gotten a lot of press in the past years. Machine learning refers to when computers learn from data. It relates to the intersection of computer science and statistics, in which algorithms are applied to do a given task without being explicitly trained; instead, they identify patterns in data and make predictions as new data enters. A basic linear regression method is an example of a standard machine learning algorithm. Consider the case where you wish to forecast your earnings based on your years of higher education. You must first define a function, such as income = y + x * years of study. Then provide a set of training data to your algorithm. This might be a basic table providing information on people's years of higher education and earnings. Then, using an ordinary least squares (OLS) regression, let your algorithms design the line for you. You may now provide the computer some test data, such as your own years of higher education, and it will forecast your earnings. Now, test the data using the algorithm, e.g. allow it to predict your earnings based on your personal years of higher education. Deep learning algorithms may be thought of as an advanced and mathematically complicated development of machine learning algorithms. Recently, the field has gotten a lot of attention and with good reason: recent breakthroughs have led to results that were previously impossible. As Sarker and I.H (2021) stated that deep learning has shown to be a great response to a wide range of real-world issues[9]. Deep learning refers to algorithms that evaluate data using a logic structure comparable to that of a person. This can occur through both supervised and unsupervised learning. Deep learning applications do this by employing a layered structure of algorithms known as an artificial neural network (ANN). The architecture of ANN is based on the biological neural network of the human brain, resulting in a significantly more competent learning process than typical machine learning models. Deep learning is applied in a variety of areas. Deep learning is used in autonomous driving to identify things like STOP signs and pedestrians[10-11]. Then, military also use deep learning to identify objects detect from satellites[12]. The effort in[13] discusses the statistical formula used in convolutional neural network (CNN) and the how the issues of execution time and memory usage is solved in[14-15] by having Eclat algorithm as the based model.

A study[16] proposed a convolutional neural network (CNN) method to analyze hostile ductal carcinoma tissue zones in whole slide images (WSI) for automatic detection of breast cancer. Researchers have compared three different CNN architectures and the CNN 3 model proposed by the system achieves 87% accuracy. The five CNN layers in Model 3 are best suited for this task even though they are "deeper" than Models 1 and 2. All three architectures were built based on a large dataset of approximately 275,000, 50 × 50 pixels of RGB image tamping. The results of the comparison between the proposed model and the machine learning (ML) algorithm show that the ability of the proposed model to improve the accuracy by 8% compared to the results of the ML algorithm. Therefore, the proposed model has succeeded in obtaining better accuracy, which allows the reduction of human error in the diagnosis process as well as reducing the cost of cancer diagnosis.

This paper[17] emphasizes various machine learning such as K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Decision tree, Naïve Bayes Logistic Regression, Random Forest and many more are used to predict breast cancer on datasets that have been taken from the Kaggle repository. ie Wisconsin. Each accuracy achieved by all these algorithms was measured and compared. All techniques are coded in python and implemented in Google Colab, which is a Scientific

Python Development Environment. Based on the experiments that have been done, SVM and Random Forest Classifier show the best accuracy of 96.5%. To implement deep learning in an effort to maximize prediction accuracy, algorithms such as ANN and CNN have been used and the results received are very satisfactory with 99.3% and 97.3% accuracy respectively. At the end of the study, the researcher concluded that deep learning models are able to produce better accuracy compared to machine learning algorithms. In addition, the output from deep learning can be probabilistically predicted using the ReLu Activation function, which is not possible with machine learning algorithms.

This study[18] highlights a model that uses deep neural networks based on multi-omic data, DeepMO such as mRNAdata, DNA methylation data and copy number variation (CNV) data collected from The Cancer Genome Atlas (TCGA). This model has been used to classify breast cancer subtypes. Each omics data is fed into a deep neural network consisting of a coding subnet and a classification subnet after data pre-processing and feature selection. DeepMO results based on multi-omics on binary classification are better than other methods in terms of accuracy and area under the curve (AUC). The classification accuracy results show that the proposed model, the integration of multi-omic data sets is able to improve the performance of the model compared to using single omics data in classifying breast cancer subtypes. Furthermore, the proposed model also shows better performance compared to latest methods on binary classification and multi-binary classification. In addition, several biological explanations for distinguishing breast cancer subtypes have also been analyzed, followed by the provision of guidance for exploring biological models.

3. ANALYSIS

3.1 Exploring data

The public WBCD dataset[6] was used in this research. The data collection includes the results of a standard breast cancer screening, which allows the illness to be identified and treated before symptoms appear. The purpose of the dataset is to conduct classification analysis so that you can anticipate which of the sub-populations a new observation belongs to base on metrics that you choose. To put it another way, we will be able to forecast if a patient has a benign or malignant cancer after analysing the cancer diagnostic information. The following characteristics for each cell nucleus were generated and will be utilised as inputs to the new deep learning model, as defined in the WBCD dataset: Marginal Adhesion (MA), Single Epithelial Cell Size (SECSize), Bare Nuclei (BN), Bland Chromatin (BC), Normal Nucleoli (NN), Mitoses, Clump Thickness (CT), Uniformity of Cell Size (UCSize), Uniformity of Cell Shape (UCShape). To acquire a statistical summary of the dataset, use the following code:

```
headers = ["ID","CT","UCSize","UCShape","MA","SECSize","BN","BC","NN","Mitoses","Diagnosis"]

data = pd.read_csv('breast-cancer-wisconsin.csv', na_values='?', header=None, index_col=['ID'], names = headers)

data = data.reset_index(drop=True)
```

First run, we encounter some problem that got 16 entries missing for Bare Neclui (indicated as? in the original WBCD dataset). After checking back the data and rerun, there is no missing data as shown in Figure 1.

	CT	UCSize	UCShape	MA	SECSize	BN	BC	NN	Mitoses	Diagnosis
count	699.000000	699.000000	699.000000	699.000000	699.000000	699.000000	699.000000	699.000000	699.000000	699.000000
mean	4.417740	3.134478	3.207439	2.806867	3.216023	3.463519	3.437768	2.866953	1.589413	2.689557
std	2.815741	3.051459	2.971913	2.855379	2.214300	3.640708	2.438364	3.053634	1.715078	0.951273
min	1.000000	1.000000	1.000000	1.000000	1.000000	0.000000	1.000000	1.000000	1.000000	2.000000
25%	2.000000	1.000000	1.000000	1.000000	2.000000	1.000000	2.000000	1.000000	1.000000	2.000000
50%	4.000000	1.000000	1.000000	1.000000	2.000000	1.000000	3.000000	1.000000	1.000000	2.000000
75%	6.000000	5.000000	5.000000	4.000000	4.000000	5.000000	5.000000	4.000000	1.000000	4.000000
max	10.000000	10.000000	10.000000	10.000000	10.000000	10.000000	10.000000	10.000000	10.000000	4.000000

Figure 1. Statistical summary of the original WBCD database.

Missing data is one of the most typical problems with datasets, and the WBCD dataset is no exception. Then, after partitioning the dataset into training and testing subsets, the WBCD dataset has only 699 samples, which is too small for deep learning. Distinct features have different ranges of feature values. The most typical function values range from 1.5 to 4.5. The number of fact samples at various labels is not evenly distributed. Indeed, the total number of facts samples

classified as B (458) substantially doubles the overall number of data samples classified as M (241). (Section 3.2 for the visualization of these numbers). Skewness is another common issue with datasets.

3.2 Visualizing data

Using the code below, Figure 2 presents the various mean value ranges of different features:

```
import seaborn as sns

data_mean = data.describe().loc[‘mean’]

data_mean.plot(kind=‘bar’, figsize=(14,6))
```

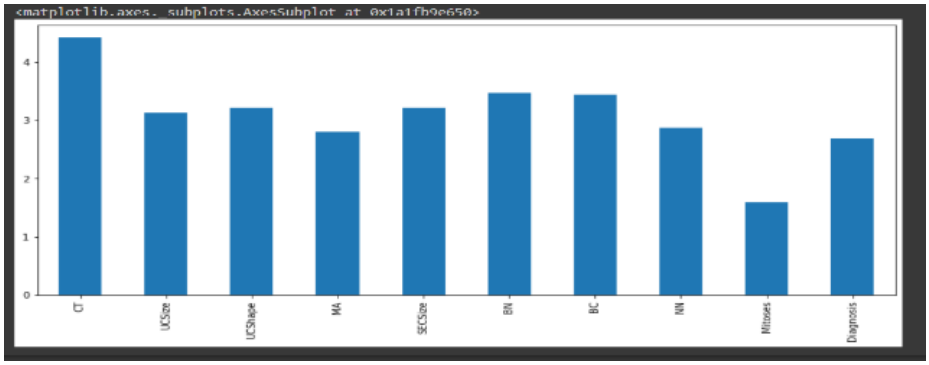

Figure 2. Various ranges of features value.

Figure 2 shows the overall number of samples labelled as B and M using the following code to demonstrate the imbalanced dataset issue:

```
data_B = data[data[‘Diagnosis’] == 2]
```

	CT	UCSize	UCShape	MA	SECSize	BN	BC	NN	Mitoses	Diagnosis
count	458.000000	458.000000	458.000000	458.000000	458.000000	458.000000	458.000000	458.000000	458.000000	458.0
mean	2.956332	1.325328	1.443231	1.364629	2.120087	1.305677	2.100437	1.290393	1.063319	2.0
std	1.674318	0.907694	0.997836	0.996830	0.917130	1.182666	1.080339	1.058856	0.501995	0.0
min	1.000000	1.000000	1.000000	1.000000	1.000000	0.000000	1.000000	1.000000	1.000000	2.0
25%	1.000000	1.000000	1.000000	1.000000	2.000000	1.000000	1.000000	1.000000	1.000000	2.0
50%	3.000000	1.000000	1.000000	1.000000	2.000000	1.000000	2.000000	1.000000	1.000000	2.0
75%	4.000000	1.000000	1.000000	1.000000	2.000000	1.000000	3.000000	1.000000	1.000000	2.0
max	8.000000	9.000000	8.000000	10.000000	10.000000	10.000000	7.000000	9.000000	8.000000	2.0

Figure 3. Samples from dataset labelled as B.

```
data_M = data[data[‘Diagnosis’] == 4]
```

	CT	UCSize	UCShape	MA	SECSize	BN	BC	NN	Mitoses	Diagnosis
count	241.000000	241.000000	241.000000	241.000000	241.000000	241.000000	241.000000	241.000000	241.000000	241.0
mean	7.195021	6.572614	6.560166	5.547718	5.298755	7.564315	5.979253	5.863071	2.589212	4.0
std	2.428849	2.719512	2.562045	3.210465	2.451606	3.180182	2.273852	3.350672	2.557939	0.0
min	1.000000	1.000000	1.000000	1.000000	1.000000	0.000000	1.000000	1.000000	1.000000	4.0
25%	5.000000	4.000000	4.000000	3.000000	3.000000	5.000000	4.000000	3.000000	1.000000	4.0
50%	8.000000	6.000000	6.000000	5.000000	5.000000	10.000000	7.000000	6.000000	1.000000	4.0
75%	10.000000	10.000000	9.000000	8.000000	6.000000	10.000000	7.000000	10.000000	3.000000	4.0
max	10.000000	10.000000	10.000000	10.000000	10.000000	10.000000	10.000000	10.000000	10.000000	4.0

Figure 4. Samples from dataset labelled as M.

```
B_M_data = {'B': [data_B.shape[0]], 'M': [data_M.shape[0]]}
B_M_df = pd.DataFrame(data=B_M_data)
B_M_df.plot(kind='bar', figsize=(10,4))
```

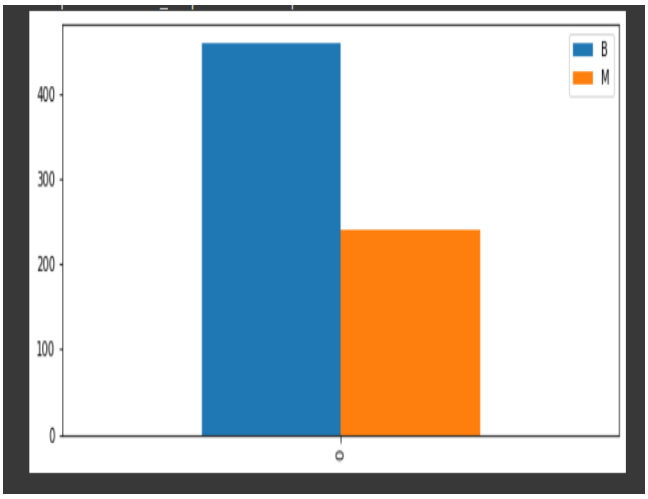

Figure 5. Total numbers of data samples labelled as B (blue) and M (Orange) in the original dataset.

4. DEEP LEARNING MODEL

As previously stated, this article treats the breast cancer diagnostic problem as a two-class (benign or malignant) classification problem. The classification is done with a new supervised deep learning model. Figure 6 below shown the architecture of new deep learning model. This model inherits the approaches tested by Stahl (2017) and Geekette (2016). Actually, very much like Stahl (2017), new technique, in particular, leverages the previously discovered characteristics, centres and scales used for pre-processes data and diagnosis problem of breast cancer treats as two class (benign or malignant) classification problem. Then, similarly to Geekette (2016), the new model technique inherits multiple hidden layer of DNN.

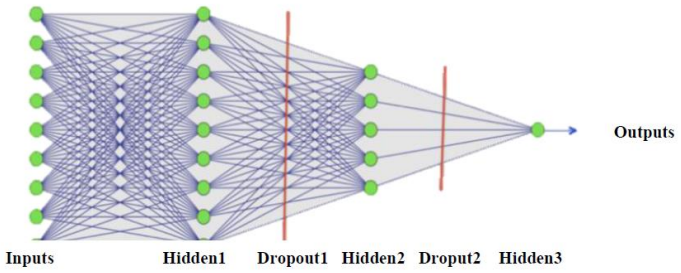

Figure 6. The architecture of the new deep learning model.

To correspond to the number of input features and output classes, the first hidden layer has 9 neurons and the last hidden layer uses 1 neuron. In addition, between hidden layers, dropout layers (Srivastava et al., 2014) with a 50% dropout rate are included to reduce overfitting and perhaps increase the accuracy of the new deep learning model. In hidden layers 1 and 2, the common rectified linear unit (ReLU) activation function relu is utilised. In the final hidden layer 3, the widely used sigmoid activation function is utilised to provide continuous output in the range of (0, 1). 0.5 of threshold is applied to produce binary diagnosis outcome from continuous output. Then, for the loss function, we used binary cross entropy loss because the WBCD classification is binary classification problem. Hence, the Adam (Adaptive Moment Estimation) (Kingma et al., 2016) is used. As for metric function, accuracy is chosen rather than binary accuracy because in model training, its support accuracy history, while binary accuracy does not. One of the popular open source Keras deep learning (Lee, H., & Song, J., 2019) is used for the execution of the new deep learning model. Furthermore, we need to verify the robustness of the new deep learning model performing a 10-fold validation, result in Section 6.2.

5. XGBOOST

Non-deep learning approaches, XGBoost (Abdulkareem et al., 2021; Chen et al., 2016) is a popular and efficient open-source version of the gradient boosted trees technique. Gradient boosting is a supervised learning technique that combines the estimates of a collection of smaller, weaker models to accurately predict a target variable. For building classification model, the XGBClassifier is needed and loaded from XGBoost. The XGBClassifier (Chen et al., 2016) class implements the Scikit-Learn interface for using XGBoost for classification. That means that it has the familiar fit method as well as predict and score. The data that have been prepared uses to define the configuration of XGBClassifier model. Next, using fit () function to train the model on training data (Section 6), and the data will be examined by the XGBoost algorithm, which will seek for connections between the characteristics and the target variable. Then, the trained data model can make a prediction, predict (), and evaluate using accuracy_score () function. Two values needed, original data that have actual result and prediction data contain predicted result. The model's accuracy score is computed by dividing the number of correct predictions by the total number of predictions (Section 6.3). XGBoost being a competitive machine learning algorithms because it can produce comparable results. As Abdulkareem et al. (2021) stated that XGBoost outperformed all other classifiers.

6. RESULT & DISCUSSION

To implement data processing and the new deep learning model, the following standard open source libraries are required, numpy, pandas, scikit-lean, Keras. As previously stated, the values of various attributes have varying ranges, with the most of those values falling outside of the [0, 1] range. However, for improved performance, the new deep learning model requires input values in the [0, 1] range. The dataset will be split into 2 parts, 75% for model training and 25% for model testing. To avoid dataset key mismatch problem, the data subsets result from splitting dataset must be re-indexed. After that, the result is Pandas DataFrames structure need to convert into Numpy arrays because Keras API does not support Pandas DataFrame. Then, the new variables are vital for XGBoost and the new deep learning model prediction.

6.1 Refinement

6.1.1 Dropout. Dropout layers (Srivastava, 2014) in principle, should assist to prevent overfitting and hence enhance deep learning model accuracy. However, as the number of epoch reaches 500 or more, utilising Dropouts at a rate of 50% yields no significant performance boost, according to the testing results. When Dropouts are implemented, a dropout rate of 50%

appears to be a sweet spot, as indicated in the Table 1 below. Changing dropout to 30% produced a prediction accuracy of about 97%, while 50% dropout got 98.09% prediction accuracy.

Table 1. Dropout.

Hidden Layers	Dropouts	Epochs	Batch Size	Accuracy
9,5,1	0.5	500	16	0.9809
9,5,1	0.3	500	16	0.9766

6.1.2 Batch Size. The batch size has an impact on the accuracy of the new deep learning model. A batch size of 16 appears to work best, as seen in the table below.

Table 2. Batch size.

Hidden Layers	Dropouts	Epochs	Batch Size	Accuracy
9,5,1	0.5	500	10	0.9723
9,5,1	0.5	500	16	0.9809
9,5,1	0.5	500	32	0.9766
9,5,1	0.5	500	64	0.9745

6.2 Model evaluation and validation result

The dataset is divided into two portions, 75 percent for training and the remaining 25% for testing, as in Geekette (2016). As previously stated, the prediction accuracy is employed as the primary assessment statistic, comparable to the approach used by Geekette (2016). A 10-fold cross validation with 500 epochs and batch size 16 was done to evaluate the resilience of the new deep learning model in the event of slight dataset changes, and the results are tabulated in Table 3.

Table 3. Model evaluation and validation.

Cross Validation No.	Prediction Accuracy
1	99.29%
2	97.87%
3	97.87%
4	97.87%
5	95.74%
6	97.16%
7	98.58%
8	97.87%
9	98.58%
10	97.16%
Average Score	97.80% (+/- 0.92%)

As seen in the table above, the accuracy performance of the new deep learning model remains rather stable in the face of dataset changes.

6.3 Benchmark result

To quantity comparative performance, the prediction accuracy of the new supervised deep learning model is compared to the outcome of implementing machine learning, XGBoost. The outcomes of implementing the 25% testing dataset to the trained XGBoost and new deep learning model are summarized in Table 4.

Table 4. Result.

Machine Learning Model	Accuracy
Deep Neural Network	0.9809
XGBoost	0.9809

The accuracy of the new deep learning model is comparable to XGBoost, as seen in the table 3 above. Both model attained same accuracy prediction rate of 98.09%. For new deep leaning model, this result outperformed the result testified by Stahl (2017) and Geekette (2016).

7. CONCLUSION

This experiment can conclude that XGBoost is competitive with the new deep learning model that using 3 hidden layer thus include dropout layer (Srivastava, 2014) and sigmoid activation function in the final hidden layer in the classification of the WBCD dataset. Numpy, Pandas, scikit-learn, Keras, Matplotlib, and other well-known open source libraries were used to create the new deep learning model in Python. From this result also we can conclude that using XGBoost as machine learning giving same better result than an advance deep learning model in case of WBCD dataset. Furthermore, deep learning has the capabilities to solve new challenge in the future because of its flexibility that many different application and data kinds can benefit from the same deep learning technique. This research will be integrated with our research in the combination of Eclat (Equivalence Class Transformation) from WA. Bakar et al. (2018) and Man et al. (2018).

ACKNOWLEDGEMENT

Authors wish to thank FRGS of Kementerian Pengajian Tinggi Malaysia (KPT) for providing financial support for this study with grant code (FRGS/1/2020/ICT06/UNISZA/03/1) and all team members especially collaborators from UMT for morale and technical support. Authors also appreciate all faculty members' help in checking for spelling errors and synchronisation inconsistencies, as well as their thoughtful comments and suggestions.

REFERENCES

[1] ACR, Breast Cancer. [online] Radiologyinfo.org. Available at: <https://www.radiologyinfo.org/en/info/breast-cancer> Accessed 9 January 2022, (2020).
[2] Cancer.org, Breast Cancer Statistics | How Common Is Breast Cancer? [online] Available at: <https://www.cancer.org/cancer/breast-cancer/about/how-common-is-breast-cancer.html> [Accessed 9 January 2022], (2022).
[3] Stahl, K., Wisconsin Breast Cancer Diagnosis Deep Learning. RPubs. Retrieved January 2, 2022, from https://rpubs.com/kstahl/wdbc_ann, (2017).
[4] Geekette, D., "Breast Cancer data: Machine Learning & amp; Analysis. RPubs." Retrieved January 2, 2022, from https://rpubs.com/elena_petrova/breastcancer, (2016).
[5] Nguyen, C., Wang, Y. and Nguyen, H. N., "Random forest classifier combined with feature selection for breast cancer diagnosis and prognostic," Journal of Biomedical Science and Engineering, 06, 551-560(2013).
[6] William, H., Wolberg, W., Nick, S. and Olvi, L. M., "Breast cancer Wisconsin (diagnostic) data set. UCI Machine Learning Repository," (1992).

[7] Abdulkareem, S. A. and Abdulkareem, Z. O., "An Evaluation of the Wisconsin Breast Cancer Dataset using Ensemble Classifiers and RFE Feature Selection Technique. International Journal of Sciences: Basic and Applied Research (IJSBAR)," 55 (2), 67-80(2021).
[8] Chen, T. and Guestrin, C., "XGBoost: A Scalable Tree Boosting System," Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, (2016).
[9] Sarker, I. H., "Deep Learning: A Comprehensive Overview on Techniques," Taxonomy, Applications and Research Directions. SN COMPUT, SCI, 2, 420(2021).
[10] Muhammad, K., Ullah, A., Lloret, J., Ser, J.D. and Albuquerque, V.H., "Deep Learning for Safe Autonomous Driving: Current Challenges and Future Directions," IEEE Transactions on Intelligent Transportation Systems, 22, 4316-4336(2021).
[11] Srivastava, N., Hinton, G.E., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R., "Dropout: a simple way to prevent neural networks from overfitting". J. Mach. Learn. Res., 15, 1929-1958(2014).
[12] Kingma, D. P. and Ba, J., "Adam: A Method for Stochastic Optimization," CoRR, abs/1412.6980(2015).
[13] Lee, H. and Song, J. "Introduction to convolutional neural network using Keras; an understanding from a statistician," Communications for Statistical Applications and Methods, 26, 591-610(2019).
[14] Bakar, W. A., Jalil, M.M., Man, M., Abdullah, Z. and Mohd, F., "Postdiffset: an Eclat-like algorithm for frequent itemset mining," International journal of engineering and technology, 7, 197(2018).
[15] Man, M., Bakar, W.A., Jalil, M.M. and Jusoh, J.A., "Postdiffset Algorithm in Rare Pattern: An Implementation via Benchmark Case Study," International Journal of Electrical and Computer Engineering, 8, 4477-4485(2018).