https://doi.org/10.31272/jae.i149.1449
https://admics.uomustansiriyah.edu.iq
P-ISSN: 1813–6729 E-ISSN: 2707-1359

Applying Machine Learning Algorithms to Determine Therapy Types for the Brain

Nigar Abdulla Ghafur
Dept. of Statistics and Informatics, College of Administration and Economics, University of Sulaimani, Sulaimaniyah, Iraq.
Email: nigarghafur@yahoo.com , ORCID: https://orcid.org/0009-0009-5417-0033
Abbas Gulmurad Beg Murad
Dept. of Statistics and Informatics, College of Administration and Economics, University of Sulaimani, Sulaimaniyah, Iraq.
Email: abbas.beg@univsul.edu.iq, ORCID: https://orcid.org/0009-0003-4785-699X

Article Information

Article History:
Received: 26 / 06 / 2025
Revised: 02 / 08 / 2025
Accepted: 10 / 08 / 2025
Available Online: 01 / 09 / 2025
Pages no: 121 – 132

Keywords:
Machine learning, K-nearest neighbour, Accuracy, Diagnosis Subtype, Brain tumour therapy.

Correspondence:
Researcher name:
Nigar Abdulla Ghafur

Email: nigarghafur@yahoo.com

Abstract

Artificial intelligence has already been utilised in medical treatments for a while, but its progress over the past ten years is impressive. In medicine, both machine learning and deep learning approaches are becoming more popular since there is more health-related data available, these techniques are better at handling massive datasets, and computers are getting more powerful.

The purpose of this paper is to investigate how to use supervised learning, more specifically, classification machine learning and deep learning algorithms, to discover the best way to classify brain tumour therapy using K-Nearest Neighbour and recurrent neural network algorithms. Employed algorithms for machine learning, K-nearest neighbour, and the recurrent neural network to see how risk factors changed the way drugs were classified for patients. Both algorithms worked well. The results suggest that forecasting methods can group drugs and are useful for making decisions. This development is a step toward using machines and deep learning in medical treatment.

1. Introduction

The development and usage of Machine learning (ML) techniques in medicine have been limited over the past century. Reasons for this limitation include the initial focus on non-ML, AI approaches in medicine, and the need for large amounts of data to discover hidden patterns [3]. ML automatically identifies significant patterns in data. Making programs learn and adapt is the goal of machine learning [17]. The K-nearest neighbour (K-NN) algorithm is based on supervised learning. Supervised learning is a type of ML. Supervised learning can be understood as allowing a supervisor to act as a teacher. Supervised learning involves training the machine using the supervisor's data or labelled data [23].

The K-NN algorithm is a widely used, simple, and straightforward non-parametric method used for classification and regression problems. It is a lazy learning algorithm that calculates distances between testing examples and training data to identify nearest neighbours and produce classification output [5]. Regulatory agencies have approved clinical tools that use algorithms based on machine learning to help physicians make treatment decisions, especially in imaging, radiology, and pathology. Because more experienced clinicians can use these tools successfully, their integration in healthcare depends on user endorsement and trust [27].

Iqbal et al explained that artificial intelligence (AI) and machine learning can improve cancer detection and treatment, tackling challenges that traditional methods cannot. AI can analyse large

amounts of information and integrate biological data to find early-stage genetic alterations and abnormal protein interactions. It helps physicians interpret complex medical data and make real-time decisions, diagnosing cancer, forecasting progression, and customising therapies despite challenges like data confidentiality and the categorisation of over 100 cancer types [6]. To address growing expenses and diminishing research rates within the healthcare industry, in addition to speeding up production and cutting expenses, machine learning algorithms streamline protocol expansion, patient management, data analysis, and regulatory compliance [7].

Alnuaimi and Albaldawi discussed ML, a crucial part of AI, including reinforcement learning, semi-supervised learning, supervised learning, and unsupervised learning. Understanding these techniques can help develop machine learning and its applications [1]. Maleki Varnosfaderani and Forouzanfar's study investigates the potential of AI in healthcare. To create solutions powered by AI that put human needs first, customised medicine and pharmaceutical development, however, require substantial stakeholders' involvement [12].

The objective of this study was to compare and evaluate the predictive models of the machine learning algorithm for classifying four types of drugs. This paper is mainly about the main issues that need to be solved to compare and analyse how well the k-nearest neighbour and recurrent neural network for therapy classification work for brain cancer disease.

2. Machine Learning Algorithm

Machine learning (ML) is a branch of artificial intelligence that uses computer systems to analyse large datasets, identify patterns, and solve problems. It includes reinforcement learning, supervised learning, and unsupervised learning. Supervised learning uses input and outcome data to predict disease outcomes, while unsupervised learning uses input data to discover targets and illnesses. Python is the preferred program for statistical analysis in the medical field [4].

2.1. Supervised Machine Learning

A well-liked machine learning technique called supervised learning (SL) uses labelled data to train algorithms to forecast outcomes that experts determine. Algorithms, mathematics and employ configured functions, classifiers, and regressors. By combining model results, ensemble approaches improve accuracy [5], for clarity of the processes of supervised machine learning, as shown in Figure (1).

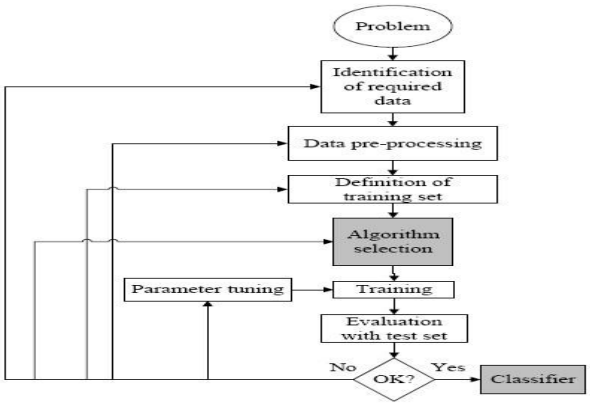

Figure (1): How supervised learning via machines works [17].

The preceding flowchart emphasises the categorisation of machine learning techniques as well as the identification of a highly effective method exhibiting optimal precision and accuracy. Additionally, this involves evaluating the efficacy of various approaches on large and small datasets to ensure accurate classification and provide knowledge about the construction of supervised machine learning algorithms [17]. By analysing the relationship between the variable in question and the available data, the statistical model known as logistic regression produces a discrete range with consistent probabilistic values, which is used in machine learning with supervision (SML) to predict

problems with binary classification [8]. It is useful for classification problems and decision-making in various situations, as illustrated in Figure (2).

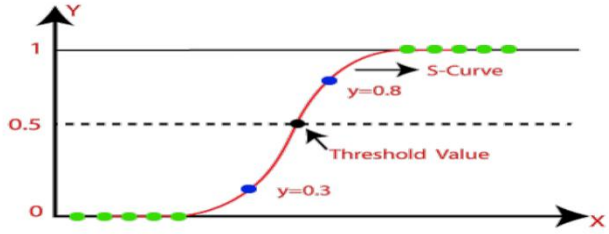

Figure (2): Image showing the output of the Logistic Regression algorithm [22].

Figure 2 classifies the investigation's results as geometry-based issues, indicating a 70% probability of failure. The next step is the history quiz, with a 30% probability of finding a solution. In detecting spam emails, logistic regression is not optimal due to constraints on classification. It can be divided into binary, multinomial, and ordinal types [22].

2.2. K-Nearest Neighbour Classifier in Supervised Machine Learning

The primary concept of the suggested approach involves determining the category labelling of the data based on K valid data points from the training set [18]. the KNN method is a straightforward, non-parametric classification technique. From the traditional length, the most widely used distance evaluation method, it forecasts choices regarding class based on distance measurements and nearby rewards [15]. There are different ways to measure how well an algorithm works, and hidden examples are often used during the training process, including "Precision, Specificity, Accuracy, Recall/Sensitivity, and F1-Score" to measure the performance of the algorithms. Thus, the higher the accuracy of an algorithm, the greater the chance of creating perfect predictions [23]. The steps in the classification of the KNN are:

A. Enter the computed combined imaging quantity.
B. Sets k (nearer surrounding areas).
C. Determine how closely the Euclidean distance (E.D.) model fits, for example, between two points of the data using the following formula:

$$\text{E.D. (A and B)} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (1)$$

D. Sort ascending space outcomes from highest to lowest.
E. Counting every class using the k-closest neighbours.
F. The testing information uses almost every class. This study measured performance metrics using the confusion matrix to estimate the accuracy of both correct and incorrect classifications. Below is the confusion matrix procedure:

A. Precision measures the accuracy of the system's equation in responding to the user's request.

$$\text{Precision} = TP \div (TP + FP) \quad (2)$$

B. Recall measures the system's knowledge retrieval success in the equation as follows:

$$\text{Recall} = TP \div (TP + FN) \quad (3)$$

C. Accuracy helps measure technique performance with this equation [14].

$$\text{Accuracy} = (TP + TN) \div (TP + TN + FP + FN) \quad (4)$$

D. The harmonic Mean of the (2) and (3) expressions defines the F1 score.

$$\text{F-1score} = 2 * \frac{Precsion.Recall}{Precsion+Recall} \quad (5)$$

E. Some applications emphasize precision or recall. To achieve this during training, adjust Fβ as follows:

$$F_\beta = \frac{(1+\beta^2)\,TP}{(1+\beta^2)\,TP+\beta FN+FP} \quad (6)$$

Whereby, true positive, or TP for short, is the proportion of attacks that are correctly detected. The proportion of valid traffic that is categorised as such is known as the true negative. The proportion of valid traffic that is classified as an attack is known as a false positive (FP). The percentage of legitimate traffic that was categorised as intrusions—FN [21].

3. An Introduction to Cancer Disease

Different organs and tissues harbour more than 100 types of cancer, rendering it a global disease. Precision and personalised healthcare (PPM) aims to address this heterogeneity by identifying genetic variations in tumours and correlating them with effective treatments [10].

3.1. Brain Tumour

Malignant brain tumours represent a principal cause of mortality worldwide. Glioma is the primary variant, distinguished by restricted treatment options. Advancements in molecular biology have not significantly improved outcomes for cancer patients [11]. Different symptoms, location of the malignancy, and tumour size make it difficult to detect brain tumours in their early stages. Headaches, vomiting, and anorexia are common symptoms that should be looked into right away for possible brain tumours [24]. The blood–brain barrier (BBB) is made up of the walls of tiny blood vessels, along with their supporting membranes and nearby cells called glial cells and pericytes, as shown in Figure 3.

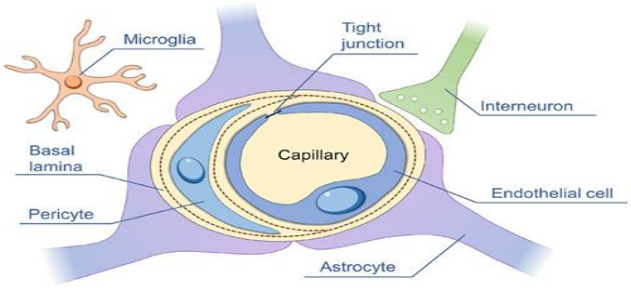

Figure (3): A diagram of a BBB fill structure, showing the capillaries, blood vessels, tight junctions, normal lamination, pericytes, and also neurons, which are microglia, as well as interneurons [13].

The endothelium, which serves as a selective wall that keeps blood substances out of the brain, maintains the capillaries in the brain, also referred to as hem capillaries. While astrocytes help chemicals move between capillaries and neurons, microglia defend the immune system [13].

3.2. Risk factors of a Brain Tumor

This paper discusses four hypotheses regarding the pathogenesis of primary brain tumours. The first hypothesis is that the grade of cancer significantly affects patient outcomes. The next factor is the diagnosis subtype, followed by the patient's age, which influences the development of brain tumours; finally, understanding the size of the malignancy is crucial for determining the best treatment.

3.3. Symptoms of a Brain Tumour

Brain cancer impacts motor skills, leading to convulsions, loss of feeling, and problems with mobility. Additionally, it impairs cognitive and emotional abilities, resulting in difficulties with speech, mood swings, focus, and hearing. Different symptoms make early detection difficult [20].

3.4. Treatments of Brain Tumour

There are types of treatment, such as surgery, chemotherapy, radiation therapy, targeted therapy, and immunotherapy [22].

4. The Development of Drugs

The creation of drugs is exhausting and uncertain, but big data has changed strategies in response to ongoing pharmaceutical difficulties. The collection of multiple omics data and the use of high-throughput devices have enhanced the clinical translation of basic research [19]. Based on the study, creating a new medication costs more than $2.5 billion [21]. In this paper, drugs will be classified into five categories, which are: Temozolomide, Bevacizumab, Vincristine, Etoposide, and Temozolomide& Bevacizumab, as illustrated in Figure 4.

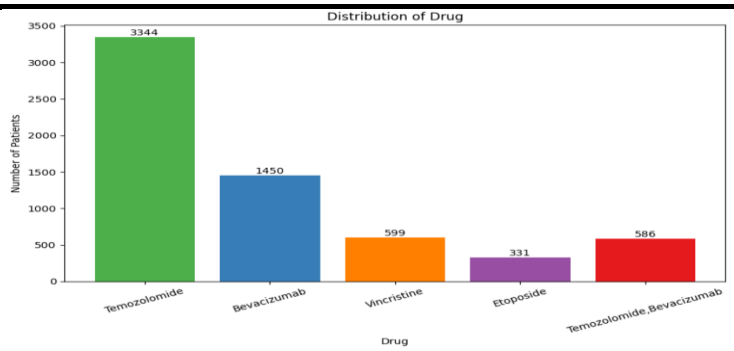

Figure (4): Bar chart for drugs.

As shown in Figure 4, temozolomide is the most frequently used drug among patients compared to others.

5. Descriptive Data and Data Collection

This paper includes a sample of 536 patients diagnosed with brain tumours, gathered from 2018 to 2024 over a three-month duration. The statistics and IT unit at Hiwa Hospital in Sulaymaniyah, which assists both pharmacy and oncology physicians, collected this data. Each patient is required to possess five specific pieces of information as outlined in Table 1.

Table 1 represents descriptive features of brain tumour patients.

Variables	Classes or Name	Frequency	Proportion
Age	<18	903	14.3
	19-40	1766	28
	41-60	2680	42.5
	>60	961	15.2
Cancer Grade	1	1008	16
	2	1953	31
	3	1639	26
	4	1710	27
Diagnosis Subtype	Chordoma	1581	25.1
	Ependymoma	853	13.5
	Glioblastoma	1000	15.8
	Medulloblastoma	938	14.9
	Oligodendroglioma	955	15.1
	Astrocytoma	983	15.6
Tumor Size	<2cm	1419	22.5
	2-5cm	1490	23.6
	5-10cm	2120	33.6
	>10cm	1281	20.3
Drug	Temozolomide	3344	53
	Bevacizumab	1450	23
	Vincristine	599	9.5
	Etoposide	331	5.2
	Temozolomide & Bevacizumab	586	9.3

6. Statistical Data Analysis

The scikit-learn library in Python programming will be used for classification using the KNN and RNN algorithms, which considers four risk factors in brain tumor: age, gender, diagnosis subtype, and tumor size, along with the labeled variable drug, each risk factor will have a learning curve, test accuracy, and confusion matrix alongside explanations of them, respectively, as follows:

6.1. Age

Figures 5-1 and 5-2 show the K-Nearest Neighbours (KNN) classifier; the learning curve for age classification shows that the performance of the model increases with the size of the training set. The accuracy of the test is 96%, and that is a good generalisation. For the confusion matrix, the model predicts most accurately, especially in (19-40 and 41-60). Alongside misclassification are low rates,

which means strong performance. The classification report shows precision, recall, F1-score, and support for four classes that have high scores, indicating a reliable and accurate model for classifying age. The model displays strong learning and generalisation capabilities for age group classification, with high accuracy, and the confusion matrix indicates most predictions are correct across all age categories.

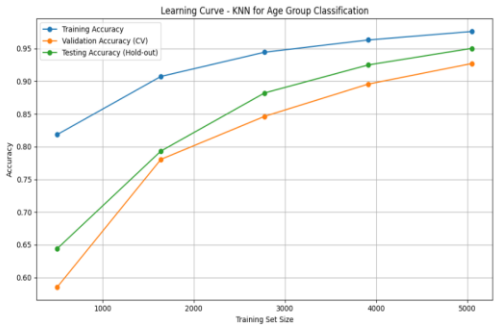

Figure (5-1): learning curve.

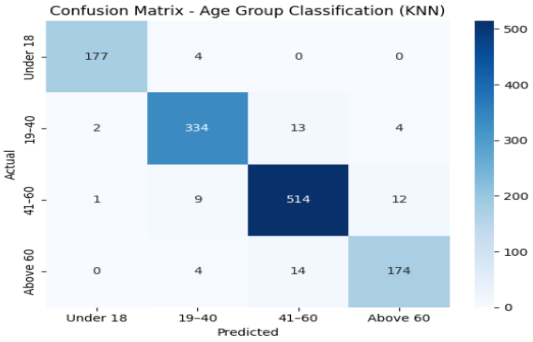

Figure (5-2): confusion matrix.

For RNN classifier, accuracy and loss over epochs demonstrate the accuracy of the validation, testing, and training records over time. It shows that the training accuracy continues to improve and levels out at about 0.85, while the validation accuracy rises and levels off at about 0.82. The test's accuracy is roughly 87.48%. As training goes on, losses for both groups go down, showing that the system is learning efficiently, and an ambiguity matrix shows the accuracy of the model's predictions for the various age groups. It demonstrates that most of the expectations were correct for each group, especially the 41–60 age group. The classification report shows the model is successful at separating items due to its high recall and precision scores, especially for the 41–60 group. The technique works effectively for people of all ages along with various parameters, with high accuracy, plus an appropriate compromise between precision and recall.

6.2. Cancer grade

Figures 6-1 and 6-2 display a learning curve and categorisation for the KNN model used to classify cancer grades. Larger training sets result in higher accuracy curves, with testing accuracy reaching 0.95, validation accuracy increasing to 0.92, and training accuracy approaching 0.96. The classifier, grade 2, has the most accurate predictions with 382 correctly classified samples. Despite frequent misclassifications between adjacent grades, the model accurately classifies cancer grades, with learning curves suggesting it may stabilise with more data. Its predictions are vital across all grades, and the KNN classifier is effective, although there is some confusion between adjacent grades.

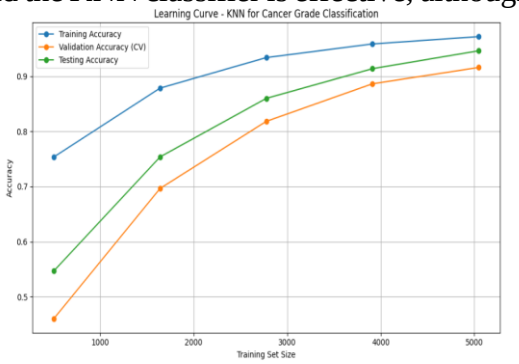

Figure (6-1): learning curve.

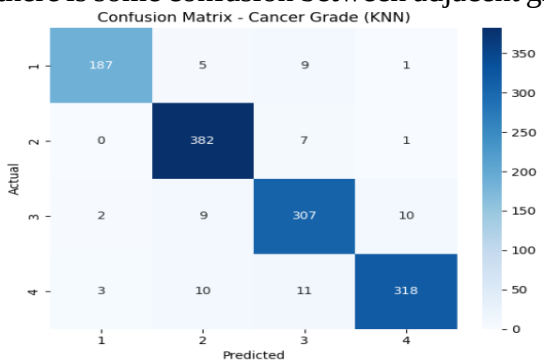

Figure (6-2): confusion matrix.

The RNN model's ability to generalise is demonstrated by the training accuracy approaching 1.0 and the validation accuracy remaining high. The loss passes rapidly during training, resulting in the model learning faster and working better in the long run. A majority of the estimates are correct, especially for groups two and four. But there are some mistakes at close levels, like first and second grade, getting mixed up. The categorisation report says that the general outcome was good, with an accuracy rate of 88.91% on the test. The algorithm does a fantastic job at estimating cancer levels, with only a little bit of uncertainty within classes that are close to each other.

6.3. Diagnosis Subtype

Figure (7-1 and 7-2) shows that the KNN model's efficiency is measured by an average of weights and a macro average with 93% accuracy. The model works well since there are significant numbers along the whole diagonal, which means that the classifications are correct. Diagonal values, like 296 for chordoma, show cases that were correctly classified. Most of the diagnoses have high diagonal values, which means they were correctly classified. There is substantial uncertainty over the exact rates of classification between subtypes, especially between glioblastoma and medulloblastoma. More samples may increase the model's performance, according to learning curves. The classification report confirms the accuracy of predicting diagnosis subtypes. Even with occasional inaccuracies from comparisons or overlaps across subtypes, the confusion matrix reveals that the model correctly identifies diagnosis subtypes.

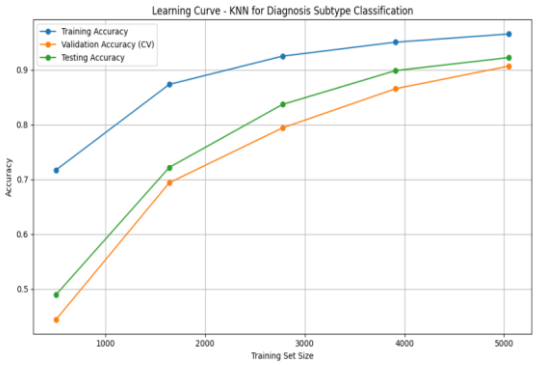

Figure (7-1):learning curve.

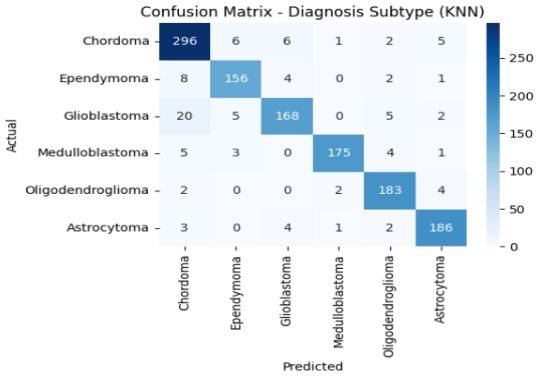

Figure (7-2):confusion matrix.

Both the validation and learning accuracy increase throughout epochs before plateauing in RNN algorithms. The validation accuracy closely follows this instruction accuracy. The framework's prediction improves with learning, as evidenced by the accuracy of the loss, and this decreases over epochs across the training and validation sets. The diagnostic findings for six distinct diagnoses are displayed in an ambiguous matrix. While entries off the diagonal display incorrect classifications, entries on the diagonal reflect correct classifications. For example, glioblastoma and ependymoma, having the same type, are occasionally misclassified, but real chordoma is correctly identified 265 times. Overall, the model's accuracy is 93.99%. Each subtype's precision, recall, and F1 score are likewise high, typically above 0.88, indicating that the classes are performing better and the model is still successful in classifying this diagnosis in RNN algorithms.

6.4. Tumour Size

KNN on tumour size learning curve, Figure (8-1 and 8-2) illustrates training and validation/testing accuracy improving with training set size. When accuracy is 0.94, the model is learning and generalising well. The lines' closeness suggests excellent model performance concerning minimum overfitting. 94% accuracy with all metrics at 0.94, the macro average—an average of precision, recall, plus F1-score across classes—indicates constant performance across all categories—weighted Average: 0.94, including class support. The confusion matrix's diagonal cells (268, 283, 400, 238) indicate the number of correctly classified instances. The off-diagonal cells represented instances where the model predicted a different size category than the actual one. 268 tumours classified as less than 2 cm have the highest number of accurate classifications. Found 283 tumours measuring 2–5 cm. The study has accurately classified 400 tumours that measure between 5 and 10 cm and found 238 tumours larger than 10 cm. The model achieves an excellent test accuracy of 94.45% in classifying tumour sizes, the classification report shows accurate predictions for every size category, and the confusion matrix indicates most cases are correctly classified. The model's overall success is attributed to its ability to predict tumour sizes consistently.

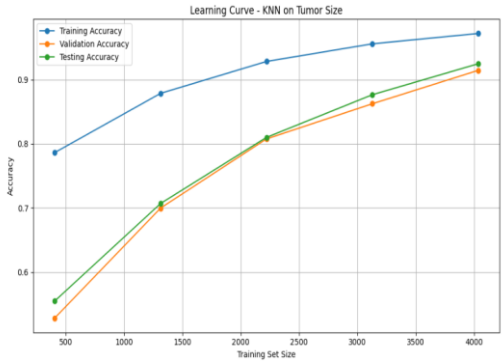

Figure (8-1): learning curve.

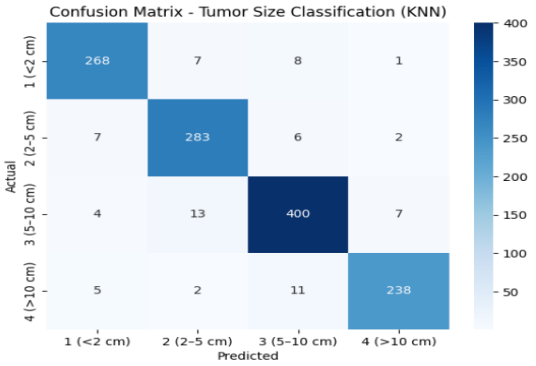

Figure (8-2): confusion matrix.

For RNN, a model's effective learning and strong extrapolation to new data are demonstrated by the increasing precision of both training and validation over time, which plateaus at extremely high values of about 0.9. The confusion matrix compares the model's predictions to the actual tumour size groups. The diagonal displays the correct classifications: 267 tumours were correctly classified as less than 2 cm, 305 as 2–5 cm, 401 as 3–10 cm, and 218 as larger than 10 cm. Incorrect classifications highlight a few similarities in size groupings, such as the assumption that six tumours with sizes ranging from 2 to 5 cm should be smaller than 2 cm. With a total accuracy of 94.37% along with excellent precision, recall, and F1-score for each class, the classification report indicates that the model has a strong ability to classify tumour sizes. For tumours under 2 cm, for instance, the F1-score is 0.97, the recall is 0.95, and the precision is 0.99.

6.5. Drug

The Figure (9-1) illustrates model accuracy improving with training set size. High training accuracy (ranging from 0.75 to 0.87) indicates that the algorithm accurately matches the training data. Testing with validation accuracy increased to 0.75–0.78 with more data. Precision (0.81) plus recall (0.88) make temozolomide a useful predictor. In Figure (9-2), most items are correctly identified when the numbers on the diagonal are large (for instance, 591 temozolomide were correctly recognised as temozolomide, 177 bevacizumab were accurately predicted as bevacizumab, and so on). Misclassifications: Off-diagonal numbers indicate prediction errors. For instance, we mistakenly identified 86 real cases of bevacizumab as temozolomide. Mistakenly identified 47 cases of temozolomide as bevacizumab. Particularly, when a class called "Temozolomide, Bevacizumab" is likely an integrated or unclear class that predicts itself as other classes, confusion results; the performance of the KNN classifier improves as more training data is added. On unseen data, it achieves an accuracy of approximately 78%, and model performs reasonably well but has some misclassification, especially between closely related classes or ambiguous cases.

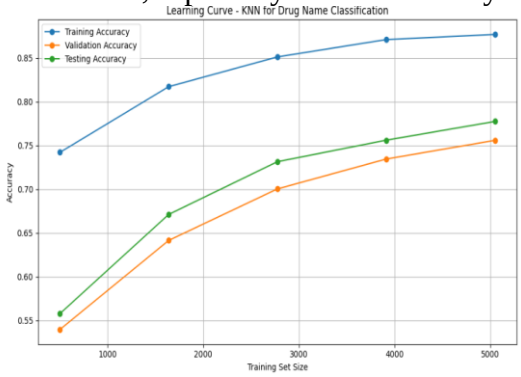

Figure (9-1): learning curve.

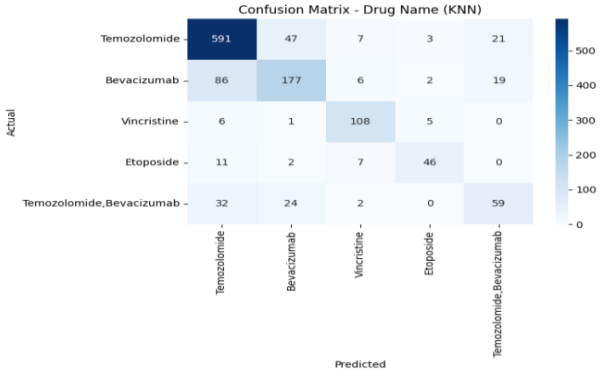

Figure (9-2): confusion matrix.

The RNN model fits the training data and achieves a high accuracy rate (~90%). The loss plot shows a consistent decrease in training, validation, and test accuracy. In the confusion matrix, the diagonal is classified correctly, and misclassifications are such that vincristine has 689 true positives and temozolomide has 455. The model is nearly 87.7%. It also provides precision, recall, and F1-score for individual classes, which reflect the model's ability to identify each class and its correctness

in predictions. The macro and weighted averages for these metrics are both around 0.88, demonstrating balanced performance across all the classes being analysed. The RNN algorithms execute well and have a balanced matrix that's reliable for classes.

6.6. Drug with four risk factors

As shown in Figures 10-1 and 10-2, the model's precision and recall metrics are balanced by the F1-score, which represents the harmonic mean of precision and recall. The weighted average, macro-average, and overall accuracy scores reflect the model's overall performance. As training size increases, the model achieves a high F1-score on the training set, while the F1-score on the validation set improves as the dataset size increases. The confusion matrix's diagonal elements, represented by darker blue, indicate accurate predictions. For instance, vincristine and etoposide were correctly classified, with 95% and 94% accuracy, respectively. Bevacizumab and temozolomide showed high accuracy, with 76% to 82%, respectively. Misclassifications between bevacizumab and temozolomide are less common, and normalisation helps compare performance across classes, despite class imbalances.

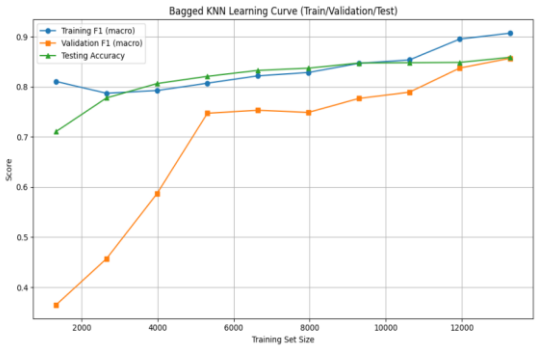

Figure (10-1): learning curve.

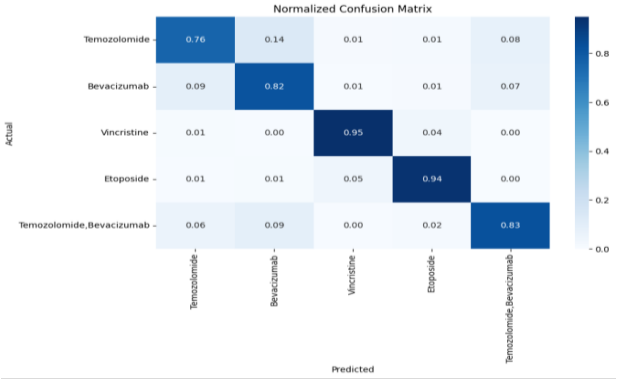

Figure (10-2): confusion matrix.

For the RNN model, the accuracy of the test, validation, and training has been steadily going up throughout 200 epochs. The accuracies for validation and testing stay close to each other, settling at a little over 0.78. The accuracy for training goes up to almost 0.80. The validity and test loss curves on the correct plot overlap and flatten below 0.5. The training, validation, and test losses are all going down at the same time. The normalised confusion matrix shows how well the model can distinguish between different types of drugs. The diagonal elements show instances that were correctly predicted. The accuracy of every drug type's classification is as follows: Vincristine: 91 per cent, Etoposide: 89 per cent, Temozolomide: 59 per cent, and Bevacizumab: 70 per cent. 78% of people who took bevacizumab and temozolomide along most of the off-diagonal numbers are relatively low. However, there is considerable misunderstanding between vincristine and etoposide and between temozolomide and the mixture of temozolomide and bevacizumab (16 per cent). This report models performance for each class: Vincristine had the highest performance: Precision: 0.86, Recall: 0.91, F1-score: 0.89. Etoposide also performed well with an F1 score of 0.87. Temozolomide and bevacizumab had relatively lower scores (F1-scores of 0.65 and 0.70, respectively). Temozolomide + Bevacizumab had decent results: F1-score: 0.75. Overall accuracy across all drug classes was 0.77, with a macro and weighted average of 0.77 for all metrics. Finally, the RNN Works well most of the time, especially with drug interactions and time patterns.

7. Conclusion

The KNN classifier demonstrates exceptional accuracy and reliability, with a testing accuracy of 92-94%, training accuracy of 94-96%, and validation 91-93%. And, the RNN learned well, with a testing accuracy of 78-93%, training accuracy of 80-100%, and validation accuracy of 78-94%. Particularly in classes that were adequately represented, models achieved high precision, recall, and F1-scores for specific groups, such as age and diagnostic subtype. This paper demonstrates how machine learning algorithms can accurately identify these trends and how medication prescriptions are strongly correlated with unique patient characteristics. Even if a few categories were identical or

close to each other (for example, cancer levels one and two or drug types that were comparable, such as temozolomide and temozolomide + bevacizumab), the most unique classes had few errors. The results reveal that two algorithms tend to manage unequal classes effectively, generalise well, and do so without overfitting. In general, the results confirm the notion that predictive algorithms taught regarding clinical trial risk factors can sort drug kinds and that these risk factors play a significant role in the process of making decisions. This gives us a powerful basis for employing machine learning in personalised healthcare and developing treatments.

8. Recommendation

Incorporate the classification algorithm KNN and RNN into a hospital's electronic medical record system to help physicians with prognostic monitoring, treatment planning, and initial diagnosis—particularly in situations where quick and precise classification is essential and using image classification techniques, such as CT scans and MRIs, for diagnosing brain tumours. Next, I recommend writing a history for every patient to make it easier for the doctor to avoid asking about treatment at each clinic visit. Additionally, I recommend taking another risk factor, such as the disease diagnosis history of patients. Finally, Patients should see one doctor; there's no need to switch if it's not necessary.

9. Supplementary material

(None).

10. Author's Contributions

Nigar Abdulla Ghafur designed the research, wrote and edited, Abbas Gulmurad Beg Murad conducted the analyses, and interpreted the results.

11. Funding

(None).

12. Data Availability Statement

The daily recorded dataset from Hiwa hospital includes a sample of 536 patients diagnosed with brain tumours, gathered from 2018 to 2024 over a three-month duration.

13. Acknowledgements

The authors would like to thank the Hiwa hospital for supplying some brain tumour data.

14. Conflict of interest

The authors declare no conflict of interest.

References

[1] Alnuaimi, A. F., & Albaldawi, T. H. (2024). "An overview of machine learning classification techniques". In BIO Web of Conferences (Vol. 97, p. 00133). EDP Sciences. 10.1051/bioconf/20249700133

[2] Bottrighi, A., & Pennisi, M. (2023). "Exploring the state of machine learning and deep learning in medicine: A survey of the italian research community". Information, 14(9), 513. 10.3390/info14090513

[3] Chavda, V. P., Anand, K., & Apostolopoulos, V. (Eds.). (2023). "Bioinformatics tools for pharmaceutical drug product development". John Wiley & Sons. 10.1002/9781119865728

[4] Hu, L. Y., Huang, M. W., Ke, S. W., & Tsai, C. F. (2016). "The distance function effect on k-nearest neighbor classification for medical datasets". SpringerPlus, 5, 1-9. https://doi.org/10.1186/s40064-016-2941-7

[5] Iqbal, M. J., Javed, Z., Sadia, H., Qureshi, I. A., Irshad, A., Ahmed, R., ... & Sharifi-Rad, J. (2021). "Clinical applications of artificial intelligence and machine learning in cancer diagnosis: looking into the future". Cancer cell international, 21(1), 270. 10.1186/s12935-021-01981-1

[6] Jain, G. (2022). "Application of Machine Learning in Drug Discovery and Development Lifecycle". Int J Med Phar Drug Re, 16. 10.22161/ijmpd.6.6.4
[7] Khaldi, B., Babelhadj, Z., & Khemgani, S. (2021). "Classification of brain tumor using some approaches of machine learning*"* (Doctoral dissertation, University of Kasdi Merbah Ouargla). https://dspace.univ-ouargla.dz/jspui/handle/123456789/35039
[8] Krzyszczyk, P., Acevedo, A., Davidoff, E. J., Timmins, L. M., Marrero-Berrios, I., Patel, M., ... &Yarmush, M. L. (2018). "The growing role of precision and personalized medicine for cancer treatment". Technology, 6(03n04), 79-100.10.1142/s2339547818300020
[9] Kumar, L. A., Satapathy, B. S., Pattnaik, G., Barik, B., Patro, C. S., & Das, S. (2021). "Malignant brain tumor: Current progresses in diagnosis, treatment and future strategies". Ann Rom Soc Cell Biol, 25(6), 16922-32. (PDF) Malignant brain tumor: Current progresses in diagnosis, treatment and future strategies
[10] Maleki Varnosfaderani, S., & Forouzanfar, M. (2024). "The role of AI in hospitals and clinics: transforming healthcare in the 21st century". Bioengineering, 11(4), 337. https://doi: 10.3390/bioengineering11040337
[11] Mitusova, K., Peltek, O. O., Karpov, T. E., Muslimov, A. R., Zyuzin, M. V., & Timin, A. S. (2022). "Overcoming the blood–brain barrier for the therapy of malignant brain tumor: current status and prospects of drug delivery approaches". Journal of nanobiotechnology, 20(1), 412. 10.1186/s12951-022-01610-7
[12] Nababan, A., Khairi, M., & Harahap, B. (2022). "Implementation of K-Nearest Neighbors (KNN) algorithm in classification of data water quality". J. Mantik, 6(1), 30-35. https://iocscience.org/ejournal/index.php/mantik/article/download/2130/1669
[13] Naimi, A., Deng, J., Shimjith, S. R., & Arul, A. J. (2022). "Fault detection and isolation of a pressurized water reactor based on neural network and k-nearest neighbor". IEEE Access, 10,17113-17121. 10.1109/ACCESS.2022.3149772 https://doi.org/10.2174/1570159X18666200626204005
[14] Osisanwo, F. Y., Akinsola, J. E. T., Awodele, O., Hinmikaiye, J. O., Olakanmi, O., & Akinjobi,J. (2017). "Supervised machine learning algorithms: classification and comparison". International Journal of Computer Trends and Technology (IJCTT), 48(3), 128-138. 10.14445/22312803/IJCTT-V48P126
[15] Parvin, H., Alizadeh, H., & Minaei-Bidgoli, B. (2008). "MKNN: Modified k-nearest neighbor". In Proceedings of the world congress on engineering and computer science (Vol. 1). Newswood Limited. Microsoft Word - Final CameraReady_MKNN_WCECS2008.doc
[16] Qian, T., Zhu, S., & Hoshida, Y. (2019). "Use of big data in drug development for precision medicine: an update". Expert review of precision medicine and drug development, 4(3), 189-200. https://doi.org/10.1080/23808993.2019.1617632
[17] Raghavapudi, H., Singroul, P., & Kohila, V. (2021). "Brain tumor causes, symptoms, diagnosis and radiotherapy treatment". Current Medical Imaging Reviews, 17(8), 931-942. https://doi.org/10.2174/1573405617666210126160206
[18] Sarwar, A., Ali, M., Manhas, J., & Sharma, V. (2020). "Diagnosis of diabetes type-II using hybrid machine learning based ensemble model". International Journal of Information Technology, 12, 419-428. https://doi.org/10.1007/s41870-018-0270-5
[19] Shetty, S. H., Shetty, S., Singh, C., & Rao, A. (2022). "Supervised machine learning: algorithms and applications". Fundamentals and methods of machine and deep learning: algorithms, tools and applications, 1-16. http://dx.doi.org/10.1002/9781119821908.ch1
[20] Suyal, M., & Goyal, P. (2022). "A review on analysis of k-nearest neighbor classification machine learning algorithms based on supervised learning". International Journal of Engineering Trends and Technology, 70(7), 43-48 / https://doi.org/10.14445/22315381/IJETT-V70I7P205
[21] Taha, A. M., Ariffin, D. S. B. B., & Abu-Naser, S. S. (2023). "A systematic literature review of deep and machine learning algorithms in brain tumor and meta-analysis". Journal of Theoretical and Applied Information Technology, 101(1), 21-36. www.jatit.org . Microsoft Word - 3 47528 44252851 jatit-A Systematic Literature Review-ashraf -28-11-2022
[22] Woodman, R. J., & Mangoni, A. A. (2023). "A comprehensive review of machine learning algorithms and their application in geriatric medicine: present and future". Aging Clinical and Experimental Research, 35(11), 2363-2397. 10.1007/s40520-023-02552-2

https://doi.org/10.31272/jae.i149.1449
https://admics.uomustansiriyah.edu.iq
P-ISSN: 1813–6729 E-ISSN: 2707-1359

تطبيق خوارزميات التعلم الآلي لتحديد أنواع العلاج لمرضى سرطان الدماغ

نيكار عبد الله غفور
قسم الإحصاء والمعلوماتية، كلية الإدارة والاقتصاد، جامعة السليمانية، السليمانية، العراق.
Email: nigarghafur@yahoo.com , ORCID: https://orcid.org/0009-0009-5417-0033

عباس كولمراد بك مراد
قسم الإحصاء والمعلوماتية، كلية الإدارة والاقتصاد، جامعة السليمانية، السليمانية، العراق.
Email: abbas.beg@univsul.edu.iq, ORCID: https://orcid.org/0009-0003-4785-699X

المستخلص

تم استخدام الذكاء الاصطناعي بالفعل في العلاجات الطبية لفترة من الوقت ، لكن تقدمه على مدى السنوات العشر الماضية مثير للإعجاب. في الطب ، أصبحت كل من أساليب التعلم الآلي والتعلم العميق أكثر شيوعا نظرا لتوفر المزيد من البيانات المتعلقة بالصحة ، وهذه التقنيات أفضل في التعامل مع مجموعات البيانات الضخمة ، وأصبحت أجهزة الكمبيوتر أكثر قوة.

هدف هذه الرسالة هو البحث في طرق استخدام التعلم تحت الإشراف، لا سيما خوارزميات التصنيف في تعلم الآلة والتعلم العميق ، لإيجاد أفضل طريقة لاستخدام خوارزميات ك-الأقرب الجار والشبكة العصبية المتكررة لتصنيف علاجات الأورام الدماغية. تم استخدام خوارزميات تعلم الآلة، ك- الأقرب الجار ، والشبكة العصبية التكرارية لرؤية كيفية تغيير عوامل الخطر لطريقة تصنيف الأدوية للمرضى. لقد عملوا الخوارزميات بشكل جيد. تشير النتائج إلى أن طرق التوقع يمكن أن تجمع الأدوية وتكون مفيدة في اتخاذ القرارات. هذه التطورات هي خطوة نحو استخدام الآلات والتعلم العميق في العلاج الطبي.

معلومات البحث

تواريخ البحث:
التقديم: 26 / 06 / 2025
المراجعة: 02 / 08 / 2025
قبول النشر: 10 / 08 / 2025
نشر الكتروني: 01 / 09 / 2025
تسلسل الصفحات: 121 - 132

الكلمات المفتاحية:
التعلم الآلي، *ك-الأقرب الجار*، الدقة، النوع الفرعي للتشخيص، علاج أورام الدماغ.

المراسلة:
أسم الباحث: نيكار عبد الله غفور

Email: nigarghafur@yahoo.com