Faculty of Natural and Applied Sciences Journal of Scientific Innovations
Print ISSN: 2814-0877 e-ISSN: 2814-0923
www.fnasjournals.com
Volume 5; Issue 2; December 2023; Page No. 1-16.

GLASS RECOGNITION AS A CUTTING-EDGE MACHINE LEARNING APPROACH FOR IDENTIFICATION AND CLASSIFICATION

[*1]Nanwin, D.N, & [2]Ofor, W.D.
[1]Department of Computer Science, Ignatius Ajuru University of Education, Port Harcourt, Nigeria
[2]Department of Computer Science, Rivers State University, Port Harcourt, Nigeria

***Corresponding author email:** kakusman@yahoo.com

Abstract
Crime scene investigations are integral to the justice system, and the efficacy of such investigations can be significantly heightened through precise identification and classification of glass recovered at crime scenes. This research endeavours to address the persistent demand for accurate glass categorization in legal proceedings. The study introduces a comprehensive glass identification and classification model employing a combination of deep neural networks and decision tree algorithms. Utilizing a glass dataset sourced from the UCI machine learning data repository, the research aims to showcase the capabilities and analyze the effectiveness of the developed models. The identification and classification model demonstrated commendable performance, achieving an accuracy of 68% for the training dataset and 60% for the test dataset. The minimal loss of 4.376815 and mean absolute error of 1.412363 underscore the robustness of the proposed model. The investigation employs a variety of tools and test techniques, with a particular emphasis on a confusion matrix table to illustrate misclassifications. Additionally, a tabulated probability distribution of the model provides a nuanced understanding of the percentage distribution in classified data. This research contributes valuable insights into the intersection of forensic science and machine learning, showcasing the potential for advanced technologies to enhance the accuracy of crime scene glass analysis. The findings emphasize the need for continued refinement and development of such models in forensic investigations, paving the way for more sophisticated tools in legal proceedings.

Keywords: Glass Classification, Identification, Machine learning, Deep Neural Network, Data Visualization.

Introduction

Glass is used for many functional purposes, such as providing light to buildings, but it also serves creative purposes. We wouldn't have mirrors without glass, and driving would become less safe. Digital screens, mobile phone screens, and television screens are all made of glass. Corrective lenses are glasses used to rectify vision. According to Faizal and Adarsh (2022), the glass industry is widely regarded as one of the most important in the world. Glass is used in a wide range of applications, from car glass and window glass to water bottles and X-Ray and Gamma Ray shielding. The glass is an amorphous solid that is transparent and non-crystalline and is used in a variety of industries. It can be used in a variety of ways, and investigators should be able to determine which type of glass is present during a crime scene examination.

To aid in the criminal investigation, a glass classification problem study was carried out. If the remaining glass is properly identified, it can be used as evidence in the event of a crime. The categorization of glass from the crime scene and glass particles found associated with the crime is a constant need for a lawsuit. These glass particles are typically very small. Viewing and comparing these small pieces of glass that will be important during a forensic context is required (Mathur & Surana, 2020). Every type of glass is composed of various elements with varying unit measurements and Refractive indices. Glass properties, particularly the refractive index, are affected by the composition and treatment of the glass. Sodium (Na), Magnesium (Mg), Calcium (Ca), Barium (Ba), Iron (Fe), Silicon (Si), Aluminium (Al), and Potassium (K) are the elements used in the manufacture of various types of glasses. According to recent studies, 60% of glass-related cases provided some positive evidence, and 40% of these cases provided strong evidence. Depending on the circumstances, the findings may also refute a person's involvement in a crime (El-Khatib et al., 2019).

Machine Learning (ML) is a vast scientific discipline that draws on concepts from computer science, statistics, cognitive science, engineering, optimization theory, and many other mathematical and scientific disciplines. There are lots of applications for machine learning, but data mining is the most important. Machine learning can be divided into two broad categories: supervised machine learning and unsupervised machine learning (Soofi & Awan, 2017). Machine learning algorithms use the same set of features to represent every instance in any dataset. The characteristics can be continuous, categorical, or binary. In contrast to unsupervised learning, which occurs when instances are given with unknown labels (the corresponding correct outputs), supervised learning occurs when instances are given with known labels (Kotsiantis et al., 2006). Machine learning is a technique used to teach machines how to handle data more efficiently and accurately. In some cases, even after viewing the data, we are unable to understand the behaviour or obtain information from it. In this case, we use machine learning techniques to classify the data. A large number of datasets are available from various sources, and there is a demand for machine learning. Machine learning is being used in a variety of industries, from medicine to the military, to extract relevant information from available datasets. Machine learning's primary goal is to learn from existing data. A large number of algorithms are being developed to teach machines to learn on their own (Reddy & Babu, 2018). Hence, the development of the glass identification and classification model with machine learning technique.

According to Blum (2007), Machine Learning is concerned with the development of programs that can learn rules from data, adapt to changes, and improve performance over time. Machine Learning has become critical as computers are expected to solve increasingly complex problems and become more integrated into our daily lives, in addition to being one of the initial dreams of Computer Science. Martens and Baesens (2010) in their paper Building an Acceptable Classification Model in Data Mining opined that classification is an important data mining task, where the value of a discrete (dependent) variable is predicted, based on the values of some independent variables. Classification models should provide correct predictions on new unseen data instances. Coleman (2002) in her view explained that scientific models are critical, analytical tools and objects that can be included in digital libraries. Her paper presents a preliminary classification scheme for the cataloguing of scientific models. Supervised classification is one of the tasks most frequently carried out by intelligent systems. This paper describes various Supervised Machine Learning (ML) classification techniques, compares various supervised learning algorithms as well as determines the most efficient classification algorithm based on the data set, the number of instances and variables (Osisanwo et al., 2017).

The glass sector is regarded as one of the world's most important. Glass is used in a multitude of scenarios, ranging from water bottles to X-ray and gamma-ray shielding. This is a transparent, non-crystalline, amorphous solid. Glass is used in a variety of ways, and investigators should be able to identify which type of glass is present during a crime scene assessment. However, every product and equipment needs a unique glass, for that reason within the industrial world, there are differing kinds of glass that may make your domestic and industrial life easier (Faizal &Adarsh, 2022). The design of new glasses is often plagued by poorly efficient Edisonian “trial-and-error” discovery approaches. As an alternative route, the Materials Genome Initiative has largely popularized new approaches relying on artificial intelligence and machine learning for accelerating the discovery and optimization of novel, advanced materials (Liu et al., 2019). An overview of structural glass engineering applications with artificial intelligence was done by these authors. Big data and the use of 'Artificial Intelligence' (AI) is currently advancing due to the increasing and even cheaper data collection and processing capabilities. Social and economic change is predicted by numerous company leaders, politicians and researchers (Kraus & Drass, 2020).

There exist different types of glasses that are being used for various purposes. These glasses are made up of different properties that can help distinguish them from one another. The identification of these glasses is a function of identifying the constituent of the glass being examined so that it can be classified properly. Investigating a crime scene and an individual that has these pieces/ particles of glass in their body, one will tend to match them together if the same pieces of glass in the suspect body are the same pieces of glass in the crime scene. The problem statement of the project work is an identification problem, being able to classify from the different types of glass which one is which. That is, to identify at a crime scene or any given time which glass is a mirror, window, windshield, television, tumbler etc. To be able to distinguish this kind of scenario forms the problem statement of this study. The study aims to develop a glass identification and classification model using machine learning techniques. The following objectives are used to achieve this study. To acquire the required glass dataset and pre-processing the same for feature collection and analysis, training the model with the acquired dataset using machine learning techniques, identify and classify the result of the trained model of the glass and carry out the evaluation of the accuracy and performance of the result by creating a confusion matrix.

Materials and Methods

The proposed glass identification and classification model employs machine learning techniques with a primary focus on identifying and categorizing various types of glass in specific scenarios such as crime scenes or decision-making events. The model utilizes artificial intelligence methods to compare the refractive indexes (RI) and constituent elements of different glass types, including mirrors, bottles, windshields, and television screens. The process involves training the model with a collected glass dataset containing seven distinct types of glass, allowing the system to analyze and classify new glass samples based on their features.

The methodology begins with loading the dataset into the system, followed by comprehensive data preprocessing to enhance classification accuracy and overall performance. Standardization and preparation of the data are crucial steps in this phase. Data visualization techniques are then applied to unveil intricate features and identify outliers, contributing to improved model performance. The dataset is subsequently transformed into training and testing sets. The training set is fed into both a deep neural model and a decision tree model, enabling the model to learn and generalize from the provided data. The evaluation of the model is conducted using the testing set, and performance metrics such as accuracy, loss, and mean absolute error are assessed. This methodology ensures that the glass identification and classification model is robust, well-trained, and capable of accurately categorizing various types of glass in real-world scenarios, thereby contributing to advancements in crime scene investigations and decision-making processes.

The architectural framework of the system illustrates the various modules of the proposed system and their flow of operation. This project work has followed a specific model in the identification and classification of glass which includes the training and testing of the dataset. Figure 1 is the model and architectural framework of the proposed system.

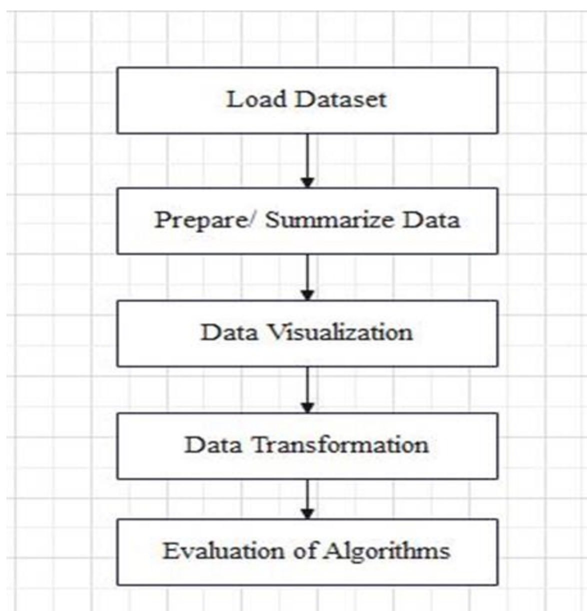


Figure 1: Architectural Model for the Proposed Glass Identification and Classification Model using Machine Learning Technique

The architectural framework in Figure 1 is used to illustrate the process and flow of operation of the proposed system. The framework has five (5) modules of operation and each module plays an important role in the development of the model. When the data is loaded into the development environment, it is prepared summarized and later visualized in different ways. The data is transformed into a form that is useable by the machine learning algorithms and evaluations are done with the algorithms. These modules will be looked into in detail.

A. Load Dataset

The acquired glass dataset from the UCI machine learning Data Repository is loaded into the development environment so that other manipulations can be done on the data. This dataset consists of 214 observations collected at the Home Office Forensic Science Laboratory, Birmingham.

B. Data Preparation and Summarization

The data preparation and summarization have to do with exposing the unknown and underlying structure of the problem to the learning algorithms. Since raw data cannot be used directly, the preparation of the data becomes necessary to make it fit or suitable for model evaluation with the machine learning algorithm. The preparation and summarization is channelled towards the goal of the project and the type of algorithm being used. This may involve data cleaning (identifying errors in the data and correcting them before use), feature selection (the relevant input variables are identified), feature engineering (from the available data, new variables are derived) and creating compact projections through dimensionality reduction of the data. Statistical methods are used to summarize collected data.

C. Data Visualization

Visualization of data is the term used to explore data for qualitative understanding and extraction of information about a dataset. Through data visualization techniques, outliers, corrupt data, and relationships between data and patterns can be identified. Data visualization is a graphical representation of data to express the hidden information. Different types of plots such as Univariate plots, bivariate plots, and multivariate plots are used to understand trends, patterns and outliers in data.

D. Data Transformation

The transformation of data is all about the conversion of data in raw form to a form or structure that is more suitable for building of model and data discovery. There is a likelihood that an algorithm will be biased if it is not transformed into the same scale. When data is transformed into the same scale, it enables the algorithm to compare the relationships among data points. There are three types of transformation, they are clipping method, log transformation and data scaling.

E. Evaluation of Algorithms

Algorithm evaluation is an essential aspect of machine learning and selecting the right algorithm is the foremost in this process. To evaluate a model, different metrics of evaluation are used to understand the model's performance, strengths and weaknesses. This is important in other to assess how efficient a model is during analysis of the research. The various evaluation metrics one can use to assess model efficiency are confusion matrix, classification accuracy, mean square error (MSE), mean absolute error (MAE), logarithmic loss, F1 score and area under the curve.

This study has explored 2 machine learning algorithms in the analysis of the glass dataset for identification and classification model. They are deep neural network (DNN) and decision tree algorithms.

I. Deep Neural Network Algorithm

A deep neural network can be seen as an artificial neural network with multiple hidden layers stacked between the input and output layers. The DNN carries out sophisticated math modelling during data processing and increases the model performance in accuracy. The deepness of the neural network only helps to increase the learning ability and performance metrics of the network which means as the task to solve becomes harder, the learning becomes deeper. Hence, employing this model in the glass dataset can improve the performance accuracy in the identification and classification model.

II. Decision Tree Algorithm

A decision tree algorithm is a hierarchical, non-parametric and supervised machine learning algorithm which consists of a node (known as the root node), branches, internal nodes and leaf nodes. It is employed in classification and regression tasks and its learning ability employs a divide and conquer technique through greedy search in the identification of optimal split points in the tree. Top-down-recursive manner is repeated in the splitting process until the majority or all the data have been classified into specific class labels.

Results

The implementation of the glass identification and classification model has been carried out using the deep neural network (DNN) and decision tree algorithm. The dataset of 214 observation data and 10 attributes has been used in the training and analysis of the model. Table 4.1 describes the glass dataset and its frequencies.

Table 1: Glass Identification Dataset Class Distribution

Class	DESCRIPTION	FREQUENCY	PERCENTAGE
1	Building Windows – Float Processed	70	32.710280
2	Building Windows – Non-Float Processed	76	35.514019
3	Vehicle Windows – Float Processed	17	7.943925
4	Vehicle Windows – Non-Float Processed	Null	Null
5	Containers	13	6.074766
6	Tableware	9	4.205607
7	Headlamps	26	13.551402

Table 1 contains the description, frequency and percentage of the dataset. From the frequency, we can see that building windows (float-processed and non-float-processed) have more frequency than any other glass type. Other glass types are fully represented in the dataset but there are no vehicle windows (non-float processed) in this dataset.

The Data Visualization of Glass Dataset

The glass dataset has been presented in four different types of visualization to demonstrate their constituent information and frequency. This information for each attribute in the dataset has been translated using different shades of charts like boxplot, histogram plot, density plot, correlation plot and scatter plot for each attribute by class value.

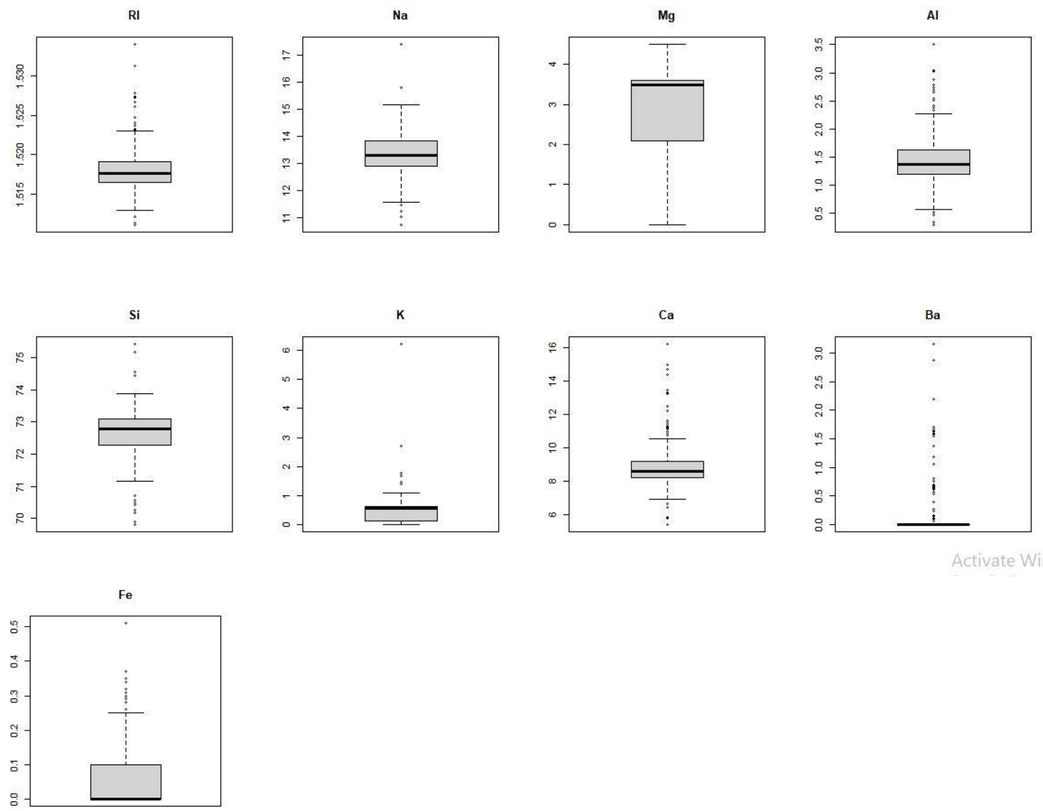


Figure 2: Box Plot for Each attribute by Class Value

Figure 2 is the box plot analysis by class with the presence of outliers in some of the classes provides valuable insights into the distribution of data within different categories. Each box plot visually represents the spread and central tendency of a specific class, with the box indicating the interquartile range (IQR) and the median. Outliers, depicted as individual points beyond the whiskers, highlight instances where data points deviate significantly from the overall pattern within a particular class. The presence of outliers can signify unique characteristics or anomalies within specific groups, potentially uncovering interesting patterns or discrepancies.

Careful consideration of these outliers is crucial for a comprehensive understanding of the dataset, as they may indicate important variations or errors that impact the interpretation of class-specific trends and distributions.

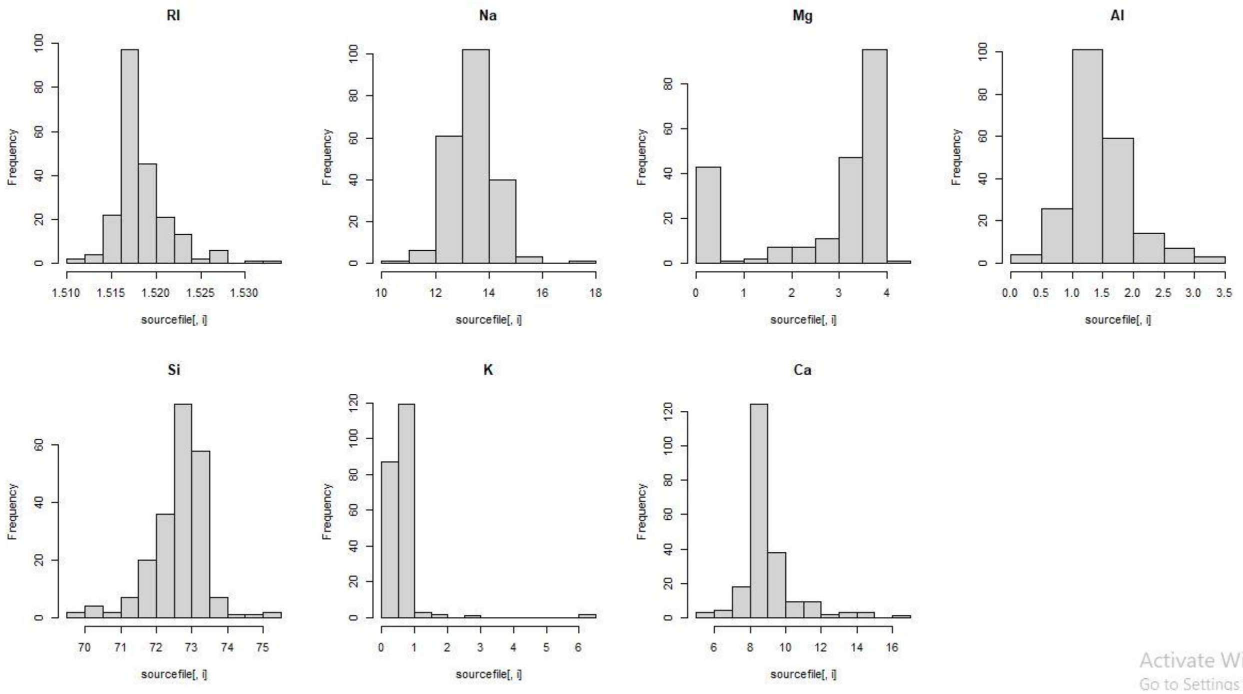
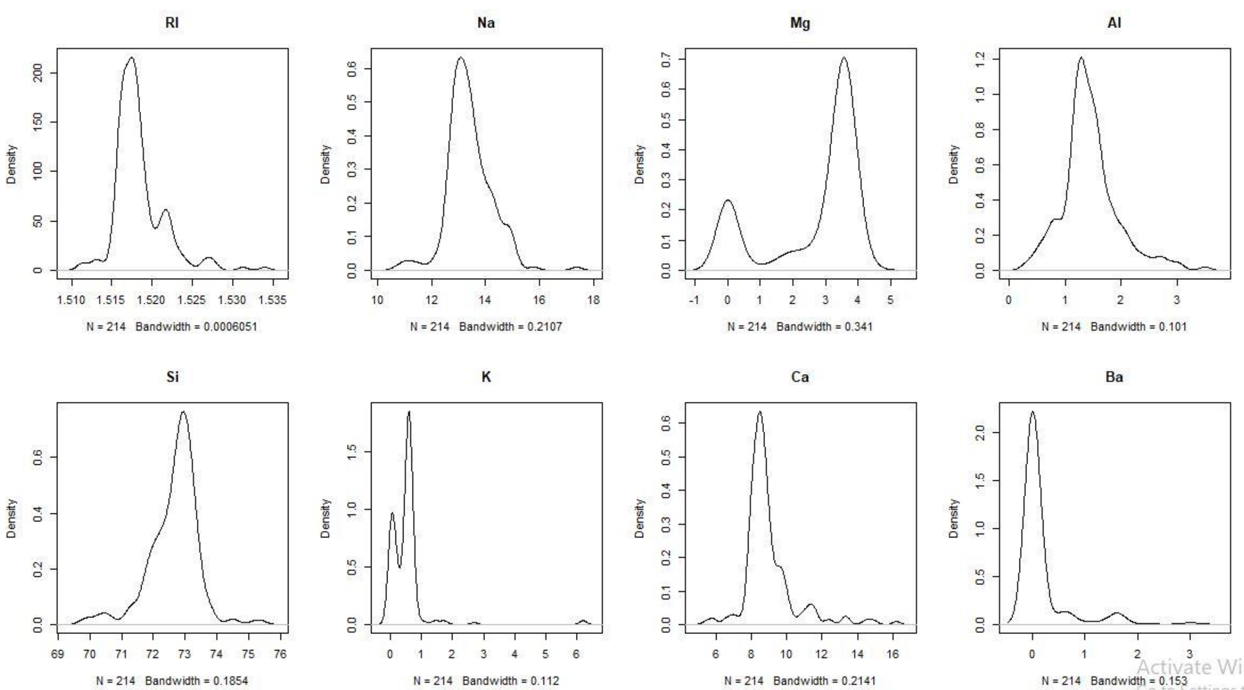


Figure 3: Histogram Plot for Each attribute by Class Value

Figure 3 depicts individual histogram plots for each attribute based on class values. The histograms offer a concise visual representation of the distribution of data within different classes, providing insights into attribute-specific patterns and variations across the dataset.



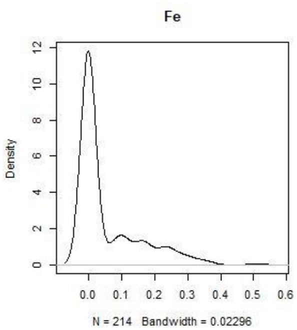


Figure 4: Density Plot for Each attribute by Class Value

In Figure 4, density plots illustrate attribute distributions by class values. These plots offer a smooth representation, highlighting the probability density of data points. The visualization aids in discerning class-specific patterns, showcasing the continuous distribution of attributes across different classes in a concise and informative manner.

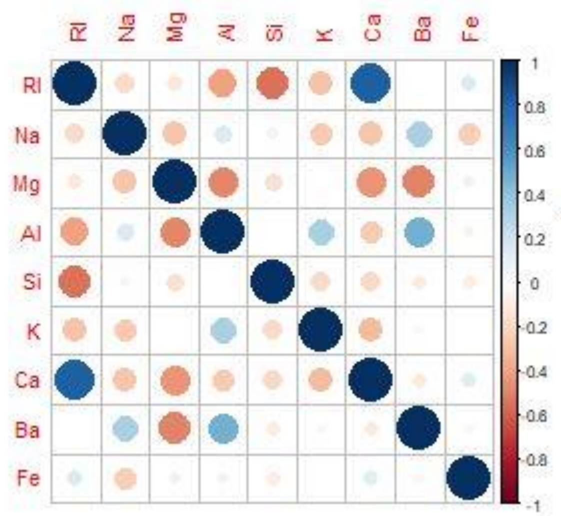


Figure 5: Heat-map Correlation Plot for Each attribute by Class Value

Figure 5 displays a heat-map correlation plot, visually capturing the relationships between attributes and class values. The colour-coded map reveals the strength and direction of correlations, providing a concise overview of attribute-class associations. This aids in identifying key patterns and influences within the dataset.

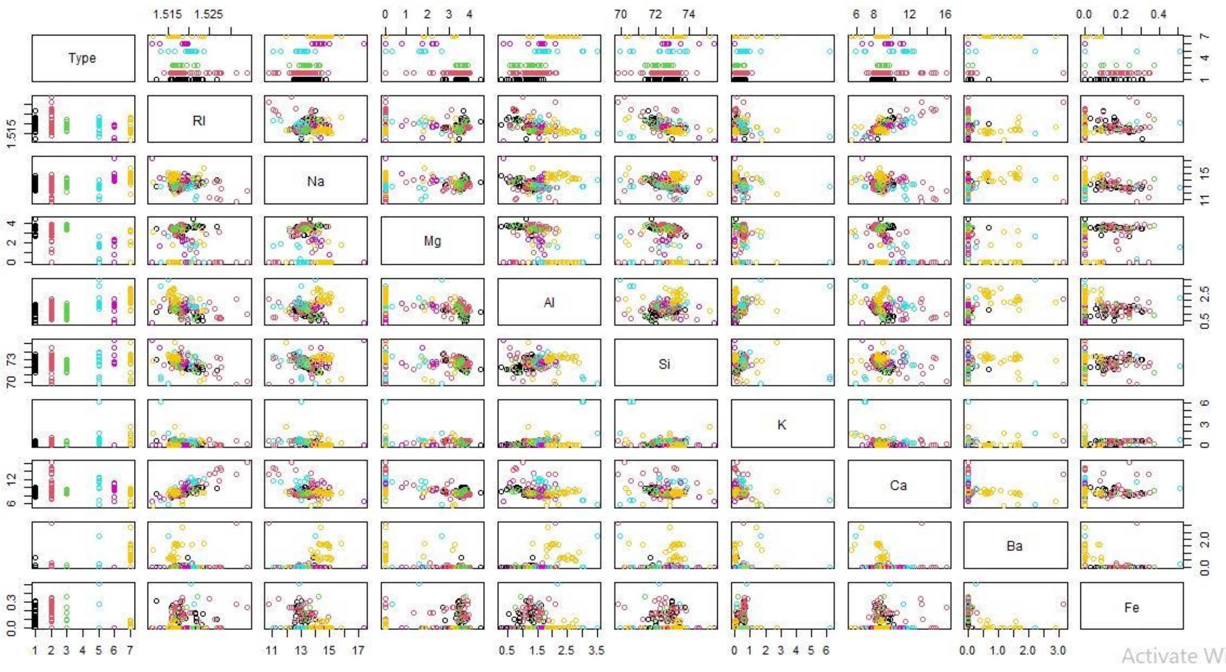


Figure 6: Scatter Plot for Each attribute by Class Value

In Figure 6, scatter plots depict relationships between attributes and class values. Each plot represents a clear visualization of data points, offering insights into attribute variations across different classes. These scatter plots facilitate the identification of patterns and trends within the dataset, aiding in comprehensive data analysis.

Deep Neural Network Evaluation

The dataset was partitioned into two with 70% for training and 30% for testing (or evaluation). The model was first developed using one hidden layer with 5 neurons and an output layer and 9 variables were used as the input variables at the input layer. The activation function used was the rectified linear unit (relu) with 100 iterations (epoch) and 124 samples from the training set were used to train the model. The result of fitting this model is shown in Figure 7.

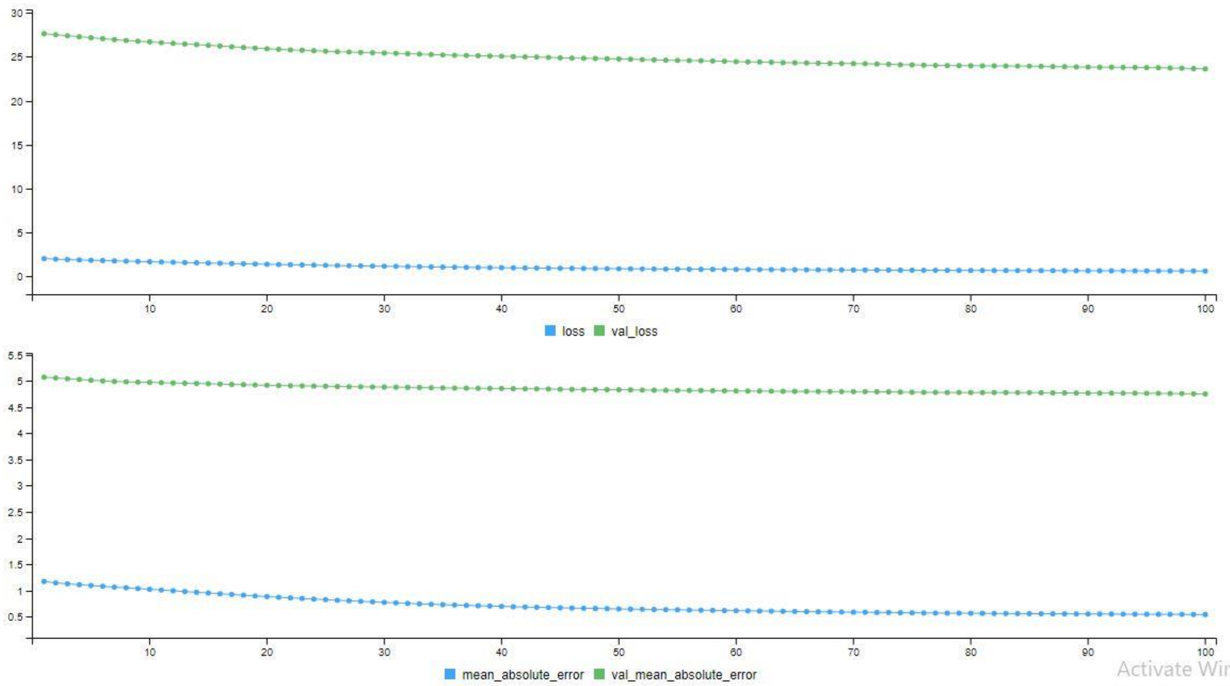


Figure 7: Validation and Mean Absolute Error Analysis for 5 Neurons

From Figure 7, we can see that the gap between the lines of the loss (i.e. loss and validation loss) and error (i.e. mean absolute error and validation mean absolute error) after 100 epochs is very wide. This means, there is no worries for the curve fittings problem but the loss margin and the mean absolute error margin (MAE) were quite high with about 8.612888 loss (which is also known as mean square error) and MAE of 1.923535. We see from the scatter plot when the test target was plotted against the prediction of the model, this gives the graph as seen in Figure 8.

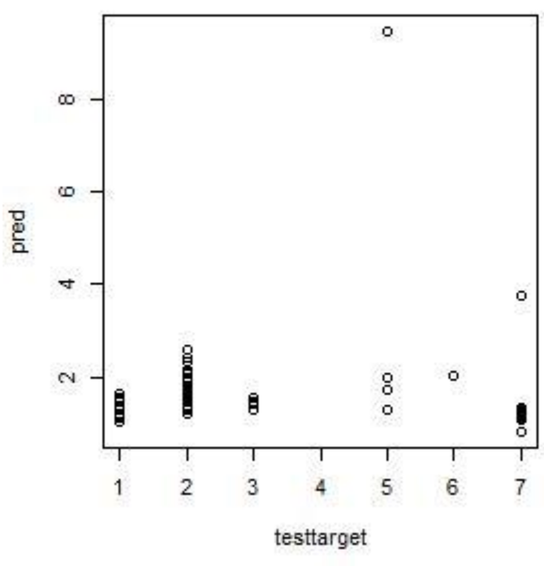


Figure 8: Dataset Graphical Representation for the Model

The dataset graphical representation in Figure 8 is showing the different performances of the various classes of glasses identification dataset. The model shows that the dataset according to their frequencies all lined up in their respective class which tended to zero. Class 4 is empty because it is not represented in this dataset. We can see that class 7 (which is headlamp) is closer to zero (0) likewise class one (which is building windows - float processed). This shows that the prediction was good but can be improved. Therefore, the hidden layers were increased to 2 with

10 neurons on the first hidden layer and 5 neurons on the second hidden layer. The result of the analysis is shown in Figure 4.9.

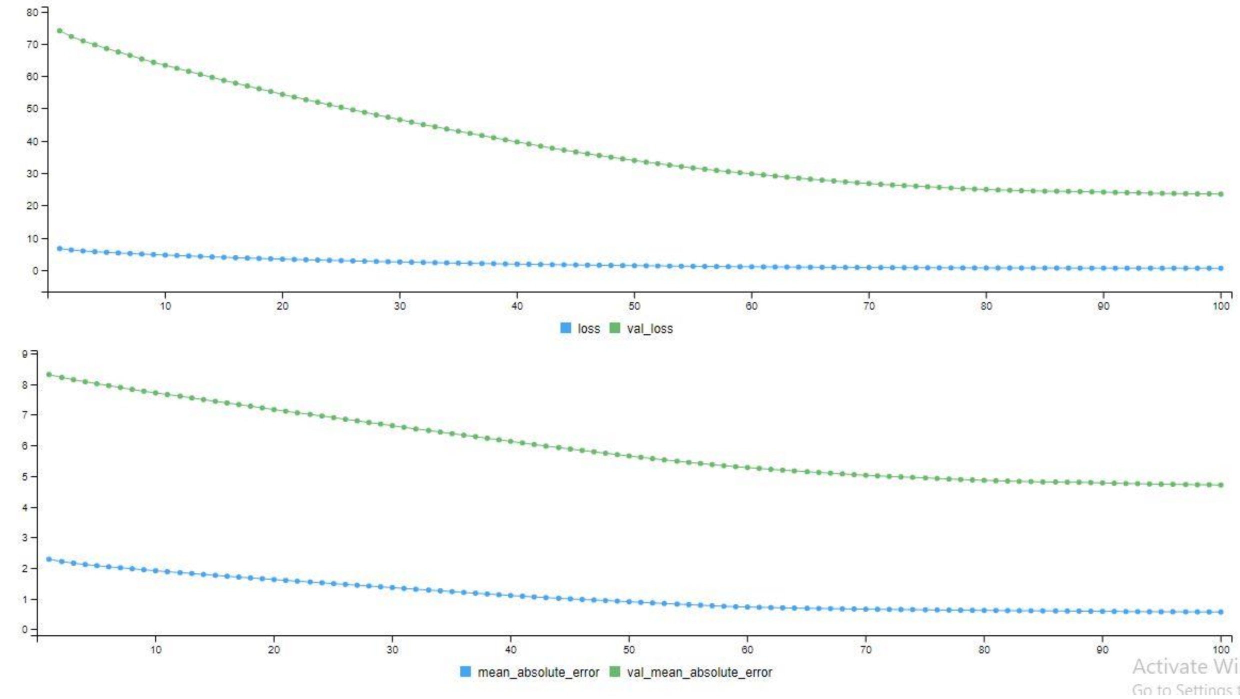
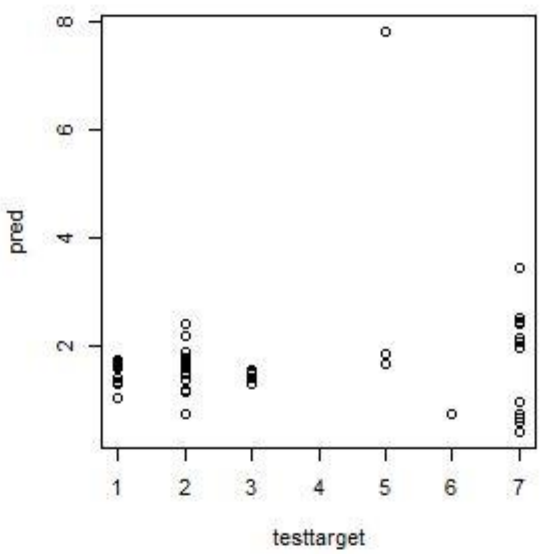


Figure 9: Validation and Mean Absolute Error graphical Representation for 2 hidden layers of 10 and 5 Neurons

Increasing the hidden layer as well as the neurons has further improved the model as we can see in Figure 9. Here, the gap between the loss and validation loss as well as the gap between the MAE and validation MAE have reduced. We can see from the top graph that the loss and val_loss dropped drastically when compared to the val_loss graph in Figure 8, likewise the mean_absolute_error and val_mean_absolute_error dropped looking at the scale. This gives us a loss of 7.771609 MSE and a mean absolute error validation of 1.881396 MAE. This drop is further reflected in the scatter plot in Figure 10.



1.

Figure 10: Validation and Mean Absolute Error Graphical Representation for 5 Neurons

We can see that when we increase the hidden layers as well as the neurons, the resultant reduction in loss and MAE was very obvious through the scatter plot in Figure 10. We can see it from the 'pred' scale and the loss reduction. But, to explore more fine-tuned options, further improvement was made on the neurons and hidden layers. The hidden layers were increased to 3 with 3 dropout layers. The dropout layers mean that during training, some percentage of the neurons are dropped to zero. The first hidden layer has 100 neurons with a 40% dropout, the second hidden layer has 50 neurons with a 30% dropout layer and the third hidden layer has 20 neurons with a 20% dropout. The result displayed in Figure 11 is the resultant effect of the major adjustment in the hidden layers and increase in neurons.

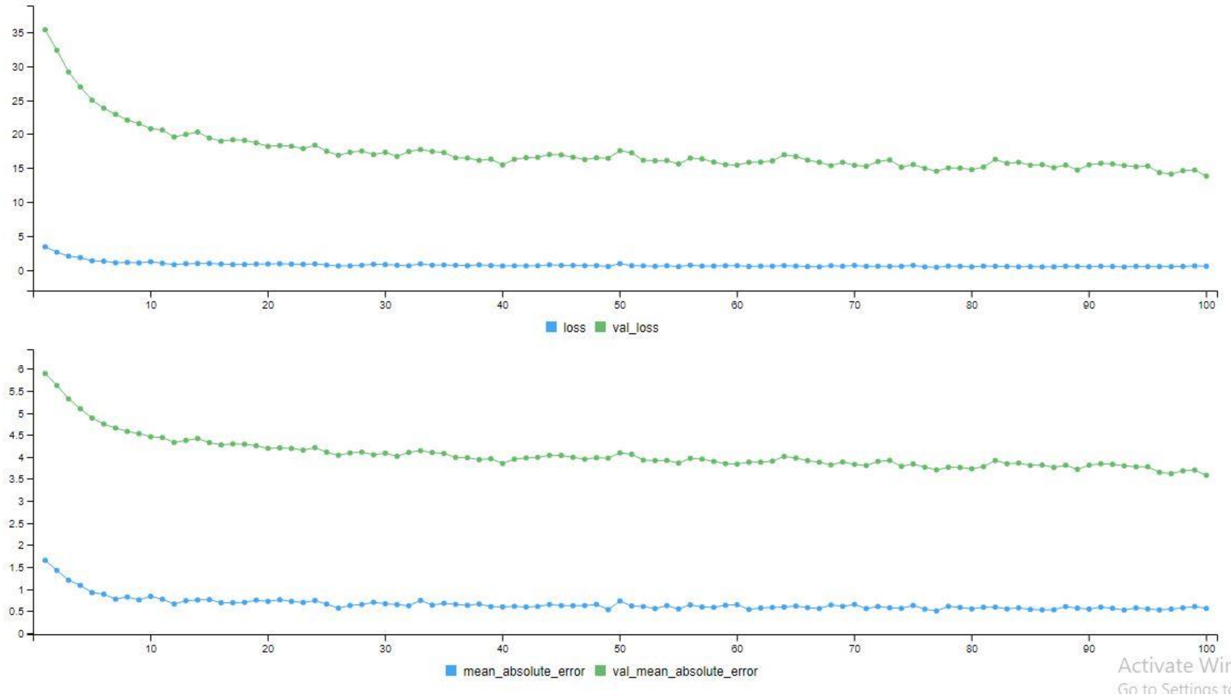


Figure 11: Validation and Mean Absolute Error Graphical Representation for 5 Neurons

From Figure 11 we can see that there was no curve fitting at all during the model implementation and the gap loss (MSE) – val_loss and the MAE – val_MAE reduced. During the training, 40%, 30% and 20% of the neurons were dropout which also resulted in a significant fluctuation. The scatter plot in Figure 12, judging from the previous plot, looked different from the previous plots and may seem like the loss increased. But, the evaluation of the model shows that, there was a significant reduction in loss and MAE as it records a 4.376815 loss and a mean absolute error of 1.412363.

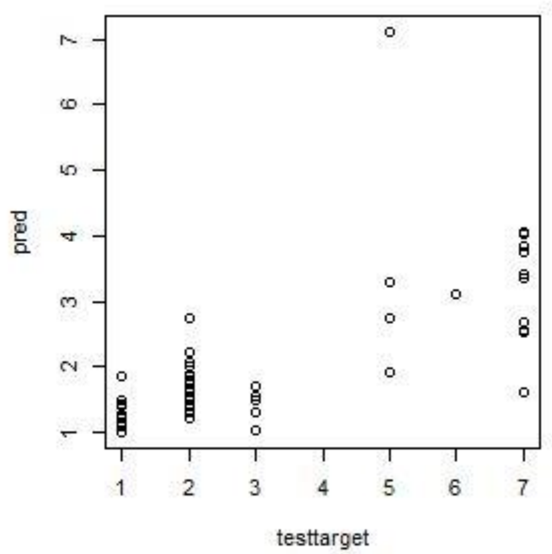


Figure 12: Validation and Mean Absolute Error Graphical Representation for 5 Neurons

Decision Tree Evaluation

The decision tree algorithm has been used also to identify and classify the glass dataset collected for evaluation and analysis. The dataset was preprocessed and partitioned for training with 80% of the data used for training the model and 20% of the data for evaluation of the model. The tree has 7 terminal nodes with 174 observations. The results of the decision tree evaluations are presented thus:

From Table 1, we can see that classes 1 and 2 carry the highest percentage of 32 and 35 percent respectively and classes 5 and 6 have the lowest frequencies with 6 and 4 percent respectively. The table explains the frequencies of each class excluding the fourth class (i.e. class 4 – vehicle window – non-float processed). This can be seen in Figure 13 below.

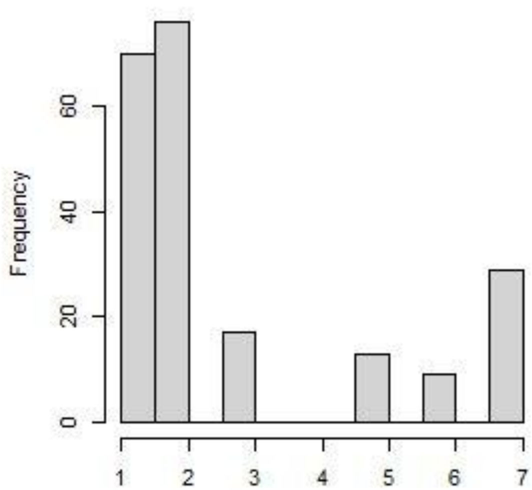


Figure 13: Frequency Distribution Table of the Dataset

Cite this article as:
Nanwin, D.N., & Ofor, W.D. (2023). Glass recognition as a cutting-edge machine learning approach for identification and classification. *FNAS Journal of Scientific Innovations, 5*(2), 1-16.

Table 2: probability distribution of each predicted class of the Test dataset

1	2	3	5	6	7
0.66666667	0.21794872	0.10256410	0.00000000	0.00000000	0.01282051
0.66666667	0.21794872	0.10256410	0.00000000	0.00000000	0.01282051
0.66666667	0.21794872	0.10256410	0.00000000	0.00000000	0.01282051
0.66666667	0.21794872	0.10256410	0.00000000	0.00000000	0.01282051
0.66666667	0.21794872	0.10256410	0.00000000	0.00000000	0.01282051
0.1282051	0.762308	0.1025641	0.00000000	0.00000000	0.00000000

From Table 2 we can see that, the first six results extracted from the list of decision tree probability analysis have identified the first five results to be building windows – float processed (1) and the sixth result to be building window – non-float processed. There are other results displayed but, we will be looking at it from Table 3.

Table 3: Misclassification Table for Train Data

	1	2	3	5	6	7
1	**52**	17	8	0	0	1
2	7	**41**	5	5	2	1
3	0	0	**0**	0	0	0
5	0	1	0	**5**	2	2
6	1	1	0	0	**4**	1
7	0	0	0	2	0	**16**

Table 3 is the misclassification analysis from the model that demonstrates how the model performed and the data that were misclassified. The results shown in the diagonal of the table represent the correct predictions as the actual dataset and the model predicted the same for each class in the dataset. Other numbers not in the diagonal are regarded as misclassifications since the actual says differently from the model. For example, from the table, we can see that the model has correctly classified class 1 to be 52% as the actual classification is also 52%. But, we can see that the model predicts 17% of data to be class 2 while the actual prediction is class 1. It also predicted 7% to be class 1 while the actual classification is to be in class 2. This led to the means square error of 0.3218391 with a performance accuracy of 0.6781609 for the trained model.

Table 4.4: Misclassification Table for Test Data

	1	2	3	5	6	7
1	**9**	6	4	0	0	0
2	1	**9**	0	1	0	1
3	0	0	**0**	0	0	0
5	0	1	0	**0**	0	0
6	0	0	0	0	**0**	0
7	0	0	0	0	1	**7**

Table 4 is the misclassification analysis from the model that demonstrates how the test or evaluation dataset performed. The results shown in the diagonal of the table represent the correct predictions as the actual dataset and the model predicted the same for each class in the dataset. We can see that from the table, the model has correctly classified class 1 to be 9% as the actual classification is also 9%. But, we can see that the model predicts 6% of data to be class 2 while the actual prediction is class 1. It also predicted 4% to be class 3 while the actual classification is to be in class 1. This led to the means square error of 0.375 with a performance accuracy of 0.625 for the test dataset.

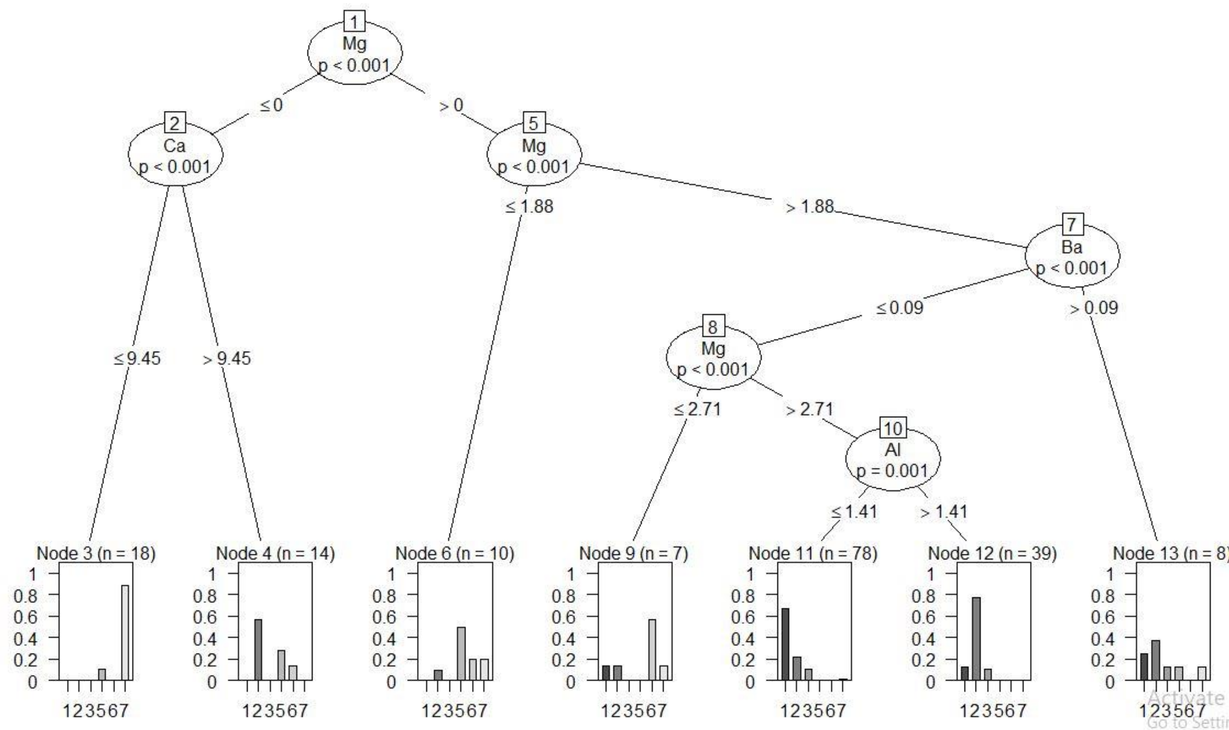


Figure 14: Decision Tree Model of the Developed System

Figure 14 is the decision tree of the trained model of glass identification and classification analysis. We can see that when the chemical constituent analysis of the glass is mg (Magnesium) and is less than or equal to zero and the Calcium (Ca) is less than or equal to 9.45 the model identified the glass to be headlamps (i.e. class 7) which has about 90% of the data and when Ca is greater than 9.45 it identified the glass to be building window – non-float processed with about 58%. When the chemical constituent of magnesium is greater than zero but less than or equal to 1.88 the model identifies and classifies them as glass containers (i.e. class 5). When Mg is greater than 1.88 and the Ba component is greater than 0.09, the model identifies and classifies them as Building window – non-float processed and so on. It is clear from Figure 14 that class 4 is nowhere in the dataset.

Cite this article as:
Nanwin, D.N., & Ofor, W.D. (2023). Glass recognition as a cutting-edge machine learning approach for identification and classification. *FNAS Journal of Scientific Innovations, 5*(2), 1-16.

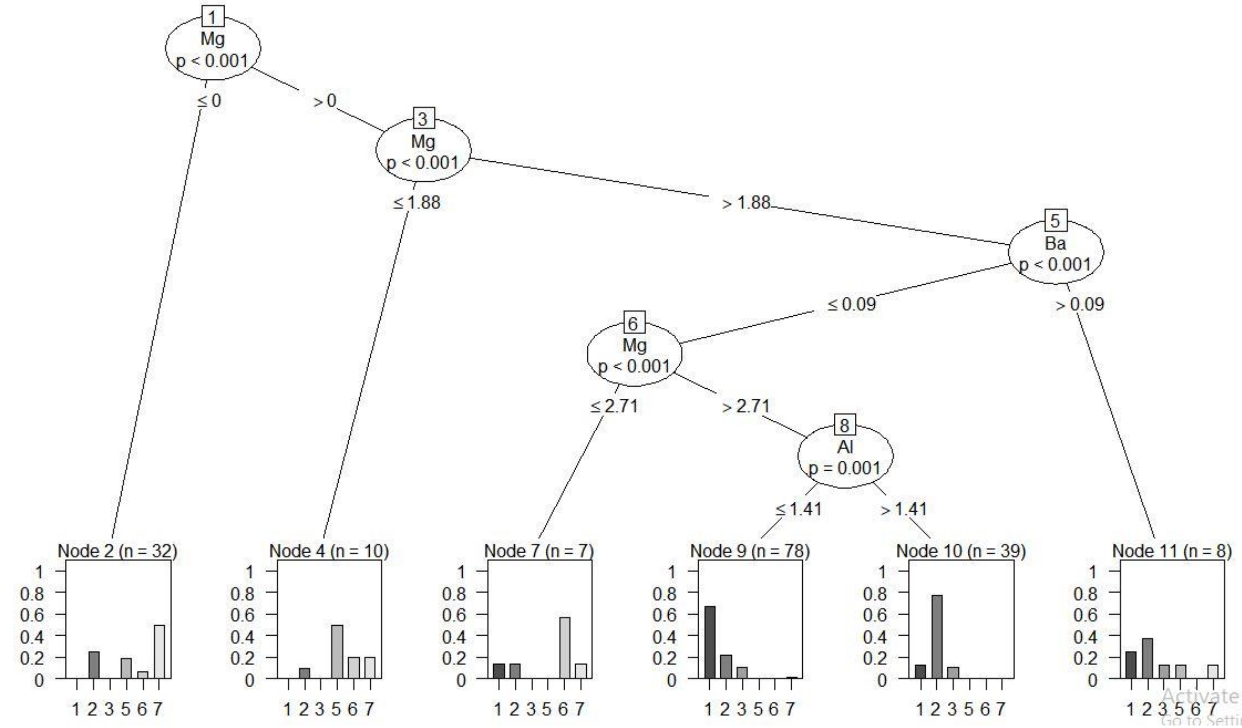


Figure 15 Pruned Decision Tree Model of the Developed System

Figure 4.15 is the decision tree of the trained model of glass identification and classification analysis. For clarity, we can see that, when the chemical constituent analysis of the glass is mg (Magnesium) and is less than zero the model identified the glass to be headlamps (i.e. class 7) which has about 50% of the data. If the chemical constituent of magnesium is greater than zero but less than or equal to 1.88 the model identifies and classifies them as glass containers (i.e. class 5). If Mg is greater than 1.88 and the Ba component is greater than 0.09, the model identifies and classifies them as Building window – non-float processed and so on. Other figures are in Appendix B for this documentation.

Discussion

The result of the deep neural network (DNN) and the decision tree algorithm has been presented. The DNN was implemented three times with different numbers of hidden layers to optimize the result of the glass identification and classification model. All the libraries needed to run the implementation were initialized and the data was loaded to the work environment of the RStudio IDE with 214 observations and 10 variables. The data was prepared and summarized for visualization. Graphical representations of the dataset were shown after which the data was transformed for the evaluation with DNN and decision tree model.

Conclusion

In conclusion, this research has successfully developed a glass identification and classification model aimed at enhancing the effectiveness of crime scene evidence classification. Employing a combination of deep neural network and decision tree algorithms, the model demonstrated commendable performance, albeit with room for improvement in accurately classifying the test dataset. The utilization of the UCI machine learning data repository allowed for the effective demonstration of the system. The DNN model training displayed a substantial reduction in loss and evaluation loss, reaching a minimum of 4.376815 and a mean absolute error of 1.412363. The trained model exhibited a mean square error of 0.3218391 with a performance accuracy of 68%, while the test dataset yielded a mean square error of 0.375 with a performance accuracy of 62.5%. The confusion matrix provided a clear overview of correctly and incorrectly classified data. Validation with the test dataset revealed a 60% performance accuracy and a validation loss of 3.45, as evident in the confusion matrix. It is important to note that the images acquired, though vandalized pipeline images, did not precisely align with the model's intended purpose. To enhance results, future efforts should focus on obtaining clear images that better align with the model's objectives. This study contributes to knowledge by introducing a robust glass identification and classification model, integrating deep neural networks and decision tree algorithms for crime scene evidence. The findings highlight advancements in forensic technology, emphasizing improved accuracy in classifying diverse glass types, crucial for investigative processes.

References

Blum, A. (2007). Machine learning theory. *Carnegie Melon Universit, School of Computer Science*, 26.

Coleman, A. S. (2002). A classification of models. *Academia*: University of Arizona, Tucson.

El-Khatib, M. J., Abu-Nasser, B. S., & Abu-Naser, S. S. (2019). Glass Classification using Artificial Neural Network. *International Journal of Academic Pedagogical Research (IJAPR), 3*(2), 25-31.

Faizal, N. & Adarsh, G. (2022). Comparative study of Glass Type Classification using Machine Learning. *YMER International Journal*, *21*(3), 111-126

Kotsiantis, S. B., Zaharakis, I. D., & Pintelas, P. E. (2006). Machine Learning: A Review of Classification and Combining Techniques. *Artificial Intelligence Review*, *26*(3), 159-190.

Kraus, M. A., & Drass, M. (2020). Artificial intelligence for Structural Glass Engineering Applications—Overview, Case Studies and Future Potentials. *Glass Structures & Engineering*, 5(3), 247-285.

Liu, H., Fu, Z., Yang, K., Xu, X., & Bauchy, M. (2019). Machine Learning for Glass Science and Engineering: A Review. *Journal of Non-Crystalline Solids*: X, 4, 100036.

Martens, D., & Baesens, B. (2010). Building Acceptable Classification Models. In Data Mining (pp. 53-74). *Springer*, Boston, MA.

Mathur, H., & Surana, A. (September, 2020). Glass Classification Based on Machine Learning Algorithms. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, *9*(11), 139-142.

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.

Reddy, R. V. K., & Babu, U. R. (2018). A Review on Classification Techniques in Machine Learning. *International Journal of Advance Research in Science and Engineering*, *7*(3), 40-47.

Soofi, A. A., & Awan, A. (2017). Classification Techniques in Machine Learning: Applications and Issues. *Journal of Basic & Applied Sciences*, *13*, 459-465.