GSJ: Volume 10, Issue 12, December 2022
ISSN 2320-9186

GSJ: Volume 10, Issue 12, December 2022, Online: ISSN 2320-9186
www.globalscientificjournal.com

WINE QUALITY PREDICTION BY USING MACHINE LEARNING ALGORITHMS

Muhammad Aslam

Superior University Lahore

phcs-f21-010@superior.edu.pk

GSJ: Volume 10, Issue 12, December 2022
ISSN 2320-9186

Abstract

Today is the era of computer technology ,and every stuff is shifted to computer technology for future use of things. These days people attempt to lead a high priced existence. They have a tendency to use the stuff either for display off or to show off to the peoples each day. Now a days the intake of crimson wine could be very common place to all. In this regard it have become essential to investigate the feature of wine prior to its intake to grip on human physical condition. Consequently this studies was a forward way to the excellent calculation of the wine the use of its varied attributes. the basic data are the dataset is taken from the resources and the methods are inclusive of support Vector system and Naïve Bayes are implemented. Frequent trial are premeditated where as the cost are as compared amongst data set and trying out set and for this reason the excellent out of the two strategies are implemented and may depends on set whose consequences is expected. Higher effects can be discovered and the best features out from different strategies are determined and merged with each other to increase the accuracy and efficiency value of these .

Keywords: Quality; Naïve Bayes; Support Vector Machine; quality, Extreme data.

1. Introduction

Machine learning (ML) is a topic of study focused on comprehending and developing "learning" methods, or methods that use data to enhance performance on a certain set of tasks. It is considered to be a component of artificial intelligence. Without being expressly taught to do so, machine learning algorithms create a model using sample data, also referred to as training data, in order to make predictions or judgments.

Artificial neural network is a replica build on a set of unified "artificial neurons," that loose models of the neurons in a human being mind. Similar to the mindset of a human brain, each connection allows in rank

GSJ: Volume 10, Issue 12, December 2022
ISSN 2320-9186

to travel from one synthetic neuron to another, or a "signal." After processing a signal, an artificial neuron can signal for other artificial neurons that are associated with each other. In conventional ANN using the production of each artificial neuron is calculated by some non-linear occupation of the summation of its input, and the signal at a link between artificial neurons is a real numeral. "Edges" are the associations between synthetic neurons. Artificial edges and neurons frequently have weights that change as learning progresses.

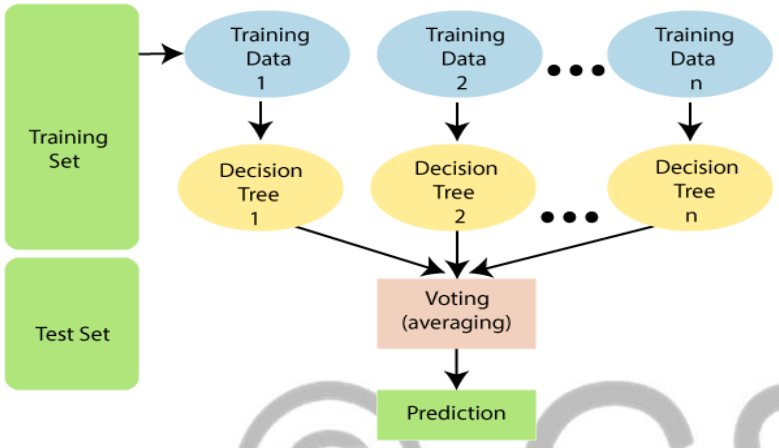


Diagram 1. Machine prediction

SVMs, sometimes referred to as support-vector networks, are a cluster of related supervised learning techniques used for regression and classification. A support vector machine preparation algorithm creates a mold that predict whether a new instance fits different catagories. collection of training examples that have each been labeled as belonging to one from different categories[7]. There are many ways to implement classifier such as linear or binary.

A naïve Bayes classifier is a technique that uses Bayes' theorem classification. Strong or naive independence between the properties of data points is an assumption made by naive Bayes classifiers. Examples of typical uses for naive Bayes classifiers include spam filters, text categorization, and medical diagnosis. These classifiers are used in machine learning because they are simple to use. Today industries are using certification to sell and purchase their products, and some clients appreciate wine to an ever-increasing extent. trying and tasting of wine is a personal experience for each of us. Many people from us considering that wine is something passive that rarely requires attention, but it still can provide joy. But the other group

GSJ: Volume 10, Issue 12, December 2022
ISSN 2320-9186

of people may choose to become active participants when they try. Ask questions about wine to get a deeper understanding and gratitude. To determine the general quality of wine, you must ask the following questions.

What does the wine look like?

What does the wine smell like?

What will be the taste of wine?

I have analyzed a variety of wine components, so I have more questions. For example, Will all pieces cooperate in presenting a perfect and pure wine? Is it simple or complicated? Can I drink now or for a long time? I wish it is not the best past. In addition, think that the price is good? Once you know the answer to all these questions, you can determine the general quality of the wine. Is it wrong, poor, acceptable, well, very good, or excellent? You must use your previous observation to maintain your conclusions.

For wine making and for the offering of lower development of industry of wine the wine business is investigating new developments for both[1]. To confirm the wine assessment physiochemical and tactile assessments are utilized . The separation of wines is not a easy process inferable from the headspace's convolution and heterogeneity. The association of wines is marvelous in soft of the truth for numerous reasons[2]. These motives are an economic judgment for the wine items, to impervious and certification the pure nature of wines, to forestall dishonesty of wines, and also to maintain the stimulant preparation. Data mining improvements have been applied to diagram wine quality. The point of machines getting to know methods for quite a several purposes is to provide a suitable model from particulars to predict wine quality[3].

Three cultivars from Italy were acquired by the UCI store in 1991 using a "Wine" data index with 178 occurrences and estimations for thirteen different synthetic components, including alcohol and magnesium. This data has been heavily used as a benchmark for new information mining classifiers. because The separation of the two is surprisingly easy for the characterization of the wine is indicated by the geological zone , the main component analysis(PCA) was used[4]. 33 Greek wines with physicochemical components were among the data they used in their analysis. A different study of wine classification used physicochemical information[5]. These numbers pertain to wine odor chromatograms that were calculated using a Fast GC Analyzer. The final analysis compares three representation strategies, including Naive Bayes, Random Forest, and Support Vector Machines

GSJ: Volume 10, Issue 12, December 2022
ISSN 2320-9186

(SVM), along with how well they are displayed in a two-organized architecture. Some have suggested using a few statistical mining frameworks to improve wine. nice evaluation [6]. A style desire framework was suggested by Cortez et al. The goal of Shanmug Anathan's approach was to predict the impact of the local climate and season on wine production and wine fines[7]. The Wine informatics structure, as demonstrated by Chen et al., represented the flavour and characteristics of wine using standard language standards. They employed cutting-edge clustering techniques and affiliation we regulations[8]. In the search article, the authors compared different computer learning algorithms on data from cardiotocography, including Naive Bayes, Decision Trees, and Support Vector Machines, to see whether they are good algorithms[9].

In the previous research, it has been made to use special machine mastering procedures and characteristic decision methods for the wine data. Er and Atasoy projected an approach to categorize the paramount wines with the usage of three distinctive classifiers like guide vector machines, accidental wooded area, and k-nearest neighborhood. more over they have been used main component analysis for characteristic resolution and they found proper results with the use of Random forest algorithm [10]. An technique that uses human flavour reviews and wine grade prediction was put out by Chen et al. They analysed the reviews and predicted the wine grade using the hierarchical cluster method and the organization rule algorithm, and they found an accuracy of 85.25%[4]. Appalasamy et al. suggested a method to forecast wine flavour using data from physiochemical tests. They have emphasized how the categorization approach enhances the first class of wine at some point in the production process[12].

To suggest the product, Reddy and Govindara julu employed a user-centric clustering technique. For the aim of the survey, they employed a set of statistics about purple wine. Based only on the literature assessment, they assigned relative vote casting to the qualities. The Gaussian Distribution Process was then used to weigh the qualities. Based on the consumer want group, they evaluated the first class[11]. The past work motivated us to attempt specific characteristic determination algorithms as well as exceptional classifiers to evaluate the overall presentation of metrics.

The knowledge above is intended to give you a fundamental foundation for accurately judging wines. Your ability to taste wines will improve and be strengthened by using as many different wines as you can when practicing these stages. To improve on this basis, tasters will benefit from

GSJ: Volume 10, Issue 12, December 2022
ISSN 2320-9186

enrolling in classes provided by an organization like the Wine & Spirits Education Trust (WSET) or the court of master vintners..

2. Literature Review

Wine quality prediction is the task of predicting the quality of wine on a scale from 1 (very poor) to 10 (excellent). This can be done using machine learning, which is a type of artificial intelligence that can learn from data and make predictions. There are many different types of machine learning algorithms, but they all share a common goal[1].

Some machine learning algorithms are better at finding patterns than others, and some are better at making predictions than others[13]. The best algorithm for wine quality prediction will depend on the data that is available. The data used to train the machine learning algorithm can be divided into two types: features and labels. Features are the characteristics of the wine that will be used to make predictions, such as the type of grape, the region where the wine is from, and the year the wine was made.

Labels are the wine quality ratings that will be predicted, such as 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10[2].

To train the machine learning algorithm, we need to have a dataset of wine quality ratings and the corresponding features. This dataset can be created manually[15].

A machine learning algorithm is a set of instructions that a computer program uses to improve its performance on a given task. There are many different types of machine learning algorithms, each designed for a specific purpose. Some of the most common machine learning algorithms include:

Supervised learning algorithms: These algorithms are used to learn from labeled training data. The most common supervised learning algorithms are regression and classification algorithms.

Unsupervised learning algorithms: These algorithms are used to learn from unlabeled data. The most common unsupervised learning algorithms are clustering algorithms.

GSJ: Volume 10, Issue 12, December 2022
ISSN 2320-9186

Reinforcement learning algorithms: These algorithms are used to learn from a reinforcement signal, such as a reward or punishment. The most common reinforcement learning algorithms are Q-learning and SARSA.

Data mining is a method for finding fresh instances that can be used to distinguish high-quality data from massive storehouses of information. It includes a variety of metrics, machine learning, and database organization. The main goal is to separate massive records from massive databases and then transform the crucial information into something that may be used in further study. Data mining is typically included in Knowledge Discovery in Databases (KDD) as a crucial investigative stage. In addition to analysis, it also includes complex considerations, analysis of large data sets, pre- and post-evaluation of the data, and finally, the discovery of new facts and subsequent updating. Information analysis frequently just considers the hypotheses and models based on the in sequence, giving little thought to the actual content of the in order[23].

Factual analysis and machine learning are combined in data mining. The practise of identifying patterns and information from huge datasets is known as "data mining." Data mining is an automated method for finding patterns in vast amounts of data, spotting anomalies, and ultimately figuring out what you want to happen. Several statistics mining strategies are integrated with their best traits for improved outcomes that produce accurate impacts with fewer errors and improved efficiency. The method of constructing new hypotheses based on a larger body of data is known by a number of names, including statistics fishing records dredging and data peeping. The wine manufacturing companies are trying to improve advancements in both winemaking and offering structures to lower back up this development[1]. Wine confirmation is evaluated using physicochemical and tactile tests. The complexity and heterogeneity of its atmosphere indicate that the segmentation of wines is not a straightforward process. The association of wines is tremendous in mild of the truth for several reasons[2].

Data mining improvements have been applied to diagram wine quality. Creating models out of data to predict wine quality is the goal of machines learning approaches for a variety of purposes[3].

Using a "Wine" informative index that includes 178 occurrences and estimations for thirteen different synthetic chemicals, including alcohol and magnesium, the UCI shop ordered three different types of wine from Italy in 1991. Since it is remarkably easy to separate, this data has been widely used as a benchmark for new statistics mining classifiers. [4].

There are thirty three Greek wines with physicochemical components were among the data they used in their analysis. The physicochemical information was utilized in another wine grouping study[5]. These statistics are based on Fast GC Analyzer estimates of wine odor chromatograms.

GSJ: Volume 10, Issue 12, December 2022
ISSN 2320-9186

A few frameworks for using statistics mining to evaluate wine have been suggested[6]. A style desire framework was suggested by Cortez et al. A "Support Vector Machine, Naive Bayes, and Random Forest " has been used to engineer the assessment of wines in their taste expectation framework. The goal of Shanmug Anathan's technique was to predict how the local climate and the current season will affect wine yields and fines [7]. The Wine informatics framework, as established by Chen et al., represented the flavor and characteristics of wine using standard language standards. They employed revolutionary clustering and affiliation we regulations[8]. The authors of the lookup paper compared various computer learning methods, including Naive Bayes, Decision Trees, and Support Vector Machines, to anticipate they[9].

Recently, attempts have been made to apply various machine learning algorithms and characteristic selection methods to the wine dataset. Three distinct classifiers, including assist vector machines, random forest, and k-nearest neighbourhood, were offered by Er and Atasoy as a method to categorise the quality of wines[10].

They chose features using prominent factor analysis and discovered that the Random Forest algorithm produced the desired results. An technique that uses human flavour reviews to predict wine grade was put forth by Chen et al. To method the evaluations and estimate the wine grade, they employed the hierarchical clustering technique and association rule algorithm, and they discovered an accuracy of 85.25%. A method to estimate wine quality based solely on the results of physiochemical tests was proposed by Appalasamy et al. They have emphasised how the classification technique enhances the quality of wine throughout production[11][12].

An approach to categorise wines based on their aroma chromatograms was proposed by Beltrán et al. They used PCA to reduce dimensionality, wavelet transform to extract characteristics, and classifiers like neural networks, linear discriminant analysis, and assist vector computers. They found that the performance of the guide vector machine with wavelet transforms was superior to that of other classifiers. [13][14][15]. Thakkar et al. employed the analytical hierarchy process (ahp) to rank the attributes before applying random forest and guide vector computers to find accuracy of 70.33% and 66.54%, respectively, in machine learning classifiers. Reddy and Govindarajulu promoted the product using a user-centric clustering method. For the purpose of the survey, they have used the Crimson Wine statistics collection. Based only on the literature evaluation, they assigned relative balloting to the qualities. The qualities were then given weights using the Gaussian Distribution Process[17]. According to the individual desire team, they evaluated the quality. Due to its complex genetic makeup, Pinot Noir is prone to point mutations that, even on the same plant, can produce various clones of the grape. There have been identified 40 different Pinot noir clones in all. 15 of them are known for producing grapes of a better calibre. Clone

GSJ: Volume 10, Issue 12, December 2022
ISSN 2320-9186

selection is influenced by a variety of factors, including soil composition, temperature, and the winegrower's goals. It is not unusual to find one or more vines with a single branch on the same plant that have different characteristics in Pinot noir vineyards[20]. If the newly discovered clone's entire crop exhibits the same traits as the original shoot following mutation, it may be considered a new variety of Pinot noir. Grape varieties including Pinot Gris, Pinot Franc, and Meunier are the result of Pinot noir. Fruit colour, fruit flavour, and wine aroma variations are all discernible[14]. Here is a detailed description of the machine learning algorithms research that has already been done.

Author name	**Year of publication**	**Journal name**	**Summary of work**
Sunny Kumar et all	Jan 2020 INDIA	Conference paper	Dataset is obtained from the sources, and methods like Nave Bayes, Support Vector Machine, and Random Forest are used. The best of the three techniques is projected based on the outcomes of the training set after various metrics have been calculated and the results compared between the training set and testing set. The best qualities from various techniques can be extracted and combined to produce better outcomes that are more accurate and effective.
Satyabrata Aich et all	May 2018	ICACT Transactions on Advanced Communications Technology	For this study project, we used quality datasets for red and white wines. To test the performance of the prediction, we explored a variety of feature selection techniques, including simulated annealing (SA) and genetic algorithm (GA) based feature selection. We have employed probabilistic, linear, and nonlinear classifiers. We have discovered that feature selection-based feature sets are more accurate in predicting performance than feature-based feature sets that take into account all features.
Yogesh gupta et all	December 2017	Science direct	This paper explores the machine learning technique such as linear regression support vector machine and neural network for product quality in two ways. firstly, determined the dependant variable and second the value of predicting variable. This paper proves that the selected variables shows more accuracy rather than the whole variables.
Joanna E. Jones et all	August 2014	American journal of Analogy	Studies on viticulture management, especially those which can modify group temperature and experience to occurrence of light, are likely to best educate production methods that lead to

GSJ: Volume 10, Issue 12, December 2022
ISSN 2320-9186

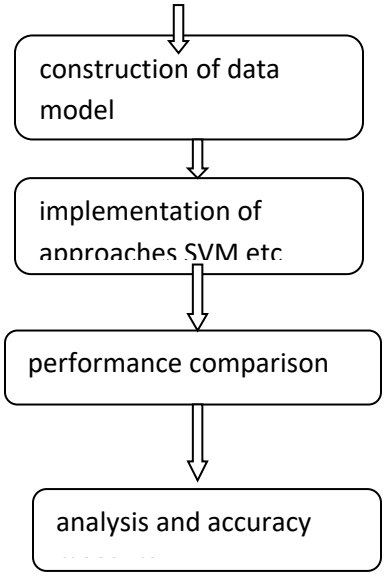
			fruit quality 26 appropriate for the creation of iconic luminous wines.
Yasem Er et all	September 2016	www.ijisae.org	In this study the author used two different data sets for the quality of wine . He used the basic three algorithms of machine learning as known as K nearest neighbor ,support vector and Random forest .Random forest algorithm was found as good with the comparison of others.
Ayten Atasoy	September 2016	IJISAE	In this study the author used two different data sets for the quality of wine . He used the basic three algorithms of machine learning as known as K nearest neighbor ,support vector and Random forest .Random forest algorithm was found as good with the comparison of others.
S.Kallithraka	November 2ooo	ELSEVIER	In this research various instrumental and sensory method are used in conjunction with statistical analysis. This research classify the quality of wine product.
A.Mustapha	2012	Asian network for science information	In this research the author used two approaches Naïve Bayes and decision tree algorithms used and their performance is measured and compared.

3. Methodology

Data is collected from UCI ML repository. Data have 1599 occurrence with 12 variables. The input is taken and conclude with the red wine quality. The quality of this data set is predicted between 3-8. 3 predicts that the quality of red wine is low and 8 predicts a red wine of excellent quality. The most prominent aspects include fixed acid properties, citric acid, volatile acidity, residual sugar, chloride, thickness, sulfur dioxide, sulfur dioxide, pH, alcohol and sulfate. Consumables have a pH scale between 3 and 4. The amount of salt represents the wine chloride content. The objective of the information files is to predict the corresponding evaluation of the wine test teacher. For example, use the scope of physical chemical characteristics, such as acidity and Sake characteristics. As a result of safety and strategic problems, simply the use of physical chemical products (input) and the output coefficient are available. "In the field of automatic learning, the confusion matrix is a frequently used table to describe the presentation of group models related to many test information known for its true quality. It allows the recognition of the presentation of the calculation. In this study, we basically use the red wine data set, then calculate the confusion matrix, related performance measurements and finally compare different automatic learning algorithms based on the precision provided for in this data set".

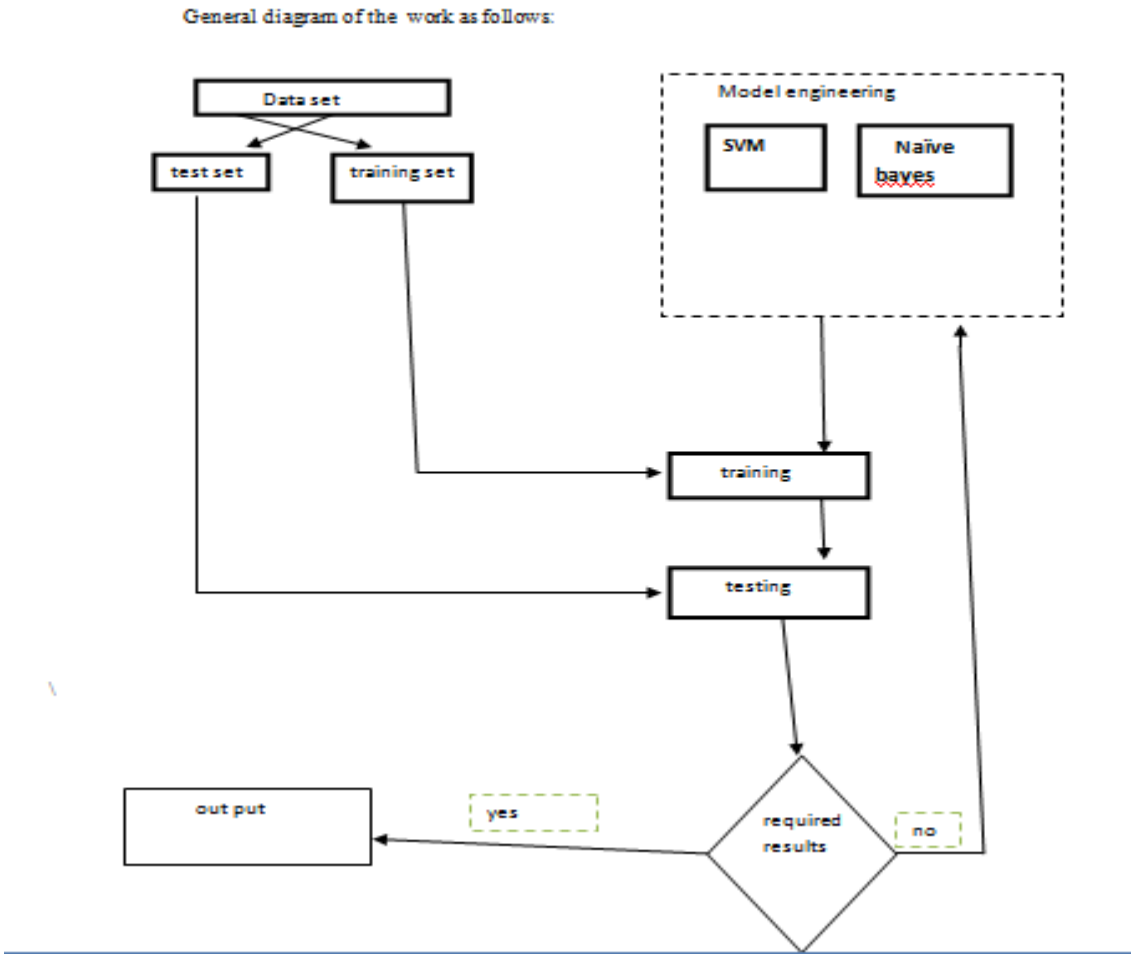
collection of data

GSJ: Volume 10, Issue 12, December 2022
ISSN 2320-9186



General Flow chart diagram

GSJ: Volume 10, Issue 12, December 2022
ISSN 2320-9186



4. Results

Today, people consumed red wine as needs or to show off. This allows health loss. Therefore, it is crucial to assess the quality of red wine before consuming it in order to maintain human health. As a result, data on red wine that was taken from the database used to forecast wine quality is included in this survey, data sets for the study programme carries out several algorithms for automated learning. In a certain data set, the precision is computed. During use, the data is divided into a training set with a probability of 0.7 and a set of testing with a probability of 0.3,

GSJ: Volume 10, Issue 12, December 2022
ISSN 2320-9186

respectively. As a result, using the Bayes ship method, the precision attained in the training set and test set is 55.91% and 55.89%, respectively, while using the random approach, it is 67.25% and 68.64% in each case. 65.83% and 65.46%, respectively, are both. It is demonstrated that the accuracy of the test training set can provide more precision than the random forest algorithm and the most recent Bayes ship algorithm because to the high likelihood of the division of the training set. When research is done to develop approaches utilizing the three algorithms, the outcomes can also be altered. Results are more effectively obtained when the SVM algorithm's Hyper plane modifications are done properly, a precise balanced tree is used, and the right probability is used.

	Training data					
	Wine quality3	**Wine quality 4**	**Wine quality 5**	**Wine quality 6**	**Wine quality 7**	**Wine quality 8**
precession	0	0.333333	0.9899	0.9830508	0.9367	0
Recall	0	0.25	0.9999	1	1	0
Specificity	0	0.9981	0.9999	0.9955	0.65886	0.9244
f-measure	0	0.2857	0.9881	0.991453	0	0
Accuracy(%)	0.6583588					
Error(%)	0.3416149					

Table. 2 training data set

	Testing data					
	Wine quality 3	**Wine quality 4**	**Wine quality 5**	**Wine quality 6**	**Wine quality 7**	**Wine quality 8**
precession	0	0	0.9799	0.9830508	0.769367	0
Recall	0	0	1	1	1	0

GSJ: Volume 10, Issue 12, December 2022
ISSN 2320-9186

Specificity	0	1	0.9599	0.7855	0.9765886	0.98446
f-measure	0	0	0.9981	0.991453	0	0
Accuracy(%)	0.6864407					
Error(%)	0.31355					

Table. 3 Performance measures testing set using Naïve Bayes .

I. Matrices for Performance measures

```
> matrix

p1     3    4    5    6    7    8
  3    6    0    0    0    0    0
  4    0   38    9    5    1    0
  5    0    0  476    0    2    0
  6    0    0    5  445    0    0
  7    0    0    0    0  140    0
  8    0    0    0    0    0   14
```

Diagram 2. Error matrix for training set using Naïve Bayes

```
> matrix

p1     3    4    5    6    7    8
  3    4    0    0    0    0    0
  4    0   14    1    1    0    0
  5    0    1  185    0    0    0
  6    0    0    0  183    1    0
  7    0    0    4    2   55    0
  8    0    0    1    2    0    4
```

Diagram 3. Error matrix using Naïve Bayes.

GSJ: Volume 10, Issue 12, December 2022
ISSN 2320-9186

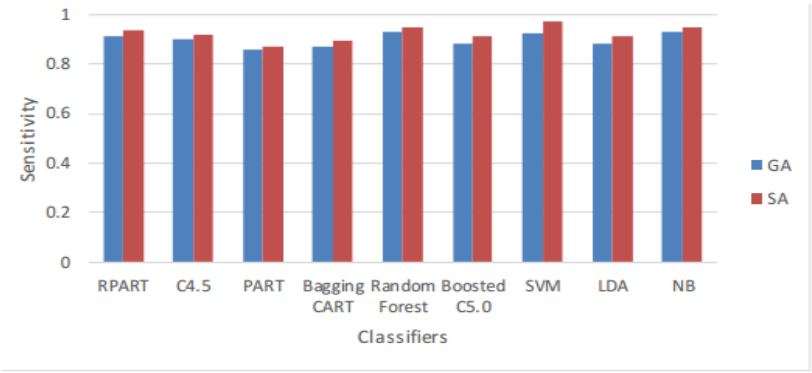


Diagram 4. displays the classifiers' sensitivity plots for the two separate feature sets.

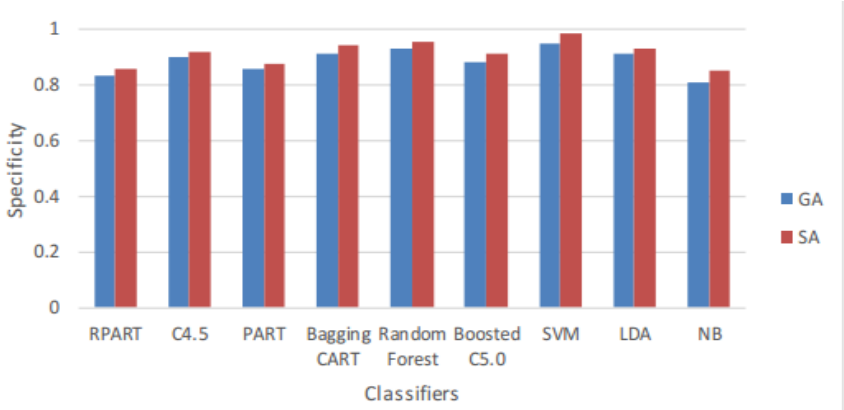


Diagram 5. SVM classifier maximum specificity among all the classifiers.

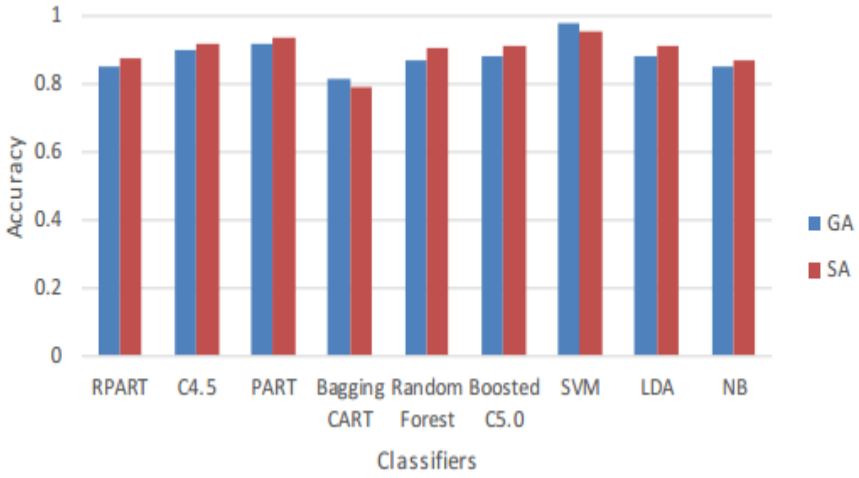


Diagram 6. Comparison of Accuracy

GSJ: Volume 10, Issue 12, December 2022
ISSN 2320-9186

Diagram7. Comparison of Accuracy

5. Conclusion and future work

Data mining in recent times is most vital technique which Is applied for investigation of the information. It seems on the Statistics and produces the specified yield. With the headway. In the innovation it facilitates in gambling the sound take a look at within the marketplace For this reason blessings the purchaser. Accuracy, precision, typographical errors, F-score, consideration, and specificity are addressed in this investigation. Since almost 70% of the records from the unique dataset are included in the training dataset, the Support Vector Machine is shown to have the best accuracy (67.25%) when used to predict the quality of purple wine using the RStudio software, followed by Random Woodland. Finally, the Naive Bayes Set of rules come in with an accuracy of 55.91% and a precision of 65.88%. Higher algorithms that combine the best elements of all available data mining approaches may emerge in the future. Much better accuracy can be found if specific adjustments to the hyper plane, a balanced tree approach, and the appropriate amount of chance are applied.

GSJ: Volume 10, Issue 12, December 2022
ISSN 2320-9186

References

[1] P. Cortez, A. Cerderia, F. Almeida, T. Matos, and J. Reis, "Modelling wine preferences by data mining from physicochemical properties," *In Decision Support Systems, Elsevier*, 47 (4): 547-553. ISSN: 0167-9236.
[2] S. Ebeler, "Linking Flavour Chemistry to Sensory Analysis of Wine," in *Flavor Chemistry,* Thirty Years of Progress, Kluwer Academic Publishers, 1999, pp. 409-422.
[3] V. Preedy, and M. L. R. Mendez, "Wine Applications with Electronic Noses," in *Electronic Noses and Tongues in Food Science,* Cambridge, MA, USA: Academic Press, 2016, pp. 137-151.
[4] A. Asuncion, and D. Newman (2007), UCI Machine Learning Repository, University of California, Irvine, [Online]. Available: http://www.ics.uci.edu/~mlearn/MLRepository.html
[5] S. Kallithraka, IS. Arvanitoyannis, P. Kefalas, A. El-Zajouli, E. Soufleros, and E. Psarra, "Instrumental and sensory analysis of Greek wines; implementation of principal component analysis (PCA) for classification according to geographical origin," Food Chemistry, 73(4): 501-514, 2001.
[6] N. H. Beltran, M. A. Duarte- MErmound, V. A. S. Vicencio, S. A. Salah, and M. A. Bustos, "Chilean wine classification using volatile organic compounds data obtained with a fast GC analyzer," Instrum. Measurement, IEEE Trans., 57: 2421-2436, 2008.
[7] S. Shanmuganathan, P. Sallis, and A. Narayanan, "Data mining techniques for modelling seasonal climate effects on grapevine yield and wine quality," IEEE International Conference on Computational Intelligence Communication Systems and Networks, pp. 82-89, July 2010.
[8] B. Chen, C. Rhodes, A. Crawford, and L. Hambuchen, Wineinformatics: applying data mining on wine sensory reviews processed by the computational wine wheel," IEEE International Conference on DataMining Workshop, pp. 142-149, Dec. 2014.
[9] K. Agrawal and H. Mohan, "Cardiotocography Analysis for Fetal State Classification Using Machine Learning Algorithms," 2019 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, Tamil Nadu, India, 2019, pp. 1-6.
[10] K. Agrawal and H. Mohan, "Text Analysis: Techniques, Applications and Challenges," presented in 2019 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, Tamil Nadu, India, 2019.
[11]UCI Machine Learning Repository, Wine quality data set, [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Wine+Quality.
[12] J. Han, M. Kamber, and J. Pei, "Classification: Advanced Methods," in Data Mining Concepts and Techniques, 3rd ed., Waltham, MA, USA Morgan Kaufmann, 2012, pp. 393-443.
[13]. Er Y (2016) The classification of white wine and red wine according to their physicochemical qualities. Int J Intell Syst Appl Eng 4(1):23–26
[14]. Executive Summary, Wine Process Monitoring, Wine Quality, Wine Safety, and Wine Complexity (2016) Wine analysis :from 'Grape to Glass' an analytical testing digest of the wine manufacturing process
[15]. Palmer J, Chen B (2018) Wine informatics : regression on the grade and price of wines through their sensory attributes
[16]. Tajini B, Paris OC (2017) BadrTajini—On campus Paris—DSTI 2017 47(4):547–553
[17]. Ghosh A (2018) Project report : red wine quality analysis final 3. An empirical red wine quality analysis of the Portuguese 'Vinho Verde' wine (2017, 2018)

GSJ: Volume 10, Issue 12, December 2022
ISSN 2320-9186

[18]. A. Co “Final report,” Apr (2001)
[19]. Gupta Y (2018) Selection of important features and predicting wine quality using machinelearning techniques. Procedia Comput Sci 125:305–312
[20]. Predictive Model, Linear Regression, and RapidMiner Studio (2018) Building and evaluating a predictive model w/ linear regression in RapidMiner studio
[21]. Cortez P (2010) Wine quality dataset. https://archive.ics.uci.edu/ml/datasets/Wine+Quality
[22]. S. Aeberhard, D. Coomans and O. de Vel, Comparison of Classifiers in High Dimensional Settings, Tech. Rep. no. 92-02, (1992), Dept. of Computer Science and Dept. of Mathematics and Statistics, James Cook University of North Queensland.
[23]. S. Aeberhard, D. Coomans and O. de Vel, "THE CLASSIFICATION PERFORMANCE OF RDA" Tech. Rep. no. 92-01, (1992), Dept. of Computer Science and Dept. of Mathematics and Statistics, James Cook University of North Queensland.
[24] Ebeler S. (1999) “Flavor Chemistry — Thirty Years of Progress: chapter Linking flavour chemistry to sensory analysis of wine”. *Kluwer Academic Publishers*, 409–422.
[25] Legin, Rudnitskaya, Luvova, Vlasov, Natale and D'Amico. (2003) “Evaluation of Italian wine by the electronic tongue: recognition, quantitative analysis and correlation with human sensory perception”. *Analytica Chimica Acta* **484 (1):** 33–34.
[26] Sun, Danzer and Thiel. (1997) “Classification of wine samples by means of artificial neural networksn and discrimination analytical methods”. *Fresenius Journal of Analytical Chemistry* **359 (2)** 143–149.
[27] Vlassides, Ferrier and Block. (2001) “Using historical data for bioprocess optimization: modeling wine characteristics using artificial neural networks and archived process information”. *Biotechnology and Bioengineering* **73 (1)** 55-68.
[28] Moreno, Gonzalez-Weller, Gutierrez, Marino, Camean, Gonzalez and Hardisson. (2007) “Differentiation of two Canary DO red wines according to their metal content from inductively coupled plasma optical emission spectrometry and graphite furnace atomic absorption spectrometry by using Probabilistic Neural Networks”. *Talanta* **72** 263–268.
[29] Yu, Lin, Xu, Ying, Li and Pan. (2008) “Prediction of Enological Parameters and Discrimination of Rice Wine Age Using Least-Squares Support Vector Machines and Near Infrared Spectroscopy”. *Agricultural and Food Chemistry* **56** 307–313.
[30] Beltran, Duarte-Mermoud, Soto Vicencio, Salah and Bustos. (2008) “Chilean Wine Classification Using Volatile Organic Compounds Data Obtained With a Fast GC Analyzer”. *IEEE Transactions on Instrumentation and Measurement* **57** 2421-2436.
[31] Cortez, Cerdeira, Almeida, Matos and Reis. (2009) “Modeling wine preferences by data mining from physicochemical properties”. *Decision Support Systems* **47** 547-553.
[32] Jambhulkar and Baporikar. (2015) “Review on Prediction of Heart Disease Using Data Mining Technique with Wireless Sensor Network”. *International Journal of Computer Science and Applications* **8 (1)** 55-59.
[33] Zaveri, and Joshi. (2017) “Comparative Study of Data Analysis Techniques in the domain of medicative care for Disease Predication”. *International Journal of Advanced Research in Computer Science* **8 (3)** 564-566.
[34] Portuguese Wine - Vinho Verde. Comissa˜o de Viticultura da Regia˜o dos Vinhos Verdes (CVRVV), http://www.vinhoverde.pt, July 2008.

GSJ: Volume 10, Issue 12, December 2022
ISSN 2320-9186