UNDERSTANDING 21ST CENTURY BORDEAUX WINES FROM WINE REVIEWS THROUGH NATURAL LANGUAGE PROCESSING AND CLASSIFICATIONS

By

Zeqing Dong

A thesis presented to the Department of Computer Science
and the Graduate School of the University of Central Arkansas
in partial fulfillment of the requirements for the degree of

Master of Science
in
Computer Science

Conway, Arkansas
May 2020

TO THE OFFICE OF GRADUATE STUDIES:

The members of the Committee approve the thesis of

_____ presented on

_____.

_____
Committee Chairperson

_____
Committee Member

_____
Committee Member

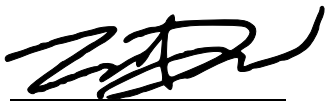
PERMISSION

Title Understanding 21st Century Bordeaux Wines from Wine Reviews through Natural Language Processing and Classification

Department Computer Science

Degree Master of Science

In presenting this thesis/dissertation in partial fulfillment of the requirements for a graduate degree from the University of Central Arkansas, I agree that the Library of this University shall make it freely available for inspections. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my thesis/dissertation work, or, in the professor's absence, by the Chair of the Department or the Dean of the Graduate School. It is understood that due recognition shall be given to me and to the University of Central Arkansas in any scholarly use which may be made of any material in my thesis/dissertation.



Zeqing Dong

April 03, 2020

ABSTRACT

Wine has been popular with the public for centuries; in the market, there are a variety of wines to choose from. Among all, Bordeaux, France, is considered as the most famous wine region in the world. In this research, we try to understand Bordeaux wines made in the 21st century through Wineinformatics study. We developed and studied two datasets: the first dataset is all the Bordeaux wine in 21st century (from 2000 to 2016); and the second one is all wines listed in a famous collection of Bordeaux wines, 1855 Bordeaux Wine Official Classification, in 21st century (from 2000 to 2016). A total of 14,349 wine reviews are collected in the first dataset, and 1359 wine reviews in the second dataset. In order to understand the relation between wine quality and characteristics, Naïve Bayes classifier, a white box classification classifier, is applied to predict the qualities (90+/89−) of wines; Support Vector Machine (SVM), a black box classification classifier, is also applied as a comparison. In the first dataset, SVM classifier achieves the best accuracy of 86.97%; in the second dataset, Naïve Bayes classifier achieves the best accuracy of 84.62%. Precision, recall, and f-score are also used as our measures to describe the performance of our models. Meaningful characteristics associate with high quality 21st century Bordeaux wines are able to be presented through this research. Also, a novel voting system based on advanced NLP is designed in the research to further improve the model performance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

1.1 Data Science

Science has its origins in attempts to understand the world in an empirically verified manner. To understand the world, one relies on testing it. Ancient natural philosophers, such as Democritus, stopped incurring the supernatural to understand natural phenomena. They instead posited material answers to understand what otherwise seemed unexplainable. This is not necessarily science though, most of what can be considered "science" are not just looking for material causes to understand phenomena, but also, an attempt to test the ideas that you have about the world.

In order to test ideas, you often need data. Each data point is a discrete record or a property of events that occurred in the world. Data is empirical, which allows for falsifiability. You have an idea about the world, but data will allow you to know to what degree your idea about the world is true or false. The paradigm shift is that the rise of computers along with Moore's law being held relatively constant has allowed humans to store and utilize a lot more data. Now, with the rise of the internet, data has become abundant. What has also allowed for the rise of data science is the advance in the application of statistically based algorithms to make sense of these larger quantities of data. These algorithms have the ability to "learn" about the data on their own, to find patterns and to predict phenomena. This allows for falsifiability. Given X data, your algorithms can or cannot predict certain phenomena at varying degrees. What is potentially interesting to non-scientists, is that this sort of method can be applied in business contexts as well.

Within this emerging field there are several different types of learning algorithms that provide utility. Depending on the type of problem they are solving, the type of input and output data, and its method, there are two main types of learning algorithms: **supervised learning [1]** and **unsupervised learning [2]**.

Supervised learning has a feature that one is trying to predict with. If one is trying to predict house prices, then one would attempt to get data on say square footage or the number of rooms. The house prices are the feature that one is attempting to predict using all the other categories of data. If the house prices is a continues feature. Then the task we are doing is **Regression [3]**. The advantage to the utilization of such algorithms is that one can discover which features account for the variance of the outcome feature one is attempting to predict. If both features are solid predictors, one could see which beta coefficient accounts for a greater amount of variance in house prices, and from there could deduce causality. It is important to note that the outcome feature does not necessarily need to be necessarily a continuous variable. It also could be a binary or categorical variable. In house price example, we can categorize the price into three groups, “cheap”, “moderate”, or “expensive.” Then try to predict which price category the house will fall into. This task is known as **Classification [4]**. Classification is a question of determining which of a set of categories a new instance belongs to base on a training dataset containing instances of known member categories. Both Regression and Classification are considered as supervised learning.

Unlike supervised learning, **unsupervised learning** has no feature/label that one is trying to predict with during the data mining process. The algorithm does not look for the features that are highly correlated to the outcome feature since there is no outcome

feature in the unsupervised learning. Unsupervised learning is to group data into categories based on similarity or distance. It can be the goal of discovering hidden patterns in the data, or it can be a method of extracting useful features.

After science, data, and learning algorithms are discussed, the missing part of data science is domain knowledge. Domain knowledge is the key concept of data science, which directly affects the quality of the data product [5,6]. We can build models on any dataset, however, telling the story of the dataset requires domain knowledge. Data science is not all about how good your model is, it is also about how good you can tell the novel information of your dataset and make sense to non-scientists. Based on different domain knowledge, data science has been generally applied to issues in society like manufacturing engineering, financial banking, fraud detection, bioinformatics, etc. [7]. It is the analysis of large observational datasets to find an unsuspected relationship and to summarize the data in novel ways that are understandable and useful to the data owner.

1.2 Wineinformatics

The ancient beverage, wine, has remained popular in modern times. While the ancients had mostly wine available from neighboring vineyards, the number and variety of wines available for purchase have exploded in modern times. According to OIV, International Organization of Wine and Vine, who is the world's authority on wine statistics, in the year of 2018, 293 million hectoliters of wine were produced across 36 countries. 17% increase in wine production from 2017 to 2018 [8]. Consumers are assaulted with an endless number of varieties and flavors. Some examples include red wine, white wine, rose wine, starch-based wine, etc., which are then also based on a variety of grapes, fruits like apples, and berries. For a non-expert, unfamiliar with the

various nuances that make each brand distinct, the complexity of decision making has vastly increased. In such a competitive market, wine reviews and rankings matter a lot since they become part of the heuristics that drive consumers' decision making. Producers of wine gain a competitive advantage by knowing what factors contribute the most to quality as determined by rankings. What has also changed is the amount of data available. Moore's law and other advances in computing have allowed for the collection and analysis of vast amounts of data.

Wineinformatics is the application of data science techniques to the advancement of wine production and quality. The sources of data can include physicochemical laboratory data and wine reviews [9]. Physicochemical laboratory data usually relates to the physicochemical composition analysis, such as acidity, residual sugar, alcohol, etc., to characterize wine. However, physicochemical analysis cannot express the sensory quality of wine. Wine reviews are produced by sommeliers, people who specialize in wine. These wine reviews usually include aroma, flavors, tannins, weight, finish, appearance, and the interactions related to these wine sensations [10]. The physicochemical laboratory data is easy to read and apply analytics to, while wine reviews' data involves natural language processing and a degree of human bias. However, wine review data is still valuable in its ability to directly generate significant consumer insights. Because of this, the information is potentially useful for producers, distributors, and other consumers.

1.3 Bordeaux Wine

Since the mid-1st century, the Romans have introduced the wine to Bordeaux, France, for local consumption, Bordeaux has been a region that continuously produced wine [11]. In the 12th century, Bordeaux wine gained popularity in England because of

the marriage of Eleanor, Duchess of Aquitaine, and Henry Plantagenet, the future King of England [12]. Bordeaux is the biggest wine delivering district in France and one of the most influential wine districts in the world. A Bordeaux wine is any wine produced in Bordeaux, France. Bordeaux has four different classifications covering different areas of the region. The most famous and oldest classification of Bordeaux is the official Bordeaux wine classification in 1855, which was developed at the request of Emperor Napoleon III, and aims to showcase the best Bordeaux wines in France to tourists around the world [13]. The official Bordeaux wine classification in 1855 was ranked according to the château's reputation and trading price, when the reputation and trading price of the wine were related to quality directly. There are five growths in red wine: Premiers Crus, Deuxièmes Crus, Troisièmes Crus, Quatrièmes Crus, and Cinquièmes Crus, where Premiers Crus are one of the most expensive wines in the world. All of the red wines in the list are from the Médoc, except for one: Château Haut-Brion from Graves. The white wine was less important than red wine, so there are only three growths: Premier Cru Supérieur, Premiers Crus, and Deuxièmes Crus [14].

There is a lot of research that focuses on the price and vintage of Bordeaux wines [15,16,17] from historical and economic data. Shanmuganathan et al. applied decision tree and statistical methods for modeling seasonal climate effects on grapevine yield and wine quality [18]. Noy et al. developed the ontology on Bordeaux wine [19,20]. Most of these Bordeaux or wine related data mining research applied their work on small to medium sized wine datasets [21-24]. However, to the best of our knowledge, there is almost no research utilizing data mining to determine the quality and character of various vintages of wines in Bordeaux comparable to the size of our dataset. In addition, we also

developed a dataset that contains all the available wines listed in the official Bordeaux wine classification in 1855 to uncover the important characteristics from this famous classification.

1.4 Goal of the Research

To study 21st century Bordeaux wines based on our previous work, we developed two new datasets through the Computational Wine Wheel related to Bordeaux. For the first dataset, we collected all the available Bordeaux wine reviews on the latest vintage (2000–2016) from the Wine Spectator [25]. These reviews are then converted from human language format into computer encoded codes through the computational wine wheel proposed in the research of Wineinformatics [26-29]. For the second dataset, we are interested in a famous collection of Bordeaux wines, 1855 Bordeaux Wine Official Classification. The quality of the wine in both datasets was determined by experts in a blind taste test. This was based upon an interval scale from 50–100 in which 100 was determined to be the highest quality while 50 being the wine that was not recommended due to quality issues. We will train algorithms on both datasets and see which one is most effective at classifying it in the 90+ category or 89− category through Naïve Bayes and SVM. If the algorithms are effective, we can potentially uncover the words most predictive of wine quality and enlighten producers on how to maintain and/or improve the quality of their wine allowing them to succeed in such a competitive environment.

In previous research [26-29], the usage of the Computational Wine Wheel was focused on two columns, SPECIFIC_NAME and NORMALIZED_NAME. The column SPECIFIC_NAME contains the keywords/tasting notes that appeared in Wine Spectator's Top 100 Wines from 2003 to 2013. The column NORMALIZED_NAME

contains the keywords normalized from the column SPECIFIC_NAME. The process of normalization involves the categorization of different words into their root form. For example, “freshly cut apple”, “ripe apple” and “apple” were normalized into “apple” since they all represent the same flavor. We expanded upon previous research by including two other columns that were also derived from the Computational Wine Wheel (there were a total of 4). These two additional columns were CATEGORY_NAME and SUBCATEGORY_NAME. For example, apple will be under “tree fruit” in the SUBCATEGORY_NAME column and “fruity” in the CATEGORY_NAME column. We decided to add these 2 new columns because we are interested in how the word frequency in the categories would affect our research. Will the additional column that includes the word frequencies from the category of “fruity” have a positive influence on the wine review? Would the number words in the subcategories “tropical fruit” and “dried fruit” affect the perception of wine quality amongst experts?

CHAPTER 2 WINE DATA

Data science is the study of data. The source of the data, the pre-processing of the data and the creation of the data, are all major factors to the quality of the data. In this study, we are trying to better understand the words used in professional wine reviews so that parties of interest can potentially find ways to improve their wine ranking and consumers can find their preferences. However, wine reviews are in human language format with a ranking score; in order to make computers understand wine reviews, Natural Language Processing (NLP) is needed. The domain knowledge is the major gateway to build the NLP program. There are many words in the targeted wine reviews, which words are irrelevant for analysis? Should the word "the" be included as a potential determinant of wine rankings? A computer does not know this but, in perhaps an unusual case, a human unfamiliar with the English language and wine may also be unable to clean the data appropriately thus harming the final analysis and predictions. Therefore, domain knowledge is often necessary for the application of data science. Ultimately, we will see, based on our cleaning the text data, whether and to what degree our statistical learning algorithms can predict wine quality rankings. These truths may help wine businesses and consumers optimize their decision making.

2.1 Data Source: Wine Spectator

The performance of data mining research relies on the quality of the data. In this research work, we focus on the wine reviews in human language format with a score as a verdict to wine. A lot of research points out the inconsistency between wine judges as well as the bias in taste [30,31]. The small group of wine experts may not agree with each other while they taste research designated wines. Every wine expert might have their own

tasting palate, wine preference, choice of word, etc. [32-37]. Therefore, we focused on a single reputable wine review organization: Wine Spectator to gather thousands of wine reviews as the research input dataset. The Wine Spectator is a wine evaluation company in which experts give credible reviews to those most interested in wine quality. Since its inception, the company has published a total of around 400,000 wine reviews. The Wine Spectator magazine publishes 15 issues a year with each containing 400 to 1000 reviews [25]. Figure 1 is the example of wine reviews on WineSpectator.com.

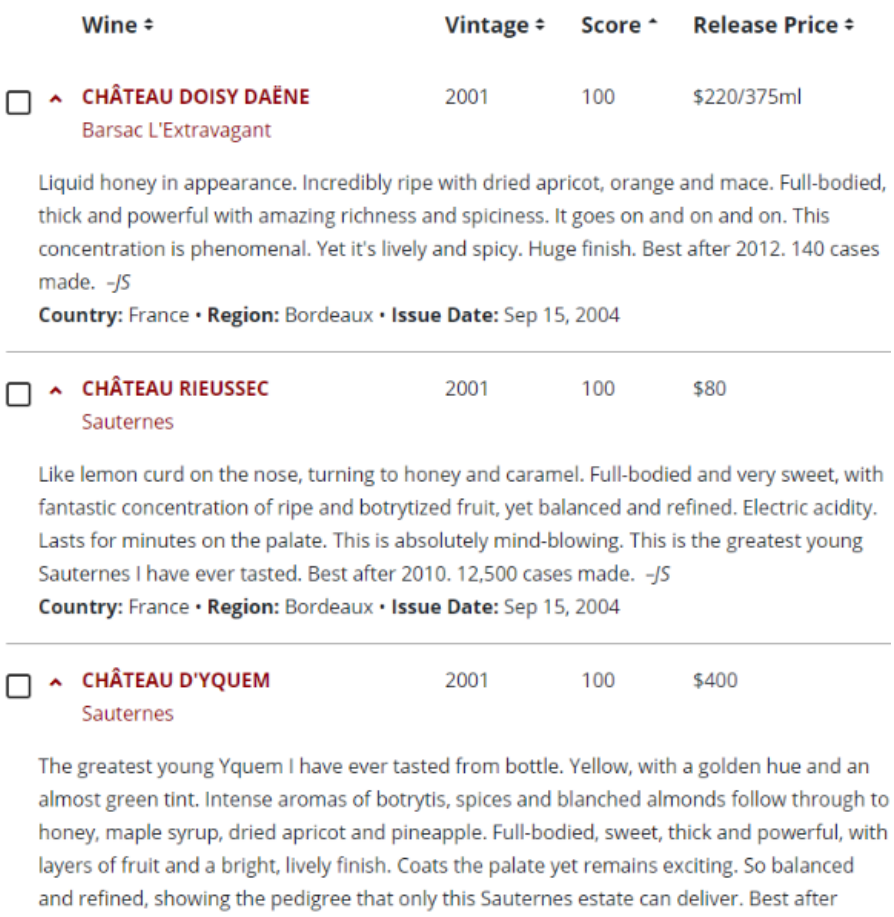


Figure 1. The example of wine reviews on WineSpectator.com

There is a 50–100 score scale on the evaluation of the wine used by Wine Spectator. Details of the score scale can be found in table1.

Table 1. Wine Spectator's 100-Point Scale

CLASSIFICATION	SCORE	COMMENT
CLASSIC	95–100	A GREAT WINE
OUTSTANDING	90–94	A WINE OF SUPERIOR CHARACTER AND STYLE
VERY GOOD	85–89	A WINE WITH SPECIAL QUALITIES
GOOD	80–84	A SOLID, WELL-MADE WINE
MEDIOCRE	75-79	A DRINKABLE WINE THAT MAY HAVE MINOR FLAWS
NOT RECOMMENDED	50-74	

Although there are some challenges on Wine Spectator's rating, ranking, and comments [38-40], based on the previous research [26-29], the correlation between wine reviews and their grading are strong. To predict a wine's quality, it receives a 90+ or 90− score based on Wine Spectators' wine reviews, built on a dataset with more than 100,000 wine reviews achieved 87.21% and 84.71% accuracy through Support Vector Machine (SVM) and Naïve Bayes model, accordingly. The regression model built on the same dataset to predict a wine's actual score based on Wine Spectator's wine reviews was only 1.6 points away on Mean Absolute Error (MAE) evaluation [28]. These findings support that the large amount of Wine Spectators' reviews are consistent and suitable for our data mining research.

2.2 Dataset Creation and Cleaning

There are a total of 14,349 wine reviews for all the Bordeaux wines made in the 21st century (2000–2016) that are available on WineSpectator.com. The easiest way to

collect the reviews is copying and pasting manually. However, it is not feasible to do it for 14,349 reviews with hundreds of web pages (15 reviews in each page in WineSpectator.com). This is where web scraping comes into play. Web scraping is an automated method of extracting large amounts of data from websites, which can then be saved as a file or spreadsheet on your local computer [41]. Figure 2 shows the flowchart of web scraping.

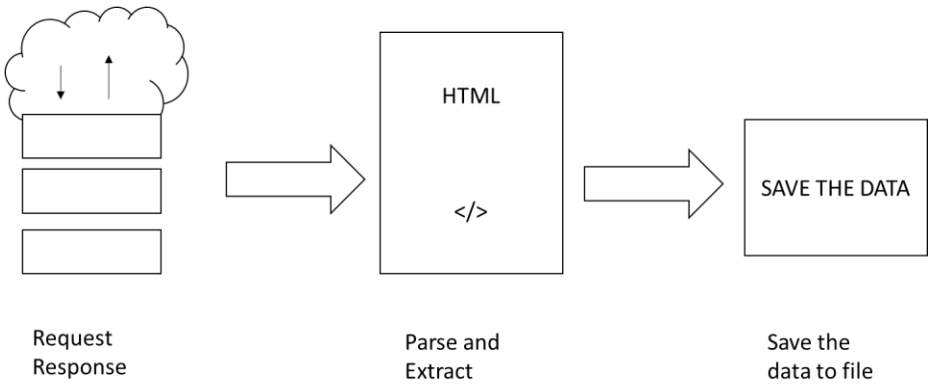


Figure 2. The flowchart of web scraping

The whole process was done in Python. Three external packages were used to perform the web scraping task. “Requests” package was used for requesting the contents of the URL in HTML format. “BeautifulSoup” was for pulling desired data out of HTML. And “re” for extracting some data inside the text with patterns. Figure 3 is an example of the wine reviews in the web page and HTML code.

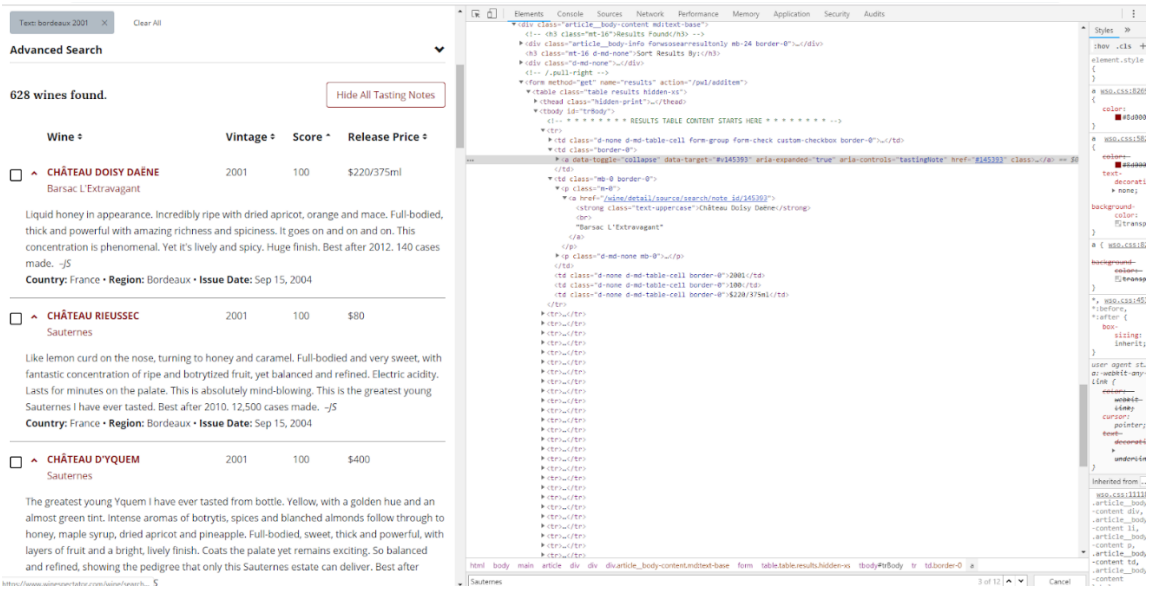


Figure 3. The example of the reviews in the web page and JavaScript code.

In order to complete our task, we need the name, review, year, score, and price of each wine. Figure 4 is a piece of HTML code that contains the name of the wine. Our goal is to extract the name out of the HTML code. We need to load the HTML code into BeautifulSoup first, then call the BeautifulSoup to look for all <p> with the class name of "m-0."

```
▼<p class="m-0">
  ▼<a href="/wine/detail/source/search/note_id/145393">
      <strong class="text-uppercase">Château Doisy Daëne</strong>
      <br>
      "Barsac L'Extravagant"
    </a>
  </p> == $0
```

Figure 4. A piece of HTML code that contains the name of a wine.

```
soup = BeautifulSoup(html, 'html.parser')
#print(num_wine)
#Get the name of the wine
wine_names = soup.find_all('p', 'm-0')
print(wine_names[0].get_text(' ','br/'))
```

Figure 5. A piece of python code to extract the name of a wine

After all the information was retrieved, we found out the scores on some reviews had the value of “BT”. By looking back at the reviews with “BT” value on scores, we found out the reviewers did give the score, but it was inside of the review notes. Figure 6 is an example of the wine review with score “BT.” To solve this issue, “re” package was used to extract the numbers with certain patterns. In this case, the pattern will be the number right after “Score range:” and “- “. We then assigned the average of those two values as the final score.

CHÂTEAU GUIRAUD Sauternes
Shows good apricot and botrytis spice character. Medium- to full-bodied, with a thick texture and a long, fruity, sweet finish. Score range: 89-91 –JS
Country: France • Region: Bordeaux • Issue Date: Web Only - 2001
2000 BT $NA

Figure 6. An example of the wine review with score “BT

After we retrieved all the data we need, we saved them into a txt file. The final file looked like figure 7.

```
Château Léoville Las Cases St.-Julien
Absolutely fantastic. This is one of the most exciting young reds I have
tasted in a long, long time. It shows intense aromas of berries,
currants and minerals, with hints of mint. Full-bodied and packed with
fruit and tannins, its long finish is refined and silky. A benchmark for
the vintage. Las Cases has always wanted to make first-growth quality in
a top-notch vintage, and it certainly did in 2000. Best after 2012.
15,000 cases made.   -JS
Country: France
2000    100 $170
Château Latour Pauillac
The fruit here is still very much in the primary phase, with a decidedly
racy feel to the raspberry coulis, cassis and blackberry reduction notes
that are streaked with violet, iron and graphite flavors. The superlong
finish alternates between a tug of sweet earth and a velvety feel, as
the fruit and grip are still melding together, but there's so much
vivacity here, there's no concern with waiting it out. The wait may be a
while though. Rather stunning that this can separate itself so clearly
from the rest of 2000's high-class field.--Blind 2000 Bordeaux
retrospective (December 2015). Best from 2020 through 2040. 14,167 cases
made.    -JM
Country: France
2000    99  $475
Château Lafite Rothschild Pauillac
This is remarkably young, with a deep well of succulent black currant,
fig and blackberry fruit notes that feel 10 years younger than most
peers, carried by wave upon wave of velvety tannins. Despite the density
and heft, there's glorious length and finesse too, with alluring black
tea, smoldering charcoal and warm paving stone notes just starting to
emerge. Awesome wine.--Blind 2000 Bordeaux retrospective (December
2015). Best from 2018 through 2043. 16,000 cases made.  -JM
Country: France
2000    98  $400
```

Figure 7. The format of wine reviews

2.3 The Computational Wine Wheel

Since the wine reviews are stored in human language format, we must convert the reviews into something that is machines understandable. The Computational Wine Wheel was developed to capture keywords/tasting notes in the wine reviews. The keywords/tasting notes were developed based on the reviews of the Wine Spectator's Top 100 Wines from 2003 to 2013. Table 2 is a simplified Computational Wine Wheel. Currently, there are four columns in the computational wine wheel: it has 14 attributes

under category column; 34 attributes under subcategory column; 1881 attributes which are the specific keywords/tasting notes under the original column; and 985 normalized attributes that were normalized from the specific keywords/tasting notes. For example, the specific keywords/tasting notes, fresh-cut apple, apple, and ripe apple are normalized into "Apple" since they represent the same flavor; yet, green apple belongs to "Green Apple" since the flavor of green apple is different from apple.

Table 2. Simplified computational wine wheel

CATEGORY	SUBCATEGORY	ORIGINAL	NORMALIZED
FRUITY	TREE FRUIT	FRESH-CUT APPLE	APPLE
FRUITY	TREE FRUIT	RIPE APPLE	APPLE
FRUITY	TREE FRUIT	APPLE	APPLE
OVERALL	TANNINS	DENSE TANNINS	TANNINS_HIGH
OVERALL	TANNINS	CRISP TANNINS	TANNINS_HIGH
OVERALL	TANNINS	SOFT TANNINS	TANNINS_LOW
HERBS/VEGETABLES	FRESH	MINTY	MINT

The Computational Wine Wheel works as a dictionary to one-hot encoding in order to convert words into vectors. The attributes under NORMALIZED are the dictionary we used. For example, there are some words that contain fruits such as apple, blueberry, plum, etc. If the wine matches the attribute in the computation wine wheel, it will be 1, otherwise, it will be 0. More examples can be found in Figure 8. Many other wine characteristics are included in the Computational Wine Wheel, such as descriptive adjectives (balance, beautifully…etc.) and body of the wine (acidity, level of tannin…etc.).

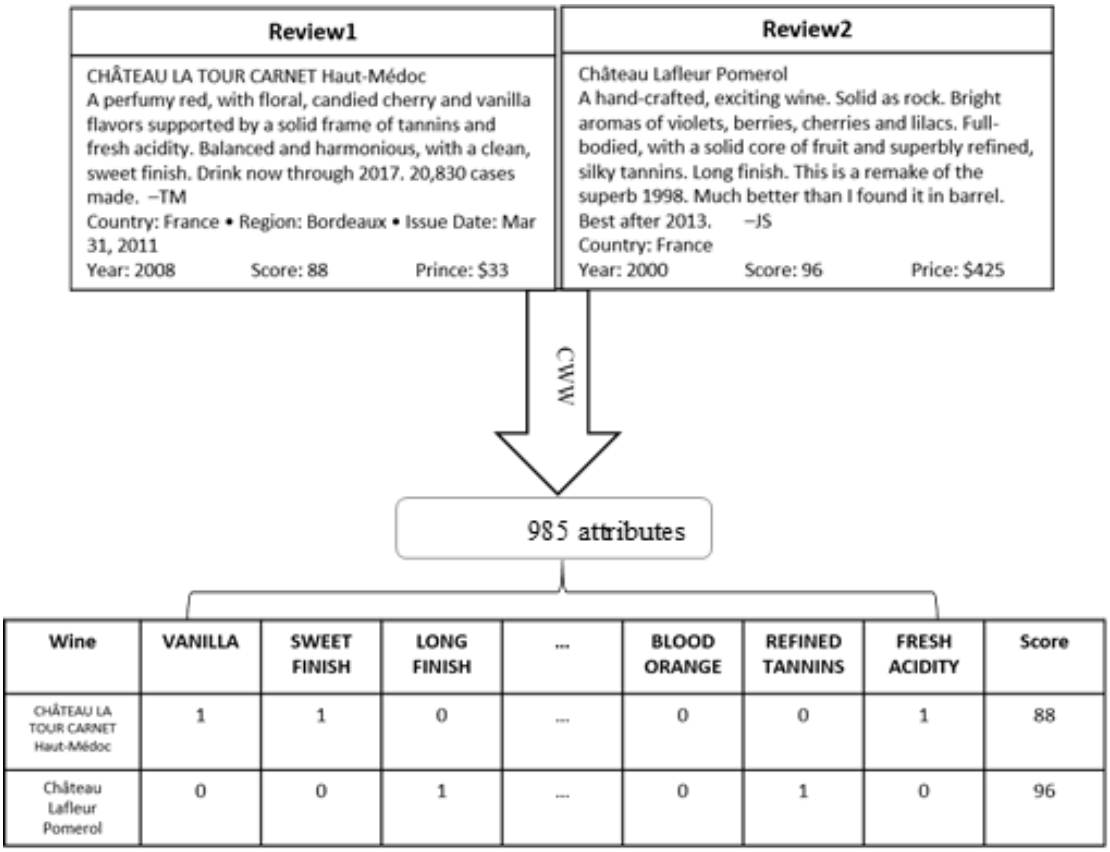


Figure 8. The flowchart of coveting reviews into machines understandable through the computation wine wheel.

2.4 Datasets

Two datasets were being studied in this research. The first one is the reviews for all the Bordeaux wines made in the 21st century. The second one is the reviews for a famous collection of Bordeaux wines, 1855 Bordeaux Wine Official Classification, which was made in the 21st century as well. In this research, we use 90 points as a cutting point. If a wine receives a score equals/above 90 points out of 100, we mark the label as a positive (+) class to the wine. Otherwise, the label would be a negative (-) class. There are some wines that scored a ranged score, such as 85-88. We use the average of the ranged score to decide and assign the label. The second dataset is a subset of the first

dataset. All the available wine reviews were collected from Wine Spectator. Details of each dataset will be discussed as follows.

2.4.1 All Bordeaux Wine Dataset

A total of 14,349 wines had been collected. There are 4263 90+ wines and 10086 89- wines. There are more 89- wines than there are 90+ wines. The score distribution is given in figure 9. Most wines score between 86 and 90. Therefore, they fall into the category of "Very Good" wine. In Figure 10, the line chart is used to represent the trend of the number of wines that have been reviewed in each year. The chart also reflects the quality of vintages. More than 1,200 wines were reviewed in 2009 and 2010, which indicates that 2009 and 2010 are good vintages in Bordeaux. Winemakers are more willing to send their wines to be reviewed if their wines are good.

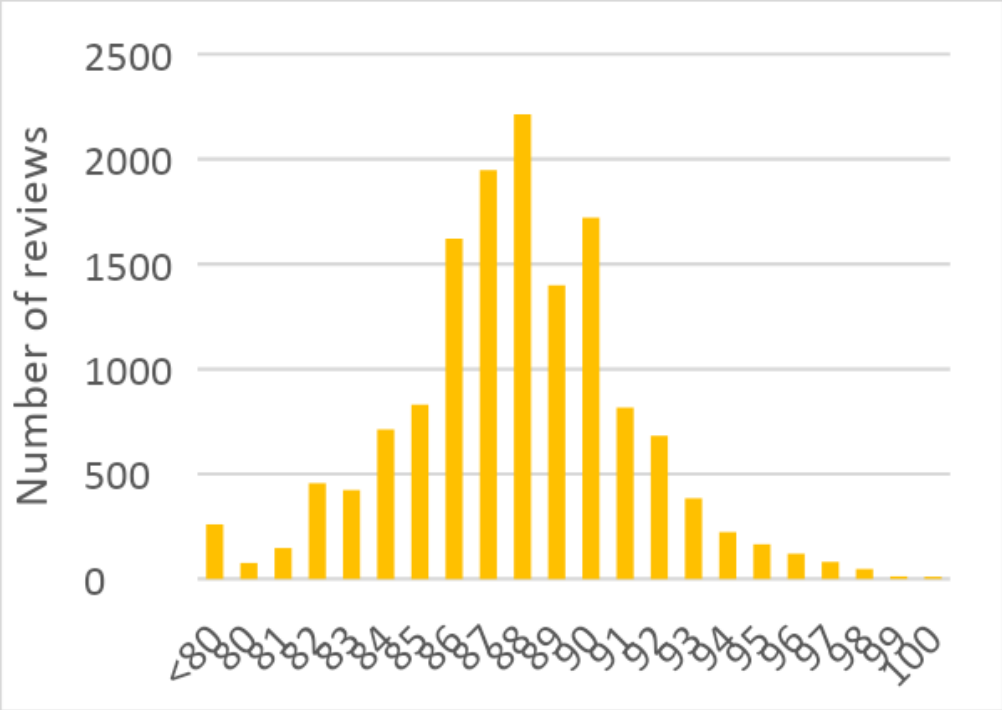


Figure 9. The score distribution of all Bordeaux Wines

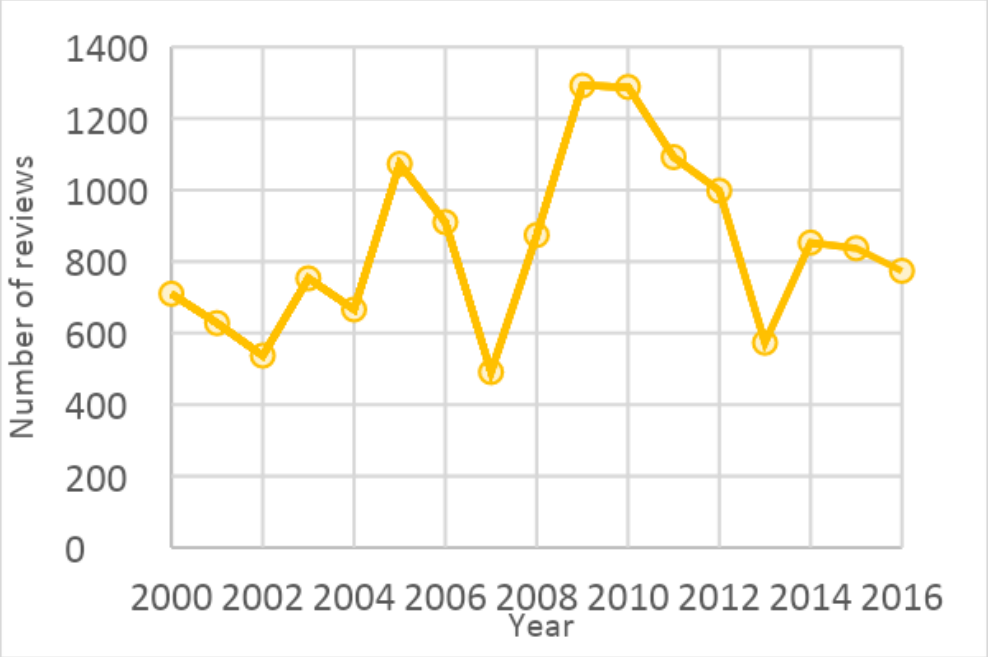


Figure 10. The number of wines reviewed annually

2.4.2 Bordeaux Wine Official Classification in 1855

A total of 1359 wines have been collected. In this dataset, we have 882 90+ wines and 477 89− wines. The score distribution is given in Figure 11. Unlike the data distribution of the first dataset, which has much more 89− wines than 90+ wines, in The Wine Spectator, the wines selected in this research are elite choices based on Bordeaux Wine Official Classification in 1855 (a complete list of Bordeaux Wine Official Classification in 1855 is given in Appendix A). Therefore, classic (95+ points) and outstanding (90–94 points) wines compose the majority of this dataset. The number of wines reviewed annually is given in Figure 12. Since Bordeaux Wine Official Classification in 1855 is a famous collection of Bordeaux wines, wine makers send their wine for review almost every year. Therefore, the line chart remains stable, which is very different from Figure 10. Regardless, some wines listed in Bordeaux Wine Official

Classification in 1855 may be still missing their wine reviews in Wine Spectator. A complete list of wines and vintages we cannot find within this dataset's scope is listed in Appendix B.

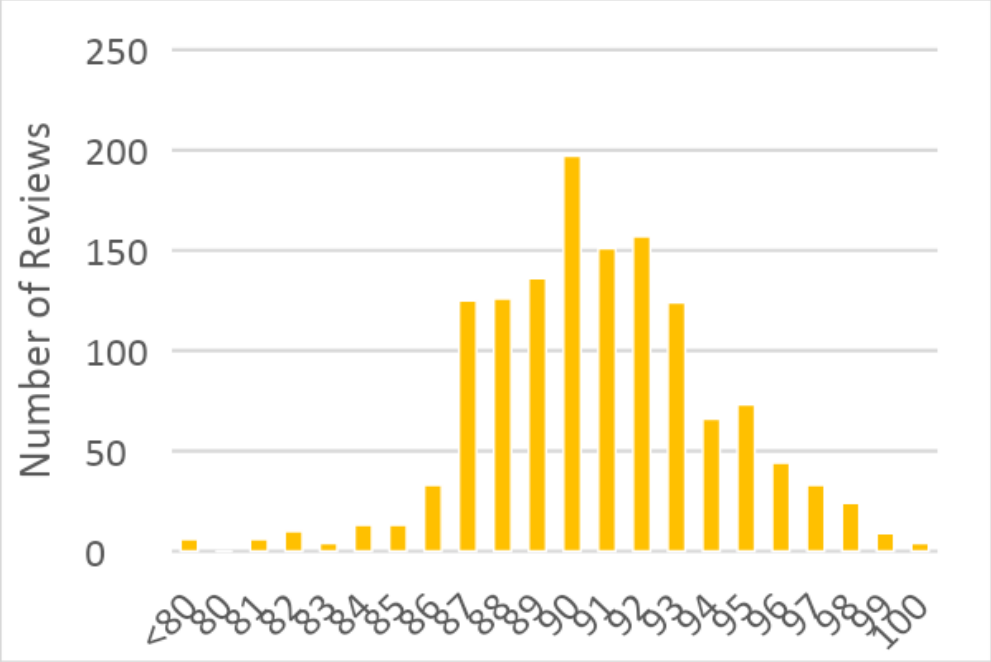


Figure 11. The score distribution of Bordeaux Wine Official Classification in 1855

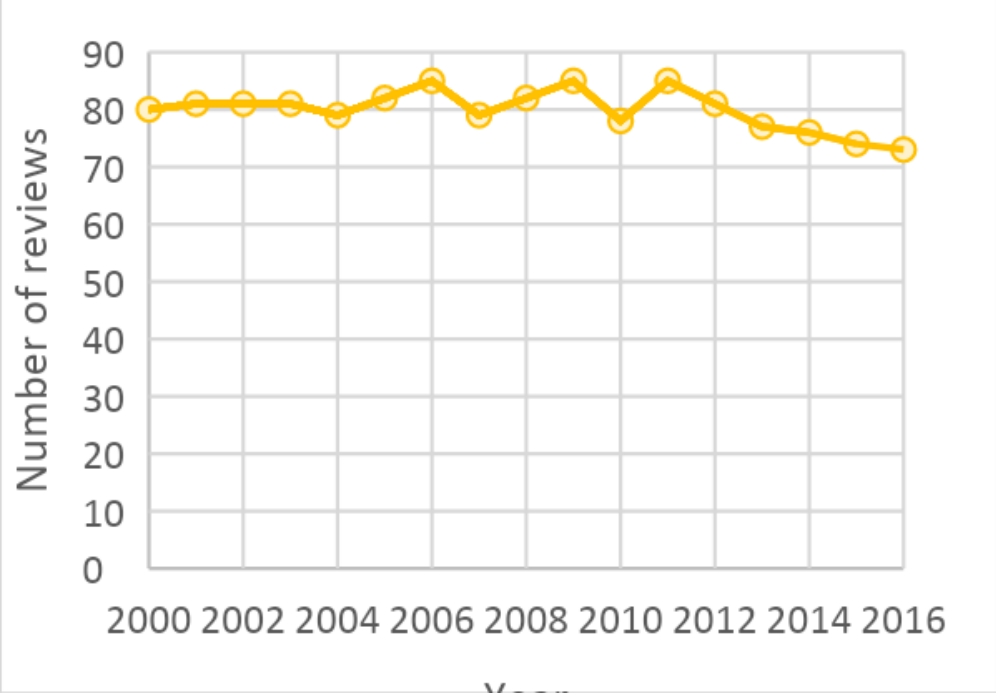


Figure 12. The number of wines reviewed annually

CHAPTER 3 PREDICTION METHODOLOGY AND EVALUATIONS

Supervised learning occurs when the training dataset is labelled, where the training dataset consists of input features and an output label. If the label is a continuous variable, it is known as **Regression**. If the label is a discrete variable, it is known as **Classification**. In this research, we are trying to understand 21st century Bordeaux wine, especially, the characteristics of classic (95+) and outstanding (90–94) wine. In order to achieve this goal, 90 points was chosen as the cutting point. If a wine receives a score equal/above 90 points out of 100, we mark the label as a positive (+) class to the wine. Otherwise, the label would be a negative (−) class. Therefore, the task we are doing is Classification in Supervised learning. The evaluation metrics, accuracy, precision, recall, and f-score were used to evaluate the performance of models with five cross-validation.

3.1 Classification Algorithms

Our goal of this research is to find out the important wine characteristics/attributes toward 21st century general Bordeaux wines. Applying white-box classification algorithms is a way to achieve the goal. Based on the previous research, Naïve Bayes classifier algorithm achieved the best accuracy among all applied white box classification algorithms (Decision Tree, k-NN, and Naïve Bayes) [27]. Support Vector Machine (SVM) classifier algorithm, which is from black box classification algorithms family, has high utility to solve classification problems. Fern’andez-Delgado applied 179 classifiers on 121 datasets from UCI database, he found out SVM is considerably effective in many datasets [42].Therefore, in this research, we applied Naïve Bayes classifier algorithm to find out the important wine characteristics/attributes toward 21st century general

Bordeaux wines. Then we applied SVM classifier as a comparison to evaluate the goodness of Naïve Bayes classifier.

3.1.1 Naïve Bayes Classifier

Throughout Wineinformatics research, the Naive Bayes classifier has been considered as the most suitable white-box classification algorithm. A Naïve Bayes classifier is a simple probabilistic classifier by applying Bayes' theorem with two assumptions: 1) there is no dependence between attributes. For example, the word APPLE appears in the review has nothing to do with the appearance of the word FRUITY even though that word also appears in the review. 2) In terms of the importance of the label/outcome, each attribute is treated the same. For example, the word APPLE and the word FRUITY have equal importance in influencing the prediction of wine quality. The assumptions made by Naïve Bayes are often incorrect in real-world applications. As a matter of fact, the assumption of independence between attributes is always wrong, but Bayes still often works well in practice [43].

Bayes' Theorem:

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}$$

$P(X|Y)$: The posterior probability of Y belongs to a particular class when X happens.

P(Y): prior probability of Y.

P(X): prior probability of X.

Naïve Bayes Classifier:

$$P(Y|X_1, X_2, \dots, X_n) = \frac{P(X_1, X_2, \dots, X_n|Y)P(Y)}{P(X_1, X_2, \dots, X_n)} = \frac{P(X_1|Y)P(X_2|Y) \dots P(X_n|Y)P(Y)}{P(X_1, X_2, \dots, X_n)}$$

$P(Y|X_1, X_2, \ldots, X_n)$: Compute all posterior probability of all values in X for all values in Y. Naïve Bayes classifier makes the prediction based on the maximum of posterior probability.

There are two main Naïve Bayes classifiers based on the type of attributes: **Gaussian Naive Bayes classifier** for continuous attributes, and **Bernoulli Naive Bayes classifier** for binary attributes.

3.1.1.1 Gaussian Naive Bayes classifier

In Gaussian Naive Bayes classifier, it is assumed that the continuous values associated with each attribute are distributed according to a Gaussian distribution. A Gaussian distribution is also known as Normal distribution.

$$P(X_i|Y) = \frac{1}{6_y\sqrt{2\pi}} \exp\left(-\frac{(X_i-\mu_y)^2}{26_y{}^2}\right)$$

μ_y: sample mean

6_y:sample standard deviation

When a value of X never appears in the training set, the prior probability of that value of X will be 0. If we do not use any techniques, $P(Y|X_1, X_2, \ldots, X_n)$ will be 0, even when some of the other prior probability of X are very high. This case does not seem fair to other X. Therefore, we assign the smallest μ_y and 6_y among all attributes to the X to handle zero prior probability.

3.1.1.2 Bernoulli Naive Bayes Classifier

In Bernoulli Naive Bayes Classifier, the attributes are binary variables. The frequency of a word occurs in the reviews is used as the probability.

$$P(X_i|Y) = \frac{N_{ic}}{N_C}$$

N_{ic}: The number of samples/reviews having attribute X_i and belongs to class Y

N_C: The number of samples/reviews belongs to class Y

Laplace Smoothing:

$$P(X_i|Y) = \frac{N_{ic} + 1}{N_C + c}$$

c: number of values in Y

Laplace smoothing was used to handle zero prior probability in Bernoulli Naive Bayes Classifier.

3.1.2 Support Vector Machines

Support vector machines (SVMs) are a set of related supervised learning algorithms used for classification and regression analysis. The SVMs have high utility to solve classification problems [42]. In the SVM for classification, each instance is plotted as a point in the n-dimensional space with their corresponding attribute values, where n represents the number of attributes. Then, the SVM looks for a line/hyperplane with maximum margin to separate the instances from different classes, where margin is the distance between the hyperplane and the edge points from different classes [44]. Those edge points are also called the "support vectors." To show how SVMs work for the linearly separable binary set, we plotted each instance as a point in the 2-dimensional graph with their corresponding attribute values, X_1 and X_2, and data points were divided into two classes, blue class and green class, in the figure 13. Several different lines were drawn to separate the data points from blue classes and green class. The best choice will be the line that leaves the maximum margin from both classes. The black line in figure 14

will be the most preferable choice. The black points in figure 14 will be the "support vectors" for this sample data determining the orientation of the line [45].

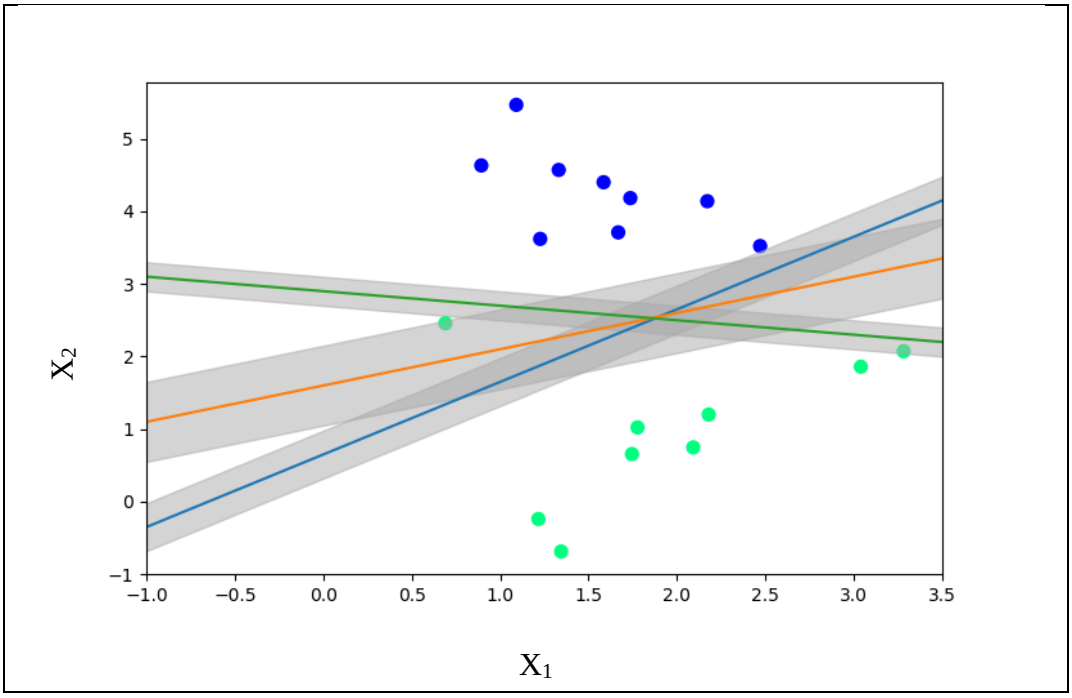


Figure 13. SVM demonstrates the process of looking for a hyperplane/line with maximum margin

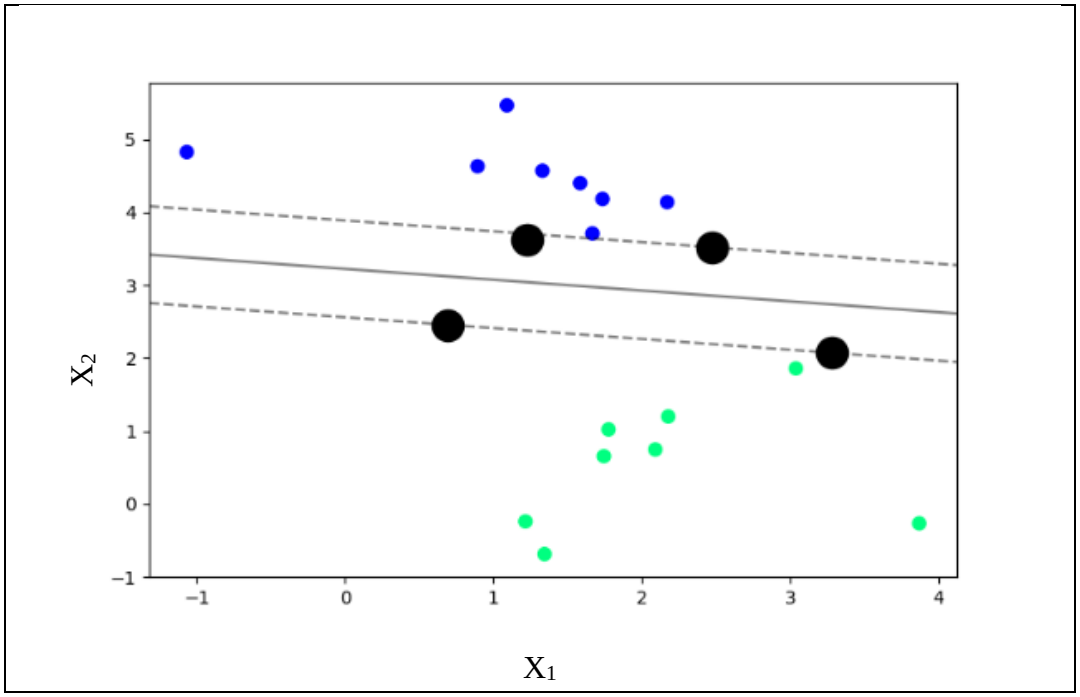


Figure 14. The hyperplane/line with maximum margin in SVM

SVM light [46] with linear kernel is the version of SVM that was used to perform the classification in this research. The linear kernel is one of the most commonly used kernels in practice, especially in text classification, where the dataset is usually in a high-dimensional, sparse feature space and linearly separable state [47, 48].

3.2 Cross-validation

Five-fold cross-validation, illustrated in Figure 15 and Figure 16, is used to evaluate the predictive performance of our models, especially in the performance of the model for new data, which can reduce overfitting to a certain extent. First, we shuffle the dataset randomly. Second, we group 90+/89- wines. Third, we split the 90+ wine group and the 89- wine group into 5 subsets separately. Fourth, we combined the first subset from the 90+ wine group and the first subset from 89- wine group into a new set, and then we repeated the same process for the rest of the data. In this way, we split our dataset into 5 subsets with the same distribution as the original dataset.

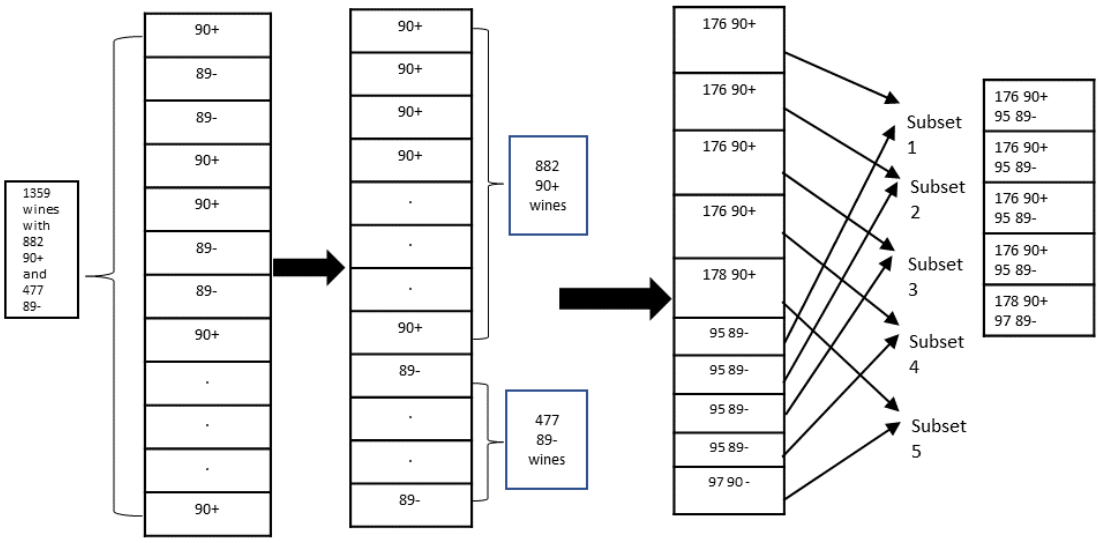


Figure 15. This figure demonstrates data splitting in 5-fold cross-validation.

After data splitting, we use the subset 1 as the test set and the rest of subsets as the training set as fold 1; we use subset 2 as the test set, and the rest of the subsets as the training set which is set as fold 2; We repeated the same process for the rest.

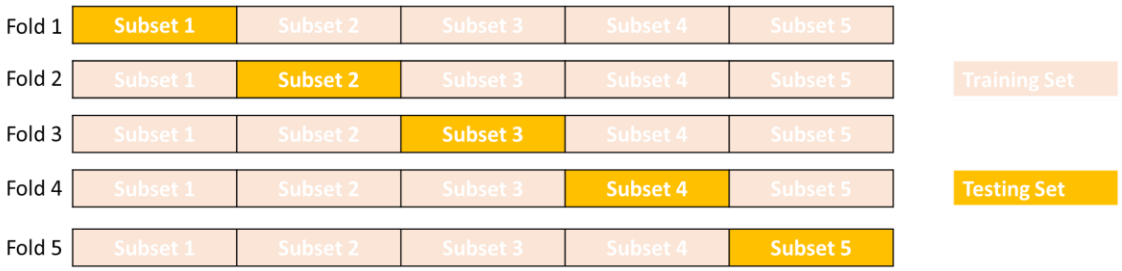


Figure 16. This figure demonstrates training and testing sets assigning in 5-fold cross-validation.

3.3 Evaluation Metrics

To evaluate the effectiveness of the classification model, several standard statistical evaluation metrics are used in this paper. First of all, we need to define True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) as:

TP: The real condition is true (1) and predicted as true (1); 90+ wine correctly classified as 90+ wine;

TN: The real condition is false (-1) and predicted as false (-1); 89- wine correctly classified as 89- wine;

FP: The real condition is false (-1) but predicted as true (1); 89- wine incorrectly classified as 90+ wine;

FN: The real condition is true (1) but predicted as false (-1); 90+ wine incorrectly classified as 89- wine;

If we use 90 points as a cutting point, to describe TP is this research's perspective would be "if a wine scores equal/above 90 and the classification model also predicts it as equal/above 90". In this research, we include the following evaluation metrics:

Accuracy: The proportion of wines that has been correctly classified among all wines. Accuracy is a very intuitive metric.

$$Accuracy = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}$$

Recall: The proportion of 90+ wines was identified correctly. Recall explains the sensitivity of the model to 90+ wine.

$$Recall = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$

Precision: The proportion of predicted 90+ wines was actually correct.

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$$

F-score: The harmonic mean of recall and precision. F-score takes both recall and precision into account, combining them into a single metric.

$$F - score = 2 \times \frac{precision * recall}{(precision + recall)}$$

CHAPTER 4 UNDERSTANDING 21ST CENTURY BORDEAUX WINES

In this chapter, we will discuss the results from both Bordeaux datasets by using Naïve Bayes classifier and SVM classifier. With the benefit of using Naïve Bayes classifier, the top 20 keywords/tasting notes in both 90+ and 89-classes were extracted to analyze the important characteristics in 21st century Bordeaux wines. A visualization technique based on Naïve Bayes was also developed to explain the decision making in Naïve Bayes.

4.1 ALL Bordeaux Wine

In ALL Bordeaux wine datasets, both Naïve Bayes Laplace classifier and SVM classifier achieved 85% accuracy or above. SVM classifier achieved the highest accuracy of 86.97%, which is from the black-box classification algorithms family. In terms of precision, SVM classifier had a much better performance than Naïve Bayes Laplace classifier, which indicates that SVM classifier has a lower false-positive rate than Naïve Bayes Laplace classifier or/and Naïve Bayes Laplace classifier has a lower true positive rate than SVM classifier. Diametrically opposed to its recall, Naïve Bayes Laplace classifier had a much better performance, which indicates that Naïve Bayes Laplace classifier had a lower false negative rate than SVM classifier or/and Naïve Bayes Laplace classifier has a higher true positive rate than SVM classifier. Naïve Bayes classifier and SVM classifier have very similar f-scores, but SVM classifier is slightly better. Overall, SVM has a slightly better performance in terms of accuracy and f-score. Details can be found in table 3. Figure 17 is generated from the results listed in Table 3.

Table 3. Accuracy, precision, recall, and f-score in different classifiers

	Accuracy	Precision	Recall	F-Score
Naïve Bayes Laplace	85.17%	73.22%	**79.03%**	76.01%
SVM	**86.97%**	**80.68%**	73.80%	**77.10%**

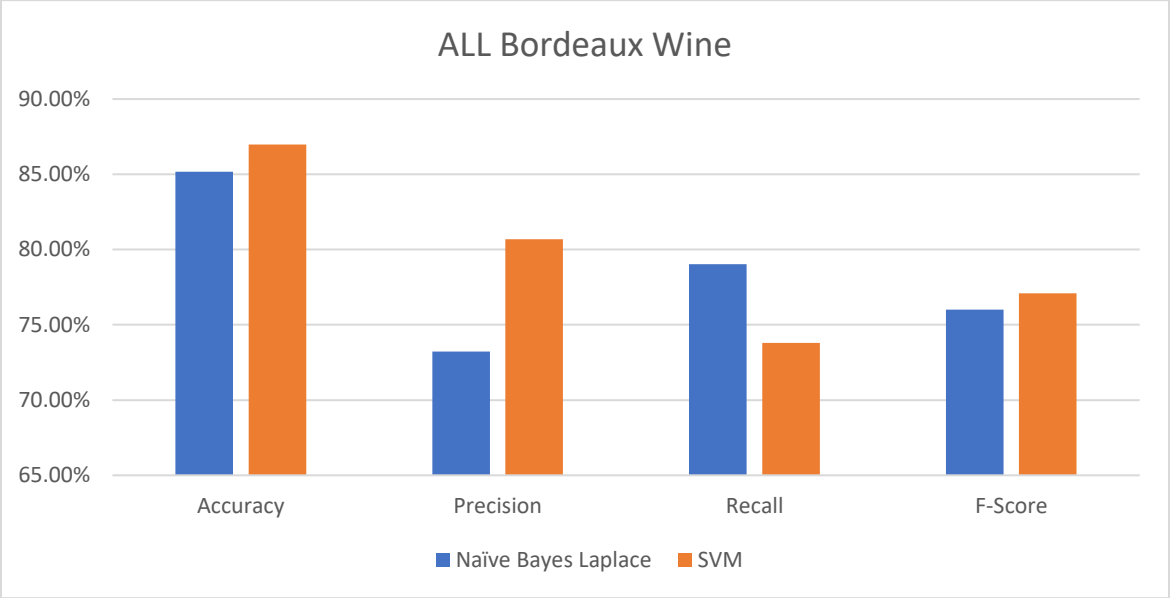


Figure 17. Accuracy, precision, recall, f-score in Naïve Bayes Laplace and SVM from ALL Bordeaux wine dataset

4.2 Bordeaux Wine Official Classification in 1855

In the Bordeaux Wine Official Classification in 1855 dataset, both the Naïve Bayes Laplace classifier and SVM classifier are able to achieve 81% accuracy or above. Naive Bayes Laplace classifier achieves the best accuracy of 84.62%. In terms of precision, Naïve Bayes Laplace classifier and SVM classifier all achieve around 86%. In terms of recall, Naive Bayes Laplace classifier achieves the best recall of 90.02%, which is around 6% higher than SVM classifier. Naïve Bayes Laplace classifier has much better sensitivity than SVM classifier. In the combination of precision and recall, Naive Bayes

Laplace classifier has the highest F-score of 88.38%. Overall, Naïve Bayes Laplace had a better performance than the SVM classifier in this specific Bordeaux wine dataset. Details can be found in table 4. Figure 14 is generated from the results listed in Table 4.

Table 4. Accuracy, precision, recall, and f-score in different classifiers

	Accuracy	Precision	Recall	F-Score
Naïve Bayes Laplace	**84.62%**	86.79%	**90.02%**	**88.38%**
SVM	81.38%	**86.84%**	84.12%	85.46%

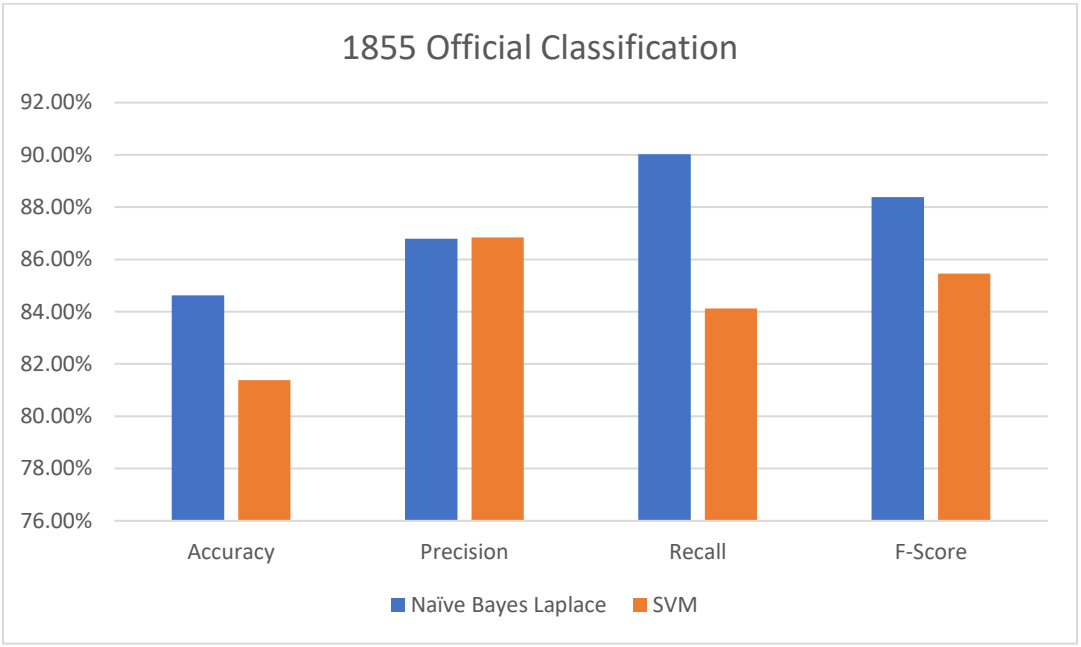


Figure 18. Accuracy, precision, recall, f-score in Naïve Bayes Laplace and SVM from 1855 Bordeaux Wine Official Classification dataset

4.3 Visualization of Bordeaux Wine Official Classification in 1855

With the benefit of using Naïve Bayes in a small dataset, we developed a visualized classification result from Naïve Bayes for the Bordeaux Wine Official Classification in 1855 dataset in Figure 19. In the figure, we have the probability that the

sample is 90+ for the horizontal axis, and the probability that the sample is 89- for the vertical axis. According to Bayes' theorem, the sample belongs to the class with a bigger probability. Therefore, a line y = x is drawn as the decision boundary. Any samples in the area that are below the line are predicted as positive classes and vice versa. The points in blue are actually from 89- class, orange is 90+ class. By seeing this figure, we can tell the numbers of misclassified samples, and can more easily see the false positive and false negative samples. Figure 20 is a zoom in picture of Figure 19 to understand the dense area of the figure. These figures demonstrate that most miss-classified wines are very close to the boundary. These visualizations provide the insight of the challenges in Wineinformatics.

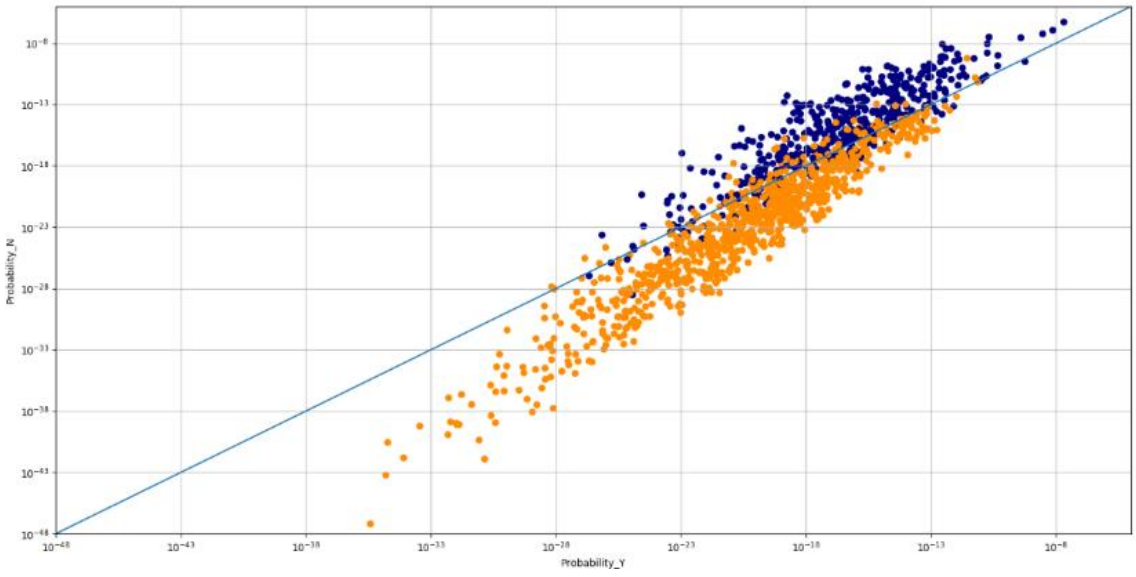


Figure 19. The visualization of the whole dataset by Naïve Bayes. The horizontal axis indicates the probability that the sample is 90+, and the vertical axis indicates the probability that the sample is type 90 -

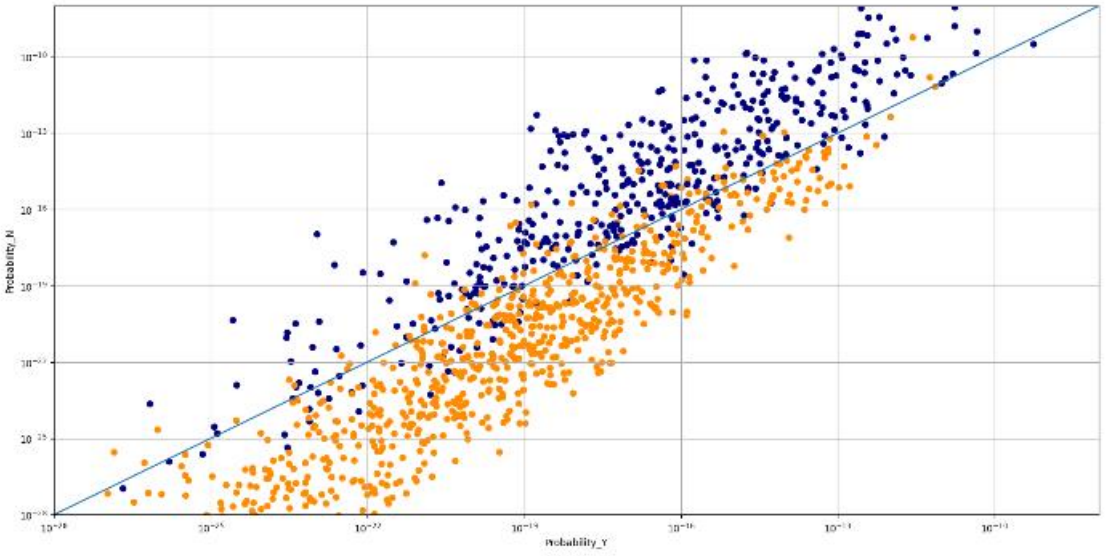


Figure 20. Magnification of Figure 19

4.4 Top 20 keywords

SVM is considered as a black-box classifier, since the classification processes are unexplainable. Naïve Bayes, on the other hand, is a white-box classification algorithm, since each attribute has its own probability to contribute to positive case and negative case. We extract keywords with 20 highest positive probabilities toward 90+ and 89− classes from both datasets.

In ALL Bordeaux Wine dataset, there are 11 common keywords that appear in both 90+ and 89− wines. Details can be found in Table 5. These common keywords represent the important wine characteristics/attributes toward 21st century general Bordeaux wines. Furthermore, our goal is to understand the important wine characteristics/attributes toward 21st century classic and outstanding Bordeaux wines. Therefore, finding out the distinct keywords between 90+ and 89− is our final goal. Details about the distinct keywords between 90+ and 89− from ALL Bordeaux Wine dataset can be found in Tables 6 and 7. According to Table 6, fruity characters including

"BLACK CURRANT", "APPLE", "RASPBERRY", and "FIG" are favorable flavors to 21st century Bordeaux. Since Bordeaux is also famous for red wines that can age for many years, "SOLID" (shows in Table 6 Body category) is preferred over "MEDIUM-BODIED" and "LIGHT-BODIED" (shows in Table 7 Body category).

Table 5. Common keywords between 90+ and 89− wines from ALL Bordeaux Wine dataset.

CATEGORY	90+ WINES AND 89− WINES			
FLAVOR/DESCRIPTORS	GREAT	FLAVORS		
FRUITY	FRUIT	PLUM	BLACKBERRY	CURRENT
BODY	FULL-BODIED	CORE		
FINISH	FINISH			
HERBS	TOBACCO			
TANNINS	TANNINS_LOW			

Table 6. Distinct keywords between 90+ and 89− wines from ALL Bordeaux Wine dataset in 90+ wines.

CATEGORY	90+ WINES			
FLAVOR/DESCRIPTORS	LONG	RANGE	RIPE	
FRUITY	BLACK CURRANT	APPLE	RASPBERRY	FIG
BODY	SOLID			
SPICE	LICORICE			

Table 7. Distinct keywords between 90+ and 89− wines from ALL Bordeaux Wine dataset in 89− wines.

CATEGORY	89− WINES		
FLAVOR/DESCRIPTORS	CHARACTER	FRESH	GOOD
FRUITY	CHERRY	BERRY	
BODY	MEDIUM-BODIED	LIGHT-BODIED	
TANNINS	TANNINE_MEDIUM		

In the 1855 Bordeaux Wine Official Classification dataset, there are 11 common keywords that appear in both 90+ and 90− wines. Details can be found in Table 8. Comparing the common keywords with ALL Bordeaux Wine dataset, 10 out of 11 are the same keywords. “TANNINES_LOW” only appears in ALL Bordeaux Wine dataset, and “SWEET” only appears in the 1855 Bordeaux Wine Official Classification dataset. Details about the distinct keywords between 90+ and 89− from 1855 Bordeaux Wine Official Classification dataset can be found in Tables 9 and 10.

Comparing the distinct keywords between 90+ and 89− wines from both datasets in 90+ wines, “LONG”, “BLACK CURRANT”, “APPLE”, and “FIG” appear in both datasets; “RANGE”, “RIPE”, “RASPBERRY”, “SOLID”, and “LICORICE” only appear in ALL Bordeaux Wine dataset; “STYLE”, “LOVELY”, “IRON”, “TANNINS_LOW”, and “SPICE” only appear in 1855 Bordeaux Wine Official Classification.

Table 8. Common keywords between 90+ and 89− wines from 1855 Bordeaux Wine Official classification dataset.

CATEGORY	**90+ WINES AND 89− WINES**			
FLAVOR/DESCRIPTORS	GREAT	FLAVORS	SWEET	
FRUITY	FRUIT	PLUM	BLACKBERRY	CURRENT
BODY	FULL-BODIED	CORE		
FINISH	FINISH			
HERBS	TOBACCO			

Table 9. Distinct keywords between 90+ and 89− wines from 1855 Bordeaux Wine Official classification dataset in 90+ wines.

CATEGORY	90+ WINES		
FLAVOR/DESCRIPTORS	LONG	STYLE	LOVELY
FRUITY	BLACK CURRENT	FIG	APPLE
EARTHY	IRON		
TANNINS	TANNINS_LOW		
SPICE	SPICE		

Table 10. Distinct keywords between 90+ and 89− wines from 1855 Bordeaux Wine Official classification dataset in 89− wines.

CATEGORY	89− WINES			
FLAVOR/DESCRIPTORS	CHARACTER	FRESH	RANGE	GOOD
FRUITY	BERRY			
BODY	MEDIUM-BODIED	LIGHT-BODIED		
TANNINS	TANNIS_MEDIUM			

CHAPTER 5 NATURAL LANGUAGE PROCESSING

In previous research [26-29], the usage of the Computational Wine Wheel was focused on two columns, SPECIFIC_NAME and NORMALIZED_NAME. The column SPECIFIC_NAME contains the keywords/tasting notes that appeared in Wine Spectator's Top 100 Wines from 2003 to 2013. The column NORMALIZED_NAME contains the keywords normalized from the column SPECIFIC_NAME. The process of normalization involves the categorization of different words into their root form. For example, "freshly cut apple", "ripe apple" and "apple" were normalized into "apple" since they all represent the same flavor. We expanded upon previous research by including two other columns that were also derived from the Computational Wine Wheel (there were a total of 4). These two additional columns were CATEGORY_NAME and SUBCATEGORY_NAME. For example, apple will be under "tree fruit" in the SUBCATEGORY_NAME column and "fruity" in the CATEGORY_NAME column. We decided to add these 2 new columns because we are interested in how the word frequency in the categories would affect our research. Will the additional column that includes the word frequencies from the category of "fruity" have a positive influence on the wine review? Would the number words in the subcategories "tropical fruit" and "dried fruit" affect the perception of wine quality amongst experts?

5.1 Data Preprocessing

In the latest version of the computational wine wheel, there are four columns CATEGORY_NAME, SUBCATEGORY_NAME, SPECIFIC_NAME, NORMALIZED_NAME. 14 attributes are in the CATEGORY_NAME, 34 attributes are

in the SUBCATEGORY_NAME; 1881 attributes are in the SPECIFIC_NAME; and 985 attributes are in the NORMALIZED_NAME. The details can be found in Table 11.

Table 11. Summary of all WINE ATTRIBUTES

CATEGORY_NAME	SUBCATEGORY_NAME	SPECIFIC_NAME	NORMALIZED_NAME
CARAMEL	CARAMEL	71	40
CHEMICAL	PETROLEUM	9	5
	SULFUR	11	10
	PUNGENT	4	3
EARTHY	EARTHY	72	31
	MOLDY	2	2
FLORAL	FLORAL	61	39
FRUITY	BERRY	49	28
	CITRUS	37	23
	DRIED FRUIT	67	60
	FRUIT	22	9
	OTHER	25	18
	TREE FRUIT	39	31
	TROPICAL FRUIT	48	27
FRESH	FRESH	41	29
	DRIED	25	21
	CANNED/COOKED	16	15
MEAT	MEAT	25	13
MICROBIOLOGICAL	YEASTY	5	4
	LACTIC	14	6
NUTTY	NUTTY	20	15
OVERALL	TANNINS	90	4
	BODY	50	23
	STRUCTURE	40	2
	ACIDITY	40	3
	FINISH	184	5
	FLAVOR/DESCRIPTORS	649	432
OXIDIZED	OXIDIZED	1	1
PUNGENT	HOT	3	2
	COOL	1	1
SPICY	SPICE	83	44
WOODY	RESINOUS	24	9
	PHENOLIC	6	4
	BURNED	47	26

In chapter 4, 985 attributes in the NORMALIZED_NAME are used as the dictionary to look over the whole text reviews. In this chapter, not only 985 NORMALIZED attributes are used, we also mapped the 985 attributes into 34 attributes in the SUBCATEGORY_NAME and 14 attributes in the CATEGORY_NAME by counting the occurrences of each attributes. Figure 21 shows us the flowchart of converting reviews into machines understandable through the computational wine wheel into NORMALIZED_NAME, SUBCATEGORY_NAME and CATEGORY_NAME.

Figure 21. The flowchart of converting reviews into machines understandable through the computation wine wheel into NORMALIZED_NAME, SUBCATEGORY_NAME and CATEGORY_NAME.

At the end, 14 category attributes and 34 subcategory attributes were added into both datasets. There are a total of 1033 attributes on both datasets now. However, the attributes in SUBCATEGORY_NAME and CATEGORY_NAME dataset are continuous values instead of Binary values; and the ranges of each attribute are various. The learning algorithm will give one feature more weights than the other if we do not do normalization with our attributes especially when the attributes in NORMALIZED_NAME are binary. Normalization can help speed up the learning, as well as avoid numerical problems, such as loss of accuracy due to arithmetic overflow [49]. In this research, we used Min-Max normalization to scale the attribute values to 0-1. The formula is as follows:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{10}$$

x is an original value

x' is the normalized value

$\min(x)$ is the minimum value among all x

$\max(x)$ is the maximum value among all x

5.2 Voting System

There is a total of 1033 attributes now. Should we treat 14 category attributes and 34 subcategory attributes the same as the 985 normalized attributes? Will both of them help us improve the model performance? And how to design a system to integrate these three different types of attribute representations?

Ensemble learning has gained popularity recently. Ensemble learning combines multiple learners/models to improve the final model predictions, especially the accuracy and robustness. If the accuracy of all learners is at least 50% and each learner is

independent, then by using majority voting to ensemble learners, the likelihood of an error will decrease as the overall number of learners increases [50,51]. Therefore, we designed a voting system that ensembles different models. Each model/voter represents a dataset consisting of a group of attributes derived from the categories, subcategories, and that which was normalized. Details of each model can be found on table 12.

Table 12. Selected Attributes in each Models

Model/Voter	Attributes
1	14 category attributes
2	34 subcategory attributes
3	985 normalized attributes
4	14 category attributes+34 subcategory attributes+985 normalized attributes
5	14 category attributes+34 subcategory attributes
6	34 subcategory attributes+985 normalized attributes
7	14 category attributes +985 normalized attributes

There are seven models/voters built based on these seven datasets. These seven models/voters will be our initial voters. The quality of the predictions of the models built upon these voters is a determinant of how the voters are selected. 1) at least 80% accuracy; 2) if there are even numbers of voters passed with 80% accuracy, we will drop the one with least accuracy. The final decision will be generated based on the majority vote of those final voters meeting the criteria of the voter selection mechanisms. The details can be found in figure 22.

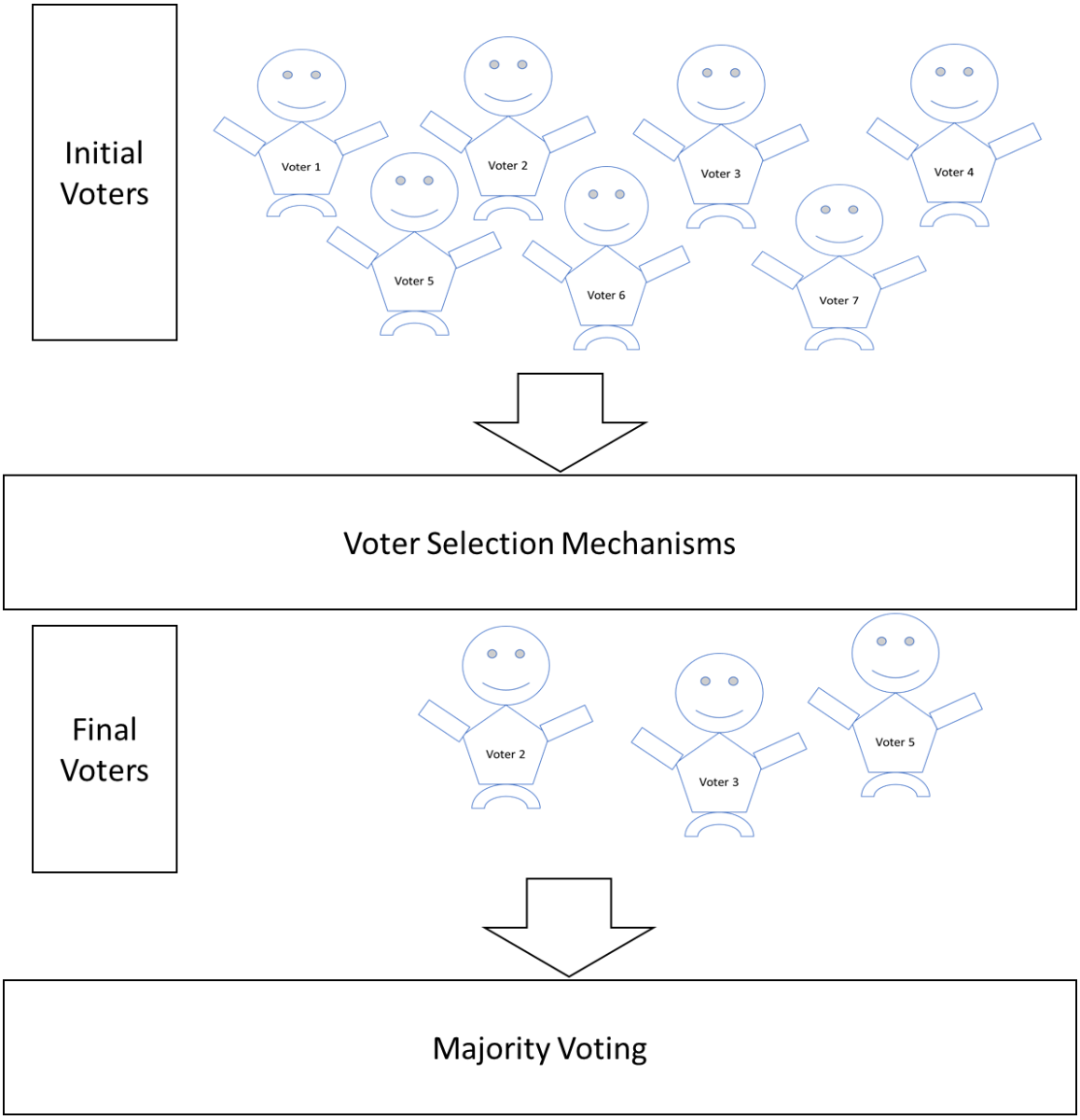


Figure 22. Voting System Design

5.3 Classifications and Results

Based on the voting system, we applied the Naïve Bayes classifier and SVM classifier on both datasets. The voting system will be useful if there are some improvements in the prediction performances of the Naïve Bayes classifier or/and SVM

classifier. We can also look at the voting system to find out whether the newly added attributes are useful or not. The results will be discussed below.

5.3.1 ALL Bordeaux Wine

Table 13 shows the results for each voter using the Naive Bayes classifier. Voter 3 is the one who uses 985 normalized attributes, which is the same model we built in chapter 4. Based on the 80% accuracy threshold, four voters can participate in the final vote. However, four is an even number, and in some cases, we might get equal votes of each class. To avoid this problem, we discarded the one with the lowest accuracy. Therefore, we have voters 3, 6 and 7 to make the final decision. Voter 3 is the one who uses 985 normalized attributes, which is the same model we built in chapter 4. As all the previous research only used the attributes that voter 3 used as well. To see if our model has improved, Voter 3 will be a good comparison. The accuracy of the final model achieved 88.10%, which is 3 % higher than Voter 3. The final model improved 10% in precision. However, the recall has fallen by 5%. Therefore, there is only 2% improvement in f-score. Overall, our voting system generated a better result than chapter 4 (only using 985 normalized attributes) with Naïve Bayes in the all Bordeaux wine dataset.

Table 13. Accuracy, precision, recall, and f-score with Naïve Bayes Classifier in All Bordeaux Wine

Voter/Model	Accuracy	Precision	Recall	F-score
1	74.39%	61.64%	36.48%	45.83%
2	74.72%	61.17%	40.86%	48.98%
3	85.17%	73.22%	79.03%	76.01%
4	82.37%	77.65%	57.10%	65.80%
5	74.93%	62.09%	40.11%	48.73%
6	84.79%	81.32%	63.38%	71.22%
7	87.32%	81.94%	73.52%	77.49%
Final (v3, v6, v7)	88.10%	83.62%	74.55%	78.82%

The final model consisted of voters 3, 6 and 7. Voter 3 was derived from the 985 normalized attributes; the voter 6 was derived from 34 subcategory attributes and 985 normalized attributes; the voter 7 was derived from 14 category attributes and 985 normalized attributes. This further proved that the newly added attributes helped us improve the model performance with Naïve Bayes in the all Bordeaux wine dataset.

Table 14 shows the results for each voter using the SVM classifier. All the seven voters passed the threshold of 80% accuracy. Therefore, all seven of them participated in the final vote. The result from our voting system is really close to the result from chapter 4(voter 3). It seems that the newly added attributes do not have the positive effect on the SVM classifier in the all Bordeaux wine dataset. However, the recall among some voters

is as low as 53.81% even though they passed the 80% accuracy threshold. We might generate a better result if we put the thresholds on precision and recall as well.

Table 14. Accuracy, precision, recall, and f-score with SVM Classifier in All Bordeaux Wine

Voter/Model	Accuracy	Precision	Recall	F-score
1	80.46%	73.31%	53.81%	62.06%
2	82.09%	75.58%	58.67%	66.06%
3	86.97%	80.68%	74.80%	77.10%
4	87.00%	80.35%	74.50%	77.31%
5	82.12%	75.53%	58.88%	66.17%
6	87.00%	80.31%	74.53%	77.30%
7	86.92%	80.13%	74.45%	77.18%
Final (all voters)	86.99%	80.38%	74.38%	77.26%

5.3.2 Bordeaux Wine Official Classification in 1855

Table 15 shows the results for each voter using the Naive Bayes classifier in the Bordeaux Wine Official Classification in 1855 dataset. There were 4 voters passed the 80% accuracy threshold. Voter 3 was discarded since it had the lowest accuracy in the final voters' group. Therefore, voters 4, 6 and 7 ensembled to make the final decision. Compared to voter 3, the final model of the voting system had a 2% improvement on accuracy, 0.22% on improvement on precision, 4% improvement on recall, and 2% improvement on f-score. Overall, the performance of Naive Bayes classifier in the Bordeaux Wine Official Classification in 1855 dataset was improved, which further proves that the newly added attributes are useful when using Naïve Bayes.

Table 15. Accuracy, precision, recall, and f-score with Naïve Bayes Classifier in 1855 Bordeaux Wine

Voter/Model	Accuracy	Precision	Recall	F-score
1	70.36%	77.15%	78.21%	77.37%
2	71.97%	75.10%	85.73%	79.91%
3	84.62%	86.79%	90.02%	88.38%
4	86.18%	85.92%	94.33%	89.89%
5	76.02%	76.08%	92.63%	83.43%
6	86.17%	86.54%	93.31%	89.78%
7	85.88%	91.34%	86.72%	88.84%
Final (v4, v6, v7)	87.06%	87.01%	94.22%	90.45%

Table 16 shows the results for each voter using the SVM classifier in the Bordeaux Wine Official Classification in 1855 dataset. Among all the seven initial voters, there were five of them passed the 80% accuracy threshold. Therefore, the final decision was made upon voter 3, 4, 5, 6 and 7. Compared to voter 3, the final model of the voting system had 4.63% improvement on accuracy, 1.82% on precision, 5.9% on recall, and 80.77% on f-score. Overall, the performance of the SVM classifier in the Bordeaux Wine Official Classification in 1855 dataset has been improved, which further proved that the newly added attributes are useful with the SVM classifier.

Table 16 Accuracy, precision, recall, and f-score with SVM Classifier in 1855 Bordeaux Wine

Voter/Model	Accuracy	Precision	Recall	F-score
1	71.22%	73.94%	86.17%	79.55%
2	79.18%	82.17%	86.74%	84.37%
3	81.38%	86.84%	84.12%	85.46%
4	86.38%	89.42%	89.68%	89.53%
5	85.58%	86.47%	92.29%	89.26%
6	82.48%	86.67%	86.28%	86.46%
7	85.57%	89.05%	88.77%	88.89%
Final (v3, v4, v5, v6, v7)	86.01%	88.66%	90.02%	89.31%

CHAPTER 6 CONCLUSION AND FUTURE WORK

6.1 Conclusion

In the first part of our research, we developed and studied two datasets: the first dataset is all the Bordeaux wine from 2000 to 2016; and the second one is all wines listed in a famous collection of Bordeaux wines, 1855 Bordeaux Wine Official Classification, from 2000 to 2016. We used a Naïve Bayes classifier and an SVM classifier to make wine quality predictions based on the wine reviews in both data sets. Overall, the Naïve Bayes classifier performed better than the SVM in the 1855 Bordeaux Wine Official Classification dataset, slightly worse than SVM in the ALL Bordeaux Wine dataset. Also, with the benefit of using Naïve Bayes classifier, we were able to find the important wine characteristics/attributes for the 21st century classic and outstanding Bordeaux wines. The list of common attributes in Table 5 and Table 8 identifies the general wine characteristics in the Bordeaux dataset, while the list of dominant attributes in Table 6 and Table 9 (Table 7 and Table 10) show the preferable characteristics for 90+ (90–) wines. Those characteristics/attributes can help producers improve the quality of their wines allowing them to concentrate on producing wine with positive characteristics and avoid producing wines with unwanted characteristics during the winemaking process.

In the second part of the research, we added the attributes in category and subcategory columns in the computational wine wheel into our original datasets. In the end, we have 1033 attributes, including 14 category attributes, 34 subcategory attributes and 985 normalized attributes, which added more information into our dataset. We designed a novel voting system that fully used the new information with the Naïve Bayes classifier and SVM classifier on the All Bordeaux wine dataset and the Bordeaux Wine

Official Classification in 1855 dataset. When using Naïve Bayes, the novel voting system had better performances on both datasets. The novel voting system with SVM classifier had better performances on the Bordeaux Wine Official Classification in the 1855 dataset but there was no significant improvement with the SVM classifier on the All Bordeaux wine dataset. Looking into the details of each voter with SVM classifier on the All Bordeaux wine dataset, the recall among some voters is as low as 53.81% even though they passed the 80% accuracy threshold. We might generate a better result if we put the thresholds on precision and recall as well. Overall, the novel voting system designed based on the newly added attributes had helped improve the model performances on Naïve Bayes classifier and, to some degree, the SVM classifier.

We would like to address the limitation of our current research. Since the computational wine wheel was developed from Wine Spectators' Top 100 lists, the proposed research might have optimal results in the dataset collected from Wine Spectators' review. While several other wine experts in the field such as Robert Parker Wine Advocate [52], Wine Enthusiast [53], and Decanter [54] may not agree with each other's comments, they can still agree in the overall score of the wine. The legendary Chateau Latour 2009 gives a great example [55]; every reviewer scores the wine either 100 or 99 and their testing notes are very different with each other. This would be our ultimate challenge in Wineinformatics research that involves the true human language processing topic.

6.2 Future Work

For future research, three follow up questions can be raised: 1. Instead of dichotomous (90+ and 90−) analysis, one could use a finer label (classic, outstanding,

very good, and good) to categorize these Bordeaux wines and perform the classification. 2. What characteristics/attributes make the Bordeaux wines be categorized as classic (95+) instead of outstanding (90–94)? 3. Apart from Naïve Bayes and SVM, are there any other classifiers that will perform well using the datasets?

The first question can be studied as a multi-class problem in data science since the computational model will be built into four different classes and produce important characteristics for each class. The second question is related to a fairly common highly unbalanced problem in data science. The number of wines scores 95+ is a lot less than 95− wines. The regular computational model such as SVM and Naïve Bayes will not be able to identify the boundary between the two classes and predict all tested wines as the majority class. How to take a balanced look at both classes rather than focusing on information gleaned from the majority class is a substantial challenge in this type of question. The third question is an eternal question in the field of data science. There are hundreds of classifiers available nowadays and testing each classifier on the dataset can be very time-consuming. Based on Fern´andez-Delgado's research, he applied 179 classifiers from 17 families on 121 datasets, and the random forest classifier stood. It had the highest accuracy among all classifiers [42]. It will be interesting to see how random forests would perform in this dataset, especially since the decision tree failed in our previous Wineinformatics study.

REFERENCES

[1] R. Caruana and A. Niculescu-Mizil, "An empirical comparison of supervised learning algorithms." in *Proceedings of the 23rd International Conference on Machine Learning*, ser. ICML '06, New York, NY, USA: ACM, pp. 161–168, 2006.

[2] T. Hastie, R. Tibshirani, and J. Friedman, *Unsupervised Learning*. New York, NY: Springer New York, pp. 485–585, 2009.

[3] J. Blakeman, "On tests for linearity of regression in frequency distributions." *Biometrika*, vol. 4, no. 3, pp. 332–350, 1905.

[4] E. Alpaydin, *Introduction to Machine Learning*. MIT Press, p. 9, ISBN 978-0-262-01243-0, 2010.

[5] U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, "From data mining to knowledge discovery in databases." *AI Magazine*, 17, 37–54, 1996.

[6] X.D. Wu, X.Q. Zhu, G.-Q. Wu, W. Ding, "Data mining with big data." *IEEE Transactions on Knowledge and Data Engineering*, 26, 97–107, 2014.

[7] S. Sumathi, "Introduction to data mining and its applications." *New Delhi: Springer*, 2009.

[8] P. Karlsson, "World wine production reaches record level in 2018, consumption is stable." *BKWine Magazine*, 2019.

[9] P. Cortez, A. Cerdeira, F. Almeida, T. Matos, J. Reis, "Modeling wine preferences by data mining from physicochemical properties." *Decision Support Systems*, 47, 547–553, 2009.

[10] A. Edelmann, J. Diewok, K.C. Schuster, B. Lendl, “Rapid method for the discrimination of red wine cultivars based on mid-infrared spectroscopy of phenolic wine extracts.” *Journal of Agricultural and Food Chemistry*, 49, 1139–1145, 2001.

[11] H. Johnson, “London: Octopus Publishing Group Ltd.” *World Atlas of Wine*, 4th edition, p. 13, 1994.

[12] History. Retrieved from https://www.bordeaux.com/us/Our-know-how/History

[13] J. Robinson, *The Oxford Companion to Wine*, Third Edition, pp. 175–177, Oxford University Press, 2006.

[14] Bordeaux Wine Official Classification of 1855. Available online: https://www.bordeaux.com/us/Our-Terroir/Classifications/Grand-Cru-Classes-en-1855.

[15] P. Combris, S. Lecocq, M. Visser, “Estimation of a hedonic price equation for Bordeaux wine: Does quality matter?” *Economic Journal*, 107, 389–402, 1997.

[16] J.M. Cardebat, J. Figuet, “What explains Bordeaux wine prices?” *Applied Economics Letters*, 11, 293–296, 2004.

[17] O. Ashenfelter, “Predicting the quality and prices of Bordeaux wine.” *Economic Journal*, 118, F174–F184, 2008.

[18] S. Shanmuganathan, P. Sallis, A. Narayanan, “Data mining techniques for modelling seasonal climate effects on grapevine yield and wine quality.” *In Proceedings of the 2010 2nd International Conference on Computational Intelligence, Communication Systems and Networks*, Liverpool, UK, 28–30, pp. 84–89, 2010.

[19] F.N. Noy, M. Sintek, M., S. Decker, M. Crubézy, R.W. Fergerson, M.A. Musen, "Creating semantic web contents with protege-2000." *IEEE Intelligent Systems,* 16, 60–71, 2001.

[20] F.N. Noy, D.L. McGuinness, "Ontology Development 101: A Guide to Creating Your First Ontology." *Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report* SMI-2001-0880, 2001.

[21] R.E. Quandt, "A note on a test for the sum of rank sums." *Journal of Wine Economics*, 2, 98–102, 2007.

[22] R.H. Ashton, "Improving experts' wine quality judgments: Two heads are better than one." *Journal of Wine Economics,* 6, 135–159, 2011.

[23] R.H. Ashton, "Reliability and consensus of experienced wine judges: Expertise within and between?" *Journal of Wine Economics*, 7, 70–87, 2012.

[24] J.C. Bodington, "Evaluating wine-tasting results and randomness with a mixture of rank preference models." *Journal of Wine Economics,* 10, 31–46, 2015.

[25] Wine Spectator. Available online: https://www.winespectator.com

[26] B. Chen, C. Rhodes, A. Crawford, L. Hambuchen, "Wineinformatics: Applying data mining on wine sensory reviews processed by the computational wine wheel." *In Proceedings of the 2014 IEEE International Conference on Data Mining Workshop,* Shenzhen, China, 14–14. pp. 142–149, 2014.

[27] B. Chen, C. Rhodes, A. Yu, V. Velchev, "The Computational Wine Wheel 2.0 and the TriMax Triclustering in Wineinformatics." *In Industrial Conference on Data Mining, Springer,* Cham, Switzerland, pp. 223–238, 2016.

[28] B. Chen, V. Velchev, J. Palmer, T. Atkison, “Wineinformatics: A Quantitative Analysis of Wine Reviewers.” *Fermentation* ,4, 82, 2018.

[29] J. Palmer, B. Chen, “Wineinformatics: Regression on the Grade and Price of Wines through Their Sensory Attributes.” *Fermentation*, 4, 84, 2018.

[30] J.M. Cardebat, F. Livat, “Wine experts’ rating: A matter of taste?” *International Journal of Wine Business Research*, 28, 43–58, 2016.

[31] J.M. Cardebat, J.M. Figuet, E. Paroissien, “Expert opinion and Bordeaux wine prices: An attempt to correct biases in subjective judgments.” *Journal of Wine Economics*, 9, 282–303, 2014

[32] J. Cao, L. Stokes, “Evaluation of wine judge performance through three characteristics: Bias, discrimination, and variation.” *Journal of Wine Economics*, 5, 132–142, 2010.

[33] J.M. Cardebat, E. Paroissien, “Standardizing expert wine scores: An application for Bordeaux en primeur.” *Journal of Wine Economics*, 10, 329–348, 2015.

[34] R.T. Hodgson, “An examination of judge reliability at a major US wine competition.” *Journal of Wine Economics*, 3, 105–113, 2008.

[35] R.T. Hodgson, “An analysis of the concordance among 13 US wine competitions.” *Journal of Wine Economics*, 4, 1–9, 2009.

[36] R. Hodgson, J. Cao, “Criteria for accrediting expert wine judges.” *Journal of Wine Economics*, 9, 62–74, 2014.

[37] H. Hopfer, H. Heymann, “Judging wine quality: Do we need experts, consumers or trained panelists?” *Food Quality and Preference*, 32, 221–233, 2014.

[38] O. Ashenfelter, R. Goldstein, C. Riddell, “Do expert ratings measure quality? The case of restaurant wine lists.” *In Proceedings of the 4th Annual AAWE Conference* at *the University of California at Davis*, Davis, CA, USA, 2010.

[39] J.M. Cardebat, P. Corsinovi, D. Gaeta, “Do Top 100 wine lists provide consumers with better information?” *Economics Bulletin*, 38, 983–994, 2018.

[40] J. Reuter, “Does advertising bias product reviews? An analysis of wine ratings.” *Journal of Wine Economics*, 4, 125–151, 2009.

[41] Web Scraping: The Comprehensive Guide for 2020. Retrieved from https://prowebscraper.com/blog/what-is-web-scraping/

[42] M. Fern´andez-Delgado, E. Cernadas, S. Barro, D. Amorim, “Do we Need Hundreds of Classifiers to Solve Real World Classification Problems?” *Journal of Machine Learning Research*, 15, 3133-3181, 2014.

[43] V. Metsis, I. Androutsopoulos, G. Paliouras, “Spam Filtering with Naive Bayes - Which Naive Bayes?” In *CEAS*, 2018.

[44] K.J.A., Suykens, J. Vandewalle, “Least squares support vector machine classifiers.” *Neural Processing Letters*, 293–300, 1999.

[45] O Favorov, J Macdonald, O Kursun, "SVM-Based Analysis of NMR Spectra in Metabolomics: Development of Procedures." *The Journal of Science and Medicine*, 1(2), 2019.

[46] J. Thorsten, Svmlight: Support Vector Machine. Available online: https://www.researchgate.net/profile/Thorsten_Joachims/publication/243763293_SVMLight_Support_Vector_Machine/links/5b0eb5c2a6fdcc80995ac3d5/SVMLight-Support-Vector-Machine.pdf

[47] E. Leopold, J. Kindermann, “Text categorization with support vector machines how to represent texts in input space?” *Machine Learning*, 46 (1–3), 423–444, 2002

[48] T. Joachims, “Text categorization with Support Vector Machines: Learning with many relevant features.” *Machine Learning: ECML-98 Lecture Notes in Computer Science*, 137–142, 1998.

[49] W. li, Z, Liu, “A method of SVM with Normalization in Intrusion Detection.” *Procedia Environmental Sciences, 11*, 256–262, 2011.

[50] L. Hansen, P. Salamon “Neural network ensembles.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:993–1001, 1990.

[51] C.O. Sakar, O. Kursun, F. Gurgen, "Ensemble canonical correlation analysis." *Applied Intelligence*, 40 (2), 291-304, 2014.

[52] Robert Parker Wine Advocate. Available online: https://www.robertparker.com/

[53] Wine Enthusiast. Available online: https://www.wineenthusiast.com/

[54] Decanter. Available online: https://www.decanter.com/

[55] Chateau Latour 2009 Wine Reviews. Available online: https://www.wine.com/product/chateau-latour-2009/119875

APPENDIX A. THE 1855 CLASSIFICATION, REVISED IN 1973

APPENDIX A.1. RED WINES

PREMIERS CRUS

Château Haut-Brion, Pessac, AOC Pessac-Léognan

Château Lafite-Rothschild, Pauillac, AOC Pauillac

Château Latour, Pauillac, AOC Pauillac

Château Margaux, Margaux, AOC Margaux

Château Mouton Rothschild, Pauillac, AOC Pauillac

DEUXIÈMES CRUS

Château Brane-Cantenac, Cantenac, AOC Margaux

Château Cos-d'Estournel, Saint-Estèphe, AOC Saint-Estèphe

Château Ducru-Beaucaillou, Saint-Julien-Beychevelle, AOC Saint-Julien

Château Durfort-Vivens, Margaux, AOC Margaux

Château Gruaud-Larose, Saint-Julien-Beychevelle, AOC Saint-Julien

Château Lascombes, Margaux, AOC Margaux

Château Léoville-Barton, Saint-Julien-Beychevelle, AOC Saint-Julien

Château Léoville-Las-Cases, Saint-Julien-Beychevelle, AOC Saint-Julien

Château Léoville-Poyferré, Saint-Julien-Beychevelle, AOC Saint-Julien

Château Montrose, Saint-Estèphe, AOC Saint-Estèphe

Château Pichon-Longueville-Baron-de-Pichon, Pauillac, AOC Pauillac

Château Pichon-Longueville-Comtesse-de-Lalande, Pauillac, AOC Pauillac

Château Rauzan-Ségla, Margaux, AOC Margaux

Château Rauzan-Gassies, Margaux, AOC Margaux

TROISIÈMES CRUS

Château Boyd-Cantenac, Cantenac, AOC Margaux

Château Calon-Ségur, Saint-Estèphe, AOC Saint-Estèphe

Château Cantenac-Brown, Cantenac, AOC Margaux

Château Desmirail, Margaux, AOC Margaux

Château Ferrière, Margaux, AOC Margaux

Château Giscours, Labarde, AOC Margaux

Château d'Issan, Cantenac, AOC Margaux

Château Kirwan, Cantenac, AOC Margaux

Château Lagrange, Saint-Julien-Beychevelle, AOC Saint-Julien

Château La Lagune, Ludon, AOC Haut-Médoc

Château Langoa-Barton, Saint-Julien-Beychevelle, AOC Saint-Julien

Château Malescot-Saint-Exupéry, Margaux, AOC Margaux

Château Marquis-d'Alesme, Margaux, AOC Margaux

Château Palmer, Cantenac, AOC Margaux

QUATRIÈMES CRUS

Château Beychevelle, Saint-Julien-Beychevelle, AOC Saint-Julien

Château Branaire-Ducru, Saint-Julien-Beychevelle, AOC Saint-Julien

Château Duhart-Milon, Pauillac, AOC Pauillac

Château Lafon-Rochet, Saint-Estèphe, AOC Saint-Estèphe

Château Marquis-de-Terme, Margaux, AOC Margaux

Château Pouget, Cantenac, AOC Margaux

Château Prieuré-Lichine, Cantenac, AOC Margaux

Château Saint-Pierre, Saint-Julien-Beychevelle, AOC Saint-Julien

Château Talbot, Saint-Julien-Beychevelle, AOC Saint-Julien

Château La Tour-Carnet, Saint-Laurent-de-Médoc, AOC Haut-Médoc

CINQUIÈMES CRUS

Château d'Armailhac, Pauillac, AOC Pauillac

Château Batailley, Pauillac, AOC Pauillac

Château Belgrave, Saint-Laurent-de-Médoc, AOC Haut-Médoc

Château Camensac, Saint-Laurent-de-Médoc, AOC Haut-Médoc

Château Cantemerle, Macau, AOC Haut-Médoc

Château Clerc-Milon, Pauillac, AOC Pauillac

Château Cos-Labory, Saint-Estèphe, AOC Saint-Estèphe

Château Croizet-Bages, Pauillac, AOC Pauillac

Château Dauzac, Labarde, AOC Margaux

Château Grand-Puy-Ducasse, Pauillac, AOC Pauillac

Château Grand-Puy-Lacoste, Pauillac, AOC Pauillac

Château Haut-Bages-Libéral, Pauillac, AOC Pauillac

Château Haut-Batailley, Pauillac, AOC Pauillac

Château Lynch-Bages, Pauillac, AOC Pauillac

Château Lynch-Moussas, Pauillac, AOC Pauillac

Château Pédesclaux, Pauillac, AOC Pauillac

Château Pontet-Canet, Pauillac, AOC Pauillac

Château du Tertre, Arsac, AOC Margaux

APPENDIX A.2. WHITE WINES

PREMIER CRU SUPÉRIEUR

Château d'Yquem, Sauternes, AOC Sauternes

PREMIERS CRUS

Château Climens, Barsac, AOC Barsac

Clos Haut-Peyraguey, Bommes, AOC Sauternes

Château Coutet, Barsac, AOC Barsac

Château Guiraud, Sauternes, AOC Sauternes

Château Lafaurie-Peyraguey, Bommes, AOC Sauternes

Château Rabaud-Promis, Bommes, AOC Sauternes

Château Rayne-Vigneau, Bommes, AOC Sauternes

Château Rieussec, Fargues-de-Langon, AOC Sauternes

Château Sigalas-Rabaud, Bommes, AOC Sauternes

Château Suduiraut, Preignac, AOC Sauternes

Château La Tour-Blanche, Bommes, AOC Sauternes

DEUXIÈMES CRUS

Château d'Arche, Sauternes, AOC Sauternes

Château Broustet, Barsac, AOC Barsac

Château Caillou, Barsac, AOC Barsac

Château Doisy-Daëne, Barsac, AOC Barsac

Château Doisy-Dubroca, Barsac, AOC Barsac

Château Doisy-Védrines, Barsac, AOC Barsac

Château Filhot, Sauternes, AOC Sauternes

Château Lamothe (Despujols), Sauternes, AOC Sauternes

Château Lamothe-Guignard, Sauternes, AOC Sauternes

Château de Malle, Preignac, AOC Sauternes

Château de Myrat, Barsac, AOC Barsac

Château Nairac, Barsac, AOC Barsac

Château Romer-du-Hayot, Fargues-de-Langon, AOC Sauternes

Château Romer, Fargues-de-Langon, AOC Sauternes

Château Suau, Barsac, AOC Barsac

APPENDIX B. THE LIST OF WINE AND VINTAGES WE CANNOT FIND

CHÂTEAU PÉDESCLAUX Pauillac (2005,2004,2003,2002,2001)

CHÂTEAU CLIMENS Barsac (2000)

CHÂTEAU RABAUD-PROMIS Sauternes (2016,2015,2014,2010,2008)

CHÂTEAU RIEUSSEC Sauternes (2012)

CHÂTEAU SUDUIRAUT Sauternes (2012)

CHÂTEAU LA TOUR BLANCHE Sauternes (2000)

CHÂTEAU BROUSTET Barsac (2012,2008,2007,2005,2004,2000)

CHÂTEAU CAILLOU Barsac(2016,2015,2014,2010,2008,2000)

CHÂTEAU LAMOTHE-DESPUJOLS Sauternes

(2016,2015,2014,2013,2012,2011,2010,2009,2006,2005,2004,2002,2000)

CHÂTEAU NAIRAC Barsac (2016,2000)

CHÂTEAU ROMER DU HAYOT Sauternes (2016,2015,2014,2010)

CHÂTEAU ROMER Sauternes (2016,2010,2008,2006,2004,2002,2001,2000)

CHÂTEAU SUAU Barsac (2014,2010,2007)

CHÂTEAU D'YQUEM Sauternes (2012)

CHÂTEAU D'ARCHE Sauternes (2016,2015,2014,2012,2010)

Château Durfort-Vivens Margaux (2016,2015,2014)

Château Pichon-Longueville-Baron-de-Pichon, Pauillac, AOC Pauillac (Château Pichon-Longueville Baron Pauillac Les Griffons de Pichon Baron

(2016,2015,2013,2011,2010,2009,2008,2007,2006,2005,2004,2003,2002,2001,2000))

Château Pichon-Longueville-Comtesse-de-Lalande, Pauillac, AOC Pauillac (Château Pichon Longueville Lalande Pauillac Réserve de la Comtesse (2013,2007))

Château Rauzan-Gassies Margaux (2007,2004)

Château Boyd-Cantenac Margaux (2016,2015,2014,2013,2012)

Château Desmirail Margaux (2007,2006,2005,2004,2003,2002,2001,2000)

CHÂTEAU MARQUIS D'ALESME BECKER Margaux (2004)

CHÂTEAU BEYCHEVELLE St.-Julien Amiral de Beychevelle

(2013,2011,2004,2003,2002,2001)

CHÂTEAU MARQUIS DE TERME Margaux (2003)

CHÂTEAU POUGET Margaux (2016,2015,2014,2013,2012)

CHÂTEAU DE CAMENSAC Haut-Médoc (2016,2015,2014,2008)

Château La Lagune Haut-Médoc (2016,2015,2013)

CHÂTEAU COS LABORY St.-Estèphe (2016,2015,2014,2013)

CHÂTEAU CROIZET-BAGES Pauillac (2007)

Château d'Issan, Cantenac, AOC Margaux (not Found)

Château Doisy-Dubroca, Barsac (not found)

Château Lamothe-Guignard, Sauternes (2016)