Heaven, Earth, and Man:

**How Weather Changes Affect the Flavor and Quality of Wine**

The Impact of Climate Changes on Wine Quality and Flavor

ABSTRACT

The quality of wine is commonly attributed to three key factors: heaven, earth, and man. Heaven refers to climatic conditions, including sunlight, temperature, and rainfall—environmental elements that play a crucial role in grape cultivation. Earth encompasses the type of soil, topography, and geographic location where the grapes are grown, all of which profoundly affect grape maturity and flavor. Man represents the winemaker’s skill, decisions, and expertise throughout the winemaking process, from harvest to fermentation and aging. Together, these three factors shape the overall quality and characteristics of wine.  
  
This study aims to explore how weather variations in grape-growing regions influence the quality of wine and to develop models that predict wine quality. The research focuses on how climatic factors, such as temperature, precipitation, and sunlight duration, impact grape maturity and flavor during different growing seasons, ultimately affecting the wine’s quality. Through the application of data analysis and machine learning techniques, predictive models will be created to estimate wine quality with precision. The study aims to provide practical insights that can help grape growers and winemakers optimize their cultivation and production strategies to enhance the overall quality of wine.

INTRODUCTION

The quality and flavor of wine are influenced by various factors, which are traditionally summarized as heaven (天), earth (地), and man (人) in winemaking theory. Heaven refers to climatic conditions such as sunlight, temperature, and rainfall—natural elements that play a crucial role in shaping grape growth. Earth refers to the soil, topography, and geographical location where the grapes are cultivated, which impact the grapes' flavor and ripeness. Man refers to the winemaker’s techniques and expertise, influencing every step from grape cultivation to the winemaking process and ultimately determining the final quality of the wine.  
  
Among these factors, climate change has a particularly significant effect on wine. As global climate patterns shift—with rising temperatures and altered rainfall patterns—these weather changes directly impact the grape growing cycle, sugar content, and acidity balance, shaping the overall flavor profile of the wine. In recent years, more and more studies have aimed to analyze weather data to predict wine quality, helping winemakers adapt to climate changes and improve their winemaking techniques.  
  
This study seeks to investigate how climate variations in grape-growing regions affect wine quality by observing and analyzing weather patterns. Additionally, the study will develop predictive models to provide insights for future grape cultivation and winemaking strategies. By integrating the traditional winemaking philosophy of heaven, earth, and man, the research aims to identify the specific connections between climate change and wine flavor, with the goal of applying these findings to modern winemaking practices to improve overall wine quality and competitiveness in the market.

RELATED WORK

In the study of wine quality and flavor, existing literature can be broadly categorized into two main directions: one focuses on analyzing wine reviews and ratings, and the other on analyzing the composition of wines to predict quality.  
  
**1. Research Based on Reviews and Ratings**

The first category of research focuses on collecting consumer evaluations of wine and using these data to build machine learning models to predict wine ratings. For example, one study scraped data from Vivino.com, creating a dataset that includes various types of wine (red, white, rosé, and sparkling), with features such as the number of ratings, rating scores, and wine type. The researchers found that for certain wines, the limited number of ratings led to a lack of representativeness, making it difficult to reflect their true quality. Based on this observation, they developed machine learning models to predict ratings for wines with fewer reviews but higher potential. This highlights the importance of consumer feedback in the wine market and indicates that new products often suffer from a lack of attention due to insufficient reviews, which can result in potential sales losses. Predicting wine ratings not only helps consumers make better purchasing decisions but also aids producers in increasing the visibility and market demand for new products.  
  
**2. Research Based on Ingredient Analysis**

The second category of research focuses on the physical and chemical composition of wine to assess its quality and flavor. These studies typically analyze the components of wine, such as alcohol content (ABV), grape varieties, flavor characteristics, vintage, and region, to predict its final quality. One study’s dataset included wines from around the world, documenting details such as wine name, description, price, grape variety, cork type, alcohol content, region, style, and flavor characteristics. The goal of this research was to build predictive models that could evaluate the flavor, aroma, and market value of wine based on these features. By analyzing the composition of wine, these studies provide more accurate predictions about its quality and market value, which is crucial for wine producers in terms of market positioning and pricing strategies.  
  
In conclusion, whether based on reviews and ratings or ingredient analysis, both directions of research aim to use data-driven approaches to predict and enhance wine quality and market performance. However, most of these studies focus on evaluating the outcomes of the wine itself, while there has been less exploration into external natural conditions, such as weather changes, that affect grape growth. This study aims to fill that gap by analyzing how climate changes in grape-producing regions impact wine flavor and quality and by building predictive models based on these observations.

PROPOSED WORK

This study aims to explore how weather changes affect the flavor and quality of wine by integrating wine rating data with meteorological data. Specifically, the research will utilize two datasets containing wines from different regions (red and white wines) and, based on the geographic locations of each region, retrieve historical weather data for relevant analysis. The detailed steps of the study are as follows:  
  
**1. Dataset Integration**

The study will begin by cleaning and integrating the two datasets to ensure data consistency and usability. Each dataset includes key variables such as wine name, country, region, winery name, rating, number of ratings, price, and year. These variables will be used for subsequent statistical analysis and model training.  
  
**2. Geographical Coordinates Retrieval**

To match wine regions with meteorological data, the study will use the region names provided in the datasets and employ geocoding techniques to retrieve the latitude and longitude coordinates of each region.  
  
**3. Weather Station Matching and Data Retrieval**

Next, using the latitude and longitude of the wine regions, the study will gather historical weather data from relevant weather stations. The specific weather variables to be collected include:

* Daily average temperature
* Rainfall
* Daily sunlight hours

These weather variables will correspond to the growing seasons of the grapes, and their impact on wine quality will be further analyzed.  
  
**4. Data Analysis and Model Construction**

After integrating the data, the study will use statistical analysis and machine learning models to explore the relationship between weather variables and wine ratings. The specific analysis steps include:

* Correlation Analysis: Investigate the correlation between different weather variables (e.g., temperature, rainfall) and wine ratings.
* Regression Models: Build multiple regression models to analyze the predictive power of weather variables on wine ratings.
* Machine Learning Models: Use Random Forest machine learning methods to predict wine quality based on weather conditions.

**5. Result Validation and Discussion**

Finally, the study will validate and discuss the prediction results of the models, analyzing the specific impact of weather conditions on wine quality. Based on these findings, recommendations will be made for adjusting winemaking and grape cultivation techniques to improve wine quality in response to climate changes.

The results of this study will assist winemakers and grape growers in understanding how weather changes affect wine quality, providing a scientific basis for adapting to future climate variations.

EVALUATION AND VALIDATION

****1. Lasso Regression Results****

During the Lasso regression analysis, we selected the seasonal weather data for different categories of wine for model training. The results indicated that the key features were mainly concentrated on sunlight duration (tsun), including Spring\_tsun, Summer\_tsun, and Fall\_tsun. The performance of the Lasso model in predicting different types of wine was as follows:

* Red wine: Final MSE = 0.0903.
* White wine: Final MSE = 0.0615.
* Rosé wine: Final MSE = 0.0805.
* Sparkling wine: Final MSE = 0.0711.

From the Lasso regression results, we found that sunlight duration (tsun) has a significant impact on wine ratings. This is likely because sunlight duration directly affects the ripening and flavor of grapes, which in turn influences the quality ratings of the wine.

**2. Multicollinearity (VIF)**

Based on the Variance Inflation Factor (VIF) results, weather variables from different seasons, such as Spring\_tavg and Summer\_tavg, exhibited very high VIF values. This suggests a strong linear correlation between these variables, likely due to the seasonal similarities in weather data, particularly among temperature and precipitation features in the same season.

To address this issue, we employed PCA (Principal Component Analysis) to reduce dimensionality and minimize feature similarity, and we used Lasso regression for feature selection. By applying these methods, we successfully mitigated the effects of multicollinearity while retaining the weather variables that had the greatest impact on predicting wine ratings.

**3. Wine Rating Prediction Model**

In the wine rating prediction project, three different models were employed to assess the impact of weather on wine ratings: Principal Component Analysis (PCA), Random Forest regression, and a Deep Learning model. Each method was validated using cross-validation to ensure the robustness of the models.

**3.1 Performance Metrics**

The main metric used to evaluate model performance was the Mean Squared Error (MSE). The MSE values for the four types of wine were as follows:

* **PCA model:** 0.067/0.056/0.079/0.055,
* **Random Forest model:** 0.068/0.055/0.091/0.049
* **Deep Learning model:** 0.040/0.057/0.287/0.285.

Comparatively, the **Random Forest** and **PCA** models demonstrated superior performance.

**3.2 Efficiency Evaluation**

The efficiency of the three models varied when processing large datasets. **Random Forest** and **PCA** were computationally faster, while the Deep Learning model required more computational resources and time. Although the Deep Learning model can capture more complex non-linear relationships, its performance in this project was not as prominent, likely due to the strong linear relationships between the data features.

**3.3 Cross-validation**

A 5-fold cross-validation was used to validate the performance of the models and avoid overfitting. The cross-validation MSE for the **Random Forest model** was 0.072/0.061/0.067/0.051, while the **PCA model** had a cross-validation MSE of 0.071/0.056/0.079/0.055, both demonstrating robust model performance.

**DATA VISUALIZATION**

**1. Sunshine Duration vs. Ratings**

* The sunshine duration in spring and summer shows that longer sunshine hours (over 30,000 seconds) are positively correlated with higher ratings (4.4 to 4.8).
* In autumn, the impact of different sunshine durations on ratings appears relatively balanced, with no obvious trends, although red wine tends to have slightly higher ratings under longer sunshine.

**2. Mean Temperature vs. Ratings**

* In spring and summer, the correlation between temperature and ratings is not easily discernible, with temperatures ranging from 5°C to 25°C.
* The impact of autumn temperatures on the ratings of various wines seems more uniform, with most ratings concentrated between 4.2 and 4.6.

**3. Precipitation vs. Ratings**

* During spring and summer, lower precipitation levels (about 0 to 5mm) result in slightly higher ratings for most types of wine (especially red wine), but once precipitation exceeds a certain amount (>5mm), the rating trend becomes flatter.
* The changes in precipitation during autumn do not show significant impacts on ratings, but when precipitation is below 5mm, the overall ratings seem to be slightly higher, particularly affecting red wine.

**TIMELINE AND CHALLENGES**  
  
**Timeline Setup**

* Data Organization and Preprocessing: 2 weeks
* Model Selection and Adjustment: 2 weeks
* Cross-Validation and Result Analysis: 1 week
* Report Writing and Submission: 1 week

**Estimated Challenges**

**1. Data Feature Engineering:**

There is a high linear correlation among weather data across different seasons, which can lead to multicollinearity issues, affecting the accuracy and interpretability of the model. To address this, we utilized PCA dimensionality reduction techniques combined with Lasso for feature selection, filtering out the weather features most correlated with wine ratings.

**2. Acquisition of Historical Weather Data:**

Due to limitations in weather station locations and data storage, we may face issues with incomplete or unavailable historical weather data. This will directly impact the accuracy of data analysis and the reliability of the model. To tackle this challenge, the project may consider using multiple sources of weather data or relying on machine learning imputation techniques to fill in missing data.

**CONCLUSION**

This project successfully constructed several models to predict wine ratings through the analysis of weather data. After comparison, the **Random Forest Regression** and **Principal Component Analysis (PCA)** models achieved a good balance between performance and computational resources, showing lower Mean Squared Error (MSE) values.  
  
In the correlation analysis, it was found that no single feature has a high correlation with the ratings; only the region of Sparkling Wine shows a higher correlation with ratings (0.42). Visual observation of the relationship between ratings and region reveals that the Champagne region's ratings are significantly higher than those of other regions, leading to a high correlation.  
  
In the feature selection for weather characteristics, three regression analysis methods were compared comprehensively, with the following results: **Lasso Regression**

* Sunshine duration (tsun) has a significant influence on wine ratings.

**Principal Component Analysis (PCA)**

* The temperatures (tavg/tmin/tmax) in spring and autumn have a high influence, explaining about 35% of the variance.
* Summer temperatures (tavg/tmin/tmax) and sunshine duration (tsun) have the next highest influence, explaining about 26% of the variance.

**Random Forest Regression**

* Summer and autumn sunshine duration (tsun) significantly affects the ratings of red and white wines.
* Spring sunshine duration (tsun) and autumn precipitation (prcp) have a larger impact on rosé wines.
* Summer precipitation (prcp) has the greatest effect on sparkling wines, holding the highest weight.
* Temperature (tavg/tmin/tmax) has a lesser influence on all types of wines compared to sunshine and precipitation.

**Summary**

* Sunshine duration has a consistent and significant influence on the ratings of all types of wines. Moderate sunshine (8-11 hours) is needed in spring, high sunshine (10-13 hours) in summer, and low sunshine (6-9 hours) in autumn.
* The impact of temperature on different wine types is not obvious, but temperatures in spring and autumn (5-20°C) significantly influence wine quality, while summer temperatures also have some effect on ratings.
* Precipitation shows a strong correlation with specific types of wine ratings, with sparkling wines and rosé wines requiring low precipitation (0-2mm) in summer and moderate precipitation (0.5-4mm) in autumn.

****FUTURE WORK****

**1. Analyzing Climate Impact on Grape Growth**

Climate factors such as temperature, precipitation, and sunlight duration directly affect grape ripeness, sugar content, and acidity, all of which influence the flavor and quality of wine. The study should first establish a relationship model between climate change and grape quality, similar to the PCA analysis and regression models already adopted, to identify key climatic variables (such as sunlight duration and precipitation in spring, summer, and autumn) related to wine quality.

**2. Climate Prediction and Vineyard Management** Through meteorological data and long-term climate forecasts, vineyards can take proactive measures to adapt to climate changes. Specific methods include:

* Adjusting planting areas: Selecting grape varieties that adapt to current climate conditions or relocating vineyards to more suitable areas. For example, climate warming may favor moving vineyards to higher altitudes or cooler regions.
* Adjusting planting schedules: Modifying the grape growth cycle based on climate changes, such as planting earlier or later to avoid ripening under extreme weather conditions (e.g., excessive summer heat).

**3. Adjusting Winemaking Techniques**

The winemaking process can also be adjusted based on changes in grape characteristics:

* Fermentation temperature control: If excessive sugar content leads to high alcohol levels, reducing fermentation temperature can help minimize the risk of over-fermentation, maintaining the wine's balance.
* Acidity adjustment: As temperatures rise, grape acidity may decrease. Techniques such as adding acidity supplements can be used to preserve the flavor and mouthfeel of the wine.
* Oak barrel aging time: Adjusting the aging time of wine in oak barrels based on grape ripeness and climatic influences to optimize the wine's aroma and structure.

**4. Application of New Technologies and Methods**

**Through da**ta analysis, vineyard owners and winemakers can adopt new technologies to address the challenges posed by climate change:

* Smart agriculture technology: Using sensors, drones, and artificial intelligence to monitor weather and soil conditions, enabling precise vineyard management and real-time adjustments to water and nutrient supply.
* Variety breeding: Selecting and breeding grape varieties that are heat-tolerant, drought-resistant, or high-acid to cope with the challenges brought by climate change.

**5. Long-Term Data Monitoring and Adjustment Strategies**

Continuously collecting and analyzing meteorological data and wine quality ratings, regularly adjusting planting and winemaking strategies. As climate change intensifies, the accumulation of long-term data and continuous improvement are crucial for enhancing wine quality.

**6. Collaboration with Scientific Research Institutions**

Finally, collaboration with research institutions in agriculture, meteorology, and brewing science can provide access to the latest technologies and research findings, utilizing climate models and analysis techniques to support vineyard management and winemaking practices.

Through these methods, the research can help formulate specific technical and management measures to address the challenges posed by climate change on grape growth and winemaking processes, ultimately enhancing wine quality.

Conference Name:ACM Woodstock conference

Conference Short Name:WOODSTOCK’18

Conference Location:El Paso, Texas USA

ISBN:978-1-4503-0000-0/18/06

Year:2018

Date:June

Copyright Year:2018

Copyright Statement:rightsretained

DOI:10.1145/1234567890

RRH: F. Surname et al.

Price:$15.00