##### **NIT6150**

##### **Advanced Project**

##### Project Proposal

##### <Next-Generation Wine Recommendations: Integrating AI and Machine Learning>

**Team Leader:** Dr. Wenjie Ye (wenjie.ye@vu.edu.au)

**Team Members:** Julia Xu(s8139499)

Kai Xing (s8143230)

Linyang Wang (s8109316)

Sihua Tang(s8107067)

**Client:** Dr. Rachid Hamadi   
VU Sydney

**Supervisor:** Dr. Wenjie Ye

VU Melbourne

**Coordinator:** Dr. Quan Vu

College of Engineering and Science

Victoria University, Melbourne

Table of Contents

NIT6150 0

Advanced Project 0

Project Proposal 0

<> 0

1. Introduction 2
2. Background and Client Profile 2
3. Purpose and Objectives 3
4. Scope and Exclusion 3
5. Assumptions and Constraints 4
6. Deliverables 4
7. Schedule 5
8. Budget 5
9. Resources Roles and Responsibility 6
10. References 7
11. Introduction

The wine industry has seen a significant shift towards digital transformation, with consumers increasingly seeking personalised experiences in their wine selections. Traditionally, wine recommendations have been provided by professionals such as sommeliers or wine dealers, relying on their expertise to guide consumers. While these traditional methods are valuable, they can be limited by availability and accessibility. This project aims to design an AI model that analyses wine data and characteristics to provide personalised wine recommendations to users. The model will be trained using a Wine Reviews database from Kaggle which contains 130,000 wine reviews. Wine is examined from several aspects, including country, description of flavour, WineEnthusiast rating, price, and type of grape.

By leveraging advanced data analytics and machine learning techniques, the software will offer tailored suggestions, helping users discover wines that align with their tastes and preferences. This digital approach not only simplifies the selection process but also provides simple access to expert-level wine recommendations, allowing consumers to make informed choices from the comfort of their homes, thereby enhancing their overall wine experience.

1. Background and Client Profile

The Australia wine market is highly competitive, with consumers often overwhelmed by the large variety of available wines. Traditional methods of recommendation, such as internet research, expert reviews and social influences, are unable to balance the cost of energy, time, and resources versus the quality of advice and level of personalisation. Consumers today wish for more personalised interactions, and the one-size-fits-all approach no longer suffices.This gap presents an opportunity to apply AI technology in data analysis and machine learning to deliver highly personalised wine recommendations that go beyond generic advice, catering to the preferences of each user.

The client is a leading figure in the wine retail and enthusiast community, seeking to improve user engagement and satisfaction through innovative technological solutions. The client’s platform caters to a diverse audience, including casual wine drinkers, serious connoisseurs, and industry professionals. With a vast catalog of wines and a dynamic user base, the client recognises the need for a personalised recommendation system that can navigate the complexity of wine preferences and offer precise suggestions.

The client's platform is renowned for its extensive selection of wines from various regions, catering to different price ranges and taste profiles. The client has successfully built a strong reputation by offering high-quality products and exceptional customer service. However, as the platform grows, ability to meet unique preferences of individual customers becomes a complex issue.

With this project’s design, the team wishes to aid the client in enhancing customer satisfaction and loyalty, driving repeat business and strengthening their position in the market. The personalised wine recommendation system will not only simplify the decision-making process for consumers but also introduce them to new wines that they might not have discovered otherwise, enriching their overall experience and deepening their connection with the platform.

1. Purpose and Objectives

**Purpose:** This project aims to develop an intelligent wine recommendation system that accurately suggests wines based on user preferences using the Kaggle Wine Reviews dataset. The objectives include applying Natural Language Processing (NLP) techniques to analyze and clean wine descriptions and using heatmap analysis and Principal Component Analysis (PCA) to understand data distribution and key features. The project will design a deep learning model using Recurrent Neural Networks (RNN) or Convolutional Neural Networks (CNN) to extract valuable information from wine descriptions. Additionally, Random Forest algorithms will be used to combine text features with numerical data like ratings and price.

**Objectives:** The model will be trained using the cleaned dataset, with hyperparameters optimized through cross-validation and grid search techniques. A user-friendly interface will be developed to allow users to input their preferences and receive personalized recommendations. The model’s performance will be evaluated using metrics such as confusion matrix, accuracy, recall, and F1 score, with ongoing optimization based on these results. User testing and feedback will be integral to refining the system.

1. Scope and Exclusion

**Scope:** The project includes extracting and cleaning the Kaggle Wine Reviews dataset, designing and developing a machine learning model using NLP and deep learning algorithms (RNN/CNN), and integrating Random Forest for combined analysis of text and numerical features. The model will be trained and optimized through cross-validation and grid search. A user-friendly interface will be developed for inputting preferences and receiving recommendations. Comprehensive testing, user feedback collection, and system refinement are also part of the scope.

**Exclusion:** The project excludes the actual sale of wine and transaction handling, recommendations for beverages other than wine, and legal responsibilities beyond ensuring data privacy and security compliance.

1. Assumptions and Constraints

**Assumptions**: The project assumes that the Kaggle Wine Reviews dataset is accurate, complete, and sufficient for training the model. It is also assumed that users will provide genuine feedback to help improve the recommendation system. Additionally, it is assumed that the client has the necessary technical infrastructure to support the deployment and operation of the AI model and application. The project also assumes that data analysis and machine learning techniques will effectively process and analyze wine descriptions to generate valuable recommendations.

**Constraints**: The project is constrained by a predefined budget and timeline, which may limit the extent of model refinement and user testing. The accuracy of the recommendations may be influenced by the diversity and depth of the dataset. Ensuring the privacy and security of user data is paramount, which may introduce additional constraints on data handling and storage. These constraints must be managed to ensure the project stays on track and meets its objectives.

1. Deliverables

The project will yield several key deliverables to showcase the work on the red wine recommendation system. The project will start with the creation of a Project Allocation Form to clearly outline each team member's roles and responsibilities. The comprehensive Project Proposal will set the stage by detailing the objectives, methodology, and anticipated outcomes. As the project progresses, the team will develop thorough System Analysis and Design Documentation, capturing all system requirements, data flows, algorithm designs, and database structures. To ensure quality, the project will include a Project Evaluation Report, which will include extensive software testing results. For the end-users, the team will craft a detailed User Manual covering all aspects of system use. To visually represent the work, the project will include an Electronic Poster suitable for academic presentations. Finally, the complete red wine recommendation system will be packaged, including all source code, executables, and necessary dependencies.

1. Schedule

The project is divided into four main phases. The first phase, ending on August 7, 2024, covers project initiation and proposal writing. System analysis and design follows, concluding on August 18. The longest phase, from August 19 to September 15, focuses on development, testing, and documentation. The final phase, ending September 16, is for submission preparation and presentation refinement. Regular meetings are scheduled throughout: weekly team meetings on Mondays, bi-weekly progress reviews on alternate Fridays, and weekly supervisor meetings via Zoom on Wednesdays. This structure allocates appropriate time for each project component while maintaining consistent progress monitoring.

1. Budget

The project operates within a minimal budget, leveraging existing university resources. Hardware requirements are met using the university's computer lab equipment. Software needs are addressed through free licenses provided by the university for tools such as Python, R, and MATLAB. Data acquisition relies on publicly available wine datasets or web scraping techniques. Potential additional expenses include specialized software licenses (AUD100-AUD200) and incentives for small-scale user testing (AUD50-AUD100). The total budget is capped at AUD300. Resource management will be conducted carefully, with the option to seek additional university support if needed. Any approved extra expenses will be equally shared among team members.

1. Resources Roles and Responsibility

The project team consists of Julia Xu (s8139499), Kai Xing (s8143230), Linyang Wang (s8109316), and Sihua Tang (s8107067), each contributing to specific sections and roles within the project. Julia Xu is responsible for the Introduction, Background and Client Profile, and Purpose and Objectives sections, focusing on project planning and ensuring clear communication of the project's scope. Kai Xing handles the Assumptions and Constraints and Scope and Exclusion sections, with a focus on data extraction, preprocessing, and feature engineering, particularly using Natural Language Processing (NLP) techniques for text analysis. Linyang Wang oversees the Deliverables and Schedule sections, leading the development and optimization of the deep learning model, including the integration of the Random Forest algorithm for combined analysis. Sihua Tang is tasked with the Budget and Resources, Roles, and Responsibility sections, developing the user interface and ensuring system usability through user testing and feedback.

Together, the team works collaboratively to refine the project's deliverables, ensuring they meet high-quality standards. Dr. Rachid Hamadi, the client, provides essential project requirements and feedback, helping align the project's outcomes with industry standards. The project is supervised by Dr. Wenjie Ye, who offers guidance on both technical and academic aspects, while Dr. Quan Vu serves as the coordinator, ensuring the project's alignment with academic requirements and providing necessary administrative support and resources. This structure and division of roles enable a comprehensive and well-rounded approach to developing the intelligent wine recommendation system.

1. References

Abdelaziz Sami, n.d. Wine Quality Dataset. [online] Available at: https://www.kaggle.com/datasets/abdelazizsami/wine-quality [Accessed 7 August 2024].

Ashenfelter, O. and Storchmann, K., 2010. Using Hedonic Models to Study the Economics of Wine and Enology. Journal of Wine Economics, 5(1), pp.2-13. Available at: https://www.jwe.vc.vt.edu/ [Accessed 7 August 2024].

Smith, S.P., 2015. Wine Informatics: Using Data Science to Explore the World of Wine. IEEE Transactions on Computational Social Systems, 2(1), pp.48-62. Available at: https://ieeexplore.ieee.org/ [Accessed 7 August 2024].