Universiti Teknologi MARA

KUALA TERENGGANU HOMESTAY RECOMMENDER SYSTEM USING CONTENT BASED FILTERING ALGORITHM

Nur Atierah Binti Nor Azman

Thesis submitted in fulfilment of the requirements for
Bachelor of Computer Science (Hons.)
College of Computing, Informatics and Mathematics

January 2024

ACKNOWLEDGEMENT

The study was successfully conducted within the designated timeframe, and I express gratitude to Allah for His magnificence and the abundant blessings bestowed upon me. Firstly, I would like to extend my appreciation to my supervisor for her invaluable guidance and support throughout this process. This research would not have been achievable without the invaluable support and guidance provided by Miss Nik Marsyahariani Binti Nik Daud. Her assistance not only facilitated the completion of this study but also demonstrated her unwavering commitment, time and dedication. Furthermore, I would like to extend my sincere appreciation to Madam Ummu Fatihah, who fulfilled the role of my instructor for CSP600 and CSP650. Throughout the entire semester, she generously imparted her expertise, offered unwavering support, and served as a source of inspiration. Other than that, I would like to thanks all of my lecturers that has been consistently reminding, encouraging and ideas sharing throughout the journey.

My family members also have played a crucial role in supporting me throughout every step of this project. I would like to express my profound gratitude towards my parents for their unwavering provision of support either by physical and mental aspects, as well as their consistent encouragement and financial aid during times of utmost necessity. I also would like to extend my gratitude to my esteemed friends for their invaluable assistance and unwavering emotional support, which played a pivotal role in the successful completion of my final year project. Last but not least, may this project become one of my practices where I can share my knowledge to the readers, to the best of my knowledge.

ABSTRACT

Kuala Terengganu's tourism industry is expanding quickly, which has increased demand for acceptable hospitality and lodging alternatives. However, finding the ideal homestay that aligns with user preference can be time taking task. Hence, Content-Based filtering is capable to find the matching homestay based on the user preferrable criteria of the homestay such as price, locations and availability. Then in order to identify the distinguishing features or term of the homestay, TF-IDF will be calculating the importance term in the homestay review or description. By analyzing this textual content, TF-IDF will creates a representation of each homestay in the form of a corpus vector. It is required to find and compares the similarity between the user's preferences and the characteristics of the homestay. The similarity measure needs to be done in order to generate the recommendations. Cosine Similarity and TF-IDF will allow the prototype to identify the homestays that align closely with the user's desired features. The system's performance is meticulously evaluated through precision at 79.3%, recall at 93%, F1-score at 86%, RMSE at 9.95 and MAE at 9.42. These results demonstrate the system capability to offer accurate and relevant homestay suggestions. The project has successfully achieved all of the objectives and able to demonstrated the practical application of content-based filtering algorithm.

TABLE OF CONTENTS

CHAPTER THREE: PROJECT METHODOLOGY