

SENTIMENT ANALYSIS ON HOTEL REVIEWS USING MACHINE
LEARNING

NURFATINI ATIQA BINTI HAMIDI

UNIVERSITI TEKNOLOGI MALAYSIA

CHAPTER 1

INTRODUCTION

1.1 Introduction

As a traveller, booking a hotel is part of the planning. Usually, other people's review will influence a user's decision to find and book the best hotel to stay. Reviews available on the internet are more relevant, actual and detailed than the reviews found in hotel brochures (Walter Kasper & Mihaela Vela, 2012). There are various sources of online platform to find the reviews to get a better insight about the hotel's reputation. For example, Google Review, Agoda and Booking.com. So, customer's reviews play an important part for both user and business owner for decision making and improvement for the services.

Sentiment analysis also known as "opinion mining" or "emotion AI" is a technique used to gather and examine user thoughts, opinions, and responses to a particular topic (Zahid & Linköping University, 2020). Sentiment analysis is frequently performed using text mining with Natural Language Processing (NLP) tools to examine evaluations and reactions (Zahid & Linköping University, 2020). It involves analysing the review in form of text to determine whether it expresses positive, negative, or neutral sentiment.

There is a variety of machine learning algorithms such as Naive Bayes, Logistic Regression, Support Vector Machines (SVM) and Random Forests that can be employed to the categorization of the sentiment analysis of hotel reviews. In this project, TF-IDF calculation is used. TF stands for term frequency and IDF stands for inverse document frequency. In a simple word, it provides those keywords, using which some specific documents can be identified or categorized. Random Forest is a popular machine learning algorithm that can be used for both classification and regression problems in Machine Learning.

1.2 Problem Background

In this digital era, online reviews have become a key factor influencing consumer decisions making in the hospitality industry. User rely much based on reviews from previous customers to gauge the hotel quality. The review also becomes a vital aspect of a hotel's online reputation and competitive edge. Platforms such as TripAdvisor, Agoda, and Booking.com received thousands of hotel reviews from their customer that reflecting the customer's experience and emotion towards their stay. This huge amount of data contains crucial opinion related information that can be used to benefit for businesses and other aspects of commercial and scientific industries (Rezwanul et al., 2017). To manually tracking and pulling out all the data is challenging. Traditional methods are very time-consuming and unable to capture the depth and exact details of customer sentiments expressed in text form. Customer often convey mixed emotions within a single review, such as positive remarks about room quality paired with complaints about service. Thus, sentiment analysis is one of the techniques that can be used to capture and interpret the review into simple categorization.

1.3 Problem Statement

Hotel reviews have become an important thing to the business owner because it can be source of income to the business. However, due to the high volume of reviews generated on daily basis, it is difficult to the business owner to respond in a timely manner. Traditional analysis method is very time consuming. For example, business owners need to read one by one the mix of good and bad reviews. It can lead to misleading the details from the cust

Apart from that, the inconsistency between review and rating from the customer, customers give a good review but rate for one star. For example, the review "Good place to stay" but the rating is one star that represents not good. By implement the sentiment analysis, it can help the business owner to give accurate interprets of the emotion and satisfaction levels from the customer reviews.

1.4 Research Objectives

The proposed project aims to achieve the following objectives:

- (a) To conduct exploratory data analysis to identify patterns of hotel reviews.
- (b) To design and implement sentiment analysis that predict the review either positive or negative.
- (c) To conduct comprehensive evaluations on the developed predictive model and build an interactive dashboard.

1.5 Scope of Study

The proposed of this project is sentiment analysis on Hotel Review. The sentiment analysis prediction based on the customer review (text) and the prediction is being categorized into two types which are positive and negative. In this project, the dataset used downloaded from Kaggle and word cloud will be used to interpret the review. The dashboard will be developed using data visualization tools in PowerBI to provide an insight of the positive and negative review.