**Project Management**

**Topic**

Hotel rating based on customer review and classification & sentimental analysis of review.

**Team Members**

Mdakbar Sarkar

Robert Fajardo

Rajesh Pahari

Platini Dacheu

**Objectives**

The main objective of this project is to build a datacentric application that not only helps the customer but also identifies key areas to improve from the provider’s perspective. After completion of the project, we will have a solid understanding of NLP and how we use the concept in the practical field.

In this management document, we will be giving an overview of what have been accomplished so far in this increment and a summary of what is to be done in the second move.

**Accomplished work**

In this section, we are uncovering what we have already done on our project. Briefly, the following points were addressed:

* Background by **Rajesh Pahari**
* Exploratory data analysis by **Robert Fajardo**
* Feature Extraction by **Mdakbar Sarkar**
* Project management by **Platini Dacheu**

1. **Background**

Research or project topics are usually not formulated for the first time. Thus, the often build on previous works either in the same domain or some close domain. They can even be inspired by a totally different branch of science. Our topic does not escape this generality. To understand the state of the art of the problem, we have researched previous works accomplished in the area. This helped us more precisely fix the boundaries of our topics. Therefore, areas like sentiment analysis, product rating and review have been deeply explored to help us accomplish our task.

1. **Exploratory Data Analysis**

All data science projects have this one fuel in common: data. For our project, we will be using the TripAdvisor Hotel dataset. Every data-based work starts by understanding the data at hand. We have therefore thoroughly “searched” our data to get the more sense of it. To that end, we have massively used visualization as the mean to uncover our data. Suitably chosen charts have made that gives us big to detailed pictures of the data we will be using and will guide us through the next steps of our work.

1. **Feature Extraction**

Successful machine learning model rely on data. Data need to be in good quantity as well as quality. Unfortunately, raw data is almost never appropriate to train a model. In order to make it suitable for model training, data needs to go through a series of steps called globally data processing. In this process, data is cleaned and prepare for model building. One important thing in these steps is the extraction of features. Data are defined by a lot of features or attributes, but not all of them are important to build a specific model. Each model has its requirements and specifications. When it comes to text data as the one we are using, there is a huge amount of words in a language vocabulary. However, not all of them are relevant in the transmission of information. The idea here is to get rid of unnecessary words and keep only those that are meaningful in transmitting information. That is what this section have been focusing on.

**Upcoming work**

A lot have already been accomplished in our project, but we still have a road to go. We have basically focused here on what we could call a “pre-work”, kind of preliminaries. The next increment will be focusing on the actual model building. We aim to build three models for our project:

* Text classification model
* Sentiment analysis model
* Topic modeling

Ultimately, this will help us make better suggestions to travelers when it comes to choosing where they want to accommodate during their stay.

**Block Diagram:**

Sentiment Score

Evaluate Model

Training Model

Word Embedding

Data Preprocessing

Prediction

Data Labeling

Word Embedding

Data Preprocessing

New Review

Hotel Reviews

**Reference**

1. Sentimental Analysis: <https://www.kaggle.com/code/emirkocak/in-depth-series-sentiment-analysis-w-transformers>
2. Hotel Rating: <https://www.kaggle.com/code/nidhaypancholi/predicting-rating-from-reviews-trip-advisor>
3. Sentiment Analysis in Hotel Reviews: https://www.altexsoft.com/blog/sentiment-analysis-hotel-reviews/