1. Project Idea:

- Sentiment Analysis on Hotel Reviews Using Machine Learning Techniques
- The hospitality industry relies heavily on customer feedback to improve services and enhance customer satisfaction. However, manually analyzing vast amounts of online reviews can be time-consuming and inefficient. This project aims to develop a sentiment analysis model using machine learning techniques to automatically classify hotel reviews as positive, negative, or neutral.

2. Relevance to Sustainable Development Goals (SDGs):

- This project aligns with SDG 8, which promotes decent work and economic growth. By automating the analysis of hotel reviews, businesses can gain valuable insights into customer sentiment, allowing them to make informed decisions that improve customer satisfaction, attract more guests, and ultimately boost their economic growth.

3. Literature Examples:

- "Sentiment Analysis of Hotel Reviews Using Deep Learning" by Wu et al. (2021): This study developed a deep learning model to classify hotel reviews as positive or negative. The model achieved high accuracy and demonstrated the potential of deep learning for sentiment analysis in the hospitality industry.
- "A Comparative Study of Machine Learning Techniques for Sentiment Analysis of Hotel Reviews" by Li et al. (2020): This paper compared the performance of various machine learning algorithms, including decision trees, logistic regression, and Naïve Bayes, for sentiment analysis of hotel reviews. The results showed that the decision tree algorithm achieved the highest accuracy.

4. Describe Your Data:

- The data for this project will be the Tripadvisor Hotel Reviews dataset, which is available from the Kaggle platform. The dataset contains over 20,491 hotel reviews, each with a rating (1-5 stars) and a review text. The target variable is the sentiment of the review, which can be classified as positive, negative, or neutral.

5. Approach (Machine Learning or Deep Learning):

- Given the size and complexity of the dataset, a machine learning approach will be used. Specifically, three machine learning algorithms will be employed: decision tree, logistic regression, and Naïve Bayes. These algorithms are well-suited for text classification tasks and have been shown to be effective for sentiment analysis.