**Project Title:** Beer Type Prediction Model Deployment

**Project Author:** Ksenia Arkhipova

**GitHub Repository:** https://github.com/arkksenia/adv\_dsi\_assign2\_beer\_predict.git

**API URL:** https://advdsiassign2beerpredict-nfkrkyxdv9xwtb3kjseadd.streamlit.app

**Executive Summary**

This report outlines the development and deployment of a machine learning model for predicting the type of beer based on user ratings in four categories: appearance, aroma, palate, taste. The project includes building a web application to serve real-time predictions via an API. This report provides comprehensive information on data collection, preprocessing, model selection, training, evaluation, deployment, API structure, and instructions for running predictions.

**1. Introduction**

The primary objective of this project is to create a robust machine learning model capable of predicting the type of beer based on users' ratings. The model's deployment will be facilitated through a web application, enabling real-time predictions.

**2. Data Collection and Preprocessing**

The dataset acquired consists of beer types and corresponding user ratings in various categories. Data preprocessing was carried out to ensure data quality and model compatibility. This phase encompassed data cleaning, handling missing values, and the selection of relevant features. The following features were chosen for prediction: brewery names, user ratings on appearance, aroma, palate, taste, alcohol by volume. Categorical variables were one-hot encoded, and numerical features were standardized to facilitate model training.

**3. Model Selection**

The K-Nearest Neighbours (KNN) algorithm was selected for this project. KNN is a non-parametric, instance-based learning algorithm. It was chosen for its versatility in handling multi-class classification tasks, as well as its capacity to work effectively with a high-dimensional feature space.

**Feature Set**

The model uses the following features for prediction:

* Brewery name
* Appearance rating
* Aroma rating
* Palate rating
* Taste rating
* Alcohol by volume

**4. Model Training and Evaluation**

The dataset was split into training and testing sets, with 80% used for training and 20% for testing. The KNN Classifier was trained on the training dataset.

Model performance was assessed using the accuracy score evaluation metric, however, for better evaluation other metrics, such as precision, recall and f1 score, should be used in the future.

**Results:** The model achieved an accuracy of 0.535. This metric indicates the model achieved a moderate level of accuracy in predicting beer types. While not highly accurate, it may still provide some valuable insights into what the name of the beer might be based on the user’s input. Users might find it useful for making more informed decisions, even if it's not a definitive predictor.

**5. Model Deployment**

For model deployment, a web application was developed using Streamlit. Streamlit offers a user-friendly way to serve the model and provides a platform for real-time predictions.

**6. API Structure**

The API comprises a single endpoint for making predictions.

**7. Instructions for Running Predictions**

To execute predictions utilizing the deployed model, follow these steps.

Clone the GitHub repository containing the web application:

bash

Git clone [repository\_url]

Install the essential Python packages, using

pip install -r requirements.txt

Initiate the Streamlit application:

Beer\_predict\_app.py

The web application can be accessed [API\_url].

Input the requested data in the application:

1. Select a brewery name from a drop-down menu
2. Input beer ratings for five categories: appearance, aroma, palate, taste, and alcohol by volume.
3. Click the “predict Beer Type” button

The response will contain the predicted beer type.

**8. Conclusion**

This project successfully deployed a machine learning model capable of predicting beer types based on user ratings, brewery name and alcohol by volume input. The model delivered moderate accuracy, and its real-time prediction service is accessible through a Streamlit web application. This application enhances user experience by assisting in beer selection based on individual preferences.

**9. Future Enhancements**

Several opportunities for future enhancements exist:

Explore other models or experiment with different parameters, particularly k-parameter, to improve model accuracy and gain better predictions.

Integrate user feedback to continuously improve the model's accuracy and predictions.

This project represents a foundation for building a robust beer recommendation system and can serve as a platform for further developments and improvements in the future.