

**Project Proposal (proposed Solution)**

|  |  |
| --- | --- |
| Date | 9 July 2024 |
| Team ID | team-739821 |
| Project Title | Precise Coffee Quality Prediction |
| Maximum Marks | 3 Marks |

**Project Proposal (Proposed Solution) template**

This project proposal outlines a solution to address a specific problem. With a clear objective, defined scope, and a concise problem statement, the proposed solution details the approach, key features, and resource requirements, including hardware, software, and personnel.

|  |  |
| --- | --- |
| **Project Overview** | |
| Objective | Develop an advanced machine learning-based system to provide precise and consistent coffee quality predictions, ensuring fair pricing, maintaining product quality, and enhancing customer satisfaction |
| Scope | 1. **Data Collection and Preprocessing:** Gather and clean coffee sample data.  2. **Model Development:** Implement and optimize machine learning models.  3. **Validation and Testing:** Validate models with test data and pilot tests.  4. **Deployment:** Create a user-friendly interface and deploy the system.  5. **Training and Support:** Offer training and support for users. 6. **Continuous Improvement:** Collect feedback and update the  system continuously. |
| **Problem Statement** | |
| Description | The current methods for assessing coffee quality are subjective and inconsistent, leading to unreliable quality evaluations and economic disparities in the coffee industry. |
| Impact | Solving this problem will ensure fair pricing for farmers, consistent product quality for roasters, and improved customer satisfaction, ultimately leading to a more reliable and profitable coffee industry. |





|  |  |
| --- | --- |
| **Proposed Solution** | |
| Approach | Utilize best model to predict coffee bean health based on provided attributes. This involves training the model on a dataset containing features such as moisture content, bean color, aroma intensity, and acidity levels. Employ cross-validation to optimize model parameters and ensure robustness. |
| Key Features | **Random Forest Model:** Chosen for its ability to handle complex datasets with multiple features and provide robust predictions. **Feature Importance:** Analyze feature importance to understand which attributes most significantly influence coffee bean health predictions. **Scalability:** Ensure the solution can handle varying dataset sizes and future scalability needs of the website. **Interpretability:** Provide insights into model predictions to enhance user understanding and trust in the predictions. |

**Resource Requirements**

|  |  |  |
| --- | --- | --- |
| **Resource Type** | **Description** | **Specification/Allocation** |
| **Hardware** | | |
| Computing Resources | CPU/GPU specifications, number of cores | e.g., 2 x NVIDIA V100 GPUs |
| Memory | RAM specifications | e.g., 16 GB |
| Storage | Disk space for data, models, and logs | e.g., 1 TB SSD |
| **Software** | | |
| Frameworks | Python frameworks | e.g., Flask |
| Libraries | Additional libraries | e.g, scikit-learn, pandas, numpy,seaborn,matplotlib.py plot. |
| Development Environment | IDE, version control | e.g, Google Colab |
| **Data** | | |
| Data | Source, size, format | e.g., Kaggle dataset, ,github,10,000 images |