ForkFinder -A Smart Restaurant recommendation system

Milestone 1: Project Initialization and Planning Phase

The "Project Initialization and Planning Phase" marks the project's outset, defining goals, scope. This crucial phase establishes project parameters, identifies key elements, allocates resources, and outlines a realistic timeline. It also involves risk assessment and mitigation planning. Successful initiation sets the foundation for a well-organized and efficiently executed machine learning project, ensuring clarity, alignment, and proactive measures for potential challenges.

### Activity 1: Define Problem Statement

Problem Statement: Develop a personalized restaurant recommendation system that effectively addresses the challenges faced by diners in discovering and choosing the right dining experiences. Improve user satisfaction by providing relevant and engaging restaurant recommendations.

Activity 2: Project Proposal (Proposed Solution) :

The proposal report aims to transform restaurant recommendation system using machine learning, boosting efficiency and accuracy. It tackles system inefficiencies, promising better operations, reduced risks, and happier customers. Key features include a machine learning-based model and real-time decision-making.Addressing inaccuracies and inefficiencies in the current restaurant recommendation system adversely affects operational efficiency and customer satisfaction.

Activity 3: Initial Project Planning

Initial Project Planning involves outlining key objectives, defining scope, and identifying restaurants and their reviews for user. It encompasses setting timelines, allocating resources, and determining the overall project strategy. During this phase, the team establishes a clear understanding of the dataset, formulates goals for analysis, and plans the workflow for data processing. Effective initial planning lays the foundation for a systematic and well-executed project, ensuring successful outcomes.
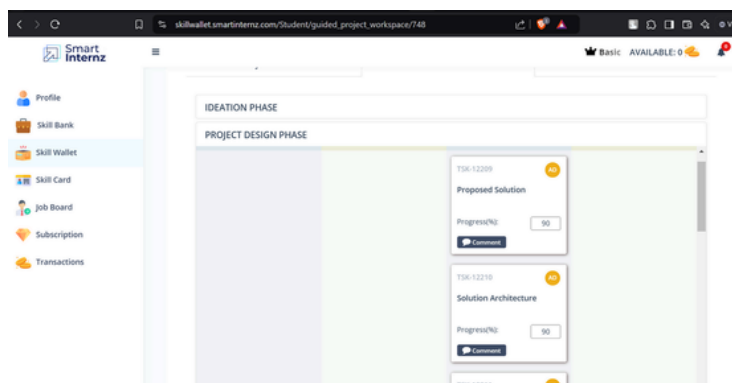
Milestone 2: Data Collection and Preprocessing Phase

The Data Collection and Preprocessing Phase involves executing a plan to gather relevant restaurant

data from Kaggle, ensuring data quality through verification and addressing missing values. Preprocessing tasks include cleaning, encoding, and organizing the dataset for subsequent exploratory analysis and machine learning model development.

Activity 1: Data Collection Plan, Raw Data Sources Identified, Data Quality Report

The dataset for "ForkFinder -A Smart Restaurant recommendation system" is sourced from Kaggle. It includes restaurant details and reviews and other important data. Data quality is ensured through thorough verification, addressing missing values, and maintaining adherence to ethical guidelines, establishing a reliable foundation for predictive modeling.Search for datasets related to recommendation,statistical information, and restaurant details. Prioritize datasets with diverse demographic information.

Activity 2: Data Quality Report

The dataset for "ForkFinder -A Smart Restaurant recommendation system" is sourced from Kaggle. It includesrestaurant details and reviews and other important data. Data quality is ensured through thorough verification, addressing missing values, and maintaining adherence to ethical guidelines, establishing a reliable foundation for predictive modeling.

Activity 3: Data Exploration and Preprocessing

Data Exploration involves analyzing the restaurant dataset to understand patterns, preferrences, and diversity. Preprocessing includes handling missing values, scaling, and encoding categorical variables. These crucial steps enhance data quality, ensuring the reliability and effectiveness of subsequent analyses .

Milestone 3: Model Development Phase

The Model Development Phase entails crafting a predictive model for restaurant recommendation. It encompasses strategic feature selection, evaluating and selecting models (Random Forest, Decision Tree, KNN, XGB), initiating training with code, and rigorously validating and assessing model performance for informed decision-making in the recommendation process.

Activity 1: Feature Selection Report

The Feature Selection Report outlines the rational behind choosing specific features (e.g., Location, Type of restro, Review,Location,Price) for recommendation model. It evaluates relevance, importance, and impact on predictive accuracy, ensuring the inclusion of key factors influencing the model's ability to different users.

Activity 2: Model Selection Report

The Model Selection Report details various models will be outlined, detailing their descriptions, hyperparameters, and performance metrics, including Accuracy or F1 Score. This comprehensive report will provide insights into the chosen models and their effectiveness. It considers each model's strengths in handling complex relationships, interpretability, adaptability, and overall predictive performance, ensuring an informed choice aligned with project objectives.

Activity 3: Initial Model Training Code, Model Validation and Evaluation Report

The Initial Model Training Code employs selected algorithms on the restaurant dataset, setting the foundation for predictive modeling. The subsequent Model Validation and Evaluation Report rigorously assesses model performance, employing metrics like accuracy and precision to ensure reliability and effectiveness in predicting outcomes.

```python
#importing and building the KNN model
def KNN(X_train,X_test,y_train,y_test):
    model = KNeighborsClassifier()
    model.fit(X_train,y_train)
    y_tr = model.predict(X_train)
    print(accuracy_score(y_tr,y_train))
    yPred = model.predict(X_test)
    print(accuracy_score(yPred,y_test))


#printing the train accuracy and test accuracy respectively
KNN(X_train,X_test,y_train,y_test)
```

Milestone 4: Model Optimization and Tuning Phase

The Model Optimization and Tuning Phase involves refining machine learning models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

Activity 1: Hyperparameter Tuning Documentation

The Gradient Boosting model was selected for its superior performance, exhibiting high accuracy during hyperparameter tuning. Its ability to handle complex relationships, minimize overfitting, and optimize predictive accuracy aligns with project objectives, justifying its selection as the final model.

Activity 2: Performance Metrics Comparison Report

The Performance Metrics Comparison Report contrasts the baseline and optimized metrics for various models, specifically highlighting the enhanced performance of the Gradient Boosting model. This assessment provides a clear understanding of the refined predictive capabilities achieved through hyperparameter tuning.

Activity 3: Final Model Selection Justification

The Final Model Selection Justification articulates the rationale for choosing Gradient Boosting as the ultimate model. Its exceptional accuracy, ability to handle complexity, and successful hyperparameter tuning align with project objectives, ensuring optimal predictions.It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

```python
# Define the Decision Tree classifier
dt_classifier = DecisionTreeClassifier()

# Define the hyperparameters and their possible values for tuning
param_grid = {
    'criterion': ['gini', 'entropy'],
    'splitter': ['best', 'random'],
    'max_depth': [None, 10, 20, 30, 40, 50],
    'min_samples_split': [2, 5, 10],
    'min_samples_leaf': [1, 2, 4]
}
```

```python
fine the Random Forest classifier
lassifier = RandomForestClassifier()

fine the hyperparameters and their possible values for tunin
n_grid = {
    'n_estimators': [50, 100, 200],
    'criterion': ['gini', 'entropy'],
    'max_depth': [None, 10, 20, 30],
    'min_samples_split': [2, 5, 10],
    'min_samples_leaf': [1, 2, 4],
```

Milestone 5: Project Files Submission and Documentation

For project file submission in Github, Kindly check my git hub profile

Milestone 6: Project Demonstration

In the upcoming module called Project Demonstration, individual records a video by sharing their screens. Explain of project and demonstration of its execution during the presentation.