**Innovation Phase : Predicting IMDb scores**

Step 1: Data Collection and Understanding

**Data Collection:**

The IMDb scores of various movies and the other relevant information are gathered from the provided kaggle link.

**Data Understanding:**

Analyze the dataset to understand the features, categories and the possible patterns that can be identified.

Step 2: Data Preprocessing

**Data Cleaning:**

To remove the irrelevant information, missing values and blank spaces, outliers etc to ensure consistency and data quality. The missing values are replaced by proper techniques such that it does not cause any significant impact on the dataset.

Step 3: Feature Engineering

Create new features from the existing patterns such as from genre, duration etc and encoding the categorical variables into meaningful features for machine learning models.

Step 4: Model Selection and Training

**Model Selection:**

Based on the dataset's characteristics a suitable machine learning algorithm (linear regression, SVM, Random forest etc) is implemented.

**Data Splitting:**

Split the dataset into training and testing sets (e.g., 80/20 split) to train and validate the model.

Step 5: Model Training and Validation

**Model Training:**

Train the selected model using the training data to learn the patterns in the IMDb scores in relation to the various criteria given in the dataset.

**Model Validation:**

Validate the trained model using the testing data to evaluate its accuracy and generalization capability.

Step 6: Model Evaluation and Optimization

**Evaluation Metrics:**

Use appropriate regression metrics like Mean Absolute Error (MAE) or Root Mean Squared Error (RMSE) to evaluate the model's performance.

**Optimization:**

Fine-tune hyperparameters and features to optimize the model for better predictions.

Step 7: Documentation and Deployment

**Documentation:**

Document the entire analysis process, including data preprocessing, feature engineering, model selection, training, evaluation and optimization.

Provide clear explanation of chosen algorithms, parameters and evaluation metrics along with interpretation of the results.

**Deployment:**

Deploy the machine learning model as a user friendly application such that it recommends movies with highest rating based on their preferences.

Output:

**Predictive Model:**

A validated and optimized machine learning model to predict the IMDb scores based on the features like genre, premier date, runtime and language is built such that the user experience is enhanced and improved.