**Ideation Phase**

**Defining the Problem Statements**

|  |  |
| --- | --- |
| **Date** | **26-09-2023** |
| **Team ID** | **3866** |
| **Project Name** | **IMDb Score Prediction System Development** |

**IMDb Score Prediction System Development**

**Problem Definition and Design Thinking**

**Introduction**

The objective of this project is to create an IMDb score prediction system that efficiently and accurately predicts movie ratings based on various features. This system aims to address concerns such as inefficiency, accuracy, and user-friendliness in the IMDb score prediction process. It will provide valuable insights to movie enthusiasts and professionals in the film industry.

In this document, we will outline the problem statement, the phases involved in solving it, and the design thinking approach guiding our project.

**Problem Statement**

Objective: Develop an IMDb score prediction system with high efficiency, accuracy, and user-friendliness.

Data: We will gather movie-related data, including historical IMDb scores, movie metadata, user ratings, and reviews. This data will be cleaned and prepared for analysis and prediction.

**Key Concerns and Requirement**:

1. Inefficiency: Manual data collection and prediction processes are inefficient and time-consuming.

2. Accuracy: Improve prediction accuracy by exploring relationships between various features and IMDb scores.

3. User-friendliness: Create a user-friendly interface for easy access and interaction.

4. Reliability: Implement a feedback loop for user input to enhance reliability.

5. Customization: Allow customization based on individual preferences.

6. Integration: Seamlessly integrate with other movie-related apps and services.

7. Testing and Quality Assurance: Rigorously test the system for accuracy and reliability.

8. Monitoring and Maintenance: Continuously monitor and maintain the system after deployment.

**Design Thinking Approach**

**Empathize:**

Before diving into solving the problem, it's essential to empathize with users and understand their needs. In this case, our primary users are movie enthusiasts and professionals in the film industry. We need to gather insights into what factors are essential to them when predicting IMDb scores and how an accurate prediction system can benefit them.

**Actions:**

- Conduct surveys or interviews with movie enthusiasts and film professionals to gather their perspectives.

- Analyse historical IMDb score trends to identify critical predictive factors.

- Seek feedback from domain experts in the film industry.

**Define:**

Based on our understanding of the problem and users' needs, we will define clear objectives and success criteria for our project.

**Objectives:**

- Develop an IMDb score prediction model with a Mean Absolute Error (MAE) of less than X on the test data.

- Create an intuitive user interface for users to input movie details and receive score predictions.

**Ideate:**

Brainstorm potential solutions and approaches to address the problem. This phase involves considering various machine learning algorithms and techniques for IMDb score prediction.

**Actions:**

- Explore different machine learning algorithms, such as regression, decision trees, and neural networks.

- Experiment with feature engineering techniques to improve prediction accuracy.

- Consider incorporating additional data sources, such as movie genre, director, and cast, to enhance predictions.

**Prototype**

Create a prototype of the IMDb score prediction system, including the machine learning model and the user interface.

**Actions:**

- Develop a Python script for data preprocessing, model training, and evaluation.

- Create a user-friendly web interface using frameworks like Flask or Django for inputting movie details and receiving predictions.

- Test the prototype with a subset of the dataset to ensure it meets performance objectives.

**Test**

Evaluate the model's performance using appropriate metrics and gather user feedback on the interface.

**Actions:**

- Split the dataset into training and testing sets.

- Train the model on the training data and evaluate it on the testing data using metrics such as MAE.

- Collect user feedback on the web interface for usability and accuracy.

**Implement**

- Once the prototype meets defined objectives and receives positive feedback, proceed with full implementation.

**Actions:**

- Train the final machine learning model using the entire dataset.

- Deploy the IMDb score prediction system as a production-ready web application.

- Conduct thorough testing to ensure the system is robust and user-friendly.

**Iterate**

Continuous improvement is crucial. Gather user feedback and iterate on the model and interface to enhance accuracy and usability.

**Actions:**

- Monitor the system's performance and retrain the model periodically with updated data.

- Address user feedback and make necessary improvements to the web interface.

- Stay informed about advancements in machine learning and movie rating models for potential enhancements.

**Conclusion**

In this document, we've outlined our approach to solving the problem of IMDb score prediction using data science and design thinking. We've defined the problem, identified key concerns and requirements, and presented a design thinking approach that includes empathizing with users, defining objectives, ideating solutions, prototyping, testing, implementing, and iterating.

Our ultimate goal is to develop an efficient, accurate, and user-friendly IMDb score prediction system that provides valuable insights for movie enthusiasts and professionals in the film industry. By following this structured approach, we aim to create a reliable tool that contributes positively to the world of movie rating prediction.