Mark Myers

C964 Task 2 Part B

Movie Recommendation Project Proposal

Problem Statement

Customer feedback shows that they are unsatisfied with our current recommendation system. The movies that are recommended have little similarity or interest. Feedback from customers shows that they are not interested in continuing to use our services since they have a hard time finding something to watch. The ability to have a high-quality recommendation system in place is tantamount to broadening our customer base.

Customer Summary

Data retrieved from user profiles has shown a broad spectrum in ages and preferred movie genres. As one would expect, action and adventure movies are popular among men, but, surprisingly, family movies are also popular. Conversely, romance and family movies have shown to be popular among women. But there too is the surprise that horror movies are also amongst the most popular genres.

We have sent out questionnaires to customers and received the feedback that poor recommendations are their biggest issue with our service. The proposed application will be able to either take a movie as a search parameter and return high-quality recommendations that are similar to the movie that was searched. Fixing this issue will both satisfy our current customers and help to expand our customer base.

Existing System Analysis

Currently, our services are provided through an app download on a smartphone or on a computer or tablet. A downfall to being limited to these applications is that many people prefer to watch movies on their television’s larger screen compared to the small screen of a smartphone, tablet, or computer.

The project proposal is the first of two projects, with the second project of creating an app for televisions to use our service. This project will improve the services provided to the customers, which mean happier customer and referrals to friends and family to sign up with us. Being able to generate a solid customer base that has continual expansion will show senior management that investing the time and resources into the second project will be well worth it in the long run.

Data

The dataset ‘movies.csv’ has columns more movie title, rating, genre, score, director, writer, lead actor, and runtime. As more movies get added to the library, each of these categories will be checked to ensure that they have the appropriate data for the specific movie that is being entered. Some movies that are entered will have information missing such as a rating or a director. Our team will work to find the correct information to be able to enter accurate data into the dataset. Movies that have no ratings listed for them will be changed to ‘Unrated’.

Eventually, we will most likely add television programs that will have different ratings compared to the movies. We will have to ensure that the program can read this information easily without having to constantly go in and change coding.

Project Methodology

During the development of this application, we will follow the waterfall methodology. Using this method, we will be able to plan, develop, and test each step of the process to ensure that the current step meets our standards and will set us up for success in subsequent steps. While there are downfalls to this method, such as flexibility, using this method will be the most cost effective for our organization to ensure that we don’t need to go back and rework parts of the application.

Initially, we will be going with a basic interface to get the application launched and get customer satisfaction boosted as soon as possible. In the future, we will work on developing a more attractive interface with additional features that we discover will be useful from utilization and customer feedback.

The project will follow the SEMMA methodology for data mining.

• Sample: An initial questionnaire will be provided to the users to help us get a better understanding of what types of movies they enjoy watching.

• Explore: The data collected from the questionnaires will be reviewed to discover trends and popular genres.

• Modify: Trends amongst users with similar likes will be assessed to help determine if a user might enjoy a recommendation that they had not previously considered.

• Model: Trends that have been discovered will be modeled to show what genres of movies are typically liked by users with similar likes.

• Assess: Using the models, testing the recommendation system out on a small group will be done to assess the usefulness of the data collected and the models that have been prepared.

Project Outcomes

Upon completion of the project, the application will be launched to the customer base. There will be a basic interface for the user to interact with. They will have the options to be able to search for movies that have a common genre, director, or lead actor. They will also be able to use a title search that will list similar movies. In addition, studies have shown that customers like to figure out what movies are the highest rated so that they can watch them. We will have a feature to show these movies to the user. We will also have an option for the user to view the lowest rated movies, since studies have shown that people also enjoy watching these movies.

The interface will be easy to understand and should not need too much explanation. However, a user guide will be created to help ensure users can find what they are looking for quickly rather than flipping through menus searching for the option they want.

Implementation Plan

Initially, we will start in the planning phase. As stated in the Project Methodology category, we will follow the SEMMA method for data mining. We will send out questionnaires to our customers to see what kind of features they want to see. It will help us understand, as well, what kind of movies they enjoy to see if we can find a commonality amongst various types of movies.

We will design the interface and the menus that will be given. We will then work to on how best to go about implementing these menus and how best to display the results. Testing will be done to ensure that the menu works as it should. Incorrect entries will be tested to ensure that the input checks are working correctly to avoid creating errors.

Upon completion of the interface, assignments will be given out to team members on which part of the project they should be working. Some members will work on creating and implementing the search features, where they can search for a genre or key words in the movies title, director, or lead actor. Other members will work on creating the lists of high and low rated movies. In addition, they will work on visuals to find common trends among ratings, genres, and scores.

Another team of members will work on the recommendation features. One of the recommendation features will be the ability to get a recommendation based on title, of which the title feature will contain machine learning to get highly accurate results. There will also be additional recommendation features including the highest rated movies for a specific genre, a specific director, or a specific lead actor.

After these steps have been completed, we will combine the different aspects of the program together to form the entire application. We will offer a beta test to users that sign up. After the beta testing, the full-scale launch will occur. Forms for user feedback will be sent out to gather ideas on what needs improvement or changed entirely.

Evaluation Plan

During development of the application, continual testing will be done to ensure that the process is ready for the next step, and we will not have to go back to change anything. We will do this by ensuring they processes meet our project requirements and the expected outcomes. Overall, the project will be considered successful if we achieve at least a 90% success rate for our movie recommendations. We will get this information from if the feedback of the users, as well as internal data seeing if they user chooses to watch one of the movies that has been recommended.

Resources and Cost

|  |  |  |
| --- | --- | --- |
| **Resource** | **Description** | **Cost** |
| Hardware | Org. owned computers | $0 |
| Software | Org. owned software | $0 |
| Cloud service | Storage of data | $50,000 |
| Programmers | Standard hours | $300,000 |
| Marketing | Standard hours | $100,000 |
|  | **Total** | $450,000 |

Timeline and Milestones

|  |  |  |  |
| --- | --- | --- | --- |
| **Phase** | **Start** | **End** | **Tasks** |
| 1 | Aug 1, 2022 | Aug 31, 2022 | Develop questionnaire |
| 2 | Sept 1, 2022 | Nov 30, 2022 | Develop recommendation system  Milestones: completion of the interface, completion of the search functions, completion of the different view, completion of the recommendation system |
| 3 | Dec 1, 2022 | Dec 31, 2022 | Testing and debugging |
| 4 | Jan 1, 2023 | Jan 31, 2023 | Test will small group of customers  Milestone: customer satisfaction of at least 90% |
| 5 | Feb 1, 2023 | Feb 14, 2023 | Go live with new system |