PROJECT REPORT ON



**MOVIE RECOMMENDATION SYSTEM USING**

**COSINE SIMILARITY APPROACH**

Submitted by**:**

Samrajya Chand

Trinity International College

03/22/2024

**ABSTRACT**

“Movie Recommendation using cosine similarity approach” is the implementation of cosine similarity algorithm approach that recommends the user similar movies in the basis of input provided by the user. Each movie is represented as a high-dimensional vector capturing its intrinsic characteristics like genre, plot summary, cast, and crew. By employing cosine similarity, which measures the cosine of the angle between two non-zero vectors, the system efficiently computes the similarity between movie vectors, enabling the identification of thematically and conceptually similar movies to a given input. We used python as an IDE and Streamlit for hosting the web interface. By using python libraries like Pandas we worked with the dataset and array of dataset provided.

**Keywords:** Cosine similarity, movie recommendation, vector representation, difflib, streamlit

TABLE OF CONTENT

## INTRODUCTION

#### Introduction

#### Problem Statement

#### Objectives

#### Scope and Limitation

## LITERATURE REVIEW AND RESEARCH METHODOLOGY

# Literature Review

# Framework of the Model

# Methodology

# SYSTEM DESIGN

# Requirement Collection

# Functional Requirement

# Non-Functional Requirement

# Feasibility Study

# System Design

1. Process Design

# Structuring System Requirements

1. Process Modeling

# IMPLEMENTATION

# RESULT AND ANALYSIS

# CONCLUSION

# REFERENCES

# APPENDIX

**CHAPTER 1**

**INTRODUCTION**

* 1. Introduction

With the ever-increasing volume of movies available across various platforms, providing personalized and relevant recommendations has become a crucial challenge in the entertainment industry. Traditional collaborative filtering techniques, which rely on user interactions and ratings, often suffer from limitations such as the cold-start problem, where recommendations cannot be generated for new or niche movies with limited data. Additionally, these methods can struggle with data sparsity issues, leading to inaccurate or biased recommendations.

To address these challenges, this project proposes a novel approach to movie recommendations by leveraging cosine similarity, a technique widely used in information retrieval and text mining. Each movie is represented as a high-dimensional vector capturing its intrinsic characteristics through natural language processing and feature engineering. Employing cosine similarity, which measures the angle between vectors, the system computes the similarity between movie vectors, enabling the identification of thematically similar movies. This approach addresses the cold-start problem, facilitates niche movie discovery, and explores dimensionality reduction and clustering for scalability and real-time recommendations.

* 1. Problem Statement

Traditional movie recommendation systems like collaborative filtering suffer from limitations such as the cold-start problem for new or niche movies, data sparsity issues, popularity bias, and scalability challenges as the number of users and movies grows. There is a need for an innovative approach that can provide accurate and personalized recommendations, even for new content, while ensuring scalability and efficiency for large datasets.

The proposed solution aims to develop a movie recommendation system using cosine similarity, a technique from information retrieval and text mining. By representing movies as high-dimensional vectors based on their characteristics (genre, plot, cast, etc.), the system can compute similarities between movies and recommend conceptually similar ones to a given input. This approach can overcome the cold-start problem, facilitate niche movie discovery, and mitigate popularity bias.

Additionally, the project explores dimensionality reduction and clustering techniques to enhance scalability and enable real-time recommendations for large-scale movie databases while maintaining accuracy and relevance.

* 1. Objectives

The objectives of this system are:

1. Providing Personalized and Relevant Recommendation
2. Enhancing system’s scalability
3. Overcoming cold start problem
4. Enabling discovery of thematically similar movie