PROJECT REPORT

On

Movie Review System

Semester VII

(2020-2021)

Department of Information Technology

BABU BANARASI DAS INSTITUTE OF TECHNOLOGY &

MANAGEMENT, LUCKNOW

Submitted to:

Prof. Girjesh Kumar Mishra

Department of Information
Technology

Submitted By:

Alisha Parveen - 1705413011

Satyam Seth - 17054130

Hashit Dixit -17054130

Ankit Gupta -17054130

Abhishek Kumar Yadav -
16054130

ACKNOWLEDGEMENT

It is indeed with a great pleasure and immense sense of gratitude that we acknowledge the help of these individuals, for the facilities provided to accomplish this main project.

We would like to thank our **Prof. Manuj Darbari**, Head of the Department of Information technology, BBD-ITM , for this constructive criticism throughout our project.

We feel elated in manifesting our sense of gratitude tour internal project guide **Prof. Girjesh Mishra, Department of Information Technology.** He has been a constant source of inspiration for us and we are very deeply thankful to him for his support and valuable advice.

We extremely grateful to our Departmental staff members, Lab technicians and Non-teaching staff members for their extreme help throughout our project.

Finally we express our heart full thanks to all of our friends who helped us in successful completion of this project.

<div align="center">

Project associates:

Satyam Seth

Harshit Dixit

Ankit Gupta

Alisha Parveen

Abhishek Kumar Yadav

</div>

ABSTRACT

The purpose of the project entitled as "**Movie Review System**" is to computerize and recommend movies based on the genre preferences which is user friendly simple, fast, and cost – effective. It deals with the collection of various movie databases provided through movie review module. Traditionally, it was done by critics and the choices of users wasn't taken into account. The main function of the system is register and store movie details and retrieves these details as and when required, and also to use these details meaningfully System input contains user login modules, movie review modules, admin control and system design . The Movie Review System can be entered using a username and password. It is accessible either by an administrator or registered user. Only the admin can add data into the database. The data can be retrieved easily. The data are well protected for personal use and makes the data processing very fast.

CHAPTER 1

INTRODUCTION

1.1 Introduction:

We usually come across movie rating websites where users are allowed to rate ad comment on movies online. These ratings are provided as input to the website rating system. The admin then checks reviews, critic's ratings and displays an online rating for every movie. Here we propose an online system that automatically allows users to post reviews and stores them to rate movies based on user sentiments. The system now analyzes this data to check for user sentiments associated with each comment. Our system consists of a sentiment library designed for English as well as Hindi sentiment analysis. The system breaks user comments to check for sentimental keywords and predicts user sentiment associated with it. Once the keywords are found it associates the comment with a sentiment rank. The system now gathers all comments for a particular movie and then calculates an average ranting to score it. This score is generated for every movie in the system. The system also sorts and displays top rating movies as per analysis and calculates a top ten list automatically. This provides an automated movie rating system based on sentiment analysis.

Objective:-

1) Define Movie
2) Recording information about the movies that are released.
3) Generating reviews.
4) Recording information related to both critic and user review.
5) Keeping record yearly of all movies.
6) Keeping information about users who register and use the recommendation system.

1.2 Literature Survey

2.1 There are numerous IEEE papers which provides plan concerning the Recommender Systems. For our system we've got referred papers as follows:

A Review Classification of Recommender Systems analysis the primary paper we tend to referred may be an analysis paper. In this have known 164 articles on recommender systems, that area unit printed from 2001 to 2009 to know the trend of recommender systems analysis and to produce practitioners and researchers with insight and future direction on recommender systems. Conjointly we've got tacit some important points. There are a unit solely VI articles for image, moving-picture show and television program recommendation. So, a lot of researches area unit needed to be studied for this. The approaches exploitation social network analysis ought to be developed within the recommender systems as recently social network analysis has been employed in the varied applications.

1.2.Scope

- Can be additional extended for generating reviews associated with the product in on-line Shops.
- Can also be used for generating reviews for the net videos.
- Can also be used for generating reviews associated with the universities throughout admission.
- Can also be used for generating reviews of the candidates within the election.

1.3 MODULES:

We use MySQL database.

The entire project mainly consists of 3 modules, which are

❖ Admin module
❖ User module (Registration)

❖ Movie Review Module

1.3.1 Admin module:

1. Dashboard: In this section, admin can view total number of users, movies according to years and their genres.
2. Movies: In this section, admin can add any movies specialization and mange according to requirement. (Add/Update).
3. Users: In this section, admin can view users detail and also have right to delete irrelevant user.
4. Contact us Queries: In this section, admin can view queries which are send by users.
5. User Session Logs: In this section, admin can see login and logout time of user.

Admin can also change his/her own password.

1.3.2 User module (Registered):

1. Dashboard: In this section, patients can view his/her profile, and previously watched movies.
2. User can update his/her profile, change the password and recover the password.
3. They can see their favorites and saved movies.

CHAPTER 2

REQUIREMENT-SPECIFICATION

2.1 INTRODUCTION:

To be used efficiently, all computer software needs certain hardware components or the other software resources to be present on a computer. These pre-requisites are known as (computer) system requirements and are often used as a guideline as opposed to an absolute rule. Most software defines two sets of system requirements: minimum and recommended. With increasing demand for higher processing power and resources in newer versions of software, system requirements tend to increase over time. Industry analysts suggest that this trend plays a bigger part in driving upgrades to existing computer systems than technological advancements.

2.2 HARDWARE REQUIREMENTS:

The most common set of requirements defined by any operating system or software application is the physical computer resources, also known as hardware. A hardware requirements list is often accompanied by a hardware compatibility list (HCL), especially in case of operating systems. An HCL lists tested, compatibility and sometimes incompatible hardware devices for a particular operating system or application. The following sub-sections discuss the various aspects of hardware requirements.

HARDWARE REQUIREMENTS FOR PRESENT PROJECT:

PROCESSOR   :   Intel dual Core, i3

RAM   :   1 GB

HARD DISK   :   80 GB

2.3 SOFTWARE REQUIREMENTS:

Software Requirements deal with defining software resource requirements and pre-requisites that need to be installed on a computer to provide optimal functioning of an application. These requirements or pre-requisites are generally not included in the software installation package and need to be installed separately before the software is installed.

SOFTWARE REQUIREMENTS FOR PRESENT PROJECT:

OPERATING SYSTEM   :   Windows 7/ XP/8/10

FRONT END   :   HTML, CSS, JavaScript, Bootstrap

BACK END   :   Python, Django

SERVER   :   Apache Tomcat Server

DATABASE   :   MySQL

CHAPTER 3

ANALYSIS

3.1 EXISTING SYSTEM:

Currently movies are recommended only taking in the pre-provided critic reviews. The current system requires numerous sources, with data stores spread throughout. Often information is incomplete. Vital information and keywords amiss so it prevents the algorithm from searching movies suited to the needs of the user.

3.2 PROPOSED SYSTEM:

The Movie Recommendation System is designed for an easier and smoother way to browse through to replace their existing system. The new system is to control and add the information through data scrapping which will be fully automated. These services are to be provided in an efficient manner, with the goal of reducing the time and resources currently required for such tasks.

3.3 SOFTWARE SPECIFICATION

HTML:

HTML or **Hypertext Markup Language** is the standard markup language used to create web pages.

HTML is written in the form of HTML elements consisting of *tags* enclosed in angle brackets (like <html>). HTML tags most commonly come in pairs like <h1> and </h1>, although some tags represent *empty elements* and so are unpaired, for example <imp>. The first tag in a pair is the *start tag*, and the second tag is the *end tag* (they are also called *opening tags* and *closing tags*). Though not always necessary, it is best practice to append a slash to tags which are not paired with a closing tag.

The purpose of a web browser is to read HTML documents and compose them into visible or audible web pages. The browser does not display the HTML tags, but uses the tags to interpret the content of the page. HTML describes the structure of a website

semantically along with cues for presentation, making it a markup language rather than a programming language.

HTML elements form the building blocks of all websites. HTML allows images and objects to be embedded and can be used to create interactive forms. It provides a means to create structured documents by denoting structural semantics for text such as headings, paragraphs, lists, links, quotes and other items. It can embed scripts written in languages such as JavaScript which affect the behavior of HTML web pages.

CASCADING STYLE SHEETS (**CSS**):

It is a style sheet language used for describing the look and formatting of a document written in a markup language. While most often used to style web pages and interfaces written in HTML and XHTML, the language can be applied to any kind of XML document, including plain XML, SVG and XUL. CSS is a cornerstone specification of the web and almost all web pages use CSS style sheets to describe their presentation.

CSS is designed primarily to enable the separation of document content from document presentation, including elements such as the layout, colors, and fonts.[1] This separation can improve content accessibility, provide more flexibility and control in the specification of presentation characteristics, enable multiple pages to share formatting, and reduce complexity and repetition in the structural content.

CSS can also allow the same markup page to be presented in different styles for different rendering methods, such as on-screen, in print, by voice (when read out by a speech-based browser or screen reader) and on Braille-based, tactile devices. It can also be used to allow the web page to display differently depending on the screen size or device on which it is being viewed. While the author of a document typically links that document to a CSS file, readers can use a different style sheet, perhaps one on their own computer, to override the one the author has specified. However if the

author or the reader did not link the document to a specific style sheet the default style of the browser will be applied.

MySQL:

MySQL is developed, distributed, and supported by Oracle Corporation. MySQL is a database system used on the web it runs on a server. MySQL is ideal for both small and large applications. It is very fast, reliable, and easy to use. It supports standard SQL. MySQL can be compiled on a number of platforms.

The data in MySQL is stored in tables. A table is a collection of related data, and it consists of columns and rows. Databases are useful when storing information categorically.

FEATURES OF MySQL:

Internals and portability:

- Written in C and C++.

- Tested with a broad range of different compilers.

- Works on many different platforms.
- Tested with Purify (a commercial memory leakage detector) as well as with Val grind, a GPL tool.
- Uses multi-layered server design with independent modules.

Security:
- A privilege and password system that is very flexible and secure, and that enables host-based verification.

- Password security by encryption of all password traffic when you connect to a server.

Scalability and Limits:

- Support for large databases. We use MySQL Server with databases that contain 50 million records. We also know of users who use MySQL Server with 200,000 tables and about 5,000,000,000 rows.

- Support for up to 64 indexes per table (32 before MySQL 4.1.2). Each index may consist of 1 to 16 columns or parts of columns. The maximum index width is 767 bytes for **InnoDB** tables, or 1000 for **MyISAM**; before MySQL 4.1.2, the limit is 500 bytes. An index may use a prefix of a column for **CHAR**, **VARCHAR**, **BLOB**, or **TEXT** column types.

WHY TO USE MySQL:

- Leading open source RDBMS
- Ease of use – No frills
- Fast
- Robust
- Security
- Multiple OS support
- Free
- Technical support
- Support large database– up to 50 million rows, file size limit up to 8 Million TB

JAVASCRIPT:

JavaScript is the scripting language of the Web. All modern HTML pages are using JavaScript. A scripting language is a lightweight programming language. JavaScript code can be inserted into any HTML page, and it can be executed by all types of web browsers. JavaScript is easy to learn.

WHY TO USE JAVASCRIPT:

JavaScript is one of the 3 languages all web developers must learn:

1. HTML to define the content of web pages
2. CSS to specify the layout of web pages
3. JavaScript to specify the behavior of web pages

Example

x = document.getElementById("demo"); //Find the HTML element with id="demo"
x.innerHTML = "Hello JavaScript";   //Change the content of the HTML element

document.getElementById() is one of the most commonly used HTML DOM methods.

JAVASCRIPT CODE:

- JavaScript code (or just JavaScript) is a sequence of JavaScript statements.
- Each statement is executed by the browser in the sequence they are written.
- This example will manipulate two HTML elements:
- Example
- document.getElementById("demo").innerHTML="Hello Dolly";
  document.getElementById("myDIV").innerHTML="How are you?";

JAVASCRIPT PROPERTIES:

- Properties are the values associated with a JavaScript object.
- A JavaScript object is a collection of unordered properties.
- Properties can usually be changed, added, and deleted, but some are read only.
-

PYTHON:

What is Python?

Python is a popular programming language. It was created by Guido van Rossum, and released in 1991.

It is used for:

- Web development (server-side),
- Software development,
- Mathematics,
- System scripting.

What can Python do?

- Python can be used on a server to create web applications.
- Python can be used alongside software to create workflows.
- Python can connect to database systems. It can also read and modify files.
- Python can be used to handle big data and perform complex mathematics.
- Python can be used for rapid prototyping, or for production-ready software development.

Why Python?

- Python works on different platforms (Windows, Mac, Linux, Raspberry Pi, etc).
- Python has a simple syntax similar to the English language.
- Python has syntax that allows developers to write programs with fewer lines than some other programming languages.
- Python runs on an interpreter system, meaning that code can be executed as soon as it is written. This means that prototyping can be very quick.
- Python can be treated in a procedural way, an object-oriented way or a functional way.

print("Hello, World!")

Output

Hello, World!

BOOTSTRAP:

WHAT IS BOOTSTRAP?

Bootstrap is the most popular **CSS Framework** for developing responsive and mobile-first websites.

Bootstrap is a potent front-end framework used to create modern websites and web apps. It's open-source and free to use, yet features numerous HTML and CSS templates for UI interface elements such as buttons and forms. Bootstrap also supports JavaScript extensions.

WHY BOOTSTRAP?

Software engineers use Bootstrap for a number of different reasons:

- It is easy to set up and master
- It has a lot of components and a good grid system.
- Styling for many HTML elements ranging from typography to buttons, as well as support of JavaScript plugins, makes it very flexible.

WHAT ARE THE USES OF BOOTSTRAP?

- Bootstrap is great for creating layouts, as its responsive CSS is designed to conform to different devices.

- It can be employed to ensure consistency, eliminate cross-browser issues, and so on.

Chapter 4

DATABASE DESIGN

The data in the system has to be stored and retrieved from database. Designing the database is part of system design. Data elements and data structures to be stored have been identified at analysis stage. They are structured and put together to design the data storage and retrieval system.

A database is a collection of interrelated data stored with minimum redundancy to serve many users quickly and efficiently. The general objective is to make database access easy, quick, inexpensive and flexible for the user. Relationships are established between the data items and unnecessary data items are removed. Normalization is done to get an internal consistency of data and to have minimum redundancy and maximum stability. This ensures minimizing data storage required, minimizing chances of data inconsistencies and optimizing for updates. The MS Access database has been chosen for developing the relevant databases.

Movie Recommendation System (MRS) MySQL tables:

Admin table Structure: This table store the login details of admin.

#	Name	Type	Collation	Attributes	Null	Default	Comments	Extra
1	id 🔑	int(11)			No	None		AUTO_INCREMENT
2	username	varchar(255)	latin1_swedish_ci		No	None		
3	password	varchar(255)	latin1_swedish_ci		No	None		
4	updationDate	varchar(255)	latin1_swedish_ci		No	None		

User log table Structure: This table store the user's login and personal details.

#	Name	Type	Collation	Attributes	Null	Default	Comments	Extra
1	id 🔑	int(11)			No	None		AUTO_INCREMENT
2	uid	int(11)			Yes	NULL		
3	username	varchar(255)	latin1_swedish_ci		Yes	NULL		
4	userip	binary(16)			Yes	NULL		
5	loginTime	timestamp			Yes	current_timestamp()		
6	logout	varchar(255)	latin1_swedish_ci		Yes	NULL		
7	status	int(11)			Yes	NULL		

Users table Structure: This table store the user's login and personal details.

| #	Name	Type	Collation	Attributes	Null	Default	Comments	Extra
1	id 🔑	int(11)			No	None		AUTO_INCREMENT
2	fullName	varchar(255)	latin1_swedish_ci		Yes	NULL		
3	address	longtext	latin1_swedish_ci		Yes	NULL		
4	city	varchar(255)	latin1_swedish_ci		Yes	NULL		
5	gender	varchar(255)	latin1_swedish_ci		Yes	NULL		
6	email 🔑	varchar(255)	latin1_swedish_ci		Yes	NULL		
7	password	varchar(255)	latin1_swedish_ci		Yes	NULL		
8	regDate	timestamp			Yes	current_timestamp()		
9	updationDate	timestamp			Yes	NULL		ON UPDATE CURRENT_TIMESTAMP()

tblcontactus table Structure: This table store the contact us query details .

#	Name	Type	Collation	Attributes	Null	Default	Comments	Extra
1	id 🔑	int(11)			No	None		AUTO_INCREMENT
2	fullname	varchar(255)	latin1_swedish_ci		Yes	NULL		
3	email	varchar(255)	latin1_swedish_ci		Yes	NULL		
4	contactno	bigint(12)			Yes	NULL		
5	message	mediumtext	latin1_swedish_ci		Yes	NULL		
6	PostingDate	timestamp			Yes	current_timestamp()		
7	AdminRemark	mediumtext	latin1_swedish_ci		Yes	NULL		
8	LastupdationDate	timestamp			Yes	NULL		ON UPDATE CURRENT_TIMESTAMP()
9	IsRead	int(11)			Yes	NULL		

CHAPTER 5

SAMPLE SCREENSHOTS

HOME PAGE
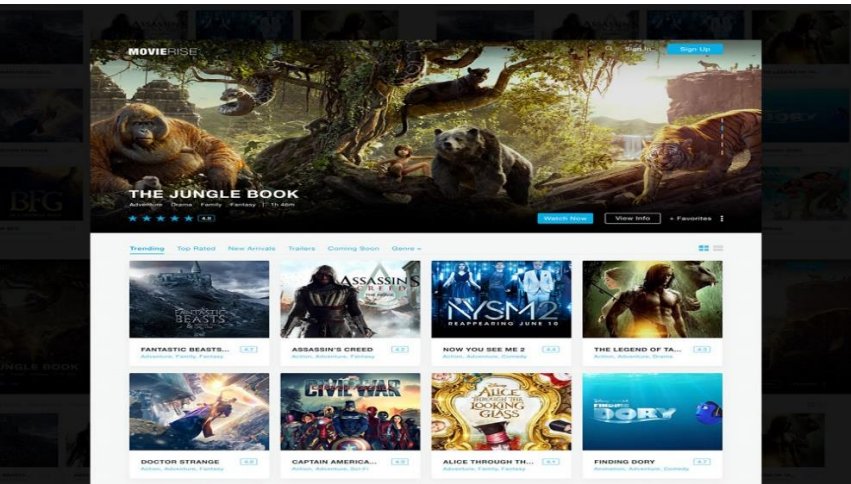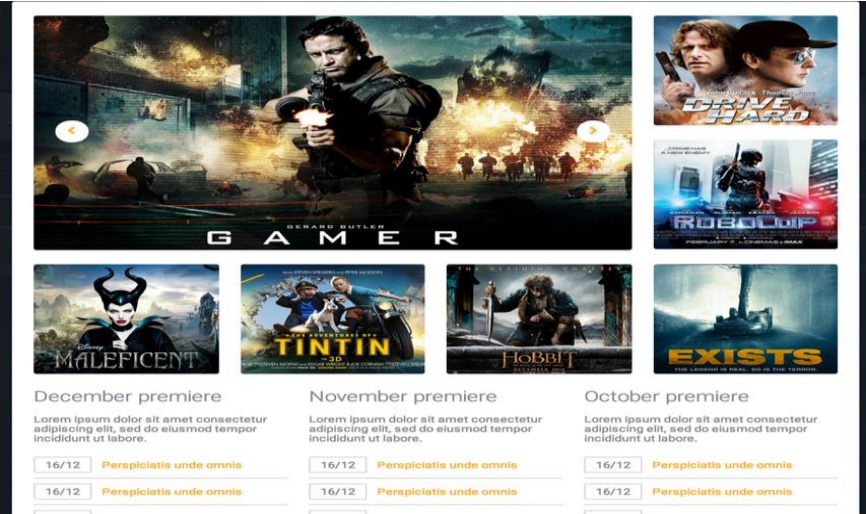
DATA COLLECTION
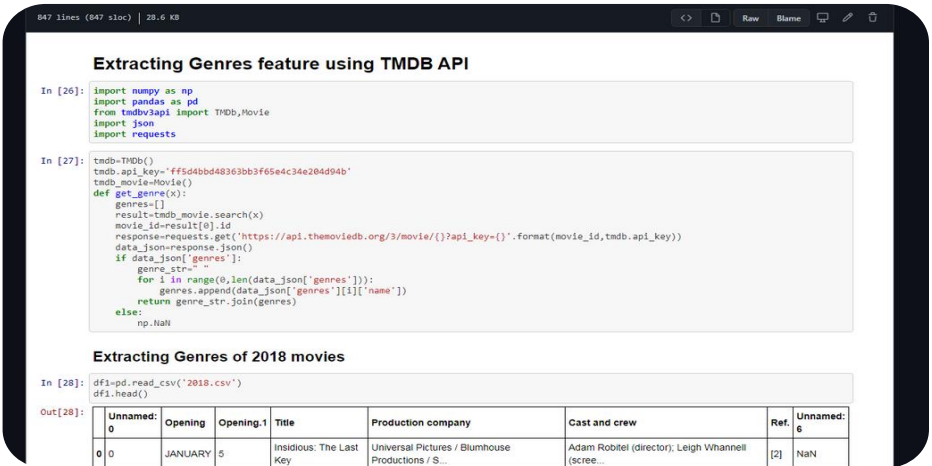
DATA PRE-PROCESSING

```
In [31]: from google.colab import drive
         drive.mount('/content/gdrive')

         Drive already mounted at /content/gdrive; to attempt to forcibly remount, call drive.mount("/content/gdrive", force_remount=True).

In [32]: import pandas as pd
         import numpy as np
         import matplotlib.pyplot as plt
         import seaborn

In [33]: df = pd.read_csv('/content/gdrive/My Drive/College_project/movie_metadata.csv')

In [34]: df.head()
```

Out[34]:

	color	director_name	num_critic_for_reviews	duration	director_facebook_likes	actor_3_facebook_likes	actor_2_name	actor_1_facebook_likes	gross
0	Color	James Cameron	723.0	178.0	0.0	855.0	Joel David Moore	1000.0	7605058
1	Color	Gore Verbinski	302.0	169.0	563.0	1000.0	Orlando Bloom	40000.0	30940418
2	Color	Sam Mendes	602.0	148.0	0.0	161.0	Rory Kinnear	11000.0	2000741
3	Color	Christopher Nolan	813.0	164.0	22000.0	23000.0	Christian Bale	27000.0	4481306
4	NaN	Doug Walker	NaN	NaN	131.0	NaN	Rob Walker	131.0	NaN

```
Out[35]: Index(['color', 'director_name', 'num_critic_for_reviews', 'duration',
                'director_facebook_likes', 'actor_3_facebook_likes', 'actor_2_name',
                'actor_1_facebook_likes', 'gross', 'genres', 'actor_1_name',
                'movie_title', 'num_voted_users', 'cast_total_facebook_likes',
                'actor_3_name', 'facenumber_in_poster', 'plot_keywords',
                'movie_imdb_link', 'num_user_for_reviews', 'language', 'country',
                'content_rating', 'budget', 'title_year', 'actor_2_facebook_likes',
                'imdb_score', 'aspect_ratio', 'movie_facebook_likes'],
               dtype='object')

         For movie Recomendation we need only few columns like

            1. Director_name
            2. actor_1_name
            3. actor_2_name
            4. actor_3_name
            5. genres
            6. movie_list

In [36]: df = df.loc[:,['director_name','actor_1_name','actor_2_name','actor_3_name','genres','movie_title']]

In [37]: df.head()
```

Out[37]:

	director_name	actor_1_name	actor_2_name	actor_3_name	genres	movie_title
0	James Cameron	CCH Pounder	Joel David Moore	Wes Studi	Action|Adventure|Fantasy|Sci-Fi	Avatar
1	Gore Verbinski	Johnny Depp	Orlando Bloom	Jack Davenport	Action|Adventure|Fantasy	Pirates of the Caribbean: At World's End
2	Sam Mendes	Christoph Waltz	Rory Kinnear	Stephanie Sigman	Action|Adventure|Thriller	Spectre
3	Christopher Nolan	Tom Hardy	Christian Bale	Joseph Gordon-Levitt	Action|Thriller	The Dark Knight Rises
4	Doug Walker	Doug Walker	Rob Walker	NaN	Documentary	Star Wars: Episode VII - The Force Awakens ...

```
In [38]: df['director_name'].value_counts()
Out[38]: Steven Spielberg    26
         Woody Allen         22
```

actor_1_name	actor_2_name	actor_3_name	genres
CCH Pounder	Joel David Moore	Wes Studi	Action Adventure Fantasy science friction
Johnny Depp	Orlando Bloom	Jack Davenport	Action Adventure Fantasy
Christoph Waltz	Rory Kinnear	Stephanie Sigman	Action Adventure Thriller
Tom Hardy	Christian Bale	Joseph Gordon Levitt	Action Thriller
Doug Walker	Rob Walker		Documentary
Daryl Sabara	Samantha Morton	Polly Walker	Action Adventure science friction
J.K. Simmons	James Franco	Kirsten Dunst	Action Adventure Romance
Brad Garrett	Donna Murphy	M.C. Gainey	Adventure Animation Comedy Family Fantasy I
Chris Hemsworth	Robert Downey Jr.	Scarlett Johansson	Action Adventure science friction
Alan Rickman	Daniel Radcliffe	Rupert Grint	Adventure Family Fantasy Mystery
Henry Cavill	Lauren Cohan	Alan D. Purwin	Action Adventure science friction
Kevin Spacey	Marlon Brando	Frank Langella	Action Adventure science friction
Giancarlo Giannini	Mathieu Amalric	Rory Kinnear	Action Adventure

Sobhita DhulipalaMrunal ThakurAvinash TiwaryJanhvi KapoorSurekha SikriRaghuvir YadavGulshan DevaiahVijay VarmaPavail GulatiJyou

Sunny KaushalRukshar DhillonShriya Pilgaonkar

Akshaye KhannaRiva KishanPriyank Sharma

Hema MaliniRajkummar RaoRakul Preet Singh

Deepika PadukoneVikrant Massey

Ajay DevgnSaif Ali KhanKajolSharad Kelkar

Abhijeet SinghArindita Kalita

Sunny SinghSonnalli SeygallPoonam DhillonSupriya Pathak

Kangana RanautJassi GillRicha ChadhaNeena Gupta

Varun DhawanShraddha KapoorPrabhu DevaNora FatehiAparshakti KhuranaMurali SharmaRaghav JuyalDharmesh YelandePunit Pathak

Reem ShaikhAtul KulkarniDivya DuttaOm PuriArif ZakariaMukesh RishiAbhimanyu SinghPankaj Tripathi

Himesh ReshammiyaSonia Mann

Saif Ali KhanTabuAlaya Furniturewala

6.   CONCLUSION:

Project highlights

UNCOVERED KEY RESEARCH TRENDS

This solution is carefully analysed and after the evaluation of the possible solutions, the most feasible solution for this project is identified and selected, so the project turns to be cost-effective, vital and practical.

Thus processing information will be faster. It guarantees accurate maintenance of details. It easily reduces the book keeping task and thus reduces the human effort and increases accuracy speed.