Technology Review

Bianca Zlavog
Florencia Marcaccio
Mansi Rathod
Sanjana Gupta

Outline

- Project Overview
- Use Cases
- Visualization Tools
- Tech Comparison
- Python Libraries

Project Overview

With so many streaming platforms available, it can be overwhelming to browse every catalog to find a movie that you'll like.

Our goal is to provide a personalized streaming experience based on the user's preferences with a movie recommendation tool.

Data Sources:

- Netflix Dataset: Training_set.rar, movie_titles.txt (https://www.kaggle.com/netflix-inc/netflix-prize-data)
- IMDB Dataset: title.basics.tsv.gz, title.ratings.tsv.gz (https://www.imdb.com/interfaces/)

Use Case 1

The user wants to watch a movie similar to a movie of their choice.

User: User inputs his/her liked movie.

Tool: Based on the user's liked movie, a list of recommended movies will be displayed.

Recommended movies based on *Lilo and Stitch*
Ice Age
The Emperor's New Groove
A Bug's Life
Monsters
Atlantis: The Lost Empire
Finding Nemo (Widescreen)
The Lion King: Special Edition
Shrek (Full-screen)
Aladdin: Platinum Edition
Brother Bear (Theatrical Widescreen Version)

Use Case 2

The user wants to see the top 10 movies of a particular year.

User: User inputs a particular year.

Tool: A list of the top 10 most recommended movies will be displayed.

Use Case 3

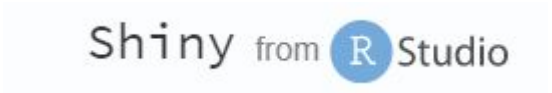
The user wants to see the top 10 movies from a particular genre.

User: User inputs the genre from a drop-down menu.

Tool: A list of the top 10 most recommended movies from that genre will be displayed.

Visualization Tools

- Bokeh (https://docs.bokeh.org/en/latest/index.html) : Interactive Visualizations
- Shiny (https://shiny.rstudio.com/) : Interactive Visualizations
- Dash (http://dash.plotly.com/) : Interactive Visualizations



Technical Comparison

Bokeh

Pros:

- Compatible with Jupyter notebooks
- Wider range of interactions

Cons:

- No proper documentation and well-supported platforms
- Slower for big datasets
- Interaction inconsistency

Dash

Pros:

- Multiple tutorials available
- Extensive documentation
- Implemented purely in Python
- Plotly compatible

Cons:

- No experience with it beyond testing
- Limited customization

Technical Comparison

Shiny

Pros:

- Can be hosted for free on shinyapps.io
- Have experience working with it
- More flexible with customizations
- More mature than dash - has a larger community following
- No knowledge of HTML/CSS/JavaScript required

Cons:

- Could have compatibility issues with Python

Python Libraries Used

- Pandas (https://pandas.pydata.org/)
- Scikit-learn (https://scikit-learn.org/)
- Numpy (https://numpy.org/)
- Scipy (https://docs.scipy.org/)

THANK YOU!