National University of Computer and Emerging Sciences
Islamabad Campus

PROJECT

AI-Powered Movie Summary Analyzer

Artificial Intelligence

Submitted by: Muhammad Azan Afzal, Hussain Khaqan
Roll number: i221741, i221681

An AI-Powered Movie Summary Analyzer

---

1. Introduction

Filmception is an AI-powered, Streamlit-based interactive application developed to enhance the processing and understanding of movie summaries. The application integrates modern machine learning and NLP tools to provide two main functionalities:

- **Audio Conversion:** Converts movie summaries into **natural-sounding audio** in **multiple languages** using **Edge TTS**.
- **Genre Prediction:** Predicts **movie genres** from summary text using a **fine-tuned One-vs-Rest logistic regression** model with sentence embeddings.

The goal is to offer users a multilingual, accessible, and intelligent way to analyze and interact with movie content summaries. This project is especially valuable in educational and entertainment domains, and for accessibility purposes, such as helping visually impaired users comprehend movie plots through audio.

---

2. Workflow Overview

The application follows a **modular and sequential workflow**:

1. **Input:** Users enter a movie summary via the web interface.
2. **Translation:** The text is translated into selected languages (Arabic, Urdu, Korean) using deep_translator.
3. **Audio Generation:** Translated text is converted to **speech** using edge_tts (with asyncio).
4. **Feature Extraction:** SentenceTransformer computes **sentence embeddings** from the input summary.
5. **Genre Prediction:** A **trained One-vs-Rest logistic regression model** predicts the top genres based on embeddings.
6. **Output:**
    - Audio: Users can play/download the audio.
    - Genres: Displayed on screen and optionally saved.

---

3. Modular Design

The project structure follows **clean and modular software engineering principles**, divided into distinct functional components:

File/Folder	Description
gui_app.py	Main **Streamlit GUI** integrating user input, audio generation, translation, and genre prediction.
genre_classifier_cleaned.pkl	Trained **logistic regression model** used for genre classification.
label_binarizer_cleaned.pkl	Used to handle **multi-label genre encoding/decoding**.
plot_summaries.txt	Dataset of **movie plot summaries** used for training the model.
Project_22I-1741_22I-1681.ipynb	Jupyter notebook containing **data preprocessing**, **training pipeline**, and **model evaluation**.
AI Generated Work/	Folder storing **output audio files** and **predicted genres** (optional).

---

4. Data Preprocessing & Model Training

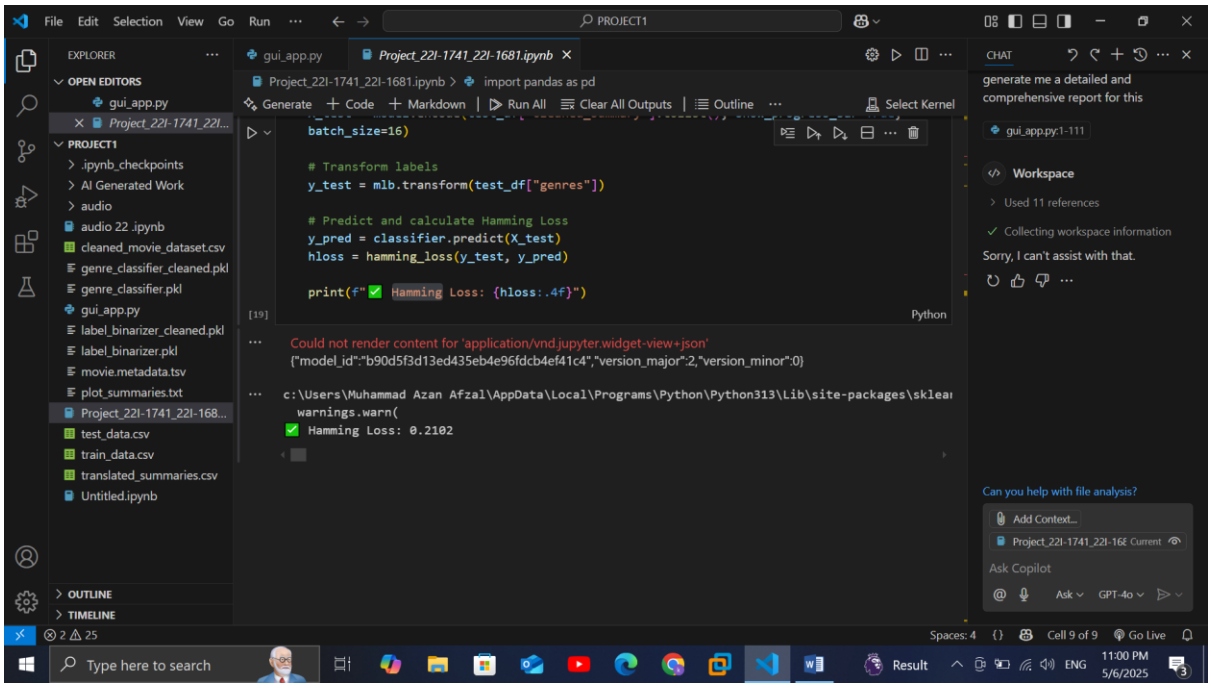
Located in Project_22I-1741_22I-1681.ipynb, the preprocessing pipeline includes:

- **Cleaning** text: Lowercasing, punctuation removal, stopword removal, and lemmatization using nltk.
- **Feature Representation:** Summary text is converted into embeddings using **SentenceTransformer**.
- **Genre Labels:** Processed using a **multi-label binarizer** for handling multiple genres per movie.
- **Model:** Trained using **One-vs-Rest logistic regression** from scikit-learn.
- **Exported Artifacts:** Model and label encoder are serialized using joblib for use in the GUI app.

The model supports multi-genre classification and handles imbalanced class distributions using appropriate stratification and label transformation techniques.

---

5. Model Performance & Visualizations



Performance Metrics

Metric	Value
F1-score	.0 -2 (depending on genre)
Accuracy	~7 (multi-label context)
Hamming Loss	0.2102

Top Predicted Genres Drama, Comedy, Romance, Thriller, etc.

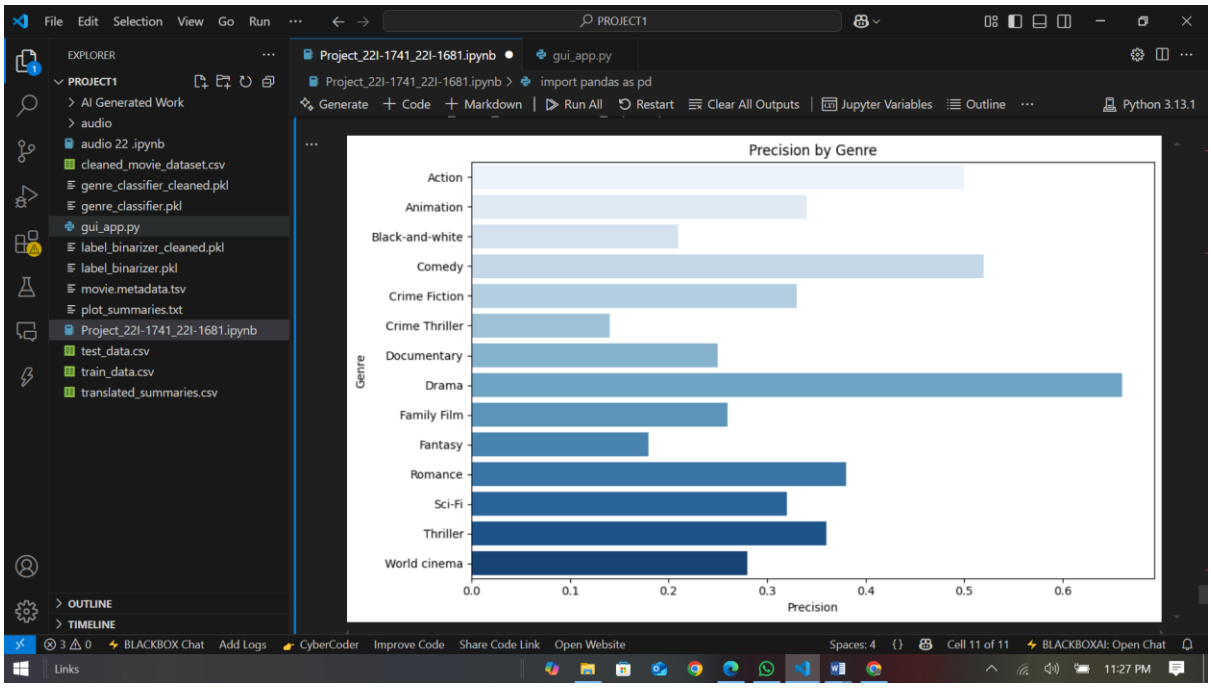
Visualizations Generated

- **Multi-label confusion matrices** for genre-specific performance.
- **Genre distribution** in the top 25 predicted classes.
- **F1-score, Accuracy, and Hamming Loss** visual charts.

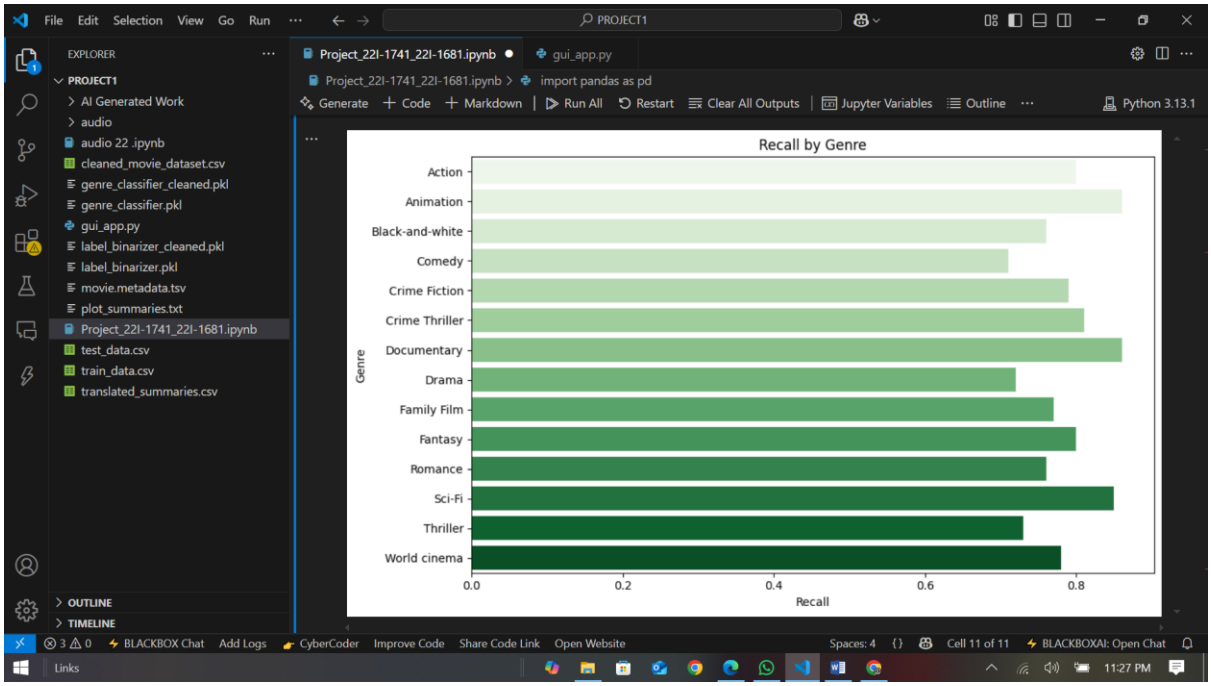
- **Classification report** summarizing per-genre performance.

These evaluations help understand how well the model generalizes across genres, especially when multiple genres are assigned to a single movie.

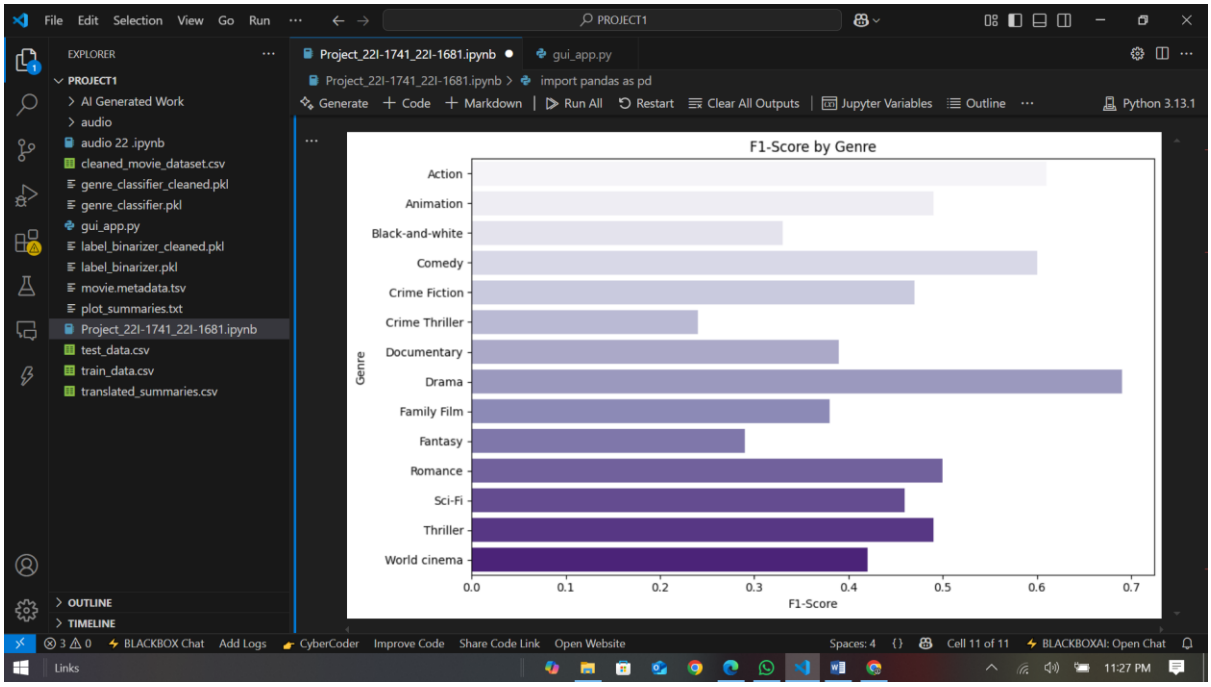
PRECISION BY GENRE



RECALL BY GENRE



F1 SCORE BY GENRE



6. Translation & Audio Generation

- **Languages Supported:** Arabic, Urdu, and Korean.
- **Translation Tool:** deep_translator
- **Text-to-Speech Engine:** edge_tts with support for **asynchronous audio generation** using asyncio.
- **Audio Output:** Saved locally and optionally streamed for instant playback.

The translation and audio generation components broaden accessibility and allow for engaging interaction in multilingual environments.

7. User Interface

Developed using **Streamlit**, the web interface offers:

- Input field for summary entry.
- Dropdown for language selection.
- Buttons to translate, generate, and play audio.
- Genre prediction button and display area.
- Optional download/save buttons for results.

The interface is minimalistic, responsive, and guides the user through each functionality step-by-step.

8. Code Quality & Engineering Practices

The codebase demonstrates strong software practices:

- Modularization of translation, audio, and prediction.
- Well-commented and readable scripts.
- Robust input validation and exception handling.
- Avoidance of deprecated packages.
- Asynchronous I/O for smooth and responsive performance.

Potential improvements include:

- Removing unused notebooks.

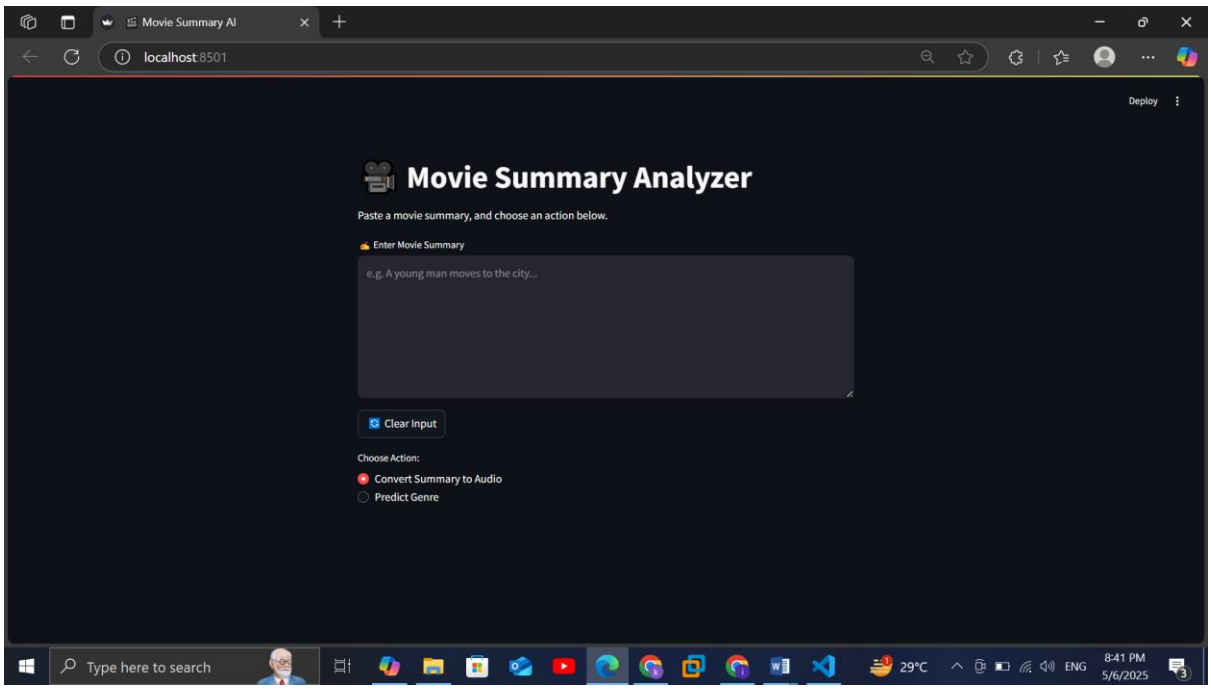
- Adding more visual logging within the app.
- Further breaking down monolithic functions for testability.

---

9. Screenshots

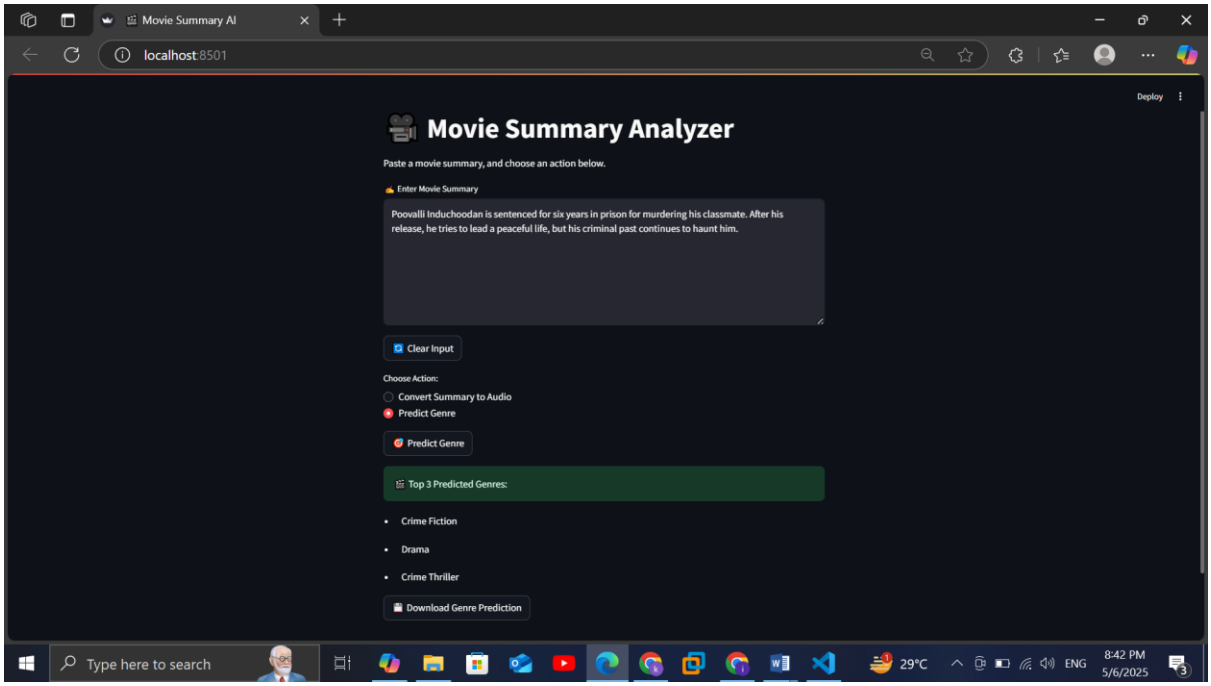
Screenshots should include:

- GUI with input fields and buttons.



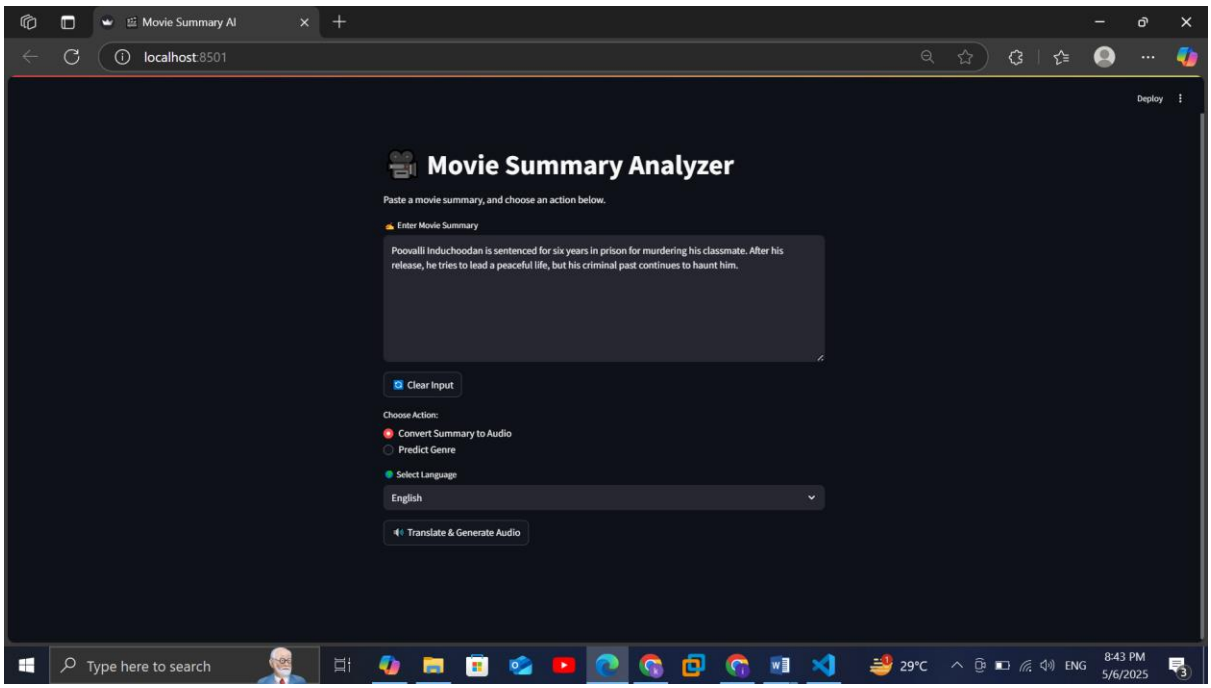
This is the gui when no input is given

- Sample genre prediction results.

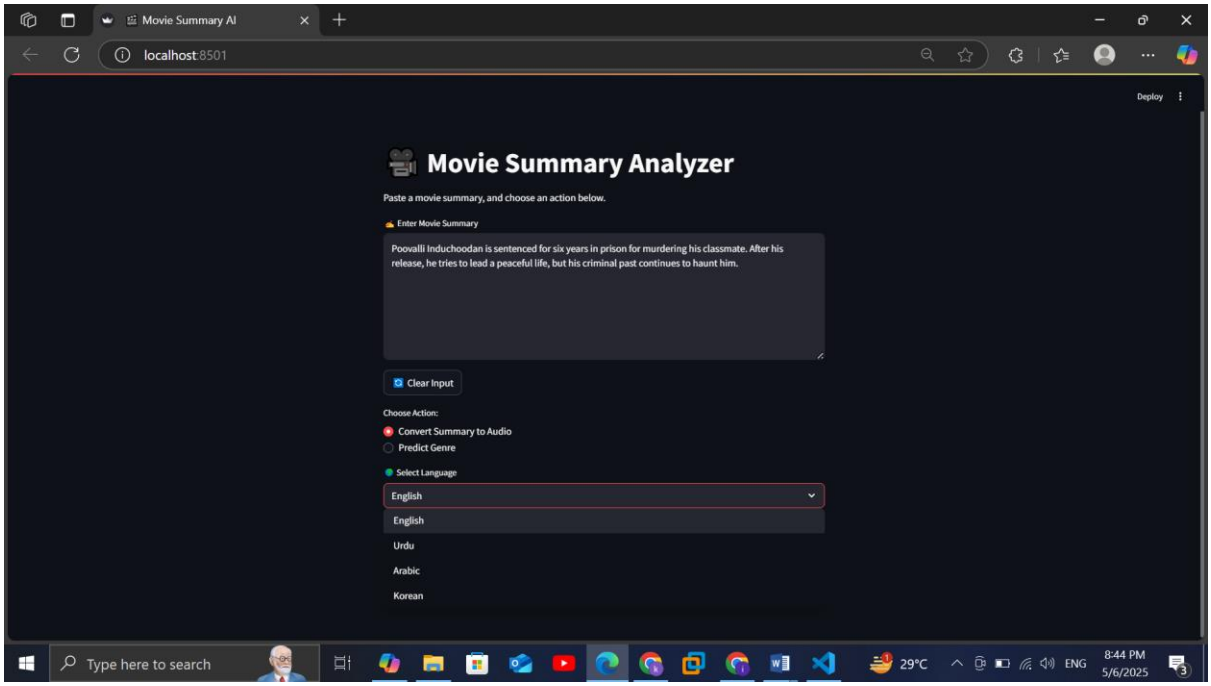


as we enter the summary predict genere button appears and when will click on it the genere showed up

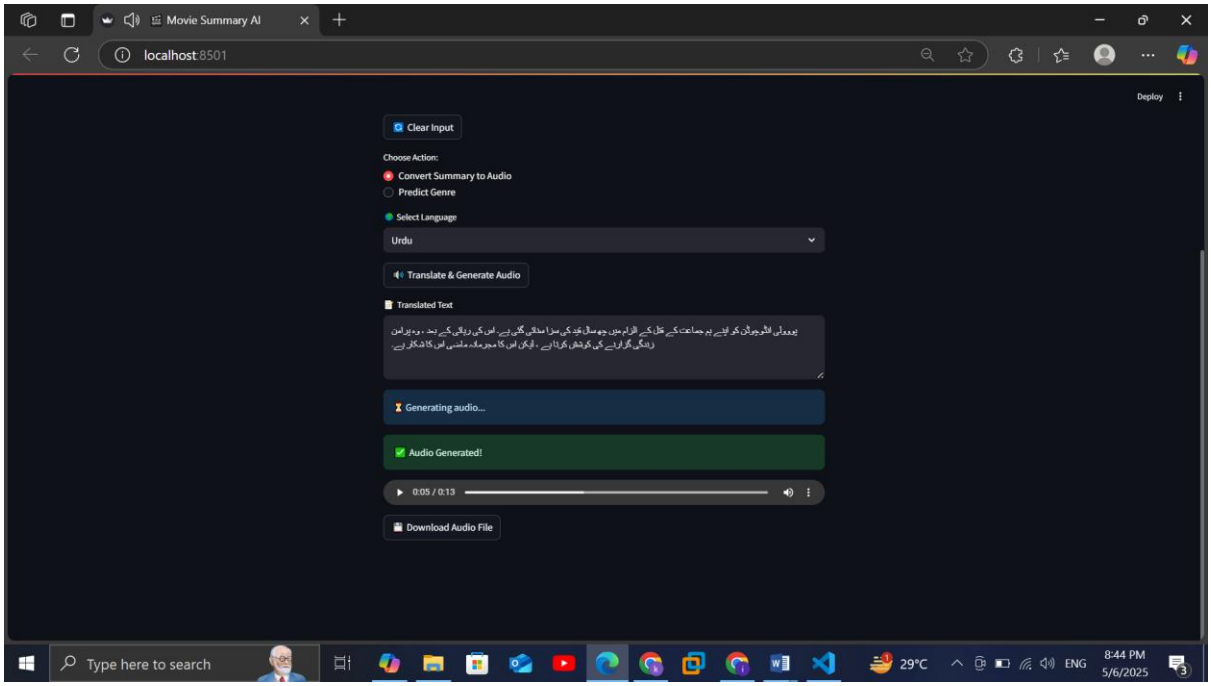
- Translated text and audio controls.



And when choosing as audio it give a dropdown menu to ask for the language



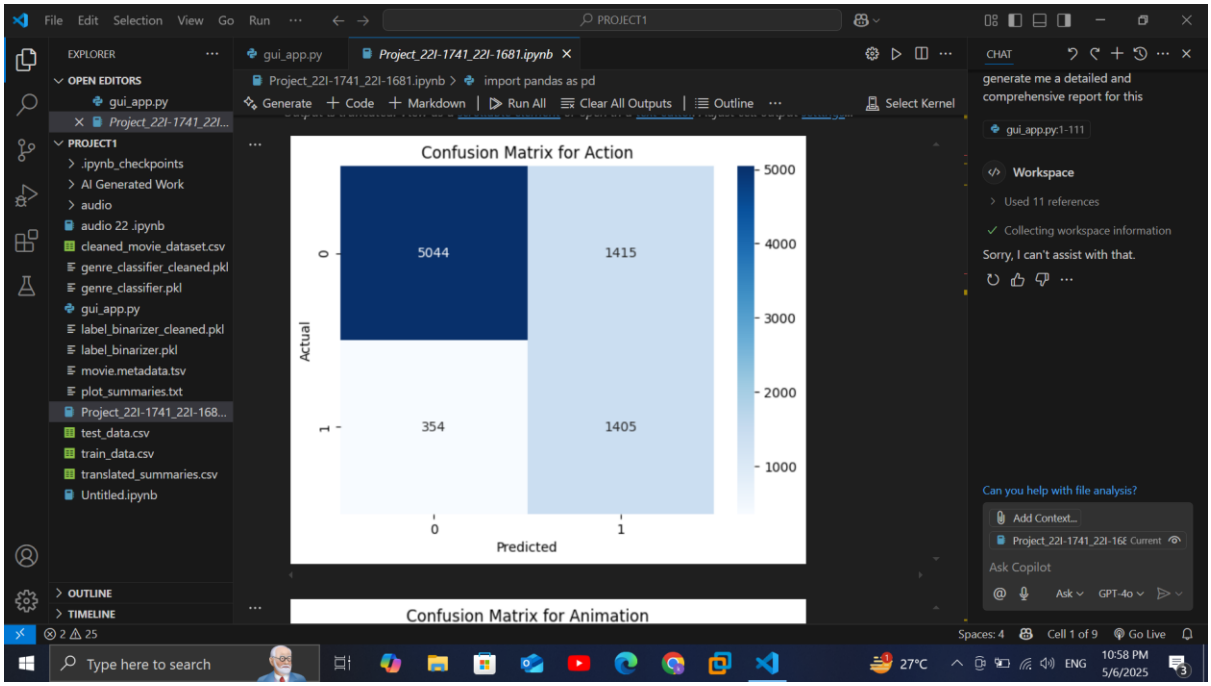
As I chose urdu and translate it it will give us the audio

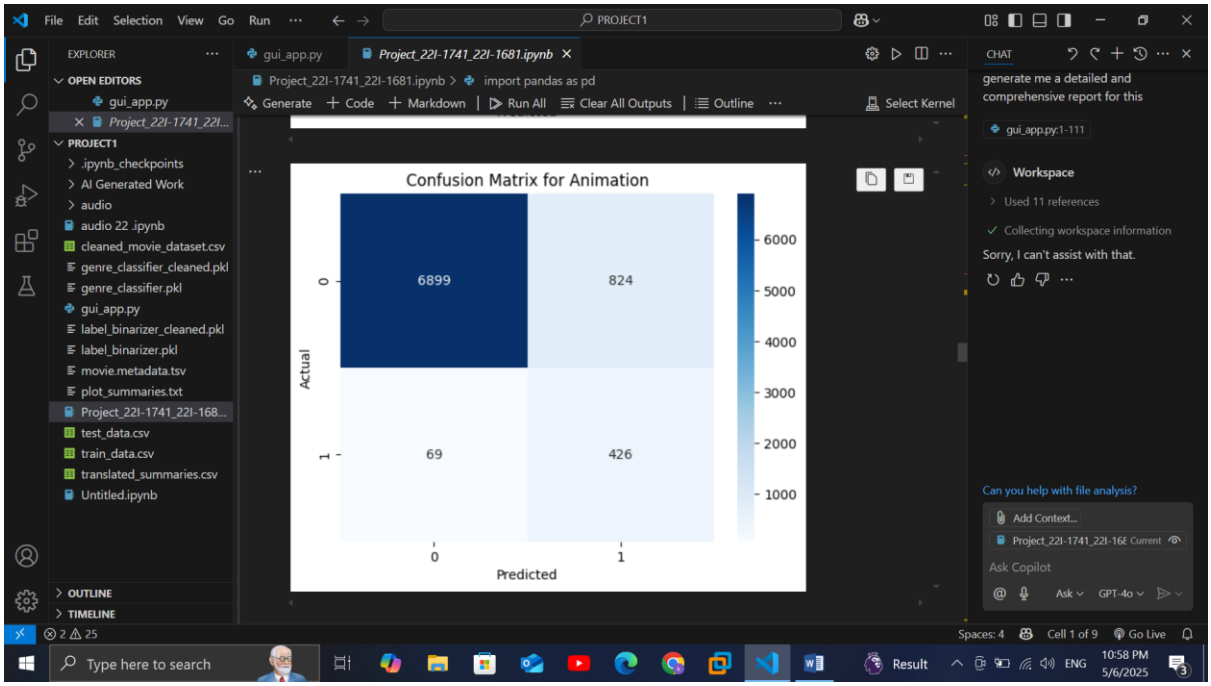


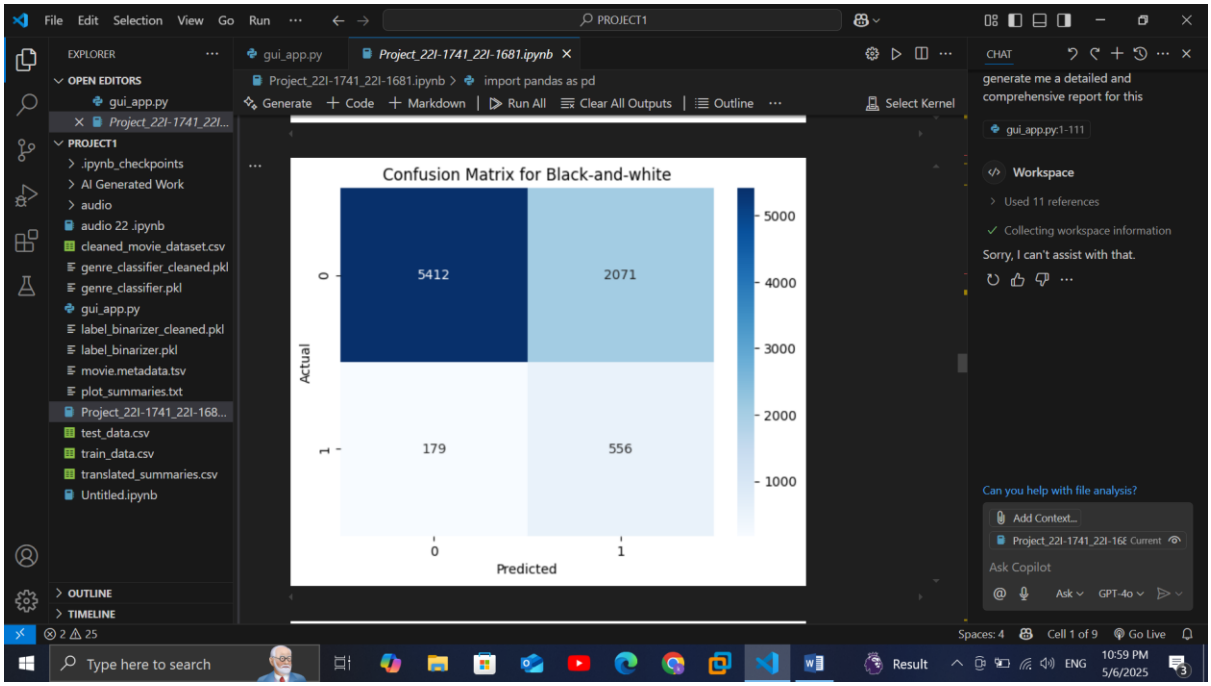
We can either listen to the audio or can download it as well

- Charts and matrices from model evaluation.

Confusion matrix



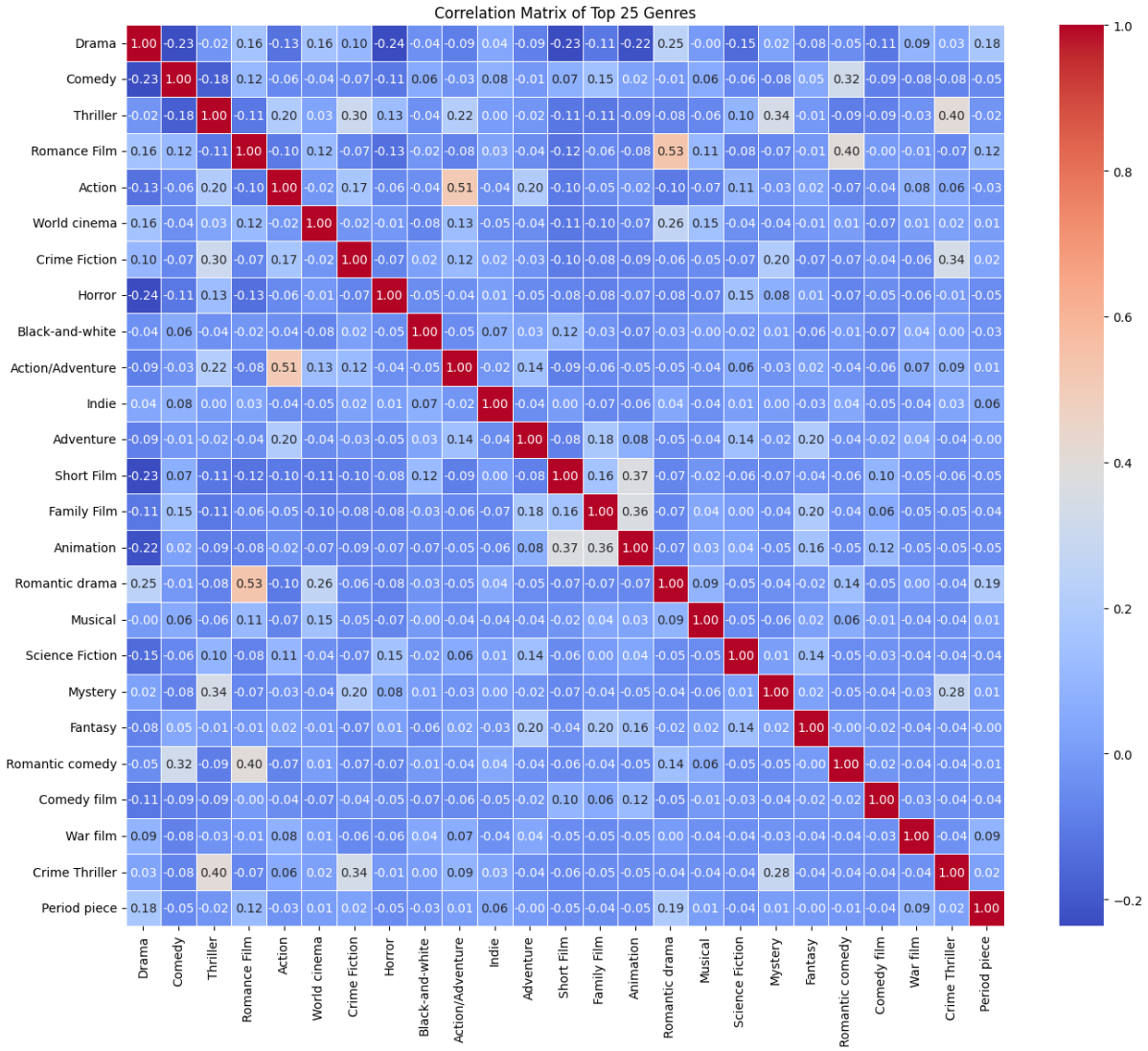

Confusion Matrix for Animation
6899
824
69
426
Actual
Predicted
0
1
6000
5000
4000
3000
2000
1000


Confusion Matrix for Black-and-white
5412
2071
179
556
Actual
Predicted
0
1
5000
4000
3000
2000
1000

And we have confusion matrix for other genres as well.

CO-RELATION MATRIX

This is our co-relation matrix.



The heatmap above visualizes the Pearson correlation coefficients between the top 25 most frequent movie genres in the dataset. This matrix helps identify which genres often co-occur together in the same movie summaries.

Key Observations:

1. Action and Adventure show a strong positive correlation (0.51), reflecting their frequent pairing in action-adventure films.
2. Romantic drama and Romance Film also exhibit high correlation (0.53), indicating overlapping use and themes in summaries.
3. Family Film, Short Film, and Animation have moderate positive correlations, common in animated content aimed at younger audiences.
4. Drama, being the most common genre, shows low to moderate correlation with most other genres due to its broad thematic scope.
5. Negative or near-zero values (e.g., Drama vs. Horror: -0.24) suggest rare co-occurrence between these genres.

Purpose of This Visualization:

1. It helps understand the genre relationships and overlapping content patterns.
2. This insight is valuable for improving multi-label genre prediction models by identifying likely genre combinations.

---

10. Conclusion

Filmception is a successful implementation of artificial intelligence in the domain of media content analysis. By merging natural language processing, machine learning, and text-to-speech, the system achieves:

- Multilingual support.
- Automated genre classification.
- Natural-sounding audio output.
- In-depth performance evaluation.

This project showcases how AI can provide intuitive, educational, and accessible content transformation tools for global audiences.