GenAI

Project – 41: Movie Genre Classification using Zero-Shot Learning

Name: Anshul Banda

SRN: PES2UG23CS081

SECTION: B

Purpose of the project:

The purpose of this project is to build a movie genre classifier that predicts whether a film belongs to the Horror, Comedy, or Action genre based on its plot summary. The project addresses the problem of classifying text without labeled training data by using a pre-trained language model. To achieve this, zero-shot classification is used, which allows the model to assign genres it was not explicitly trained on by comparing the semantic meaning of the plot with the given genre labels.

Tools and Models used:

This project uses the transformers library by Hugging Face, which provides access to pre-trained transformer models and high-level pipelines for natural language processing tasks. The library simplifies tasks such as text classification by handling tokenization, model loading, and inference.

The model used in this project is facebook/bart-large-mnli, a transformer-based model trained on natural language inference data. This training enables the model to understand relationships between text and labels, making it well-suited for zero-shot classification tasks.

Zero-shot learning was chosen because it allows text to be classified into categories without requiring labeled training data. This approach is efficient, flexible, and ideal for situations where creating or collecting labeled datasets is impractical.

Methodology:

The project was completed in multiple steps to understand how different zero-shot classification setups affect the model's predictions.

- Initial Zero-Shot Classification:
  In the first attempt, simple genre labels such as Horror, Comedy, and Action were used, and the model was forced to choose only one genre for each plot. This resulted in inaccurate predictions, with the model often selecting “Action” even when it was not the best fit. This happened because the labels were broad and overlapped in meaning, causing the model to favor a more general category.
- Improved Label Design:
  To improve the results, more descriptive labels were introduced along with a hypothesis template to give the model better context. This helped the model understand the genres more clearly and increased confidence in the predictions. While the results were better, the model was still limited by having to select only one genre.
- Multi-Label Classification (Final Solution):
  In the final step, multi-label classification was enabled so that each genre could be evaluated independently. This removed the need for the labels to compete with each other and allowed the model to decide whether a genre applied on its own. As a result, the predictions became more accurate and consistent, making this approach the most effective solution.

Results & Observations:

The final multi-label zero-shot classification approach produced accurate and consistent results across all test plots. Each movie plot was correctly matched with its intended genre, along with a clear confidence score. This confirmed that evaluating genres independently leads to more reliable predictions.

Plot Type	**Correct Genre**	**Model Prediction**
Horror Plot	Horror	Horror (96%)
Comedy Plot	Comedy	Comedy (54%)
Action Plot	Action	Action (93%)

Overall, the model performed best on clearly defined genres such as Horror and Action, while Comedy received a lower confidence score due to its more subtle cues. Despite this, the correct genre was still identified in all cases.

Conclusion:

This project showed how powerful zero-shot models can be for text classification, even without any task-specific training data. It also highlighted how much model behavior depends on label design and prompt structure. Simple labels and forced single-choice classification led to biased results, while clearer labels and multi-label evaluation produced

more accurate predictions. Overall, the project demonstrates that zero-shot learning is a flexible and practical approach, but careful design of labels and prompts is essential for reliable performance.