

UNIVERSITÀ DI PISA

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

Master's Degree in Artificial Intelligence and Data Engineering

CineSense

Decoding Audience Insights, Shaping Better Films

Group:

Martina Fabiani

Tommaso Falaschi

Rossana Antonella Sacco

Project GitHub Repository:

<https://github.com/LeBonWskii/FilmReviewsAnalysis-BusinessProject>

ACADEMIC YEAR 2024/2025

Contents

1	Introduction	3
1.1	Addressing Industry Challenges	3
1.2	The Innovation	4
1.2.1	LLMs and Review Analysis for Film Industry Enhancement	4
2	The Application	5
2.1	Key Features and Functionalities	5
2.2	PDF report	7
2.3	User Interface	11
2.3.1	Genre Selection & Processing	12
2.3.2	Real-Time Processing Log	13
2.3.3	Report Generation & Download	13
2.3.4	Data Visualization & Insights	13
2.3.5	CineSense Benefits	13
3	Implementation	14
3.1	Dataset Preprocessing	14
3.2	Gemini Long Context API	15
3.2.1	Gemini-1.5-Flash	15
4	Results	17
4.1	Validation of Sentiment Analysis and Topic Extraction	17
4.1.1	Sentiment Analysis	17
4.1.2	Topic Extraction	18
5	Competitor Analysis	19
5.1	Existing Competitors	19
5.2	The CineSense Advantage	19
6	Conclusion	21
	Bibliography	22

Chapter 1

Introduction

CineSense is an innovative application designed to support the film production and distribution process by leveraging the power of artificial intelligence. By integrating advanced **Large Language Models (LLMs)** and **sentiment analysis** techniques, CineSense analyzes audience reviews to extract valuable insights into the key factors that drive a film's success, tailored to its specific genre.

In an industry where audience reception can make or break a movie, understanding viewer sentiment is crucial. **CineSense** processes large volumes of user-generated feedback, identifying trends, strengths, and potential weaknesses. This data-driven approach helps filmmakers, producers, and distributors make more informed decisions, optimizing both creative and marketing strategies.

With its ability to identify patterns in audience perception, CineSense offers a **revolutionary** tool for the film industry, bridging the gap between viewer expectations and production choices. Whether in the early development phase, post-production, or marketing campaigns, CineSense provides actionable insights to enhance the overall success of a film.

1.1 Addressing Industry Challenges

We have examined the **existing** sentiment analysis and audience feedback evaluation techniques within the film industry. While similar methodologies have been developed, their practical application remains **limited** in real-world scenarios.

The failures of projects like "**PapMusic**" [4] and "**The Lone Ranger**" [3] highlight how a lack of audience understanding and ineffective strategic planning can **significantly** contribute to a film's downfall. PapMusic, for instance, struggled due to a misalignment between its artistic vision and audience expectations, leading to poor reception. Similarly, The Lone Ranger suffered from overinflated production costs, misguided marketing, and a failure to resonate with modern audiences, resulting in a substantial financial loss.

Our goal is to develop a solution that helps **prevent** similar pitfalls by equipping industry professionals with AI-driven tools to optimize decision-making. By

leveraging advanced sentiment analysis and feedback interpretation techniques, our approach ensures that films are not only creatively compelling but also strategically aligned with audience preferences, **increasing their chances of success**.

1.2 The Innovation

The film industry increasingly demands a fusion of traditional methodologies with modern **AI-driven approaches**. In an era where audience reception dictates a movie's success, collecting and interpreting viewer feedback has become essential. Historically, **Machine Learning** techniques have been employed in this field with two primary objectives:

- **Sentiment Analysis:** To assess the overall audience perception—whether positive or negative—based on movie reviews.
- **Topic Extraction:** To identify the most frequently discussed themes, allowing filmmakers and distributors to make data-driven decisions on production, marketing, and audience targeting, tailored to a specific genre.

Recent advancements have integrated **Machine Learning** and **Natural Language Processing (NLP)** to enhance audience analysis further. This combination enables a deeper understanding of the **key drivers** that influence a film's reception, not only by detecting trending topics but also by evaluating the emotional and critical aspects of audience responses. By doing so, **production studios** can anticipate potential challenges, refine creative choices, and tailor marketing efforts to maximize engagement.

1.2.1 LLMs and Review Analysis for Film Industry Enhancement

The advent of **Large Language Models (LLMs)** has revolutionized sentiment analysis, making it more precise and adaptable to the needs of the film industry. LLMs, extensively used for **review analysis** in various sectors, provide filmmakers and studios with an **unprecedented ability** to extract meaningful insights from vast amounts of audience feedback.

By leveraging this technology, **CineSense** goes beyond traditional reporting—it identifies key strengths and weaknesses in how a film is perceived, offering a **comprehensive sentiment breakdown**. This approach not only highlights **positive audience reception** but also pinpoints **critical aspects** that may require attention, such as pacing issues, character development, or marketing misalignment.

With CineSense, studios can **refine their storytelling**, adjust **promotional strategies**, and make data-driven improvements before and after a film's release. This AI-powered approach ensures that every creative and strategic decision aligns with audience expectations, ultimately increasing a film's potential for success.

Chapter 2

The Application

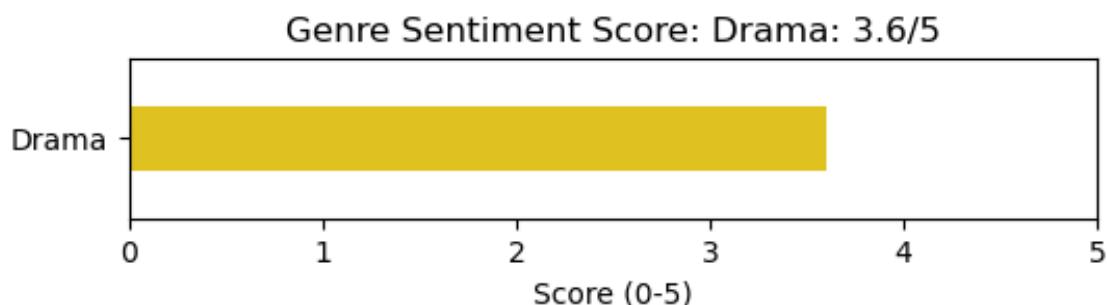
CineSense is responsible for analyzing and compiling user reviews left on many films released from the **years 2000 to today**, with the aim of offering the user a preview through numerous charts together that are complemented by a downloadable report containing many additional detailed information that allows to better understand the development context of what we are looking for and allowing have a better understanding of the audience's taste and placement than our film product may have in today's market.

2.1 Key Features and Functionalities

At its core, CineSense is an intelligent data-driven analysis tool that leverages Natural Language Processing (NLP) and sentiment analysis to transform movie reviews into valuable insights.

Once the dataset is processed, users can explore the following visualizations:

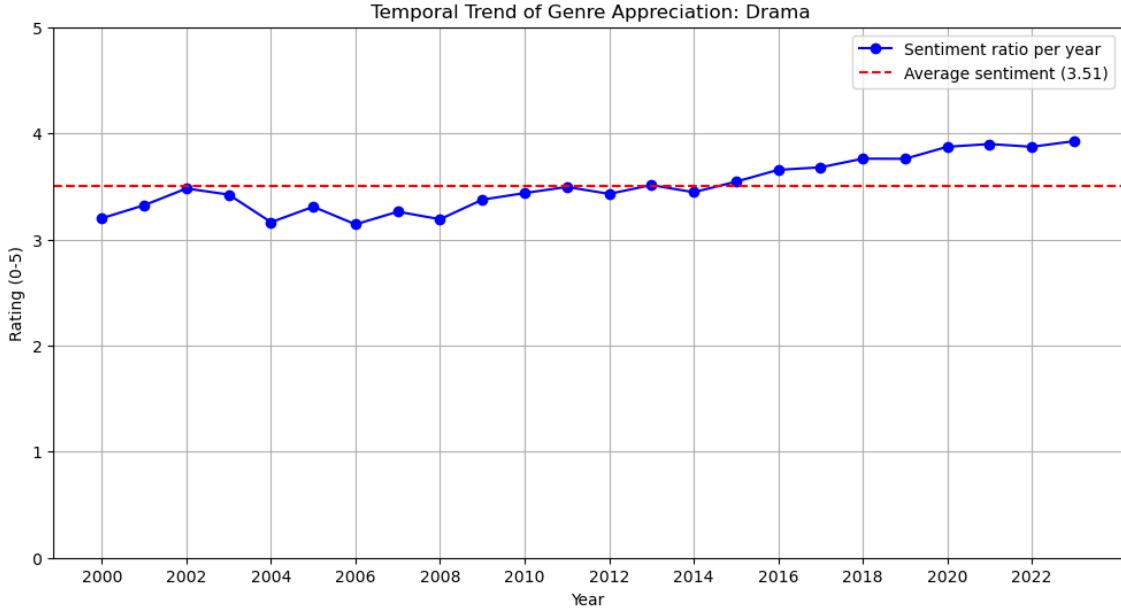
- **Audience Sentiment Overview:** A bar chart displaying the proportion of positive and negative reviews for each film or genre, providing a quick pulse check on audience reception.



- **Word Cloud Visualization:** A dynamic word cloud highlighting the most frequently used words in reviews, helping users identify key themes and recurring audience opinions while filtering out generic terms.

In particular, three wordclouds are displayed. A general word cloud that takes into analysis all the reviews of that genre which can help to have an overall view of what influences that genre. Two other wordclouds are also displayed, in which are taken into analysis only the reviews with positive sentiment or only the reviews with negative sentiment, so as to understand what is influenced positively or negatively by that genre.

- **Sentiment Trend Analysis Over Time:** A line chart illustrating how sentiment towards a specific genre has evolved year by year. This feature helps industry professionals detect shifting audience preferences and emerging trends.



Beyond the visual insights, CineSense generates a detailed PDF report, offering an in-depth exploration of sentiment trends, key audience concerns, and genre-specific reception. This report serves as a strategic guide for industry professionals seeking to align their creative and marketing efforts with audience expectations.

2.2 PDF report

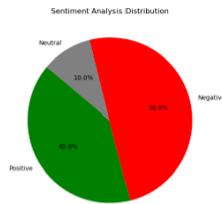
The downloadable PDF report provides in-depth insights derived from the LLM analysis, utilizing advanced techniques such as BERTopic for topic extraction.

The key sections of the report include:

- **Sentiment Analysis:** This section classifies the overall sentiment of the reviews into three categories: positive, neutral, or negative. By analyzing the emotional tone of user feedback, businesses can gauge audience satisfaction and identify common pain points or praised features.

****1. Sentiment Analysis:****

A manual sentiment analysis was performed on each review, classifying them as positive, negative, or neutral. The following table summarizes the results:

****Main Points of Praise:****

- **Thrilling/Action-packed sequences:**** Positive reviews often highlight exciting and well-choreographed action scenes.
- **Engaging plot/characters:**** Captivating storylines, well-developed characters, and strong performances receive positive attention.
- **Originality/Innovation:**** Films that break from genre conventions or offer fresh perspectives are praised.
- **Humor:**** Well-placed humor and witty dialogue contribute to positive reviews.

****Main Points of Criticism:****

- **Predictable/Clichéd plots:**** Many reviews criticize formulaic storylines, particularly in sequels and adaptations.
- **Weak characters:**** Lack of character development, unconvincing motivations, and stereotypical portrayals are frequent complaints.
- **Excessive/meaningless violence:**** While action is expected, many find the violence gratuitous, repetitive, and lacking purpose.
- **Over-reliance on CGI/Special Effects:**** Several reviews criticize the overuse of CGI, particularly when it overshadows story and character.
- **Poor dialogue/Writing:**** Weak dialogue, clunky writing, and excessive exposition are also common criticisms.
- **Strong visuals/cinematography:**** Striking visuals and impressive cinematography receive positive remarks.

****Common Patterns:****

- **Positive Reviews:**** Often focus on the visceral thrill of the action, engaging characters, and emotional resonance.
- **Negative Reviews:**** Tend to highlight plot predictability, weak character development, and excessive reliance on visual effects.

Figure 2.2: Action Movie Review Analysis report

- Keyword Extraction:** This part of the report focuses on identifying the most relevant keywords from the reviews. By extracting frequently mentioned terms, it helps highlight the primary themes discussed by users. This enables businesses to understand what aspects of a product or service are most talked about, whether positively or negatively.

****2. Keyword Extraction:****

The following tables present the most frequent words and phrases for positive and negative reviews (excluding stop words):

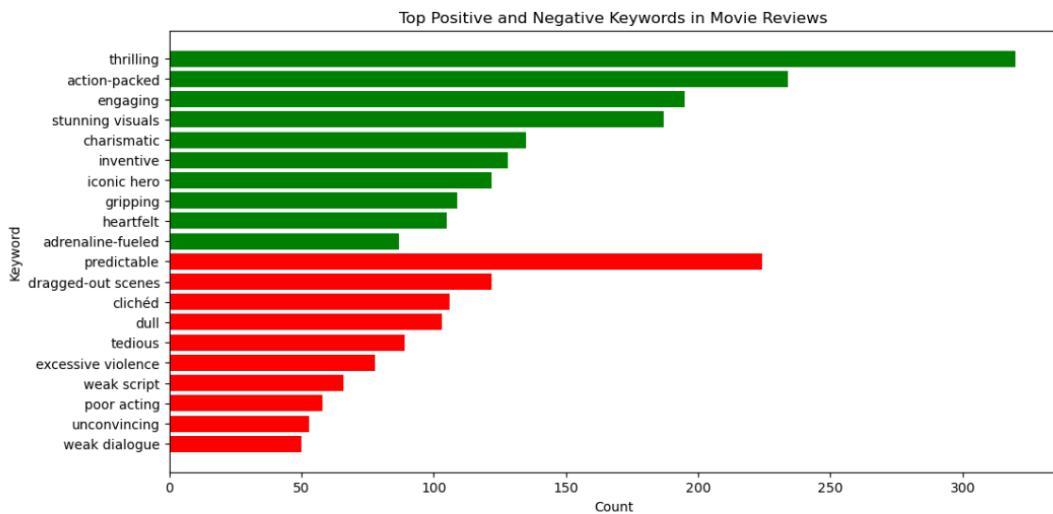


Figure 2.3: Action Movie Review Analysis report

- **Topic Modeling with BERTopic:** BERTopic is an advanced topic modeling technique that combines transformer-based embeddings with clustering algorithms to extract meaningful topics from textual data. This section groups user feedback into major topics, allowing businesses to categorize reviews efficiently. It provides an overview of key themes and trends, helping stakeholders understand what users discuss the most.

****3. Topic Modeling:****

BERTopic was used to identify the main topics in the reviews. The following table summarizes the top 5 topics:

1. ****Superhero Fatigue:**** *Description:* Discusses the oversaturation of superhero films, criticizing predictable plots, weak characters, and excessive reliance on CGI.
Keywords: superhero, Marvel, DC, CGI, predictable, formulaic.
2. ****Action Spectacle vs. Substance:**** *Description:* Explores the balance between impressive action sequences and meaningful storytelling, with some reviews praising thrilling visuals and others criticizing a lack of narrative depth.
Keywords: action, spectacle, CGI, plot, characters, story, thrilling, boring.
3. ****Reboots/Remakes/Sequels:**** *Description:* Focuses on the challenges of reviving or continuing franchises, with criticism of derivative plots and praise for innovative approaches.
Keywords: reboot, remake, sequel, original, predictable, fresh.
4. ****Cultural Representation:**** *Description:* Discusses the portrayal of different cultures and identities in action films, criticizing stereotypes and praising authenticity.
Keywords: diversity, representation, stereotype, authentic, racist, sexist.
5. ****Low-Budget Brilliance vs. Incompetence:**** *Description:* Examines the potential for low-budget action films to offer creative brilliance or suffer from technical limitations.
Keywords: low-budget, creative, inventive, amateurish, cheesy, flawed.

Figure 2.4: Action Movie Review Analysis report

- **Marketing Insights:** This section translates the findings from the sentiment analysis, keyword extraction, and topic modeling into actionable business insights. By analyzing user preferences and expectations, especially across different film genres, businesses can refine their marketing strategies, tailor content, and enhance their offerings to better align with audience demands.

****4. Marketing Insights Summary:****

- * ***Highlight unique selling propositions:** Emphasize originality and innovation in action sequences, characters, or storytelling. Differentiate from the glut of superhero films by focusing on other action subgenres (e.g., martial arts, spy thrillers).
- * ***Address criticisms of predictability:** In trailers and marketing materials, showcase plot twists, unexpected character development, or thematic depth to counter the perception of formulaic storytelling.
- * ***Balance action spectacle with substance:** While highlighting thrilling action in trailers, also showcase engaging character moments and narrative hooks to attract a broader audience.
- * ***Target specific demographics:** Use language and imagery that resonate with different audiences. For example, for action purists, emphasize practical effects and stuntwork; for those seeking more depth, highlight compelling characters and thought-provoking themes.
- * ***Leverage positive reviews:** Use quotes from positive reviews that highlight specific strengths (e.g., "thrilling action," "engaging characters," "surprising twists").
- * ***Monitor and respond to negative feedback:** Analyze common criticisms in negative reviews to identify areas for improvement in future films and marketing campaigns. Address concerns about excessive violence or reliance on CGI by showcasing a more balanced approach to filmmaking.
- * ***Optimize movie descriptions:** Use keywords identified in the analysis to create more engaging and accurate movie descriptions that target specific audiences.

Figure 2.5: Action Movie Review Analysis report

Each of these sections contributes to a comprehensive understanding of user feedback, enabling data-driven decision-making for businesses.

2.3 User Interface

The CineSense application provides a **user-friendly** interface designed to facilitate the analysis of audience reviews for different movie genres. The UI is structured to guide users through **data processing, sentiment analysis, and visualization of insights** to optimize film production and marketing strategies.

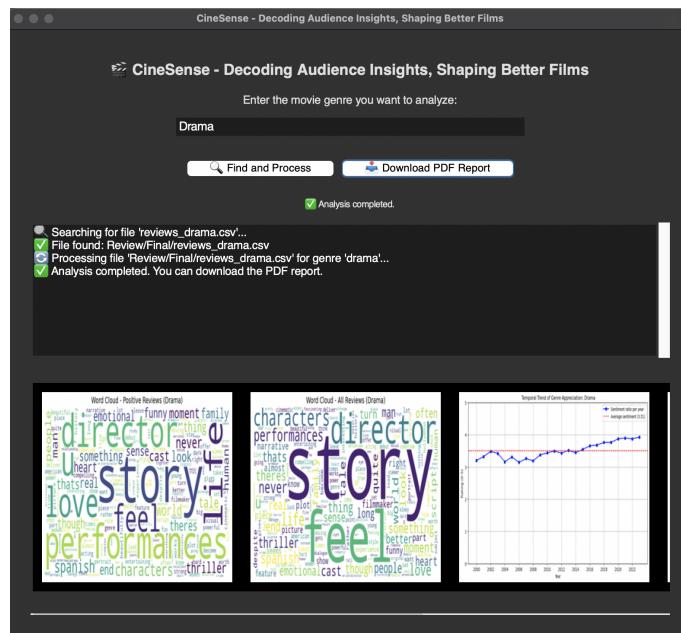


Figure 2.6: UI Home Page

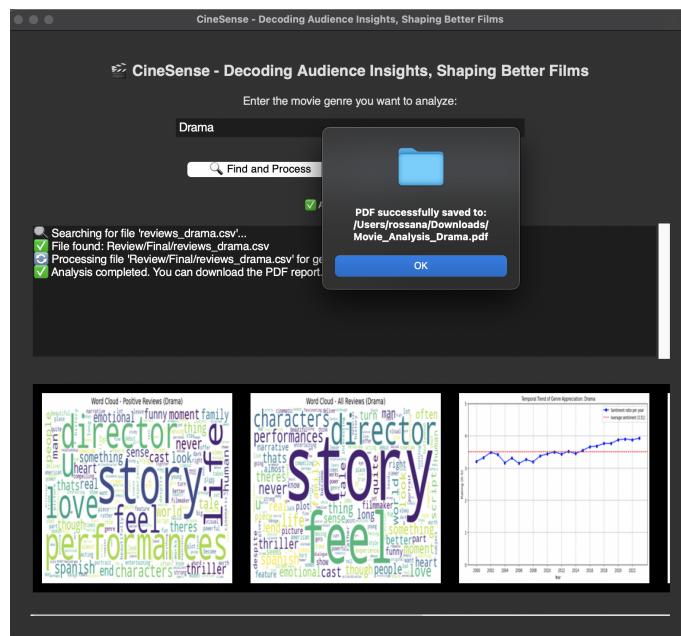


Figure 2.7: UI, report saved successfully

2.3.1 Genre Selection & Processing

Users can enter the desired movie genre (e.g., *Drama*) in the input field. Clicking "**Find and Process**" initiates the search for the corresponding dataset and starts processing the audience reviews.

2.3.2 Real-Time Processing Log

A progress log below the buttons displays real-time updates on the steps performed:

- **Searching** for the corresponding review file
- **Confirming** file discovery and processing status
- **Indicating** when the analysis is completed and the PDF report is ready for download.

2.3.3 Report Generation & Download

Once the analysis is complete, the "**Download PDF Report**" button allows users to save a detailed report containing insights into the selected genre.

2.3.4 Data Visualization & Insights

The lower section of the interface displays key visual elements that provide actionable insights:

- **Word Cloud - Positive Reviews (Left)**: Highlights frequently used words in positive audience reviews, helping to understand key strengths of films in this genre.
- **Word Cloud - All Reviews (Center)**: Shows the most commonly mentioned words across all reviews, providing an overall sentiment map.
- **Temporal Trend of Genre Appreciation (Right)**: A line graph tracking sentiment evolution over time, allowing studios to observe how audience perception of the genre has changed.

The user can visualize the complete analysis by downloading the PDF report.

2.3.5 CineSense Benefits

This intuitive UI makes it easy for filmmakers, producers, and distributors to extract meaningful insights from audience feedback.

Chapter 3

Implementation

Before developing the full-scale application, the core algorithm was designed and trained within the *Gemini Module*. To ensure meaningful data for analysis, we first performed a join operation with the film metadata dataset to associate each review with its corresponding film genre. Once the genre information was integrated, the data underwent a thorough preprocessing phase to improve quality and relevance. This step included removing empty rows and eliminating special characters such as emojis and foreign letters, ensuring only high-quality data was retained for analysis. We kept only the columns relevant to our study: `creationDate`, `reviewText`, `scoreSentiment`, `title`, `runtimeMinutes`, `genre`, and `originalLanguage`.

3.1 Dataset Preprocessing

To test the functionalities of our application, we selected three genres: **Action**, **Drama**, and **Comedy**, as these are the most popular genres with the highest number of reviews. Before sending the data to the Gemini API for further processing, two crucial preprocessing steps were implemented:

- **Filtering Phase:** A custom filtering process was executed using the `filter.py` module. This module is designed to clean up text data by removing or replacing certain elements that could interfere with analysis or machine learning models. For example, it removes emojis, hashtags, mentions, and special characters that are not needed for understanding the core content of a review. Additionally, it replaces common abbreviations and internet slang (such as "u" for "you" or "lol" for "laughing out loud") with their full forms, making the text more readable and analyzable. By processing text in this way, it becomes cleaner and more standardized, helping businesses or developers extract more meaningful insights, especially when working with large amounts of user-generated content like social media posts or customer reviews. This type of preprocessing is essential for ensuring that text data is ready for further analysis, such as sentiment analysis or feedback categorization.

- **Tokenization Phase:** After filtering, the comments are transformed into arrays of tokens—essential words. These arrays were also used to create the word cloud graphics.

In the CineSense project, once the initial text mining preprocessing is completed, the Gemini API is utilized to extract key topics from user reviews and conduct sentiment analysis, classifying feedback as positive, neutral, or negative.

At a subsequent stage, Generative AI analyzes the extracted insights to identify recurring patterns and critical issues in audience feedback. Based on these findings, it suggests strategic recommendations to address concerns and enhance user satisfaction.

Finally, the system compiles the AI-generated insights into a downloadable PDF report, providing businesses with actionable intelligence to refine marketing strategies, optimize content offerings, and better align with audience expectations.

3.2 Gemini Long Context API

CineSense leverages cutting-edge artificial intelligence technologies to analyze vast amounts of movie reviews, providing production companies with actionable insights. A key component of this process is the Gemini Long Context API [2], specifically **Gemini-1.5-Flash**, which allows us to handle extensive datasets efficiently. This chapter explores the advantages of using the Gemini Long Context API and how it enhances CineSense's ability to extract meaningful patterns from audience feedback.

3.2.1 Gemini-1.5-Flash

Gemini-1.5-Flash stands out as a highly capable **Large Language Model (LLM)** designed to handle long-context data efficiently, making it the optimal choice for CineSense. Traditional LLMs often face constraints in processing large amounts of text, leading to fragmented analyses and loss of contextual coherence. In contrast, Gemini-1.5-Flash overcomes these limitations by enabling the processing of datasets containing millions of tokens in a single query. This capability allows CineSense to analyze entire review corpora without excessive truncation or preprocessing, ensuring a holistic understanding of audience sentiment.

Another key advantage of Gemini-1.5-Flash lies in its **speed** and **cost-effectiveness**. Optimized for high-speed processing, it delivers real-time insights while maintaining computational efficiency, allowing CineSense to handle high-volume datasets with minimal hardware demands. Additionally, its ability to retain **contextual information across multiple queries** enhances the model's adaptability, enabling incremental learning and refinement. This ensures that trends in audience sentiment are captured more accurately, leading to precise topic extraction and better predictions of a movie's success.

The integration of Gemini-1.5-Flash into CineSense provides a significant edge in analyzing reviews, not only by extracting relevant topics but also by interpreting

complex relationships within audience feedback. Unlike conventional models that require constant reprocessing of data, Gemini-1.5-Flash preserves narrative coherence across discussions, enabling CineSense to offer more insightful and actionable recommendations to film studios. The model’s superior ability to understand, process, and analyze large-scale datasets makes it a critical component in revolutionizing how the film industry leverages audience feedback to refine and improve cinematic productions.

An important factor in selecting Gemini-1.5-Flash for CineSense was its superior ability to **handle long-context processing efficiently**. The following image, taken from a comparative study [1], illustrates the cumulative average string similarity score as a function of context length **over 2000 instances of the MRCR task**. This benchmark demonstrates how Gemini 1.5 Pro and Gemini 1.5 Flash outperform competitors such as GPT-4 Turbo and Claude 3 models when processing extensive text sequences. The results indicate that after 32K tokens, Gemini 1.5 Pro surpasses Claude 3 Opus and GPT-4 Turbo, while Gemini 1.5 Flash maintains superior performance over Claude 3 Sonnet from 5K tokens onwards. Both Gemini 1.5 Pro and Gemini 1.5 Flash exhibit smaller performance drops as context length increases, making them ideal for large-scale text analysis, such as processing vast movie review datasets in CineSense.

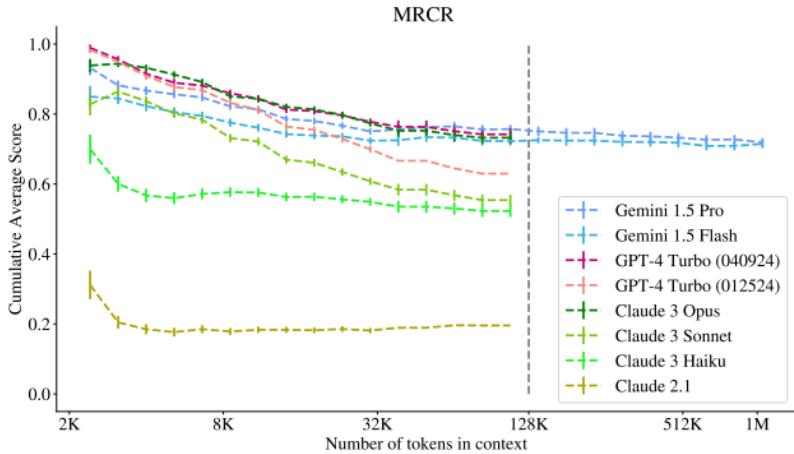


Figure 3.1: Performance comparison of various LLMs on the MRCR task

The Gemini Long Context API is a game-changer for CineSense, enabling it to handle vast datasets with precision and efficiency. By leveraging Gemini-1.5-Flash, CineSense ensures that every piece of audience feedback is analyzed in-depth and in context, providing production companies with unparalleled insights. The ability to process long-context data without fragmentation sets CineSense apart from traditional review analysis tools, making it an essential AI-driven solution for the future of film production.

Chapter 4

Results

4.1 Validation of Sentiment Analysis and Topic Extraction

To ensure the accuracy and reliability of our sentiment analysis and topic extraction approach, we validated the algorithm using multiple datasets containing movie reviews from different genres: **Drama**, **Action**, and **Comedy**. These datasets were specifically selected to capture a broad spectrum of audience feedback, both in terms of review volume and linguistic characteristics.

Some datasets, such as those for **Drama films**, contained *detailed, in-depth reviews*, where viewers elaborated on elements like character development and cinematography. Others, like those for **Action movies**, featured a *higher number of short and direct comments*, often emphasizing aspects such as pacing and special effects. The **Comedy genre**, on the other hand, exhibited a mix of structured critiques and casual reactions, posing additional challenges in topic identification and sentiment classification.

In addition, the audience demographics and expectations differ significantly across genres. **Drama films** typically attract audiences that engage deeply with storytelling and emotional depth, while **Action movies** often receive feedback centered around spectacle and excitement. **Comedy films**, being highly subjective, tend to generate a diverse range of sentiments based on individual humor preferences.

4.1.1 Sentiment Analysis

For the validation of the sentiment analysis performed by Gemini we used the **sentiment label** present within our dataset associated with each review. This label can take positive or negative value, and what we did was pass the reviews to Gemini, without passing this label, so that it managed to rank these reviews with a negative or positive sentiment. The results obtained were excellent and highly accurate, in particular by taking the three reference genres already specified above, **93%** accuracy for the drama genre, **97%** for the comedy genre, and finally **89%** for the action

genre.

4.1.2 Topic Extraction

For **topic extraction**, automated validation was not feasible. Instead, we cross-referenced the extracted topics with *common themes discussed in movie reviews and professional critic analyses*. The results showed strong consistency with industry insights:

- **Drama films:** Negative reviews often focused on *pacing issues* and *overly complex narratives*, while positive reviews highlighted *strong character arcs* and *emotional depth*.
- **Action movies:** Viewers frequently praised *intense action sequences* but criticized *predictable plots* and *lack of character development*.
- **Comedy films:** While humor was the central theme, negative reviews frequently mentioned *forced jokes* and *inconsistent tone*, whereas positive reviews appreciated *smart writing and comedic timing*.

These findings confirm that our approach effectively captures **the core themes and sentiments that define audience reception** across different film genres. The insights generated by our system can serve as a **valuable resource for filmmakers, studios, and content strategists**, helping them better understand audience expectations and refine their creative decisions.

CineSense provides valuable insights into audience sentiment and preferences, making it a powerful tool for filmmakers, producers, and marketers in the film industry. By leveraging advanced AI-driven sentiment analysis and topic modeling, CineSense deciphers audience reviews to highlight key trends, strengths, and potential weaknesses of a film. These insights can be used in various ways: during pre-production, filmmakers can align their creative choices with audience expectations; in post-production, they can refine elements such as pacing or character development based on real feedback; and in marketing campaigns, studios can tailor promotional strategies to resonate with specific target demographics. Additionally, the detailed PDF reports generated by CineSense serve as a strategic guide, providing industry professionals with data-driven recommendations to optimize a film's success in an increasingly competitive market.

Chapter 5

Competitor Analysis

The film industry is increasingly incorporating AI-driven insights to refine production and marketing strategies. Several companies have entered this space, offering solutions that focus on predictive analytics, audience demand measurement, or script optimization. However, these tools often lack the ability to extract deep, actionable insights directly from audience reviews. CineSense sets itself apart by combining Natural Language Processing (NLP), Topic Modeling, and Strategic Business Intelligence, enabling production companies to leverage audience feedback in a more meaningful way.

5.1 Existing Competitors

Competitors like *Cinelytic* focus on financial projections and casting decisions, while *Parrot Analytics* tracks global audience demand but lacks in-depth review analysis. *StoryFit* specializes in pre-production script optimization, missing post-release audience insights. Traditional review aggregators such as *IMDb*, *Rotten Tomatoes*, and *Metacritic* provide sentiment scores but do not analyze the deeper themes within audience feedback. These solutions, while valuable, do not offer the level of contextual understanding and strategic guidance that CineSense provides.

5.2 The CineSense Advantage

CineSense distinguishes itself by going beyond simple sentiment analysis. Utilizing **Gemini-1.5-Flash** and **BERTopic modeling**, it can process massive datasets, maintaining context across extensive reviews to extract meaningful trends. Unlike competitors that focus on historical data, CineSense continuously tracks evolving audience preferences, helping studios make informed, data-driven decisions.

Its approach to predictive audience engagement analysis allows studios to understand not just whether a film is well-received but why certain aspects resonate with viewers. Additionally, CineSense offers specific recommendations for optimizing future productions based on detailed, genre-specific insights, something other

platforms fail to provide.

CineSense redefines how the film industry utilizes AI by integrating advanced NLP, topic extraction, and predictive analytics to deliver strategic business intelligence. Unlike existing solutions that either focus on financial forecasting or broad sentiment trends, CineSense provides a comprehensive framework for understanding audience engagement at a granular level. This unique capability positions it as a game-changer in **AI-driven filmmaking analysis**.

Chapter 6

Conclusion

The **CineSense** project represents a significant advancement in the analysis of audience feedback within the film industry, leveraging state-of-the-art **Large Language Models** (LLMs) to generate actionable insights.

Through a structured pipeline, CineSense efficiently processes vast amounts of user reviews, extracting sentiment trends, key topics, and audience concerns. The core steps include:

- **Data Collection and Preprocessing** – Gathering and cleaning user reviews to ensure high-quality input.
- **Topic and Sentiment Analysis** – Utilizing the Gemini API to classify sentiment and identify discussion topics.
- **Pattern Recognition and Strategic Insights** – Employing Generative AI to uncover underlying issues and suggest improvements.
- **Automated Report Generation** – Delivering structured, downloadable PDF reports containing insights that businesses can act upon.

Innovation Through LLMs

A key innovation of CineSense lies in the integration of LLMs like the **Gemini API**, which go beyond traditional rule-based or statistical methods. Unlike conventional analytics, LLMs dynamically interpret user sentiment, detect nuanced patterns in reviews, and generate strategic recommendations based on deep contextual understanding. This enhances the decision-making process for businesses, enabling them to refine marketing strategies, improve content offerings, and better align with audience expectations.

By harnessing the power of AI-driven automation and deep learning, CineSense provides a cutting-edge solution for film industry professionals, bridging the gap between audience perception and business strategy.

Bibliography

- [1] GeminiTeam. Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context. Available at: <https://arxiv.org/abs/2403.05530>, December 2024.
- [2] Google. Gemini api: Long context. Available at: <https://ai.google.dev/gemini-api/docs/long-context>.
- [3] IlPost. Il flop estivo di disney, the lone ranger. Available at: <https://www.ilpost.it/2013/08/07/the-lone-ranger-johnny-deep-disney/>, 2013.
- [4] IlPost. Il caso papmusic - l'assurdo film di animazione italiano. Available at: <https://www.ilpost.it/2024/10/03/pap-music-animation-for-fasion-film-animazione/>, 2024.