

FILMIN

Monitoring Brand Perception on Trustpilot

Ana Sánchez

Applied AI Bootcamp
AllWomen - April 2025



What's the Project About?

The brand

Filmin is a Spanish streaming platform specializing in independent, classic and European cinema.

Objective

Analyze and monitor user reviews of Filmin on Trustpilot to understand brand perception.

Why Trustpilot?

Leading platform for customer reviews.
Strong influence on brand reputation and consumer decisions.

Business Importance

Streaming market is highly competitive.
Customer trust and satisfaction are crucial for retention and growth.
Real-time feedback enables quick response to issues.

Key Questions

What are customers saying about Filmin?
How can we automate the monitoring of feedback?

Techniques Used

Data Collection

Web scraping with
Python from
Trustpilot website

Data Preprocessing

Cleaning text (removing URLs,
special characters, stopwords).
Tokenization, lemmatization,
standardizing ratings and locations.

Data Handling

CSV files managed in
Google Colab and
AWS - Jupyter Lab

Sentiment Analysis

Used DistilBERT for Spanish,
English and Catalan.
Handles context, sarcasm
and informal language.

Topic Modeling

Used BERTopic to extract
main themes and topics
from reviews.

Dashboard

Gradio prototype for
real-time visualization.

Process Steps & Decision Making

Data Collection

1

- Scraped latest 1,500 reviews from Trustpilot (out of ~6,000).
- Managed scraping limits and missing data.

Data Preprocessing

2

- Converted ratings to integers, handled missing values and outliers.
- Standardized location codes and cleaned text columns.
- Tokenized and lemmatized text, removed brand names and contractions.

Sentiment Analysis

3

- Addressed informal language.
- Evaluated performance: accuracy, precision or recall.
- Visualized results with bar charts, score bar plots, and line plots over time.

Topic Modeling

4

- Reviewed coherence and diversity of topics.
- Manually checked interpretability of results.
- Visualized topics with word clouds, 2D topic maps or topic distribution charts.

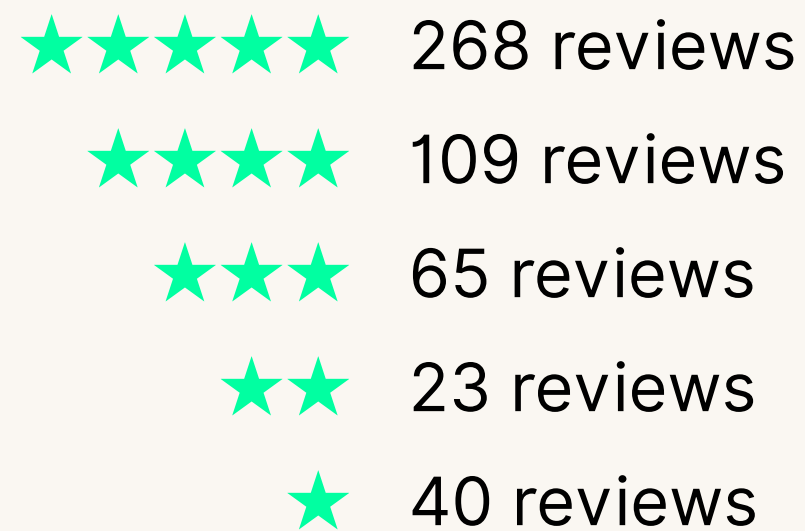
Evaluation & Adjustments

5

- Adjusted preprocessing and model parameters for better results.
- Addressed class imbalance and unclear topics.

Key Results & Insights

Sentiment Distribution



Average rating

4.46 / 5

Main Topics Identified

- Classic and unique film selection
- User experience and interface
- Exclusive content and documentaries
- Technical issues (app crashes, streaming problems)
- Subtitles and language options

Example Insights

- Most users praise the unique catalog and European films.
- Negative reviews focus on technical issues and extra charges.
- Subtitles and language options are frequently mentioned.

Conclusion & Next Steps



Challenges

- Scraping limits and missing data from Trustpilot.
- Handling Spanish language nuances and informal expressions.
- Technical issues integrating Gradio in Colab.

Project Impact

- Enables real-time, automated monitoring of Filmin's brand perception.
- Helps the company respond quickly to user feedback and improve services.

Future Improvements

- Expand to more data sources and languages.
- Test with more advanced NLP models.
- Enhance dashboard features for deeper insights.

The prototype

Content is no longer available

Thank you!

