Movie Review Sentiment Analyzer

An NLP-based Sentiment Classification Project

Presented by Aryan Aman

Introduction: Setting the Scene

The Movie Buff's Dilemma

Millions of movie reviews, endless opinions. How do we make sense of it all quickly and consistently?
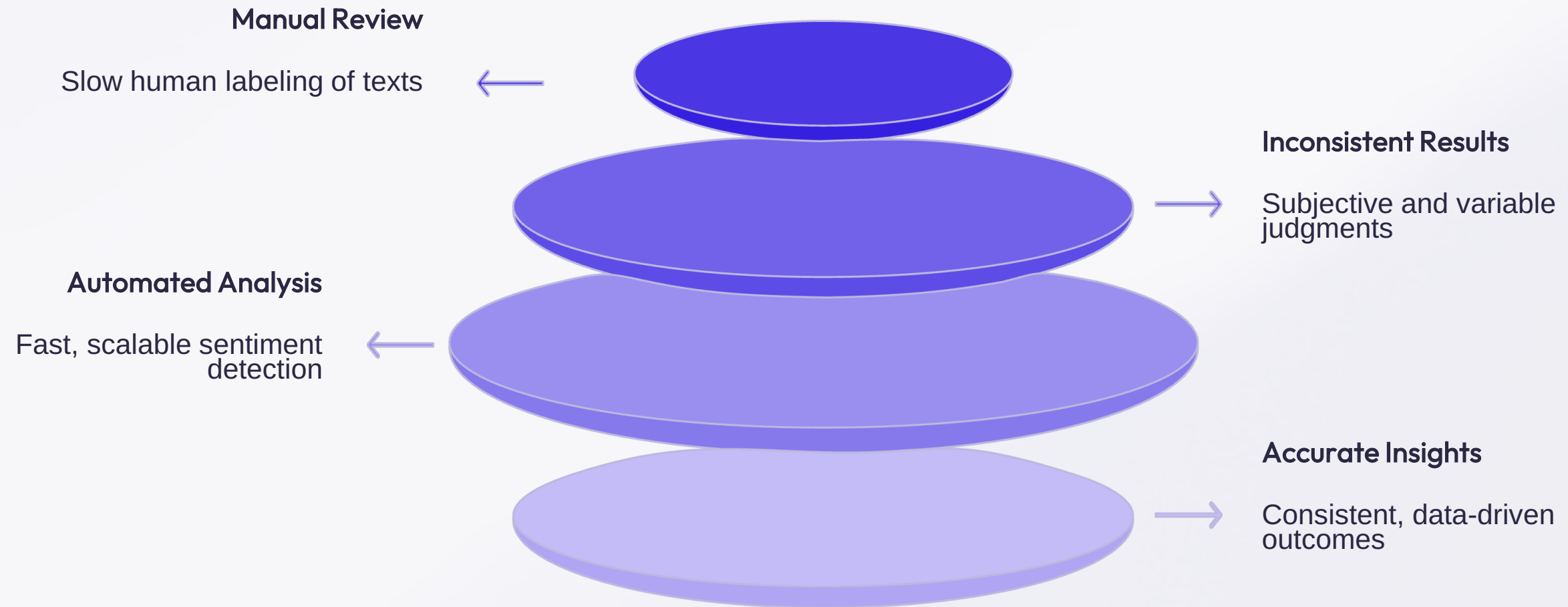
Beyond Manual Analysis

Manually reading and classifying reviews is slow, subjective, and prone to inconsistency at scale.

This project aims to automate the classification of movie review sentiment, transforming raw text into actionable insights.

The Challenge: Automating Sentiment

Manual Review

Slow human labeling of texts ←

Inconsistent Results

Subjective and variable judgments →

Automated Analysis

Fast, scalable sentiment detection ←

Accurate Insights

Consistent, data-driven outcomes →

Our core challenge is to replace tedious, inconsistent manual review classification with a robust, automated system using Natural Language Processing (NLP).

Our Objective: Precision & Speed

Classify Sentiment

Develop a model capable of accurately classifying movie reviews as either Positive or Negative.

Leverage AI

Utilize NLP and Machine Learning techniques to ensure rapid, scalable, and highly accurate sentiment analysis.

The goal is to provide instant insights into audience perception, eliminating the guesswork.

The Dataset: IMDB Movie Reviews

Positive	"A truly **masterful** and **engaging** cinematic experience!"
Negative	"The plot was **confusing** and the acting utterly **dreadful**."
Positive	"Visually **stunning** with an **unforgettable** soundtrack."

Our model is trained on the extensive IMDB Movie Review dataset, meticulously labeled as Positive or Negative. This rich source of real-world opinions forms the backbone of our analysis.

Methodology: From Text to Prediction

Preprocessing

Cleaning raw text: removing noise, stop words, and standardizing case for consistency.

TF-IDF Vectorization

Converting text into numerical features, representing word importance within reviews.

Logistic Regression

Applying a robust classification algorithm to predict sentiment based on the vectorized data.
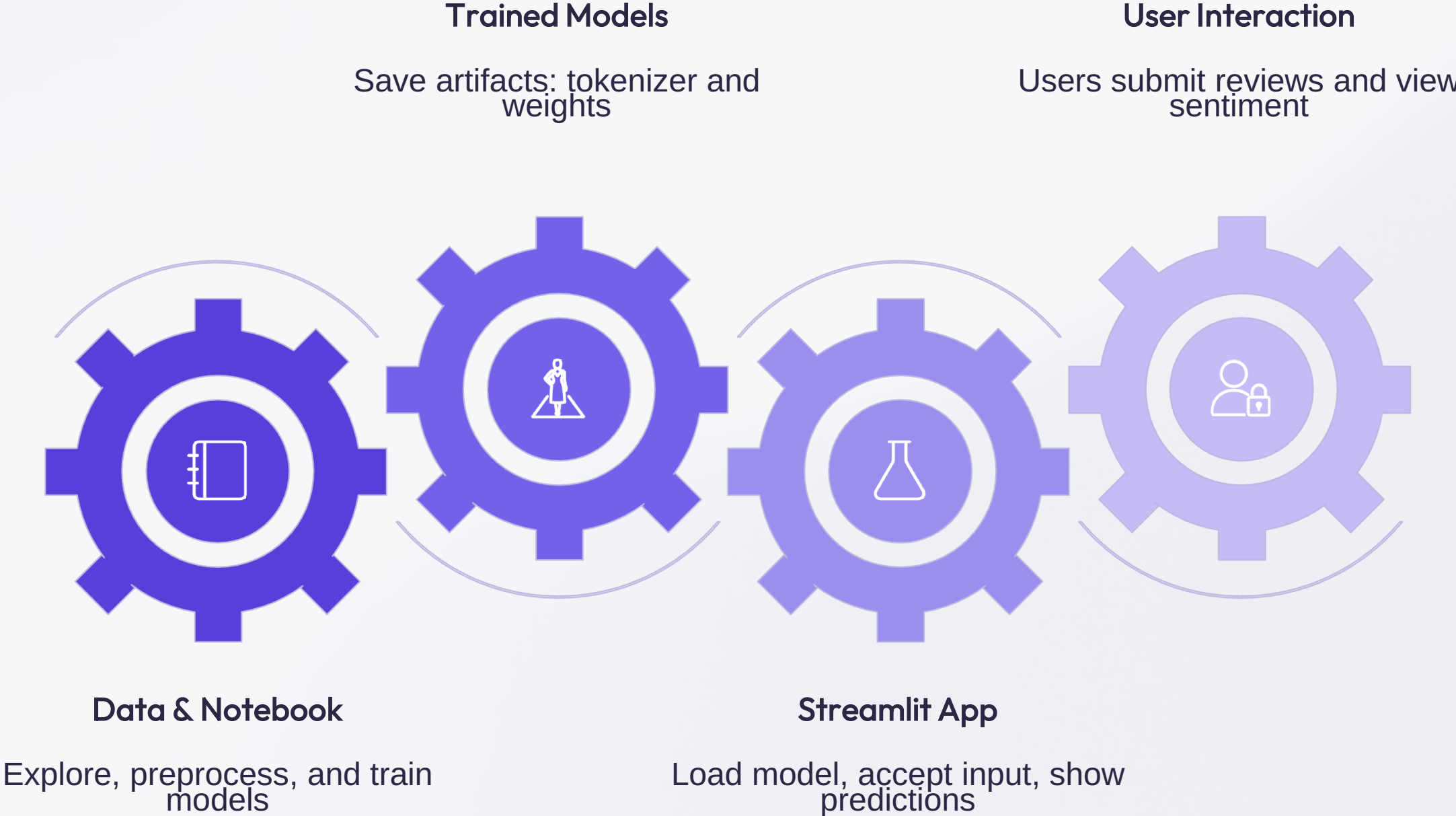
Evaluation

Assessing model performance using key metrics: Accuracy, Confusion Matrix, and F1-score.

This streamlined workflow ensures efficient and accurate sentiment classification.

Project Architecture

Trained Models

Save artifacts: tokenizer and weights

User Interaction

Users submit reviews and view sentiment

Data & Notebook

Explore, preprocess, and train models

Streamlit App

Load model, accept input, show predictions

Results: Performance & Prediction

~89%

Accuracy Achieved

Our model demonstrates a strong capability in correctly classifying review sentiment.

Example Input:
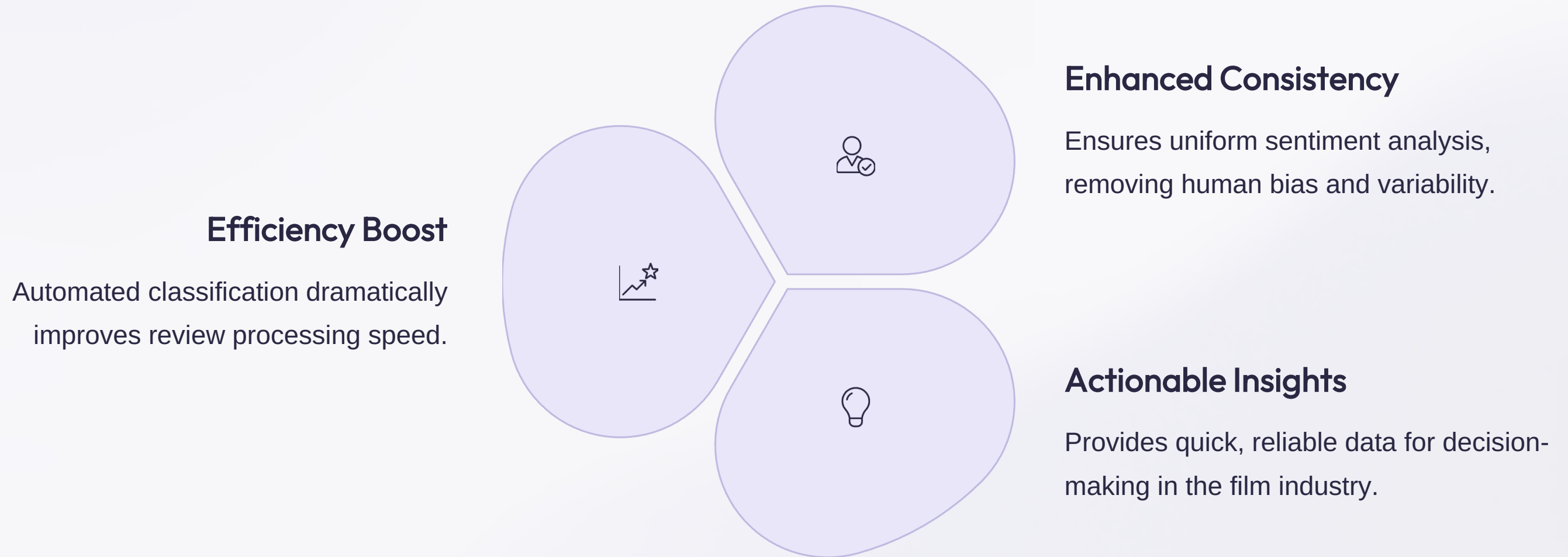
"The story was dull and disappointing."

Predicted Sentiment:

Negative

The model consistently predicts sentiment with high accuracy, validating our approach.

Conclusion: Impact & Innovation

Enhanced Consistency

Ensures uniform sentiment analysis, removing human bias and variability.

Efficiency Boost

Automated classification dramatically improves review processing speed.

Actionable Insights

Provides quick, reliable data for decision-making in the film industry.

This project demonstrates the power of NLP and Machine Learning in transforming raw data into valuable intelligence.

Thank you. Aryan Aman