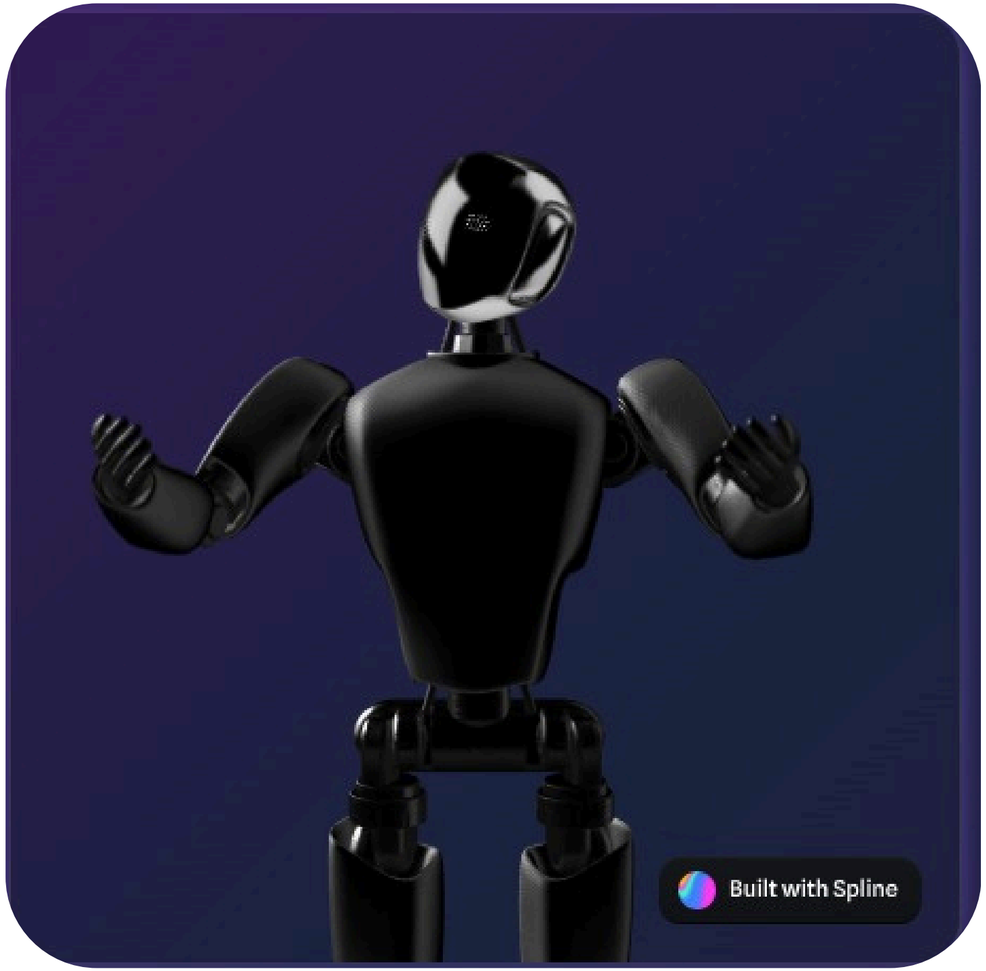
INFLUENCERIQ

Team Members:
Harshada Bhandare (23102B2003)
Sakshi Mane (23102B2004)
Krishna Mahorkar (22102B0002)
Vedant Mantri (22102B0066)



INTRODUCTION

Influencer marketing is now one of the most powerful tools for brand promotion.

But metrics like likes and followers don’t show authentic engagement.

InfluencerIQ solves this by using NLP, and sentiment analysis to evaluate influencer credibility and audience quality.

PROBLEM STATEMENT

With the rapid rise of digital creators on platforms like YouTube and Instagram, brands struggle to identify genuine influencers due to fake engagement and inauthentic audiences. Traditional tools rely on basic metrics such as likes, views, and followers, which overlook engagement quality and audience sentiment.

To address this, a data-driven, AI-powered system is needed that:

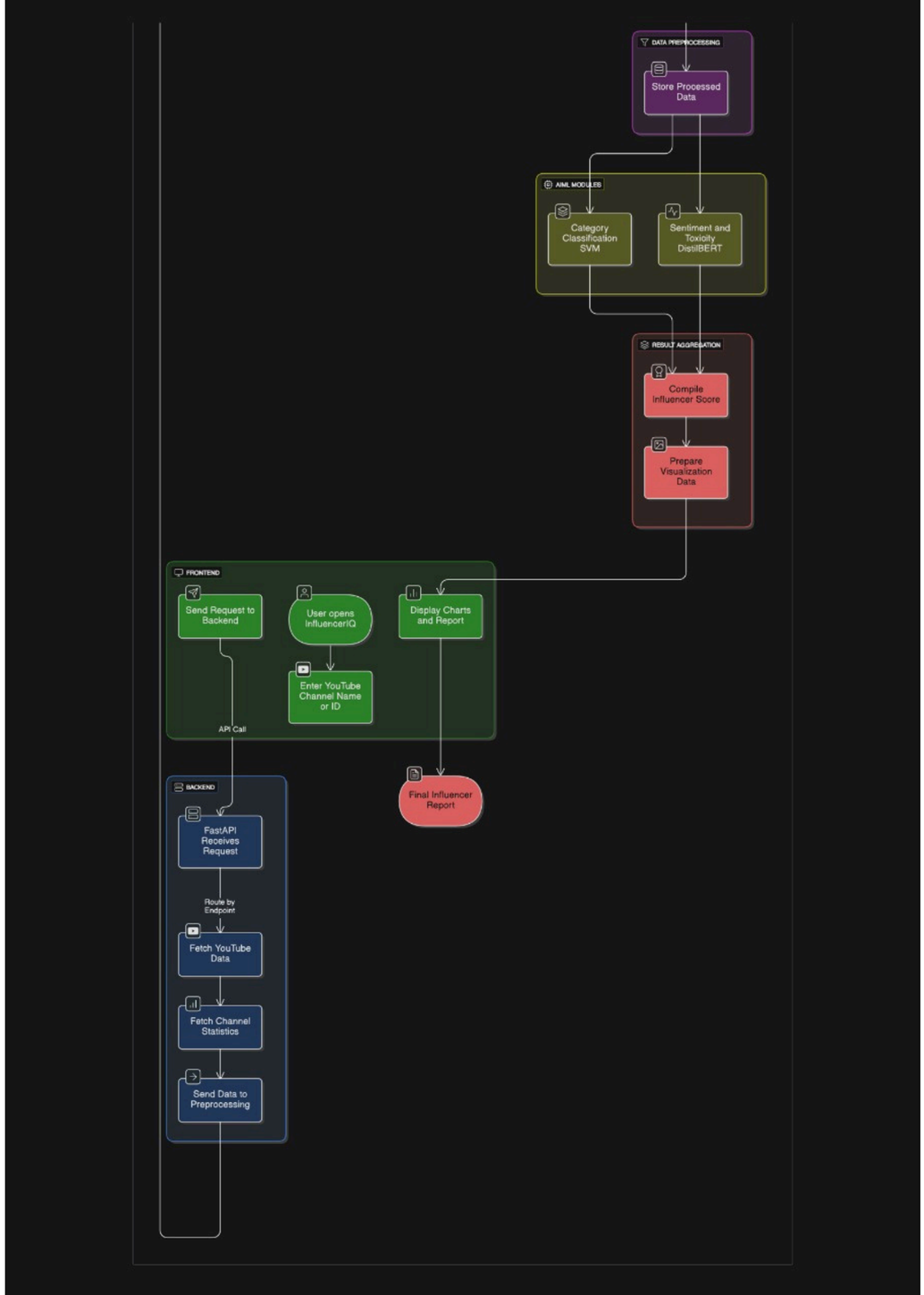
- Analyzes audience sentiment to gauge positivity and credibility.
- Detects toxicity and spam-driven interactions.
- Generates an objective Influencer IQ Score combining engagement, sentiment, and authenticity.
- Provides an interactive dashboard for real-time insights and decision-making.

OBJECTIVE

- Perform sentiment and toxicity analysis on YouTube and Instagram comments to measure audience positivity, negativity, and overall community health.
- Design an automated scoring system (Influencer IQ Score) that integrates engagement rate, growth velocity, sentiment tone, and toxicity penalties into a single, objective rating.
- Build an interactive web dashboard to visualize influencer analytics — including engagement trends, sentiment distribution, and category-wise performance — for easy interpretation by brands and agencies.
- Enable data-driven decision-making by providing insights that help brands identify authentic, high-impact influencers and avoid fake or toxic profiles.

SYSTEM ARCHITECTURE

- **Data Acquisition:** Collected influencer data and comments from YouTube API and Kaggle datasets for analysis.
- **Preprocessing:** Used spaCy NLP to clean text by removing noise, tokenizing, and lemmatizing comments for consistency.
- **Model Training:** Trained SVM for large-scale sentiment detection and DistilBERT for deeper contextual understanding.
- **Scoring Module:** Combined engagement, sentiment, and toxicity metrics to generate an overall Influencer IQ Score.
- **Visualization:** Built an interactive React.js dashboard with Recharts to display insights and influencer performance.



INFLUENCERIQ SCORING

Engagement Score

Measures audience interaction through likes, comments, and views.

Formula:

$(0.7 \times ((\text{Likes} + \text{Comments}) / (\text{Views} + 1)) + 0.3 \times (\log_{1+}(\text{Likes} + \text{Comments}) / \log_{1+}(\text{Subscribers} + 10))) \times 100$

📈 Bonus: +10% for ≥7 posts/week, +5% for ≥3 posts/week

Base Rating

Evaluates overall influence based on engagement, likes, views, and subscribers.

Formula:

$(0.35 \times S + 0.25 \times E + 0.25 \times L + 0.15 \times V) \times 5$

Where

$S = \log(\text{Subscribers} + 1) / \log(10{,}000{,}000)$ $L = \text{Likes} / (\text{Likes} + \text{Dislikes} + 1)$

$E = (\text{Likes} + \text{Dislikes}) / (\text{Views} + 1)$ $V = \min(1, \text{Views} / (\text{Subscribers} \times 10 + 1))$

Toxicity Score

Penalizes for negative/toxic comments.

Formula:

Base Penalty = $(0.7 \times \text{Severity} + 0.3 \times \text{Balance}) \times 5$

Penalty Multiplier:

50% negative → ×2.0 30% negative → ×1.5

20% negative → ×1.2 10% negative → ×1.0

Final Sponsorship Rating

Combines all metrics and adjusts for sponsorships and toxicity.

Formula:

Raw Final Rating = Base Rating + Sponsorship Penalty − Toxicity Penalty

Final Rating = clipped between 1–5

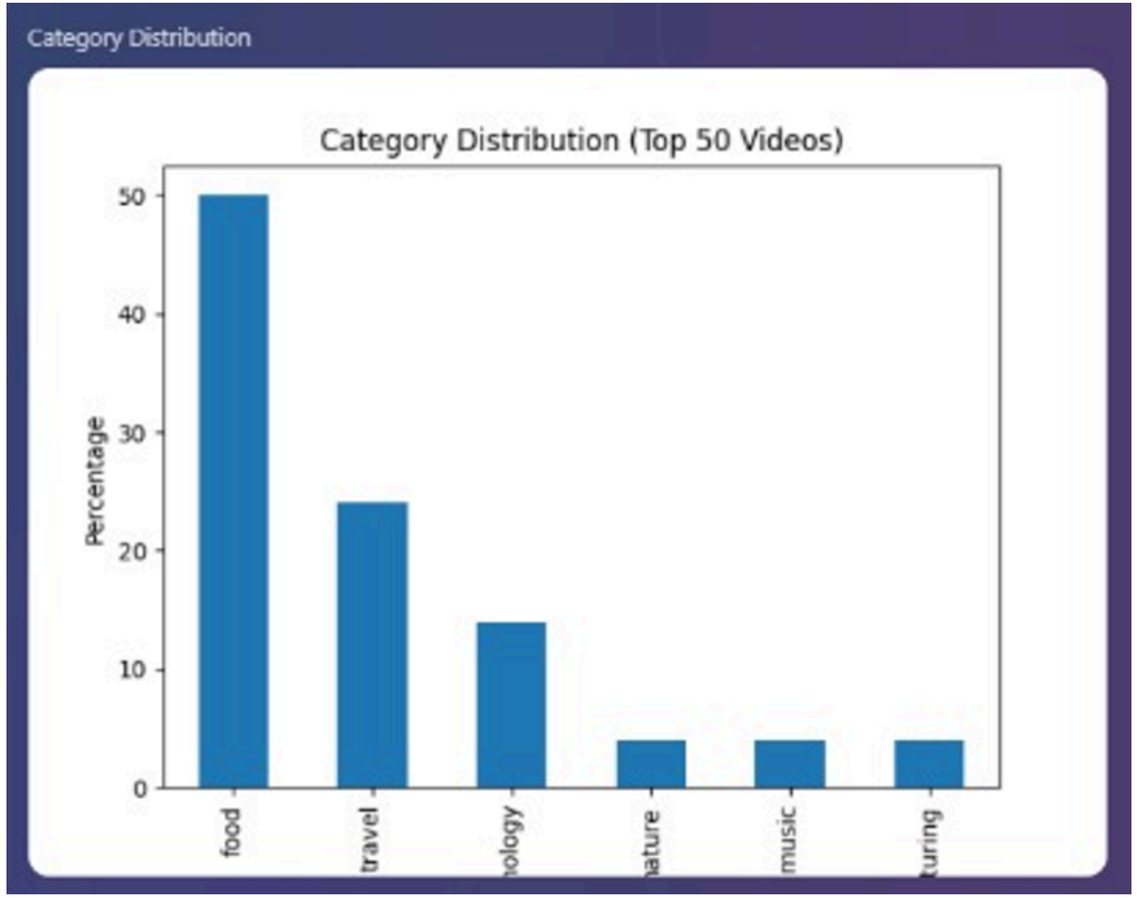
🔹 Sponsorship keywords (e.g., #ad, use code) add +1 penalty

DASHBOARD & RESULTS

A Logistic Regression was used to group videos into major content categories based on themes, keywords, and engagement metrics. This helps identify what type of content the influencer primarily produces.

Key Insights:

- Categorization helps in understanding content diversity and audience interest areas.
- Enables brands to align sponsorships with the influencer’s strongest performing categories.
- Supports data-driven decisions for content strategy and niche targeting.
- Reveals trends in audience engagement across categories, aiding future planning.



DASHBOARD & RESULTS

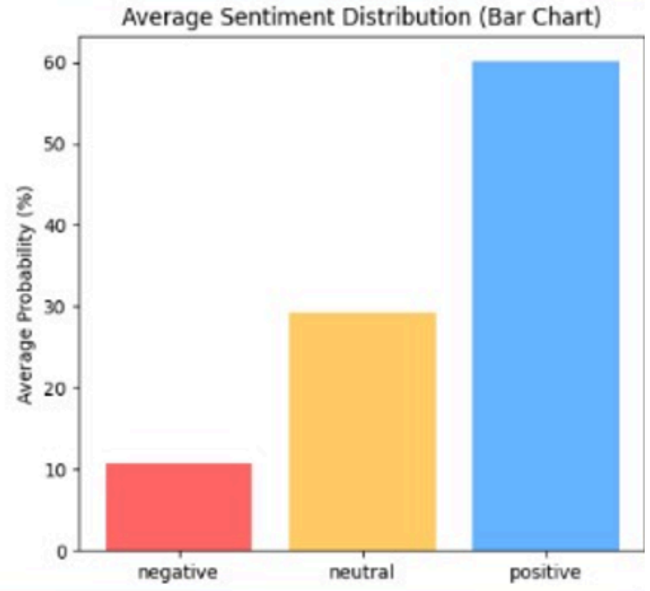
Sentiment analysis was performed using SpaCy for text processing and an SVM model for classifying comments into positive, neutral, and negative categories.

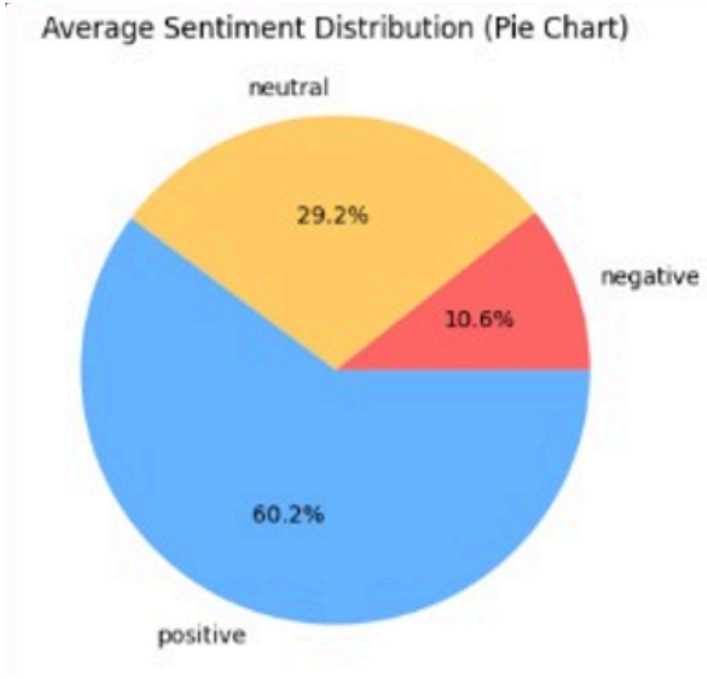
Approach:

- SpaCy used for preprocessing (tokenization, lemmatization).
- SVM trained for sentiment classification.
- Sarcasm handled using custom rule-based filtering.

Insights:

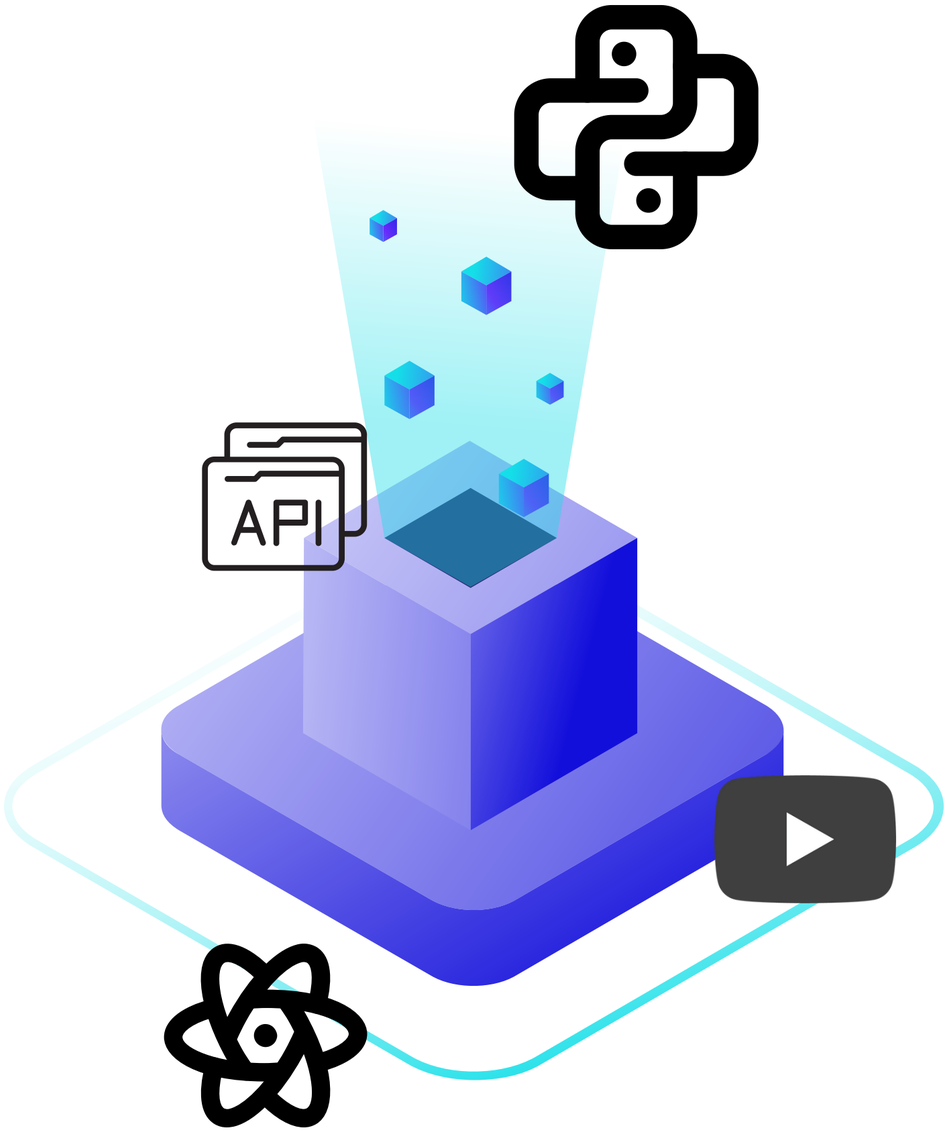
- Shows overall audience perception and engagement tone.
- Useful for tracking brand sentiment and content reception.





TOOLS & TECHNOLOGIES

Python – A versatile, high-level programming language used for backend logic, data processing, and model integration.
React – A powerful frontend library for building dynamic and interactive user interfaces.
Tailwind CSS – A utility-first CSS framework that allows for fast and responsive UI design.
Postman – A tool used for testing and debugging APIs efficiently.
YouTube API – Enables fetching and displaying video content dynamically from YouTube.
FastAPI – A modern, high-performance Python framework for building fast, scalable, and reliable APIs.
Frameworks: Apache Spark, Streamlit



CONCLUSION

- InfluencerIQ is a data-driven, AI-powered platform for evaluating social media influencers.
- It combines NLP, sentiment analysis, and interactive dashboards to deliver actionable

FUTURE ENHANCEMENT

- API Integration: Add Instagram Graph and Twitter APIs for multi-platform analytics.
- Emotion Detection: Expand beyond sentiment to capture emotions like joy, anger, and surprise.
- Advanced Models: Implement deep learning architectures (BERT / RoBERTa) for improved sentiment accuracy and contextual understanding.

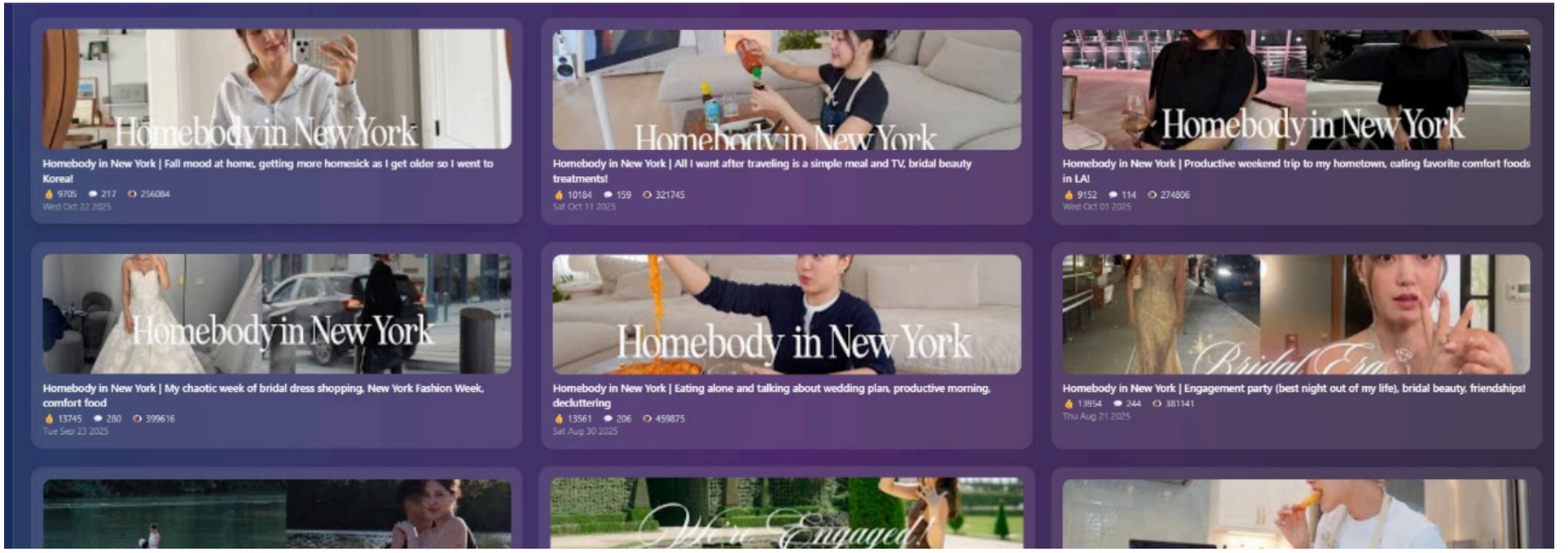
INFLUENCERIO

UNLOCK THE SECRET TO VIRAL MARKETING

Find, analyze, and partner with the perfect influencers for your brand. Get real-time insights, competitor intel, and AI-powered recommendations—so you never waste a dollar on the wrong collab.

YOUTUBE

- 500+ Brands Trust Us
- AI-Powered Matching



YOUTUBE INFLUENCER SEARCH

michel choi

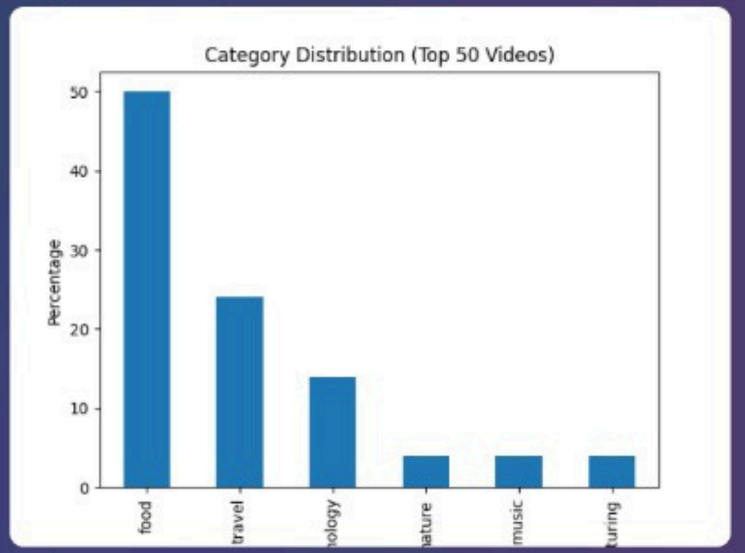
Search

Michelle Choi
2,310,000 Subscribers

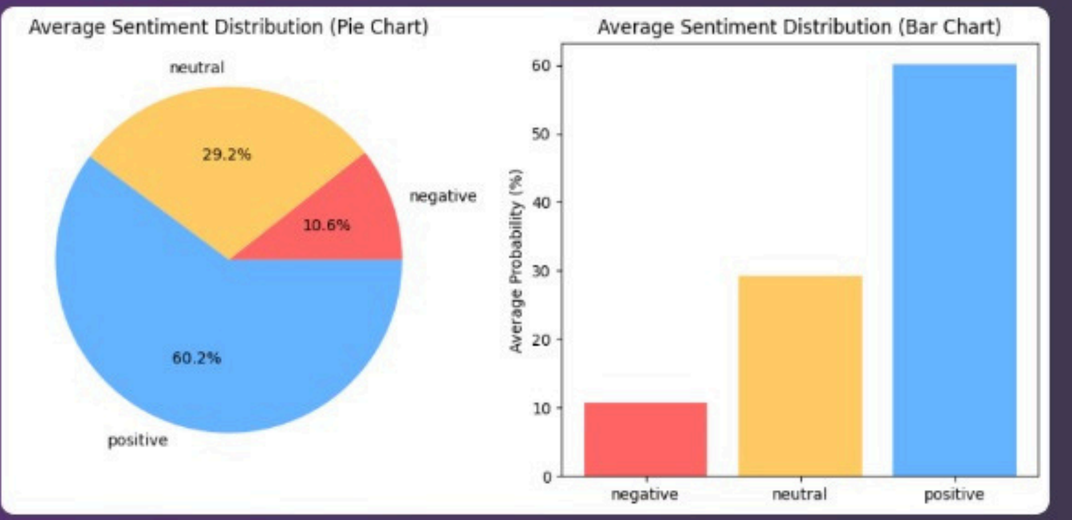
Hi! My name is Michelle. I share my experiences on how I'm navigating my 20s through conversations about mental health, wellness, beauty, and life in New York. I also have a brand called Little Puffy. You can check it out here: www.littlepuffy.com Email: michellechoi@select.co

Engagement Rate	Base Rating	Sponsorship Penalty	Toxicity Penalty	Final Rating	Posts per Week	Posting Regularity	Avg Video Length (sec)	Growth Velocity (%)
34.01%	**3.17**	**1**	**0**	**4**	**0.61**	**7.81**	**610**	**88.42**

Category Distribution



Sentiment Distribution



THANK YOU