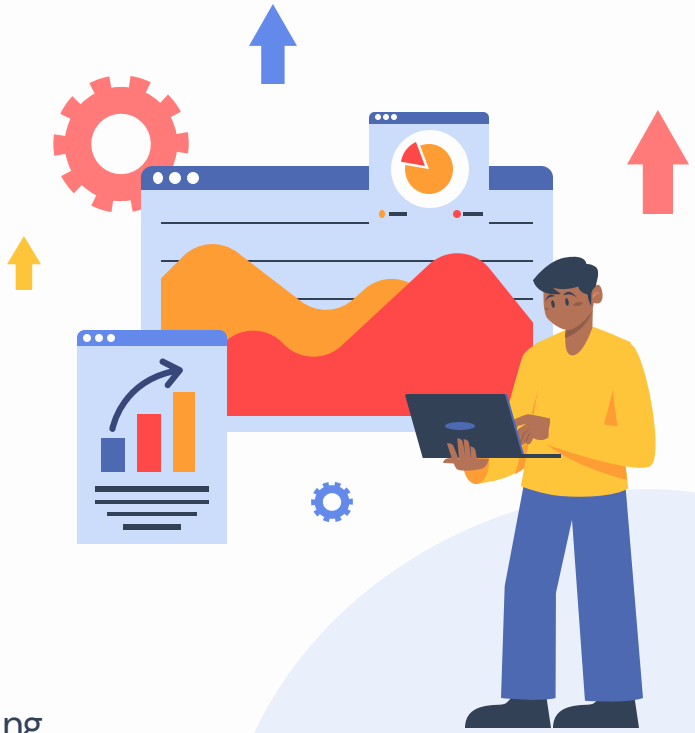
Meme Analytics Bot

AI -powered Meme Scraping, Analysis & Reporting

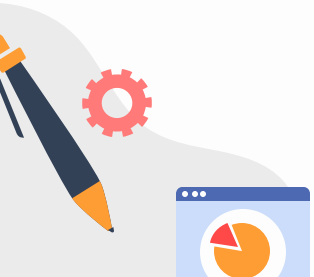
By: Shaune Ang



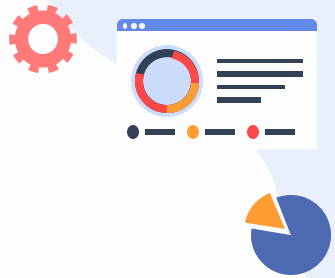
Guiding Objective

Give creators an insight for them to have the best chance of getting popular

Table of contents



Key Decisions



Fetching Frequency

Once every hour to obtain an even spread of data

NoSQL Database

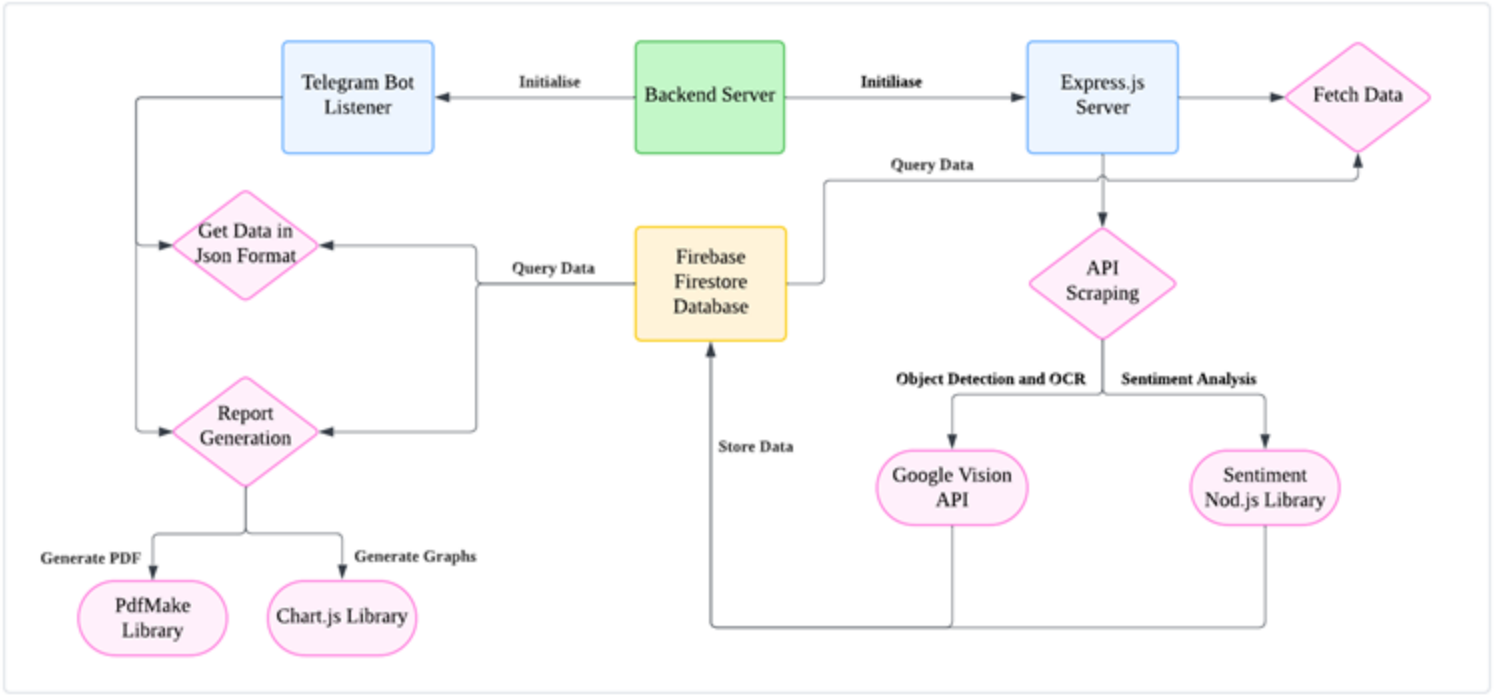
Flexible semi-structured data

Quick implementation

Popularity Metric

Up vote count used to determine popularity

System Overview



Techstack

Backend	Purpose
Node.js + Express.js	Backend API & server logic
Firebase Firestore	NoSQL database for meme storage
Reddit API	Fetching trending memes

Analysis	Purpose
Google Vision AI	Image recognition & OCR
Sentiment	Sentiment analysis
Compromise NLP	Text analysis & keyword extraction

Report	Purpose
Chart.js + ChartJS Node Canvas	Data visualization
pdfmake	PDF report generation
Telegram Bot API	User interaction

Core Features

Divide the class into small groups of 3-5 students. Assign each group a **real-life application** for which they will have to use **math concepts** to solve

1 Periodically Scrape Memes

Up to date database

2 Graph-Based Insights & Trends

- Meme popularity over time
- Vote ratios & comment trends
- Text & image correlation

Topics

3 Telegram Report Generation

Easy user interaction

4 AI & NLP Meme Analysis

- Image Text Extraction (OCR)
- Sentiment & Keyword Analysis
- Object Detection

Report Content

Section	Purpose
Top Voted Memes	Overall statistics on the top 20 up voted memes. With information such as *Author, Upvotes, Downvotes, Upvote Ratio, Comments*.
Ranked Meme Creators	To highlight whether there are creators that are consistently on top, so that creators can follow these authors and try to emulate them.
Top Keywords Extracted	Shows any trending keywords that appear in top memes.
Meme Timestamp Distribution	Identifies peak meme activity hours and the best times to post for maximum engagement.
Comments vs. Upvotes	Determine if highly upvoted memes also attract more discussion, indicating higher engagement
Votes vs. Upvote Ratio	Determine if highly voted memes are well-received or controversial
Post Format Distribution	Track which meme formats are most popular and how users engage with different media types
Sentiment Distribution	Analyse whether trending memes are uplifting, controversial, or neutral in tone.

Challenges and Pitfalls

Categorisation of Memes

Difficult to categorise memes based on individual pieces of information (e.g. image, text, OCR)

Solution:
Use CLIP to train a model to detect meme categories based on multiple inputs

OCR for Image Text Inaccurate

When using Tesseract OCR for recognising the text in images, there were inaccuracies in extracted text.

Solution:
Tested with Google Vision API and used it instead

Keyword Extraction Relevance

Keywords extracted from OCR and Title were not very relevant

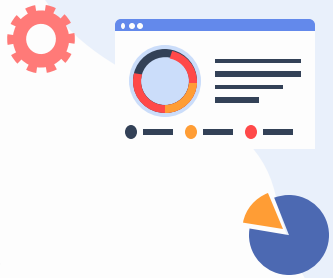
Solution:
Used Google Vision API's Object Detection to detect objects in the image instead

Additional Remedy:
Reprocess keywords before storing

Future Developments

These are some developments that could have been made given more time.

1. Categorical Analysis of Top Performing Memes
2. Web dashboard with interactive meme analytics, live charts and data



Thank you!

AI -powered Meme Scraping, Analysis & Reporting

By: Shaune Ang