Me: *uses machine learning*
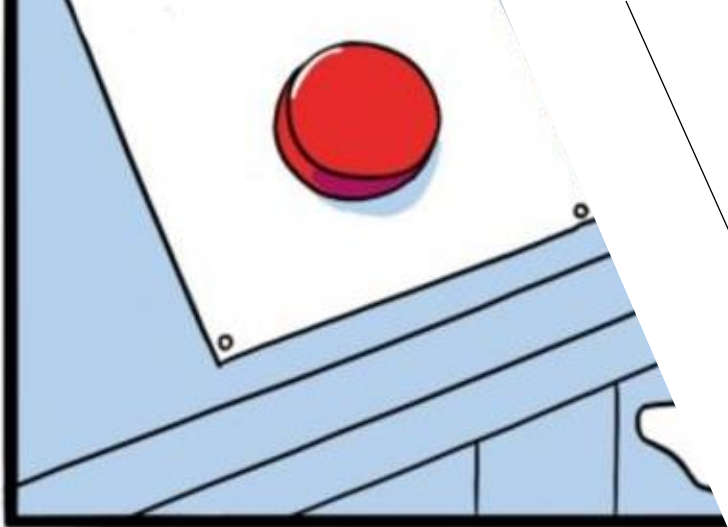
Machine: *learns*

Me:

MEMEMATCH: YOUR AI MEME RECOMMENDATION ASSISTANT

Tri An Le

I. PROBLEMS

- Multimodal Meme and Meme Template Retrieval and Recommendation.

- Text-to-Image Relevance Evaluation.

- Information Retrieval Metrics and Analysis.
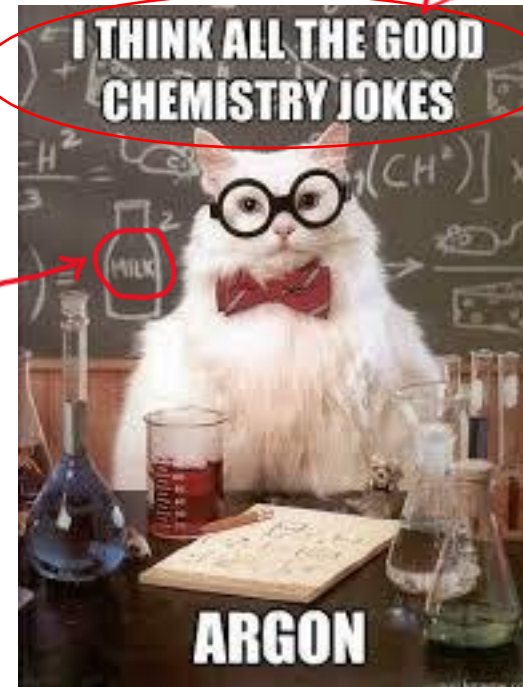
II. DATA

BACKGROUND

Dataset:

- ~301K image-with-text memes from Reddit and ImageFlip.

- ~2100 meme templates from ImageFlip.

- Metadata includes: upvotes, titles, urls, created time.

- 2 parts: local context and global context.

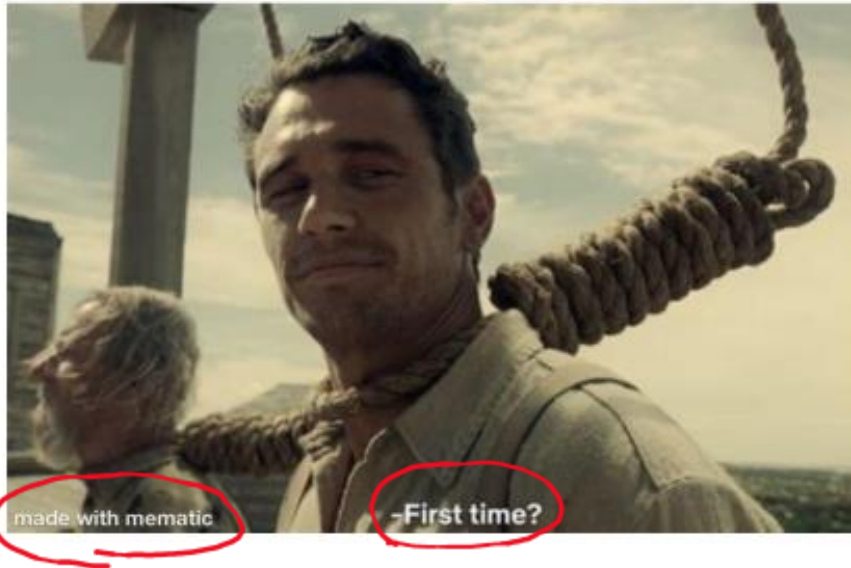

Part of local context

Part of global context

III. DATA PREPROCESSING, METHODOLOGY & RESULTS

1. Text Extraction

"3DS owners: OH NOI We'll have to pirate games nowl Wii owners: made with mematic First time?"

"Corporate needs you to find the differences between this picture and this picture: made with mematic They're the same picture. Fat free milk water"

OCR result

-> Cannot be used because of improper format (watermark, global context, delimiter issue, etc.)

1. Text Extraction

Local Context

"3DS owners: OH NOI We'll have to pirate games nowl Wii owners:"

"Fat free milk water"

OCR result after data cleaning
-> Good to use!

EasyOCR

	id	score	title	body	url	image_url	post_link	created_utc
…								
755700	meme_submissions_1287583	24	Darkness		https://i.redd.it/zb9g8eu3kwi81.jpg	https://preview.redd.it/zb9g8eu3kwi81.jpg?auto...	/r/meme/comments/swqxzy/darkness/	2022-02-20 02:52:39

	id	score	title	body	url	image_url	post_link	created_utc
…								
738792	meme_submissions_1249423	27	No difference whatsoever		https://i.redd.it/wytrjxdyn2581.jpg	https://preview.redd.it/wytrjxdyn2581.jpg? auto...	/r/meme/comments/reknze/no_difference_whatsoever/	2021-12-12 08:27:02

→ Final local context: "Fat free milk water No difference whatsoever"

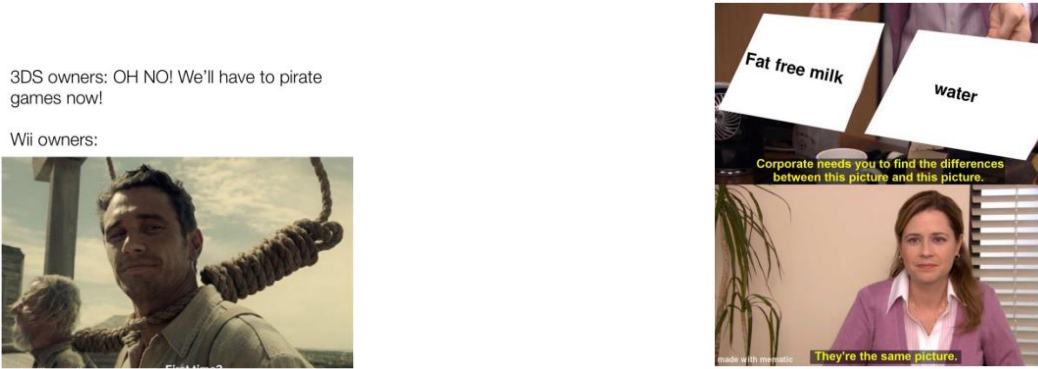# 3. Image Captioning

"a man with a rope around his neck"
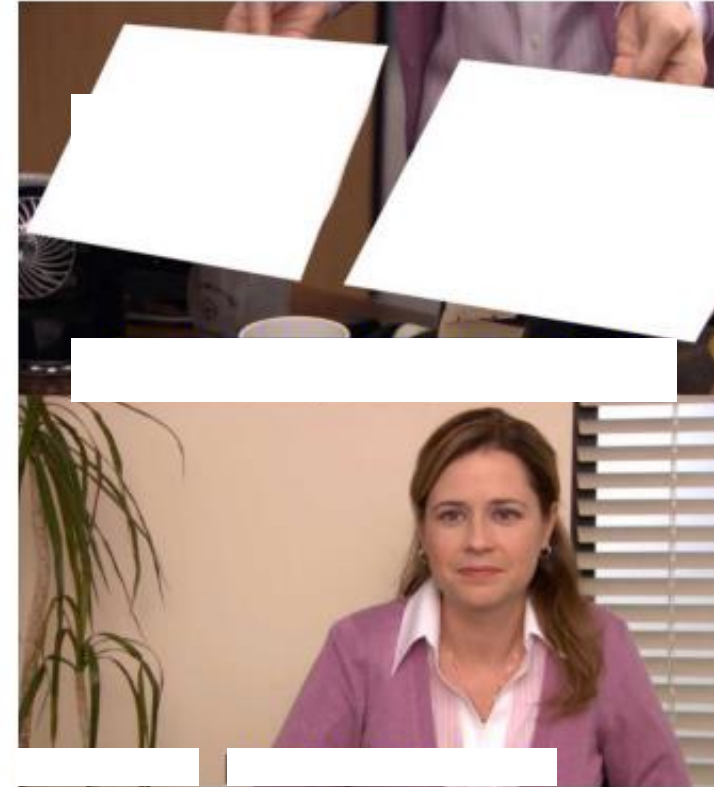
"a woman in a purple jacket is holding a piece of paper"

- PaddleOCR for white masking
- "Salesforce/blip-image-captioning-base" for Image Captioning

4. Sentiment Analysis

2 RoBERTa models ⟹ 14 sentiments and their corresponding scores

- **Anger** 😡🔥
- **Anticipation** ⏳🤔
- **Disgust** 🤢🤧
- **Fear** 😨😱
- **Joy** 😃🎉
- **Love** ❤️😍
- **Optimism** 😊🌞
- **Pessimism** 😞🌧️
- **Sadness** 😢💔
- **Surprise** 😲😮
- **Trust** 🤝🫱🫲
- **Negative** 🚫👎
- **Neutral** 😐🤷
- **Positive** ✅👍

0.479241	0.593545
0.19945	0.054121,
0.647908	0.673541
0.066868	0.010768
0.032594	0.033441
0.002387	0.003806
0.016223	0.021999
0.075464	0.022189
0.217513	0.673541
0.060239	0.022189
0.00559	0.00443
0.768447	0.079837
0.215292	0.720003
0.016261	0.20016

"cardiffnlp/twitter-roberta-base-emotion-multilabel-latest" & "cardiffnlp/twitter-roberta-base-sentiment-latest"

3DS owners: OH NO! We'll have to pirate games now!

Wii owners:

made with mematic   –First time?

Sarcasm or Irony|
Confusion or Disbelief|
Reaction or Reply Meme|
Parody or Spoof|
Popular Meme Format|
Self-Deprecation|
Absurd or Random Humor|
Comparison or Contrast|
Emotional Frustration|
Product or Brand Promotion|
Media or Brand Critique|
Hyperbole or Exaggeration|
Sadness or Disappointment|
Punchline or Joke

Fat free milk   water

Corporate needs you to find the differences between this picture and this picture.

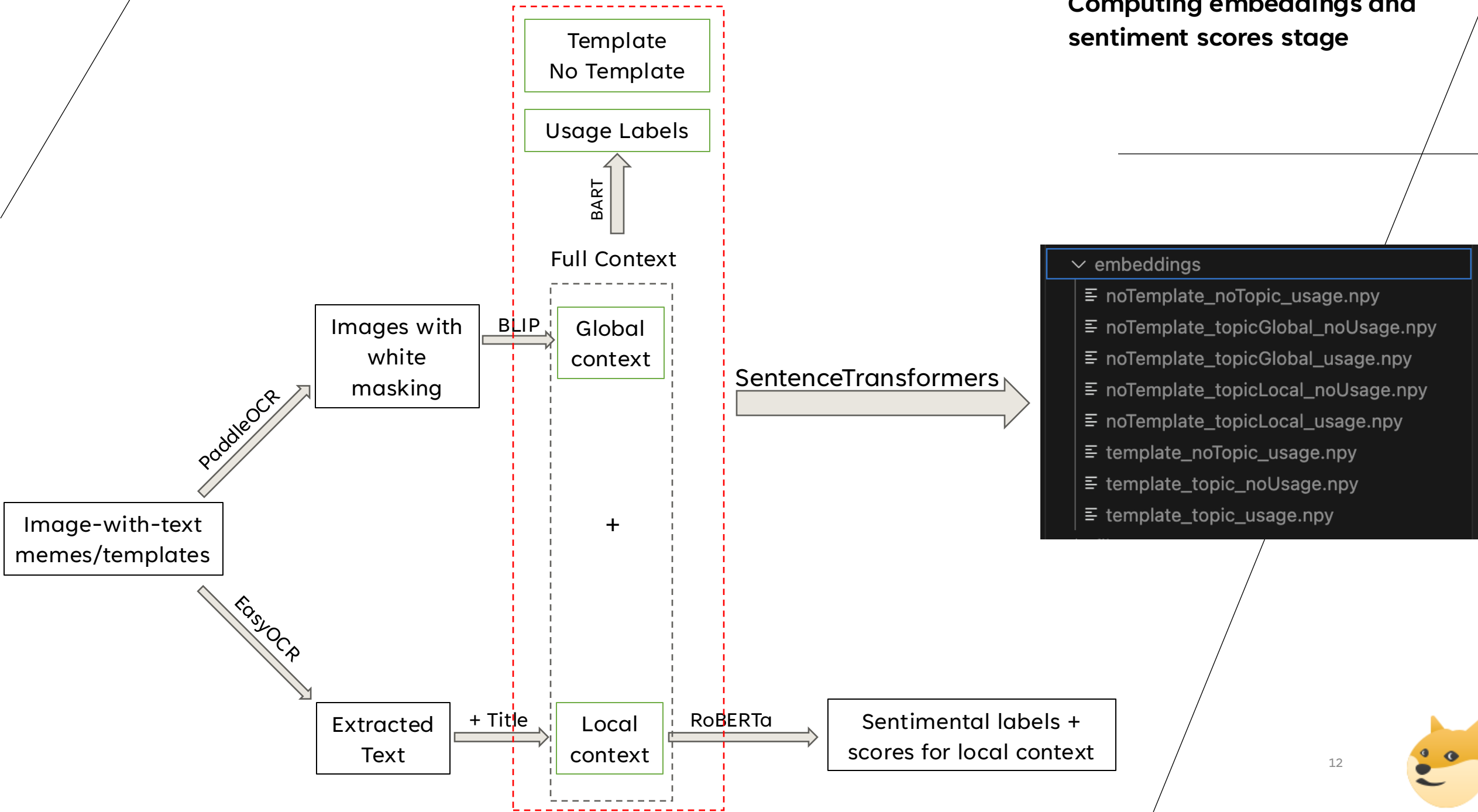made with mematic   They're the same picture.

Parody or Spoof|
Absurd or Random Humor

-> Methodology: Zero-shot classification with Hugging Face Transformers (BART-Large-MNLI) and using custom usage labels.

Template
No Template

Usage Labels

BART

Full Context

Images with white masking

BLIP

Global context

SentenceTransformers

embeddings
- noTemplate_noTopic_usage.npy
- noTemplate_topicGlobal_noUsage.npy
- noTemplate_topicGlobal_usage.npy
- noTemplate_topicLocal_noUsage.npy
- noTemplate_topicLocal_usage.npy
- template_noTopic_usage.npy
- template_topic_noUsage.npy
- template_topic_usage.npy

PaddleOCR

Image-with-text memes/templates

EasyOCR

Extracted Text

+ Title

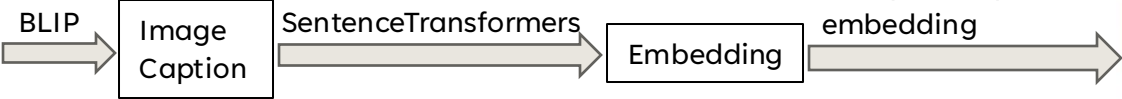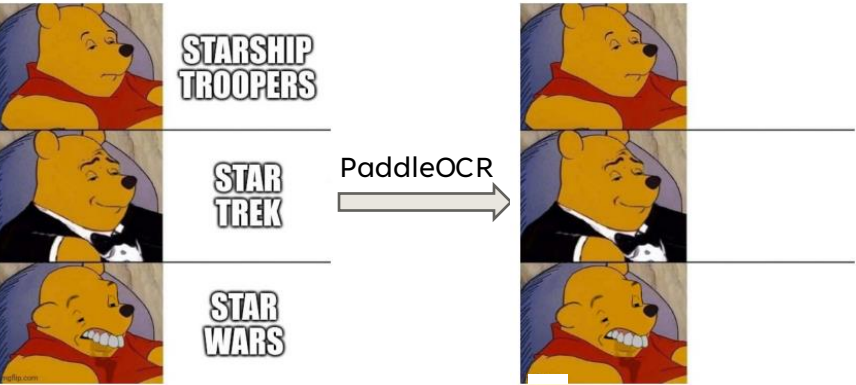Local context

RoBERTa

Sentimental labels + scores for local context

12

Feature 1: Find similar memes based on Global Context (Image-based)

Output: Top n similar memes based on image

Uploaded Meme:

PaddleOCR → BLIP → Image Caption → SentenceTransformers → Embedding → Cosine Similarity with precomputed embedding
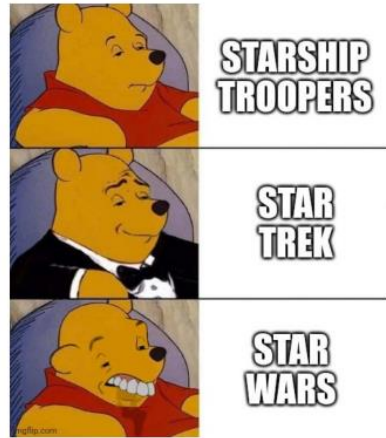
```
∨ embeddings
   ☰ noTemplate_noTopic_usage.npy
   ☰ noTemplate_topicGlobal_noUsage.npy
   ☰ noTemplate_topicGlobal_usage.npy
   ☰ noTemplate_topicLocal_noUsage.npy
   ☰ noTemplate_topicLocal_usage.npy
   ☰ template_noTopic_usage.npy
   ☰ template_topic_noUsage.npy
   ☰ template_topic_usage.npy
```

*Embeddings are generated by SentenceTransformer("all-mpnet-base-v2")

Feature 2: Find similar memes based on Local Context (Text-based)

Output: Top n similar memes based on text

Uploaded Meme:

EasyOCR → Extracted Text → SentenceTransformers → Embedding → Cosine Similarity with precomputed embedding →
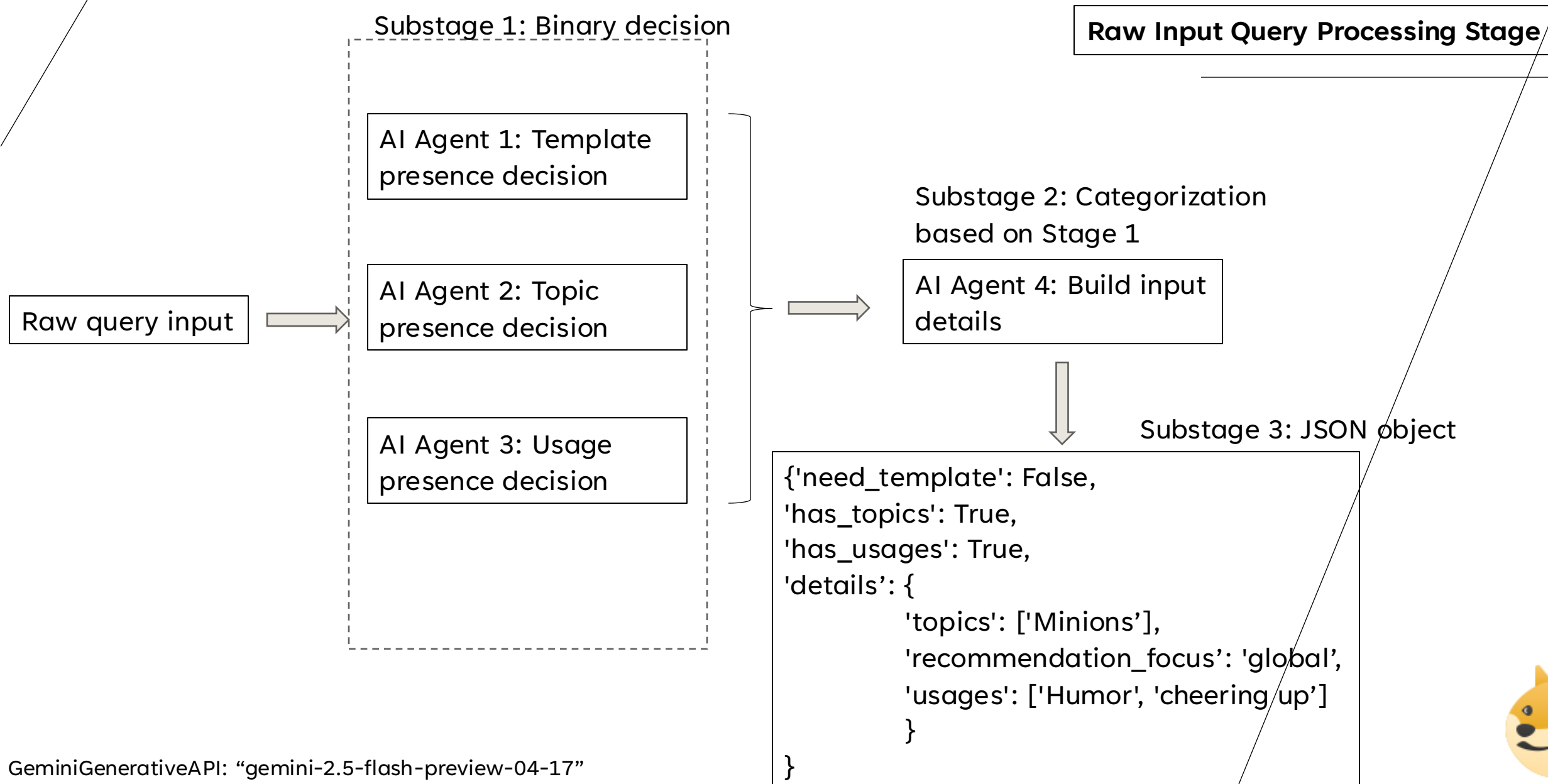
Star Wars

Star Trek

Battlestar Galactica

embeddings
- noTemplate_noTopic_usage.npy
- noTemplate_topicGlobal_noUsage.npy
- noTemplate_topicGlobal_usage.npy
- noTemplate_topicLocal_noUsage.npy
- noTemplate_topicLocal_usage.npy
- template_noTopic_usage.npy
- template_topic_noUsage.npy
- template_topic_usage.npy

*Embeddings are generated by SentenceTransformer("all-mpnet-base-v2")

Feature 3: Natural Language Query

Query: My day has been really bad, I want a meme that can cheer me up. I really like the movie Despicable Me, give me some Minions memes.

Substage 1: Binary decision

Raw Input Query Processing Stage

AI Agent 1: Template presence decision

Substage 2: Categorization based on Stage 1

AI Agent 4: Build input details

Raw query input

AI Agent 2: Topic presence decision

Substage 3: JSON object

AI Agent 3: Usage presence decision

```
{'need_template': False,
'has_topics': True,
'has_usages': True,
'details': {
        'topics': ['Minions'],
        'recommendation_focus': 'global',
        'usages': ['Humor', 'cheering up']
        }
}
```

GeminiGenerativeAPI: "gemini-2.5-flash-preview-04-17"

Feature 3: Natural Language Query

Query: My day has been really bad, I want a meme that can cheer me up. I really like the movie Despicable Me, give me some Minions memes.

Semantic Searching & Ranking Stage

```
{'need_template': False,
'has_topics': True,
'has_usages': True,
'details': {
        'topics': ['Minions'],
        'recommendation_focus': 'global',
        'usages': ['Humor', 'cheering up']
        }
}
```

SentenceTransformers → User input embedding

Cosine Similarity with precomputed case-based embedding → Top n Recommendations

∨ embeddings
≡ noTemplate_noTopic_usage.npy
≡ noTemplate_topicGlobal_noUsage.npy
→ ≡ noTemplate_topicGlobal_usage.npy
≡ noTemplate_topicLocal_noUsage.npy
≡ noTemplate_topicLocal_usage.npy
≡ template_noTopic_usage.npy
≡ template_topic_noUsage.npy
≡ template_topic_usage.npy

Feature 3: Natural Language Query

Query: I just want to see some memes.

```
{'need_template': False,
'has_topics’: False,
'has_usages’: False,
'details’: {
           'sentiment_preference': 'neutral'
           }
}
```
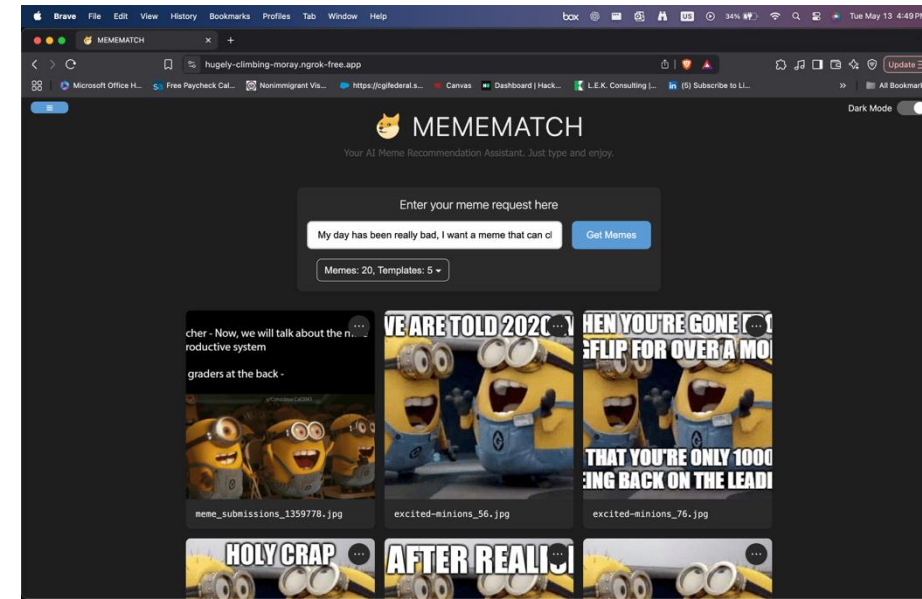
...

Sentiment score ranking

Top n
Recommendations

Frontend

Deployment

Web Interface

- **Frontend**: Built with HTML, CSS, and JavaScript for the user interface.
- **Backend**: Powered by FastAPI to handle HTTP requests and responses via RESTful endpoints.
- **Tunneling**: Ngrok is used to expose the local FastAPI server to the internet through a secure tunnel.

Android App

- **UI Layer**.
- **ViewModel Layer**.
- **Backend Communication**: Retrofit is used to interact with a Python-based backend API.
- **Firebase Integration:** Firebase Authentication, Firestore Database.
- **Extra features**: Favorites, sharing to social media platforms.
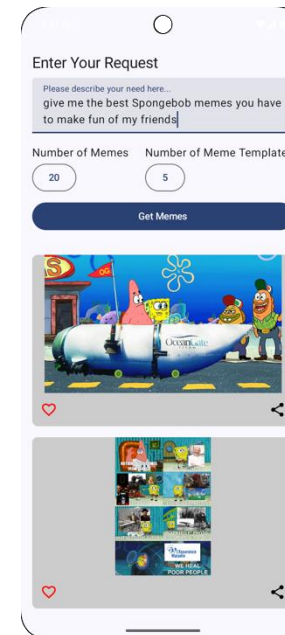
Demo!

III. PERFORMANCE EVALUATION & ANALYSIS
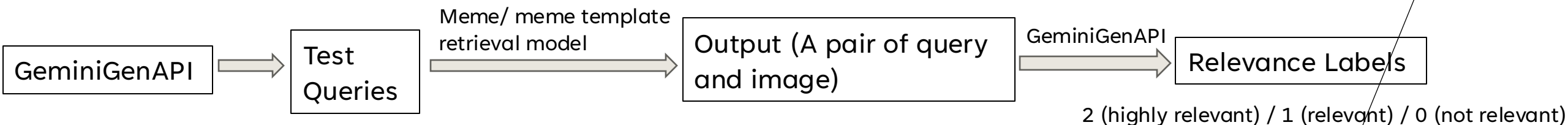
1. Semi-automated Test Label Generation

How relevant are the meme/meme template retrieval results and the query?

- Generate **~200 meme queries** and **~200 meme template queries** by GeminiGenAPI that are diverse in lengths, topics, and expressions.

- **20 results** per meme query.

- **10 results** per meme template query.

```
GeminiGenAPI  →  Test        →  Meme/ meme template  →  Output (A pair of query  →  GeminiGenAPI  →  Relevance Labels
                 Queries         retrieval model          and image)
```

2 (highly relevant) / 1 (relevant) / 0 (not relevant)

GeminiGenerativeAPI: "gemini-2.5-flash-preview-04-17"

```
memes mocking the idea of networking on LinkedIn with strangers,2,test_images/socially-awesome-awkward-penguin_12.png
memes mocking the idea of networking on LinkedIn with strangers,2,test_images/Victoria_David_Beckham_Be_Honest_1.png
memes mocking the idea of networking on LinkedIn with strangers,1,test_images/meme_submissions_1330569.png
memes mocking the idea of networking on LinkedIn with strangers,0,test_images/Expectation_vs_Reality_5.png
memes mocking the idea of networking on LinkedIn with strangers,2,test_images/meme_submissions_1362907.png
memes mocking the idea of networking on LinkedIn with strangers,0,test_images/meme_submissions_1217150.png
memes mocking the idea of networking on LinkedIn with strangers,1,test_images/pepe-silvia_8.png
memes mocking the idea of networking on LinkedIn with strangers,1,test_images/meme_submissions_1393434.png
memes mocking the idea of networking on LinkedIn with strangers,2,test_images/Peter_Parker_Reading_Book__Crying_9.png
memes mocking the idea of networking on LinkedIn with strangers,2,test_images/meme_submissions_1226418.png
memes mocking the idea of networking on LinkedIn with strangers,1,test_images/meme_submissions_1249349.png
memes mocking the idea of networking on LinkedIn with strangers,1,test_images/meme_submissions_1478706.png
memes mocking the idea of networking on LinkedIn with strangers,0,test_images/squidward-pointing_9.png
memes mocking the idea of networking on LinkedIn with strangers,2,test_images/meme_submissions_1386498.png
memes mocking the idea of networking on LinkedIn with strangers,1,test_images/meme_submissions_1458924.png
memes mocking the idea of networking on LinkedIn with strangers,1,test_images/neo-dodging-a-bullet-matrix_10.png
memes mocking the idea of networking on LinkedIn with strangers,2,test_images/meme_submissions_1293717.png
memes mocking the idea of networking on LinkedIn with strangers,1,test_images/meme_submissions_1345435.png
memes mocking the idea of networking on LinkedIn with strangers,1,test_images/meme_submissions_1266685.png
memes mocking the idea of networking on LinkedIn with strangers,2,test_images/math-ladyconfused-lady-135.png
```

2. Evaluation Result

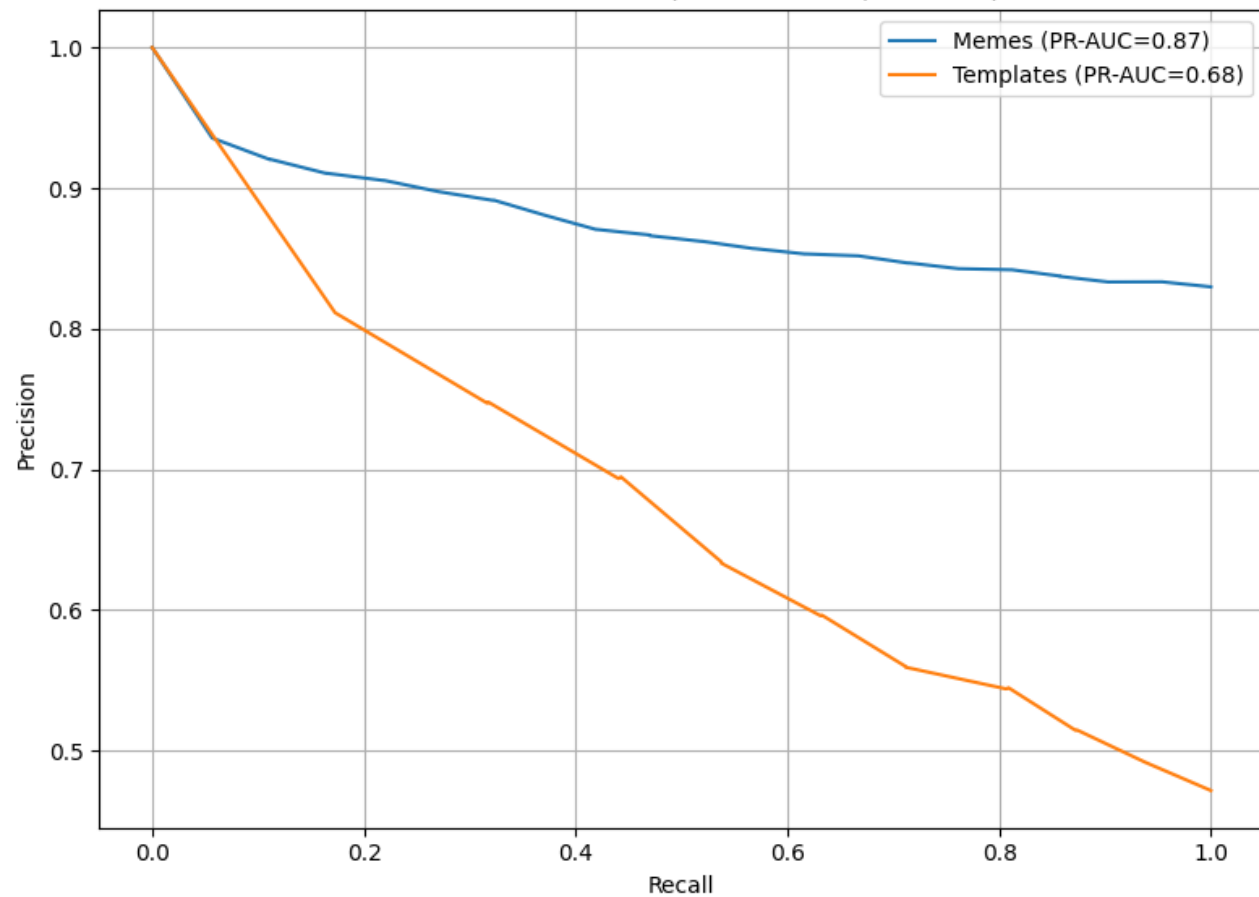- Threshold = 1, **relevant**

Evaluation results for meme retrieval	Evaluation results for meme template retrieval
nDCG@5: 0.6383 Precision@5: 0.8465 Recall@5: 0.2590 nDCG@10: 0.6837 Precision@10: 0.8446 Recall@10: 0.5151 nDCG@20: 0.8411 Precision@20: 0.8255 Recall@20: 1.0000 MAP: 0.8635 MRR: 0.9163	nDCG@5: 0.5008 Precision@5: 0.4960 Recall@5: 0.5233 nDCG@10: 0.7029 Precision@10: 0.4697 Recall@10: 1.0000 MAP: 0.5982 MRR: 0.6558

o The meme retriever is very strong. It's precise, retrieves all relevant memes, and ranks them well.

o The **meme template retriever finds all relevant results** (great recall).

o Compared to meme retrieval, **template performance is weaker**, especially in **early precision** and **ranking sharpness** (due to the nature of meme template and the size of template dataset).

2. Evaluation Result

- Threshold = 2, **highly relevant**

Evaluation results for meme retrieval	Evaluation results for meme template retrieval
nDCG@5: 0.6383 Precision@5: 0.4901 Recall@5: 0.2503 nDCG@10: 0.6837 Precision@10: 0.4871 Recall@10: 0.4893 nDCG@20: 0.8411 Precision@20: 0.4869 Recall@20: 0.9802 MAP: 0.5623 MRR: 0.6578	nDCG@5: 0.5008 Precision@5: 0.3714 Recall@5: 0.5021 nDCG@10: 0.7029 Precision@10: 0.3417 Recall@10: 0.9657 MAP: 0.4944 MRR: 0.5509

o Model does **very well in finding** highly relevant memes. There's **room to improve early ranking** (P@5/10).

o The template retriever **finds the right items**, but doesn't **rank them aggressively enough at the top**.

o The performance of the **meme template retriever** is **decent but weaker than meme retriever**, especially on **early precision**.
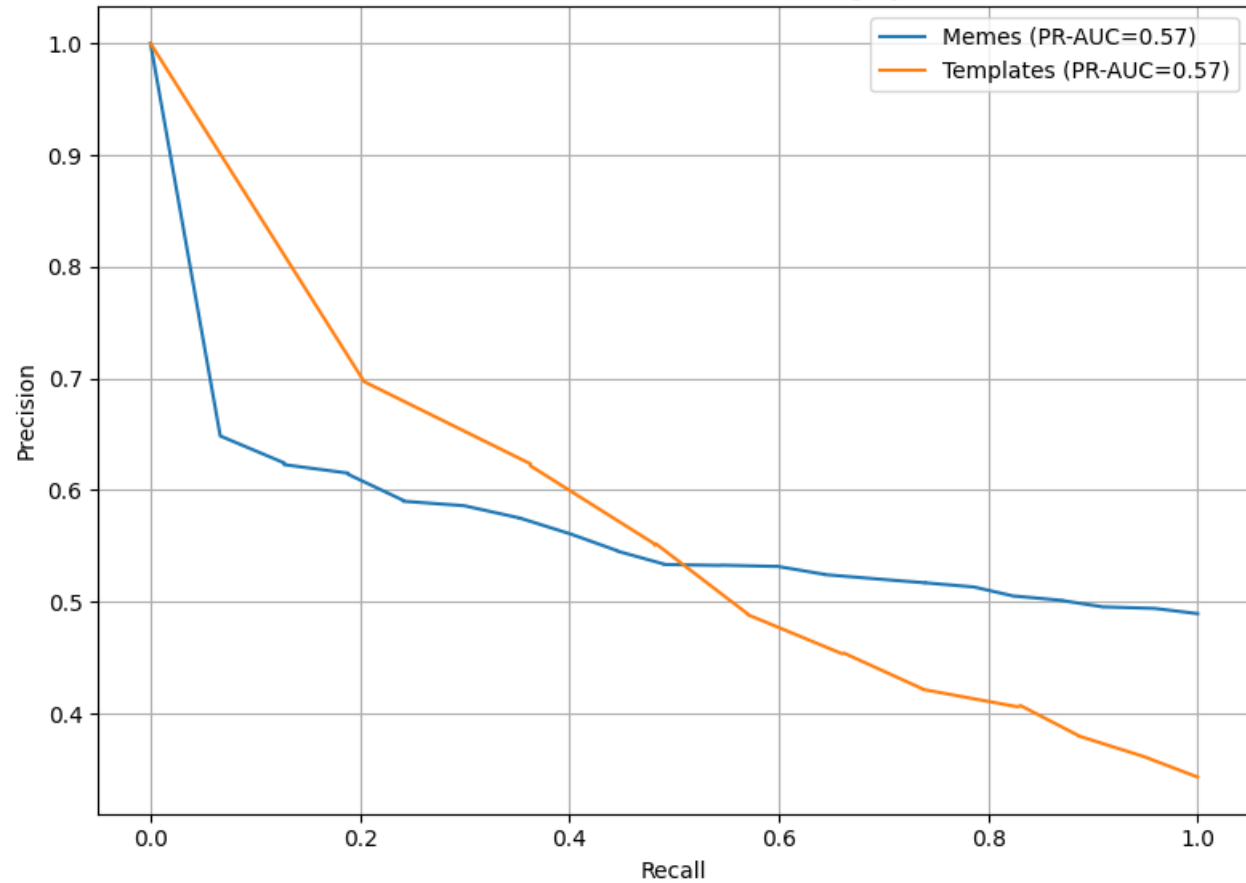
Precision-Recall Curve (threshold = 1, relevant)

- Memes (PR-AUC=0.87)
- Templates (PR-AUC=0.68)
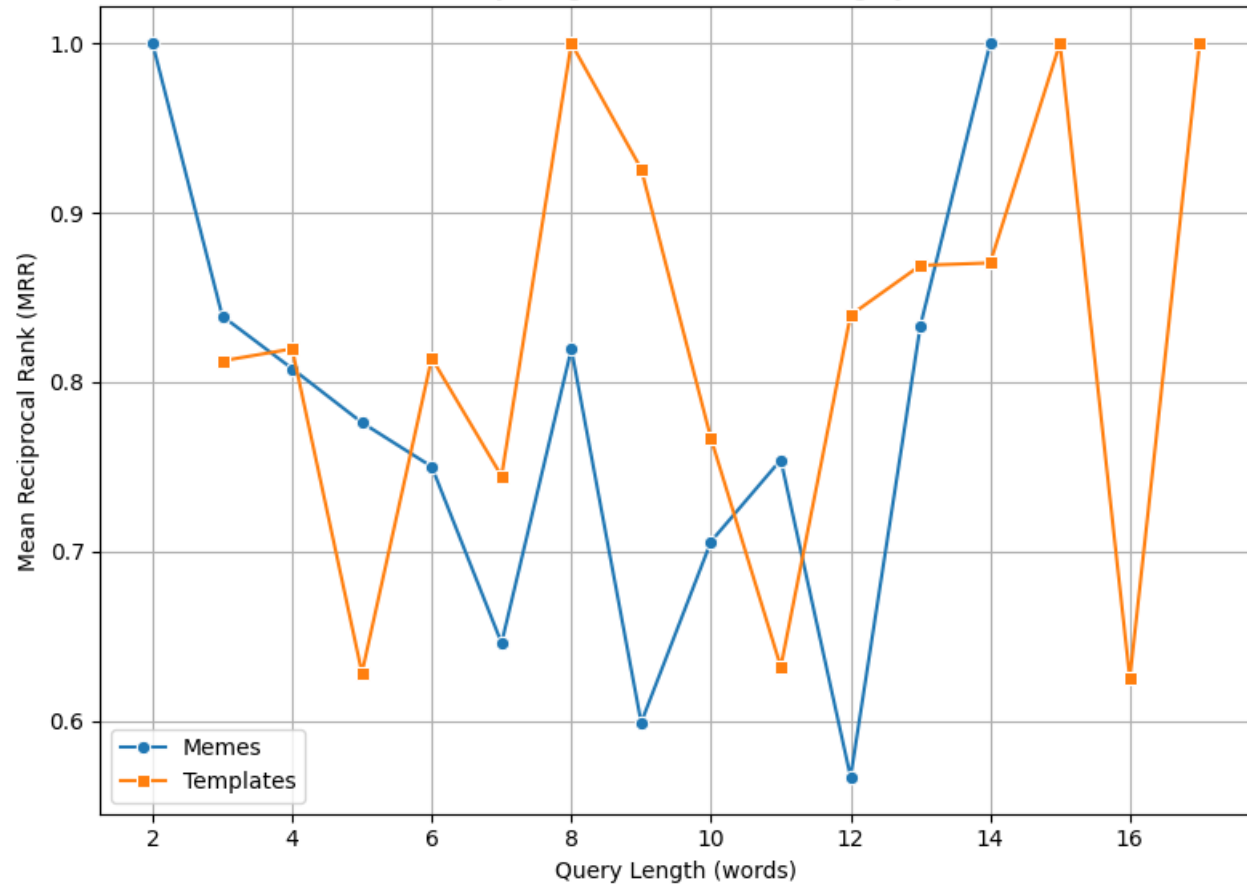
MRR vs. Query Length (threshold = 1, relevant)
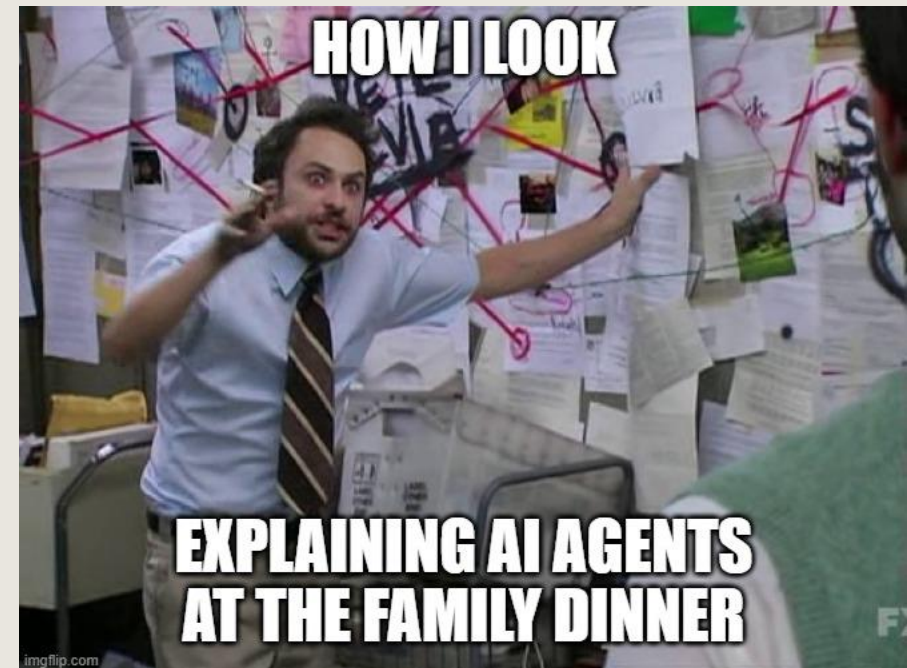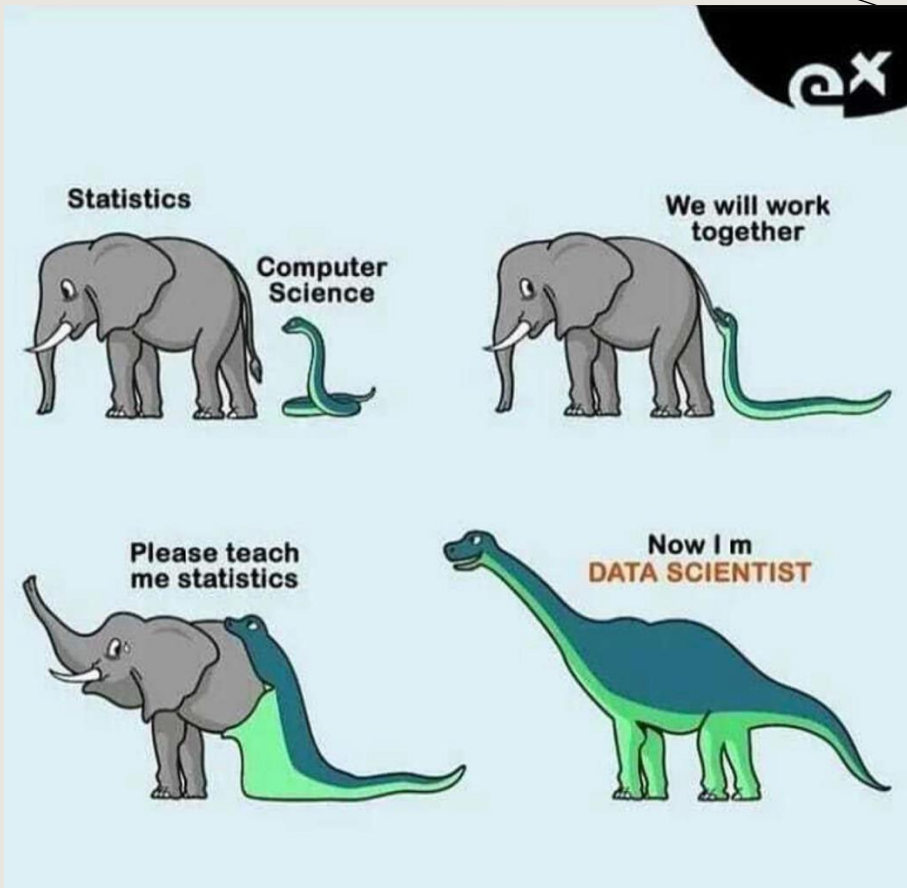
- Memes
- Templates
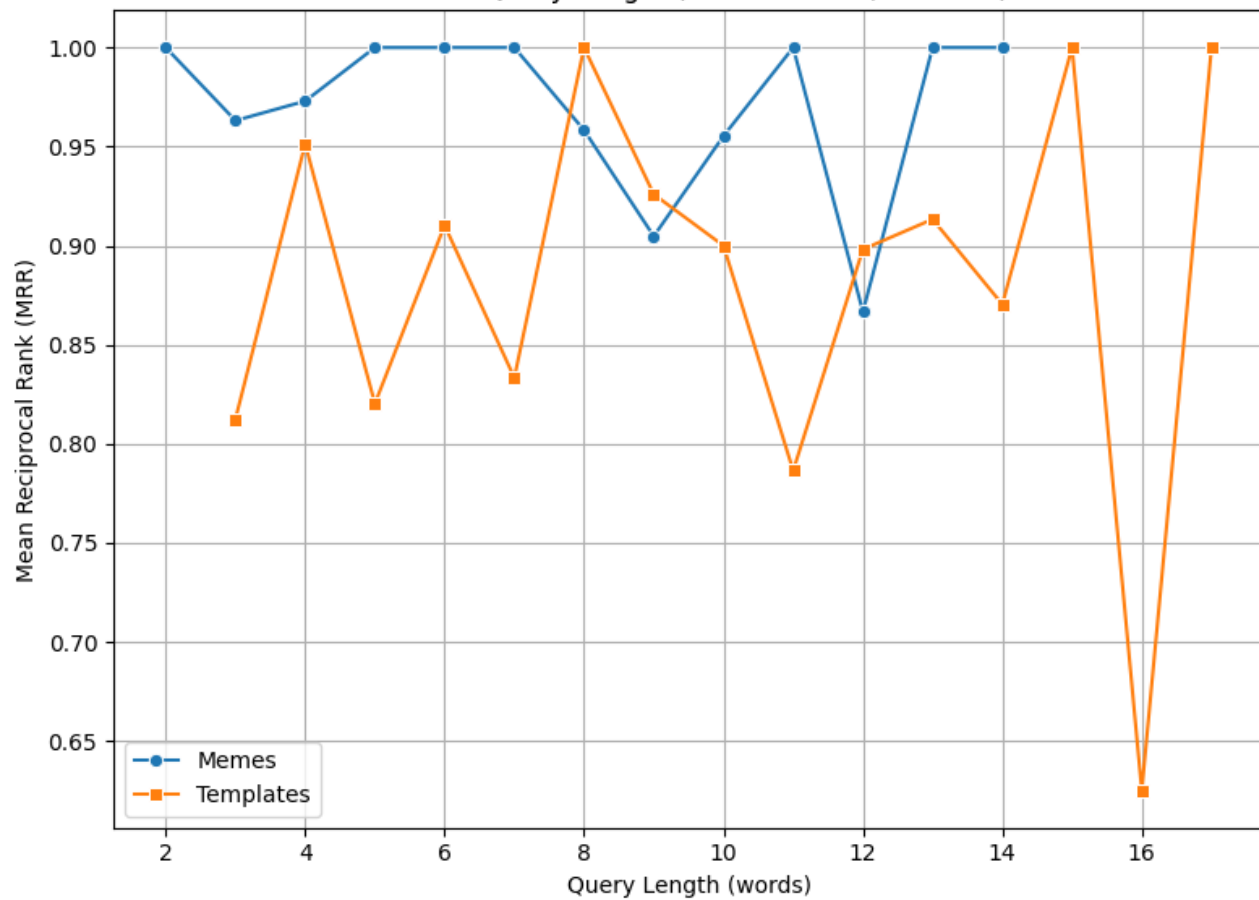
Precision-Recall Curve (threshold = 2, highly relevant)

MRR vs. Query Length (threshold = 2, highly relevant)

IV. FUTURE WORK

III. Future Work

- Update and increase the meme dataset.

- Extend the scope of the project to GIFs dataset.

- Improve the performance of the existing text-to-image model.

- Manually checking the text-to-image model performance on small batch.

- IOS deployment.

How Neural Networks work?
Neurons:
I know a guy who knows a guy
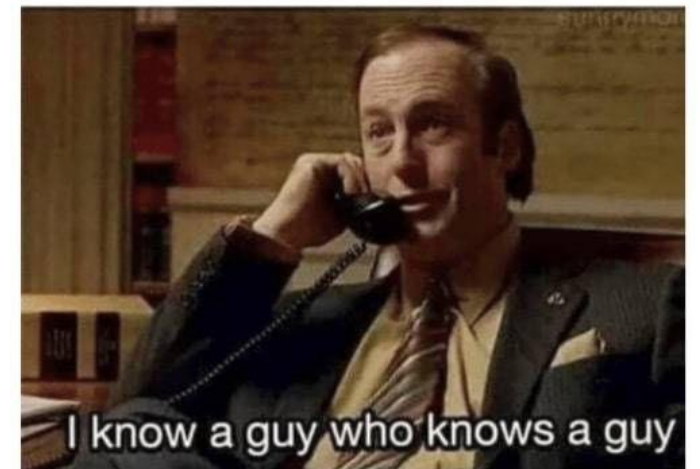
MemeMatch finds the perfect meme for every moment.

THANK YOU

Tri An Le

Emails: triandole@gmail.com

Contact: +36 20 492 3602