

Intelligent Reasoning Systems Project



Mango - An Academic & Knowledge Platform

Date: Oct. 28th, 2023

Copyright ❤️ Group 24 and Project Mango Team

Group Members

- Goh Min Hua A0285810A
- Li Jiacheng A0285823W
- Mao Zhihong A0285799X

Introduction



Mango is an academic knowledge platform for college students and researchers.

It has two main components -

- A **Search Engine** for essays with their summaries.
- An **Academic Papers Recommender** for the latest and valuable academic papers.

Introduction

- The platform aims to assist **young college students** and **researchers** to quickly grab a general understanding of a particular field.
- It can also help them to keep up to date on the forefront development and progress of that field.



Project Scope

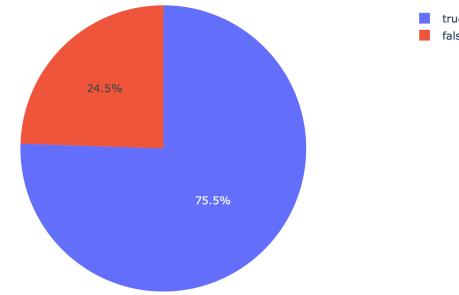
- Current approach of academic knowledge discovery tightly depends on tutors and senior students.
- The process can be ***laborious*** and ***time-wasting***.
- Users may struggle and get lost in the sea of unfamiliar concepts and academic papers.



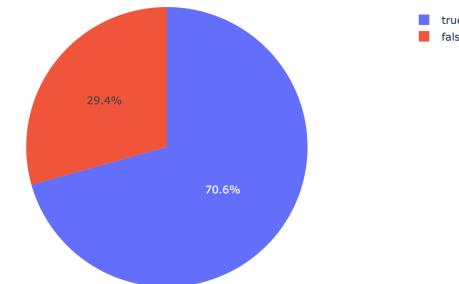
Project Scope

- Based on our survey, we found that over 75% percents of people feel like keeping up to the academic progress.
- And more than 70% are interested in a dedicated platform.

Q2: Are you interested in the latest academic progress and may check the news occasionally?

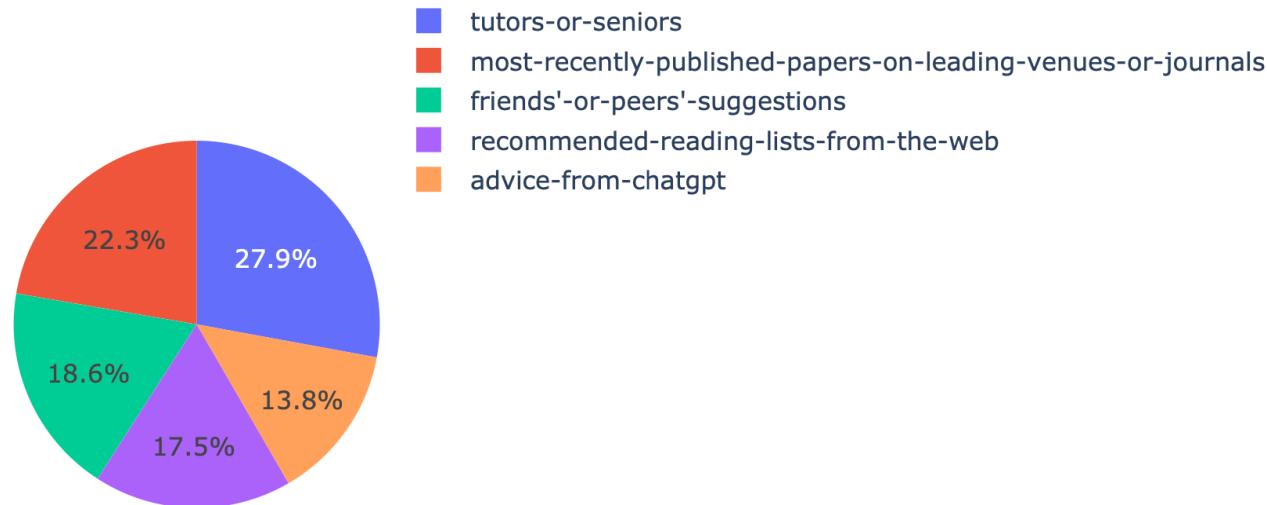


Q4: Are you interested in a platform for gathering and presenting up-to-date academic progress?



Project Scope

Q8: Who/ what do you rely on to know what papers to read?
(yes, exclude the times when you don't know what to read)



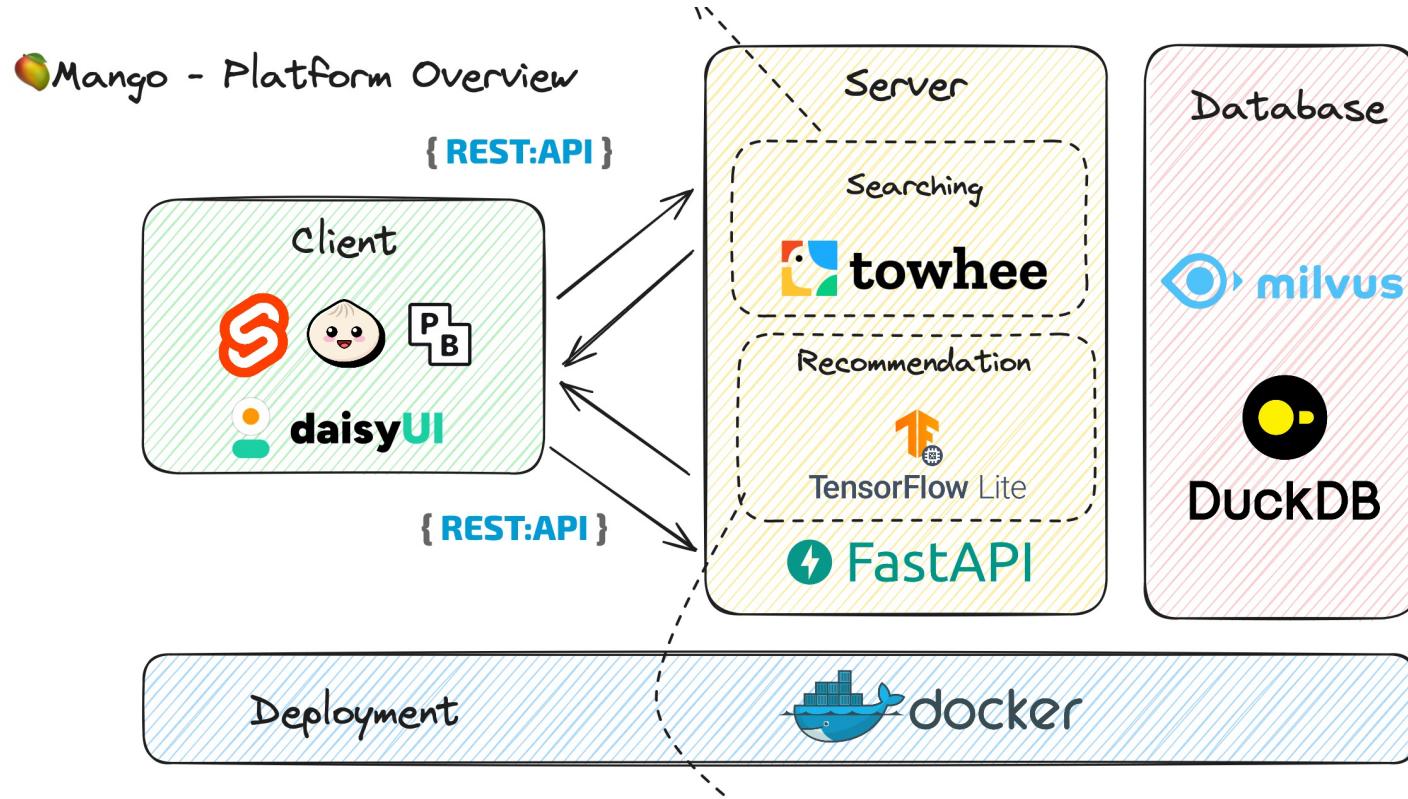
Data Sources:

- Open Source academic papers details and side information provided by **Semantic Scholar** and **Paper With Code**
- 10,4453 papers in total.

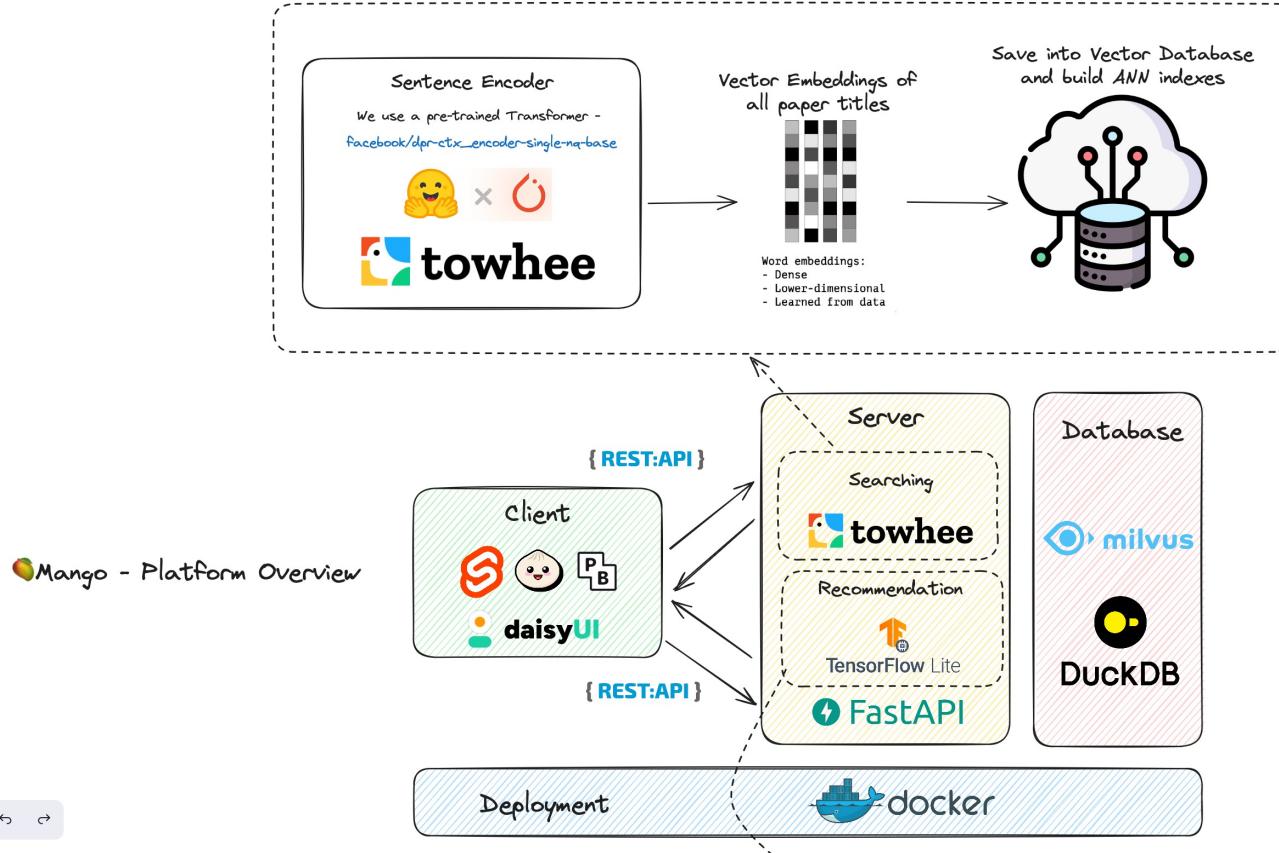
Academic environment:

- Questionnaire survey of the target population to understand the insights and corresponding needs in a specific area.

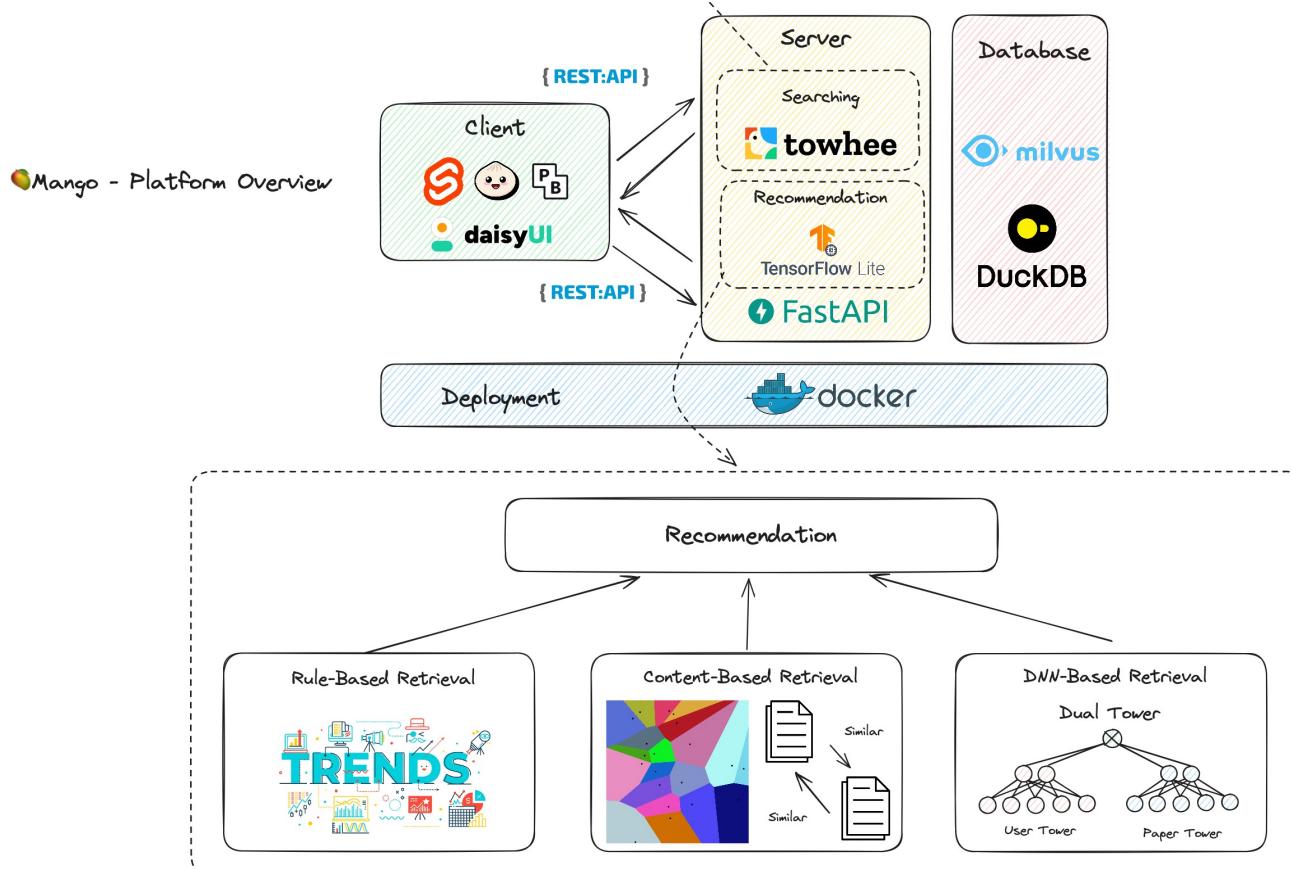
System Overview



System Design - Search Engine



System Design - Recommendation



Challenge:

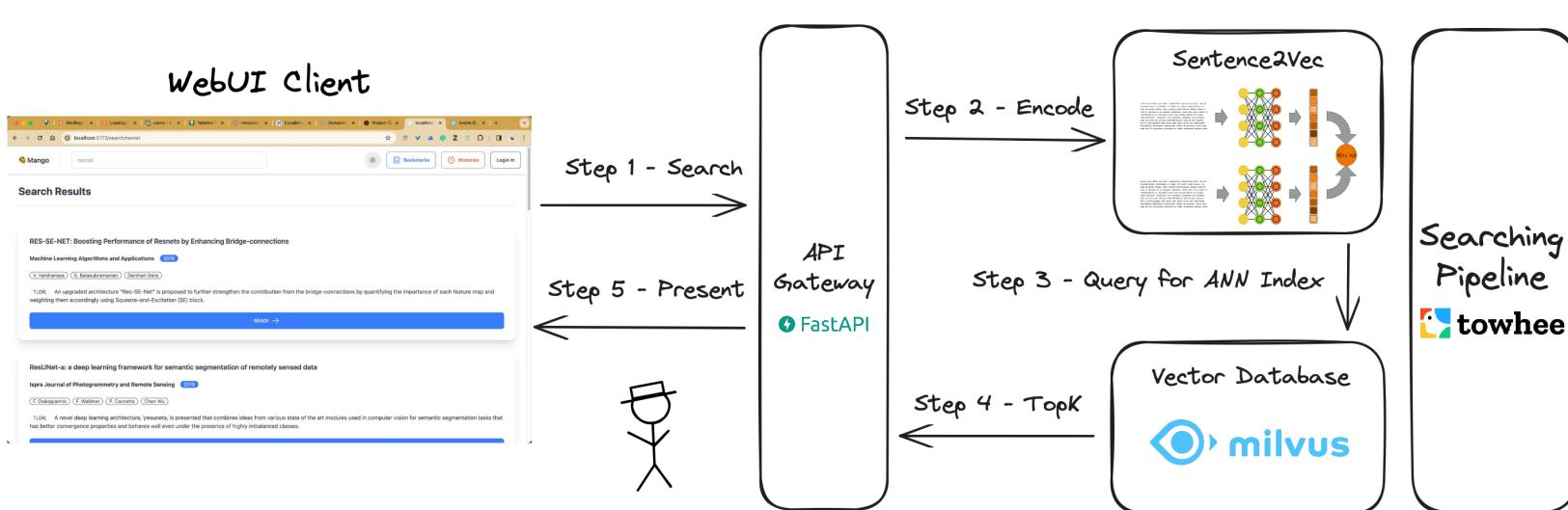
- Sparse information due to the vastness of academic fields.
- Platform is in its starting stage, which means we are lacking user interaction data.

Strategy:

- Focus on a selected fields, filter papers with rich information.
- Diversify feedback channels and employ various cold start techniques.

Implementation - Search Engine

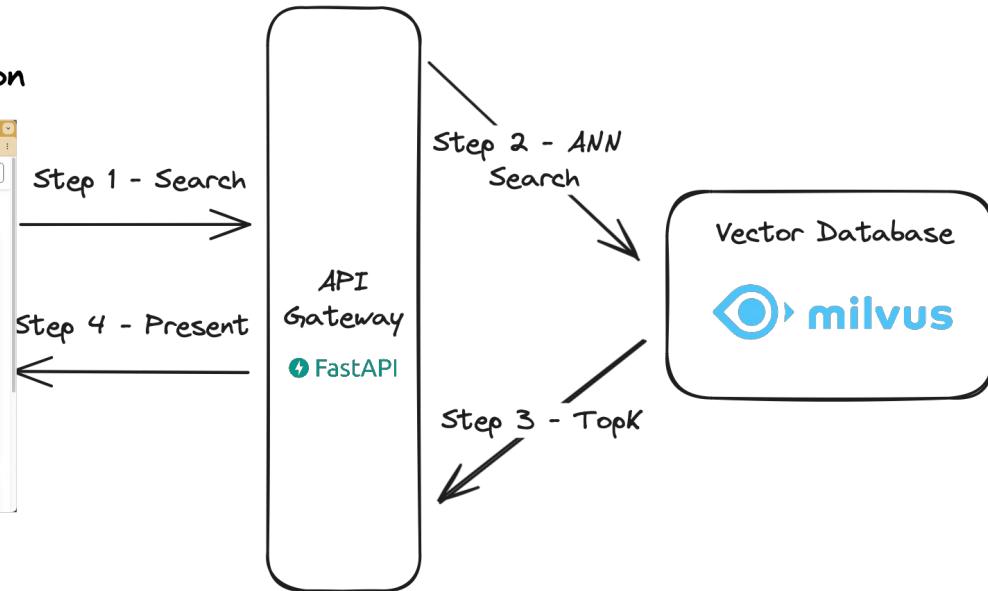
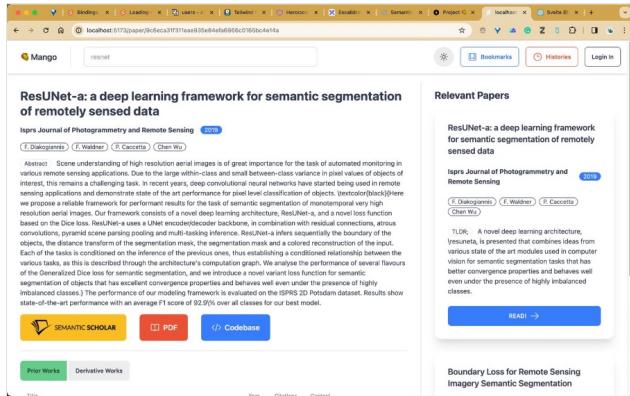
Searching Pipeline - Dense Passage Retrieval



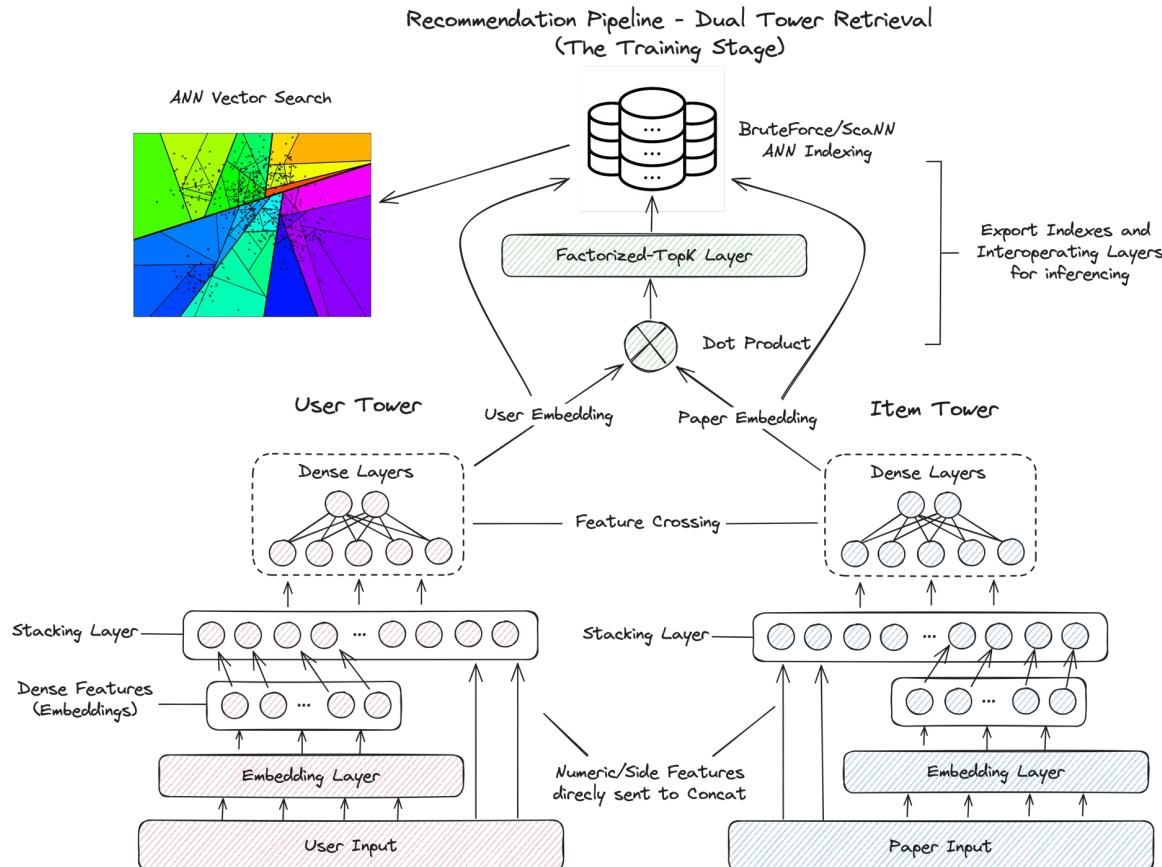
Implementation - Recommendation

Recommendation Pipeline - Embedding Retrieval (Content-Based/Similarity-Based)

Detail page needs
relevant paper recommendation

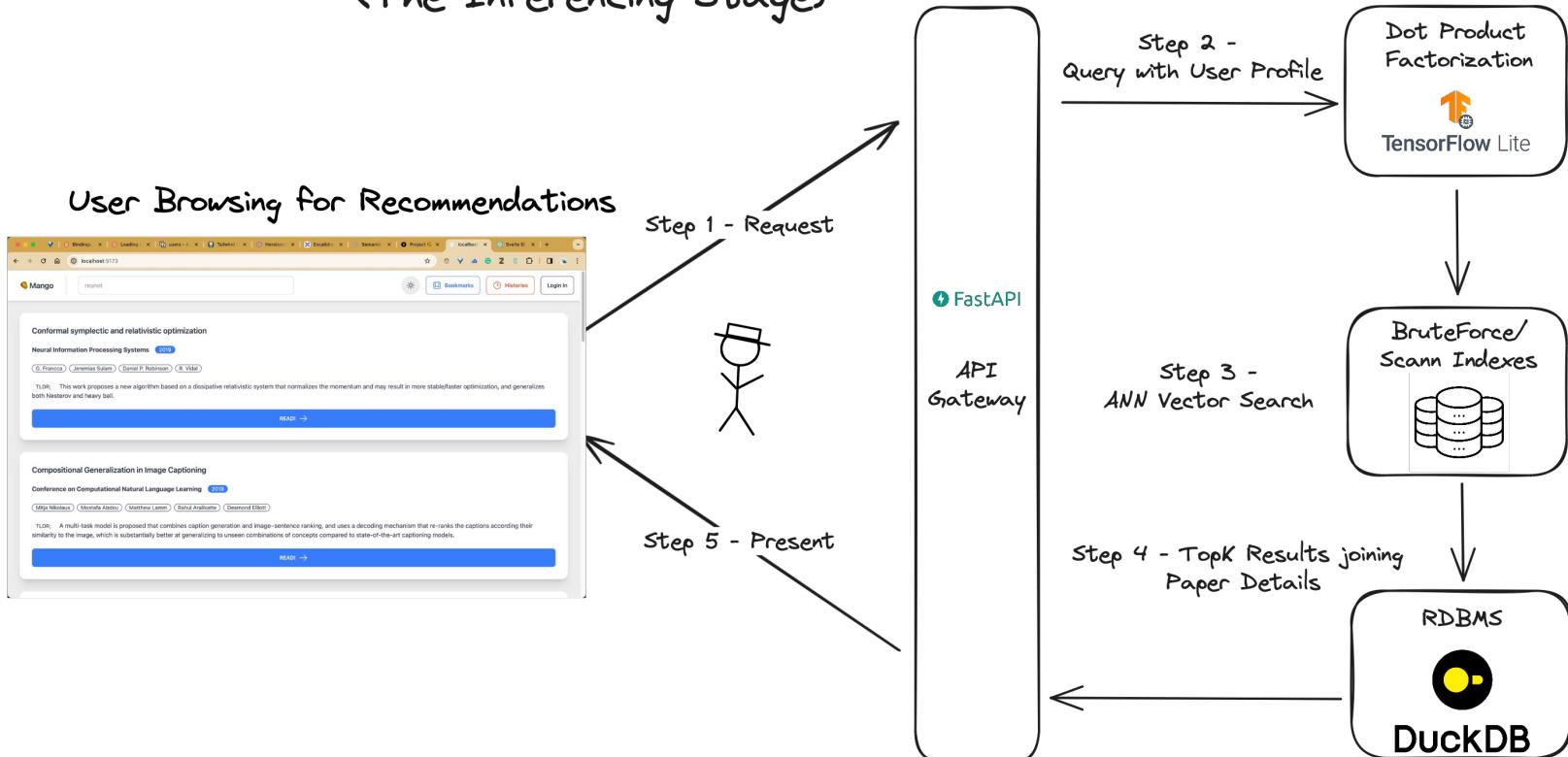


Implementation - Recommendation

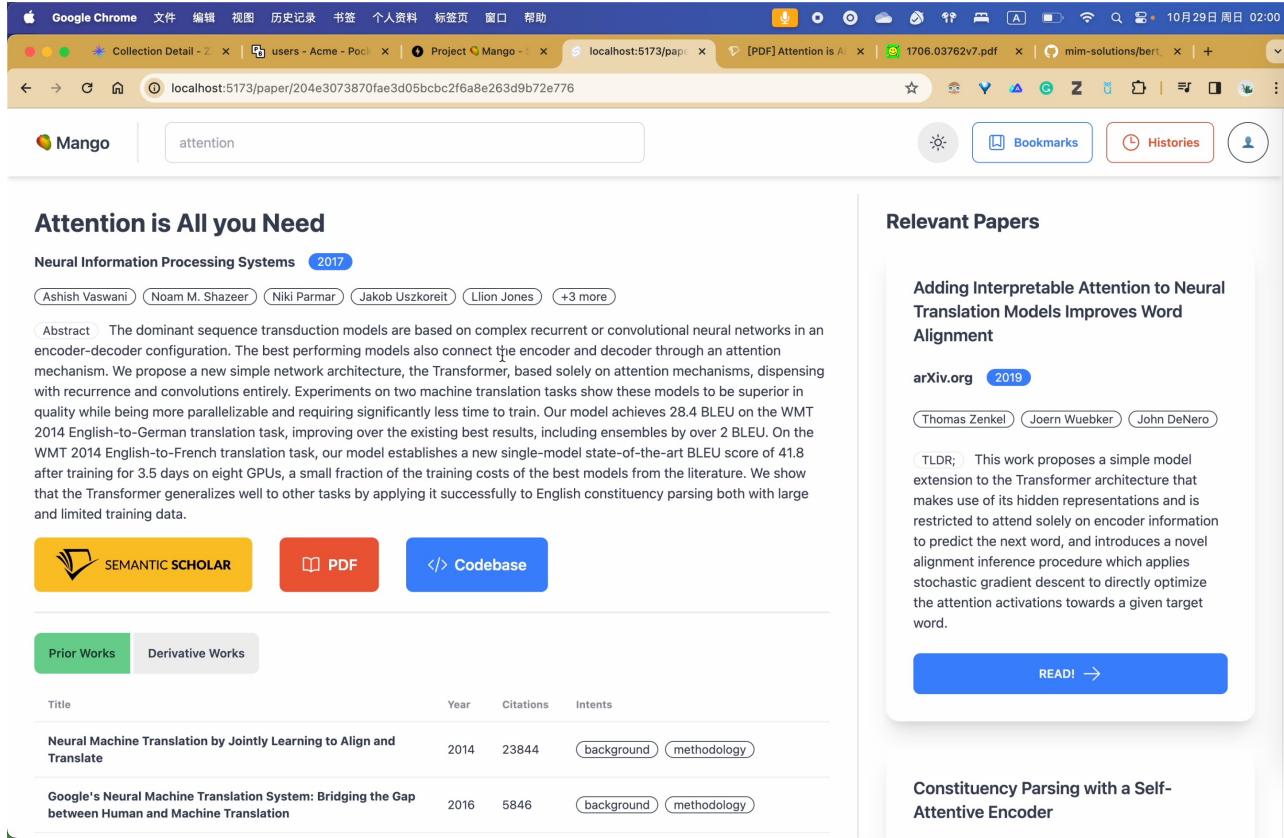


Implementation - Recommendation

Recommendation Pipeline - Dual Tower Retrieval (The Inferencing Stage)



Show Time



The screenshot shows a Google Chrome browser window with multiple tabs open. The main content is a research paper titled "Attention is All you Need" from the Neural Information Processing Systems conference in 2017. The paper's authors are Ashish Vaswani, Noam M. Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, and others. The abstract discusses the Transformer model, highlighting its superior performance and efficiency compared to previous sequence transduction models. Below the abstract are three buttons: "SEMANTIC SCHOLAR", "PDF", and "Codebase". A "Prior Works" tab is currently selected, showing two related papers: "Neural Machine Translation by Jointly Learning to Align and Translate" (2014) and "Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation" (2016). To the right, a sidebar titled "Relevant Papers" lists another paper: "Adding Interpretable Attention to Neural Translation Models Improves Word Alignment" from arXiv.org in 2019. This paper's abstract is summarized as proposing a simple extension to the Transformer architecture to attend solely on encoder information for word alignment.

Attention is All you Need

Neural Information Processing Systems 2017

(Ashish Vaswani) (Noam M. Shazeer) (Niki Parmar) (Jakob Uszkoreit) (Llion Jones) (+3 more)

Abstract The dominant sequence transduction models are based on complex recurrent or convolutional neural networks in an encoder-decoder configuration. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.

SEMANTIC SCHOLAR PDF Codebase

Prior Works Derivative Works

Title	Year	Citations	Intents
Neural Machine Translation by Jointly Learning to Align and Translate	2014	23844	(background) (methodology)
Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation	2016	5846	(background) (methodology)

Relevant Papers

Adding Interpretable Attention to Neural Translation Models Improves Word Alignment

arXiv.org 2019

(Thomas Zenkel) (Jörn Wuebker) (John DeNero)

TLDR; This work proposes a simple model extension to the Transformer architecture that makes use of its hidden representations and is restricted to attend solely on encoder information to predict the next word, and introduces a novel alignment inference procedure which applies stochastic gradient descent to directly optimize the attention activations towards a given target word.

READ! →

Constituency Parsing with a Self-Attentive Encoder

Future Work

- Improve the recommendation system by adding a dedicated **Ranking layer** to perform smarter ranking.
- Involve **Graph Embeddings** to get better retrieval accuracy.
- Set up **automation pipeline** of implicit and explicit feedback collection.
- Keep polishing **user interface**.

Conclusion

- In essence, Mango is not just a platform; it's a ***transformative*** tool designed to empower an entire generation of researchers.
- The overwhelmingly positive user feedback further emphasizes the platform's significance in bridging the gap between ***academic research*** and ***real-world applications***.

Thank you for listening!

Date: Oct. 28th, 2023

Copyright ❤️ Group 24 and Project Mango Team