

# Text-to-Image Recommendation System

A Harvard Capstone Project with The Associated Press



Karina Huang, Dianne Lee, Abhimanyu Vasisht, Phoebe Wong

## INTRODUCTION

The Associated Press (AP) is a cooperative, not-for-profit news agency that hosts an image database of over 35 million photographs for internal and external publishing services.

We build an interactive user-interface to recommend AP images for any input article using the article headline, text and tags, exonerating editors from having to manually search for relevant images using the following models (and an ensemble of the models):

- Ranked Tag Recommendation with TextRank<sup>1</sup> and Tag Importance Ranking<sup>2</sup>
- Tag-to-Tag Recommendation with k-Nearest Neighbors and Word2vec Embeddings
- Text-to-Text Recommendation with GloVe and Universal Sentence Encoder (USE) Embeddings

## DATA

### Articles

22,602 news articles published between Feb – Sep 2019

**Headline:** Essence Fest marks 25 years of celebrating black culture

**Tags:** New Orleans, Mary J.Blige, African American, Fourth of July ...

**Summary:** The Essence Festival draws thousands to New Orleans during the fourth of July week and is celebrating 25 years of bringing African American women of all ages together for thought-provoking conversation and performances from top musical acts. ...The festival is July 5-7 and will include a performance by Mary J. Blige.

### Images

78,021 images associated with news articles

**Headline:** Essence Festival 25

**Tags:** United States, Fairs and festivals, Recreation and leisure, Lifestyle

**Caption:** ... H.E.R. performs at the Coachella Music & Arts Festival at the Empire Polo Club in Indio, Calif. ...



## RANKED TAG RECOMMENDATION SYSTEM

### I. Article Tag Ranking: TextRank

- A graph-based ranking model that ranks tags with "importance"
- Multiplies word co-occurrence matrix with words' weights
- A word's weight increases if it occurs with other important words

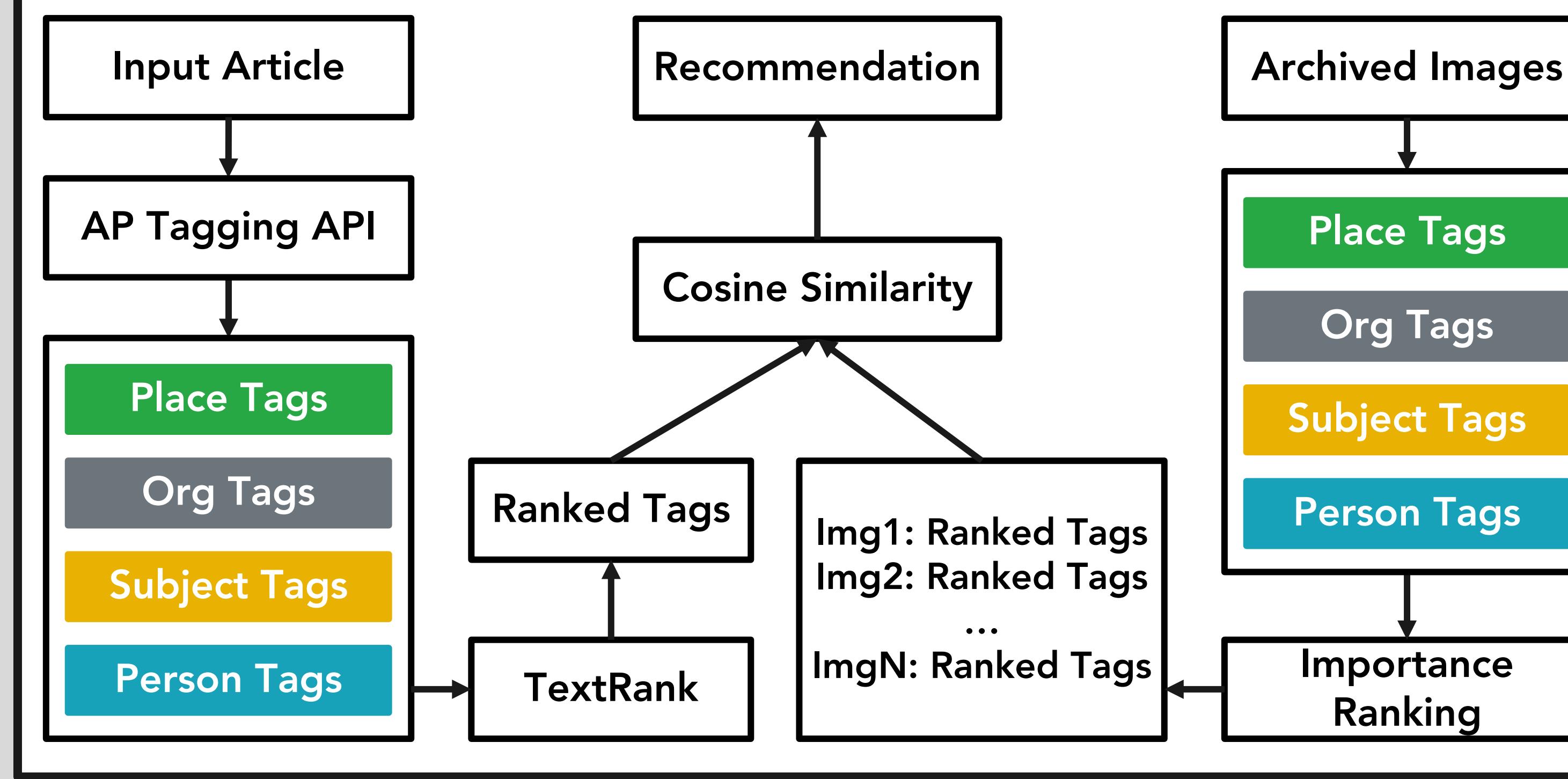
TextRank Bags of Words	TextRank Score	Named Entities
organization Nevada Policy Research Ins...	.023	Nevada
Supportive Nevada lawmaker	.014	Nevada
Speaker Barbara Buckley	.008	Barbara Buckley
State Legislatures	.005	Legislature
Carson City	.004	Carson City

### II. Image Tag Ranking: Tag Importance Ranking

- Discount object tags [person] importance by
  - occurrence
  - number of descriptions
- Upvote scene tags [org, place, subject] importance if
  - tag appear in sentence
  - tag is the subject of the sentence

### III. Ranked Tag Recommendation System

- (✓) Prioritize relevant and semantically related content
- (✗) Long training time; rely on quality of tags



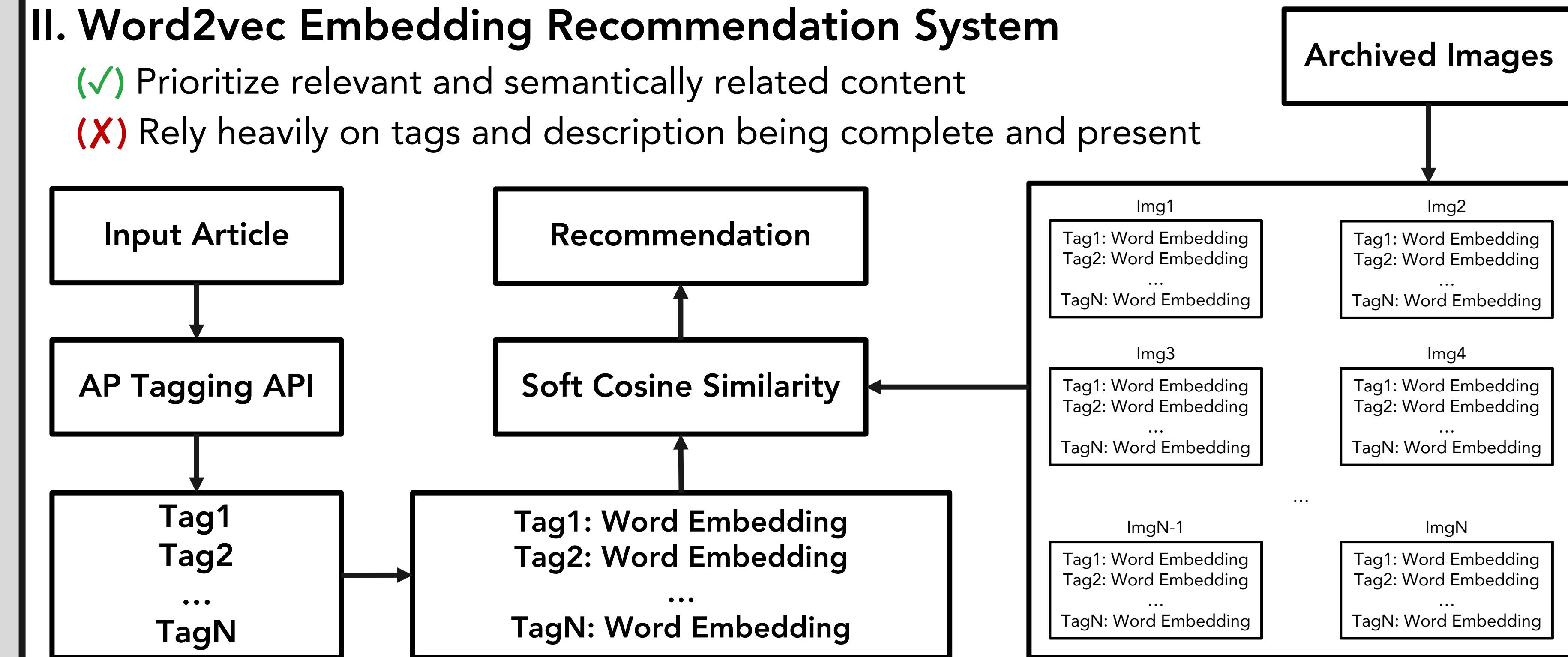
## TAG-TO-TAG RECOMMENDATION SYSTEM

### I. K-Nearest Neighbors Recommendation System [Baseline Algorithm]

- (✓) Exact tag overlap between article and image tags
- (✗) All tags weighted equally; rely on quality of tags

### II. Word2vec Embedding Recommendation System

- (✓) Prioritize relevant and semantically related content
- (✗) Rely heavily on tags and description being complete and present



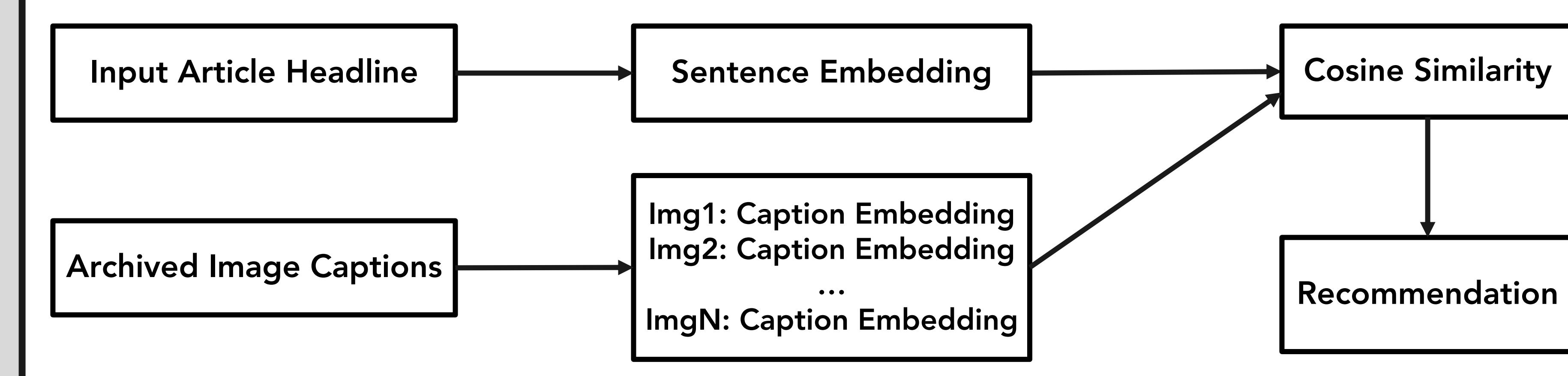
## TEXT-TO-TEXT RECOMMENDATION SYSTEM

### I. GloVe Embedding Recommendation System

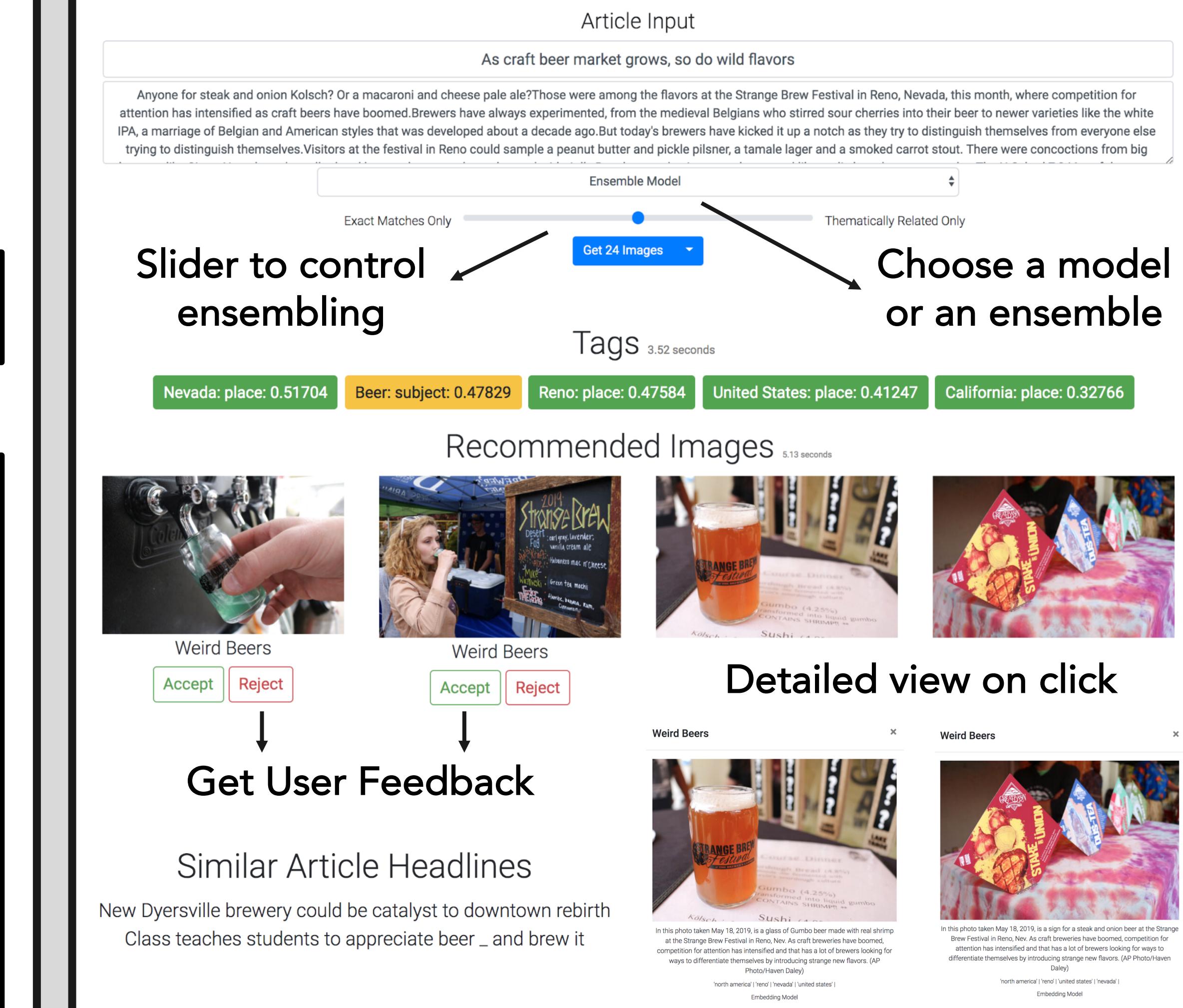
- (✓) Fast and captures semantic similarity between words
- (✗) Does not consider the order of words in headline

### II. USE Embedding Recommendation System

- (✓) Fast, simple and context-aware sentence encoding
- (✗) Lower interpretability compared with other approaches

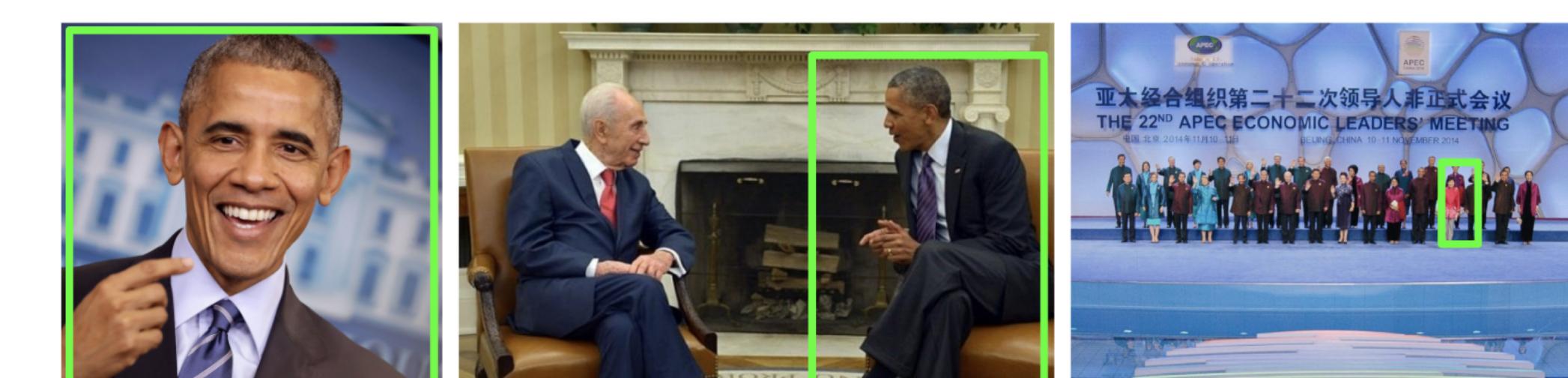


## USER INTERFACE



## DISCUSSION

- Using visual information from photos to improve predictions using Faster RCNN. This was not scalable (7-10s per image).



- Word Mover's Distance used to directly measure similarity between article text and image headlines (15 minutes per article).
- SpaCy Named Entity Recognition used in place of AP's tagging API but this was not as comprehensive.