Algorithms for Mastodon

Stanford | ENGINEERING
Computer Science

Jason Chon[RA] Max Rodriguez[RA] Will Schulz[Mentor]

Computer Science, Stanford

Introduction

The feed ranking algorithm is the core of large social media platforms. It dictates the type of content that billions of users are immediately exposed to, orchestrating how users might behave and think within and outside the virtual space [1].

Working with the Algorithms for Mastodon team, we implemented an end-to-end social media platform using an ML-based engagement prediction feed ranking algorithm. We use humanistic-AI bots and a dataset consisting of tens of millions of posts (being updated as you read this) to create a realistic and engaging server-wide virtual environment.

Research question: What are the qualities of a highly successful feed-ranking algorithm?

System Overview

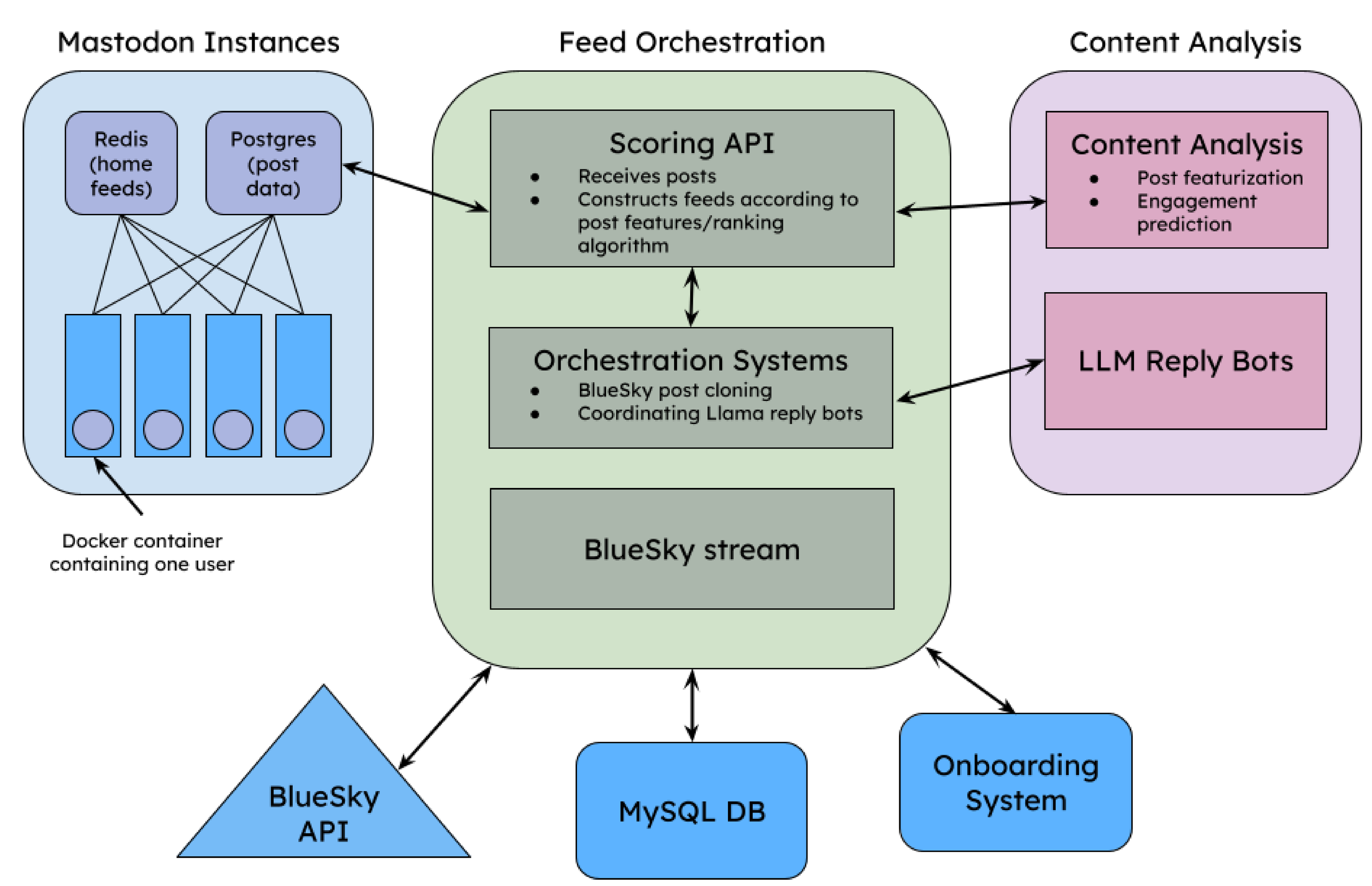

Figure 1. Overview of our system based on Mastodon.

Research Goals

Develop a Useful System

We are creating a tool that can be used by any social-media researcher. To that end, our system is modular and designed to provide as much control as desired over users' experiences on the platform, allowing researchers to run rigorous studies in a realistic social media environment.

Testing Our Algorithms

Using the components we've created, we've begun developing our own content-ranking algorithms to later apply to real human participants and gather concrete behavioral data.

A Fully-Configurable Social Media Platform

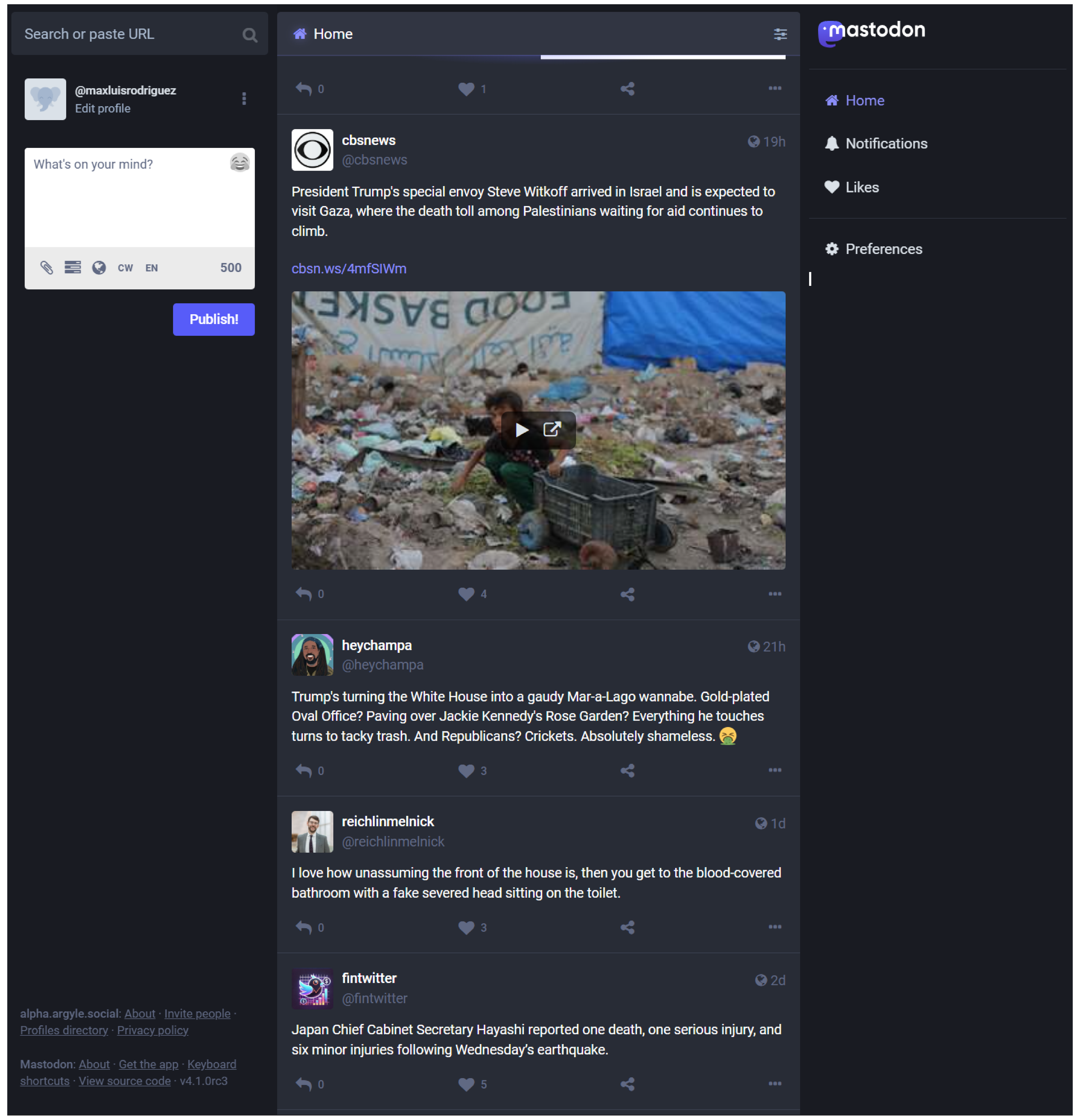

Figure 2. Example home feed resulting from a basic political up-ranking algorithm.

A High-Level Overview

Tools we developed to control our social media environment include:

- Content analysis pipeline that featurizes posts along several dimensions: political content, emotional sentiment, quality, etc.
- Custom APIs for managing many Mastodon instances, scoring posts, updating feeds, and more
- CLI tools and modular scripts for defining custom feed-ranking algorithms
- Changes to the Mastodon front and back ends
- A first-pass at content ranking based on engagement prediction

References

[1] H. K. Azzaakiyyah.
The impact of social media use on social interaction in contemporary society.
Technology and Society Perspectives (TACIT), 1(1):1–9, 2023.

Engagement Prediction Algorithm

We use a deep multilayer perceptron (MLP) to predict whether a given user will interact with a novel social media post.

$$\mathcal{L}_{\text{BCE}}(y,\hat{y}) = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)\right]$$

Figure 3. BCE loss function used for training the MLP. Here, $y_i \in \{0,1\}$ denotes the true label, where 1 indicates the user liked the post and 0 indicates the user did not like the post. $\hat{y}_i$ indicates the predicted probability, and N the batch size.

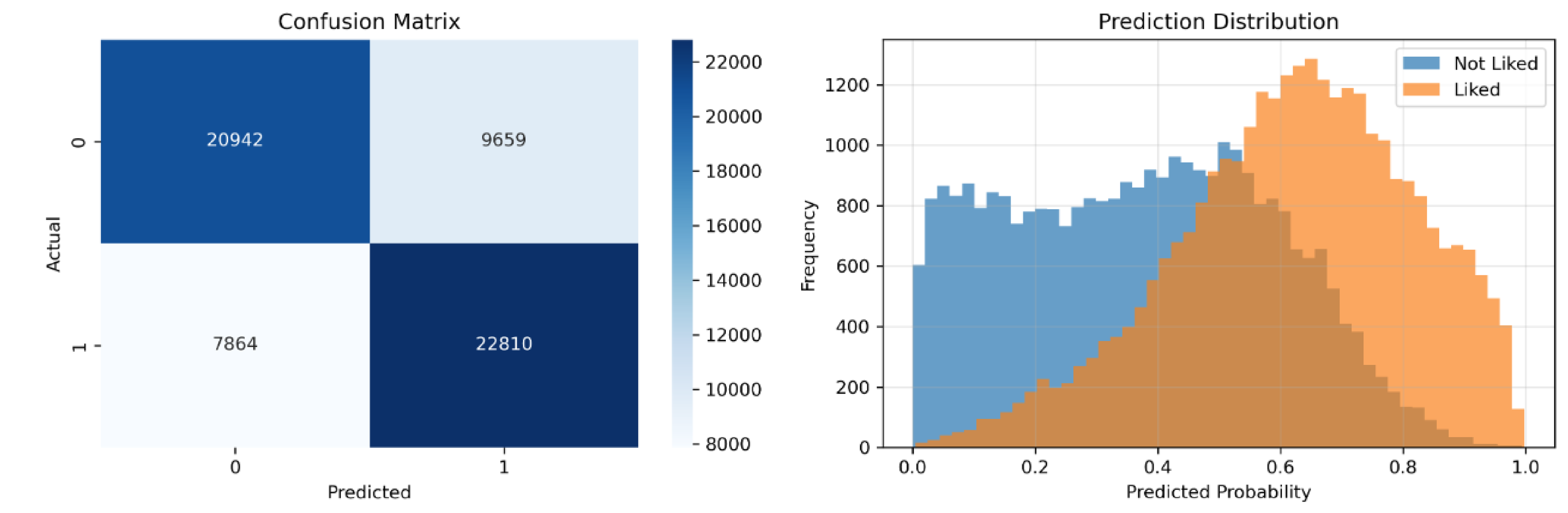

Figure 4. MLP prediction results on our holdout dataset prior to content re-leveling.

Train Data Feature Embedding

Each training sample combines a Bluesky user's engagement history with a candidate post labeled as liked or not. Engagement history is encoded using Sentence-BERT for text and ResNet18 for images, applied to 50% of the user's liked posts. The remaining liked posts and an equal number of non-liked posts serve as prediction targets.

Data Rebalancing for Inclusive Content

To prevent the over-representation of majority users' interests during feed generation, we use an iterative Kmeans + Gini coefficient minimization down-sampling loop to effectively maximize content diversity.

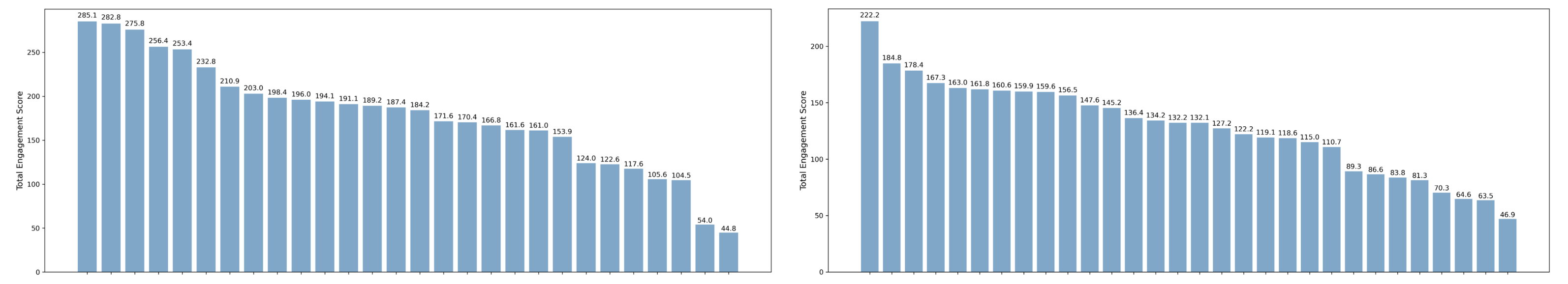

Figure 5. Content clusters before re-leveling (left) and after 5 iterations (right).

Discussion & Future Works

Our research sheds light on the potential inner workings of engaging, large-scale social media platforms. With billions of active users worldwide, understanding how these platforms shape online behavior and shape thoughts and actions will remain of utmost importance.

Moving forward, we plan to:

- Fine-tune engagement prediction and fully integrate our ML-based algorithm into the pipeline
- Run a live study analyzing the effects of our algorithms on real people