Data Analytics Practice Opportunity 2024/25

Data Analysis of Weibo Rumors based on Bert Model

WENG Ya Ru (ISTM/1)
CHANG Yu (ISTM/1)

Purpose

This study aims to systematically investigate the dissemination mechanisms and socio-psychological impacts of online rumors through computational social science methods, focusing on COVID-19-related misinformation as a case study. By analyzing structural propagation patterns, linguistic strategies, and sentiment dynamics, the research seeks to uncover how misinformation cascades differ functionally and affectively from factual corrections. The ultimate goal is to generate actionable insights for three critical applications:
(1) early detection of high-risk misinformation cascades through network topology signatures;
(2) optimization of fact-checking interventions via emotion-aware rebuttals that neutralize anxiety-inducing narratives; and
(3) improvement of platform content moderation systems by integrating propagation pattern recognition algorithms to prioritize multi-hub diffusion clusters.

What is BERT model

The BERT (Bidirectional Encoder Representation from Transformers)[1] model is a pre-trained language model based on Transformers, using a bidirectional encoder structure to learn language representations through pre-training[2]. The BERT model has achieved significant results in the field of natural language processing and is widely used in tasks such as text classification, named entity recognition, and sentiment analysis.

Methodology

The BERT model is built on the Transformer architecture and uses stacked encoder layers to achieve bidirectional language understanding. Each encoder applies a multi-head self-attention mechanism to capture relationships among all words simultaneously, ensuring that each token's context includes both preceding and following words[3]. Additionally, a feedforward neural network refines these representations. BERT is pre-trained on two tasks: Masked Language Modeling (MLM), where some tokens are masked and predicted based on context, and Next Sentence Prediction (NSP), which trains the model to discern if one sentence follows another. This design results in rich, contextualized representations that excel in NLP applications[4].

Visualization

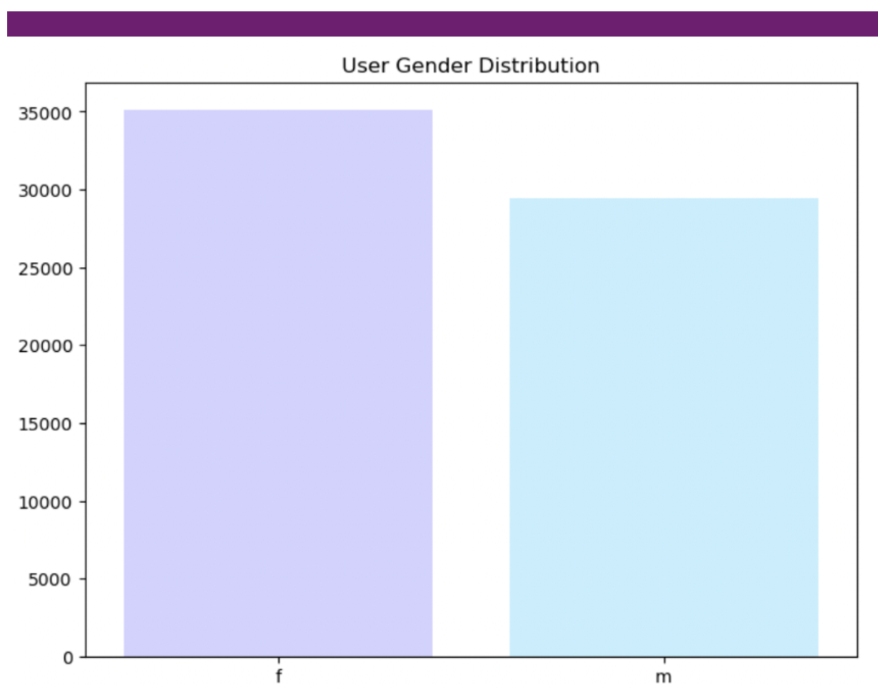

Figure 1 User Gender Distribution

- Female-majority user base
- Gender imbalance likely shapes popular content types
- Information spread may vary by topics that resonate differently with each gender

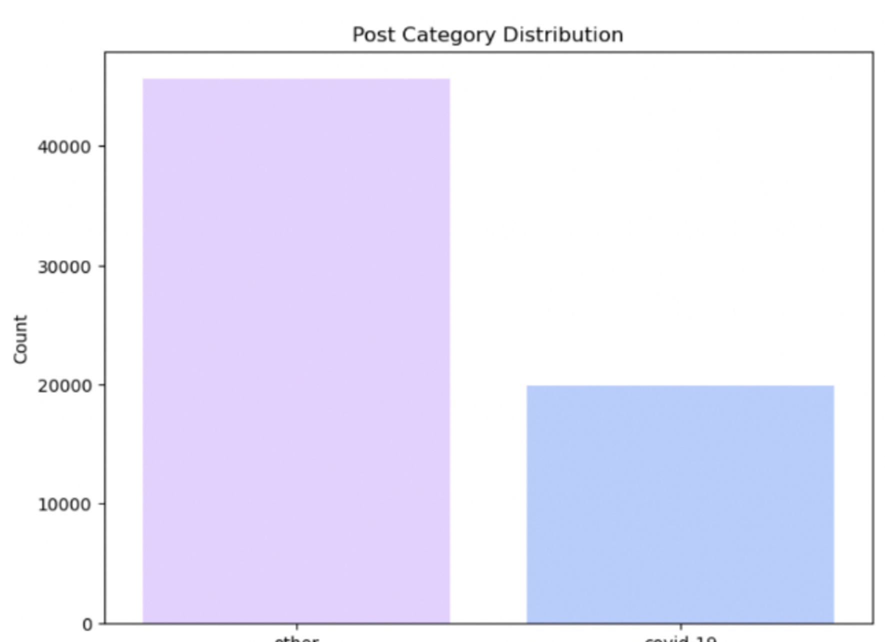

Figure 2 Post Category Distribution

- "Other" category leads in post count
- "COVID-19" posts still substantial despite lower count
- Pandemic remains a major discussion topic on the platform

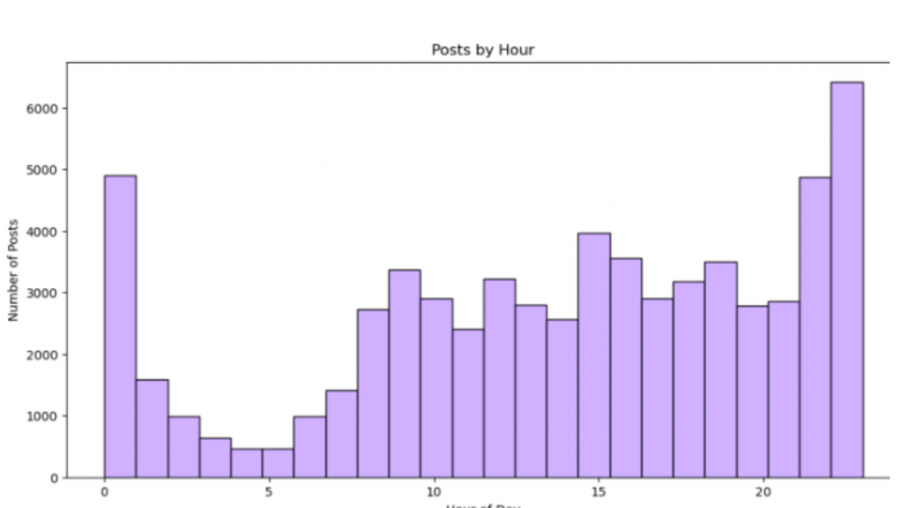

Figure 3 Post by Time

- Two activity peaks: 0–1 AM and 8–11 PM
- Positive posts peak 10 AM–3 PM (~5 000)
- Night lull 0–5 AM (<1 000)
- Daytime breaks (work/lunch) drive higher engagement
- "Golden window" during peaks for spreading positive/scientific info
- Time-sensitive strategy: push counter-rumor content in active hours

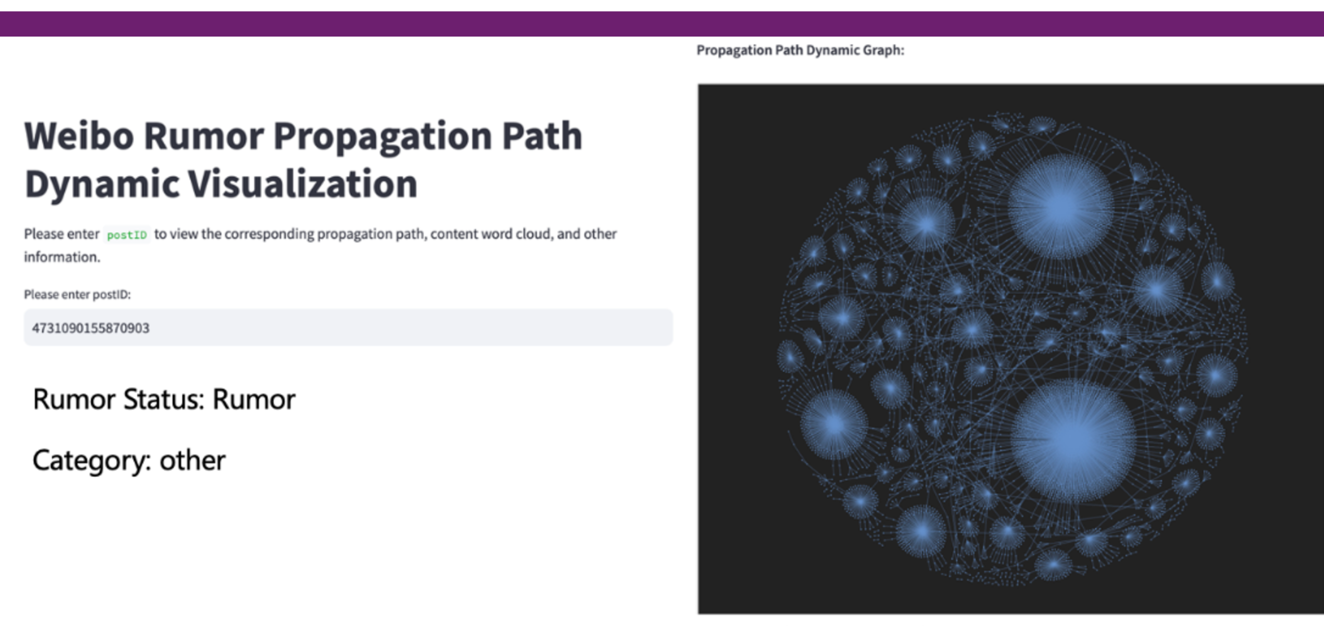

Figure 4 Propagation Path Dynamic Graph of other Rumor

In Figure 4,

- Developed an interactive Streamlit dashboard to analyze online rumor spread and content features
- Visualized diffusion via a Propagation Path Dynamic Graph
- Identified high-impact central nodes with radial dissemination patterns
- Confirmed multi-level spread: core propagators → secondary sharers → peripheral audiences
- Employed text mining to cluster posts thematically and sentiment analysis to gauge emotional polarity by tier

- Combined network topology and content semantics to pinpoint superspreaders, trace paths, and monitor sentiment shifts
- Enabled real-time filtering by node centrality, propagation depth, and sentiment thresholds for dynamic investigation

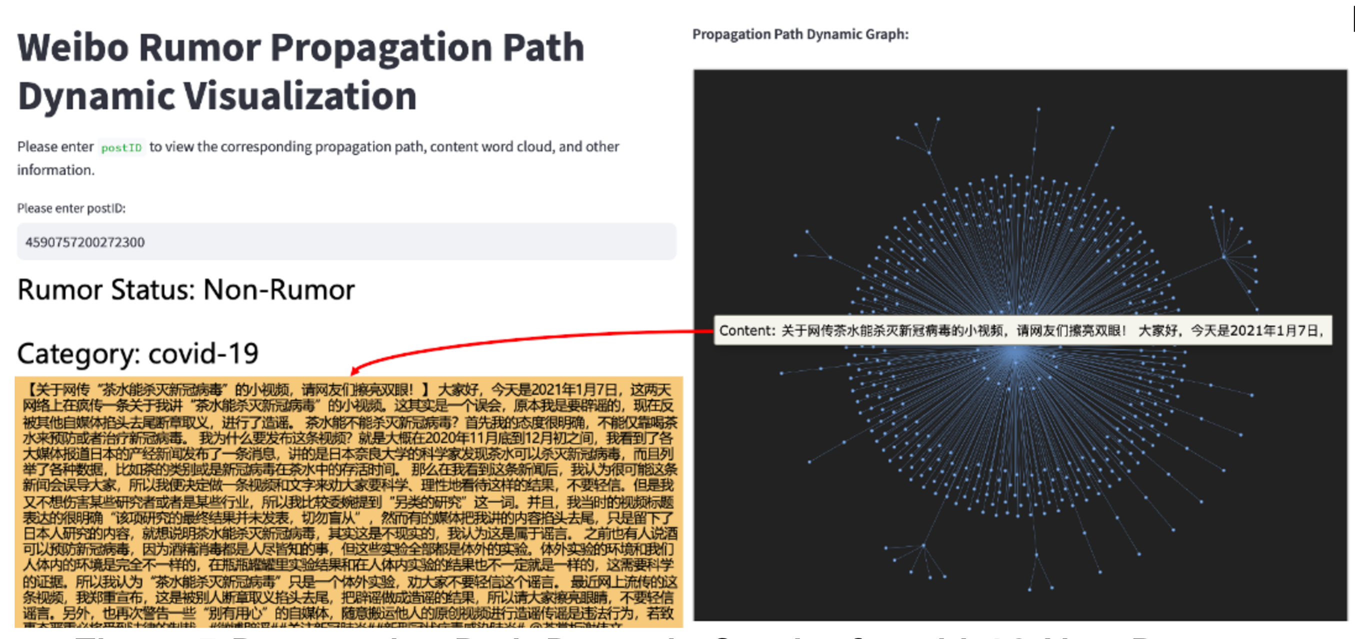

Figure 5 Propagation Path Dynamic Graph of covid-19 Non-Rumor

In Figure 5,

- Non-hoax COVID-19 post focused on debunking misinformation and public health education
- Content comprised factual corrections, not emotional or sensational narratives
- Streamlit dashboard showed a single-core diffusion: shares mainly from authoritative institutional accounts
- Rumor exhibited multi-hub propagation with numerous influential amplifiers
- Decentralized networks drive rumors' rapid, broad spread; centralized credibility yields slower, controlled fact-sharing
- Network topology and content type jointly shape information virality

Sentiment Score

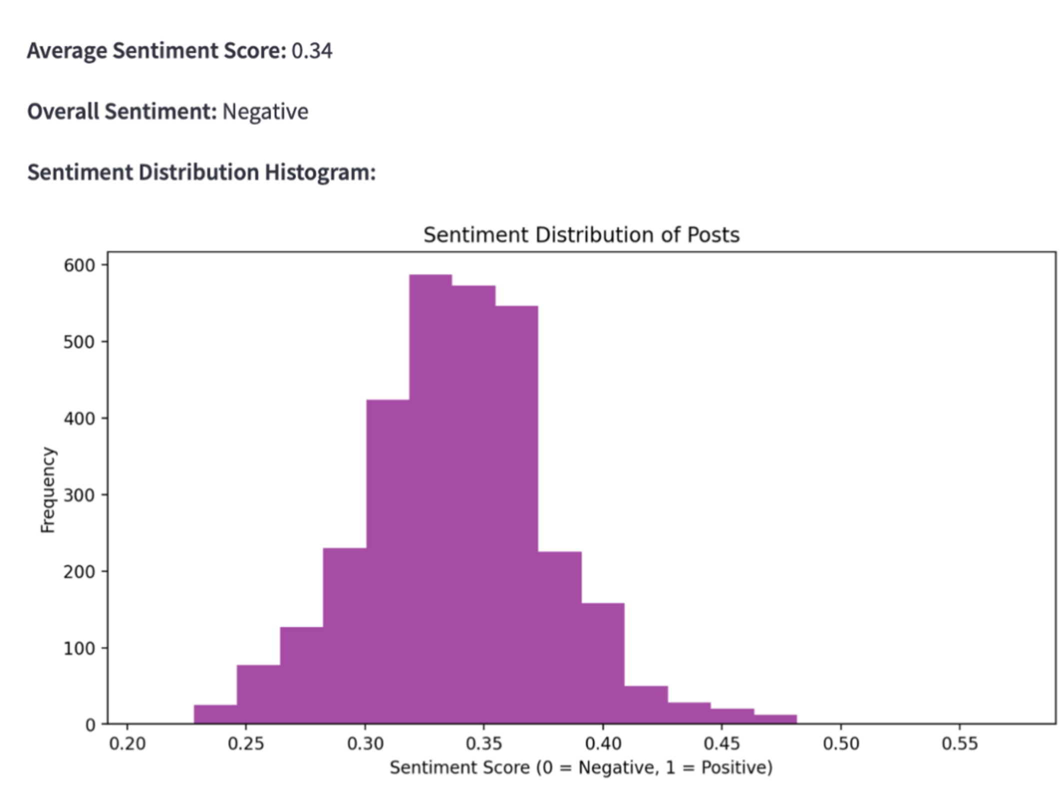

Figure 6 Sentiment Score Distribution of Rumor Post

We can see that in Figure 6:

- Average sentiment score: 0.34 (near neutral but classed "negative")
- Distribution nearly normal, peaking at 0.34
- 62.1% of scores in 0.3–0.4 (moderately negative)
- 23.7% in 0.2–0.3 (strongly negative)
- Only 14.2% exceed 0.45 (limited positive polarity)
- Lower-end skew reinforces overall negative classification

To operationalize these findings, we recommend:

- Identifying dates with sentiment minima (e.g., scores ≤0.2) through temporal granularity analysis.
- Conducting content audits on these critical periods to isolate linguistic or thematic triggers of extreme negativity.
- Developing mitigation strategies such as targeted counter-messaging or platform-level sentiment dampening interventions.

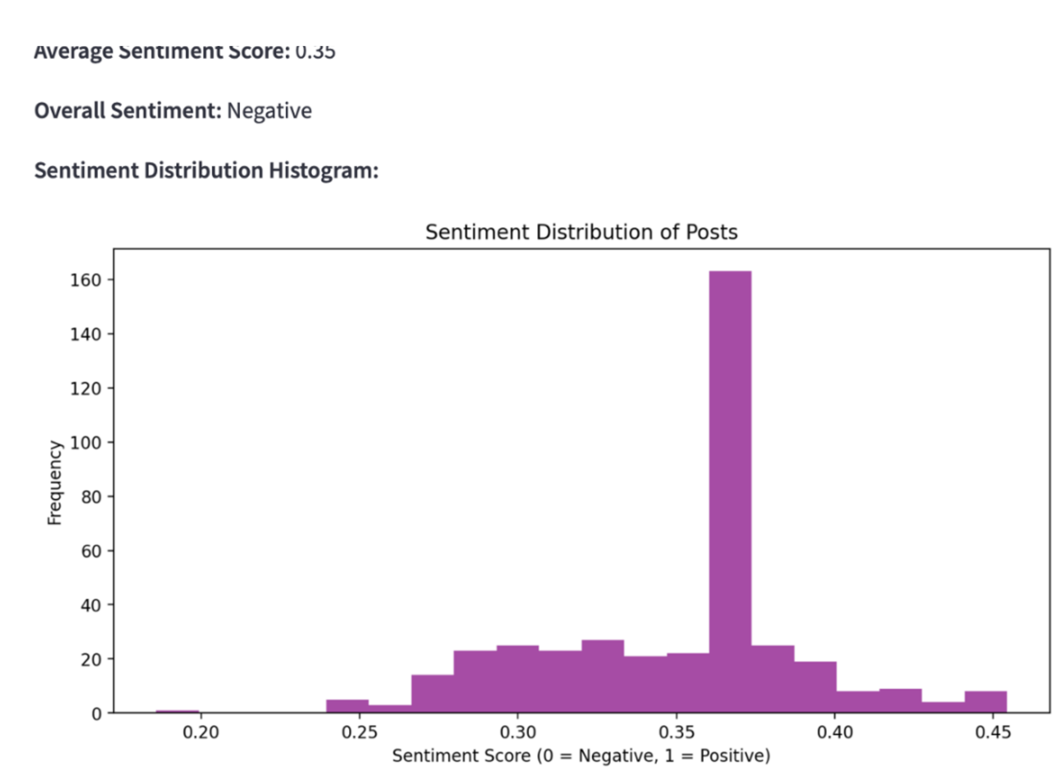

Figure 7 Sentiment Score Distribution of Non-Rumor Post

We can see that in Figure 7:

- Sentiment centers at 0.36, labeled "negative"
- Paradox: debunking posts skew negative because of refutation language and sensitive topics
- Analysis used a BERT-based model pretrained on general corpora
- BERT beats lexicon tools in semantics but misses Weibo-specific nuances (colloquial negations, sarcasm)
- Neutral/educational debate phrases may inherit unintended negativity due to lack of platform context

Bibliography

[1] Li M, Chen L, Zhao J, et al. Sentiment analysis of Chinese stock reviews based on BERT model[J]. Applied Intelligence, 2021, 51: 5016-5024.
[2] Zhang T, Xu B, Thung F, et al. Sentiment analysis for software engineering: How far can pre-trained transformer models go?[C]//2020 IEEE International Conference on Software Maintenance and Evolution (ICSME). IEEE, 2020: 70-80.
[3] Gu X, Liu L, Yu H, et al. On the transformer growth for progressive bert training[J]. arXiv preprint arXiv:2010.12562, 2020.
[4] Won D Y, Lee Y J, Choi H J, et al. Contrastive representations pre-training for enhanced discharge summary BERT[C]//2021 IEEE 9th International Conference on Healthcare Informatics (ICHI). IEEE, 2021: 507-508.

Acknowledgement

The dataset was constructed via the Weibo API covering November 2019 – March 2022 and comprises 4,184 source posts alongside 961,975 microblogs (retweets and comments) with complete diffusion cascades and user profiles. True news (1,407 items) was confirmed through official accounts, the Weibo Community Management Center, and the China Internet Joint Rumor Debunking Platform, while 768 COVID-19–related rumors were identified via the Community Management Center's reporting service. All microblog content was encoded using TF-IDF vectorization on the top 5,000 words.

More about this project: https://dsprojects.lib.cuhk.edu.hk/en/projects/weibo-rumors/introduction/
If you would like to reuse the data, please contact **data@cuhk.edu.hk**.