SentiFM: Epidemic Forecasting with Text Sentiment Analyses using Foundation Models

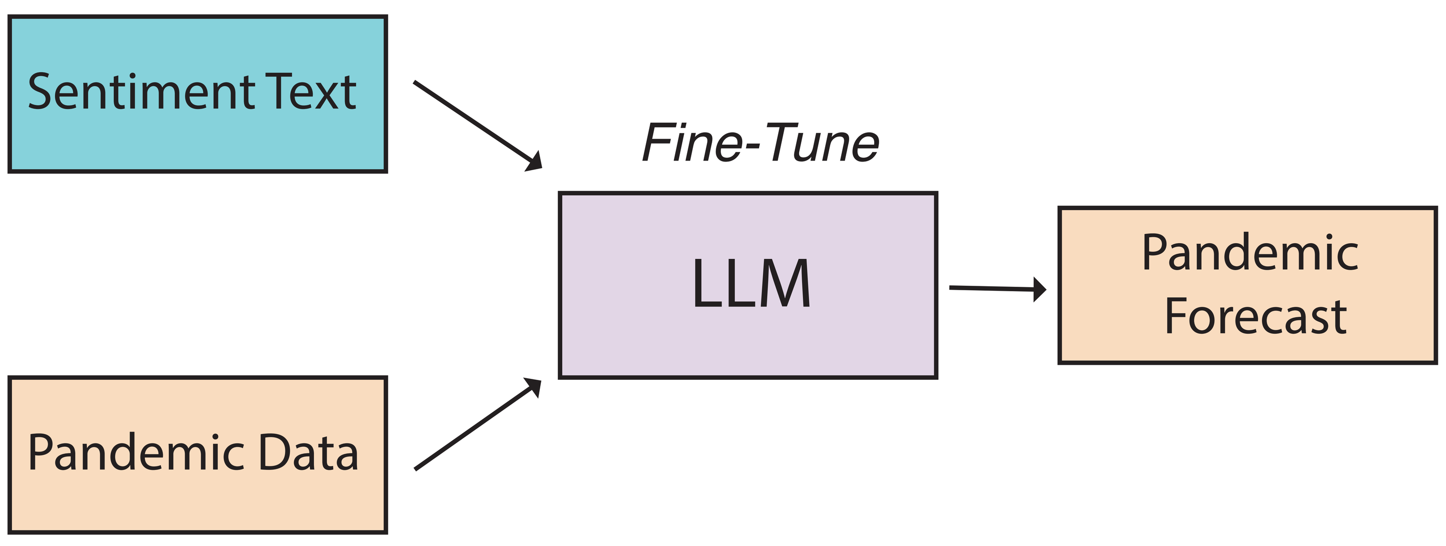
Junyang He Huanzhong Jia Yule Wang
Georgia Tech

Introduction

— Traditional epidemic forecasting models (e.g., SIR/SEIR, time-series regressions) rely on lagging indicators such as confirmed cases and hospitalizations. These indicators reflect disease spread only after significant delays.

— However, disease transmission is deeply behavior-driven. Public compliance with interventions is shaped by collective sentiments like fear, optimism, or fatigue.

— SentiLLM proposes a sentiment-aware forecasting framework that leverages Foundation Models (LLMs) to extract real-time public sentiment signals from Twitter, integrating them with COVID-19 epidemiological data to enhance prediction accuracy.



Method

— We evaluate multiple strategies to incorporate sentiment into epidemic forecasting:

1. **Naive Concatenation:** Aggregate sentiment features (e.g., emotion intensity, sentiment ratio, tweet volume) are concatenated with historical COVID data and fed into LSTM.

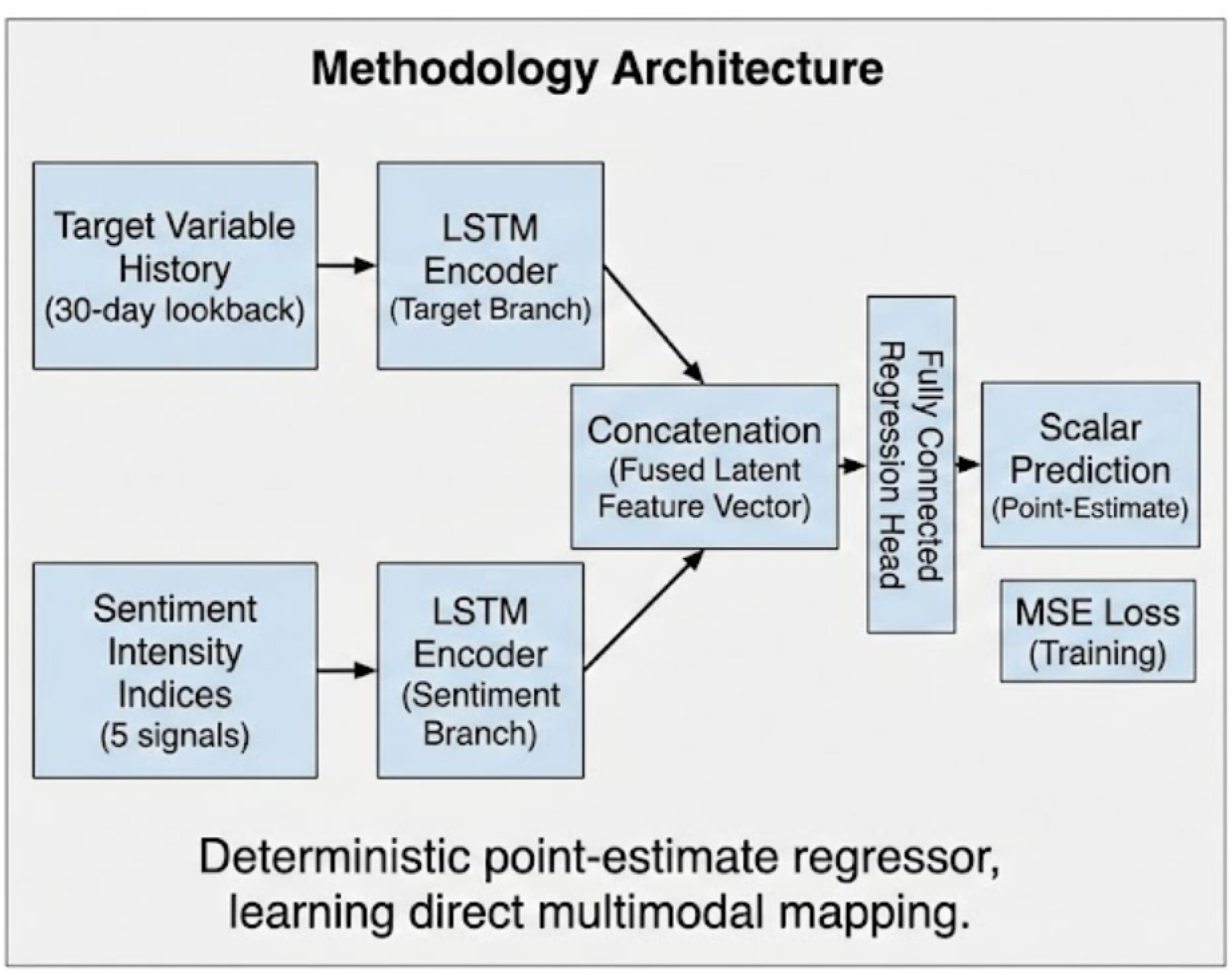
2. **In-Context Learning (Chronos-2):** Chronos-2 accepts sentiment time series as covariates. Group Attention dynamically learns correlations between sentiment and COVID-19 time-series data trends.

3. **Multimodal Fusion:** A dual-branch architecture processes epidemiological and sentiment streams separately via LSTM encoders, then fuses them via an MLP for prediction.

— Model Training Details

— Accounts for delayed public response. Sentiment time series is shifted forward by 4 days before being used as input to reflect behavioral latency.

— Datasets:
A. Epidemiological: WHO COVID-19 dataset (cases, deaths, testing, vaccination)
B. Sentiment: COVID-19 Twitter dataset with emotion labels (fear, happiness, sadness, etc.)
C. Timeframe: Dec 31, 2020 – Sep 1, 2021 (USA)



Experiment Results

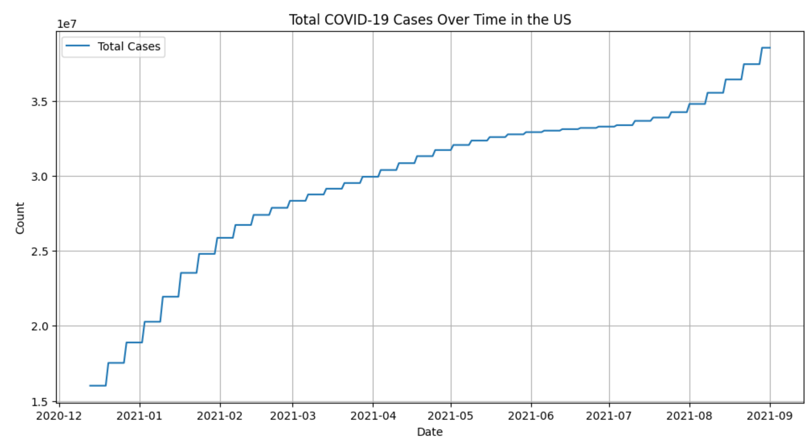
COVID feature	Without sentiment	With sentiment
People vaccinated	0.0254	0.0211
Total tests	0.3379	0.3218
Total cases	0.2930	0.2919
Total deaths	0.489	0.151

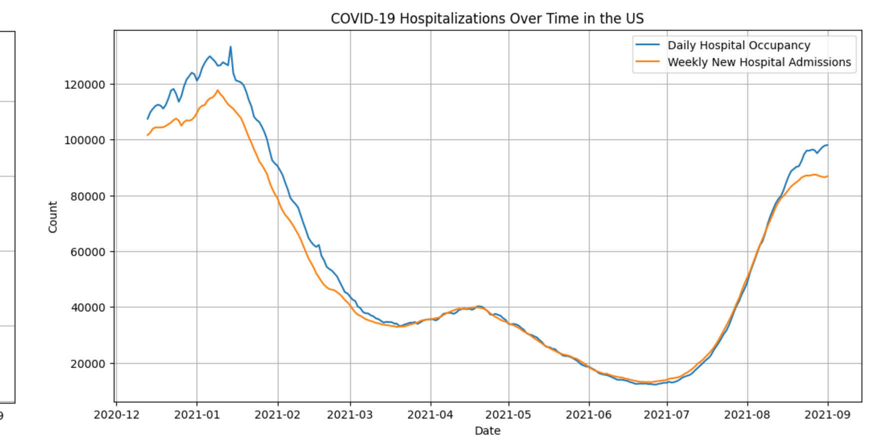
Effect of Naively Concat

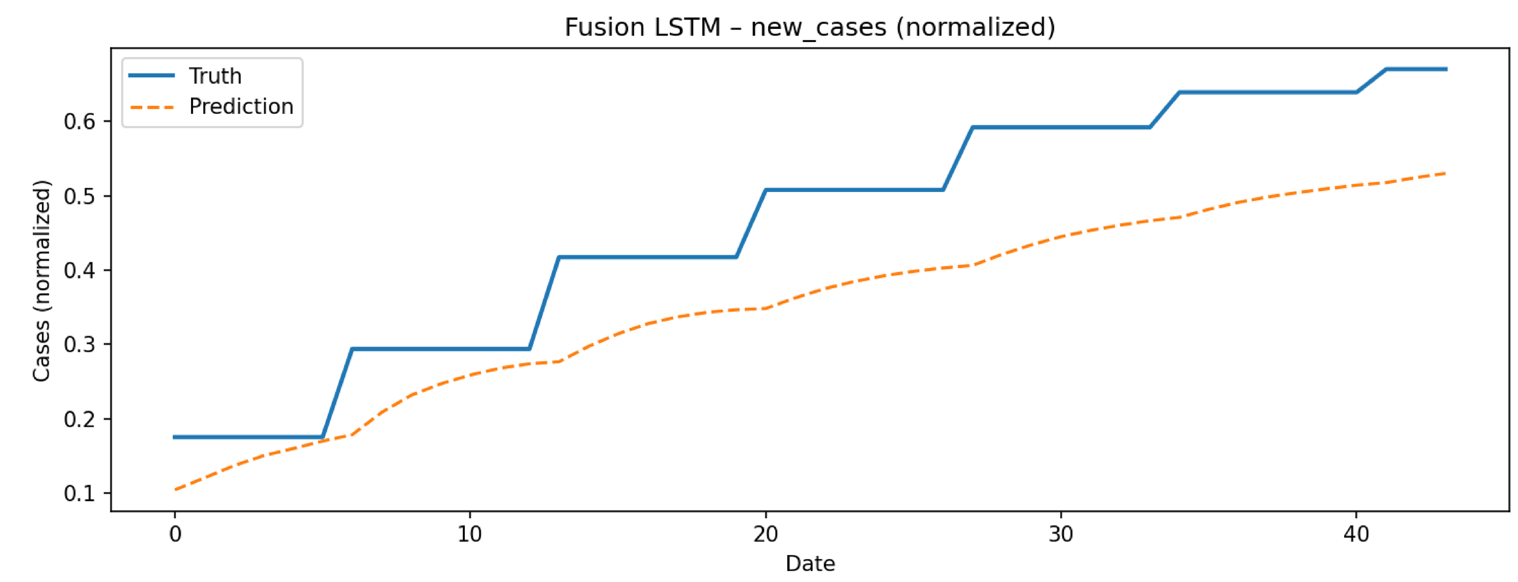
COVID feature	Without sentiment	With sentiment
People vaccinated	0.00421	0.00240
Total tests	0.00898	0.00628
Total cases	0.0677	0.0702
Total deaths	0.0286	0.0297

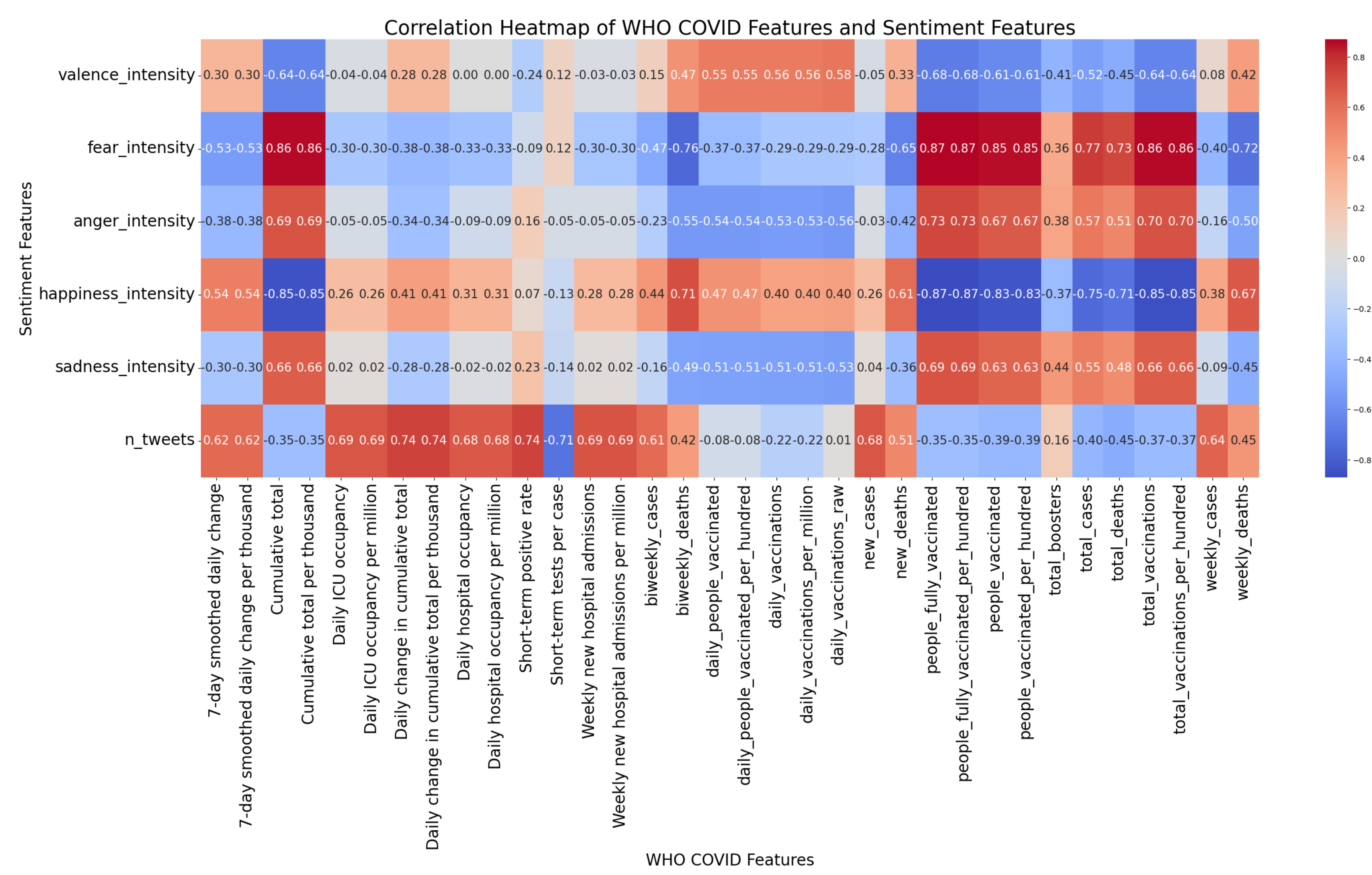
Effect of Naively Concat

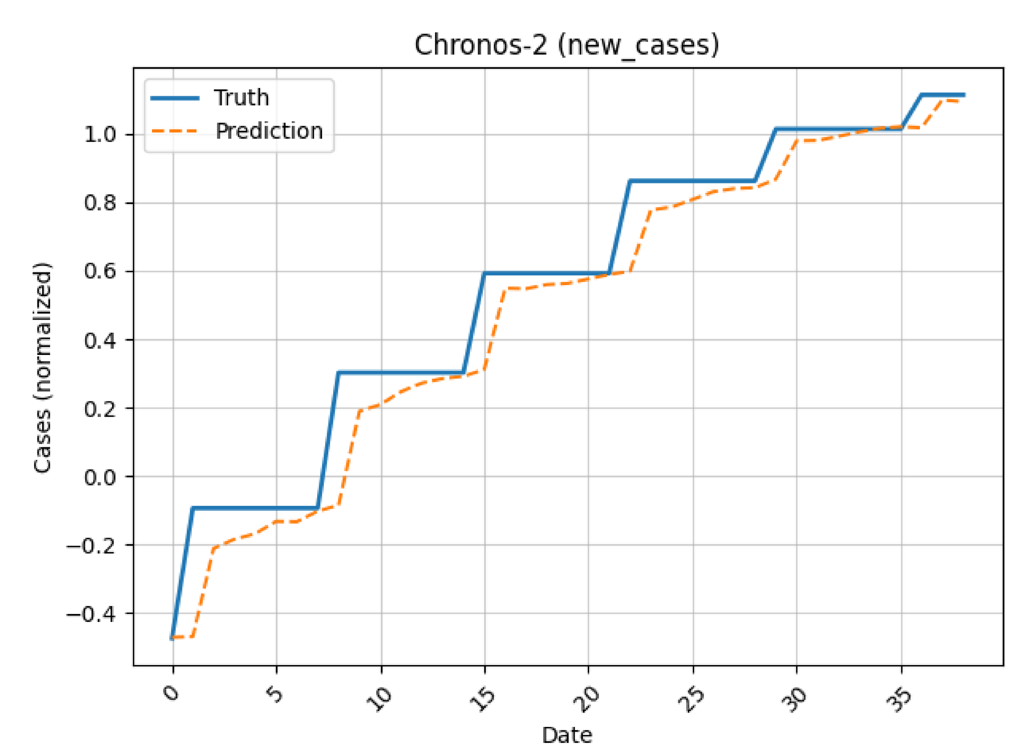
Experiment Results











Chronos-2 total cases (ICL with sentiment data).

Model Architecture	New Cases	New Deaths	Vaccinations
Classical Baselines			
SARIMAX (Univariate)	0.1795	0.0508	0.0031
SARIMAX (Raw Sentiment)	0.1673	0.0492	**0.0003**
SARIMAX (Senti-Shift $\tau = 14$)	0.1583	0.0445	0.0004
Generative AI			
Llama-3 (Zero-Shot Base)	0.0754	0.0435	0.0018
Llama-3 (QLoRA Fine-Tuned)	0.0713	0.0410	0.0015
Proposed Method			
LSTM (Multimodal Fusion)	**0.0067**	**0.0125**	**0.0003**

Cross-Architecture Performance Evaluation (Scaled MSE)

COVID feature	Without sentiment	+ Valence intensity
People vaccinated	0.00421	0.00216
Total tests	0.00898	0.00457
Total cases	0.0677	0.0681
Total deaths	0.0286	0.0288

+ Fear intensity	+ Anger intensity	+ Happiness intensity	+ Sadness intensity
0.00210	0.00213	0.00206	0.00209
0.00458	0.00506	0.00418	0.00467
0.0690	0.0693	0.0681	0.0691
0.0293	0.0293	0.0288	0.0294

Effect of each sentiment intensity feature on epidemiological forecast with Chronos-2

Conclusion

— Adding sentiment improves accuracy, especially for vaccination and testing forecasts.
— Foundation models show potential with well-structured prompts and QLoRA tuning, but face challenges in precise numerical prediction.