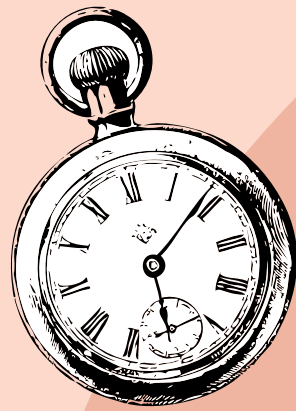
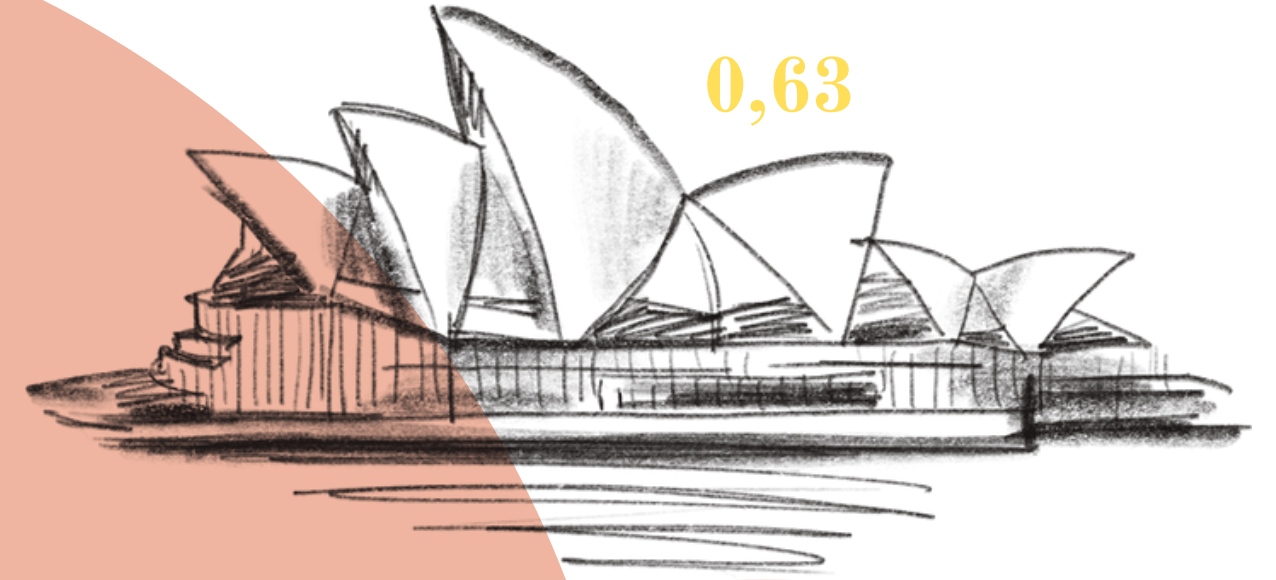


0,45



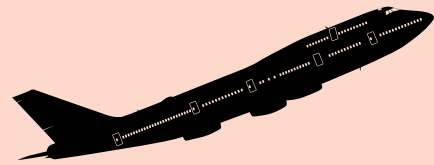
0,63



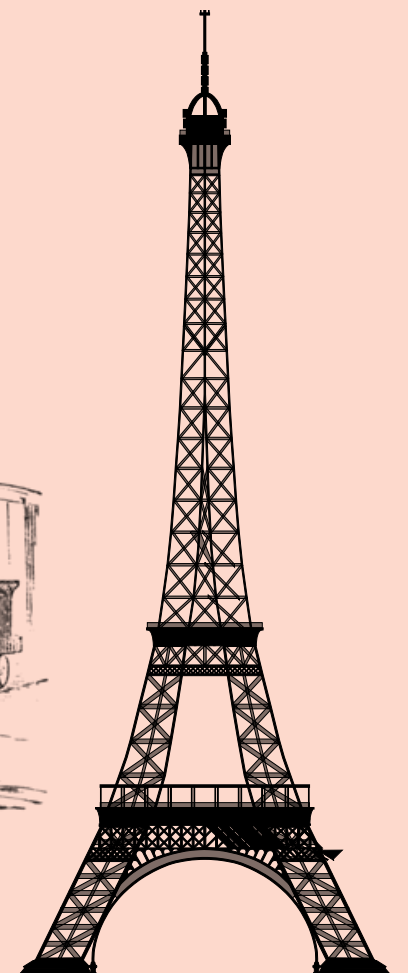
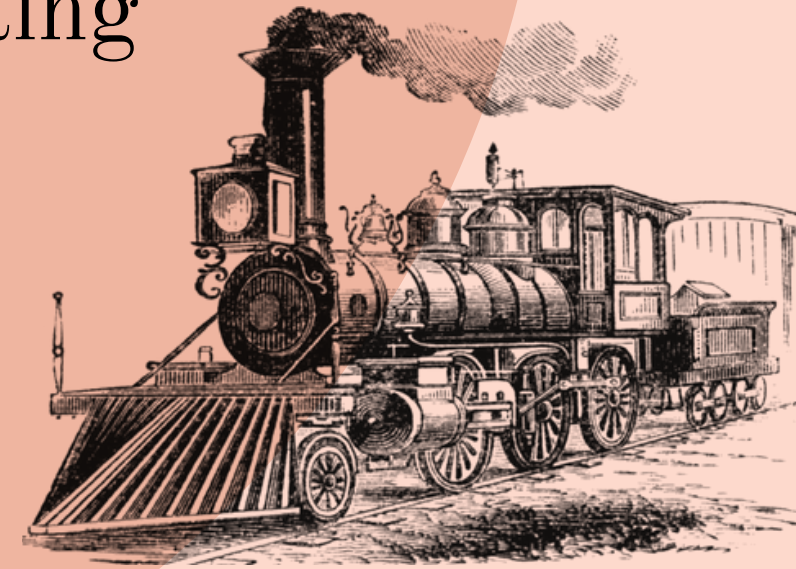
Urban Mood Forecaster

Time Series Analysis and
Forecasting of Global Sentiment
DS 207, Time Series Forecasting

0,78



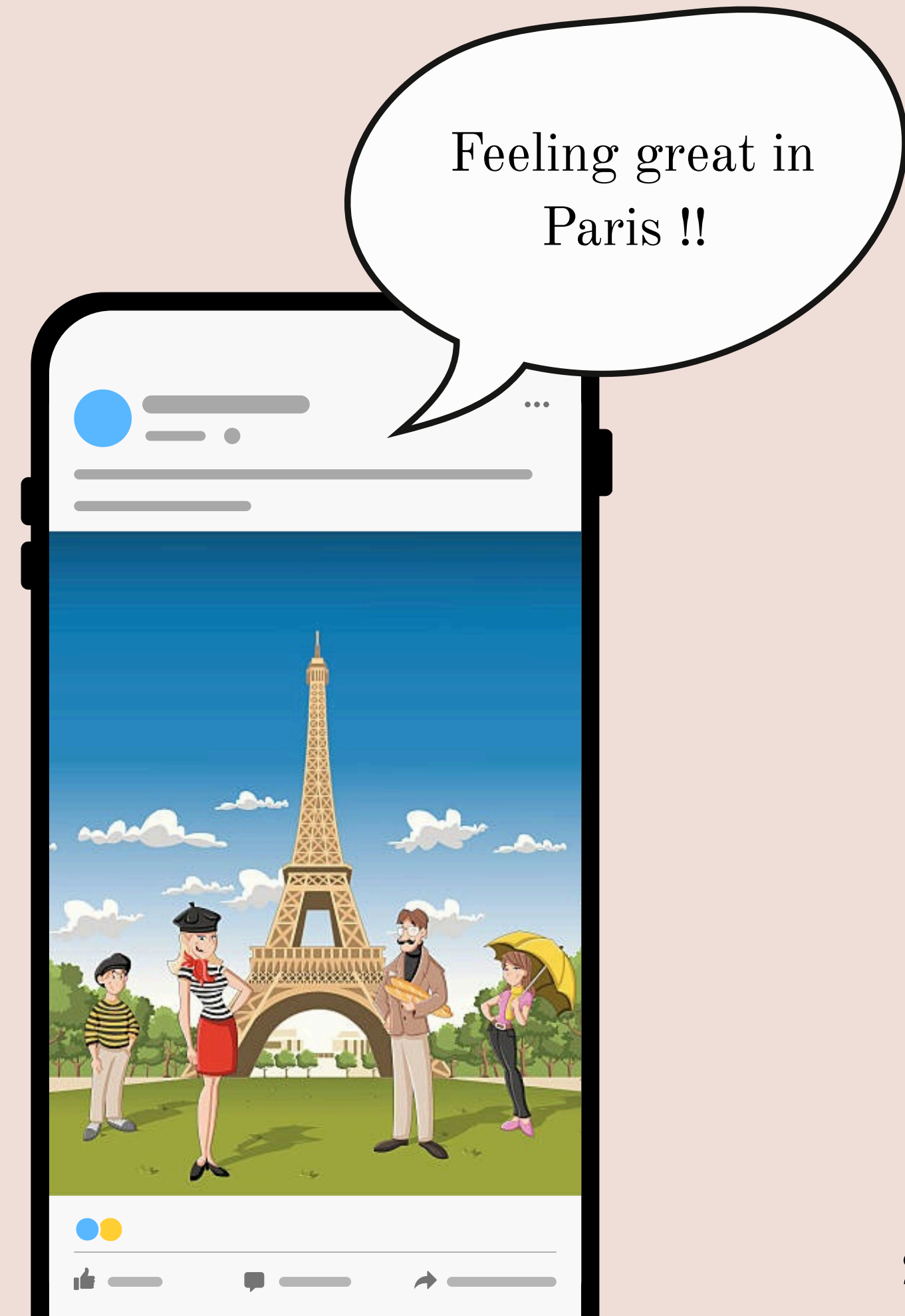
0,75



- Team Members:
- Shushan Meyroyan (AUA ID:), DS
- Shushan Gevorgyan (AUA ID: A09220116), DS
- Armen Madoyan (AUA ID:), DS
- Davit Sargsyan (AUA ID:), DS

Motivation: Why do we care about mood?

- UN Sustainable Development Goals include subjective well-being.
- Traditional indicators like GDP and unemployment miss how people actually feel.
- Social media (Twitter/X) provides high-frequency, geo-tagged signals of public mood.
- Question: Can we turn these signals into a systematic “urban mood” monitor and forecaster?

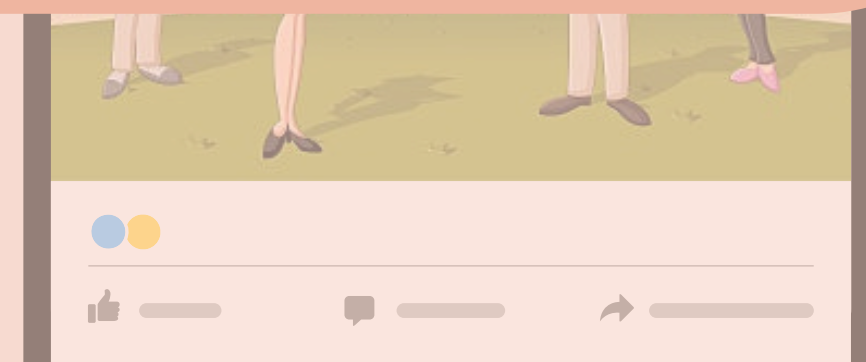


Motivation: Why do we care about mood?

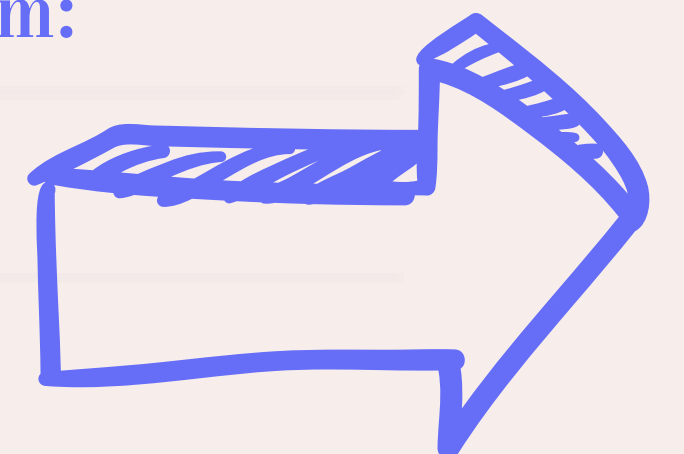
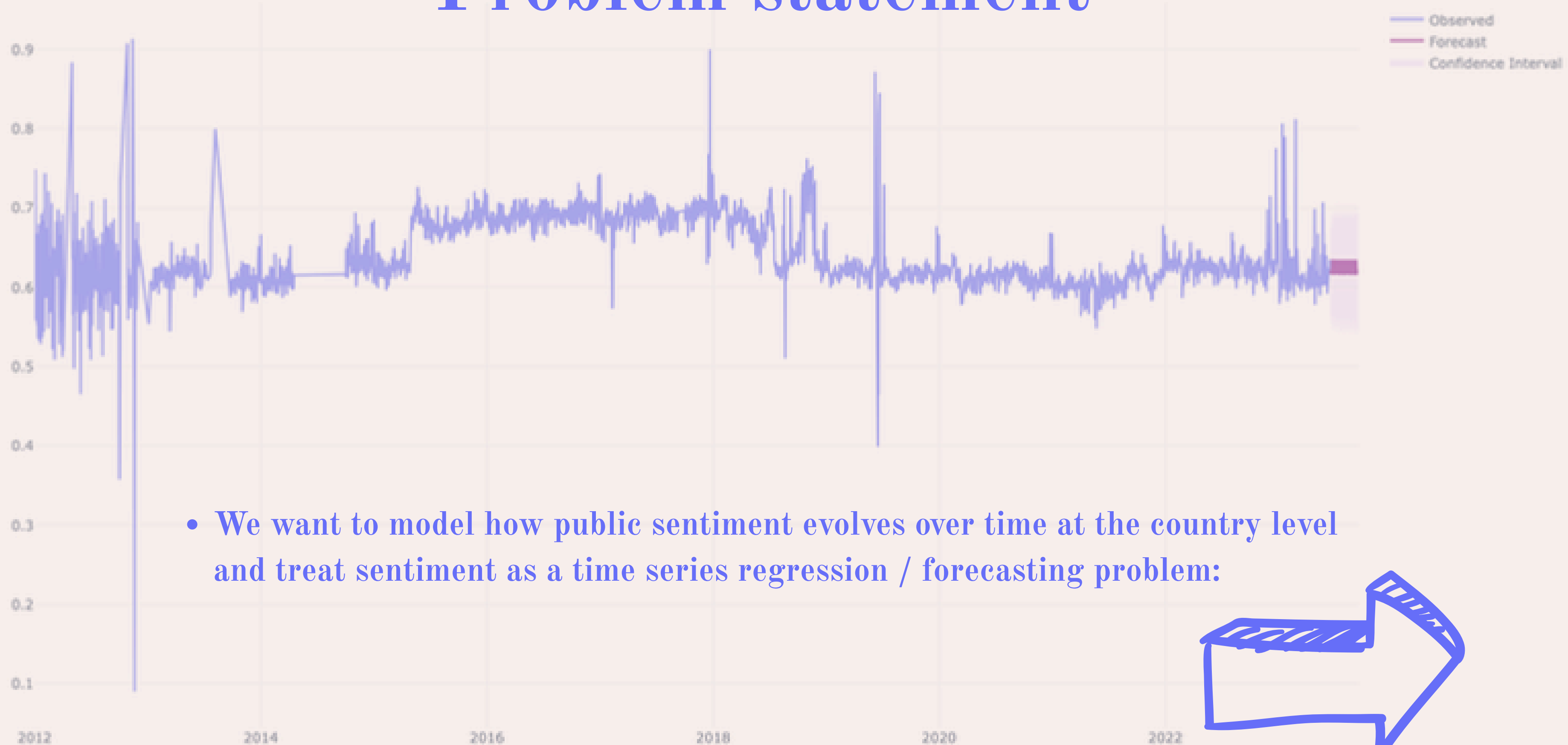
- UN Sustainable Development Goals: economic growth, subjective well-being
- Traditional indicators: GDP, unemployment rate
- Social media: sentiment analysis, frequency, geographic distribution
- Question: Can we predict economic growth with a systematic “mood” forecaster?

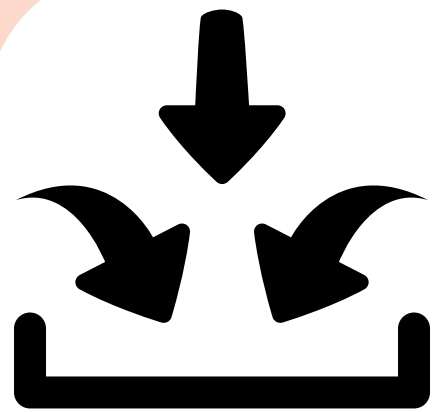


There is a gap between economic metrics and emotional reality



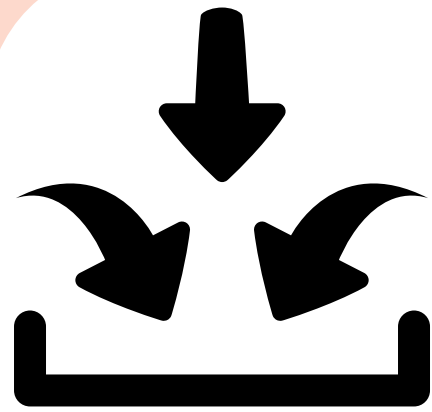
Problem statement





past daily sentiment scores...

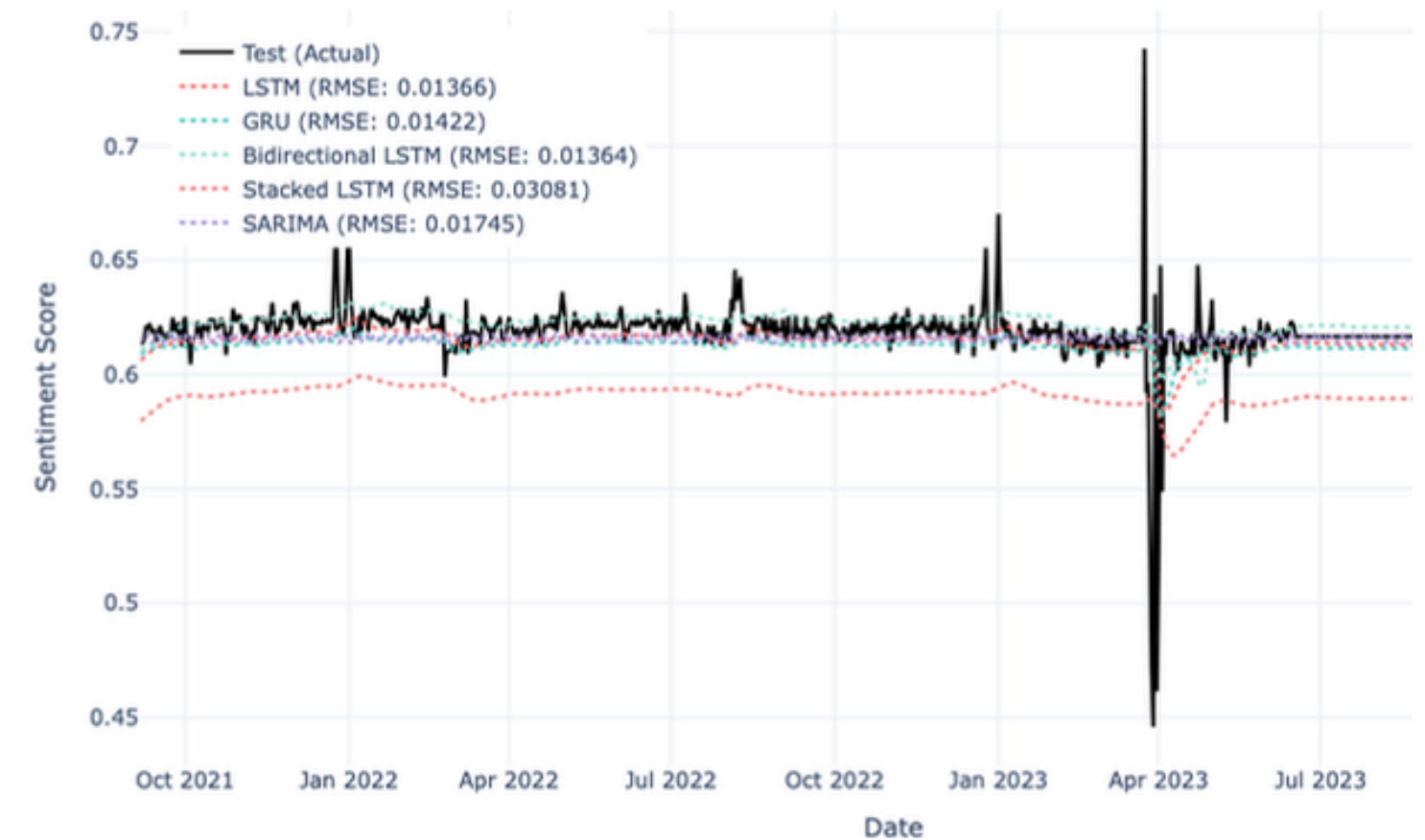
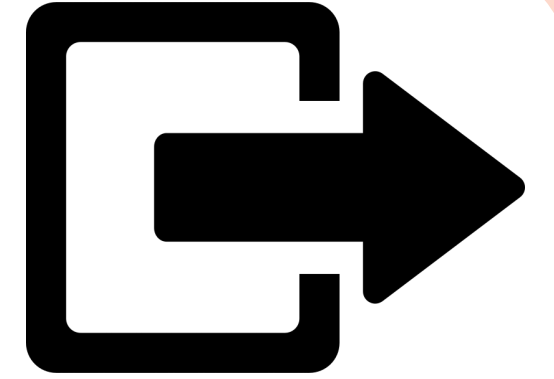
DATE	country	N	SCORE	year	
196	03/01/...	Afghanista	2	0.388336	2012
302	04/01/...	Afghanista	1	0.719559	2012
1266	13/01/...	Afghanista	1	0.613117	2012
1371	14/01/...	Afghanista	2	0.53986	2012
1588	16/01/...	Afghanista	2	0.481106	2012
...
412047	02/06/...	Zimbabwe	3338	0.618421	2023
412196	03/06/...	Zimbabwe	503	0.662444	2023
412346	04/06/...	Zimbabwe	252	0.629384	2023
412494	06/06/...	Zimbabwe	293	0.602974	2023
412642	07/06/...	Zimbabwe	195	0.61463	2023



past daily sentiment scores...

	DATE	country	N	SCORE	year
196	03/01/...	Afghanista	2	0.388336	2012
302	04/01/...	Afghanista	1	0.719559	2012
1266	13/01/...	Afghanista	1	0.613117	2012
1371	14/01/...	Afghanista	2	0.53986	2012
1588	16/01/...	Afghanista	2	0.481106	2012
...
412047	02/06/...	Zimbabwe	3338	0.618421	2023
412196	03/06/...	Zimbabwe	503	0.662444	2023
412346	04/06/...	Zimbabwe	252	0.629384	2023
412494	06/06/...	Zimbabwe	293	0.602974	2023
412642	07/06/...	Zimbabwe	195	0.61463	2023

... future sentiment values



Dataset details: TSGI overview

- Dataset -> Twitter Sentiment Geographical Index (TSGI).
- Source: Sustainable Urbanization Lab (MIT) + Center for Geographic Analysis (Harvard).
- Domain: Social media–based subjective well-being.
- Temporal coverage: 2012–2023
- Granularity: global, country, state/province, county/city.
- We use: country-level daily indices.

index	date	country	N of posts that day	score	year
196	03/01/2012	Afghanistan	2	0.388336	2012
302	04/01/2012	Afghanistan	1	0.719559	2012
1266	13/01/2012	Afghanistan	1	0.613117	2012
1371	14/01/2012	Afghanistan	2	0.53986	2012
1588	16/01/2012	Afghanistan	2	0.481106	2012
...
412047	02/06/2023	Zimbabwe	3338	0.618421	2023
412196	03/06/2023	Zimbabwe	503	0.662444	2023
412346	04/06/2023	Zimbabwe	252	0.629384	2023
412494	06/06/2023	Zimbabwe	293	0.602974	2023
412642	07/06/2023	Zimbabwe	195	0.61463	2023

Data organization:
One CSV per year: TSGI_country_2012.csv,
In total: X years (say 2012–2023 → 12 files),
thousands of country–day records.

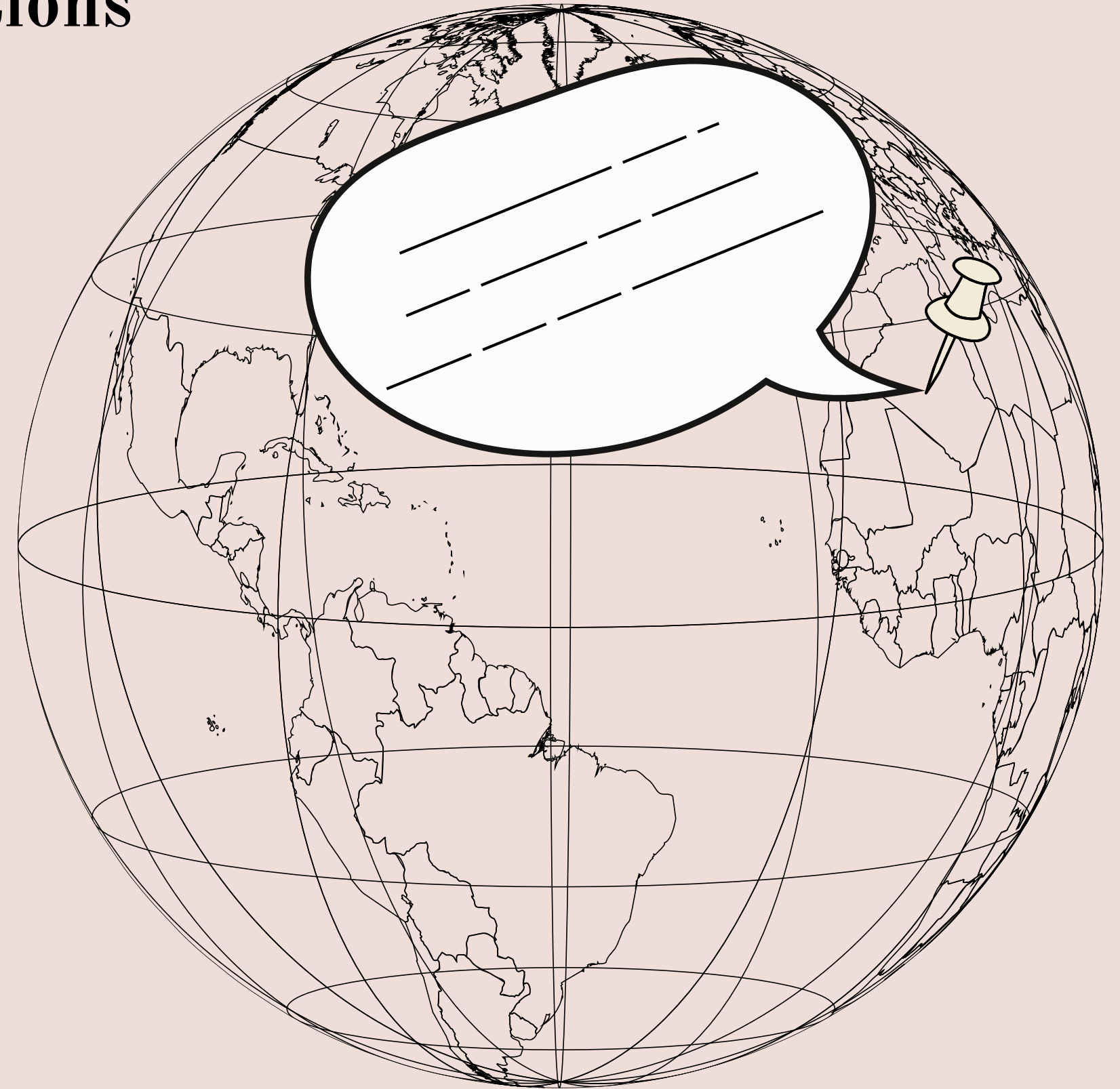
Research methods and important questions

RQ1: How well can classical time series models (AR, MA, ARIMA, SARIMA, Holt–Winters) forecast sentiment?

RQ2: Do deep RNN models (LSTM, GRU, Bidirectional LSTM, Stacked LSTM) improve forecasting accuracy?

RQ3: Which seasons are “happiest” for each country?

RQ4: Which countries are happiest in a given season, and can we use this for travel recommendations?



Methodology overview: model families and metrics

Classical Statistical
Models:
AR, MA, ARIMA,
and SARIMA

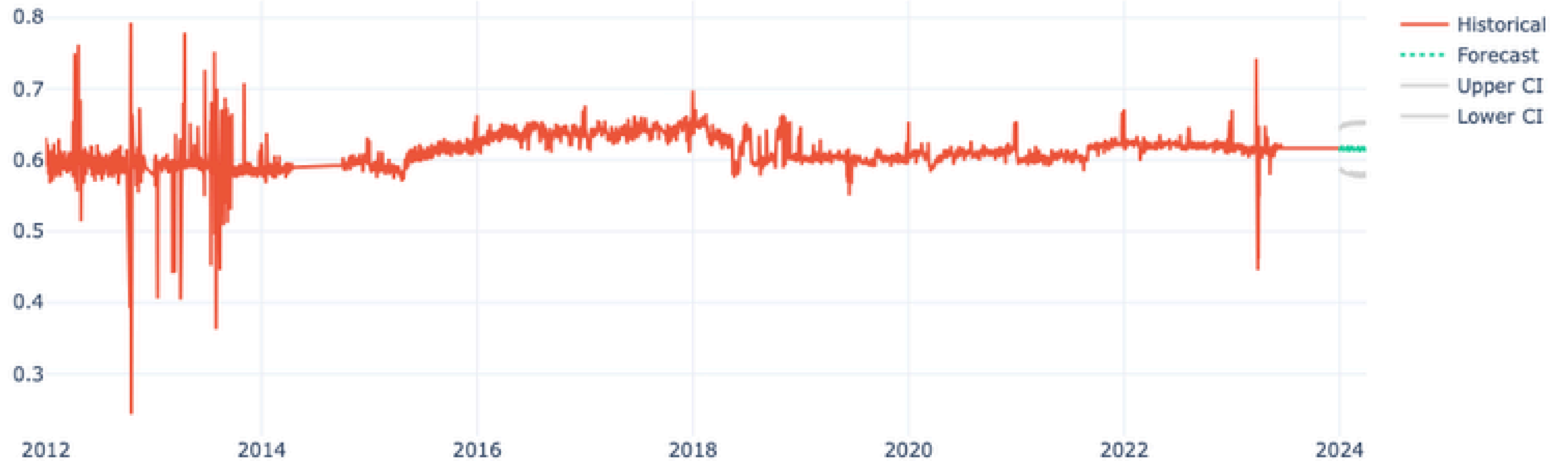
Baseline Models :
Naive forecasting &
rolling mean

Deep learning
approaches:
LSTM, GRU,
Bidirectional, and
Stacked LSTM.

*Evaluation metrics:
Main: RMSE on held-out test set.
Also: MAE, AIC/BIC for
ARIMA/SARIMA during selection.

Best SARIMA order: (1, 0, 1) seasonal: (0, 1, 1, 7)

SARIMA Forecast for Global Sentiment (Next 90 Days)



Baseline Model Forecast Comparison



Baseline Model RMSEs:

Naïve → RMSE: **0.01815**

Mean → RMSE: **0.01603**

Rolling(7) → RMSE: **0.01988**

Holt-Winters → RMSE: **0.01977**

Deep Learning models:

LSTM

(Long Short-Term Memory)

- What it does

Designed to learn long-term dependencies in sequential data.

Uses three gates (input, forget, output) and a cell state for controlled memory.

- Strengths

Excellent at capturing temporal patterns, even when relationships span many timesteps.

Reduces vanishing gradient problems common in vanilla RNNs.

- Weaknesses

More parameters → slower to train than GRU.

More memory-heavy.

Deep Learning models:

LSTM

(Long Short-Term Memory)

- What it does

Designed to learn long-term dependencies in sequential data. Uses three gates (input, forget, output) and a cell state for controlled memory.

- Strengths

Excellent at capturing temporal patterns, even when relationships span many timesteps.

Reduces vanishing gradient problems common in vanilla RNNs.

- Weaknesses

More parameters → slower to train than GRU.

More memory-heavy.

GRU

(Gated Recurrent Unit)

- What it does

A simplified version of LSTM with two gates (update, reset).

No separate cell state → only uses the hidden state.

- Strengths

Faster training, fewer parameters.

Often performs similarly to LSTM on many tasks.

Works well with smaller datasets.

- Weaknesses

Slightly less expressive than LSTM because of fewer gates.

Deep Learning models:

LSTM

(Long Short-Term Memory)

- What it does

Designed to learn long-term dependencies in sequential data. Uses three gates (input, forget, output) and a cell state for controlled memory.

- Strengths

Excellent at capturing temporal patterns, even when relationships span many timesteps.

Reduces vanishing gradient problems common in vanilla RNNs.

- Weaknesses

More parameters → slower to train than GRU.

More memory-heavy.

GRU

(Gated Recurrent Unit)

- What it does

A simplified version of LSTM with two gates (update, reset).

No separate cell state → only uses the hidden state.

- Strengths

Faster training, fewer parameters.

Often performs similarly to LSTM on many tasks.

Works well with smaller datasets.

- Weaknesses

Slightly less expressive than LSTM because of fewer gates.

Bidirectional LSTM

- What it does

Processes the sequence forward and backward. Concatenates both directional outputs.

- Why useful

Captures patterns where the future context within a window also matters.

More powerful than a standard one-direction LSTM for fixed-window forecasting.

- Strengths

Learns richer patterns (past + “future” inside the window).

Often provides the best accuracy in time series with local symmetric patterns.

- Weaknesses

More computationally expensive (two LSTMs running in parallel).

Cannot be used where true future values must be unknown during training (but works for our sliding-window setup).

Deep Learning models:

Stacked LSTM \\ (Multi-layer LSTM)

- What it does

Multiple LSTM layers stacked on top of each other.

Each layer learns increasingly abstract representations of the sequence.

- Strengths

Higher learning capacity for complex, nonlinear patterns.

In theory, more expressive than a single LSTM layer.

- Weaknesses

Very heavy: more parameters → more likely to overfit (which happened in your results).

Needs more data and more tuning.
Slowest to train.

GRU (Gated Recurrent Unit)

- What it does

A simplified version of LSTM with two gates (update, reset).

No separate cell state → only uses the hidden state.

- Strengths

Faster training, fewer parameters.

Often performs similarly to LSTM on many tasks.

Works well with smaller datasets.

- Weaknesses

Slightly less expressive than LSTM because of fewer gates.

Bidirectional LSTM

- What it does

Processes the sequence forward and backward.
Concatenates both directional outputs.

- Why useful

Captures patterns where the future context within a window also matters.

More powerful than a standard one-direction LSTM for fixed-window forecasting.

- Strengths

Learns richer patterns (past + “future” inside the window).

Often provides the best accuracy in time series with local symmetric patterns.

- Weaknesses

More computationally expensive (two LSTMs running in parallel).

Cannot be used where true future values must be unknown during training (but works for our sliding-window setup).

Data pipeline

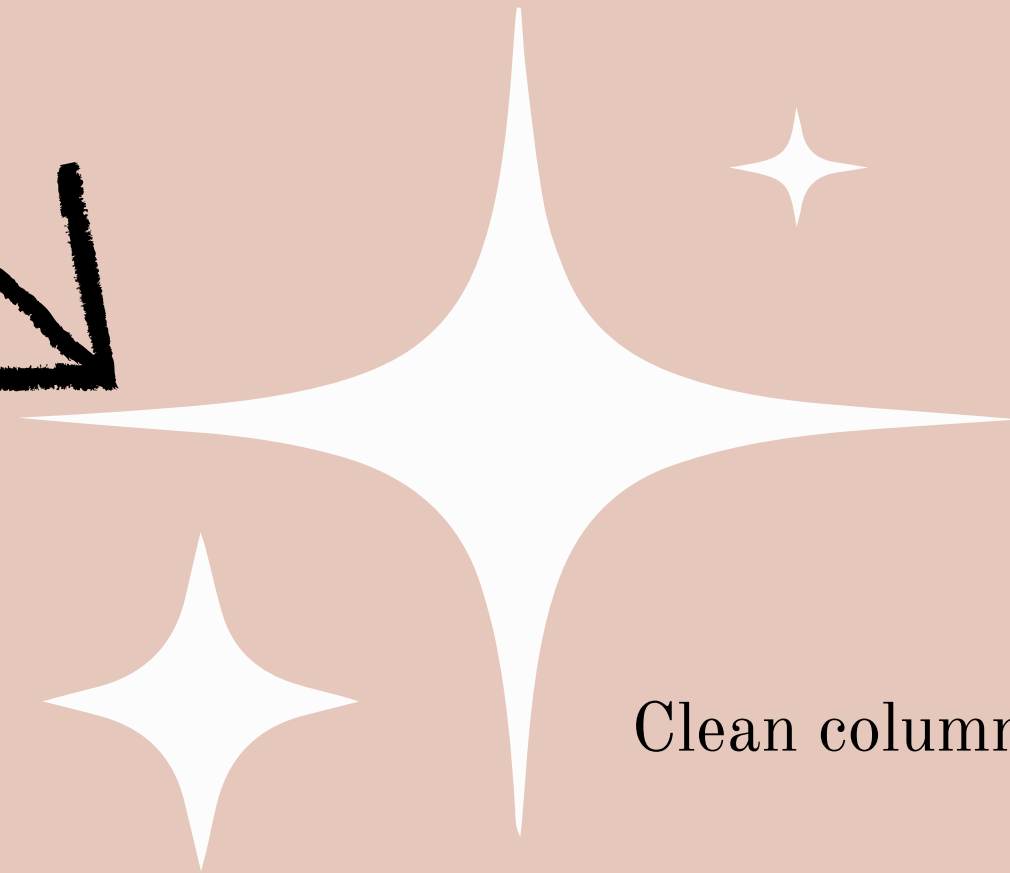
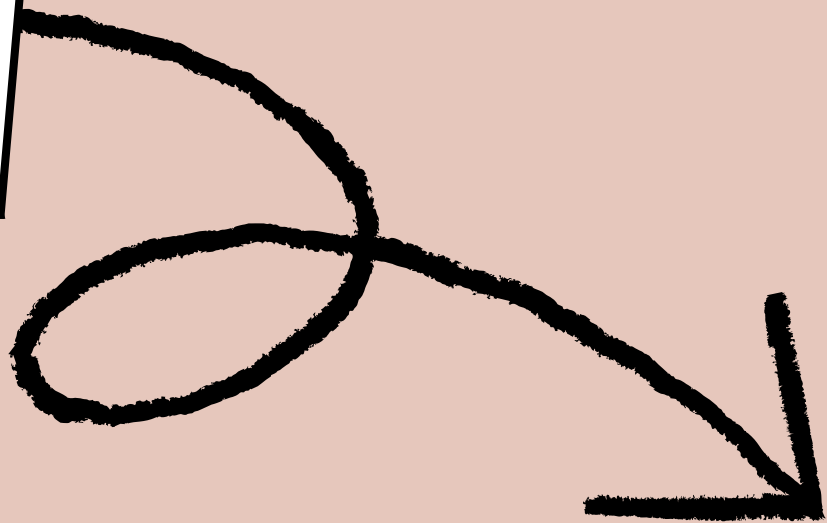


Combine yearly files

Data pipeline



Combine yearly files



Clean columns

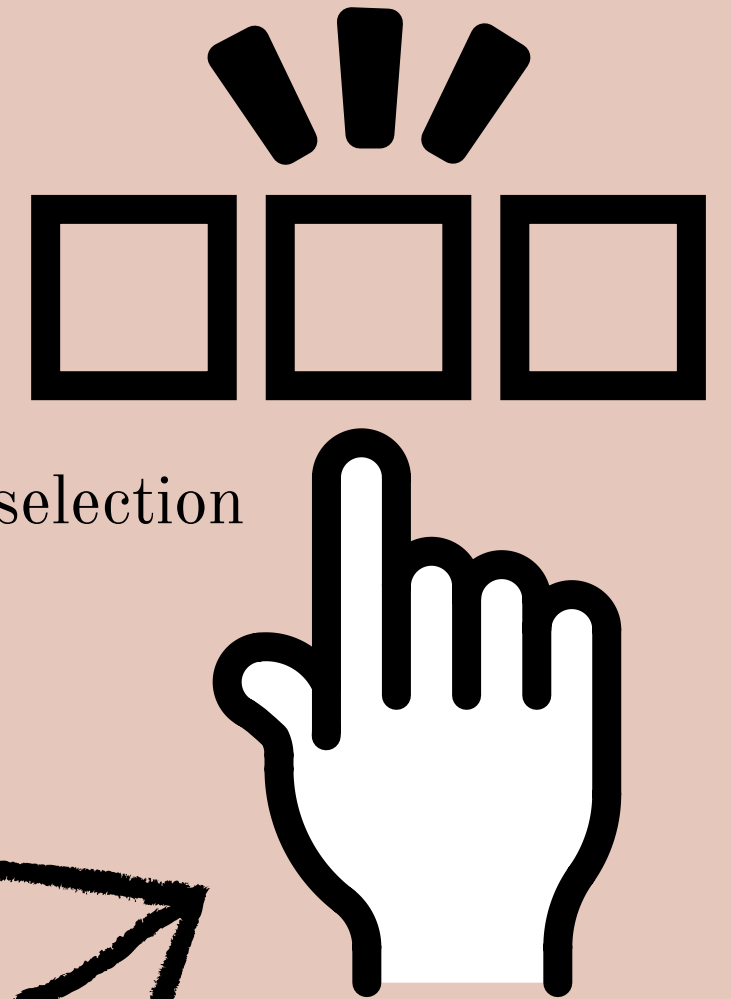
Data pipeline



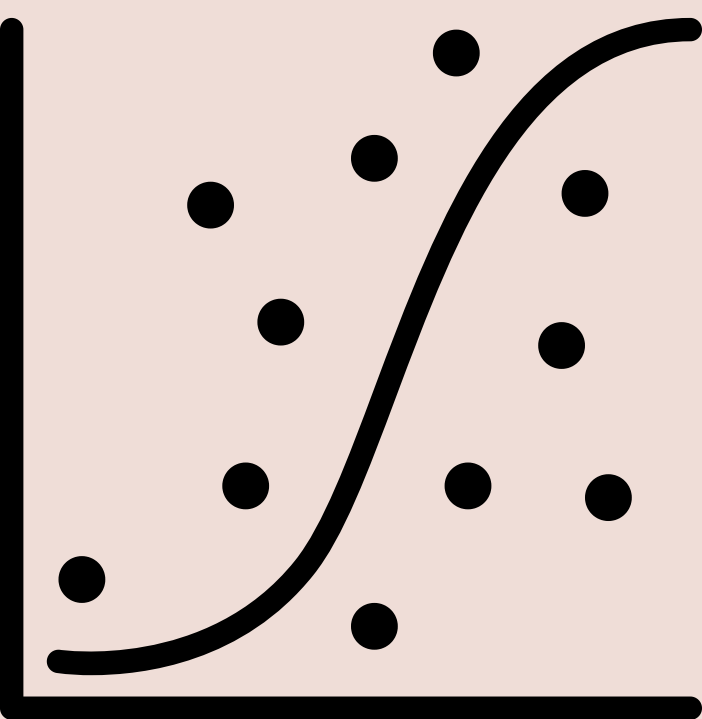
Combine yearly files



Country selection



Clean columns



Time series construction and missing values

Enforce daily frequency:

(`.asfreq('D')` to make the index every calendar day.)

Handle missing days/scores:

(Linear interpolation of SCORE between known dates.)

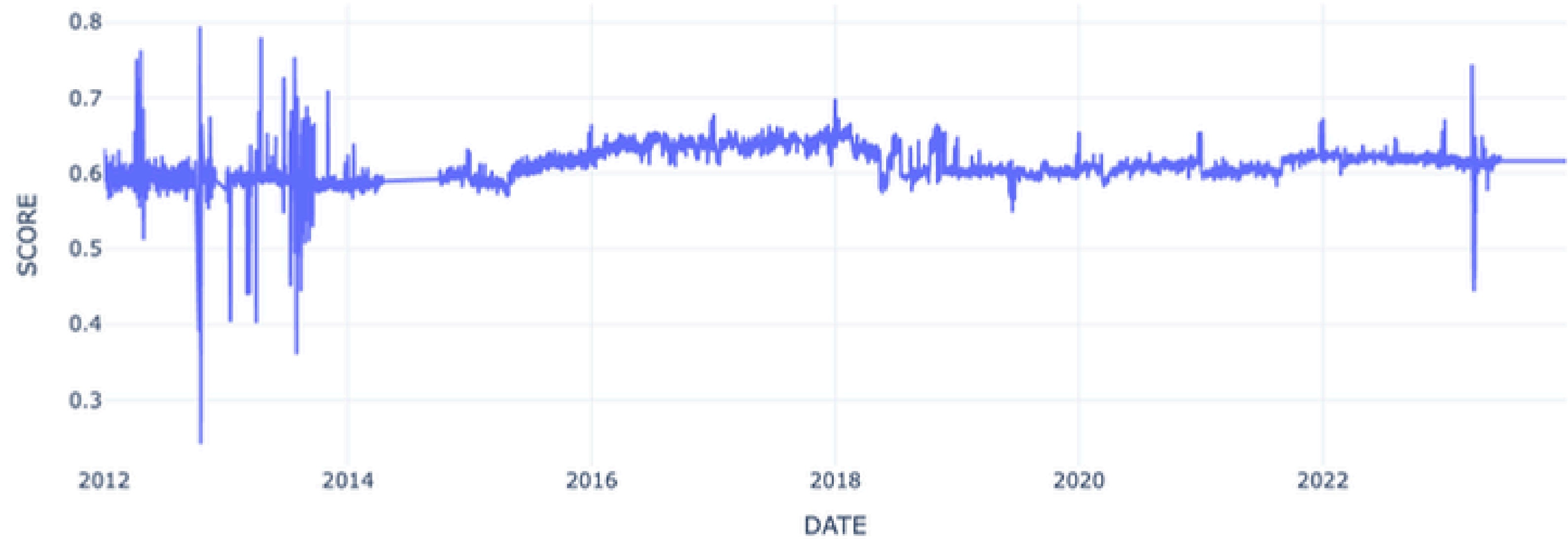
Train–test split:

(First 80% of days → train; last 20% → test (chronological split).)

For global experiments:

Average SCORE over all countries per day → global sentiment series.

Global Sentiment with Interpolated Missing Days



Global Sentiment with Interpolated Missing Days

==== Stationarity Tests on Global Sentiment (Daily) ====

ADF Test:

Test Statistic : -4.4823

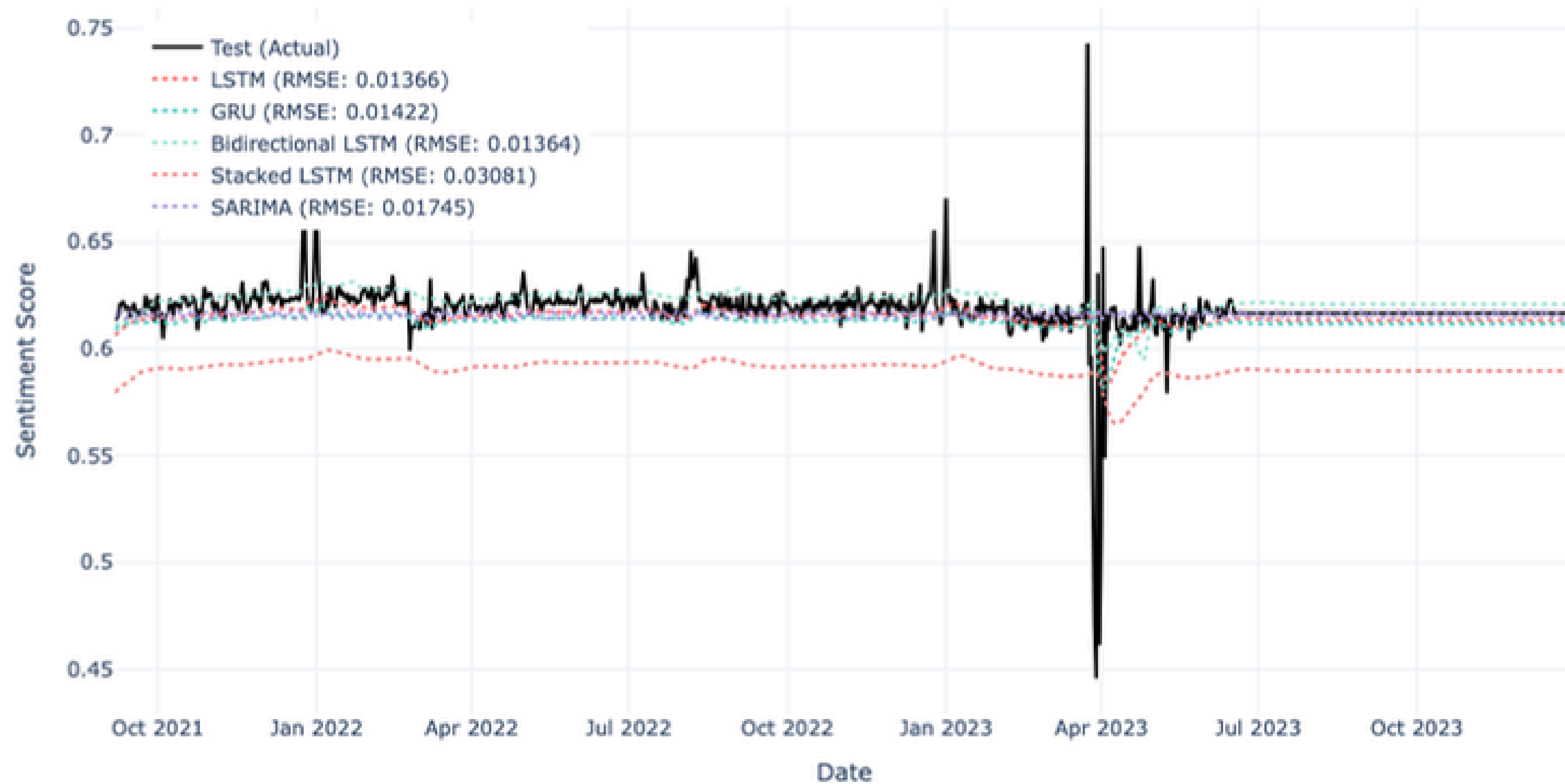
p-value : 0.0002

Critical Values: {'1%': -3.4318510679201006, '5%': -2.8622032882383506, '10%': -2.5671230728180445}

\Rightarrow Reject $H_0 \rightarrow$ Series is likely stationary.

KPSS Test: Test Statistic : 2.3451 p-value : 0.0100 Critical Values: {'10%': 0.347, '5%': 0.463, '2.5%': 0.574, '1%': 0.739} Reject $H_0 \rightarrow$ Series is likely non-stationary.

RNN Models vs SARIMA: Forecast Comparison

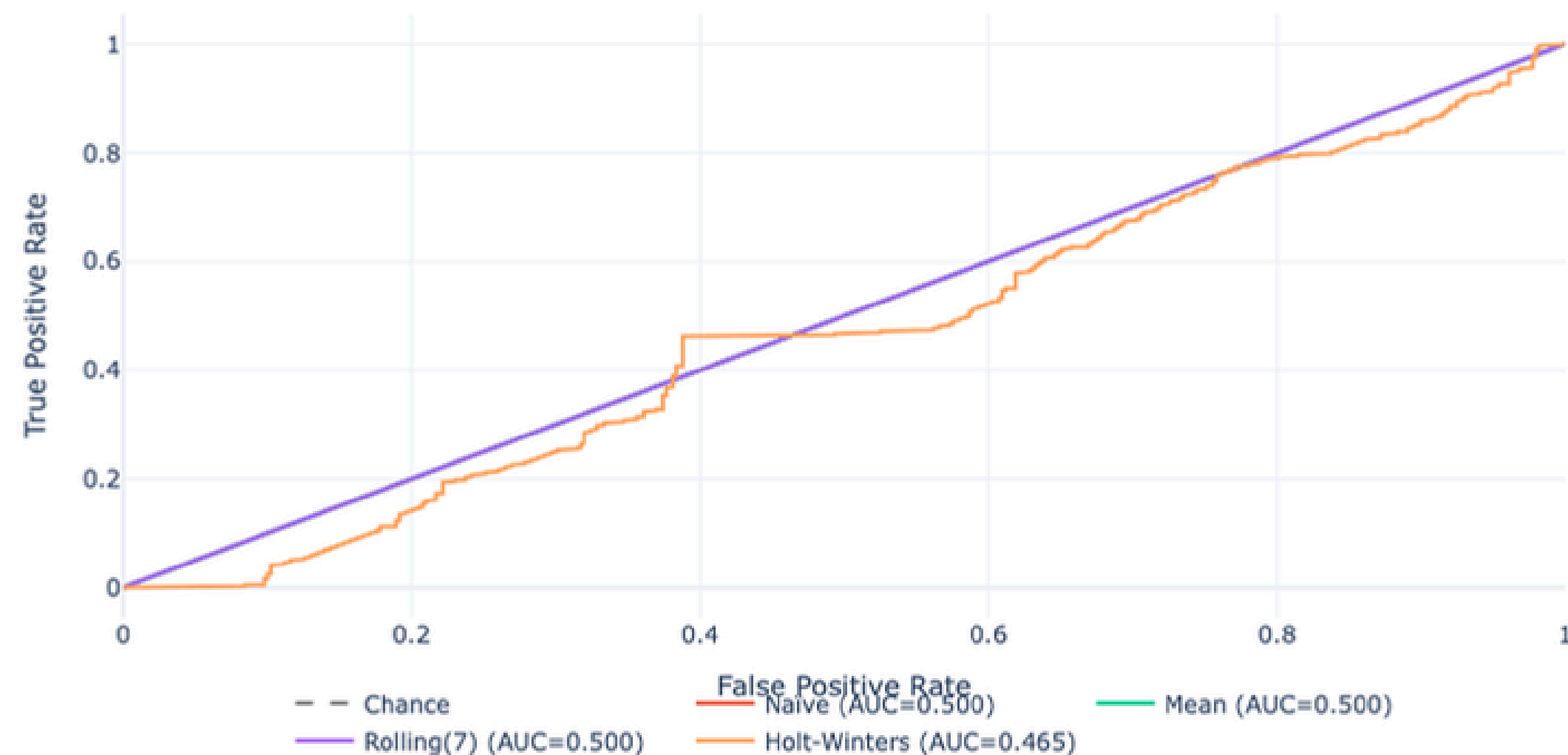


RNN Models vs SARIMA: Forecast Comparison



	Model	Accuracy	Precision	Recall	F1 Score	ROC AUC
0	Naïve	0.49943	0.0	0.0	0.0	0.500000
1	Mean	0.49943	0.0	0.0	0.0	0.500000
2	Rolling(7)	0.49943	0.0	0.0	0.0	0.500000
3	Holt-Winters	0.49943	0.0	0.0	0.0	0.464739

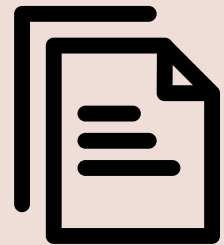
ROC Curves Across Models



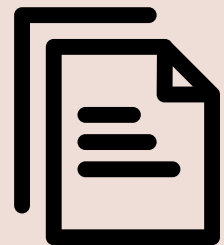
Project Structure – What We Did & Expected Outcomes



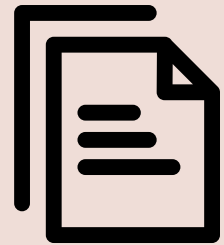
time_series_urban_mood_analysis



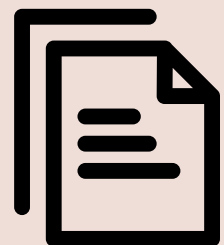
app.py → Urban Mood Forecaster Streamlit app



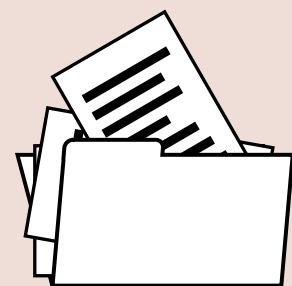
backend.py → Classical and deep learning forecasting models



user_input.py → Terminal-based application for forecasting



Urban_Mood_Analysis.ipynb → EDA & modeling experiments



dataverse_files → sentiment scores by country/state/country

Achieved Outcomes:

- Accurate daily sentiment forecasts for any country
- A clear comparison between classical models vs. RNNs
- Discovery of seasonal mood patterns across countries
- An interactive forecasting & travel advisory app
- A reusable forecasting backend with saved models
- Insights into how global sentiment behaves over time

Conclusion, Key Findings and Limitations of the work:

Urban Mood Forecaster demonstrates that global and country-level public sentiment can be modeled and forecasted reliably using a combination of classical time-series models and modern recurrent neural networks !

Our experiments show that:

- Bidirectional LSTM and LSTM achieved the lowest RMSE, capturing short-term fluctuations and non-linear dynamics better than classical models.
- SARIMA remains a strong, interpretable baseline, accurately modeling weekly seasonality and general trends.
- The system powers a Streamlit application that provides forecasts, confidence intervals, seasonal mood summaries, and travel recommendations in an accessible, interactive way.

Limitations :

- Sentiment comes only from geo-tagged tweets, not full populations.
- Models use just one feature (daily sentiment), limiting complexity.
- Seasonal analysis is simplified and may miss rare events.

References

- Chai, Y. (2022). Twitter Sentiment Geographical Index (Version V3) [dataset]. Harvard Dataverse. <https://doi.org/10.7910/DVN/3IL00Q>
- Kienmayer, M. (2022). Geospatial timeseries imputation using deep neural networks (Bachelor's thesis, Uppsala University). Uppsala University.
- Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and practice (2nd ed.). OTexts. <https://otexts.com/fpp2/>

Thank
you