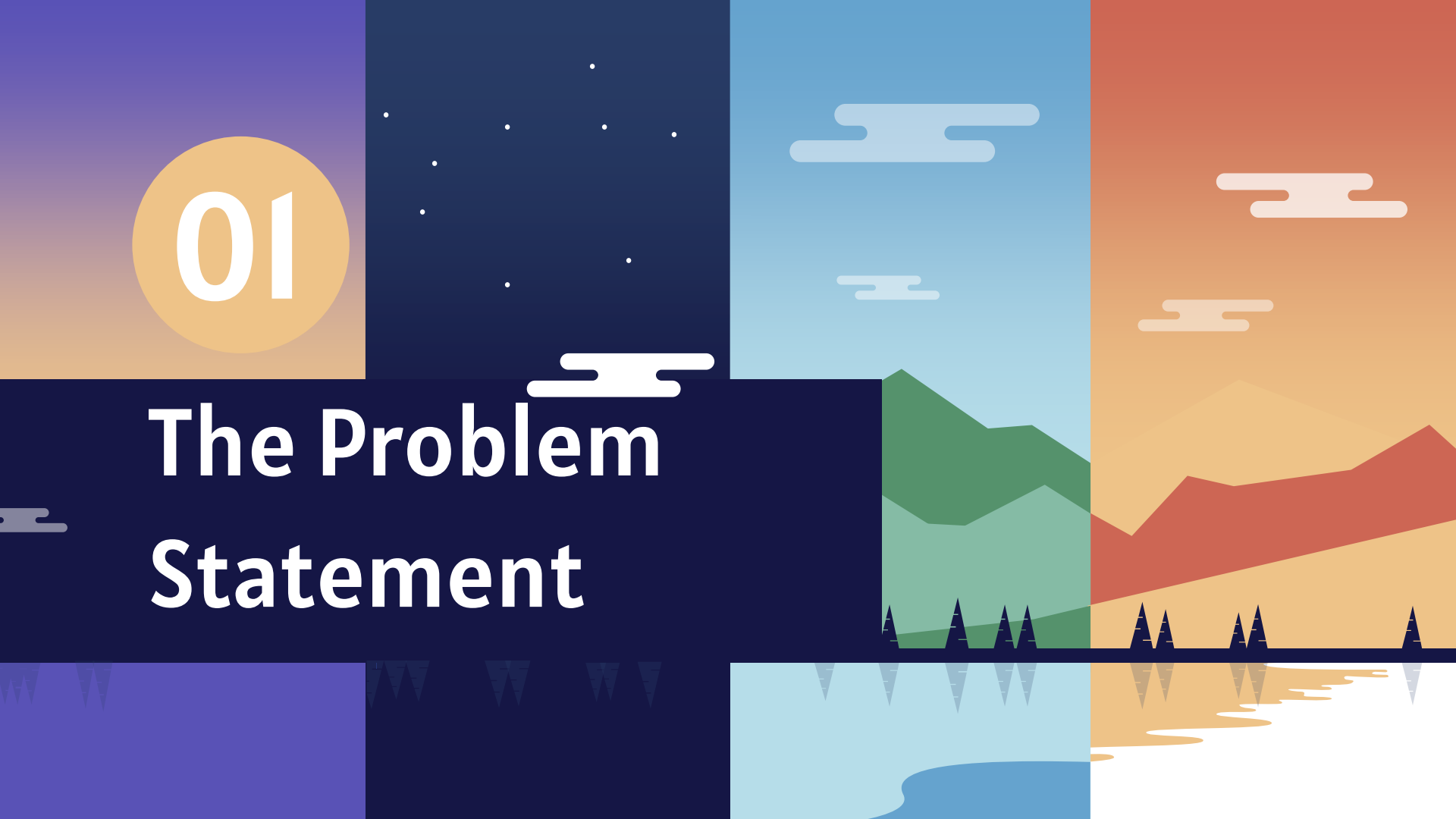
CAPSTONE SPRINT 3:

Unpredictable Vancouver Weather Predicted with Data Science.

Why is it so hot or cold without warning?

Table of contents


01
The Problem Statement

How can we better predict high and low temperatures?

Based solely on the weather statistics already available, can Machine Learning be used to predict a day with very hot or very cold weather?

Solution: Use LSTM to train a model on the weather statistics and predict the temperature.

Potential Impact:

- Help people be more prepared
- Help businesses sell seasonal products


02
Dataset and
Preprocessing

https://vancouver.weatherstats.ca/

Data Overview:

Hourly Vancouver Weather Statistics (July 1st, 2013 - June 30, 2023)

1. Filling null values
2. Feature Engineering for EDA
3. Feature Selection through EDA

date_time_local	pressure_station	pressure_sea	wind_dir	wind_speed	wind_gust	relative_humidity	dew_point	temperature	windchill	humidex	visibility
2013-07-01 00:00:00	101.18	101.16	SSE	7	0.0	91	18.2	19.7	0.0	0.0	32200.0
2013-07-01 01:00:00	101.22	101.21	SE	6	0.0	89	17.8	19.6	0.0	0.0	32200.0
2013-07-01 02:00:00	101.26	101.24	E	11	0.0	88	16.7	18.7	0.0	0.0	32200.0
2013-07-01 03:00:00	101.26	101.25	E	4	0.0	84	16.5	19.2	0.0	0.0	32200.0
2013-07-01 04:00:00	101.30	101.28	NNW	5	0.0	87	15.7	17.9	0.0	0.0	32200.0


03
Key Insights

Highest and Lowest Air Temperatures

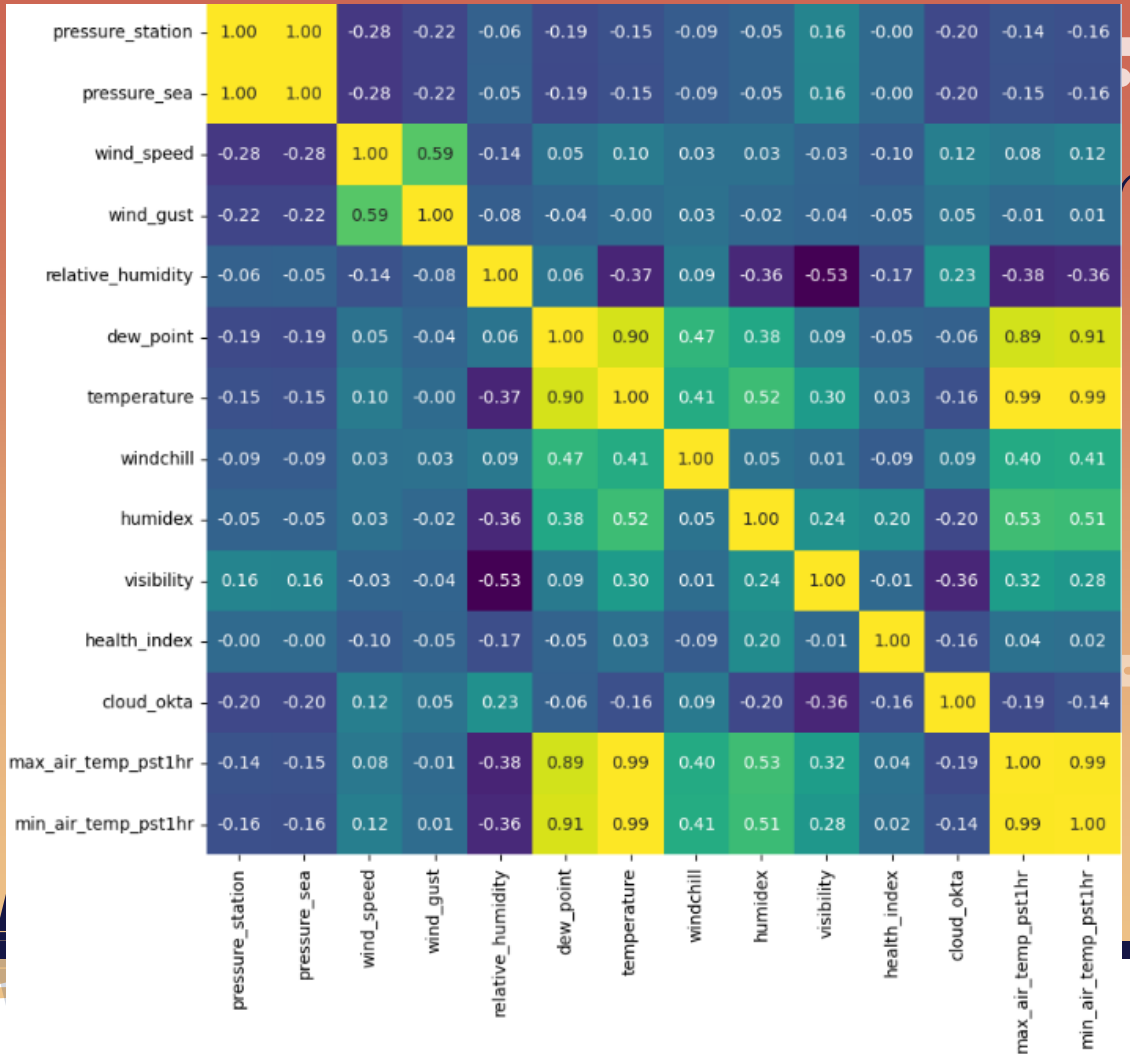
Lowest temperature in the past 10 years:

--15.3°C(not -14.5°C)
December 27, 2021, at 05:00

Highest temperature in the past 10 years:

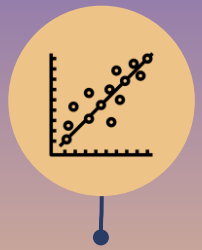
32.6°C (not 32.1°C)
June 29, 2021, at 15:00

Little to no Correlation to Temperature




04
Model Comparison / Interpretation

Models Fitted to Data

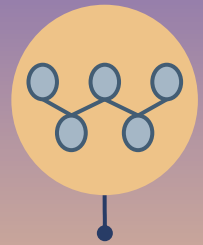


Linear Regression

R^2 = 99.64%
MAPE = 7.22%

FB Prophet

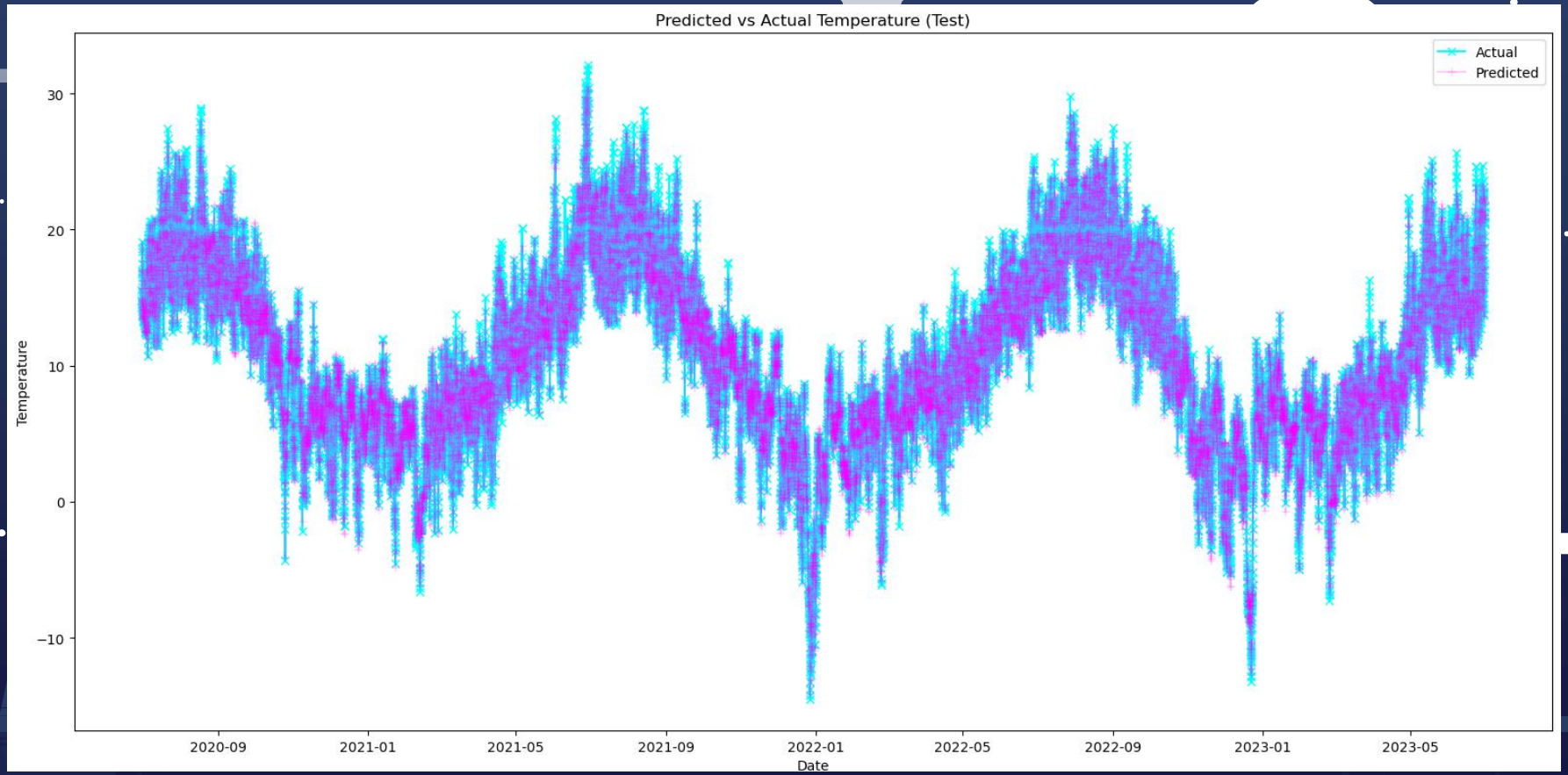
R^2 = 99.53%
MAPE = 8.33%



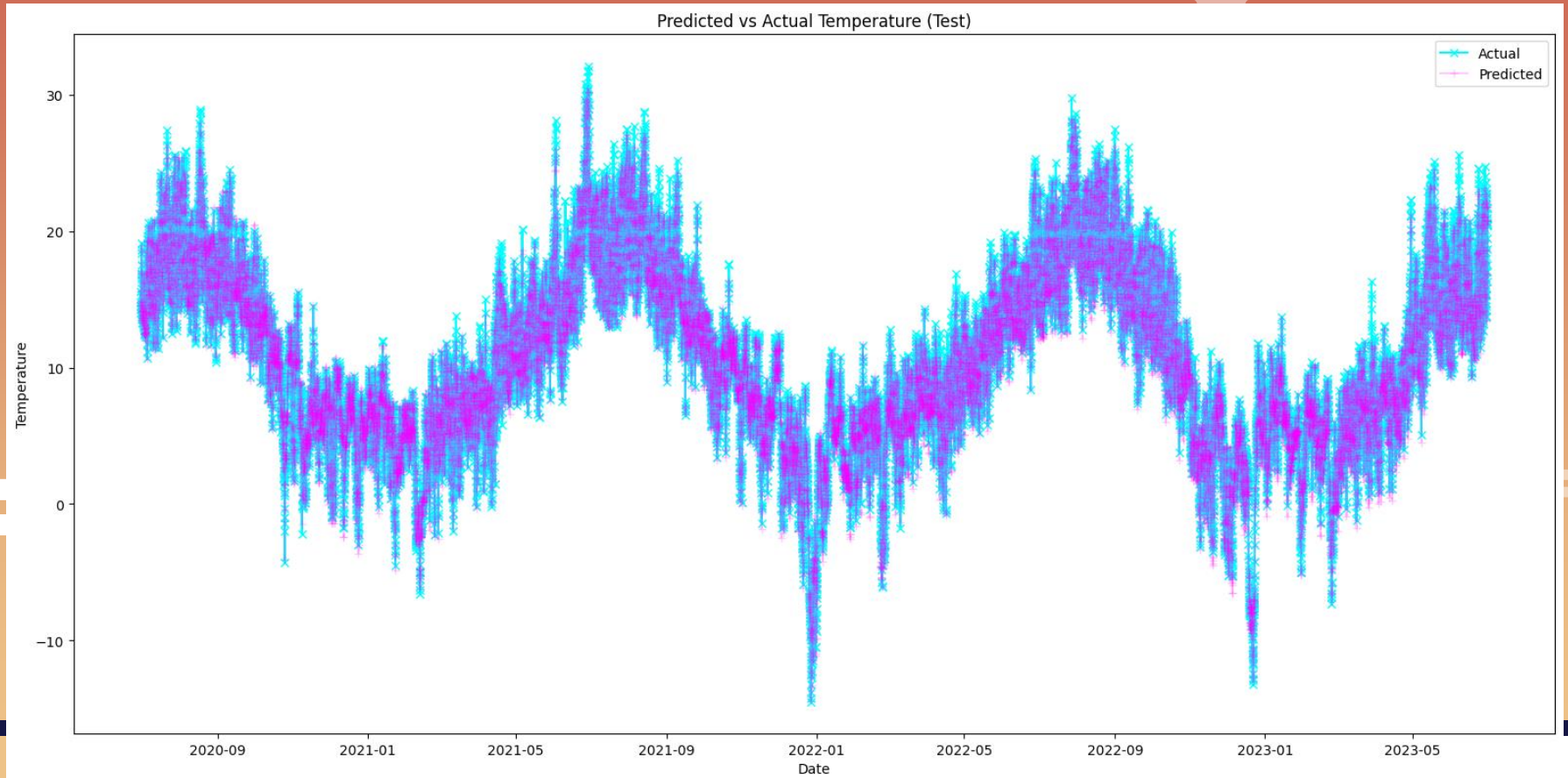
LSTM

R^2 = 99.96%
RMSE = 7.22%

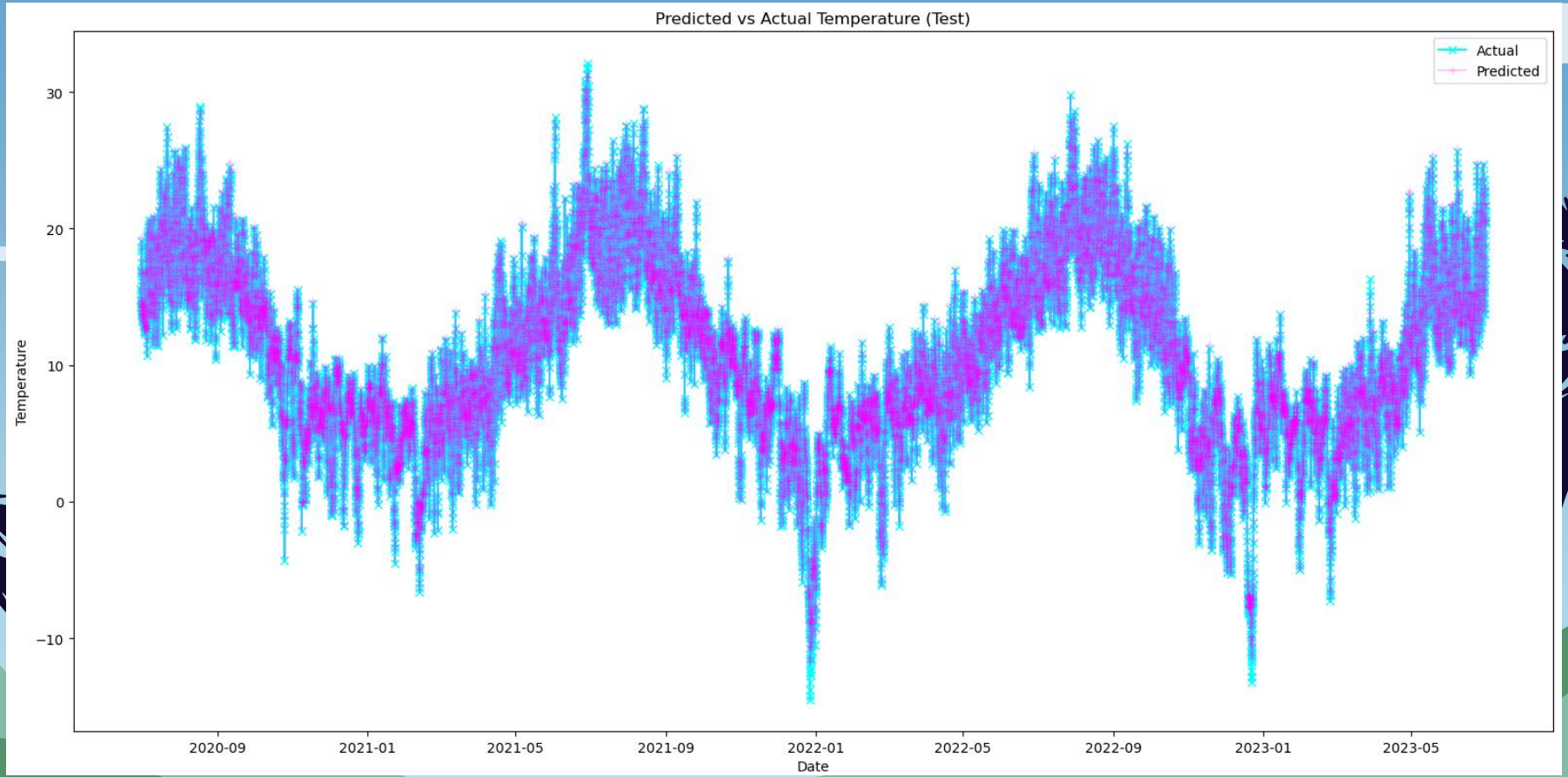
Linear Regression



Prophet


Predicted vs Actual Temperature (Test)
Actual
Predicted
Temperature
30
20
10
0
−10
2020-09
2021-01
2021-05
2021-09
2022-01
2022-05
2022-09
2023-01
2023-05
Date

LSTM

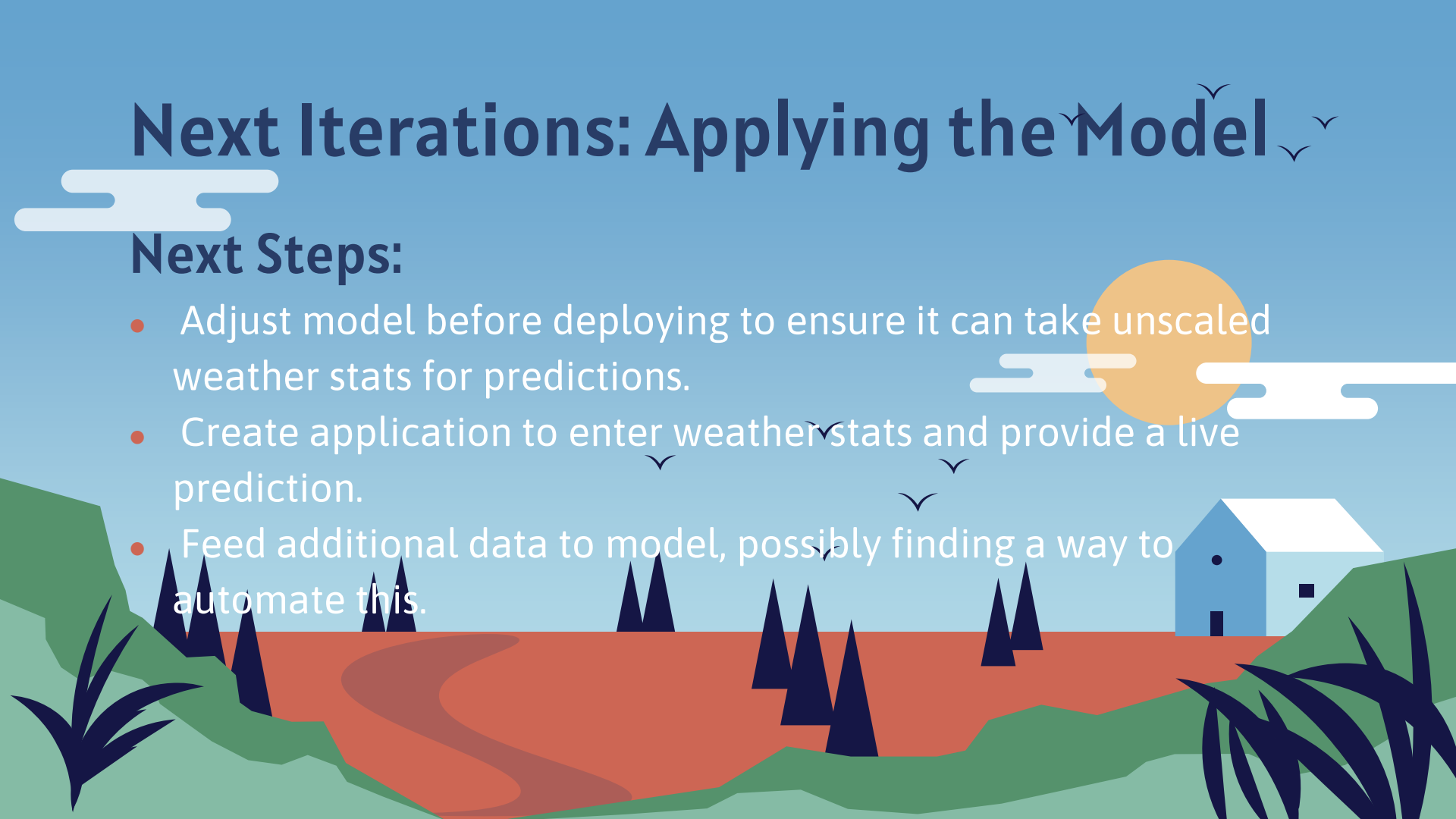

Predicted vs Actual Temperature (Test)
Actual
Predicted
Temperature
30
20
10
0
−10
2020-09
2021-01
2021-05
2021-09
2022-01
2022-05
2022-09
2023-01
2023-05
Date


05
Next Steps

Next Iterations: Applying the Model

Next Steps:

- Adjust model before deploying to ensure it can take unscaled weather stats for predictions.
- Create application to enter weather stats and provide a live prediction.
- Feed additional data to model, possibly finding a way to automate this.



Thanks!

Pedro A. Montano