Omdena Sillicon Valley Chapter

Xtreme Weather Forecast

Chapter Lead
Nishrin Kachwala

PRESENTED BY
Md. Anas Mondol

THE PROBLEM TO BE SOLVED

Extreme swings in temperatures and precipitation drive the need for accurate long-term forecasts.

Physics-based models dominate short-term weather forecasting. But these models have a limited forecast horizon.

With the availability of meteorological data, data scientists can improve weather forecasting by blending it with physics-based forecasting.

Data

- The data is provided by WIDS, in collaboration with Climate Change AI (CCAI).
- There is weather and climate information for a number of US locations, for a number of start dates for the two-week observation, as well as the forecasted temperature and precipitation from a number of weather forecast models.
- Each row in the data corresponds to a single location and a single start date for the two-week period.

TASK: The prediction task involving forecasting sub-seasonal temperatures (temperatures over a two-week period) within the United States, for each location and start date.

Tasks Performed

1. Data Preprocessing
2. Exploratory Data Analysis
3. Feature Engineering
7. Model Evaluation

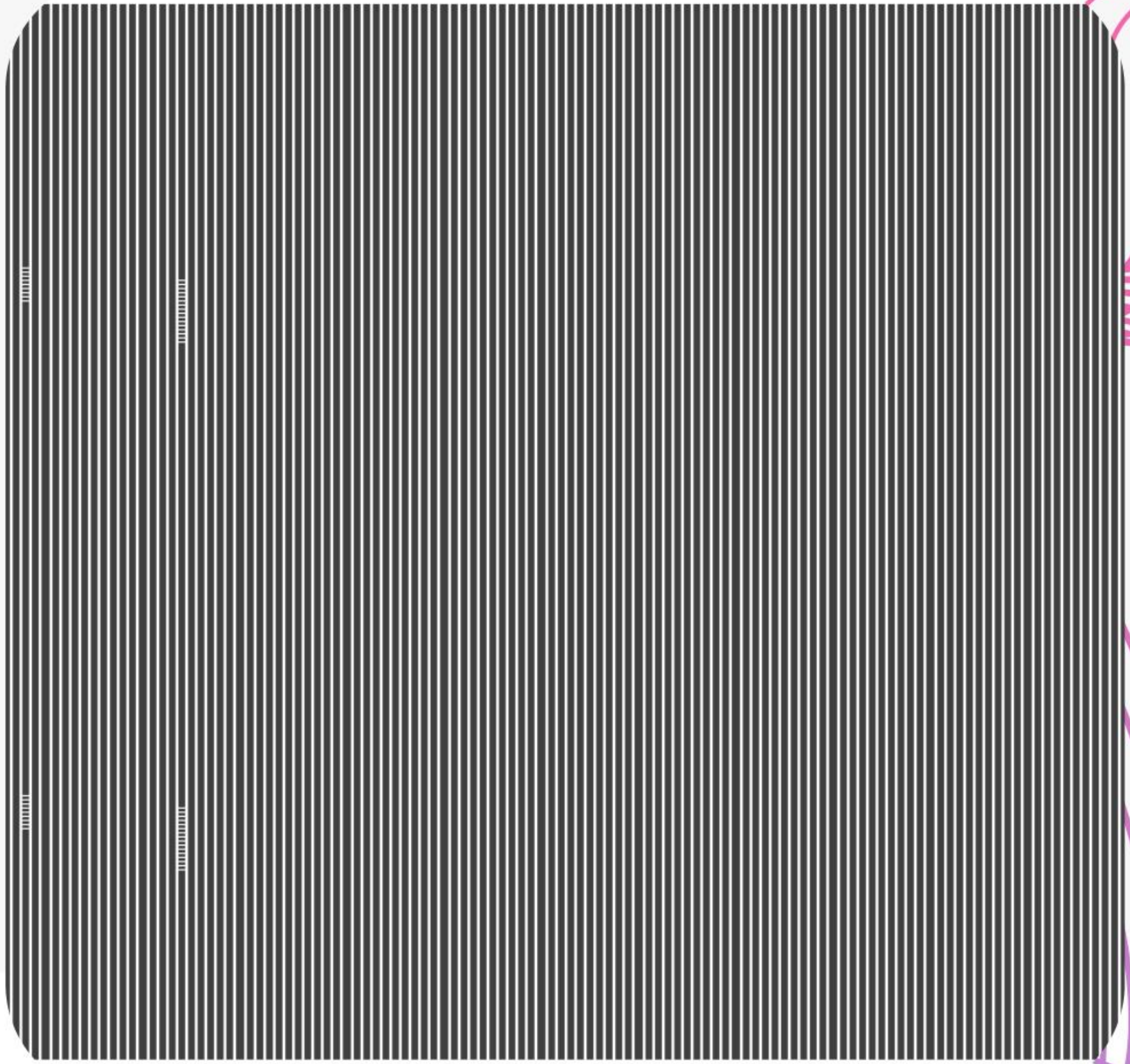
DATA PREPROCESSING

Null Values

OBTAINING NULL VALUES BY IMPUTATION AND OTHER APPROACHES

Processing

PROCESSING THE DATA AND MAKING THE DATASET READY FOR ANALYSIS



Data Preprocessing

- There were a **few missing values** in the dataset.
- Most of the missing values were of specific columns and timeframe (here referred to **months**).
- Thus it was the case of **NMAR** (Not Missing at Random)

Methodology

- We used many Imputers like Mean, Median, Most Frequent and KNN Imputer to extract the null values.

Proposed Solution for Data Preprocessing

Step I

Checked in which columns are the values missing.

Step II

Checked on which indices are the missing values changing.

Step III

Evaluating the null values by the help of various Imputers.

Omdena SILICON VALLEY

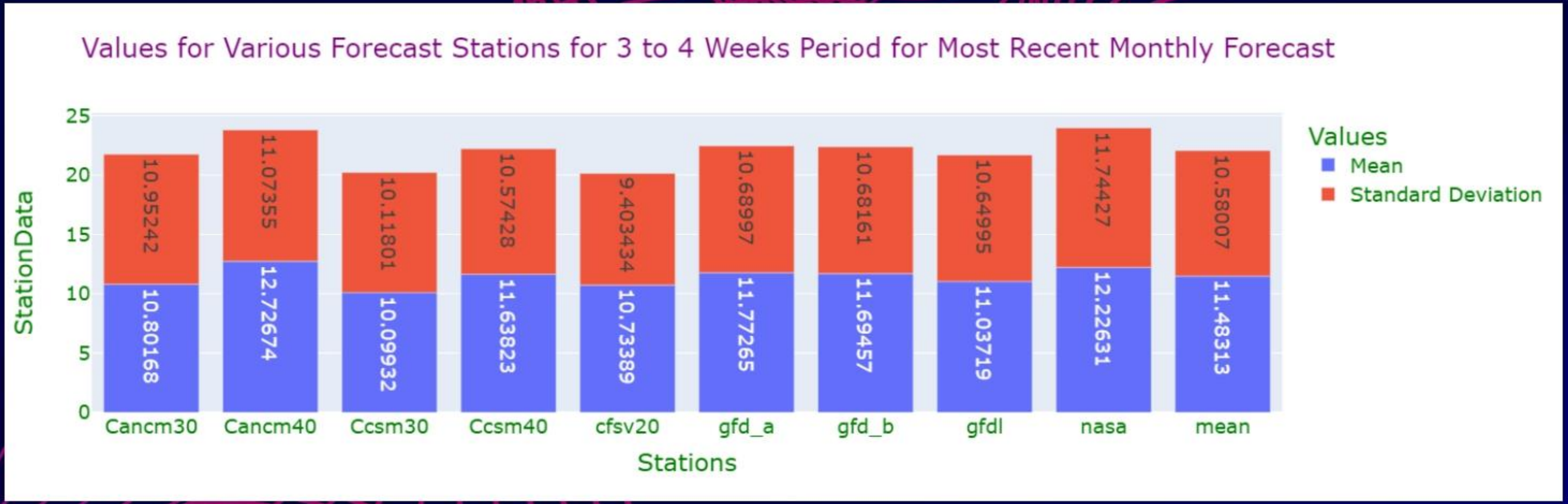
EXPLORATORY DATA ANALYSIS

Analysis

ANALYSIS OF DATASET IS CRUCIAL FOR MODEL DEVELOPMENT AND TRAINING

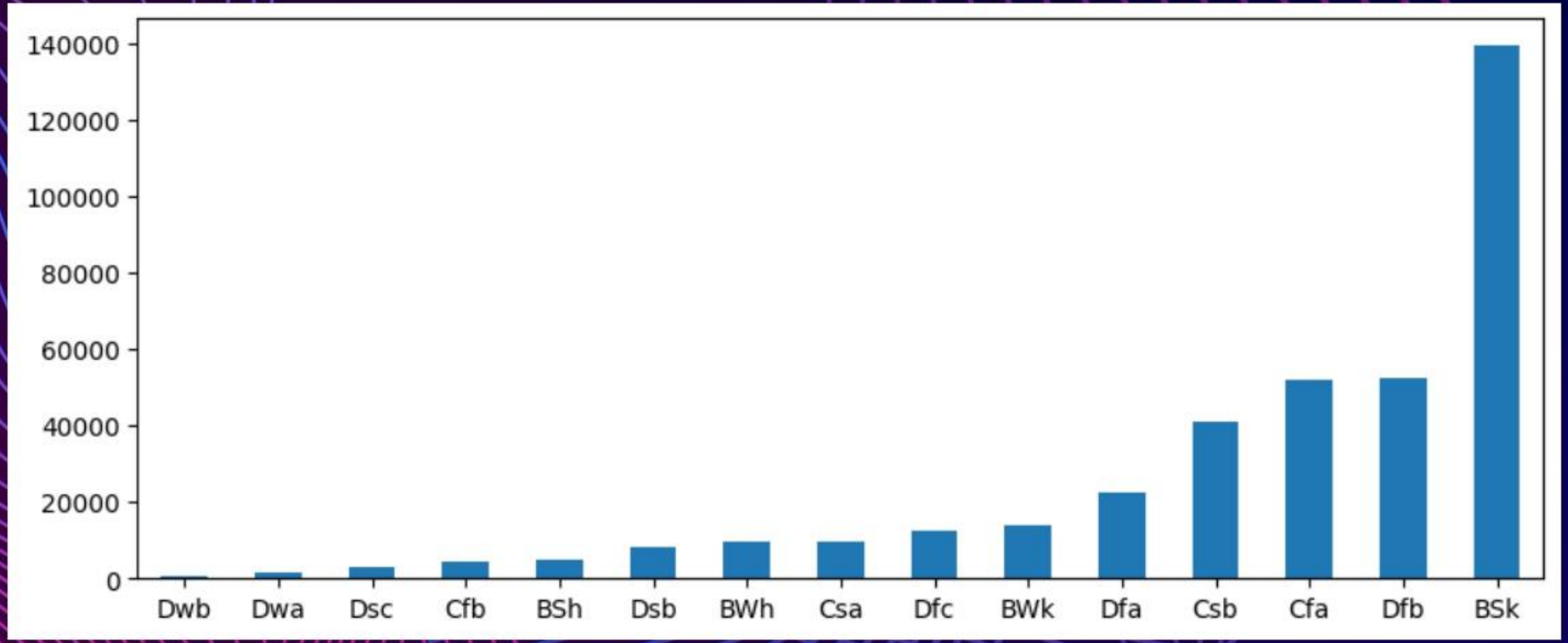
Visualization

ANALYSIS IS ONLY COMPLETE WHEN THE VISUALIZATION IS PROPER AND ROBUST



Distribution of Climatic Regions

The distribution of number of data points in terms of Regions is uneven.

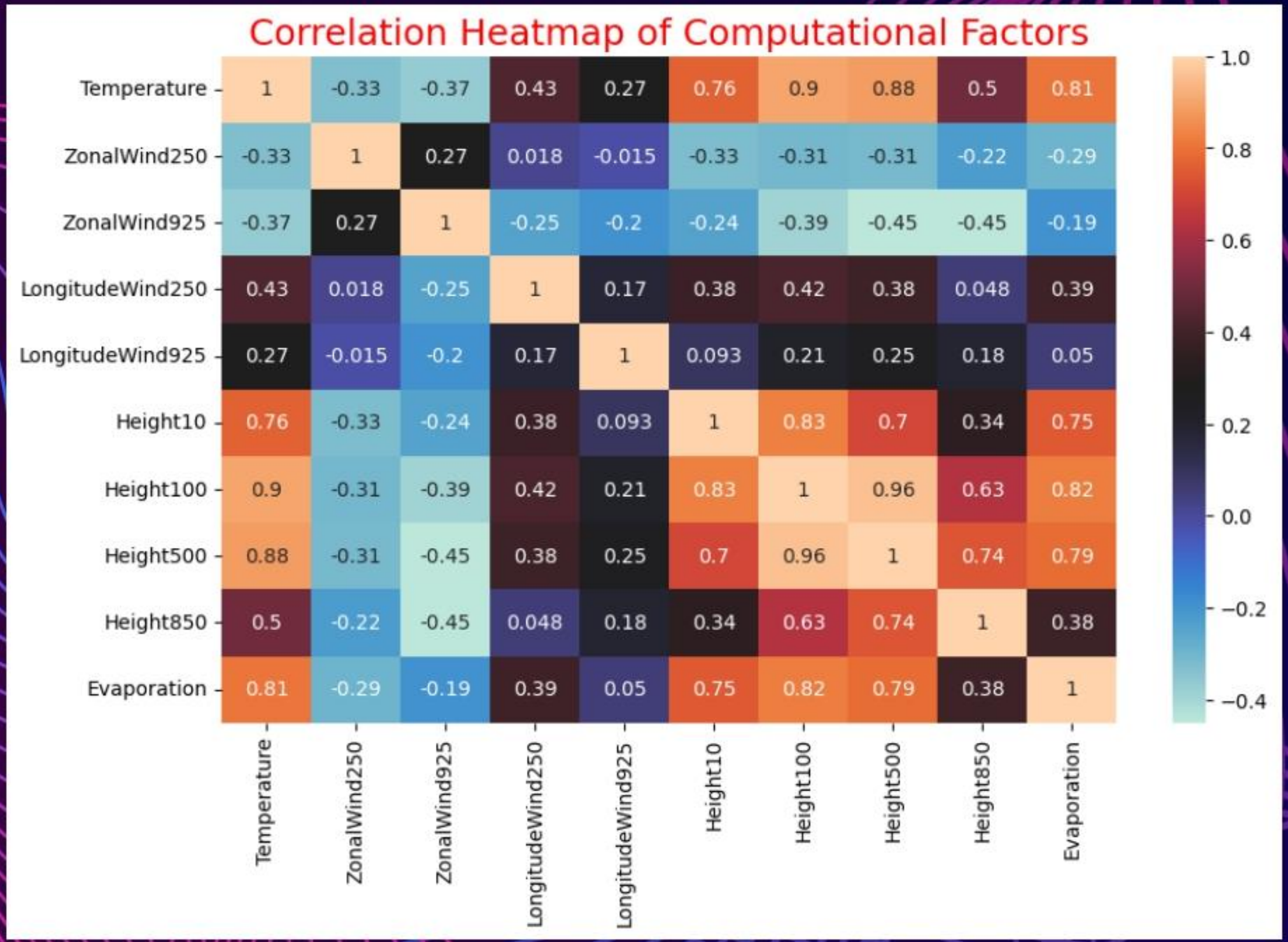


Analysis Of Weather Station Temperatures

The mean and deviations of all Weather stations is almost similar. They computed the temperature with almost same accuracy.

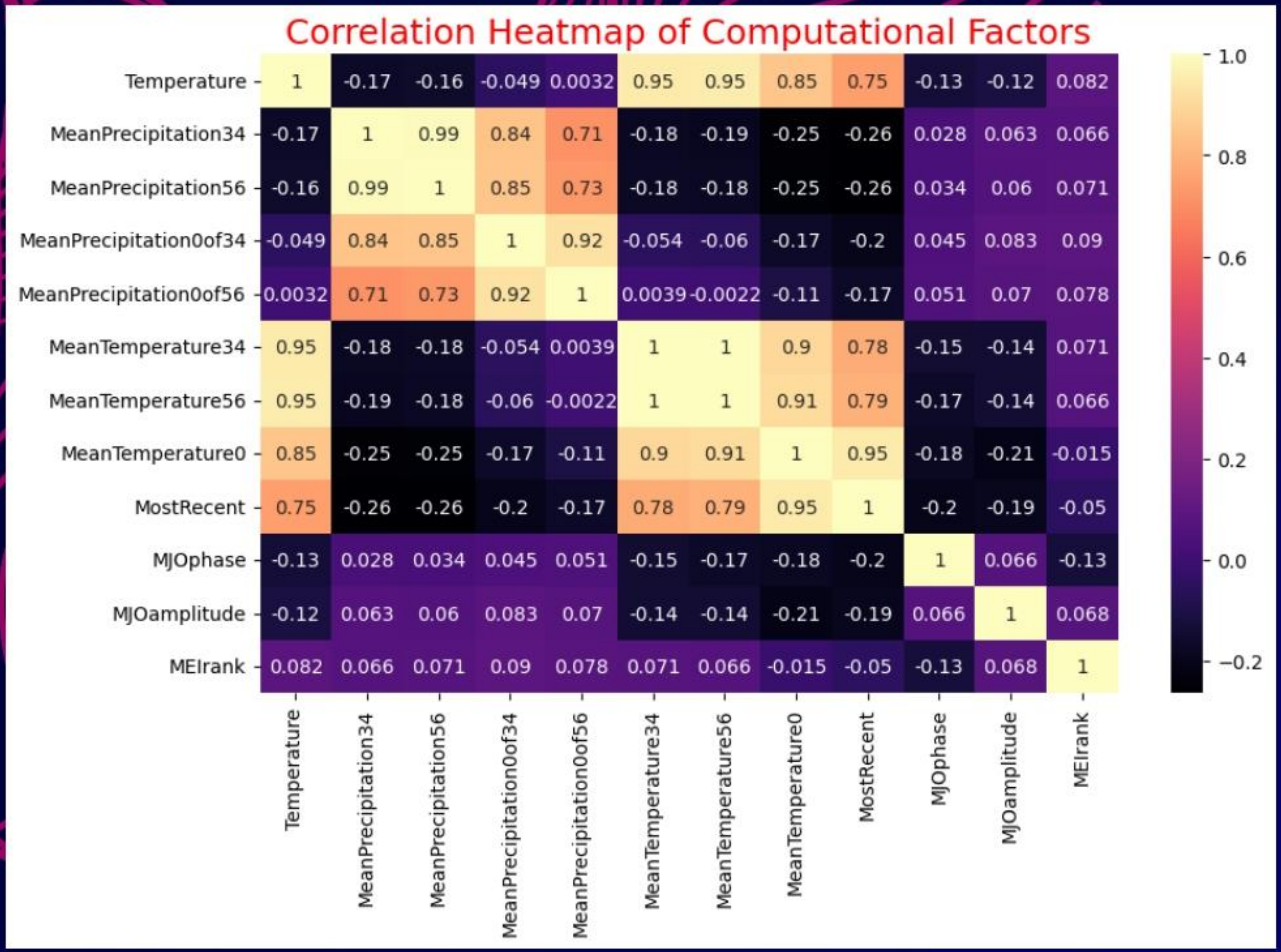
Correlation Matrix

Matrices were also used to compute the irrelevant factors as well.



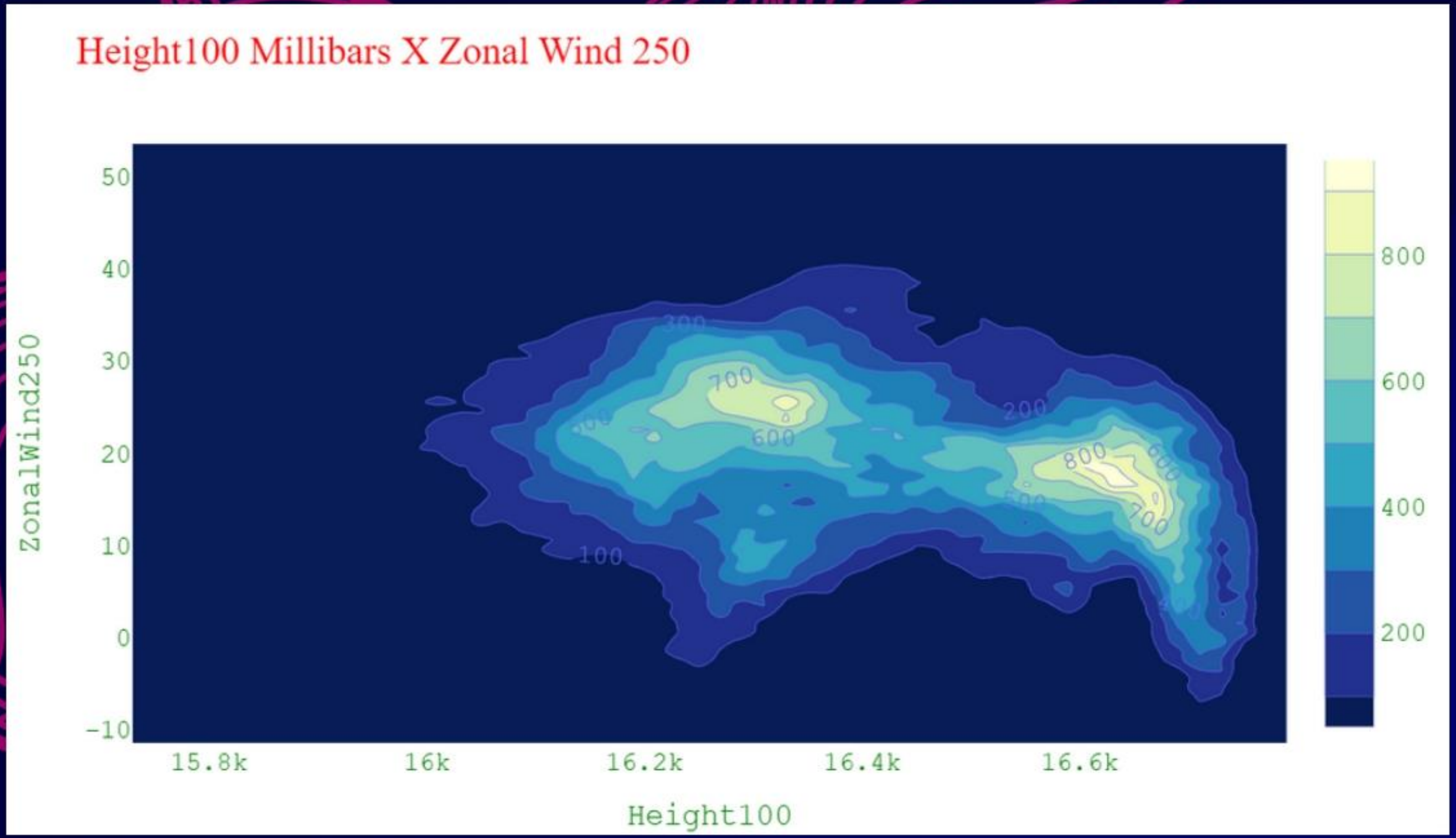
Correlation Matrix

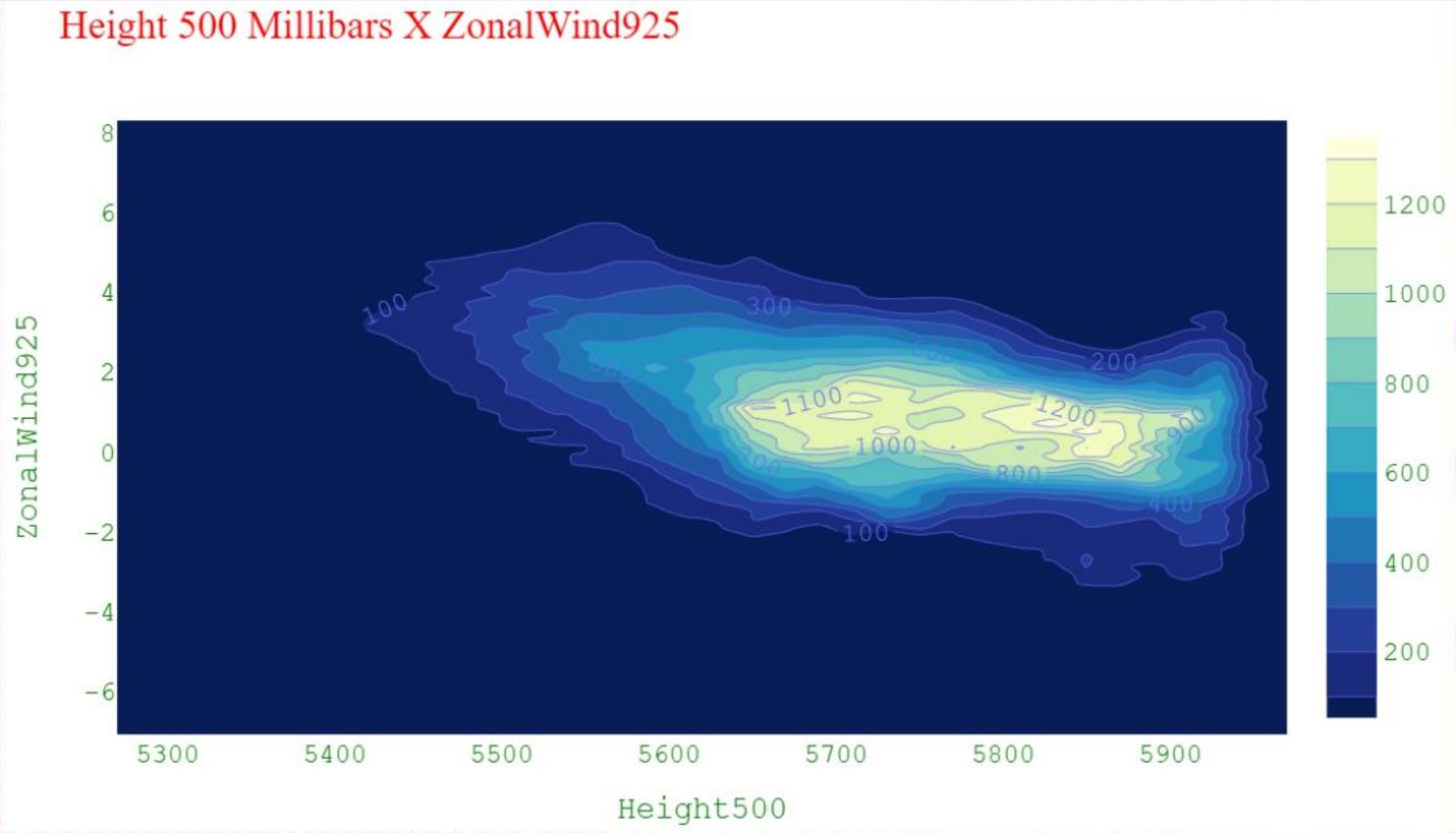
The Correlation matrices were used to get the correlated factors with respect to the target variable.



Contour Maps

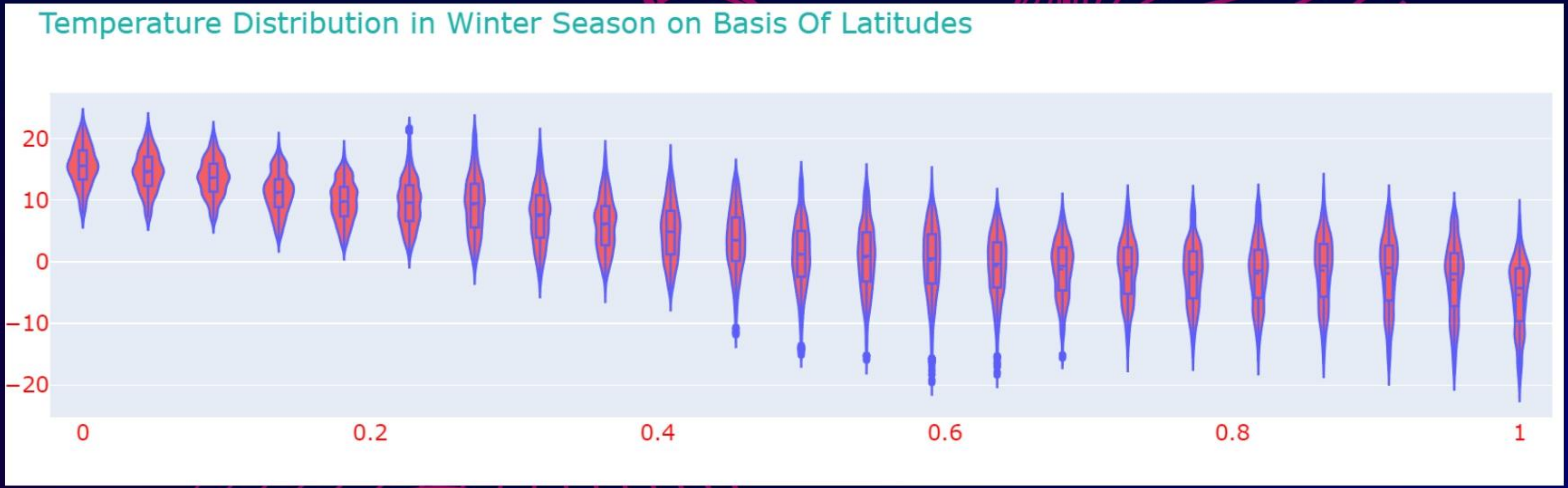
Contour Maps to check the Intensity points and distribution of variables.





Contour Maps

Contour Maps were also used to get the density of data points among the variables.



Dataset Splitting

To perform more influential EDA we performed it by splitting the dataset by **various influential factors** such that the dataset was splitted into the following clusters:-

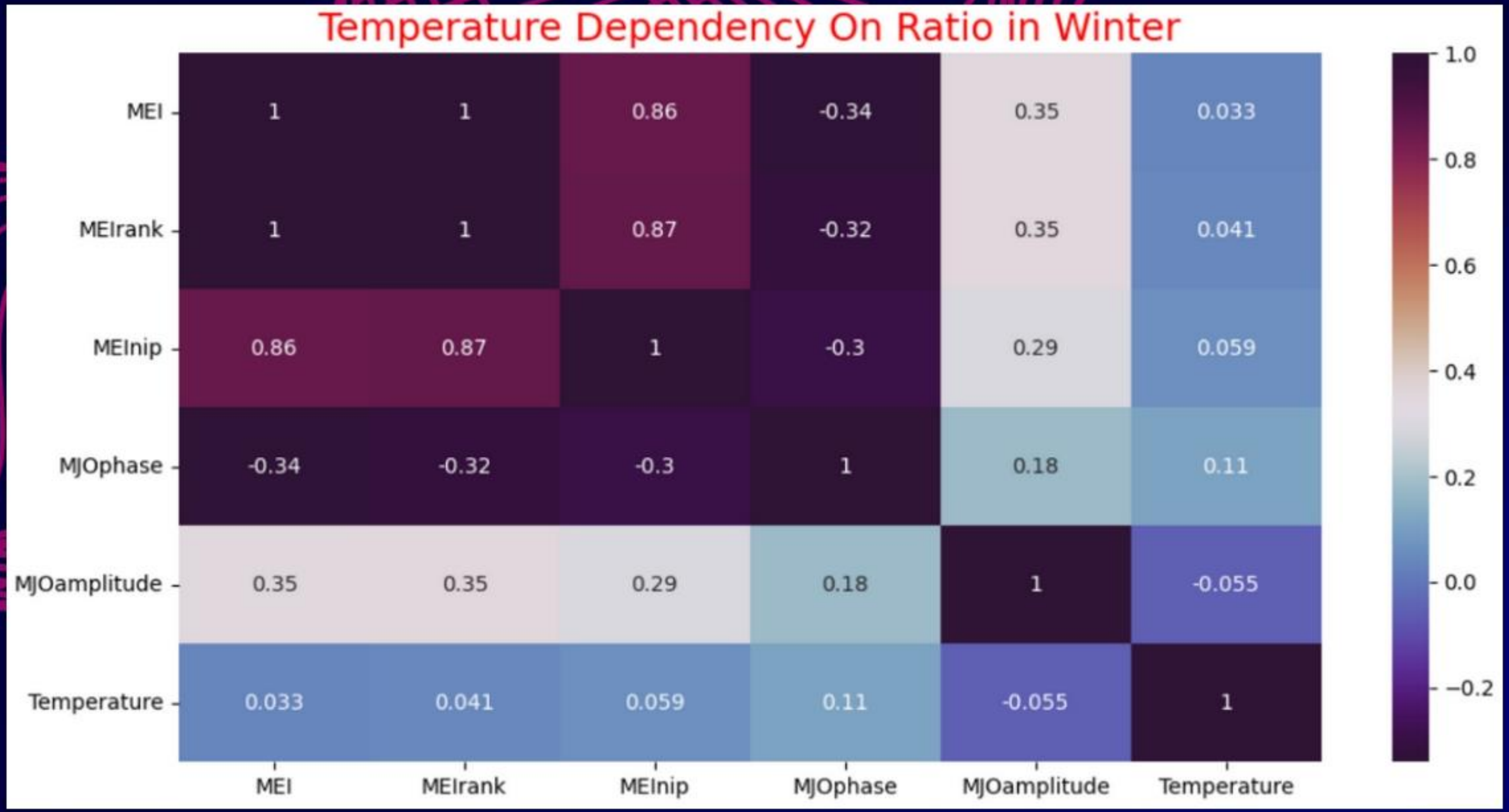
1. **Seasons**
2. **Year**
3. **Regions**

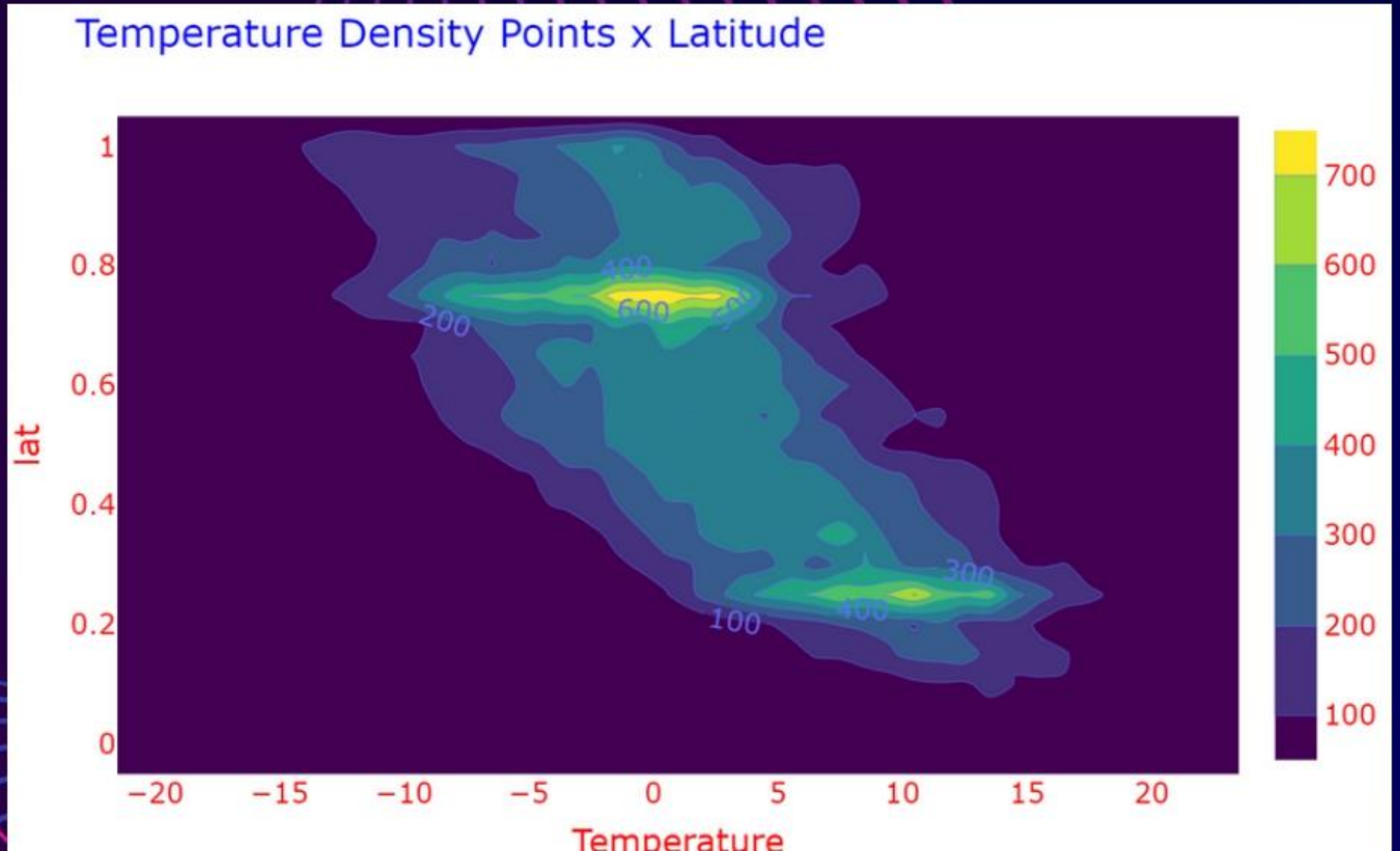
Violin Plots

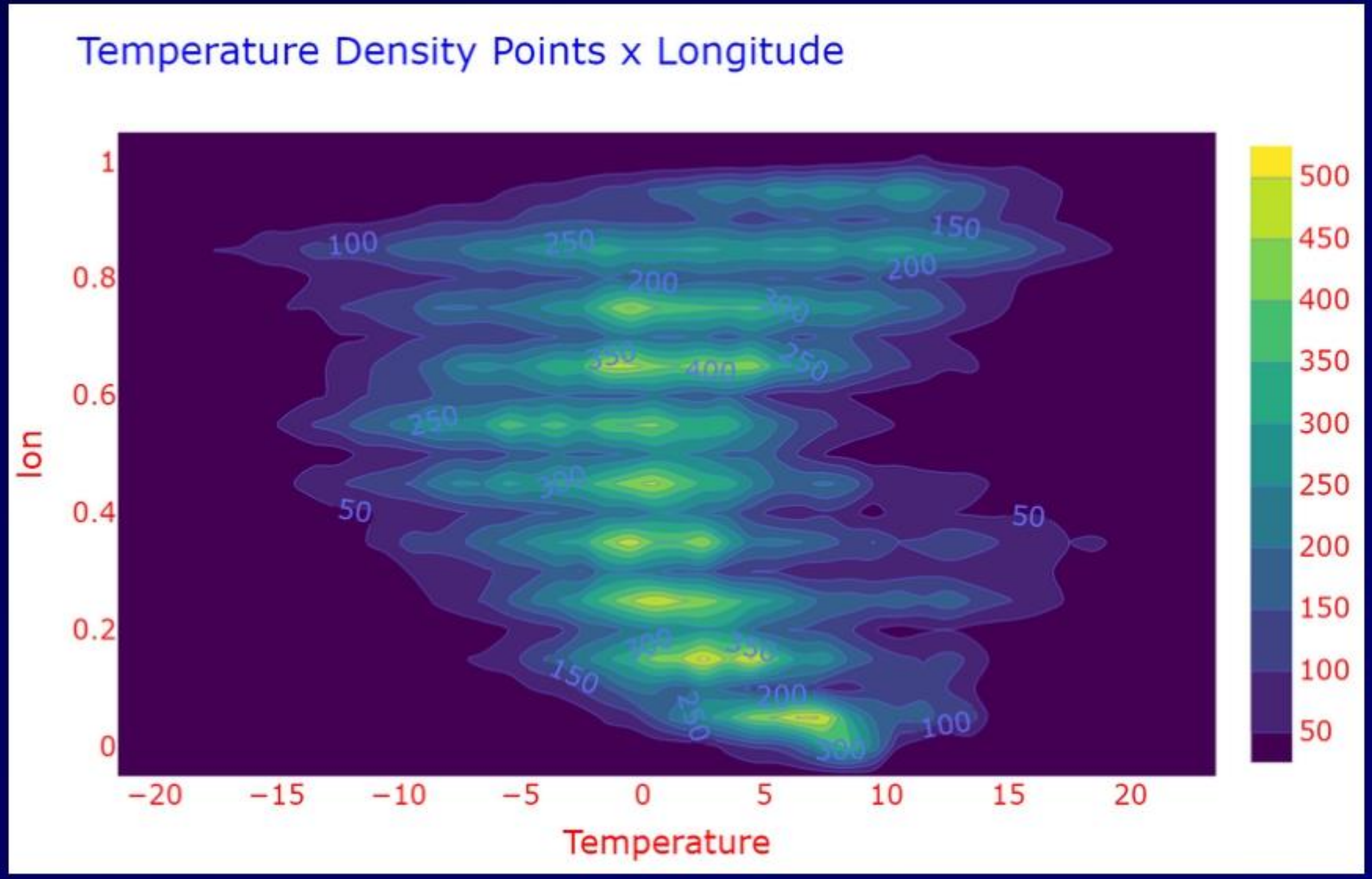
The violin plots were used to check the outliers in the given data, so the insights could be provided to feature engineering team as well.

Feature Extraction

Heatmaps and Contour Maps were widely used for Feature-extraction and eradicating irrelevant columns.







MODELING

Literature review	TEMPERATURE PREDICTION PROBLEM
Model selection	TIME SERIES PROBLEM REGRESSION PROBLEM LSTM
Model development	USING TUTORIALS
Model evaluation	RMSE, LOSS FUNCTION

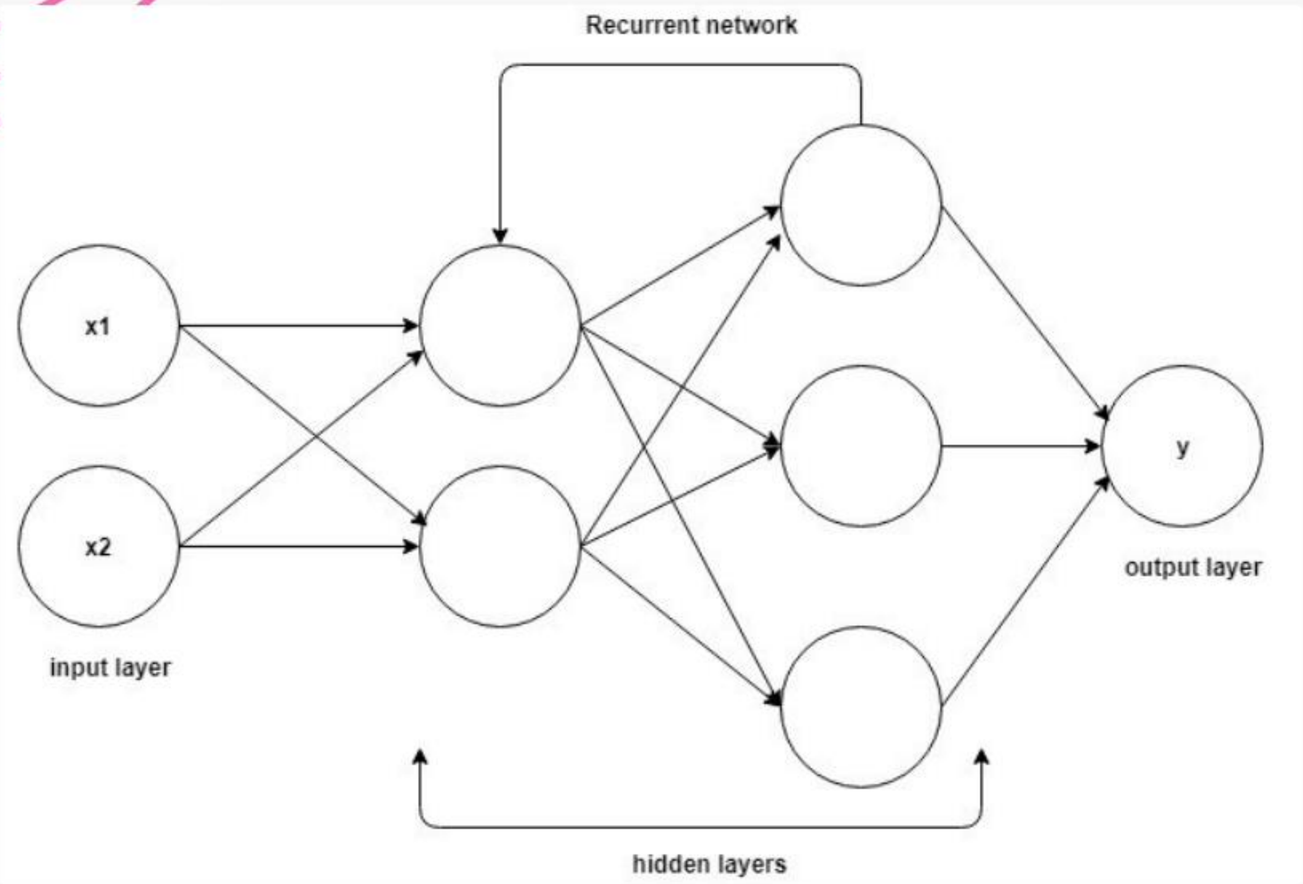
Why LSTM?

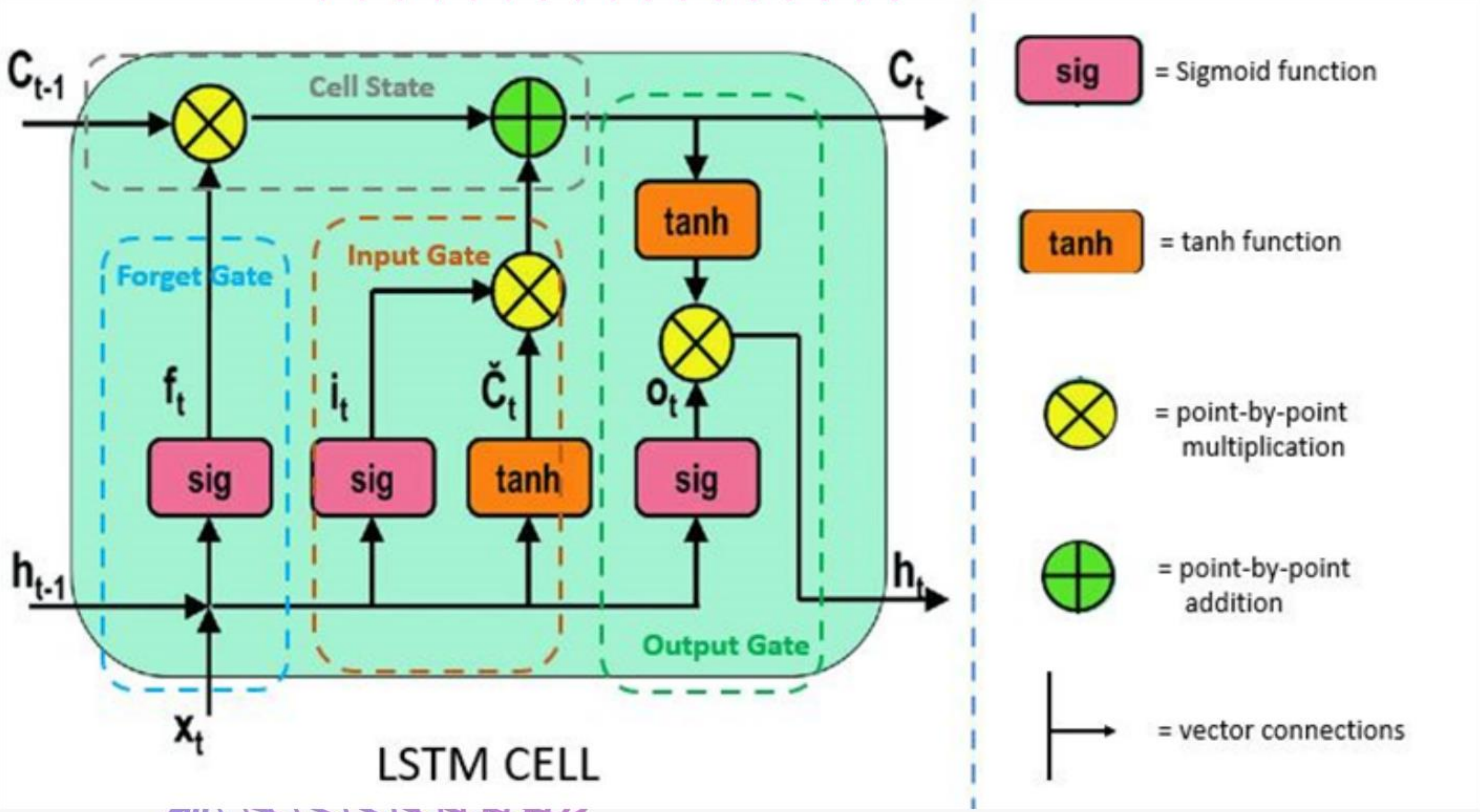
- **Literature review for the temperature prediction problem**
- **The most popular model: 11 publications**
- **Available tutorials**

LSTM	Introduction to Long Short Term Memory (LSTM)		https://www.analyticsvidhya.com/blog/2021/03/introduction-to-long-short-term-memory-lstm/#:~:text=Long%20Short%20Term%20Memory%20Network,is%20used%20for%20persistent%20memory	tutorial
LSTM	Weather forecasting		https://blog.paperspace.com/weather-forecast-using-ltsm-networks/	tutorial
LSTM	Time series forecasting		https://www.kaggle.com/code/mineshjethva/time-series-forecasting-with-lstm-for-uni-multivar	tutorial
Bidirectional LSTM	Time series forecasting with Bidirectional LSTM and other DL models		https://medium.com/@dave.cote.msc/hands-on-advanced-deep-learning-time-series-forecasting-with-tensors-7facae522f18	tutorial
Bidirectional LSTM, CNN	Temperature prediction with Bidirectional LSTM		https://github.com/PotatoThanh/Bidirectional-LSTM-and-Convolutional-Neural-Network-For-Temperature-Prediction	git example

Bidirectional LSTM	Stack layer & Bidirectional Layer Long Short - Term Memory (LSTM) Time Series Model with Intermediate Variable for weather Prediction	2021	https://ieeexplore.ieee.org/document/9752357
LSTM	Global Land Temperature Forecasting Using Long Short-Term Memory Network	2020	https://ieeexplore.ieee.org/document/9191567
LSTM	Weather Prediction Using LSTM Neural Networks	2022	https://ieeexplore.ieee.org/document/9824268
Bidirectional LSTM	Forecasting of Temperature by using LSTM and Bidirectional LSTM approach: Case Study in Semarang, Indonesia	2021	https://ieeexplore.ieee.org/document/9617495
LSTM	Prediction Of Temperature And Rainfall In Bangladesh Using Long Short Term Memory Recurrent Neural Networks	2020	https://ieeexplore.ieee.org/document/9254585
LSTM	A Method for Weather Forecasting Using Machine Learning	2021	https://ieeexplore.ieee.org/document/9672403

LSTM background





- **At forget gate: sigmoid function**
- **At the input gate: both tanh and sigmoid function**
- **At the output gate: tanh is applied to new cell state and sigmoid to current input and previous hidden state**

IMAGES SOURCES: HTTPS://WWW.PLURALSIGHT.COM/GUIDES/INTRODUCTION-TO-LSTM-UNITS-IN-RNN HTTPS://TOWARDSDATASCIENCE.COM/MACHINE-LEARNING-RECURRENT-NEURAL-NETWORKS-AND-LONG-SHORT-TERM-MEMORY-LSTM-PYTHON-KERAS-EXAMPLE- 86001CEAAEBC

LSTM architecture

for the temperature prediction problem

LITERATURE REVIEW

- 1-5 HIDDEN LAYERS (1 LAYER FOR MAJORITY) AND 5-16 UNITS
- ACTIVATION FUNCTION: SIGMOID, RELU
- OPTIMIZER: RMSPROP, NOT SPECIFIED FOR MAJORITY
- LEARNING RATE: 0.001-0.2
- BATCH SIZE: 32-50, NOT SPECIFIED FOR MAJORITY
- EPOCH: 10, NOT SPECIFIED FOR MAJORITY

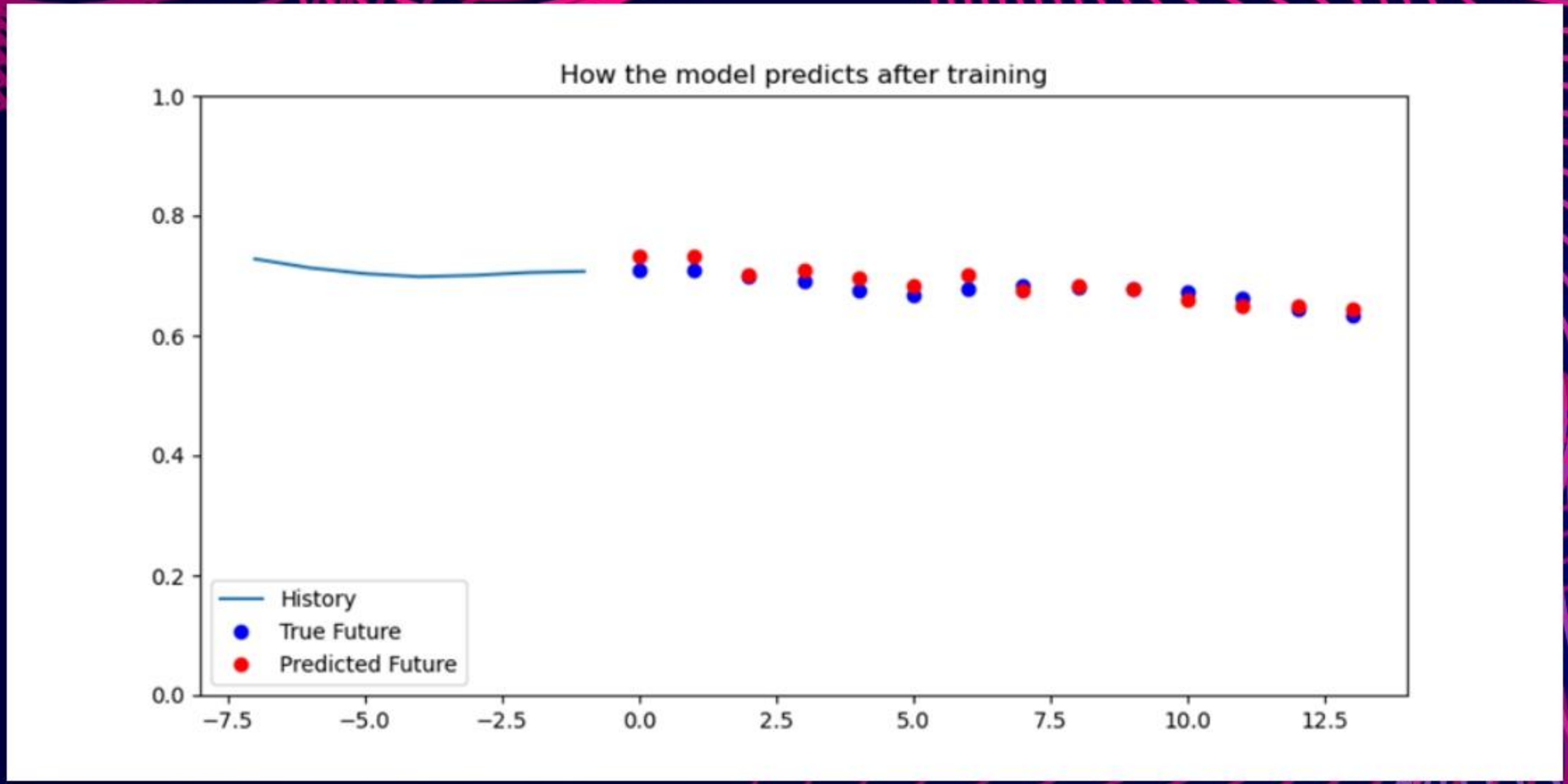
TEAM EXPERIMENTS

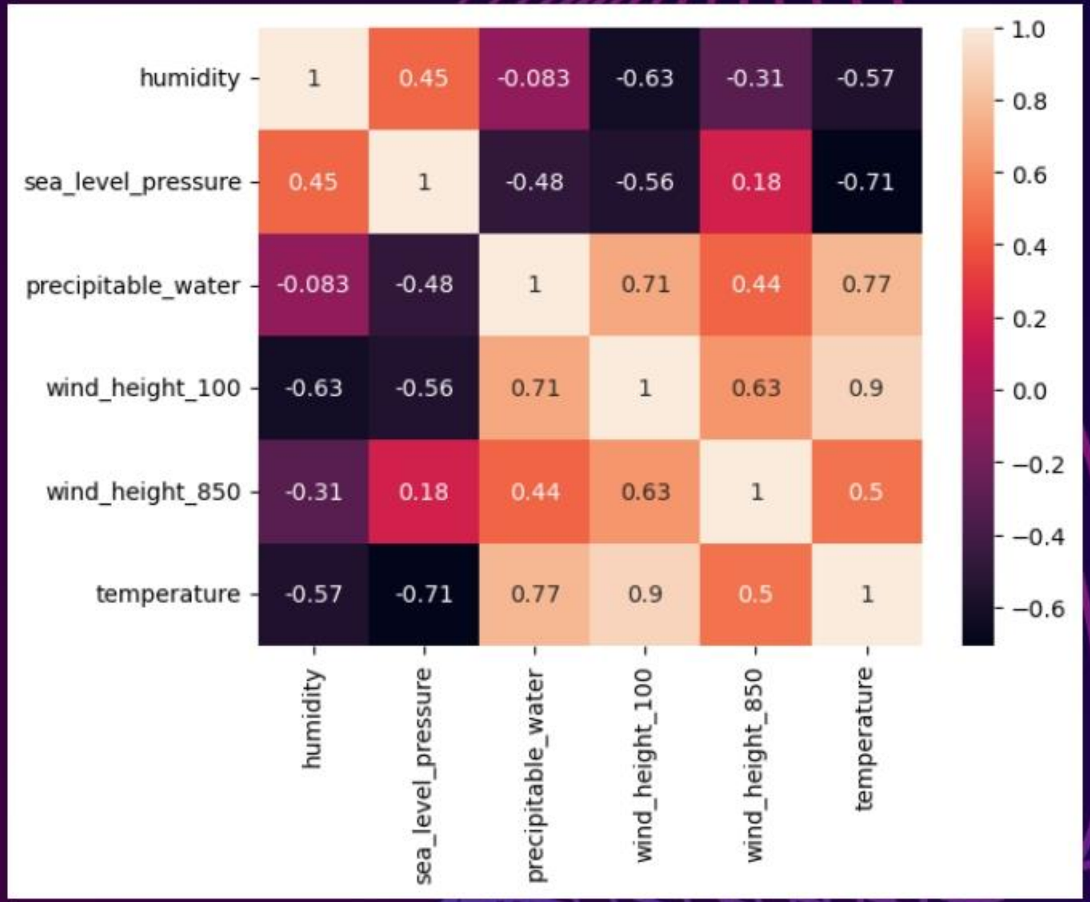
- 1-3 HIDDEN LAYERS AND 20-128 UNITS
- ACTIVATION FUNCTION: TAHN, SIGMOID, RELU, SOFTMAX
- OPTIMIZER: RMSPROPY
- LEARNING RATE: 0.01
- EPOCH: 10-30
- WINDOW SIZE: 7-21
- BATCH SIZE: 16-256

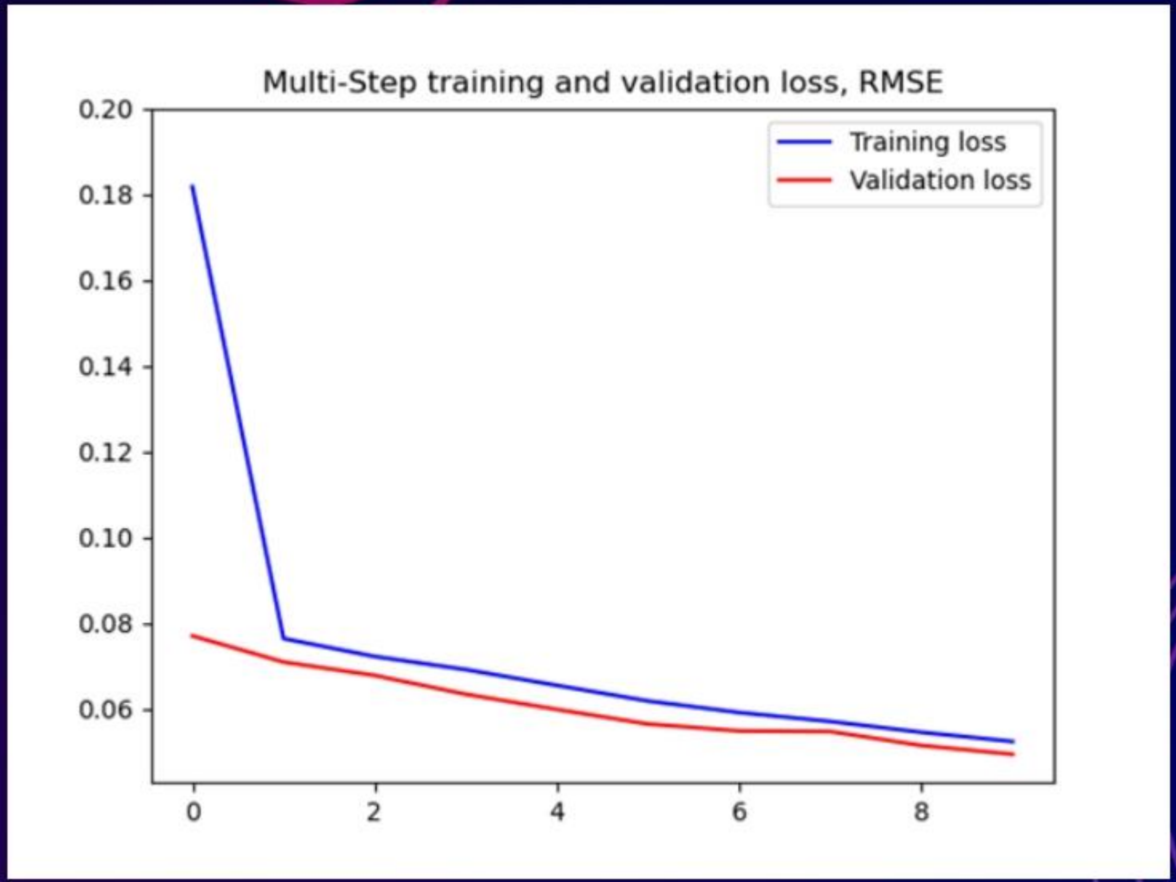
Model evaluation

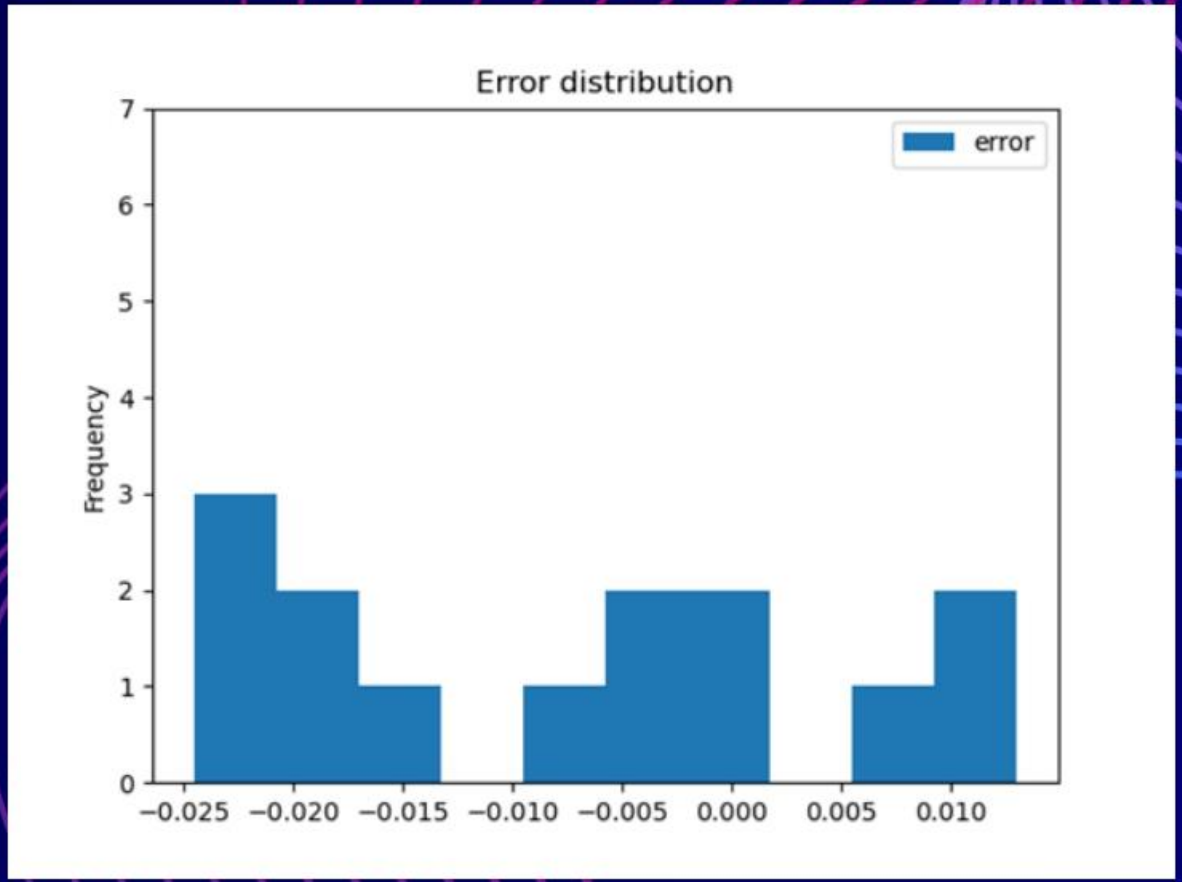
LSTM CONFIGURATION

- WINDOW SIZE: 7
- INPUT LAYER: 6 FEATURES
- THE FIRST LAYER: 32 NEURONS, ACTIVATION ‘TANH’
- THE SECOND LAYER: 16 NEURONS, ACTIVATION 'SIGMOID‘
- OUTPUT LAYER: 14 NEURONS (PREDICTIONS)
- OPTIMIZER: RMSPROP
- EPOCH: 10









The observations from changing the architecture

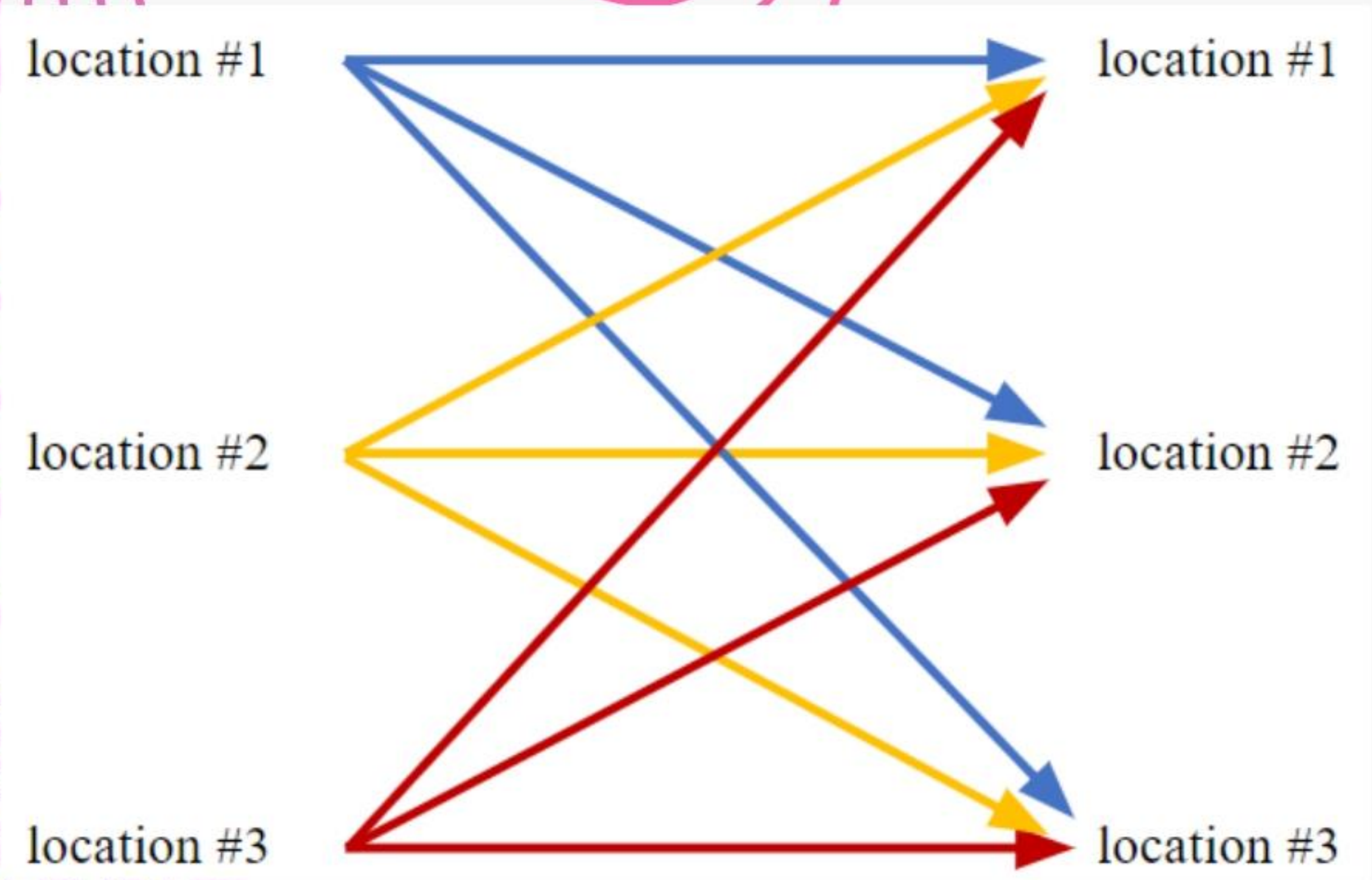
- **LAYERS**: INCREASE THE ACCURACY, BUT NO MORE THAN 3
- **NEURONS**: 50 IS OPTIMAL FOR 1-LAYER, DO NOT INFLUENCE 2-LAYER AND 3-LAYER
- **ACTIVATION FUNCTIONS**: RELU IS OPTIMAL FOR 1-LAYER, DIFFERENT ACTIVATION FUNCTIONS FOR DIFFERENT LAYERS PERFORM BETTER THAN THE SAME

- THE BEST RESULT OF 1-LAYER: 50 NEURONS, RELU AS ACTIVATION FUNCTION, 30 EPOCHS, 0.04 RMSE, 62% ACCURACY
- THE OBSERVATION OF 2-LAYERS: 10 EPOCHS, 0.04 RMSE, 41% ACCURACY, CHANGING THE NUMBER OF NEURONS DO NOT GIVE GOOD ADVANTAGES COMPARING THE LOSS FUNCTION, THE SAME ACTIVATION FUNCTIONS OF TWO LAYERS SHOW WORSE RESULTS THAN DIFFERENT ONE
- THE OBSERVATION OF 3-LAYERS: 20 EPOCHS, 0.04 RMSE, 52% ACCURACY, THE CHANGE IN NUMBER OF THE LAYERS IS ABLE TO INCREASE THE ACCURACY (BUT NO MORE THAN 3 LAYERS), CHANGING ACTIVATION FUNCTIONS AND THE NUMBER OF NEURONS DOES NOT INFLUENCE THE ACCURACY AND LOSS FUNCTION MUCH

How could the model be improved?

WEATHER MODELS

CANCM3 CANCM4
CCSM3
CCSM4
CFSV2
GFDL
GFDL-FLOR-A
GFDL-FLOR-B
NASA



- **to add features of several locations***
- **to include weather models predictions with different weights***
- **to proceed experiments with LSTM architecture for new features configuration**

***recommendations are provided by the domain expert Licheng Geng**

Acknowledgment

TEAM,
THANK YOU VERY MUCH FOR YOUR TIME, SUPPORT, CURIOSITY, AIM TO LEARN, PARTICIPATING IN THE MEETINGS, COLLABORATION IN SLACK, HARD WORK ON THE MODEL, SHARING OBSERVATIONS AND OPINIONS!

VISHU KALIER, YIRUI ZHU, LICHENG GENG, NISHA MENON, TEKLE, AND OTHERS!

THANK YOU FOR YOUR CONTRIBUTION AND EFFORT!
WE COULD MOVE THE PROJECT TOGETHER!