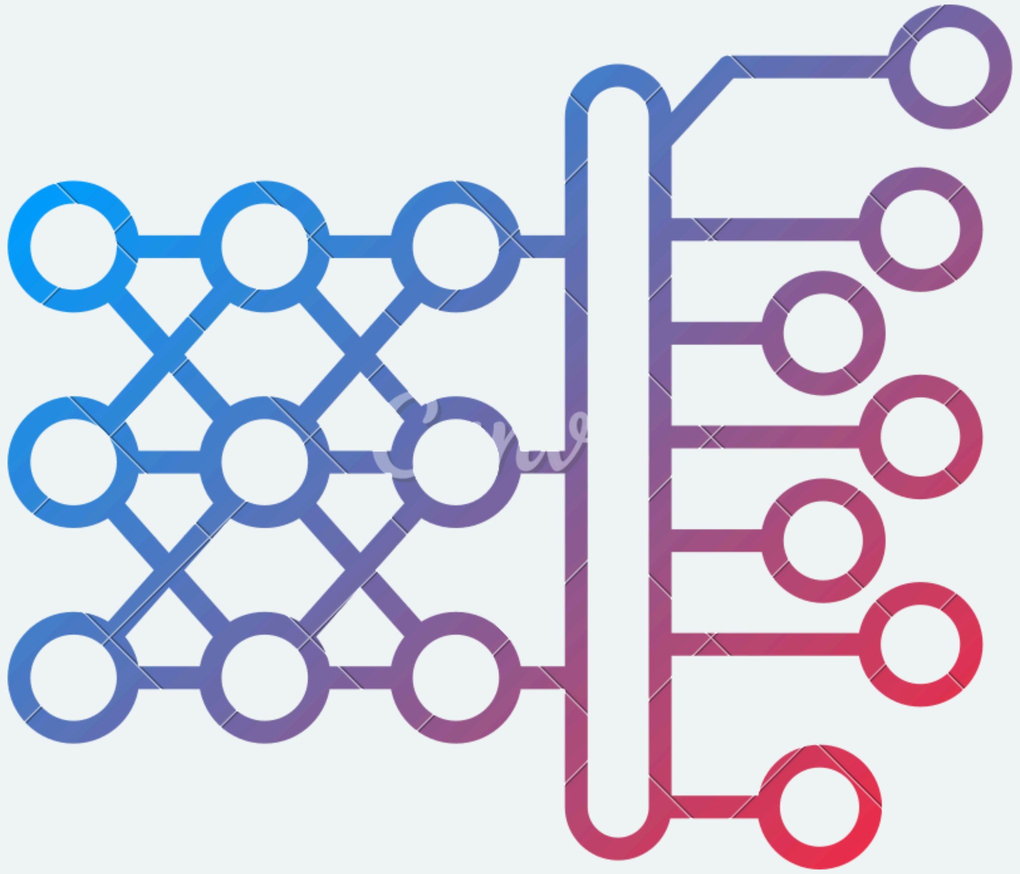


DEEP LEARNING PROJECT

DATA-DRIVEN WEATHER FORECAST

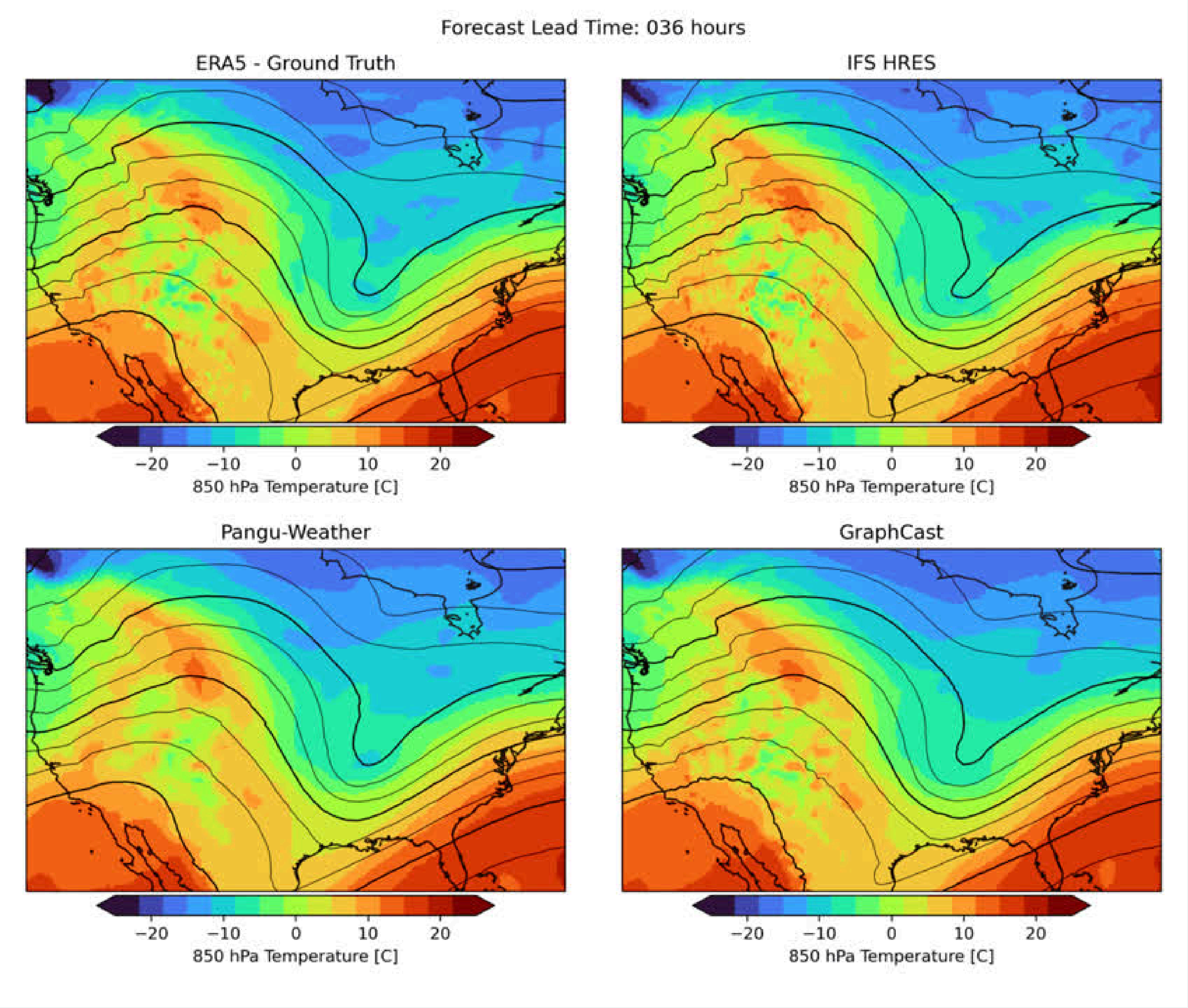
PRESENTED BY:

SALAM ALKAISSI
MAYRA SUAREZ

WeatherBench 2

What is WeatherBench 2.0?

- An advanced global weather forecasting benchmark designed to push the boundaries of data-driven weather modeling.
- Builds upon the foundation of WeatherBench 1.0 (introduced in 2020), which standardized the evaluation of AI-driven global weather models.
- Purpose: Provide a common framework for assessing machine learning models and comparing them to traditional numerical weather prediction (NWP)



WeatherBench 2

Model / Dataset	Source	Method	Type	Initial conditions	Horizontal resolution **
ERA5	ECMWF	Physics-based	Reanalysis		0.25°
IFS HRES	ECMWF	Physics-based	Forecast (deterministic)	Operational	0.1°
IFS ENS	ECMWF	Physics-based	Forecast (50 member ensemble)	Operational	0.2° *
Pangu-Weather (operational)	Huawei	ML-based	Forecast (deterministic)	Operational IFS	0.25°
GraphCast (operational)	Google DeepMind	ML-based	Forecast (deterministic)	Operational IFS	0.25°
ERA5 forecasts	ECMWF	Physics-based	Hindcast (deterministic)	ERA5	0.25°
Keisler (2022)	Ryan Keisler	ML-based	Forecast (deterministic)	ERA5	1°
Pangu-Weather	Huawei	ML-based	Forecast (deterministic)	ERA5	0.25°
GraphCast	Google DeepMind	ML-based	Forecast (deterministic)	ERA5	0.25°
FuXi	Fudan University, Shanghai	ML-based	Forecast (deterministic)	ERA5	0.25°
SphericalCNN	Google Research	ML-based	Forecast (deterministic)	ERA5	1.4x0.7°
NeuralGCM 0.7°	Google Research	Hybrid	Forecast (deterministic)	ERA5	0.7°
NeuralGCM ENS	Google Research	Hybrid	Forecast (Ensemble)	ERA5	1.4°

PROJECT OBJECTIVE

The purpose of this project is to build a new advanced data-driven model to enhance the accuracy and reliability of weather forecasts, particularly for high-impact extreme weather events

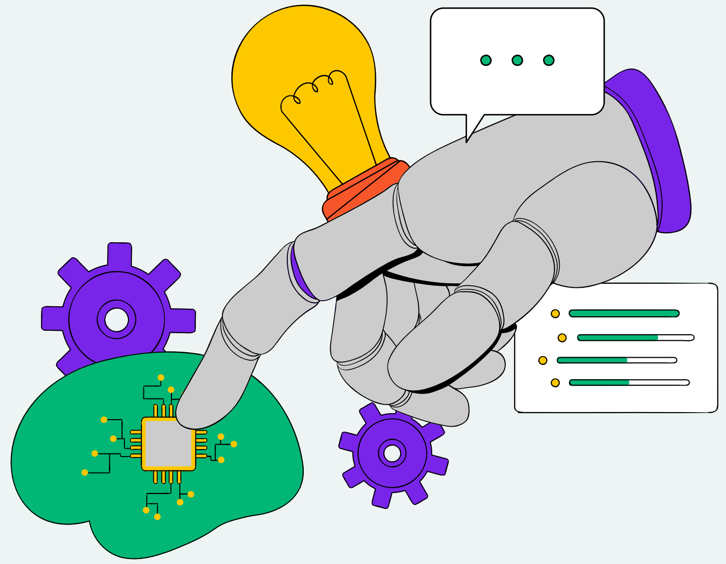
- Try to build more timely and accurate predictions of extreme weather events to mitigate damage and protect lives.
- Try to enhance regional and high-resolution forecasts to provide valuable insights for agriculture, disaster response, and urban planning.

CHALLENGES

- Achieving high-resolution forecasts over smaller regions requires large amounts of data and significant computational resources. The model must be able to generalize from global data but provide accurate localized predictions.

- **Short-Term Probabilistic Forecasting:** Create a 1-3 day probabilistic forecasting model to predict possible weather outcomes, addressing uncertainties and supporting better risk management and emergency planning.

- Many current models are deterministic, so moving to probabilistic predictions will require ensemble techniques or models that can directly generate probability distributions

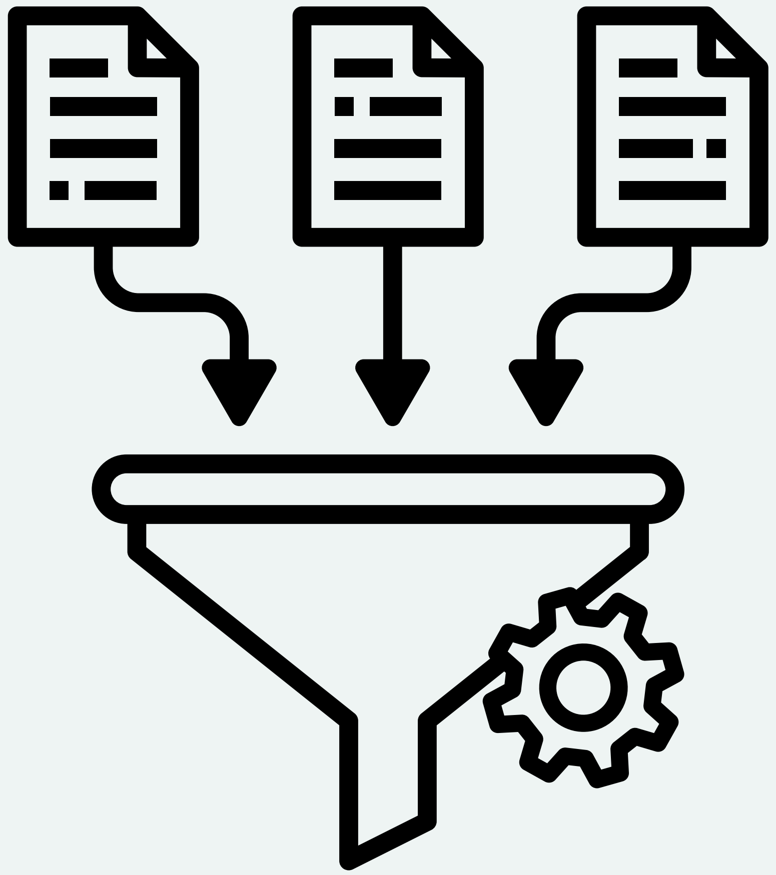


DATA PREPARATION

Data Access:

- Retrieved datasets from ECMWF and WeatherBench repositories in NetCDF(Network Common Data Form) format.
- Used Python libraries such as xarray ,netCDF4 and cfgrib were utilized for efficient data loading, subsetting, and manipulation, particularly for handling multi-dimensional arrays like time, latitude, and longitude.
- Additionally, the ECMWF OpenData client was used to programmatically retrieve forecast data, ensuring flexibility in accessing specific parameters over defined regions and time periods.

ERA5 data was processed to extract relevant features for modeling, while ECMWF served as the ground truth for evaluation. The datasets were resampled to align temporal and spatial resolutions, ensuring consistency across features and target variables. This streamlined workflow enabled seamless integration of both data sources for model training and robust evaluation.



DATA SOURCES

DATASETS :

- **ERA5 :** Offers global reanalysis data with a horizontal resolution of 0.25° making it suitable for high-resolution weather forecasting and climate analysis.
  For this study, we specifically focus on **2-meter temperature (t2m)**, which represents the air temperature at 2 meters above ground level.

DIMENSIONS:

– **time:** 2105 time points, which could represent timestamps for each weather forecast or measurement (likely hourly data, given the format datetime64[ns]).
– **latitude:** 36 latitude points, spanning a specific geographical area, each representing a latitude value from 51.05° to 42.3°.
– **longitude**: 59 longitude points, representing the longitudinal span from -5.1° to 9.4°.

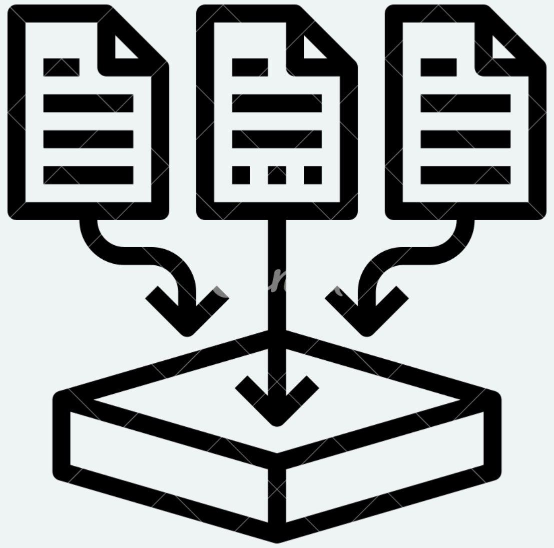
These three dimensions (time, latitude, longitude) allow the dataset to model weather or climate data across a spatial and temporal grid.

- **ECMWF : (GROUND TRUTH)**

These forecasts are generated from numerical weather prediction models and provide future temperature predictions at various time steps.
These forecasts are used for comparison against our model's predictions. The extracted ECMWF ground truth data provides us with the actual historical temperatures for the evaluation period and serves as the ground truth for model comparison.

DIMENSIONS:

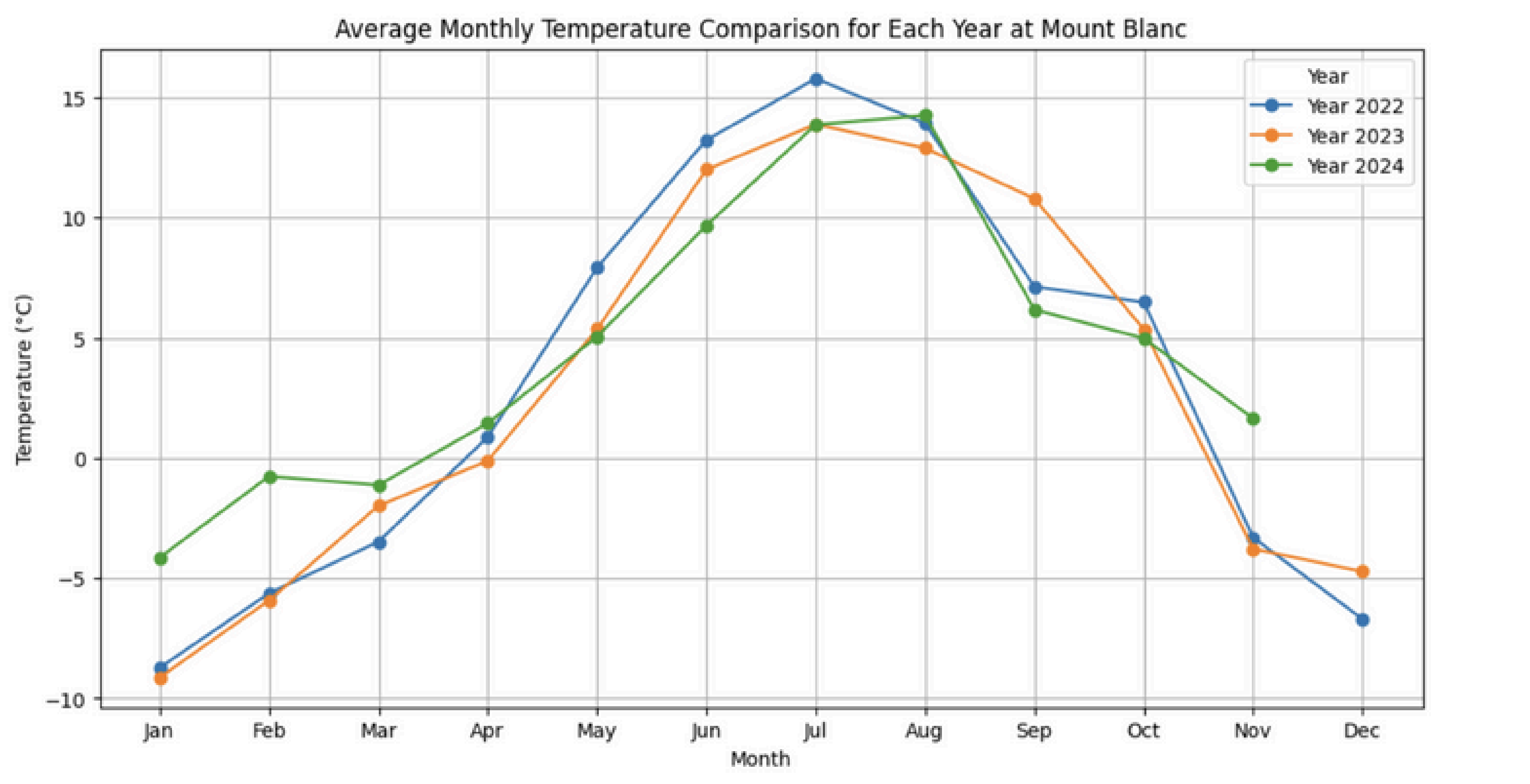
– **time:** 3 time points (typically for different dates or time intervals).
– **latitude:** 721 latitude points (from +90° to -90°, covering the entire globe).
– **longitude:** 1440 longitude points (covering the full 360° range from -180° to 180°).
Contains datetime values ranging from 2024-12-09 to 2024-12-11.
This 3 points are used for comparison



CHOSEN LOCATION: MONT BLANC

Mont Blanc, Alps (France/Italy border)
Coordinates: Approximately 45.832° N, 6.865° E.

- Mont Blanc, with its elevation of 4,809 meters (15,777 feet), is one of the highest peaks in Europe. It experiences extreme weather conditions that make it an interesting study area, especially in the context of weather forecasting or climate studies.



DATA PREPARATION

- **Handling Missing Values:**

  Imputed missing values with the mean across latitude, longitude, and time dimensions to eliminate gaps in the dataset.

- **Normalization:**

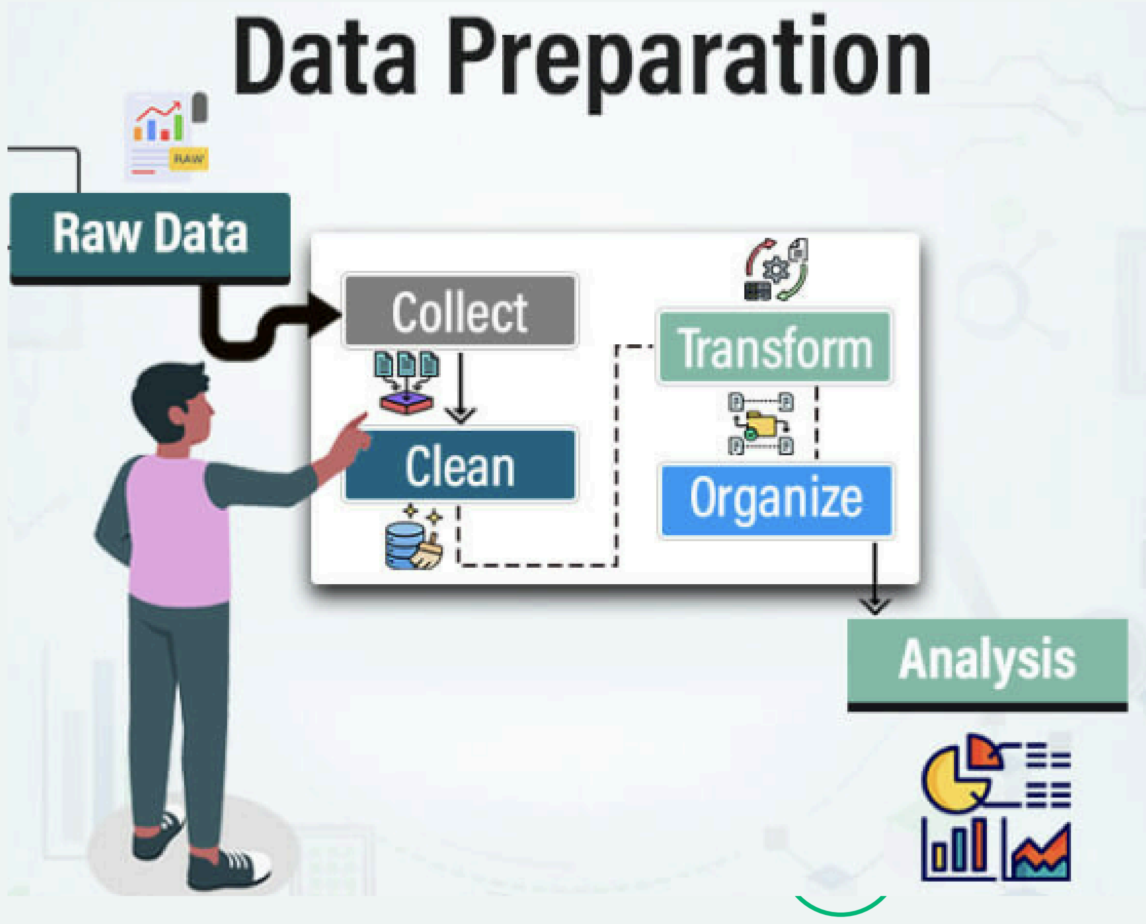
- Scaled temperature data using Min-Max normalization, transforming values into the range [0, 1].
- This preprocessing step ensured consistency and facilitated faster convergence during training.

- **Feature Engineering:**

- Applied a sliding window approach to generate sequences of historical temperature values.
- These sequences served as input features for time-series modeling, helping the model learn temporal patterns.

- **Train-Test Split:**

- As a final step before model training, the dataset was split into training (70%), validation (20%), and testing (10%) subsets.
- This split ensured proper evaluation of the model while preventing data leakage.



DEEP LEARNING MODELS

LSTM (Long Short-Term Memory)

- LSTM is a type of RNN designed to capture long-term dependencies in sequential data. It uses a memory cell and gating mechanisms to regulate information flow, enabling it to learn complex temporal patterns.
- **Strength:** Ideal for modeling long-term weather trends, such as seasonal patterns or extended forecasts, while effectively handling vanishing gradient issues in long time series.

GRU (Gated Recurrent Unit):

- GRU is a simplified RNN that processes sequential weather data, such as temperature by using gating mechanisms to manage the flow of information. It requires fewer parameters than LSTMs, making it faster and computationally efficient.
- **Strength**: Captures short- to medium-term temporal dependencies in weather patterns while mitigating vanishing gradient issues common in traditional RNNs.

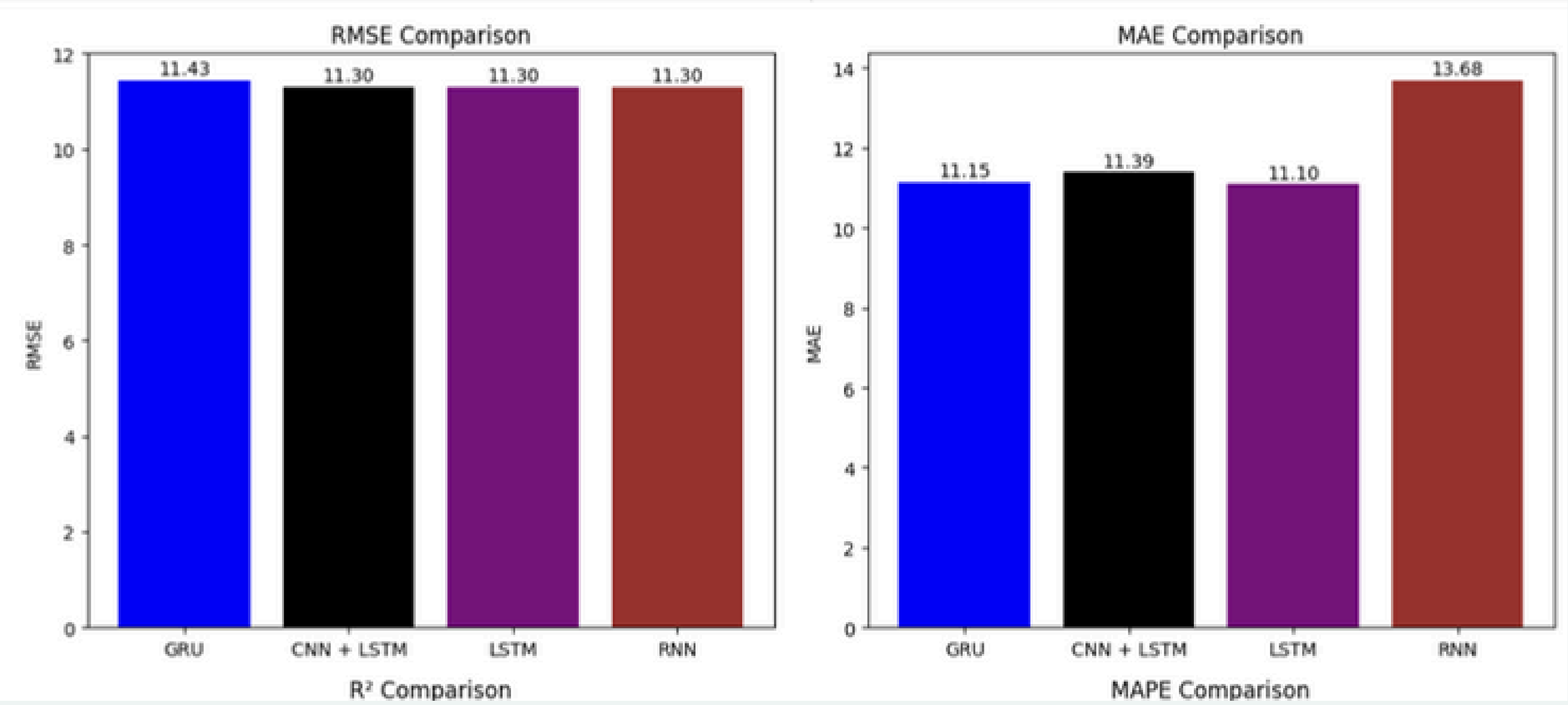
Hybrid LSTM + CNN: (Long Short-Term Memory & Convolutional Neural Networks)

- This architecture integrates LSTM which specialize in learning long-term temporal dependencies, with CNNs which excel in detecting spatial patterns. LSTM handles the evolving temporal dynamics of weather variables over time, while CNN captures localized spatial relationships like temperature gradients.
- **Strength**: Combines the strengths of both models to process spatiotemporal weather data, enabling accurate and holistic predictions.

EVALUATION METRICS & RESULTS

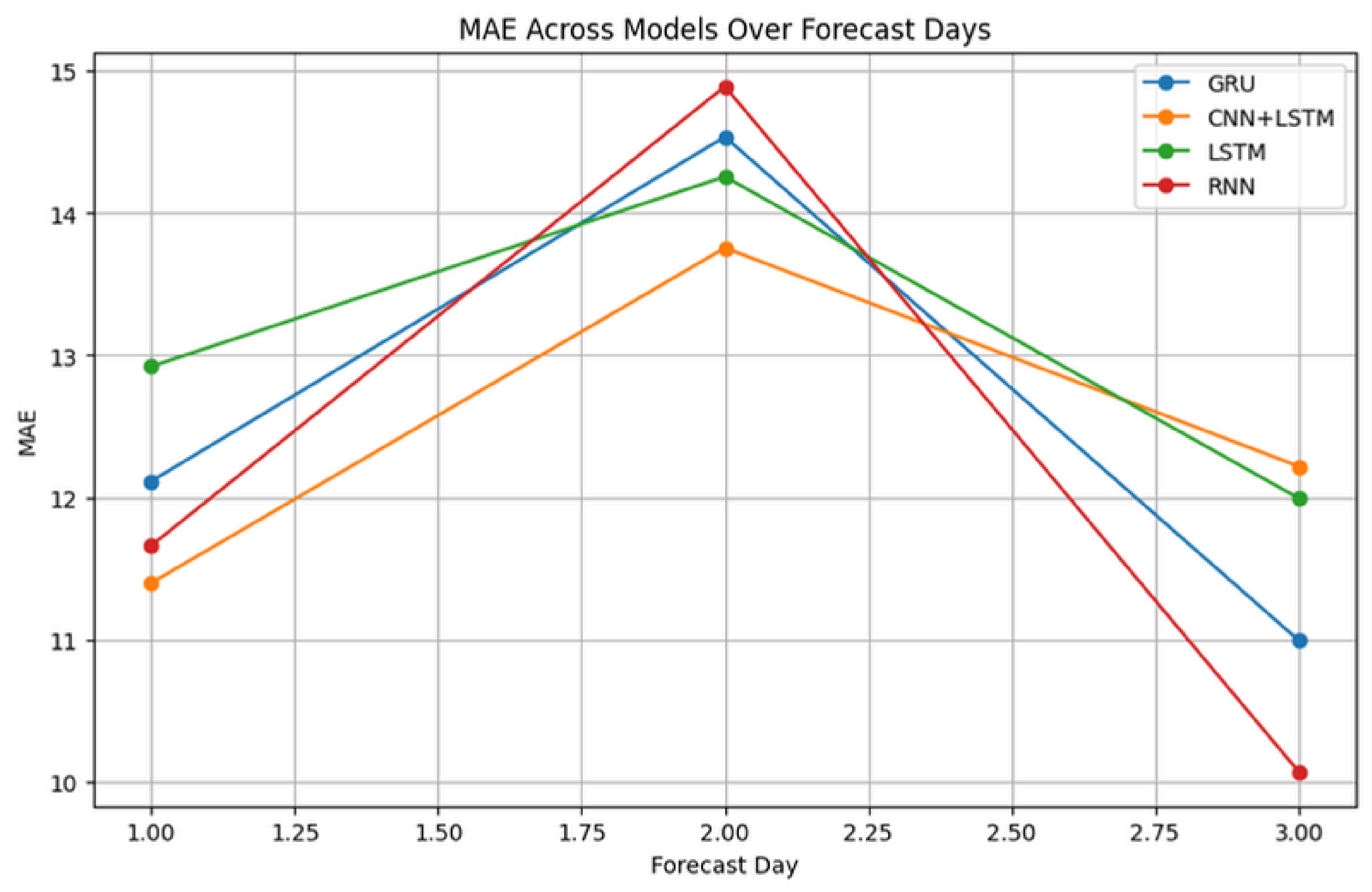
The chosen model is CNN+LSTM cause with 11.304 units has the lowest RMSE (Root Mean Squared Error) and also the lowest MAE (Mean Absolute Error) with 129.2 units, both were the metrics chosen for evaluation. This indicates significant enhancements in predictive accuracy and error minimization.

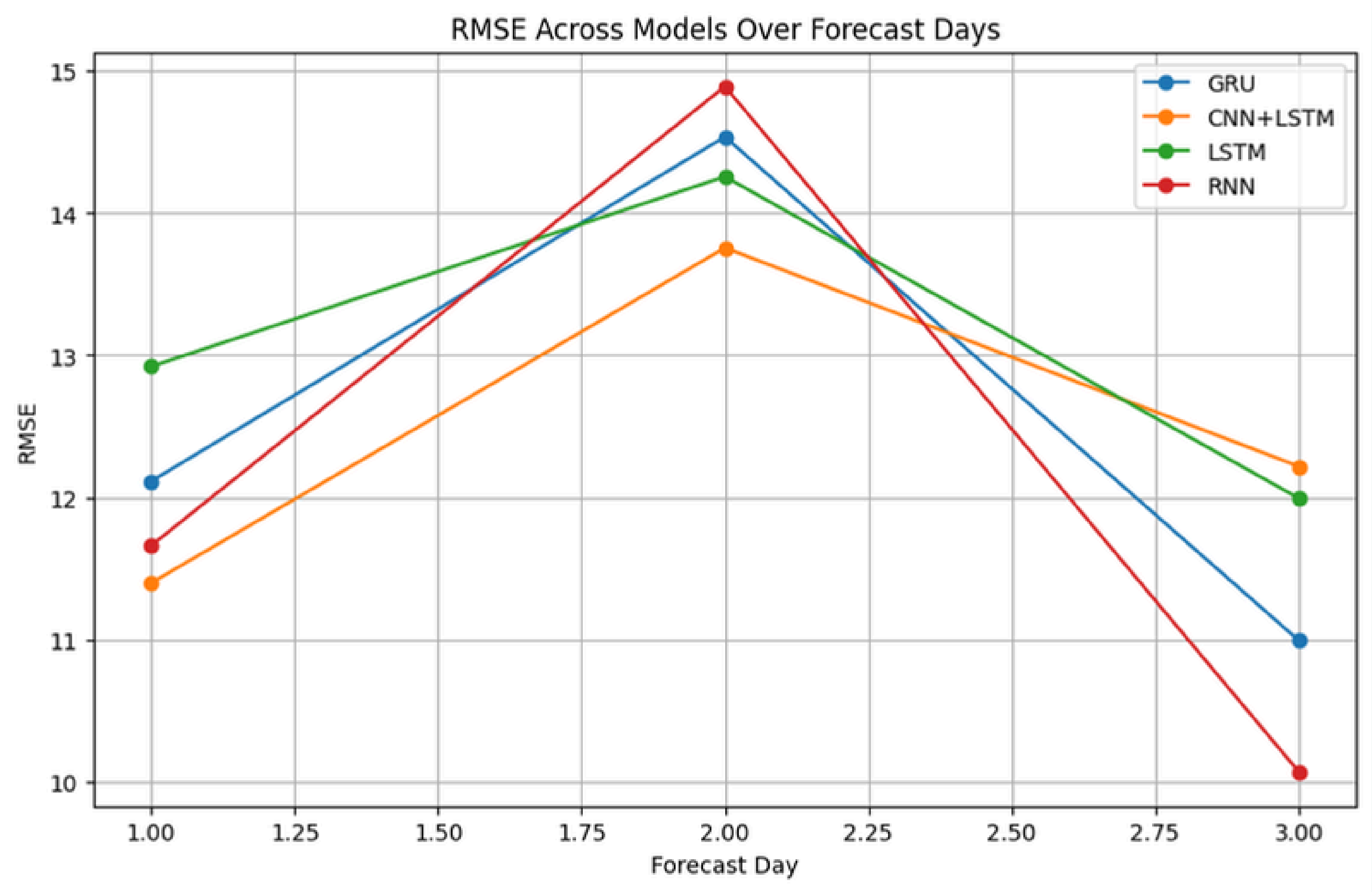
	LSTM	GRU	CNN+LSTM	RNN
RMSE	11.30	11.43	11.30	11.30
MAE	129.2	130.63	139.91	197.94
OVERALL	70.2925	71.03	75.605	104.62



NORMALIZED METRIC= METRIC VALUE / MAX METRIC VALUE

EVALUATION METRICS & RESULTS





CONCLUSIONS

The project successfully achieved the following objectives:

• Developed and evaluated predictive models using RMSE and MAE as primary metrics.
• Improved model performance through optimization, resulting in a notable reduction in error metrics.
• Validated the robustness of the model against ECMWF ground truth data.