1. Introduction

The objective of this project is to predict **Wind_Velocity_Avg** using deep learning techniques. A time-series forecasting approach with an optimized **LSTM (Long Short-Term Memory)** model is implemented to capture temporal dependencies in wind velocity data.

2. Dataset Overview

The dataset consists of daily wind velocity observations with the following attributes:

- **Observation_Date**: Date of observation (dtype: datetime).
- **Wind_Velocity_Avg**: Target variable (average wind speed per day, dtype: float64).
- **Wind_Direction_Avg**: Average wind direction in degrees (dtype: float64).
- **Wind_Velocity_Max**: Maximum recorded wind speed (dtype: float64).
- **Wind_Velocity_Min**: Minimum recorded wind speed (dtype: float64).

Dataset Summary:

- Total Records: 3,652
- No missing values
- Wind velocity values range from **1.38** to **48.73**
- Data is structured as a time series, with observations recorded daily.

3. Model Architecture

To forecast wind velocity, an **LSTM-based deep learning model** is employed. The model is optimized to improve accuracy by:

- **Increasing the number of LSTM layers** (3 layers with 100, 75, and 50 neurons, respectively).
- **Adding Batch Normalization layers** to stabilize training.
- **Applying Dropout layers** (30%) to prevent overfitting.
- **Using the Adam optimizer with a learning rate of 0.001**.

Model Structure:

1. **LSTM Layer (100 neurons, return sequences=True)**
2. **Batch Normalization**
3. **Dropout (30%)**
4. **LSTM Layer (75 neurons, return sequences=True)**
5. **Batch Normalization**
6. **Dropout (30%)**
7. **LSTM Layer (50 neurons, return sequences=False)**
8. **Batch Normalization**

9. **Dropout (30%)**
10. **Dense Layer (50 neurons, ReLU activation)**
11. **Dense Layer (25 neurons, ReLU activation)**
12. **Dense Output Layer (1 neuron)**

Hyperparameters:

- **Sequence Length:** 10 days of past data
- **Epochs:** 100
- **Batch Size:** 32
- **Loss Function:** Mean Squared Error (MSE)
- **Optimization Method:** Adam optimizer (learning rate = 0.001)

4. Results and Performance

- **Validation MSE:** `0.015558892522007227`
- **Validation RMSE:** `0.12485561749505815`
- **Training Loss vs. Validation Loss Plot** (Figure 1): Shows model performance over epochs.
- **Actual vs. Predicted Wind Velocity Plot** (Figure 2): Highlights how well the model aligns with real wind velocity data.
- **Scatter Plot of Actual vs. Predicted Values** (Figure 3): Illustrates model predictions compared to actual values.

5. Conclusion

The optimized LSTM model successfully predicts wind velocity with improved performance. Future improvements could include:

- Tuning hyperparameters further to reduce validation loss.
- Incorporating additional meteorological features.
- Using external datasets to enhance generalization.

6. Figures

Figure 1: Training Loss vs. Validation Loss Over Epochs *(Generated during training and saved as `loss_curve.png`)*

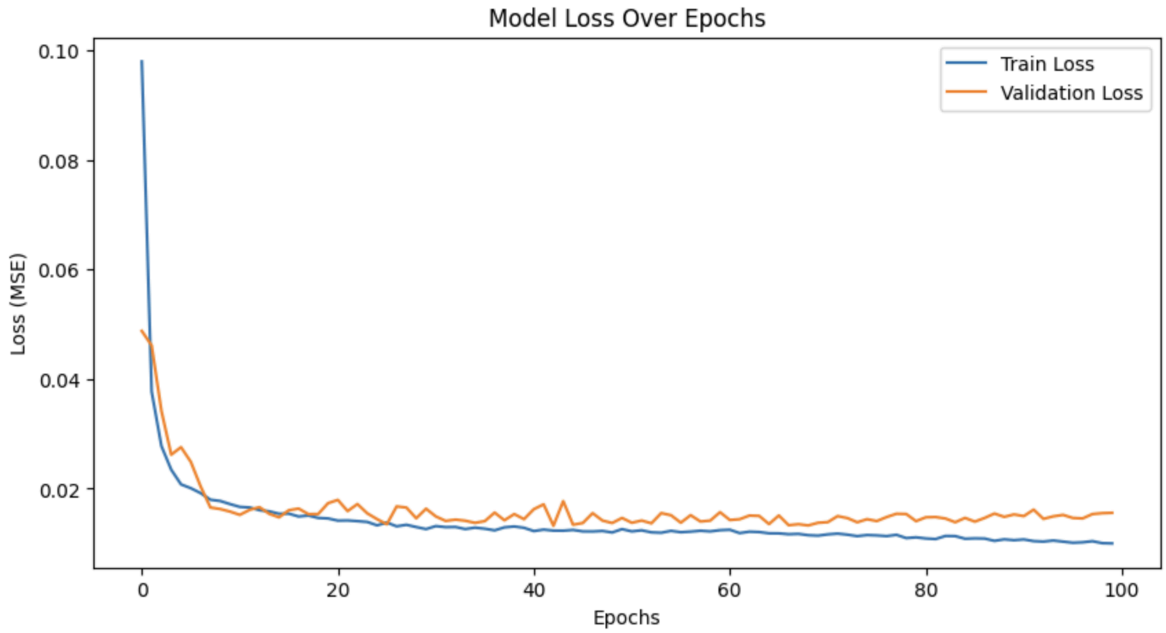


Figure 2: Actual vs. Predicted Wind Velocity *(Saved as `prediction_comparison.png`)*

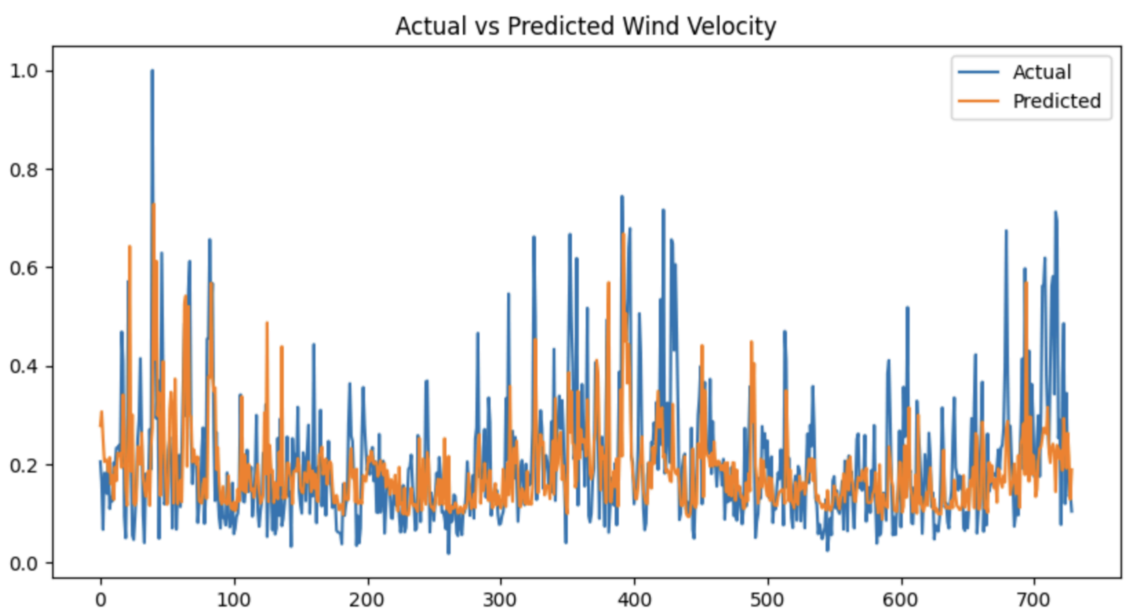
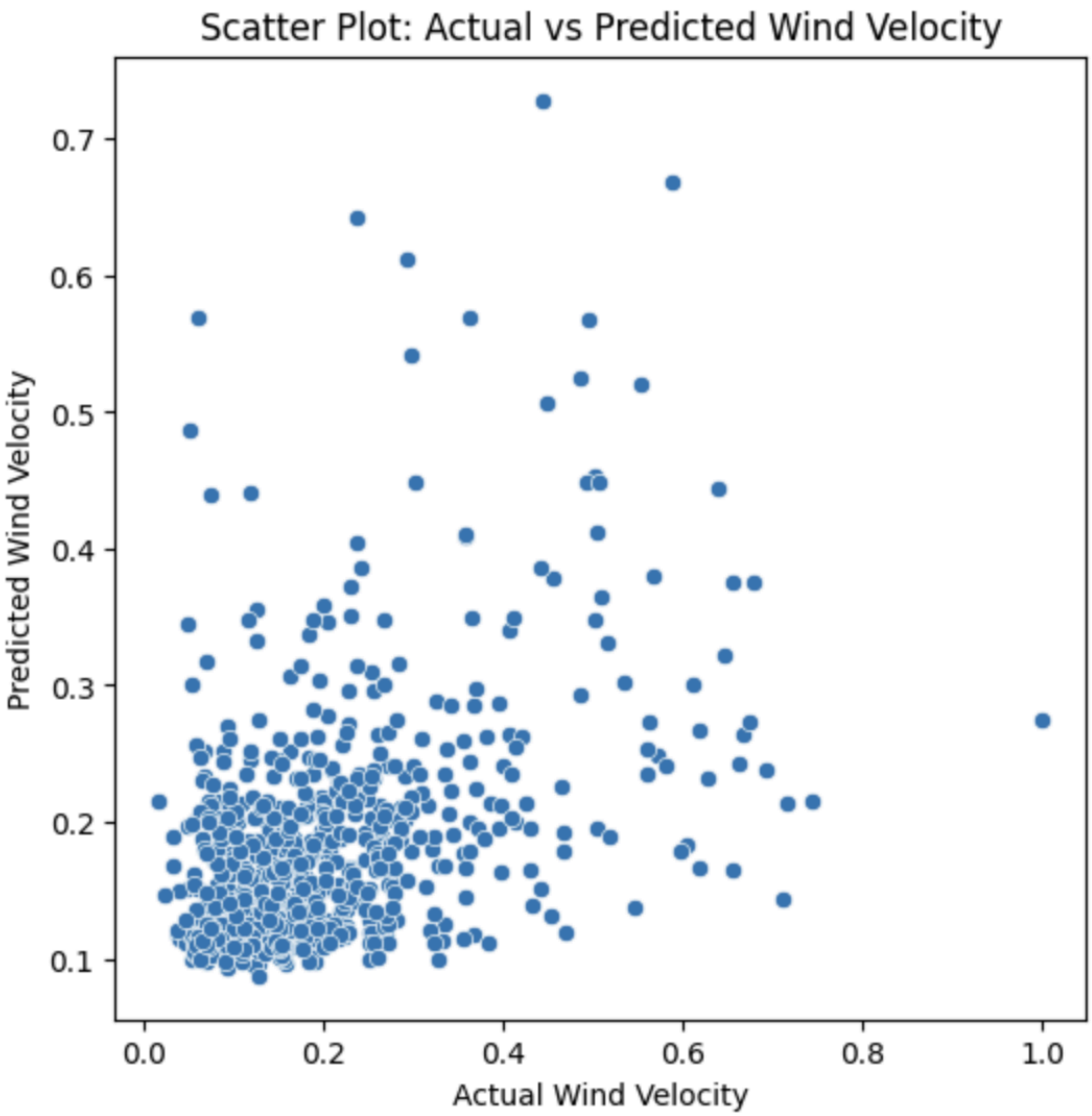


Figure 3: Scatter Plot - Actual vs. Predicted Wind Velocity *(Saved as `scatter_plot.png`)*


Scatter Plot: Actual vs Predicted Wind Velocity
Predicted Wind Velocity
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.2
0.4
0.6
0.8
1.0
Actual Wind Velocity