Next-Day Average Temperature Prediction in Rwanda Using Machine Learning and Deep Learning

GAHIGI Robert
Student ID: 101101

February 2026

Abstract

Accurate temperature forecasting supports agriculture, planning, and climate monitoring in Rwanda. This study predicts next-day average temperature using historical weather observations and compares three regression algorithms: Linear Regression, Random Forest Regressor, and a Multi-Layer Perceptron (MLP) Regressor (deep learning). Temporal feature engineering was applied using one-day and two-day lag temperatures, alongside meteorological predictors such as relative humidity, sunshine duration, and wind speed. Models were evaluated using MAE, RMSE, and R^2. Linear Regression achieved the best test performance (RMSE = 1.270, $R^2 = 0.708$). Random Forest performance improved after hyperparameter tuning (RMSE reduced from 1.384 to 1.294). TimeSeriesSplit cross-validation on the tuned model produced mean RMSE 1.340 (std 0.148), indicating stable generalization over time. As supporting analyses to meet midterm requirements, hot-day classification and K-means clustering were performed to characterize weather patterns.

1 Introduction

Weather variability affects food production, water management, and public safety in Rwanda. Reliable short-term temperature prediction helps decision-making for agriculture and operations. Classical statistical forecasting can struggle when relationships are nonlinear or when multiple interacting weather variables contribute to outcomes. Machine learning methods offer flexible alternatives that can capture complex patterns.

This project focuses on next-day average temperature prediction in Rwanda and applying multiple algorithms and comparing their performance, including at least one deep learning approach.

2 Methods

2.1 Dataset and Variables

The dataset contains daily weather observations from Rwanda, including temperature and supporting meteorological variables. The analysis uses the following predictors for the main regression task: one-day lag temperature (Temp_lag1), two-day lag temperature (Temp_lag2), relative humidity (RH), sunshine duration (SUNSHNE), and wind speed (Windspeed). The target variable is next-day average temperature (Temp_next_day).

2.2 Data Preparation and Feature Engineering

Data were cleaned by ensuring consistent formatting, handling missing values, and creating a chronological ordering required for time-aware modeling. Two lagged temperature features were created:

$$\text{Temp\_lag1} = \text{Temp}_{t-1}, \quad \text{Temp\_lag2} = \text{Temp}_{t-2}$$

The prediction target was defined as:

$$\text{Temp\_next\_day} = \text{Temp}_{t+1}$$

2.3 Modeling section

Primary task (Regression). Three regression algorithms were trained and evaluated:

- Linear Regression
- Random Forest Regressor
- MLP Regressor (deep learning)

Evaluation metrics. Models were compared using MAE, RMSE, and R^2:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i|, \quad \text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Model improvement. Randomized hyperparameter tuning was applied to Random Forest and compared against the baseline model.

Robust validation. Since the data are temporal, TimeSeriesSplit cross-validation was used to estimate generalization to future periods.

Supporting tasks. Two supporting analyses were included:

- Classification: predicting *Hot_tomorrow* (binary hot-day indicator)

Table 1: Regression model performance for next-day temperature prediction.

Model	MAE	RMSE	R2
Linear Regression	0.969747	1.269627	0.707571
MLP Regressor	0.961565	1.273104	0.705967
Random Forest Regressor	1.053918	1.383700	0.652662

Table 2: Random Forest performance before and after hyperparameter tuning.

Model	MAE	RMSE	R2
Random Forest (Before)	1.053918	1.383700	0.652662
Random Forest (After)	0.975784	1.294062	0.696207

- Clustering: K-means clustering of weather patterns and silhouette evaluation

3 Results and Interpretation

3.1 Regression Results

Table 1 summarizes test-set performance. Linear Regression achieved the best RMSE and R^2, suggesting that lag temperature features explain much of the next-day temperature variation. The MLP Regressor performed similarly to Linear Regression, while the Random Forest baseline was weaker before tuning.

3.2 Model Improvement (Random Forest Tuning)

After tuning, Random Forest improved meaningfully (Table 2). RMSE decreased from 1.384 to 1.294, confirming improved performance.

3.3 Time-Series Cross-Validation Summary

TimeSeriesSplit RMSE scores across five folds were:

$$[1.55644724, 1.2630835, 1.42467043, 1.11896543, 1.33914839]$$

The mean RMSE was 1.340 with standard deviation 0.148, indicating relatively consistent performance across time-based splits.

3.4 Feature Importance

Feature importance from the tuned Random Forest (Table 3) shows Temp_lag1 and Temp_lag2 as dominant predictors. This confirms strong temporal dependency in Rwanda

Table 3: Feature importance from the tuned Random Forest model.

Feature	Importance
Temp$_{l}ag1$	0.491142
Temp$_{l}ag2$	0.357236
RH	0.101174
SUNSHNE	0.047325
Windspeed	0.003124

Table 4: Classification model performance (Hot_tomorrow).

Model	Accuracy	F1	ROC$_{A}UC$
MLP Classifier	0.841477	0.860894	0.917089
Logistic Regression	0.839015	0.857095	0.915761
Random Forest	0.825568	0.846115	0.895498

temperature dynamics, while RH and SUNSHNE provide additional information.

3.5 Hot-Day Classification

Table 4 shows that the MLP Classifier achieved the best ROC-AUC (0.917), with Logistic Regression close behind. This indicates that the engineered features can also support a practical hot-day warning task.

3.6 Clustering

K-means clustering identified three clusters ($k = 3$) with silhouette score 0.317. This indicates moderate separation between weather regimes and can help summarize typical daily conditions.

4 Discussion and Recommendations

The results show that simple linear models can be highly competitive for next-day temperature prediction when lag features are included. Random Forest benefited from tuning, demonstrating that ensemble methods can improve with careful parameter selection. The feature importance analysis emphasizes the role of recent temperature history as the primary driver, while humidity and sunshine contribute additional signal.

Recommendations.

- Add more seasonal and calendar features (month, day-of-year) to capture seasonality.

- Consider time-series deep learning models (e.g., LSTM) if longer histories are available.
- Expand to multi-step forecasting (e.g., 3–7 days ahead) for operational planning.

5 Conclusion

This study applied three regression algorithms (including a deep learning model) to predict next-day average temperature in Rwanda and compared performance using standard metrics. Linear Regression achieved the strongest test performance (RMSE 1.270, R^2 0.708). Random Forest tuning improved RMSE from 1.384 to 1.294, and TimeSeriesSplit validation suggested stable generalization. Supporting classification and clustering analyses further demonstrated the dataset's usefulness for additional decision-support tasks.

6 References

- Scikit-learn documentation: `https://scikit-learn.org/`
- Géron, A. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow.* O'Reilly Media.