An Internet of Things Enabled Machine Learning Based Drone Mounted Weather Prediction System

| Devendra Parihar, Shreehitha Peddoju, Ajay Chaudhary and Kavitha Kadarla

OVERVIEW

- INTRODUCTION
- PROBLEM STATEMENT
- OBJECTIVES
- METHODOLOGY
- RESULT
- CONCLUSION

INTRODUCTION

The agriculture sector is vital for income and food security in developing economies, yet limited arable land demands innovative solutions. Technologies like IoT, machine learning, and cloud computing enhance agricultural production by optimizing resources such as water, fertilizers, and pest management.

IoT sensors gather environmental data, enabling real-time decision-making for better resource management. Machine learning models predict weather patterns, weeds, and pests, aiding precision agriculture and improving crop yields significantly.

Cloud computing supports large-scale data processing, while edge nodes, such as drone-mounted gateways, reduce latency and enhance IoT efficiency. This synergy between technologies ensures timely decisions and boosts agricultural productivity.

PROBLEM STATEMENT

Agriculture is heavily reliant on accurate weather forecasts, yet traditional prediction methods often fail to provide timely and localized insights. This limitation hinders farmers' ability to make informed decisions, impacting productivity and resource management.

The integration of IoT with machine learning holds promise, but existing systems face challenges like high latency and computational demands. IoT sensors typically rely on cloud-based processing, which delays real-time decision-making, reducing their practical utility in precision farming.

Furthermore, the lack of accessible, cost-effective, and scalable solutions for small and medium-scale farmers exacerbates the problem. A localized, drone-mounted system could address these gaps, enabling real-time, data-driven farming practices to enhance crop yields and resource efficiency.

OBJECTIVES

Primary Aim: Design an IoT-enabled, drone-mounted weather prediction system

Specific Goals:

- Utilize ML algorithms for accurate weather forecasting.
- Reduce latency in decision-making through edge computing.
- Enable precision agriculture practices to improve resource utilization and crop yield.

ABOUT DATASETS(AUSTRALIA)

- Key Question: Will it rain tomorrow?
- Goal: Predict next-day rain using classification models on the target variable `RainTomorrow`.
- Dataset:

  Size: 145,460 rows × 23 columns

  Coverage: ~10 years of daily weather observations from various Australian locations.
- Source:

  Data from the Australian Bureau of Meteorology (BoM)

  Accessible via: [BoM Climate Data](http://www.bom.gov.au/climate/data)
- Acknowledgments:

  Adapted from BoM's Climate Data Online.
  Copyright: Commonwealth of Australia, 2010

ABOUT DATASETS(BIKANER)

- Key Question: Will it rain tomorrow?
- Goal: Predict next-day rain using classification models on the target variable `Rain`.
- Dataset:

  Size: 14,611 rows × 28 columns

  Coverage: ~ daily weather observations of 1982 to 2021 .
- Source:

  Data from the NASA Power Project, it provide solar and meteorological datasets

  Accessible via: [NASA Power Project website](https://power.larc.nasa.gov)
- Acknowledgments:

  Adapted from : NASA Power Project

METHODOLOGY

System Design:

- IoT sensors collect environmental and soil data.
- Drone-mounted edge nodes process data locally to reduce latency.

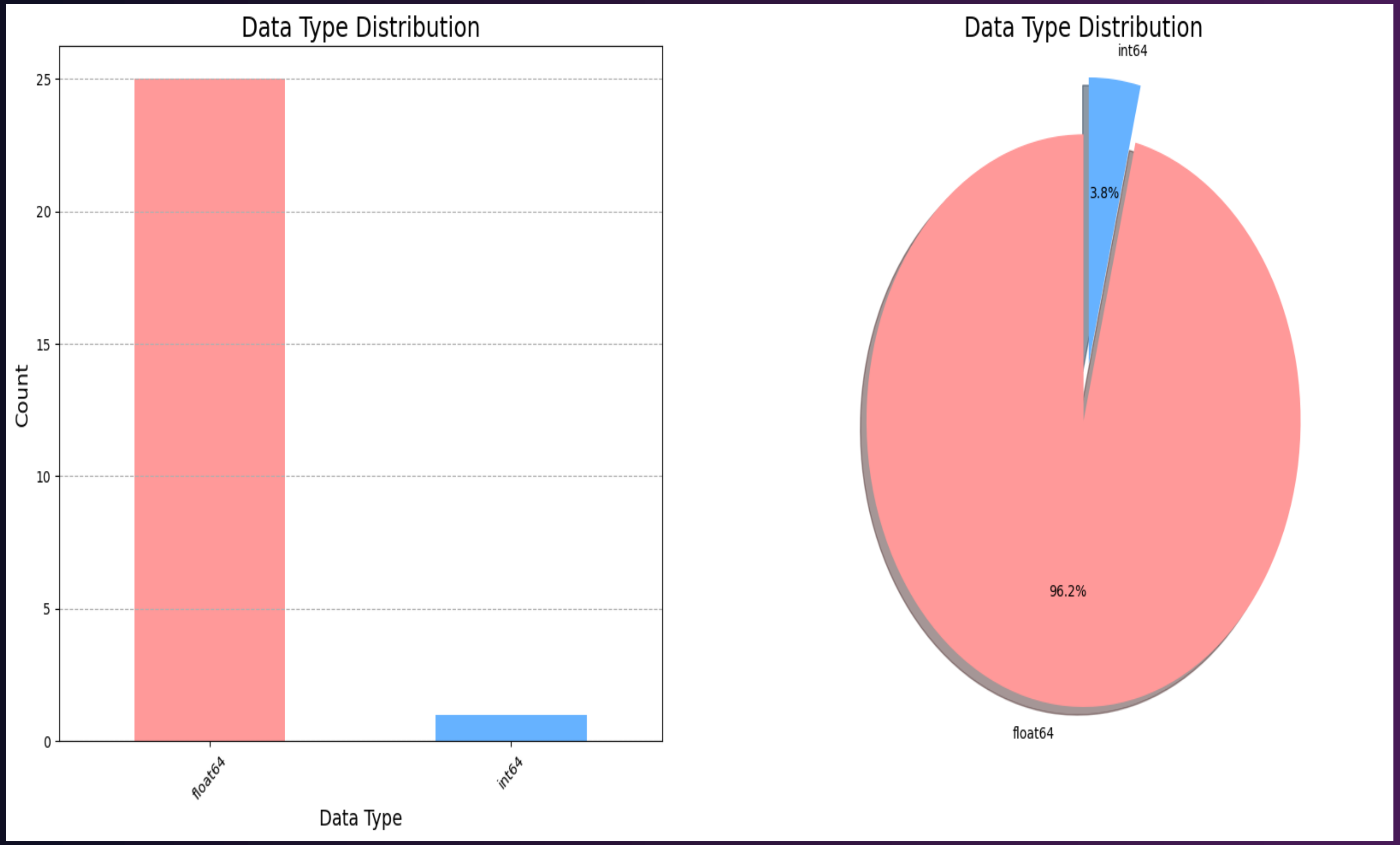
Machine Learning Algorithms:

- Tested models: Logistic Regression, Decision Tree, Random Forest, Gradient Boosting, XGBoost, LightGBM, CatBoost.
- Finalized models based on accuracy and efficiency.

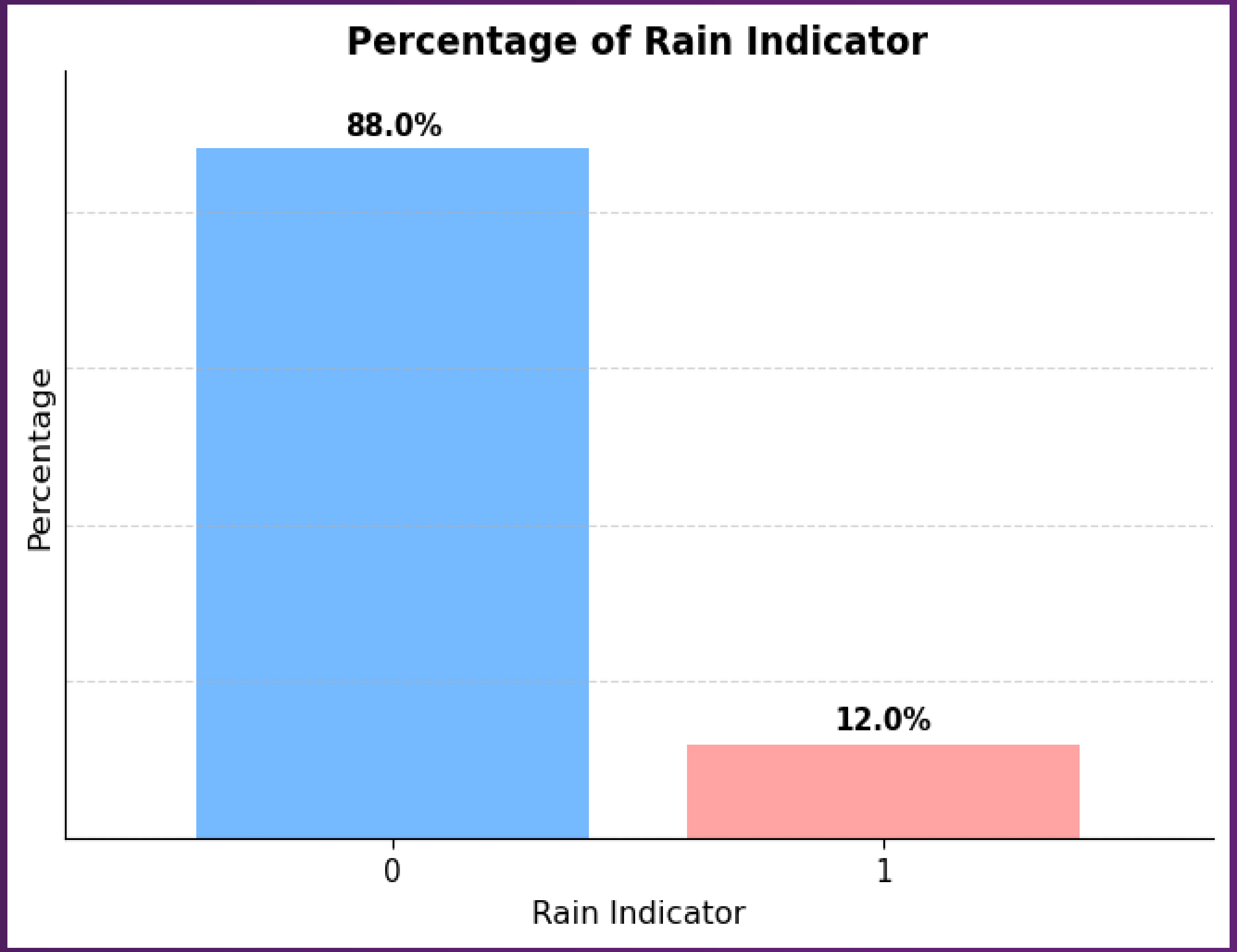
Data Preprocessing:

- Cleaning, integration, and transformation of datasets.
- Historical data training on high-computation facilities.

PERFORMED EDA ON DATA



Distribution of data type in dataset



Distribution of Rain in Dataset

MODELS FOR TRAINING

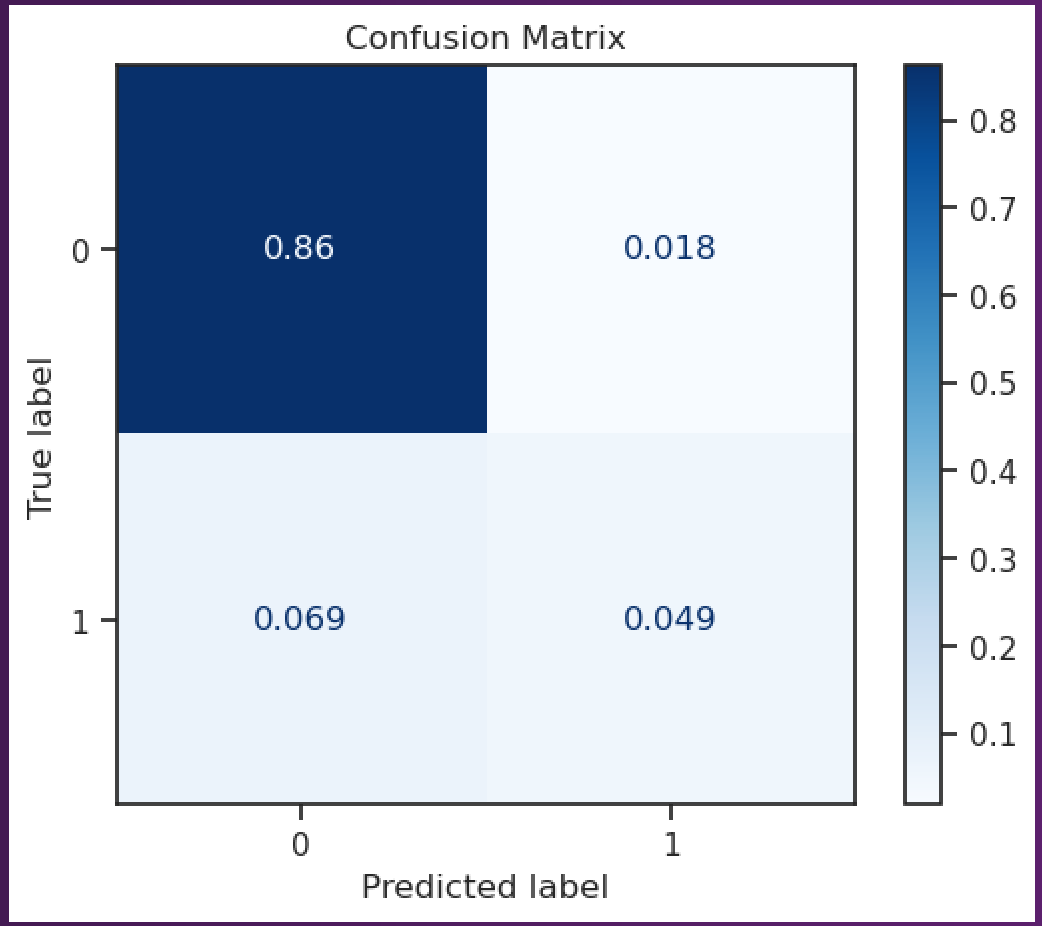
1. Logistic Regression: by predicting outcomes using probabilities.

   - Accuracy : 91.29%

   - Precision : 92.63%

   - Recall : 97.92%

   - F1 Score : 95.21



MODELS FOR TRAINING

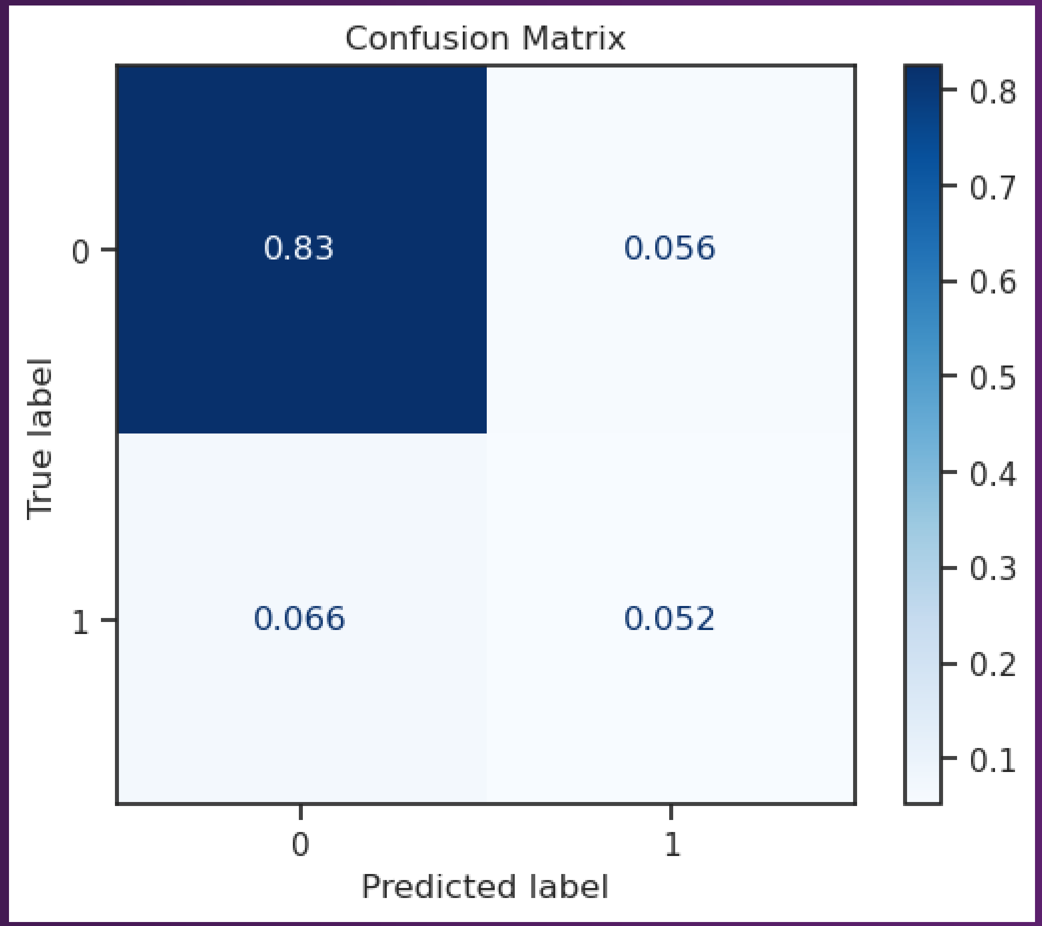
2. Decision Tree: Tree-based decision-making model..

    - Accuracy : 88.28%

    - Precision : 93.20%

    - Recall : 87.07%

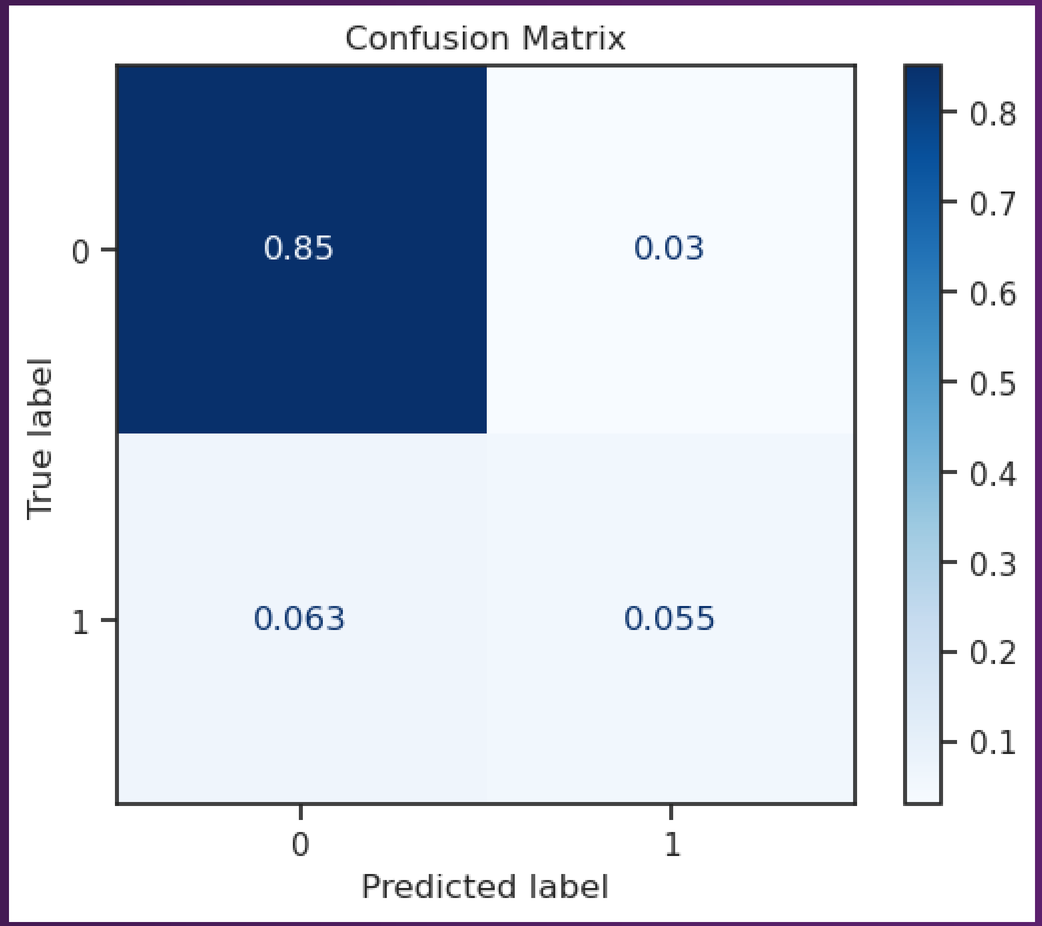
    - F1 Score : 87.07



MODELS FOR TRAINING

3. Random Forest: Ensemble of decision trees.

   - Accuracy : 90.77%
   - Precision : 93.15%
   - Recall : 96.65%
   - F1 Score : 94.87



MODELS FOR TRAINING

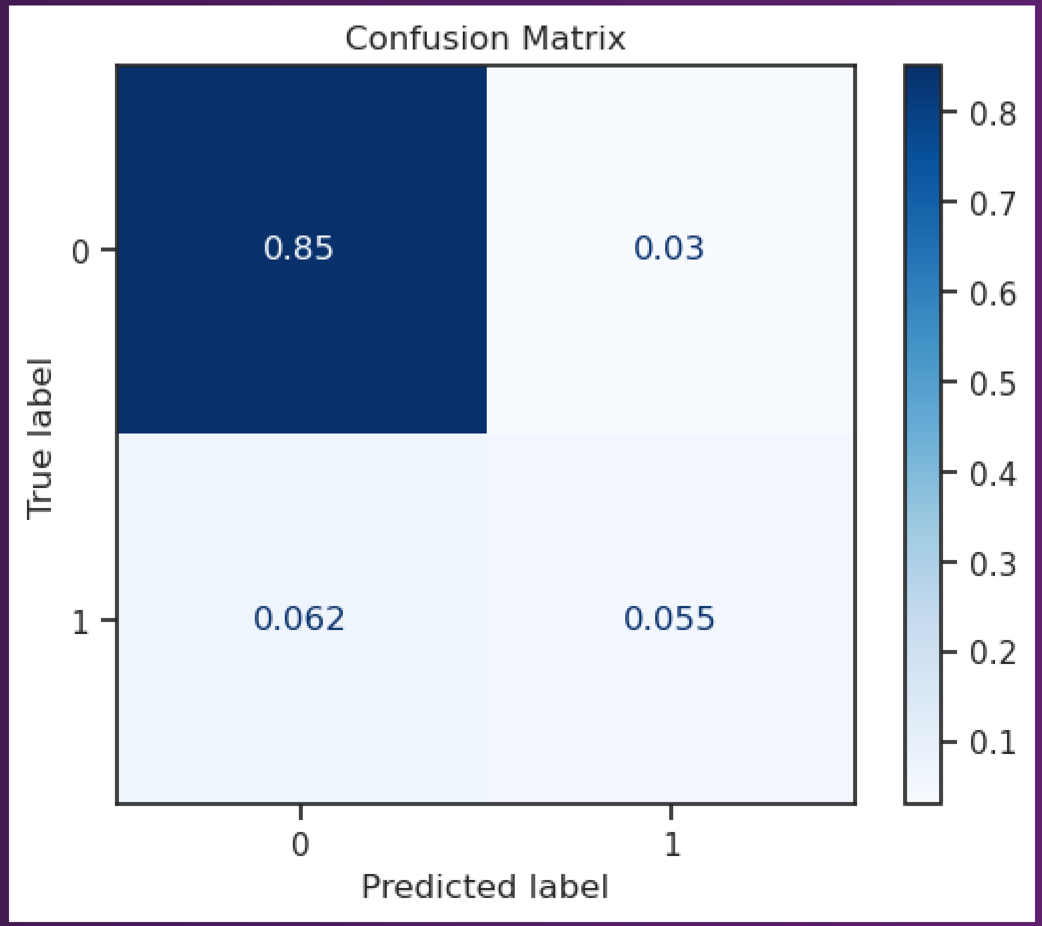
4. LGBM : Gradient boosting framework.

    - Accuracy : 90.83%

    - Precision : 93.21%

    - Recall : 96.65%

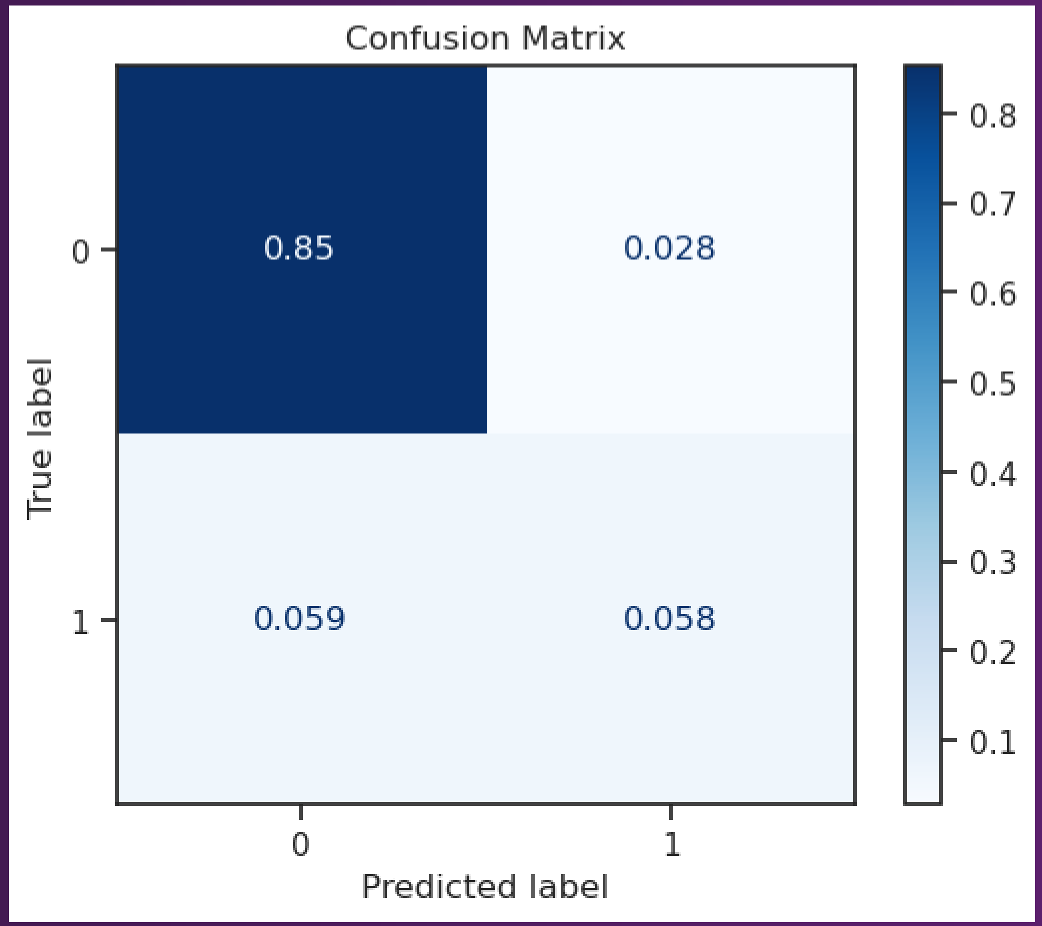
    - F1 Score : 94.40



MODELS FOR TRAINING

5. Cat Boost: Efficient, fast, gradient boosting, categorical features.

- Accuracy : 91.24%
- Precision : 93.50%
- Recall : 96.81%
- F1 Score : 95.12



MODELS FOR TRAINING

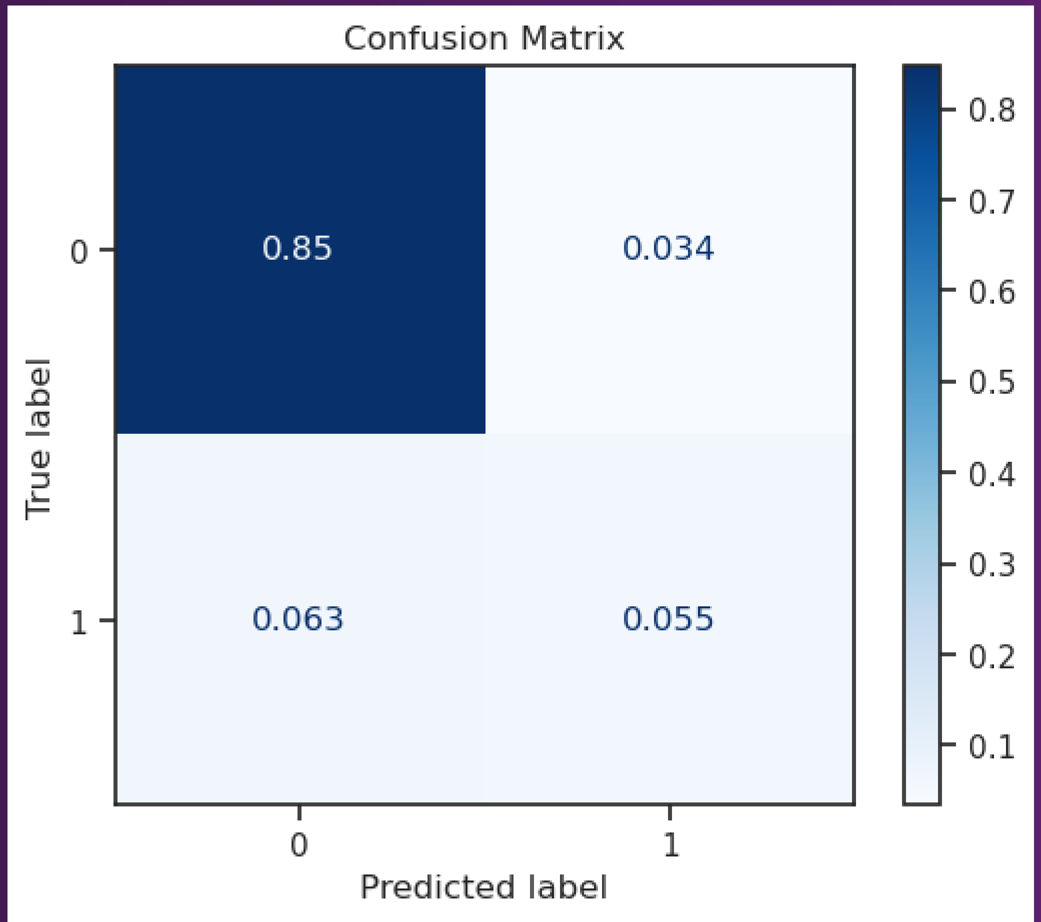
6. XGBOOST : High-performance, gradient boosting algorithm.

   - Accuracy : 90.31%

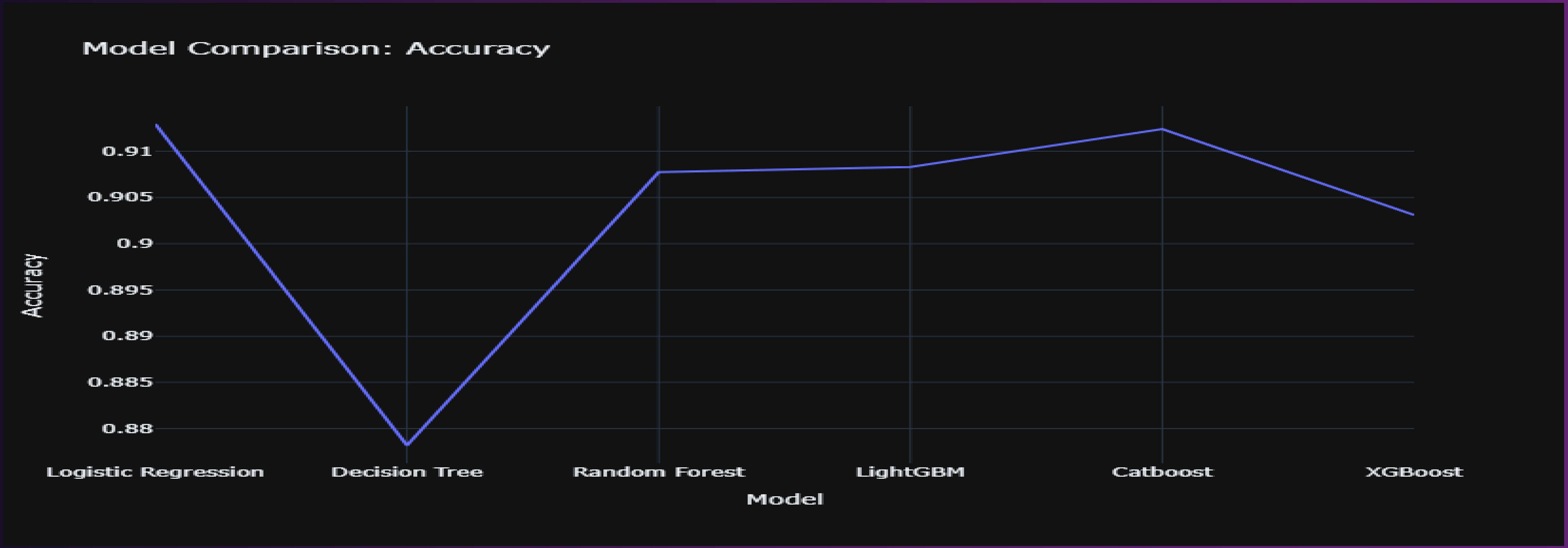
   - Precision : 93.12%

   - Recall : 96.12%

   - F1 Score : 94.60



MODEL COMPARISON



Results and Conclusion

1. Key Findings:

- Random Forest, XGBoost, and CatBoost models achieved the highest accuracy.
- XGBoost demonstrated a precision of 91.5% and recall of 97.61% on Australian rainfall data.
- Logistic Regression achieved the highest accuracy (91.29%) for the Bikaner dataset.

Measures	LR	DT	RF	LGBM	Catboost	XGB
Train Time(Sec)	1.93	0.474	20.38	1.627	129.1	49.78
Coheńs Kappa	0.582	0.720	0.728	0.497	0.877	0.898
Accuracy (%)	79.52	86.12	95.0	86.62	93.92	94.96
precision (%)	78.22	82.0	89.89	85.38	90.0	91.50
recall(%)	74.12	87.72	93.93	84.0	96.92	97.61
F1-score	76.13	84.77	91.85	84.69	93.35	94.46

TABLE I: Comparison of Various Machine Learning Models on Australian Rainfall Weather Data dataset

Measures	LR	DT	RF	LGBM	Catboost	XGB
Train Time(Sec)	0.10	0.02	0.91	0.10	3.81	0.65
Coheńs Kappa	0.48	0.42	0.49	0.49	0.52	0.47
Accuracy (%)	91.29	88.28	90.77	90.83	91.24	90.31
precision (%)	92.63	93.20	93.15	93.21	93.50	93.12
recall (%)	97.92	87.7	96.65	96.65	96.81	96.12
F1-score	95.21	87.7	94.87	94.90	95.12	94.60

TABLE II: Comparison of Various Machine Learning Models on Bikaner, India dataset

Results and Conclusion

Conclusion:

- The proposed system effectively predicts weather conditions, reducing dependency on traditional forecasting methods.
- Real-time decision-making supports efficient agricultural practices.

Future Work:

- Scaling the system to include more variables like pest prediction.
- Integration with AI-powered automation for complete farming cycles

IEEE InGARSS 2024 Goa

THANK YOU

Q & A

- Precision: The proportion of true positive predictions among all positive predictions made : (True Positives / (True Positives + False Positives)).
- Recall: The proportion of true positive predictions among all actual positive instances (True Positives / (True Positives + False Negatives)).
- F1-Score: The harmonic mean of precision and recall, providing a balance between the two (2 * (Precision * Recall) / (Precision + Recall)).
- Roc-Currve: it plots the True Positive Rate versus the False Positive Rate at various thresholds for a binary classifier.