# Rainfall Prediction Using Classification Algorithms

## PREDICTING TOMORROW'S RAIN WITH WEATHER DATA


**GROUP 15**

Adyanshi Singh (202401100300023)
Aditya Singh (202401100300023)
Abhishek Deewakar (202401100300006)
Achin Sahu (202401100300011)

# Rainfall Prediction Using Classification Algorithms

## Problem Statement:

Rainfall prediction is critical for agriculture, disaster management, and daily planning. This project aims to classify whether it will rain the next day based on today's weather conditions.

## OBJECTIVE

To build a machine learning model that predicts whether it will rain tomorrow using historical weather data.

# Challenges Tackled by the Model

### Uncertainty in Weather Patterns

By analyzing historical weather data, the model detects patterns to make accurate predictions.

### Decision-Making for Agriculture

Helps farmers plan better by providing timely rain predictions.

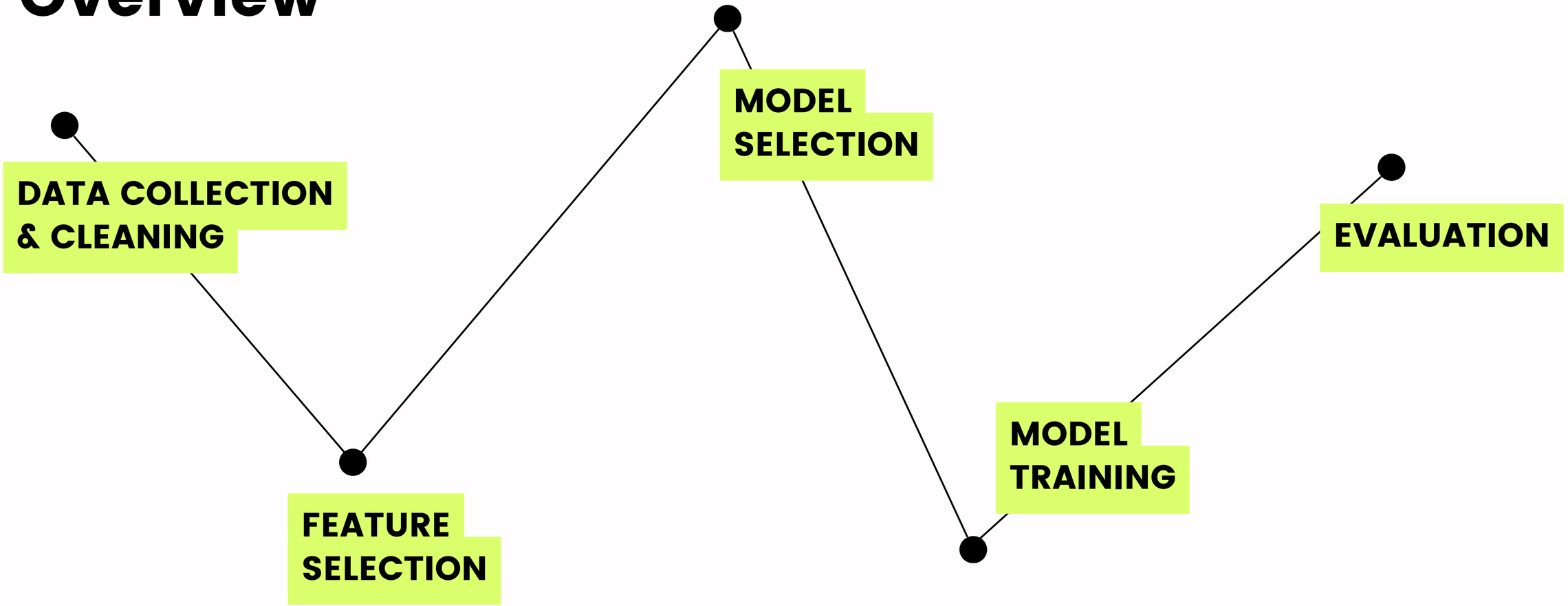
### Urban Flood Management & Public Safety

Enables early warnings to prevent floods and manage city infrastructure.

### Need for Real-Time Predictions

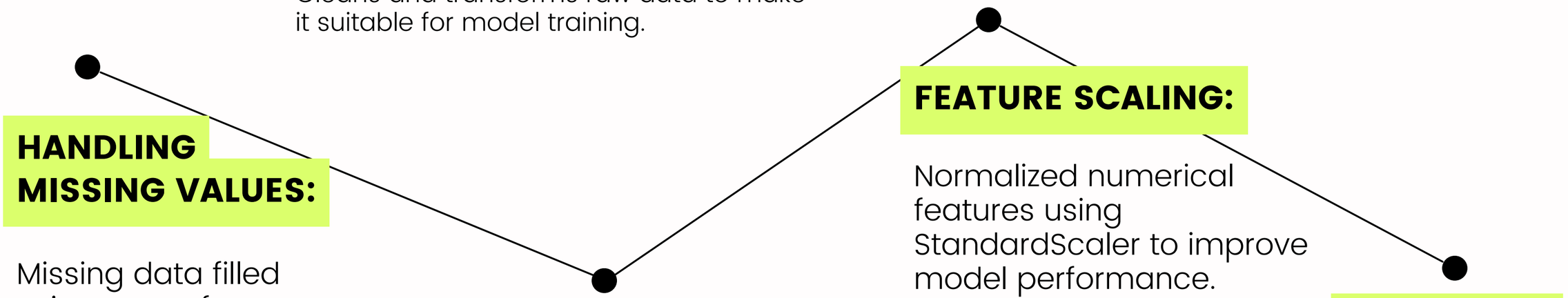
Delivers quick predictions using current weather inputs after training.

# Methodology Overview



# Data Preprocessing

Cleans and transforms raw data to make it suitable for model training.



**HANDLING MISSING VALUES:**

Missing data filled using mean for numerical and mode for categorical features.

**ENCODING CATEGORICAL VARIABLES:**

Converted text labels into numbers using Label Encoding for model compatibility.

**FEATURE SCALING:**

Normalized numerical features using StandardScaler to improve model performance.

**TRAIN-TEST SPLIT:**

Dataset split into 80% training and 20% testing sets for evaluation.

# Dataset Description

## Dataset Name:

Rain in Australia

## Source:

Kaggle

## Time Period:

2008–2017

## Records:

Approximately 145,460 daily observations

## Geographic Coverage:

49 weather stations across Australia

# Key Features in the Dataset:

**Date of Observation**

**Name of the weather station**

## Rainfall:

Amount of rain recorded today; key indicator of tomorrow's rain.

## RainToday:

Rain on the current day increases the chance of rain tomorrow.

## Humidity3pm:

Higher afternoon humidity often signals upcoming rainfall.

## Pressure3pm:

A drop in pressure suggests unstable, rainy weather.

## Temp3pm:

Afternoon temperature affects cloud and rain formation.

## WindGustSpeed:

Strong wind gusts may indicate stormy or rainy conditions.

# Algorithms Used:

## LOGISTIC REGRESSION:

- A simple linear model used for binary classification tasks.

## DECISION TREE:

- A tree-based model that splits data into branches based on feature values.

## RANDOM FOREST:

- An ensemble of decision trees that improves accuracy and reduces overfitting.

# Performance Matrix

| Model | Accuracy | Precision | Recall | F1 Score |
|---|---|---|---|---|
| Logistic Regression | 82% | 79% | 76% | 77% |
| Decision Tree | 84% | 81% | 79% | 80% |
| Random Forest | 87% | 85% | 83% | 84% |

# References


- Kaggle Dataset: Rain in Australia
- Scikit-learn Documentation: Geographic Coverage:
- Python Data Science Handbook by Jake VanderPlas
- Research papers on Rainfall Prediction using ML

# THANK YOU!