

Model Development Phase Template

Date	24 April 2024
Team ID	Team-738169
Project Title	Rainfall Prediction Using Machine Learning
Maximum Marks	5' Marks

Feature Selection Report Template

In the forthcoming update, each feature will be accompanied by a brief description. Users will indicate whether it's selected or not, providing reasoning for their decision. This process will streamline decision-making and enhance transparency in feature selection.

Feature	Description	Selected (Yes/No)	Reasoning
Date	serves as a timestamp capturing the temporal aspect of the dataset.	Yes	<ul style="list-style-type: none"> Provides a chronological framework for the rainfall prediction dataset. Facilitates time-based analysis, trend identification, and pattern recognition.

Location	It facilitates the identification of spatial patterns, trends, and correlations in rainfall data, which are essential for model development, validation, and interpretation.	Yes	The primary purpose of the "Location" column is to provide spatial context to the rainfall data. By associating each data point with a specific location, it enables the analysis of rainfall patterns across different geographical regions
MinTemp	The "MinTemp" column likely refers to the minimum temperature recorded for a given period, such as a day or an hour.	Yes	Lower temperatures can lead to the condensation of moisture in the atmosphere, potentially resulting in rainfall. Thus, tracking the minimum temperature can provide insights into the likelihood of precipitation.
MaxTemp	This temperature value represents the highest recorded temperature during that period.	Yes	High temperatures can indicate the presence of warm air masses, which often contribute to the formation of convective clouds and thunderstorms, both of which are associated with rainfall.
Rainfall	The Rainfall column in the database records the amount of precipitation, typically in millimeters or inches, that occurs within a specific time frame at a particular location.	Yes	The Rainfall column provides the essential ground truth data necessary for training rainfall prediction models. By capturing historical rainfall measurements, the model can learn patterns, trends, and relationships between various meteorological variables and rainfall occurrence.
Evaporation	Contains data related to the rate at which water changes from a liquid to a vapor	Yes	Evaporation is a key component of the hydrological cycle, where water evaporates from the Earth's surface into the atmosphere, eventually forming clouds and leading to precipitation

	state in a given area over a specific period, usually measured in millimeters per day or some equivalent unit.		events like rainfall.
Sunshine	Records the duration or amount of sunshine received at a particular location during a specific time interval.	Yes	Sunshine directly affects the amount of solar radiation received by the Earth's surface, which, in turn, influences atmospheric temperature and pressure gradients. These factors are significant drivers of weather and precipitation patterns.
WindGustDir	records the direction from which the maximum wind gust was observed during a given time interval. It typically contains cardinal or ordinal directions	Yes	Wind direction plays a crucial role in influencing weather patterns, including rainfall. Different wind directions can bring varying moisture levels and atmospheric conditions, which can directly impact rainfall patterns in a particular region.
WindGustSpeed	records the maximum wind speed observed during a specific time period at a given location. This speed is typically measured in units such as kilometers per hour (km/h) or miles per hour (mph).	Yes	Wind speed is a crucial factor affecting rainfall patterns. Higher wind speeds can impact the movement and distribution of clouds, moisture, and atmospheric pressure systems, which in turn influence rainfall intensity and distribution
WindDir9am	denotes the direction from which the wind is blowing at 9 AM. It is typically measured in	Yes	Wind direction at a specific time, such as 9 AM, provides valuable information about the prevailing atmospheric conditions during the morning hours. Changes in wind direction over time can indicate shifts in weather patterns, such

	degrees clockwise from true north.		as the approach of a storm system or the onset of a frontal boundary, which may affect rainfall.
WindDir3pm	represents the direction of the wind at 3:00 PM local time. This attribute records the compass direction from which the wind is blowing at that specific time of the day.	Yes	The wind direction at 3:00 PM is an important meteorological parameter that can significantly influence weather patterns and, consequently, rainfall prediction. For instance, winds blowing from the ocean onto the land may carry moisture, potentially leading to increased rainfall in coastal areas.
WindSpeed9am	records the speed of the wind at 9 a.m. on each day. This measurement is typically expressed in units like kilometers per hour (km/h) or meters per second (m/s), indicating how fast the air is moving at that specific time.	Yes	Wind speed plays a crucial role in shaping weather patterns. It affects the movement of air masses, the formation of clouds, and the distribution of moisture in the atmosphere.
WindSpeed3pm	records the speed of the wind measured at 3 PM local time. It typically represents the afternoon conditions when atmospheric dynamics may have significant effects on weather patterns, including the formation and movement of clouds and	Yes	Wind speed can be correlated with rainfall patterns. Higher wind speeds may indicate the presence of weather systems capable of generating precipitation, such as storms or frontal boundaries. By incorporating WindSpeed3pm into the model, you can account for the influence of local and regional wind patterns on rainfall distribution.

	precipitation.		
Humidity9 am	captures the relative humidity level at 9 AM, typically measured as a percentage. This metric indicates the amount of moisture present in the air relative to the maximum amount the air can hold at that temperature.	Yes	Humidity at 9 AM serves as an early indicator of atmospheric moisture content. It provides insight into the initial conditions of the atmosphere, which can significantly influence weather patterns throughout the day.
Humidity3 pm	records the relative humidity level at 3:00 PM local time. Relative humidity refers to the amount of water vapor present in the air compared to the maximum amount the air could hold at that temperature	Yes	Relative humidity is closely related to the formation of clouds and subsequent rainfall. Higher humidity levels indicate more moisture in the atmosphere, increasing the likelihood of precipitation.
Pressure9a m	contains data representing the atmospheric pressure measured at 9:00 AM at a particular location. Atmospheric pressure is typically measured in hectopascals (hPa) or millibars (mb) and is a key meteorological parameter that	Yes	Atmospheric pressure plays a significant role in shaping weather patterns. Changes in pressure can indicate the presence of high or low-pressure systems, which in turn influence the movement of air masses, cloud formation, and the likelihood of precipitation, including rainfall. Recording atmospheric pressure specifically at 9:00 AM provides a snapshot of the atmospheric state at the start of the day.

	affects weather conditions.		
Pressure3pm	represents the atmospheric pressure readings taken at 3:00 PM. Atmospheric pressure is a crucial meteorological parameter that influences weather patterns, including the formation and movement of clouds, as well as the likelihood of precipitation.	Yes	Atmospheric pressure is not a static parameter; it fluctuates throughout the day due to various meteorological factors. By specifically capturing the pressure reading at 3:00 PM, you're accounting for the atmospheric conditions during the afternoon period, which is often crucial for understanding weather patterns leading up to rainfall events.
Cloud9am	refers to a variable indicating the cloud cover at 9:00 AM. Including this variable in your dataset is crucial for modeling rainfall because cloud cover has a significant impact on weather conditions and precipitation patterns.	Yes	Cloud cover can vary significantly throughout the day due to factors like diurnal temperature changes, atmospheric instability, and weather systems. Recording cloud cover specifically at 9:00 AM captures a snapshot of atmospheric conditions during the morning hours, which can be critical for understanding daily weather patterns and predicting subsequent rainfall events.
Cloud3pm	refers to a feature related to cloud cover at 3 pm, a variable commonly used in weather forecasting and related prediction tasks such as rainfall prediction.	Yes	Cloud cover at 3 pm reflects the atmospheric conditions during the peak daytime heating period, which significantly impacts the development of convective clouds and subsequent rainfall events. Including this feature allows the model to capture the temporal dynamics of cloud formation and its relationship with rainfall

			occurrence.
Temp9am	records the temperature measured at 9:00 AM local time. It represents one of the key meteorological variables that can significantly influence rainfall patterns and occurrences.	Yes	By capturing the temperature at a specific time point (9:00 AM), you ensure that your model considers the nuanced variations in temperature that occur throughout the day, providing a more detailed understanding of the atmospheric conditions preceding rainfall events.
Temp3pm	recorded at 3 PM, a crucial time slot in weather forecasting due to its proximity to the peak heating period of the day	Yes	At 3 PM, most regions experience the highest temperature of the day due to solar radiation reaching its zenith. This peak heating period influences various atmospheric processes such as evaporation, convection, and cloud formation, all of which are pivotal factors in rainfall occurrence.
RainToday	RainToday is a binary variable that denotes whether rainfall occurred on a specific day or not. Typically, it's represented as a categorical variable with two values: "Yes" or "No", indicating the presence or absence of rainfall, respectively	Yes	Understanding whether it rained on the previous day provides valuable historical context for predicting future rainfall. Rainfall patterns often exhibit temporal dependencies, where the occurrence of rain on one day can influence the likelihood of rain on subsequent days. RainToday serves as an indicator of the current moisture content in the atmosphere and the prevailing weather conditions
RainTomorrow	Indicates whether it will rain tomorrow or not. It serves as the target	Yes	"RainTomorrow" serves as the target variable for machine learning models. By predicting whether it will rain tomorrow, these models can provide

	variable for rainfall prediction models, with a value of 1 denoting the occurrence of rain and 0 indicating no rain.		valuable insights for planning and decision-making in various sectors such as agriculture, transportation, and emergency management.
Date Month	represents the month component of the date when rainfall data is recorded or predicted. It typically ranges from 1 to 12, representing January to December, respectively.	Yes	Rainfall patterns often exhibit seasonal variations influenced by factors such as temperature, humidity, and atmospheric conditions. By including the Date Month variable, the model can capture these seasonal patterns, enabling it to make more accurate predictions for specific months.
Date Day	represents the calendar day on which a particular set of rainfall data was collected or recorded. It typically consists of numerical values representing the day of the month, ranging from 1 to 31, depending on the specific month.	Yes	Including the "Date Day" enables the analysis of temporal patterns in rainfall data. By tracking rainfall occurrences over specific days of the year, you can identify seasonal trends, monthly variations, and even daily patterns that might influence rainfall prediction.