Predict precipitation or not

Kim YU

Mar 15 , 2021

1 Introduction

The probability forecast of precipitation is that the forecasters predict the percentage of precipitation chance in a certain area and a certain period of time according to various meteorological data after sorting, analysis, judgment and discussion. It's just a "chance" to predict precipitation, but it's not guaranteed to be 100%! It means that there will be precipitation of 0.1 cm or more in the next 36 hours (every 12 hours is a period), which has nothing to do with precipitation time and area. For example, the 70% probability of rainfall forecast in Wenzhou means that there is a 70% chance of rainfall in Wenzhou. It may only rain in Longwan District or only rain for 10 minutes. After knowing the probability of precipitation forecast, we can take measures to reduce the inconvenience of life for the weather that may rain. If you think it doesn't matter to take rain, it's troublesome to take rain gear, so you don't have to take rain gear when the probability of precipitation is 70% or 80%; but for older people or people who are afraid of rain, even if the probability of precipitation is only 30% or 40%, you must prepare rain gear!

To predict whether the weather is precipitation, we need to judge the energy, water vapor and dynamic conditions of the day.

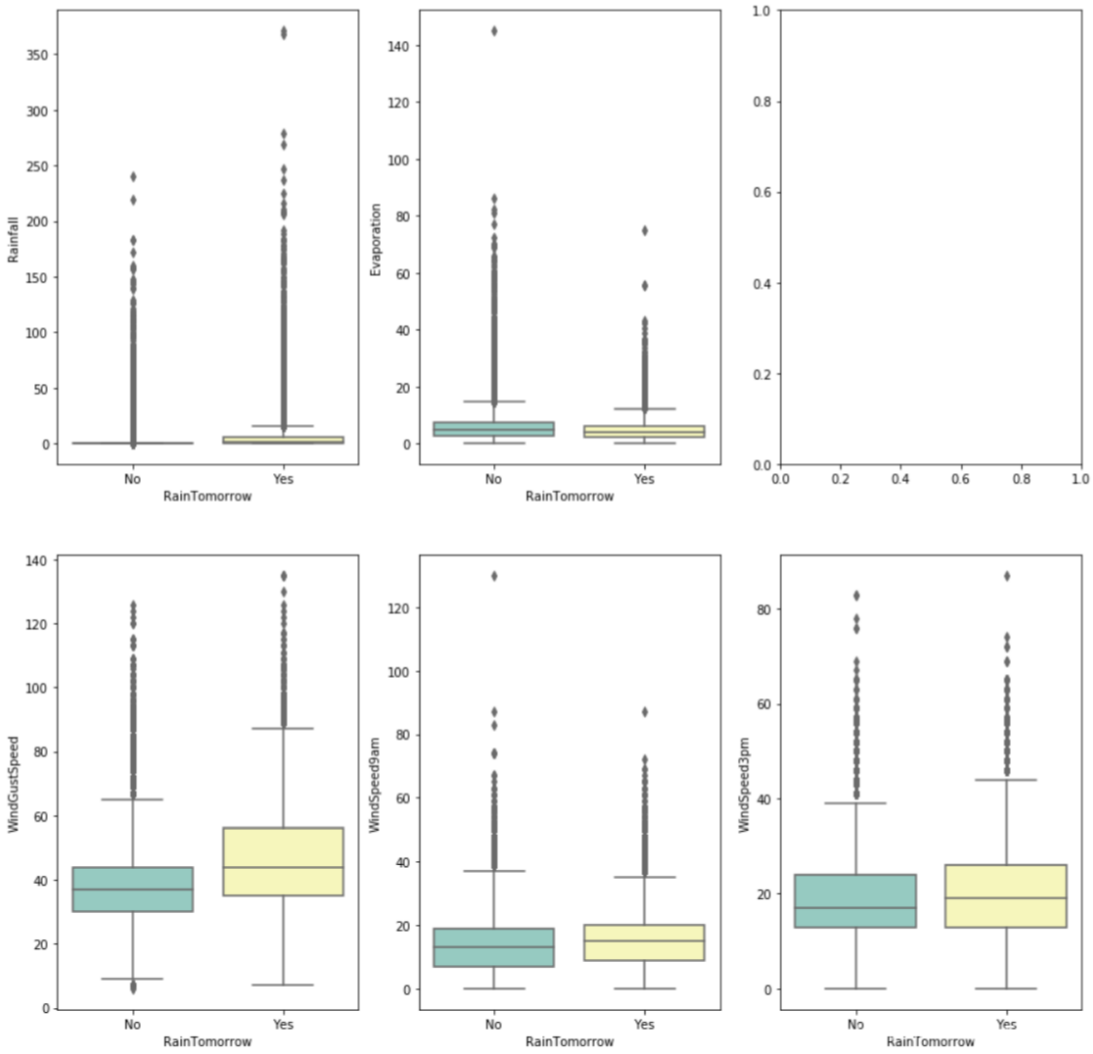
Accurately predicting whether precipitation is conducive to the development of many industries, including agriculture, water conservancy and transportation, etc. The most intuitive way is to guide the public whether to carry umbrellas when traveling.

2 Data acquisition and cleaning

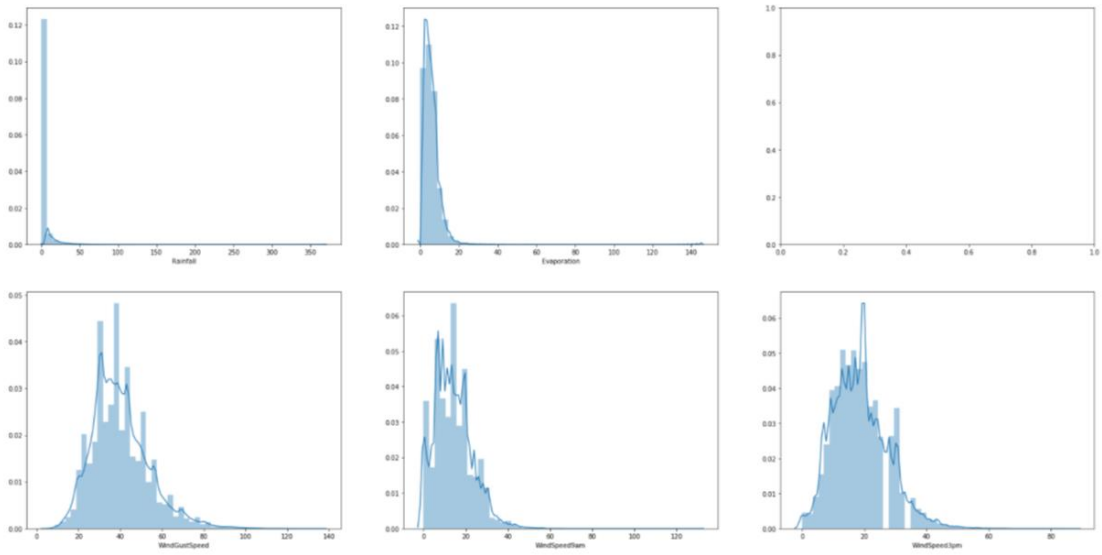
The data used is historical data of Australian weather conditions obtained through Kaggle. https://www.kaggle.com/muthuj7/weather-dataset

3 Exploratory Data Analysis

Box line analysis:Rainfall、Evaporation、WindGustSpeed、WindSpeed9am、WindSpeed3pm.

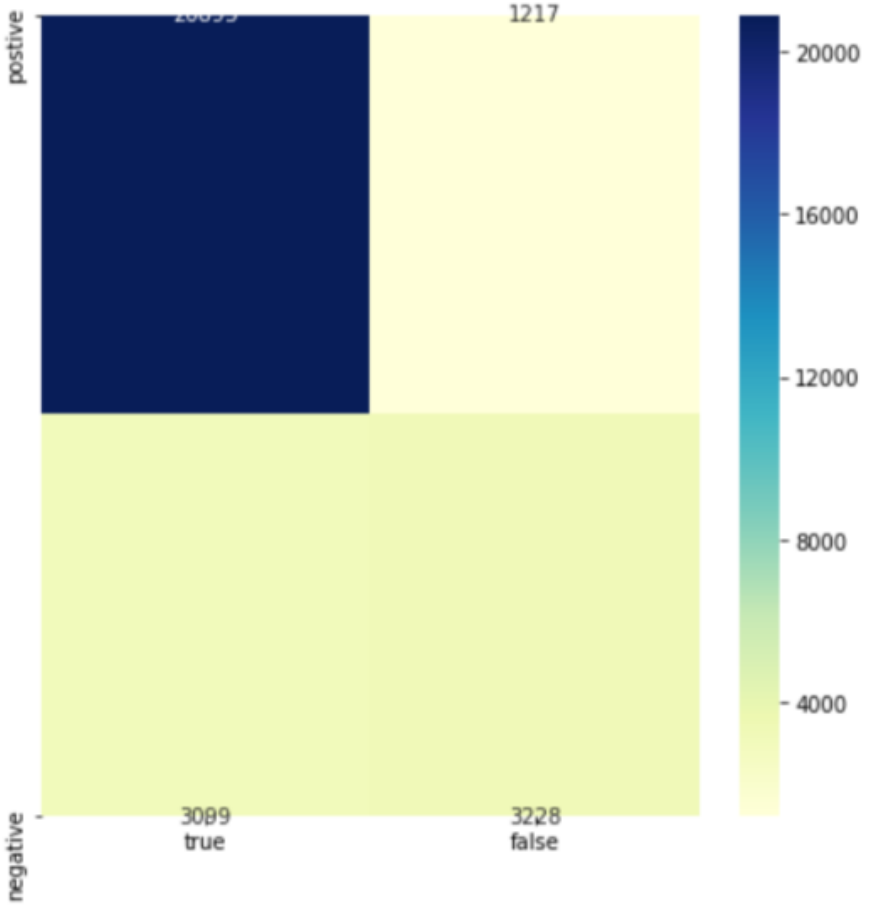


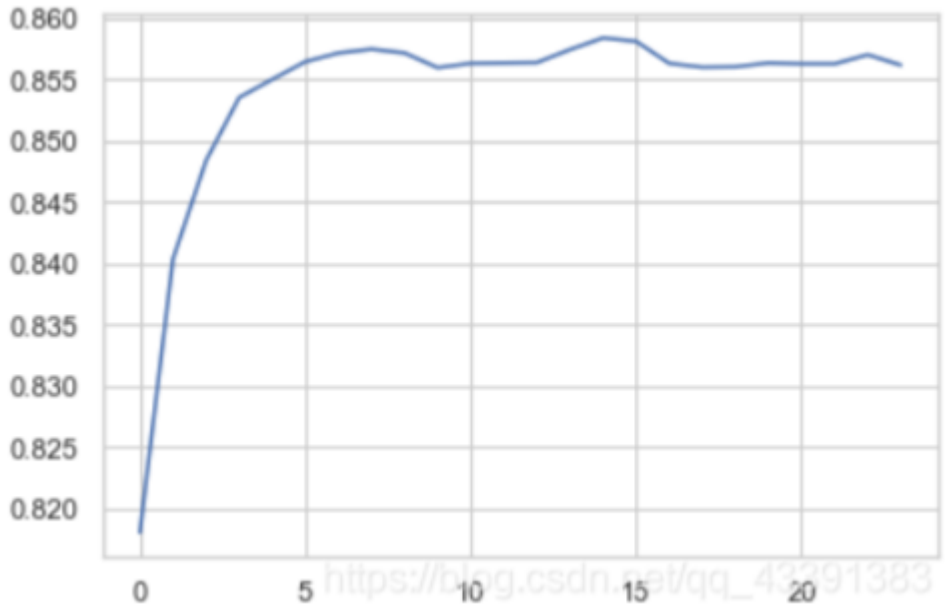
Continuous value analysis



4 Predictive Modeling

For logical regression analysis of precipitation, SKLEARN database was used for training and test data was carried out.





5 Conclusions

In this project, I analyzed various variables affecting precipitation, and analyzed different conditions to predict precipitation through logistic regression, and the accuracy rate was about 80%.

6 Future directions

On the regression problem, I can get about 50% improvement from the precipitation forecast model, and about 80% accuracy classification problem in the test. There are still significant differences that cannot be predicted. The model is studied. I think the model can improve the precipitation characteristics at the capture layer.