



## Final Project: Building a Rainfall Prediction Classifier

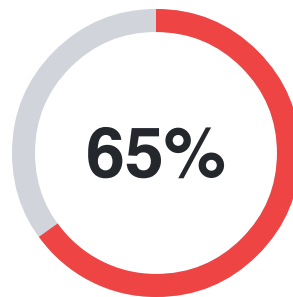
### Solid Effort!

Keep going! Mistakes are opportunities to learn and grow. Review your answers and try\_again to achieve your best result.

Required passing grade: 70%

Status: **Failed**

Final Score: 13 / 20 (65%)



[Provide Assignment Feedback](#)

### Question 1

Score: **8/10**

Upload the duly completed Jupyter notebook for "Final Project : Building a Rainfall Prediction Classifier". Please ensure that you have completed all 14 coding exercises in the ipynb file before sharing the same.

- FinalProject\_AUSWeather.ipynb [View Content](#)





The Jupyter notebook is mostly completed with 14 coding exercises, demonstrating a good understanding and application of the concepts. The code is generally well-organized and structured, with the use of headings and comments for clarity. Visuals are included and relevant to the analysis, enhancing the overall presentation. However, there is a minor issue with Exercise 15, where the 'model' variable is not defined, leading to a `NameError`. This affects the completeness of the final exercise. To improve, ensure that all variables are properly defined and consider adding more detailed comments to further enhance readability and understanding. Overall, the notebook is well-done with only minor areas for improvement.

## Question 2

Score: 2/3

Which of the following features would be inefficient in predicting whether it will rain tomorrow or not? (Select all that apply)

☒ Evaporation



☒ WindGustDir



☒ MaxTemp



☒ Humidity9am



Evaporation cannot be used efficiently since the total amount of water evaporated cannot be known till the day is over..

WindGustDir cannot be used efficiently since the direction of the strongest wind gust of the day cannot be known till the day is over..

MaxTemp cannot be used efficiently since the maximum temperature cannot be known till the day is over..

Humidity9am can be used for predicting the rainfall of tomorrow since it is defined over a specific time..

The correct option(s) were: MaxTemp, WindGustDir, Evaporation.

### Question 3

Score: 0/2

What is the True Positive rate of the RandomForestClassifier based on the confusion matrix from Exercise 13 in the assignment notebook?

Note: The answer should be expressed as a percentage, rounded to the nearest whole number.

86



The answer you provided, '86', is not relevant to the question. The question asked for the True Positive rate of the RandomForestClassifier based on the confusion matrix from Exercise 13, which should be expressed as '51' or '51%'. Please review the confusion matrix from Exercise 13 and provide the correct True Positive rate as specified. Make sure to express your answer as a percentage, rounded to the nearest whole number.

### Question 4

Score: 2/2

Identify the most important feature for predicting whether it will rain based on the feature importance bar graph from the project lab.

☒ Humidity3pm



☐ Sunshine

☐ Temp3pm

☐ Pressure3pm



Correct! Humidity3pm is the most important feature for predicting rain according to the feature importance bar graph

## Question 5

Score: 1/3

In bullets or as a numbered list, answer the following.

Comment on the accuracy of the LogisticRegression and Random Forest Classifier models.

Comment on the true positive rate of the two models.

Overall, which one of the two is a better predictor of whether it will rain tomorrow or not?

- [83, 84]
- [86,86]
- Random Forest Classifier

**Relevance:** The response partially aligns with the question, as it attempts to address the accuracy and true positive rates of the models and chooses the better predictor. However, it lacks detail and context.

**Completeness:** The submission does not fully address all parts of the question. The learner provides numerical values without explanation or context, and the response lacks a comparison or analysis of the metrics.

✦ **Quality of Execution:** The technical quality is low because the response does not provide any reasoning or context for the numbers given. It fails to compare the models as requested.

**Originality:** The response does not demonstrate any unique or creative approach, as it simply lists numbers without explanation.

**Clarity:** The response lacks clarity due to the absence of explanations or context for the numerical values provided.

**Adaptability:** To improve, the learner should provide explanations for the numbers, compare the models based on the metrics, and give a clear rationale for choosing the Random Forest Classifier as the better predictor.

[Report an issue](#)

[Retake Assignment](#)

