Rainfall Prediction Variables

Mengyu Jackson

Summary

26 Weather Stations record all core and optional variables

23 Weather Stations record "Core" variables + some Optional variables

Australian government wants to know if recording more variables will improve rainfall forecasts at these 23 locations.

We found only Wind Gust Speed is important to record.

———

Outline

- Business Problem
- Data
- Results
- Recommendation
- Conclusions
- Source

Business Problem

Rainfall forecasting is used for many critical purposes.

All weather stations record Core variables, but about half do not collect all Optional variables

The Australian government wants to know: will recording "Optional" variables at all stations improve rainfall forecasting?

Data

- 3 to 10 years of daily weather observations
- 49 locations across Australia
- 14 "Core" variables
- 7 "Optional" variables
- 26 locations collect all variables (Core + Optional)
- 23 locations never collect at least 1 Optional variable

Data
Core variables

- Rainfall Today
- Min + Max Temperature
- Temperature (9 am + 3 pm)
- Wind Gust Direction
- Wind Direction (9 am + 3 pm)
- Wind Speed (9 am + 3 pm)
- Humidity (9 am + 3 pm)

Data
Optional variables

- Cloud (9am + 3pm)
- Evaporation
- Pressure (9am + 3pm)
- Sunshine
- Wind Gust Speed

Results

Best Model Predicts Rainfall Without Optional Features

- Trained 3 Model Types (Linear, Tree Based, Time Series)
- Predictions using Core + Wind Gust Speed was comparable to Core + all Optional features in all models
- Best model (Time Series) predicted rainfall at "Core + Wind" sites equally well as "Full" sites
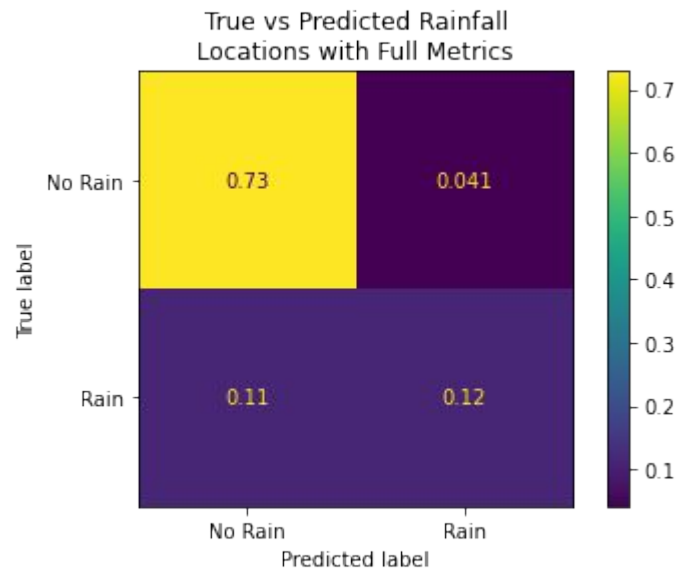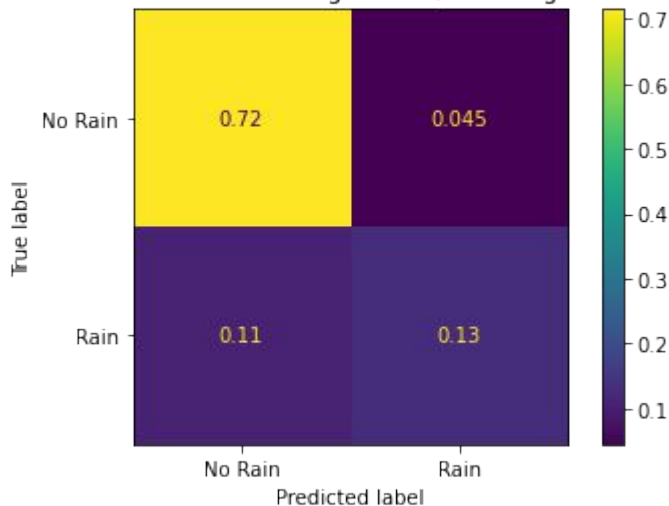
Results

High Impact:

- Wind Gust Speed

Low Impact:

- Cloud (9 am + 3 pm)
- Evaporation
- Pressure (9 am + 3 pm)
- Sunshine

True vs Predicted Rainfall

Recommendation

- Newcastle and Albany should record Wind Gust Speed
- Newcastle is the worst-predicted location by the Time Series Model
- Albany is in the bottom 25%

Conclusions

Recording Wind Gust Speed at all Locations will help improve rainfall forecasting at those locations.

Other Optional features can be helpful for some model types, but the best models do not need them.

Source

Observations were drawn from numerous weather stations. The daily observations are available from http://www.bom.gov.au/climate/data.

An example of latest weather observations in Canberra: http://www.bom.gov.au/climate/dwo/IDCJDW2801.latest.shtml

Definitions adapted from http://www.bom.gov.au/climate/dwo/IDCJDW0000.shtml

Data source: http://www.bom.gov.au/climate/dwo/ and http://www.bom.gov.au/climate/data.

Thank You For Watching!

- Mengyu Jackson