Dengue and Data

A study on the Dengue outbreaks in Malaysia

The Dataset

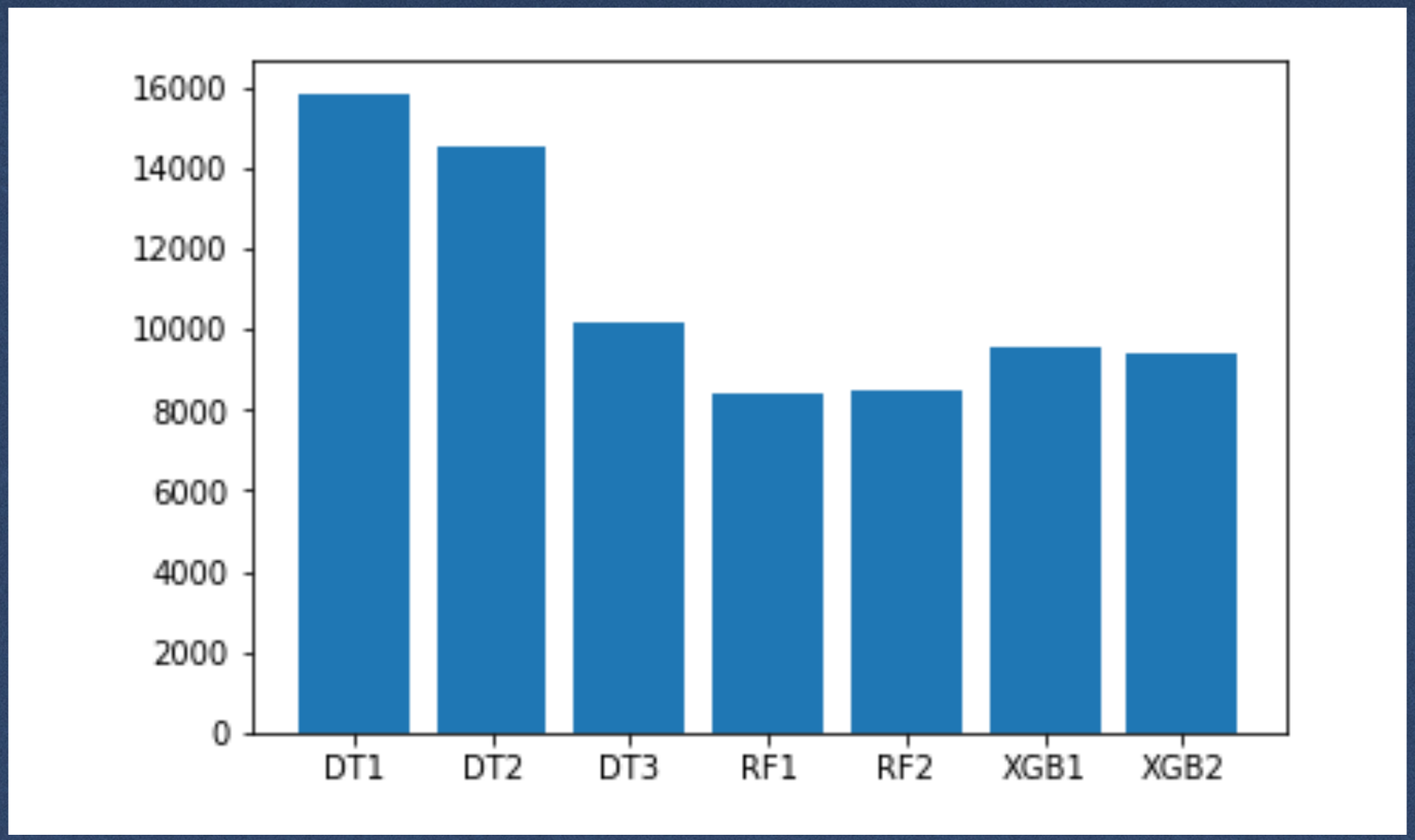
- *City observation is made*
- *Month of observation*
- *Year of observation*
- *Cloud cover*
- *Wind speed*
- *Humidity*
- *Max temperature*
- *GDP (whole country, that year)*
- *Population(whole country that year*

Sources:

Weather: https://darksky.net/dev

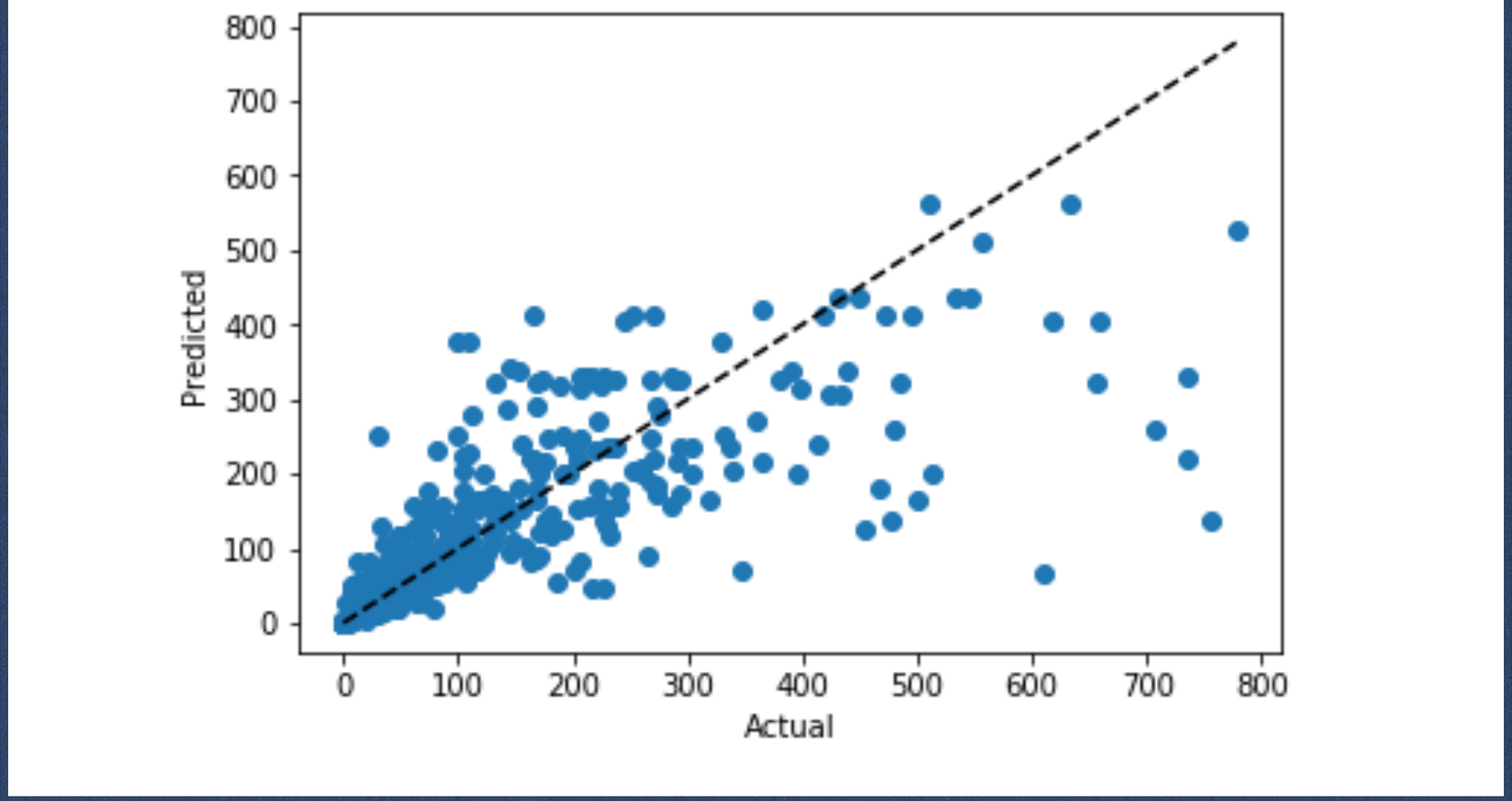
GDP, Population: https://data.worldbank.org/country/malaysia

Reported cases: ://www.tycho.pitt.edu/



The models and their performance

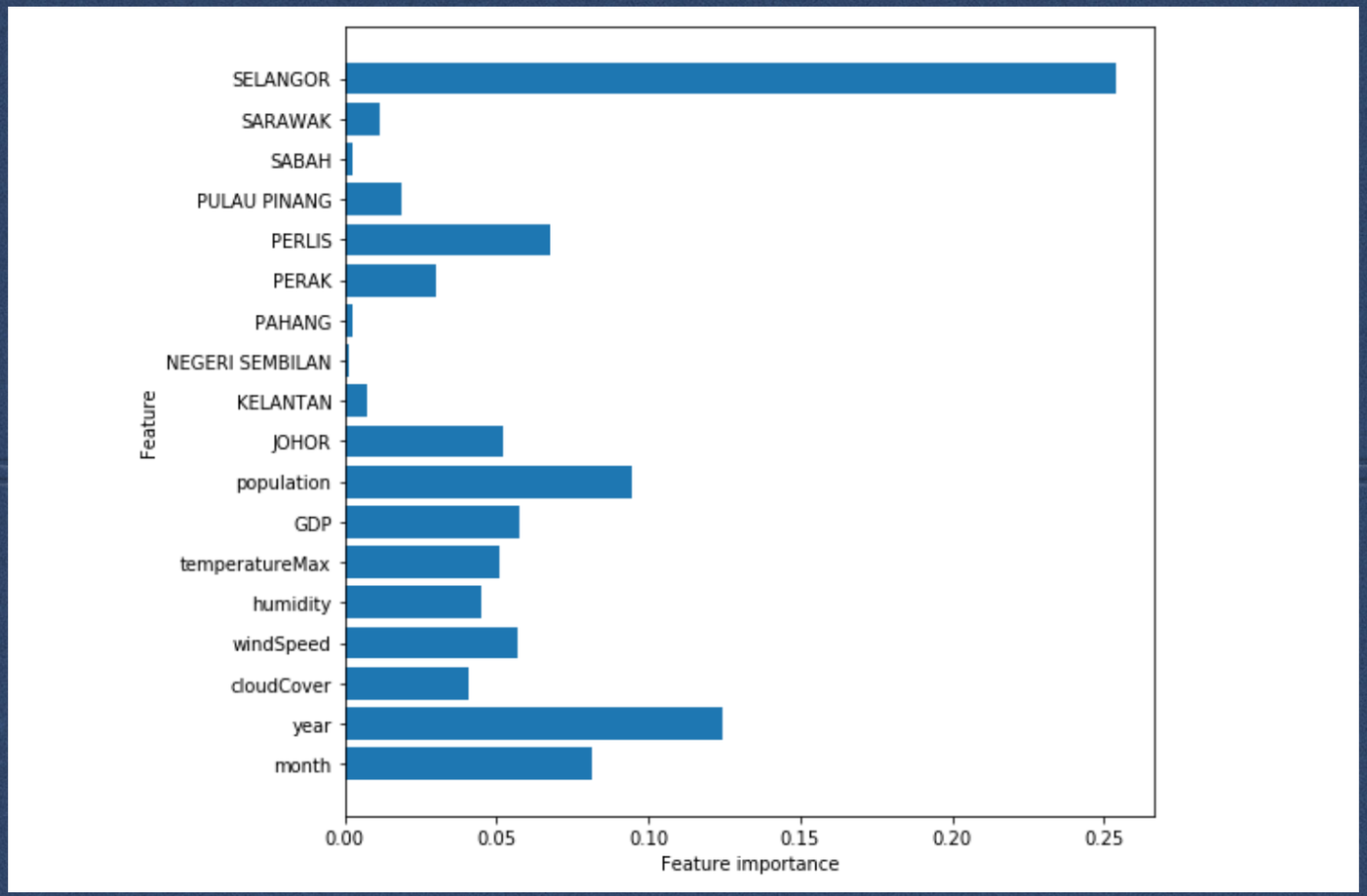
metric: mean squared error



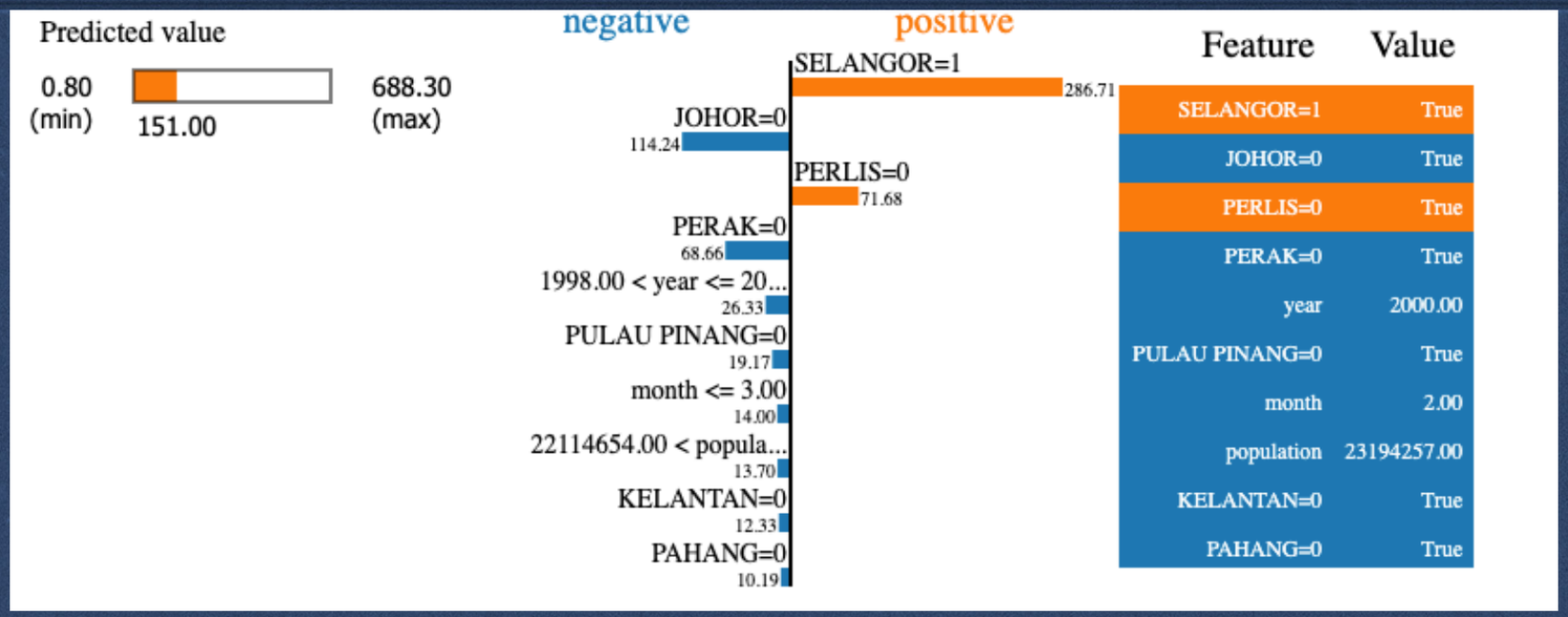
Random Forest Model

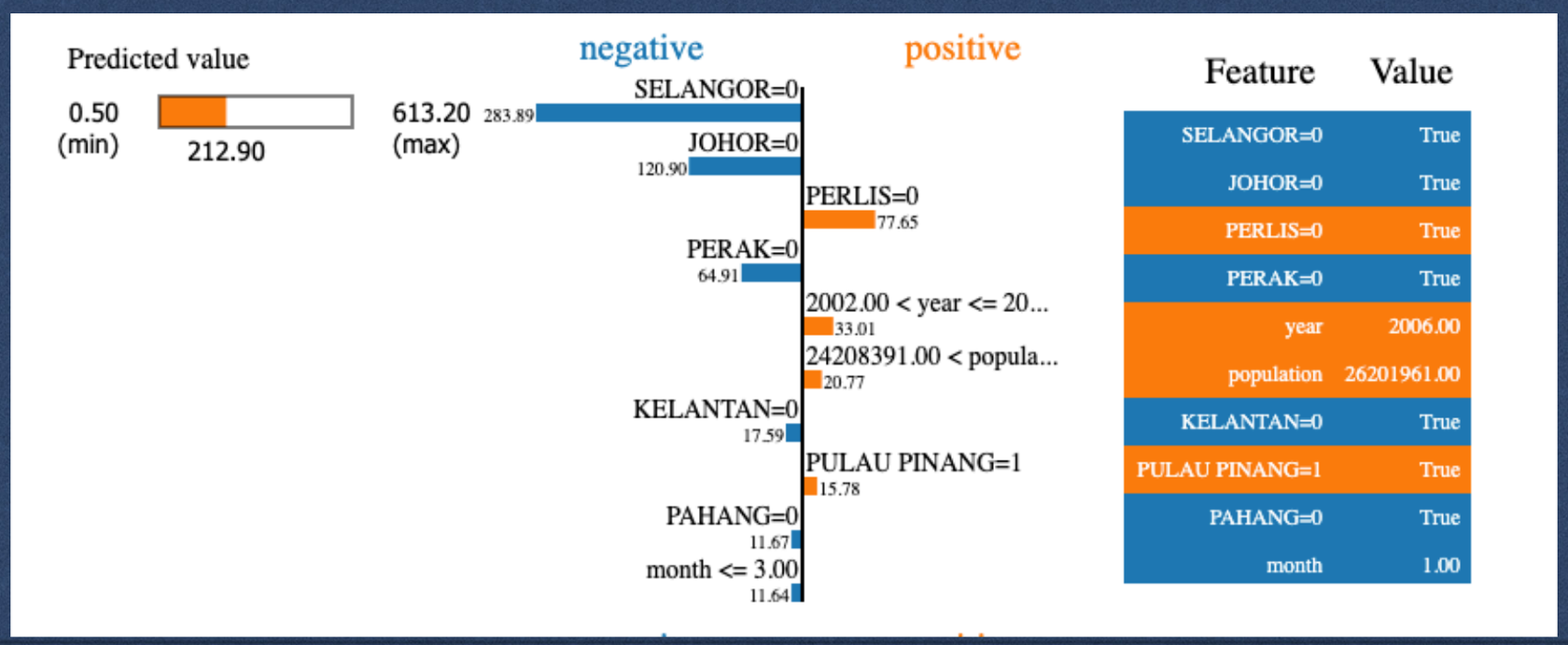
Mean Squared Error: 8373.59

Feature Importance



Prediction Calculations





In Conclusion

- *Residing in Selangor, Johor, or Perak was linked to higher instances of Dengue contraction*
- *Perlis was safer and was linked to lower instances of Dengue*
- *Populations > 2.4 million exhibited more cases of Dengue*
- *Observations after 2002 generally had higher observations of Dengue*