



Labor Condition Application (LCA) Case Status Prediction

Will your application to hire non-immigrant foreign employees get approved?

Developer: Tong Yin
Product Owner: Joe Gilbert
QA: Grace Cui

Motivation



Background: In order to hire foreign employees, employers in the United States have to first file a Labor Condition Application (LCA) to the Office of Foreign Labor Certification (OFLC).

Objective: This web app helps employers:

- **forecast** their petition result (Yes for Approved, No for Denied)
- **test** with inputs to understand better the process of getting approval



Data Description

Source: United States Department of Labor Disclosure Data (10/01/2016 - 09/30/2017)

<https://www.foreignlaborcert.doleta.gov/performance/data.cfm>

Details: Raw data: 50+ feature numerical/categorical variables, 600,000+ observations

 Processed data used in app: ~15 predictor variables, ~50,000 observations

Final predictors: number of total workers, prevailing wage, H1-B dependence, etc.

Final response: case status code (1 for TRUE, 0 for FALSE)



Random Forest Classifier

Why Random Forest?

Binary response variable; Many categorical predictors

How?

Trained and evaluated on final processed data using **pandas** and **scikit-learn**

Successful Criterion?

Misclassification Rate ~ 0.33

Given a user input, call the trained RF object and it will return 1 (Approved) or 0 (Denied)



Interesting Insight

Serious imbalance problem in data: 90% of observations are APPROVED

➡ Balanced data by downsampling to build more robust predictions



Thank you for watching!

Questions?

Please contact me at
tongyin2018@u.northwestern.edu
or visit <https://github.com/Tong-Yin>