5741 Project Proposal: H-1B Petition Analysis

Yueran Yang, Zongyuan Yuan

October 2021

1 Background

The H-1B is a visa in the United States that allows U.S. employers to temporarily employ foreign workers in specialty occupations. In the first stage of the application, the employer files a Labor Condition Application (LCA) for the employee that includes information such as wages and job titles to the Department of Labor. After approval for LCA, the process goes into a random lottery based on different wages. However, the new policy would change the current wage level requirements.

2 Question

- The propose of our project is to help the company to judge whether they should hire a foreign worker for a position considering visa status. As dealing with H1B applications would bring the company additional cost and also losing a talent due to failure in H1B petition can cause more money on hiring new person. We hope that we can help the company to find a optimal prediction on those risks.
- We also want to analyze which kind of jobs or companies should the foreign students to apply to maximize the chance in H1B application.

3 Data

The data using in this project is H1B Disclosure Dataset in Kaggle shared by user Charmi. The entries in the data contains wages, positions, companies, companies tax number, status, etc.. Thus, we can use each entries features to train a classifier for predicting the status. Using different features can help us achieving different goals (as we mentioned in the questions above).

4 Importance

The importance of the project is that we can help the company to determine whether they should hire a foreign worker considering the randomness in H1B application, potential costs in lawyers and potential loss in the foreign workers, which can help the companies to save money in training new workers and hiring application lawyers working for the foreign workers. Also, after standing in the companies' shoes, we can help a foreigner to better plan their careers.