PROJECT DOCUMENT

Topic: Predictive Modelling for H1b Visa Approval Using IBM Watson

Introduction

Over 2 million visa petitions are filed by the employers each year and only 65000 petitions are approved. So, the goal is to explore the petitions filed and their outcomes for the past six years i.e., from 2011 to 2016, and to find a pattern to predict the outcome by using a predictive model developed using Machine Learning techniques.

In the Guided Project, our goal is to predict the outcome of H-1B visa applications that are filed by many professional foreign nationals every year. Here, we framed the problem as a classification problem and applied it in order to output a predicted case status of the application. The input to our algorithm is the attributes of the applicant. H-1B is a type of non-immigrant visa in the United States that allows foreign nationals to work in occupations that require specialized knowledge and a bachelor's degree or higher in the specific specialty. This visa requires the applicant to have a job offer from an employer in the US before they can file an application to the US immigration service (USCIS). We believe that this prediction algorithm could be a useful resource both for the future H-1B visa applicants and the employers who are considering sponsoring them.

In order to predict the case status of the applicants, we will be feeding the model with the dataset which contains the required fields by which the machine can classify the case status as certified or denied.

Overview

Companies file approximately 2 million visa applicants each year, however only 65000 of them are authorized. So, the idea is to look at all of the petitions that have been submitted and their outcomes over the last six years, from 2011 to 2016, and see if there is a pattern that can be used to estimate the outcome using a predictive model constructed using Machine Learning techniques.

Purpose

The objective is to forecast the outcome of H-1B visa applications, which are filed by a large number of professional foreign nationals each year. In this instance, we framed the challenge as a classification problem and used it to anticipate the application's case state.

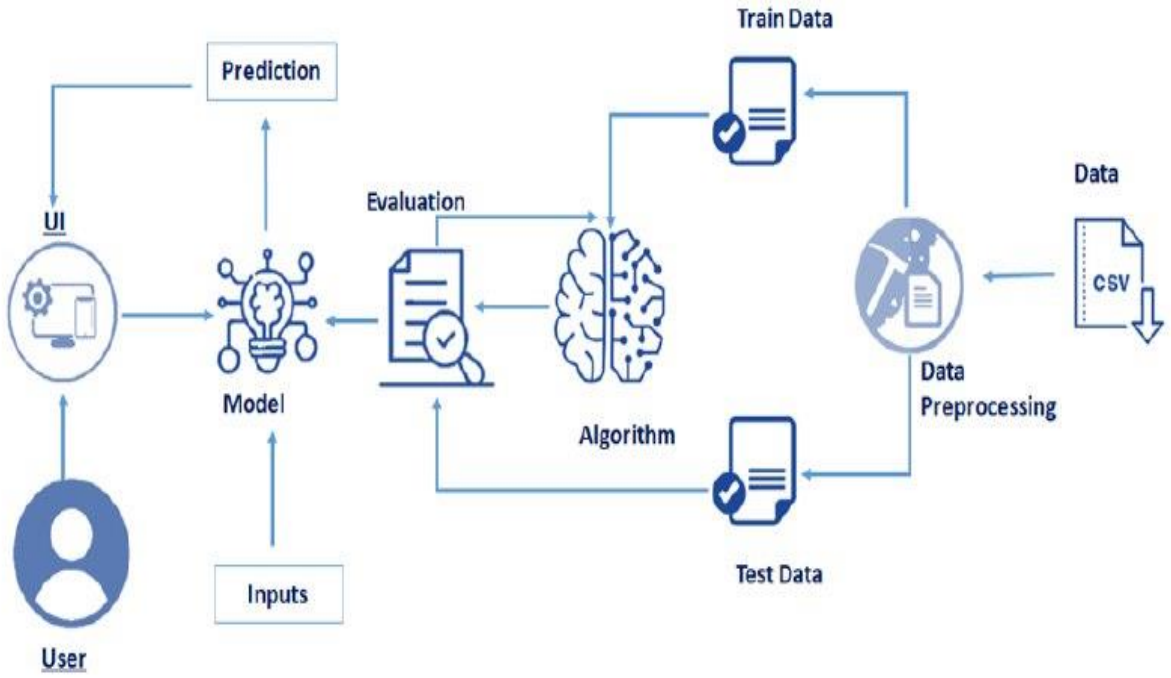
Existing problem

Due to the limited number of granted Visas, assessing the acceptance or denial of the same is challenging. As a result, we can easily predict the conclusion using machine learning techniques.

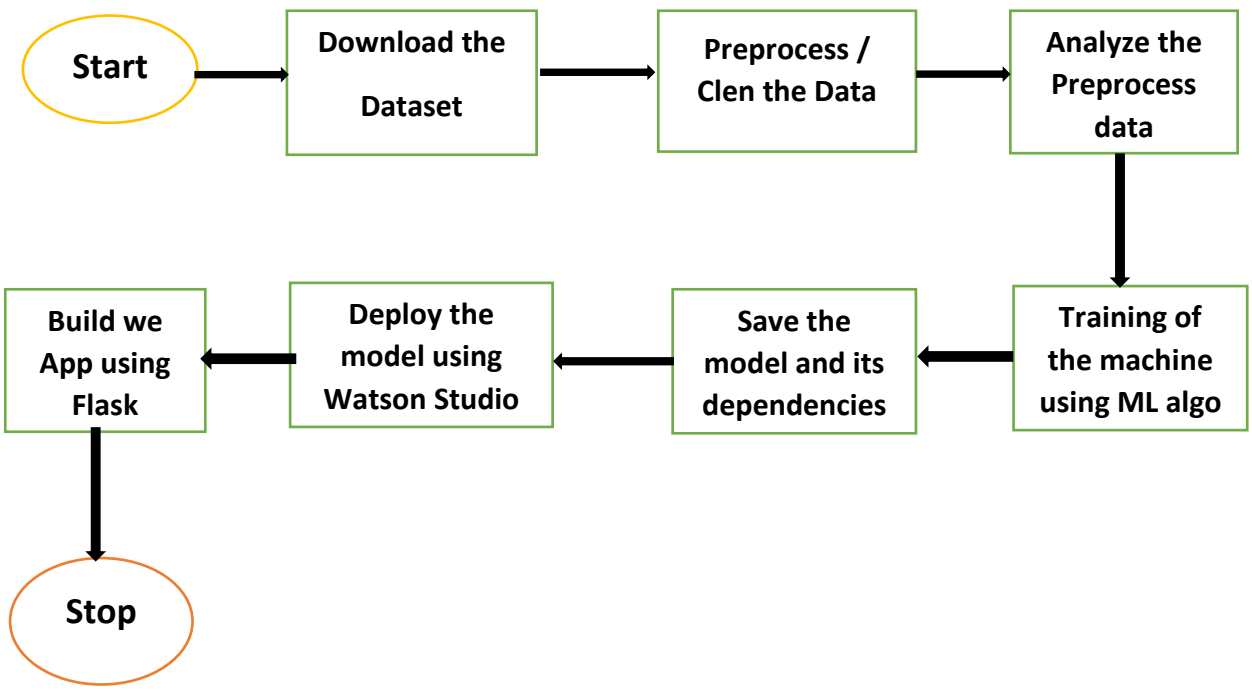
Proposed solution

For prediction, we use conventional machine learning methods such as classification, random forest, and decision trees. For quick user interaction, the model will be deployed on Flask.

Block Diagram:

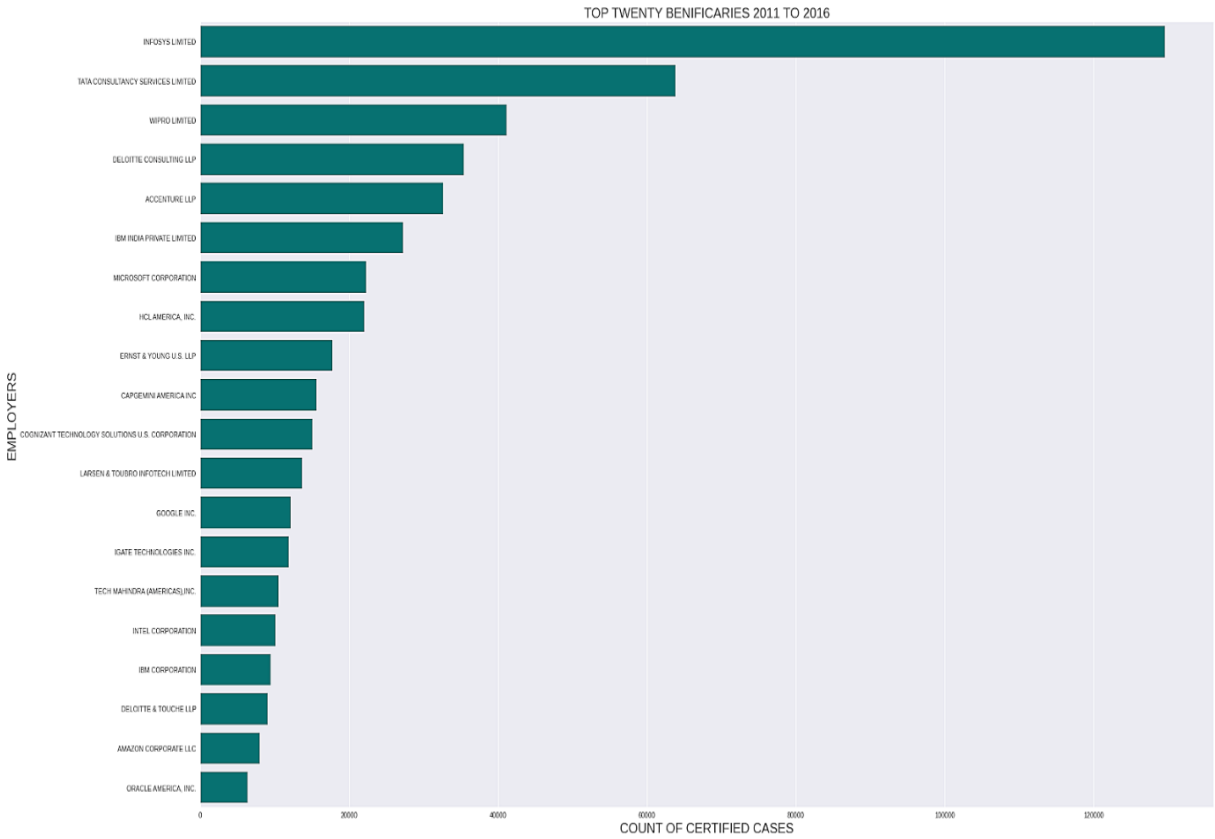


Flowchart:



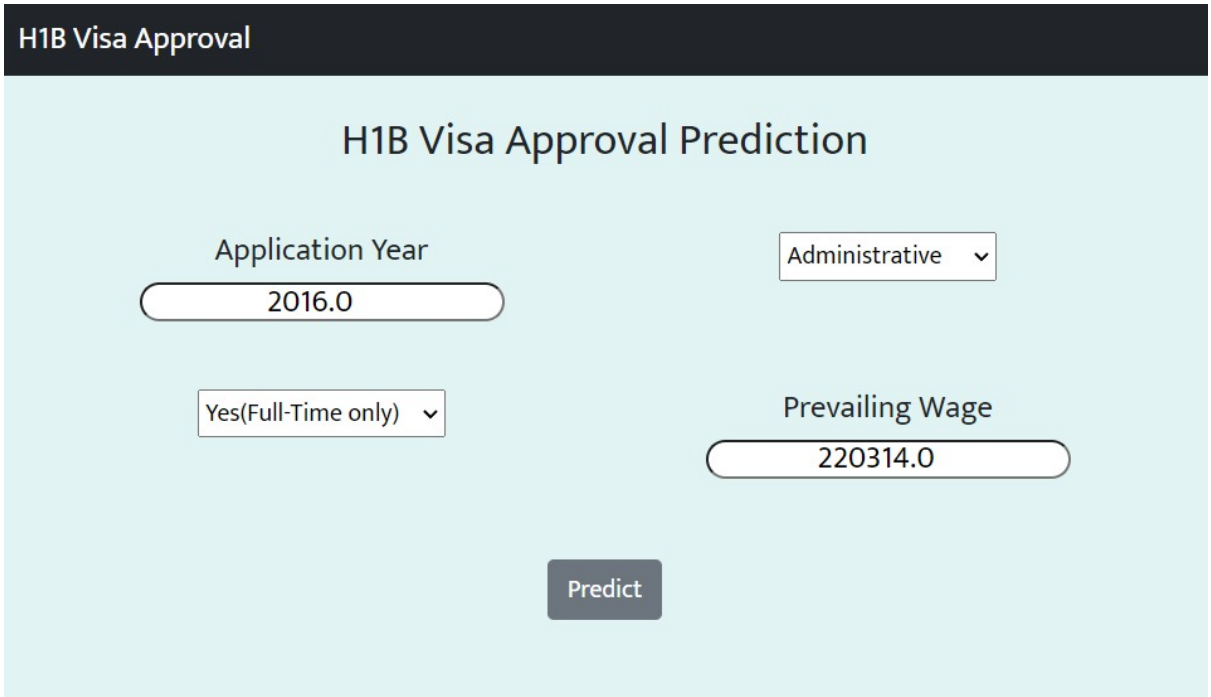
Experimental Investigations

- Out of the seven fields, certified visa is the most prevalent.
- The majority of visas granted are for full-time employment.
- During 2011 to 2016, we see a progressive rise in the number of accepted H1B applications.
- Infosys one of the biggest proponents of H1.



Result

A random forest model is used to predict visa status with 86 percent accuracy when we input information such as occupational code, position, prevailing wage, and year of application.



Advantages and Disadvantages

Knowing the status of the visa will encourage applicants in compiling documentation if it is to be approved, and in putting more effort to upgrade their application if it is to be denied.

The application's lack of robustness is a key drawback.

Applications

The aim is to estimate the outcome of H-1B visa applications, which are filed regularly by a high number of professional foreign people. We presented the problem as a classification task in this example and utilized it to determine the application's case state.

Conclusion

To anticipate the visa status, a model that focuses on Random Forest is developed.

Future Scope

- Extending the dataset's capacity and introducing additional features.
- The user experience (UX) and User Interface (UI) can be upgraded.

Bibliography

https://smartinternz.com/

Appendix

Source Code: Here