

CSI 5155 Project Proposal Group 7

- Project Title: The prediction of case status under H1B visa application
- Project Type: application oriented
- Names of Group Members:
 1. Lingfeng Zhang,300134245
 2. Biliang Wang,300105002
- Description of the Dataset:
 - Dataset Name: H1B Disclosure Dataset
 - Available Link:
 1. <https://www.kaggle.com/abishekanbarasan1995/h1b-case-status-prediction>
 2. <https://www.kaggle.com/trivedicharmi/h1b-disclosure-dataset>
 3. https://www.foreignlaborcert.doleta.gov/docs/Performance_Data/Disclosure/FY15-FY16/H-1B_FY16_Record_Layout.pdf
 - Detailed Illustration:
 - Dataset Size: Around 500,000 from year 2015 to 2018 in total
 - Feature Number: 40
 - Class Number: 4
 - Our group plan to use some methods to reduce the number of features and the number of dataset(e.g. selecting the data in one year)
- Project Description: Many internationals hope to get working permission in the USA for varied reasons. In this project, we plan to classify the H1B visa case status. Our group is eager to mining covered insights to help the applicants understand the process of getting an H1B visa. Receiving the approval of H1B visa needs to meet many criteria, and different applicants have their own special characters. This is the reason why the case status is tricky to be predicted for every single individual. By analyzing the H1B visa application dataset, our group hopes to search which kinds of criterion affect the case status more in the H1B visa application than others. Base on the exploration, hopefully, many H1B visa applicants can success by using this "guideline". Furthermore, we will compare all differences of several classical machine learning algorithms completely, understand how these algorithms work and construe the pros and cons of these various models.

- Statement: I hereby certify I will not use these dataset(s) and algorithm(s) in any other courses, past, present or future. I understand that submitting overlapping material for more than one course is not allowed. This work is also not part of my thesis or capstone project research. I understand that any such action on my part will lead to a grade of INC being awarded for CSI5155 and may lead to academic sanctions imposed by the faculty of Engineering.