Written page written

By Ritchie Ling

I use random forest as the model which is recommended by professor Charles.

First clean data, encoding on category variables. Ranking from 8 to 1 by the importance of category. For example, IF is ranked 8 and I1 is ranked 1 because IF is the most rare one and I1 is the most usual one.

No other special need to say, coding in class to make mind clear.