
# H-1B AND PERMANENT VISA APPLICATION CASES ANALYSIS

Data Mining Report (Part III)


## ABSTRACT

Using classification to create a suitable model for predicting the certification result of a specific applicant's case.



HUA, MENGLI/ MA, LIANG

CSE7331 Data Mining

# Contents

# II. Executive Summary

Data classification is the process of analyzing data by organizing them into categories for its most effective and efficient use. In machine learning and statistics, classification is the problem of identifying to which of a set of categories (sub-populations) a new observation belongs, on the basis of a training set of data containing observations (or instances) whose category membership is known.[1] It is an example of pattern recognition, same as the clustering analyzing. The main difference between these two modeling methods is classification can be used to learn where a training set of correctly identified observations is available (which is considered supervised learning) while clustering only involves grouping data into categories based on some measure of inherent similarity or distance (which is considered unsupervised learning).

As a supervised learning, classification can be seen as a function that maps an input to an output based on example input-output pairs.[2] In other words, it can be used to predict the result of a input case based on the known examples (the training datasets). Thus, the goal of classification is to map new examples based on the training examples. Determining a suitable classifier for a given problem is however still more an art than a science. There is no single classifier that works best on all given problems, because that the performance of classifier depends greatly on the characteristics of the chosen data.[3]

In this project, we will still work on the U.S. Permanent Visa dataset, implement several classification algorithms and answer following questions:

- ❖ Which algorithm could give the best model for this case?
- ❖ Predict whether an applicant can get certified?
- ❖ How could this classification result be used in practice?

And we will also discuss that is the most important factor of certification rate differ between states?

# III. Data Preparation

To answer the question about predict the certified or not of a specific case, we defined a boolen value called certified based on the *case_status*. If the *case_status* equals to “Certified”, the correspond value of *certified* is TRUE; if not, the value will be set to FALSE. Then, transfer the *certified* into factor. And the values of *certified* (TRUE and FALSE) are the class variables for this question. We also defined *SOC_NAME_short* as the first 6 character of each value of the *pw_soc_title*, for better aggregating the similar SOC name. The features used to answer the question are *citizenship* (defined as the clean data[a] of *country_of_citizenship*), *pw_amount* (defined as the clean data of *pw_amount_9089*), *class_of_ad* (defined as the clean data of *class_of_admission*), *state* (defined as the clean data of *job_info_work_state*), and the *SOC_NAME_short*.

a. Each data is cleaned in project 1 and 2.

Thus, the whole dataset used for classification is containing *certified*, *SOC_NAME_short*, *citizenship*, *pw_amount*, *class_of_ad*, and *state*.

# IV. Modeling

For Q1, we first choose a balanced classification method, classification tree, as the classifier.

## 1. Classification Tree

Classification tree analysis is a kind of decision tree analysis when the predicted outcome is the class to which the data belongs.[4] By using this method, we can implement the visualization of the classification model as a tree structure. Via the simple tree structure, it is not only easier to understand or analyze the data, but also easier to process the nominal variables even the variable with missing value.[5] Additionally, decision tree model has the lowest training time cost.[6]

The generated classification tree is as follow (Figure 4.1):

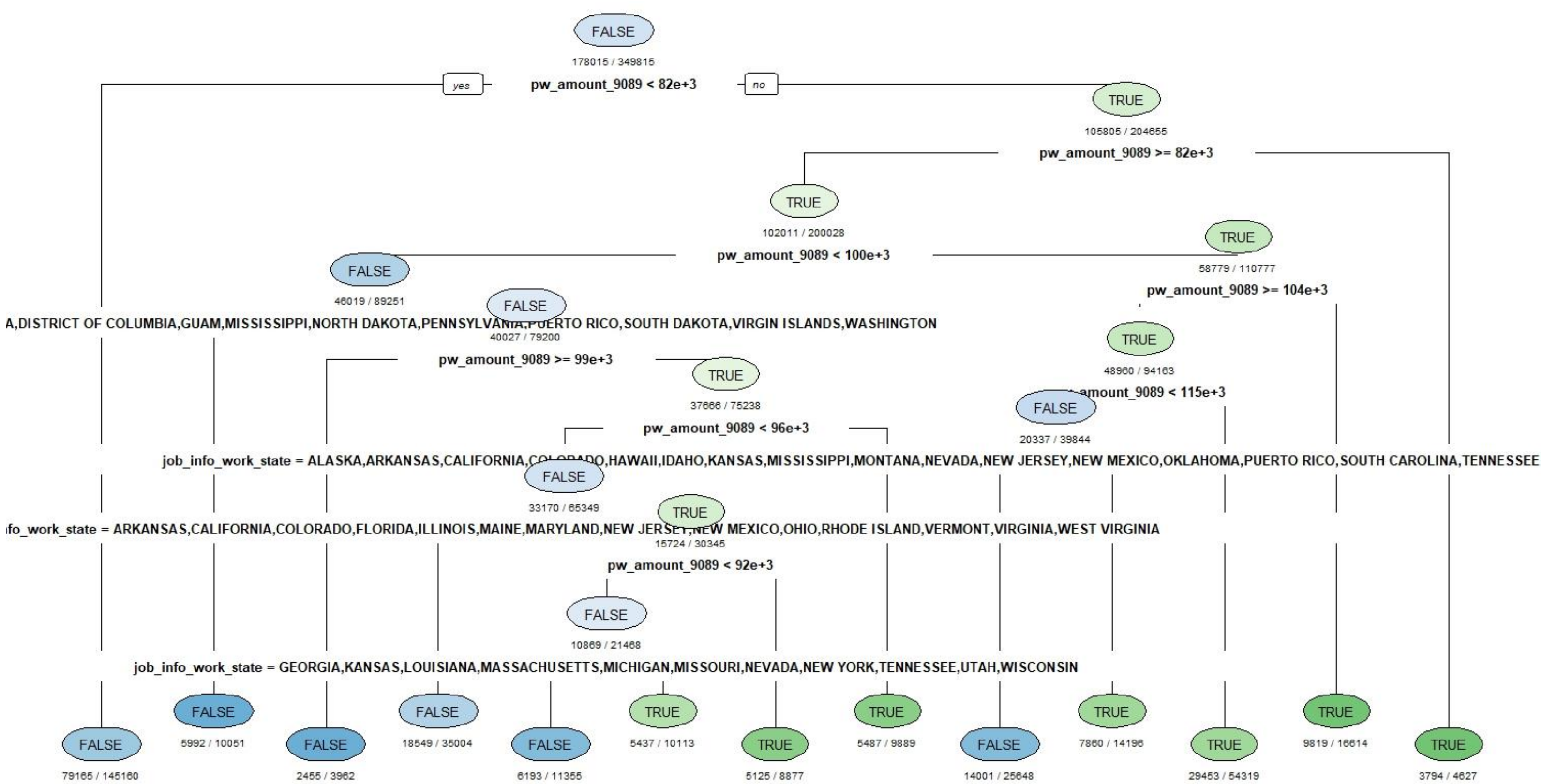

Figure 4.1

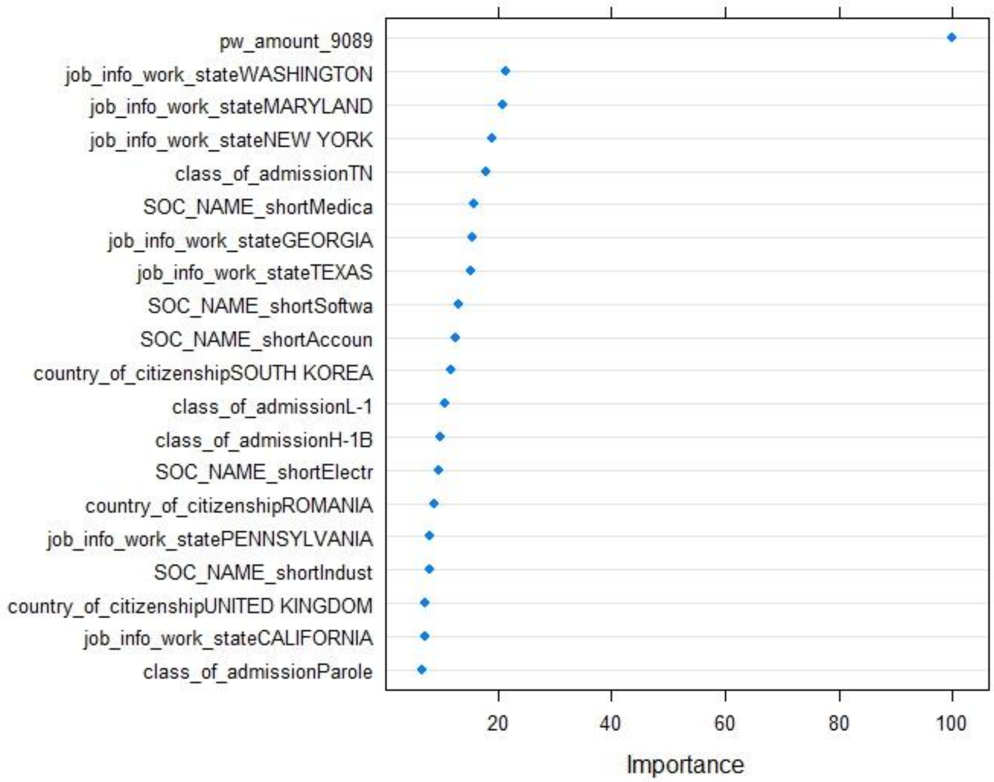


Figure 4.2

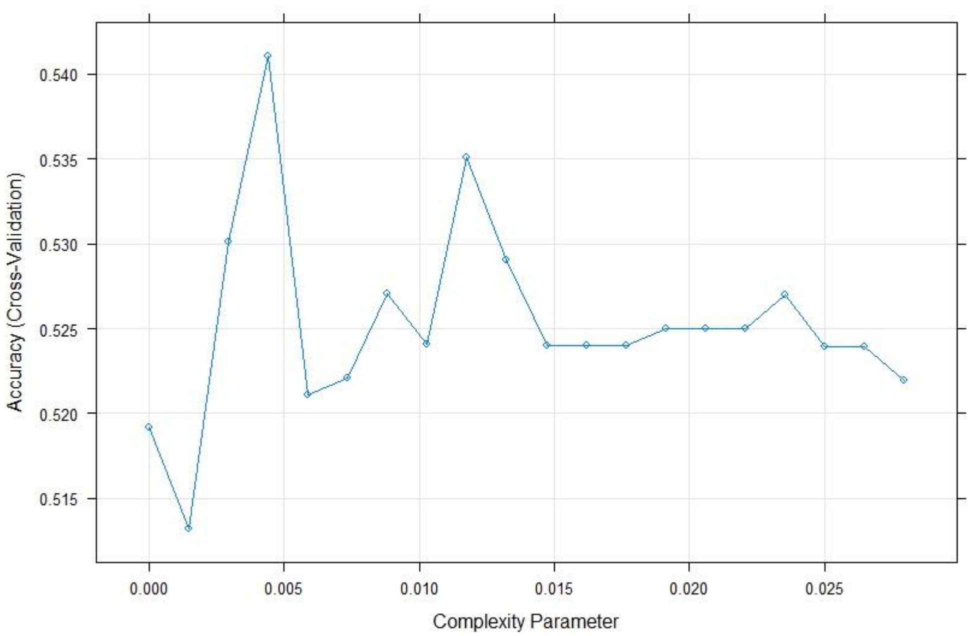

Figure 4.3

From the tree graph above (Figure 4.1), it is easily to find that the most important factor of this model is pw_amount. If an applicant's wage is lower than 82e+3, he/ she is more likely to fail his/ her application. (not certified: 79165/145160≈0.55) The second one is state. And this finding can also be proven by the importance distribution plot (Figure 4.2). The accuracy of this decision tree is shown as the Figure 4.3, seems that the highest value is almost 0.54, not so useful when predicting. This lower accuracy may be caused by the exist noises and the complexity of the dataset.

However, from the Figure 4.1, we can find out that for different state, the range of the wage that will affect the certification rate is different. For example, for states such as New Jersey, New Mexico, California, when their applicants whose wage is higher than 115e+3, they are less likely to get certified than other states' applicants (not certified: 14001/25648≈0.55).

The evaluation result of this model is shown as Table 4.1.

**Table 4.1**

| **Training error** | 0.5286321 |
|---|---|
| **Generalization error** | 0.5290235 |

**split into 2/3 training and 1/3 testing data*

This evaluation help proving that this decision tree is not the best model for this case.

Then, we use several unbalanced methods and compared all these models to find out which one is better for this specific case. The first one is the PART (Rule-based classifier).

**Because of the hardware restriction, we could only use a smaller random sample (1,000 rows) for following models training.*

## 2. PART (Rule-Based Classifier)

Rule-based classifier is a method using a set of IF-THEN rules for classification.[7] This method is easy to understand. The expressive ability of a rule set is almost equivalent to a decision tree because decision trees can be represented by mutually exclusive and exhaustive rule sets. Both the rule-based classifier and the decision tree classifier make a linear division of the attribute space and assign classes to each division. However, if rule-based classifiers allow one record to trigger multiple rules, a more complex decision boundary can be constructed. Rule-based classifiers are often used to produce descriptive models that are easier to interpret, and the model's performance is comparable to decision tree classifiers.[8]

Through observing the whole result of the PART (rule-based classifier), the 182 rules, we could find out that the key factor, which has the highest appearance rate, is the worksite state. This is different from the decision tree. This different may be caused by their different feature: for decision tree, the order is one of its feature, and duplication may exist in subtrees; on the other hand, for rule-based classifier, the order is meaningless and because the rules are mutual exclusion and exhaustion, no duplication is allowed. And because of the rules are more detail than tree's node, their most important factors are different from each other are reasonable.

The evaluation of this one is showing as Table 4.2.

**Table 4.2**

| threshold | pruned | Accuracy | Kappa |
|---|---|---|---|
| 0.0100 | yes | 0.7000226 | 0.3973009 |
| 0.0100 | no | 0.7520662 | 0.5057001 |
| 0.1325 | yes | 0.7069030 | 0.4124645 |
| 0.1325 | no | 0.7520662 | 0.5057001 |
| 0.2550 | yes | 0.7492244 | 0.4969353 |
| 0.2550 | no | 0.7520662 | 0.5057001 |
| 0.3775 | yes | 0.7351842 | 0.4686577 |
| 0.3775 | no | 0.7520662 | 0.5057001 |
| 0.5000 | yes | 0.7351842 | 0.4686577 |
| 0.5000 | no | 0.7520662 | 0.5057001 |

**Resampling: cross-validated (10 fold). The final values used for the model were threshold = 0.5 and pruned = no.*

The accuracy has obviously improved when compared to the decision tree. The improvement may thanks to the two crucial features, mutual exclusion and exhausted. These two properties work together to ensure that each record is covered by only one rule. No duplication, higher accuracy.

Then, we used the linear support vector machine.

## 3. Linear SVM

Linear SVM is the newest extremely fast machine learning (data mining) algorithm for solving multiclass classification problems from ultra large data sets that implements an original proprietary version of a cutting plane algorithm for designing a linear support vector machine. It creates an SVM model in a CPU time which scales linearly with the size of the training data set. Compared to other SVM methods, it shows a superior performance when high accuracy is required. This method is efficiency in dealing with extra-large data sets. It is a solution of multiclass classification problems with any number of classes and can working with high dimensional data (thousands of features, attributes) in both sparse and dense format. There is no need for expensive computing resources (personal computer is a standard platform) when use linear SVM.[9]

The data of final model is shown as the Table 4.3.

**Table 4.3**

| | |
|---|---|
| **SV type** | C-svc (classification) |
| **parameter** | cost C = 1 |
| **Number of Support Vectors** | 534 |
| **Objective Function Value** | -2678365 |
| **Training error** | 0.467 |

Less training error than the decision tree.

The evaluation of the linear SVM is shown as Table 4.4.

**Table 4.4**

| | |
|---|---|
| **Accuracy** | 0.5089615 |
| **Kappa** | 0.01014418 |

**Resampling: cross-validated (10 fold). Tuning parameter 'C' was held constant at a value of 1.*

From Table 4.3 and 4.4, it is easy to figure out that though the linear SVM model for this case has less training error than decision tree, but its accuracy is the worst one within the three models. This may due to the SMV is more suitable for small samples, grand samples may prone to "overfitting". Another reason may be that the increasing amount of data resulting in skewed data, which causes the classifier to produce errors.[10]

The key factors of this model are showing as Figure 4.4. The top three are SOC Name, worksite state, and country of citizenship.

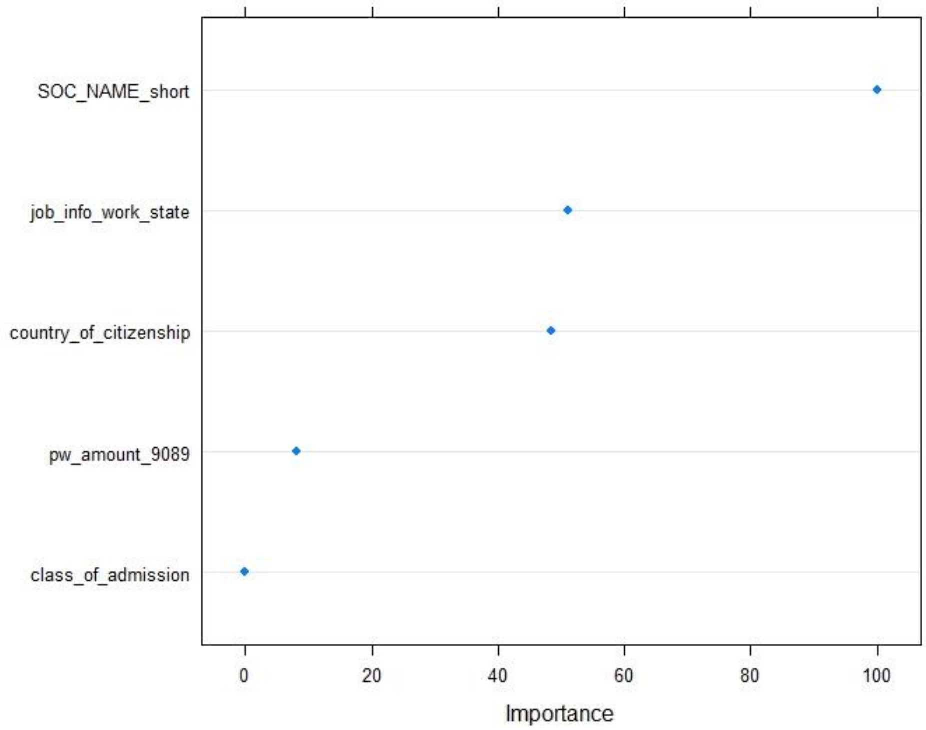


Figure 4.4

It is neither same to decision tree nor PART. But we can also find out that all of them shows the importance of worksite state.

Next, we used the artificial neural network.

## 4. Artificial Neural Network

Artificial neural networks (ANNs) or connectionist systems are computing systems vaguely inspired by the biological neural networks that constitute animal brains.[11] Such systems "learn tasks by considering examples, generally without task-specific programming. ANN is nonlinear model that is easy to use and understand compared to statistical methods. It is also non-paramateric model while most of statistical methods are parametric model that need higher background of statistic. This method can be used to describe relatively simple problems. It is often used to treat super complicated problems, in which too many variables to be simplified in a model.[12]

The final model is a 788-1-1 network with 791 weights.

The evaluation result is listed as Table 4.5.

**Table 4.5**

| size | decay | Accuracy | Kappa |
|---|---|---|---|
| **1** | 0e+00 | 0.5200112 | 0.006554205 |
| **1** | 1e-04 | 0.5321023 | 0.032440627 |
| **1** | 1e-03 | 0.5260417 | 0.019556913 |
| **1** | 1e-02 | 0.5500041 | 0.071125699 |
| **1** | 1e-01 | 0.5220013 | 0.011478416 |
| **3** | 0e+00 | NaN | NaN |
| **3** | 1e-04 | NaN | NaN |
| **3** | 1e-03 | NaN | NaN |
| **3** | 1e-02 | NaN | NaN |
| **3** | 1e-01 | NaN | NaN |
| **5** | 0e+00 | NaN | NaN |
| **5** | 1e-04 | NaN | NaN |
| **5** | 1e-03 | NaN | NaN |
| **5** | 1e-02 | NaN | NaN |
| **5** | 1e-01 | NaN | NaN |
| **7** | 0e+00 | NaN | NaN |
| **7** | 1e-04 | NaN | NaN |
| **7** | 1e-03 | NaN | NaN |
| **7** | 1e-02 | NaN | NaN |
| **7** | 1e-01 | NaN | NaN |
| **9** | 0e+00 | NaN | NaN |
| **9** | 1e-04 | NaN | NaN |

| | | | |
|---|---|---|---|
| **9** | 1e-03 | NaN | NaN |
| **9** | 1e-02 | NaN | NaN |
| **9** | 1e-01 | NaN | NaN |

**Resampling: cross-validated (10 fold). Accuracy was used to select the optimal model using the largest value. The final values used for the model were size = 1 and decay = 0.01.*

The accuracy is little higher than the decision tree, but obviously it not better than rule-based classifier. The low accuracy may be caused by the same reason as the linear SVM as larger training set leading to overfitting. There is another possible reason is that there are some similar training data which may resulting in overtraining.[13]

The crucial factors of the model are showing as Figure 4.5. It shows detailly, and we can still find that the top influent one is state. This is similar to the conclusion we reached via compared the past three models.

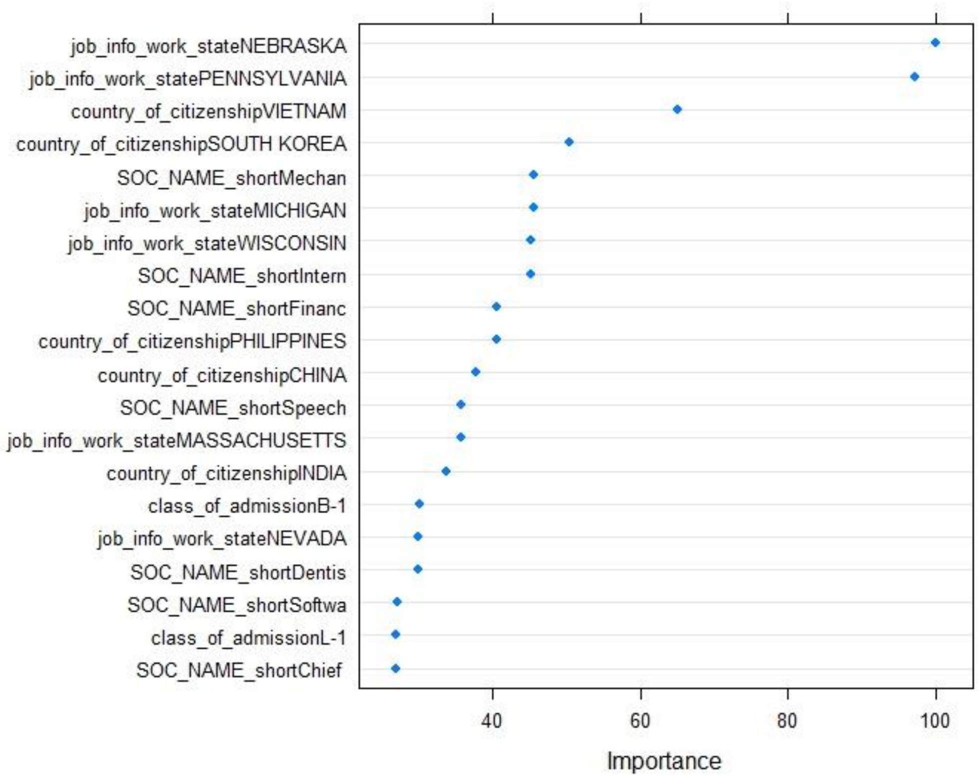


Figure 4.5

The last one we chose is random forests.

## 5. Random Forest

Random forests is an ensemble learning method for classification, regression and other tasks, that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees.[14] Random decision forests correct for decision trees' habit of overfitting to their training set.[15] It is a good algorithm to use for complex classification tasks. The main advantage of this algorithm is that the model created can easily be interrupted.[16] Random Forests train each tree independently, using a random sample of the data. This randomness helps to make the model more robust than a single decision tree, and less likely to overfit on the training data.[17]

The final model of the random forest can be shown in Table 4.6.

**Table 4.6**

| **Call** | randomForest(x=x, y=y, mtry=param$mtry) |
|---|---|
| **Type of random forest** | Classification |
| **Number of trees** | 500 |
| **No. of variables tried at each split** | 787 |
| **OOB estimate of error rate** | 46.9% |

**Confusion matrix**

| | FALSE | TRUE | class. error |
|---|---|---|---|
| FALSE | 252 | 231 | 0.4782609 |
| TRUE | 238 | 279 | 0.4603482 |

The result of evaluation can be listed as Table 4.7.

**Table 4.7**

| **mtry** | **Accuracy** | **Kappa** |
|---|---|---|
| 2 | 0.5170009 | 0.0000000 |
| 8 | 0.5170009 | 0.0000000 |
| 39 | 0.7339836 | 0.4617983 |
| 176 | 0.8950039 | 0.7894549 |

| | | |
|---|---|---|
| 787 | 0.9240255 | 0.8478434 |

**Resampling: cross-validated (10 fold). Accuracy was used to select the optimal model using the largest value. The final values used for the model were mtry=787.*

It seems the best model for this specific case! The Highest accuracy (almost 0.93) with the highest kappa value. This is because that random forest is an ensemble models. Single-model classification models are often inaccurate, such as decision trees, and are prone to overfitting problems. Therefore, combining multiple single-category models to improve prediction accuracy is called classifier ensemble method. And the ensemble model is the result of classifier integration.[18][19]

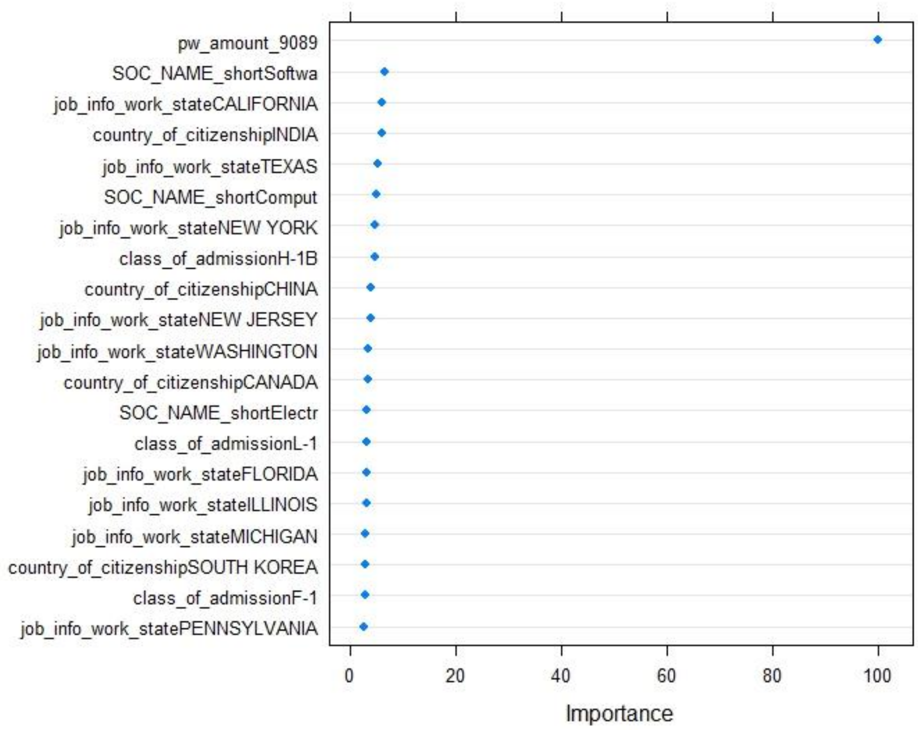


Figure 4.6

The evaluation result is shown as Figure 4.6. It is same to the decision tree, the top one is wage. And worksite state is also the top 3.

# V. Evaluation

From the result of each classification, we can safely achieve the following conclusions:

- The most important factor for certification rate is state and prevailing wage.
- There are different factor influences between different states.

❖ The best model for this specific case is the random forest.

This can also be proven by comparing all models. The result of this comparing is shown as Table 4.8 and 4.9 (a, b).

**Table 4.8**

| **Call** | Summary.resamples (object=resamps) |
|---|---|
| **Models** | ctree, SVM, rules, NeuralNet, randomForest |
| **Number of resamples** | 10 |

**Table 4.9 (a)**

**Accuracy**

| | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. | Na's |
|---|---|---|---|---|---|---|---|
| **ctree** | 0.5148515 | 0.5149265 | 0.5151515 | 0.5170009 | 0.5200000 | 0.5200000 | 0 |
| **SVM** | 0.4700000 | 0.4900000 | 0.5000500 | 0.5089615 | 0.5297780 | 0.5643564 | 0 |
| **rules** | 0.6900000 | 0.7196535 | 0.7587879 | 0.7520662 | 0.7764026 | 0.8080808 | 0 |
| **NeuralNet** | 0.5049505 | 0.5211881 | 0.5300000 | 0.5500041 | 0.5505051 | 0.6732673 | 0 |
| **randomForest** | 0.9100000 | 0.9129663 | 0.9240255 | 0.9240255 | 0.9305198 | 0.9400000 | 0 |
| **Kappa** | | | | | | | |
| | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. | Na's |
| **ctree** | 0.00000000 | 0.000000000 | 0.0000000000 | 0.00000000 | 0.0000000 | 0.0000000 | 0 |
| **SVM** | -0.05633803 | -0.018142418 | -0.0008158063 | 0.01014418 | 0.03610229 | 0.1090617 | 0 |
| **rules** | 0.37751004 | 0.442841486 | 0.5181519700 | 0.50570013 | 0.55272222 | 0.6174497 | 0 |
| **NeuralNet** | -0.01998335 | 0.005250404 | 0.0224618858 | 0.07112570 | 0.07838796 | 0.3337997 | 0 |
| **randomForest** | 0.81956696 | 0.825913916 | 0.8497262896 | 0.84784345 | 0.86110322 | 0.8794212 | 0 |

**Table 4.9 (b) Difference Between Each Two Models**

**Accuracy**

| | ctree | SVM | rules | NeuralNet | randomForest |
|---|---|---|---|---|---|
| **ctree** | | 0.008039 | -0.235065 | -0.033003 | -0.407025 |
| **SVM** | 1.0000 | | -0.243105 | -0.041043 | -0.415064 |
| **rules** | 1.696e-07 | 4.136e-08 | | 0.202062 | -0.171959 |
| **NeuralNet** | 0.8882 | 0.2242 | 1.874e-06 | | -0.374021 |
| **randomForest** | 2.859e-14 | 2.189e-10 | 4.951e-07 | 2.909e-08 | |
| | | | **Kappa** | | |
| | **ctree** | **SVM** | **rules** | **NeuralNet** | **randomForest** |
| **ctree** | | -0.01014 | -0.50570 | -0.07113 | -0.84784 |
| **SVM** | 1.0000 | | -0.49556 | -0.06098 | -0.83770 |
| **rules** | 6.289e-08 | 9.486e-08 | | 0.43457 | -0.34214 |
| **NeuralNet** | 0.8398 | 1.0000 | 1.533e-06 | | -0.77672 |
| **randomForest** | 1.286e-14 | 1.074e-10 | 5.309e-07 | 4.193e-08 | |

From the qualification requirements of the U.S. Permanent visa, it is easily to find that the international students and workers, the companies and organizations who want to hire international employees, and the responding agency of the U.S. government are related to predict the certification rate of the Green Card application: applicants and companies are need to use the predict to know whether there is qualified and how to increase their competition; the department members need to use the predict to preprocess all of the applications, which can reduce their workload—millions of applications are created every year.

*Extra analysis: A survey from Goldman Sachs shows that 900k to 1mn individuals are working under H-1B visas in the US, and that about two-thirds of qualified H-1B visa holders eventually apply for a* *green card.*[20] *This data vividly performing that H-1B visa has connection with the Permanent Resident visa; also, H-1B workers play a key role in the U.S. society.*

*A total of 190098 H1B visa applications were received this year, a decrease of 9,000 from last year. It seems that the probability of a lottery is getting larger.*[21] *Actually, combined with the recent Trump administration's policies and attitude toward immigration, we can predict that the difficulty of applying for H1B will continue to increase each year. The number of applicants will continue to decline. Since green card holders holding H1B visas account for about 75% of all green card applicants, it can be predicted that the application status of H1B visas will also have a great impact on green card applications. The number of applicants who meet the green card application conditions will be reduced, so the number of applicants is also expected to decrease. As the application review becomes stricter, the certified probability is also predicted to decrease. However, those who are waiting for the priority date in employment-Based cases will not be affected for the time being (because they have been scheduled for a long time).*

**A. FINAL ACTION DATES FOR FAMILY-SPONSORED PREFERENCE CASES**

On the chart below, the listing of a date for any class indicates that the class is oversubscribed (see paragraph 1); "C" means current, i.e., numbers are authorized for issuance to all qualified applicants; and "U" means unauthorized, i.e., numbers are not authorized for issuance. (NOTE: Numbers are authorized for issuance only for applicants whose priority date is **earlier** than the final action date listed below.)

| Family-Sponsored | All Chargeability Areas Except Those Listed | CHINA-mainland born | INDIA | MEXICO | PHILIPPINES |
|---|---|---|---|---|---|
| F1 | 08APR11 | 08APR11 | 08APR11 | 15NOV96 | 22JAN06 |
| F2A | 01JUN16 | 01JUN16 | 01JUN16 | 22APR16 | 01JUN16 |
| F2B | 15MAY11 | 15MAY11 | 15MAY11 | 01DEC96 | 15DEC06 |
| F3 | 01FEB06 | 01FEB06 | 01FEB06 | 01SEP95 | 01APR95 |
| F4 | 01OCT04 | 01OCT04 | 01MAR04 | 08JAN98 | 01FEB95 |

### A. FINAL ACTION DATES FOR EMPLOYMENT-BASED PREFERENCE CASES

On the chart below, the listing of a date for any class indicates that the class is oversubscribed (see paragraph 1); "C" means current, i.e., numbers are authorized for issuance to all qualified applicants; and "U" means unauthorized, i.e., numbers are not authorized for issuance. (NOTE: Numbers are authorized for issuance only for applicants whose priority date is **earlier** than the final action date listed below.)

| **Employment-based** | **All Chargeability Areas Except Those Listed** | **CHINA-mainland born** | **EL SALVADOR GUATEMALA HONDURAS** | **INDIA** | **MEXICO** | **PHILIPPINES** | **VIETNAM** |
|---|---|---|---|---|---|---|---|
| 1st | C | 01JAN12 | C | 01JAN12 | C | C | C |
| 2nd | C | 01SEP14 | C | 22DEC08 | C | C | C |
| 3rd | C | 01JUN15 | C | 01MAY08 | C | 01JAN17 | C |
| Other Workers | C | 01MAY07 | C | 01MAY08 | C | 01JAN17 | C |
| 4th | C | C | 15DEC15 | C | 22OCT16 | C | C |
| Certain Religious Workers | C | C | 15DEC15 | C | 22OCT16 | C | C |
| 5th Non-Regional Center (C5 and T5) | C | 22JUL14 | C | C | C | C | 22JUL14 |
| 5th Regional Center (I5 and R5) | C | 22JUL14 | C | C | C | C | 22JUL14 |

*From these data, we can see that for those who apply for a green card (mainly Chinese, Indian, Mexican, and Filipino), another major challenge after the application is to wait for their priority date to become current. For employment-Based cases, CHINA-mainland born is 09/01/14(EB2), 06/01/15(EB3), India is 12/22/08(EB2), 05/01/08(EB3), so for Chinese and Indian applicants, they need to wait a long time.[22] Maybe they can consider FAMILY-SPONSORED. As a spouse of a U.S. citizen, a green card does not need to wait. The priority date for the spouse's application for the green card as the holder of the green card is 09/22/17. It will save 2-3 years for the Chinese and 9 years for the Indian.*

# VI. Reference

[1] Statistical classification (https://en.wikipedia.org/wiki/Statistical_classification)

[2] Stuart J. Russell, Peter Norvig (2010) Artificial Intelligence: A Modern Approach, Third Edition, Prentice Hall ISBN 9780136042594.

[3] Evaluation of classification (https://en.wikipedia.org/wiki/Statistical_classification#Evaluation)

[4] Decision tree types (https://en.wikipedia.org/wiki/Decision_tree_learning#Decision_tree_types)

[5] Advantage of decision tree (https://zhuanlan.zhihu.com/p/31691985)

[6] Xiaoqiang Guang, Research on the Classifying Algorithm based on Decision Tree (http://www.docin.com/p-118475292.html)

[7] Rule-based classifier (http://mlwiki.org/index.php/Rule-Based_Classifier)

[8] Features of rule-based classifier (http://www.cnblogs.com/zengzhihua/p/5458373.html)

[9] Te-Ming Huang, Vojislav Kecman, “Linear Support Vector Machine”, v. 3.0, 2009. (http://linearsvm.com/)

[10] "Artificial Neural Networks as Models of Neural Information Processing | Frontiers Research Topic". Retrieved 2018-02-20. (https://www.frontiersin.org/research-topics/4817/artificial-neural-networks-as-models-of-neural-information-processing)

[11] Relationship between Classification Accuracy and Sample Size of Support Vector Machines (http://muchong.com/html/201303/5628410.html)

[12] What are the advantages of using Artificial Neural Network compared to other approaches? (https://www.researchgate.net/post/What_are_the_advantages_of_using_Artificial_Neural_Network_compared_to_other_approaches)

[13] Error, Accuracy of Neural Network Forecast Foreign Exchange (https://wenku.baidu.com/view/b2741e671ed9ad51f01df241.html)

[14] Ho, Tin Kam (1995). Random Decision Forests (PDF). Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995. pp. 278–282. (http://ect.bell-labs.com/who/tkh/publications/papers/odt.pdf)

[15] Hastie, Trevor; Tibshirani, Robert; Friedman, Jerome (2008). The Elements of Statistical Learning (2nd ed.). Springer. ISBN 0-387-95284-5.

[16] Advantages of random forests (http://www.nickgillian.com/wiki/pmwiki.php/GRT/RandomForests)

[17] What are the advantages/disadvantages of using Gradient Boosting over Random Forests? (https://www.quora.com/What-are-the-advantages-disadvantages-of-using-Gradient-Boosting-over-Random-Forests)

[18] Machine Learning: Ensemble Algorithms (Random Forest, Adaboost) (https://blog.csdn.net/chem0527/article/details/51784621)

[19] Ensemble Model—Random Forest (https://blog.csdn.net/zm714981790/article/details/51262103)

[20] James Schneider, Alec Phillips, Heather Bellini, Heath P. Terry, Simona Jankowski, Mohammed Moawalla, Lara Fourman, Julia McCrimlisk& Love Ghotra (2017 Feb. 3). The H-1B Visa debate: A global FAQ for investors. (pp. 4-5)

[21] USCIS Completes the H-1B Cap Random Selection Process for FY 2019 (https://www.uscis.gov/news/alerts/uscis-completes-h-1b-cap-random-selection-process-fy-2019)

[22] Visa Bulletin For May 2018 (https://travel.state.gov/content/travel/en/legal/visa-law0/visa-bulletin/2018/visa-bulletin-for-may-2018.html)