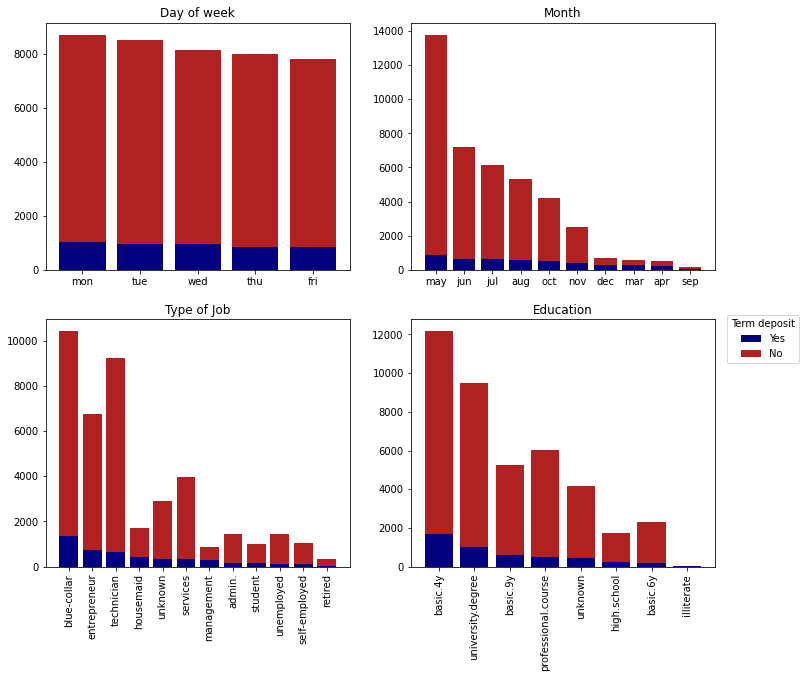
1. Analysis Overview

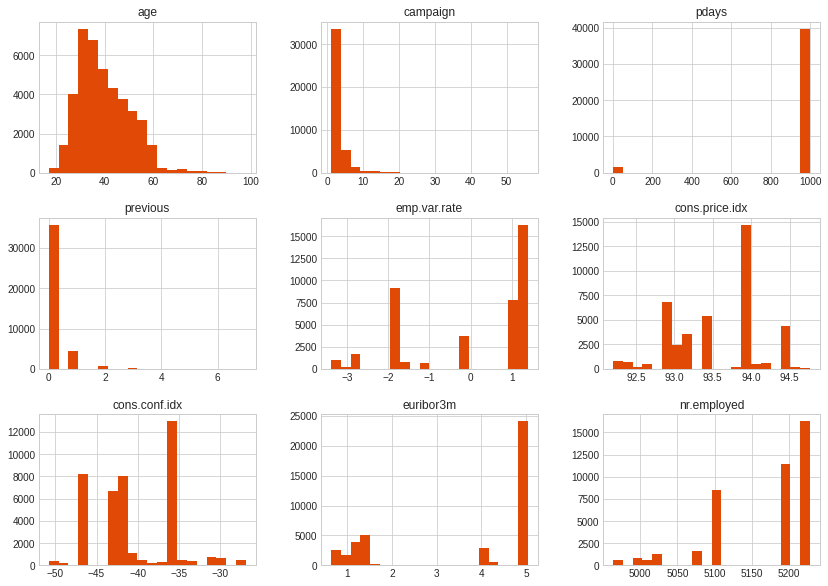
The main goals of this project were to find which factors would contribute the most and the least to customers purchasing a term deposit. Another goal of this project was to determine what would contribute to people not purchasing a term deposit, so that those mistakes will not be made.

1. Targeted Marketing Campaign Models

Checking data if the data is balanced



*These represent how data is balanced for day of the week, month, type of job, and education level*



*These also represent the balance of the data according to their corresponding labels*

As you can see, the data is imbalanced.

* ***Cross Validation***

|  |  |  |
| --- | --- | --- |
|  | **Classifiers** | **Crossval Mean Scores** |
| **5** | Rand FC | 0.864522 |
| **3** | Dec Tree | 0.838216 |
| **2** | KNN | 0.804562 |
| **4** | Grad B CLF | 0.794164 |
| **0** | Logistic Reg. | 0.780483 |
| **6** | Neural Classifier | 0.745929 |
| **7** | Naives Bayes | 0.745929 |
| **1** | SVC | 0.597788 |

## Effects of National Economic Indicators on Campaign Models

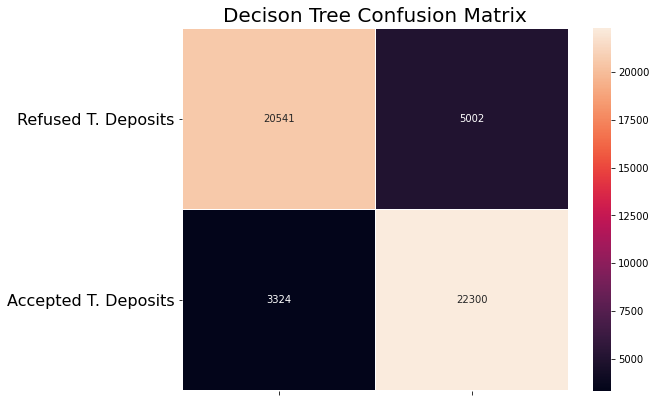
The national economic indicators have significant impacts on the success of campaign to convince customers to accept a term deposit.

Analysis first began by balancing the data to ensure that it would give us unbiased results. I did this by using a couple of data science techniques. The first was balancing data via a method known as SMOTE (Synthetic Minority Oversampling Technique) and cross validation.

***Cross Validation***

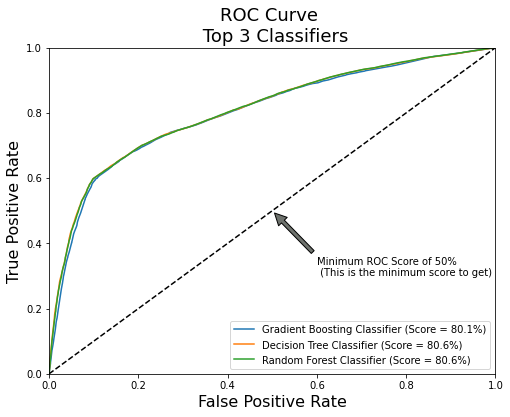
|  |  |  |
| --- | --- | --- |
|  | **Classifiers** | **Crossval Mean Scores** |
| **3** | Dec Tree | 0.749370 |
| **5** | Rand FC | 0.748979 |
| **4** | Grad B CLF | 0.747240 |
| **0** | Logistic Reg. | 0.719702 |
| **1** | SVC | 0.719702 |
| **2** | KNN | 0.703891 |
| **6** | Neural Classifier | 0.560149 |
| **7** | Naives Bayes | 0.560149 |

The next step was to run clean data through a number of models in order to train and prep models to make proper predictions. Here you can see what I call a “confusion matrix”. This specific one relates to the decision tree model and how well it will be able to predict positive and negative outcomes for customers accepting a term deposit.



***Top-left:*** *True Negative* ***Top-Right:*** *False Negative* ***Bottom-Left:*** *False Positive* ***Bottom-Right:*** *True Positive*

Once that training was complete, I looked at the accuracy of each of models to see which will predict the most correct number of accepted term deposits based on the economic factors. I came out with three that I used later to create an ensemble of models. This will be used to verify and enhance models.



*ROC Curves for three different types of classifier models.*

The following graph is one that shows the various economic factors and the impact they have on whether someone will accept a term deposit or not. As shown, the number of employed people in the country has the largest effect while the Euribor 3 month rate is the second most important.



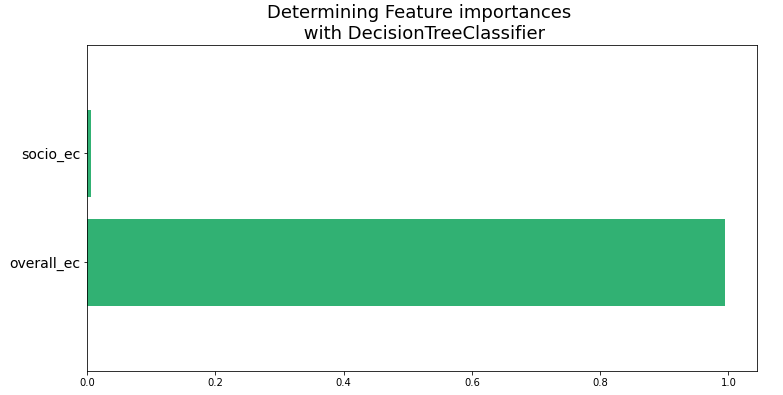
*Decision tree importances where nr.employed is greatest and emp.var.rate is least.*

All of this data modeling allowed us to make the following predictions: the economic factors in play mean that nearly 14000 people will reject a term deposit while just over 8000 will say yes.



*Prediction of how many people will accept a term deposit based on economic factors.*

Based on the graph above, I can conclude that the economic factors do have a significant role when trying to get customers to accept a term deposit. That being said, as shown above, the split is about 2 to 1 in favor of rejecting the term deposit. This is not nearly as high as the raw data, around an 89% rejection rate, which goes to show that many other factors play a role in the decisions made by customers on whether or not to accept a term deposit.



*Description Overall Economy is very high importance and Socio-economic is very low importance*

It appears that overall economic health (99.4) is a much greater predictor for term deposits compared with economic strength of the individual (.004).

1. Prior Marketing Campaign Analysis

From the analysis of the previous marketing campaign, in particular analyzing the decision tree, I realized that the feature that led to the most subscriptions was repeated contact.

Another thing I noticed was that being married or single had almost no effect on whether a customer would buy a term deposit. However, those with lower levels of education, such as a high school diploma or less were more frequent in purchasing.

1. Conclusion and Caveats

From the tests, it appears that overall economic health is the greatest indicator of term deposit acceptance. During times of economic prosperity, blue-collar, basic-4-year-educated workers are most likely to sign up for a term deposit. This is the demographic where I am likely to see the most success from campaign efforts. Twice contacted people are more likely to agree to a term deposit. This reflects the strategy I should use when contacting target demographic.