Building an

Optimal Premium Model in an insurance company

Tasks to be achieved

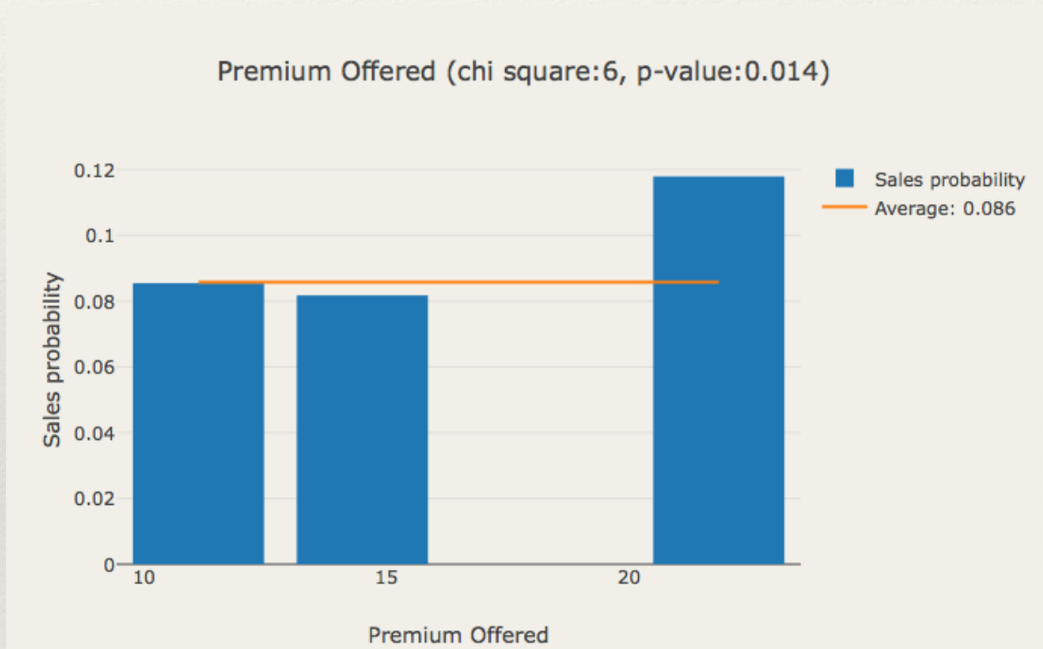
- Finding the ideal target: people who are more likely to contract the insurance.
- Obtaining the optimal price that should be offered to the clients.
- Calculating the difference between offering the premium randomly and optimally using the information obtained in the model.

Data available

- Two databases with clients information are available.
- In the first one we have the information of 20.000 clients which have already been contacted; 9% of them have contracted the product.
- Important data is included such as the premium offered, the number of products that they have already bought, the number of years that they have been clients of the company and the socioeconomic status.
- In the second database of non - previously contacted clients, we have the same information of 10.000 clients but only 5.000 are going to be contacted due to mechanical restrictions.

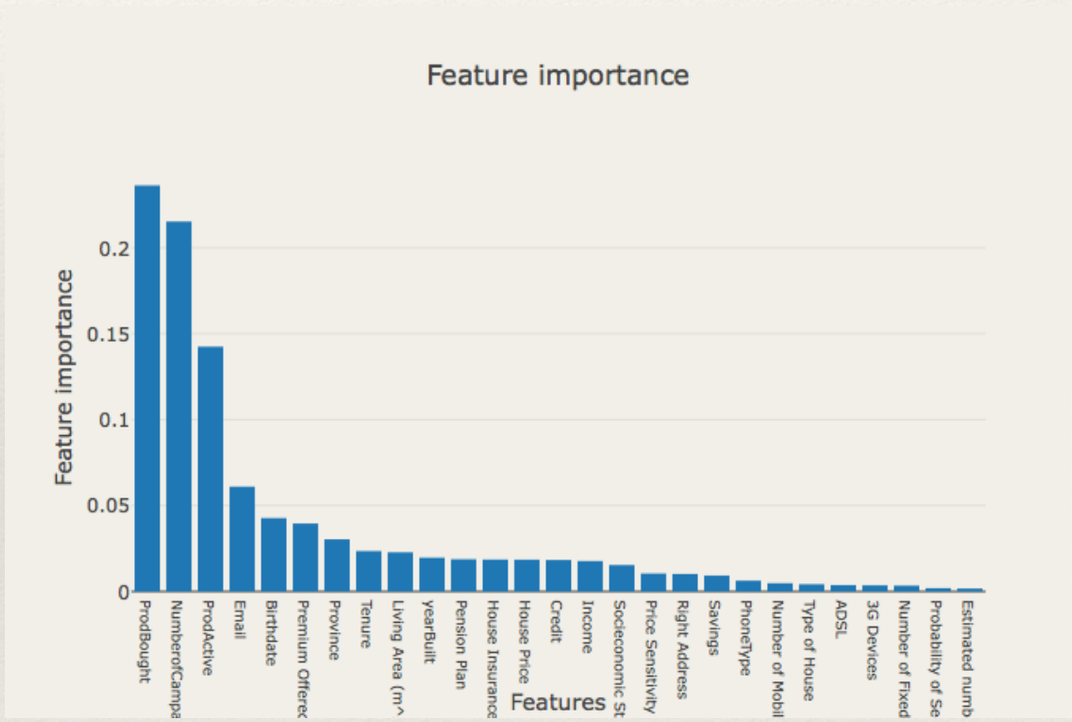
How important is the premium?

- Intuition tells that the higher the premium, the lower the sales.
- But data suggest otherwise: the clients that are more probable to buy the product don't seem to care much about price.
- A more detailed analysis will confirm this.

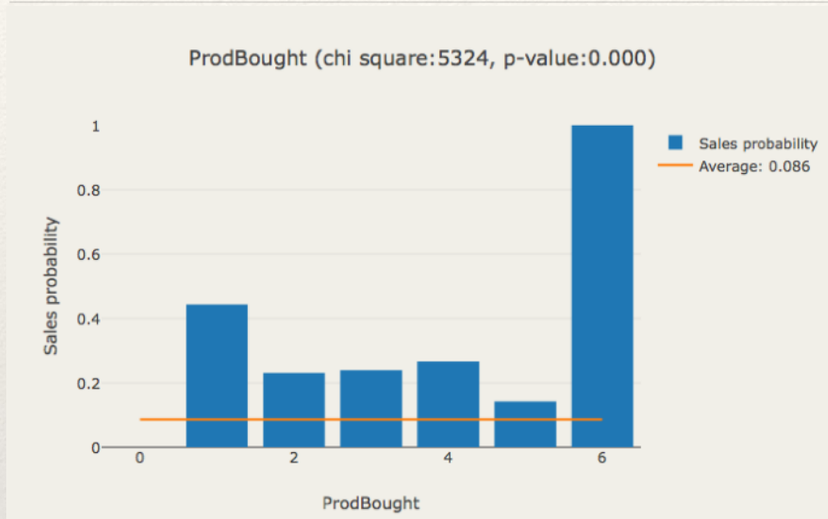


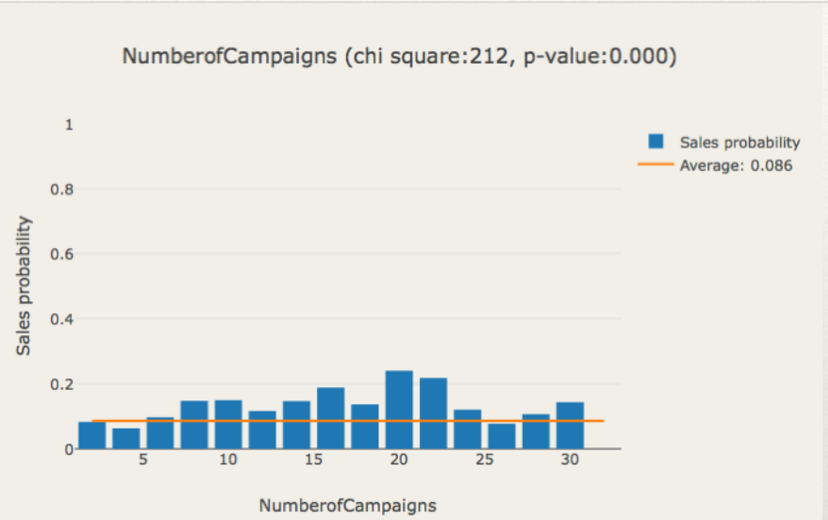
Feature importance

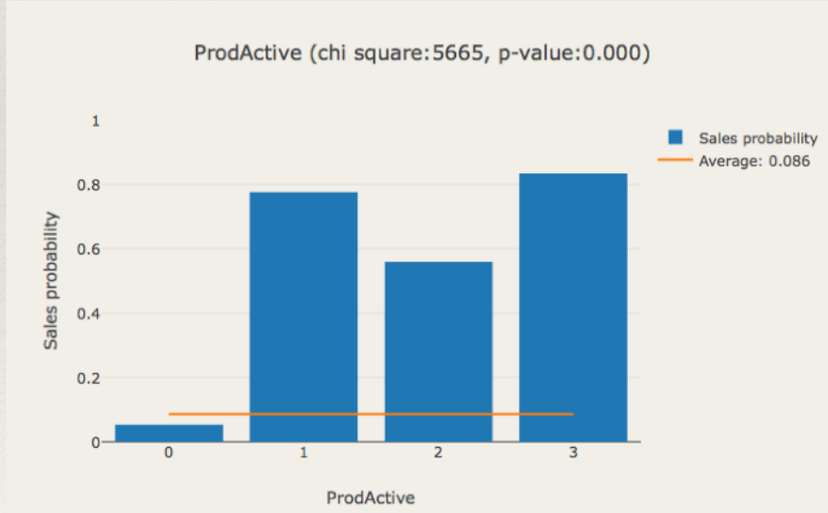
- As a first step, we fit a random forest model with all data, to analyze variable importance, and select only a subset of the features to enter the final predictive model.

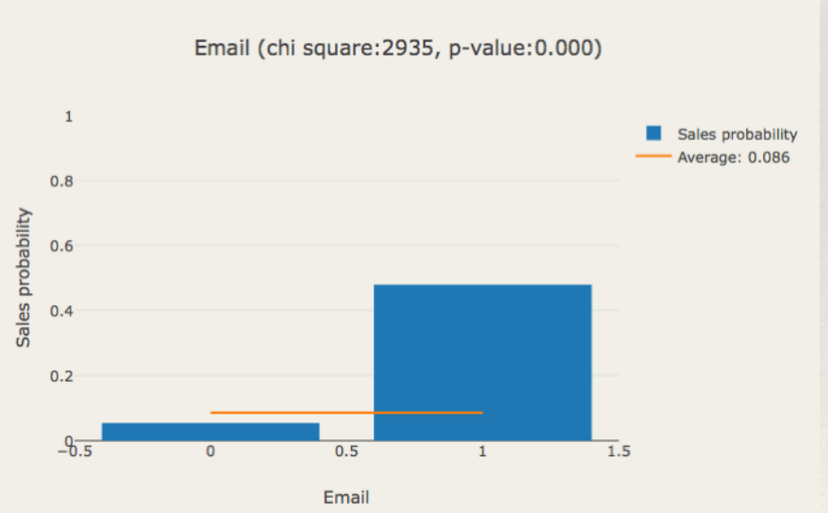


The top four features









Using a neural network to predict sales

- We select the 17 most important variables to fit a neural network to predict the probability of sales.
- The model performs well in a test set, with a 97% of accuracy.
- We will use the model to predict the expected sales conversion, and the target set of clients to call.

Predicting sales

- We can use the model to predict sales probability and expected income for different premiums.
- Sales probability decreases with premium, but too slowly, so we can go with the maximum premium to obtain a maximum expected benefit.
- Calling the subset of clients suggested by the model instead of a random set produces a notable increase in expected profit.

Optimal results

- Using the model to selects a set of users to be called (the ones that the model predicts as having a bigger sales probability), instead of choosing them randomly allow us to double the expected income.
- The reason is that we can discard all users whose sales probability is negligible, and call the ones with a real sales probability.

	Random calls	Selected users
Optimal premium	21,85	21,85
Sales percentage	5,12 %	10,24 %
Income per call	1,12	2,24

Could the premium be even higher?

- A last use of the model would be to predict how far the premium could be increased, where the real maximum income lies.
- The model predicts a maximum of 2.40 around a premium of 28.0.
- Nevertheless, the model has no data in this region of the premium, and its predictive power may be lower, so we would not advice to offer this higher premium without further field studies.