

Fraud Detection



Issues & Objective

- A major insurance company in Singapore used to manually examine each travel insurance claim to identify potentially fraudulent one
- Suspicious claims were subject to a more detailed investigation
- The project objective was to develop a score to identify potentially fraudulent claims which would be subject to greater scrutiny.



Solution

- Adaboost - short for Adaptive Boosting, a powerful machine learning algorithm was used for detecting potentially fraudulent cases
- Substantial lift demonstrated. – on the below test data set it sufficed to examine 7.75% of all claims to identify 91.67% of all fraudulent claims



Challenges

- Data included 77,445 claim records of which only 120 had been determined to be potentially fraudulent
- So identified potentially fraudulent claims are rare events (0.15%) and therefore hard to detect
- It was however expected that there could be a large number of undetected fraudulent claims



Results

Results				
Score (Probability of Claim being Fraudulent)	Potentially Fraudulent	Not Fraudulent	Cumulative (Potentially Fraudulent)	Cumulative (Total)
>50%	0	1	0.00%	0.00%
40-50%	0	19	0.00%	0.09%
30-40%	4	163	33.33%	0.83%
20-30%	7	1,551	91.67%	7.75%
10-20%	1	12,536	100.00%	63.41%
0-10%	0	8,243	100.00%	100.00%