



# Credit Card Default Prediction

---

Descriptive Statistics and Machine Learning

# Content

---

- What are the best client profiles?  
And the worst?
- Can we predict credit default?
- Predictive models.



# Dataset

- Bank from Taiwan;
- 30,000 clients;
- Payments and bills during 6 months;
- Credit card limits;
- Personal Data (Gender, Marital Status, Age, Education)



# Client Risk Profile

---

- Levels of Education seems not to affect the risk;
- Women are less likely to default than men;
- Customers with higher credit card limits have less probability of default;



# Best Client Profile

---



**Gender:** Female

**Marital Status:** Single

**Age:** 31 – 40

**Education:** Bachelor degrees or higher  
qualification



# Worst Client Profile

---

**Gender:** Male

**Marital Status:** Married

**Age:** 51 – 60

**Education:** Bachelor degrees



Can we predict credit  
card default in order to  
avoid it?



# Predictive Models

---

- Random Forest;
- Decision Tree;
- Logistic Regression;
- Support vector machines;
- KNN;
- Gaussian Process Classification;
- AdaBoostClassifier;
- BaggingClassifier;
- ExtraTreesClassifier;
- GradientBoostingClassifier;
- Neural Network Models;





# Accuracy

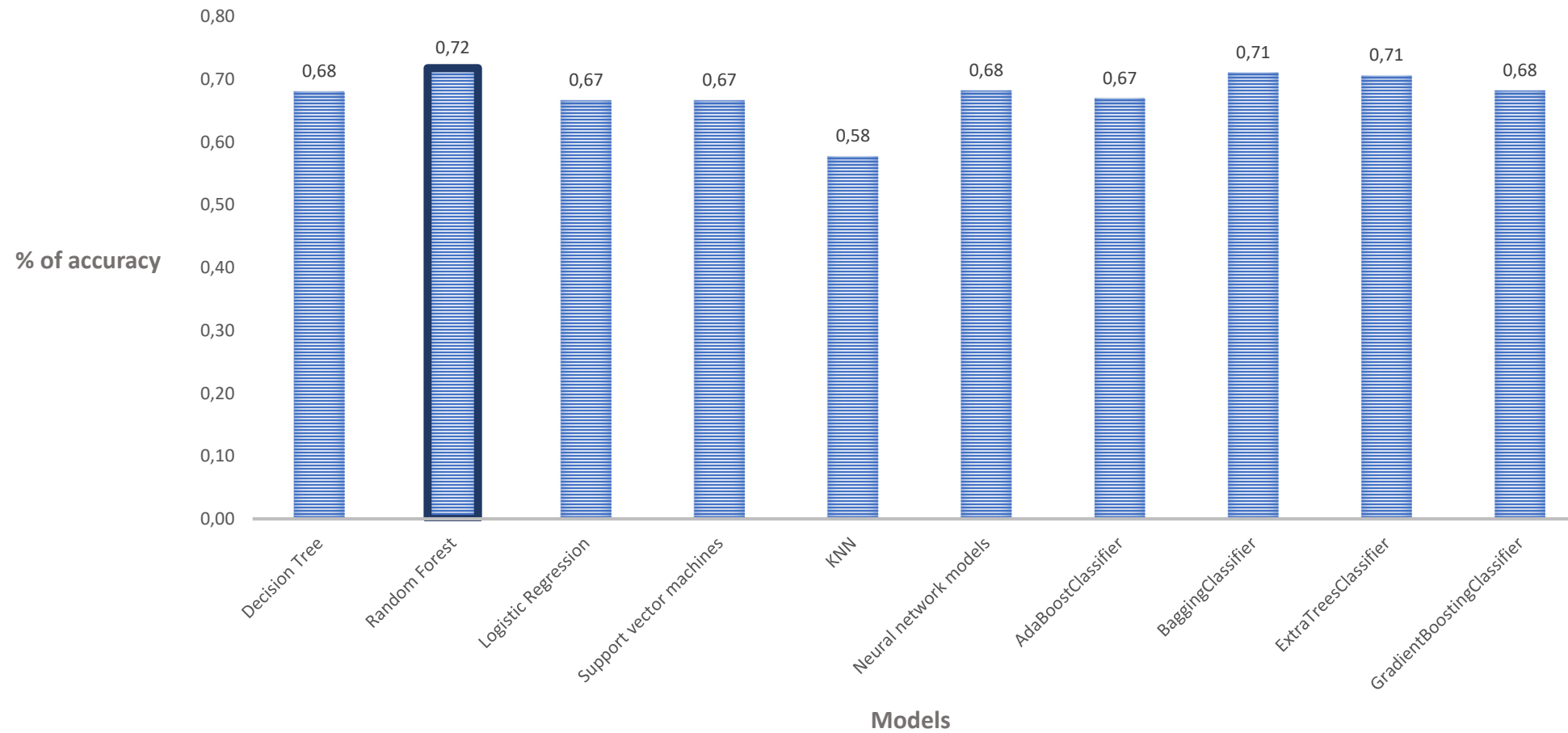
---

## Random Forest

It is possible to predict 72% of the people who are going to default.

# Predictive Models Results

---



# Limitations & Improvements

---

*“We don't have better algorithms, we just have more data” – Peter Norvig (Google)*





CTMX	0.45	▲	+0.45%
FTR	-0.23	▼	-2.34%
CSCO	-1.01	▼	-1.80%
CHK	0.02	▲	
AAPL	+2		
PRTG			
AMZN			
TSLA			
AVGO			
SIRI	0.65		

Thank you for  
your attencion!