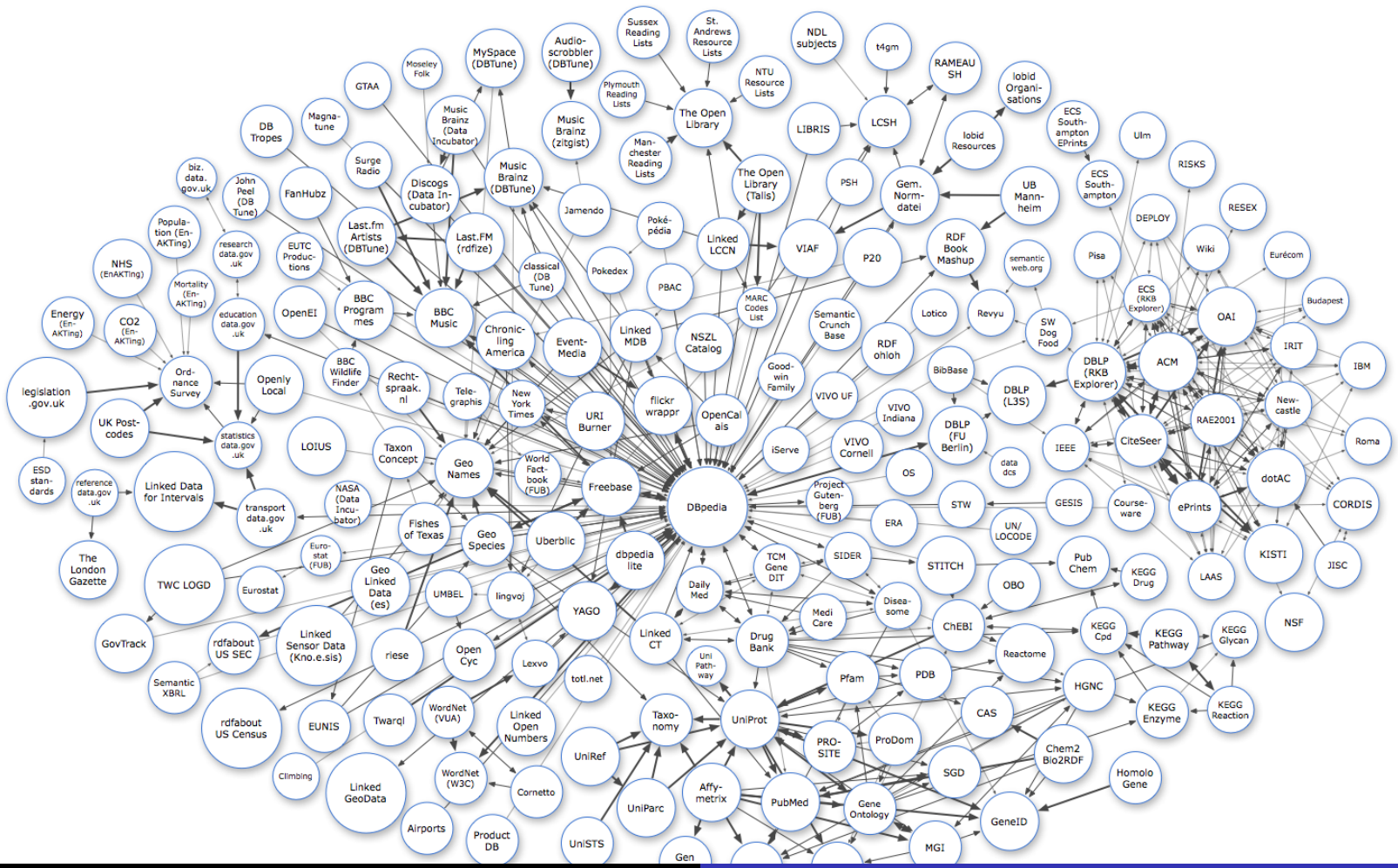
No BI without Machine Learning

Francis Pieraut
francis@qmining.com
http://fraka6.blogspot.com/

10 March 2011
MTI-820 ETS

Too Much Data



Why a talk about machine learning and BI?

- Machine Learning $\Rightarrow$ Data-Mining $\Rightarrow$ BI
- Prediction or Clutering $\Rightarrow$ Patterns $\Rightarrow$ Patterns (revenus \$\$ $\Uparrow$)

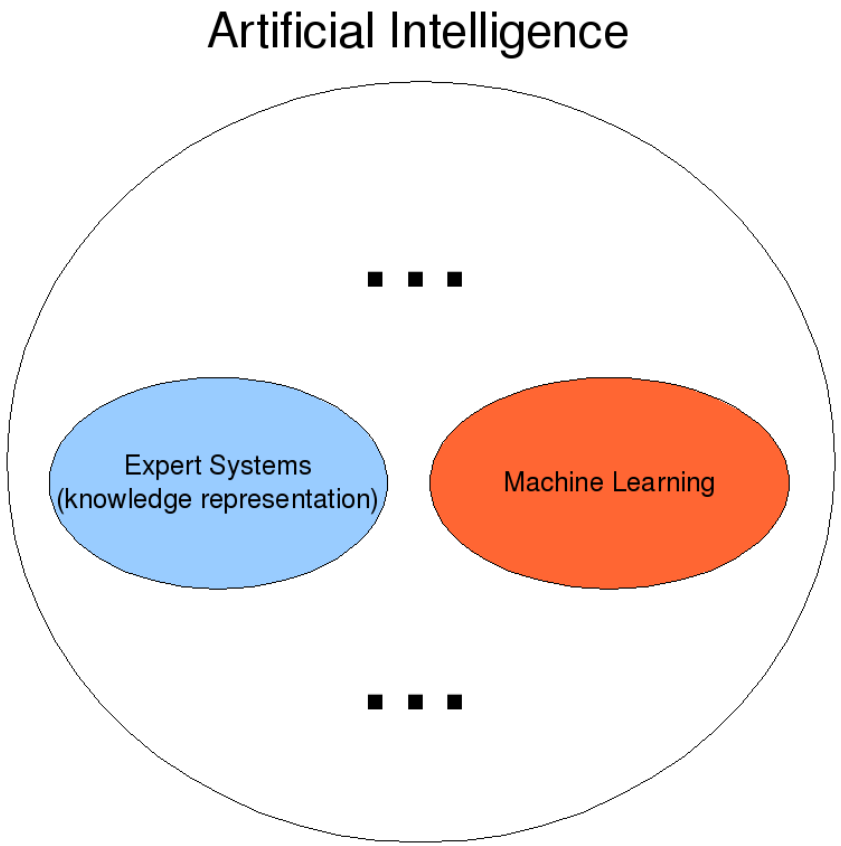
Speaker: Francis Pieraut, P.Eng. M.Sc.A.

- Master@LISA - Statistical Machine Learning - udm
  (flayers: C++ Neural Networks lib)

- Industry - 7 years in Machine Learning/AI startups
  (mlboost: Python Machine Learning Boost lib)

- Founder QMining 

AI and Machine Learning - Data-mining

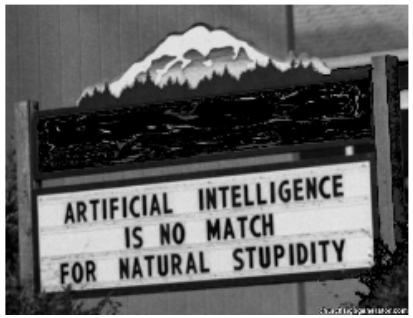


Machine Learning and Data-Mining

- Machine Learning: learn from data
- Data-mining: extracting patterns from data
- Machine Learning use extracted patterns to do prediction

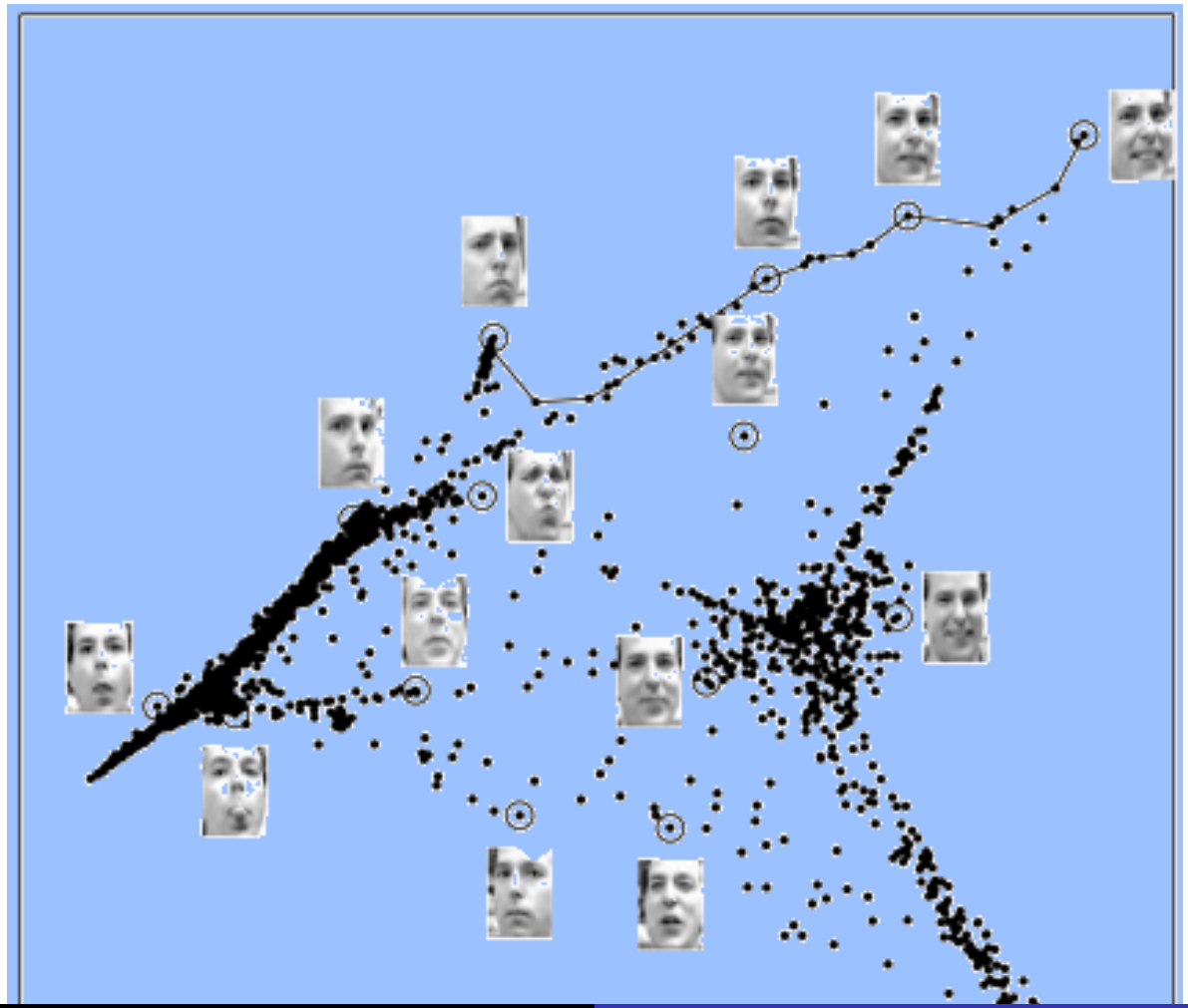
Machine Learning

- Learning from data
- Classification vs Clustering
- Applications: Attrition, Rank Customer (approve loans and credit card),Fraud detection, Target-Marketing, Risk Analysis (insurance) etc.



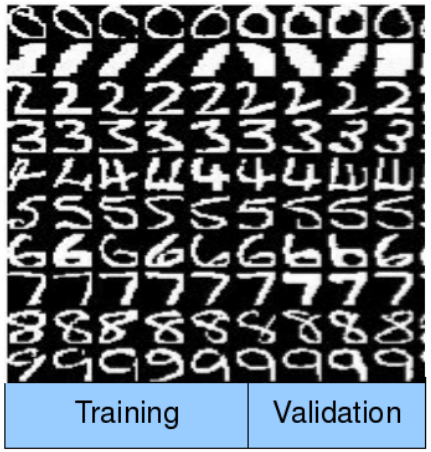
Supervised Learning (need class tag for each example)

Unsupervised Learning - dimension reduction/clustering

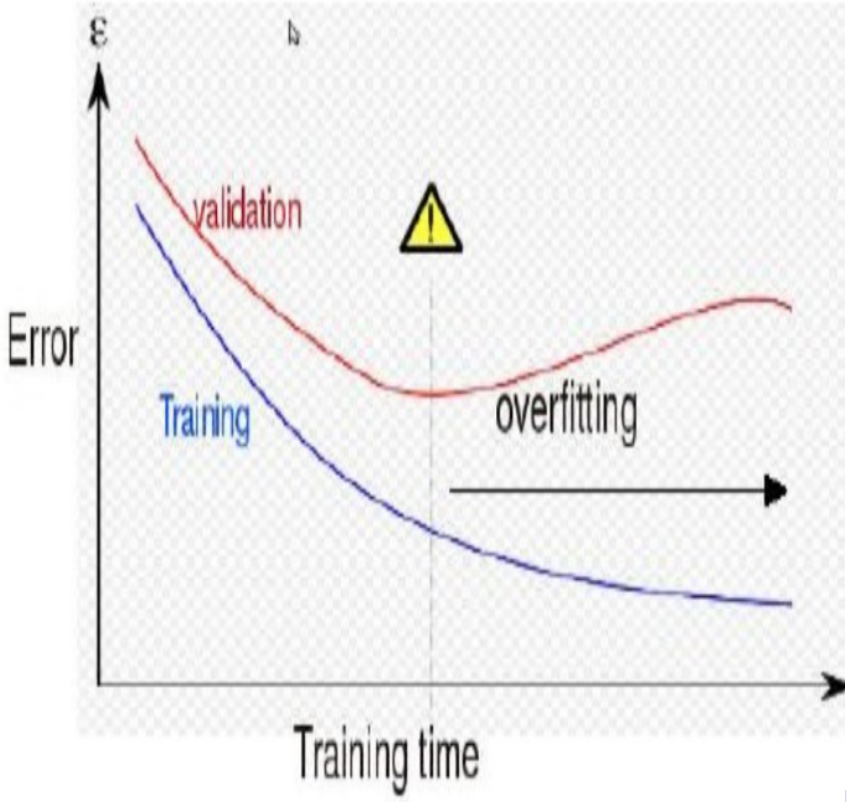


Learning Process

Dataset split



Classification error over time



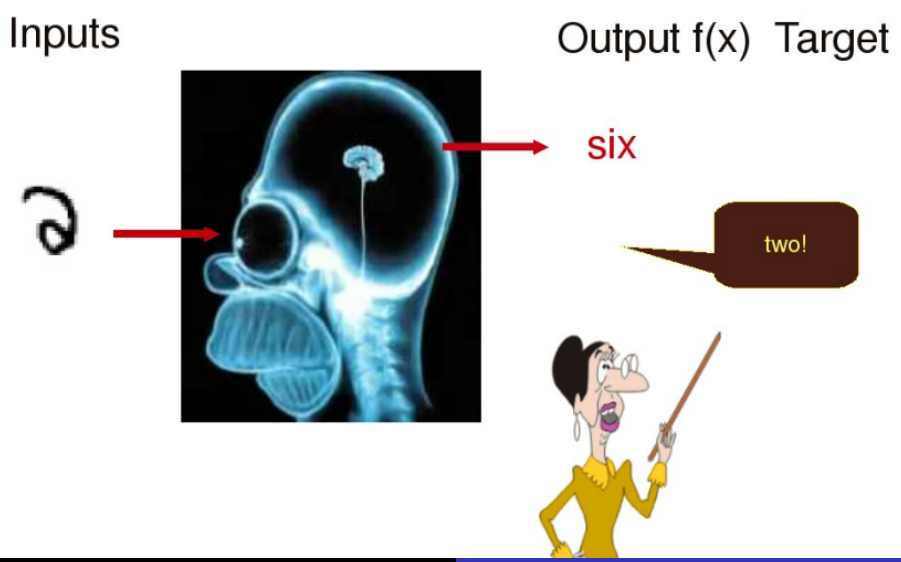
Tanks in the desert

- Using ML requires insights
- An algo is only goods as its data

Important Concepts

- Datasets (features + class)
- Generalization vs Overfitting
- Classification vs Clustering
- Features Quality *(invariant and informative)*

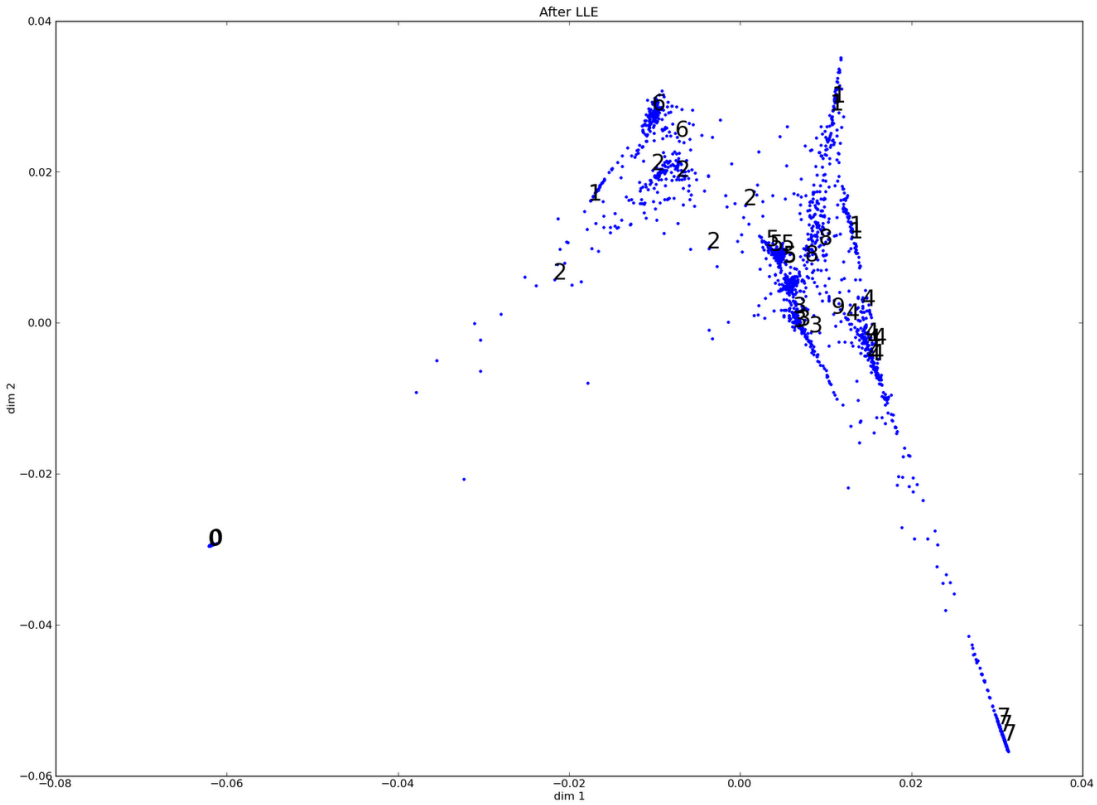
Example: inputs + Target



Bell Canada

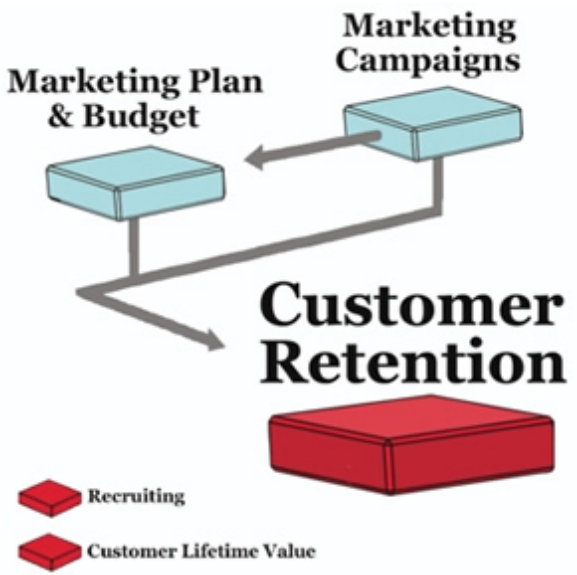
- Find most probable interested clients
- 10000 most likely to buy
- google mail

Microcell labs (Fido-Rogers)



PivotalPayments

- Find most probable clients to quit



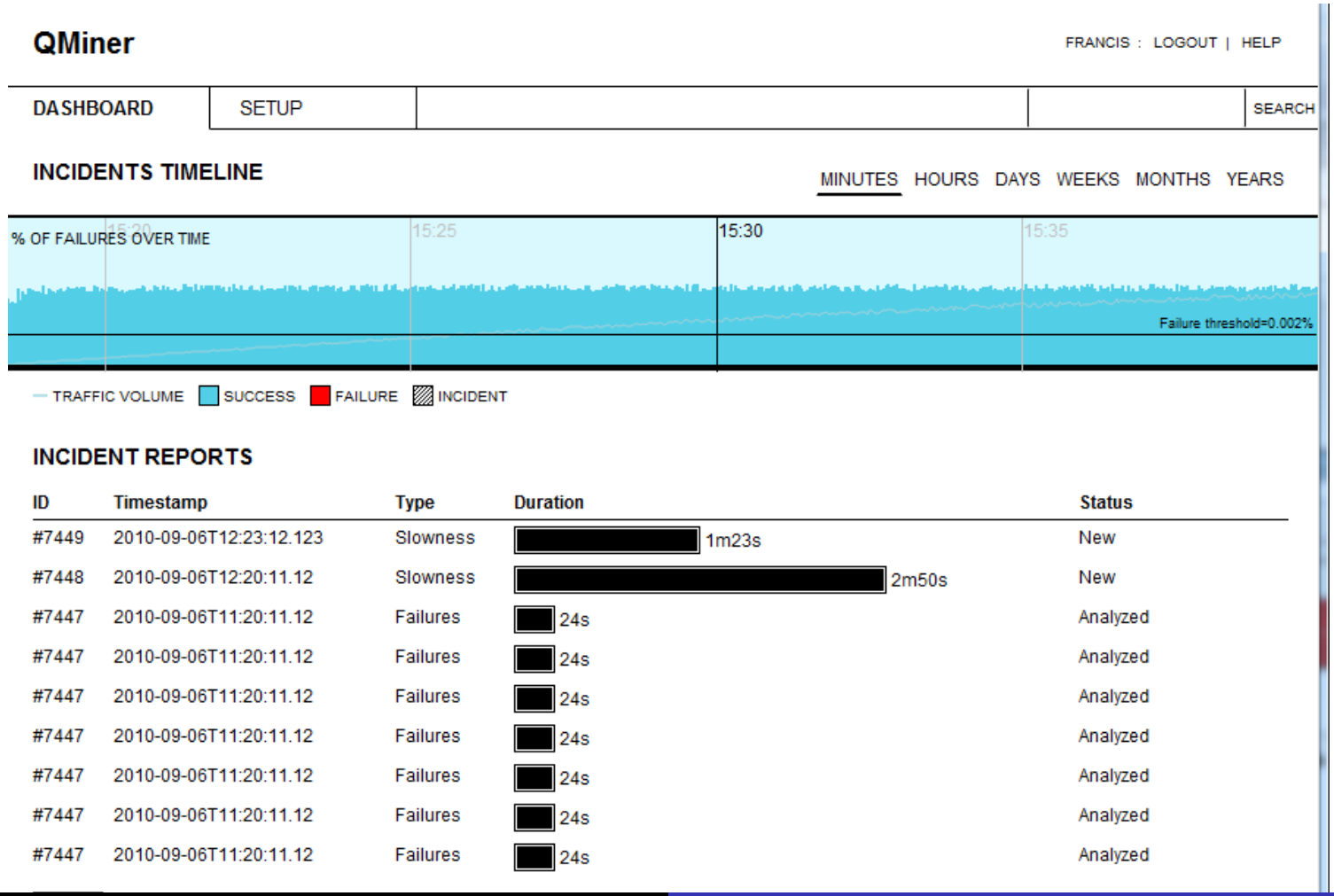
Insurance

- Score customer risk of making a claim

Ubisoft-QMonitor

- Find incidents on servers
- Find patterns (network, server, etc.)

QMiner - Global User Experience Incident Mining



What you should remember?

- No BI without Machine Learning
- Machine learning algorithms applications ⇑
- goal = generalization⇒good prediction (DON'T OVERFIT)
- 80-90% pre or post-processing + data visualization
- Python provide amazing integration
- **QMining is looking for intership students
- BI for Business User http://www.qlikview.com/

Any questions?

...

intership 2011 $\Rightarrow$ francis@qmining.com

http://fraka6.blogspot.com/

..

Thanks,

Francis Pieraut

francis@qmining.com