

# HOW TO TRAIN YOUR **MACHINE** using python™



**By: Ashish Pagey**

# ABOUT ME



# ABOUT MERRILL CORPORATION




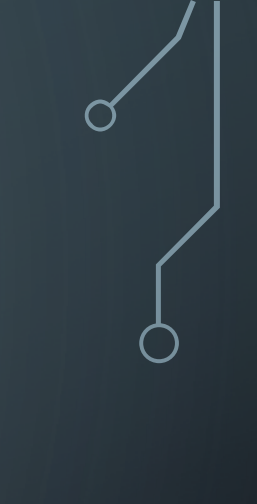
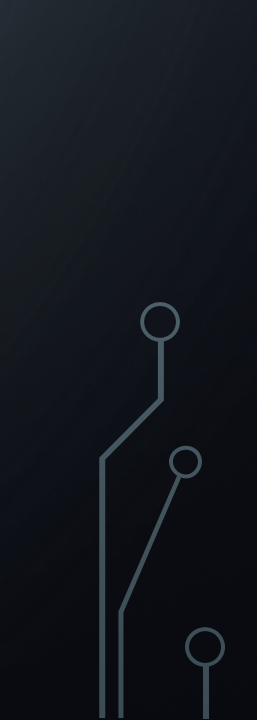
- 50 year history! Now, SaaS + Service!
- Financial Transactions & Reporting
- Marketing & Communications for Regulated Industry
- Customer Content & Collaboration

**psst We're looking for development, DevOps and product management talent! Are you awesome? Contact us at [Tami.Gold@MerrillCorp.com](mailto:Tami.Gold@MerrillCorp.com)**





# QUICK POLL

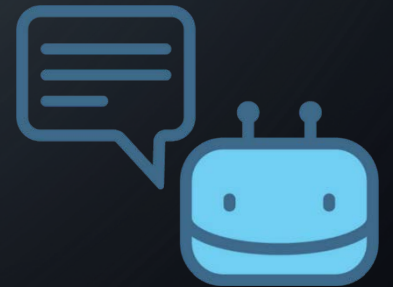
- What is your level of understanding about Machine learning?
    - Heard the Buzz
    - Played with some demo code
    - Used is actively on a project
    - Wrote a book on it
- 
- 
- 

# THE BUZZ

amazon alexa



Siri



Chatbots

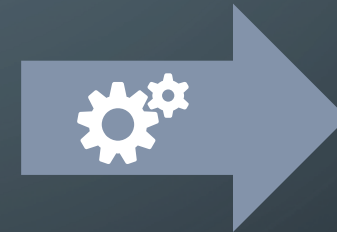
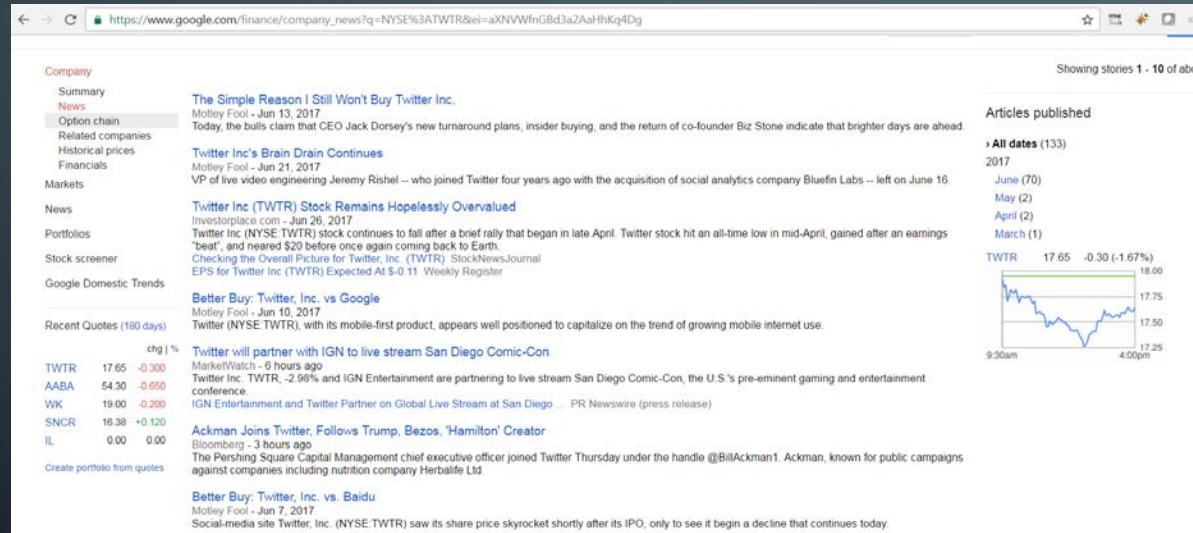
# WHY CARE ABOUT MACHINE LEARNING

- Developers love solving problems. ML enables you to tackle a whole new set of problems
- Increasingly the data you work with won't be in a nice structured RDBMS
- ML is not just for Data Scientist. Programmers should learn it just as they learnt structured programming, object oriented, client server, n-tier, Web 2.0, cloud,...



# TODAY'S OBJECTIVE

Train a learning model to make buy/hold/sell recommendation



buy

hold

sell

DISCLAIMER: This is a demo only. It is not quite reliable to bet your retirement !

# LET'S TRAIN THE MACHINE!





# MACHINE LEARNING APPROACHES

## Supervised learning

**Supervised learning** is the machine **learning** task of inferring a function from labeled training data. The training data consist of a set of training **examples**. In **supervised learning**, each example is a pair consisting of an input object (typically a vector) and a desired output value (also called the supervisory signal)

**Examples:** classification, regression

## Unsupervised learning

Unsupervised machine learning is the [machine learning](#) task of inferring a function to describe hidden structure from "unlabeled" data (a classification or categorization is not included in the observations). Since the examples given to the learner are unlabeled, there is no evaluation of the accuracy of the structure that is output by the relevant algorithm—which is one way of distinguishing unsupervised learning from [supervised learning](#).

**Examples:** clustering, anomaly detection

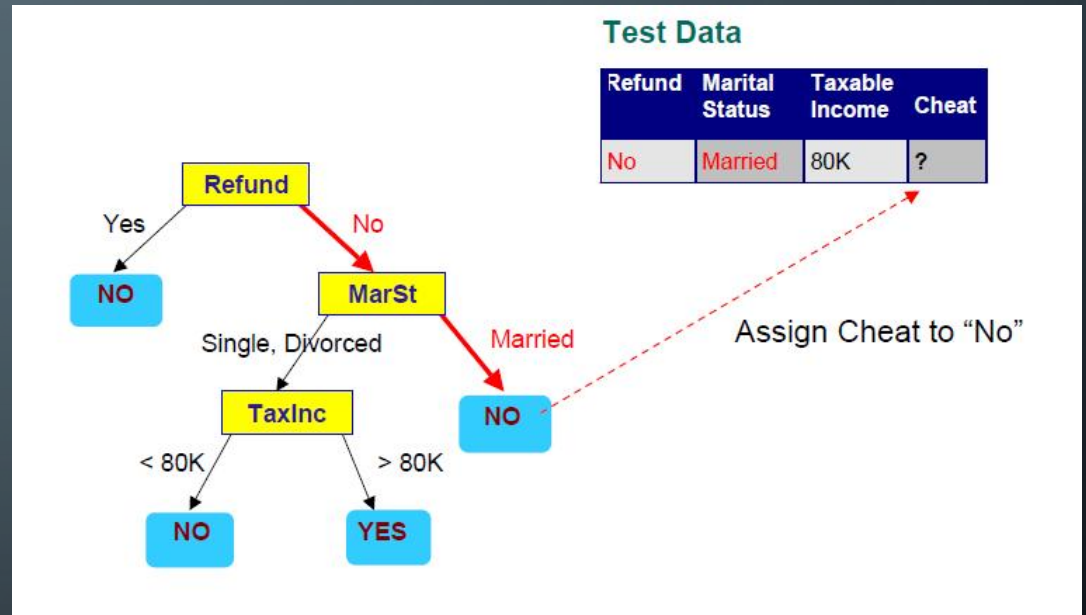
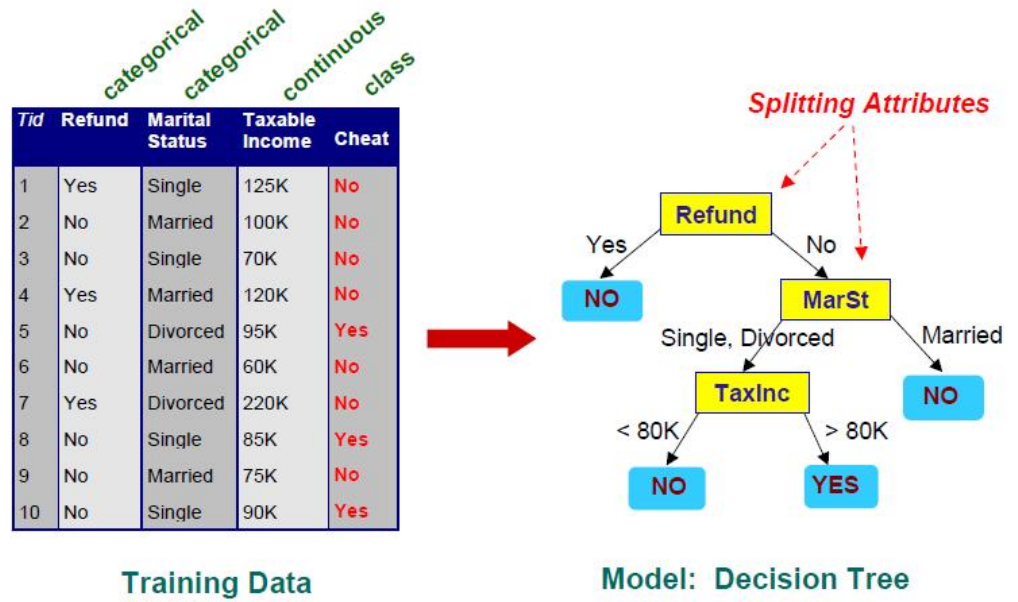
# NAÏVE BAYES

Example: a fruit may be considered to be an apple if it is red, round, and about 3" in diameter.

A Naive Bayes classifier considers each of these “features” (red, round, 3" in diameter) to contribute independently to the probability that the fruit is an apple, regardless of any correlations between features. Features, however, aren't always independent which is often seen as a shortcoming of the Naive Bayes algorithm and this is why it's labeled “naive”.

source: <http://blog.aylien.com/naive-bayes-for-dummies-a-simple-explanation>

# DECISION TREE CLASSIFIER



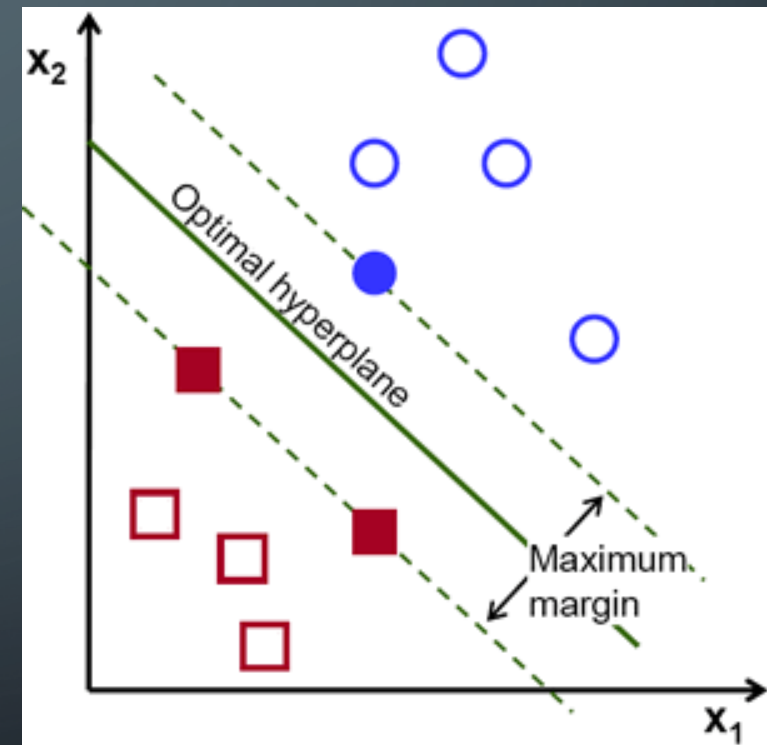
Source:

[http://mines.humanoriented.com/classes/2010/fall/csci568/portfolio\\_exports/lguo/decisionTree.html](http://mines.humanoriented.com/classes/2010/fall/csci568/portfolio_exports/lguo/decisionTree.html)





# SUPPORT VECTOR MACHINE

A **Support Vector Machine (SVM)** is a discriminative classifier formally defined by a separating hyperplane. In other words, given labeled training data (supervised learning), the algorithm outputs an optimal hyperplane which categorizes new examples.





# IMPROVING ACCURACY

- More training data
  - Cleaner training data
  - Refine feature extraction
  - Tune model parameters
  - Try different models
- 
- 

# CONTINUE THE EXPLORATION...

- SK Learn : <http://scikit-learn.org/stable/>
- Tensorflow: <https://www.tensorflow.org/>
- MonkeyLearn: <https://monkeylearn.com/>
- Stanford CoreNLP: <https://stanfordnlp.github.io/CoreNLP/>



THANKS FOR LISTENING

