Decoding Machine Learning Algorithms

Tvisha Gupta, Satraajeet Mukherjee, Prabakar Rajasekaran
Ernst & Young, LLP

About the speakers

- **Tvisha Gupta**
- Actuarial Analytics Consultant
- Tvisha works in the Actuarial Analytics team in EY Bangalore, applying analytics in various areas like cyber risk, claims, etc.

- **Ernst & Young (EY) LLP**
- EY is a leading provider of actuarial and analytics consulting services to insurance companies and other firms worldwide
- EY is recognized as a leader in Gartner's 2018 Magic Quadrant, in providing Analytics & Artificial Intelligence services globally

About the speakers

- **Satraajeet Mukherjee**
- Lead – Actuarial Analytics
- Satraajeet leads the Actuarial Analytics team in EY Bangalore and is currently leading analytics projects from EY New York

- **Ernst & Young (EY) LLP**
- EY is a leading provider of actuarial and analytics consulting services to insurance companies and other firms worldwide
- EY is recognized as a leader in Gartner’s 2018 Magic Quadrant, in providing Analytics & Artificial Intelligence services globally

What is Machine Learning?

- An application of artificial intelligence that provides systems the ability to automatically learn, act and improve from experience (just like humans) without being explicitly programmed.
- Relies on data and information fed in the form of observations and real world interactions.

Introduction

- Machine learning vs Statistics
  - Learning ability
  - Similarities
  - Differences
  - Black Box?

Why Machine Learning?

- Problems that are difficult to be formulated in well defined algorithms.
- Problems involving analysis of large and complex datasets, with different kinds of relationships amongst the variables.
- Problems involving the need for adaptive models.

Need to decode Machine Learning

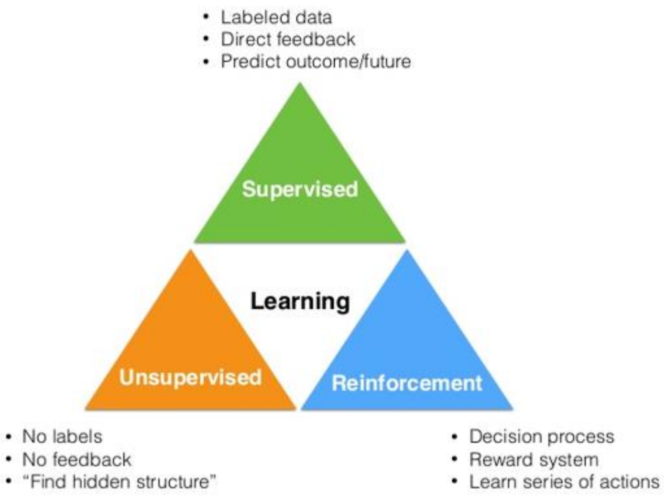
- Why is there a need to decode Machine Learning Algorithms?
  - Interpretability – what is the formula?
  - How was an observation predicted?
  - What features matter the most?

How to decode Machine Learning

- Understand the data
- Understand how the machine learns!

Types of Machine Learning

- Three broad categories of machine learning:
  - Supervised Learning
  - Unsupervised Learning
  - Reinforcement Learning



Supervised Learning

- Most common form of machine learning used in practice
- Framework of supervised learning
  - Determine the appropriate training set.
  - Determine the features that represent the outcome in the training set.
  - Determine the structure of the learned function and the learning algorithm.
  - Complete the design and run the algorithm on the training set.
  - Judge the model performance.

Supervised Learning

- **No Free Lunch Theorem**
  - No perfect model or algorithm exists that works well for all datasets.
  - In other words, for every learner, there is a task on which it fails, even though there is some other learner that works well on it.
  - Consequence – While choosing an appropriate algorithm, assumptions about the target variable need to made. In the absence of any assumption, no scientific method performs better than the random choice.

Supervised Learning

- Tradeoffs and choices to be made
  - Bias-variance tradeoff
  - Function complexity vs amount of training data
  - Choice of error
  - Noise in output
  - Bias-complexity tradeoff
  - Redundancy in the data
  - Feature selection and dimensionality reduction

A bit more to consider in supervised learning

- Parameter tuning
  - Model is defined by the parameters it uses!
  - Difference between parameters and hyperparameters
- Ensembling
  - One predictor is not enough. So multiple predictors are brought together to perform a better job.
  - Bagging vs boosting

Some popular ML algorithms

ICA CIA BERLIN 2018

- Random Forest
- Support vector machines
- Gradient boosting method

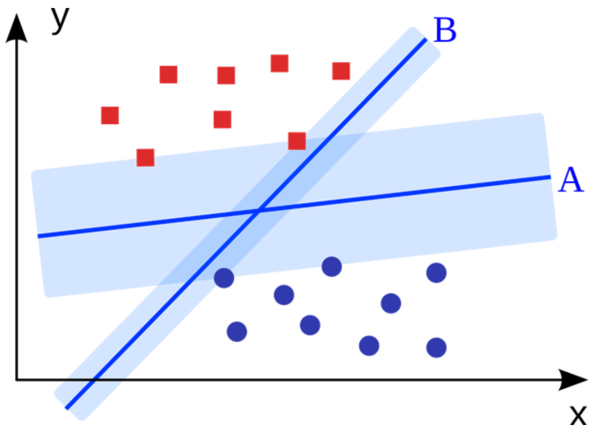
Supervised Learning

- Clustering
- Principal Component Analysis
- Expectation maximization
- Anomaly detection

Unsupervised Learning

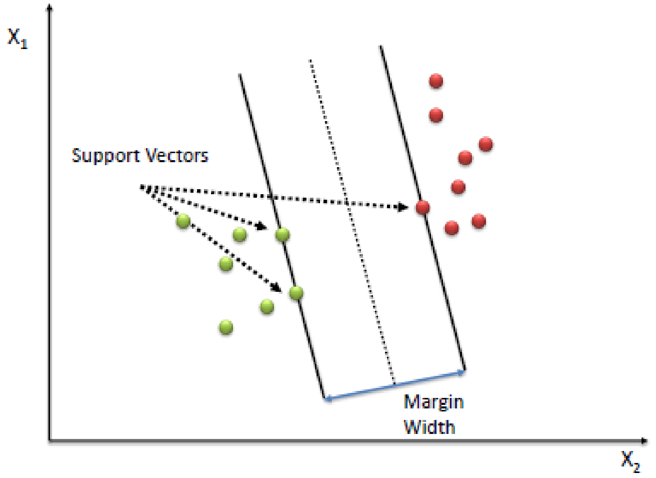
Case Study: Support Vector Machines

- What is a support vector machine?
  - A supervised learning algorithm that works on the principles of creating a margin and separating data with a large gap.



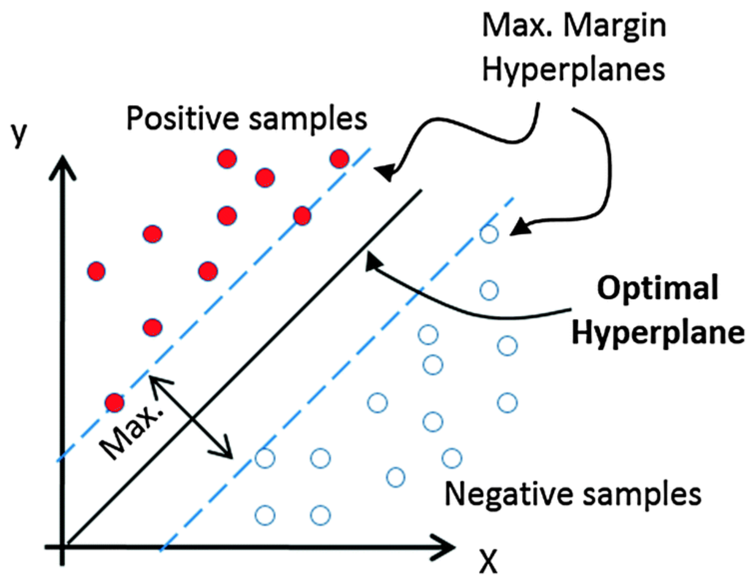
Case Study: Support Vector Machines

- Decoding Support Vector Machines:
  - Support Vectors
  - Feature Selection



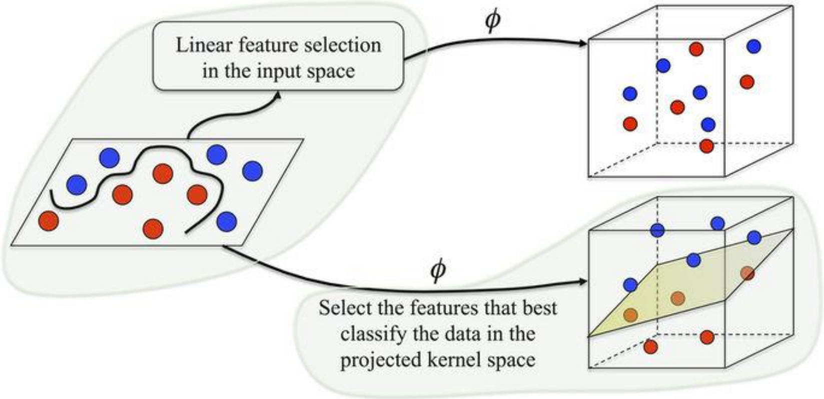
Case Study: Support Vector Machines

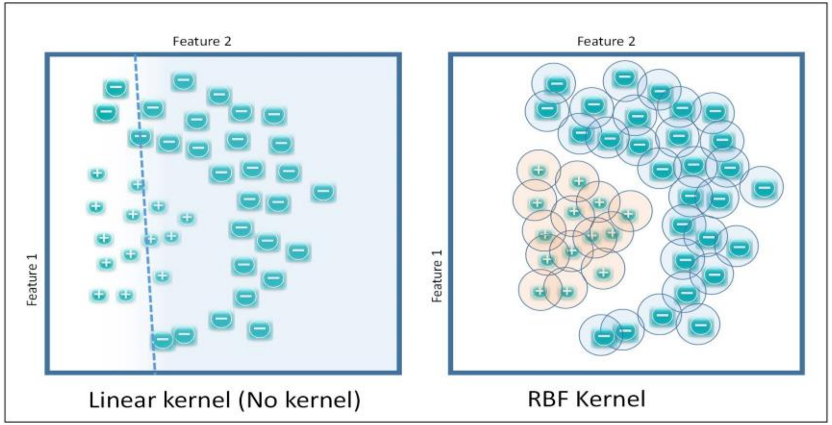
- Decoding Support Vector Machines:
  - Separating hyperplanes
  - Maximum margin



Case Study: Support Vector Machines

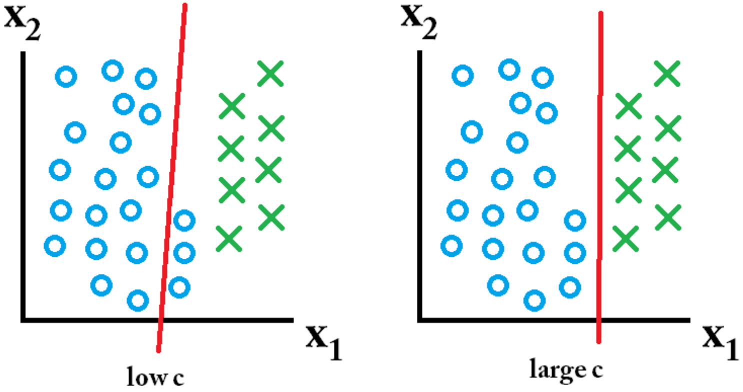
- Decoding Support Vector Machines:
  - Separating data non-separable by a linear line – use kernels.





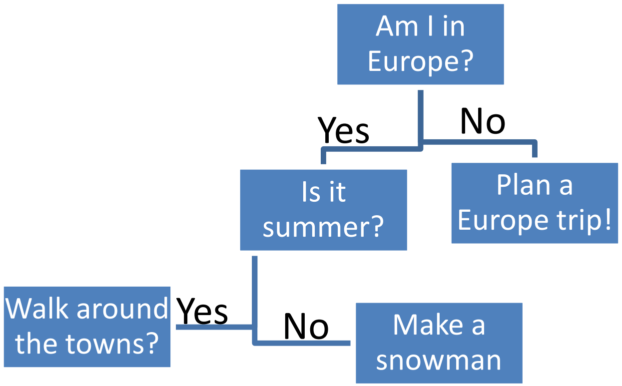
Case Study: Support Vector Machines

- Decoding Support Vector Machines:
  - When perfect separation is not possible – accept some error and tune the acceptance limit.



Case Study: Tree based ML algorithms

- Trees are a popular component of ML algorithms.
- Used widely in the form of decision trees or as a part of random forest and gradient boosting method.

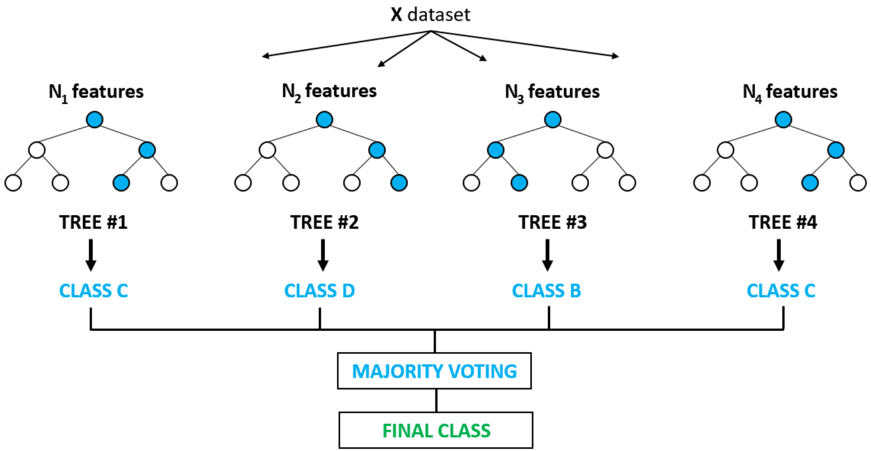


Case Study: Tree based ML algorithms

- Constructing a tree
  - Determining the optimal split
  - Gain measure
- Variable importance
- Pruning a tree

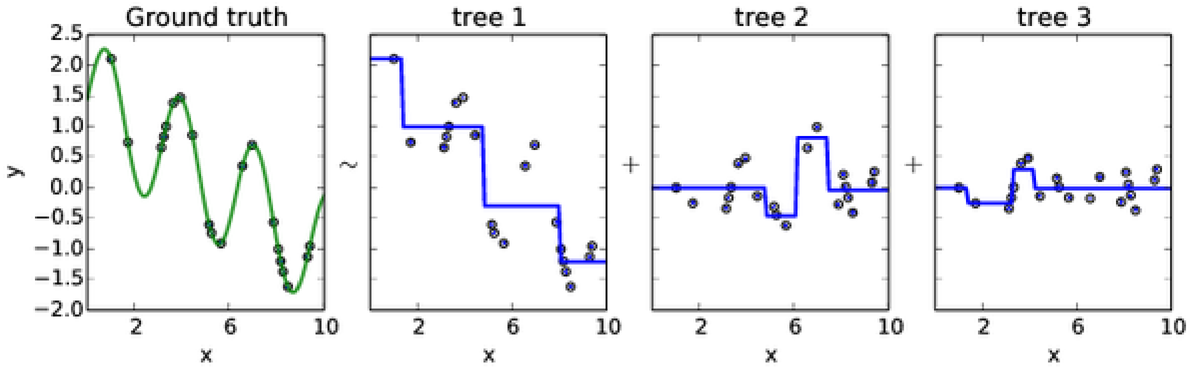
Case Study: Tree based ML algorithms

- Tree in a random forest



Case Study: Tree based ML algorithms

- Tree in a GBM



Appropriateness of Machine Learning algorithms

- Each model has its pros and cons. Hence it is essential to understand:
  - Relationships of variables used in the model.
  - Extent of accuracy, training time, number of parameters, number of features needed in the model.
  - Need for closed form solution.
  - What each method can do well and what it cannot.

Thank you very much for your attention!

Satraajeet Mukherjee

Address: 5 Times Square
New York, USA
Phone: +1 646 239 0961
Mail: Satraajeet@gmail.com
Web: www.ica2018.org

Tvisha Gupta

Address: EY LLP, Brigade Metropolis,
Bangalore, India
Phone: +91 9591295063
Mail: Tvishagupta92@gmail.com
Web: www.ica2018.org