Let's Learn Machine-Learning (LLML)

Yogesh Haribhau Kulkarni

Outline

YHK

About Me

Yogesh Haribhau Kulkarni

Bio:

- 20+ years in CAD/Engineering software development
- Got Bachelors, Masters and Doctoral degrees in Mechanical Engineering (specialization: Geometric Modeling Algorithms).
- Currently doing Coaching in fields such as Data Science, Artificial Intelligence Machine-Deep Learning (ML/DL) and Natural Language Processing (NLP).
- Feel free to follow me at:
  - Github (github.com/yogeshhk)
  - LinkedIn (www.linkedin.com/in/yogeshkulkarni/)
  - Medium (yogeshharibhaukulkarni.medium.com)
  - Send email to yogeshkulkarni at yahoo dot com

Office Hours: Saturdays, 2 to 5pm (IST); Free-Open to all; email for appointment.

Introduction to AI

Data Science, Artificial Intelligence is critical in bringing intelligent automation

What are Data Sciences?

ie What is Artificial Intelligence? Machine Learning? Deep Learning?

Data Science

- Science of Data (obviously)
- Use of Data for Applications
- Some parts of AI uses Data to find patterns and insights which are helpful in multiple applications
- Machine and Deep Learning that part of AI that leverages data.

So, more on AI-ML here . . .

What is the Core Idea?

What's the core idea?'

- behind problem solving?
- behind writing software algorithms?
- solving research problems?

YHK

Desire

- To find a “function”
- To find a relation
- To find a transformation
- To build a model
- From given inputs to desired outputs.

That’s it.

YHK

Functions

- Some functions are straight forward
- *"In summer, ice-cream sale goes up"*
- Cause and effect
- Relation (function, Mathematical model) is found out
- Here, simple rule based programming suffices

Functions

- But some functions are complex
- *"More you put efforts, your business flourishes."*
- Cause and effect again, but the relation is far to complex
- Too many variables
- Here, simple rule based programming not humanly possible.
- Lots of research needed to come up with equations.

Functions

- $E = mc^2$
- What's this? a function?
- Input variable(s)?
- Output variable(s)?
- Parameters?
- How's the relation? linear?

YHK

Functions

- But most real-life functions are not deterministic
- Some are probabilistic, some non-linear.
- *"Detecting if the tumor is benign or malignant"*
- *"At any state in the game of chess, whats the next move?"*

YHK

Chess: next move?

- Needs extreme expertise
- Needs "intelligence"
- How do you get that?
  - Built by lots of training.
  - By studying lots of past games.
- This is how Humans build intelligence

Intelligence

- Can machine (software/program) also do the same?
- Can it play chess?
- Can it build intelligence?
- By looking at past experiences (data),
- Training Data: games played, moves used, etc.

Yes, it can!! Thats Artificial Intelligence.

What is AI?

What is Artificial Intelligence (AI)?

My definition:
"If machines (or computer programs) start doing some/all of these "intelligent" tasks, then that's Artificial Intelligence"

Intelligence: the differentiation

- Ability to think various domains
- Ability produce something new
- Ability to detect the unseen
- Ability to enhance knowledge (rules, patterns)

All these, AI has started doing. The AI era has arrived!!

Everyday usage

Artificial intelligence seems to have become ubiquitous.

- Replying to our emails on Gmail
- Learning how to drive our cars,
- Sorting our holiday photos.
- etc.

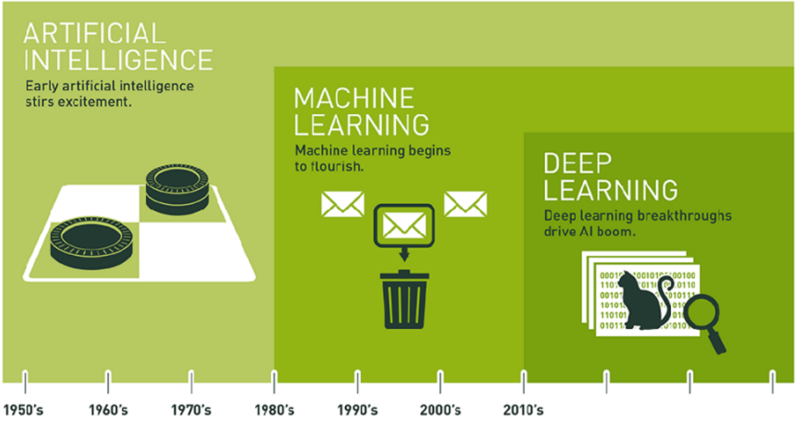
Too good to be true, isn't it, sort of Magical !!

But then ...

- When its too good, you start suspecting
- Is it for real!!
- How can such thing happen?
- How far will it go?

The next thing you know, people are worrying about exactly how and when AI is going to doom humanity.

Relationship between AI, ML, DL

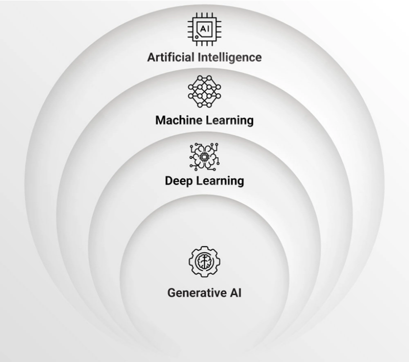


(Ref: https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/)

YHK

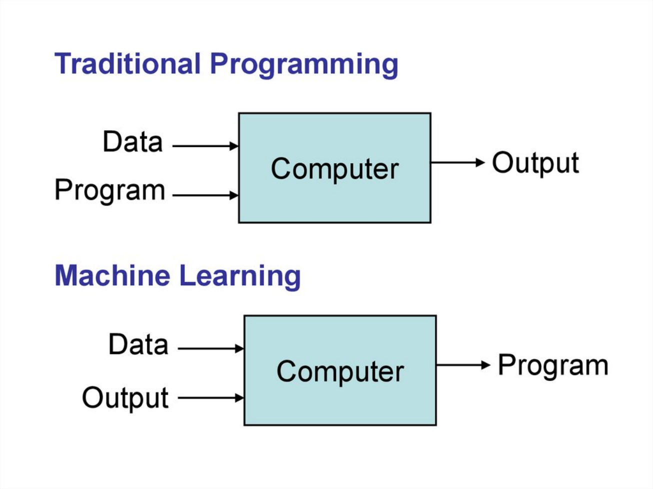
The Modern AI Hierarchy

- **Artificial Intelligence (AI):** Machines mimicking human intelligence.
- **Machine Learning (ML):** Learning from data without explicit programming.
- **Deep Learning (DL):** Neural networks with many layers.
- **Generative AI (GenAI):** DL models that can *create* new content (text, code, images) rather than just analyzing existing data.



(Ref: AI, ML, DL, and Generative AI Face Off: A Comparative Analysis - Synoptek)

Traditional vs. Machine Learning?



Why Machine Learning?

- Problems with High Dimensionality
- Hard/Expensive to program manually
- Techniques to model ‘ANY’ function given ‘ENOUGH’ data.
- Job $$$

YHK

Why now?

- Flood of data (Internet, IoT)
- Increasing computational power
- Easy/free availability of algorithms
- Increasing support from industries

Is AI a threat?

Intro ML

Is AI a threat?

If you believe in what Elon Musk says, then YES.

(Ref: What is Artificial Intelligence — Artificial Intelligence Tutorial For Beginners — Edureka)

Intro ML

Is AI a threat?

If you believe in these movies, then YES.

Well, AI based War robots are not impossible anymore.

(Ref: What is Artificial Intelligence — Artificial Intelligence Tutorial For Beginners — Edureka)

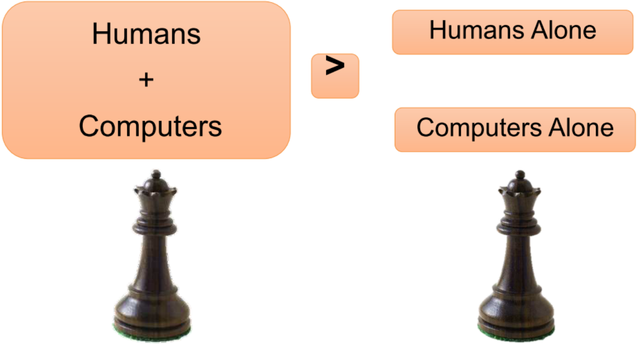
YHK

Fear: Are we being replaced?

- Yes. in tasks that are repetitive
- But not which require complex thinking and creativity

Mostly

Technology Enhancing (Not Replacing) Humans



(Ref: "Artificial Intelligence Overview" - Harry Surden)

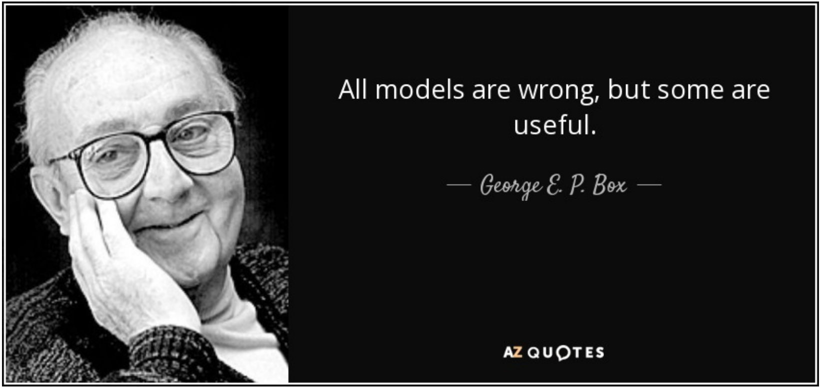
YHK

Limits on Artificial Intelligence

- Many things still beyond the realm of AI
- No thinking computers
- No Abstract Reasoning
- Often AI systems Have Accuracy Limits
- Many things difficult to capture in data
- Sometimes Hard to interpret Systems

After all this - The Truth . . .

Intro ML


All models are wrong, but some are useful.
— George E. P. Box —
AZ QUOTES

Introduction to Machine Learning

How do we learn?

- What do we do when we have to prepare for an examination?
- Study. Learn. Imbibe. Take notes. Practice mock papers.
- Thus, prepare for the unseen test.

What is Learning?

"Learning is any process by which a system improves performance from experience."

\- Herbert Simon, Turing Award 1975, Nobel in Economics 1978.

What is Machine Learning?

Machine learning is a type of artificial intelligence (AI) which:

- Learns function without being explicitly programmed.
- Can grow and change when exposed to new data.

Quick Definition

Machine learning is a field of study that gives computers the ability to learn, without being explicitly programmed.
- Arthur Samuel, 1959

So, rather than coding each step explicitly, you give computers just some examples, and it figures out the steps.

Another Definition of Machine Learning

Machine learning is manifestation of statistical learning: implemented through software.

WellPosed Definition

A computer program is said to learn from experience E with respect to some class of tasks T and some performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.

- T. Mitchell's book Machine Learning (1997)

In the various problem settings T, P, and E can refer to completely different things.

Intuition

- For any process (Machine ie Artificial or Biological), need to define a task, say Spam Detection. Machine can also do it and humans can also detect it.
- Experience means multiple observations, runs, practices. So in our examples, its more Spam/Non-Spam examples.
- Performance defined is How many we got right ie Accuracy.
- Now if Task of Spam detection improves Accuracy if we look at more samples, then this is Machine Learning.
- Its also called Inductive Learning ("Learning by Experience"). Based on evidence (not Facts), so its suggestive.
- Another type is deductive (inferences based on facts/rules. So, more deterministic. Eg I went to Movie today. Today is Saturday. Inference: I went to movie on Saturday.)

YHK

Intuition

- Mitchell's definition, though looks OK for Machine Learning, its not very rigorous.
- Contrary example: Say your task is Bike driving, experiences is more miles you drive, performance is smoothness of drive (vibrations).
- As you put more miles, due to smoothening, performance improves (vibrations go down), is it Learning?
- No!!!
- But for Machine Learning domain, it looks fine.

Tasks

Tasks T in machine learning

- Classification of an instance to one of the categories based on its features;
- Regression – prediction of a numerical target feature based on other features of an instance;
- Clustering – identifying partitions of instances based on the features of these instances so that the members within the groups are more similar to each other than those in the other groups;
- etc

YHK

Experience

- Experience E refers to data (we can't go anywhere without it).
- For example, to predict loan defaults based on the data accumulated about our clients.
- Here, the experience E is the available training data: a set of instances (clients), a collection of features (such as age, salary, type of loan, past loan defaults, etc.) for each,
- And a target variable (whether they defaulted on the loan) is (1 or 0),
- This is a (binary) classification problem.

Performance

- Metric of the algorithm's performance evaluation P
- Such metrics differ for various problems and algorithms
- A simple metric for classification algorithms, the proportion of correct answers – accuracy – on the test set.

So, What is Machine Learning?

- Ability of computers to “learn” from “data”
- Learn: Discover patterns, underlying structure
- Data: Comes from sensors, transactions, etc.

Goal of Statistical learning

Dependent variables need to predicted or estimated in terms of independent variables.

- Data in control: independent variables.
- Data not in control: dependent variables.

Example

Goal: to measure sales based on the advertising budget allocated for TV, Radio, and Print.

- Can control: budgets of TV, Radio, and Print.
- Cannot control: how they will impact the sales.
- Express dependent (sales) as function of independent (advertising budget).
- Want to uncover this hidden relationship.
- Statistical learning reveals hidden data relationships.
- Relationships between dependent and independent data.

YHK

Mathematical Definition of Machine Learning

Machine Learning comes up with a Model given inputs and targets.

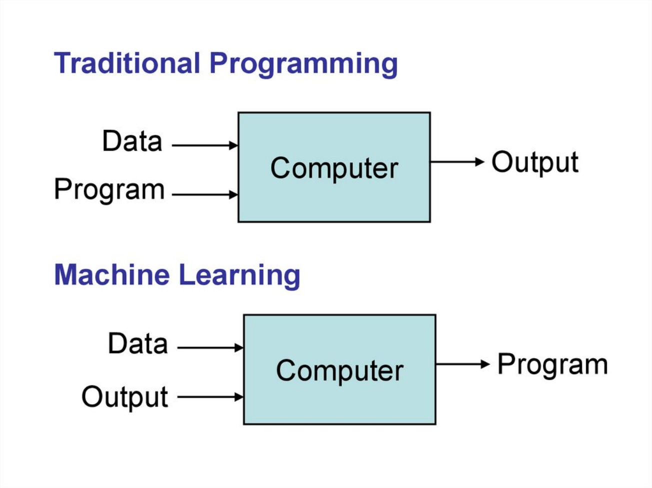
- Input data is available.
- Input data is transformed to get output.
- Output: something that needs to be predicted or estimated.
- Transformation engine is called Model or function.

Model entities

For $income = c + \beta_0 \times education + \beta_1 \times experience$

- Inputs: Education and experience, also called as features or attributes or dimensions or variables.
- Mathematical entities added to input data, are Parameters.. β_0 and β_1 are parameters
- Income is target, also called as outcome or class.

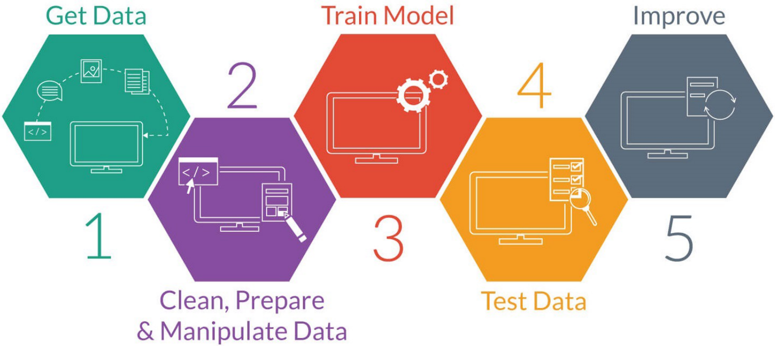
Traditional vs. Machine Learning?



Why Machine Learning?

- Problems with High Dimensionality
- Hard/Expensive to program manually
- Techniques to model 'ANY' function given 'ENOUGH' data.
- Job $$$

Machine Learning Process



(Reference: The Role of Big Data in Strengthening Machine Learning Projects - Techno FAQ)

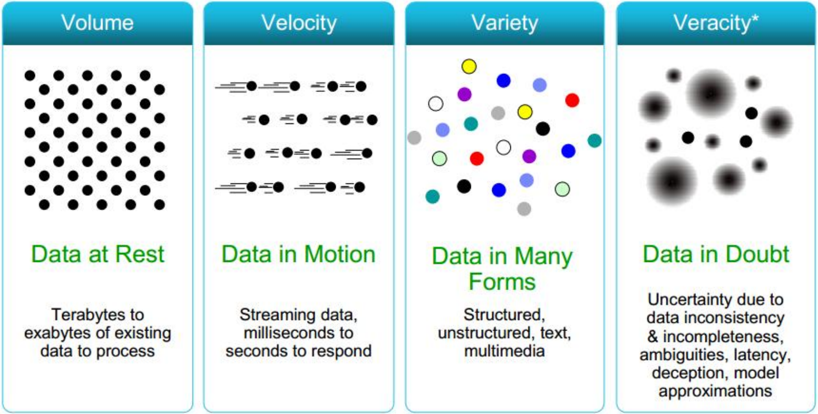
Why now?

- Flood of data (Internet, IoT)
- Increasing computational power
- Easy/free availability of algorithms
- Increasing support from industries

The storm: The Big Data is coming

- In 2012, HBR put Data Scientists on the radar
- "The Sexiest Job of the 21st Century".
- Industry, trying to be data-driven, than manual.

(Big) Data Characteristics



(Image Credit: http://www.rosebt.com/blog/data-veracity)

What's the answer?

AI-ML-DL

- Machines showing intelligence of Humans
- Machine Learning: part of AI
- Logic is not programmed by hand,
- Gets emerged in training with data.

A Puzzle

How different is Machine Learning?

Maths Puzzle

Math Quiz #1 - Teacher's Answer Key

1)	2	4	5	= 3		5)	6	2	2	= 10	
2)	5	2	8	= 2		6)	3	1	1	= 2	
3)	2	2	1	= 3		7)	5	3	4	= 11	
4)	4	2	2	= 6		8)	1	8	1	= 7	

Maths Puzzle

- Letting the computer work out that relationship for you.
- ‘Learn’ to solve such problems,
- ‘Test’ with any other problem of the same type!

Types of Machine Learning

Two kinds of learning

- Supervised
- Unsupervised

Supervised

- Training data with correct answers
- Both used to train the model
- Then apply unseen data on model

Unsupervised

- Training data with no answers
- Extract patterns, groups

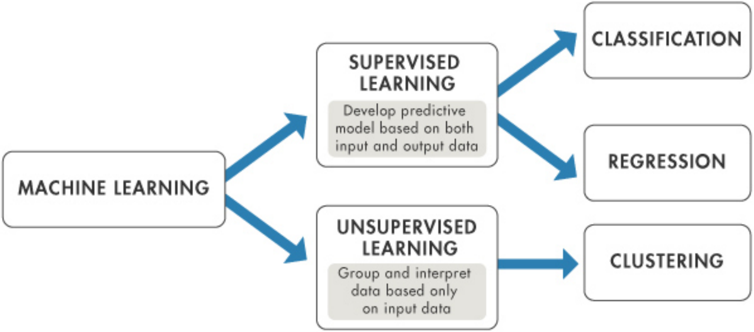
YHK

Some types of algorithms

- Prediction: predicting a continuous variable from data
- Classification: assigning records to predefined groups
- Clustering: splitting records into groups based on similarity
- Association learning: seeing what often appears together

YHK

Machine Learning Learning Algorithms



(Reference: Machine Learning in MATLAB - MATLAB & Simulink - MathWorks)

Machine Learning Learning Algorithms

- Is this A or B? : Classification algorithms
- Is this weird? : Anomaly detection algorithms
- How much—or—How many? : Regression algorithms
- How is this organized? : Clustering algorithms, Dimensionality reduction
- What should I do next? : Reinforcement learning algorithms

(Ref: Brandon Rohrer's breakdown of the "5 questions data science answers")

Classification

- **Description:** Identifying the category an object belongs to.
- **Applications:** Spam detection, Image recognition.
- **Algorithms:** SVM, nearest neighbors, random forest, Logistic Regression

Regression

- **Description:** Predicting a continuous-valued attribute associated with an object.
- **Applications:** Drug response, Stock prices.
- **Algorithms:** Linear Regression

Clustering

- **Description:**Automatic grouping of similar objects into sets.
- **Applications:** Customer segmentation, Grouping experiment outcomes
- **Algorithms:** k-Means

Dimensionality Reduction

- **Description:** Reducing the number of random variables to consider.
- **Applications:** Visualization, Increased efficiency
- **Algorithms:** PCA, Singular Value Decomposition

YHK

Popular Algorithms in Machine Learning

- Linear, Logistic Regression
- Decision Trees
- SVM - Support Vector Machines, Naive Bayes
- K-Means

Popular Algorithms in Machine Learning

Recommendation: Is Gender the decider or Age?

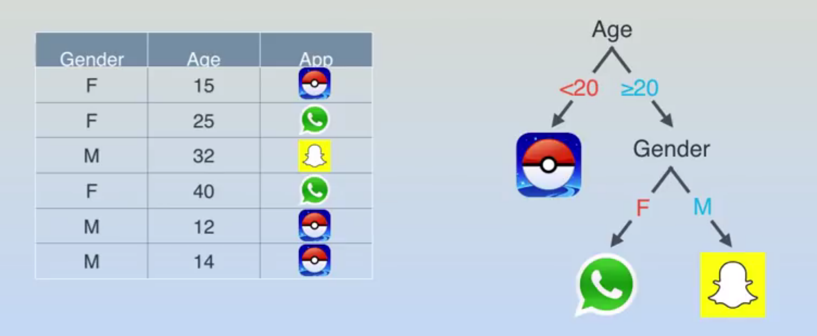
Gender is not much, but all below 20 years downloaded Pokemon Go.Age splits data best.

(Image Credit: A Gentle Introduction To Machine Learning; SciPy 2013 Presentation - Kastner, Kyle)

YHK

Popular Algorithms in Machine Learning

Decision Tree.



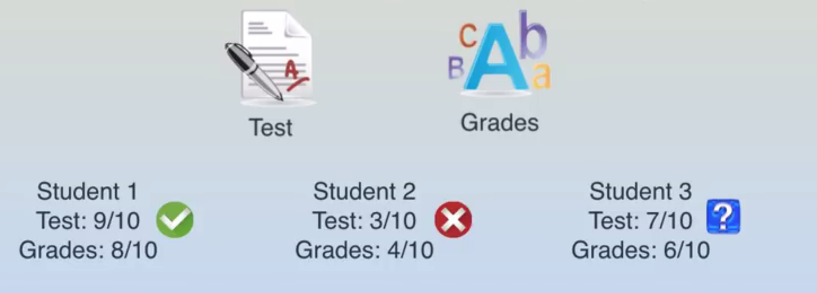
Any new person can walk through the Tree and predict.

(Image Credit: A Gentle Introduction To Machine Learning; SciPy 2013 Presentation - Kastner, Kyle)

YHK

Popular Algorithms in Machine Learning

Deciding acceptance to Univ based on Test scores and grade.

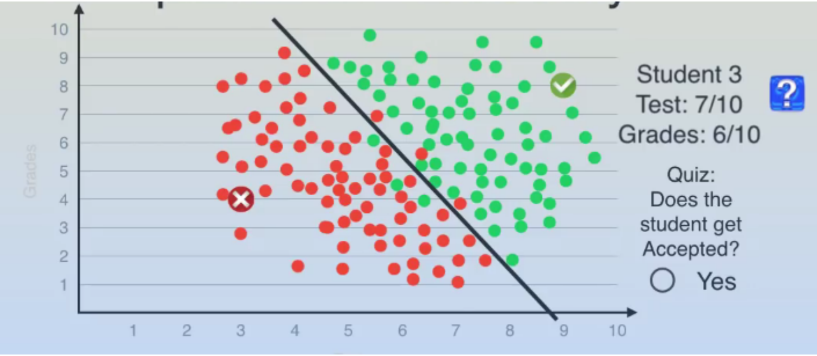


Will Student 3 get accepted?

(Image Credit: A Gentle Introduction To Machine Learning; SciPy 2013 Presentation - Kastner, Kyle)

Popular Algorithms in Machine Learning

Putting it in a grid. Test score as X and Grades as Y. Prev data shown with acceptance colored.



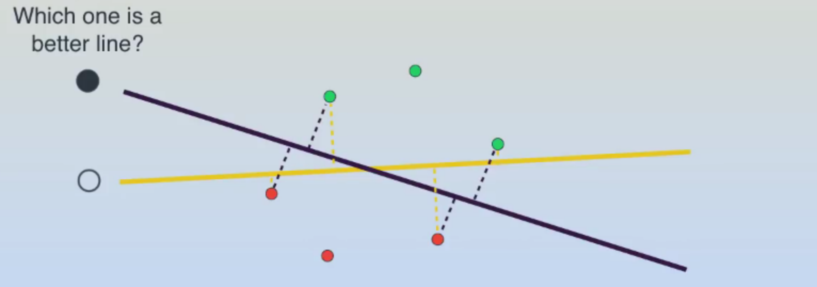
After separating line, Student 3's fate can be predicted. That's Logistic Regression.

(Image Credit: A Gentle Introduction To Machine Learning; SciPy 2013 Presentation - Kastner, Kyle)

YHK

Popular Algorithms in Machine Learning

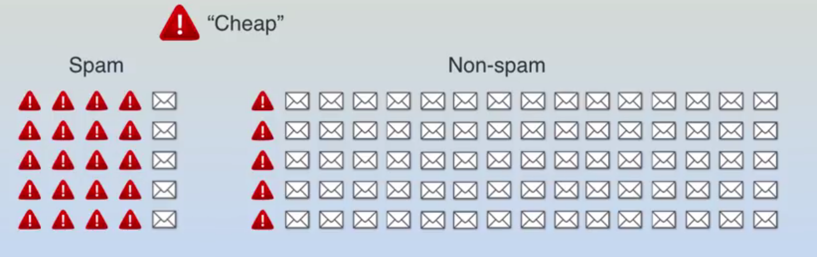
Which line separates best?



The one with max separation in the middle. That's Support Vector Machine.
(Image Credit: A Gentle Introduction To Machine Learning; SciPy 2013 Presentation - Kastner, Kyle)

Popular Algorithms in Machine Learning

Classification: Detecting if mail is spam or not.

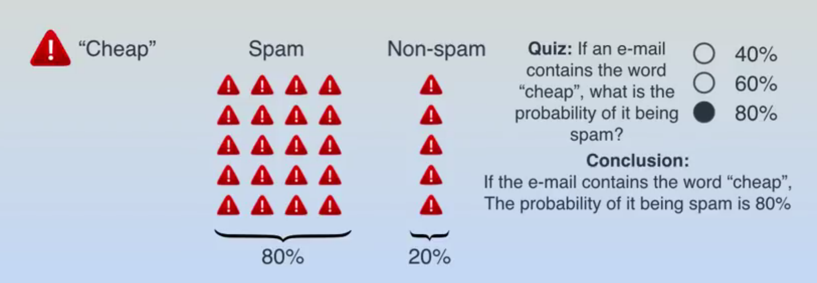


Features: mail containing "cheap".
Most past spam mails contain it.

(Image Credit: A Gentle Introduction To Machine Learning; SciPy 2013 Presentation - Kastner, Kyle)

Popular Algorithms in Machine Learning

Easy to compute probability of being a Spam, if it contains "cheap".

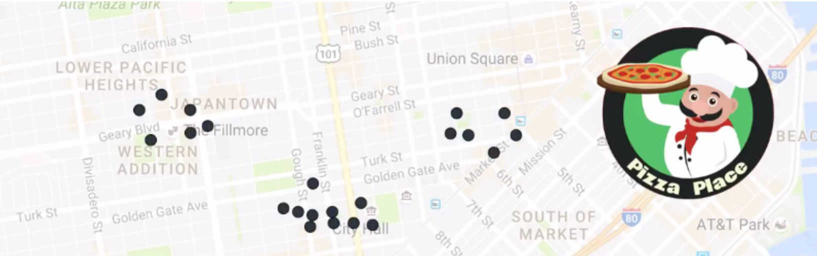


That's Naive Bayes.

(Image Credit: A Gentle Introduction To Machine Learning; SciPy 2013 Presentation - Kastner, Kyle)

Popular Algorithms in Machine Learning

Wish to put 3 pizza shops in a city. Customers are plotted. What are the best locations?

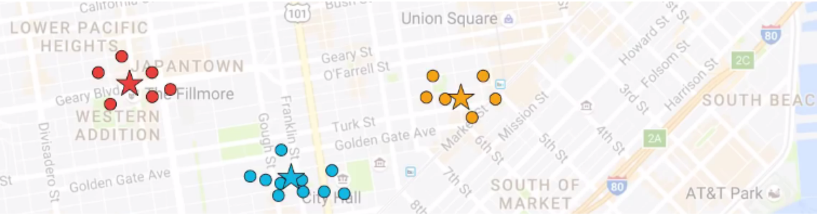


Start with random locations and set ownership. Update locations. Repeat.

YHK

Popular Algorithms in Machine Learning

Locations settle at the centroids of the clusters.



That's K-means.

(Image Credit: A Gentle Introduction To Machine Learning; SciPy 2013 Presentation - Kastner, Kyle)

Example: Predicting House Price

(Ref: Machine Learning is Fun! - Adam Geitgey)

Numerical way

Bedrooms	Sq. feet	Neighborhood	Sale price
3	2000	Normaltown	$250,000
2	800	Hipsterton	$300,000
2	850	Normaltown	$150,000
1	550	Normaltown	$78,000
4	2000	Skid Row	$150,000

Bedrooms	Sq. feet	Neighborhood	Sale price
3	2000	Hipsterton	???

Machine Learning Type

- This is supervised learning.
- Knew how much each house sold for,
- So, knew the answer to the problem
- Need work backwards to figure out the logic.

YHK

Traditional way

BUT, how to decide which numbers to put? PRAY!!!

```
def estimate_house_sales_price(num_of_bedrooms, sqft, neighborhood):
  price = 0
  price_per_sqft = 200

  if neighborhood == "hipsterton":
    price_per_sqft = 400
  elif neighborhood == "skid row":
    price_per_sqft = 100

  price = price_per_sqft * sqft
  if num_of_bedrooms == 0:
    price = price - 20000
  else:
    price = price + (num_of_bedrooms * 1000)

  return price
```

Intro ML

Prayer

- Wouldn't it be better if computer figures out?
- Treat it as black box
- Feed Inputs and outputs
- That's it!!

```python
def estimate_house_sales_price(num_of_bedrooms, sqft, neighborhood):
    price = <computer, plz do some math for me>
    return price
```

Prayer granted!!

- Notice the magic numbers
- .841, 1231.123, 2.324, and 201.234.
- These are weights.
- Better the weights - better the prediction!
- Done!!

```python
def estimate_house_sales_price(num_of_bedrooms, sqft, neighborhood):
  price = 0
  price += num_of_bedrooms * .841
  price += sqft * 1231.123
  price += neighborhood * 2.324
  price += 201.234
  return price
```

How to figure out? A dumb way

Step 1: Start with each weight set to 1.0:

```python
def estimate_house_sales_price(num_of_bedrooms, sqft, neighborhood):
  price = 0
  price += num_of_bedrooms * 1.0
  price += sqft * 1.0
  price += neighborhood * 1.0
  price += 1.0
  return price
```

How to figure out? A dumb way

Step 2: Guess/predict for all houses

Bedrooms	Sq. feet	Neighborhood	Sale price	My Guess
3	2000	Normaltown	$250,000	$178,000
2	800	Hipsterton	$300,000	$371,000
2	850	Normaltown	$150,000	$148,000
1	550	Normaltown	$78,000	$101,000
4	2000	Skid Row	$150,000	$121,000

Predictions NOT good, right?

What to do?

- Actual \$250,000, but guessed \$178,000
- Off by \$72,000 for that single house.
- Diff can be positive or negative, so square it
- Add squared diffs of all houses.
- Total: \$86,123,373.
- Thats whole error in the system.
- That's how "wrong" your function currently is.

What Next?

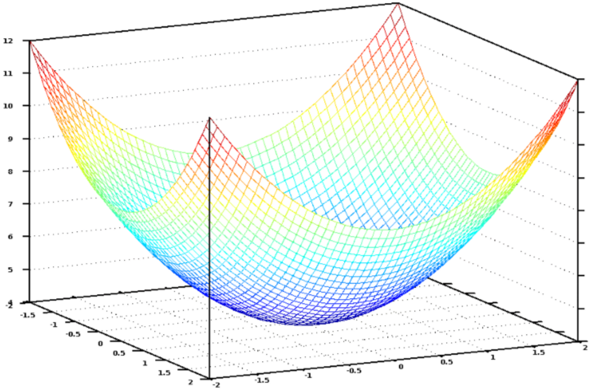
- Average per house error is “cost”.
- Get cost to be zero by playing with the weights.
- Thats the Goal!!!

Mathematically

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

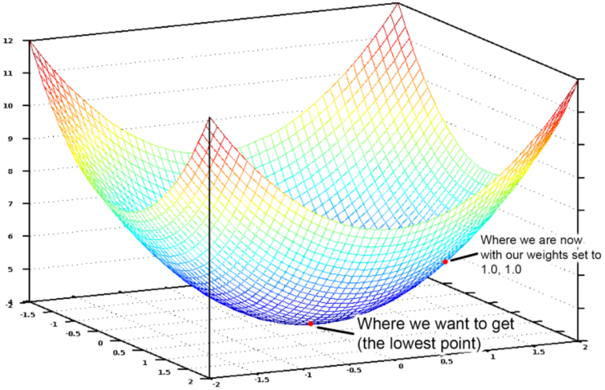
Graphically

Plotting cost values, for all possible ranges of weights for number_of_bedrooms and sqft:



Graphically

Cost is lowest at lowest point of the surface. Ideal.
Weights at that point are the answers!



How to find the lowest cost point?

- Start somewhere.
- Find direction (slope? Derivative? Partial?)
- Derivative: tells us which way is downhill for any given point on our graph.
- Move in slope direction.
- Adjust our weights to get to next point
- "walking down hill" towards the lowest point.
- That's gradient descent.
- Scikit Learn does this for you, hushsh!!

Calculus, anybody?

Repeat until convergence {

$$\theta_j \leftarrow \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta)$$

}

Applications of Machine Learning

Everyday Applications of Machine Learning

- Face Recognition (Facebook)
- Spam recognition in Emails
- Recommender Systems
- Feelings Analysis, Sentiments
- Natural language: Translate a sentence from Hindi to English, question answering, etc.
- Speech: Recognise spoken words, speaking sentences naturally
- Game playing: Play games like chess
- Robotics: Walking, jumping, displaying emotions, etc.
- Driving a car, flying a plane, navigating a maze, etc.

Cool-down: Summary

SO . . .

- What is Machine learning, after-all?
- Its usage in your domain?

YHK

References

Many publicly available resources have been refereed for making this presentation. Some of the notable ones are:

- Introduction to Machine Learning with scikit.learn - West of Ireland Data Science
- STAT 365/665: Data Mining and Machine Learning - Taylor Arnold
- CSC 600: Data Mining - Richard Burns
- Data Science Simplified - Pradeep Menon
- Learn Data Science - Nitin Borwankar
- IAML: Decision Trees - Victor Lavrenko and Charles Sutton
- Data Science Notebooks
- Analytics Vidhya Blogs
- Machine Learning - Brett Wujek , SAS Institute
- Introduction to Entropy for Data Science - Mike Schulte

YHK

 Intro ML

Thanks ...

- Office Hours: Saturdays, 3 to 5 pm (IST); Free-Open to all; email for appointment to yogeshkulkarni at yahoo dot com
- Call + 9 1 9 8 9 0 2 5 1 4 0 6

(https://www.linkedin.com/in/yogeshkulkarni/)

(https://medium.com/@yogeshharibhaukulkarni)

(https://www.github.com/yogeshhk/)

YHK