# MACHINE LEARNING

# The A Team



## Presented By:

Suvojit Manna
Pronab Mukherjee
Barun Gupta
Somnath Rakshit

# Contents in Brief

# Let’s Get Started

# Machine Learning — What ?



Field of study that gives computers the ability to learn without being explicitly programmed.

Instead of writing code, you feed data to the generic algorithm and it builds its own logic based on the data.

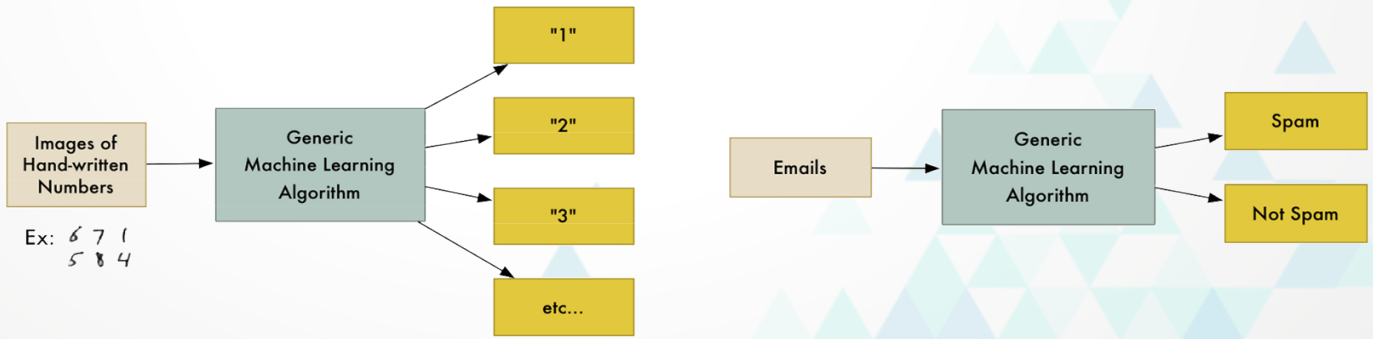



Figure: Classification Algorithms

# Case Studies — Supervised Learning



| Bedroom | Sq.Ft | Neighbourhood | Price |
|---|---|---|---|
| 3 | 2000 | Uptown | $350,000 |
| 2 | 800 | Downtown | $200,000 |
| 2 | 850 | City Centre | $150,000 |
| 1 | 550 | Suburbs | $75,000 |
| 4 | 2000 | Suburbs | $200,000 |

| Bedroom | Sq.Ft | Neighbourhood | Price |
|---|---|---|---|
| 3 | 2000 | Uptown | ??? |

## Definiton

Supervised learning is the machine learning task of inferring a function from labeled training data.

# Case Studies — Supervised Learning



| **Math's Exam - Answer Keys** | |
|---|---|
| 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) 2 2 4 = 6 | 8) 1 8 1 = 7 |

- The training data consist of a set of training examples.
- Training Data :
  - Input Object : Set of Features
  - Desired Output : Supervisory Signal
- A supervised learning algorithm produces an inferred function.
- An analogus task in human and animal phsycology : Concept Learning.

# Case Studies — Unsupervised Learning



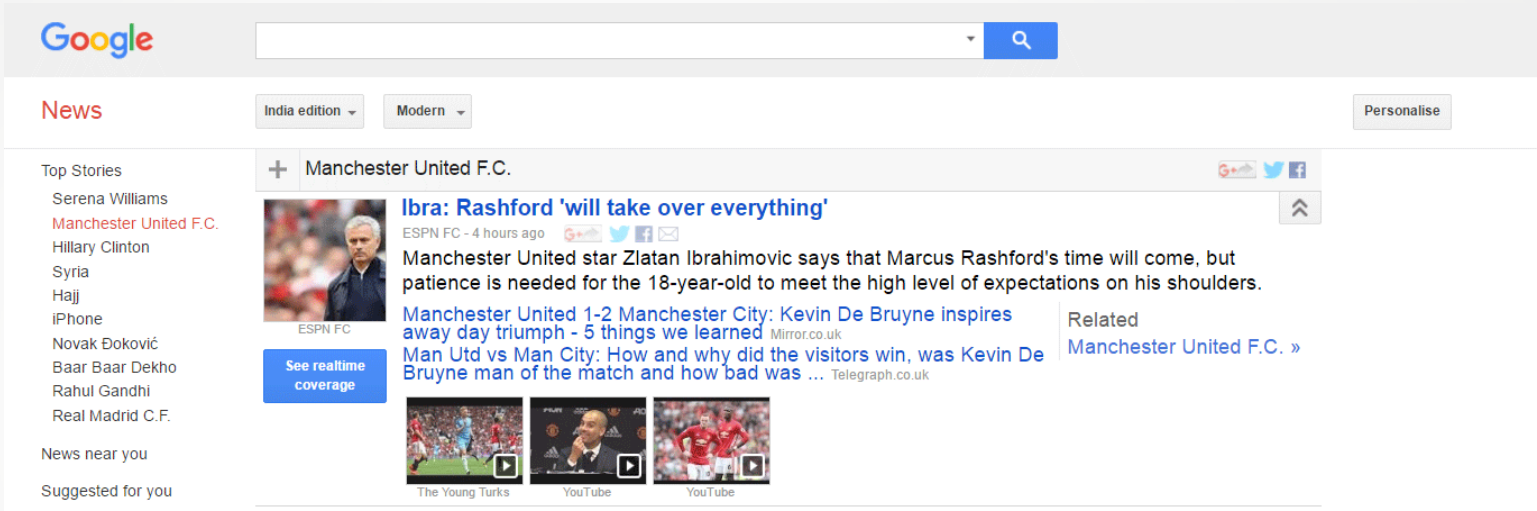



Figure: Google News grouping similar stories together.

## Definiton

Unsupervised learning is the machine learning task of inferring a function to describe hidden structure from unlabeled data.

# Cocktail Party Problem — Unsupervised Learning



Sound from :

- *Microphone 1*
- *Microphone 2*

Output from Learning Algorithm :

- *Output 1*
- *Output 2*

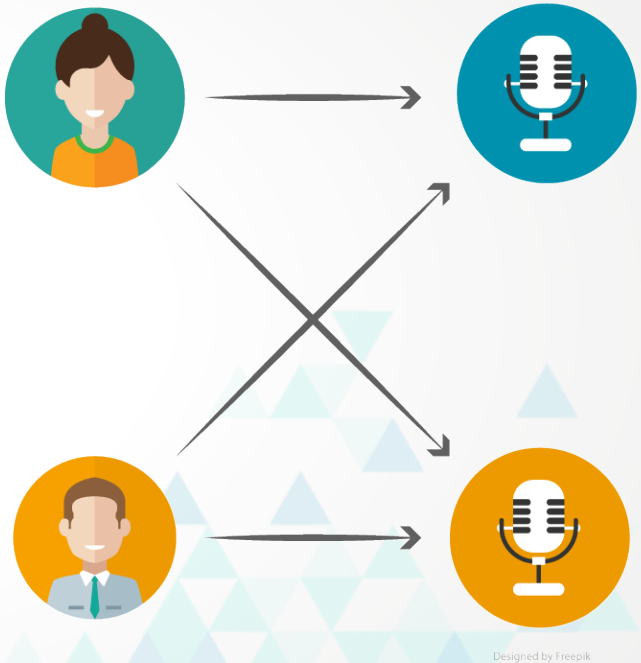



Figure: Overlapped Recordings

# Case Studies — Unsupervised Learning



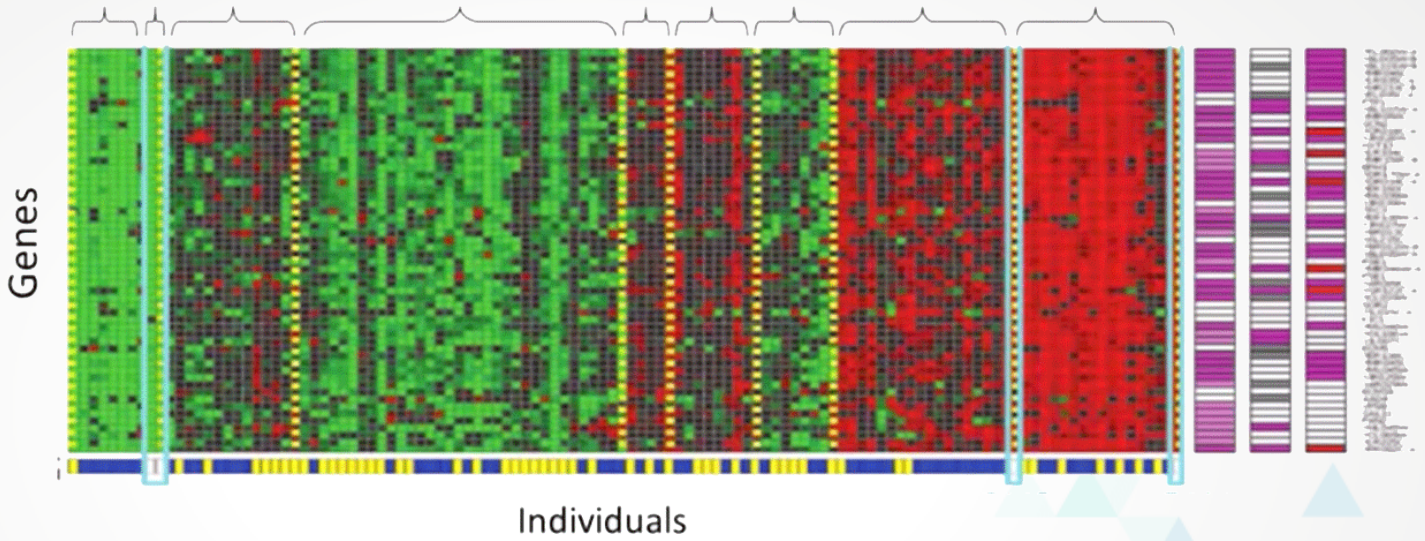



Figure: Gene Clustering

- Training Data given to the learner is unlabeled.
- No error or reward signal to evaluate a potential solution.
- Closely related to density estimation in statistics.

# Machine Learning — Formal Definiton



The field of machine learning is concerned with the question of how to construct computer programs that automatically improve with experience.

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

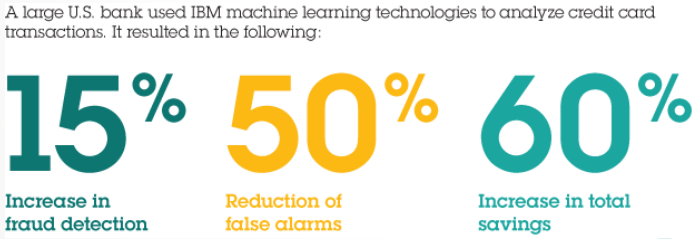
Evolved from :

- Pattern Recognition
- Computational Learning Theory
- Artificial Intelligence

# Industry Trends



Google Chauffer : Self Driving Car by Google




IBM Research : Credit Card Fraud Detection


Mail Services : Spam Filtering

# Industry Trends



Completed • $1,000

The Online Privacy Foundation

**Psychopathy Prediction Based on Twitter Usage**

Mon 14 May 2012 – Fri 29 Jun 2012 (4 years ago)

Dashboard

Private Leaderboard - Psychopathy Prediction Based on Twitter Usage

This competition has completed. This leaderboard reflects the final standings.

See someone using multiple accounts? Let us know.

| # | Δrank | Team Name * in the money | Score | Entries | Last Submission UTC (Best – Last Submission) |
|---|---|---|---|---|---|
| 1 | ↑5 | y_tag * | 0.86997 | 12 | Tue, 26 Jun 2012 12:46:19 |
| 2 | ↑9 | Bruce Cragin | 0.86745 | 10 | Fri, 29 Jun 2012 22:28:17 (-47.6h) |
| 3 | ↑21 | Indy Actuaries | 0.86700 | 6 | Fri, 29 Jun 2012 03:40:38 (-3.4d) |

Figure: Kaggle Challenge : Psychopathy Prediction

The aim of the competition is to determine to what degree it's possible to predict people with a sufficiently high degree of Psychopathy based on Twitter usage and Linguistic Inquiry.

# Entertainment — Machine Learning

# Applications — Machine Learning

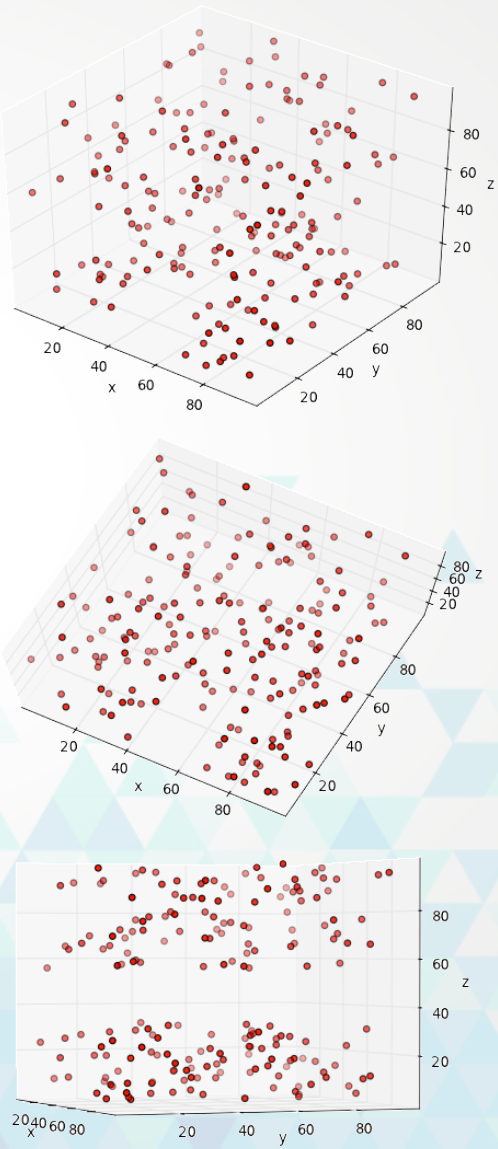


- Adaptive websites
- Classifying DNA sequences
- Computer vision
- Internet fraud detection
- Natural language processing
- Online advertising
- Recommender systems
- Search engines
- Sentiment analysis
- Speech and handwriting recognition

# Machine Learning — Why ?



- Can work with huge amount of data.
- Can make intelligent decisions by taking into account multiple features.
- Can find patterns in large amount of data which is almost impossible for human beings.
- These algorithms are self-modifying in nature, they get better over time as the usage increases.



# Introduction — Regression

# Usages — Regression

# Benefits — Regression

# Example Cases — Regression

# Introduction — Classifications

# Usages — Classifications

# Example Cases — Classifications

# Introduction — Deep Learning

# Neural Networks — Deep Learning

# Meaning — Deep Learning

# Usages — Deep Learning

# Advantages — Deep Learning

# The pain is almost over

# Bibliography



- Phil Simon (March 18, 2013). Too Big to Ignore: The Business Case for Big Data. Wiley. p. 89. ISBN 978-1-118 -63817-0.
- Mitchell, T. (1997). Machine Learning, McGraw Hill. ISBN 0-07-042807-7
- Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar (2012) Foundations of Machine Learning, The MIT Press ISBN 9780262018258.
- Jordan, Michael I.; Bishop, Christopher M. (2004). "Neural Networks". In Allen B. Tucker. Computer Science Handbook, Second Edition (Section VII: Intelligent Systems). Boca Raton, FL: Chapman & Hall/CRC Press LLC. ISBN 1-58488-360-X.
- https: //www.kaggle.com/c/twitter-psychopathy-prediction
- Mastering the game of Go with deep neural networks and tree search (2016), D. Silver et al.

Now that was very interesting!

The End