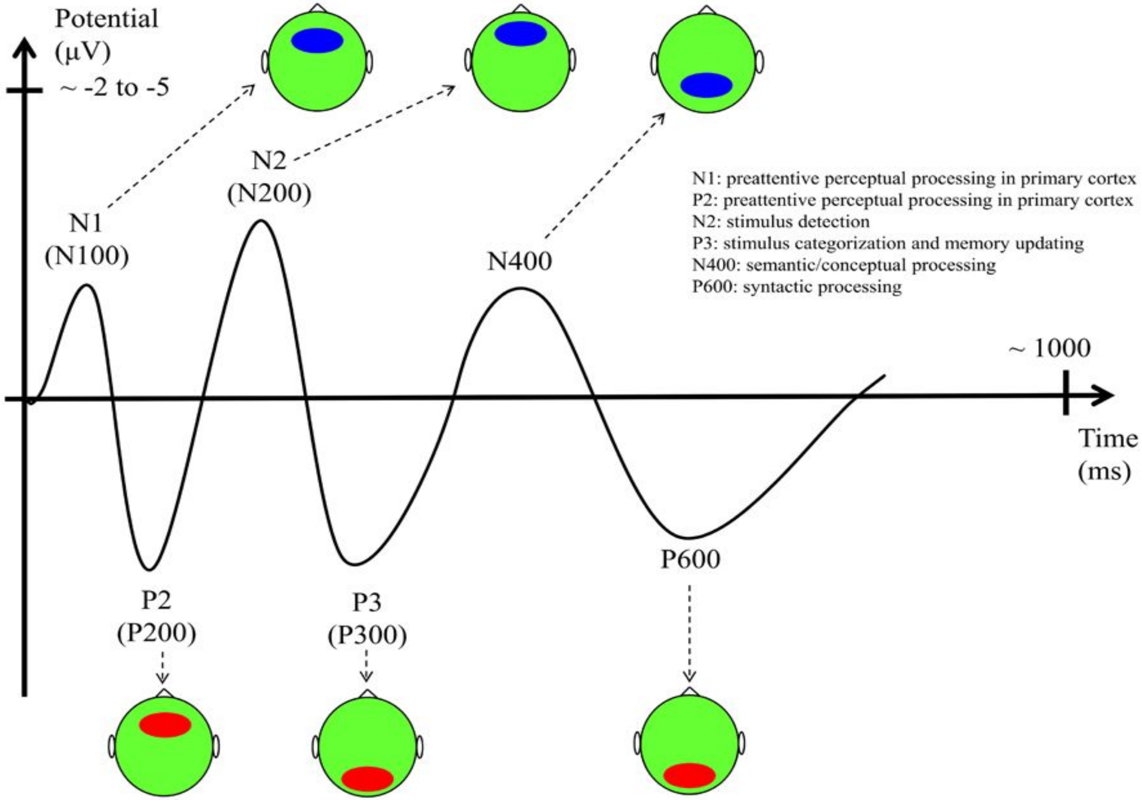
NEUROTECHUOFT: INTERMEDIATE WORKSHOPS
MACHINE LEARNING IN NEUROTECHNOLOGY

Agenda

- Event-related potentials
- Why Machine Learning?
- Statistics: Review
- Intro to ML

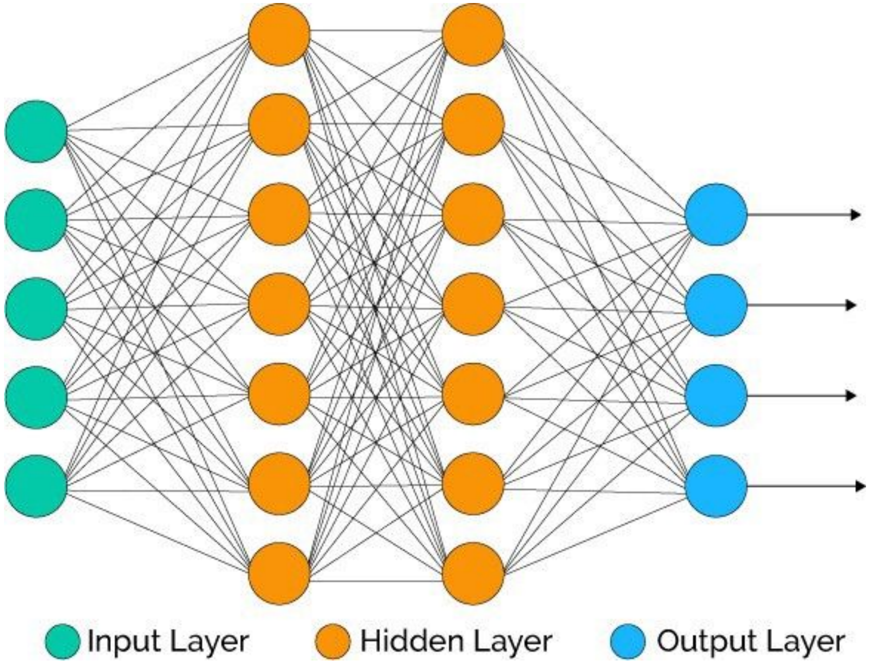
Event related potentials (Motivation)

- Different
  pattern of electrical activity corresponding to the "event"/process
- How can we classify these events automatically?



Why Machine Learning?

- Adaptability
- Learn non-trivial patterns in data
- Observe patterns in real time
- Make sense of a lot of data
- Fast & scalable



Quick Review: Statistics

- **Population**: a collection of objects of interest
- **Sample:** a subset of a population
- **Parameter:** a value that describes a feature of the population

What is an estimator?

- **Estimator:** functions of observations that estimate some parameter

Unbiasedness

An estimator $\hat{\theta}$ of an unknown parameter θ is **unbiased** if and only if

$$E(\hat{\theta}) = \theta$$

Consistency

An estimator $\hat{\theta}$ of an unknown parameter θ is **consistent** if

$$\lim_{n \to \infty} P(|\hat{\theta}_n - \theta| \geq \epsilon) = 0$$

Pictures from STA248 slides

Quick Review: Partial Derivatives

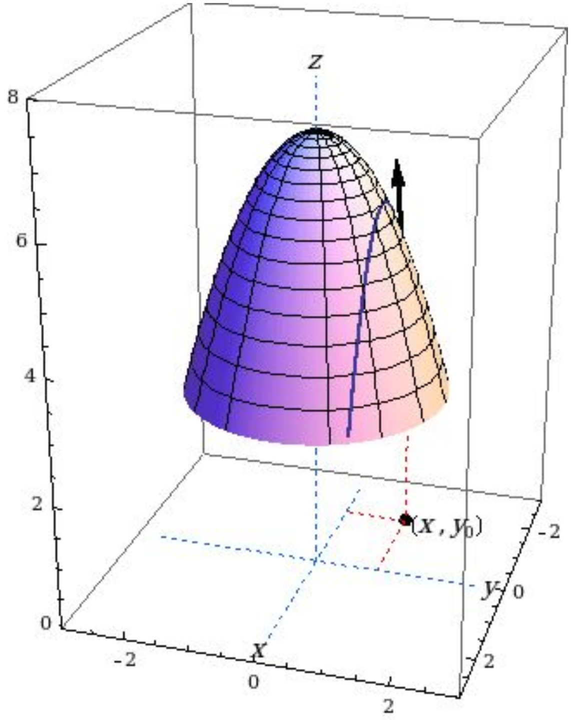
- Regular derivative: **if we change x by a bit**, how does **y change**?

$$y = f(x)$$

- Partial derivative: **if we change ONLY x by a bit**, how does **y change**?

$$y = f(v, w, x)$$

Partial Derivatives in GIF form

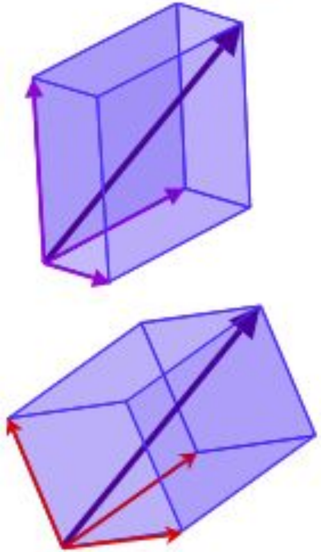


Maximum Likelihood Estimation

- $L(\theta) = f(x_1 \dots x_N|\theta) \approx \prod_{i=1}^{N} f(x_i|\theta)$
- Take logarithm and derivative and set to zero and you’re blessed!

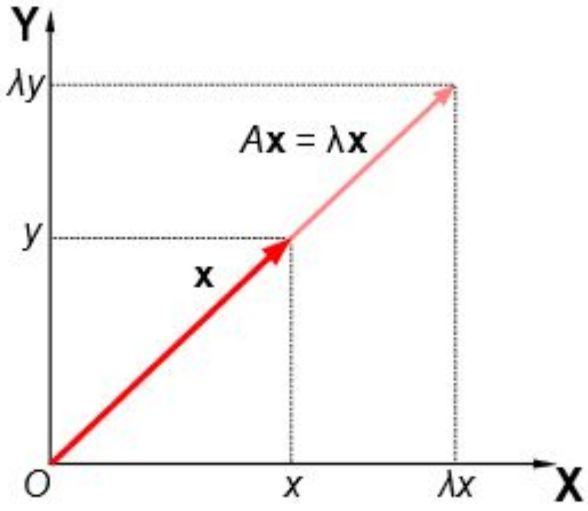
Linear Algebra Review

- Basis and components



From Wikipedia

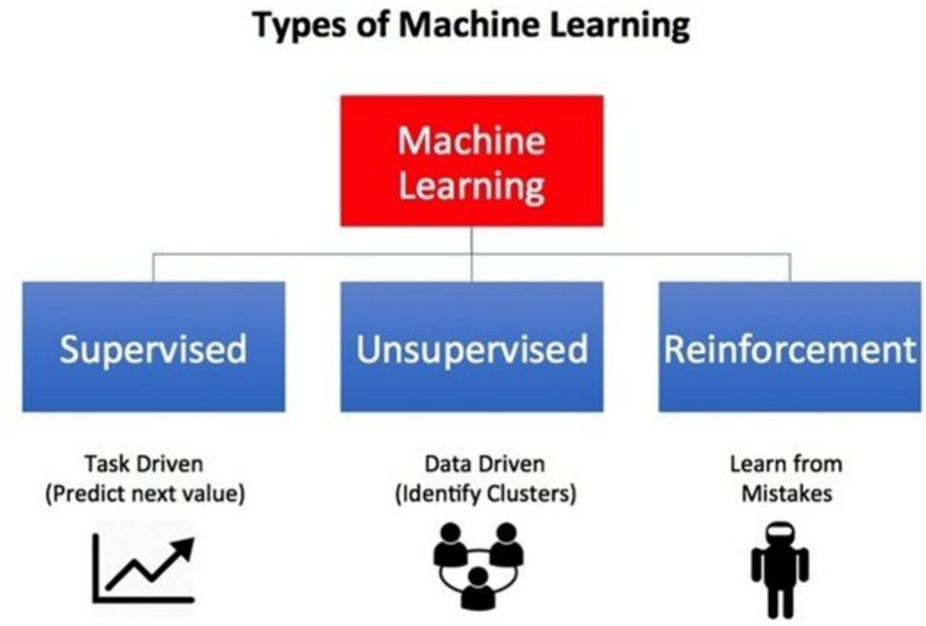
- Eigenvectors and eigenvalues



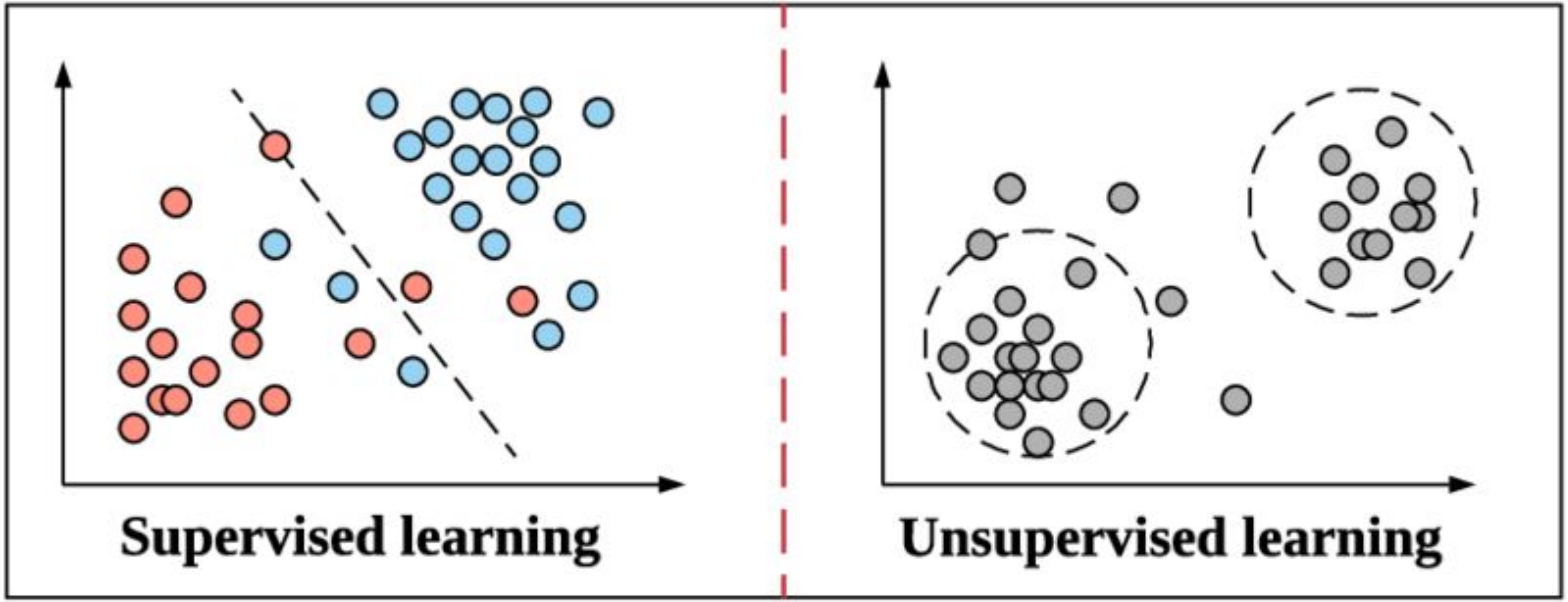
From Wikipedia

Machine Learning comes in different forms...

- Supervised learning
- Unsupervised learning
- Reinforcement learning
- Semi-supervised learning

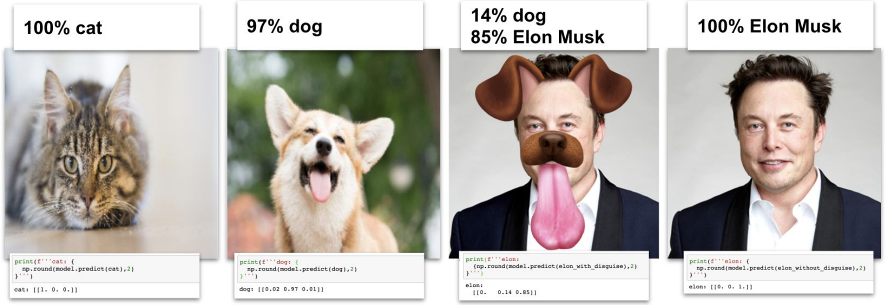


Supervised vs. Unsupervised



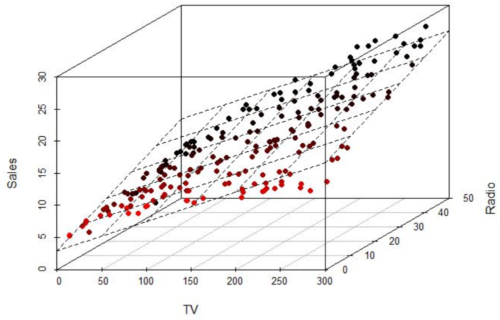
Supervised Learning

- When is it used?
  - When you are dealing with **labelled data** (for classification)
  - Making predictions on **unseen** data -- can be classification or regression
- Example: Say I have a dataset of images, can be either cat or dog
- **Train** a supervised learning model to **learn** to distinguish the two by feeding it **labelled** images of both categories
- How to **test** the model?
  - Give the model a cat or dog image that it **has not seen before**



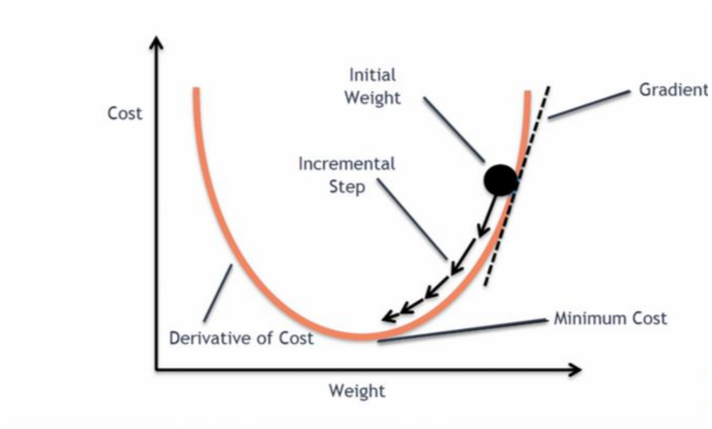
Regression

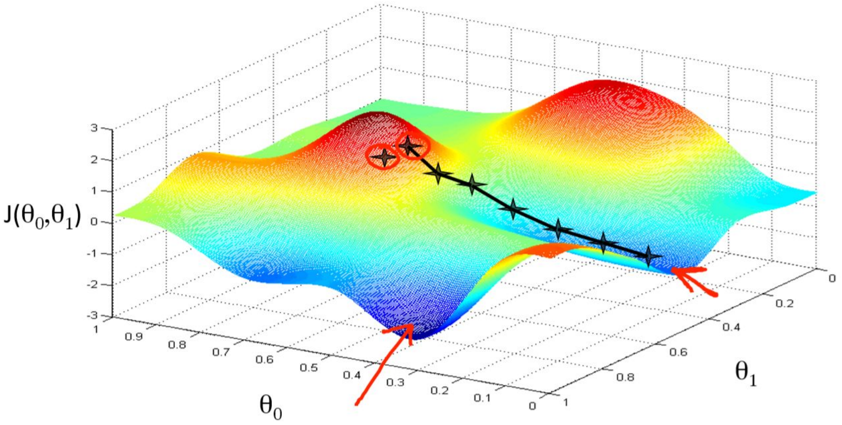
- Simple linear regression: $y = w*X + B$
  - W and B are unknown
  - What is the goal?
- Optimal w?
- Direct solution (we are deriving this)



Gradient Descent

- Iterative Algorithm
- Minimizing a cost function
- Learning rate

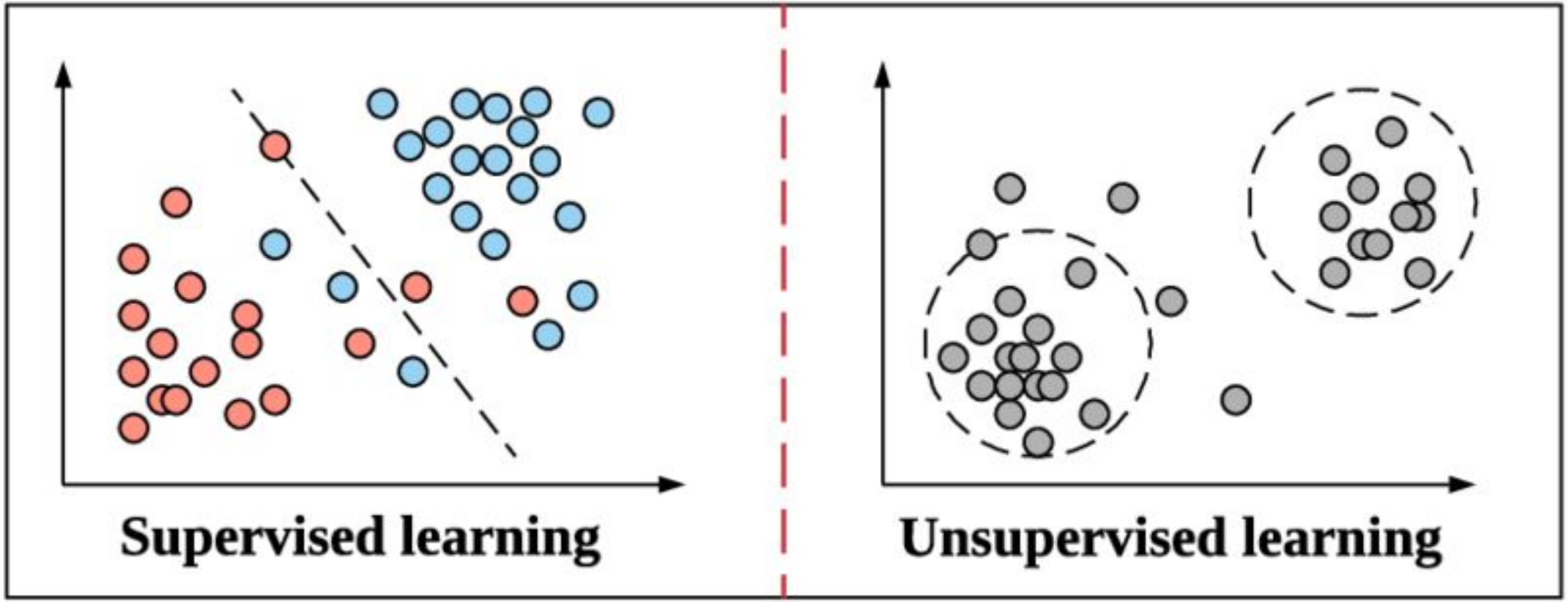




Unsupervised Learning

- When is it used?
  - When you are dealing with **unlabelled data** that contains some underlying structure
  - Making predictions on **unseen** data -- can be classification or regression
- Example: sample of patients with a particular disease
  - Identify subtypes of the disease?
- Train an unsupervised learning model to learn underlying structure of data
- How to **test** the model? (Can be hard)
  - Test on a dataset that contains labels
  - Experts manually label data

Supervised vs. Unsupervised



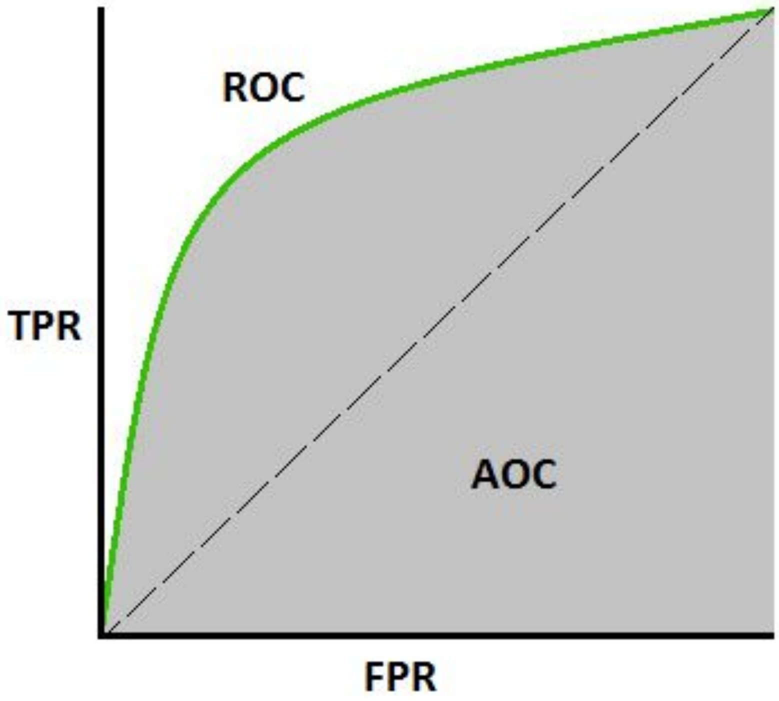
Performance Metrics

AUC - Area Under Curve

ROC - Receiver Operator Characteristic

TPR - True positive rate

FPR - False positive rate



Getting your hands dirty with ML: Libraries/Frameworks

- Data processing/manipulation: Numpy and Pandas
- General ML toolbox: Scikit Learn
- Neural Networks: TensorFlow, PyTorch, Keras
- Data visualization: Matplotlib, Seaborn
- **Note:** These are specific to Python (we will be using Python mainly)

Now you’re ready to go through the ML Colab Notebook in breakout groups :)

EXTRA

Regression Estimator with MLE: Solution