# Make a Prediction

The main focus of machine learning is making decisions or predictions based on data

# Problem of Induction

Why do we think that previously seen data will help us predict the future?

Stated in another way, what will enable us to predict a solution based on previous data ?

In some cases if we can assume that the data observes some properties such that if all the data belongs to the same distribution & we take random data , then we can make a successful prediction (IID Independent and Identical Distribution )

# Human role in ML

## Define the problem

A human still has to frame the problem in ML terms as defined further below .

### Supervised Learning Problem

Dn (Training data) ∈ {(x (1) , y (1) ), . . . ,(x (n) , y (n) )}

where x (i) represents an object to be classified

x (i)∈ R D  ( a d-dimensional vector of real and/or discrete values)

#### Classification

y (i) is an element of a *discrete set* of values, for e.g y (i) ∈ {0,1} (Binary classification)

#### Regression

y(i) ∈ R k

## **Get the data**

Acquire and organize data .Assumptions about the data (IID , Markov, Adversarial ( think chess ) )

## Solutions Space

Here we define our Hypothesis Class which we feel should give us the solution .

### Without Models

This can be one that produces no model (like nearest neighbour)

### With Models

We can create a model (linear & logistic regression).

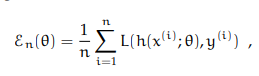
The model will be some hypothesis or prediction rule **y = h(x; θ)** for some functional form h

The idea is that **θ is a vector of one or more parameter values that will be determined** by fitting the model to the training data and then be held fixed.

Given a new **x(n+1)**, we would then make the prediction **h(x(n+1); θ)**

The fitting process is often articulated as an **optimization problem** :

So, we would seek θ that minimizes



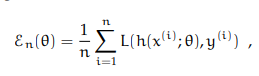
where the loss function L(g, a) measures how bad it would be to make a guess of g when the actual value is a.

## Generate Predictions

Apply the algorithm to the data , run the data through the model /algo and see the result.

## Evaluate the solution

Evaluation Criteria ( 0 - 1 loss , linear loss , squared loss, Asymmetric loss)



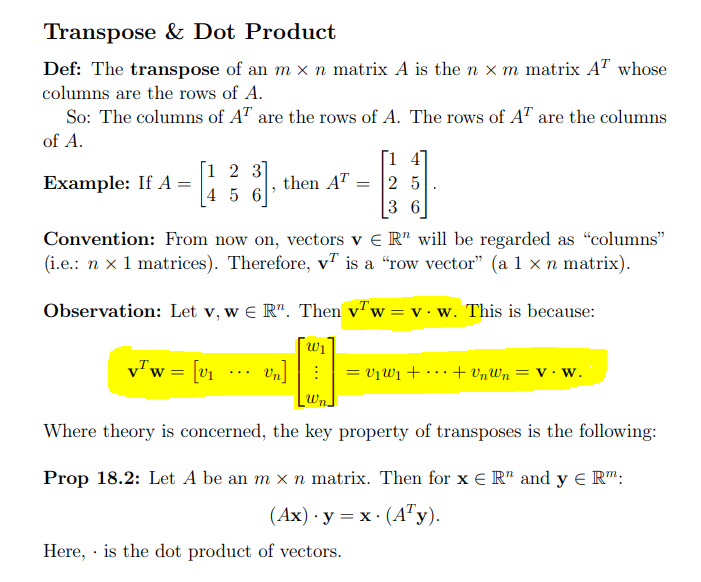
## Learning Algorithm - Optimizing the Error

Once we have a solution & a measure to gauge the performance , then we can devise a strategy to minimize the Error . Sometimes we can use software that was designed, generically, to perform optimization. In many other cases, we use algorithms that are specialized for machine-learning problems, or for particular hypotheses

# References

## Relation between Dot and Vector Product

<http://math.stanford.edu/~jmadnick/R3.pdf>



## Distance of point from plane

