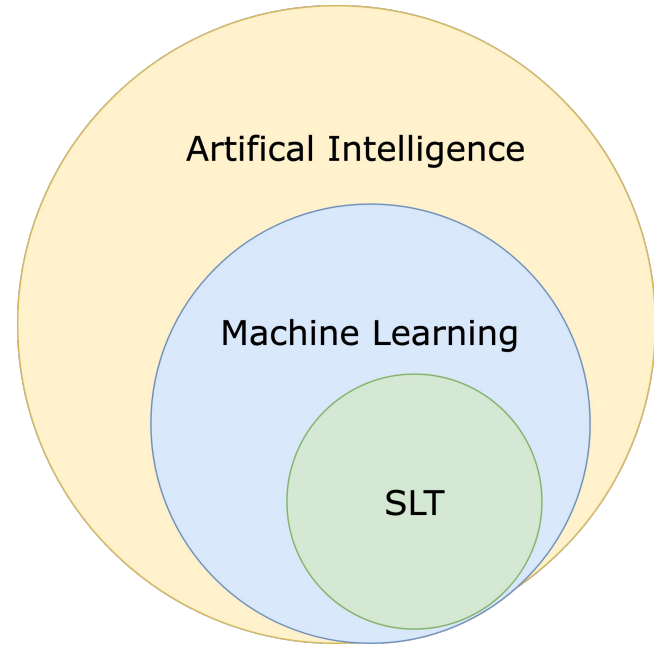# Artificial Intelligence: Statistical Learning Theory

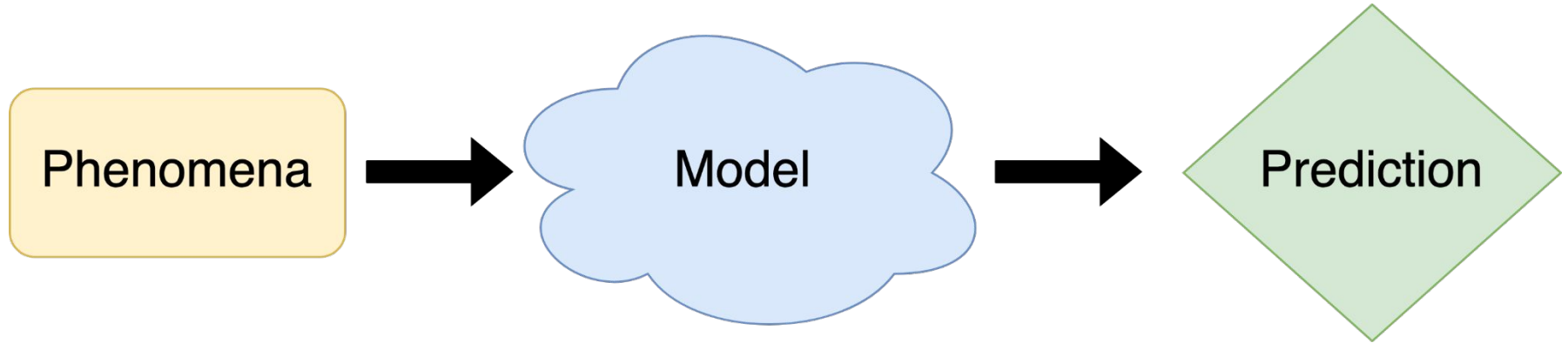# What is Statistical Learning Theory? (SLT)

What is SlT?

- Artificial intelligence covers a wide range of applications
- Machine learning is a particular approach
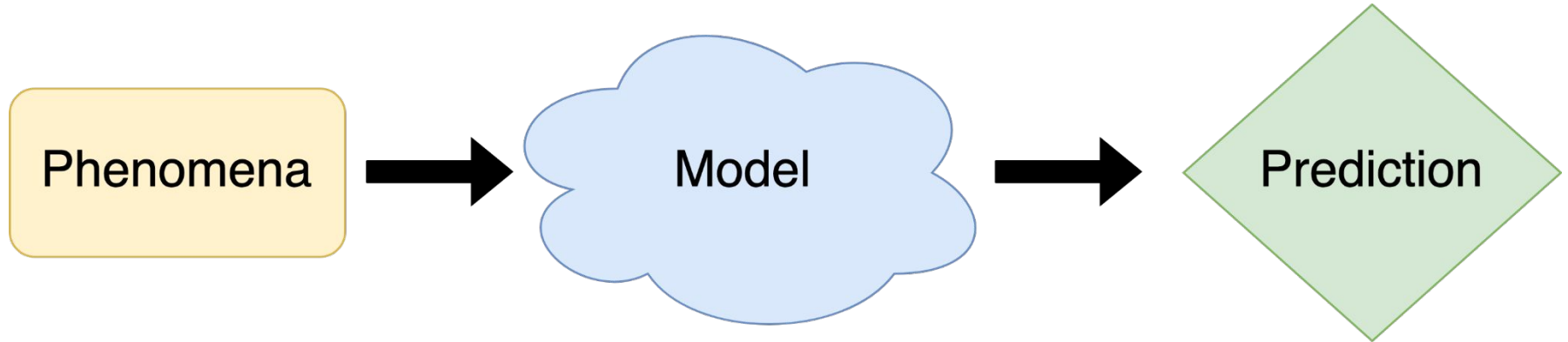- SLT governs part of the machine learning theory

Artifical Intelligence

Machine Learning

SLT

Not all statistics are equal!

- Classical
- Frequentist/Descriptive(< -- We are here)
- Bayesian/Inferential(< -- And here)
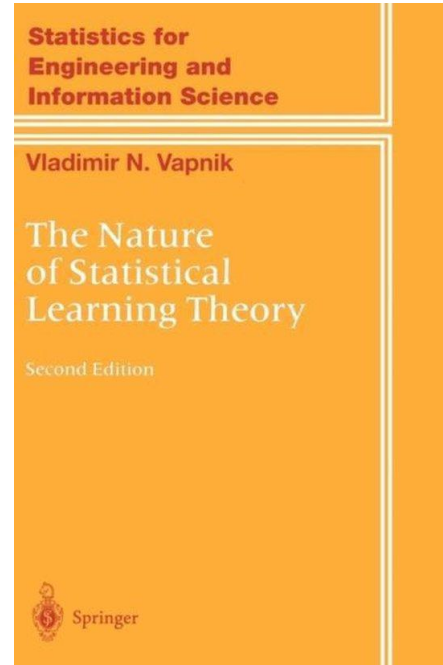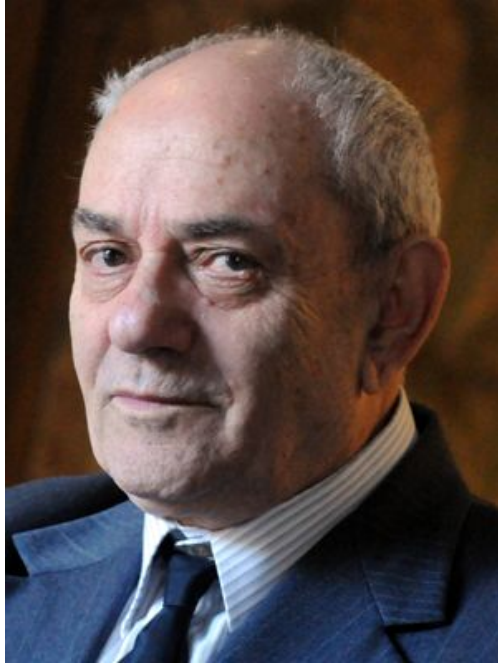- Ergodic

The basic idea

The basic idea

Phenomena → Model → Prediction

We want to automate this process!

In Machine learning terms

Data → Training → Testing

We want to automate this process!

Vapnik's Stastistical Learning

Statistics for Engineering and Information Science

Vladimir N. Vapnik

The Nature of Statistical Learning Theory
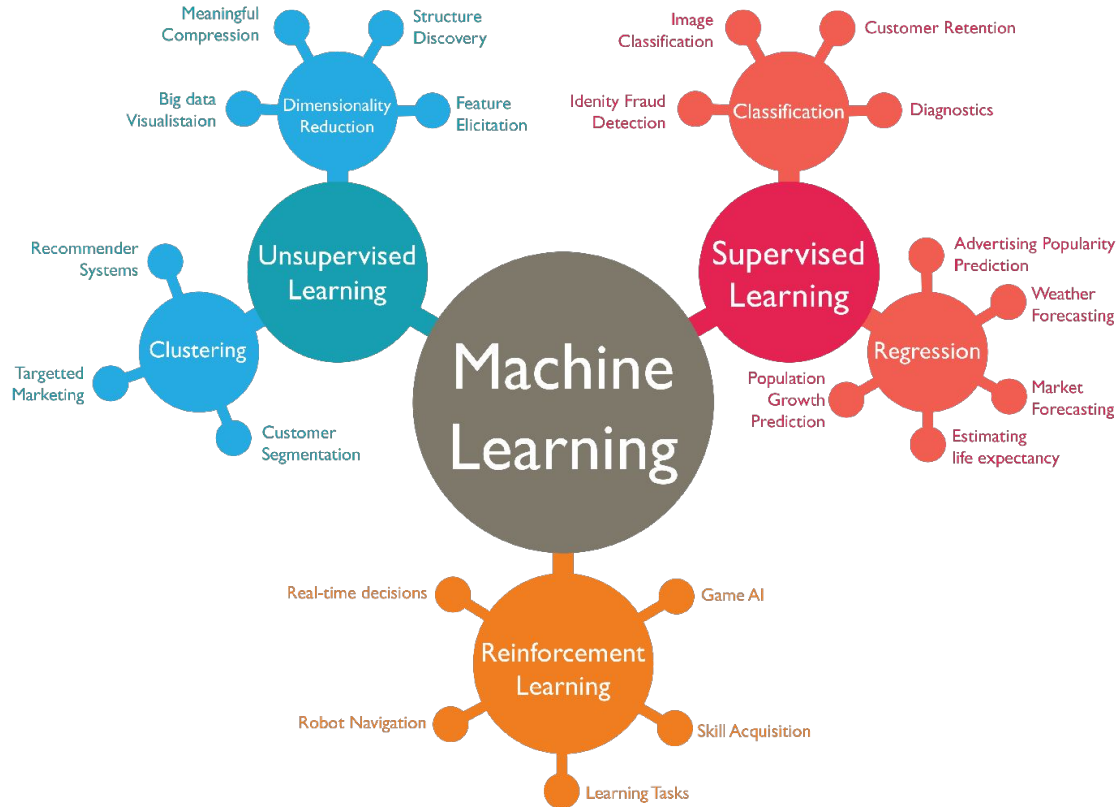
Second Edition

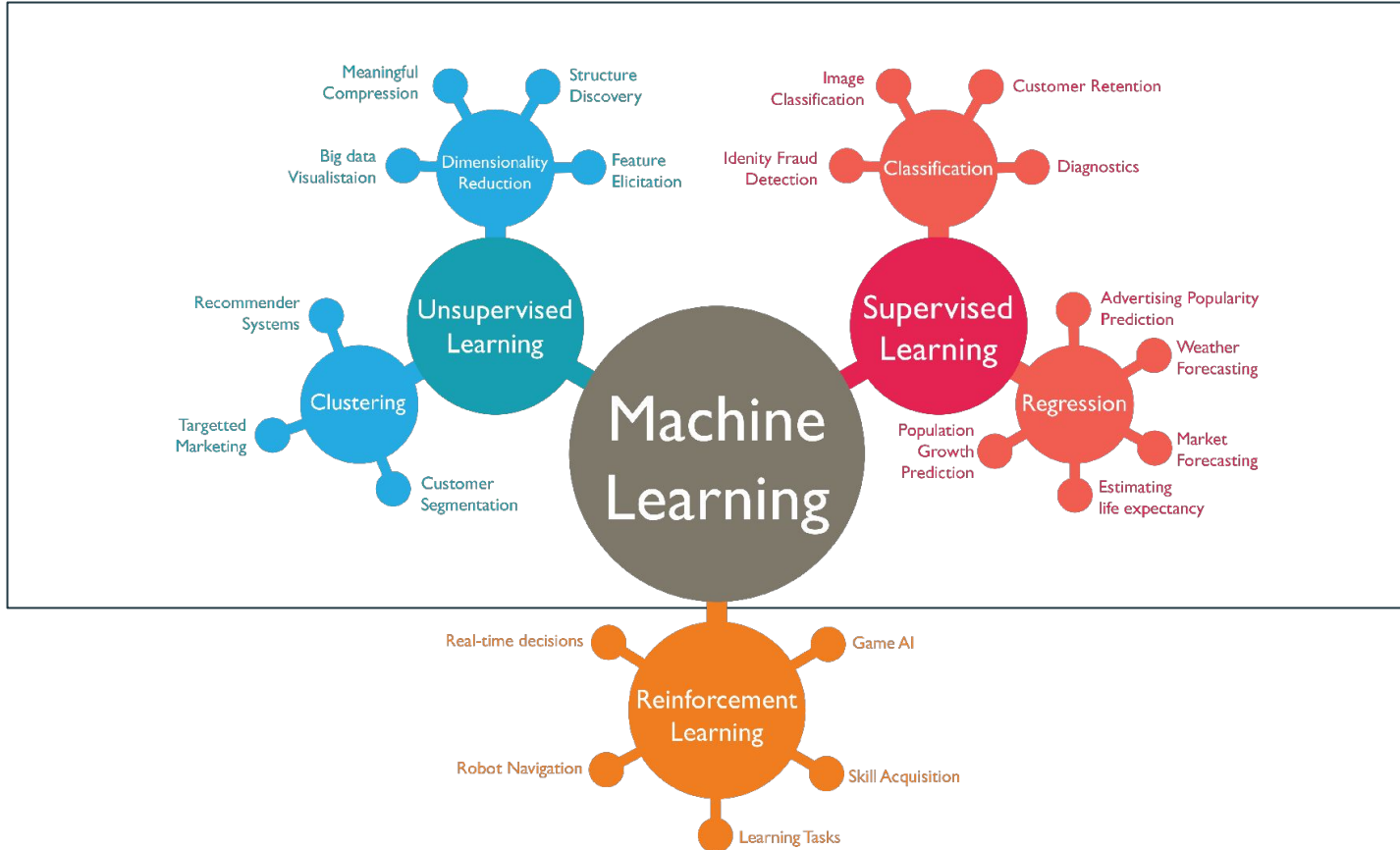Springer

Bayesian Statistics

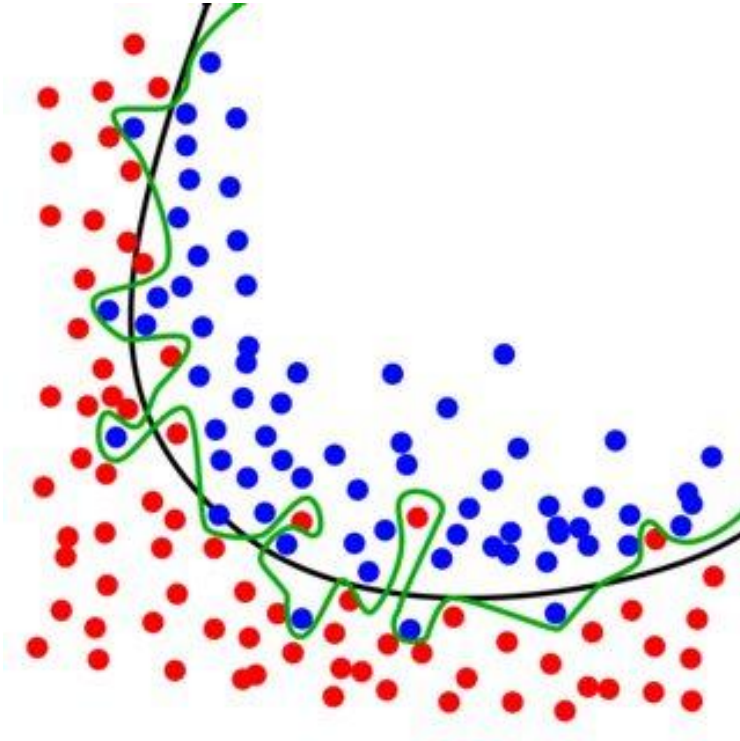$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}$$
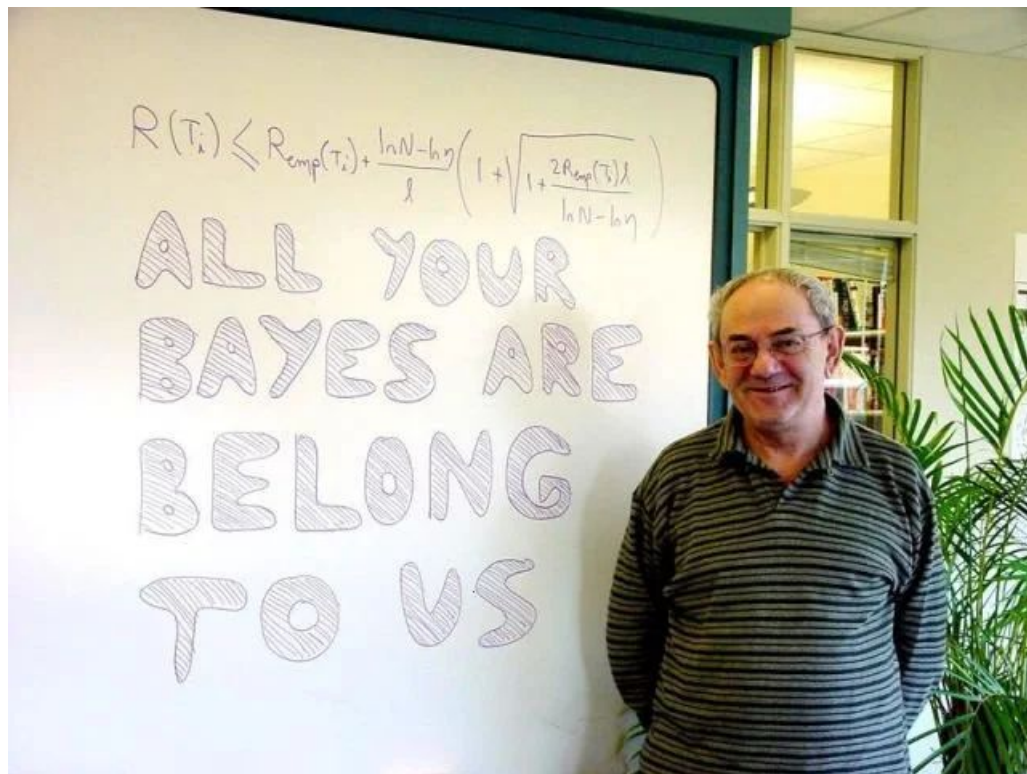
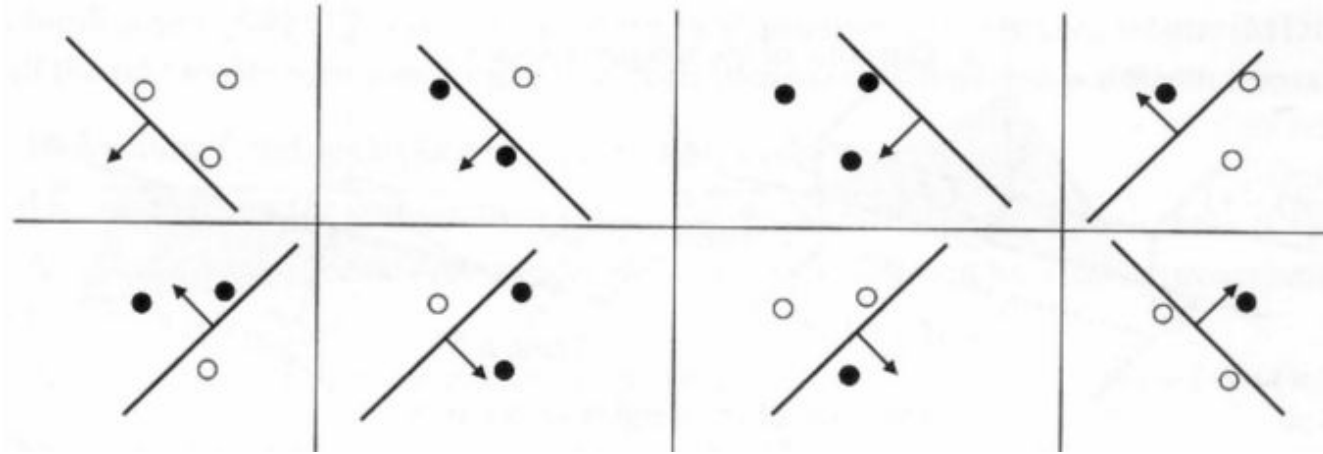# Statistical Learning in AI

Statistical Learning in AI

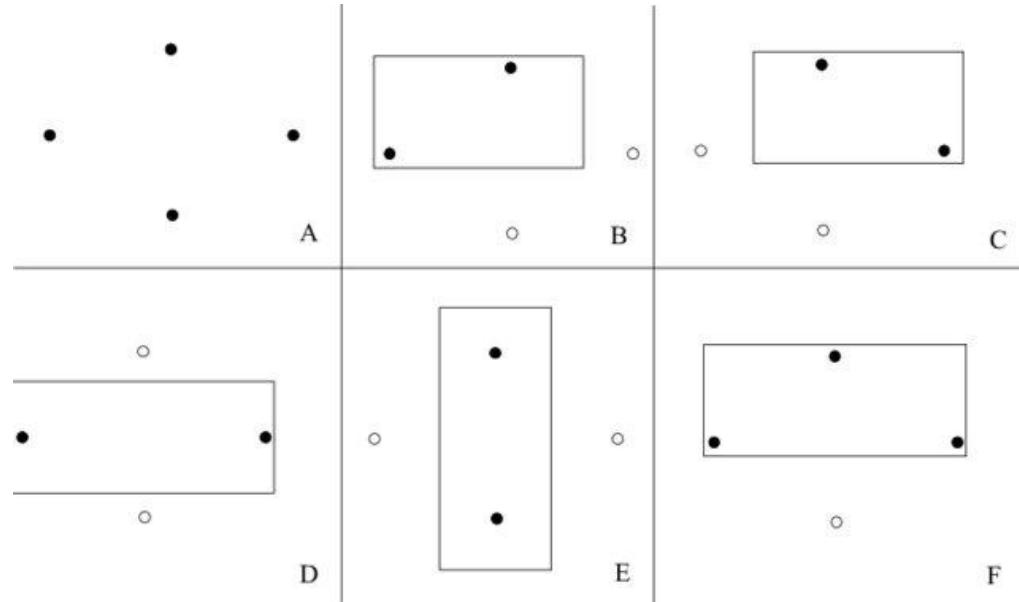# Empirical risk theory

Empirical risk theory

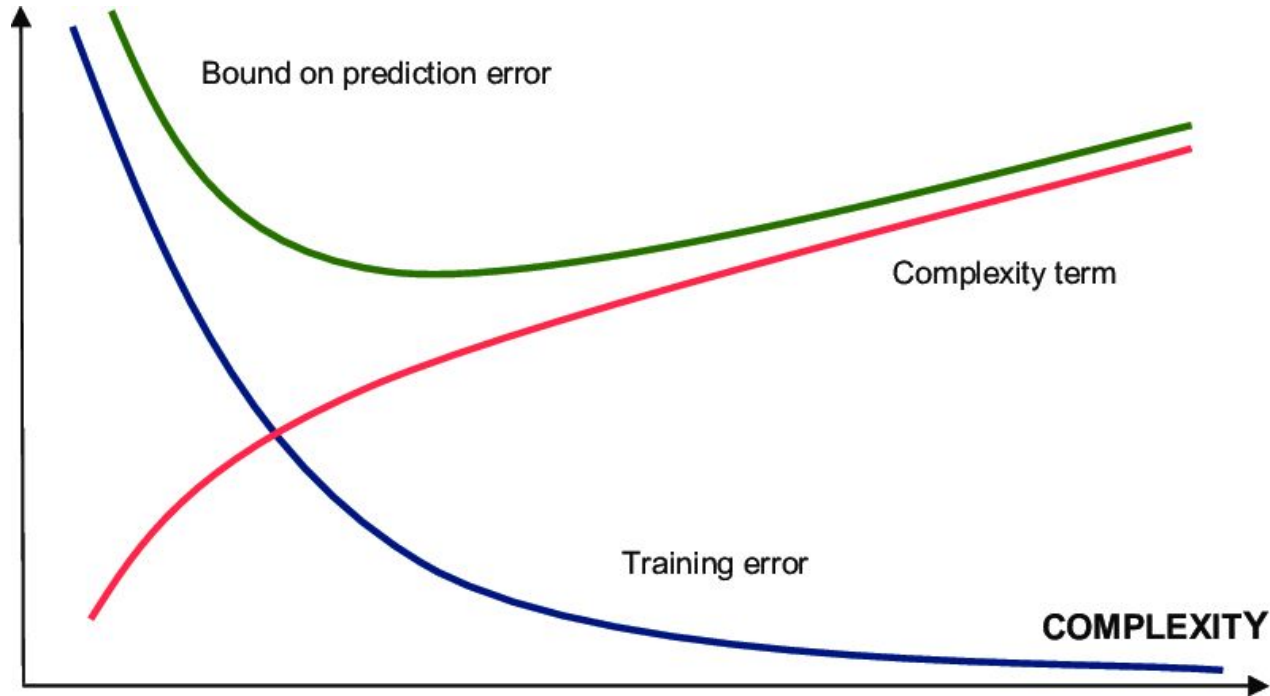# Vapnik Chervonenkis Dimensions

Ref: Theory of VC Dimensions, Medium.com
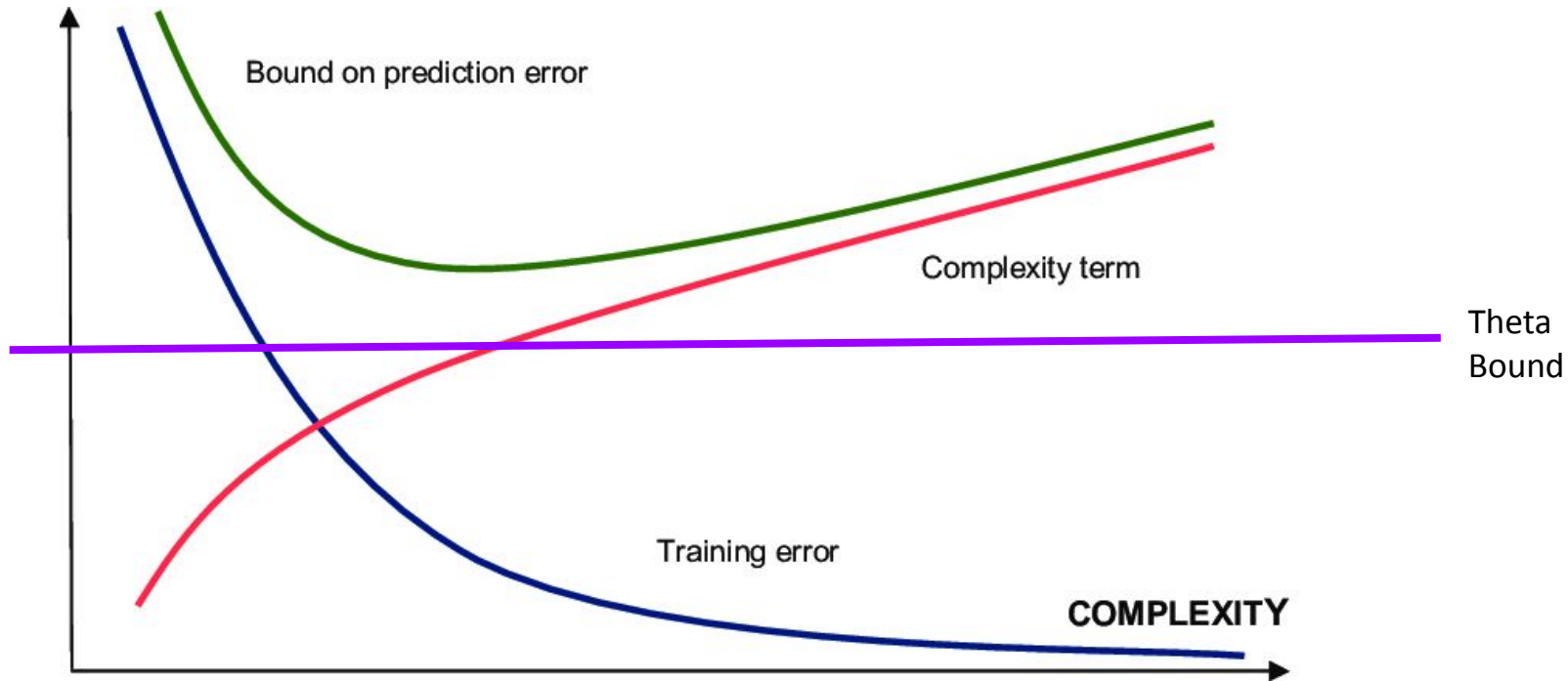
Vapnik Chervonenkis Dimensions

Ref: Kernel-Based Machines for Abstract and Easy Modeling of Automatic Learning,(2011)

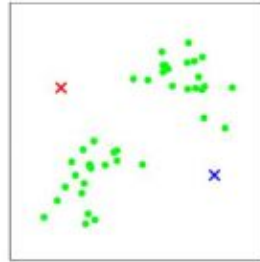# Structural risk minimization
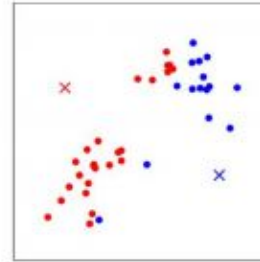
Structural risk minimization

Bound on prediction error

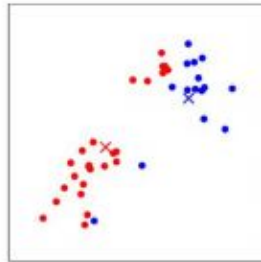Complexity term

Theta Bound

Training error

COMPLEXITY

Some Machine Learning Examples

K means
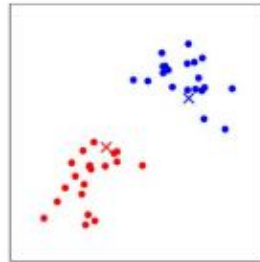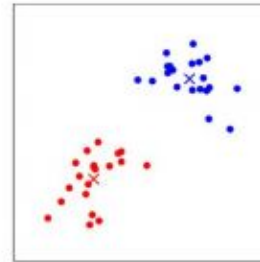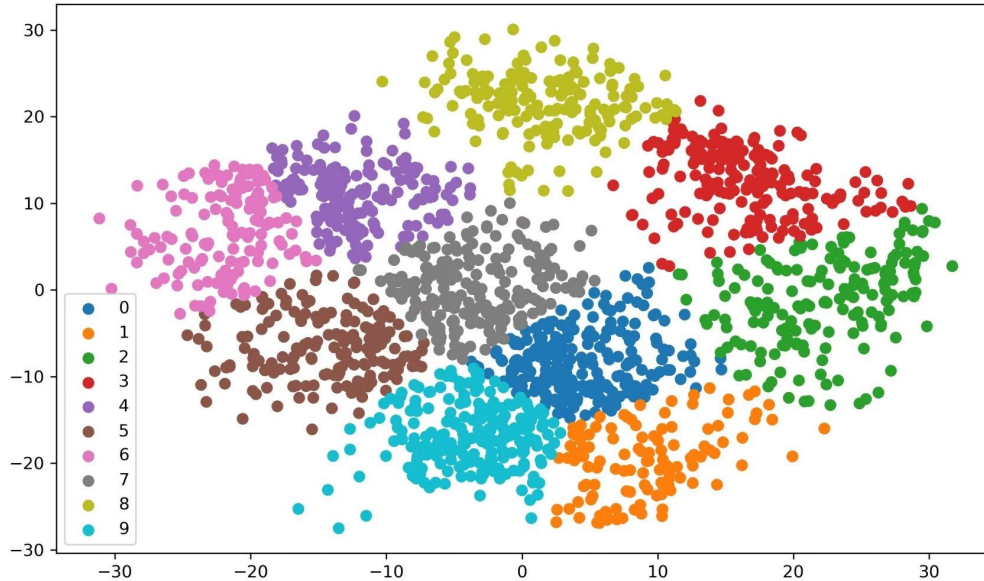
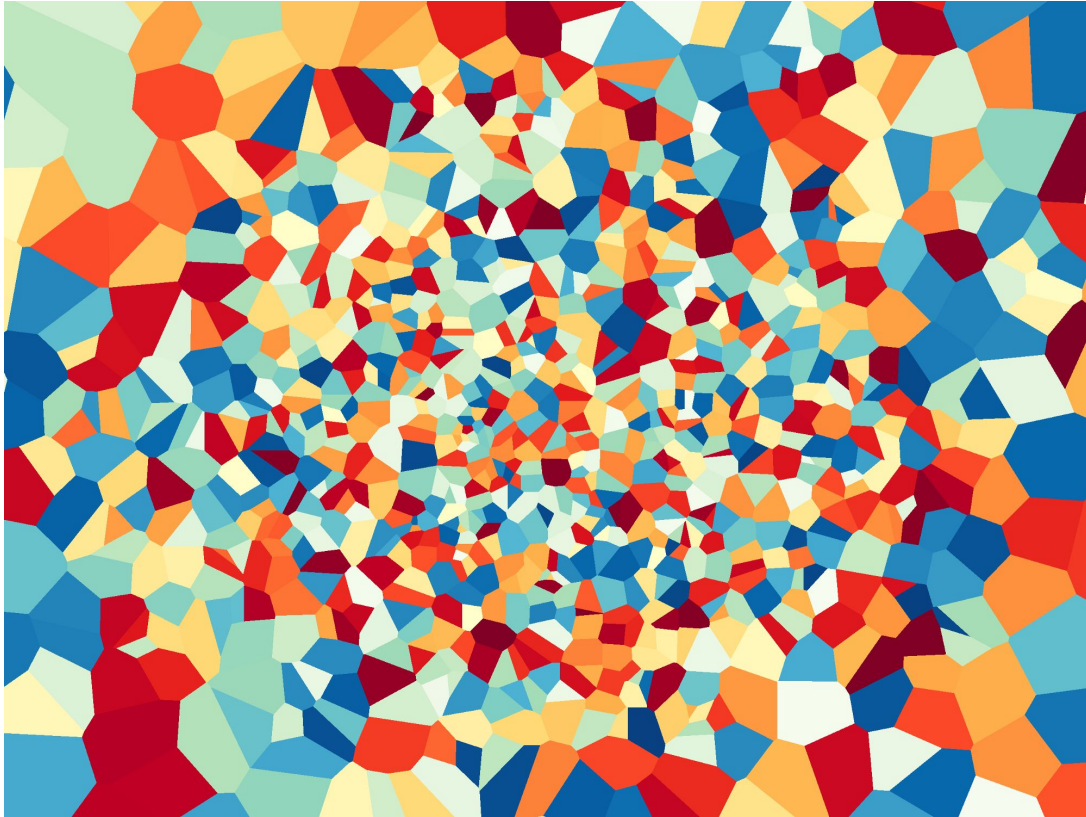

From Stanford's CS221 course

K means extended
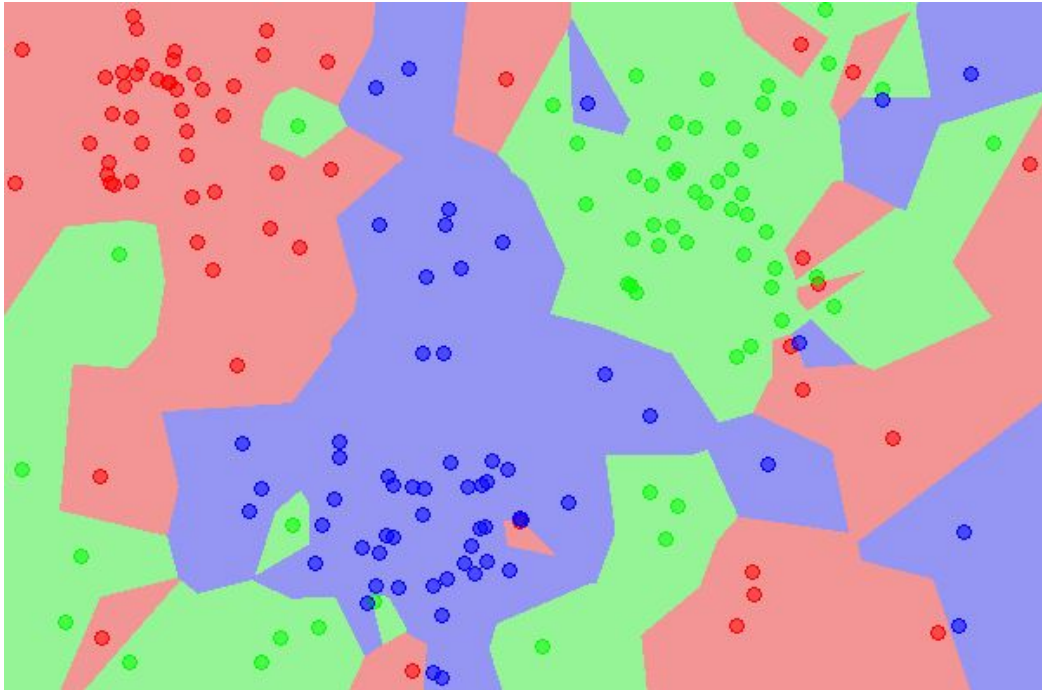

From Askpython's K-Means example implementation

Voronoi spaces

K-nearest neighbours (NN=1)

From **E.M. Mirkes, _KNN and Potential Energy: applet._**
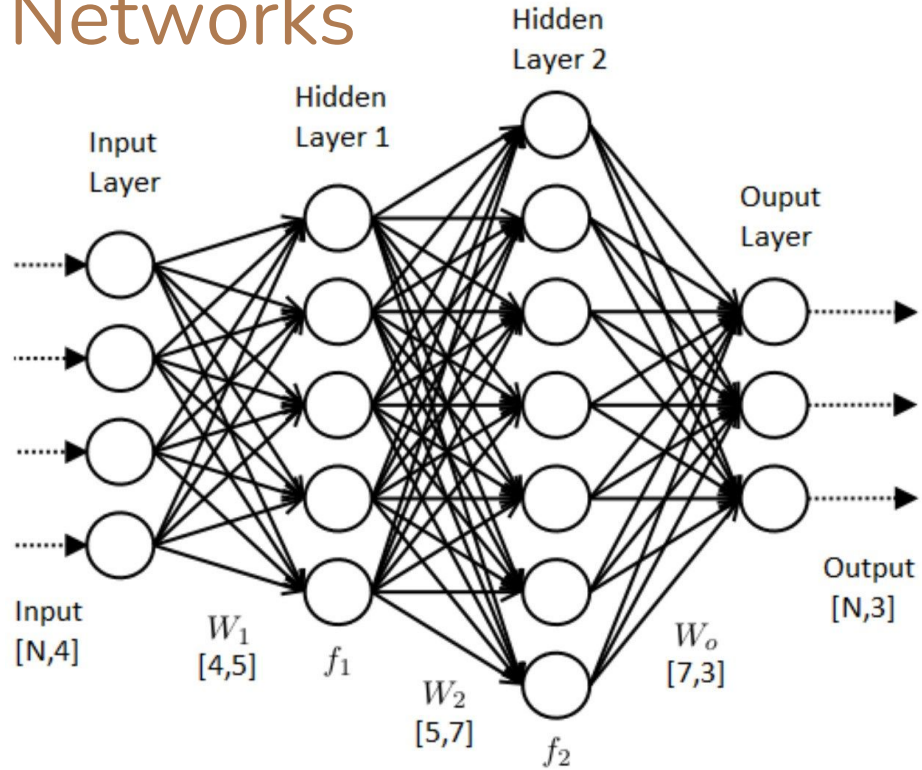**University of Leicester, 2011**

K-nearest neighbours (NN=5)
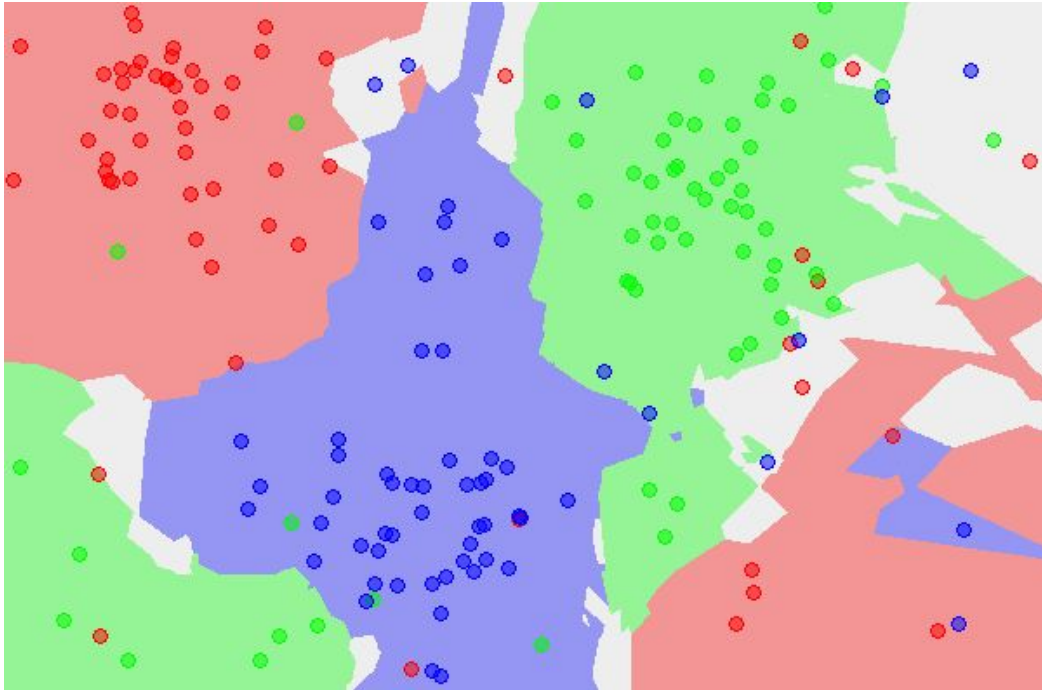
From **E.M. Mirkes, _KNN and Potential Energy: applet._
University of Leicester, 2011**

Support Vector Machines
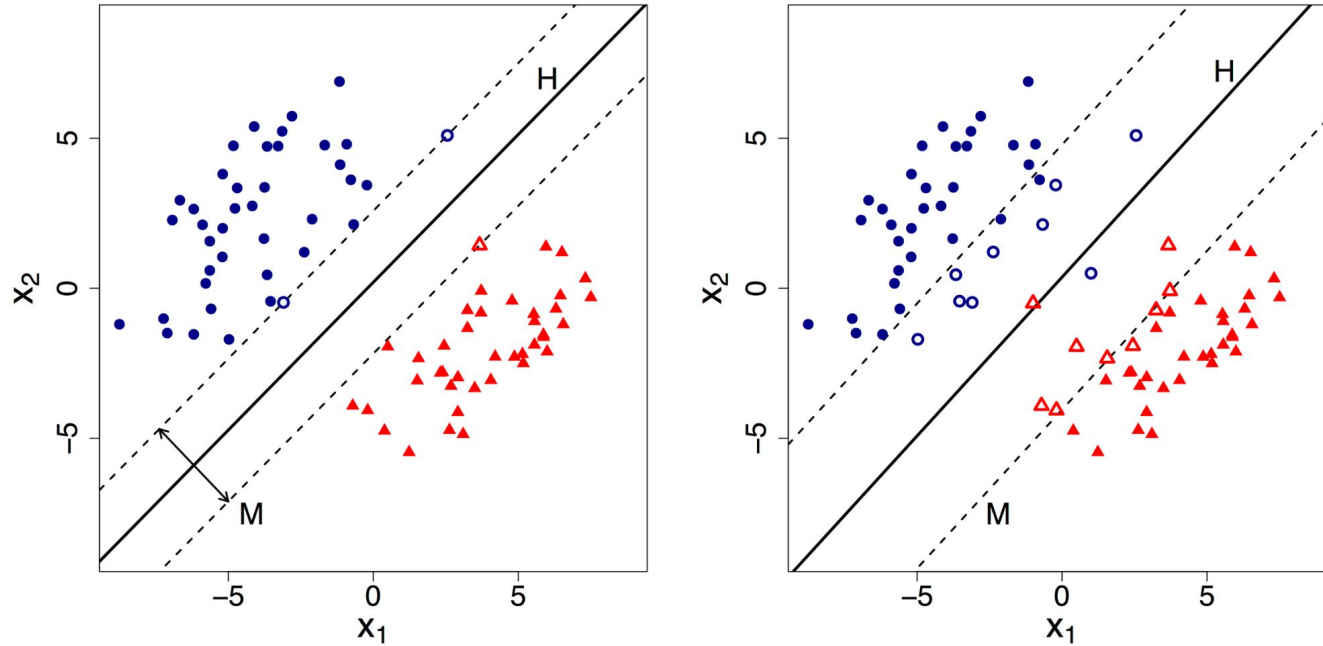
From: Kirchner, Antje, and Curtis S. Signorino. 2018. "Using Support Vector Machines for Survey Research." *Survey Practice* 11

Neural Networks

Neural Networks (1 Perceptron)
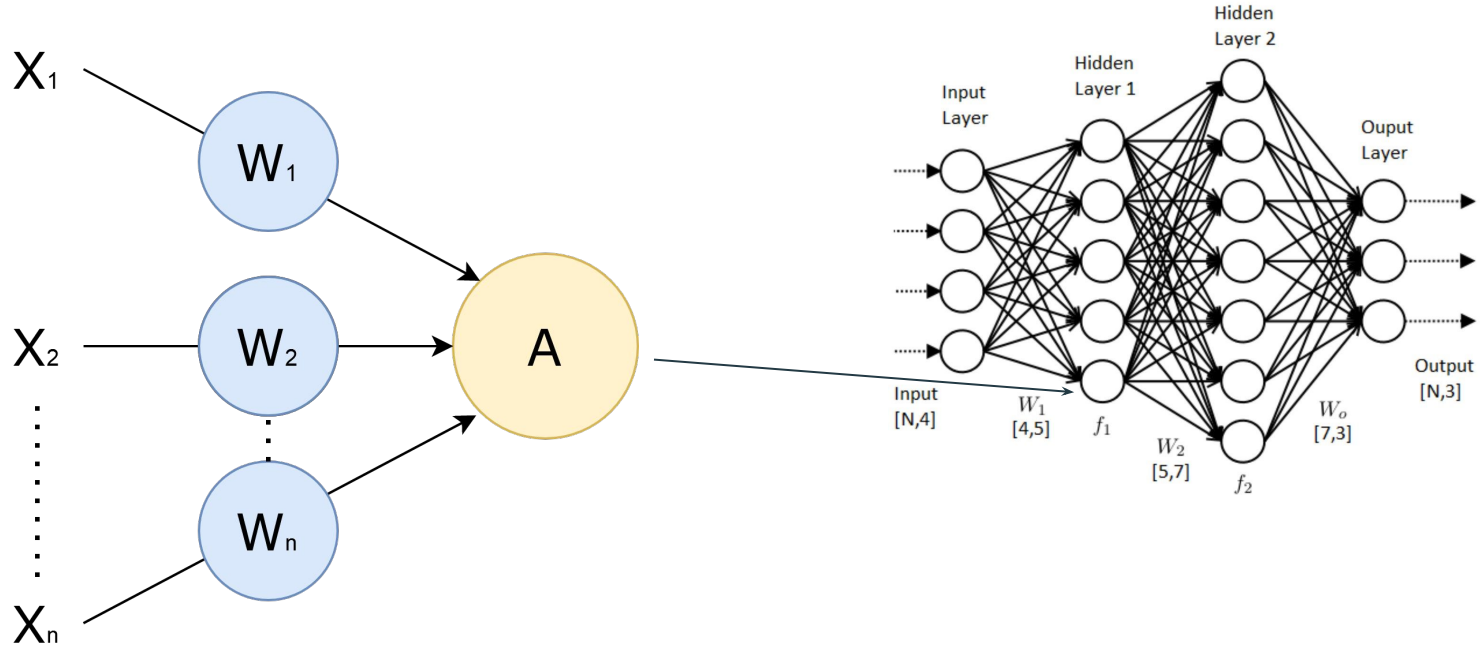
Neural Networks (activation functions)

(a) Plot of identity (linear) function
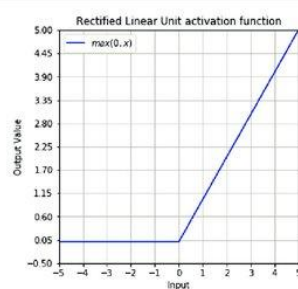
(b) Plot of rectified linear unit function
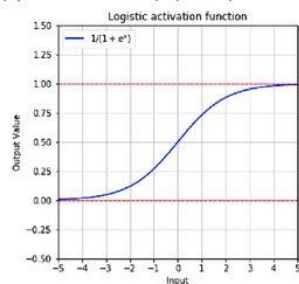
(c) Plot of hyperbolic tangent function
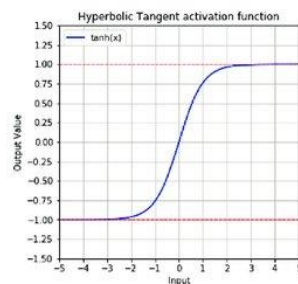
(d) Plot of logistic (sigmoid) function

From: Baressi Šegota, Sandi & Lorencin, Ivan & Musulin, Jelena & Štifanić, Daniel & Car, Zlatan. (2020).

Neural Networks (1 Perceptron)

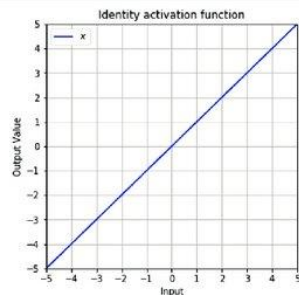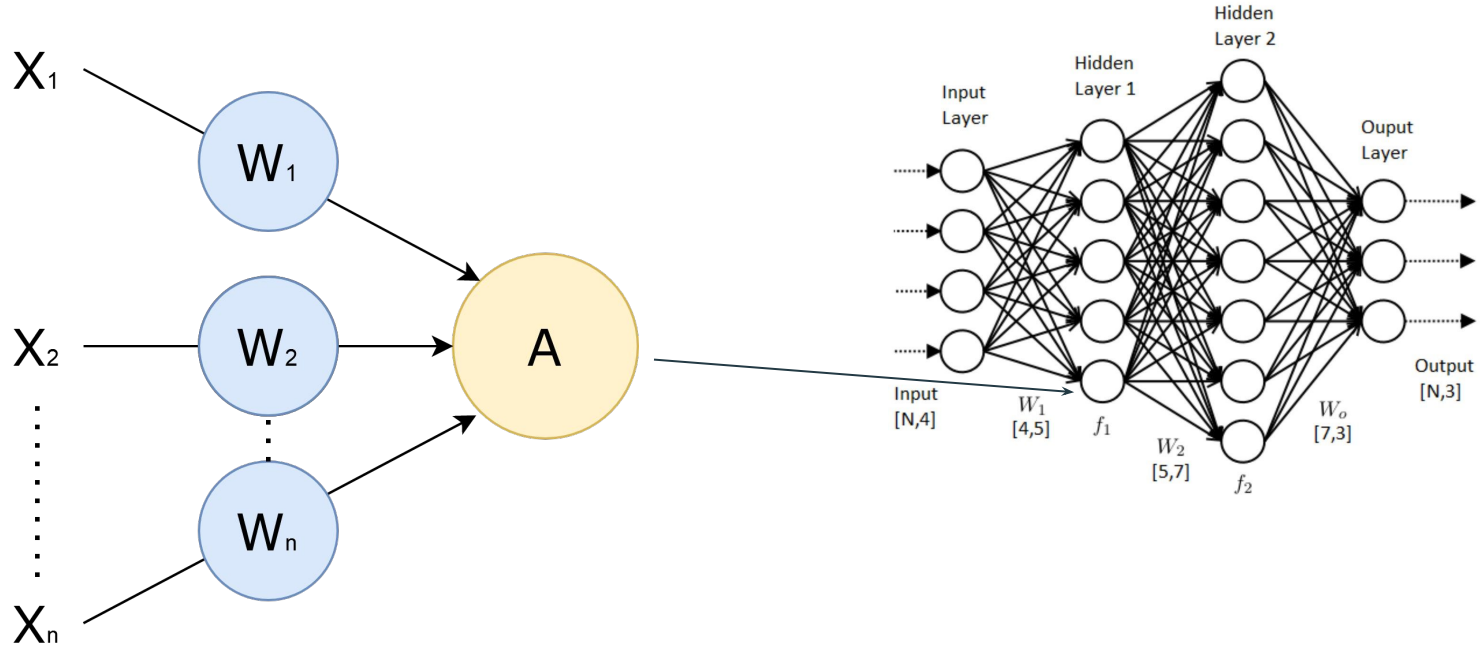# Neural Networks application

Prediction

Classification

Generation

Ref: https://stackoverflow.com

Ref: https://towardsdatascience.com

Ref :https://deepart.io/
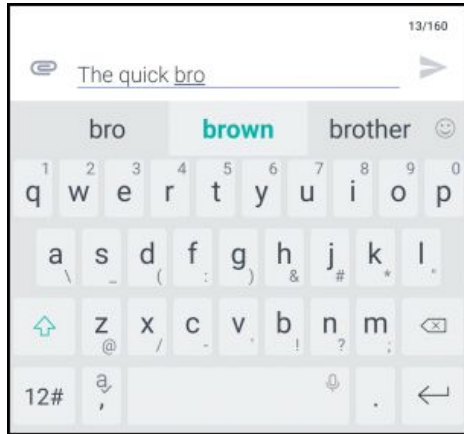
Neural Networks application

Alphafold

de Almeida, A.F., Moreira, R. & Rodrigues, T(2019)

Some takeaways

- Every machine learning algorithm is a balance between ERM and SRM

- We therefore look at a balance between accuracy and simplicity

- Think about your problem in visual space (if at all possible)

Literature

Statistics for Engineering and Information Science

Vladimir N. Vapnik
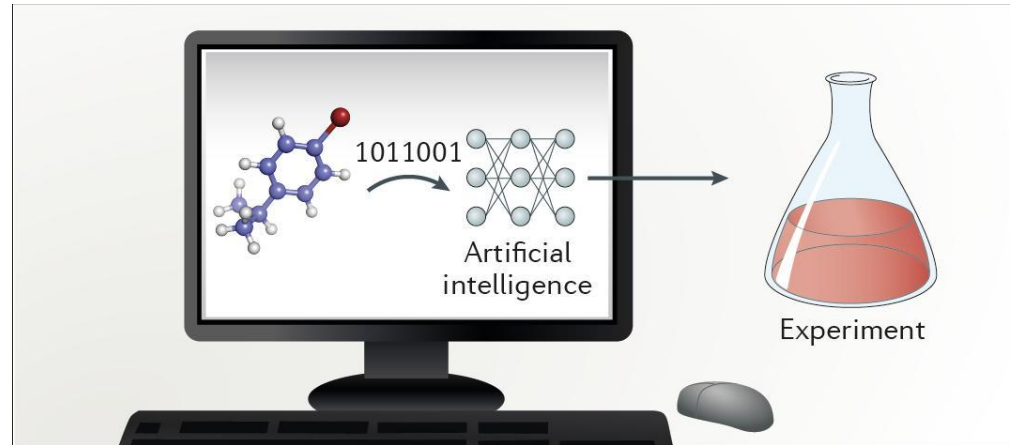
The Nature of Statistical Learning Theory

Second Edition

Springer

Reliable Reasoning

INDUCTION AND STATISTICAL LEARNING THEORY

GILBERT HARMAN AND SANJEEV KULKARNI

Kevin Gurney

AN INTRODUCTION TO NEURAL NETWORKS

Artificial Intelligence
A Modern Approach
SECOND EDITION

The Intelligent Agent Book

Stuart Russell • Peter Norvig

Prentice Hall Series in Artificial Intelligence