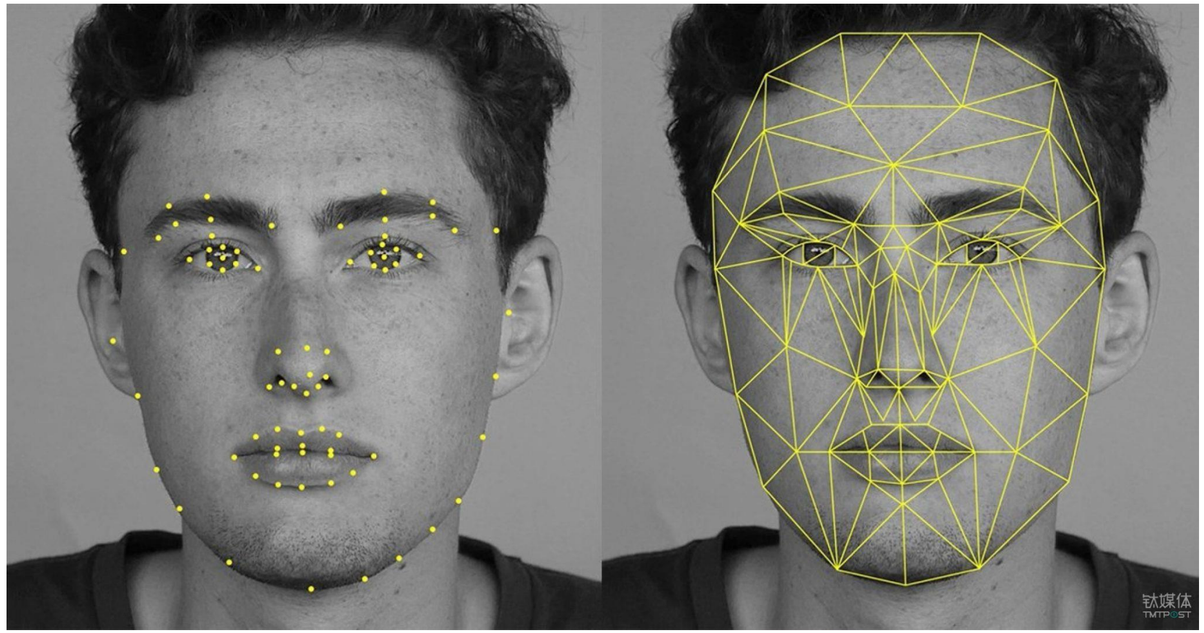
Learning feasibility

By : Dr. Mustafa M. Shiple

Introduction

Learning is used when

- A pattern exists
- We cannot pin it down mathematically
- We have data on it



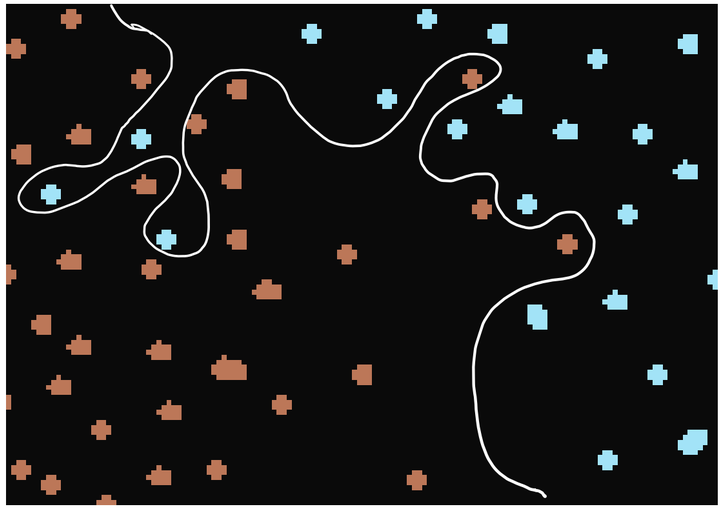
❑**Input space (χ):** patient's medical history, symptoms, personal health information etc.

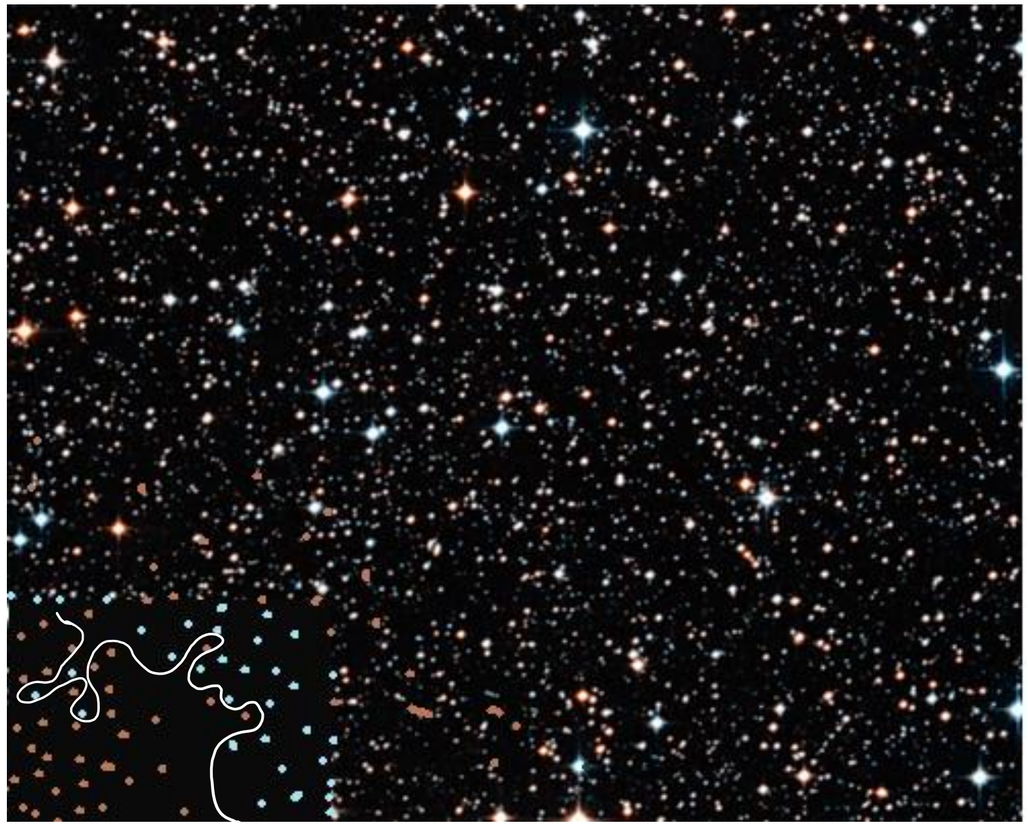
❑**Output space (y):** all possible diseases

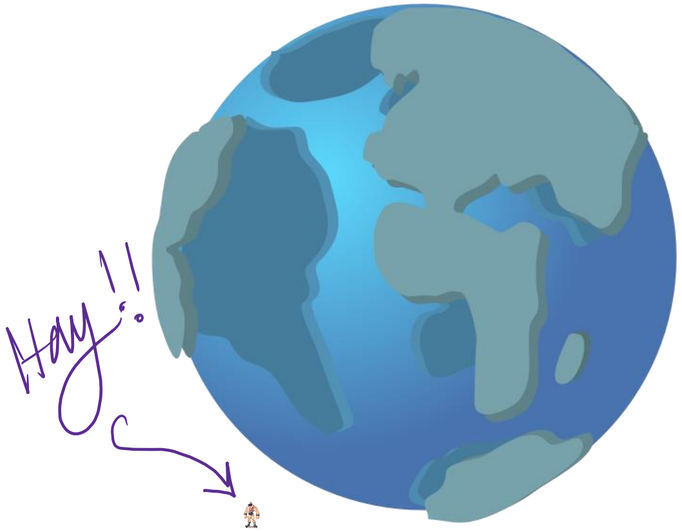
❑**Target function $f : \chi \rightarrow y$** : ideal formula to identify a patient's problem

❑**Data set:** All available patients' information and their corresponding correct problem diagnostic.

whole image



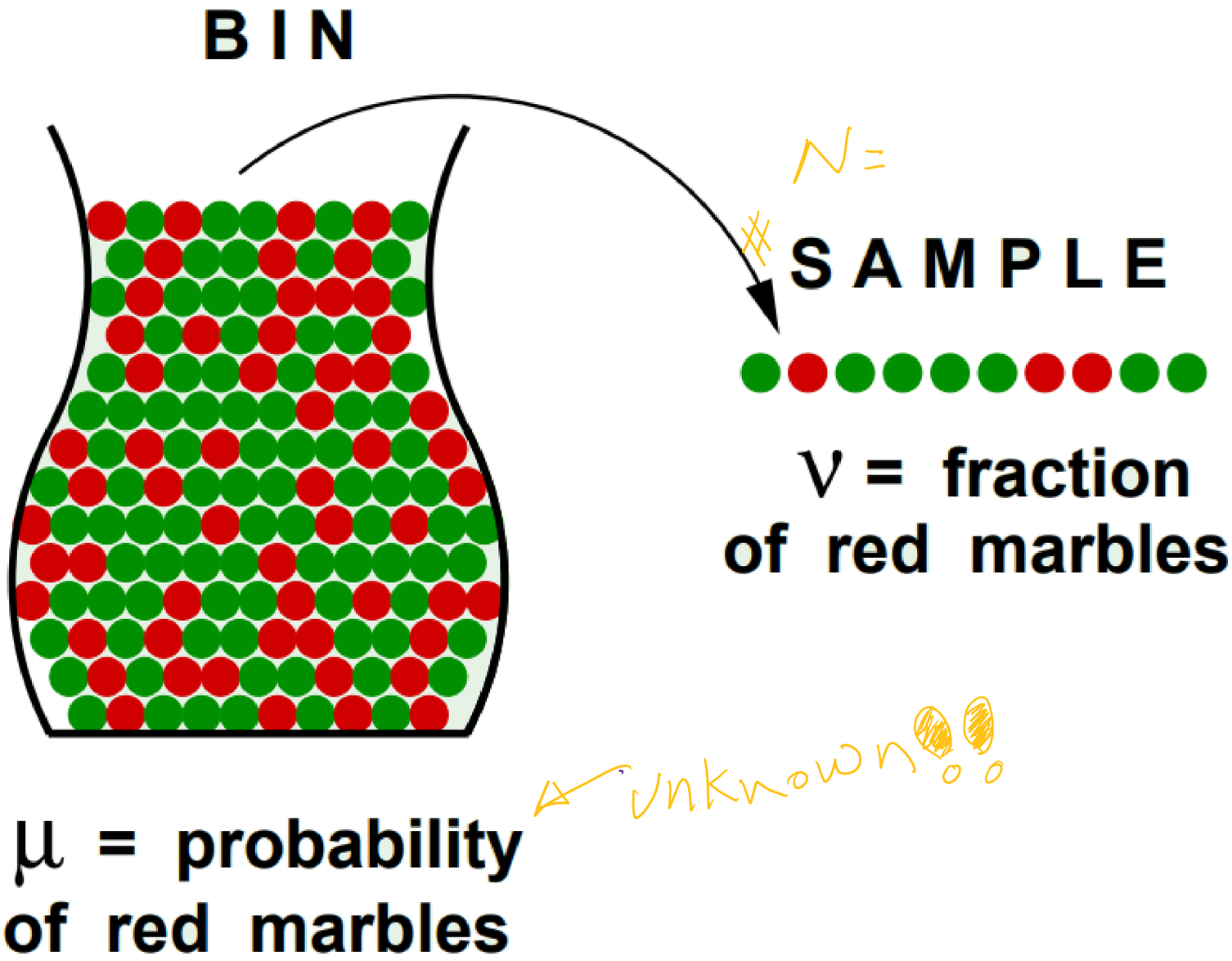




Feasibility of Learning

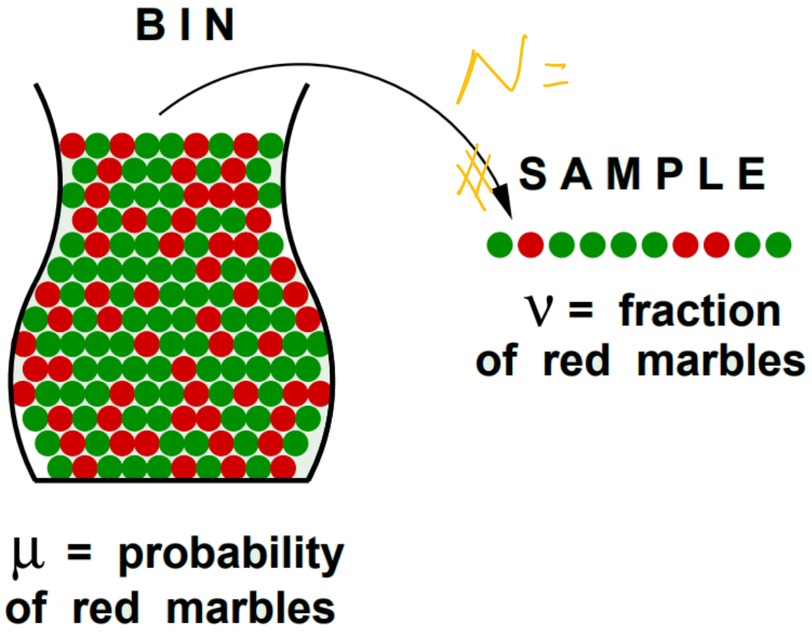
Introduction

- Consider a 'bin' with red and green marbles.
- ℙ [picking a red marble] = μ
- ℙ[picking a green marble] = 1 — μ



*Learning from data, Yaser S. Abu-Mostafa

Introduction



$$\mathbb{P}[|\nu - \mu| > \epsilon] \leq$$

If N is large => $\nu = \mu \Rightarrow \varepsilon = 0$

But if N is not that much ☹

$$\mathbb{P}[|\nu - \mu| > \epsilon] \leq e^{-N}$$

IS that all what we have !!!!

$$\mathbb{P}[|\nu - \mu| > \epsilon] \leq e^{-\epsilon^2 N}$$

Hoeffding's inequality

$$\mathbb{P}[|\nu - \mu| > \epsilon] \leq 2e^{-2\epsilon^2 N}$$

*Learning from data, Yaser S. Abu-Mostafa

Hoeffding's inequality

$$\mathbb{P}[|\nu - \mu| > \epsilon] \leq 2e^{-2\epsilon^2 N}$$

$$\mu = \nu$$

P.A.C (Probable approximate correct)

Not dependent on μ

trade off $\epsilon \; and \; N$

*Learning from data, Yaser S. Abu-Mostafa

Thank you