

01 - Introduction

Ethan Graham

20th June 2023

Supervised Classification

This is when we directly look at input and output labels. We represent data as **feature vectors**

example: *feature vector:* What films have you watched? Did you like them?
predictor: List of films to propose

Error function

Machine learning often boils down to minimizing an error function. For example: We try to split seabasses and salmons by brightness. The error function becomes $E(w) = n_{salmons \text{ with brightness } < T} + n_{seabasses \text{ with brightness } \geq T}$

We minimize $E(W)$ w.r.t to w to find the optimal parameters. (w is the model parameter vector)

we denote $y(x, w)$ the predicted label. We want $y(x, w) > 0$: Salmon, $y(x, w) < 0$: Bass. This gives us **decision boundary** $C = \{x \in R^2, y(x, w) = 0\}$

commonly $y(x, w) = w_0 + w_1x_1 + w_2x_2 + \dots + w_nx_n$