

Introduction to Machine Learning

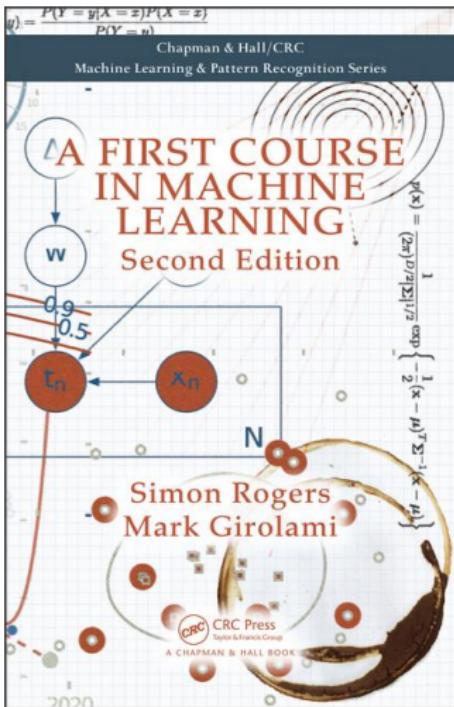
Mauricio A. Álvarez

Foundations of Machine Learning
The University of Manchester

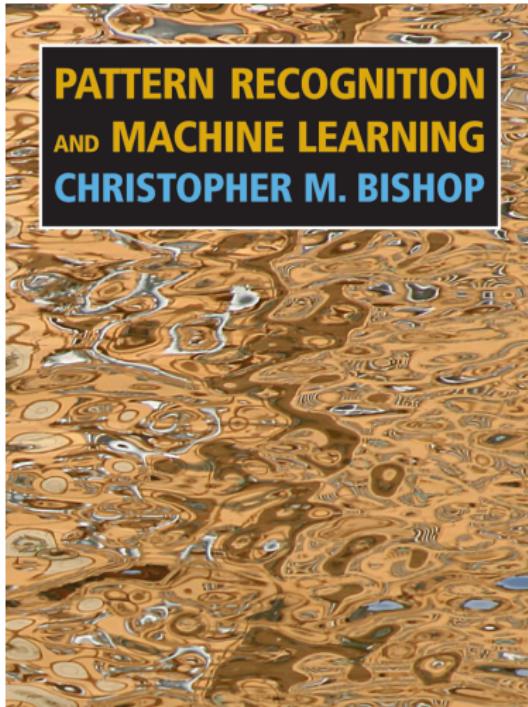


The University of Manchester

Textbooks

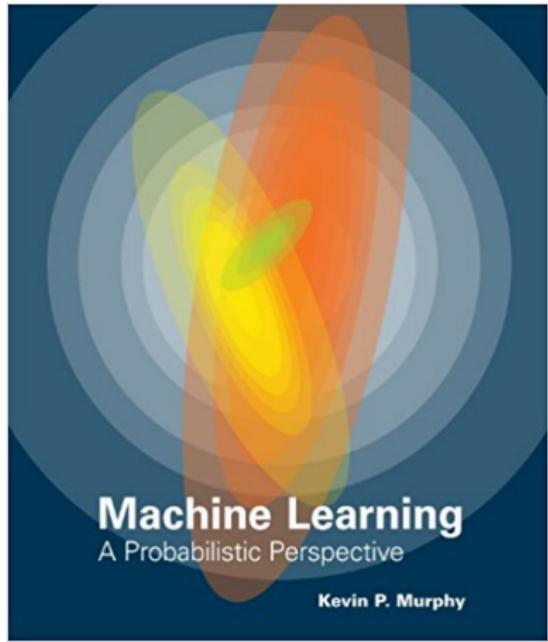


Rogers and Girolami, *A First Course in Machine Learning*, Chapman and Hall/CRC Press, 2nd Edition, 2016.

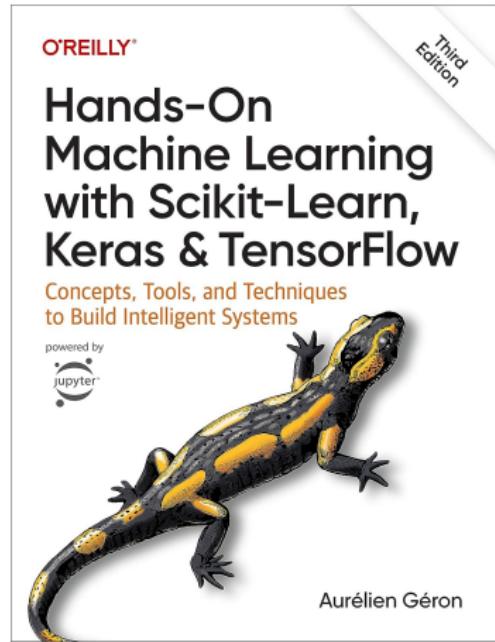


Bishop, *Pattern Recognition and Machine Learning*, Springer-Verlag, 2006.

Textbooks

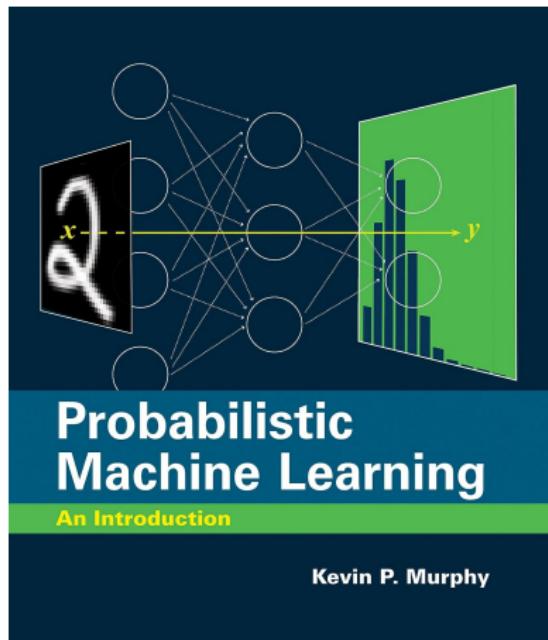


Murphy, *Machine Learning: A Probabilistic Perspective*, MIT Press, 2012.

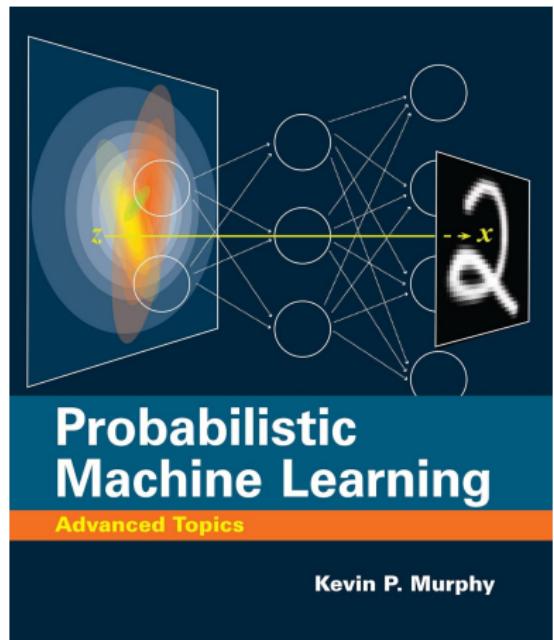


Géron, *Hands-On Machine Learning with Scikit-Learn, Keras and TensorFlow*, O'Reilly, 3rd Edition, 2022.

Textbooks

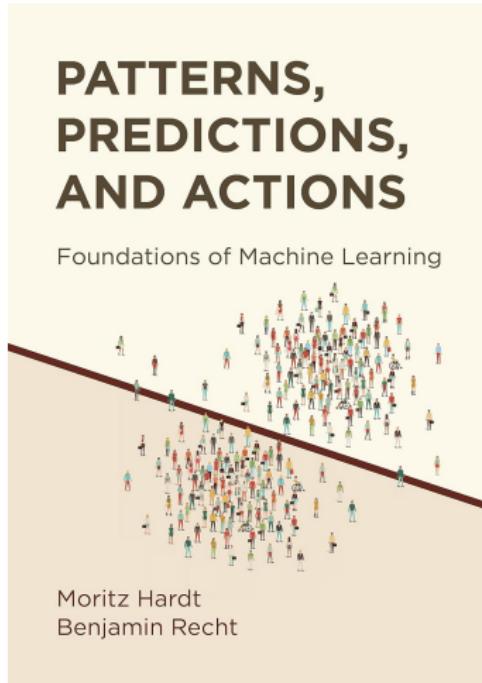


Murphy, *Probabilistic Machine Learning: An Introduction*, MIT Press, 2022.

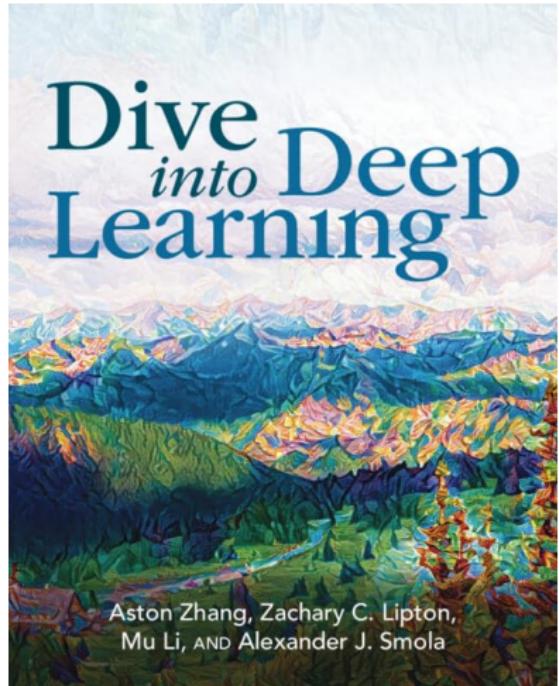


Murphy, *Probabilistic Machine Learning: Advanced Topics*, MIT Press, 2023.

Textbooks



Hardt and Recht, *Patterns, Predictions, and Actions: Foundations of Machine Learning*, Princeton University Press, 2022.



Zhang et al., *Dive into Deep Learning*, Cambridge University Press, 2023.

Contents

Machine learning

Definitions

An example of a predictive model

Machine learning or Statistical Learning

- We would like to design an algorithm that help us to solve different prediction problems.
- The algorithm is designed based on a mathematical model or function, and a dataset.
- Extract knowlegde from data.

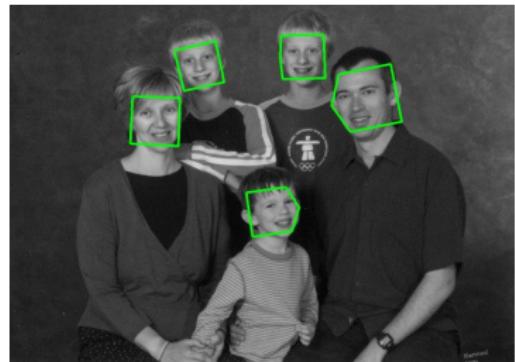
Examples of ML problems

Handwritten digit recognition



Examples of ML problems

Face detection and face recognition



From Murphy (2012).

Examples of ML problems

Predicting the age of a person looking at a particular YouTube video.



Examples of ML problems

Stock market



Examples of ML problems

Clustering: segmenting customers in e-commerce



Examples of ML problems

Recommendation systems

Customers Who Bought This Item Also Bought

Page 1 of 17

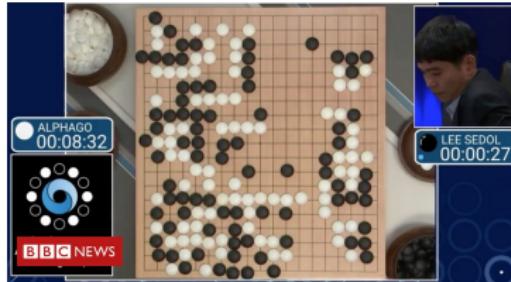
The screenshot shows a row of six book covers from an Amazon search result. Each book has a 'LOOK INSIDE' button at the top right. The books are:

- Machine Learning: A Probabilistic... by Kevin P. Murphy (4 stars, 35 reviews, \$81.71 Prime)
- The Elements of... by Trevor Hastie (4 stars, 40 reviews, #1 Best Seller in Bioinformatics)
- Probabilistic Graphical Models: Principles and... by Daphne Koller (4 stars, 26 reviews, Hardcover, \$99.75 Prime)
- Machine Learning with R by Brett Lantz (4 stars, 26 reviews, Paperback, \$49.49 Prime)
- An Introduction to... by Gareth James (4 stars, 37 reviews, #1 Best Seller in Mathematical & Statistical...)
- Reinforcement Learning: An Introduction... by Richard S. Sutton (4 stars, 17 reviews, Hardcover, \$64.60 Prime)

Navigation arrows are visible on the left and right sides of the row.



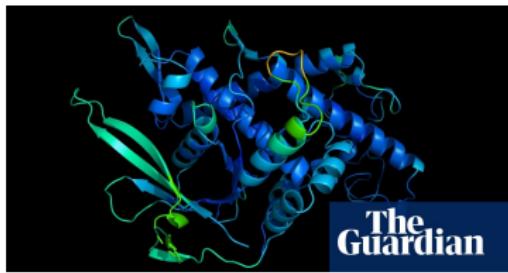
ML has contributed to advances in AI



AlphaGo



Autonomous driving



AlphaFold

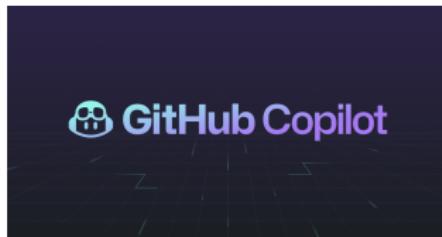
Generative AI



DALL-E



ChatGPT



Github Copilot

Contents

Machine learning

Definitions

An example of a predictive model

Basic definitions

- Handwritten digit recognition



- Variability
- Each image can be transformed into a vector \mathbf{x} (feature extraction).
- An instance is made of the pair (\mathbf{x}, y) , where y is the label of the image.
- Objective: find a function $f(\mathbf{x}, \mathbf{w})$.

Basic definitions

- **Training set:** a set of N images and their labels $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)$, to fit the predictive model.
- **Estimation or training phase:** process of getting the values of \mathbf{w} of the function $f(\mathbf{x}, \mathbf{w})$, that best fit the data.
- **Generalisation:** ability to correctly predict the label of new images \mathbf{x}_* .

Supervised and unsupervised learning

- **Supervised learning:**
 - Variable y is discrete: *classification*.
 - Variable y is continuous: *regression*.
- **Unsupervised learning:** from the set $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)$, we only have access to $\mathbf{x}_1, \dots, \mathbf{x}_N$
 - Find similar groups: *clustering*.
 - Find a probability function for \mathbf{x} : *density estimation*.
 - Find a lower dimensionality representation for \mathbf{x} : *dimensionality reduction and visualisation*.
- **Other types of learning:** reinforcement learning, semi-supervised learning, active learning, multi-task learning.

Contents

Machine learning

Definitions

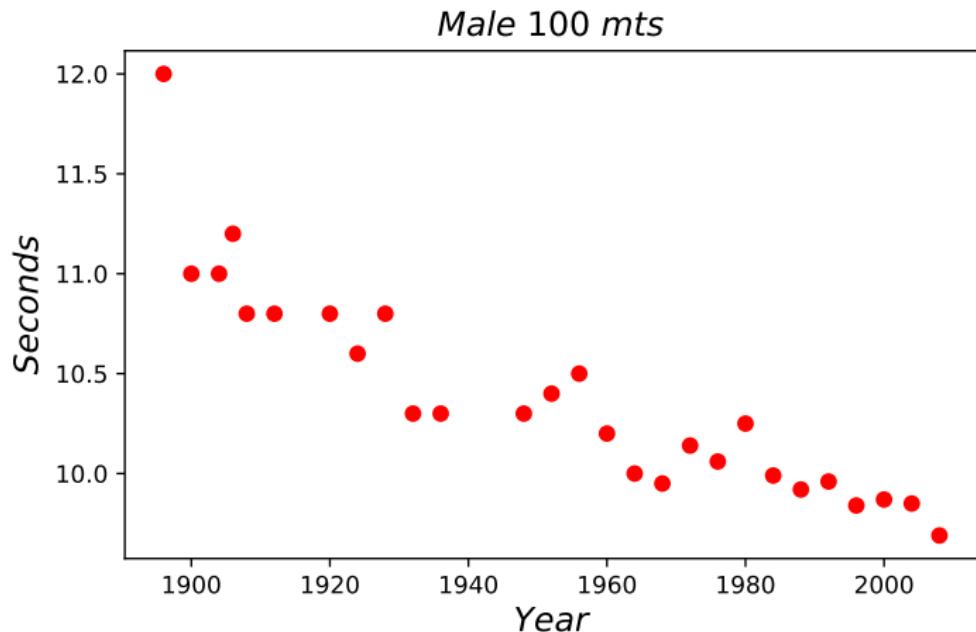
An example of a predictive model

Olympic 100m Data



Image from Wikimedia Commons <http://bit.ly/191adDC>.

Dataset



Model

- We will use a linear model $f(x, \mathbf{w})$, where y is the time in seconds and x the year of the competition.
- The linear model is given as

$$y = w_1 x + w_0,$$

where w_0 is the intercept and w_1 is the slope.

- We use \mathbf{w} to refer both to w_0 and w_1 .

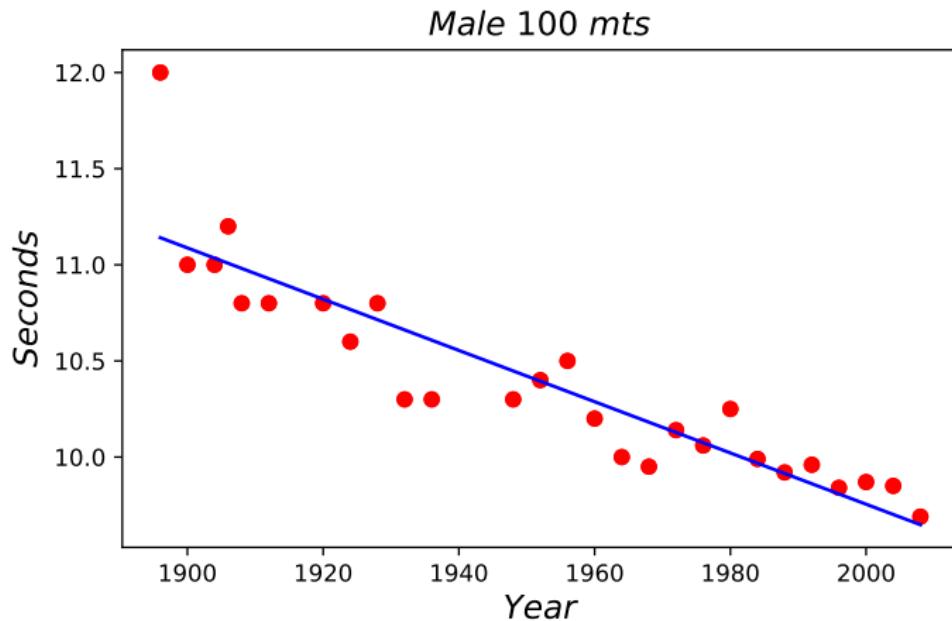
Objective function

- We use an objective function to estimate the parameters w_0 and w_1 that best fit the data.
- In this example, we use a least squares objective function

$$E(w_0, w_1) = \sum_{\forall i} (y_i - f(x_i))^2 = \sum_{\forall i} [y_i - (w_1 x_i + w_0)]^2.$$

- By minimising the error with respect to \mathbf{w} , we get the solution as $w_0 = 36.4$ and $w_1 = -1.34 \times 10^{-2}$.

Data and model



Predictions

- We can now use this model for making predictions.
- For example, what does the model predict for 2012?
- If we say $x = 2012$, then

$$\begin{aligned}y &= f(x, \mathbf{w}) = f(x = 2012, \mathbf{w}) \\&= w_1 x + w_0 = (-1.34 \times 10^{-2}) \times 2012 + 36.4 = 9.59.\end{aligned}$$

- The actual value was 9.63.

Main challenges of machine learning

- Insufficient quantity of training data.
- Nonrepresentative training data.
- Poor-quality data.
- Irrelevant features.
- Overfitting the training data.
- Underfitting the training data.