

# Introduction to Machine Learning

Maheśan Nirānjan  
University of Southampton

DiscNet Summer School

June 2025

## Modules We Offer in Southampton

- Foundations of Machine Learning
  - Machine Learning Technologies
  - Advanced Machine Learning
  - Differential Programming and Deep Learning
  - Deep Learning Technologies
  - Bayesian, Active & Reinforcement Learning
  - Natural Language Processing
  - Causal Reasoning and Machine Learning
  - Computational Biology
  - Computer Vision
  - Computational Finance
- 
- This week is a random sample from magenta

# Machine Learning: Good employment prospects!



## Data Scientist - Financial Investment / Risk

Intellect UK Group Limited City of London (EC1) Contract Published: 5 days ago  
From £200 to £300 per day

Apply >

Save

Join a leading hedge fund with a strong reputation for innovation and success in the financial market leverage cutting-edge technology and data-driven strategies to deliver exceptional results for our clients. As we continue to expand our capabilities, we are seeking a talented Data Scientist with exp in Axioma and risk analytics to join our dynamic team.

This role will sit outside IR35

Data scientist in quantitative finance  
Intellect UK Group Ltd.  
Hedge Fund  
150 - 200K



## Research Associate (Fixed Term)

University of Cambridge - MRC Biostatistics Unit

Apply

An opportunity has arisen for a talented computational statistician or probabilistic machine learning methods developer to join Dr Paul Kirk's group at the MRC Biostatistics Unit, Cambridge University, within the Biostatistical Machine Learning research theme.

Biostatistical Machine Learning (BML) is a new research theme within the MRC Biostatistics Unit, jointly led by Paul Kirk and Sach Mukherjee. BML focuses on cross-cutting, methodological research in machine learning (ML), artificial intelligence (AI), and high-dimensional statistics. The overarching aim is to combine flexible and scalable AI and ML approaches with the need for robustness, interpretability and scientific understanding that is essential in biostatistical applications.

We are seeking an ambitious and motivated individual to contribute to Paul Kirk's group and the broader BML research theme. Depending

Location: Cambridge  
Salary: £36,024 to £44,263  
Hours: Full Time  
Contract Type: Fixed-Term/Contract  
Placed On: 30th August 2024  
Closes: 22nd September 2024  
Job Ref: SL43016

Researcher in Biostatistics  
Medical Research Council, Cambridge  
Cancer Research  
36 - 44K

Standard disclaimers apply!

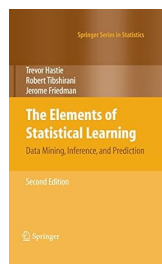
## Foundations of Machine Learning ... draws from several disciplines

- Function approximation
- Linear algebra
  - Low rank approximation, Singular Value Decomposition
- Probability and Statistics
  - Multi-variate Gaussian, Information Theory, Bayes' Formula
- Optimization
- Software Engineering
- Parallel Computing
- Simulation, Experiment Design

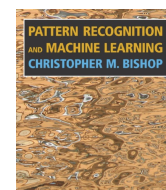
# Machine Learning: Data-driven Modelling

Data	$\{\mathbf{x}_n, \mathbf{t}_n\}_{n=1}^N \quad \{\mathbf{x}_n\}_{n=1}^N$
Function Approximator	$\mathbf{t} = f(\mathbf{x}, \boldsymbol{\theta}) + v$
Parameter Estimation	$E_0 = \sum_{n=1}^N \{\ \mathbf{t}_n - f(\mathbf{x}_n; \boldsymbol{\theta})\ \}^2$
Prediction	$\hat{\mathbf{t}}_{N+1} = f(\mathbf{x}_{N+1}, \hat{\boldsymbol{\theta}})$
Regularization	$E_1 = \sum_{n=1}^N \{\ \mathbf{t}_n - f(\mathbf{x}_n)\ \}^2 + r(\ \boldsymbol{\theta}\ )$
Modelling Uncertainty	$p(\boldsymbol{\theta}   \{\mathbf{x}_n, \mathbf{t}_n\}_{n=1}^N)$
Probabilistic Inference	$\mathbf{E}[g(\boldsymbol{\theta})] = \int g(\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta} = \frac{1}{N_s} \sum_{n=1}^{N_s} g(\boldsymbol{\theta}^{(n)})$
Sequential Estimation	$\boldsymbol{\theta}(n-1 n-1) \longrightarrow \boldsymbol{\theta}(n n-1) \longrightarrow \boldsymbol{\theta}(n n)$ Kalman & Particle Filters; Reinforcement Learning

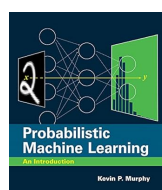
## Good Books To study the fundamentals



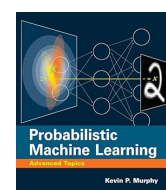
T. Hastie, R. Tibshirani & J. Friedman  
Elements of Statistical Learning



C.M. Bishop  
**PRML** : Pattern Recognition and Machine Learning



K. Murphy  
Probabilistic Machine Learning



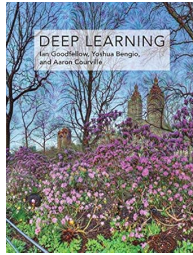
Advanced Topics

*"There is nothing to be learnt from a professor, which is not to be met with in books"*  
- David Hume (1711-1776)

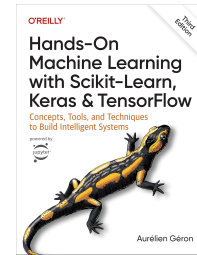
(**Wikipedia**: "Hume had little respect for the professors of his time [...] He did not graduate")

## More books...

For practical aspects of the subject...

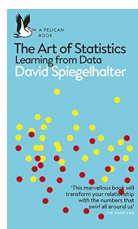


I. Goodfellow *et al.*  
Deep Learning

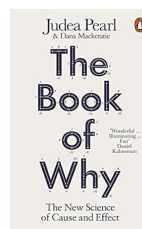


A. Geron  
Hands-On Machine Learning with Scikit-Learn

For some general, societally important, aspects of inference from data...



D. Spiegelhalter  
The Art of Statistics



J. Pearl  
The Book of Why



S.J. Gould  
Mismeasure of Man

Mahesan Niranjan

Machine Learning

June 2025

7 / 10

## Machine Learning: Data-driven Modlling

Data	$\{\mathbf{x}_n, \mathbf{t}_n\}_{n=1}^N$ $\{\mathbf{x}_n\}_{n=1}^N$
Function Approximator	$\mathbf{t} = f(\mathbf{x}, \boldsymbol{\theta}) + v$
Parameter Estimation	$E_0 = \sum_{n=1}^N \{\ \mathbf{t}_n - f(\mathbf{x}_n; \boldsymbol{\theta})\ \}^2$
Prediction	$\hat{\mathbf{t}}_{N+1} = f(\mathbf{x}_{N+1}, \hat{\boldsymbol{\theta}})$
Regularization	$E_1 = \sum_{n=1}^N \{\ \mathbf{t}_n - f(\mathbf{x}_n)\ \}^2 + r(\ \boldsymbol{\theta}\ )$
Modelling Uncertainty	$p(\boldsymbol{\theta}   \{\mathbf{x}_n, \mathbf{t}_n\}_{n=1}^N)$
Probabilistic Inference	$\mathbf{E}[g(\boldsymbol{\theta})] = \int g(\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta} = \frac{1}{N_s} \sum_{n=1}^{N_s} g(\boldsymbol{\theta}^{(n)})$
Sequential Estimation	$\boldsymbol{\theta}(n-1 n-1) \longrightarrow \boldsymbol{\theta}(n n-1) \longrightarrow \boldsymbol{\theta}(n n)$ Kalman & Particle Filters; Reinforcement Learning

Mahesan Niranjan

Machine Learning

June 2025

8 / 10

# Machine Learning

## Some Recent Challenges, Trends

- Function approximator: Shallow versus deep neural architectures
- Several tricks of parameter estimation of large (and deep) models
- Several tricks to induce generalization
- Automatic differentiation
- Problem formulation
  - Supervised & unsupervised versus Self-supervised learning
  - Human-in-the-loop Reinforcement Learning
  - Combining Mechanistic and Data-driven models – e.g. Physics Informed Neural Networks
- Interesting problems in various domains.
- Large amounts of data (sometimes illegally acquired!)
- Large amounts of computing (only a few can afford!)

## Lab One: Multi-variate Gaussian

- Univariate Gaussian

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \frac{(x - m)^2}{\sigma^2} \right\}$$

- Multivariate Gaussian  $\mathbf{x} \in \mathcal{R}^D$

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{D/2}(\det \Sigma)^{1/2}} \exp \left\{ -\frac{1}{2}(\mathbf{x} - \mathbf{m})^t \Sigma^{-1} (\mathbf{x} - \mathbf{m}) \right\}$$

Mean  $\mathbf{m}$  is a vector

Covariance,  $\mathbf{C}$ , matrix: symmetric, positive definite!

$$\mathbf{x} \sim \mathcal{N}(\mathbf{m}, \Sigma), \mathbf{y} = \mathbf{A}\mathbf{x} \implies \mathbf{y} \sim \mathcal{N}(\mathbf{A}\mathbf{m}, \mathbf{A}\Sigma\mathbf{A}^T)$$

- Sampling from a given density
- Projections of data onto a subspace