Recommended resources: ML/DL

This document brings together a selection of books, courses, and online resources that members of our ML and research teams have found genuinely useful. It is not meant as a checklist or a body of knowledge you are expected to master. Instead, it offers a glimpse into how we think about machine learning and deep learning, and where someone preparing for an interview with our ML/DL team might want to start.

At Qminers, we apply ML and DL to complex, noisy financial data, where signal is subtle and assumptions matter. We therefore value solid foundations in probability, statistics, and modelling, as well as a good understanding of how modern ML and deep learning methods actually work in practice. The recommended materials reflect this balance between theory and hands-on experience.

If you are interested in machine learning beyond standard benchmarks, and in models that have to perform under real-world constraints, you should find these resources relevant.

What does Ondřej recommend?

Ondřej Skácel
Quantitative Analyst

Software Engineering

1 Fluent Python Book (or alternative)

II Data Structures
2. An Array of Sequences
3. Dictionaries and Sets

III Functions as Objects
5. First-Class Functions
7. Function Decorators and Closures

IV Object-Oriented Idioms
8. Object References, Mutability, and Recycling
9. A Pythonic Object
10. Sequence Hacking, Hashing, and Slicing

V Control Flow
14. Iterables, Iterators, and Generators
15. Context Managers and else Blocks
16. Coroutines*
18. Concurrency with asyncio*

2 **Basics of NumPy, Matplotlib, Pandas*, PyTorch and Pytest.**

3 **Theoretical computer science basics (hopefully you already know from university)**

- Algorithmic complexity (big-O notation, amortized vs average vs worst case complexity)
- Common algorithms and data structures
- A few days of LeetCode or similar algorithmic problems/competitions

4 **Basics of computer architectures***

5 **Unix/Linux fundamentals***

Materials marked by asterisk* are required only for the ML engineer position.

Quant Researcher Brainteasers

Get used to problem solving.

https://brainstellar.com/
https://openquant.co/questions
https://www.janestreet.com/puzzles/archive/index.html

Machine Learning

1 **Probability, Statistics and Linear Algebra**

- Working with probability distributions (CDF, PDF, E, Var, addition and multiplication)
- Conditional probability, Total probability
- Basics statistics, significance, Confidence intervals, Central limit theorem
- E.g. Statistical Inference Book
- Matrices, eigenvalues and eigenvectors
- Differentiation and integration

2 **Elements of Statistical Learning Book**

- 3. Linear Methods for Regression
- 10. Boosting and Additive Trees
- 16. Ensemble Learning

3 **Alternatively, parts of Understanding Machine Learning Book**

4 **Modelling basics** (from Caltech's Machine Learning Course by Professor Yaser Abu-Mostafa)

- Bias-Variance Tradeoff
- Overfitting

5 Neural Network Basics

6 **Deep Learning Background and Architectures**

E.g. https://www.bishopbook.com/,
(optionally https://www.deeplearningbook.org/)

- Gradient, popular loss functions (MSE, Cross Entropy, ...)
- Training algorithms (SGD, Momentum, RMSProp, Adam)
- Regularization and normalization techniques (for model, data, and training algorithm)
- Convolutional nets, LSTM, VAE, GAN, SSM
- Transformers

7 Google Research Deep Neural Network Tuning Handbook

8 Karpathy's Recipe for Training Neural Networks