Recommended reading

These materials are not required to pass the exam. But they are worth looking into if you're struggling with one of the subjects, or if you just want to learn more.

Introduction

- **A few useful things to know about machine learning.** (https://homes.cs.washington.edu/~pedrod/papers/cacm12.pdf) Pedro Domingos.
- **Machine Learning crash course.** (https://developers.google.com/machine-learning/crash-course/) From Google.
    - The **glossary** (https://developers.google.com/machine-learning/crash-course/glossary) may be particularly useful if you get stuck on an unfamiliar concept.

Linear Models 1

- A good way to **visualize squared error loss.** (https://twitter.com/chrisalbon/status/969636660462813184)
- **Simple introduction** (https://towardsdatascience.com/linear-algebra-for-deep-learning-f21d7e7d7f23) to the basics of linear algebra.

Methodology 1

- Derived features: **https://developers.google.com/machine-learning/crash-course/feature-crosses/video-lecture** (https://developers.google.com/machine-learning/crash-course/feature-crosses/video-lecture)

Methodology 2

- **Steven Strogatz explains eigenvectors (short video).** (https://www.youtube.com/watch?v=AXk12z-NGPI&t=40s)
- Another **explanation of the singular value decomposition** (http://andrew.gibiansky.com/blog/mathematics/cool-linear-algebra-singular-value-decomposition/).
- Andrew Ng's **introduction to machine learning** (https://www.youtube.com/watch?v=UzxYlbK2c7E) at Stanford (video lecture).

Probabilistic models 1

- **A visual, interactive introduction** (http://students.brown.edu/seeing-theory/) to probability theory.
    - There's also a **draft of a book** (http://students.brown.edu/seeing-theory/doc/seeing-theory.pdf).
- **A visual explanation of Bayes' rule** (https://www.youtube.com/watch?v=BrK7X_XlGB8), and how to use it in everyday thinking (short video).

Deep Learning 1

- A lecture on neural networks by Patrick Winston (video). (https://www.youtube.com/watch?v=uXt8qF2Zzfo)
- Introduction to backpropagation: reader (http://www.cs.toronto.edu/~rgrosse/courses/csc321_2018/readings/L06%20Backpropagation.pdf) , slides (http://www.cs.toronto.edu/~rgrosse/courses/csc321_2018/slides/lec06.pdf) . Part of this course (http://www.cs.toronto.edu/~rgrosse/courses/csc321_2018/) .
- CS231n (http://cs231n.stanford.edu/) , a very famous deep learning course focusing on computer vision. Videos on youtube (https://www.youtube.com/watch?v=vT1JzLTH4G4&list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv) .
- All the matrix calculus (https://arxiv.org/abs/1802.01528) needed for Deep Learning. In the lectures, we stuck to partial derivatives, which is enough to explain the basic concepts. Matrix calculus makes this more practical, by allowing you to take all the partial derivatives of a matrix or vector in one go.
    - A more gentle introduction (http://cs231n.stanford.edu/vecDerivs.pdf) (of just 7 pages) from the C231n notes.
- What do neural net loss functions look like (video)? (http://www.ipam.ucla.edu/abstract/?tid=14548&pcode=DLT2018) An attempt to visualize the very high-dimensional loss landscapes of deep learning models in two dimensions.
- https://blog.slavv.com/37-reasons-why-your-neural-network-is-not-working-4020854bd607 (https://blog.slavv.com/37-reasons-why-your-neural-network-is-not-working-4020854bd607)
-

Linear Models 2

- How to express the width of the margin (https://www.youtube.com/watch?v=eUfvyUEGMD8) of an SVM (short video).

Probabilistic Models 2

Deep Learning 2

Tree Models

- Trevor Hastie - Gradient Boosting Machine Learning (video) (https://www.youtube.com/watch?v=wPqtzj5VZus)

Models for Sequential Data

- Markov models: https://blog.codinghorror.com/markov-and-you/ (https://blog.codinghorror.com/markov-and-you/)
- http://colah.github.io/posts/2015-08-Understanding-LSTMs/ (http://colah.github.io/posts/2015-08-Understanding-LSTMs/)
- http://karpathy.github.io/2015/05/21/rnn-effectiveness/ (http://karpathy.github.io/2015/05/21/rnn-effectiveness/)

- Skip-gram Word2Vec: http://mccormickml.com/2016/04/19/word2vec-tutorial-the-skip-gram-model/ (http://mccormickml.com/2016/04/19/word2vec-tutorial-the-skip-gram-model/)
-

Matrix Models

- Matrix factorization techniques for recommender systems (article). (https://canvas.vu.nl/courses/29266/files/502943/download?wrap=1) (https://canvas.vu.nl/courses/29266/files/502943/download?wrap=1)
- http://alexhwilliams.info/itsneuronalblog/2018/02/26/crossval/?mlreview (http://alexhwilliams.info/itsneuronalblog/2018/02/26/crossval/?mlreview)
- http://alexhwilliams.info/itsneuronalblog/2016/03/27/pca/#tldr (http://alexhwilliams.info/itsneuronalblog/2016/03/27/pca/#tldr)

Reinforcement Learning

Review

- Differentiable programming. (https://www.edge.org/response-detail/26794)