Data Architecture Notes

Contents

I Mathematics for Machine Learning **9**
0.1 Linear Algebra . . . 11
0.2 Calculus . . . 11
0.3 Probability and Statistics . . . 11

II Mathematics for Machine Learning **13**
0.4 Mathematics for Machine Learning: Linear Algebra . . . 15
0.5 Mathematics for Machine Learning: Multivariate Calculus . . . 15
0.6 Mathematics for Machine Learninig: PCA . . . 15

III Linear Algebra for Machine Learning and Data Science **17**
0.7 Systems of Linear equations . . . 19
0.8 Solving systems of linear equations . . . 19
0.9 Vectors and Linear Transformations . . . 19
0.10 Determinants and Eigenvectors . . . 19

IV Machine Learning Specialisation **21**
0.11 Supervised Machine Learning: Regression and Classification . . . 24
0.11.1 Overview of machine learning . . . 24
0.11.2 Supervised vs Unsupervised Machine Learning . . . 24
0.11.3 Regression model . . . 24
0.11.4 Train the model with gradient descent . . . 24
0.11.5 Multiple Linear Regression . . . 24
0.11.6 Gradient descent in practice . . . 24
0.11.7 Classification with logistic regression . . . 24
0.11.8 Cost function with logistic regression . . . 24
0.11.9 Gradient descent for logistic regression . . . 24
0.11.10 The problem overfitting . . . 24
0.12 Advanced Learning Algorithms . . . 24
0.12.1 Neural networks intuition . . . 24
0.12.2 Neural network model . . . 24
0.12.3 TensorFlow implementation . . . 24
0.12.4 Neural network implementation in Python . . . 24
0.12.5 Speculations on artificial general intelligence (AGI) . . . 24
0.12.6 Vectorisation (Optional) . . . 24
0.12.7 Neural Network Training . . . 24
0.12.8 Activation Functions . . . 24
0.12.9 Multiclass Classification . . . 24
0.12.10 Additional Neural Network Concepts . . . 24

0.12.11 Back Propagation . . . 24
0.12.12 Advice for applying machine learning . . . 24
0.12.13 Advice for applying machine learninig . . . 24
0.12.14 Bias and variance . . . 24
0.12.15 Machine learning development process . . . 24
0.12.16 Skewed datasets . . . 24
0.12.17 Decision Trees . . . 24
0.12.18 Decision Tree Learning . . . 24
0.12.19 Tree ensembles . . . 24
0.13 Unsupervised Learning, Recommenders, Reinforcement Learning . . . 24
0.13.1 Clustering . . . 24
0.13.2 Anomaly Detection . . . 24
0.13.3 Collaborative filtering . . . 24
0.13.4 Recommender systems implementation details . . . 24
0.13.5 Content-based filtering . . . 24
0.13.6 Principal Component Analysis . . . 24
0.13.7 Reinforcement Learning introduction . . . 24
0.13.8 State-action value function . . . 24
0.13.9 Continuous state spaces . . . 24

V Deep learning specialisation 25
0.14 Neural Networks and Deep Learning . . . 27
0.14.1 What is deep learning? . . . 27
0.14.2 Logistic Regression as a Neural Network . . . 27
0.14.3 Python and Vectorisation . . . 27
0.14.4 Shallow Neural Network . . . 27
0.14.5 Deep Neural Network . . . 27
0.15 Improving Deep Neural Networks: Hyperparameter Tuning, Regularization and Optimization . . . 27
0.15.1 Structuring Machine Learning Projects . . . 27
0.15.2 Setting up a Machine Learning Application . . . 27
0.15.3 Regularising a Neural Network . . . 27
0.15.4 Setting up a Optimisation Problem . . . 27
0.15.5 Optimisation Algorithms . . . 27
0.15.6 Hyperparameter Tuning . . . 27
0.15.7 Batch Normalisation . . . 27
0.15.8 Multi-Class Classification . . . 27
0.15.9 Introduction to Programming Frameworks . . . 27
0.16 Convolution Neural Networks . . . 27
0.16.1 Convolutional Neural Networks . . . 27
0.16.2 Deep Convolutional Models: Case Studies . . . 27
0.16.3 Detection Algorithms . . . 27
0.16.4 Face Recognition . . . 27
0.16.5 Neural Style Transfer . . . 27
0.17 Sequence Models . . . 27
0.17.1 Recurrent Neural Networks . . . 27
0.17.2 Introduction to Word Embeddings . . . 27
0.17.3 Learning Word Embeddings: Word2vec and GloVe . . . 27
0.17.4 Applications using Word Embeddings . . . 27
0.17.5 Sequence Models and Attention Mechanism . . . 27

0.17.6 Transformer Network 27

VI Generative Adverserial Networks (GANs) Specialisation 29
0.18 Build Basic Generative Adversarial Networks 31
0.18.1 Intro To GANs 31
0.18.2 Deep Convolutional GANs 31
0.18.3 Wasserstein GANs with Gradient Penalty 31
0.18.4 Conditional GAN and Controllable Generation 31
0.19 Build Better Generative Adversarial Networks 31
0.19.1 Evaluation of GANs 31
0.19.2 GAN Disadvantages and Bias 31
0.19.3 StyleGAN and Advancements 31
0.20 Apply Generative Adverserial Networks 31
0.20.1 GANs for Data Augmentation and Privacy 31
0.20.2 Image-to-Image Translation with Pix2Pix 31
0.20.3 Unpaired Translation with CycleGAN 31

VII Natural Language Processing Specialisation 33
0.21 Natural Language Processing with Classification and Vector Spaces 35
0.21.1 Sentiment Analysis with Logistic Regression 35
0.21.2 Sentiment Analysis with Naive Bayes 35
0.21.3 Vector Space Models 35
0.21.4 Machine Translation and Document Search 35
0.22 Natural Language Processing with Probabilitsic Models 35
0.22.1 Autocorrect 35
0.22.2 Part of Speech Tagging and Hidden Markov Models 35
0.22.3 Autocomplete and Language Models 35
0.22.4 Word embeddings with neural networks 35
0.23 Natural Language Processing with Sequence Models 35
0.23.1 Recurrent Neural Network for Language Modeling 35
0.23.2 LSTMs and Named Entity Recognition 35
0.23.3 Siamese Networks 35
0.24 Natural Language Processing with Attention Models 35
0.24.1 Neural Machine Translation 35
0.24.2 Text Summarisation 35
0.24.3 Question Answering 35

VIII AI for Scientific Research Specialisation 37
0.25 Introduction to Data Science and scikit-learn in Python 39
0.25.1 Introduction to Python Programming for Hypothesis Testing 39
0.25.2 Creating a Hypothesis: Numpy, Pandas and Scikit-Learn 39
0.25.3 Scikit-Learn: ML for Hypothesis Testing 39
0.25.4 Using Classification to Predict the Presence of Heart Disease 39
0.26 Machine Learning Models in Science 39
0.26.1 Before the AI: Preparing and Preprocessing Data 39
0.26.2 Foundational AI Algorithms: K-Means and SVM 39
0.26.3 Advanved AI: Neural Netoworks and Decision Trees 39
0.27 Neural Networks and Random Forests 39
0.27.1 Introduction to Neural Networks 39

0.27.2 Deep Dive into Neural Networks 39
0.27.3 Exploring Random Forests 39
0.28 Project 39

List of Figures

Python: https://scikit-learn.org/stable/
Scala https://spark.apache.org/mllib/

Part I

Mathematics for Machine Learning

0.1 Linear Algebra

0.2 Calculus

0.3 Probability and Statistics

Part II

Mathematics for Machine Learning

0.4 Mathematics for Machine Learning: Linear Algebra

0.5 Mathematics for Machine Learning: Multivariate Calculus

0.6 Mathematics for Machine Learninig: PCA

Part III

Linear Algebra for Machine Learning and Data Science

0.7 Systems of Linear equations

0.8 Solving systems of linear equations

0.9 Vectors and Linear Transformations

ó

0.10 Determinants and Eigenvectors

Part IV

Machine Learning Specialisation

0.11 Supervised Machine Learning: Regression and Classification

0.11.1 Overview of machine learning

0.11.2 Supervised vs Unsupervised Machine Learning

0.11.3 Regression model

0.11.4 Train the model with gradient descent

0.11.5 Multiple Linear Regression

0.11.6 Gradient descent in practice

0.11.7 Classification with logistic regression

0.11.8 Cost function with logistic regression

0.11.9 Gradient descent for logistic regression

0.11.10 The problem overfitting

0.12 Advanced Learning Algorithms

0.12.1 Neural networks intuition

0.12.2 Neural network model

0.12.3 TensorFlow implementation

0.12.4 Neural network implementation in Python

0.12.5 Speculations on artificial general intelligence (AGI)

0.12.6 Vectorisation (Optional)

0.12.7 Neural Network Training

0.12.8 Activation Functions

0.12.9 Multiclass Classification

0.12.10 Additional Neural Network Concepts

0.12.11 Back Propagation

0.12.12 Advice for applying machine learning

0.12.13 Advice for applying machine learninig

0.12.14 Bias and variance

0.12.15 Machine learning development process

0.12.16 Skewed datasets

0.12.17 Decision Trees

0.12.18 Decision Tree Learning

0.12.19 Tree ensembles

0.13 Unsupervised Learning, Recommenders, Reinforcement Learning

Part V

Deep learning specialisation

0.14 Neural Networks and Deep Learning

0.14.1 What is deep learning?

0.14.2 Logistic Regression as a Neural Network

0.14.3 Python and Vectorisation

0.14.4 Shallow Neural Network

0.14.5 Deep Neural Network

0.15 Improving Deep Neural Networks: Hyperparameter Tuning, Regularization and Optimization

0.15.1 Structuring Machine Learning Projects

0.15.2 Setting up a Machine Learning Application

0.15.3 Regularising a Neural Network

0.15.4 Setting up a Optimisation Problem

0.15.5 Optimisation Algorithms

0.15.6 Hyperparameter Tuning

0.15.7 Batch Normalisation

0.15.8 Multi-Class Classification

0.15.9 Introduction to Programming Frameworks

0.16 Convolution Neural Networks

0.16.1 Convolutional Neural Networks

0.16.2 Deep Convolutional Models: Case Studies

0.16.3 Detection Algorithms

0.16.4 Face Recognition

0.16.5 Neural Style Transfer

0.17 Sequence Models

0.17.1 Recurrent Neural Networks

0.17.2 Introduction to Word Embeddings

0.17.3 Learning Word Embeddings: Word2vec and GloVe

0.17.4 Applications using Word Embeddings

0.17.5 Sequence Models and Attention Mechanism

0.17.6 Transformer Network

Part VI

Generative Adverserial Networks (GANs) Specialisation

0.18 Build Basic Generative Adversarial Networks

0.18.1 Intro To GANs

0.18.2 Deep Convolutional GANs

0.18.3 Wasserstein GANs with Gradient Penalty

0.18.4 Conditional GAN and Controllable Generation

0.19 Build Better Generative Adversarial Networks

0.19.1 Evaluation of GANs

0.19.2 GAN Disadvantages and Bias

0.19.3 StyleGAN and Advancements

0.20 Apply Generative Adverserial Networks

0.20.1 GANs for Data Augmentation and Privacy

0.20.2 Image-to-Image Translation with Pix2Pix

0.20.3 Unpaired Translation with CycleGAN

Part VII

Natural Language Processing Specialisation

0.21 Natural Language Processing with Classification and Vector Spaces

0.21.1 Sentiment Analysis with Logistic Regression

0.21.2 Sentiment Analysis with Naive Bayes

0.21.3 Vector Space Models

0.21.4 Machine Translation and Document Search

0.22 Natural Language Processing with Probabilitsic Models

0.22.1 Autocorrect

0.22.2 Part of Speech Tagging and Hidden Markov Models

0.22.3 Autocomplete and Language Models

0.22.4 Word embeddings with neural networks

0.23 Natural Language Processing with Sequence Models

0.23.1 Recurrent Neural Network for Language Modeling

0.23.2 LSTMs and Named Entity Recognition

0.23.3 Siamese Networks

0.24 Natural Language Processing with Attention Models

0.24.1 Neural Machine Translation

0.24.2 Text Summarisation

0.24.3 Question Answering

Part VIII

AI for Scientific Research Specialisation

0.25 Introduction to Data Science and scikit-learn in Python

0.25.1 Introduction to Python Programming for Hypothesis Testing

0.25.2 Creating a Hypothesis: Numpy, Pandas and Scikit-Learn

0.25.3 Scikit-Learn: ML for Hypothesis Testing

0.25.4 Using Classification to Predict the Presence of Heart Disease

0.26 Machine Learning Models in Science

0.26.1 Before the AI: Preparing and Preprocessing Data

0.26.2 Foundational AI Algorithms: K-Means and SVM

0.26.3 Advanved AI: Neural Netoworks and Decision Trees

0.27 Neural Networks and Random Forests

0.27.1 Introduction to Neural Networks

0.27.2 Deep Dive into Neural Networks

0.27.3 Exploring Random Forests

0.28 Project