Workshop on High Performance Machine Learning

23-25 Aug 2019

Learning material:

Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow, 2nd Edition

<https://www.oreilly.com/library/view/hands-on-machine-learning/9781492032632/>

<https://github.com/ageron/handson-ml2>

Day 1: ML Fundamentals (scikit-learn)

Registration at HIT Shenzhen

AM - Chapter 2 [WANG Qiang]:

pipeline of machine learning project, with a focus on data preprocessing, e.g., splitting the data set, missing data, how to handle numerical data and categorical data (one-hot coding)

LUNCH

PM-I - Chapter 3 [HE Xin]:

binary classification, with a focus on performance evaluation metrics, e.g., cross-validation, confusion matrix, precision, recall (sensitivity), F1-score, precision-recall tradeoff, ROC (Receiver Operating Characteristic) Curve and AUC (area under the curve)

multiclass classification: one-versus-all (OvA) and one-versus-one (OvO)

PM-II - Chapter 4 and Chapter 7 [TANG Zhenheng]

GD/SGD/minibatch SGD/momentum/convergence / Learning Curves

DINNER

Day 2: Deep Learning Fundamentals (Keras):

BREAKFAST

AM-I: Chapter 10 [WANG Yuxin]

sequential model, functional API, subclassing API, callbacks, TensorBoard

AM-II: Chapter 12 [WANG Qiang]

Custom models and training

LUNCH

PM-I: Chapter 14 - CNN [HE Xin]:

PM-II: Chapter 15-16 - RNN & LSTM [WANG Yuxin]

DINNER

Day 3: Advanced Deep Learning

BREAKFAST

AM-I: Chapter 11 [SHI Shaohuai]

SGD, Batch Normalization, Faster Optimizers, Regularization

AM-II: Chapter 19 [SHI Shaohuai]

Distributed Training (Data Parallelism)

LUNCH

Go back to HK