Introduction to Data Science

Description

clustering;

frequent pattern mining;

linear classification and regression model: model parameters and hyper-parameters, validation, overfitting-underfitting and the bias-variance trade-off;

introduction to prediction techniques (as black-box functions);

data quality and pre-processing: noise, missing values, data transformation, normalization;

the CRISP-DM process;

recommendation techniques;

Literature

- Peter Flach (2012). Machine Learning: The Art and Science of Algorithms that Make Sense of Data. Cambridge University Press.

- Jiawei Han, Micheline Kamber, Jian Pei (2011). Data Mining: Concepts and Techniques. Morgan Kaufmann.

- Pang-Ning Tan, Michael Steinbach, Vipin Kumar (2005). Introduction to Data Mining. Addison Wesley