

Lecture schedule

1. Introduction
(Tan 1.1-1.4)

Data: Feature extraction and visualization

2. Data and feature extraction
(Tan 2.1-2.3 + B1 (+ A))
3. Measures of similarity and summary statistics
(Tan 2.4 + 3.1-3.2 + C1-C2)
4. Data visualization
(Tan 3.3)

Supervised learning: Classification and regression

5. Decision trees and linear regression
(Tan 4.1-4.3 + D)
6. Overfitting and performance evaluation
(Tan 4.4-4.6)
7. Nearest neighbor, naive Bayes, and artificial neural networks
(Tan 5.2-5.4)

8. Ensemble methods and multi class classifiers
(Tan 5.6-5.8)

Unsupervised learning: Clustering and density est.

9. K-means and hierarchical clustering
(Tan 8.1-8.3+8.5.7)

10. Mixture models and association mining (Tan 9.2.2 + 6.1-6.3)

11. Density estimation and anomaly detection
(Tan 10.1-10.4)

Machine learning and data modelling in practice

12. Putting it all together: Summary and overview
13. Mini project