

6. Learning to Classify Text

6.1 Supervised Classification

- 1.1 Gender Identification
- 1.2 Choosing The Right Features
- 1.3 Document Classification
- 1.4 Part-of-Speech Tagging
- 1.5 Exploiting Context
- 1.6 Sequence Classification
- 1.7 Other Methods for Sequence Classification

6.2 Further Examples of Supervised Classification

- 2.1 Sentence Segmentation
- 2.2 Identifying Dialogue Act Types
- 2.3 Recognizing Textual Entailment
- 2.4 Scaling Up to Large Datasets

6.3 Evaluation

- 3.1 The Test Set
- 3.2 Accuracy
- 3.3 Precision and Recall
- 3.4 Confusion Matrices
- 3.5 Cross-Validation

6.4 Decision Trees

- 4.1 Entropy and Information Gain

6.5 Naive Bayes Classifiers

- 5.1 Underlying Probabilistic Model
- 5.2 Zero Counts and Smoothing
- 5.3 Non-Binary Features
- 5.4 The Naivete of Independence
- 5.5 The Cause of Double-Counting

6.6 Maximum Entropy Classifiers

- 6.1 The Maximum Entropy Model
- 6.2 Maximizing Entropy
- 6.3 Generative vs Conditional Classifiers

6.7 Modeling Linguistic Patterns

- 7.1 What do models tell us?

6.8 Summary

6.9 Further Reading

6.10 Exercises