

Table of Contents

第一章 文本数据分析	1
1.1 文本数据基础	2
1.2 文本预处理	2
1.3 特征提取	2
1.4 可视化	2
1.5 实践案例	2

第一章 文本数据分析

1.1 文本数据基础

1.1.1 什么是文本数据

1.1.2 文本数据的常见来源

1.2 文本预处理

1.2.1 分词

1.2.2 去除停用词

1.2.3 大小写转换

1.2.4 标点符号处理

1.2.5 词干提取与词形还原

1.3 特征提取

1.3.1 词袋模型 (Bag of Words)

1.3.2 TF-IDF

1.3.3 词嵌入 (Word Embedding)

1.3.4 文本分析方法

1.3.5 词频统计