

综述：

Text feature extraction based on deep learning: a review 2017

(引用 36 次)

基于朴素贝叶斯的文本分类研究综述_贺鸣 2016

文本分类研究综述 2019 汪岚 刘柏嵩

半监督文本分类综述 2011

中文文本分类方法综述 2019

文本相似度计算方法研究综述 2019

自然语言处理发展及应用综述_赵京胜 2019

最早自动分类： On relevance, probabilistic indexing and
information retrieval (1960)

经典博士论文： Representation and Learning in Information
Retrieval (1992)

机器学习系列：

特征选择比较： A comparative study on feature selection in text
categorization (1997)

A comparison of event models for naive bayes text

classification (1998)

Transductive inference for text classification using support vector machines (1999)

Using maximum entropy for text classification (1999)

Text classification from labeled and unlabeled documents using EM (2000)

贝叶斯反垃圾邮件 An evaluation of naive bayesian anti-spam filtering (2000)

Support vector machine active learning with applications to text classification (2001)

Text Classification using String Kernels 2002 (引用次数 1627)

Text classification using string kernels (2002)

特征选择: An extensive empirical study of feature selection metrics for text classification (2003)

基于图的: Graph-based text classification: learn from your neighbors (2006)

短文本分类: Short text classification in twitter to improve information filtering (2010)

单词共现: Word co-occurrence features for text classification (2011)

Baseline Needs More Love: On Simple Word-Embedding-Based Models and Associated Pooling Mechanisms

Using TF-IDF to Determine Word Relevance in Document Queries

Chinese Segmentation and New Word Detection using Conditional

Random Fields: 2004 (引用 496 次)

A novel probabilistic feature selection method for text classification (2012)

Learning to Classify Text from Labeled and Unlabeled Documents (1998)

A Tutorial on Automated Text Categorisation (1999)

A comparative study of TF*IDF, LSI and multi-words for text classification 2011 (引用 399 次)

Feature Selection and Feature Extraction for Text Categorization 1992 (引用次数 621)

深度学习系列:

Text feature extraction based on deep learning: a review 2017 (引用 36 次)

Adversarial Multi-task Learning for Text Classification (2017)

Deep Learning for Chinese Word Segmentation and POS Tagging 2013 (引用 226)

Relation classification via convolutional deep neural network (2014)

A method for chinese text classification based on apparent semantics and latent aspects (2015)

基于Word2vec和改进型TF-IDF的卷积神经网络文本分类模型(2019)

文本分类TF-IDF算法的改进研究(2019)

TextCNN : Convolutional Neural Networks for Sentence Classification(EMNLP 2014)

GRU: Document Modeling with Gated Recurrent Neural Network for Sentiment Classification (EMNLP 2015)

RCNN : Recurrent Convolutional Neural Networks for Text Classification(AAAI 2015)

char-CNN, 没有使用词向量: Character-level Convolutional Networks for Text Classification (NIPS 2015)

Recurrent Neural Network for Text Classification with Multi-Task Learning (IJCAI 2016)

HAN : Hierarchical Attention Networks for Document Classification (NAACL 2016)

fastText: Bag of Tricks for Efficient Text Classification (EACL 2017)

DPCNN: Deep Pyramid Convolutional Neural Networks for Text Categorization (ACL 2017)

DeepMoji : Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion

and sarcasm (EMNLP 2017)

标签量非常多 Deep learning for extreme multi-label text classification (2017)

胶囊网络: Investigating Capsule Networks with Dynamic Routing for Text Classification (EMNLP 2018)

RNN-Capsule: Sentiment Analysis by Capsules (WWW 2018)

UMFiT 迁移学习: Universal Language Model Fine-tuning for Text Classification (2018)

GCN: Graph Convolutional Networks for Text Classification (AAAI 2019)

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding (NAACL 2019)