

基于字符级联合网络特征融合的中文文本情感分析

王丽亚, 刘昌辉, 蔡敦波, 赵彤洲, 王 梦

(武汉工程大学 计算机科学与工程学院, 湖北 武汉 430205)

摘要: 针对传统卷积神经网络(CNN)同层神经元之间信息不能互传,无法充分利用同一层次上的特征信息,以及无法提取长距离上下文相关特征的问题.该文针对中文文本,提出字符级联合网络特征融合模型进行情感分析,在字符级的基础上采用 BiGRU 和 CNN-BiGRU 并行的联合网络提取特征,利用 CNN 的强学习能力提取深层次特征,再利用双向门限循环神经网络(BiGRU)进行深度学习,加强模型对特征的学习能力.另一方面,利用 BiGRU 提取上下文相关的特征,丰富特征信息.最后在单方面上引入注意力机制进行特征权重分配,降低噪声干扰.在数据集上进行多组对比实验,该方法取得 92.36%的 F1 值,结果表明本文提出的模型能有效的提高文本分类的准确率.

关键词: 卷积神经网络; BiGRU; 注意力机制; 中文文本情感分析

Chinese text sentiment analysis based on joint network and attention model

WANG Li-ya, LIU Chang-hui, CAI Dun-bo, ZHAO Tong-zhou, WANG Meng

(College of Computer Science and Engineering, Wuhan Institute of Technology, Wuhan 430205, China)

Abstract: For the traditional convolutional neural network (CNN), the information between the same layer of neurons cannot be transmitted to each other, and the feature information at the same level cannot be fully utilized, and the problem of long-distance context-related features cannot be extracted. In this paper, based on the Chinese text, this paper proposes a model of character-level joint network feature fusion for sentiment analysis. Based on the character level, BiGRU and CNN-BiGRU parallel joint network are used to extract features, and CNN's strong learning ability is used to extract deep features and reuse. The two-way threshold cyclic neural network (BiGRU) performs deep learning and enhances the model's ability to learn features. On the other hand, BiGRU is used to extract context-related features to enrich feature information. Finally, the attention mechanism is introduced unilaterally to perform feature weight distribution to reduce noise interference. Multi-group comparison experiments were carried out on the dataset. The method obtained 92.36% F1 value. The results show that the proposed model can effectively improve the accuracy of text classification.

Key words: convolutional neural network; BiGRU; attention; Chinese text sentiment analysis

作者简介:

王丽亚 女, (1994-), 硕士研究生.研究方向为文本挖掘.

E-mail:1578938844@qq.com.

刘昌辉 男, (1965-), 博士, 教授, 硕士生导师.研究方向为机器人控制技术、软件系统与智能计算、信息处理.

蔡敦波 男, (1981-), 博士, 教授, 硕士生导师.研究方向为人工智能的智能规划、自动推理、约束优化、文本挖掘.

赵彤洲 女, (1970-), 硕士, 副教授, 硕士生导师.研究方向为机器视觉、数据分析、文本挖掘.

王 梦 女, (1994-), 硕士研究生.研究方向为数据挖掘.