摘要

隨著互聯網的不斷普及和高速發展，各式各樣的網路平臺已經走入人們的生活。網路用戶在社交類平臺分享個人生活的態度，在銷售類平臺留下對商品服務的評論，在資訊類平臺留下對社會事件的看法。這些帶有個人情感資訊的網路文本無疑是極具價值的。於企業而言，分析消費者的評論可以幫助其瞭解產品的不足，從而改善產品性能；於政府而言，通過總結線民對事件的看法，以引導事件的發展方向。因此，情感分析日益受到社會各界的廣泛關注。

傳統的基於情感詞典和傳統機器學習的情感分析方法在語義資訊特徵提取方面表現不夠理想，難以應對網路上抽象、簡化的文字表達，更加生活化的語言和不斷湧現的網路新詞，本文基於知識增強語義表示（Enhanced Representation through Knowledge Integration，ERNIE）和雙通道模型設計並實現了微博文本情感分析系統，本文主要工作及結論如下：

（1）通過編寫爬蟲程式收集網上用戶的評論文本，經過人工篩選、文本清洗後，使用 jieba 分詞工具完成分詞工作。

（2）提出了基於知識增強語義表示預訓練模型和神經網路微調的微博文本情感分析方案。在 Windows 平臺下進行仿真實驗，通過實驗驗證，對比其他已公開發表的優秀模型的性能，從而確定系統的可行性和優越性。

（3）設計並實現一個新浪微博文本情感分析系統，對文本進行預處理之後輸入模型進行情感分析，隨後將所獲取的批量文本分析結果可視化。

在 Windows 和 Ubuntu 雙系統下搭配進行仿真實驗，在 Windows 系統訓練好的模型參數保存；在 Ubuntu 系統下搭建前端和後端的 API 架構，並將 Windows 保存下來的模型參數接入 Ubuntu 中後端 API 以便調用。

關鍵字：微博文本，情感分析，神經網路，ERNIE 模型

Abstract

With the continuous popularization and rapid development of the Internet, various types of online platforms have entered people's lives. Internet users share their personal lives on social platforms, leave comments on goods and services on sales platforms, and express their opinions on social events on information platforms. These online texts with personal emotional information are undoubtedly of great value. For enterprises, analyzing consumer comments can help them understand product shortcomings and improve product performance. For governments, summarizing netizens' opinions on events can guide the direction of event development. Therefore, sentiment analysis has increasingly received widespread attention from all sectors of society.

Traditional sentiment analysis methods based on sentiment dictionaries and traditional machine learning have not performed well in semantic information feature extraction, and are difficult to deal with abstract and simplified text expressions, more life-like language, and constantly emerging internet neologisms. This paper proposes a Weibo text sentiment analysis system based on Enhanced Representation through Knowledge Integration (ERNIE) and a dual-channel model. The main work and conclusions of this paper are as follows:

（1）By writing web crawlers, online user comments were collected, and after manual selection and text cleaning, the segmentation work was completed using the jieba segmentation tool.

（2）A Weibo text sentiment analysis solution based on knowledge-enhanced semantic representation pre-training models and neural network fine-tuning was proposed. Simulation experiments were conducted on the Windows platform, and through experimental verification and performance comparison with other excellent models already published, the feasibility and superiority of the system were determined.

（3） A Sina Weibo text sentiment analysis system was designed and implemented. After preprocessing the text, it was input into the model for sentiment

analysis, and the batch text analysis results obtained for the same topic were then visualized. Simulation experiments were conducted on a Windows and Ubuntu dual system, with the model parameters trained on the Windows system saved. The API architecture for the frontend and backend was built on the Ubuntu system, and the saved model parameters from Windows were integrated into the backend API in Ubuntu for easy access.

Keywords: Weibo text, sentiment analysis, neural network, ERNIE model.