applied sciences

MDPI

*Editorial*

# Natural Language Processing: Recent Development and Applications

**Kuei-Hu Chang**

Department of Management Sciences, R.O.C. Military Academy, Kaohsiung 830, Taiwan; evenken2002@yahoo.com.tw

Natural Language Processing (NLP) can be categorized into the subfields of artificial intelligence (AI) and linguistics. NLP mainly explores how to make machines understand, process, reproduce, and generate natural languages. In today's Internet of Things (IoT) era, digital devices and cyberspace help people solve their daily needs and problems. They are not only personal assistants in our daily lives but also play an important role in sharing humans' burdens. In other words, the developments of NLP determine the progress of machines' developments. In terms of NLP applications, they embrace corpus analysis, information retrieval (IR), machine translation (MT), part-of-speech (POS) tagging, semantic analysis, text categorization, and so on. Those applications are closely related to the processing efficiency of big textual information and directly affect the advances of information technology (IT). Therefore, to enable machines to understand humans' intentions and assist humans in more efficient ways and to enhance the quality of human-computer interactions, proper tools and methods must be developed to improve the efficiency of NLP.

This Special Issue will focus on technologies and methodologies in the fields of NLP issues, also considering their combined use with algorithms, corpus-based approaches, IoT devices, internet resources, and AI methods. This Special Issue contains a total of 15 papers. Its content covers topics such as corpus design, analysis, and processing; the application of artificial intelligence methods for natural language processing; NLP applications; text classification; and corpus linguistics.

In terms of corpus design, analysis, and processing, several notable contributions stand out. Alfraidi et al. [1] introduced the Saudi Novels Corpus to advance the study of corpus stylistics in Arabic. Zhang [2] conducted a corpus-based systemic functional analysis to investigate the syntactic and semantic properties of 'on the contrary' in British university student essay writing. Chen et al. [3] utilized the lexical threshold theory and a corpus-based word classification method to assess the difficulty level of a military online English proficiency test.

In terms of applying artificial intelligence methods for natural language processing, several notable studies stand out. Ahmad et al. [4] employed an enhanced deep learning model to detect rumors on social media networks. Lin et al. [5] combined the analytic hierarchy process method with the artificial neural network method to address risk-related issues in construction projects. Chandio et al. [6] utilized the deep recurrent architecture of bidirectional long-short-term memory for sentiment analysis in Roman Urdu.

In terms of NLP applications, Alrumayyan and Al-Yahya [7] utilized language modeling and neural embeddings to support the task of jurisprudence principles. Al-Ghamdi et al. [8] employed the fine-tuning of Arabic bidirectional encoder representations from transformer-based models to develop Arabic pre-trained models. Alanazi [9] used cryptocurrency-related Twitter text to classify pure and compound sentiments. Nie and Li [10] introduced a self-adaptive learning framework to propose an automatic label correction mechanism for identifying mislabeled data and processing label corrections.



**Citation:** Chang, K.-H. Natural Language Processing: Recent Development and Applications. *Appl. Sci.* **2023**, *13*, 11395. https://doi.org/10.3390/app132011395

Received: 6 October 2023
Accepted: 10 October 2023
Published: 17 October 2023

In terms of text classification, Hu et al. [11] combined existing unlabeled data with prompt learning methods to classify small amounts of Chinese text data. Palomino and Aider [12] demonstrated that text preprocessing can enhance the accuracy of social media sentiment analysis classifiers.

In terms of corpus linguistics, Tantos and Kosmidis [13] employed network analysis to compare and analyze the C58 and STAC annotated corpora. Drawing from the broadcast news speech corpus and the Bible domain speech corpus, Brhanemeskel et al. [14] used a text-based query system with an automatic speech recognition module to facilitate Amharic speech search. Wang et al. [15] integrated the attention mechanism and adversarial training to propose a cross-lingual entity recognition method, effectively addressing the issue of semantic dilution.

While submissions for this Special Issue have closed, in-depth research on the development and application of natural language processing continues to address various challenges in natural language processing and artificial intelligence encountered in daily life.

**Funding:** This research received no external funding.

**Acknowledgments:** Thanks to all the authors and peer reviewers for their valuable contributions to this Special Issue, "Natural Language Processing: Recent Development and Applications". I would also like to express my gratitude to all the staff and people involved in this Special Issue.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

1. Alfraidi, T.; Abdeen, M.A.R.; Al-Thubaity, A. The Saudi novel corpus: Design and compilation. *Appl. Sci.* **2022**, *12*, 6648. [CrossRef]
2. Zhang, Y. Syntactic and semantic properties of on the contrary in British university student essay writing: A corpus-based systemic functional analysis. *Appl. Sci.* **2022**, *12*, 10635. [CrossRef]
3. Chen, L.C.; Chang, K.H.; Yang, S.C.; Chen, S.C. A corpus-based word classification method for detecting difficulty level of English proficiency tests. *Appl. Sci.* **2023**, *13*, 1699. [CrossRef]
4. Ahmad, T.; Faisal, M.S.; Rizwan, A.; Alkanhel, R.; Khan, P.W.; Muthanna, A. Efficient fake news detection mechanism using enhanced deep learning model. *Appl. Sci.* **2022**, *12*, 1743. [CrossRef]
5. Lin, C.L.; Fan, C.L.; Chen, B.K. Hybrid analytic hierarchy process–artificial neural network model for predicting the major risks and quality of Taiwanese construction projects. *Appl. Sci.* **2022**, *12*, 7790. [CrossRef]
6. Chandio, B.A.; Imran, A.S.; Bakhtyar, M.; Daudpota, S.M.; Baber, J. Attention-based RU-BiLSTM sentiment analysis model for Roman Urdu. *Appl. Sci.* **2022**, *12*, 3641. [CrossRef]
7. Alrumayyan, N.; Al-Yahya, M. Neural embeddings for the elicitation of jurisprudence principles: The case of Arabic legal texts. *Appl. Sci.* **2022**, *12*, 4188. [CrossRef]
8. Al-Ghamdi, S.; Al-Khalifa, H.; Al-Salman, A. Fine-tuning BERT-based Pre-trained models for Arabic dependency parsing. *Appl. Sci.* **2023**, *13*, 4225. [CrossRef]
9. Alanazi, S.A. Robust sentimental class prediction based on cryptocurrency-related tweets using tetrad of feature selection techniques in combination with filtered classifier. *Appl. Sci.* **2022**, *12*, 6070. [CrossRef]
10. Nie, B.L.; Li, C.Y. Distantly supervised named entity recognition with self-adaptive label correction. *Appl. Sci.* **2022**, *12*, 7659. [CrossRef]
11. Hu, T.K.; Chen, Z.Q.; Ge, J.K.; Yang, Z.X.; Xu, J.C. A Chinese few-shot text classification method utilizing improved prompt learning and unlabeled data. *Appl. Sci.* **2023**, *13*, 3334. [CrossRef]
12. Palomino, M.A.; Aider, F. Evaluating the effectiveness of text pre-processing in sentiment analysis. *Appl. Sci.* **2022**, *12*, 8765. [CrossRef]
13. Tantos, A.; Kosmidis, K. From discourse relations to network edges: A network theory approach to discourse analysis. *Appl. Sci.* **2023**, *13*, 6902. [CrossRef]
14. Brhanemeskel, G.M.; Abate, S.T.; Ayall, T.A.; Seid, A.M. Amharic speech search using text word query based on automatic sentence-like segmentation. *Appl. Sci.* **2022**, *12*, 11727. [CrossRef]
15. Wang, H.; Zhou, L.K.; Duan, J.Y.; He, L. Cross-lingual named entity recognition based on attention and adversarial training. *Appl. Sci.* **2023**, *13*, 2548. [CrossRef]