

Editorial

Machine Learning Advances and Applications on Natural Language Processing (NLP)

Leonidas Akritidis *  and Panayiotis Bozanis *

Department of Science and Technology, International Hellenic University, 57001 Thessaloniki, Greece

* Correspondence: lakritidis@ihu.gr (L.A.); pbozanis@ihu.gr (P.B.)

The recent technological advances in the research field of machine learning have played a crucial role in the improvement of Natural Language Processing (NLP). Today, state-of-the-art models and algorithms are allowing machines to understand, interpret, and generate human language of unprecedented quality. These advances allowed researchers to introduce effective tools and solutions for a wide variety of applications, including sentiment analysis, machine translation, conversational AI, question answering, named entity recognition, and others.

Deep learning stands at the heart of most modern NLP applications. Initially, the introduction of Recurrent Neural Networks (RNNs) [1] and their improved variants (i.e., Long-Short Term Memory–LSTM [2], Gated Recurrent Units–GRUs [3], etc.) allowed the effective processing of sequential data and the capture of contexts in text. Despite their inherent problems (i.e., unstable training due to the phenomenon of the vanishing and exploding gradients), these models have largely managed to overcome the severe limitations of the traditional NLP approaches.

Another milestone in the development of NLP was the introduction of word embeddings. Algorithms such as Word2Vec [4], GloVe [5], and FastText [6] have been designed to transform each word into a vector representation in a continuous space, while capturing the semantic relationships among words. These embeddings significantly improved the performance of the aforementioned NLP models across various tasks, replacing the sparse, non-informative TF-IDF vector representations [7].

The introduction of the Transformer architecture by Vaswani et al. in 2017 marked the beginning of the revolution era of NLP [8]. Unlike Convolutional Neural Networks (CNNs) [9] and RNNs, Transformers are not based on convolution or recurrence operations. Instead, they rely entirely on an innovative attention mechanism that enables the modeling of long-range dependencies in text. More specifically, the mechanism weighs the importance of different tokens in a sequence relative to a specific token. This is achieved by computing three vectors for each token: Query, Key, and Value. Each vector is obtained by multiplying the input embeddings with learned weight matrices.

Through its attention mechanism, the Transformer model became the building block of powerful pre-trained language models that revolutionized the area of NLP. In particular, BERT (Bidirectional Encoder Representations from Transformers) introduced a Transformer-based architecture that processes the entire sequence of words simultaneously, considering the context from both directions (namely, left-to-right and right-to-left) [10]. This bidirectional nature allows BERT to capture deeper semantic relationships in text.

After the introduction of BERT, numerous variants have been developed to enhance its performance and adapt it to different use cases. RoBERTa (Robustly optimized BERT approach) improves BERT by using a larger training corpus, eliminating the next-sentence prediction task, and training for longer periods [11]. On the other hand, DistilBERT is



Received: 1 August 2025

Accepted: 12 August 2025

Published: 19 August 2025

Citation: Akritidis, L.; Bozanis, P. Machine Learning Advances and Applications on Natural Language Processing (NLP). *Electronics* **2025**, *14*, 3282. <https://doi.org/10.3390/electronics14163282>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

a smaller and faster variant of BERT that is considered to be more suitable when the underlying computational resources are limited [12].

In contrast, GPT (Generative Pre-trained Transformer) adopts a different approach compared to BERT. More specifically, GPT is a unidirectional language model trained to predict the next word in a sentence, generating text based on the left-to-right context. While BERT excels in tasks requiring deep understanding of bidirectional context, GPT is designed for tasks such as text generation, language modeling, and conversational AI [13,14].

This Special Issue explores the most recent machine learning advancements in the research field of NLP. It includes ten original articles that systematically study popular NLP problems and introduce novel technologies, models, and algorithms to address them.

Sentiment analysis is a traditional NLP problem that focuses on the identification of the emotional tone behind a body of text. It is frequently treated as a typical classification problem that classifies the content of a text as positive, negative, or neutral. The relevant techniques can be applied to various data sources, including product reviews, social media posts, user comments, and customer feedback. These elements render them particularly important, since they provide the businesses and organizations with tools that allow them to gain insight into public opinion, customer satisfaction, and brand perception.

Motivated by the significance of the sentiment analysis techniques, the present Special Issue published five articles on the topic. More specifically, Y. Fu et al. (Contributor 5) introduced Self-HCL, a new method for multimodal sentiment analysis. Self-HCL first enhances the unimodal features using a unimodal feature enhancement module, and then, it jointly trains both multimodal and unimodal tasks. The proposed framework integrates a hybrid contrastive learning strategy with the aim of improving multimodal fusion and performance, even when unimodal annotations are lacking.

On the other hand, Faria et al. (Contributor 8) studied the emerging problem of sentiment analysis for memes in under-resourced languages. In this context, they developed three deep learning-based approaches: (i) a text-based model that uses Transformer architectures; (ii) an image-based model leveraging visual data for sentiment classification; and (iii) SentimentFormer, a hybrid model that integrates both text and image modalities. The authors evaluated the three models with the MemoSen dataset and concluded that the hybrid SentimentFormer model was the most effective. Moreover, Papageorgiou et al. (Contributor 1) investigated stock market prediction using reinforcement learning (specifically, a double deep Q-network), combined with technical indicators and sentiment analysis. The proposed model predicts short-term stock movements of NVIDIA, using data from Yahoo Finance and StockTwits. The results indicate that the inclusion of sentiment analysis elements in the prediction improves profitability and decision making.

Apart from the articles that introduce original models and techniques, this Special Issue also contains survey and investigation papers on sentiment analysis. More specifically, Kampatzis et al. (Contributor 2) conducted a survey that examined sentiment classification techniques in texts containing scientific citations. The authors explored various methods (from lexicon-based to machine and deep learning approaches) and highlighted the importance of interpreting both the emotional tone and intent behind citations. In another study, Kang et al. (Contributor 9) explored the use of GPT and FinBERT for sentiment analysis in the finance sector. The investigation focuses on the impact of news and investor sentiment on market behavior, and compares the performance of GPT and FinBERT, using a refined prompt design approach to optimize GPT-4o.

Text classification is not just limited to sentiment analysis applications. It is extended to cover other downstream tasks, such as named entity classification (e.g., news, products, articles, etc.), document categorization, acceptability of linguistic quality, and others. This

Special Issue includes two studies related to the generic field of text classification. The first one is the work of Kalogeropoulos et al. (Contributor 3), which enhances the Graphical Set-based model by integrating node and word embeddings in its edges. In particular, the proposed technique employs the well-established Word2Vec, GloVe, and Node2Vec algorithms with the aim of generating vector representations of the text. Subsequently, it utilizes these representations to augment the edges of the model in order to improve its classification accuracy. The second study was authored by Guarasci et al. (Contributor 4) and introduced a new methodology for automatically evaluating linguistic acceptability judgements using the Italian Corpus of Linguistic Acceptability. By leveraging the ELEC-TRA language model, the proposed approach outperformed the existing baselines and exhibited a capability in addressing language-specific challenges.

Named Entity Recognition (NER) constitutes another fundamental NLP task. Given a corpus of text, the goal of NER is to automatically identify and classify named entities into predefined categories. In other words, NER facilitates the recognition of key pieces of information within unstructured text. This is often proved to be crucial for tasks, such as information retrieval, question answering, and text summarization. In this spirit, Gao et al. (Contributor 6) presented a NER framework for extracting entities from Chinese equipment fault diagnosis texts. The framework integrates the following three models: RoBERTa-wwm-ext for extracting context-sensitive embeddings, a Bidirectional LSTM for capturing context features, and CRF for improving the accuracy of sequence labeling.

Finally, two research groups presented original studies on other interesting topics. More specifically, the work of Hirota et al. (Contributor 7) explored the use of descriptive text as an alternative to visual features in Visual Question Answering (VQA) tasks. Instead of relying on visual features, the proposed approach employs a language-only Transformer model to process description–question pairs. The authors also investigate strategies for data augmentation, with the aim of improving the diversity of the training set and reducing statistical bias.

Furthermore, Fernandes et al. (Contributor 10) evaluated the performance of 16 LLMs in automating engineering tasks related to Low-Power Wide-Area Networks. The main focus is whether lightweight, locally executed LLMs can generate correct Python code for these tasks. The models were compared with state-of-the-art models, such as GPT-4 and DeepSeek-V3. The evaluation revealed that while GPT-4 and DeepSeek-V3 consistently provided correct solutions, smaller models like Phi-4 and LLaMA-3.3 also performed well.

The diversity of the studies of this Special Issue indicates that NLP-related research constantly improves and evolves toward the introduction of models that truly understand the meaning of text. However, there are still many challenges on the way, including ambiguity and context understanding, performance improvement for low-resource languages, model explainability, and multimodal integration. Addressing these challenges is crucial for building more accurate and generalizable NLP systems.

Conflicts of Interest: The authors declare no conflicts of interest.

List of Contributions

1. Papageorgiou, G.; Gkaimanis, D.; Tjortjis, C. Enhancing Stock Market Forecasts with Double Deep Q-Network in Volatile Stock Market Environments. *Electronics* **2024**, *13*, 1629. <https://doi.org/10.3390/electronics13091629>.
2. Kampatzis, A.; Sidiropoulos, A.; Diamantaras, K.; Ougiaroglou, S. Sentiment Dimensions and Intentions in Scientific Analysis: Multilevel Classification in Text and Citations. *Electronics* **2024**, *13*, 1753. <https://doi.org/10.3390/electronics13091753>.
3. Kalogeropoulos, N.-R.; Ioannou, D.; Stathopoulos, D.; Makris, C. On Embedding Implementations in Text Ranking and Classification Employing Graphs. *Electronics* **2024**, *13*, 1897. <https://doi.org/10.3390/electronics13101897>.

4. Guarasci, R.; Minutolo, A.; Buonaiuto, G.; De Pietro, G.; Esposito, M. Raising the Bar on Acceptability Judgments Classification: An Experiment on ItaCoLA Using ELECTRA. *Electronics* **2024**, *13*, 2500. <https://doi.org/10.3390/electronics13132500>
5. Fu, Y.; Fu, J.; Xue, H.; Xu, Z. Self-HCL: Self-Supervised Multitask Learning with Hybrid Contrastive Learning Strategy for Multimodal Sentiment Analysis. *Electronics* **2024**, *13*, 2835. <https://doi.org/10.3390/electronics13142835>.
6. Gao, F.; Zhang, L.; Wang, W.; Zhang, B.; Liu, W.; Zhang, J.; Xie, L. Named Entity Recognition for Equipment Fault Diagnosis Based on RoBERTa-wwm-ext and Deep Learning Integration. *Electronics* **2024**, *13*, 3935. <https://doi.org/10.3390/electronics13193935>.
7. Hirota, Y.; Garcia, N.; Otani, M.; Chu, C.; Nakashima, Y. A Picture May Be Worth a Hundred Words for Visual Question Answering. *Electronics* **2024**, *13*, 4290. <https://doi.org/10.3390/electronics13214290>.
8. Faria, F.T.J.; Baniata, L.H.; Baniata, M.H.; Khair, M.A.; Bani Ata, A.I.; Bunternghit, C.; Kang, S. SentimentFormer: A Transformer-Based Multimodal Fusion Framework for Enhanced Sentiment Analysis of Memes in Under-Resourced Bangla Language. *Electronics* **2025**, *14*, 799. <https://doi.org/10.3390/electronics14040799>.
9. Kang, J.-W.; Choi, S.-Y. Comparative Investigation of GPT and FinBERT's Sentiment Analysis Performance in News Across Different Sectors. *Electronics* **2025**, *14*, 1090. <https://doi.org/10.3390/electronics14061090>.
10. Fernandes, D.; Matos-Carvalho, J.P.; Fernandes, C.M.; Fachada, N. DeepSeek-V3, GPT-4, Phi-4, and LLaMA-3.3 Generate Correct Code for LoRaWAN-Related Engineering Tasks. *Electronics* **2025**, *14*, 1428. <https://doi.org/10.3390/electronics14071428>.

References

1. Schuster, M.; Paliwal, K.K. Bidirectional Recurrent Neural Networks. *IEEE Trans. Signal Process.* **1997**, *45*, 2673–2681. [[CrossRef](#)]
2. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
3. Cho, K.; Van Merriënboer, B.; Bahdanau, D.; Bengio, Y. On the Properties of Neural Machine Translation: Encoder-Decoder Approaches. *arXiv* **2014**, arXiv:1409.1259. [[CrossRef](#)]
4. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient Estimation of Word Representations in Vector Space. *arXiv* **2013**, arXiv:1301.3781. [[CrossRef](#)]
5. Pennington, J.; Socher, R.; Manning, C.D. Glove: Global Vectors for Word Representation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014; pp. 1532–1543.
6. Bojanowski, P.; Grave, E.; Joulin, A.; Mikolov, T. Enriching Word Vectors with Subword Information. *Trans. Assoc. Comput. Linguist.* **2017**, *5*, 135–146. [[CrossRef](#)]
7. Akritidis, L.; Bozanis, P. How Dimensionality Reduction Affects Sentiment Analysis NLP Tasks: An Experimental Study. In Proceedings of the IFIP International Conference on Artificial Intelligence Applications and Innovations, Hersonissos, Greece, 17–20 June 2022; Springer: Berlin/Heidelberg, Germany, 2022; pp. 301–312.
8. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is All you Need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*. Available online: https://papers.nips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html (accessed on 1 July 2025).
9. LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
10. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Minneapolis, MN, USA, 2–7 June 2019; Volume 1 (Long and Short Papers), pp. 4171–4186.
11. Liu, Y.; Ott, M.; Goyal, N.; Du, J.; Joshi, M.; Chen, D.; Levy, O.; Lewis, M.; Zettlemoyer, L.; Stoyanov, V. RoBERTa: A Robustly Optimized BERT Pretraining Approach. *arXiv* **2019**, arXiv:1907.11692.
12. Sanh, V.; Debut, L.; Chaumond, J.; Wolf, T. DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter. *arXiv* **2019**, arXiv:1910.01108.

13. Radford, A.; Wu, J.; Child, R.; Luan, D.; Amodei, D.; Sutskever, I. Language models are unsupervised multitask learners. *OpenAI Blog* **2019**, *1*, 9.
14. Brown, T.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.D.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; et al. Language Models are Few-Shot Learners. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 1877–1901.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.