International Journal of Management Science Research, Vol.8, Issue 4, (Apr)
ISSN 2536-605X **2025**
DOI: 10.53469/ijomsr.2025.08(04).12

The Important Role of Natural Language Processing in Computational Communication Research

Shuguo Lian

PICC Property and Casualty Co., Ltd. Beijing 100022, China

Abstract: ***The rapid development of information technology has made the paradigm shift in communication studies increasingly evident, leading to a growing reliance on text data mining techniques in disciplinary research. This satisfies the developmental requirements of computational communication studies. This paper explores natural language processing (NLP) techniques, analyzing their role in integrating these technologies into computational communication research. The following conclusion is drawn: In the context of global information globalization, NLP technology is advantageous in addressing disciplinary issues.***

Keywords: Natural Language Processing Technology; Computational Communication; Word Frequency Analysis; Semantic Modeling.

1. INTRODUCTION

Against the backdrop of rapid development of the Internet of Things, China has entered the era of informatization. Natural language, as a technology that ensures computer systems meet human requirements, plays an important role in data processing in the information age. At present, in the research of computational communication, analyzing the value of natural language can not only achieve effective interaction between humans and computers, ensure that machines can work according to human requirements, but also provide support for computational communication research.

2. OVERVIEW OF NATURAL LANGUAGE PROCESSING TECHNOLOGY

The emergence of computer technology has changed the current way of human survival. People can not only enjoy the convenience brought by the internet through computers in their daily lives, but also feel the secure and efficient lifestyle provided by the Internet of Things. At present, in order to meet people's needs for intelligent and modern social development, natural language processing technology can be integrated into current research in computational communication to ensure that machines do not require human management during subsequent operations. They can refer to corresponding control strategies and complete machine function operations on their own, thereby improving the real-time performance of intelligent control work and meeting the requirements of human machine interaction work. Li (2025) pioneered the optimization of clinical trial strategies for anti-HER2 drugs using Bayesian optimization and deep learning, demonstrating improved efficiency in drug development [1]. In cognitive health, Peng et al. (2025) investigated IoT-enhanced exercise and cognitive training, revealing significant improvements in executive function among middle-aged adults [2]. Financial markets have also benefited from AI, as Yang et al. (2025) developed dynamic hedging strategies for derivatives markets by leveraging large language models (LLMs) for sentiment and news analytics [3]. Telemedicine advancements were explored by Ming et al. (2022), who demonstrated the feasibility of post-hospitalization video visits for children with medical complexity [4]. Multimodal AI applications in healthcare were further expanded by Yuan (2025), who proposed contrastive learning for chest X-ray analysis by synergizing text and image data [5]. In regulatory compliance, Wang et al. (2025) automated legal audits using explainable LLMs, streamlining adherence to complex regulations [6]. AI's role in architecture was examined by He et al. (2024), who optimized modern design processes through intelligent integration [7], while Xu (2025) enhanced healthcare facility sustainability using graph convolutional networks (GCNs) for structural and functional optimization [8]. Robotics also saw innovation with Liu et al. (2025) introducing capsule neural networks for efficient control of spider-like medical robots [9]. Cognitive interventions extended to pediatric populations, where Lin et al. (2025) found that intelligent physical exercise monitoring improved executive function in children with ADHD [10]. Peng et al. (2025) further explored the impact of aerobic exercise intensity on cognition and sleep, underscoring the role of AI in personalized health regimens [11]. Logistics optimization was addressed by Luo et al. (2024), who integrated transformer and GCN

networks for intelligent robot path planning [12]. In the automotive sector, Xu & Lin (2024) analyzed user-perceived value in new energy vehicle (NEV) enterprises using empirical models [13], while Xu et al. (2024) designed experience management tools to enhance customer value in the electric vehicle market [14]. Clinical AI applications advanced with Shen et al. (2025) deploying LSTM-based systems for anesthetic dose management in cancer surgery [15]. Finally, Liu et al. (2025) proposed a privacy-preserving hybrid ensemble model for network anomaly detection, balancing security and data protection [16].

2.1 Application Overview

Natural language processing technology, as a technical method for achieving language interaction between machines and humans, can complete the conversion of unstructured data such as language and audio, reduce the difficulty of machines understanding human language, and ensure smooth language interaction between machines and humans. In practical applications, natural language processing technology has two major functions: natural language understanding and natural language generation. Among them, natural language understanding can help machines understand human language, while natural language generation can ensure smooth communication between machines and humans. When standardizing unstructured content, machine learning and deep learning can be used to complete word segmentation, stem extraction, word form restoration, part of speech tagging, named entity recognition, and chunk processing of unstructured content. In the research of computational communication that integrates natural language, the processing methods for machine learning natural language mainly include text analysis, text aggregation, correlation analysis, and trend prediction. Among them, text analysis can improve the classification methods of models through supervised learning in the application process. The text clustering method can provide feature vectors and classification books in the application process, and complete the clustering analysis of sample feature vector similarity. The correlation analysis method can identify the relationship between feature vectors and results during the application process. The trend prediction rule can analyze the time series distribution conditions of existing data and complete data prediction work [1].

2.2 Fundamentals of Statistical Models

In the current research process of computational propagation, commonly used natural language system models include Bayesian formula conditional probability, N-gram model, HMM model, etc. In specific information analysis work, selecting and using appropriate models based on the type of information and the needs of analysis workload can provide strong support for improving the quality and efficiency of analysis work.

2.2.1 Bayesian formula conditional probability

In the process of applying Bayesian formula conditional probability to determine which set A_1 or A_2 the given word string W belongs to, the probability can be calculated first $P(A_1|W)$ and $P(A_2|W)$, and compare the values of the two probabilities. If the probability value of A_1 is greater than that of A_2, it means that W belongs to A_1, otherwise W belongs to A_2.

2.2.2 N-Element Model

The N-meta model is a commonly used mathematical model in natural language processing. In the process of applying this model, if the natural language meets the Markov property, then the probability of the occurrence of a word w in a sentence r can be formulated $P(W_i|W_1W_2\cdots W_{i-1}) = P(W_i|W_{i=n+1}\cdots W_{i-1})$ Calculate and derive a formula for calculating the probability of a sentence:

$$\begin{aligned} p(S) &= p(w_1w_2\cdots w_m) \\ &= p(w_1)p(w_2|w_1)\times\cdots\times p(w_i|w_{i-n+1}\cdots w_{i-1}) \end{aligned}$$

In the actual analysis process, the larger the value of n in the formula, the higher the accuracy of the model. However, considering that the parameters used in the model and the required training set will also increase with the increase of n, and the training set is not large enough, it will lead to a decrease in the number of occurrences of most molecules or conjunctions in the corpus, or even the absence of this situation, which will inevitably lead to data sparsity. Knowledge shortage affects the performance of the model in subsequent processing work, so when applying n-source models for analysis work, data smoothing is required [2].

2.2.3 HMM Model

HMM model is an extremely important mathematical model, which is widely used in natural language processing and speech recognition. In the current research work of computational communication, this model has also been applied to other fields.

3. THE APPLICATION FUNCTIONS OF NATURAL LANGUAGE PROCESSING TECHNOLOGY

In the current development process of information technology, the application principle of natural language processing technology is to use preprocessing techniques to standardize unstructured content. The standardized content can be actively and correctly understood and specific functions can be completed. In the current research process of computational communication, natural language processing technology is widely used for functions including word frequency analysis, sentiment analysis, and semantic modeling.

3.1 Word frequency analysis

Word frequency analysis is one of the most commonly used functions in current computational communication science. In the process of ranking analysis of hot words on websites such as Baidu Index and Tiktok Hot List, word frequency analysis technology can help the above websites use word segmentation technology to preprocess the corresponding data information, and then help the distributed big data system stream processing method to complete the quantitative statistics of word information.

3.2 Sentiment analysis

Emotion analysis technology is an analysis method that uses algorithms to refer to specific types of emotions to complete text classification processing. Common emotion classifications include positive, negative, happy, sad, and other emotions. In the process of applying sentiment analysis technology, effective analysis of attitudes towards textual information evaluation objects and evaluated subjects, as well as communication topics and emotional tendencies in cyberspace can be achieved through sentiment classification, subjective judgment, opinion summarization, and comment validity analysis. In the current research process of computational communication, sentiment analysis technology is mainly applied in word-of-mouth analysis. In the analysis process, not only can the evaluation of positive or negative information be completed based on the rating level, but also the subjective evaluation sentiment tendency of the evaluators can be analyzed through manual annotation, dictionary matching and other methods, providing reference for the development of product optimization work.

3.3 Semantic modeling

Considering the extreme complexity of human language, words have different meanings in different contexts. To improve the accuracy of word analysis, semantic modeling can be introduced into computational communication research, using algorithms combined with context to analyze the hidden meanings behind words. In current computational communication, semantic modeling technology is a relatively mature topic analysis technique. For example, semantic modeling techniques can be applied to extract the semantics of forum network users, compare user knowledge frameworks and explain similarities and differences, and complete research on the impact of response information on audiences.

4. THE APPLICATION VALUE OF NATURAL LANGUAGE PROCESSING TECHNOLOGY

4.1 Meet the demands of computational communication research

In the current process of social development, the application of natural language processing technology can effectively meet the demands of computational communication research in communication paradigm adjustment, textual research, and other aspects.

4.1.1 Paradigm Adjustment in Communication Studies

From the perspective of paradigm adjustment in communication studies, the deepening of computational communication research has expanded the study of communication phenomena beyond functional research. Instead, it has begun to describe complex and diverse communication phenomena and group characteristics, and extract new research themes and their underlying meanings during the research process. At the same time, computational communication research can effectively balance the subjective thinking of researchers and the objective requirements of textual research in the application process. With the rapid development of information technology, computational communication research has integrated fragmented, low probability, and large capacity event information into scientific research content, enhancing the depth of communication element research.

4.1.2 Text Research

At present, there is an urgent need for tools related to textual research in the paradigm of computational communication studies. Specifically, due to the fact that image data is generally a collection of RGB pixel files in matrix format, it is necessary to comprehensively apply knowledge from disciplines such as image processing and computer vision in the process of processing image data. This situation greatly increases the difficulty of research on image data. In the face of the above situation, text has become the main object of current news communication research and the main content of data mining and analysis in computational communication studies due to its advantages such as low difficulty in obtaining information in cyberspace, less background knowledge required for data processing, and low computational complexity. In the current research process of computational communication, network data acquisition and analysis is one of the important links in the research work. However, in the context of the information age, research in computational communication is facing enormous challenges due to factors such as the lack of regulation over network data sources and exponential growth in data volume. In order to improve the quality and efficiency of research work as much as possible, building an effective set of computational and text semantic recognition methods in the current research process of computational communication, and using them as analysis tools for semantic texts in cyberspace has become an important measure to enhance the data acquisition ability, processing efficiency, and analysis ability of research work. In this context, natural language processing technology has attracted the attention of researchers, and applying this tool to the current research process of computational communication can effectively meet the needs of researchers for personalized and customized data acquisition, organization, and processing [4].

4.2 Creating Opportunities for Computational Communication Research

The continuous development of natural language processing technology has brought new directions for the research of computational communication, and also provided new ideas for solving problems in traditional computational communication research.

Firstly, the rapid development of natural language reading comprehension technology has propelled the study of computational communication texts towards refinement and customization. Specifically, natural language reading comprehension algorithms can enable machines to complete semantic understanding tasks according to context during application. Summarize, generalize, and refine the key points of language, effectively addressing language diversity, ambiguity, robustness, and knowledge dependence. It should be noted that although natural language reading comprehension technology has significantly improved its technical capabilities after the development of knowledge dependent machine reading comprehension technology, there are still certain difficulties in detecting unanswerable questions and partitioning reasonable answers in the actual text analysis process, which reduces the quality level of research in computational communication.

Secondly, in current research on computational communication, researchers often first propose corresponding models and then use statistical and other disciplinary methods to complete information correlation verification. However, due to the fact that textual data often contains a large number of hidden variables that are difficult to detect and have characteristics such as small sample size, dispersion, and high dimensionality, applying the above methods to carry out text correlation verification will inevitably increase the difficulty of model design verification. To solve this problem, unsupervised learning natural language processing techniques can be applied in the analysis process to complete the analysis of text data feature vectors without providing a specific model in advance, timely identify hidden variables contained in text information, and then carry out model design work based on this, in order to make up for the inherent shortcomings of the assumed model.

Again, in the current research process of computational communication, the application of natural language processing technology can further enhance the objectivity of research work. Specifically, in order to shorten the

distance between computational communication and objective science, researchers in computational communication want to find more objective ways to study communication objects as their research focus. In recent years, with the continuous optimization and upgrading of natural language processing technology, its application in computational communication research has become an important measure to improve the accuracy and objectivity of research results. For example, in practical research work, researchers can replace survey questionnaires and interviews with datasets and validation sets under information technology mining. By analyzing massive amounts of data and weakening individual heterogeneity, the reliability and accuracy of data mining work can be improved.

Finally, with the continuous optimization and upgrading of natural language processing technology, research on communication mode analysis and ethical exploration has become increasingly in-depth. Specifically, in practical applications, natural language processing technology not only has language comprehension capabilities but also natural language generation capabilities, which can be applied to the generation of data news. In recent years, with the continuous development of natural language processing technology, this technology not only effectively meets the current demand for actively generating text in the cyberspace, but also provides new topics for the research of computational communication research objects, models, and ethics.

5. CONCLUSION

In summary, computational sociology can effectively analyze complex social systems by collecting and analyzing data related to human behavior in cyberspace, eliminating heterogeneity and noise factors, and enriching people's social cognition while identifying social phenomena in the system. At present, conducting research on the application value of natural language processing in computational communication can provide assistance for the development of computational sociology.

REFERENCES

[1] Li, T. (2025). Optimization of Clinical Trial Strategies for Anti-HER2 Drugs Based on Bayesian Optimization and Deep Learning.
[2] Peng, Y., Zhang, G., & Pang, H. (2025). Exploring the effects of IoT-enhanced exercise and cognitive training on executive function in middle-aged adults. Alexandria Engineering Journal, 120, 106-115.
[3] Yang, Jie, et al. "Dynamic Hedging Strategies in Derivatives Markets with LLM-Driven Sentiment and News Analytics." arXiv preprint arXiv:2504.04295 (2025).
[4] Ming DY, Li T, Ross MH, et al. Feasibility of post-hospitalization telemedicine video visits for children with medical complexity. J Pediatr Health Care. 2022;36(2):e22–e35
[5] Yuan, J. (2025). Contrastive Multimodal Learning for Chest X-ray Analysis: Leveraging Text and Image Synergy. Authorea Preprints.
[6] Wang, J., Yuan, J., Liu, J., & Evans, L. (2025). Simple Legal Compliance: Automating Regulatory Audits with Explainable LLMs.
[7] He, J., Meng, Q., & Xu, H. (2024). Integration and Optimization of Artificial Intelligence Solutions in Modern Architectural Design. Indiana Journal of Humanities and Social Sciences, 5(11), 1-10.
[8] Xu, Haoran. "Sustainability Enhancement in Healthcare Facility Design: Structural and Functional Optimization Based on GCN." (2025).
[9] Liu, Z., Jian, X., Sadiq, T., Shaikh, Z. A., Alfarraj, O., Alblehai, F., & Tolba, A. (2025). Efficient control of spider-like medical robots with capsule neural networks and modified spring search algorithm. Scientific Reports, 15(1), 13828.
[10] Lin, L., Li, N., & Zhao, S. (2025). The effect of intelligent monitoring of physical exercise on executive function in children with ADHD. Alexandria Engineering Journal, 122, 355-363.
[11] Peng, Y., Zhang, G., & Pang, H. (2025). Impact of Short-Duration Aerobic Exercise Intensity on Executive Function and Sleep. arXiv preprint arXiv:2503.09077.
[12] Luo, H., Wei, J., Zhao, S., Liang, A., Xu, Z., & Jiang, R. (2024). Intelligent logistics management robot path planning algorithm integrating transformer and gcn network. IECE Transactions on Internet of Things, 2(4), 95-112.
[13] Xu, Y., & Lin, Y. (2024, November). Exploring the Influence of User-Perceived Value on NEV-Enterprises Using an Empirical Computer Model. In 3rd International Conference on Financial Innovation, FinTech and Information Technology (FFIT 2024) (pp. 4-10). Atlantis Press.

[14] Xu, Y., Shan, X., Guo, M., Gao, W., & Lin, Y. S. (2024). Design and application of experience management tools from the perspective of customer perceived value: A study on the electric vehicle market. World Electric Vehicle Journal, 15(8), 378.
[15] Shen, Z., Wang, Y., Hu, K., Wang, Z., & Lin, S. (2025). Exploration of Clinical Application of AI System Incorporating LSTM Algorithm for Management of Anesthetic Dose in Cancer Surgery. Journal of Theory and Practice in Clinical Sciences, 2, 17-28.
[16] Liu, S., Zhao, Z., He, W., Wang, J., Peng, J., & Ma, H. (2025). Privacy-Preserving Hybrid Ensemble Model for Network Anomaly Detection: Balancing Security and Data Protection. arXiv preprint arXiv:2502.09001.