Research Statement

WANG Zhaoxia
School of Computing and Information Systems, Singapore Management University
Tel : (65) 6828-1347 ; Email : zxwang@smu.edu.sg
31 (Day) 10 (Month) 2025 (Year)

Background

In recent years, AI has emerged to be a powerful force, both in academia and in commercial deployment. In terms of commercial values, in 2017 PwC projected that in 2030, global GDP will be higher by 14% - which translates to US$15.7 TRILLION - because of the deployment of AI. I have set my sight to continue to develop my research along the line that I have always pursued: BIG DATA and AI techniques, especially **next generation AI techniques as well as their applications**.

Even though existing AI technologies, such as machine translation system, Chat Robots and other intelligent robots have performed to some level of satisfaction, **true language understanding for the purpose of directing an agent (e.g., a robot) to perform tasks is still impossible.** Currently, the most advanced robots are just showing talents in science fiction movies, rather than in the real-world.

ChatGPT is a remarkable model developed by **OpenAI**. It can comprehend and generate coherent, contextually relevant responses in conversational contexts. This powerful NLP tool offers numerous advantages. However, one of its shortcomings is the lack of true language understanding or real human-like understanding. Other large language models (LLMs), such as **o1-preview (OpenAI), LLaMA (Meta), Mistral (Mistral AI), Qwen (Alibaba), Orca (Microsoft), and InternLM (Shanghai AI Lab),** share the same shortcoming of lacking true human-like reasoning and understanding capabilities.

Therefore, to address this gap, there is a need to focus on **Next Generation AI** and the applications. True language understanding or real human-like understanding is certainly one important research direction of Next Generation AI. To demonstrate true language understanding, embodiment is a critical step. **Next-generation AI systems** are an interdisciplinary field that benefits from inspiration across a variety of domains, such as linguistics, computer science, psychology, neuroscience, statistics, engineering, sociology, economics, and more. This interdisciplinary approach to AI is currently being pursued by researchers from various institutes, such as researchers and professors at SMU, NTU, NUS, A*STAR, Peking University, and the Beijing Institute for General AI (BIGAI). I plan to collaborate with them continuously.

My research creating next generation AI systems is to **build AI systems which can simulate the capability of humans**, and the main research topics of the targeted next generation AI are as followings:

1. Natural Language Understanding (through embodiment)
2. Learning (simulating learning capability of humans)
3. Knowledge Representation, Management and Reasoning including Logiic Analysis and Planning, Causality, etc. (simulating the capability of knowledge management and logic analysis of humans)
4. Computer Vision (simulating vision capability of humans)
5. Robotics (integrating the achievements from all the 4 research topics above to build a next generation embodied AI system)

6. **Real-world applications and deployments of the above AI research** (e.g., AI in Education, AI in transportation and supply chain, AI in Finance, AI in insurance, AI in maritime applications, etc.)

Specifically, the particular emphasis on one set of important topics of next generation AI, **Human-like Natural Language Understanding (NLU),** will greatly facilitate human-machine interactions through embodiment, text mining, sentiment analysis, machine learning and their applications as laid out in detail below. These areas are important as research topics of next generation AI and in terms of engendering valuable commercial applications and deployments.

Research Areas and Summary

Research Areas: Next generation AI, focused on Natural Language Processing (NLP) and Natural Language Understanding (NLU), Machine Learning, Data Analytics, Deep Language Understanding and Intelligent Robots, Casual Reasoning and Language Understanding, Image Processing, Text Mining, Fine-grained Sentiment and Emotion Analysis, Social Computing, Social Media Content Mining and Analysis, Knowledge Representation, Management and Reasoning, Artificial Intelligence (AI) & Computational Intelligence (CI) and their real-world applications and deployments.

Research Summary: I have more than 15 years of research and development work experience as a scientist or senior scientist in the AI research areas. I also have a few of years of teaching experience in the university as an associate professor since 2004, teaching courses related to AI, e.g., machine learning, data mining, NLP for smart assistants, etc. I enjoyed my work very much. Doing research gives me the opportunity to propose new ideas and to develop new algorithms to make contributions to those areas. Teaching in the universities gives me the opportunity to make contributions to the cultivation of the young generation, helping them to grow and develop rapidly to be good persons including excellent professionals.

I have already accumulated rich research and development experiences in big data and AI technologies with 70+ papers published in international journals and conferences and 5+ patents/Technical Disclosures (TDs) filed as the first inventor. Evaluation licenses or commercial licenses have been signed by more than 10 companies on my patents. As a highly motivated team player and an expert in fine-grained sentiment and emotion analysis, social data analytics, AI, and data mining, I have led and successfully delivered 10+ research and industry projects as well.

I will enhance the methods which I have developed to further develop new NLP techniques to make contributions to **next generation AI technologies** (e.g., human-like AI, explainable AI, rather than existing black-box unexplainable AI). My detailed research areas are described as follows which I have organized into 5 aspects below:

- **Human-Like Natural Language Understanding**
- **Enhancements and Applications of Foundation Models (e.g., various learning-based models, LLMs)**
- **Knowledge Representation for Next Generation AI**
- **Research on Applications and Deployments of AI Techniques (e.g., AI in Education, AI in transportation and supply chain, AI in Finance, AI in insurance, AI in maritime applications, etc.)**
- **Cultivating Young Generations**

Research Aspects

1. Human-Like Natural Language Understanding

I would like to continue my research work on **social media sentiment analysis** and **natural language processing** to develop new data analysis and AI technologies toward true **human language understanding** and **true human like AI technologies,** which are very **important research topics** for **next generation AI techniques**.

The existing methods on sentiment analysis and natural language processing can only achieve simple opinion classification analysis or sentiment analysis. They are still unable to realize real understanding of language in the same way as that understood by human beings. What I have been doing, human-like language understanding, is the ultimate goal of researchers in this field.

2. Enhancements and Applications of Foundation Models (e.g., various learning-based models, LLMs)

Even though the latest input-output learning methods have performed to some level of satisfaction, true language understanding for the purpose of directing an agent (e.g., a robot) to perform tasks is still impossible. Currently, the most advanced robots are just showing talent in science fiction movies, rather than in the real world.

Therefore, there is a need to do research on the enhancement of the **foundation models** such as the various existing learning-based models and LLMs. One of the ways to **enhance existing learning methods**, such as neural network-based learning algorithms, is to incorporate knowledge models to develop human-like explainable learning methods. The other way is to **enhance existing foundation models, such as LLMs, and applying them to actionable solutions** that can be implemented in real-world settings, while creating technologies and strategies that directly address societal needs.

3. Knowledge Representation for Next Generation AI

Research and development on next generation AI (also named human-like AI) is one of my research plans and goals. I would like to develop human-like AI technologies that will enable future AI robots to truly understand human language and better serve humans (for example, to enable machines to understand human language which can be applied to the care of the elderly: nursing robots for elderly people, companion robots, chat robots, etc.).

We know that atoms are the basic units of matter and the defining structure of the elements in the physical world. Concepts that are “atomic” in nature (named “atomic concepts”) can be used to form the foundation of much more complex, high-level concepts, much like complex molecules are built from atoms. Just like when we are building skyscrapers, a solid foundation is the most important step toward its success.

To build next generation AI techniques, we assume that “atomic” basic concepts are the smallest and basic elements of the NLP or NLU domain. Therefore, we propose “atomic” basic concepts (also named ground concepts) as well as “atomic” basic concept representations (also named ground concept representations) to enhance NLP research, and the other AI-related research. The overall idea/design is shown in Fig. 1, which includes four important modules:

- Module (A) and (B) are constructed through a process of crowd sourcing, automatic online sourcing and machine learning.
- Module (C) converts a language-dependent surface sentential structure into a language independent deep-level predicate representation which is related to our physical world.
- Module (D) converts the predicate representation into grounded real-world references and constructs that enable a robot to carry out the language instructions accordingly.

As shown in Fig. 1, an "atomic" basic concept (also named ground concept) as well as an "atomic" basic concept representation (also named ground concept representation) are proposed based on the existing efforts. The "atomic" basic concepts are the smallest and most basic elements in the NLP and NLU world [1]. Atomic concepts can then form the foundation of much more complex, high-level concepts, much like complex molecules built from atoms. Such "atomic" basic concepts build the foundation for NLP and NLU, especially for true human-like explainable NLP and NLU [1].

My current works, such as **fine-grained sentiment and emotion analysis algorithms**, MiMuSA—**mimicking human language understanding for fine-grained multi-class sentiment analysis**, and **AI-based intelligent fine-grained multi-level** social media opinion extraction and public opinion analysis methods [1-19] lay a solid foundation for achieving the goals of this research plan. The initial idea is to combine knowledge-based expert system and machine learning technology to develop true human-like AI technologies which will entail new methods.

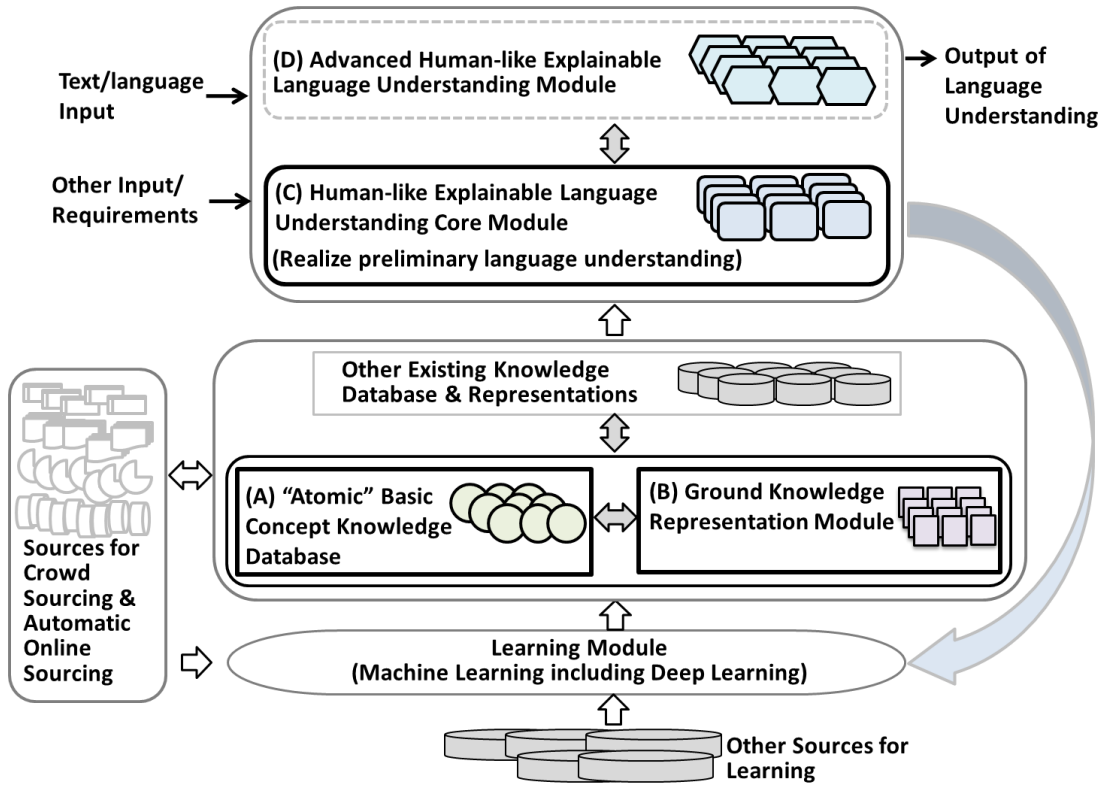

Fig. 1. The Overall Idea/Design of the Proposed Method –
"Human-like Explainable Language Understanding with Atomic Basic Concept Representation" for implementation of next generation AI techniques [1]

4. Research on Applications and Deployments of AI Techniques

4.1 Application Research 1 (AI in Finance): Opinion-based Financial Analysis

My experience in social media analytics and opinion analysis work, such as sentiment and emotion analysis, laid a solid foundation for my stock market trending analysis work. The stock market has been characterized as dynamic and non-linear. This is because the relationship between stock market parameters and target stock market prices is not linear. Political stability, economic performance, even some events relevant to the companies may have an influence on the trending of the stock market prices.

Different events may affect public sentiments and emotions differently, which may in turn have an effect on the trend of stock market prices. Therefore, news articles and social media data are very useful and important in financial analysis. However, currently no good methods exist that can take these social media into consideration to provide better analysis of the financial market. Recently, I have pioneered a method to overcome these limitations by leveraging my previous work on time series analysis and sentiment analysis.

I had several years' experience working on time series data analysis by leveraging AI-based methods, such as AI learning-based network time series traffic analysis [20]. I have extended the network traffic analysis methodology to the finance domain to analyze stock market time series data [21-22]. Incorporating my research experience in social media sentiment analysis [1-14] and leveraging the AI based time series analysis technology [20], my latest work on stock market analysis overcomes the limitations of traditional financial analysis technologies by incorporating news articles, event and social media analysis into stock market analysis. This work has been published [5] [21] [22] [31] and my related work on comparing learning-based methods for stock market prediction has won the best paper award in ICCCS (International Conference on Cloud Computing and Security) 2018 [22].

In summary, my existing works on time series analysis, social media analysis, sentiment and emotion analysis, machine learning, stock market prediction and trending analysis, intelligent fine-grained multi-level social media opinion extraction and public opinion analysis methods, natural language processing and natural language understanding, and other AI research works as well as their applications and deployments lay a solid foundation for achieving the goals of this research.

4.2 Application Research 2 (AI in Education, AI in transportation and supply chain, AI in insurance, etc.): Multidisciplinary Application of AI Technologies

AI is an interdisciplinary subject that benefits from inspiration from a variety of domains such as linguistics, psychology, neuroscience, statistics, computer science, engineering, sociology, etc. Its areas of applications are also very wide, ranging from social media, information security, finance, transportation, energy, earth science, urban mobility, insurance etc.

In addition to the specific areas mentioned in the research plans above, as a data and AI scientist, I would also like to extend big data analysis and AI technologies into additional application domains, as shown in Fig. 2, in the future. This will involve collaborations with colleagues and researchers both within my research area and from other domains. The core technologies of big data and AI (e.g., various learning-based models, LLMs), can be applied across various fields [22-39], contributing to new research and applications in areas such as smart finance [4, 21-22], intelligent control, intelligent transportation, intelligent robotics, fault detection and diagnosis [23-27], smart image processing [28-29], sentiment and emotion sensing [1-4, 30-32], and more.

Moreover, I would like to explore industrial collaborations with companies as well. The collaborations with industry will enable us to transfer the core technologies of big data analysis and AI technologies into businesses or services, so as to allow our research achievements to go beyond academic papers and create impact on society (e.g., produce real-world intelligent robots to benefit society).

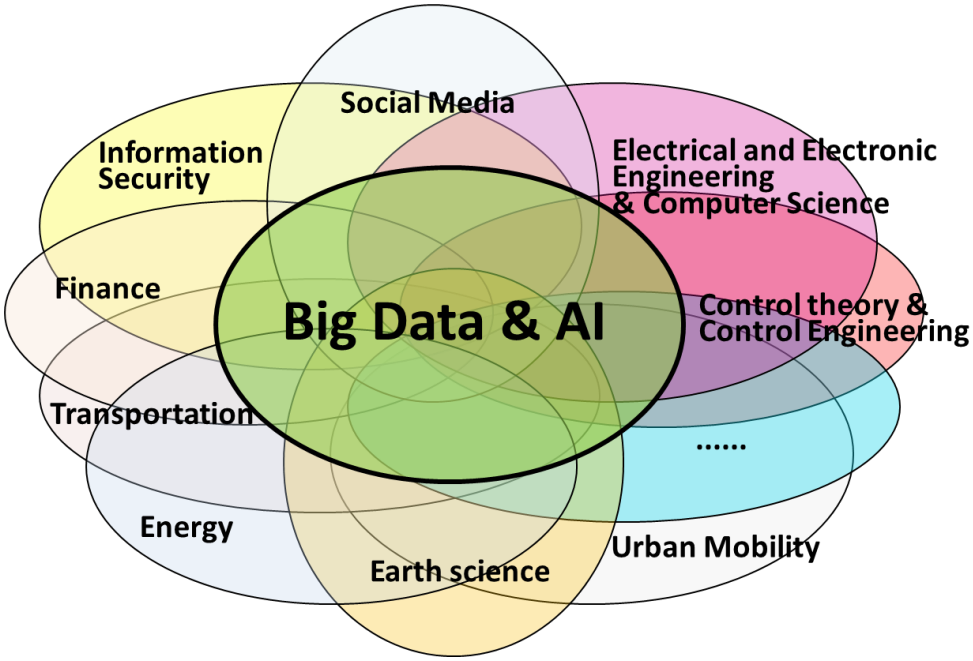

Fig. 2. Big Data and Artificial Intelligence
(To carry out interdisciplinary research collaborations with other researchers from institutes, universities, or companies in industry. The core technologies of big data analysis and AI will be applied to solve the problems from different domains.)

5. An AI-Care Lab

In line with the current sentiment that AI should serve the interests of the public, we believe it is important that AI technology be human-centric. At the core of being human-centric is one essential principle: to care. This means caring about the many aspects of life—human needs, human activities, and human well-being.

Therefore, we are building a laboratory around the theme of “AI Care”, named **AI-Care Lab.** Over the past years, we have already conducted substantial preliminary work aligned with this theme, and many of our efforts have been published, as shown below [1-39]. Our work spans a wide range of topics, including Video Monitoring to Protect Children (VideoCare), Emotion-Aware AI Assistant (EmoCare), Education Companion (EduCare), Financial Assistant (FinCare), Library Assistant (LibCare), MusicCare, Healthcare Assistant, FakeCare, MaritimeCare, and more.

I have extensive experience in conducting, publishing, and patenting research in affective computing, sentiment analysis, and AI-Care related works, which provides a strong foundation for developing care-oriented AI systems.

6. Cultivating Young Generations

At the same time, in addition to the above research plan and work, I hope to cultivate the young generation (undergraduates as well as graduate students), specifically cultivating postdoctoral fellows, young researchers who are interested in big data and artificial intelligence, etc. I would also like to have the opportunity to train international postdoctoral fellows. Outstanding postdoctoral fellows will have the opportunity to work in our group to contribute outstanding achievements to strengthen the reputation of our university, SMU.

Selected Publications and Outputs

1. **Z. Wang,** Z. Hu, S-B. Ho, E. Cambria, A. H. Tan, MiMuSA—mimicking human language understanding for fine-grained multi-class sentiment analysis. *Neural Computing and Applications*, vol.35, 1-15, 2023.
2. J. Cui, **Z. Wang,** S-B. Ho, E. Cambria, Survey on sentiment analysis: evolution of research methods and topics. Artificial Intelligence Review, vol.56, 8469–8510, 2023.
3. Z. Hu, **Z. Wang**, Y. Wang, A. H. Tan, MSRL-Net: A multi-level semantic relation-enhanced learning network for aspect-based sentiment analysis, *Expert Systems with Applications*, 119492, 2023.
4. Z. Hu, **Z. Wang,** Y. Wang, A. H. Tan, Aspect Sentiment Triplet Extraction Incorporating Syntactic Constituency Parsing Tree and Commonsense Knowledge Graph. *Cognitive Computation*, vol.15, no. 1, 337-347, 2023.
5. **Z. Wang,** Z. Hu, F. Li, S-B. Ho, E. Cambria, Learning-based stock trending prediction by incorporating technical indicators and social media sentiment, *Cognitive Computation*, vol 15, no.3, 1092-1102, 2023.
6. **Z. Wang**, S-B. Ho, E. Cambria, A Review of Emotion Sensing: Categorization Models and Algorithms, *Multimedia Tools and Applications*, vol. 79, no. 47, pp.35553-35582, 2020.
7. **Z. Wang** and Z. Lin, Optimal Feature Selection for Learning-Based Algorithms for Sentiment Classification, *Cognitive Computation*, vol. 12, no. 1, pp. 238-248, 2020.
8. S.-B. Ho, **Z. Wang**, Language and Robotics: Complex Language Understanding. *Proceedings of the 12th International Conference on Intelligent Robotics and Applications, Part IV*, Shenyang, China, August 8-10, pp. 641-654, 2019.
9. S.-B. Ho, **Z. Wang**, On True Language Understanding. *Proceedings of the 5th International Conference on Artificial Intelligence and Security, Part II*, New York, July 26-28, 2019, pp. 87-99, Switzerland: Springer Nature.
10. **Z. Wang**, J.C. Tong, P. Ruan, F. Li, Lexicon Knowledge Extraction with Sentiment Polarity Computation, *The IEEE International Conference on Data Mining series (ICDM 2016), SENTIRE*, Barcelona, Spain, Dec. 12-15, 2016.
11. **Z. Wang**, C.S. Chong, L. Lan, Y. Yang, S-B Ho and J.C. Tong, Fine-grained Sentiment Analysis of Social Media with Emotion Sensing, *IEEE Future Technologies Conference 2016 (FTC 2016),* San Francisco, United States, Dec. 6-7, 2016.
12. **Z. Wang**, Y. Parth, Extreme Learning Machine for Multi-class Sentiment Classification of Tweets, *The International Conference on Extreme Learning Machines (ELM2015)*, Hangzhou, China, Dec. 15-17, 2015. In Proceedings of ELM-2015, vol.1, pp. 1-11. Springer International Publishing, 2016.
13. **Z. Wang**, J.C. Tong, D. Chan, Issues of Social Data Analytics with a New Method for Sentiment Analysis of Social Media Data, *6th IEEE International Conference on Cloud Computing Technology and Science (IEEE cloudcom 2014),* pp. 899-904, Singapore, Dec. 15-18, 2014.
14. **Z. Wang**, J.C. Tong, H.C. Chin, Enhancing Machine-learning Methods for Sentiment Classification of Web Data, Information Retrieval Technology, pp. 394-405, 2014. Springer. *10th Asia Information Retrieval Society Conference (AIRS) 2014*, Sarawak, Malaysia, Dec. 3-5, 2014.
15. **Z. Wang**, J.C. Tong, X. Xin, H.C. Chin, Anomaly Detection through Enhanced Sentiment Analysis on Social Media Data, *6th IEEE International Conference on Cloud Computing Technology and Science (IEEE cloudcom 2014),* pp. 917-922, Singapore, Dec. 15-18, 2014.
16. Z. Chen, **Z. Wang**, Z. Lin, T. Yang, Comparing ELM with SVM in the Field of Sentiment Classification of Social Media Text Data, *The International Conference on Extreme Learning Machines (ELM2017),* Yantai, China, Oct. 4-7, 2017.
17. **Z. Wang,** J.C. Tong, S.-B. Ho, "Intelligent Sensing System-Method and System of Intelligent Sentiment and Emotion Sensing with Adaptive Learning", Patent Cooperation Treaty (PCT) International Application No.: PCT/SG2017/050172, International Filing Date: 30 March 2017.

18. **Z. Wang,** R. Goh, Y. Yang, “A method and system for sentiment classification and emotion classification”, US Patent Application No.: 15/523,201, International Filing Date: 24 November 2015, National Phase Entry Date: 28 April 2017. (Singapore Patent Application No.: 11201704150W, International Filing Date: 24 November 2015, National Phase Entry Date: 22 May 2017).
19. **Z. Wang**, J.C. Tong, “Fine-Scoring Sensing system Multi-Level Fine-Scaled Sentiment and Emotion Sensing Method and System”, Technique Disclosure (TD): IHPC-TD-SCC-2016-014.
20. **Z. Wang,** Y. Sun, Z. Chen and Z. Yuan, The Research on Predicting Network Traffic Using Neural Networks, *Journal of Opto-electronics•Laser,* vol.17, no.10, pp.1255-1258, 2006 (in Chinese).
21. **Z. Wang**, Z. Lin, S-B. Ho, Stock Market Prediction Analysis by Incorporating Social and News Opinion and Sentiment, *The IEEE International Conference on Data Mining series* (ICDM 2018), *SENTIRE*, pp.1375-1380, Singapore, Nov. 21-23, 2018.
22. **Z. Wang**, Adrian Tan, F. Li, S-B. Ho, Comparisons of Learning-Based Methods for Stock Market Prediction, *The 4th International Conference on Cloud Computing and Security (ICCCS 2018), Haikou*, China, Jun. 8 - 10, 2018 (**The Best Paper)**.
23. **Z. Wang** and C.S. Chang, A Feature Based Frequency Domain Analysis Algorithm for Fault Detection of Induction Motors, *the 6th IEEE Conference on Industrial Electronics and Applications (ICIEA2011),* pp. 27-32, Beijing, China, Jun. 21- 23, 2011.
24. **Z. Wang** and C.S. Chang, Online Fault Detection of Induction Motors Using Frequency Domain Independent Components Analysis, *IEEE International Symposium on Industrial Electronics*, pp. 2132-2137, Gdansk, Poland, Jun. 27-30, 2011.
25. **Z. Wang,** C.S. Chang, X. German and W. W. Tan, Online Fault Detection of Induction Motors Using Independent Component Analysis and Fuzzy Neural Network, *The 8th IET International Conference on Advances in Power System Control, Operation and Management*, pp.321-325, Hong Kong, China, Nov. 8-11, 2009.
26. **Z. Wang**, C.S. Chang, F. Yang and W. W. Tan, Adaptive Type-2 Fuzzy Maintenance Advisor for Offshore Power Systems, *IEEE International Conference on Systems, Man, and Cybernetics*, pp. 4520-4526, Texas, USA, Oct. 11-14, 2009.
27. T. Yang, H. Pen, **Z. Wang,** C.S. Chang, Feature Knowledge Based Fault Detection of Induction Motors Through the Analysis of Stator Current Data, *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 3, pp. 549-558, IEEE, 2016.
28. **Z. Wang,** H. Pen, T. Yang, Q. Wang, Structure-priority Image Restoration through GA Optimization, *IEEE Access*, vol. 8, no. 1, pp. 90698-90708, 2020.
29. H. Pen, Q. Wang, **Z. Wang,** Boundary precedence image inpainting method based on Self-organizing Maps, *Knowledge-Based Systems,* p.106722, 2021.
30. **Z. Wang**, S-B. Ho, E. Cambria, Multi-Level Fine-Scaled Sentiment Sensing with Ambivalence Handling, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol.28, no.4, pp.683-697, 2020.
31. Z. Hu, **Z. Wang,** S-B Ho, A-H Tan, Stock Market Trend Forecasting Based on Multiple Textual Features: A Deep Learning Method, *The 33rd IEEE International Conference on Tools with Artificial Intelligence (ICTAI2021),2021*.
32. **Z. Wang,** D. Huang, J. Cui, X. Zhang, S-B. Ho, E, Cambria, A Review of Chinese Sentiment Analysis: Subjects, Methods, and Trends, Artificial Intelligence Review, vol. 58, no. 3, pp. 75, 2025.
33. D Huang, **Z Wang,** Evaluation of orca 2 against other LLMs for retrieval augmented generation, RAFDA workshop, *Pacific-Asia Conference on Knowledge Discovery and Data Mining* (PAKDD) 2024, 3-19, 2024.
34. X Zhang, **Z Wang,** H Wang, J Xiang, C Wu, G Cao, CausVSR: Causality inspired visual sentiment recognition, *International Joint Conference on Artificial Intelligence* (*IJCAI) 2024*, Jeju, Aug. 3-9, 2024.
35. D. Huang, **Z. Wang**, Explainable Sentiment Analysis with DeepSeek-R1: Performance, Efficiency, and Few-Shot Learning. *IEEE Intelligent Systems*, 2025. doi: 10.1109/MIS.2025.3614967.
36. X. Zhang, **Z. Wang**, H. Wang, G. Cao, Counterfactual thinking driven emotion regulation for image sentiment recognition. In Proceedings of the *34th International Joint Conference on Artificial Intelligence (IJCAI) 2025*, Aug 16 (pp. 4273-4281).
37. D. Huang, **Z. Wang**, LLMs at the Edge: Performance and Efficiency Evaluation with Ollama on Diverse Hardware, *International Joint Conference on Neural Networks (IJCNN)* 2025, IEEE, 2025, pp. 1–8.
38. Y.S. Tan, **Z. Wang**, Explainable Multimodal Sentiment Analysis of Social Media Visual Content for Child Safety, *IEEE Intelligent Systems*, vol. 40, no. 4, pp.64-72, 2025.
39. D. Huang, T. S. Nguyen, F. Liausvia, **Z. Wang**, RAP: A Metric for Balancing Repetition and Performance in Open-Source Large Language Models. In Proceedings of the 2025 *Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics (NAACL):* Human Language Technologies (Volume 1: Long Papers) (pp.1479-1496). 2025.