

ITAI 2373 – Natural Language Processing

Student: Natalia Solorzano W207818526

Professor: Narges DeBary

Spring 2025 Course Number: 16579

A09: Language Models and Transformers in NLP

Natural Language Processing (NLP) has rapidly evolved thanks to two core innovations: language models and transformers. These technologies have enabled computers to handle language-based tasks more accurately and contextually. This paper briefly examines the development of language models and how transformer-based systems have reshaped NLP.

Language Models in NLP

Language models help systems predict what words are likely to appear in a sentence, based on what comes before. They are widely used in tools like autocomplete, speech-to-text, and online translation. Older models like n-grams calculated word probabilities but had trouble understanding longer or more complex phrases.

Modern models now use artificial neural networks that represent words in a numerical format, allowing them to learn patterns, associations, and relationships between words in a more flexible way. This shift paved the way for more complex and effective models.

Transformers: Revolutionizing NLP

Transformer models brought a major change to NLP by introducing a new way to understand language. They don't look at sentences word by word like older approaches. Instead, they analyze entire sequences at once, which helps them understand context better. A key part of this system is called self-attention, which helps the model figure out which parts of the sentence are most important when interpreting meaning.

This approach allows the model to recognize word relationships across long distances in text, making it more accurate in tasks that require deep understanding of meaning and context.

Impact on NLP Applications

Today, transformers are used in many AI systems, including:

- Translating text between languages with higher accuracy.
- Creating short summaries of articles or documents.
- Answering questions in customer support or virtual assistants.
- Generating content for creative or academic writing tasks.

Their ability to process large amounts of information quickly while keeping track of meaning has made them the foundation for advanced tools like BERT and ChatGPT.

Conclusion

Language models and transformers have made it possible for computers to understand human language more like we do. With improved accuracy and the ability to capture context, they support everything from translation to AI chat tools. As this field continues to grow, we can expect even more useful and responsive applications in everyday life.

References

- TechTarget. (2023). *What Is Language Modeling?* Retrieved from <https://www.techtarget.com/searchenterpriseai/definition/language-modeling>
- Wikipedia contributors. (2023). *Language model*. Wikipedia, The Free Encyclopedia. Retrieved from https://en.wikipedia.org/wiki/Language_model
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). *Attention Is All You Need*. Retrieved from <https://arxiv.org/abs/1706.03762>
- DataCamp. (2022). *How Transformers Work: A Detailed Exploration of the Architecture in NLP*. Retrieved from <https://www.datacamp.com/tutorial/how-transformers-work>
- Hugging Face. (2023). *How do Transformers work?*. Retrieved from <https://huggingface.co/learn/nlp-course/en/chapter1/4>
- Data Science Dojo. (2023). *Transformer Models: NLP's New Powerhouse*. Retrieved from <https://datasciencedojo.com/blog/transformer-models/>