

Foundation Models for Language

PAULO CAVALIN

IBM Research, Brazil, pcavalin@br.ibm.com

Keywords: neural networks, language models, foundation models

In this presentation I will talk about the recently-famous and large neural language models such as BERT, GPT-3, and the like, which have been bringing great impact on NLP applications. More specifically, I will cover the different types of models, their architecture, and how they are developed. With this talk I not only expect to help expanding the portfolio of techniques for those already working with NLP, but also opening up the minds for those designing ML systems for other domains, given the potential of such models as a new paradigm for reuse of ML models.

Bibliography

- [1] JACOB DEVLIN ET AL., *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, Volume 1 (Long and Short Papers), pages 4171-4186, Minneapolis, Minnesota. Association for Computational Linguistics, 2019.
- [2] TOM BROWN ET AL., *Language Models are Few-Shot Learners*, Part of Advances in Neural Information Processing Systems 33 (NeurIPS 2020), 2020
- [3] RISHI BOMMASANI ET AL., *On the Opportunities and Risks of Foundation Models*, <https://arxiv.org/abs/2108.07258>, 2021.