



Contribution ID: 225

Tür: **Oral Presentation**

Large Language Models

19 Aralık 2024 Perşembe 17:30 (50 dakika)

Language models have been a staple component of systems for natural language processing for about 35 years and have been applied in applications such as spelling correction, speech recognition, and statistical machine translation early on, to name a few. Recently however, there has been a remarkable surge in the capabilities and applications of language models. These modern “large language models”, enabled by advances in deep learning and in hardware technology and by availability of huge amounts of data, have led to rapid implementation of many applications ranging from text generation to sentiment analysis, question answering, enhanced machine translation, and many others, some including multiple modalities. Clearly many more applications are in our future.

This lecture will be an accessible introduction of the basic ideas behind language models and their evolution from the earlier models into the large models currently being used today. It will discuss how changes in language representation about a decade ago led to the use of “semantics” instead of superficial representations.

Presentation language / Sunum Dili

EN (English)

Disciplines / Disiplinler

E-mail / E-posta

ko@andrew.cmu.edu

ORCID ID

0000-0002-4977-0079

Institution / Affiliation / Kurum

Carnegie Mellon University, School of Computer Science, Language Technologies Institute

Country / Ülke

Başlıca yazarlar:: Kemal Oflazer (Carnegie Mellon University)

Sunu yapanlar: Kemal Oflazer (Carnegie Mellon University)

Session Classification: Keynote Session 4