Teaching Machines to Communicate with People using Deep Learning
Zornitsa Kozareva, Google

Abstract:
Over the years there has been a paradigm shift on how humans interact with machines. Today's users are no longer satisfied with seeing a list of relevant web pages, instead they want to complete tasks and take actions. This raises the questions on "*How do we teach machines to communicate with people and be more useful in a human-centered environment?*" and "*How do we build machines that help us organize our daily lives, arrange travel and become aware of our preferences and habits?*". In this talk, I will describe these challenges in the context of conversational assistants. Then, I will delve into deep learning algorithms for natural language understanding, user intent prediction and question answering.

While these deep neural networks achieve state-of-art performance by running on high performance computing hardware such as GPUs and TPUs on cloud, yet it remains challenging to run them on devices with tiny memory footprint and limited compute capacity such as mobile phones, smart watches and IoT. Next, I will highlight our new efforts for on-device deep learning for natural language processing by introducing novel on-device Self-Governing Neural Networks (SGNNs), which learn compact projection vectors with local sensitive hashing. The key advantage of SGNNs over existing work is that they surmount the need for pre-trained word embeddings and complex networks with huge parameters. Finally, I will conclude with results from extensive evaluations on wide range of natural language tasks such as dialog act classification, user intent prediction, sentiment analysis, emotion recognition among others and show how SGNNs achieved significant improvement over state-of-the-art results, while yielding fast and tiny models.