Graduate School of Science and Technology Master's Thesis Abstract

Laboratory name (Supervisor)	Natural Language Processing (WATANABE TARO (Professor))		
Student ID	2111311	Submission date	2023 / 1 / 19
Name	LI HUAYANG		
Thesis title	Residual Learning of Neural Text Generation with n-gram Language Model		

Abstract

N-gram language models (LM) have been largely superseded by neural LMs as the latter exhibits better performance. However, we find that n-gram models can achieve satisfactory performance on a large proportion of testing cases, indicating they have already captured abundant knowledge of the language with relatively low computational cost. With this observation, we propose to learn a neural LM that fits the residual between an n-gram LM and the real-data distribution. The combination of n-gram and neural LMs not only allows the neural part to focus on the deeper understanding of language but also provides a flexible way to customize an LM by switching the underlying n-gram model without changing the neural model. Experimental results on three typical language tasks (i.e., language modeling, machine translation, and summarization) demonstrate that our approach attains additional performance gains over popular standalone neural models consistently. We also show that our approach allows for effective domain adaptation by simply switching to a domain-specific n-gram model, without any extra training.