香港浸會大學理學院
HKBU Faculty of Science

DEPARTMENT OF COMPUTER SCIENCE

PhD Degree Oral Presentation

PhD Candidate:	Mr Dong QIAN
Date	20 December 2022 (Tuesday)
Time:	10:00 am – 12:00 pm (35 mins presentation and 15 mins Q & A)
Venue:	ZOOM (Meeting ID: 998 2889 5054) (The password and direct link will only be provided to registrants)
Registration:	https://bit.ly/bucs-reg (Deadline: 6:00 pm, 19 December 2022)

Neural Sequence Modeling with Deep Generative Models

Abstract

A fundamental goal of probabilistic sequence modeling is to learn a joint probability distribution over a sequence of discrete symbols, such as words in a sentence. Neural autoregressive models have been a standard method for modeling sequences, yielding the state-of-the-art performance over a variety of tasks. Despite the success, neural autoregressive modeling has weaknesses: (1) training autoregressive models corresponds to one-step predictions, making long-term dependencies hard to be captured; (2) there exist an objective mismatch and a distribution discrepancy between the training and evaluation. In this thesis, we propose a series of improved training approaches for generative models.

First, we apply variational autoencoders (VAEs) with autoregressive models for modeling sequences, where the encoder learns holistic semantics into a latent variable, while the decoder applies the learned variable to reconstruct low-level details. To alleviate posterior collapse encountered in training VAEs, we propose to add a mutual information term, parameterized via neural networks, to VAEs' objective. Experimental results show that our method outperforms baselines and generates reasonable sentences by linear interpolation in the latent space.

Second, we propose to incorporate a hierarchical structure into VAEs to learn multiple levels of more abstract representations, where a sequence of latent variables is organized with a Markovian structure, and the decoder is parameterized via autoregressive decoders. We analyze the role of the encoder and decoder inside HVAEs. Together with mutual information maximization, experimental results on text and image datasets show that our method yields competitive performance and shapes a coarse-to-fine hierarchical organization.

Third, we propose a VAE-based approach with a sequence of latent variables to ensure the contextual coherence throughout multi-turn dialogues. To this end, we encourage consecutive representations to be similar by mutual information maximization, such that relevant factors can be extracted to decide if latent variables are encoded. Experimental results show that our method outperforms baselines and generates coherent dialogue responses with context transitions being detected.

Finally, to optimize a non-differentiable sequence-level metric, we apply a recently proposed training paradigm, called generative flow networks (GFlowNets), which sample a diverse set of sequences with probability proportional to a metric via a constructive sequence of actions. We develop a new variant for training GFlowNets with the goal of further enhancing the exploration capability. The preliminary results on text completion show that our method can facilitate exploration for more rewarding words to alleviate degenerate behavior. Also, it can achieve the least empirical L1 error and visit all the modes quickly in hypergrid environments, where modes occupy a tiny fraction of the total space, and are far from each other, separated by low-probability areas.

***** ALL INTERESTED ARE WELCOME *****