Review of Bachelor's Thesis

Student: Bolshakova Ksenia
Title: Stylized Natural Language Generation in Dialogue Systems (id 22587)
Reviewer: Kesiraju Santosh, DCGM FIT BUT

1.   Assignment complexity   **considerably demanding assignment**
The work explores generating stylized text in an open-ended dialogue system.
The task of generating natural reply in a dialogue system is itself a challenging task. Introducing style into the conversations make it even more challenging.

2.   Completeness of assignment requirements   **assignment fulfilled**
The student reviewed the current literature, datasets and state-of-the-art (SotA) methods and found an interesting task of stylized text generation for DS. The student used SotA ta pre-trained models for the experiments and analysis.

3.   Length of technical report   **in usual extent**
The thesis work is substantial.

4.   Presentation level of technical report   **87 p. (B)**
The thesis is very well organized into chapters that are form a continuity.
The initial chapters describe the problems and discusses methods with their limitations. These form the basis for the later chapters that introduce the proposed approach and experiments.
Some of the Equations are hard to follow. Certain symbols are not used without defining them. V_l in Eq. (3.9)
Eq. (7.1) cross entropy is the training objective, but it is not clear what $p(x)$ and $q(x)$ represents? Is the model training on predicting the next word, just like any language model training? Or is it conditioned on something else

5.   Formal aspects of technical report   **83 p. (B)**
The thesis is well written and understandable mostly. However the explanations in the Chapter 8 are little difficult to follow.

6.   Literature usage   **95 p. (A)**
The student has a done a good review of overall literature, explaining the techniques and their drawbacks.

7.   Implementation results   **87 p. (B)**
The student used existing open source software (PyTorch) and publicly avaliable pre-trained models (BART, GPT) and built on top of it. This makes it relatively easier for others to compare or re-implement.
However certain details of training and hyper-parameter configurations are not fully described in the paper.

8.   Utilizability of results
The work is very new in terms of research and has a potential for expansion and further investigation. The student has run several experiments and reported many examples. However, there are few experimental results which I feel are missing:
1. Different weighting for SST and poetic styles -> would it improve the generation?
2. All the examples compare two AI systems. How about a human in the conversation?

9.   Questions for defence
Q1. What is the intuition behind switching the models (GPT, BART) for every 3 words ?
Why do you think it is better?
Q2. How much influence does the weighting have in the generation of the dialogues. In the thesis, you have used 0.3 for pre-trained model and 0.7 for the stylized model. Does different stylized models (jokes, poetry,) require different weights?

10.   Total assessment   **90 p. excellent (A)**
1. Stylized natural dialogue generation is a challenging research work and the student had managed to make some progress.
2. The work has definite industrial applications which makes it more interesting.
3. The thesis also covers the literature in quite detail which makes it self-contained.

In Brno 24 June 2020

 Kesiraju Santosh
reviewer