ChiNet

Nil Adell **Joseph Cornelius** **Alexander Nedergaard** **Lama Saouma**

Feel free to add more sections but those listed here are strongly recommended.

1 Introduction

You can keep this short. Ideally you introduce the task already in a way that highlights the difficulties your method will tackle.

2 Methodology

Your idea. You can rename this section if you like. Early on in this section – but not necessarily first – make clear what category your method falls into: Is it generative? Discriminative? Is there a particular additional data source you want to use?

3 Model

The math/architecture of your model. This should formally describe your idea from above. If you really want to, you can merge the two sections. **Xander is on this**

We define our sentence RNN as

$$h_i^s = GRU(s_i; h_{i-1}^s)$$

where s denotes the embedded sentence. We denote the final hidden state of the sentence RNN as r^s.
We then define our document RNN as

$$h_i^d = GRU(r_i^s; h_{i-1}^d)$$

and similarly denote the final hidden state of the document RNN as r^d.
Attention is defined as

$$\tilde{r}_i^s = r_t^s W_A (r_i^s)^T$$

where r_t^s denotes the target sentence and W_A is the attention matrix.
Now, we define our generator RNN as

$$h_i^g = GRU(y_i; h_{i-1}^g)$$

where y_i is the word generated at the previous time step. We initialize y_0 as the embedded <start> word. Unlike the sentence and document RNN, where the initial hidden states h_0^s and h_0^d are set to 0, we initialize our generator hidden state as

$$h_0^g = r^d + z$$
$$z \sim \mathcal{N}(0, 1)$$

We determine the generated word from the generator hidden state using the Gumbel-Softmax

$$h- > y.equation$$

Our discriminator score is defined as

$$D = \sigma(r^d W_{d \to s} (r_t^s)^T)$$

where $W_{d \to s}$ denotes a transformation matrix from document space to sentence space and σ is the sigmoid function.

4 Training

What is your objective? How do you optimize it? **Xander is on this**

5 Experiments

This **must** at least include the accuracy of your method on the validation set.

6 Conclusion

You can keep this short, too.