Introduction

- Common NLP benchmarks often neglect performance on out-of-distribution data, which can become problematic as inference data can often be different than training data.
- We extend a method proposed by Lee et al. whose novel idea is to jointly train a question-answering (QA) model together with a discriminative model that forces the encoder to learn domain-agnostic embeddings.
- We include an additional classifier module to handle datasets that include questions without any answer.
- We also use enrich the training procedure with data augmentation techniques such as backtranslation.

Method

QA model loss:

$$
L_{QA} = -\frac{1}{N} \sum_{k=1}^{K} \sum_{i=1}^{N_i} \log P_q(y_i^{(k)} | h_i^{(k)}) + \log P_e(y_i^{(k)} | h_i^{(k)})
$$

$$
L_{Adv} = \frac{1}{N} \sum_{k=1}^{K} \sum_{i=1}^{N_i} KL(D_i(h_i^{(k)}) || P_D(h_i^{(k)}))
$$

Discriminator loss:

$$
L_D = -\frac{1}{N} \sum_{i=1}^{N_i} \sum_{k=1}^{K} \log P_D(y_i^{(k)} | h_i^{(k)})
$$

Backtranslation:

- (a) ... published by Sony Computer Entertainment of America, released on PlayStation 2.
- (b) ... published by Sony Computer Entertainment of America, released on PlayStation 2.
- (c) ... published by Sony Computer Entertainment of America, released on PlayStation 2.
- (d) ... published by Sony Computer Entertainment of America, released on PlayStation 2.
- (e) ... published by Sony Computer Entertainment of America, released on PlayStation 2.
- (f) ... published by Sony Computer Entertainment of America, released on PlayStation 2.

References