< What to talk about and how? Selective Generation using LSTMs with Coarse-to-Fine Alignment >

Hongyuan Mei, NAACL 2015

Pros

1. The paper treats selective generation as end-to-end learning without the need for an external aligner or ground-truth selection labels.
2. The paper divides the search complexity into pre-selection and refinement stages which avoids performing exhaustive search over the full record set. And they also present model ablations to elucidate those different components.
3. The authors attach the record-word alignment heat map in qualitative analysis, which illustrates the ‘garbage collection’ clearly.

Cons

1. As it learns content selection and surface realization jointly, the ground-truth selection labels are not visualized. The ablation analysis part should clarify which type of records are aligned incorrectly by heat-map or diagram. The scenarios in weather forecasting dataset contains 36 weather records, which is not too difficult to evaluate.
2. One-hot vector used in event record representation is too sparse which could lead to the `garbage collection` potentially. In order to retain its binning-style information between nearby numbers, we could use vector representation for numerical attributes to replace one-hot.
3. They achieve further improvement (0.65 in sBLEU) on ‘WEATHERGOV’ via a k-nearest neighbor beam filter under the setting (beam width M: 2, nearest neighbor K: 1). They should attach an explanation about the physical meaning of M/K to present their motivation of this filter instead of showing the performance comparison directly.

< Learning when to trust distant supervision: An application to low-resource POS tagging using cross-lingual projection>

Meng Fang, CoNLL 2016

Pros

1. The author is so good at telling story that the motivation and main contribution can be easy to understand.
2. Explicitly model the bias affecting cross-lingual projected annotations instead of capturing the correlations between tag sets.
3. The training objective combining two cross-entropy terms for gold supervision and projected data is inspiring. I will try this function to augment our NCRF model by treating high-confidence data as gold label, other as projected label.

Cons

1. The cross-lingual projection between low-resource language and projected language is a temporary solution. It highly relies on the quality of translation quality.
2. They should provide a curve about how many annotated data used to do initialization and correspond performance, as the pre-train process strongly influence the model’s optima.

3. The bias transformation matrix between POS tags and projected outputs in Kinyarwanda has many high magnitude off-diagonal elements. It may be caused by the transition process for too many unaligned tokens. However, author didn't give a reason or explanation about this.

<Autoencoding beyond pixels using a learned similarity metric>

Anders Boesen Lindbo Larsen, ICML 2016

Pros

1. The paper is clearly written and is fairly easy to follow, especially in the diagram.
2. The authors cleanly combined two very different approaches to generative modeling. The resulting model does indeed produce sharper imaged than VAEs.
3. The paper provides useful tips in training the proposed network followed by comprehensive empirical comparison with competing methods via visual inspection.

Cons

1. As with GANs, it is unclear how to evaluate such models properly, though the authors do a reasonable job in the paper. Still, the paper does not quite convince me that this interesting development is a substantial advance.
2. The experimental validation is mostly based on the visual inspection, while lack of control experiment. It might be worth discussing which evaluation metrics are available for VAE/GAN compared to plain VAEs and GANs.
3. The experimental section could be made stronger by performing a hyper parameter search for each method.