Exploring Explainable Selection to Control Abstractive Summarization
Wang Haonan¹, Gao Yang¹, Bai Yu¹, Mirella Lapata², Huang Heyan¹
¹School of Computer Science and Technology, Beijing Institute of Technology
²Institute for Language, Cognition and Computation, School of Informatics, University of Edinburgh

1 Introduction
Rely on selecting informative content (extractor) as well as aggregating into a summary in line with linguistic expression (abstractor). The extractors are largely black-box decisions without a rationale of what is informative content. We need methods for identifying the sentence informativeness, identifying whether a sentence is relevant to a document and, if so, to what extent. Another importance influence is the novelty of the contribution a sentence makes to a summary. Therefore, to reveal more of the inner workings of these black-box models so as to inject a level of control into the substance and integrity of the final summary, we developed a novel select-and-generate framework, called ESCA, that focuses on explain-ability.

2 Contributions
- Proposed an explainable content selection module for document summarization.
- Ability to extract the appropriate content for generating a desired summary based on explicit and quantified measures of informativeness, relevance and novelty to the final summary.
- Automatically creating synthetic datasets w.r.t novelty and relevance for exercising controllable inference without the need to retrain the entire system.

3 Proposed Model
- We present a novel hybrid framework for document summarization ESCA, equipped with a pair-wise ranking extractor that connects with an abstractor armed with a sentence-level attention pointer.
- ESCA is designed to explicitly explain why sentences are marked for extraction and to control which sentences are extracted according to novelty and relevance score.
- The relationship between sentence pairs contributes more to the degree of sentence importance than the relationship between sentence and document.
- Abstractive summary is a trade-off between controllability and quality of summary, high-quality summary should focus on more attributes information.

4 Crucial formula
- Pair-wise Learning Loss for Extractor
\[ L_{ext} = -\sum_{i=1}^{m} \sum_{j=1}^{m} \left( \hat{p}_{ij} \log r_{ij} + (1 - \hat{p}_{ij}) \log (1 - r_{ij}) \right) \]
- Construct interaction matrix of directional sentence pair relation.
\[ q_{ij}(h_i, a_i, h_j, a_j, d) = \sigma(W_{i}h_i + h_{i}W_{d}d + h_{i}W_{j}h_j - h_{i}W_{n}\tanh(a_i) + b_{m}) \]
- Hidden state & accumulated summary representation
\[ h_{t} = a_{t}Z_{c} \]
\[ a_{t} = \frac{1}{\sum_{t=1}^{T} h_{t} \cdot q_{tk}} \]
- Output distribution of decoder
\[ P_{vocab}(w) = P(y_1 \leq t, x; \theta) = \softmax(W_{2}(W_{1}[s_{0}, h_{i}^t] + b_{1}) + b_{2}) \]
- Controllable Inference
\[ Q^{2} = Q^{1} \odot M, \text{ where } M_{ij} = \begin{cases} 1, & \text{val} \geq \epsilon \\ 0, & \text{val} < \epsilon \end{cases} \]

5 Experiment

6 Conclusion & Case Study
- A novel hybrid framework for abstractive summarization ESCA, equipped with a pair-wise ranking extractor that connects with an abstractor armed with a sentence-level attention pointer.
- ESCA is designed to explicitly explain why sentences are marked for extraction and to control which sentences are extracted according to novelty and relevance score.
- The relationship between sentence pairs contributes more to the degree of sentence importance than the relationship between sentence and document.
- Abstractive summary is a trade-off between controllability and quality of summary, high-quality summary should focus on more attributes information.