



CSE442
NATURAL LANGUAGE-DRIVEN UI/UX
Spring 2024

Name: Jannatul Ferdoshi
ID: 20301193

Paper Title:

Parameter-Efficient Sparse Retrievers and Rerankers using Adapters

Paper Link:

<https://paperswithcode.com/paper/parameter-efficient-sparse-retrievers-and?fbclid=IwAR3fBRmre8ZFVBwmyagJnReLMWPKjnqZoEIHk4RiQgHYETSciAWCJgC8SwI>

1 Summary**1.1 Motivation**

In order to train neural retrieval models such as SPLADE in Information Retrieval tasks, adapters are investigated in this research as a potentially economical solution. Its objectives are to maximize effectiveness, examine information exchange across retrieval phases, assess domain adaptation, and assess the impact of adapters. The idea behind adapters is that they can improve the efficacy and efficiency of models in IR activities.

1.2 Contribution

In order to train neural retrieval models such as SPLADE in information retrieval tasks, adapters are presented in this research as an affordable solution. The usefulness of adapters is demonstrated, their influence on retrieval effectiveness is examined, domain adaptation is assessed, and information transfer throughout retrieval stages is investigated.

1.3 Methodology

Ablation analyses on adapter layers, comparing first-stage retrievers and rerankers' knowledge transfer strategies, analyzing parameter sharing with adapters, assessing adapter performance against existing techniques, evaluating adapter-tuned SPLADE across benchmark and out-of-domain datasets, and exploring adapter-tuning for generalizability are all part of the study's methodology. With regard to training neural retrieval models across a range of tasks and domains, these methods offer a comprehensive knowledge of adaptor efficacy.

1.4 Conclusion

According to the study's findings, adapters are a practical and affordable substitute for complete fine-tuning when training neural retrieval models like SPLADE for information retrieval tasks. The retrieval process might be improved by adapters since they have promise in increasing efficacy, efficiency, and generalizability across several domains.

2 Limitations

2.1 First Limitation

One potential weakness of the work is its concentration on a particular neural retrieval model, namely SPLADE, which might restrict the applicability of the results to other retrieval models. The efficiency of employing adapters may vary among retrieval models because to differences in their designs, training methods, and features. As such, not all brain retrieval models in the field of information retrieval may directly benefit from the findings and conclusions of this study.

2.2 Second Limitation

The study's primary evaluation of adapter utilization using SPLADE, a sparse retriever model, is another factor associated with the first constraint. Compared to thick retrievers or other retrieval models that are frequently employed in information retrieval, sparse retrievers may have different requirements and special qualities. Therefore, the results about the efficiency and efficacy of adapters with SPLADE might not apply directly to other retrieval models or dense retrievers. In order to give more thorough insights on adapters' efficacy and generalizability in information retrieval tasks, this limitation emphasizes the need for more study to examine their applicability across a wider range of retrieval models.

3 Synthesis

Broad implications for information retrieval and natural language processing stem from the paper's results on adapter-based sparse retrieval models.

First off, by cutting training time without compromising performance, adapter tuning's efficiency provides useful advantages for time-sensitive applications like recommendation systems or search engines. Adapters have potential in domain adaptation tasks as well, allowing for rapid adaptation to other domains or user preferences, hence improving the relevancy of search results and suggestions. Finally, investigating knowledge transfer between rankers and rerankers offers ways to optimize model topologies and boost the performance of multi-stage retrieval systems.

To sum up, the concepts presented in this study have potential applications across several domains and offer intriguing avenues for future research in the fields of NLP and information retrieval.