The Journal of Open Source Software

gobbli: A uniform interface to deep learning for text in Python

Jason Nance[1] **and Peter Baumgartner**[1]

1 RTI International

DOI: 10.21105/joss.02395

Software

- Review
- Repository
- Archive

Editor: Arfon Smith

Reviewers:

- @w4ngatang
- @ljvmiranda921
- @sisco0

Submitted: 26 May 2020
Published: 24 June 2021

Summary

Machine learning has long been used to address natural language processing (NLP) tasks like sentiment analysis (Pang & Lee, 2008) and document classification (Aggarwal & Zhai, 2012). Traditional approaches to these tasks require numerous labeled examples from the specific domains in which they will be applied. Such algorithms can only use the available training data, which is often limited in size and diversity, to learn to understand natural language. In the last few years, transfer learning (Weiss et al., 2016) has caused a paradigm shift in NLP. Rather than training distinct task-specific models from scratch, a transfer learning model first learns the rules of language from a large, diverse text corpus during an extensive self-supervised training regimen. Self-supervised tasks are formulated such that unlabeled data can be used to train supervised models (Raina et al., 2007); an example is masked language modeling, where a subset of words in a document are masked out, and the model predicts the masked words (Devlin et al., 2018). The transfer learning model thus learns a rich representation of language, which can be fine-tuned to solve specific problems. According to Torrey & Shavlik (2010), this approach mimics how humans reuse their general understanding of language across tasks. Transfer learning has not only rapidly advanced the state of the art in classification but has enabled near-human performance on more advanced tasks like question answering (Raffel et al., 2019; Rajpurkar et al., 2016) and natural language inference (Bowman et al., 2015; Rajpurkar et al., 2016; Williams et al., 2018).

While the performance gains on benchmark tasks are undeniable, applied researchers face challenges using transfer learning models to solve new problems. A wide variety of models are being developed by disparate research teams using different technologies (Liu et al., 2019; Raffel et al., 2019; Sun et al., 2019). A practitioner may therefore be required to learn a new programming language, a deep learning library, a containerization technology, and a model interface whenever they want to evaluate the feasibility of a new model on a custom task. `gobbli` was developed to address this problem.

`gobbli` is a Python library intended to bridge state-of-the-art research in natural language processing and application to real-world problems. The library defines a simple interface for training classification models, producing predictions, and generating embeddings. Several models implementing the interface are available using programmatically-created Docker containers to abstract away differences in underlying deep learning libraries and model hyperparameters. This approach allows users to easily evaluate models and compare performance across model types without spending time adapting their dataset and use case to each model. Compared to other deep learning libraries used for NLP like `transformers` (Wolf et al., 2019) and `fastai` (Howard & others, 2018), `gobbli` is designed to emphasize simplicity and interoperability rather than customization and performance in order to make deep learning more accessible to applied researchers.

Beyond its model wrappers, `gobbli` provides several helpful utilities for NLP practitioners. Data augmentation has emerged as a popular technique to improve model performance when

training data is limited. Multiple methods for data augmentation are implemented in gobbli, including backtranslation (Shleifer, 2019), word replacement (Wei & Zou, 2019), and contextual augmentation (Kobayashi, 2018). These methods can be used independently of gobbli models for interoperability with other modeling libraries. gobbli also bundles a set of interactive web applications built using Streamlit (Teixeira et al., 2018) which can be used to explore a dataset, visualize embeddings, evaluate model performance, and explain model predictions without writing any code.

gobbli was developed from experiences on client contracts and internal projects at RTI International. It is intended for use by anyone solving problems using applied NLP, including researchers, students, and industry practitioners.

Acknowledgments

Work on gobbli was funded by RTI International.

References

Aggarwal, C. C., & Zhai, C. (2012). A survey of text classification algorithms. In *Mining text data* (pp. 163–222). Springer. https://doi.org/10.1007/978-1-4614-3223-4_6

Bowman, S. R., Angeli, G., Potts, C., & Manning, C. D. (2015). A large annotated corpus for learning natural language inference. *arXiv Preprint arXiv:1508.05326*. https://doi.org/10.18653/v1/d15-1075

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv Preprint arXiv:1810.04805*.

Howard, J., & others. (2018). *Fastai*. https://github.com/fastai/fastai; GitHub.

Kobayashi, S. (2018). Contextual augmentation: Data augmentation by words with paradigmatic relations. *arXiv Preprint arXiv:1805.06201*. https://doi.org/10.18653/v1/n18-2072

Liu, X., He, P., Chen, W., & Gao, J. (2019). Multi-task deep neural networks for natural language understanding. *arXiv Preprint arXiv:1901.11504*. https://doi.org/10.18653/v1/p19-1441

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, *2*(1-2), 1–135.

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu, P. J. (2019). Exploring the limits of transfer learning with a unified text-to-text transformer. *arXiv Preprint arXiv:1910.10683*.

Raina, R., Battle, A., Lee, H., Packer, B., & Ng, A. Y. (2007). Self-taught learning: Transfer learning from unlabeled data. *Proceedings of the 24th International Conference on Machine Learning*, 759–766.

Rajpurkar, P., Zhang, J., Lopyrev, K., & Liang, P. (2016). SQuAD: 100,000+ questions for machine comprehension of text. *arXiv Preprint arXiv:1606.05250*. https://doi.org/10.18653/v1/d16-1264

Shleifer, S. (2019). Low resource text classification with ULMFit and backtranslation. *arXiv Preprint arXiv:1903.09244*.

Sun, Y., Wang, S., Li, Y., Feng, S., Chen, X., Zhang, H., Tian, X., Zhu, D., Tian, H., & Wu, H. (2019). ERNIE: Enhanced representation through knowledge integration. *arXiv Preprint arXiv:1904.09223*.

Teixeira, T., Treuille, A., & others. (2018). *Streamlit*. https://github.com/streamlit/streamlit; GitHub.

Torrey, L., & Shavlik, J. (2010). Transfer learning. In *Handbook of research on machine learning applications and trends: Algorithms, methods, and techniques* (pp. 242–264). IGI Global.

Wei, J. W., & Zou, K. (2019). EDA: Easy data augmentation techniques for boosting performance on text classification tasks. *arXiv Preprint arXiv:1901.11196*. https://doi.org/10.18653/v1/d19-1670

Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. *Journal of Big Data*, *3*(1), 9.

Williams, A., Nangia, N., & Bowman, S. (2018). A broad-coverage challenge corpus for sentence understanding through inference. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (long Papers)*, 1112–1122. https://doi.org/10.18653/v1/n18-1101

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., & Brew, J. (2019). HuggingFace's transformers: State-of-the-art natural language processing. *ArXiv*, *abs/1910.03771*.