NLP+Vis: NLP Meets Visualization

Shafiq Joty[♦♠], **Enamul Hoque**[♣], **Jesse Vig**[♦]
[♦]Salesforce Research, [♠]Nanyang Technological University, Singapore
[♣]York University, Canada
[♣]{enamulh}@yorku.ca
[♦]{sjoty, jvig}salesforce.com

Abstract

Natural language and visualization (Vis) are two powerful modalities of human communication. The goal of this tutorial is to push forward the agenda of tightly integrating these two modalities. To this end, the tutorial will introduce NLP+Vis with a focus on two main threads of work: (*i*) *NLP for Vis:* How to develop and adapt state-of-the-art NLP models for solving various visualization tasks? and (*ii*) *Vis for NLP*: How to leverage visualization techniques to interpret and explain complex NLP models effectively? The tutorial will first motivate why NLP+Vis is an important area of research and provide an overview of research topics on combining NLP and Vis techniques. Then an overview of state-of-the-art deep learning models for NLP will be covered. Next, we will provide an overview of applying visualization techniques to help make NLP models more interpretable and explainable. In the final part, we will focus on various application tasks at the intersection of NLP and Vis. We will conclude with an interactive discussion of future challenges for NLP+Vis applications. The audience will include researchers interested in applying NLP for visualizations as well as others who focus more generally at the intersection of machine learning and visualization.

1 Tutorial Overview

Natural language and visualization are two powerful modalities of human communication. Visualizations (Vis) are pervasive as they frequently appear in research papers, textbooks, reports, news articles, and webpages in various forms such as charts, diagrams, and infographics. While visualizations can be very effective in finding patterns, trends, and outliers in data, natural language can help explain the key points in visualizations (Obeid and Hoque, 2020) and enable users to express their complex information needs about data naturally (Setlur et al., 2016). For example, recent work on Chart Question Answering (QA) has demonstrated how NLP techniques can reduce perceptual and cognitive efforts by automatically answering complex reasoning questions about charts (Kantharaj et al., 2022; Masry et al., 2022; Lee et al., 2022) or by generating natural language summaries from charts (Shankar et al., 2022; Obeid and Hoque, 2020). We also refer the interested readers to Prof. Marti Hearst's keynote (link) at IEEE Vis'22 on how NLP can help Visualization.

Likewise, visualizations also have critical applications in the NLP domain. For example, visualization techniques can be leveraged to interpret neural NLP models and to visually explain how a model makes a prediction (Chatzimparmpas et al., 2020; Belinkov and Glass, 2019; Li et al., 2016; Tenney et al., 2020; Strobelt et al., 2018; Vig, 2019), and more recently to design *prompts* (i.e., natural language instructions accompanied with zero or few demonstrations) to effectively use large language models for zero-shot and few-shot task generalization (Strobelt et al., 2022).

The proposed tutorial will be aimed at those who would like to push forward the agenda of tightly integrating state-of-the-art NLP methods with visualizations. To this end, the tutorial aims to cover two primary topics of interest: (*i*) *NLP for Vis:* How to develop and adapt state-of-the-art NLP models for solving various visualization-related downstream tasks? (*ii*) *Vis for NLP*: How to leverage visualization techniques to interpret, explain and adapt complex NLP models effectively?

An overview of the tutorial is provided below:

• In the tutorial, we will first introduce the domain of NLP+Vis and provide an overview of various downstream tasks in this domain such as question answering with charts (e.g., Lee et al. (2022); Kantharaj et al. (2022); Masry et al. (2022)), science diagrams (Kembhavi et al., 2016), and infographics (Mathew et al., 2022), as well as natural language generation for visualizations (e.g., Shankar

Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: Tutorial Abstracts, pages 1–6
December 6-10, 2023

et al. (2022)) and text-to-chart (e.g., Wang et al. (2022)).

• Next, we will introduce the state-of-the-art deep learning methods from NLP which can be leveraged for solving various computational tasks for visualization research. In this part, we will cover topics such as Seq2Seq models, attentions and Transformers, pretraining and fine-tuning of large language models (e.g., GPT, BERT, BART, T5). We will also briefly cover emerging research in multimodal NLP (e.g., vision-language, data2NLP).

• Then, we will provide an overview of applying visualization techniques for making NLP models interpretable and explainable. In particular, we will cover how interactive visualization techniques can be leveraged to understand how the NLP model internally works and to explain how a specific prediction is made (Tenney et al., 2020; Wallace et al., 2019; Li et al., 2016; Spinner et al., 2019; Strobelt et al., 2018; Vig, 2019). We will also discuss the limitations and common pitfalls of applying visualization to model interpretability. Furthermore, we will cover how visualization techniques can be incorporated within interactive machine learning (Jiang et al., 2019) as well as prompt design (Strobelt et al., 2022) for zero-shot and few-shot generalization of large language models like GPT-3.

• In the final part, we will demonstrate applications of deep learning to NLP in the areas of visualizations including visual text analytics (Liu et al., 2018), chart question answering (e.g., Kantharaj et al. (2022); Masry et al. (2022)), conversational interfaces for visualizations (e.g., Hoque et al. (2017); Setlur et al. (2020)) and automatic data-driven story generation (e.g., Shi et al. (2020)). We will also cover NLP models for enhancing chart accessiblity and visualization literacy.

• The tutorial will conclude with an overview of future challenges in the domain of NLP+Vis.

The tutorial will facilitate interactive conversations with those who participate in person as well as those who will participate virtually. A website will host the details of the tutorial including slides and other resources such as suggested readings as well as web links to related datasets and code repositories.

1.1 Relevance to ACL Community

There are rapidly increasing research papers that are being published at the intersection of Vis and NLP, but to our knowledge, there has not been any tutorial at any ACL venues. We gave a related tutorial at the IEEE Vis 2022 conference. However, considering the target audience (visualization community), we restricted the content of that tutorial to *introductory* and the *NLP for Vis* topic only. In that sense, the scope of the proposed tutorial is much broader and covers mostly cutting-edge research. Given the growing interest in combining NLP and visualization and the recent advances in state-of-the-art deep learning techniques for NLP, we believe it is a very good time to arrange a tutorial on NLP+Vis.

1.2 Type of the Tutorial

Cutting-edge

1.3 Target Audience and Prerequisites

The tutorial will provide a gentle introduction to advanced deep learning models for NLP for solving various visualization-related tasks. Familiarity with Python (using numpy and PyTorch), Calculus, Linear Algebra, Basic Probability and Statistics and Machine Learning basics are expected.

While the primary target audience includes those interested in applying NLP techniques for visualization, the tutorial may be of interest to those who are more generally interested to work at the intersection of machine learning and visualization.

2 Outline: Tutorial Structure

2.1 Introduction [20 mins]

- What is NLP?
- What is Vis?
- Why NLP+Vis?
- An overview of research topics on combining NLP and Vis techniques
- An overview of the tutorial

2.2 NLP for Vis [70 mins]

- Encoder-decoder model
- Attention mechanism
- Transformer architecture

- Self-supervised learning (e.g., BERT, GPT, BART, T5)
- Applications (QA, Summarization, Dialog)
- Multi-modal deep learning
- Huggingface library

2.3 Coffee Break

2.4 Vis for NLP [25 mins]

- Intro to Vis for Interpretability
- Vis Tools and Use Cases
- Challenges and Limitations

2.5 NLP + Vis Applications [50 mins]

- Visual text analytics
- Natural language interfaces for visualizations
- ChartNLP (e.g., Chart question answering, Text2Chart)
- Natural language generation for visualization
- Automated data-driven storytelling
- NLP for chart accessibility
- NLP+Vis for inclusions (e.g., promote visualization Literacy)

2.6 Future Challenges [15 mins]

- Building benchmarks for training and evaluation
- Data annotation challenges
- Emerging applications

3 Breadth

30 - 40% of the tutorial materials will come from the work by the tutorial presenters, and the remaining 60 - 70% will come from other researchers' work.

4 Promoting Diversity and Inclusions

The tutorial integrates diversity and inclusion-related topics into the agenda. It is well-known that the lack of understanding of the important data aggravates inequalities in access to information among different user populations ranging from vulnerable and marginalized communities (e.g., refugees and indigenous communities) to people who face various physical and cognitive challenges (e.g., blindness, dementia, autism). For example, natural language can be helpful in improving chart accessibility (Sharif et al., 2022) and supporting novice users in exploring visualizations (Setlur et al., 2016). The tutorial will highlight possible application areas of NLP+Vis for promoting inclusions and diversity.

5 Instructors

Shafiq Joty[1] is a research director at Salesforce Research, and is also an Associate Professor (on leave) at NTU, Singapore. His work has primarily focused on developing language analysis tools and NLP applications. A significant part of his current research focuses on multilingual (machine translation, cross-lingual transfer), multimodal (visual-language learning, NLP+Vis, Code+NLP) NLP, interpretability and robustness of NLP models. His research contributed to 17 patents and more than 110 papers in top-tier NLP and ML conferences and journals including ACL, EMNLP, NAACL, NeurIPS, ICML, ICLR, CVPR, ECCV, ICCV, CL and JAIR. Shafiq served (or will serve) as a PC chair of SIGDIAL'23, an S/AC for ICLR-23, ACL'22, EMNLP'21, ACL'19-21, EMNLP'19, NAACL'21 and EACL'21 and an AE for ACL-RR. He gave tutorials at IEEE Vis'22, ACL'19, ICDM'18 and COLING'18, and taught deep learning for NLP,[2] a graduate-level NLP course, and an undergraduate NLP course at NTU.

Enamul Hoque[3] is an Associate Professor at York University where he directs the Intelligent Visualization Lab. Previously, he was a postdoctoral fellow in Computer Science at Stanford University. He received the Ph.D. degree in Computer Science from the University of British Columbia. His research focuses on combining information visualization and human-computer interaction with natural language processing to address the challenges of the information overload problem. Recently, he has worked on developing natural language interfaces for visualizations (e.g., (Hoque et al., 2017; Setlur et al., 2020)), automatic chart question answering (Kim et al., 2020; Kantharaj et al., 2022; Masry et al., 2022), chart retrieval (Hoque and Agrawala, 2019) and chart summarization (Shankar

[1] https://raihanjoty.github.io/

[2] https://ntunlpsg.github.io/ce7455_deep-nlp-20/

[3] https://www.yorku.ca/enamulh/

et al., 2022; Obeid and Hoque, 2020). He has also worked on developing visual text analytics to support the user's task of exploring and analyzing conversations (e.g., Hoque and Carenini (2014, 2015, 2016); Jasim et al. (2021)). Since his research is uniquely positioned at the intersection of information visualization, NLP, and HCI, he publishes at the major venues in each of these areas such as IEEE Vis, ACL, EMNLP, CHI, and UIST. He serves as an Area Chair for the ACL Rolling Review (2021-) and as a program committee member (2018-) for the IEEE Vis. He has also been teaching the graduate-level Information Visualization course at York University for the past 3 years.

Jesse Vig[4] is a lead research scientist at Salesforce Research working on NLP, explainable AI, and HCI. Much of his research has explored novel interpretability methods, ranging from causal analysis of language models (Vig et al., 2020) to attention interpretation in protein sequence models (Vig et al., 2021b). He developed the BertViz[5] (Vig, 2019) library for visualizing attention in Transformer models, as well as the SummVis (Vig et al., 2021a) and ProVis (Vig et al., 2021b) visualization tools. His work has appeared in NeurIPS, ICLR, IUI, UIST, ACL, NAACL, FAccT, and WWW, as well as the VISxAI and BlackBoxNLP workshops. Vig's research has been recognized with a Best Paper award at the Intelligent User Interfaces conference.

6 Audience Size

We expect 75 - 100 attendees. We gave a similar tutorial at the 2022 IEEE International Conference on Visualizations (Vis 2022), a top conference in data visualization. To the best of our knowledge, there were 600 - 800 attendees at that conference. The tutorials were run before the main conference. Despite this, our tutorial attracted a good number of attendees (~ 40).

7 Preferable Venues

Our preferable venues are in the following order: (*i*) ACL, (*ii*) EMNLP, and (*iii*) EACL

8 Technical Equipment

Projector and Internet access.

[4]

[5] https://github.com/jessevig/bertviz

9 Ethical Considerations

We have considered several ethical issues related to the topics of the tutorial. To respect the intellectual property of different dataset sources, we will only use publicly available charts that comply with their terms and conditions. To promote reproducibility, we will share the relevant code repositories and datasets. Finally, we will explain any possible misuse of techniques presented in the tutorial. In particular, we foresee one possible misuse of different models presented in the tutorial which is to spread misinformation. Currently, NLP model outputs tend to contain factual errors. Hence, if such model outputs are published without being corrected, they may mislead and misinform the general public.

References

Yonatan Belinkov and James Glass. 2019. Analysis Methods in Neural Language Processing: A Survey. *Transactions of the Association for Computational Linguistics*, 7:49–72.

Angelos Chatzimparmpas, Rafael M Martins, Ilir Jusufi, and Andreas Kerren. 2020. A survey of surveys on the use of visualization for interpreting machine learning models. *Information Visualization*, 19(3):207–233.

E Hoque and G Carenini. 2014. Convis: a visual text analytic system for exploring blog conversations. In *Proc. EuroVis*, pages 221–230.

Enamul Hoque and Maneesh Agrawala. 2019. Searching the visual style and structure of d3 visualizations. *IEEE transactions on visualization and computer graphics*, 26(1):1236–1245.

Enamul Hoque and Giuseppe Carenini. 2015. ConVisIT: Interactive topic modeling for exploring asynchronous online conversations. In *Proceedings of the 20th International Conference on Intelligent User Interfaces*, IUI '15, page 169–180, New York, NY, USA. Association for Computing Machinery.

Enamul Hoque and Giuseppe Carenini. 2016. MultiConvis: A visual text analytics system for exploring a collection of online conversations. In *Proc. IUI*, pages 96–107.

Enamul Hoque, Vidya Setlur, Melanie Tory, and Isaac Dykeman. 2017. Applying pragmatics principles for interaction with visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):309–318.

Mahmood Jasim, Enamul Hoque, Ali Sarvghad, and Narges Mahyar. 2021. Communitypulse: Facilitating community input analysis by surfacing hidden

insights, reflections, and priorities. In *Designing Interactive Systems Conference 2021*, pages 846–863.

Liu Jiang, Shixia Liu, and Changjian Chen. 2019. Recent research advances on interactive machine learning. *Journal of Visualization*, 22(2):401–417.

Shankar Kantharaj, Xuan Long Do, Rixie Tiffany Ko Leong, Jia Qing Tan, Enamul Hoque, and Shafiq Joty. 2022. Opencqa: Open-ended question answering with charts. In *Proceedings of EMNLP (to appear)*.

Aniruddha Kembhavi, Mike Salvato, Eric Kolve, Minjoon Seo, Hannaneh Hajishirzi, and Ali Farhadi. 2016. A diagram is worth a dozen images. In *European conference on computer vision*, pages 235–251. Springer.

Dae Hyun Kim, Enamul Hoque, and Maneesh Agrawala. 2020. Answering questions about charts and generating visual explanations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13.

Kenton Lee, Mandar Joshi, Iulia Turc, Hexiang Hu, Fangyu Liu, Julian Eisenschlos, Urvashi Khandelwal, Peter Shaw, Ming-Wei Chang, and Kristina Toutanova. 2022. Pix2struct: Screenshot parsing as pretraining for visual language understanding. *arXiv preprint arXiv:2210.03347*.

Jiwei Li, Xinlei Chen, Eduard Hovy, and Dan Jurafsky. 2016. Visualizing and understanding neural models in nlp. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 681–691.

Shixia Liu, Xiting Wang, Christopher Collins, Wenwen Dou, Fangxin Ouyang, Mennatallah El-Assady, Liu Jiang, and Daniel A Keim. 2018. Bridging text visualization and mining: A task-driven survey. *IEEE transactions on visualization and computer graphics*, 25(7):2482–2504.

Ahmed Masry, Do Long, Jia Qing Tan, Shafiq Joty, and Enamul Hoque. 2022. ChartQA: A benchmark for question answering about charts with visual and logical reasoning. In *Findings of the Association for Computational Linguistics: ACL 2022*, pages 2263–2279.

Minesh Mathew, Viraj Bagal, Rubèn Tito, Dimosthenis Karatzas, Ernest Valveny, and CV Jawahar. 2022. Infographicvqa. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 1697–1706.

Jason Obeid and Enamul Hoque. 2020. Chart-to-text: Generating natural language descriptions for charts by adapting the transformer model. In *Proceedings of the 13th International Conference on Natural Language Generation*, pages 138–147. Association for Computational Linguistics.

Vidya Setlur, Sarah E. Battersby, Melanie Tory, Rich Gossweiler, and Angel X. Chang. 2016. Eviza: A natural language interface for visual analysis. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, UIST 2016, pages 365–377, New York, NY, USA. ACM.

Vidya Setlur, Enamul Hoque, Dae Hyun Kim, and Angel X. Chang. 2020. Sneak pique: Exploring autocompletion as a data discovery scaffold for supporting visual analysis. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, UIST '20, page 966–978, New York, NY, USA. Association for Computing Machinery.

Kantharaj Shankar, Leong Rixie Tiffany Ko, Lin Xiang, Masry Ahmed, Thakkar Megh, Hoque Enamul, and Joty Shafiq. 2022. Chart-to-text: A large-scale benchmark for chart summarization. In *In Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL), 2022*.

Ather Sharif, Olivia H Wang, Alida T Muongchan, Katharina Reinecke, and Jacob O Wobbrock. 2022. Voxlens: Making online data visualizations accessible with an interactive javascript plug-in. In *CHI Conference on Human Factors in Computing Systems*, pages 1–19.

Danqing Shi, Xinyue Xu, Fuling Sun, Yang Shi, and Nan Cao. 2020. Calliope: Automatic visual data story generation from a spreadsheet. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):453–463.

Thilo Spinner, Udo Schlegel, Hanna Schäfer, and Mennatallah El-Assady. 2019. explainer: A visual analytics framework for interactive and explainable machine learning. *IEEE transactions on visualization and computer graphics*, 26(1):1064–1074.

Hendrik Strobelt, Sebastian Gehrmann, Hanspeter Pfister, and Alexander M. Rush. 2018. Lstmvis: A tool for visual analysis of hidden state dynamics in recurrent neural networks. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):667–676.

Hendrik Strobelt, Albert Webson, Victor Sanh, Benjamin Hoover, Johanna Beyer, Hanspeter Pfister, and Alexander M Rush. 2022. Interactive and visual prompt engineering for ad-hoc task adaptation with large language models. *IEEE transactions on visualization and computer graphics*.

Ian Tenney, James Wexler, Jasmijn Bastings, Tolga Bolukbasi, Andy Coenen, Sebastian Gehrmann, Ellen Jiang, Mahima Pushkarna, Carey Radebaugh, Emily Reif, and Ann Yuan. 2020. The language interpretability tool: Extensible, interactive visualizations and analysis for NLP models.

Jesse Vig. 2019. A multiscale visualization of attention in the transformer model. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 37–42,

Florence, Italy. Association for Computational Linguistics.

Jesse Vig, Sebastian Gehrmann, Yonatan Belinkov, Sharon Qian, Daniel Nevo, Yaron Singer, and Stuart Shieber. 2020. Investigating gender bias in language models using causal mediation analysis. In *Advances in Neural Information Processing Systems*, volume 33, pages 12388–12401. Curran Associates, Inc.

Jesse Vig, Wojciech Kryscinski, Karan Goel, and Nazneen Rajani. 2021a. SummVis: Interactive visual analysis of models, data, and evaluation for text summarization. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: System Demonstrations*, pages 150–158, Online. Association for Computational Linguistics.

Jesse Vig, Ali Madani, Lav R. Varshney, Caiming Xiong, richard socher, and Nazneen Rajani. 2021b. BERTology meets biology: Interpreting attention in protein language models. In *International Conference on Learning Representations*.

Eric Wallace, Jens Tuyls, Junlin Wang, Sanjay Subramanian, Matt Gardner, and Sameer Singh. 2019. Allennlp interpret: A framework for explaining predictions of nlp models. pages 7–12.

Yun Wang, Zhitao Hou, Leixian Shen, Tongshuang Wu, Jiaqi Wang, He Huang, Haidong Zhang, and Dongmei Zhang. 2022. Towards natural language-based visualization authoring. *IEEE Transactions on Visualization and Computer Graphics*.