hallisky@uw.edu — **Skyler Hallinan** — 360-286-5645
linkedin.com/in/skyler-hallinan skylerhallinan.com github.com/shallinan1

Research Summary

Experienced natural language processing researcher passionate about creatively enhancing and leveraging the capabilities of language models for meaningful problems, including commonsense reasoning, question answering, debiasing, and rewriting.

Education

University of Washington *Seattle, WA*
M.S. in Computer Science (GPA: 3.99) June 2023

University of Washington *Seattle, WA*
B.S. in Computer Science, Applied Math, and Bioengineering; Cum Laude (GPA: 3.86) June 2021

- Honors: Outstanding Senior Award (computer science), Levinson Emerging Scholar, Dean's Medal Nominee

Experience

xlab, Paul G. Allen School of Computer Science & Engineering *Sep 2020 – Present*
Undergraduate Researcher, Advisor: Yejin Choi Seattle, WA

Unified Style Transfer with Reinforcement Learning ∼ *in progress for EMNLP 2023*
- Led project on a novel reinforcement learning algorithm for natural language generation, which leverages massive off-policy exploration and filtering, followed by reward-quantitized online learning
- Demonstrated success of this method on *unified* style transfer: rewriting text from any arbitrary style to a target style

Improving QA Models with Knowledge Augmentation ∼ *EMNLP 2022, main conference*
- Introduced a model to dynamically generate commonsense knowledge, which improves downstream QA models
- Utilized reinforcement learning with PPO to optimize the knowledge-generation model for fluency and helpfulness, based on feedback from an inference QA model as the reward

Rewriting Toxic Text ∼ *in submission to ACL 2023*
- Led project on rewriting text to reduce toxicity by identifying and replacing toxic spans while preserving meaning
- Developed novel method to identify salient tokens to edit by comparing logits between a domain-adapted toxic and non-toxic denoising model

Misinfo Reaction Frames ∼ *ACL 2022, main conference*
- Utilized novel commonsense framework to assess propaganda and misinformation techniques in media based on reader reactions and improved upon existing baselines for misinformation detection

H2Lab, Paul G. Allen School of Computer Science & Engineering *Jun 2021 – Present*
Undergraduate Researcher, Advisor: Hannaneh Hajishirzi Seattle, WA

Biases in QA Datasets and Models
- Led project assessing the underlying biases contained in question answering (QA) models and corresponding datasets
- Demonstrated that many open-domain QA models/datasets are biased towards entities from US culture

Noah's ARK, Paul G. Allen School of Computer Science & Engineering *Sep 2020 – Jun 2021*
Undergraduate Researcher, Advisor: Noah Smith Seattle, WA

Identifying Subtle Biases and Microagressions in News Headlines
- Proposed a new commonsense framework of subtle biases by incorporating *environmental context*
- Utilized a generative language model to automatically identify subtle biases in news headlines

Bloomberg *Jun 2021 – Aug 2021*
Software Engineering Intern, BLAW Search New York, NY

Zero Shot Tax-Type Classification
- Developed and implemented two novel, zero-shot classification approaches for short text using LEGAL-BERT, achieving 67% performance on the test set (accuracy) compared to human annotators

Featured Publications

[a.2] J. Liu, **S. Hallinan**, X. Lu, P. He, S. Welleck, H. Hajishirzi, and Y. Choi, "Rainier: Reinforced Knowledge Introspector for Commonsense Question Answering", Proceedings of EMNLP (Long), 2022

[a.1] S. Gabriel, **S. Hallinan**, M. Sap, P. Nguyen, F. Roesner, E. Choi, and Y. Choi, "Misinfo Reaction Frames: Reasoning about Readers Reactions to News Headlines", Proceedings of ACL (Long), 2022

Preprints

[p.1] **S. Hallinan**, A. Liu, Y. Choi, M. Sap, "Detoxifying Text with MaRCo: Controllable Revision with Experts and Anti-Experts", in submission to ACL 2023 (Short)

Skills

Languages: Python, Java, C/C++, LaTeX, HTML/CSS/JS; **Frameworks:** PyTorch, HuggingFace