Ivan Montero

www.ivamon.com | ivanspmontero@gmail.com

Education

University of Washington *Seattle, WA*

M.Sc. Computer Science *Sept. 2021 - June 2022*

- GPA: 3.98
- Advisor: Noah A. Smith

University of Washington *Seattle, WA*

B.Sc. Computer Science *Sept. 2017 - June 2021*

- GPA: 3.90
- Advisor: Noah A. Smith

Publications

Michael Hassid, Hao Peng, Daniel Rotem, Jungo Kasai, **Ivan Montero**, Noah A. Smith, Roy Schwartz, "How Much Does Attention Actually Attend? Questioning the Importance of Attention in Pretrained Transformers", In *Findings of the 2022 Conference on Empirical Methods in Natural Language Processing* (Findings of EMNLP), 2022. URL `https://arxiv.org/abs/2211.03495/`

Ivan Montero, Nikolaos Pappas, Noah A. Smith, "Sentence Bottleneck Autoencoders from Transformer Language Models", In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing* (EMNLP), 2021. Oral presentation. URL `https://aclanthology.org/2021.emnlp-main.137/`

Ivan Montero, Shayne Longpre, Ni Lao, Andrew J. Frank, Christopher DuBois, "Pivot Through English: Reliably Answering Multilingual Questions without Document Retrieval", In *Proceedings of the Workshop on Multilingual Information Access at the 2022 Conference on Empirical Methods in Natural Language Processing* (MIA @ EMNLP), 2022. URL `https://aclanthology.org/2022.mia-1.3/`

Florian Mai, Nikolaos Pappas, **Ivan Montero**, Noah A. Smith, James Henderson, "Plug and Play Autoencoders for Conditional Text Generation", In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing* (EMNLP), 2020. URL `https://www.aclweb.org/anthology/2020.emnlp-main.491/`

Professional Experience

Sept. 2022 - Present	**Machine Learning Engineer**, Apple Question answering for Siri.
June 2021 - Sept. 2021	**Research Intern**, Apple Open-domain question answering improvements through document-level representation learning.
Sept. 2020 - Dec. 2020	**Software Engineering Intern**, Facebook Image understanding improvements to Photo Search on the Visual Search Relevance team.
June 2020 - Sept. 2020	**Software Engineering Intern**, Google Embedding retrieval optimizations on the Machine Learning Google Research team.
March 2020 - June 2020	**Research Intern**, Apple Pivot Through English: Reliably Answering Multilingual Questions without Document Retrieval
Sept. 2019 - March 2020	**Teaching Assistant**, University of Washington Machine Learning (2x), Deep Learning, Natural Language Processing, and Computer Vision
June 2019 - Sept. 2019	**Software Engineering Intern**, Google Street View Billboard Detection And Physical Metric Inference on the Ads team.
June 2018 - Sept. 2018	**Engineering Practicum Intern**, Google Image Clustering Pipeline design/implementation on Image Understanding Google Research.

Teaching Experience

Spring 2022	**Computer Vision**, Teaching Assistant
Winter 2022	**Natural Language Processing**, Teaching Assistant
Autumn 2021	**Deep Learning**, Teaching Assistant
Winter 2020	**Machine Learning**, Teaching Assistant
Autumn 2019	**Machine Learning**, Teaching Assistant
Spring 2019	**Software Design and Implementation**, Teaching Assistant

Research Experience

University of Washington – Noah's ARK *Seattle, WA*

Advisor: Noah Smith, Mentors: Nikolaos Pappas (2019-2021), Hao Peng & Jungo Kasai (2021) *Aug. 2019 - Present*

- **Efficient Attention Distillation** (2021)
  We modify the knowledge distillation framework, which learns a smaller student model from a larger teacher that achieves the same performance, to experiment with efficient linear attention variants in the student by explicitly matching the standard quadratic attention distribution of the teacher.
- **Multilingual Embeddings from Monolingual Pretrained Transformers** (2021)
  Explore using a fixed English BERT model with a new trainable embedding table to perform masked language modeling in a non-English language, and explore the extents of English representation transferability to other languages
- **Sentence Bottleneck Autoencoders from Transformer Language Models** (2021)
  Explore the construction of a sentence-level autoencoder from a pretrained, frozen transformer language model. The sentence representations discovered by our model achieve better quality than previous methods that extract representations from pretrained transformers on single-sentence similarity, generation, and classification tasks.
- **Sequence Generation with Learnable Continuous Outputs** (2020)
  Explore a sequence generation model with learnable target continuous outputs which leverages a word autoencoder to a avoid the computationally expensive softmax prediction layer.
- **Plug and Play Autoencoders for Conditional Text Generation** (2020)
  Explore a sequence-to-sequence framework that learns a task-specific continuous mapping between the latent representations of sequence autoencoders. Our pre-training of autoencoders reduces transfer learning for other NLP tasks to simply learning a continuous translation, leading to up to four times faster evaluation and more parameter-efficient training.

Apple – Siri Web Answers *Seattle, WA*

Advisor: Chris DuBois, Mentors: Shayne Longpre (2020), Ni Lao (2021) *Aug. 2019 - Sept. 2021*

- **Unsupervised Representation Learning for Web-Scale Document Retrieval** (2021)
  Open-Domain Question Answering improvements through document-level representation learning. Explored phrase-level and contextualized exact methods to improve semantic retrieval.
- **Pivot Through English: Reliably Answering Multilingual Questions without Document Retrieval** (2020)
  Perform research experiments on the most effective, unified manner to reliably transfer knowledge from English question answering systems to lower resource languages by leveraging multilingual paraphrase detection.

Seattle Children's Research Institute *Seattle, WA*

Advisor: Peter J. Myler, Mentor: Aakash Sur *Sept. 2018 - May 2020*

- **Recognizing Base J from Single Molecule Real Time (SMRT) Sequencing** (2018)
  Explore machine learning and signal processing methods to construct a genome-wide mapping of modified bases in infectious organisms from polymerase pauses during sequencing. Presented at the UW's 22nd Annual Undergraduate Research Symposium.

Awards, Fellowships, & Grants

2020	**John and JoAnne Wisniewski Endowed Scholarship**, University of Washington
2019	**Microsoft Endowed Scholarship**, Microsoft
2018	**Washington State Opportunity Scholarship**, WSOS
2017	**Paul G. Allen School Direct Admission**, University of Washington
	Edward Jones Maple Valley Scholarship, Edward Jones
	Public School Employee Union Scholarship, Tahoma School District

Service

2019 **UW Research Computing Club**, Undergraduate Liaison *Seattle, WA*
2018 **UW HCDE Alternative Spring Break**, Instructor *Neah Bay, WA*
2017 **Washington Trails Association**, Trail Maintenance Volunteer *Seattle, WA*

Miscellaneous

- **Languages:** Native proficiency in English. Limited working proficiency in Chinese and Spanish.