Laura (Wendlandt) Burdick

Research Statement

My research interests lie broadly in the field of Natural Language Processing (NLP). I am particularly interested in word embeddings and multilingual research, with wider interests in computational social science, machine learning techniques, and multimodal problems combining vision and language.

Experience

"Black-box" methods are becoming increasingly common in NLP and machine learning. These methods have clear inputs and outputs, but the mapping between inputs and outputs is complex and not well understood. One example is algorithms that create word embeddings, vectors that capture useful linguistic (semantic and syntactic) information about individual words. My dissertation focuses on understanding the limits and properties of word embedding algorithms, as well as how these affect usage in downstream applications.

We introduced the metric of stability, which measures the consistency of local neighborhoods across two or more embedding spaces. We showed that common English word embedding spaces are surprisingly unstable, even across simple variations such as the change in random initialization of an algorithm. We additionally extended this work to over 100 languages, where we considered how linguistic typology is related to stability. Finally, we showed how instability affects downstream tasks, such as part-of-speech tagging and word similarity.

In addition to my thesis work, I have also contributed to other research projects in the areas of NLP, computational social science, and data science. I have particular experience with multimodal research combining insights from text, images, and video. I was the lead researcher in an interdisciplinary collaboration with psychologists at the University of Texas at Austin, where we looked at how an individual's personality is expressed through images and text, resulting in two publications. Moving beyond images to videos, I contributed to another project studying how to identify human actions in online videos using both textual and visual cues. Finally, I led a research group for the DARPA-funded AIDA project, where we built tools for understanding and representing multimodal input. I was the lead author on two publications resulting from this work, and I represented our team at two PI meetings.

My work has been published in both top NLP venues (NAACL, ACL), as well as interdisciplinary venues (SocInfo, Cognitive Computation). On the NLP side, my 2018 NAACL paper on instability in common word embedding spaces has been well-received, with over twenty citations in under two years. One interdisciplinary venue that I particularly enjoyed publishing in was Bloomberg Data for Good Exchange, where I worked with the Michigan Data Science Team and the U-M University Musical Society to analyze customer data and interests in musical performances. In addition to publishing, I am active in the NLP community, both internationally and locally. I co-chaired the Student Research Workshop at NAACL 2019, and I am participating in a committee to select the first ACL equity director. Locally, I co-organized a2-dlearn, an Ann Arbor deep learning symposium, and served as a session facilitator at the Michigan AI symposium.

Goals

As faculty, I am excited to lead an undergraduate research group, incorporating students into every aspect of the research process. During my Ph.D., I was a research mentor

for four undergraduate students and three graduate students, and I believe that this experience will allow me to continue being a successful mentor. Currently, I am working with two students on a project that we expect to publish in Spring 2020. I have also had experience leading and designing Explore CS Research, a program to introduce undergraduates to research for the first time by pairing them with mentors and giving them the opportunity to complete a small research project.

My research will continue to focus on natural language processing and its applications. I am particularly interested in continuing multilingual research. I have previously explored how linguistic properties are related to the performance of word embeddings, but other NLP tools also have varying success rates for different languages. I am interested in exploring if linguistic properties of these languages helps explain why tools work well in some languages, but not others. On the application side, this knowledge would allow us to improve the performance of tools in low-resource languages. Research on this topic could be undertaken, with guidance, by third or fourth year undergraduates with previous machine learning experience. These projects would allow students to learn about different NLP tools, explore evaluation metrics and techniques, and learn how to do detailed comparisons between algorithms.

In addition to multilingual research, I am interested in questions that can be answered using social media data. These questions are often multidisciplinary in nature, and I look forward to collaborating with faculty from other departments. For instance, how are common idioms used on different social networking platforms? Are they used differently by people with different demographic characteristics, or in different locations? This is a project that a first or second year undergraduate could successfully work on. It would involve data collection, deploying pre-made machine learning tools, and data analysis, skills that can be developed before taking a machine learning class.

Mentoring Students

When mentoring students on research, I believe that it is important to help students develop the skills they need to successfully pursue a research career. These involve timely and effective communication habits, public speaking in front of both technical and non-technical audiences, and clear writing. While teaching students the technical skills involved in research, I would also ensure that they had opportunities to practice these communication skills.

Doing research with undergraduates is a form of teaching. A successful mentoring relationship is one where the mentor passes on important skills, both technical and non-technical, to a student, in order to prepare them for future success. Allowing students to develop as researchers is a key motivation of doing research as faculty, and a responsibility that I am excited to undertake.