Work Experience

June 2016 - Present **Research Scientist**, *FiscalNote*, Washington, D.C.
Developing a variety of applications on legislative and regulatory data using ML and NLP techniques. Projects have ranged from rapidly exploring new datasets to collaborating with Product team to prototype new features to creating production systems for serving out analytics.

Key Accomplishments:

- Designed and built the "NLP-Engine": a framework for applying text analysis with unified procedures for collecting, cleaning and storing the processed data, as well as, an API to support processing the data on remote servers. The system significantly reduced the overhead for the creating new modules (e.g a sentiment classifier) and sharing the analysis.
- Ran a project to create a Knowledge Graph for exploring connections between federal bills, laws and regulations.
  - Coordinated efforts of two additional engineers for the data ingestion and storage aspects of the project.
  - Designed methods to interact with the Neo4j database. Created a custom web visualization tool using Vis.js.
- Implemented a tool for search query creation and expansion using statistically generated key terms
- Created a system for Named Entity Extraction and Matching.
  - Supervised two interns on the initial subproject to evaluate how out-of-the-box SpaCy model performs on Congressional Bills and Reports.
  - Designed an algorithm for mapping extracted entities to canonical database entries; created a service to enable other team's to use this functionality.
  - Applied the system to build a prototype for recommending tags (people, organizations) based on users' free-form input notes.
- Led a small team in the creation of a custom taxonomy of topics for legislation/policy. Implemented a suite of methods for reviewing "key terms" associated with each topic using custom embedding models and outlier detection methods.
- Redesigned a key platform service that predicts the likelihood that a bill will pass and provides client-facing explanations. Created an online architecture to replace a legacy batch system, trained new models that are easy and efficient to maintain, and designed more intuitive prediction explanations.
- Experimented with using topic models and embedding clustering to identify key themes in comments on proposed federal regulations.

Summer 2015 **Data Science Intern**, *Khan Academy*, Mountain View, CA.

Summer 2014 **Data Engineering Intern**, *Pinterest*, San Francisco, CA.

Education

2012-2016 **BS in Computer Science, Carnegie Mellon University**, *Pittsburgh, PA.*
Minor: Language Technologies. Phi Beta Kappa

Research

Summer 2019 **BillSum: A corpus for automatic summarization of US legislation**, *Anastassia Kornilova and Vlad Eidelman*, Workshop on New Frontiers in Automatic Summarization at EMNLP.
Created a corpus for automatic summarization of US and California Legislation and established baselines.

Summer 2018 **Party matters: Enhancing legislative embeddings with author attributes for vote prediction**, *Anastassia Kornilova, Daniel Argyle and Vlad Eidelman*, In Proceedings of ACL.
Developed a novel neural algorithm for predicting how a legislator will vote on a bill that achieved a 4% boost in accuracy over the previous state-of-the-art.

Summer 2018 **How Predictable is Your State? Leveraging Lexical and Contextual Information for Predicting Legislative Floor Action at the State Level**, *Vlad Eidelman, Anastassia Kornilova and Daniel Argyle*, In Proceedings of COLING.
Presented methods for modeling the likelihood of a bill receiving floor action across all 50 states and D.C.

Teaching Experience

May 2017 - Spring 2018 **Teaching Assistant**, *Technology Education and Literacy in Schools*.
Assisted high school teacher of AP Computer Science Principles.

Spring 2016 **Teaching Assistant - Introduction to CyberPhysical System**, *Carnegie Mellon*, Pittsburgh, PA.
Graded exams, created homeworks and ran office hours for an upper-level computer science course

Technical skills

Python, SQL, Keras, Gensim, Scikit-Learn, Pandas, Elasticsearch, AWS, Git, LaTeX

(703) 899 1388 • anastassia.kornilova@gmail.com • akornilo.com