BILJANA ROLIH, PH.D.

MACHINE LEARNING ENGINEER

Full stack Machine Learning Engineer with a passion for Natural Language Processing and experience in developing and maintaining NLP and ML models in production.

EDUCATION

University of California, Los Angeles
PhD Computational Chemistry 2015

University of California, Santa Barbara
BS Physics 2006

SKILLS

RELEVANT SKILLS:
Machine Learning, Natural Language Processing, Productionalization of Models, Data analysis, Recommender Systems

SKILLED LANGUAGES/TOOLS:
Python, Docker, Pandas&Numpy, Language Models, BERT, NLTK, SpacY

PROFICIENT LANGUAGES/TOOLS:
Google Cloud Platform, Containerized development, Kubernetes, Google PubuSub, AWS (EC2 &RDS), SQL, Mathematica, Matlab/Octave, MongoDB

TEACHING EXPERIENCE:
Statistical Mechanics, Quantum Mechanics, Chemistry (basic)

billie.rolih@gmail.com
408-497-1159
San Francisco, CA
biljana-rolih
billiescodes

EXPERIENCE

Qordoba
San Francisco, CA
Machine Learning Engineer
Sept. 2018 to Mar. 2020

Key achievements:
- Built an in-house version of a toxic speech detector from proof-of-concept to production, with an accuracy and speed comparable to AutoML models
- Improved latency of services from 1s to 300ms for gender bias, formality, plain language, hate speech and liveliness services, making the Qordoba platform experience seamless
- Collaborated with the product team to create gender bias and formality services, based on aggregate geometries of word vectors, as well as a passive voice detector in English. The gender bias service was integral to the content for the majority of our users, and specifically praised by content teams at Twitter.

Main responsibilities:
- Worked closely with the product team to develop new NLP features for analyzing and improving content on the Qordoba platform
- Built proof-of-concept models for analyzing style and Brand Voice. Services that fall under Brand Voice are content liveliness, hate speech detection, gender bias, and formality of language.
- Collaborated with product to develop services that check for plain language and when appropriate suggest a rewrite. Plain language is a key requested service by our customers.
- Responsible for end-to-end pipeline of services, including deploying, productionalizing, and upkeep of machine learning microservices
- Improved code performance and scalability for optimal customer experience
- Developed content suggestions for all services, grammar, Brand Voice, etc. At Qordoba the Data Science team closely monitored and rapidly responded to customer feedback on word suggestion quality on a weekly basis
- Experienced in working with key NLP libraries, statistical and deep learning models, including SpacY, NLTK, language models, BERT, word embeddings (word2vec)

Coinbase
San Francisco, CA
Data Science Intern
June 2017 to Dec. 2017

- Analyzed key metrics in company paid acquisition program; optimized business metrics (e.g. cost of user acquisition) as well as usage metrics (e.g. user retention, engagement)
- Built end-to-end data pipelines for importing daily data on user demographics, sources and traffic. Additionally integrated the production-quality code to company platform
- Designed SQL tables and KPI dashboards for solving business problems related to user insights
- Optimized marketing and email campaign strategies for user engagement, revenue and acquisition; Obtained an average of 3% lift in user acquisitions from marketing campaigns.
- Delivered data insights on customer segmentations, and implemented lifetime value models

Metis
San Francisco, CA
Data Scientist Fellow
Jan. 2017 to Apr. 2017

I was a fellow in a 12 week immersive data science program, which covers topics in Statistics, Machine Learning, Programming, Communication, and Design. In that time I developed a variety of data science projects with business applications , some of which are outlined below:
- Building a diagnostic recommender for Medical Coders: A recommender system for simplifying the transcription of diagnostics, and medical services by medical coding professionals. Using topic modeling and data from patient diagnoses and medical chart events to create clusters in diagnostic space. These clusters were used to generate specific diagnostic code recommendations for subsequent medical records.
- Predicting mortality rates in the ICU based on diagnoses: Using the MIMIC III database on patients, explored factors that influence patient mortality during hospital stays. Developed a model for mortality during hospital stays based on chart data, and found reliable predictors for mortality in Intensive Care using only key diagnoses.
- Building a content-based recommender of New York Times articles: Applied topic modeling on select New York Times articles from 2016, to categorize into topic clusters from which to build a user interest map. Built a user interface where the user is prompted with several articles to choose from until a profile of interests can be built. The article choices are also used to recommend new articles to the user.

SF Brigade
San Francisco, CA
Volunteer Coder in the Data Science group
Sept. 2016 to Jan. 2017

- Performed data studies with real life impact in a team with other volunteers, passionate about improving the local community. Worked with the California Justice Department on uncovering potential latent bias in the treatment of juveniles detained at police stations across the State.

UCLA
Los Angeles, CA
Graduate Researcher and Teaching Assistant
Sept. 2009 to June 2014

- Research Project: Computational modeling of kinetic and thermodynamic properties of solid metal hydrides, intended for applications in hydrogen storage for fuel-cell vehicles
- Research experience in computational methods , Monte Carlo modeling methods, data mining in materials
- Teaching assistant within the Department of Chemistry
- Assisted in a range of theory and laboratory courses in physical chemistry, including two graduate courses on statistical mechanics.