

Anna Melnikov

(773) 647-5985 – anna.me@hey.com

Machine Learning engineer with a focus on NLP and big data. Experienced with leading projects, organizing work across remote and distributed teams with diverse areas of expertise. Focused on delivering high-value products with clear metrics to reach concrete business goals for organizations and their clients.

Highlights

- **Over 5 years of professional experience with Python**
- **Machine learning**
 - Pytorch, Tensorflow, sklearn
- **Data processing and analysis:** PySpark, Pandas, Dask
- **Cloud:**
 - AWS (S3, EMR, Lambda, Athena, Sagemaker, CloudWatch, and others)
 - Google Cloud Platform (Kubernetes Engine, Container Registry, Google Storage, Firestore, Stackdriver)
- **Containers:** Docker, Kubernetes
- **Data:** MongoDB, Athena (PrestoDB), Hive, PostgreSQL
- **Web:** Flask, async HTTP

Experience

Bazaarvoice, Chicago

March 2018 - October 2020

Staff Engineer, Machine Learning and NLP

- Lead developer of an application that uses pre-trained NLP models and rule-based algorithms to extract actionable information from hundreds of millions of product reviews across multiple languages. Insights from this application increase the holistic value of BV's SaaS platform
 - Built a Spark pipeline on AWS EMR that utilises Spacy NLP engine and RAKE keyword extraction algorithm to find key phrases to identify common themes in reviews
 - Contributed majority of unit and integration tests for the application, and used Docker, Google Cloud Build and Git to automate testing and deployment
 - Co-developed an AWS Lambda service to automate the triggering and monitoring of the application
- Core engineer on a team that built an ML-based application to categorize customer-generated product reviews based on criteria specified by clients, including whether they are fit to be published. This application provides an API that serves millions of requests per month, and makes close to real-time predictions on customer content, which has allowed Bazaarvoice to both significantly reduce cost and increase the speed of moderating reviews for BV's clients
 - Worked on preparing and analysing data, and training ML models (both Deep Learning and "Classic")
 - Combined ML models and heuristic-based algorithms to address specific challenges in making predictions on imbalanced and partially unlabelled data
 - Built key parts of the software stack that serves those models and algorithms

- Led deployment to Google Kubernetes Engine, after investigating multiple options for bringing the service to production
- Led the development of automated test and CI/CD pipeline. Found ways to simplify deployment process and reduce risk of failures in production
- Principal developer of a system that uses a bespoke Bi-LSTM model to map hundreds of millions of clients' products into BV's universal catalog. This system continues to play a crucial role in enabling new features of the SaaS platform, including those described above
 - With this system, we increased coverage of product categories to almost 90%, a nearly fivefold increase in coverage over previous (manual) mappings, at a fraction of the cost and time it took to make those mappings
 - This application is deployed at scale using Spark, AWS EMR, Amazon Sagemaker, and AWS Lambda
- Project lead for an internal toolkit to train deep learning or traditional ML models in Kubernetes
 - This toolkit makes it easier for software engineers and data scientists to run ML experiments, reduces technical barriers to using cloud hardware to develop ML algorithms, enables faster iteration, provides repeatable results, and simplifies path to production

AddStructure, Chicago (acquired by Bazaarvoice February 2018)

December 2016 - March 2018

Machine Learning/NLP Software Engineer

- Contributed key parts of a recurrent neural network (RNN) classifier for predicting sentiment of short documents that exceeds the accuracy of the previous production system by $\approx 7\%$
- Built an RNN classifier to detect if a review is representative of a given product, improving the accuracy over the previously used system by more than 5%
- Used Tensorflow and Spacy to build a system that analyses semantic and syntactic relations in reviews and question in order to extract product features referenced in reviews, and provides clients with quick summaries of customer content
- Using word vector embeddings and Approximate Nearest Neighbors, built a system that, given textual information about a product, found other closely related products in a retailer's inventory

Other skills

- **Other languages and technologies:** Experience with JavaScript, Kotlin, Swift, C and several other languages. Extensive experience with with Flask, gunicorn, Sanic, aiohttp
- **Bilingual** in English and Russian, proficient in German, conversational Persian

Education

University of Chicago, Chicago, Illinois, USA

Graduated with General Honors, BA Linguistics, 2015