Anna Orosz

+1 (215) 206-2244| aorosz@sas.upenn.edu | linkedin.com/in/anorosz | https://anna.orosz.pro

Education

University of Pennsylvania | **Bachelor's** degree in **Mathematics** and **Computer Science** |Philadelphia, PA *May 2021*

Relevant Coursework: Machine Learning, NLP, Algorithms & Data Structures, Unix/Linux, Python, Computer Architecture, Graph Theory and Algorithms, Abstract Algebra, Analysis, Game Theory, Calculus, Language and Automata, Economics

Professional Experience

LogMeIn, Inc. | **Machine Learning** Intern | Budapest, Hungary *January 2018 – July 2019*

- Developed award-winning **Natural Language Processing** project training QA systems using state-of-the-art Deep Neural Networks
  - Won 1st prize Tech Innovation + Audience favorite prize + 2nd place Best Business Value Innovation at *HackIn* hackathon
- Researched term weighting methods and class hierarchy models to facilitate email automation for a Top 5 Bank in India
- Enriched text corpus in multiple languages and ensured stable performance by measuring for *Nanorep*, *Bold360 AI* Software
- Specialized in **NLP Applied Research**, worked in *Python* with *Keras, TensorFlow, scikit-learn*

Facebook, Inc. | **Software Engineering** Intern | Menlo Park, CA *May 2017 – August 2017*

- Built internal client-library for next-generation tool `Hyperloop` – which we designed to complete bulk data transfers for my team's - the **Data Transfer Infrastructure** Department – clientele, a.k.a. for *Facebook*'s Data Science Division
- Remodeled internal tool `COPTA` - used to copy huge directories across *Hadoop Distributed File System (HDFS)* clusters and data centers - to depend on Hyperloop in addition to previous dependencies
- Engineered detailed *Scuba* tables for `Hyperloop` and `COPTA`
- Operated with multi-tenancy, network utilization, scheduling, cross-dc connection-pooling in *C++*

RapidMiner, Inc. | **Developer** Intern | Budapest, Hungary *May 2016 – August 2016*

- Used **Data Science** to program operators in `Studio` designed to process, analyze and alter data locally
- Designed Operators in `Radoop` - using *Spark* scripts - to process data on large scale with *Apache Hadoop*
- Repaired bugs and designed new features in *Java* while maintaining a stable infrastructure

Projects

QA System *November 2018 – March 2019*

Fine-tuned and published the trained *BERT* model as a service to build a Question Answering system that is able to act as a real-life operator for *LogMeIn*'s *Bold360 AI*'s clients. (won 3 prizes at regional Hackathon) **Python, TensorFlow**

Youtube Spam Detector *November 2017*

Classified comments as ham or spam with Naïve Bayes and SVM models by using TF-IDF transformation. **Python, scikit-learn**

County Election Predictor *September 2017*

Built Neural Network model to predict if Clinton or Trump wins in counties. **Python, Keras, TensorFlow**

Image Classification *May 2017*

Retrained a TensorFlow neural network Machine Learning model trained on the *ImageNet* Large Visual Recognition Challenge dataset to classify images containing the *UPenn* logo by populating training data through scripts. **Python, TensorFlow**

Wiki Parser *April 2017*

Formed complex graph models by utilizing in-links and out-links to create a directed graph representation of the Wikipedia network and used BFS, DFS, Dijkstra's and Kosaraju's algorithm to research the network. **Java**

Raspberry Pi Security System *May 2016*

Utilized motion detector to initiate the security system to take a picture and trigger email response system. **Python, Shell**.

Leadership Experience

The Daily Pennsylvanian: *Principal Machine Learning Engineer* at UPenn school newspaper's Analytics department

AΩE Engineering Sorority: *Mentor* to students in a professional engineering sorority established for girls' advancement in STEM

Tutoring Center @UPenn: *Tutor* to fellow undergraduates for Computer and Information Science & Mathematics college courses

Science is a Woman Thing: *Initiator & Organizer* for Hungary's 1st *STEM conference* for several hundred high school female students

Proficiencies and Passions

Programming Languages: Java, Python, C++, C, OCaml, Bash/Shell, HTML, Assembly Language

Platforms/Tools/Modules: Linux/Unix, Git/Mercurial, scikit-learn, Keras, TensorFlow, JupyterHub/Lab

Languages: German (Fluent), Hungarian (Native Speaker), French (Intermediate)

Interests: Machine Learning, Artificial Intelligence, Big Data, Open-Source

Activities & Hobbies: Tango and Swing dancing, riding Vespa's, horseback-riding, skating & skiing, cinematography