Saul Kohn, PhD

646-787-7602 | saul.aryeh.kohn@gmail.com | Philadelphia, PA
linkedin.com/in/saul-kohn | github.com/SaulAryehKohn | devpost.com/SaulAryehKohn

Experience

Deep Learning Research Scientist | *Proscia, Inc.* | Philadelphia, PA *08/2019 – Present*

- Implemented, developed, trained and analyzed Tensorflow models for detection and localization of skin, colon and prostate cancer in gigapixel microscope images. Several patent filings submitted.
- Led the development of multi-task neural networks that detected malignant with state-of-the-art accuracy. Abstract accepted by the proceedings of the European Society for Digital Pathology.
- Created and operationalized a weakly-supervised neural network that performed data quality control, removing artifacts such as pen ink and air bubbles from microscope slide images.

NLP Data Scientist | *The Vanguard Group* | Malvern, PA *08/2018 – 08/2019*

- Created a question-answering neural network using transfer learning and sentence embeddings. Launched the model as a chatbot backend on the highest-traffic page of the Vanguard website.
- Assisted in the transformation of the 1000+ person call center by building graph-based machine learning models that predicted the reason a client was calling, based on their web activity.
- Developed a Python library to parse millions of highly-unstructured emails into a SQL database. Analyzed the dataset using dynamic topic modeling, granting new insights into client behavior.

Fellow | *Insight Data Science* | New York, NY *06/2018 – 08/2018*

- Created a Python-based chatbot framework, combining named entity recognition and topic modeling to generate concise text summaries of chat threads, delivered in real time to the user.
- Integrated the tool with Slack using AWS and Flask. See github.com/SaulAryehKohn/robo-recall

PhD Researcher | *University of Pennsylvania* | Philadelphia, PA *08/2014 – 05/2018*

- Developed an open source Python library to map atmospheric density using public data from worldwide GPS beacons. See github.com/UPennEoR/radionopy
- Designed and executed a pipeline in SQL, Python and Bash to compress and analyze 100 TB of radio telescope readings. Quality assured the data using a custom-built convolutional neural net.
- Built-out a suite of digital signal processing tools to perform state estimation of telescope electronics.
- Managed the cluster used by our collaboration of 60 scientists at institutions worldwide.
- Author or co-author of more than 30 academic papers, with over 800 total citations.

Education

PhD in Physics & Astronomy, *University of Pennsylvania***,** Philadelphia, PA **2014 – 2018**
MPhys with Honors Astrophysics, *University of Edinburgh*, Edinburgh, UK **2009 – 2014**

Skills

Languages: Python, SQL, Bash | *Some Familiarity*: Golang, C, C++, Fortran, Java, Javascript, Solidity
Scientific Tools: numpy, scipy, scikit-learn, Tensorflow, pyTorch, pandas, NLTK, gensim, spaCy, Tableau
Development Techniques: AWS, Google Cloud, git, versioning, unit testing, continuous integration, Hadoop