YISHAI RASOWSKY
Portfolio | Github | Linkedin | 054 751 6040 | yishairasowsky@gmail.com

I made aliyah from the United States 14 years ago. I have 2 years of work experience in ML and NLP. The main Python libraries I use are: Pandas, NumPy, and Scikit-learn for basic ML; Spacy, RegEx, Gensim, and NLTK for NLP; and Tensorflow or PyTorch for neural networks. I speak Hebrew fluently.

SKILLS

- Developing and deploying ML models
- Analytics, Pandas, Visualization
- NLP, Scikit-Learn, Tensorflow, PyTorch

- Agile, AWS, MongoDB, Django
- Industrious, organized, clean coding
- Learning new technologies quickly

WORK EXPERIENCE

Data Scientist (Independent contractor), Shamaym, 2021-22

- Created NLP topic models using Pandas and Scikit-Learn to classify debriefs; precision 95% and recall 50%.
- Developed NLP text generation using OpenAI API for auto-completion and text summarization.
- Built from scratch Python web app modules using Flask; deployed to production server using Linux.

Data Scientist (Independent contractor), Streamline Verify, 2021-22

- Created and deployed NLP record linkage algorithm with Pandas and Scikit-Learn to deduplicate professional records of medical professionals, and achieved >95% completeness and homogeneity.
- Implemented resampling procedure using the bootstrap method and k-fold cross validation in order to optimally leverage a limited data set.
- Engaged in feature engineering and feature enrichment using Pandas and geocoder to enhance the business value of each insight to be gleaned from my data.

INTERNSHIPS

Data Scientist (Internship), Sefaria, 2020

- Created NLP multi-label classifier using Scikit-Learn to assign topic labels to text bilingual passages (ancient Hebrew and English), and achieved >95% precision, 36% recall. Designed specifically for high precision.
- Implemented hierarchical classification using a knowledge graph in order to enrich the insights gleaned from each parent topic and its children.
- Developed an end-to-end NLP pipeline using Pandas, Scikit-Learn, Scikit-MultiLearn, MatPlotLib, Seaborn, and Kubernetes in order to automatically provide the potential user with desired topic labels.

Python Developer (Independent contractor), Smrt, 2019

- Created NLP entity extractor using rule-based logic for extraction of key entities from PDF office lease contracts, which achieved accuracy of >90% for several entities and boosted precision and recall by 20%.
- Developed interactive confusion matrix using mplcursors and to evaluate statistics of predictions.
- Created the end-toend AI portion of the product's web app using Pandas, Scikit-learn, SpaCy, Amazon Web Services, MatPlotLib, and Seaborn to enable fine-resolution entity idenitifcation.

PERSONAL PROJECTS

- **BERT Classification**: Modules fine-tune BERT for SOTA NLP classification and named entity recognition (NER).
- **Neural Network**: Built from scratch without third party libraries. Displays superiority on non-linear dataset.
- **Quantum Machine Learning**: Coding and analysis of loss function for quantum circuits on iris classification.
- **Screen Capture**: Surveillance system that sends screenshots in real time via email and WhatsApp.
- **Story Illustrator**: PyPI package that produces slideshow with images and captions from user's input text.

EDUCATION

Honors BA Math and Physics, Amherst College, 2004-08

- Quantifying entanglement thesis; outstanding score on comprehensive math exam.

Advanced Talmud, Ohr Somayach, 2008-16

- Gemara with encyclopedic breadth and deep analysis; granted semicha.

CERTIFICATIONS

Data Science and Machine Learning Certificate, IBM, 2018

- Executed data-driven solutions to increase efficiency and accuracy.
- Created data regression models with visualization using predictive data modeling.

Programming Course, Udacity, 2018

- Achieved proficiency in Python, Numpy, Pandas, and Github.
- Implemented knowledge acquired in SQL, HTML, CSS, Javascript.