Nicolás Kuzak

(786) 282-3587 • nicokuzak95@gmail.com •
505 W 37th Street, Apt. 1907, New York, NY 10018

>> WORK EXPERIENCE

PriceWaterhouseCoopers (PwC)	New York City, NY
Machine Learning Scientist, Senior Associate, EmTech AI	July 2019 – Present
Data Scientist, Associate, AI Accelerator	October 2017 – June 2019

NLP
- Led 3 data science workstreams in $4 million engagement with a Fortune 100 consumer bank to parse global laws and regulations, find actionable obligations, and prove coverage of obligations within internal documents
- Created a CNN text classification model and supporting PowerBI visualization that contributed to ~30,000 hours of reduced manual labor needed to sift through internal bank documents
- Containerized text classification model pipeline to leverage Kubeflow Pipelines and shorten retraining time
- Parsed documents from global regulatory sources into a tree structure that depicts an inter- and intra-document hierarchy for simplification of manipulating and mapping regulations
- Designed a baseline cosine similarity system to assess coverage of federal regulations by large consumer bank
- Monitored, reviewed, and performed bi-weekly pull requests of Git repository for team of 10 data scientists

Computer Vision
- Deployed digit- and face-recognition models in Nao Robot for external use at conferences and lab visits

Anomaly Detection
- Placed second in the 2019 Airbus Hackathon by implementing clustering methods, LSTMs, and AutoEncoders to solve the unsupervised anomaly detection problem of finding irregularities in unidentified airplane sensors

MLOps
- Created a CI/CD pipeline on Azure DevOps for international PwC teams to leverage when productionizing models that creates an endpoint with certifications and API Keys
- Produced an instructional GitHub repo for hosting flask apps with security, load balancing, and authentication on Google Cloud Platform

SQL Generation
- Built a SQL Generator chatbot by joining a NetworkX graph interpretation of SQL with a client database to take in a user's English request and translate it to SQL, return reports, and perform analysis
- Managed 2 offshore data scientists and collaborated with international user experience team to containerize the application for a working live demo at the JPMorgan Healthcare Conference

Miscellaneous ML Innovation
- Applied AnyLogistix to create various economic supply chain simulations of COVID-19 disruption
- Assembled a docker container and surveyed the usability of AutoKeras for the Emerging Technology team

>> LEADERSHIP

- Coordinated an internal Social Equity Hackathon that brought together 70 members of PwC's Emerging Technology group, Artificial Intelligence Lab, Diversity & Inclusion Council, and CEO Action Committee to act upon 10 datasets dealing with social inequity in education, voting rights, healthcare, and other domains
- Hosted 2 Lunch and Learn series (one on GitHub and one on academic papers) and numerous AI Demo Days to market models, results, teams, and algorithms to broader firm and upskill fellow employees
- Organized AI Academy, a full-day event with 8 lectures exploring AI techniques to further knowledge of team
- Taught ML classes for non-Emerging Tech at PwC and created GitHub markdown blogs for MLOps instructions
- Advised Carnegie Mellon University intern group that was implementing BERT for banks

>> EDUCATION

Cornell University College of Engineering	Ithaca, NY
GPA: 3.82/4.00 • Magna Cum Laude • Dean's List (all 8 semesters)	May 2017
Bachelor of Science in Operations Research Engineering • Dyson Business Minor for Engineers	

>> SKILLS AND INTERESTS

Programming: Python • Deep Learning • NLP • Docker • Kubernetes • (Kube + Air + ML + Git) flow • GCP • Azure
Data Science Libraries: sklearn • skorch • PyTorch • pandas • Keras • Flask • matplotlib • bs4 • numpy • nltk
Languages: Native fluency in oral and written English and Spanish
Certifications: Azure AI Fundamentals (AI-900)
Personal Interests: NYT Crossword & Spelling Bee • Boca Juniors • Travel • Guitar
Public Online Learning: Udacity Intro to (Machine Learning + Deep Learning) • Stanford (CS231n: CNNs for Visual Recognition + CS224n: NLP with Deep Learning)