Pengcheng Ding

(667) 803 0823 | pding@fastmail.com | linkedin.com/in/pchding | github.com/pchding| about.pding.dev

SKILLS

Programming Languages: Python, R, Git, SQL, Bash
Tools: scikit-learn, PyTorch, torchtext, spaCy, NLTK, Pandas, MLflow, Elasticsearch, Streamlit
Applications: machine learning (CRF, LDA, ensemble learning), deep learning (BERT, LSTM, CNN), game theory, stats/optimization (stochastic approximation/programming, bi-level optimization)

EXPERIENCE

Data Science Fellow, *Insight,* New York May 2020 - Aug 2020

- Helped Alphabet Health increase their clients' accessibility to Multiple Sclerosis literature by building keyphrases extraction models for literature abstracts
- Implemented LSTM and Bi-LSTM-CRF sequence tagging models for extraction. Utilized multiple embeddings for improving performance. Built Elasticsearch backend for storage and clio-lite based Streamlit front-end to conduct contextual search for certain key phrases
- Enabled the company to train extraction models on any search term on Pubmed and augment the abstracts with extracted key phrases through an end-to-end model implemented in MLflow. Decreased company's reliance on Power BI by deploying Kibana and Elasticsearch

Research Assistant, *Johns Hopkins University,* Baltimore Jul 2015- Present

- Simulated effects of grid charges on the electric market and found possible ameliorations to correct distortions from such charges with game theoretic optimization models. Utilized time series forecast and clustering techniques to generate input scenarios for the stochastic optimization model and solved the model on HPC cluster
- Collaborated with ISO-NE to improve unit commitment models to better account for forecasting uncertainties; Implemented hybrid robust and stochastic optimization framework; optimized performance for ISO system
- Organized and designed workshop to help graduate engineering students learn how to build power system optimization models in GAMS

Research Intern, *World Bank,* Washington, D.C. Apr 2018- Dec 2019

- Analyzed benefits of coordinated trading and system planning in Central Asia under different cooperation patterns in optimization model
- Utilized cooperative game theory to allocate system benefits in Python to ensure incentives for each country to cooperate; suggested cost allocations arrangements based on benefits received with further cooperation
- Mentored master students on modelling and game theory; guided and created tutorials for Bank employees to adapt model for other cases

PROJECTS

Autoencoder based recommendation engine for HackerNews post (pchding.github.io/HN_bubble)

- Recommended unusual posts that may contain new ideas or tools that are not well accepted yet under chosen topics to help user to break the filter bubble
- Categorized HackerNews posts combining LDA and BERT embeddings; created LSTM with attention and CNN based auto-encoder models to find unusual posts under each topic for recommendation

EDUCATION

Johns Hopkins University Expected Dec 2020
Ph.D. Environmental Engineering (Power market modelling and optimization)
École Polytechnique Jan 2014
M.S Energy and Environment
Nanjing University Jun 2012
B.S. Physics