Fei Liang

+1 (617)697-8732 | fl443@cornell.edu | https://www.linkedin.com/in/feiliang3196

OVERVIEW

Prospective Data Scientist with 2 years of hands-on experience working on various high impact projects on logistics optimization, customer perception analysis, weather forecasting, and image segmentation. Dedicated in delivering high quality outcome-oriented solutions to facilitate stakeholders making informed decisions.

EDUCATION

Cornell University, College of Engineering *Expected Dec. 2023*

Master of Engineering in Operations Research and Information Engineering, Data Analytics Ithaca, NY

- **GPA: 4.0**; Activities: ORIE M.Eng Student Leadership Committee; Teaching Assistant in Probability and Inference

Pennsylvania State University, Schreyer Honors College *Aug. 2018 - May. 2021*

B.S. in Computational Mathematics, Minor in Statistics, Minor in Information Science State College, PA

- **GPA: 4.0**; Activities: Vice President & Founder of Penn State Robo X club, 3rd place in Robomaster Competition

Related Coursework: Machine Learning, Database Management, Optimization, System Modeling and Design

Certificates and Awards: IBM Data Science Certificate, Dataiku Core Designer Certificate, MCM Honorable Award

SKILLS

Programming: Proficient in Python; familiar with ML and DL libraries like Scikit-learn, PyTorch, and TensorFlow

Developer Tools: SQL and NoSQL databases; Git version control; Bash scripts; Linux OS

Hard Skills: Machine Learning algorithms; Statistical Modeling; Data Visualization; Optimization Theorem; etc.

Soft Skills: Fast learning; Written and Oral Communication; Team collaboration; etc.

RELATED EXPERIENCE

Data Science Intern, Pitney Bowes *Jun. 2023 - present*

- **Optimization Algorithm**: Enhanced the efficiency of company's logistics system by developing an optimization algorithm in container shipping, which achieved a 15% improvement in delivery lateness.
- **Stochastic Simulation**: Guided facility operation upgrades by building a stochastic simulation model using Python (SimPy) to assess the performance of various algorithms under different scenarios.
- **Professional Communication**: Facilitated stakeholder and facility managers' decision making by devising a comprehensive improvement plan and communicating with stakeholders on a weekly basis.

Data Science Analyst, Johnson & Johnson & Cornell University Capstone Project *Dec. 2022 - May 2023*

- **Model Development**: Assessed thousands of comments by developing an NLP analysis model that drives the improvement company's prescription management platform, resulting in a 20% increase in employee satisfaction rate.
- **Data Pipeline Implementation**: Enhanced free-text data quality and reliability through segmentation, lemmatization, and tokenization by implementing a robust NLP pre-processing data pipeline using Amazon S3 and Dataiku.
- **NLP Algorithms Design**: Assessed employees' perception of the prescriptive management system by implementing sentiment, emotion, and topic analysis using NLP algorithms like BERT transformers, HDBSCAN, and LDA.
- **Insight Delivery**: Enabled stakeholders to make data-driven decisions by developing a Tableau data dashboard and devising a comprehensive analysis report to deliver the insights of employees' free-text feedbacks.

Deep Learning Researcher, Beijing Institute of Technology *Jan. 2022 - Jun. 2022*

- **Forecasting Platform**: Guided air quality researchers daily activities by building a nationwide real-time air pollution forecast platform, which achieved 90% accuracy in forecasting air pollutant density levels for a 12-hour window.
- **LSTM Model Development**: Collaborated with cross-functional teams to design and develop a Long Short-Term Memory (LSTM) model in Python PyTorch, facilitating informed decision-making based on real-time air quality data.
- **Dashboard Integration**: Enhanced data usability for stakeholders by embeding an interactive dashboard into the platform.

Machine Learning Engineer Intern, Anzhen Hospital *Sep. 2020 - Jul. 2021*

- **Image Segmentation**: Developed a U-Net deep neural network algorithm for cardiac scar segmentation using Python PyTorch, achieving 75% labeling accuracy and reducing doctors' diagnosis time by 30%.
- **Image Dataset Enhancement**: Applied computer vision pre-processing techniques, like gray scale normalization and data augmentation, boosting algorithm robustness and increasing overall segmentation accuracy by 20%.
- **Clinical Collaboration**: Worked closely with medical professionals and radiologists to adapt the algorithm for clinical use, enhancing workflow efficiency by 30% and improving patient care through faster, more accurate diagnoses.