An Jiang, Ph.D.

jiangan0808@gmail.com
801-960-0637
Lehi, UT 84043

Work History

Imagine Learning - Machine Learning Engineer | Data Scientist
Provo, United States
03/2018 - Current

- Gen AI Application Development:

- Developed and deployed a retrieval augmented generation (RAG) Q & A Slackbot with OpenAI GPT models using LangChain and Elasticsearch vector database, enhancing internal knowledge sharing and documentation access within the organization.
- Developed an innovative web-based essay writing assistant utilizing React for the frontend and integrating GPT models for AI-driven, real-time writing feedback and prompt engineering. Designed to significantly enhance students' essay writing skills through instant and continuous feedback mechanisms.
- Led the organization of a Gen AI workshop during the company's hack week, aimed at educating employees on Large Language Models (LLMs) and Generative AI via a series of Jupyter Notebooks. Developed a curriculum on AWS Bedrock and Hugging Face, emphasizing prompt engineering, decoding methods, and LangChain agents, covering text summarization, RAG, chatbot creation, stable diffusion, and code generation.

- Learning Analytics & Sparse Factor Analysis:

- Led the research and development of Sparse Factor Analysis on student test responses to uncover underlying capabilities and curriculum design alignment, utilizing Maximum Likelihood Estimation (MLE), Bayesian statistics and Markov Chain Monte Carlo (MCMC) sampling to develop robust statistical models.
- Translated complex statistical models into an efficient Python toolkit for educators, facilitating easy application of SFA in learning analytics to inform teaching strategies and curriculum improvements, ultimately aiming to provide a more personalized learning experience for students.

- Predictive analytics application development:

- Executed a comprehensive customer churn analysis across multiple data domains using SQL in Snowflake, driving retention strategies with a predictive modeling pipeline. Identified key churn factors with SHAP (SHapley Additive exPlanations), significantly reducing churn rates.
- Spearheaded an NLP analysis project on customer support chats, incorporating sentiment analysis, key phrase and entity extraction, and topic modeling, initially utilizing AWS Comprehend and Lambda for an event-driven serverless architecture. This evolved into a Streamlit application hosted on Snowflake, where I further enhanced the project by training and deploying custom machine learning models using Snowpark User-Defined Functions (UDFs).

- AI in Competitive Gaming:

Skills

- **Probabilistic Modeling techniques** : Factor Analysis, Bayesian Analysis, Markov Chain Monte Carlo (MCMC), Optimization
- **Generative AI models** : GPT, LLMs, Reinforcement Learning
- **NLP techniques** : Text Embeddings, Retrieval-Augmented Generation (RAG), Sentiment Analysis
- **Deep Learning Frameworks** : PyTorch, TensorFlow, Keras
- **Programming Languages** : Python, C++, C, MATLAB, Octave
- **NLP Tools and Libraries** : OpenAI, LangChain, spaCy, Gensim, Hugging Face Transformers
- **Data Science Libraries** : Numpy, SciPy, Pandas, Scikit-learn, Matplotlib, Seaborn
- **Cloud-based Platforms** : AWS , Snowflake
- **Others** : SQL, Flask, Git, Jira

Education

08/2016
Brigham Young University
Provo, UT
Ph.D.: Mathematics

- Advanced study in stochastic calculus, option pricing theory, differential equations, dynamical systems, and numerical analysis of stochastic partial differential equations (PDEs).
- Enhanced Black-Scholes-Merton model by incorporating stochastic interest rates to address real-world market

- Developed non-linear reward functions for AWS Deep Racer competition that clinched Hack-to-the-Future award at Imagine Learning.
- Trained AI agents to excel in video games, and employing a variety of neural network architectures (CNN, RNN, LSTM) for reinforcement learning using PyTorch.

Imagine Learning - Junior Data Scientist
Provo, United States
11/2017 - 03/2018
Developed machine learning models to forecast student performance on Texas Assessments of Academic Readiness (STAAR), demonstrating positive impact of Imagine Learning's educational products. Utilizing Python libraries such as Numpy, SciPy, Pandas, Matplotlib, Seaborn, XGBoost, and LightGBM, this project involved sophisticated feature engineering methods, including imputation, one-hot encoding, and data normalization, alongside time series cross-validation and grid search for hyper-parameter optimization. Subsequently, this model was integrated into Flask web application, enabling teachers to proactively identify and support students at risk, thereby enhancing educational outcomes through targeted intervention.

Goldman Sachs - Client Data Analyst
Salt Lake City, United States
10/2016 - 09/2017
Participated in remediation project for Private Wealth Management Division, focusing on regulatory compliance with US PATRIOT Act. This involved conducting detailed reconciliations and audits of PWM accounts. Analyzed over 5,000 complex domestic and international client accounts, including those associated with hedge funds, mutual funds, and private investment vehicles. Additionally, conducted enhanced due diligence for accounts in Onshore China, EMEA, Asia, as well as politically exposed and high-risk accounts to fulfill Anti-Money Laundering (AML) requirements, ensuring comprehensive compliance and risk management.

Additional Information

- Awards: Hack-to-the-future award; AWS Deep Racer competition champion - Imagine Learning
- SOA actuary exams passed: Probability, Financial Mathematics
- Language: Chinese - Native, English - Full professional proficiency

complexities.
- Derived PDEs for American spread option pricing under variable interest rates, offering an analytical solution.
- Implemented finite difference methods for efficient numerical computation of options' PDE.
- Validated model accuracy through extensive Monte Carlo simulations, employing regression-based, Longstaff and Schwartz, and dual methods.
- Demonstrated improved accuracy and practical relevance of the PDE-based option pricing model in dynamic financial markets, providing insights for risk management and investment strategies.

06/2011
Sichuan University
Chengdu, China
Bachelor of Science: Mathematics
Enrolled in Wu Yuzhang Honors College.