WINSTON WANG

mail@winston-wang.com | (312) 292-7535 | winston-wang.com | linkedin.com/in/win-wang

SUMMARY

Curiosity-driven Data Scientist with 3+ years of experience, achieved a 45% increase in predictive model accuracy for TOGO Car-sharing, and led the development of a SQL Generator project at AbbVie, enhancing data accessibility for non-technical stakeholders. Expertise in Python, machine learning, recommender systems, and artificial intelligence algorithms utilized to significantly improve data-driven decision-making and operational efficiency across healthcare, e-commerce, and transportation sectors. Keen to harness data science to enhance customer experiences and operational strategies, seeking to contribute to forward-thinking teams.

EXPERIENCE

Machine Learning Scientist **Jul 2022 – Present**
TOGO Car-sharing Chicago, IL

- Streamlined interviews with business partners to redefine the 'Car and flows distribution issue,' which led to the development of a machine learning model and performance metrics designed to address operational realities.
- Conducted a comprehensive assessment of machine learning algorithms (SVM, ARIMA, KNN) and deep learning models (CNN) to determine the most effective approach, leading to a 20% improvement in revenue forecasting accuracy.
- Pioneered the development and application of a cutting-edge deep learning model, fusing 1D CNN with LSTM, resulting in a remarkable reduction in RMSE by one to two orders of magnitude and outperforming the accuracy of the open-source model by approximately 45%.
- Developed and launched a robust MLOps dashboard to continuously monitor model performance and input and output metrics. Streamlined concept drift and data drift detection, resulting in a 40% reduction in model retraining time.
- Proposed an innovative architecture combining M-GCN and LSTM models, enhanced with an attention mechanism that could revolutionize traffic management approaches.

Data Scientist **Aug 2021 – Jul 2022**
TUST Biochemical Processes & Tech Lab Chicago, IL

- Implemented genomic prediction for *Aspergillus niger* using unsupervised algorithms integrated with Hidden Markov Model (HMM) and Hidden Semi-Markov Model (HSMM), enhancing gene identification efficiency.
- Achieved over 80% reduction in erroneous gene notation, providing clear directions for genetic engineering research through combined prediction methods.
- Altered research approach to increase target gene prediction accuracy by 90%, resulting in considerable time and budget savings and directing research toward actionable genetic engineering insights.

Junior Data Scientist **Jun 2021 – Aug 2021**
AbbVie Chicago, IL

- Engineered a user-centric SQL Generator and Data Retriever tool, simplified access for users unfamiliar with Optum/Truven datasets to easily retrieve needed data; reduced dependency on data engineers by 80%.
- Orchestrated end-to-end development of an advanced pipeline, merging real-world medical and prescription claim data into a user-friendly interface.
- Leveraged A/B testing to identify and implement the optimal interface, enhancing user experience and driving data accessibility and insights. Resulting in a 60% reduction in data processing time.
- Built the prototype that verifies the feasibility of this tool's technical route and capabilities. This tool could save $1 billion in license fees per year for the company once done.
- Collaborated on crafting the ML algorithm for extracting insights from patients' comorbidity and medicine usage, enabling healthcare providers to identify high-risk patients with 94% accuracy.

Lead Data Scientist **Jan 2021 – Jun 2021**
InstaHub Philadelphia, PA

- Crafted the development of a strategic blueprint for the department, emphasizing AI-powered energy waste detection in building HVAC operations, guided development priorities, established standards and protocols, and directed the project lifecycle from inception to completion.
- Implemented Extended Kalman Filter and Markov-based recurrent neural networks for occupancy prediction, tailored to user requirements. This innovation led to an average of 20% reduction in energy usage across pilot buildings, showcasing effective energy management through predictive analytics.
- Collaborated with cross-functional teams to design and deploy an analytics dashboard for real-time monitoring and analysis of heat transfer data, improving operational efficiency by 25% and enabling proactive maintenance strategies.

EDUCATION

Master of Science in Data Science; Computational Methods **Jun 2022**
DePaul University Chicago, IL

- Awards: Graduate with Distinction (GPA: 3.98/4.0)
- Leadership: Data Science Group (President)

Master of Science in Biology; Cellular and Molecular Biology **Dec 2019**
Illinois Institute of Technology Chicago, IL

Bachelor of Engineering in Bioengineering; Biochemical Engineering **Jun 2014**
Tianjin University of Science and Technology Tianjin, China

- Awards: The 8th "Challenge Cup" Fosun National College Students Business Plan Competition - Bronze Prize
- Project manager of a National Undergraduate Training Program for Innovation and Entrepreneurship

SKILLS

Programming and Software Development

- Python, R, SQL, Java, Perl
- Git, Docker, Nginx, WordPress
- Waterfall, Agile Development
- Microsoft Project

Data Analysis and Data Visualization

- Pandas, NumPy, JAGS
- Tableau, PowerBI, Matplotlib, Seaborn, ggplot2, Qlik Sense

Machine Learning and Artificial Intelligence

- SciKit Learn, TensorFlow, PyTorch, LongChain, MLX
- Supervised Learning, Unsupervised Learning, Recommender System, Reinforcement Learning
- Hyperparameter Tuning, Model Performance Analysis
- Convolutional Neural Network (CNN), Recurrent Neural Network (RNN)
- Transfer Learning, Natural Language Processing (NLP), Large Language Model Finetuning

Big Data Technologies

- Mango DB, Apache Hive
- Databricks, Apache Hadoop, Pig, Storm, Spark

Cloud and OS

- Amazon Web Services (AWS), Azure, Snowflake, Linux