Hye Won Hwang

+1 872 888 5568 | nicole.hw.hwang@gmail.com

www.linkedin.com/in/hye-won-nicole-hwang | www.github.com/hyeniii | hyeniii.github.io | www.huggingface.co/hyewonee96

EDUCATION

M.S. Machine Learning and Data Science (formerly MSiA), *Northwestern University, Illinois, USA* 12/2023
- **Relevant Courses:** Scalable Software Architecture, Cloud Engineering, Analytics for Big Data, Text Analytics (NLP), Deep Learning, Data Mining, Predictive Analytics, Social Network Analysis, Databases and Info Retrieval, Data Warehousing
- **Competition:** 7th place at the nationwide Humana Mays Case Competition

B.S. & M.S. Biotechnology, *Yonsei University, Seoul, South Korea* 08/2021
- **Published Paper:** Hye Won Hwang, et al. "A genomics-based semi-rational approach for expanding the postbiotic potential of collagen peptides using *Lactobacillaceae*." Journal of Agricultural and Food Chemistry 70.27 (2022): 8365-8376.

EXPERIENCES

Data Scientist | Opportunity AI | Remote, USA 02/2024 – current
- Enhanced neural network (NN) training pipelines on AWS SageMaker by integrating feature reduction techniques and Bayesian hyperparameter tuning, improving model efficiency and performance.
- Automated data loading and preprocessing processes from S3 to Athena tables using AWS Glue, ensuring efficient and scalable data management for large datasets.
- Utilize Jira to track and share work progress with teammates, ensuring transparency, collaboration, and efficient project management.
- Collaborate directly with SVP to streamline communication of complex modeling processes, enabling clearer insights into model outputs tailored for client-facing reports.

Data Science Intern | TransUnion | Chicago, IL, USA 06/2023 – 08/2023
- Utilized Spark to query data from a 1-billion user database, conducting extensive data analysis and preprocessing to ensure data quality for model ingestion.
- Developed a recommendation engine, boosting performance metrics by 5-20% across clusters compared to previous models.
- Designed modular python pipelines, enabling automation using bash scripts from data pre-processing to model evaluation.
- Implemented segmentation strategies to delineate customer cohorts, enhancing model performance.

Data Science Consultant | Honda R&D | Northwestern University Practicum 10/2022 – 06/2023
- Developed and optimized cost prediction models, reducing mean absolute error by 95% for more accurate financial planning.
- Collaborated with cross-functional teams to identify and resolve misalignments between end-users and senior leadership, ensuring effective model implementation.

PROJECTS

Sheldon-bot | github.com/hyeniii/sheldon-bot 2023
Developed an artificial intelligence (AI) chatbot mimicking Sheldon Cooper's conversational style from "The Big Bang Theory," leveraging DialoGPT for personalized dialogue generation, tracking model performance metrics using TensorBoard.

Property Price Predictor & Description Generator | github.com/hyeniii/aws-mlops 2023
Created a Streamlit web application for estimating rental prices and generating AI-based property listings, utilizing AWS serverless architecture (AWS Lambda, API Gateway, Step Functions) for scalability and cost-efficiency, along with automated ML model training and generative AI for content creation.

Phished | github.com/hyeniii/phishing-webpage-detection 2023
Developed a web application that detects phishing URLs using machine learning model, implemented in React.js with a FastAPI backend, and supported by AWS services (ECS, ECR, S3) for deployment and storage.

Nutri-AI | github.com/Nutri-AI 2022
Developed end-to-end Flutter-based mobile application through transfer learning of object detection model (YOLO v3) on over 30,000 images for meal logging and supplement recommendation. The application features a seamless integration of Docker for containerization, deployed on AWS EC2 with a FastAPI backend and Flutter frontend for an enhanced user experience. It employs DynamoDB (NoSQL) for data management, offering a comprehensive solution for nutritional tracking and guidance.

LANGUAGE AND SOFTWARE SKILLS

Programming: Python, SQL, JavaScript, Bash
Packages: PyTorch, Scikit-learn, Pandas, Numpy, Matplotlib, LangChain, MLflow, HuggingFace
Cloud Computing: AWS (S3, RDS, DynamoDB, Lambda, Step Functions, ECS, EC2, SageMaker, Athena)
Certification: AWS Certified Cloud Practitioner (2023)
Software: Spark, Docker, Git, Fast-API, Node.js, Express, React.js, Tableau, MySQL