Hye Won Hwang

+1 872 888 5568 | nicole.hw.hwang@gmail.com | linkedin.com/in/hye-won-nicole-hwang | github.com/hyeniii | hyeniii.github.io

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 Engineer | Opportunity AI | Evanston, USA 02/2024 – current

- Coordinate and act as a **key intermediary** between technical teams to understand and streamline the data pipeline, ensuring seamless integration from data preparation to neural network (NN) model application.
- Spearheaded the design and development of advanced reporting metrics to enhance model performance visibility and experimental outcomes.
- Collaborate closely with data scientists in the **experimentation and evaluation of NN models in SageMaker**, contributing to the refinement of modeling strategies and experiment design.

Data Science Intern | TransUnion | Chicago, USA 06/2023 – 08/2023

- Utilized **Spark** to efficiently query data from a **+1 billion user database** and conducted extensive data analysis and pre-processing operations to ensure data quality and accuracy.
- Developed a **multiclass classification model** to recommend the user's next purchase, **increasing forecasting metrics by 5~20%** across various clusters relative to the predecessor model.
- Streamlined data pipelines with modular components for full automation from data preprocessing to evaluation.
- Implemented segmentation strategies to delineate distinct customer cohorts to enhance model efficiency.

Data Science Consultant | Honda R&D | Northwestern University Practicum 10/2022 – 06/2023

- Achieved a significant reduction in error, cutting **mean absolute error by ~95%** against the predecessor model.
- Developed and optimized cost prediction models for enhanced future cost and budget planning.
- **Collaborated closely with cross-functional teams**, leveraging diverse expertise and perspectives, to successfully present strategic initiatives to key stakeholder leadership, driving alignment and fostering innovation.

PROJECTS

Sheldon-bot| github.com/hyeniii/sheldon-bot 2023

Developed an artificial intelligence 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, enhancing web security through effective phishing detection.

Nutri-AI | github.com/Nutri-AI 2022

Developed a Flutter-based mobile application utilizing a fine-tuned 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, Unix Shell Scripting, Java
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

s