Neil John D. Ortega

neiljdo | neiljdo

neiljohn.ortega@gmail.com
+81-80-4629-2192

Skills and Competencies

Soft Skills: Strong work ethic, Strong verbal and written communication skills, Leadership and teamwork skills, Highly adaptable, Problem solver.
Machine Learning and Deep Learning: AutoGluon, AWS SageMaker, PyTorch, fastai, Keras, Tensorflow, scikit-learn, nltk, pandas, numpy, OpenCV, etc.
Big Data: Spark, Airflow, Redis, Kafka, Hive, Hadoop, etc.
Full-Stack Development: fastAPI, OpenAPI, Django, django-rest-framework, Flask, Shiny, D3.js, Front-end frameworks, etc.
DevOps: Docker, Kubernetes, Drone, Prometheus, Grafana, Jenkins, Ansible, AWS, etc.
Scripting: Python, R, SQL, Bash, etc.
Collaboration and Project Management: Git, JIRA, Confluence, Agile methodology, Scrum, etc.

Experience

Data Scientist

Nov 2021 — Present

AI Cross Inc. *Tokyo, Japan*

- **Machine Learning Platform-as-a-Service**
  - Currently designing and implementing the core ML module of the platform that would allow users to train models and retrieve inference results via APIs.
- **Business Marketing Platform-as-a-Service**
  - Developed prototype for generating and delivering AI-based recommendations in 1 month, allowing the team to pitch and present to potential clients and validate the value provided by the service.
  - Created the system architecture for the product based on the initial prototype.
  - Implemented data ingestion, model training, and recommendations generation components of the system, working across the entire stack (frontend, backend, ML module).
  - Constantly communicated with the sales team to learn the client feedback and re-prioritize the development milestones accordingly.
  - Implemented A/B testing feature that would allow businesses to experiment with their strategies and choose the best option within statistical bounds.
  - Took the charge and self-learned about UI/UX to further improve the design of the platform.
  - Collaborated with the CTO to create the pipeline for operationalizing the system.

Machine Learning Engineer

Sep 2019 — Nov 2021

LINE Fukuoka Corporation *Fukuoka, Japan*

- **Detection of High-Profile Individuals from Caricatures**
  - Benchmarked the performance of the existing tool we wanted to improve upon using Python scripts.
  - Developed a detection approach based on the WarpGAN paper.
  - Implemented Spark, Airflow batch-processing pipeline that handles 1M image/s throughput.
  - Developed and deployed APIs for real-time inference and visualization of batch detection results.
  - Added service monitoring via Prometheus and Grafana to tighten the feedback loop in detecting target events.
  - Wrote the entire deployment pipeline that allowed for faster iteration.
- **Content Suggestion via Image-Based Emotion Recognition**
  - Coordinated with product team to scope the requirement.
  - Designed model and solution architecture based on existing literature on emotion recognition, recommendation systems.
  - Trained baseline models for recognizing emotion from input images.
  - Developed and deployed APIs for real-time inference.
  - Implemented service monitoring via Prometheus and Grafana.
- **Machine Learning Tooling and Knowledge-Sharing Tasks**
  - Worked with other members of the ML team to improve existing and/or create new ML/MLOps tools.

- Presented to the team the following areas of interest: meta-learning, model quantization 🔗, auto-ML 🔗, causal inference 🔗, secure AI 🔗, and explainable AI 🔗.

Data Scientist Nov 2016 — Sep 2019
Teradata Corporation *Manila, Philippines*

- **Data Extraction from Scanned Documents**
  - Trained random forest models and CNNs to improve the accuracy of predicting the location of fields of interest in documents.
  - Developed ad-hoc scripts for rule-based detection and extraction of field values from documents.
  - Improved image input to OCR service using OpenCV image preprocessing scripts.
  - Extracted training data from scanned images using Tesseract and Microsoft OCR.
- **Survey Response Classification**
  - Trained LSTM + word-embedding models using Keras with Tensorflow and scikit-learn for the classification of survey responses to easily diagnose problems the end-users are facing.
  - Developed the end-to-end (E2E) solution architecture, closely collaborating with principal data scientist.
  - Created an app for data ingestion and model results visualization using R, Python, and SQL that reduced manual labor and errors done by client.
  - Streamlined both local and production deployment by writing a Docker pipeline.
  - Hosted learning sessions for technologies the rest of the team are not familiar with.
- **Customer Behavior Modeling**
  - Operationalized the models developed by data scientists via an ad-hoc parallelized Python pipeline.
  - Tracked and fixed issues with the client's infrastructure, providing support to the client's Cloudera Cluster administrators.
  - Benchmarked and optimized performance for both Hive data store and Teradata Aster compute engine that reduced E2E processing time from several days to 2-3 hours.
- **Model Monitoring and Maintenance**
  - Created a Spark, Shiny application that allows users to run a scoring model against datasets in HDFS and visualize model health metrics.
  - Determined high-value tasks and scoped the project, closely working with the client's analytics team.
  - Deployed the entire solution via Ansible and Cloudera Manager on the customer's AWS cluster.
  - Solution streamlined the customer's processes and reduced E2E processing time from 1 week to 1 hour.

Senior Full-Stack Engineer Sep 2011 — Nov 2016
icannhas Inc. *Manila, Philippines*

Education

Master of Science in Physics 2021 — Present
University of the Philippines Diliman, Quezon City. Expected completion: 2024

Bachelor of Science in Civil Engineering, *magna cum laude* 2006 — 2011
University of the Philippines Diliman, Quezon City
Awardee, Engineering Top 100 in Academics, 2011 (out of 5000 students)
10th Place, ACM International Collegiate Programming Contest Asia Manila Regional Contest, 2009
Honorable Mention, 37th International Physics Olympiad, 2006

Certificates

2022 Qiskit Global Summer School on Quantum Simulations 2022
Online course. Expected completion: July 29, 2022

2021 Qiskit Global Summer School on Quantum Machine Learning 2021
Online course.