Neil John D. Ortega

linkedin.com/neiljdo | github.com/neiljdo

neiljohn.ortega@gmail.com
+81 080-4629-2192

Skills and Competencies

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

Experience

Machine Learning Engineer

Sep 2019 — Present

LINE Fukuoka Corporation

- **Detection of High-Profile Individuals from Caricatures**
  - Wrote benchmarking scripts to determine the accuracy of the tool we want to improve.
  - Experimented with an approach based on 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.
- **Emotion-Based Content Suggestion**
  - Coordinated with product team to scope the requirement.
  - Read existing literature on emotion recognition, recommendation systems to design model and solution architectures.
  - Trained baseline models for emotion recognition.
  - Developed and deployed APIs for real-time inference.
  - Implemented service monitoring via Prometheus and Grafana.
- **Machine Learning Team Tools**
  - Added new features to in-house Ansible, Airflow, and Kubernetes-based MLOps tools.
- **Knowledge-Sharing Tasks**
  - Made two presentations on meta-learning papers and one on model quantization techniques.

Data Scientist

Nov 2016 — Sep 2019

Teradata Corporation

- **Data Extraction from Scanned Documents**
  - Developed ad-hoc rule-based detection and extraction of field values from documents.
  - Implemented image preprocessing scripts using OpenCV to improve image input to OCR service.
  - Wrote scripts using Tesseract and Microsoft OCR to extract information from scanned images.
  - Used random forest and CNNs for predicting the location of fields of interest in documents to improve accuracy.
- **Survey Response Classification**
  - Collaborated with principal data scientist in developing the solution architecture.
  - Created app for data ingestion and model results visualization using R, Python, and SQL that reduced manual labor and errors done by client.
  - Trained LSTM + word-embedding models for the classification task using Keras with Tensorflow and scikit-learn.
  - Created a Docker pipeline for the team for efficient deployment to both local and production environments.
  - Acted as a domain expert for technologies the rest of the team are not familiar with.

- **Customer Behavior Modeling**
  - Created a parallelized Python pipeline to operationalize the model developed by data scientists.
  - Coordinated with client's Cloudera Cluster administrators in tracking and debugging issues with their infrastructure.
  - Benchmarked and optimized performance for both Hive data store and Teradata Aster compute engine that reduced end-to-end (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.
  - Worked with the client's analytics team to properly scope the project and to determine the pain points of their workflow for improvement.
  - Implemented dashboards for visualizing model health metrics.
  - 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.
- **Real-Time Social Media Discovery**
  - Created real-time sentiment analysis application made using Apache NiFi, Kafka, Spark Streaming, and SparkML to showcase the capabilities of the Analytics Team in a conference.
- **Internal Consultant Directory**
  - Wrote API for internal directory of consultants connected to a recommender system meant for improving resource allocation and management.
- **On-boarding Materials**
  - Trained members of the ETL team with the basics of Apache NiFi, Kafka, Storm, and Cloudera environment.

Senior Full-Stack Engineer Sep 2011 — Nov 2016

icannhas Inc.

- **Stock Market Buy-or-Sell Recommender**
  - Implemented near real-time prediction backend using scikit-learn and Django.
  - Implemented frontend using plot.ly and ReactJS.
- **Consulting and (Internal) Product R&D**
  - Worked on multiple web-based applications for both local and international clients.
  - Coordinated with stakeholders in scoping projects, prioritization of features, etc.
  - Coordinated with project team members in budget estimation, scheduling, etc.
  - Implemented backend solutions mainly in Python, Django, django-celery, django-rest-framework, etc.
  - Implemented frontend solutions using HTML5, CSS3, JS, and miscellaneous JS frameworks.

Education

Bachelor of Science in Civil Engineering, *magna cum laude* 2006 — 2011

University of the Philippines, Diliman, Quezon City

Activities and Organizations

University of the Philippines Tennis Society

Secretary-General for Special Events 2009 — 2010

- Headed committee responsible for teaching tennis to future and current members and in handling the logistics of events.
- Spearheaded annual university-wide tournament hosted by the organization. Lead the entire organization of 50+ members, coordinated with sponsors and participants to ensure the success of the annual event.