

# Neil John D. Ortega

[linkedin.com/neiljdo](https://www.linkedin.com/in/neiljdo) | [github.com/neiljdo](https://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**

- Internal tool for identifying high-profile individuals from highly-stylized versions of their images.
- Wrote benchmarking scripts to determine the accuracy of the existing tool.
- Experimented with an approach based on WarpGAN paper.
- Developed API for the final solution.
- Automated and optimized deployment pipeline.
- Increased processing speed by 14x using Redis-based optimizations.
- Implemented Spark, Airflow batch-processing pipeline to increase throughput even further.
- Created and deployed web-based UI to show detection results.

- **Emotion-Based Content Suggestion**

- Internal tool for suggesting content based on detected emotion from user photos.
- Read existing literature on emotion recognition, recommendation systems to design model and solution architectures.
- Trained baseline models for emotion recognition.
- Developed API for the solution.
- Automated and optimized deployment pipeline.
- Automated load testing of API.

- **Machine Learning Team Tools**

- Miscellaneous internal libraries used by the ML team to make work easier.
- Added new features to in-house Ansible, Airflow, and Kubernetes-based MLOps tools.

- **Knowledge-Sharing Tasks**

- Quarterly presentation reviewing existing research, ideas, etc.
- to the entire ML team.
- Wrote 3 presentations that focus on meta-learning and model quantization techniques.

### Data Scientist

Nov 2016 – Sep 2019

*Teradata Corporation*

- **Data Extraction from Scanned Documents**

- Pipeline for automatic extraction of desired field values from scanned (images and PDF files) version of 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 ML and DL techniques for predicting the location of fields of interest in documents to improve accuracy.

- **Survey Response Classification**

- AI-based app for automated tagging of free-form customer survey responses to speed up market defect discovery.
- Participated in developing the solution architecture.
- Wrote backend and frontend code for data ingestion and model results visualization using R, Python, and SQL.
- Trained and tuned LSTM + word-embedding models for the classification task using Keras with Tensorflow and scikit-learn.
- Created a Docker-based 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**

- Hidden Markov model for predicting the behavior of telecom subscribers for location-based and time-based targeting.
- 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 parts of their workflow to be improved.
- Implemented dashboards for visualizing metrics indicating the health of the models.
- 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.

- **Load Demand Forecasting and Disaggregation**

- Created an interactive Shiny application with a custom map UI via Leaflet that allows users to visualize both actual and forecasted electrical load demand and disaggregation.
- Presented the solution on-site.

- **Real-Time Social Media Discovery**

- Real-time sentiment analysis application made using Apache NiFi, Kafka, Spark Streaming, and SparkML (Machine Learning).
- Implemented both backend and frontend, using NodeJS and MongoDB, and Covalent/Angular and in-house chart widgets, respectively.

- **Internal Consultant Directory**

- Custom internal directory for consultants connected to a recommender system meant for improving resource allocation and management.
- Wrote the REST API backend in Django and django-rest-framework.

- **On-Site Search Improvements**

- Integrated pandas and numpy into Flask backend to improve search performance.
- Updated C3.js and D3.js-based visualizations.

- **Plugin Development**

- Developed an internal reusable R plugin for creating customizable plots and widgets for Shiny applications.

- **On-boarding Materials**

- Created customized demo and exercise materials for Apache NiFi, Kafka, and Storm, integrating the three into a single streaming dataflow application in a Cloudera environment.
- Handled training of members from another department.

- **Stock Market Buy-or-Sell Recommender**

- Machine learning-based approach for predicting stock market trends and recommending buy or sell orders.
- 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.
- 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 of the organization. Coordinated with sponsors, org members, and participants for the entire event.