# Mark Bajo

Machine Learning Engineer


**Status:** Machine Learning Engineer at Ridge-i
**Fields:** Machine Learning, Computer Vision, LLMs
**Languages:** English, Japanese(N2), Filipino
**Activities:** Lifelong Learning, Competitive Programming, Mathematics

Tokyo, Japan
markbajo.io
markarjaybajo@gmail.com
+81 80 5936 7734


## Projects

### LLM - Detect AI Generated Text - github.com/mmbajo/LLM_detect

· Currently engaged in a Kaggle competition focused on classifying essays as LLM-generated, where I've achieved a score of 0.940 AUC, closely approaching the current top score of 0.964 AUC. Some of the task I did in this competition includes: dataset creation by prompting LLMs, creating text preprocessing functions, fine tuning LLMs using peft and LoRA etc.

· Made a simple application for a friend who works as a school teacher to help her discern if her students use ChatGPT on their homeworks. The model used in the application is efficient and can be run on a CPU, and scores around 0.932 AUC on the leaderboard.

Python | TensorFlow | HuggingFace | scikit-learn | pandas | numpy | XGBoost | lightgbm | Streamlit

### PROBA-V Super Resolution Challenge - github.com/mmbajo/PROBA-V

· Joined a competition hosted by European Space Agency wherein competitors are tasked to fuse multiple low resolution images and predict its high resolution version.

· Placed top 2 as of April 4, 2020. The other competitors are prestigious research laboratories around the world.

Python | TensorFlow | PyTorch | numpy | OpenCV | Linux

## Experience

### Machine Learning Engineer - Ridge-i — 2020 - Present

- Chick Health Anomaly Detection Using Images
  - Currently a member of a 7-person team developing a system to assess the health of chicks using images captured from four different angles. This project involves creating various models, including: chick segmentation model to estimate its weight, dirt segmentation model to estimate if the chick is clean or not, malformation segmentation models to detect different kinds of malformation.
  - Personally responsible for creating the model pipeline, building experiment tools for data collection, model training and experimentation, and dataset maintenance.
- Color Differential Model for Paint Coating Quality Control
  - Innovated a tailored color differential model to assist human operators in assessing incoming paint coatings by using spectral reflectance values, a scarcely researched area. The model employed a variety of pre-trained feature extractors and unique input feature representations, products of meticulous feature engineering.
  - Personally responsible for training and experimenting on model architecture and feature engineering techniques, maintaining datasets and experiment history, and building inference API services. Achieved 0.90 F1-score, allowing partial-automation on the production line.
- Optical Character Recognition (OCR) for VIN Decoding
  - Utilized easy-ocr architecture as the foundation for our OCR system and leveraged a Trie-based VIN verification system to validate prediction outputs and capture VIN prefix frequencies, meeting specific client requirements.
  - Personally responsible for developing model training/inference, and post-processing pipelines. Achieved an image accuracy rate of over 99.9%, exceeding client expectations.
- Rapid-prototyping, Product demo, and DevRules
  - Worked with scientists to turn ML prototypes, like 3D bin packing simulators and satellite-imagery object detection systems, into client-ready demos. Collaborated with sales to quickly build tech solutions, securing key contracts and bolstering market share.
  - Streamlined Project Onboarding: Created and managed template repositories, complete with GitHub Actions, Docker templates, issue/PR templates, and sample unit tests, simplifying project initiation for engineers.

Python | JavaScript | HTML | PyTorch | pytest | numpy | sklearn | pandas | Detectron2 | anomalib | Optuna | OpenCV | Flask | mlflow | Django | Docker | EC2 | AWS Video Stream | S3 | Spark | Airflow | Terraform | DVC | Git

**Research and Development Engineer** - Anritsu Infivis, Inc. 2015 - 2020

- Engineered Advanced Vibration Control for High-Precision Devices: Led the research and application of mathematical models to drastically reduce vibration noise by 80% in high-precision weighing systems, thereby cutting down response times by 50%.
- Spearheaded AI-Driven Contaminant Detection: Led the end-to-end creation and deployment of a YOLO-based machine learning model for contaminant identification in food and pharmaceuticals. Successfully validated the solution's commercial viability while outlining key product specifications.

Python | Matlab | LabView | C/C++ | ONNX | OpenVINO | CMake | TensorFlow | numpy | OpenCV | Git

## Education

**B.Sc. Mechanical Engineering** - University of the Philippines 2010 - 2015

· Achieved a high academic standing, graduating in the top 10 of my cohort from the University of the Philippines.

· Coursework: Applied Mathematics, Numerical Simulation, Statistics, System Identification, Robotics, Control Systems