Koyilbek Valiev

E-mail: valievkoyiljon112@gmail.com **Phone:** +(82)-010-2253-3010
Github: github.com/koyilbek **Website:** valiev-koyiljon.github.io/Web

Summary

AI Engineer with 2+ years of hands-on experience building vision-language systems and machine learning pipelines. Proven track record in video anomaly detection, time series forecasting, and multimodal retrieval using VLMs. Actively exploring VLM/LLM-based methods, retrieval-augmented generation, and Vision-Language-Action (VLA) models, with a growing interest in real-world robotics applications.

Experiences

Multimodal AI Engineer Intern *Pia Space, Seoul, South Korea*
March 2025 - mid June 2025

- **AI Video Anomaly Detection :** Collected and annotated diverse anomaly video datasets (e.g., violence, falls) using CVAT and proprietary annotation tools. Optimized lightweight VLMs (<1B params) for video anomaly classification with dynamic sampling and prompt engineering in few-/zero-shot settings. Conducted both qualitative and quantitative analysis to evaluate and refine model performance.
- **Multimodal Retrieval for Video Anomaly Detection:** Co-developed **MACS 3.0**, a prompt-driven multimodal system for CCTV anomaly detection, achieving an overall model performance improvement of approximately **24%** over previous versions. Built a benchmark dataset with manually captioned videos for video-text and video-video retrieval. Implemented and evaluated cross-modal retrieval models (video-to-text, text-to-video, video-to-video) to enhance scalable multimodal video understanding. Adapted a research method on multimodal explanation maps for vision-language models, generating visualizations that highlight key text prompt words and corresponding video regions to improve model interpretability and alignment.

AI Engineer | Lead Intern *Recs Innovation Ltd, Naju, South Korea*
Feb 2024 – March 2025

- **Bidding Amount Prediction:** Developed a dynamic ML pipeline that analyzes incoming data distributions and selects the best-performing algorithm. Implemented data cleaning, anomaly detection, feature engineering, and clustering, achieving a 0.0017% prediction error and a 15× increase in bid win rate.
- **Solar Power Forecasting:** Built a 48-hour ahead forecasting system based on a hybrid neural network combining CNN, LSTM, GRU, and Transformer layers to capture spatial, temporal, and long-range dependencies. Leveraged forecasted weather data to reach an average 6% prediction error. The solution was integrated and deployed on the Sun-Q EMS platform for real-time energy optimization.
- **Photovoltaic Sensor Anomaly Detection:** Led development of an unsupervised AI solution using sensor data from solar plants. Employed models including LSTM Autoencoder, LSTM-VAE, TranAD, and VAE to detect anomalies, achieving over 70% F1-score on real-world data. The solution is deployed in the Sun-Q EMS platform for continuous monitoring and fault detection.
- **Documentation & Team Leadership:** Authored comprehensive technical documentation on system architecture and data workflows. Managed a small AI team using Notion and ClickUp for project planning, task delegation, and progress tracking, enhancing team efficiency and delivery.

AI Research Intern *Capstone Project, Woosong University*
Sep 2023 - Dec 2023

- Coordinated team of 4 engineers to develop a real-time trash-bag optimization system for NetVision.
- Developed YOLOv8 to detect, classify, segment, track & size-estimate bags in video.
- Participated with this project in Woosong University's 2023 Capstone Competition with permission from NetVision, winning 1st place award.

AI Intern *Sequus PTY LTD, Australia (Remote)*
Jun 2023 - Aug 2023.

- Automated conversion of bounding-box outputs into YOLO, Pascal VOC, and custom formats.
- Annotated hundreds of architectural drawings (LabelImg) to improve model accuracy.

Projects: Refer portfolio koyilbek for more details and other projects

- ***Multimodal Effective Video Search and Interaction with RAG:*** Retrieval-augmented video Q&A using Whisper, LLaVA, BridgeTower, LanceDB, LangChain & Gradio. The project effectively integrated state-of-the-art retrieval techniques to enhance user interaction with video content. More details in: ***Project Link***
- **Neural Conversational Chatbot**: Built a seq2seq chatbot in PyTorch with a bidirectional GRU encoder and GRU decoder enhanced by Luong's global attention, trained on 220K+ Cornell Movie-Dialogs pairs. Optimized batching and hyperparameters (500 hidden units, 2 layers, 0.1 dropout) for efficient training, and added attention-weight visualization and a modular codebase. More details: ***Project Link***

Education

Woosong University
Bachelor's in AI & Big Data | *2021 – 2025*

Key courses: Data Visualization, Artificial Intelligence, Machine Learning Theory & Lab, Computer Vision, Statistics, Linear Algebra, Discrete Mathematics, Capstone Project

Skills

Python, PyTorch, Hugging Face Transformers, LangChain, LlamaIndex, FastAPI, OpenCV, Ollama, Unsloth, MLflow, Optuna, Git, NumPy, Pandas, scikit-learn, Matplotlib, Docker

Awards

- 1st Place, Woosong University Capstone Competition among 200 teams (*Dec 2023)*
- President's Award, Woosong University (*Dec 2023)*
- 100% Merit Scholarship, Woosong University (*Sep 2023 – Dec 2023)*
- 1st Place, IoT Learning Concert (*Sep 2023 – Dec 2023)*
- 2nd Place, Machine Learning Lab Learning Concert (*Sep 2023 – Dec 2023)*