Hyunwoo Joo

21.Mar.1992 (32, Male)

68, Hwagok-ro 58-gil,
Gangseo-gu, Seoul, Republic of Korea
(82) 10-8346-8299
stevepaulljobs@gmail.com

EXPERIENCE (6 years and 6 months)

DEEP.FINE, Seoul — *AI Engineer / Team lead*

Sep 2023 - PRESENT

[Head of Advanced Research Team]

- Led the adoption of Vision and Generative AI technologies, driving innovation in enterprise projects.

- Managed technical showcases at CES 2024, MWC, and other global exhibitions, strengthening the company's AI presence.

- Directed AI projects for government and corporate clients, ensuring seamless delivery and execution.

- Authored technical proposals and analyzed RFPs to secure competitive contracts.

- Planned GPU cloud support projects, delivering live demos and performance reports for stakeholders.

[MLOps Developer]

MLOps System for License Plate Recognition (YOLO & PaddleOCR)

***Objective*: Process 50,000–70,000 vehicle images daily with sub-2-second response time.**

- Asynchronous Pipeline

1. Built an OCR queue using Redis and Celery, achieving over 10x throughput improvement compared to single-request handling.
2. Automated background preprocessing to reduce user waiting time.

- Kubernetes Zero-Downtime Deployment

1. Configured YAML-based deployments with Canary Release strategies for real-time traffic distribution, minimizing downtime.
2. Ensured high availability with Ingress Controller and TorchServe load balancing, even in the event of a single-node failure.

- Performance Optimization

1. Reduced batch processing time to <200ms with parallel inference using `asyncio.gather`.

SKILLS

Python / C++ / Ruby
Pytorch / Transformers / TorchServe
FastAPI / Flask / Django
Docker / Docker-compose
Kubernetes / ArgoCD / Helm
SQL / NoSQL / SQLAlchemy / PGVector / Langchain
ELK (Elasticsearch / Logstash / Kibana) / Github Actions
Nginx / Redis / Celery / Kafka
Linux / MacOS
Shell Script
AWS / GCP /Azure
Typescript / Javascript
Vue.js / Express.
Streamlit / Netify / Vercel

ACTIVITIES

Weakly Spatial AI : Reviewing latest research and tech news related with Visual SLAM and Spatial AI (2024.09 ~ Present)

Reading Deep Learning papers : Led NLP's papers (2020.01 ~ 2023.04)

LANGUAGES

Native: Korean, Mandarin

Business level: English

2. Enhanced PaddleOCR with dynamic padding, increasing inference speed by 20%.

- **Monitoring & Reliability**

1. Deployed Prometheus + Grafana for real-time system monitoring; maintained <0.1% error rate and 95%+ SLA.

- **Key Contributions**

1. Fully developed Redis/Celery pipelines and batch inference logic.
2. Optimized PaddleOCR models and contributed to GUI/CI/CD setup.

[AI Developer]

Development of Multi-Layout OCR System

***Objective*: Enable AI OCR to automatically recognize thousands of document types with varying layouts, without requiring predefined templates.**

- **Dynamic Layout Text Detection**

1. Used TRACE to detect text out-boundary regions for robust text candidate identification across diverse form structures.
2. Refined detected regions with CRAFT to generate horizontal and vertical bounding boxes, achieving over 90% recognition accuracy without prior knowledge of templates.

- **TrOCR Recognizer Tuning and Inference Optimization**

1. Integrated special tokens into bounding box embeddings for unified sequence training, resulting in an F1 score above 90%.
2. Applied dynamic padding for transformer-based models, reducing GPU memory usage and optimizing inference speed.

- **LLM-Based Key-Value Matching System**

1. Fine-tuned Llama2-7B with LoRA to automatically convert recognized text into JSON format.
2. Employed Sentence Transformers to measure key embedding similarity, mapping specific fields (e.g., name, amount, date) within documents.
3. Validated and restructured JSON output into structured data, improving downstream analysis and database integration.

- **Business Impact**

1. Enabled automatic processing of 2,000–3,000 document types without predefined layouts, increasing OCR adoption by over 30%.
2. Improved recognition accuracy by 15% and processing speed by 5x compared to traditional form-based OCR systems.
3. Reduced end-to-end processing time (from scanning to database ingestion) from 20 minutes to under 2 minutes, significantly

cutting operational costs.

- **Key Contributions**

1. Designed the data pipeline, integrating TRACE and CRAFT detectors (50%).
2. Tuned the TrOCR model, optimizing transformer structures and inference speed (100%)
3. Developed a prototype for Llama2-7B with LoRA for key-value matching, integrated with Sentence Transformers (100%).

EO Studio, Seoul — *AI Engineer / Team member*

Aug 2022 - Aug 2023

[Platform Development]

Enhancing a Startup Community Platform

***Objective*: Increase user engagement and satisfaction by improving author recommendation and document search features.**

- **Author Recommendation System with Collaborative Filtering**

1. Designed recommendation metrics using tabular features (e.g., likes, views).
2. Improved click-through and subscription rates by recommending authors users are likely to enjoy, boosting retention by ~20% (Google Analytics).

- **Elasticsearch-Powered Search Engine**

1. Reduced average query response time to under 1 second, improving data discovery efficiency.
2. Enhanced user search satisfaction and reduced post-click bounce rates.

[LMS Platform Optimization]

***Objective*: Reduce churn rates and increase user engagement by providing real-time updates on lecture schedules and assignment deadlines.**

- **Real-Time Schedule Push Notification Service**

1. Developed Social Media/SMS integration streaming APIs for dynamic message templates (e.g., user name, course details) using Apache Kafka and Redis.
2. Increased attendance rates and reduced deadline inquiries by 30%.

[B2C AI Service Planning]

***Objective*: Validate and prototype an LLM-based backend automation service for early-stage startups.**

1. Conducted One-pager ad campaigns using Meta Ads and created landing pages with no-code tools.
2. Performed A/B testing, user persona definition, and PMF analysis during a 6-week campaign.
3. Analyzed campaign performance (CTR, conversion rates) to define MVP direction, achieving 100+ early adopters and setting the foundation for official service launch.

- Key Contributions

1. Designed automation scenarios and user stories leveraging LLM (100%).
2. Created ad content and managed Meta Ads campaigns, including targeting and creatives (100%).

Sysmetic, Seoul — *AI Engineer / Team Lead*

Oct 2020 - Aug 2022

[AI Team Lead]

- Oversaw technical reviews and development strategies for key in-house AI projects.

- Contributed to government grant proposals through documentation and presentations, securing project funding.

- Improved team collaboration processes, managed project timelines, and performed code reviews as a technical lead.

[News Recommendation System Development]

Collected large-scale text data from ~200 Korean and English news agencies for AI analysis.

***Objective*: Address limitations of existing RSS/simple parsing methods by developing a more precise crawler and text estimation algorithms.**

- Large-Scale Crawling Program

1. Built a Selenium-based crawler engine, parsing `robots.txt` and `sitemap.xml` for valid links and handling dynamic page rendering.
2. Prioritized SEO tags (OG Graph) to preprocess main content regions (title, description, etc.).
3. Designed a DOM tree BFS algorithm with weighted nodes to detect content patterns
    a. Learned patterns of classes and IDs likely to contain the main content, using text-to-HTML ratio for precise identification.
    b. Improved crawling efficiency by caching exploration times, node weights, and text lengths for faster revisits.
    c. Re-initialized weights and re-evaluated nodes when the

original text length decreased or content quality degraded.

- **Key Results**

1. Automated the collection of over 100,000 news articles daily in mixed Korean and English.
2. Improved article body extraction precision by ~10%, significantly reducing false positives.
3. Reduced crawling time per iteration by 30% through caching and optimized algorithms.

- **Key Contributions**

1. Designed the crawler architecture and implemented the BFS DOM analysis algorithm (100%).
2. Developed Selenium distributed execution, caching logic, and error handling for 404 and permission issues (50%).
3. Managed deployment, cron job scheduling, and log monitoring processes with the team (50%).

[NLP Development: Keyword Extraction, Document Summarization, and Entity Recognition]

***Objective 1*: Quickly extract key phrases and summaries from large-scale news articles for decision-making reports and search engines.**

***Objective 2*: Build an NER module to automatically recognize company names and stock tickers for financial and industrial analysis.**

- **Keyword Extraction and Extractive Summarization**

1. Fine-tuned a BERT model using SEO tags (title, description) as ground truth and generated interim labels with PageRank for missing tags.
2. Measured similarity between news headlines/descriptions and full text to extract top-k sentences for summarization.

- **BERT NER for Company Names and Tickers**

1. Cached major company names/tickers in a database and fine-tuned the model using similar tokens as ground truth.
2. Trained a multi-label classification model with NER datasets and added weights to surrounding entities, improving recall and reducing emissions.

- **Key Results**

1. Achieved 80–90% precision/recall for keyword extraction while reducing summary length by 20% without significant information loss.
2. Improved company/ticker recognition recall from ~70% to over 90%, enhancing financial domain analysis accuracy.
3. Enabled faster internal/external report generation and automated large-scale document processing, reducing resource costs.

4. Demonstrated results for government projects, contributing to additional funding and contracts.

- Key Contributions

1. Designed the fine-tuning pipeline for BERT (keyword extraction, NER) and implemented interim label generation with PageRank (100%).
2. Led model evaluation metrics (Precision/Recall/F1) and hyperparameter tuning (70%).
3. Expanded the multi-label structure (company names/tickers) and improved DB caching logic for entity recognition (100%).
4. Collaborated with QA and testing teams to refine outputs and deploy production-ready services (50%).

Good Intelligence, Seoul — *AI Engineer / Team member*

Oct 2019 - Oct 2020

[AI Developer]

- Designed and implemented an automated medical journal collection program.
- Developed a program to organize disease and drug-related terms and visualize topologies.
- Improved medical text analysis performance by fine-tuning a Bio-BERT-based model.

Ubiforum, Seoul — *Web Developer / Team Member*

Aug 2018 ~ Oct 2019

[Web Developer]

- Developed and maintained MES/ERP frontend systems using Angular.js.
- Designed and implemented SPC data visualization programs using Chart.js.
- Built MES/ERP functionalities and data processing logic using Java Spring.

EDUCATION

Sogang University, Seoul — Master of Engineering (Graduated)

Sep 2019 – Feb 2022

Major: AI & Data Science

Thesis : Tile-Aware BERTSUM for Korean Newspaper Extractive Summerization

Tsinghua University, Beijing, China — Bachelor of Arts (Graduated)

Sep 2011 – Jun 2016

Major: Chinese Language & Literature

Internship: *Newsis Korea-China Translation & Intern Journalist*

Sep 2014 – Jul 2015

Beijing No.80 Middle School, Beijing, China — High School Diploma (Graduated)

Sep 2008 – Jul 2011

PROJECTS

Naver Real Estate Image Review System — Image Inpainting

Mar 2023 – Mar 2024

- Developed a Kubernetes-based AI inference engine deployment system.
- Designed and implemented an inference pipeline using YOLO, SAM, and ViT for image inpainting.
- Optimized architecture for AWS S3 data integration tailored to a B2C web service.

Medical Journal Language Modeling System — BERT Pretraining

Mar 2022 – Nov 2022

- Built a pretraining dataset for medical journal language models.
- Developed a BERT-Base pretraining module, including embedding models and tokenizer training.
- Created a BERT-based downstream task program for SQuAD.

CERTIFICATIONS

Language: HSK Level 6 (275 points) – Feb 2015

Language: ACTFL OPIc IH Level – May 2018

Certification: Engineer Information Processing Certificate – Nov 2018

Certification: ADSP (Advanced Data Analytics Semi-Professional) – Jul 2020

Course: Coursera - *Natural Language Processing with Classification and Vector Spaces* – Oct 2020

2020.10

MILITARY SERVICE

Completed: Republic of Korea Army, Discharged as Sergeant (Full Term)

Jun 2016 – Mar 2018