

Hai Do Manh

✉ haimanhdo1999@gmail.com | ☎ (+84)949109533 | [in linkedin.com/in/dsqt1999](https://www.linkedin.com/in/dsqt1999)
Hanoi, VietNam

OBJECTIVE

Always strive to learn, develop and cultivate the latest knowledge and technology in AI, leveraging my expertise to develop innovative solutions to address enterprise challenges.

EDUCATION

Hanoi University of Science and Technology (HUST) <i>Engineer of Control Engineering and Automation;</i> Specialization: Control techniques and intelligent systems	Hanoi, Vietnam Aug 2017 – Apr 2022
Vinh Phuc High School for the Gifted <i>Specialized in Information Technology;</i> Runner-up in the Provincial High School Informatics Contest for Grades 11 and 12; National high school exam: 28.85/30	Vinh Phuc, Vietnam Aug 2014 – Jun 2017

EXPERIENCE

Oxiitek <i>AI Engineer</i>	Hanoi, VietNam Aug 2024 – Present
<ul style="list-style-type: none">Enterprise Automation Platform: Architected and deployed comprehensive automation solutions using n8n, implementing automated CRM workflows, approval processes, customer feedback collection systems, and seamless Jira task generation for development teams.AI Voice Bot for Smart Speakers: Engineered an intelligent agent-based voice bot enabling natural language interaction for smart home device management and control with advanced conversational AI capabilities.VoiceQA - AI Call Quality Evaluation System: Architected and developed a comprehensive AI-powered call quality evaluation system utilizing Gemini 2.5 Flash for intelligent analysis of customer service calls. Implemented RESTful API with FastAPI, integrated AWS S3 for audio storage, FAISS vector database for semantic search, and AWS Transcribe for speech-to-text processing. The system provides automated scoring against predefined quality criteria with real-time audio processing capabilities and Docker-based deployment infrastructure.MCP for IoT Integration: Developed a comprehensive MCP (Model Context Protocol) server for office API ecosystem integration, enabling seamless IoT device control and management through LLM Agents with real-time responsiveness.Customer Service Automation Platform: Built an end-to-end automation platform featuring speech-to-text transcription, intelligent conversation summarization, and automated support ticket generation with contextual analysis.OCR Document Processing Solutions: Deployed production-ready Optical Character Recognition solutions for document processing, implementing privacy-preserving personal information redaction and automated product label counting in exhibition imagery.AWS Data Pipeline and Analytics Platform: Designed and implemented comprehensive ETL data pipeline architecture managing data flow from DynamoDB to Data Lake and S3 storage. Orchestrated automated data processing workflows using AWS Glue with scheduled cronjobs for real-time data synchronization. Built analytical datasets using Amazon Athena for SQL-based querying and created interactive business intelligence dashboards with AWS QuickSight, enabling data-driven decision making and real-time performance monitoring across enterprise operations.	
Job3s <i>AI Engineer</i>	Hanoi, VietNam Apr 2023 – Aug 2024
<ul style="list-style-type: none">LLM Content Generation System: Developed an intelligent content generation system utilizing Large Language Models for automated job description creation and resume optimization, improving recruitment content quality and consistency.Multi-Site Web Scraping Pipeline: Engineered robust data pipeline infrastructure for multi-site web scraping using BeautifulSoup, Scrapy, and Selenium, processing and storing large-scale datasets to AWS S3 with automated data validation.AI Career Recommendation Engine: Created AI-powered career recommendation engine using Elasticsearch for intelligent candidate-job matching and salary prediction algorithms, significantly enhancing recruitment process efficiency and accuracy.Resume Parsing System: Built comprehensive resume parsing system extracting 20+ structured fields (personal information, education, experience, skills) through hybrid PDF processing and OCR techniques using VietOCR for Vietnamese text recognition.Semantic Similarity Matching: Implemented semantic similarity matching system calculating resume-job description relevance using PhoBERT-base embeddings and cosine similarity algorithms, achieving high-precision candidate ranking.	

VinBigData, Vingroup Company

NLP Engineer

Hanoi, VietNam

Oct 2022 – Apr 2023

- **Aspect-Based Sentiment Analysis Dataset:** Constructed large-scale aspect polarity-based sentiment analysis dataset for Vinpearl customer reviews, implementing comprehensive data labeling, preprocessing pipelines, and Vietnamese DeBERTa model evaluation.
- **ABSA Research with Context-Aware Attention:** Conducted advanced research on Aspect-Based Sentiment Analysis comparing BERT, BERT-pair, DeBERTa, and Context-aware attention mechanisms on Vietnamese sentiment datasets.
Research Impact: Context-aware attention demonstrated 6% performance improvement over traditional self-attention in ABSA tasks across VLSP16 and Vinpearl datasets.
- **GPT-Powered Product Chatbot:** Designed and preprocessed 400K+ Shopee product information dataset for GPT-powered chatbot system, enabling efficient product information retrieval and customer query resolution.

VinBigData Institute

AI Engineer Training Scholar

Hanoi, VietNam

Jul 2022 – Oct 2022

- **AI Specialization Program:** Completed intensive AI specialization program covering advanced machine learning, deep learning, computer vision, and natural language processing with hands-on industry projects.
- **Professional Development:** Enhanced professional competencies in teamwork, technical leadership, and presentation skills through comprehensive corporate-sponsored training curriculum.
- **Multi-Domain AI Projects:** Delivered multiple AI projects including medical polyps segmentation and classification, violence detection in video streams, and intelligent question-answering systems, demonstrating versatility across AI domains.

Smart Technology for Industry (ST4I)

Computer Vision Engineer - Intern

Hanoi, Vietnam

May 2021 – Jul 2021

- **Template Matching Optimization:** Optimized template matching algorithms across Python, C++, and C# platforms, achieving significant improvements in processing speed and accuracy for industrial data processing applications.

HUST

Research Assistant

Hanoi, VietNam

May 2020 – April 2022

- **ML/CV Research:** Conducted cutting-edge research in machine learning, computer vision, and deep learning methodologies, contributing to advancements in edge detection algorithms and facial/object recognition systems.
- **SpO2 Prediction System:** Developed non-contact real-time SpO2 prediction system using camera-based computer vision and machine learning techniques for remote health monitoring applications.

PROJECTS

AI-Powered Advertisement Banner Classification System | TensorFlow, OpenCV, FastAPI, PostgreSQL

Jan 2025 – Present

- **Computer Vision Pipeline:** Architected and developed a comprehensive deep learning system for automatic classification of advertisement banners from elevator video content using ResNet50/101/152 and MobileNetV3Large architectures with transfer learning techniques.
- **Advanced Video Processing:** Implemented intelligent frame extraction using OpenCV and FFmpeg with SSIM-based scene change detection, achieving optimal frame selection for banner identification with 95% accuracy.
- **Production-Ready API:** Built scalable FastAPI backend with asynchronous processing, PostgreSQL integration, and Docker deployment supporting real-time video analysis and campaign mapping for advertising effectiveness tracking.
- **Data Augmentation & Optimization:** Designed custom data augmentation pipeline including random resize, light spot effects, and scene-specific transformations to simulate real elevator lighting conditions, improving model robustness by 15%.

WrenAI - Enterprise Natural Language to SQL Platform | Python, FastAPI, Qdrant, BigQuery, Docker

April 2025 – Present

- **Multi-Tenant AI Architecture:** Led development of enterprise-grade multi-tenant AI service supporting isolated data processing, tenant-specific LLM configurations, and secure API authentication with complete data isolation using Qdrant vector database collections.
- **Dynamic Data Source Integration:** Engineered seamless BigQuery integration with automatic schema discovery, dynamic SQL generation, and intelligent data profiling using AI-powered analysis for enterprise data warehouse connectivity.
- **RAG-Enhanced Query Generation:** Implemented advanced Retrieval-Augmented Generation pipeline converting natural language queries to optimized SQL with context-aware processing, supporting multiple LLM providers (OpenAI, Azure OpenAI, Ollama) and achieving 90% query accuracy.
- **Microservices Architecture:** Designed and deployed scalable microservices ecosystem including AI service, UI components, data engine, and launcher with Docker orchestration, supporting 100+ concurrent tenants with sub-second response times.

- **Intelligent Document Processing:** Architected comprehensive RAG (Retrieval-Augmented Generation) chatbot system enabling natural language interaction with enterprise document repositories. Implemented intelligent document upload, organization by topics, and fast search capabilities across large document collections.
- **Multi-Language Support:** Developed conversational AI platform with advanced natural language processing supporting Vietnamese and English languages, featuring context-aware responses and intelligent conversation summarization for enhanced user experience.
- **Scalable Web Architecture:** Built production-ready Django web application with Docker containerization, featuring user authentication, admin panel for content management, and responsive web interface optimized for both development and production environments.
- **Enterprise Integration:** Designed modular architecture supporting easy integration with existing enterprise systems, automated deployment pipelines, and comprehensive API documentation for seamless organizational adoption.

AI Customer Service Automation Suite | Python, NLP, FastAPI, PostgreSQL, Docker

Jan 2025 – Present

- **Intelligent Customer Support Chatbot:** Developed comprehensive AI-powered customer service chatbot utilizing advanced NLP techniques for automated query resolution, ticket classification, and intelligent routing to appropriate support channels with 85% first-contact resolution rate.
- **Aviation Information Collection Bot:** Engineered specialized chatbot for aviation industry collecting flight declaration information, validating regulatory compliance, and automating data entry processes for aviation authorities with real-time validation and error handling.
- **Automated PDF Report Generation:** Built intelligent system for generating comprehensive PDF reports for agricultural drone flight registration processes, including automated form filling, regulatory compliance checking, and digital signature integration with government aviation databases.
- **Multi-Channel Integration:** Implemented omnichannel chatbot deployment supporting web interfaces, mobile applications, and API integrations with existing CRM systems, enabling seamless customer experience across all touchpoints.

PERSONAL COMPETENCES

- **Programming Languages:** Python, C++, C#, JavaScript, HTML, CSS, SQL
- **AI/ML Frameworks:** TensorFlow/Keras, PyTorch, LangChain, LangIndex, Hugging Face Transformers, OpenAI API
- **Web Frameworks:** FastAPI, Django, Flask, RESTful APIs, Unicorn
- **Databases & Storage:** PostgreSQL, MySQL, MongoDB, FAISS Vector Database, Qdrant, AWS S3, Elasticsearch
- **Cloud & DevOps:** AWS (S3, Glue, Athena, QuickSight, Transcribe, DynamoDB), Docker, Docker Compose, Microservices
- **AI/ML Specializations:** RAG (Retrieval-Augmented Generation), NLP, Computer Vision, OCR, Semantic Search, LLM Integration
- **Data Processing:** ETL Pipelines, Data Lake Architecture, Business Intelligence, PDF Processing, Audio Processing
- **Project Management:** Git, GitHub, GitLab, Trello, Confluence, Agile
- **Others:** Linux, Postman, n8n Automation, IoT Integration, MCP (Model Context Protocol)