Ms. Hu Xinjie

+971 0585200913 · kosam0913@gmail.com · Portfolio · UAE Resident (Family-Sponsored Visa)

SUMMARY

7+ years of experience in workflow and pipeline development across multinational corporations and agile environments, including 2+ years focused on AI iteration workflow development and MLOps. Passionate about building robust, human-centric AI solutions that drive meaningful and positive impact

PROFESSIONAL EXPERIENCE

miHoYo, Shanghai, China | Motion Technical Director **Nov 2022 – Dec 2024**

- Led the end-to-end motion capture AI model training workflow, including data management, model training, and feedback integration, accelerating training cycles by 20× via distributed compute cluster (render farm), reducing iteration time from days to hours
- Established the MLflow-based model lifecycle management system, enabling reproducible training, transparent benchmarking, and deployment performance logging across 100+ experiments
- Applied advanced feature engineering to improve model robustness under noisy or occluded input, increasing accuracy by 15% in live settings
- Conducted data validation and statistical analysis to reduce overfitting, established training data version control and detailed experiment tracking system, ensuring consistent and high reproducibility of models
- Collaborated with researchers and engineers to align data flow with production needs, driving continuous pipeline development and improvement

miHoYo, Shanghai, China | Pipeline Technical Director **Aug 2021– Nov 2022**

- Designed and deployed a robust motion capture data pipeline from the architectural level, automating upload, conversion and synchronization, the pipeline has been running stably in production since launch
- Developed and enhanced general production pipelines using the Shotgun framework and Rez packages across 10+ departments, streamlining workflows and increasing productivity
- Designed and implemented a production delivery system integrated with linked notification workflows, improving cross-team communication and project visibility for 100+ artists across all departments, resulting in more timely updates and reduced delivery delays
- Built distributed computing cluster resource allocation and render retry system using Kafka and Linux services, optimizing resource usage and reducing manual intervention by over 90%
- Designed a specialized Unreal Engine launcher for live-streaming requirements and automated the Git CI/CD workflow, ensuring seamless integration and optimal performance

Industrial Light & Magic, Singapore | Associate Production Engineer **Jul 2017 – Jul 2021**

- Provided daily technical support across departments in studio, worked closely with CG supervisor/ production/ CG artists to ensure multiple projects run smoothly and get delivered as scheduled
- Developed render data extraction scripts and integrated metric collection with Elasticsearch, establishing real-time data visualization via Graphite and Grafana on cluster with 20,000+ CPU cores
- Designed Clarisse resource allocation scripts maximizing license and hardware utilization while meeting daily render targets
- Developed 3D software plugins/scripts to automate routines to increase workflow efficiency of artists, improved in-house pipeline tools to expedite workflow based on project-specific requirements
- Enrolled in post-production projects: Red Notice, Aquaman, Solo: A Star Wars Story, Ready Player One

PROJECTS

Notion Tree View Chrome Plugin (link) **Apr 2025 – Jun 2025**

- Identified user pain points in Notion's navigation experience and developed a Chrome plugin that adds a structured tree view to pages, improving usability and visual hierarchy for power users
- Built with React and TypeScript, featuring a modular architecture and dynamic DOM parsing for compatibility and performance
- Ranked #9 on Product Hunt Daily Launch, receiving strong user feedback and market validation

UCG – WeChat Mini Program for the UAE Chinese Community (link) **Feb 2025 – Jun 2025**

- Independently designed and developed a WeChat Mini Program for the Chinese-speaking community in the UAE, delivering community news and service directories

- Implemented frontend using native HTML, CSS, and JavaScript, fully integrated into the WeChat Mini Program environment
- Built backend using Flask, designed and managed a MySQL relational database with advanced schema modeling and query optimization
- Engineered efficient database access via indexing, transaction management, and optimized relational joins
- Deployed on a self-managed Ali Cloud Linux server, configured Nginx as reverse proxy and used Linux systemd services to manage persistent backend processes
- Completed closed beta testing with positive user feedback; currently in final stage of Tencent Mini Program approval

Ride Price Factors and Ride-Sharing Demand Forecasting (link) **Mar 2025 – May 2025**

University of Pennsylvania - MCIT Capstone Project

- Developed end-to-end data pipelines for preprocessing and modeling ride-sharing datasets, including data ingestion, cleaning, feature engineering, and model evaluation.
- Performed in-depth exploratory data analysis to identify root causes of price fluctuation and demand spikes, leveraging statistical techniques and domain-specific segmentation.
- Extracted features from datetime, geospatial, and categorical data to enhance model interpretability and predictive performance
- Applied and compared regression models such as Random Forest, XGBoost, and Ridge Regression for forecasting ride prices and demand, providing actionable business insights

EDUCATION

University of Pennsylvania, USA (Part-time Online) **Jan 2025 – Present**
MSE. Data Science, School of Engineering

Nanyang Technological University, Singapore (CGPA: 4.00 / 5.00) **Jul 2016 – Jun 2017**
MSc. Digital Media Technology, School of Computer Science and Engineering
Specialization: 3D Modelling & Advanced Computer Graphics

Wuhan University, China, Second Upper Honors (CGPA: 3.42 / 4.00) **Sep 2012 – Jun 2016**
BSc. Geographic Information System, School of Resource and Environmental Science
Specialization: Geospatial Analysis & Computer Graphics
Research Project: China National Hyperspectral Research Project (2015-2016): *Developed sparse representation model with SVM-optimized dictionaries, achieving 70% data reduction via spectral-spatial fusion*

CERTIFICATIONS

- Machine Learning Engineering for Production (MLOps), DeepLearning.AI (Coursera)
- TensorFlow Developer Certificate, DeepLearning.AI (Coursera)
- Deep Learning Specialization, DeepLearning.AI (Coursera)

SKILLS

Programming & Scripting
Python, SQL, Linux Shell scripting, Git

Machine Learning & Data Engineering
TensorFlow (Keras), MLflow, Data Version Control (DVC), ETL/ELT Pipelines, CI/CD, Kafka

Data Processing & Analysis
Pandas, Numpy, Apache Spark

Cloud & Infrastructure
AliCloud, Nginx, Linux Systemd