Yuanyuan Xu

Wuxi, Jiangsu, China
+86 15600159007 • xuyuanyuan102888@outlook.com
cavalryxyy.github.io/BudingBlog/ • cavalryxyy

I have over 6 years of working experience in both AI application, data science and Python developer, which shows excellent performance and teamwork ability. I am a quick learner and I want to explore data patterns that can drive meaningful results for the business.

Experience

- **Bekaert - Global Advanced Digital Solution** **2023 Jan~now** **Remote**
  Data Scientist & AI Engineer

  Project 1: Industrial Process Optimization (MLOps)
  - Migrated archived SAS repositories to Python ML models with PySpark ETL processing.
  - Transformed manufacturing data using Azure Synapse and PySpark for real-time analytics.
  - Designed and developed Machine Learning models for real-time prediction.

  Technical Stack: Azure Synapse, Python, PySpark, SAS CLI, PowerBI, Azure DevOps

  Project 2: Sentiment Analysis (AI Application)
  - Built pipeline to extract financial metrics and forward-looking business strategy statements.
  - Implemented RAG system with GPT-4o for sentiment scoring.
  - Created regularization algorithms for sentiment score smoothing and validation.

  Technical Stack: Prompt Engineering, RAG, Financial Analytics

  Project 3: Patent Landscaping (AI Application)
  - Developed GenAI bot system with LangChain for automated patent analysis and categorization.
  - Implemented GPT-4, Gemini, and AutoGen integration for intelligent patent processing.
  - Built keyword clustering and priority assessment algorithms for landscape analysis.

  Technical Stack: Prompt Engineering, AutoGen, LangChain

- **Bekaert Technology Research Development Co.** **2020 Jan ~2022 Dec** **Wuxi**
  Data Engineer & Data Analyst

  Project 1: Portfolio Analysis (Business Analytics / MLOps)
  - Built comprehensive ETL pipeline using Azure Synapse Analytics for portfolio data processing.
  - Developed SQL-based data warehouse with interactive PowerBI dashboards for business insights.
  - Created KPI metrics and performance analytics for investment decision support.

  Technical Stack: Azure Synapse Analytics, SQL, PowerBI, Data Warehousing

  Project 2: Tensile Strength Predictive Model (MLOps)
  - Developed time-series forecasting model using Scikit-learn for manufacturing quality prediction.
  - Built model evaluation dashboard with cross-validation and performance metrics.
  - Deployed FastAPI service with Azure DevOps for real-time predictions.

  Technical Stack: Python, Pandas, Scikit-learn, Plotly, FastAPI, Azure DevOps

  Project 3: Die Driver Hypothesis Test & Predictive Model (MLOps)
  - Built predictive ML model with hypothesis testing for manufacturing optimization.
  - Implemented data preprocessing pipeline with feature engineering and quality metrics.
  - Developed Plotly dashboards for performance visualization and manufacturing insights.

- Created FastAPI service with Azure DevOps deployment for production integration.
Technical Stack: Python, Pandas, Scikit-learn, Plotly, FastAPI, Azure DevOps

Project 4: Vehicle Trajectory Tracking (Data Analytics)
- Processed GPS trajectory data with Pandas and coordinate transformation algorithms.
- Built trajectory analysis engine with distance calculations and acceleration analysis.
- Developed FastAPI service with Plotly visualizations for real-time tracking dashboards.
- Implemented RDP algorithm for trajectory optimization and multi-track processing.
Technical Stack: Python, Pandas, Plotly, FastAPI, Azure DevOps, GPS Data Processing

○ **Bekaert Technology Research Development Co.** **2019 Sep ~2020 Dec** **Wuxi**
Python Developer

Project 1: Industrial IoT & Manufacturing Optimization (Computer Vision)
- Developed Python/OpenCV backend pipeline for real-time magnetic point detection and monitoring.
- Implemented FastAPI REST services for data collection and process optimization.
- Built computer vision algorithms for automated quality control in manufacturing processes.
- Implemented real-time data processing with Azure Stream Analytics for production monitoring.
- Built FastAPI service with 20-minute batch processing for predictive maintenance.
- Architected IoT solution connecting IR thermometers and PLCs via Azure IoT Hub.
Technical Stack: Azure Synapse Analytics, Python, OpenCV, FastAPI, Computer Vision Algorithms

Project 2: Web Calculator (Web Applications & Cloud Infrastructure)
- Developed React/FastAPI web app serving 200+ users across different time zones, featuring dynamic form builder with Dash components for custom quote modeling, authentication with hierarchical roles (Admin/Sales/Developer), and multi-currency engine with auto-rate fetching.
- Architected hybrid storage: SQLite for transactional data, Azure Blob for contract PDFs.
- Implemented CI/CD via Azure Pipelines (Docker > Azure WebApp) with zero-downtime deployments.
Technical Stack: FastAPI, React, Docker, Azure Web Apps, SQLite, Azure Blob Storage

Project 3: Calculate Internal Lamellar Spacing (Computer Vision)
- Converted MATLAB scripts to Python-based end-to-end image processing pipeline.
- Developed automated microstructure analysis for material science applications.
- Implemented OpenCV algorithms for precise lamellar spacing measurements.
Technical Stack: MATLAB, Python, OpenCV, Image Processing

Project 4: Micro Structure (Data Analytics / Machine Learning)
- Built comprehensive backend pipeline for microstructure image classification and analysis.
- Implemented data extraction, cleaning, and feature engineering for material characterization.
- Developed machine learning models for automated microstructure classification.
Technical Stack: Python, Scikit-learn, OpenCV, Data Analytics

Internship

Siemens Industrial Turbomachinery AB **Finspång**
Parameter Study of Blade-Disc Attachment (Thesis Internship) *Apr 2018–Oct 2018*

- Validated a novel design of the blade-disc attachment in a gas turbine.
- Studied the effects of the contact geometry, thermal stress, and manufacturing tolerances on the life cycle of the turbine rotor.

Research Areas:
Finite element modeling and simulation, fatigue and topology optimization.

Education

- **KTH Royal Institute of Technology, School of Engineering Mechanics** — **Stockholm**
  M.S. in Solid Mechanics — *Sep 2016–Oct 2018*
- **Beijing Institute of Technology, School of Mechatronical Engineering** — **Beijing**
  B.S. in Engineering Mechanics — *Sep 2012–Aug 2016*

Skills

- **Programming Languages & Frameworks:**
  Proficient in Python ecosystem: OpenCV, Pandas, Scikit-learn, Scipy, FastAPI, Plotly, Dash.
  Familiar with SQL, R, SAS, MATLAB, LaTeX, React, Docker, Azure DevOps.
- **Cloud & Infrastructure:**
  Azure Synapse Analytics, Azure Machine Learning, Azure Web Apps, Azure Blob Storage, Azure Pipeline.
- **AI & Machine Learning:**
  Prompt Engineering, RAG (Retrieval-Augmented Generation), LangChain, AutoGen, Computer Vision Algorithms, Image Processing.
- **Data Analytics & Visualization:**
  PowerBI, Data Warehousing, GPS Data Processing, Business Analytics, Financial Analytics.
- **Languages:**
  Native in Chinese. Fluent in English.