Y.GU

Yanying Gu

Data Scientist with 9 years of experience designing, building, and deploying intelligent agents and ML systems that automate data-driven decisions in financial environments. Deep expertise in Python, software design patterns, and scalable architectures. Skilled at integrating ML solutions with enterprise systems and delivering production-grade solutions with measurable business impact. Known for translating strategy into code, owning full ML lifecycles from prototype to production, and driving efficiency and value across fast-paced, high-performance teams.

CONTACT

+31657143815
Yanying.gu@gmail.com
Alkmaar (NL)
https://www.linkedin.com/in/yanying-gu/

COMPETENCE

- ✓ Creative & Innovative
- ✓ Collaborativ
- ✓ Accountable
- ✓ Growth-Oriented
- ✓ Impact-Driven

Education

2006 – 2010
PhD Telecom
Electrical Engineering, Mathematics and Computer Science (EEMCS)
TuDelft

2004 – 2006
MSC Telecom
TuDelft
Cum Laude, top 5%

LANGUAGE

English – fully professional
Dutch – NT2 (C2)

Experience

2022 - 2025
Data scientist, MLE, | Ayvens/LeasePlan, Société Générale | Amsterdam

- Project: Automated Decision Engine
  Designed and deployed Python-based ML solution on AWS SageMaker to streamline approval workflows and operational decisions. Integrated models with enterprise systems to automate 50% of approvals, enhancing processing speed, scalability, and regulatory compliance.

- Project: ML Enhancement in AWS
  Redesigned ML pipelines for forecasting by migrating ETL to Snowflake and implementing SageMaker Pipelines with automated CI/CD. Delivered a robust platform enabling accurate predictions and faster deployment of inference and risk models across teams.

- Project: ML/AI solution migration to Azure
  Developed end-to-end engine for risk detection on Azure Databricks, coordinating cross-functional teams to maintain production reliability. Scaled ML workflows for operational monitoring, supporting proactive mitigation of exceptions and regulatory risks.

- Project: Intelligent Operational Agent
  Designed AI agents to automate approval and execution workflows using LLMs, LangChain, and RAG pipelines. Integrated real-time operational data streams,

enabling data-driven decisions while maintaining compliance and operational reliability.

- Project: ML-Driven Vehicle Price Estimation
  Implemented ML pipelines for asset valuation using SageMaker, Airflow, and Snowflake. Delivered predictive pricing models across 23 markets, enabling faster, data-driven decisions and reducing manual workload by 60%. The solution automated pricing based on vehicle and market data, with built-in monitoring and retraining, and was fully integrated into business operations.

- Project: Proactive Retention & Marketing Optimization Engine
  Designed a reproducible data science framework for proactive customer retention, combining churn prediction, survival modeling, customer value modeling, and causal inference to optimize marketing interventions under budget and compliance constraints. Enabled test-and-learn decision-making through uplift modeling, A/B experimentation, and ROI-driven targeting suitable for regulated financial services.

2018 - 2022
Data scientist, MLE | LeasePlan Digital & IT | Amsterdam

- Project: Mileage Prediction – Predictive Operational Analytics
  Developed production-grade ML pipelines for forecasting asset usage, enabling proactive detection of maintenance risks. Integrated batch inference, monitoring, and alerting to support predictive decision-making and reduce operational downtime for 5M+ assets weekly.
- Project: Customer Satisfaction & Retention Risk Analysis (B2B Customer Relationship Analytics)
  Applied statistically analysis and NLP-based text mining on structured surveys and open-text feedback to identify key drivers of customer retention risk, balancing interpretability and data constraints.

2010 - 2018
Data scientist | KPN | Den Haag

- Project: IoT Performance Monitoring
  Use Case: Predictive analytics and time series analysis to optimize operational network performance and reliability
  Tools/Stack: Python, Pandas, NumPy, Scikit-learn, RNNs, CNNs, modular design patterns

Big Data Engineer, Architect | KPN | Den Haag
- Project: Big Data Governance and Operations
  Use Case: Data management, governance, and compliance for large-scale operations
  Tools/Stack: Hadoop, Hive, GDPR/PDPA compliance, Spark, SQL, Python 3.x, modular data pipelines, and enterprise-grade security frameworks.

Innovator, Project manager | KPN Mobile Innovation Voice & Signaling | Den Haag
- Relational database and relational data modelling
- LTE/4G and roaming, MVNO, Diameter signaling