Adam Malinowski
Expert Data Scientist/ML Engineer

SUMMARY

Engineer with 10+ years' experience in AI, excelling in NLP, GenAI, computer vision, and model deployment. Expertise in Python, SQL, PySpark, and cloud platforms. Established record of enhancing AI-driven services, improving workflows, and boosting efficiency. Proven capabilities in developing robust ML pipelines, integrating state-of-the-art technologies like LLMs and RAG pipelines, and delivering solutions across diverse industries. Advanced in web and full-stack development, data engineering, analysis, and DevOps practices, underpinned by solid formal education in computer science.

TECHNICAL SKILLS

Main Technical Skills	Python, SQL, PySpark, NLP, GenAI
Programming Languages	Python
AI & Machine Learning	Azure Cognitive Search, Faiss, GenAI, LangChain, LangGraph, Mlflow, n8n, NLP, NumPy, OCR, OpenAI, OpenCV, PandasAI, Pinecone, Scikit-learn, Xgboost
Python Libraries and Tools	Matplotlib, NumPy, PySpark, Scikit-learn, Seaborn
Data Analysis and Visualization Technologies	Apache Airflow, Data visualization, Jupyter Notebook, PandasAI, Power BI, Tableau
Databases & Management Systems / ORM	MongoDB, PostgreSQL, SQL
UI Frameworks, Libraries, and Browsers	Dlib
Azure Cloud Services	Azure Cognitive Search
Google Cloud Platform	Google BigQuery
Deployment, CI/CD & Administration	Automation deployment CI, CD, CD DevOps pipelines, CI/CD, Jenkins
Methodologies, Paradigms and Patterns	Clean Architecture, FDD

Virtualization, Containers and Orchestration	Docker, Kubernetes, Terraform
Version Control	Github Actions
Platforms	Google Colaboratory
Message/Queue/Task Brokers	Kafka
SDK / API and Integrations	RESTful API
Other Technical Skills	AutoGen, CRNN, Unreal Engine, YOLOv7

WORK EXPERIENCE

GenAI / LLM Developer, GenAI/LLM Development

Duration: Feb 2024 – Present

Summary: Built LangChain agents with guardrails, retries, and n8n orchestration, significantly reducing hallucinations and improving workflow automation.

Responsibilities: Developing LangChain agents, integrating LangGraph flows and serverless solutions for document processing and summarization, and monitoring GenAI model performances with interactive analytics.

Technologies: LangChain, LlamaIndex, OpenAI GPT-4, Hugging Face, PyTorch, AWS Lambda, S3, Step Functions, Power BI, Snowflake, n8n

ML Engineer, Machine Learning Engineering

Duration: Apr 2020 – Feb 2024

Summary: Developed retrieval augmented generation pipelines and cloud-native pipelines for real-time document processing, enabling significant improvements in chatbot accuracy and AI access.

Responsibilities: Designing RAG pipelines, building Telegram bots, fine-tuning language models, orchestrating CI/CD workflows, and deploying OCR and face recognition systems.

Technologies: Pinecone, Claude 2, Vertex AI, BigQuery, Docker, Flask, Django, Terraform, GitHub Actions, TensorFlow/Keras, BERT, FAISS, YOLOv7, CRNN, MobileBERT, OpenCV, Dlib, React, TypeScript, Elasticsearch, Gradio, Streamlit, Power BI, ELK, Prometheus, Grafana

Data Scientist, Data Science

Duration: Oct 2015 – Apr 2020

Summary: Analyzed large volumes of customer data for actionable insights, built ML models for prediction and segmentation, and developed ETL pipelines to support scalable real-time analytics.

Responsibilities: Conducting data analysis and predictive modeling, developing sentiment analysis models, and building scalable ETL pipelines in collaboration with product and engineering teams.

Technologies: SQL, R, Excel, Scikit-learn, Pandas, NumPy, Matplotlib, PySpark, Apache Spark, BigQuery, NLTK, SpaCy, AWS SageMaker, XGBoost, Plotly, Seaborn, Tableau

EDUCATION

- **BSc Computer Science**
- 2010 – 2014
- **MSc Computer Science**
- 2014 – 2015