

VAHID SHARIFI

Data Scientist ~ AI R&D

+44 77 6091 2928

vahidsharifi7676@gmail.com

/in/vahid-sharifi

https://github.com/vahidsharifi

SKILLS

Programming Languages: Python, R, SQL, HTML, JavaScript.

Tools and Technologies: LangChain, OpenAI, HuggingFace, GCP, Scikit-learn TensorFlow, BeautifulSoup, Docker, WanDB, FastAPI, NLTK, SpaCy, Gensim, Keras, Pandas, NumPy.

EXPERIENCE

- 2025 – Present **Researcher & Generative AI Consultant** **University of Gloucestershire**
- Conducting literature reviews on AI storytelling, assessing feasibility, and advising on research directions.
 - Providing technical consultancy on Generative AI models and benchmarks for creative writing.
 - Assisting in methodology design for AI-powered storytelling frameworks and project development.
 - Presenting research findings in an accessible format for non-technical stakeholders, using infographics to enhance understanding.
- Google API / Gemini / ChatGPT / Claude / DeepSeek / OpenAI
- 2024 – Present **AI Quality & Performance Analyst** **Outlier**
- Monitoring and analysing AI model outputs to enhance accuracy and relevance.
 - Developed targeted prompts to evaluate and optimise model performance.
 - Providing feedback as an AI and mathematics expert for reinforce learning from human feedback(RLHF).
- 2024 – 2025 **Research Assistant (AI in Education) & AI Solutions Developer** **University of Gloucestershire**
- Developed AI-powered tools for enhancing teaching materials and decolonising the curriculum.
 - Employed image-to-text LLMs to extract embedded text out of the images.
 - Designed a human in the loop evaluation procedures for assessing the tools performance.
 - Applied NLP techniques to quantify the qualitative research, generating actionable insights from unstructured data.
 - Conducted systematic literature reviews for development of AI-driven methodologies.
- Python / LangChain / OpenAI / LangSmith / HuggingFace
- 2023 – 2024 **Researcher (AI in Healthcare) & NLP Engineer** **University of Gloucestershire**
- Led a research project to develop an AI-driven solution for generating nutritionally accurate dietary plans by integrating LLMs with external databases.
 - Engineered datasets by compiling data via APIs and LLMs for in-depth nutritional analysis.
 - Developed a data retrieval pipeline using RAG to improve nutrition data accuracy by 90%.
 - Employed statistical modelling and analysis for data validation and quality assurance.
- Python / LangChain / OpenAI / Gemini / HuggingFace / Pandas / Numpy / Tensorflow / Scikit-learn / SciPy
- 2022 – 2023 **Project Tube-LLM: YouTube In-Depth Transcript Search Engine** **[GitHub Link](#) / [GCP Deployed](#)**
- Implemented transcript retrieval from YouTube videos using vector databases.
 - Utilised Google Cloud Platform to to deploy the project
 - Utilised WanDB and NoSQL database such as Firestore for development and evaluation.
- Python / LangChain / GCP / LangSmith / WanDB / Docker / FastAPI / OpenAI
- 2020 – 2022 **ML Engineer, Co-Founder** **[Novin Sim](#)**
- Designed and implemented a hybrid recommender system to enhance user experience.
 - Deployed firebase for record and analysing user activity.
 - Applied NLP techniques to automate customer feedback analysis.
 - Developed and maintained financial dashboards to track and visualise KPIs such as profit, expenses, and revenue, aiding in data-driven decision-making for stakeholders.
 - Developed and maintained the business website.
- Python / R / JavaScript / HTML / SQL / PowerBI

EDUCATION

- 1/2023 - 6/2024 **Master of Science in Data Science** **University of Gloucestershire**
Achieved "Distinction" honor at 78%
- 8/2016 - 8/2021 **Bachelor of Science in Industrial Engineering** **Iran University of Science and Technology**
Relevant courseworks on statistical quality control.

LANGUAGES

English - IELTS 7, Persian - native