Yulin SHI

✉ eliane.shiyulin@gmail.com   📞 0766082632   📍 Paris   in yulin-shi   🔗 shi-yulin.com

EDUCATION

09/2022 – 09/2024 Paris, France	**Master's Degree in Computer Science (DAC),** Sorbonne University • Specialized in Data, Learning, and Knowledge. • Courses: Databases, Statistical Learning, Machine Learning, Deep Learning, Reinforcement Learning, Optimization, Symbolic Learning
09/2019 – 06/2022 Paris, France	**Bachelor's Degree in Computer Science,** Sorbonne University

EXPERIENCE

11/2025 – Present Remote	**AI Engineer,** Freelance Develop an interactive voice-driven web agent • Multi-agent orchestration: Coordinate specialized agents to qualify visitors via voice and autonomously navigate the website to relevant product sections. • Real-time web automation: Translate user intent into direct browser actions to visually present specific features. • Grounded RAG: Utilize a document-based knowledge base to provide reliable, fact-checked answers during the live demonstration. Voice AI, Multi-agent Systems, Browser Automation, RAG
08/2025 – Present Remote	**Adjunct Lecturer,** University Metropolitan Tirana Delivered lectures and practical lab sessions in Machine Learning and Natural Language Processing for Master's students in Software Engineering.
05/2025 – 07/2025 Paris, France	**NLP engineer** Develop an automated academic paper aggregation system (FeedPaper ⬈) • Multi-agent: Scrape real-time arXiv RSS feeds and generate concise LLM-based explanations of why papers match user-defined topics. • Hybrid search: Use vector and keyword retrieval to surface highly relevant papers. • Automated backend: Use Celery to orchestrate scraping, indexing, and daily report generation. LLMs, Multi-agent systems, Information retrieval, Hybrid search

09/2024 – 01/2025
Paris, France
	Freelance, NLP engineer
	• Developed an advanced search platform for African legal documents (Legomnia ⬈)
	   • Hybrid Search: Integrated hybrid search capabilities for efficient information retrieval.
	   • Metadata Extraction: Automated extraction of document metadata for filtering purposes.
	Embeddings Models, Information retriever, Hybrid search, Metadatas extractions, Web development

03/2024 – 08/2024
Paris, France
	RTE, Internship in AI/NLP
	• Detection of Equipment Failures from Operational Reports
	   • Designed a complete NLP pipeline: data acquisition, model training, performance evaluation, and deployment.
	   • Benchmarked models (CamemBERT, Llama, Vigostral,etc), fine-tuning for domain-specific tasks.
	   • Created a user-friendly web interface for end-user interaction and report analysis.
	LLM, QLoRA, Finetuning, classification

02/2023 – 05/2023	**Freelance,** NLP engineer
Paris, France	• Implemented methods from the article "Webpage text extraction using long short-term memory network based deep learning method" to extract the main content from HTML files.
	• Evaluated text length, number of punctuations, and content of each tag to identify the main content.
	• Achieved an F1 score of 87% using the LSTM model with FastText embedding.
	Embedding Model, LSTM, Web

HACKATHONS

12/2024	**4th Prize Winner, NASA Breath Diagnostics Challenge**
	• Developed a classification model to analyze data from NASA E-Nose for distinguishing between COVID-positive and COVID-negative breath samples.
	• Applied advanced data preparation and AI techniques to optimize performance despite limited sample size.
	Link to the competition ⬈
	Machine learning, Classification

06/2024	**First Prize Winner, Cryptocurrency Price Prediction Challenge**
	• Conducted data analysis and preprocessing to ensure quality and relevance.
	• Optimized prediction accuracy through model selection and concatenation and developed a high-performance predictive model to capture market trends.
	Machine learning, Decision tree

05/2024	**Finalist in the Mistral AI Paris Hackathon 2024: Development of a guitar Tablature Generation**
Paris	• We developed a model that generates guitar tablature and MIDI files from initial measures, desired style, and key. Using an open-source dataset of guitar recordings, we fine-tuned the Mistral-7B model for coherent musical continuations. This work was presented at the Mistral AI Paris Hackathon and Mistral Online Hackathon.
	• Link to our solution ⬈
	LLM, Fine-Tuning, QLoRA, Quantization

ACADEMIC PROJECTS

02/2022 – 06/2022	**Next-generation blockchain PoC implementation**
	• Implementation of a traditional blockchain and development of a new proposal where different types of information can coexist within the same blockchain in Java.
	• Link to the paper ⬈

SKILLS

Programming — Python, Java, C++, SQL, Shell • **Libraries & Frameworks** — PyTorch, TensorFlow, scikit-learn, numpy, pandas, LangChain • **Tools** — Git, LaTeX, Cloud (GCP) • **Web development**

LANGUAGES

| French | ● ● ● ● ○ | English | ● ● ● ● ○ |
| Chinese | ● ● ● ● ● | | |

NLP & LLMs

RAG systems, LLM fine-tuning (LoRA/QLoRA) • Prompt engineering, Instruction tuning •

Vector databases (FAISS / Weaviate) • Embeddings, text classification • Knowledge graph integration