Mohamed Sriha

Chicago, IL | 415-xxx- xxxx | mohamedsriha.com | medsriha@gmail.com | LinkedIn | GitHub

WORK EXPERIENCE

deepset.ai | *Senior AI Engineer – Solution Engineering* | *Mar 2024 - Present* | **Remote**

- Led end-to-end development of a multi-agent LLM-based chatbot for a market research customer, driving 5× ROI within 1 year, a 25% increase in user engagement, and production launch in under 4 months—saving 2–3 FTEs through streamlined LLMOps and data integration.
- Integrated Snowflake into deepset-cloud within weeks of joining, enabling agents-based pipelines to seamlessly access large-scale enterprise data through the platform without the need for data migration.
- Developed a RAG-based chatbot system, processing over 15 billion structured and unstructured data points to deliver insights, reducing reliance on external business intelligence and copilot tools.
- Designed and implemented a RAG-based agent system for a financial customer to negotiate missed car loan payments, leveraging multiple agents to automate negotiation workflows and improve customer engagement.
- Contributed to deepset open-source framework "Haystack" by improving core functionalities and adding new components
- Worked collaboratively with product and engineering teams to identify and prioritize platform improvements, ensuring they align with customer needs.
- Partnered with product managers, software engineers, and operations teams to identify opportunities for business impact, understand and prioritize requirements, and drive engineering decisions.

BMO US | *Lead Machine Learning Engineer – Anti-Money Laundering* | *May 2019 - Feb 2024* | **Chicago, Illinois**

- Led a team to develop an NLP automation tool for regulatory narrative generation, reducing case investigation time by 65%, saving $1M+ annually, and earning recognition from the Chief Risk Officer.
- Built an anomaly detection model using Isolation Forest to detect suspicious activities within Hong Kong and Singapore trade finance instruments; defined candidate variables, applied feature engineering, and established performance metrics; reduced false positives by 35%.
- Developed a supervised model aimed at identifying the likelihood of cash structuring using the CART algorithm, resulting in a 75% monthly precision.
- Defined best practices to standardize team coding and operating procedure access to the model, resulting in a more efficient process with fewer errors.
- Led the implementation of a comprehensive Machine Learning model explainability framework, expanding its adoption across external teams to meet regulatory requirements.
- Thought leader and internal consultant resource for cross-departmental teams that could benefit from ML solutions.

BMO US | *Data Scientist – Credit Risk* | *Aug 2018 - Apr 2019* | **Chicago, Illinois**

- Prototyped several credit default models combining DBSCAN clustering for customer segmentation with ensemble methods (Random Forest, XGBoost) for default prediction; designed as candidates to replace legacy rule-based systems, demonstrating a 20% improvement in accuracy ratio.
- Designed and implemented rigorous statistical validation frameworks to ensure model stability, fairness, and robustness across diverse customer segments and economic scenarios.
- Extracted and transformed terabytes of customer data across disparate database schemas, implementing comprehensive bias detection methodologies to ensure fair lending practices.

EDUCATION

Carnegie Mellon University	**Pittsburgh, PA**
Master's in Computer Science and Information Technology	*June 2018*
Institut Supérieur de Gestion	**Tunis, Tunisia**
Bachelor's in Economics	*Apr 2012*

SKILLS

- Python, SQL, Java, JavaScript, PyTorch, Haystack, LangChain
- Flask, Spring, FastAPI, Git, Docker, Kubernetes, CI/CD
- AWS, Hadoop ecosystem, Apache Spark, Snowflake

- Natural language processing
- Software engineering fundamentals
- Predictive modeling and anomaly detection