Hamza Sallemi

+216 92326403 aslamihamza@gmail.com linkedin.com/in/hamza-sallemi github.com/sallamihamza

PROFESSIONAL SUMMARY

Engineering specialist in Applied Mathematics and Modeling with expertise in Artificial Intelligence and Machine Learning. Demonstrated experience in designing and deploying AI solutions for risk prediction, customer behavior analysis, and medical image processing. Proven skills in full-stack development for integrating ML models into production. Passionate about transforming complex business challenges into innovative and scalable technical solutions.

Areas of Expertise: *Data Science | Machine Learning | Deep Learning | Risk Prediction | Predictive Analytics | API Development*

EDUCATION

Higher National Engineering School of Tunis **2023 – Present**
Applied Mathematics and Modeling Engineering

Higher Institute of Informatics and Mathematics of Monastir **2022 – 2023**
Master's Degree in Data Science

Higher Institute of Informatics and Mathematics of Monastir **2019 – 2022**
Bachelor's Degree in Mathematics

PROFESSIONAL EXPERIENCE

CODIX TN — Data Scientist Intern ***Jun 2025 – Jul 2025***
Centre Urbain Nord, Tunis, Tunisia | Hybrid

- Designed and implemented a credit risk prediction system leveraging a 45,000-record customer database, increasing predictive accuracy by 15% over traditional methods.
- Led comparative evaluation of 3 ML algorithms (Random Forest, XGBoost, Linear Regression) to identify optimal solution, establishing Random Forest as production model.
- Architected and deployed a REST API with Flask enabling real-time risk predictions with sub-200ms response time.
- Orchestrated development of responsive React user interface, improving user adoption by 40% and accelerating credit decisions.
- Automated data preprocessing pipeline for 13 critical variables, reducing input errors by 35% and ensuring data consistency.

Gafsa Phosphate Company — Software Development Intern ***Jun 2024 – Jul 2024***
Cité Bayech, Gafsa, Tunisia | On-site

- Developed a comprehensive intern management application in C# integrating tracking, evaluation, and reporting, reducing administrative workload by 30% and saving 15 hours/week.
- Collaborated with 5-member IT team to document technical solutions and create internal knowledge base.

MAJOR TECHNICAL PROJECTS

Credit Risk Prediction App | *Python, Flask, React, scikit-learn, RandomForest, Pandas* **Jun 2025**

- Designed and implemented a full-stack credit risk scoring application, leveraging Random Forest to achieve a 15% improvement in prediction accuracy.
- Created a Flask REST API for real-time scoring, seamlessly integrating with a responsive React front-end for user interaction.
- Automated data preprocessing and validation, ensuring high data consistency and accuracy across 13 critical input features.

Customer Churn Prediction for a Bank | *Python, scikit-learn, XGBoost, Pandas, SHAP, Power BI* **Mar 2025**

- Built a churn prediction model for a retail bank and improved classification performance using XGBoost.
- Engineered features from customer account activity and resolved class imbalance using scale_pos_weight and SMOTE.
- Applied SHAP to interpret model outputs and identify factors driving customer churn decisions.
- Exported predictions and key metrics to Power BI and created an interactive dashboard to monitor churn trends and high-risk customer groups.

Natural Language to SQL Converter | *OpenAI, Microsoft Azure, Streamlit, SQLite, Pandas, LangChain* **Oct 2024**

- Streamlined deployment processes by hosting the application on Streamlit and Azure, resulting in a 40% reduction in deployment costs and improved operational efficiency
- Ensured high data integrity and user satisfaction with SQL factuality of 89.39%, semantic similarity of 99.98%, and criteria match of 95.66%.
- Enhanced user engagement by 30% through the implementation of adaptive database handling and intuitive data visualization.

COVID-19 Chest X-Ray Classification | *TensorFlow, Keras, CNN, OpenCV, Python* **Aug 2024**

- Developed a CNN model with TensorFlow for COVID-19 detection from chest X-rays, achieving 87% accuracy and facilitating rapid diagnosis.

TECHNICAL SKILLS

Languages: Python, C/C++, Java, R, MATLAB, Microsoft Excel, Power BI.
Web Technologies: React.js, Tailwind CSS, Node.js.
Cloud & DevOps: Azure, Docker, GitHub Actions.
Backend Frameworks: Flask, FastAPI.
Databases: SQLite, MySQL, PostgreSQL
AI/ML: TensorFlow, PyTorch, scikit-learn, OpenAI GPT, Keras, CNN, Streamlit.

CERTIFICATIONS

Datacamp Data Scientist Associate certificate *Credential ID: DSA0016615442509*
Datacamp Data Engineer Associate certificate *Credential ID: DEA0018901421109*
Datacamp Data Analyst Associate certificate *Credential ID: DAA0018380597111*
Datacamp AI Engineer for Data Scientists Associate certificate *Credential ID: AEDS0010408968802*
Microsoft Azure AI Fundamentals *AI-900* *Credential ID: wbruu-H9uQ*
Microsoft Certified: Power Platform Fundamentals PL-900 *Credential ID: UYnh-sFVH*
GOMYCODE: The Data Science Bootcamp *Credential ID: 00011770*

LANGUAGES

- **Arabic : Native**
- **French : Professional (C1)**
- **English : Professional (C1)**