Sarah Hanaa Ben Abidallah

Liverpool, UK

+44 7399514637| sarah.benabidallah.sh@gmail.com | linkedin.com/in/sarah-benabidallah | github.com/SarahBenabidallah

EDUCATION

LIVERPOOL JOHN MOORES UNIVERSITY **Liverpool**, UK

Masters of Science in Data Science *Sep. 2022-Oct 2023*

- **Grade:** Distinction (72%)

  Relevant Coursework: Machine Learning, Data Mining, Efficient Algorithms For Complex Datasets, Big Data Computing, Statistical Methods in R, Data Analytics.

UNIVERSITY OF SCIENCE AND TECHNOLOGY HOUARI BOUMEDIENE **Algiers** ,Algeria

Masters of Engineering in Software Engineering *Sept 2015-2021*

- **Grade:** First Class Honors

  Relevant Coursework: Advanced Algorithms and Complexity, Modeling and evaluation of system performance, Information Systems and Software Engineering, Database Architecture and Administration, Foundations of Artificial Intelligence

WORK EXPERIENCE

Liverpool Centre for Cardiovascular Science. **Liverpool,** UK

Data Analyst *May. 2023 – August 2023*

- Led the project funded by the UK Health Security Agency to predict antimicrobial resistance in lung infection with cystic fibrosis using electronic health record data from Liverpool Heart & Chest Hospital. Utilized 20,035 training and 5,360 test observations.
- Established high accuracy in predicting multi-drug resistance in patients, demonstrating an accuracy of 97% exceeding standard benchmarks and rivaling established methodologies.
- Reduced decision-making time in antibiotic prescriptions by 48 to 96 hours, enhancing patient care efficiency.
- Improved antibiotic selection accuracy in up to 284 people living with cystic fibrosis through the implementation of predictive models.

PROJECTS

IDENTIFYING BIOMARKERS FOR PARKINSON'S DISEASE PROGRESSION *Feb 2023 - Sept. 2023*

- Identified key biomarkers crucial for early detection and intervention strategies in Parkinson's Disease progression by analyzing 37,772 records from the PPMI database.
- Attained a 14% enhancement in predictive model accuracy, elevating it from 70% to 84%, through optimization of hyperparameters using grid search and cross-validation techniques.
- Implemented innovative ensemble learning methods, combining SVM, Random Forest, and XGBoost models through soft and hard voting, resulting in a notable 10% increase in prediction accuracy compared to individual models, fostering more resilient and reliable predictions.

HUMAN ACTIVITY RECOGNITION PREDICTION *Feb 2023 – April 2023*

- Led a team of three to Craft advanced machine learning solutions, such as Random Forest and neural networks to analyze accelerometer data for human activity recognition. Improved model accuracy by 18% on the WISDM Lab dataset through hyperparameter tuning and feature engineering, achieving 92% accuracy rate.
- Leveraged dimensionality reduction techniques, such as PCA, T-SNE, and UMAP, to select the most informative features.
- Achieved superior model performance with a reduced feature set, preserving 90% of the original information.
- Spearheaded the team in conducting efficient hyperparameter tuning, resulting in a 15% reduction in model training time.

PREDICTING INCOME FROM ADULT CENSUS DATA USING NEURAL NETWORKS *March 2023 – April 2023*

- Directed the design and implementation of a machine learning algorithm to predict income by utilizing demographic and socio-economic data sourced from the expansive Adult Census dataset comprising 48,842 instances; facilitated the generation of accurate income projections, enabling strategic policy planning on income equity.
- Orchestrated the implementation of GPU-accelerated computing on Google Colab and Kaggle, effectively parallelizing model training processes; optimized computational resources and slashed processing time by 40%
- Secured an accuracy rate of 89% by implementing Deep Neural Networks and Convolutional Neural Networks, surpassing industry standards and showcasing the project's ability to deliver precise income predictions.
- Extracted crucial data on income levels to shape strategic policies; pinpointed key factors influencing income distribution, guiding policymakers in formulating effective strategies to tackle economic disparities head-on.

SKILLS

• **Data Science & Miscellaneous Technologies:** A/B testing, ETL, Data science pipeline (cleansing, wrangling, visualization, modeling, interpretation), Statistics, Time series, Experimental design, Hypothesis testing, OOP, OOD, APIs, Excel, Git, Azure, AWS, PowerBI

• **Big Data & Machine Learning:** Spark, Hadoop, MongoDB, Python (eg. scikit-learn, NumPy, pandas, seaborn, matplotlib, TensorFlow, Keras).

• **Programming Languages:** Python, R, Java, SQL, MySQL, OLAP, C, JavaScript, HTML, JSON, CSS, Node.js, React.