

ALI ESTIFA

Junior Data Scientist | Machine Learning | Python | Health Data

alilab.ir | github.com/Nsport441 | LinkedIn | estifa@gmail.com

PROFESSIONAL SUMMARY

Entry-level Data Scientist with a multidisciplinary background in health research, sports physiology, and nutrition. Hands-on experience in Python, data analysis, machine learning, and deep learning, including predictive modeling on clinical data and image classification. Interested in applied AI, healthcare analytics, and data-driven research.

TECHNICAL SKILLS

Programming & Analysis: Python, Pandas, NumPy, Jupyter Notebook, data cleaning, exploratory data analysis

Machine Learning: Scikit-learn, Random Forest, SVM, Gradient Boosting, KNN, cross-validation, ROC-AUC, feature importance, hyperparameter tuning

Deep Learning: TensorFlow, Keras, neural networks, CNNs, Sequential API, Functional API, KerasTuner

Data & Tools: SQL, PostgreSQL, Excel, Power BI, Git, GitHub, Docker, FastAPI, VS Code, Google Colab, n8n

SELECTED PROJECTS & RESEARCH

Sex-Specific Machine Learning Models for NAFLD Prediction | Python | Scikit-learn | Clinical Data

- Analyzed data from 306 overweight/obese adults and built Random Forest, SVM, Gradient Boosting, and KNN classifiers.
- Compared general and sex-specific models using accuracy and ROC-AUC; the male Random Forest reached 0.75 accuracy and 0.72 AUC.
- Examined feature importance across demographic, anthropometric, biochemical, and sleep-related variables.

Breast Cancer Classification with Deep Learning | TensorFlow | Keras

- Built and tuned neural-network classifiers on 569 observations and 30 features using dropout, K-fold validation, and KerasTuner; optimized validation accuracy reached approximately 99%.
- Compared baseline and regularized architectures and evaluated model performance across validation folds.

Chest X-Ray Pneumonia Classification | CNN | TensorFlow | Keras

- Developed image-classification CNNs with Sequential and Functional Keras architectures; improved validation accuracy to approximately 95%.
- Applied model evaluation and iterative architecture refinement to improve generalization on chest X-ray data.

Machine Learning for Drug-Induced Malnutrition | Research Review | Biomedical AI

- Conducted literature screening on ML approaches for early detection of medication-associated malnutrition and nutrient depletion using Rayyan.

EDUCATION & TRAINING

M.Sc. Sports Physiology - Nutrition Focus | 2025 - Research focus included health/clinical data and machine learning for fatty liver disease prediction.

Bachelor's Degree - Sports Nutrition - Academic foundation in nutrition, physiology, health, and human performance.

IBM Applied Data Science Capstone - Coursera - Applied data science project work and end-to-end analytical workflow.

ONGOING PROFESSIONAL DEVELOPMENT

Current learning: Data Analysis, Mathematics for Machine Learning, PostgreSQL for analytics, and practical deployment workflows with Docker/FastAPI.

Interests: Data Science | Machine Learning | Healthcare AI | Predictive Analytics | Biomedical Data