SEDJARI Yassine

Portfolio Github Linkedin Contact

Education

ENSIAS, Mohammed V University Sep 2021 - July 2024

Bachelor of Science in Computer Science and Applied Mathematics, Major: Artificial intelligence

- Coursework includes: Statistics, Machine Learning, Deep learning, Optimization, Theoretical Computer Science, Random Event Modeling, Operations Research, Game Theory, Data Analysis, High Performance Computing.

Lycée Al Khansaa Sep 2019 - July 2021

Classes préparatoires, Mathematics, Physics, Engineering Science

- Two-year post secondary program in advanced mathematics and physics leading to nationwide entrance examinations to the Grandes Ecoles for scientific studies.
- Coursework includes: Convex Optimization, Mathematical Analysis, Vector Calculus and Parametric Equations, Integral Calculus and Differential Equations, Probability Theory and Random Variables, Linear Algebra and Matrices.

Internship Experience

Oracle Labs, Morocco (On-site) Juin 2023 - Sept 2023

Research Assistantship *Supervised by Mr. Ali SEYFI*

- Enhanced the AUC score of Oracle AutoML API on image classification task by 2.73%.
- Introduced a new display format for the summary results of the Oracle AutoML pipeline, enhancing the user experience.

Technocolabs Software, India (Remote) July 2022 - Aug 2022

Machine Learning Engineer Intern *Supervised by Mr. Yasin SHAH*

- Trained and deployed a Logistic Regression model which measures the creditworthiness of borrowers in the context of Peer To Peer Lending Market.

Research Works

Computing the Vapnik Chervonenkis Dimension for Non-Discrete Settings Preprint version

- Under the guidance of Prof. Anselm Blumer and in collaboration with colleagues, we conducted research on Vapnik-Chervonenkis (VC) computing, aiming to explore its dimensions in non-discrete settings for presentation at NeurIPS 2023. However, despite our rigorous efforts, our paper was unfortunately desk-rejected.
- Recognizing limitations from our initial assumptions on infinite H and X, we're refining our approach based on Linial et al.'s finite H and X insights. We aim to introduce more balanced assumptions: H as infinite, X as finite.

High Performance Computing Applied to Logistic Regression: A CPU and GPU Implementation Comparison IEEExplore paper

- We successfully devised and executed a GPU-accelerated rendition of the Gradient Descent Logistic Regression algorithm. This accomplishment was acknowledged by its acceptance at The IEEE International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings).
- Our innovative GPU-accelerated algorithm employs model parallelism for faster execution than sklearn, striking a balance between efficiency and competitive f1 scores. We've also introduced a Python package "SwiftLogisticReg" to facilitate testing and exploration.

Projects

Top 2%: Twitter Sentiment Analysis Oracle Competition Project Report

- Trained a Multinomial Naive Bayes model to classify sports and politics tweets, surpassing the benchmark solution offered by Oracle.

Unveiling the Twittersphere: Community Detection Analysis Github repository

- Applied Edge-based and Feature-based approaches, on Stanford Network Analysis Project (SNAP) Twitter graph data to generate meaningful clusters with labeled themes, such as 'Social Media','Gaming' and 'Music'.

Arabic Speech To Moroccan Sign Language Web Application Github repository

- Implemented a Python-based Arabic Speech (AS) to Moroccan Sign Language (MSL) speech-to-video translator using Automatic Arabic Speech Recognition.
- Developed a Messenger-like web app with Flask, HTML, CSS, and JavaScript, incorporating NLP techniques for Arabic speech preprocessing and an OpenCV-based MSL video retriever and concatenator.

Open Domain Arabic-English Question Answering System Github repository

- Developed a system utilizing TF-IDF for Document Retrieval, a DistilBert-based Document Reader and Answer Scorer all fine-tuned via knowledge distillation on SQuAD v1.1. This system efficiently answers factual questions by leveraging Wikipedia's corpus.
- Developped a web application resembling a Messenger interface, providing users with the QA system.

Training Neural Networks with Firefly Optimization Algorithm GitHub Repository

- Conducted a comparative analysis between the Firefly Optimization Algorithm (FOA) and the classical approach of gradient descent in optimizing a Neural Network.

Machine Learning From Scratch Github repository

- Implemented diverse machine learning algorithms, including Single Layer Perceptron, Pocket Perceptron, Adaline Algorithm, Linear Regression, Logistic Regression, Polynomial Regression, One-vs-All and One-vs-One Classifiers, along with algorithms featuring Dense Random and Hamming decoding. Furthermore, the package encompasses Cross Validation utilizing K-Fold and Regularization techniques for Linear and Logistic Regression.

PyOptim: Python Package for Numerical Analysis and Optimization Github repository

- Developed a Python package that integrates approximately ten unconstrained optimization algorithms, featuring 2D and 3D visualizations for comparative analysis, and includes matrix operations such as inverse, decomposition and solutions for linear systems.

Technical Skills

Programming Languages: Proficient : Python. Prior knowledge: C/C++, Java, JavaScript, SQL, HTML, CSS.
Software Tools: PyTorch, scikit-learn, TensorFlow, Keras, Cuda, Transformers, Flask, Git.
Languages: Fluent: English (TOEFL iBT Score: 105), French. Native: Arabic
Independent. Passion for Problem Solving. Critical Reading of Research Papers.

Volunteer Work and Awards

ENSIAS | *Head of research and development cell* 2022-2023
As a member of the Artificial Intelligence club at ENSIAS, I co-founded the (R&D) cell, where I spearheaded weekly sessions to elucidate data science and machine learning concepts for freshman students.

National Merit Scholarship (ISTIHQAQ) 2019-2022
"ISTIHQAQ" is an academic merit scholarship for children of Education-Training staff, awarded annually to the top baccalaureate graduates.