Zeyad El-Sayed

zeyadusf | zeyadusf | huggingface.co/zeyadusf | ziayd.usf@gmail.com | 01023504597
Zagazig, Egypt | Military Service: Completed

Technical Summary

AI/ML Engineer with experience in building and fine-tuning NLP, LLM, computer vision, and GANs, performing data analysis, preprocessing, handling imbalanced datasets, and implementing Autoencoders and DL algorithms.

Skills

Programming Languages: Python, C++
Frameworks & Libraries: PyTorch, TensorFlow/Keras, Hugging Face, Scikit-learn, Pandas, NumPy, OpenCV, Matplotlib, PyCaret
Backend & APIs: FastAPI, RESTful APIs ,Django (Basic)
DevOps & Deployment: Docker
Developer Tools: Git, Jupyter, VS Code
Machine Learning: Supervised/Unsupervised ML, Feature Engineering, Model Evaluation, Imbalanced Data Handling, PEFT (LoRA, QLoRA)
Deep Learning: CNNs, RNNs, LSTMs, Transformers, Autoencoders, DCGANs
NLP & LLMs: Tokenization, Attention, Transformers, Fine-Tuning, Embeddings
Computer Vision: Image Denoising, GANs
Data Engineering & BI: SQL, Tableau, Power BI, Excel

Experience

Information Technology Institute (ITI) Jul 2023 – Sep 2023
Business Intelligence Trainee
- Learned SQL, data modeling, and data warehouse concepts.
- Built dashboards using Tableau and Power BI.
- Applied BI tools to real-world analytics tasks.

Projects

Catch the AI | *Python, Django, React, PostgreSQL* GitHub | *Project Website*
- Built ViT and EfficientNet models to detect AI-generated images with **96%** accuracy for media verification.
- Built an NLP-based detector (DeBERTa, RoBERTa) with **95%** accuracy in spotting AI-generated text.
- Implemented Wav2Vec 2.0 for speech detection, achieving **90%** accuracy in identifying AI-generated audio.

LLMs from Scratch | *PyTorch, NLP* GitHub
- Built (LLM) from scratch with tokenization, attention mechanisms, and transformer blocks.
- Implemented end-to-end training pipelines, including optimization and evaluation.
- Acquired deep understanding of modern NLP architectures and transformer-based design principles.

Fine-Tuning LLMs | *PyTorch, PEFT techniques, HuggingFace, NLP* GitHub | *HF Spaces*
- Curated a fine-tuning repository and rapidly deployed models on Hugging Face Spaces for text classification, generation, and summarization.
- Applied LoRA and other PEFT techniques to optimize NLP models efficiently.
- Evaluated models using ROUGE-L, BLEU, and additional metrics, achieving 95%+ performance across tasks.

Imbalanced Data Treatment | *Python,imblearn sklearn, matplotlib, seaborn* GitHub
- Performed data analysis on imbalanced datasets and benchmarked models before applying sampling.
- Applied multiple sampling techniques (Over/Under, SMOTE, TomekLinks, SMOTE-Tomek)
- Tested KNN, RF, SGD, XGB, DT before/after sampling, achieving clear boosts in recall and F1-score.

Credit Score Classification | *Python, PyCaret, Tableau, sklearn* GitHub
- Cleaned and analyzed data with heavy missing/wrong values using custom code, statistical methods, and Tableau for visualization.
- Built a solid preprocessing pipeline that improved data quality and model performance.
- Trained models used PyCaret library to train models (Extra Trees, RF, XGBoost, KNN, DT).

Education

Benha University Oct 2020 – Jun 2024
B.S. in Computers & Artificial Intelligence *GPA: 3.4*