

Summary

I am a third-year Information Technology student with a strong interest in Data Engineering. I have good knowledge of data management, data analysis, and programming. I have worked on different projects using tools and technologies like Apache Hadoop, and SQL. I have also built machine learning models using CNN and ANN for image classification tasks. I am always willing to learn new technologies to improve my skills and become better in my work.

Skills

- **Programming Languages:** Python, C++
 - **Query Languages:** SQL(SQL Server, MySQL)
 - **Big Data Technologies:** Hadoop Ecosystem
 - **Database Management:** Relational Databases (SQL Server, MySQL), NoSQL (MongoDB)
- **Data warehousing:** Snowflake
 - **Data Visualization:** Power BI
 - **Machine Learning:** Feature engineering, data preprocessing (Scikit-learn, TensorFlow basics)
 - **Deep Learning:** CNN, ANN
 - **Spreadsheet Tools:** Microsoft Excel

Projects

REAL-TIME FACE RECOGNITION - PERSONAL PROJECT

Mar 2025 - Apr 2025

- Description: Developed a real-time face recognition system using webcam video input.
- Technologies: OpenCV (video capture), MTCNN (face detection), FaceNet (embedding), SVM (classification).
- Responsibilities: Captured live video using OpenCV; detected and cropped faces with MTCNN; extracted 512-vector dimensional embeddings via FaceNet; trained SVM classifier with 94% test accuracy, improved to 97% after tuning C and gamma.
- Outcome: Accurately identified and labeled known individuals in real-time video stream.
- Project link: https://github.com/Huysdfghf/project_facial_recognition

BRAIN CANCER CLASSIFICATION - TEAM PROJECT (3 MEMBERS)

Dec 2024 - Jan 2024

- Dataset: Brain Tumor MRI dataset from Kaggle
- Description: Built a CNN model based on the Xception architecture to classify brain MRI images into four types: glioma, meningioma, pituitary, and no tumor.
- Responsibilities:
- Data Preprocessing & Augmentation: Processed 7,022 MRI images with resizing, denoising, and pixel normalization. Applied augmentation techniques (rotation, flipping, elastic deformation) using TensorFlow/Keras.
- Model Design & Optimization: Implemented a CNN based on Xception architecture. Tuned hyperparameters (learning rate, batch size) via GridSearchCV and EarlyStopping, improving accuracy from 92% to 98.9%.
- Training & Evaluation: Trained on 5,712 augmented images over 30 epochs using Adam optimizer. Achieved 99.2% validation accuracy; precision and recall exceeded 98% across all classes (evaluated with confusion matrix and classification report).
- Outcome: Achieved 98% accuracy on the validation set.
- Project link: <https://github.com/Huysdfghf/brain-cancer-classification>

Education

VAN LANG UNIVERSITY - HCM, VIETNAM
OCT 2022 — PRESENT
Majors : Information Technology - Data
GPA: 3.2

Reference

MSc. Nguyễn Thái Anh
Lecturer at Van Lang University