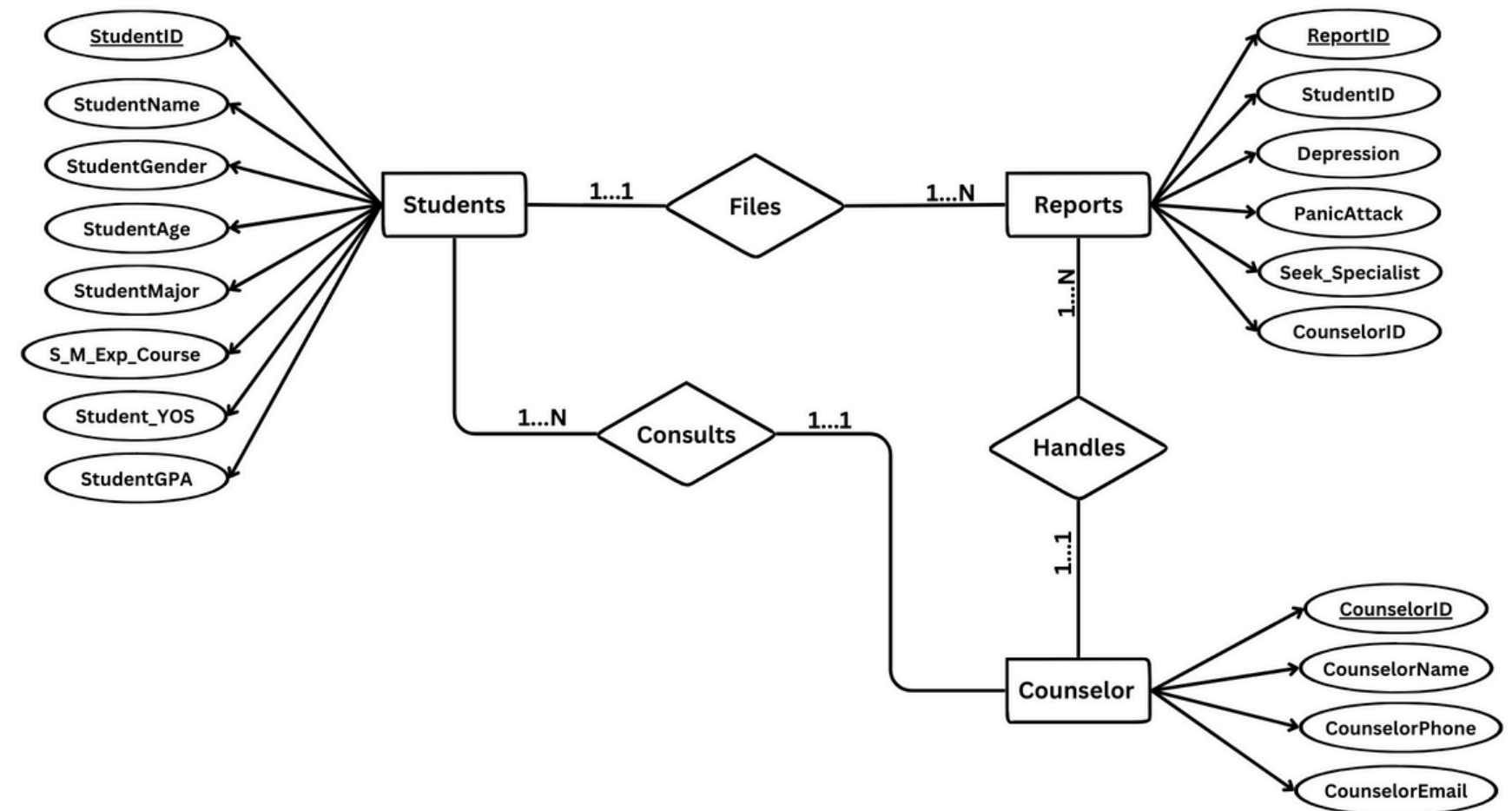# BRYAN MYER SETIAWAN'S PORTOFOLIO

ABOUT ME

Hello, my name is Bryan Myer Setiawan. I am a 6th semester Computer Science student at BINUS University with current GPA of 3.84/4. I am a student with a strong work ethic, dedication, and responsibility to develop myself in the areas of soft and hard skills. Focus on lecture studies and additional experience in the world of organizations. Interested in becoming an Artificial Intelligence Engineer, Data Scientist, and Data Analyst.

CURRICULUM VITAE

DATABASE SYSTEMS

Designing a reliable, efficient, and secure database that is proper and easy to use. This is done by eliminating functional dependencies on the database.

Project Report Link

DATA MINING

Learning to analyze and summarize information through given dataset. In this project I collaborate in a group project to merge 2 different dataset and get new insights on the newly-merged data.

Final Project Report Link

DATASET

insurance_claims.csv

Key Features:
- Categorical: policy_state, insured_sex, incident_type, fraud_reported.
- Numerical: age, policy_deductable, policy_annual_premium, total_claim_amount.
- Dates: policy_bind_date, incident_date.

Notes:
- Dataset 1 has missing values (to be dropped).
- Missing data is represented by ? in some categorical columns (e.g., property_damage).
- total_claim_amount ranges from small to significant values; outliers may exist.
- The fraud_reported column is a key focus, enabling fraud detection analysis.

carsData.csv

Key Features:
- Categorical: model, transmission, fuelType, Manufacturer.
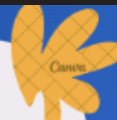- Numerical: year, price, mileage, mpg, engineSize.

Notes:
- Price Distribution is likely skewed, with luxury models influencing upper price limits.
- mpg varies by fuelType, likely higher for hybrids or smaller engines.
- Outliers in numeric columns like price and mileage since older cars generally have higher mileage, but exceptions exist for collector or luxury models.

CONCLUSION

- FRAUD DETECTION IN VEHICLE INSURANCE CLAIMS IS CHALLENGING DUE TO IMBALANCED DATASETS AND VARIED FEATURES.
- DATA PREPROCESSING TECHNIQUES, SUCH AS HANDLING MISSING VALUES AND NOISE, AND BALANCING THE DATASET WITH SMOTE, SIGNIFICANTLY IMPROVED MODEL PERFORMANCE.
- XGBOOST DEMONSTRATED THE BEST PERFORMANCE OVERALL, ACHIEVING THE HIGHEST ACCURACY AND BALANCED DETECTION ACROSS FRAUD AND NON-FRAUD CASES.
- THE APPLICATION OF SMOTE HELPED MITIGATE CLASS IMBALANCE, ENHANCING THE RELIABILITY AND FAIRNESS OF ALL MODELS.
- DECISION TREES AND RANDOM FORESTS ALSO SHOWED IMPROVED METRICS AFTER SMOTE, EMPHASIZING THEIR APPLICABILITY FOR FRAUD DETECTION.
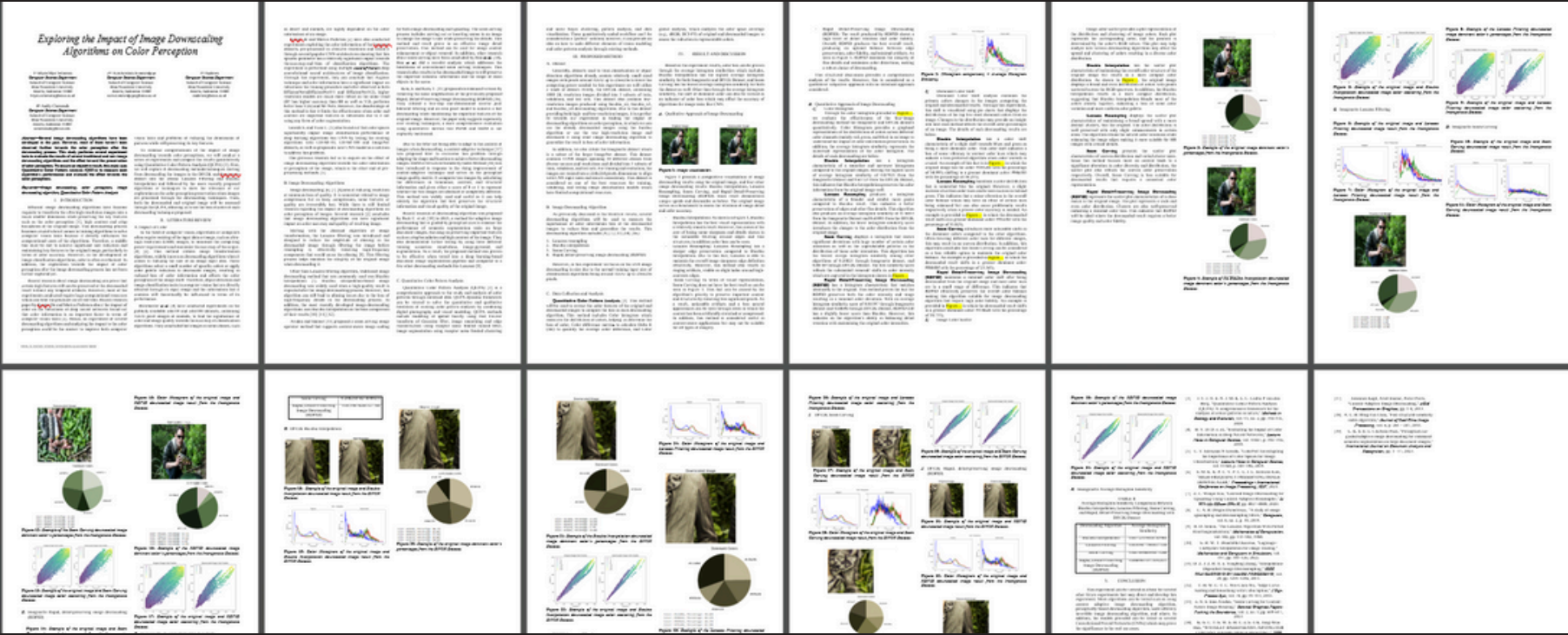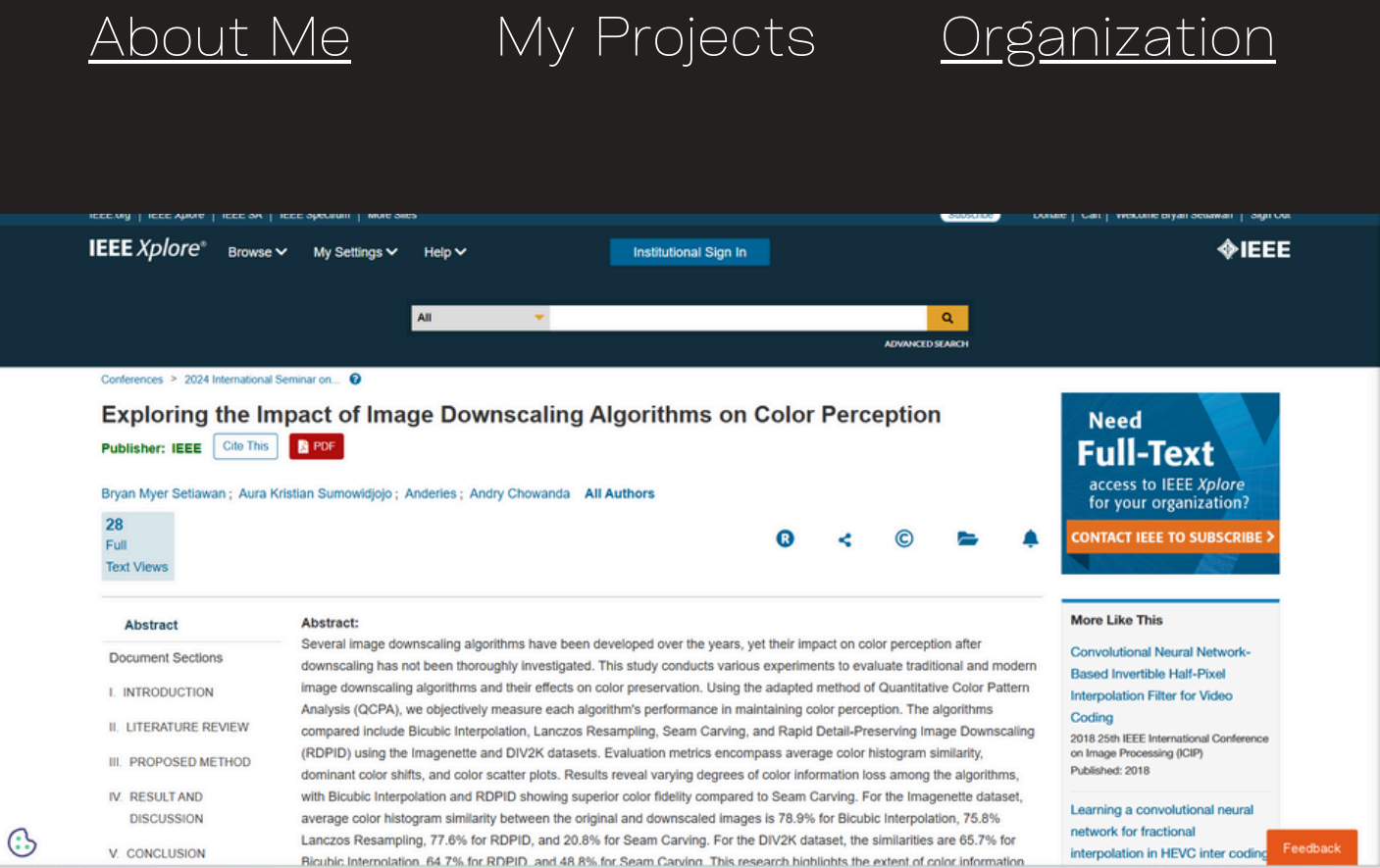
CLOUD COMPUTING

Certified Alibaba Cloud Associate that reflects on my understanding in cloud infrastructure and alibaba cloud products.

RESEARCH PAPER

Research paper on image downscaling algorithms and the effect on color perceptions. This paper was then accepted and published in IEEE website through isemantic 2024 Conference.
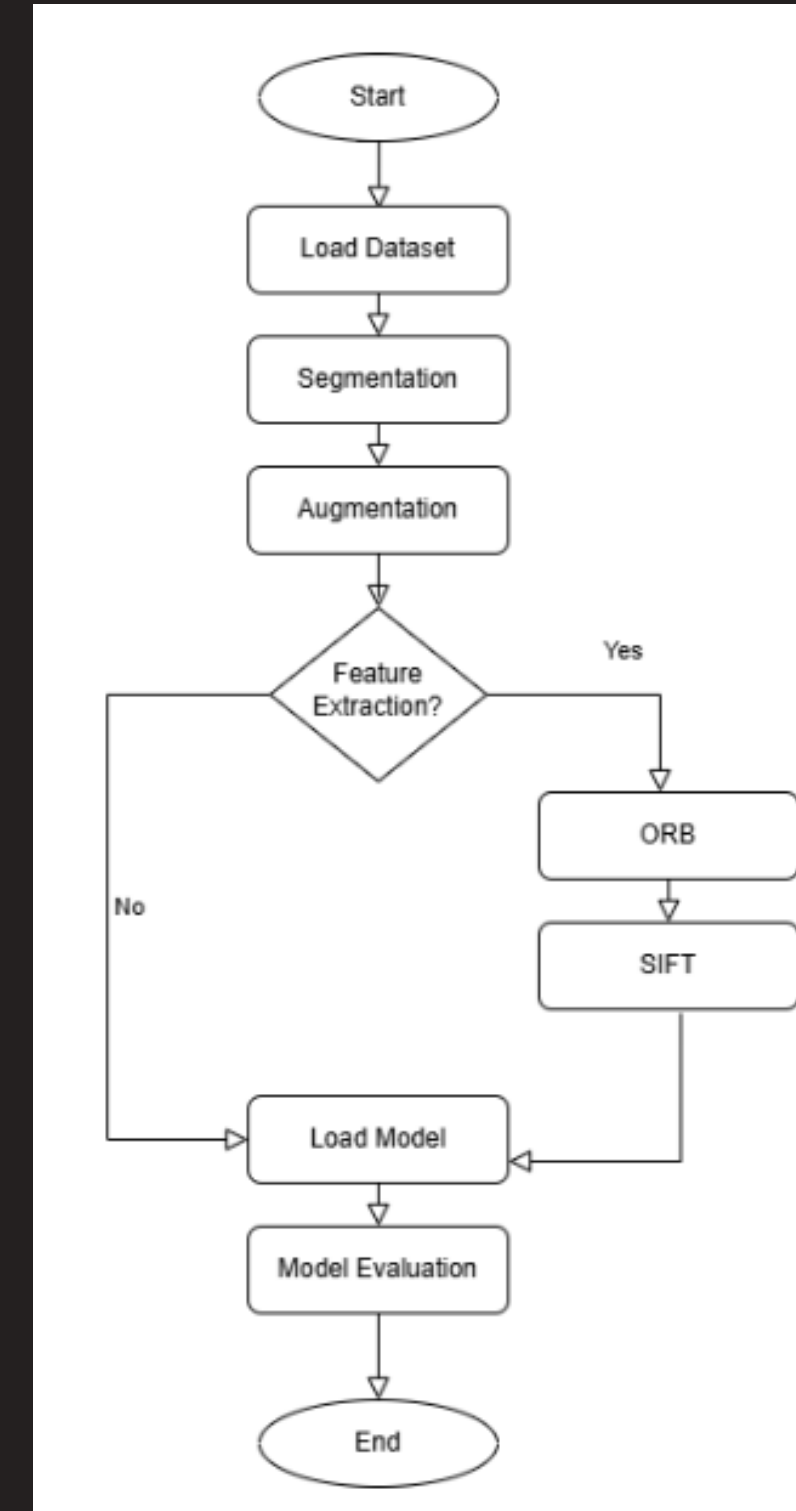
Research Paper Link

COMPUTER VISION

Utilizing Computer Vision Techniques to enhance Machine Learning Algorithms and data intepretation. In this project I contributed in the dataset finding, and report writing
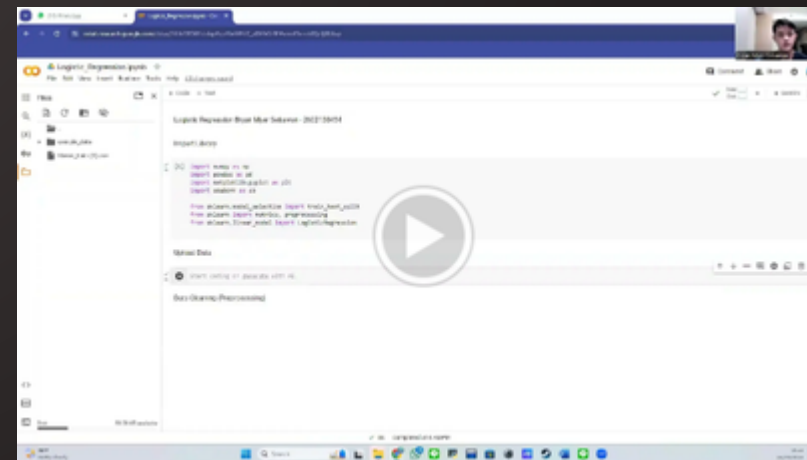
Final Project Report Link

ORGANIZATIONAL EXPERIENCE

BSLC General Secretary and Core Board of Directors (2024-2025)

Academic Scholarship Mentor (2024-2025)

BSLC Benchmarking Project Manager 2023

THANK YOU

Contact Details:

📞 +62818409888

✉️ bryanmyersetiawan@gmail.com

in Bryan Myer Setiawan