



BT4222 Mining Web Data for Business Insights

Source Code Links for Fraud Detection Project

Group No: Group 3

Academic Year 2023/2024 Semester 1

Name	Matriculation Number
Chan Tse Ee	
Cheang Xing Wei	
Mohamed Jasim	
Steffi Lim En Qi	
Lin Jiayong	

Source Code Links

Item	Description	Link
Project Proposal	Description of the project on detecting fraudulent transactions	https://docs.google.com/document/d/1S7c8dQ2mLMfh8F8MOEYkh3j-TEzZpaUswq_04c7bFUw/edit?usp=sharing
Datasets	Kaggle dataset source: Bank Account Fraud Dataset Suite (NeurIPS 2022)	https://www.kaggle.com/datasets/sjpgjesus/bank-account-fraud-dataset-neurips-2022
Selected Datasets	For our project, we are using only 3 of 6 datasets available: Base, Variant I and Variant II focusing on the group size and prevalence disparity	https://drive.google.com/drive/folders/1EMmF7BP5UzXtkKSZxNAPrzn6x_yKMS2?usp=drive_link
Source Code: Main	Notebook for all our work done excluding Post-Processing and In-Processing Bias Algorithms	https://colab.research.google.com/drive/1bx9VuKkU4Z2FlwdNijACa3m18FFDt9WA?usp=sharing
Source Code: Post-Processing	Notebook for Post-Processing Bias Algorithm	https://colab.research.google.com/drive/1TPH0JA-WKfKt2CiLVe-zaQxowJ8OcEfm#scrollTo=vsQWpdbv8Qld
Source Code: In-Processing	Notebook for In-Processing Bias Algorithm	https://colab.research.google.com/drive/116ELrJ8Rv7Jw7Br1cqK_T44aDo4NVjzH#scrollTo=42nwrrsHAzdw