8/10
Good job. You need to rethink how you will achieve evolving. One potential is using the latent space vectors and combining them. You can introduce new samples that way.

Student No:219013044
Name: Enock Onkarabile Buys
Honors year project Research Proposal
Academy of Computer Science and Software Engineering
Faculty of Science

Topic: Generative AI for Real-Time Fraudulent Transaction simulation

## 1.Problem at hand:

Existing Banking AI fraud detection models are having a hard time evolving while the increase in digital banking has made to fraud constantly increase because they are trained on past fraud cases(Nguyen et al., 2022) and are not smart enough to anticipate any new possible ways they might effectively use for future refences while policies and lack of data make it hard to train robust detection systems(Buczak and Guven 2015).

Generative adversarial networks offer a good solution by creating realistic fraudulent transactions that can magnify fraud detection models.

I aim to address certain gaps in this phenomenon to make my reseach stand out:

- Generate realistically evolving fraud patterns: not rely only on past fraud data
- Generative Adversarial networks instability and mode collapse: imitate real-world fraud variance
- Make explainable generated fraud transactions so that it can be trusted by the banks.

## 2.Task Type

- Generative modelling for synthetic fraud data augmentation
- Anomaly detection to Identify fraudulent transactions
- Benchmarking to compare fraud detection models trained with real data versus models trained with synthetic fraud data.

## 3.proposed solution

### 3.1 pre-processing

- Feature engineering using transactions timestamps , categories , geolocation and peoples spending behavioral patterns.

- Handling class imbalances by oversampling with synthetic data
- Prevent mode collapse in GANs by normalization techniques

### 3.2 Feature extraction

- Use deep learning-based and statistical feature extraction for fraud analysis
- Fix categorial transaction features into continuous space using Word2Vec/Autoencoders

### 3.3 implementing the GNAs

- Baseline GAN: Standard fraud data generation
- Conditional GAN: generation that is conditional on merchant types, location and transaction type
- Time-series GAN: get sequential fraud transaction patterns(Yoon, Jarrett et al. 2019)

### 3.4 Fraud classifier implementation

- A supervised fraud detection classifier will be implemented to validate the usefulness of synthetic fraud data.
- Options for classifiers include XGBoost, Random Forest, CNN, or RNN, trained using both real and synthetic fraud data.
- Model evaluation will assess generalization capabilities when synthetic fraud data is introduced.

### 3.5 Anomaly detection and benchmarking

- Train the models on real and synthetic fraud transactions
- Evaluate performance against existing fraud detection methods using precision-recall.
- compare different data augmentation approaches

It would be nice to know what features the dataset offers for training, how many samples etc

## 4. dataset(s)

- IEEE-CIS Fraud Detection Dataset (Kaggle)
- Synthetic Financial Datasets For Fraud Detection (PaySim)
- European Credit Card Fraud Dataset (Kaggle)
- Synthetic Data Vault (MIT SDV)
- Brazilian E-Commerce Public Dataset by Olist
- Elliptic Dataset (Bitcoin transaction fraud detection)
- Open Payments Fraud Dataset (Medicare Fraud)
- DSSG Financial Transaction Dataset
- Xente Fraud Detection Dataset (Mobile Money Transactions)

5.Evaluation matrix
GAN Metrics:

- Frechet Inception Distance (FID): Measures similarity between real and synthetic data.
- Kernel Inception Distance (KID): Evaluates the stability and diversity of generated fraud data.
- Precision & Recall for Distributions: Checks the quality and diversity of synthetic samples.
- Mode Score: Ensures that the GAN captures multiple fraud patterns.

Fraud Detection Classifier Metrics:

- Precision, Recall, F1-score, ROC-AUC to measure classifier accuracy.
- False Positive Rate (FPR) & False Negative Rate (FNR): Critical in financial fraud detection.
- Log Loss & Brier Score: Measures model confidence in fraud classification.

References:

Nguyen, G., Tran, M., & Pham, D. (2022). AI-based fraud detection in digital banking. Journal of Financial Security, 15(1), 45-60.

Buczak, A. L. and E. Guven (2015). "A survey of data mining and machine learning methods for cyber security intrusion detection." IEEE Communications surveys & tutorials **18**(2): 1153-1176.

Yoon, J., et al. (2019). "Time-series generative adversarial networks." Advances in neural information processing systems **32**.