Detection of anomalous behavior in Credit Card transactions with an One-Class Generative Adversarial Network

Finatix

Agenda

- Deep Learning for Anomaly detection
- Introduction Dataset
  - Feature Engineering
  - Dimensionality Reduction
- Approach: One-Class Generative Adversarial Network (OCGAN)
  - Autoencoder
  - Generative Adversarial Network
  - OCGAN Framework
- Training Process

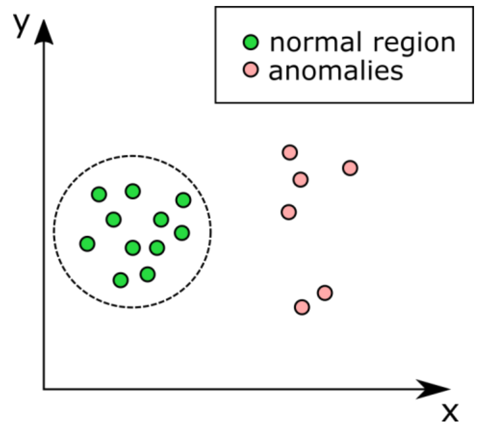
Why Anomaly Detection and Deep Learning?

- Research problem: Identification of suspicious transaction behavior

- Major approaches: Misuse detection vs. Anomaly detection

- Issues:
  - Highly-imbalanced datasets
  - Real-life transaction data

- Approach with Deep Learning:
  - Definition of a „normal region“
  - Generate synthetic fraud samples for training the model



Finatix

What are our data?

- Anonymized Credit card transactions from German and Austrian account holders which were obtained from a real-world application
- Highly-imbalanced: 248,612 transactions with 3252 fraud records (~1.3%)
- 42 primary attributes:
  - 26 non-numeric variables (21 categoricals & 5 datetime)
  - 6 numeric variables
- Problem: Feature may not be meaningful enough to depict the complex transaction behavior

Finatix

What is the added value of Feature Engineering?

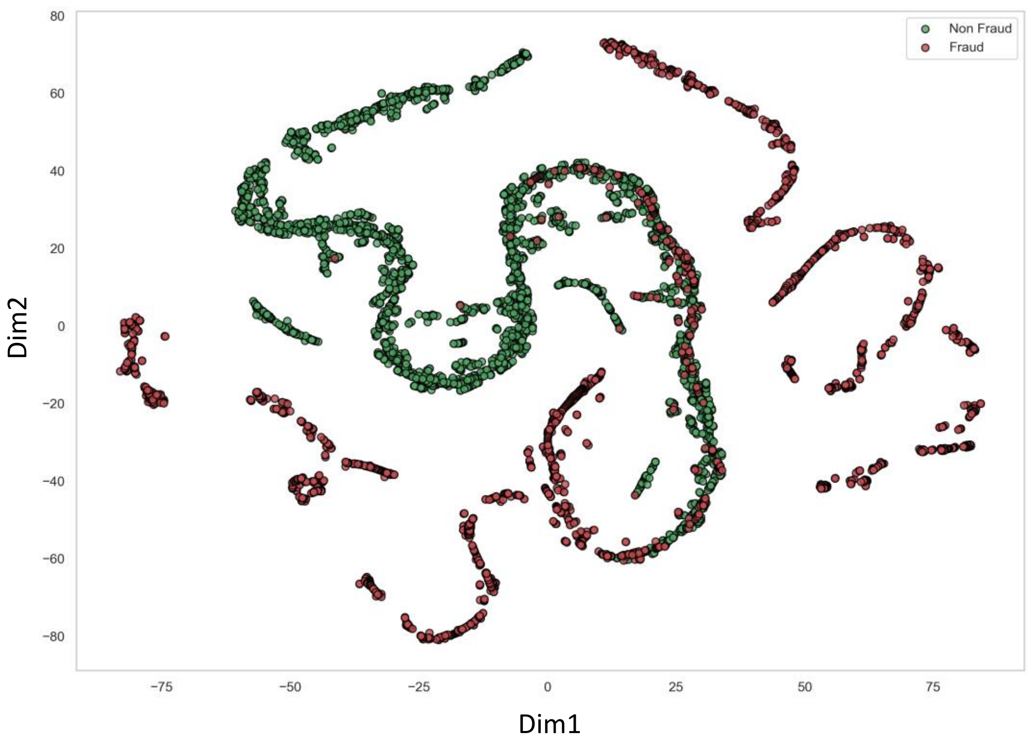
Attributes	Description
Year of birth	Birth year of the account holder
Provider	Associated provider to the registration email address
Payment transaction year	Year of the individual transaction
Payment transaction month	Month of the individual transaction
Payment transaction day	Day of the individual transaction
Payment transaction hour	Hour of the individual transaction
Minutes between transactions	Difference in minutes between transactions of the same account
Count of transaction last 24h	Number of all transactions of an account within the last 24 hours
Transaction frequency same merchant	Number of all transactions of an account with the same merchant
Load frequency last 24h	Number of all account recharges within the last 24 hours
Loadings above historic median	Number of all account recharges above the historic median value
Transaction amount over month	Average amount per transaction during the last month
Transaction amount over past three months	Average weekly transaction amount over the last quarterly period
Transactions with same merchant over month	Number of transactions with same merchant over last month
Daily transaction amount same merchant	Average daily transaction amount spent with same merchant over last month
Transactions with merchant same day	Number of transactions for a specific merchant on same month day

- Applying domain knowledge onto the data
- Support the model's learning by providing derived attributes
- Decompose features to reduce categorical variance
- Mapping of transaction behavior

Finatix

t-SNE: who's your neighbor?

2D-Dimensionality Reduction t-SNE

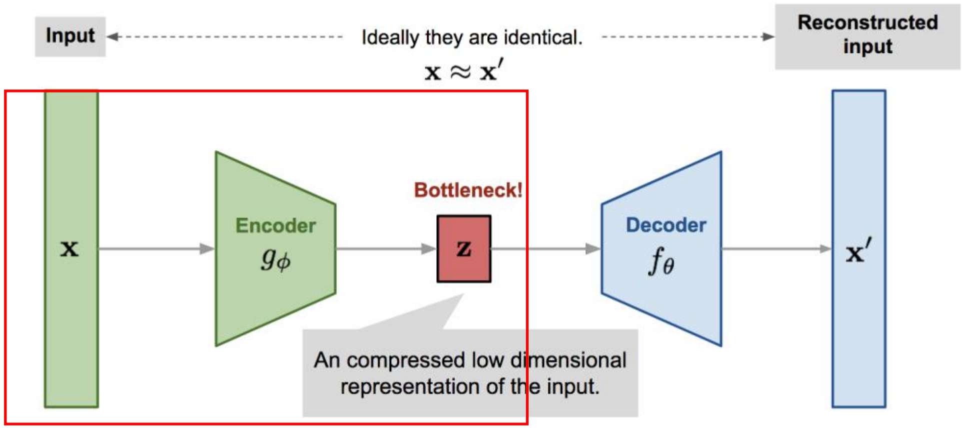


- Dimensionality Reduction for high-dimensional feature spaces
- Findings:
  - Frauds location in the outer regions
  - Significant diffusion of fraudulent and normal instances
- <u>How to separate them?</u>

Finatix

Using an Autoencoder to separate classes

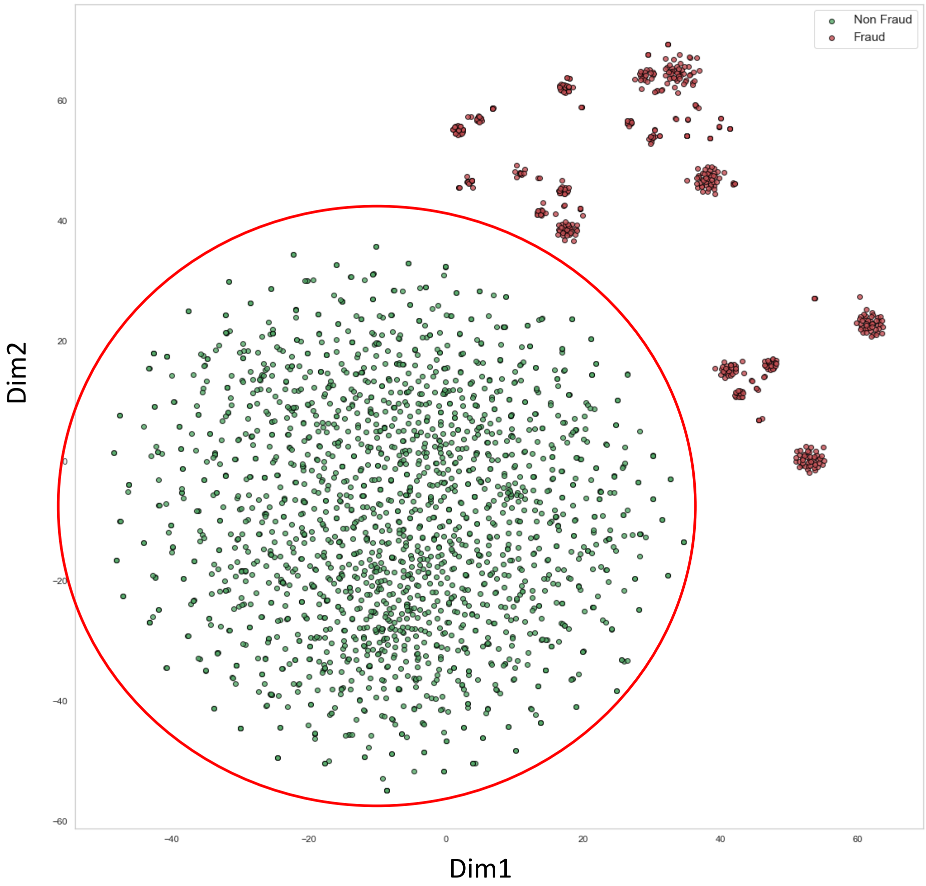
- Standard-Feedforward Neural Network → transformation of feature vectors into representations
- Learning representations of the latent structures in the given data
- Well-suited for the definition of an area in which only normal instances are located



[see https://lilianweng.github.io/lil-log/2018/08/12/from-autoencoder-to-beta-vae.html]

Learned Representations by the Autoencoder

2D-Dimensionality Reduction t-SNE

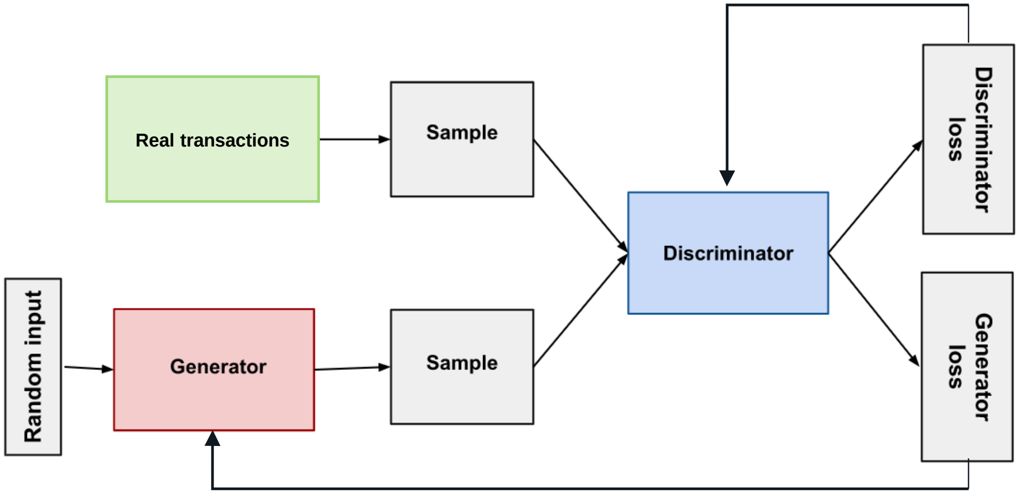


- Visualization of latent representations
- Encoded instances allow proper distinction
- <u>Solved: Definition of a normal region – done!</u>

Finatix

Handling class imbalances with Generative Adversarial Network

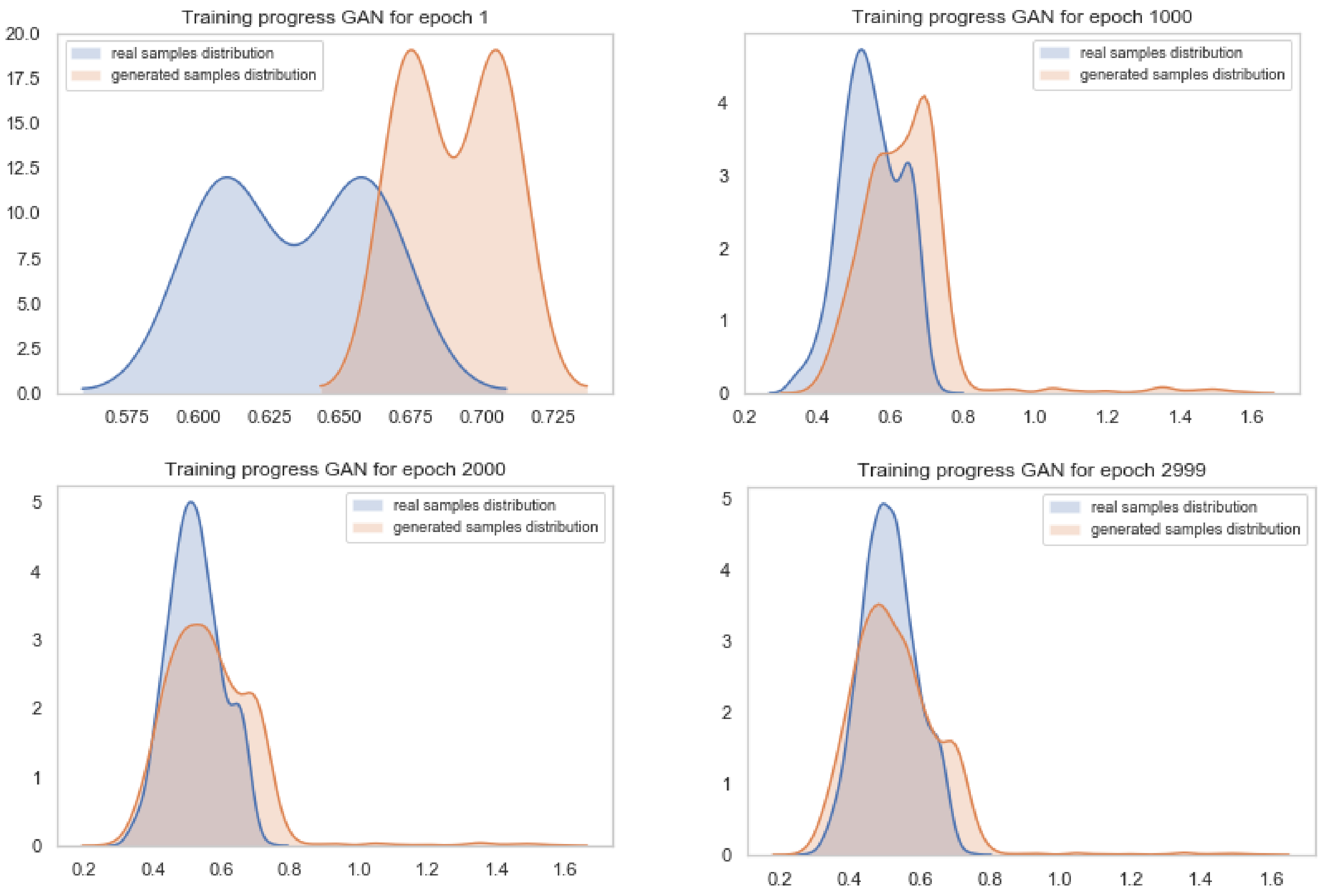
- GAN's architecture consists of two adversary Neural Networks
- Generator and Discriminator competing against each other in a minimax-game [Nash, 1951]
- Discriminator Network as a simple classifier | Generator Network creates synthetic samples



[see https://developers.google.com/machine-learning/gan/gan_structure]

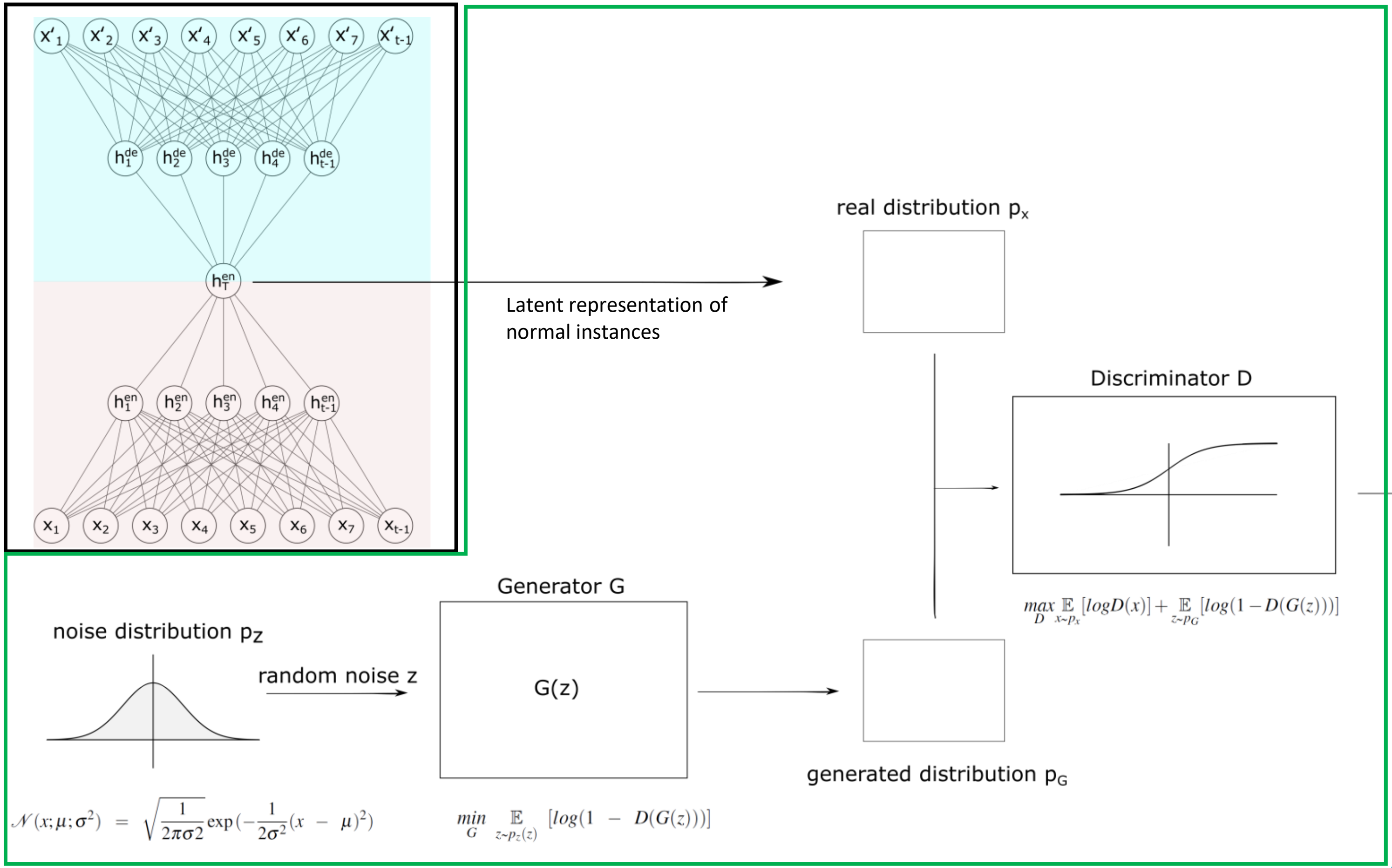
Finatix

OCGAN's generated synthetic frauds



Finatix

Design of the OCGAN-Framework



Finatix

Thanks for your attention!

Tim Schmidt

Position: Software Engineer

E-Mail: t.schmidt@finatix.de

Finatix GmbH

Barfußgäßchen 12 | 04109 Leipzig

Finatix