Credit Card Fraud Detection

Final Presentation (01.02.2024)

Module: Deep Learning Lab
Supervisor: M.Sc. Rodrigo Lopez Portillo Alcocer
Students: Hasan Evci and Tareq Abu El Komboz

Outline

Introduction

- Motivation
- Anomaly Detection
- Dataset

Motivation

- The harm caused by credit card fraud cases is increasing worldwide
- **Goal:** Distinguish malicious transactions from normal ones
- **Challenge:** High similarity between them

Source:
Report et al., 2021

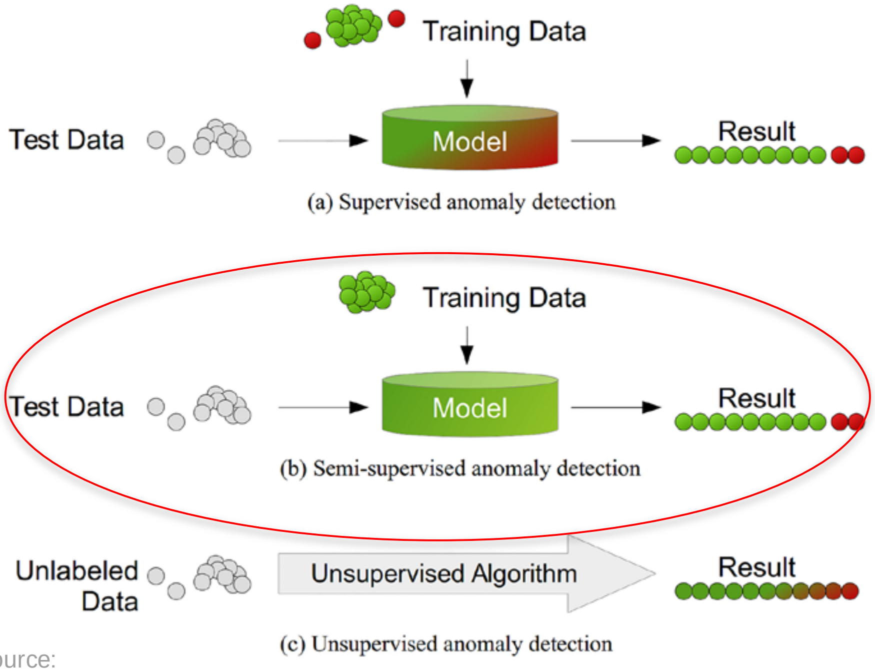
Anomaly Detection

*"An anomaly is an observation that **deviates considerably** from some **concept of normality.**"*

- Data Space $\mathcal{X} \subseteq \mathbb{R}^D$
- Ground-Truth Law of normal behavior $\mathbb{P}^+$ on $\mathcal{X}$
- Set of anomalies $\mathcal{A} = \{\boldsymbol{x} \in \mathcal{X} \mid p^+(\boldsymbol{x}) \leq \tau\}, \quad \tau \geq 0$

Dataset

- Credit Card Fraud Detection Dataset 2023
    - Over 550,000 records
    - Binary classification
- Challenges
    - High-dimensional
    - Anonymized
    - Semi-supervised

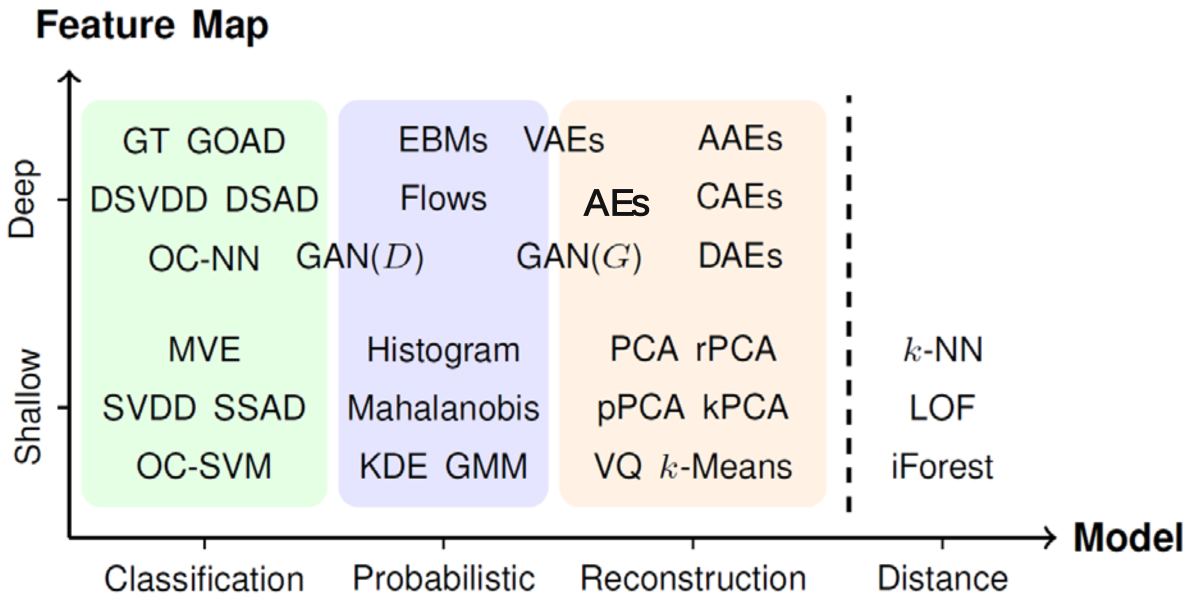


Source:
Goldstein and Uchida., 2016

Methods

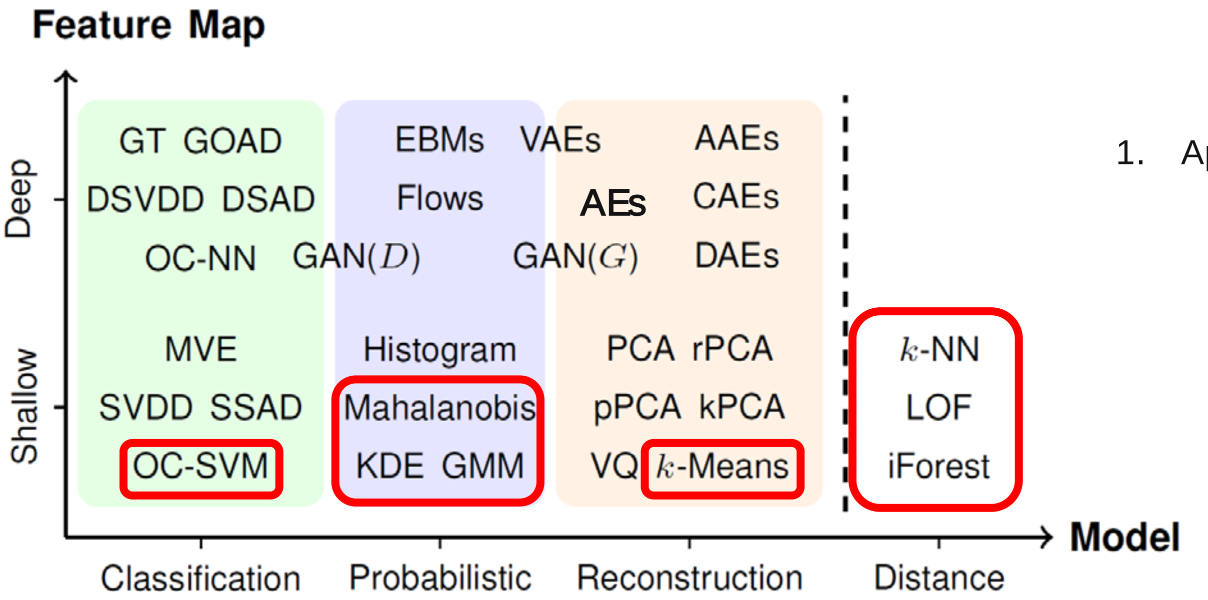
- Approach
- Autoencoder for Anomaly Detection
- One-Class Neural Network

Our approach



Source: Lukas Ruff et al., A unifying review of deep and shallow anomaly detection., Proc. IEEE, 109(5):756–795, 2021.

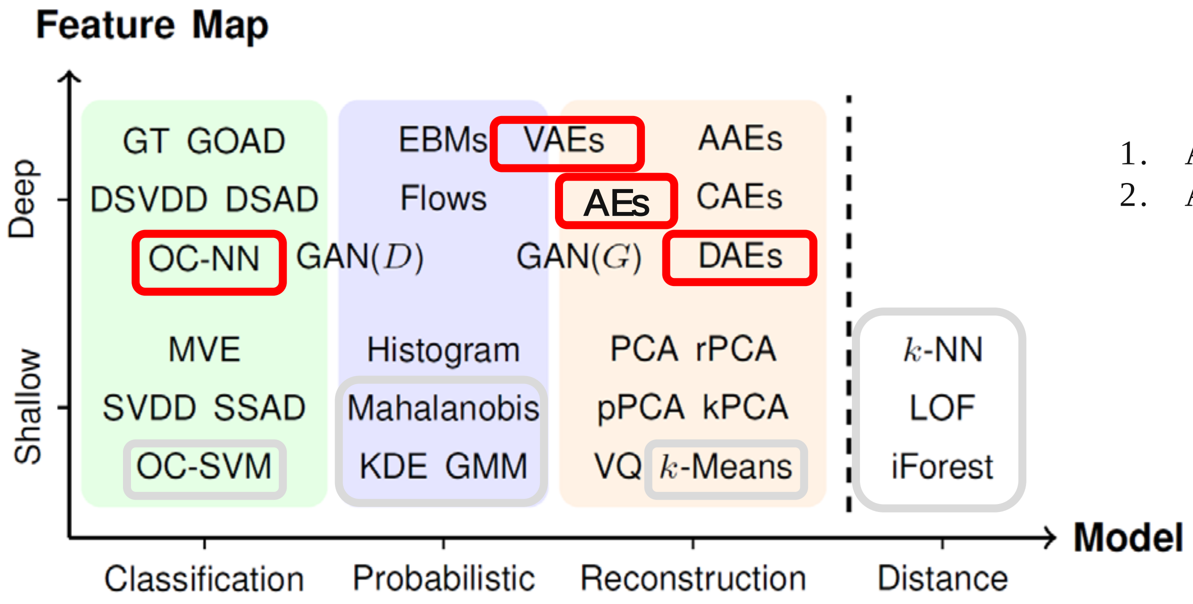
Our approach



1. Apply baselines

Source: Lukas Ruff et al., A unifying review of deep and shallow anomaly detection., Proc. IEEE, 109(5):756–795, 2021.

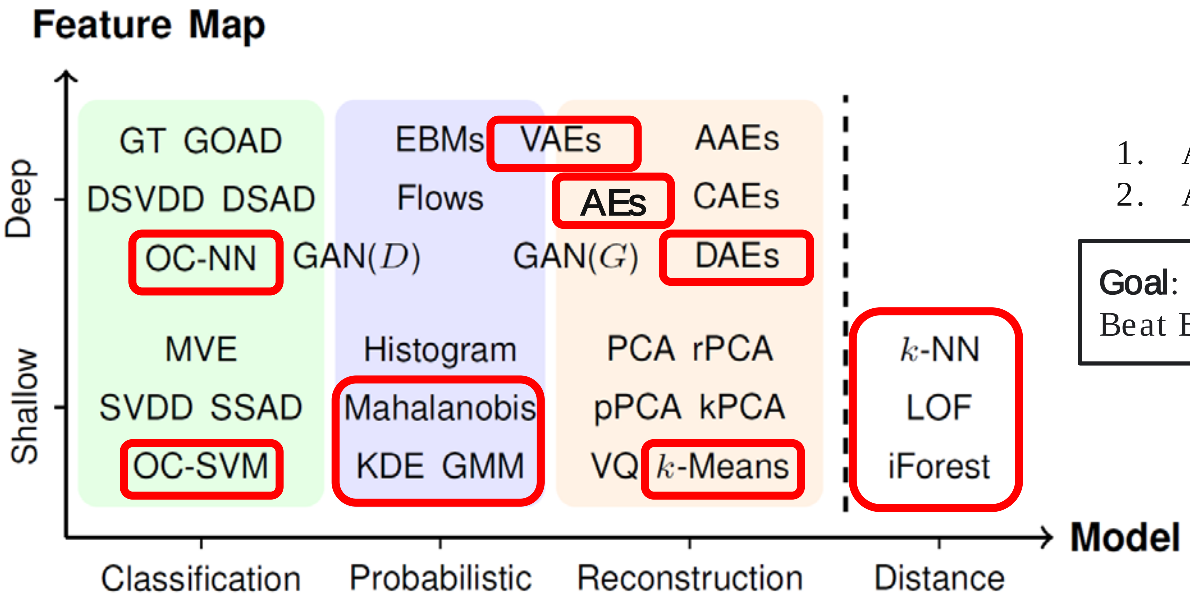
Our approach



1. Apply baselines
2. Apply DL models

Source: Lukas Ruff et al., A unifying review of deep and shallow anomaly detection., Proc. IEEE, 109(5):756–795, 2021.

Our approach



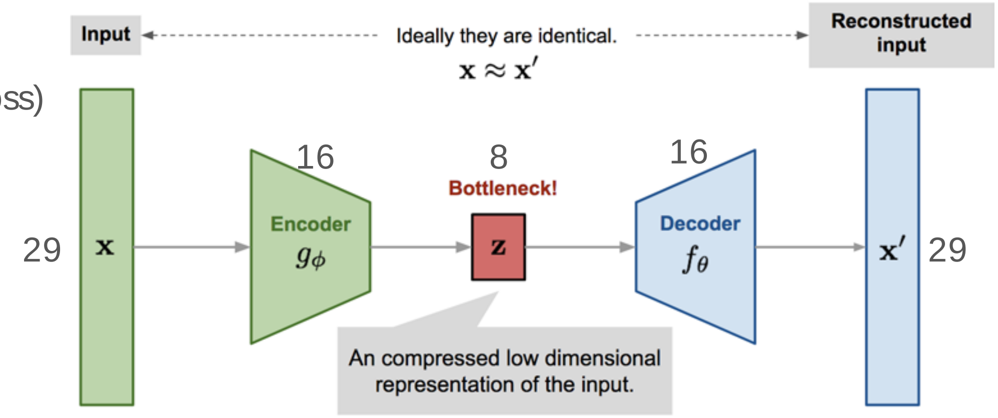
1. Apply baselines
2. Apply DL models

Goal:
Beat Baselines with DL

Source: Lukas Ruff et al., A unifying review of deep and shallow anomaly detection., Proc. IEEE, 109(5):756–795, 2021.

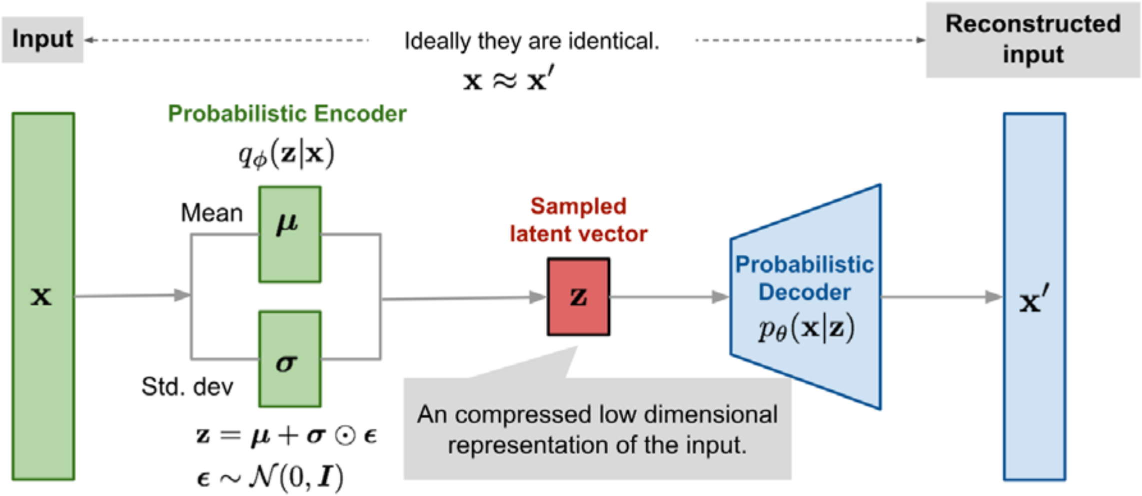
Autoencoder for Anomaly Detection

- **Idea**: Learn to reconstruct normal data patterns and identify anomalies as data points with significantly higher reconstruction errors
- **Encoder-Decoder Architecture**
  - ELU activation for all layers
- **Loss function**: MSE (Reconstruction loss)
- **Regularization**:
  - Early Stopping
  - Batch Normalization
- **Optimization method**: Adam



Source: https://lilianweng.github.io/posts/2018-08-12-vae/

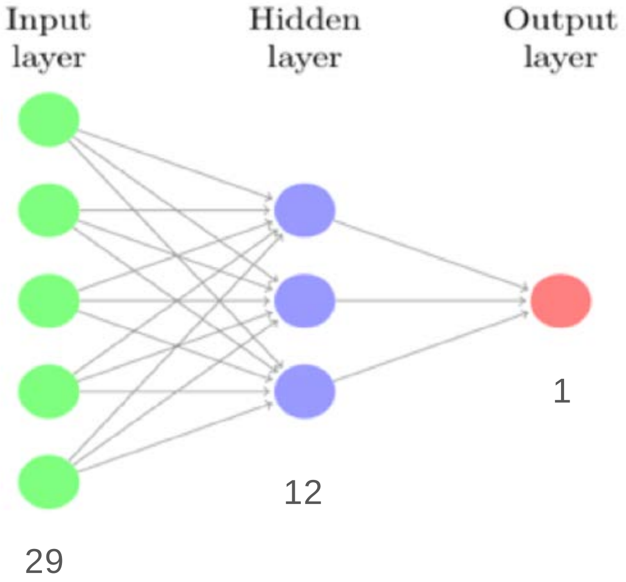
Variational AE for Anomaly Detection



Source: https://lilianweng.github.io/posts/2018-08-12-vae/

One-Class Neural Network (OC-NN) for Anomaly Detection

- **Idea**: learn feature space where normal and anomalous data are easily separable
  - cluster normal data points together
  - high scores to normal data, low scores to anomalies
- **Architecture**: Feed-Forward Neural Network
  - Sigmoid activation
- **Loss function**: Similar to OC-SVM's loss function



Source: Raghavendra Chalapathy, Aditya Krishna Menon, and Sanjay Chawla. Anomaly detection using one-class neural networks. CoRR, abs/1802.06360, 2018.

OC-SVM vs. OC-NN

Problem of (hybrid) OC-SVM:

- Separates feature learning (autoencoder) and anomaly detection (OC-SVM)

Solution of OC-NN:

- Combines deep learning for feature extraction with anomaly detection

OC-SVM Loss function:

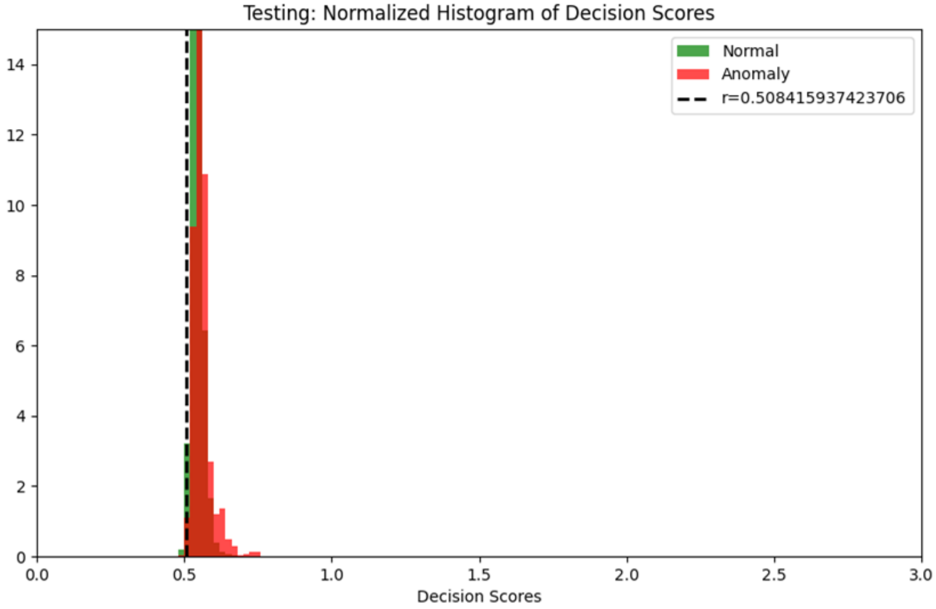
$$\min_{w,r} \frac{1}{2}\|w\|_2^2 + \frac{1}{\nu} \cdot \frac{1}{N}\sum_{n=1}^{N} \max(0, r - \langle w, \Phi(\mathbf{X}_{n:})\rangle) - r$$

OC-NN Loss function:

$$\min_{w,V,r} \frac{1}{2}\|w\|_2^2 + \frac{1}{2}\|V\|_F^2 + \frac{1}{\nu} \cdot \frac{1}{N}\sum_{n=1}^{N} max(0, r - \langle w, g(VX_{n:})\rangle) - r$$

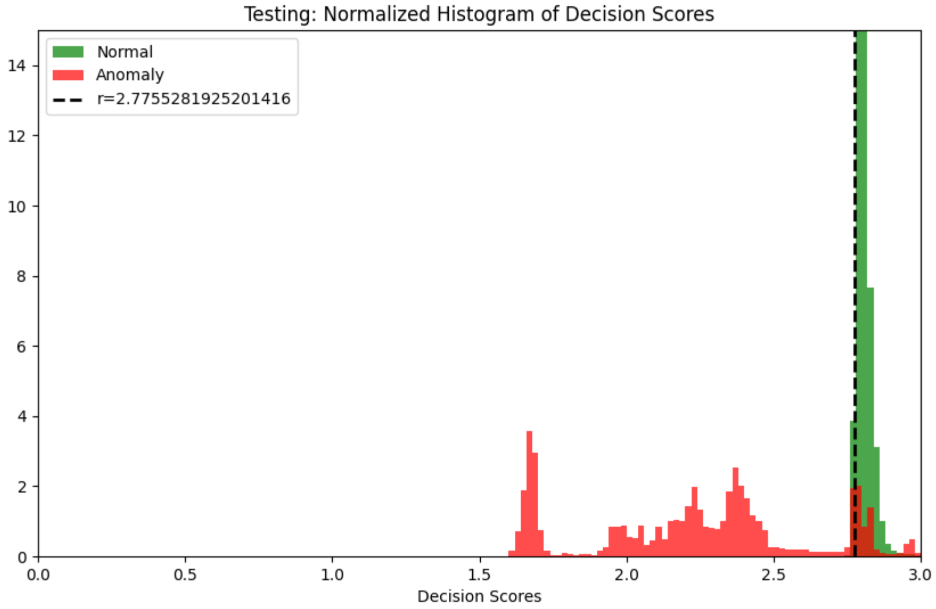
OC-NN: Histogram of Decision Scores

decision scores
after 1 training epoch



OC-NN: Histogram of Decision Scores

decision scores
after 5 training epochs



Results

- Model Performances
- Explainability: Shapley Values

Metrics

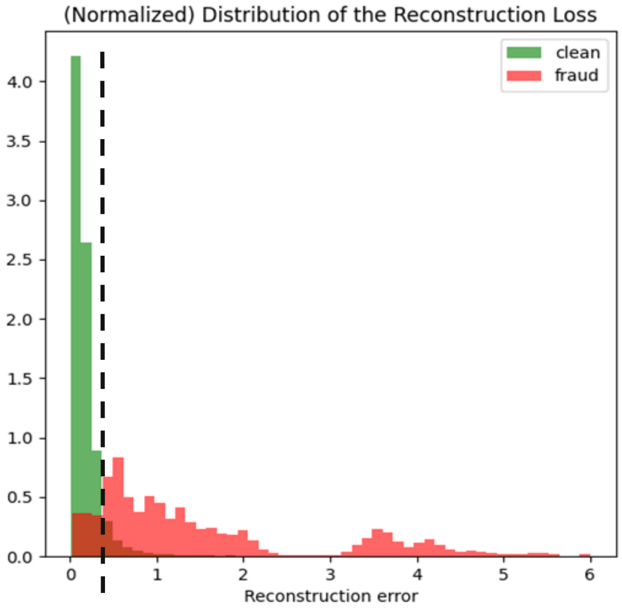
- Precision = $\frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$
- Recall = $\frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$
- F1-score = $2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$
- MCC $= \frac{\mathrm{TP} \times \mathrm{TN} - \mathrm{FP} \times \mathrm{FN}}{\sqrt{(\mathrm{TP} + \mathrm{FP})(\mathrm{TP} + \mathrm{FN})(\mathrm{TN} + \mathrm{FP})(\mathrm{TN} + \mathrm{FN})}}$
- AUC-ROC (Area under the ROC Curve)
- AUC-PRC (Area under Precision-Recall Curve)

		True Class	
		Fraud	Normal
Predicted Class	Fraud	TP	FP
	Normal	FN	TN

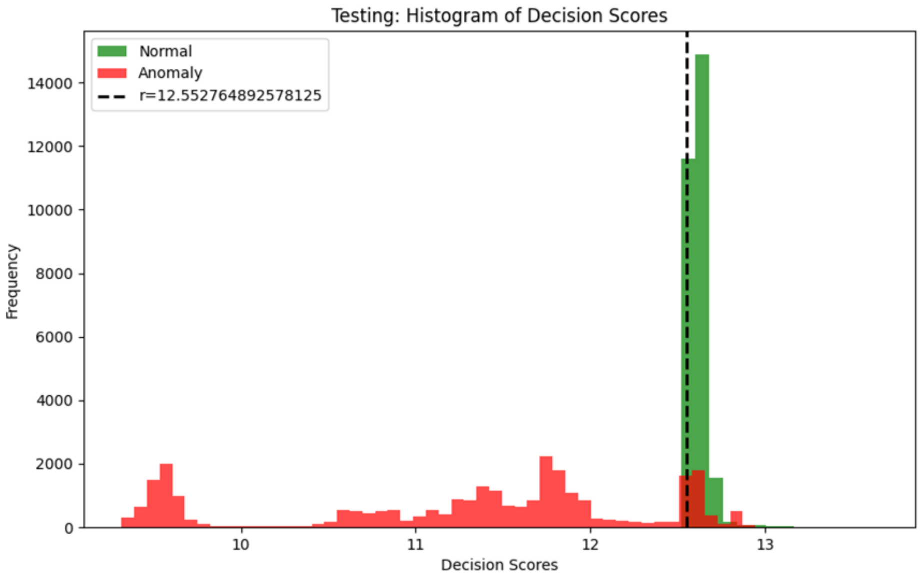
Model Performances

	Model	Precision	Recall	F1-Score	ROC-AUC	MCC
Baselines	Naive	0.5	1.0	0.66	0.5	0
	K-Means (k=2)	**1.0**	0.764	0.866	–	0.786
	OC-SVM	0.653	0.937	0.77	0.89	0.49
	SGD OC-SVM	0.682	**0.981**	0.805	**0.973**	0.582
	Isolation Forest	0.946	0.64	0.763	0.917	0.638
	kNN distance	0.937	0.842	0.887	0.949	0.79
	LOF	0.424	0.424	0.424	0.396	-0.152
	Mahalanobis Distance	0.945	0.89	0.917	0.947	0.84
	GMM	0.958	0.859	0.905	0.958	0.825
	KDE	0.682	0.978	0.803	0.941	0.579
DL Models	Autoencoder	0.899	0.895	0.896	0.941	0.793
	DAE	0.896	0.882	0.889	0.93	0.78
	VAE	0.891	0.828	0.859	0.918	0.729
	OC-NN	0.974	0.871	**0.92**	0.929	**0.853**

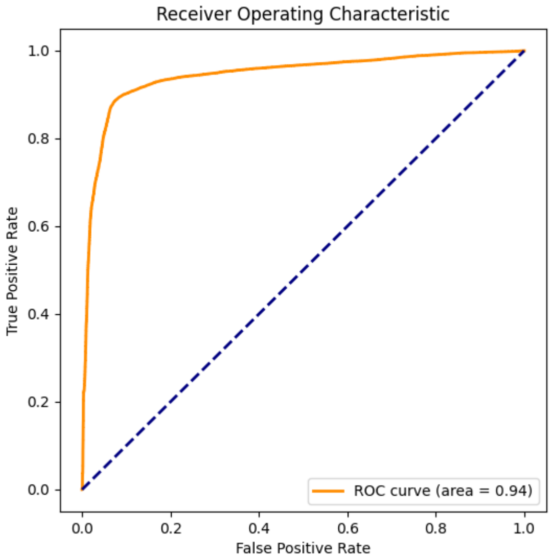
Autoencoder and OC-NN: Decision Boundary

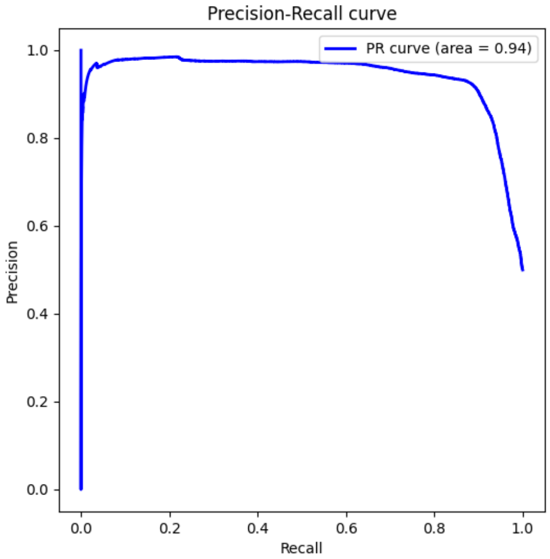


90% percentile

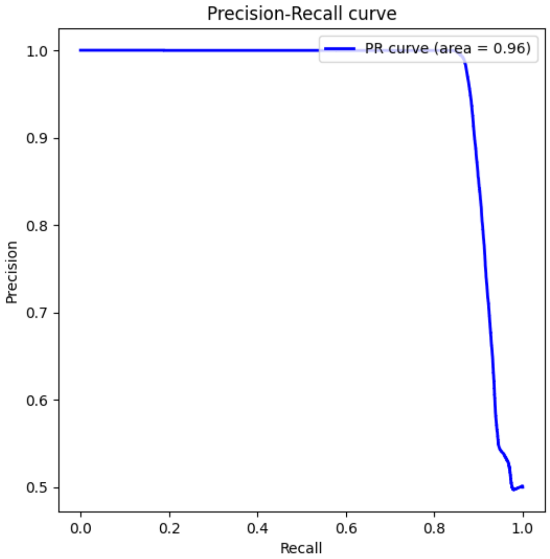


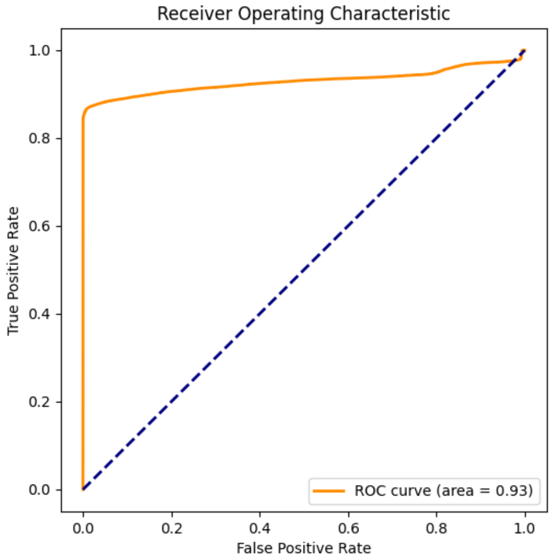
Autoencoder: AUROC and AUPRC



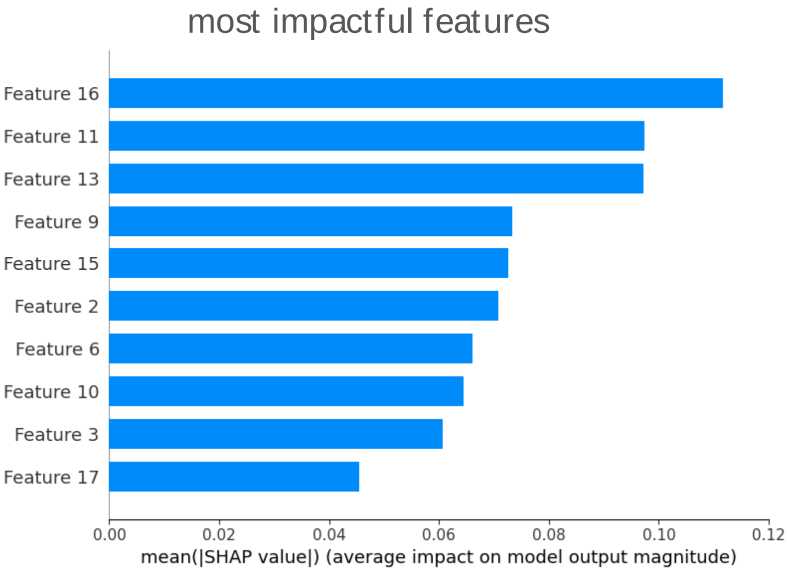


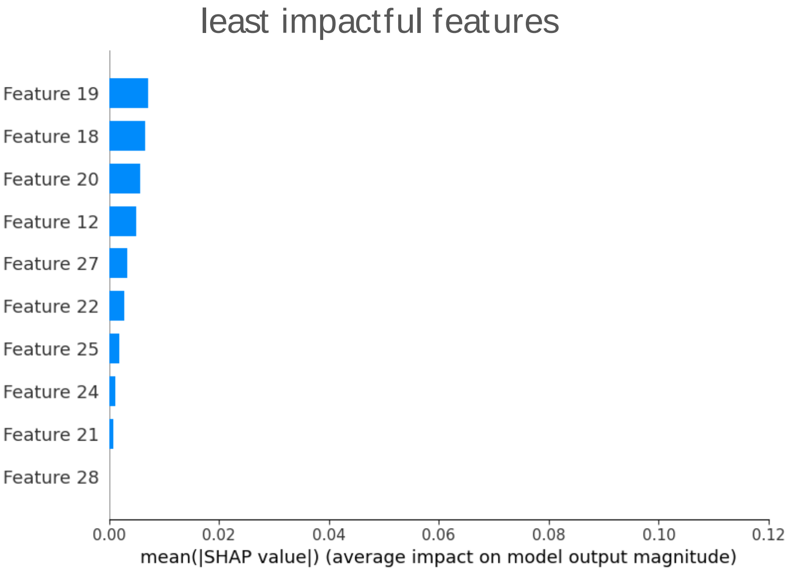
OC-NN: AUROC and AUPRC





Explainability: Shapley values
OC-NN Model





Results: Hybrid approach

	Model	Precision	Recall	F1-Score	ROC-AUC	MCC
Baselines	Naive	0.5	1.0	0.66	0.5	0
	K-Means (k=2)	0.257	0.237	0.247	–	-0.451
	OC-SVM	0.579	**0.684**	**0.627**	0.617	0.19
	SGD OC-SVM	0.077	0.017	0.028	0.21	-0.299
	Isolation Forest	0.655	0.192	0.297	0.628	0.129
	kNN distance	**0.866**	0.369	0.518	**0.827**	**0.381**
	LOF	0.423	0.423	0.423	0.386	-0.155
	Mahalanobis Distance	0.604	0.077	0.136	0.601	0.054
	GMM	0.675	0.057	0.098	0.66	0.069
	KDE	0.662	0.25	0.363	0.635	0.157
DL Model	OC-NN	0.003	0	0	0.214	-0.08

Conclusion

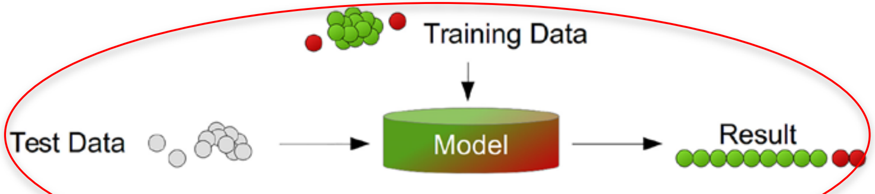
Conclusion

Takeaways

- OC-NN performs the best
- DL models are more robust for high-dimensional data
- Simpler baseline models still of importance

Future Work

- Examine hybrid approach further
- Different Training Scenario
    - Normal and fraud data during training



(a) Supervised anomaly detection

Thank you for your attention!

Credit Card Fraud Detection - Final presentation (01.02.2024)

Module: Deep Learning Lab
Supervisor: M.Sc. Rodrigo Lopez Portillo Alcocer
Students: Hasan Evci and Tareq Abu El Komboz

References

- C. C. Aggarwal, “An Introduction to Outlier Analysis.”, pp. 1–34. Cham: Springer International Publishing, 2017.
- L. Ruff, J. R. Kauffmann, R. A. Vandermeulen, G. Montavon, W. Samek, M. Kloft, T. G. Dietterich, and K. Müller, “A unifying review of deep and shallow anomaly detection.”, Proc. IEEE, vol. 109, no. 5, pp. 756–795, 2021.
- M. Goldstein and S. Uchida, “A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data.” ,PloS one, vol. 11, no. 4, 2016.
- Raghavendra Chalapathy, Aditya Krishna Menon, and Sanjay Chawla. Anomaly detection using one-class neural networks. CoRR, abs/1802.06360, 2018.
- V. N. Dornadula and S. Geetha, “Credit Card Fraud Detection using Machine Learning Algorithms.”, in Procedia Computer Science, vol. 165, 2019, pp. 631-641.
- C. V. Priscilla and D. P. Prabha, "Credit Card Fraud Detection: A Systematic Review.", in Intelligent Computing Paradigm and Cutting-edge Technologies: Proceedings of the First International Conference on Innovative Computing and Cutting-edge Technologies (ICICCT 2019), Istanbul, Turkey, October 30-31, 2019, vol. 1, pp. 290-303, Springer International Publishing, 2020.
- G. E. Hinton and R. R. Salakhutdinov, “Reducing the dimensionality of data with neural networks.”, Science, vol. 313, no. 5786, pp. 504-507, 2006.
- N. Report, “Global credit card fraud statistics.”, The Nilson Report, no. 1232, 2021. (*Accessed on 2023/11/26*).

Appendix

Autoencoder: Decision Boundary

Different strategies to set **decision boundary:**

- Just beyond overlap of normal and fraud data (**balanced precision and recall, high f1 score**)
- “At the end” of normal data distribution (**high precision**)
- “At the beginning” of fraud data distribution (**high recall**)
- Based on statistical measures (e.g. 95% percentile of normal data)