MSc Applied Artificial Intelligence 2021/22

Advanced Practice – Reflective Report

Analysing Feature Contributions of Model Outputs: Credit Card Default & Phishing Dataset

Submission Date: 17 December 2021
Word count: 2841 (excluding references)

Table of Contents

1. Introduction

This reflective report will demonstrate and detail my critical reflection of my experience through the internal project undertaken with the MSc Applied Artificial Intelligence Advanced Practice module. In addition to the analysis of my personal learning and development, I will demonstrate and structure my experience using the Gibbs Reflective Cycle (Fig.1).

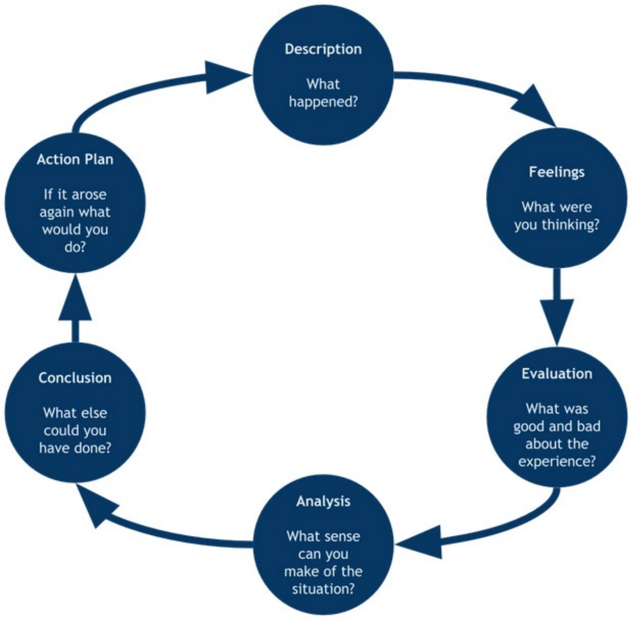


Fig 1. Gibbs Reflective Cycle (James Cook University, 2021)[1]

The project carried out was the exploration of statistical distribution of dataset features based on their feature importance and classification results. The objective was to explore the attributes with highest feature importance and determine whether its contribution is relevant to the model output.

The project had a timescale from October 1st to the reflective report submission date of December 17th 2021. Throughout the project, I found that the structured weekly deadlines (Appendix 1) guided my project completion well rather than feel overcome by pressure.

2. Activities Undertaken

2.1 Design

During the first supervisor meeting, we shared our coding and technical experiences as a class. The task for the first week was dedicated to identifying a research area of interest and to present our individual proposals during the second supervisor meeting.

I looked forward to drafting a proposal that would cater to a niche research area, Explainable Artificial Intelligence. My initial research was based on my project in the previous semester on Explainable Artificial Intelligence in Credit Risk Management. The aim was not only to incorporate short-term, but also longer-term research goals. During the second week, some of my proposed research topics included:

- Decision making algorithms in life insurance with interpretable Machine Learning.

- Artificial Intelligence in market movement prediction.

- Risk prediction of the mortality of patients with end stage kidney disease.

My ideas stemmed from a combination of surveys and research papers read in the previous semester and also included articles from the Alan Turing Institute website. The overall experience of producing a project idea was positive with the benefit of my experience in my previous module. I was aware of the current technologies in place and used that to identify a niche topic of research.

However, when I had presented my ideas to my supervisor, he had pointed out the advantages and disadvantages of each idea and made me think of how the results can be useful – or not in some cases. A challenge I struggled with was the practicality of my project objectives. Receiving the feedback from my supervisor was helpful in tapping into his technical knowledge and how beneficial certain projects may be. Discussing further, we had agreed that I would work individually on addressing biases across datasets within specific families of algorithms.

Further to this, I had planned to expand the project in the future by being able to map specific inputs that have a stronger influence between the input and output of a model to visualise and explain their relationships. If I were to encounter such a process in the future, I would not have spent too much time reading every research paper I come across. I doubted my skillset in the beginning of the process of identifying a research topic and did not branch out to more complex methods of achieving my objective.

2.2 Literature Review & Datasets

A range of literature was reviewed between Weeks 2 and 5, initially to identify a research topic for the internal project then to explore existing literature based on the selected research topic. I approached this task with my experience from the Research Methods module from the previous semester and utilised several libraries including IEEE Xplore Database, Science Direct and Google Scholar. A list of the included and excluded literature can be found in Appendix 2.

It was beneficial to have had experience from my previous semester to guide the exploration of literature because I found the relevant information with ease. I struggled again with practicality because the majority of my idea exploration was far too complex for the scope and timeframe of the project. I also struggled with time management in this period because in fear of missing out any key information, I had spent too much time reviewing every piece of literature I came across.

However, I was pleased with the outcome of the literature review because I had come to the decision of the algorithms I would utilise and then searched Kaggle and the UCI Repository for datasets relevant to key areas of research in Artificial Intelligence. Moving forward, I understand now to invest my time in the relevant techniques relating to the research topic through the direct implementation within the project.

2.3 Implementation & Results

At the end of October, I summarised and presented my selected methodologies and datasets to my supervisor. I chose to implement Decision Trees, Random Forests, Gradient Boosting and XGBoost algorithms on the Credit Card Default[2] and Phishing Detection[3] datasets. The datasets were cleaned and feature engineering was performed on the Credit Card Default dataset. For each dataset, I created three other adjusted datasets. One had Principal Component Analysis applied and the other two had 50% and 20% of the original training set that were positive.

The Principal Component Analysis dataset was created after my supervisor had reviewed my initial results and suggested that it would be useful to analyse the feature importance.

I first carried out a blind run on each dataset to determine baseline results. The next step was using GridSearchCV first with 'Accuracy' as the scoring to tune the hyperparameters. I then repeated the previous step using 'F1 Score' as the scoring. I chose to run GridSearchCV on both scoring methods because the Phishing Detection dataset was well-balanced and suited 'Accuracy' better but the 'F1 Score' was better-suited for the Credit Card Default data. The results are listed in Appendix 3.

The project objective as an entirety was to analyse feature contributions towards the resulting classification in a dataset. Combining this semester-long research project with an extension through to my dissertation to conduct further analysis will produce findings I believe will make a significant impact in the way we look at the relevance of certain features in a dataset.

The next step in my model was to visualise the feature importance across each of the dataset (and their adjusted versions). I used both the classifier's feature importance attribute and the Mutual Information scoring to check for any consistencies. My final step was determining the statistical distribution of the attribute with the highest feature importance and check its consistencies across the amended datasets.

I noted that despite high accuracy and F1 scores on the original datasets, more often than not, the adjusted datasets with either 50% or 20% of the original positive input entries had almost as high scoring as the original dataset. The feature with the highest impact was also consistent across the different adjusted datasets.

The overall experience of the implementation was positive because I was able to structure my code in line with the project's objectives, from cleaning and feature engineering to the tuning of algorithm hyperparameters. I paid attention to feedback from my supervisor at each meeting to drive improvements to the implementation of my methodology. Some of the feedback include creating a bias in the dataset by altering the input examples to examine changes in feature importance.

There was no doubt though that I experienced "imposter syndrome" and doubted my work at various points of the project as my background in coding and Artificial Intelligence is limited to what I have had learned on the course. I felt that I was not performing to the required standards. If I were to do this again, I would attempt to explore a more thorough explorations of algorithms to identify a better explanation of attributes influencing model outputs.

3. Conclusion

3.1 Reflection

Conducting the reflective report and summarising my results have increased my professional level of competency in many ways. The first is that I have developed a critical mindset towards approaching machine learning problems as a result of conducting my project.

My mentor made it clear that it was important that my project's objective would contribute to the future development in the field of Explainable AI. For this reason, I revised my project goal several times until my supervisor was satisfied with the level of complexity and contributing power.

I tried to meet the recommended 20 hours a week of work through both time utilising the library facilities and my own computer off-campus. As a result, I feel a sense of ownership over the work that I have produced, this reflective report included. This ownership has encouraged a great deal of enthusiasm for my future career in the industry. My attendance to the majority of my classes increased my optimism for my progress in coding and understanding of Machine Learning as I communicated with both my supervisor and other students on the course.

I found that at the end of the project, my understanding of algorithms have improved alongside my Python coding skills. I am able to work with dataframes with little reference to documentation and learned to successfully implement to several new concepts like Mutual Information of target variables and Principal Component Analysis.

However, a recurring challenge I faced throughout the project was the practicality of my objectives and meeting timeframes. To improve my time management, I have tried to respect the importance of scheduling tasks, minimising any distractions and avoiding tasks that are not a priority. And to address the practicality of my objectives, I paid attention to the feedback from my supervisor and picked up his advice like generalising my objectives for the benefit of future research rather than having it be business are specific.

In the future, should I encounter a similar project I will do the following things differently. The first would be limiting the scope of my research proposal but paying attention on key features of the research's contribution. An example would be developing a meta-algorithm that explains specific input contributions across any form of dataset and algorithm. I would also learn on-the-go by implementing existing methodologies alongside my project rather than invest too much time in learning the concepts independent of my work. I also aim to be more confident in my capabilities and to become more vocal of my opinions and questions.

3.2 Personal Benefits

The research experience was a useful extension from experience with the Research Methods module and combining with the technical skills I have developed throughout the project has increased my motivations for the course, making more extensive plans for upcoming job opportunities and carrying out the necessary measures and initiatives to ensure I accomplish these plans. I am grateful for the guidance of my mentor throughout this project for all the practical tips that can be applied in a workplace setting.

My level of self-confidence has also increased because I was able to meet the project objectives and identify my strengths and weaknesses through the completion of this reflective report. I find that self-confidence is key in being accomplished in any aspect of life because it allows us to set ambitious plans and utilise all the available resources efficiently in order to achieve these plans.

My time-management skills have also greatly improved since the start of the project as I had to adopt to meeting the set weekly or bi-weekly goals to stay on track with the project. As a related positive note, I was able to spend more of my time exploring career-relevant reports that I have used to broaden my knowledge of new implementations in the industry.

Moving forward, I now have the opportunity to expand on this internal project by further exploring how I can improve and integrate this method of feature analysis into the field of Explainable Artificial Intelligence. I would like to create the opportunity to develop a meta-algorithm that can be applied across all datasets and algorithms. This may prove to be key in securing my standing within the industry in the future.

3.3 Professional Values & Behaviour

Although I found the project to be challenging and confusing initially, I was able to appreciate the value of the critical analysis during the drafting of my reflective report. The skills of critical analysis developed and applied throughout this project can be easily applied when real business issues would need to be resolved in my future career capacity.

Adaptability, communication and critical thinking are two key skills areas that I have honed over the period of the project. There were a number of times that I had to modify my model's objectives to meet the scope of the project. These changes were made in response to feedback from my supervisor. I now better understand the importance of producing alternative solutions when the preferred method is rejected. Moving forward, I will make a conscious effort to accept that the objective of a project can be dynamic as I move through its stages.

As I progressed through the implementation of my models, I was able to critically understand my supervisor's critiques of my chosen methodologies and as mentioned, adapt my model suitably. However, I will have to improve my capability to scrutinise the information I take in and judgements made from it. I will also have to become less self-conscious in asking more questions because in the long run, this will benefit my personal and professional development in the workplace by distinguishing new perspectives on matters.

My communication skills have also benefitted from this project as the weekly meetings with my supervisor gave me the opportunity to share my findings, ideas and proposed approaches. One downfall to be improved is my failure to clearly articulate my ideas relating to the project. I relate this partly to my limited technical knowledge and not being able to fully comprehend my supervisor's critique of my work. Moving forward, I will have to pursue more projects to increase both my technical knowledge and self-confidence in communicating with peers in the industry.

This Advanced Practice module as a whole has bettered myself as both an individual and professional, having been able to complete the semester-long project and met the objectives I have set with my supervisor. I am grateful to have had the opportunity to critically evaluate my progress to grow into my future career.

4. Appendices

Appendix 1. Weekly Project Schedule with Tasks Agreed with Supervisor

TASK	September 27th	October 4th	October 11th	October 18th	October 25th	November 1st	November 8th	November 15th	November 22nd	November 29th	December 6th	December 13th
Phase 1 - Topic Identification & Research												
Week 1-2: Identifying 3 key topics of interest **Outcome:** Think of 10 potential topics, pick 3 (identify challenges, data, etc) and justify choice												
Week 3-4: Research on allocated topic (XAI) **Outcome:** 7 papers ranging from surveys to domain specific researches were identified. A summary document produced covering the definition of XAI, how and why it is used and what future work is. This will be presented in Week 3. 1 survey and 2 key research papers were identified detailing useful XAI methodology and evaluation with respect to reducing discrimination.												
Phase 2 - Methodology & Implementation												
Week 5-6: Methodology Review **Outcome:** The 2 surveys in Phase 1 were used as a point of reference of existing algorithms to be studied to design an XAI algorithm that identifies key input enteries that influences model outcomes. An evaluation measure is yet to be determined.												
Week 7: Methodology Development & Implementation **Outcome:** More relevant literature was identified inlcuding InterpretML by Microsoft that implements a glassbox model (suitable for addressing most algorithms' transparency). However, this analysis was too deep for me to personally complete. A discussion with my supervisor helped me decide to use statistical analysis between features of each classified input as a basis of identifying algorithm bias.												
Week 8-9: Model Implementation & Checking **Outcome:** 3-4 algorithms were selected to be run on three different datasets surrounding different topics (Safety, Finance, Security). These models have been run and model results now have to be interpreted with respect to the presence of algorithmic bias. There was a pause in work in Week 8 because of personal circumstances.												
Phase 3 - Results, Discussion & Conclusion												
Week 10: Result Interpretation & Discussion **Outcome:** The reflective report has been updated. Further exploration into PCA and Mutual Info Classifier was implemented onto the dataset to identify importance of features on the classification.												
Phase 4 - Documentation & Further Work												
Week 11: The reflective report was further updated. The dataset was adjusted to cheeck how it affects feature importance on the classification outcome.												
Week 12: Complete and submit reflective report.												

Appendix 2. Literature Inclusion/Exclusion Evaluation

	Research Paper Title	Include / Exclude	Reasoning
1	XAI Evaluation: Evaluating Black-Box Model Explanations for Prediction	Include	Relevant to the understanding black-box models
2	The effects of explainability and causability on perception, trust and acceptance: Implications for explainable AI	Include	Relevant to the explainable aspect of decision-making models
3	Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)	Include	Relevant to understanding existing algorithms
4	Neuroscope: An Explainable AI Toolbox for Semantic Segmentation and Image Classification of Convolutional Neural Nets	Exclude	Too complex for the project scope
5	From Machine Learning to Explainable AI	Include	Relevant to understanding existing algorithms
6	Explainable AI in Fintech Risk Management	Include	Relevant to the use of XGB models
7	A Multidisciplinary Survey and Framework for Design and Evaluation of Explainable AI Systems	Exclude	Not relevant to the project topic
8	Development and application of consumer credit scoring models using profit-based classification measures	Exclude	Not relevant to the project topic
9	Explainable AI: A Review of Machine Learning Interpretability Methods	Include	Relevant to the understanding of existing models
10	Fairness in credit scoring: Assessment, implementation and profit implications	Exclude	Not relevant to the project topic
11	Measuring discrimination in algorithmic decision making	Include	Appropriate to understanding the impact of sensitive attributes
12	InterpretML: A Unified Framework for Machine Learning Interpretability	Exclude	Too complex for the project scope

Appendix 3. Results

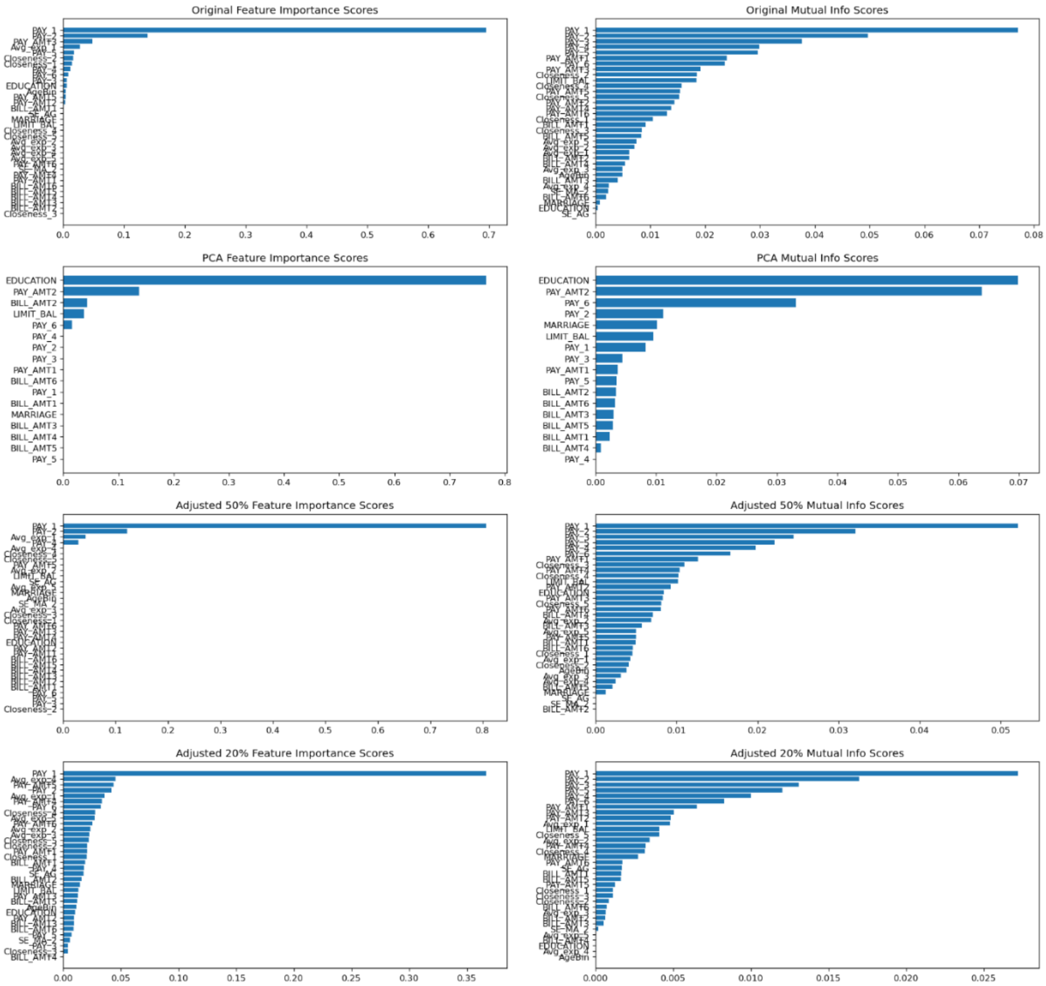
Dataset 1: Credit Card Default

A. Algorithm: Decision Tree

1. GridSearchCV (F1 Score)

	Best Estimator	Accuracy	F1 Score	TP	TN	FP	FN
Original Dataset	DecisionTreeClassifier(max_depth=5, max_leaf_n...	0.817833	0.471725	488	4419	268	825
PCA (Original Dataset)	DecisionTreeClassifier(max_depth=4, max_leaf_n...	0.802167	0.471269	529	4284	403	784
Adjusted Dataset(50%)	DecisionTreeClassifier(criterion='entropy', ma...	0.882706	0.417132	224	4487	201	425
Adjusted Dataset(20%)	DecisionTreeClassifier(max_depth=9, max_leaf_n...	0.938247	0.15978	29	4605	70	235

2. Comparison of Feature Importance



3. Comparison of Statistical Distribution (Adjusted Dataset 50%)

Class == 1

	PAY_1
count	81.000000
mean	0.283951
std	0.617142
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	2.000000

Class == 0

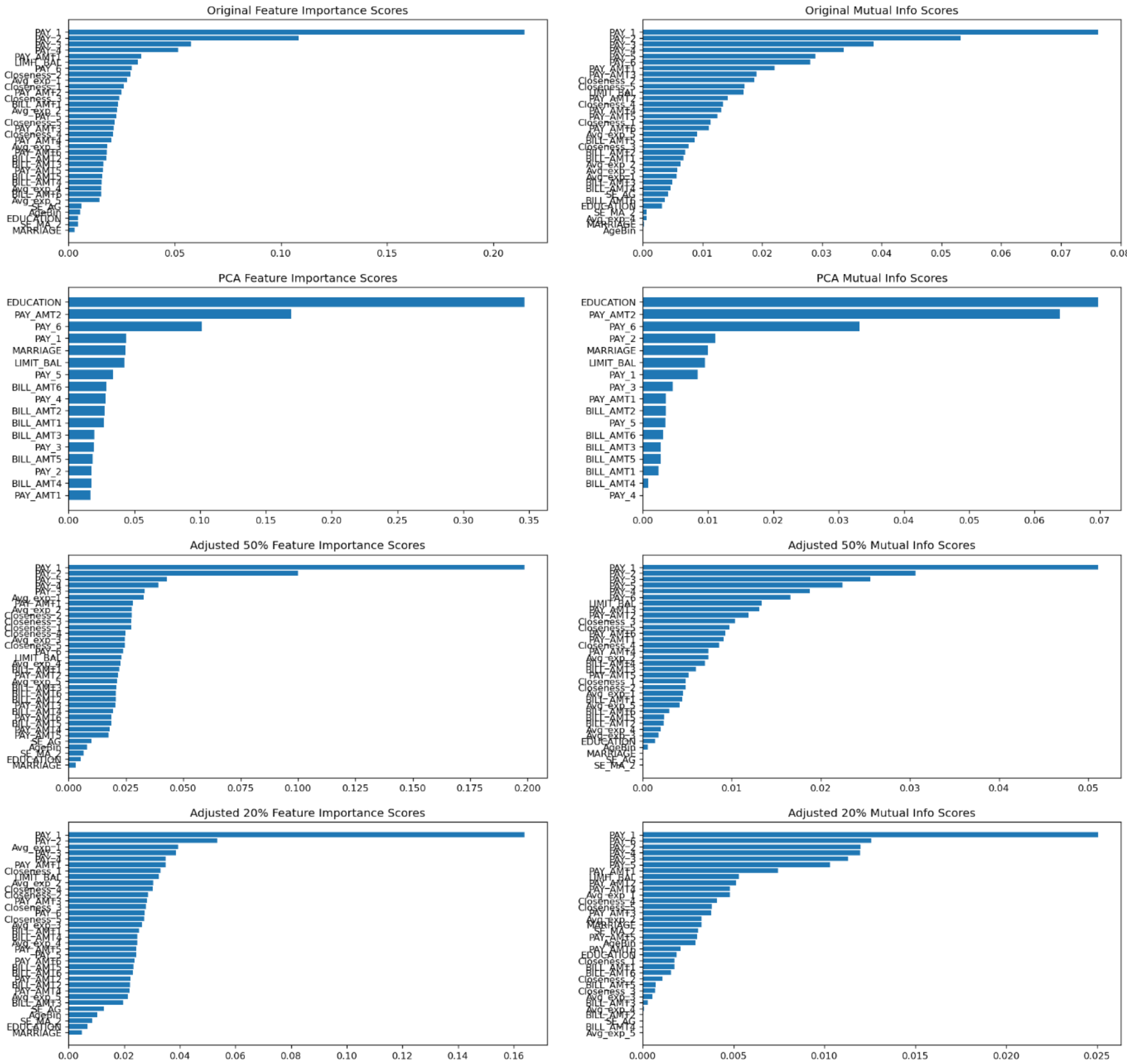
	PAY_1
count	5256.000000
mean	0.285008
std	0.662874
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	8.000000

B. Algorithm: Random Forest

1. GridSearchCV (F1 Score)

	Best Estimator	Accuracy	F1 Score	TP	TN	FP	FN
Original Dataset	(DecisionTreeClassifier(max_depth=9, max_featu...	0.821667	0.458502	453	4477	210	860
PCA (Original Dataset)	(DecisionTreeClassifier(max_depth=9, max_featu...	0.815333	0.449304	452	4440	247	861
Adjusted Dataset(50%)	(DecisionTreeClassifier(max_depth=9, max_featu...	0.883455	0.237745	97	4618	70	552
Adjusted Dataset(20%)	(DecisionTreeClassifier(max_depth=9, max_featu...	0.946548	0.029412	4	4671	4	260

2. Comparison of Feature Importance



3. Comparison of Statistical Distribution (Adjusted Dataset 20%)

Class == 1

	PAY_1
count	571.0
mean	1.0
std	0.0
min	1.0
25%	1.0
50%	1.0
75%	1.0
max	1.0

Class == 0

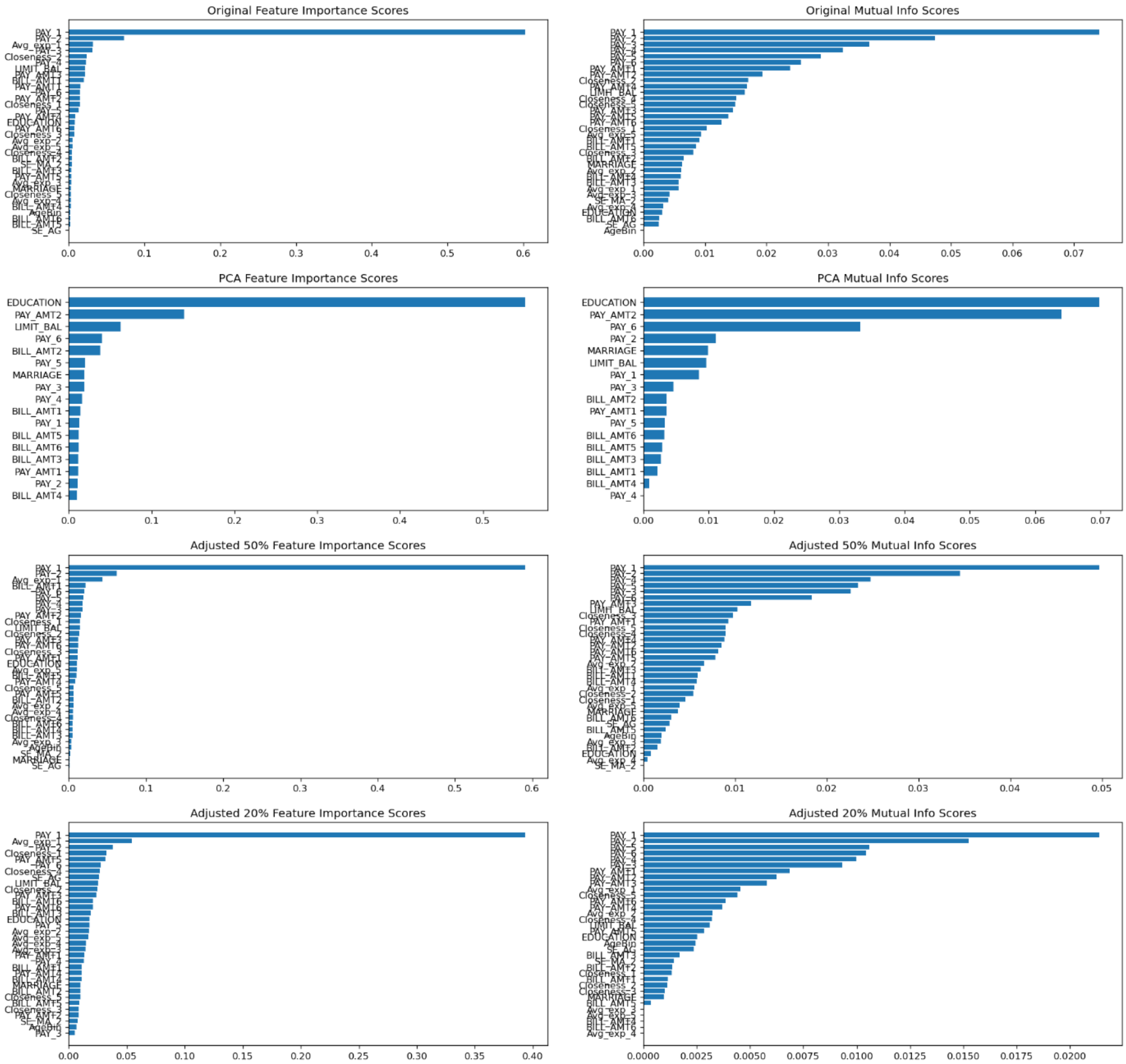
	PAY_1
count	4043.0
mean	0.0
std	0.0
min	0.0
25%	0.0
50%	0.0
75%	0.0
max	0.0

C. Algorithm: Gradient Boosting

1. GridSearchCV (F1 Score)

	Best Estimator	Accuracy	F1 Score	TP	TN	FP	FN
Original Dataset	([DecisionTreeRegressor(criterion='friedman_ms...	0.821	0.464606	466	4460	227	847
PCA (Original Dataset)	([DecisionTreeRegressor(criterion='friedman_ms...	0.813333	0.457364	472	4408	279	841
Adjusted Dataset(50%)	([DecisionTreeRegressor(criterion='friedman_ms...	0.889264	0.359697	166	4580	108	483
Adjusted Dataset(20%)	([DecisionTreeRegressor(criterion='friedman_ms...	0.945333	0.014599	2	4667	8	262

2. Comparison of Feature Importance



3. Comparison of Statistical Distribution (Adjusted Dataset 50%)

Class == 1

	PAY_1
count	631.0
mean	1.0
std	0.0
min	1.0
25%	1.0
50%	1.0
75%	1.0
max	1.0

Class == 0

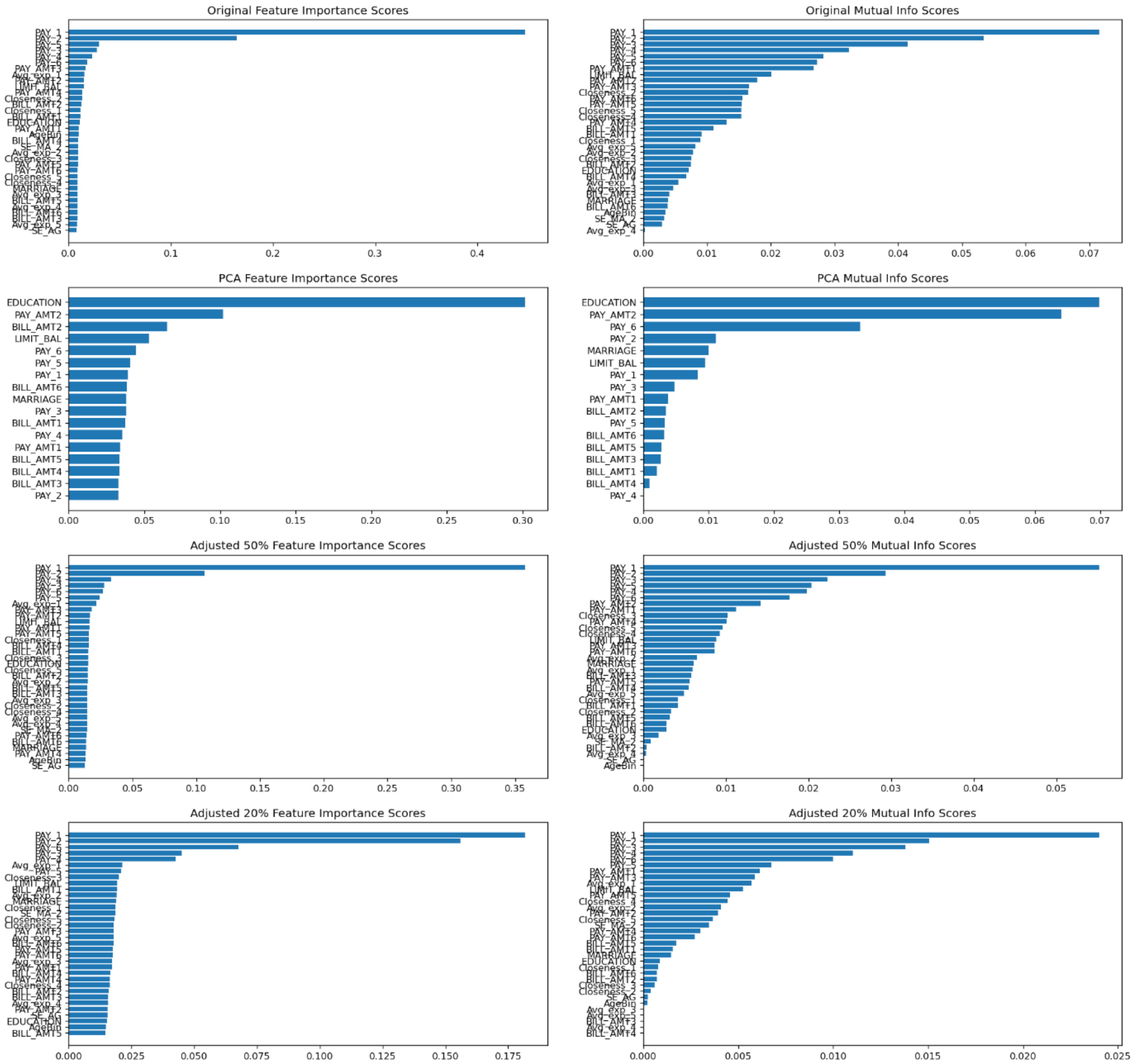
	PAY_1
count	4299.0
mean	0.0
std	0.0
min	0.0
25%	0.0
50%	0.0
75%	0.0
max	0.0

D. Algorithm: XGB

1. GridSearchCV (F1 Score)

	Best Estimator	Accuracy	F1 Score	TP	TN	FP	FN
Original Dataset	XGBClassifier(base_score=0.5, booster='gbtree'...	0.817667	0.457879	462	4444	243	851
PCA (Original Dataset)	XGBClassifier(base_score=0.5, booster='gbtree'...	0.812667	0.444115	449	4427	260	864
Adjusted Dataset(50%)	XGBClassifier(base_score=0.5, booster='gbtree'...	0.887015	0.349515	162	4572	116	487
Adjusted Dataset(20%)	XGBClassifier(base_score=0.5, booster='gbtree'...	0.945131	0.181269	30	4638	37	234

2. Comparison of Feature Importance



3. Comparison of Statistical Distribution (Adjusted Dataset 50%)

Class == 1

	PAY_1
count	631.0
mean	1.0
std	0.0
min	1.0
25%	1.0
50%	1.0
75%	1.0
max	1.0

Class == 0

	PAY_1
count	4299.0
mean	0.0
std	0.0
min	0.0
25%	0.0
50%	0.0
75%	0.0
max	0.0

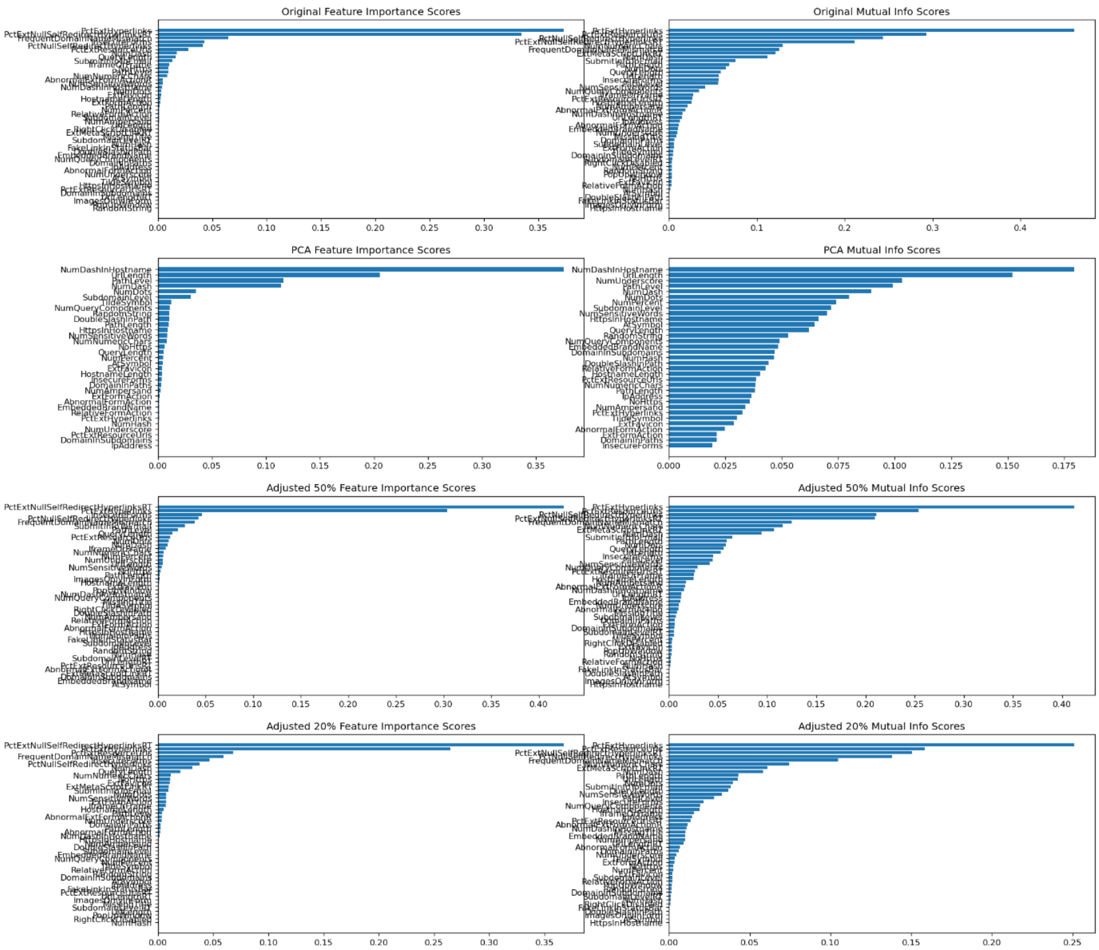
Dataset 2: Phishing Detection

A. Algorithm: Decision Tree

1. GridSearchCV with Accuracy

	Best Estimator	Accuracy	F1 Score	TP	TN	FP	FN
Original Dataset	DecisionTreeClassifier(criterion='entropy', ma...	0.9705	0.971007	988	953	35	24
PCA (Original Dataset)	DecisionTreeClassifier(max_depth=9, max_leaf_n...	0.9095	0.909905	914	905	83	98
Adjusted Dataset(50%)	DecisionTreeClassifier(max_depth=9, max_leaf_n...	0.96	0.939148	463	977	42	18
Adjusted Dataset(20%)	DecisionTreeClassifier(criterion='entropy', ma...	0.9775	0.930946	182	991	15	12

2. Comparison of Feature Importance



3.

3.Comparison of Statistical Distribution (Adjusted Dataset 20%)

Class == 1

	PctExtNullSelfRedirectHyperlinksRT	PctExtHyperlinks
count	4.00000	4.000000
mean	0.50000	0.211675
std	0.57735	0.247588
min	0.00000	0.000000
25%	0.00000	0.000000
50%	0.50000	0.187500
75%	1.00000	0.399175
max	1.00000	0.471698

Class == 0

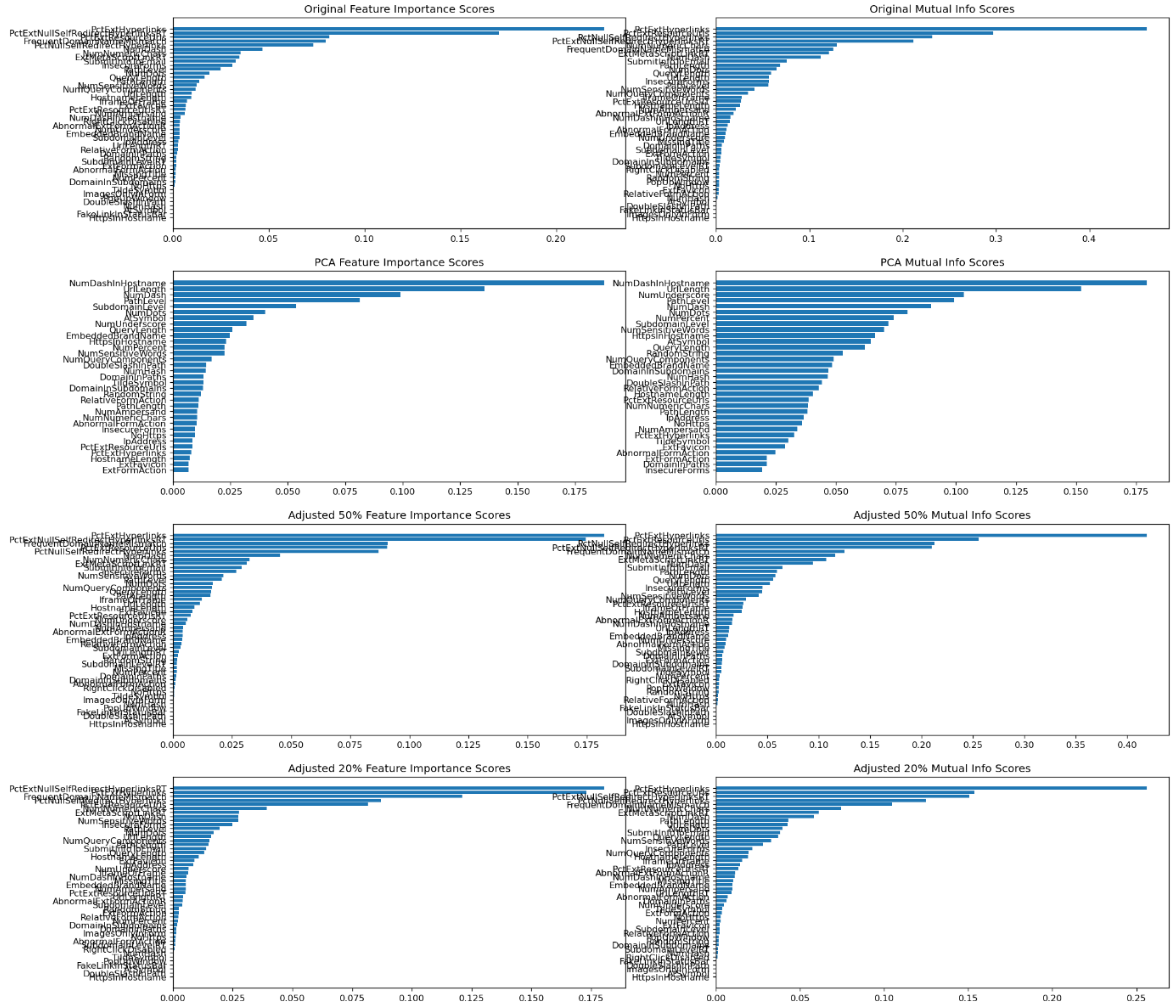
	PctExtNullSelfRedirectHyperlinksRT	PctExtHyperlinks
count	1196.000000	1196.000000
mean	0.652174	0.171457
std	0.666788	0.231925
min	-1.000000	0.000000
25%	1.000000	0.024845
50%	1.000000	0.090652
75%	1.000000	0.203914
max	1.000000	1.000000

B. Algorithm: Random Forest

1. GridSearchCV with Accuracy

	Best Estimator	Accuracy	F1 Score	TP	TN	FP	FN
Original Dataset	(DecisionTreeClassifier(max_depth=9, max_featu...	0.973	0.973267	983	963	25	29
PCA (Original Dataset)	(DecisionTreeClassifier(max_depth=9, max_featu...	0.9475	0.947683	951	944	44	61
Adjusted Dataset(50%)	(DecisionTreeClassifier(criterion='entropy', m...	0.97	0.95288	455	1000	19	26
Adjusted Dataset(20%)	(DecisionTreeClassifier(max_depth=9, max_featu...	0.984167	0.950131	181	1000	6	13

2. Comparison of Feature Importance



3. Comparison of Statistical Distribution (Adjusted Dataset 20%)

Class == 1

	PctExtNullSelfRedirectHyperlinksRT	PctExtHyperlinks
count	5.00000	5.000000
mean	0.20000	0.369340
std	0.83666	0.412633
min	-1.00000	0.000000
25%	0.00000	0.000000
50%	0.00000	0.375000
75%	1.00000	0.471698
max	1.00000	1.000000

Class == 0

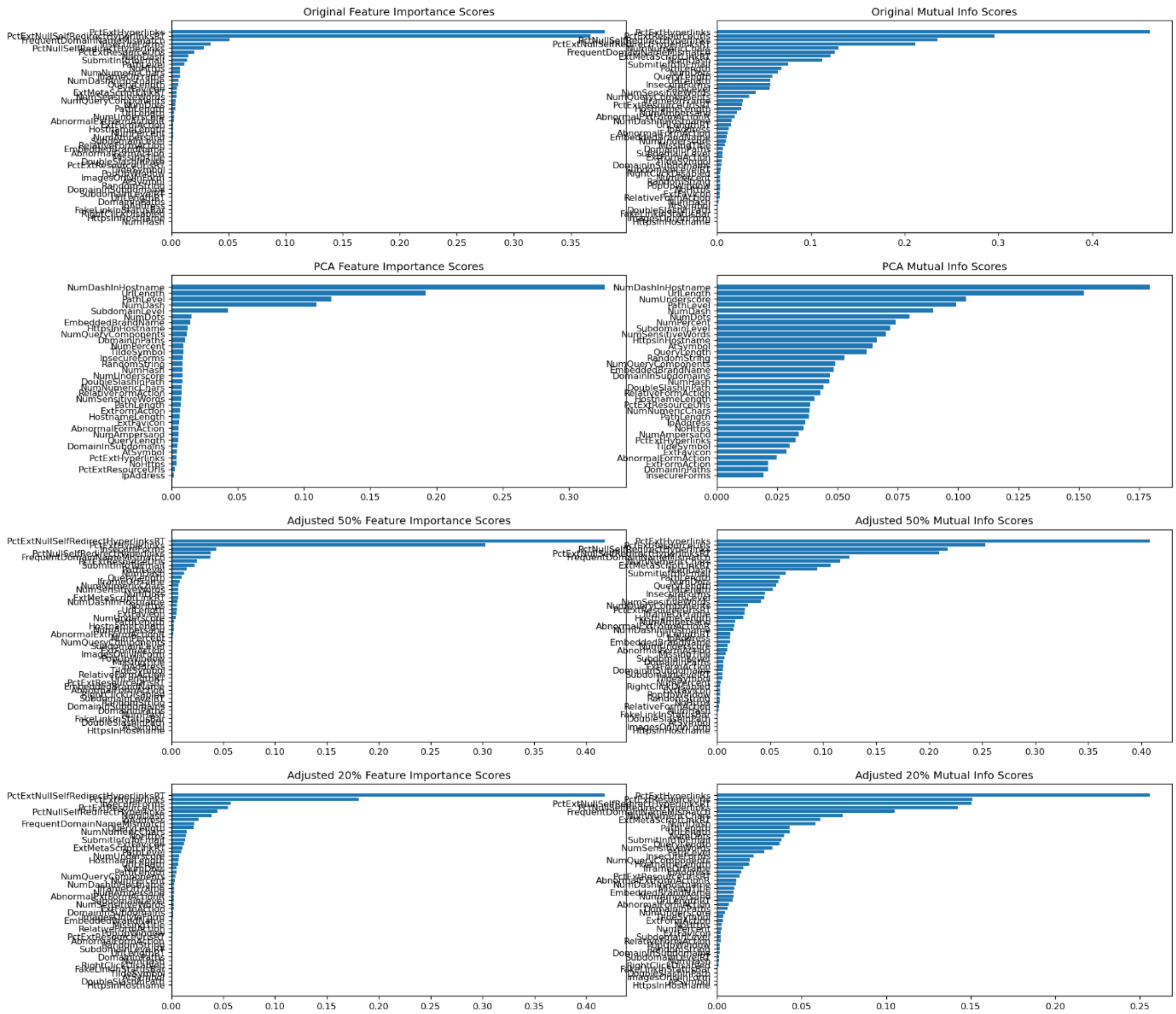
	PctExtNullSelfRedirectHyperlinksRT	PctExtHyperlinks
count	1195.000000	1195.000000
mean	0.653556	0.170764
std	0.665350	0.230779
min	-1.000000	0.000000
25%	1.000000	0.024814
50%	1.000000	0.090395
75%	1.000000	0.202813
max	1.000000	1.000000

C. *Algorithm: Gradient Boosting*

1. GridSearchCV with Accuracy

	Best Estimator	Accuracy	F1 Score	TP	TN	FP	FN
Original Dataset	([DecisionTreeRegressor(criterion='friedman_ms...	0.9865	0.986693	1001	972	16	11
PCA (Original Dataset)	([DecisionTreeRegressor(criterion='friedman_ms...	0.955	0.955357	963	947	41	49
Adjusted Dataset(50%)	([DecisionTreeRegressor(criterion='friedman_ms...	0.98	0.96888	467	1003	16	14
Adjusted Dataset(20%)	([DecisionTreeRegressor(criterion='friedman_ms...	0.985833	0.956743	188	995	11	6

2. Comparison of Feature Importance



3. Comparison of Statistical Distribution (Adjusted Dataset 50%)

Class == 1

	PctExtNullSelfRedirectHyperlinksRT	PctExtHyperlinks
count	43.000000	43.000000
mean	0.093023	0.336910
std	0.971350	0.432041
min	-1.000000	0.000000
25%	-1.000000	0.000000
50%	1.000000	0.000000
75%	1.000000	0.854167
max	1.000000	1.000000

Class == 0

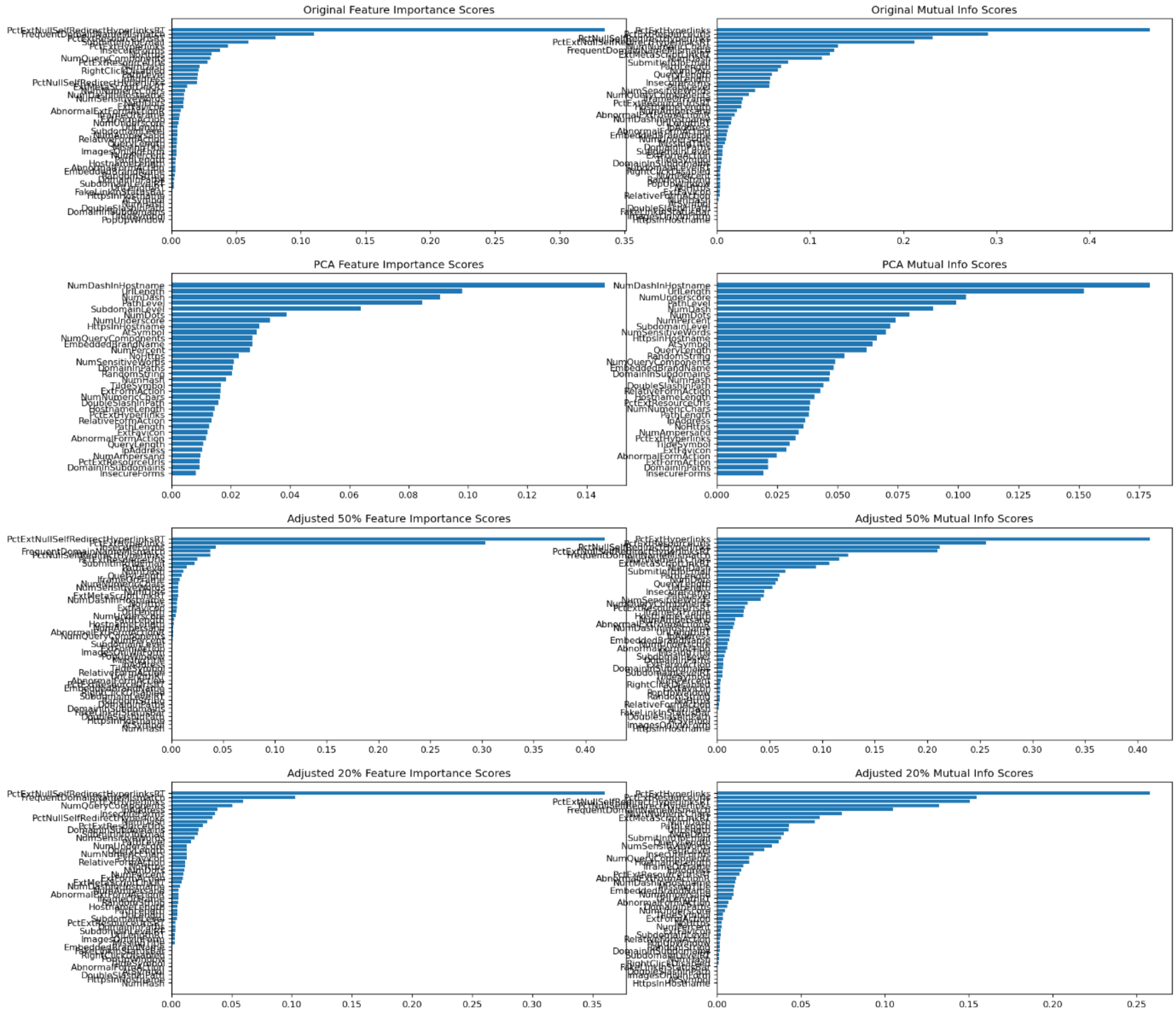
	PctExtNullSelfRedirectHyperlinksRT	PctExtHyperlinks
count	1457.000000	1457.000000
mean	0.523679	0.199622
std	0.785966	0.289593
min	-1.000000	0.000000
25%	0.000000	0.003226
50%	1.000000	0.083333
75%	1.000000	0.222222
max	1.000000	1.000000

D. Algorithm: XGB

1. GridSearchCV with Accuracy

	Best Estimator	Accuracy	F1 Score	TP	TN	FP	FN
Original Dataset	XGBClassifier(base_score=0.5, booster='gbtree'...	0.988	0.988189	1004	972	16	8
PCA (Original Dataset)	XGBClassifier(base_score=0.5, booster='gbtree'...	0.966	0.966337	976	956	32	36
Adjusted Dataset(50%)	([DecisionTreeRegressor(criterion='friedman_ms...	0.98	0.969008	469	1001	18	12
Adjusted Dataset(20%)	XGBClassifier(base_score=0.5, booster='gbtree'...	0.990833	0.971722	189	1000	6	5

2. Comparison of Feature Importance



3. Comparison of Statistical Distribution (Adjusted Dataset 20%)

Class == 1

	PctExtNullSelfRedirectHyperlinksRT	PctExtHyperlinks
count	6.000000	6.000000
mean	0.000000	0.461629
std	0.894427	0.432802
min	-1.000000	0.000000
25%	-0.750000	0.093750
50%	0.000000	0.423349
75%	0.750000	0.810232
max	1.000000	1.000000

Class == 0

	PctExtNullSelfRedirectHyperlinksRT	PctExtHyperlinks
count	1194.000000	1194.000000
mean	0.654941	0.170134
std	0.663903	0.229846
min	-1.000000	0.000000
25%	1.000000	0.024783
50%	1.000000	0.090359
75%	1.000000	0.201778
max	1.000000	1.000000

5. References

[1] Arts Edge Guide: Some Models of Reflection (James Cook University Australia, 2021). Accessed: 5 December 2021 [Online]
Available: https://libguides.jcu.edu.au/ba3000/reflection/models-of-reflection

[2] Credit Card Default Dataset (UCI Repository)
Accessed: 27 October 2021 [Online]
Available: https://archive.ics.uci.edu/ml/datasets/default+of+credit+card+clients

[3] Phishing Dataset for ML (Kaggle)
Accessed: 27 October 2021 [Online]
Available: https://www.kaggle.com/shashwatwork/phishing-dataset-for-machine-learning/

[4] Shin, D. (2021) The effects of explainability and causability on perception, trust and acceptance: Implications for explainable AI, International Journal of Human-Computer Studies

[5] Zhang, Y., Xu, F.,Zou, J., Petrosian, O. (2021) XAI Evaluation: Evaluating Black-Box Model Explanations for Prediction

[6] Adadi, A. and Berrada, M. (2018) Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI), IEEE Access, vol. 6, pp. 52138-52160

[7] Holzinger, A. (2018) From Machine Learning to Explainable AI, 2018 World Symposium on Digital Intelligence for Systems and Machines (DISA), pp. 55-66

[8] Bussmann, N., Paolo, G., Dimitri, M., Jochen, P. (2020) Explainable AI in FinTech Risk Management

[9] Lindardatos, P., Papastefanopulos, V., Kotsiantis, S. (2020) Explainable AI: A Review of Machine Learning Interpretability Methods

[10] Zliobaite, I. (2017) Measuring Discrimination in Algorithmic Decision Making