cisdi Journal

Computing, Information Systems & Development Informatics Journal
Vol. 10 No. 3, September, 2019 - www.cisdijournal.org

An Improved Hybrid System for The Prediction of Debit and Credit Card Fraud

Nwogu E. R. & Nwachukwu, E. O. PhD
Department of Computer Science
University of Port-Harcourt, Nigeria.

Ejiofor V. E. PhD
Department of Computer Science
Nnamdi Azikiwe University
Awka, Nigeria

Corresponding E-mail: nwogu.emeka@mouau.edu.ng

ABSTRACT

This work presents a design for the prediction of debit card fraud using a hybrid approach that involves generating synthetic data from a sample of real transaction data using the RtoSynR approach. The generated synthetic data is combined with the available sample of real data to form RtoSynR data and used to train a Genetic Algorithm optimized Random Forests model. The built model is then used to predict the genuineness of incoming transactions. The system used Object Oriented Analysis and Design and was designed with Python programming language while the synthetic data generation module was implemented by integrating R into a python designed interface. The various tests conducted with the model showed an improvement in the accuracy of the prediction over usual Random Forests model with real data bringing down the total number of misclassifications in the system from 2504 to 9.

Keyword: Hybrid, RtoSynR, GAORF, Genetic Algorithm, Synthetic, Debit card

CISDI JournalReferenceFormat
Nwogu E. R., Nwachukwu, E. O. & Ejiofor V. E (2019): An Improved Hybrid System for the Prediction of Debit and Credit Card Fraud. Computing,InformationSystems,DevelopmentInformatics&AlliedResearchJournal.Vol 10No 3, Pp 87-100. Availableonlineatwww.cisdijournal.org. DOI Affix - https://doi.org/10.22624/AIMS/CISDI/V10N3P8

1. INTRODUCTION

Online banking has continuously positioned itself as an unavoidable means of financial transaction in our technology oriented world. It involves conducting financial transactions over a secure website (Nwogu and Odoh, 2014). At the center of online banking is credit/debit card use, which has become a flexible service-on-the-go medium for financial transaction in our modern day society; resulting in increased credit/debit card fraud as the case may be. Despite the huge benefits of deploying and using online banking systems, the system has been plagued by enormous vulnerability which has been mainly linked directly to credit/debit card use. Richard & David (2002) in their worked posits that rate of credit card fraud has been on the increase. This view is also shared by David (2010) who posits that credit card fraud increased enormously between 2005 and 2007. Consequently, banks, merchants and card issuers are constantly looking for a better improved fraud prevention system for credit/debit card transactions. Due to a general belief that Information Technology systems can go a long way in checkmating and curtailing credit/debit card fraud, banks and merchants have invested lots of resources in fraud prevention research and systems.

Koosha and Osmar (2012) shows that there has been an increase in research on fraud detection in the financial market since 2008.Many credit/debit card fraud detection and prevention techniques have been discussed and implemented in many literature; some have also been deployed successfully in many systems. Although these systems have been successful at different times, yet some problems have remained a challenge in the development of such systems. This has been identified as the unavailability of real data for fraud prevention research. Other important problems that have been identified are high class imbalance, the availability of few transaction labels by fraud investigators, together with confidentiality issues with the few available data. These problems as identified in Sonepat and Sonepat (2011), Haibo and Edwardo (2009), Andrea et al. (2014 and Zhang etal. (2015) have unarguably militated research on the development of fraud detection systems. In the absence of credit or debit card transaction data, the research community have sought a way to implement a synthetic data driven research on fraud detection systems as described in Nwogu et al. (2019), Emilie and Erland (2003), Neha et al. (2016) and Lopez-Rojas and Axelsson (2012). These authors have supported strongly the use of synthetic data in credit card fraud detection and prevention research claiming there is no significant difference between the two types of data.

Similarly, machine learning approaches to fraud detection and prevention have been adjudged to be the most reliable approach as they have the potentials of also bringing down the cost of fraud detection and prevention systems; and as posited by Jon and Sriganesh (2008), the cost of transaction screening should not be higher than or close to the possible loss as a result of the fraud. Emphasis has now shifted to machine learning approaches to credit card fraud detection. It is imperative to use expert rules and statistical based models (basically Machine Learning) to process credit and debit card transactions and run a check to distinguish the genuine transactions from the potentially fraudulent transactions. With the use of Machine Learning (ML) techniques which is more advanced than predictive models, we can discover fraudulent patterns and predict transactions that are most likely to be fraudulent efficiently (Christopher, 2006). This technique basically infers a prediction model on the basis of a set of examples. The model in most cases is a parametric function, and allows predicting the likelihood of fraud in a transaction given a set of features describing the transaction.

2. RELATED LITERATURE

According to Raghavendra and Lokesh (2011), Credit card fraud can be defined as “Unauthorized account activity by a person for which the account was not intended. Potamitis (2013) citing Richard and David (2002) and Business Wire (2011) claimed that the total credit card fraud in the United Kingdom in the year 2000 stood at £286 million, while the United States had a total loss of $3.56 billion in the year 2009, a sharp increase of 10.2% from the previous year’s figure. Adrian (2015) reports that economic crime and fraud remains an intractable problem for global companies. To mitigate the huge losses due to fraudulent activities in credit and debit card transaction systems, a couple of approaches have been used, especially using machine learning approaches.

The work by Sahin and Duman (2011) implemented a system that used support vector machine to improve the classification accuracy of a decision tree model. This system monitored each account, and used a suspicion score generated by the system to flag each transaction as either genuine or fraudulent transaction. RamaKalyani and UmaDevi (2012) implemented a system that used genetic algorithm to reduce the number of false alerts, ultimately improving the classification accuracy of the system. Vijayshree etal. (2016) implemented a hybrid system that integrates support vector machine with decision tree and used for monitoring user accounts and classifying credit card transactions. Ishu etal. (2016) proposed a system that used genetic algorithm to determine whether an ongoing transaction request is fraudulent or genuine.

Their system included a data warehouse that contains the customer data which is taken through a rules engine to extract some important derived parameters for classifying incoming transactions. Similarly, Bharathidason & Venkataeswaran (2014) proposed Enhanced Random Forest (ERF) algorithm that used *K* in-of-bag data subsets to build *K* uncorrelated trees. They posit that this system can improve the accuracy of Random Forests algorithms in a card fraud detection system.

They used the Area-under-curve metric to select the best trees, and subsequently, the correlations between these trees are computed. From these *K* trees, *Q* most unrelated trees are then selected to form a Forest. Their result showed better accuracy as against normal Random Forests. Oberoi (2017) designed a Genetic Algorithm based system where the Genetic Algorithm is used to make decision about the network topology, number of hidden layers, and number of nodes that can be used in designing a neural network based fraud detection model.

3. METHODOLOGY

Our methodology involves a hybrid approach that combines RtoSynR model described in Nwogu et al. (2019) and Genetic Algorithm Optimized Random Forests described in Nwogu et al. (2019). A small sample of transaction data provided by an anonymous bank is used to simulate more transaction data using multivariate data generation approach and bootstrap approach which involves the mathematical modeling and extraction of parameters from the available real dataset. This approach is called Real-to-Synthetic-Real model; which is a four-step recursive model as shown in figure 1 that produces RtoSynR dataset as its output. The RtoSynR dataset is used to train a Random Forests model optimized with Genetic Algorithm. The essence of the optimization is to produce more accurate trees and forest for the prediction of incoming transactions. The optimization eliminates error due to poor convergence and greedy search in decision trees and Random Forests in general. figure 2 shows the macro-model of the RtoSynR architecture, while figure 3 shows the expanded architectural overview of the proposed Fraud Prevention system.

3.1 Modeling of the system

To achieve the proposed system, Object Oriented Analysis and Design paradigm was used. The unified modeling language (UML) was used to analyze the interactions between the various objects of the system. The UML tools used include the use case and sequence diagrams shown in figures 4 and 5. While the use case diagram shows us the different activities that go on in the new system, the sequence diagram shows us the flow of message between the objects and activities in the new system.

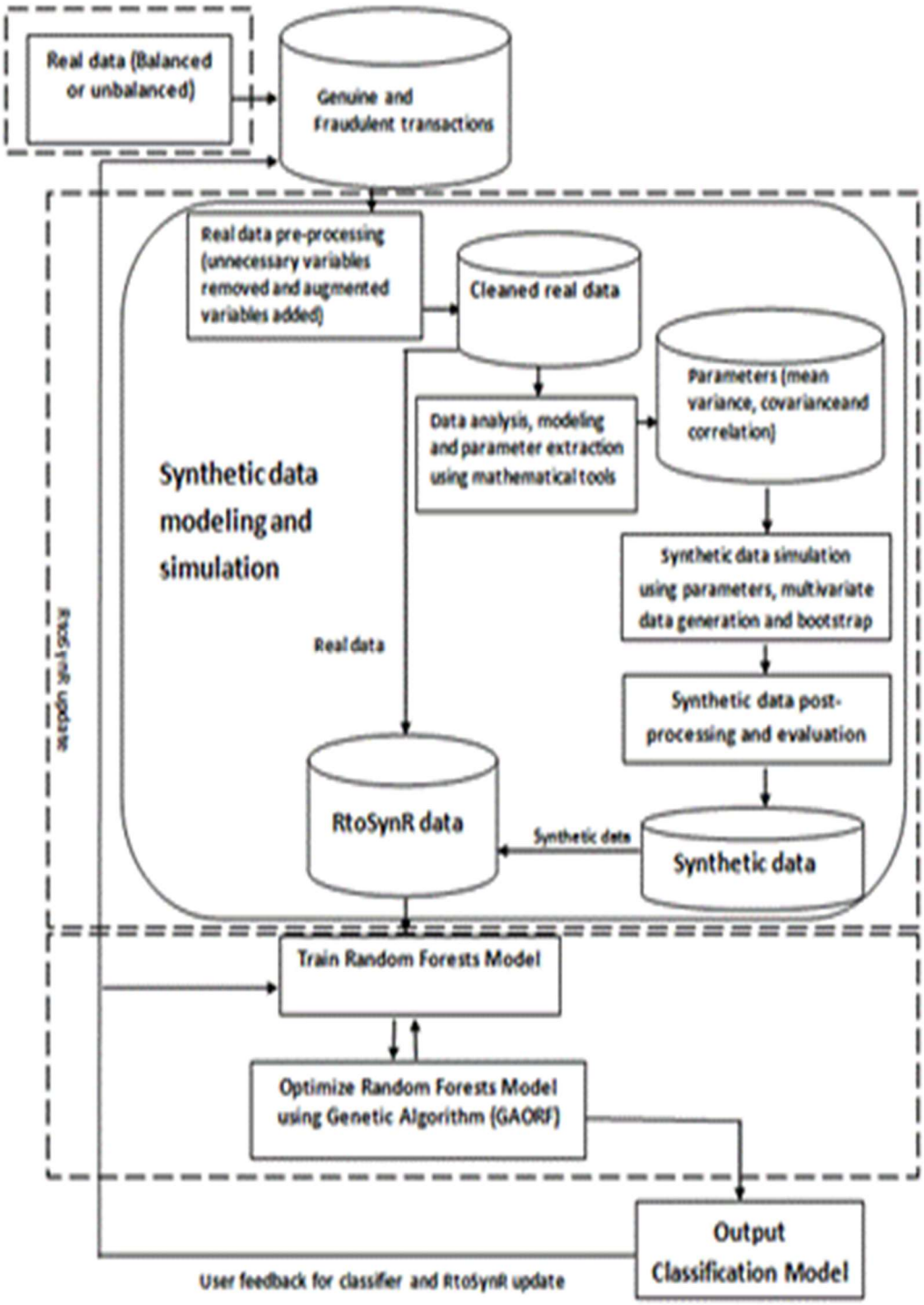


Figure 1: The general architecture of the proposed system

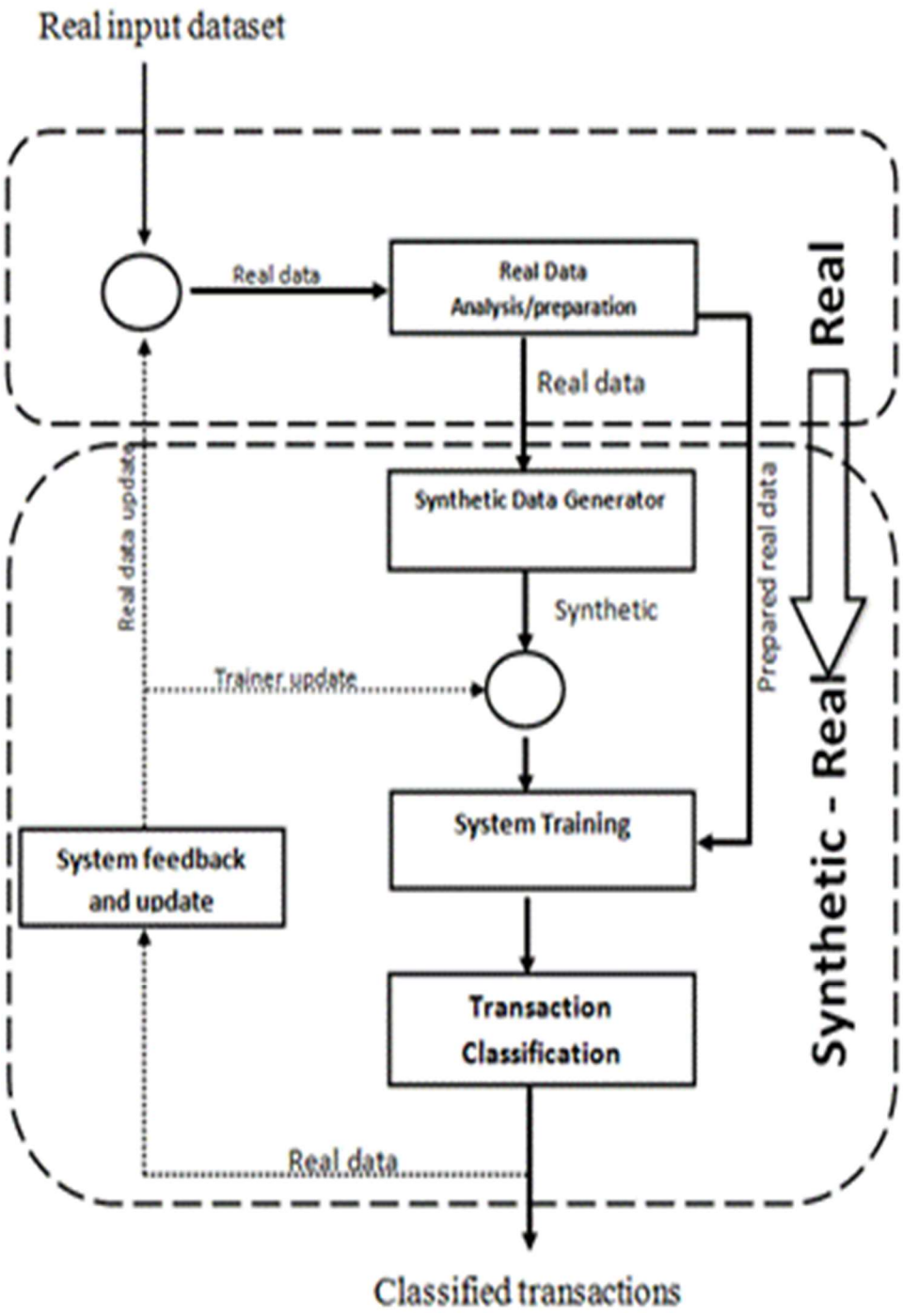


Figure 2: The macro-model of the RtoSynR architecture (source: Nwogu et al, 2019)

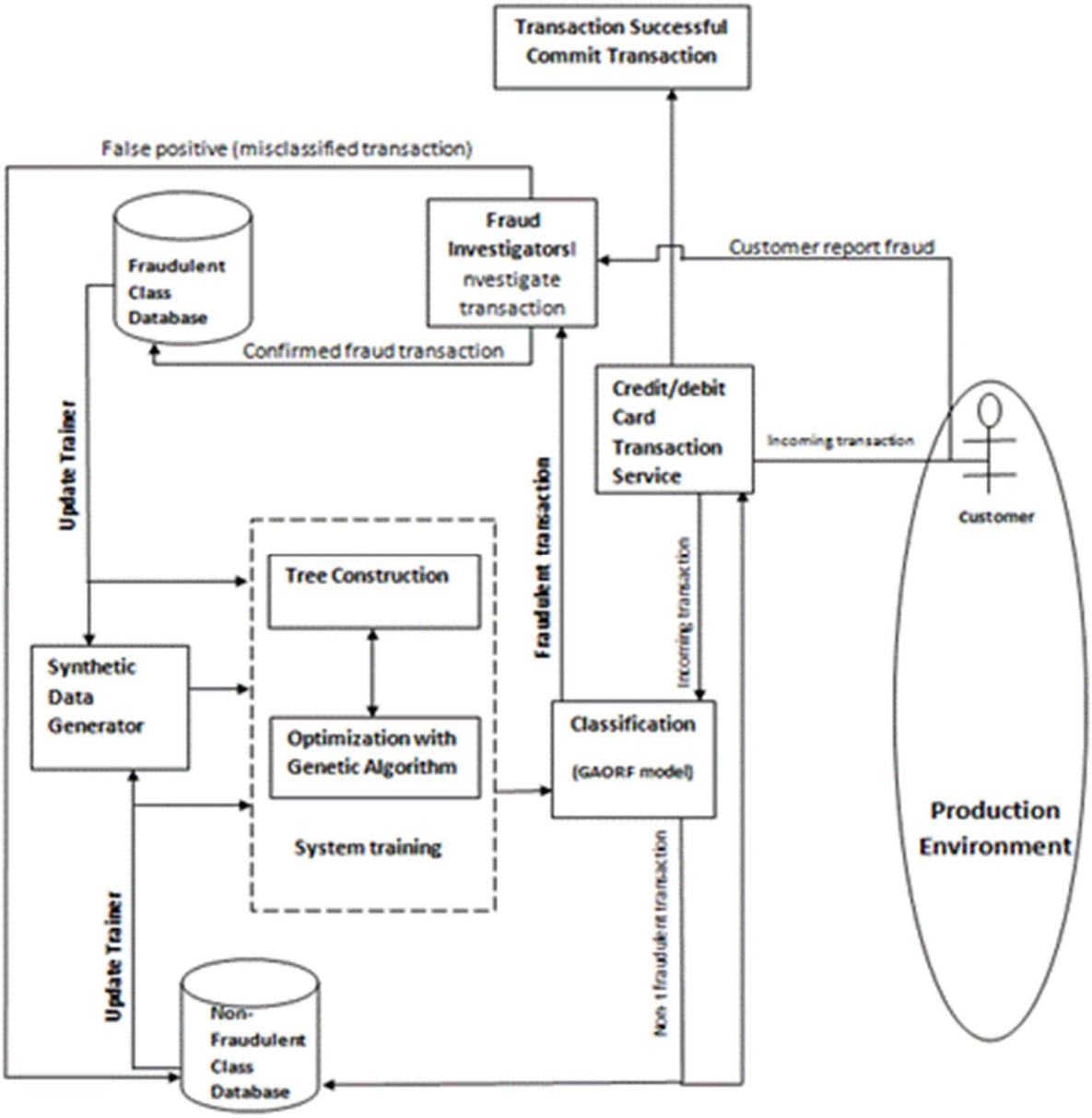


Figure 3: The expanded architectural overview of the proposed Fraud Prevention system

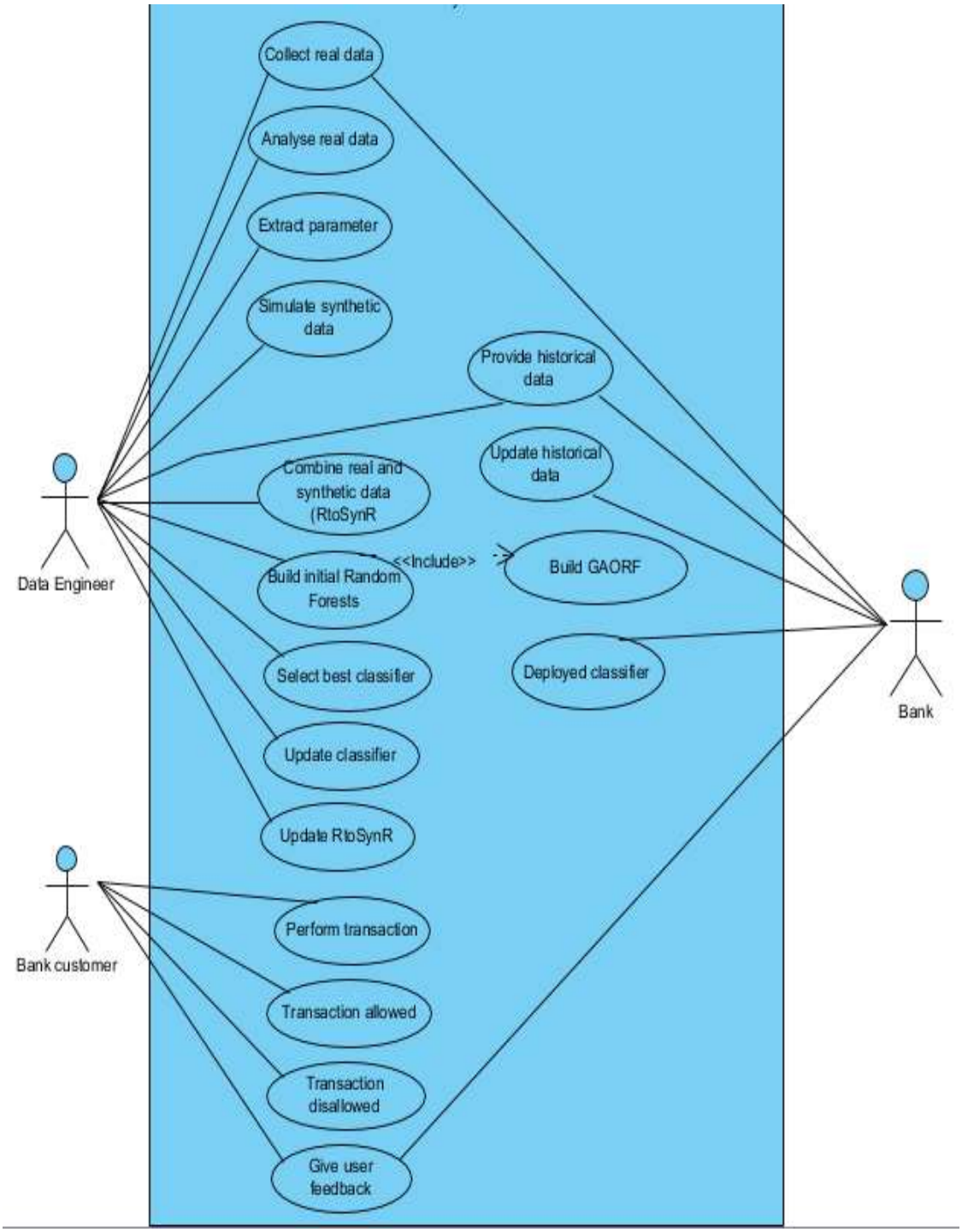


Figure 4: The use case diagram of the proposed system

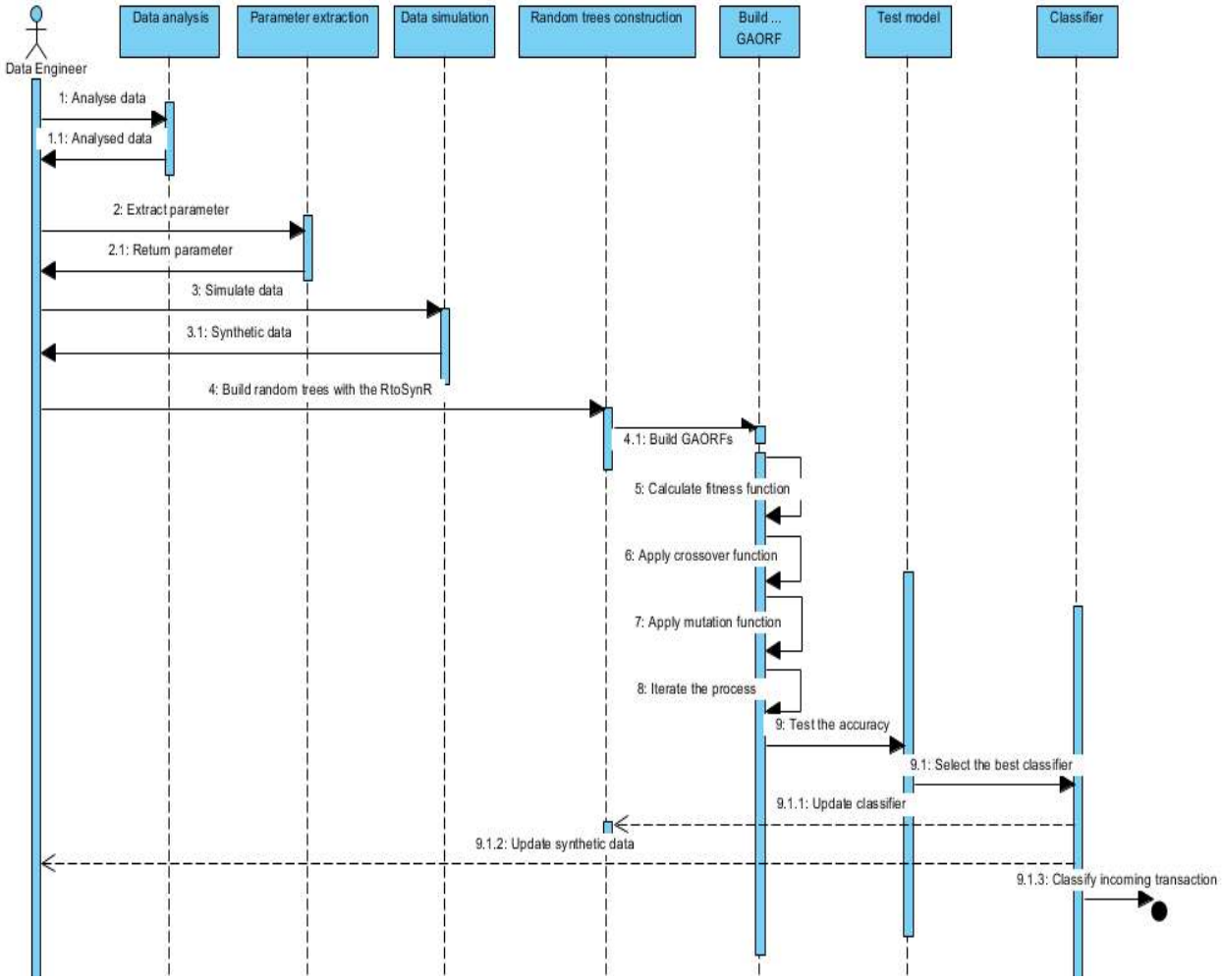


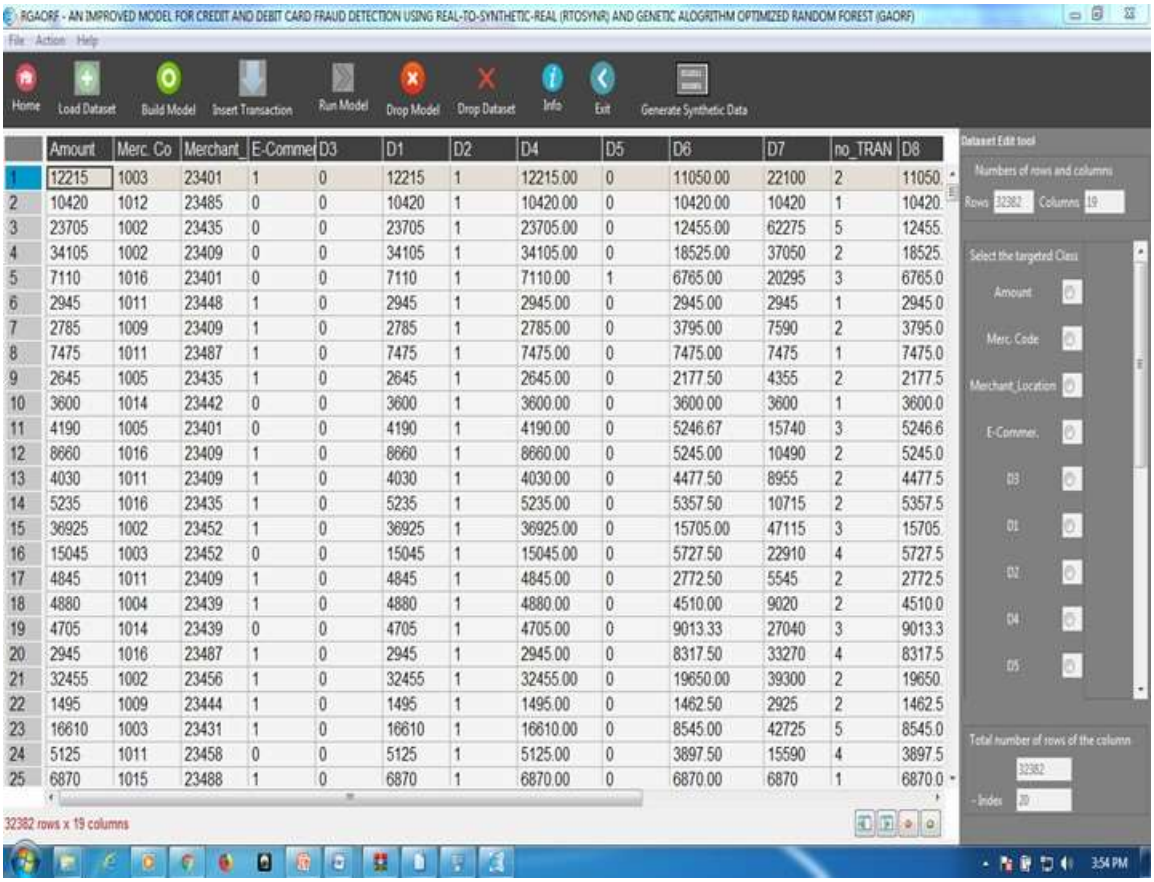
Figure 5: The sequence diagram of the proposed system

3.2 Dataset

The real dataset used in this work was provided by an anonymous commercial bank in Nigeria. The dataset was a select number of transactions that contained genuine and fraudulent transactions as two different datasets. The fraudulent transactions were reported fraudulent transactions from 300 customers (cardholders) in a 17 month period from February 2nd, 2016 to June 29th, 2017 numbering 300 records. The genuine transactions were unreported transactions from 1000 customers (cardholders) that occurred in a 6 month period from January 6th, 2017 to June 29th, 2017 numbering 26,082 records. In the course of the implementation of the work, derived features were added to the available features in order to provide more parameter features in the dataset for the training of the system.

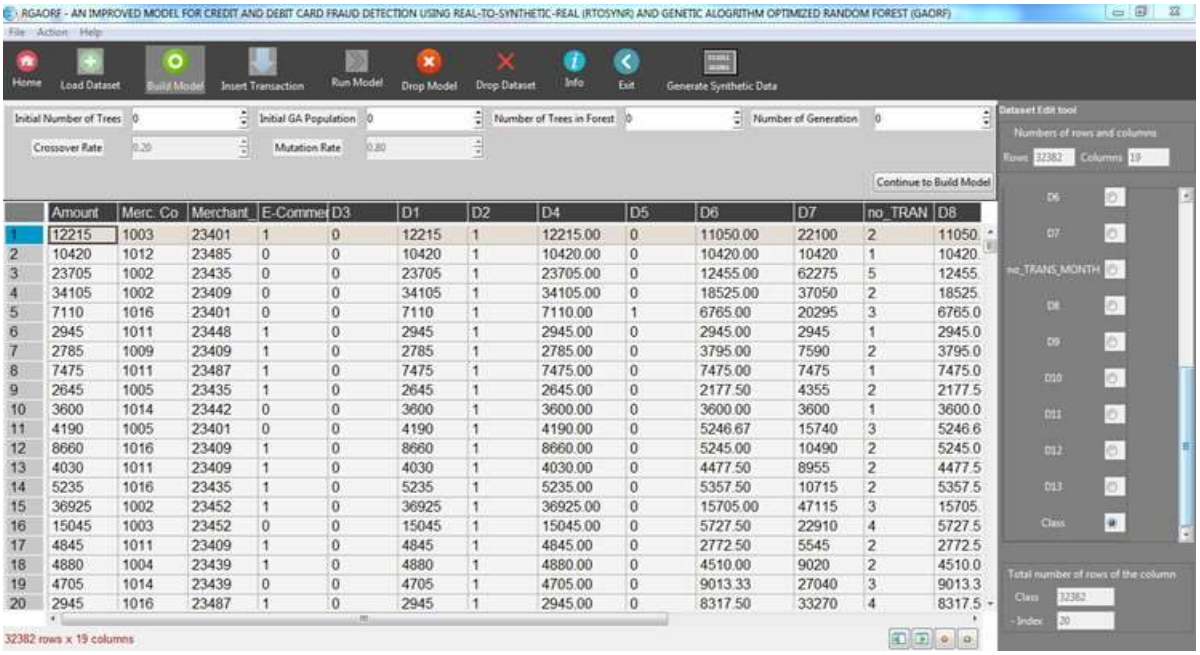
4. SYSTEM IMPLEMENTATION

The data simulation module was implemented on R programming language (R Core Team, 2018). The Genetic Algorithm Optimized Random Forests was developed in Python 3.6 platform using the Science Kit Learn Library available on the Anaconda 5.2 data science package for windows with Jupyter Notebook 5.5.0 used as the code editor to edit and test the codes during the developmental stage. The Graphical User Interface was developed using the Python Tkinter package, which is a toolkit for the programming of Graphical User Interface. Figures 6 through 10 show the screenshots of the system interface and output screen.



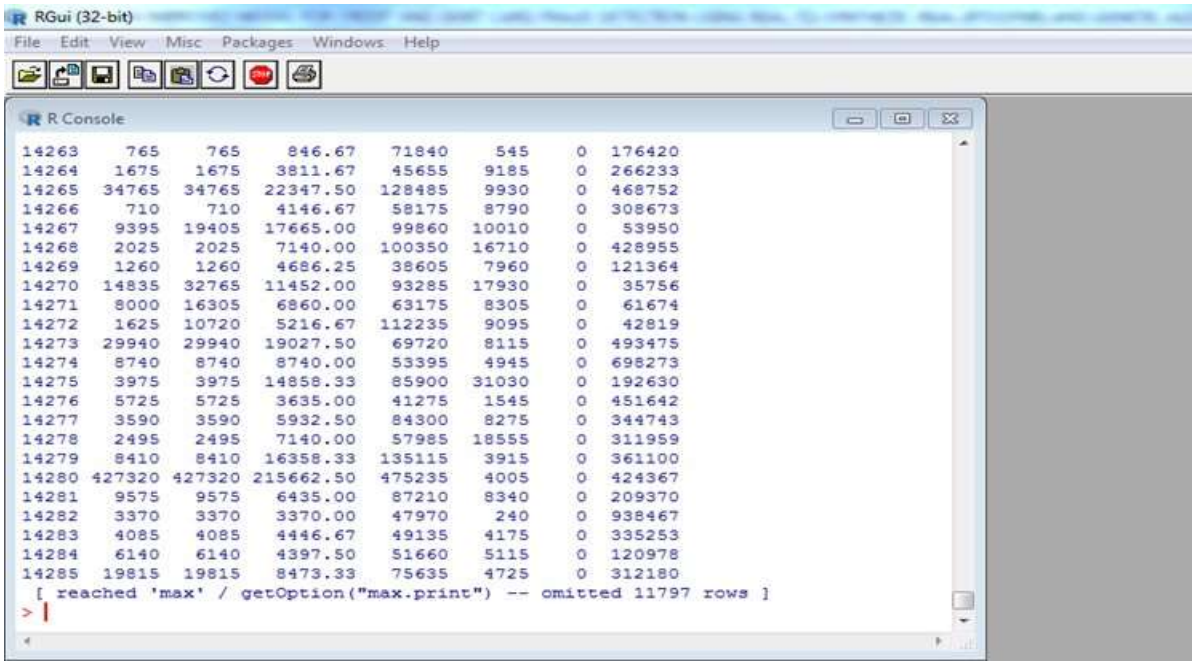
	Amount	Merc. Co	Merchant_	E-Commer	D3	D1	D2	D4	D5	D6	D7	no_TRAN	D8
1	12215	1003	23401	1	0	12215	1	12215.00	0	11050.00	22100	2	11050.
2	10420	1012	23485	0	0	10420	1	10420.00	0	10420.00	10420	1	10420.
3	23705	1002	23435	0	0	23705	1	23705.00	0	12455.00	62275	5	12455.
4	34105	1002	23409	0	0	34105	1	34105.00	0	18525.00	37050	2	18525.
5	7110	1016	23401	0	0	7110	1	7110.00	1	6765.00	20295	3	6765.0
6	2945	1011	23448	1	0	2945	1	2945.00	0	2945.00	2945	1	2945.0
7	2785	1009	23409	1	0	2785	1	2785.00	0	3795.00	7590	2	3795.0
8	7475	1011	23487	1	0	7475	1	7475.00	0	7475.00	7475	1	7475.0
9	2645	1005	23435	1	0	2645	1	2645.00	0	2177.50	4355	2	2177.5
10	3600	1014	23442	0	0	3600	1	3600.00	0	3600.00	3600	1	3600.0
11	4190	1005	23401	0	0	4190	1	4190.00	0	5246.67	15740	3	5246.6
12	8660	1016	23409	1	0	8660	1	8660.00	0	5245.00	10490	2	5245.0
13	4030	1011	23409	1	0	4030	1	4030.00	0	4477.50	8955	2	4477.5
14	5235	1016	23435	1	0	5235	1	5235.00	0	5357.50	10715	2	5357.5
15	36925	1002	23452	1	0	36925	1	36925.00	0	15705.00	47115	3	15705.
16	15045	1003	23452	0	0	15045	1	15045.00	0	5727.50	22910	4	5727.5
17	4845	1011	23409	1	0	4845	1	4845.00	0	2772.50	5545	2	2772.5
18	4880	1004	23439	1	0	4880	1	4880.00	0	4510.00	9020	2	4510.0
19	4705	1014	23439	0	0	4705	1	4705.00	0	9013.33	27040	3	9013.3
20	2945	1016	23487	1	0	2945	1	2945.00	0	8317.50	33270	4	8317.5
21	32455	1002	23456	1	0	32455	1	32455.00	0	19650.00	39300	2	19650.
22	1495	1009	23444	1	0	1495	1	1495.00	0	1462.50	2925	2	1462.5
23	16610	1003	23431	1	0	16610	1	16610.00	0	8545.00	42725	5	8545.0
24	5125	1011	23458	0	0	5125	1	5125.00	0	3897.50	15590	4	3897.5
25	6870	1015	23488	1	0	6870	1	6870.00	0	6870.00	6870	1	6870.0

Figure 6: Loaded dataset on the *GAORF-RtoSynR* model interface



	Amount	Merc. Co	Merchant_	E-Commer	D3	D1	D2	D4	D5	D6	D7	no_TRAN	D8
1	12215	1003	23401	1	0	12215	1	12215.00	0	11050.00	22100	2	11050.
2	10420	1012	23485	0	0	10420	1	10420.00	0	10420.00	10420	1	10420.
3	23705	1002	23435	0	0	23705	1	23705.00	0	12455.00	62275	5	12455.
4	34105	1002	23409	0	0	34105	1	34105.00	0	18525.00	37050	2	18525.
5	7110	1016	23401	0	0	7110	1	7110.00	1	6765.00	20295	3	6765.0
6	2945	1011	23448	1	0	2945	1	2945.00	0	2945.00	2945	1	2945.0
7	2785	1009	23409	1	0	2785	1	2785.00	0	3795.00	7590	2	3795.0
8	7475	1011	23487	1	0	7475	1	7475.00	0	7475.00	7475	1	7475.0
9	2645	1005	23435	1	0	2645	1	2645.00	0	2177.50	4355	2	2177.5
10	3600	1014	23442	0	0	3600	1	3600.00	0	3600.00	3600	1	3600.0
11	4190	1005	23401	0	0	4190	1	4190.00	0	5246.67	15740	3	5246.6
12	8660	1016	23409	1	0	8660	1	8660.00	0	5245.00	10490	2	5245.0
13	4030	1011	23409	1	0	4030	1	4030.00	0	4477.50	8955	2	4477.5
14	5235	1016	23435	1	0	5235	1	5235.00	0	5357.50	10715	2	5357.5
15	36925	1002	23452	1	0	36925	1	36925.00	0	15705.00	47115	3	15705.
16	15045	1003	23452	0	0	15045	1	15045.00	0	5727.50	22910	4	5727.5
17	4845	1011	23409	1	0	4845	1	4845.00	0	2772.50	5545	2	2772.5
18	4880	1004	23439	1	0	4880	1	4880.00	0	4510.00	9020	2	4510.0
19	4705	1014	23439	0	0	4705	1	4705.00	0	9013.33	27040	3	9013.3
20	2945	1016	23487	1	0	2945	1	2945.00	0	8317.50	33270	4	8317.5

Figure 7: Genetic Algorithm parameter specification interface



14263	765	765	846.67	71840	545	0	176420
14264	1675	1675	3811.67	45655	9185	0	266233
14265	34765	34765	22347.50	128485	9930	0	468752
14266	710	710	4146.67	58175	8790	0	308673
14267	9395	19405	17665.00	99860	10010	0	53950
14268	2025	2025	7140.00	100350	16710	0	428955
14269	1260	1260	4686.25	38605	7960	0	121364
14270	14835	32765	11452.00	93285	17930	0	35756
14271	8000	16305	6860.00	63175	8305	0	61674
14272	1625	10720	5216.67	112235	9095	0	42819
14273	29940	29940	19027.50	69720	8115	0	493475
14274	8740	8740	8740.00	53395	4945	0	698273
14275	3975	3975	14858.33	85900	31030	0	192630
14276	5725	5725	3635.00	41275	1545	0	451642
14277	3590	3590	5932.50	84300	8275	0	344743
14278	2495	2495	7140.00	57985	18555	0	311959
14279	8410	8410	16358.33	135115	3915	0	361100
14280	427320	427320	215662.50	475235	4005	0	424367
14281	9575	9575	6435.00	87210	8340	0	209370
14282	3370	3370	3370.00	47970	240	0	938467
14283	4085	4085	4446.67	49135	4175	0	335253
14284	6140	6140	4397.50	51660	5115	0	120978
14285	19815	19815	8473.33	75635	4725	0	312180

[reached 'max' / getOption("max.print") -- omitted 11797 rows]

Figure 8: Simulated data on the data generation module

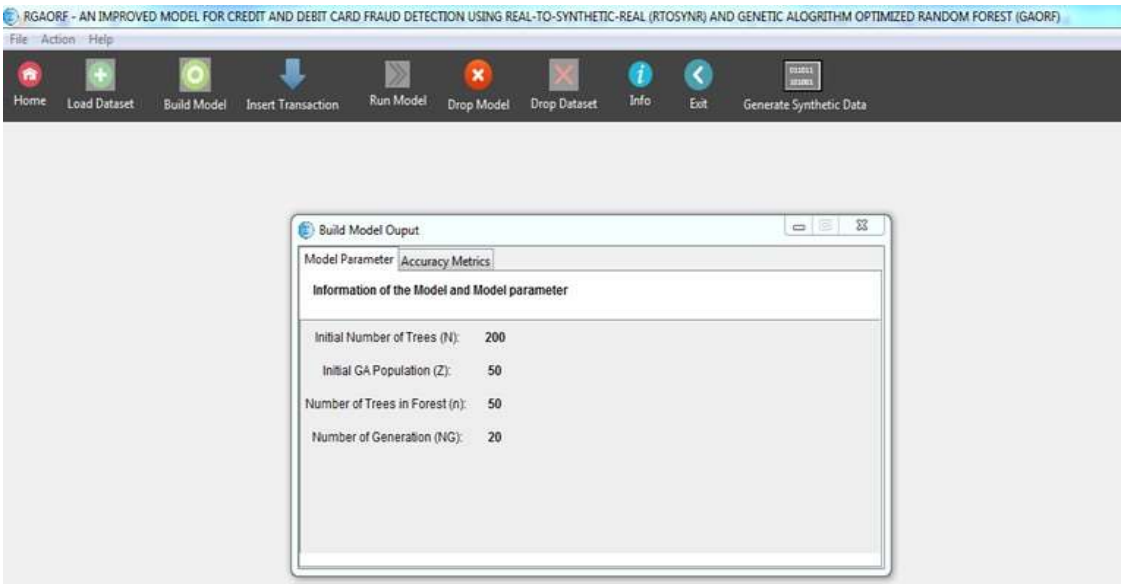


Figure 9:Output of *GAORF* training

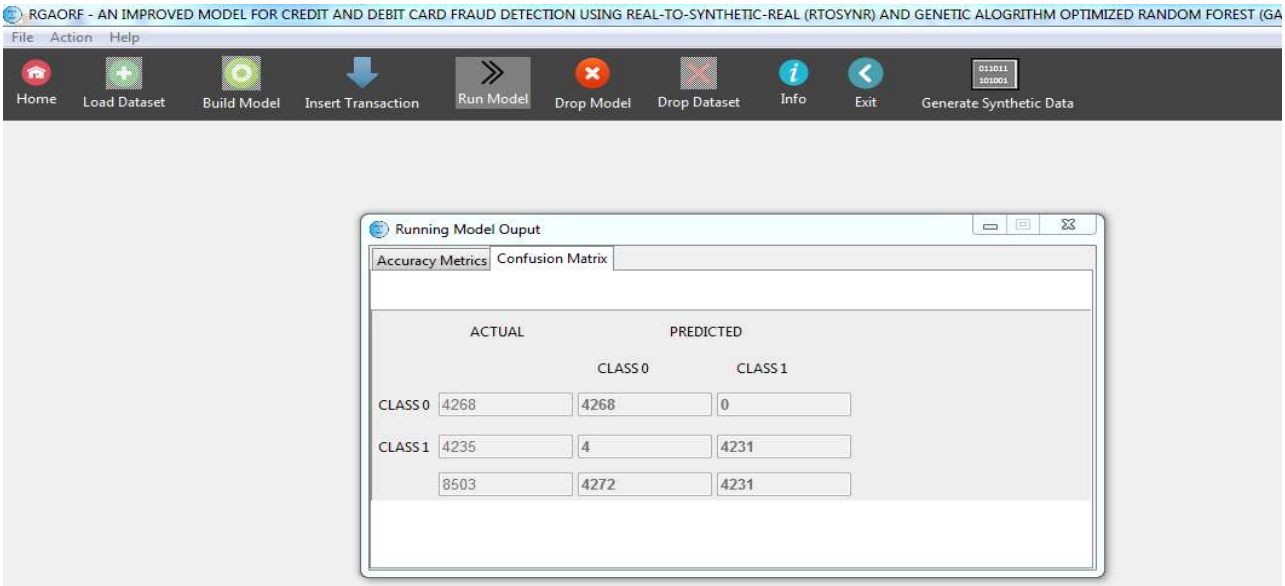


Figure 10: Confusion matrix output of the predicted transactions

5. RESULTS AND DISCUSSION

To test the accuracy of the developed system which is the *GAORF-RtoSynR* model over the existing Random Forests-Real model, we compared the results of *GAORF-RtoSynR* model against the Random Forests-Real model implemented on Salford Predictive Modeler 8.2. The results show that the *GAORF-RtoSynR* model was able to improve the accuracy of the fraud detection system, by reducing the misclassification from 1081 to 4 for all test set proportion sizes in a forest size of 100 trees. Similarly, the number of misclassification was reduced from 1423 to 5 for all test set proportions sizes for a forest size of 200 trees. Table 1 shows the overall accuracy metrics for *GAORF-RtoSynR* model over the existing Random Forests-Real model. Similarly, figures 11 and 12 show the accuracy plots of *GAORF-RtoSynR* model against Random Forests-Real model for the forests of sizes 100 and 200 trees respectively plotted on Python 3.6 platform using the matplotlib. The figures show a great improvement on accuracy in the *GAORF-RtoSynR* model over a Random Forests-Real model.

Table 1: Accuracy metrics of Random Forests-Real model of 100 and 200 trees compared with *GAORF-RtoSynR* model of equivalent tree sizes using genetic algorithm generations value of 20, mutation rate of 0.2 and crossover rate of 0.8

Test proportion	Test parameter	No of Trees			
		100 Trees		200 Trees	
		Random Forests-Real dataset	*GAORF-RtoSynR*	*Random Forests-Real dataset*	*GAORF-RtoSynR*
0.1	Accuracy	99.02%	99.97%	98.50%	99.97%
	Specificity	99.01%	99.96%	98.48%	99.96%
	Sensitivity/ Recall	100.00%	100.00%	100.00%	100.00%
	Precision	51.85%	99.84%	41.18%	99.84%
	F1 Satistics	68.29%	99.92%	58.33%	99.92%
0.2	Accuracy	90.26%	99.98%	89.81%	100.00%
	Specificity	90.15%	99.98%	89.69%	100.00%
	Sensitivity/ Recall	100.00%	100.00%	100.00%	100.00%
	Precision	10.38%	99.92%	9.97%	100.00%
	F1 Satistics	18.81%	99.96%	18.13%	100.00%
0.3	Accuracy	99.21%	100.00%	98.85%	99.98%
	Specificity	99.20%	100.00%	98.84%	99.99%
	Sensitivity/ Recall	100.00%	100.00%	100.00%	99.95%
	Precision	61.01%	100.00%	51.87%	99.95%
	F1 Satistics	75.78%	100.00%	68.31%	99.95%
0.4	Accuracy	96.84%	99.99%	96.31%	99.99%
	Specificity	96.80%	99.99%	96.27%	99.99%
	Sensitivity/ Recall	100.00%	100.00%	100.00%	100.00%
	Precision	28.08%	99.96%	25.05%	99.96%
	F1 Satistics	43.84%	99.98%	40.06%	99.98%
0.5	Accuracy	98.92%	99.99%	97.25%	99.99%
	Specificity	98.91%	99.99%	97.22%	99.99%
	Sensitivity/ Recall	100.00%	100.00%	100.00%	100.00%
	Precision	51.54%	99.97%	29.43%	99.97%
	F1 Satistics	68.02%	99.98%	45.48%	99.98%

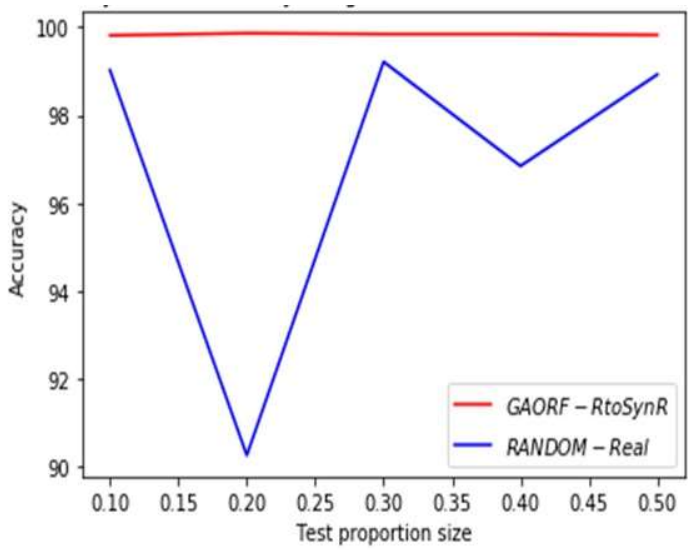


Figure 11: Accuracy plots of *RtoSynR* dataset trained on GAORF against Real dataset trained on Random Forests with 100 trees in the forests

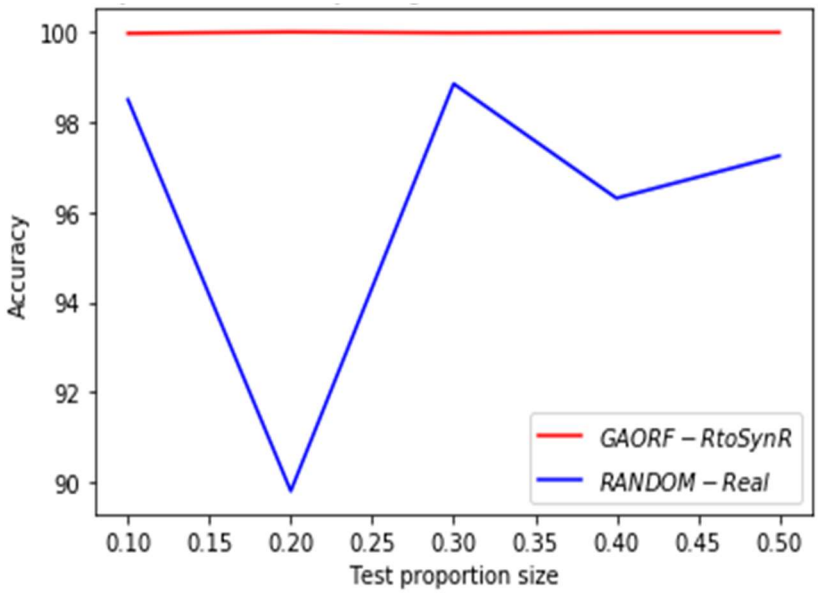


Figure 12: Accuracy plots of *RtoSynR* dataset trained on GAORF against Real dataset trained on Random Forests with 200 trees in the forests

6. CONCLUSION

This work has implemented a hybrid system for the prediction of credit and debit card transaction frauds in an online environment. The use of Real-to-Synthetic-Real and Genetic Algorithm Optimized Random Forests models has improved the classification accuracy of this system. This can help solve the problem of lack of transaction data in the domain of credit card fraud detection research and poor optimization and convergence of Random Forests algorithms. The system has good improvement bringing down the total misclassifications from 2504 to 9 for all tests proportions and tree sizes of 100 and 200 in forests.

REFERENCES

1. Adrian, B. (2015). Emerging Markets Queries in Finance and Business Detecting and Preventing Fraud with Data Analytics. *Elsevier Procedia Economics and Finance, 32*, 1827 – 1836.
2. Andrea, D. P., C. Olivier, L. B. Yann-Ael, W. Serge and B. Gianluca (2014). Learned lessons in Credit Card Fraud Detection From a Practitioners Perspective. *Expert Systems with Applications, 41*(10), 4915–4928.
3. Bharathidason, S., & Venkataeswaran, C. J. (2014). Improving Classification Accuracy Based on Random Forest Model with Uncorrelated High Performing Trees. *International Journal of Computer Applications, 101*(13).
4. Business Wire. (2011). U.S. Leads the World in Credit Card Fraud, States. *The Nilson Report*. Retrieved January 10th, 2017, from http://www.businesswire.com/news/home/20111121005121/en/U.S.-Leads-World-Credit-Card-Fraud-states
5. Christopher, M. B. (2006). *Pattern Recognition and Machine Learning.* New York: Springer.
6. David, J. H. (2010). Fraud Detection in Telecommunications and Banking Discussion of Becker, Volinsky, and Wilks (2010) and Sudjianto et al. *Technometrics, 52*(1), 34-38.
7. Emilie, L. B., and J. Erland (2003). Synthesizing Test Data for Fraud Detection Systems. *Proceedings of Nineteenth Annual Computer Security Applications Conference.* Las Vegas, Nevada. Retrieved January 3rd, 2017, from https://www.acsac.org/2003/papers/74.pdf
8. Haibo, H., and A. G. Edwardo (2009). Learning from Imbalanced Data. *Knowledge and Data Engineering, 21*(9), 1263–1284.
9. Ishu, T., M. Mrigya and Monika. (2016). Credit Card Fraud Detection. *International Journal of Advanced Research in Computer and Communication Engineering, 5*(1).
10. Jon, T., and M. Sriganesh (2008). Real-Time Credit Card Fraud Detection Using Computational Intelligence. *Expert Systems with Applications , 35*(4), 1721–1732.
11. Koosha, G., and R. Z. Osmar (2012). *Data Mining Applications for Fraud Detection in Securities Market.* Department of Computing Science University of Alberta, Edmonton. Retrieved February 4th, 2017, from https://webdocs.cs.ualberta.ca/~zaiane/postscript/EISIC12.pdf
12. Lopez-Rojas, E. A., and S. Axelsson (2012). Multi Agent Based Simulation of Financial Transaction. *Proceedings of the Nordic Conference on Secure IT systems.* Karlskrona, Sweden. Retrieved from http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A834702anddswid=1933
13. Neha, P., W. Roy and V. Kalyan (2016). The Synthetic Data Vault. *proceedings of th2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA).* doi:10.1109/DSAA.2016.49
14. Nwogu, E., and M. Odoh (2014). Security Issues Analysis on Online Banking Implementations in Nigeria. *International Journal of Computer Science and Telecommunications, 6*(1), 20-27.
15. Nwogu, E. R., E. O. Nwachukwu and V. E. Ejiofor (2019). On Real - to - Synthetic (*Rtosyn*) Approach for Generating Synthetic Transaction Data for Fraud Detection Using Hybrid Generative Model, in press.

16. Nwogu, E. R., E. O. Nwachukwu and V. E. Ejiofor (2019). Genetic Algorithm Optimized Random Forests (*GAORF*): a Model for Optimizing and Improving the Accuracy of Random Forests Models, in press.
17. Oberoi, R. (2017). Credit – Card Fraud Detection System: Using Genetic Algorithm. *International Journal of Computer & Mathematical Sciences, 6* (6).
18. Potamitis, G. (2013). *Design and Implementation of a Fraud Detection Expert System using Ontology Based Techniques.* Masters Thesis. Retrieved January 10th, 2017, from https://studentnet.cs.manchester.ac.uk/resources/.../Potamitis-Giannis-fulltext.pdf
19. R Core Team. (2018). *R: A language and environment for statistical.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/
20. Raghavendra, P., and S. Lokesh (2011). Credit Card Fraud Detection Using Neural Network. *International Journal of Soft Computing and Engineering (IJSCE), 1*.
21. RamaKalyani, K., and D. UmaDevi (2012). Fraud Detection of Credit Card Payment System by Genetic Algorithm. *International Journal of Scientific and Engineering Research, 3*(7).
22. Richard, J. B., and J. H. David (2002). *Unsupervised Profiling Methods for Fraud Detection.* Retrieved December 16th, 2016, from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.24.5743andrep=rep1andtype=pdf
23. Richard, J. B., and J. H. David (2002). Statistical Fraud Detection :A Review. *Statistical Science, 17*(3), 235-249.
24. Sahin, Y., and E. Duman (2011). Detecting Credit Card Fraud by Decision Trees and Support Vector Machines. *Proceedings of the International MultiConference of Engineers and Computer Scientist 2011.* Honkong. Retrieved February 10th, 2018, from www.iaeng.org/publication/IMECS2011/IMECS2011_pp442-447.pdf
25. Sonepat, S. H., and R. H. Sonepat (2011). Analysis on Credit Card Fraud Detection Methods. *International Journal of Computer Trends and Technology (IJCTT), 8*(1).
26. Vijayshree, B. N., S. K. Poonam, V. Dipali, W. Kunal and P. D. Bhagyashree (2016). Fraudulent Detection in Credit Card System Using SVM and Decision Tree. *International Journal of Scientific Development and Research (IJSDR), 1*(5).
27. Zhang, X., F. Yanwei, Z. Andi, S. Leonid and A. Gady (2015). *Learning Classifiers from Synthetic Data Using a Multichannel Autoencoder.* Retrieved December 10th, 2016, from https://pdfs.semanticscholar.org/2071/7f1cb12ab208458c0f2505b237d8f061f97a.pdf