Fair Value : Jurnal Ilmiah Akuntansi dan Keuangan
Volume 5, Number 8, 2023
P-ISSN: 2622-2191 E-ISSN : 2622-2205
Open Access: https://journal.ikopin.ac.id/index.php/fairvalue

Utilization of machine learning to detect the possibility of suspicious financial transactions

Dina Anggraeni[1], Siti Nurwahyuningsih Harahap[2]
[1,2] University of Indonesia
dina.anggraeni@ui.ac.id

Article Info

Article history:
Received Jan 12th 2023
Revised Feb 18th 2023
Accepted Mar 4th 2023

Keyword:
KNIME; Machine learning; Detect suspicious transactions; Financial transactions; Supervised learning algorithm

ABSTRACT

In order to prevent money laundering and terrorism financing, it is critical for banks to develop an effective mechanism to detect suspicious transactions. Nowadays, one of the most widely developed methods is machine learning. This article aims to discuss the best algorithm model in machine learning to detect possibilities for Suspicious Financial Transactions in XYZ Bank. This research uses quantitative research methods. The machine learning method used is supervised machine learning, with three models compared: Decision Tree, Gradient Boosting, and Random Forest. The tool used is The Konstanz Information Miner (KNIME). According to the findings of the study, the best model for detecting the possibility of SFT in bank XYZ is random forest with an accuracy rate of 99,98%. Based on this level of accuracy, this study reveals that a machine learning approach using historical company data makes a significant contribution to XYZ bank in detecting Suspicious Financial Transactions.

INTRODUCTION

Money laundering and financing of terrorism are important issues for economies and financial institutions around the world. Criminals and terrorists exploit financial institutions as a means of carrying out organized, large-scale money laundering (Krishnapriya & Prabakaran, 2014). In the end, this puts financial institutions, like banks, in a difficult position when it comes to complying with regulations, preserving financial security, preserving reputation, and averting operational risks like liquidity crises and legal action. The Head of the Indonesian Financial Transaction Reports and Analysis Center (PPATK), Kiagus Ahmad Badaruddin, stated that the world is entering the "digital money laundering era". He also stated that, "Revenues generated from 11 transnational crimes, such as drug trafficking, illicit arms trafficking, and human trafficking, are estimated to range from US$1.6 trillion to US$2.2 trillion per year" (Wicaksono, 2020).

Banks, as parties directly involved in this matter, play an important role in detecting SFT. Banking is a business that relies heavily on customer trust and is strictly regulated by the government (Latif, 2020). Therefore, it is crucial for banks to constantly improve services in terms of customer protection while also ensuring that the company's operations are in accordance with applicable regulations. In 2021, Bank XYZ's Corporate Assurance has identified several primary focus areas of risks, including cybercrime and suspicious financial transactions. This is consistent with a study conducted by ORX (Operational Risk Horizon) in November 2021, which stated that information security risk, including cyber security, is the top risk today (Johnson, 2021). As a result, it is essential for Bank XYZ to improve the bank's ability to detect SFT.

Currently, Bank XYZ's SFT detection system is using a rule-based system to identify SFT. However, currently the use of the system is considered to have several challenges in dealing with SFT. Some of the challenges faced are insufficient number of AML compliance officers. Traditional rule-based systems can produce a large number of unnecessary alerts, which makes them extremely inefficient (Hossain et al., 2019). The time-consuming task of thoroughly checking each transaction for which the system generated an alert requires a sizable number of compliance officers and analysts. This manual work makes traditional solutions difficult to scale. Inability to scale would definitely aggravate clients and potentially interfere with their experience. A poor customer experience results from unnecessarily holding up or delaying payments because of an increased workload (Panjaitan, 2022).

Secondly, there is a chance of false positive cases; the old transaction monitoring systems generate alerts whenever they detect suspicious activity, based on a threshold. Nevertheless, since every system makes mistakes, not every alert identifies money laundering. The last reason is a rule-based system requires a lot of time to maintain in a world where laws and regulations are constantly changing. To be able to comply with the legislation, all entities must revise and update their system rules (Ayunita et al., 2019).

Previously there were 3 (three) studies regarding methods of classifying money laundering transactions at financial institutions. The accuracy of the algorithm model was examined in research by Kumar et al. (2021), Raiter (2021), and Vassallo et al. (2021) that compared various money laundering detection algorithms. However, the outcomes are also dissimilar. Based on these researches, three algorithms—Decision Tree, Random Forest, and Gradient Boosting—are thought to be the most effective at spotting money laundering in financial institutions. This leads one to the conclusion that the data pattern significantly determines how accurate the algorithm is. Thus, based on the algorithm in earlier research that is thought to be the most accurate, the optimal algorithm for this study will be chosen.

The goal of this research is to find the best model for predicting and detecting SFT possibilities in transactions at Bank XYZ by utilizing machine learning algorithms with three different algorithm models: Decision Tree, Gradient Boosting, and Random Forest. In determining the model, this research uses the tools “The Konstanz Information Miner” (KNIME) and adopts the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework which is commonly used for data analysis (Provost & Fawcett, 2013). The findings of this study are expected to help Bank XYZ in identifying SFT using machine learning, allowing the SFT identification process to be completed in a shorter and more precise time.

RESEARCH METHODS

This research uses quantitative research methods. Quantitative research is defined as research that uses a lot of numbers, starting from data collection, data analysis, and data presentation (Priadana & Sunarsi, 2021). This study is initiated to select the best model of machine learning algorithm in order to predict and detecting SFT possibilities in transactions of Bank XYZ. The three different algorithm models used in this study are Decision Tree, Gradient Boosting, and Random Forest (Rong et al., 2020). In the process of determining the model using the codification of the data mining process the Cross Industry Standard Process for Data Mining (CRISP-DM) with The Konstanz Information Miner (KNIME) as the tool. This study uses primary data that will be used transactions data during the period January 1, 2019 to December 30, 2021 in Bank XYZ by population and workflows of data mining that will be used is briefly described from the diagram below:

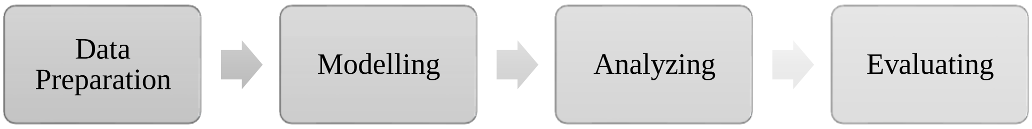
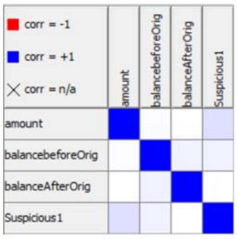


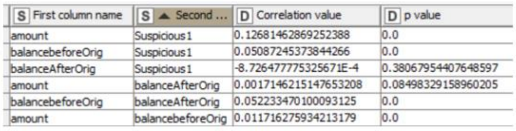
Figure 1 Research Workflow

RESULTS AND DISCUSSION

Data Preparation

Data Preparation is done in order to prevent data inconsistencies. Since the dataset being employed is an excel file, the data is imported at this point using the Excel Reader node to KNIME. Then, we found that some data is numerical but is interpreted as text by the software, so that we changed the data type from String to Number. After that, we used correlation testing, which examines the relationship between the variables in the data. The correlation testing result is shown from the Figure 2.





S First column name	S ▲ Second ...	D Correlation value	D p value
amount	Suspicious1	0.12681462869252388	0.0
balancebeforeOrig	Suspicious1	0.05087245373844266	0.0
balanceAfterOrig	Suspicious1	-8.726477775325671E-4	0.38067954407648597
amount	balanceAfterOrig	0.0017146215147653208	0.08498329158960205
balancebeforeOrig	balanceAfterOrig	0.052233470100093125	0.0
amount	balancebeforeOrig	0.011716275934213179	0.0

Figure 2 Linear Correlation

Based on the results of the linear correlation which can be seen from Figure 2, the variable 'amount' is the variable that has the most significant influence compared to other variables which can be seen from the correlation value of 0.1268 with a p-value of 0. Therefore, the 'amount' variable is used as a parameter to predict suspicious financial transactions.

Modelling

In this step, tests were run on a number of models, namely Random Forest, Decision Tree, and Gradient Boosting, in order to compare and determine which model is most effective at identifying indications of suspicious transactions in transactions (Leo et al., 2019). In the modelling process, 80% of the data is partitioned for learning and the remaining 20% is used for testing methods based on the Pareto Principle.

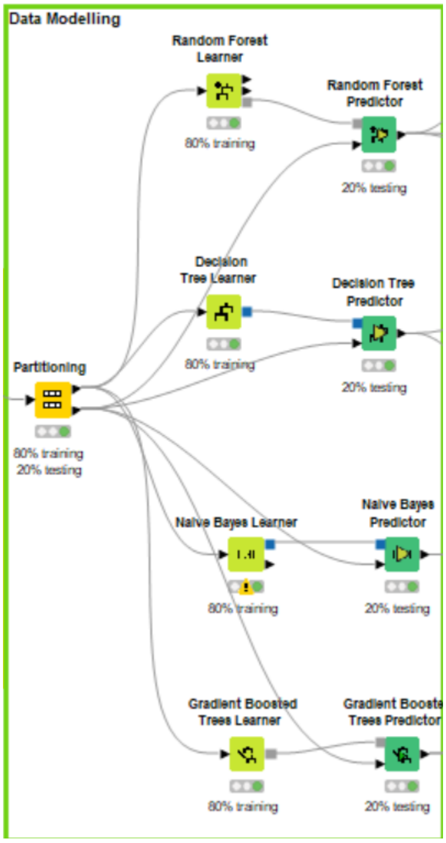
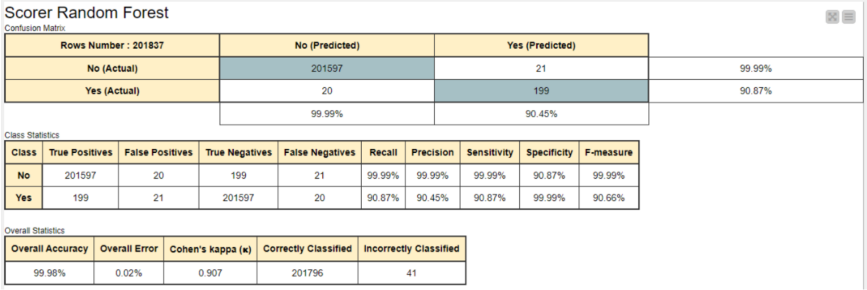


Figure 3 Modelling Workflow

Analyzing

In the analyzing process, the learning process for model accuracy is carried out by comparing actual data and predictions made by machine learning (Medar et al., 2017). Each model generates an accuracy value, which is displayed as a score. The best model with the most suitable level of accuracy will be examined based on the score. The results of the accuracy of each model can be seen from the scoreboard presented in Figures 4 to 6.

Random Forest



Scorer Random Forest

Confusion Matrix

Rows Number : 201837	No (Predicted)	Yes (Predicted)	
No (Actual)	201597	21	99.99%
Yes (Actual)	20	199	90.87%
	99.99%	90.45%	

Class Statistics

Class	True Positives	False Positives	True Negatives	False Negatives	Recall	Precision	Sensitivity	Specificity	F-measure
No	201597	20	199	21	99.99%	99.99%	99.99%	90.87%	99.99%
Yes	199	21	201597	20	90.87%	90.45%	90.87%	99.99%	90.66%

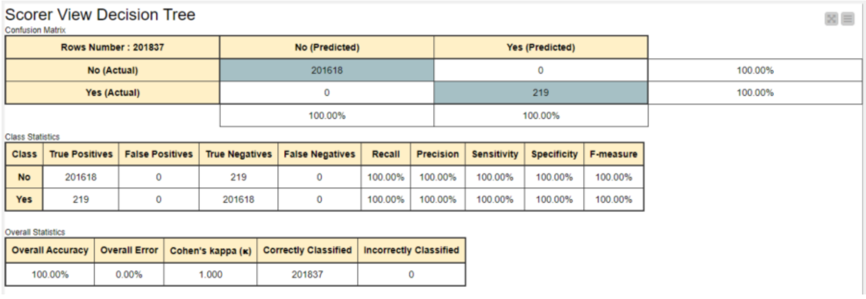
Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
99.98%	0.02%	0.907	201796	41

Figure 4 Random Forest Scoreboard

The random forest scoreboard in Figure 4 shows that the overall level of accuracy is 99.98% with an overall error rate of 0.02% which indicates a very high level of accuracy. Cohen's Kappa value is 0.907. Cohen's Kappa value in the Random Forest model is very good because it is close to +1, which means there is high consistency, reliability, and similarity in measurement. Out of a total of 201,837 transactions tested by the Random Forest Predictor, only 41 transactions were misclassified. Judging from the sensitivity value is also very good, namely 99.99%. Based on the assessment, the Random Forest model has a very good level of accuracy and reliability for predicting the probability of SFT occurring.

Decision Tree



Scorer View Decision Tree

Confusion Matrix

Rows Number : 201837	No (Predicted)	Yes (Predicted)	
No (Actual)	201618	0	100.00%
Yes (Actual)	0	219	100.00%
	100.00%	100.00%	

Class Statistics

Class	True Positives	False Positives	True Negatives	False Negatives	Recall	Precision	Sensitivity	Specificity	F-measure
No	201618	0	219	0	100.00%	100.00%	100.00%	100.00%	100.00%
Yes	219	0	201618	0	100.00%	100.00%	100.00%	100.00%	100.00%

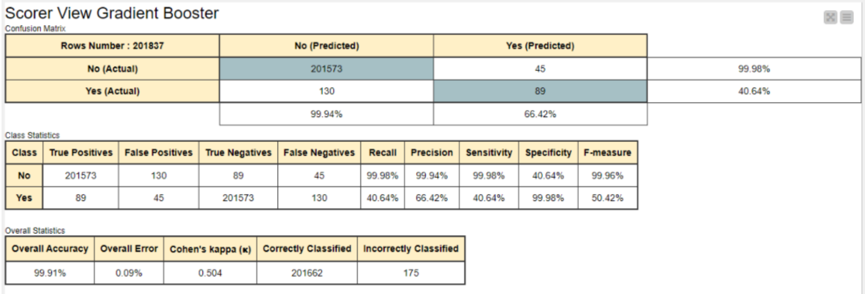
Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
100.00%	0.00%	1.000	201837	0

Figure 5 Decision Tree Scoreboard

Based on the scoreboard of the Decision Tree Model shown in Figure 5, the overall level of accuracy is 100%, there is no error in classification and Kappa Cohen is 1. A very perfect score for a model. Overall, the performance of the Decision Tree model looks very good. However, in choosing a model, it is necessary to consider whether there is a possibility of overfitting. Overfitting occurs when a model predicts variables very well in the original data, but does not estimate accurately in new data sets, causing the model to be too sensitive and predict poorly for other data. In the Decision Tree model, this can be seen from the sensitivity level of 100%. So, it is very possible that this model will be very sensitive and have very narrow criteria for predicting a transaction. Therefore, a model that is indicated as overfitting is a model that needs to be avoided.

Gradient Boosting



Scorer View Gradient Booster

Confusion Matrix

Rows Number : 201837	No (Predicted)	Yes (Predicted)	
No (Actual)	201573	45	99.98%
Yes (Actual)	130	89	40.64%
	99.94%	66.42%	

Class Statistics

Class	True Positives	False Positives	True Negatives	False Negatives	Recall	Precision	Sensitivity	Specificity	F-measure
No	201573	130	89	45	99.98%	99.94%	99.98%	40.64%	99.96%
Yes	89	45	201573	130	40.64%	66.42%	40.64%	99.98%	50.42%

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
99.91%	0.09%	0.504	201662	175

Figure 6 Gradient Boosting Scoreboard

Gradient Boosting’s scoreboard in Figure 6 reveals a 99.91% overall accuracy level and a 0.09% overall error level, both of which indicate exceptionally high levels of accuracy. The Kappa for Cohen is 0.504, which indicated unbalanced in this model. The accuracy score of 99.91% is largely impacted by NO predictions that are accurate, however there are a lot of misclassifications for YES predictions. This is evident from the 45 (66.42%) and 130 (40.64%) False Positive and False Negative cases, respectively, of the total transactions that should be projected as SFT. This is in line with the sensitivity level of YES category which only shows 40.46%.

Evaluating

Based on the explanation of the three models used in this study, it can be concluded that the best model for detecting SFT at XYZ Bank is the Random Forest. The application of machine learning has significantly aided Bank XYZ in its efforts to increase the effectiveness of SFT detection through study.

Efficiency of the compliance officer review process

Table 1 of the research's data shows that the old system's data analysis generated alerts for 2,390 transactions to be identified and investigated by Bank XYZ's compliance officers. Only 219 of the 2,390 transactions—or around 9.16% of the old system alerts—could truly be classified as suspicious, according to the Compliance Officer's analysis. The outcomes of the identification and inquiry are comparable to the 220 alerts, or around 9.21%, generated using machine learning predictions.

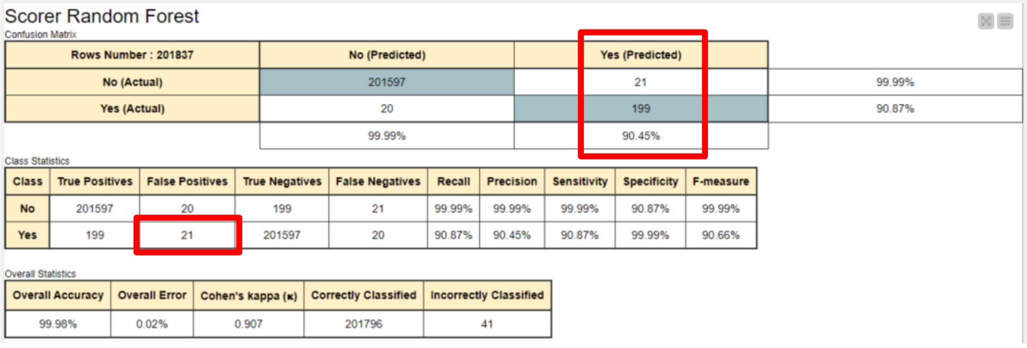
Table 1. Alert Comparison

Total *Alert* from Old System	Actual SFT	Prediction SFT
2.390 transactions	219 transactions	220 transactions

This can prove that the use of machine learning is more efficient than the use of the old system, because the prediction of alerts generated by machine learning is closer to the number of suspicious transactions from the identification and investigation results of compliance officers. As a result, compliance officers will be able to focus more on identifying transactions that are actually suspicious because their work will be more effective and efficient.

Minimize False Positive Cases

As previously stated, there are a lot of instances of false positives in the outdated system that XYZ Bank used, with false positive cases constituting up to 99% of the alerts produced. Contrary to the old system, using machine learning can reduce the possibility of false positive cases. This is seen in Figure 8 which demonstrates how false positive cases might be reduced.



Scorer Random Forest

Confusion Matrix

Rows Number : 201837	No (Predicted)	Yes (Predicted)	
No (Actual)	201597	21	99.99%
Yes (Actual)	20	199	90.87%
	99.99%	90.45%	

Class Statistics

Class	True Positives	False Positives	True Negatives	False Negatives	Recall	Precision	Sensitivity	Specificity	F-measure
No	201597	20	199	21	99.99%	99.99%	99.99%	90.87%	99.99%
Yes	199	21	201597	20	90.87%	90.45%	90.87%	99.99%	90.66%

Overall Statistics

Overall Accuracy	Overall Error	Cohen's kappa (κ)	Correctly Classified	Incorrectly Classified
99.98%	0.02%	0.907	201796	41

Figure 7 False Positive from Random Forest Scoreboard

In transactions that indicated SFT, the number of false positive cases is 21 out of 220 transactions (9.5% of the total transactions). Therefore, it can be concluded that the use of machine learning can reduce false positive cases in the future in the SFT prediction process at Bank XYZ.

A more adaptive SFT prediction system

The previous system that XYZ Bank utilized was a rule-based system, where the number of rules that needed to be defined depended on how many actions the system wanted to be able to handle. For example, 20 ‘rules’ require manual coding of 20 rules. The rule-based system used by XYZ Bank ultimately has low performance and is risky because the rules made will not change or update themselves if they are not processed or updated by the owner. According to the rules that have been implemented, the system will read as if it were 100% correct or 100% inaccurate, which has the potential to result in a false positive situation in relation to the previous point. Whereas when using machine learning is more adaptive because the model will learn, interpret, and group SFT based on transaction anomalies.

CONCLUSION

In this study, big data is used from transactions at XYZ Bank to detect suspicious financial transactions by utilizing machine learning as a tool to classify normal and suspicious transactions. The analysis was carried out by comparing 3 algorithm models: Decision Tree, Random Forest, and Gradient Boosting. Based on the research results, it can be concluded that after comparing the accuracy, sensitivity, and recall values, the Random Forest model has better performance than other models. This model give contribution to XYZ Bank regarding SFT prediction by reducing alerts which will increase the efficiency of the review process by compliance officers, the algorithm model produced by machine learning also can minimize false positive cases. The prediction system generated by machine learning is made based on behavioural pattern analysis from data not based on certain rules that purposely created to catch transactions that are suspected of being suspicious. Therefore, the model produced by machine learning will be more adaptive to changes in data and situations compared to the old system used by Bank XYZ. Given the limited data and the fact that this research can only use the parameter of the number of daily (regular) transactions to forecast SFT, it is recommended that future research be expanded with more data sets and alternative machine learning approaches, such as unsupervised learning.

REFERENCES

Ayunita, Y., Yahanan, A., & Syaifuddin, M. (2019). Perlindungan Hukum Terhadap Pengemudi Taksi (Mitra) Berbasis Online Pada PT. Grab Indonesia. *Lex Lata*, *1*(1).

Hossain, S., Sarma, D., Tuj-Johora, F., Bushra, J., Sen, S., & Taher, M. (2019). A belief rule based expert system to predict student performance under uncertainty. *2019 22nd International Conference on Computer and Information Technology (ICCIT)*, 1–6.

Johnson, S. (2021). *Top Risk Review November 2021*. https://managingrisktogether.orx.org/

Krishnapriya, G., & Prabakaran, M. (2014). Money laundering analysis based on time variant behavioral transaction patterns using data mining. *Journal of Theoretical and Applied Information Technology*, *67*(1), 12–17.

Kumar, A., Das, S., Tyagi, V., Shaw, R. N., & Ghosh, A. (2021). Analysis of classifier algorithms to detect anti-money laundering. *Computationally Intelligent Systems and Their Applications*, 143–152.

Latif, A. (2020). Konsep Hukum Sumber Dana dari Nasabah Penyimpan pada Bank Buku I di Indonesia dalam Menghindari Money Laundry. *Repertorium: Jurnal Ilmiah Hukum Kenotariatan*, *9*(1), 1–10.

Leo, M., Sharma, S., & Maddulety, K. (2019). Machine learning in banking risk management: A literature review. *Risks*, *7*(1), 29.

Medar, R., Rajpurohit, V. S., & Rashmi, B. (2017). Impact of training and testing data splits on accuracy of time series forecasting in machine learning. *2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, 1–6.

Panjaitan, H. T. (2022). *Pengaruh Skeptisisme Profesional, Keahlian Forensik, Tekanan Waktu, Dan Beban Kerja Terhadap Kemampuan Mendeteksi Kecurangan (Studi Pada Bpkp Perwakilan*

Provinsi Sumatera Utara). Universitas Atma Jaya Yogyakarta.

Priadana, M. S., & Sunarsi, D. (2021). *Metode Penelitian Kuantitatif*. Pascal Books.

Provost, F., & Fawcett, T. (2013). *Data Science for Business: What you need to know about data mining and data-analytic thinking*. " O'Reilly Media, Inc."

Raiter, O. (2021). Applying supervised machine learning algorithms for fraud detection in anti-money laundering. *Journal of Modern Issues in Business Research*, *1*(1), 14–26.

Rong, G., Alu, S., Li, K., Su, Y., Zhang, J., Zhang, Y., & Li, T. (2020). Rainfall induced landslide susceptibility mapping based on Bayesian optimized random forest and gradient boosting decision tree models—A case study of Shuicheng County, China. *Water*, *12*(11), 3066.

Vassallo, D., Vella, V., & Ellul, J. (2021). Application of gradient boosting algorithms for anti-money laundering in cryptocurrencies. *SN Computer Science*, *2*, 1–15.

Wicaksono, A. (2020). *PPATK Sebut "Virtual Currency" Bisa Mendanai Terorisme*. www.cnnindonesia.com