V8 list of defined criteria
KPB3 – Design of ML algorithms; type – list; month of submission – M12

Funded by the European Union NextGenerationEU | RECOVERY AND RESILIENCE | GOVERNMENT OFFICE OF THE SLOVAK REPUBLIC | UNIVERSITY OF ŽILINA Faculty of Management Science and Informatics | BRAIN:IT

List of defined criteria

	criterion name	brief description	explanation of point assignment	extended description
1	Anomaly detection accuracy (Precision, Recall, F1, Accuracy)	How accurately the algorithm detects anomalies regardless of class	1 – F1 below 0.7; 2 – F1 0.7–0.8; 3 – F1 0.8–0.9; 4 – F1 above 0.9; 5 – exceptional accuracy, even with unlabelled data.	Evaluation of overall accuracy using metrics such as Precision, Recall, F1-score, Accuracy or AUC-ROC.
2	False alarm rate (False Positives/Negatives)	How often the algorithm generates false positives/negatives	1 – FP/FN above 20%; 2 – FP/FN 10–20%; 3 – FP/FN 5–10%; 4 – below 5%; 5 – virtually zero false alarms.	Reducing the false alarm rate is key to usability in a SOC. Ideal models have a very low number of FP/FN.
3	Support for learning without labels (unsupervised)	Ability to learn from input data without the need for annotations	1 – supervised only; 2 – semi-supervised; 3 – basic unsupervised; 4 – advanced unsupervised models; 5 – advanced with a hybrid approach.	Not all models require labelled data. Unsupervised models are advantageous particularly with large amounts of unlabelled data.
4	Computational demands (CPU/GPU, memory)	Resources required for model training and inference	1 – demanding for both training and inference; 2 – high load; 3 – medium; 4 – optimized; 5 – lightweight deployment with low load.	The load on computational resources is assessed, including GPU and RAM demands and training duration.
5	Inference speed (detection time)	Response speed during detection in production	1 – more than 1s; 2 – hundreds of ms; 3 – tens of ms; 4 – below 10ms; 5 – real-time for large volumes.	How quickly the model can provide a result on new data. Critical for real-time deployment.
6	Robustness to changes in data distribution (drift, noise)	Resistance to noise, changes in data or attacks	1 – susceptible to noise; 2 – not very resistant; 3 – average robustness; 4 – stable; 5 – resistant even to adversarial inputs.	The model must be resistant to unexpected changes, fluctuations or noise, otherwise it will stop working correctly.
7	Ability to work with time-sequential data	Suitability for data evolving over time (e.g. logs, network traffic)	1 – unsuitable; 2 – only through workarounds; 3 – handles sequences; 4 – natively supports time series; 5 – optimized for sequences.	Ability to capture temporal patterns and anomalies in time series. Architectures such as LSTM and Transformer are an advantage.
8	Flexibility of input data (logs, metrics, flow data)	How many types of data the model can handle without the need to modify the architecture	1 – only one type; 2 – 2 types; 3 – basic flexibility; 4 – multiple types without modifications; 5 – universal model.	The model must be able to work with both structured and unstructured inputs – logs, metrics, network flows.
9	Possibility of adaptive/online learning	Support for updating the model during operation without retraining	1 – no possibility; 2 – via retraining; 3 – batch-based update; 4 – partial update; 5 – continual learning.	Ability to update the model without the need for complete retraining, e.g. by means of online learning.
10	Possibility of model explainability (interpretability)	Possibility of explaining the detection output to the analyst	1 – completely uninterpretable; 2 – poor explainability; 3 – partially explainable; 4 – good support; 5 – highly explainable with visualizations.	Possibility of providing the analyst with an understandable explanation of why the anomaly was detected.
11				
12				
13				
14				
15				

Points:
the point assignment logic for all criteria must be set up so that points can be assigned from 1 to 5, where 1 = the lowest level of fulfilment of the criterion and 5 = full fulfilment of the criterion.

Assignment of weights to individual criteria

	criterion name	weight assigned to the criterion
1	Anomaly detection accuracy (Precision, Recall, F1)	0,07
2	False alarm rate (False Positives/Negatives)	0,10
3	Support for learning without labels (unsupervised)	0,09
4	Computational demands (CPU/GPU, memory)	0,10
5	Inference speed (detection time)	0,15
6	Robustness to changes in data distribution (drift, noise)	0,09
7	Ability to work with time-sequential data	0,10
8	Flexibility of input data (logs, metrics, flow data)	0,11
9	Possibility of adaptive/online learning	0,09
10	Possibility of model explainability (interpretability)	0,10
11		
12		
13		
14		
15		
	sum of weights:	**1**

Assignment of points to individual criteria

available solutions	Criteria for the ML tool														
	Presnosť detekcie anomálií (Precision, Recall, F1, Accuracy)	Miera falošných poplachov (False Positives/Negatives)	Podpora učenia bez labelov (unsupervised)	Výpočtová náročnosť (CPU/GPU, pamäť)	Rýchlosť inferencie (čas detekcie)	Robustnosť voči zmene distribúcie dát (drift, noise)	Schopnosť práce s časovo-sekvenčnými dátami	Flexibilita vstupných údajov (logy, metriky, tokové dáta)	Možnosť adaptívneho/online učenia	Možnosť vysvetliteľnosti modelu (interpretability)	0	0	0	0	0
Random Forest	4	3	1	2	4	3	2	2	2	3					
Isolation Forest	4	4	5	3	4	4	2	4	2	2					
Autoencoder	4	4	5	3	3	4	3	4	3	2					
LSTM	4	4	3	2	2	4	5	4	4	3					
Transformer	5	5	5	1	2	4	5	5	5	1					
One-Class SVM	3	3	4	3	3	3	2	2	2	2					
GAN	4	3	5	1	2	4	3	3	2	1					
k-Nearest Neighbors (k-NN)	3	2	1	2	2	2	2	2	2	3					
Naive Bayes	2	2	1	2	4	2	1	1	1	3					
XGBoost	5	5	2	4	4	4	3	4	3	3					
Decision Tree	3	3	1	2	4	2	2	2	2	4					

Recalculation taking into account the weights of individual criteria

available solutions	Kritériá pre nástroj ML															resulting total for the given solution
	Presnosť detekcie anomálií (Precision, Recall, F1, Accuracy)	Miera falošných poplachov (False Positives/Negatives)	Podpora učenia bez labelov (unsupervised)	Výpočtová náročnosť (CPU/GPU, pamäť)	Rýchlosť inferencie (čas detekcie)	Robustnosť voči zmene distribúcie dát (drift, noise)	Schopnosť práce s časovo-sekvenčnými dátami	Flexibilita vstupných údajov (logy, metriky, tokové dáta)	Možnosť adaptívneho/online učenia	Možnosť vysvetliteľnosti modelu (interpretability)	0	0	0	0	0	
weights	0,07	0,10	0,09	0,10	0,15	0,09	0,10	0,11	0,09	0,10	0,00	0,00	0,00	0,00	0,00	
Random Forest	0,28	0,30	0,09	0,20	0,60	0,27	0,20	0,22	0,18	0,30						2,64
Isolation Forest	0,28	0,40	0,45	0,30	0,60	0,36	0,20	0,44	0,18	0,20						3,41
Autoencoder	0,28	0,40	0,45	0,30	0,45	0,36	0,30	0,44	0,27	0,20						3,45
LSTM	0,28	0,40	0,27	0,20	0,30	0,36	0,50	0,44	0,36	0,30						3,41
Transformer	0,35	0,50	0,45	0,10	0,30	0,36	0,50	0,55	0,45	0,10						3,66
One-Class SVM	0,21	0,30	0,36	0,30	0,45	0,27	0,20	0,22	0,18	0,20						2,69
GAN	0,28	0,30	0,45	0,10	0,30	0,36	0,30	0,33	0,18	0,10						2,70
k-Nearest Neighbors (k-NN)	0,21	0,20	0,09	0,20	0,30	0,18	0,20	0,22	0,18	0,30						2,08
Naive Bayes	0,14	0,20	0,09	0,20	0,60	0,18	0,10	0,11	0,09	0,30						2,01
XGBoost	0,35	0,50	0,18	0,40	0,60	0,36	0,30	0,44	0,27	0,30						3,70
Decision Tree	0,21	0,30	0,09	0,20	0,60	0,18	0,20	0,22	0,18	0,40						2,58