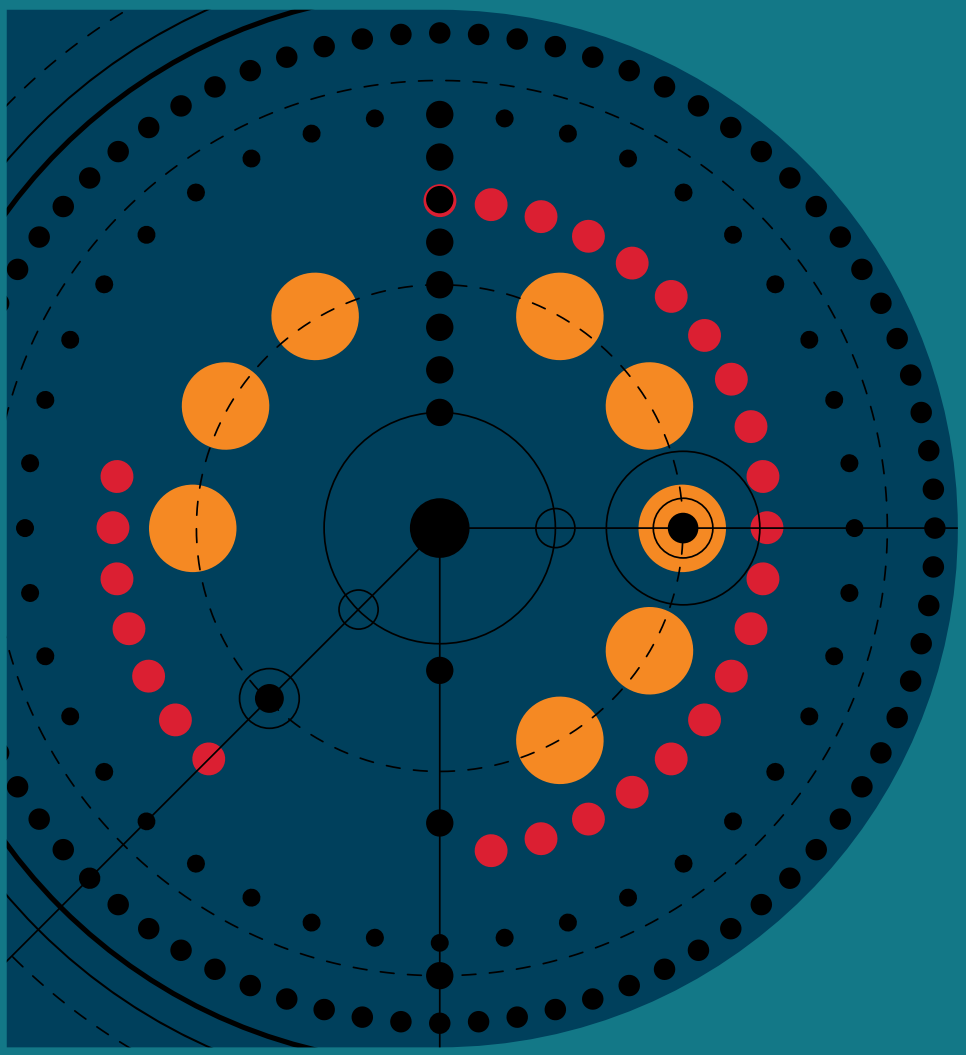
Advanced Analytics Platform

The power of enhanced analytics and machine learning



Digital
Intelligence

BAE SYSTEMS

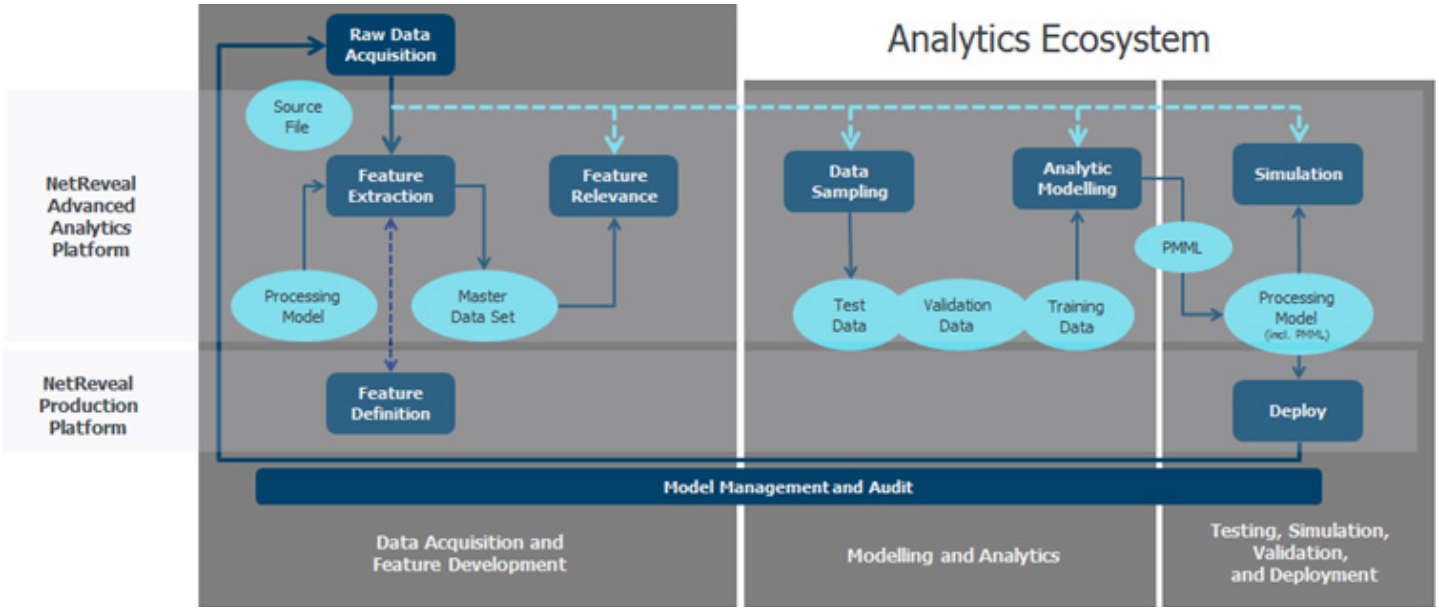
NetReveal Advanced Analytics Platform

The Advanced Analytics Platform (AAP) is NetReveal's specialist financial crime analytics and machine learning engine. It is driven by data science workflows essential to deliver improved detection performance in terms of increased effectiveness and efficiency. AAP is easy to use, does not require users to have deep data science domain skills, and requires no complex code or scripts to be written.

AAP adds enhanced analytics and machine learning across the NetReveal product suite and has two main interactions with the core platform:

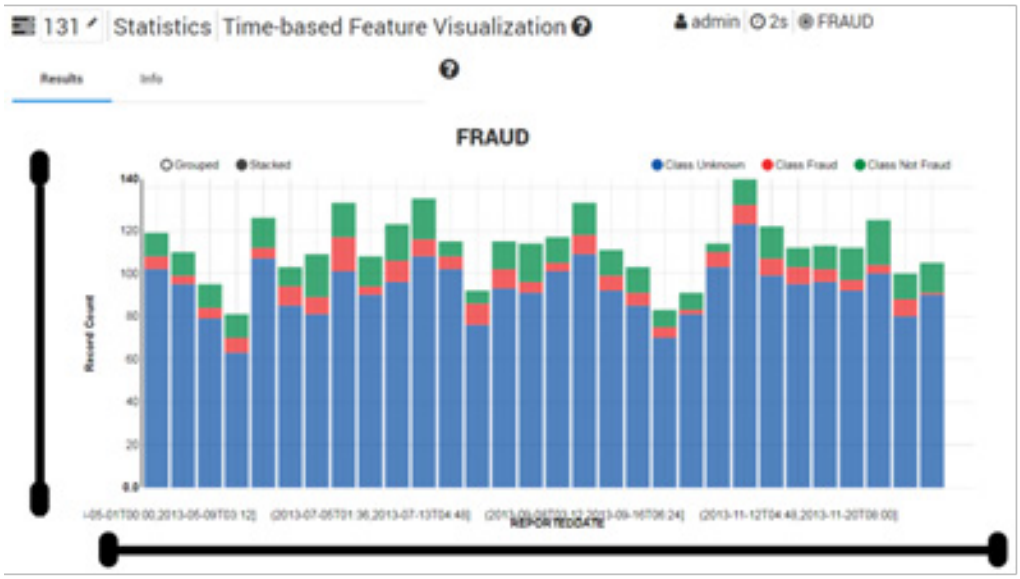
- AAP contains NetReveal's big data engine for feature development, specifically to build enhanced datasets for modelling
- AAP exports its model outputs in the form of standard Predictive Modelling Markup Language (PMML) files which plug directly into the NetReveal Scoring Manager

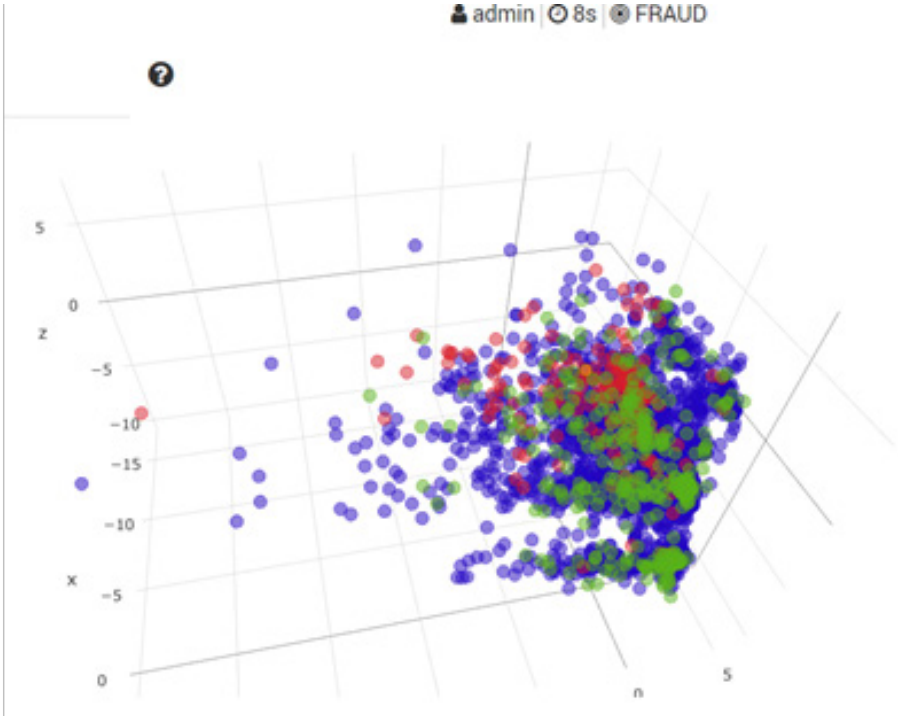
The full analytics ecosystem containing AAP and the NetReveal production platform is shown below:



Data Acquisition and Feature Development

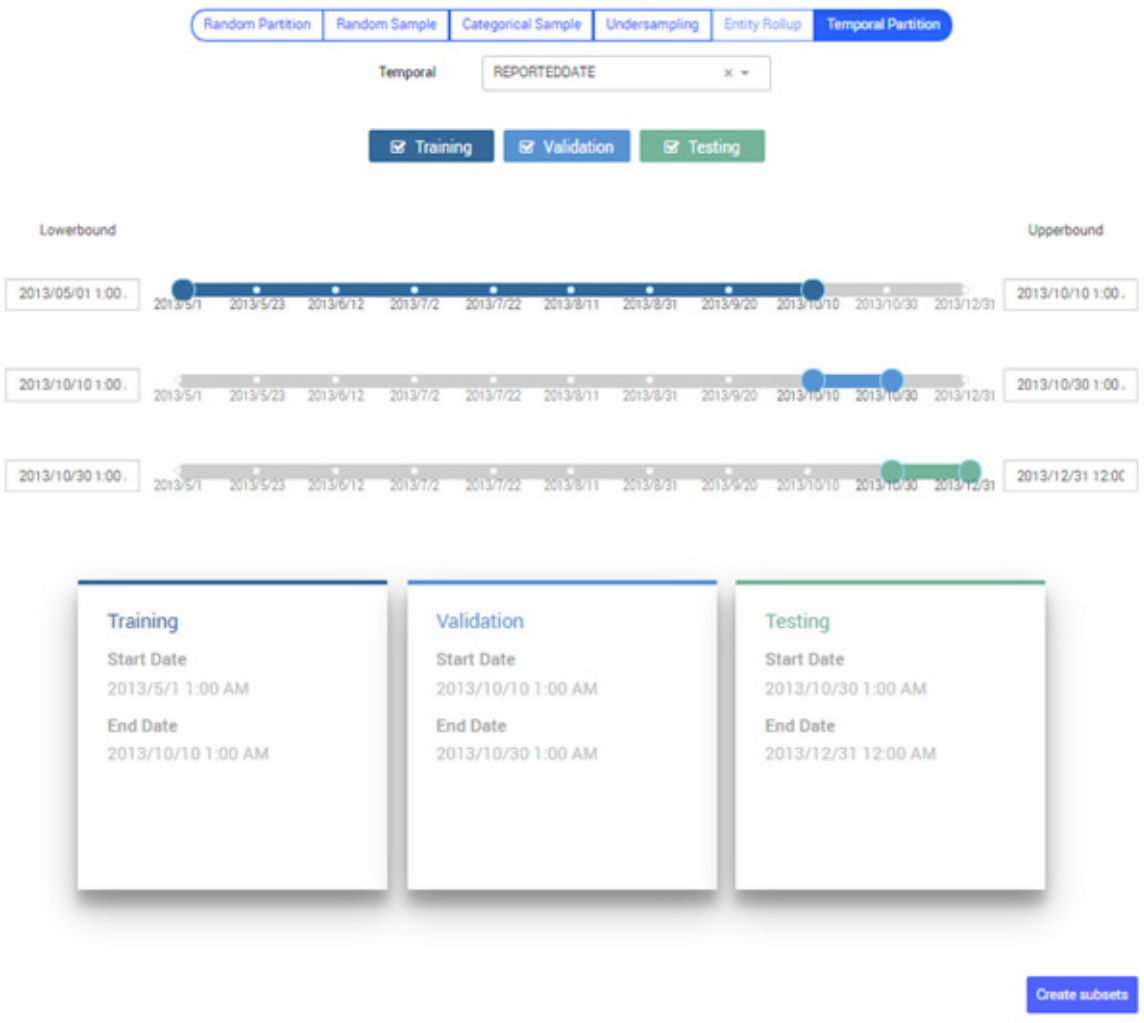
AAP ingests flat files containing customer, account, and transaction data in addition to a processing model configuration file from NetReveal that defines what fields are to be profiled and with respect to what time periods, etc. It is then very straightforward within AAP to extract behavioural profile features from the raw data. AAP contains a range of statistical and visualisation-based functions that explore these features, in particular to remove correlated or low importance features. Illustrative screenshots from this part of the ecosystem are show below:





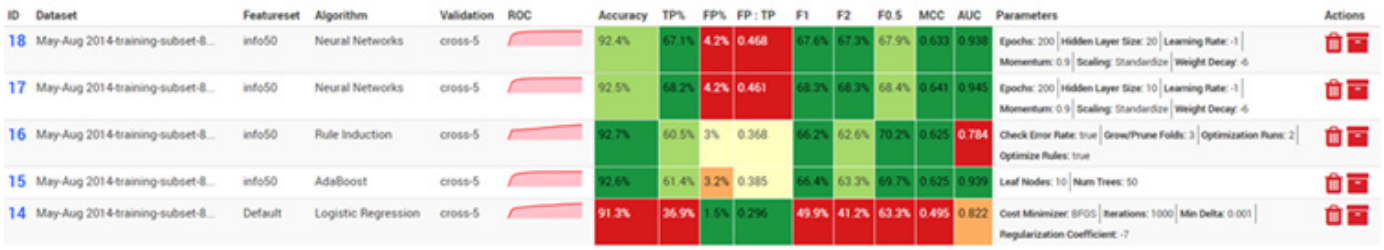
Modelling and Analytics

AAP contains a powerful set of machine learning analytics, both supervised and unsupervised models. In addition AAP simplifies the crucial steps required to partition, sample, and normalise the data before these models are applied. For example, AAP employs a slider bar control for splitting data into training, validation and test sets and a UI to filter the data based on a particular entity (e.g. account) or a particular customer attribute (e.g. segment). A screenshot is displayed below.

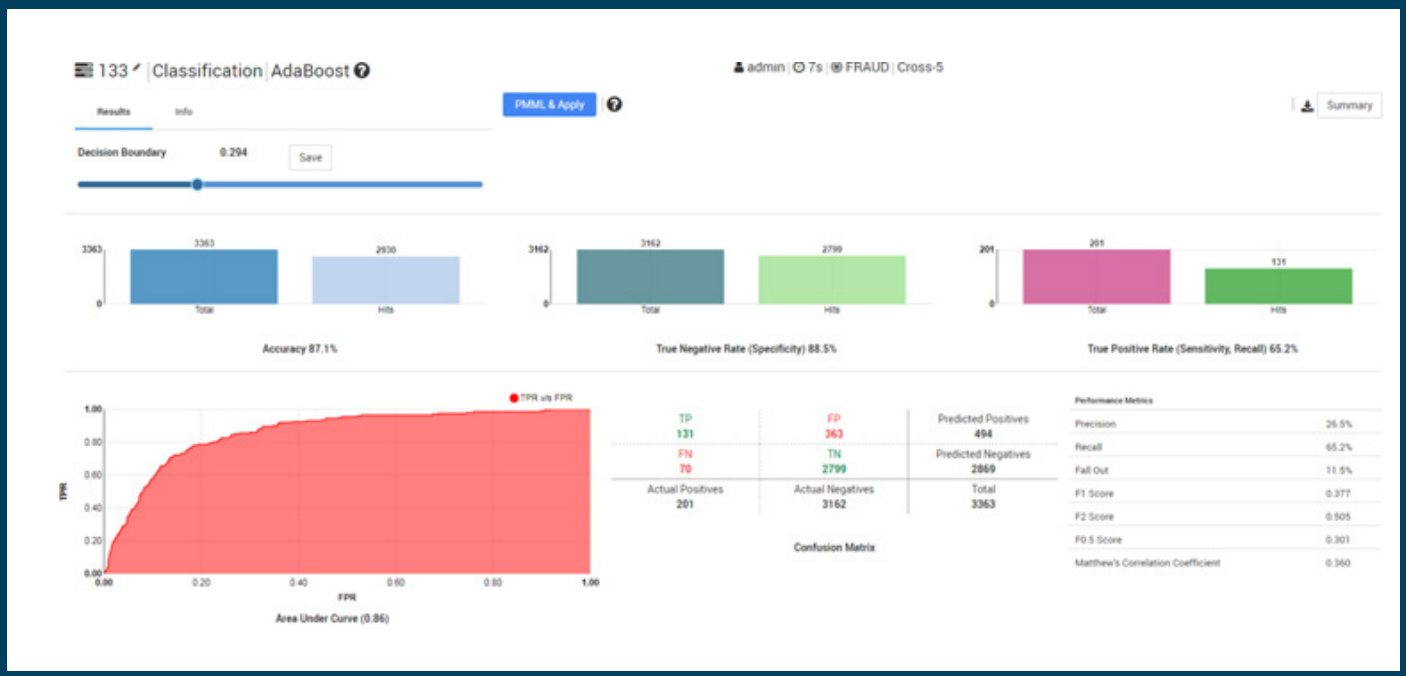


AAP's supervised machine learning analytics, for classifying data into suspicious or not-suspicious classes include: Rule Induction, Logistic Regression, Random Forest, Support Vector Machine, Adaboost, Gradient-Boosted Tree, and more. It is important to note that the non-suspicious class significantly overwhelms the suspicious class in typical datasets. This has to be carefully accounted for in any detection analytics and in AAP this is achieved by state-of-art under-sampling and over-sampling techniques. Crucially also, AAP relieves burden on the user associated with tuning the so-called hyper-parameters of its supervised machine learning models in order to achieve optimal detection performance. Specifically, AAP includes the option to auto-optimise these parameters.

AAP provides intuitive dashboard views that display the performance of multiple detection models and allow the user to rank them against a range of standard detection metrics, e.g. number of false positives, number of true positives, area under ROC curve, F1 score. This is shown in the screenshot below and the user may download this output as a png or csv file for inclusion in audit reports.



AAP also includes a drill-down view into any of the models to assess their individual performance in detail. An example is shown in the screenshot below. This shows a breakdown of the model's performance in terms of accuracy against detecting known suspicious instances (true positives) in a validation dataset and detecting known non-suspicious instances (true negatives) in the validation set. In any detection problem there is a trade-off between the number of true positives and the number of false positives. This trade-off is controlled by the decision boundary on the model's predictions and AAP exposes this to the user in the form a slider bar. AAP automatically optimises the decision boundary against a standard detection metric, but the user may override this value.



When AAP has built (or trained) a number of models they can be tested on fresh (or test) data. It is important to note that AAP generates white-box and auditable explanations for each suspicious event its machine learning models predict in the test dataset. This is highlighted in the screenshot below. Shown is a sample of predicted suspicious events and the features / feature values that the model has determined to be the most discriminating ones together with their importance weights.

Top Predictions

RowID	Key	Target Class	Probability	Reasons
146	TXN-387-20090120	S_AML_AL_002	0.89	+
177	TXN-237-20090121	S_AML_AL_002	0.87	+
677	TXN151599	S_AML_AL_002	0.87	+
254	TXN-208-20090123	S_AML_AL_002	0.87	+
592	TXN151416	S_AML_AL_002	0.86	+

Reasons

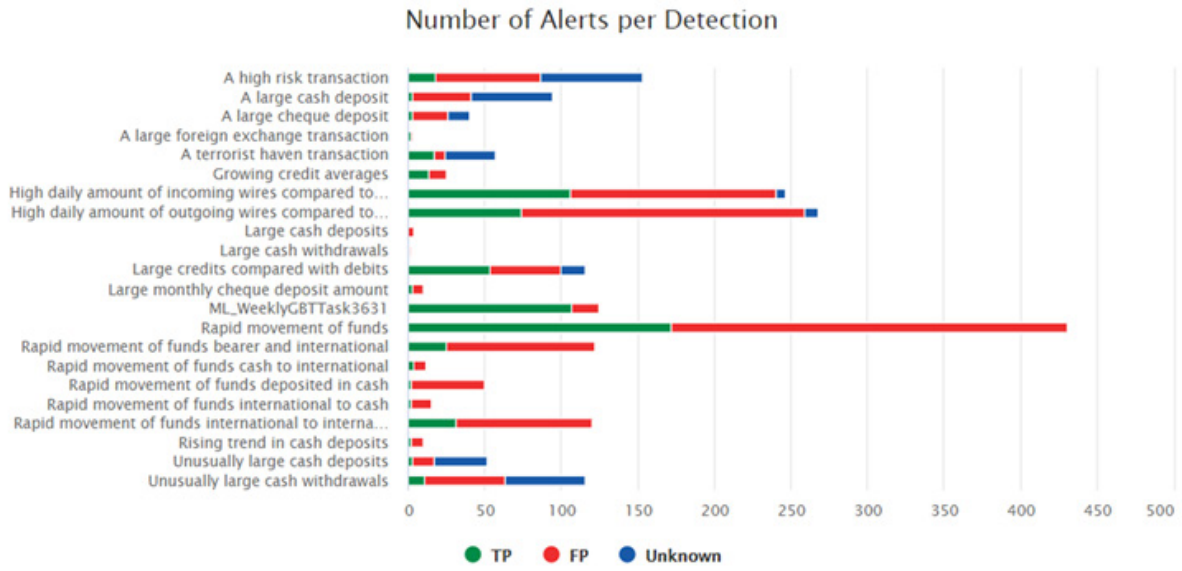
-	Feature	Value	Weight
	ExternalCreditMaximum.max.month.present	3048041.26	17.49
	ExternalCreditAmount.acc.week.present	3048041.26	17.43
	CashCreditAmount14.acc.day.total	3048041.26	16.27

AAP's unsupervised machine learning analytics, for segmenting data into groups of clusters, e.g. customers who exhibit similar transactional patterns of behaviour, include: K-Means, DBSCAN, Gaussian Mixture Model, Self-Organising Maps, and more. An example of a set of behavioural clusters that AAP discovered in a dataset is shown below. In AAP it is possible to not only determine the clusters (in this case using the K-Means analytic) but also to visualise them via a 2-D embedding of the data (here using the t-SNE analytic).

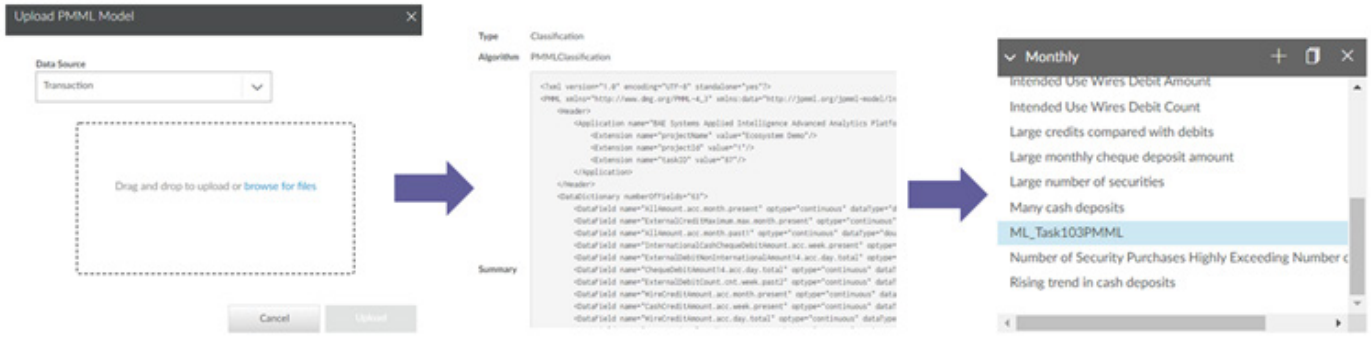
AAP's unsupervised machine learning analytics, for detecting outlier events in data, include: Isolation Forest, K-Nearest Neighbours, Local Outlier Factor, and Local Outlier Probabilities. These analytics discover statistical anomalies in data and provide details with regard the underlying features / feature values that gave rise to each anomaly. The use case for these analytics is the so called "unknown unknowns" – patterns of financial crime that have not been seen before or even anticipated and therefore not represented by any of the existing detection scenarios.

Testing, Simulation, Validation and Deployment

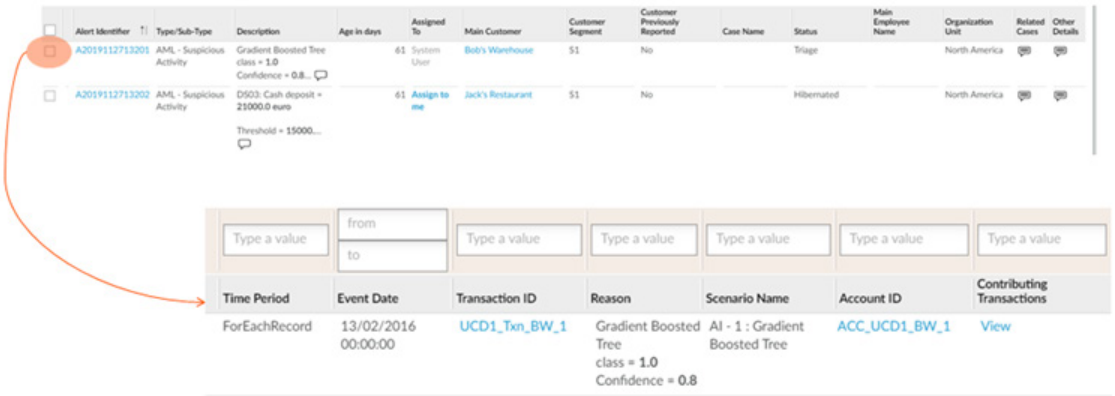
AAP enables one-click simulation-based testing of any of its trained predictive models alongside the detection scenarios contained within NetReveal. This results in a further dashboard view comparing their respective performances with respect to true and false positives.



As indicated earlier, AAP's machine learning models are exported in the form of PMML files. These can be simply dragged-and-dropped into NetReveal's user interface for deployment in the detection engine. This process is illustrated in the screenshot below:



An example of a user alert in NetReveal that has been scored by an AAP machine learning model (Gradient Boosted Tree) is shown in the screenshot below:



We are Digital Intelligence

BAE Systems Digital Intelligence is home to 4,800 digital, cyber and intelligence experts. We work collaboratively across 16 countries to collect, connect and understand complex data, so that governments, nation states, armed forces and commercial businesses can unlock digital advantage in the most demanding environments. Launched in 2022, Digital Intelligence is part of BAE Systems, and has a rich heritage in helping to defend nations and businesses around the world from advanced threats.

BAE Systems
Surrey Research Park
Guildford
Surrey GU2 7RQ
United Kingdom
T: +44 (0) 1483 816000

BAE Systems
8000 Towers Crescent Drive
13th Floor
Vienna, VA 22182
USA
T: +1 720 696 9830

BAE Systems
Level 1
14 Childers St
Canberra, ACT 2601
Australia
T: +61 1300 027 001

BAE Systems
Suite 905 Arjaan Office Tower,
Dubai Media City
Dubai
T: +971 (0) 4556 4700

BAE Systems
1 Raffles Place #42-01, Tower 1
Singapore 048616
Singapore
T: +65 6499 5000

BAE Systems, Surrey Research Park, Guildford, Surrey, GU2 7RQ, UK

E: learn@baesystems.com

W: baesystems.com/digital

linkedin.com/company/baesystemsdigital

twitter.com/BAES_digital

BAE SYSTEMS