ALTAIR

GUIDE TO USING DATA ANALYTICS TO PREVENT FINANCIAL FRAUD

INTRODUCTION

Financial fraud takes countless forms and involves many different aspects of business including; insurance and government benefit claims, retail returns, credit card purchases, under and misreporting of tax information, and mortgage and consumer loan applications.

Combating fraud requires technologies and business processes that are flexible in their construct, can be understood by all who are involved in fraud prevention, and are agile enough to adapt to new attacks without needing to be rebuilt from scratch.

Data from all business units and functional silos must be included to create a holistic view of the success of fraud operations. Insight found in predictive machine learning (ML) models combined with the expertise of fraud analysts can foster an evidenced-based method to quickly respond to new attempts to commit financial crimes.

Tools that can rapidly segment data, remove biases, and generate and import business rules, are key components to this approach.

Armed with advanced data analytics, firms and government agencies can identify the subtle sequences and associations in massive amounts of data to identify trends, patterns, anomalies, and exceptions within financial transaction data. Specialists can use this insight to concentrate their attention on the cases that are most likely fraud.

47% of companies experienced fraud in the past 24 months, with an average of six per company. The most common types were customer fraud, cybercrime, and asset misappropriation.

According to PwC's 2020 Global and Economic Crime and Fraud Survey

Tools to Combat Financial Fraud

In this guide, you will learn how machine learning (ML) and data visualization can help you reduce losses by making better informed, risk-based decisions quickly.

About 13% of the companies who had experienced fraud reported losses of more than $50 million.

According to PwC's 2020 Global and Economic Crime and Fraud Survey

FRAUD DETECTION

Raw data arriving in PDF or text-based reports from clients and third-party systems can create confusion due to double-payments, cash or billing schemes, or other types of corporate fraud.

Fraud Detection Using Data Transformation

Altair's data preparation solutions can automate the extraction and transformation of data from many data formats and apply advanced fraud detection techniques such as Benford's law or Gestalt tests. Organizations can reuse data models built around these detection techniques, which saves valuable time when questions about new types of frauds arise fraud.

Altair's solutions can automate daily, weekly, and/or monthly data runs through a series of data transformation models designed to identify potential instances of fraud including (but not limited to):

- Benford's law (1 digit, 2 digits, summation)
- Relative size factor test
- Gestalt element link test
- Even amounts test
- Same, same, same test
- Same, same, different test

Fraud is an unnecessary cost because much of it can be pre-empted.

Benford's law, also called the Newcomb-Benford law, the law of anomalous numbers, or the first-digit law, is an observation about the frequency distribution of leading digits in many real-life sets of numerical data. The law states that in many naturally occurring collections of numbers, the leading significant digit is likely to be small. For example, in sets that obey the law, the number 1 appears as the leading significant digit about 30% of the time, while 9 appears as the leading significant digit less than 5% of the time. If the digits were distributed uniformly, they would each occur about 11.1% of the time. Benford's law also makes predictions about the distribution of second digits, , digit combinations, and so on.

A relative size factor test identifies anomalies where the largest amount for subsets in a given key is outside the norm for those subsets. This test compares the top two amounts for each subset and calculates the relative size factor for each. For example, to identify potential fraud in payments data, the test compares the ratio between the largest and the second-largest amounts in sets of data grouped by vendor.

Gestalt element link tests help detect relationships or links within a data file that serve as indicators of potential fraud. The tests establish possible links between two selected fields over the entire data set. One use case for this type of test is bribery detection. Since most organizations do not have access to the records of payer firms, they can use this type of test to analyze their own data to find patterns that indicate improper relationships.

Even amounts tests help flag potentially fraudulent payment activities. People who commit these frauds tend to use even amounts, which means there are no pennies in the invoices they have created. They also create invoices that are just below the manager's approval limit.

Same, same, same tests help reduce the number of false positives. This test identifies cases in which the same person pays the same supplier on the same day the same amount as well as variations on this theme.

Same, same, different tests look for cases in which a different person pays the same supplier on the same day the same amount.

Using these techniques, a credit union found that seven out of ten flagged accounts were fraudulent. They halted lending activity for those accounts, which saved $1.4 million in the first month after deploying Altair® Monarch® data preparation software.

Learn more about financial services applications for the Altair RapidMiner platform.

WHAT ARE BUSINESS RULES AND WHY ARE THEY IMPORTANT IN FRAUD DETECTION?

Building Business Rules with Analytics Systems

Using analytics tools to generate business rules offers three key benefits:

- **Flexible** – Platforms such as a payment processor with a fraud detection solution built into it often disallows custom businesses rules or analytic models from being added on top of these platforms. With the right analytic tools, businesses can enhance the capabilities of their existing fraud detection systems by building their own business rules.
- **Understandable** – Creating rules and building models are iterative in nature. You need to know what is influencing the decision, and this understanding is more apparent in a business rule. This recognition available in analytics tools can identify biases that might drift into models, allowing you to easily exclude certain demographics or other variables from a rule that can be pushed back into models.
- **Agile** – If hit by a massive fraud attack, businesses do not have the time to build, test, train, and deploy new models that have a specific target activity in mind based on previous experience of something that has happened multiple times. Business rules can quickly respond to unusual behavior. A sudden deviation in normal patterns can activate a rule that will stop something immediately. With analytics tools, this can be done without having historic data on target data.

Visual Fraud Detection using Altair Decision Trees

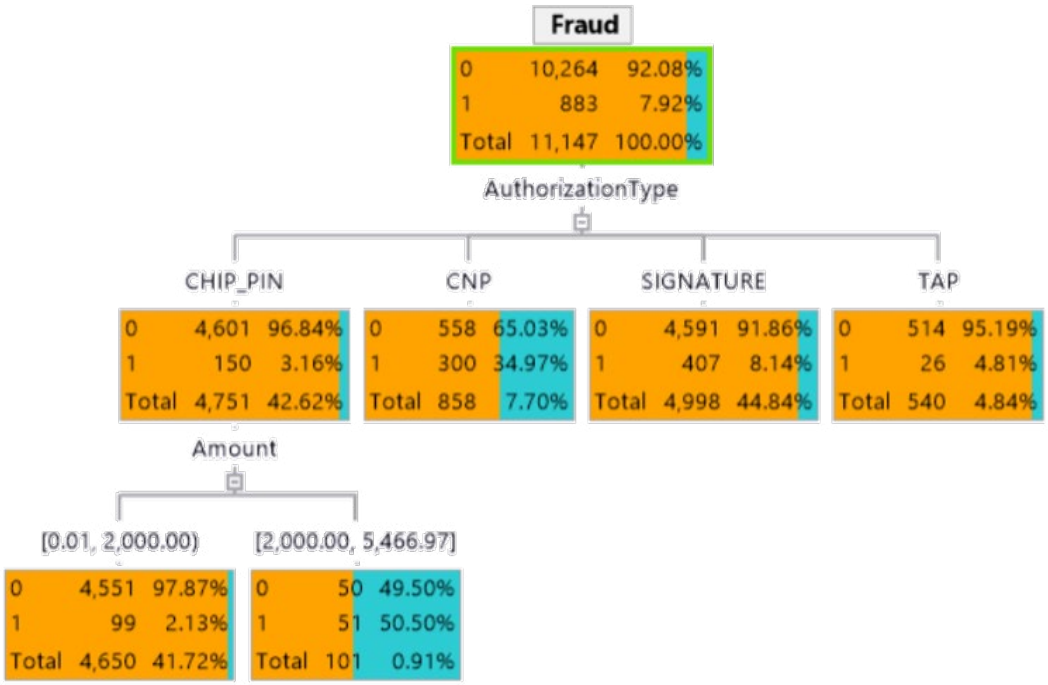
Altair RapidMiner's ML and artificial intelligence (AI) solution is well known for its ability to visualize data using intuitive workflows. Its wizard-driven interface helps users build predictive and prescriptive analytic models quickly without writing any code. Its industry-leading, patented decision tree technology enables business users without special training to build complex predictive models easily and quickly, the results of which can be converted into sets of business rules. The flexibility of Altair decision trees makes it easy to segment, profile and explore groups of customers, transaction types, transaction amounts, third party data, and other data sources to identify fraudulent activities. These capabilities are within reach for users who understand the business, rather than something only accessible to people with years of data science experience.

Defining Business Rules

A business rule engine allows financial service organizations to define, test, and execute rules that are meant to prevent fraudulent acts based on previously identified activities. A business rule engine may interact with point of sales solutions and transaction applications, monitoring credit or debit purchases, applications for lines of credit, and other similar activities. Business rules are continuously updated as they adapt to changes in both fraudulent and non-fraudulent behavior.

Each node in an Altair decision tree is essentially an IF/AND/THEN statement. In predictive fraud detection modeling, Altair allows users to convert the entire tree or parts of the tree into a language supported by the business rule application. For example, a business rule might include this logic:

IF transaction = $2,000 and IF the credit card is used and IF the credit card was used at location in another state THEN block the transaction.



In this hypothetical example, fraud has been detected in 7.92% of the records analyzed, and the instances of fraud are all related to debit card transactions.

The bank's fraud team built a decision tree with this rule: "If Authorization Type is equal to CHIP_PIN and Amount is greater than or equal to 2000 then there is a 49.505% chance that Fraud will be 0 and a 50. 495% that Fraud will be 1."

In plain English, there is approximately a 50% chance that card taps for amounts greater than $200 will be fraudulent.

A decision tree may have hundreds of nodes. Altair makes it very easy to convert the tree into a series of rules that can then be deployed into a rules engine.

English Rule Code Output

English Language Rule # 1: There is a 92.0786 percent chance that Fraud will be 0 and a 7.92141 percent chance that Fraud will be 1.

English Language Rule # 2: If Region is equal to CANADA or USA then There is a 92.9948 percent chance that Fraud will be 0 and a 7.00516 percent chance that Fraud will be 1.

English Language Rule # 3: If Region is equal to CHINA or UK then There is a 7.90514 percent chance that Fraud will be 0 and a 92.0949 percent chance that Fraud will be 1.

English Language Rule # 4: If Region is equal to MEXICO then There is a 96.2018 percent chance that Fraud will be 0 and a 3.79819 percent chance that Fraud will be 1.

English Language Rule # 5: If Region is equal to CANADA or USA and Amount is greater than or equal to 0.01 and is less than 319.88 then There is a 97.6132 percent chance that Fraud will be 0 and a 2.38676 percent chance that Fraud will be 1.

English Language Rule # 6: If Region is equal to CANADA or USA and Amount is greater than or equal to 319.88 and is less than 756.97 then There is a 95.2446 percent chance that Fraud will be 0 and a 4.75537 percent chance that Fraud will be 1.

English Language Rule # 7: If Region is equal to CANADA or USA and Amount is greater than or equal to 756.97 and is less than 1141.96 then There is a 90.6757 percent chance that Fraud will be 0 and a 9.32432 percent chance that Fraud will be 1.

English Language Rule # 8: If Region is equal to CANADA or USA and Amount is greater than or equal to 1141.96 and is less than or equal to 5466.97 then There is a 66.0904 percent chance that Fraud will be 0 and a 33.9096 percent chance that Fraud will be 1.

English Language Rule # 9: If Region is equal to CANADA or USA and Amount is greater than or equal to 0.01 and is less than 319.88 and AuthorizationType is equal to CHIP_PIN then There is a 99.7234 percent chance that Fraud will be 0 and a 0.276625 percent chance that Fraud will be 1.

English Language Rule # 10: If Region is equal to CANADA or USA and Amount is greater than or equal to 0.01 and is less than 319.88 and AuthorizationType is equal to CNP then There is a 90.6404 percent chance that Fraud will be 0 and a 9.35961 percent chance that Fraud will be 1.

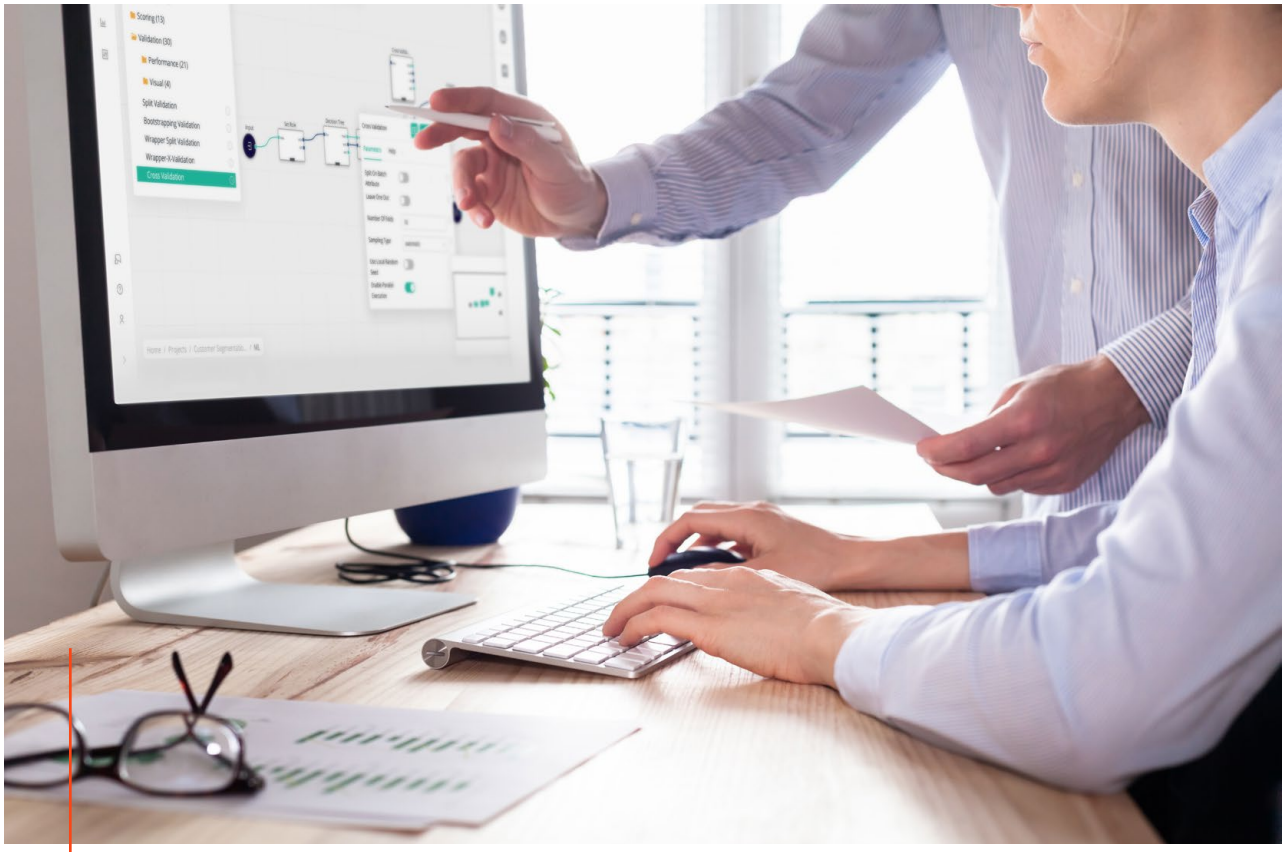
English Language Rule # 11: If Region is equal to CANADA or USA and Amount is greater than or equal to 0.01 and is less than 319.88 and AuthorizationType is equal to SIGNATURE or TAP then There is a 97.4932 percent chance that Fraud will be 0 and a 2.50681 percent chance that Fraud will be 1.

English Language Rule # 12: If Region is equal to CANADA or USA and Amount is greater than or equal to 0.01 and is less than 319.88 and AuthorizationType is equal to CHIP_PIN and MerchantType is equal to AIRLINE, CHARITY, CONTRACTOR, FOOD/LIQUOR, FUEL/GAS, MEDICAL, NONPRESENT, PERSONAL_SERVICES, RESTAURANT, TELEMARKET, TELEPHONE or TRANSPORTATION then There is a 100 percent chance that Fraud will be 0 and a 0 percent chance that Fraud will be 1.

English Language Rule # 13: If Region is equal to CANADA or USA and Amount is greater than or equal to 0.01 and is less than 319.88 and AuthorizationType is equal to CHIP_PIN and MerchantType is equal to BUSINESS_SERVICES, HOTEL, RETAIL or VEHICLE_RENTAL then There is a 98.8981 percent chance that Fraud will be 0 and a 1.10193 percent chance that Fraud will be 1.

English Language Rule # 14: If Region is equal to CANADA or USA and Amount is greater than or equal to 0.01 and is less than 319.88 and AuthorizationType is equal to CNP and MerchantType is equal to AIRLINE, CHARITY, FUEL/GAS, MEDICAL, PERSONAL_SERVICES, RESTAURANT, TELEPHONE or TRANSPORTATION then there is a 86.4486 percent chance that Fraud will be 0 and a 13.5514 percent chance that Fraud will be 1.

English Language Rule # 15: If Region is equal to CANADA or USA and Amount is greater than or equal to 0.01 and is less than 319.88 and AuthorizationType is equal to CNP and MerchantType is equal to BUSINESS_SERVICES, CONTRACTOR, FOOD/LIQUOR, HOTEL, NONPRESENT, RETAIL, TELEMARKET or VEHICLE_RENTAL then there is a 95.3125 percent chance that Fraud will be 0 and a 4.6875 percent chance that Fraud will be 1.



Organizations use Altair to automate the extraction and transformation of data and apply advanced fraud detection techniques to the resulting information. They can then reuse data models built around these detection techniques, which saves valuable time when questions about new types of fraud arise.

APPLYING DEEP LEARNING TO FRAUD DETECTION

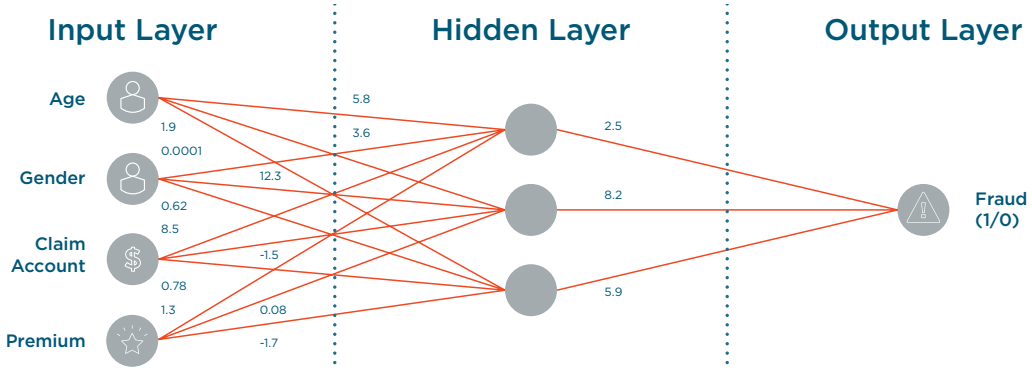
The deep learning approach (also referred to as neural networks) is another powerful method for detecting fraud in large volumes of transactions. Deep learning tools are sometimes used with decision trees to improve their accuracy, but they can perform well on their own in many circumstances.

Adapting Automatically to New Fraudulent Behaviors

Deep learning algorithms offers an important benefit over decision trees: They can process vast amounts of data and create new business rules automatically if needed. Decision trees, while very powerful and highly customizable, require a fair amount of manual work to build, deploy, and test. In contrast, deep learning algorithms are able to learn from changing data trends and adapt. When it comes to fraud applications, this means they can often begin predicting fraud faster than any other approach.

A common disadvantage of deep learning is that complex models produce results that are not visually interpretable or transparent in how their predictions were derived. These types of models are described as 'black box' models; they simply accept data inputs and produce predictive results. They can process massive amounts of data and adapt as trends within the data change, but their visual output is not explainable or interpretable.

Altair RapidMiner is unique in that it leverages the power of deep learning but exposes how these models were configured, and maps how the data in the model was used to make a predictive outcome. Altair is fully committed to the principles of responsible and transparent AI and gives analytic teams and financial analysts the confidence that their predictive modeling processes can be easily understood, interpreted, and used in decision making.



Simple Feed-Forward network with three layers: input, hidden and output.

Altair RapidMiner supports up to ten layers of predictive models. The platform's deep learning algorithms can model complex relationships between inputs and outputs efficiently and find patterns in large amounts of data.

STREAMING ANALYTICS APPLIED TO HIGH FREQUENCY TRADING FRAUD

Today's complex electronic markets allow for new methods of market manipulation. Because of the fast pace of innovation and sometimes overwhelming complexity behind trading activity, it is essential for firms engaged in trading to have real-time visibility into trading activity. The volume and velocities for trading data are enormous. For example, the U.S. markets alone make on the order of 9.5 billion trades every day, with each trade comprised of a large number of "ticks" like orders, cancellations, executions, and requests for quotes (RFQs). Most firms engaged in electronic trading handle between 2.5 billion to 6 billion ticks per day in equities data alone.

All this activity leaves plenty of room for fraudulent activity to take place unnoticed — unless compliance officers have the right tools available.

Types of Trading Fraud

There are numerous ways that traders in high frequency trading (HFT) environments can manipulate market and trading activity in fraudulent ways, including:

Spoofing: In a spoofing fraud, a trader places limit orders and then removes them before they are executed. By spoofing orders, the fraudsters hope to distort other traders' perceptions of market demand and supply. For example, a large bid order might be placed with the intention of being canceled before it is executed; the spoofer seeks to trade profitably on the prices that rise as a result of the spoofed order.

Layering: Layering is similar to spoofing except that the fraudster enters multiple orders with no intention of executing them, but at many different prices. This causes the midpoint of the price spread to move, and the same trader can then execute a profitable real trade based on the manipulated price. Both spoofing and layers are illegal in the U.S., but nonetheless, compliance officers must still watch out for these activities.

Quote Stuffing: This is a relatively new technique and involves overwhelming the market with huge numbers of simultaneous quotes (on the order of tens of thousands of orders per second) in order to confuse rival traders into thinking there is a large amount of activity on a particular stock.

Wash Trading: This illegal activity involves giving the appearance that authentic purchases and sales are being made. Here, the fraudster buys and sells the same securities simultaneously, which can increase the trading volume and make the security appear as though a number of buyers are active in the market.

Momentum Ignition: With this tactic, traders cause sharp price movements with a series of trades that look like trades from other HFT firms, with the intention of attracting more trades for the targeted security. The fraudster knows that after the artificially created

rapid price movements will end, the price will move back to its baseline. They know they can make a profit by taking a position early on and then moving out of that position before the activity dies down. As with everything else in HFT, all this can happen in a matter of a few seconds.

This is not an exhaustive list of possible trading frauds, of course, and determined criminals are always trying to develop new strategies given the velocity and volumes data involved.

Streaming Analytics Provides Real-Time Visibility

The Altair Panopticon™ streaming analytics platform enables compliance officers in investment banks to maintain a bird's eye view of all trading activity throughout the trading day by trader, instrument, asset class, office, desk, and any other dimension of interest.

Compliance officers can monitor trading activity in real-time in as much detail as needed — down to the nanosecond timestamps if necessary — to spot potentially fraudulent trading activity within their own firms, and potentially by rival firms in the market. They can also view streams of trading activity or historic trading data.

The system can handle any number of real-time streams of trading and market data, process them with extremely low latency, and present the data using a set of high density visualizations that enable users to spot outliers, anomalies, clusters, and trends in seconds.

Applications Include:

Trade Surveillance: Identify cases of potential spoofing, quote stuffing, wash trading, and other fraudulent activity. Playback through a series of trades tick-by-tick to gain a full understanding of exactly what happened.

Alert Consolidation: Intraday and historic alert reporting and trend analysis consolidating alerts across traders, regions, and asset classes.

Holistic Surveillance: Analyze alerts generated from trade and comms surveillance. Correlate the different components of fraudulent behavior and support trade reconstruction as outlined in Dodd Frank regulations. Add critical context to existing alerts and produce higher value alerts, which supports more efficient and effective investigation.

Behavioral Risk Profiling: Holistic surveillance of traders together with trading and risk positions and security alerting. Score traders based on the risk they take on and compare their current performance to past performance and their peers. Investigate trader interaction networks to speed understanding of past activity and identify behavioral abnormalities

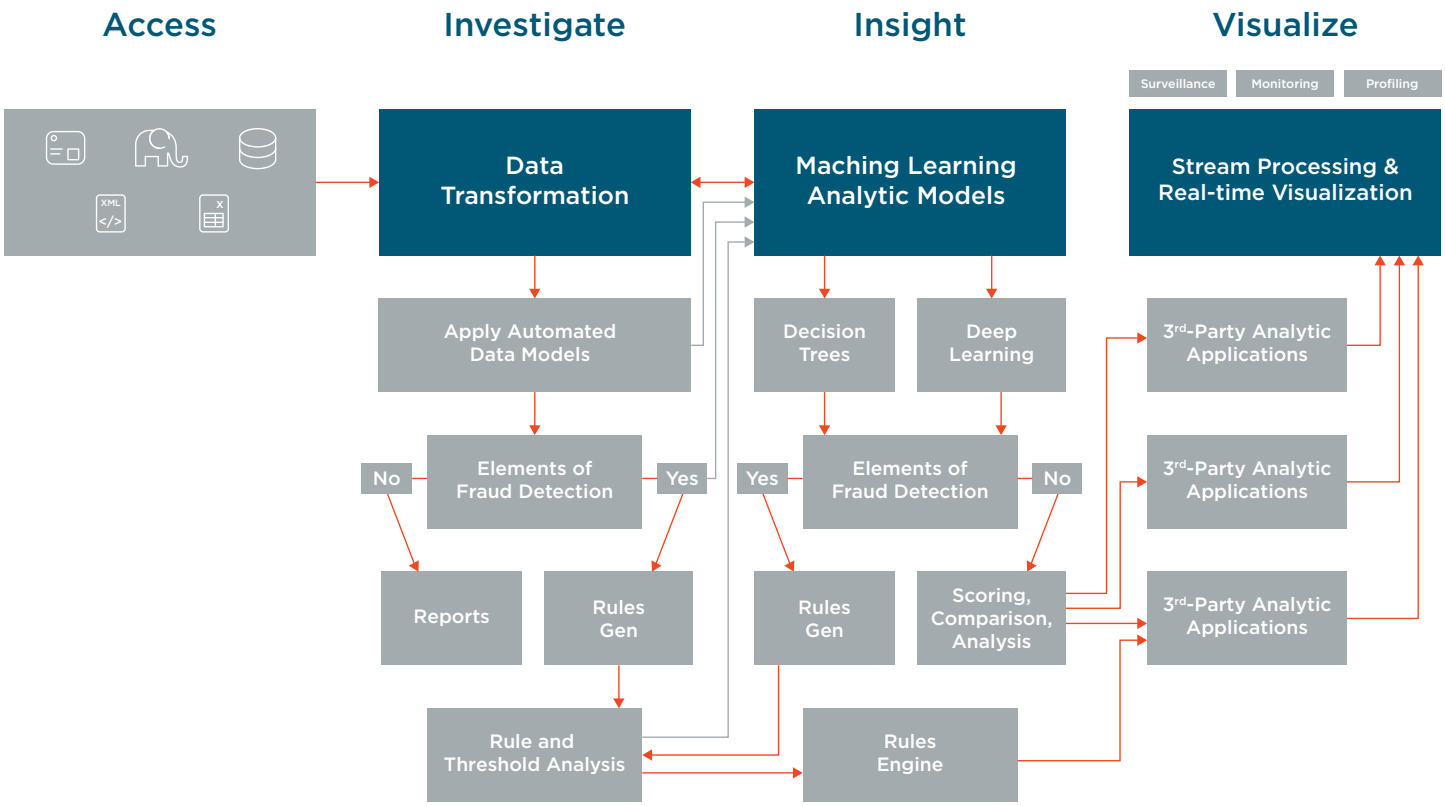
WITH ALTAIR

Altair data preparation solutions can automate the extraction and transformation of data from multiple data formats and apply advanced fraud detection techniques such as Benford's law or the Gestalt tests. You can easily generate and deploy business rules using Altair predictive analytics to trigger against probable fraudulent activities. Altair's deep learning is very efficient at modeling complex relationships between inputs, outputs, and finding fraudulent patterns in large amounts of data.

Fraud Analytics Workflow

Altair supports the complete data analytics workflow, including data preparation solutions that enable connections between disparate data sources and formats, advanced ML and AI to help make better decisions, and visual analytics tools that provide transparency in high data volume/velocity environments.

Altair's code-optional development environment enables data science teams to build models using combinations of SAS language, Python, R, and SQL code.



Where Altair Helps Financial Services

Data Extraction

- From core banking platform reports
- From credit card/consumer loan/mortgage servicing platform reports
- From vendor invoices and statements

Data Transformation

- Daily settlement for credit card, debit card, ATM
- General ledger reconciliation
- Mergers and acquisition: trial balance mapping and account conversion
- RPA data provisioning
- Loan sales enablement
- Incentive plans enablement
- FFIEC/NCUA call report automation
- CFPB audit preparation
- Systems and data migrations
- Credit card rewards analytics

Machine Learning

- Credit risk analytics
- Marketing analytics
- Membership analytics
- Fraud detection and prevention
- Behavioral analytics
- Real time risk aggregation
- Forex RFQs and hit ratios

Data Democracy

- Active data catalog
- Data curation
- Team driven collaboration
- Governance, lineage, security
- Centralized dataset and data model repository

Cultivating a Data Driven Team

Altair's 35-year history working with financial service organizations including many of the largest global banks, credit unions, and mortgage service providers means we understand how analytics can help address risk mitigation, regulatory oversights, new customer engagement channels, operational insight, and more. Our easy to use, no code data transformation, ML, and real-time data visualization and stream processing platform enables executives, financial analysts, and data scientists to collaboratively use insight using governed, trusted, and accurate data.

Learn more about how we help financial service customers operationalize their data analytics to drive efficiencies and reduce risk at altair.com/financial-services

Altair is a global leader in computational science and artificial intelligence (AI) that provides software and cloud solutions in simulation, high-performance computing (HPC), data analytics, and AI. Altair enables organizations across all industries to compete more effectively and drive smarter decisions in an increasingly connected world – all while creating a greener, more sustainable future.

For more information, visit www.altair.com