TRANSPARENT AI AND MACHINE LEARNING: ALTAIR® KNOWLEDGE STUDIO®

Knowledge Studio delivers explainable artificial intelligence (AI) and automates machine learning tasks to enable people to make fully informed decisions based on massive amounts of data. The software displays all the details of a model's configuration so it is easy to understand how it generates predictions.

Analysts who may not be familiar with modeling or AI processes can quickly uncover insights to help solve complicated problems. Data scientists can fine tune model parameters and develop highly sophisticated models using a drag-and-drop interface with no coding required.

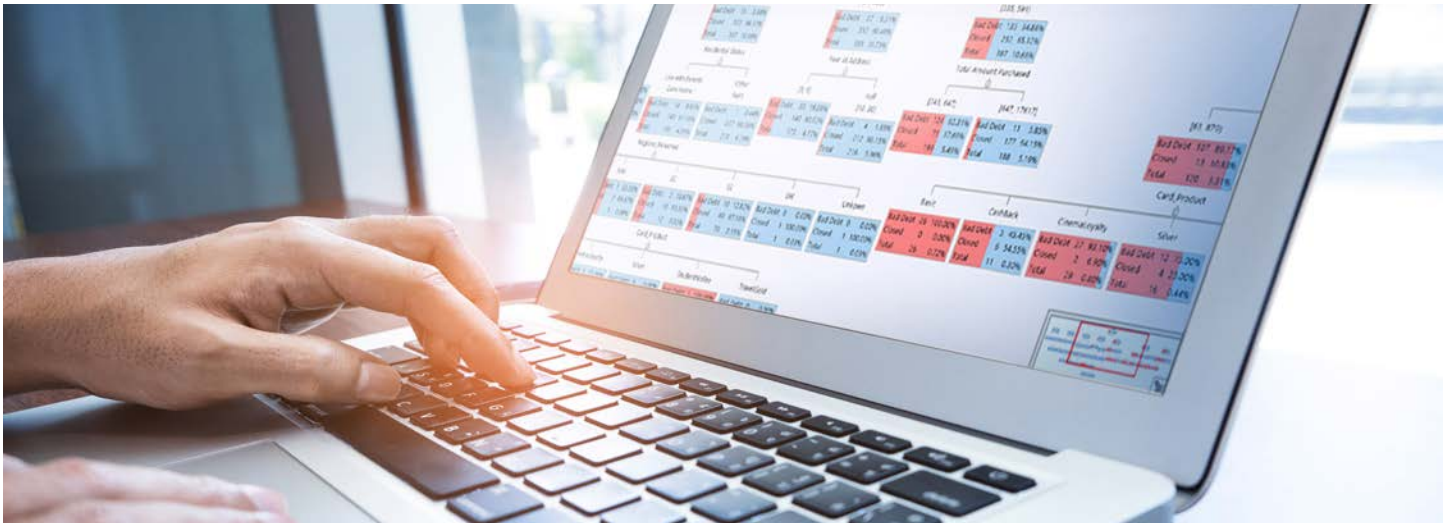
Applications Throughout the Enterprise

Altair's open, flexible predictive analytics platform is designed for data scientists and business analysts alike. Its industry-leading visual approach to analytic modeling enables data science teams to create high quality machine learning (ML) and AI models. Our collaborative approach to machine learning enables data scientists and business users to minimize repetitive tasks across the enterprise.

Knowledge Studio is a perfect fit for managing credit and fraud risk, performing marketing analytics, managing product lifecycles, and designing customer loyalty programs. From healthcare and financial services to telecommunications and manufacturing, Knowledge Studio enables analytics teams to gain useful, actionable insight from their data.

Build Complex AI Models with No Coding

Knowledge Studio's intuitive visual interface enables users to connect to data sources, transform disparate data formats into usable datasets, and build models without writing a single line of code. Teams throughout your organization can generate insights with a wide range of modeling techniques and algorithms, including Altair's patented decision trees and strategy trees, regression models, and neural networks, without extensive training or specialized knowledge. The software's interactive UI also allows data scientists to save enormous amounts of time developing complex models so they can begin delivering useful predictions quickly. Your teams can complete complex projects in minutes or hours, not weeks or months.



Generate Confidence with Explainable AI

The ability to understand how models generate their output is critical to building confidence in the reliability and value your AI investments. With Knowledge Studio, it's easy for model designers, analysts, and decision-makers to understand exactly how models work. The software's highly visual approach to model construction maintains high levels of machine learning performance and prediction accuracy while allowing people to fully understand, trust, and manage their AI models.

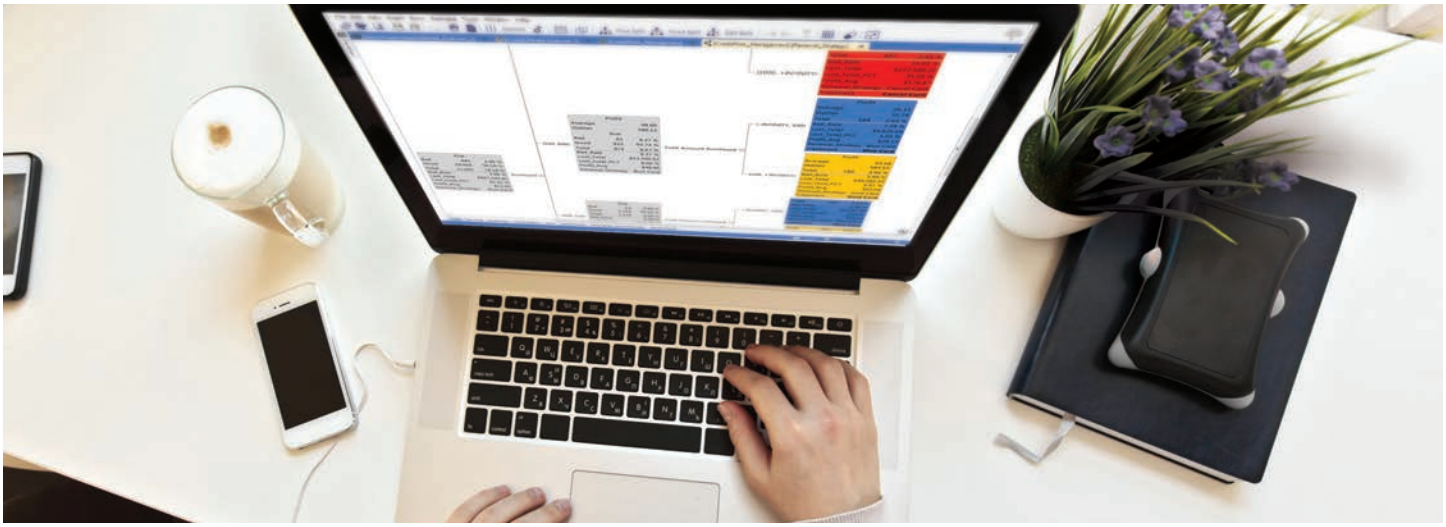
Knowledge Studio's ease of use and ability to automated critical steps in the modeling process enable data scientists to develop AI and ML models faster than coding them from libraries or using other tools.

Knowledge Studio supports a transparent model-building process. Users, including people who are not data scientists, can see how models gather and process data to generate predictions. Everyone involved in using and acting on the predictions generated by the models can feel confident in the quality and validity of those results.

Save Time with Automated Machine Learning (AutoML)

Anyone who has developed AI models from scratch can appreciate how time-consuming it can be to build and test accurate models using traditional coding tools. Fast-moving business environments, however, must accelerate the development, testing, and deployment of accurate models. More efficient methods enable businesses to begin accruing the benefits of AI and ML in less time. AutoML also allows analysts and business users — people without specialized data science training — to build effective models. This is increasingly important given the tight market for AI experts.

Knowledge Studio's pre-configured design exposes users to best practices and takes the guesswork out of modeling. Users can select models, tweak configurations, tune performance, and generate results for multiple predictive models in minutes within in a single node. They can create predictive models and generate deployment code within a single workflow using more than 80 preconfigured wizard-based nodes and share insights with business intelligence (BI) applications or export models to decision engines, databases, or real-time scoring systems.



Knowledge Studio enables data scientists and business analysts to assign actions to ML models and make quick comparisons between numerous strategic plans based on their predicted results.

Select Source Data and Build Models in Minutes

Knowledge Studio allows users to profile and segment data before building any models. Its visual interface simplifies the data evaluation process to save time during model development. They can then select from a large variety of algorithms including regression models, cluster analysis, bagging and boosting, neural networks, scorecards, and other model types. Designers can evaluate the suitability of models to their tasks and their data sources with built-in lift charts, receiver operating characteristic (ROC) curves, Kolmogorov Smirnov (KS) charts, profit charts, confusion matrices, and more. Users can also advance from predictive to prescriptive analytics by optimizing the predictive output of models under realistic business constraints.

Import and Export Python and R Code

Knowledge Studio generates model code in Python, R, SAS, SQL, PMML, and other programming languages with the simple click of a mouse. You can also import and execute R and Python code within Knowledge Studio workflows. This capability allows you to accelerate model building and deployment by leveraging the user-curated custom library of previously written code.

Knowledge Studio supports R version 3.0.3 and above and Python version 3.5 and above.
This allows users to:

- Import R and Python datasets, write R and Python programs, and explore R and Python datasets with Altair data visualizations.
- Develop decision trees, strategy trees, and tree ensemble models using R and Python datasets.
- Compare R and Python models with other models using the system's model analyzer node.
- Generate R, Python, SQL, and PMML code from Knowledge Studio workflows.
- Add R and Python packages to manage data manipulation and data mining and apply statistical algorithms.
- Create linear and nonlinear models, classical statistical tests, time-series analysis, classification, clustering, and more.

Patented Decision Trees Accelerate Model Design

Knowledge Studio's decision trees enable model builders to segment, profile, identify, and rank the relationships between the variables used in a model against the business objective or hypothesis. They make no assumptions about the source data and enable model designers to explore unfamiliar datasets and identify variables with high predictive value in minutes.

Knowledge Studio's decision trees enable users to:

- Leverage a powerful set of statistical algorithms to devise complex AI models without writing code.
- Pinpoint the most significant variables in source data and understand which rules to apply to those variables.
- Pan, zoom, collapse, and expand views of data within a fully visual interface.

- Compare tree structures in side-by-side views.
- Incorporate business knowledge and policies into models.
- Fine tune parameters and attributes as needed with an intuitive visual interface.
- Show relationships between variables or manually determine where the model should display a variable relationship.
- Understand the best indicators of predictive behavior quickly.

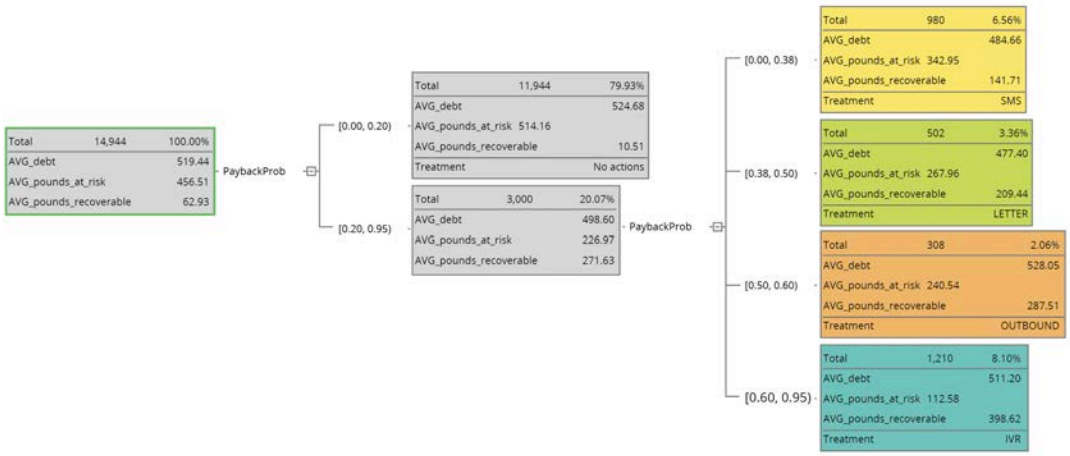
Knowledge Studio's automatic code generation tools can export decision trees as SQL, SAS, SPSS, Java, R, or Python code to deploy them in other analytical environments.

Common applications for decision trees include evaluation of marketing campaigns, financial risk measurement, fraud detection, and detection of cross-selling/up-selling strategies involving multiple offerings and distribution channels. For example, a decision tree-based model can predict the likelihood of a customer's willingness to purchase a new product, cancel a subscription, or default on a loan.

Strategy Trees

Knowledge Studio strategy trees are built upon Altair's patented decision tree technology and enable analysts to assign actions and treatments to predictive models, compare strategic plans, and evaluate approaches to resolving business challenges.

Data science teams can use strategy trees to combine multiple models, such as propensity and customer lifetime value, to prioritize customer segments and actions. Strategy trees allow users to analyze an unlimited number of business rules within a single graphical view and eliminate the need to create multiple models. They can gain insight from hundreds or even thousands of dimensions that are normally invisible, apply business knowledge and policies to data profiles, and understand the impact treatments have on different subsets of the data.
As with decision trees, users can export strategy trees to SQL, SAS, SPSS, Java, R, or Python code to utilize them in other analytical environments.



Knowledge Studio strategy trees can analyze an unlimited number of business rules within a single graphical tree view, eliminating the need to create multiple models.

Strategy trees support:

- Automatic treatment assignment through optimization and rule-based decisioning.
- Visualization of data subsets and actions to facilitate better understanding of models and prescriptive strategies.
- Single, holistic views across data subsets and treatments for faster time to insight.
- Automated calculations across an unlimited number of target variables for more accurate analyses.
- Exploration of "what if" scenarios prior to deployment.
- Testing and validation of strategic models within a visual environment.
- Automatic code generation.

Businesses can use strategy trees to prevent lost revenue opportunities by predicting a customer's propensity to respond to different sales and marketing campaigns, predict the likelihood of credit defaults, identify the customers most likely to commit fraud, and improve employee retention rates detecting factors that may cause people to leave the firm.

Generalized Linear Model (GLM) Nodes

GLM-based ML models are an advanced statistical technique that allow the use of dependent variables that do not follow normal distributions. Analysts can build ML models in which the relationship between independent and dependent variables is not linear; the dependent variable can be discrete and categorical rather than continuous.

Simpson's Paradox

In simple terms, Simpson's Paradox occurs when a trend appears in subgroups but disappears or is reversed when subgroups are combined into a single dataset. Knowledge Studio can detect this statistical phenomenon automatically and increase the precision of machine leaning models.

Imbalanced Classes

Most ML algorithms assume there are equal numbers of examples for each class in the source data; however, many datasets contain substantially different numbers of records for important classes — resulting in an imbalanced class problem. Failure to handle this properly results in models with poor predictive performance. Knowledge Studio has a node specifically built to handle imbalanced class issues. It can identify and correct imbalanced class problems.

Substitute Missing Values

Datasets often have missing values due to file corruption, failure to record data points, or other causes. Handling missing data values correctly is critical to developing accurate predictive models. Knowledge Studio makes it easy to identify datasets containing missing values and generate new substitute values based on a variety of substitution methods.

Autoregressive Integrated Moving Average (ARIMA) Models

ARIMA is a simple yet powerful method for making skillful time series forecasts, often incorporating seasonal and other types of semi-regular variations. For example, you can use ARIMA models to forecast electricity and raw materials utilization in a factory, output volumes in an oil refinery, fuel consumption for truck fleet, rail, and seaborne shipping companies, patient churn and intake volumes in hospitals, or sales, income, and expenses in any kind of business. Knowledge Studio's Auto ARIMA function estimates values for ARIMA parameters using grid search and step-wise algorithms.

Auto-Feature Engineering

Knowledge Studio allows users to specify relationships between multiple tables and automatically generate transformations and aggregations for every available combination of model features. This enables users to save time by generating multiple candidate features from a dataset automatically.

EXAMPLE USE CASES

Credit Scoring

Scoring is the process of running new data through the chosen champion model to predict an outcome. Lenders use Knowledge Studio to build and test predictive models that make decisions about the creditworthiness of loan applicants. The models are trained on historical data as well as on third-party providers delinquency scores, payment ratings, demographic attributes, and current account activity. Lenders then are free to score new applications for risk-based products and based on the outcome, lenders will then determine whether a risk-based product should be granted or not. Predictive AI models improve the ability to manage risk and optimize resources and allow the company to assess and optimize each stage of the customer credit lifecycle.

Tool Condition Monitoring (TCM) and Remaining Useful Life (RUL) Analysis

Tool wear in metal cutting operations has a direct impact on the quality and accuracy of finished surfaces. Replacing a tool at the optimum time — while it is still performing up to spec but just before its degradation begins to cause damage, reduce output quality, or increase scrap rates — is highly desirable. The large amounts of data produced by sensors combined with human inspections of finished pieces can be used to train machine learning algorithms to identify the "sweet spot" and proactively alert operators when a tool is approaching time for replacement.

Machine Failure Prediction

ML technology leverages historical and real-time data from sensors mounted to production equipment as well as PLCs, SCADA, and other sources and can accurately flag potential failures of whole machines and/or critical components before they can cause downtime. Failures may be binary in nature; that is, either a failure occurred or not. Failures can also be multi-class and fall into several different categories, including reduced speed, throughput, or quality. Obviously, the more complex the machine (or system), the more machine learning models can help prevent failures that can impact productivity.

Warranty Risk Profile Analysis

Most manufacturers must handle large numbers of warranty claims related to a variety of products and components. The volume of claims can easily run to millions per year for consumer goods manufacturers. It is critical to prioritize and understand which issues deserve high priority responses and detect patterns within the claims that indicate emerging quality or design problems that requiring immediate attention. Warranty risk profile analysis, sometimes referred to as quality issue prioritization, is a vital part of any ongoing quality improvement process. The data from warranty claims, once cleansed and sorted, is one of the most valuable parts of the feedback loop that enables companies to improve the reliability and customer satisfaction for their products.

Reduce Accounts Receivable Days Outstanding

Days sales outstanding (DSO) is a critical performance measure for many businesses, and anything that can reduce the firm's DSO will improve the bottom line. A common challenge involves the many different systems that a firm may use to manage its inventory, production processes, shipments, sales, and accounting. Reconciling data from such disparate systems, which were often implemented by different teams and different times with different objectives, is a nontrivial task. Machine learning algorithms can anticipate potential slow payers and increasing DSO numbers before they can affect the business.

Cross-Selling to Increase Loyalty and Revenue

Knowledge Studio can create dynamic customer profiles and aid in applying relevant product and service treatments to customer segments. Understanding how customers use products and services over time or knowing when not to position an offering to a customer can refine cross-selling strategies that lead to improved customer loyalty and increased repeat business. Customer success teams can build predictive models with Knowledge Studio that examine the strength of patterns and relationships in sales data. They can gain insight into purchasing and usage patterns to create more effective cross-selling campaigns and develop marketing strategies that take multiple distribution channels into account.

To learn more, please visit **altair.com/knowledge-studio**

Knowledge Studio is efficient and easy to use for business analysts and data scientists alike. It enables development of sophisticated AI and machine learning models with no coding required. Data science doesn't have to be complicated to be powerful.

Learn More: **altair.com/machine-learning**