The Forrester Wave™: Multimodal Predictive Analytics And Machine Learning, Q3 2020

The 11 Providers That Matter Most And How They Stack Up

by Mike Gualtieri and Kjell Carlsson, PhD

September 10, 2020

Why Read This Report

In our 26-criterion evaluation of multimodal predictive analytics and machine learning (PAML) providers, we identified the 11 most significant ones — Altair, Alteryx, BigML, Dataiku, IBM, KNIME, Minitab, RapidMiner, Samsung SDS, SAS, and TIBCO Software — and researched, analyzed, and scored them. This report shows how each provider measures up and helps application development and delivery (AD&D) professionals select the right one for their needs.

Key Takeaways

IBM, SAS, RapidMiner, Dataiku, And TIBCO Software Lead The Pack

Forrester's research uncovered a market in which IBM, SAS, RapidMiner, Dataiku, and TIBCO Software are Leaders; Alteryx and KNIME are Strong Performers; Samsung SDS and BigML are Contenders; and Altair and Minitab are Challengers.

Automation, ModelOps, And Product Roadmap Should Factor Prominently In Buy Decisions

As the demand for AI skyrockets, vendors must provide data science and extended AI teams with more automation to increase productivity, model operations for smooth deployment, and a product roadmap that makes breakneck machine learning innovations accessible.

The Forrester Wave™: Multimodal Predictive Analytics And Machine Learning, Q3 2020

The 11 Providers That Matter Most And How They Stack Up

by Mike Gualtieri and Kjell Carlsson, PhD
with Srividya Sridharan and Robert Perdoni

September 10, 2020

Table Of Contents

Related Research Documents

The Forrester Wave™: Notebook-Based Predictive Analytics And Machine Learning, Q3 2020

Introducing AI-Powered, Human-Controlled Digital Decisioning Platforms

Introducing ModelOps To Operationalize AI

Now Tech: Predictive Analytics And Machine Learning, Q2 2020

FORRESTER®

Forrester Research, Inc., 60 Acorn Park Drive, Cambridge, MA 02140 USA
+1 617-613-6000 | Fax: +1 617-613-5000 | forrester.com

Multimodal PAML Will Make Or Break Enterprise AI Ambitions

For enterprises serious about AI, implementing the right PAML solutions will be the most important decision they make. That's because AI is fundamentally composed of machine learning (ML) models, and the pace at which data science teams can build and deploy them is directly correlated to the nuggets and then the mother lode of intelligence that models can infuse within business processes and customer engagement applications. As a reminder, Forrester segments PAML into three segments: multimodal, notebook-based, and automation-focused. This Forrester Wave evaluates vendors in the multimodal segment.

Multimodal PAML customers should look for providers that offer:

- **AutoML that isn't just a checkbox.** Automated machine learning (AutoML) is not about replacing data scientists. It is about making them dramatically more productive by providing them with tools that run repetitive tasks in the model development lifecycle, such as feature engineering, training, and evaluation. AutoML is no longer solely in the purview of automation-focused PAML vendors.[1] It is solidly mainstream due to a combination of innovations and its essentiality to meet enterprise demand for more models, faster. Some multimodal PAML vendors offer AutoML capabilities that rival automation-focused PAML vendors, such as DataRobot and H2O.ai, which are not in this evaluation because Forrester categorizes them as automation-focused PAML. Most multimodal PAML vendors offer some degree of AutoML but must invest much more to stay competitive.
- **Model operations (ModelOps) to operationalize ML at scale.** A top complaint of data science teams and, increasingly, line-of-business leaders is the challenges in deploying machine learning models in production. Frantic handoffs, manual monitoring, and loose governance impede organizations' ability to deploy more business-worthy AI use cases faster. Multimodal PAML vendors have traditionally focused their product development on analytical tools and ML methods. That's still important, but now that AI has gone more mainstream, enterprises want more machine learning, faster. ModelOps capabilities within PAML products enable cross-functional AI teams to efficiently deploy, monitor, retrain, and govern AI models in production systems.[2]
- **A roadmap that is as bold as your AI ambitions.** Most multimodal PAML vendor products can help data scientists get the job done. You can use these solutions to acquire data, build a data transformation pipeline, apply transformations, visualize results, choose from among many analytical methods, and deploy models. However, the importance and scope of PAML has expanded to become strategic versus tactical, because ML models are the fundamental building blocks of AI, which will become a key force in the digital future. Enterprises must now choose a multimodal PAML solution to satisfy their AI ambitions. That definitely means looking at the strength of the vendor's current offering, but it also means examining the vendor's roadmap and ability to execute on a vision of an AI platform that 1) integrates with other tools and technologies and 2) expands beyond data science teams to also serve broader AI teams.

Evaluation Summary

The Forrester Wave evaluation highlights Leaders, Strong Performers, Contenders, and Challengers. It's an assessment of the top vendors in the market and does not represent the entire vendor landscape. You'll find more information about this market in our reports on predictive analytics and machine learning.

We intend this evaluation to be a starting point only and encourage clients to view product evaluations and adapt criteria weightings using the Excel-based vendor comparison tool (see Figure 1 and see Figure 2). Click the link at the beginning of this report on Forrester.com to download the tool.

FIGURE 1 Forrester Wave™: Multimodal Predictive Analytics And Machine Learning, Q3 2020

FIGURE 2 Forrester Wave™: Multimodal Predictive Analytics And Machine Learning Scorecard, Q3 2020

	Forrester's weighting	Altair	Alteryx	BigML	Dataiku	IBM	KNIME	Minitab	RapidMiner	Samsung SDS	SAS	TIBCO Software
Current offering	50%	1.88	3.14	2.62	3.24	4.36	3.21	1.00	4.00	2.55	4.26	3.17
Data	14%	1.50	3.00	2.00	4.00	4.50	3.50	1.00	3.00	2.50	3.50	2.50
Modeling	14%	1.67	3.67	1.67	3.67	4.33	3.00	1.00	5.00	3.00	3.67	4.33
Collaboration	14%	1.67	4.33	2.33	4.33	3.00	4.33	1.00	4.33	3.67	4.33	3.67
Model evaluation	14%	2.33	1.67	3.00	1.67	5.00	2.33	1.00	5.00	1.00	5.00	2.33
Model operations (ModelOps)	14%	3.00	2.33	2.33	3.00	3.67	2.33	1.00	3.67	1.67	4.33	2.33
Methods and algorithms	14%	2.00	4.00	2.00	3.00	5.00	4.00	1.00	4.00	3.00	4.00	4.00
Platform infrastructure	14%	1.00	3.00	5.00	3.00	5.00	3.00	1.00	3.00	3.00	5.00	3.00
Strategy	50%	1.00	3.50	2.00	4.00	4.50	2.50	1.50	4.00	2.50	4.00	4.00
Ability to execute	25%	1.00	3.00	3.00	5.00	5.00	3.00	1.00	5.00	3.00	5.00	5.00
Solution roadmap	25%	1.00	3.00	1.00	5.00	5.00	3.00	1.00	5.00	1.00	3.00	3.00
Enablement	25%	1.00	3.00	3.00	3.00	5.00	1.00	3.00	3.00	5.00	5.00	5.00
Partners	25%	1.00	5.00	1.00	3.00	3.00	3.00	1.00	3.00	1.00	3.00	3.00
Market presence	0%	1.00	4.33	1.00	3.00	4.33	3.67	1.00	3.00	2.33	5.00	3.00
Customer adoption	33%	1.00	5.00	1.00	3.00	5.00	5.00	1.00	3.00	3.00	5.00	5.00
Evaluated product revenue	33%	1.00	3.00	1.00	1.00	3.00	1.00	1.00	1.00	3.00	5.00	1.00
Market awareness	33%	1.00	5.00	1.00	5.00	5.00	5.00	1.00	5.00	1.00	5.00	3.00

All scores are based on a scale of 0 (weak) to 5 (strong).

Vendor Offerings

Forrester included 11 vendors in this assessment: Altair, Alteryx, BigML, Dataiku, IBM, KNIME, Minitab, RapidMiner, Samsung SDS, SAS, and TIBCO Software (see Figure 3).

FIGURE 3 Evaluated Vendors And Product Information

Vendor	Product evaluated	Product version evaluated
Altair	Altair Knowledge Studio, Altair Knowledge Studio for Apache Spark, Knowledge Manager	
Alteryx	Alteryx APA (Analytic Process Automation) Platform	2020.2
BigML	BigML	BigML Enterprise
Dataiku	Dataiku Data Science Studio (DSS)	Dataiku DSS 7.0.2
IBM	IBM Watson Studio	
KNIME	KNIME Software (KNIME Analytics Platform, KNIME Server)	4.2
Minitab	Minitab's Salford Predictive Modeler	Version 8
RapidMiner	RapidMiner Platform	RapidMiner Studio, RapidMiner AI Hub 9.7
Samsung SDS	Brightics AI	V3.7
SAS	SAS Visual Data Mining and Machine Learning (VDMML), SAS Model Manager	VDMML 8.5; MM 15.3
TIBCO Software	TIBCO Data Science	

Vendor Profiles

Our analysis uncovered the following strengths and weaknesses of individual vendors.

Leaders

› **IBM is packed with AI lifecycle services everywhere — public, private, and on-prem.** What do you get when you combine a full stack of data analytics capabilities — from data management to PAML to business intelligence — in a microservices framework that can run seamlessly on-premises, in the private cloud, and in multiple public clouds? IBM Cloud Pak for Data, an offering that makes capabilities across the PAML lifecycle available when and where your users need them. The crown jewel of Cloud Pak for Data is Watson Studio, a PAML offering that combines easy-to-use, SPSS-inspired workflow capabilities with open source ML libraries and notebook-based interfaces. IBM continues to add innovations from IBM Research like fairness monitoring, bias mitigation, AutoML, and federated learning.

IBM offers a compelling, scalable, increasingly integrated, and harmonized platform with differentiated capabilities that spans the entire PAML lifecycle and can be deployed anywhere across any cloud. Users will have to navigate some technical previews and quickly evolving features, but that will save them from otherwise having to stitch together a mismatch of proprietary solutions and open source software.

› **SAS opens up with Viya.** Viya is SAS's completely reengineered platform that underpins all of the vendor's solutions, including SAS Visual Analytics, SAS Visual Data Mining and Machine Learning, and SAS Model Manager. Viya cloud enables SAS products, unifies user interfaces, and provides interoperability, integration, and extensibility with APIs. The result? SAS can innovate faster, offer cloud solutions, and integrate with external services and open source tools.

SAS offers strengths across the board, including exceptionally well-integrated AutoML and other guided analytics capabilities. SAS increasingly supports open source programming languages including Python and R, allowing data science teams that use these tools to leverage the SAS engine. With SAS Model Manager, it now supports the operationalization of models not developed in SAS and is moving closer to becoming a holistic PAML platform that supports all your data science needs.

› **RapidMiner blends ease of use with data science rigor.** Your friendly neighborhood data science tool has put on a business suit and is looking good in it. Enterprise AI teams composed of seasoned data scientists and data-savvy engineers, analysts, and business users will value RapidMiner's blend of expansive functionality and attention to statistical rigor with ease of use and automation. It has some of the most productivity-enhancing capabilities for automated data preparation (Turbo Prep) and model development (Auto Model) in the multimodal market, along with one of the most comprehensive visual tools for building data and ML pipelines. RapidMiner also supports Python-loving data scientists who prefer notebooks, and it enables everyone to operationalize and manage their models on a common platform.

In the near term, RapidMiner might not just have something for everyone; it could have everything for everyone. But first it needs to round out its already impressive capabilities supporting notebooks and ModelOps and finish transitioning its IDE to the web; and, like most multimodal PAML offerings, it wouldn't hurt if its deep learning capabilities for text and computer vision use cases got more investment too.

› **Dataiku ascends to the crème de la crème.** It's hard to believe that Dataiku is only seven years old. In that short time, the vendor has been on an unstoppable march toward democratizing its enterprise PAML for a host of different roles within an enterprise, from business analysts to data scientists. It also offers persona-targeted training and enablement programs that help enterprises enable a large number of users with data science capabilities.

Enterprises looking to spark or accelerate a data and ML movement in their organizations will find their needs play to Dataiku's raison d'être. Hardcore data scientists might sniff that Dataiku doesn't support every latest avant-garde ML framework, but even they will appreciate its built-in support for Jupyter notebooks and external IDEs (PyCharm, RStudio, Sublime, and VS Code), new ModelOps capabilities, and behind-the-scenes automation of critical tasks such as infrastructure provisioning. Dataiku also needs to catch up to its competitors when it comes to AutoML and monitoring capabilities, but the vendor's always aggressive roadmap should help it continue to grow and thrive even among much larger competitors.

› **TIBCO Data Science tackles complex, consequential use cases.** While most vendor machine learning solutions focus on ML models, TIBCO's focus is on AI applications. TIBCO Data Science supports ML model creation through the entire lifecycle, but it also integrates elegantly with the company's other platform products to create insights and/or AI apps. For insights, TIBCO Data Science integrates with TIBCO Spotfire. For AI solutions, the ML model-building lifecycle seamlessly intersects with the rest of TIBCO's application portfolio, such as TIBCO Streaming. TIBCO Data Science is used by its customers to build a wide range of AI apps, from customer engagement use cases to silicon manufacturing, as well as a host of internet-of-things (IoT) apps.

TIBCO's strengths are its data exploration, data preparation, security, and modeling tools. Its AutoML tools are designed to augment analysts (with seamless integration in both Spotfire and TIBCO Data Science) and data scientists (by outputting reusable pipelines and Python code), and more automation features are on the way. To be more competitive, TIBCO should invest more in automated feature engineering, model evaluation, and experiment tracking.

Strong Performers

› **Alteryx's customers come for the data, stay for the machine learning.** Alteryx may be best known to data scientists as the data prep tool they wish they had, but the company has been on a continuous mission to span the full data and ML lifecycle. It has long had capabilities for

building predictive models and has added capabilities for AutoML, notebooks, and capabilities for deploying models. Particularly exciting is its acquisition of Feature Labs, which promises, when integrated, to bring a host of automated feature-engineering capabilities to the Alteryx platform.

Professional photographers often say that the best camera is the one you carry with you. For the ever-increasing number of citizen data scientists, the best PAML platform might just well be the data platform they already have, know how to use, and will need as part of the PAML lifecycle anyway. However, to win over hardcore data science teams, Alteryx will need to develop its modeling and model operationalization capabilities.

› **KNIME continues to foster a vibrant community of open source contributors.** Thanks to its community contributors, the KNIME Analytics Platform includes approximately 4,000 analytical, statistical, data transformation, and ML methods downloadable to all for free. That's a lot of value for individual users. The company supports its open source community stewardship by offering small, medium, and large subscriptions to the KNIME Server, which provides data science teams with additional features for collaboration, automation, deployment, and management.

KNIME offers strengths in visual modeling tools, breadth of analytical methods, automation, and applications. Recent useful additions drive collaboration and speed to deployment by allowing users to package and parameterize parts of pipelines into reusable, shareable "components" and then tag parts of model development pipelines to build immediate production pipelines. The KNIME Analytics Platform's sweet spot is for noncoding data science teams. KNIME should invest more in coding interfaces to broaden its appeal to coding data scientists.

Contenders

› **Samsung SDS's unsung platform goes broad and deep (learning).** Given the vast array of Samsung's products, from ships to 5G networking infrastructure, it should come as no surprise that it offers a full-fledged PAML platform. You might not have heard of Samsung's Brightics AI platform, but it has everything you need to develop and deploy a suite of ML and AI apps, including data prep, a visual environment for building data and ML pipelines, AutoML capabilities for guided analytics and hyperparameter optimization, and capabilities for deploying your models. It'll come as less of a surprise when you hear that Brightics AI is the platform that Samsung SDS — the $9.2 billion in revenue, 23,000-plus employee strong IT services firm — uses internally to deliver digital transformation projects ranging from demand forecasting to chemotherapy treatment recommendations.

Samsung's top strengths are the ease and speed with which users can create straightforward ML applications, from data ingestion to simple analytics apps — and it has a few tricks up its sleeves, especially for deep learning applications (e.g., pretrained models and features for accelerating data labeling). Business users can quickly become citizen data scientists and citizen developers. However, to become the enterprise PAML platform of choice, Samsung will need to build out richer capabilities across the PAML lifecycle.

› **BigML is instant, elegant ML.** With a few clicks on bigml.com, professional or aspiring data scientists can train and deploy machine models using BigML's sophisticated yet intuitive graphical user interface. Founded in 2011, BigML's stated mission has been to make machine learning "easy and beautiful for everyone." Mission accomplished. Web access and free-to-low entry-point pricing has resulted in more than 120,000 users worldwide. In recent years, BigML has expanded its focus on individual users and small teams to also include the needs of large enterprises, including on-premises deployments.

BigML's strengths are its overall ease of use as well as its support for unsupervised learning techniques (e.g., clustering, anomaly detection, topic modeling, and association discovery) and AutoML capabilities. However, to become more competitive in the enterprise market, BigML must find a perfect balance between keeping the solution easy to use and expanding the features that enterprises expect and competitors already offer. Access to open source, more comprehensive data-wrangling tools and collaboration features, and stronger model monitoring capabilities will move the dial the most for BigML.

Challengers

› **Altair makes data science visually joyful.** Altair Knowledge Studio is truly designed for all skill levels, from subject matter experts with no formal data science training to seasoned data scientists diligently working to squeeze more performance out of an existing model. Knowledge Studio supports many popular machine techniques, but it differentiates by providing a highly functional yet simple-to-use visual interface for creating, exploring, and modifying decision trees and strategy trees.

Altair Knowledge Studio offers strengths in data exploration, visual modeling interfaces, governance, and ModelOps. To be more competitive, Altair should invest in automation features, integrate more tightly open source methods, and become fully cloud native.

› **Minitab makes authoritative ML methods available to the Minitab masses.** Salford Predictive Modeler (SPM) is best known for its implementation of specific algorithms: CART, MARS, random forests, and TreeNet. Most other vendor solutions have these methods and many more, but Minitab's methods are implemented and fine-tuned by their inventors — including Jerome Friedman, a professor of statistics at Stanford University. Aspiring citizen data scientists using Minitab can now access numerous SPM capabilities, including prebuilt modeling pipelines that are based on data science best practices and targeted for both vertical and horizontal use cases.

SPM's novel approach of implementing specific algorithms better than anyone else was successful in the past and will be a welcome addition for Minitab users. However, Minitab must invest faster across the board to compete in today's market.

Evaluation Overview

We evaluated vendors against 26 criteria, which we grouped into three high-level categories:

- **Current offering.** Each vendor's position on the vertical axis of the Forrester Wave graphic indicates the strength of its current offering. Key criteria for these solutions include data, modeling, collaboration, model evaluation, model operations, methods and algorithms, and platform infrastructure.
- **Strategy.** Placement on the horizontal axis indicates the strength of the vendors' strategies. We evaluated their ability to execute, solution roadmap, enablement, and partners.
- **Market presence.** Represented by the size of the markers on the graphic, our market presence scores reflect each vendor's customer adoption, evaluated product revenue, and market awareness.

Vendor Inclusion Criteria

Forrester included 11 vendors in the assessment: Altair, Alteryx, BigML, Dataiku, IBM, KNIME, Minitab, RapidMiner, Samsung SDS, SAS, and TIBCO Software. Each of these vendors has:

- **A multimodal PAML, as identified by Forrester.** Vendors included in this evaluation must offer a multimodal PAML solution as defined above and as represented in the latest Now Tech report on this market.[3]
- **A comprehensive, differentiated multimodal PAML solution.** The vendors included in the evaluation must offer a solution that can operate on large data sets and provide capabilities for data acquisition and preparation, statistical and ML algorithms, a differentiated user interface to build models, and ModelOps features.
- **A standalone multimodal PAML marketed to enterprise data science/AI teams.** Forrester included only solutions marketed toward enterprise data science and/or AI teams that use ML algorithms and other analytical algorithms to build predictive models and surface insights. PAML solutions that Forrester deemed to be technologically embedded into any particular application, business intelligence, data prep, applications, ETL, or middleware stacks were not included in this evaluation.[4]
- **Minimum install base and revenue requirements.** The vendor must have at least three paying, named enterprise customers using the PAML solution. The vendor must have provided Forrester with three customer references that were willing to speak with us and fill out a survey. Please note that reference calls and surveys will be kept under nondisclosure. Included vendors must also have a trailing 12-month revenue of at least $4 million.
- **Sparked client inquiries that put the vendor on Forrester's radar.** Forrester clients often discuss the vendors and products through inquiries; alternatively, the vendor may, in Forrester's judgment, warrant inclusion or exclusion in this evaluation because of technology trends, market presence, lack of client interest, or loss of market momentum.

Supplemental Material

Online Resource

We publish all our Forrester Wave scores and weightings in an Excel file that provides detailed product evaluations and customizable rankings; download this tool by clicking the link at the beginning of this report on Forrester.com. We intend these scores and default weightings to serve only as a starting point and encourage readers to adapt the weightings to fit their individual needs.

The Forrester Wave Methodology

A Forrester Wave is a guide for buyers considering their purchasing options in a technology marketplace. To offer an equitable process for all participants, Forrester follows The Forrester Wave™ Methodology Guide to evaluate participating vendors.

In our review, we conduct primary research to develop a list of vendors to consider for the evaluation. From that initial pool of vendors, we narrow our final list based on the inclusion criteria. We then gather details of product and strategy through a detailed questionnaire, demos/briefings, and customer reference surveys/interviews. We use those inputs, along with the analyst's experience and expertise in the marketplace, to score vendors, using a relative rating system that compares each vendor against the others in the evaluation.

We include the Forrester Wave publishing date (quarter and year) clearly in the title of each Forrester Wave report. We evaluated the vendors participating in this Forrester Wave using materials they provided to us by June 12, 2020, and did not allow additional information after that point. We encourage readers to evaluate how the market and vendor offerings change over time.

In accordance with The Forrester Wave™ Vendor Review Policy, Forrester asks vendors to review our findings prior to publishing to check for accuracy. Vendors marked as nonparticipating vendors in the Forrester Wave graphic met our defined inclusion criteria but declined to participate in or contributed only partially to the evaluation. We score these vendors in accordance with The Forrester Wave™ And The Forrester New Wave™ Nonparticipating And Incomplete Participation Vendor Policy and publish their positioning along with those of the participating vendors.

Integrity Policy

We conduct all our research, including Forrester Wave evaluations, in accordance with the Integrity Policy posted on our website.

Endnotes

[1] See the Forrester report "Now Tech: Predictive Analytics And Machine Learning, Q2 2020."

[2] See the Forrester report "Introducing ModelOps To Operationalize AI."

[3] See the Forrester report "Now Tech: Predictive Analytics And Machine Learning, Q2 2020."

[4] ETL: extract, transform, load.

We work with business and technology leaders to drive customer-obsessed vision, strategy, and execution that accelerate growth.

PRODUCTS AND SERVICES

- Research and tools
- Analyst engagement
- Data and analytics
- Peer collaboration
- Consulting
- Events
- Certification programs

Forrester's research and insights are tailored to your role and critical business initiatives.

ROLES WE SERVE

Marketing & Strategy Professionals

CMO
B2B Marketing
B2C Marketing
Customer Experience
Customer Insights
eBusiness & Channel Strategy

Technology Management Professionals

CIO
- Application Development & Delivery

Enterprise Architecture
Infrastructure & Operations
Security & Risk
Sourcing & Vendor Management

Technology Industry Professionals

Analyst Relations

CLIENT SUPPORT

For information on hard-copy or electronic reprints, please contact Client Support at +1 866-367-7378, +1 617-613-5730, or clientsupport@forrester.com. We offer quantity discounts and special pricing for academic and nonprofit institutions.