sparkcognition™

The Darwin™ Difference: Why Darwin Stands Out From the AutoML Pack

How Darwin Stacks Up to Its Competitors

Darwin™ is an automated model building product that allows you to go from data to model in less time than traditional methods, enabling the rapid prototyping of scenarios and productive extraction of insights. Despite the recent growth of automated model building products on the market, Darwin remains unmatched in the speed, scale, and accuracy it delivers.

		AutoML Competitors		ML Toolkits
	DARWIN™	DATAROBOT	H20	OPEN SOURCE
AUTOML CAPABILITIES				
AUTOMATED DATA VISUALIZATION	✓	✓	✓	✗
GUIDED DATA PREPARATION	✓	✗	✗	✗
AUTOMATED DATA CLEANING & FEATURE GENERATION	✓	✓	✓	✗
AUTOMATED MODEL OPTIMIZATION*	✓	✓	✓	✗
PREDICTION EXPLANATION	✓	✓	✓	✓
SUPPORTED LEARNING MODES				
SUPERVISED LEARNING	✓	✓	✓	✓
TIME SERIES SUPPORT	✓	✓	✓	✓
UNSUPERVISED LEARNING	✓	✗	✗	✓
SEMI-SUPERVISED LEARNING	✓	✗	✗	✓
DYNAMIC DEEP LEARNING	✓	✗	✗	✓
DEPLOYMENT CAPABILITIES				
CLOUD AND ON-PREMISES	✓	✓	✓	✗
RUN-TIME ENGINE (DEVICE)	✓	✓	✗	✗
AUTOMATED MODEL MAINTENANCE (SDK/API)	✓	✓	✓	✗

Darwin Outperforms in Time Series Data and Complex Problems

	DARWIN™	AUTOSKLEARN	H20	RANDOM FOREST
ELECTRIC DEVICES CLASSIFICATION	0.58	0.13	0.33	0.47
UCI BEIJING WEATHER TIME-SERIES REGRESSION	0.55	0.29	-0.29	0.14
UCI EEG EYE STATE CLASSIFICATION	0.84	0.47	0.61	0.61
UCI OZONE CLASSIFICATION	0.99	0.98	0.98	0.98
MNIST DIGIT CLASSIFICATION	0.97	-	0.97	0.87
BOSTON HOUSING REGRESSION	0.72	0.72	0.78	0.73

From the Darwin Efficacy Report

Darwin Use Cases Collection

Darwin allows organizations to efficiently scale machine learning applications, delivering business-wide impact by making machine learning a first-class citizen, ubiquitous and transparent across all teams.

ACCURACY	APPLICATION
92%	Identification of customers at risk for loan defaults and delinquency
94%	Identification of customers who are at risk of churning
80%	Identification of transactions at risk of being past due
80%	Detection of fraudulent activity on electronic transactions
100%	Classify subterranean drill-head operational states
90%	Predict automotive sub-component quality during assembly
83%	Identify degredation in commercial aircraft components
90%	Detect impurities during iron ore manufacturing
92%	Identifying emotions based on images of facial expressions
80%	Predicting NBA teams' win/loss record based on player and team statistics
84%	Evaluating political bias on online articles
72%	Predicting the likelihood of adoption for a shelter animal
96%	Predicting matches on speed dating events based on features of participants
91%	Approval of mortgages based on demographics and historical credits
96%	Predicting the total cost generated by disasterous events like hurricanes
94%	Identifying whether or not a patient has cardiovascular disease

Automated model building is not the enemy of data scientists. We can use it as an ally in order to spend more time in more complex problems, converting data into knowledge.

— JAIME CASTELLANOS
Global Hitss Data Scientist

The Business Case for Darwin

Darwin enables data science and analyst groups across your organization to achieve fast, tangible results that impact your bottom line. Darwin's core patented technology lies in automating the model-building process, which allows data professionals to drive higher efficiencies and discover new revenue opportunities within a line of business.

INCREASED TEAM PRODUCTION

Patented approach centered around productive automation workflows that reduce the work of months to days

STAY AT FRONTLINE OF INNOVATION

State-of-the-art automated model building technology in the hands of all your teams future-proofing your operations

PLUG-AND-PLAY FOR IMMEDIATE VALUE

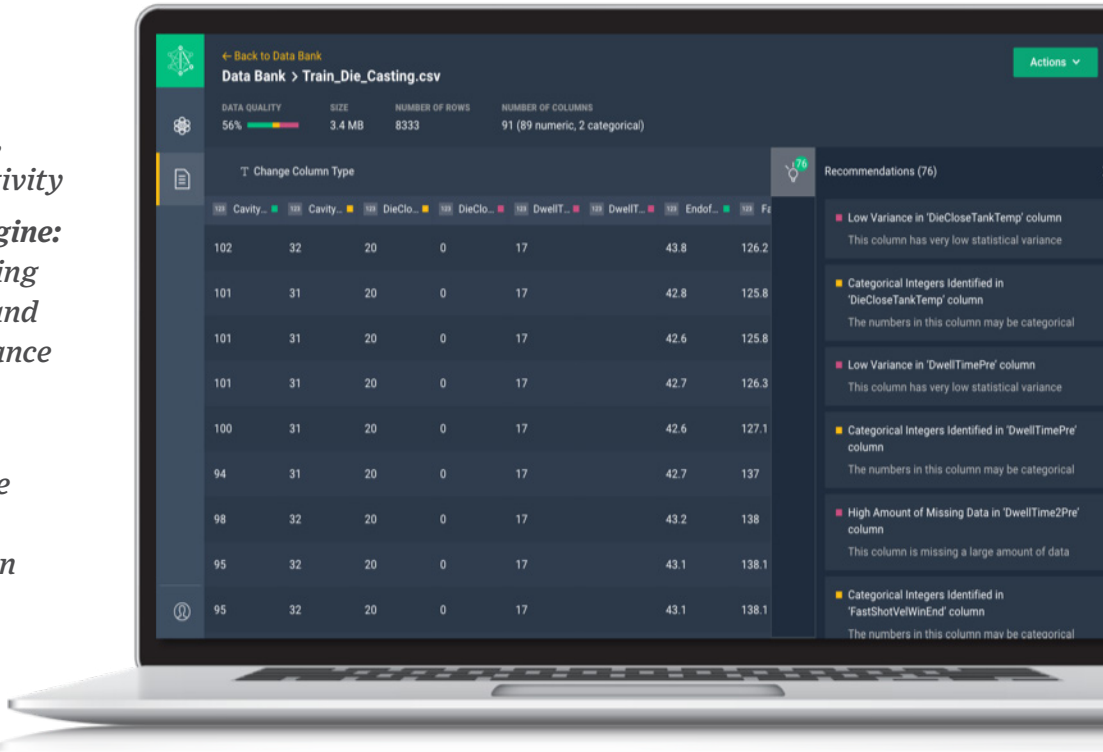
Seamless and rapid integration with your existing systems to provide value from day one

Global HITSS, a subsidiary of America Movil, saw a 30% increase in productivity in one case study, reducing the amount of time spent on a project from 30 days to only 10 minutes.

From a User Perspective

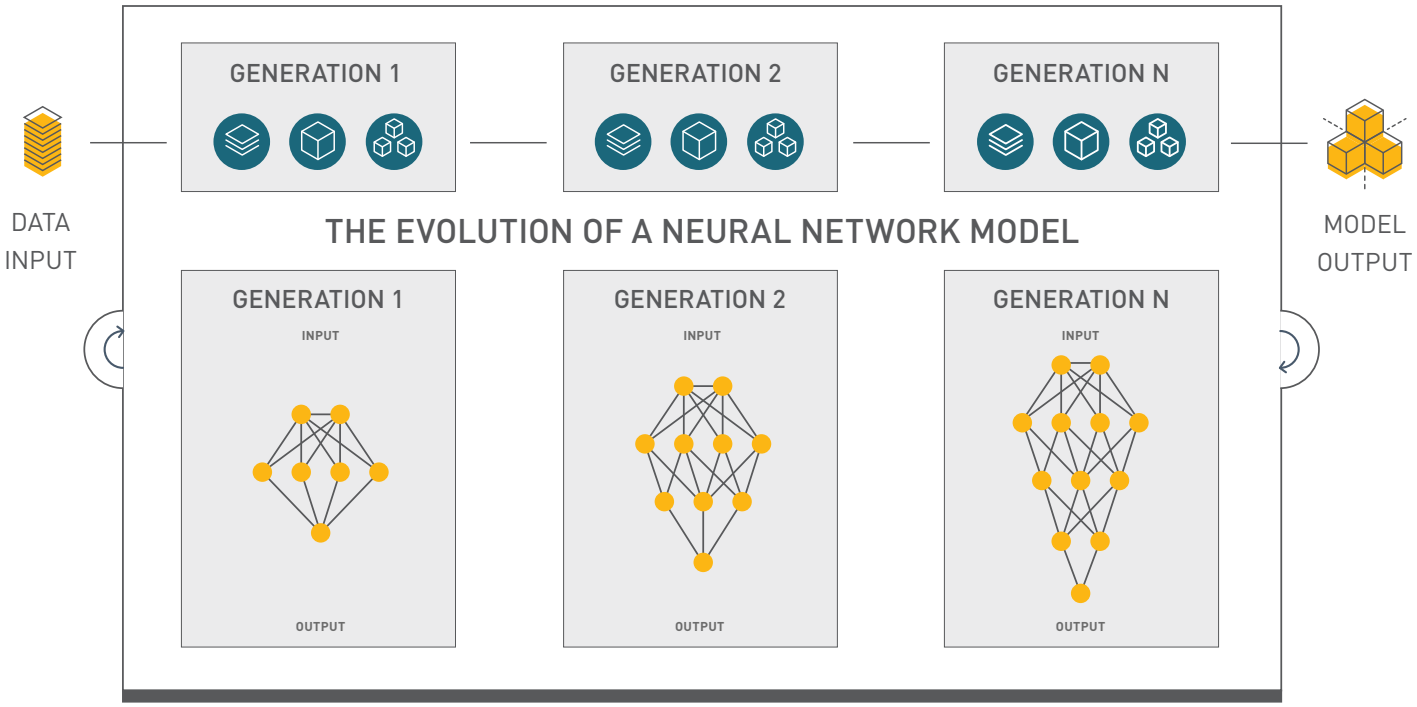
Darwin empowers data professionals with an intuitive environment that takes them from zero to model in less time than traditional methods, no programming required.

- ***Intuitive Data Preparation and Modeling Paradigms:*** *Guided approach that eases the complexity of the science, positively impacting productivity*
- ***Quality Driven Modeling Engine:*** *Models are tailored considering the intracacies of your data and evaluated based on performance and accuracy*
- ***Streamlined Deployment Pipelines:*** *APIs and Run-Time Engine options to interface with your existing application environment*



A Deeper Look at the Science Behind Darwin

Darwin's approach focuses on the acceleration of data preparation and model-building workflows in the data science process. Darwin uses a patented neuroevolutionary approach that specializes in discovering novel, elegant network architectures (also supporting hyperparameter search for common algorithms such as Random Forest and XGBoost). This approach is used when creating solutions for complex data sets with intricate relationships, or long time series problems, given the power and flexibility neural networks offer. During this process Darwin cleans your data to make it usable for algorithmic development, generating dozens of features to enrich your dataset and ultimately kicking off the creation and optimization of thousands of models.



Why Does Neuroevolution Matter?

Rather than simply choosing the best performer from a predefined list of algorithms, Darwin uses a blend of evolutionary and deep learning methods to iteratively find the optimal model tailored to your data. This automated model building process effectively creates unique solutions that precisely and accurately generate predictions for your data problems.

- ***Unparalleled Accuracy Through Deep Learning:*** *Neuroevolutionary process specialized in the search and auto tuning of neural architectures based on the intracacies of your data*
- ***Handle Complex Temporal Relationships*** *LSTM and TCN architectures to capture complex relationships over time and exploit then to make more accurate predictions*
- ***True Generalization to Address the Unknown:*** *Neuroevolution quickly adapts to dynamic circumstances, evolving architectures to mantain complexity, efficacy, and efficiency*