What is **Data Science?**

Transforming Data into Value

+1 813 374 0027 ask@nix-united.com nix-united.com

Contents

01. Introduction 2

02. Data Science In Plain Words 3

03. How Data Science Delivers Business Value 11

04. How Data Science Empowers Industries 16

05. 8 Steps Before Starting A Data Science Project 30

06. How To Choose A Data Science Vendor Wisely 36

07. How NIX United Provides Data Science Services 40

01. Introduction

Organizations are overflowing with data. But the real value of data is not in their accumulation, but in their selection and effective use. Today, with companies constantly collecting huge amounts of data, they need to be analyzed and transformed to produce actionable insights—otherwise, it's a waste of valuable information.

Right now, data science is something of hype, and everyone is talking about what magic it can do. But not everyone understands what it is, what benefits the effective use of data can bring to your business specifically, and if you need it.

> ***"We're entering a new world in which data may be more important than software."***
>
> *—Tim O'Reilly, founder of O'Reilly Media*

After reading this document, you will have a basic understanding of what data science is, how it can benefit different industries, and where to start.

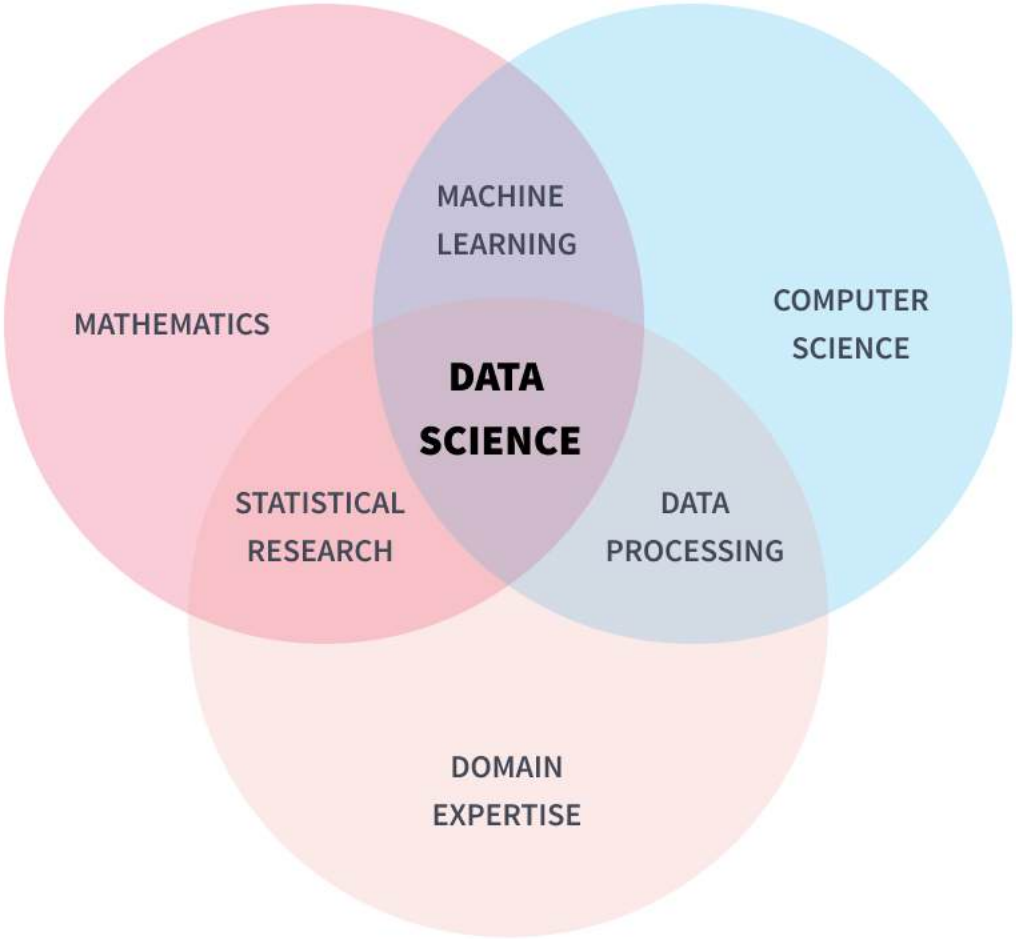
02. Data Science In Plain Words

Data science is a broad area of study that deals with vast volumes of data using scientific methods, processes, algorithms, and systems to extract knowledge from structured and unstructured data. The data used for analysis can come from many different sources and are presented in various formats. Data science uses the power of:

- Artificial Intelligence (AI)
- Machine Learning (ML)
- Business Intelligence (BI)

According to IBM, data science combines the scientific method, math and statistics, specialized programming, advanced analytics, AI, and even storytelling to uncover and explain business insights buried in data.

> ***"Data science is the transformation of data using mathematics and statistics into valuable insights, decisions, and products"***
>
> *—John W. Foreman, Chief Data Scientist for MailChimp*



Data scientists combine specialized tools and use mathematical analysis to identify data patterns and trends that can be put at the core of useful data models. As the science of data evolves and new tools are acquired over time, the primary goal of business remains focused on finding valuable patterns and extracting business-boosting insights from data.

Big Data

Businesses today are collecting increasingly larger data sets that contain various and typically unstructured records of customers, employees, and

corporate assets—literally anything that can be tracked digitally or entered manually. However, these unstructured data sets are not yet big data.

> ***"Collecting doesn't mean discovering"***
>
> *—Tim O'Reilly, founder of O'Reilly Media*

The term "big data" is used to describe data characterized by high volume, high speed, and high variety. Sources of data are becoming more complex than those for traditional data because they are being driven by artificial intelligence, mobile devices, social media, and the Internet of Things (IoT). This requires creating well-tuned, efficient, and secure big data management solutions to properly visualize data and uncover its essence.

> ***"Without big data analytics, companies are blind and deaf, wandering out onto the web like deer on a freeway."***
>
> *—JGeoffrey Moore, management consultant and author of Crossing the Chasm*

Data Engineering

Data engineering implies the collection, validation, storing, and accessibility of data, serving as a foundation for analytics and powering BI and AI solutions to make data-driven decisions. This includes real-time data processing, cloud-based big data management, and data system modernization to reach ultimate business goals faster.

nix CASE

How SaaS Solution for Efficient Healthcare Data Management and Executive-level Insights helped the client deliver a more personalized user experience and advanced business intelligence opportunities, which resulted in becoming one of the top-line healthcare providers on the market.

Robust data engineering solutions can process huge amounts of data, including audio, video, and other types of unstructured information, in seconds. Board members and key decision-makers can access information immediately, rather than waiting months for a dedicated manager or legacy systems to produce a report.

Data engineering powers BI and AI solutions to make data-driven decisions.

The use of data engineering services and technologies leads to the consistent management of big data across the organization, allowing consumers to meet their information needs better, faster, and even autonomously.

Data Analytics

Data analytics is the process of examining data sets to identify trends and draw conclusions about the information they contain. Data analytics

solutions sift through large volumes of data to answer business-related questions and improve organizational performance.

> ***"We are surrounded by data, but starved for insights."***
>
> *—Jay Baer, marketing and customer experience expert*

Data analytics allows organizations to digitally transform their business and culture, becoming more innovative and forward-thinking in their decision making. Going beyond traditional KPI monitoring and reporting and finding hidden patterns in data, algorithm-driven organizations become new innovators and business leaders.

Business Intelligence

BI uses software to transform vast amounts of raw data into valuable information and helps organizations make smart choices and informed decisions and set priorities. Business intelligence prioritizes descriptive analytics, which provides a summary of historical and present data to answer "what" and "how" questions.

> ***"BI is about providing the right data at the right time to the right people so that they can take the right decisions"***
>
> *—Nic Smith with Microsoft BI Solutions Marketing*

Comprehensive BI services use data mining, descriptive analytics, statistical analysis and visualization to uncover unseen trends and anomalies, compare

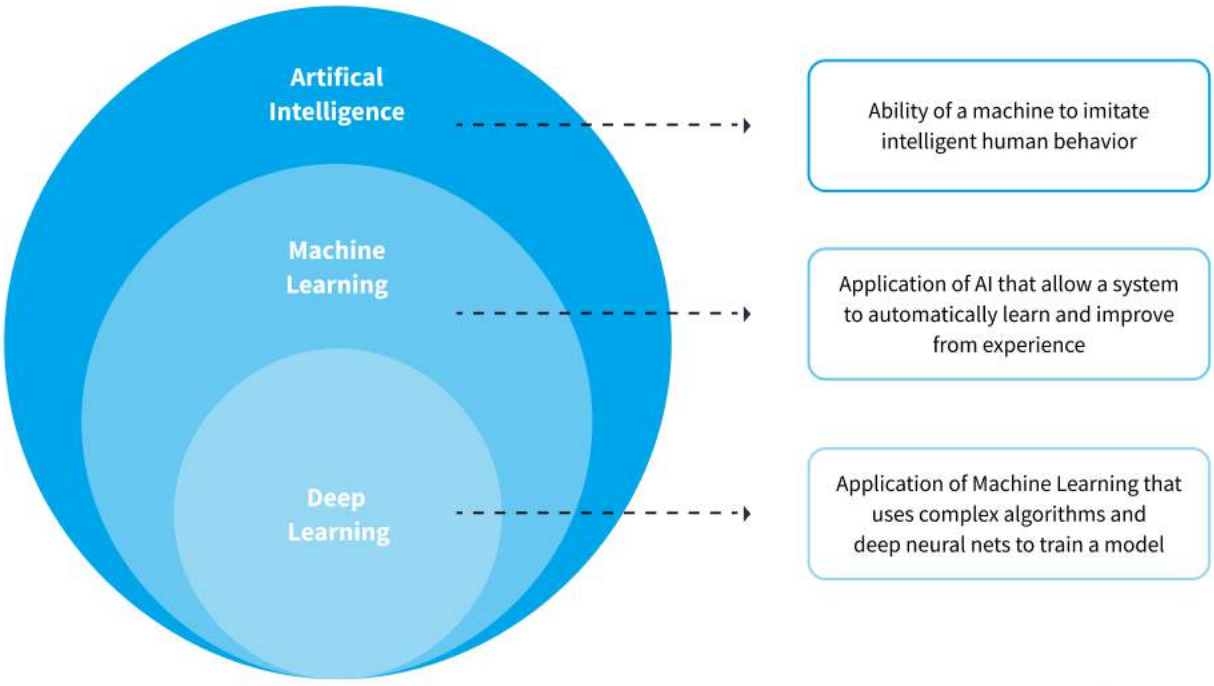
markets and discover new growth opportunities. By simultaneously analyzing historical and real-time data, BI solutions enable organizations to better assess situations, plan resources more effectively, and respond quickly to changing market conditions.

Artificial Intelligence

> *"AI is the science and engineering of making intelligent machines."*
>
> *—John McCarthy*

Artificial intelligence is the theory and development of computer systems that can perform tasks normally requiring human intelligence, automating routine tasks and delivering high efficiency.

Sophisticated AI solutions include natural language processing, computer vision and robotic process automation, which help businesses solve business problems more efficiently, increase retention and lifetime value (LTV), and increase bottom-line profits. AI creates new value for software products, accelerates internal operations, improves efficiency and provides actionable insights that ensure market dominance.



Machine Learning

"Machine learning is the field of study that gives computers the ability to learn without being explicitly programmed."

—Arthur Samuel

Machine learning allows the creation of software that brings desirable outcomes, without having to find out exactly how to achieve them. By extracting insights from data sets, artificial intelligence solutions utilize ML algorithms to unveil broad opportunities for automation, efficiency and precision improvement, and error minimization. ML solutions are able to solve such basic tasks as segmentation, recommendation engines, anomaly detection, prediction and probability, and others.

What do algorithms do?

	Machine learning	Recipe Analogy
Task	An algorithm is a step-by-step instruction set or formula for solving a problem or completing a task	1. Take the chicken out 2. Salt and season 3. Bake it
Main objective	Minimize errors or some sort of "loss function" to attain the best approach to solve a task	Minimize the number of things/ steps needed to take in order to serve the dish
Insight/ result	The algorithm learns from its mistakes/errors, finds the best approach and generates insights and rules that can be used to make predictions	Learn from your mistakes the next time you attempt the recipe

Machine learning algorithms extract valuable insights heavily focused on continuous use in dynamically changing environments and emphasize adjustments, retraining, and updating of algorithms based on previous experiences. The goal of machine learning is to constantly adapt to new data and discover new patterns or rules in it. Sometimes it can be realized without human guidance and explicit reprogramming.

03. How Data Science Delivers Business Value

Data science solutions generate additional value for products, speed up operations, increase efficiency, and highlight data insights that bring game-changing advantages.

Competitive Advantage For Clients

There are a lot of similar-looking digital products with common functionality that try to stand out with fancy wrappers and bigger marketing budgets. Applying data science solutions to an existing technology product creates a tangible difference for end-users, refines relevance, increases engagement, improves retention and lifetime value (LTV), and ultimately grows the bottom line. At the end of the day, satisfied and recurring customers are those who actually promote your brand.

nix CASE

How AI-powered chatbot for the 420,000 customer base allows for personalized healthcare patient assistance and boosts engagement

Accelerated Internal Operations

The vast majority of industries use legacy systems in their internal operations. Automating even the smallest parts of the workflow results in dramatic increases in data processing speed, expands capabilities, and ultimately reduces the cost of each operation. Data science solutions are designed to complement humans in areas where human involvement becomes a bottleneck, such as detecting rush hour traffic violations in a big city, demand forecasting, and the automation of document management.

nix CASE

RPA Enterprise Solution that empowers large enterprises to take automation to the next level, accelerate digital strategies, and transform workflows.

Powerful Analytics

Data science empowers companies with a better understanding of their business operations and better highlights the needs of customers. For example, customer analytics help better define a customer portrait, personalize an offer, give personalized recommendations, determine the best next step, and predict churn.

Marketing analytics empowered with data science enable calculation and comparison of ROI between channels, selection of the most effective sales techniques, personalization of marketing, competitor analysis, and dynamic pricing.

Intelligent Data-Driven Decisions

Humans are unable to analyze dozens of data sources in milliseconds and instantly produce a balanced decision; we need more time. Moreover, sometimes human biases impact our performance and lead to ingenuous mistakes. Data science solutions can be a helpful tool or an autonomous system that highlights critical insights and produces smart decisions instantly, increasing overall business efficiency.

nix CASE

Data engineering solution to enhance hospital performance with in-depth research and tailored reports.

Make Generated Data Work For You

Today we collect and store more data than ever—some are well-organized and valuable, and others are useless without human input. Employing data science solutions to work with huge amounts of diverse historical data creates a foundation for intelligent forecasting and optimization. Unlike traditional statistical methods, data science solutions processes have many more dependencies and provide more accurate predictions, allowing more flexibility and rapid adaptation to complex market conditions.

nix CASE

Multi-tenant SaaS platform that empowers clients with a powerful tool for health data analysis and forecasting, simplifying management through high-speed data processing and accurate output.

Measure And Keep Track Of Everything

Capture and analyze your organization's online and offline data streams in real-time via a single environment. Using data visualization and tailored dashboards, you can deconstruct complex goals, and shift employees' focus on determinants that drive tangible results.

Business Insight Mining

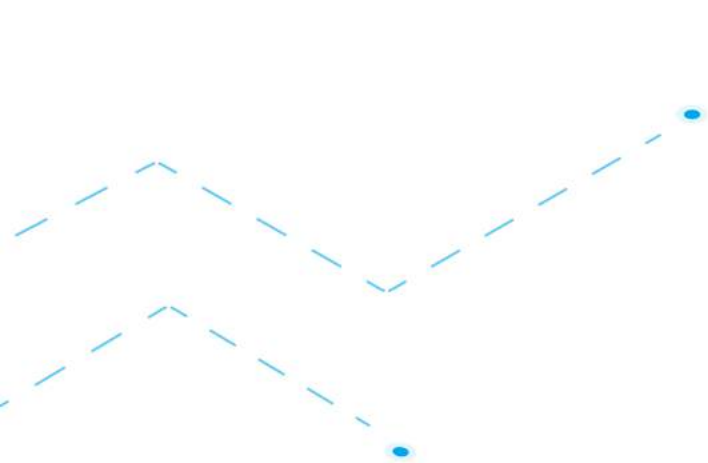
Analyze and benchmark organization operations, identify patterns, and unobvious correlations, understand reasons for your ups and downs and reveal opportunities and threats. Through comprehensive data analytics, you can gather actionable insights on customer behavior, employee performance and supplier and partner interactions.

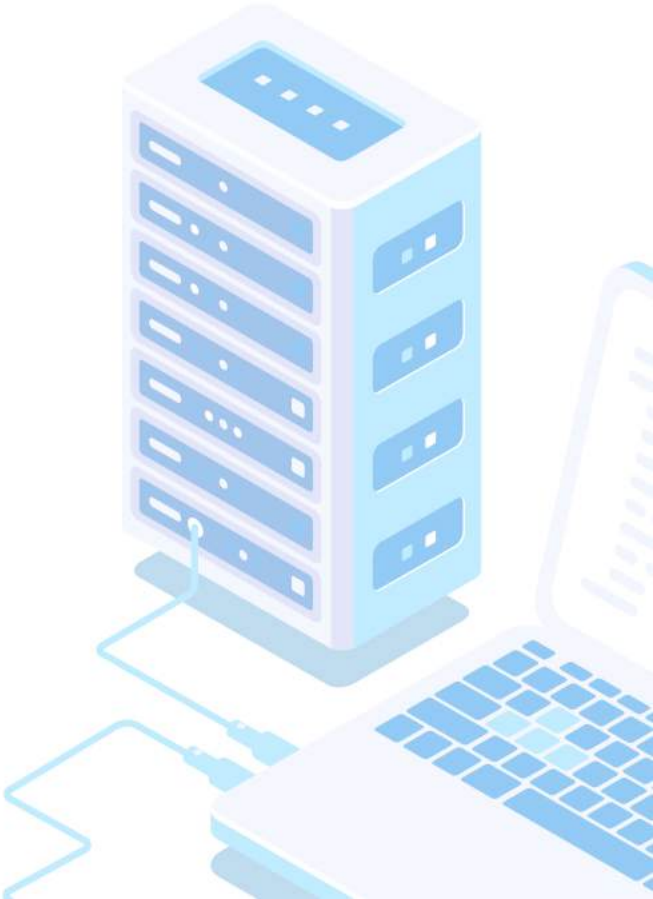
nix CASE

ML-based Platform with predictive models for healthcare insights mining that allows hospitals and insurance companies to analyze and predict the likelihood of cost efficiency of treatments compared within the industry.

Trends Capturing And Forecasting

Analyze historical and real-time data simultaneously for better estimation, resource planning, and rapid response to changing market conditions. In contrast to traditional analysis approaches, intelligent data analytics has much higher accuracy and faster response time which enriches you with up-to-date insights when relevant and applicable.





04. How Data Science Empowers Industries

Retail & eCommerce

Automate warehouse and store operations, optimize merchandising and product assortment.

• Customer Segmentation & Profile

Whether you are looking to determine new product offerings or develop a personalized marketing campaign, customer segmentation is the principal basis for allocating resources and extracting maximum value from both high- and low-profit customers.

• Customer Lifetime Value

Understanding the customer life cycle and customer lifetime value (CLV) offers benefits in budgeting, segmentation, and prioritization. CLV ultimately helps you in forecasting and improving the overall health of your company.

• Cross-Selling & Upselling

Analytical tools and techniques analyze the patterns in the customer's past behavior, correlate this information with similar customers, and then identify potential service or product opportunities at each contact point with the customer.

Personalization

Today's consumers tend to become smart buyers and prefer better value and customer experience. AI solutions fulfill clients' expectations even further by treating each customer specially and providing unique personalized experiences quickly, every time they need it—and even before they notice the need—through smart recommendations and conversational commerce.

Next Best Action

Mining historical and transactional customer data to identify common behaviors prior to a specific event (e.g., making a purchase or subscribing to a service) and predicting demand trends—based on these patterns, the predictive model determines whether behaviors match those indicating a future event and what should be offered to increase the likelihood of targeted action.

nix CASE

How Building an ETL Process into SaaS Increased Overall Performance from 1k to 1M Entries and optimized workflow management in medical organizations.

Data science helps to derive and implement suitable strategies for churn prevention. A customer churn prediction model using data on customer behavior can evaluate how likely it is for each customer to leave the product or service. Businesses can then use targeted marketing to encourage clients most at risk of discontinuing services.

Marketing

Data science solutions help deliver the right experience at the right time using the right channels and provide personalized experiences for your customers.

• Market Sizing & Opportunity Analysis

Data science solutions help to identify market potential, the most profitable client segments, market trends based on syndicated reports, secondary data, primary research studies, social media intelligence, industry experts, internal data, and other sources.

nix CASE

AWS-based BI Platform for Data Visualization and Marketing Insights that enhanced the efficiency of marketing activities and allowed the client to better understand target audiences, perform thorough data analysis, and optimize overall marketing campaigns.

• Market Entry Strategy

Applying data science in marketing helps to develop a market entry strategy more efficiently, determine the need for product modifications, and highlight changes in pricing policy. With smart data centralization across marketing channels into a single repository and aggregation of big data, businesses can extract valuable dependencies and insights.

• Customer Loyalty Analytics

Using data science solutions allows businesses to identify the most profitable target audience, and determine the necessary positioning and interjections to induce users to buy the product or service. Businesses can explore customer age, gender, location, and behavior to make highly targeted campaigns, evaluating how much customers are prone to make a purchase or leave.

Finance & Banking

> ***"Today you need to reduce the liability in terms of the losses and find the financial criminals more effectively"***
>
> *—Trace Fooshee, Senior Analyst at Aite Group, former Head of Enterprise Fraud Strategy at SunTrust Bank*

• Fraud Detection

Applied data science solutions in the financial industry help to detect any type of fraudulent transaction faster. Integrating data science into fraud analysis helps identify suspicious activity in real-time, quickly extract incomplete and inaccurate data, and provide a higher return on investment.

• Credit Risk Management

Banks rely on knowledge of customers' risk ratings and financial behavior. Data science solutions use internal data, such as information about previous loans and defaults, to determine the risk of a potential customer. This makes it much easier for companies to assess a customer's profile in need of credit.

- ## Real-Time Analytics

Real-time analytics empowered with data science in finance and banking applications allows you to track customer behavior and provide critical real-time data in the format the customer wants. It also allows for immediate changes in strategy (e.g., providing amazing deals or discounted loan opportunities) as a smart response to business system swings (e.g., declining customer loyalty or frequency of purchases). Furthermore, adding business intelligence features to your existing software provides actionable insights and enables better management of the company's product and customer experience.

- ## Customer Data Management & Personalized Services

Data science offers new ways to use customer data to deliver personalized experiences. Sophisticated data solutions help collect the right types of data to realize a shared customer vision, turn customer data into cross-departmental company knowledge through a customer data platform (CDP), and create a unified view of the customer to align channel usage and monitor performance in real-time. This allows you to anticipate customers' needs, offering real-time personalization and AI-driven communication with the customer.

Insurance

- ## Customer Segmentation

Data science solutions make it easy to segment insurance agency customers based on their financial assets, age, location, or other demographic

characteristics. Classifying customers into different groups after identifying similarities in their attitudes, preferences, behaviors or personal information allows insurance companies to develop attractive and valuable products for each group. This leads to the possibility of targeted cross-selling and launching personalized products that can be marketed effectively.

• Recommendation Engines

We've already learned that data science can help insurers develop personalized products that are more appealing to customers. A recommendation engine algorithm can identify customer preferences and choices based on their account activity and instantly recommend customized products to increase upselling and cross-selling revenue.

• Risk Assessment

Risk assessment can significantly reduce losses in insurance. An underwriter's ability to identify the risks associated with insuring a customer or asset directly impacts the business. Data science can pave the way for artificial intelligence and cognitive analytics-enabled systems that can analyze customer policy documents and determine the optimal premium and coverage amount that can be offered for a given policy. This will significantly improve the efficiency of underwriters and allow low-risk policies to be processed quickly.

• Claims Segmentation And Triage Analytics

Claims segmentation and triage analytics is the process of analyzing the complexity of each claim and assigning points to them based on their level of complexity. This dramatically helps insurers reduce the claims processing time by expediting low complexity claims and transferring more complex

claims to an appropriate specialist. This solution will also help insurers use claims adjusters efficiently.

• Lifetime Value Forecasting

Modern predictive analytics systems perform a comprehensive and in-depth examination of multiple data points to make intelligent pricing and policy decisions. Algorithms combine and process all the data to make a forecast. This allows them to predict customer behavior and attitudes, policy retention, and abandonment. In addition, CLV prediction can be valuable for developing marketing strategies because it provides insight into the customer.

Healthcare

• Data-Driven Clinical Decision Making

Predictive analytics is a powerful ally for health care providers, allowing them to gain insight into things such as treatment effectiveness weeks or even months earlier than previously possible. Today's data science solutions can greatly accelerate physicians' decision-making process, enabling them to replace an ineffective course of treatment with one that better meets the patient's needs. This is especially vital when it comes to improving outcomes for chronic diseases such as cancer and heart disease. It's also significantly more cost-effective than the traditional trial-and-error model.

nix CASE

How NIX helped leverage data science and artificial intelligence for digital transformation of assisted reproduction field.

Drug Discovery

Data science has the potential to completely redefine the methods used to find new treatments, especially when combined with AI and ML algorithms. Sophisticated data science solutions are helping to improve the speed and momentum of drug trials, enabling safer screening and the use of AI to create new drug safety protocols.

Remote Patient Monitoring

Data Science solutions enable remote patient monitoring via IoT, in the form of wearable technology and other devices. Real-time patient monitoring allows matching physicians to the most appropriate patients and providing access to care for those in remote areas. Wearable Bluetooth-enabled devices can record key patient health indicators such as blood sugar, blood pressure, and heart rate by connecting to devices such as smartphones and sending the data directly to doctors for analysis, diagnosis, and treatment.

nix CASE

IoT Telehealth solution powered with AI for diagnostics that contributes to the early detection of health problems, increase treatment quality by intelligent assistance and remote tracking capabilities, and reduce costs for patients and healthcare providers.

• Data Collection

One of the most important responsibilities of a physician is to properly collect the patient's medical history. This can often be difficult because the patient is not a specialist and does not know what data is appropriate to disclose. Using ML in health care management, healthcare providers can identify the most relevant questions they should ask the patient based on various metrics. This will help collect relevant data and, at the same time, get a prediction of the most likely conditions.

• Predictive Care For At-Risk Patients

In addition to chronic patients, other at-risk patient groups can benefit from predictive care. This is especially true for the elderly and patients recently discharged from hospitals after invasive surgery. With the benefits of telemedicine and predictive analytics, these patients can avoid adverse events or get crisis care as quickly as possible. By processing historical data, the software can even predict an elderly patient's fall, thereby saving them from possible injury and re-hospitalization.

• Supply Chain Management

The supply chain for any hospital is a complex system as supplies depend on patient load and the specifics of each case.

Using predictive analytics can help a hospital make future purchasing decisions based on accurate data. This makes purchasing more efficient and cost-effective by reducing unnecessary purchases and equipment waste.

Supply Chain & Logistics

• Data-Driven Supply Chain

Multi-layered supply chains create large amounts of data and metadata—suppliers, transportation providers and other stakeholders contribute data through countless IT systems with multiple interfaces. Added to this data are product-specific factors. The ability to store and process these vast amounts of data can unlock new information and potential.

• Predictive Maintenance

Automotive AI takes the maintenance process from preventive to predictive—instead of depending on an event or time approaching for maintenance, AI can provide actionable ideas for vehicle maintenance in real-time. By analyzing historical and contextual data, AI sensors and algorithms can offer a real-time alert system that provides condition-based vehicle maintenance requirements.

nix CASE

AI-enabled Web Application for Vehicle Monitoring and Preventing Cyberattacks that provides real-time vehicle insights to automakers and fleet owners and helps them reduce warranty costs, minimize vehicle downtime, decrease the total cost of ownership, and optimize vehicle performance.

"AI-powered hardware can visually inspect and provide superior quality control" on various products, such as machined parts, painted car bodies, textured metal surfaces and more"

—*McKinsey Global Institute.*

Real-Time Tracking

The Blockchain's immutable ledger allows real-time tracking of goods within the supply chain. Companies can queue events in the supply chain—for example, distributing goods that have recently arrived in port to different shipping containers. Sharing data and tracking the entire supply chain can increase transparency, predictability of logistics operations, and cost savings by ensuring error-free processes.

"An incorruptible record of where a diamond was sourced, its supply chain and its certification, could eradicate counterfeiting and the proliferation of blood diamonds"

—Graham Richter, managing director at Accenture Technology.

• Demand Forecasting

With the ability to integrate more data at a high level of detail, companies can use predictive and prescriptive analytics to improve the accuracy of demand forecasting. This includes refinements to find optimal levels of forecast aggregation, optimizing for different forecast horizons, or using advanced forecasting algorithms.

Education

• Learning Analytics

Data science in education helps process the vast amounts of data that are generously piled up across educational institutions and e-learning platforms. This data is so huge that a human expert could not process it in a few years, even if he wanted to. As a result, data science solutions allow us to gain insight into valuable patterns that cannot be detected using our eyes and brains alone. Such analytics can be used for a variety of purposes—to gain a more complete picture of the specifics of education, optimize key processes, and measure absolute and comparative performance.

• Predictive Analytics

Predictive analytics usually means getting a practical view of possible future events with the clear intention of preventing undesirable moments or promoting favorable ones. In order for the ML system to do this successfully, it tracks student progress, analyzes behavioral patterns, and assesses performance to help students reach their full potential.

Personalized Learning

Personalized learning remains one of the best and most valuable uses of machine learning in education right now. Before the advent of ML applications, it was nearly impossible to process and adjust instructional material for every student in a classroom unless it was private one-on-one lessons. Now there is a more targeted approach that allows students or staff —during the induction or qualification process—to go through the material at their own pace or even to choose their own course flow and preferences.

nix CASE

Powerful SDK and robust BI solution that empowers the client's product with an improved user experience and better marketing and sales performance.

Educational Chatbots

Intelligent chatbots are being used by schools to solve the problem of widespread absenteeism. For example, an artificial intelligence-driven two-way text messaging system has been developed to help children who frequently miss classes by allowing teachers to contact a student's family. It also offers personalized 24/7 support for students who are having difficulty learning.

Predicting Graduation And Dropout Rates

Higher education institutions are increasingly using data science in education and machine learning solutions to predict scenarios such as the likelihood that students will enroll, graduate, and be prepared for careers in their chosen fields of study. These capabilities also help educational institutions track student dropout rates and related demographic and educational factors to predict potential future dropouts so they can proactively intervene and allocate resources to prevent them.

05. 8 Steps Before Starting A Data Science Project

So you are ready to improve your data—where should you start?

1 Ascertain If Data Science Is What You Need

It makes sense to turn to data specialists when it is evident that the business has difficulties with data processing. This complexity can be in several planes, for example:

- The volume of data is too huge to use traditional processing methods—you need specialists who will design a solution for big data collection, analysis, and storing it in a suitable form for further use.
- The data is unstructured, has a variable format, and does not follow pre-programmed algorithms (searching for negative user reviews, recognizing familiar faces in videos, translating text into another language)—you need a unique approach that can train algorithms to understand new data without explicit instructions, perform the task at hand, and still be retrained to become more accurate over time.
- There is an enormous amount of heterogeneous data, and if you need to analyze a sales report, draw conclusions on a series of events, or understand how some events affect others, it can be too complex for

human capabilities. Data science algorithms can find similarities, trends, or patterns in the data and provide insight in an understandable, easy-to-read format.

2 Be Aware Of Business Digitalization Maturity

It's crucial to consider the capabilities of the business itself and its readiness to implement data-driven solutions. Digitization of business processes means businesses at various stages must generate relevant data to be used as material for data science solutions. This can be purchase history, customer information, a record of balances, logs of employees' actions, tracked logistics and statistics, and many others. If there are problems with this, the first step should be digitalizing the company's fundamental business processes.

3 Determine Your Business Goals

Decide what insights you need to run your business and what issues should be solved. It is necessary to involve the direct participants of the process—employees, customers, and partners—to define the objective, and describe the problems and constraints. If direct involvement is not possible, you can get feedback through questionnaires, surveys, or interviews.

The problem statement should always include the business goal the solution should strive for. In other words, why is the business willing to invest in the development, and what should be the return on investment? For example, to

increase margins by cutting costs, to improve the efficiency of operations, or to increase the ROI of marketing campaigns.

4 Set Up Data Collection

At the base of the pyramid is data collection. At this stage, the goal is to determine what data is needed, what data is available, and how data flows through the system. Is there a reliable flow/ETL process in place? Where is the data stored and how easy is it to access and analyze?

Once you start automating the data flow, you can add additional components to complete the picture. Consider the tools available to collect, manipulate, analyze, and provide the necessary information. Without data, no machine learning or AI solution can learn or predict outcomes.

5 Define KPIs

At the base of the pyramid is data collection. At this stage, the goal is to determine what data is neededR what data is available, and how data flows through the system. Is there a reliable flow/ETL process in place? Where is the data stored and how easy is it to access and analyze?

Your product's goals should inform your product success metrics based on the end user's goals. Also, you can divide success metrics into distinct time periods according to when you can and should be measuring them.

6 Perform Risk Assessment

When deciding to implement data science solutions, it's better to foresee challenges you may face—which may include finding the right talent or solving basic issues revolving around getting the raw data organized, unknown security vulnerabilities, and financial risks. Heterogeneous sources often make it difficult for data scientists to understand and gather meaningful insights.

The risk of inability to identify and manage data can lead some managers to delay the adoption of the techniques and thus prevent them from realizing their full potential.

7 Define Your Toolset

At this stage, it's essential to decide whether you need a ready-made solution or custom-tailored to your specific business goals.

When it is better to buy:

- The out-of-the-box software fits your business process perfectly
- Your business is sustainable and processes change very rarely or never
- You have tight deadlines and need a ready-made solution as soon as possible, thus can't wait for a custom one

When it is better to develop from scratch:

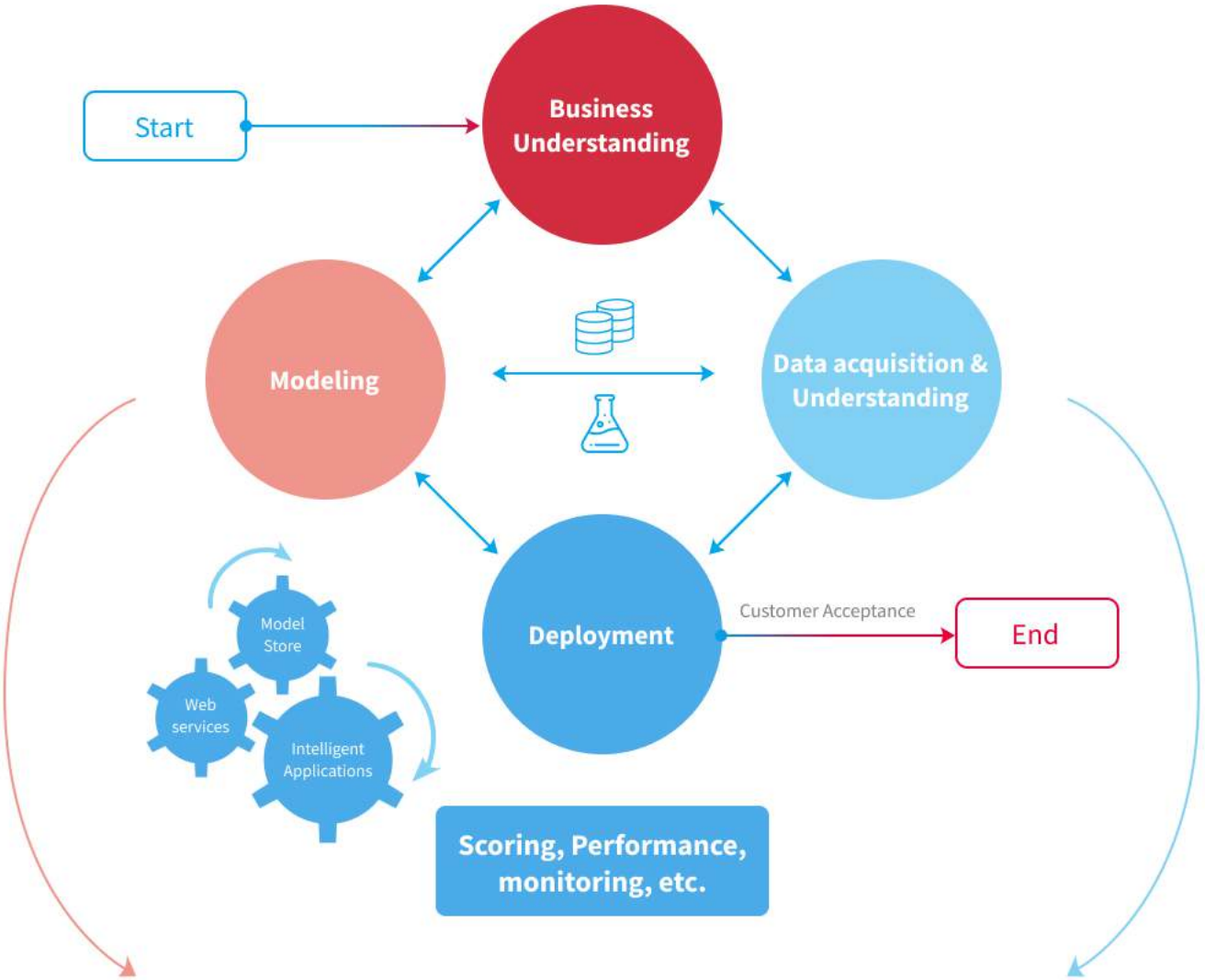
- The unprofitability of an off-the-shelf solution, especially in the long term (e.g the more users, the more requests, the more data, the more expensive)
- The out-of-the-box software does not fit perfectly into the business process and creates issues either with fine-tuning the software or adapting business processes
- The off-the-shelf solution limits the further development of the company due to a lack of functionality

8 Assemble The Team

As a rule, a team is formed based on the project objectives, when there is a vision of what should be done, required skills and roles. This kind of work requires experience in project management, team management or product development. If you don't have such a specialist in your in-house team, you will need to turn to an outsourcing provider who will help you to shape the project vision and offer a solution with the appropriate set of technologies, tools, necessary team, and deadlines.

This information should be enough to calculate the project budget and assess its rationality, taking into account the planned KPI, payback time, and technical risks. Once the decision to launch the project has been made, you can move on to the next stage of data collection and preparation for analysis.

Data Science Lifestyle



Transform, Binning Temporal, Text, Image Feature Selectiont	**Feature Engineering**	**Data Source**	On-premises vs Cloud Database vs Files
Algorithms, Ensemble Parameter Tuning Retraining Model management	**Model Training**	**Pipeline**	Streaming vs Batch Low vs High Frequency
Cross Validation Model Reporting A/B Testing	**Model Evaluation**	**Environment**	On-premises vs Cloud Database vs Data Lake vs Small vs Medium vs Big Data
		Wrangling, Exploration & Cleaning	Structured vs Unstructured Data Validation and Cleanup Visualization

06. How To Choose A Data Science Vendor Wisely

Proven Expertise

• Domain Experience

In addition to experience in software development, the vendor should be familiar with your business. Domain experience is not always a critical factor, but it can save you some effort—from understanding specific terms and concepts to knowing common issues, solutions, bottlenecks, and workarounds. This type of experience will help to avoid shortcomings and wasted time.

• Relevant Cases In The Portfolio

If a vendor states expertise in a certain field, there are likely case studies to back this up. You can use them to study similar experiences and even get ideas for your project.

Client Testimonials

• References

If you need to check the vendor's background, you can ask for references from their clients—preferably in the exact location and business domain.

This way, you can learn about the general approaches to operations, meeting deadlines, and how effectively the work process is built.

• Call The Experts

Suppose your provider can't share relevant experience in the form of case studies. In that case, a call with an expert who has hands-on experience in your industry is an excellent opportunity to find out how deeply they understand your business, what problems they encountered, and what solutions were provided.

Offshore Vs Nearshore

• Time Zone

If the time zone difference is more than 4-5 hours, that's a reason to discuss whether the vendor has a well-established process for operating in different time zones. What percentage of customers are in your or similar time zones? How do they monitor the efficiency of communication and operations?

• Distributed Teams

The vendor should have a practice of constantly working with a distributed team where all processes are fine-tuned and well-managed. It would be best if the team is within a couple of time zones of each other. Otherwise, their time difference will be superimposed on the difference with you, and here even an experienced project manager will find it challenging to effectively organize the development process.

On-Site Visits

For some projects, yes, it's essential. Especially if you're looking to automate the business processes of your enterprise, where understanding production and learning all the operations on site is key. In that case, it's worth arranging a visit for the key team members to understand the process from the inside and get immersed in the specifics of the business faster. But for most projects, this is not necessary, given the COVID pandemic and the plentitude of tools for remote collaboration.

Price Formation

Location Affects Price

The company's expenses are included in the cost of services. Accordingly, a client pays for a stylish office in a prestigious area.

The More Middlemen, The More Expensive

Each task requires specific skills and a number of people. Freelancers or small agencies can usually cope with small projects, but when it comes to complex projects that require different areas and technologies, you need to assemble a large team of specialists with relevant expertise. However, it is also worth separating companies that hire people or subcontractors for you and companies that already have on board a large team of specialists with diverse skills because, in the first case, you will overpay for intermediaries and increase the final cost of the project.

Right Proposal Preparation Process

Deep Dive Analysis First

An experienced vendor always starts with an in-depth analysis of the project to get a clear understanding of the task, goals, and scope of work, and offer a comprehensive solution, rather than rushing to close the deal with a raw estimate.

For this purpose, the vendor conducts pre-project analysis and risk assessment involving experts with knowledge in the field and technology. In most cases, such analysis is free of charge, and its result is a well-elaborated roadmap, project estimated budget, tech stack, and team composition.

Complex Projects Require More Time

The more complex the project, the more time is needed to analyze the task and prepare a proposal. Generally, it is 1-2 weeks with a couple of calls, including a meeting with the expert team. The vendor who prepares a proposal the day after the call is likely to be done hastily, without immersion in the problem and elaboration of solutions, and therefore with questionable accuracy.

The Right Proposal

The prepared proposal should contain not only the project cost estimation, rates, and team composition, but also a vision of the proposed solution, architecture design, planned tech stack, and limitations.

07. How NIX United Provides Data Science Services

End-To-End Data Science Software Development

We leverage our deep knowledge of data science and take on all aspects of the data lifecycle— from the requirements phase to the launch and maintenance of a market-ready solution.

Extension Of Your Data Science Team

We provide a top-notch team assembled explicitly according to your business needs to enhance you with capabilities that you lack, accelerating the software development process and reducing time to market for standalone data science solutions.

Data Science Consulting

Our data science experts with deep technical and business knowledge dive deep into your ideas, prove them feasible, and help you find the most appropriate solution for your specific business needs.

Read more about our data science services

We hope this document helps you gain a basic understanding of what data science is, how it can benefit different industries, where to start and how to choose data science vendor wisely.
In the meantime, we offer you to know more about us—click links to review how our engineers created various software solutions for different industries. Possibly, the right data science vendor that you were looking for is here.

- **Case Studies**

- **They Trust Us**

- **Recognitions & Awards**

Contact our experts

About NIX

NIX is a global software development company that delivers services to leading companies from different business verticals, dynamically contributing to their technological growth. Since 1994, we create software solutions that boost companies, transforming business capabilities into tangible advantages and measurable values and efficiently addressing present-day business challenges across technologies, data, and processes.

Where technology meets ingenuity!

+1 813 374 0027

ask@nix-united.com

nix-united.com