

White Paper

# Driving Better Business Results with As-a-Service Machine Learning

JEAN S. BOZMAN

**Cloud** ARCHITECTS

 **SambaNova**  
SYSTEMS

# Introduction

Businesses need to deal with rapidly growing data in their central locations – and at remote sites. They want to harvest the patterns in their data by applying artificial intelligence (AI) and machine learning (ML) to solve complex challenges. But that can only be done if they have the expertise in machine learning and data modeling to discover the actionable insights that could drive their business to the next level.

In today’s rapidly changing business conditions, many companies are looking for the competitive advantage that efficient technology can bring. These companies include organizations in the healthcare, life sciences, financial services, retail, logistics, manufacturing, and energy industries.

Data is growing 42% per year, according to IDC. The rapid growth is so fast, and the amount of data is so great, that it costs customers valuable time and money just to prepare the data for fast and accurate processing.



Large amounts of data, in the petabyte and exabyte range, can be found outside the core data center, stretching to the hybrid cloud and even farther to the Edge — where the data originates for many business processes. Edge locations include car manufacturing plants, offshore rigs exploring oil and gas reserves, and pharmaceutical research labs designing next-generation drugs to fight cancer.

**“ In today’s rapidly changing business conditions, many companies are looking for the competitive advantage that efficient technology can bring. These companies include organizations in the healthcare, life sciences, financial services, retail, logistics, manufacturing, and energy industries.”**

# Data is Driving the Next Wave of Business Success

Many organizations know that the ability to achieve accurate, efficient, and actionable insights is directly linked to their future business success. However, most businesses quickly realize that they simply don't have enough in-house expertise to take full advantage of advanced AI and ML.

The expense and complexity of launching a major machine learning initiative can often prevent these projects from achieving success. As a result, many ML projects fail due to a lack of data modeling expertise.

There must be another way to gain the business benefits of ML without hiring dozens, or hundreds, of AI programmers and ML experts. Building large data models, then refining them and maintaining them repeatedly over time, is time-consuming and labor-intensive — often taking many months to complete.

What if the power and efficiency of a highly tuned and scalable ML engine could be brought to your business as a data service? That would be a game-changer.

**“ Many organizations know that the ability to achieve accurate, efficient, and actionable insights is directly linked to their future business success.**

**However, most businesses quickly realize that they simply don't have enough in-house expertise to take full advantage of advanced AI and ML.”**

## The SambaNova Solution

SambaNova Systems Inc. of Palo Alto, California, has a new approach to accelerating and delivering AI and ML workflows to speed customers' time-to-results. SambaNova's Dataflow-as-a-Service™ connects enterprises with powerful computing services in a subscription-based, pay-as-you-go model. This means that the work of configuring, managing, and maintaining the AI/ML system is done by SambaNova — allowing its customers to focus on the data results that drive their business.

The company was founded by Stanford University computer scientists who designed a new type of system that is entirely focused on making AI and ML technologies more powerful to solve a wide range of business challenges. They designed a new kind of processor that allows computations to scale up across many processors in a seamless way, so that system resources keep pace with the amount of computing that must be done. The company is well-funded; funders include SoftBank, BlackRock, Intel Capital, GV (formerly known as Google Ventures), and more.

With SambaNova's Dataflow-as-a-Service, services are scaled up without any need for reconfiguring or slowing in-house data modeling projects. Dataflow-as-a-Service seamlessly adds more system resources to scale up the processing resources powering each data model — and delivers the results as a service to its customers.

**“ With SambaNova’s Dataflow-as-a-Service, services are scaled up without any need for reconfiguring or slowing in-house data modeling projects. Dataflow-as-a-Service seamlessly adds more system resources to scale up the processing resources (that are) powering each data model — and delivers the results as a service to its customers.”**

## Data Services Delivered in a Subscription Model

SambaNova's Dataflow-as-a-Service is a subscription-based service that delivers pre-trained and pre-optimized ML models trained to accuracy directly to the business. This saves costs related to model building, training, and maintenance — in sharp contrast to assembling clusters of GPUs and CPUs to build up data models for unique hardware configurations.

The service is focused on delivering results directly to customers. It connects enterprises with scalable ML models by leveraging a subscription model. With this approach, SambaNova's experts take on the work of configuring, managing, and maintaining compute infrastructure and clusters.

SambaNova's sophisticated use of ML technology means that its customers do not have to hire master data modelers to "tune" their growing data models. Instead, the expertise needed to refine data models is included in the Dataflow-as-a-Service offering.

**“ Services deliver results directly to customers. (Dataflow-as-a-Service) connects enterprises with scalable ML models by leveraging a subscription model. With this approach, SambaNova's experts take on the work of configuring, managing, and maintaining the compute infrastructure and clusters.”**

# SambaNova's Dataflow-as-a-Service

SambaNova answers the questions many customers have about using AI and ML to gain competitive advantage by analyzing the data generated by their business.

- **How can I put AI and ML to work to address my business problems?**
- **How can I accelerate the business transformation to meet pandemic-related business conditions?**
- **What if my IT personnel don't have advanced ML and data modeling skills?**
- **What if I don't have enough data scientists to build and optimize complex deep-learning models?**

**Let's take a closer look at how this happens:**

Customers pay through a subscription agreement on a pay-as-you-go basis. With Dataflow-as-a-Service, organizations have the option to deploy the service inside the data center and behind the corporate firewall, or remotely, because it is fully managed and maintained by SambaNova.



## Leveraging the As-a-Service Subscription Model

Responding to the pandemic crisis and economic dislocations worldwide, the world looks to gain deep insights, powered by ML, to inform business strategies in fast-changing markets.

Businesses are increasingly looking to AI to discover, to manage, to protect and to secure vast data assets spanning their organization, from Core to Cloud to Edge locations.

SambaNova's deep research into AI and ML, and the world of scalable cloud computing, has resulted in innovations that provide real business value, including:

- Modeling expertise that guides the data-optimization process.
- Powerful AI/ML software that navigates through petabytes (PB) of data.
- Data service that delivers results directly to a company's business units.

SambaNova delivers AI/ML services directly to its business customers, allowing them to reach business results much faster than before.

## Bringing ML Expertise to Business Challenges

Core attributes of SambaNova's AI/ML computing service include:

- **Workload optimization.** SambaNova Systems technology optimizes AI/ML workloads by design, leveraging a deep reservoir of ML expertise.
- **Support for large datasets.** Fast and accurate training of extremely large datasets.
- **Access to large memory.** SambaNova systems have large memory surrounding the processors, so analytics do not run out of "runway" as large datasets are processed.
- **Security and compliance.** Systems are remotely managed by SambaNova experts or installed on-site to comply with security or regulatory policies.
- **Predictable pricing model.** Fixed-rate pricing simplifies budgeting for complex and scalable data modeling and ML projects aimed at finding actionable business results.

SambaNova Systems' onboard analytics can be accessed as pure-play ML services without installing any hardware systems. Because it is based on a subscription model, SambaNova's highly scalable AI/ML engine is available on a pay-as-you-go basis, providing business value at less cost than many hardware-based AI/ML solutions.

By choosing the Dataflow-as-a-Service subscription, customers pay for the resources they use rather than acquiring and managing a scalable SambaNova DataScale® hardware system on their own. Now, with Dataflow-as-a-Service, organizations have the flexibility to deploy in their location of choice, whether that is inside the datacenter and behind the corporate firewall or fully managed and maintained by SambaNova.

## Three Key Areas of Focus for Machine Learning

SambaNova's fully integrated hardware and software stack is aligned to address three of the most demanding uses for ML technology in the world: natural language processing (NLP), high-resolution imaging, and recommendation. All three of these workloads listed below are needed by today's businesses and organizations, and all of them demand highly scalable hardware to host large-scale ML models.

- **Natural Language Processing.** NLP uses AI to translate, process, understand, and generate voice, text, and documents. NLP enables autonomous multilingual voice assistants for customer service, real-time credit checking and application processing, and for text extraction from volumes of research for document summarization.
- **True Resolution Computer Vision.** Images range in quality from simple smartphone photos and videos – now at the 4K level – moving to 8K for professional uses. Applications include healthcare, medical images, media/entertainment, remote sensing, and video surveillance. Medical images can be viewed by software, informed by AI algorithms, to rapidly identify medical conditions for follow-up treatment.
- **Recommendation Systems.** Familiar from their use in the world of online commerce and social media sentiment analysis, recommendation engines analyze a pattern of use by consumers. Recommendation engines are software-based systems that generate lists of recommendations for similar products or services (e.g., online retail websites, and media and entertainment websites for film, TV, and books).

When working with these three important use cases, customers using other ML technologies have said they would rather avoid launching too many data “round-trips” or “hops”— from processor to memory and back again. But that’s exactly what other ML approaches — many of them deployed on GPU-based clusters — require for large, extensible ML workloads.

There is another way to reach ML insights — and to do so more quickly.

When analyzing petabytes (PB) or exabytes (EB) of data, SambaNova’s Dataflow-as-a-Service taps a scalable set of system resources, supplying the processing power needed for large data models. Using the SambaNova service, the hardware components of the solution look like a unified resource rather than a collection of individual chips that must be clustered together before processing completes.

Building on a strong foundation of its DataScale hardware systems, SambaNova’s Dataflow-as-a-Service delivers powerful ML directly to the business customers and managers who pay for it. The AI/ML tasks remain within the scalable computer system’s architecture.

## Finding Patterns in the Data at the Speed of Business

Double-digit worldwide growth rates for the worldwide AI market show that organizations view AI and ML technologies as important tools for competitive advantage. Very often, AI and ML technologies are used to build data models that generate powerful deep learning results based on customers’ extensive data resources.

According to IDC’s semi-annual AI Intelligence Tracker, revenues for AI – including software, hardware, and services – are forecast to grow 16.4% year over year in 2021 to \$327.5 billion (USD) worldwide. By 2024, the worldwide market is expected to reach \$554.3 billion (USD) worldwide, with a five-year compound annual growth rate (CAGR) of 17.5%.

**Source: IDC Press Release, IDC Forecasts Improved Growth for Global AI Market in 2021, February, 2021.**

By applying AI and ML technology, enterprises are viewing the underlying dynamics of their business, revealing changes they can make in their business models, and in their products. Virtuous cycles start with growth in one area of the business that then sparks more growth in other areas.

These virtuous cycles allow the overall business to accelerate its revenue growth and its profits. In many cases, these cycles are interrupted or blocked by incomplete or faulty analyses of the data across an enterprise. Often, the data is trapped inside data “silos” so that the overall analysis is skewed, or biased, or incorrect. The ability to view larger datasets, and to process them quickly, improves the business results of timely analytics.

## SambaNova's Value-Add

- **SambaNova's Modeling Expertise is Built into the Service.** Achieving business goals may require ML and engineering skill sets that can be costly and hard to find. Enterprises and organizations may not have enough deep-learning engineering skill-sets in-house to address challenging ML projects. Many are finding that hiring personnel with deep data expertise and advanced ML skills is expensive and that any projects must be built from the ground up, taking many months to prepare.
- **Reduce Systems Costs for Hardware and Software.** Many organizations lack the processing power to tackle the increasingly steep mountains of data they are gathering from ongoing operations day after day. Or they cannot afford to finance the expensive supercomputers or clustered GPUs that could tackle the most demanding AI/ML tasks.
- **Achieve Faster Results.** Building and integrating infrastructure for AI/ML processing can be costly and difficult to configure. AI and ML services give customers the ability to tap the processing power of systems that are pre-defined to run deep learning workloads. This saves customers valuable time — and money. Automation of repetitive tasks and efficient management will speed processing—and reduce costs associated with AI/ML analysis of customers' production data.

# Machine Learning is Transforming the World's Businesses

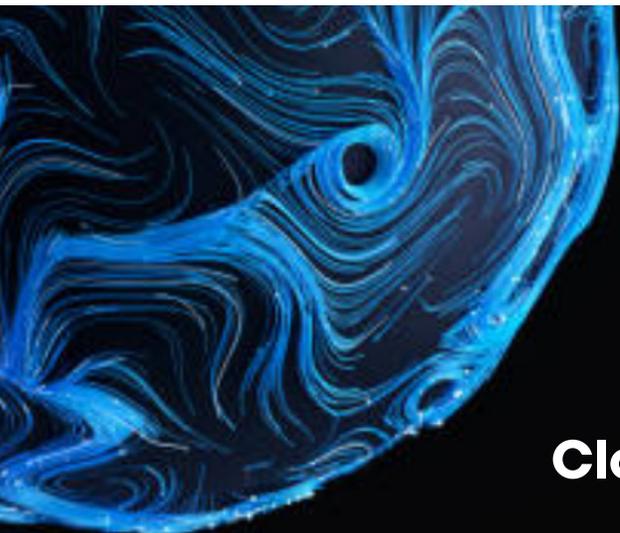
Technology is changing in significant ways: processing power is becoming progressively less expensive, over time — and ML software is becoming progressively more powerful. And what about that lack of ML skill sets? That is being addressed by subscription-based services that allow customers to immediately gain the knowledge of AI/ML experts who can configure, deploy, and manage data models.

Revealing the patterns that are contained inside an organization's production data is essential to recovery from the twin crises of COVID-19 and associated business dislocations worldwide. Mapping products and starting services that meet emerging consumer requirements are vital to building business volume back to pre-pandemic levels — or better.

Analyzing the economic fallout from the pandemic is difficult. Some organizations are hiring, while others have been forced to reduce staff during the pandemic. Some geographic areas are returning to near-normal business conditions, while others are experiencing lockdowns and severely impacted business conditions. Some are seeing higher demand for their products and services, while others are seeing a fraction of the pre-pandemic demand (e.g., hospitality, airline and railroad travel, and restaurants).

This time of great change is the right time to apply powerful ML services to solve the complex challenges ahead. SambaNova can help you accelerate your business.

**To learn more, visit SambaNova Systems on the web at [SambaNova.ai](https://SambaNova.ai).**



**Cloud** ARCHITECTS

 **SambaNova**<sup>®</sup>  
S Y S T E M S