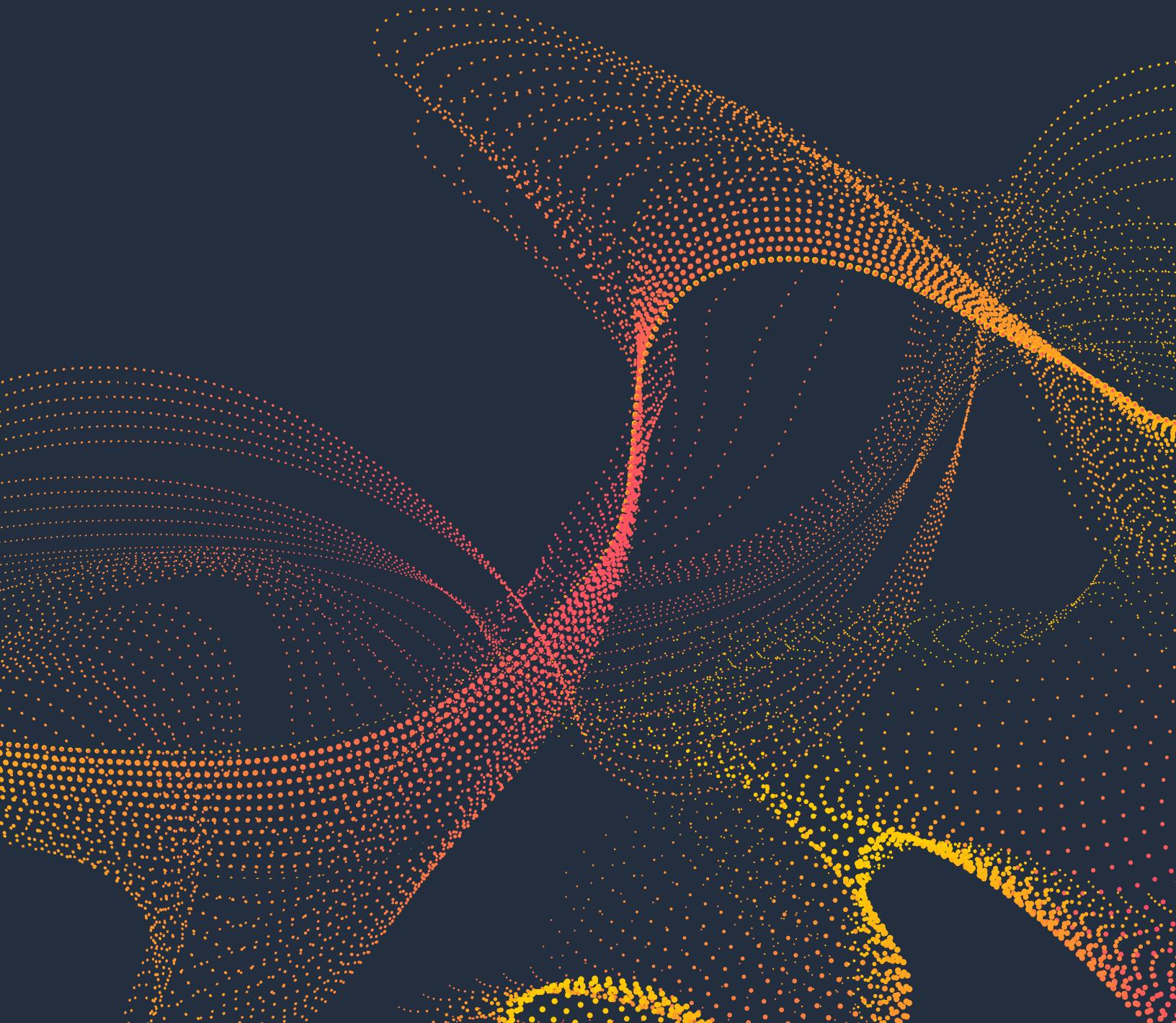


The total cost of ownership (TCO) of Amazon SageMaker



Amazon SageMaker TCO

[Amazon SageMaker](#) is a modular, end-to-end machine learning (ML) service to build, train, and deploy models at scale and at significantly lower costs compared to other ML options. The total cost of ownership (TCO) of Amazon SageMaker over a three year horizon is over 54% lower compared to other cloud-based ML options such as self-managed [Amazon EC2](#) (EC2) options and AWS managed [Elastic Kubernetes Service](#) (EKS) on AWS. In addition to the lower TCO, Amazon SageMaker's fully managed and integrated features let you put ML ideas into production faster and improve data scientist productivity by up to 10 times.



Coinbase uses ML models on Amazon SageMaker to help with fraud prevention, identity verification, and large-scale compliance. **Using Amazon SageMaker reduced model training times from 20 hours to 10 minutes.**



Intuit developed ML models that can pull a year's worth of bank transactions to find deductible business expenses for customers. **Using Amazon SageMaker, Intuit reduced ML deployment time by 90%, from 6 months to 1 week.**

Teams of all sizes benefit from significantly lower TCO when using Amazon SageMaker. For example, over 3 years, a small team of 5 data scientists can expect a 96% lower TCO using Amazon SageMaker versus building and maintaining their own ML platform on EC2 or EKS. Medium-sized teams of 15 data scientists can expect 90% lower TCO with Amazon SageMaker. Large teams of 50 data scientists can expect 78% lower TCO with Amazon SageMaker compared to EC2, and 65% lower TCO compared to EKS on AWS. Even larger teams of 250 data scientists can expect a 72% lower TCO with Amazon SageMaker compared to EC2, and 54% lower TCO compared to EKS on AWS.

Teams Using Amazon SageMaker			
OVERALL SUMMARY		Amazon SageMaker 3 year TCO Savings	
		Compared to EC2	Compared to EKS
Small Scenario	5 Data Scientists	-96%	-96%
Medium Scenario	15 Data Scientists	-91%	-90%
Large Scenario	50 Data Scientists	-78%	-65%
Extra-Large Scenario	250 Data Scientists	-72%	-54%



Calculating the TCO for Amazon SageMaker included evaluating costs for each of the three ML phases:

Build

Explore, preprocess data, experiment with ML frameworks and algorithms using notebooks

Train

Train ML models and tune model hyperparameters on large datasets

Deploy

Put ML models into production and make inferences on unseen data

Each of these phases included the cost of running infrastructure (compute, storage, and network), operational costs, and costs of building security and compliance. Compute, storage, and network infrastructure costs are measured in terms of usage. The operational costs for operating, monitoring, and maintaining the infrastructure, are measured in terms of the salary of engineers. ML workloads are inherently data-dependent across the entire lifecycle. Organizations incur costs to secure ML workloads, achieve compliance with regulatory standards, and to maintain security and compliance. The security and compliance costs are measured as the cost of engineering resources and their salaries.

Typically, the TCO for Amazon SageMaker is lower in the first year compared to EC2 or EKS options due to the higher investments in building required security and compliance features (which come out-of-the-box on Amazon SageMaker) but need to be built on these other platforms. The TCO for Amazon SageMaker continues to remain significantly lower than other platforms in subsequent years. This is due to the much greater utilization Amazon SageMaker achieves compared to EC2 or EKS, and the ongoing costs of maintaining the core infrastructure, security, and compliance when choosing self-managed EC2 or managed EKS options—all of which Amazon SageMaker provides out-of-the-box.

Services considered for the TCO analysis



The TCO analysis in this post has considered the following options:

1

Self-managed on Amazon EC2 – This option gives you a DIY approach to ML using AWS services such as [AWS Batch](#) and [Amazon ECS](#). You take on the responsibility of provisioning and managing Amazon EC2 instances, including instance failure recovery, patching, and autoscaling. You use the DLAMI with the ML frameworks and libraries pre-built but need to optimize data access to get high throughput, optimize your setup for scale, and enable distributed training. In addition, you need to build and maintain the required security and compliance features for your ML workloads.

2

Managed Kubernetes on AWS – Services such as [Amazon EKS](#) makes it easy to deploy, manage, and scale containerized workloads on Amazon EC2. However, you need to take on the additional cost overhead of managing your own cluster and tuning the performance and utilization based on the memory, compute, and network requirements for your workloads. In addition, you need to build the right level of security, compliance, and availability for your ML workloads. While there are open-source tools such as Kubeflow that make running ML workloads easier, you still incur infrastructure management costs because the expertise required to build, manage, and secure a Kubernetes cluster is much higher and less readily available.

3

Fully managed ML service – [Amazon SageMaker](#) fully managed so you don't need to build, manage, or maintain any infrastructure or tooling to build, train, or deploy ML models. You don't have to manage any instances, patching, or upgrades, or build autoscaling or load-balancing capabilities to scale your workloads. This reduces your required operational support to less than one-tenth compared to self-managed solutions. You can select compute, memory, or graphics optimized instances and get high instance utilization. In addition, you pay for storage and network based on your usage. Amazon SageMaker has built-in security and compliance for ML workloads and does not require you to invest in additional security for your workloads.

Security and compliance costs



ML workloads rely on large volumes of data, much of it sensitive or proprietary, to build and train models. When the models are in production, the ML models need to secure incoming inference requests as well as responses.

Though the underlying platforms for EC2 and EKS enable security and compliance, you still need to build the security and compliance for your ML setup, which incurs additional engineering costs when you use self-managed EC2 or EKS solutions. This is because you need to secure data, control access, and maintain compliance when you use notebooks, train across multiple nodes, and run inferences on customer data. In contrast, Amazon SageMaker is designed to run ML applications at scale with best-in-class security, compliance, and world-wide availability so you don't need to build these capabilities.

Security costs

Amazon SageMaker is fully secure with end-to-end encryption at rest and in transit, including encryption of the root volume and EBS volume, [Amazon VPC](#) support, [AWS Private Link](#), customer managed keys, [IAM](#) fine-grained access control, [AWS CloudTrail](#) audits, inter-node encryption for training, tag-based access control, network isolation, and Interactive Application Proxy. Amazon SageMaker saves large and extra-large businesses more than 89 development months of engineering effort equivalent to \$926K in costs over a 3 year period by providing the security features out-of-the-box as compared to self-managed solutions on EC2 and EKS, in which you need to build these capabilities yourself. For modeling purposes, it is assumed that the security requirements for small and medium businesses are lower, needing approximately 25% and 50% of the features provided by Amazon SageMaker.

Compliance costs

While EC2 and EKS are secure and compliant, Amazon SageMaker is compliant with HIPAA, PCI, SOC, GDPR, ISO, and FEDRAMP for ML workloads. Amazon SageMaker also supports FIPS endpoints. While most companies have a central audit and compliance team that lays down the requirements for compliance and manages the compliance process, there is additional work for the service owner to prepare compliance artifacts, build service level tooling to demonstrate compliance, and maintain the compliance on an ongoing basis. With Amazon SageMaker, this is provided out-of-the-box at no additional cost. Amazon SageMaker saves large and extra-large businesses 60 development months of engineering effort for compliance equivalent to \$627K in costs over a 3 year period, as compared to building these capabilities yourself in EC2 or EKS. For modeling purposes, it is assumed that the compliance requirements for small and medium business are lower, needing approximately 25% and 50% of these costs, respectively.

Scenarios considered

While every business is unique, this post assumes four scenarios to determine the TCO.

Scenario	Small	Medium	Large	X-Large
Operating in the AWS cloud	Yes	Yes	Yes	Yes
Number of data scientists in the company	5	15	50	250
Yearly growth of ML workloads	15%	10%	5%	2.5%
Regions the company operates in	Multiple	Multiple	Multiple	Multiple
Mix of workloads: Classic ML (such as Scikit-Learn or XGBoost)	95%	85%	50%	50%
Mix of workloads: Deep learning (such as PyTorch or TensorFlow)	5%	15%	50%	50%
Duration of training job per data scientist (hours)	0.1	0.5	2.5	5
Total duration of training jobs per business day (hours)	0.5	7.5	250	1250
Total number of models deployed in production	2	10	50	150
S3 storage (GB/month)	25	250	5,000	25,000
EBS storage (GB/month)	15	150	1,000	5,000
Data transferred over network (GB/month)	1	10	500	2,500
Security and compliance requirements	Low (25%)	Med (50%)	High (100%)	High (100%)
Ratio of engineers to data scientists (self-managed)	1:10	1:20	1:30	1:40
Ratio of engineers to data scientists (Amazon SageMaker)	One tenth of the self-managed options			
Annual salary for full-time engineer	\$125K	\$125K	\$125K	\$125K

Scenario 1: Small-sized business

The TCO of Amazon SageMaker for a small sized business with 5 data scientists is 96% lower than EC2 and 96% lower than EKS. The TCO savings are 97% in Year 1, 94% in Year 2, and 94% in Year 3. The following table presents a summary of the analysis.

3 year TCO Analysis					
Small Scenario	Amazon SageMaker	DIY options on Amazon EC2	Managed K8s on AWS	SM TCO v/s DIY EC2	SM TCO v/s Managed K8s
TOTAL COSTS (Build + Train + Deploy)	\$54,831	\$1,322,754	\$1,284,710	-96%	-96%
Total Build Costs	\$16,124	\$352,814	\$352,814	-95%	-95%
Total Infrastructure Costs	\$8,889	\$6,356	\$6,356	40%	40%
Operations Costs	\$7,234	\$217,031	\$217,031	-97%	-97%
Total Security & Compliance Costs	\$ -	\$129,427	\$129,427	-100%	-100%
Total Train Costs	\$14,681	\$575,602	\$573,446	-97%	-97%
Total Infrastructure Costs	\$7,447	\$229,144	\$226,988	-97%	-97%
Operations Costs	\$7,234	\$217,031	\$217,031	-97%	-97%
Total Security & Compliance Costs	\$ -	\$129,427	\$129,427	-100%	-100%
Total Deploy Costs	\$24,026	\$394,337	\$358,450	-94%	-93%
Total Infrastructure Costs	\$16,792	\$47,879	\$11,991	-65%	40%
Operations Costs	\$7,234	\$217,031	\$217,031	-97%	-97%
Security & Compliance Costs	\$ -	\$129,427	\$129,427	-100%	-100%

The following are the build details:

- There are 5 data scientists on the team, each running 1 notebook instance of T3.medium.
- The cost per instance per hour \$0.0582 in Amazon SageMaker and \$0.0416 in EC2 and EKS.



The following are the training details:

- 0.1 hours of training per data scientist per business day. Training instances are a mix of M5.xlarge and P3.2xlarge based on the mix of classic ML and deep learning workloads for the scenario.
- Amazon SageMaker does not require you to manage any clusters, and you only pay for the duration for which the training jobs are running. For EC2, each data scientist runs a dedicated instance. With EKS, you can share 1 instance across 4 data scientists, though the cluster needs to run continuously.
- The blended cost per instance per hour is \$0.87 in Amazon SageMaker and \$0.62 in EC2 and EKS.

The following are the deployment details:

- 2 ML models deployed in production. Each model deployed runs on 2 instances. Instances are a mix of M5.xlarge and P3.2xlarge based on the mix of classic ML and deep learning workloads.
- You can deploy 4 models on 1 instance with Amazon SageMaker multi-model endpoints, and 4 models on 1 instance with EKS. Each model deployed in EC2 requires dedicated instances.
- The blended cost per instance per hour is \$0.87 in Amazon SageMaker and \$0.62 in EC2 and EKS.

The following details are applicable to all phases:

- Security and compliance costs are approximately 25% of the costs incurred by large businesses. This number is lower because small businesses have lower security and compliance requirements.
- S3 storage costs are \$0.023 per GB per month.
- EBS volume costs are \$0.14 in Amazon SageMaker and \$0.10 in EC2 and EKS.
- Network costs are \$0.016 in Amazon SageMaker and \$0.001 in EC2 and EKS.

Scenario 2: Medium sized business

The TCO of Amazon SageMaker for a medium-sized business with 15 data scientists is 91% lower than EC2 and 90% lower than EKS. The TCO savings are 94% in Year 1, 86% in Year 2, and 85% in Year 3. The following table presents a summary of the analysis.

3 year TCO Analysis					
Medium Scenario	Amazon SageMaker	DIY options on Amazon EC2	Managed K8s on AWS	SM TCO v/s DIY EC2	SM TCO v/s Managed K8s
TOTAL COSTS (Build + Train + Deploy)	\$213,233	\$2,501,227	\$2,189,631	-91%	-90%
Total Build Costs	\$36,013	\$587,536	\$587,536	-94%	-94%
Total Infrastructure Costs	\$25,669	\$18,369	\$18,369	40%	40%
Operations Costs	\$10,344	\$310,313	\$310,313	-97%	-97%
Total Security & Compliance Costs	\$-	\$258,854	\$258,854	-100%	-100%
Total Train Costs	\$26,313	\$943,737	\$932,525	-97%	-97%
Total Infrastructure Costs	\$15,969	\$374,570	\$363,359	-96%	-96%
Operations Costs	\$10,344	\$310,313	\$310,313	-97%	-97%
Total Security & Compliance Costs	\$-	\$258,854	\$258,854	-100%	-100%
Total Deploy Costs	\$150,906	\$969,954	\$669,570	-84%	-77%
Total Infrastructure Costs	\$140,563	\$400,788	\$100,404	-65%	40%
Operations Costs	\$10,344	\$310,313	\$310,313	-97%	-97%
Security & Compliance Costs	\$-	\$258,854	\$258,854	-100%	-100%

The following are the build details:

- There are 15 data scientists on the team, each running 1 notebook instance of T3.medium.
- The cost per instance per hour \$0.0582 in Amazon SageMaker and \$0.0416 in EC2 and EKS.



The following are the training details:

- 0.5 hours of training per data scientist per business day. Training instances are a mix of M5.xlarge and P3.2xlarge based on the mix of classic ML and deep learning workloads for the scenario.
- Amazon SageMaker does not require you to manage any clusters, and you only pay for the duration for which the training jobs are running. For EC2, each data scientist runs a dedicated instance. With EKS, you can share 1 instance across 4 data scientists, though the cluster needs to run continuously.
- The blended cost per instance per hour is \$0.87 in Amazon SageMaker and \$0.62 in EC2 and EKS.

The following are the deployment details:

- 150 ML models deployed in production. Each model deployed runs on 2 instances. Instances are a mix of M5.xlarge and P3.2xlarge based on the mix of classic ML and deep learning workloads.
- You can deploy 4 models on 1 instance with Amazon SageMaker multi-model endpoints, and 4 models on 1 instance with EKS. Each model deployed in EC2 requires dedicated instances.
- The blended cost per instance per hour is \$0.87 in Amazon SageMaker and \$0.62 in EC2 and EKS.

The following details are applicable to all phases:

- Security and compliance costs are approximately 50% of the costs incurred by large businesses. This number is lower because medium businesses have low security and compliance requirements.
- S3 storage costs are \$0.023 per GB per month.
- EBS volume costs are \$0.14 in Amazon SageMaker and \$0.10 in EC2 and EKS.
- Network costs are \$0.016 in Amazon SageMaker and \$0.001 in EC2 and EKS.

Scenario 3: Large-sized business

The TCO of Amazon SageMaker for a large sized business with 50 data scientists is 78% lower than EC2 and 65% lower than EKS. The TCO savings are 73% in Year 1, 62% in Year 2, and 51% in Year 3. The following table presents a summary of the analysis.

3 year TCO Analysis					
Large Scenario	Amazon SageMaker	DIY options on Amazon EC2	Managed K8s on AWS	SM TCO v/s DIY EC2	SM TCO v/s Managed K8s
TOTAL COSTS (Build + Train + Deploy)	\$2,062,111	\$9,502,477	\$5,866,850	-78%	-65%
Total Build Costs	\$105,571	\$1,234,703	\$1,234,683	-91%	-91%
Total Infrastructure Costs	\$83,679	\$60,224	\$60,204	39%	39%
Operations Costs	\$21,892	\$656,771	\$656,771	-97%	-97%
Total Security & Compliance Costs	\$ -	\$517,708	\$517,708	-100%	-100%
Total Train Costs	\$277,127	\$2,364,422	\$2,273,403	-88%	-88%
Total Infrastructure Costs	\$255,235	\$1,189,943	\$1,098,924	-79%	-77%
Operations Costs	\$21,892	\$656,771	\$656,771	-97%	-97%
Total Security & Compliance Costs	\$ -	\$517,708	\$517,708	-100%	-100%
Total Deploy Costs	\$1,679,412	\$5,903,351	\$2,358,764	-72%	-29%
Total Infrastructure Costs	\$1,657,520	\$4,728,872	\$1,184,285	-65%	40%
Operations Costs	\$21,892	\$656,771	\$656,771	-97%	-97%
Security & Compliance Costs	\$ -	\$517,708	\$517,708	-100%	-100%

The following are the build details:

- There are 50 data scientists on the team, each running 1 notebook instance of T3.medium.
- The cost per instance per hour \$0.0582 in Amazon SageMaker and \$0.0416 in EC2 and EKS.



The following are the training details:

- 2.5 hours of training per data scientist per business day. Training instances are a mix of M5.xlarge and P3.2xlarge based on the mix of classic ML and deep learning workloads for the scenario.
- Amazon SageMaker does not require you to manage any clusters, and you only pay for the duration for which the training jobs are running. For EC2, each data scientist runs a dedicated instance. With EKS, you can share 1 instance across four scientists, though the cluster needs to run continuously.
- The blended cost per instance per hour is \$0.87 in Amazon SageMaker and \$0.62 in EC2 and EKS.

The following are the deployment details:

- 50 ML models deployed in production. Each model deployed runs on 2 instances. Instances are a mix of M5.xlarge and P3.2xlarge based on the mix of classic ML and deep learning workloads.
- You can deploy 4 models on 1 instance with Amazon SageMaker multi-model endpoints, and 4 models on 1 instance with EKS. Each model deployed in EC2 requires dedicated instances.
- The blended cost per instance per hour is \$0.87 in Amazon SageMaker and \$0.62 in EC2 and EKS.

The following details are applicable to all phases:

- Security and compliance costs incurred are 100% of the costs.
- S3 storage costs are \$0.023 per GB per month.
- EBS volume costs are \$0.14 in Amazon SageMaker and \$0.10 in EC2 and EKS.
- Network costs are \$0.016 in Amazon SageMaker and \$0.001 in EC2 and EKS.

Scenario 4: Extra-Large sized business

The TCO of Amazon SageMaker for an extra-large-sized business with 250 data scientists is 72% lower than EC2 and 54% lower than EKS. The TCO savings are 55% in Year 1, 43% in Year 2, and 43% in Year 3. The following table presents a summary of the analysis.

3 year TCO Analysis					
X-Large Scenario	Amazon SageMaker	DIY options on Amazon EC2	Managed K8s on AWS	SM TCO v/s DIY EC2	SM TCO v/s Managed K8s
TOTAL COSTS (Build + Train + Deploy)	\$7,738,746	\$27,551,979	\$17,003,445	-72%	-54%
Total Build Costs	\$488,285	\$3,214,319	\$3,214,218	-85%	-85%
Total Infrastructure Costs	\$408,191	\$293,778	\$293,678	39%	39%
Operations Costs	\$80,094	\$2,402,832	\$2,402,832	-97%	-97%
Total Security & Compliance Costs	\$ -	\$517,708	\$517,708	-100%	-100%
Total Train Costs	\$2,436,221	\$7,924,226	\$7,485,380	-69%	-67%
Total Infrastructure Costs	\$2,356,126	\$5,003,686	\$4,564,840	-53%	-48%
Operations Costs	\$80,094	\$2,402,832	\$2,402,832	-97%	-97%
Total Security & Compliance Costs	\$ -	\$517,708	\$517,708	-100%	-100%
Total Deploy Costs	\$4,814,240	\$16,413,434	\$6,303,846	-71%	-24%
Total Infrastructure Costs	\$4,734,146	\$13,492,894	\$3,383,305	-65%	40%
Operations Costs	\$80,094	\$2,402,832	\$2,402,832	-97%	-97%
Security & Compliance Costs	\$ -	\$517,708	\$517,708	-100%	-100%

The following are the build details:

- There are 250 data scientists on the team, each running 1 notebook instance of T3.medium.
- The cost per instance per hour \$0.0582 in Amazon SageMaker and \$0.0416 in EC2 and EKS.



The following are the training details:

- 5 hours of training per data scientist per business day. Training instances are a mix of M5.xlarge and P3.2xlarge based on the mix of classic ML and deep learning workloads for the scenario.
- Amazon SageMaker does not require you to manage any clusters, and you only pay for the duration for which the training jobs are running. For EC2, each data scientist runs a dedicated instance. With EKS, you can share 1 instance across 4 data scientists, though the cluster needs to run continuously.
- The blended cost per instance per hour is \$0.87 in Amazon SageMaker and \$0.62 in EC2 and EKS.

The following are the deployment details:

- 150 ML models deployed in production. Each model deployed runs on 2 instances. Instances are a mix of M5.xlarge and P3.2xlarge based on the mix of classic ML and deep learning workloads.
- 4 models can be deployed on 1 instance with Amazon SageMaker Multi-model endpoints. 4 models can be deployed on 1 instance with EKS. Each model deployed in EC2 requires dedicated instances.
- Blended Cost per instance per hour \$0.87 in Amazon SageMaker and \$0.62 in EC2 and EKS.

The following details are applicable to all phases:

- Security and compliance costs incurred are 100% of the costs.
- S3 Storage costs are \$0.023 per GB per month.
- EBS volume costs are \$0.14 in Amazon SageMaker and \$0.10 in EC2 and EKS.
- Network costs are \$0.016 in Amazon SageMaker and \$0.001 in EC2 and EKS.

Data scientist productivity gains



Data scientists and ML developers are not that easy to find these days. It's not only expensive to hire the right candidate with ML expertise, but it can also be extremely difficult to retain them due to the high demand for these roles. Moreover, the maturity of the ML infrastructure and development processes often limits a data scientist's productivity. You spend a large portion of your time on manual tasks such as setting up infrastructure and environments, repetitive tasks with low automation, iterating on model versions, managing model lifecycles, and rewriting code to make it production-ready. Amazon SageMaker provides pre-built components and fully managed features that boost data scientist productivity by up to 10 times.

Amazon SageMaker makes it easy to build, train, tune, and deploy models at scale. It brings together all the tools to perform, debug, and automate the end-to-end steps of building and managing an ML model. You can write code, launch and track ML experiments, debug models, and monitor model quality—all through the single unified interface, thereby improving productivity. This post makes the following assumptions about each phase.

Build phase

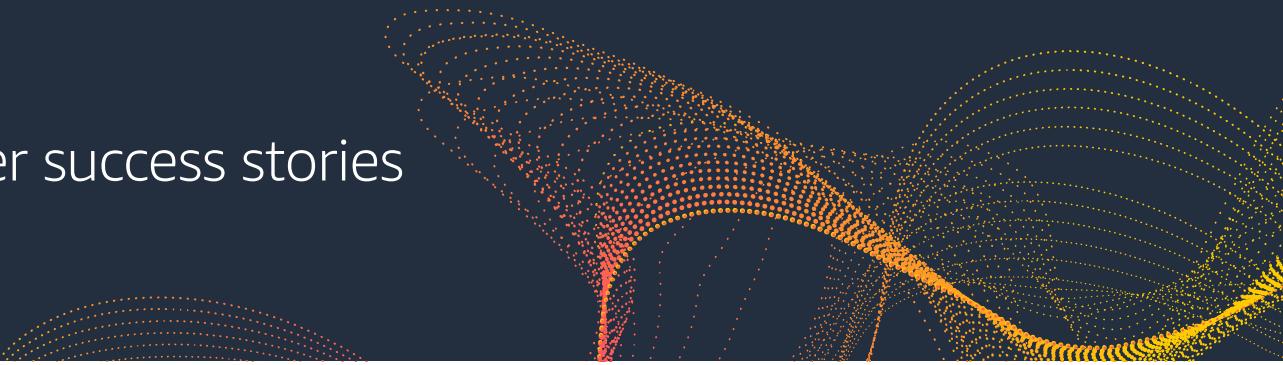
During the build phase, you can explore data, experiment with various ML frameworks, algorithms, and build models on small samples of data. Amazon SageMaker provides pre-built and fully managed Jupyter Notebooks that need zero setup, are pre-configured with the latest and most popular ML libraries, and are easy to share with colleagues. Data preparation is a key challenge and time-consuming process for data scientists. Amazon SageMaker provides visual data prep tools as well as the ability to run ad-hoc analysis, ETL, and Spark jobs on a fully managed and distributed infrastructure. In addition, to reduce the time it takes to build high-quality models, Amazon SageMaker provides AutoML capabilities and 17 pre-built ML algorithms that cover the most common ML use cases in supervised, unsupervised, and reinforcement learning.

Train phase

During the training phase, you can experiment with data processing techniques, various ML algorithms, and model hyperparameters. The number of iterations and training jobs required before deploying a model in production successfully can range from hundreds to thousands. Amazon SageMaker's automatic model tuning capabilities use ML to find the best model based on customer-defined objectives and reduce the time needed to get to high-quality models. In addition, the debugging capabilities in Amazon SageMaker let you detect issues in training jobs early, better introspect deep learning models, and get more insights into the training performance. Finally, as the number of projects, trials, and experiments in the organization increase to the order of thousands, Amazon SageMaker's experiment management capabilities and ML IDE make it easy to organize, track, and manage ML experiments and training jobs across time and organizations.

Deploy phase

During the deployment phase, you can put your ML models in production and make inferences on unseen data. Amazon SageMaker enables one-click model deployment without the need for any code changes. You can train the models in Amazon SageMaker or different environments, and build them using Amazon SageMaker built-in models or custom-built models. You can deploy the same models to make inferences in low latency in real time or for batches of data. In addition, you can deploy new models seamlessly without any impact on availability or loss of performance. Finally, Amazon SageMaker automatically saves inference requests and responses and analyzes the data collected periodically to detect model drift and data quality issues in production. This continuous monitoring and alerting enable you take preventative actions such as retraining models without building any tooling.



Customer success stories

The following are success stories on how Amazon SageMaker customers have saved costs and improved productivity.



Coinbase uses ML models on Amazon SageMaker to help with fraud prevention, identity verification, and large-scale compliance. **Using Amazon Sagemaker reduced model training times from 20 hours to 10 minutes.**



Intuit developed ML models that can pull a year's worth of bank transactions to find deductible business expenses for customers. **Using Amazon SageMaker, Intuit reduced ML deployment time by 90%, from 6 months to 1 week.**



Using Amazon SageMaker, NuData Security prevents credit card fraud by analyzing anonymized user data to detect anomalous activity before a fraudulent transaction occurs. **With Amazon SageMaker, NuData reduced ML development time by 60%, simplified their ML architecture by 95%, and worked with a large bank to passively block nearly 100% of fraudulent attempt traffic within the bank's consumer friction tolerance.**



Using Amazon SageMaker, Voodoo can decide in real time which ad to show to their players and invoke their endpoint over 100 million times by over 30 million users daily, representing close to a billion predictions per day. **With AWS machine learning, Voodoo put an accurate model into production in less than a week**, supported by a small team, and has built on top of it continuously as their team and business grow.



Using TensorFlow on Amazon SageMaker, Siemens Financial Services developed an NLP model to extract critical information to accelerate investment due diligence, **reducing time to summarize diligence documents from 12 hours to 30 seconds.**



Celgene uses Apache MXNet on Amazon SageMaker for toxicology prediction to analyze biological impacts of potential drugs virtually, without putting patients at risk. **A model that previously took 2 months to train can now be trained in 4 hours.**



ADP uses AWS machine learning, including Amazon SageMaker, to quickly identify workforce patterns and predict outcomes before they happen, for example employee turnover or the impact of an increase in compensation. **ADP reduced the time to deploy machine learning models from 2 weeks to just 1 day.**

Conclusion

Amazon SageMaker is a modular, fully managed, end-to-end ML service that lets you build, train, tune, and deploy models at scale. It eliminates the overhead costs of provisioning hardware and systems software, managing infrastructure and operations, and building security and compliance for ML workloads. This enables you to focus on your business problems, develop with new techniques, and take ML projects from idea to production faster. The total cost of ownership of Amazon SageMaker over a 3-year horizon is over 54% lower compared to other cloud-based ML options such as self-managed EC2 options and managed Kubernetes (K8s) options on AWS such as EKS. In addition to the lower TCO, Amazon SageMaker's fully managed and integrated features lets you put ML ideas into production faster and improve data scientist productivity by up to 10 times.