

nClouds | AWS Case Studies

Onriva

How nClouds helped Onriva support rapid business growth by implementing machine learning.

About Onriva

Onriva is an industry-transforming travel platform that is leveraging the latest standards and technologies (NDC), AI, and machine learning to provide a personalized one-stop shopping experience for all travel, including air, hotels, cars, packages, and more. To learn more, go to www.onriva.com

Benefits Summary



Faster build of highly accurate training datasets.



Reduced costs and complexity, increased accuracy.



Rapid extract, transform, load (ETL) of data.



Industry

Online travel services

Location

Foster City, CA

Challenge

Support rapid business growth by implementing machine learning to enable personalized travel itineraries that drive increased booking rates.

Featured Services

Machine learning (Amazon SageMaker), Amazon Simple Storage Service (Amazon S3), AWS Glue, AWS Lambda

CHALLENGE

Support rapid business growth by implementing machine learning to enable personalized travel itineraries that drive increased booking rates.

In a 2017 Accenture study, 67% of respondents said that they want brands to use previous travel information to help them make better travel decisions. So, it's not surprising that companies in the travel industry have embraced machine learning to deliver new insights and provide customers with enhanced user experience.

By analyzing large datasets, machine learning-infused travel systems can predict travelers' needs and provide them with personalized recommendations. As a result, the traveler benefits from getting relevant information at the right time, while the travel provider benefits from new customer acquisition, maximizing revenues, and increased customer engagement and loyalty.

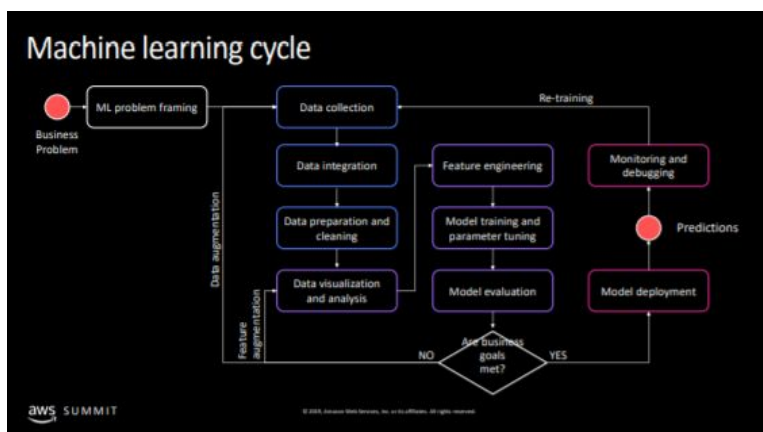


InClouds brought a deep expertise in designing and implementing big data infrastructure and helped (with) incorporating machine learning and AI for our recommender platform. Really appreciate their help at the very early stage of model development and getting our data ready for machine learning."

— **Irina I Kuznetsova,**
Chief Product Officer,
Onriva

Onriva wanted to apply machine learning to develop a powerful travel itinerary recommendation algorithm. To do so, they needed to enhance their AWS infrastructure to quickly and easily build and train machine learning models, and then directly deploy them into a production-ready hosted environment.

Onriva asked nClouds to help them build a machine learning model. The model would analyze user searches to make predictions on user patterns. The AWS machine learning model cycle is shown below:



Source: Build Machine Learning Models with Amazon SageMaker, talk at AWS Summit Paris with Tarkett by Julien Simon, Global Technical Evangelist, Artificial Intelligence & Machine Learning, Amazon Web Services, published April 3, 2019

nClouds began the project by creating Amazon SageMaker notebook instances. These are fully-managed, machine learning EC2 compute instances running the Jupyter Notebook App, an open-source web application to create and share documents that contain live code, equations, visualizations, and narrative text. Jupyter notebook use cases include machine learning, data cleaning and transformation, numerical simulation, statistical modeling, and data visualization.

As anyone who has tried to implement a predictive model via machine learning knows, getting to a comprehensive training dataset is 90% of the work. nClouds worked with Onriva to understand the key features driving bookings and designed a training dataset that exhibited all those features. nClouds then implemented the nightly generation of an updated training dataset leveraging 1.5 years' worth of historical data consisting of disparate sources formatted in deeply nested JavaScript Object Notation (JSON) along with relational structures. nClouds' expertise in technologies like AWS Glue and Athena enabled them to complete this task in several weeks. Once the dataset was in place, Onriva could start building their own machine learning models using SageMaker to perform predictive analysis and generate personalized travel itinerary recommendations for their users.

Here is an illustration of the steps to using Amazon SageMaker to build, train, and deploy machine learning models at scale:



Source: Build Machine Learning Models with Amazon SageMaker, talk at AWS Summit Paris with Tarkett by Julien Simon, Global Technical Evangelist, Artificial Intelligence & Machine Learning, Amazon Web Services, published April 3, 2019



Why AWS and nClouds

Onriva turned to nClouds, a Premier Consulting Partner in the Amazon Web Services Partner Network (APN), to partner with them in the build phase of implementing machine learning. nClouds applied their AWS technical expertise to create a dataset ready to support machine learning.

Onriva leveraged several Amazon Web Services:

- **Amazon Athena** - An interactive query service that makes it easy to analyze data in Amazon S3 using standard SQL. Athena is serverless, so there is no infrastructure to manage, and Onriva pays only for the queries that are run.
- **Amazon Elastic Compute Cloud (EC2)** - A web service that provides Onriva with secure, resizable compute capacity in the cloud.
- **Amazon Relational Database Service (Amazon RDS)** - Enables Onriva to easily set up, operate, and scale a relational database in the cloud.
- **Amazon SageMaker** - A fully-managed service that covers the entire machine learning workflow to label and prepare Onriva's data, choose an algorithm, train the algorithm, tune and optimize it for deployment, make predictions, and take action.
- **Amazon Simple Storage Service (Amazon S3)** - A flexible way to store and retrieve data, providing Onriva with cost optimization, access control, and compliance.
- **AWS Glue** - A fully-managed extract, transform, and load (ETL) service that makes it easy for Onriva to prepare and load their data for analytics.
- **AWS Lambda** - Enables Onriva to run code without provisioning or managing servers.

Onriva's solution stack also includes additional, essential third-party tools and services:

- **Apache Spark** - A unified analytics engine for large-scale data processing.
- **Microsoft Power BI** - A business analytics solution that lets Onriva visualize their data.

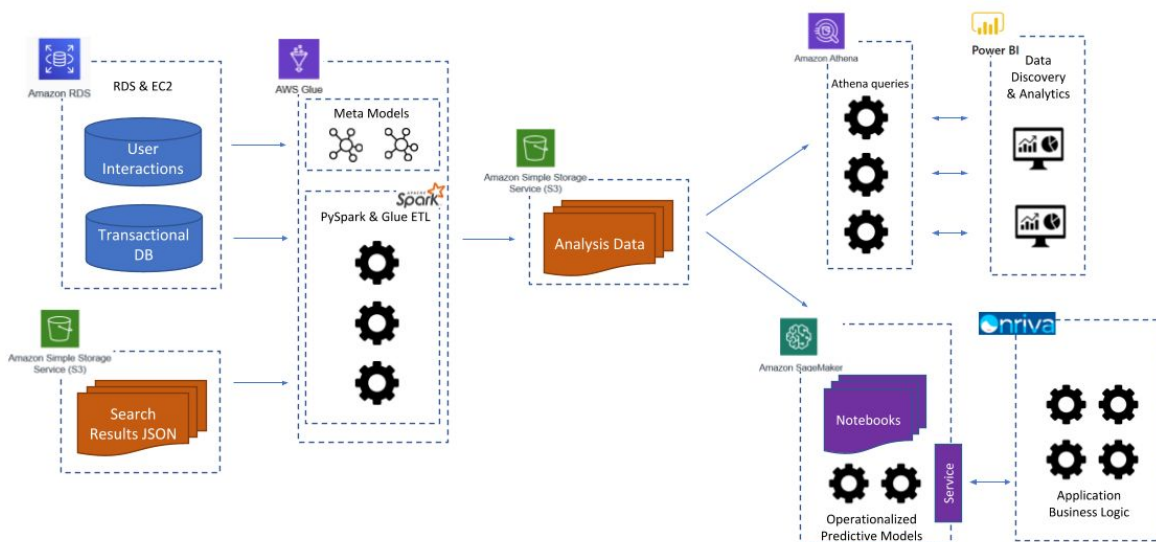
nClouds' Solution Architecture for Onriva

Onriva engaged with nClouds to help them with the build phase of implementing machine learning.

Users' search results are provisioned by AWS Lambda before being incorporated in a transactional database. Onriva's transactional database uses Amazon RDS, and their user interaction database uses Amazon EC2. Search results in JSON format are stored in Amazon S3.

AWS Glue ETL service prepares and loads Onriva's data for analytics. That data is then applied to metamodels designed for fast machine learning and Apache Spark for data processing. The resulting analysis data is sent to Jupyter notebooks for training data exploration and preprocessing as well as to Amazon Athena for queries. And, as mentioned earlier, Amazon SageMaker performs predictive analysis and generates personalized travel itinerary recommendations for their users.





High-level architecture diagram

The Benefits

Teaming with nClouds, Onriva began the build phase of developing a machine learning model, by creating a dataset ready to support machine learning. This phase of the project has yielded numerous benefits:



Faster build of highly accurate training datasets.

nClouds created Jupyter notebooks for training data exploration and preprocessing, which in turn enabled a faster build of highly accurate training datasets. Now Onriva has one source of truth for their machine learning environment.



Reduced costs and complexity, increased accuracy.

nClouds implemented Amazon SageMaker to reduce cost and complexity, while also increasing the accuracy of data labeling by bringing together machine learning with a human labeling process called active learning.



Rapid extract, transform, load (ETL) of data.

Using AWS Glue, Onriva's data is immediately searchable, queryable, and available for ETL. It automates much of the effort in building, maintaining, and running ETL jobs. AWS Glue is serverless, so there is no infrastructure to provision or manage.

Endnotes

Accenture. (2017). Accenture Interactive Personalized Marketing Index: The New Travel Experience. Retrieved from https://www.accenture.com/_acnmedia/PDF-61/Accenture-Interactive-Personalized-Marketing-Index_v2.pdf#zoom=50

Simon, Julien. (2019). Build Machine Learning Models with Amazon SageMaker. [SlideShare]. Retrieved from <https://www.slideshare.net/JulienSIMON5/build-machine-learning-models-with-amazon-sagemaker-april-2019/8>

About nClouds

nClouds is a certified, award-winning provider of AWS and DevOps consulting and implementation services. We partner with our customers, as extensions of their teams, to build and manage modern infrastructure solutions that deliver innovation faster. We leap beyond the status quo.

Copyright © 2022 nClouds, Inc. All rights reserved

