USE CASE LIBRARY

External Customer Stories

Go to the Snowflake ML Compass for all available resources

REV 01.05.24

Safe Harbor and Disclaimers

Other than statements of historical fact, all information contained in these materials and any accompanying oral commentary (collectively, the “Materials”), including statements regarding (i) Snowflake’s business strategy, plans or priorities, (ii) Snowflake’s new or enhanced products, services, and technology offerings, including those that are under development or not generally available, (iii) market growth, trends, and competitive considerations, (iv) our vision for Snowpark, the Data Cloud, and industry-specific Data Clouds, including the expected benefits and network effects of the Data Cloud; and (v) the integration, interoperability, and availability of Snowflake’s products, services, or technology offerings with or on third-party platforms or products, are forward-looking statements. These forward-looking statements are subject to a number of risks, uncertainties and assumptions, including those described under the heading “Risk Factors” and elsewhere in the Annual Reports on Form 10-K and the Quarterly Reports on Form 10-Q that Snowflake files with the Securities and Exchange Commission. In light of these risks, uncertainties, and assumptions, the future events and trends discussed in the Materials may not occur, and actual results could differ materially and adversely from those anticipated or implied in the forward-looking statements. As a result, you should not rely on any forwarding-looking statements as predictions of future events. Forward-looking statements speak only as of the date the statements are first made and are based on information available to us at the time those statements are made and/or management's good faith belief as of that time. Except as required by law, we undertake no obligation, and do not intend, to update the forward-looking statements in these Materials.

Any future product or roadmap information (collectively, the “Roadmap”) is intended to outline general product direction. The Roadmap is not a commitment, promise, or legal obligation for Snowflake to deliver any future products, features, or functionality; and is not intended to be, and shall not be deemed to be, incorporated into any contract. The actual timing of any product, feature, or functionality that is ultimately made available may be different from what is presented in the Roadmap. The Roadmap information should not be used when making a purchasing decision. In case of conflict between the information contained in the Materials and official Snowflake documentation, official Snowflake documentation should take precedence over these Materials. Further, note that Snowflake has made no determination as to whether separate fees will be charged for any future products, features, and/or functionality which may ultimately be made available. Snowflake may, in its own discretion, choose to charge separate fees for the delivery of any future products, features, and/or functionality which are ultimately made available.

The Materials may contain information provided by third-parties. Snowflake has not independently verified this information, and usage of this information does not mean or imply that Snowflake has adopted this information as its own or independently verified its accuracy.

Table of Contents

Scene+
Zscaler
Cloudbeds
Fidelity
CHG Healthcare
IGS Energy
Swire Coca Cola
Spark New Zealand
Decile
INVISTA
Ecolab
Cooke Aquaculture
BAMA
Lessmore
Paytronix
Keysight
Avios

SpartanNash
Snowflake Data Science

ARCHITECTURE

Snowflake ML

Integrated set of capabilities for end-to-end ML in Snowflake

Data Science Agent PR soon Snowflake Notebooks External
IDEs
Batch Data
Apps

Snowflake ML APIs
Experiment Tracking PR soon

ML Jobs GA soon

Model Registry
Feature Store

Model Serving

ML Observability
Explainability
Streams
Dashboards
Container
Runtime
GPU CPU + pip install more
ML-Optimized
Warehouse
CPU
Any Model
APIs
PR: Private Preview | PU: Public Preview

Customers Building on Snowflake ML

Customer-Facing Slides

Snowflake ML Customer Deck

CUSTOMER REFERENCES

For Data Scientists and ML Engineers

Customer Highlight:

Faster and cheaper dev to production for ML workflows in Snowflake

Sagemaker
Preprocessing
Scoring
& Inference
Post-processe
d scores

- Costly & slow to move EDW data to S3 for SageMaker, especially with 60M+ rows
- Separate ML platform caused friction
- Platform overhead

Snowflake
data
Data
Egress
Data
Ingress
Snowflake
data

Feature
Engineering &
Model Training
Training
Dataset
Snowflak
e data
Feature
Store
Scoring/Inference
Model
Registry

- Cut time to production on feature engineering, training & inference by over 60%
- 36% cost savings on compute and operation
- 30+ models built in under 6 months using Snowflake ML

Single Table to CDP
Snowflake Notebook/Task
Inference
Dataset

Scene+ leverages machine learning to deliver relevant member experiences across our properties. This requires working with a vast amount of data. Leveraging the straightforward Snowflake Feature Store drove a 66%

reduction in processing time; we can join the model universe with the features with just four blocks of code. Previous methods required writing extensive Python scripts, input files and additional dependency scripts."
Aasma John
Data Science Manager | Scene+
Blog link

[11]

Customer Highlight:

Easy, governed sales trend predictions across countries in Snowflake Notebooks

Developer View:

E2E Lifecycle is easier in Snowflake than Sagemaker

Admin/Governance View:

ML is cheaper + easier to govern in Snowflake

Build a model ~2 weeks setup
(set up connectors +
permissions to pull in data)
0 set up, so you
can start building
immediately
Permissions
set up
Time consuming
Must manage permission
set up and connectivity in
Snowflake + Sagemaker
None Needed
E2E flows use only
Snowflake RBAC
Productionize
the model
~3 weeks setup
(Separate IT team had to
create image pipeline,
schedule to run in 3rd party
tool, and write predictions
to Snowflake)
<1 hour
(Data Scientist can use
GIT to push from DEV to
PROD account; PROD
account DevOps team
will schedule)
Data
transfer
Costly
Data in-flight between AWS
<> Snowflake (also potential
data security issues)
None Needed
Data fully within
Snowflake
Visibility &
collaboration
Iterate & update
Notebooks
~1 day
for changes to go live
(Separate IT had to
push image)
<1 hour
(Dev Ops team is

notified of changes via
email, and pulls in new
updates quickly)
Difficult
Siloed local-machine
buildings; hard to leverage
each other's work for
projects = longer time
to delivery
Easy to govern
Fully visibility into ML
projects; easier to reuse
each other's notebooks
and flows

Cloudbeds - Achieves 95% accuracy across a 20,000-property dataset with Snowflake ML

Cloudbeds is a global hospitality management platform that uses ML to forecast performance across 20,000 global properties
Link to blog
Difficulty scaling forecasting across a large dataset with 20k properties.
Long 12+ hour training cycles, reducing experiments the team could run each week.
Accelerated experimentation cycles, achieving 95% accuracy and efficiency gains of over 90% within a six-month forecasting window.
Experiments that took 12+ hours were reduced to 30 minutes

Customer Story:

95% forecasting accuracy with 24x reduced training time for 20k+ properties in Snowflake

Before After with Snowflake ML
Training Cycles 12+ hours per experiment
3 experiments per week
30 mins per experiment
5 experiments per day
Read the blog!
Large datasets Multiple data stores Streaming data into Snowflake
Experimentation
tracking Limited Complete across all properties
Reproducibility Manual Full tracked and automated

Fidelity Investments

Feature Engineering with Snowflake ML

Language of Choice on a Single Platform

Data is not duplicated nor transferred across the network.

Scalability without Operational Complexity

Handles large data volumes
Scales both vertically and horizontally
Simple to use
Lazy evaluation

No Governance and Security Trade-offs

Leverages extensive RBAC controls, enabling tightly managed security

Learn More

MinMax Scaler One Hot Encoding Pearson Correlation

28 million rows 100 million rows 162k rows

536 7 77x
654 13 50x
702 41 17x

CHG Healthcare

Classification models in Snowflake ML for healthcare staffing

"Using GPUs from Snowflake Notebooks on Container Runtime turned out to be the most cost-effective solution for our machine learning needs. We appreciated the ability to take advantage of Snowflake's parallel processing with libraries beyond Snowflake ML, offering flexibility and improved efficiency for our workflows."
— Data Scientist, CHG Healthcare

Model Development in Snowflake ML
Physician
Data
Python SDK
Predicted
Outcomes
Worksite
Assess Model
Data Structured
Data
Preprocessing
Train & Test Split
EDA Model
Training
Feature
Importance
& SHAP
Performance

SQL API
Predicted
Error & Feature
Outcomes
Drift Monitoring
3rd Party
Data
Snowpark
Container
Services
Application
Endpoint
Email

Learn More

IGS Energy - ML Migration from Databricks to Snowflake ML

IGS is a retail energy provider in the Midwest that offers various sustainable energy solutions.

Their business requires demand forecasting for transactions ranging from long term-decisions around buying and selling power to trades in the “day-ahead energy market”
IGS’ ML forecasting models require massive amounts of data, up to 40-50 billion rows exceeding a terabyte
Using one predictive model per customer account which came with a lot of overhead when scaled to all accounts
Link to case study
Moving from hundreds of thousands of individual forecasting models in Databricks to one unified model in Snowflake helped IGS achieve 75% cost savings in training — without sacrificing accuracy.
Training models and generating predictions previously took a half hour vs minutes in Snowflake to generate forecasts for hundreds of thousands of customers

“We can more easily build predictive models and mock up data products all in the Snowflake ecosystem because the data is all there.”
- Dan Shah, Manager of Data Science, IGS Energy

Swire Coca Cola - Managed Spark Migration for AI/ML

Optimization of planned logistic routes for Swire, a local bottler for Coca-Cola and other beverage brands
Link to blog
Managed Spark platform had multiple billing processes and additional cloud provider costs that made it difficult to manage and track costs
Data was moved in and out of Snowflake for feature engineering and model training
Managing complex infrastructure diverted away time and resources from building models
Governance, data integrity and security challenges
Model deployment accelerated by weeks with Snowpark and Snowpark ML, resulting in significant cost savings
The seamless transition was facilitated by the inherent

advantages of Snowpark and Snowpark ML, which exhibited a syntactic similarity to Spark and SparkML

Spark New Zealand - Unified Marketing Analytics Platform

End-to-end marketing analytics solution using from Snowpipe to Snowpark ML

Create machine learning models with Snowpark ML to better understand customers' needs and preferences

Data science team was using virtual machines for processing with deployment in docker containers involved unnecessary data transfers, increased costs and complexity

Link to blog

Queries are executed lazily, improving performance, and the absence of data movement enhances governance. This streamlined methodology eliminates the necessity for intricate deployment pipelines and additional computing resources, resulting in significant cost and time savings.

Building Predictive Purchase Models with Snowflake ML

9.2x

Easy Management

Speed up with Snowflake ML compared to Spark

Performance / Scale

Speed

Eliminate Complexity

Optimize

Launch

Decile - Propensity to Purchase Models in Snowflake ML

Decile is an all in one Customer Data + Analytics Platform that helps marketing and leadership teams make data-driven decisions to help their brands grow profitably.

Built Propensity to Purchase models in Snowflake ML using the Snowpark

ML Modeling APIs and Model
Registry
Previously used managed Spark for
ML which caused several challenges:
- Data transfers
- Multiple levels of resource administration, complicating tuning
- Balancing resources / timing required sampling data for input into models
- Additional code for caching

Link to webinar
9.2x speed up with Snowflake ML compared to Spark (reduced time from 60 to 6.5 mins)
Intuitive developer experience with familiar SKLearn and XGBoost APIs
Facilitates backtesting, QA, and monitoring
Can easily share and collaborate with other team members with Snowflake permissions model
Uses the Model Registry to manage model versions and easily perform inference with SQL on a daily basis

"By bringing familiar modeling capabilities to Snowflake, Snowpark ML has enabled us to more rapidly iterate on our models, improving accuracy and operational efficiency."
- SVP of Engineering at Decile in webinar

INVISTA - Unified and Governed MLOps in Snowflake ML

Supply chain forecasting
Link to webinar
Link to Summit presentation
Cross-Team Dependencies
Multiple platforms and points of failure
No structure in experimentation → deployment
Less measurables per / business vertical / project
Multiple architectures for ML & DL deployment
DevOps style model re-training
Universal structure in experimentation → deployment
Compute to the data
Great visibility for business on usage & cost metrics / vertical
Custom alerts for failure
Ease in orchestrating dags

"Doing MLOps on Snowflake will reduce our deployment time from months to days, even hours, all while reducing development cloud spend and ensuring model tracking and visibility."
- Data Science Leader at INVISTA in webinar

Ecolab - Sales Opportunity Predictions with Snowflake ML

Extend value delivery to Ecolab's customers by recommending product solutions that similar customers purchase using Snowpark ML KMeans clustering
Diverse customer base driving accuracy issues with existing solution
PaaS offerings required significant additional development. SaaS offerings pre-Snowflake offered limited customization
71x speed up for processing with Snowflake ML vs native Python on 10.1GB of data (12.6 seconds vs. 15 minutes with native Python)
Saved ~ 4 weeks of development time with Snowflake ML
7 lines of code to create a custom model, train it and register it for reusability with Snowpark ML Modeling APIs
The Snowflake Model Registry stores the ML model as a first-class schema-level object that anyone in our org can use

Cooke Aquaculture - Rapid ML Insights with Snowflake ML

Cooke Aquaculture is the largest private family-owned seafood company in the world that uses Snowflake ML to generate predictions for production metrics
Scalability - building models on over ten years of raw data from over seventy sites stored in Snowflake
Link to blog
Unified Development Environment: Everything was contained within Snowflake from raw data, data pipelines and ML models.
Ability to scale compute resources on demand: No need to worry about local processing constraints.
Snowflake's powerful ML tools: Snowflake ML includes integrated capabilities for faster time to production ML.

Predicting Sales Opportunities with Snowflake ML

71x

Speed up for
preprocessing and
training with Snowflake
ML compared to
native Python
Saved an estimated
4 weeks of development
time with Snowflake ML
solution on 10 GB of Data

BAMA - Financial Forecasting Model with Snowflake ML

Bama-Gruppen AS, Norway's largest private distributor of fruit and vegetables, predicts invoice due dates at the time of purchase
BAMA is exposed to currency risk - they buy their products in USD, EUR and GBP, but sell them in NOK
Solution with Snowflake ML reduces cash flow-related losses by 35%

"The due date prediction model has significantly reduced our exposure to currency exchange rate risk, enhancing financial stability and forecasting accuracy"

- Jarle Gjerde, Group CFO @ BAMA

Predicting Invoice Due Dates with Snowflake ML

35%

Reduction in
cash-flow
related losses

"The due date prediction model
has significantly reduced our
exposure to currency exchange
rate risk, enhancing financial
stability and forecasting accuracy"
Jarle Gjerde
Group CFO @ BAMA

Classy - Rapid ML Insights with Snowflake ML

Classy is a giving platform that enables nonprofits to connect supporters with the causes they care about
Classy uses Snowflake ML to harness the power of intelligence to help nonprofits understand their supports including:

- Better understanding

donors and how they discovered their organization
● Ability to tailor communications based on this knowledge
● Understanding the right opportunities for upgrade

A small Data team building a ML service at a massive scale

Modern data stack doesn't readily support MLOps

Limited people to support new architectures

Donation forms using the Intelligent Ask Amounts ML solutions saw 11% greater donation revenue in testing

Now able to use compute and storage within Snowflake during the whole model development process in Hex notebooks

Quickly train and evaluate models in a user-friendly syntax

Able to register models to the Snowflake Model Registry

Powerful hyperparameter tuning in only a few lines of code

S&P Global -

●

75% time savings by moving from PySpark on Databricks to Snowflake ML

At Lessmore, leveraging the Snowflake Model Registry has transformed our model development and experimentation process for our customer lifetime value forecasts. This shift has not only accelerated our innovation cycle but also reduced our costs by a factor of 10, while enhancing efficiency."

—Moritz Schöne, Head of Data Science, Lessmore

Moritz Schöne

Head of Data Science | Lessmore

Blog link

Using Snowflake ML has enabled us to hit a big milestone in our data and AI vision to efficiently deliver true, one-to-one, personalized experiences to our customers. We've been able to achieve a 70% cost reduction and enhanced agility by moving from running hour-long inference jobs to predictions in near real-time."

Stefan Kochi
CTO | Paytronix
Blog link

31

Having tried Snowflake Notebooks on the Container Runtime, we can say the experience has been remarkable. The flexible container infrastructure supported by distributed processing on both CPUs and GPUs, optimized data loading, and seamless integration with Model Registry have improved our workflow efficiency."
Krishna Moleyar
Analytics & Automation for IT Global Applications | Keysight
Blog link

32

I have really enjoyed using Snowflake Notebooks on Container Runtime for the flexibility and speed they offer. I am able to run my code without worrying about it time-ing out or variables being forgotten. Enabling PyPI integration, I also have the added benefit of using a wider range of Python packages, making my analysis and data science tasks more flexible."
Krishna Moleyar
Data Scientist | Avios
Blog link

33

For Analysts

SpartanNash - Retail Forecasting with ML Functions

Use historical sales data stored in Snowflake to produce year-long sales forecasts for 2023 – divided into 13 fiscal periods
Manual, Excel-based process
Prediction accuracy relied on institutional knowledge
Lower granularity predictions not possible (i.e., at department or item level)
Time commitment: For 183 locations and 10 districts – 400 hours per period, 5,200 hours per year

Fully automated process
Predictions rely on data, not passed-down knowledge
Lower granularity predictions possible
Model training takes 5 minutes per week, on automated schedule
Accuracy improved from 71% to 88%

Slide deck
We've been using Snowflake's ML-based forecasting function for three months now and have saved hours of effort while generating more accurate forecasts than our previous process. Given our experience, we're exploring more ways to use Snowflake ML functions throughout SpartanNash.

Efficiently & accurately predicts daily sales across locations with ML Functions

88% The Snowflake ML Forecasting function now runs automatically each week using Snowflake tasks, replacing 100 hours of manual effort. average forecast accuracy using Snowflake's ML Forecasting function, up from 71% accuracy from a previous method.*

Snowflake Data Science - Lead Scoring with ML Functions

Snowflake needs help determining which prospects to assign to sales teams during annual planning – to maximize Snowflake revenue
To do so, the Snowflake Sales Data Science team calculates a "propensity score" indicating likelihood of a prospect to convert to a customer – based on the historical success with similar prospects
Pulling technographic, firmographic, and emplographic data in from multiple sources – only to move it out of Snowflake

again for modeling
Using multiple third party tools that required moving data out of Snowflake – e.g., Jupyter Notebooks, an ML vendor, Tableau for visualizing results
By knitting together Snowflake Notebooks, ML Functions and Streamlit:
- 5x higher conversion rate with target accounts
- 1K hours of development saved
- 150K accounts assigned through strategic market expansion
- 70% efficiency improvements thanks to reliance only on Snowflake tools (no third party), no data movement, simplified ML development

Slide deck

Propensity Score Methods Evolution with ML Functions

Feature Gathering Modeling Deployment Tracking

Gather features from internal and data providers
Preprocessing data and training models
Infer trained model to make prediction on new data
Monitor model performance and impact

Q2’22 Jupyter Notebook
Snowpark UDFs
Snowpark Stored Proc Tableau

* WiP: Full model completion is aiming to be tested and completed in 2024 Q2

Business Impact for Propensity Scoring Evolution

Growth & Operational Excellence

*Numbers are from 2023

Appendix

End-to-End ML Workflow with Snowflake ML

Data lake Training dataset Snowflake
Model Registry

Predictions

NAME OF COMPANY - Brief Use Case

Quote [optional]

- Title at company with link to asset