

亚马逊云科技宣布推出新一代Amazon SageMaker

为数据、分析和AI提供统一平台

北京2024年12月10日 /美通社/ -- 亚马逊云科技在2024 re:Invent全球大会上，宣布推出新一代Amazon SageMaker，将客户所需的快速SQL分析、PB级大数据处理、数据探索和集成、模型开发和训练以及生成式人工智能(AI)等功能统一到一个集成平台上。

全新的Amazon SageMaker Unified Studio使客户能够轻松查找和访问整个组织的数据，并汇集了亚马逊云科技专门构建的分析、机器学习（ML）和AI功能，客户可以在Amazon Q Developer的协助下，为各种类型的常见数据用例选择最合适的工具处理数据。Amazon SageMaker Catalog和内置的治理功能确保合适的用户以正确目的访问合适的数据、模型和开发组件。全新的Amazon SageMaker Lakehouse统一了数据湖、数据仓库、运营数据库和企业应用程序中的数据，使客户可以轻松访问和处理Amazon SageMaker Unified Studio中的数据，并使用其熟悉的AI和ML工具或与Apache Iceberg兼容的查询引擎。全新zero-ETL与领先的软件即服务（SaaS）应用程序的集成，让客户无需复杂的数据管道，即可在Amazon SageMaker Lakehouse和Amazon Redshift中访问第三方SaaS应用程序的数据，并进行分析或机器学习。众多客户和合作伙伴包括Adatastra、Confluence、Etleap、idealista、Informatica、Lennar、Natera、NatWest Group、NTT Data、罗氏、Tableau、丰田北美公司等已经在使用新一代Amazon SageMaker，整合他们的数据、分析和AI项目。

亚马逊云科技人工智能和数据副总裁Swami Sivasubramanian博士表示："分析和AI正在融合，从历史分析到ML模型训练和生成式AI应用程序，客户以越来越互联的方式使用数据。为了支持这些工作负载，许多客户已经在使用我们专门构建的分析和ML工具组合，例如，已成为处理数据和构建ML模型事实标准的Amazon SageMaker、Amazon EMR、Amazon Redshift、Amazon S3数据湖和Amazon Glue的组合。新一代Amazon SageMaker将这些功能汇集到一起，同时还加入了一些令人兴奋的新特性，为客户提供数据处理、SQL分析、ML模型开发和训练以及构建生成式AI所需的所有工具。"

使用Amazon SageMaker Unified Studio更快地协作和构建

如今，数十万客户使用Amazon SageMaker来构建、训练和部署ML模型。许多客户还依赖亚马逊云科技提供的一整套专门构建的分析服务来支持各种工作负载，包括SQL分析、搜索分析、大数据处理和流式分析。越来越多的客户不再孤立地使用这些工具，相反，他们正在将分析、ML和生成式AI相结合来获取洞察并为用户提供新体验。这些客户将受益于一个统一的环境——汇集了客户熟悉的亚马逊云科技分析、ML和生成式AI工具，不但可以轻松访问他们的所有数据，还能够与团队或组织的其他成员轻松协作处理数据项目。

新一代Amazon SageMaker包括一个新的、统一的工作室，为客户提供一个单一的数据和AI开发环境，用户可以在其中查找和访问其组织中的所有数据，为各种常见的数据用例选择最佳的工具，并将数据和AI项目扩展至团队内及不同分工角色以实现协作。Amazon SageMaker Unified Studio整合了在Amazon Bedrock、Amazon EMR、Amazon Redshift、Amazon Glue和现有Amazon SageMaker Studio中客户喜欢使用的一系列独立"工作室"、查询编辑器和可视化工具的功能和工具。这使客户可以轻松访问和使用这些功能来发现和准备数据、编写查询或代码、处理数据以及构建ML模型。Amazon Q

Developer全程协助支持开发任务，包括数据发现、编码、SQL生成和数据集成等。例如，用户可以询问Amazon Q："我应该使用哪些数据来更好地了解产品销售情况？"或"生成SQL来计算按照产品类别分类的总收入。"用户可以安全地发布数据、模型、应用程序和其他构件并与其团队或组织成员共享，从而加快对数据资产的可发现性和使用率。借助Amazon SageMaker Unified Studio中与Amazon Bedrock集成的开发环境（IDE），用户可以使用Amazon Bedrock精选的高性能基础模型和工具（例如Agents、Guardrails、Knowledge Bases和Flows），快速、轻松地构建和部署生成式AI应用程序。Amazon SageMaker Unified Studio内置数据发现、共享和治理功能，因此分析师、数据科学家和工程师可以轻松搜索和找到其用例所需的正确数据，同时应用所需的安全控制和权限，维护访问控制，并保护他们的数据安全。

NatWest Group是英国一家领先的银行，为超过1900万客户提供服务，使用多种工具进行数据工程、SQL分析、ML和生成式AI工作负载。借助Amazon SageMaker Unified Studio，NatWest Group将在整个组织内拥有一个统一的环境来支持这些工作负载，并预计其数据用户访问分析和AI功能所需的时间将减少50%，从而使他们花更少的时间管理多个服务，将更多的时间用于客户创新。

通过Amazon SageMaker数据和AI治理满足企业安全需求

新一代Amazon SageMaker简化了整个组织内数据和AI的发现、治理和协作。借助基于Amazon DataZone构建的Amazon SageMaker Catalog，管理员可以使用具有精细控制的单一权限模型，定义和实施一致的访问策略，这样跨团队的数据工作人员可以安全地发现和访问经过批准的数据和模型，这些数据和模型包含由生成式AI创建的业务上下文元数据。管理员可以轻松地定义和实施跨模型、工具和数据源的权限，而定制的安全措施有助于确保AI应用程序的安全性和合规性。客户还可以通过Amazon SageMaker中的数据分类、毒舌检测（toxicity detection）、防护栏（guardrails）和负责任的AI策略来保护其AI模型。

Amazon SageMaker Lakehouse减少数据孤岛并统一数据

如今，超过一百万个数据湖构建在Amazon Simple Storage Service（Amazon S3）上，使客户能够集中其数据资产，并通过亚马逊云科技分析、AI和ML工具获取价值。数据湖使客户能够按原样存储数据，从而轻松组合来自多个来源的数据。客户的数据可能分布在多个数据湖以及数据仓库中，用一种简单的方法统一所有这些数据将让客户受益。

Amazon SageMaker Lakehouse可统一访问存储在Amazon S3数据湖、Redshift数据仓库和联合数据源中的数据，无论数据的存储方式和物理位置在哪，都可以减少数据孤岛并让查询更容易。借助Amazon SageMaker中与Apache Iceberg兼容的全新数据湖仓功能，客户可以从Amazon SageMaker Unified Studio中，使用他们熟悉的与Apache Iceberg开放标准兼容的AI和ML工具以及查询引擎，访问和处理他们的所有数据。现在，无论数据以何种方式存储以及存储位置，客户可以使用他们喜欢的分析和ML工具对数据进行处理，以支持包括SQL分析、即席查询、数据科学、ML和生成式AI在内的用例。Amazon SageMaker Lakehouse提供集成的精细访问控制，这些访问控制一致地应用于Lakehouse中所有分析和AI工具中的数据，客户只需定义一次权限即可在整个组织中安全地共享数据。

罗氏（Roche）是一家制药和诊断领域的先锋企业，致力于推进科学进步以改善人们的生活。该公司将使用Amazon SageMaker Lakehouse统一来自Amazon Redshift和Amazon S3数据湖的数据，以消除数据孤岛，增强团队之间的协作，并允许用户无缝利用数据，无需昂贵的数据移动或重复的安全访问控制。借助Amazon SageMaker Lakehouse，Roche预计数据处理时间将减少40%，这让他们减少数据管理工作，而将精力更多用于推动业务发展。

全新zero-ETL集成SaaS应用程序，可快速、轻松地访问SaaS数据

为了在运营中真正利用数据，企业需要无缝访问所有数据，无论这些数据位于何处。这就是亚马逊云科技不断发展zero-ETL的原因。zero-

ETL使数据集成不再是繁琐的手动工作，客户可以轻松地在地方的获取数据。这包括Amazon Aurora MySQL和PostgreSQL、Amazon RDS for MySQL以及Amazon DynamoDB与Amazon Redshift的zero-ETL集成，帮助客户快速轻松地访问在Amazon Redshift和Amazon SageMaker Lakehouse中的常用的关系和非关系数据库中的数据，并用于分析和ML。除了运营数据库和数据湖，许多客户还将关键企业数据存储在SaaS应用程序中，客户将从轻松访问所有这些数据进行分析 and ML中受益。

客户使用与SaaS应用程序集成的全新zero-ETL，可以轻松地访问Amazon SageMaker Lakehouse中的Zendesk和SAP等应用程序以及Amazon Redshift中的数据，并用于分析和AI。这消除了对数据管道的需求，数据管道的构建本身就具有挑战性且成本高昂，而且数据管道管理复杂，容易出错而让客户无法及时获得想要的见解。集成SaaS应用程序的zero-ETL包含数据同步、增量更新和删除检测以及目标merge的最佳实践。

来自不同行业、各种规模的组织机构，包括Infosys、Intuit和Woolworths，已经受益于亚马逊云科技zero-ETL集成，无需构建和管理数据管道，即可快速轻松地连接和分析数据。例如，通过集成SaaS应用程序的zero-ETL，在线房地产平台idealista将能够简化其数据提取和摄取流程，无需多个管道来访问存储在第三方SaaS应用程序中的数据，并使他们的数据工程团队能够专注于从数据中获得可操作的见解，而不是构建和管理基础设施。

新一代Amazon SageMaker现已可用，Amazon SageMaker Unified Studio目前推出预览并很快可用。

原文地址：<http://www.china-nengyuan.com/news/218647.html>