aws

用户指南

Amazon Fraud Detector

版本 latest

Amazon Fraud Detector: 用户指南

Table of Contents

什么是 Amazon Fraud Detector？ 1
优势 1
核心概念和术语 2
Amazon Fraud Detcort 的工作原 5
使用 Amazon Fraud Detector 检测欺诈行为 6
访问亚马逊 Fraud Detector 8
可用性 8
接口 8
定价 8
为 Amazon Fraud Detector 做好准备 9
报名参加 AWS 9
注册获取 AWS 账户 9
创建具有管理访问权限的用户 10
设置访问亚马逊 Fraud Detector 界面的权限 11
设置用于访问 Amazon Fraud Detector 的接口 12
访问亚马逊 Fraud Detector 控制台 12
设置 AWS CLI 12
设置 AWS SDK 13
Amazon Fraud Detecto 14
获取并上传示例数据集 14
教程：开始使用亚马逊Fraud Detector 控制台 16
A 部分：构建、训练和部署亚马逊Fraud Detector 模型 16
B 部分：生成欺诈预测 19
教程：入门AWS SDK for Python (Boto3) 24
先决条件 24
开始使用 24
（可选）使用 Jupyter (IPython) 笔记本浏览亚马逊Fraud Detector API 33
后续步骤 33
事件数据集 35
事件数据集结构 35
使用数据模型浏览器获取事件数据集要求 36
数据模型浏览器 36
收集事件数据 37
数据集验证 42

数据集存储 43
事件类型 44
创建事件类型 44
在亚马逊欺诈检测器控制台中创建事件类型 45
使用创建事件类型 AWS SDK for Python (Boto3) 46
删除事件或事件类型 46
事件数据存储 49
使用 Amazon S3 在外部存储您的事件数据 49
创建 CSV 文件 50
将事件数据上传到 Amazon S3 存储桶 52
使用亚马逊Fraud Detector 在内部存储您的事件数据 53
准备用于存储的事件数据 54
使用批量导入存储事件数据 55
使用 GetEventPredictions API 操作存储事件数据 66
使用 SendEvent API 操作存储事件数据 66
获取存储的事件数据的详细信息 68
查看存储的事件数据集的指标 68
活动编排 69
设置事件编排 70
在 Amazon Fraud Detector 中启用事件编排 70
在 Amazon Fraud Detector 控制台中启用事件编排 70
使用启用事件编排 AWS SDK for Python (Boto3) 71
在 Amazon Fraud Detector 中禁用事件编排 71
在 Amazon Fraud Detector 控制台中禁用事件编排 71
使用禁用事件编排 AWS SDK for Python (Boto3) 72
模型 73
选择模型类型 73
在线欺诈洞察 73
交易欺诈见解 75
账户接管见解 76
构建模型 81
使用训练和部署模型 AWS SDK for Python (Boto3) 81
模型分数 83
对性能指标进行建模 83
模型变量重要性 85
使用模型变量重要性值 86

评估模型变量重要性值 87
查看模型变量重要性排名 87
了解模型变量重要性值的计算方式 88
导入 SageMaker 模型 88
使用导入 SageMaker 模型 AWS SDK for Python (Boto3) 88
删除模型或模型版本 89
探测器 92
创建探测器 92
在亚马逊欺诈检测器控制台中创建检测器 92
使用创建探测器AWS SDK for Python (Boto3) 95
创建探测器版本 95
规则执行模式 96
使用创建探测器版本AWS SDK for Python (Boto3) 96
删除探测器、探测器版本或规则版本 97
资源 99
Variables 99
数据类型 99
默认值 100
变量类型 100
变量丰富 118
创建变量 125
删除变量 127
Labels 128
创建标签 128
更新标签 129
更新存储在 Amazon Fraud Detector 中的事件数据中的事件标签 129
删除标签 130
规则 131
规则语言参考 131
创建规则 136
更新规则 138
列表 139
创建列表 139
在列表中添加条目 141
为列表分配变量类型 142
删除列表 143

从列表中删除条目 144
从列表中删除所有条目 144
结果 145
创建结果 145
删除结果 147
实体 147
创建实体类型 148
删除实体类型 148
使用以下方法管理资源AWS CloudFormation 149
创建on on etFraud Detector 模板 150
管理on on etFraud Detector or 堆栈 150
了解 on on etFraud Detector CloudFormation r 参数 151
AzFAWS CloudFormation raud Detector on 151
了解有关 AWS CloudFormation 的更多信息 152
欺诈预测 153
实时预测 154
实时欺诈预测的工作原理 154
获得实时欺诈预测 154
批量预测 155
批量预测的工作原理 156
输入与输出文件 156
获取批量预测 156
有关 IAM 角色的指南 157
使用获取批量欺诈预测 AWS SDK for Python (Boto3) 158
预测解释 159
查看预测解释 160
了解预测解释是如何计算的 162
安全性 163
数据保护 163
静态加密 164
传输中加密 164
密钥管理 164
VPC 端点 (AWS PrivateLink) 166
选择退出 168
Identity and Access Management 169
受众 169

使用身份进行身份验证 170
使用策略管理访问 172
Amazon Fraud Detector 如何与 IAM 协作 174
基于身份的策略示例 177
混淆代理问题防范 185
故障排除 187
监控亚马逊 Fraud Detector 189
合规性验证 189
韧性 190
基础设施安全性 191
监控亚马逊 Fraud Detector 192
使用监控 CloudWatch 192
使用 Amazon Fraud Detector 的 CloudWatch 指标。 192
Amazon Fraud Detector 指标 195
使用记录亚马逊 Fraud Detector API 调用 AWS CloudTrail 198
Amazon Fraud Detector 中的信息 CloudTrail 198
了解 Amazon Fraud Detector 日志文件条目 199
故障排除 201
对训练数据问题进行故障排除 201
给定数据集中的欺诈率不稳定 202
数据不足 202
缺少或不同的 EVENT_LABEL 值 204
缺少或错误的 EVENT_TIMESTAMP 值 205
未摄取数据 206
变量不足 207
变量类型缺失或不正确 207
缺少变量值 207
唯一变量值不足 208
变量表达式不正确 208
唯一实体不足 210
配额 211
Amazon FFraud Detector d D 211
亚马逊Fraud Detector 检测器/变量/结果/规则 211
Amazon FFraud Detector d D 212
文档历史记录 213
..... ccxvi

什么是 Amazon Fraud Detector？

Amazon Fraud Detector 是一项完全托管的欺诈检测服务，可自动检测潜在的在线欺诈活动。这些活动包括未经授权的交易和创建虚假账户。Amazon Fraud Detector 的工作原理是使用机器学习来分析您的数据。它以亚马逊20多年欺诈检测经验丰富的专业知识为基础来实现这一目标。

您可以使用 Amazon Fraud Detector 来构建自定义的欺诈检测模型，添加决策逻辑来解释模型的欺诈评估，并为每项可能的欺诈评估分配结果，例如通过或发送以供审查。有了 Amazon Fraud Detector，您无需机器学习专业知识即可检测欺诈活动。

首先，请收集并准备您在组织中收集的欺诈数据。然后，Amazon Fraud Detector 使用这些数据代表您训练、测试和部署自定义欺诈检测模型。在此过程中，Amazon Fraud Detector 使用从中学习到欺诈模式的机器学习模型AWS和亚马逊自己的欺诈专业知识来评估您的欺诈数据并生成模型分数和模型性能数据。您可以配置决策逻辑来解释模型的分数，并分配如何处理每项欺诈评估的结果。

优势

Amazon Fraud Detector 提供以下好处。这些优势使您能够快速发现欺诈行为，而无需投入传统上构建和维护欺诈管理系统所需的时间和资源。

自动创建欺诈模型

Amazon Fraud Detector 的欺诈检测模型是为满足您的特定业务需求而定制的全自动机器学习模型。您可以使用 Amazon Fraud Detector 模型识别任何在线交易中的潜在欺诈行为，例如创建新账户、在线支付和访客结账。

由于欺诈模型是通过自动化流程创建的，因此您可以放弃与创建和训练模型相关的许多步骤。这些步骤包括数据验证和丰富、特征工程、算法选择、超参数调整和模型部署。

要使用 Amazon Fraud Detector 创建欺诈检测模型，您只需上传公司的历史欺诈数据集并选择模型类型即可。然后，Amazon Fraud Detector 会自动为您的用例找到最合适的欺诈检测算法并创建模型。您无需懂编码或拥有机器学习专业知识即可创建欺诈检测模型。

不断演变和学习的欺诈模型

欺诈检测模型必须不断发展，以跟上不断变化的欺诈格局。Amazon Fraud Detector 通过计算账户年限、自上次活动以来的时间和活动计数等信息来自动执行此操作。结果是，您的模型了解了经常进行交易的可信客户与欺诈者典型的持续尝试之间的区别。这有助于在两次再训练之间更长时间地保持模型的性能。

欺诈模型性能可视化

使用您提供的数据对模型进行训练后，Amazon Fraud Detector 会验证您的模型性能。它还提供可视化工具供您评估性能。对于您训练的每个模型，您可以看到模型性能分数、分数分布图、混淆矩阵、阈值表以及您提供的所有输入按其对模型性能的影响进行排名。使用这些性能工具，您可以了解模型的性能以及哪些输入正在推动模型性能。如果需要，您可以调整模型以提高其整体性能。

欺诈预测

Amazon Fraud Detector 会为您的组织的业务活动生成欺诈预测。欺诈预测是对业务活动的欺诈风险评估。Amazon Fraud Detector 使用预测逻辑以及与活动相关的数据生成预测。您在创建欺诈检测模型时提供了这些数据。您可以实时获取单项活动的欺诈预测，也可以离线获取一组活动的欺诈预测。

欺诈预测解释可视化

作为欺诈预测过程的一部分，Amazon Fraud Detector 会生成预测解释。预测解释可以深入了解用于训练模型的每个数据元素如何影响模型的欺诈预测分数。使用表格和图表等可视化工具提供预测解释。您可以使用这些工具直观地确定每个数据元素对预测分数的影响程度。然后，您可以使用这些信息来分析数据集中的欺诈模式并检测偏见（如果有）。最后，您还可以在手动欺诈调查过程中使用预测解释来确定主要风险指标。这可以帮助您缩小导致误报预测的根本原因。

基于规则的操作

训练完欺诈检测模型后，您可以添加规则以对评估的数据采取行动，例如接受数据、发送数据以供审查或收集更多数据。规则是指告亚马逊 Fraud Detector 在欺诈预测期间如何解释数据的条件。例如，您可以创建一条规则，将可疑客户账户标记为待审核。您可以将此规则设置为在检测到的模型分数均高于您预先确定的阈值且账户付款的授权码 (AUTH_CODE) 无效时启动。

核心概念和术语

以下是 Amazon Fraud Detector 中使用的核心概念和术语列表：

事件

事件是指贵组织的业务活动，经过欺诈风险评估。Amazon Fraud Detector 会生成事件的欺诈预测。

标签

标签将单个事件归类为欺诈事件或合法事件。标签用于在 Amazon Fraud Detector 中训练机器学习模型。

实体

实体表示正在执行事件的对象。您提供实体 ID 作为贵公司欺诈数据的一部分，以指明实施该事件的特定实体。

事件类型

事件类型定义发送到 Amazon Fraud Detector 的事件的结构。这包括作为事件一部分发送的数据、执行事件的实体（例如客户）以及对事件进行分类的标签。示例事件类型包括在线支付交易、账户注册和身份验证。

实体类型

实体类型对实体进行分类。示例分类包括客户、卖家或账户。

事件数据集

事件数据集是贵公司的特定业务活动或事件的历史数据。例如，贵公司的活动可能是在线账户注册。来自单个事件（注册）的数据可能包括关联的 IP 地址、电子邮件地址、账单地址和事件时间戳。您可以向 Amazon Fraud Detector 提供事件数据集来创建和训练欺诈检测模型。

模型

模型是机器学习算法的输出。这些算法是在代码中实现的，并根据您提供的事件数据运行。

模型类型

模型类型定义了模型训练期间使用的算法、丰富和特征转换。它还定义了训练模型的数据要求。这些定义的作用是针对特定类型的欺诈行为优化您的模型。您可以指定创建模型时要使用的模型类型。

模型训练

模型训练是使用提供的事件数据集创建可以预测欺诈事件的模型的过程。模型训练过程中的所有步骤都是完全自动化的。这些步骤包括数据验证、数据转换、特征工程、算法选择和模型优化。

模特分数

模型分数是贵公司历史欺诈数据的评估结果。在模型训练过程中，Amazon Fraud Detector 会评估数据集中的欺诈活动，并得出介于 0 到 1000 之间的分数。对于这个分数，0 代表低欺诈风险，而 1000 代表最高的欺诈风险。分数本身与误报率 (FPR) 直接相关。

模型版本

模型版本是训练模型的输出。

模型部署

模型部署是激活模型版本并使其可用于生成欺诈预测的过程。

Amazon SageMaker 模型终端节点

除了使用 Amazon Fraud Detector 构建模型外，您还可以选择在 Amazon SageMaker Fraud Detector 评估中使用托管的模型终端节点。

有关在中构建模型的更多信息 SageMaker，请参见使用训练模型Amazon SageMaker。

探测器

探测器包含检测逻辑，例如您要评估是否存在欺诈的特定事件的模型和规则。您可以使用模型版本创建探测器。

探测器版本

探测器可以有多个版本，每个版本的状态为`DraftActive`、或`Inactive`。一次只能有一个探测器版本处于`Active`状态。

Variable

变量表示与您要在欺诈预测中使用的事件关联的数据元素。变量可以作为欺诈预测的一部分随事件一起发送，也可以派生变量，例如 Amazon Fraud Detector 模型的输出或Amazon SageMaker。

规则

规则是一种条件，它告诉 Amazon Fraud Detector 在欺诈预测期间如何解释变量值。规则由一个或多个变量、一个逻辑表达式和一个或多个结果组成。规则中使用的变量必须是探测器评估的事件数据集的一部分。此外，每个探测器必须至少有一个与之关联的规则。

结果

这是欺诈预测的结果或输出。欺诈预测中使用的每条规则都必须指定一个或多个结果。

欺诈预测

欺诈预测是对单个事件或一系列事件的欺诈行为进行评估。Amazon Fraud Detector 通过同步提供模型分数和基于规则的结果，实时生成单个在线事件的欺诈预测。Amazon Fraud Detector 会为一系列离线事件生成欺诈预测。您可以使用预测进行离线操作，也可以每小时 proof-of-concept、每天或每周对欺诈风险进行回顾性评估。

欺诈预测解释

欺诈预测解释可以深入了解每个变量如何影响模型的欺诈预测分数。它提供了有关每个变量如何影响风险评分的信息，包括幅度（从 0 到 5，其中 5 为最高）和方向（使分数变高或降低）。

Amazon Fraud Detcort 的工作原

Amazon Fraud Detector 构建了一个机器学习模型，该模型经过自定义，可以检测您企业中潜在的欺诈性在线活动。要开始操作，您可以提供您的业务用例。根据您的业务用例，Amazon Fraud Detector 会推荐一种模型类型，用于为您创建欺诈检测模型。此外，它还提供有关您需要作为企业历史数据一部分提供的数据元素的见解。Amazon Fraud Detector 使用历史数据集自动为您创建和训练自定义模型。

自动模型训练过程包括选择一种机器学习算法，该算法可以针对您的特定业务用例检测欺诈行为，验证您提供的数据，以及执行数据操作以提高模型性能。训练模型后，Amazon Fraud Detector 会生成模型分数和其他模型性能指标。您可以使用分数和性能指标来评估模型性能。如果需要，您可以从为训练提供的数据集中添加或移除数据元素，并重新训练模型以提高模型分数。

在创建、训练和激活模型后，您需要配置决策逻辑（也称为规则），该逻辑告诉模型如何解释业务生成的数据，并为如何处理每项活动的解释分配结果。结果可以代表诸如批准或审查活动之类的行动，也可以代表活动的风险级别，例如高风险、中等风险和低风险。

探测器是一个容器，用于存放您的模型和相关规则。您需要创建、测试探测器并将其部署到您的生产环境中。

部署在生产环境中的探测器为您的业务应用程序提供欺诈检测功能。为了进行欺诈评估，该模型将来自您的业务活动的所有传入数据与业务的历史数据进行比较，并使用其复杂的机器学习算法和您创建的规则来分析结果并分配结果。借助 Amazon Fraud Detector，您可以实时评估来自单个业务活动的数据，也可以离线评估来自多个业务活动的数据。

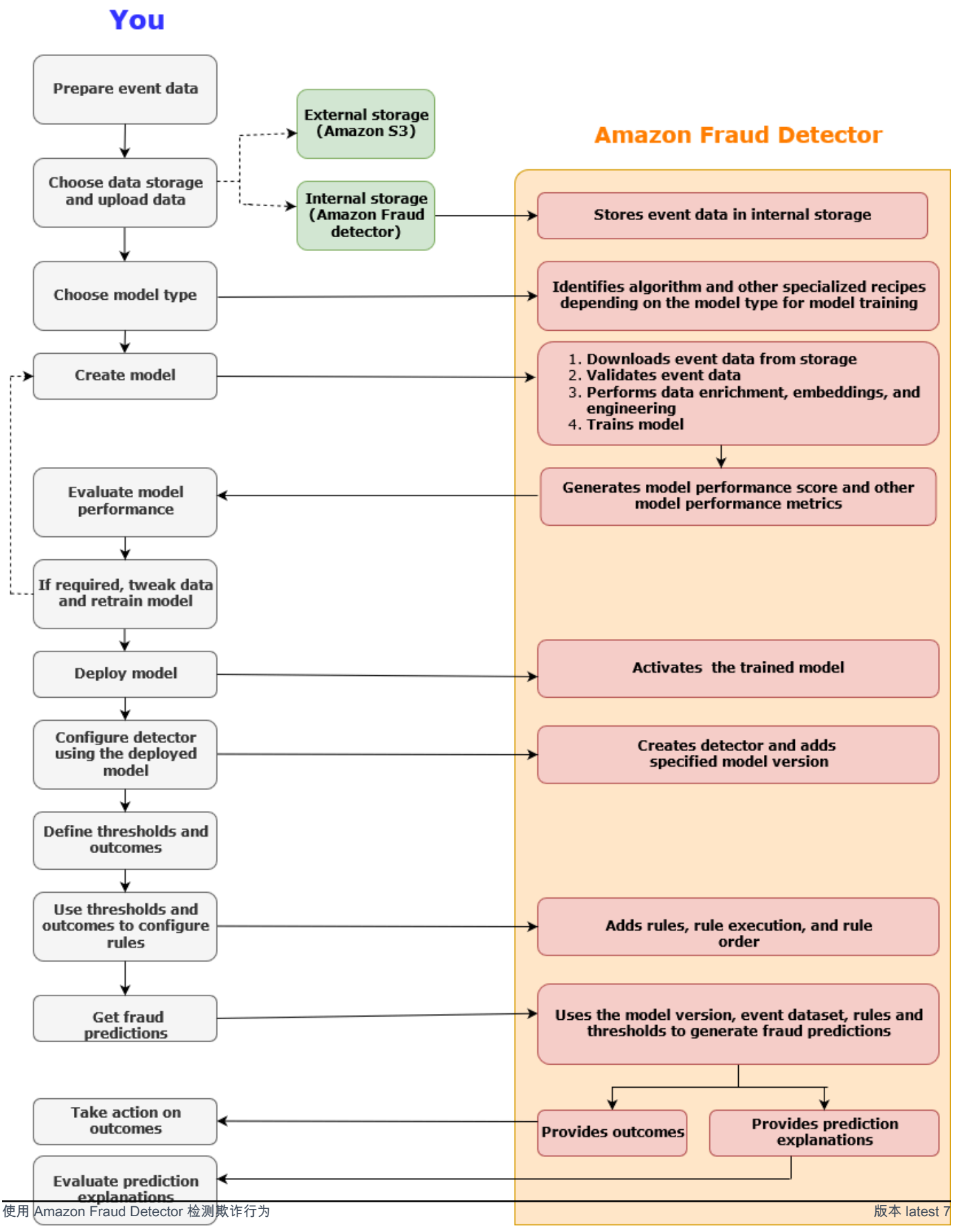
假设您有一家将在线资金转账作为其活动之一的企业。您想使用 Amazon Fraud Detector 来实时检测欺诈性转账请求。首先，您需要先向 Amazon Fraud Detector 提供过去的资金转账请求中的数据。Amazon Fraud Detector 使用这些数据来创建和训练一个模型，该模型经过自定义，可以检测欺诈性的资金转账请求。然后，您可以通过添加模型和配置模型解释数据的规则来创建探测器。在线资金转账活动规则的一个例子是，如果资金转账请求来自哪里xyz@example.com电子邮件地址，发送审核请求。在企业的生产环境中，当收到资金转账请求时，模型会分析请求附带的数据，并使用规则来分配结果。然后，您可以根据分配的结果对请求采取行动。

Amazon Fraud Detector 使用训练数据集、模型、探测器、规则和结果等组件为您的企业提供欺诈评估逻辑。

有关使用 Amazon Fraud Detector 检测欺诈时将使用的工作流程的信息，请参阅使用 Amazon Fraud Detector 检测欺诈行为

使用 Amazon Fraud Detector 检测欺诈行为

本节介绍使用 Amazon Fraud Detector 检测欺诈的典型工作流程。它还总结了如何完成这些任务。下图提供了使用 Amazon Fraud Detector 检测欺诈行为的工作流程的高级视图。


You
Prepare event data
Choose data storage and upload data
External storage (Amazon S3)
Internal storage (Amazon Fraud detector)
Choose model type
Create model
Evaluate model performance
If required, tweak data and retrain model
Deploy model
Configure detector using the deployed model
Define thresholds and outcomes
Use thresholds and outcomes to configure rules
Get fraud predictions
Take action on outcomes
Evaluate prediction explanations
Amazon Fraud Detector
Stores event data in internal storage
Identifies algorithm and other specialized recipes depending on the model type for model training
1. Downloads event data from storage
2. Validates event data
3. Performs data enrichment, embeddings, and engineering
4. Trains model
Generates model performance score and other model performance metrics
Activates the trained model
Creates detector and adds specified model version
Adds rules, rule execution, and rule order
Uses the model version, event dataset, rules and thresholds to generate fraud predictions
Provides outcomes
Provides prediction explanations

欺诈检测是一个持续的过程。部署模型后，请务必根据预测说明评估其性能分数和指标。通过这样做，您可以确定主要风险指标，缩小导致误报的根本原因，分析数据集中的欺诈模式并检测偏见（如果存在）。为了提高预测的准确性，您可以调整数据集以包含新的或修订的数据。然后，您可以使用更新的数据集重新训练模型。随着更多可用数据的出现，您可以继续重新训练模型以提高准确性。

访问亚马逊 Fraud Detector

Amazon Fraud Detector 有多种AWS 区域版本，可以通过AWS接口进行访问。

可用性

Amazon Fraud Detector 已在美国东部（弗吉尼亚北部）、美国东部（俄亥俄州）、美国西部（俄勒冈）、欧洲（爱尔兰）、亚太地区（新加坡）和亚太地区（悉尼）上市AWS 区域。

接口

您可以使用以下任何接口创建、训练、部署、测试、运行和管理欺诈检测模型和检测器：

AWS Management Console-Amazon Fraud Detector 提供了一个基于 Web 的用户界面，即亚马逊欺诈探测器控制台。如果您注册了AWS 账户，则可以访问 Amazon Fraud Detector 控制台。有关更多信息，请参阅设置 Amazon Fraud Detector。

AWS Command Line Interface(AWS CLI)-提供一个界面，您可以使用命令行外壳中的命令与包括 Amazon Fraud Detector 在内的各种用户进行交互。AWS 服务 AWS CLIAmazon Fraud Detector 的命令实现的功能与亚马逊 Fraud Detector 控制台提供的功能相同。

AWSSDK-提供特定语言的 API 并管理许多连接细节，例如签名计算、请求重试处理和错误处理。如需了解更多信息，请前往构建工具AWS页面，向下滚动到 SDK 部分，然后选择加号 (+) 展开该部分。

AWS CloudFormation-提供可用于定义 Amazon Fraud Detector 资源和属性的模板。有关更多信息，请参阅《AWS CloudFormation用户指南》中的 Amazon Fraud Detector 资源类型参考。

定价

使用 Amazon Fraud Detector，您只需为实际用量付费。没有最低费用或预付费用。我们会根据用于训练和托管模型的计算时间、使用的存储量以及您所做的欺诈预测数量向您收费。有关更多信息，请参阅 Amazon Fraud Detector 定价。

为 Amazon Fraud Detector 做好准备

要使用 Amazon Fraud Detector，您首先需要一个亚马逊网络服务 (AWS) 账户，然后必须设置允许您 AWS 账户 访问所有接口的权限。稍后，当您开始创建 Amazon Fraud Detector 资源时，您需要授予权限，允许 Amazon Fraud Detector 访问您的账户，代表您执行任务并访问您拥有的资源。

完成本节中的以下任务，为使用 Amazon Fraud Detector 做好准备：

- 注册 AWS.
- 设置允许您访问 Amazon Frau AWS 账户 d Detector 界面的权限。
- 设置你想用来访问 Amazon Fraud Detector 的接口。

完成这些步骤后，请参阅Amazon Fraud Detecto继续开始使用 Amazon Fraud Detector。

报名参加 AWS

当您注册亚马逊 Web Services (AWS) 时，系统会自动注册所有服务，包括 Amazon Fra AWS 账户 u AWS d Detector。您只需为使用的服务付费。如果您已经有了 AWS 账户，请跳到下一个任务。

注册获取 AWS 账户

如果您没有 AWS 账户，请完成以下步骤来创建一个。

要注册 AWS 账户

1. 打开 https://portal.aws.amazon.com/billing/signup。
2. 按照屏幕上的说明进行操作。

   在注册时，将接到一通电话，要求使用电话键盘输入一个验证码。

   当您注册时 AWS 账户，就会创建AWS 账户根用户一个。根用户有权访问该账户中的所有 AWS 服务 和资源。作为安全最佳实践，请为用户分配管理访问权限，并且只使用根用户来执行需要根用户访问权限的任务。

AWS 注册过程完成后会向您发送一封确认电子邮件。在任何时候，您都可以通过转至 https://aws.amazon.com/ 并选择我的账户来查看当前的账户活动并管理您的账户。

创建具有管理访问权限的用户

注册后，请保护您的安全 AWS 账户 AWS 账户根用户 AWS IAM Identity Center，启用并创建管理用户，这样您就不会使用 root 用户执行日常任务。

保护你的 AWS 账户根用户

1. 选择 Root 用户并输入您的 AWS 账户 电子邮件地址，以账户所有者的身份登录。AWS Management Console在下一页上，输入您的密码。

   要获取使用根用户登录方面的帮助，请参阅《AWS 登录 用户指南》中的以根用户身份登录。
2. 为您的根用户启用多重身份验证 (MFA)。

   有关说明，请参阅 I A M 用户指南中的为 AWS 账户 根用户启用虚拟 MFA 设备（控制台）。

创建具有管理访问权限的用户

1. 启用 IAM Identity Center

   有关说明，请参阅《AWS IAM Identity Center 用户指南》中的启用 AWS IAM Identity Center。
2. 在 IAM Identity Center 中，为用户授予管理访问权限。

   有关使用 IAM Identity Center 目录 作为身份源的教程，请参阅《用户指南》 IAM Identity Center 目录中的使用默认设置配置AWS IAM Identity Center 用户访问权限。

以具有管理访问权限的用户身份登录

- 要使用您的 IAM Identity Center 用户身份登录，请使用您在创建 IAM Identity Center 用户时发送到您的电子邮件地址的登录网址。

  有关使用 IAM Identity Center 用户登录的帮助，请参阅AWS 登录 用户指南中的登录 AWS 访问门户。

将访问权限分配给其他用户

1. 在 IAM Identity Center 中，创建一个权限集，该权限集遵循应用最低权限的最佳做法。

   有关说明，请参阅《AWS IAM Identity Center 用户指南》中的创建权限集。
2. 将用户分配到一个组，然后为该组分配单点登录访问权限。

有关说明，请参阅《AWS IAM Identity Center 用户指南》中的添加组。

设置访问亚马逊 Fraud Detector 界面的权限

要使用亚马逊 Fraud Detector，请设置访问亚马逊 Fraud Detector 控制台和 API 操作的权限。

按照安全最佳实践，创建一个 AWS Identity and Access Management (IAM) 用户，其访问权限仅限于 Amazon Fraud Detector 的操作，并具有所需权限。您可以根据需要添加其他权限。

以下政策提供了使用 Amazon Fraud Detector 所需的权限：

- `AmazonFraudDetectorFullAccessPolicy`

  您可以执行以下操作：
  - 访问所有 Amazon Fraud Detector 资源
  - 列出并描述中的所有模型端点 SageMaker
  - 列出账户中的所有 IAM 角色
  - 列出所有 Amazon S3 存储桶
  - 允许 IAM Pass 角色将角色传递给 Amazon Fraud Detector
- `AmazonS3FullAccess`

  允许完全访问 Amazon Simple Storage Service。如果您需要将训练数据集上传到 Amazon S3，则需要这样做。

下面介绍如何创建 IAM 用户并分配所需权限。

创建用户并分配所需权限

1. 登录 AWS Management Console 并打开 IAM 控制台，网址为 https://console.aws.amazon.com/iam/。
2. 在导航窗格中，选择用户，然后选择添加用户。
3. 对于 User name（用户名），输入 **`AmazonFraudDetectorUser`**。
4. 选中 “AWS 管理控制台访问权限” 复选框，然后配置用户密码。
5. （可选）默认情况下，AWS 要求新用户在首次登录时创建新密码。您可以清除 User must create a new password at next sign-in（用户必须在下次登录时创建新密码）旁边的复选框以允许新用户在登录后重置其密码。

6. 选择下一步: 权限。
7. 选择创建组。
8. 在组名中输入**AmazonFraudDetectorGroup**。
9. 在政策列表中，选中AmazonFraudDetectorFullAccessPolicy和 AmazonS3 FullAccess 的复选框。选择创建组。
10. 在组列表中，选中您的新组所对应的复选框。如果您在列表中看不到该群组，请选择 “刷新”。
11. 选择下一步: 标签。
12. （可选）通过以键值对的形式附加标签来向用户添加元数据。有关如何在 IAM 中使用标签的说明，请参阅标记 IAM 用户和角色。
13. 选择 “下一步：查看” 以查看新用户的用户详细信息和权限摘要。如果您已准备好继续，请选择创建用户。

设置用于访问 Amazon Fraud Detector 的接口

你可以使用 Amazon Fraud Detector 控制台或 AWS SDK 访问亚马逊 Frau AWS CLI d Detector。在使用它们之前，请先设置 AWS CLI 和 S AWS DK。

访问亚马逊 Fraud Detector 控制台

您可以通过访问 Amazon Fraud Detector 控制台和其他 AWS 服务 AWS Management Console。您的 AWS 账户，授予您访问权限 AWS Management Console。

要访问亚马逊 Fraud Detector 控制台，

1. 前往https://console.aws.amazon.com/并登录您的 AWS 账户。
2. 导航到亚马逊 Fraud Detector。

借助 Amazon Fraud Detector 控制台，您可以创建和管理您的模型以及欺诈检测资源，例如探测器、变量、事件、实体、标签和结果。您可以生成预测并评估模型的性能和预测。

设置 AWS CLI

您可以使用 AWS Command Line Interface (AWS CLI) 通过在命令行外壳中运行命令来与 Amazon Fraud Detector 进行交互。只需最少的配置，您就可以在终端的命令提示符下使用命令 AWS CLI 来运行与 Amazon Fraud Detector 控制台提供的功能类似的命令。

要设置 AWS CLI

下载并配置 AWS CLI。有关说明，请参阅《 AWS Command Line Interface 用户指南》中的以下主题：

- 使用 AWS 命令行界面进行设置
- 配置 AWS 命令行界面

有关 Amazon Fraud Detector 命令的信息，请参阅可用命令

设置 AWS SDK

您可以使用 AWS 软件开发工具包编写用于创建和管理欺诈检测资源以及获取欺诈预测的代码。这些软件开发 AWS 工具包支持 Amazon Fraud Detector JavaScript和 Python (Boto3) 。

要设置 AWS SDK for Python (Boto3)

您可以使用 AWS SDK for Python (Boto3) 创建、配置和管理 AWS 服务。有关如何安装 Boto 的说明，请参阅AWS 适用于 Python 的 SDK (Boto3)。确保你使用的是 Boto3 SDK 版本 1.14.29 或更高版本。

安装后 AWS SDK for Python (Boto3)，运行以下 Python 示例以确认您的环境配置正确。如果配置正确，则响应将包含探测器列表。如果未创建探测器，则列表为空。

```
import boto3
fraudDetector = boto3.client('frauddetector')

response = fraudDetector.get_detectors()
print(response)
```

设置适用于 Java 的 AWS 软件开发工具包

有关如何安装和加载的说明 AWS SDK for JavaScript，请参阅为设置软件开发工具包 JavaScript。

Amazon Fraud Detecto

开始之前，请确保您已阅读使用 Amazon Fraud Detector 检测欺诈行为并完成中的步骤为 Amazon Fraud Detector 做好准备。

使用本节中的实践教程可以了解如何使用 Amazon Fraud Detector 构建、训练和部署欺诈检测模型。在本教程中，您将扮演欺诈分析师的角色，使用机器学习模型来预测新账户注册是否具有欺诈性。必须使用账户注册中的数据对模型进行训练。Amazon Fraud Detector 为本教程提供了账户注册数据集示例。在开始教程之前，必须上传示例数据集。

您可以使用以下界面之一开始使用 Amazon Fraud Detector。在开始本教程之前，请确保您按照说明操作获取并上传示例数据集

- 教程：开始使用亚马逊Fraud Detector 控制台
- 教程：入门AWS SDK for Python (Boto3)

获取并上传示例数据集

您在本教程中使用的示例数据集提供了在线账户注册的详细信息。数据集位于使用 UTF-8 格式的逗号分隔值 (CSV) 的文本文件中。CSV 数据集文件的第一行包含标题。标题行后面有多行数据。这些行中的每一行都由来自单个账户注册的数据元素组成。为方便起见，我们将对数据进行了标记。数据集中的一列用于标识帐户注册是否是欺诈性的。

获取和上传示例数据集

1. 转到样品。

   有两个包含在线账户注册数据的数据文件—— registration_data_20K_minimum.csv 和 registration_data_20K_full.csv。该文件仅`registration_data_20K_minimum`包含两个变量：ip_address 和 e mail_address。该文件`registration_data_20K_full`包含其他变量。这些变量适用于每个事件，包括账单地址、电话号码和用户代理。这两个数据文件还包含两个必填字段：

   - 事件_TIMESTAMPHED — 定义事件发生的时间
   - EVENT_LABEL — 将事件归类为欺诈事件或合法事件

   在本教程中，您可以使用这两个文件中的任何一个。下载要使用的数据文件。

2. 创建Amazon Simple Storage Service（Amazon S3）存储桶。

   在此步骤中，您将创建用于存储数据集的外部存储。此外部存储是 Amazon S3 存储桶。有关 Amazon S3 的更多信息，请参阅 Amazon S3 是什么？

   a. 登录到AWS Management Console，然后通过以下网址打开 Amazon S3 控制台：https://console.aws.amazon.com/s3/。

   b. 在存储分区中，选择创建存储桶。

   c. 对于存储桶名称，请输入存储桶的名称。确保遵守控制台中的存储桶命名规则，并提供全局唯一的名称。我们建议您使用描述存储桶用途的名称。

   d. 对于 AWS 区域，请选择要AWS 区域在哪里创建存储桶。您选择的地区必须支持亚马逊Fraud Detector。要减少延迟AWS 区域，请选择最接近您的地理位置的。有关支持 Amazon Fraud Detector 的区域列表，请参阅全球基础设施指南中的区域表。

   e. 在本教程中，保留对象所有权的默认设置、封锁公共访问的存储段设置、存储段版本控制和标签。

   f. 对于默认加密，在本教程中选择禁用。

   g. 查看您的存储桶配置，然后选择 C reate bucket。

3. 将示例数据文件上载到 Amazon S3 存储桶。

   现在您有了存储桶，请将您之前下载的示例文件之一上传到您刚刚创建的 Amazon S3 存储桶。

   a. 在存储桶中，列出了您的存储段名称。选择存储桶。

   b. 请选择 Upload（上传）。

   c. 在文件和文件夹中，选择添加文件。

   d. 选择您在计算机上下载的示例数据文件之一，然后选择 “打开”。

   e. 保留目标、权限和属性的默认设置。

   f. 查看配置，然后选择上传。

   g. 示例数据文件已上载到 Amazon S3 存储桶中。请记下存储桶位置。在对象中，选择您刚刚上传的示例数据文件。

   h. 在对象概述中，复制 S 3 URI 下的位置。这是您的示例数据文件的 Amazon S3 位置。稍后您将使用它。此外，您还可以复制 S3 存储桶的 Amazon 资源名称 (ARN) 并将其保存。

教程：开始使用亚马逊Fraud Detector 控制台

本教程由两部分组成。第一部分描述如何构建、训练和部署欺诈检测模型。第二部分介绍如何使用该模型实时生成欺诈预测。使用您上传到 S3 存储桶的示例数据文件对模型进行训练。在本教程结束时，您将完成以下操作：

- 构建和训练亚马逊Fraud Detector 模型
- 生成实时欺诈预测

> **Important**
>
> 继续之前，请确保您已按照说明操作获取并上传示例数据集

A 部分：构建、训练和部署亚马逊Fraud Detector 模型

在 A 部分中，您可以定义业务用例、定义事件、构建模型、训练模型、评估模型的性能并部署模型。

步骤 1：选择您的业务用例

- 在此步骤中，您将使用数据模型浏览器将您的业务用例与 Amazon Fraud Detector 支持的欺诈检测模型类型进行匹配。Data Models Explorer 是一款与 Amazon Fraud Detector 控制台集成的工具，可推荐一种模型类型，用于为您的业务用例创建和训练欺诈检测模型。数据模型浏览器还可让您深入了解需要在数据集中包含的必填数据、推荐数据元素和可选数据元素。该数据集将用于创建和训练您的欺诈检测模型。

  在本教程中，您的业务用例是新账户注册。在您指定业务用例后，数据模型浏览器将推荐用于创建欺诈检测模型的模型类型，还将为您提供创建数据集所需的数据元素列表。由于您已经上传了包含新账户注册数据的示例数据集，因此无需创建新的数据集。

  a. 打开AWS管理控制台并登录您的账户。导航到 Amazon Fraud Detec
  b. 在左侧导航窗格中，选择数据模型资源管理器。
  c. 在数据模型浏览器页面的业务用例下，选择新账户欺诈。
  d. Amazon Fraud Detector 显示推荐的模型类型，用于为所选业务用例创建欺诈检测模型。模型类型定义了 Amazon Fraud Detector 将用于训练您的欺诈检测模型的算法、丰富和转换。

     请记下推荐的模型类型。稍后在创建模型时将需要此信息。
  e. 数据模型见解窗格可让您深入了解创建和训练欺诈检测模型所需的必需和推荐数据元素。

查看您下载的示例数据集，确保其中包含表中列出的所有必填数据元素和一些推荐的数据元素。

稍后，当您为特定业务用例创建模型时，您将使用提供的见解来创建数据集。

步骤 2：创建事件类型

- 在此步骤中，您可以定义要评估欺诈的业务活动（事件）。定义事件涉及设置数据集中的变量、实体启动事件以及对事件进行分类的标签。在本教程中，您将定义账户注册事件。

  a. 打开AWS管理控制台并登录您的账户。导航到 Amazon Fraud Detec

  b. 在左侧导航窗格中，选择 Events（事件）。

  c. 在事件类型页面中，选择创建。

  d. 在事件类型详细信息下，输入`sample_registration`作为事件类型名称，并根据需要输入事件的描述。

  e. 对于实体，选择创建实体。

  f. 在创建实体页面中，输入`sample_customer`作为实体类型名称。或者，输入实体类型的描述。

  g. 选择 Create entity (创建实体)。

  h. 在事件变量下，对于选择如何定义此事件的变量，选择从训练数据集中选择变量。

  i. 对于 IAM 角色，选择创建 IAM 角色。

  j. 在创建 IAM 角色页面中，输入您将示例数据上传到的 S3 存储桶的名称，然后选择创建角色。

  k. 在数据位置中，输入示例数据的路径。这是您在上传示例数据后保存的`S3 URI`路径。路径与此类似:`S3://your-bucket-name/example dataset filename.csv`.

  l. 请选择 Upload（上传）。

     Amazon Fraud Detector 从您的示例数据文件中提取标题并将其映射为变量类型。映射会显示在控制台中。

  m. 在 “标签-可选” 下的 “标签” 中，选择 “创建新标签”。

  n. 在创建标签页面中，输入`fraud`作为名称。此标签对应于代表示例数据集中欺诈性账户注册的值。

  o. 选择 “创建标签”。

p. 创建第二个标签，然后输入`legit`作为名称。此标签对应于代表示例数据集中合法账户注册的值。

q. 选择创建事件类型。

步骤 3：创建模型

1. 在模型页面上，选择添加模型，然后选择创建模型。
2. 对于步骤 1-定义模型细节，输入`sample_fraud_detection_model`作为模型名称。或者，添加模型的描述。
3. 对于模型类型，选择在线欺诈洞察模型。
4. 对于事件类型，选择 sample_registration。这是您在步骤 1 中创建的事件类型。
5. 在历史事件数据中，

   a. 在事件数据源中，选择存储在 S3 中的事件数据。

   b. 对于 IAM 角色，选择您在步骤 1 中创建的角色。

   c. 在训练数据位置中，输入示例数据文件的 S3 URI 路径。
6. 选择下一步。

步骤 4：火车模型

1. 在模型输入中，将所有复选框保持选中状态。默认情况下，Amazon Fraud Detector 使用历史事件数据集中的所有变量作为模型输入。
2. 在标签分类中，对于欺诈标签，选择欺诈，因为此标签对应于代表示例数据集中欺诈事件的值。对于合法标签，请选择 legit，因为此标签对应于代表示例数据集中合法事件的值。
3. 对于 “未标注事件” 处理，保留此示例数据集的默认选择 “忽略未标记的事件”。
4. 选择下一步。
5. 查看后，选择创建并训练模型。Amazon Fraud Detector 创建了一个模型并开始训练该模型的新版本。

   在模型版本中，状态列表示模型训练的状态。使用示例数据集的模型训练大约需要 45 分钟才能完成。模型训练完成后，状态更改为 “准备部署”。

步骤 5：查看模型性能

使用 Amazon Fraud Detector 的一个重要步骤是使用模型分数和性能指标评估模型的准确性。模型训练完成后，Amazon Fraud Detector 使用未用于训练模型的 15% 的数据来验证模型性能，并生成模型性能分数和其他性能指标。

1. 要查看模型的性能，

   a. 在 Amazon Fraud Detector 的左侧导航窗格中，选择模型。

   b. 在模型页面中，选择你刚刚训练的模型（sample_fraud_detection_model），然后选择 1.0。这是亚马逊Fraud Detector 为您的模型创建的版本。

2. 查看模型性能总分以及亚马逊Fraud Detector 为此模型生成的所有其他指标。

   要了解有关此页面上模型性能分数和性能指标的更多信息，请参阅模型分数和对性能指标进行建模。

   您可以期望所有经过训练的 Amazon Fraud Detector 模型都具有真实的欺诈检测性能指标，这些指标与您在本教程中看到的模型的性能指标类似。

步骤 6：部署模型

在您查看了经过训练的模型的性能指标并准备好使用它来生成欺诈预测之后，就可以部署该模型了。

1. 在亚马逊Fraud Detector 控制台的左侧导航窗格中，选择模型。
2. 在模型页面中，选择 sample_fraud_detection_m odel，然后选择要部署的特定模型版本。在本教程中，选择 1.0。
3. 在模型版本页面上，选择操作，然后选择部署模型版本。
4. 在模型版本中，状态显示部署的状态。部署完成后，状态更改为 “活动”。这表明模型版本已激活并可用于生成欺诈预测。B 部分：生成欺诈预测继续完成生成欺诈预测的步骤。

B 部分：生成欺诈预测

欺诈预测是对商业活动（事件）欺诈的评估。亚马逊欺诈检测器使用检测器生成欺诈预测。检测器包含您想要评估是否存在欺诈的特定事件的检测逻辑，例如模型和规则。检测逻辑使用规则告诉 Amazon Fraud Detector 如何解释与模型相关的数据。在本教程中，您将使用之前上传的账户注册示例数据集评估账户注册事件。

在第 A 部分中，您创建、训练和部署了模型。在 B 部分中，您为sample_registration事件类型构建检测器，添加已部署的模型，创建规则和规则执行顺序，然后创建并激活用于生成欺诈预测的检测器版本。

步骤 1：构建探测器

创建探测器

1. 在 Amazon Fraud Detector 的左侧导航窗格中，选择 Setec to r。
2. 选择 “创建探测器”。
3. 在定义探测器详细信息页面中，输入sample_detector探测器名称。或者，输入探测器的描述，例如my sample fraud detector。
4. 对于事件类型，选择 sample_registration。这是您在本教程的 A 部分中创建的事件。
5. 选择下一步。

步骤 2：添加模型

如果您完成了本教程的 A 部分，那么您可能已经有了 Amazon Fraud Detector 模型可以添加到您的检测器中。如果您尚未创建模型，请转到第 A 部分并完成创建、训练和部署模型的步骤，然后继续执行第 B 部分。

1. 在添加模型-可选中，选择添加模型。
2. 在添加模型页面的选择模型中，选择您之前部署的 Amazon Fraud Detector 模型名称。对于选择版本，选择已部署模型的模型版本。
3. 选择 Add model (添加模型)。
4. 选择下一步。

步骤 3：添加规则

规则是指在评估欺诈预测时指示 Amazon Fraud Detector 如何解释模型性能分数的条件。在本教程中，您将创建三条规则：high_fraud_riskmedium_fraud_risk、和low_fraud_risk。

1. 在 “添加规则” 页的 “定义规则” 下，输入high_fraud_risk规则名称，在 “描述-可选” 下输入**This rule captures events with a high ML model score**规则的描述。
2. 在 E x pression 中，使用 Amazon Fraud Detector 简化规则表达式语言输入以下规则表达式：

   ```
   $sample_fraud_detection_model_insightscore > 900
   ```

3. 在结果中，选择创建新结果。结果是欺诈预测的结果，如果在评估期间规则匹配，则返回结果。
4. 在创建新结果中，输入verify_customer作为结果名称。或者，输入描述。
5. 选择 “保存结果”。
6. 选择 “添加规则” 以运行规则验证检查器并保存规则。创建后，Amazon Fraud Detector 会使该规则可在您的检测器中使用。
7. 选择 “添加其他规则”，然后选择 “创建规则” 选项卡。
8. 再次重复此过程两次，使用以下low_fraud_risk规则详细信息创建您的medium_fraud_risk和规则：

   - 中等欺诈风险

     规则名称：medium_fraud_risk

     结果：review

     表达式：

     ```
     $sample_fraud_detection_model_insightscore <= 900 and
     ```

     ```
     $sample_fraud_detection_model_insightscore > 700
     ```

   - 低欺诈风险

     规则名称：low_fraud_risk

     结果：approve

     表达式：

     ```
     $sample_fraud_detection_model_insightscore <= 700
     ```

   这些值是本教程使用的示例。为自己的探测器创建规则时，请使用适合您的模型和用例的值，

9. 创建所有三条规则后，选择 “下一步”。

   有关创建和编写规则的更多信息，请参阅规则和规则语言参考。

步骤 4：配置规则执行和规则顺序

检测器中包含的规则的规则执行模式决定是否对您定义的所有规则进行评估，或者规则评估是否在第一个匹配的规则处停止。规则顺序决定了您希望规则的运行顺序。

默认规则执行模式为FIRST_MATCHED。

第一次匹配

第一个匹配的规则执行模式根据定义的规则顺序返回第一个匹配规则的结果。如果指定FIRST_MATCHED，Amazon Fraud Detector 会按顺序评估规则，从第一个到最后一个，在第一个匹配的规则处停止。然后，Amazon Fraud Detector 提供该单一规则的结果。

您运行规则的顺序可能会影响生成的欺诈预测结果。创建规则后，按照以下步骤对规则进行重新排序，使其按所需顺序运行：

如果您的high_fraud_risk规则还不在规则列表的顶部，请选择顺序，然后选择 1。这将移high_fraud_risk至第一个位置。

重复此过程，使您的medium_fraud_risk规则位于第二位置，而您的low_fraud_risk规则位于第三位置。

全部匹配

无论规则顺序如何，所有匹配的规则执行模式都会返回所有匹配规则的结果。如果指定ALL_MATCHED，Amazon Fraud Detector 会评估所有规则并返回所有匹配规则的结果。

选择FIRST_MATCHED本教程，然后选择 “下一步”。

步骤 5：查看并创建探测器版本

检测器版本定义了用于生成欺诈预测的特定模型和规则。

1. 在查看和创建页面中，查看您配置的探测器详细信息、模型和规则。如果需要进行任何更改，请选择相应部分旁边的 Edit（编辑）。
2. 选择 “创建探测器”。创建后，探测器的第一个版本将显示在探测器版本表中，并显示其Draft状态。

   您可以使用草稿版本来测试您的探测器。

步骤 6：测试并激活探测器版本

在 Amazon Fraud Detector 控制台中，您可以使用带有运行测试功能的模拟数据来测试检测器的逻辑。在本教程中，您可以使用示例数据集中的账户注册数据。

1. 滚动至 Det ector 版本详细信息页面底部的运行测试。
2. 对于事件元数据，输入事件发生时间的时间戳，然后输入执行事件的实体的唯一标识符。在本教程中，从日期选择器中选择一个日期作为时间戳，然后输入 “1234” 作为实体 ID。
3. 对于事件变量，输入要测试的变量值。在本教程中，您只需要`ip_address`和`email_address`字段。这是因为它们是用于训练您的亚马逊Fraud Detector 模型的输入。您可以使用以下示例值。这假设你使用了建议的变量名：

   - IP 地址：`205.251.233.178`
   - 电子邮件地址：`johndoe@exampledomain.com`
4. 选择 “运行测试”。
5. Amazon Fraud Detector 根据规则执行模式返回欺诈预测结果。如果规则执行模式为`FIRST_MATCHED`，则返回的结果对应于第一个匹配的规则。第一条规则是优先级最高的规则。如果评估为真，则表示匹配。如果规则执行模式为`ALL_MATCHED`，则返回的结果对应于所有匹配的规则。这意味着它们都被评估为真实。Amazon Fraud Detector 还会返回添加到检测器中的任何模型的模型分数。

   您可以更改输入并运行几次测试以查看不同的结果。您可以使用示例数据集中的 ip_ad dress 和 e mail_ address 值进行测试，并检查结果是否符合预期。
6. 当你对探测器的工作方式感到满意时，将其从提升`Draft`到`Active`。这样做可以使检测器可用于实时欺诈检测。

   在 Det ector 版本详细信息页面上，选择操作、发布、发布版本。这会将探测器的状态从 “草稿” 更改为 “活动”。

   此时，您的模型和相关的检测器逻辑已准备就绪，可以使用 Amazon Fraud DetectorGetEventPrediction API 实时评估在线活动是否存在欺诈。您也可以使用 CSV 输入文件和`CreateBatchPredictionJob` API 离线评估事件。有关欺诈预测的更多信息，请参阅欺诈预测

完成本教程后，您执行了以下操作：

- 将示例事件数据集上载到 Amazon S3。

- 使用示例数据集创建并训练了 Amazon Fraud Detector 欺诈检测模型。
- 查看了 Amazon Fraud Detector 生成的模型性能分数和其他性能指标。
- 部署了欺诈检测模型。
- 创建了探测器并添加了部署的模型。
- 向检测器添加了规则、规则执行顺序和结果。
- 通过提供不同的输入并检查规则和规则执行顺序是否按预期运行来测试检测器。
- 通过发布探测器将其激活。

教程：入门AWS SDK for Python (Boto3)

本教程介绍如何构建和训练 Amazon Fraud Detector 模型，然后使用该模型使用生成实时欺诈预测 AWS SDK for Python (Boto3)。使用您上传到 Amazon S3 存储桶的账户注册示例数据文件对模型进行训练。

在本教程结束时，您将完成以下操作：

- 构建和训练亚马逊Fraud Detector 模型
- 生成实时欺诈预测

先决条件

以下是本教程的先决步骤。

- 已完成为 Amazon Fraud Detector 做好准备。

  如果你已经使用了设置 AWS SDK，请确保你使用的是 Boto3 SDK 版本 1.14.29 或更高版本。
- 按照说明提交了本教程所需的获取并上传示例数据集文件。

开始使用

步骤 1：设置和验证 Python 环境

Boto 是适用于 Python 的Amazon Web Services (AWS) 软件开发工具包。您可以使用它来创建、配置和管理AWS 服务。有关如何安装 Boto3 的说明，请参阅 AWS SDK for Python (Boto3)。

安装后AWS SDK for Python (Boto3)，运行以下 Python 示例命令以确认您的环境配置正确。如果您的环境配置正确，则响应包含探测器列表。如果未创建探测器，该列表是空的。

```
import boto3
fraudDetector = boto3.client('frauddetector')

response = fraudDetector.get_detectors()
print(response)
```

步骤 2：创建变量、实体类型和标签

在此步骤中，您将创建用于定义模型、事件和规则的资源。

创建变量

变量是数据集中的数据元素，您要使用它来创建事件类型、模型和规则。

在以下示例中，CreateVariableAPI 用于创建两个变量。变量是email_address和ip_address。将它们分配给相应的变量类型：EMAIL_ADDRESS和IP_ADDRESS。这些变量是您上传的示例数据集的一部分。当您指定变量类型时，Amazon Fraud Detector 将在模型训练和获得预测时解释变量。只有具有关联变量类型的变量才能用于模型训练。

```
import boto3
fraudDetector = boto3.client('frauddetector')

 #Create variable email_address
fraudDetector.create_variable(
    name = 'email_address',
    variableType = 'EMAIL_ADDRESS',
    dataSource = 'EVENT',
    dataType = 'STRING',
    defaultValue = '<unknown>'
)

#Create variable ip_address
fraudDetector.create_variable(
    name = 'ip_address',
    variableType = 'IP_ADDRESS',
    dataSource = 'EVENT',
    dataType = 'STRING',
    defaultValue = '<unknown>'
)
```

创建实体类型

实体表示正在执行事件的对象，实体类型将实体分类。示例分类包括客户、卖家或账户。

在以下示例中，PutEntityTypeAPI 用于创建sample_customer实体类型。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.put_entity_type(
   name = 'sample_customer',
   description = 'sample customer entity type'
)
```

创建标签

标签将事件归类为欺诈事件或合法事件，并用于训练欺诈探测模型。模型学会使用这些标签值对事件进行分类。

在以下示例中，Putlabel API 用于创建两个标签、fraud和legit。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.put_label(
    name = 'fraud',
    description = 'label for fraud events'
)

fraudDetector.put_label(
    name = 'legit',
    description = 'label for legitimate events'
)
```

步骤 3：创建事件类型

使用 Amazon Fraud Detector，您可以构建模型来评估风险并针对单个事件生成欺诈预测。事件类型定义单个事件的结构。

在以下示例中，PutEventTypeAPI 用于创建事件类型sample_registration。您可以通过指定在上一步中创建的变量 (,ip_addresssample_customer)email_address、实体类型 () 和标签 (fraud,legit) 来定义事件类型。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.put_event_type (
     name = 'sample_registration',
     eventVariables = ['ip_address', 'email_address'],
     labels = ['legit', 'fraud'],
     entityTypes = ['sample_customer'])
```

步骤 4：创建、训练和部署模型

Amazon Fraud Detector 训练模型学习检测特定事件类型的欺诈。在上一步中，您创建了事件类型。在此步骤中，您将为事件类型创建和训练模型。该模型充当模型版本的容器。每次训练模型时，将会创建新的版本。

使用以下示例代码创建和训练在线欺诈洞察模型。这个模型叫做sample_fraud_detection_model。它适用于sample_registration使用您上传到 Amazon S3 的账户注册示例数据集的事件类型。

有关 Amazon Fraud Detector 支持的不同模型类型的更多信息，请参阅选择模型类型。

创建模型

在以下示例中，CreateModelAPI 用于创建模型。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.create_model (
       modelId = 'sample_fraud_detection_model',
       eventTypeName = 'sample_registration',
       modelType = 'ONLINE_FRAUD_INSIGHTS')
```

训练模型

在以下示例中，CreateModelVersionAPI 用于训练模型。'EXTERNAL_EVENTS'为trainingDataSource和指定存储示例数据集的 Amazon S3 位置以及 Amazon S3 存储桶RoleArn的存储位置externalEventsDetail。对于trainingDataSchema参数，请指定 Amazon Fraud Detector 如何解释示例数据。更具体地说，指定要包含哪些变量以及如何对事件标签进行分类。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.create_model_version (
         modelId = 'sample_fraud_detection_model',
         modelType = 'ONLINE_FRAUD_INSIGHTS',
         trainingDataSource = 'EXTERNAL_EVENTS',
         trainingDataSchema = {
            'modelVariables' : ['ip_address', 'email_address'],
            'labelSchema' : {
               'labelMapper' : {
                   'FRAUD' : ['fraud'],
                   'LEGIT' : ['legit']
        }
    }
},
         externalEventsDetail = {
              'dataLocation' : 's3://your-S3-bucket-name/your-example-data-
filename.csv',
              'dataAccessRoleArn' : 'role_arn'
}
)
```

您可以对模型进行多次训练。每次训练模型时，将会创建新的版本。模型训练完成后，模型版本状态更新为TRAINING_COMPLETE。您可以查看模型性能分数和其他模型性能指标。

查看模型性能

使用 Amazon Fraud Detector 的一个重要步骤是使用模型分数和性能指标评估模型的准确性。模型训练完成后，Amazon Fraud Detector 使用未用于训练模型的 15% 的数据来验证模型性能。它会生成模型性能分数和其他性能指标。

使用 DescribeModelVersionsAPI 查看模型性能。查看模型性能总分以及亚马逊Fraud Detector 针对该模型生成的所有其他指标。

要了解有关模型性能分数和性能指标的更多信息，请参阅模型分数和对性能指标进行建模。

您可以期望所有经过训练的 Amazon Fraud Detector 模型都具有真实的欺诈检测性能指标，这些指标与本教程中的指标类似。

部署模型

查看训练模型的性能指标后，部署模型并将其提供给 Amazon Fraud Detector 以生成欺诈预测。要部署经过训练的模型，请使用 UpdateModelVersionStatusAPI。在以下示例中，它用于将模型版本状态更新为 ACTIVE。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.update_model_version_status (
     modelId = 'sample_fraud_detection_model',
     modelType = 'ONLINE_FRAUD_INSIGHTS',
     modelVersionNumber = '1.00',
     status = 'ACTIVE'
)
```

步骤 5：创建探测器、结果、规则和探测器版本

探测器包含检测逻辑，例如模型和规则。此逻辑适用于您想要评估是否存在欺诈的特定事件。规则是指在预测期间指示 Amazon Fraud Detector 如何解释变量值的条件。而结果是欺诈预测的结果。探测器可以有多个版本，每个版本的状态为 DRAFT、ACTI V E 或 INAC TI VE。探测器版本必须至少有一条与其关联的规则。

使用以下示例代码创建检测器、规则、结果并发布检测器。

创建探测器

在以下示例中，PutDetectorAPI 用于为sample_registration事件类型创建sample_detector检测器。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.put_detector (
     detectorId = 'sample_detector',
     eventTypeName = 'sample_registration'
)
```

创造成果

为每个可能的欺诈预测结果创建结果。在以下示例中，PutOutcomeAPI 用于创建三个结果-verify_customerreview、和approve。这些结果随后被分配给规则。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.put_outcome(
     name = 'verify_customer',
     description = 'this outcome initiates a verification workflow'
    )

fraudDetector.put_outcome(
     name = 'review',
     description = 'this outcome sidelines event for review'
    )

fraudDetector.put_outcome(
     name = 'approve',
     description = 'this outcome approves the event'
)
```

创建规则

规则由数据集中的一个或多个变量、逻辑表达式以及一个或多个结果组成。

在以下示例中，CreateRuleAPI 用于创建三种不同的规则：high_riskmedium_risk、和low_risk。创建规则表达式以将模型性能分sample_fraud_detection_model_insightscore值与各种阈值进行比较。这是为了确定事件的风险级别并分配在上一步中定义的结果。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.create_rule(
     ruleId = 'high_fraud_risk',
     detectorId = 'sample_detector',
```

```
     expression = '$sample_fraud_detection_model_insightscore > 900',
     language = 'DETECTORPL',
     outcomes = ['verify_customer']
     )

fraudDetector.create_rule(
     ruleId = 'medium_fraud_risk',
     detectorId = 'sample_detector',
     expression = '$sample_fraud_detection_model_insightscore <= 900 and
 $sample_fraud_detection_model_insightscore > 700',
     language = 'DETECTORPL',
     outcomes = ['review']
     )

fraudDetector.create_rule(
     ruleId = 'low_fraud_risk',
     detectorId = 'sample_detector',
     expression = '$sample_fraud_detection_model_insightscore <= 700',
     language = 'DETECTORPL',
     outcomes = ['approve']
     )
```

创建探测器版本

检测器版本定义了用于进行欺诈预测的模型和规则。

在以下示例中，CreateDetectorVersionAPI 用于创建探测器版本。它通过提供模型版本详细信息、规则和规则执行模式 FIRST_MATCHED 来实现此目的。规则执行模式指定了评估规则的顺序。规则执行模式 FIRST_MATCHED 指在第一个匹配的规则处停止，按顺序评估规则，从第一个到最后一个，在第一个匹配的规则处停止。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.create_detector_version(
      detectorId = 'sample_detector',
      rules = [{
          'detectorId' : 'sample_detector',
          'ruleId' : 'high_fraud_risk',
          'ruleVersion' : '1'
},
```

```
{
          'detectorId' : 'sample_detector',
          'ruleId' : 'medium_fraud_risk',
          'ruleVersion' : '1'
},
{
          'detectorId' : 'sample_detector',
          'ruleId' : 'low_fraud_risk',
          'ruleVersion' : '1'
}
],
      modelVersions = [{
          'modelId' : 'sample_fraud_detection_model',
          'modelType': 'ONLINE_FRAUD_INSIGHTS',
          'modelVersionNumber' : '1.00'
}      ],
      ruleExecutionMode = 'FIRST_MATCHED'
)
```

步骤 6：生成欺诈预测

本教程的最后一步使用上一步中sample_detector创建的检测器实时生成sample_registration事件类型的欺诈预测。检测器评估上传到 Amazon S3 的示例数据。响应包括模型性能分数以及与匹配规则相关的任何结果。

在以下示例中，GetEventPredictionAPI 用于为每个请求提供来自单个账户注册的数据。在本教程中，从账户注册示例数据文件中获取数据（电子邮件地址和 ip_address）。顶部标题行之后的每一行（行）代表来自单个账户注册事件的数据。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.get_event_prediction(
      detectorId = 'sample_detector',
      eventId = '802454d3-f7d8-482d-97e8-c4b6db9a0428',
      eventTypeName = 'sample_registration',
      eventTimestamp = '2020-07-13T23:18:21Z',
      entities = [{'entityType':'sample_customer', 'entityId':'12345'}],
 eventVariables = {
      'email_address': 'johndoe@exampledomain.com',
      'ip_address': '1.2.3.4'
```

```
}
)
```

完成本教程后，您完成了以下操作：

- 将示例事件数据集上载到 Amazon S3。
- 创建了用于创建和训练模型的变量、实体和标签。
- 使用示例数据集创建和训练模型。
- 查看了 Amazon Fraud Detector 生成的模型性能分数和其他性能指标。
- 部署了欺诈检测模型。
- 创建了探测器并添加了部署的模型。
- 向检测器添加了规则、规则执行顺序和结果。
- 探测器版本已创建。
- 通过提供不同的输入并检查规则和规则执行顺序是否按预期运行来测试检测器。

（可选）使用 Jupyter (IPython) 笔记本浏览亚马逊Fraud Detector API

有关如何使用亚马逊Fraud Detector API 的更多示例，请参阅aws-fraud-detector-samples GitHub 存储库。笔记本涵盖的主题包括使用 Amazon Fraud Detector API 构建模型和检测器，以及使用`GetEventPrediction` API 发出批量欺诈预测请求。

后续步骤

现在，您已经创建了模型和探测器，您可以进行更深入的研究，开始创建模型和检测器并生成欺诈预测。

《亚马逊Fraud Detector 用户指南》中的以下部分描述了您的企业或组织如何使用亚马逊Fraud Detector 来检测欺诈。

- 准备并创建事件数据集以训练模型。
- 创建事件类型
- 创建模型
- 创建探测器
- 获取欺诈预测

- 管理您的亚马逊Fraud Detector 资源（特别是变量、实体、结果和标签）
- 配置 Amazon Fraud Detector 以实现您的安全性和合规性
- 监控亚马逊Fraud Detector 并记录亚马逊Fraud Detector API 调用
- 排查Amazon Fraud Tor

事件数据集

事件数据集是贵公司的历史欺诈数据。您将这些数据提供给亚马逊Fraud Detector 以创建欺诈检测模型。

Amazon Fraud Detector 使用机器学习模型生成欺诈预测。每个模型都使用模型类型进行训练。模型类型指定了用于训练模型的算法和转换。模型训练是使用您提供的数据集来创建可以预测欺诈事件的模型的过程。有关更多信息，请参阅 Amazon Fraud Detector Detec

用于创建欺诈检测模型的数据集提供事件的详细信息。事件是对欺诈风险进行评估的业务活动。例如，账户注册可以是一项活动。与账户注册事件相关的数据可以是事件数据集。Amazon Fraud Detector 使用此数据集来评估账户注册欺诈。

在向 Amazon Fraud Detector 提供数据集以创建模型之前，请务必定义创建模型的目标。您还需要确定要如何使用该模型，并定义您的指标，以根据您的特定要求评估该模型的性能是否良好。

例如，创建用于评估账户注册欺诈的欺诈检测模型的目标可能如下：

- 自动批准合法注册。
- 捕获欺诈性注册以供日后调查。

确定目标后，下一步是决定如何使用该模型。以下是使用欺诈检测模型评估注册欺诈的一些示例：

- 用于对每个账户注册进行实时欺诈检测。
- 每小时对所有账户注册进行离线评估。

可用于衡量模型性能的一些指标示例包括：

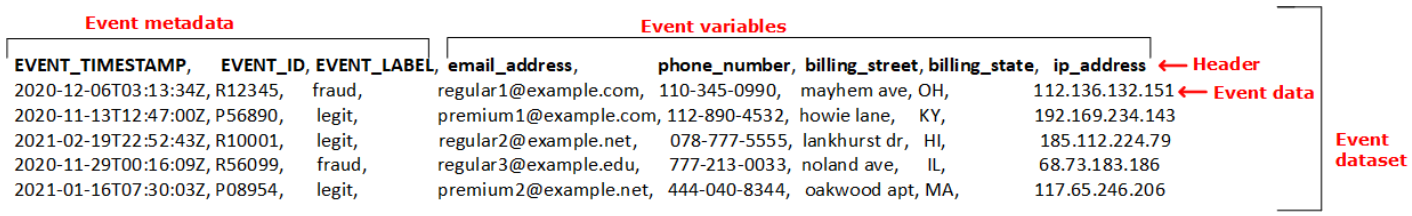
- 在生产中，性能一直优于当前的基线。
- 捕获 X% 的欺诈注册，误报率为 Y%。
- 最多可接受 5% 的自动批准的欺诈性注册。

事件数据集结构

Amazon Fraud Detector 要求您使用 UTF-8 格式的逗号分隔值 (CSV) 在文本文件中提供事件数据集。CSV 数据集文件的第一行必须包含文件标头。文件头由事件元数据和事件变量组成，这些变量描述了与事件相关的每个数据元素。标题后面是事件数据。每行由来自单个事件的数据元素组成。

- 事件元数据-提供有关事件的信息。例如，EVENT_TIMESTAMP 是指定事件发生时间的事件元数据。根据您的业务用例以及用于创建和训练欺诈检测模型的模型类型，Amazon Fraud Detector 要求您提供特定的事件元数据。在 CSV 文件标题中指定事件元数据时，请使用与 Amazon Fraud Detector 指定的相同的事件元数据名称，并且仅使用大写字母。
- 事件变量-表示特定于您的事件的数据元素，您要使用这些数据元素来创建和训练欺诈检测模型。根据您的业务用例以及用于创建和训练欺诈检测模型的模型类型，Amazon Fraud Detector 可能会要求或建议您提供特定的事件变量。您也可以选择提供事件中的其他事件变量，这些变量要包含在模型训练中。在线注册活动的事件变量的一些示例可以是电子邮件地址、IP 地址和电话号码。在 CSV 文件标题中指定事件变量名称时，请使用您选择的任何变量名称并仅使用小写字母。
- 事件数据-表示从实际事件中收集的数据。在 CSV 文件中，文件标题之后的每一行由来自单个事件的数据元素组成。例如，在在线注册事件数据文件中，每行都包含来自单个注册的数据。行中的每个数据元素都必须与相应的事件元数据或事件变量相匹配。

下面是包含账户注册事件数据的 CSV 文件文件中数据的 CSV 文件中数据的 CSV 文件夹。标题行包含大写的事件元数据和小写的事件变量，后面是事件数据。数据集中的每一行都包含与单一账户注册相关的数据元素，每个数据元素与标题对应。



使用数据模型浏览器获取事件数据集要求

您选择创建模型的模型类型定义了对数据集的要求。Amazon Fraud Detector 使用您提供的数据集来创建和训练您的欺诈检测模型。在 Amazon Fraud Detector 开始创建您的模型之前，它会检查数据集是否符合大小、格式和其他要求。如果数据集不满足要求，则模型创建和训练将失败。您可以使用数据模型浏览器来确定用于业务用例的模型类型，并深入了解已识别模型类型的数据集要求。

数据模型浏览器

数据模型浏览器是 Amazon Fraud Detector 控制台中的一个工具，可将您的业务用例与 Amazon Fraud Detector 支持的模型类型保持一致。数据模型浏览器还提供了对亚马逊欺诈检测器创建欺诈检测模型所需的数据元素的见解。在开始准备事件数据集之前，请使用数据模型资源管理器找出 Amazon Fraud Detector 推荐给您的业务使用的模型类型，并查看创建数据集所需的必需、推荐和可选数据元素列表。

要使用数据模型浏览器，

1. 打开AWS管理控制台并登录您的账户。导航到 Amazon Fraud Detec
2. 在左侧导航窗格中，选择 Manifest（数据模型资源管理器）
3. 在数据模型浏览器页面的业务用例下，选择要评估欺诈风险的业务用例。
4. Amazon Fraud Detector 显示与您的业务用例相匹配的推荐型号类型。模型类型定义了 Amazon Fraud Detector 将用于训练您的欺诈检测模型的算法、丰富和转换。

   请记下推荐的模型类型。稍后在创建模型时将需要此信息。

   > **Note**
   >
   > 如果您找不到您的业务用例，请使用描述中的 “联系我们” 链接向我们提供您的业务用例的详细信息。我们将推荐用于为您的业务用例创建欺诈检测模型的模型类型。

5. 数据模型见解窗格可让您深入了解为您的业务用例创建和训练欺诈检测模型所需的必需、推荐和可选数据元素。使用见解窗格中的信息来收集您的事件数据并创建您的数据集。

收集事件数据

收集事件数据是创建模型的重要步骤。这是因为您的模型在预测欺诈方面的性能取决于数据集的质量。当您开始收集事件数据时，请记住数据模型资源管理器为您创建数据集而提供的数据元素列表。您将需要收集所有必需（事件元数据）数据，并根据创建模型的目标决定要包括哪些推荐和可选的数据元素（事件变量）。确定要包含的每个事件变量的格式和数据集的总大小也很重要。

事件数据集质量

要为模型收集高质量数据集，我们有下列建议：

- 收集成熟数据- 使用最新数据有助于识别最新的欺诈模式。但是，要检测欺诈用例，请让数据成熟。到期期取决于您的业务，可能需要两周到三个月不等。例如，如果您的事件包括信用卡交易，则数据的到期日可能由信用卡的退款期或调查人员做出决定所花费的时间决定。

  确保用于训练模型的数据集有足够的时间根据您的业务发展成熟。
- 确保数据分布不会出现明显偏差- Amazon Fraud Detector 模型训练过程基于 EVENT_TIMESTAMP 对您的数据集进行采样和分区。例如，如果您的数据集包含最近 6 个月提取的欺诈事件，但仅包含最后一个月的合法事件，则数据分布被视为漂移且不稳定。不稳定的数据集可能会导致模型性能评估

出现偏差。如果您发现数据分布存在明显偏差，请考虑通过收集与当前数据分布类似的数据来平衡数据集。

- 确保数据集代表实现/测试模型的用例—— 否则，估计的性能可能会有偏差。假设您使用的是一个模型来自动拒绝所有室内申请人，但是您的模型是使用一个数据集进行训练的，该数据集包含先前已获得批准的历史数据/标签。那么，您的模型的评估可能不准确，因为评估基于的数据集，而该数据集没有来自被拒绝的申请人的陈述。

事件数据格式

作为模型训练过程的一部分，Amazon Fraud Detector 会将您的大部分数据转换为所需的格式。但是，您可以轻松使用一些标准格式来提供数据，这有助于避免以后在 Amazon Fraud Detector 验证您的数据集时出现问题。下表提供了有关提供推荐事件元数据的格式的指导。

> **Note**
>
> 创建 CSV 文件时，请务必以大写字母输入事件元数据名称，如下所示。

元数据名称	格式	必填
EVENT_ID	如果提供，它必须满足以下要求： • 这对于那个活动来说是独一无二的。 • 它代表了对您的业务有意义的信息。 • 它遵循正则表达式模式（例如，`^[0-9a-z_-]+$.`） • 除上述要求外，我们建议您不要向 EVENT_ID 附加时间戳。这样做可能会在更新事件时导致问题。这是因为如果您这样做，则必须提供完全相同的 EVENT_ID。	取决于模型类型

元数据名称	格式	必填
事件时间戳	• 它必须按下面的格式之一指定： • %yyyy-%mm-%ddt%HH: %mm: %ssz（仅限世界标准时间的 ISO 8601 标准，没有毫秒） 示例：2019-11-30T13：01:01 Z • %yyyy/%mm/%dd %hh: %mm: %ss (AM/PM) 示例：2019/11/30 1:01:01 下午，或 2019/11/30 13:01:01 • %mm/%dd/%yyyy %hh: %mm: %ss 示例：2019 年 11 月 30 日下午 1:01:01，2019 年 11 月 30 日 13:01:01 • %mm/%dd/%yy %hh: %mm: %ss 示例：11/30/19 1:01:01 下午，11/30/19 13:01:01 • Amazon Fraud Detector 在解析事件时间戳的日期/时间戳格式时会做出以下假设： • 如果您使用的是 ISO 8601 标准，则它必须与前面的规范完全匹配	是

元数据名称	格式	必填

	• 如果您使用的是其他格式之一，则还有额外的灵活性：
	  • 对于几个月和几天，您可以提供个位数或两位数。例如，2019 年 12 月 1 日是一个有效日期。
	  • 如果您没有 hh: mm: ss，则无需包含 hh: mm: ss（也就是说，您可以简单地提供日期）。您也可以只提供小时和分钟的子集（例如，hh: mm）。不支持仅提供小时数。也不支持毫秒。
	  • 如果您提供 AM/PM 标签，则假定时钟为 12 小时。如果没有 AM/PM 信息，则假定为 24 小时制。
	  • 您可以使用 “/” 或 “-” 作为日期元素的分隔符。假定时间戳元素为 “:”。	

实体_ID	• 它必须遵循正则表达式模式:`^[0-9A-Za-z_.@+-]+$` . • 如果实体 ID 在评估时不可用，请将实体 ID 指定为未知。	取决于模型类型
实体类型	你可以使用任何字符串	取决于模型类型

元数据名称	格式	必填
事件标签	您可以使用任何标签，例如 “欺诈”、“合法”、“1” 或 “0”。	如果包括 LABEL_TIMESTAMP，则为必填项
标签时间戳	它必须遵循时间戳格式。	如果包含事件标签，则为必填项

有关事件变量的信息，请参阅变量。

> **Important**
>
> 如果您正在创建 Account Takeover Insights (ATI) 模型，准备数据请参阅，了解有关准备和选择数据的详细信息。

空值或缺失值

EVENT_TIMESTAMP 和 EVENT_LABEL 变量不得包含任何空值或缺失值。其他变量的值可以为 null 或缺失值。但是，我们建议您对这些变量使用少量空值。如果 Amazon Fraud Detector 确定事件变量的空值或缺失值过多，它将自动从您的模型中省略变量。

最小变量

创建模型时，除了必需的事件元数据外，数据集还必须包含至少两个事件变量。这两个事件变量必须通过验证检查。

事件数据集大小

必填

您的数据集必须满足以下基本要求才能成功进行模型训练。

- 来自至少 100 个事件的数据。
- 数据集必须包含至少 50 个归类为欺诈的事件（行）。

推荐

我们建议您的数据集包含以下内容，以成功进行模型训练并获得良好的模型性能。

- 包括至少三周的历史数据，但最多包括六个月的数据。
- 包括至少 10K 个总事件数据。
- 包括至少 400 个归类为欺诈的事件（行）和 400 个归类为合法的事件（行）。
- 如果您的模型类型需要 ENTITY_ID，则包含 100 多个唯一实体。

数据集验证

在 Amazon Fraud Detector 开始创建您的模型之前，它会检查用于训练模型的数据集中包含的变量是否符合大小、格式和其他要求。如果数据集未通过验证，则不会创建模型。在创建模型之前，必须先修复未通过验证的变量。Amazon Fraud Detector 为您提供了一个数据分析器，您可以使用它来帮助您在开始训练模型之前识别和修复数据集的问题

数据分析器

Amazon Fraud Detector 提供了一种开源工具，用于分析和准备数据以进行模型训练。此自动数据分析器可帮助您避免常见的数据准备错误，并识别可能对模型性能产生负面影响的潜在问题，例如映射错误的变量类型。分析器生成数据集的直观而全面的报告，包括变量统计数据、标签分布、类别和数值分析以及变量和标签相关性。它提供了有关变量类型的指导以及将数据集转换为 Amazon Fraud Detector 要求的格式的选项。

使用数据分析器

自动数据分析器采用AWS CloudFormation堆栈构建，只需单击几下即可轻松启动该堆栈。所有代码都可以在 Github上找到。有关如何使用数据分析器的信息，请按照我们博客中的说明使用适用于 Amazon Fraud Detector 的自动数据分析器更快地训练模型

常见的事件数据集错误

以下是 Amazon Fraud Detector 在验证事件数据集时遇到的一些常见问题。运行数据分析器后，在创建模型之前，使用此列表检查数据集是否存在错误。

- CSV 文件不是 UTF-8 格式。
- 数据集中的事件数小于 100。
- 被确定为欺诈或合法的事件数量少于 50。
- 与欺诈事件相关的唯一实体数量少于 100。
- EVENT_TIMESTAMP 中超过 0.1% 的值包含空值或支持的日期/时间戳格式以外的值。
- EVENT_LABEL 中超过 1% 的值包含空值或事件类型中定义的值以外的值。

- 可用于模型训练的变量少于两个。

数据集存储

在您收集数据集后，您需要使用 Amazon Fraud Detector Detector Service (Amazon S3) 在外部存储。我们建议您根据用于生成欺诈预测的模型选择数据集的存储位置。有关模型类型的更多信息，请参阅选择模型类型。有关存储数据集的更多信息，请参阅事件数据存储。

事件类型

使用亚马逊欺诈检测器，您可以为事件生成欺诈预测。事件类型定义了发送到 Amazon Fraud Detector 的单个事件的结构。定义后，您可以构建模型和探测器来评估特定事件类型的风险。

事件的结构包括以下内容：

- 实体类型：对谁在执行事件进行分类。在预测期间，指定实体类型和实体 ID 以定义谁执行了事件。
- 变量：定义哪些变量可以作为事件的一部分发送。模型和规则使用变量来评估欺诈风险。一旦添加，就无法从事件类型中删除变量。
- 标签：将事件归类为欺诈性或合法事件。在模型训练期间使用。标签一旦添加，便无法从事件类型中移除。

创建事件类型

在创建欺诈检测模型之前，必须先创建事件类型。创建事件类型涉及定义您的业务活动（事件）以评估是否存在欺诈。定义事件包括识别数据集中用于欺诈评估的事件变量，指定实体发起事件以及对事件进行分类的标签。

创建事件类型的先决条件

在开始创建事件类型之前，请确保您已完成以下操作：

- 使用该数据模型浏览器工具深入了解了 Amazon Fraud Detector 所需的数据元素，以创建您的欺诈检测模型。
- 使用您从数据模型浏览器中获得的见解来创建事件数据集，并将您的数据集上传到 Amazon S3 存储桶。
- 已创建Variables实体、并且Labels您希望亚马逊欺诈检测器用于为此事件创建欺诈检测模型。确保您创建的变量、实体类型和标签包含在您的事件数据集中。

您可以在 Amazon Fraud Detector 控制台中使用 API、使用或使用 AWS SDK 创建您的事件类型。AWS CLI

在亚马逊欺诈检测器控制台中创建事件类型

要创建事件类型，

1. 打开AWS管理控制台并登录您的账户。导航到亚马逊欺诈检测器。
2. 在左侧导航窗格中，选择事件。
3. 在事件类型页面中，选择创建。
4. 在 “事件类型详细信息” 下，

   a. 在名称中，输入您的活动的名称。

   b. 在描述中，（可选）输入描述。

   c. 在实体中，选择您为活动创建的实体类型。
5. 在事件变量下，

   - 在 “选择如何定义此事件的变量” 中，

     - 如果您已经为此事件创建了事件变量，请从变量列表中选择选择变量，然后在变量中选择为此事件创建的变量。
     - 如果您尚未为此事件创建变量，请选择从训练数据集中选择变量，
       - 在 IAM 角色中，选择您希望 Amazon Fraud Detector 使用的 IAM 角色来访问包含您的数据集的 Amazon S3 存储桶
       - 在数据位置中输入数据集位置的路径。使用与此类似的S3 URI路径:`S3://your-bucket-name/example dataset filename.csv`.
       - 请选择 Upload（上传）。
       - 在变量下，将显示 Amazon Fraud Detector 从您的数据集文件中提取的所有事件变量名称。

         如果您希望包含用于检测欺诈的变量，请在变量类型中选择变量类型。选择 “移除”，将变量从欺诈检测中移除。对列表中的每个变量重复此步骤。
6. 在 “标签”（可选）下的 “标签” 中，选择您为此事件创建的标签。确保为欺诈和合法事件各选择一个标签。
7. 如果您想为此事件设置自动下游处理，请在 “使用亚马逊进行事件编排 EventBridge-可选” 下，打开 “使用亚马逊启用事件编排”。EventBridge有关事件协调的更多信息，请参阅活动编排。

> **Note**
>
> 您也可以在创建事件类型后稍后启用事件编排。

8. 选择创建事件类型。

使用创建事件类型 AWS SDK for Python (Boto3)

以下示例显示了 PutEventType API 的示例请求。该示例假设您已经创建了变量`ip_addresslegit`和`email_addressfraud`、标签和以及实体类型`sample_customer`。有关如何创建这些资源的信息，请参阅资源。

> **Note**
>
> 在将变量、实体类型和标签添加到事件类型之前，必须先创建变量、实体类型和标签。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.put_event_type (
name = 'sample_registration',
eventVariables = ['ip_address', 'email_address'],
labels = ['legit', 'fraud'],
entityTypes = ['sample_customer'])
```

删除事件或事件类型

当您删除某个事件时，Amazon Fraud Detector 会永久删除该事件，与该事件相关的数据将不再存储在亚马逊欺诈检测器中。

删除亚马逊欺诈检测器通过 **`GetEventPrediction`** API 评估的事件

1. 登录AWS Management Console并打开亚马逊欺诈检测器控制台，网址为 https://console.aws.amazon.com/frauddetector。
2. 在控制台的左侧导航窗格中，选择搜索过去的预测。
3. 选择要删除的事件。

4. 选择 “操作”，然后选择 “删除事件”。
5. 输入**delete**，然后选择删除事件。

> **Note**
>
> 这将删除与该事件 ID 关联的所有记录，包括发送到操作的任何事件数据以及通过该SendEventGetEventPrediction操作生成的任何预测数据。

要删除存储在 Amazon Fraud Detector 中但尚未经过评估的事件（即该事件是通过SendEvent操作存储的），您必须提出DeleteEvent请求并指定事件 ID 和事件类型 ID。如果您想同时删除事件和与该事件相关的任何预测历史记录，请将deleteAuditHistory参数的值设置为 “true”。如果将deleteAuditHistory参数设置为 “true”，则在删除操作完成后，最多可在 30 秒内通过搜索获得事件数据。

删除与某个事件类型关联的所有事件

1. 在控制台的左侧导航窗格中，选择事件类型
2. 选择要删除所有事件的事件类型。
3. 导航到 “存储的事件” 选项卡，然后选择 “删除存储的事件”

删除所有存储的事件可能需要一些时间，具体取决于该事件类型的存储事件的数量。例如，一个 1 GB 的数据集（普通客户大约需要 1-2 百万个事件）需要大约 2 小时才能删除。在此期间，您发送到该事件类型的 Amazon Fraud Detector 的新事件不会被存储，但您可以继续通过GetEventPrediction操作生成欺诈预测。

删除事件类型

您无法删除探测器或模型中使用的事件类型，也无法删除关联已存储的事件类型。在删除事件类型之前，必须删除与该事件类型关联的所有事件。

当您删除事件类型时，亚马逊欺诈检测器会永久删除该事件类型，并且数据将不再存储在亚马逊欺诈检测器中。

1. 在 Amazon Fraud Detector 控制台的左侧导航窗格中，选择资源，然后选择事件。
2. 选择要删除的事件类型。
3. 选择 “操作”，然后选择 “删除事件类型”。

4. 输入事件类型名称，然后选择删除事件类型。

事件数据存储

收集数据集后，您可以使用 Amazon S3 (Amazon S3) 将 Amazon (Amazon S3) 存储在外部存储数据集。我们建议您根据用于生成欺诈预测的模型选择数据集的存储位置。以下是这两个存储选项的详细分类。

- 内部存储- 您的数据集存储在 Amazon Fraud Detector 中。与事件相关的所有事件数据都存储在一起。您可以随时上传存储在 Amazon Fraud Detector 中的事件数据集。您可以将事件逐一传输到 Amazon Fraud Detector API，也可以使用批量导入功能导入大型数据集（最多 1GB）。当您使用 Amazon Fraud Detector 存储的数据集训练模型时，您可以指定时间范围来限制数据集的大小。
- 外部存储- 您的数据集存储在 Amazon Fraud Detector 以外的外部数据源中。目前，Amazon Frauud Detector 支持为此目的使用Amazon S3。如果您的模型位于上传到 Amazon S3 的文件上，则该文件的未压缩数据不能超过 5GB。如果不止于此，请务必缩短数据集的时间范围。

下表提供了有关模型类型及其支持的数据源的详细信息。

模型类型	兼容的训练数据源
在线欺诈洞察	外部存储，内部存储
交易欺诈洞察	内部存储
账户收购洞察	内部存储

有关使用 Amazon 简单存储服务在外部存储数据集的信息，请参阅使用 Amazon S3 在外部存储您的事件数据 。有关使用 Amazon Fraud Detector 在内部存储数据集的信息，请参阅使用亚马逊Fraud Detector 在内部存储您的事件数据。

使用 Amazon S3 在外部存储您的事件数据

如果您正在训练 Online Fraud Insights 模型，则可以选择使用 Amazon S3 在外部存储事件数据。要将事件数据存储在 Amazon S3 中，您必须先创建 CSV 格式的文本文件，添加事件数据，然后将 CSV 文件上传到 Amazon S3 存储桶。

> **Note**
>
> 交易欺诈洞察和账户收购洞察模型类型不支持 Amazon S3 外部存储的数据集

创建 CSV 文件

亚马逊Fraud Detector 要求您的 CSV 文件的第一行包含列标题。CSV 文件中的列标题必须映射到事件类型中定义的变量。有关示例数据集，请参见获取并上传示例数据集

Online Fraud Insights 模型要求训练数据集至少包含 2 个变量和最多 100 个变量。除事件变量外，训练数据集还必须包含以下标题：

- EVENT_TIMESTAMP-定义事件发生的时间
- EVENT_LABEL-将事件归类为欺诈事件或合法事件。列中的值必须对应于事件类型中定义的值。

以下 CSV 数据示例，代表在线商家的历史注册事件：

```
EVENT_TIMESTAMP,EVENT_LABEL,ip_address,email_address
4/10/2019 11:05,fraud,209.146.137.48,fake_burtonlinda@example.net
12/20/2018 20:04,legit,203.0.112.189,fake_davidbutler@example.org
3/14/2019 10:56,legit,169.255.33.54,fake_shelby76@example.net
1/3/2019 8:38,legit,192.119.44.26,fake_curtis40@example.com
9/25/2019 3:12,legit,192.169.85.29,fake_rmiranda@example.org
```

> **Note**
>
> CSV 数据文件可以包含双引号和逗号作为数据的一部分。

相应事件类型的简化版本如下所示。事件变量对应于 CSV 文件中的标题，中的值EVENT_LABEL对应于标签列表中的值。

```
(
name = 'sample_registration',
eventVariables = ['ip_address', 'email_address'],
labels = ['legit', 'fraud'],
entityTypes = ['sample_customer']
```

)

事件时间戳格式

确保您的事件时间戳采用所需格式。作为模型构建过程的一部分，Online Fraud Insights 模型类型根据事件时间戳对您的数据进行排序，并将您的数据拆分用于训练和测试目的。为了获得对性能的合理估计，模型首先在训练数据集上训练，然后在测试数据集上测试该模型。

Amazon Fraud Detector 支持模型训练`EVENT_TIMESTAMP`期间的值采用以下日期/时间戳格式：

- %yyyy-%mm-%ddt%HH: %mm: %ssz（仅限世界标准时间的 ISO 8601 标准，没有毫秒）

  示例：2019-11-30T13：01:01 Z
- %yyyy/%mm/%dd %hh: %mm: %ss (AM/PM)

  示例：2019/11/30 1:01:01 下午，或 2019/11/30 13:01:01
- %mm/%dd/%yyyy %hh: %mm: %ss

  示例：2019 年 11 月 30 日下午 1:01:01，2019 年 11 月 30 日 13:01:01
- %mm/%dd/%yy %hh: %mm: %ss

  示例：11/30/19 1:01:01 下午，11/30/19 13:01:01

Amazon Fraud Detector 在解析事件时间戳的日期/时间戳格式时会做出以下假设：

- 如果您使用的是 ISO 8601 标准，则它必须与前面的规范完全匹配
- 如果您使用的是其他格式之一，则还有额外的灵活性：
  - 对于几个月和几天，您可以提供个位数或两位数。例如，2019 年 12 月 1 日是一个有效日期。
  - 如果您没有 hh: mm: ss，则无需包含 hh: mm: ss（也就是说，您可以简单地提供日期）。您也可以只提供小时和分钟的子集（例如，hh: mm）。不支持仅提供小时数。也不支持毫秒。
  - 如果您提供 AM/PM 标签，则假定时钟为 12 小时。如果没有 AM/PM 信息，则假定为 24 小时制。
  - 您可以使用 “/” 或 “-” 作为日期元素的分隔符。假定时间戳元素为 “:”。

跨时间对数据集进行采样

我们建议您提供同一时间范围内的欺诈示例和合法样本。例如，如果您提供过去 6 个月的欺诈事件，则还应提供平均跨越同一时间段的合法事件。如果您的数据集包含非常不均匀的欺诈和合法事件分布，

您可能会收到以下错误：“随着时间的推移，欺诈分布的波动令人无法接受。无法正确拆分数据集。”通常，解决此错误的最简单方法是确保在相同的时间范围内对欺诈事件和合法事件进行均匀采样。如果您在短时间内经历了大量欺诈激增，则可能还需要删除数据。

如果您无法生成足够的数据来创建均匀分布的数据集，则一种方法是随机化事件的 EVENT_TIMESTAMP，使其均匀分布。但是，这通常会导致性能指标不切实际，因为 Amazon Fraud Detector 使用 EVENT_TIMESTAMP 根据数据集中的相应事件子集评估模型。

空值和缺失值

亚马逊Fraud Detector 处理空值和缺失值。但是，应限制变量的空值百分比。EVENT_TIMESTAMP 和 EVENT_LABEL 列不应包含任何缺失值。

文件验证

如果触发以下任一条件，Amazon Frauud Detector 将无法训练模型：

- 如果 CSV 无法解析
- 如果列的数据类型不正确

将事件数据上传到 Amazon S3 存储桶

使用事件数据创建 CSV 文件后，将文件上传到 Amazon S3 存储桶。

要上传到 Amazon S3 存储桶

1. 登录到 AWS Management Console，然后通过以下网址打开 Amazon S3 控制台：https://console.aws.amazon.com/s3/。
2. 请选择 Create bucket（创建桶）。

   此时将打开 Create bucket（创建桶）向导。
3. 在 Bucket name（桶名称）中，输入符合 DNS 标准的桶名称。

   桶名称必须满足以下要求：

   - 在所有 Amazon S3 中是唯一的。
   - 长度必须介于 3 到 63 个字符之间。
   - 不包含大写字符。

- 以小写字母或数字开头。

创建存储桶后，便无法再更改其名称。有关命名存储桶的信息，请参阅《Amazon Simple Storage 用户指南》中的存储桶命名规则。

> **Important**
>
> 避免在存储桶名称中包含敏感信息，如账号。桶名称会显示在指向桶中的对象的 URL 中。

4. 对于 Region（区域），选择要放置存储桶的AWS区域。您必须选择使用 Amazon (N. Virginia)、美国东部（俄亥俄州）、美国西部（俄勒冈）、欧洲（爱尔兰）、亚太区域（新加坡）或亚太区域（悉尼）。
5. 在 Bucket settings for Block Public Access（阻止公有访问的桶设置）中，请选择要应用于桶的 Block Public Access（阻止公有访问）设置。

   我们建议您将所有设置保留为启用状态。有关阻止公有访问的更多信息，请参阅《Amazon Service 用户指南》中的阻止公有访问您的 Amazon S3 存储空间。
6. 选择创建桶。
7. 将训练数据文件上传到 Amazon S3 存储桶。记下您的培训文件的 Amazon S3 位置路径（例如 s3://bucketname/object.csv）。

使用亚马逊Fraud Detector 在内部存储您的事件数据

您可以选择将事件数据存储在 Amazon Fraud Detector 中，稍后使用存储的数据来训练您的模型。通过将事件数据存储在 Amazon Fraud Detector 中，您可以训练使用自动计算变量的模型来提高性能、简化模型再训练并更新欺诈标签以关闭机器学习反馈循环。事件存储在事件类型资源级别，因此相同事件类型的所有事件一起存储在单个事件类型数据集中。作为定义事件类型的一部分，您可以选择通过切换 Amazon Fraud Detector 控制台中的事件摄取设置来指定是否存储该事件类型的事件。

您可以存储单个事件，也可以在 Amazon Fraud Detector 中导入大量事件数据集。单个事件可以使用 GetEventPredictionAPI 或 API 进行流式传输。SendEvent使用 Amazon Fraud Detector 控制台中的批量导入功能或使用 CreateBatchImportJobAPI，可以快速轻松地将大型数据集导入到 Amazon Fraud Detector。

您可以随时使用 Amazon Fraud Detector 控制台查看每种事件类型已存储的事件数量。

准备用于存储的事件数据

使用 Amazon Fraud Detector 内部存储的事件数据存储在Event Type资源级别。因此，来自同一事件的所有事件数据都存储在单个事件中Event Type。存储的事件稍后可用于训练新模型或重新训练现有模型。使用存储的事件数据训练模型时，您可以选择指定事件的时间范围以限制训练数据集的大小。

每次您使用亚马逊欺诈检测器控制台、API 或SendEvent API 将数据存储在 Amazon Fraud Detector 中时，Amazon Fraud Detector 都会在存储之前验证您的数据。CreateBatchImportJob如果您的数据未通过验证，则不会存储事件数据。

使用亚马逊Fraud Detector 在内部存储数据的先决条件

- 为确保您的事件数据通过验证并成功存储数据集，请确保您已使用数据模型资源管理器提供的见解来准备数据集。
- 为您要使用 Amazon Fraud Detector 存储的事件数据创建了事件类型。如果没有，请按照说明创建事件类型。

智能数据验证

当您将数据集上传到 Amazon Fraud Detector 控制台进行批量导入时，Amazon Fraud Detector 在导入数据之前使用智能数据验证 (SDV) 验证您的数据集。SDV 会扫描上传的数据文件并识别诸如数据丢失以及格式或数据类型不正确之类的问题。除了验证您的数据集外，SDV 还提供验证报告，其中列出了已发现的所有问题，并建议解决影响最大的问题的措施。在 Amazon Fraud Detector 成功导入您的数据集之前，SDV 发现的一些问题可能非常严重，必须解决这些问题。有关更多信息，请参阅智能数据验证报告：

SDV 在文件级别和数据（行）级别验证您的数据集。在文件级别，SDV 会扫描您的数据文件并识别诸如文件访问权限不足、文件大小不正确、文件格式和标题（事件元数据和事件变量）等问题。在数据级别，SDV 会扫描每个事件数据（行）并识别诸如不正确的数据格式、数据长度、时间戳格式和空值等问题。

智能数据验证目前仅在 Amazon Fraud Detector 控制台中可用，并且默认情况下验证处于启用状态。如果您不希望 Amazon Fraud Detector 在导入数据集之前使用智能数据验证，请在上传数据集时在 Amazon Fraud Detector 控制台中关闭验证。

在使用 API 或AWS SDK 时验证存储的数据

通过、或CreateBatchImportJob API 操作上传事件时SendEventGetEventPrediction，Amazon Fraud Detector 会验证以下内容：

- 该事件类型的 EventIngestion 设置为 “启用”。
- 事件时间戳无法更新。具有重复事件 ID 且不同的 EVENT_TIMESTAMP 的事件将被视为错误。
- 变量名称和值与其预期格式相符。有关更多信息，请参阅 创建变量
- 必填变量用值填充。
- 所有事件时间戳均不超过 18 个月，并且不在future。

使用批量导入存储事件数据

借助批量导入功能，您可以使用控制台、API 或 AWS SDK 在 Amazon Fraud Detector 中快速轻松地上传大型历史事件数据集。要使用批量导入，请创建一个包含所有事件数据的 CSV 格式的输入文件，将 CSV 文件上传到 Amazon S3 存储桶，然后启动导入任务。Amazon Fraud Detector 首先根据事件类型验证数据，然后自动导入整个数据集。导入数据后，即可将其用于训练新模型或重新训练现有模型。

输入与输出文件

输入 CSV 文件必须包含与关联事件类型中定义的变量相匹配的标题以及四个必需变量。参阅 准备用于存储的事件数据 了解更多信息。输入数据文件的最大大小为 20 千兆字节 (GB)，或大约 5000 万个事件。活动数量将因您的活动规模而异。如果导入任务成功，则输出文件为空。如果导入不成功，则输出文件包含错误日志。

创建 CSV 文件

Amazon Fraud Detector 仅从逗号分隔值 (CSV) 格式的文件导入数据。CSV 文件的第一行必须包含与关联事件类型中定义的变量完全匹配的列标题以及四个必需变量：EVENT_ID、EVENT_TIMESTAMP、ENTITY_ID 和 ENTITY_TYPE。您还可以选择包括 EVENT_LABEL 和 LABEL_TIMESTAMP（如果包含 EVENT_LABEL，则需要 LABEL_TIMESTAMP）。

定义强制变量

强制变量被视为事件元数据，必须以大写形式指定。模型训练时会自动包含事件元数据。下表列出了强制变量、每个变量的描述以及变量所需的格式。

名称	描述	要求
EVENT_ID	事件的标识符。例如，如果您的活动是在线交易，则	• 批量导入任务需要 EVENT_ID。

名称	描述	要求
	EVENT_ID 可能是提供给客户的交易参考号。	• 此名称对于该事件来说必须是唯一的。 • 它应该代表对您的业务有意义的信息。 • 此名称必须满足正则表达式模式（例如^[0-9a-z_-]+$.） • 我们不建议您将时间戳附加到 EVENT_ID。这样做可能会在更新事件时导致问题。这是因为如果您这样做，则必须提供完全相同的 EVENT_ID。

名称	描述	要求

事件时间戳	事件发生时的时间戳。时间戳必须采用 UTC 的 ISO 8601 标准。	• 批量导入任务需要 EVENT_TIMESTAMP。
		• 此名称必须按以下其中一种格式指定：
		  • %yyyy-%mm-%ddt%HH:%mm: %ssz（仅限世界标准时间的 ISO 8601 标准，没有毫秒）
		  示例：2019-11-30T13：01:01 Z
		  • %yyyy/%mm/%dd %hh:%mm: %ss (AM/PM)
		  示例：2019/11/30 1:01:01 下午，或 2019/11/30 13:01:01
		  • %mm/%dd/%yyyy %hh:%mm: %ss
		  示例：2019 年 11 月 30 日下午 1:01:01，2019 年 11 月 30 日 13:01:01
		  • %mm/%dd/%yy %hh:%mm: %ss
		  示例：11/30/19 1:01:01 下午，11/30/19 13:01:01
		• Amazon Fraud Detector 在解析事件时间戳的日期/时间戳格式时会做出以下假设：
		  • 如果您使用的是 ISO 8601 标准，则它必须与前面的规范完全匹配

名称	描述	要求

		• 如果您使用的是其他格式之一，则还有额外的灵活性：
		  • 对于几个月和几天，您可以提供个位数或两位数。例如，2019 年 12 月 1 日是一个有效日期。
		  • 如果您没有 hh: mm: ss，则无需包含 hh: mm: ss（也就是说，您可以简单地提供日期）。您也可以只提供小时和分钟的子集（例如，hh: mm）。不支持仅提供小时数。也不支持毫秒。
		  • 如果您提供 AM/PM 标签，则假定时钟为 12 小时。如果没有 AM/PM 信息，则假定为 24 小时制。
		  • 您可以使用 “/” 或 “-” 作为日期元素的分隔符。假定时间戳元素为 “:”。

名称	描述	要求
实体_ID	执行事件的实体的标识符。	• 批量导入任务需要 ENTITY_ID • 它必须遵循正则表达式模式:`^[0-9A-Za-z_.@+-]+$` . • 如果实体 ID 在评估时不可用，请将实体 ID 指定为未知。
实体类型	执行活动的实体，例如商家或客户	批量导入任务需要 ENTITY_TYPE
事件标签	将事件分类为`fraudulent` 或`legitimate`	如果包含 LABEL_TIMESTAMP，则需要 EVENT_LABEL
标签时间戳	上次填充或更新事件标签的时间戳	• 如果包含 EVENT_LABEL，则需要 LABEL_TIMESTAMP。 • 此名称必须遵循时间戳格式。

将 CSV 文件上载到 Amazon S3 进行批量导入

使用数据创建 CSV 文件后，将文件上传到 Amazon S3 存储桶。

要将事件数据上传到 Amazon S3 存储桶

1. 登录到 AWS Management Console，然后通过以下网址打开 Amazon S3 控制台：https://console.aws.amazon.com/s3/。
2. 请选择 Create bucket（创建桶）。

   此时将打开 Create bucket（创建桶）向导。
3. 在 Bucket name（桶名称）中，输入符合 DNS 标准的桶名称。

桶名称必须满足以下要求：

- 在所有 Amazon S3 中是唯一的。
- 长度必须介于 3 到 63 个字符之间。
- 不包含大写字符。
- 以小写字母或数字开头。

创建存储桶后，便无法再更改其名称。有关命名存储桶的信息，请参阅《Amazon Service 用户指南》中的存储桶命名规则。

> **Important**
>
> 避免在存储桶名称中包含敏感信息，如账号。桶名称会显示在指向桶中的对象的 URL 中。

4. 对于 Region（区域），选择要放置存储桶的AWS区域。您必须选择使用 Amazon (N. Virginia)、美国东部（俄亥俄州）、美国西部（俄勒冈）、欧洲（爱尔兰）、亚太区域（新加坡）或亚太区域（悉尼）。
5. 在 Bucket settings for Block Public Access（阻止公有访问的桶设置）中，请选择要应用于桶的 Block Public Access（阻止公有访问）设置。

   我们建议您将所有设置保留为启用状态。有关阻止公有访问的更多信息，请参阅《Amazon Service 用户指南》中的阻止公有访问您的 Amazon S3 存储空间。
6. 选择创建桶。
7. 将训练数据文件上传到 Amazon S3 存储桶。记下您的培训文件的 Amazon S3 位置路径（例如 s3://bucketname/object.csv）。

在亚马逊Fraud Detector 控制台中Batch 导入事件数据

您可以使用`CreateBatchImportJob` API 或 AWS SDK 在 Amazon Fraud Detector 控制台中轻松导入大量事件数据集。在继续操作之前，请确保您已按照说明将数据集准备为 CSV 文件。确保您还将 CSV 文件上传到 Amazon S3 存储桶。

使用亚马逊Fraud Detector 控制台

在控制台中批量导入事件数据

1. 打开 AWS 控制台并登录您的账户，然后导航到 Amazon Fraud Detector。
2. 在左侧导航窗格中，选择事件。
3. 选择事件类型。
4. 选择 “存储的事件” 选项卡。
5. 在存储的事件详细信息窗格中，确保事件提取处于开启状态。
6. 在导入事件数据窗格中，选择新建导入。
7. 在新事件导入页面中，提供以下信息：

   - [推荐] 保留此数据集的 “启用智能数据验证”-新设置为默认设置。
   - 对于数据的 IAM 角色，选择您为保存您计划导入的 CSV 文件的 Amazon S3 存储桶创建的 IAM 角色。
   - 在输入数据位置中，输入存放 CSV 文件的 S3 位置。
   - 如果您想指定一个单独的位置来存储导入结果，请单击 “将输入和结果的数据位置分开” 按钮，并提供有效的 Amazon S3 存储桶位置。

   > **Important**
   >
   > 确保您选择的 IAM 角色对您的输入 Amazon S3 存储桶具有读取权限，对您的输出 Amazon S3 存储桶具有写入权限。

8. 选择开始。
9. 导入事件数据窗格中的状态列显示您的验证和导入任务的状态。顶部的横幅提供了对数据集状态的高级描述，因为您的数据集首先经过验证，然后再进行导入。
10. 按照提供的指导进行操作监控数据集验证和导入作业的进度。

监控数据集验证和导入作业的进度

如果您使用 Amazon Fraud Detector 控制台执行批量导入任务，则默认情况下，Amazon Fraud Detector 会在导入之前验证您的数据集。您可以在 Amazon Fraud Detector 控制台的新事件导入页面监控验证的进度和状态并导入任务。页面顶部的横幅简要描述验证结果和导入作业的状态。根据验证结果和导入任务的状态，您可能需要采取措施来确保成功验证和导入数据集。

下表详细介绍了根据验证和导入操作的结果必须采取的操作。

横幅消息	状态	含义	我该怎么办
数据验证已开始	验证持续改善	SDV 已开始验证您的数据集	等待状态改变
由于数据集中的错误，无法继续进行数据验证。修复数据文件中的错误并开始新的导入任务。有关更多信息，请参阅验证报告	验证失败	SDV 在您的数据文件中发现了问题。要成功导入数据集，必须解决这些问题。	在导入事件数据窗格中，选择 Job ID 并查看验证报告。按照报告中的建议解决列出的所有错误。有关更多信息，请参阅使用验证报告：
数据导入已开始。验证成功完成	导入持续改善	您的数据集通过了验证。AFD 已开始导入您的数据集	等待状态改变
验证已完成，但出现警告。数据导入已开始	导入持续改善	您的数据集中的某些数据未通过验证。但是，通过验证的数据符合导入的最低数据大小要求。	监视标语中的消息并等待状态更改
您的数据已部分导入。一些数据未通过验证，未被导入。有关更多信息，请参阅验证报告。	已导入。状态显示警告图标。	您的数据文件中验证失败的某些数据未被导入。通过验证的其余	在导入事件数据窗格中，选择 Job ID 并查看验证报告。按照数据级别警告表中的建议解决列出的警告。您无需解决所有警告。但是，要成功导入，请确保您的数据集有超过

横幅消息	状态	含义	我该怎么办
		数据均已导入。	50% 的数据通过验证。解决警告后，开始新的导入任务。有关更多信息，请参阅使用验证报告：
由于处理错误，数据导入失败。开始新的数据导入任务	导入失败	由于暂时性运行时错误，导入失败	开始新的导入任务
数据已成功导入	已导入	验证和导入均成功完成	选择导入任务的Job ID 以查看详细信息，然后继续进行模型训练

> **Note**
>
> 我们建议在数据集成功导入 Amazon Fraud Detector 后等待 10 分钟，以确保它们被系统完全吸收。

智能数据验证报告

智能数据验证会在验证完成后创建验证报告。验证报告详细介绍了 SDV 在您的数据集中发现的所有问题，并建议采取哪些措施来修复最具影响力的问题。您可以使用验证报告来确定问题是什么、问题在数据集中的位置、问题的严重程度以及如何修复它们。即使验证成功完成，也会创建验证报告。在这种情况下，您可以查看报告以查看是否列出了任何问题，如果有，则决定是否要修复其中的任何问题。

> **Note**
>
> 当前版本的 SDV 会扫描您的数据集以查找可能导致批量导入失败的问题。如果验证和批量导入成功，您的数据集仍可能存在可能导致模型训练失败的问题。即使验证和导入成功，我们也建议您查看验证报告，并解决报告中列出的任何问题以成功进行模型训练。解决问题后，创建新的批量导入任务。

访问验证报告

验证完成后，您可以随时使用以下选项之一访问验证报告：

1. 验证完成后，导入任务正在进行中，在顶部横幅中选择查看验证报告。
2. 导入任务完成后，在导入事件数据窗格中，选择刚刚完成的导入任务的Job ID。

使用验证报告

导入任务的验证报告页面提供此导入任务的详细信息、发现的严重错误列表、有关数据集中特定事件（行）的警告列表（如果找到），以及数据集的简短摘要，其中包括无效值和每个变量的缺失值等信息。

- 导入任务详情

  提供导入作业的详细信息。如果导入任务失败或数据集已部分导入，请选择转到结果文件以查看导入失败事件的错误日志。

- 严重错误

  提供 SDV 确定的数据集中最具影响力的问题的详细信息。此窗格中列出的所有问题都很关键，在继续导入之前必须解决这些问题。如果您尝试在未解决关键问题的情况下导入数据集，则导入任务可能会失败。

  要解决关键问题，请遵循针对每个警告提供的建议。解决了 “严重错误” 窗格中列出的所有问题后，创建新的批量导入作业。

- 数据级别警告

  提供数据集中特定事件（行）的警告摘要。如果填充了数据级别警告窗格，则数据集中的某些事件未通过验证且未导入。

  对于每个警告，描述列显示有问题的事件数量。而且，示例事件 ID 提供了部分示例事件 ID 列表，您可以将其用作起点来查找存在问题的其余事件。使用为警告提供的建议来解决问题。也可以使用输出文件中的错误日志来获取有关该问题的更多信息。错误日志是针对所有批量导入失败的事件生成的。要访问错误日志，请在导入作业详细信息窗格中选择转到结果文件。

  > **Note**
  >
  > 如果数据集中超过 50% 的事件（行）未通过验证，则导入任务也会失败。在这种情况下，必须先修复数据，然后才能开始新的导入任务。

- 数据集摘要

提供数据集验证报告的摘要。如果 “警告数量” 列显示的警告超过 0 个，请决定是否需要修复这些警告。如果警告数列显示为 0，请继续训练您的模型。

使用适用于 Python 的 AWS 软件开发工具包 (Boto3) Batch 导入事件数据

下面的示例显示对 CreateBatchImportJobAPI 的示例请求。批量导入任务必须包含 jobID、InputPath、OutputPath eventTypeName和iamRoleArn。除非作业在 CREATE_FAILED 状态下存在，否则 joBid 不能包含与过去作业相同的 ID。输入路径和输出路径必须是有效的 S3 路径。您可以选择不在 OutputPath 中指定文件名，但是，您仍然需要提供有效的 S3 存储桶位置。eventTypeName 和 iamRoleArn 必须存在。IAM 角色必须授予输入 Amazon S3 存储桶的读取权限和输出 Amazon S3 存储桶的写入权限。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.create_batch_import_job (
jobId = 'sample_batch_import',
inputPath = 's3://bucket_name/input_file_name.csv',
outputPath = 's3://bucket_name/',
eventTypeName = 'sample_registration',
iamRoleArn: 'arn:aws:iam::************:role/service-role/AmazonFraudDetector-
DataAccessRole-*************'
)
```

取消批量导入任务

您可以随时在 Amazon Fraud Detector 控制台中使用CancelBatchImportJob API 或 AWS SDK 取消正在进行的批量导入任务。

要在控制台中取消批量导入任务，

1. 打开 AWS 控制台并登录您的账户，然后导航到 Amazon Fraud Detector。
2. 在左侧导航窗格中，选择事件。
3. 选择事件类型。
4. 选择 “存储的事件” 选项卡。
5. 在导入事件数据窗格中，选择要取消的正在进行的导入任务的作业 ID。
6. 在事件作业页面中，单击 “操作”，然后选择 “取消事件导入”。

7. 选择 “停止事件导入” 以取消批量导入作业。

使用适用于 Python 的 AWS 软件开发工具包 (Boto3) 取消批量导入任务

下面的示例显示对CancelBatchImportJob API 的示例请求。取消导入任务必须包含正在进行的批量导入作业的任务 ID。

```
import boto3
fraudDetector = boto3.client('frauddetector')
fraudDetector.cancel_batch_import_job (
    jobId = 'sample_batch'
)
```

使用 GetEventPredictions API 操作存储事件数据

默认情况下，发送到GetEventPrediction API 进行评估的所有事件都存储在 Amazon Fraud Detector 中。这意味着，当您生成预测并使用该数据近乎实时地更新计算变量时，Amazon Fraud Detector 将自动存储事件数据。您可以通过在 Amazon Fraud Detector 控制台中导航到事件类型并将事件摄取设置为关闭或使用PutEventType API 操作将 EventIngestion 值更新为 “禁用” 来禁用数据存储。有关GetEventPrediction API 操作的更多信息，请参阅欺诈预测。

> **Important**
>
> 我们强烈建议为某一事件类型启用事件提取后，将其保持启用状态。禁用相同事件类型的事件采集然后生成预测可能会导致行为不一致。

使用 SendEvent API 操作存储事件数据

您可以使用SendEvent API 操作将事件存储在 Amazon Fraud Detector 中，而无需为这些事件生成欺诈预测。例如，您可以使用该SendEvent操作上传历史数据集，以后可以使用该数据集来训练模型。

SendEvent API 的事件时间戳格式

使用SendEvent API 存储事件数据时，必须确保事件时间戳采用所需格式。亚马逊Fraud Detector 支持以下日期/时间戳格式：

- %yyyy-%mm-%ddt%HH: %mm: %ssz（仅限世界标准时间的 ISO 8601 标准，没有毫秒）

示例：2019-11-30T13：01:01 Z

- %yyyy/%mm/%dd %hh: %mm: %ss (AM/PM)

  示例：2019/11/30 1:01:01 下午，或 2019/11/30 13:01:01

- %mm/%dd/%yyyy %hh: %mm: %ss

  示例：2019 年 11 月 30 日下午 1:01:01，2019 年 11 月 30 日 13:01:01

- %mm/%dd/%yy %hh: %mm: %ss

  示例：11/30/19 1:01:01 下午，11/30/19 13:01:01

Amazon Fraud Detector 在解析事件时间戳的日期/时间戳格式时会做出以下假设：

- 如果您使用的是 ISO 8601 标准，则它必须与前面的规范完全匹配
- 如果您使用的是其他格式之一，则还有额外的灵活性：
  - 对于几个月和几天，您可以提供个位数或两位数。例如，2019 年 12 月 1 日是一个有效日期。
  - 如果您没有 hh: mm: ss，则无需包含 hh: mm: ss（也就是说，您可以简单地提供日期）。您也可以只提供小时和分钟的子集（例如，hh: mm）。不支持仅提供小时数。也不支持毫秒。
  - 如果您提供 AM/PM 标签，则假定时钟为 12 小时。如果没有 AM/PM 信息，则假定为 24 小时制。
  - 您可以使用 “/” 或 “-” 作为日期元素的分隔符。假定时间戳元素为 “:”。

以下是SendEvent API 调用示例。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.send_event(
            eventId        = '802454d3-f7d8-482d-97e8-c4b6db9a0428',
            eventTypeName  = 'sample_registration',
            eventTimestamp = '2020-07-13T23:18:21Z',
            eventVariables =  {
       'email_address' : 'johndoe@exampledomain.com',
       'ip_address' : '1.2.3.4'},
            assignedLabel  = ‘legit’,
            labelTimestamp = '2020-07-13T23:18:21Z',
            entities       = [{'entityType':'sample_customer', 'entityId':'12345'}],
```

```
)
```

获取存储的事件数据的详细信息

将事件数据存储在 Amazon Fraud Detector 中后，您可以使用 GetEventAPI 检查为事件存储的最新数据。以下示例代码检查为该sample_registration事件存储的最新数据。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.get_event(
            eventId        = '802454d3-f7d8-482d-97e8-c4b6db9a0428',
            eventTypeName  = 'sample_registration'
)
```

查看存储的事件数据集的指标

对于每种事件类型，您可以在 Amazon Fraud Detector 控制台中查看指标，例如存储的事件数量、存储事件的总大小以及最早和最新存储事件的时间戳。

要查看某一事件类型的存储事件指标，

1. 打开AWS控制台并登录您的账户。导航到 Fraud Detector。
2. 在左侧导航窗格中，选择事件。
3. 选择事件类型。
4. 选择 “存储的事件” 选项卡。
5. 存储的事件详细信息窗格显示指标。这些指标每天自动更新一次。
6. （可选）点击刷新事件指标来手动更新您的指标。

> **Note**
>
> 如果您刚刚导入了数据，我们建议您在完成数据导入后等待 5-10 分钟以刷新和查看指标。

活动编排

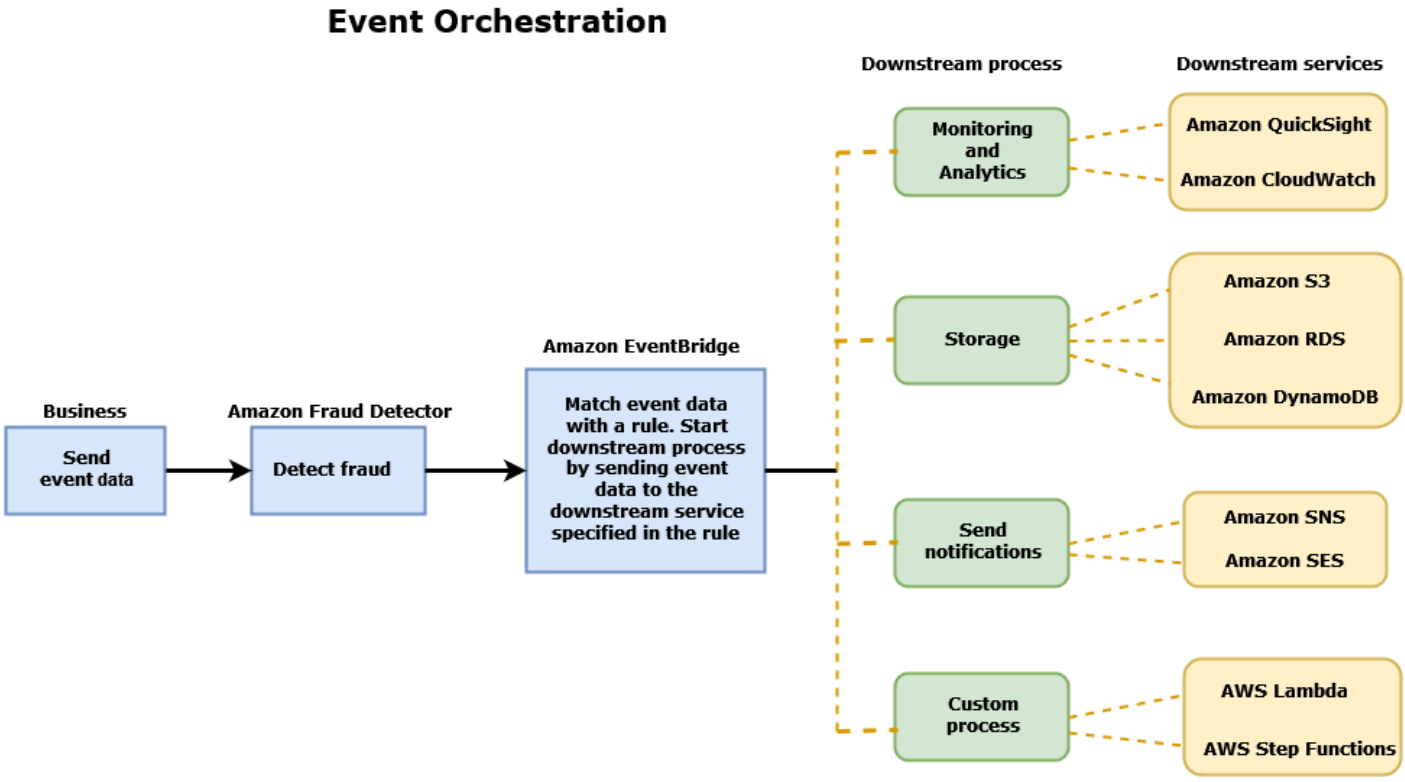
通过事件编排，您可以轻松地使用 Amazon EventBridge 将事件发送到AWS 服务下游处理。Amazon Fraud Detector 为您提供了一些简单的规则，您可以在检测到欺诈后使用这些规则自动处理事件。通过事件编排，您可以自动执行下游事件流程，例如将事件发送到仪表板以从事件数据中获取见解，根据欺诈检测结果生成通知，以及根据从欺诈检测中吸取的经验使用标签更新事件。

通过事件编排，您可以通过 Amazon EventBridge 轻松访问AWS环境中的服务。您可以将 Amazon 配置 EventBridge 为直接向AWS 服务或使用 API 目的地间接发送事件。用于协调下游流程的也称为目标。AWS 服务您可以用来协调下游处理的一些目标如下：

- 用于监控和分析 — 亚马逊 QuickSight、亚马逊 CloudWatch
- 用于存储 — 亚马逊 S3、亚马逊 RDS、亚马逊 D ynamo DB
- 用于发送通知 — 亚马逊 SNS、亚马逊 SES
- 用于自定义处理 — AWS Lambda、AWS Step Functi on s

有关 Amazon 支持的编排目标的更多信息 EventBridge，请参阅 Ama zon EventBridge 目标。

下图提供了事件编排工作原理的高级视图。



设置事件编排

为事件设置事件编排需要您在目标服务中设置流程，将 Amazon 配置 EventBridge 为接收和发送事件数据，并在 Amazon EventBridge 中创建规则来指定启动下游流程的条件。完成以下步骤来设置事件编排：

设置事件编排

1. 前往亚马逊 EventBridge 用户指南并学习如何使用亚马逊 EventBridge。请务必学习如何在 Amazon 中 EventBridge 为您的用例创建规则。
2. 按照说明进行操作在 Amazon Fraud Detector 中启用事件编排。

> **Note**
>
> 默认情况下，您的活动的活动编排处于禁用状态。

3. 将目标服务设置为接收和处理事件数据。例如，如果您的下游流程涉及发送通知，而您想使用 Amazon SNS，请前往 Amazon SNS 控制台，创建 SNS 主题，然后通过终端节点订阅该主题。
4. 按照说明创建 Amazon EventBridge 规则。

> **Important**
>
> 在 Amazon 中构建事件模式时 EventBridge，请务必`aws.frauddetector`提供来源字段和`Event Prediction Result Returned`详情类型字段。

在 Amazon Fraud Detector 中启用事件编排

您可以在创建事件类型时或在创建事件类型之后为事件启用事件编排。可以在 Amazon Fraud Detector 控制台中启用事件编排，使用`put-event-type`命令、`PutEventType` API 或使用。AWS SDK for Python (Boto3)

在 Amazon Fraud Detector 控制台中启用事件编排

此示例为已创建的事件类型启用事件编排。如果您要创建新的事件类型并希望启用编排，请按照说明进行操作。创建事件类型

启用事件编排

1. 打开AWS管理控制台并登录您的账户。导航至 Amazon Fraud Detector。
2. 在左侧导航窗格中，选择 Events（事件）。
3. 在事件类型页面中，选择您的事件类型。
4. 开启 “通过 Amazon EventBridge 启用事件编排”。
5. 继续执行步骤 3 的说明设置事件编排。

使用启用事件编排 AWS SDK for Python (Boto3)

以下示例显示了更新事件类型`sample_registration`以启用事件编排的示例请求。该示例使用 `PutEventType` API，并假设您已经创建了变量`ip_addresslegit`和`email_addressfraud`、标签和以及实体类型`sample_customer`。有关如何创建这些资源的信息，请参阅资源。

```
import boto3
fraudDetector = boto3.client('frauddetector')
fraud_detector.put_event_type(
  name = 'sample_registration',
  eventVariables = ['ip_address', 'email_address'],
  eventOrchestration = {'eventBridgeEnabled': True},
  labels = ['legit', 'fraud'],
  entityTypes = ['sample_customer'])
```

在 Amazon Fraud Detector 中禁用事件编排

您可以随时在 Amazon Fraud Detector 控制台中禁用事件的事件编排，使用`put-event-type`命令、`PutEventType` API 或使用。AWS SDK for Python (Boto3)

在 Amazon Fraud Detector 控制台中禁用事件编排

禁用事件编排

1. 打开AWS管理控制台并登录您的账户。导航至 Amazon Fraud Detector。
2. 在左侧导航窗格中，选择 Events（事件）。
3. 在事件类型页面中，选择您的事件类型。
4. 关闭 “通过 Amazon EventBridge 启用事件编排”。

使用禁用事件编排 AWS SDK for Python (Boto3)

以下示例显示了使用 PutEventType API 更新事件类型sample_registration以禁用事件编排的示例请求。

```
import boto3
fraudDetector = boto3.client('frauddetector')
fraud_detector.put_event_type(
  name = 'sample_registration',
  eventVariables = ['ip_address', 'email_address'],
  eventOrchestration = {'eventBridgeEnabled': False},
  entityTypes = ['sample_customer'])
```

模型

Amazon Fraud Detector 使用机器学习模型生成欺诈预测。每个模型都使用一种模型类型进行训练。模型类型指定用于训练模型的算法和转换。模型训练是使用您提供的数据集来创建可以预测欺诈事件的模型的过程。

要创建模型，必须先选择模型类型，然后准备并提供用于训练模型的数据。

选择模型类型

Amazon Fraud Detector 中有以下模型类型。选择适合您的用例的模型类型。

- 在线欺诈洞察

  O nline Fraud Insights 模型类型经过优化，可以在几乎没有关于被评估实体的历史数据（例如，新客户在线注册新账户）时检测欺诈。
- 交易欺诈洞察

  Transaction Fraud Insights 模型类型最适合检测欺诈用例，在这种用例中，被评估的实体可能具有互动历史记录，模型可以分析这些历史记录以提高预测准确性（例如，具有过去购买历史的现有客户）。
- 账户接管见解

  账户盗用见解模型类型可检测账户是否因网络钓鱼或其他类型的攻击而遭到入侵。被盗账户的登录数据（例如登录时使用的浏览器和设备）与与该账户关联的历史登录数据不同。

在线欺诈洞察

Online Fraud Insights 是一种受监督的机器学习模型，这意味着它使用欺诈和合法交易的历史示例来训练模型。在线欺诈洞察模型可以根据少量历史数据检测欺诈。该模型的输入非常灵活，因此您可以对其进行调整以检测各种欺诈风险，包括虚假评论、促销滥用和房客结账欺诈。

Online Fraud Insights 模型使用一系列机器学习算法进行数据扩充、转换和欺诈分类。作为模型训练过程的一部分，Online Fraud Insights使用第三方数据（例如IP地址或信用卡的发卡银行）丰富了IP地址和BIN号等原始数据元素。除第三方数据外，Online Fraud Insights还使用深度学习算法，该算法考虑了在亚马逊和亚马逊上看到的欺诈模式AWS。使用梯度树提升算法，这些欺诈模式将成为模型的输入特征。

为了提高性能，Online Fraud Insights 通过贝叶斯优化过程优化梯度树提升算法的超参数。它按顺序训练数十种具有不同模型参数（例如树木数量、树木深度和每片树叶的样本数）的不同模型。它还使用不同的优化策略，例如增加少数族裔欺诈群体的权重，以应对非常低的欺诈率。

选择数据源

在训练在线欺诈洞察模型时，您可以选择根据存储在外部（Amazon Fraud Detector 之外）或存储在 Amazon Fraud Detector 中的事件数据来训练模型。Amazon Fraud Detector 目前支持的外部存储是亚马逊简单存储服务 (Amazon S3) Service。如果您使用的是外部存储，则必须将事件数据集以逗号分隔值 (CSV) 格式上传到 Amazon S3 存储桶。在模型训练配置中，这些数据存储选项被称为 EXTERNAL_EVENTS（用于外部存储）和 INGESTED_EVENTS（用于内部存储）。有关可用数据源以及如何在其中存储数据的更多信息，请参阅事件数据存储。

准备数据

无论您选择将事件数据存储在哪里（Amazon S3 或 Amazon Fraud Detector），对在线欺诈洞察模型类型的要求都是一样的。

您的数据集必须包含列标题 EVENT_LABEL。此变量将事件归类为欺诈事件或合法事件。使用 CSV 文件（外部存储）时，必须在文件中包含每个事件的 EVENT_LABEL。对于内部存储，EVENT_LABEL 字段是可选的，但必须标记所有事件才能包含在训练数据集中。在配置模型训练时，您可以选择是忽略未标记的事件，为未标记的事件假设合法标签，还是为所有未标记的事件假设欺诈性标签。

选择数据

有关选择用于训练在线欺诈洞察模型的数据的信息，请参阅收集事件数据。

在线欺诈洞察训练流程基于事件_TIMESTAMP对历史数据进行采样和分区。无需手动对数据进行采样，这样做可能会对模型结果产生负面影响。

事件变量

除了所需的事件元数据外，Online Fraud Insights 模型至少需要两个变量，这些变量已通过模型训练的数据验证，并且每个模型最多允许 100 个变量。通常，您提供的变量越多，模型就越能更好地区分欺诈和合法事件。虽然 Online Fraud Insights 模型可以支持数十个变量，包括自定义变量，但我们建议将 IP 地址和电子邮件地址包括在内，因为这些变量通常在识别被评估的实体方面最有效。

验证数据

作为培训过程的一部分，Online Fraud Insights 将验证数据集中是否存在可能影响模型训练的数据质量问题。验证数据后，Amazon Fraud Detector 将采取适当的措施来构建尽可能好的模型。这包括针对潜

在的数据质量问题发出警告，自动删除存在数据质量问题的变量，或者发出错误并停止模型训练过程。有关更多信息，请参阅数据集验证。

交易欺诈见解

交易欺诈洞察模型类型旨在检测在线欺诈或 card-not-present交易欺诈。Transaction Fraud Insights 是一种受监督的机器学习模型，这意味着它使用欺诈和合法交易的历史示例来训练模型。

Transaction Fraud Insights 模型使用一系列机器学习算法进行数据扩充、转换和欺诈分类。它利用功能工程引擎来创建实体级和事件级聚合。作为模型训练过程的一部分，Transaction Fraud Insights使用第三方数据（例如IP地址或信用卡的发卡银行）丰富了IP地址和BIN号等原始数据元素。除了第三方数据外，Transaction Fraud Insights还使用深度学习算法，这些算法考虑了在亚马逊上看到的欺诈模式，AWS这些欺诈模式使用梯度树提升算法成为模型的输入特征。

为了提高性能，Transaction Fraud Insights通过贝叶斯优化过程优化梯度树提升算法的超参数，按顺序训练数十种不同的模型，这些模型具有不同的模型参数（例如树木数量、树木深度、每片树叶的样本数）以及不同的优化策略，例如增加少数族裔欺诈群体的权重以实现极低的欺诈率。

作为模型训练过程的一部分，Transaction Fraud 模型的特征工程引擎会计算训练数据集中每个唯一实体的值，以帮助改进欺诈预测。例如，在训练过程中，Amazon Fraud Detector 会计算并存储实体上次购买的时间，并在您每次调用`GetEventPrediction`或 `SendEvent` API 时动态更新此值。在欺诈预测期间，事件变量与其他实体和事件元数据相结合，以预测交易是否为欺诈行为。

选择数据源

交易欺诈洞察模型仅使用亚马逊欺诈探测器（INGESTED_EVENTS）内部存储的数据集进行训练。这允许 Amazon Fraud Detector 持续更新有关您正在评估的实体的计算值。有关可用数据源的更多信息，请参见 事件数据存储

准备数据

在训练交易欺诈洞察模型之前，请确保您的数据文件包含准备事件数据集中提到的所有标题。Transaction Fraud Insights 模型将收到的新实体与数据集中欺诈和合法实体的示例进行比较，因此为每个实体提供许多示例会很有帮助。

Amazon Fraud Detector 会自动将存储的事件数据集转换为正确的训练格式。模型完成训练后，您可以查看性能指标并确定是否应将实体添加到训练数据集中。

选择数据

默认情况下，Transaction Fraud Insights 会根据您选择的事件类型对存储的整个数据集进行训练。您可以选择设置时间范围以减少用于训练模型的事件。设置时间范围时，请确保用于训练模型的记录有足够的时间成熟。也就是说，已经过了足够的时间来确保正确识别合法和欺诈记录。例如，对于信用卡拒付欺诈，通常需要 60 天或更长时间才能正确识别欺诈事件。为了获得最佳模型性能，请确保训练数据集中的所有记录都已成熟。

无需选择代表理想欺诈率的时间范围。Amazon Fraud Detector 会自动对您的数据进行采样，以在欺诈率、时间范围和实体数量之间取得平衡。

如果您选择的时间范围没有足够的事件来成功训练模型，Amazon Fraud Detector 将在模型训练期间返回验证错误。对于存储的数据集，EVENT_LABEL 字段是可选的，但必须对事件进行标记才能包含在训练数据集中。在配置模型训练时，您可以选择是忽略未标记的事件，为未标记的事件假设合法标签，还是为未标记的事件使用欺诈性标签。

事件变量

除了必需的事件元数据外，用于训练模型的事件类型必须包含至少 2 个变量，这些变量已通过数据验证，最多可包含 100 个变量。通常，您提供的变量越多，模型就越能更好地区分欺诈和合法事件。尽管 Transaction Fraud Insight 模型可以支持数十个变量，包括自定义变量，但我们建议您包括 IP 地址、电子邮件地址、支付工具类型、订单价格和信用卡 BIN。

验证数据

作为训练过程的一部分，Transaction Fraud Insights 会验证训练数据集是否存在可能影响模型训练的数据质量问题。验证数据后，Amazon Fraud Detector 会采取适当的措施来构建尽可能好的模型。这包括针对潜在的数据质量问题发出警告，自动删除存在数据质量问题的变量，或者发出错误并停止模型训练过程。有关更多信息，请参阅数据集验证。

Amazon Fraud Detector 将发出警告，但如果唯一实体的数量少于 1,500 个，则会继续训练模型，因为这可能会影响训练数据的质量。如果您收到警告，请查看绩效指标。

账户接管见解

Account Takeover Insights (ATI) 模型类型通过检测账户是否通过恶意收购、网络钓鱼或凭据被盗而遭到入侵，来识别欺诈性的在线活动。Account Takeover Insights 是一种机器学习模型，它使用来自在线业务的登录事件来训练模型。

您可以在实时登录流程中嵌入经过训练的账户接管洞察模型，以检测账户是否遭到入侵。该模型评估了各种身份验证和登录类型。它们包括 Web 应用程序登录、基于 API 的身份验证和 single-sign-on

(SSO)。要使用账户接管见解模型，请在出示有效的登录凭据后调用 GetEventPredictionAPI。API 会生成一个分数，用于量化账户被盗的风险。Amazon Fraud Detector 使用您定义的分数和规则返回登录事件的一个或多个结果。结果是您配置的。根据您收到的结果，您可以对每次登录采取适当的措施。也就是说，您可以批准或质疑为登录而提供的凭据。例如，您可以通过要求提供账户 PIN 作为额外验证来质疑凭证。

您还可以使用账户接管见解模型来异步评估账户登录情况，并对高风险账户采取行动。例如，可以将高风险账户添加到调查队列中，供人工审阅者确定是否需要采取进一步行动，例如暂停该账户。

Account Takeover Insights 模型使用包含您企业历史登录事件的数据集进行训练。您提供这些数据。您可以选择将账户标记为合法账户或欺诈账户。但是，这并不是训练模型所必需的。账户接管见解模型根据账户成功登录的历史记录来检测异常情况。它还学习如何检测用户行为中的异常情况，这些异常表明恶意账户盗用事件的风险增加。例如，通常使用同一组设备和 IP 地址登录的用户。欺诈者通常使用不同的设备和地理位置登录。这种技术可以得出活动异常的风险评分，这通常是恶意账户接管的主要特征。

在训练账户接管见解模型之前，Amazon Fraud Detector 使用机器学习技术的组合来进行数据扩充、数据聚合和数据转换。然后，在训练过程中，Amazon Fraud Detector 会丰富您提供的原始数据元素。原始数据元素的示例包括 IP 地址和用户代理。Amazon Fraud Detector 使用这些元素来创建描述登录数据的额外输入。这些输入包括设备、浏览器和地理位置输入。Amazon Fraud Detector 还使用您提供的登录数据来持续计算描述过去用户行为的聚合变量。用户行为的示例包括用户从特定 IP 地址登录的次数。使用这些额外的增强功能和聚合，Amazon Fraud Detector 可以从您的登录事件中获得少量输入来生成强大的模型性能。

Account Takeover Insights 模型可以检测不良行为者访问合法账户的实例，无论不良行为者是人类还是机器人。该模型生成一个单一分数，用于指示账户被盗的相对风险。可能已被盗用的账户会被标记为高风险账户。您可以通过以下两种方式之一处理高风险账户。或者，您也可以强制执行额外的身份验证。或者，您可以将账户发送到队列进行手动调查。

选择数据源

账户接管见解模型是根据存储在内部的 Amazon Fraud Detector 中的数据集进行训练的。要使用 Amazon Fraud Detector 存储您的登录事件数据，请创建一个包含用户登录事件的 CSV 文件。对于每个事件，包括登录数据，例如事件时间戳、用户 ID、IP 地址、用户代理以及登录数据是否有效。创建 CSV 文件后，首先将文件上传到 Amazon Fraud Detector，然后使用导入功能存储数据。然后，您可以使用存储的数据训练模型。有关使用 Amazon Fraud Detector 存储事件数据集的更多信息，请参阅使用亚马逊Fraud Detector 在内部存储您的事件数据

准备数据

Amazon Fraud Detector 要求您以逗号分隔值 (CSV) 文件形式提供您的用户账户登录数据，该文件以 UTF-8 格式编码。CSV 文件的第一行必须包含文件头。文件头由描述每个数据元素的事件元数据和事件变量组成。标题后面有事件数据。事件数据中的每一行都由来自单个登录事件的数据组成。

对于 Accounts Takeover Insights 模型，您必须在 CSV 文件的标题行中提供以下事件元数据和事件变量。

事件元数据

我们建议您在 CSV 文件标题中提供以下元数据。事件元数据必须使用大写字母。

- EVENT_ID-登录事件的唯一标识符。
- ENTITY_TYPE-执行登录事件的实体，例如商家或客户。
- ENTITY_ID-执行登录事件的实体的标识符。
- EVENT_TIMESTAMP-登录事件发生的时间戳。时间戳必须采用 ISO 8601 标准（世界标准时间）。
- EVENT_LABEL（推荐）-将事件归类为欺诈或合法事件的标签。您可以使用任何标签，例如 “欺诈”、“合法”、“1” 或 “0”。

> **Note**
>
> - 事件元数据必须使用大写字母。它区分大小写。
> - 登录事件不需要标签。但是，我们建议您包含 EVENT_LABEL 元数据并为登录事件提供标签。如果标签不完整或不完整，也没关系。如果您提供标签，Amazon Fraud Detector 将使用它们来自动计算账户接管发现率，并将其显示在模型绩效图表和表格中。

事件变量

对于 Accounts Takeover Insights 模型，您必须提供必需（必须）变量和可选变量。创建变量时，请确保将变量分配给正确的变量类型。作为模型训练过程的一部分，Amazon Fraud Detector 使用与变量关联的变量类型来进行变量扩充和特征工程。

> **Note**
>
> 事件变量名称必须使用小写字母。它们区分大小写。

必填变量

训练账户接管洞察模型需要以下变量。

类别	变量类型	描述
IP 地址	IP_ADDRESS	登录事件中使用的 IP 地址
浏览器和设备	用户代理	登录事件中使用的浏览器、设备和操作系统
有效的凭证	VALIDCRED	表示用于登录的凭据是否有效

可选变量

以下变量是训练账户接管洞察模型的可选变量。

类别	类型	描述
浏览器和设备	指纹	浏览器或设备指纹的唯一标识符
会话 ID	SESSION_ID	身份验证会话的标识符
标签	事件标签	将事件归类为欺诈性或合法性的标签。您可以使用任何标签，例如 “欺诈”、“合法”、“1” 或 “0”。
Timestamp	LABEL_TIMESTAMP	标签上次更新的时间戳。如果提供了 EVENT_LABEL，则这是必需的。

Note

- 您可以为两个必填变量可选变量提供任何变量名称。必须将每个必填变量和可选变量分配给正确的变量类型。

- 您可以提供其他变量。但是，Amazon Fraud Detector 不会在训练账户接管见解模型时包含这些变量。

选择数据

收集数据是创建 “账户接管洞察” 模型的重要一步。开始收集登录数据时，请考虑以下要求和建议：

必填

- 提供至少 1,500 个用户账户示例，每个示例至少包含两个关联的登录事件。
- 您的数据集必须涵盖至少 30 天的登录事件。您可以稍后指定用于训练模型的事件的特定时间范围。

推荐

- 您的数据集包含登录失败事件的示例。您可以选择将这些失败的登录标记为 “欺诈” 或 “合法”。
- 使用跨越六个月的登录事件准备历史数据，包括 10 万个实体。

如果您没有已满足最低要求的数据集，可以考虑通过调用 SendEventAPI 操作将事件数据流式传输到 Amazon Fraud Detector。

验证数据

在创建账户接管见解模型之前，Amazon Fraud Detector 会检查您在数据集中包含的用于训练模型的元数据和变量是否符合大小和格式要求。有关更多信息，请参阅数据集验证：它还会检查其他要求。如果数据集未通过验证，则不会创建模型。要成功创建模型，请务必在再次训练之前修复未通过验证的数据。

常见的数据集错误

在验证用于训练账户接管见解模型的数据集时，Amazon Fraud Detector 会扫描这些问题和其他问题，如果遇到一个或多个问题，则会抛出错误。

- CSV 文件不是 UTF-8 格式。
- CSV 文件标头不包含以下元数据中的至少一个：`EVENT_IDENTITY_ID`、或`EVENT_TIMESTAMP`。
- CSV 文件头不包含以下变量类型的至少一个变量：`IP_ADDRESS`、`USERAGENT`、或`VALIDCRED`。
- 有不止一个变量与同一个变量类型相关联。
- 中超过 0.1% 的值`EVENT_TIMESTAMP`包含空值或支持的日期和时间戳格式以外的值。

- 从第一个事件到最后一个事件之间的天数少于 30 天。
- 该变量类型的IP_ADDRESS变量中有超过 10% 无效或为空。
- 超过 50% 的变量类型的USERAGENT变量包含空值。
- 变量类型的所有VALIDCRED变量都设置为false。

构建模型

Amazon Fraud Detector 模型学习检测特定事件类型的欺诈行为。在 Amazon Fraud Detector 中，您首先创建一个模型，该模型充当模型版本的容器。每次训练模型时，都会创建一个新版本。有关如何使用AWS控制台创建和训练模型的详细信息，请参阅步骤 3：创建模型。

每个模型都有相应的模型分数变量。当你创建模型时，Amazon Fraud Detector 会代表你创建这个变量。您可以在规则表达式中使用此变量来解释欺诈评估期间的模型分数。

使用训练和部署模型 AWS SDK for Python (Boto3)

模型版本是通过调用CreateModel和CreateModelVersion操作创建的。 CreateModel启动模型，该模型充当模型版本的容器。 CreateModelVersion启动训练过程，从而生成模型的特定版本。每当您调用 CreateModelVersion 时，就会创建一个新版本的解决方案。

以下示例显示了对 CreateModel API 的请求示例。此示例创建了 Online Fraud Insights 模型类型，并假设您已经创建了事件类型sample_registration。有关创建事件类型的更多详细信息，请参阅创建事件类型。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.create_model (
modelId = 'sample_fraud_detection_model',
eventTypeName = 'sample_registration',
modelType = 'ONLINE_FRAUD_INSIGHTS')
```

使用 CreateModelVersionAPI 训练你的第一个版本。对于TrainingDataSource和，ExternalEventsDetail请指定训练数据集的源和 Amazon S3 的位置。TrainingDataSchema请指定 Amazon Fraud Detector 应如何解释训练数据，特别是要包含哪些事件变量以及如何对事件标签进行分类。默认情况下，Amazon Fraud Detector 会忽略未标记的事件。此示例代码AUTO用于unlabeledEventsTreatment指定 Amazon Fraud Detector 决定如何使用未标记的事件。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.create_model_version (
modelId = 'sample_fraud_detection_model',
modelType = 'ONLINE_FRAUD_INSIGHTS',
trainingDataSource = 'EXTERNAL_EVENTS',
trainingDataSchema = {
    'modelVariables' : ['ip_address', 'email_address'],
    'labelSchema' : {
        'labelMapper' : {
            'FRAUD' : ['fraud'],
            'LEGIT' : ['legit']
        }
       unlabeledEventsTreatment = 'AUTO'
    }
},
externalEventsDetail = {
    'dataLocation' : 's3://bucket/file.csv',
    'dataAccessRoleArn' : 'role_arn'
}
)
```

成功请求将生成一个带有状态的新模型版本TRAINING_IN_PROGRESS。在训练期间的任何时候，您都可以通过调用UpdateModelVersionStatus并将状态更新为来取消训练TRAINING_CANCELLED。训练完成后，模型版本状态将更新为TRAINING_COMPLETE。您可以使用 Amazon Fraud Detector 控制台或致电来查看模型性能DescribeModelVersions。有关如何解释模型分数和性能的更多信息，请参阅模型分数和对性能指标进行建模。

查看模型性能后，激活该模型，使其可供探测器在实时欺诈预测中使用。Amazon Fraud Detector 将在多个可用区域部署模型以实现冗余，同时开启自动缩放功能，确保模型可根据您所做的欺诈预测数量进行扩展。要激活模型，请调用 UpdateModelVersionStatus API 并将状态更新为ACTIVE。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.update_model_version_status (
modelId = 'sample_fraud_detection_model',
modelType = 'ONLINE_FRAUD_INSIGHTS',
modelVersionNumber = '1.00',
status = 'ACTIVE'
```

)

模型分数

Amazon Fraud Detector 为不同的模型类型生成模型分数的方式不同。

对于账户接管见解 (ATI) 模型，Amazon Fraud Detector 仅使用汇总值（通过组合一组原始变量计算得出的值）来生成模型分数。新实体的第一个事件得分为 -1，表示存在未知风险。这是因为对于新实体，用于计算聚合的值将为零或空。Account Takeeover Insights (ATI) 模型为同一实体和现有实体的所有后续事件生成介于 0 到 1000 之间的模型分数，其中 0 表示欺诈风险低，1000 表示欺诈风险高。对于 ATI 模型，模型分数与挑战率 (CR) 直接相关。例如，500分对应于估计的5%的挑战率，而900分对应于估计的0.1%的挑战率。

对于在线欺诈见解 (OFI) 和交易欺诈洞察 (TFI) 模型，Amazon Fraud Detector 使用汇总值（通过组合一组原始变量计算得出的值）和原始值（为变量提供的值）来生成模型分数。模型分数可以介于 0 到 1000 之间，其中 0 表示欺诈风险低，1000 表示欺诈风险高。对于 OFI 和 TFI 模型，模型分数与误报率 (FPR) 直接相关。例如，600分数对应于估计的10%的假阳性率，而900分对应于估计的2%的假阳性率。下表详细说明了某些模型分数与估计的误报率的相关性。

模特分数	估计 FPR
975	0.50%
950	1%
900	2%
860	3%
775	5%
700	7%
600	10%

对性能指标进行建模

模型训练完成后，Amazon Fraud Detector 会使用未用于训练模型的数据的 15% 来验证模型性能。您可以期望经过训练的 Amazon Fraud Detector 模型具有与验证绩效指标相似的真实欺诈检测性能。

作为一家企业，您必须在发现更多欺诈行为和给合法客户增加更多摩擦之间取得平衡。为了帮助选择适当的平衡点，Amazon Fraud Detector 提供了以下工具来评估模型性能：

- 分数分布图 — 模型分数分布的直方图假设示例总数为 100,000 个事件。左 Y 轴代表合法事件，右 Y 轴代表欺诈事件。您可以通过单击图表区域来选择特定的模型阈值。这将更新混淆矩阵和 ROC 图表中的相应视图。
- 混淆矩阵 — 通过比较模型预测与实际结果，汇总给定分数阈值下的模型精度。Amazon Fraud Detector 假设示例事件数量为 100,000 个。欺诈和合法事件的分布模拟了您企业中的欺诈率。
  - 真正的积极方面 — 模型预测欺诈，而该事件实际上是欺诈。
  - 误报 — 模型预测欺诈，但该事件实际上是合法的。
  - 真正的负面因素 — 模型预测合法，而事件实际上是合法的。
  - 假阴性 — 模型预测是合法的，但该事件实际上是欺诈。
  - 真实阳性率 (TPR)-模型检测到的欺诈总数的百分比。也称为捕获率。
  - 误报率 (FPR)-被错误预测为欺诈的合法事件总数的百分比。
- 接收器运算符曲线 (ROC)-将真实阳性率绘制为假阳性率与所有可能的模型分数阈值的函数。通过选择 “高级指标” 查看此图表。
- 曲线下方区域 (AUC)-汇总所有可能的模型分数阈值上的 TPR 和 FPR。没有预测能力的模型的AUC为0.5，而完美模型的分数为1.0。
- 不确定性范围 — 它显示模型预期的 AUC 范围。范围越大（AUC 的上限和下限差异> 0.1）意味着模型的不确定性越高。如果不确定性范围很大 (>0.1)，请考虑提供更多标记的事件并重新训练模型。

使用模型性能指标

1. 从分数分布图开始，查看您的欺诈和合法事件的模型分数分布。理想情况下，您将明确区分欺诈和合法事件。这表明模型可以准确识别哪些事件是欺诈性的，哪些是合法的。通过单击图表区域选择模型阈值。您可以看到调整模型分数阈值如何影响您的真阳性和误报率。

   > **Note**
   >
   > 分数分布图在两个不同的 Y 轴上绘制了欺诈和合法事件。左 Y 轴代表合法事件，右 Y 轴代表欺诈事件。

2. 查看混淆矩阵。根据您选择的模型分数阈值，您可以查看基于 100,000 个事件样本的模拟影响。欺诈和合法事件的分布模拟了您企业中的欺诈率。使用这些信息在真阳性率和误报率之间找到适当的平衡。

3. 要了解更多详细信息，请选择高级指标。使用 ROC 图表来了解任何模型分数阈值的真阳性率和误报率之间的关系。ROC 曲线可以帮助你微调真阳性率和误报率之间的权衡。

> **Note**
>
> 您也可以通过选择 “表格” 来查看表格形式的指标。
> 表格视图还显示指标精度。精确度是指正确预测为欺诈事件的欺诈事件与所有预测为欺诈事件的百分比。

4. 根据您的目标和欺诈检测用例，使用绩效指标为您的企业确定最佳模型阈值。例如，如果您计划使用该模型将新账户注册分为高、中或低风险，则需要确定两个阈值分数，这样您就可以起草三个规则条件，如下所示：

   - 分数 > X 为高风险
   - 分数 < X but > Y 为中等风险
   - 分数 < Y 表示风险较低

模型变量重要性

模型变量重要性是 Amazon Fraud Detector 的一项功能，它可以对模型版本中的模型变量进行排名。根据每个模型变量对模型整体性能的相对重要性为其提供一个值。值最高的模型变量对模型来说比该模型版本的数据集中的其他模型变量更为重要，并且默认情况下列在顶部。同样，默认情况下，值最低的模型变量列在底部，与其他模型变量相比，其重要性最小。使用模型变量重要性值，您可以深入了解哪些输入正在推动模型的性能。

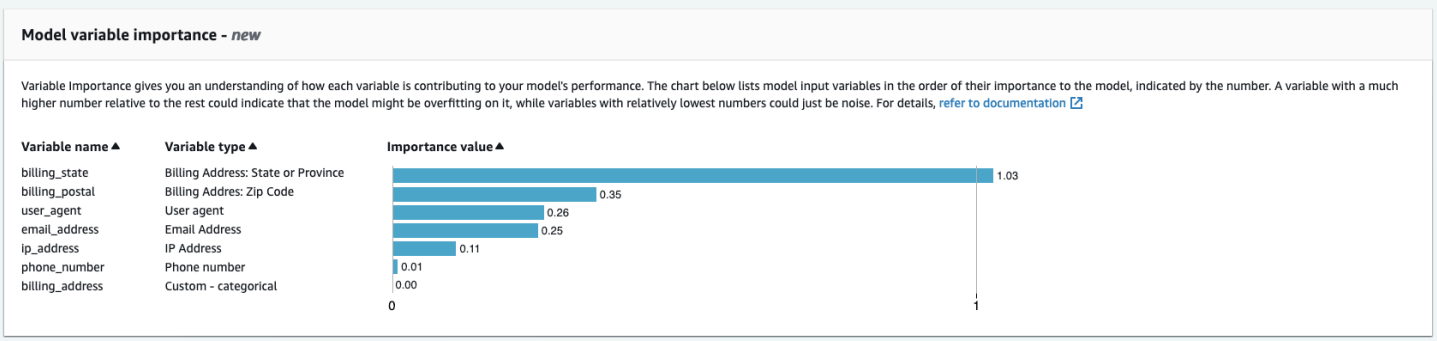
您可以在 Amazon Fraud Detector 控制台中或使用 DescribeModelVersionAPI 查看经过训练的模型版本的模型变量重要性值。

模型变量重要性为用于训练模型版本的每个变量提供以下一组值。

- 变量类型：变量的类型（例如，IP 地址或电子邮件）。有关更多信息，请参阅变量类型：对于账户接管见解 (ATI) 模型，Amazon Fraud Detector 为原始变量和聚合变量类型提供了可变重要性值。原始变量类型将分配给您提供的变量。聚合变量类型被分配给一组原始变量，Amazon Fraud Detector 已将这些变量组合在一起，计算出聚合的重要性值。
- 变量名称：用于训练模型版本的事件变量的名称（例如、`ip_address`、`email_address`、`are_creadentials_valid`）。对于聚合变量类型，将列出用于计算聚合变量重要性值的所有变量的名称。

- 变量重要性值：一个数字，表示原始变量或聚合变量对模型性能的相对重要性。典型范围：0—10

在 Amazon Fraud Detector 控制台中，在线欺诈见解 (OFI) 或交易欺诈洞察 (TFI) 模型的模型变量重要性值显示如下。除了原始变量的重要性值外，账户接管洞察 (ATI) 模型还将提供聚合的变量重要性值。可视化图表可以很容易地看到变量之间的相对重要性，垂直虚线提供了排名最高的变量的重要性值的参考。



Amazon Fraud Detector 可以为每个 Fraud Detector 模型版本生成可变的重要性值，无需支付额外费用。

> **Important**
>
> 2021 年 7 月 9 日之前创建的模型版本没有可变的重要性值。必须训练模型的新版本才能生成模型变量重要性值。

使用模型变量重要性值

您可以使用模型变量重要性值来深入了解推动模型性能向上或向下的因素，以及哪些变量的贡献最大。然后调整模型以提高整体性能。

更具体地说，要提高模型性能，请根据您的领域知识检查变量重要性值，并调试训练数据中的问题。例如，如果使用账户 ID 作为模型的输入，并且它列在顶部，请查看其变量重要性值。如果变量重要性值明显高于其余值，则您的模型可能过于适合特定的欺诈模式（例如，所有欺诈事件都来自同一个账户 ID）。但是，如果变量依赖于欺诈标签，也可能存在标签泄露的情况。根据您基于领域知识的分析结果，您可能需要移除变量并使用更加多样化的数据集进行训练，或者保持模型原样。

同样，看看排在最后的变量。如果变量重要性值明显低于其余值，则此模型变量在训练模型时可能没有任何重要性。你可以考虑移除该变量来训练更简单的模型版本。如果您的模型变量很少，例如只有两个变量，则 Amazon Fraud Detector 仍会提供变量重要性值并对变量进行排名。但是，在这种情况下，见解将是有限的。

Important

1. 如果您发现模型变量重要性图表中缺少变量，则可能是由于以下原因之一。考虑修改数据集中的变量并重新训练模型。
   - 训练数据集中变量的唯一值计数小于 100。
   - 训练数据集中缺少大于 0.9 的变量值。
2. 每次要调整模型的输入变量时，都需要训练一个新的模型版本。

评估模型变量重要性值

我们建议您在评估模型变量重要性值时考虑以下因素：

- 必须始终将变量重要性值与领域知识结合起来进行评估。
- 检查模型版本中某个变量相对于其他变量的变量重要性值的变量重要性值。不要单独考虑单个变量的变量重要性值。
- 比较同一模型版本中变量的变量重要性值。不要比较不同模型版本中相同变量的变量重要性值，因为模型版本中变量的变量重要性值可能不同于不同模型版本中相同变量的值。如果您使用相同的变量和数据集来训练不同的模型版本，则不一定会生成相同的变量重要性值。

查看模型变量重要性排名

模型训练完成后，您可以在 Amazon Fraud Detector 控制台中或使用 DescribeModelVersionAPI 查看训练过的模型版本的模型变量重要性排名。

要使用控制台查看模型变量重要性排名，

1. 打开AWS控制台并登录您的账户。导航至 Amazon Fraud Detector。
2. 在左侧导航窗格中，选择 Models (模型)。
3. 选择您的型号，然后选择您的模型版本。
4. 确保已选择 “概览” 选项卡。
5. 向下滚动查看模型变量重要性窗格。

了解模型变量重要性值的计算方式

完成每个模型版本训练后，Amazon Fraud Detector 会自动生成模型变量重要性值和模型的性能指标。为此，Amazon Fraud Detector 使用了 Shapley 添加剂解释 (SHA P)。SHAP 本质上是考虑所有模型变量的所有可能组合后模型变量的平均预期贡献。

SHAP 首先分配每个模型变量的贡献以预测事件。然后，它汇总这些预测以创建模型级别的变量排名。为了为预测分配每个模型变量的贡献，SHAP 会考虑所有可能的变量组合之间模型输出的差异。通过包括包括或删除特定变量集以生成模型输出的所有可能性，SHAP 可以准确地访问每个模型变量的重要性。当模型变量彼此高度相关时，这一点尤其重要。

在大多数情况下，机器学习模型不允许您移除变量。相反，您可以将模型中已删除或缺失的变量替换为一个或多个基准中的相应变量值（例如，非欺诈事件）。选择合适的基准实例可能很困难，但是 Amazon Fraud Detector 通过将此基线设置为人口平均值来简化这一操作。

导入 SageMaker 模型

您可以选择将 SageMaker托管模型导入到 Amazon Fraud Detector。与模型类似，可以将 SageMaker 模型添加到探测器中，并使用 `GetEventPrediction` API 生成欺诈预测。作为`GetEventPrediction`请求的一部分，Amazon Fraud Detector 将调用您的 SageMaker 终端节点并将结果传递给您的规则。

您可以将 Amazon Fraud Detector 配置为使用作为`GetEventPrediction`请求一部分发送的事件变量。如果您选择使用事件变量，则必须提供输入模板。Amazon Fraud Detector 将使用此模板将您的事件变量转换为调用 SageMaker 终端节点所需的输入有效负载。或者，您可以将 SageMaker 模型配置为使用作为请求一部分发送的 ByteBuffer。`GetEventPrediction`

Amazon Fraud Detector 支持导入使用 JSON 或 CSV 输入格式以及 JSON 或 CSV 输出格式的 SageMaker 算法。支持的 SageMaker 算法示例包括 XGBoost、Linear Learner 和 Random Cut Forest。

使用导入 SageMaker 模型 AWS SDK for Python (Boto3)

要导入 SageMaker 模型，请使用 `PutExternalModel` API。以下示例假设 SageMaker 终端节点`sagemaker-transaction-model`已部署，处于`InService`状态，并且使用 XGBoost 算法。

输入配置指定将使用事件变量来构造模型输入（设置`useEventVariables`为`TRUE`）。鉴于 xgBoost 需要 CSV 输入，输入格式为 TEXT_CSV。 csvInputTemplate 指定如何根据作为`GetEventPrediction`请求的一部分发送的变量构建 CSV 输入。此示例假设您已经创建了变量`order_amt`、`prev_amt`、`hist_amt`和`payment_type`。

输出配置指定 SageMaker 模型的响应格式，并将相应的 CSV 索引映射到 Amazon Fraud Detector 变量`sagemaker_output_score`。配置完成后，即可在规则中使用输出变量。

> **Note**
>
> SageMaker 模型的输出必须映射到带源的变量`EXTERNAL_MODEL_SCORE`。您无法使用变量在控制台中创建这些变量。改为在配置模型导入时必须创建它们。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.put_external_model (
modelSource = 'SAGEMAKER',
modelEndpoint = 'sagemaker-transaction-model',
invokeModelEndpointRoleArn = 'your_SagemakerExecutionRole_arn',
inputConfiguration = {
    'useEventVariables' : True,
    'eventTypeName' : 'sample_transaction',
    'format' : 'TEXT_CSV',
    'csvInputTemplate' : '{{order_amt}}, {{prev_amt}}, {{hist_amt}}, {{payment_type}}'
},

outputConfiguration = {
    'format' : 'TEXT_CSV',
    'csvIndexToVariableMap' : {
        '0' : 'sagemaker_output_score'
    }
},

modelEndpointStatus = 'ASSOCIATED'
)
```

删除模型或模型版本

您可删除 Amazon Fraud Detector 中的模型和模型版本，前提是它们未与探测器版本关联。当您删除模型时，Amazon Fraud Detector 会永久删除该模型，并且该数据将不再存储在 Amazon Fraud Detector 中。

您也可以删除Amazon SageMaker 模型，前提是它们未与探测器版本关联。删除 SageMaker 模型会将其与 Amazon Fraud Detector 断开连接，但该模型在中仍然可用 SageMaker。

删除模型版本

您只能删除处于Ready to deploy状态的模型版本。要将模型版本从更改ACTIVE为Ready to deploy状态，请取消部署模型版本。

1. 登录到AWS Management Console并打开 Amazon Fraud Detector，，，，，，，，，，，https://console.aws.amazon.com/frauddetector
2. 在亚马逊Fraud Detector 控制台，选择模型。
3. 选择包含您要删除的模型版本的模型。
4. 选择要删除的模型版本。
5. 选择 Actions，然后选择 Delete。
6. 输入模型版本名称，然后选择删除模型版本。

取消部署模型版本

您无法取消部署任何检测器版本正在使用的模型版本 (ACTIVE,INACTIVE,DRAFT)。因此，要取消部署探测器版本正在使用的模型版本，请先从探测器版本中移除该模型版本。

1. 在亚马逊Fraud Detector 控制台，选择模型。
2. 选择包含要取消部署的模型版本的模型。
3. 选择要删除的模型版本。
4. 选择 “操作”，然后选择 “取消部署模型版本”。

删除模型

在删除模型之前，您必须先删除与该模型关联的所有模型版本。

1. 在亚马逊Fraud Detector 控制台，选择模型。
2. 选择要删除的模型。
3. 选择 Actions，然后选择 Delete。
4. 输入模型名称，然后选择删除模型。

移除亚马逊 SageMaker 模特

1. 在亚马逊Fraud Detector 控制台，选择模型。
2. 选择要删除的 SageMaker 模型。

3. 选择 “操作”，然后选择 “移除模型”。
4. 输入模型名称，然后选择删除 SageMaker模型。

探测器

检测器是包含欺诈检测逻辑（例如模型和规则）的容器，用于您要评估的特定业务事件是否存在欺诈。首先，您可以通过指定已经定义的事件来创建检测器，然后可以选择添加已由 Amazon Fraud Detector 为该事件创建和训练的模型版本。

然后，向检测器添加规则和规则执行顺序，以创建检测器版本。检测器版本定义了规则，也可以定义模型，该模型将作为生成欺诈预测的请求的一部分运行。您可以将检测器中定义的任何规则添加到检测器版本中。您还可以将根据评估的事件类型训练的任何模型添加到探测器版本中。一个探测器可以有多个版本，每个版本都有不同的规则和规则执行顺序，以满足多个用例。

每个探测器版本的状态必须为`DRAFT`，`ACTIVE`，或`INACTIVE`。只能有一个探测器版本`ACTIVE`一次处于状态。亚马逊欺诈检测器将检测器版本与`ACTIVE`状态以生成欺诈预测。

创建探测器

您可以通过指定已定义的事件类型来创建检测器。您可以选择添加已由 Amazon Fraud Detector 训练和部署的模型。如果您添加模型，则可以在创建规则时在规则表达式中使用由 Amazon Fraud Detector 生成的模型分数（例如`$model score < 90`）。

您可以在 Amazon Fraud Detector 控制台中创建检测器，使用PutDetectorAPI，使用put 探测器命令，或者使用AWSSDK。如果您使用 API、命令或 SDK 创建探测器，则在创建探测器后，请按照说明进行操作创建探测器版本。

在亚马逊欺诈检测器控制台中创建检测器

此示例假设您已经创建了事件类型，还创建并部署了要用于欺诈预测的模型版本。

步骤 1：构建探测器

1. 在亚马逊欺诈检测器控制台的左侧导航窗格中，选择探测器。
2. 选择创建探测器。
3. 在定义探测器详细信息页面，输入`sample_detector`用于输入探测器名称。（可选）输入探测器的描述，例如`my sample fraud detector`。
4. 对于事件类型，选择您为欺诈预测创建的事件类型。
5. 选择下一步。

步骤 2：添加已部署的模型版本

1. 请注意，这是一个可选步骤。您无需向探测器添加模型。要跳过此步，选择 Next（下一步）。
2. 在添加型号-可选，选择添加模型。
3. 在添加模型页面，用于选择型号，选择您之前部署的亚马逊欺诈检测器型号名称。对于选择版本，选择已部署模型的模型版本。
4. 选择 Add model (添加模型)。
5. 选择下一步。

步骤 3：添加规则

规则是一个条件，用于告知 Amazon Fraud Detector 在评估欺诈预测时如何解释变量值。此示例将使用模型分数作为变量值创建三条规则：`high_fraud_risk`，`medium_fraud_risk`，以及`low_fraud_risk`。要创建自己的规则、规则表达式、规则执行顺序和结果，请使用适合您的模型和用例的值。

1. 在添加规则页面，下方定义规则，输入`high_fraud_risk`用于规则名称及下方描述-可选，输入**`This rule captures events with a high ML model score`**作为规则的描述。
2. 在表情，使用 Amazon Fraud Detector 简化规则表达式语言输入以下规则表达式：

   ```
   $sample_fraud_detection_model_insightscore > 900
   ```

3. 在成果，选择创造新结果。结果是欺诈预测的结果，如果在评估期间规则匹配，则返回结果。
4. 在创造新结果，输入`verify_customer`作为结果名称。（可选）输入描述。
5. 选择保存结果。
6. 选择添加规则运行规则验证检查器并保存规则。创建后，Amazon Fraud Detector 会将该规则用于您的检测器。
7. 选择添加另一条规则，然后选择创建规则选项卡。
8. 再重复两次这个过程来创建你的`medium_fraud_risk`和`low_fraud_risk`使用以下规则详细信息的规则：

   - medium_fraud_risk

     规则名称：`medium_fraud_risk`

     结果：`review`

表情：

```
$sample_fraud_detection_model_insightscore <= 900 and

$sample_fraud_detection_model_insightscore > 700
```

- 低欺诈风险

  规则名称：low_fraud_risk

  结果：approve

  表情：

  ```
  $sample_fraud_detection_model_insightscore <= 700
  ```

9. 为用例创建所有规则后，选择下一步。

   有关创建和编写规则的更多信息，请参阅规则和规则语言参考。

步骤 4：配置规则执行和规则顺序

检测器中包含的规则的规则执行模式决定了是否对您定义的所有规则进行了评估，或者规则评估是否在第一个匹配的规则处停止。规则顺序决定了你希望规则的运行顺序。

默认规则执行模式为FIRST_MATCHED。

第一次匹配

首次匹配的规则执行模式根据定义的规则顺序返回第一个匹配规则的结果。如果指定 FIRST_MATCHED，Amazon Fraud Detector 会按顺序评估规则，从第一个到最后一个，在第一个匹配的规则处停止。然后，亚马逊欺诈检测器会提供该单一规则的结果。

您运行规则的顺序可能会影响最终的欺诈预测结果。创建规则后，请按照以下步骤重新排序规则以按所需顺序运行它们：

如果你的high_fraud_risk规则还不在规则列表的顶部，选择订购，然后选择1。这会移动high_fraud_risk到第一个位置。

重复这个过程，这样你的medium_fraud_risk规则在第二位，而你的low_fraud_risk规则在第三位。

全部匹配

无论规则顺序如何，所有匹配的规则执行模式都会返回所有匹配规则的结果。如果你指定ALL_MATCHED，Amazon Fraud Detector 会评估所有规则并返回所有匹配规则的结果。

选择FIRST_MATCHED对于本教程，然后选择下一步。

步骤 5：查看并创建探测器版本

检测器版本定义了用于生成欺诈预测的特定模型和规则。

1. 在查看并创建页面，查看您配置的探测器详细信息、模型和规则。如果您需要进行任何更改，请选择编辑在相应的部分旁边。
2. 选择创建探测器。创建后，您的探测器的第一个版本将出现在探测器版本表中Draft状态。

   你使用草稿用于测试您的探测器的版本。

使用创建探测器AWS SDK for Python (Boto3)

以下示例显示了请求的示例PutDetectorAPI。探测器充当您的探测器版本的容器。这个PutDetectorAPI 指定探测器将评估的事件类型。以下示例假设您已经创建了事件类型sample_registration。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.put_detector (
detectorId = 'sample_detector',
eventTypeName = 'sample_registration'
)
```

创建探测器版本

检测器版本定义了规则、规则执行顺序以及可选的模型版本，这些版本将用作生成欺诈预测的请求的一部分。您可以将检测器中定义的任何规则添加到检测器版本中。您还可以添加根据评估的事件类型训练的任何模型。

每个探测器版本的状态为DRAFT，ACTIVE，或INACTIVE。只能有一个探测器版本ACTIVE一次处于状态。在这期间GetEventPrediction请求，亚马逊欺诈检测器将使用ACTIVE如果没有，则检测器DetectorVersion已指定。

规则执行模式

亚马逊欺诈检测器支持两种不同的规则执行模式：FIRST_MATCHED和ALL_MATCHED。

- 如果规则执行模式是FIRST_MATCHED，Amazon Fraud Detector 会按顺序从头到尾评估规则，在第一个匹配的规则处停止。然后，亚马逊欺诈检测器会提供该单一规则的结果。如果某条规则的计算结果为 false（不匹配），则对列表中的下一条规则进行评估。
- 如果规则执行模式是ALL_MATCHED，则评估中的所有规则都将并行执行，无论其顺序如何。Amazon Fraud Detector 会执行所有规则并返回每条匹配规则的定义结果。

使用创建探测器版本AWS SDK for Python (Boto3)

以下示例显示了请求的示例CreateDetectorVersionAPI。规则执行模式设置为FIRST_MATCHED，因此 Amazon Fraud Detector 将按顺序评估规则，先到最后一个，在第一个匹配的规则处停止。然后，Amazon Fraud Detector 会提供该单一规则在此期间的结果GetEventPrediction response。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.create_detector_version(
detectorId = 'sample_detector',
rules = [{
    'detectorId' : 'sample_detector',
    'ruleId' : 'high_fraud_risk',
    'ruleVersion' : '1'
},
{
    'detectorId' : 'sample_detector',
    'ruleId' : 'medium_fraud_risk',
    'ruleVersion' : '1'
},
{
    'detectorId' : 'sample_detector',
    'ruleId' : 'low_fraud_risk',
    'ruleVersion' : '1'
```

```
}
],
modelVersions = [{
    'modelId' : 'sample_fraud_detection_model',
    'modelType': 'ONLINE_FRAUD_INSIGHTS',
    'modelVersionNumber' : '1.00'
}],
ruleExecutionMode = 'FIRST_MATCHED'
)
```

要更新探测器版本的状态，请使用UpdateDetectorVersionStatusAPI。以下示例从更新了探测器版本状态DRAFT到ACTIVE。在... 期间GetEventPrediction请求，如果未指定检测器 ID，亚马逊欺诈检测器将使用ACTIVE探测器的版本。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.update_detector_version_status(
detectorId = 'sample_detector',
detectorVersionId = '1',
status = 'ACTIVE'
)
```

删除探测器、探测器版本或规则版本

在删除 Amazon Fraud探测器中的探测器之前，您必须先删除与该探测器关联的所有探测器版本和规则版本。

当您删除检测器、检测器版本或规则版本时，Amazon Fraud Detector 会永久删除该资源，数据将不再存储在 Amazon Fraud Detector 中。

删除探测器版本

只能删除处于DRAFT或INACTIVE状态的探测器版本。

1. 登录AWS Management Console并打开 Amazon FecydFraud Detector 控制台，网址为 https://console.aws.amazon.com/frauddetector。
2. 在亚马逊Fraud Detector 控制台的左侧导航窗格中，选择要探测器的。
3. 选择包含要删除的探测器版本的探测器。
4. 选择要删除的探测器版本。

5. 选择 Actions，然后选择 Delete。
6. 输入**delete**，然后选择 “删除检测器”。

删除规则版本

只有在任何版本ACTIVE或INACTIVE检测器版本均未使用规则版本时，才能将其删除。如有必要，在删除规则版本之前，首先将ACTIVE检测器版本移至INACTIVE，然后删除INACTIVE检测器版本。

1. 在亚马逊Fraud Detector 控制台的左侧导航窗格中，选择要探测器的。
2. 选择要删除的规则版本的探测器版本。
3. 选择 R ule（关联规则）选项卡，然后选择要删除的规则。
4. 选择要删除的规则版本。
5. 选择 “操作”，然后选择 “删除规则版本”。
6. 输入**delete**，然后选择 “删除版本”。

删除探测器

在删除某个探测器之前，您必须先删除与该探测器关联的所有探测器版本和规则版本。

1. 在亚马逊Fraud Detector 控制台的左侧导航窗格中，选择要探测器的。
2. 选择要删除的探测器。
3. 选择 “操作”，然后选择 “删除检测器”。
4. 输入**delete**，然后选择 “删除检测器”。

资源

模型、规则和检测器使用变量、结果、标签、列表和实体等资源来评估事件中的欺诈风险。本部分提供有关创建和管理资源的信息。

主题

- Variables
- Labels
- 规则
- 列表
- 结果
- 实体
- 使用管理亚马逊Fraud Detector 资源AWS CloudFormation

Variables

变量表示您要在欺诈预测中使用的数据元素。这些变量可以取自您为训练模型准备的事件数据集、您的 Amazon Fraud Detector 模型的风险评分输出或亚马逊SageMaker模型。有关取自事件数据集的变量的更多信息，请参阅使用数据模型浏览器获取事件数据集要求。

在创建事件类型时，必须先创建要在欺诈预测中使用的变量，然后将其添加到事件中。必须为创建的每个变量分配一个数据类型、一个默认值和一个可选的变量类型。Amazon Fraud Detector 丰富了您提供的一些变量，例如 IP 地址、银行识别码 (BIN) 和电话号码，以创建额外的输入并提高使用这些变量的模型的性能。

数据类型

变量必须具有该变量所代表的数据元素的数据类型，并且可以选择为其分配预定义的数据元素之一变量类型。对于分配给变量类型的变量，数据类型是预先选择的。可能的数据类型包括以下类型：

数据类型	描述	默认值	示例值
字符串	字母、整数或两者的任意组合	<empty>	abc，123，1D3B
整数	正整数或负整数	0	1，-1

数据类型	描述	默认值	示例值
布尔值	对还是错	False	真的，错误的
DateTime	仅以 ISO 8601 标准 UTC 格式指定的日期和时间	<empty>	2019-11-30T 13:01:01 Z
Float	带小数点的数字	0.0	4.01、0.10

默认值

变量必须有默认值。当 Amazon Fraud Detector 生成欺诈预测时，如果亚马逊欺诈检测器未收到变量的值，则使用此默认值来运行规则或模型。您提供的默认值必须与选定的数据类型相匹配。在 AWS 控制台中，Amazon Fraud Detector `0` 为整数、布尔值、`false`浮点数分配默认值为，`0.0`为字符串分配默认值（空）。您可以为这些数据类型中的任何一个设置自定义默认值。

变量类型

创建变量时，可以选择将变量分配给变量类型。变量类型表示用于训练模型和生成欺诈预测的常见数据元素。只有具有关联变量类型的变量才能用于模型训练。作为模型训练过程的一部分，Amazon Fraud Detector 使用与变量关联的变量类型来执行变量丰富、特征工程和风险评分。

Amazon Fraud Detector 预定义了以下可用于分配给您的变量的变量类型。

类别	变量类型	描述	数据类型	示例	
会话	IP_ADDRES S	活动期间收集的 IP 地址	字符串	192.0.2.0 注意：亚马逊欺诈检	

类别	变量类型	描述	数据类型	示例		
				测器丰富了这些数据。有关更多信息，请参阅地理定位增强		

类别	变量类型	描述	数据类型	示例
	用户代理	活动期间收集的用户代理	字符串	Mozilla 5.0（Windows NT 10.0、Win64、x64、rv:68.0）Gecko 20100101
	指纹	用于事件的设备的唯一标识符	字符串	sadfow987u234
	SESSION_ID	事件活动会话的会话 ID	字符串	sid123456789
	证书是否有效	表示用于活动登录的凭证是否有效	布尔值	True
用户	电子邮件地址	活动期间收集的电子邮件地址	字符串	abc@domain.com

类别	变量类型	描述	数据类型	示例
	PHONE_NUMBER	活动期间收集的电话号码	字符串	+1 555-0100 注意：亚马逊欺诈检测器丰富了这些数据。有关更多信息，请参阅充实电话

类别	变量类型	描述	数据类型	示例
				号码
计费	账单名称	与账单地址相关的名称	字符串	John Doe

类别	变量类型	描述	数据类型	示例
	计费电话	与账单地址关联的电话号码	字符串	+1 555-0100 注意：亚马逊欺诈检测器丰富了这些数据。有关更多信息，请参阅充实电话

类别	变量类型	描述	数据类型	示例
				号码
	账单地址_L1	账单地址的第一行	字符串	任何街道
	账单地址_L2	账单地址的第二行	字符串	任何单位 123
	账单_CITY	账单地址中的城市	字符串	任何城市
	账单状态	账单地址中的州或省	字符串	任何州或省

类别	变量类型	描述	数据类型	示例		
	计费国家	账单地址中的国家	字符串	任何国家 注意：亚马逊欺诈检测器丰富了这些数据。有关更多信息，请参阅地理		

类别	变量类型	描述	数据类型	示例		
				定位增强		

类别	变量类型	描述	数据类型	示例
	BILL_ZIP	账单地址中的邮政编码	字符串	01234 注意：亚马逊欺诈检测器丰富了这些数据。有关更多信息，请参阅地理定位

类别	变量类型	描述	数据类型	示例
				增强
配送	配送_名称	与送货地址相关的名称	字符串	John Doe

类别	变量类型	描述	数据类型	示例
	配送电话	与送货地址关联的电话号码	字符串	+1 555-0100 注意：亚马逊欺诈检测器丰富了这些数据。有关更多信息，请参阅充实电话

类别	变量类型	描述	数据类型	示例
				号码
	配送地址_L1	配送地址的第一行	字符串	123 Any Street
	配送地址_L2	送货地址的第二行	字符串	第 123 单元
	航运城市	配送地址中的城市	字符串	任何城市
	配送状态	配送地址中的州或省	字符串	任何州

类别	变量类型	描述	数据类型	示例	
	配送国家	配送地址中包含的国家/地区	字符串	任何国家 注意：亚马逊欺诈检测器丰富了这些数据。有关更多信息，请参阅地理	

类别	变量类型	描述	数据类型	示例		
				定位增强		

类别	变量类型	描述	数据类型	示例
	SPIPPING_ZIP	配送地址中的邮政编码	字符串	01234 注意：亚马逊欺诈检测器丰富了这些数据。有关更多信息，请参阅地理定位

类别	变量类型	描述	数据类型	示例
				增强
付款	订单_ID	交易的唯一标识符	字符串	LUX60
	价格	订单总价	字符串	560.00
	CURRENCY COD	ISO 4217 货币代码	字符串	USD
	付款类型	活动期间用于付款的付款方式	字符串	信用卡
	身份验证码	由信用卡发卡机构或发卡银行发送的字母数字代码	字符串	0000
	AVS	来自卡处理器的地址验证系统 (AVS) 响应码	字符串	是
产品	产品类别	订单商品的产品类别	字符串	厨房
自定义	NUMERIC	任何可以表示为实数的变量	Floa	1.224

类别	变量类型	描述	数据类型	示例
	CATEGORICAL	描述类别、区段或组的任何变量	字符串	大型
	自由表单文本	作为活动一部分捕获的任何自由格式文本（例如，客户评论或评论）	字符串	自由格式文本输入的示例

将变量分配给变量类型

如果您计划使用变量来训练模型，请务必选择正确的变量类型来分配给该变量。不正确的变量类型分配可能会对模型性能产生负面影响。以后更改赋值也可能变得非常困难，尤其是在多个模型和事件都使用了该变量的情况下。

您可以为变量分配任何一种预定义的变量类型或一种自定义变量类型 — `FREE_FORM_TEXT` `CATEGORICAL`、或`NUMERIC`。

将变量分配给正确的变量类型的重要注意事项

1. 如果变量与预定义的变量类型之一匹配，请使用它。确保变量类型与变量相对应。例如，如果您将 ip_add ress 变量分配给`EMAIL_ADDRESS`变量类型，则 ip_address 变量将无法获得 ASN、ISP、地理位置和风险评分等丰富内容。有关更多信息，请参阅变量丰富：
2. 如果变量与任何预定义的变量类型都不匹配，请按照下面列出的建议分配一种自定义变量类型。

3. 将CATEGORICAL变量类型分配给通常没有自然排序的变量，可以分为类别、段或组。你用来训练模型的数据集可能有 ID 变量，例如 m erchant_id、campaign_i d 或 policy _i d。这些变量代表群组（例如，所有具有相同 policy_id 的客户代表一个群组）。必须为具有以下数据的变量分配 CATEGORICAL 变量类型-
   - 包含诸如 customer_ID、segment_ID、color _ID 、depart_code 或 product_ID 等数据的变量。
   - 包含具有真、假或空值的布尔数据的变量。
   - 可以分组或类别的变量，例如公司名称、产品类别、卡片类型或推荐媒介。

   > **Note**
   >
   > ENTITY_ID是亚马逊欺诈检测器用于分配给 ENTITY_ID 变量的保留变量类型。ENTITY_ID 变量是启动您要评估的操作的实体的 ID。如果您正在创建交易欺诈洞察 (TFI) 模型类型，则需要提供 ENTITY_ID 变量。您需要决定数据中的哪个变量可以唯一标识启动操作的实体，并将其作为 ENTITY_ID 变量传递。将 CATEGORICAL 变量类型分配给数据集中的所有其他 ID，前提是它们存在且您正在使用它们进行模型训练。在您的数据集中不是实体的其他 ID 的示例可以是 m erchant_ID、p olicy_ID 和 campaign_ ID。

4. 为包含文本块的变量分配FREE_FORM_TEXT变量类型。FREE_FORM_TEXT 变量类型的示例有 — 用户评论、评论、日期和推荐代码。FREE_FORM_TEXT 数据包含多个由分隔符分隔的标记。分隔符可以是字母数字和下划线符号以外的任何字符。例如，用户评论和评论可以用 “空格” 分隔符，日期和推荐代码可以使用连字符作为分隔符来分隔前缀、后缀和中间部分。亚马逊欺诈检测器使用分隔符从 FREE_FORM_TEXT 变量中提取数据。
5. 将 NUMER IC 变量类型分配给实数且具有固有顺序的变量。数值变量的示例包括每周天数、事件严重程度、客户评级。尽管您可以为这些变量分配 CATEGORICAL 变量类型，但我们强烈建议将所有具有固有顺序的实数变量分配给 NUMERIC 变量类型。

变量丰富

Amazon Fraud Detector 丰富了您提供的一些原始数据元素，例如 IP 地址、银行识别码 (BIN) 和电话号码，以创建额外的输入并提高使用这些数据元素的模型的性能。丰富信息有助于识别潜在的可疑情况，并帮助模型捕获更多欺诈行为。

充实电话号码

Amazon Fraud Detector 使用与地理位置、原始承运人和电话号码有效性相关的其他信息丰富了电话号码数据。对于在 2021 年 12 月 13 日当天或之后接受培训且电话号码包含国家/地区代码 (+xxx) 的所

有模特，将自动启用电话号码增强功能。如果您在模型中加入了电话号码变量并在 2021 年 12 月 13 日之前对其进行了训练，请重新训练您的模型，使其能够利用这种扩充功能。

我们强烈建议您对电话号码变量使用以下格式，以确保成功丰富您的数据。

变量	格式	描述
PHONE_NUMBER	E.164 标准	确保在电话号码中包含国家/地区代码 (+xxx)。
计费电话和配送电话	E.164 标准	确保在电话号码中包含国家/地区代码 (+xxx)。

地理定位增强

从 2022 年 2 月 8 日起，亚马逊欺诈检测器将计算您为活动提供的 IP_ADDRESS、BILLING_ZIP 和 SHIPPING_ZIP 值之间的物理距离。计算出的距离用作欺诈检测模型的输入。

要启用地理位置增强功能，您的事件数据必须包含三个变量中的至少两个：IP_ADDRESS、BILLING_ZIP 或 SHIPPING_ZIP。此外，每个 BILLING_ZIP 和 SHIPPING_ZIP 值必须分别具有有效的 BILLING_COUNTRY 代码和 SHIPPING_COUNTRY 代码。如果您的模型在 2022 年 2 月 8 日之前经过训练且包含这些变量，则必须重新训练该模型以启用地理位置增强功能。

如果由于数据无效，亚马逊欺诈检测器无法确定与事件的 IP_ADDRESS、BILLING_ZIP 或 SHIPPING_ZIP 值关联的位置，则改用特殊的占位符值。例如，假设某个事件具有有效的 IP_ADDRESS 和 BILLING_ZIP 值，但是 SHIPPING_ZIP 值无效。在这种情况下，仅对 IP_ADDRESS —> BILLING_ZIP 进行扩充。IP_ADDRESS—>SHIPPING_ZIP 和 BILLING_ZIP—>SHIPPING_ZIP 的丰富功能尚未完成。取而代之的是，使用占位符值来代替它们。无论您的模型是否启用了地理位置丰富，模型的性能都不会改变。

你可以通过将你的 BILLING_ZIP 和 SHIPPING_ZIP 变量映射到 CUSTOM_CATEGORICAL 变量类型来选择退出地理位置增强功能。更改变量类型不会影响模型的性能。

地理位置变量格式

我们强烈建议您对地理位置变量使用以下格式，以确保成功丰富您的位置数据。

变量	格式	描述
IP_ADDRESS	IPv4 地址	例如-1.1.1.1
账单_ZIP 和 shipping_ZIP	指定国家/地区的 ISO 3166-1 alpha-2 邮政编码	有关更多信息，请参阅本主题中的国家和地区代码部分。
计费国家/地区和配送国家	ISO 3166-1 alpha-2 由两个字母组成的标准国家/地区代码	有关更多信息，请参阅本主题中的国家和地区代码部分。Amazon Fraud Detector 尝试将一个国家/地区名称的所有常见变体与其 ISO 3166-1 双字母标准国家/地区代码进行匹配。但是，我们不能保证它们会被正确匹配。

国家和地区代码

下表提供了 Amazon Fraud Detector 支持用于丰富地理位置的国家和地区的完整列表。每个国家和地区都有指定的国家/地区代码（特别是 ISO 3166-1 alpha-2 由两个字母组成的国家/地区代码）和邮政编码。

邮政编码格式

- 9-数字
- a-字母
- [X]-X 是可选的。例如，Guersney “GY9 [9] 9aa” 表示 “GY9 9aa” 和 “GY99 9aa” 均有效。使用一种格式。
- [X/XX]-可以使用 X 或 XX。例如，百慕大 “aa [aa/99]” 表示 “aa aa” 和 “aa 99” 均有效。使用这些格式中的任何一种，但不要同时使用这两种格式。
- 一些国家有固定的前缀。例如，安道尔的邮政编码为 AD999。这意味着国家/地区代码必须以字母 AD 开头，然后是三个数字。

代码	名称	邮政编码
广告	安道尔	AD999
AR	荷属安的列斯	9999
AT	奥地利	9999
AU	澳大利亚	9999
AZ	阿塞拜疆	AZ 9999
BD	孟加拉国	9999
是	比利时	9999
BG	保加利亚	9999
BM	百慕大	aa [aa/99]
BY	白俄罗斯	999999
CA	加拿大	a9a 9a9
CH	瑞士	9999
CL	智利	9999999
CO	哥伦比亚	999999
CR	哥斯达黎加	99999
CY	塞浦路斯	9999
CZ	捷克	999 99
德国	德国	99999
DK	丹麦	9999
DO	多米尼加共和国	99999

代码	名称	邮政编码
DZ	阿尔及利亚	99999
EE	爱沙尼亚	99999
ES	西班牙	99999
如果	芬兰	99999
FM	密克罗尼西亚联邦	99999
FO	法罗群岛	999
FR	法国	99999
GB	英国	a [a] 9 [a/9] 9aa
GG	根西岛	GY9 [9] 9aa
GL	格陵兰	9999
GP	瓜德罗普	99999
GT	危地马拉	99999
GU	关岛	99999
人力资源	克罗地亚	99999
呼	匈牙利	9999
IE	爱尔兰	a99 [a/9] [a/9] [a/9] [a/9]
IM	马恩岛	IM9 [9] 9aa
IN	印度	999999
IS	冰岛	999
它	意大利	99999

代码	名称	邮政编码
JE	泽西岛	JE9 [9] 9aa
日本	日本	999-9999
KR	大韩民国	99999
LI	列支敦士登	9999
LK	Sri Lanka	99999
LT	立陶宛	99999
LU	卢森堡	L9999
LV	拉脱维亚	LV-9999
MC	摩纳哥	99999
MD	摩尔多瓦共和国	9999
MH	马绍尔群岛	99999
MK	北马其顿	9999
MP	北马里亚纳群岛	99999
MQ	马提尼克	99999
公吨	马耳他	aaa 9999
MX	墨西哥	99999
我的	马来西亚	99999
NL	荷兰	9999 aa
否	挪威	9999
NZ	新西兰	9999

代码	名称	邮政编码
PH	菲律宾	9999
PK	巴基斯坦	99999
PL	波兰	99-999
PR	波多黎各	99999
PT	葡萄牙	9999-999
PW	帕劳群岛	99999
回复	重聚	99999
RO	罗马尼亚	999999
RU	俄罗斯联邦	999999
SE	瑞典	999 99
SG	新加坡	999999
是	斯洛文尼亚	9999
SK	斯洛伐克	999 99
SM	圣马力诺	99999
第	泰国	99999
TR	土耳其	99999
UA	乌克兰	99999
US	美国	99999
伙计	乌拉圭	99999
六	美属维尔京群岛	99999

代码	名称	邮政编码
WF	瓦利斯和富图纳群岛	99999
还有	马约特岛	99999
ZA	南非	9999

增强用户代理

如果您创建 Account Takeoer Insights (ATI) 模型，则必须在数据集中提供useragent变量类型的变量。此变量包含登录事件的浏览器、设备和操作系统数据。Amazon Fraud Detector 使用诸如user_agent_familyOS_family和之类的附加信息丰富了用户代理数据。device_family

创建变量

您可以在 Amazon Fraud Detector 控制台中创建变量，使用 create- variable 命令 CreateVariable，使用或使用 AWS SDK for Python (Boto3)

使用亚马逊欺诈检测器控制台创建变量

此示例创建了两个变量email_address和ip_address，并将它们分配给相应的变量类型（EMAIL_ADDRESS和IP_ADDRESS）。这些变量用作示例。如果您要创建用于模型训练的变量，请使用数据集中适合您的用例的变量。在创建变量变量丰富之前变量类型，请务必阅读和内容。

要创建变量，

1. 打开AWS管理控制台并登录您的账户。
2. 导航到 Amazon Fraud Detec tor，在左侧导航栏中选择变量，然后选择创建。
3. 在新变量页面中，email_address作为变量名称输入。（可选）输入变量的描述。
4. 在变量类型中，选择电子邮件地址。
5. Amazon Fraud Detector 会自动为该变量类型选择数据类型，因为此变量类型是预定义的。如果您的变量没有被自动分配变量类型，请从列表中选择一个变量类型。有关更多信息，请参阅变量类型：
6. 如果要为变量提供默认值，请选择定义自定义默认值并为变量输入默认值。如果您正在遵循此示例，请跳过此步骤。
7. 选择创建。

8. 在 email_address 概述页面中，确认您刚刚创建的变量的详细信息。

   如果您需要更新，请选择编辑并提供更新。选择保存更改。

9. 重复该过程创建另一个变量ip_address，然后为变量类型选择 IP 地址。

10. 变量页面显示新创建的变量。

> **Important**
>
> 我们建议您从数据集中创建任意数量的变量。您可以稍后在创建事件类型时决定要包含哪些变量来训练模型以检测欺诈和生成欺诈检测。

使用创建变量 AWS SDK for Python (Boto3)

以下示例显示了对 CreateVariableAPI 的请求。该示例创建了两个变量email_address和ip_address，并将它们分配给相应的变量类型（EMAIL_ADDRESS和IP_ADDRESS）。

这些变量用作示例。如果您要创建用于模型训练的变量，请使用数据集中适合您的用例的变量。在创建变量变量丰富之前变量类型，请务必阅读和内容。

一定要指定变量源。它有助于确定变量值的派生位置。如果变量源是 E VEN T，则变量值将作为GetEventPrediction请求的一部分发送。如果变量值为MODEL_SCORE，则由亚马逊欺诈检测器填充。如果EXTERNAL_MODEL_SCORE，则变量值由导入的SageMaker模型填充。

```
import boto3
fraudDetector = boto3.client('frauddetector')

 #Create variable email_address
   fraudDetector.create_variable(
     name = 'email_address',
     variableType = 'EMAIL_ADDRESS',
     dataSource = 'EVENT',
     dataType = 'STRING',
     defaultValue = '<unknown>'
     )

#Create variable ip_address
   fraudDetector.create_variable(
     name = 'ip_address',
```

```
    variableType = 'IP_ADDRESS',
    dataSource = 'EVENT',
    dataType = 'STRING',
    defaultValue = '<unknown>'
    )
```

删除变量

当您删除变量时，亚马逊欺诈检测器会永久删除该变量，并且数据将不再存储在亚马逊欺诈检测器中。

您无法在 Amazon Fraud Detector 中删除事件类型中包含的变量。你必须先删除与变量关联的事件类型，然后删除变量。

您无法手动删除 Amazon Fraud Detector SageMaker 模型输出变量和模型输出变量。当您删除模型时，Amazon Fraud Detector 会自动删除模型输出变量。

您可以在 Amazon Fraud Detector 控制台中删除变量，使用 delete- variable CLI 命令，使用 DeleteVariableAPI 或使用 AWS SDK for Python (Boto3)

使用控制台删除变量

要删除变量，

1. 登录AWS Management Console并打开亚马逊欺诈检测器控制台，网址为 https://console.aws.amazon.com/frauddetector。
2. 在 Amazon Fraud Detector 控制台的左侧导航窗格中，选择资源，然后选择变量。
3. 选择要删除的变量。
4. 选择 Actions，然后选择 Delete。
5. 输入变量名称，然后选择删除变量。

使用删除变量 AWS SDK for Python (Boto3)

以下代码示例使用 API 删除变量 c ustomer_nam e。DeleteVariable

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.delete_variable (

name = 'customer_name'
```

```
)
```

Labels

标签将事件归类为欺诈事件或合法事件。标签与事件类型相关联，用于在 Amazon Fraud Detector 中训练机器学习模型。如果您计划训练 Online Fraud Insights (OFI) 或交易欺诈洞察 (TFI) 模型，则您的训练数据集中必须至少有 400 个事件归类为欺诈或合法事件。您可以使用任何标签（例如欺诈、合法、1 或 0）对训练数据集中的事件进行分类。训练完成后，经过训练的模型评估事件是否存在欺诈，并使用这些值将事件归类为欺诈事件或合法事件。

您必须首先使用训练数据集中使用的值创建标签，然后将标签与用于构建和训练欺诈检测模型的事件类型相关联。

创建标签

您可以使用 put-label 命令、PutLabelAPI 或使用 Amazon Fraud Detector 控制台创建标签AWS SDK for Python (Boto3)。

使用亚马逊Fraud Detector 控制台创建标签

要创建标签，

1. 打开AWS管理控制台并登录您的账户。
2. 导航到 Amazon Fraud Detector，在左侧导航栏中选择 “标签”，然后选择 “创建”。
3. 在创建标签页面中，输入欺诈事件的标签名称作为标签名称。标签名称必须与代表训练数据集中欺诈活动的标签相对应。或者，根据需要输入的说明。
4. 选择 “创建标签”。
5. 创建第二个标签并为合法事件输入标签名称。确保标签名称对应于代表训练数据集中合法活动的值。

使用创建标签AWS SDK for Python (Boto3)

以下AWS SDK for Python (Boto3)示例代码使用 PutLabelAPI 创建了两个标签（欺诈、合法）。创建标签后，您可以将其添加到事件类型以对特定事件进行分类。

```
import boto3
```

```
fraudDetector = boto3.client('frauddetector')

fraudDetector.put_label(
name = 'fraud',
description = 'label for fraud events'
)

fraudDetector.put_label(
name = 'legit',
description = 'label for legitimate events'
)
```

更新标签

如果您的事件数据集存储在 Amazon Fraud Detector 中，则可能需要为存储的事件添加或更新标签，例如当您对事件进行离线欺诈调查并想要关闭机器学习反馈循环时。

您可以使用update-event-label命令、使用 UpdateEventLabelAPI 或使用AWS SDK for Python (Boto3)

以下AWS SDK for Python (Boto3)示例代码添加了与使用`UpdateEventLabel` API 注册事件类型相关的标签欺诈。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.update_event_label(
            eventId        = '802454d3-f7d8-482d-97e8-c4b6db9a0428',
            eventTypeName  = 'registration',
            assignedLabel  = 'fraud',
            labelTimestamp = '2020-07-13T23:18:21Z'
)
```

更新存储在 Amazon Fraud Detector 中的事件数据中的事件标签

您可能需要为已存储在 Amazon Fraud Detector 中的事件添加或更新欺诈标签，例如当您对某个事件进行离线欺诈调查并想要关闭机器学习反馈回路时。要更新已存储在 Amazon Fraud Detector 中的事件的标签，请使用`UpdateEventLabel` API 操作。下面介绍一个 UpdateEventLabel API 调用的示例。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.update_event_label(
            eventId        = '802454d3-f7d8-482d-97e8-c4b6db9a0428',
            eventTypeName  = 'sample_registration',
            assignedLabel  = 'fraud',
            labelTimestamp = '2020-07-13T23:18:21Z'
)
```

删除标签

当您删除标签时，Amazon Fraud Detector 会永久删除该标签，数据将不再存储在 Amazon Fraud Detector 中。

无法删除 Amazon Fraud Detector 中事件类型所包含的标签。而且无法删除分配给事件 ID 的标签。必须先删除相关的事件 ID。

您可以使用 delet e-label 命令、使用 DeleteLabelAPI 或使用 Amazon Fraud Detector 控制台删除标签AWS SDK for Python (Boto3)

使用控制台删除标签

要删除标签，请执行以下操作：

1. 登录到AWS Management Console并打开 Amazon Fraud Detector 控制台，网址：https://console.aws.amazon.com/frauddetector。
2. 在 Amazon Fraud Detector 控制台的左侧导航窗格中，选择 “资源”，然后选择 “标签”。
3. 选择要删除的标签。
4. 选择 Actions，然后选择 Delete。
5. 输入标签名称，然后选择 “删除标签”。

使用删除标签AWS SDK for Python (Boto3)

以下AWS SDK for Python (Boto3)示例代码使用 DeleteLabelAPI 删除了合法标签。

```
import boto3
```

```
fraudDetector = boto3.client('frauddetector')

fraudDetector.delete_event_label (
    name = 'legit'
)
```

规则

规则是指告亚马逊欺诈检测器在欺诈预测期间如何解释变量值的条件。规则是检测器逻辑的一部分，它由以下元素组成：

- 变量或列表-变量代表事件数据集中要用于欺诈预测的数据元素。列表是事件数据集中变量的一组输入数据元素。规则中使用的变量必须在评估的事件类型中预定义，并且规则中使用的列表必须与变量类型相关联。有关更多信息，请参阅 Variables 和 列表。
- 表达式-规则中的表达式捕捉了您的业务逻辑。如果您在规则中使用变量，则使用变量、比较运算符（如 >、<、<=、>=. == 和值）构造一个简单的规则表达式。如果您使用的是列表，则规则表达式构造为列表条目和列表名称。in有关更多信息，请参阅规则语言参考：您可以使用and和将多个表达式组合在一起or。所有表达式的计算结果必须为布尔值（真或假），并且长度小于 4,000 个字符。不支持 if-else 类型条件。
- 结果 — 结果是匹配规则时亚马逊欺诈检测器返回的响应。结果表明了欺诈预测的结果。您可以为每种可能的欺诈预测创建结果并将其添加到规则中。有关更多信息，请参阅结果：

探测器必须至少有一个关联规则。一条规则最多可以有 3 个列表，一个检测器最多可以有 30 个列表。您可以在检测器创建过程中创建规则。您还可以创建新规则并将其与现有检测器关联。

规则语言参考

以下部分概述了 Amazon Fraud Detector 中的表达式（即规则编写）功能。

使用变量

您可以使用在评估的事件类型中定义的任何变量作为表达式的一部分。使用美元符号表示变量：

```
$example_variable < 100
```

使用清单

您可以使用与变量类型关联并填充条目的任何列表作为规则表达式的一部分。使用美元符号表示列表条目值：

```
$example_list_variable in @list_name
```

比较、成员资格和身份运营商

亚马逊欺诈检测器包括以下比较运算符：>、>=、<、<=、！=，==，输入，不在

示例如下：

示例：<

```
$variable < 100
```

示例：输入，不在

```
$variable in [5, 10, 25, 100]
```

示例：！=

```
$variable != "US"
```

示例：==

```
$variable == 1000
```

操作员表

操作符	亚马逊欺诈探测器操作员
等于	==
不等于	!=
大于	>
小于	<
大于或等于	>=
小于或等于	<=

操作符	亚马逊欺诈探测器操作员
In	in
And	and
或者	or
非	!

基础数学

可以在表达式中使用基本的数学运算符（例如，+、-、*、/）。一个典型的用例是在评估期间需要合并变量时。

在下面的规则中，我们将变量$variable_1与相加$variable_2，并检查总数是否小于 10。

```
$variable_1 + $variable_2 < 10
```

基本数学表数据

操作符	亚马逊欺诈探测器操作员
再加上	+
减去	-
Multiply	*
Divide	/
模数	%

正则表达式（regex）

你可以使用正则表达式来搜索特定的模式，作为表达式的一部分。如果您想匹配某个变量的特定字符串或数值，这尤其有用。Amazon Fraud Detector 仅在使用正则表达式时支持匹配（例如，根据提供的字符串是否与正则表达式匹配，它会返回 True/False）。亚马逊 Fraud Detector 的正则表达式支持基于

java 中的.matches ()（使用 RE2J 正则表达式库）。互联网上有几个有用的网站可用于测试不同的正则表达式模式。

在下面的第一个示例中，我们首先将变量email转换为小写。然后我们检查模式@gmail.com是否在email变量中。请注意，第二个句点是转义的，因此我们可以明确检查字符串.com。

```
regex_match(".*@gmail\.com", lowercase($email))
```

在第二个示例中，我们检查变量是否phone_number包含国家/地区代码，+1以确定电话号码是否来自美国。加号是经过转义的，这样我们就可以明确检查字符串+1。

```
regex_match(".*\+1", $phone_number)
```

正则表达式表

操作符	亚马逊欺诈检测器示例
匹配任何以开头的字符串	regex_match（“^mystring”，$变量）
精确匹配整个字符串	regex_match（“mystring”，$variable）
匹配除换行符之外的任何字符	regex_match (”。“，$变量）
匹配除'mystring'前面的换行符之外的任意数量的字符	regex_match (”。 *mystring”，$变量）
逃出特殊字符	\

检查缺失值

有时检查该值是否丢失是有益的。在亚马逊欺诈检测器中，这由 null 表示。您可以使用以下语法来执行此操作：

```
$variable != null
```

同样，如果你想检查某个值是否不存在，你可以执行以下操作：

```
$variable == null
```

多种条件

您可以使用and和将多个表达式组合在一起or。当找到单个真值时，Amazon Fraud Detector 会在OR表达式中停止；当找到单个假值AND时，它会在表达式中停止。

在下面的示例中，我们使用条件检查两个and条件。在第一条语句中，我们正在检查变量 1 是否小于 100。在第二个中，我们检查变量 2 是否不是 US。

假设规则使用and，则两个条件都必须为 TRUE，整个条件才会计算为 TRUE。

```
$variable_1 < 100 and $variable_2 != "US"
```

您可以使用圆括号对布尔运算进行分组，如下所示：

```
$variable_1 < 100 and $variable_2 != "US" or ($variable_1 * 100.0 > $variable_3)
```

其他表达式类型

DateTime函数

函数	描述	示例
获取当前日期时间 ()	以 ISO8601 UTC 格式提供规则执行的当前时间。你可以使用 getepochmilseconds (getcurrentdatetime ()) 来执行其他操作	getcurrentdatetime () == “2023-03-28T 18:34:02 Z”
在之前 (DateTime 1, DateTime 2)	如果调用方 DateTime 1 在 2 之前，则返回布尔值（真/假）DateTime	isbefore（getcurrentdatetime ()，“2019-11-30T 01:01 Z”）== “False” isbefore（getcurrentdatetime ()，“2050-11-30T 01:05:01 Z”）== “True”
isafter (DateTime 1, DateTime 2)	如果调用者 DateTime 1 在 2 之后，则返回布尔值（真/假）DateTime	isafter（getcurrentdatetime ()，“2019-11-30T 01:01 Z”）== “True”

函数	描述	示例
		isafter（getcurrentdatetime（），“2050-11-30T 01:05:01 Z”）== “False”
getepoch milliseconds () DateTime	取 a DateTime 并以纪元毫秒为单位返回。DateTime对于对日期执行数学运算很有用	getepochmilliseconds (“2019-11-30T 01:01:01 Z”) == 1575032461

字符串运算符

操作符	示例
将字符串转换为大写	大写（$变量）
将字符串转换为小写	小写（$变量）

其他

操作符	Comment
添加评论	# 我的评论

创建规则

您可以在 Amazon Fraud Detector 控制台中创建规则，使用 creat e-rule 命令，使用 CreateRuleAPI 或使用。AWS SDK for Python (Boto3)

每条规则必须包含一个能够捕捉您的业务逻辑的单一表达式。所有表达式的计算结果必须为布尔值（真或假），并且长度小于 4,000 个字符。不支持 if-else 类型条件。表达式中使用的所有变量都必须在评估的事件类型中预定义。同样，表达式中使用的所有列表都必须是预定义的，与变量类型相关联并用条目填充。

以下示例为现有探测器创建规则`high_riskpayments_detector`。该规则将表达式和结果`verify_customer`与规则相关联。

先决条件

要执行下面提到的步骤，请确保在继续创建规则之前完成以下操作：

- 创建探测器
- 创建结果

如果您要为用例创建检测器、规则和结果，请将示例检测器名称、规则名称、规则表达式和结果名称替换为与用例相关的名称和表达式。

在亚马逊欺诈检测器控制台中创建新规则

1. 打开AWS管理控制台并登录您的账户。导航到亚马逊欺诈检测器。
2. 在左侧导航窗格中，选择探测器，然后选择您为用例创建的检测器，例如 payments_ detector。
3. 在 payments_detec tor 页面中，选择关联规则选项卡，然后选择创建规则。
4. 在新规则页面中，输入以下内容：
   a. 在名称中，输入规则的名称，例如 **high_risk**
   b. 在 “描述-可选” 中，可选择输入规则描述，示例，**This rule captures events with a high ML model score**
   c. 在表达式中，使用表达式快速参考指南为您的用例输入规则表达式。示例 `$sample_fraud_detection_model_insightscore >900`
   d. 在结果中，选择您为用例创建的结果，例如 verify _customer。结果是欺诈预测的结果，如果在评估期间规则匹配，则返回结果。
5. 选择 “保存规则”

您为探测器创建了新规则。这是规则的第 1 版，亚马逊欺诈检测器会自动将其提供给检测器使用。

使用创建规则 AWS SDK for Python (Boto3)

以下示例代码使用 CreateRuleAPI `high_risk` 为现有探测器创建规则`payments_detector`。示例代码还向规则添加了规则表达式和结果`verify_customer`。

先决条件

要使用示例代码，请确保在继续创建规则之前完成以下操作：

- 创建探测器

- 创建结果

如果您要为用例创建检测器、规则和结果，请将示例检测器名称、规则名称、规则表达式和结果名称替换为与用例相关的名称和表达式。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.create_rule(
ruleId = 'high_risk',
detectorId = 'payments_detector',
expression = '$sample_fraud_detection_model_insightscore > 900',
language = 'DETECTORPL',
outcomes = ['verify_customer']
)
```

您已经创建了规则的第 1 版，亚马逊欺诈检测器会自动将其提供给检测器使用。

更新规则

您可以随时通过添加或更新规则描述、更新规则表达式或添加或删除规则结果来更新规则。更新规则时，会创建新的规则版本。

您可以使用update-rule-version命令、使用 UpdateRuleVersionAPI 或使用 AWS SDK 更新亚马逊欺诈检测器控制台中的规则。

更新规则后，请务必更新您的检测器版本以使用新的规则版本。

更新亚马逊欺诈检测器控制台中的规则

要更新规则，

1. 打开AWS管理控制台并登录您的账户。导航到亚马逊欺诈检测器。
2. 在左侧导航窗格中，选择探测器。
3. 在探测器窗格中，选择与要更新的规则关联的检测器。
4. 在检测器页面中，选择关联规则选项卡，然后选择要更新的规则。
5. 在规则页面中，选择操作，然后选择创建版本。
6. 请注意，版本已更改。输入更新的描述、表达式或结果。

7. 选择 “保存新版本”

使用更新规则 AWS SDK for Python (Boto3)

以下示例代码使用 UpdateRuleVersionAPI 将规则的阈值`high_risk`从 900 更新为 950。此规则与探测器相关联`payments_detector`。

```
fraudDetector.update_rule_version(
rule = {
    'detectorId' : 'payments_detector',
    'ruleId' : 'high_risk',
    'ruleVersion' : '1'
},
expression = '$sample_fraud_detection_model_insightscore > 950',
language = 'DETECTORPL',
outcomes = ['verify_customer']
)
```

列表

列表是事件数据集中变量的一组输入数据。您在与探测器相关的规则中使用输入数据。规则是指在欺诈期间指示 Amazon Fraud Detector 如何解释输入数据的条件。例如，您可以创建 IP 地址列表，然后创建规则以在列表中存在特定 IP 地址时拒绝访问。使用列表的规则以 `in$ip_address_value@list_name` 格式表示。

使用 Amazon Fraud Detector，您可以通过添加或删除数据来管理列表，而无需更新相关规则。与您的列表关联的规则会自动合并新添加或删除的数据。

一个列表最多可以包含 100,000 个唯一条目，每个条目的长度可达 320 个字符。默认情况下，您在规则中使用的每个列表都与亚马逊欺诈检测器的变量类型 FREE_FORM_TEXT 相关联。您可以随时为列表分配变量类型。您最多可以在一个规则中使用 3 个列表。

您可以使用 API、或使用AWS SDK 在 Amazon Fraud Detector 控制台中创建列表、向列表中添加条目、删除列表或删除列表中的一个或多个条目，或者为列表分配变量类型。AWS CLI

创建列表

您可以创建包含事件数据集中变量的输入数据（条目）的列表，并在规则表达式中使用该列表。无需更新使用列表的规则即可动态管理列表中的条目。

要创建列表，您必须先指定一个名称，然后选择将该列表与 Amazon Fraud Detector变量类型 支持的列表相关联。默认情况下，亚马逊Fraud Detector 假定该列表为 FREE_FORM_TEXT 变量类型。

您可以使用 API、或使用AWS SDK 在 Amazon Fraud DetectoAWS CLI r 控制台中创建列表。

使用亚马逊Fraud Detector 控制台创建列表

要创建列表

1. 打开AWS管理控制台并登录您的账户。导航至Amazon Fraud Detec
2. 在左侧导航窗格中，选择列表。
3. 在 “列表详情” 下
   a. 在 Li st名称中，输入列表的名称。
   b. 在 Desc ri ption（描述）中，输入说明。
   c. （可选）在变量类型中，为列表选择变量类型。

   > **Important**
   > 如果您的列表包含 IP 地址，请确保选择 IP_ADDRESS 作为变量类型。如果您未选择变量类型，则 Amazon Fraud Detector 会假定该列表为 FREE_FORM_TEX T 变量类型。

4. 在添加列表数据中，添加列表条目，每行一个条目。您也可以从电子表格中复制和粘贴条目。

   > **Note**
   > 确保条目不使用逗号分隔，并且在列表中是唯一的。如果输入两个相同的条目，则只会添加一个。

5. 选择创建。

使用创建列表AWS SDK for Python (Boto3)

您可以通过指定列表名称来创建列表。创建列表时，您可以选择提供描述、关联变量类型或向列表中添加条目。或者，您可以稍后通过添加条目或描述来更新列表。如果您在创建列表时尚未分配变量类型，则可以稍后为该列表分配变量类型。分配列表后，无法更改列表的变量类型。

> **Important**
>
> 如果您的列表包含 IP 地址，请确保将 IP_ADDRESS 指定为变量类型。如果您未分配变量类型，则 Amazon Fraud Detector 会假定该列表为 FREE_FORM_TEX T 变量类型。

以下示例使用 CreateListAPI 操作通过提供描述、变量类型和添加四个列表条目来创建列表。`allow_email_ids`

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.create_list (
     name = 'allow_email_ids',
     description = 'legitimate email_ids'
     variableType = 'EMAIL_ADDRESS',
     elements = ['emailId _1', 'emailId_2', 'emailId_3','emailId_4']
     )
```

在列表中添加条目

创建列表后，您可以随时在列表中添加或追加条目。当您在列表中添加或追加条目时，您无需更新与该列表相关的规则。该规则会自动合并新添加的条目。

您的列表最多可包含 100,000 个唯一条目，每个条目最多可包含 320 个字符。

您可以使用 API、使用或使用AWS SDK 在 Amazon Fraud DetectoAWS CLI r 控制台中添加条目。

使用亚马逊Fraud Detector 控制台在列表中添加条目

要在列表中添加一个或多个条目

1. 打开AWS管理控制台并登录您的账户。导航至Amazon Fraud Detec
2. 在左侧导航窗格中，选择列表。
3. 在列表页面中，选择要向其添加条目的列表。
4. 在列表详细信息页面中，选择列表数据选项卡，然后选择添加数据。
5. 在添加列表数据框中，在每行添加一个条目或从电子表格中复制并粘贴条目。确保不要使用逗号分隔条目。

6. 选择 Add（添加）。

使用在列表中添加条目AWS SDK for Python (Boto3)

以下示例使用 UpdateListAPI 操作在`allow_email_ids`列表中添加两个新条目。确保您要添加的条目在列表中是唯一的。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.update_list (
     name = 'allow_email_ids',
     updateMode = 'APPEND'
     elements = ['emailId_11','emailId_12']
```

为列表分配变量类型

您在规则中使用的每个列表都必须与亚马逊欺诈检测器的变量类型变量类型相关联。默认情况下，亚马逊Fraud Detector 假定该列表为 FREE_FORM_TEXT 变量类型。请务必注意，包含 IP 地址的列表必须与 IP_ADDRESS 变量类型相关联。

您可以在创建列表时或之后的任何时候将列表与变量类型相关联。如果您已经将列表与变量类型关联并希望稍后对其进行更改，则必须创建一个新列表。您无法更改列表的变量类型。

您可以使用 API、使用或使用AWS SDK 在 Amazon Fraud DetectorAWS CLI 控制台中分配变量类型。

使用亚马逊Fraud Detector 控制台为列表分配变量类型

为列表分配变量类型

1. 打开AWS管理控制台并登录您的账户。导航至Amazon Fraud Detec
2. 在左侧导航窗格中，选择列表。
3. 在列表页面中，选择要为其分配变量类型的列表。
4. 在您的列表详细信息页面中，选择操作并选择编辑列表。
5. 在编辑列表框中，为列表选择变量类型。
6. 选择保存。

使用将变量类型分配给列表AWS SDK for Python (Boto3)

以下示例使用 UpdateListAPI 操作为`allow_ip_address`列表分配变量类型。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.update_list (
     name = 'allow_ip_address',
     variableType = 'IP_ADDRESS'
)
```

删除列表

您可以删除任何规则中均未使用的列表。当您删除列表时，Amazon Fraud Detector 会永久删除该列表和列表中的所有条目。

您可以使用AWS CLI或AWS SDK 在 Amazon Fraud Detector 控制台中使用 API 删除列表。

使用亚马逊Fraud Detector 控制台删除列表

要删除列表

1. 打开AWS管理控制台并登录您的账户。导航至Amazon Fraud Detec
2. 在左侧导航窗格中，选择列表
3. 在 Li st s（列表）页面中，选择要删除的列表。
4. 在您的列表详细信息页面中，选择操作并选择删除列表。
5. 选择 “删除列表”。

使用删除列表AWS SDK for Python (Boto3)

以下示例使用 DeleteListAPI 操作进行删除`allow_email_ids`。

```
import boto3
                        fraudDetector = boto3.client('frauddetector')
fraudDetector.delete_list(
   name = 'allow_email_ids'
)
```

从列表中删除条目

您可以随时从列表中删除一个或多个条目。删除列表中的条目时，无需更新列表所关联的规则。该规则会自动合并更新的列表。

您可以使用 API、AWS CLI或AWS SDK 从 Amazon Fraud Detector 控制台的列表中删除条目。

使用亚马逊Fraud Detector 控制台从列表中删除条目

要从列表中删除一个或多个条目

1. 打开AWS管理控制台并登录您的账户。导航至Amazon Fraud Detec
2. 在左侧导航窗格中，选择列表
3. 在 Li st s（列表）页面中，选择包含要删除的条目的列表。
4. 在列表详细信息页面中，选择列表数据选项卡，然后选择要删除的条目。
5. 选择 “删除”，然后再次选择 “删除” 进行确认。

使用从列表中删除条目AWS SDK for Python (Boto3)

在以下示例中，UpdateListAPI 操作从`allow_email_ids`列表中删除条目。

```
import boto3
                        fraudDetector = boto3.client('frauddetector')
fraudDetector.update_list(
   name = 'allow_email_ids',
   updateMode = 'REMOVE',
   elements = ['emailId_4', 'emailId_12']
)
```

从列表中删除所有条目

如果规则中未使用列表，则可以删除列表中的所有条目。您可以删除列表中的所有条目，然后在同一列表中添加条目。

您可以使用 API、AWS CLI或AWS SDK 从 Amazon Fraud Detector 控制台的列表中删除条目。

使用亚马逊Fraud Detector 控制台从列表中删除所有条目

从列表中删除所有条目

1. 打开AWS管理控制台并登录您的账户。导航至Amazon Fraud Detec
2. 在左侧导航窗格中，选择列表
3. 在 Li st s（列表）页面中，选择包含要删除的条目的列表。
4. 在您的列表详细信息页面中，选择列表数据选项卡，然后选择全部删除。
5. 在 “全部删除” 框中，键`delete all`入确认，然后选择 “删除所有列表数据”。

使用删除列表中的所有条目AWS SDK for Python (Boto3)

在以下示例中，UpdateListAPI 操作从`allow_email_ids`列表中删除所有条目。

```
import boto3
                        fraudDetector = boto3.client('frauddetector')
fraudDetector.update_list(
   name = 'allow_email_ids',
   updateMode = 'REPLACE',
   elements = []
)
```

结果

结果是欺诈预测的结果。您可以为每个可能的欺诈预测结果创建结果。例如，您可能希望结果代表风险级别（high_risk、medium_risk 和 low_risk）或操作（批准、审查）。在创建一个结果后，您可以在一个规则中添加一个或多个结果。作为GetEventPrediction响应的一部分，Amazon Fraud Detector 会返回任何匹配规则的定义结果。

创建结果

您可以在 Amazon Fraud Detector 控制台中使用 put- outfuting 命令、使用 PutOutcomeAPI 或使用 AWS SDK for Python (Boto3)。

使用亚马逊Fraud Detector 控制台创建结果

要创建一个或多个结果，

1. 打开AWS管理控制台并登录您的账户。导航到 Amazon FFraud Detector eto
2. 在左侧导航窗格中，选择结果。
3. 在 “结果” 页面中，选择 “创建”。
4. 在 Ne w 结果页面中，输入以下内容：
   a. 在结果名称中，输入结果的名称。
   b. 在结果描述中，您可以在其中输入描述。
5. 选择 “保存结果”。
6. 重复步骤 2 到 5 以创建其他结果。

使用创建结果AWS SDK for Python (Boto3)

以下示例使用PutOutcome API 创建三个结果。它们是verify_customerreview、和approve。创建结果后，您可以将其分配给规则。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.put_outcome(
name = 'verify_customer',
description = 'this outcome initiates a verification workflow'
)

fraudDetector.put_outcome(
name = 'review',
description = 'this outcome sidelines event for review'
)

fraudDetector.put_outcome(
name = 'approve',
description = 'this outcome approves the event'
)
```

删除结果

无法删除规则版本中使用的结果。

当您删除结果时，Amazon Fraud Detector 会永久删除该结果，并且该数据将不再存储在 Amazon Fraud Detector 中。

您可以在 Amazon Fraud Detec tor 控制台中使用删除结果命令、使用 DeleteOutcomeAPI 或使用AWS SDK for Python (Boto3)

在亚马逊Fraud Detector 控制台中删除结果

删除结果

1. 登录到AWS Management Console并打开 Amazon Fraud Detector 控制台，网址：https://console.aws.amazon.com/frauddetector。
2. 在 Amazon Fraud Detector 控制台的左侧导航窗格中，选择 “资源”，然后选择 “结果”。
3. 选择要删除的结果。
4. 选择 Actions，然后选择 Delete。
5. 输入结果名称，然后选择删除结果。

使用删除结果AWS SDK for Python (Boto3)

以下示例使用 DeleteOutcomeAPI 删除`verify_customer`结果。删除结果后，您无法再将其分配给规则。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.delete_outcome(
name = 'verify_customer'
)
```

实体

实体代表正在执行该事件的人或事物。实体类型对实体进行分类。示例分类包括客户、卖家、用户或账户。作为事件数据集的一部分，您提供实体类型 (ENTITY_TYPE) 和实体标识符 (ENTITY_ID)，以指示执行该事件的特定实体。

Amazon Fraud Detector 在为事件生成欺诈预测时使用实体类型来指明谁实施了该事件。您要在欺诈预测中使用的实体类型必须首先在 Amazon Fraud Detector 中创建，然后在创建事件类型时将其添加到事件中。

创建实体类型

您可以使用put-entity-type命令、使用 PutEntityTypeAPI 或使用 Amazon Fraud Detector 控制台创建实体类型AWS SDK for Python (Boto3)。使用适用于 Python (Boto3) 的 SDK fr Python (Boto3) 生成实体类型`customer`。如果您要创建与用于训练欺诈检测模型的事件类型关联的实体类型，请使用事件数据集中适合您的用例的实体类型。

使用亚马逊Fraud Detector 控制台创建实体类型

要创建实体类型，

1. 打开AWS管理控制台并登录您的账户。
2. 导航到 Amazon Fraud Detector，在左侧导航栏中选择实体，然后选择创建。
3. 在创建实体页面中，输入 cu st omer 作为实体类型名称。或者，选择输入实体的描述。
4. 选择 Create entity (创建实体)。

使用创建实体类型AWS SDK for Python (Boto3)

以下AWS SDK for Python (Boto3)代码示例使用`PutEntityType` API 创建实体类型`customer`。如果您要创建与用于训练欺诈检测模型的事件类型关联的实体类型，请使用事件数据集中适合您的用例的实体。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.put_entity_type(
name = 'customer',
description = 'customer'
)
```

删除实体类型

在亚马逊Fraud Detector 中，无法删除事件类型中包含的实体类型。您必须先删除与该实体关联的事件类型，然后删除该实体类型。

当您删除实体类型时，Amazon Fraud Detector 会永久删除该实体类型，数据将不再存储在 Amazon Fraud Detector 中。

可以在 Amazon Fraud Detector 控制台中使用delete-entity-type命令、使用 DeleteEntityTypeAPI 或使用AWS SDK for Python (Boto3)

在亚马逊Fraud Detector 控制台中删除实体类型

要删除实体类型，

1. 登录到AWS Management Console并打开 Amazon Frazon Frazon FFraud Detector zon Frazon Fra zon F https://console.aws.amazon.com/frauddetector
2. 在 Amazon Fraud Detector 控制台的左侧导航窗格中，选择 “资源”，然后选择 “实体”。
3. 选择要删除的实体类型。
4. 选择 Actions，然后选择 Delete。
5. 输入实体类型名称，然后选择删除实体类型。

使用删除实体类型AWS SDK for Python (Boto3)

以下AWS SDK for Python (Boto3)示例代码使用 DeleteEntityTypeAPI 删除了实体类型客户。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.delete_entity_type (

name = 'customer'

)
```

使用管理亚马逊Fraud Detector 资源AWS CloudFormation

Azon Fetector 集成AWS CloudFormation，后者是一项服务，可帮助您对 Amazon FetFraud Detector 资源进行建模和设置，这样您只需花较少的时间来创建和管理资源与基础设施。您可以创建一个描述所需的全部 Amazon FetFraud Detector ector 资源的模板（例如探ector），然后将为您预置和配置这些资源。 EntityType EventTypeAWS CloudFormation您可以重复使用该模板，在多个 AWS 账户和区域中以一致的方式重复预置和配置资源。

使用 AWS 不收取任何额外费用 CloudFormation。

创建on on etFraud Detector 模板

要为 Amazon FetFraud Detector r 预置和配置资源，您必须了解AWS CloudFormation模板。模板是 JSON 或 YAML 格式的文本文件。这些模板描述要在 AWS CloudFormation 堆栈中调配的资源。如果您不熟悉 JSON 或 YAML，可以在 AWS CloudFormation Designer 的帮助下开始使用 AWS CloudFormation 模板。有关更多信息，请参阅 AWS CloudFormation 用户指南中的什么是 AWS CloudFormation Designer？。

您还可以使用AWS CloudFormation模板创建、更新和删除您的 Amazon FFraud Detector tor 资源。有关更多信息（包括您的资源的 JSON 和 YML 模板示例），请参阅AWS CloudFormation用户指南中的Amazon FetFraud Detector tor 资源类型参考。

如果您已经在使用 CloudFormation，则无需管理其他 IAM 策略或 CloudTrail 日志记录。

管理on on etFraud Detector or 堆栈

您可以通过 CloudFormation 控制台或 AWS CLI 创建、更新和删除您的 Amazon Fraud Detector 堆栈。

要创建堆栈，您必须拥有描述 AWS CloudFormation 将在堆栈中包含哪些资源的模板。您还可以通过将已经创建的 Amazon Fraud Detector 资源导入到新的或现有的堆栈中，将它们导入到 CloudFormation 管理中。

有关管理堆栈的详细说明，请参阅AWS CloudFormation用户指南以了解如何创建、更新和删除堆栈。

整理您的 on on etFraud Detector or 堆栈

你整理AWS CloudFormation堆栈的方式完全取决于你。通常，最佳做法是按生命周期和所有权组织堆栈。这意味着按资源的变化频率或负责更新的团队对资源进行分组。

您可以选择通过为每个探测器及其检测逻辑（例如，规则、变量等）创建堆栈来组织堆栈。如果您使用其他服务，则应考虑是否要将 Amazon Fraud Detector 资源与其他服务的资源堆叠在一起。例如，您可以创建一个包含帮助收集数据的 Kinesis 资源和处理数据的 Amazon Fraud Detector 资源的堆栈。这可能是确保欺诈团队的所有产品协同作用的有效方法。

了解 on on etFraud Detector CloudFormation r 参数

除了所有 CloudFormation 模板中可用的标准参数外，Amazon Fraud Detector 还引入了另外两个参数来帮助您管理部署行为。如果您不包括这两个参数中的一个或两个参数，则 CloudFormation 将使用如下所示的默认值。

参数	值	默认值
DetectorVersionStatus	活动：将新的/更新的探测器版本设置为活动状态 草稿：将新的/更新的探测器版本设置为草稿状态	草案
内联	TRU@@ E： CloudFormation 允许在创建/更新/删除堆栈时创建/更新/删除资源。 FALSE： CloudFormation 允许验证对象是否存在，但不允许对该对象进行任何更改。	TRUE

AzFAWS CloudFormation raud Detector on

以下是用于管理探测器和关联的探测器版本的示例AWS CloudFormation YAML 模板。

```
# Simple Detector resource containing inline Rule, EventType, Variable, EntityType and
 Label resource definitions
Resources:
  TestDetectorLogicalId:
    Type: AWS::FraudDetector::Detector
    Properties:
      DetectorId: "sample_cfn_created_detector"
      DetectorVersionStatus: "DRAFT"
      Description: "A detector defined and created in a CloudFormation stack!"

      Rules:
        - RuleId: "over_threshold_investigate"
          Description: "Automatically sends transactions of $10000 or more to an
 investigation queue"
          DetectorId: "sample_cfn_created_detector"
          Expression: "$amount >= 10000"
```

```
          Language: "DETECTORPL"
          Outcomes:
            - Name: "investigate"
              Inline: true
        - RuleId: "under_threshold_approve"
          Description: "Automatically approves transactions of less than $10000"
          DetectorId: "sample_cfn_created_detector"
          Expression: "$amount <10000"
          Language: "DETECTORPL"
          Outcomes:
            - Name: "approve"
              Inline: true
      EventType:
        Inline: "true"
        Name: "online_transaction"
        EventVariables:
          - Name: "amount"
            DataSource: 'EVENT'
            DataType: 'FLOAT'
            DefaultValue: '0'
            VariableType: "PRICE"
            Inline: 'true'
        EntityTypes:
          - Name: "customer"
            Inline: 'true'
        Labels:
          - Name: "legitimate"
            Inline: 'true'
          - Name: "fraudulent"
            Inline: 'true'
```

了解有关 AWS CloudFormation 的更多信息

要了解有关 AWS CloudFormation 的更多信息，请参阅以下资源：

- AWS CloudFormation
- AWS CloudFormation 用户指南
- AWS CloudFormation API 参考
- AWS CloudFormation 命令行界面用户指南

欺诈预测

您可以使用 Amazon Fraud Detector 实时获取单个事件的欺诈预测，也可以离线获取一组事件的欺诈预测。要为单个事件或一组事件生成欺诈预测，您需要向亚马逊Fraud Detector 提供以下信息：

- 欺诈预测逻辑
- 事件元数据

欺诈检测逻辑

欺诈预测逻辑使用一个或多个规则来评估与事件相关的数据，然后提供结果和欺诈预测分数。您可以使用以下组件创建欺诈预测逻辑：

- 事件类型-定义事件的结构
- 模型-定义预测欺诈的算法和数据要求
- 变量-表示与事件相关的数据元素
- 规则-在Fraud Detector 指在欺诈预测期间如何解释变量值
- 结果-欺诈预测生成的结果
- 探测器版本-包含特定事件的欺诈预测逻辑

有关用于创建欺诈检测逻辑的组件的更多信息，请参阅亚马逊Fraud Detector 概念。在开始生成欺诈预测之前，请确保您已经创建并发布了包含您的欺诈预测逻辑的检测器版本。您可以使用欺诈检测器控制台或 API 创建和发布检测器版本。有关如何使用控制台的说明，请参阅入门（控制台）。有关使用 API 的说明，请参阅创建检测器版本。

事件元数据

事件元数据提供正在评估的事件的详细信息。您要评估的每个事件都必须包含与您的探测器版本相关的事件类型中每个变量的值。此外，您的事件元数据必须包括以下参数：

- EVENT_ID — 事件的标识符。例如，如果您的事件是在线交易，则 EVENT_ID 可能是提供给客户的交易参考号。

  有关 EVENT_ID 的重要注意事项

  - 该事件必须是唯一的

- 应代表对您的业务有意义的信息
- 必须满足正则表达式模式：`^[0-9a-z_-]+$.`
- 必须保存。EVENT_ID 是事件的引用，用于对事件执行操作，例如删除事件。
- 不建议将时间戳附加到 EVENT_ID，因为这可能会导致稍后要更新事件时出现问题，因为您需要提供完全相同的 EVENT_ID。

- ENTITY_TYPE — 执行事件的实体，例如商家或客户。
- ENTITY_ID-执行事件的实体的标识符。ENTITY_ID 必须满足以下正则表达式模式：`^[0-9a-z_-]+$`。如果在评估时 ENTITY_ID 不可用，则传递未知字符串。
- EVENT_TIMESTAMP-事件发生时的时间戳。该时间戳必须采用 ISO 8601 标准（世界标准时间）。

实时预测

您可以通过调用`GetEventPrediction` API 实时评估在线活动是否存在欺诈行为。您在每个请求中提供有关单个事件的信息，并根据与指定检测器相关的欺诈预测逻辑同步接收模型分数和结果。

实时欺诈预测的工作原理

`GetEventPrediction`API 使用指定的探测器版本来评估为该事件提供的事件元数据。在评估期间，Amazon Fraud Detector 首先为添加到检测器版本的模型生成模型分数，然后将结果传递给规则进行评估。规则按照规则执行模式的指定执行（请参阅创建检测器版本）。作为响应的一部分，Amazon Fraud Detector 提供模型分数以及与匹配规则相关的任何结果。

获得实时欺诈预测

要获得实时欺诈预测，请确保您创建并发布了包含您的欺诈预测模型和规则的检测器，或者仅包含规则集。

您可以通过使用AWS命令行接口 (AWSCLI) 或 Amazon Fraud Detector 软件开发工具包调用 GetEventPredictionAPI 操作来实时预测事件。

要使用 API，请在每个请求中提供单个事件的信息。作为请求的一部分，您必须指`detectorId`定 Amazon Fraud Detector 将用于评估事件。您也可以指在其中指示`detectorVersionId`。如果未指定，则亚马逊Fraud Detect`detectorVersionId` o`ACTIVE` r 将使用检测器的版本。

您可以选择通过在字段中传递数据来发送数据以调用 SageMaker 模型`externalModelEndpointBlobs`。

使用以下方法进行欺诈预测AWS SDK for Python (Boto3)

要生成欺诈预测，请调用GetEventPrediction API。以下示例假设您已完成B 部分：生成欺诈预测。作为回复的一部分，您将收到模型分数以及任何匹配的规则和相应的结果。您可以在aws-fraud-detector-samples GitHub 存储库中找到其他GetEventPrediction请求示例。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.get_event_prediction(
detectorId = 'sample_detector',
eventId = '802454d3-f7d8-482d-97e8-c4b6db9a0428',
eventTypeName = 'sample_registration',
eventTimestamp = '2020-07-13T23:18:21Z',
entities = [{'entityType':'sample_customer', 'entityId':'12345'}],
eventVariables = {
    'email_address' : 'johndoe@exampledomain.com',
    'ip_address' : '1.2.3.4'
}
)
```

批量预测

您可以在 Amazon Fraud Detector 中使用批量预测任务来获得对一组不需要实时评分的事件的预测。例如，您可以创建批量预测任务以执行离线操作proof-of-concept，或者以每小时、每天或每周为单位回顾性评估事件风险。

您可以使用亚马逊Fraud Detector 控制台创建批量预测任务，也可以使用AWS命令行接口 (AWSCLI) 或亚马逊Fraud Detector SDK 调用 CreateBatchPredictionJobAPI 操作。

主题

- 批量预测的工作原理
- 输入与输出文件
- 获取批量预测
- 有关 IAM 角色的指南
- 使用获取批量欺诈预测 AWS SDK for Python (Boto3)

批量预测的工作原理

CreateBatchPredictionJobAPI 操作使用指定的检测器版本根据位于 Amazon S3 存储桶中的输入 CSV 文件中提供的数据进行预测。然后 API 将生成生成的 CSV 文件

Batch 预测作业以与GetEventPrediction操作相同的方式计算模型分数和预测结果。类似于GetEventPrediction，要创建批量预测作业，您首先创建事件类型，可选择训练模型，然后创建用于评估批处理作业中事件的检测器版本。

批量预测作业评估的事件风险评分的定价与 GetEventPrediction API 创建的分数的定价相同。有关详细信息，请参阅亚马逊Fraud Detector 定价。

您一次只能运行一次。

输入与输出文件

输入 CSV 文件应包含与所选探测器版本关联的事件类型匹配的标头。输入数据文件的最大大小为 1GB。活动数量将因您的活动规模而异。

除非您为输出数据指定单独的位置，否则 Amazon Fraud Detector 会在与输入文件相同的存储桶中创建输出文件。输出文件包含来自输入文件的原始数据和以下附加列：

- MODEL_SCORES— 详细说明与所选探测器版本相关的每个模型中事件的模型分数。
- OUTCOMES— 详细说明所选探测器版本及其规则评估的事件结果。
- STATUS— 表示事件是否已成功评估。如果未成功评估事件，则此列显示失败的原因代码。
- RULE_RESULTS— 基于规则执行模式的所有匹配规则的列表。

获取批量预测

以下步骤假设您已经创建了事件类型，使用该事件类型（可选）训练了模型，并为该事件类型创建了探测器版本。

要想的权限，请参阅。

1. 登录AWS Management Console并打开 Amazon Fraad Detector 控制台（https://console.aws.amazon.com/frauddetector）。
2. 在 Amazon Fraud Detector 控制台的左侧导航窗格中，选择Batc h 预测，然后选择新建批量预测。

3. 在 Job 名称中，为您的批量预测作业指定一个名称。如果您未指定姓名，亚马逊Fraud Detector会随机生成一个任务名称。
4. 在检测器中，选择用于此批量预测的检测器。
5. 在探测器版本中，为该批量预测选择探测器版本。您可以在任何状态下选择探测器版本。如果您的探测器处于检测器版本`Active`状态，则会自动选择该版本，但您也可以根据需要更改此选择。
6. 在 IAM 角色中，选择或创建一个对您的输入和输出 Amazon S3 存储桶具有读写权限的角色。参阅 有关 IAM 角色的指南 了解更多信息。

   要获得批量预测，调用该`CreateBatchPredictionJob`操作的 IAM 角色必须拥有对您的输入 S3 存储桶的读取权限和对输出 S3 存储桶的写入权限。有关存储桶权限的更多信息，请参阅 Amazon S3 用户指南中的用户策略示例。
7. 在输入数据位置中，指定输入数据的 Amazon S3 位置。如果您想将输出文件放在不同的 S3 存储桶中，请选择单独的数据位置进行输出，并提供输出数据的 Amazon S3 位置。
8. （可选）为您的批量预测作业创建标签。
9. 选择开始。

   Amazon Fraud Detector 会创建批量预测任务，该任务的状态为`In progress`。Batch 预测作业的处理时间因事件数量和探测器版本配置而异。

要停止正在进行的批量预测作业，请转到批量预测作业详细信息页面，选择操作，然后选择停止批量预测。如果您停止批量预测作业，则不会收到该作业的任何结果。

当批量预测任务的状态更改为时`Complete`，您可以从指定的输出 Amazon S3 存储桶中检索任务的输出。输出文件的名称采用格式`batch prediction job name_file creation timestamp_output.csv`。例如，名为的作业的输出文件`mybatchjob`是`mybatchjob_1611170650_output.csv`。

要搜索由批量预测作业评估的特定事件，请在 Amazon Fraud Detector 控制台的左侧导航窗格中选择搜索过去的预测。

要删除已完成的批量预测作业，请转到批量预测作业详细信息页面，选择操作，然后选择删除批量预测。

有关 IAM 角色的指南

要获得批量预测，调用该CreateBatchPredictionJob操作的 IAM 角色必须拥有对您的输入 S3 存储桶的读取权限和对输出 S3 存储桶的写入权限。有关存储桶权限的更多信息，请参阅Amazon S3 用户指

南中的用户策略示使用的用户策略示使用的用户策略示使用的用户策略示使用的用户策略 在 Amazon Fraud Detector 控制台上，您可以通过三个选项为Batch 预测选择 IAM 角色：

1. 在创建新的Batch 预测作业时创建角色。
2. 选择您之前在亚马逊Fraud Detector 控制台中创建的现有 IAM 角色。在执行此步骤之前，请务必为角色添加`S3:PutObject`权限。
3. 为先前创建的 IAM 角色输入自定义 ARN。

如果收到与 IAM 角色相关的错误，请验证以下几点：

1. 您的 Amazon S3 输入与输出存储桶位于同一区域。
2. 您使用的 IAM 角色拥有您的输入 S3 存储桶的权限和输出 S3 存储桶的`s3:PutObject`权限。`s3:GetObject`
3. 您正在使用的 IAM 角色具有服务主体的信任策略`frauddetector.amazonaws.com`。

使用获取批量欺诈预测 AWS SDK for Python (Boto3)

以下示例演示使用的示例演示使用的示例演示使用的示例CreateBatchPredictionJob请求。批量预测任务必须包含以下现有资源：检测器、探测器版本和事件类型名称。以下示例假设您已经创建了事件类型`sample_registrationsample_detector`、检测器和探测器版本`1`。

```
import boto3
fraudDetector = boto3.client('frauddetector')

fraudDetector.create_batch_prediction_job (
   jobId = 'sample_batch',
   inputPath = 's3://bucket_name/input_file_name.csv',
   outputPath = 's3://bucket_name/',
   eventTypeName = 'sample_registration',
   detectorName = 'sample_detector',
   detectorVersion = '1',
   iamRoleArn = 'arn:aws:iam::**:role/service-role/AmazonFraudDetector-DataAccessRole-
**'
)
```

预测解释

预测解释可以深入了解每个事件变量如何影响模型的欺诈预测分数，并作为欺诈预测的一部分自动生成。每项欺诈预测的风险评分都介于 1 到 1000 之间。预测解释详细介绍了每个事件变量对风险评分的影响，包括幅度（0-5，5 表示最高）和方向（驾驶分数提高或降低）。您也可以对以下任务使用预测解释：

- 在将事件标记为待审核时，在手动调查期间确定最重要的风险指标。
- 缩小导致误报预测的根本原因（例如，合法事件的高风险评分）。
- 分析事件数据的欺诈模式并检测数据集中的偏差（如果有）。

> **Important**
>
> 预测解释是自动生成的，仅适用于在 2021 年 6 月 30 日当天或之后训练的模型。要获得在 2021 年 6 月 30 日之前训练的模型的预测说明，请重新训练这些模型。

预测解释为用于训练模型的每个事件变量提供以下一组值。

相对影响

提供变量对欺诈预测分数的影响的可视参考。相对影响值包括欺诈风险的星级（0-5，5 表示最高）和方向（增加/减少）影响。

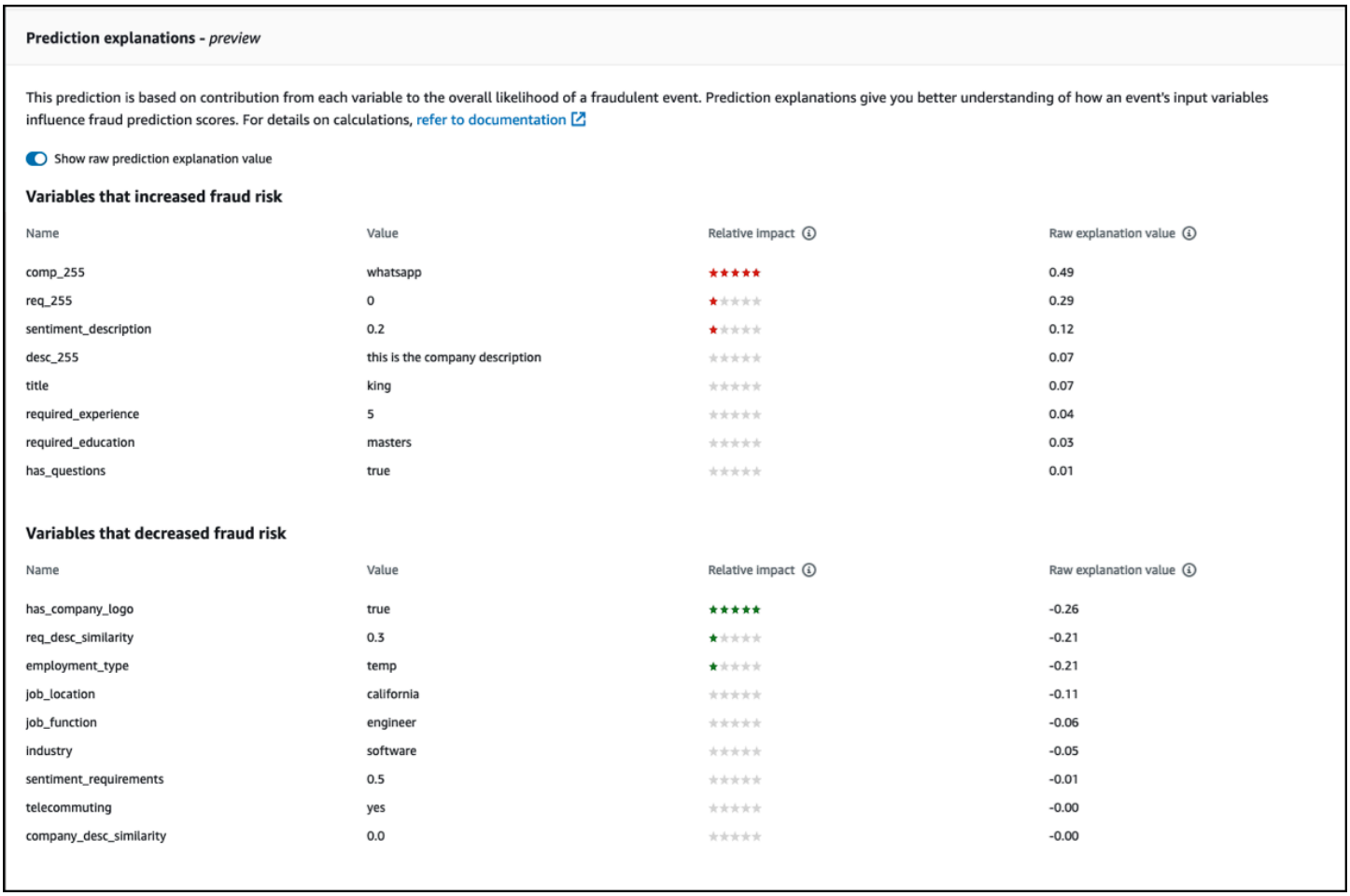
- 增加欺诈风险的变量用红色星星表示。红色星星的数量越多，变量越能提高欺诈分数并增加欺诈的可能性。
- 降低欺诈风险的变量由绿色星星表示。绿色开局次数越多，变量就越能降低欺诈风险评分，欺诈的可能性就会降低。
- 所有变量的零星表示这些变量本身都没有显著改变欺诈风险。

原始解释值

提供未解释的原始值，表示为欺诈的对数赔率。这些值通常介于 -10 到 +10 之间，但范围从-无穷大到 + 无穷大。

- 正值表示该变量推动了风险评分的提高。
- 负值表示该变量降低了风险评分。

在 Amazon Fraud Detector 控制台中，预测解释值显示如下。彩色星评级和相应的原始数值可以很容易地看到变量之间的相对影响。



Prediction explanations - *preview*

This prediction is based on contribution from each variable to the overall likelihood of a fraudulent event. Prediction explanations give you better understanding of how an event's input variables influence fraud prediction scores. For details on calculations, refer to documentation

Show raw prediction explanation value

Variables that increased fraud risk

Name	Value	Relative impact	Raw explanation value
comp_255	whatsapp	★★★★★	0.49
req_255	0	★☆☆☆☆	0.29
sentiment_description	0.2	★☆☆☆☆	0.12
desc_255	this is the company description	☆☆☆☆☆	0.07
title	king	☆☆☆☆☆	0.07
required_experience	5	☆☆☆☆☆	0.04
required_education	masters	☆☆☆☆☆	0.03
has_questions	true	☆☆☆☆☆	0.01

Variables that decreased fraud risk

Name	Value	Relative impact	Raw explanation value
has_company_logo	true	★★★★★	-0.26
req_desc_similarity	0.3	★☆☆☆☆	-0.21
employment_type	temp	★☆☆☆☆	-0.21
job_location	california	☆☆☆☆☆	-0.11
job_function	engineer	☆☆☆☆☆	-0.06
industry	software	☆☆☆☆☆	-0.05
sentiment_requirements	0.5	☆☆☆☆☆	-0.01
telecommuting	yes	☆☆☆☆☆	-0.00
company_desc_similarity	0.0	☆☆☆☆☆	-0.00

查看预测解释

生成欺诈预测后，您可以在 Amazon Fraud Detector 控制台中查看预测说明。要使用 AWS SDK 中的 API 查看预测解释，必须先调用 `ListEventPrediction` API 以获取事件的预测时间戳，然后调用 `GetEventPredictionMetadata` API 以获取预测解释。

使用 Amazon Fraud Detector 控制台查看预测解释

要使用控制台查看预测说明，

1. 打开AWS控制台并登录您的账户。导航到亚马逊 Fraud Detector。
2. 在左侧导航窗格中，选择 “搜索过去的预测”。
3. 使用 “属性”、“运算符” 和 “值” 筛选器来选择要查看的预测。
4. 在顶部的筛选器窗格中，确保选择生成要查看的预测的时间段。

5. 结果窗格显示在指定时间段内生成的所有预测的列表。单击预测的事件 ID 以查看预测解释。
6. 向下滚动到 “预测解释” 窗格。
7. 将显示原始预测解释值按钮设置为打开以查看所有变量的原始预测解释值。

使用适用于 Python 的 AWS 开发工具包 (Boto3) 查看预测解释

以下示例显示了使用`ListEventPredictions`和 AWS SDK 中的`GetEventPredictionMetadata` API 查看预测解释的示例请求。

示例 1：使用 **`ListEventPredictions`** API 获取最新预测列表

```
import boto3
fraudDetector = boto3.client('frauddetector')
fraudDetector.list_event_predictions(
  maxResults = 10,
  predictionTimeRange = {
     end_time: '2022-01-13T23:18:21Z',
     start_time: '2022-01-13T20:18:21Z'
    }
 )
```

示例 2；使用 **`ListEventPredictions`** API 获取过去对事件类型 “注册” 的预测列表

```
import boto3
fraudDetector = boto3.client('frauddetector')
fraudDetector.list_event_predictions(
   eventType = {
      value = 'registration'
    }
   maxResults = 70,
   nextToken = "10",
   predictionTimeRange = {
     end_time: '2021-07-13T23:18:21Z',
     start_time: '2021-07-13T20:18:21Z'
    }
 )
```

示例 3：使用 **`GetEventPredictionMetadata`** API 获取在指定时间段内生成的指定事件 ID、事件类型、检测器 ID 和检测器版本 ID 的过去预测的详细信息。

为该请求predictionTimestamp指定的值是通过首先调用 ListEventPredictions API 获得的。

```
import boto3
fraudDetector = boto3.client('frauddetector')
fraudDetector.get_event_prediction_metadata (
   detectorId = 'sample_detector',
   detectorVersionId = '1',
   eventId = '802454d3-f7d8-482d-97e8-c4b6db9a0428',
   eventTypeName = 'sample_registration',
   predictionTimestamp = '2021-07-13T21:18:21Z'
 )
```

了解预测解释是如何计算的

Amazon Frau d Detector 使用 SHAP（ShapeLey 加法解释）通过计算用于模型训练的每个事件变量的原始解释值来解释单个事件的预测。原始解释值由模型在生成预测时作为分类算法的一部分进行计算。这些原始解释值表示每个输入对欺诈几率对数的贡献。使用映射将原始解释值（从-infinity到+infinity）转换为相对影响值（-5 到 +5）。从原始解释值中得出的相对影响值表示欺诈（正面）或合法（负）几率增加的次数，因此更容易理解预测解释。

Amazon Fraud Detector 中的安全

云安全 AWS 是重中之重。作为 AWS 客户，您可以受益于专为满足大多数安全敏感型组织的要求而构建的数据中心和网络架构。

安全是双方共同承担 AWS 的责任。责任共担模式将其描述为云的安全性和云中的安全性：

- 云安全 — AWS 负责保护在 AWS 云中运行 AWS 服务的基础架构。 AWS 还为您提供可以安全使用的服务。作为AWS 合规计划合规计划合规计划合的一部分，第三方审计师定期测试和验证我们安全的有效性。要了解适用于 Amazon Fraud Detector 的合规计划，请参阅按合规计划划分的 AWS 范围内的服务 服务。
- 云端安全-您的责任由您使用的 AWS 服务决定。您还需要对其他因素负责，包括您的数据的敏感性、您公司的要求以及适用的法律法规。

本文档可帮助您了解在使用 Amazon Fraud Detector 时如何应用分担责任模型。以下主题向您展示了如何配置 Amazon Fraud Detector 以实现您的安全和合规目标。您还将学习如何使用其他 AWS 服务来帮助您监控和保护您的 Amazon Fraud Detector 资源。

主题

- Amazon Fraud Detector 中的数据保护
- Amazon Fraud Detector 的身份和访问管理
- 在 Amazon Fraud Detector 中记录和监控
- Amazon Fraud Detector 的合规性验证
- Amazon Fraud Detector 中的弹性
- Amazon Fraud Detector 中的基础设施安全

Amazon Fraud Detector 中的数据保护

分担责任模型 AWS 分担责任模型适用于 Amazon Fraud Detector 中的数据保护。如本模型所述 AWS ，负责保护运行所有内容的全球基础架构 AWS Cloud。您负责维护对托管在此基础设施上的内容的控制。您还负责您所使用的 AWS 服务 的安全配置和管理任务。有关数据隐私的更多信息，请参阅数据隐私常见问题。有关欧洲数据保护的信息，请参阅 AWS 安全性博客 上的 AWS 责任共担模式和 GDPR 博客文章。

出于数据保护目的，我们建议您保护 AWS 账户 凭证并使用 AWS IAM Identity Center 或 AWS Identity and Access Management (IAM) 设置个人用户。这样，每个用户只获得履行其工作职责所需的权限。我们还建议您通过以下方式保护数据：

- 对每个账户使用多重身份验证 (MFA)。
- 使用 SSL/TLS 与资源通信。 AWS 我们要求使用 TLS 1.2，建议使用 TLS 1.3。
- 使用设置 API 和用户活动日志 AWS CloudTrail。
- 使用 AWS 加密解决方案以及其中的所有默认安全控件 AWS 服务。
- 使用高级托管安全服务（例如 Amazon Macie），它有助于发现和保护存储在 Amazon S3 中的敏感数据。
- 如果您在 AWS 通过命令行界面或 API 进行访问时需要经过 FIPS 140-2 验证的加密模块，请使用 FIPS 端点。有关可用的 FIPS 端点的更多信息，请参阅《美国联邦信息处理标准（FIPS）第 140-2 版》。

我们强烈建议您切勿将机密信息或敏感信息（如您客户的电子邮件地址）放入标签或自由格式文本字段（如名称字段）。这包括您使用控制台、API 或软件开发工具包 AWS 服务 使用 Amazon Fraud AWS Detector AWS CLI或其他软件开发工具包的情况。在用于名称的标签或自由格式文本字段中输入的任何数据都可能会用于计费或诊断日志。如果您向外部服务器提供网址，强烈建议您不要在网址中包含凭证信息来验证对该服务器的请求。

加密静态数据

Amazon Fraud Detector 使用您选择的加密密钥对您的静态数据进行加密。您可以选择以下任一种密钥：

- AWS 拥有的 KMS 密钥。如果您未指定加密密钥，则默认使用此密钥对您的数据进行加密。
- 客户管理的 KMS 密钥。您可以使用密钥策略控制对客户托管的 KMS 密钥的访问权限。有关创建和管理客户托管 KMS 密钥的信息，请参阅密钥管理。

加密传输中数据

Amazon Fraud Detector 会将数据从您的账户中复制出来，并在内部 AWS 系统中进行处理。默认情况下，Amazon Fraud Detector 使用带有 AWS 证书的 TLS 1.2 来加密传输中的数据。

密钥管理

Amazon Fraud Detector 使用以下两种密钥之一对您的数据进行加密：

- AWS 拥有的 KMS 密钥。这是默认模式。
- 客户管理的 KMS 密钥。

创建客户托管的 KMS 密钥

您可以使用 KMS 控制台或 CreateKeyAPI 创建客户托管的 AWS KMS 密钥。创建密钥时，请确保

- 选择对称加密客户管理的 KMS 密钥，Amazon Fraud Detector 不支持非对称 KMS 密钥。有关更多信息，请参阅《密钥管理服务开发人员指南》 AWS KMS中的非对称 AWS 密钥。
- 创建单个区域 KMS 密钥。Amazon Fraud Detector 不支持多区域 KMS 密钥。有关更多信息，请参阅《密钥管理服务开发人员指南》 AWS KMS中的多区域 AWS 密钥。
- 提供以下密钥政策，向 Amazon Fraud Detector 授予使用密钥的权限。

```
        {
    "Effect": "Allow",
    "Principal": {
        "Service": "frauddetector.amazonaws.com"
    },
    "Action": [
        "kms:Encrypt",
        "kms:Decrypt",
        "kms:ReEncrypt*",
        "kms:GenerateDataKey*",
        "kms:DescribeKey",
        "kms:CreateGrant",
        "kms:RetireGrant"
    ],
    "Resource": "*"
}
```

有关密钥策略的信息，请参阅《密钥管理服务开发人员指南》中的在 AWS KMS 中使用 AWS 密钥策略。

使用客户托管的 KMS 密钥加密数据

使用 Amazon Fra ud Detector 的 PutKMS EncryptionKey API 使用客户管理的 KMS 密钥加密您的静态亚马逊欺诈探测器数据。您可以随时使用 PutKMSEncryptionKey API 更改加密配置。

有关加密数据的重要说明

- 设置客户托管的 KMS 密钥后生成的数据经过加密。在设置客户托管的 KMS 密钥之前生成的数据将保持未加密状态。
- 如果更改了客户管理的 KMS 密钥，则使用先前加密配置加密的数据将不会被重新加密。

查看数据

当您使用客户托管的 KMS 密钥加密您的 Amazon Fraud Detector 数据时，无法使用亚马逊欺诈检测器控制台的 “搜索过去的预测” 区域中的筛选条件搜索使用此方法加密的数据。为确保搜索结果完整，请使用以下一个或多个属性来筛选结果：

- 事件 ID
- 评估时间戳
- 探测器状态
- 探测器版本
- 模型版本
- 模型类型
- 规则评估状态
- 规则执行模式
- 规则匹配状态
- 规则版本
- 可变数据源

如果客户管理的 KMS 密钥已被删除或计划删除，则您的数据可能不可用。有关更多信息，请参阅删除 KMS 密钥。

Amazon Fraud Detector 和接口 VPC 终端节点 (AWS PrivateLink)

您可以通过创建接口 VPC 终端节点在您的 VPC 和 Amazon Fraud Detector 之间建立私有连接。接口终端节点由一项技术提供支持 AWS PrivateLink，该技术使您无需互联网网关、NAT 设备、VPN 连接或 AWS Direct Connect 连接即可私下访问 Amazon Fraud Detector API。您的 VPC 中的实例不需要公有 IP 地址即可与 Amazon Fraud Detector API 通信。您的 VPC 和 Amazon Fraud Detector 之间的流量不会离开亚马逊网络。

每个接口端点均由子网中的一个或多个弹性网络接口表示。

有关更多信息，请参阅 A mazon VPC 用户指南中的接口 VPC 终端节点 (AWS PrivateLink)。

Amazon Fraud Detector VPC 终端节点的注意事项

在为 Amazon Fraud Detector 设置接口 VPC 终端节点之前，请务必查看亚马逊 VPC 用户指南中的接口终端节点属性和限制。

Amazon Fraud Detector 支持从你的 VPC 调用其所有 API 操作。

Amazon Fraud Detector 支持 VPC 终端节点策略。默认情况下，允许通过终端节点完全访问 Amazon Fraud Detector。有关更多信息，请参阅《Amazon VPC 用户指南》中的使用 VPC 端点控制对服务的访问权限。

为 Amazon Fraud Detector 创建接口 VPC 终端节点

您可以使用亚马逊 VPC 控制台或 AWS Command Line Interface (AWS CLI) 为 Amazon Fraud Detector 服务创建 VPC 终端节点。有关更多信息，请参阅《Amazon VPC 用户指南》中的创建接口端点。

使用以下服务名称为 Amazon Fraud Detector 创建 VPC 终端节点：

- com.amazonaws.*region*.frauddetector

例如，如果您为终端节点启用私有 DNS，则可以使用该区域的默认 DNS 名称向 Amazon Fraud Detector 发出 API 请求`frauddetector.us-east-1.amazonaws.com`。

有关更多信息，请参阅《Amazon VPC 用户指南》中的通过接口端点访问服务。

为 Amazon Fraud Detector 创建 VPC 终端节点策略

您可以为 Amazon Fraud Detector 的接口 VPC 终端节点创建策略，以指定以下内容：

- 可执行操作的主体
- 可执行的操作
- 可对其执行操作的资源

有关更多信息，请参阅《Amazon VPC 用户指南》中的使用 VPC 端点控制对服务的访问权限。

以下 VPC 终端节点策略示例，允许所有有权访问 VPC 接口终端节点的用户访问名为的 Amazon Fraud Detector 探测器`my_detector`。

```
{
  "Statement": [
      {
          "Action": "frauddetector:*Detector",
          "Effect": "Allow",
          "Resource": "arn:aws:frauddetector:us-east-1:123456789012:detector/
my_detector",
          "Principal": "*"
      }
  ]
}
```

在本示例中，会拒绝以下操作：

- 其他 Amazon Fraud Detector API 操作
- 调用 Amazon Fraud Detec `GetEventPrediction` tor API

> **Note**
>
> 在此示例中，用户仍然可以从 VPC 外部执行其他 Amazon Fraud Detector API 操作。有关如何将 API 调用限制为该 VPC 中执行的那些调用的信息，请参阅 Amazon Fraud Detector 基于身份的政策。

选择不使用您的数据来改善服务

您为训练模型和生成预测而提供的历史事件数据仅用于提供和维护您的服务。这些数据还可用于提高亚马逊 Fraud Detector 的质量。您的信任、隐私和内容安全是我们的首要任务，并确保我们的使用符合我们对您的承诺。有关更多信息，请参阅数据隐私常见问题解答

您可以通过访问 AWS O rganizations 用户指南中的人工智能服务选择退出政策页面并按照其中说明的流程选择不使用您的事件数据来开发或提高 A mazon Fraud Detecto r 的质量。

> **Note**
>
> 您的 AWS 账户需要由 AWS Organizations 集中管理，这样您才能使用选择退出政策。如果您尚未为您的 AWS 账户创建组织，请访问创建和管理组织页面并按照此处说明的流程进行操作。

Amazon Fraud Detector 的身份和访问管理

AWS Identity and Access Management (IAM) AWS 服务 可帮助管理员安全地控制对 AWS 资源的访问权限。IAM 管理员控制谁可以进行身份验证（登录）和授权（有权限）使用 Amazon Fraud Detector 资源。您可以使用 IAM AWS 服务 ，无需支付额外费用。

主题

- 受众
- 使用身份进行身份验证
- 使用策略管理访问
- Amazon Fraud Detector 如何与 IAM 协作
- Amazon Fraud Detector 基于身份的策略示例
- 混淆代理问题防范
- 对 Amazon Fraud Detector 身份和访问进行故障排

受众

您的使用方式 AWS Identity and Access Management (IAM) 会有所不同，具体取决于您在 Amazon Fraud Detector 中所做的工作。

服务用户 — 如果您使用 Amazon Fraud Detector 服务完成工作，则您的管理员会为您提供所需的凭证和权限。当您使用更多的 Amazon Fraud Detector 功能来完成工作时，您可能需要额外的权限。了解如何管理访问权限有助于您向管理员请求适合的权限。如果您无法访问 Amazon Fraud Detector 中的某项功能，请参阅对 Amazon Fraud Detector 身份和访问进行故障排。

服务管理员 — 如果你负责公司的 Amazon Fraud Detector 资源，那么你可能拥有对 Amazon Fraud Detector 的完全访问权限。您的工作是确定您的服务用户应该访问哪些 Amazon Fraud Detector 功能和资源。然后，您必须向 IAM 管理员提交请求以更改服务用户的权限。请查看该页面上的信息以了解 IAM 的基本概念。要详细了解贵公司如何将 IAM 与 Amazon Fraud Detector 配合使用，请参阅Amazon Fraud Detector 如何与 IAM 协作。

IAM 管理员 — 如果您是 IAM 管理员，则可能需要详细了解如何编写策略来管理 Amazon Fraud Detector 的访问权限。要查看您可以在 IAM 中使用的基于身份的 Amazon Fraud Detector 基于身份的策略示例，请参阅。Amazon Fraud Detector 基于身份的策略示例

使用身份进行身份验证

身份验证是您 AWS 使用身份凭证登录的方式。您必须以 IAM 用户身份或通过担 AWS 账户根用户任 IAM 角色进行身份验证（登录 AWS）。

您可以使用通过身份源提供的凭据以 AWS 联合身份登录。 AWS IAM Identity Center （IAM Identity Center）用户、贵公司的单点登录身份验证以及您的 Google 或 Facebook 凭据就是联合身份的示例。当您以联合身份登录时，您的管理员以前使用 IAM 角色设置了身份联合验证。当你使用联合访问 AWS 时，你就是在间接扮演一个角色。

根据您的用户类型，您可以登录 AWS Management Console 或 AWS 访问门户。有关登录的更多信息 AWS，请参阅《AWS 登录 用户指南》中的如何登录到您 AWS 账户的。

如果您 AWS 以编程方式访问，则会 AWS 提供软件开发套件 (SDK) 和命令行接口 (CLI)，以便使用您的凭据对请求进行加密签名。如果您不使用 AWS 工具，则必须自己签署请求。有关使用推荐的方法自行签署请求的更多信息，请参阅 IAM 用户指南中的签署 AWS API 请求。

无论使用何种身份验证方法，您可能需要提供其他安全信息。例如， AWS 建议您使用多重身份验证 (MFA) 来提高账户的安全性。要了解更多信息，请参阅《AWS IAM Identity Center 用户指南》中的多重身份验证和《IAM 用户指南》中的在 AWS中使用多重身份验证（MFA）。

AWS 账户 root 用户

创建时 AWS 账户，首先要有一个登录身份，该身份可以完全访问账户中的所有资源 AWS 服务 和资源。此身份被称为 AWS 账户 root 用户，使用您创建账户时使用的电子邮件地址和密码登录即可访问该身份。强烈建议您不要使用根用户执行日常任务。保护好根用户凭证，并使用这些凭证来执行仅根用户可以执行的任务。有关要求您以根用户身份登录的任务的完整列表，请参阅《IAM 用户指南》中的需要根用户凭证的任务。

用户和组

I AM 用户是您 AWS 账户 内部对个人或应用程序具有特定权限的身份。在可能的情况下，我们建议使用临时凭证，而不是创建具有长期凭证（如密码和访问密钥）的 IAM 用户。但是，如果您有一些特定的使用场景需要长期凭证以及 IAM 用户，建议您轮换访问密钥。有关更多信息，请参阅《IAM 用户指南》中的对于需要长期凭证的使用场景定期轮换访问密钥。

IAM 组是一个指定一组 IAM 用户的身份。您不能使用组的身份登录。您可以使用组来一次性为多个用户指定权限。如果有大量用户，使用组可以更轻松地管理用户权限。例如，您可能具有一个名为 IAMAdmins 的组，并为该组授予权限以管理 IAM 资源。

用户与角色不同。用户唯一地与某个人员或应用程序关联，而角色旨在让需要它的任何人代入。用户具有永久的长期凭证，而角色提供临时凭证。要了解更多信息，请参阅《IAM 用户指南》中的何时创建 IAM 用户（而不是角色）。

IAM 角色

I AM 角色是您内部具有特定权限 AWS 账户 的身份。它类似于 IAM 用户，但与特定人员不关联。您可以 AWS Management Console 通过切换角色在中临时担任 IAM 角色。您可以通过调用 AWS CLI 或 AWS API 操作或使用自定义 URL 来代入角色。有关使用角色的方法的更多信息，请参阅《IAM 用户指南》中的 使用 IAM 角色。

具有临时凭证的 IAM 角色在以下情况下很有用：

- 联合用户访问 – 要向联合身份分配权限，请创建角色并为角色定义权限。当联合身份进行身份验证时，该身份将与角色相关联并被授予由此角色定义的权限。有关联合身份验证的角色的信息，请参阅《IAM 用户指南》中的 为第三方身份提供商创建角色。如果您使用 IAM Identity Center，则需要配置权限集。为控制您的身份在进行身份验证后可以访问的内容，IAM Identity Center 将权限集与 IAM 中的角色相关联。有关权限集的信息，请参阅《AWS IAM Identity Center 用户指南》中的权限集。
- 临时 IAM 用户权限 – IAM 用户可代入 IAM 用户或角色，以暂时获得针对特定任务的不同权限。
- 跨账户存取 – 您可以使用 IAM 角色以允许不同账户中的某个人（可信主体）访问您的账户中的资源。角色是授予跨账户访问权限的主要方式。但是，对于某些资源 AWS 服务，您可以将策略直接附加到资源（而不是使用角色作为代理）。要了解用于跨账户访问的角色和基于资源的策略之间的差别，请参阅《IAM 用户指南》中的 IAM 角色与基于资源的策略有何不同。
- 跨服务访问 — 有些 AWS 服务 使用其他 AWS 服务服务中的功能。例如，当您在某个服务中进行调用时，该服务通常会在 Amazon EC2 中运行应用程序或在 Amazon S3 中存储对象。服务可能会使用发出调用的主体的权限、使用服务角色或使用服务相关角色来执行此操作。
  - 转发访问会话 (FAS) — 当您使用 IAM 用户或角色在中执行操作时 AWS，您被视为委托人。使用某些服务时，您可能会执行一个操作，然后此操作在其他服务中启动另一个操作。FAS 使用调用委托人的权限以及 AWS 服务 向下游服务发出请求的请求。 AWS 服务只有当服务收到需要与其他 AWS 服务 或资源交互才能完成的请求时，才会发出 FAS 请求。在这种情况下，您必须具有执行这两个操作的权限。有关发出 FAS 请求时的策略详情，请参阅转发访问会话。
  - 服务角色 - 服务角色是服务代表您在您的账户中执行操作而分派的 IAM 角色。IAM 管理员可以在 IAM 中创建、修改和删除服务角色。有关更多信息，请参阅《IAM 用户指南》中的创建向 AWS 服务委派权限的角色。
  - 服务相关角色-服务相关角色是一种链接到的服务角色。 AWS 服务服务可以代入代表您执行操作的角色。服务相关角色出现在您的中 AWS 账户 ，并且归服务所有。IAM 管理员可以查看但不能编辑服务相关角色的权限。

- 在 A@@ mazon EC2 上运行的应用程序 — 您可以使用 IAM 角色管理在 EC2 实例上运行并发出 AWS CLI 或 AWS API 请求的应用程序的临时证书。这优先于在 EC2 实例中存储访问密钥。要向 EC2 实例分配 AWS 角色并使其可供其所有应用程序使用，您需要创建附加到该实例的实例配置文件。实例配置文件包含角色，并使 EC2 实例上运行的程序能够获得临时凭证。有关更多信息，请参阅《IAM 用户指南》中的 使用 IAM 角色为 Amazon EC2 实例上运行的应用程序授予权限。

要了解是使用 IAM 角色还是 IAM 用户，请参阅《IAM 用户指南》中的何时创建 IAM 角色（而不是用户）。

使用策略管理访问

您可以 AWS 通过创建策略并将其附加到 AWS 身份或资源来控制中的访问权限。策略是其中的一个对象 AWS ，当与身份或资源关联时，它会定义其权限。 AWS 在委托人（用户、root 用户或角色会话）发出请求时评估这些策略。策略中的权限确定是允许还是拒绝请求。大多数策略都以 JSON 文档的 AWS 形式存储在中。有关 JSON 策略文档的结构和内容的更多信息，请参阅《IAM 用户指南》中的 JSON 策略概览。

管理员可以使用 AWS JSON 策略来指定谁有权访问什么。也就是说，哪个主体 可以对什么资源执行操作，以及在什么条件下执行。

默认情况下，用户和角色没有权限。要授予用户对所需资源执行操作的权限，IAM 管理员可以创建 IAM 策略。管理员随后可以向角色添加 IAM 策略，用户可以代入角色。

IAM 策略定义操作的权限，无关乎您使用哪种方法执行操作。例如，假设您有一个允许 `iam:GetRole` 操作的策略。拥有该策略的用户可以从 AWS Management Console AWS CLI、或 AWS API 获取角色信息。

基于身份的策略

基于身份的策略是可附加到身份（如 IAM 用户、用户组或角色）的 JSON 权限策略文档。这些策略控制用户和角色可在何种条件下对哪些资源执行哪些操作。要了解如何创建基于身份的策略，请参阅《IAM 用户指南》中的创建 IAM 策略。

基于身份的策略可以进一步归类为内联策略或托管式策略。内联策略直接嵌入单个用户、组或角色中。托管策略是独立的策略，您可以将其附加到中的多个用户、群组和角色 AWS 账户。托管策略包括 AWS 托管策略和客户托管策略。要了解如何在托管式策略和内联策略之间进行选择，请参阅《IAM 用户指南》中的在托管式策略与内联策略之间进行选择。

基于资源的策略

基于资源的策略是附加到资源的 JSON 策略文档。基于资源的策略的示例包括 IAM 角色信任策略和 Simple Storage Service（Amazon S3）存储桶策略。在支持基于资源的策略的服务中，服务管理员可以使用它们来控制对特定资源的访问。对于在其中附加策略的资源，策略定义指定主体可以对该资源执行哪些操作以及在什么条件下执行。您必须在基于资源的策略中指定主体。委托人可以包括账户、用户、角色、联合用户或 AWS 服务。

基于资源的策略是位于该服务中的内联策略。您不能在基于资源的策略中使用 IAM 中的 AWS 托管策略。

访问控制列表（ACL）

访问控制列表（ACL）控制哪些主体（账户成员、用户或角色）有权访问资源。ACL 与基于资源的策略类似，尽管它们不使用 JSON 策略文档格式。

Amazon S3 和 Amazon VPC 就是支持 ACL 的服务示例。 AWS WAF要了解有关 ACL 的更多信息，请参阅《Amazon Simple Storage Service 开发人员指南》 中的访问控制列表（ACL）概览。

其他策略类型

AWS 支持其他不太常见的策略类型。这些策略类型可以设置更常用的策略类型向您授予的最大权限。

- 权限边界 - 权限边界是一个高级功能，用于设置基于身份的策略可以为 IAM 实体（IAM 用户或角色）授予的最大权限。您可为实体设置权限边界。这些结果权限是实体基于身份的策略及其权限边界的交集。在 `Principal` 中指定用户或角色的基于资源的策略不受权限边界限制。任一项策略中的显式拒绝将覆盖允许。有关权限边界的更多信息，请参阅《IAM 用户指南》中的 IAM 实体的权限边界。
- 服务控制策略 (SCP)-SCP 是 JSON 策略，用于指定组织或组织单位 (OU) 的最大权限。 AWS Organizations AWS Organizations 是一项用于对您的企业拥有的多 AWS 账户 项进行分组和集中管理的服务。如果在组织内启用了所有功能，则可对任意或全部账户应用服务控制策略 (SCP)。SCP 限制成员账户中的实体（包括每个 AWS 账户根用户实体）的权限。有关 Organizations 和 SCP 的更多信息，请参阅《AWS Organizations 用户指南》中的 SCP 的工作原理。
- 会话策略 – 会话策略是当您以编程方式为角色或联合用户创建临时会话时作为参数传递的高级策略。结果会话的权限是用户或角色的基于身份的策略和会话策略的交集。权限也可以来自基于资源的策略。任一项策略中的显式拒绝将覆盖允许。有关更多信息，请参阅《IAM 用户指南》中的会话策略。

多个策略类型

当多个类型的策略应用于一个请求时，生成的权限更加复杂和难以理解。要了解在涉及多种策略类型时如何 AWS 确定是否允许请求，请参阅 IAM 用户指南中的策略评估逻辑。

Amazon Fraud Detector 如何与 IAM 协作

在使用 IAM 管理对 Amazon Fraud Detector 的访问权限之前，您应该了解哪些可用于 Amazon Fraud Detector 的 IAM 功能。要全面了解 Amazon Fraud Detector 和其他 AWS 服务如何与 IAM 配合使用，请参阅 IAM 用户指南中的与 IAM 配合使用的AWS 服务。

主题

- Amazon Fraud Detector 基于身份的政策
- Amazon Fraud Detector 基于资源的政策
- 基于亚马逊 Fraud Detector 标签的授权
- 亚马逊 Fraud Detector IAM 角色

Amazon Fraud Detector 基于身份的政策

通过使用 IAM 基于身份的策略，您可以指定允许或拒绝的操作和资源以及允许或拒绝操作的条件。Amazon Fraud Detector 支持特定的操作、资源和条件密钥。要了解在 JSON 策略中使用的所有元素，请参阅《IAM 用户指南》 中的 IAM JSON 策略元素参考。

要开始使用 Amazon Fraud Detector，我们建议创建一个只能访问亚马逊 Fraud Detector 操作并具有所需权限的用户。您可以根据需要添加其他权限。以下政策提供了使用 Amazon Fraud Detector 所需的权限：`AmazonFraudDetectorFullAccessPolicy`和`AmazonS3FullAccess`。有关使用这些政策设置 Amazon Fraud Detector 的更多信息，请参阅为 Amazon Fraud Detector 做好准备。

操作

管理员可以使用 AWS JSON 策略来指定谁有权访问什么。也就是说，哪个主体 可以对什么资源 执行操作，以及在什么 条件 下执行。

JSON 策略的 `Action` 元素描述可用于在策略中允许或拒绝访问的操作。策略操作通常与关联的 AWS API 操作同名。有一些例外情况，例如没有匹配 API 操作的仅限权限 操作。还有一些操作需要在策略中执行多个操作。这些附加操作称为相关操作。

在策略中包含操作以授予执行关联操作的权限。

Amazon Fraud Detector 中的策略操作在操作前使用以下前缀：`frauddetector:`。例如，要使用 Amazon Fraud Detecto `CreateRule` r API 操作创建规则，您需要在策略中包含该`frauddetector:CreateRule`操作。策略语句必须包含 `Action` 或 `NotAction` 元素。Amazon Fraud Detector 定义了自己的一组操作，这些操作描述了您可以使用此服务执行的任务。

要在单个语句中指定多项操作，请使用逗号将它们隔开，如下所示：

```
"Action": [
      "frauddetector:action1",
      "frauddetector:action2"
```

您也可以使用通配符（*) 指定多个操作。例如，要指定以单词 `Describe` 开头的所有操作，包括以下操作：

```
"Action": "frauddetector:Describe*"
```

要查看 Amazon Fraud Detector 操作列表，请参阅 IAM 用户指南中的 A mazon Fraud Detector 定义的操作。

资源

管理员可以使用 AWS JSON 策略来指定谁有权访问什么。也就是说，哪个主体 可以对什么资源 执行操作，以及在什么条件 下执行。

`Resource` JSON 策略元素指定要向其应用操作的一个或多个对象。语句必须包含 `Resource` 或 `NotResource` 元素。作为最佳实践，请使用其 Amazon 资源名称（ARN）指定资源。对于支持特定资源类型（称为资源级权限）的操作，您可以执行此操作。

对于不支持资源级权限的操作（如列出操作），请使用通配符 (*) 指示语句应用于所有资源。

```
"Resource": "*"
```

Amazon Fraud Detector 定义的资源类型列出了所有 Amazon Fraud Detector 资源 ARN。

例如，要在语句中指定`my_detector`检测器，请使用以下 ARN：

```
"Resource": "arn:aws:frauddetector:us-east-1:123456789012:detector/my_detector"
```

有关 ARN 格式的更多信息，请参阅 A mazon 资源名称 (ARN) 和 AWS 服务命名空间。

要指定属于特定账户的所有探测器，请使用通配符 (*)：

```
"Resource": "arn:aws:frauddetector:us-east-1:123456789012:detector/*"
```

某些 Amazon Fraud Detector 操作，例如用于创建资源的操作，无法对特定资源执行。在这些情况下，您必须使用通配符 （ *)。

```
"Resource": "*"
```

要查看 Amazon Fraud Detector 资源类型及其 ARN 的列表，请参阅 IAM 用户指南中的 A mazon Fraud Detector 定义的资源。要了解您可以为每种资源指定哪些操作的 ARN，请参阅 A mazon Fraud Detector 定义的操作。

条件键

管理员可以使用 AWS JSON 策略来指定谁有权访问什么。也就是说，哪个主体可以对什么资源执行操作，以及在什么条件下执行。

在 Condition 元素（或 Condition 块）中，可以指定语句生效的条件。Condition 元素是可选的。您可以创建使用条件运算符（例如，等于或小于）的条件表达式，以使策略中的条件与请求中的值相匹配。

如果您在一个语句中指定多个 Condition 元素，或在单个 Condition 元素中指定多个键，则 AWS 使用逻辑 AND 运算评估它们。如果您为单个条件键指定多个值，则使用逻辑OR运算来 AWS 评估条件。在授予语句的权限之前必须满足所有的条件。

在指定条件时，您也可以使用占位符变量。例如，只有在使用 IAM 用户名标记 IAM 用户时，您才能为其授予访问资源的权限。有关更多信息，请参阅《IAM 用户指南》中的 IAM 策略元素：变量和标签。

AWS 支持全局条件密钥和特定于服务的条件密钥。要查看所有 AWS 全局条件键，请参阅 IAM 用户指南中的AWS 全局条件上下文密钥。

Amazon Fraud Detector 定义了自己的一组条件键，还支持使用一些全局条件键。要查看所有 AWS 全局条件键，请参阅 IAM 用户指南中的AWS 全局条件上下文密钥。

要查看 Amazon Fraud Detector 条件密钥列表，请参阅 IAM 用户指南中的 A mazon Fraud Detector 条件密钥。要了解您可以使用条件键的操作和资源，请参阅 Amazon Fraud Detector 定义的操作。

示例

要查看 Amazon Fraud Detector 基于身份的政策示例，请参阅。Amazon Fraud Detector 基于身份的策略示例

Amazon Fraud Detector 基于资源的政策

Amazon Fraud Detector 不支持基于资源的政策。

基于亚马逊 Fraud Detector 标签的授权

您可以将标签附加到亚马逊 Fraud Detector 资源，也可以在请求中将标签传递给亚马逊 Fraud Detector。要基于标签控制访问，您需要使用 `aws:ResourceTag/key-name`、`aws:RequestTag/key-name` 或 `aws:TagKeys` 条件键在策略的条件元素中提供标签信息。

亚马逊 Fraud Detector IAM 角色

I AM 角色是您的 AWS 账户中具有特定权限的实体。

在 Amazon Fraud Detector 中使用临时证书

可以使用临时凭证进行联合身份验证登录，分派 IAM 角色或分派跨账户角色。您可以通过调用AssumeRole或之类的 AWS STS API 操作来获取临时安全证书GetFederationToken。

Amazon Fraud Detector 支持使用临时证书。

服务相关角色

服务相关角色允许 AWS 服务访问其他服务中的资源以代表您完成操作。服务相关角色显示在 IAM 账户中，并归该服务所有。IAM 管理员可以查看但不能编辑服务相关角色的权限。

Amazon Fraud Detector 不支持与服务相关的角色。

服务角色

此功能允许服务代表您担任服务角色。此角色允许服务访问其他服务中的资源以代表您完成操作。服务角色显示在您的 账户中，并由该账户拥有。这意味着管理员可以更改此角色的权限。但是，这样做可能会中断服务的功能。

Amazon Fraud Detector 支持服务角色。

Amazon Fraud Detector 基于身份的策略示例

默认情况下，用户和 IAM 角色无权创建或修改 Amazon Fraud Detector 资源。他们也无法使用 AWS Management Console AWS CLI、或 AWS API 执行任务。管理员必须创建 IAM policy，以便为用户和

角色授予权限以对所需的指定资源执行特定的 API 操作。然后，管理员必须将这些策略附加到需要这些权限的用户或组。

要了解如何使用这些示例 JSON 策略文档创建 IAM 基于身份的策略，请参阅《IAM 用户指南》中的在 JSON 选项卡上创建策略。

主题

- 策略最佳实践
- AWS 托管（预定义）的 Amazon Fraud Detector 政策
- 允许用户查看他们自己的权限
- 允许完全访问亚马逊 Fraud Detector 资源
- 允许对 Amazon Fraud Detector 资源进行只读访问
- 允许访问特定资源
- 使用双模式 API 时允许访问特定资源
- 根据标签限制访问权限

策略最佳实践

基于身份的策略决定了某人是否可以在您的账户中创建、访问或删除 Amazon Fraud Detector 资源。这些操作可能会使 AWS 账户产生成本。创建或编辑基于身份的策略时，请遵循以下准则和建议：

- 开始使用 AWS 托管策略并转向最低权限权限 — 要开始向用户和工作负载授予权限，请使用为许多常见用例授予权限的AWS 托管策略。它们在你的版本中可用 AWS 账户。我们建议您通过定义针对您的用例的 AWS 客户托管策略来进一步减少权限。有关更多信息，请参阅《IAM 用户指南》中的 AWS 托管策略或工作职能的AWS 托管策略。
- 应用最低权限 – 在使用 IAM 策略设置权限时，请仅授予执行任务所需的权限。为此，您可以定义在特定条件下可以对特定资源执行的操作，也称为最低权限许可。有关使用 IAM 应用权限的更多信息，请参阅《IAM 用户指南》中的 IAM 中的策略和权限。
- 使用 IAM 策略中的条件进一步限制访问权限 – 您可以向策略添加条件来限制对操作和资源的访问。例如，您可以编写策略条件来指定必须使用 SSL 发送所有请求。如果服务操作是通过特定的方式使用的，则也可以使用条件来授予对服务操作的访问权限 AWS 服务，例如 AWS CloudFormation。有关更多信息，请参阅《IAM 用户指南》中的 IAM JSON 策略元素：条件。
- 使用 IAM Access Analyzer 验证您的 IAM 策略，以确保权限的安全性和功能性 – IAM Access Analyzer 会验证新策略和现有策略，以确保策略符合 IAM 策略语言 (JSON) 和 IAM 最佳实践。IAM

Access Analyzer 提供 100 多项策略检查和可操作的建议，以帮助您制定安全且功能性强的策略。有关更多信息，请参阅《IAM 用户指南》中的 IAM Acess Analyzer 策略验证。

- 需要多重身份验证 (MFA)-如果 AWS 账户您的场景需要 IAM 用户或根用户，请启用 MFA 以提高安全性。若要在调用 API 操作时需要 MFA，请将 MFA 条件添加到您的策略中。有关更多信息，请参阅《IAM 用户指南》中的 配置受 MFA 保护的 API 访问。

有关 IAM 中的最佳实操的更多信息，请参阅《IAM 用户指南》中的 IAM 中的安全最佳实操。

AWS 托管（预定义）的 Amazon Fraud Detector 政策

AWS 通过提供由创建和管理的独立 IAM 策略来解决许多常见用例 AWS。这些 AWS 托管策略为常见用例授予必要的权限，这样您就不必调查需要哪些权限。有关更多信息，请参阅AWS Identity and Access Management 管理用户指南中的 AWS 托管策略。

以下 AWS 托管政策仅适用于 Amazon Fraud Detector，您可以将其附加给账户中的用户：

`AmazonFraudDetectorFullAccess`: 授予对 Amazon Fraud Detector 资源、操作和支持的操作的完全访问权限，包括：

- 列出并描述 Amazon 中的所有型号终端节点 SageMaker
- 列出账户中的所有 IAM 角色
- 列出所有 Amazon S3 存储桶
- 允许 IAM Pass 角色将角色传递给 Amazon Fraud Detector

此策略不提供不受限制的 S3 访问权限。如果您需要将模型训练数据集上传到 S3，则还需要`AmazonS3FullAccess`托管策略（或范围缩小的自定义 Amazon S3 访问策略）。

您可以通过登录 IAM 控制台并按策略名称进行搜索来查看策略的权限。您还可以创建自己的自定义 IAM 策略，根据需要授予 Amazon Fraud Detector 操作和资源的权限。您可以将这些自定义策略附加到需要它们的用户或组。

允许用户查看他们自己的权限

该示例说明了您如何创建策略，以允许 IAM 用户查看附加到其用户身份的内联和托管式策略。此策略包括在控制台上或使用 AWS CLI 或 AWS API 以编程方式完成此操作的权限。

```
{
    "Version": "2012-10-17",
    "Statement": [
```

```
        {
            "Sid": "ViewOwnUserInfo",
            "Effect": "Allow",
            "Action": [
                "iam:GetUserPolicy",
                "iam:ListGroupsForUser",
                "iam:ListAttachedUserPolicies",
                "iam:ListUserPolicies",
                "iam:GetUser"
            ],
            "Resource": ["arn:aws:iam::*:user/${aws:username}"]
        },
        {
            "Sid": "NavigateInConsole",
            "Effect": "Allow",
            "Action": [
                "iam:GetGroupPolicy",
                "iam:GetPolicyVersion",
                "iam:GetPolicy",
                "iam:ListAttachedGroupPolicies",
                "iam:ListGroupPolicies",
                "iam:ListPolicyVersions",
                "iam:ListPolicies",
                "iam:ListUsers"
            ],
            "Resource": "*"
        }
    ]
}
```

允许完全访问亚马逊 Fraud Detector 资源

以下示例允许用户 AWS 账户 完全访问所有 Amazon Fraud Detector 资源和操作。

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "frauddetector:*"
            ],
            "Resource": "*"
```

```
        }
    ]
}
```

允许对 Amazon Fraud Detector 资源进行只读访问

在本示例中，您授予用户对您的 Amazon Fraud Detector 资源的 AWS 账户 只读访问权限。

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "frauddetector:GetEventTypes",
                "frauddetector:BatchGetVariable",
                "frauddetector:DescribeDetector",
                "frauddetector:GetModelVersion",
                "frauddetector:GetEventPrediction",
                "frauddetector:GetExternalModels",
                "frauddetector:GetLabels",
                "frauddetector:GetVariables",
                "frauddetector:GetDetectors",
                "frauddetector:GetRules",
                "frauddetector:ListTagsForResource",
                "frauddetector:GetKMSEncryptionKey",
                "frauddetector:DescribeModelVersions",
                "frauddetector:GetDetectorVersion",
                "frauddetector:GetPrediction",
                "frauddetector:GetOutcomes",
                "frauddetector:GetEntityTypes",
                "frauddetector:GetModels"
            ],
            "Resource": "*"
        }
    ]
}
```

允许访问特定资源

在此资源级策略示例中，您可以向用户授予 AWS 账户 访问除一个特定 Detector 资源之外的所有操作和资源的访问权限。

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "frauddetector:*"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Deny",
            "Action": [
                "frauddetector:*Detector"
            ],
            "Resource": "arn:${Partition}:frauddetector:${Region}:${Account}:detector/
${detector-name}"
        }
    ]
}
```

使用双模式 API 时允许访问特定资源

Amazon Fraud Detector 提供双模式获取 API，既可用作 “列出” 操作，也可用作 “描述” 操作。在不带任何参数的情况下调用双模式 API 时，会返回与您的关联的指定资源的列表 AWS 账户。使用参数调用双模式 API 时，会返回指定资源的详细信息。资源可以是模型、变量、事件类型或实体类型。

双模式 API 支持 IAM 策略中的资源级权限。但是，只有在请求中提供一个或多个参数时，才会应用资源级别权限。例如，如果用户调用 GetVariablesAPI 并提供变量名称，如果变量资源或变量名称附加了 IAM 拒绝策略，则用户将收到AccessDeniedException错误。如果用户调用 GetVariables API 但未指定变量名，则会返回所有变量，这可能会导致信息泄露。

要仅允许用户查看特定资源的详细信息，请在 IAM 拒绝NotResource策略中使用 IAM 策略元素。将此策略元素添加到 IAM 拒绝策略后，用户只能查看NotResource区块中指定的资源的详细信息。有关更多信息，请参阅 IAM 用户指南 NotResource中的 IAM JSON 策略元素：。

以下示例策略允许用户访问 Amazon Fraud Detector 的所有资源。但是，NotResource策略元素用于将 GetVariablesAPI 调用限制为仅限前缀为user*job_*、和var*的变量名。

```
{
 "Version": "2012-10-17",
```

```
 "Statement": [
  {
    "Effect": "Allow",
    "Action": "frauddetector:*",
    "Resource": "*"
  },
 {
    "Effect": "Deny",
    "Action": "frauddetector:GetVariables",
    "NotResource": [
       "arn:aws:frauddetector:*:*:variable/user*",
       "arn:aws:frauddetector:*:*:variable/job_*",
       "arn:aws:frauddetector:*:*:variable/var*"
    ]
  }
 ]
}
```

响应

对于此示例策略，响应表现出以下行为：

- 不包含变量名的 GetVariables 调用会导致`AccessDeniedException`错误，因为该请求映射到 Deny 语句。
- 如果 GetVariables 调用包含不允许的变量名，则会导致`AccessDeniedException`错误，因为变量名未映射到`NotResource`块中的变量名。例如，使用变量名的 GetVariables 调用会`email_address`导致`AccessDeniedException`错误。
- 如果 GetVariables 调用包含与`NotResource`块中的变量名相匹配的变量名，则按预期返回。例如，包含变量名的 GetVariables 调用`job_cpa`会返回`job_cpa`变量的详细信息。

根据标签限制访问权限

此示例策略演示了如何根据资源标签限制对 Amazon Fraud Detector 的访问权限。此示例假设：

- 在你的游戏中， AWS 账户 你定义了两个不同的小组，分别是 Team1 和 Team2
- 您已经创建了四个探测器
- 你想允许 Team1 的成员在 2 个探测器上进行 API 调用
- 你想允许 Team2 的成员在其他 2 个探测器上进行 API 调用

控制对 API 调用的访问（示例）

1. 向 Team1 使用的探测器添加带有键`Project`和值`A`的标签。
2. 向 Team2 使用的探测器添加带有键`Project`和值`B`的标签。
3. 创建一个 IAM 策略，其`ResourceTag`条件是拒绝访问带有密钥`Project`和值的标签的探测器`B`，并将该策略附加到 Team1。
4. 创建一个 IAM 策略，其`ResourceTag`条件是拒绝访问带有密钥`Project`和值的标签的探测器`A`，并将该策略附加到 Team2。

以下是一个政策示例，该政策拒绝对任何带有密钥为`Project`且值为的标签的 Amazon Fraud Detector 资源执行特定操作`B`：

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "frauddetector:*",
      "Resource": "*"
    },
    {
      "Effect": "Deny",

      "Action": [

        "frauddetector:CreateModel",
        "frauddetector:CancelBatchPredictionJob",
        "frauddetector:CreateBatchPredictionJob",
        "frauddetector:DeleteBatchPredictionJob",
        "frauddetector:DeleteDetector"
      ],

      "Resource": "*",
      "Condition": {
        "StringEquals": {
          "aws:ResourceTag/Project": "B"
        }
      }
    }
  ]
}
```

混淆代理问题防范

当无权执行某项操作的实体可以强迫特权更高的实体执行该操作时，就会出现混乱的副手问题。 AWS 如果您向第三方（称为跨账户）或其他 AWS 服务（称为跨服务）提供对账户中资源的访问权限，则这些工具可帮助您保护自己的账户。

当一个服务（呼叫服务）呼叫另一个服务（被叫服务）时，可能会出现跨服务混淆的代理问题。可以操纵调用服务以使用其权限对另一个客户的资源进行操作，否则该服务不应有访问权限。为防止出现这种情况，您可以创建策略来保护所有服务的数据，这些服务主体已被授予对您的服务资源的访问权限。

Amazon Fraud Detector 支持在您的权限策略中使用服务角色来允许服务代表您访问其他服务的资源。角色需要两个策略：一个角色信任策略，用于指定允许代入角色的主体，另一个权限策略用于指定可以对角色执行的操作。当服务代表您担任角色时，必须允许服务主体在角色信任策略中执行 `sts:AssumeRole` 操作。当服务调用时`sts:AssumeRole`，会 AWS STS 返回一组临时安全证书，服务主体使用这些证书来访问该角色的权限策略允许的资源。

为了防止出现跨服务混淆的副手问题，Amazon Fraud Detector 建议在角色信任策略中使用`aws:SourceArn`和`aws:SourceAccount`全局条件上下文密钥，将对角色的访问权限限制为仅限由预期资源生成的请求。

`aws:SourceAccount`指定账户 ID，`aws:SourceArn`指定与跨服务访问关联的资源的 ARN。`aws:SourceArn`必须使用 ARN 格式指定。在同一个政策声明中使用时，请确保`aws:SourceAccount`和`aws:SourceArn`使用相同的账户 ID。

防范混淆代理问题最有效的方法是使用 `aws:SourceArn` 全局条件上下文键和资源的完整 ARN。如果您不知道资源的完整 ARN 或者要指定多个资源，请使用带有通配符 (*) 的`aws:SourceArn`全局上下文条件密钥来表示 ARN 的未知部分。例如，`arn:aws:`*`servicename`*`:*:`*`123456789012`*`:*`。有关 Amazon Fraud Detector 资源和可在权限策略中使用的操作的信息，请参阅 Amazon Fraud Detector 的操作、资源和条件密钥。

以下角色信任策略示例在`aws:SourceArn`条件键中使用通配符 (*) 允许 Amazon Fraud Detector 访问与账户 ID 关联的多个资源。

```
{
        "Version": "2012-10-17",
        "Statement": [
             {
               "Effect": "Allow",
```

```
                "Principal": {
                "Service": [
                    "frauddetector.amazonaws.com"
                    ]
                },
                "Action": "sts:AssumeRole",
                "Condition": {
                "StringEquals": {
                    "aws:SourceAccount": "123456789012"
                },
                  "StringLike": {
                     "aws:SourceArn": "arn:aws:frauddetector:us-west-2:123456789012:*"
        }
      }
    }
  ]
}
```

以下角色信任策略仅允许 Amazon Fraud Detector 访问`external-model`资源。注意条件块中的`aws:SourceArn`参数。资源限定符是使用为进行 `PutExternalModel` API 调用而提供的模型端点构建的。

```
    {
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": [
          "frauddetector.amazonaws.com"
        ]
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "StringEquals": {
          "aws:SourceAccount": "123456789012"
        },
        "StringLike": {
          "aws:SourceArn": "arn:aws:frauddetector:us-west-2:123456789012:external-
model/MyExternalModeldoNotDelete-ReadOnly"
        }
      }
```

```
    }
  ]
}
```

对 Amazon Fraud Detector 身份和访问进行故障排

使用以下信息来帮助您诊断和修复在使用 Amazon Fraud Detector 和 IAM 时可能遇到的常见问题。

主题

- 我无权在 Amazon Fraud Detector 中执行任何操作
- 我无权执行 iam：PassRole
- 我想允许 AWS 账户以外的人访问我的 Amazon Fraud Detector 资源
- Amazon Fraud Detector 无法担任给定角色

我无权在 Amazon Fraud Detector 中执行任何操作

如果 AWS Management Console 告诉您您无权执行某项操作，则必须联系管理员寻求帮助。管理员是向您提供登录凭证的人。

当`mateojackson`用户尝试使用控制台查看有关`####`详细信息但没有`frauddetector:GetDetectors`权限时，会出现以下示例错误。

```
User: arn:aws:iam::123456789012:user/mateojackson is not authorized to perform:
 frauddetector:GetDetectors on resource: my-example-detector
```

在这种情况下，Mateo 请求他的管理员更新其策略，以允许他使用 `frauddetector:GetDetectors` 操作访问 `my-example-detector` 资源。

我无权执行 iam：PassRole

如果您收到错误消息，说您无权执行该`iam:PassRole`操作，则必须更新您的政策，以允许您将角色传递给 Amazon Fraud Detector。

有些 AWS 服务 允许您将现有角色传递给该服务，而不是创建新的服务角色或服务相关角色。为此，您必须具有将角色传递到服务的权限。

当名为的 IAM 用户`marymajor`尝试使用控制台在 Amazon Fraud Detector 中执行操作时，会出现以下示例错误。但是，服务必须具有服务角色所授予的权限才可执行此操作。Mary 不具有将角色传递到服务的权限。

```
User: arn:aws:iam::123456789012:user/marymajor is not authorized to perform:
 iam:PassRole
```

在这种情况下，必须更新 Mary 的策略以允许她执行 `iam:PassRole` 操作。

如果您需要帮助，请联系您的 AWS 管理员。您的管理员是提供登录凭证的人。

我想允许 AWS 账户以外的人访问我的 Amazon Fraud Detector 资源

您可以创建一个角色，以便其他账户中的用户或您组织外的人员可以使用该角色来访问您的资源。您可以指定谁值得信赖，可以担任角色。对于支持基于资源的策略或访问控制列表（ACL）的服务，您可以使用这些策略向人员授予对您的资源的访问权。

要了解更多信息，请参阅以下内容：

- 要了解 Amazon Fraud Detector 是否支持这些功能，请参阅Amazon Fraud Detector 如何与 IAM 协作。
- 要了解如何提供对您拥有的资源的访问权限 AWS 账户 ，请参阅 IAM 用户指南中的向您拥有 AWS 账户 的另一个 IAM 用户提供访问权限。
- 要了解如何向第三方提供对您的资源的访问权限 AWS 账户，请参阅 IAM 用户指南中的向第三方提供访问权限。 AWS 账户
- 要了解如何通过身份联合验证提供访问权限，请参阅《IAM 用户指南》中的为经过外部身份验证的用户（身份联合验证）提供访问权限。
- 要了解使用角色和基于资源的策略进行跨账户存取之间的差别，请参阅《IAM 用户指南》中的 IAM 角色与基于资源的策略有何不同。

Amazon Fraud Detector 无法担任给定角色

如果您收到错误消息 Amazon Fraud Detector 无法担任给定角色，则必须更新指定角色的信任关系。通过将 Amazon Fraud Detector 指定为可信实体，该服务可以担任该角色。当您使用 Amazon Fraud Detector 创建角色时，会自动设置这种信任关系。您只需要为不是由 Amazon Fraud Detector 创建的 IAM 角色建立这种信任关系。

为现有角色与 Amazon Fraud Detector 建立信任关系

1. 通过以下网址打开 IAM 控制台：https://console.aws.amazon.com/iam/。
2. 在导航窗格中选择 “角色”。
3. 选择要修改的角色的名称，然后选择信任关系选项卡。

4. 选择编辑信任关系。
5. 在 Policy Document 下，粘贴以下内容，然后选择 Update Trust Policy。

```
{
    "Version": "2012-10-17",
    "Statement": [ {
        "Effect": "Allow",
        "Principal": {
            "Service": "frauddetector.amazonaws.com"
        },
        "Action": "sts:AssumeRole"
    } ]
}
```

在 Amazon Fraud Detector 中记录和监控

AWS 提供以下监控工具，用于监视 Amazon Fraud Detector，在出现问题时进行报告，并在适当时自动采取行动：

- Amazon 会实时 CloudWatch监控您的 AWS 资源和您运行 AWS 的应用程序。有关的更多信息 CloudWatch，请参阅 Amazon CloudWatch 用户指南。
- AWS CloudTrail捕获由您的账户或代表您的 AWS 账户进行的 API 调用和相关事件，并将日志文件传输到您指定的 Amazon S3 存储桶。有关更多信息 CloudTrail，请参阅《AWS CloudTrail 用户指南》。

有关监控 Amazon Fraud Detector 的更多信息，请参阅监控亚马逊 Fraud Detector。

Amazon Fraud Detector 的合规性验证

作为多个合 AWS 规计划（例如 SOC、PCI、FedRAMP 和 HIPAA）的一部分，第三方审计师评估 AWS 服务的安全性和合规性。

要了解是否属于特定合规计划的范围，请参阅AWS 服务 “按合规计划划分的范围” ”，然后选择您感兴趣的合规计划。 AWS 服务 有关一般信息，请参阅AWS 合规计划AWS。

您可以使用下载第三方审计报告 AWS Artifact。有关更多信息，请参阅中的 “下载报告” 中的 “ AWS Artifact。

您在使用 AWS 服务 时的合规责任取决于您的数据的敏感性、贵公司的合规目标以及适用的法律和法规。 AWS 提供了以下资源来帮助实现合规性：

- 安全与合规性快速入门指南 — 这些部署指南讨论了架构注意事项，并提供了在这些基础上 AWS 部署以安全性和合规性为重点的基准环境的步骤。
- 在 A@@ mazon Web Services 上构建 HIPAA 安全与合规架构 — 本白皮书描述了各公司如何使用 AWS 来创建符合 HIPAA 资格的应用程序。

> **Note**
>
> 并非所有 AWS 服务 人都符合 HIPAA 资格。有关更多信息，请参阅符合 HIPAA 要求的服务参考。

- AWS 合规资源AWS — 此工作簿和指南集可能适用于您所在的行业和所在地区。
- AWS 客户合规指南 — 从合规角度了解责任共担模式。这些指南总结了保护的最佳实践， AWS 服务并将指南映射到跨多个框架（包括美国国家标准与技术研究院 (NIST)、支付卡行业安全标准委员会 (PCI) 和国际标准化组织 (ISO)）的安全控制。
- 使用AWS Config 开发人员指南中的规则评估资源 — 该 AWS Config 服务评估您的资源配置在多大程度上符合内部实践、行业准则和法规。
- AWS Security Hub— 这 AWS 服务 可以全面了解您的安全状态 AWS。Security Hub 通过安全控件评估您的 AWS 资源并检查其是否符合安全行业标准和最佳实践。有关受支持服务及控件的列表，请参阅 Security Hub 控件参考。
- Amazon GuardDuty — 它通过监控您的 AWS 账户环境中是否存在可疑和恶意活动，来 AWS 服务检测您的工作负载、容器和数据面临的潜在威胁。 GuardDuty 通过满足某些合规性框架规定的入侵检测要求，可以帮助您满足各种合规性要求，例如 PCI DSS。
- AWS Audit Manager— 这 AWS 服务 可以帮助您持续审计 AWS 使用情况，从而简化风险管理以及对法规和行业标准的合规性。

Amazon Fraud Detector 中的弹性

亚马逊云科技 全球基础设施围绕亚马逊云科技区域和可用区构建。亚马逊云科技区域提供多个在物理上独立且隔离的可用区，这些可用区通过延迟低、吞吐量高且冗余性高的网络连接在一起。利用可用

区，您可以设计和操作在可用区之间无中断地自动实现失效转移的应用程序和数据库。与传统的单个或多个数据中心基础设施相比，可用区具有更高的可用性、容错性和可扩展性。

有关 AWS 区域和可用区的更多信息，请参阅 AWS 全球基础设施。

Amazon Fraud Detector 中的基础设施安全

作为一项托管服务，Amazon Fraud Detecto AWS r 受到全球网络安全的保护。有关 AWS 安全服务以及如何 AWS 保护基础设施的信息，请参阅AWS 云安全。要使用基础设施安全的最佳实践来设计您的 AWS 环境，请参阅 S AWS ecurity Pillar Well-Architected Fram ework 中的基础设施保护。

您可以使用 AWS 已发布的 API 调用通过网络访问 Amazon Fraud Detector。客户端必须支持以下内容：

- 传输层安全性协议 (TLS)。我们要求使用 TLS 1.2，建议使用 TLS 1.3。
- 具有完全向前保密 (PFS) 的密码套件，例如 DHE（临时 Diffie-Hellman）或 ECDHE（临时椭圆曲线 Diffie-Hellman）。大多数现代系统（如 Java 7 及更高版本）都支持这些模式。

此外，必须使用访问密钥 ID 和与 IAM 委托人关联的秘密访问密钥来对请求进行签名。或者，您可以使用 AWS Security Token Service（AWS STS）生成临时安全凭证来对请求进行签名。

监控亚马逊 Fraud Detector

监控是维护 Amazon Fraud Detector 和您的其他 AWS 解决方案的可靠性、可用性和性能的重要组成部分。AWS 提供以下监控工具，用于监视 Amazon Fraud Detector，在出现问题时进行报告，并在适当时自动采取行动：

- Amazon 会实时 CloudWatch监控您的 AWS 资源和您运行 AWS 的应用程序。您可以收集和跟踪指标，创建自定义的控制平面，以及 设置警报以在指定的指标达到您指定的阈值时通知您或采取措施。有关更多信息，请参阅 Amazon CloudWatch 用户指南。
- AWS CloudTrail捕获由您的账户或代表您的 AWS 账户进行的 API 调用和相关事件，并将日志文件传输到您指定的 Amazon S3 存储桶。您可以标识哪些用户和账户调用了 AWS、发出调用的源 IP 地址以及调用的发生时间。有关更多信息，请参阅AWS CloudTrail 《用户指南》。

主题

- 使用亚马逊监控 Amazon Fraud Detector CloudWatch
- 使用记录亚马逊 Fraud Detector API 调用 AWS CloudTrail

使用亚马逊监控 Amazon Fraud Detector CloudWatch

您可以使用监控 Amazon Fraud Detector CloudWatch，它收集原始数据并将其处理为可读的近乎实时的指标。这些统计数据会保存 15 个月，从而使您能够访问历史信息，并能够更好地了解您的 Web 应用程序或服务的执行情况。此外，可以设置用于监测特定阈值的警报，并在达到相应阈值时发送通知或执行操作。有关更多信息，请参阅 Amazon CloudWatch 用户指南。

主题

- 使用 Amazon Fraud Detector 的 CloudWatch 指标。
- Amazon Fraud Detector 指标

使用 Amazon Fraud Detector 的 CloudWatch 指标。

要使用指标，您必须指定以下信息：

- 指标命名空间。命名空间是 Amazon Fraud Detector 用来发布其指标的 CloudWatch 容器。如果您使用 CloudWatch ListMetricsAPI 或 list-metrics 命令来查看 Amazon Fraud Detector 的指标，请指定命名`AWS/FraudDetector`空间。

- 指标维度。维度是一种名称/值对，可帮助您唯一标识指标，例如，DetectorId可以是维度名称。指定指标维度是可选的。
- 指标名称，如 GetEventPrediction。

您可以使用 AWS Management Console、或 CloudWatch API 获取 Amazon Fraud Detector 的监控数据。 AWS CLI您也可以通过亚马逊 AWS 软件开发套件 (SDK) 或 CloudWatch API 工具使用 API。CloudWatch 控制台根据来自 CloudWatch API 的原始数据显示一系列图表。根据您的需求差异，您可能倾向于使用控制台中显示的图表，也可能倾向于检索自 API 的图表。

下面的列表显示这些指标的一些常见用途。这些是入门建议，并不全面。

如何？	相关指标
如何跟踪已执行的预测数量？	监控 GetEventPrediction 指标。
如何监控GetEventPrediction 错误？	使用GetEventPrediction 5xxError 和GetEventPrediction 4xxError 指标。
我如何监控 GetEventPrediction 调用的延迟？	使用 GetEventPredictionLatency 指标。

您必须拥有相应的 CloudWatch 权限才能监控 Amazon Fraud Detector CloudWatch。有关更多信息，请参阅 Amazon 的身份验证和访问控制 CloudWatch。

访问 Amazon Fraud Detector 指标

以下步骤展示了如何使用 CloudWatch 控制台访问 Amazon Fraud Detector 指标。

要查看指标（控制台）

1. 打开 CloudWatch 控制台，网址为 https://console.aws.amazon.com/cloudwatch/。
2. 选择 “指标”，选择 “所有指标” 选项卡，然后选择 F raud Detector。
3. 选择指标维度。
4. 从列表中选择所需的指标，然后为图表选择时间段。

创建警报

您可以创建一个 CloudWatch 警报，在警报状态发生变化时发送亚马逊简单通知服务 (Amazon SNS) Simple Notification Service 消息。告警会监控您指定的时间段内的某个指标。它在多个时间段内根据相对于给定阈值的指标值，执行一项或多项操作。操作是一个发送到 Amazon SNS 主题或 Auto Scaling 策略的通知。

警报仅针对持续的状态变化调用操作。 CloudWatch 警报不会仅仅因为它们处于特定状态就调用操作。该状态必须改变并在指定数量的时间段内一直保持。

设置警报（控制台）

1. 登录 AWS Management Console 并打开 CloudWatch 控制台，网址为 https://console.aws.amazon.com/cloudwatch/。
2. 在导航窗格中，选择警报，然后选择创建警报。这将打开 “创建警报向导”。
3. 选择选择指标。
4. 在 “所有指标” 选项卡中，选择 F raud Detec tor。
5. 选择 “按探测器 ID”，然后选择GetEventPrediction指标。
6. 选择绘成图表的指标选项卡。
7. 对于 Statistic（统计数据），选择 Sum（总计）。
8. 选择选择指标。
9. 对于 “条件”，为 “阈值” 类型选择 “静态”，为 “随时...” 选择 “更大”，然后输入您选择的最大值。选择下一步。
10. 要将警报发送到现有 Amazon SNS 主题，对于发送通知到：，请选择现有 SNS 主题。要为新的电子邮件订阅列表设置名称和电子邮件地址，请选择新建列表。 CloudWatch 保存列表并将其显示在字段中，这样您就可以用它来设置 future 的警报。

    > **Note**
    >
    > 如果您使用新列表创建一个新的 Amazon SNS 主题，则必须先验证电子邮件地址，然后目标收件人才能接收通知。Amazon SNS 仅在警报进入警报状态时发送电子邮件。如果警报状态变化发生在电子邮件地址验证之前，则目标收件人不会收到通知。

11. 选择下一步。为您的闹钟添加名称和可选描述。选择下一步。
12. 选择创建警报。

Amazon Fraud Detector 指标

Amazon Fraud Detector 向发送以下指标 CloudWatch。所有指标都支持以下统计信息：Average、Minimum、Maximum、Sum。

指标	描述
GetEventPrediction	GetEventPrediction API 请求的数量。 有效维度：DetectorID
GetEventPredictionLatency	响应来自请求的客户端请求所用的时间间隔。GetEventPrediction 有效维度：DetectorID 单位：毫秒
GetEventPrediction4XXError	Amazon Fraud Detector 返回 4xx HTTP 响应代码的 GetEventPrediction 请求数量。对于每个 4xx 响应，将发送 1 个。 有效维度：DetectorID
GetEventPrediction5XXError	Amazon Fraud Detector 返回 5xx HTTP 响应码的 GetEventPrediction 请求数量。对于每个 5xx 响应，将发送 1 个。 有效维度：DetectorID
Prediction	预测的数量。如果成功则发送 1。 有效尺寸：DetectorID ，DetectorVersionID
PredictionLatency	预测操作所用的时间间隔。 有效尺寸：DetectorID ，DetectorVersionID 单位：毫秒

指标	描述
PredictionError	Amazon Fraud Detector 遇到错误的预测次数。如果遇到错误，则发送 1。 有效尺寸：DetectorID ，DetectorVersionID
VariableUsed	使用变量作为评估一部分的 GetEventPrediction 请求数。 有效尺寸：DetectorID 、DetectorVersionID 、VariableName
VariableDefaultReturned	变量未作为事件属性的一部分存在的 GetEventPrediction 请求数，因此在评估期间使用了变量的默认值。 有效尺寸：DetectorID 、DetectorVersionID 、VariableName
RuleNotEvaluated	由于先前的规则匹配而未评估规则的 GetEventPrediction 请求数。 有效尺寸：DetectorID 、DetectorVersionID 、RuleID
RuleEvaluateTrue	规则触发为 True 且返回规则结果的 GetEventPrediction 请求数。 有效尺寸：DetectorID 、DetectorVersionID 、RuleID
RuleEvaluateFalse	规则评估为 False 的 GetEventPrediction 请求数。 有效尺寸：DetectorID 、DetectorVersionID 、RuleID

指标	描述
RuleEvaluateError	规则评估出错的 GetEventPrediction 请求数 有效尺寸：DetectorID 、DetectorVersionID 、RuleID
OutcomeReturned	返回指定结果的 GetEventPrediction 呼叫数。 有效尺寸：DetectorID 、DetectorVersionID 、OutcomeName
ModelInvocation (Amazon SageMaker model endpoint)	在评估过程中调用 SageMaker 模型端点的 GetEventPrediction 请求数。 有效尺寸：DetectorID 、DetectorVersionID 、ModelEndpoint
ModelInvocationError (Amazon SageMaker model endpoint)	在评估期间，被调用的 SageMaker 模型端点返回错误的 GetEventPrediction 请求数。 有效尺寸：DetectorID 、DetectorVersionID 、ModelEndpoint
ModelInvocationLatency (Amazon SageMaker model endpoint)	从 Amazon Fraud Detector 中查看的导入模型响应所花费的时间间隔。此间隔仅包括模型调用。 有效尺寸：DetectorID 、DetectorVersionID 、ModelEndpoint 单位：毫秒
ModelInvocation	在评估过程中调用模型的 GetEventPrediction 请求数。 有效尺寸：DetectorID 、DetectorVersionID 、ModelType 、ModelID

指标	描述
ModelInvocationError	评估期间，Amazon Fraud Detector 模型返回错误的 GetEventPrediction 请求数量。 有效尺寸：DetectorID 、DetectorVersionID 、ModelType 、ModelID
ModelInvocationLatency	从 Amazon Fraud Detector 中可以看出，亚马逊欺诈探测器模型做出响应所花费的时间间隔。此间隔仅包括模型调用。 有效尺寸：DetectorID 、DetectorVersionID 、ModelType 、ModelID 单位：毫秒

使用记录亚马逊 Fraud Detector API 调用 AWS CloudTrail

Amazon Fraud Detector 与 AWS CloudTrail一项服务集成，该服务可记录用户、角色或 AWS 服务在 Amazon Fraud Detector 中采取的操作。 CloudTrail 将所有 Amazon Fraud Detector 的 API 调用捕获为事件，包括来自亚马逊欺诈探测器控制台的调用以及从代码调用到亚马逊 Fraud Detector API 的调用。

如果您创建跟踪，则可以允许将 CloudTrail 事件持续传输到 Amazon S3 存储桶，包括 Amazon Fraud Detector 的事件。如果您未配置跟踪，您仍然可以在 CloudTrail 控制台的 “事件历史记录” 中查看最新的事件。通过收集的信息 CloudTrail，您可以确定向 Amazon Fraud Detector 发出的请求、发出请求的 IP 地址、谁提出请求、何时提出请求以及其他详细信息。

要了解更多信息 CloudTrail，请参阅《AWS CloudTrail 用户指南》。

Amazon Fraud Detector 中的信息 CloudTrail

CloudTrail 在您创建 AWS 账户时已在您的账户上启用。当 Amazon Fraud Detector 中发生活动时，该活动会与其他 AWS 服务 CloudTrail 事件一起记录在事件历史记录中。您可以在自己的 AWS 账户中查看、搜索和下载最近发生的事件。有关更多信息，请参阅使用事件历史记录查看 CloudTrail 事件。

要持续记录您的 AWS 账户中的事件，包括 Amazon Fraud Detector 的事件，请创建跟踪。跟踪允许 CloudTrail 将日志文件传输到 Amazon S3 存储桶。默认情况下，在控制台中创建跟踪记录时，此跟踪

记录应用于所有 AWS 区域。跟踪记录 AWS 分区中所有区域的事件，并将日志文件传送到您指定的 Amazon S3 存储桶。此外，您可以配置其他 AWS 服务，以进一步分析 CloudTrail 日志中收集的事件数据并对其采取行动。有关更多信息，请参阅下列内容：

- 创建跟踪概述
- CloudTrail 支持的服务和集成
- 配置 Amazon SNS 通知 CloudTrail
- 接收来自多个区域的 CloudTrail 日志文件和接收来自多个账户的 CloudTrail日志文件

Amazon Fraud Detector 支持将每个操作（API 操作）作为事件 CloudTrail 记录在日志文件中。有关更多信息，请参阅操作。

每个事件或日记账条目都包含有关生成请求的人员信息。身份信息有助于您确定以下内容：

- 请求是使用根凭证还是用户凭证发出的。
- 请求是使用角色还是联合用户的临时安全凭证发出的。
- 请求是否由其他 AWS 服务发出。

有关更多信息，请参阅CloudTrail 用户身份元素。

了解 Amazon Fraud Detector 日志文件条目

跟踪是一种配置，允许将事件作为日志文件传输到您指定的 Amazon S3 存储桶。 CloudTrail 日志文件包含一个或多个日志条目。事件代表来自任何来源的单个请求，包括有关所请求的操作、操作的日期和时间、请求参数等的信息。 CloudTrail 日志文件不是公共 API 调用的有序堆栈跟踪，因此它们不会按任何特定顺序出现。

以下示例显示了演示该`GetDetectors`操作的 CloudTrail 日志条目。

```
{
    "eventVersion": "1.05",
    "userIdentity": {
        "type": "IAMUser",
        "principalId": "principal-id",
        "arn": "arn:aws:iam::user-arn",
        "accountId": "account-id",
        "accessKeyId": "access-key",
        "userName": "user-name"
```

```
    },
    "eventTime": "2019-11-22T02:18:03Z",
    "eventSource": "frauddetector.amazonaws.com",
    "eventName": "GetDetectors",
    "awsRegion": "us-east-1",
    "sourceIPAddress": "source-ip-address",
    "userAgent": "aws-cli/1.11.16 Python/2.7.11 Darwin/15.6.0 botocore/1.4.73",
    "requestParameters": null,
    "responseElements": null,
    "requestID": "request-id",
    "eventID": "event-id",
    "eventType": "AwsApiCall",
    "recipientAccountId": "recipient-account-id"
}
```

故障排除

以下各节可帮助您解决在使用 Amazon Fraud Detector 时可能遇到的问题

对训练数据问题进行故障排除

使用本节中的信息来帮助诊断和解决您在训练模型时可能在 Amazon Fraud Detector 控制台的模型训练诊断窗格中看到的问题。

模型训练诊断窗格中显示的问题分类如下。解决问题的要求取决于问题的类别。

- 错误-导致模型训练失败。必须解决这些问题，模型才能成功训练。
- 警告-导致模型训练继续，但是，在训练过程中，某些变量可能会被排除在外。请查看本节中的相关指南，以提高数据集的质量。
- 信息（信息）-对模型训练没有影响，所有变量都用于训练。我们建议您查看本节中的相关指南，以进一步提高数据集的质量和模型性能。

主题

- 给定数据集中的欺诈率不稳定
- 数据不足
- 缺少或不同的 EVENT_LABEL 值
- 缺少或错误的 EVENT_TIMESTAMP 值
- 未摄取数据
- 变量不足
- 变量类型缺失或不正确
- 缺少变量值
- 唯一变量值不足
- 变量表达式不正确
- 唯一实体不足

给定数据集中的欺诈率不稳定

问题类型：错误

描述

随着时间的推移，给定数据中的欺诈率过于不稳定。请确保在一段时间内对您的欺诈和合法事件进行统一抽样。

原因

如果数据集中的欺诈和合法事件分布不均匀且来自不同的时段，则会发生此错误。Amazon Fraud Detector 模型训练流程根据 EVENT_TIMESTAMP 对您的数据集进行采样和分区。例如，如果您的数据集包含从过去 6 个月提取的欺诈事件，但仅包括最后一个月的合法事件，则该数据集被视为不稳定。不稳定的数据集可能会导致模型性能评估出现偏差。

解决方案

确保提供来自同一时段的欺诈和合法事件数据，欺诈率不会随着时间的推移而发生显著变化。

数据不足

1. 问题类型：错误

   描述

   少于 50 行被标记为欺诈事件。确保欺诈事件和合法事件均超过最小计数 50，然后重新训练模型。

   原因

   如果您的数据集中标记为欺诈的事件少于模型训练所需的数量，则会发生此错误。Amazon Fraud Detector 需要至少 50 个欺诈事件才能训练您的模型。

   解决方案

   确保您的数据集至少包含 50 个欺诈事件。如果需要，您可以通过覆盖更长的时间来确保这一点。
2. 问题类型：错误

   描述

   标记为合法事件的行少于 50 行。确保欺诈事件和合法事件均超过最低限额 $threshold，然后重新训练模型。

原因

如果您的数据集中标记为合法的事件少于模型训练所需的事件，则会发生此错误。Amazon Fraud Detector 需要至少 50 个合法事件才能训练您的模型。

解决方案

确保您的数据集至少包含 50 个合法事件。如果需要，您可以通过覆盖更长的时间来确保这一点。

3. 问题类型：错误

   描述

   与欺诈相关的唯一实体数量少于 100 个。考虑加入更多欺诈实体的示例，以提高绩效。

   原因

   如果您的数据集包含欺诈事件的实体少于模型训练所需的实体，则会发生此错误。交易欺诈洞察 (TFI) 模型要求至少 100 个存在欺诈事件的实体，以确保最大限度地覆盖欺诈领域。如果所有欺诈事件都由一小部分实体执行，则该模型可能无法很好地概括。

   解决方案

   确保您的数据集包含至少 100 个存在欺诈事件的实体。如果需要，您可以确保这涵盖更长的时间。

4. 问题类型：错误

   描述

   与合法实体关联的唯一实体数量少于 100。考虑加入更多合法实体的示例，以提高绩效。

   原因

   如果您的数据集包含合法事件的实体少于模型训练所需的实体，则会发生此错误。交易欺诈洞察 (TFI) 模型要求至少 100 个实体拥有合法事件，以确保最大限度地覆盖欺诈领域。如果所有合法事件都由一小部分实体执行，则模型可能无法很好地概括。

   解决方案

   确保您的数据集包含至少 100 个具有合法事件的实体。如果需要，您可以确保这涵盖更长的时间。

5. 问题类型：错误

   描述

数据集中少于 100 行。确保总数据集中超过 100 行，且至少 50 行被标记为欺诈。

原因

如果您的数据集包含的记录少于 100 条，则会发生此错误。Amazon Fraud Detector 需要来自数据集中至少 100 个事件（记录）的数据进行模型训练。

解决方案

确保您的数据集中包含来自 100 多个事件的数据。

缺少或不同的 EVENT_LABEL 值

1. 问题类型：错误

   描述

   大于 1% 的 EVENT_LABEL 列是空值，或者是模型配置中定义的值以外的值。**$label_values**确保 EVENT_LABEL 列中的缺失值少于 1%，并且这些值是模型配置中定义的值。**$label_values**

   原因

   出现此错误是由于以下原因之一：

   - 在包含您的训练数据的 CSV 文件中，超过 1% 的记录在 EVENT_LABEL 列中存在缺失值。
   - 在包含您的训练数据的 CSV 文件中，超过 1% 的记录在 EVENT_LABEL 列中的值与与您的事件类型关联的值不同。

   在线欺诈洞察 (OFI) 模型要求在每条记录的 EVENT_LABEL 列中填入与您的事件类型（或，映射到）关联的标签之一。`CreateModelVersion`

   解决方案

   如果此错误是由缺少 EVENT_LABEL 值造成的，请考虑为这些记录分配适当的标签或从数据集中删除这些记录。如果此错误是因为某些记录的标签不在其中**label_values**，请确保将 EVENT_LABEL 列中的所有值添加到事件类型的标签中，并在模型创建时映射到欺诈或合法（欺诈、合法）。

2. 问题类型：信息

   描述

您的 EVENT_LABEL 列包含与模型配置中定义的值以外的空值或标签值。**`$label_values`**在训练之前，这些不一致的值被转换为 “非欺诈”。

原因

您之所以获得此信息，是因为以下原因之一：

- 在包含训练数据的 CSV 文件中，只有不到 1% 的记录在 EVENT_LABEL 列中存在缺失值
- 在包含您的训练数据的 CSV 文件中，EVENT_LABEL 列中只有不到 1% 的值与与您的事件类型关联的值不同。

在这两种情况下，模型训练都将成功。但是，那些缺少标签值或未映射标签值的事件的标签值会转换为合法的。如果您认为这是一个问题，请按照下面提供的解决方案进行操作。

解决方案

如果您的数据集中缺少 EVENT_LABEL 值，请考虑从数据集中删除这些记录。如果未映射为这些 EVENT_LABELS 提供的值，请确保每个事件的所有这些值都映射到欺诈或合法（欺诈、合法）。

缺少或错误的 EVENT_TIMESTAMP 值

1. 问题类型：错误

   描述

   您的训练数据集包含带有不符合可接受格式的时间戳的 EVENT_TIMESTAMP。确保格式是可接受的日期/时间戳格式之一。

   原因

   如果 EVENT_TIMESTAMP 列包含的值不符合 Amazon Fraud Detector 支持的时间戳格式，则会发生此错误。

   解决方案

   确保为 EVENT_TIMESTAMP 列提供的值符合支持的时间戳格式。如果 EVENT_TIMESTAMP 列中有缺失的值，则可以使用支持的时间戳格式回填这些值，也可以考虑完全删除事件，而不是输入、或之类的字符串。`none` `null` `missing`

2. 问题类型：错误

您的训练数据集包含缺失值的 EVENT_TIMESTAMP。确保没有缺失值。

原因

如果数据集中的 EVENT_TIMESTAMP 列缺少值，则会发生此错误。Amazon Fraud Detector 要求您的数据集中的 EVENT_TIMESTAMP 列具有值。

解决方案

确保数据集中的 EVENT_TIMESTAMP 列具有值，并且这些值符合支持的时间戳格式。如果 EVENT_TIMESTAMP 列中有缺失的值，则可以使用支持的时间戳格式回填这些值，也可以考虑完全删除事件，而不是输入、或之类的字符串。`none` `null` `missing`

未摄取数据

问题类型：错误

描述

未找到用于训练的已摄取事件，请检查您的训练配置。

原因

如果您使用存储在 Amazon Fraud Detector 中的事件数据创建模型，但在开始训练模型之前没有将数据集导入 Amazon Fraud Detector，则会发生此错误。

解决方案

使用 Amazon Fraud Detector 控制台中的 `CreateBatchImportJob` API 操作、API 操作或批量导入功能，首先导入您的事件数据，然后训练您的模型。`SendEvent`有关更多信息，请参阅存储的事件数据集。

> **Note**
>
> 我们建议您在完成数据导入后等待 10 分钟，然后再使用它来训练模型。

您可以使用 Amazon Fraud Detector 控制台来检查每种事件类型已经存储的事件数量。有关更多信息，请参阅查看存储事件的指标。

变量不足

问题类型：错误

描述

数据集必须包含至少 2 个适合训练的变量。

原因

如果您的数据集包含的适合模型训练的变量少于 2 个，则会发生此错误。只有当变量通过所有验证后，Amazon Fraud Detector 才会认为该变量适合模型训练。如果变量验证失败，则会在模型训练中将其排除在外，您将在模型训练诊断中看到一条消息。

解决方案

确保您的数据集至少有两个用值填充的变量并通过所有数据验证。请注意，您提供列标题（EVENT_TIMESTAMP、EVENT_ID、ENTITY_ID、EVENT_LABEL 等）的事件元数据行不被视为变量。

变量类型缺失或不正确

问题类型：警告

描述

的预期数据类型**$variable_name**是数字。查看和更新**$variable_name**您的数据集，然后重新训练模型。

原因

如果变量被定义为 NUMERIC 变量，但是在数据集中，它的值无法转换为 NUMERIC，则会收到此警告。因此，该变量被排除在模型训练中。

解决方案

如果要将其保留为 NUMERIC 变量，请确保您提供的值可以转换为浮点数。请注意，如果变量包含缺失值，请不要使用诸如`none`、`nenenull`、或之类的字符串填充它们`missing`。如果变量确实包含非数字值，请将其重新创建为 CATEGORICAL 或 FREE_FORM_TEXT 变量类型。

缺少变量值

问题类型：警告

描述

您的训练数据集中**$variable_name**缺少大于的**$threshold**值。考虑修改**$variable_name**数据集并重新训练以提高性能。

原因

如果由于缺失值太多而丢弃了指定的变量，则会收到此警告。Amazon Fraud Detector 允许变量缺失值。但是，如果一个变量有太多的缺失值，则它对模型的贡献不大，并且该变量会在模型训练中被删除。

解决方案

首先，确认这些缺失值不是由于数据收集和准备中的错误造成的。如果它们是错误，那么你可以考虑将它们从模型训练中删除。但是，如果您确实认为这些缺失值很有价值，但仍想保留该变量，则可以在模型训练和实时推理中手动用常量填充缺失值。

唯一变量值不足

问题类型：警告

描述

的唯一值计数小**$variable_name**于 100。查看和更新**$variable_name**您的数据集，然后重新训练模型。

原因

如果指定变量的唯一值数小于 100，则会收到此警告。阈值因变量类型而异。由于唯一值很少，数据集有可能不够通用，无法覆盖该变量的特征空间。因此，该模型可能无法很好地概括实时预测。

解决方案

首先，确保变量分布代表真实的业务流量。然后，你可以采用更多经过精细训练且基数更高的变量，例如使用`full_customer_name`代替`first_name`和`last_name`单独使用，也可以将变量类型更改为 CATEGORICAL，这样可以降低基数。

变量表达式不正确

1. 问题类型：信息

   描述

大于 50% 的**$email_variable_name**值与预期的正则表达式 http://emailregex.com 不匹配。考虑修改**$email_variable_name**数据集并重新训练以提高性能。

原因

如果您的数据集中超过 50% 的记录的电子邮件值不符合正则电子邮件表达式，因此未通过验证，则会显示此信息。

解决方案

格式化电子邮件变量值以符合正则表达式。如果缺少电子邮件值，我们建议将其留空，而不是用nonenull、或之类的字符串填充missing。

2. 问题类型：信息

描述

大于 50% 的**$IP_variable_name**值与 IPv4 或 IPv6 地址的正则表达式不匹配 https://digitalfortress.tech/tricks/top-15-commonly-used-regex/。考虑修改**$IP_variable_name**数据集并重新训练以提高性能。

原因

如果您的数据集中超过 50% 的记录的 IP 值不符合正则 IP 表达式，因此验证失败，则会显示此信息。

解决方案

格式化 IP 值以符合正则表达式。如果缺少 IP 值，我们建议将其留空，而不是用nonenull、或之类的字符串填充missing。

3. 问题类型：信息

描述

大于 50% 的**$phone_variable_name**值与基本的电话正则表达式 /$pattern/ 不匹配。考虑修改**$phone_variable_name**数据集并重新训练以提高性能。

原因

如果您的数据集中超过 50% 的记录的电话号码不符合常规电话号码表达式，因此未通过验证，则会显示此信息。

解决方案

格式化电话号码以符合正则表达式。如果缺少电话号码，我们建议将其留空，而不是用nonenull、或之类的字符串填充missing。

唯一实体不足

问题类型：信息

描述

唯一实体的数量小于 1500。考虑添加更多数据以提高性能。

原因

如果您的数据集的唯一实体数量少于建议的数量，则会显示此信息。交易欺诈洞察 (TFI) 模型使用时间序列汇总和通用交易功能来提供最佳性能。如果您的数据集的唯一实体太少，则大多数通用数据（例如 IP_ADDRESS、EMAIL_ADDRESS）可能没有唯一值。然后，还有一个风险，即该数据集不够通用，无法覆盖该变量的特征空间。因此，该模型可能无法很好地概括来自新实体的交易。

解决方案

包括更多实体。如果需要，可以延长训练数据的时间范围。

配额

您的对于每个 Amazon WS 服务都AWS 账户具有默认配额（以前称为限制）。除非另有说明，否则，每个配额是区域特定的。对于下表中提及的所有可数，您可以请求增加配额。有关更多信息，请参阅请求增加配额

下表按组件概述了亚马逊Fraud Detector 配额。

Amazon FFraud Detector d D

配额名称	默认配额	可调整
训练数据大小	5GB	否
每个账户的模型数	50	否
每个模型的版本	200	否
每个账户数	5	否
每个账户数	3	否
每个模型的并发训练作业	1	否

亚马逊Fraud Detector 检测器/变量/结果/规则

配额名称	默认配额	可调整
每个账户的变量	5000	否
每个 账户的规则数	5000	否
每个规则数	3	否
每个账户的结果	5000	否
每个账户的探测器	100	否

配额名称	默认配额	可调整
每个数	30	否
每个数	100	否
每个探测数	10	否
每个账户的标签	100	否
每个账户数	100	否
每个账户数	100	否

Amazon FFraud Detector d D

配额名称	默认配额	可调整
GetEventPrediction 每秒 API 数	200 TPS	是
每次 GetEventPrediction API 调用的有效负载大小	256 KB	否
每次 GetEventPrediction API 调用的输入数量	5000	否

文档历史记录

下表描述了亚马逊欺诈检测器用户指南中的重要更改。我们还经常更新《亚马逊欺诈检测器用户指南》，以解决您发送给我们的反馈。

变更	说明	日期
新的变量和数据类型	Amazon Fraud Detector 引入了新的变量类型和可用于提取有用信息的数据类型。	2023年6月5日
活动编排	事件编排使您可以使用 Amazon EventBridge 轻松将事件发送到AWS 服务进行下游处理。	2023 年 5 月 30 日
清单	列表资源允许您作为规则的一部分引用一组值，例如 IP 地址或电子邮件地址。在规则中使用列表来允许或拒绝访问或交易。	2023 年 2 月 14 日
数据模型浏览器	数据模型浏览器可深入了解亚马逊欺诈检测器创建欺诈检测模型所需的数据元素。在准备事件数据集之前，请使用数据模型浏览器。	2022 年 12 月 15 日
账户收购洞察模型	使用账户接管洞察 (ATI) 模型检测因恶意接管、网络钓鱼或凭证被盗而遭到入侵的账户。	2022 年 7 月 21 日
章节更新	更新了介绍章节，增加了有关亚马逊欺诈检测器的更多信息	2022年4月11日
变量丰富	启用对您提供的部分原始数据的扩充功能，以提高使用这些	2022 年 2 月 8 日

	数据元素并在 2022 年 2 月 8 日之前训练的模型的性能。	
选择退出政策	使用选择退出政策选择不将您的事件数据用于开发或提高 Amazon Fraud Detector 的质量。	2022 年 1 月 6 日
混乱的副手预防	创建策略以防止第三方或跨服务实体操纵有权代表其行事的实体以获取对您账户中资源的访问权限。	2021 年 12 月 6 日
创建事件数据集	使用创建事件数据集中提供的指导来准备和收集用于训练模型的数据。	2021 年 11 月 22 日
预测解释	使用预测解释来深入了解每个事件变量如何影响模型的欺诈预测分数。	2021 年 11 月 10 日
故障排除	使用训练数据问题疑难解答中的信息来帮助诊断和解决训练模型时可能在 Amazon Fraud Detector 控制台中看到的问题。	2021 年 10 月 11 日
交易欺诈洞察模型	使用交易欺诈洞察 (TFI) 模型检测在线或card-not-present交易欺诈。	2021 年 10 月 11 日

存储的事件	将您的事件数据存储在 Amazon Fraud Detector 中，然后使用存储的数据来训练您的模型。通过将事件数据存储在 Amazon Fraud Detector 中，您可以训练使用自动计算变量的模型来提高性能、简化模型再训练并更新欺诈标签以关闭机器学习反馈循环。	2021 年 10 月 11 日
模型变量重要性	使用模型变量重要性来深入了解哪些因素推动了模型性能的上升或下降，以及哪些模型变量的贡献最大。然后调整模型以提高整体性能。	2021 年 7 月 9 日
与 AWS CloudFormation 集成	AWS CloudFormation用于管理您的亚马逊欺诈检测器资源。	2021 年 5 月 10 日
批量预测	使用批量预测来获得对一组不需要实时评分的事件的预测。	2021 年 3 月 31 日
章节重做	重做 “入门” 和其他部分	2020 年 7 月 17 日
首次发布	首次发布	2019 年 12 月 2 日

本文属于机器翻译版本。若本译文内容与英语原文存在差异，则一律以英文原文为准。