sift

Sift Services

An explanation for the end user

Contents

Introduction

Online fraudsters can cause a great deal of harm to consumers and businesses by taking money, account credentials, and identity information. At Sift, we help protect thousands of websites and apps from various types of fraud and abuse (e.g., payment fraud, account takeover, fake accounts, scams and spam content) so those businesses can focus on safely delivering their services to end users like you. You can learn more about our mission to preserve trust in the internet here.

We want to explain how our Digital Trust & Safety services work, especially for those of you who are end users of businesses that use Sift (i.e., an end user of a Sift customer). From a technology perspective, detecting online fraud and abuse is extremely difficult—fraudsters are constantly evolving and adapting their techniques. At Sift, we've determined that the best way to protect consumers and online businesses from these bad actors is to leverage artificial intelligence (AI) and machine learning (ML). When we say we use AI or machine learning, we mean that we apply mathematical algorithms (e.g., logistic regression) that use statistics *to find patterns in large amounts of data.* Based on those patterns, we develop predictions as to whether an action or event is fraudulent.

In this ebook, we'll begin by describing the types of data that we process to predict fraud and how we obtain that data. We will also explain how we prepare the data for analysis and the techniques and algorithms we apply to predict fraud. And finally, we'll explain how our customers use our service and how this might impact you.

Before digging into the details, we'd like to make clear that we only use your data to provide online fraud prevention services to our customers—we do not sell, share, or use your data for any other purpose.

What data does Sift use to detect fraud?

First, machine learning is nothing without data. We need high-quality, relevant data to be truly effective at detecting and preventing fraud. Our models learn from **real-time** data that we receive from thousands of websites and apps (i.e., our customers) around the world. We analyze a variety of data types and formats that together create a holistic picture of emerging fraud threats across our global network of customers.

The data that we use is described in our Service Privacy Notice (see this section). We use data points such as a user's contact information (e.g., email address you use to place an order or create an account), device information (e.g., IP address, mobile/desktop properties), transaction information (e.g., information about orders you've purchased

on a website), and customer site communication information (e.g., listings you may have posted on a website). Our customers decide what data to send to us—this will depend on the particular service they offer and the type of fraud we are tasked with preventing.

We receive most data from our customers through an Application Programming Interface (API). However, we also collect data directly from users through standard tracking technologies (like our JavaScript code or SDK), which Sift customers embed on their websites and apps. You can learn more about the technologies we use to collect data in our Service Privacy

How does Sift prepare the data?

Next, we prepare the data by standardizing it so that it can be understood by our machine learning algorithms. Below, we explain some of the preparation techniques we employ: data normalization and feature engineering.

Data normalization

Fraudsters are constantly hunting for new methods to get around existing system controls and rules. One way to address this is by using *data normalization*—structuring and organizing data so that it is cohesive. Here is an example of one of our data normalization techniques:

Address normalization

An online business might use a rule to block a potentially fraudulent email address, e.g., *johndoe123@gmail.com*. In response, a fraudster may create a similar-looking email address, e.g., *johndoe124@gmail.com*, to circumvent the controls. When we spot minor variations to a known fraudulent email address, we can accurately match and then flag similar email addresses.

This technique is also used for physical addresses:

Variations of the same physical shipping addresses used by fraudsters

John Doe	John D	John C Doe
123, Smith Ln	Smith Lane, #123	#123, Smith Ln
San Francisco, CA	San Francisco, California	San Francisco, CA

Our address normalization technique assists in extracting the key substrings in the data field to identify *repeatable data patterns* and then applies probabilistic models to weigh the likelihood that two data inputs are correlated with each other.

Feature engineering

Feature engineering transforms raw data into structured, machine-processable formats. This allows us to create building blocks that are powerful fraud indicators. A type of feature engineering that is critical to building an effective machine learning system is called *feature extraction*. With feature extraction, we aim to derive the most useful characteristics from the raw data sent to Sift.

Fraud isn't static, and new patterns emerge daily. The more features we can extract, the more accurate our service is. Here are some of the types of features that we extract:

Types of features

While the features we use constantly evolve based on new and existing fraud patterns, and available data from our customers, they can be broadly classified into the following categories:

Event features	Properties of a user's most recent event *Examples: content posted, credit card type, billing ZIP code*
State features	Properties of a user's current state *Examples: country, time of day, browser type, IP address*
Temporal features	Number of actions associated with an event type or data point *Examples: number of transactions in the past hour, number of other user logins associated with the same IP address*
Graph features	How a user relates to others on the site and other sites *Example: number of times an IP address has been labeled as fraudulent across the Sift network*
Identity features	Features that are associated with the identity of a user *Example: name, email address, or phone number*
Behavioral features	Features that are associated with a type of behavior *Example: type of purchase*
Velocity features	Rate of an action *Example: number of failed transactions within a given time period, rate at which a user logs in from a particular country per day/week/month*

How does Sift use machine learning?

Once the data has been prepared, we can begin the analysis. In addition to using unsupervised machine learning models, Sift also uses *supervised machine learning*—meaning that the models we use learn from a training dataset before they are deployed for predicting.

Model training

To train our models, we use historical data and customer feedback to find correlations between "inputs" and "outputs." "Inputs" are user-generated events and associated metadata that we receive from our customers. An event is a discrete action that a user has taken on a website or app—for instance, completing a transaction, publishing a piece of content, or logging in to an online account.

"Outputs" refer to the Sift Score that we calculate using our machine learning models to reflect the probability that a given event is fraudulent. Using historical data from our customers allows our modeling to quickly determine fraud patterns that are unique to a specific customer's business. In addition, when a customer confirms to Sift whether an event was fraudulent or not, the feedback provides further training for our models.

Sift's predictive models

Once our models have been trained, we can begin using them to predict fraud. Sift uses two types of predictive models. Our "global" model is trained with a general understanding of fraud patterns across our network of customers, while our "custom" models are tailored to the data of a particular customer. This ensemble of models allows us to accurately score an individual event while taking a holistic approach when analyzing risk.

Global model

The Sift global model incorporates data from across all of our customers. Fraudulent users could have accounts on multiple websites, and spotting fraudulent behavior on one site helps to identify it on other sites as well. This means that we leverage the insights across our customer base to predict fraud and abuse.

Cohort model

Sometimes, Sift uses what we call "cohort models" or "threat clusters" - these are models that incorporate data from a group of customers which share a similar attribute (e.g., industry, geography). Certain cohorts of businesses experience similar fraud patterns - e.g. a crypto marketplace may have different types of fraud trends as compared to a quick service restaurant. By using a more targeted cross-customer model, we can more quickly and efficiently help our customers detect and prevent fraud and abuse.

Custom model

Each business is unique, and so we customize our approach and build predictive models to catch fraud that is specific to a customer's business. For example, if a Sift customer is a shoe company, we might recognize that buying size-15 shoes at a particular time of day is associated with fraudulent activity.

Sift's machine learning model stack

There's no world where one single algorithm can fix everything. Sift operates across multiple vectors of fraud (e.g., payment fraud, account takeover, fake accounts, scams and spam content), which each require a unique approach to identifying fraudulent behavior. Therefore, we build different model stacks for different vectors of fraud, applying various machine-learning algorithms that together deliver a "meta model." Combining these models together far outperforms any individual model. Below is a summary of the main types of models that we use: logistic regression, decision forests, deep learning (RNNs), XGBoost, Naive Bayes, and DBSCAN.

Logistic regression

Logistic regression models the probability of a particular state existing (e.g., whether a transaction is fraudulent or not) and is particularly useful in cases where only a limited set of information is available—such as a guest checkout experience while shopping.

Gradient boosted decision trees

Decision trees can model nonlinear interactions among different features. They easily incorporate domain knowledge and can be thought of as a large system of automatically learned rules.

RNNs

RNNs (Recurrent Neural Networks) are deep-learning algorithms that look at the time series of particular events (e.g., transactions performed, content published, or user-account logins) for all of a customer's users and then use that information to create a model that predicts which sequences of events are correlated with fraudulent or legitimate behavior.

XGBoost

XGBoost takes a very similar approach as gradient boosted decision trees. The main advantage of XGBoost is that it is computationally more efficient (both in terms of memory and speed). It enables us to solve fraud-detection problems in a fast and accurate way.

Naive Bayes

Naive Bayes is a very fast model and works well in cases of limited training data or labels. It provides us with the ability to easily learn online with every new piece of data that the customer has confirmed as associated with fraud, speeding up our ability to react to new fraud patterns and onboard new customers.

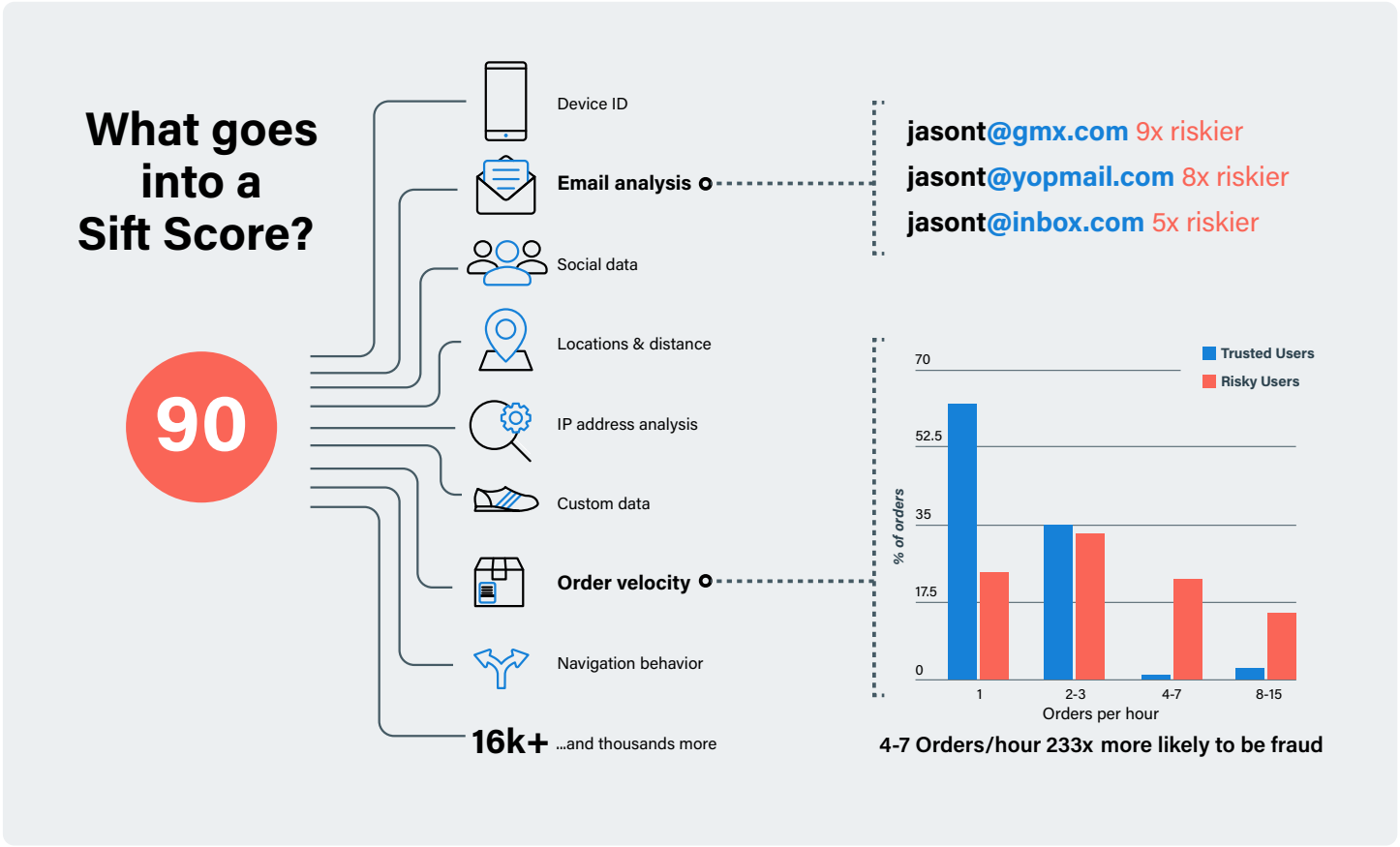
DBSCAN

DBSCAN or Density-Based Spatial Clustering of Applications with Noise is a model that finds core samples in regions of high density and expands clusters from them. It is a density-based algorithm that groups together points that are close to each other based on a density criterion. Points that are not part of any cluster are considered noise

Applying our models

Each time we analyze an event for a customer, we extract features related to that event and apply a dynamic combination of the machine-learning models described above. These features are then weighed against historical fraud we've seen both on the customer's site and within our global network of customers to determine a Sift Score. A Sift Score is a number between 0 and 100 that reflects the probability that *an event* is fraudulent. A score of 0 indicates a low likelihood of fraud, while a score of 100 indicates a high likelihood of fraud. We also represent this score as a probability between 0-1 in our systems (e.g., Sift Score of 50 or 0.5).

Importantly, there are no Sift Scores for you (or any user) because we don't score users; we score *user interactions on a specific website for a specific type of fraud*. We calculate the likelihood of whether actions you have taken on a Sift customer site are associated with specific types of fraud (e.g., payment fraud, account abuse). However, these interactions do not add up into a single Sift Score about you.



Ensuring accuracy

An integral part of Sift's approach is our real-time machine learning that helps ensure the accuracy of Sift Scores. This approach includes real-time learning, proactive scoring, and rescoring.

Real-time learning

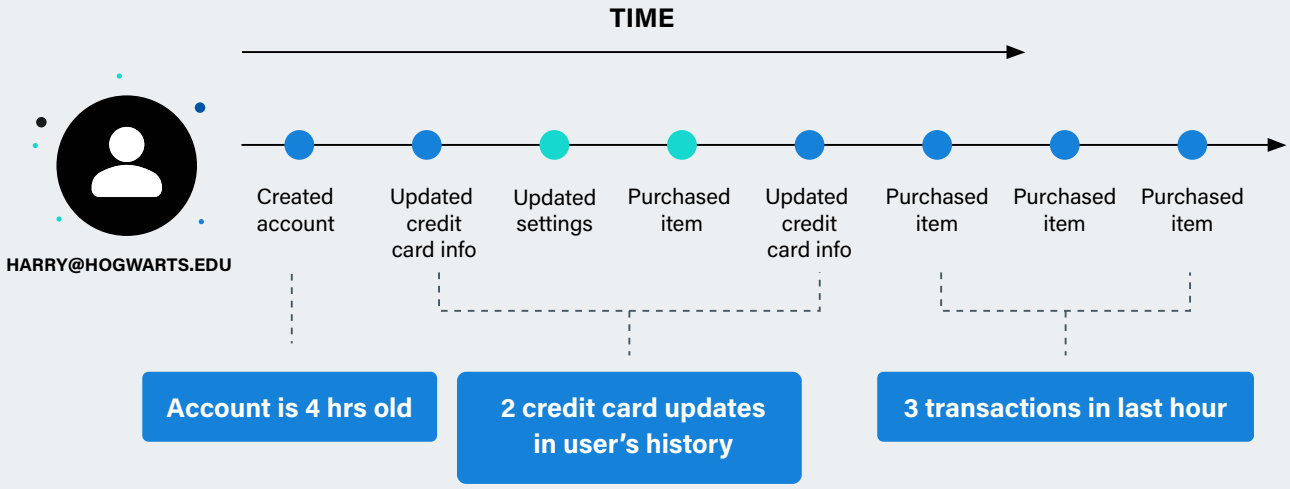
We use *real-time learning*, which means that new information is pushed in *real time* to update our cross-customer models. For example, a fraudster is likely to have tested stolen cards on several different sites. With Sift, where such card-testing behavior is identified, the fraudster's information can inform different customers' models.

Proactive scoring

Sift applies *proactive scoring*, which means that we calculate the Sift Score every time a user interacts on a customer's site, based on the knowledge of actions and patterns within the customer's platform and our network.

Rescoring

We also *rescore* events as we learn new information, e.g., if a new action is taken on the website or we receive new customer feedback. If a customer confirms that an event was fraudulent or legitimate, that feedback also gets looped back into our system and feeds into our models. This is part of the reason why our system, and all computations that we do, are constantly evolving.



How our customers use Sift

Sift works with businesses across a range of industries and sectors, from retailers and online travel agencies to on-demand apps and marketplaces, providing services that help customers manage every aspect of fraud prevention. Through our unified SaaS platform, customers can build and automate workflows, access insights and analysis, and make data-driven decisions to protect their business.

Our customers decide how to use the Sift Scores that we provide them. Customers can build and manage business logic around Sift Scores and set up automatic workflows for particular events on their platforms. For example, depending on how they have configured our platform, customers can use the Sift Score to automatically accept or block a particular event (such as a transaction), initiate a two-factor authentication request, or flag an event for manual review by their fraud analyst team. The customer sets and controls the thresholds for these automated actions, and the thresholds will depend on factors like the customer's industry, business objectives, and risk tolerance.

Beyond Sift Scores, we may also provide aggregated reporting and insights derived from data across our entire customer network. These insights, which include valuable metrics—such as the frequency of blocks associated with a particular identifier (e.g., an email address)—offer customers broader context to enhance manual decision-making and gain a wider perspective on potential risks. In addition, we may include synthesized summaries and risk assessments based on a combination of structured data analysis and automated content generation technologies All of these features are accessible within the Sift Console, a dashboard that enables customers to create, monitor, and adjust their workflows.

Ultimately, this type of real-time fraud and abuse analysis benefits you as an end user because it helps ensure the safety and security of services that you use online. It ensures that any given event relates to a genuine user and that the user is who they say they are. For example, if your credentials or payment information has been stolen and a fraudster tries to use it on the site of a Sift customer, we aim to stop that fraudulent use.

If you want to learn more about how a business you interact with uses Sift, or want to contest a decision that's been made by a business that uses Sift, we encourage you to reach out to that online business to assist you. If they confirm that a transaction was legitimate, then the feedback will be provided to our models.

We hope this provides you a better understanding of how Sift works. To learn more about how we process your data, please see our Service Privacy Notice.