



API Usage Analytics Service

Client Contact:	Jack Kelly (jack@fullcontact.com) Andy Pai (andy@fullcontact.com)
Team Size:	3 - 4
Location:	Tabor Center, Downtown Denver
Skillset:	RESTful APIs, Statistical Analysis, Web Design
IP Assignment	To client (FullContact)

About Us:

FullContact is a privacy-safe Identity Resolution company building trust between people and brands. We deliver the capabilities needed to create tailored customer experiences by unifying data and applying insights in the moments that matter. We got our start in TechStars Boulder 2011, and are now an international team headquartered in LoDo (Lower Downtown Denver) with a satellite office in Kochi, India. Our engineering department in Denver has over a dozen members focused on APIs, infrastructure, and data processing.

The Problem:

We have over one thousand customer accounts with access to our APIs, making tens of millions of calls daily. Not all of our customers utilize the same amount of API access - some might make a few dozen calls in a day, some might make millions. Our sales team does their best to set customers up with the right size of API quota, but sometimes they miss the mark or customer needs change. We would like to be able to better predict when a customer's usage of our APIs will be noticeably above or below their current quota.

We currently have usage data for all of our customers flowing into an internal metrics-tracking system. We need a new system which can pull that data out and provide

predictive analytics to identify customer accounts which are at-risk of over- or under-utilizing their access quotas.

The Project:

You will design a web application to pull usage data, apply predictive analytics, and expose the resulting predictions to internal users at FullContact. You will also work with our back office team to productionize and deploy your application into our cloud infrastructure so that it can be continuously available. Your application will need to operate on over a thousand data streams, with tens of millions of data points per day.

The service must be able to provide two key predictions based on usage data for our customers:

1. For a given customer, does their historical usage indicate their current quota is not at an appropriate level? Whether too high or too low, we want to identify such cases quickly and get them on an appropriate plan.
2. Given aggregate data over multiple customers, or even all accounts, are there any macro-level usage trends? Insights at this level will allow us to adapt to larger infrastructure needs before they are necessary.

You will also need to expose the results of these predictions in convenient visual format, suitable for non-engineers to interact with. We want some kind of web UI that allows users to explore the various predictions and predictive models. Where possible, the information presented to the user should be calculated in real time, or else scheduled for periodic updates so that it remains relevant.

Depending on the progress made by the team during the project, we have additional features that might be worked on:

1. More advanced predictive models, including any supporting infrastructure if those models require periodic training on larger amounts of historical data.
2. Predictive analysis of our cloud infrastructure, to help us discover unwanted usage patterns driving one of our largest business expenses.
3. Multi-stream correlation analysis, to assist us in identifying fraudulent or abusive accounts, particular abuses of our free-tier plan.

You will have the support of multiple engineers with many years of industry experience. You will be issued a company laptop configured for secure access to our internal systems. We cover reasonable transportation expenses to and from our office, however, availability of the physical office location may be impacted by social distancing

measures at the request of health officials. We will work with you to set up an appropriate remote work solution as necessary.

What Success Looks Like

- We can predict and identify accounts whose usage patterns do not match their current access quota
- Such predictions are available in a web UI suitable for use by non-technical employees at FullContact
- The delivered system keeps its predictions up to date based on current usage data
- The delivered system is stable and documented well enough that our engineers can take confident ownership of the codebase and deploy pipelines
- Because the system is expected to integrate tightly with FullContact's internal infrastructure, we will require a contact to assign ownership to FullContact.

Skills Required

- HTTP interaction
- RESTful API design and implementation
- A programming language suitable for building the requested system
- Web UI design and implementation
- Git for version control

Summer Internship:

Many of our full time engineering staff are graduates of CSM and participated in field session projects that we hosted.

A limited number of paid summer internships may be available upon course completion.