



FullContact

LOCATION

18th & Blake
(LoDo)

Near many great restaurants, breweries, and bars.

Right next to Union Station for easy travel.

IDEAL TEAM SIZE

2-4

Bot or Not

About Us:

FullContact is solving the world's contact information problem. We're a 2011 TechStars Boulder graduate that provides APIs and applications that allow businesses and consumers to keep their contact information up to date automatically. We're a team of 200+ people located across the globe in Denver, Texas, Latvia, and India, headquartered in LoDo (downtown Denver) down the street from Union Station.

Background:

FullContact has a stated mission to become the premier contact management system of record for individuals and businesses. This is done by keeping all their contact information in one place and automatically up to date and making this contact information available everywhere they need it. We'd like to be able to better and more quickly identify bots before those sign ups can affect our business metrics and abuse our free features.

Project:

As we've grown as a software as a service (SaaS) company, we've seen an increase in automated (bot) sign ups that are fraudulent in nature. These bot sign ups are particularly troublesome as they affect internal business metrics, abuse free features and increase cost / noise when trying to segment customers for marketing.

We've done initial research to automate the identification of bot sign ups via a number of features such as (not limited to); email address, sign up date, source ip and prominence of email address in our identity graph. We intend to utilize data analysis and machine learning techniques to determine whether or not an account sign up was indeed a "bot or not".

Of course, the fun doesn't stop there.

A good portion of our API customers are SaaS companies (Just like us). Every SaaS company has a customer database that is utilized for customer relationship purposes. An intriguing aspect of this problem is to approach "bot or not" in a generalized form so our customers could deduce a higher identity match rate, blacklist "bots" from their own systems, and improve market segmentation.

Solution:

A successful project will be able to:

- Train or derive an automated system that can identify bot accounts within FullContact's entire customer database within a reasonable time frame (minutes).
- Train an automated system that can identify bot accounts from new sign ups in our customer database in a reasonable time frame (real-time, hourly).
- Build an interface to manually review questionable account(s).
- **Stretch Goal** - Generalize "bot or not" intelligence to an upcoming api ("subscriptions") that allows customers to store data / receive updates from us.
 - Upload an entire customer database to our subscriptions api.
 - Derive / Apply a "bot or not" model on the entire customer database.
 - Apply a "bot or not" model to new sign ups in the customer database.
 - Send updates to customer contacts with a "bot" likelihood.

A successful project will likely be rolled into production.

Summer Internship:

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

Required Skills:

- Willing and eager to solve hard problems
- Enjoy working with algorithms
- Interested in learning data analysis technologies.

Useful Skills:

- Machine learning (artificial intelligence) algorithms
- Java