Active-Learning driven Testing of RESTful web APIs

A. Giuliano Mirabella
amirabella@us.es

SCORE Lab, I3US Institute
Universidad de Sevilla

Contents

Contents

Context

REST APIs

- Context
  - REST APIs
  - Dependencies
  - ML-based prediction of requests validity
- Problem
- Approach
  - Example
- Evaluation
  - PI1
  - PI2
- Future Work

Context

Dependencies

`videoDuration`	**`string`** The **`videoDuration`** parameter filters video search results based on their duration. If you specify a value for this parameter, you must also set the **`type`** parameter's value to **`video`**.

`custom` object

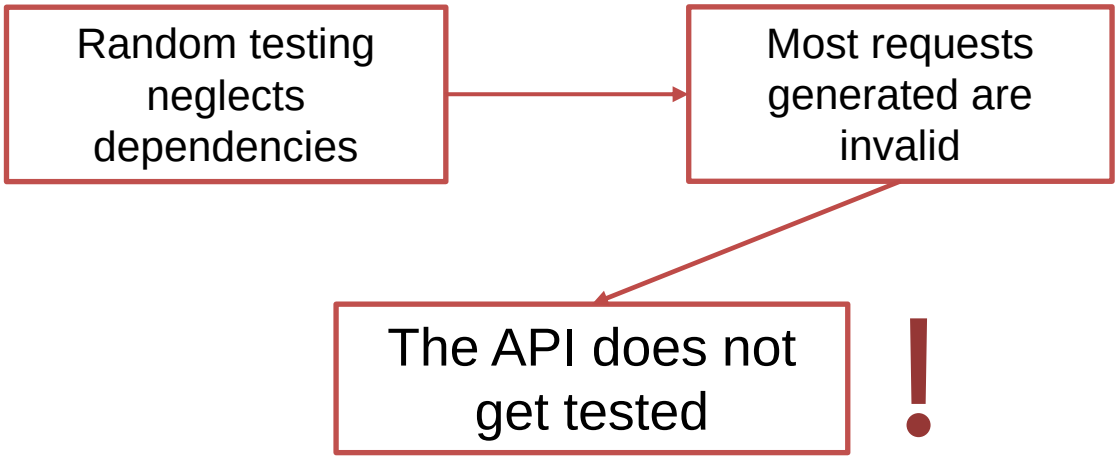
The custom amount to apply to an invoice. If you include a label, you must include a custom amount.

browse	Find venues within a given area. Unlike the `checkin intent`, browse searches an entire region instead of only finding venues closest to a point. A region to search can be defined by including either the `ll` and `radius` parameters, or the `sw` and `ne`. The region will be circular if you include the ll and radius parameters, or a bounding box if you include the sw and ne parameters.

- Context
  - REST APIs
  - Dependencies
  - ML-based prediction of requests validity
- Problem
- Approach
  - Example
- Evaluation
  - PI1
  - PI2
- Future Work

Context

Dependencies



98% faulty test cases in YouTube

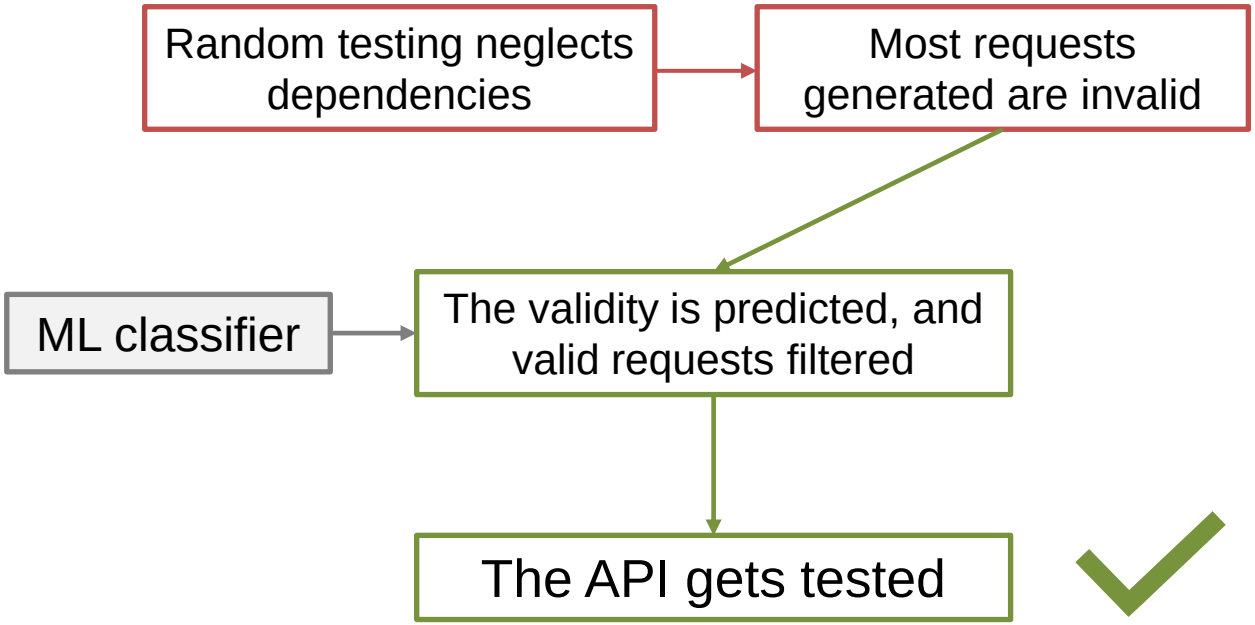
Context

Dependencies

Automated test case generation
for
RESTful APIs with unspecified dependencies

Context

ML-based prediction of requests validity



99% valid test cases in YouTube

Context

ML-based prediction of requests validity

Deep Learning-Based Prediction of Test Input Validity for RESTful APIs

A. Giuliano Mirabella
SCORE Lab, I3US Institute
Universidad de Sevilla
Seville, Spain
amirabella@us.es

Alberto Martin-Lopez
SCORE Lab, I3US Institute
Universidad de Sevilla
Seville, Spain
alberto.martin@us.es

Sergio Segura
SCORE Lab, I3US Institute
Universidad de Sevilla
Seville, Spain
sergiosegura@us.es

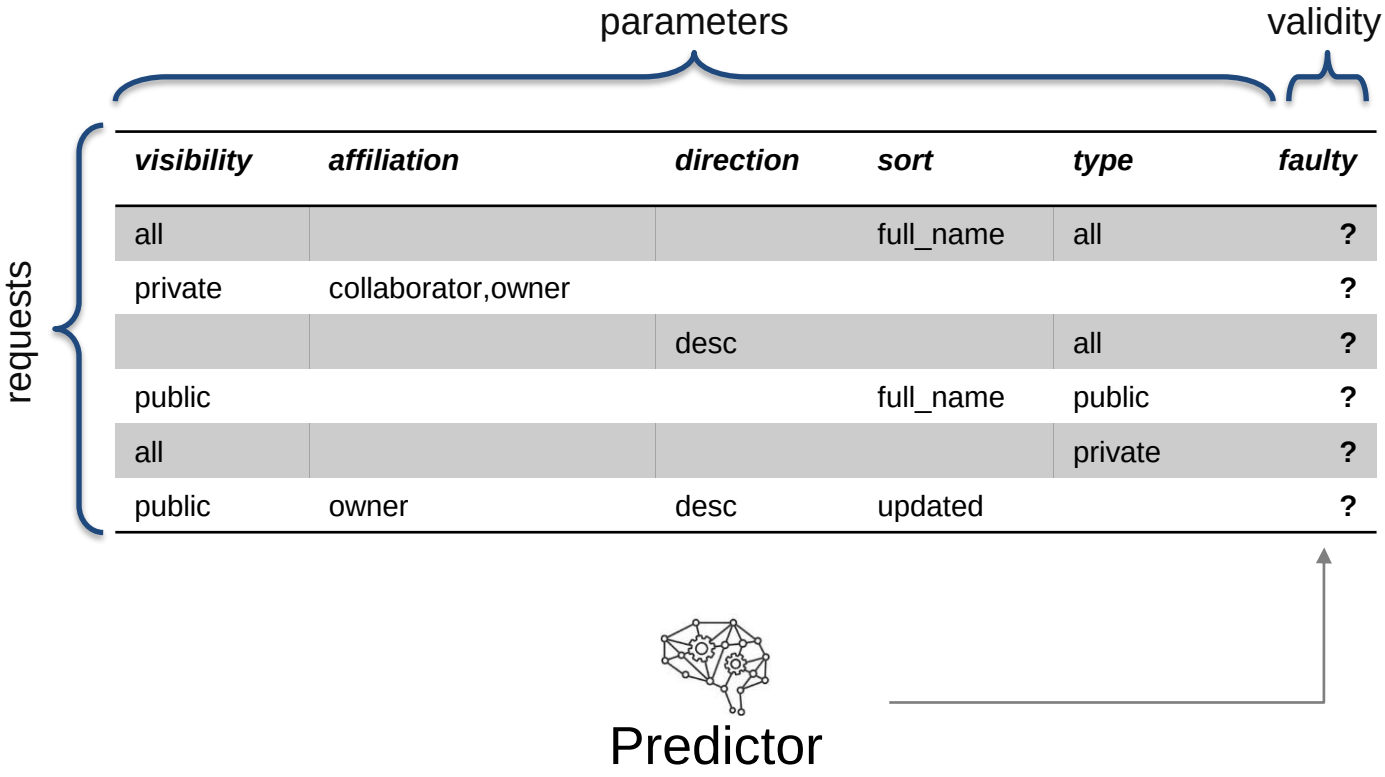
Luis Valencia-Cabrera
SCORE Lab, I3US Institute
Universidad de Sevilla
Seville, Spain
lvalencia@us.es

Antonio Ruiz-Cortés
SCORE Lab, I3US Institute
Universidad de Sevilla
Seville, Spain
aruiz@us.es

International Workshop on Testing for Deep Learning and Deep Learning for Testing, 2021.

Context

ML-based prediction of requests validity



Context

ML-based prediction of requests validity

parameters

validity

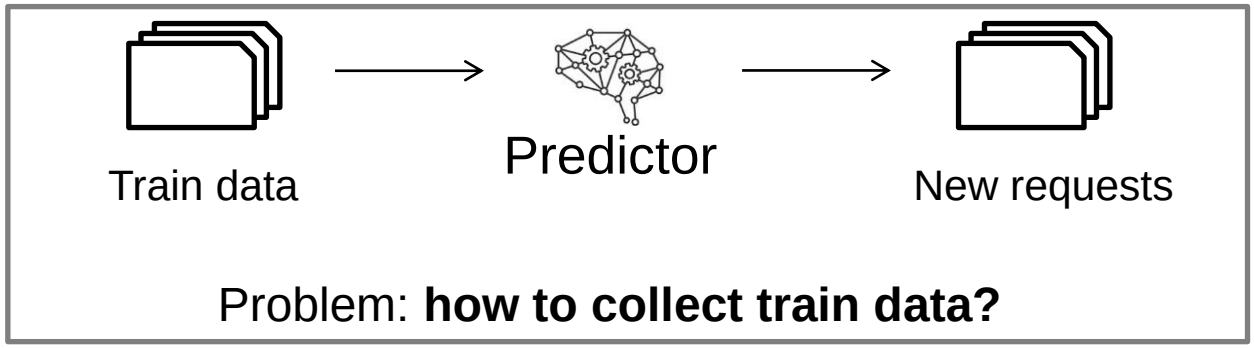
requests

visibility	*affiliation*	*direction*	*sort*	*type*	*faulty*
all			full_name	all	**True**
private	collaborator,owner				**False**
		desc		all	**False**
public			full_name	public	**True**
all				private	**True**
public	owner	desc	updated		**False**

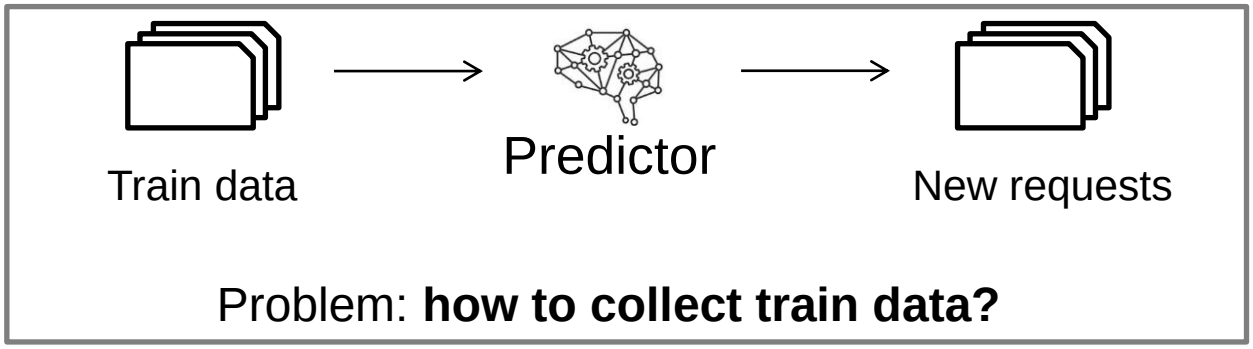
Predictor

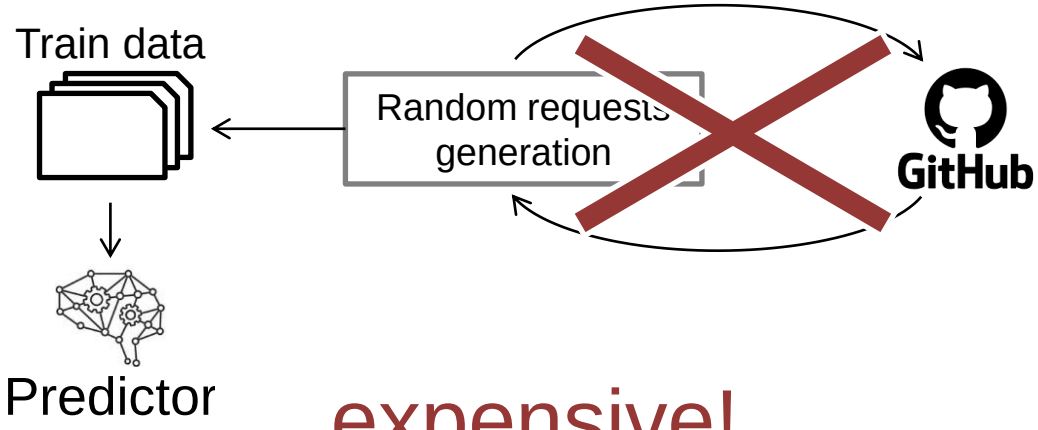
- Context
  - REST APIs
  - Dependencies
  - ML-based prediction of requests validity
- Problem
- Approach
  - Example
- Evaluation
  - PI1
  - PI2
- Future Work

Problem



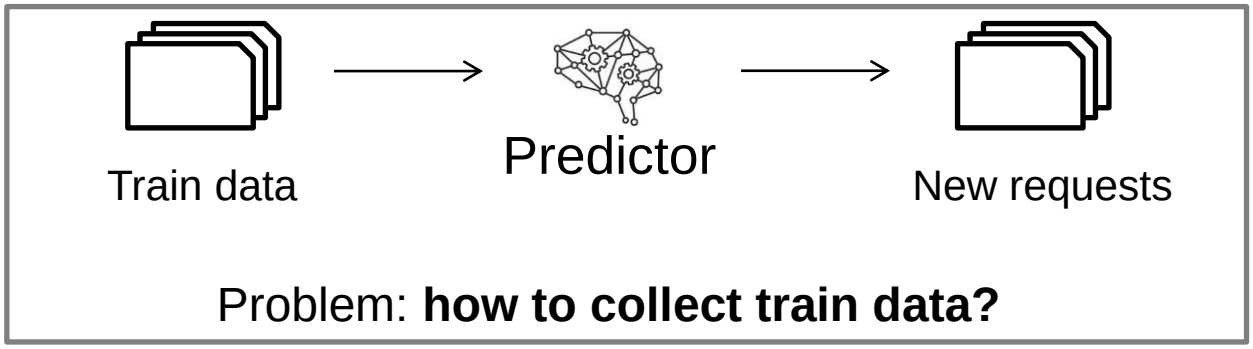
Problem

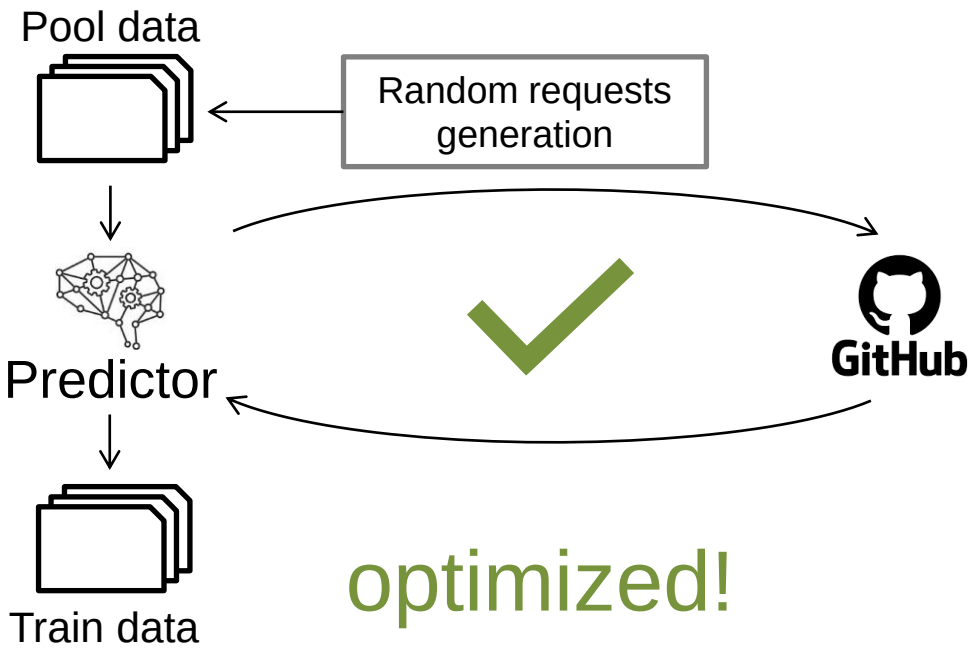




expensive!

Approach





Approach

Example

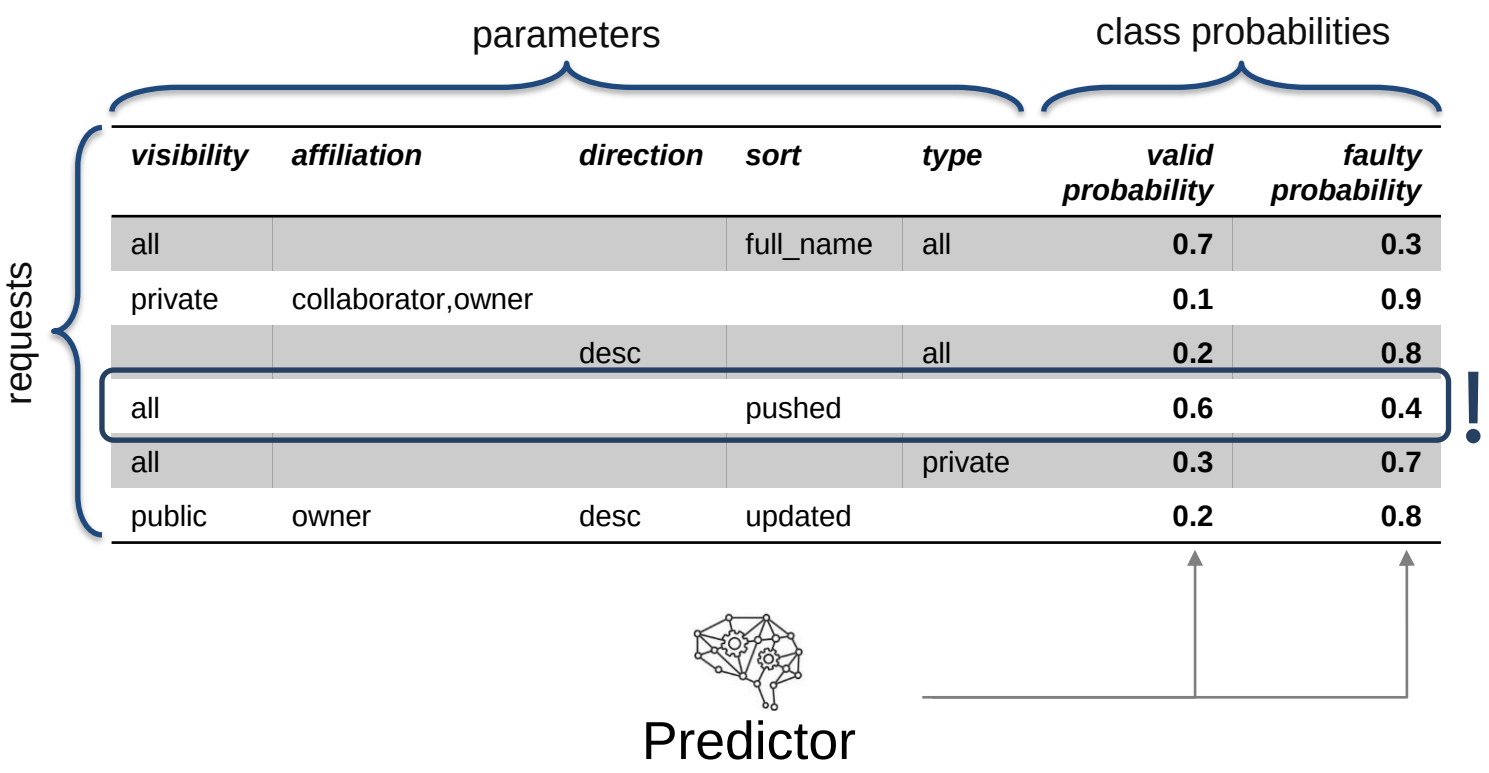
parameters

requests

visibility	*affiliation*	*direction*	*sort*	*type*
all			full_name	all
private	collaborator,owner			
		desc		all
all			pushed	
all				private
public	owner	desc	updated	

Approach

Example



Approach

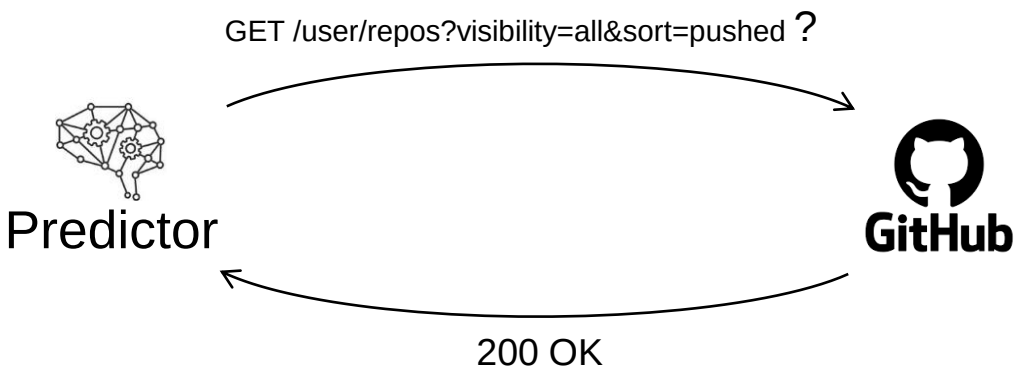
Example

parameters | class probabilities

requests

visibility	*affiliation*	*direction*	*sort*	*type*	*valid probability*	*faulty probability*
all			full_name	all	**0.7**	**0.3**
private	collaborator,owner				**0.1**	**0.9**
		desc		all	**0.2**	**0.8**
all			pushed		**0.6**	**0.4**
all				private	**0.3**	**0.7**
public	owner	desc	updated		**0.2**	**0.8**

!



- Context
  - REST APIs
  - Dependencies
  - ML-based prediction of requests validity
- Problem
- Approach
  - Example
- Evaluation
  - PI1
  - PI2
- Future Work

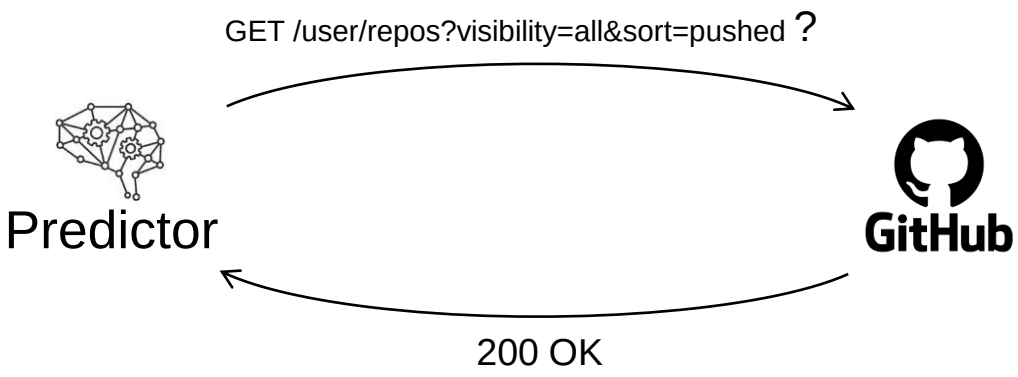
Approach

Example

parameters | class probabilities

requests

visibility	*affiliation*	*direction*	*sort*	*type*	*valid probability*	*faulty probability*
all			full_name	all	**0.7**	**0.3**
private	collaborator,owner				**0.1**	**0.9**
		desc		all	**0.2**	**0.8**
all			pushed		**1**	**0**
all				private	**0.3**	**0.7**
public	owner	desc	updated		**0.2**	**0.8**



- Context
  - REST APIs
  - Dependencies
  - ML-based prediction of requests validity
- Problem
- Approach
  - Example
- Evaluation
  - PI1
  - PI2
- Future Work

Evaluation

RQ1: How effective is this approach in generating valid requests compared to a random testing baseline?

RQ2: What is the fault-detection capability of this approach compared to a random testing baseline?

Evaluation

RQ1

RQ1: How effective is this approach in generating valid requests compared to a random testing baseline?

Service	Valid requests (%)	
	Random testing	AL-driven testing
GitHub	62.1	98.7
Stripe-CC	13.7	97.4
Stripe-CP	55.8	99.3
Yelp	44.2	83.7
YouTube-GCT	13.4	85.0
YouTube-VID	25.3	99.0
YouTube-SRC	3.0	89.2
Mean	**31.3**	**93.2**

*The ratio of valid requests obtained is **93%**, three times more than random testing baseline (31%).*

Evaluation

RQ2

RQ2: What is the fault-detection capability of this approach compared to a random testing baseline?

Service	Specification Faults	
	Random testing	AL-driven testing
GitHub	313	1159
Stripe-CC	0	0
Stripe-CP	104	284
Yelp	29	60
YouTube-GCT	0	262
YouTube-VID	45	442
YouTube-SRC	30	742
Mean	**74**	**421**

*The number of faults detected is **421**, more than five times the faults detected with random testing baseline (74).*

Future Work

Human-readable dependencies inference

IF type=='private' THEN NOT visibility=='all' ;

~~IF type=='public' THEN visibility ;~~

IF type=='public' THEN NOT visibility=='private' ;

} ZeroOrOne(type, visibility)

YouTube

`videoDuration` — `string`

The `videoDuration` parameter filters video search results based on their duration. If you specify a value for this parameter, you must also set the `type` parameter's value to `video`.

`custom` object

The custom amount to apply to an invoice. If you include a label, you must include a custom amount.

PayPal

FOURSQUARE

browse

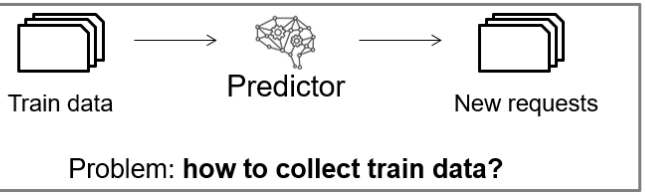
Find venues within a given area. Unlike the `checkin intent`, browse searches an entire region instead of only finding venues closest to a point. A region to search can be defined by including either the `ll` and `radius` parameters, or the `sw` and `ne`. The region will be circular if you include the ll and radius parameters, or a bounding box if you include the sw and ne parameters.

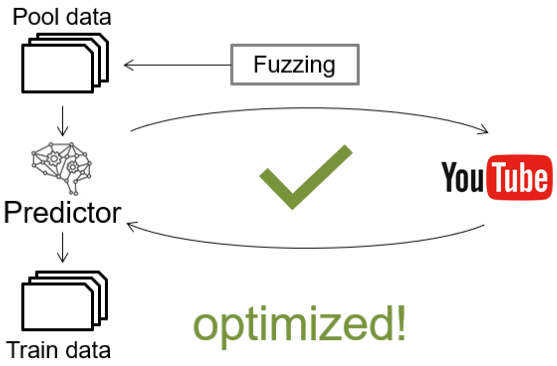
parameters — validity

requests

visibility	*affiliation*	*direction*	*sort*	*type*	*faulty*
all			full_name	all	**True**
private	collaborator,owner				**False**
		desc		all	**False**
public			full_name	public	**True**
all				private	**True**
public	owner	desc	updated		**False**

Predictor





PI2: What is the fault-detection capability of this approach compared to a fuzzing baseline?

Service	OAS Faults	
	Random testing	AL-driven testing
GitHub	313	1159
Stripe-CC	0	0
Stripe-CP	104	284
Yelp	29	60
YouTube-GCT	0	262
YouTube-GV	45	442
YouTube-S	30	742
Mean	**74**	**421**

*The number of faults detected is **421**, more than five times the faults detected with random testing techniques (74).*

Active-Learning driven Testing of RESTful web APIs

A. Giuliano Mirabella
amirabella@us.es

SCORE Lab, I3US Institute
Universidad de Sevilla