

craigslist

create a posting

my account

search craigslist

"the best AUD" team

team members

kai-duan chang
li-ci chuang
su-tien lee

yen-tsz huang
yi-hsuan hsu
meghan harris

final group project

12/8/2021

"the best AUD"

team members



kai



lilian



yen



meghan



sandy



su-tien

craigslist

create a posting

my account

overview

search overview

- ★ background
- ★ business analysis
- ★ data analysis
- ★ validation
- ★ conclusion

background



what is Craigslist?

- ★ one of the most popular classified advertisements website used for viewing and posting advertisements
- ★ allows people to easily find, buy, or sell anything
- ★ typically used regionally

craigslist

create a posting

my account

background

west lafayette, in / los angeles, ca

community

activities

artists

childcare

classes

events

general

groups

local news

lost+found

missed

connections

musicians

pets

politics

rants & raves

rideshare

volunteers

craigslist

create a posting

my account

background

west lafayette, in / los angeles, ca

community

activities

artists

childcare

classes

events

general

groups

local news

lost+found

missed

connections

musicians

pets

politics

rants & raves

rideshare

volunteers

craigslist

create a posting

my account

background

west lafayette, in / los angeles, ca

community

activities

artists

childcare

classes

events

general

groups

local news

lost+found

missed

connections

musicians

pets

politics

rants & raves

rideshare

volunteers

lost & found

search lost & found

- ☐ search titles only
- ☐ has image
- ☐ posted today
- ☐ bundle duplicates
- ☐ include nearby areas

MILES FROM LOCATION

miles

from zip

[use map...](#)

[reset](#)

[update search](#)



Nov 13

LOST SD/MEMORY CARD (LAFAYETTE)



Nov 9

lost black kitten (St Joe addition off teal rd) [pic](#)



Nov 4

*****Lost abandoned cat***** (Lafayette) [pic](#)



Oct 22

key fob found mazda 2014 red (Lafayette)

lost & found

search lost & found

- ☐ search titles only
- ☐ has image
- ☐ posted today
- ☐ bundle duplicates
- ☐ include nearby areas

MILES FROM LOCATION

miles

from zip

use map...

reset

update search



Nov 13

LOST SD/MEMORY CARD (LAFAYETTE)



Nov 9

lost black kitten (St Joe addition off teal rd) pic



Nov 4

Lost abandoned cat (Lafayette) pic



Oct 22

key fob found mazda 2014 red (Lafayette)

lost black kitten (St Joe 2 addition off teal rd)

image 1 of 2



7 to 9 month black kitten has shown up.

lost black kitten (St Joe 2 addition off teal rd)

post title

image 1 of 2



7 to 9 month black kitten has shown up.

post content

craigslist's lost+found section has
no option to automatically label and
sort posts by lost or found.

business analysis



project objectives:

1

provide users with the ability to search for all ads via lost or found.

2

provide users with quick links to specific posts based on most popular topics.

3

identify images based on keyword.

provide users with the
ability to search for all ads
via lost or found.

how do we determine if a
post is lost or found?

objective #1 outline

- ★ **text mine** based on **post content** instead of post title.

- ★ some titles are labeled the opposite of what they are.

- example: someone found a cat and posted “lost cat” in terms of user to find

- ★ **identify key words** associated with “lost” and “found” that could be used in post instead.

ultimate goal for objective #1

VISUAL EXAMPLE

lost & found

search lost & found

- ☐ search titles only
- ☐ has image
- ☐ posted today
- ☐ bundle duplicates
- ☐ include nearby areas

MILES FROM LOCATION

miles from zip

use map

SORT BY

- ☐ lost
- ☐ found

reset

update search

★ Nov 13 LOS

★ Nov 9 lost

★ Nov 4 ***

★ Oct 22 key

provide users with quick links to
specific posts based on
most popular topics

which topics within
lost+found are the
most popular?

objective #2 question

objective #2 outline

★ **identify **key topics** that appear within lost+found .**

★ users who use this page are looking for **something specific**.

Linking the most popular topics allows a user to quickly identify if craigslist has what they are looking for, thereby **improving user experience**.

ultimate goal for objective #2

VISUAL EXAMPLE

lost & found

search lost & found

- ☐ search titles only
- ☐ has image
- ☐ posted today
- ☐ bundle duplicates
- ☐ include nearby areas

MILES FROM LOCATION

miles from zip

use map

CATEGORY

- ☐ topic #1
- ☐ topic #2
- ☐ topic #3

reset

update search

★ Nov 13 LOS

★ Nov 9 lost

★ Nov 4 ***

★ Oct 22 key

identify images
based on keyword

objective #3

what keywords have the
most images and can be
further distinguished from
each other?

objective #3 outline

★ **identify **top topic keywords** that have the most images to further distinguish posts from each other**

★ users have to sift through a lot of posts to find what they might be looking for.

To improve user-experience, include more keywords about the posts through **image recognition**.

ultimate goal for objective #3

VISUAL EXAMPLE

lost & found

search lost & found

- ☐ search titles only
- ☐ has image
- ☐ posted today
- ☐ bundle duplicates
- ☐ include nearby areas

MILES FROM LOCATION

miles

from zip

use map

- ▼ topic #1

☐ image keyword #1

☐ image keyword #2

☐ all
- ▼ topic #2

☐ image keyword #1

☐ image keyword #2

☐ all
- ▼ topic #3

 Nov 13

LOS

 Nov 9

lost

 Nov 4

 Oct 22

key

data analysis



data analysis

web scraping

craigslist

create a posting

my account

search craigslist

event calendar

M	T	W	T	F	S	S
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31	1	2

community

activities

artists

childcare

classes

events

general

groups

local news

lost+found

missed

connections

musicians

pets

politics

rants & raves

rideshare

volunteers

data analysis

web scraping

CL 90036 + 120 miles ▾ > c

Lost & Found

☐ search titles only
☐ has image
☐ posted today
☐ bundle duplicates

MILES FROM LOCATION

120 90036 ↻

use map...

▼ lost or found?
☐ lost
☐ found

reset update search

CL 47906 + 250 miles ▾ > c

Lost & Found

☐ search titles only
☐ has image
☐ posted today
☐ bundle duplicates

MILES FROM LOCATION

250 47906 ↻

use map...

▼ lost or found?
☐ lost
☐ found

reset update search

data analysis

los angeles
595

web scraping

CL 90036 + 120 miles ▾ > c

Lost & Found

☐ search titles only
☐ has image
☐ posted today
☐ bundle duplicates

MILES FROM LOCATION

120 90036 ↻

use map...

▼ lost or found?
☐ lost
☐ found

reset update search

CL 47906 + 250 miles ▾ > c

Lost & Found

☐ search titles only
☐ has image
☐ posted today
☐ bundle duplicates

MILES FROM LOCATION

250 47906 ↻

use map...

▼ lost or found?
☐ lost
☐ found

reset update search

west lafayette
490

data analysis

web scraping

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1		Title	Content												
2		0 key fob dc	found key fob found by shell gas station of off green bush												
3		1 lost black	7 to 9 month black kitten has shown up.												
4		2 LOST SD	Hi, I know this is a long shot but I lost my memory card in Lafayette near the army recruiter office/mccallisters area. Maybe even at he												
5		3 *** Lost &	I rehomed 2 cats last year to a family as barn cats, they were former ferals that lived at my jobsite, they were not safe at my work, and												
6		4 Found He	Found on Saturday on 86th near Brownsburg a heavy case that must have fallen off someone? truck. Please describe contents and we												
			Loki has been missing since 10/18. He is a 65 lb black lab mix with a white spot on his chest and black 5 LOST DC birthmark on his tongue.												

data analysis

web scraping

	A	B	C	D	E	F	G	H	I	J
	Y									
	0: found									
1	1: lost	D	Title	Content						
2	0		0 key fob down	found key fob found by shell gas station of off green bush						
3	0		1 lost black kit	7 to 9 month black kitten has shown up.						
4	1		2 LOST SD/MHi,	I know this is a long shot but I lost my memory card in Lafayette near the army recruiter						
5	1		3 *** Lost aba	I rehomed 2 cats last year to a family as barn cats, they were former ferals that lived at my j						
6	0		4 Found Heavy	Found on Saturday on 86th near Brownsburg a heavy case that must have fallen off someone						
7	1		5 LOST DOG	Loki has been missing since 10/18. He is a 65 lb black lab mix with a white spot on his chest						
8	1		6 West side lur	I am looking for my friend Maria, on the west side. We would get together for lunch sometime						
9	1		7 lost Pomeran	Has three masses on him. (One on cheek, one on chin/neck area, one on lower abdomen) 08/08/21						
10	1		8 lost blue hear	Lost single small blue hearing aid. In/around my path from Midtown up north Keystone ave						
11	1		9 Lost Diamon	REWARD ** Art Deco setting with pear-shaped orange sapphire with 2 emerald cut diamonds						
12	1		10 Lost white ar	Missing male cat on 11/9/21, mostly white with orange spots on his head and tail. He is about						

data analysis

classification model

- Naïve Bayes model: 74.92%
- DT model: 70.57%
- Random Forest model: 66.22%
- SVM model: 77.93%

data analysis

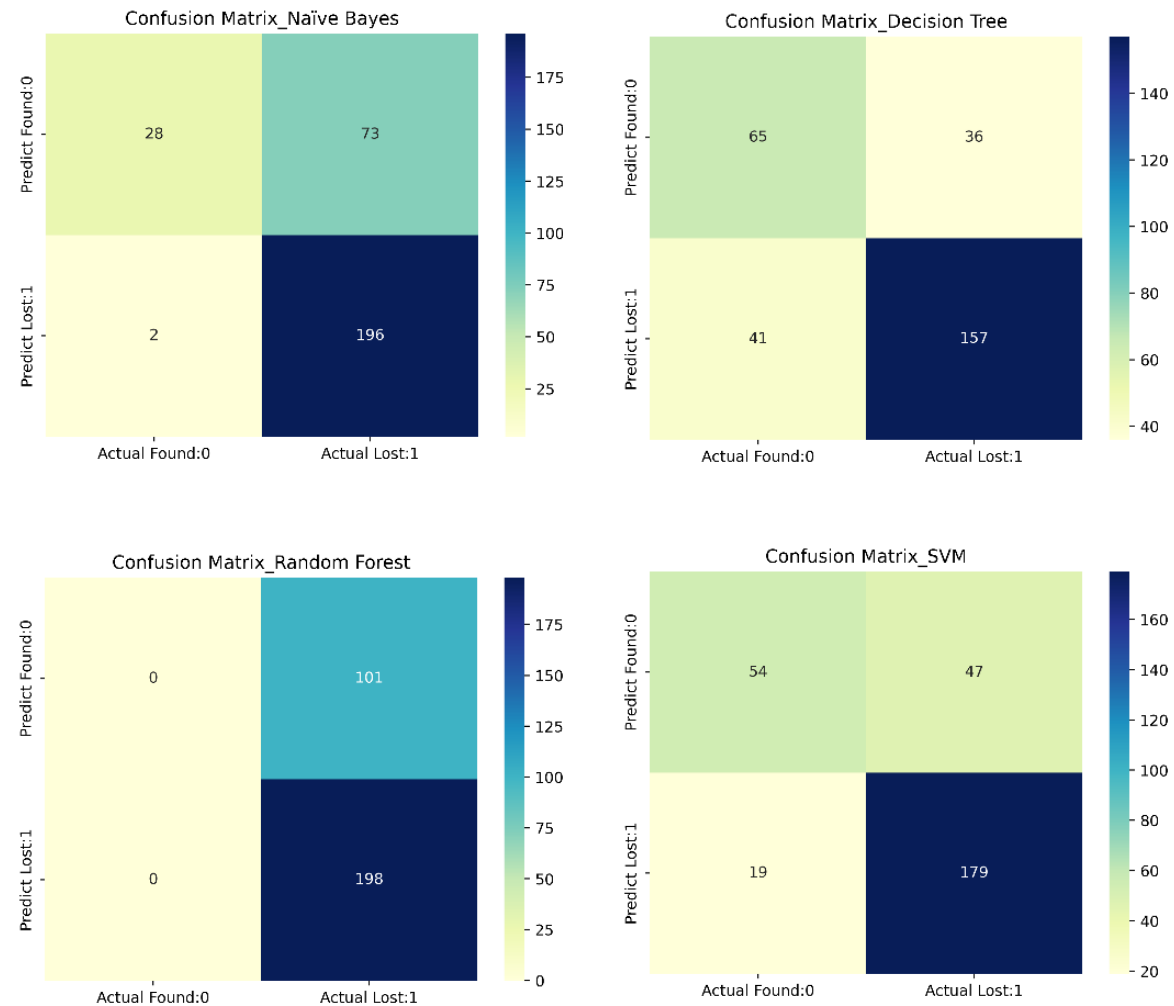
classification model

- Naïve Bayes model: 74.92%
- DT model: 70.57%
- Random Forest model: 66.22%
- SVM model: 77.93%



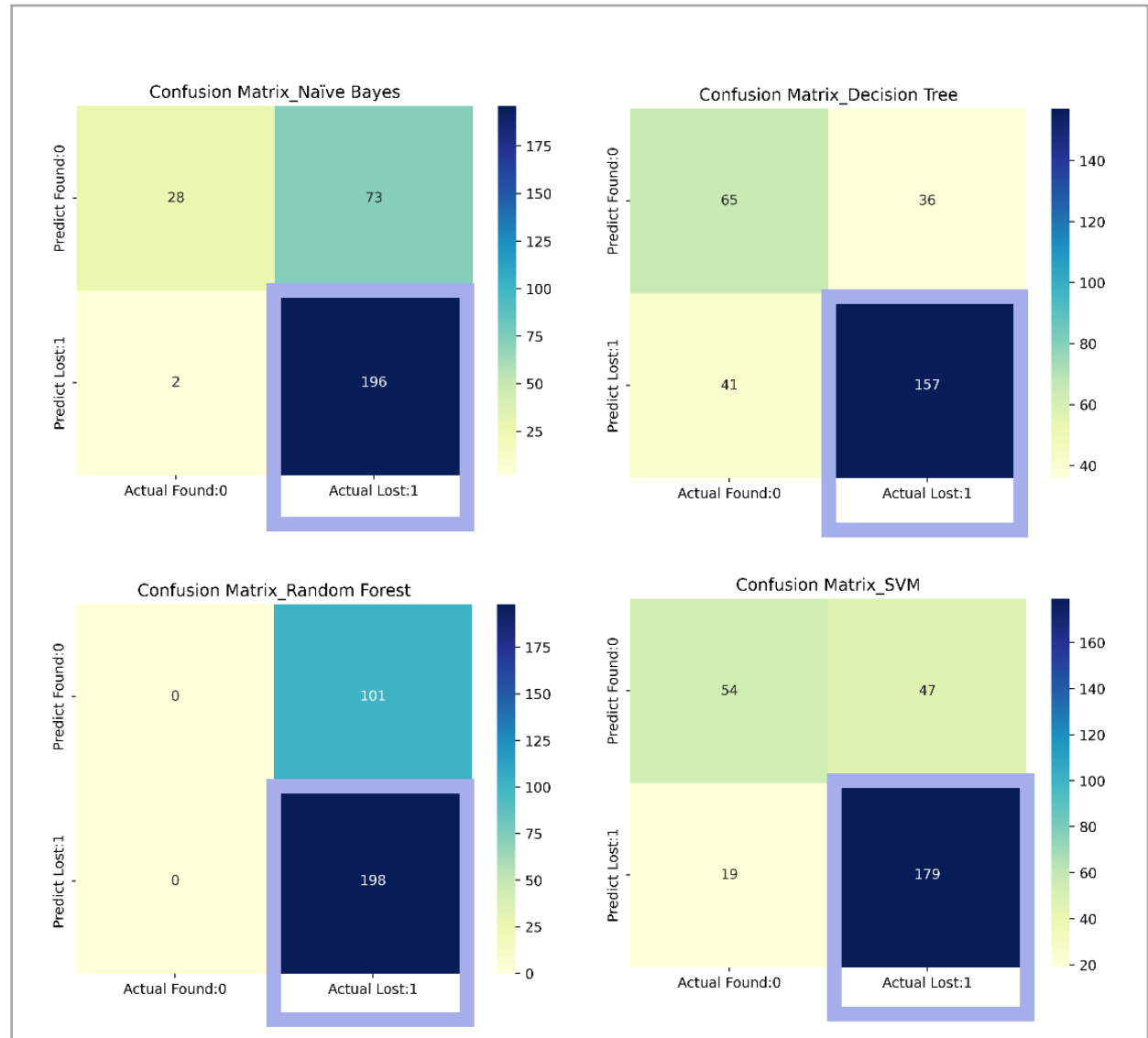
data analysis

classification model



data analysis

classification model



data analysis

LDA topic model

Topic 1:

cat / old / dog / female / name

Topic 2:

near / cat / back / dog / one

Topic 3:

key / dog / wallet / around / street

data analysis

VGG image recognition



- lakeland_terrier
- airedale
- toy_poodle
- miniature_poodle
- teddy

data analysis

VGG image recognition



- egyptian_cat
- quilt
- mouse
- siamese_cat
- space_heater

validation



validation approach

- ★ manually coded prediction column for each datapoint
- ★ split data into 70% train and 30% test
 - ★ then split train data into 75% train and 25% validation
- ★ evaluated each model with train, test, and validation accuracy scores
 - ★ SVM model was the best model and score improved 20% with validation dataset

```
Performance of SVM model  
train set:  0.9643835616438357  
test set:   0.9425837320574163  
validation set: 0.9180327868852459
```

conclusion



problem:

craigslist's lost+found section has no option to automatically label and sort posts by lost or found.

craigslist's lost+found section has no option to automatically label and sort posts by lost or found.

OBJECTIVE #1

provide users with the ability to search for all ads via lost or found.

OBJECTIVE #2

provide users with quick links to specific posts based on most popular topics

OBJECTIVE #3

identify images based on keyword

CONCLUSION

lost & found

☐ search titles only☐ has image☐ posted today☐ bundle duplicates☐ include nearby areas

MILES FROM LOCATION

[use map...](#)

Nov 13

LOST SD/MEMORY CARD (LAFAYETTE)



Nov 9

lost black kitten (St Joe addition off teal rd) [pic](#)



Nov 4

*****Lost abandoned cat***** (Lafayette) [pic](#)



Oct 22

key fob found mazda 2014 red (Lafayette)

lost & found

search lost & found

- ☐ search titles only
- ☐ has image
- ☐ posted today
- ☐ bundle duplicates
- ☐ include nearby areas

MILES FROM LOCATION

miles







from zip

use map...

SORT BY

- ☐ lost
- ☐ found

objective #1

-  Nov 13 **LOST SD/MEMORY CARD** (LAFAYETTE)
-  Nov 9 **lost black kitten** (St Joe addition off teal rd) 
-  Nov 4 *****Lost abandoned cat***** (Lafayette) 
-  Oct 22 **key fob found mazda 2014 red** (Lafayette)

lost & found

search lost & found

- ☐ search titles only
- ☐ has image
- ☐ posted today
- ☐ bundle duplicates
- ☐ include nearby areas

MILES FROM LOCATION

miles

from zip

[use map...](#)

SORT BY

- ☐ lost
- ☐ found

CATEGORY

- ☐ cat
- ☐ dog
- ☐ keys/wallet

objective #2



Nov 13

LOST SD/MEMORY CARD (LAFAYETTE)



Nov 9

lost black kitten (St Joe addition off teal rd) [pic](#)



Nov 4

*****Lost abandoned cat***** (Lafayette) [pic](#)



Oct 22

key fob found mazda 2014 red (Lafayette)

lost & found

- ☐ search titles only
- ☐ has image
- ☐ posted today
- ☐ bundle duplicates
- ☐ include nearby areas

MILES FROM LOCATION

miles

from zip

use map...

SORT BY

- ☐ lost
- ☐ found

CATEGORY

▼ cat

- ☐ egyptian
- ☐ siamese
- ☐ all

▼ dog

- ☐ lakeland terrier
- ☐ poodle
- ☐ all

▼ keys/wallet

search lost & found



Nov 13

LOST SD/MEMORY CARD (LAFAYETTE)



Nov 9

lost black kitten (St Joe addition off teal rd) pic



Nov 4

Lost abandoned cat (Lafayette) pic



Oct 22

key fob found mazda 2014 red (Lafayette)

objective #3

lost & found

- ☐ search titles only
- ☐ has image
- ☐ posted today
- ☐ bundle duplicates
- ☐ include nearby areas

MILES FROM LOCATION

[use map...](#)

SORT BY

- ☐ lost
- ☐ found

CATEGORY

▼ cat

- ☐ egyptian
- ☐ siamese
- ☐ all

▼ dog

- ☐ lakeland terrier
- ☐ poodle
- ☐ all

▼ keys/wallet

search lost & found



Nov 13

LOST SD/MEMORY CARD (LAFAYETTE)



Nov 9

lost black kitten (St Joe addition off teal rd) [pic](#)



Nov 4

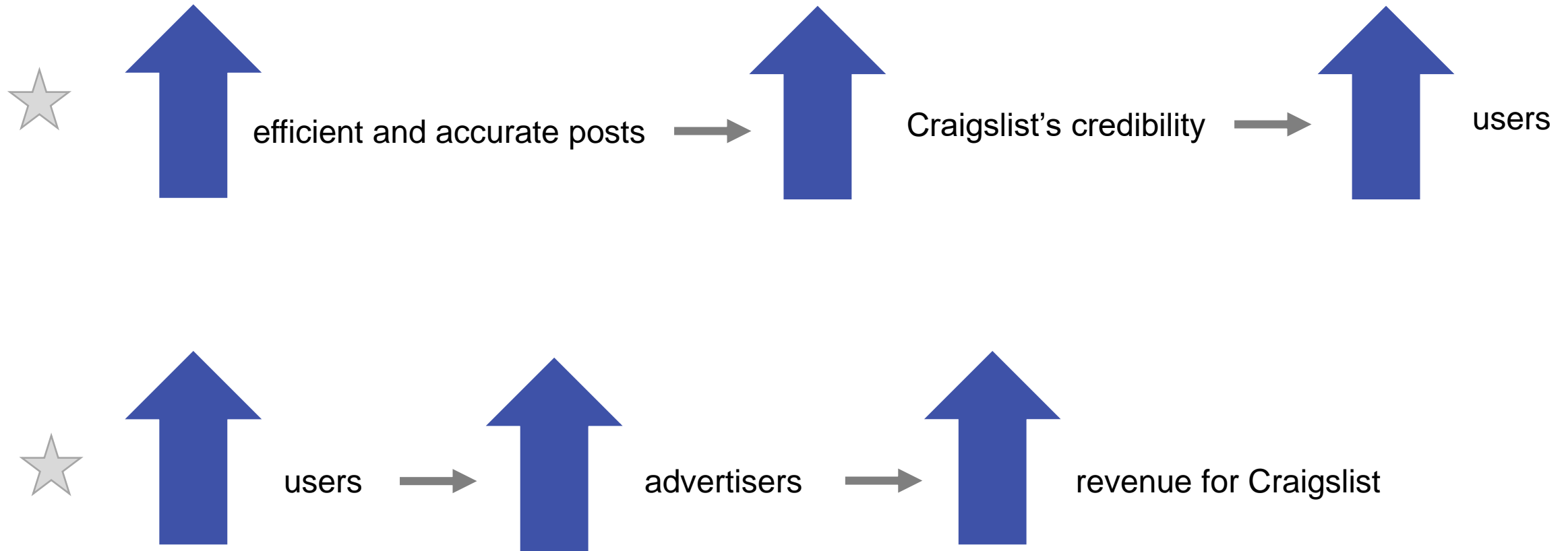
Lost abandoned cat (Lafayette) [pic](#)



Oct 22

key fob found mazda 2014 red (Lafayette)

How suggested improvements bring value to users and Craigslist:



thank you

