Involuntary 'Spotted by Locals' using Geotagged Photos

Maarten de Jonge (6286658)
Anna Keune (6056547)
Merijn van Wouden (6306632)
Supervisor: Frank Nack

May 6, 2012

The links used in document that point to the *iavconcepts* domain are meant for an online view of the results. The links will be valid until at least January 1, 2013. On request, the source code and extracted Flickr data will always be available.

1 Introduction

Millions of tourists come to the city of Amsterdam each year; most of them however flock around the already crowded city centre. It's a beautiful place for sure, but there's plenty to see outside the centre too. The local government would rather have the tourists spread more evenly across the city, to avoid huge crowds.

The 'Spotted by Locals' site provides an online city guide made by locals with their favourite places.[2] The idea of this project is to use social media (more specifically, pictures obtained from Flickr[1]) to automatically find places of interest outside the city centre. This way interesting places that locals go to can become visible, and the city could promote them as interesting tourist locations.

2 Method

The photos were downloaded from Flickr using the Flickr API, saving them to our own database. Starting with the most recent ones, restricting them to be in the border we set in Amsterdam using the latitude and longitude connected to all the (geotagged) photos (non-geotagged ones are not taken into consideration at all). To determine what photos tourists and locals took our approach was to look at the time zones that the users had filled in. There is also a field for the user's country, but it's not as often filled in as the time zones. The disadvantage of this method is that a user could

also be from Germany, making it classified as a local. Though, contrary to GMT time zones, tourists originating from the same longitudinal range as Holland, but latitudinally living far away (e.g. some country in West-Africa) are not classified as local, as the time zones are not specified as GMT, but as blobs of cities close to each other.

Since there are many photos of Amsterdam on Flickr and that number keeps growing, we wanted to make a classifier that can predict whether a photo is relevant or not. We try to map density of photo amounts taken in all the areas to the area's popularity among this group of people (either locals or tourists). Hence, irrelevant examples must not be taken into consideration (a photo of someone's home toilet, newly coloured nails, a daisy, etc). A photo is relevant if there is some publicly accessible place of interest where someone could go to, view or experience. Some examples of very relevant photos: photos of musea, typical Amsterdam tourist attractions, restaurants, concerts and art. Less relevant photos are those of a bicycle against a wall, a photo of someone's shoes or a photo inside a home or an office. To make a classifier we annotated more than 1000 photos from 1 to 3, where 3 is very relevant and 1 is irrelevant. All ungraded photos have a relevance rating of 0. A screen-shot of the annotation tool can be seen in fig 1.

So after the process of annotating part of the dataset, it was necessary to know what made some photos relevant and which not (concerning characteristics extractable with our programming skills). Some information is available through the Flickr API (photo title, tags, description, a user's total amount of photos uploaded yet, it's profile description). The only processing that we did on the photo itself was to make a brightness histogram consisting of 8 bins. These bin values were also used in the machine learning process. This all happened to be insufficient to reliably predict whether a photo is relevant or not. We discovered that the words used in the title, description and tags are necessary (and sufficient) to make a much better performance. So, additional to our basic photo database table, there are two other tables: one containing all found words once, together with its average relevance (taken over all its occurrences, see table 1 and table 2), and another table joining all atomic words separately to a photo id. This way it is possible to say something about the relevance of a bag of words belonging to a photo (so no actual language processing is used). This gave a large boost to the prediction results.

Table 1: Words most linked to relevance

```
SELECT * FROM 'wordcount' WHERE rel_avg > 2.5 AND
count > 50 ORDER BY rel_avg DESC
```

word	count	relcount	rel_avg
peter	126	376	2.98413

Table 1: Words most linked to relevance (continued)

word	count	relcount	rel_avg
singer	100	297	2.97
13th	52	154	2.96154
boots	52	154	2.96154
kathleen	52	154	2.96154
stage	84	248	2.95238
female	87	255	2.93103
little	58	168	2.89655
paradiso	280	809	2.88929
royal	53	153	2.88679
concert	274	770	2.81022
4th	122	342	2.80328
live	249	685	2.751
13	58	158	2.72414
melkweg	163	438	2.68712
march	138	370	2.68116
beautiful	66	175	2.65152
museum	129	342	2.65116
zoo	51	134	2.62745
red	86	220	2.55814
february	115	293	2.54783
sony	150	378	2.52
photography	166	417	2.51205

Table 2: Words least linked to relevance

```
SELECT * FROM `wordcount` WHERE rel_avg < 1.5 AND
count > 50 ORDER BY rel_avg ASC
```

word	count	relcount	rel_avg
jaar	93	94	1.01075
februari	77	79	1.02597
rokin	68	72	1.05882
centraal	61	66	1.08197
aan	53	61	1.15094
station	77	89	1.15584
met	75	88	1.17333
old	54	68	1.25926
het	133	173	1.30075
it	54	72	1.33333
urban	58	78	1.34483

Table 2: Words least linked to relevance (continued)

word	count	relcount	rel_avg
amaro	71	96	1.35211
earlybird	77	105	1.36364
op	129	178	1.37984
this	94	133	1.41489
xproii	71	101	1.42254
iphoneography	645	921	1.42791
uploadedbyinstagram	626	894	1.42812
instagramapp	627	896	1.42903
squareformat	634	911	1.43691
taken	121	175	1.44628
for	88	128	1.45455
normal	91	134	1.47253
square	662	975	1.47281
to	135	201	1.48889

Why people take photos at a place would be really interesting to know, because it would be easier to predict what good photos are and you would be able to make a recommender system that can recommend for example all the places in Amsterdam that have good street art. We tried to find reasons why a photo was interesting by looking at the tags and analysing the image, but without results. The tags are mostly generic and since they're optional they're not always filled in. Image processing is also really hard, because you actually want to know what objects are on the photo and this is far from easy. Histograms that can be made to see how dark or light a photo is don't reveal why that picture was taken, but those values can be used in the decision tree learner.

3 Results

The graded photos were used as a training set in Weka and the ungraded set as a test set. With a decision tree (J48) the ungraded photos were now graded. The entire decision tree can be found in Appendix A. The accuracy of the training set with cross-validation was 70.493% with a F-measure of 0.701. As we classify into three numeric classes, these statistics must be interpreted in a different way as with a binary classification. Classifying a 'bad' photo as a 'good' one is rare, it would rather be classified as 'neutral' when misclassified (which still counts as an error).

Now for the final visualization, the map of Amsterdam is split up in X by X bins (where X is a parameter), and all photos taken inside a bin count for the total amount with a certain weight. The photos with a 'bad' relevance

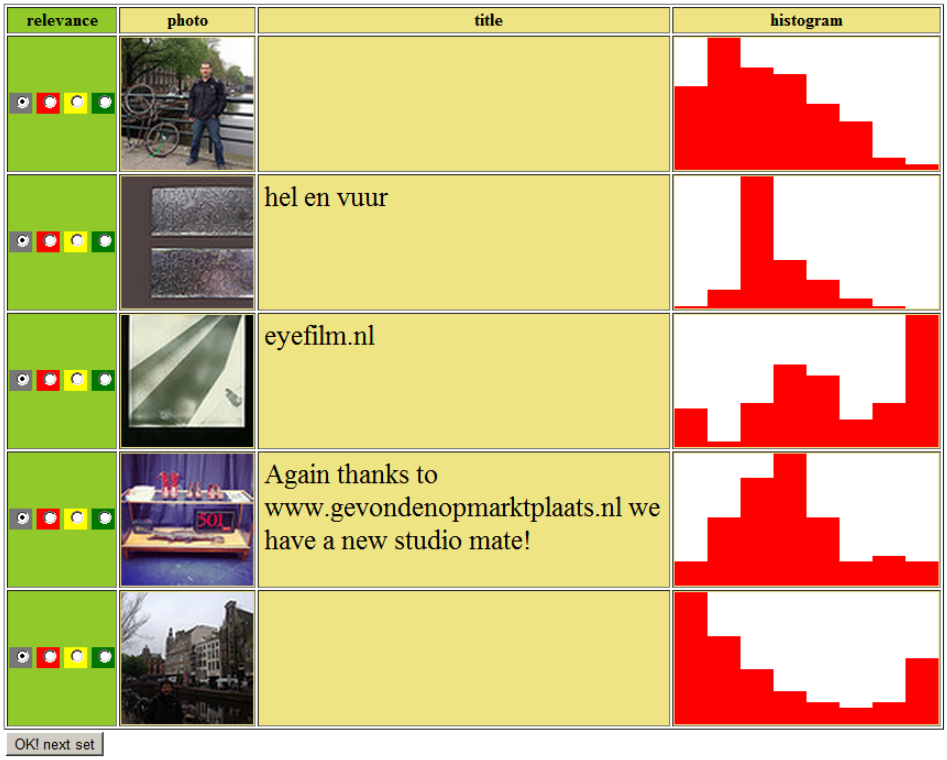

Figure 1: The annotation tool that we made. Can be found on: `http://iavconcepts.com/flickr/annotationtool.php`

have weight 0 (don't count). The 'neutral' ones have weight 1. The 'good' ones have weight 3. These weights are chosen arbitrarily.

The total amount of weights are split between the tourists, and the locals. The ratio between these numbers determine the colour of the bin; See fig 2. If a bin has a greater content of green than red, it is more popular among locals. The more red, the more popular among tourists and less among locals. Photo density is linked to transparency. If a bin is very transparent, it has a low total amount of photos. This scale is taken naturally logarithmic, as the very central bins have a huge density compared to the outskirts. This also means that the bins closer to the centre give more reliable information! As visible and expected, the very central part of the town centre is dominated by tourists, and around it, locals find it more attractive.

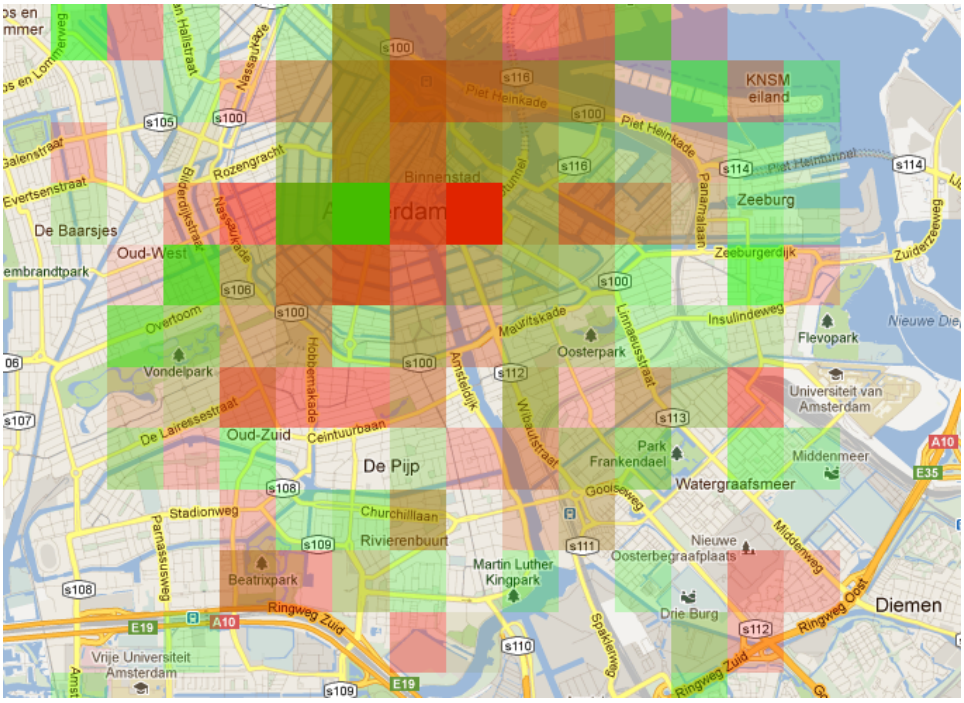

Figure 2: Green stands for the popularity amongst locals and red for the popularity amongst tourists. `http://iavconcepts.com/flickr/amsterdamsect.php`

4 Conclusion

Our method works, as the red areas significantly correlate with the touristic areas whereas the green ones significantly correlate with the ones more popular among locals. In our case this is not very clear, because to the

extent in which we have collected data, we are still bothered by those users that upload a large amount of photos of the same type on the same location (and we didn't make an exception for these people in our approach). This produces an error that can be smoothed by both using more data and by pruning the amount of data gathered from a single user. The more data used, the more the map converges to a realistic one and the smaller the bins can be taken.

5 Discussion

For future research it would be interesting to pinpoint the relevant places more exactly and find the reasons why those places are visited by people.

References

[1] Yahoo! Inc. Flickr. `www.flickr.com`.

[2] SpottedByLocals.com. Spotted by locals - experience cities like a local. `http://www.spottedbylocals.com`.

6 Appendix A

```
J48 pruned tree
------------------
keysval <= 1.95272
| keysval <= 1.57014
| | keysval <= 0
| | | lon <= 4.86076:  1 (95.0/2.0)
| | | lon > 4.86076
| | | | hist8 <= 9
| | | | | hist2 <= 1
| | | | | | user_desclength <= 32:  1 (16.0)
| | | | | | user_desclength > 32
| | | | | | | lon <= 4.92133:  2 (3.0/1.0)
| | | | | | | lon > 4.92133:  1 (3.0)
| | | | | hist2 > 1
| | | | | | hist7 <= 10
| | | | | | | lat <= 52.3783
| | | | | | | | user_timezone = 0
| | | | | | | | | hist5 <= 10
| | | | | | | | | | hist1 <= 42
| | | | | | | | | | | user_photocount <= 6489
| | | | | | | | | | | | lon <= 4.87383
| | | | | | | | | | | | | user_desclength <= 11:  1 (3.53/1.72)
| | | | | | | | | | | | | user_desclength > 11
| | | | | | | | | | | | | | hist4 <= 4:  1 (2.0)
| | | | | | | | | | | | | | hist4 > 4:  2 (6.0/1.0)
| | | | | | | | | | | | | lon > 4.87383:  3 (70.51/33.91)
| | | | | | | | | | | user_photocount > 6489
| | | | | | | | | | | | hist6 <= 3:  2 (2.0)
| | | | | | | | | | | | hist6 > 3:  1 (3.0/1.0)
| | | | | | | | | | hist1 > 42
| | | | | | | | | | | lat <= 52.3674:  2 (4.0/1.0)
| | | | | | | | | | | lat > 52.3674:  1 (2.36)
| | | | | | | | | hist5 > 10:  1 (63.17/28.1)
| | | | | | | | user_timezone = 1
| | | | | | | | | hist2 <= 6
| | | | | | | | | | hist2 <= 5
| | | | | | | | | | | hist1 <= 4:  1 (10.2)
| | | | | | | | | | | hist1 > 4
| | | | | | | | | | | | hist3 <= 5:  1 (4.0)
| | | | | | | | | | | | hist3 > 5:  3 (3.0)
| | | | | | | | | | hist2 > 5:  3 (10.92/1.64)
| | | | | | | | | hist2 > 6:  1 (249.32/102.85)
```

```
| | | | | | | lat > 52.3783
| | | | | | | | hist2 <= 6:  2 (3.0)
| | | | | | | | hist2 > 6
| | | | | | | | | use_photocount <= 3893
| | | | | | | | | | use_photocount <= 78:  3 (3.0/1.0)
| | | | | | | | | | use_photocount > 78
| | | | | | | | | | | hist6 <= 12:  1 (34.0/3.0)
| | | | | | | | | | | hist6 > 12:  2 (2.0)
| | | | | | | | | use_photocount > 3893:  2 (3.0/1.0)
| | | | | | hist7 > 10
| | | | | | | hist5 <= 20
| | | | | | | | hist1 <= 24
| | | | | | | | | lon <= 4.93189
| | | | | | | | | | hist5 <= 15
| | | | | | | | | | | lat <= 52.3706
| | | | | | | | | | | | hist2 <= 5:  2 (6.0)
| | | | | | | | | | | | hist2 > 5
| | | | | | | | | | | | | lat <= 52.3596:  3 (11.0/2.0)
| | | | | | | | | | | | | lat > 52.3596
| | | | | | | | | | | | | | hist3 <= 11
| | | | | | | | | | | | | | | hist8 <= 5
| | | | | | | | | | | | | | | | use_photocount <= 287:  2 (3.0/1.0)
| | | | | | | | | | | | | | | | user_photocount > 287:  1 (3.0)
| | | | | | | | | | | | | | | hist8 > 5:  2 (2.0)
| | | | | | | | | | | | | | hist3 > 11
| | | | | | | | | | | | | | | hist6 <= 9:  3 (3.0)
| | | | | | | | | | | | | | | hist6 > 9:  2 (7.0/1.0)
| | | | | | | | | | lat > 52.3706
| | | | | | | | | | | hist7 <= 12
| | | | | | | | | | | | hist1 <= 8:  2 (3.0/1.0)
| | | | | | | | | | | | hist1 > 8:  3 (5.0/1.0)
| | | | | | | | | | | hist7 > 12
| | | | | | | | | | | | user_desclength <= 308
| | | | | | | | | | | | user_desclength <= 12
| | | | | | | | | | | | | lat <= 52.3773
| | | | | | | | | | | | | | lon <= 4.91481:  3 (5.0)
| | | | | | | | | | | | | | lon > 4.91481:  1 (3.0/1.0)
| | | | | | | | | | | | | lat > 52.3773:  1 (6.0/1.0)
| | | | | | | | | | | | user_desclength > 12:  1 (2.0)
| | | | | | | | | | | user_desclength > 308:  3 (2.0)
| | | | | | | | | hist5 > 15
| | | | | | | | | | lat <= 52.3648:  3 (4.0/1.0)
| | | | | | | | | | lat > 52.3648
| | | | | | | | | | | user_desclength <= 634:  1 (10.0/1.0)
```

```
| | | | | | | | | | | | user_desclength > 634:  2 (4.0/1.0)
| | | | | | | | | | lon > 4.93189:  1 (5.0)
| | | | | | | | | hist1 > 24:  1 (7.0/1.0)
| | | | | | | | hist5 > 20:  3 (12.0)
| | | | hist8 > 9
| | | | | hist3 <= 4
| | | | | | hist3 <= 0:  3 (2.0)
| | | | | | hist3 > 0
| | | | | | | hist6 <= 7:  1 (7.0)
| | | | | | | hist6 > 7
| | | | | | | | lon <= 4.90061:  1 (3.0)
| | | | | | | | lon > 4.90061:  3 (3.0)
| | | | | hist3 > 4
| | | | | | hist7 <= 26
| | | | | | | hist4 <= 4:  3 (10.0/1.0)
| | | | | | | hist4 > 4
| | | | | | | | lat <= 52.356
| | | | | | | | | hist6 <= 4:  1 (2.0)
| | | | | | | | | hist6 > 4
| | | | | | | | | | lat <= 52.3512:  3 (4.0)
| | | | | | | | | | lat > 52.3512:  2 (13.0/1.0)
| | | | | | | | lat > 52.356
| | | | | | | | | user_desclength <= 23
| | | | | | | | | | hist6 <= 4:  3 (14.0/3.0)
| | | | | | | | | | hist6 > 4
| | | | | | | | | | | user_photocount <= 3893
| | | | | | | | | | | | hist8 <= 31
| | | | | | | | | | | | | lat <= 52.372:  3 (29.0/15.0)
| | | | | | | | | | | | | lat > 52.372
| | | | | | | | | | | | | | hist1 <= 12:  3 (11.0/2.0)
| | | | | | | | | | | | | | hist1 > 12:  2 (4.0)
| | | | | | | | | | | | hist8 > 31:  1 (2.0)
| | | | | | | | | | | user_photocount > 3893:  1 (5.0/1.0)
| | | | | | | | | user_desclength > 23
| | | | | | | | | | lat <= 52.3766
| | | | | | | | | | | hist8 <= 23
| | | | | | | | | | | | hist4 <= 22:  1 (17.0/4.0)
| | | | | | | | | | | | hist4 > 22:  3 (2.0)
| | | | | | | | | | | hist8 > 23:  3 (2.0)
| | | | | | | | | | lat > 52.3766:  2 (2.0)
| | | | | hist7 > 26:  3 (12.0)
| | keysval > 0:  1 (1010.0/89.0)
| keysval > 1.57014
| | lon <= 4.85773
```

```
| | | keysval <= 1.88179
| | | | user_photocount <= 2793:  1 (2.0)
| | | | user_photocount > 2793
| | | | | lon <= 4.85383:  3 (2.0)
| | | | | lon > 4.85383:  2 (2.0)
| | | keysval > 1.88179:  2 (25.0)
| | lon > 4.85773
| | | keysval <= 1.68903
| | | | hist2 <= 27
| | | | | lat <= 52.3487
| | | | | | hist6 <= 4:  2 (2.0)
| | | | | | hist6 > 4:  1 (13.0)
| | | | | lat > 52.3487
| | | | | | hist5 <= 18
| | | | | | | hist4 <= 19
| | | | | | | | lat <= 52.3606
| | | | | | | | | user_photocount <= 156:  2 (3.0)
| | | | | | | | | user_photocount > 156
| | | | | | | | | | hist4 <= 5:  3 (6.0/1.0)
| | | | | | | | | | hist4 > 5
| | | | | | | | | | | lat <= 52.3585
| | | | | | | | | | | | hist1 <= 11
| | | | | | | | | | | | | hist4 <= 14:  1 (12.0/3.0)
| | | | | | | | | | | | | hist4 > 14:  3 (2.0)
| | | | | | | | | | | | hist1 > 11:  3 (8.0/1.0)
| | | | | | | | | | | lat > 52.3585:  1 (5.0)
| | | | | | | | lat > 52.3606
| | | | | | | | | hist6 <= 4
| | | | | | | | | | hist4 <= 7
| | | | | | | | | | | user_timezone = 0
| | | | | | | | | | | | hist8 <= 0
| | | | | | | | | | | | | hist3 <= 4:  3 (2.0)
| | | | | | | | | | | | | hist3 > 4:  2 (4.0/1.0)
| | | | | | | | | | | | hist8 > 0
| | | | | | | | | | | | | user_desclength <= 1300:  1 (2.44/0.72)
| | | | | | | | | | | | | user_desclength > 1300:  2 (2.36/1.0)
| | | | | | | | | | | user_timezone = 1
| | | | | | | | | | | | hist8 <= 0:  1 (4.0)
| | | | | | | | | | | | hist8 > 0
| | | | | | | | | | | | | hist2 <= 9
| | | | | | | | | | | | | | user_desclength <= 3477:  3 (4.92/1.64)
| | | | | | | | | | | | | | user_desclength > 3477:  1 (3.0)
| | | | | | | | | | | | | hist2 > 9:  2 (7.28/1.64)
| | | | | | | | | | hist4 > 7
```

```
| | | | | | | | | | keysval <= 1.62368:  2 (9.0/1.0)
| | | | | | | | | | keysval > 1.62368
| | | | | | | | | | | hist3 <= 14:  1 (3.0)
| | | | | | | | | | | hist3 > 14:  2 (8.0/2.0)
| | | | | | | | | hist6 > 4
| | | | | | | | | | user_photocount <= 543
| | | | | | | | | | | user_desclength <= 146
| | | | | | | | | | | | user_timezone = 0:  1 (7.67/2.0)
| | | | | | | | | | | | user_timezone = 1
| | | | | | | | | | | | | hist5 <= 8:  2 (5.0)
| | | | | | | | | | | | | hist5 > 8
| | | | | | | | | | | | | | lon <= 4.88582:  2 (3.0)
| | | | | | | | | | | | | | lon > 4.88582:  1 (7.33/2.0)
| | | | | | | | | | | user_desclength > 146
| | | | | | | | | | | | user_photocount <= 221
| | | | | | | | | | | | | lon <= 4.89955:  3 (2.0)
| | | | | | | | | | | | | lon > 4.89955:  1 (2.0)
| | | | | | | | | | | | user_photocount > 221:  1 (4.0)
| | | | | | | | | | user_photocount > 543
| | | | | | | | | | | hist4 <= 6
| | | | | | | | | | | | lat <= 52.3657:  3 (3.0/1.0)
| | | | | | | | | | | | lat > 52.3657:  1 (13.0/1.0)
| | | | | | | | | | | hist4 > 6
| | | | | | | | | | | | hist4 <= 7
| | | | | | | | | | | | | lon <= 4.90901
| | | | | | | | | | | | | | hist5 <= 5:  1 (2.0/1.0)
| | | | | | | | | | | | | | hist5 > 5:  2 (6.0)
| | | | | | | | | | | | | lon > 4.90901:  1 (3.0)
| | | | | | | | | | | | hist4 > 7
| | | | | | | | | | | | | hist4 <= 10
| | | | | | | | | | | | | | hist3 <= 16
| | | | | | | | | | | | | | | hist8 <= 5
| | | | | | | | | | | | | | | | lat <= 52.3654:  1 (2.0)
| | | | | | | | | | | | | | | | lat > 52.3654:  3 (6.0/1.0)
| | | | | | | | | | | | | | | hist8 > 5
| | | | | | | | | | | | | | | | lon <= 4.89909:  3 (8.0/1.0)
| | | | | | | | | | | | | | | | lon > 4.89909
| | | | | | | | | | | | | | | | | hist3 <= 10:  2 (4.0/1.0)
| | | | | | | | | | | | | | | | | hist3 > 10:  3 (2.0)
| | | | | | | | | | | | | | | hist3 > 16:  1 (2.0)
| | | | | | | | | | | | | | hist4 > 10
| | | | | | | | | | | | | | | hist7 <= 3:  3 (3.0)
| | | | | | | | | | | | | | | hist7 > 3
| | | | | | | | | | | | | | | | hist2 <= 17
```

```
| | | | | | | | | | | | | | | hist5 <= 12
| | | | | | | | | | | | | | | | user_timezone = 0
| | | | | | | | | | | | | | | | | hist6 <= 8:  1 (2.0/1.0)
| | | | | | | | | | | | | | | | | hist6 > 8:  3 (3.42/0.42)
| | | | | | | | | | | | | | | | | user_timezone = 1:  1 (7.58)
| | | | | | | | | | | | | | | | hist5 > 12
| | | | | | | | | | | | | | | | | hist1 <= 4:  1 (2.0)
| | | | | | | | | | | | | | | | | hist1 > 4
| | | | | | | | | | | | | | | | | user_timezone = 0:  2 (2.0)
| | | | | | | | | | | | | | | | | user_timezone = 1
| | | | | | | | | | | | | | | | | | lon <= 4.89153:  2 (5.0/1.0)
| | | | | | | | | | | | | | | | | | lon > 4.89153:  3 (5.0/1.0)
| | | | | | | | | | | | | | | hist2 > 17:  2 (7.0/1.0)
| | | | | | | hist4 > 19
| | | | | | | | lon <= 4.91083:  1 (14.0/2.0)
| | | | | | | | lon > 4.91083:  3 (3.0)
| | | | | | hist5 > 18
| | | | | | | hist1 <= 1:  2 (3.0)
| | | | | | | hist1 > 1
| | | | | | | | user_desclength <= 528
| | | | | | | | | lat <= 52.3766
| | | | | | | | | | hist4 <= 14:  1 (4.0/1.0)
| | | | | | | | | | hist4 > 14:  2 (8.0/1.0)
| | | | | | | | | lat > 52.3766:  1 (3.0)
| | | | | | | | user_desclength > 528:  1 (12.0/1.0)
| | | | hist2 > 27
| | | | | hist4 <= 7:  3 (3.0)
| | | | | hist4 > 7:  2 (4.0/1.0)
| | | keysval > 1.68903
| | | | hist4 <= 11
| | | | | user_photocount <= 22185
| | | | | | user_timezone = 0
| | | | | | | lat <= 52.3664
| | | | | | | | user_desclength <= 28
| | | | | | | | | hist7 <= 5:  2 (10.78/2.61)
| | | | | | | | | hist7 > 5
| | | | | | | | | | user_photocount <= 183:  3 (2.61/0.87)
| | | | | | | | | | user_photocount > 183
| | | | | | | | | | | hist1 <= 11:  2 (2.43/0.43)
| | | | | | | | | | | hist1 > 11:  1 (2.3/0.43)
| | | | | | | | user_desclength > 28
| | | | | | | | | hist4 <= 8
| | | | | | | | | | hist3 <= 18
| | | | | | | | | | | hist7 <= 23
```

```
| | | | | | | | | | | | hist1 <= 27
| | | | | | | | | | | | | hist3 <= 9:  1 (8.87/0.43)
| | | | | | | | | | | | | hist3 > 9
| | | | | | | | | | | | | | lat <= 52.3583:  2 (2.43)
| | | | | | | | | | | | | | lat > 52.3583:  1 (5.0/1.0)
| | | | | | | | | | | | hist1 > 27
| | | | | | | | | | | | | user_desclength <= 359:  3 (2.0)
| | | | | | | | | | | | | user_desclength > 359:  2 (4.43/1.43)
| | | | | | | | | | | | hist7 > 23:  3 (2.0)
| | | | | | | | | | | hist3 > 18:  3 (2.43)
| | | | | | | | | | hist4 > 8:  1 (6.87)
| | | | | | | | lat > 52.3664
| | | | | | | | hist3 <= 18
| | | | | | | | | lon <= 4.91224
| | | | | | | | | lon <= 4.88667
| | | | | | | | | | hist4 <= 7
| | | | | | | | | | | user_desclength <= 264
| | | | | | | | | | | | lon <= 4.88327:  1 (3.43/1.43)
| | | | | | | | | | | | lon > 4.88327:  3 (10.0/1.0)
| | | | | | | | | | | user_desclength > 264:  1 (2.0)
| | | | | | | | | | hist4 > 7:  3 (7.04/0.87)
| | | | | | | | | lon > 4.88667
| | | | | | | | | | hist3 <= 8
| | | | | | | | | | | hist8 <= 12
| | | | | | | | | | | | hist8 <= 8
| | | | | | | | | | | | | hist7 <= 3
| | | | | | | | | | | | | | hist1 <= 64:  2 (6.43)
| | | | | | | | | | | | | | hist1 > 64
| | | | | | | | | | | | | | | user_photocount <= 471:  1 (2.0/1.0)
| | | | | | | | | | | | | | | user_photocount > 471:  3 (2.0)
| | | | | | | | | | | | | hist7 > 3
| | | | | | | | | | | | | | hist4 <= 7:  1 (15.43/4.43)
| | | | | | | | | | | | | | hist4 > 7
| | | | | | | | | | | | | | | hist3 <= 6:  1 (3.0/1.0)
| | | | | | | | | | | | | | | hist3 > 6:  3 (3.3/0.43)
| | | | | | | | | | | | hist8 > 8:  2 (8.0/1.0)
| | | | | | | | | | | hist8 > 12
| | | | | | | | | | | | hist6 <= 6
| | | | | | | | | | | | | lon <= 4.89308:  3 (2.0)
| | | | | | | | | | | | | lon > 4.89308:  1 (2.0)
| | | | | | | | | | | | hist6 > 6:  3 (2.87/0.43)
| | | | | | | | | | hist3 > 8
| | | | | | | | | | | lon <= 4.89667
| | | | | | | | | | | | hist2 <= 24
```

```
| | | | | | | | | | | | | | lat <= 52.3725:  1 (10.43/1.43)
| | | | | | | | | | | | | | lat > 52.3725
| | | | | | | | | | | | | | | hist7 <= 2
| | | | | | | | | | | | | | | | keysval <= 1.83784:  1 (2.0)
| | | | | | | | | | | | | | | | keysval > 1.83784:  2 (4.0/1.0)
| | | | | | | | | | | | | | | hist7 > 2
| | | | | | | | | | | | | | | | hist6 <= 6:  3 (2.43)
| | | | | | | | | | | | | | | | hist6 > 6:  2 (2.17/0.87)
| | | | | | | | | | | | | | hist2 > 24:  3 (6.43/1.43)
| | | | | | | | | | | | | | lon > 4.89667
| | | | | | | | | | | | | | hist2 <= 15:  3 (12.17/0.87)
| | | | | | | | | | | | | | hist2 > 15
| | | | | | | | | | | | | | hist6 <= 1:  3 (3.0)
| | | | | | | | | | | | | | hist6 > 1:  2 (5.43/2.0)
| | | | | | | | | lon > 4.91224:  3 (6.0)
| | | | | | | | hist3 > 18
| | | | | | | | | hist2 <= 27:  3 (7.87/1.43)
| | | | | | | | | hist2 > 27:  2 (2.0)
| | | | | | user_timezone = 1
| | | | | | | hist8 <= 11
| | | | | | | | user_photocount <= 5033
| | | | | | | | | user_desclength <= 271
| | | | | | | | | | hist8 <= 0:  3 (12.09/4.7)
| | | | | | | | | | hist8 > 0
| | | | | | | | | | | hist3 <= 4
| | | | | | | | | | | | user_photocount <= 871:  3 (2.0)
| | | | | | | | | | | | user_photocount > 871:  1 (15.0/2.0)
| | | | | | | | | | | hist3 > 4:  3 (97.3/53.91)
| | | | | | | | | user_desclength > 271
| | | | | | | | | | hist1 <= 23
| | | | | | | | | | | hist8 <= 4
| | | | | | | | | | | | hist1 <= 9
| | | | | | | | | | | | | user_desclength <= 897:  1 (2.13/0.57)
| | | | | | | | | | | | | user_desclength > 897:  3 (2.0)
| | | | | | | | | | | | hist1 > 9
| | | | | | | | | | | | | hist6 <= 1:  1 (2.57/1.0)
| | | | | | | | | | | | | hist6 > 1:  2 (9.7/1.57)
| | | | | | | | | | | hist8 > 4
| | | | | | | | | | | | hist4 <= 9
| | | | | | | | | | | | | hist3 <= 6:  2 (3.0)
| | | | | | | | | | | | | hist3 > 6:  1 (4.13/1.0)
| | | | | | | | | | | | hist4 > 9:  2 (7.0)
| | | | | | | | | | hist1 > 23
| | | | | | | | | | | user_photocount <= 1014
```

```
| | | | | | | | | | | | hist7 <= 5:  3 (6.13/1.0)
| | | | | | | | | | | | hist7 > 5:  1 (2.57/0.57)
| | | | | | | | | | | user_photocount > 1014
| | | | | | | | | | | | hist8 <= 1
| | | | | | | | | | | | | lon <= 4.88685:  2 (2.0)
| | | | | | | | | | | | | lon > 4.88685:  1 (2.13/0.57)
| | | | | | | | | | | | hist8 > 1:  1 (6.0)
| | | | | | | | user_photocount > 5033:  3 (19.7/3.57)
| | | | | | | hist8 > 11
| | | | | | | | hist8 <= 33
| | | | | | | | | hist4 <= 10
| | | | | | | | | | hist7 <= 10
| | | | | | | | | | | hist3 <= 9:  3 (16.26/3.0)
| | | | | | | | | | | hist3 > 9
| | | | | | | | | | | | hist1 <= 11:  2 (4.57)
| | | | | | | | | | | | hist1 > 11:  3 (3.26/0.57)
| | | | | | | | | hist7 > 10
| | | | | | | | | | hist4 <= 4:  3 (2.0)
| | | | | | | | | | hist4 > 4:  2 (13.26/2.13)
| | | | | | | | hist4 > 10:  3 (2.57/0.57)
| | | | | | | | hist8 > 33:  1 (4.0/1.0)
| | | | | user_photocount > 22185:  3 (11.0/1.0)
| | | | hist4 > 11
| | | | | hist8 <= 0
| | | | | | hist2 <= 3:  1 (2.0)
| | | | | | hist2 > 3
| | | | | | | hist4 <= 14
| | | | | | | | hist7 <= 1:  2 (4.0)
| | | | | | | | hist7 > 1:  3 (5.0/1.0)
| | | | | | | hist4 > 14:  3 (15.0/1.0)
| | | | | hist8 > 0
| | | | | | hist1 <= 1:  2 (5.0)
| | | | | | hist1 > 1
| | | | | | | hist7 <= 12
| | | | | | | | hist7 <= 5
| | | | | | | | | hist7 <= 4
| | | | | | | | | | lat <= 52.3563
| | | | | | | | | | | user_photocount <= 5107:  1 (6.0/1.0)
| | | | | | | | | | | user_photocount > 5107:  3 (2.0)
| | | | | | | | | | lat > 52.3563
| | | | | | | | | | | hist8 <= 12
| | | | | | | | | | | | hist6 <= 4
| | | | | | | | | | | | | hist7 <= 3
| | | | | | | | | | | | | | lon <= 4.89179
```

```
| | | | | | | | | | | | | | | user_photocount <= 1447:  3 (5.0/1.0)
| | | | | | | | | | | | | | | user_photocount > 1447
| | | | | | | | | | | | | | | | hist3 <= 23
| | | | | | | | | | | | | | | | | user_desclength <= 28:  1 (5.0/1.0)
| | | | | | | | | | | | | | | | | user_desclength > 28:  2 (3.0/1.0)
| | | | | | | | | | | | | | | | hist3 > 23:  2 (4.0)
| | | | | | | | | | | | | | | lon > 4.89179:  2 (6.0)
| | | | | | | | | | | | | | | hist7 > 3:  3 (2.0)
| | | | | | | | | | | | | | hist6 > 4
| | | | | | | | | | | | | | | lon <= 4.88634:  2 (11.0/1.0)
| | | | | | | | | | | | | | | lon > 4.88634
| | | | | | | | | | | | | | | hist7 <= 3
| | | | | | | | | | | | | | | | user_desclength <= 565:  3 (15.0/2.0)
| | | | | | | | | | | | | | | | user_desclength > 565
| | | | | | | | | | | | | | | | | lat <= 52.3725:  3 (2.0)
| | | | | | | | | | | | | | | | | lat > 52.3725:  2 (3.0)
| | | | | | | | | | | | | | | hist7 > 3
| | | | | | | | | | | | | | | | hist1 <= 4:  3 (3.0)
| | | | | | | | | | | | | | | | hist1 > 4:  2 (9.0/1.0)
| | | | | | | | | | | hist8 > 12:  3 (4.0)
| | | | | | | | | | hist7 > 4
| | | | | | | | | | | hist6 <= 8
| | | | | | | | | | | | keysval <= 1.82383:  3 (4.0)
| | | | | | | | | | | | keysval > 1.82383:  2 (3.0)
| | | | | | | | | | | hist6 > 8:  2 (9.0)
| | | | | | | | | hist7 > 5
| | | | | | | | | | hist6 <= 7
| | | | | | | | | | | lon <= 4.88265:  3 (4.0)
| | | | | | | | | | | lon > 4.88265
| | | | | | | | | | | | hist4 <= 13:  3 (4.0/1.0)
| | | | | | | | | | | | hist4 > 13:  1 (8.0/1.0)
| | | | | | | | | | hist6 > 7
| | | | | | | | | | | hist4 <= 19
| | | | | | | | | | | | keysval <= 1.79692
| | | | | | | | | | | | | user_timezone = 0
| | | | | | | | | | | | | | hist7 <= 7:  1 (2.5)
| | | | | | | | | | | | | | hist7 > 7
| | | | | | | | | | | | | | | user_desclength <= 36
| | | | | | | | | | | | | | | | lon <= 4.89469:  1 (2.0)
| | | | | | | | | | | | | | | | lon > 4.89469:  3 (2.0)
| | | | | | | | | | | | | | | user_desclength > 36:  3 (2.0)
| | | | | | | | | | | | | | user_timezone = 1:  3 (8.5/1.5)
| | | | | | | | | | | | | | keysval > 1.79692
| | | | | | | | | | | | | | | hist3 <= 20
```

```
| | | | | | | | | | | | user_timezone = 0
| | | | | | | | | | | | | user_desclength <= 277:  2 (7.32/0.44)
| | | | | | | | | | | | | user_desclength > 277
| | | | | | | | | | | | | | lat <= 52.3592:  2 (2.0/1.0)
| | | | | | | | | | | | | | lat > 52.3592:  1 (3.0)
| | | | | | | | | | | | | user_timezone = 1:  2 (15.68/6.56)
| | | | | | | | | | | | hist3 > 20:  3 (4.0)
| | | | | | | | | | | hist4 > 19
| | | | | | | | | | | | lat <= 52.3718:  1 (4.0)
| | | | | | | | | | | | lat > 52.3718:  2 (3.0)
| | | | | | | hist7 > 12
| | | | | | | | hist4 <= 16
| | | | | | | | | hist2 <= 5:  2 (3.0)
| | | | | | | | | hist2 > 5
| | | | | | | | | | lat <= 52.3596
| | | | | | | | | | | user_photocount <= 5336:  2 (3.0)
| | | | | | | | | | | user_photocount > 5336:  3 (2.0)
| | | | | | | | | | lat > 52.3596:  3 (10.0)
| | | | | | | | hist4 > 16:  2 (4.0)
keysval > 1.95272
| keysval <= 2.12094
| | user_desclength <= 7791
| | | use_photocount <= 1991
| | | | hist8 <= 32:  3 (116.0/16.0)
| | | | hist8 > 32:  1 (2.0/1.0)
| | | use_photocount > 1991
| | | | use_desclength <= 356
| | | | | use_desclength <= 187
| | | | | | hist7 <= 9:  3 (33.0/5.0)
| | | | | | hist7 > 9
| | | | | | | lon <= 4.8969
| | | | | | | | hist5 <= 6:  3 (3.0/1.0)
| | | | | | | | hist5 > 6:  2 (6.0)
| | | | | | | lon > 4.8969:  1 (2.0/1.0)
| | | | | use_desclength > 187:  2 (96.0/36.0)
| | | | use_desclength > 356
| | | | | hist2 <= 16:  3 (25.0/3.0)
| | | | | hist2 > 16
| | | | | | hist7 <= 7
| | | | | | | hist8 <= 3
| | | | | | | | hist8 <= 2
| | | | | | | | | lon <= 4.8977:  3 (9.0/1.0)
| | | | | | | | | lon > 4.8977:  2 (2.0)
| | | | | | | hist8 > 2:  2 (2.0)
```

```
| | | | | | | hist8 > 3:  3 (8.0)
| | | | | | | hist7 > 7:  2 (2.0)
| | use_desclength > 7791
| | | hist7 <= 2
| | | | hist4 <= 1:  1 (4.0)
| | | | hist4 > 1:  3 (5.0/1.0)
| | | hist7 > 2:  1 (4.0)
| keysval > 2.12094:  3 (1002.0/23.0)
Number of Leaves :  245
Size of the tree :  489
```