@MikeIzbicki

Follow

Estoy en CSUN para explicar

How to geolocate social media messages

314159 27182 161803 (42)

Case study: Power and Kibell (2017)

911 operator receives report of thick black smoke somewhere outside Melbourne

Case study: Power and Kibell (2017)

911 operator receives report of thick black smoke somewhere outside Melbourne

A “social media intelligence analyst” searches Twitter for more information

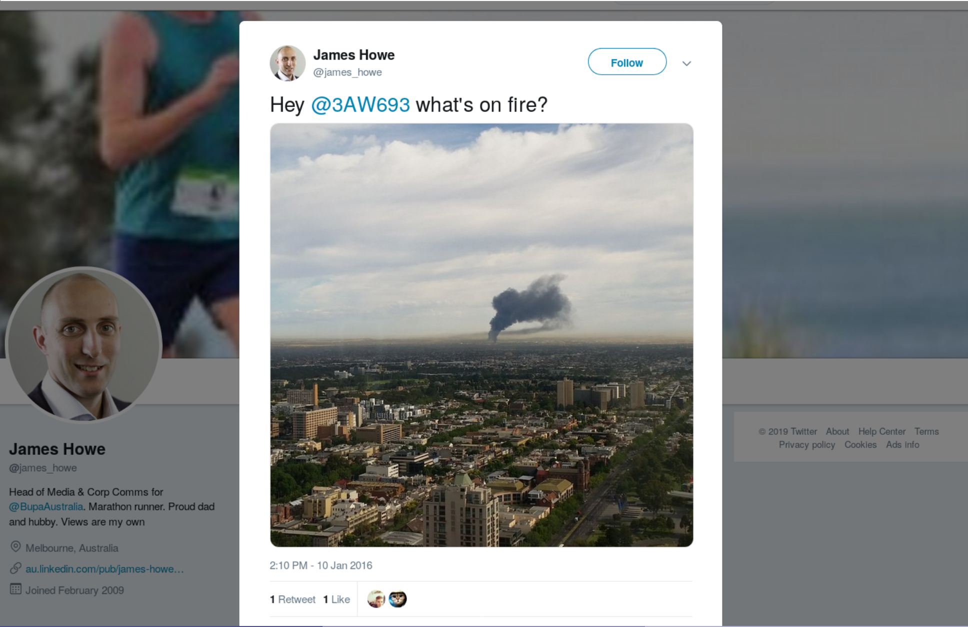
Case study: Power and Kibell (2017)

911 operator receives report of thick black smoke somewhere outside Melbourne

A "social media intelligence analyst" searches Twitter for more information

They find two important tweets:

Case study: Power and Kibell (2017)



James Howe
@james_howe

Hey @3AW693 what's on fire?

2:10 PM - 10 Jan 2016

1 Retweet 1 Like

James Howe
@james_howe

Head of Media & Corp Comms for @BupaAustralia. Marathon runner. Proud dad and hubby. Views are my own

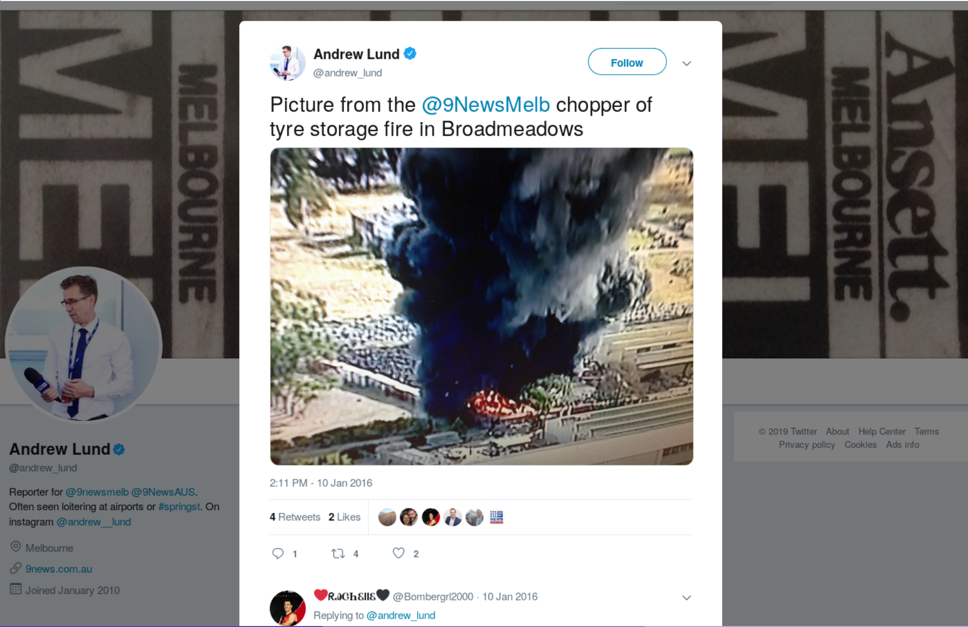
Melbourne, Australia

au.linkedin.com/pub/james-howe…

Joined February 2009

© 2019 Twitter About Help Center Terms
Privacy policy Cookies Ads info

Case study: Power and Kibell (2017)



Andrew Lund
@andrew_lund

Reporter for @9newsmelb @9NewsAUS. Often seen loitering at airports or #springst. On instagram @andrew__lund

Melbourne

9news.com.au

Joined January 2010

Andrew Lund
@andrew_lund

Follow

Picture from the @9NewsMelb chopper of tyre storage fire in Broadmeadows

2:11 PM - 10 Jan 2016

4 Retweets 2 Likes

1 4 2

RᴥCh٤ll٤ @Bombergrl2000 · 10 Jan 2016
Replying to @andrew_lund

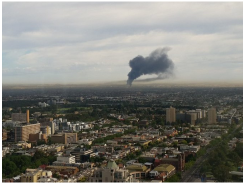
© 2019 Twitter About Help Center Terms
Privacy policy Cookies Ads info

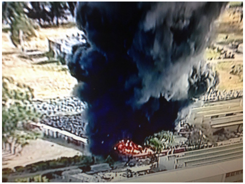
Case study: Power and Kibell (2017)

911 operator receives report of thick black smoke somewhere outside Melbourne

A "social media intelligence analyst" searches Twitter for more information

They find two important tweets:





Case study: Power and Kibell (2017)

911 operator receives report of thick black smoke somewhere outside Melbourne

A “social media intelligence analyst” searches Twitter for more information

They find two important tweets:

The fire chief uses these photos to dispatch firefighters

- Tire fires are dangerous and difficult to extinguish
- Require prompt response to prevent spread
- Require bulldozers and aircraft

Case study: Power and Kibell (2017)

911 operator receives report of thick black smoke somewhere outside Melbourne

A "social media intelligence analyst" searches Twitter for more information

They find two important tweets:

The fire chief uses these photos to dispatch firefighters

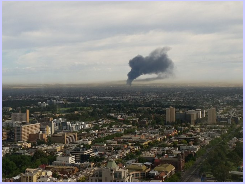
- Tire fires are dangerous and difficult to extinguish
- Require prompt response to prevent spread
- Require bulldowzers and aircraft

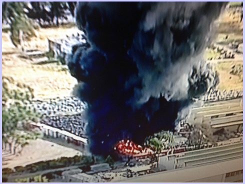
Intelligence from Twitter enabled a prompt response, preventing disaster

Case study: Power and Kibell (2017)

A "social media intelligence analyst" searches Twitter for more information

They find two important tweets:

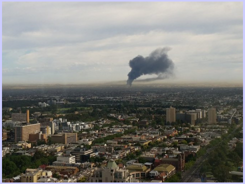


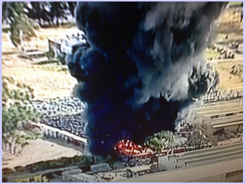


Case study: Power and Kibell (2017)

A "social media intelligence analyst" searches Twitter for more information

They find two important tweets:





How does this happen?!

Case study: Power and Kibell (2017)

A “social media intelligence analyst” searches Twitter for more information

They find two important tweets:

How does this happen?!

Need software that

- understands the physical GPS coordinates of messages
- understands the topic of messages
- searches massive datasets efficiently

Case study: Power and Kibell (2017)

A "social media intelligence analyst" searches Twitter for more information

They find two important tweets:

How does this happen?!

Need software that

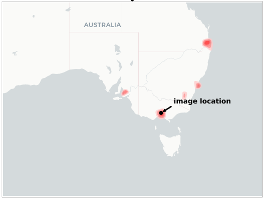
- understands the physical GPS coordinates of messages
- understands the topic of messages
- searches massive datasets efficiently

The geolocation problem

or

The @9NewsMelb news chopper is flying above a tyre storage fire in Broadmeadows

Machine Learning



probability distribution over the earth

The geolocation problem

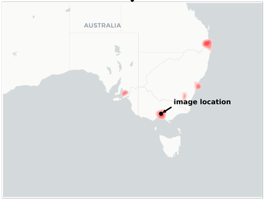
or

The @9NewsMelb news chopper is flying above a tyre storage fire in Broadmeadows

This talk →

Machine Learning

probability distribution over the earth

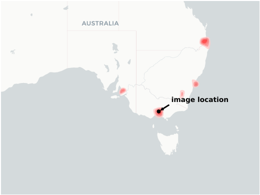


The geolocation problem

or

The @9NewsMelb news chopper is flying above a tyre storage fire in Broadmeadows

This talk → Machine Learning



probability distribution over the earth

Many other applications:

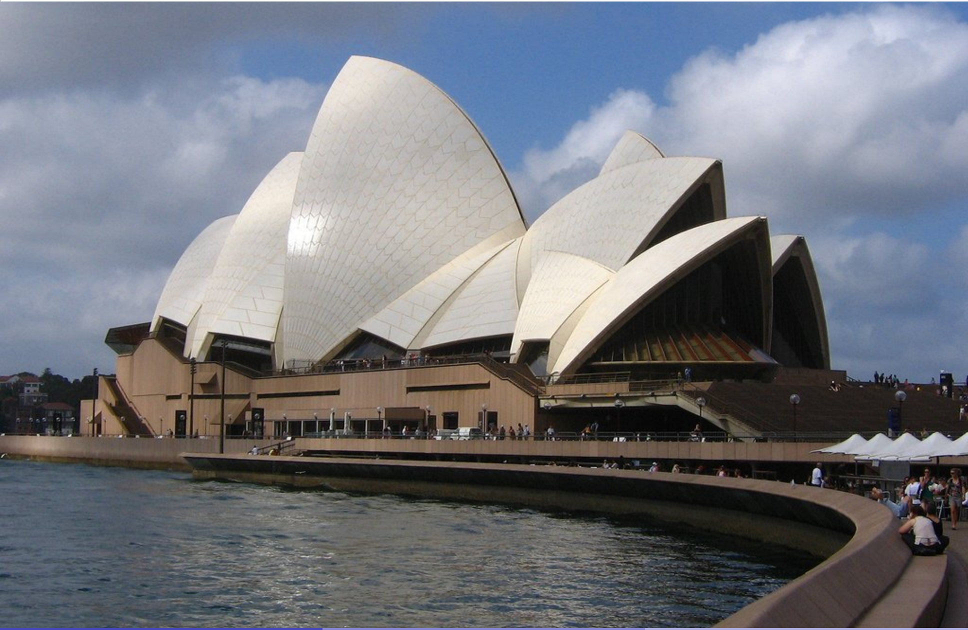
- identify witnesses to a crime (Truelove et al., 2017)
- map the spread of influenza (Paul et al., 2014)
- estimate unemployment rates (Antenucci et al., 2014)
- monitor climate change (Wentz et al., 2014)

... and more ...

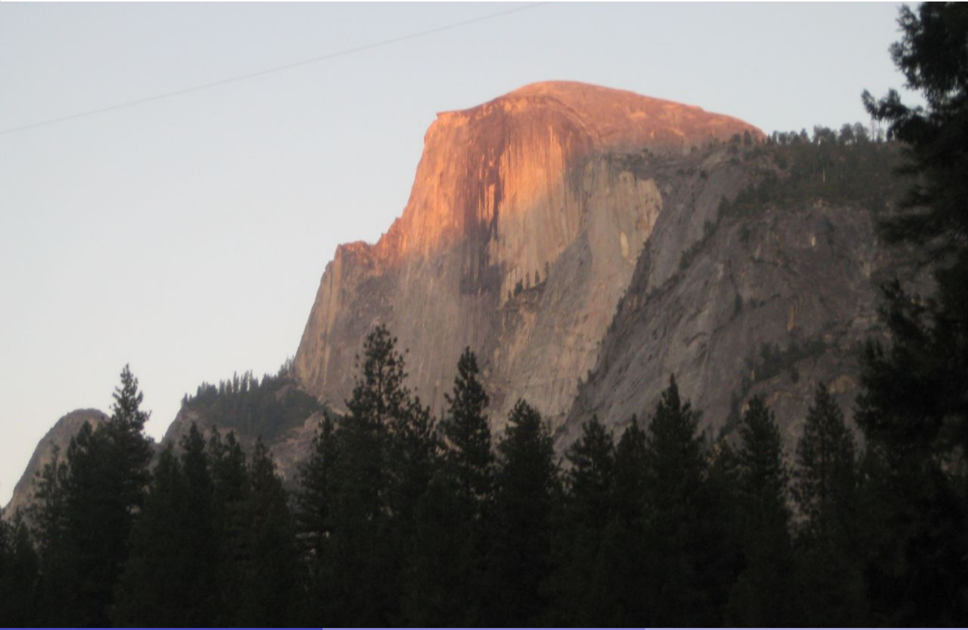
Outline

- Geolocating images with deep learning
  - Examples
  - Deep learning review
  - The PlaNet method (Weyand et al., 2016)
  - The mixture of von Mises-Fisher distribution
- Geolocating text with deep learning
  - Overview of Twitter
  - Examples
  - Word-based methods
  - UnicodeCNN, a character based method
- Future research directions

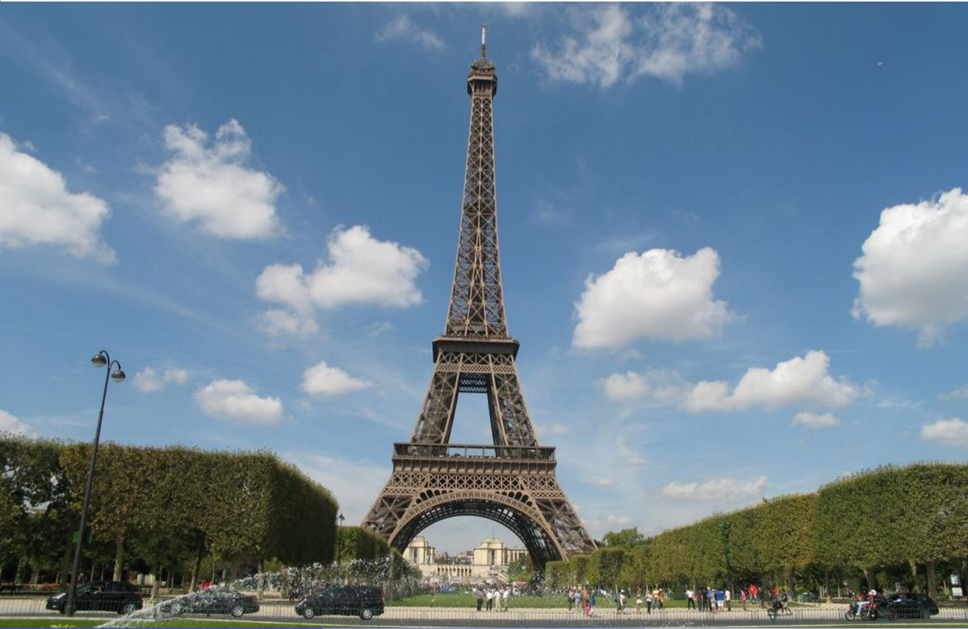
Easy Examples



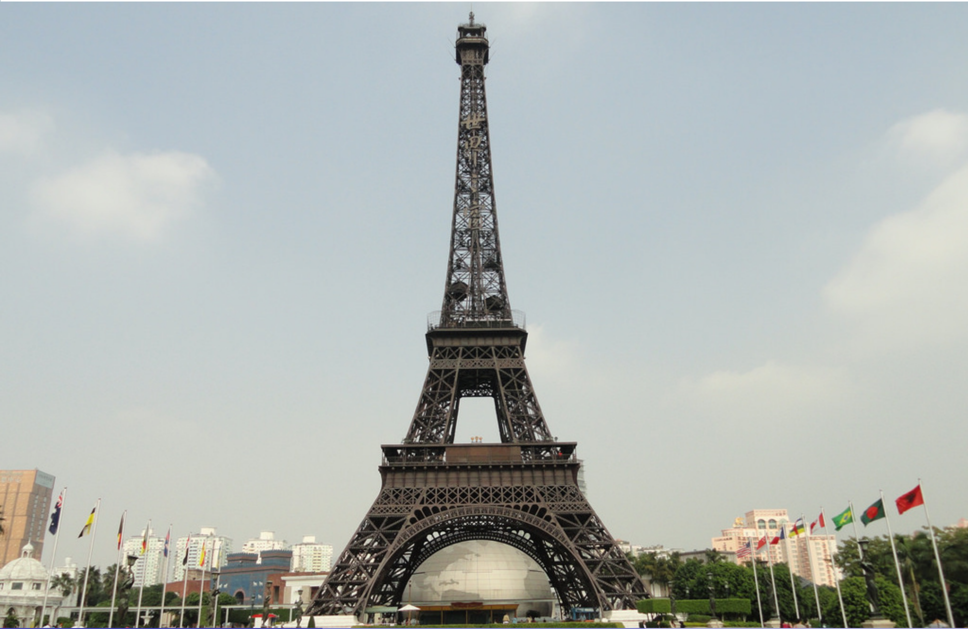
Easy Examples



Easy Examples



Easy Examples



Easy Examples

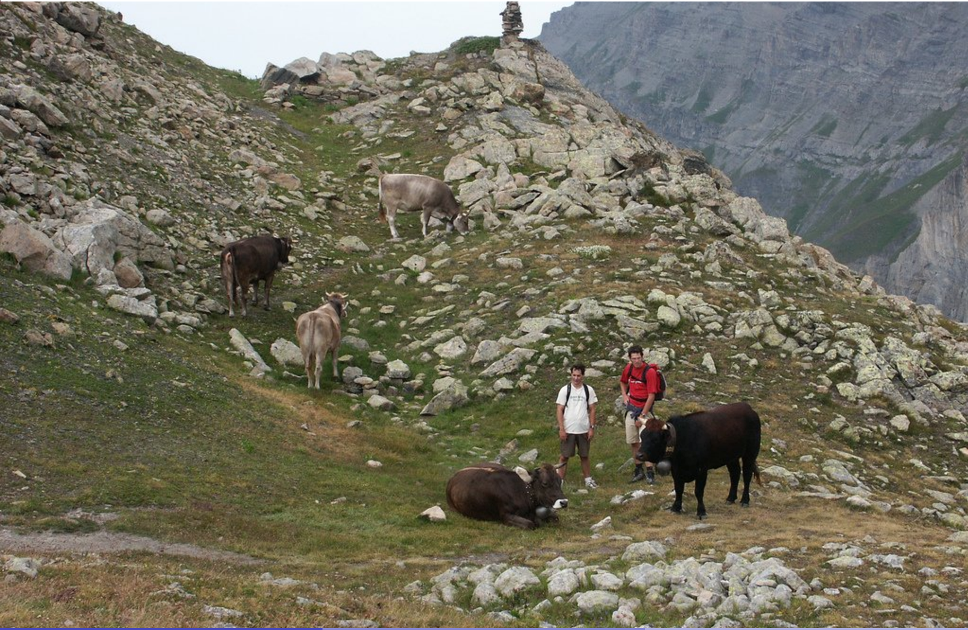
Image contains sufficient information for exact geolocation

Trained humans can geolocate <1km

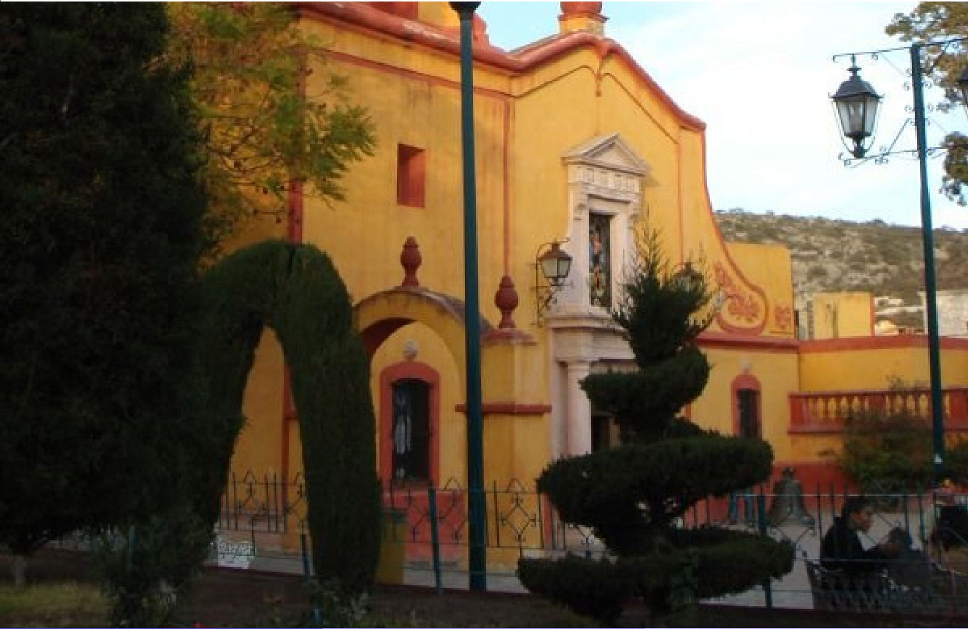
Many algorithms also have great accuracy

- Hays and Efros (2008)
- Hays and Efros (2015)
- Weyand et al. (2016)
- Vo et al. (2017)
- ... and more ...

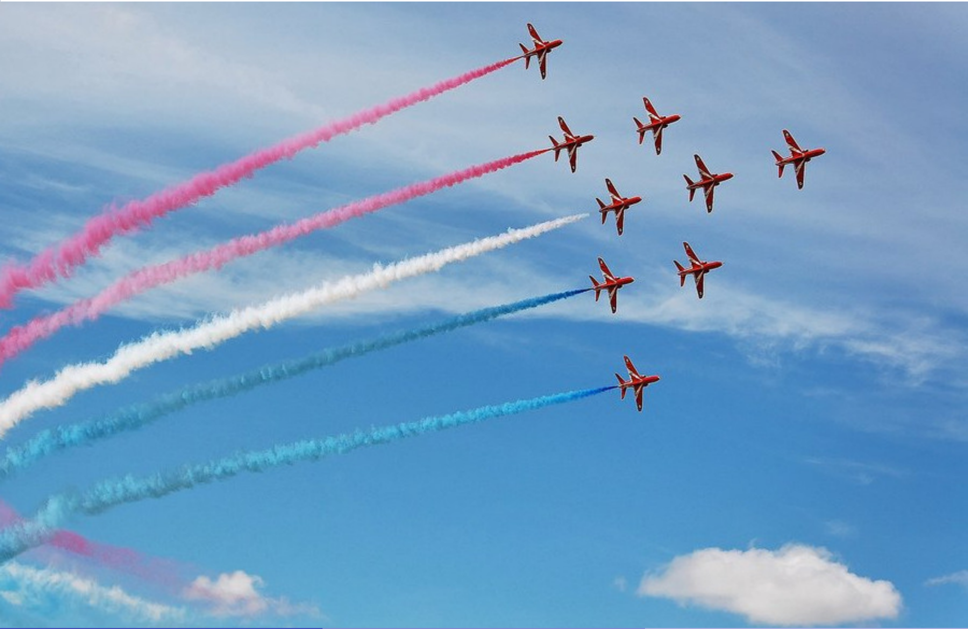
Hard Examples



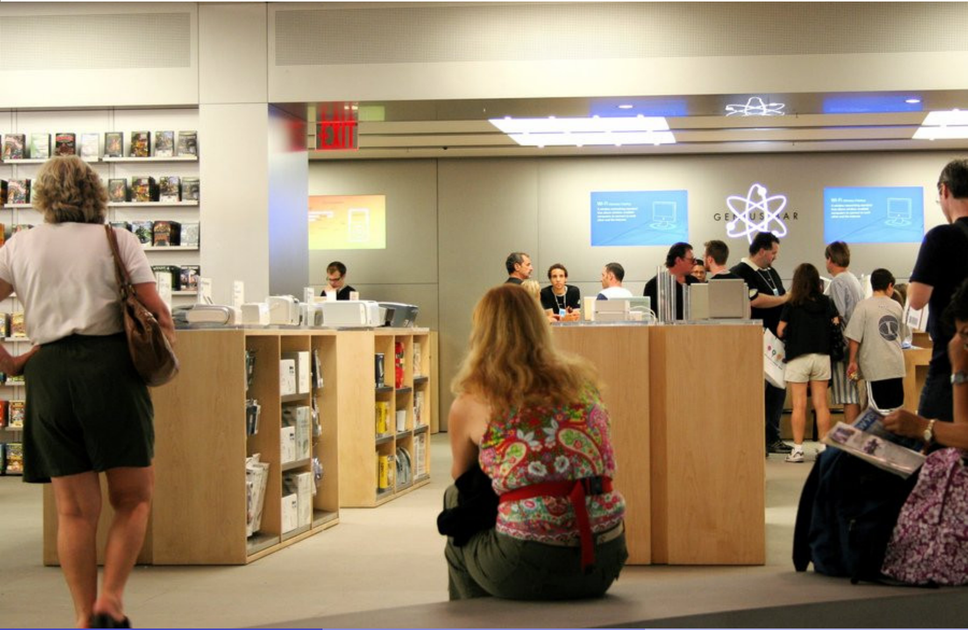
Hard Examples



Hard Examples



Hard Examples



Hard Examples

Images do not contain sufficient information for exact geolocation

Trained humans can determine general regions

Existing algorithms are “overconfident” and fail for these images

Most images fall into this category

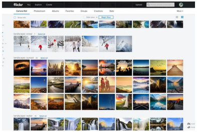
The data

5 million geotagged images provided by Mousselly-Sergieh et al. (2014)

Geotags generated by

cellphones

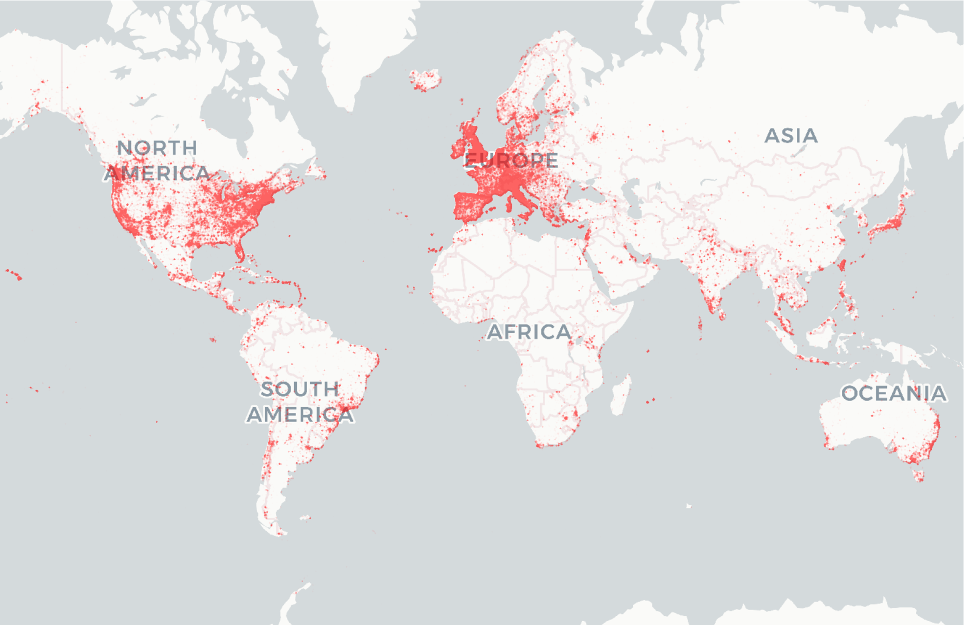
GPS enabled
DSLR cameras



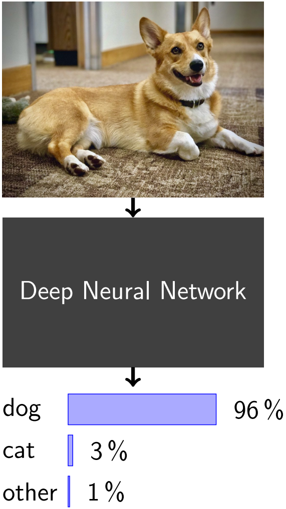
manual entry

Most images accurate to within ∼10m, some outliers

The data (visualized)



Deep learning for image classification



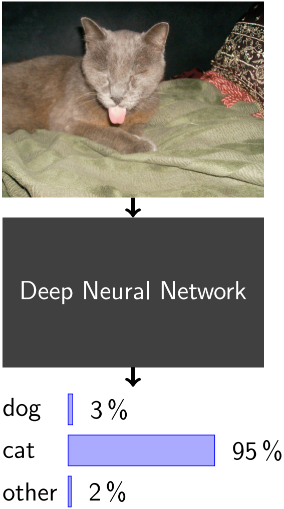
Deep Neural Network

dog 96 %

cat 3 %

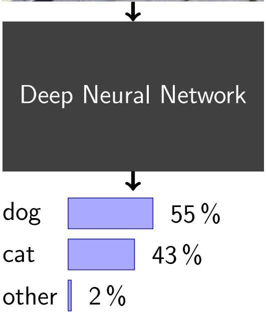
other 1 %

Deep learning for image classification



Deep Neural Network

Deep learning for image classification



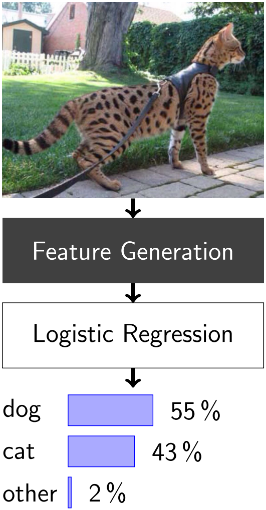
Deep Neural Network

dog 55 %

cat 43 %

other 2 %

Deep learning for image classification



Feature Generation

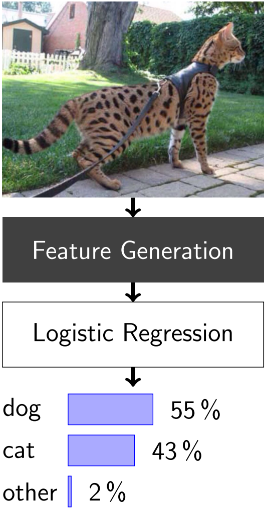
Logistic Regression

dog 55 %

cat 43 %

other 2 %

Deep learning for image classification



Feature Generation

Logistic Regression

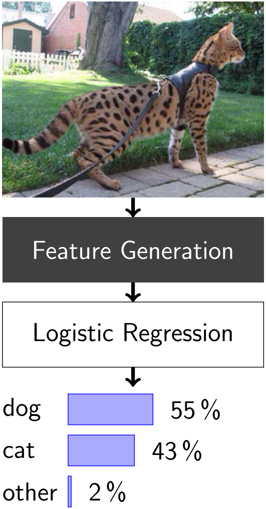
dog 55 %

cat 43 %

other 2 %

LeNet (LeCun et al., 1998)
AlexNet (Krizhevsky et al., 2012)
VGG (Simonyan and Zisserman, 2014)
Inception v1 (Szegedy et al., 2015)
Inception v2 (Ioffe and Szegedy, 2015)
Inception v3 (Szegedy et al., 2016)
Inception v4 (He et al., 2016)
ResNet (Witten et al., 2016)
ResNet2 (He et al., 2016)
WideResNet (Zagoruyko and Komodakis, 2016)
SqueezeNet (Iandola et al., 2016)
MobileNet (Howard et al., 2017)
DenseNet (Huang et al., 2017)
ResNext (Lin et al., 2018)
NASNet (Zoph et al., 2018)
PNASNet (Liu et al., 2018)

Deep learning for image classification

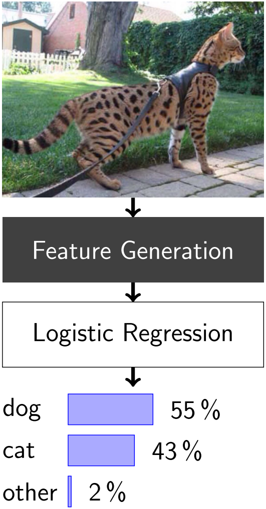


LeNet (LeCun et al., 1998)
AlexNet (Krizhevsky et al., 2012)
VGG (Simonyan and Zisserman, 2014)
Inception v1 (Szegedy et al., 2015)
Inception v2 (Ioffe and Szegedy, 2015)
Inception v3 (Szegedy et al., 2016)
Inception v4 (He et al., 2016)
ResNet (Witten et al., 2016)
ResNet2 (He et al., 2016)
WideResNet (Zagoruyko and Komodakis, 2016)
SqueezeNet (Iandola et al., 2016)
MobileNet (Howard et al., 2017)
DenseNet (Huang et al., 2017)
ResNext (Lin et al., 2018)
NASNet (Zoph et al., 2018)
PNASNet (Liu et al., 2018)

better than humans in practice

treat as black box

Deep learning for image classification



Feature Generation

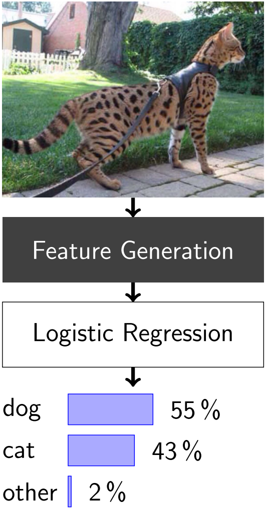
Logistic Regression

dog 55 %

cat 43 %

other 2 %

Deep learning for image classification



Feature Generation

Logistic Regression

dog 55 %

cat 43 %

other 2 %

Known since the 1800s, good theoretical properties

Let d be number of features,
c be number of classes,
n be number of data points,

then the generalization error $= \Theta\left(\sqrt{\frac{cd}{n}}\right)$

So more classes requires more data

Prior methods for image geolocation

Step 1: Divide the world into classes

Step 2: Classify using a deep neural network

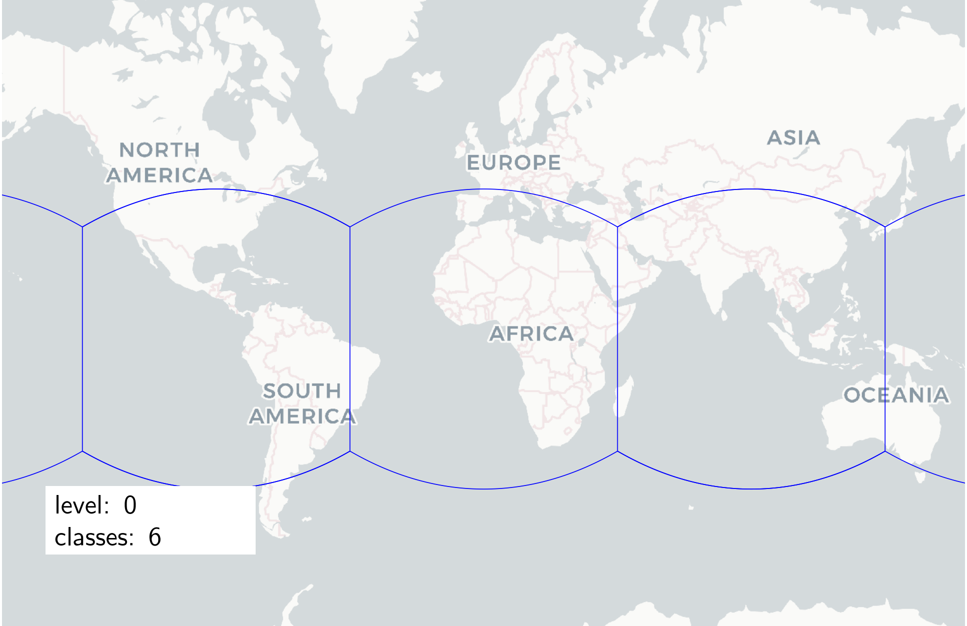
Prior methods for image geolocation

Step 1: Divide the world into classes

Google's PlaNet method (Weyand et al., 2016)

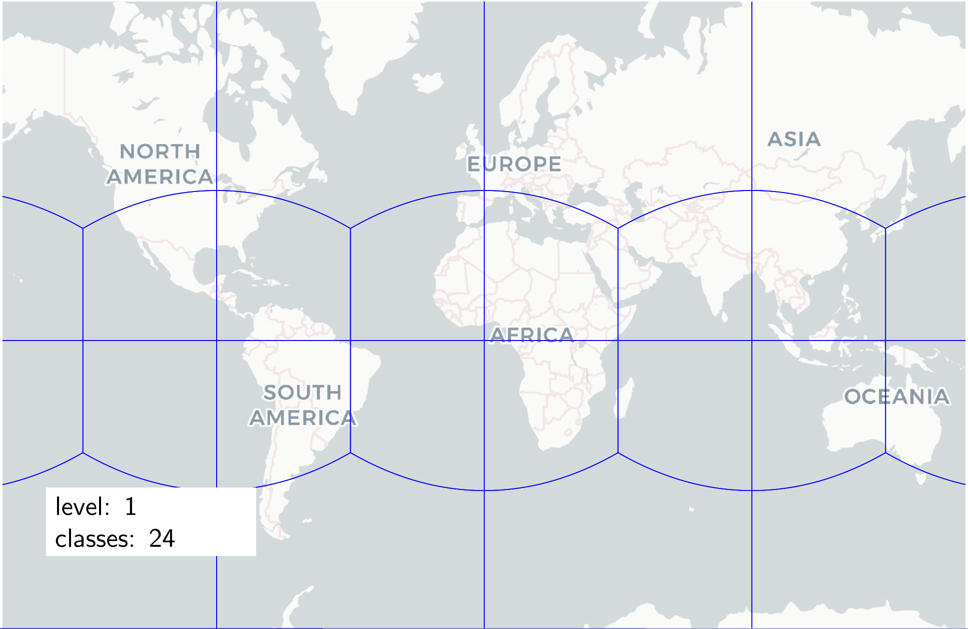
Step 2: Classify using a deep neural network

Generating classes using S2 geometry



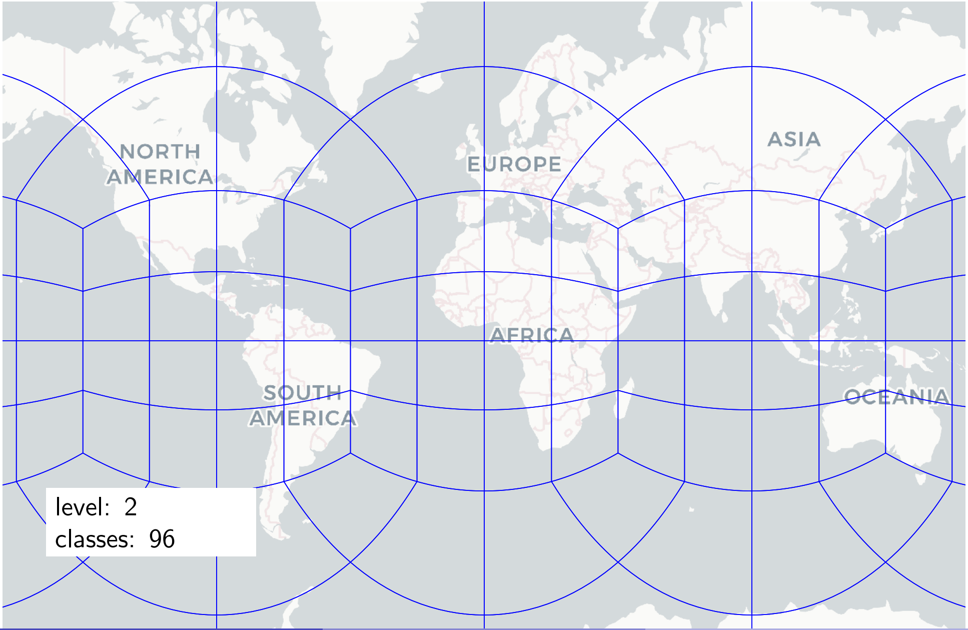
level: 0
classes: 6

Generating classes using S2 geometry



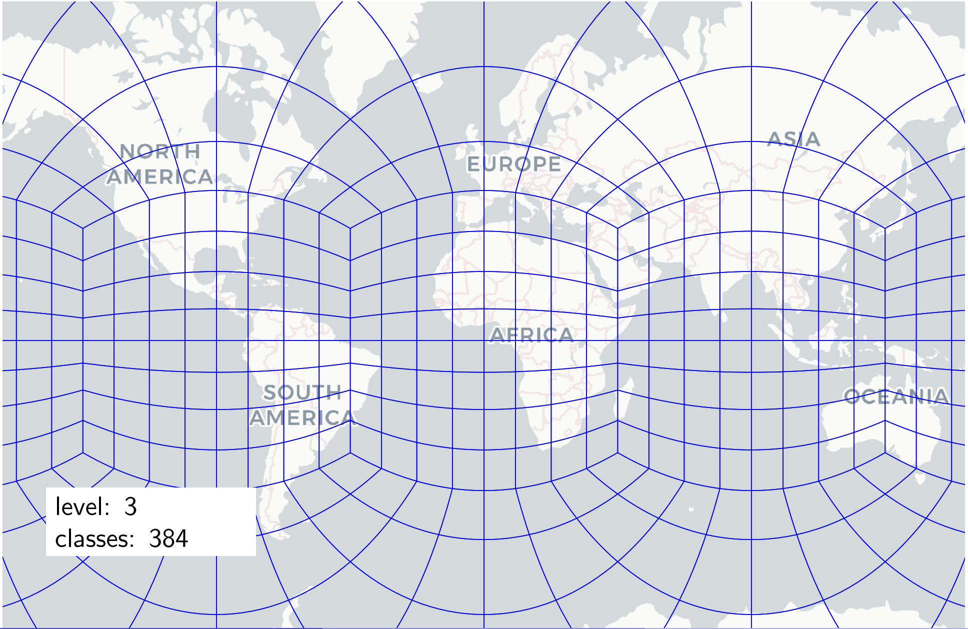
level: 1
classes: 24

Generating classes using S2 geometry



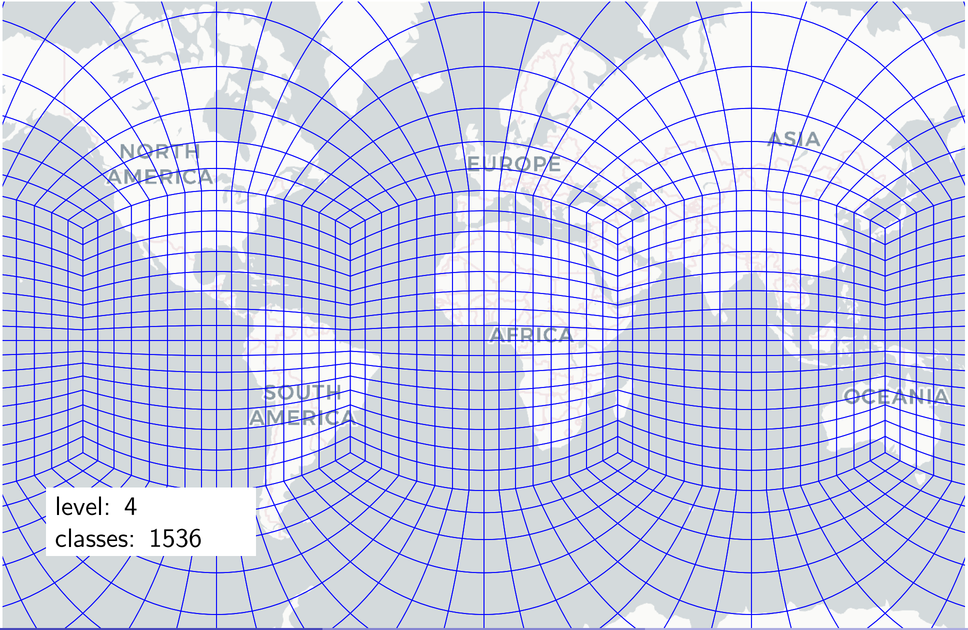
level: 2
classes: 96

Generating classes using S2 geometry



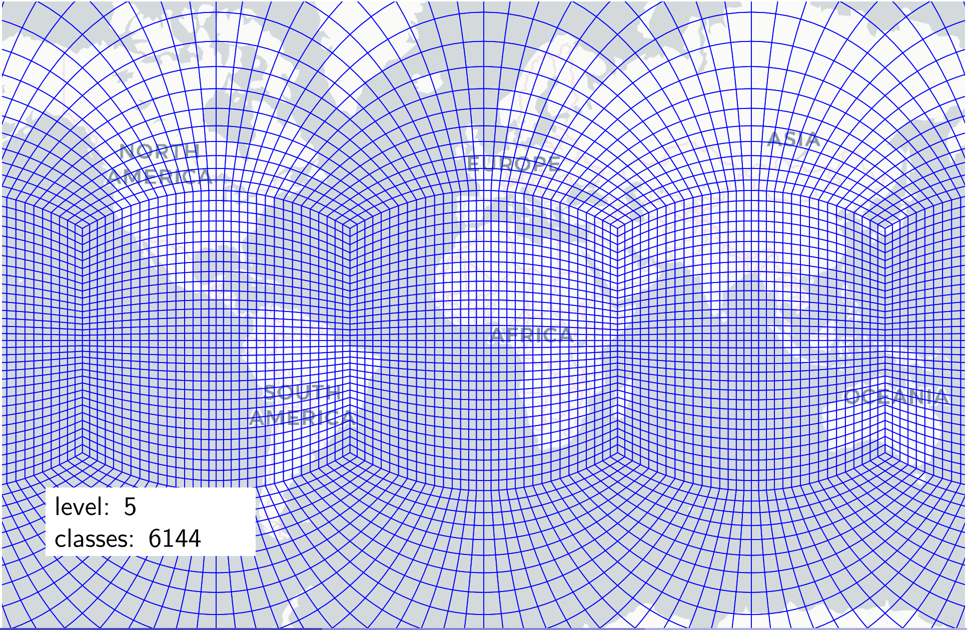
level: 3
classes: 384

Generating classes using S2 geometry



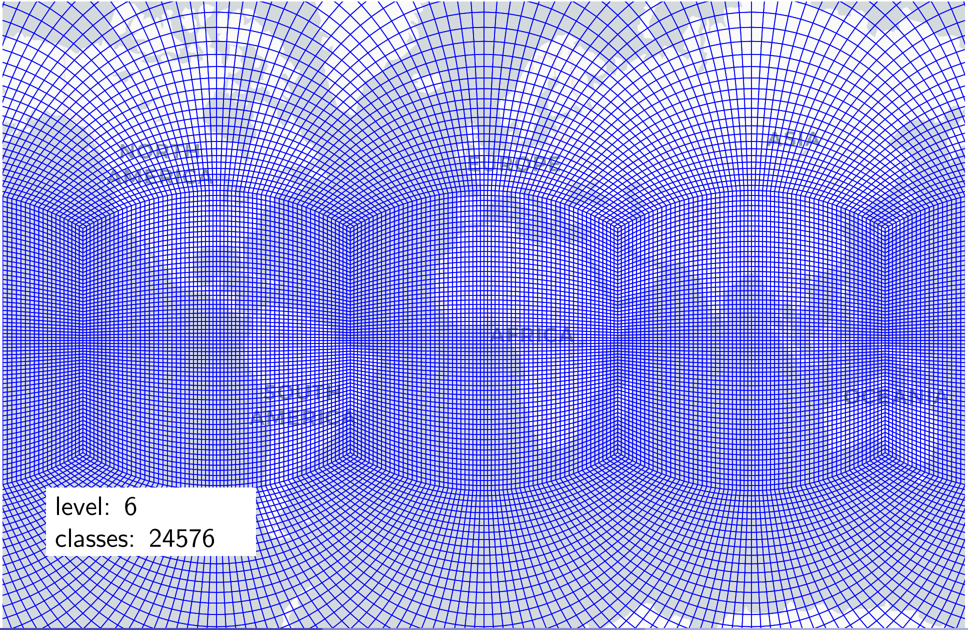
level: 4
classes: 1536

Generating classes using S2 geometry



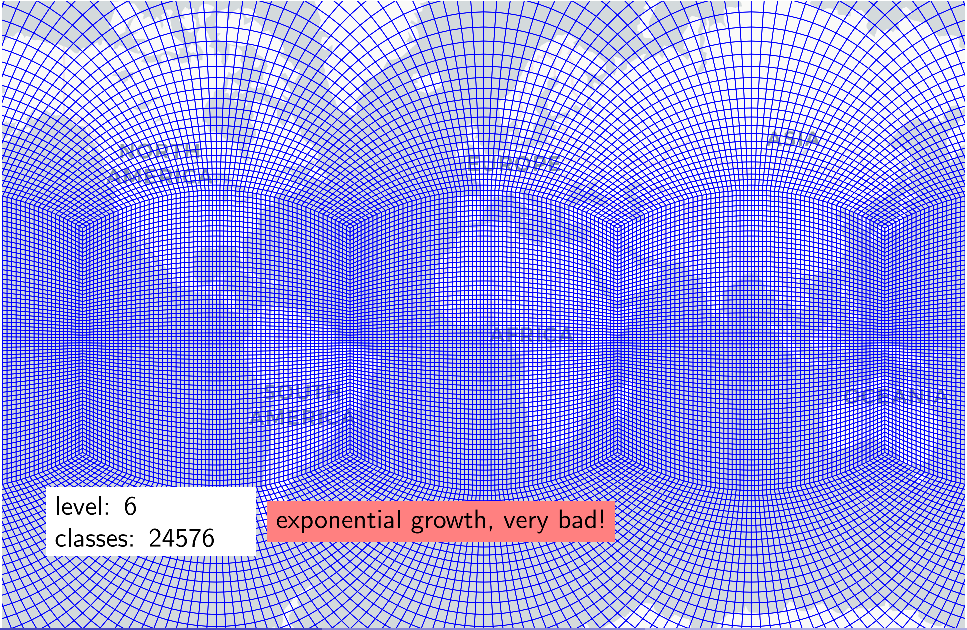
level: 5
classes: 6144

Generating classes using S2 geometry



level: 6
classes: 24576

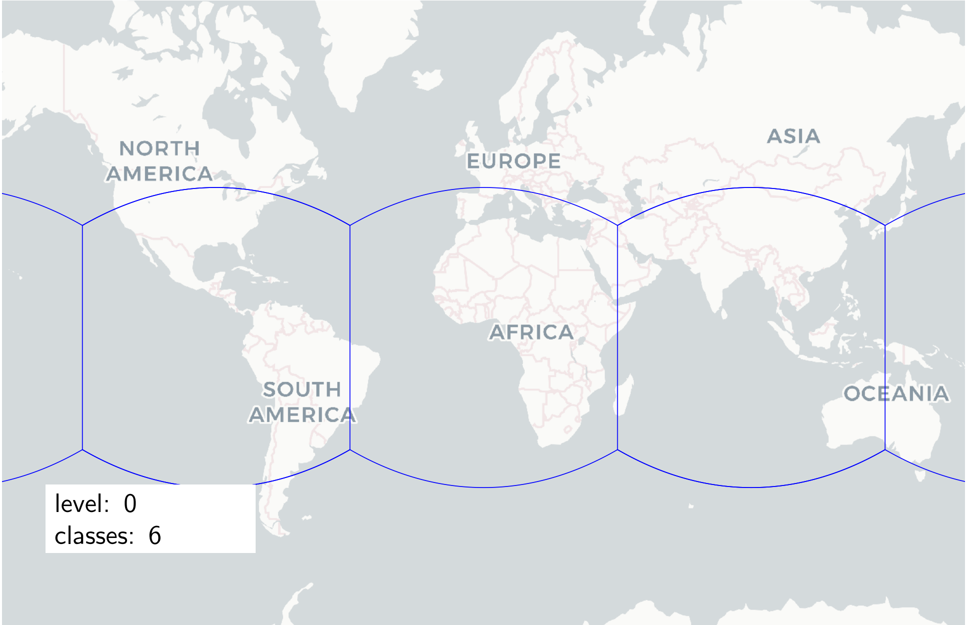
Generating classes using S2 geometry



level: 6
classes: 24576

exponential growth, very bad!

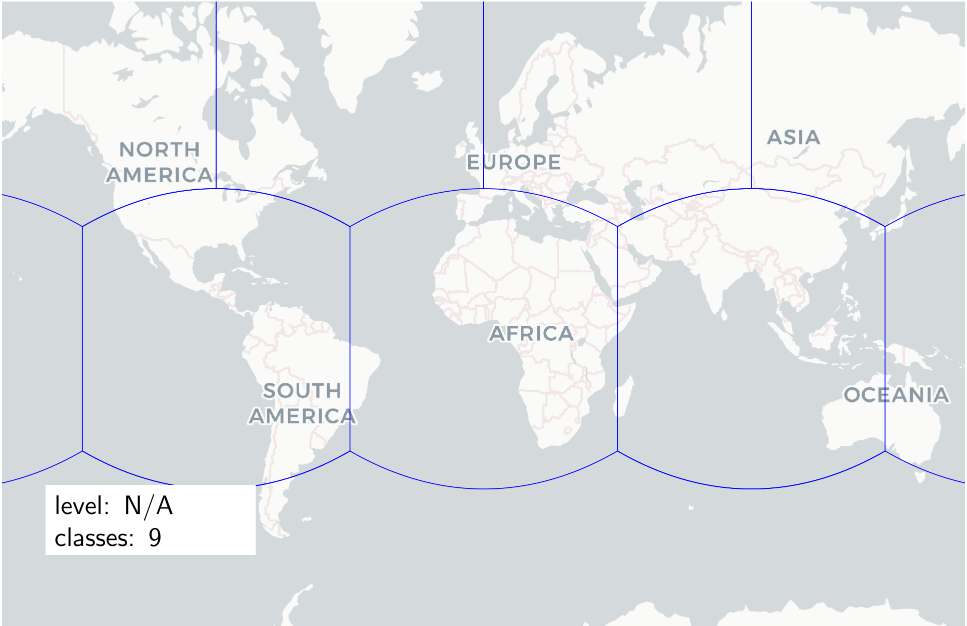
Generating classes using S2 geometry (and data)



level: 0

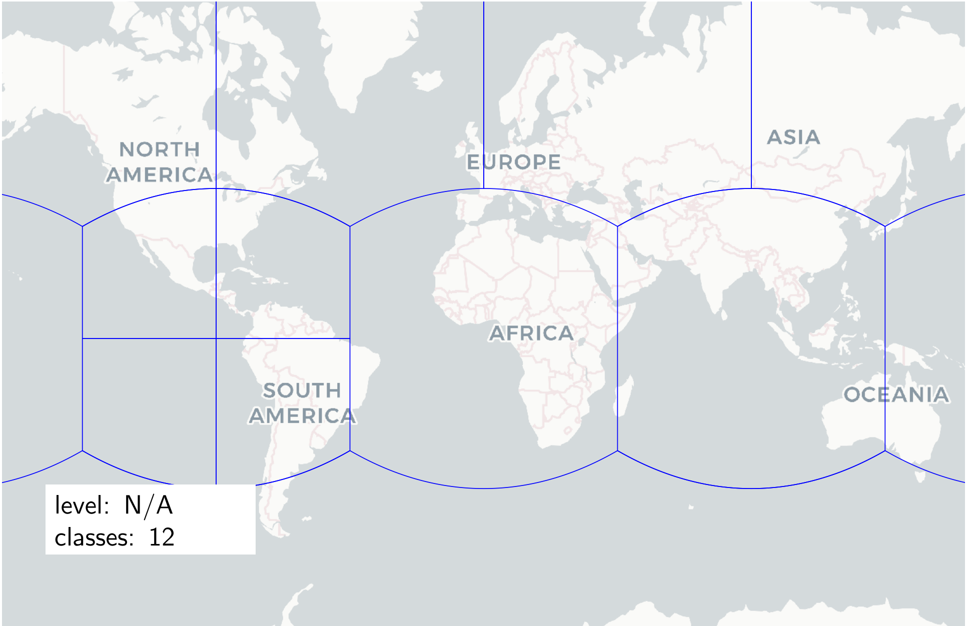
classes: 6

Generating classes using S2 geometry (and data)



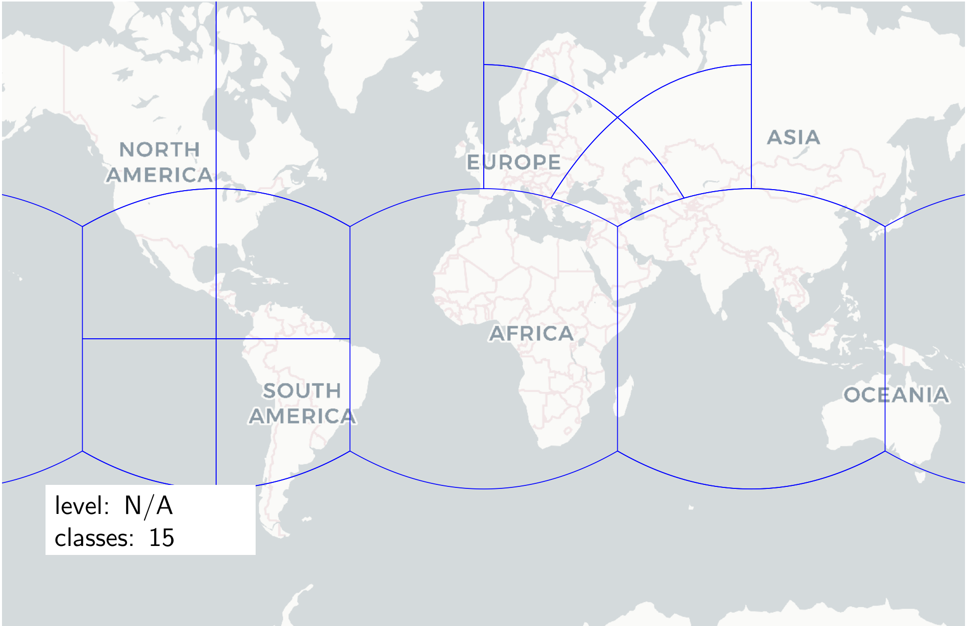
level: N/A
classes: 9

Generating classes using S2 geometry (and data)



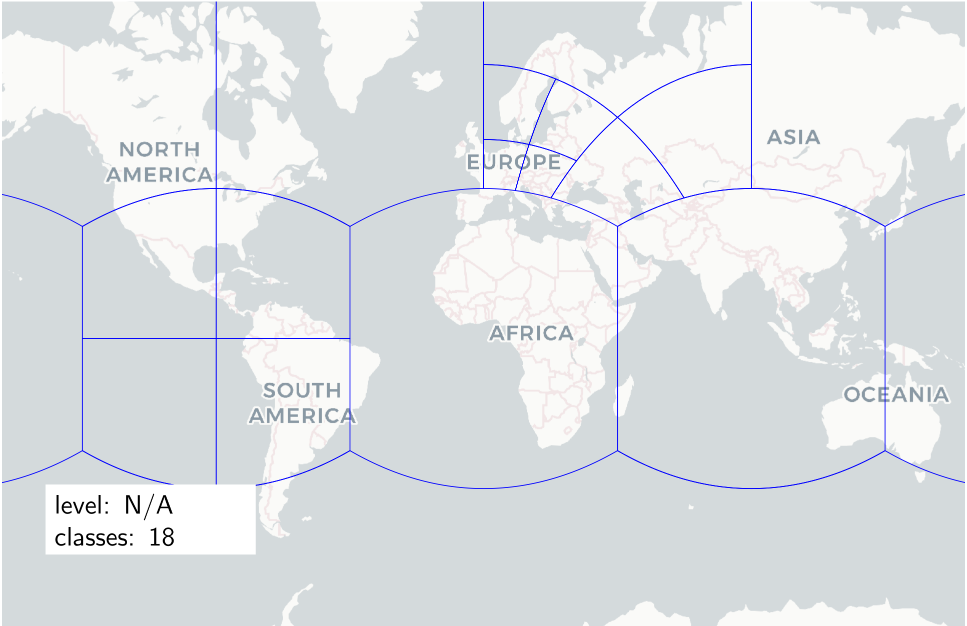
level: N/A
classes: 12

Generating classes using S2 geometry (and data)



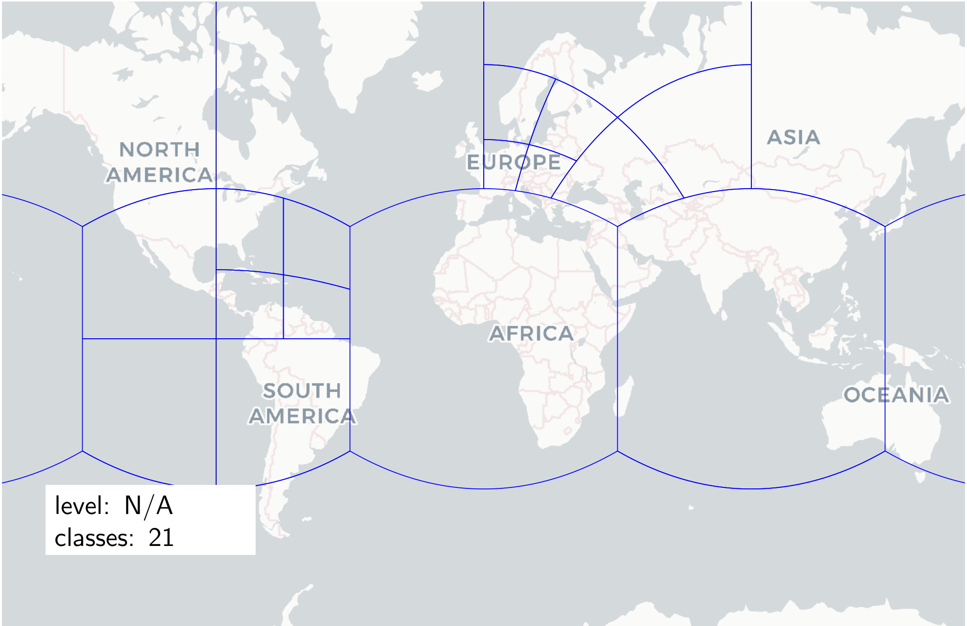
level: N/A
classes: 15

Generating classes using S2 geometry (and data)



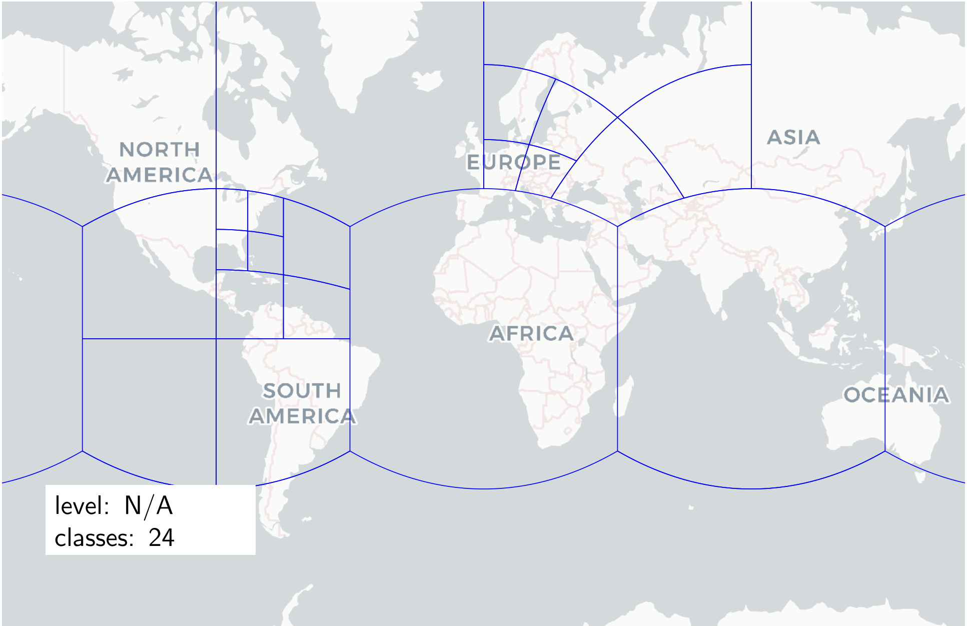
level: N/A
classes: 18

Generating classes using S2 geometry (and data)



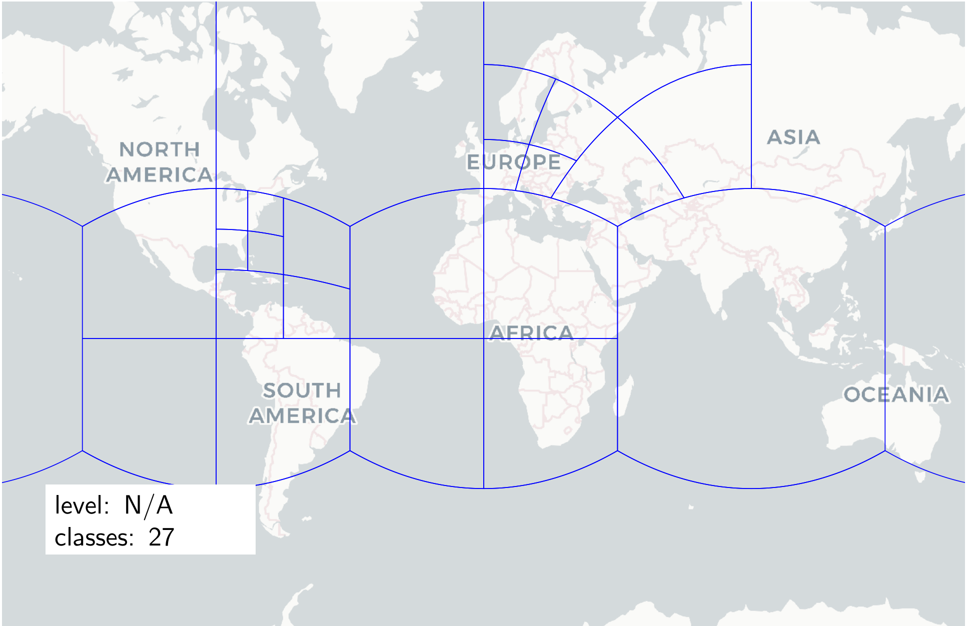
level: N/A
classes: 21

Generating classes using S2 geometry (and data)



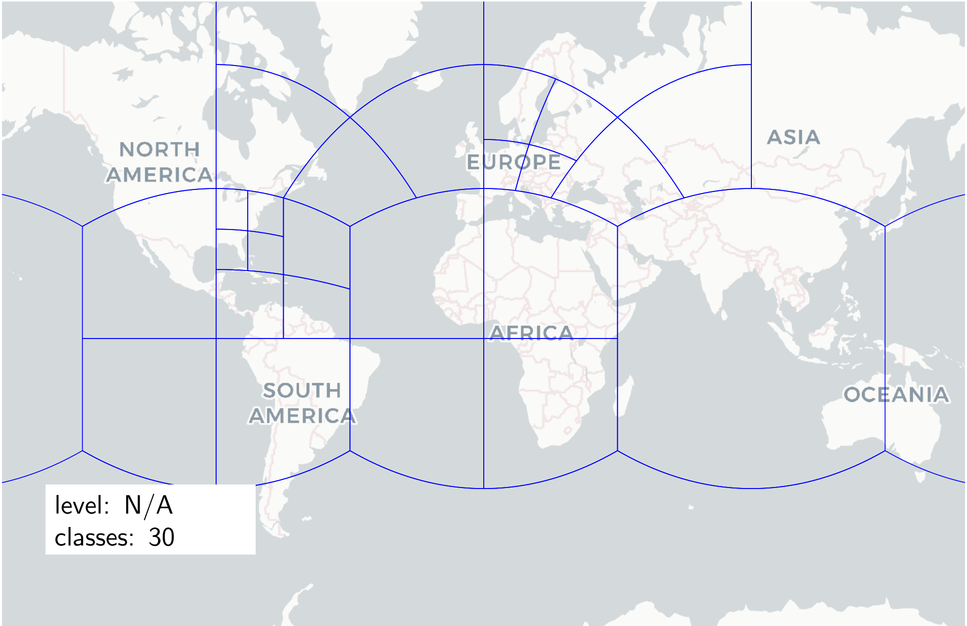
level: N/A
classes: 24

Generating classes using S2 geometry (and data)



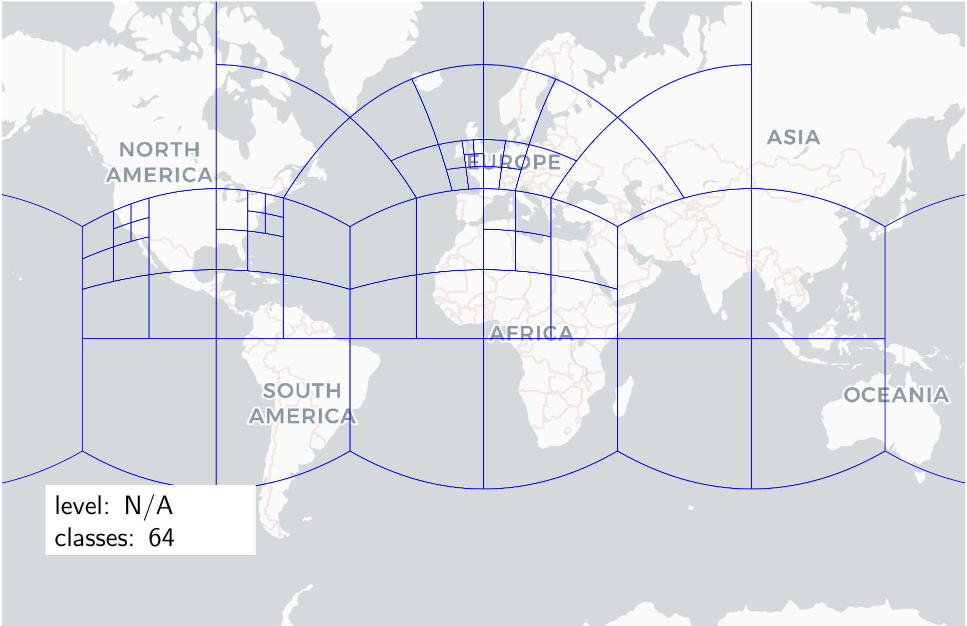
level: N/A
classes: 27

Generating classes using S2 geometry (and data)



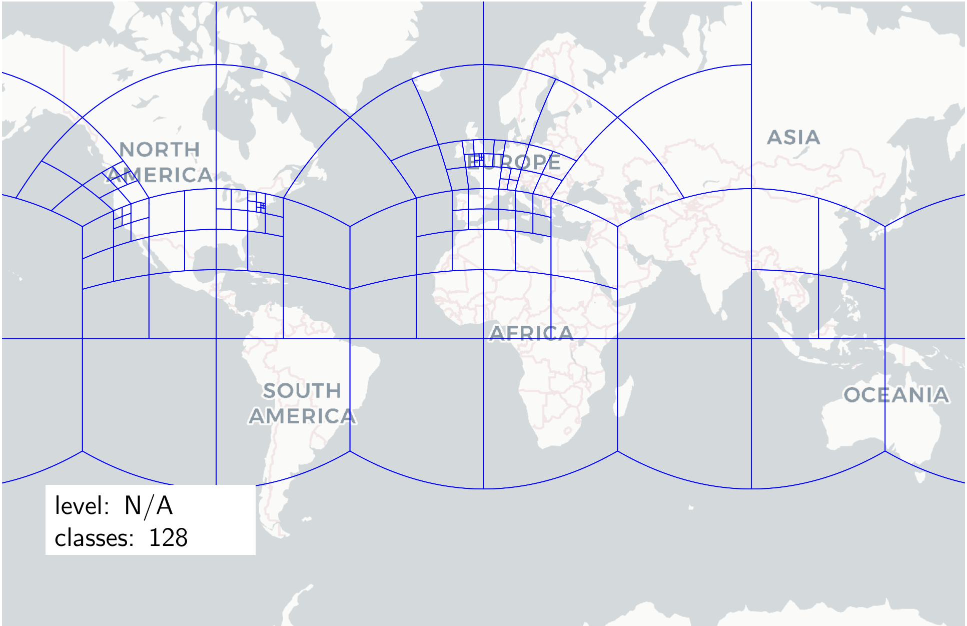
level: N/A
classes: 30

Generating classes using S2 geometry (and data)



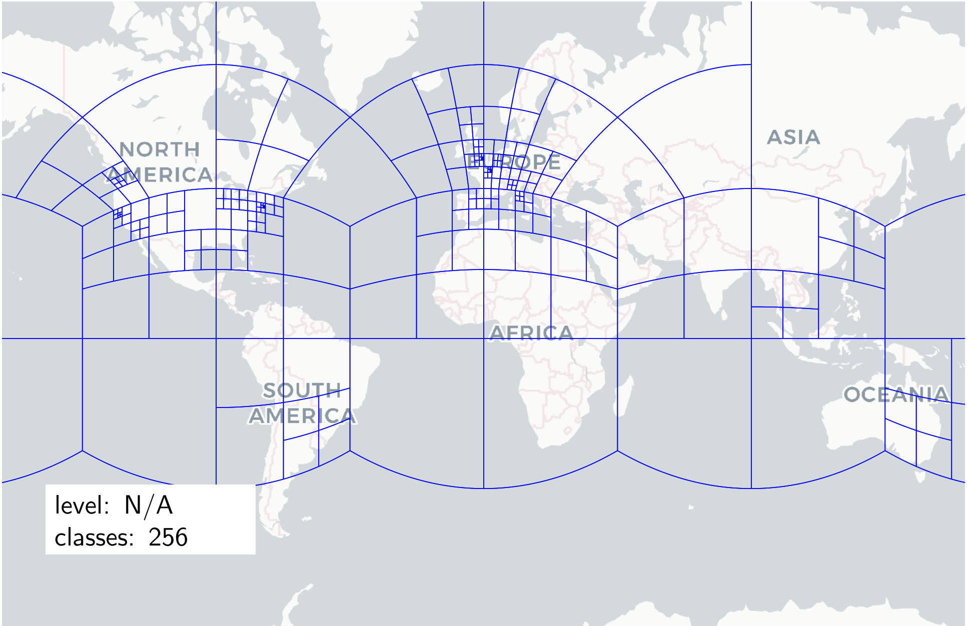
level: N/A
classes: 64

Generating classes using S2 geometry (and data)



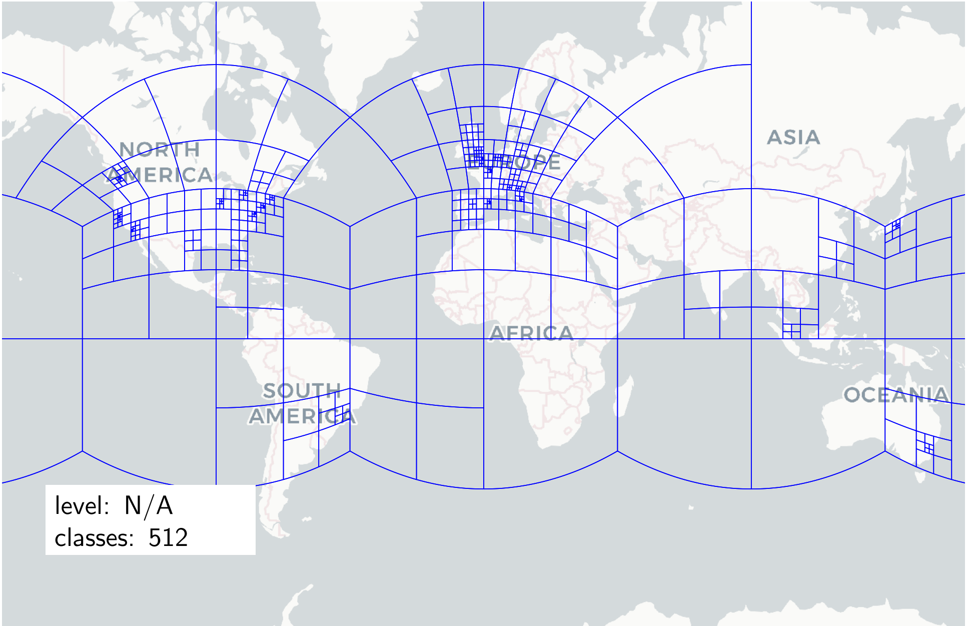
level: N/A
classes: 128

Generating classes using S2 geometry (and data)



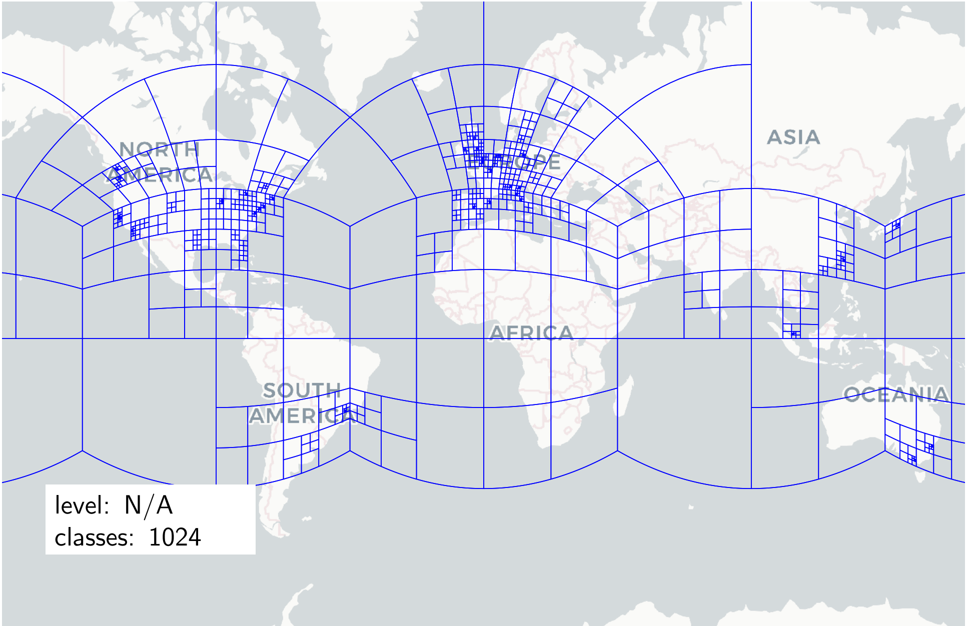
level: N/A
classes: 256

Generating classes using S2 geometry (and data)



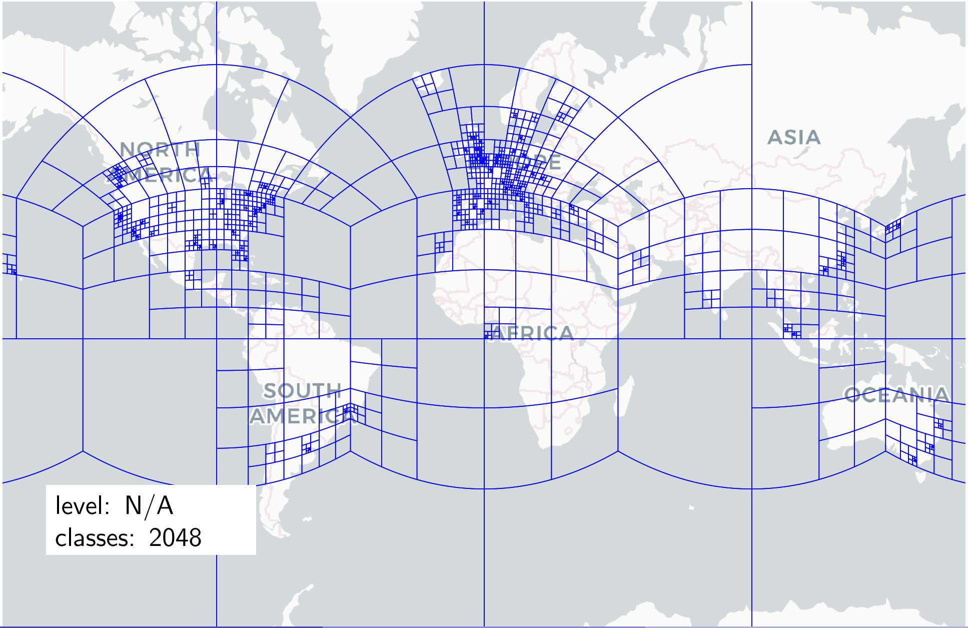
level: N/A
classes: 512

Generating classes using S2 geometry (and data)



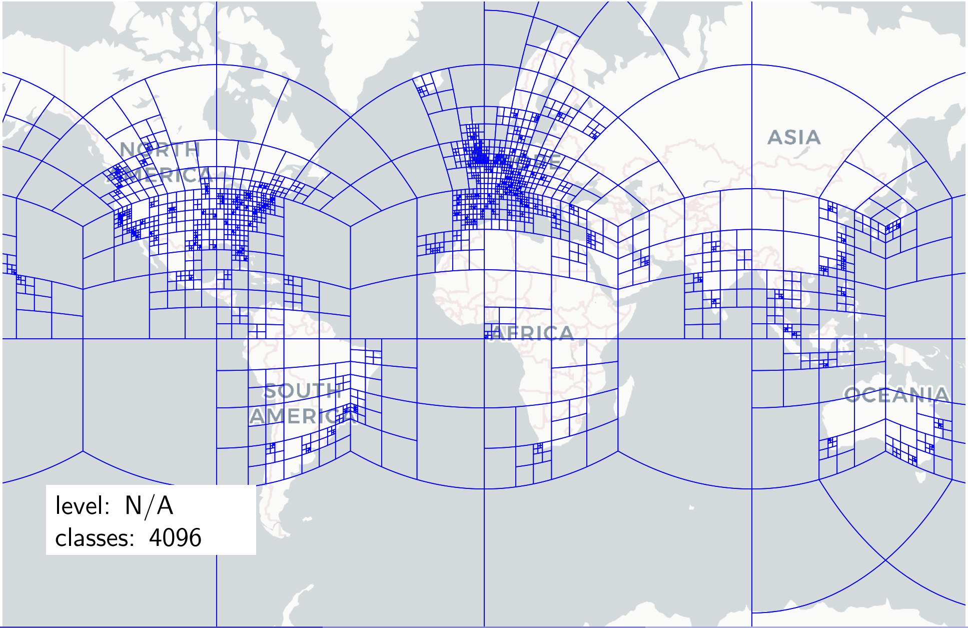
level: N/A
classes: 1024

Generating classes using S2 geometry (and data)



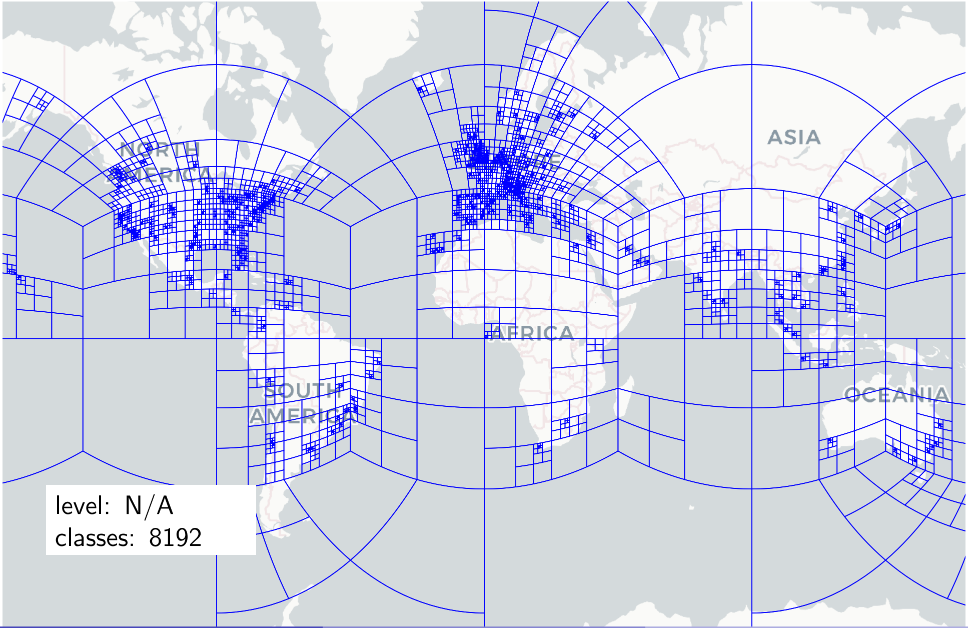
level: N/A
classes: 2048

Generating classes using S2 geometry (and data)



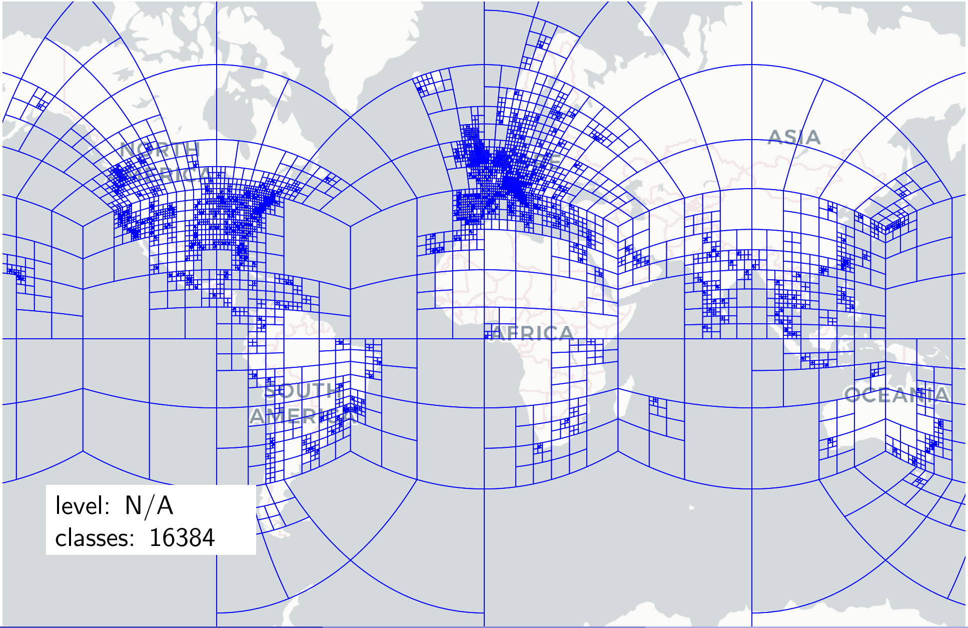
level: N/A
classes: 4096

Generating classes using S2 geometry (and data)



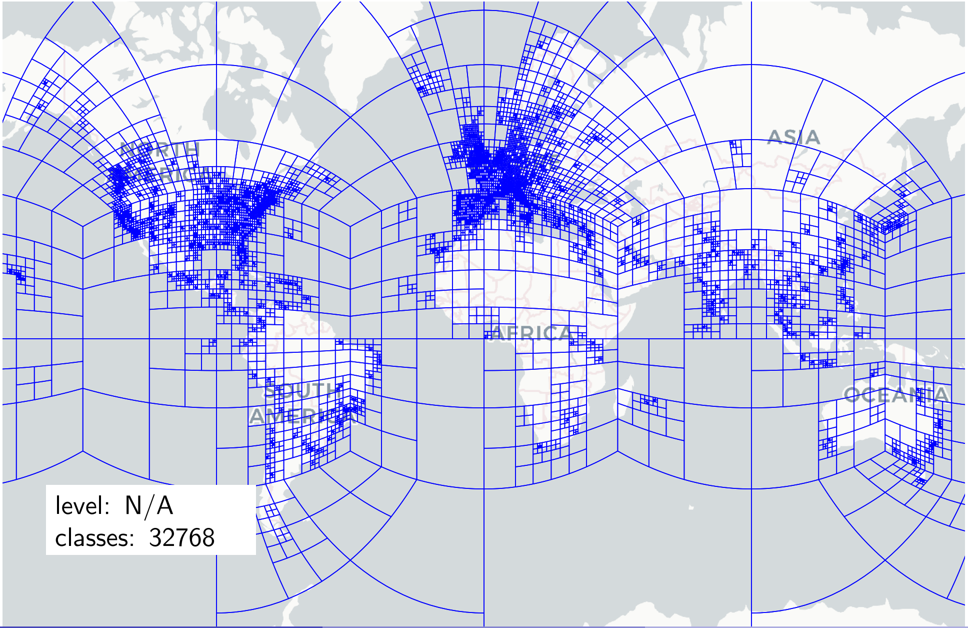
level: N/A
classes: 8192

Generating classes using S2 geometry (and data)



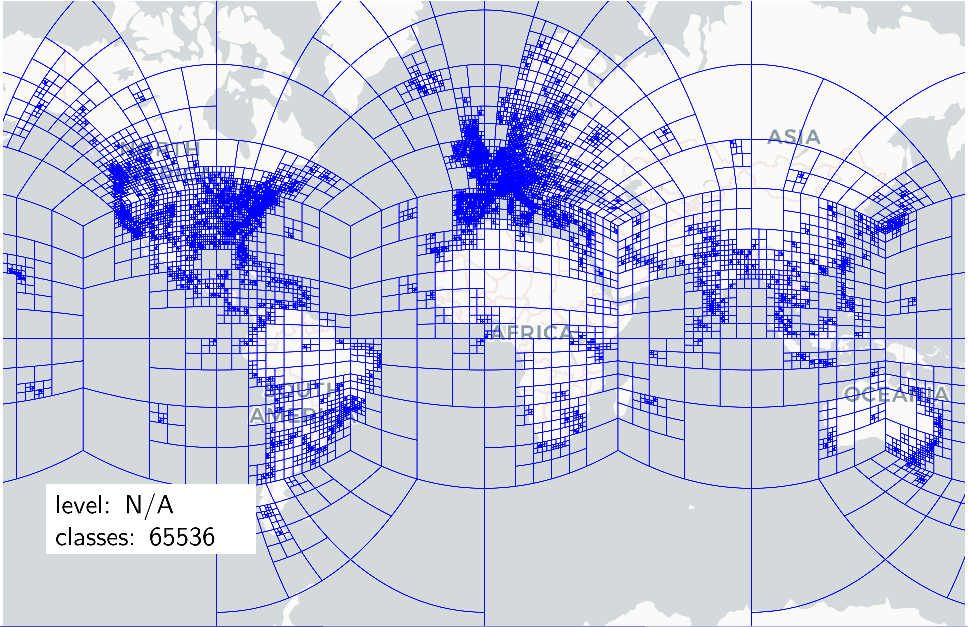
level: N/A
classes: 16384

Generating classes using S2 geometry (and data)



level: N/A
classes: 32768

Generating classes using S2 geometry (and data)

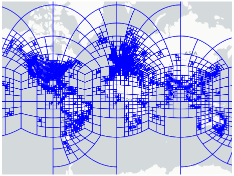


level: N/A
classes: 65536

Prior methods for image geolocation

Step 1: Divide the world into classes

Google's PlaNet method (Weyand et al., 2016)

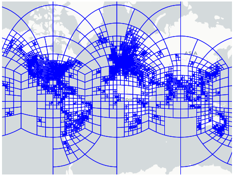


Step 2: Classify using a deep neural network

Prior methods for image geolocation

Step 1: Divide the world into classes

Google's PlaNet method (Weyand et al., 2016)



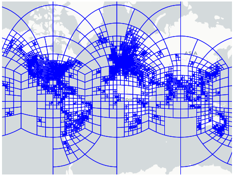
Step 2: Classify using a deep neural network

Advantage: easy to implement

Prior methods for image geolocation

Step 1: Divide the world into classes

Google's PlaNet method (Weyand et al., 2016)

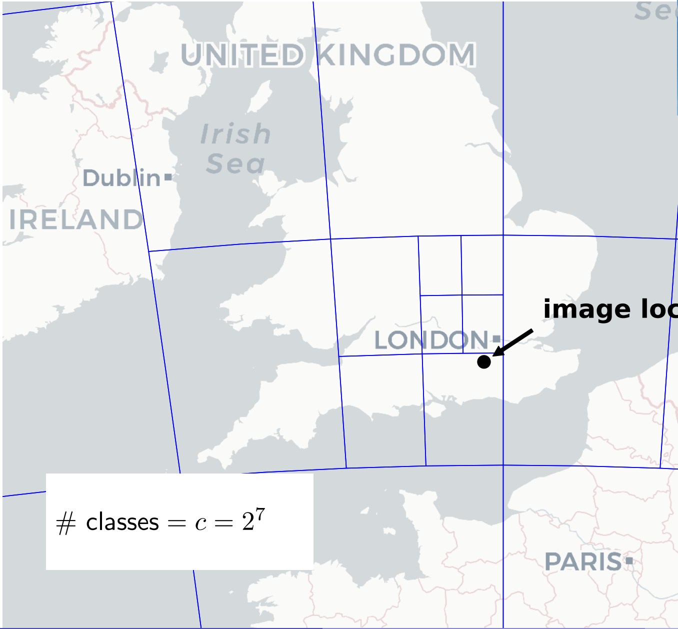


Step 2: Classify using a deep neural network

Advantage: easy to implement

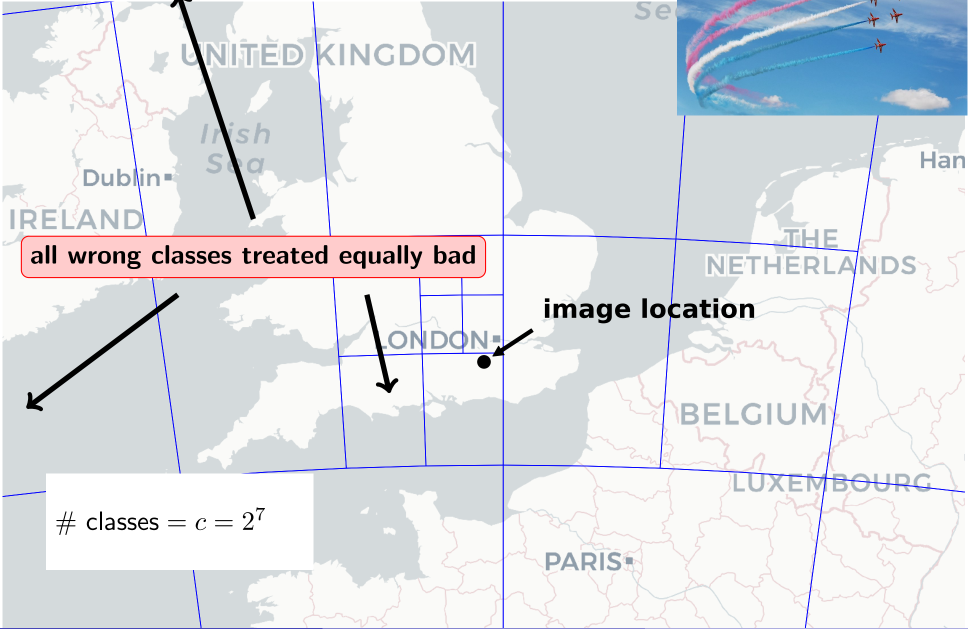
Problem: does not understand the earth's geometry

Does not understand earth's geometry



# classes $= c = 2^7$

Does not understand earth's geometry



all wrong classes treated equally bad

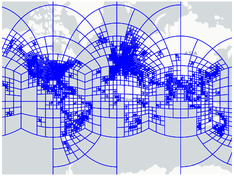
image location

$\#$ classes $= c = 2^7$

Prior methods for image geolocation

Step 1: Divide the world into classes

Google's PlaNet method (Weyand et al., 2016)



Step 2: Classify using a deep neural network

Advantage: easy to implement

Problem: does not understand the earth's geometry

Problem: optimal choice of c unclear

if $c \uparrow$, then the resolution ↑

if $c \uparrow$, then the generalization error ↑

How to choose c?

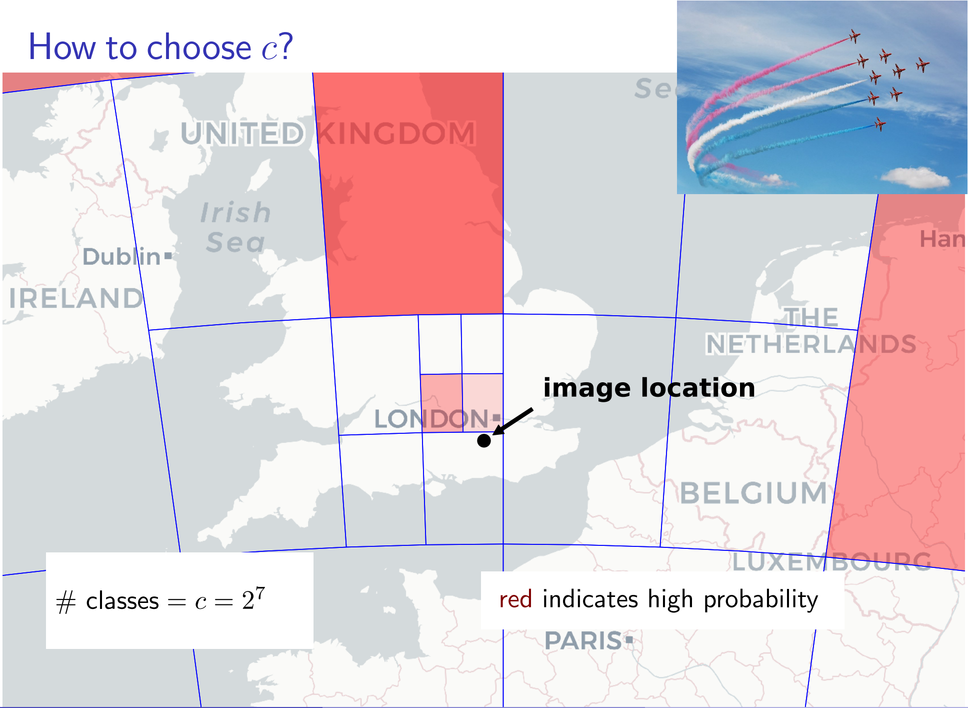
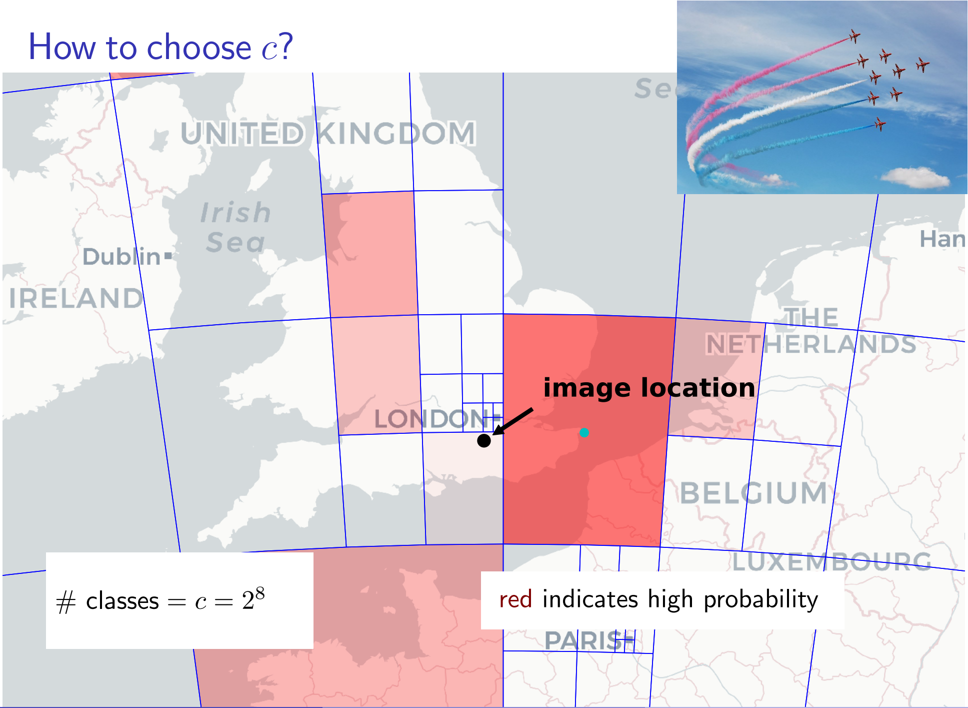


image location

$\#$ classes $= c = 2^7$

red indicates high probability

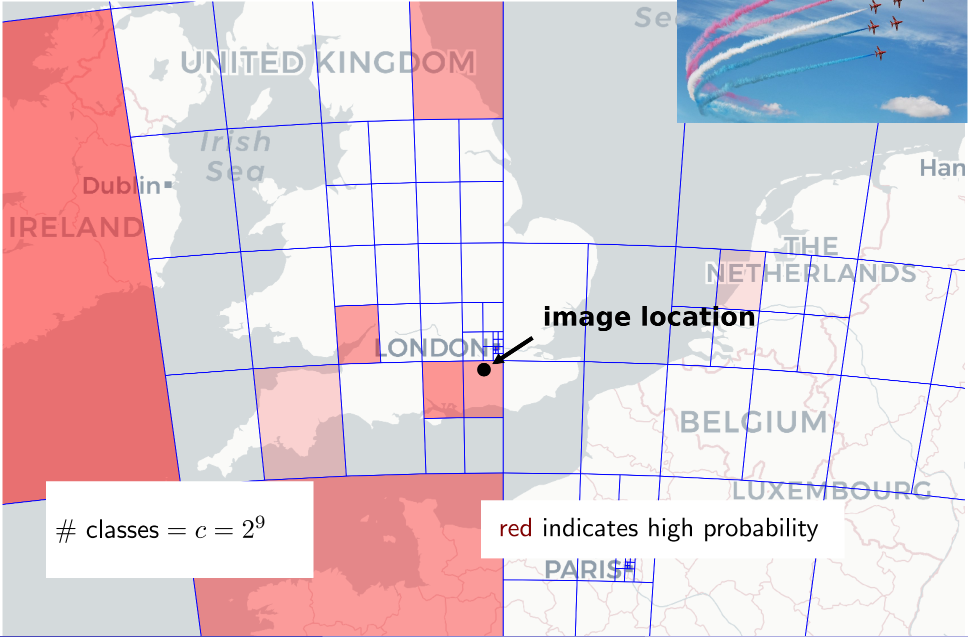
How to choose c?



$\#$ classes $= c = 2^8$

red indicates high probability

How to choose c?



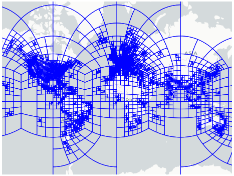
$\#$ classes $= c = 2^9$

red indicates high probability

Prior methods for image geolocation

Step 1: Divide the world into classes

Google's PlaNet method (Weyand et al., 2016)



Step 2: Classify using a deep neural network

Advantage: easy to implement

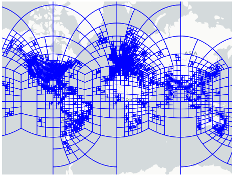
Problem: does not understand the earth's geometry

Problem: optimal choice of c unclear

if $c\uparrow$, then the resolution $\uparrow$

if $c\uparrow$, then the generalization error $\uparrow$

Prior methods for image geolocation



Step 1: Divide the world into classes

Google's PlaNet method (Weyand et al., 2016)

Step 2: Classify using a deep neural network

Advantage: easy to implement

Problem: does not understand the earth's geometry

Problem: optimal choice of c unclear

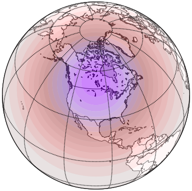
if $c \uparrow$, then the resolution $\uparrow$

if $c \uparrow$, then the generalization error $\uparrow$

The Mixture of von Mises-Fisher distribution fixes these problems

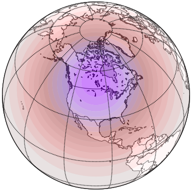
What is the Mixture of von Mises-Fisher Distribution?

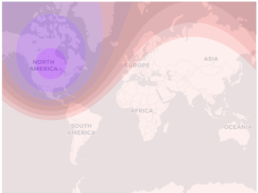
The *von Mises-Fisher* (vMF) distribution is like the Gaussian for spheres



What is the Mixture of von Mises-Fisher Distribution?

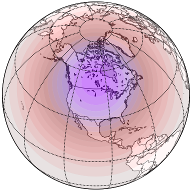
The *von Mises-Fisher* (vMF) distribution is like the Gaussian for spheres

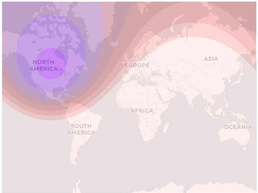




What is the Mixture of von Mises-Fisher Distribution?

The *von Mises-Fisher* (vMF) distribution is like the Gaussian for spheres



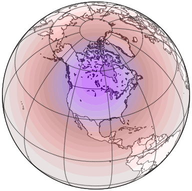


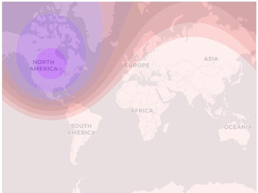
$$\text{vMF}(\mathbf{y}; \mu, \kappa) = \frac{\kappa}{\sinh \kappa} \exp(\kappa \mathbf{y}^\top \mu)$$

μ : mean direction
κ : concentration

What is the Mixture of von Mises-Fisher Distribution?

The *von Mises-Fisher* (vMF) distribution is like the Gaussian for spheres



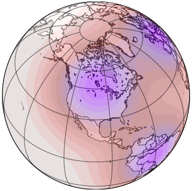


$$\text{vMF}(\mathbf{y};\mu,\kappa) = \frac{\kappa}{\sinh\kappa}\exp(\kappa\mathbf{y}^\top\mu)$$

μ : mean direction

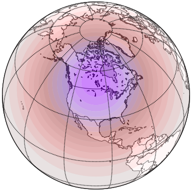
κ : concentration

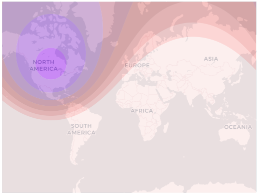
The *Mixture of vMF* (MvMF) is a weighted sum of vMF distributions



What is the Mixture of von Mises-Fisher Distribution?

The *von Mises-Fisher* (vMF) distribution is like the Gaussian for spheres

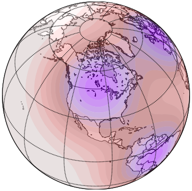


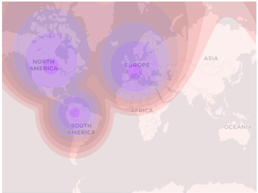


$$\text{vMF}(\mathbf{y};\mu,\kappa) = \frac{\kappa}{\sinh\kappa}\exp(\kappa\mathbf{y}^\top\mu)$$

μ : mean direction
κ : concentration

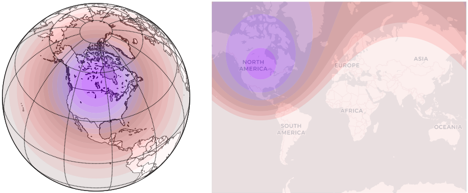
The *Mixture of vMF* (MvMF) is a weighted sum of vMF distributions





What is the Mixture of von Mises-Fisher Distribution?

The *von Mises-Fisher* (vMF) distribution is like the Gaussian for spheres

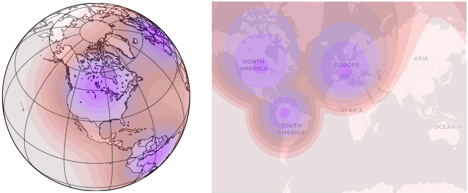


$$\text{vMF}(\mathbf{y};\mu,\kappa) = \frac{\kappa}{\sinh\kappa}\exp(\kappa\mathbf{y}^\top\mu)$$

μ : mean direction

κ : concentration

The *Mixture of vMF* (MvMF) is a weighted sum of vMF distributions

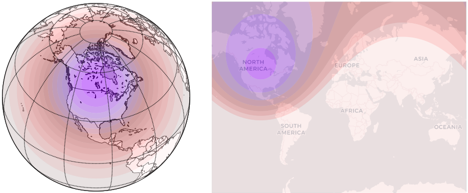


$$\text{MvMF}(\mathbf{y}) = \sum_{i=1}^{c}\mathbf{w}_i\,\text{vMF}(\mathbf{y};\mu_i,\kappa_i)$$

$\mathbf{w}_i$: mixture weights

What is the Mixture of von Mises-Fisher Distribution?

The *von Mises-Fisher* (vMF) distribution is like the Gaussian for spheres

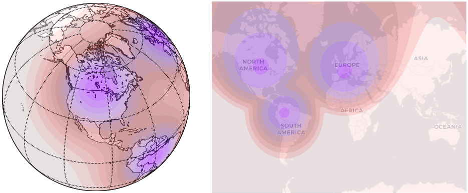


$$\text{vMF}(\mathbf{y};\mu,\kappa) = \frac{\kappa}{\sinh\kappa}\exp(\kappa\mathbf{y}^\top\mu)$$

μ : mean direction

κ : concentration

The *Mixture of vMF* (MvMF) is a weighted sum of vMF distributions



$$\text{MvMF}(\mathbf{y}) = \sum_{i=1}^{c}\mathbf{w}_i\,\text{vMF}(\mathbf{y};\mu_i,\kappa_i)$$

$\mathbf{w}_i$: mixture weights

Idea: represent each class by a vMF

Geolocation with the MvMF

Step 1: divide the world into classes

Google's PlaNet method (Weyand et al., 2016)

Step 2: initialize an MvMF distribution with μ_i determined by classes

Step 3: learn the mixture weights

Geolocation with the MvMF

Step 1: divide the world into classes

Google's PlaNet method (Weyand et al., 2016)

Step 2: initialize an MvMF distribution with μ_i determined by classes

Step 3: learn the mixture weights

Advantage: easy to implement

Geolocation with the MvMF

Step 1: divide the world into classes

Google's PlaNet method (Weyand et al., 2016)

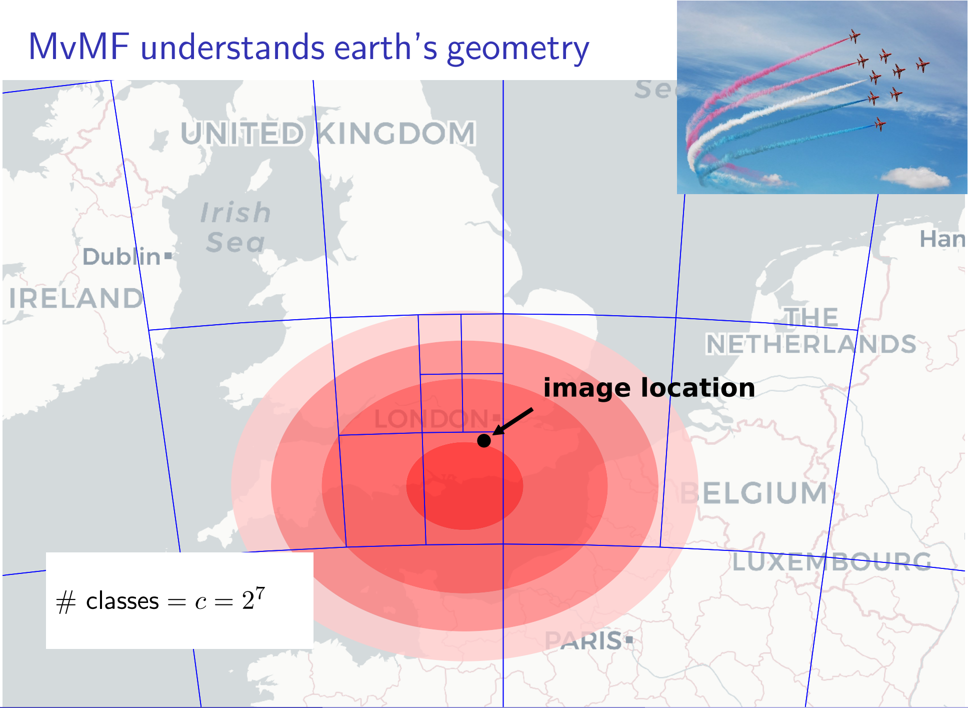
Step 2: initialize an MvMF distribution with μ_i determined by classes

Step 3: learn the mixture weights

Advantage: easy to implement

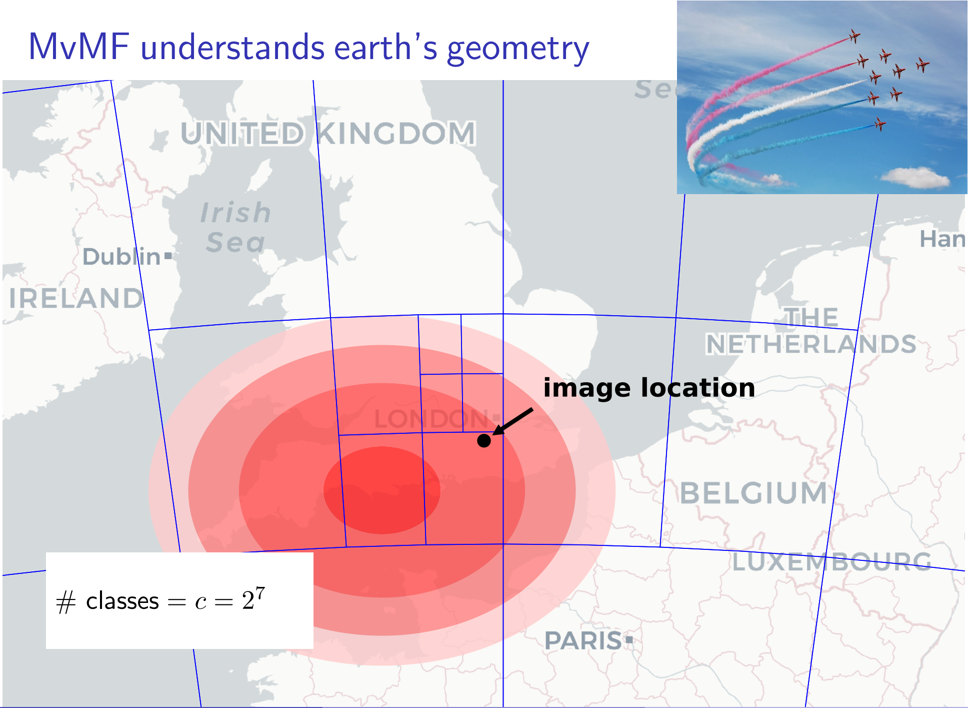
Advantage: does understand the earth's geometry

MvMF understands earth's geometry



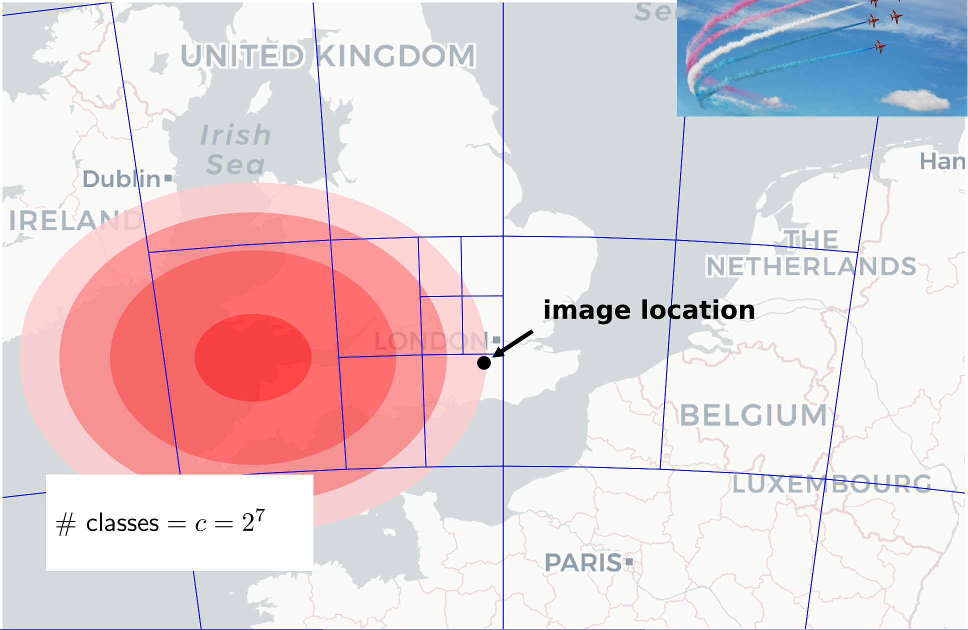
$\#$ classes $= c = 2^7$

MvMF understands earth's geometry



$\#$ classes $= c = 2^7$

MvMF understands earth's geometry



# classes $= c = 2^7$

Geolocation with the MvMF

Step 1: divide the world into classes

Google's PlaNet method (Weyand et al., 2016)

Step 2: initialize an MvMF distribution with μ_i determined by classes

Step 3: learn the mixture weights

Advantage: easy to implement

Advantage: does understand the earth's geometry

Geolocation with the MvMF

Step 1: divide the world into classes

Google's PlaNet method (Weyand et al., 2016)

Step 2: initialize an MvMF distribution with μ_i determined by classes

Step 3: learn the mixture weights

Advantage: easy to implement

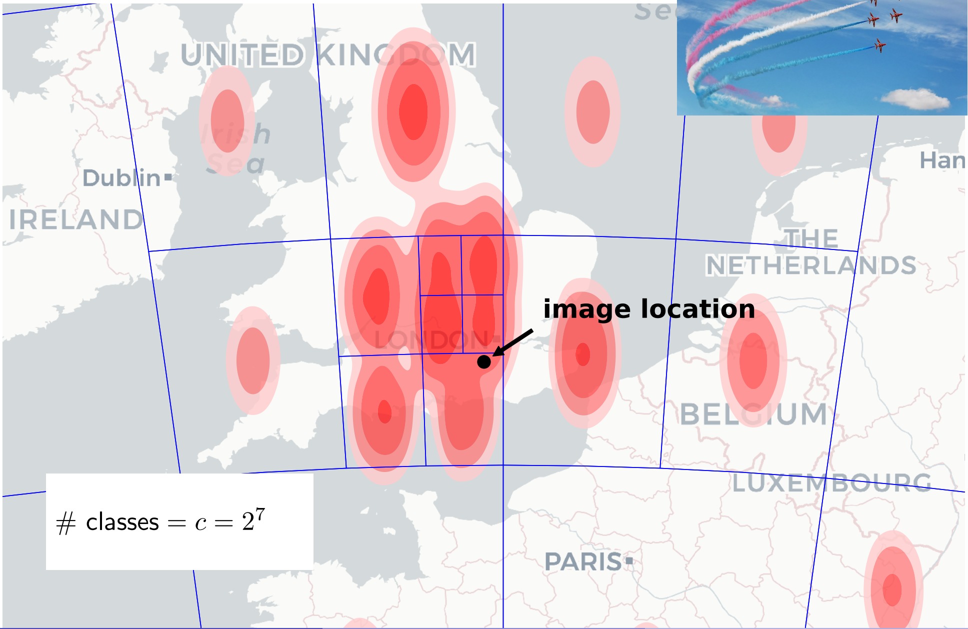
Advantage: does understand the earth's geometry

Advantage: increasing c always improves statistical performance

in practice

in theory

The MvMF and c (practice)



$\#$ classes $= c = 2^7$

The MvMF and c (practice)

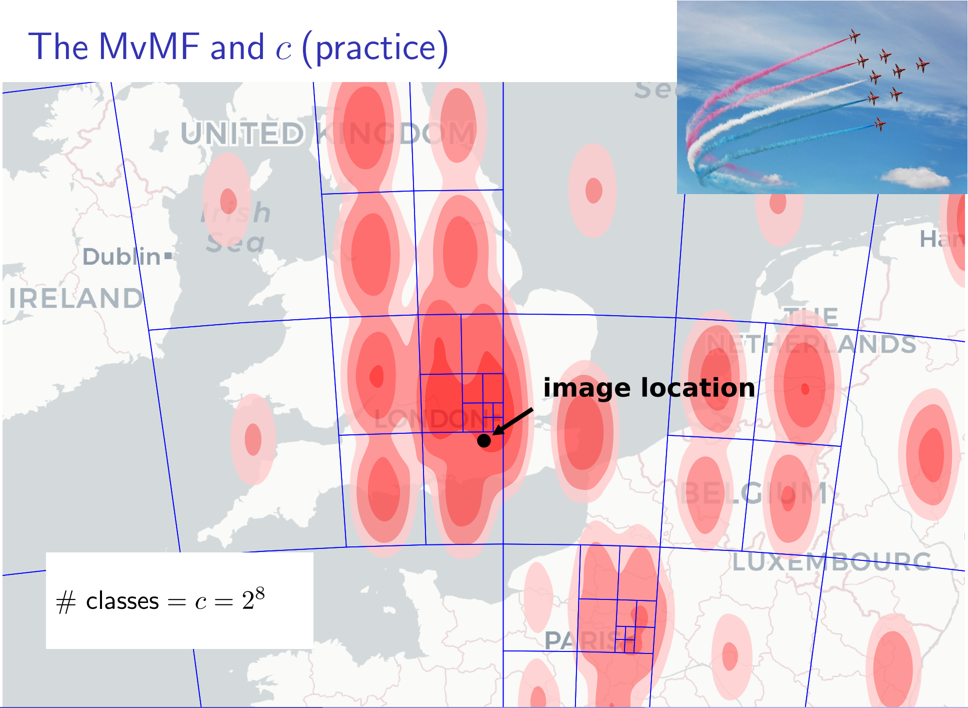


image location

$\#$ classes $= c = 2^8$

The MvMF and c (practice)

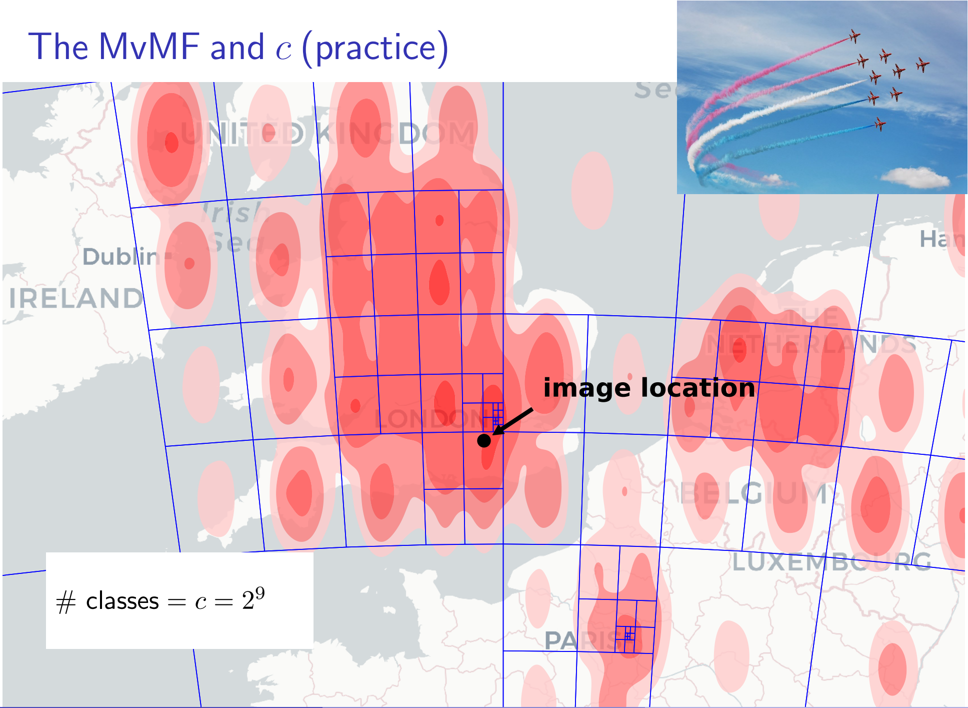


image location

$\#$ classes $= c = 2^9$

The MvMF and c (practice)

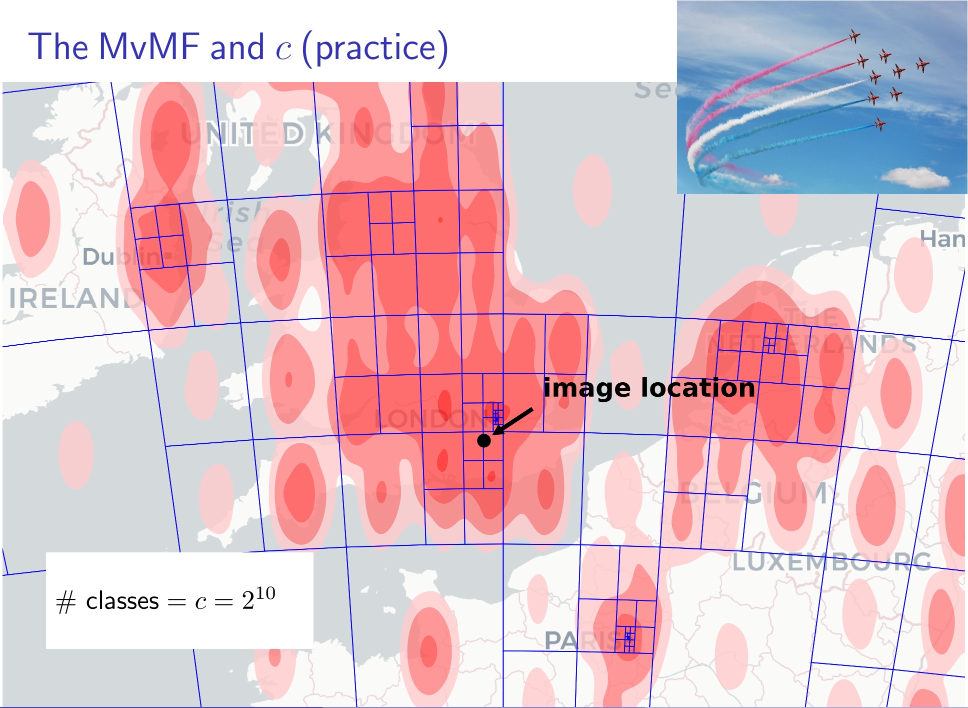


image location

$\#$ classes $= c = 2^{10}$

The MvMF and c (practice)

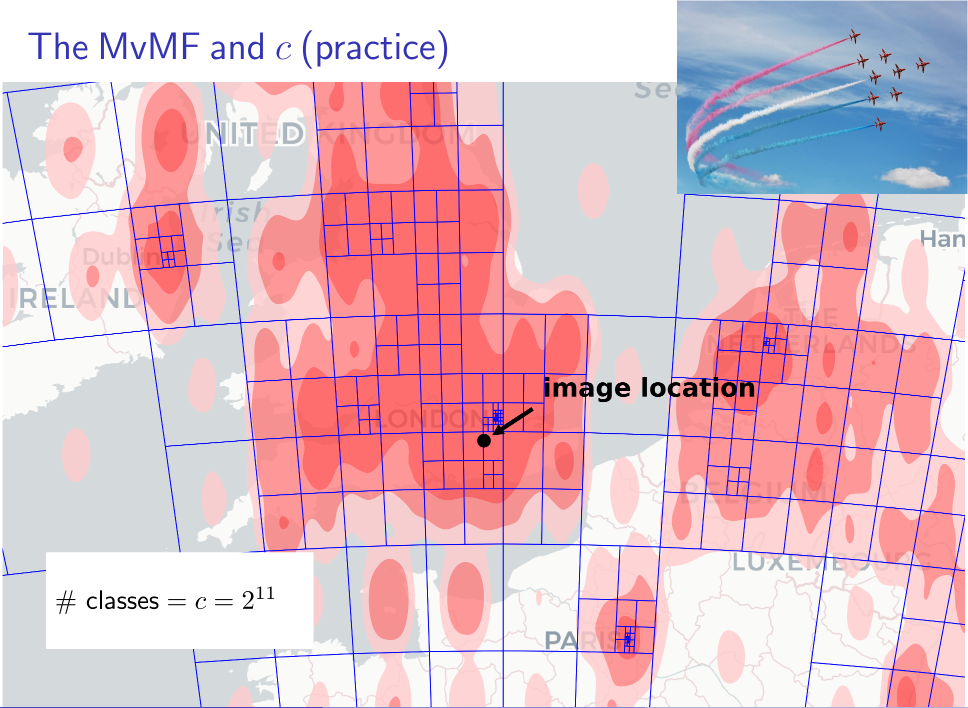


image location

# classes $= c = 2^{11}$

The MvMF and c (practice)

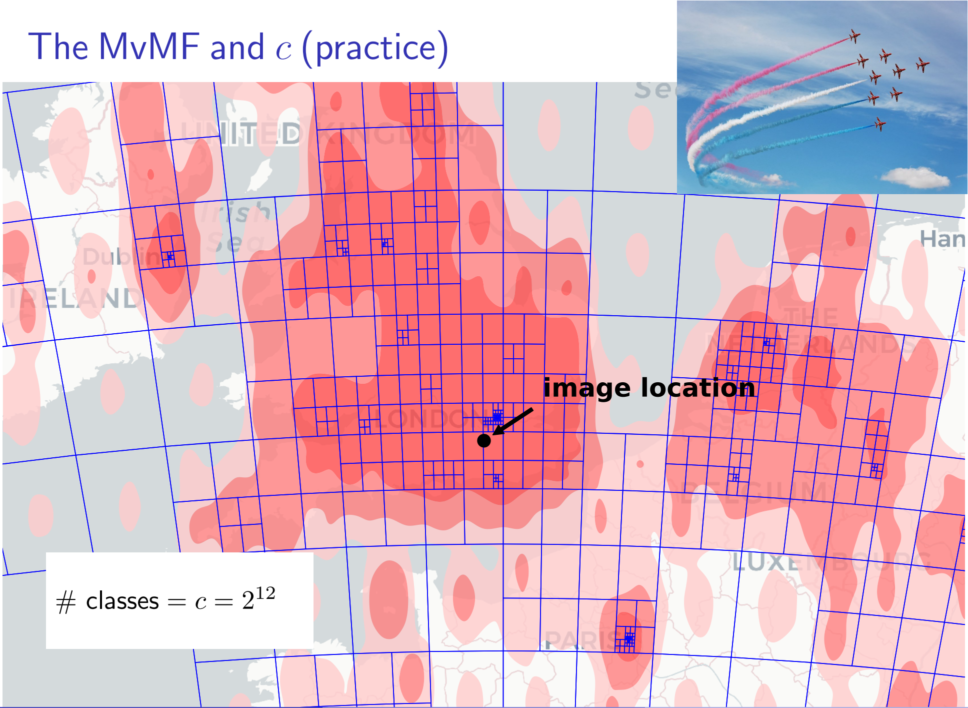
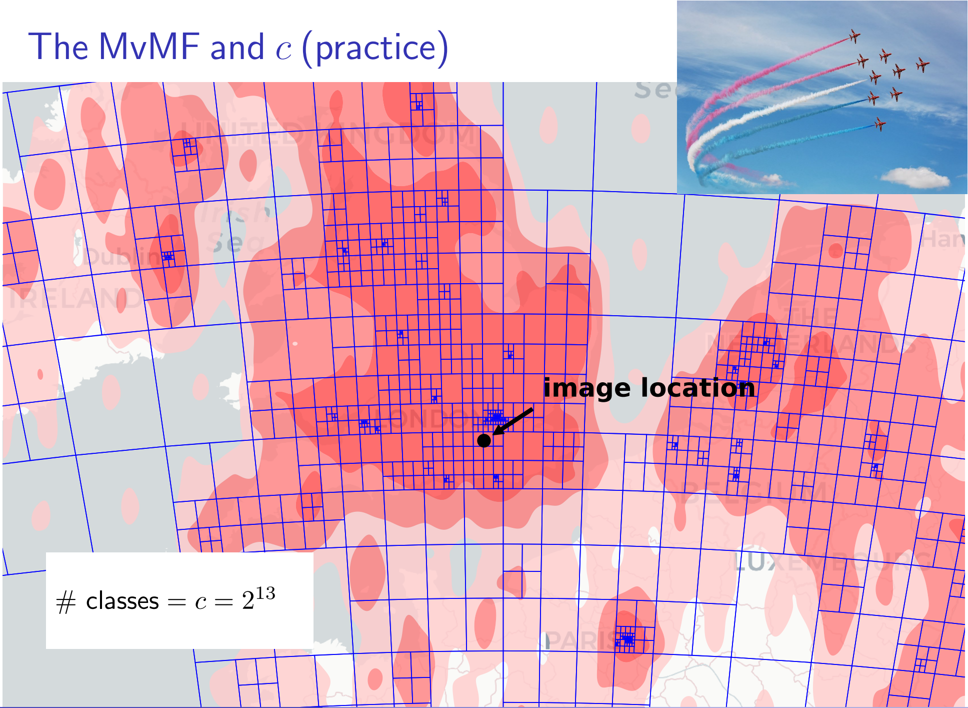


image location

# classes $= c = 2^{12}$

The MvMF and c (practice)



$\#$ classes $= c = 2^{13}$

The MvMF and c (practice)

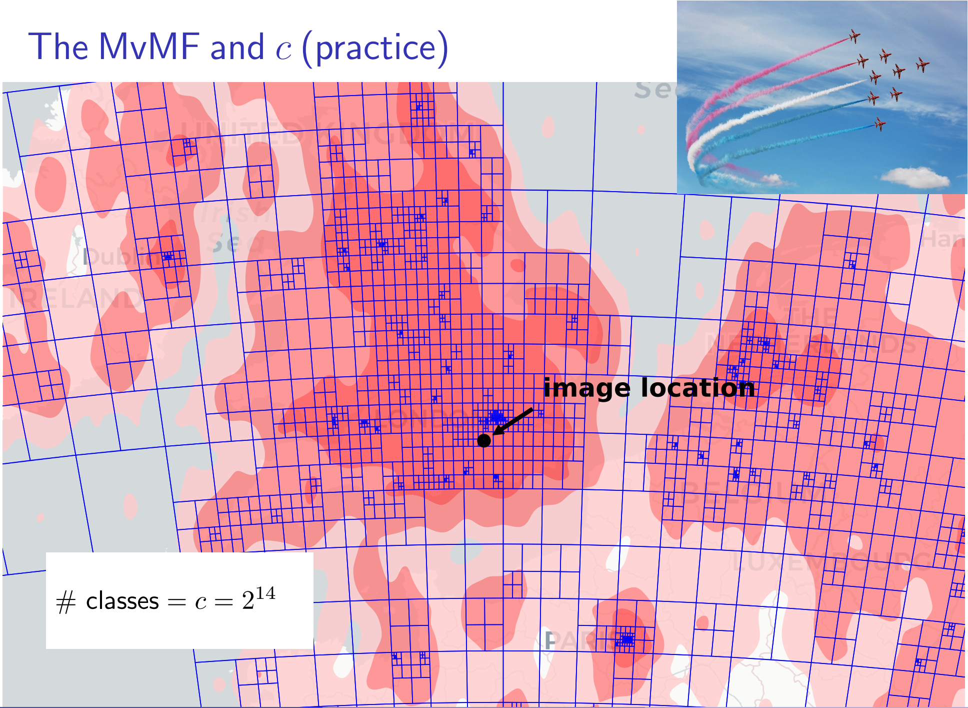


image location

$\#$ classes $= c = 2^{14}$

The MvMF and c (theory)

Theorem (Informal)

If we minimize the MvMF loss using stochastic gradient descent, then

$$\textit{generalization error} \le O\left(\sqrt{\frac{d}{n}}\right). \tag{1}$$

Notice: *no dependence on the number of mixture components c.*

The MvMF and c (theory)

Theorem (Informal)

If we minimize the MvMF loss using stochastic gradient descent, then

$$\textit{generalization error} \leq O\left(\sqrt{\frac{d}{n}}\right). \tag{1}$$

Notice: *no dependence on the number of mixture components c.*

Proof sketch. Stochastic gradient descent guarantees that for any convex loss,

$$\text{generalization error} \leq O\left(\frac{G}{\sqrt{n}}\right), \tag{2}$$

where G is the maximum gradient of the loss.

The MvMF and c (theory)

Theorem (Informal)

If we minimize the MvMF loss using stochastic gradient descent, then

$$\textit{generalization error} \leq O\left(\sqrt{\frac{d}{n}}\right). \tag{1}$$

Notice: *no dependence on the number of mixture components c.*

Proof sketch. Stochastic gradient descent guarantees that for any convex loss,

$$\text{generalization error} \leq O\left(\frac{G}{\sqrt{n}}\right), \tag{2}$$

where G is the maximum gradient of the loss.

For logistic regression, $G \leq \sqrt{dc}$.

The MvMF and c (theory)

Theorem (Informal)

If we minimize the MvMF loss using stochastic gradient descent, then

$$\textit{generalization error} \leq O\left(\sqrt{\frac{d}{n}}\right). \tag{1}$$

Notice: *no dependence on the number of mixture components c.*

Proof sketch. Stochastic gradient descent guarantees that for any convex loss,

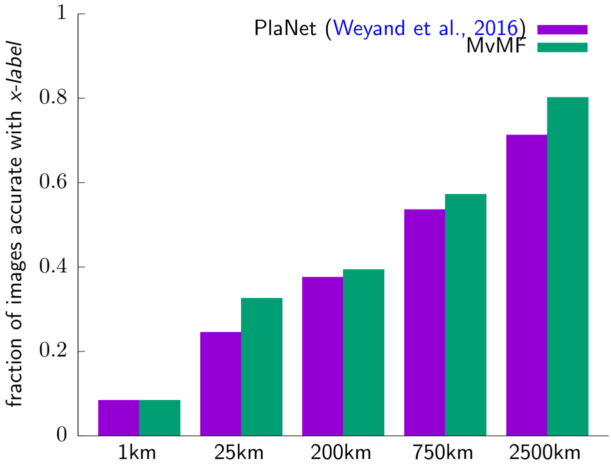
$$\text{generalization error} \leq O\left(\frac{G}{\sqrt{n}}\right), \tag{2}$$

where G is the maximum gradient of the loss.

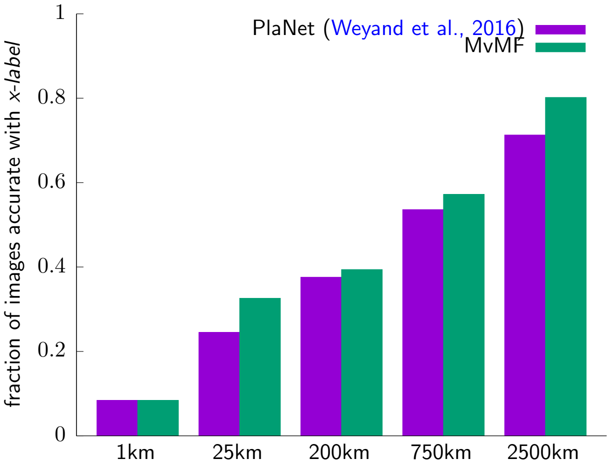
For logistic regression, $G \leq \sqrt{dc}$.

For MvMF, $G \leq \sqrt{d}$. □

Quantitative Results



Quantitative Results



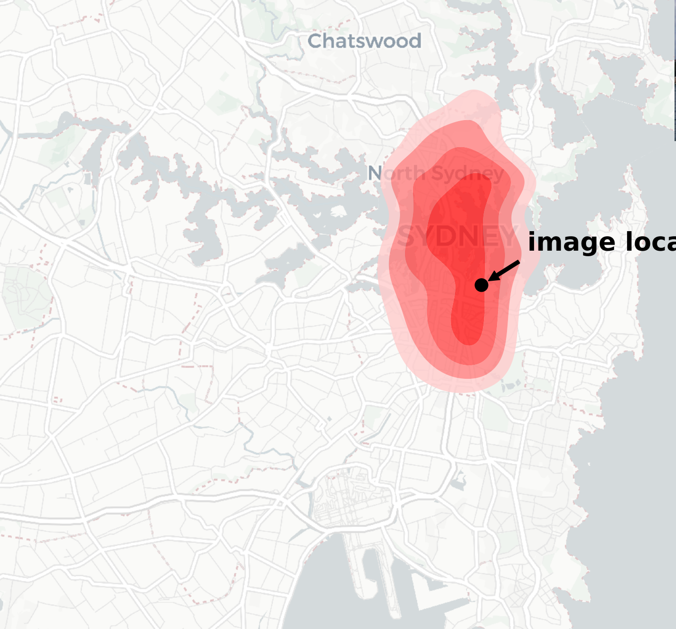
PlaNet (Weyand et al., 2016)

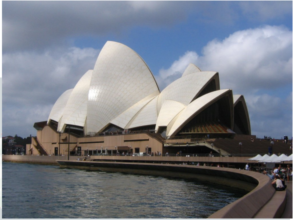
- 128 million images
- 2.5 months on 200 machine cluster

MvMF

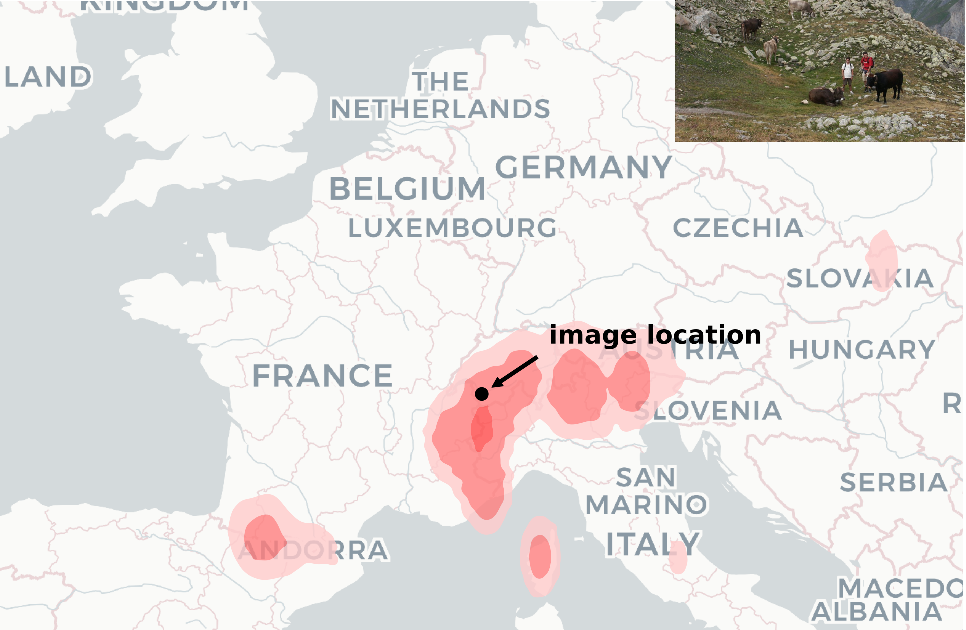
- 5 million images
- 2 months on 1 GPU

Qualitative results

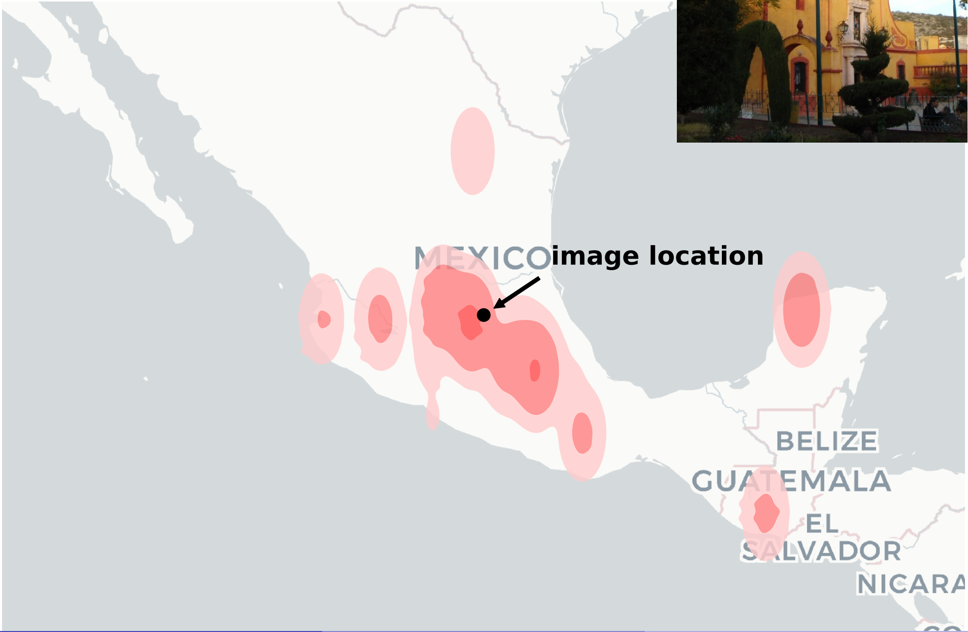




Qualitative results



Qualitative results

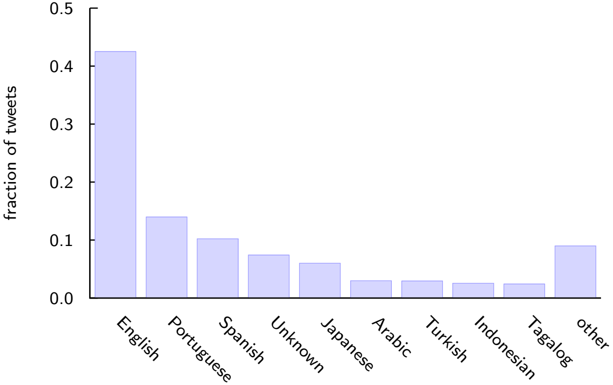


Outline

- Geolocating images with deep learning
  - Examples
  - Deep learning review
  - The PlaNet method (Weyand et al., 2016)
  - The mixture of von Mises-Fisher distribution
- Geolocating text with deep learning
  - Overview of Twitter
  - Examples
  - Word-based methods
  - UnicodeCNN, a character based method
- Future research directions

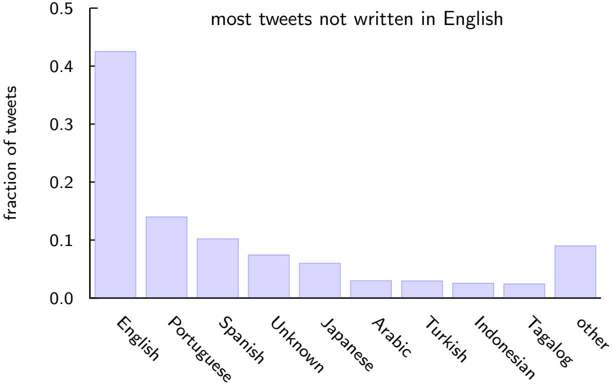
About twitter

- 500 million tweets / day
- Over 100 languages used on Twitter



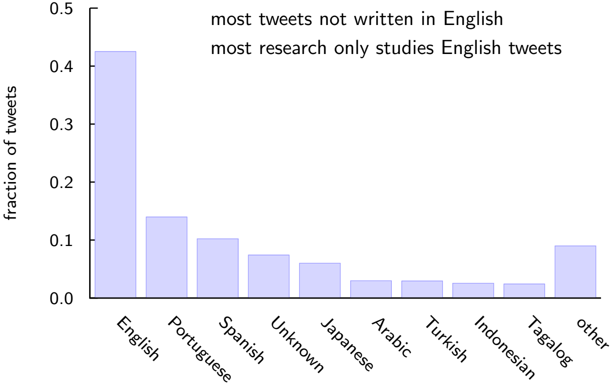
About twitter

- 500 million tweets / day
- Over 100 languages used on Twitter



About twitter

- 500 million tweets / day
- Over 100 languages used on Twitter

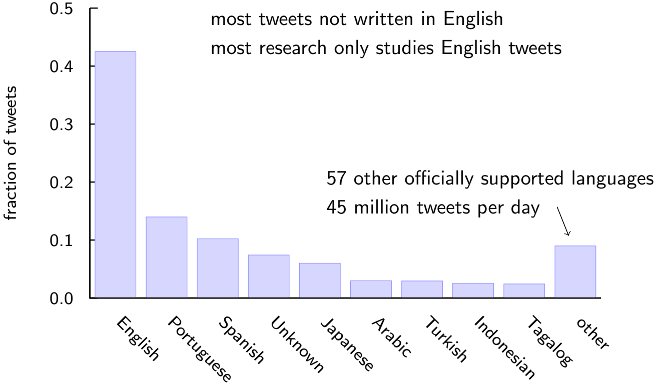


most tweets not written in English

most research only studies English tweets

About twitter

- 500 million tweets / day
- Over 100 languages used on Twitter


most tweets not written in English

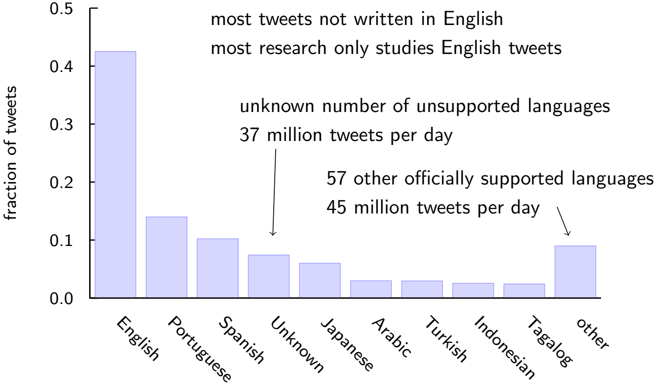
most research only studies English tweets

57 other officially supported languages

45 million tweets per day

About twitter

- 500 million tweets / day
- Over 100 languages used on Twitter



most tweets not written in English

most research only studies English tweets

unknown number of unsupported languages

37 million tweets per day

57 other officially supported languages

45 million tweets per day

The data

My dataset:

- 900 million tweets
- Every geotagged tweet sent between 26 Oct 2017 - 8 Jul 2018

The data

My dataset:

- 900 million tweets
- Every geotagged tweet sent between 26 Oct 2017 - 8 Jul 2018

Prior work:

- Cheng et al. (2010)
  Han et al. (2013)
  Compton et al. (2014)
  Maier and Gómez-Rodríguez (2014)
  Rahimi et al. (2015)
  Rahimi et al. (2017)
  Li et al. (2012)
  Mahmud et al. (2014)
  Zhang and Gelernter (2014)
  Gonçalves and Sánchez (2015)
  Dredze et al. (2016)
  Tinoco (2017)
- Filter tweets to only contain popular languages
- Filter tweets to only come from major cities/countries
- Datasets contain < 5 million tweets

The data

My dataset:

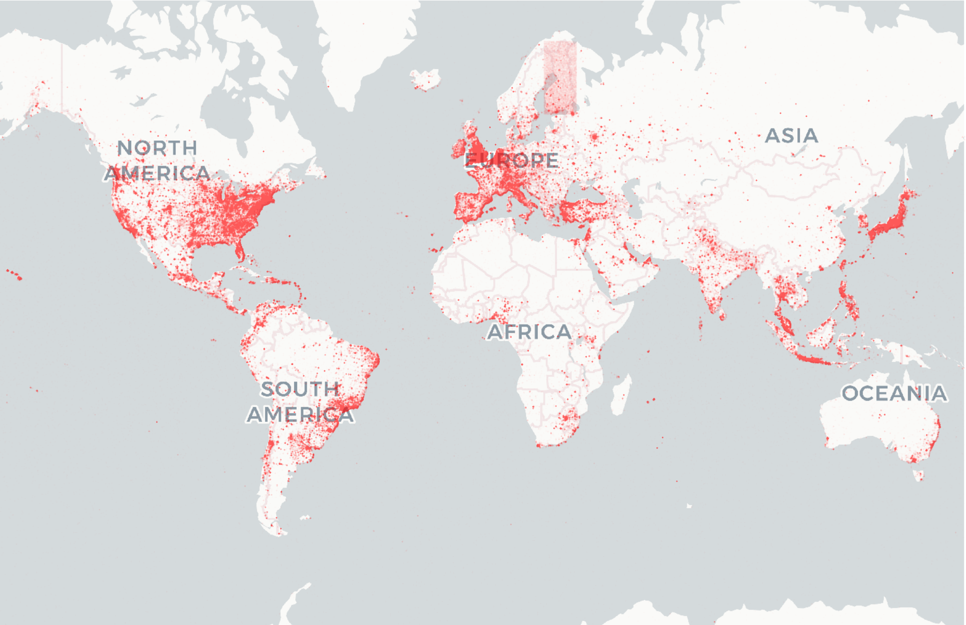
- 900 million tweets
- Every geotagged tweet sent between 26 Oct 2017 - 8 Jul 2018

Prior work:

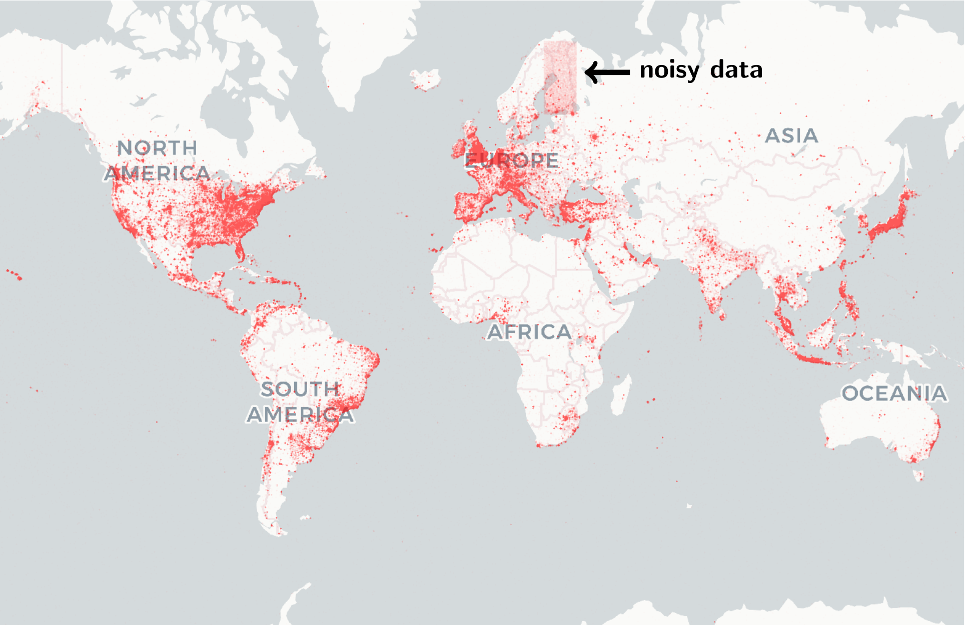
- Cheng et al. (2010) Li et al. (2012) Han et al. (2013) Mahmud et al. (2014) Compton et al. (2014) Zhang and Gelernter (2014) Maier and Gómez-Rodríguez (2014) Gonçalves and Sánchez (2015) Rahimi et al. (2015) Dredze et al. (2016) Rahimi et al. (2017) Tinoco (2017)
- Filter tweets to only contain popular languages
- Filter tweets to only come from major cities/countries
- Datasets contain < 5 million tweets

Goal: geolocate tweets in all languages sent from anywhere in the world

The data (visualized)

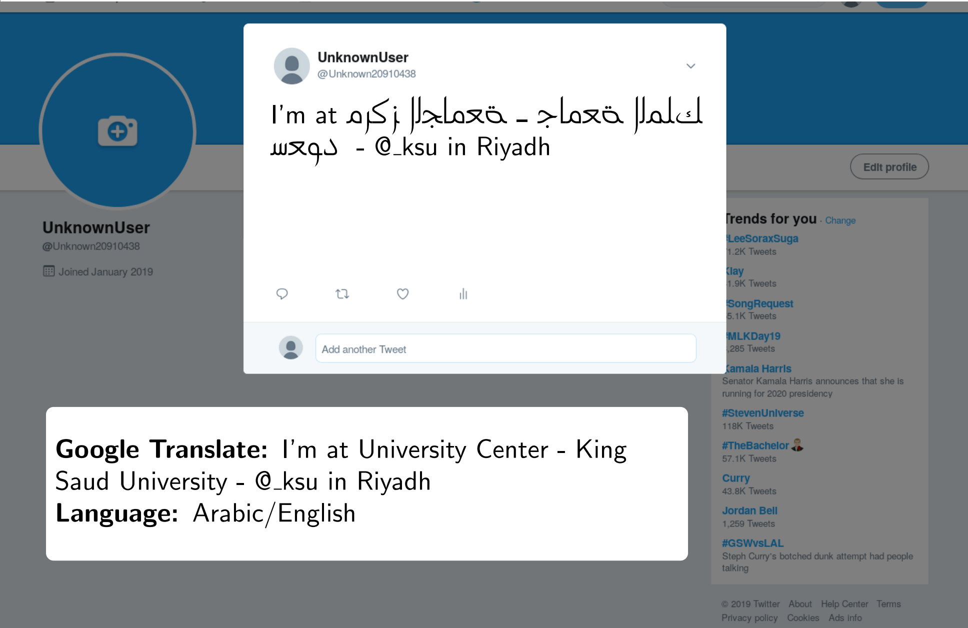


The data (visualized)



← noisy data

Few tweets are “easy”



UnknownUser
@Unknown20910438

I’m at مركز الجامعة - جامعة الملك سعود - @_ksu in Riyadh

Add another Tweet

UnknownUser
@Unknown20910438
Joined January 2019

Edit profile

Trends for you · Change

LeeSoraxSuga
1.2K Tweets

Klay
1.9K Tweets

SongRequest
5.1K Tweets

MLKDay19
,285 Tweets

Kamala Harris
Senator Kamala Harris announces that she is running for 2020 presidency

#StevenUniverse
118K Tweets

#TheBachelor
57.1K Tweets

Curry
43.8K Tweets

Jordan Bell
1,259 Tweets

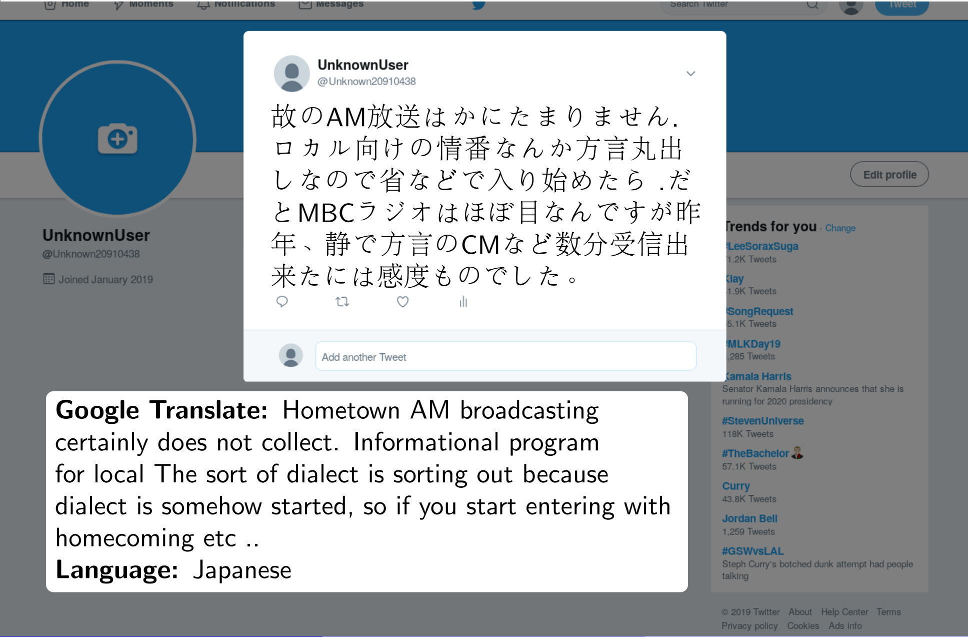
#GSWvsLAL
Steph Curry’s botched dunk attempt had people talking

© 2019 Twitter About Help Center Terms
Privacy policy Cookies Ads info

Google Translate: I’m at University Center - King Saud University - @_ksu in Riyadh

Language: Arabic/English

Most tweets are “hard”



UnknownUser
@Unknown20910438
Joined January 2019

UnknownUser
@Unknown20910438

故のAM放送はかにたまりません.ロカル向けの情番なんか方言丸出しなので省などで入り始めたら .だとMBCラジオはほぼ目なんですが昨年、静で方言のCMなど数分受信出来たには感度ものでした。

Trends for you · Change

LeeSoraxSuga
1.2K Tweets

lay
1.9K Tweets

SongRequest
5.1K Tweets

MLKDay19
,285 Tweets

amala Harris
Senator Kamala Harris announces that she is running for 2020 presidency

#StevenUniverse
118K Tweets

#TheBachelor
57.1K Tweets

Curry
43.8K Tweets

Jordan Bell
1,259 Tweets

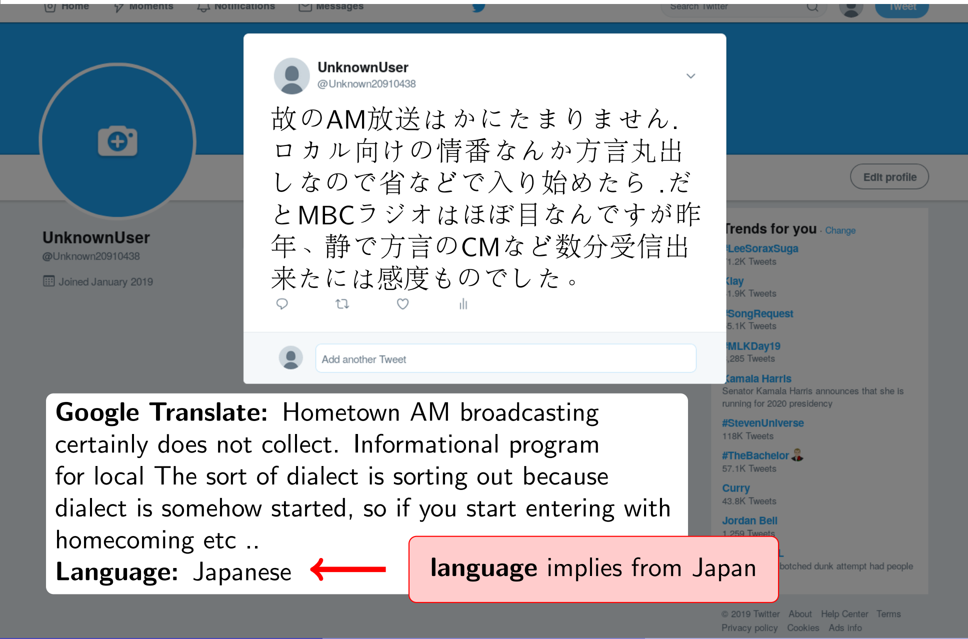
#GSWvsLAL
Steph Curry's botched dunk attempt had people talking

© 2019 Twitter About Help Center Terms
Privacy policy Cookies Ads info

Google Translate: Hometown AM broadcasting certainly does not collect. Informational program for local The sort of dialect is sorting out because dialect is somehow started, so if you start entering with homecoming etc ..

Language: Japanese

Most tweets are “hard”



UnknownUser
@Unknown20910438

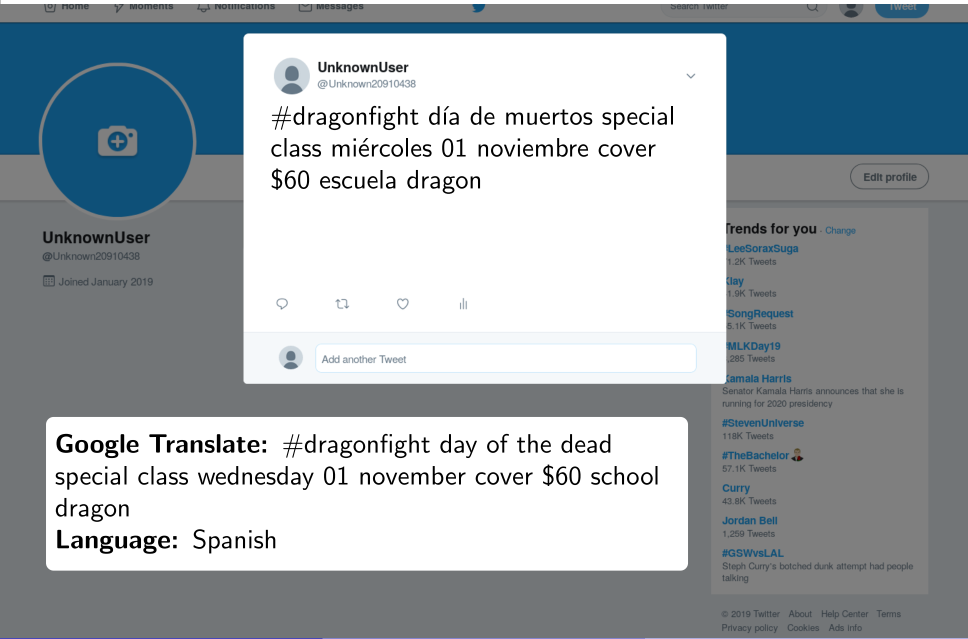
故のAM放送はかにたまりません.
ロカル向けの情番なんか方言丸出
しなので省などで入り始めたら .だ
とMBCラジオはほぼ目なんですが昨
年、静で方言のCMなど数分受信出
来たには感度ものでした。

Google Translate: Hometown AM broadcasting certainly does not collect. Informational program for local The sort of dialect is sorting out because dialect is somehow started, so if you start entering with homecoming etc ..

Language: Japanese ←

language implies from Japan

Most tweets are “hard”



UnknownUser
@Unknown20910438

#dragonfight día de muertos special class miércoles 01 noviembre cover $60 escuela dragon

Add another Tweet

UnknownUser
@Unknown20910438
Joined January 2019

Trends for you · Change

LeeSoraxSuga
1.2K Tweets

lay
1.9K Tweets

SongRequest
5.1K Tweets

MLKDay19
,285 Tweets

amala Harris
Senator Kamala Harris announces that she is running for 2020 presidency

#StevenUniverse
118K Tweets

#TheBachelor
57.1K Tweets

Curry
43.8K Tweets

Jordan Bell
1,259 Tweets

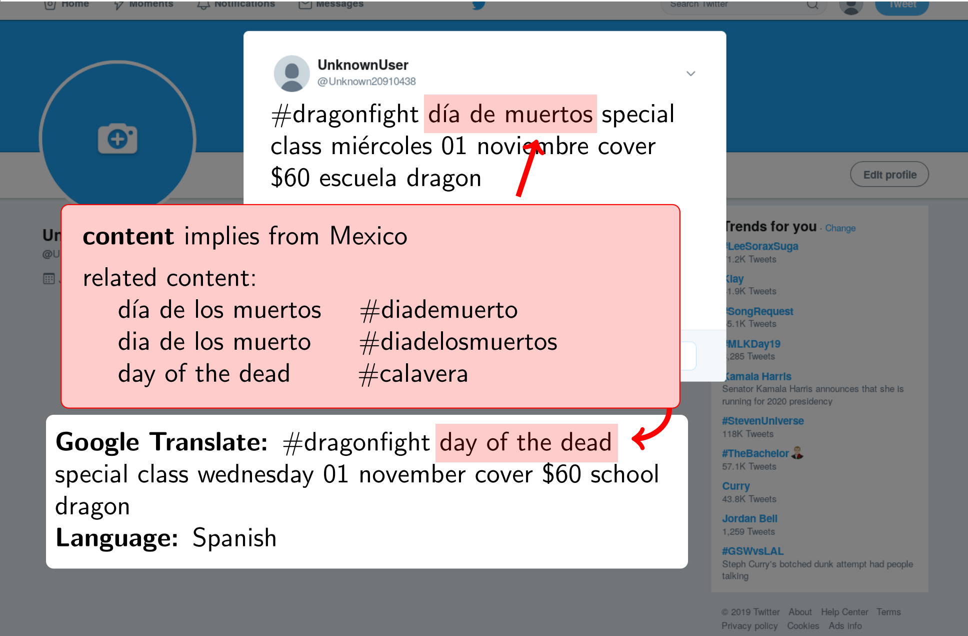
#GSWvsLAL
Steph Curry's botched dunk attempt had people talking

© 2019 Twitter About Help Center Terms
Privacy policy Cookies Ads info

Google Translate: #dragonfight day of the dead special class wednesday 01 november cover $60 school dragon

Language: Spanish

Most tweets are “hard”



UnknownUser
@Unknown20910438

#dragonfight día de muertos special class miércoles 01 noviembre cover $60 escuela dragon

content implies from Mexico

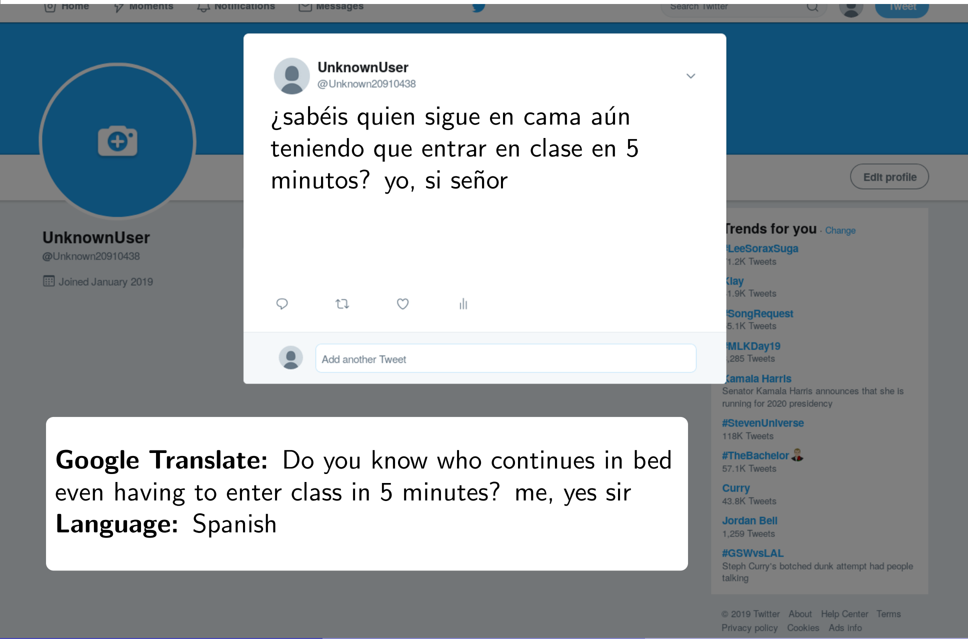
related content:

día de los muertos	#diademuerto
dia de los muerto	#diadelosmuertos
day of the dead	#calavera

Google Translate: #dragonfight day of the dead special class wednesday 01 november cover $60 school dragon

Language: Spanish

Most tweets are “hard”



UnknownUser
@Unknown20910438

¿sabéis quien sigue en cama aún teniendo que entrar en clase en 5 minutos? yo, si señor

Add another Tweet

UnknownUser
@Unknown20910438
Joined January 2019

Trends for you · Change

LeeSoraxSuga
1.2K Tweets

lay
1.9K Tweets

SongRequest
5.1K Tweets

MLKDay19
,285 Tweets

amala Harris
Senator Kamala Harris announces that she is running for 2020 presidency

#StevenUniverse
118K Tweets

#TheBachelor
57.1K Tweets

Curry
43.8K Tweets

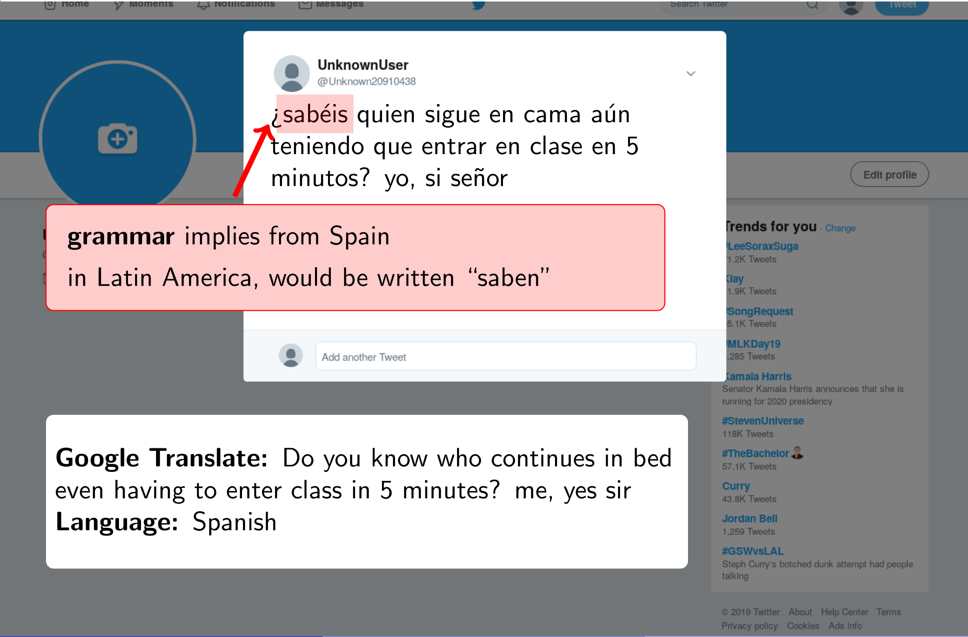
Jordan Bell
1,259 Tweets

#GSWvsLAL
Steph Curry's botched dunk attempt had people talking

© 2019 Twitter About Help Center Terms
Privacy policy Cookies Ads info

Google Translate: Do you know who continues in bed even having to enter class in 5 minutes? me, yes sir
Language: Spanish

Most tweets are “hard”



UnknownUser
@Unknown20910438

¿sabéis quien sigue en cama aún teniendo que entrar en clase en 5 minutos? yo, si señor

Add another Tweet

Trends for you · Change

LeeSoraxSuga
1.2K Tweets

Klay
1.9K Tweets

SongRequest
5.1K Tweets

MLKDay19
,285 Tweets

Kamala Harris
Senator Kamala Harris announces that she is running for 2020 presidency

#StevenUniverse
118K Tweets

#TheBachelor
57.1K Tweets

Curry
43.8K Tweets

Jordan Bell
1,259 Tweets

#GSWvsLAL
Steph Curry's botched dunk attempt had people talking

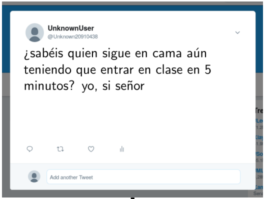
© 2019 Twitter About Help Center Terms
Privacy policy Cookies Ads info

grammar implies from Spain

in Latin America, would be written “saben”

Google Translate: Do you know who continues in bed even having to enter class in 5 minutes? me, yes sir
Language: Spanish

Machine learning for tweet geolocation



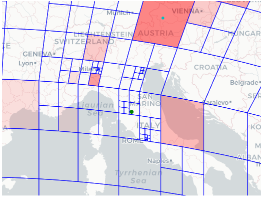
UnknownUser
@Unknown20910438

¿sabéis quien sigue en cama aún teniendo que entrar en clase en 5 minutos? yo, si señor

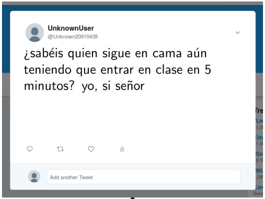
Add another Tweet

Feature Generation

Logistic Regression



Machine learning for tweet geolocation

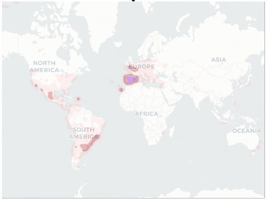


UnknownUser
@Unknown20910438

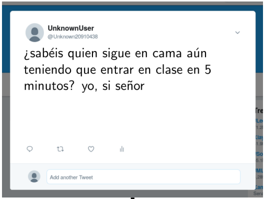
¿sabéis quien sigue en cama aún teniendo que entrar en clase en 5 minutos? yo, si señor

Feature Generation

MvMF



Machine learning for tweet geolocation



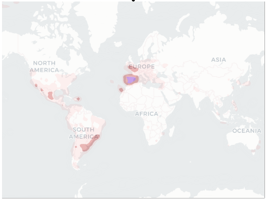
UnknownUser
@Unknown20910438

¿sabéis quien sigue en cama aún teniendo que entrar en clase en 5 minutos? yo, si señor

Add another Tweet

Feature Generation

MvMF



Previous methods: generate features from **words**

UnicodeCNN: generates features from **characters**

The features-from-words approach

Input text: ¿Sabéis quien sigue en cama?

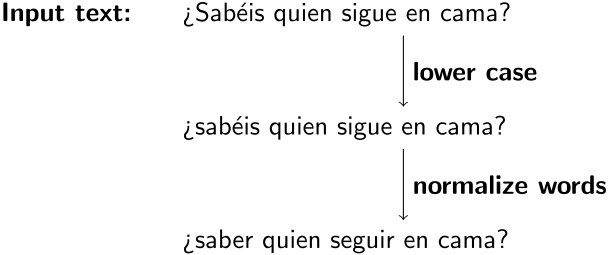
The features-from-words approach

Input text: ¿Sabéis quien sigue en cama?

↓ **lower case**

¿sabéis quien sigue en cama?

The features-from-words approach



Input text: ¿Sabéis quien sigue en cama?

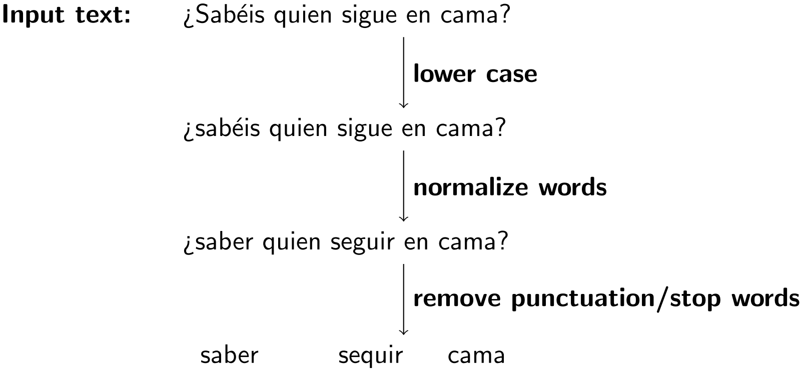
↓ **lower case**

¿sabéis quien sigue en cama?

↓ **normalize words**

¿saber quien seguir en cama?

The features-from-words approach



Input text: ¿Sabéis quien sigue en cama?

lower case

¿sabéis quien sigue en cama?

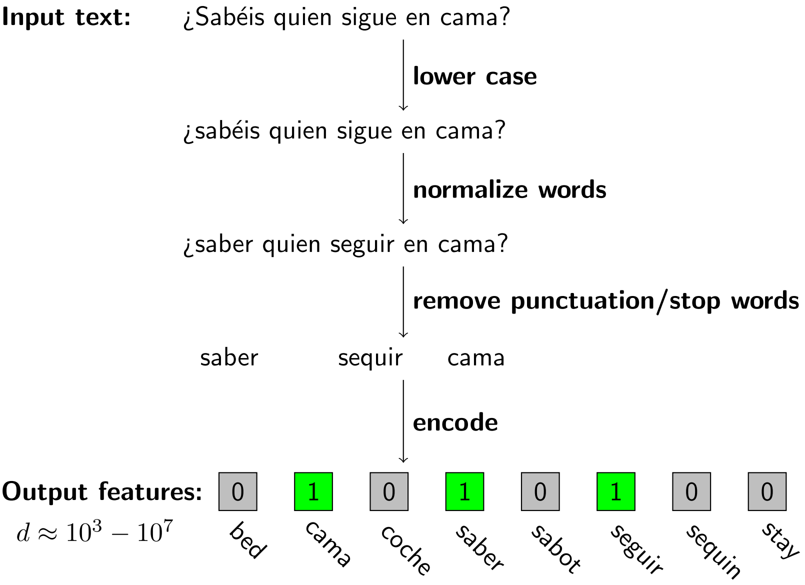
normalize words

¿saber quien seguir en cama?

remove punctuation/stop words

saber sequir cama

The features-from-words approach



Input text: ¿Sabéis quien sigue en cama?

lower case

¿sabéis quien sigue en cama?

normalize words

¿saber quien seguir en cama?

remove punctuation/stop words

saber sequir cama

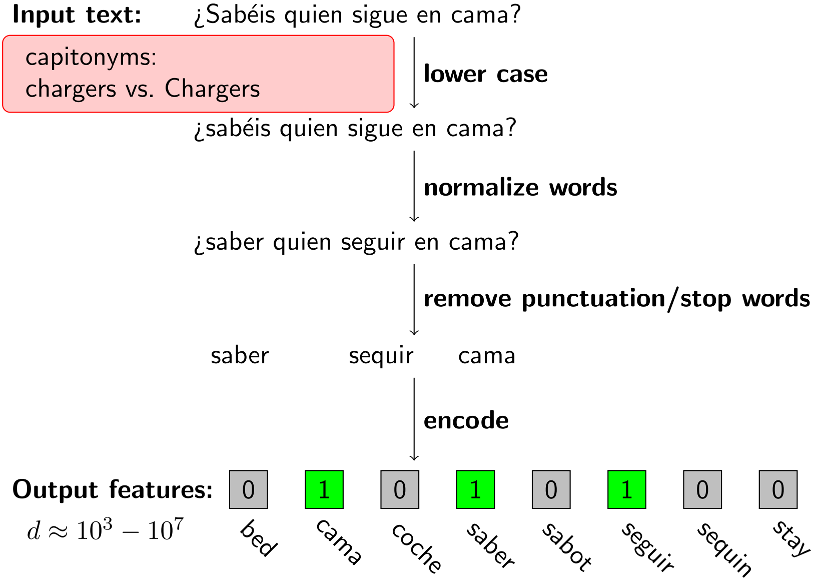
encode

Output features:

$d \approx 10^3 - 10^7$

bed	cama	coche	saber	sabot	seguir	sequin	stay
0	1	0	1	0	1	0	0

The features-from-words approach



Input text: ¿Sabéis quien sigue en cama?

capitonyms:
chargers vs. Chargers

lower case

¿sabéis quien sigue en cama?

normalize words

¿saber quien seguir en cama?

remove punctuation/stop words

saber sequir cama

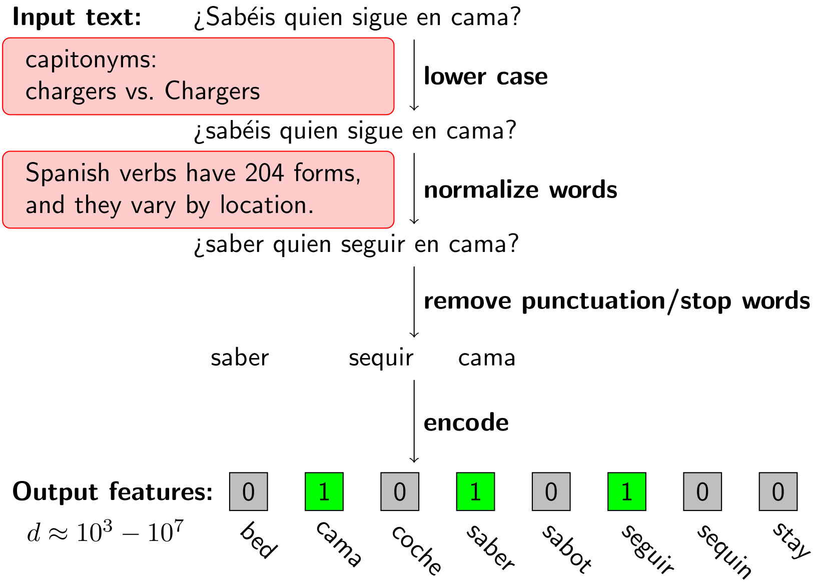
encode

Output features:

$d \approx 10^3 - 10^7$

bed	cama	coche	saber	sabot	seguir	sequin	stay
0	1	0	1	0	1	0	0

The features-from-words approach



Input text: ¿Sabéis quien sigue en cama?

capitonyms:
chargers vs. Chargers

lower case

¿sabéis quien sigue en cama?

Spanish verbs have 204 forms, and they vary by location.

normalize words

¿saber quien seguir en cama?

remove punctuation/stop words

saber sequir cama

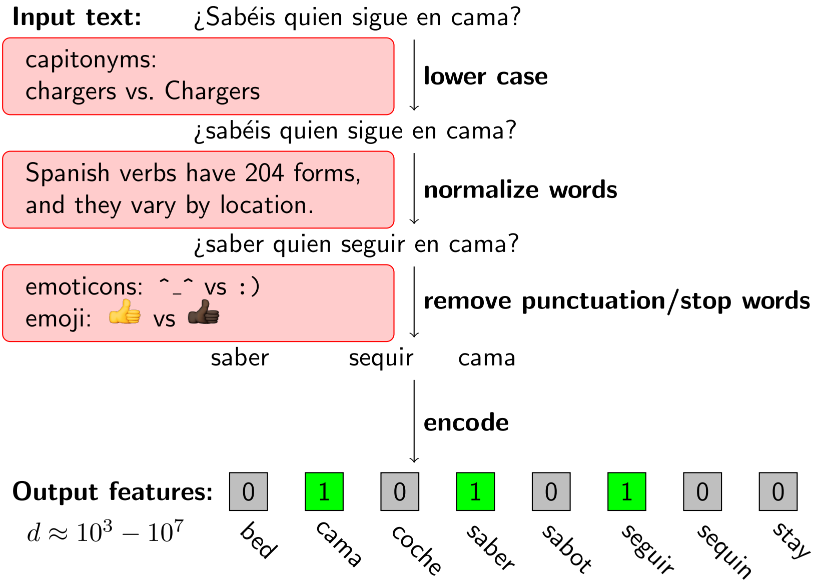
encode

Output features:

$d \approx 10^3 - 10^7$

bed	cama	coche	saber	sabot	seguir	sequin	stay
0	1	0	1	0	1	0	0

The features-from-words approach



Input text: ¿Sabéis quien sigue en cama?

capitonyms:
chargers vs. Chargers

↓ **lower case**

¿sabéis quien sigue en cama?

Spanish verbs have 204 forms, and they vary by location.

↓ **normalize words**

¿saber quien seguir en cama?

emoticons: ^_^ vs :)
emoji: 👍 vs 👍🏿

↓ **remove punctuation/stop words**

saber sequir cama

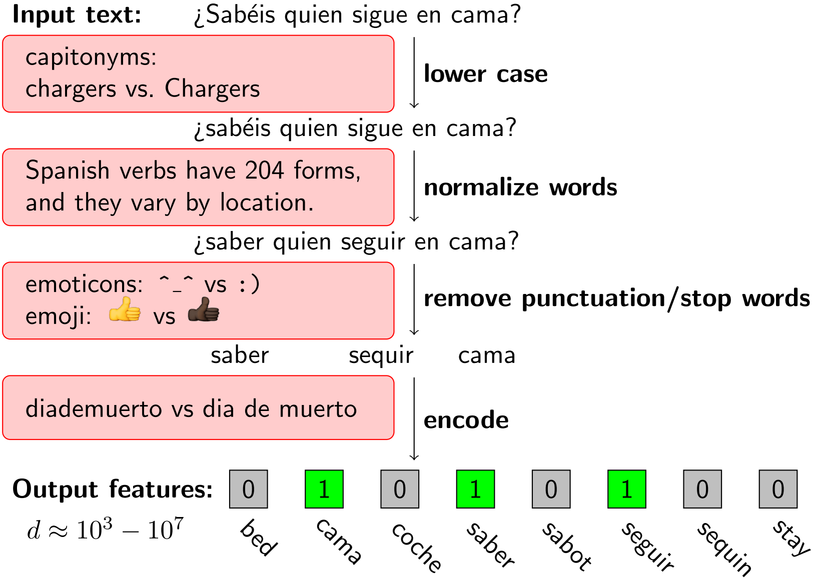
↓ **encode**

Output features:

$d \approx 10^3 - 10^7$

bed	cama	coche	saber	sabot	seguir	sequin	stay
0	1	0	1	0	1	0	0

The features-from-words approach



Input text: ¿Sabéis quien sigue en cama?

capitonyms:
chargers vs. Chargers

lower case

¿sabéis quien sigue en cama?

Spanish verbs have 204 forms, and they vary by location.

normalize words

¿saber quien seguir en cama?

emoticons: ^_^ vs :)
emoji: 👍 vs 👍🏿

remove punctuation/stop words

saber sequir cama

diademuerto vs dia de muerto

encode

Output features:

$d \approx 10^3 - 10^7$

bed	cama	coche	saber	sabot	seguir	sequin	stay
0	1	0	1	0	1	0	0

The features-from-words approach

Input text: ¿Sabéis quien sigue en cama?

capitonyms:
chargers vs. Chargers

↓ **lower case**

¿sabéis quien sigue en cama?

Spanish verbs have 204 forms, and they vary by location.

↓ **normalize words**

¿saber quien seguir en cama?

emoticons: ^_^ vs :)
emoji: 👍 vs 👍🏿

↓ **remove punctuation/stop words**

saber sequir cama

diademuerto vs dia de muerto

↓ **encode**

each language is different!

Output features:

$d \approx 10^3 - 10^7$

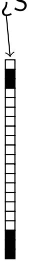
bed	cama	coche	saber	sabot	seguir	sequin	stay
0	1	0	1	0	1	0	0

The UnicodeCNN (features from characters)

Input text: ¿Sabéis quien sigue en cama?

The UnicodeCNN (features from characters)

Input text: ¿Sabéis quien sigue en cama?

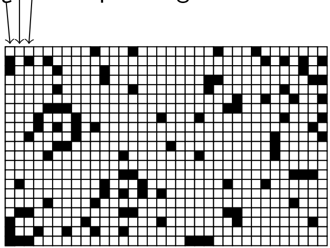


Step 1: Encode each character as a vector of bits
Each bit represents a property from the Unicode standard

- is it punctuation?
- is it whitespace?
- does it sound like 'a'?

The UnicodeCNN (features from characters)

Input text: ¿Sabéis quien sigue en cama?



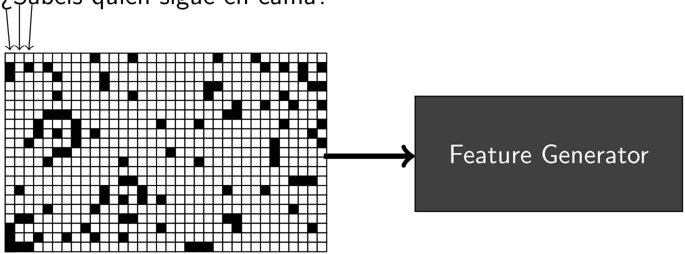
Step 1: Encode each character as a vector of bits
Each bit represents a property from the Unicode standard

- is it punctuation?
- is it whitespace?
- does it sound like 'a'?

Step 2: Concatenate the vectors to form an "image"

The UnicodeCNN (features from characters)

Input text: ¿Sabéis quien sigue en cama?



Step 1: Encode each character as a vector of bits
Each bit represents a property from the Unicode standard

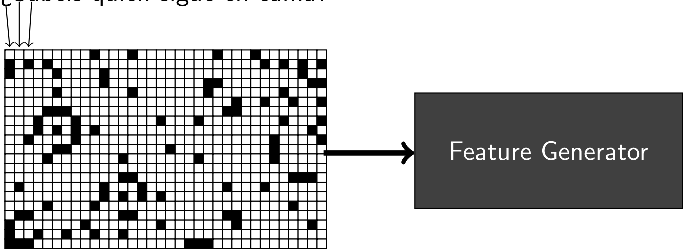
- is it punctuation?
- is it whitespace?
- does it sound like 'a'?

Step 2: Concatenate the vectors to form an "image"

Step 3: Apply image feature generation methods (i.e. CNN)

The UnicodeCNN (features from characters)

Input text: ¿Sabéis quien sigue en cama?



Step 1: Encode each character as a vector of bits
Each bit represents a property from the Unicode standard

- is it punctuation?
- is it whitespace?
- does it sound like ‘a’?

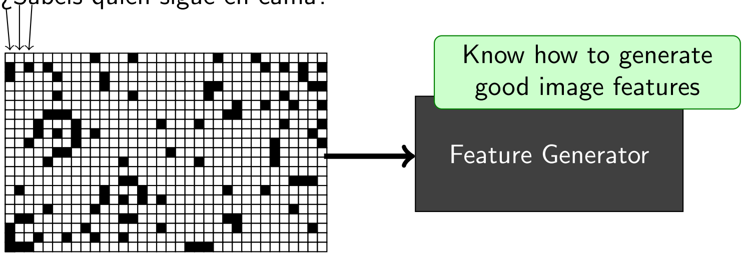
Works for all languages

Step 2: Concatenate the vectors to form an “image”

Step 3: Apply image feature generation methods (i.e. CNN)

The UnicodeCNN (features from characters)

Input text: ¿Sabéis quien sigue en cama?



Know how to generate good image features

Step 1: Encode each character as a vector of bits
Each bit represents a property from the Unicode standard

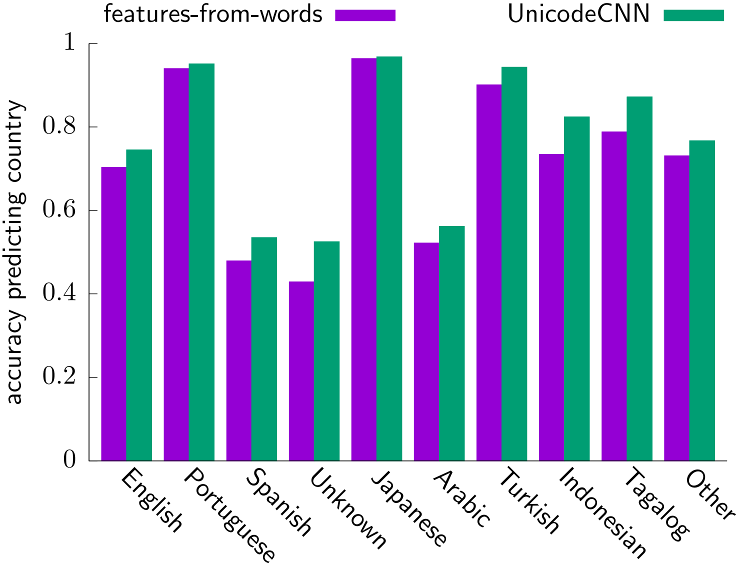
- is it punctuation?
- is it whitespace?
- does it sound like ‘a’?

Works for all languages

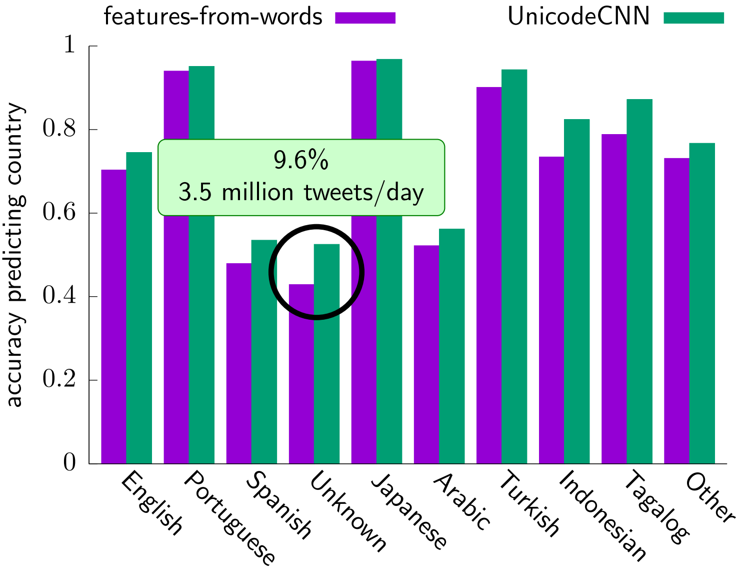
Step 2: Concatenate the vectors to form an “image”

Step 3: Apply image feature generation methods (i.e. CNN)

Quantitative Results



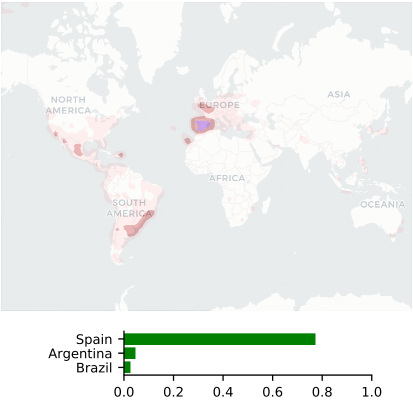
Quantitative Results



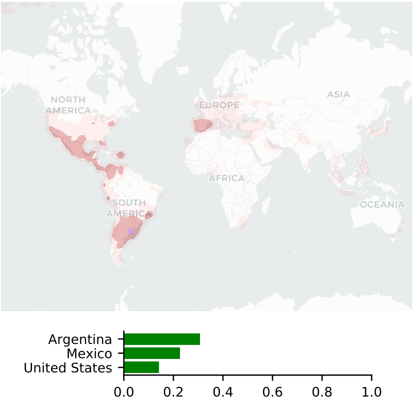
9.6%
3.5 million tweets/day

Qualitative Results: verb conjugations

no me apres**éis**



no me apres**en**



Neither the word `apreséis` nor `apresen` appears in the training data

The UnicodeCNN learned a general rule about Spanish dialects

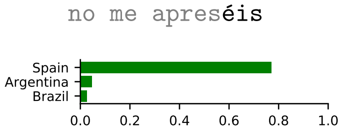
Outline

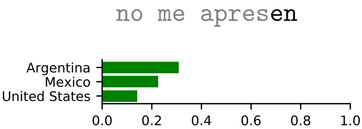
- Geolocating images with deep learning
  - Examples
  - Deep learning review
  - The PlaNet method (Weyand et al., 2016)
  - The mixture of von Mises-Fisher distribution
- Geolocating text with deep learning
  - Overview of Twitter
  - Examples
  - Word-based methods
  - UnicodeCNN, a character based method
- Future research directions

Future Work (1/3): analyzing social media data

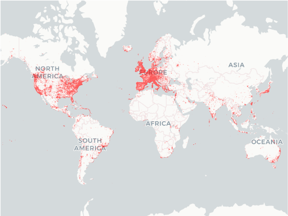
Example student projects

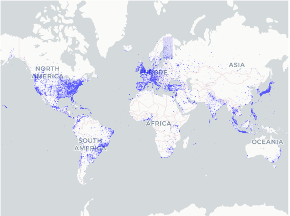
- Automatically discover, characterize, and map language dialects





- Correct for unbalanced data caused by unequal internet usage

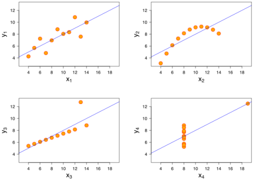




Future Work (2/3): machine learning theory ∩ practice

- statistical and computational properties of learning algorithms

robust to outliers



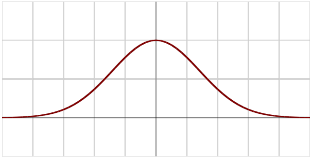
(ongoing)

distributed

(ICML2013,ongoing)

- learning with generalized metric spaces (ICML2015)

d(x,y) =

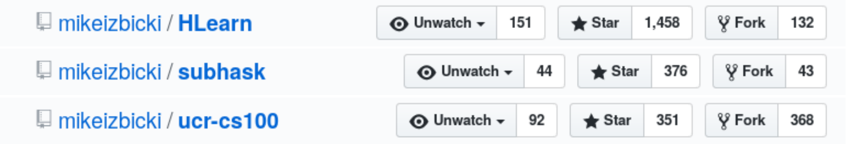


- programming language support for machine learning
  (TFP 2013, MLOSS 2013, Monad Reader 2013, DCP 2014)

Future Work (3/3): open source education

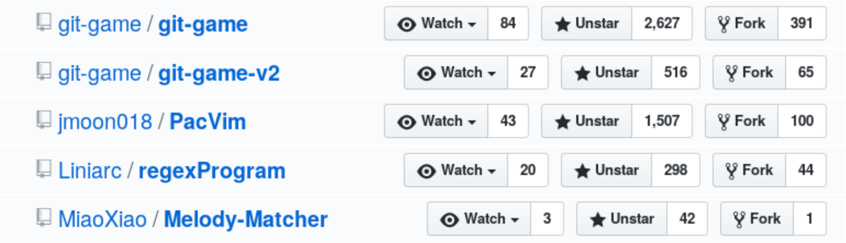
GitHub

Some of my projects:



mikeizbicki / **HLearn**	Unwatch	151	★ Star	1,458	Fork	132
mikeizbicki / **subhask**	Unwatch	44	★ Star	376	Fork	43
mikeizbicki / **ucr-cs100**	Unwatch	92	★ Star	351	Fork	368

Some student projects:

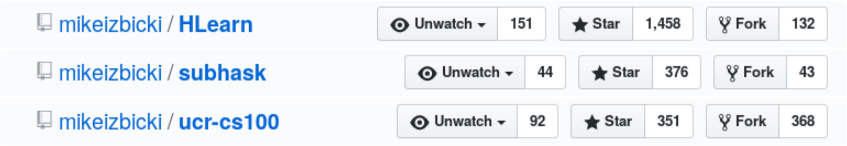


git-game / **git-game**	Watch	84	★ Unstar	2,627	Fork	391
git-game / **git-game-v2**	Watch	27	★ Unstar	516	Fork	65
jmoon018 / **PacVim**	Watch	43	★ Unstar	1,507	Fork	100
Liniarc / **regexProgram**	Watch	20	★ Unstar	298	Fork	44
MiaoXiao / **Melody-Matcher**	Watch	3	★ Unstar	42	Fork	1

ICOPUST2015, graduate student research award SIGCSE2015

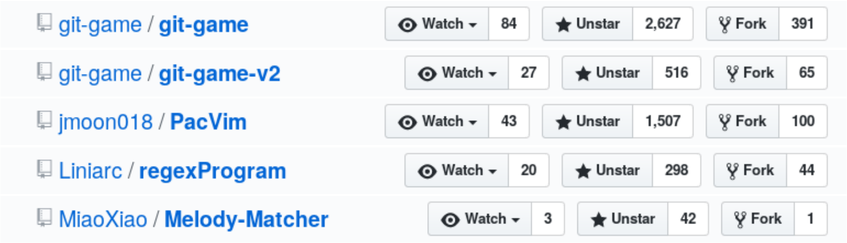
Future Work (3/3): open source education GitHub

Some of my projects:



Repository	Watch		Star		Fork	
mikeizbicki / **HLearn**	Unwatch ▾	151	★ Star	1,458	Fork	132
mikeizbicki / **subhask**	Unwatch ▾	44	★ Star	376	Fork	43
mikeizbicki / **ucr-cs100**	Unwatch ▾	92	★ Star	351	Fork	368

Some student projects:



Repository	Watch		Star		Fork	
git-game / **git-game**	Watch ▾	84	★ Unstar	2,627	Fork	391
git-game / **git-game-v2**	Watch ▾	27	★ Unstar	516	Fork	65
jmoon018 / **PacVim**	Watch ▾	43	★ Unstar	1,507	Fork	100
Liniarc / **regexProgram**	Watch ▾	20	★ Unstar	298	Fork	44
MiaoXiao / **Melody-Matcher**	Watch ▾	3	★ Unstar	42	Fork	1

ICOPUST2015, graduate student research award SIGCSE2015

First open source contributions from North Koreans (Congressional Award)

이즈비키 마이클
@ Mikelzbicki

Questions?

¿Preguntas?

Questões?

Вопросы?

الأسئلة؟

질문이 있으십니까?

質問は?

오후 2시 48분 - 2018 5월 6일

27182리트 윗 8.9k의 '좋아요'

314159 27182 161803 (42)

Bibliography I

Dolan Antenucci, Michael Cafarella, Margaret Levenstein, Christopher Ré, and Matthew D Shapiro. 2014. *Using social media to measure labor market flows*. Technical Report.

Zhiyuan Cheng, James Caverlee, and Kyumin Lee. 2010. You are where you tweet: a content-based approach to geo-locating twitter users. In *CIKM*.

Ryan Compton, David Jurgens, and David Allen. 2014. Geotagging one hundred million twitter accounts with total variation minimization. In *Big Data*.

Mark Dredze, Miles Osborne, and Prabhanjan Kambadur. 2016. Geolocation for Twitter: Timing Matters.

Bruno Gonçalves and David Sánchez. 2015. Learning about Spanish dialects through Twitter. *arXiv preprint arXiv:1511.04970* (2015).

Bo Han, Paul Cook, and Timothy Baldwin. 2013. A Stacking-based Approach to Twitter User Geolocation Prediction.

James Hays and Alexei A Efros. 2008. IM2GPS: estimating geographic information from a single image. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, 1–8.

Bibliography II

James Hays and Alexei A Efros. 2015. Large-scale image geolocalization. In *Multimodal Location Estimation of Videos and Images*. Springer, 41–62.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Identity mappings in deep residual networks. In *European conference on computer vision*. Springer, 630–645.

Andrew G Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861* (2017).

Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. 2017. Densely connected convolutional networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2261–2269.

Forrest N Iandola, Song Han, Matthew W Moskewicz, Khalid Ashraf, William J Dally, and Kurt Keutzer. 2016. Squeezenet: Alexnet-level accuracy with 50x fewer parameters and¡ 0.5 mb model size. *arXiv preprint arXiv:1602.07360* (2016).

Bibliography III

Sergey Ioffe and Christian Szegedy. 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167* (2015).

Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. 2012. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*. 1097–1105.

Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. 1998. Gradient-based learning applied to document recognition. *Proc. IEEE* 86, 11 (1998), 2278–2324.

Rui Li, Shengjie Wang, Hongbo Deng, Rui Wang, and Kevin Chen-Chuan Chang. 2012. Towards social user profiling: unified and discriminative influence model for inferring home locations. In *SIGKDD*.

Tsung-Yi Lin, Priyal Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. 2018. Focal loss for dense object detection. *IEEE transactions on pattern analysis and machine intelligence* (2018).

Bibliography IV

Chenxi Liu, Barret Zoph, Maxim Neumann, Jonathon Shlens, Wei Hua, Li-Jia Li, Li Fei-Fei, Alan Yuille, Jonathan Huang, and Kevin Murphy. 2018. Progressive neural architecture search. In *Proceedings of the European Conference on Computer Vision (ECCV)*. 19–34.

Jalal Mahmud, Jeffrey Nichols, and Clemens Drews. 2014. Home location identification of twitter users. *TIST* (2014).

Wolfgang Maier and Carlos Gómez-Rodríguez. 2014. Language variety identification in Spanish tweets. In *EMNLP*.

Hatem Mousselly-Sergieh, Daniel Watzinger, Bastian Huber, Mario Döller, Elöd Egyed-Zsigmond, and Harald Kosch. 2014. World-wide Scale Geotagged Image Dataset for Automatic Image Annotation and Reverse Geotagging. In *Proceedings of the 5th ACM Multimedia Systems Conference (MMSys '14)*. ACM, New York, NY, USA, 47–52. DOI: http://dx.doi.org/10.1145/2557642.2563673

Michael J Paul, Mark Dredze, and David Broniatowski. 2014. Twitter improves influenza forecasting. *PLoS Currents* (2014).

Robert Power and Justin Kibell. 2017. The social media intelligence analyst for emergency management. (2017).

Bibliography V

Afshin Rahimi, Trevor Cohn, and Timothy Baldwin. 2015. Twitter user geolocation using a unified text and network prediction model. *arXiv preprint arXiv:1506.08259* (2015).

Afshin Rahimi, Trevor Cohn, and Timothy Baldwin. 2017. A neural model for user geolocation and lexical dialectology. *arXiv preprint arXiv:1704.04008* (2017).

Karen Simonyan and Andrew Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014).

Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. 2015. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 1–9.

Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. 2016. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2818–2826.

Antonio Ruiz Tinoco. 2017. Variation of the second person singular of the simple past tense in twitter: hiciste vs. hicistes" you did". *Dialectologia: Revista Electrònica* (2017).

Bibliography VI

Marie Truelove, Kourosh Khoshelham, Simon McLean, Stephan Winter, and Maria Vasardani. 2017. Identifying Witness Accounts from Social Media Using Imagery. *ISPRS International Journal of Geo-Information* 6, 4 (2017), 120.

Nam Vo, Nathan Jacobs, and James Hays. 2017. Revisiting IM2GPS in the deep learning era. In *Computer Vision (ICCV), 2017 IEEE International Conference on*. IEEE, 2640–2649.

Elizabeth A Wentz, Sharolyn Anderson, Michail Fragkias, Maik Netzband, Victor Mesev, Soe W Myint, Dale Quattrochi, Atiqur Rahman, and Karen C Seto. 2014. Supporting global environmental change research: A review of trends and knowledge gaps in urban remote sensing. *Remote Sensing* 6, 5 (2014), 3879–3905.

Tobias Weyand, Ilya Kostrikov, and James Philbin. 2016. Planet-photo geolocation with convolutional neural networks. In *European Conference on Computer Vision*. Springer, 37–55.

Ian H Witten, Eibe Frank, Mark A Hall, and Christopher J Pal. 2016. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.

Sergey Zagoruyko and Nikos Komodakis. 2016. Wide residual networks. *arXiv preprint arXiv:1605.07146* (2016).

Bibliography VII

Wei Zhang and Judith Gelernter. 2014. Geocoding location expressions in Twitter messages: A preference learning method. *Journal of Spatial Information Science* (2014).

Barret Zoph, Vijay Vasudevan, Jonathon Shlens, and Quoc V Le. 2018. Learning transferable architectures for scalable image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 8697–8710.