February 24, 2016

Privacy Implications of Machine Learning

Sebastian Zimmeck

COMS W3410 — Computers and Society

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Motivation

Social Networks

Gov. Agencies

Ad. Networks

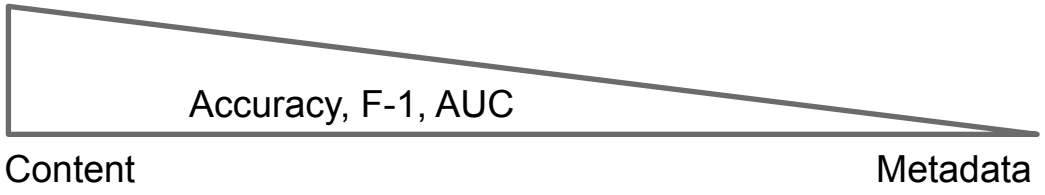
Learning of new Information

- **Direct learning**: from observing user attributes.

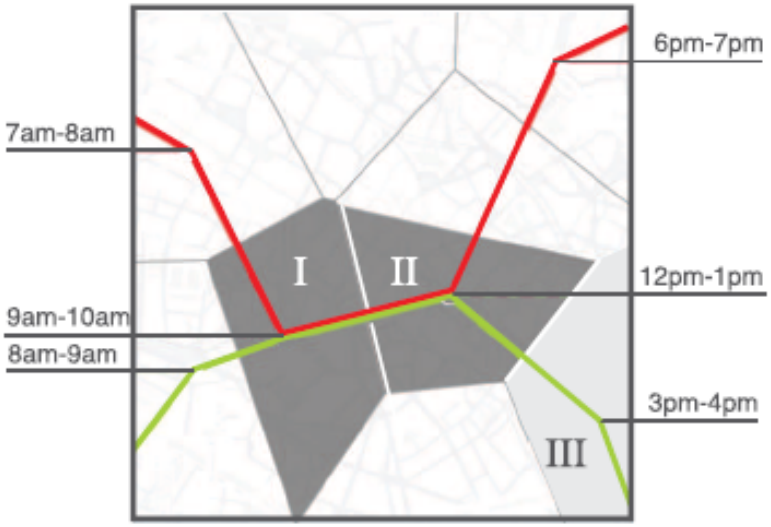
Learned Attribute	Algorithm	Features	AUC	Publication
African American or Caucasian American	Logistic Regression	Facebook Likes	0.95	Kosinski et al., PNAS ’13.
US Citizen or not	Naïve Bayes	Call patterns, locations, …	0.73	Altshuler et al., PASSAT ’12.

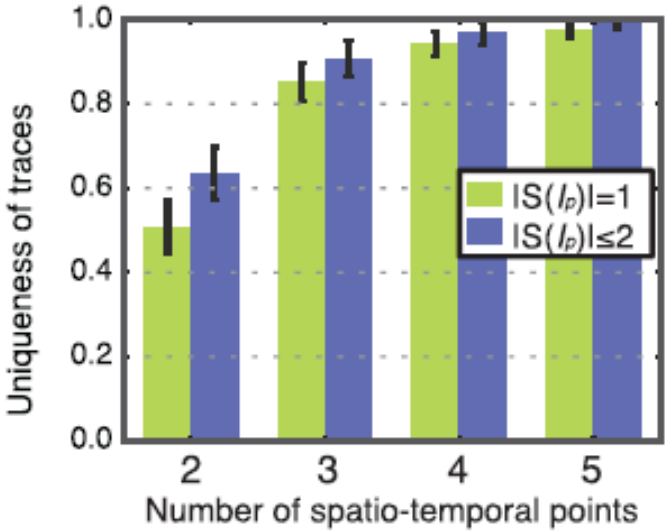
- **Indirect learning**: friends have similar attributes. A small value of conductance indicates a strong community. (Mislove et al., WSDM ’10.)

- **Content vs. Metadata**



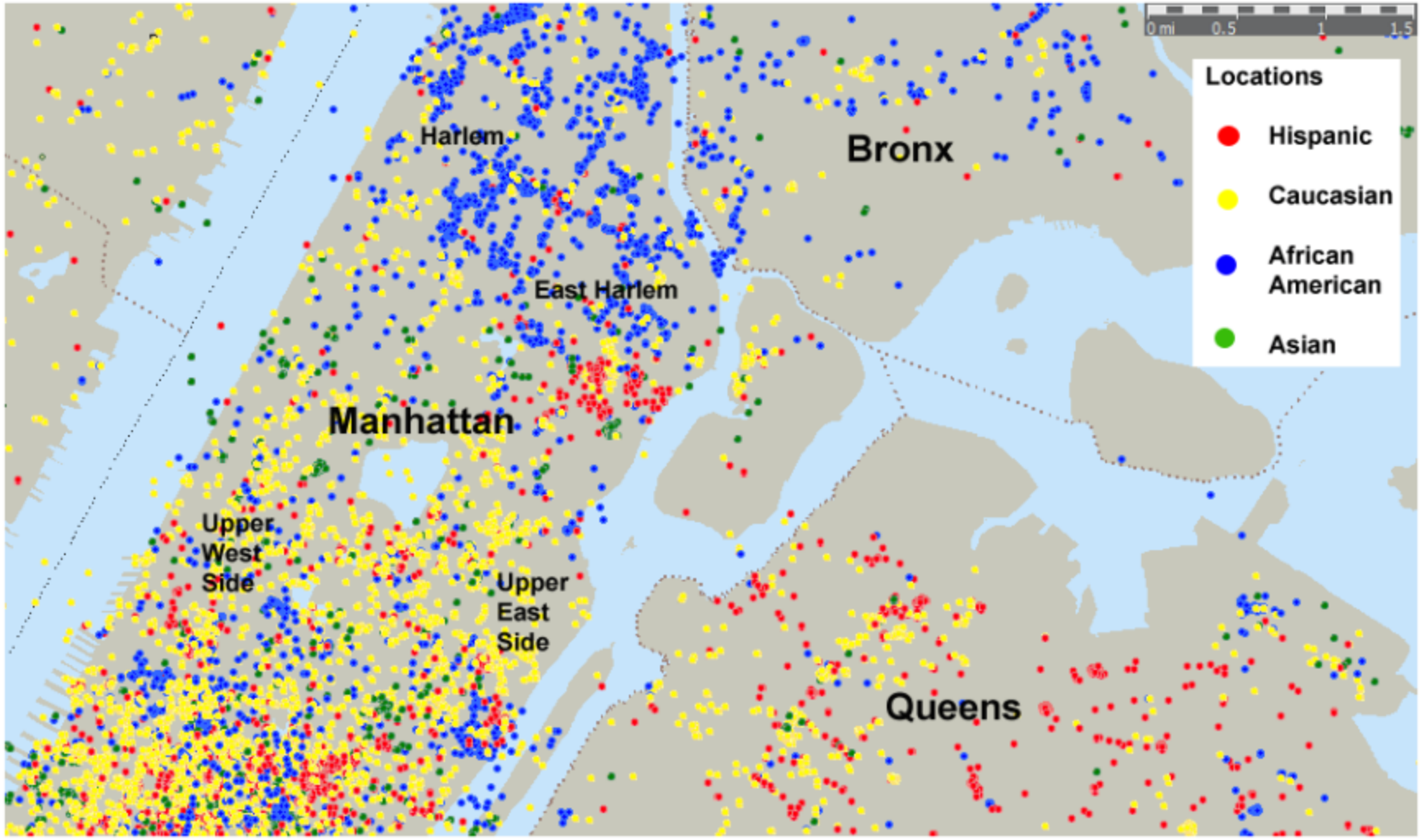
- **Locations can be often accurately predicted**: the empirically determined user entropy suggests that there is a 93% average predictability in user mobility, an exceptionally high value rooted in the inherent regularity of human behavior. (Song et al., Science ’10.)





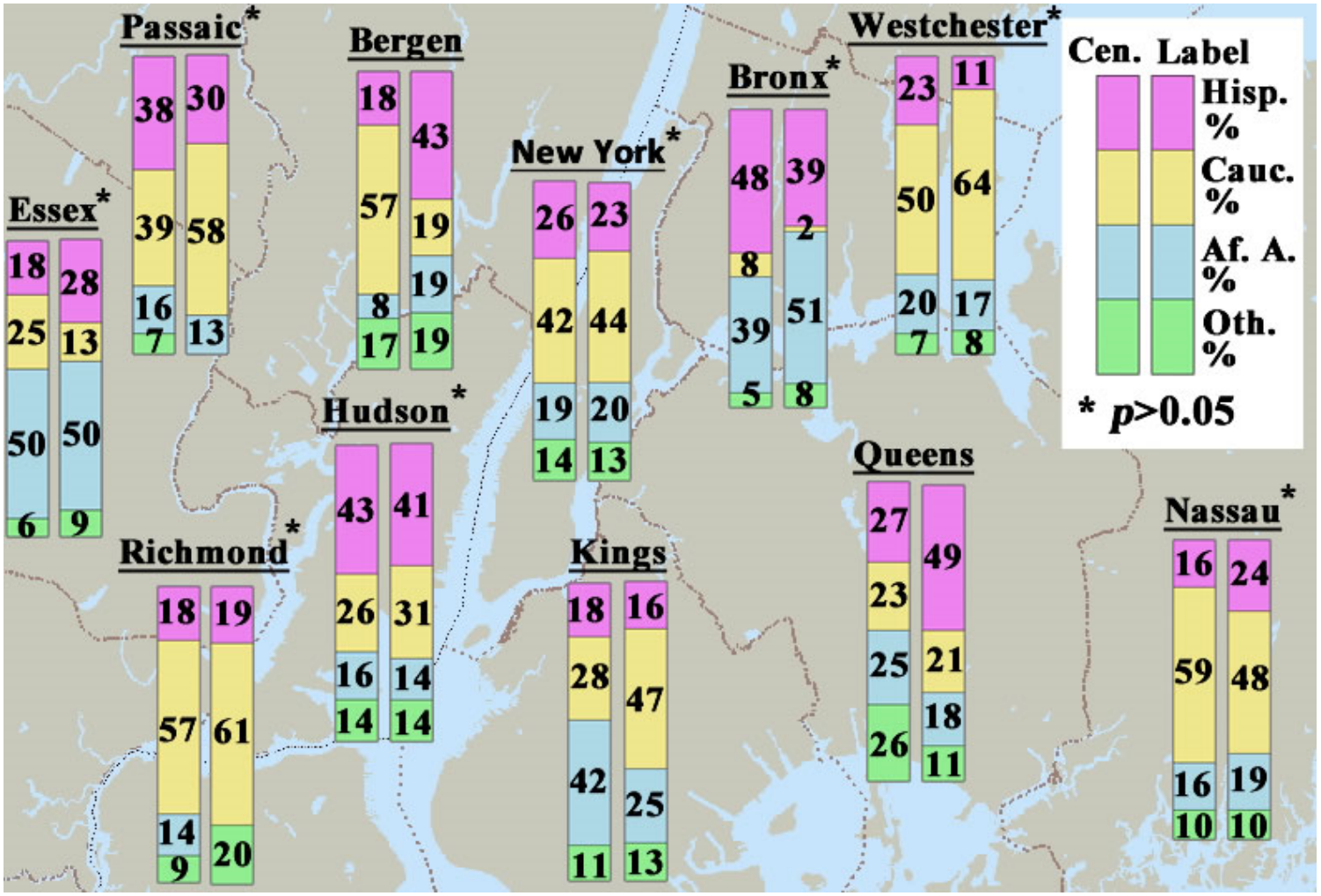
Location traces are often unique: randomly taking four spatio-temporal points is enough to uniquely characterize 95% of the population; based on 15 months of cell tower data for 1.5 M people.
(de Montjoye et al., Nature ’13.)

1. Learning
The Discriminative Power of Location Data



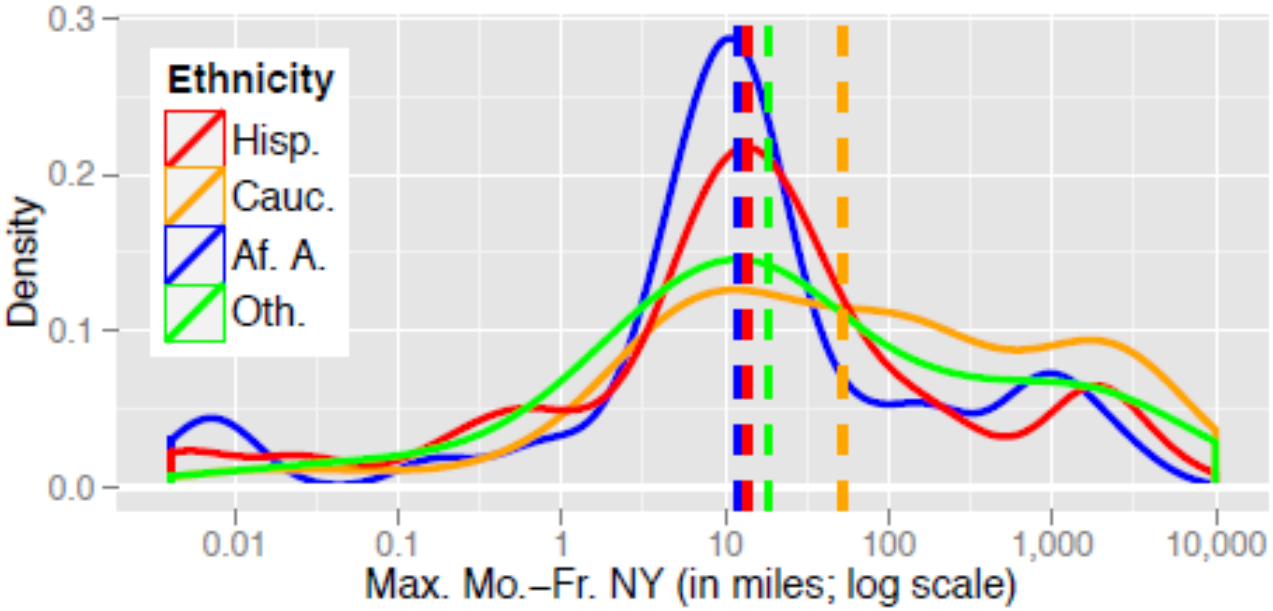
(Riederer et al., COSN '15)

1. Learning
The Discriminative Power of Location Data



(Riederer et al., COSN '15)

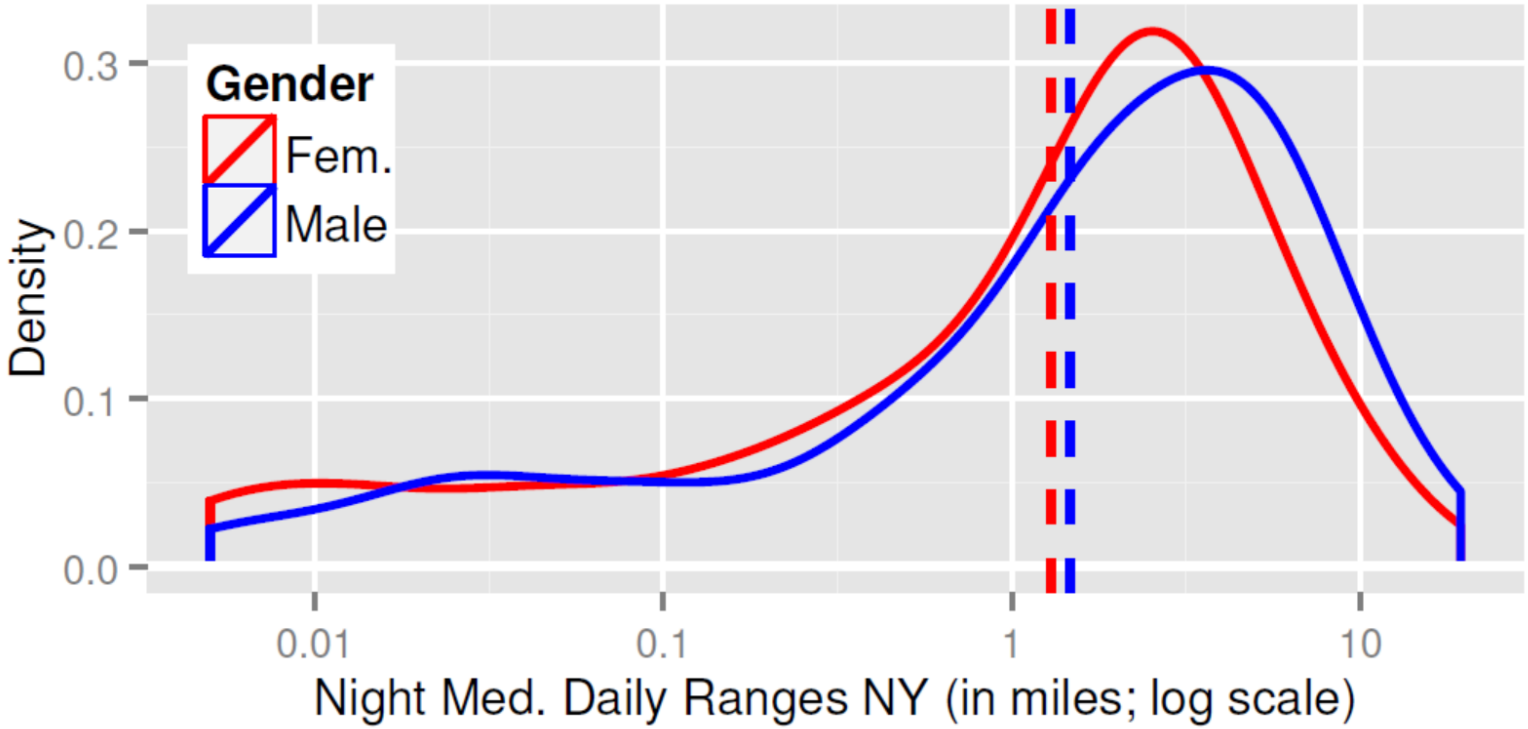
- **Daily Ranges:** different ethnic groups travel significantly different distances.



- **Visited Locations**: ethnicity can often be correctly inferred *solely* from visited locations and travel patterns. We are able to distinguish whether a user is Caucasian or has a minority background with an AUC value of 0.76.

  (Riederer et al., COSN '15)

1. Learning
The Discriminative Power of Location Data



(Riederer et al., COSN ‘15)

1. Learning
Naïve Bayes Example

Dan Jurafsky

$$\hat{P}(c) = \frac{N_c}{N}$$

$$\hat{P}(w \mid c) = \frac{count(w,c)+1}{count(c)+|V|}$$

	Doc	Words	Class
Training	1	Chinese Beijing Chinese	c
	2	Chinese Chinese Shanghai	c
	3	Chinese Macao	c
	4	Tokyo Japan Chinese	j
Test	5	Chinese Chinese Chinese Tokyo Japan	?

Priors:

$P(c) = \frac{3}{4}$

$P(j) = \frac{1}{4}$

Conditional Probabilities:

P(Chinese|*c*) = (5+1) / (8+6) = 6/14 = 3/7

P(Tokyo|*c*) = (0+1) / (8+6) = 1/14

P(Japan|*c*) = (0+1) / (8+6) = 1/14

P(Chinese|*j*) = (1+1) / (3+6) = 2/9

P(Tokyo|*j*) = (1+1) / (3+6) = 2/9

P(Japan|*j*) = (1+1) / (3+6) = 2/9

Choosing a class:

$P(c|d5) \propto 3/4 * (3/7)^3 * 1/14 * 1/14 \approx 0.0003$

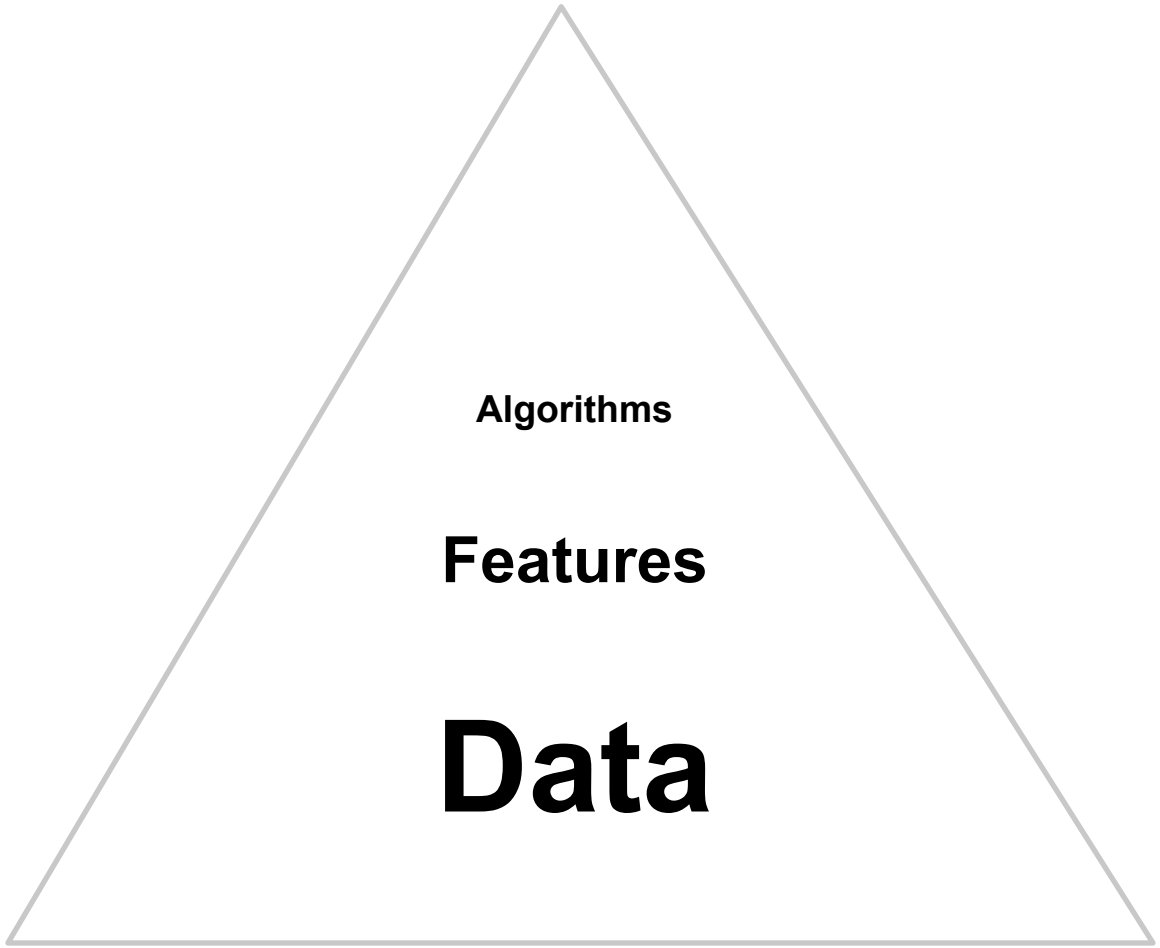
$P(j|d5) \propto 1/4 * (2/9)^3 * 2/9 * 2/9 \approx 0.0001$

https://web.stanford.edu/class/cs124/lec/naivebayes.pdf

Purpose: Create an accurate predictive model

- Overfitting vs. Underfitting
- Stopwords
- Information Gain
- TF-IDF

…


Algorithms
Features
Data

United States v. Jones, 132 SCt 945 (2012)

- "In the pre-computer age, the greatest protections of privacy were neither constitutional nor statutory, but practical. Traditional surveillance for any extended period of time was difficult and costly and therefore rarely undertaken." (Justice Alito, concurring opinion)

- Fundamental protection provided by the Fourth Amendment:

  The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no warrants shall issue, but upon probable cause, supported by oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.

- The question here: is warrantless GPS tracking *unreasonable?*

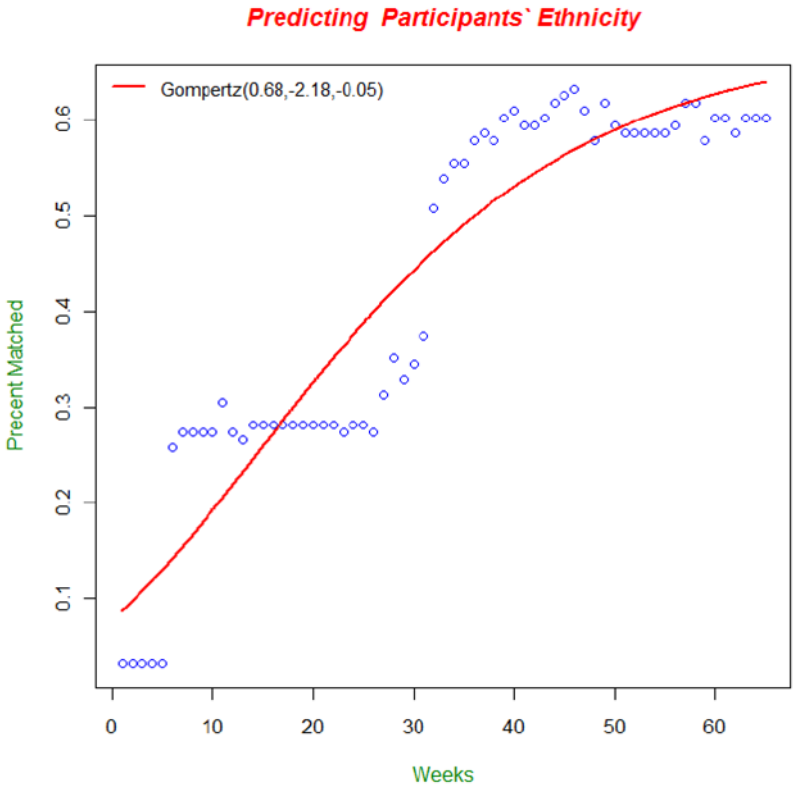
*"**[D]isclosed in [GPS] data ... [are] trips the indisputably private nature of which takes little imagination** to conjure: trips to the psychiatrist, the plastic surgeon, the abortion clinic, the AIDS treatment center, the strip club, the criminal defense attorney, the by-the-hour motel, the union meeting, the mosque, synagogue or church, the gay bar and on and on."* (United States v. Jones, Sotomayor, J., concurring, quoting People v. Weaver)

*"In this case, for four weeks, law enforcement agents tracked every movement that respondent made in the vehicle he was driving. We need not identify with precision the point at which the tracking of this vehicle became a search, for **the line was surely crossed before the 4-week mark.**"* (United States v. Jones, Alito, J., concurring.)

Before Jones, Fourth Amendment decisions had always evaluated each step of an investigation individually. Jones introduced what we might call a '**mosaic theory**' of the Fourth Amendment, […]. How should courts aggregate conduct to know when a sufficient mosaic has been created? (Kerr, Mich. '12)

Location privacy can be quantified based on the correctness of an inference attack. (Shokri et al., S&P '11.)

We made a suggestion. (Bellovin et al., N.Y.U. '14.)



(Altshuler et al., PASSAT '12)

Implications on Privacy

- Collecting data
- Learning of new information covered by law?

…

Questions?