# Multi-Label Learning with Millions of Labels for Query Recommendation

Rahul Agrawal Microsoft AdCenter

Archit Gupta
IIT Delhi

Yashoteja Prabhu Microsoft Research India Manik Varma
Microsoft Research India

#### Recommending Advertiser Bid Phrases





© 1996-2012 GEICO

geico auto insurance geico car insurance geico insurance www geico com care geicos geico com need cheap auto insurance wisconsin cheap car insurance quotes cheap auto insurance florida

all state car insurance

coupon code

#### **Query Rewriting**

"Absolutely cheapest car insurance"

geico auto insurance

geico car insurance

geico insurance

www geico com

care geicos

geico com

need cheap auto insurance

wisconsin cheap car insurance quotes

cheap auto insurance florida

all state car insurance coupon code

#### Ranking & Relevance Meta Stream



#### The GEICO Gecko

@TheGEICOGecko

The official Twitter home of the GEICO Gecko, helping people save hundreds on their car insurance.

Washington, DC · http://www.geico.com

1.542 TWEETS 1,237 **FOLLOWING**  9,113 **FOLLOWERS** 





The GEICO Gecko @TheGEICOGecko

Feb 23

Catching up on the nominees before tomorrow. I've always been a fan of the "Short Film" category. pic.twitter.com/zjHzpYkXR7

Details



The GEICO Gecko @TheGEICOGecko

Feb 18

Anyone know where the halfpipe is around here? Gotta practice my tailgrab. pic.twitter.com/fvQTy8Fu

tter.com/TheGEICOGecko/following

geico auto insurance

geico car insurance

geico insurance

www geico com

care geicos

geico com

need cheap auto insurance

wisconsin cheap car insurance quotes

cheap auto insurance florida

geico twitter

#### Recommending Advertiser Bid Phrases





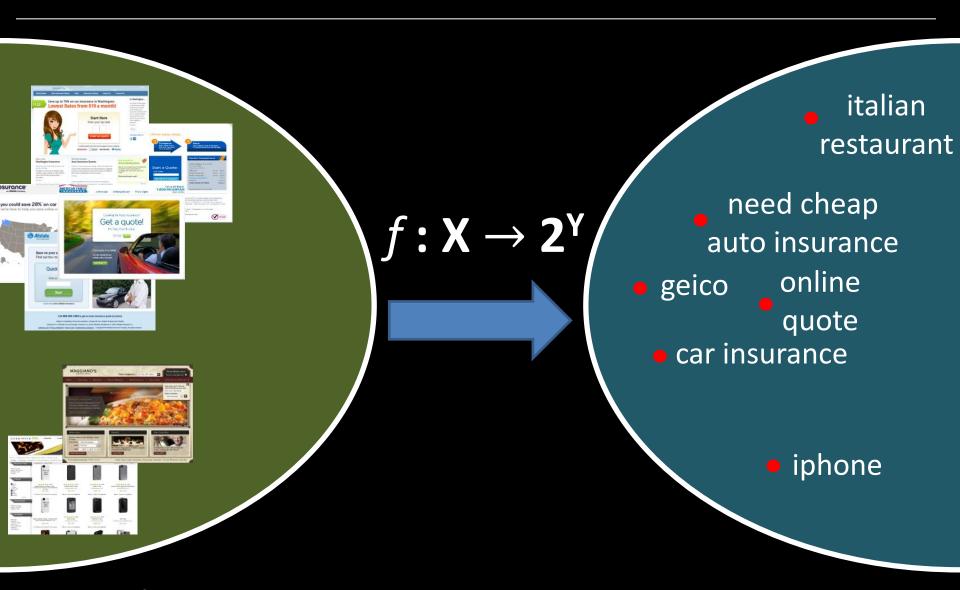
© 1996-2012 GEICO

geico auto insurance geico car insurance geico insurance www geico com care geicos geico com need cheap auto insurance wisconsin cheap car insurance quotes cheap auto insurance florida

all state car insurance

coupon code

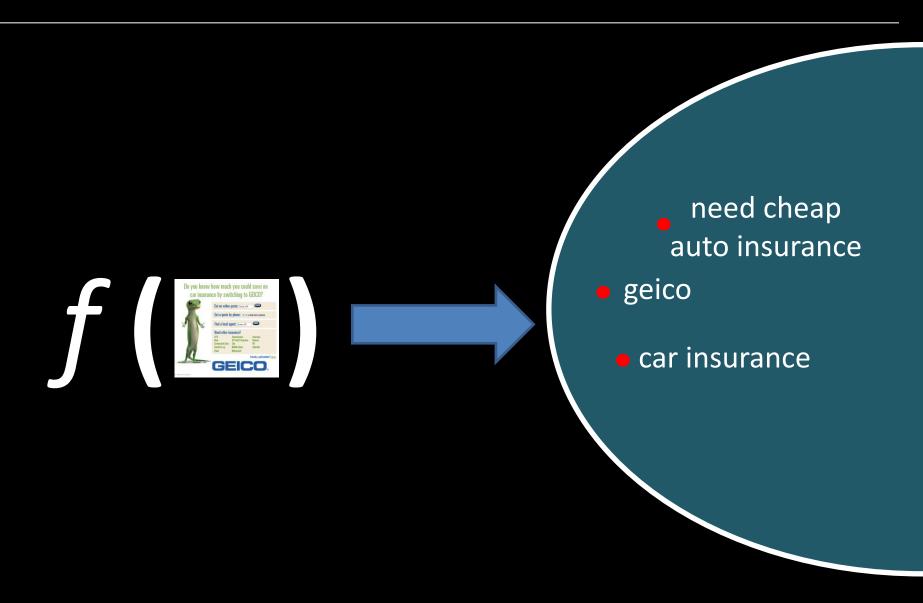
#### Learning to Predict a Set of Queries



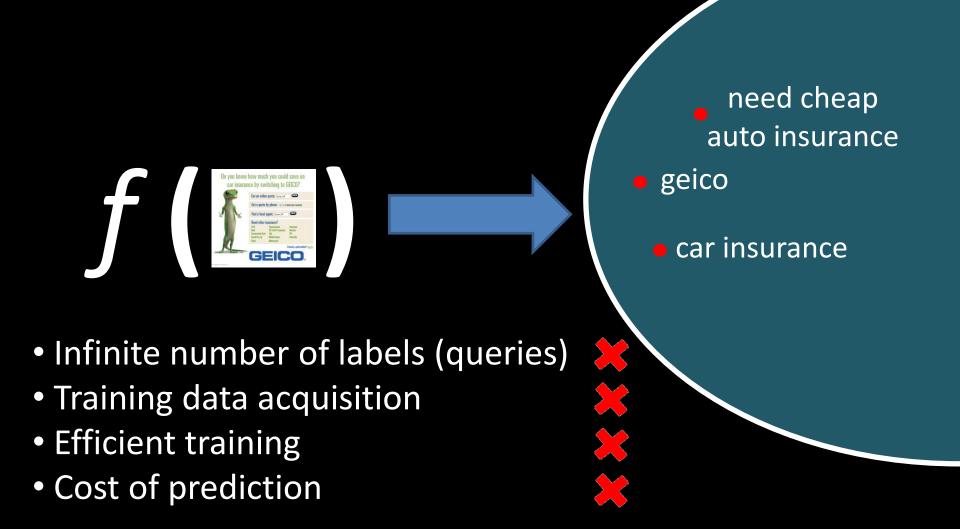
X: Ads

Y: Queries

#### Learning to Predict a Set of Queries



#### Multi-Label Learning Challenges



# Binary Classification & Ranking

$$h: (X, Y) \rightarrow \{\times, \checkmark\}$$

- Infinite number of labels (queries)
- Training data acquisition
- Efficient training
- Cost of prediction

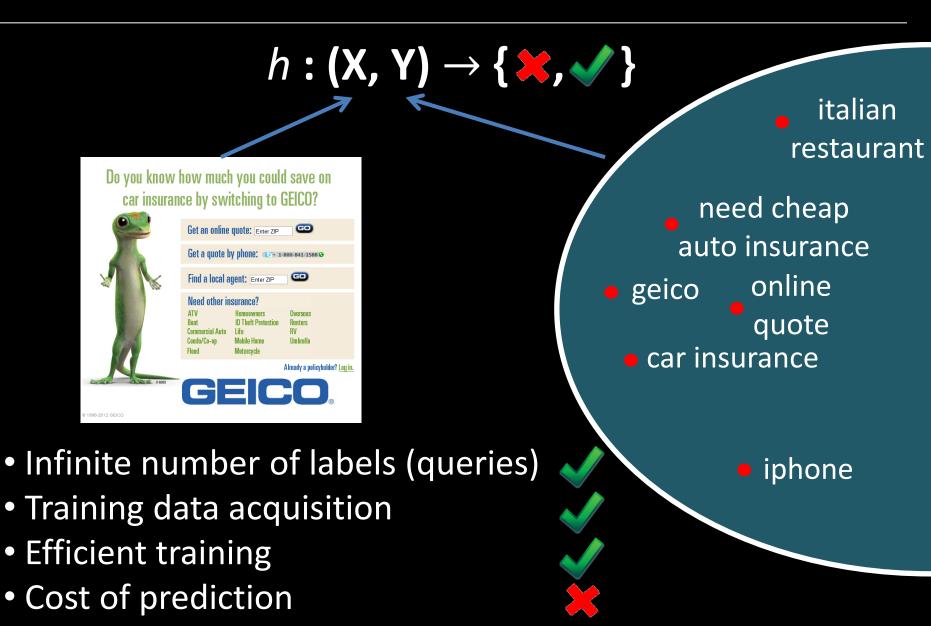








#### **Binary Classification**



#### Binary Classification – KEX







- Infinite number of labels (queries)
- Training data acquisition
- Efficient training
- Cost of prediction





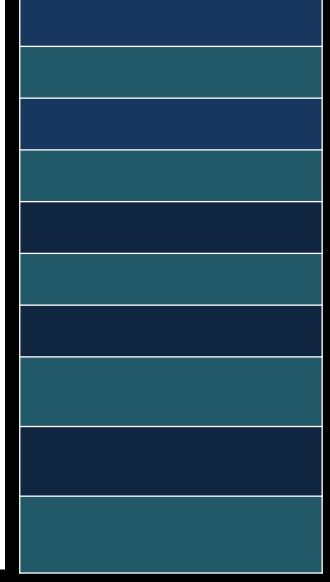




#### Query Recommendations by KEX

#### Do you know how much you could save on car insurance by switching to GEICO?





#### Query Recommendations by KEX

Do you know how much you could save on car insurance by switching to GEICO?





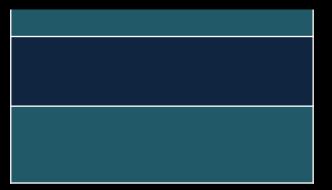
h car insurance  $\rightarrow ?$ 



holise par tens par conditions of the control of th







#### Query Recommendations by KEX

#### Simone & Sylvia

Boutique

Classes & Workshops

Plastic Pony

- Design Patternmaking Made to Measure
- Get In Touch

plastic ponies

simone

plastics

clothing and accessories

sylvia

pony clothing

couture

playground

plastic recycling

children's clothing

#### plastic pony playground couture



Our line of infant and children's clothing and accessories feature organic and natural fabrics, designed and styled with retro inspirations that will have your children showered with compliments.

Plastic Pony clothing is of superior quality. Made from ultra-comfortable, cozy materials.

Plastic Pony fashions are perfectly suited for the active lives of children. As for their parents, Plastic Pony clothes are very easy to maintain.

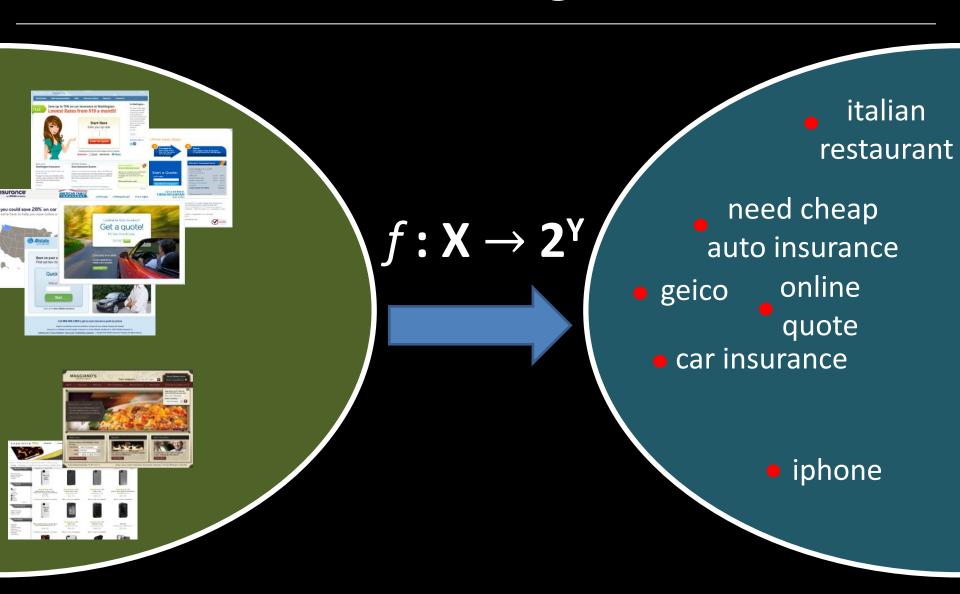
Painstaking care is also taken with respect to the patterns, ensuring great freedom of movement.

We make children's clothing starting at 3 months up to size 12.



About Simone & Sylvia

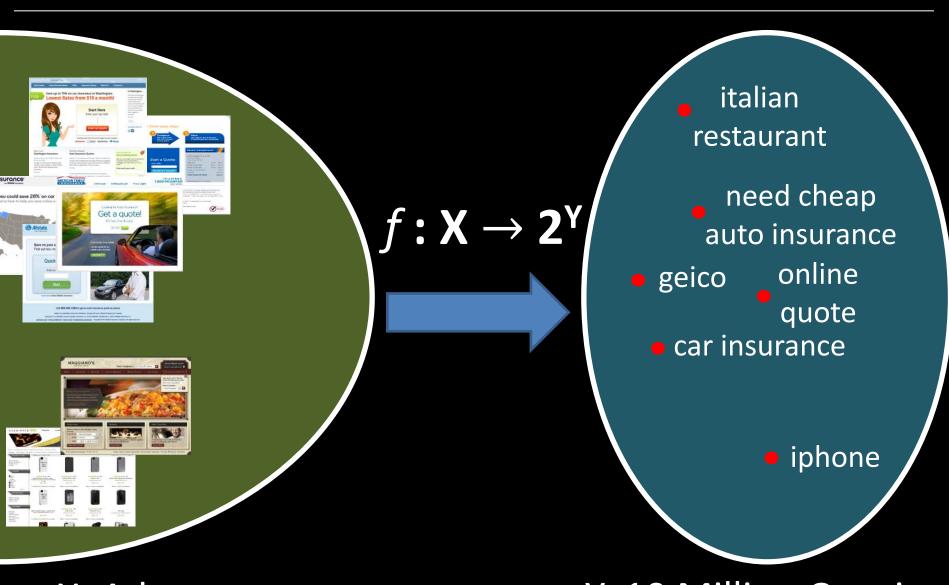
### Multi-Label Learning Formulation



X: Ads

Y: Queries

#### Learning with Millions of Labels



X: Ads

Y: 10 Million Queries

#### Multi-Label Random Forests

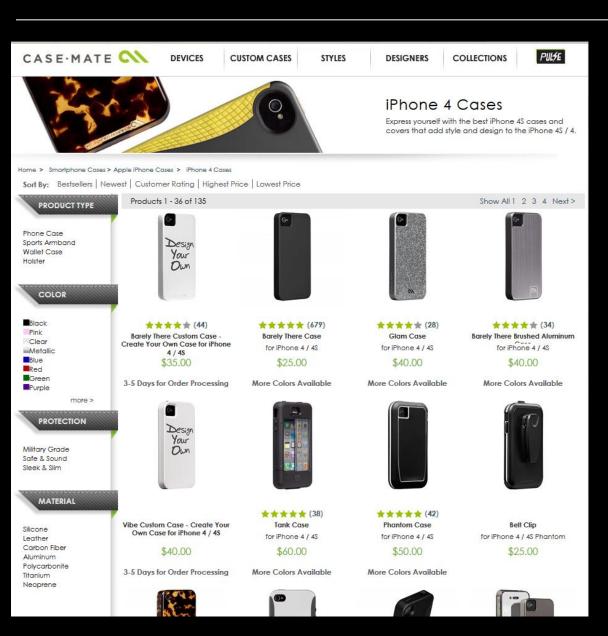
- We develop Multi-Label Random Forests with logarithmic prediction costs that make predictions in a few milliseconds.
- We train on 200 M points, 100 M categories and 10 M features in 28 hours on a grid with 1000 compute nodes.
- We develop a tree growing criterion which learns from positive data alone.
- We generate training data automatically from click logs.
- We develop a sparse SSL formulation to infer beliefs about the state of missing and noisy labels.

#### Training Data — Missing Labels

- No annotator can mark all the relevant labels for a data point.
- We have missing labels during
  - Training
  - Validation
  - Testing.
- Even fundamental ML techniques such as validation can go awry.
- One can't design error metrics invariant to missing labels.



#### Training Data and Features





TF-IDF Bag of Words Features

# **Training Labels**

C	CASE-MATE COUSTOM CASES STYLES DESIGNERS COLLECTIONS									
`	case for iphone	best iphone case	apple iphone 3g metallic slim fit case	best iphone nn4 cases						
HO S	iphone cases	best iphone cases	apple iphone 4g cases	black white premium bumper case apple iphone nn4 att						
	best iphone nn4 case	case iphone	apple iphone 4g premium soft silicone rubber black phone protector skin cover case	bunny rabbit silicone case skin iphone nn4 stand tail holder						
	iphone 3gs cases	otterbox universal defender case iphone nn4 black silicone black plastic	apple iphone nn4 cases	iphone case						
	iphone 4s case	sena iphone cases	belkin grip vue tint case iphone nn4 clear	iphone 4g cases						
	iphone case speck iphone case		best case iphone 4s	iphone 4gs cases						
77	iphone nn4 case switcheasy neo case iphone 3g black		best case iphone nn4	iphone 4s defender series case						
SS	3g iphone cases	waterproof iphone case	best iphone 3g cases	iphone case design						
	apple iphone cases	waterproof iphone cases	best iphone 4s case	iphone cases 3g						
F T N	best iphone 3g case	amazonbasics protective tpu case screen protector att verizon iphone nn4 iphone 4s clear	best iphone 4s cases	iphone cases 4g						

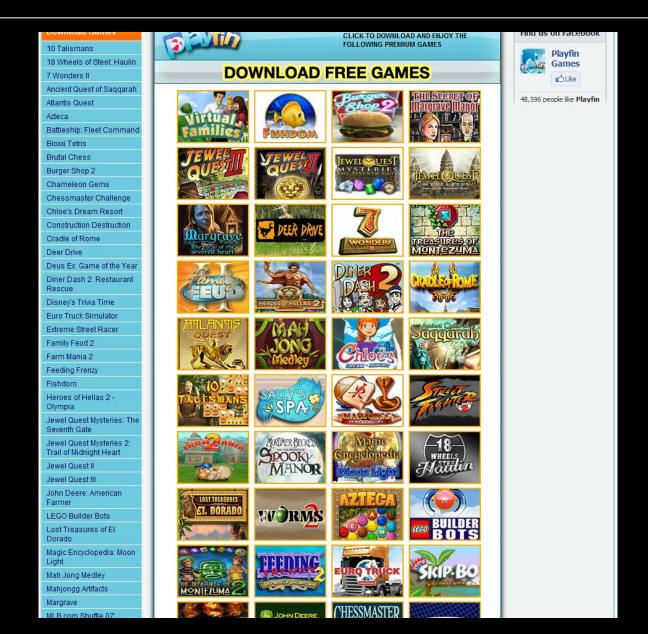
# Training Labels



# Missing and Noisy Labels

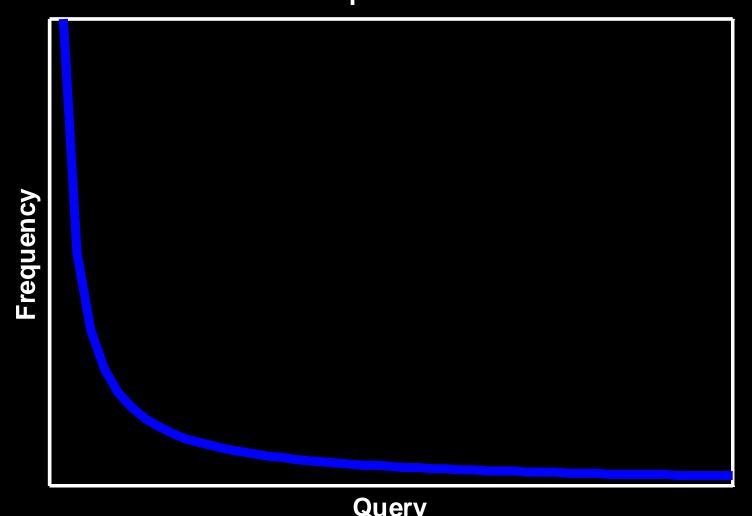
MACCIANIOS		
best italian restaurants philadelphia	italian restaurant chains	
italian restaurants	italian restaurant connecticut	
italian restaurant	italian restaurant district columbia	
italian restaurants arkansas	thai restaurant	
italian restaurants connecticut	thai restaurants	
italian restaurants idaho	restaurants	
italian restaurants phoenix	mexican restaurants	

# Missing and Noisy Labels



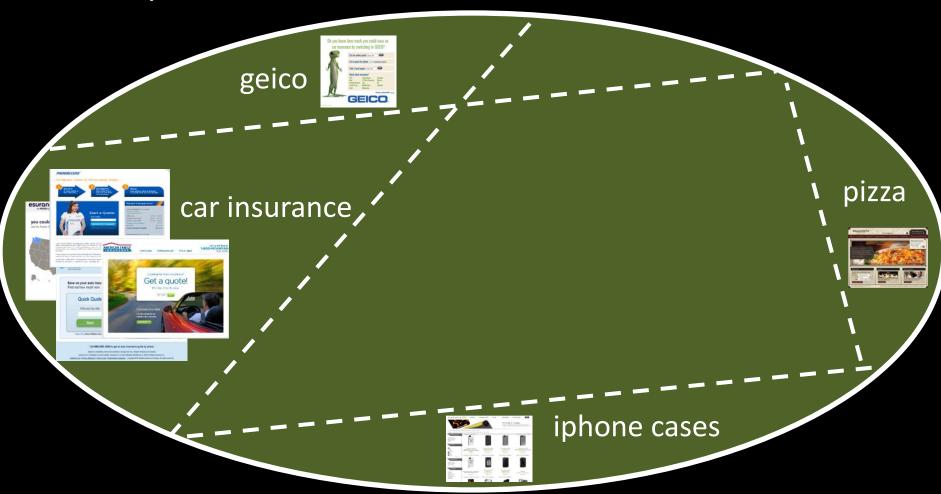
### **Biased Training Data**

Most labels will have very few positive training examples
 Zipf's Law



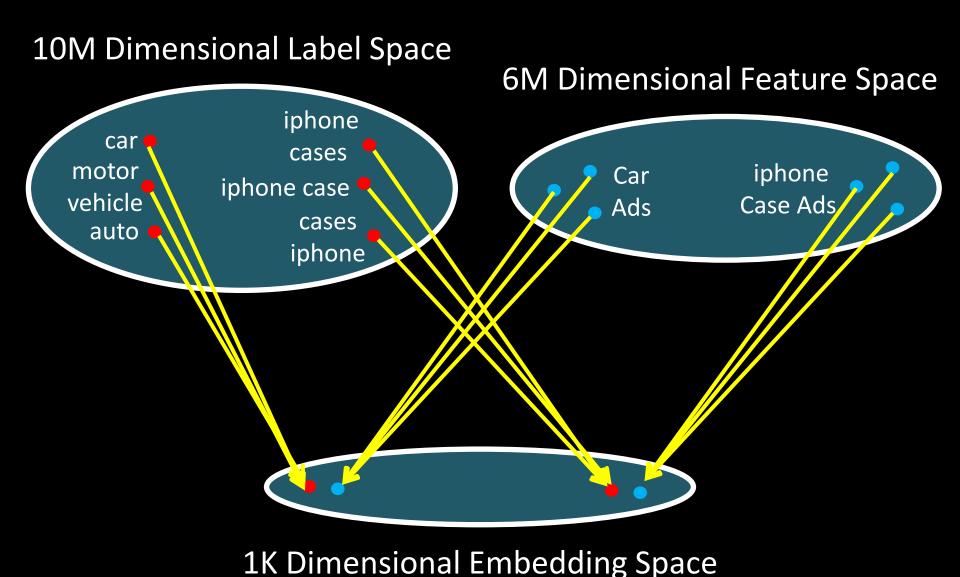
#### Multi-Label Prediction Costs

Linear prediction costs are infeasible



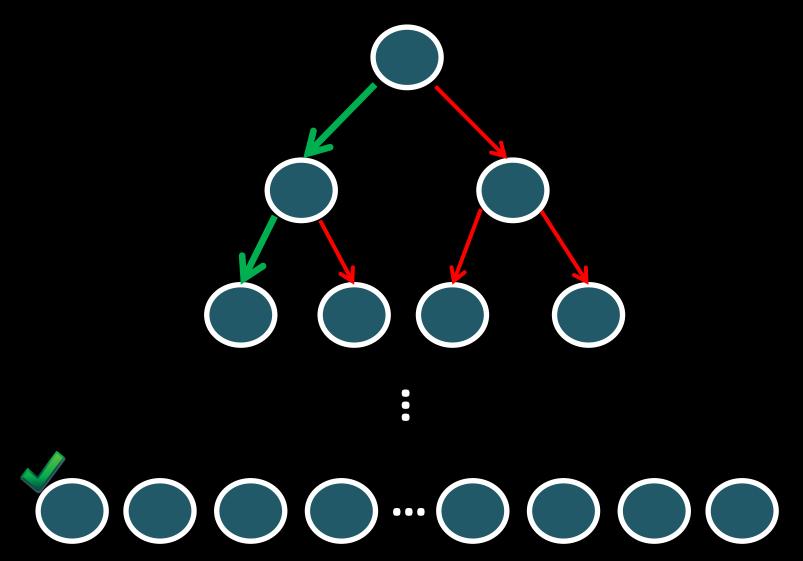
1-vs-All Classification

#### Label and Feature Space Compression



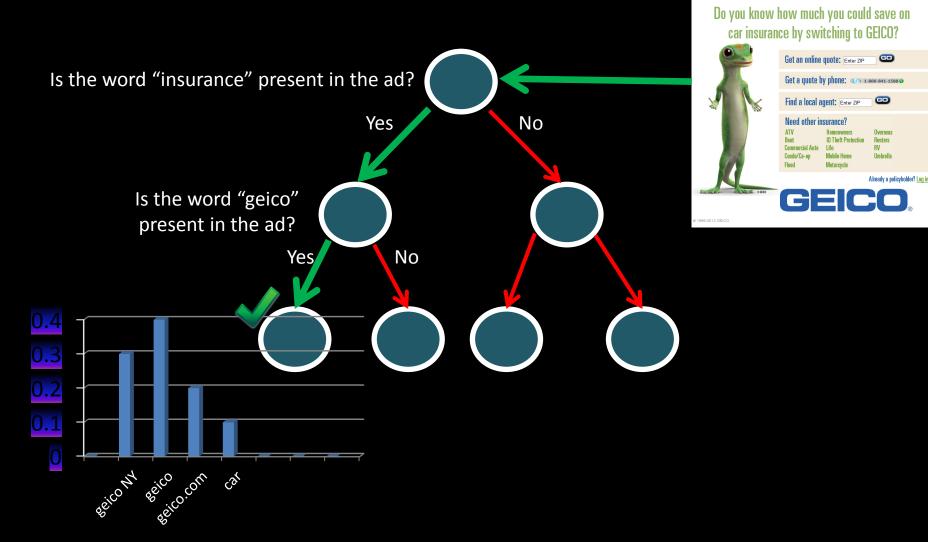
# Hierarchical Prediction

Prediction in logarithmic time

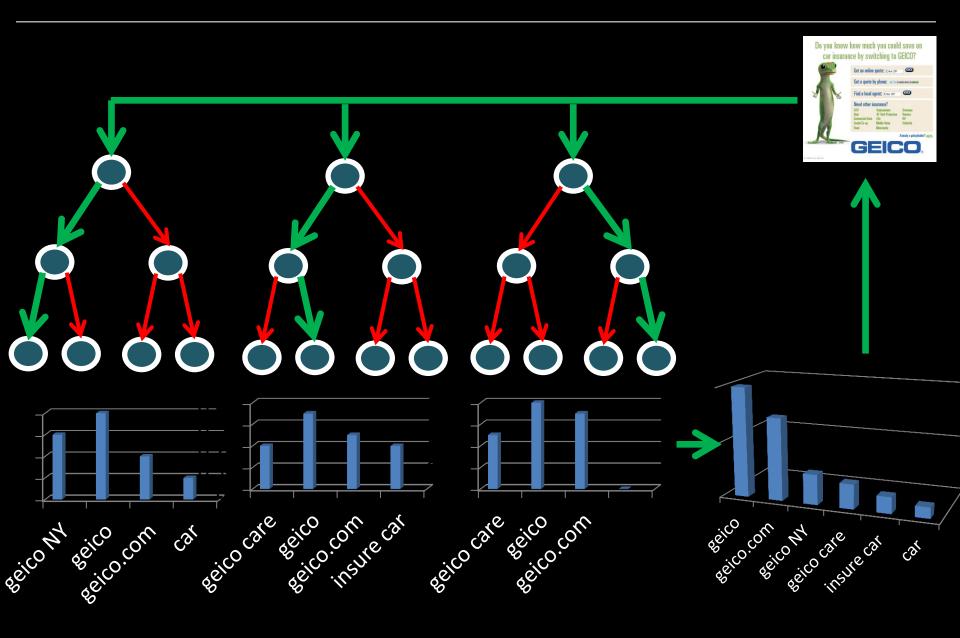


#### **Gating Tree Based Prediction**

Prediction in logarithmic time



# Ensemble of Randomized Gating Trees



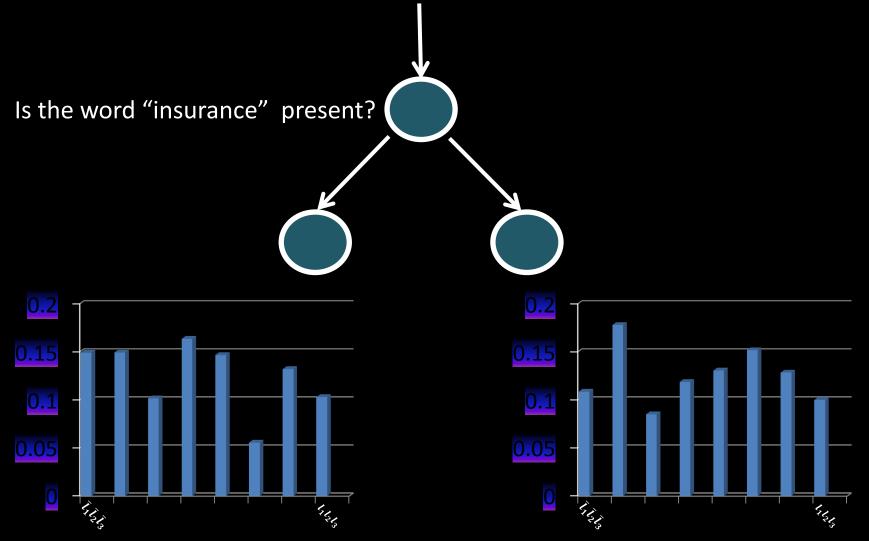
#### **Efficient Training**

- We seek classifiers and optimization algorithms that
  - Are massively parallelizable
  - Don't need to load the feature vectors (1 Tb) into RAM
  - Don't need to load the label matrix (100 Gb) into RAM

Number of training points	200 Million	
Number of labels	100 Million	
Dimensionality of feature vector	10 Million	
Number of cores	500 – 1000	
RAM per core	2 Gb	
Training time	28 hours	

#### Multi-Label Random Forests

The splitting cost needs to be calculated in a 2<sup>10M</sup> space

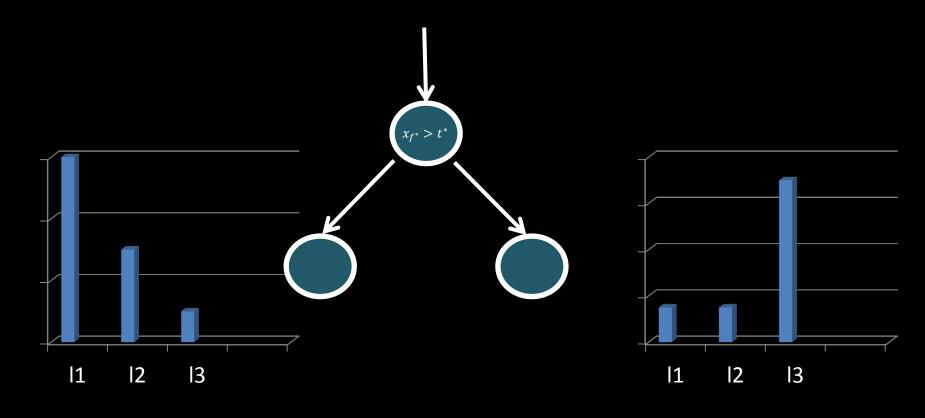


#### Learning from Positively Labeled Data

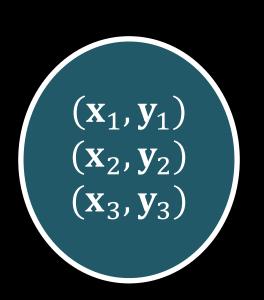
• Split condition :  $x_{f^*} > \overline{t^*}$ 

$$(f^*, t^*) = \operatorname{argmin}_{f,t} n_l \sum_k p_l(l_k) (1 - p_l(l_k)) + n_r \sum_k p_r(l_k) (1 - p_r(l_k))$$

$$p(l_k) = \sum_i p(l_k | \operatorname{ad}_i) p(\operatorname{ad}_i)$$



#### Multi-Label Random Forests

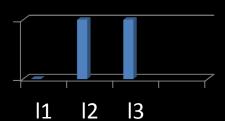


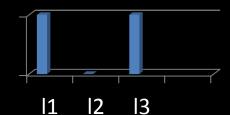
$$\mathbf{x}_1, \mathbf{y}_1 = \{l_2, l_3\}$$

$$\mathbf{x}_2, \mathbf{y}_2 = \{l_1, l_3\}$$

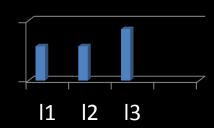
$$\mathbf{x}_3, \mathbf{y}_3 = \{l_1, l_2, l_3\}$$











# Query Recommendation Data Sets

Data set statistics

Data Set	# of Training Points (M)	# of Test Points (M)	# of Dimensions (M)	# of Labels (M)
Wikipedia	1.53	0.66	1.89	0.97
Ads1	8.00	0.50	1.58	1.22
Web	40.00	1.50	2.62	1.22
Ads2	90.00	5.00	5.80	9.70

#### Performance Evaluation – Precision@k

• We use loss functions where the penalty incurred for predicting the real (but unknown) ground truth is never more than that of predicting any other labelling

$$L(\mathbf{y}^*, \mathbf{y}_{\text{Observed}}) \le L(\mathbf{y}, \mathbf{y}_{\text{Observed}}) \quad \forall \mathbf{y} \in Y$$

Hamming Loss

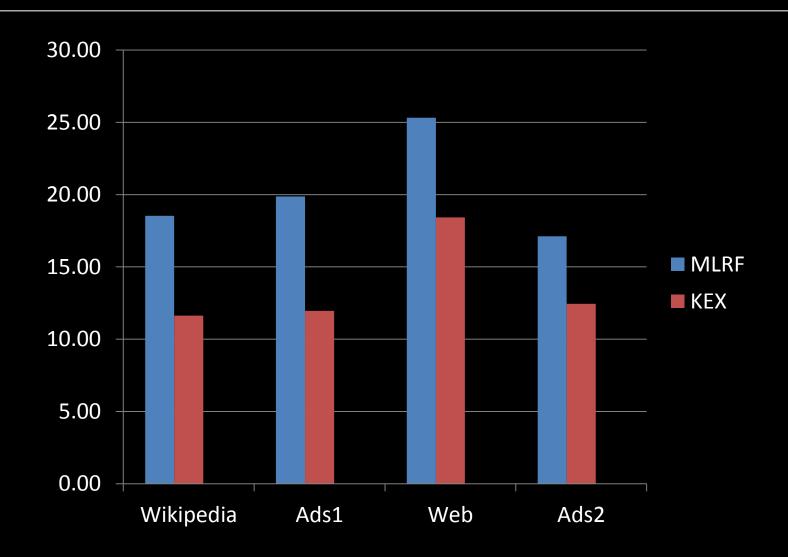


• Precision at *k* 



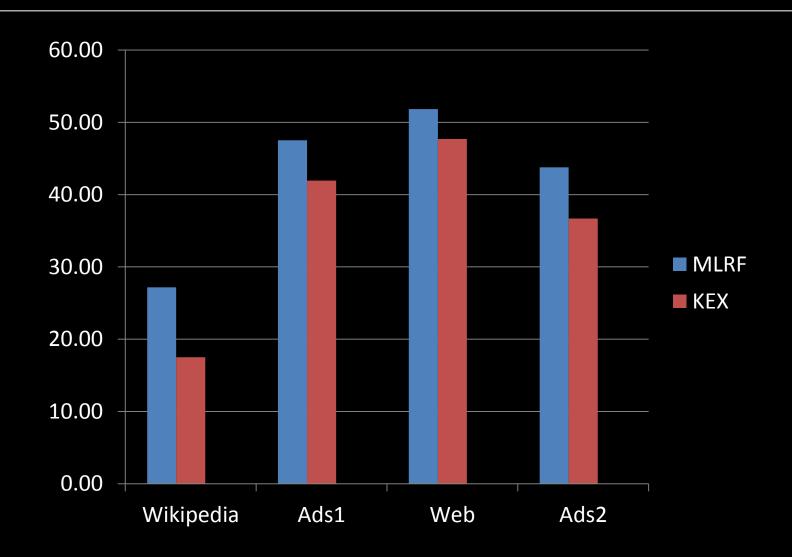
We found Precision at 10 to be robust for our application.

#### Query Recommendation Results



Percentage of top 10 predictions that were clicked queries

### Query Recommendation Results



Percentage of top 10 predictions that were relevant

# Do you know how much you could save on car insurance by switching to GEICO?



Get an online quote: Enter ZIP œ Get a quote by phone: 1-800-841-1588 6 œ Find a local agent: Enter ZIP Need other insurance? ATV Homeowners Overseas Boat ID Theft Protection Renters Commercial Auto Life RV Condo/Co-op Mohile Home Umbrella Flood Motorcycle

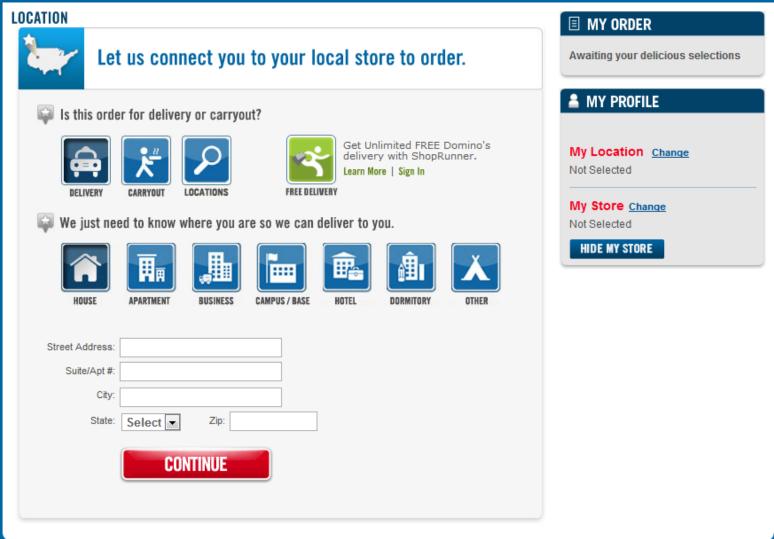
Already a policyholder? Log in.



### Geico Car Insurance

	KEX	MLRF			
		geico auto insurance			
		geico car insurance			
		geico insurance			
		www geico com			
		care geicos			
		geico com			
		need cheap auto insurance			
		wisconsin cheap car insurance quotes	ı in.		
		cheap auto insurance florida			
1996		all state car insurance coupon code			







# Domino's Pizza

HOME	KEX	MLRF	Tracker   Log In
LOCATION		dominos	iciano coloctiono
Let		dominos pizza	icious selections
DELIVERY		domino pizza	<u>Change</u>
We just nee		domino pasta bowls	<u>ge</u>
HOUSE		domino pizza coupons	
Street Address: Suite/Apt #: [ City: [		domino pizza deals	
State: [		domino pizza locations	
		domino pizza menu	
Corporate Info // Nutrition		domino pizza online	FIND US ON FACEBOOK ON FACEBOOK

#### Simone & Sylvia

- Boutique
- Classes & Workshops
- Plastic Pony
- Design Patternmaking Made to Measure
- About Simone & Sylvia

Get In Touch

#### plastic pony playground couture



Our line of infant and children's clothing and accessories feature organic and natural fabrics, designed and styled with retro inspirations that will have your children showered with compliments.

Plastic Pony clothing is of superior quality. Made from ultra-comfortable, cozy materials.

Plastic Pony fashions are perfectly suited for the active lives of children. As for their parents, Plastic Pony clothes are very easy to maintain.

Painstaking care is also taken with respect to the patterns, ensuring great freedom of movement.

We make children's clothing starting at 3 months up to size 12.



# Simone & Sylvia Kid's Clothing

	KEX	MLRF	
Boutique	plastic ponies	toddlers clothes	via
	simone	toddlers clothing	
	plastics	toddler costumes	
	clothing and accessories	children clothes sale	
	sylvia	children clothes	<b>b</b>
	pony clothing	designer children clothes	
	couture	cute children clothes	
	playground	retro clothing	
	Plastic recycling	retro baby clothes	
	children's clothing	baby clothing	



#### INUS THUMEIS

FLOWERS ♦ GOURMET BASKETS ♦ BALLOONS ♦ GIFTS
HOME GARDENING SUPPLIES & DECOR

Celebrating life...one occasion at at time...

Home | Corporate & Hospitality | Funeral/Sympathy | Gardening | Careers | About Us

# Deliveries Everyday To The Seattle Area

**ORDER NOW** 





#### Follow Us!





Local 206-722-2200 Toll Free 800-455-3701

4873 Rainier Ave South Seattle, WA 98118

**Business Hours** 

Monday-Thursday

8:30am - 5:30pm

Friday

8:30am - 6:30pm

Saturday

9am - 5pm

Sunday

10am - 2pm

# **KCS Flowers**

	KEX	MLRF	
	funeral flowers	flowers delivery	
<b>⊯</b> 342	sympathy funeral flowers	funeral arrangements	
<b>. f</b> Like	web home	birthday flowers	
20	bleitz funeral home	funeral flowers funeral planning	
<b>80</b> 487:	funeral flowers discount		
S	yarington's funeral home	flowers valentines	
M	harvey funeral home	free delivery flowers	
}	green lake funeral home	cheap flowers	
8	howden kennedy funeral home	florists	-
	arranging flowers	cheap flowers funeral	-

isiness · Home a ruminy · Air Froducts

#### **Exclusive Offers**

Become a Vistaprint Insider to receive exclusive offers and tips.

Enter e-mail address

Sign Up

#### Customizable Apparel

#### The perfect fit at a perfect price.

- · Great for on the job and off
- · Made from superior quality fabric
- Choose from 100s of designs
- No minimum orders or setup fees

#### Details and Pricing

#### **Printed T-shirts**

With our fully customizable, 100% pre-shrunk cotton T-shirts, you'll be able to advertise and promote your business wherever you go



Starting at \$11.99

Get Started

#### Printed Ladies' T-shirts

Our most feminine fit – and the most stylish way to display your logo or message.



Starting at \$13.99

**Get Started** 

#### **Embroidered Men's Polo Shirts**

100% cotton, high-quality polos you'll be proud to put your name on. They look and feel great, wash after wash.



Starting at \$19.99

Get Started

#### **Embroidered Ladies' Polo Shirts**

The same high-quality look and feel as our men's polo, but slightly more tailored for a flattering fit.



Starting at \$19.99

Get Started

# Vistaprint Designer T-Shirts

Exclusive Offers  Become a Vistaprint to receive exclusive of	KEX	MLRF	Search &
and tips.  Enter e-mail address  Sign Up	embroidered apparel	custom t shirts	
	custom apparel	funny t shirts	o display your
×	readymade apparel	hanes beefy t shirts	· ·
×	customizable	hanes t shirts	
×	apparel	long sleeve t shirts	Mildery C.
	customizable apparel	personalized t shirts	t Started )
×	leading print	printed t shirts	; polo, but
×	online business cards	retro gamer t shirts	T
×	apparel and accessories	t shirts	A+Nails Any place property
×	own text	buy custom t shirts	t Started )





#### Privacy Policy

\*Average annual savings based on savings reported by customers from many states who called our call center between 1/11 and 12/11 and switched to MetLife Auto & Home where our quoted premium was less than the disclosed prior carrier's premium. Source: MetLife Auto & Home internal research (2012).

MetLife Auto and MetLife Auto & Home are brands of Metropolitan Property and Casualty Insurance Company and its affiliates: Metropolitan Casualty Insurance Company, Metropolitan Direct Property and Casualty Insurance Company (CA Certificate of Authority: 6730; Warwick, RI), Metropolitan General Insurance Company, Metropolitan Group Property and Casualty Insurance Company (CA COA: 6393; Warwick, RI), and Metropolitan Lloyds Insurance Company of Texas, all with administrative home offices in Warwick, RI. Coverage, rates, and discounts are available in most states to those who qualify.

© 2012 MetLife Auto & Home PEANUTS © PEANUTS Worldwide, LLC L0911208373[exp 1112][All States][DC]

### Metlife Auto Insurance

Met	KEX	MLRF		
o me	etlife auto home insurance	metlife auto insurance	<u></u>	
	auto home insurance	auto Insurance		
	auto insurance	car Insurance		
SE	massachusetts	automobile Insurance		
	metlife agent	geico insurance		
	driver discount	cheap car insurance		
	additional cost	metlife auto	_	
*Average ar	saving benefits	insurance broker	re our	
quoted pren MetLife Autr Property and	car discount	insurance	in Direct urance	
available in 1  © 2012 Mett PEANUTS © L091120837	auto quote	home insurance		

Menu Coupons Gift Cert Mailing List Review Order Online Food Delivery



A Taste Your Heart Glows

This page is under construction. Please visit us at Facebook.com/WantaThai.



Mon-Thu: 11 AM - 9 PM

Fri: 11 AM - 9:30 PM Sat: 12 PM - 9:30 PM

Sun: 12 PM - 9 PM

#### WELCOME TO WANTA THAI CUISINE

12.24.2011

Wanta Thai Cuisine is a contemporary Thai eatery that transports you to friendly Thai culture. We serve healthy and quality food for a reasonable price, along with generally good service at a convenient location. Visit us to proof why we are one of the best Thai restaurant in Redmond (near Bellevue), WA.

Wanta is a Pali word meaning as showing respect. When Thai people meet, we greet and pay homage with gentle Wanta manner to each other (referred to as the Wai in Thai). Guests will appreciate the traditional Thai service and world-renowned hospitality. You will enjoy delicious Thai food along with humble service in a warm family-friendly atmosphere. We respect customer's time as we prioritize that your time is money. During the lunch time, we accommodate and welcome your rush hour. We serve quicklity food, not fast food. Whereas dinner time, we associate with fine relationship. We make the atmosphere comfy so customers can enjoy the excellent food as in their own living room.

### Wanta Thai Restaurant

M	KEX	MLRF		
	authentic thai restaurant	thai restaurant		
1	delicious thai food	thai restaurants		
	thai cuisine	mexican restaurants		
	thai restaurant	cheap hotels		
<b>)</b>	thai food	hotels		
	wanta	fast food restaurants		
	best thai restaurant	restaurants coupons		
N	thai eateries	best web hosting restaurants	We at a near pay	
s s	thai	vegetarian foods	will Thai e as rush	
,	contemporary thai	new york restaurants	tine	

MAGGIANO'S

Find a Maggiano's: City, State, ZIP, Address

ONLINE RESERVATIONS

Reserve your table now!

MENU

" Locations " Delivery " Online Ordering " Private Events " Gift Cards

E-MAIL CLUB SIGN-UP









best italian restaurants philadelphia	italian restaurant chains		
italian restaurants	italian restaurant connecticut		
italian restaurant	italian restaurant district columbia		
italian restaurants arkansas	thai restaurant		
italian restaurants connecticut	thai restaurants		
italian restaurants idaho	restaurants		
italian restaurants phoenix	mexican restaurants		

### **Thai-Italina Restaurant**

thaiitalina.com

262-514-2600

New Year, New Look, New Menu Come and Check it out

Thai/Italina Restaurant 220B E Main Street Waterford, WI 53185 United States ph: 262-514-2600

HOME MENU CONTACT US We have been in Waterford since 2004.

Best Thai food in racine county.

We make it fresh and tasty.







#### **Business Hour**

Monday 4:00pm......8:30pm

Tuesday-Friday 11:30am......8:30pm Saturday 4:00pm.....9:00pm

#### Close on Sunday

Master and Visa Card accepted

No personal check please





# Compensating for Missing Labels



Auto insurance quotes



Progressive insurance



Case-mate phone cases

0.5



American family insurance



Maggiano's restaurant



0.7

0.9

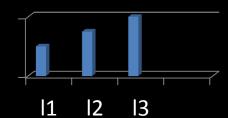
**Esurance** 

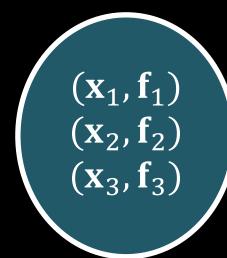


Allstate auto insurance

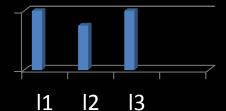
### Training on Belief Vectors

$$\mathbf{x}_1, \mathbf{y}_1 = \{l_2, l_3\}, \mathbf{f}_1$$

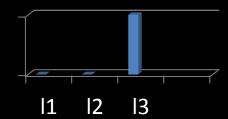




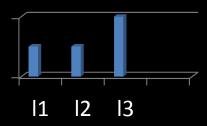
$$\mathbf{x}_2, \mathbf{y}_2 = \{l_1, l_3\}, \mathbf{f}_2$$



$$\mathbf{x}_3, \mathbf{y}_3 = \{l_1, l_2, l_3\}, \mathbf{f}_3$$



$$p(\mathbf{f})$$



### Sparse Semi-Supervised Learning

Graph-based SSL optimizes label belief smoothness and fidelity to original labels

$$\mathbf{F^*} = \operatorname{Min}_{\mathbf{F}} \frac{1}{2} \operatorname{Tr} \left( \mathbf{F}^t \left( \mathbf{I} - \mathbf{D}^{-\frac{1}{2}} \mathbf{W} \mathbf{D}^{-\frac{1}{2}} \right) \mathbf{F} \right) + \frac{\lambda}{2} ||\mathbf{F} - \mathbf{Y}||^2$$
s. t.  $|\mathbf{F}|_0 \le K$ 

of W

$\mathbf{W}_{MXM}$	Document-document similarity matrix
$\mathbf{D}_{MXM}$	Diagonal matrix representing the row sums
$\mathbf{Y}_{MXL}$	0/1 label matrix
$\mathbf{F}_{MXL}$	Real valued label belief matrix
λ	Trade-off Hyperparameter
M	Number of documents
L	Number of labels
17	

Sparsity constant

# Sparse Semi-Supervised Learning

 Graph-based SSL optimizes label belief smoothness and fidelity to original labels

$$\begin{aligned} \mathbf{F*} &= \text{Min}_{\mathbf{F}} \ \frac{1}{2} \mathbf{\Sigma}_{i=1..L} \mathbf{\Sigma}_{j=1..M} \ w_{jl} * (\frac{\mathbf{F}_{ij}}{\sqrt{\mathbf{D}_{jj}}} - \frac{\mathbf{F}_{il}}{\sqrt{\mathbf{D}_{ll}}})^2 + \frac{\lambda}{2} \ \Sigma_{i=1..M} \ (F_{ij} - Y_{ij})^2 \\ &\text{s. t.} \ |\mathbf{F}|_0 \leq K \end{aligned}$$

**W**<sub>MXM</sub> Document-document similarity matrix

 $\mathbf{D}_{MXM}$  Diagonal matrix representing the row sums of W

 $\mathbf{Y}_{MXL}$  0/1 label matrix

 $\mathbf{F}_{MXL}$  Real valued label belief matrix

 $\lambda$  Trade-off Hyperparameter

M Number of documents

L Number of labels

K Sparsity constant

### **Iterative Hard Thresholding**

Sparse SSL formulation

$$\mathbf{F}^* = \operatorname{Min}_{\mathbf{F}} J(\mathbf{F}) = \frac{1}{2} \operatorname{Tr} \left( \mathbf{F}^t \left( \mathbf{I} - \mathbf{D}^{-\frac{1}{2}} \mathbf{W} \mathbf{D}^{-\frac{1}{2}} \right) \mathbf{F} \right) + \frac{\lambda}{2} ||\mathbf{F} - \mathbf{Y}||^2$$
s. t.  $|\mathbf{F}|_0 \le K$ 

 The iterative hard thresholding algorithm converges to a global/local optimum

$$\mathbf{F}_{0} = \mathbf{Y}$$

$$\mathbf{F}_{t+\frac{1}{2}} = \frac{1}{\lambda+1} \mathbf{D}^{-\frac{1}{2}} \mathbf{W} \mathbf{D}^{-\frac{1}{2}} \mathbf{F}_{t} + \frac{\lambda}{\lambda+1} \mathbf{Y}$$

$$\mathbf{F}_{t+1} = \text{Top}_{K}(\mathbf{F}_{t+\frac{1}{2}})$$

### Iterative Hard Thresholding

- If  $Y_{ij} \in \{0, 1\}$  and **W** is positive definite then
  - The sequence  $F_0$ ,  $F_1$ , ... converges to a stationary point  $F^*$ .

• 
$$J(\mathbf{F}_0) \ge J(\mathbf{F}_1) \ge \cdots \ge J(\mathbf{F}^*)$$

- If  $|F^*|_0 < K$  then  $F^*$  is a globally optimal solution
- If  $|F^*|_0 = K$  then  $F^*$  is a locally optimal solution

$$J(F^*) - J(F^+) \le Min(\frac{\lambda}{2}(K + |Y|_0), \frac{\lambda + 1}{2}(ML - K)\alpha_K(F^*)\sqrt{|Y|_0})$$

### Semi-Supervised Learning Results

 Precision@10 as judged by automatically generated click labels as well as by human experts.

	Click Labels (%)			Human Verification (%)		
Data Set	MLRF	MLRF+ SSL	KEX	MLRF	MLRF+ SSL	KEX
Wikipedia	15.72	18.53	11.63	24.46	27.17	17.51
Ads1	18.13	19.88	11.96	45.86	47.53	41.95
Bing	22.51	25.32	18.42	50.47	51.83	47.69
Ads2	15.91	17.12	12.45	41.28	43.78	36.69

# Query Expansion Results

Query expansion techniques can help both KEX and MLRF

	Click Lab	els (%)	Human Verification (%)		
Data Set	MLRF+ SSL+KSP	KEX+KSP	MLRF+ SSL+KSP	KEX+KSP	
Wikipedia	18.01	10.81	31.48	22.14	
Ads1	21.54	12.38	51.08	43.27	
Web	26.66	19.88	53.69	48.13	
Ads2	19.24	14.35	46.77	40.07	

### Query Recommendation Results

• Edit distance [Ravi et al. WSDM 2010]

	Click Labels (%)						
Data Set	KEX	KEX+KSP	MLRF	MLRF+SSL	MLRF+SSL+ KSP		
Wikipedia	0.81	0.78	0.71	0.66	0.63		
Ads1	0.83	0.76	0.71	0.65	0.61		
Web	0.73	0.68	0.65	0.62	0.58		
Ads2	0.77	0.73	0.69	0.63	0.59		

#### Conclusions

- Query recommendation can be posed as multi-label learning.
- Learning with millions of labels can be tractable and accurate.
- Other applications
  - Query expansion.
  - Document and ad relevance and ranking.
  - Fine-grained query intent classification.

### Acknowledgements

- Deepak Bapna
- Prateek Jain
- A. Kumaran
- Mehul Parsana
- Krishna Leela Poola
- Adarsh Prasad
- Varun Singla

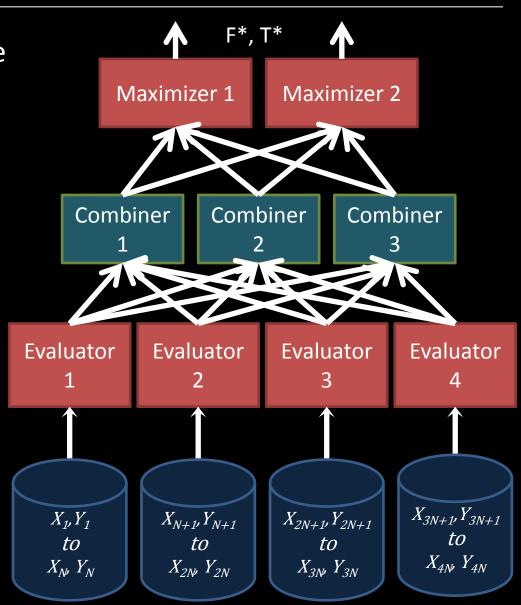
# Advantages of an ML Approach

• Can generalize to other domains such as images on Flickr or videos on YouTube.



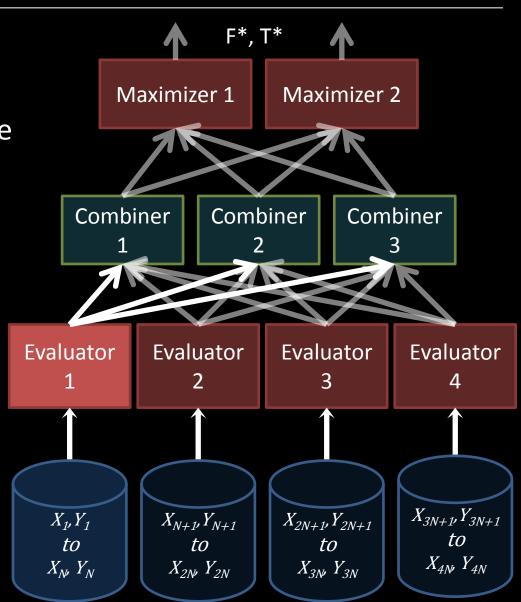
### System Architecture

- We leverage the Map/Reduce framework.
- Trees are grown in parallel breadth-wise.
- Number of compute nodes
  - Evaluators 500
  - Combiners 100
  - Maximizers 25
- Our objective is to balance the compute load across machines while minimizing data flow



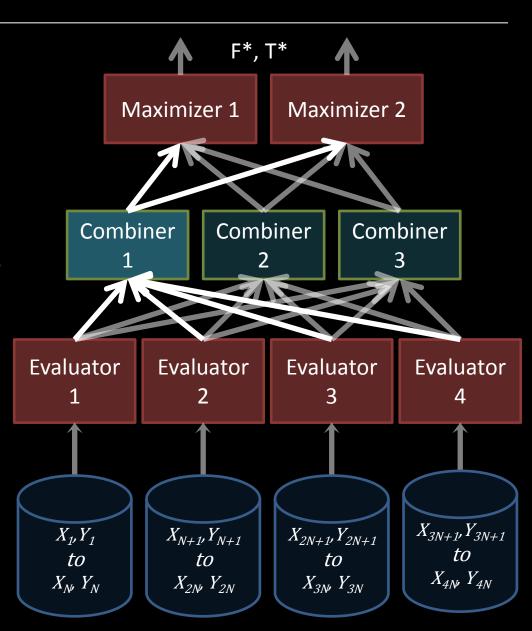
#### **Evaluators**

- Input
  - N training instances
  - Set of keys Tree ID, Node ID, Feature ID and threshold
- Output
  - Partial label distributions for the keys
- Computation
  - N \* # of keys



#### Combiners

- Input
  - Partial label distributions for assigned keys
- Output
  - Objective function values for the keys.
- Computation
  - # of keys \* Avg # of
     Evaluators / key \* # of
     labels in the distribution
     for the key.



#### Maximizers

- Input
  - Objective function values for assigned keys
- Output
  - Optimal feature and threshold for assigned nodes in trees.
- Computation
  - # of keys \* Avg # of features per key \* Avg # of thresholds per feature

