Extreme Classification

COV 878: Special Topics in Machine Learning

Manik Varma
Microsoft Research & IIT Delhi

Binary Classification

- Answer yes/no questions involving uncertainty

Is this George Washington or not?

Multi-class Classification

- Answer multiple choice questions

Which US President is present in this image?

Multi-label Classification

- Pick multiple answers in a multiple choice question

Which US Presidents are present in this image?

Traditional Classification

- Classification with a small number of choices

Outlook
Spam or not?

KINECT
< 100 gestures

Windows Hello
User or not?

Microsoft Cognitive Services
< 25K objects

msn
< 1000 topics
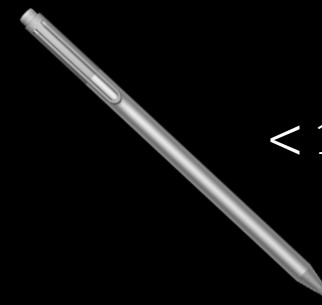
Surface Pen
< 100 characters

Windows Defender
Virus or not?

Outlook
< 1000 tags

Cortana
'Hey Cortana' or not?

Extreme Classification

- Classification with millions of labels

Ad	Predicted Bing Queries

Predicted Bing Queries:

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

MLRF: Multi-label Random Forests [Agrawal, Gupta, Prabhu, Varma WWW 2013]

Extreme Classification in Academia

- Publications at AAAI, AISTATS. ECCV, IJCAI, ICLR, ICML, KDD, NIPS, SIGIR, WSDM, WWW, *etc*.

- 8 popular workshops organized in 5 years at Dagstuhl, ECML, ICML, NIPS, WWW, *etc*.

- Code, datasets & benchmarks released on The Extreme Classification Repository
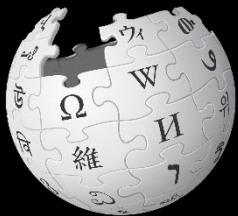
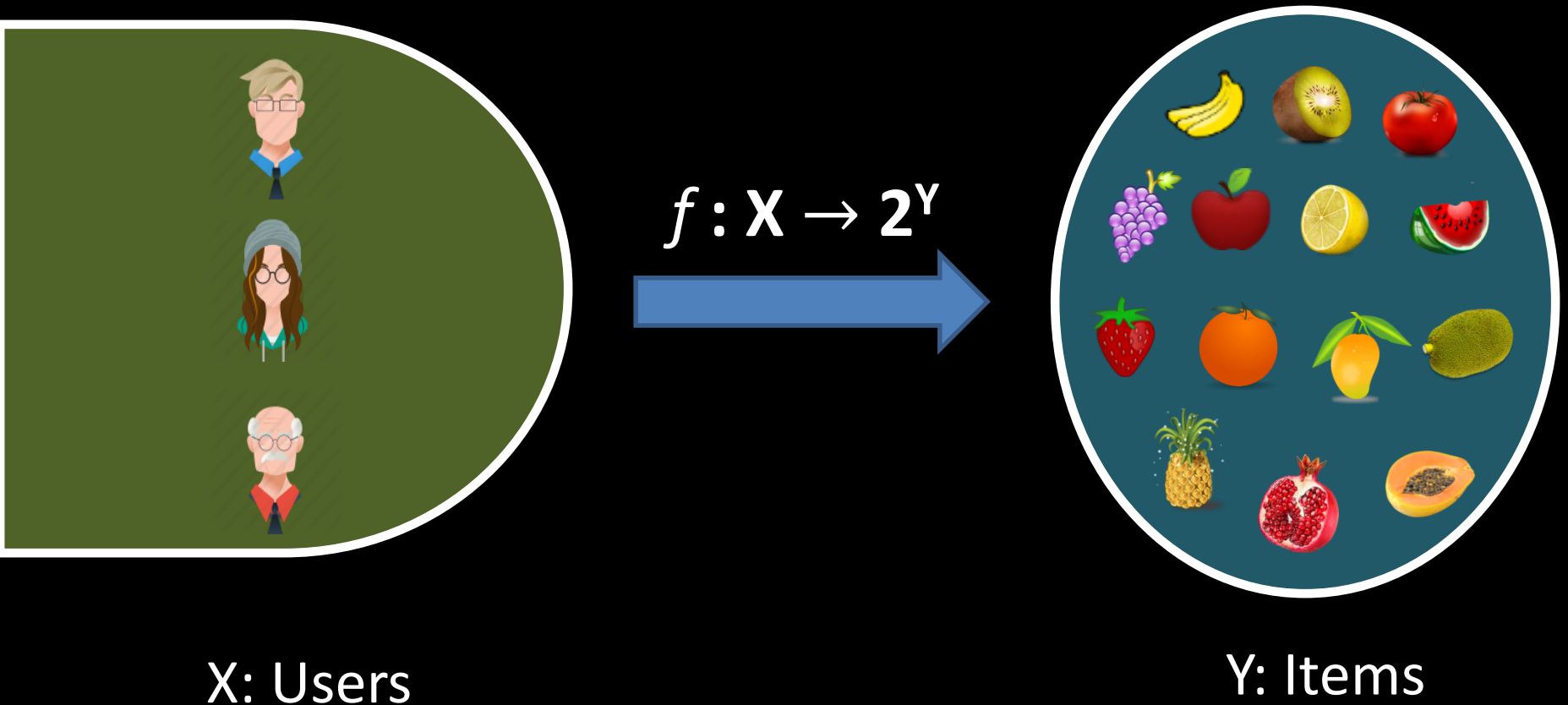- Wikipedia results have improved from 20% in 2013 to 65% in 2018

Applications

- Information Retrieval
  - Ranking for web search & advertising

- Recommender Systems
  - Item to item recommendation

- Natural Language Processing
  - Language modelling
  - Document tagging

- Computer Vision
  - Person recognition
  - Learning universal feature representations

- Bioinformatics
  - Gene function prediction

Extreme Multi-Label Classification

- Problem formulation

$$f : \mathbf{X} \rightarrow \mathbf{2^Y}$$

X: Users

Y: Items

Extreme Multi-Label Learning

- Problem formulation

f ( ) →
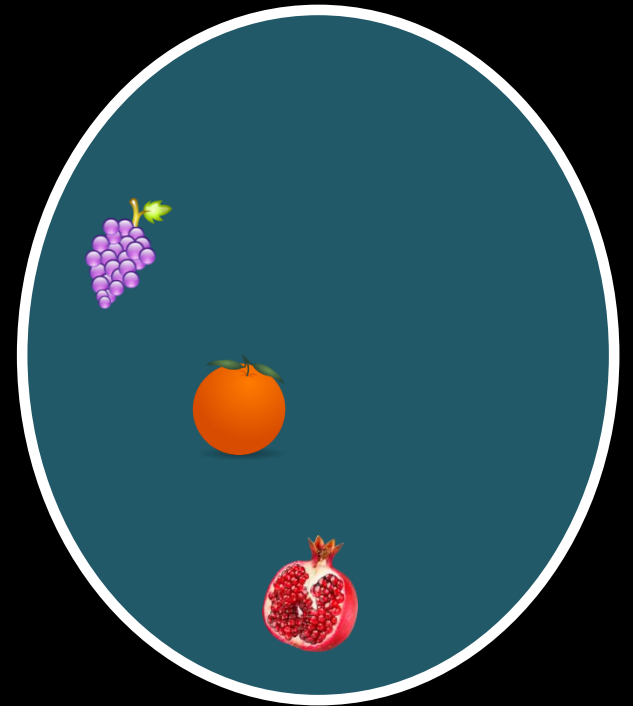
Bing Ads – Tesco's Distilled Water

Bidded Query: distilled water 5 litres

Predictions: Bing Ads vs Extreme Classification

Bing Ads

water 5

Extreme Classification

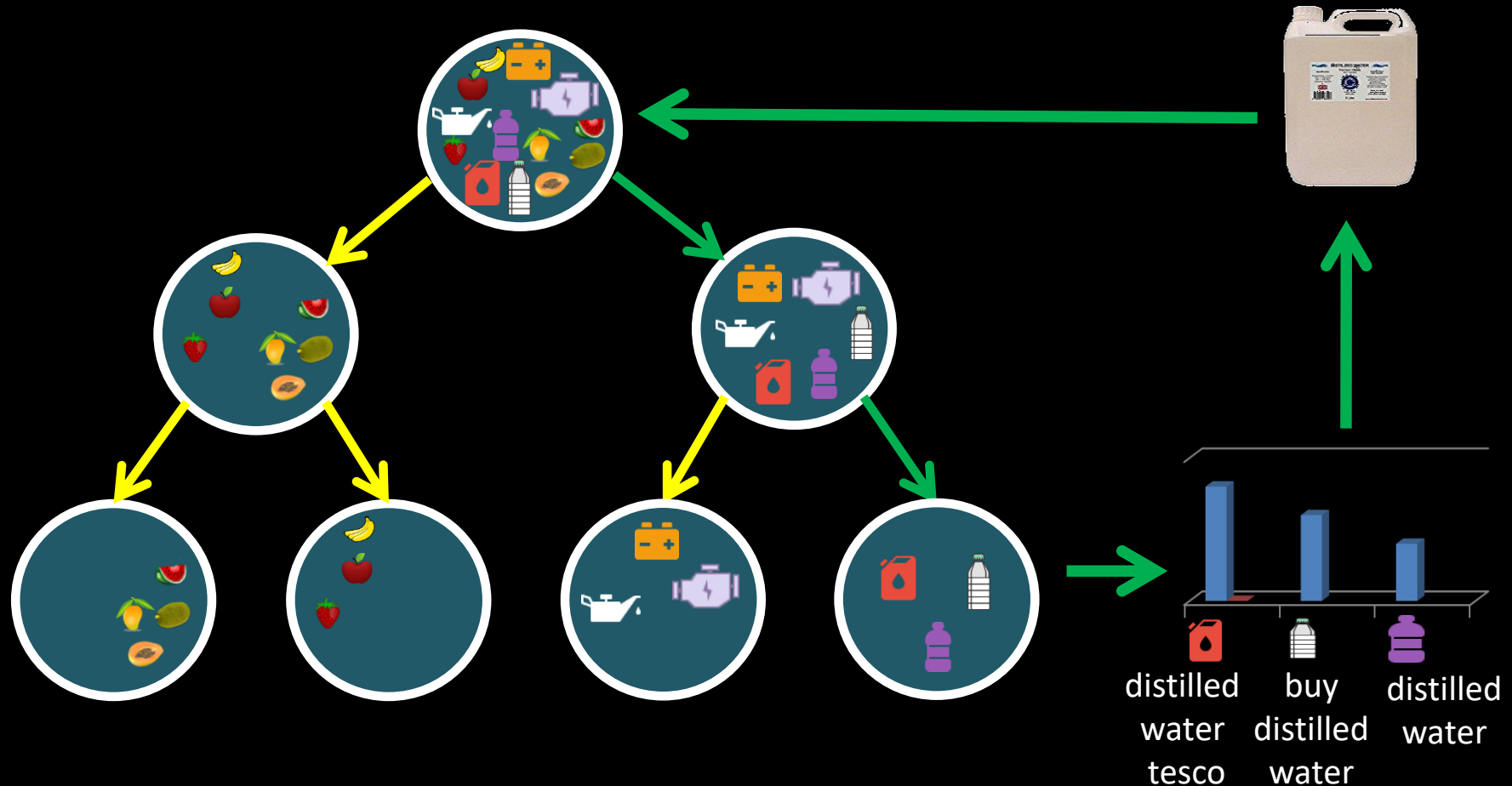| distilled water tesco |
|---|
| where buy distilled water |
| distilled water |
| buy distilled water |
| distilled water amazon |
| distilled water vs purified water |
| distilled water uk |
| distilled water delivery |
| where can I buy distilled water |
| distilled water uk supermarket |

Traditional Approach

- Reduction to binary classification

$$h : \text{(Ad, Phrase)} \rightarrow \{ \textcolor{red}{\times}, \textcolor{green}{\checkmark} \}$$

$$h(\,\text{[distilled water]}, \ \textbf{buy distilled water}\,) \rightarrow \textcolor{red}{\times}$$

$$h(\,\text{[distilled water]}, \ \textbf{water 5}\,) \rightarrow \textcolor{green}{\checkmark}$$

Extreme Classification Approach

- Efficient & accurate prediction via a learnt hierarchy

Parabel: Partitioned Label Trees [Prabhu, Kag, Harsola, Agrawal, Varma WWW 2018

Extreme Classification for Bing Ads

- Shipped in various products in international markets

UK Dynamic
Search Ads

French
Text Ads

German
Product Ads

Item-to-item Recommendation - Walmart

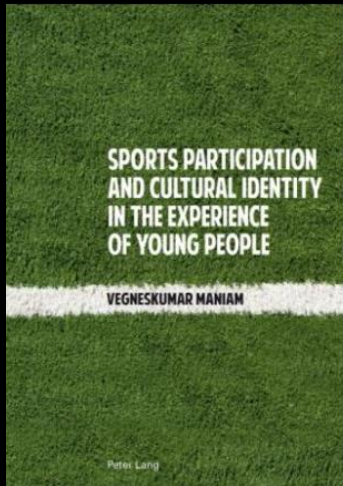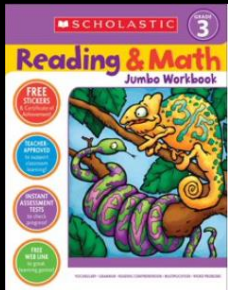Sports Participation and Cultural Identity in the Experience of Young People

★★★★★ Write a review

$48.97

Only 1 left!
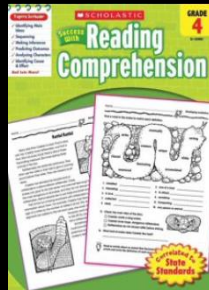
Free shipping

Arrives by Saturday, Aug 25 Options
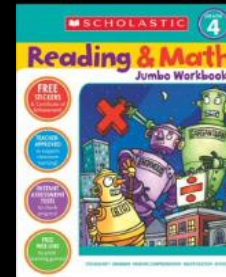
Qty:

Consider these popular products

Reading & Math Jumbo Workbook: Grade 3

Scholastic Success with Reading Comprehension, Grade 4
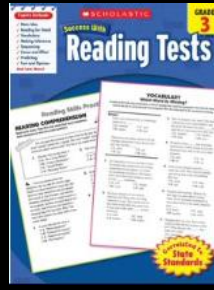
Reading Tests, Grade 3
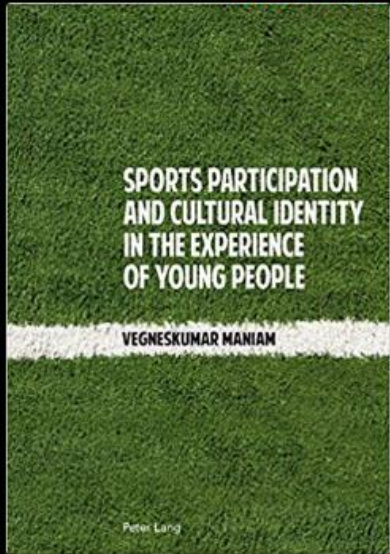
Reading & Math Jumbo Workbook: Grade 4

Return policy

...stry

Item-to-item Recommendation - Amazon

amazon Try Prime

Look inside ↓

Sports Participation and Cultural Identity in the Experience of Young People

by Vegneskumar Maniam (Author)

Be the first to review this item

See all 2 formats and editions

Kindle	Paperback
$49.27	$59.49

Read with Our **Free App**

9 Used from $25.58
13 New from $47.63

This book focuses on inclusion and exclusion in sporting activities among young people in a multicultural society. Do young people who identiy with cultural groups other than the majority experience exclusion from sporting teams, or do they find themselves readily included? Does this vary across identities, and sports? In the context of Australia, where sport is an integral part of national life and one in four of the population were born overseas, and over 270 different ancestries are acknowledged, young people were asked to write about their cultural identity and their experiences playing sport. Using a humanistic sociological approach, the inductive analysis justaposed their sense of cultural identity with their participation or non-participation in sport, and with the particular sports played. This book is important for ... ty especially acad...and inclusion of ... tion.

...view
...k reviews, editors

Sponsored products related to this item

Items related to this item

Amazon vs Walmart vs Extreme Classification

Amazon

Walmart

Reading & Math Jumbo Workbook: Grade 3
Scholastic Success with Reading Comprehension, Grade 4
Reading Tests, Grade 3
Reading & Math Jumbo Workbook: Grade 4
Scholastic Success with Multiplication Facts, Grades 3-4
Cursive Writing Practice: Inspiring Quotes
Modern Handwriting: Beginning Cursive, Grades 1 - 3
Math, Grade 3
Cursive Writing Practice: Jokes & Riddles, Grades 2-5
10 Week-By-Week Sight Word Packets

Extreme Classification

Sports in Society: Issues and Controversies
Beer and Circus: How Big-Time College Sports Is Crippling Undergraduate Education
Sport in Contemporary Society : An Anthology
Big-Time Sports in American Universities
Power at Play: Sports and the Problem of Masculinity (Men and Masculinity)
Successful Coaching
Foundations of Sport and Exercise Psychology With Web Study Guide-5th Edition
Friday Night Lights: A Town, A Team, And A Dream
Out of Play: Critical Essays on Gender and Sport

Collaborative Filtering

- Extreme classification can increase the recommendation accuracy from 6% to 36%
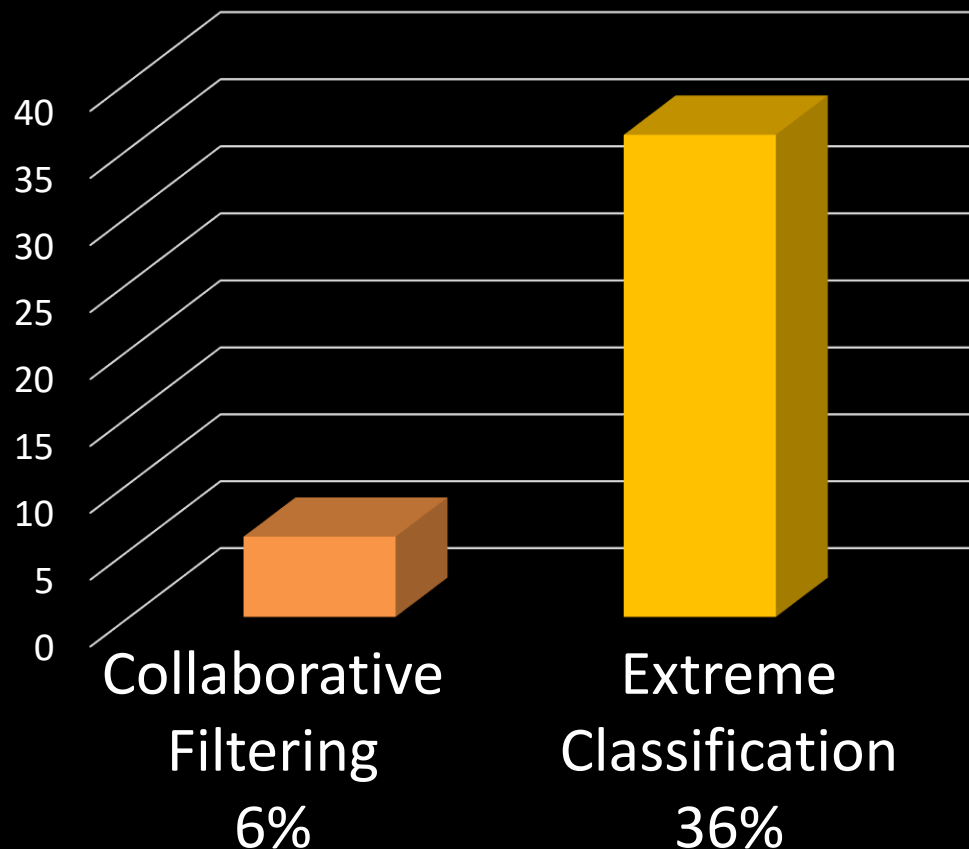
Traditional Approach

- Collaborative filtering & matrix factorization

Ratings Matrix = User Traits X Item Attributes
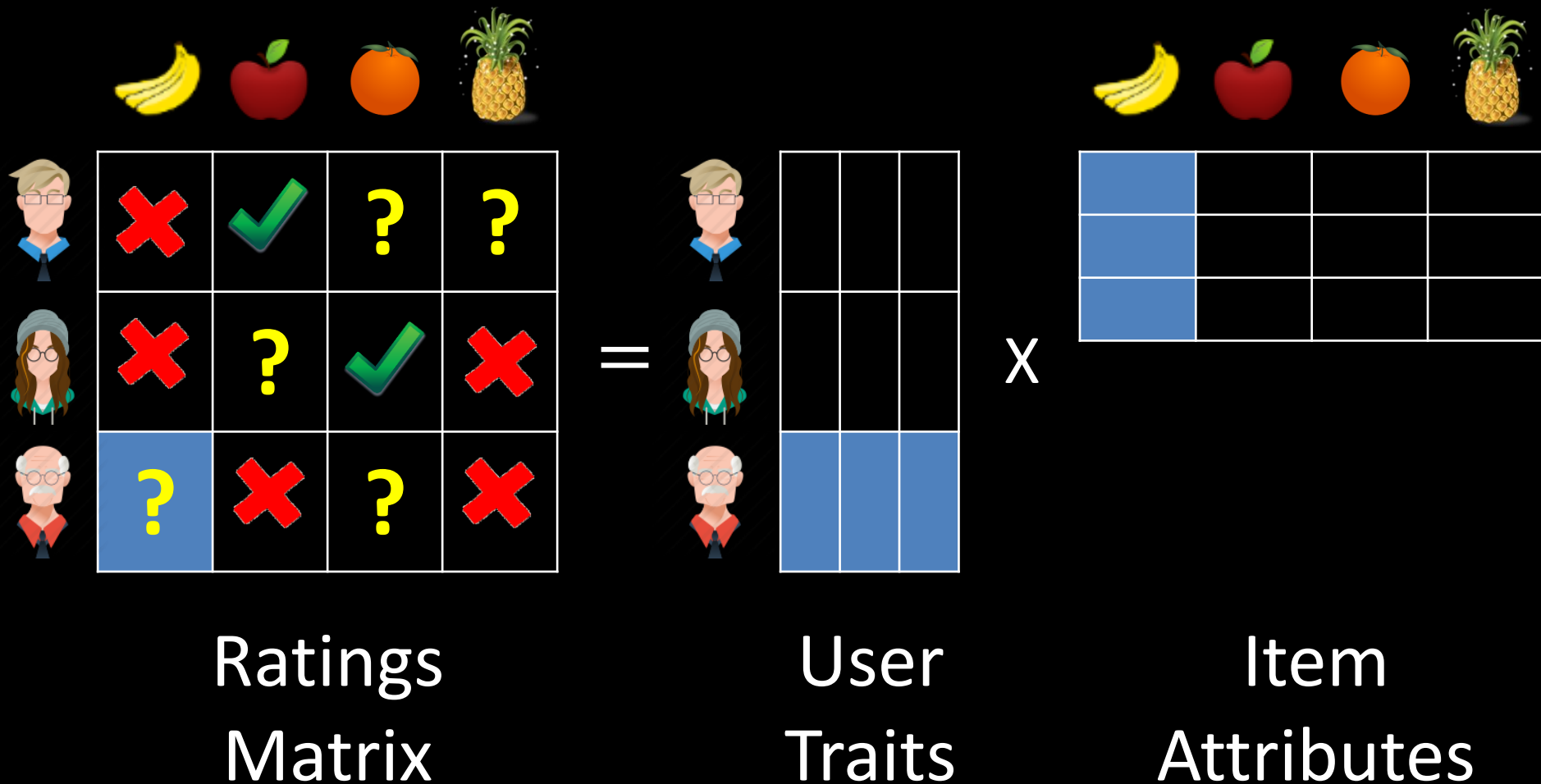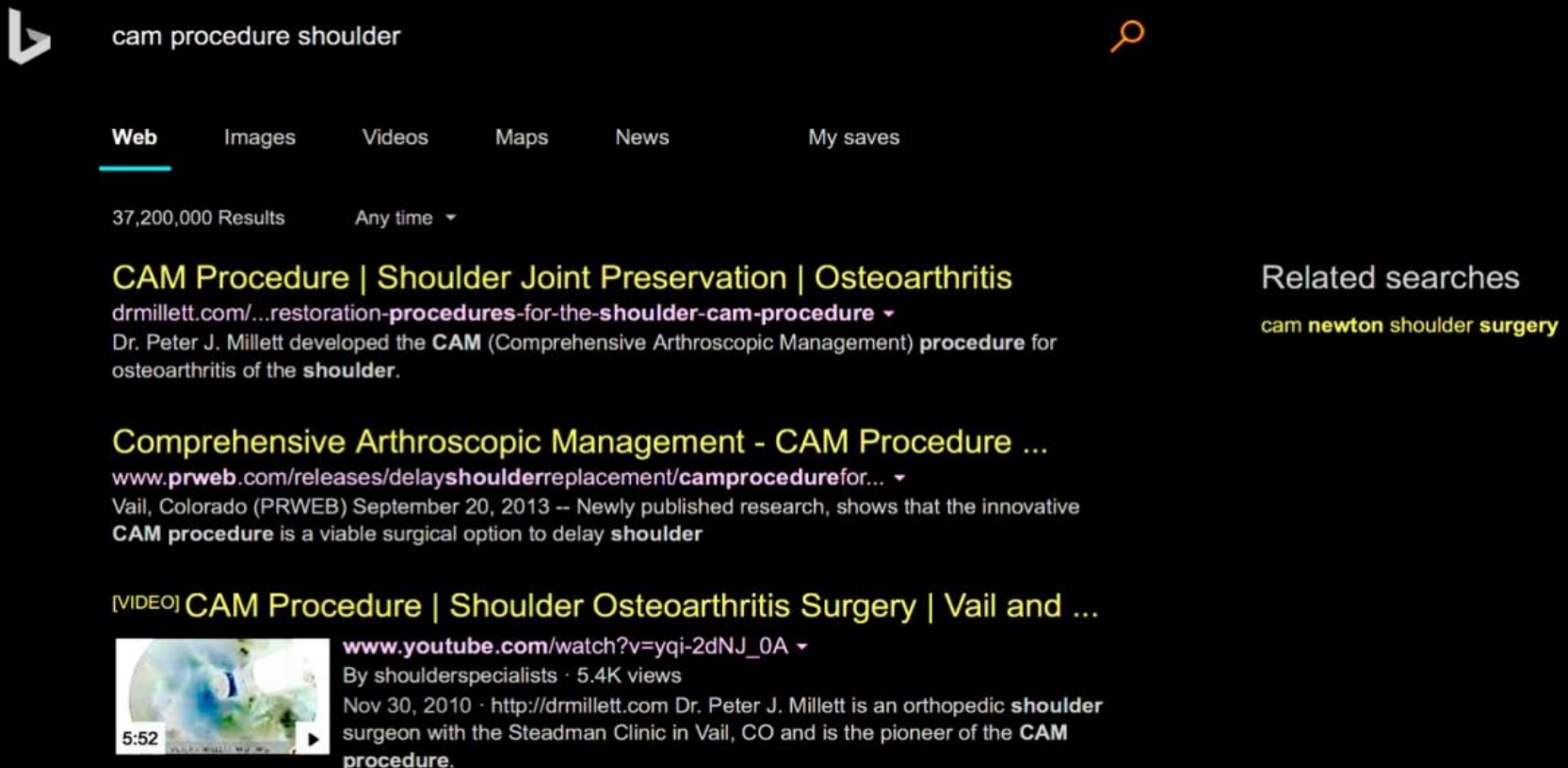
Extreme Classification Approach

- Recommendation based on user and item features

SwiftXML [Prabhu, Kag, Gopinath, Harsola, Agrawal, Varma WSDM 18]
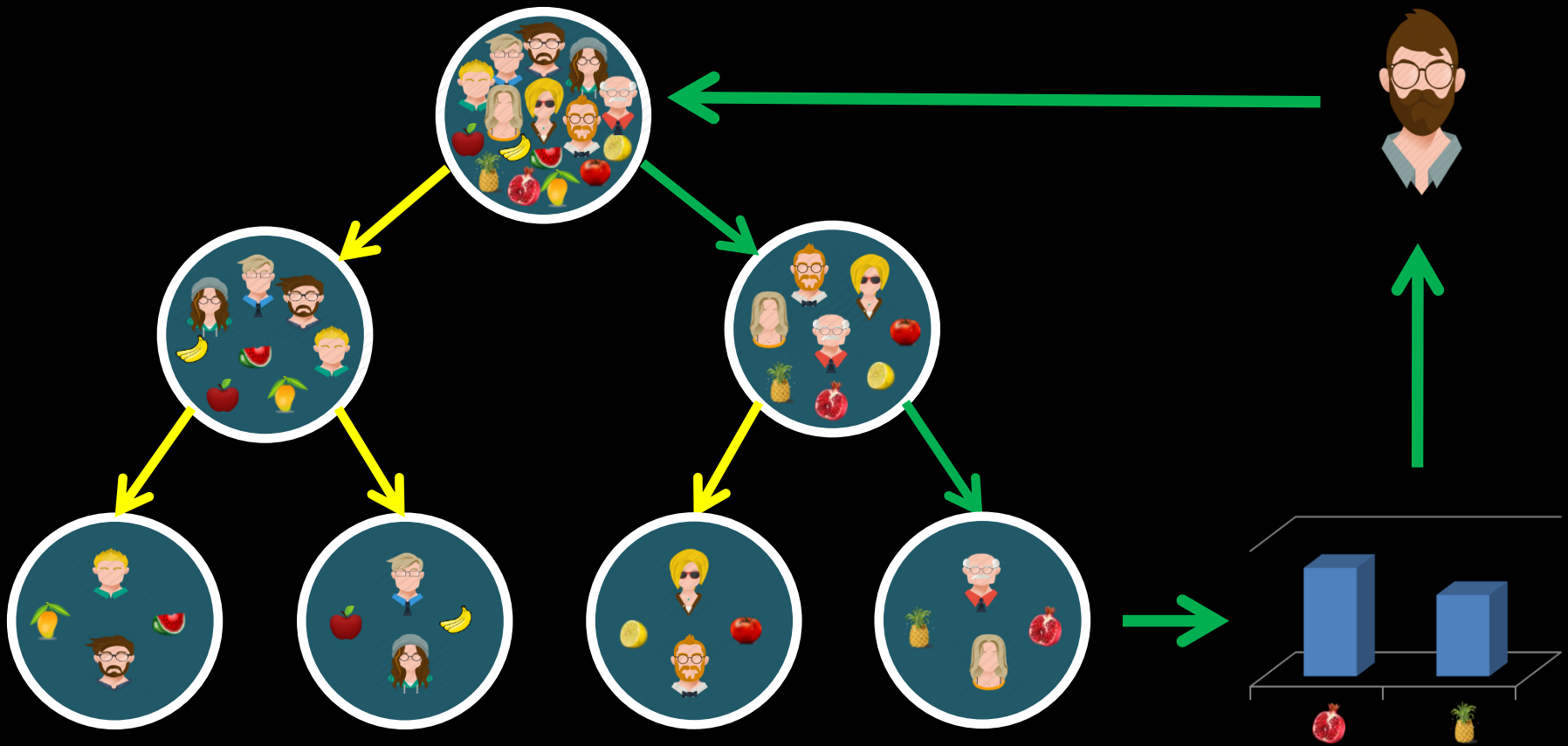
Bing RS – "cam procedure shoulder"

- Recommend related queries that might
  - Serve the user's information requirements better
  - Provide more information on the topic

Predictions: Bing vs Extreme Classification

Bing	Extreme Classification
cam newton shoulder surgery	how long off work for shoulder surgery
	shoulder surgery procedures
	recovery from arthroscopic shoulder surgery
	shoulder joint resurfacing surgery
	shoulder clean up surgery
	tenex procedure for rotator cuff
	cost of arthroscopic shoulder surgery
	shoulder replacement surgery success rate
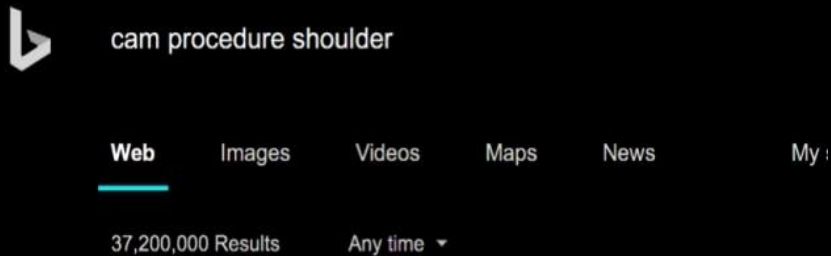
Sessions Based Approaches

- Might not work well for tail queries
- Intent changes might lead to poor suggestions

Query-URL Based Approaches

- Might not work well for tail queries
- Might lead to content drift

cam procedure shoulder

cam newton shoulder surgery

How long off work for shoulder surgery

https://drmillett.com/

https://webmd.com/
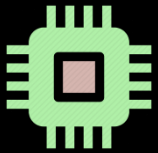
https://melbournearm
clinic.com/
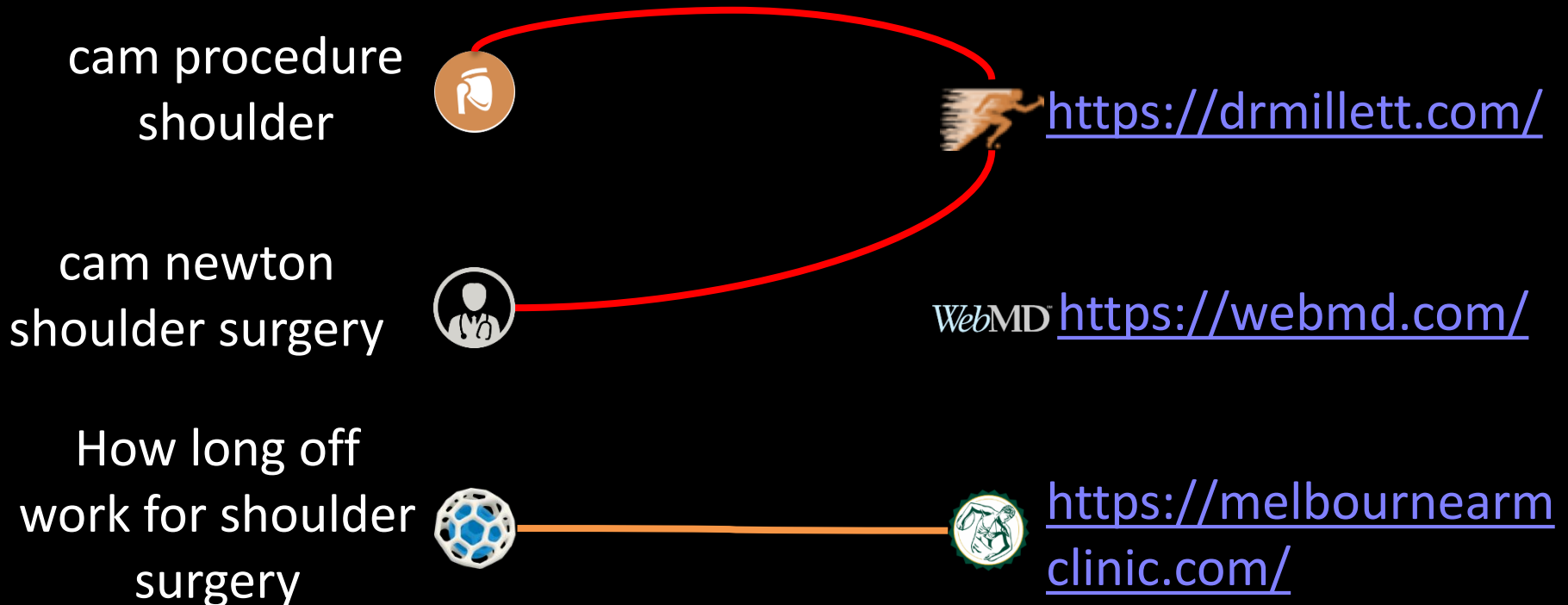
Slice [WSDM 2019]

- Specialized for low-dimensional dense unit-norm features

- Scales to 100 M labels and 240 M training points
  - Leverages label sparsity
  - Log time training based on negative sampling
  - Log time prediction using approximate NN search

- Improvements in Bing Related Searches

Trigger Coverage	Suggestion Density	Success Rate	Tail Success Rate
52.01%	33.0%	2.62%	12.62%

Tagging Wikipedia Articles

Predictions: Wiki vs Extreme Classification

Wikipedia

Works by Dante Alighieri
Divine Comedy
1321 books
1300 in Italy
Visionary poems
Epic poems in Italian
14th-century Christian texts
14th-century books
Virgil
Afterlife

Extreme Classification

Works by Dante Alighieri
Divine Comedy
1321 books
1300 in Italy
Visionary poems
Epic poems in Italian
14th-century Christian texts
14th-century books
Virgil
Dante Alighieri

Recognizing People on Facebook

Choices: Bradley Cooper, Ellen DeGeneres, Meryl Streep, Jennifer Lawrence, Channing Tatum, Julia Roberts, Kevin Spacey, Brad Pitt, Angelina Jolie, Lupita Nyong'o, Peter Nyong'o

Language Modelling

Brevity is the soul of ...

Wit   Twit   Lingerie

Conclusions

- Extreme classification
  - Tackle applications with millions of choices
  - A new paradigm for ranking & recommendation

- Algorithms & papers
  - MLRF [WWW 2013], FastXML [KDD 2014]
  - SLEEC [NIPS 2015], PfastreXML[KDD 2016]
  - SwiftXML [WSDM 2018], Parabel [WWW 2018]
  - Slice [WSDM 2019]

- The Extreme Classification Repository
  - Code & datasets
  - Benchmark results
  - Papers

Research Questions

- Applications
- Obtaining good quality training data
- Log time and space training and prediction
- Obtaining discriminative features at scale
- Extreme loss functions
- Performance evaluation
- Dealing with tail labels and label correlations
- Dealing with missing and noisy labels
- Explore/exploit for tail labels
- Statistical guarantees
- Fine-grained classification