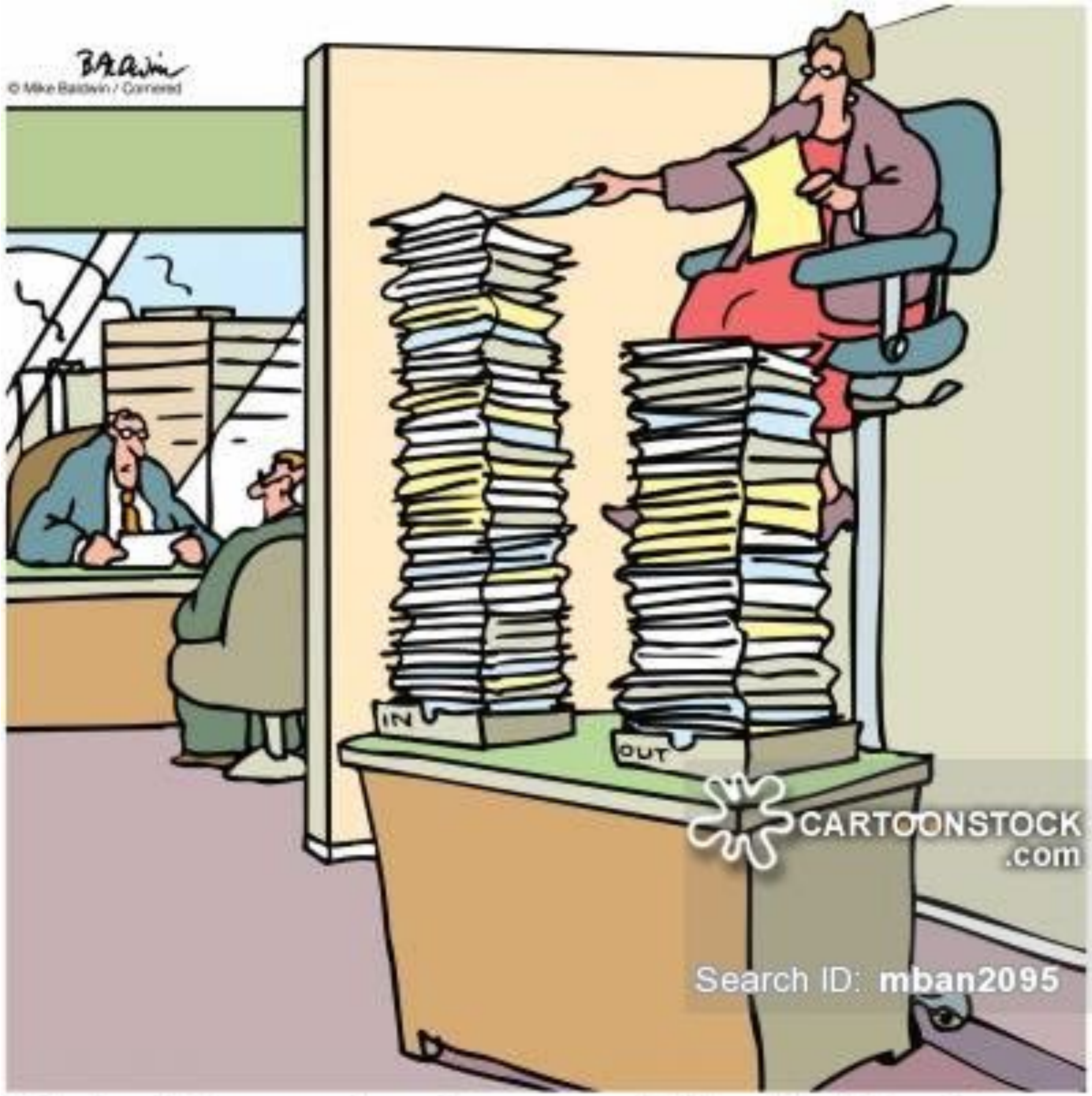
Recent Developments in Online Learning for Big Data Applications

Irwin King

Associate Dean (Education), Faculty of Engineering, CUHK
Director, Shenzhen Rich Media and Big Data Analytics and Application Key Lab, SZRI, CUHK
PI, The Knowledge and Education Exchange Platform (KEEP), CUHK

Department of Computer Science and Engineering
The Chinese University of Hong Kong

king@cse.cuhk.edu.hk
http://www.cse.cuhk.edu.hk/~king



"We look for people who can quickly adapt to changes in the workplace."

source: http://www.cartoonstock.com/directory/w/workload.asp

The Secretary Problem

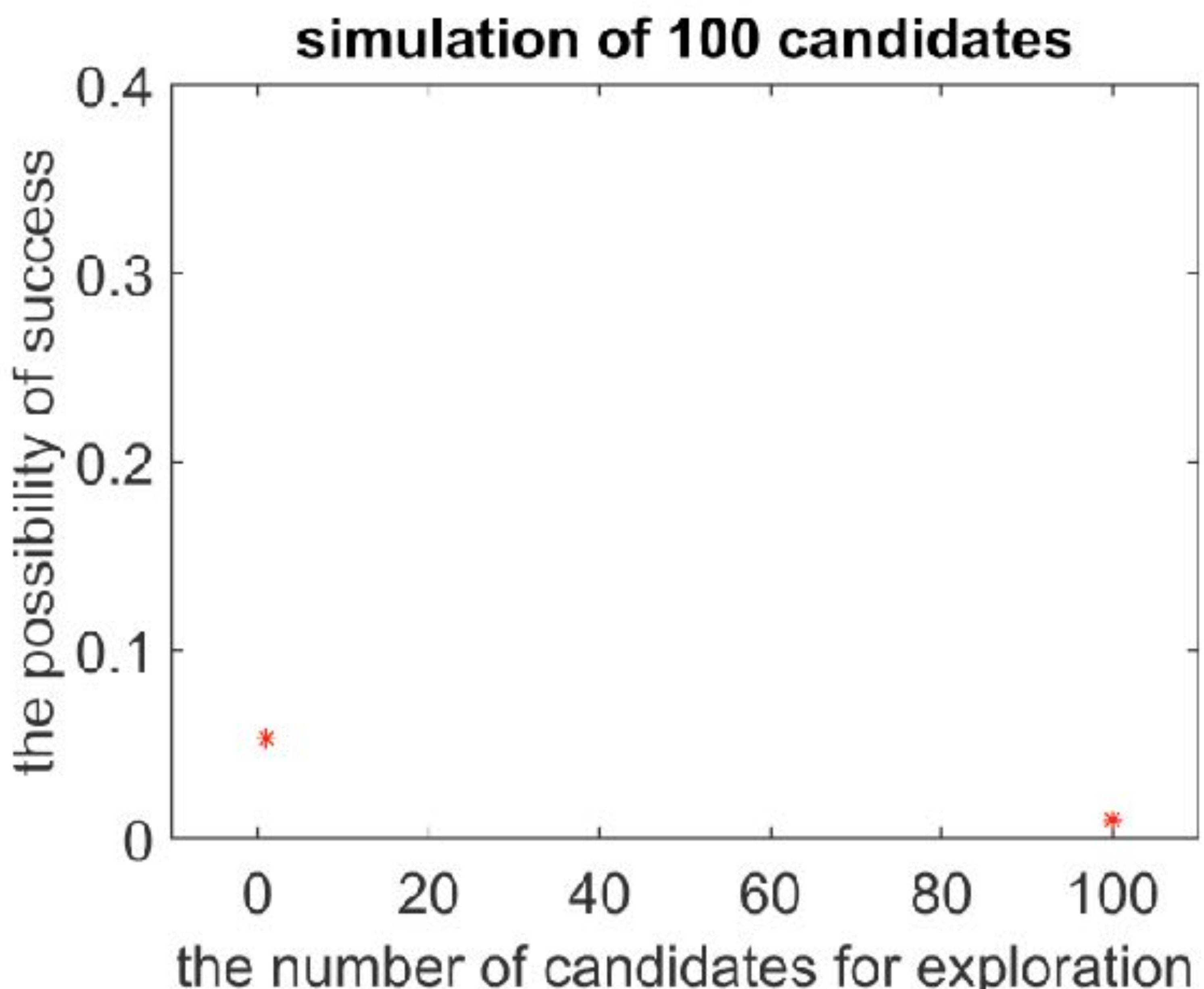
- Marriage problem, Sultan's dowry problem, googol game, optimal stopping problem, etc.
- Class of Sequential Decision Problems

- How to choose the best candidate?
- Assumptions
  - Interview candidates one by one
  - Make a decision to hire or not immediately after the interview
  - Cannot go back and hire another candidate
  - Know the total number of candidates to be interviewed

source: http://quoratopstories.tumblr.com/post/108021448078/what-are-the-most-interesting-or-popular

The Secretary Problem

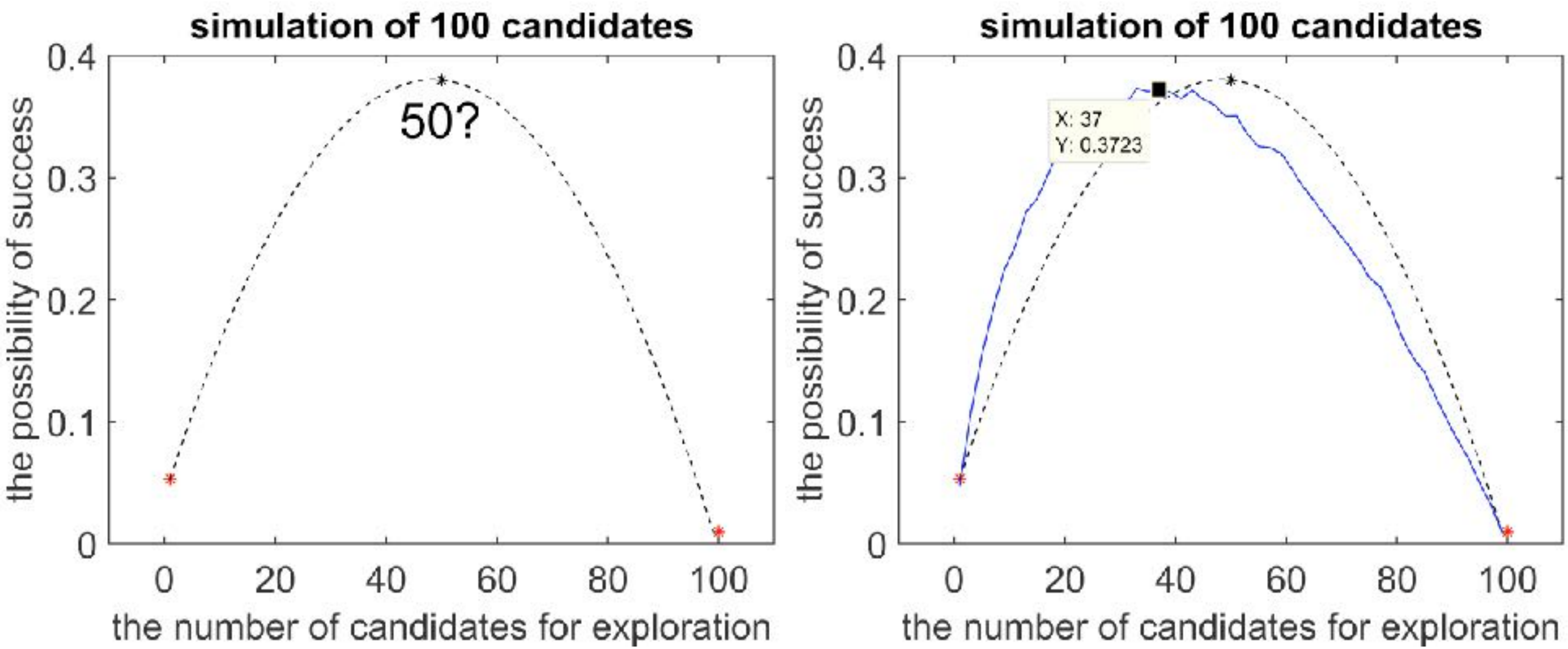
- Naive solutions
  - Interview the first candidate and set the benchmark
  - Or, interview *N*-1candidates and choose the last one



Experiment steps:

1. Permute the sequence of [1,2,…,100]
2. Set the number of candidates for exploration
3. Set the benchmark and choose the best candidate
4. Run 10,000 times repeatedly

The Secretary Problem



The Secretary Problem

- Optimal strategy [T. S. Ferguson, 1989]
  - Reject the first $N/e \approx 0.37N$ candidates categorically
  - Accept the first one above the top category after N/e
  - The highest probability is $1/e$









Multi-Armed Bandits (MAB) Problem

- Pull an arm to get a payoff of the arm for each round
- Assume each round costing one dollar, and a total budget of *N* dollars

K arms

Exploration vs. Exploitation

Multi-Armed Bandits (MAB) Problem

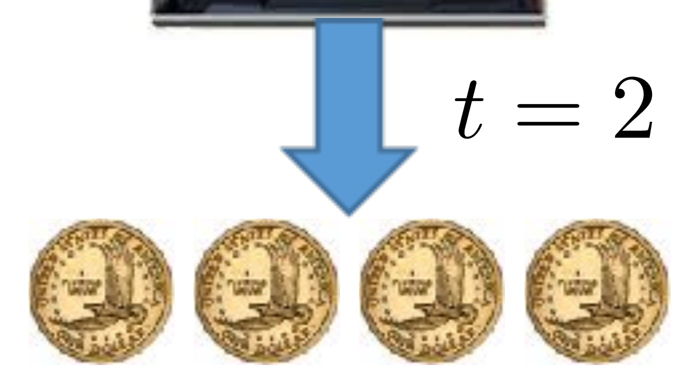
- Pull an arm to get a payoff of the arm for each round
- Assume each round costing one dollar, and a total budget of *N* dollars

Multi-Armed Bandits (MAB) Problem

- Pull an arm to get a payoff of the arm for each round
- Assume each round costing one dollar, and a total budget of *N* dollars

K arms

$t = 2$



Multi-Armed Bandits (MAB) Problem

- Mean value (MV)

	Arm 1	Arm 2	Arm 3	Arm 4	
Step 1	1/1=1	1/1=1	1/1=1	0/1=0	Play all the arms once
Step 2	(1+0)/2=0.5	1	1	0	Break ties randomly
Step 3	0.5	(1+2)/2=1.5	1	0	Play the best arm
Step 4	0.5	(3+1)/3=1.3	1	0	Play the best arm
Step 5	0.5	(4+0)/4=1	1	0	Break ties randomly
Step 6	0.5	1	(1+3)/2=2	0	Play the best arm
Step 7	0.5	1	(4+2)/3=2	0	Play the best arm
Step 8	0.5	1	(4+2)/4=1.5	0	Play the best arm

Find the best arm 3 via mean value, but never explore arm 4

Multi-Armed Bandits (MAB) Problem

- Mean value (MV) + standard deviation (SD)

	Arm 1	Arm 2	Arm 3	Arm 4
Step 1	1/1+1=2	1/1+1=2	1/1+1=2	0/1+1=1
Step 2	(1+0)/2+0.8=1.3	2	2	1
Step 3	1.3	(1+1)/2+0.8=1.8	2	1
Step 4	1.3	1.8	(1+0)/2+0.8=1.3	1
Step 5	1.3	(2+0)/3+0.6=1.27	1.3	1
Step 6	(1+0)/3+0.6=0.93	1.27	1.3	1
Step 7	0.93	(2+0)/4+0.4=0.9	1.3	1
Step 8	0.93	0.9	(1+0)/3+0.6=0.93	1
Step 9	0.93	0.9	0.93	(1+1)/2+0.8=1.8

Assumption of SD:1->0.8->0.6->0.4->0.2

Multi-Armed Bandits (MAB) Problem

- Optimal strategy [T. L. Lai & H. Robbins, 1985]
  - Play each arm once for the first *K* rounds
  - Play the explored best arm with upper confidence bound
- Result
  - Mean + upper confidence bound (UCB)

$$UCB_i = \sqrt{\frac{2 \log N}{n_i}}, \sum_i n_i = N$$

  - *N* is the number of times for selecting an arm

Multi-Armed Bandits (MAB) Problem

- ϵ -greedy strategy [N. Cesa-Bianchi & P. Fischer, 1998]
  - With probability of $1 - \epsilon_t$ to play the explored best arm
  - With probability of ϵ_t to randomly select inferior arms

Some Variants: Finding *k* Arms

- Find the top-k arms
- Find top arms in disjoint groups of arms

K arms

$$t = 1, 2, \cdots$$

Some Variants: Unknown *N*

- [J. Langford & T. Zhang, 2008]
- No knowledge of a time horizon to maximize reward

K arms

$$t = 1, 2, \cdots$$

Some Variants: Infinite Arms

- [Y. Wang, J. Y. Audibert & R. Munos, 2009]
- Online advertising tasks with infinite advertisements

●●●

Some Variants: Adversarial Bandits

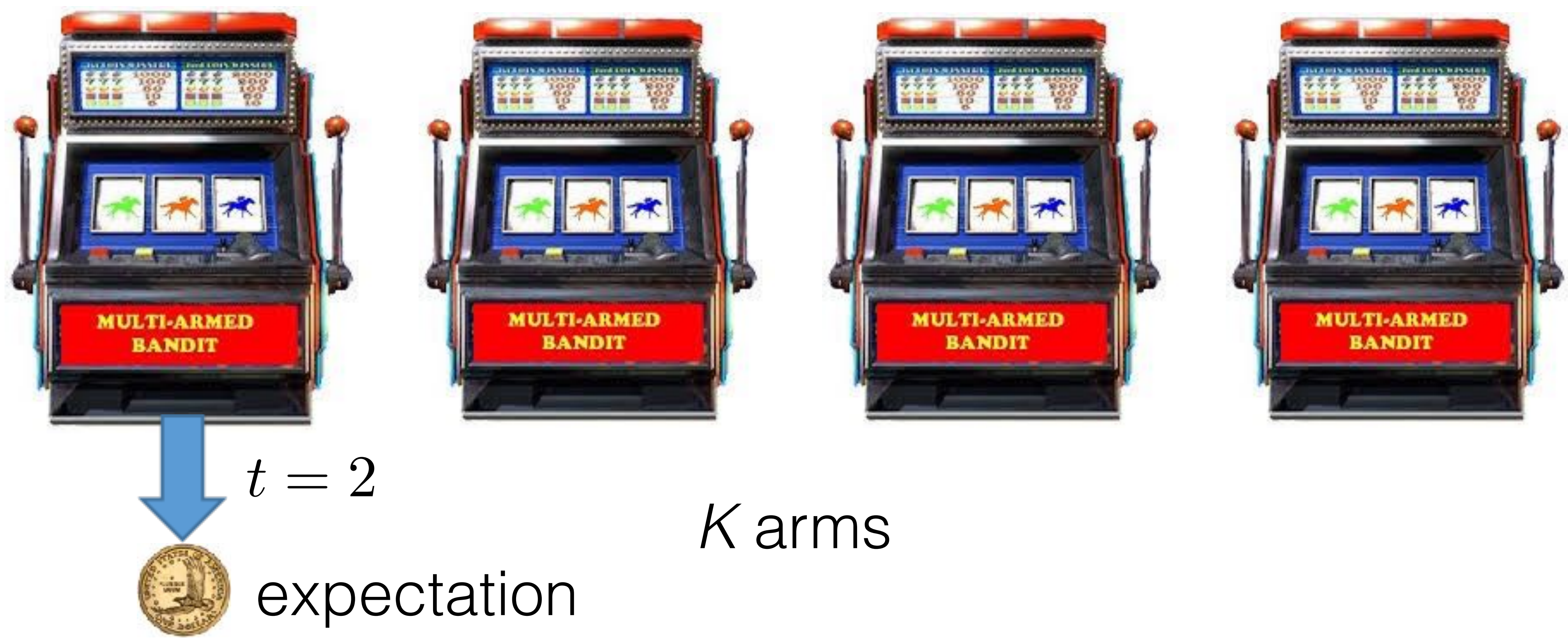
- [O. Besbes, Y. Gur & A. Zeevi, 2014]
- Example: time-varying expected payoff for bandits expectation as in online investments in financial markets

K arms

expectation

Some Variants: Adversarial Bandits

- [O. Besbes, Y. Gur & A. Zeevi, 2014]
- Example: time-varying expected payoff for bandits expectation



Some Variants: Contextual Bandit

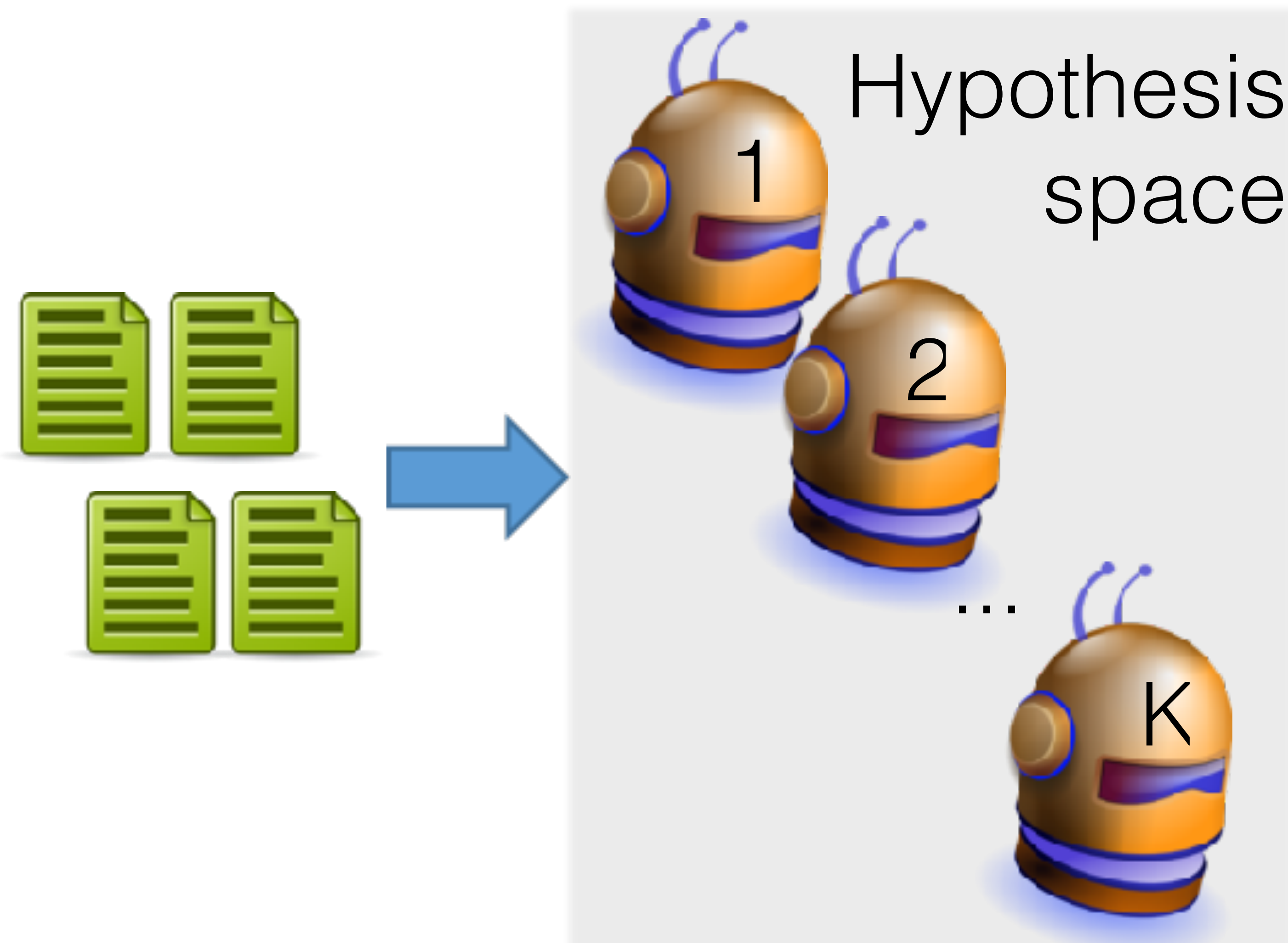
- [Li et al. 2010]
- Additional contextual information in online advertising and online recommendations

Feature vectors

$$t = 1, 2, \cdots$$

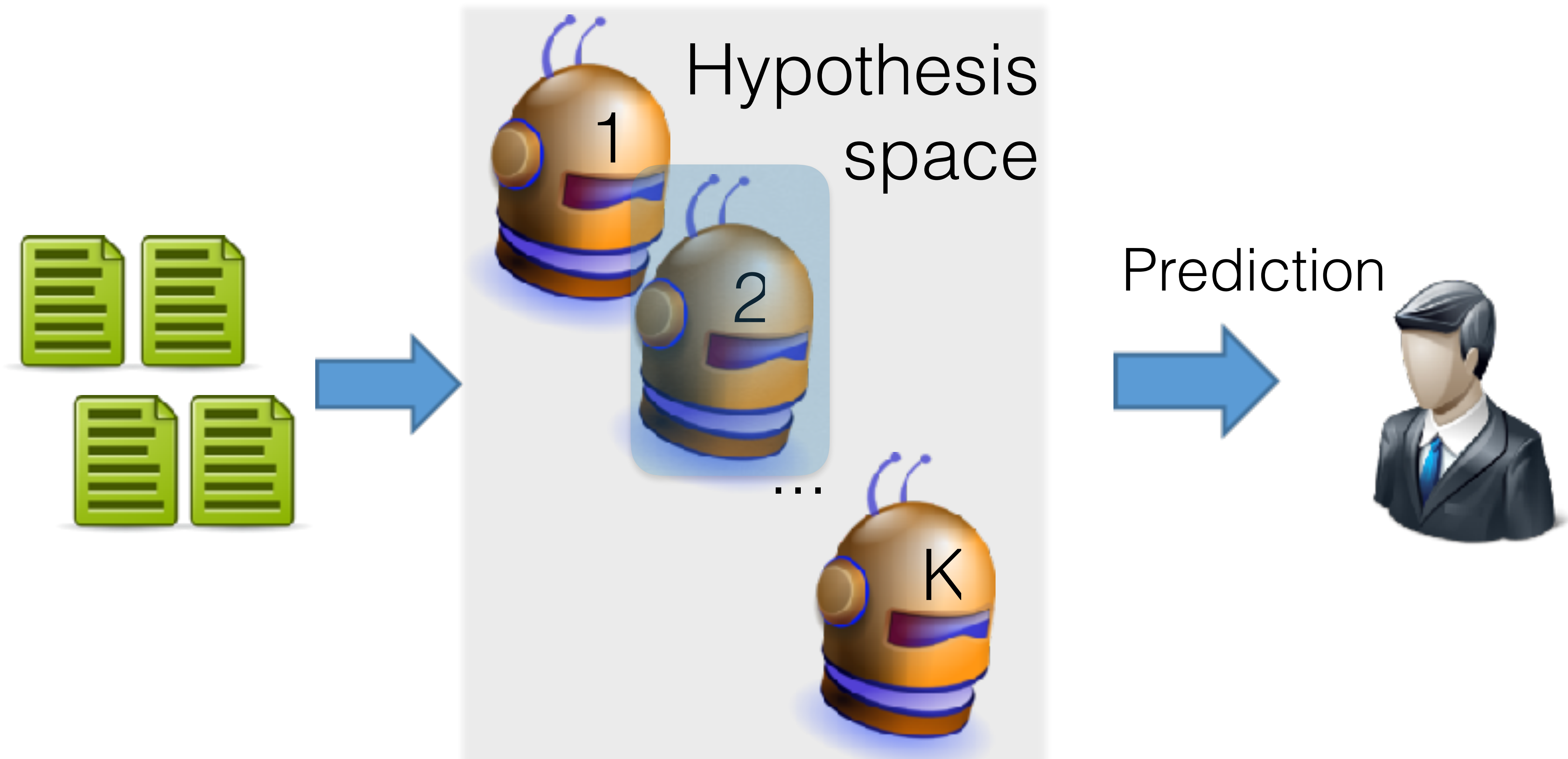
Online Learning

- How to choose the best hypothesis for data?



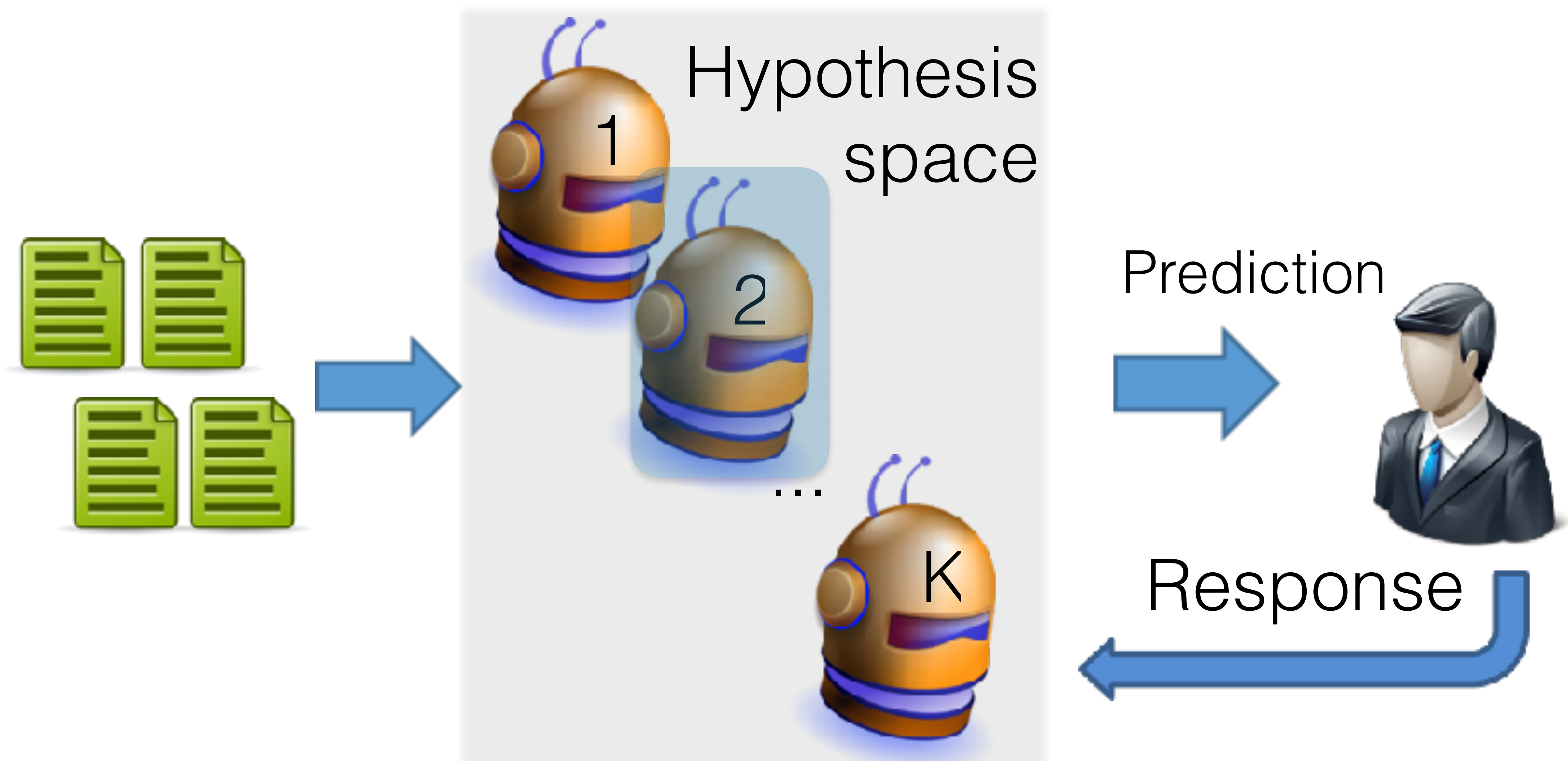
Online Learning

- How to choose the best hypothesis for data?



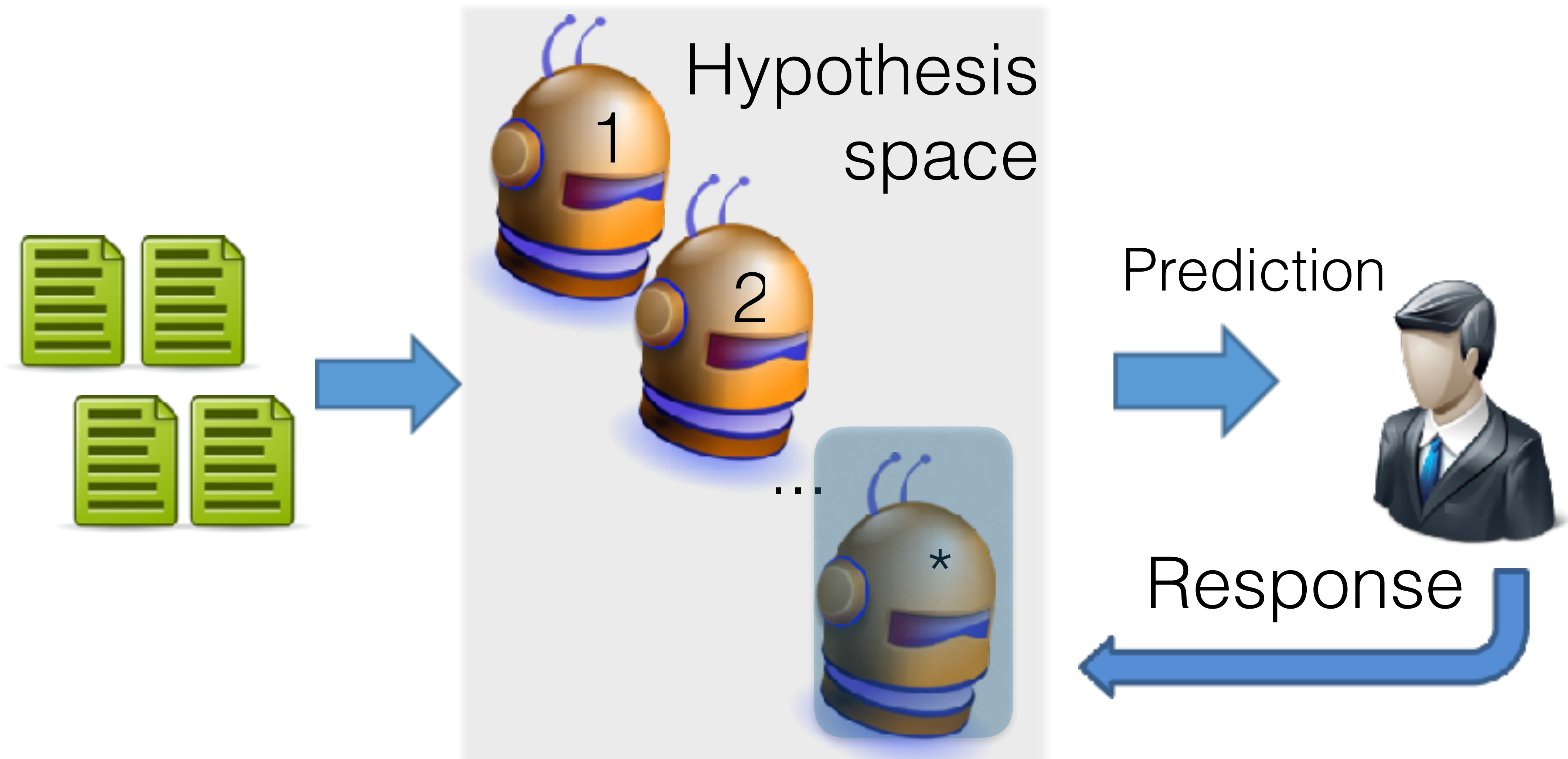
Online Learning

- How to choose the best hypothesis for data?



Online Learning

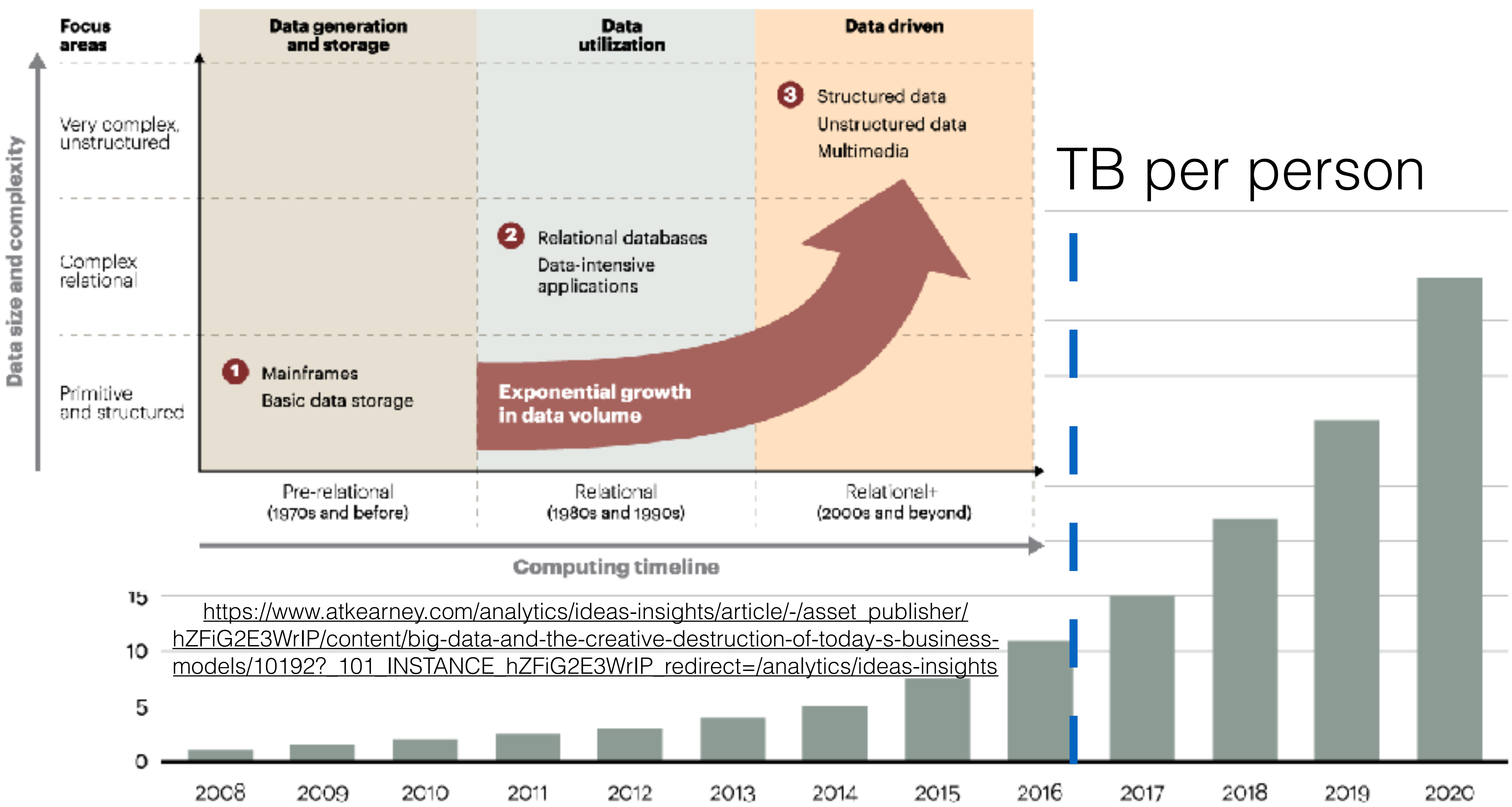
- How to choose the best hypothesis for data?



Minimize the loss between response and prediction!

Why Online Learning

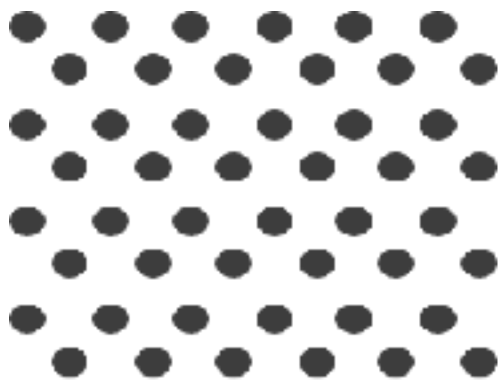
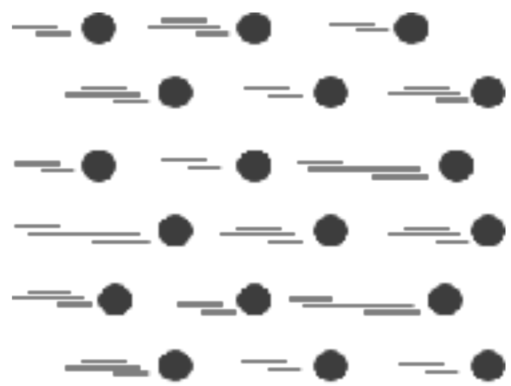
Why Online Learning

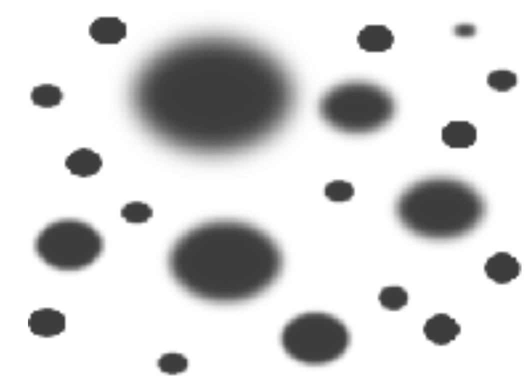


source: https://www.atkearney.com/analytics/ideas-insights/article/-/asset_publisher/hZFiG2E3WrIP/content/big-data-and-the-creative-destruction-of-today-s-business-models/10192?_101_INSTANCE_hZFiG2E3WrIP_redirect=%2Fanalytics%2Fideas-insights

Why Online Learning

Volume	Velocity	Variety	Veracity



IDC Analyze the Future

40 ZB (2020)

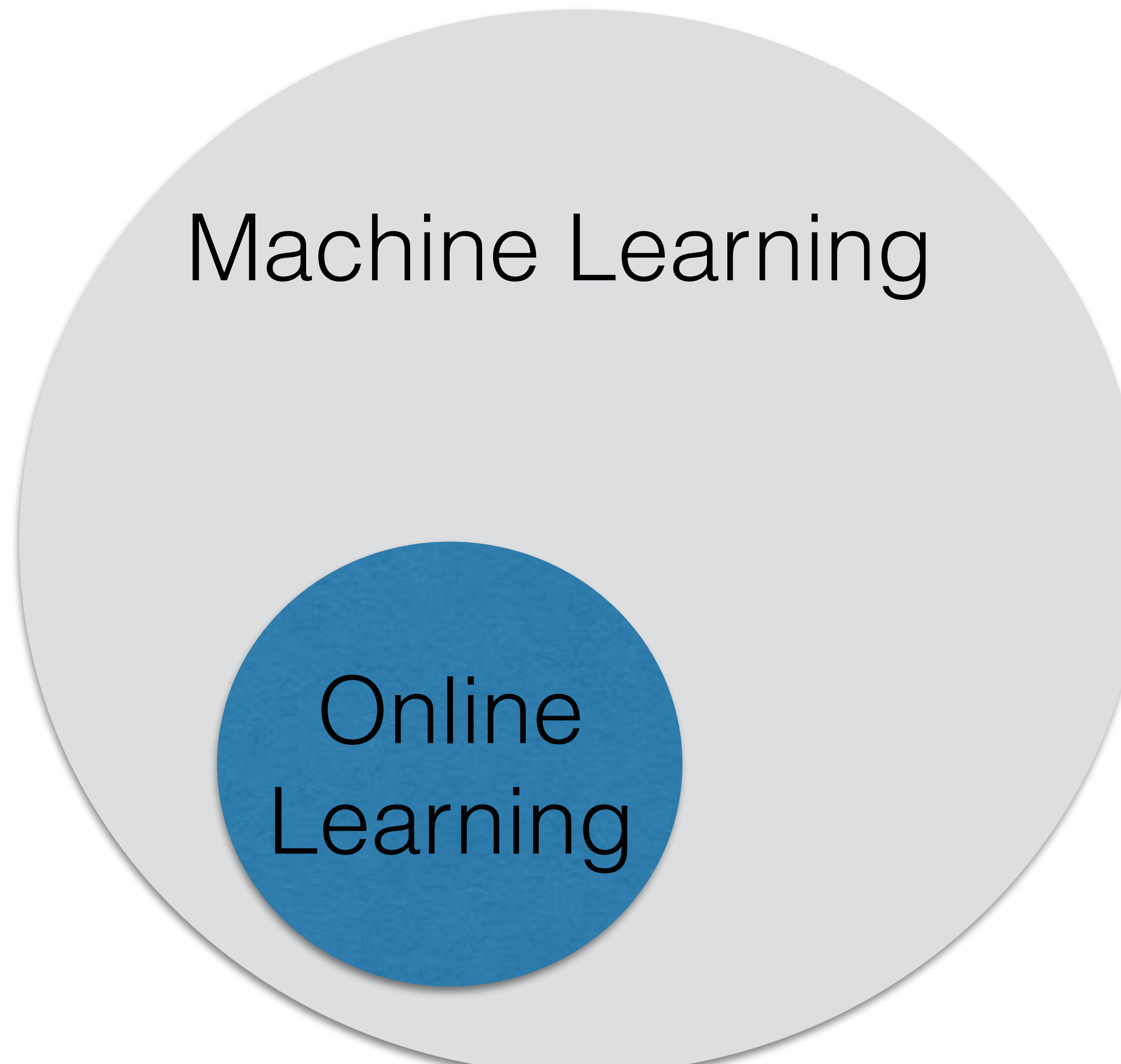
5.2 TB per person

facebook

500 TB per day new data

Online Learning for Big Data

- Online learning is to solve problems involving sequential interactions between data and environment
- Examples
  - Online classifications
  - Online advertising
  - Online investments
  - ...



Definition of Online Learning

- Machine learning problems

loss function

$$\min_{w} \sum_{t=1}^{T} l(f(x_t, w), y_t) + R(w)$$

hypothesis function, input, label, regularized term

- Online learning problems

$$\min_{\{w_1, w_2, \cdots, w_T\}} \sum_{t=1}^{T} l_t(f_t(x_t, w_{t-1}), y_t) + R(w_{t-1})$$

How To Solve Online Learning

- Statistical assumption: i.i.d. and adversarial
  - Recursive Least Squares (RLS) [H. Kushner & G. G. Yin, 2003]

$$
\begin{aligned}
F_t &= F_{t-1} - \frac{F_{t-1} x_t x_t^T F_{t-1}}{1 + x_t^T F_{t-1} x_t} \\
w_t &= w_{t-1} - F_t x_t (x_t^T w_{t-1} - y_t)
\end{aligned}
$$

  - Stochastic Gradient Descent (SGD) [M. Zinkevich, 2003]

$$w_t = w_{t-1} - \gamma x_t (x_t^T w_{t-1} - y_t)$$

  - Other online convex optimization techniques

Characteristics of Online Learning

- Memory: Full vs. Partial
  - Online learning can take on all the training data repeatedly or a subset of training data at once
- Feedback: Full vs. Partial
  - Output feedback can be partial or full
    - bandits vs. online regression
- Hypothesis: i.i.d. vs. non-stationary
  - Data generation can be stationary or adversarial
    - Regret bound: $O(\log T)$ vs. $O(\sqrt{T})$

Characteristic Taxonomy

- Online learning algorithms
  - i.i.d. & full output
    - oldest: Perceptron [F. Rosenblatt, 1957]
    - recursive least squares [H. Kushner & G. G. Yin, 2003]
    - stochastic gradient descent [M. Zinkevich, 2003]
  - convex
    - online convex optimization [S. Shalev-Shwartz, 2008]
    - follow the leader [S. Shalev-Shwartz & S. M. Kakade, 2009]
    - online subgradient descent [P. Liang, 2015]
  - partial output (bandits)
    - no input
      - adaptive allocations [T. L. Lai & H. Robbins, 1985]
      - upper confidence bound [P. Auer, N. Cesa-Bianchi & P. Fischer,2002]
    - input
      - bandits with advices [C. C. Wang etc. 2005]
      - contextual bandits [A. Slivkins, 2014]

Application Taxonomy

- Online learning algorithms
  - Classifications
    - Perceptron [F. Rosenblatt, 1957]
    - Winnow [N. Littlestone, 1989]
    - non-linear classification
      - ALMA, 2001
      - ROMMA, 2002
      - MIRA, 2003
      - PA, 2006
      - etc.

    [S. C. Hoi, J. Wang, & P. Zhao, 2014]
  - Regressions
    - kernel methods [J. Kivinen, etc., 2004] [E. Moroshko, etc., 2015]
    - online support vector regression [F. Parrella, 2007] [B. Gu, etc., 2015]
    - online regression with varying distributions [T. Hu, 2011] [M. S. Kan, etc., 2015]
  - Rankings
    - online collaborative filtering [J. Abernethy, etc., 2007] [S. Banerjee, 2016]
    - Bayesian ranking [R. C. Weng & C. J. Lin, 2011] [W. Pan, etc., 2015]

Our Recent Work

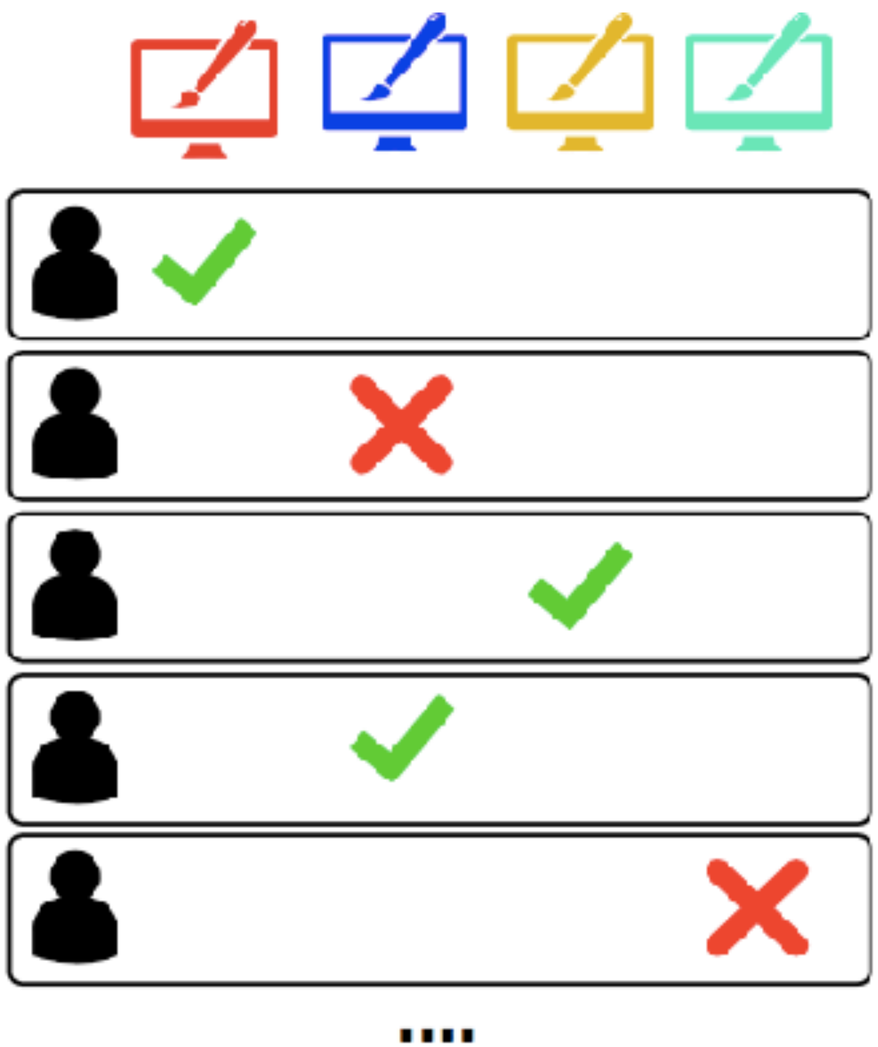
- Bandit algorithms for search and recommendation (NIPS2014, ICONIP2016, CIKM2016)
  - Combinatorial Pure Exploration of Multi-Armed Bandits [Chen et al., 2014]
  - Locality-Sensitive Linear Bandit Model for Online Social Recommendation [Zhao et al., 2016]
  - Constructing Reliable Gradient Exploration [Zhao et al., 2016]
- Online kernel classification (AAAI2015)
  - Kernelized Online Imbalanced Learning (KOIL) [Hu et al., 2016]

Bandit Algorithms for Recommendation

- Tackle the adaptive issues
  - Historical records are biased
  - User interests change over time
- Tackle the cold start problem
  - A great issue in recommender system
  - Lack of enough records/observations for new items or new users
- Our consideration
  - Using graph structures among items, e.g., spanning trees, paths, matching, etc.
  - Using graph structures among users, e.g., social networks, etc.

Combinatorial Pure Exploration in Multi-Armed Bandits

- Pure exploration of MAB in A/B testing, clinical trials, wireless network, crowdsourcing, ...
  - A fixed budget to minimize the probability of error
  - A fixed confidence value to minimize the number of rounds



- n arms = n variants
- play arm i = a page view on the i-th variant
- reward = a click on the ads
- finding the best arm = finding the variant with the highest average ads clicks

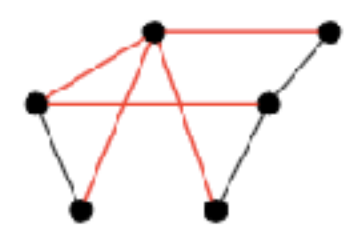
Combinatorial Pure Exploration (CPE)

- Play one arm at each round
- Find the optimal set of arms satisfying certain contents by maximizing the sum of expected rewards of arms in the set as

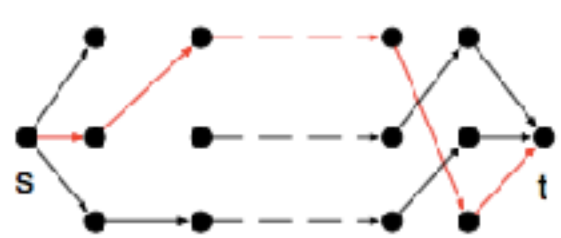
$$M_* = \underset{M \in \mathcal{M}}{\operatorname{argmax}} \sum_{i \in M} w(i)$$

size-k-sets

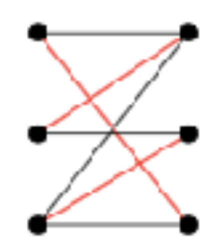
spanning trees



paths

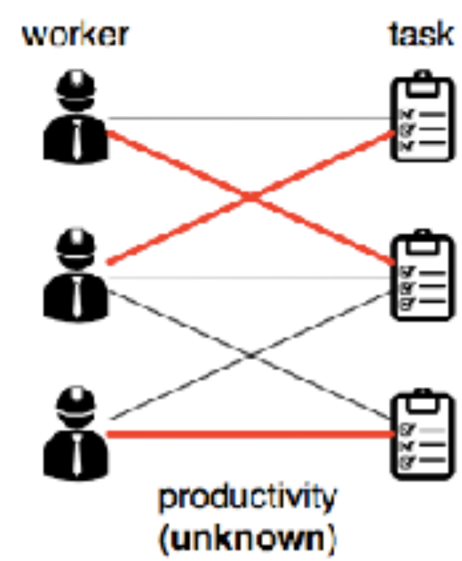


matchings



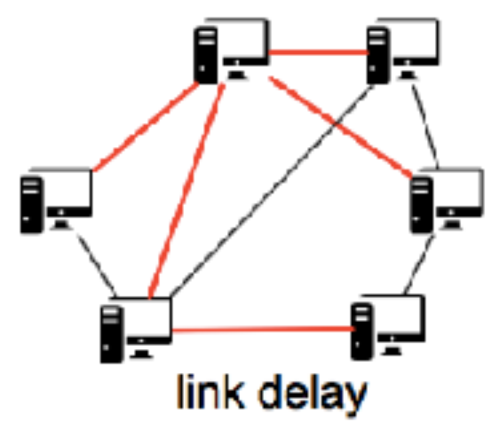
Motivating Examples

- matching



Goal:
1) estimate the productivities from tests.
2) find the optimal **1-1 assignment**.

- spanning trees and paths



Goal:
1) estimate the delays from measurements
2) find the minimum spanning tree
or shortest path.

- size-*k*-sets
  - finding the top-*k* arms.

Our Results

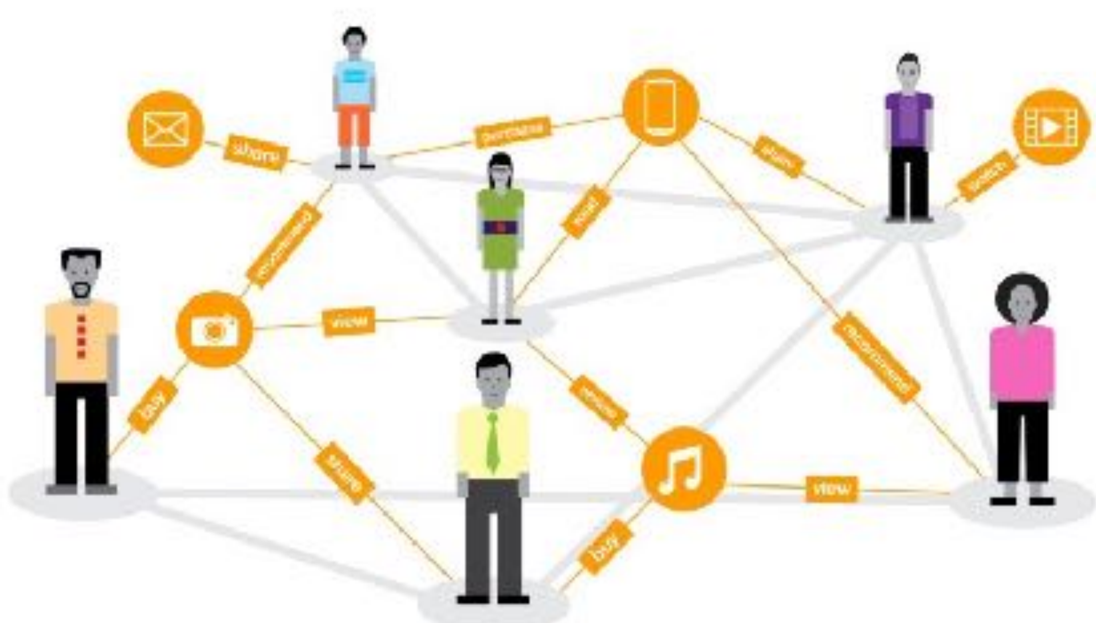
- Algorithms
  - Two general learning algorithms for a wide range of M
- Upper bounds
  - Sample complexity / probability of error
- Lower bounds
  - Algorithms are optimal (within log factors) for many types of M (in particular, bases of a matroid)
- Compared with existing work
  - The first lower bound for the top-k problem
  - The first upper and lower bounds for other combinatorial constraints

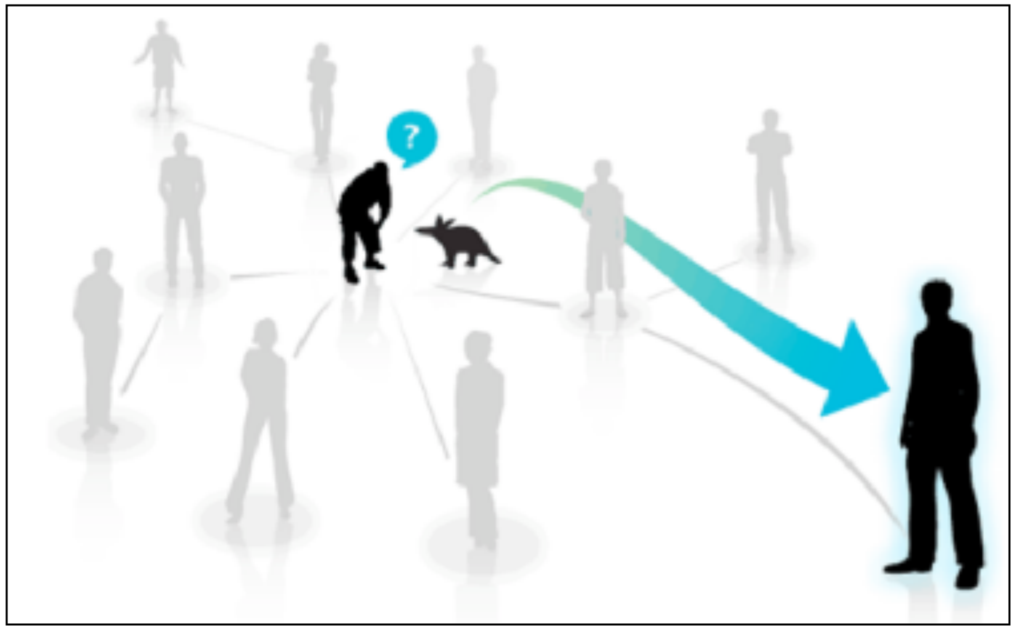
Locality-Sensitive Linear Bandit Model for Online Social Recommendation

- Motivations
  - Adaptive recommendation by incorporating social information
  - Contextual bandits for online recommendation
    - Arms ⟷ Items
    - Context ⟷ Item feature
    - Reward ⟷ Click/Purchase
    - Select ⟷ Recommend
- Applications
  - Recommender system
  - Online advertising

Locality-Sensitive Linear Bandit Model for Online Social Recommendation

- Most bandit algorithms focus on one-player modeling
- Existing social recommendation research focus on offline training
- Our goal is to integrate social network knowledge into bandit algorithms



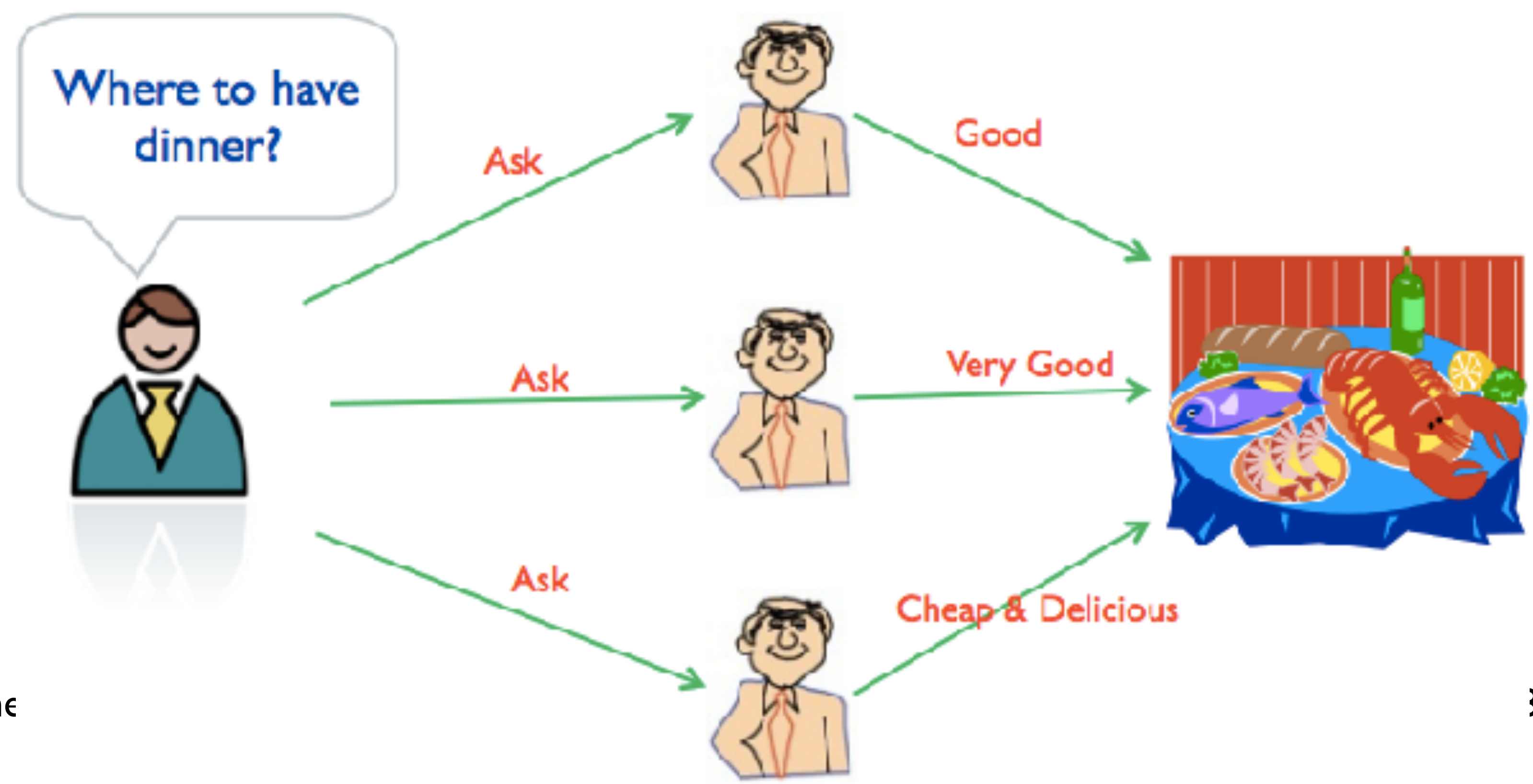


Locality-Sensitive Linear Bandit Model

- Linear Reward Assumption:
    - Given a user *u* and an item *i* with feature vector x_i at time *t*, the reward (preference) is modeled as

$$r_u(x_i) = x_i^T \theta_u^* + \eta_{i,t}$$

    - where θ_u^* is the unknown parameter of user *u* and $\eta_{i,t}$ is a sub-Gaussian noisy term.



Locality-sensitive Social Regularization Construction

- Social regularization term, summarizing local information from social relations

$$\hat{\theta}_{u'} = \sum_{v \in N(u)} \frac{\exp(n_v)\hat{\theta}_v}{\sum_{w \in N(u)} \exp(n_w)}$$

- Ridge Regression + Social Regularization term

$$\frac{1}{2}\|b_{u,t} - X_{u,t}\hat{\theta}_u\|_F^2 + \frac{\lambda}{4}\|\hat{\theta}_u\| + \frac{\lambda}{4}\|\hat{\theta}_u - \hat{\theta}_{u'}\|_F^2$$

- Closed-form solution

$$\hat{\theta}_u = (X_{u,t}X_{u,t}^T + \lambda I)^{-1}(X_{u,t}b_{u,t} + \frac{\lambda}{2}\hat{\theta}_{u'})$$

Input: $\lambda, \alpha_1, \alpha_2, ..., \alpha_T$
Initialization:
for *each user* u **do**
  $A_u^0 \leftarrow \lambda \mathbf{I}^{d\times d}, \mathbf{b}_u \leftarrow \mathbf{0}^d$
end
Simulation:
for *round* $t \leftarrow 1, ..., T$ **do**
  for *each user* u **do**
    for $v \in N_u(G)$ **do**
      $p_v \leftarrow \frac{\exp(n_v)}{\sum_{v' \in N_u(G)} \exp(n_{v'})}$
    end
    $\widehat{\theta}_{u'} \leftarrow \sum_{v \in N_u(G)} p_v \widehat{\theta}_v$
    $\widehat{\theta}_u \leftarrow A_u^{-1}(b_u + \frac{\lambda}{2}\widehat{\theta}_{u'})$
    for $i \in 1, ..., k$ **do**
      $\hat{r}_{t,a(i)} \leftarrow x_{t,a(i)}^T \widehat{\theta}_u + \boxed{\alpha_t \sqrt{x_{t,a(i)}^T A_u^{-1} x_{t,a(i)}}}$ UCB
    end
    Choose the arm $a_t^u \leftarrow \arg\max_{a(i)} \hat{r}_{t,a(i)}$
    Observe rewards r_{t,a_t^u}
    $A_u^t \leftarrow A_u^{t-1} + x_{t,a_t^u} x_{t,a_t^u}^T$
    $b_u \leftarrow b_u + r_{t,a_t^u} x_{t,a_t^u}$
    $n_u \leftarrow n_u + 1$
  end
end

Our Results

- Algorithm
  - Locality-sensitive linear bandit (LS.Lin) algorithm
- Theoretical analysis
  - Upper bounds of cumulative rewards
- Compared with existing methods
  - Only consider users' local social relations to avoid propagation of uncertainty to whole network
  - Use a softmax combination to differentiate the contribution from different social relations

Way Forward

- Adversarial environments
  - Contextual bandits with varying distributions
  - Non-linear rewards, Support approximate maximization oracles
- Non-convex assumption
  - Non-convex function for online update
- Social-related bandit algorithms
  - Explore complex structure (community, structure hole, etc.) in social network
  - Model the complex behaviors among users (cooperative vs. competitive, game theory, etc.)

Useful Links

- Tutorials
  - https://sites.google.com/site/banditstutorial/
  - http://ttic.uchicago.edu/~shai/icml08tutorial/
  - http://www.cs.princeton.edu/~ehazan/tutorial/tutorial.htm
- Workshops
  - NIPS 2015 Workshop on Non-convex Optimization for Machine Learning: Theory and Practice
  - Advances in non-convex analysis and optimization
  - NIPS 2010 Workshop: Machine Learning in Online ADvertising (MLOAD 2010)
  - Multi-armed Bandit Workshop 2016 at STOR-i, Lancaster University, UK

Useful Links

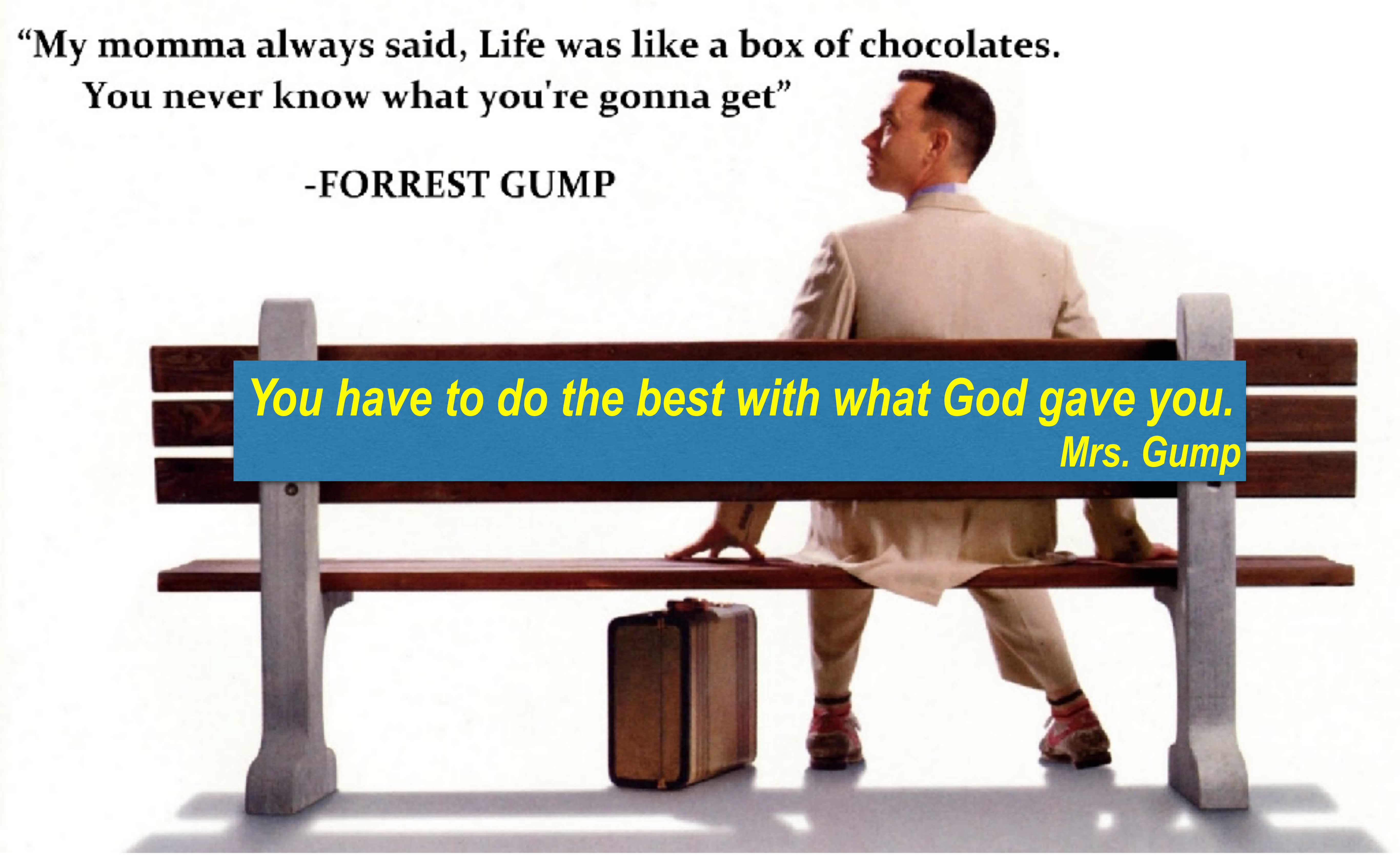
- Summer school and course
  - Online Learning Summer School (http://www.diku.dk/online-learning-summer-school-2015/)
  - Bandit Algorithms (http://banditalgs.com/)
- Library of online algorithms
  - DOGMA (Discriminative Online (**Good?**) Matlab Algorithms) (http://dogma.sourceforge.net/)
  - Vowpal Wabbit (Fast Learning) (http://hunch.net/~vw/)
  - LIBOL (A Library for Online Learning Algorithms) (http://libol.stevenhoi.org/)

Acknowledgments

- Ken Chan (Ph.D.)
- Wang Chan (Ph.D.)
- Xixian Chen (Ph.D.)
- Yaoman Li (Ph.D.)
- Han Shao (Ph.D.)
- Yuxin Su (Ph.D.)
- Yue Wang (Ph.D.)
- Xiaotian Yu (Ph.D.)
- Jichuan Zeng (Ph.D.)
- Hongyi Zhang (Ph.D.)
- Jiani Zhang (Ph.D.)
- Shenglin Zhao (Ph.D.)
- Tong Zhao (Ph.D.)
- Looking for PhD students working on machine learning, Big Data, social computing, …

Conclusion

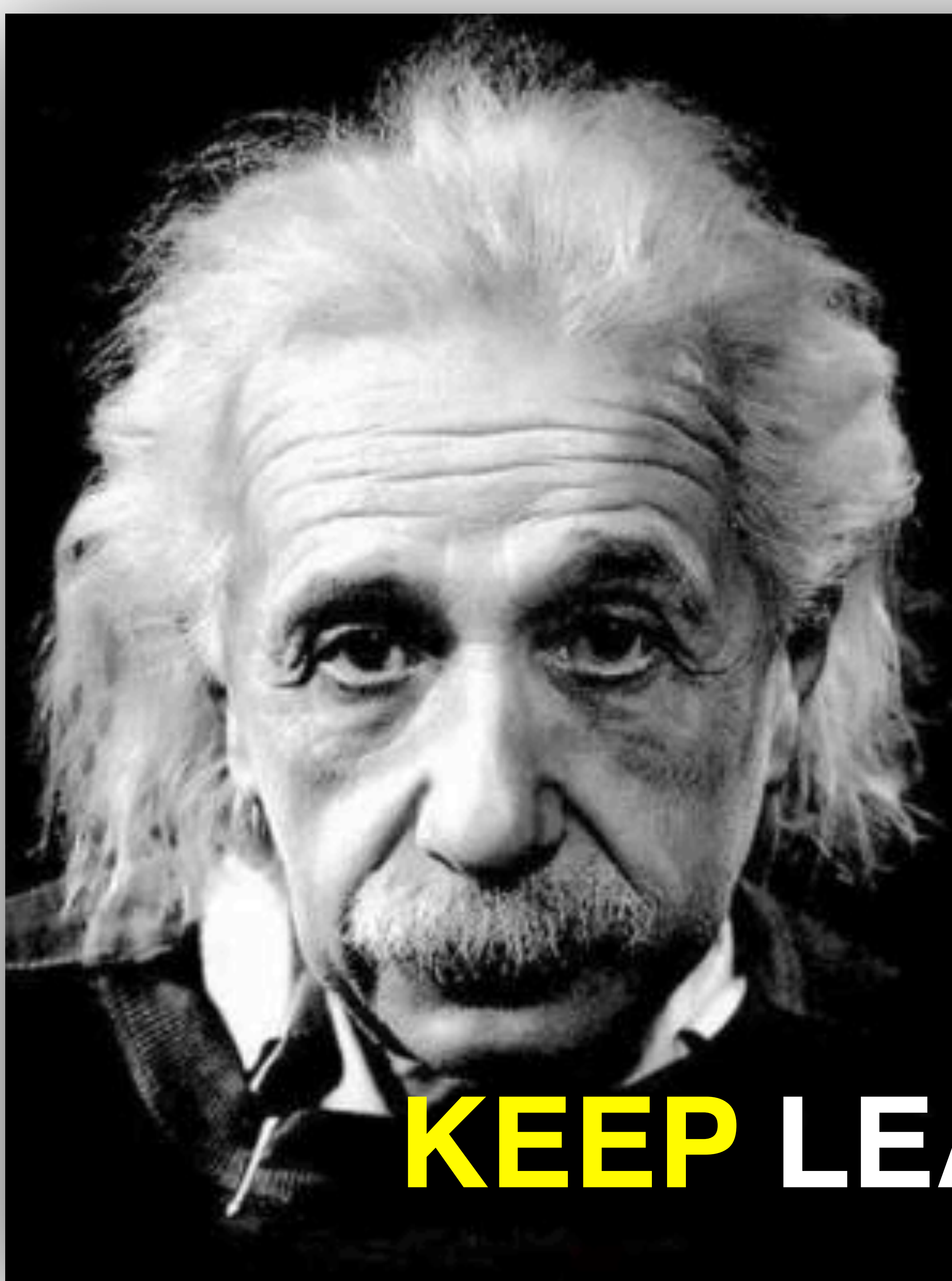
- Online learning is an effective approach to handle incoming data based on the sequential decision framework
- Reviewed literature based on characteristics and application taxonomy
- Present our recent work in designing bandit algorithms with graph structures



You have to do the best with what God gave you.

Mrs. Gump

https://www.quotesaga.com/quote/252/



Once you
stop learning,
you start
dying…

Albert Einstein

KEEP LEARNING!

KEEP

KEEP STANDS FOR

KNOWLEDGE **EDUCATION** **EXCHANGE** **PLATFORM**

PARTNERS & RELATIONSHIPS

TOP 8 HONG KONG INSTITUTIONS AND INDUSTRY ORGANIZATIONS

UNIVERSITY AND INDUSTRY PARTNERS

KEY FEATURES

TO HELP TEACHERS BECOME BETTER EDUCATORS

KEEPSearch

provides specific education related resources, including courses or events

KEEPCatalog

compiles tools and applications to create a community for educators to share best practices

KEEPCourse

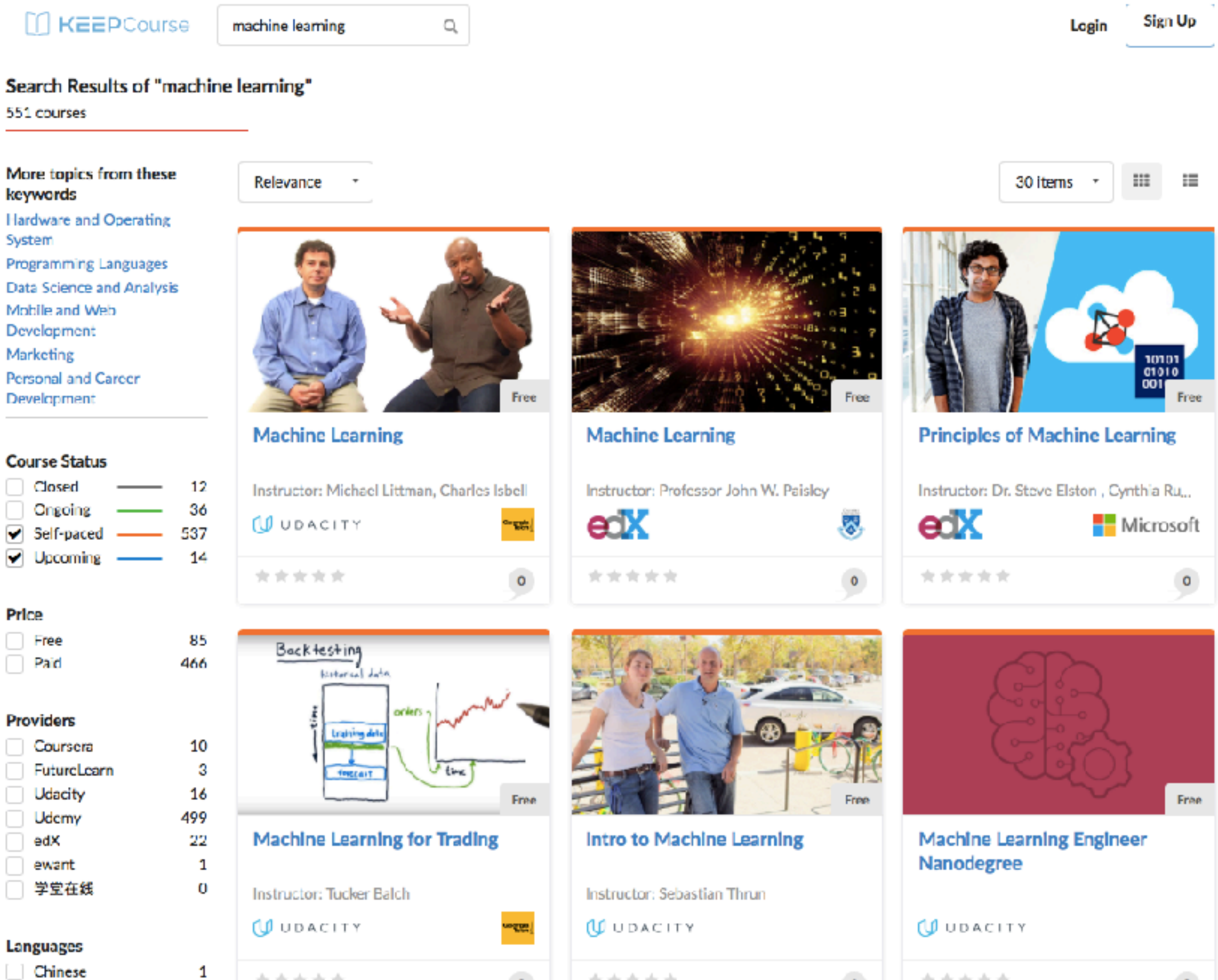
gives educators a chance to upload course content on to a private or global audience

KEEPoll

encourages in-class interaction among teachers and students to assess learning progress

KEEPAttendance

take attendance data for a large or small class by simply scanning a QR code!


KEEPCourse
machine learning
Login
Sign Up
Search Results of "machine learning"
551 courses
More topics from these keywords
Hardware and Operating System
Programming Languages
Data Science and Analysis
Mobile and Web Development
Marketing
Personal and Career Development
Course Status
Closed 12
Ongoing 36
Self-paced 537
Upcoming 14
Price
Free 85
Paid 466
Providers
Coursera 10
FutureLearn 3
Udacity 16
Udemy 499
edX 22
ewant 1
学堂在线 0
Languages
Chinese 1
Relevance
30 items
Free
Machine Learning
Instructor: Michael Littman, Charles Isbell
UDACITY
0
Free
Machine Learning
Instructor: Professor John W. Paisley
edX
0
Free
Principles of Machine Learning
Instructor: Dr. Steve Elston , Cynthia Ru...
edX
Microsoft
0
Backtesting
Free
Machine Learning for Trading
Instructor: Tucker Balch
UDACITY
0
Free
Intro to Machine Learning
Instructor: Sebastian Thrun
UDACITY
0
Free
Machine Learning Engineer Nanodegree
UDACITY
0

KEEP'S MILESTONES

What have we accomplished?

8.4M Unique Sites Crawled

4249 Courses Indexed

1345 Products Reviewed

5342 Class Poll Interactions

4500+ Users Registered

9 Platforms indexed

5 + 4 Partner Integrations

3 New Products Launched

KEEP'S DIRECTION

ANALYTICS

- Analyze user's click behavior and search patterns for **learning analytics**
- Recommend **personalized education** programs and services for stakeholders

GAMING

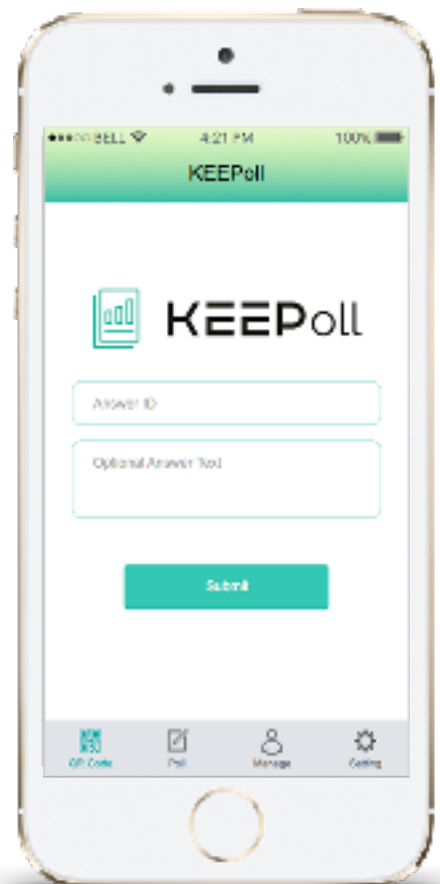
- Provide **gaming technology** to encourage engagement and participation
- Motivate and empower students through a fun and **innovative approach**

SOCIAL

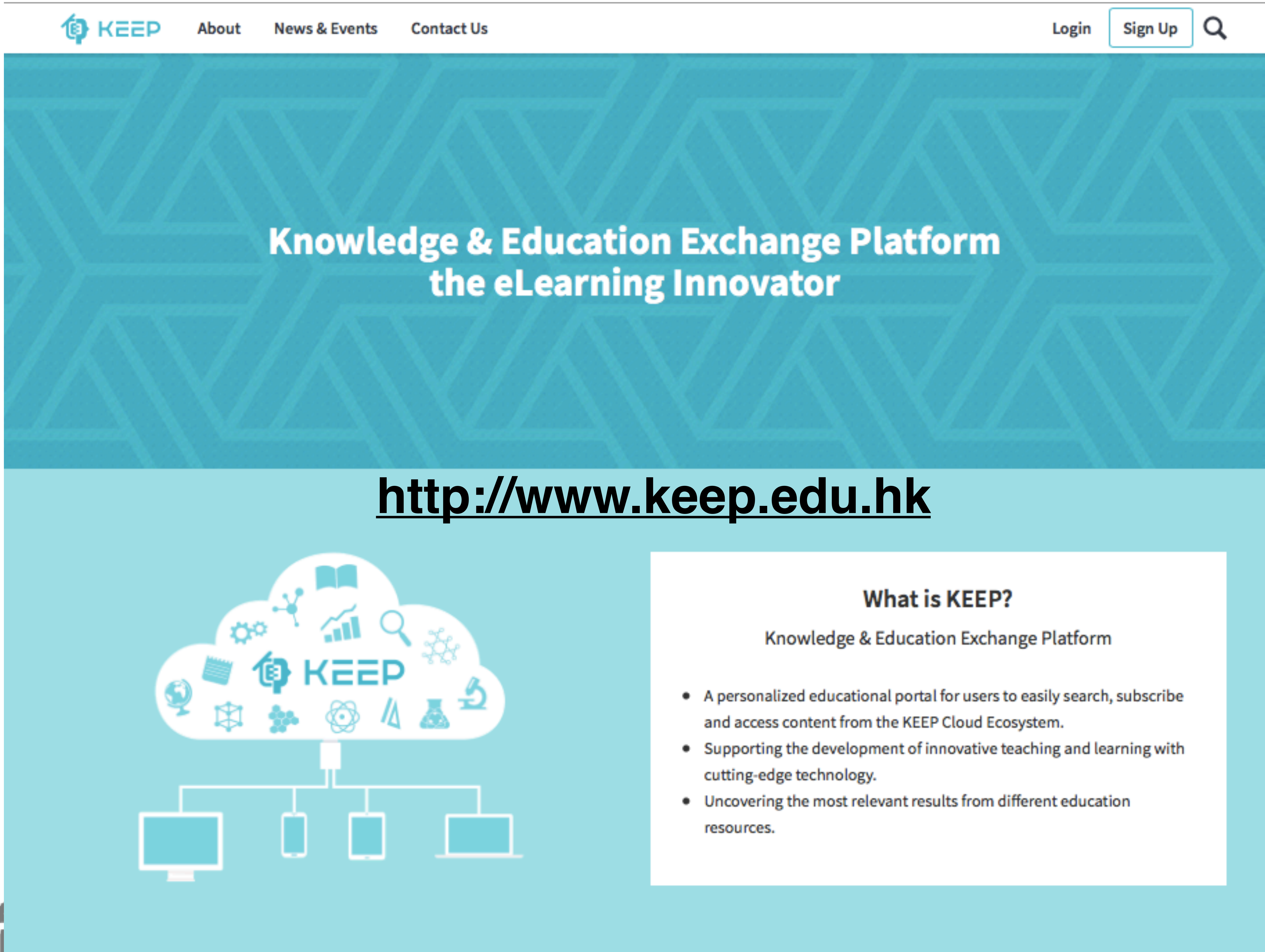
- Connect users with one another via **social networks** and special interest groups
- Promote **collaboration** by provide avenues for group/ community learning

MOBILE

- Enhance learning and teaching experiences through **personalized devices**
- Develop mobile and **wearable technology** for KEEP


KEEPoll
KEEPoll
Submit


KEEPCatalog
Login



http://www.keep.edu.hk

CUHK Excellence

- The only university in Hong Kong having Nobel Laureates as faculty with **five** Distinguished Professors-at-Large

Prof. Yang Chen-Ning, Nobel Laureate in Physics

Prof. Charles Kao Nobel Laureate in Physics

Prof. Sir James A. Mirrlees, Nobel Laureate in Economic Sciences

Prof. Yau Shing-Tung, Fields Medalist

Prof. Andrew Yao, Turing Award

- **Nine** academicians of Chinese Academy of Sciences and Chinese Academy of Engineering


The Chinese University of Hong Kong
KEEP LEARNING FOR LIFE!

Q&A