

# هوش مصنوعی و استنتاج کاتوره‌ای

سیاوش بیگدلی

کاظم داوودی، علی اکبر شمس، حسن کربلایی

# قائده استنتاج

$$\frac{A}{B} \rightarrow B$$

- قائده
- پيش فرض
- نتيجه گيري
- معيار بهينگي
- فضاي جستجو
- بهينه سازي

# تئوری کوچک فرما (Fermat)

- آیا  $N$  یک عدد اول است؟
- روش جبری : بررسی تمام ضرب‌های کوچکتر
- روش تصادفی :
- ویژگی عدد اول  $N$   $X^N - X \propto N$  به ازای همه‌ی اعداد طبیعی  $X$
- با تعدادی عدد تصادفی  $X$  احتمال آن را بسنجیم (توزیع یکنواخت)
- بدون قطعیت، جواب احتمالاتی
- بسیار سریع به طمأنینه می‌رسیم

# ایان هکینگ (Ian Hacking)

- جایگزین ضعف مدل‌های جبری (قرون ۱۷ و ۱۸): مدل‌سازی کاربردی در قمار، بیمه و تحقیقات جمعیتی

- امر تصادفی: عبور از «جهالت» به نگاه «ابزاری» و «ذاتی»

- نگاه جبری: عدم قطعیت به مثابه جهالت

- نگاه آماری: ابزارگری احتمالات حتی آنجا که دلیل اصلی ناشناخته است


- نگاه ذاتی: وقایع میتوانند ذات تصادفی داشته باشند

# استدلال تصادفی

- استدلال تصادفی در neuroscience
- تصادف در رفتار انسانی
- تصادف در طراحی الگوریتم‌ها ([How Randomness Improves Algorithms](#))

نسبت به ابزارهای موجود  
هوش مصنوعی استفاده‌ی بیشتری از  
امر تصادفی دارد در یادگیری و  
استنتاج. این باعث میشود که یک  
سری فرایندها را بهتر مدل کند.

ورود امر تصادفی به تعریف‌ها و پیش‌فرض‌ها

A dark brown camel stands in a vast, arid desert landscape. The camel is facing left, with its head slightly turned. The ground is dry, sandy, and dotted with small, low-lying green shrubs. The sky is a clear, pale blue. The camel's hump is large and rounded, and its legs are dark and sturdy.

# معاینه یعنی دیدن با احتیاط دور از شک و هیجان

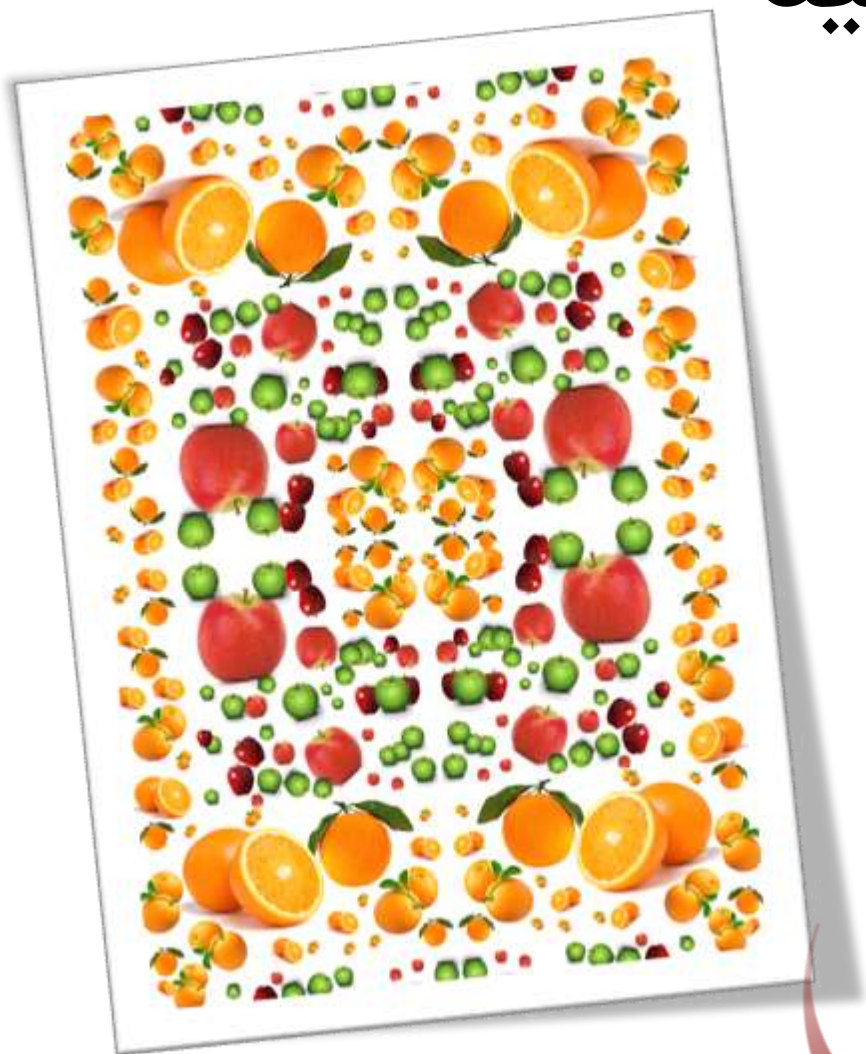
شیخ عبدالله انصاری

# تمرین عملی

- سیب و پرتقال‌ها را با یک خط مستقیم از هم جدا کنید
- از روی میوه‌ها رد نشد
- همه‌ی میوه‌ها، پشت و روی برگه
- هر چند بار بخواهید می‌توانید کاغذ را تا کنید

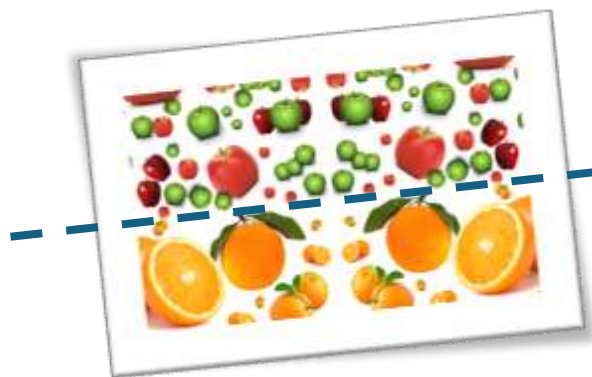


# سیب و پرتقال‌ها را از هم جدا کنید



• کاغذ را تا کنید

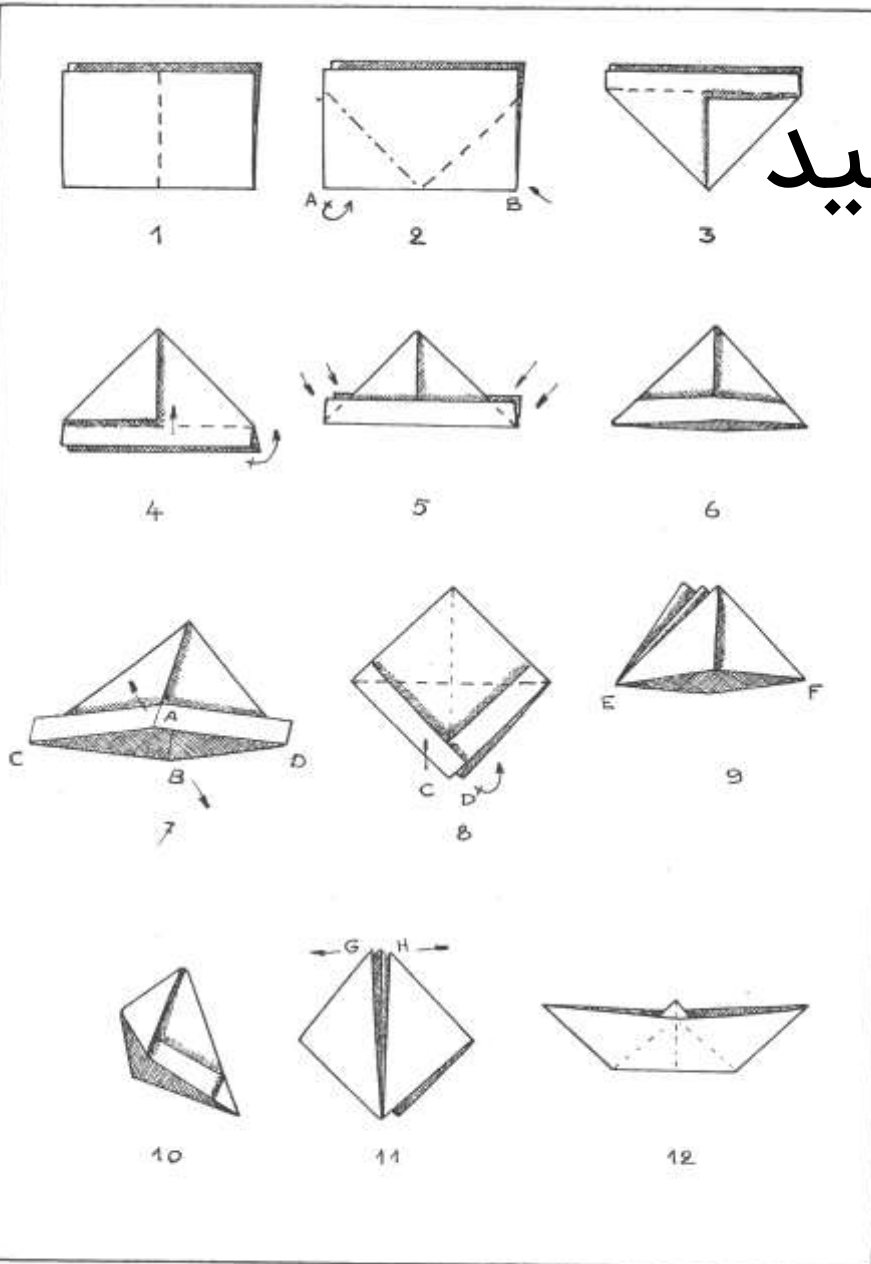
- همه‌ی سیب‌ها یک طرف، پرتقال‌ها طرف دیگر
- حتی آنهایی که پشت صفحه هستند ☺
- حتی سیب کوچولوهای دور برگه



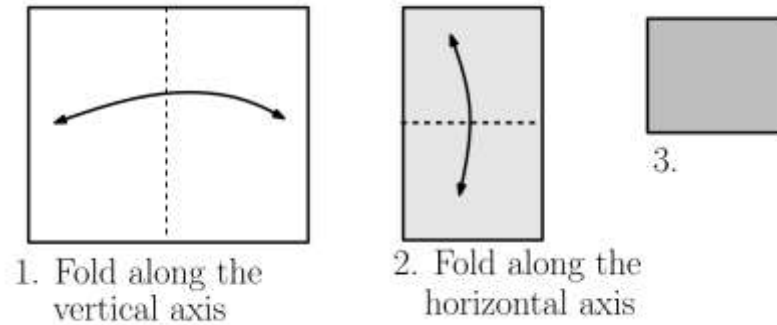
# سیب و پرتقال‌ها را از هم جدا کنید

• کاغذ را تا کنید

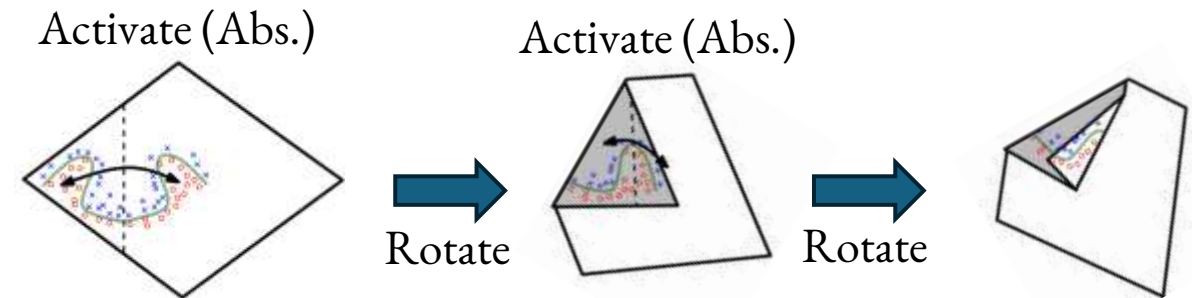
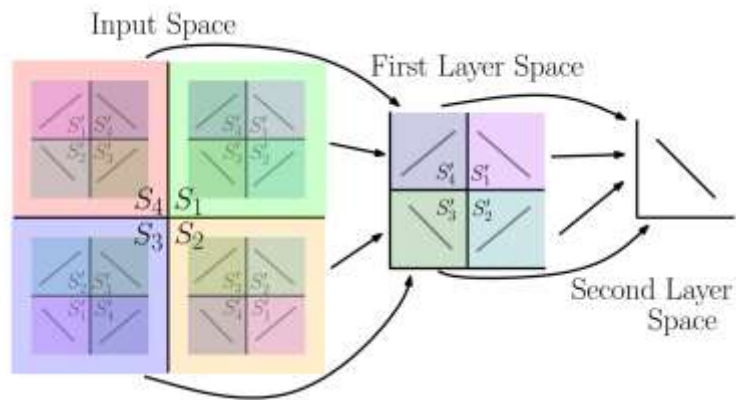
- همهی سیب‌ها یک طرف، پرتقال‌ها طرف دیگر
- حتی آنهایی که پشت صفحه هستند ☺
- حتی سیب کوچولوهای دور برگه



# شبکه‌های عصبی عمیق



- تا کردن فضا (قدر مطلق به عنوان activation)
- تبدیل (مانند fully connected)



# شبکه‌های عصبی عمیق

- مدل شما بر چه اساسی دسته بندی می‌کند؟
- چه اتفاقاتی در فرایند مدل‌سازی در Pytorch/TF می‌افتد؟
- چه نسبتی با توقعات ما دارد؟

فرض کنید شما یک تصمیم بر  
پایه‌ی هوش مصنوعی میگیرید. چه  
توجیهی برای این تصمیم دارید؟

مانند سیستم COMPAS و توجیه تصمیم برای مدافع

# شناخت‌شناسی هوش مصنوعی

• بررسی پیش‌فرض‌ها، فرایندها، محدودیت‌ها یا فواید

## Probability theory

1. *Kolmogorov axioms*
2. *Probabilities (normalized distribution)*
3. *Conditional probability*
4. *Conditional independence*
5. Probabilistic problems

## Statistics

6. *Bayes rule*
7. Identically distributed
8. Ergodic process
9. Fast mixing weakly correlated samples
10. Exchangability
11. Stationarity
12. Independent sampling
13. *Memoryless property*
14. *Time-reversible process*

## Statistical learning

15. *[Bayesian] risk minimization*
16. *Empirical Bayes method*
17. *Non-parametric learning*
18. Instance-based methods (lazy)
19. Pattern recognition
20. Representer theorem (RKHS)
21. Predictive learning (extrapolation)
22. *Probably Approximately Correct (PAC)*
23. *Statistical Inference*

## Modeling

24. *Parametrization*
25. *Selection problem*
26. *[Stochastic] optimization*
27. *Structured prediction*
28. *Exiting statistics (entering data-science)*
29. *Likelihoodism (data-independent inference)*
30. *Consequentialism (generalization capability; data-irrelevant modeling)*

# کجا امر تصادفی در استنتاج استفاده می‌شود؟ کجا این روش مقبول است؟

ورود امر تصادفی به تعریف‌ها و پیش‌فرض‌ها



تصادف را امری فرض می‌گیریم که  
از حیطه‌ی دانش ما خارج است؛  
نمی‌توانیم ربط آن را بفهمیم

ورود امر تصادفی به تعریف‌ها و پیش‌فرض‌ها



# فرایندهای تصادفی در پیش فرض های هوش مصنوعی

آماري	۸. فرآیند ارگودیک
	۹. نمونه های ترکیب سریع با همبستگی ضعیف
	۱۰. قابلیت تعویض
	۱۳. ویژگی بدون حافظه
یادگیری آماری	۱۶. بهینه سازی تجربی ریسک
	۲۲. احتمالاً تقریباً صحیح (PAC)
	۲۳. استنتاج آماری (تصادفی)
	۲۴. پارامترگذاری
مدل سازی	۲۵. مسئله انتخاب مدل
	۲۶. بهینه سازی تصادفی

# پیش فرض های آماری

۸. فرآیند ارگودیک : میانگین های زمانی یک فرآیند به میانگین توزیع اصلی همگرا می شوند. مثلاً میانگین مشتق دسته ها برابر میانگین مشتق همه ی پایگاه داده است [\[Analysis of Stochastic Gradient Descent in Continuous Time\]](#)

۹. نمونه های ترکیب سریع با همبستگی ضعیف : وابستگی بین نمونه ها به سرعت کاهش می یابد و برای تقریباً استقلال را تضمین می کند

۱۰. قابلیت تعویض : توزیع مشترک داده ها تحت جایگشت ها تغییر نمی کند، به این معنی که ترتیب نمونه ها اهمیتی ندارد

# بدون حافظه‌گی و برگشت‌پذیری زمانی

- خاصیت مارکوف: فرآیندی که بدون حافظه است، تکامل آینده به تاریخچه وابسته نیست و تنها به حالت فعلی (حال) بستگی دارد

$$p(x_t | x_{t-1}, x_{t-2}, \dots, x_0) = p(x_t | x_{t-1})$$

- فرآیند مارکوف برگشت‌پذیر است اگر احتمال‌های مشترک در جهت رفت و برگشت برای زمان‌های همسایه یکسان باشند

$$p(x_t = i, x_{t-1} = j) = p(x_t = j, x_{t-1} = i)$$

# فرایندهای تصادفی در پیش فرض های هوش مصنوعی

آماري	۸. فرآیند ارگودیک
	۹. نمونه های ترکیب سریع با همبستگی ضعیف
	۱۰. قابلیت تعویض
	۱۳. ویژگی بدون حافظه
یادگیری آماری	۱۶. بهینه سازی تجربی ریسک
	۲۲. احتمالاً تقریباً صحیح (PAC)
	۲۳. استنتاج آماری (تصادفی)
	۲۴. پارامترگذاری
مدل سازی	۲۵. مسئله انتخاب مدل
	۲۶. بهینه سازی تصادفی

# یادگیری آماری

۱۶. بهینه‌سازی تجربی ریسک

۲۲. احتمالاً تقریباً صحیح (PAC)

۲۳. استنتاج آماری (تصادفی)

# بهینه‌سازی تجربی ریسک

• همگرایی در احتمال (پایداری)

$$\lim_{n \rightarrow \infty} P(|X_n - X| > \epsilon) = 0$$

• قانون ضعیف اعداد بزرگ : میانگین دسته به میانگین توزیع همگرا در احتمال می‌شود

• قضیه حد مرکزی : جمع متغیرهای تصادفی به توزیع نرمال همگرا می‌شود

# احتمالاً تقریباً صحیح (Probably Approximately Correct)

- $x \in \{0,1\}^n$  : نمونه از توزیع احتمالی  $p$
- $c \subseteq \{0,1\}^n$  : مجموعه برچسب‌های حقیقی (concept)
- $h \subseteq \{0,1\}^n$  : مجموعه برچسب‌های فرضیه (تخمین زده شده)
- $\Delta$  : تفاوت متقارن مجموعه‌ها

• **تعریف خطا** : احتمال اختلاف فرضیه با حقیقت واقعی

$$error(h) = P_x(h(x) \neq c(x)) = \sum_{x \in h \Delta c} p(x)$$

# احتمالاً تقریباً صحیح (Probably Approximately Correct)

$$\delta < 1, \epsilon > 0$$

- الگوریتم  $A$  با پیچیدگی زمانی چندجمله‌ای
- اندازه‌ی پایگاه داده حداقل چندجمله‌ای در  $\frac{1}{\delta}, \frac{1}{\epsilon}, n$
- $C$  : مجموعه تمام برچسب‌های ممکن  $c \subseteq \{0,1\}^n, n > 1$
- $H$  : مجموعه تمام فرضیه‌های ممکن  $h \subseteq \{0,1\}^n, n > 1$

کلاس مفاهیم  $C$  توسط فضای فرضیه  $H$  قابل یادگیری احتمالاً تقریباً صحیح است اگر که الگوریتم آموزشی  $A$  وجود داشته باشد که با احتمال  $1 - \delta$  فرضیه‌ای  $h$  بازگرداند به طوری که  $error(h) < \epsilon$ . به عبارت دیگر،  $A$  یک الگوریتم یادگیری PAC برای  $C$  است.



# نظریه‌های جایگزین یادگیری محاسباتی

- یادگیری آنلایین : مدل‌ها در طول زمان با دریافت داده‌های جدید به‌روزرسانی می‌شوند
- یادگیری دقیق : مدل‌ها برای یافتن پاسخ‌های دقیق و بدون خطا به‌روزرسانی می‌شوند  
(مانند [PExact](#))
- یادگیری الگوریتمی : روش غیر احتمالاتی برای یادگیری دقیق آنلایین
- یادگیری توزیعی : متمرکز بر پردازش موازی داده‌های آموزش
- یادگیری از پرس‌وجوها : یادگیرنده برای بهینه‌تر شدن سوال جدید طراحی می‌کند
- یادگیری تعاملی

# یادگیری تعاملی

- فضای جستجو/اعمال مانند لایه‌های شبکه عصبی
- معیار بهینگی مانند سرعت و دقت
- الگوریتم جستجو مبتنی بر نویز مانند ژنتیک، یادگیری تقویتی، مدل مولد

# استنتاج آماری تصادفی

- مبتنی بر نمونه‌گیری : تخمین تابع توزیع با استفاده از دسته‌ای از نمونه‌ها (مانند [MC]MC)
- استنتاج مکرر : به‌روزرسانی تخمین بر اساس بخشی از نمونه‌ها (مانند expectation propagation)
- استنتاج تغییرات (وَرَدَشی) تصادفی : جایگزینی توزیع واقعی با توزیع‌های محاسبه‌پذیر روی دسته‌های کوچک تصادفی
- مدل‌سازی مولد تصادفی

# فرایندهای تصادفی در پیش فرض های هوش مصنوعی

آماري	۸. فرآیند ارگودیک
	۹. نمونه های ترکیب سریع با همبستگی ضعیف
	۱۰. قابلیت تعویض
	۱۳. ویژگی بدون حافظه
یادگیری آماری	۱۶. بهینه سازی تجربی ریسک
	۲۲. احتمالاً تقریباً صحیح (PAC)
	۲۳. استنتاج آماری (تصادفی)
	۲۴. پارامترگذاری
مدل سازی	۲۵. مسئله انتخاب مدل
	۲۶. بهینه سازی تصادفی

# مدل سازی

۲۴. پارامترگذاری

۲۵. انتخاب مدل

۲۶. بهینه سازی تصادفی

۲۸. خروج از احتمالات و ورود به علوم داده

# پارامترگذاری

- مقداردهی تصادفی (ثابت یا بهینه‌شونده)

- ۲٪ تفاوت دقت در تکرارهای مختلف بر CIFAR

- پیش‌افکنی تصادفی (RP) در شبکه‌های عصبی عمیق

- Johnson-Lindenstrauss lemma

- (هم‌گرایی در احتمال)

RP scheme	Matrix construction time	Projection time		Embedding quality
		Dense input	Sparse input	
Gaussian	$\mathcal{O}(dk)$	$\mathcal{O}(ndk)$	$\mathcal{O}(\text{nnz}(\mathbf{A})k)$	$\mathcal{O}(\epsilon^{-2}n)$
Achlioptas'	$\mathcal{O}(dk)$	$\mathcal{O}(ndk)$	$\mathcal{O}(\text{nnz}(\mathbf{A})k)$	$^a$
Li's	$\mathcal{O}(\sqrt{d}k)$	$\mathcal{O}(n\sqrt{d}k)$	$\mathcal{O}(\text{nnz}(\mathbf{A})k)^b$	$^a$
SRHT	$\mathcal{O}(dk + d \log d)$	$\mathcal{O}(nd \log k)$	$\mathcal{O}(nd \log k)$	$\mathcal{O}(\epsilon^{-2}(n + d) \log n)$
Count Sketch	$\mathcal{O}(d)$	$\mathcal{O}(nd)$	$\mathcal{O}(\text{nnz}(\mathbf{A}))$	$\mathcal{O}(\epsilon^{-2}n^2)$

<sup>a</sup> For Achlioptas' and Li's projections, we did not find any estimates of  $k$  for which the OSE property holds.

- شبکه‌های ویژگی تصادفی

# مسئله انتخاب مدل

- جستجوی معماری شبکه‌های عصبی (NAS) مثلاً بر پایه‌ی یادگیری تقویتی، ژنتیک، فرایند گاس+بیز
- یادگیری گروهی (مانند Random forests, Random Projections)
- نمونه‌برداری مدل (مانند Bayesian Neural Networks, Dropout)

# بهینه‌سازی تصادفی

- جستجو در الگوریتم ژنتیک و یادگیری تقویتی

- داده‌افزایی (مانند Mix-out)

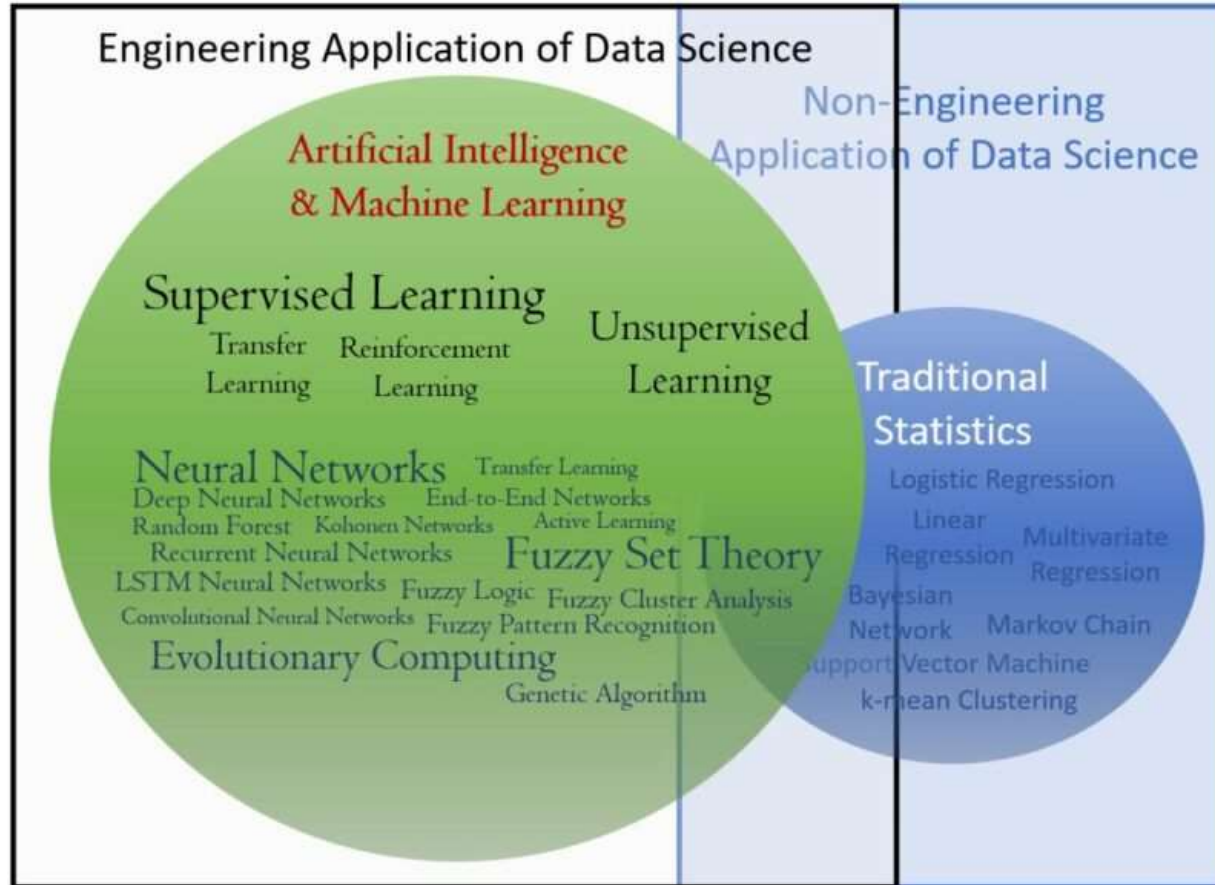
- تابع بهینه تعریف‌شده با برون‌فکنی تصادفی

(مانند [GAN Training with Multiple Random Projections](#) و [Random Linear Projections Loss](#))

- هنجارسازی (مانند Group norm, batch norm)



# خروج از احتمالات



Shahab D. Mohaghegh, Traditional Statistics vs. Artificial Intelligence and Machine Learning

آمار:

- فرضیه
- مدل سازی
- جمع آوری داده
- اعتبارسنجی
- هوش مصنوعی (علوم داده):
- داده
- فرضیه
- مدل سازی
- اعتبارسنجی

فرضیه به تابعی از داده های  
تصادفی تبدیل می شود.

# Stacking ensemble model of deep learning for plant disease recognition

بهینه‌سازی تجربی ریسک	SGD
استنتاج آماری (تصادفی)	
فرآیند ارگودیک	
قابلیت تعویض	
مقداردهی اولیه تصادفی	پارامترگذاری
یادگیری گروهی	
مسئله انتخاب مدل	جستجوی معماری شبکه‌های عصبی (NAS)
بهینه‌سازی تصادفی	داده‌افزایی
همگرایی در احتمال	احتمالاً تقریباً صحیح (PAC)
داده قبل از مدل	خروج از احتمالات

# استفاده از امر تصادفی در فرایندهای هوش مصنوعی

- شروع با تصادف و بازیابی مکرر
  - پارامترهای شبکه
  - مدل‌های بر پایه انتشار (diffusion)
- شروع با تصادف
  - شبکه‌های تصادفی برای استخراج ویژگی‌ها
  - برون‌فکنی تصادفی تابع خطا و مدل
  - دسته‌بندی برای آموزش (SGD)
- استنتاج با تصادف؛ مدل‌های مولد (نمونه‌برداری)
  - GAN, Normalizing Flows, Denoising Diffusion
  - Autoregressive generative models (e.g. LLMs)

# تصادف برای افزودن کارآمدی

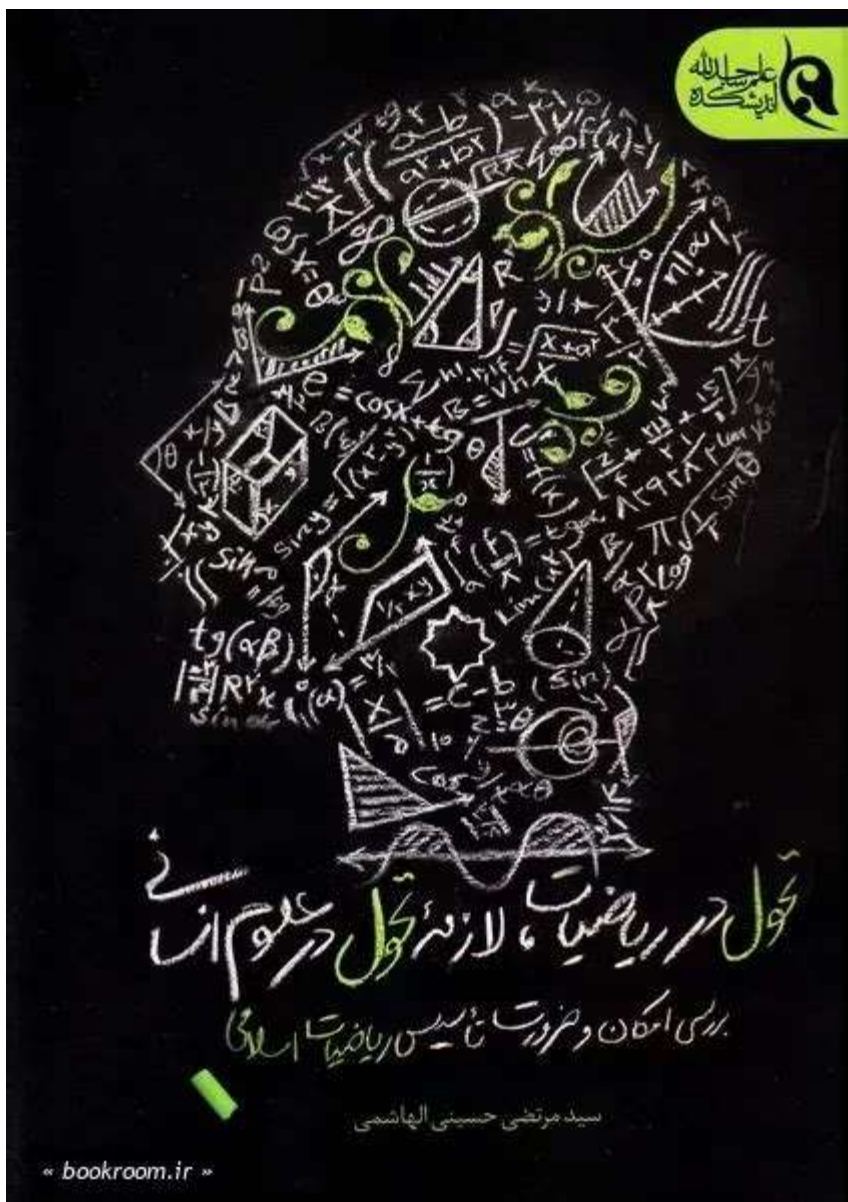
- الگوریتم جستجو :  $x$  در پایگاه داده وجود دارد؟  
 $O(d \log N)$   $p(x)$
- موتور جستجو : کدام  $x$  در پایگاه داده وجود دارد؟  
 $< O(\log d \log N)$   $p(x|y)$
- مدل مولد شرطی : نمونه  $x$  در پایگاه داده چیست؟  
 $O(d)$   $p(x|y)$

سرعت، دقت، تعمیم

# پارادیم‌های کمی‌سازی

- حجت‌الاسلام سید مرتضی حسینی الهاشمی
- پارادیم‌های ریاضی متحول می‌شوند برای مدل‌سازی بهتر
- ولی نقص در مدل‌سازی اختیار فردی و اجتماعی

- مصاحبه در آپارات



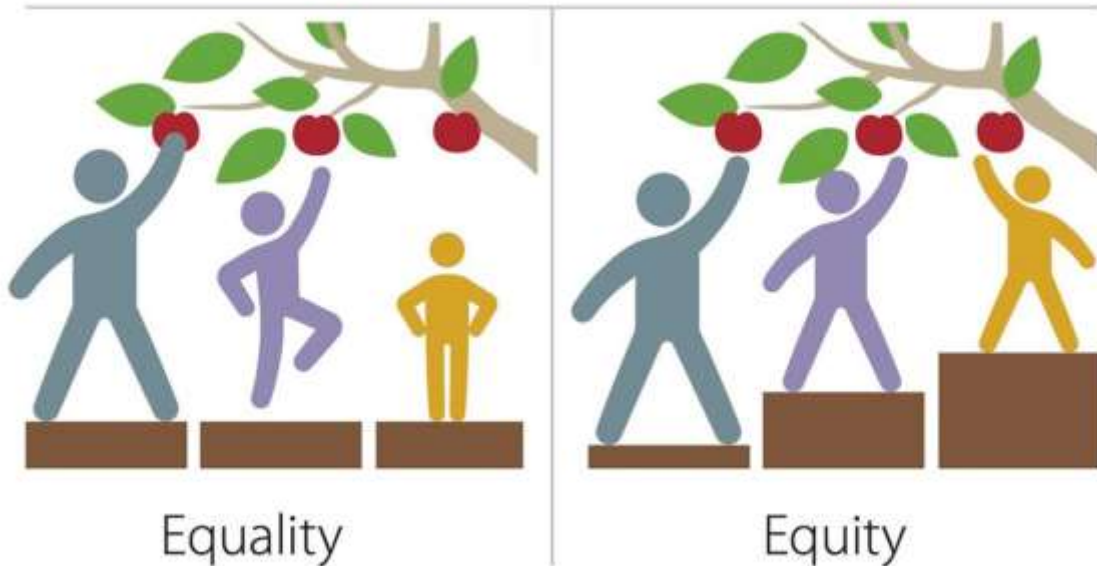
# استفاده از امر تصادفی و آگاهی

- خلاقیت بیشتر انسان
- خلق گمانه جدید
- جهت‌دهی با علاقه‌ی انسان
- آنچه نمیدانیم را توکل می‌کنیم
- آیا در رابطه‌ی بین دو نفر همیشه این را فرض گرفت؟



# استفاده از امر تصادفی و عدالت

- مناسب برخورد منصفانه (مساوات)
- منطبق با فلسفه‌ی علمی-تکتورژیک فمینیستی (feminism techno-science)
- تقلیل عدالت به انصاف (justice as fairness)
- امر تصادفی با توزیع یکسان برای همه‌ی گه‌ها
- نقض نقش اراده‌ی فردی-اجتماعی



# استفاده از امر تصادفی و آزادی

- آزادی فردی، اجتماعی، تاریخی
  - امر اجتماعی و تاریخی به مثابه امر تصادفی
  - آزادی در مقابل رهایی
  - وحدت بخشی
- رهایی انسان از زنجیره‌ی سرمایه داری
  - افزایش انصاف در امکانات شناختی
  - افزایش شکافت شناختی
- کمک به افزایش اختیار انسان
  - شیر یا خط
- کجا و چگونه می‌توان تصادف را با خواست کاربر جایگزین کرد؟



# استفاده از امر تصادفی و آزادی

- آزادی فردی، اجتماعی، تاریخی
  - امر اجتماعی و تاریخی به مثابه امر تصادفی
  - آزادی در مقابل رهایی
  - وحدت بخشی
- رهایی انسان از زنجیره‌ی سرمایه داری
  - افزایش انصاف در امکانات شناختی
  - افزایش شکافت شناختی
- کمک به افزایش اختیار انسان
  - شیر یا خط
- کجا و چگونه می‌توان تصادف را با خواست کاربر جایگزین کرد؟

# هوش مصنوعی و استنتاج کاتوره‌ای

- آیا ما این را با جبر و یا تصادف جایگزین کردیم؟
- اختیار در چه مقطعی از هوش مصنوعی می‌تواند دخیل باشد؟

