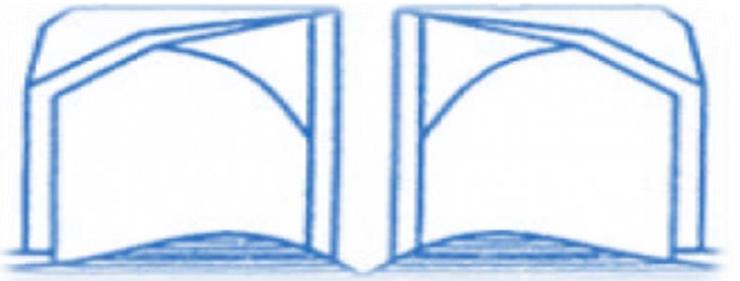




دانشگاه علوم فناوری های میان رشته ای
دانشکده سایر های به شنبه



انقلاب پردازش زبان طبیعی

نحوه عملکرد مدل‌های زبانی بزرگ و آینده هوش مصنوعی زایینده

هادی ویسی

h.veisi@ut.ac.ir

مردادماه ۱۴۰۳

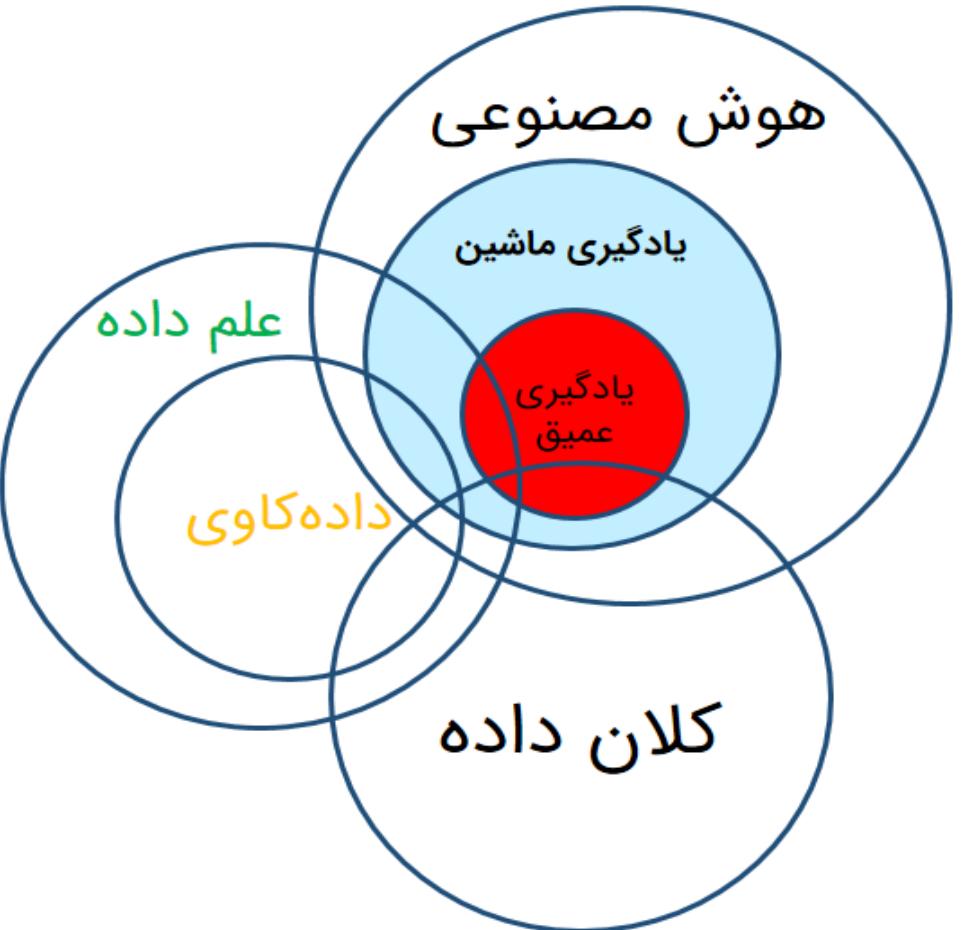
هوش مصنوعی: تعریف

▣ هوش مصنوعی (Artificial Intelligence)

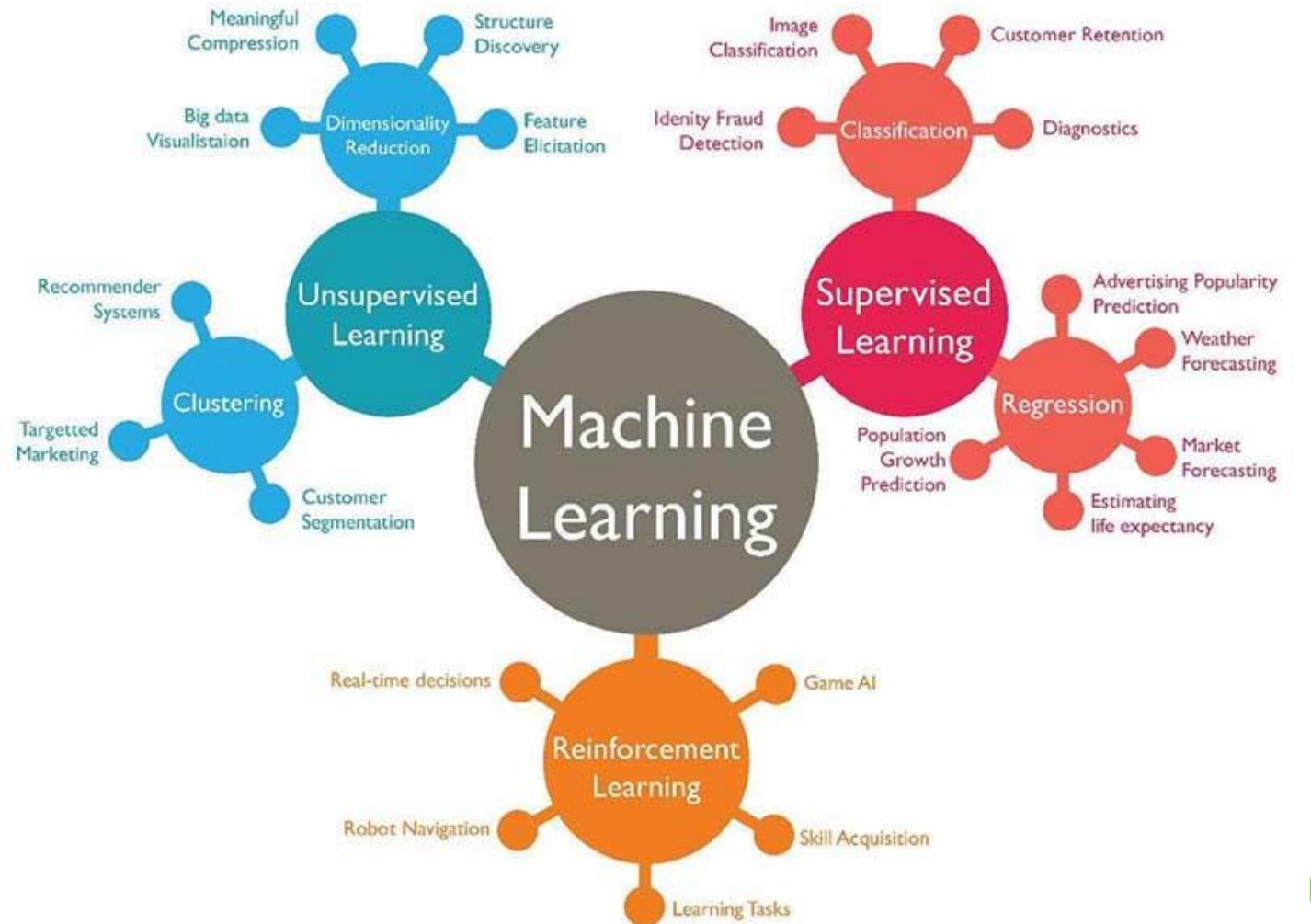
- ◀ شاخه‌ای از علوم رایانه که هدف آن تولید ماشین‌هایی است که توانایی انجام کارهایی را دارند که نیازمند هوش انسانی است



هوش مصنوعی: مفاهیم مرتب



هوش مصنوعی: انواع یادگیری



هوش مصنوعی مولد (Generative AI)

❑ نوعی از هوش مصنوعی که تولید داده انجام می دهد

❑ متن: ChatGPT

❑ کد (کامپیوتری): CodeX

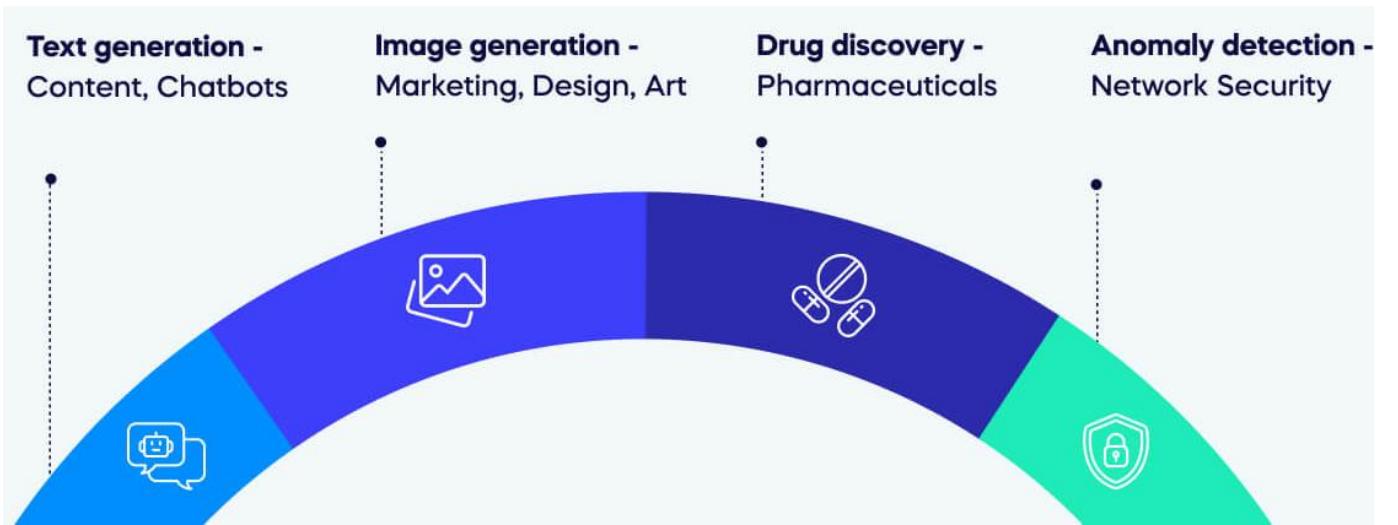
❑ صدا (گفتار): Vall-E

❑ موسیقی: SoundRaw

❑ تصویر: Dall-E

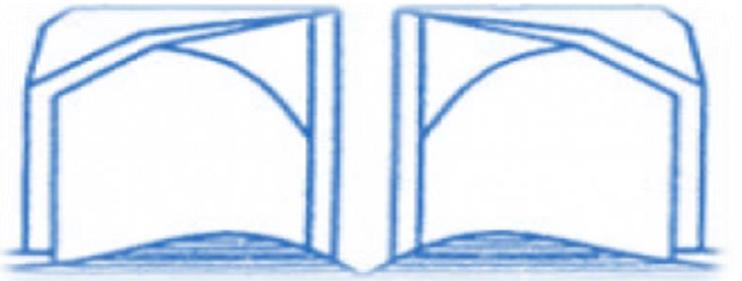
❑ ویدیو: Synthesia

... ▶





دانشگاه علوم فنی و مهندسی
دانشکده سایر های همکاری



پردازش زبان طبیعی و مدل زبانی

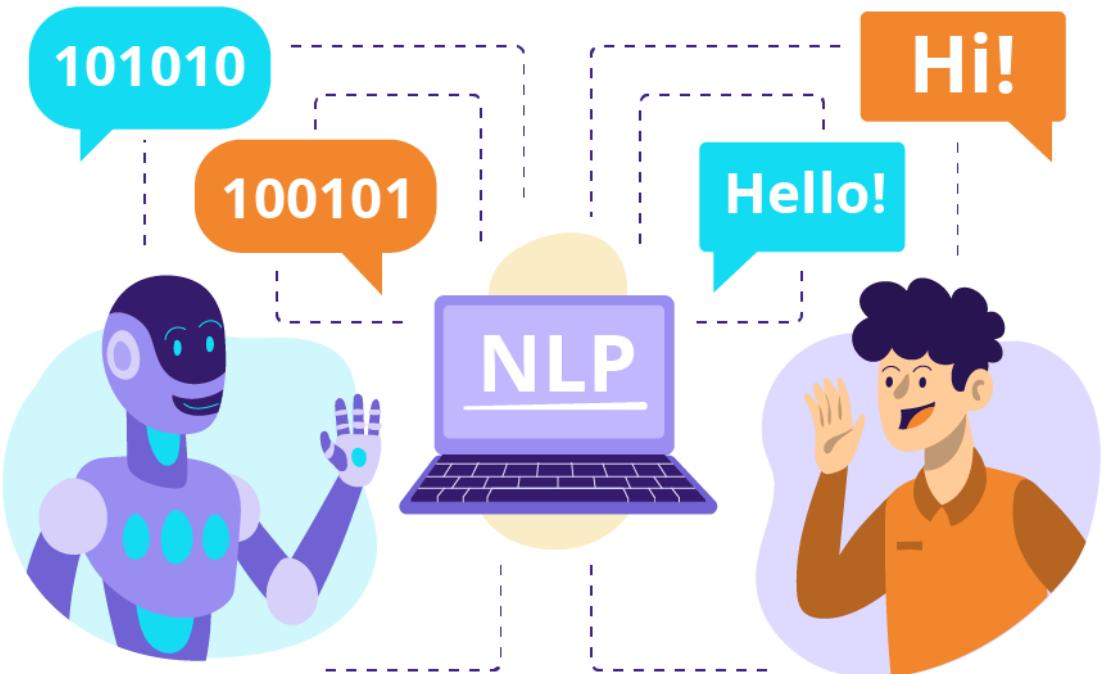
پرده نخست

زبان . . .



از حدود 3000 سال قبل از میلاد	بیش از یک میلیون سال	قدمت
نیازمند یادگیری	همه، از دو سالگی	عمومیت
کاراکتر: ض، ز، ظ، ذ	/Z/ (واج): آوا	واحد پایه
علائم سجاوندی، کارکترهای خاص	نوا، آهنگ گفتار، تپق، مکث	ویژگی‌های یکتا
کم	زیاد	تغییر در زمان
کمتر: ساختار یافته و منظم	زیاد: لهجه‌های مختلف	تنوع
	طبیعی، منعطف، بهینه، اقتصادی، تاثیرگذاری، رایج، تکامل بیشتر	برتری نسبی

پردازش زبان طبیعی (NLP) . . .

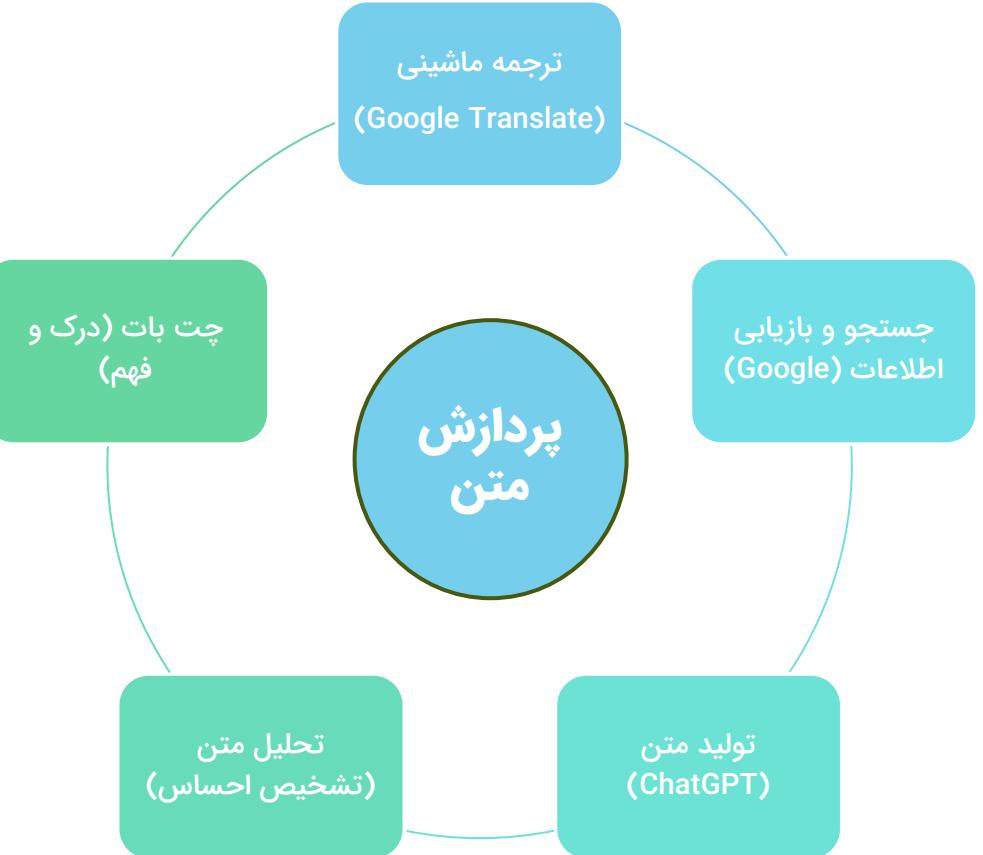
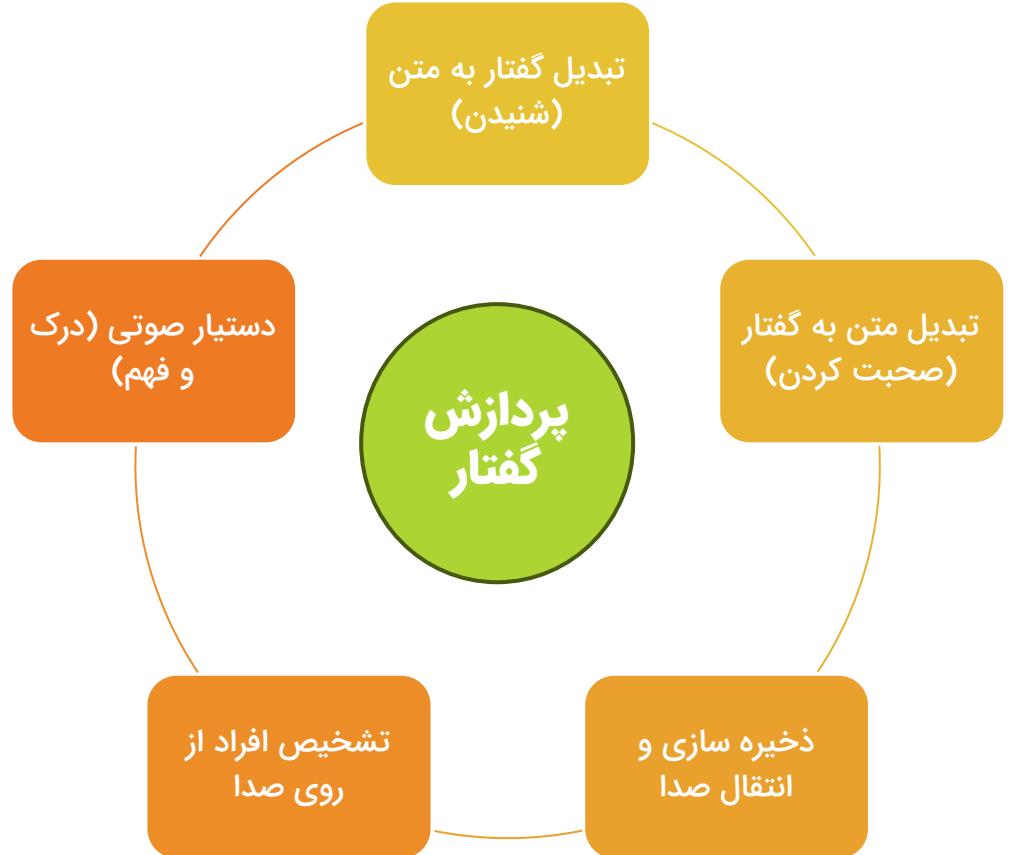


□ ایجاد توانایی های زبانی انسان برای رایانه

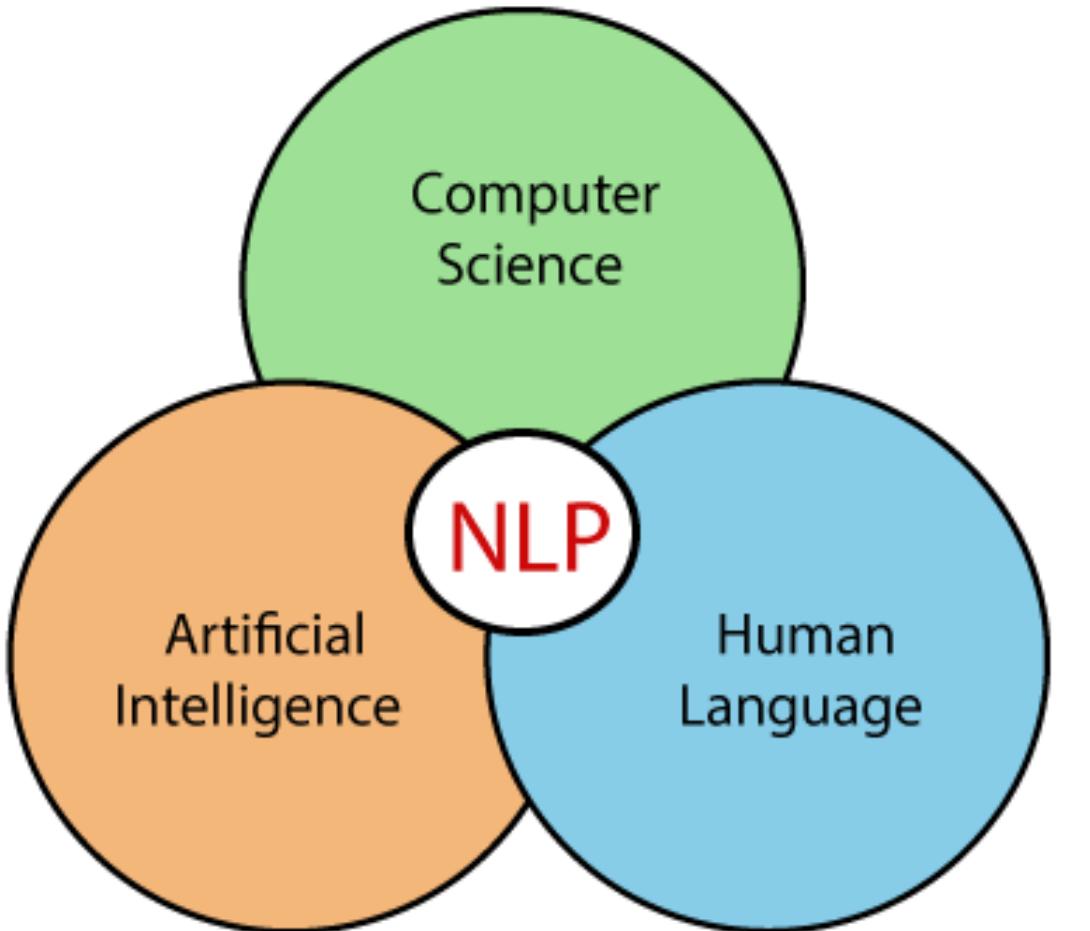
- ◀ نوشتن
- ◀ صحبت کردن
- ◀ شنیدن
- ◀ فهمیدن

□ زبان رابط اصلی بین انسان و رایانه

پردازش زبان طبیعی (NLP) . . .



پردازش زبان طبیعی (NLP) . . .



سطح پردازش زبان

کاربردشناسی (Pragmatics) و
گفتمان (Discourse)

معنای سخن در بافت های مختلف
این سخنرانی انتخاباتی است

معناشناسی (Semantics)

معنای کلمات و جملات
حیاط وحش

نحو/دستور (Syntax/Grammar)

ترکیب کلمات برای ساخت جمله/عبارت

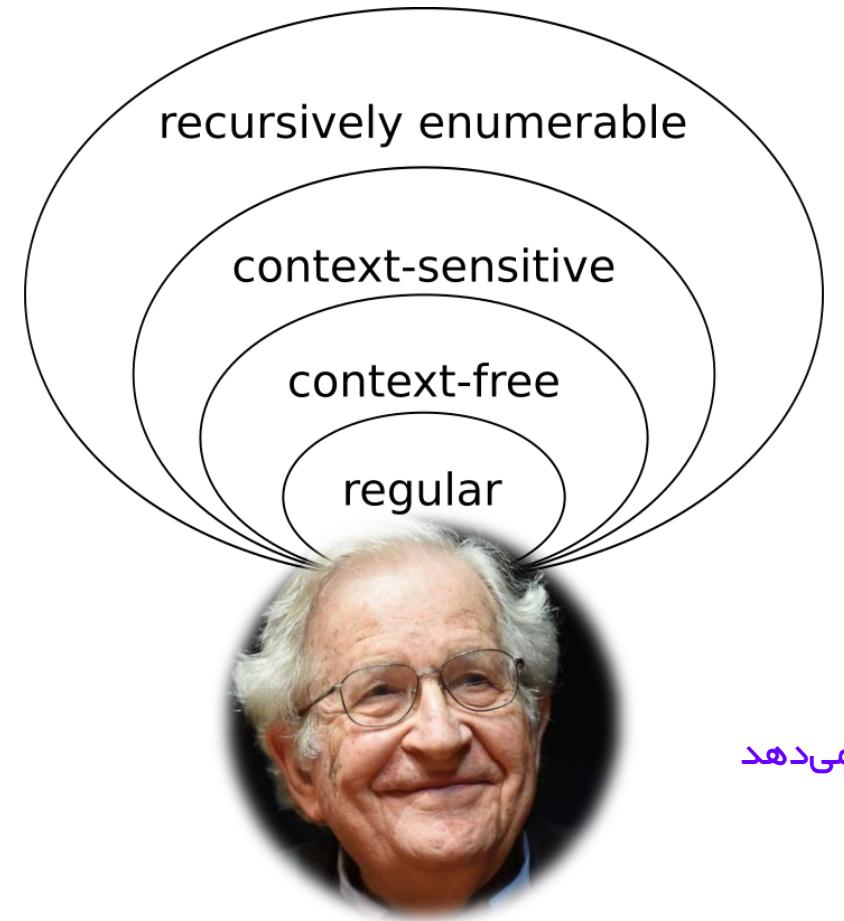
ساخت واژی (Morphology)

نحوه شکل گیری کلمات
زبانشناس = زبان + شناس

آواشناسی/واج شناسی (Phonetics/Phonology)
ظاهرشناسی (Graphology)

صداهای پایه در زبان گفتاری
شکل نمایش زبان در زبان نوشتاری (کاراکتر)

مدل سازی زبانی (Language Modeling)



□ مدل زبانی (Language Model)

◀ مدل کردن رایانه‌ای نحوه رخداد توالی کلمات در زبان

□ انواع مدل زبانی

◀ ساختاری

◀ با استفاده از تعدادی قاعده زبانی، نحوه توالی لغات را مشخص می‌کند

◀ آماری

◀ به یک دنباله از کلمات زبان مانند $W=w_1w_2\dots w_m$ یک مقدار احتمال $P(W)$ نسبت می‌دهد

● مثال: (بازیکنان تیم ملی نوجوانان جودو)

مدل سازی زبانی . . .



□ سطوح مدل سازی زبانی

- ◀ واژگانی محلی
- ◀ نحوی
- ◀ معنایی

□ کاربردهای مدل زبانی

- ◀ پیش‌بینی کلمات
- ◀ بازشناسی گفتار
- ◀ درک زبان طبیعی
- ◀ ترجمه ماشینی
- ◀ بازشناسی نویسه‌های نوری
- ... ▶

مدل سازی زبانی آماری . . .

□ چرا مدل زبانی آماری؟

◀ انتساب مقدار احتمال به یک عبارت/جمله

◀ کاربرد در ترجمه

$P(\text{high winds tonight}) > P(\text{large winds tonight})$ ◄

◀ کاربرد در خطایاب

$P(\text{about fifteen minutes from}) > P(\text{about fifteen minuets from})$ ◄

◀ کاربرد در تشخیص گفتار

$P(\text{I saw a van}) >> P(\text{eyes awe of an})$ ◄

مدل سازی زبانی آماری . . .

□ هدف

- ◀ محاسبه احتمال رخداد یک جمله (دباله‌ای از کلمات)

$$P(W) = P(w_1, w_2, w_3, w_4, w_5 \dots w_m)$$

(بازیکنان تیم ملی نوجوانان جودو)

- ◀ محاسبه احتمال کلمه بعدی (با این فرض که کلمات قبلی را داریم)

$$P(w_5 | w_1, w_2, w_3, w_4)$$

(بازیکنان تیم ملی نوجوانان اجودو)

□ مدل زبانی (LM)

- ◀ محاسبه $P(W)$ یا $P(w_m | w_1, w_2 \dots w_{m-1})$



مدل سازی زبانی آماری . . .

$$P(\text{the} \mid \text{its water is so transparent that}) = \frac{\text{Count}(\text{its water is so transparent that the})}{\text{Count}(\text{its water is so transparent that})}$$

□ نحوه محاسبه احتمال‌ها؟

◀ شمارش روی پیکره متنی

◀ این کار عملی نیست! چون ...

◀ تعداد جملات مختلف بی‌نهایت است

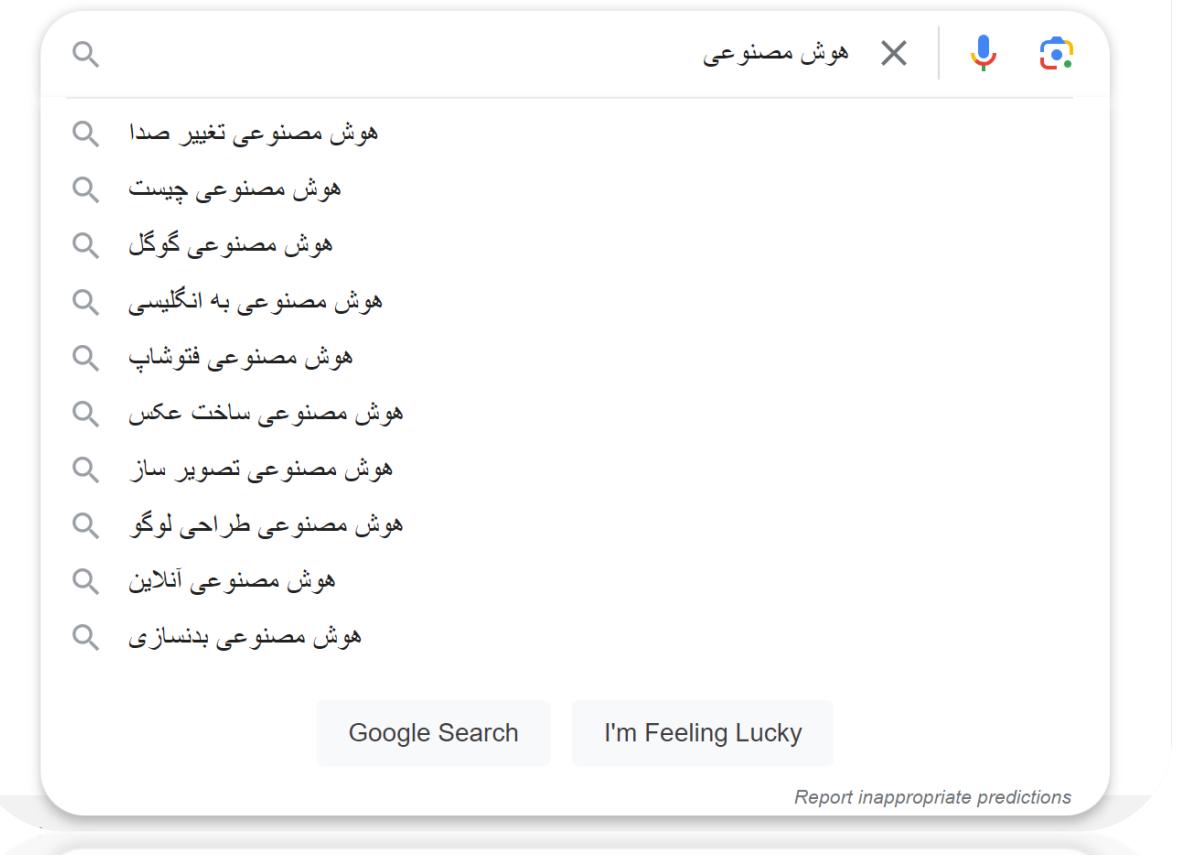
◀ هیچ مجموعه داده‌ای برای شمارش آن موجود نیست

◀ فرض مارکوف (Markov)

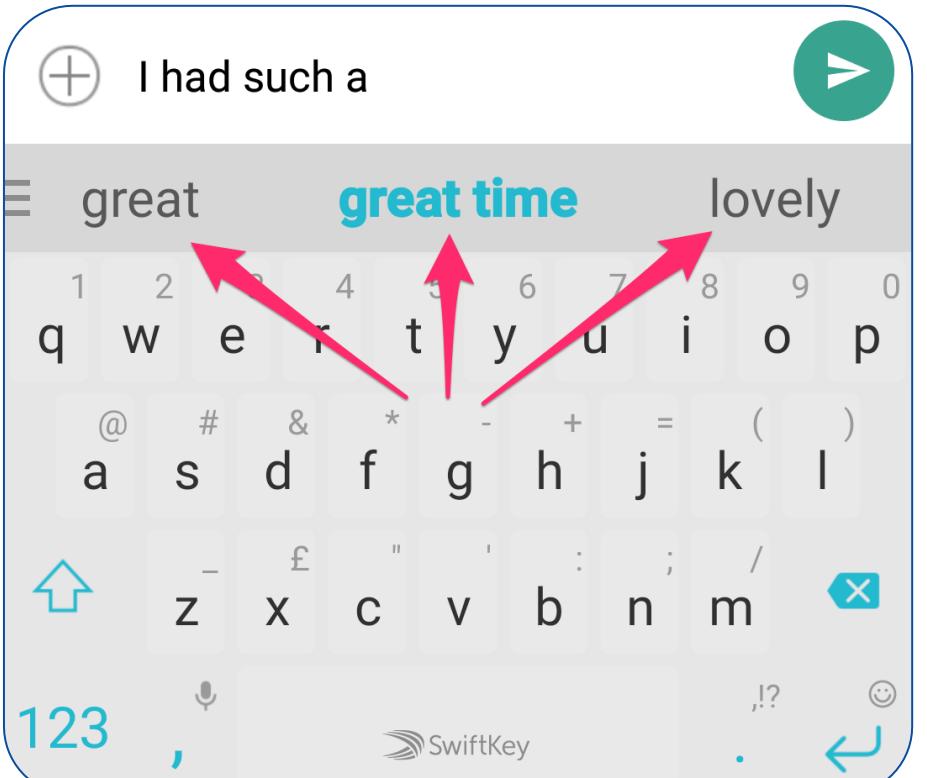
◀ مرتبه یک: فقط وابسته به یک متغیر (کلمه) قبل

◀ مرتبه دو: فقط وابسته به دو متغیر (کلمه) قبل

مدل‌سازی زبانی آماری: پیش‌بینی کلمه بعدی



مدل سازی زبانی آماری: پیش بینی کلمه بعدی



... N-Gram: مدل سازی زبانی آماری

□ ساده ترین حالت: در نظر نگرفتن وابستگی به کلمات قبل

$$P(W) = P(w_1 w_2 \cdots w_m) \cong \prod_{i=1}^m P(w_i | w_{i-(k)} \cdots w_{i-1}) \cong \prod_{i=1}^m P(w_i) = P(w_1) P(w_2) P(w_3) \cdots P(w_m)$$

k=0
k=N-1 → N=1

◀ مثال

◀ (جودو)P(نوجوانان)P(ملی)P(تیم)P(بازیکنان)= (بازیکنان تیم ملی نوجوانان جودو)

Mono-Gram
UniGram
1-Gram

$$P_{monogram}(w_i) = \frac{Count(w_i)}{Count(AllWords)} = \frac{N(w_i)}{N_{total}}$$

◀ مثال

◀ در متنی با ۱۰۰۰۰ کلمه، تعداد تکرار کلمه «ملی» ۱۰ باشد، آنگاه $P=10/10,000=0.001$

... N-Gram: مدل سازی زبانی آماری

□ مدل دوتایی (Bi-Gram): در نظر گرفتن یک کلمه قبل از هر کلمه - وابستگی دو کلمه‌ای

$$P(w_1 w_2 \dots w_m) \cong \prod_{i=1}^m P(w_i | w_{i-(k)} \dots w_{i-1}) \cong \prod_{i=1}^m P(w_i | w_{i-1}) = P(w_1) P(w_2 | w_1) P(w_3 | w_2) \dots P(w_m | w_{m-1})$$

N=2

$$P(w_i | w_1 w_2 \dots w_{i-1}) \approx P(w_i | w_{i-1})$$

يعني

مثال

= (بازیکنان تیم ملی نوجوانان جودو) p

(نوجوانان|جودو) P × (ملی|نوجوانان) P × (تیم|ملی) P × (بازیکنان|تیم) P × (بازیکنان)

$$P_{bigram}(w_j | w_i) = \frac{N(w_i w_j)}{N(w_i)}$$

محاسبه دوتایی

◀ برای محاسبه (ملی|نوجوانان) P، تعداد تکرار «ملی نوجوانان» تقسیم بر تعداد تکرار «ملی»

● مثال: اگر تعداد «ملی نوجوانان» ۵۰ باشد و تعداد «ملی» ۱۰۰ باشد، (ملی|نوجوانان) P برابر با $50/100$ یا ۰.۵ است

... N-Gram: مدل سازی زبانی آماری

 $N=3$ ◀

- مدل سه تایی (Tri-Gram): در نظر گرفتن دو کلمه قبل از هر کلمه - وابستگی سه کلمه ای

$$P(W) = P(w_1)P(w_2 | w_1)P(w_3 | w_1w_2)P(w_4 | w_2w_3)\cdots P(w_m | w_{m-2}w_{m-1})$$

- امکان توسعه به تعداد بیشتر: ۴-Gram، ۳-Gram و ...

- مقدار معمول: ۵-Gram تا ۲-Gram

- در عمل وابستگی زبانی بیشتر از ۵ کلمه است

➤ “The computer which I had just put into the machine room on the fifth floor crashed”

مدل سازی زبانی آماری . . .

□ آمارها سخن می‌گویند

◀ مدل کردن واقعیات (از Berkeley Restaurant Project)

$$P(\text{english} \mid \text{want}) = .0011 \triangleleft$$

$$P(\text{chinese} \mid \text{want}) = .0065 \triangleleft$$

◀ معنی: درخواست برای غذای چینی شش برابر انگلیسی است !

◀ مدل کردن نحو

$$P(\text{to} \mid \text{want}) = .66 \triangleleft$$

$$P(\text{eat} \mid \text{to}) = .28 \triangleleft$$

$$P(\text{food} \mid \text{to}) = 0 \triangleleft$$

$$P(\text{want} \mid \text{spend}) = 0 \triangleleft$$

$$P(\text{i} \mid \langle s \rangle) = .25 \triangleleft$$

◀ مدل کردن معنا

$$P(\text{about fifteen minutes from}) > P(\text{about fifteen minuets from}) \triangleleft$$

مدل سازی زبانی آماری: قدمتی به بلندای تاریخ هوش مصنوعی

□ ایده‌ای قدیمی اما تحول آفرین
 کlad شانون، ۱۹۴۸ ◀

3. THE SERIES OF APPROXIMATIONS TO ENGLISH

To give a visual idea of how this series of processes approaches a language, typical sequences in the approximations to English have been constructed and are given below. In all cases we have assumed a 27-symbol "alphabet," the 26 letters and a space.

1. Zero-order approximation (symbols independent and equiprobable).

XFOML RXKHRJFFUJ ZLPWCFWKCYJ FFJEYVKCQSGHYD QPAAMKBZAACIBZL-HJQD.

2. First-order approximation (symbols independent but with frequencies of English text).

OCRO HLI RGWR NMIELWIS EU LL NBNSEBYA TH EEI ALHENHTTPA OOBTTVA NAH BRL.

3. Second-order approximation (digram structure as in English).

ON IE ANTSOUTINYS ARE T INCTORE ST BE S DEAMY ACHIN D ILONASIVE TU-COOWE AT TEASONARE FUSO TIZIN ANDY TOBE SEACE CTISBE.

4. Third-order approximation (trigram structure as in English).

IN NO IST LAT WHEY CRATICT FROURE BIRS GROCID PONDENOME OF DEMONSTURES OF THE REPTAGIN IS REGOACTIONA OF CRE.

5. First-order word approximation. Rather than continue with tetragram, . . . , n -gram structure it is easier and better to jump at this point to word units. Here words are chosen independently but with their appropriate frequencies.

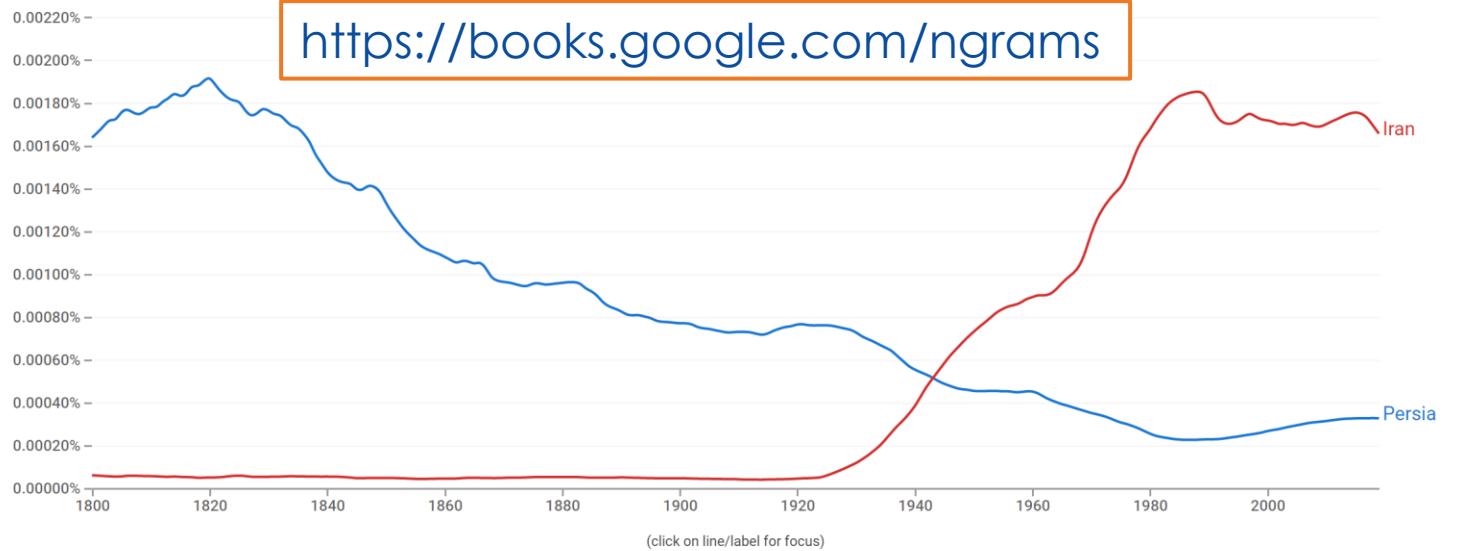
REPRESENTING AND SPEEDILY IS AN GOOD APT OR COME CAN DIFFERENT NATURAL HERE HE THE A IN CAME THE TO OF TO EXPERT GRAY COME TO FURNISHES THE LINE MESSAGE HAD BE THESE.

6. Second-order word approximation. The word transition probabilities are correct but no further structure is included.

THE HEAD AND IN FRONTAL ATTACK ON AN ENGLISH WRITER THAT THE CHARACTER OF THIS POINT IS THEREFORE ANOTHER METHOD FOR THE LETTERS THAT THE TIME OF WHO EVER TOLD THE PROBLEM FOR AN UNEXPECTED.

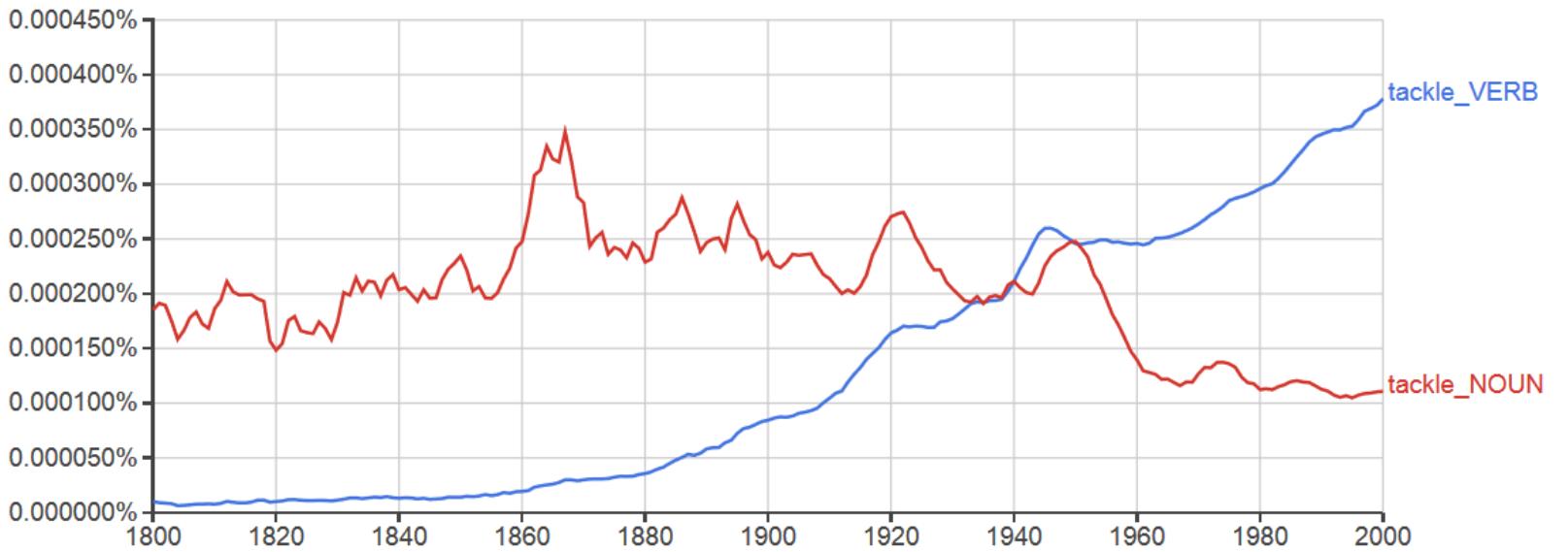
مدل سازی زبانی آماری: . . . Google Book N-grams

- استخراج شده از کتابهای کتابخانه گوگل (<http://books.google.com>)
- ◀ از سال ۱۸۰۰ تا ۲۰۱۲
- ◀ حاوی ۵ میلیون کتاب (برای ۸ زبان)
- ◀ کتابخانه گوگل در ۲۰۱۳ حاوی ۳۰ میلیون کتاب بوده است (تخمین وجود ۱۳۰ میلیون کتاب در دنیا)
- ◀ نسخه ۲ (در سال ۲۰۱۲)



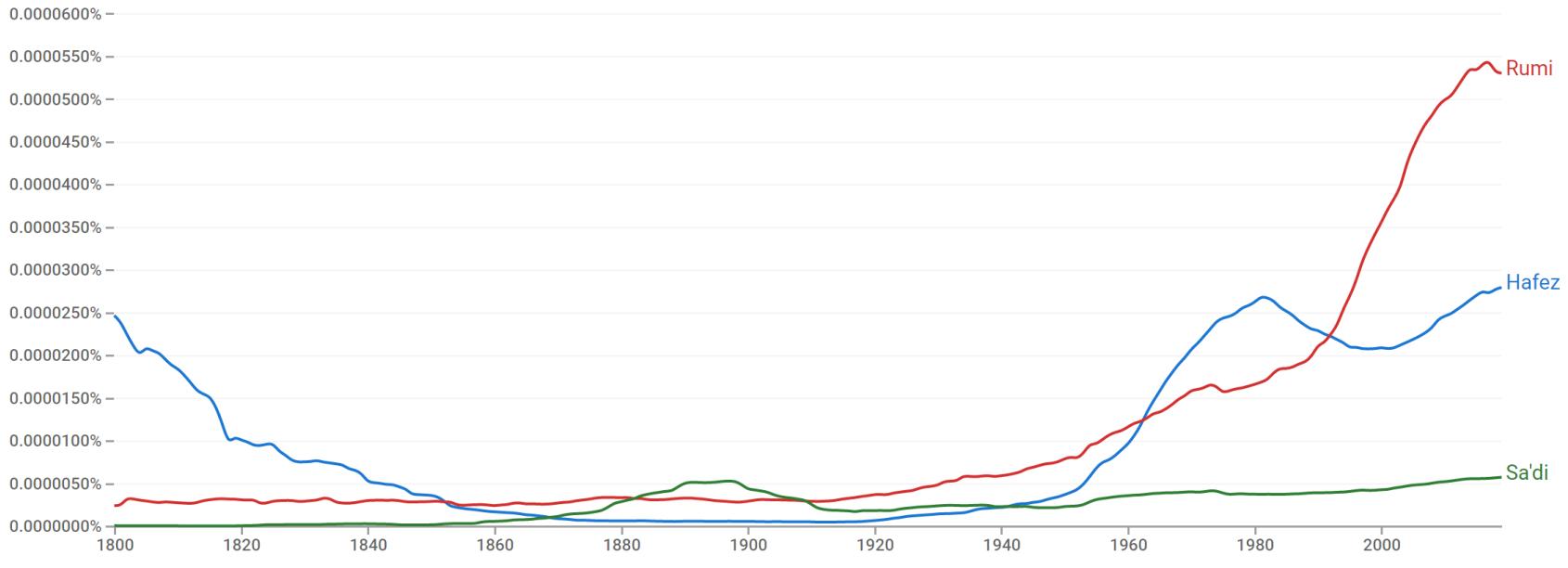
مدل سازی زبانی آماری: . . . Google Book N-grams

- بررسی تغییرات زبان و واژه ها از سال ۱۸۰۰ تا ۲۰۱۲
- ◀ استفاده از کلمه tackle به عنوان اسم یا فعل



مدل سازی زبانی آماری: Google Book N-grams

□ شاعر ایرانی شناخته شده در خارج از ایران!

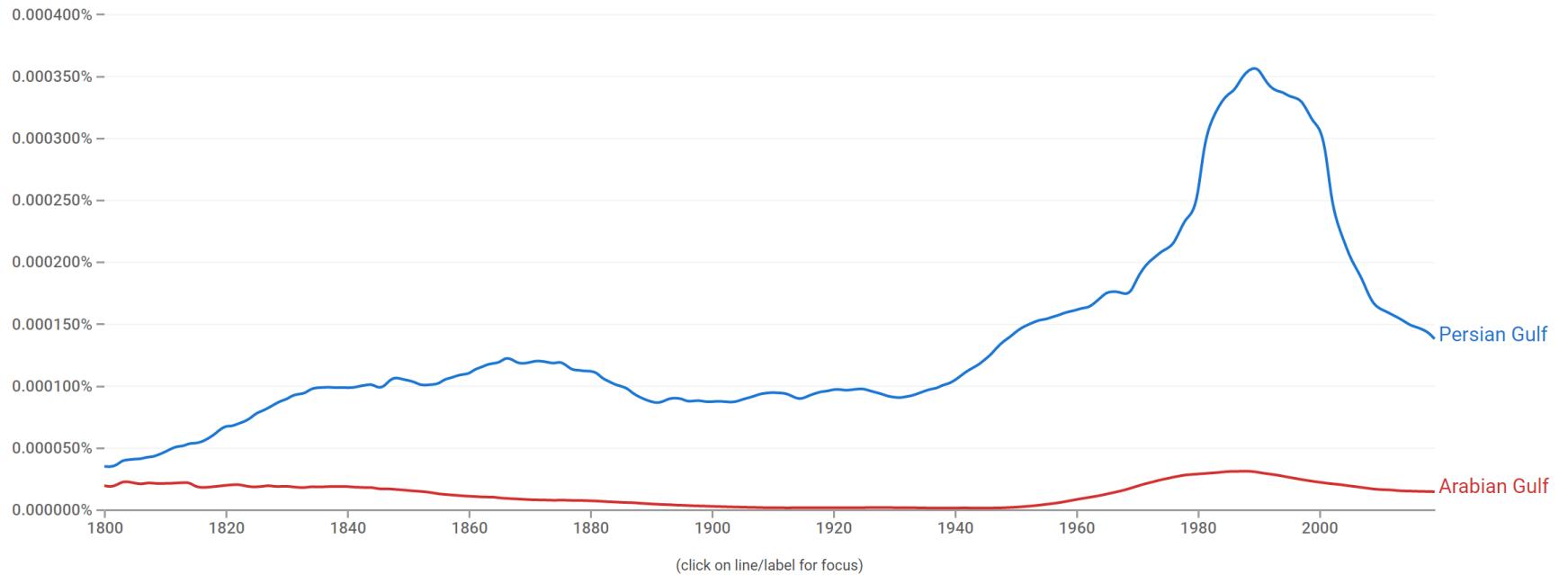


Search in Google Books

■ Hadi Veisi (h.veisi@ut.ac.ir)

مدل سازی زبانی آماری: Google Book N-grams

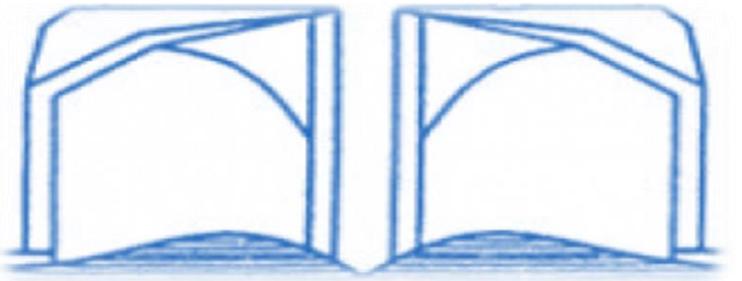
□ بدون شرح!



Search in Google Books



دانشگاه علوم فناوری های میان رشته ای
دانشکده سازمان های به شرمند



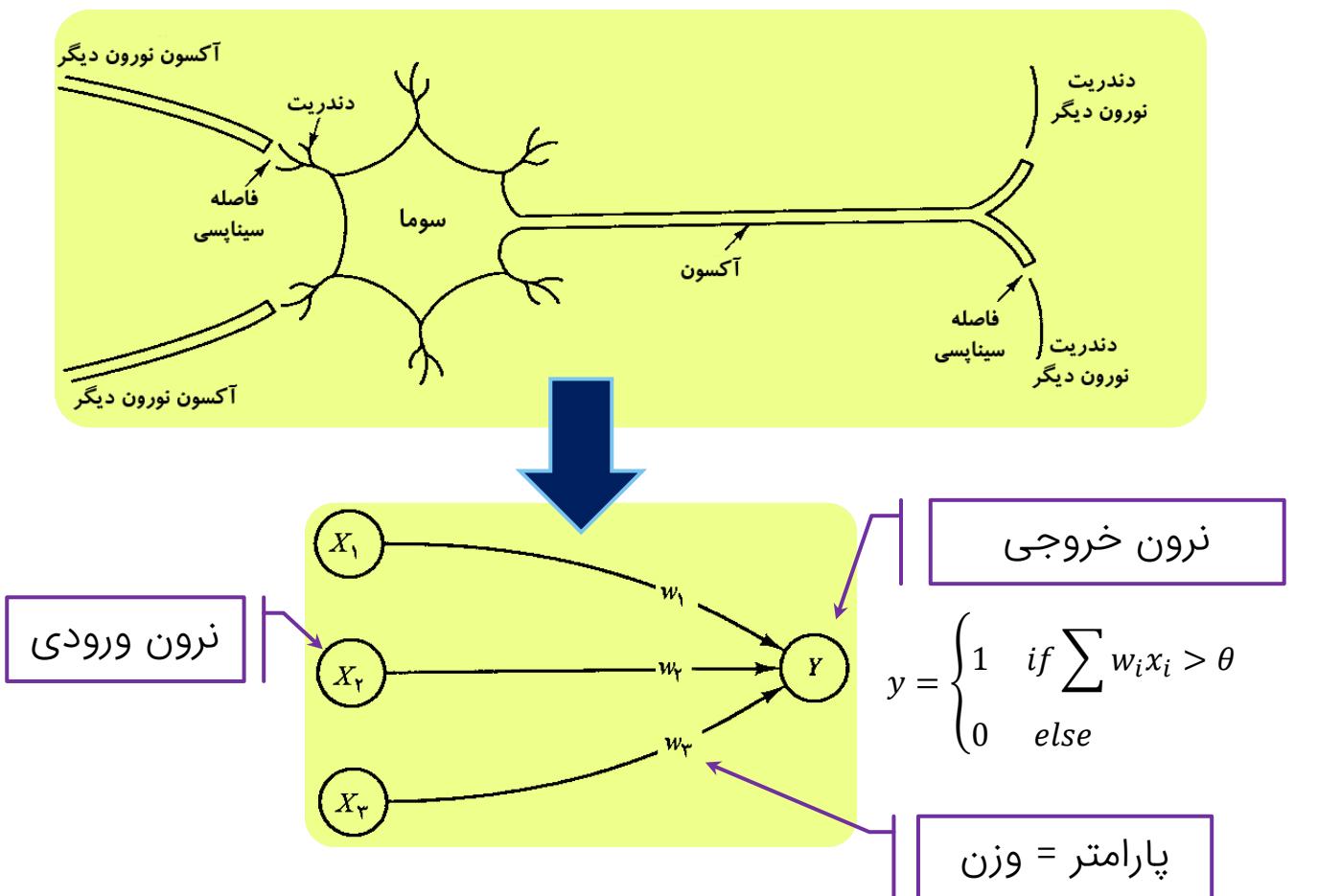
پرده دوم شبکه عصبی مصنوعی

شبکه عصبی مصنوعی



- مغز انسان = شبکه‌ای بزرگ از عصب‌ها (نرون‌ها)
 - ◀ ۱۰۰.۰۰۰.۰۰۰.۰۰۰
 - ◀ ۱۰.۰۰۰ اتصال (سیناپس) برای هر نرون
- ◀ حدود هزار تریلیون اتصال (سیناپس)
- ◀ بزرگترین مدل ساخته شده مصنوعی (Claude3): ۲ تریلیون پارامتر
- شبکه عصبی مصنوعی = شبیه‌سازی شبکه عصبی طبیعی

شبکه عصبی مصنوعی: مدل سازی



شبکه عصبی مصنوعی: مثال ساده



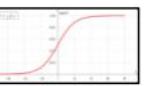
image2vector

$$\begin{pmatrix} 255 \\ 231 \\ \dots \\ 94 \\ 142 \end{pmatrix}$$
 $/255$
 $x_0^{(i)}$
 $/255$
 $x_1^{(i)}$
 \dots
 \dots
 $/255$
 $x_{12286}^{(i)}$
 $/255$
 $x_{12287}^{(i)}$
 w_0
 w_1
 w_{12286}
 w_{12287}
 $w^T x^{(i)} + b$
 σ

"it's a cat"

0.73 > 0.5

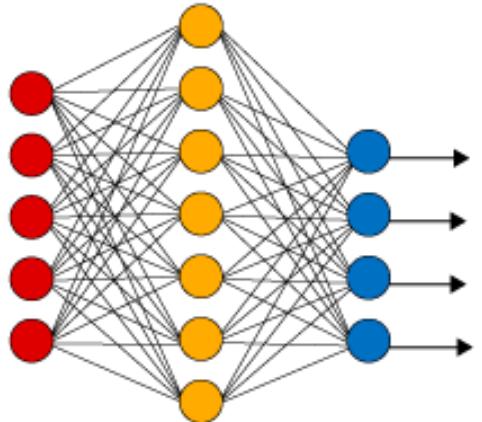
0.73



شبکه عصبی عمیق (یادگیری عمیق)

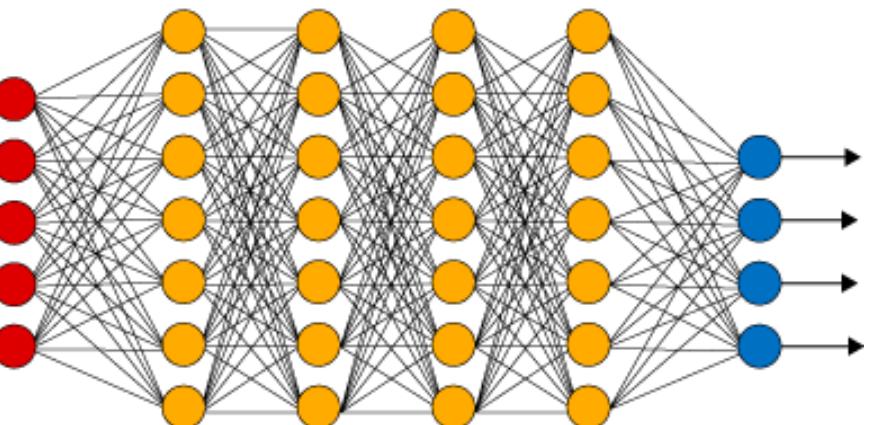
- افزایش تعداد لایه های مخفی (میانی)

Simple Neural Network



● Input Layer

Deep Learning Neural Network

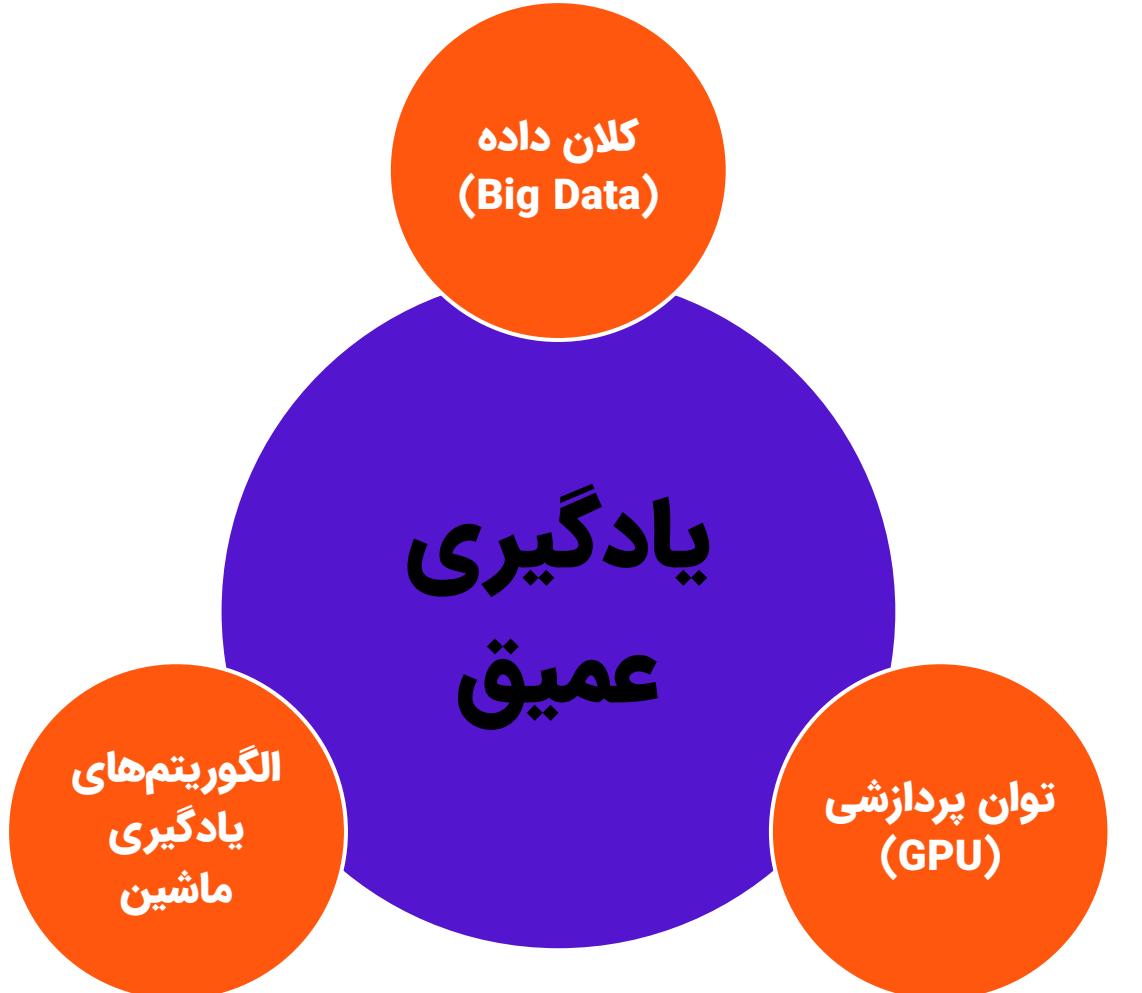


● Hidden Layer

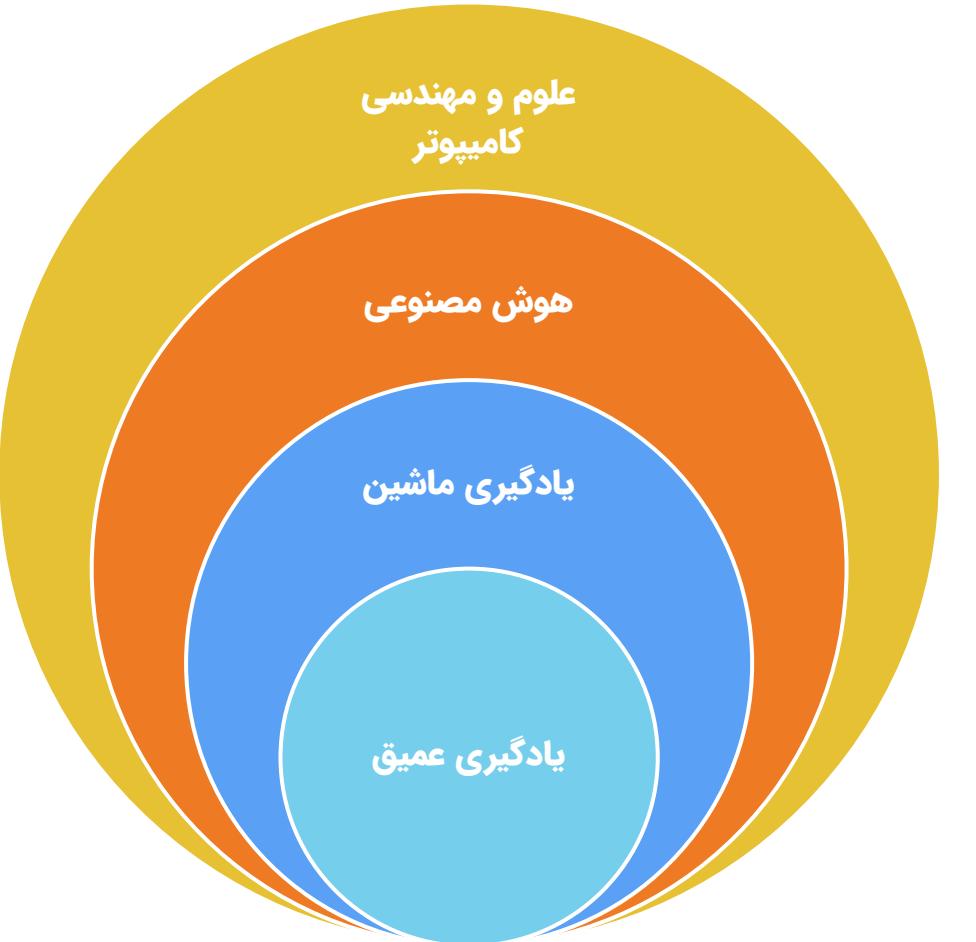
● Output Layer

- مدل پیچیده = تعداد پارامترهای زیاد = توان محاسباتی بالا

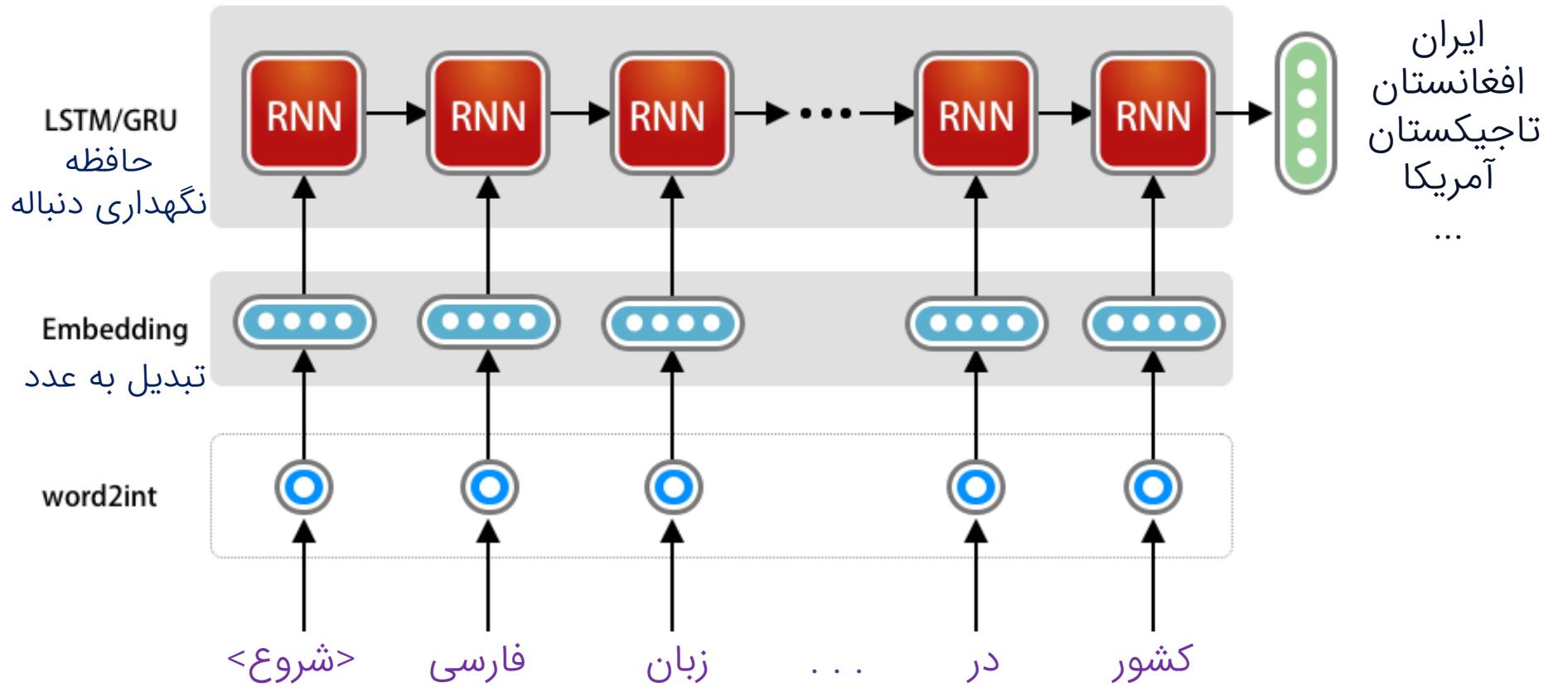
نیازمندی های یادگیری عمیق



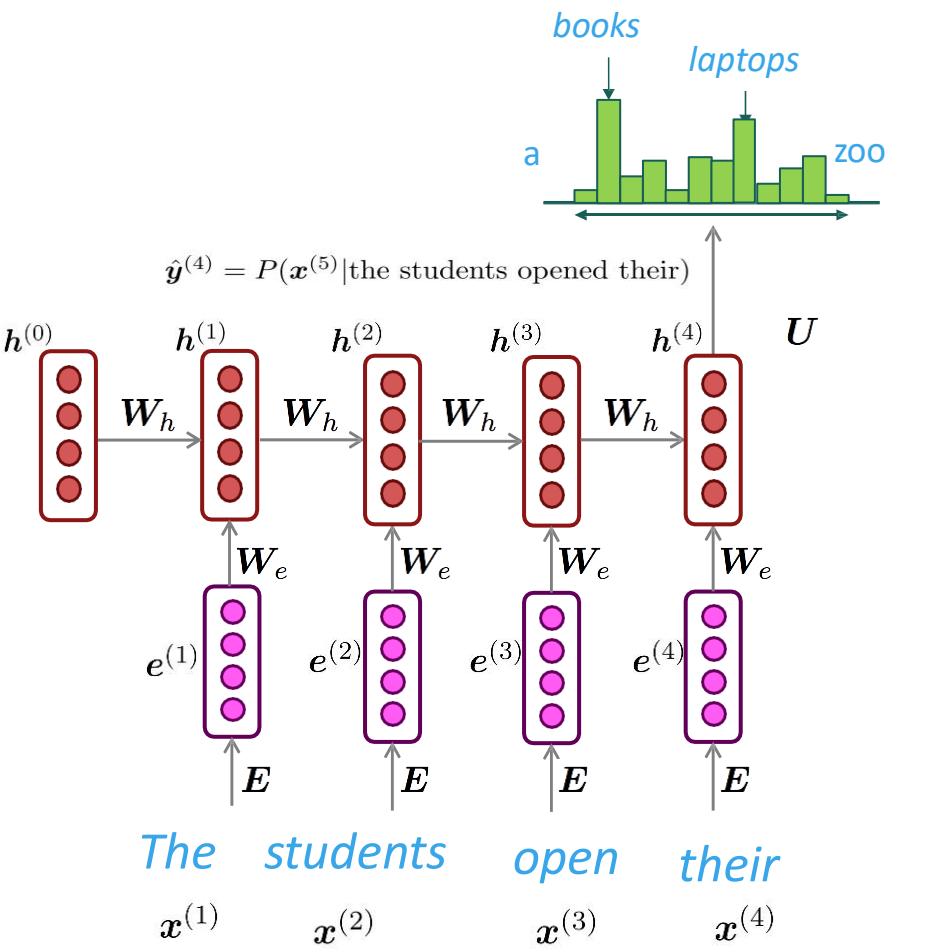
یادگیری عمیق و هوش مصنوعی



RNN: Recurrent Neural Network شبکه بازگشتی



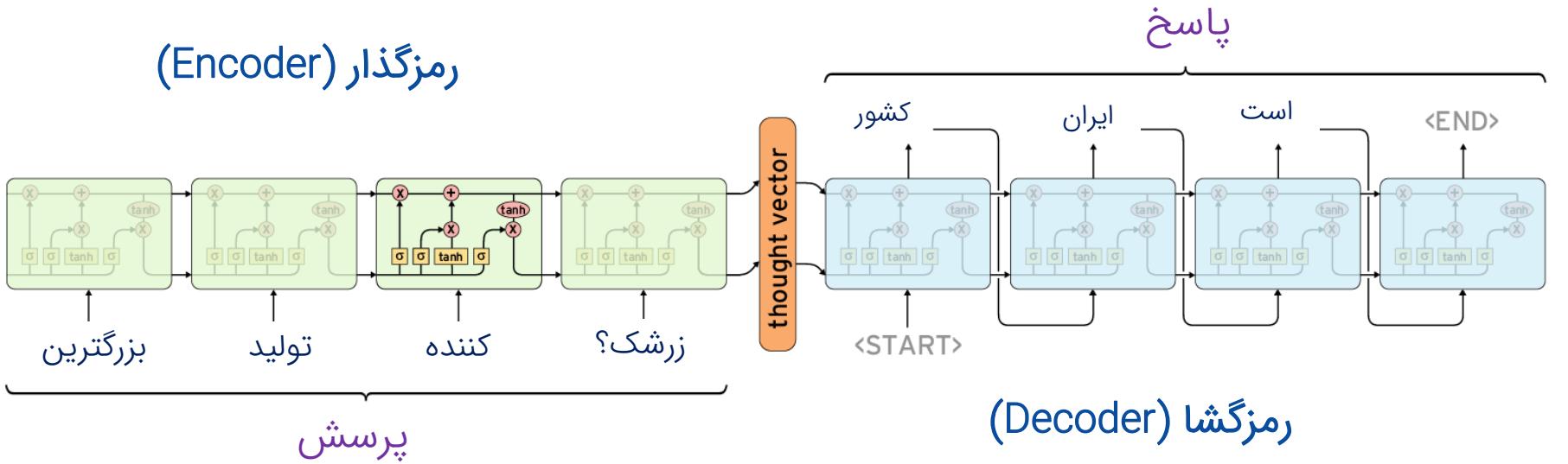
شبکه بازگشتی RNN برای مدل زبانی



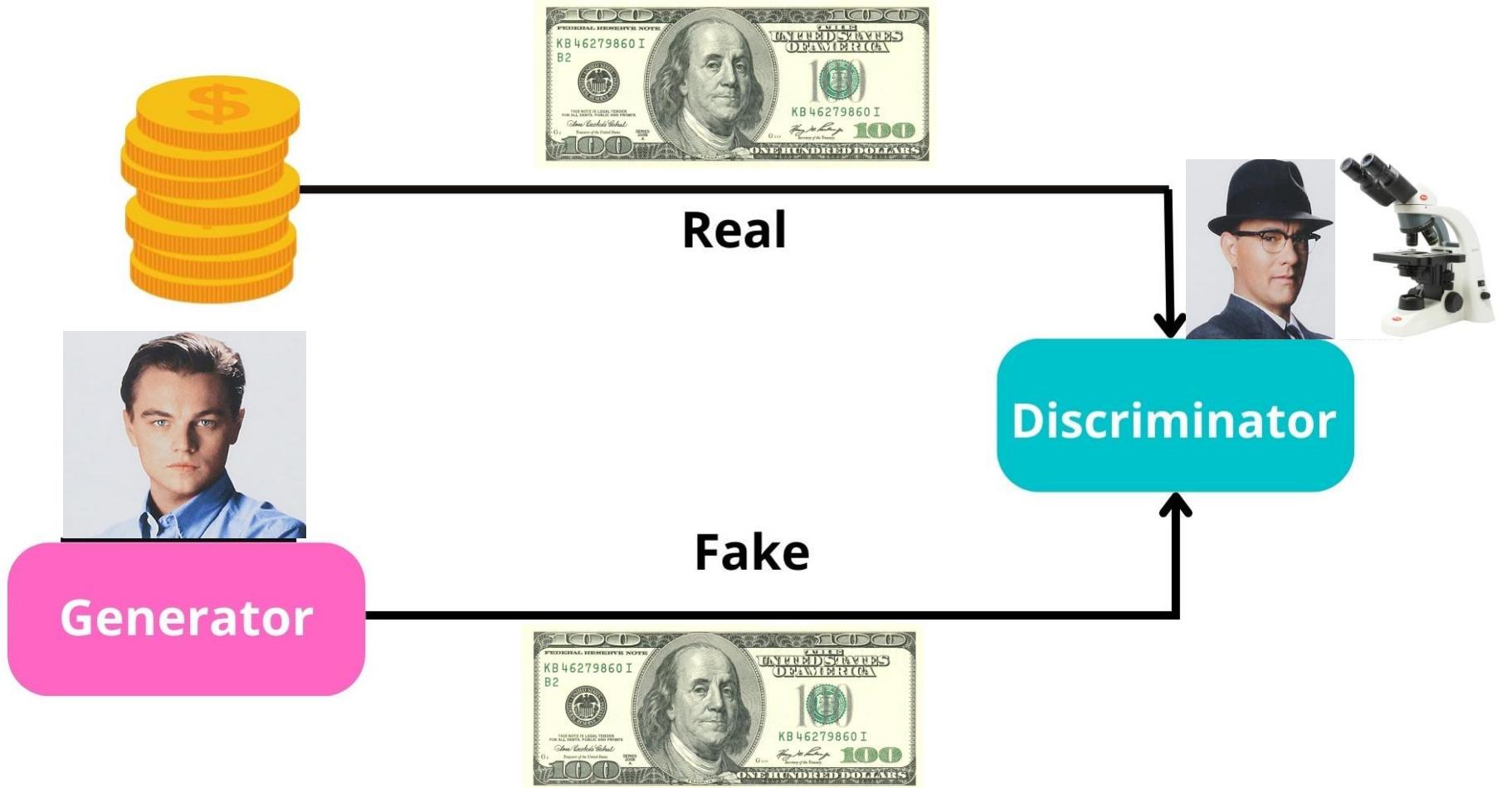
شبکه RNN: معماری رمزگذار-رمزگشا

Encoder-Decoder □ معماری

- ◀ گفتگو (چتبات)
- ◀ ترجمه ماشینی
- ... ◀



شبکه GAN: Generative Adversarial Network



شبکه GAN: تولید داده با کیفیت



2014

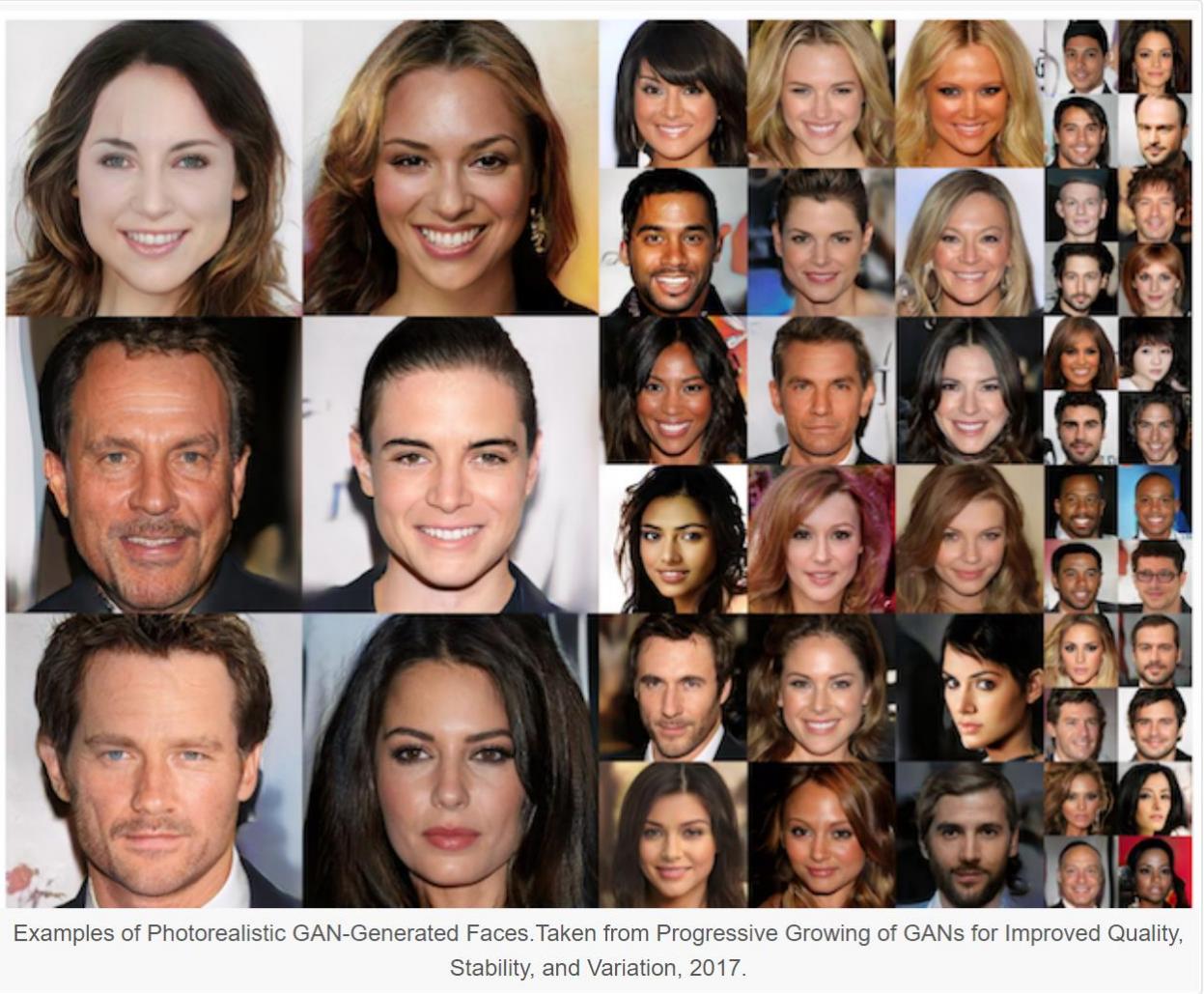
2015

2016

2017

Example of the Progression in the Capabilities of GANs from 2014 to 2017. Taken from The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation, 2018.

شبکه GAN: آغاز نگرانی



شبکه GAN: سایر کاربردها



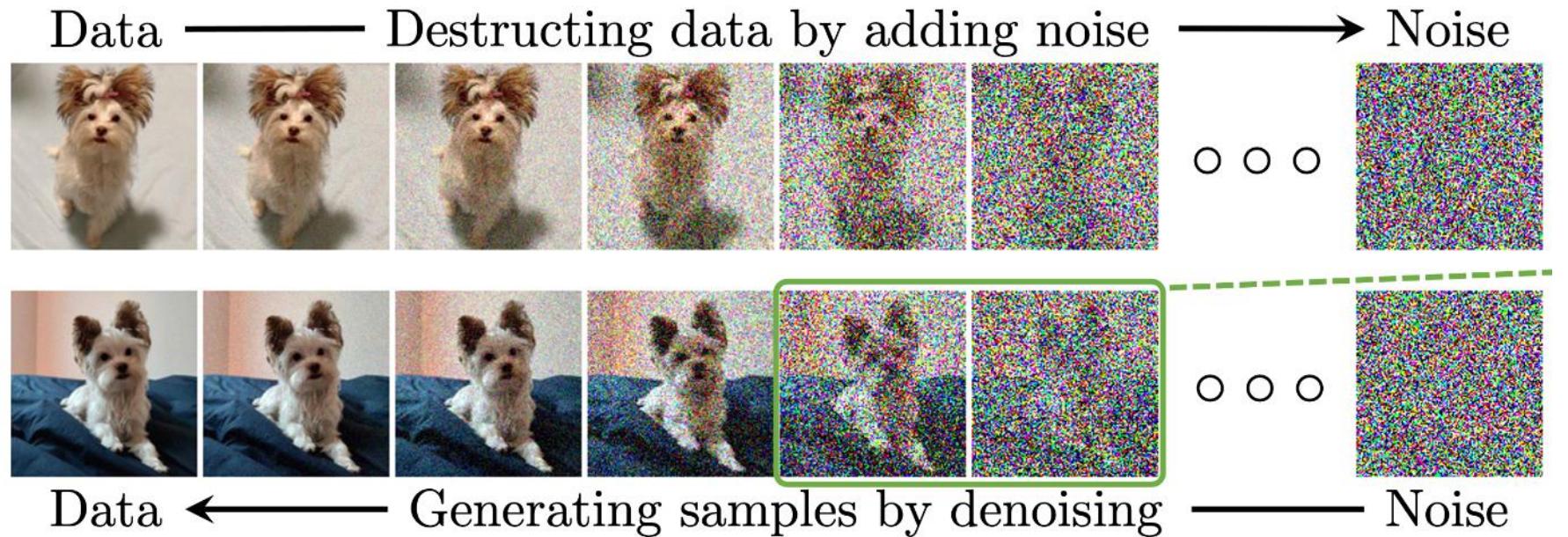
- تبدیل چهره مرد و زن به همدیگر
- تغییر سن در تصویر چهره
- تبدیل تصویر چهره به آواتار
- افزایش کیفیت تصویر
- رنگی کردن تصاویر سیاه و سفید
- تولید تصاویر سه بعدی
- ترمیم تصویر
- پیش‌بینی فریم‌ها بعدی در ویدئو
- ...

شبکه GAN: تولید متن (شعر فردوسی)

که ای نامور شاه را دل ز جای
 که از من بدین داستان بگذرد
 به دل گفت کای در جهان کینه جوی
 به نزدیک او با سپاه امدادند
 بگوییم که باشد به دل پر ز خون
 زمین را همی داشت اندر نهان
 سوی شاه گردن فراوان براند
 ز توران به درگاه شاه جهان
 بیاراست لشکر برآمد ز راه
 که ای پهلوان سر به سر بر زمین
 به پیش سپاهش به تخت و کلاه
 که بیژن نیاید به پیش سپاه
 که از کار او در جهان کارزار
 چو خورشید تابان به درگاه شاه

- به پیش سپاه اندر امد ز جای
- به پیران چنین گفت کای پرخد
- بدو گفت بهرام کای نامجوی
- برآمد ز ایوان شاه امدادند
- به کار اندر امد ز دریای خون
- بدو گفت بر من بدی در جهان
- بدین گونه با او سپه را بخواند
- به روی اندر امد ز من در جهان
- چو بشنید ازو بازگشتن به راه
- بدو گفت بازار ایران زمین
- به داد و به دیدار او شد سپاه
- بران کوه بر شد ز هر سو سپاه
- به پیش اندر امد سوی کارزار
- به ایوان بیاراست گرد سیاه

مدل Diffusion



Sohl-Dickstein, Jascha; Weiss, Eric; Maheswaranathan, Niru; Ganguli, Surya (2015-06-01). "Deep Unsupervised Learning using Nonequilibrium Thermodynamics". Proceedings of the 32nd International Conference on Machine Learning. PMLR. 37: 2256–2265.

Ho, Jonathan; Jain, Ajay; Abbeel, Pieter (2020). "Denoising Diffusion Probabilistic Models". Advances in Neural Information Processing Systems. Curran Associates, Inc. 33: 6840–6851.

تولید تصویر: Stable Diffusion

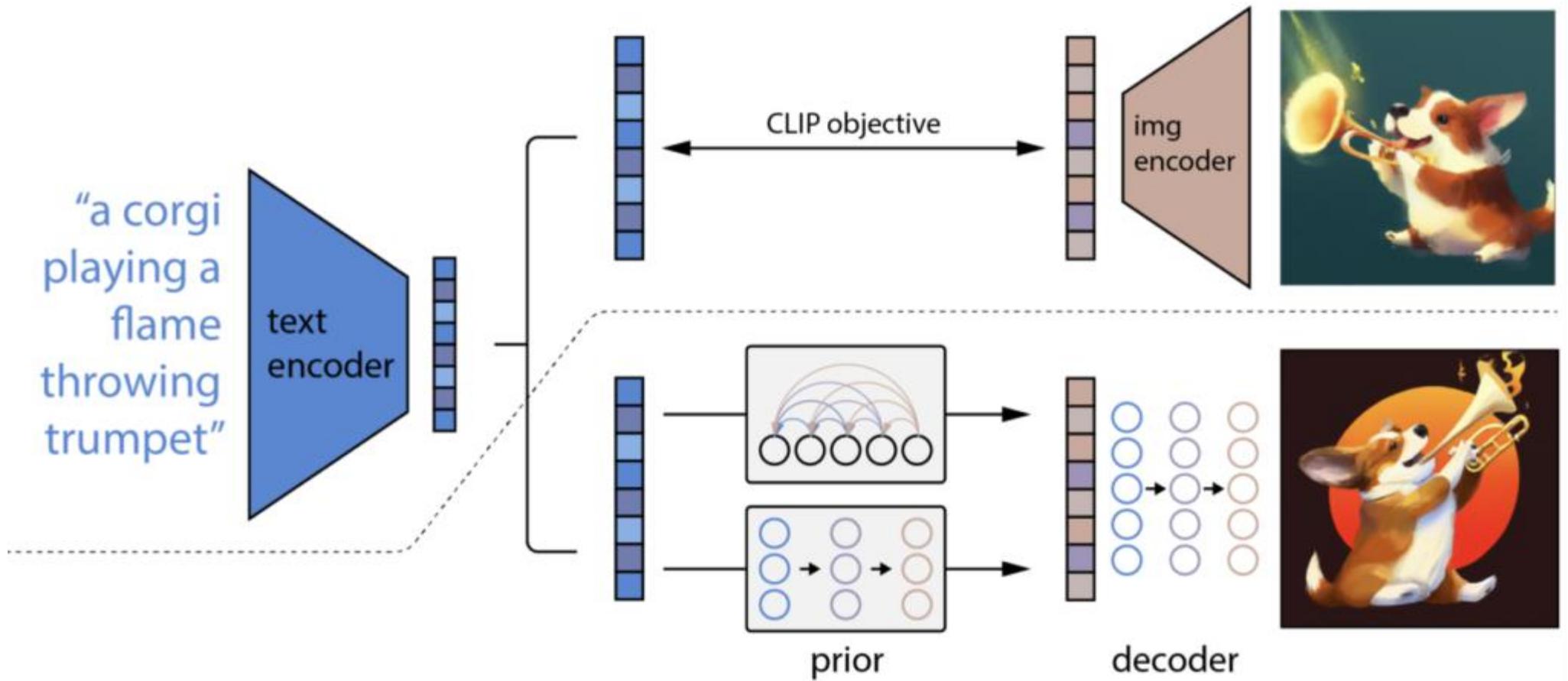


□ تبدیل متن به عکس

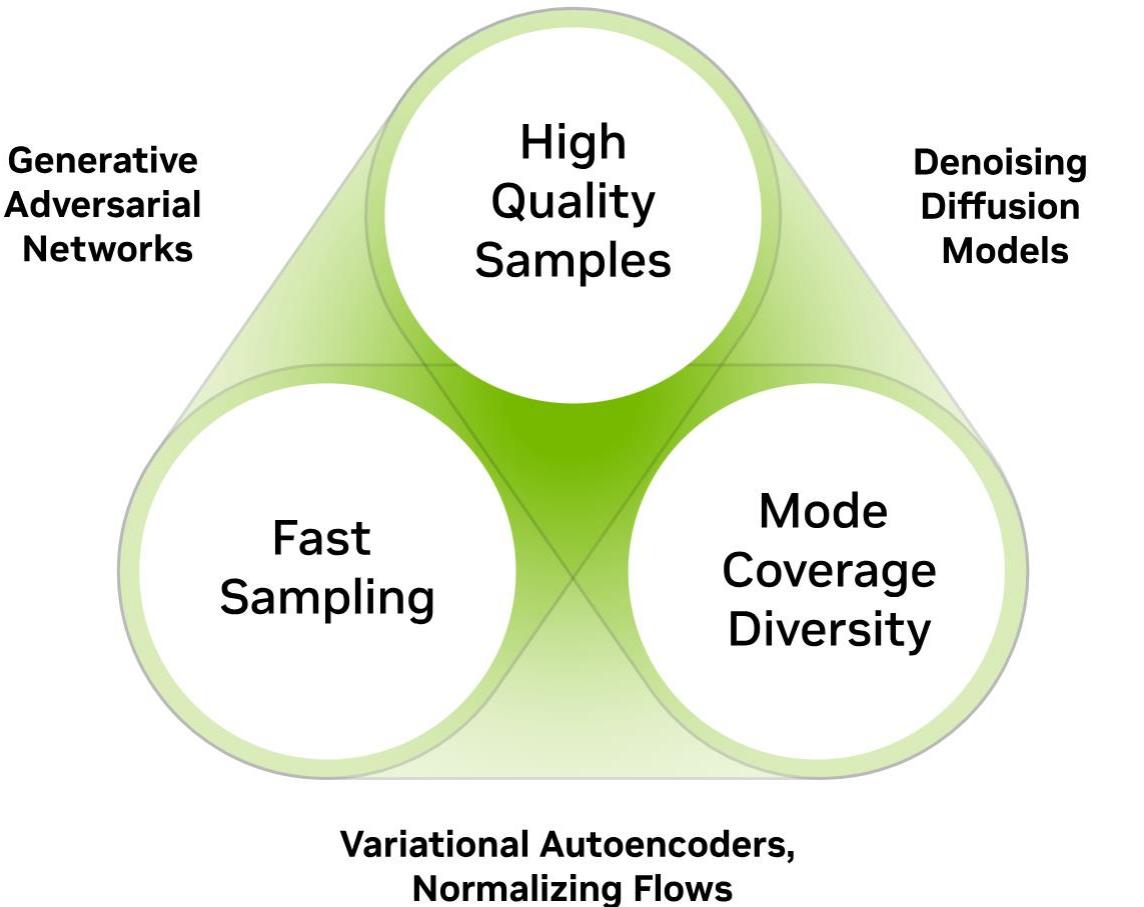
مارک زاکربرگ در حال
کباب درست کردن

■ Hadi Veisi (h.veisi@ut.ac.ir)

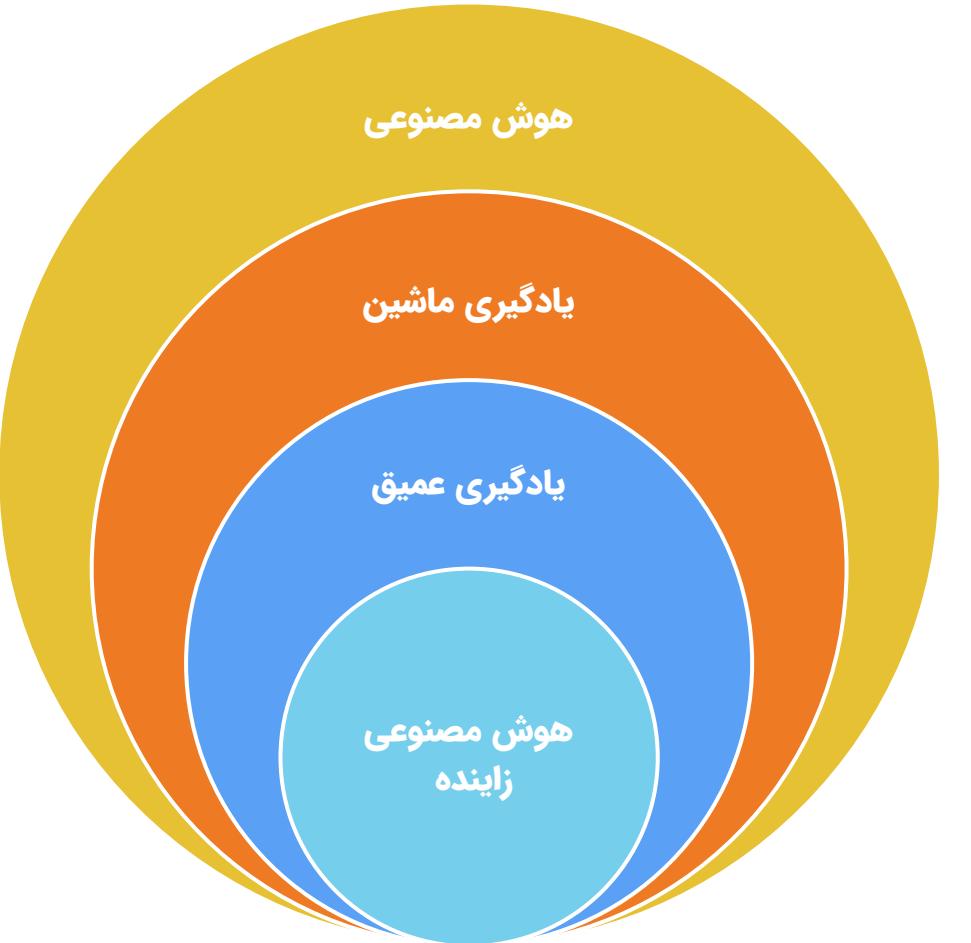
Dall-E: تولید تصویر:



مقایسه روش های هوش مصنوعی زاپینده

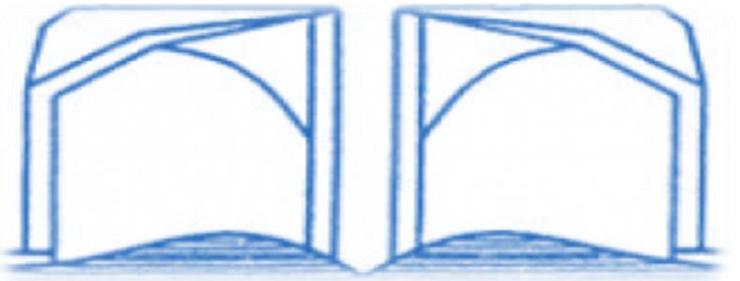


هوش مصنوعی زاینده و هوش مصنوعی





دانشگاه علوم فناوری های میان رشته ای
دانشکده سازمان های به شهند



پرده سوم شبکه عصبی مصنوعی مدرن و مدل زبانی

شبکه های بازگشتی: سازوکار توجه . . .

. . . (Attention) □ توجه

◀ توجه بیشتر انسان به اطلاعات مهم در پردازش داده (تصویر، متن و ...)

◀ مثال: عنوان گذاری تصویر



A woman is throwing a frisbee in a park.



A dog is standing on a hardwood floor.



A stop sign is on a road with a mountain in the background.

دیدگاه کاربران

پرسش و پاسخ

کاربر دیجی‌کالا ۱۴۰۰ - ۲۳ آذری بهشت

بعد از سه روز استفاده باید بگم از هر جهت گوشی کامل و بی نقصیه

آیا این دیدگاه برایتان مفید بود؟

◀ مثال: تحلیل احساس

■ Hadi Veisi (h.veisi@ut.ac.ir)

شبکه های بازگشتی: سازوکار توجه . . .

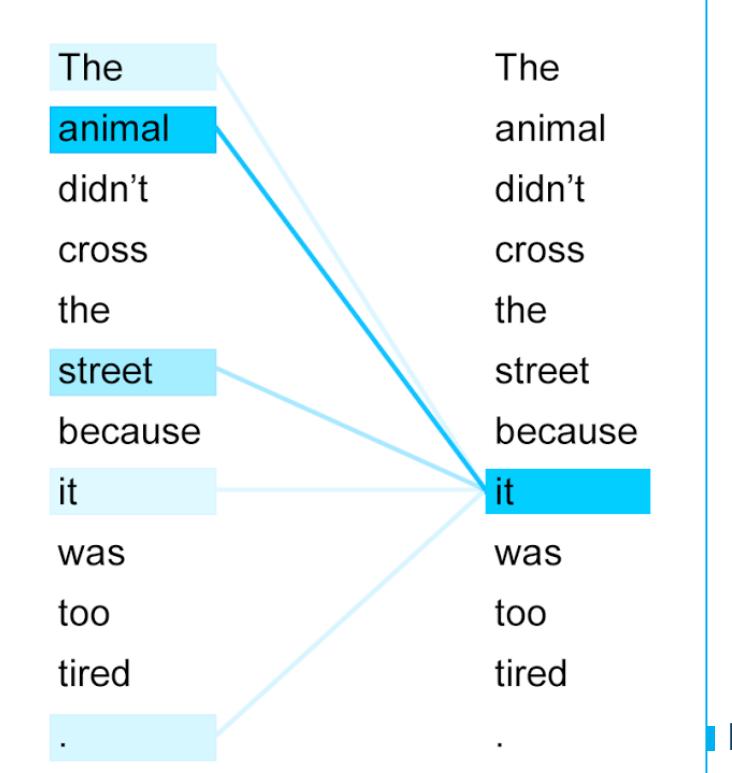
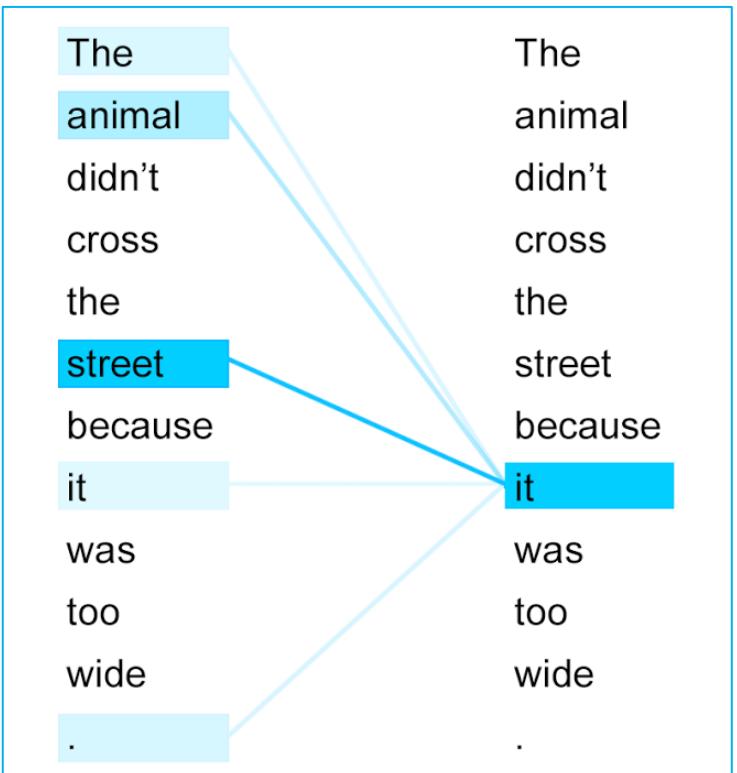
فارسی که قدمتی طولانی دارد، به عنوان زبان رسمی در کشور ایران

سابقه/بافت متن

شبکه های بازگشتی: سازوکار توجه . . .

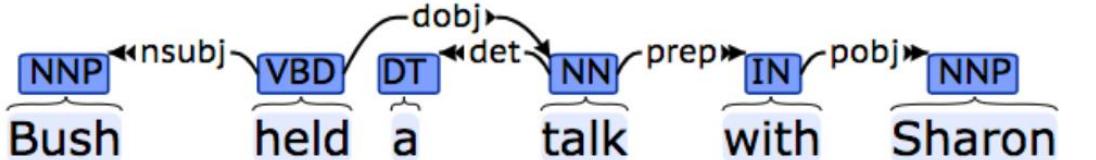
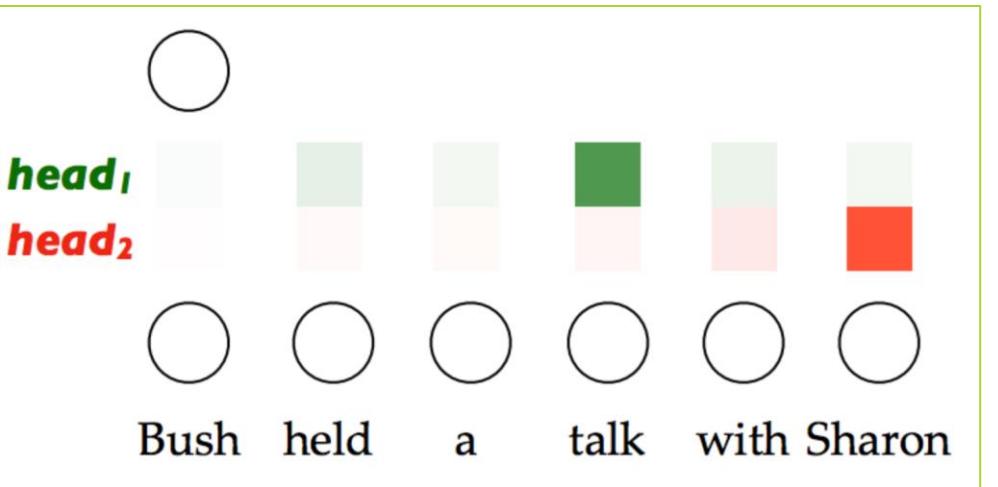
□ توجه به خود (Self-Attention)

- ◀ بیانگر میزان توجه مابین واحدهای مختلف یک دنباله (مانند کلمات یک جمله)
- ◀ یاد گرفتن اینکه *it* معادل کدام کلمه است (مرجع ضمیر)



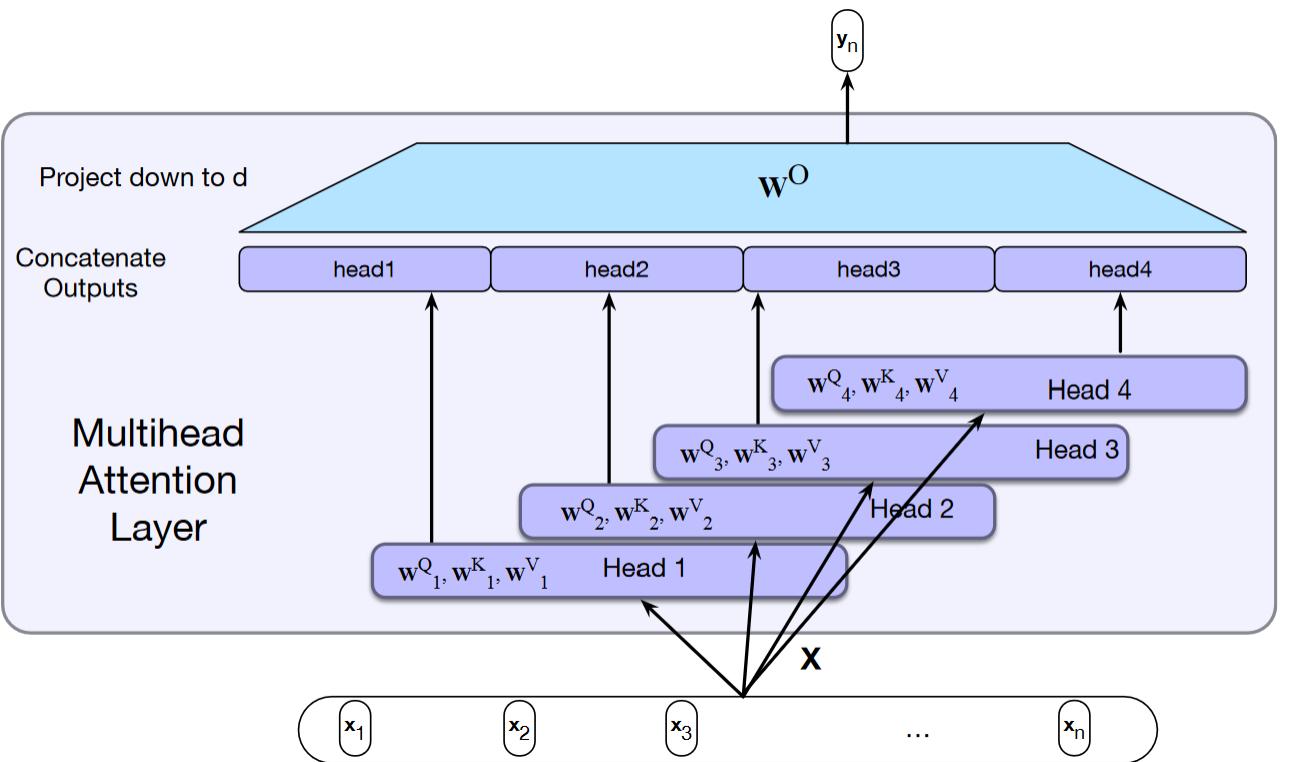
شبکه های بازگشتی: سازوکار توجه . . .

- ❑ توجه چند هسته ای (Multi-Head Attention)
- ◀ حالت توسعه یافته Self-Attention که در آن چند لایه توجه نرم به صورت موازی اجرا شده و نتایج آنها با هم ادغام می شود

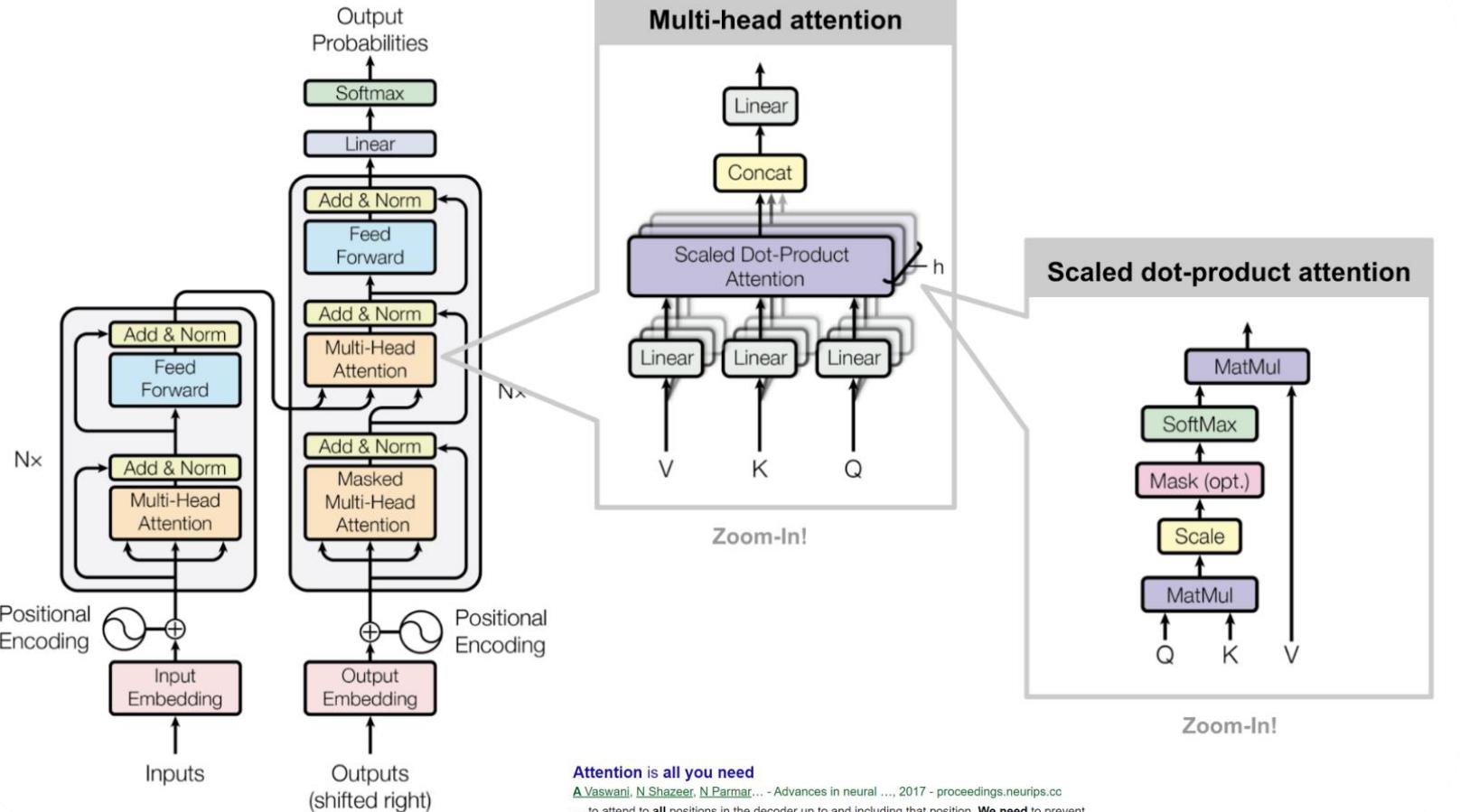


شبکه های بازگشتی: سازوکار توجه

□ توجه چند هسته ای (Multi-Head Attention): نمای کلی



شبکه عصبی مبدل (Transformer)



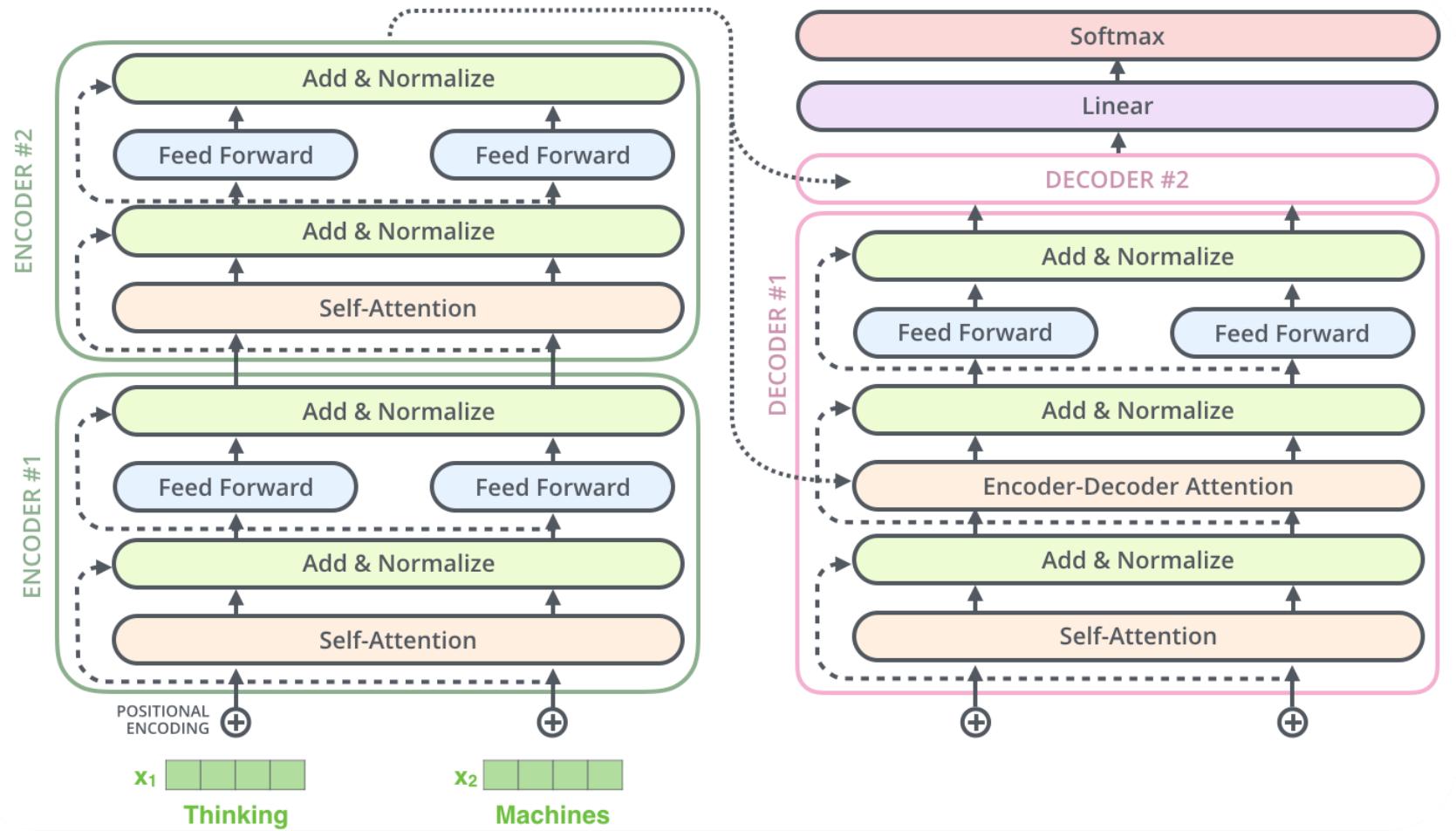
Attention is all you need

A. Vaswani, N. Shazeer, N. Parmar... - Advances in neural ... , 2017 - proceedings.neurips.cc
... to attend to all positions in the decoder up to and including that position. We need to prevent
... We implement this inside of scaled dot-product attention by masking out (setting to ∞) ...
☆ Save 55 Cite Cited by 117597 Related articles All 87 versions ☰

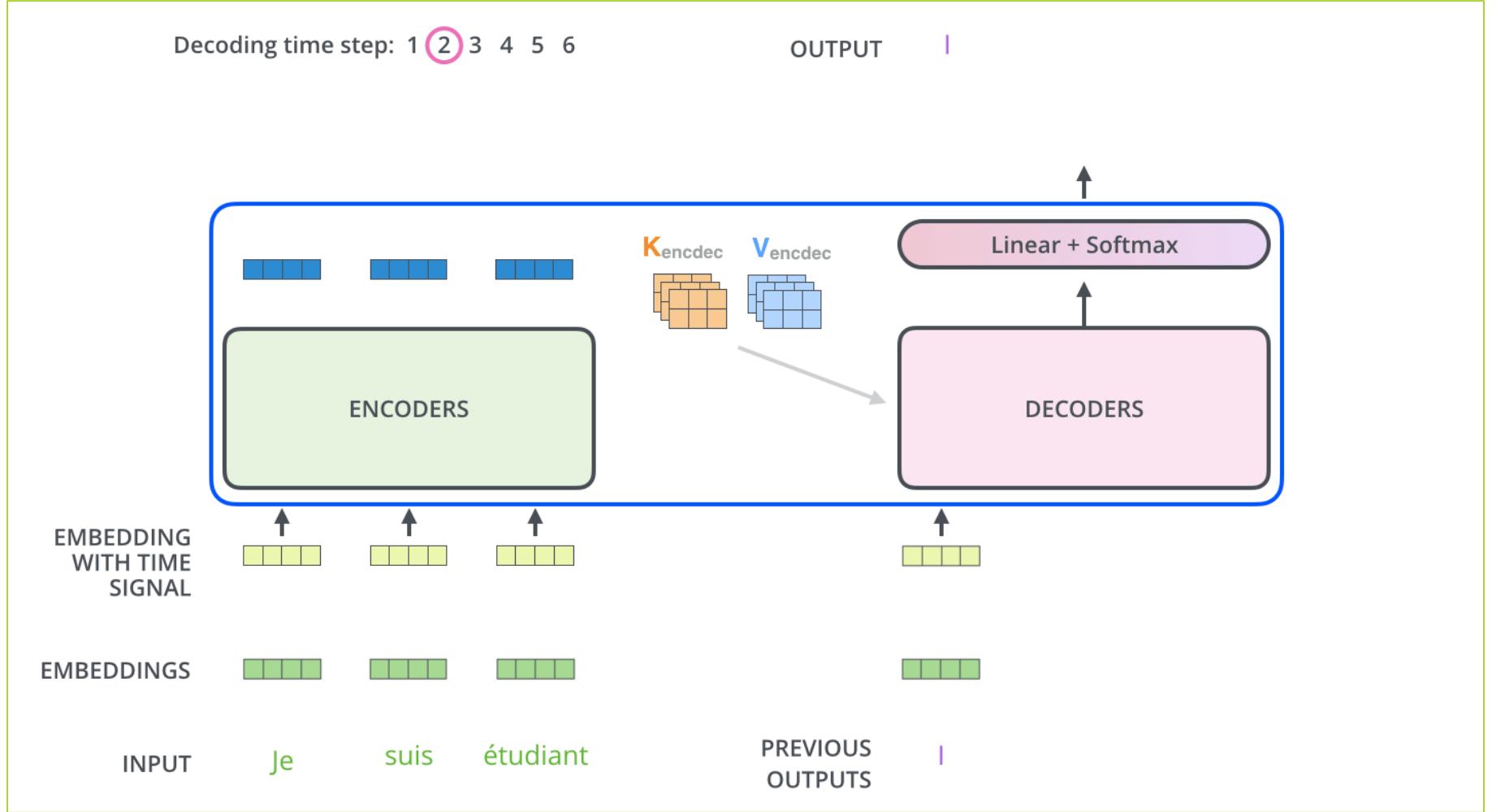
Attention is all you need

J. Uszkoreit, L. Jones, A. Gomez, L. Kaiser, I. Polosukhin - NIPS, 2017 - aiotlab.org
... Attention • Visual attention and textual attention ... Could we build a transformer-based
model whose language model looks both forward and backwards, ie "is conditioned on both ...
☆ Save 99 Cite Cited by 2234 Related articles All 2 versions ☰

شبکه عصبی مبدل (Transformer)



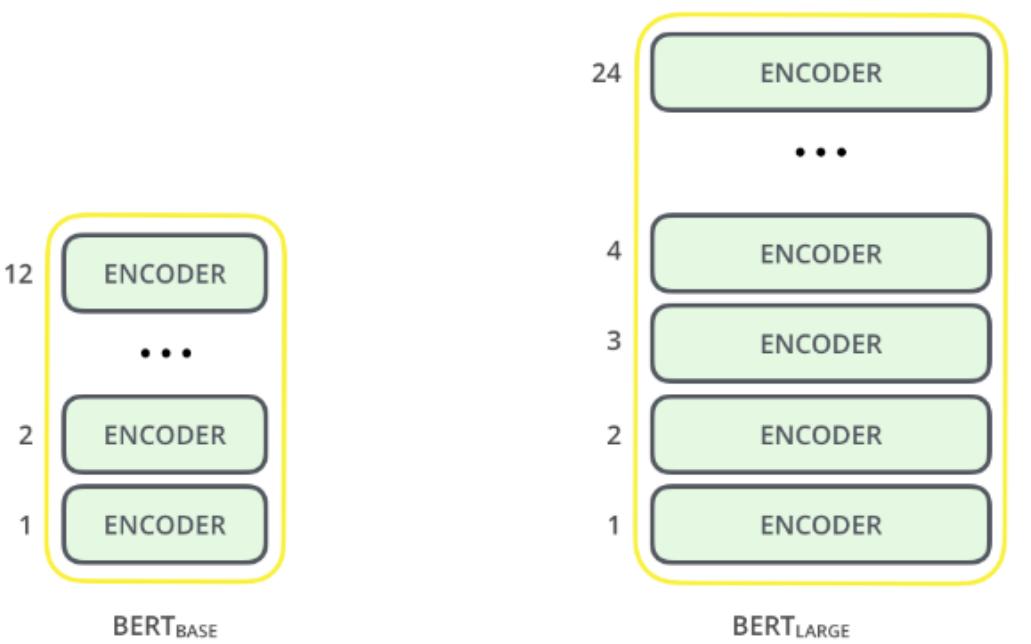
شبکه عصبی مبدل (Transformer)



مدل زبانی مبتنی بر مبدل: BERT

Bidirectional Encoder Representations from Transformers (BERT) □ مدل

◀ فقط رمزگذار = تعابیه کلمات



تعداد رمزگذارها: ۱۲
 ابعاد بردارها (لایه مخفی): ۷۶۸
 تعداد هسته‌ها در توجه چندهسته‌ای: ۱۲

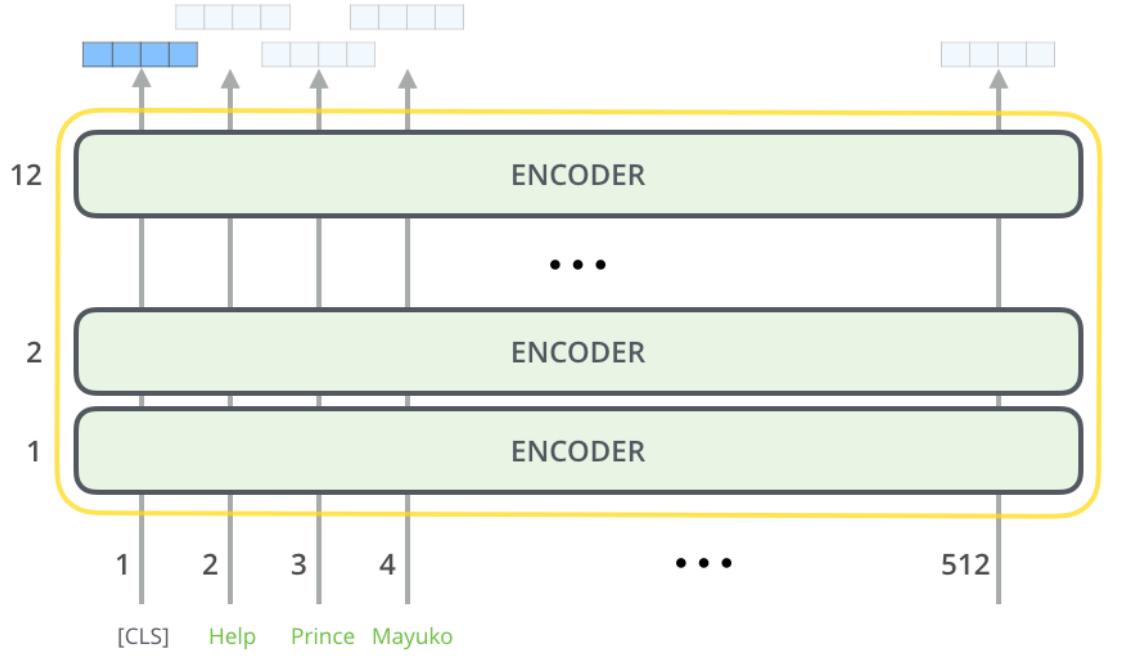
تعداد رمزگذارها: ۲۴
 ابعاد بردارها (لایه مخفی): ۱۰۲۴
 تعداد هسته‌ها در توجه چندهسته‌ای: ۱۶

مدل زبانی مبتنی بر مبدل: BERT

◀ طول دنباله ورودی: ۵۱۲ توکن (کلمه)

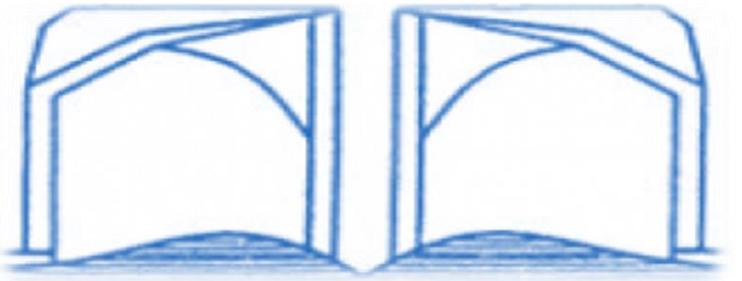
◀ تعداد لایه (رمزگذارها): ۲۴ (مدل Bert-Large)

◀ ابعاد بردارها (لایه مخفی): ۱۰۲۴ و تعداد هسته‌ها در توجه چندهسته‌ای: ۱۶





دانشگاه علوم فنی های میان رشته ای
دانشکده سازمان های به شنبه



پرده کنونی مدل زبانی بزرگ و انقلاب GPT

مدل زبانی بزرگ (LLM)

- مدلی مبتنی بر شبکه عصبی (مدل) برای پیش بینی (احتمال) کلمه بعدی

- ◀ تعداد لایه های زیاد
- ◀ پارامترهای زیاد در هر لایه
- ◀ آموزش با حجم زیاد متن

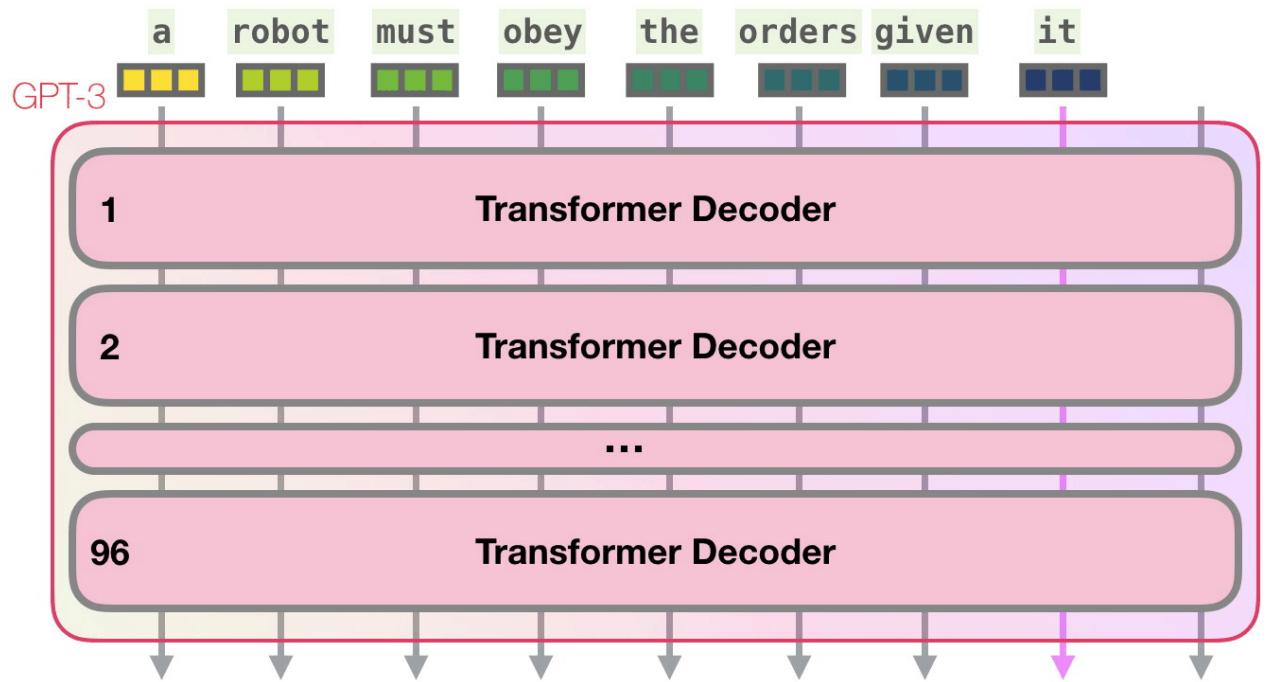


مدل زبانی مبتنی بر مبدل: GPT³

◀ طول دنباله ورودی: ۲۰۴۸ توکن (کلمه)

◀ تعداد لایه (رمزگذارها): ۹۶ (هر کدام ۱.۸ میلیون پارامتر)

◀ داده: حدود ۵۰۰ میلیارد توکن

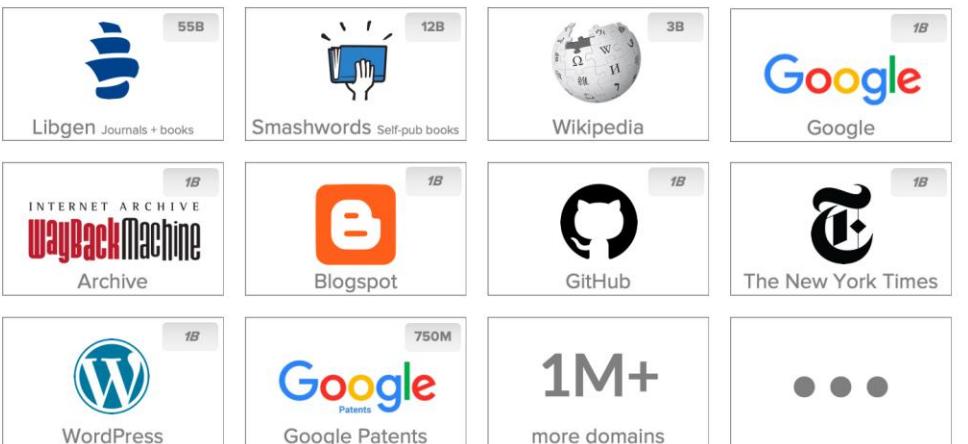


... GPT3 مدل

□ منابع دادگانی

◀ حدود ۵۰۰ میلیارد توکن

GPT-3'S TOP 10 DATASETS (BY DOMAIN/SOURCE)



Informed "best guess" only.
Sources: <http://LifeArchitect.ai/papers/>

Alan D. Thomason, Rev 3: February 2022.
<https://LifeArchitect.ai/>

 LifeArchitect.ai/models

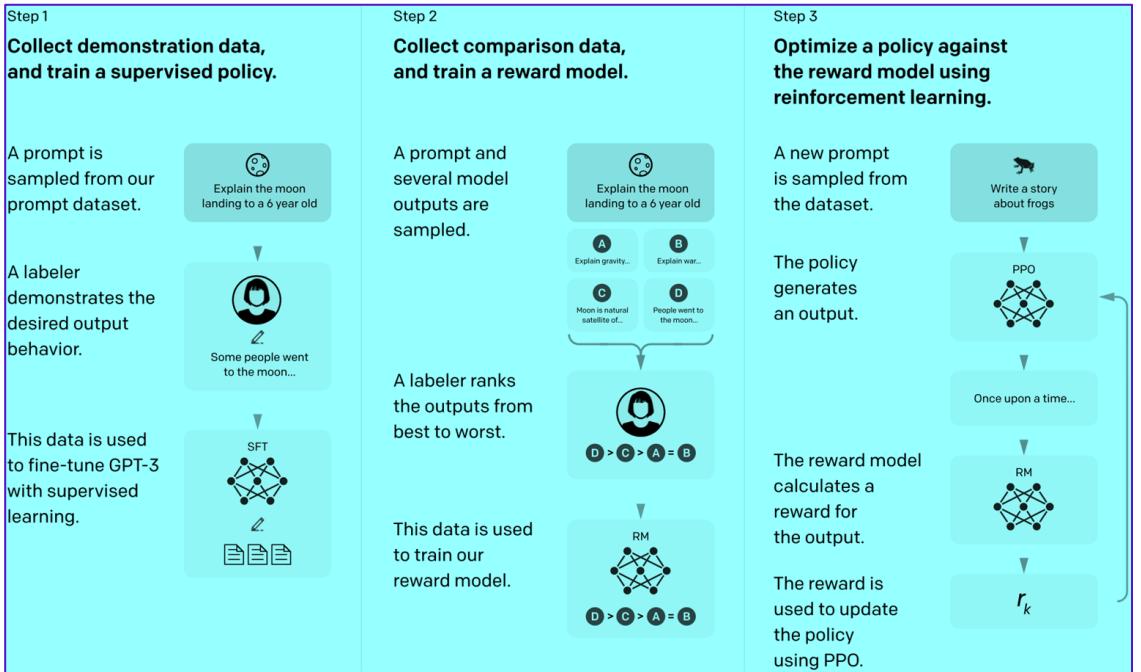
Dataset	Tokens	Assumptions	Tokens per byte	Ratio	Size
Web data	410B	–	0.71	1:1.9	570
WebText2	19B	25% > WebText	0.38	1:2.6	50
Books1	12B	Gutenberg	0.57	1:1.75	21
Books2	55B	Bibliotik	0.54	1:1.84	101
Wikipedia	3B	See RoBERTa	0.26	1:3.8	11.4
Total	499B				753.4GB

از ChatGPT به GPT3

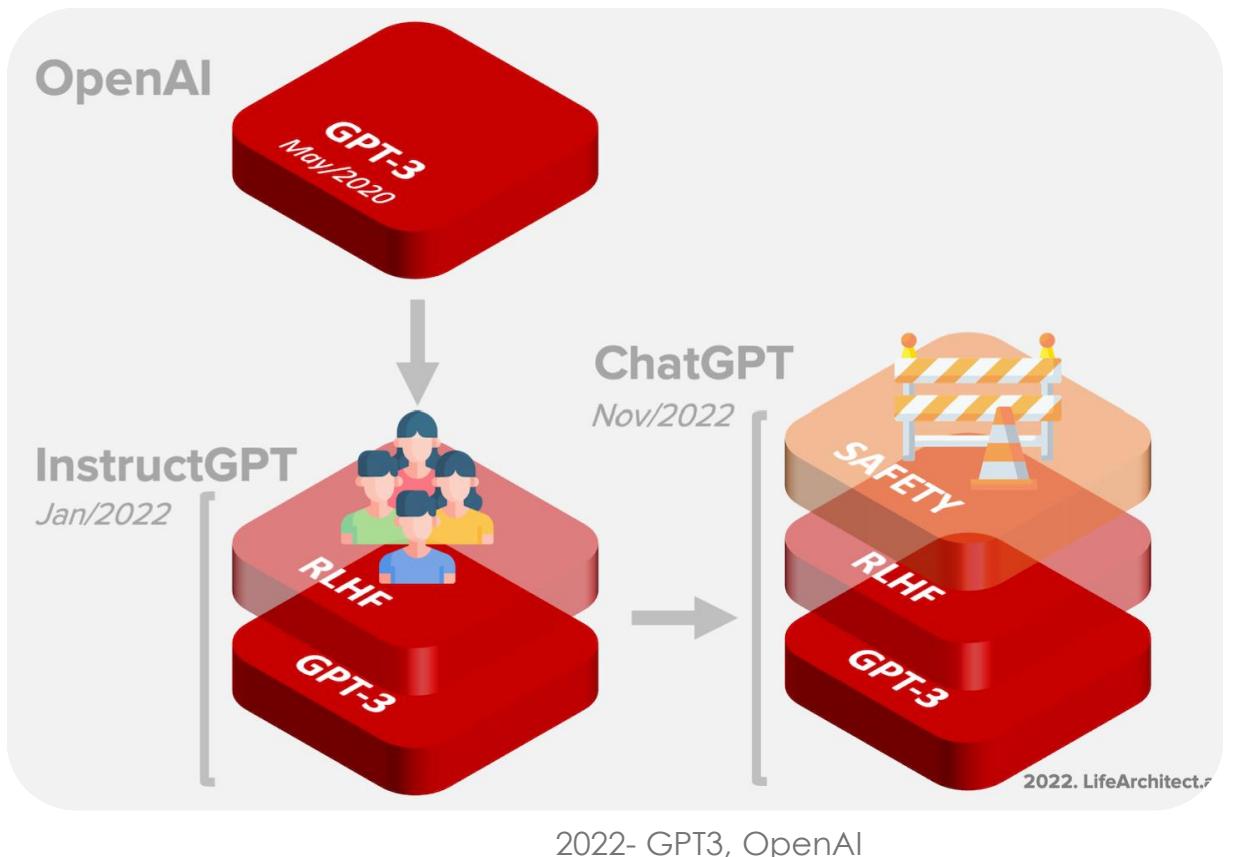
GPT-۳.۵/ InstructGPT مدل □

- ◀ طول بافت = ۴K برابر با دو برابر GPT۳
- ◀ آموزش روی همان داده GPT۳
- ◀ تنظیم دقیق (Fine-Tune) با داده انسانی سوال و جواب

RLHF: Reinforcement Learning with Human Feedback ◀



... ChatGPT به GPT3 از

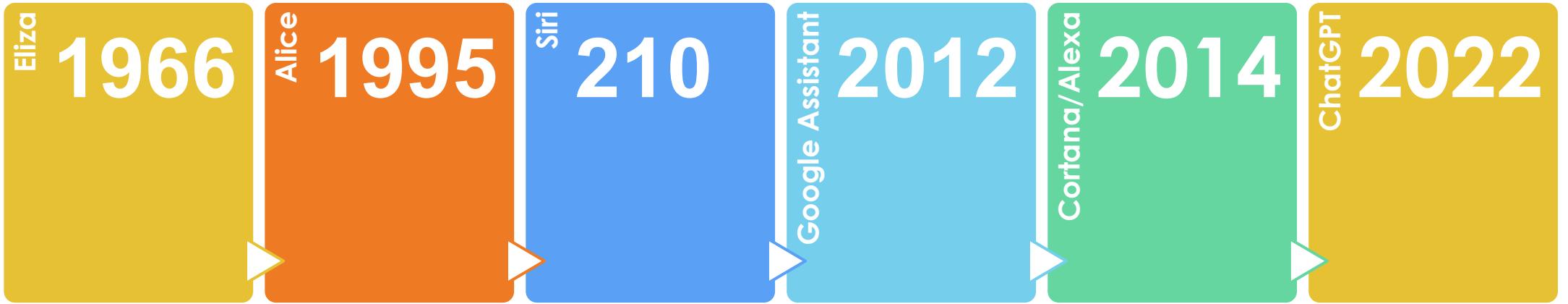


- ❑ تنظیم دقیق (Fine-Tune) دوباره
- ❑ سازگار کردن با قوانین و محدودیت‌ها
- مشابه مدل DeepMind Sparrow



دانشگاه علوم فنی های میان رشته ای
دانشکده سالانه های بومی

و انقلاب درگ زبان و تولید متن ChatGPT





دانشگاه علوم فنی های میان رشته ای
دانشکده سازه های بومی

مدل GPT4: باز هم بزرگ تر . . .

- داده آموزشی
◀ ۲۰ تریلیون توکن
- اندازه مدل
◀ یک تریلیون پارامتر
- مدلسازی وابستگی در متن
◀ ۳۲ هزار توکن

... GPT4 مدل

Count	Dataset	Percentage tokens	Raw Size (GB)	Tokens (B)
1	Common Crawl (mostly English)		1,900 ^m	506B ^m
2	Dialogue (hypothesis)		26,600	17,000B
3	Special		6,277 ⁱ	1,405B ⁱ
4	Code (GitHub)		3,100 ^m	422B ^m
5	Books		2,100 ^m	560B ^m
6	Wikipedia		48 ⁱ	13B ⁱ
7	AMPS (maths)		23	6B
8	UC Berkeley MATH*	<0.1%	30MB	8M
9	OpenAI GSM8K*	<0.1%	17MB	4.5M
Totals			40,000GB (40TB)	20,000B (20T)

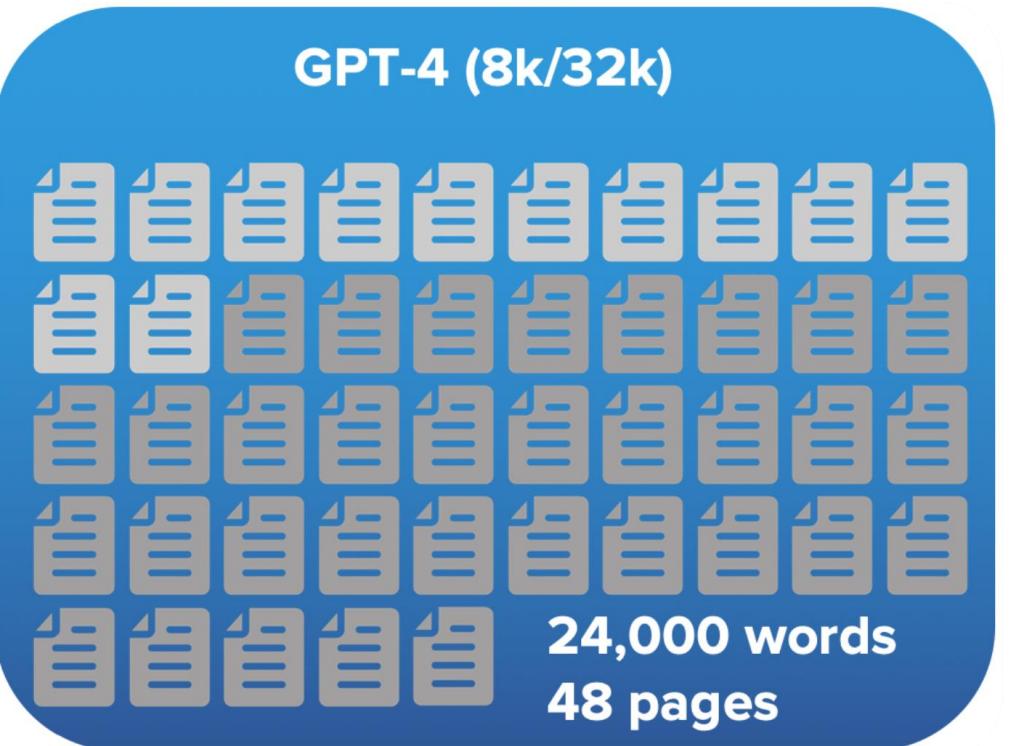
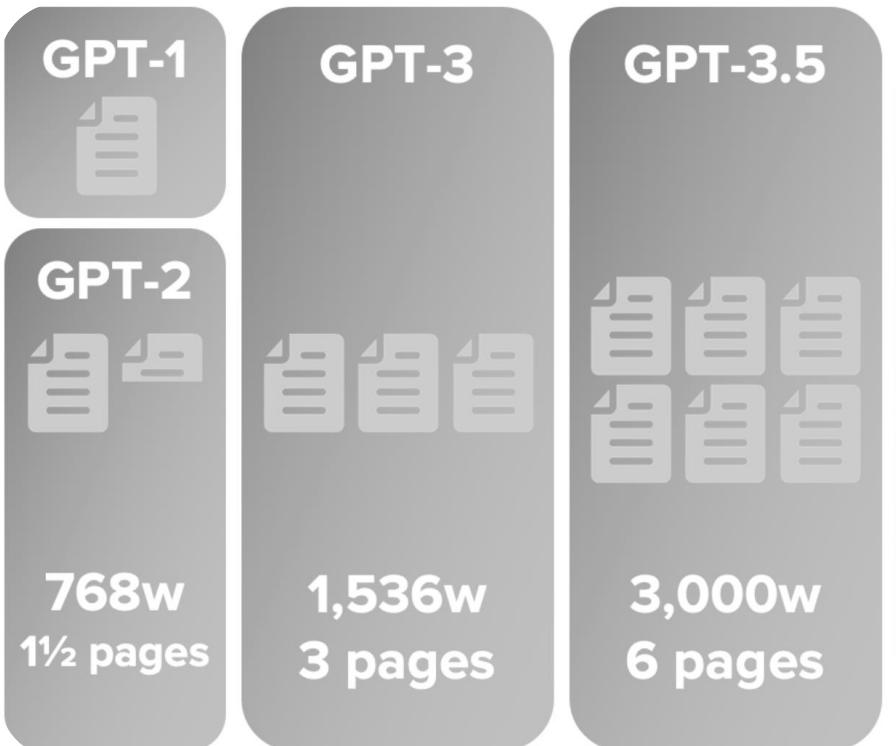
□ منابع دادگانی

◀ حدود ۲۰ تریلیون توکن

GPT3 ۴۰ برابر ◀

مدل GPT4 . . .

□ مدل سازی وابستگی زبانی (بافت متن) بزرگ تر



مدل GPT4 . . .

□ برتری به انسان در بیشتر حوزه ها!

GPT-4 VS HUMAN TESTS (APR/2023)

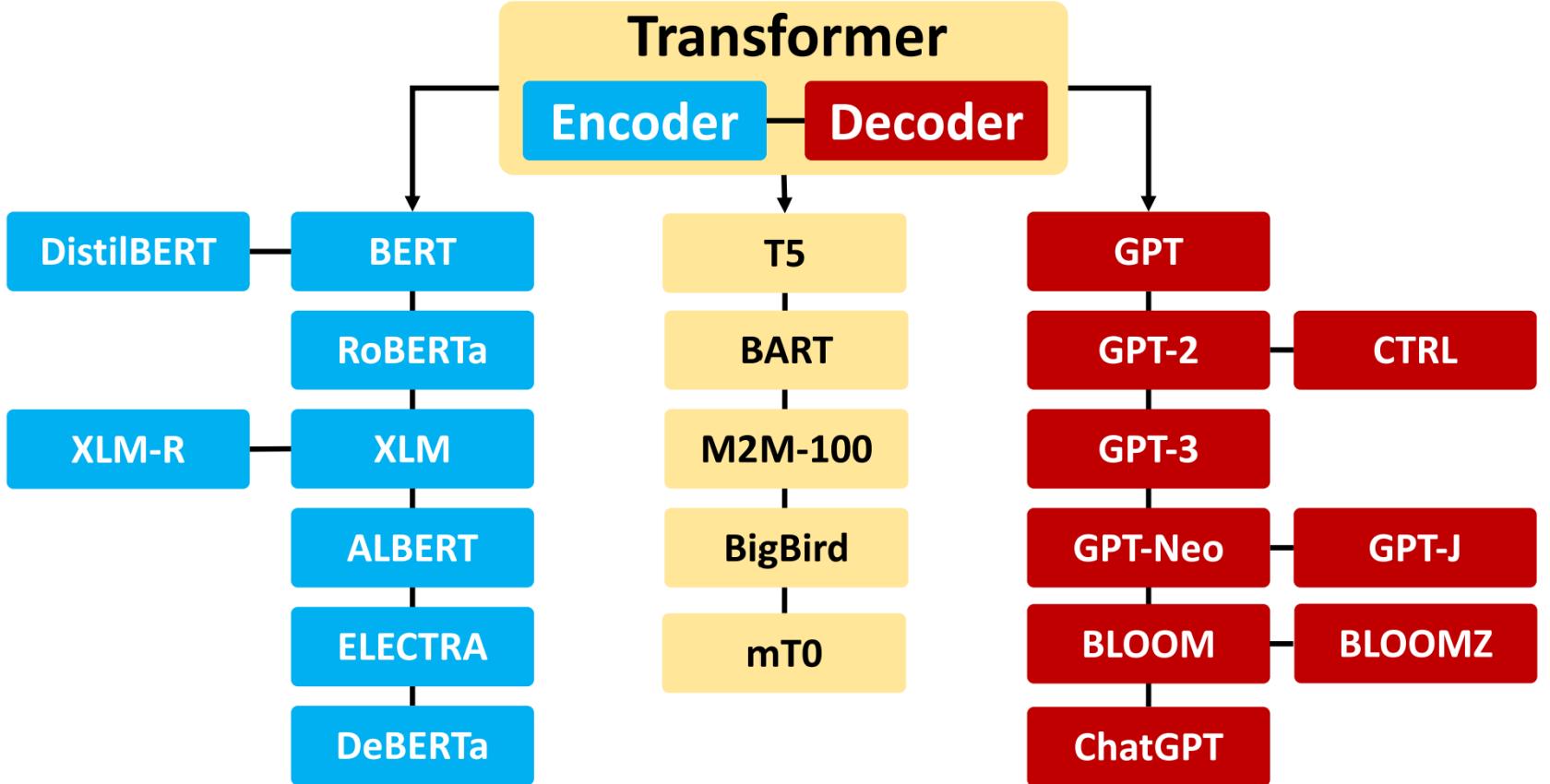


[LifeArchitect.ai/iq-testing-ai](https://lifearchitect.ai/iq-testing-ai/)

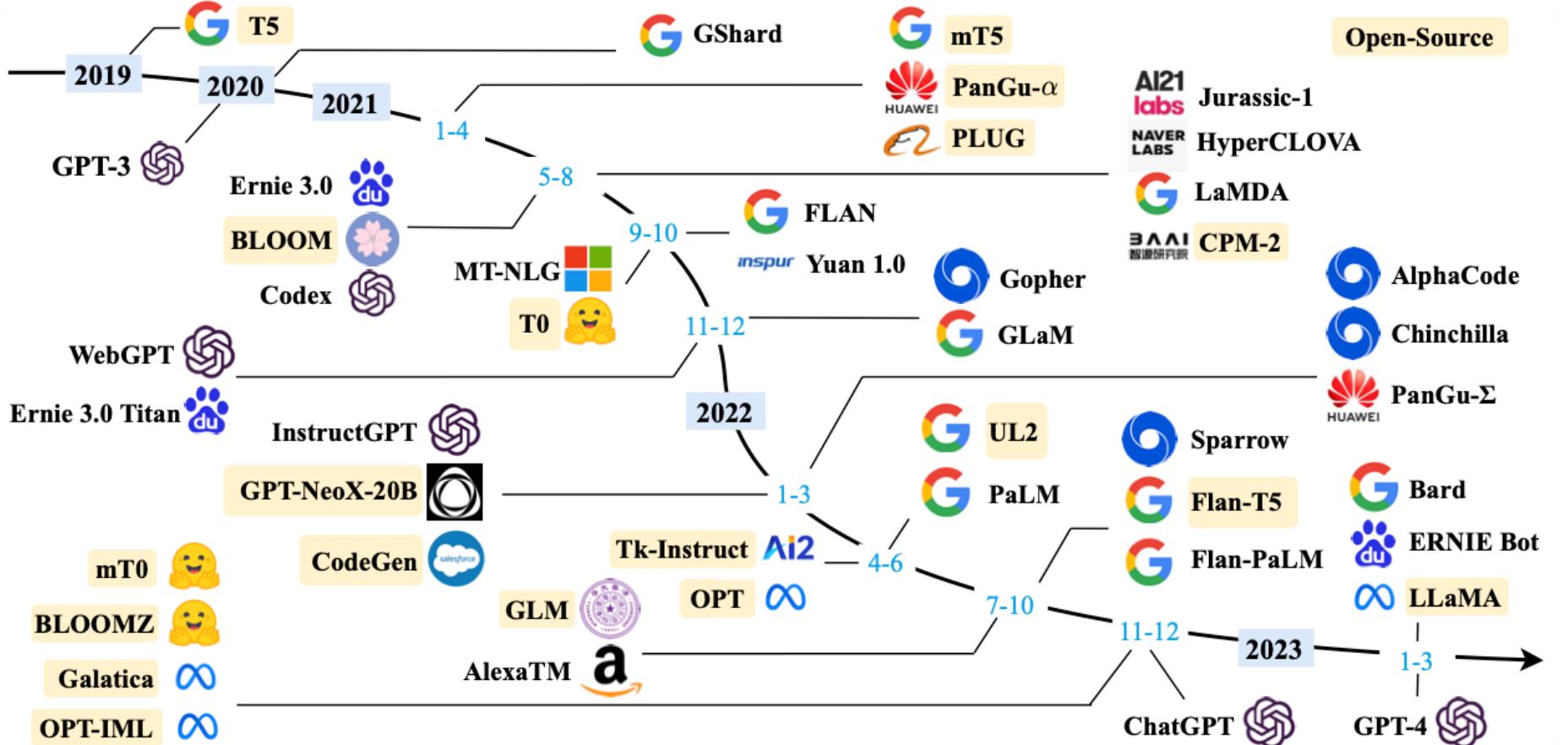
... GPT4 و GPT3 مقایسه

شاخص	GPT3	GPT4
تعداد پارامترها	175 میلیارد	1000 میلیارد (1 تریلیون) برابر GPT3 5.7
داده آموزشی (токن)	500 میلیارد	20000 میلیارد برابر GPT3 40
نوع ورودی	متن	متن (26 زبان) و عکس
طول جمله ورودی (токن)	4K (GPT3.5) و 2K (GPT3)	32K و 8K برابر GPT3.5 8
ایمنی (برتری GPT4)	%40 بهتر در تولید پاسخهای Factual %82 بهتر در تولید محتوای نامناسب	
کارایی (برتری GPT4)		دقت بالاتر در دادگان های معیار مختلف و موضوع های مختلف
فرمان پذیری (Steerability)	(system API) دستور	کنترل بیشتر در اختصاصی کردن و محدود کردن محتواها در API

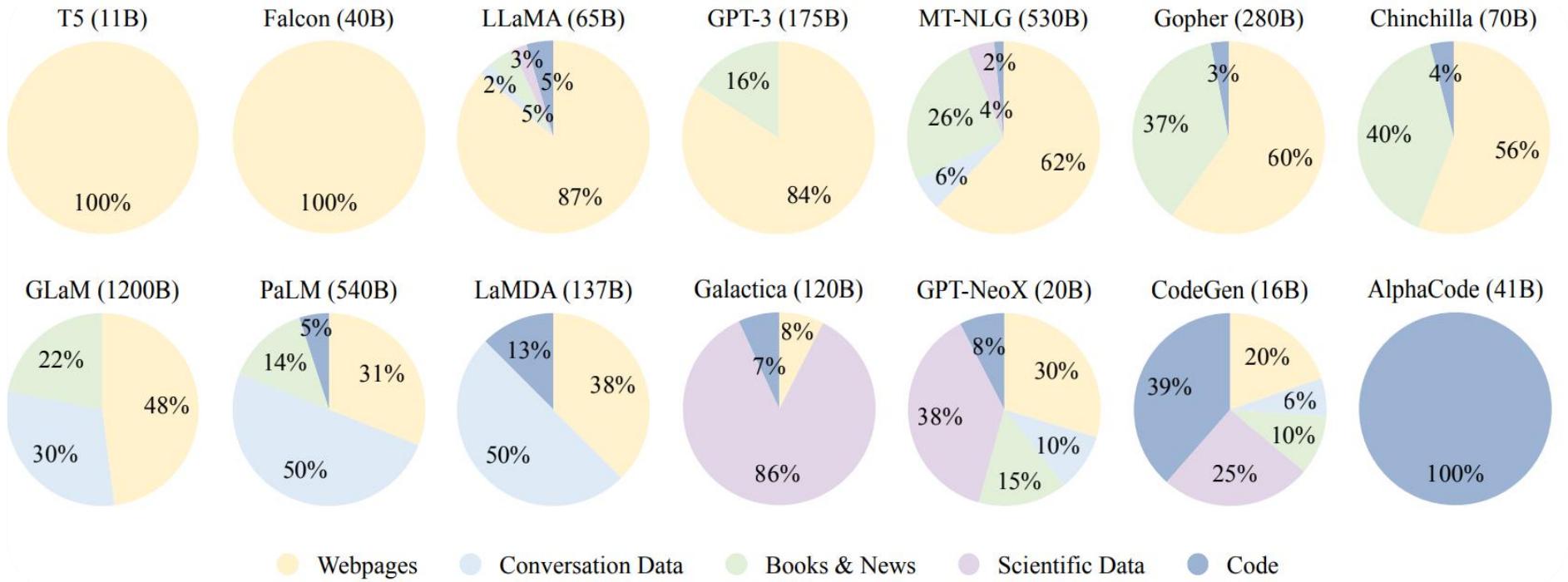
مدل زبانی های بزرگ (LLM) : انواع . . .



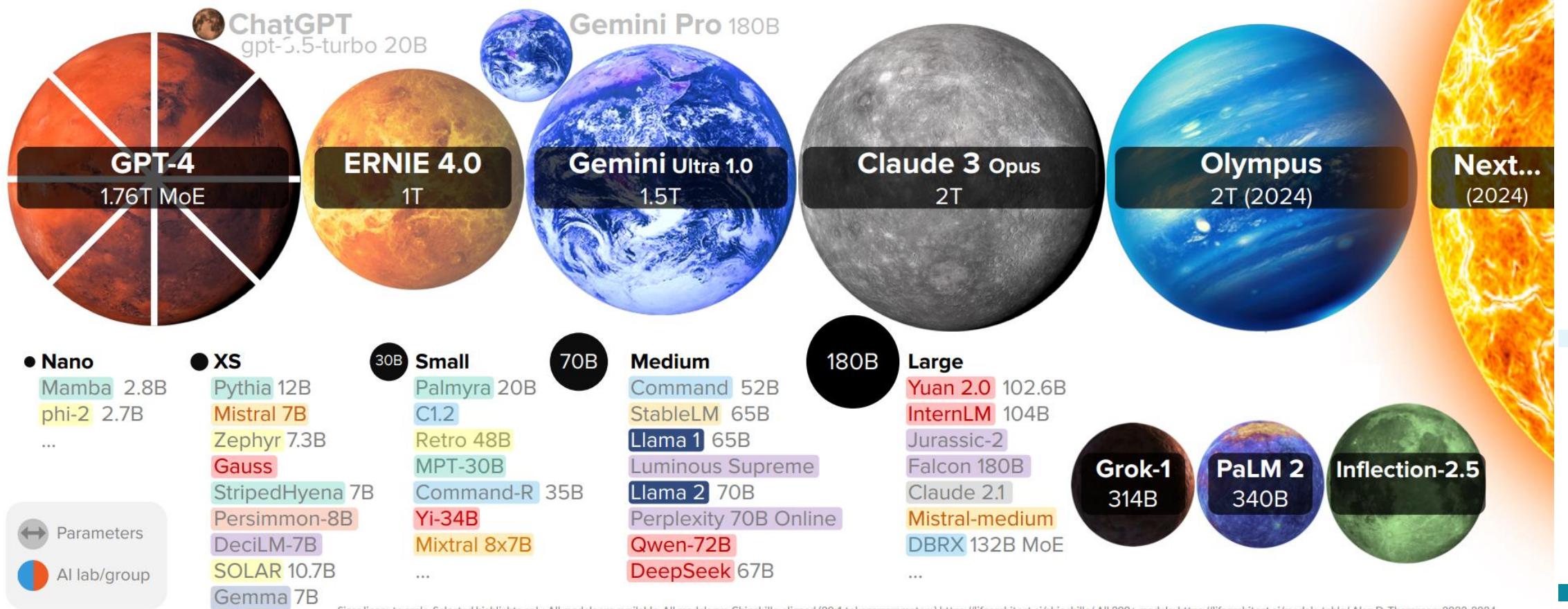
مدل زبانی بزرگ (LLM): سال های شکوفایی



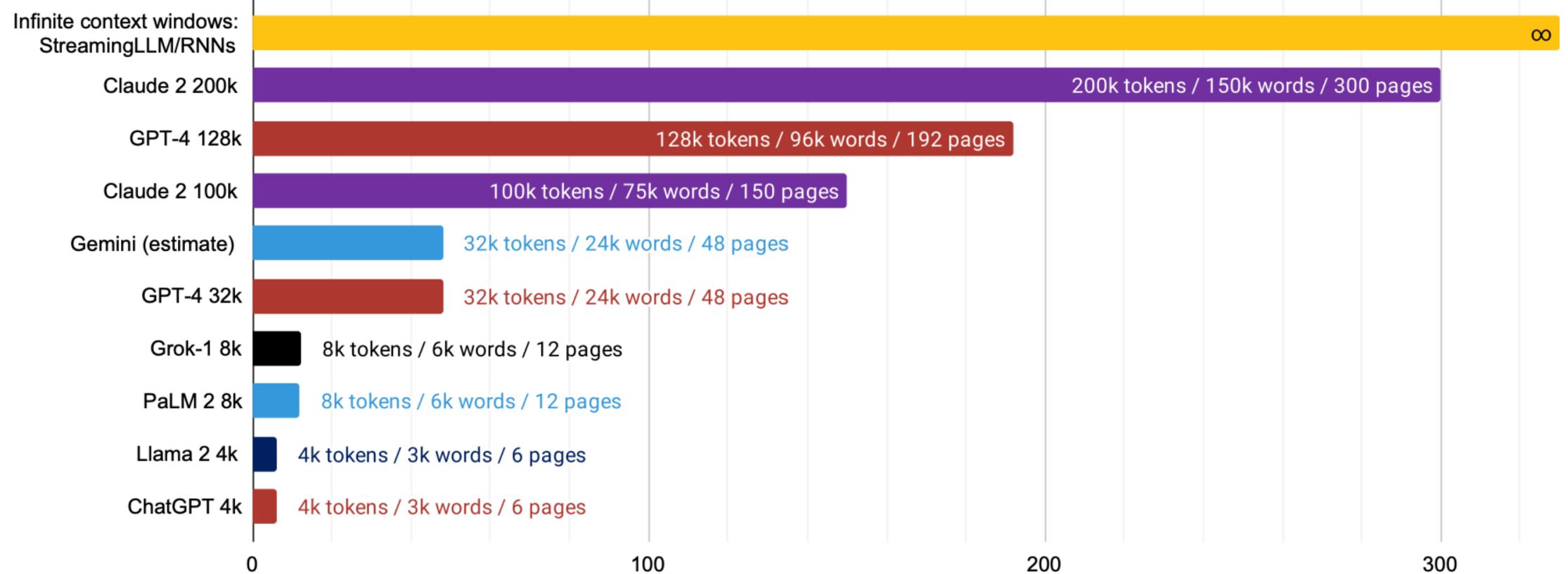
مدل زبانی بزرگ (LLM): منابع داده



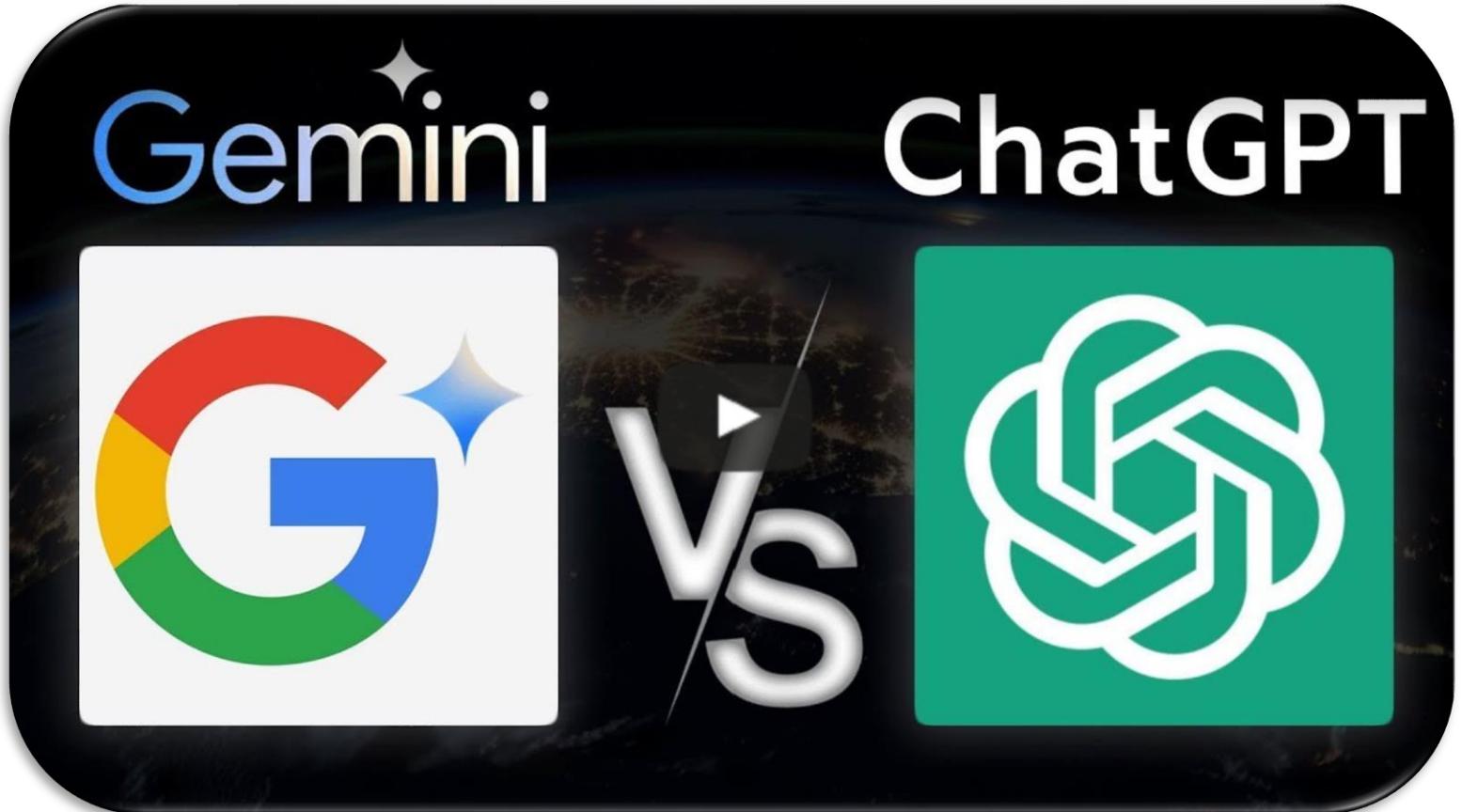
مدل زبانی بزرگ (LLM): اندازه = تعداد پارامترها



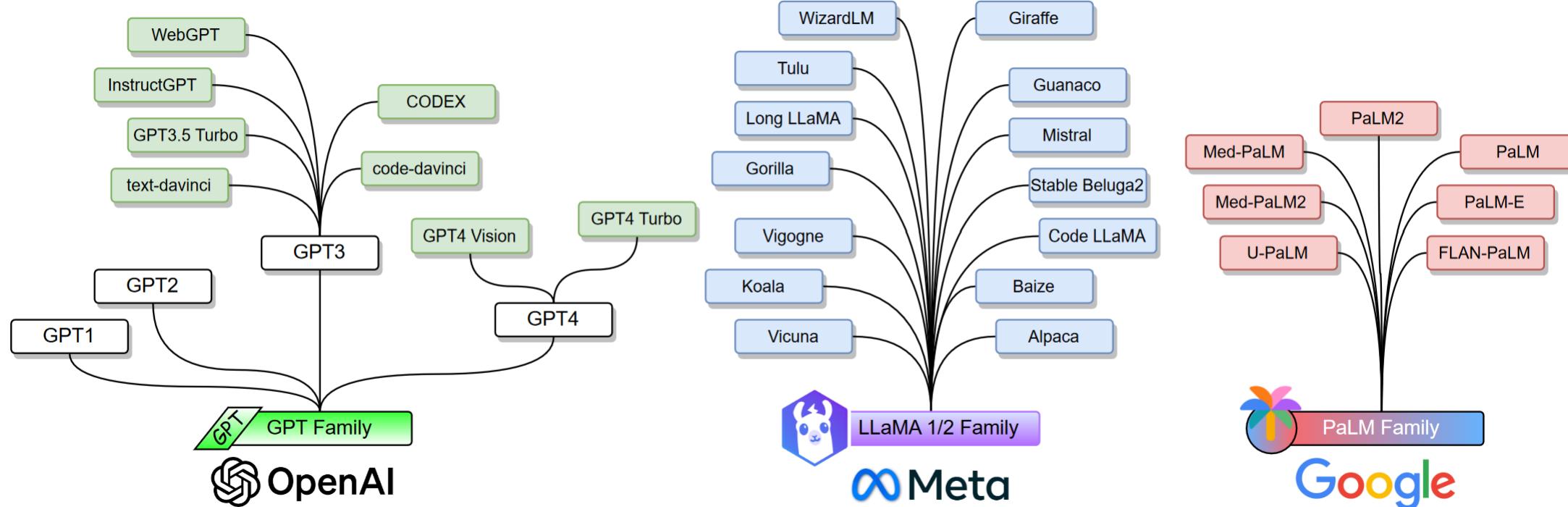
مدل زبانی بزرگ (LLM): اندازه بافت = طول متن ورودی/خروجی



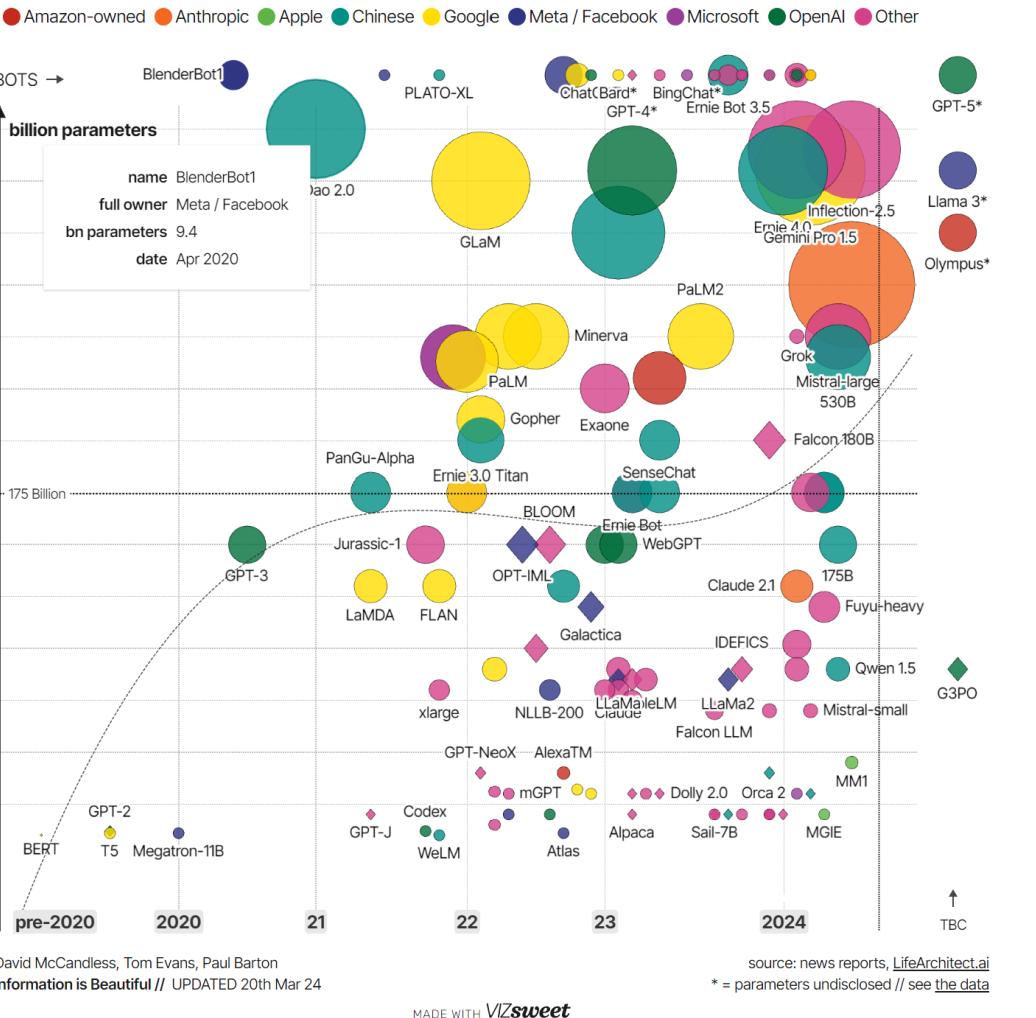
مبدل و مدل زبانی بزرگ: بازی بزرگان



مدل زبانی بزرگ (LLM): بزرگان!



مدل زبانی بزرگ (LLM): بزرگان!



<https://informationisbeautiful.net/visualizations/the-rise-of-generative-ai-large-language-models-llms-like-chatgpt/>

تولید متن: برخی کاربردها

- **تولید محتوا**
 - ◀ نوشتن نامه، مقالات، پست‌های رسانه‌های اجتماعی، توضیح محصول و انواع دیگر محتوا برای وبسایت‌ها و پلتفرم‌های رسانه‌های اجتماعی
- **چتبات‌ها و دستیارهای مجازی**
 - ◀ پاسخ به درخواست‌های مشتریان پاسخ و ارائه پشتیبانی
- **آموزش**
 - ◀ ایجاد محتواهای آموزشی مانند کتاب‌های درسی، آزمون‌ها و راهنمایها
 - ترجمه از زبانی به زبان دیگر
 - خلاصه‌سازی متن



عدم نیاز به یادگیری زبان در گفتگوی افراد با زبانهای مختلف



دانشگاه علوم فناوری های میان رشته ای
دانشکده سالانه های بومی

تولید کد کامپیوتری: CoPilot و CodeX

The image displays two web browser windows side-by-side, illustrating the capabilities of different AI-driven code generation tools.

Left Window (openai.com/blog/openai-codex):

- Header:** OpenAI
- Navigation:** Research, Product, Developers, Safety, Company
- Title:** OpenAI Codex
- Text:** We've created an improved version of OpenAI Codex, our AI system that translates natural language to code, and we are releasing it through our API in private beta starting today.
- Button:** Start using Codex ↗

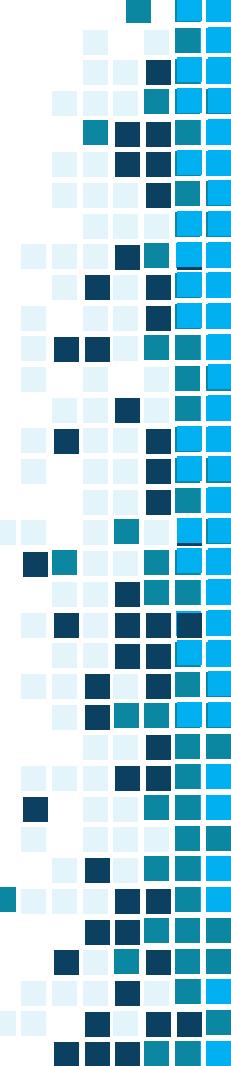
Right Window (github.com/features/copilot):

- Header:** github.com/features/copilot
- Navigation:** Product, Solutions, Open Source, Pricing
- Search Bar:** Search GitHub
- Sign In / Sign Up Buttons:**
- Section:** Features (Actions, Packages, Security, Codespaces, Copilot, Code review, Search, Issues, Discussions)
- Title:** Your AI pair programmer
- Text:** GitHub Copilot uses the OpenAI Codex to suggest code and entire functions in real-time, right from your editor.
- Buttons:** Start my free trial >, Compare plans
- Code Preview:** A snippet of code in a text editor interface, showing imports and function definitions.



تولید نرم افزار با زبان طبیعی (کامپایلرهای زبان انگلیسی)!

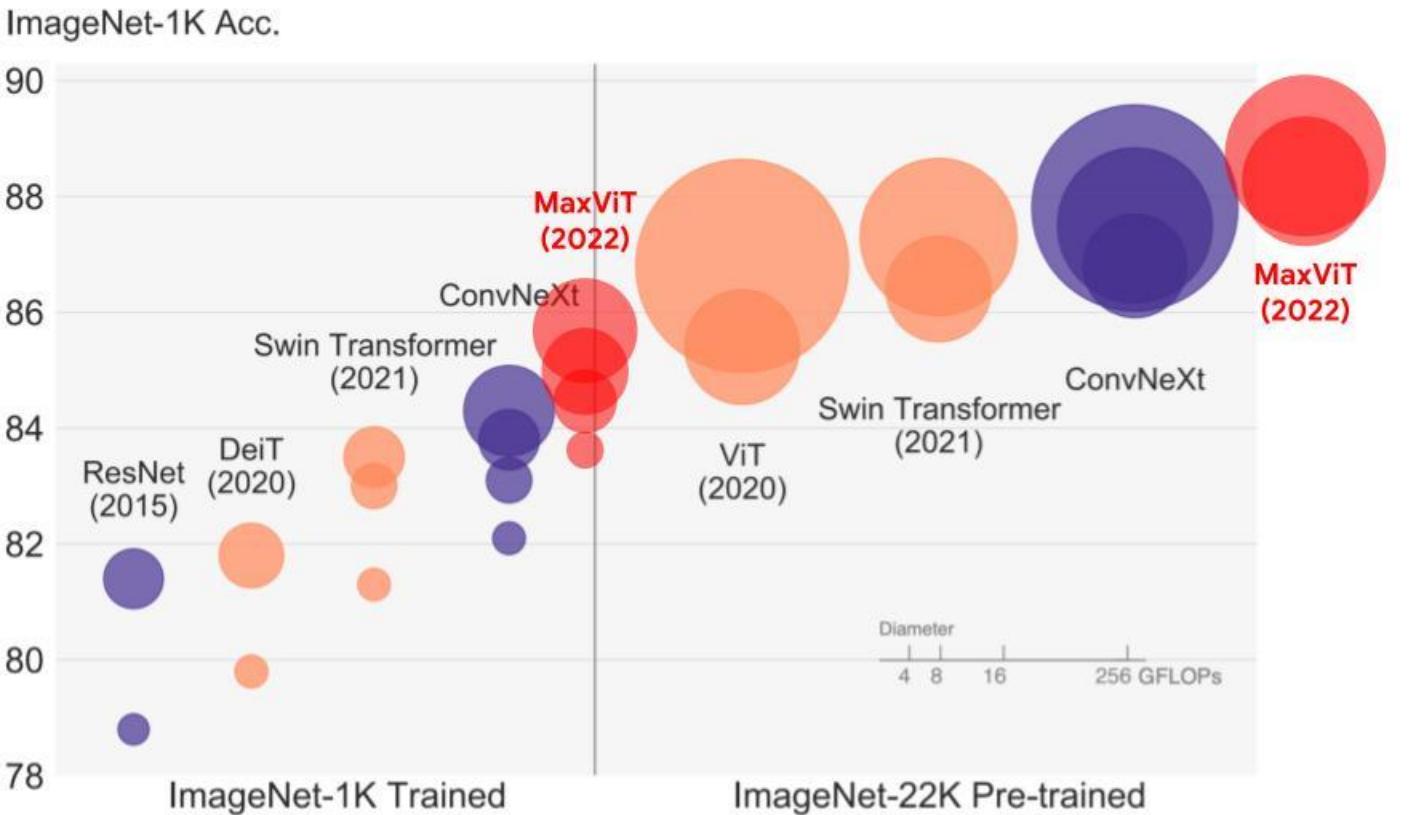
مبدل‌ها: مبدل بینایی (Vision Transformer)



Dosovitskiy et. al., (2021-06-03). "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale".
Liu et. al., (2021-03-25). "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows"

■ Hadi Veisi (h.veisi@ut.ac.ir)

مبدل ها: مبدل بینایی (Vision Transformer)



(Bing) OpenAI از DALL-E: تبدیل متن به عکس:

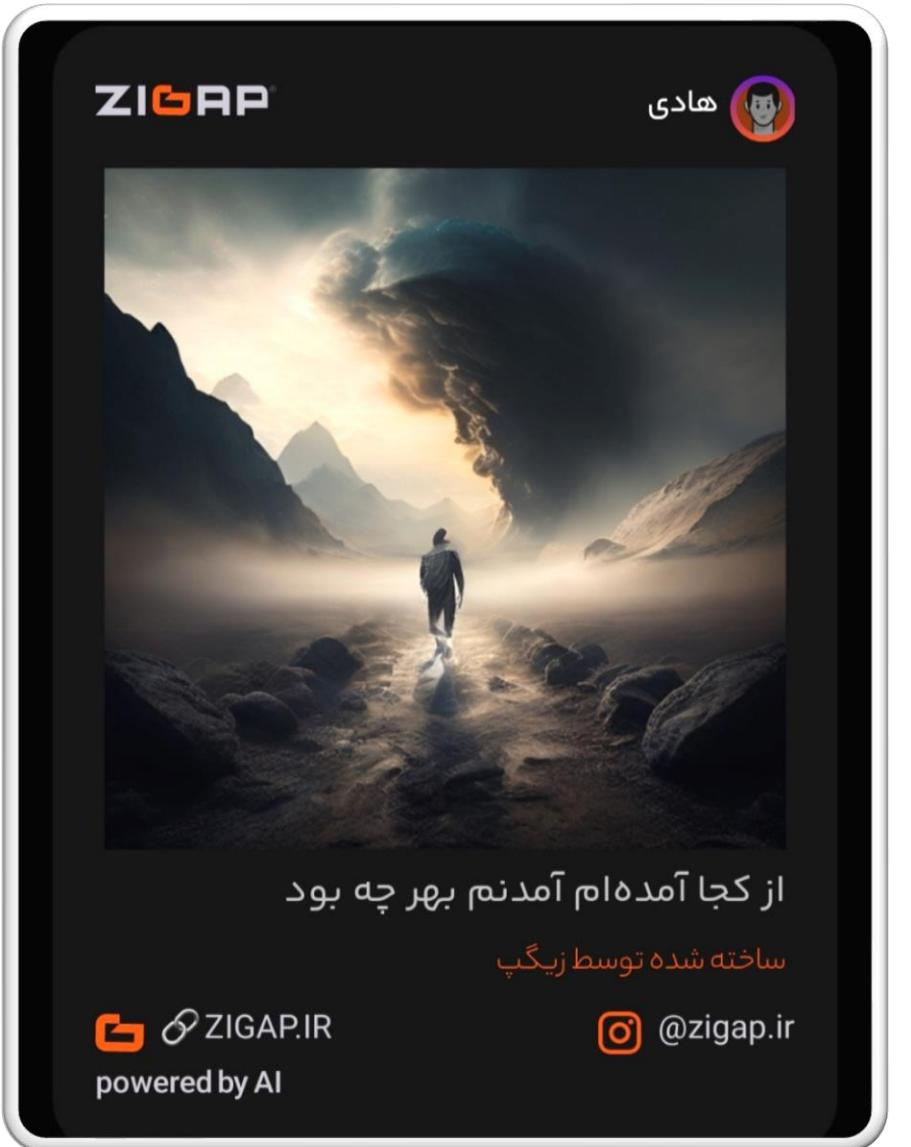
Microsoft Bing |  | **Image Creator**
powered by DALL-E | **PREVIEW**

یک جمعیت ایرانی در یک اتاق کنفرانس



The Microsoft Bing interface shows four generated images of an Iranian conference audience. The images depict large crowds of people, mostly men wearing traditional white caps, seated in rows in ornate hallways and rooms with high ceilings and gold-colored decorations. The images are generated by the Image Creator feature, which uses DALL-E technology.

تبدیل متن به عکس: Leonardo



□ اپ زیگپ (ZiGap)

◀ قابل دانلود از کافه بازار، مایکت و سایت اپ

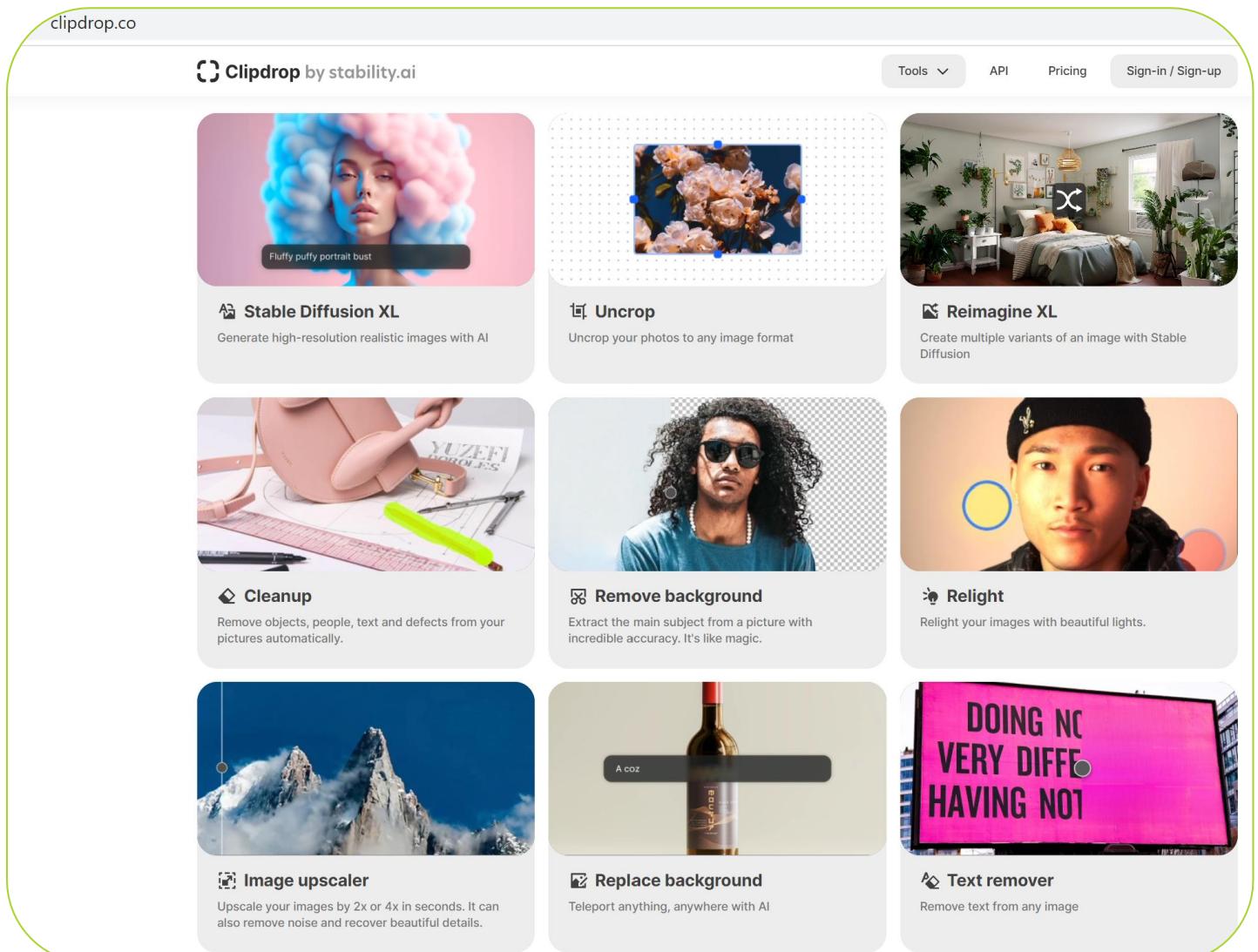
Zigap.ir ◀

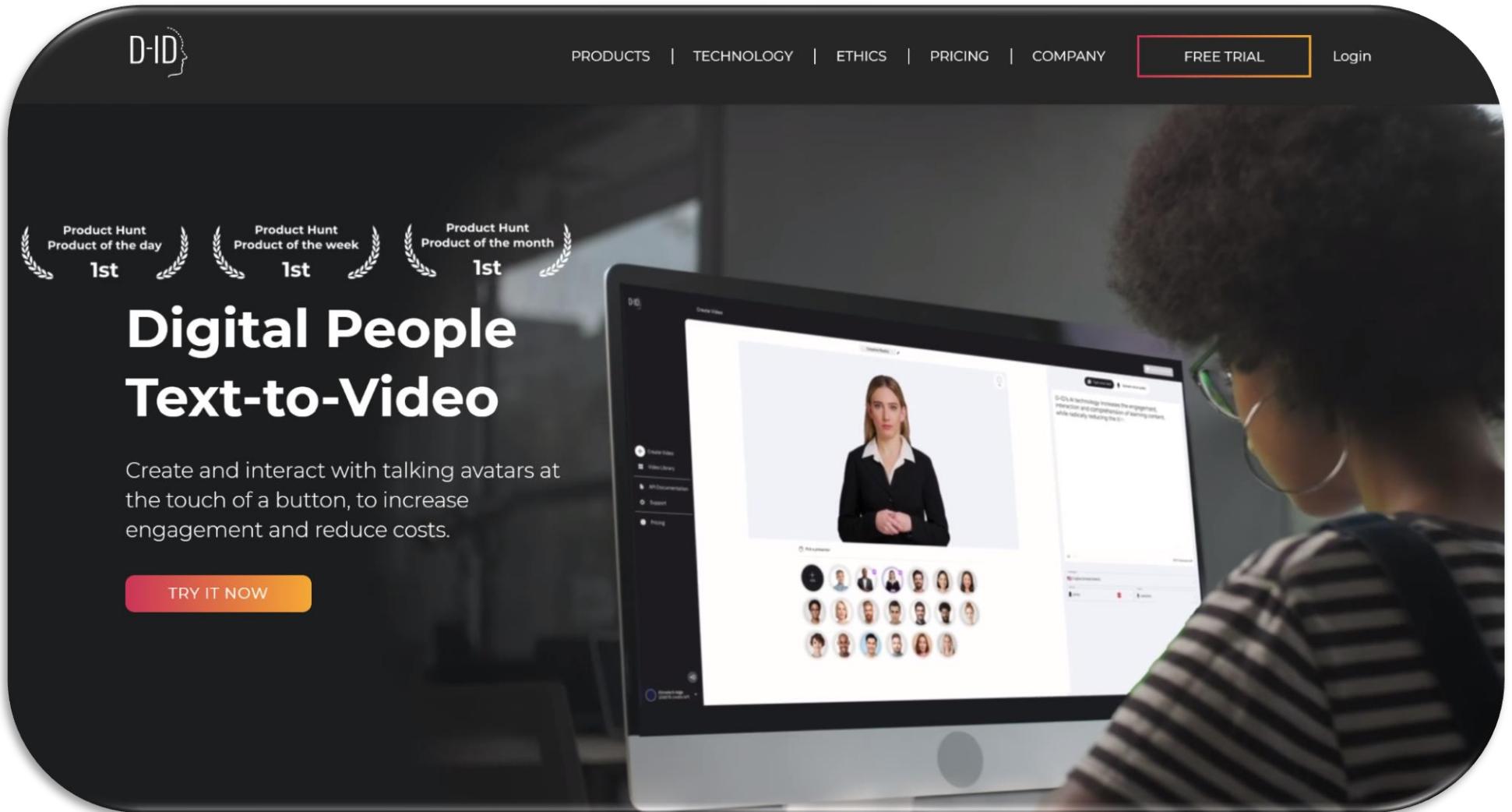
تبديل متن به تصویر: Midjourney



*Images generated on MidJourney by Henrique Centieiro and Bee Lee

ویرایش تصویر: clipdrop.co





تولید ویدئو: Sora

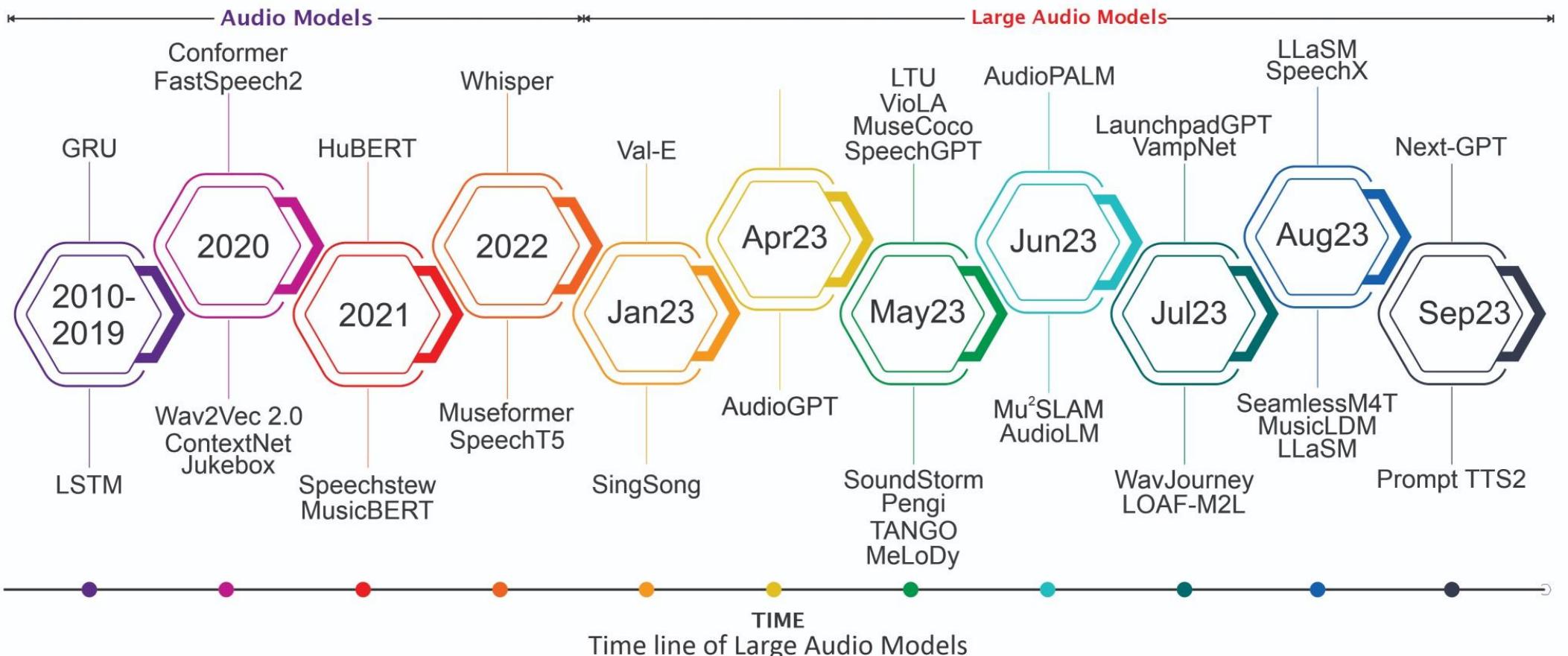


Prompt: A stylish woman walks down a Tokyo street filled with warm glowing neon and animated city signage. She wears a black leather jacket, a long red dress, and black boots, and carries a black purse. She wears sunglasses and red lipstick. She walks confidently and casually. The street is damp and reflective, creating a mirror effect of the colorful lights. Many pedestrians walk about.

نخستین فیلم سینمایی
با هوش مصنوعی



مبدل ها: مدل های بزرگ صوتی





دانشگاه علوم فنی های میان رشته ای
دانشکده سالانه های بومی

تولید موسیقی

The screenshot shows the homepage of Soundraw, an AI music generator. The URL https://soundraw.io is visible in the browser's address bar. The page features a dark background with white and yellow text. At the top, there is a navigation bar with links for Create Music, License, FAQ, API for businesses, Pricing, Sign in, and a prominent yellow Sign Up button. Below the navigation, the text "AI MUSIC GENERATOR FOR CREATORS" is displayed. The main headline reads "Stop searching for the song you need. Create it." followed by the subtext "Royalty-free music, AI generated for you". A large yellow "Create music" button is centered below the subtext. On the left side, there is a small image of a glowing blue light effect, and on the right side, there is a close-up image of a piano keyboard.



دانشگاه علوم فناوری های میان رشته ای
دانشکده سالانه های بومی

تولید گفتار (صدا)

C https://vall-e.io ☆ ↴

Vall-E Ai

Neural Codec Language Models are Zero-Shot Text to Speech Synthesizers

Chengyi Wang*, Sanyuan Chen*, Yu Wu*, Ziqiang Zhang, Long Zhou, Shujie Liu, Zhuo Chen, Yanqing Liu, Huaming Wang, Jinyu Li, Lei He, Sheng Zhao, Furu Wei

Microsoft

Trusted by

OpenAI Microsoft intel Google OpenAI



دانشگاه علوم فنی های میان رشته ای
دانشکده سازمان های پژوهشی

تبديل گفتار (صدا)



About Studio Live Demo Join Us

Find your voice for virtual worlds

Customise your online identity with studio quality AI voice overs & real-time AI voice changing.

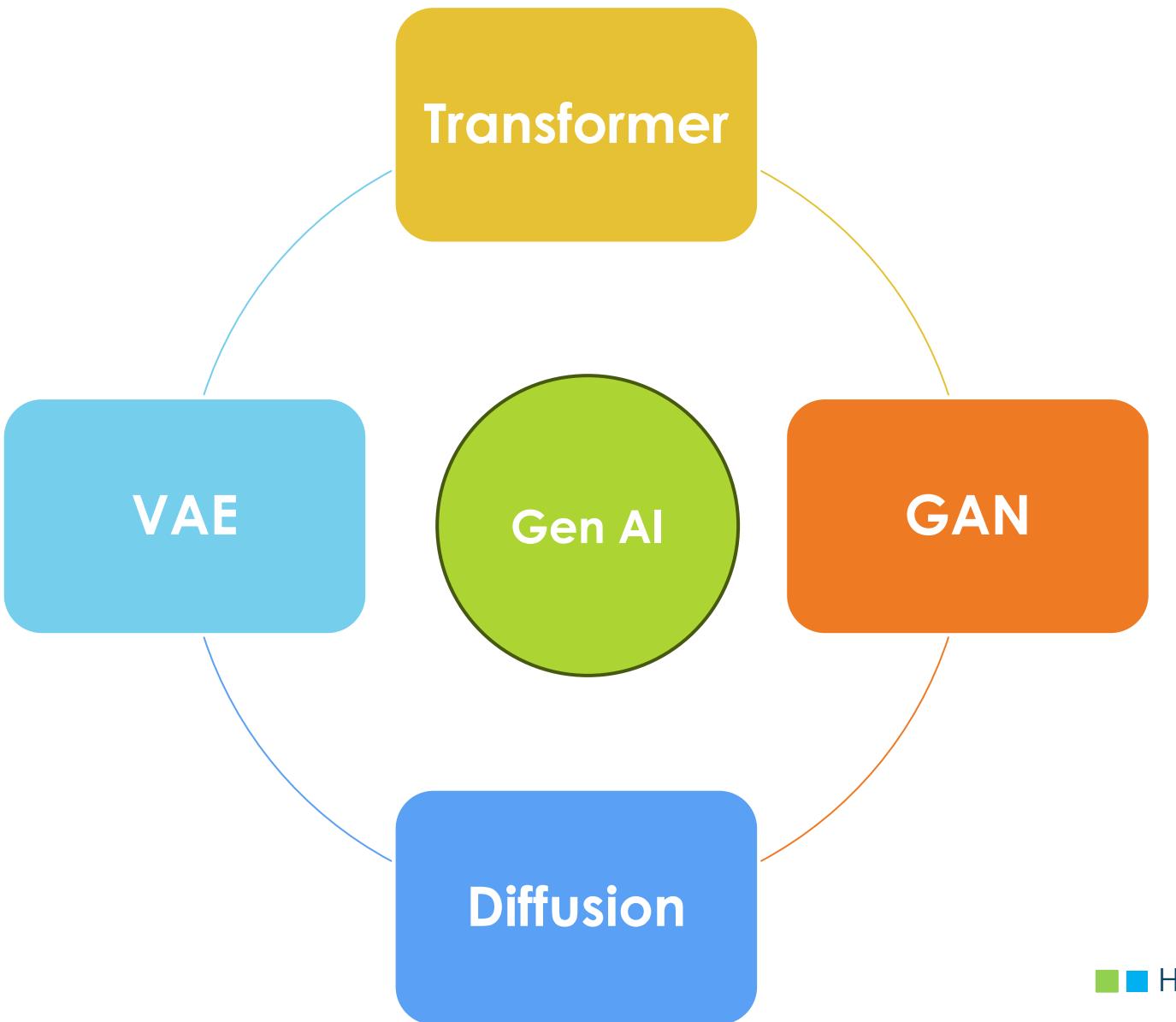
Convert Your Voice

■ Hadi Veisi (h.veisi@ut.ac.ir)

COMING
SOON

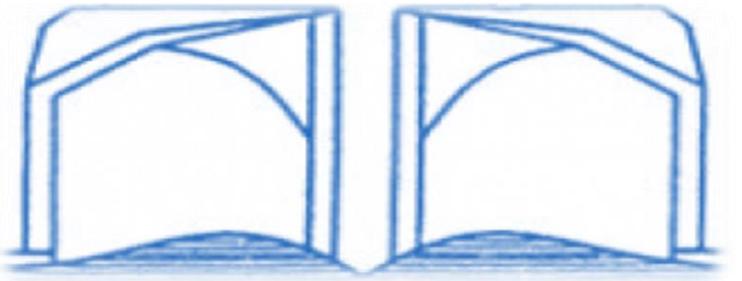
نسخه دیجیتال از مردگان

روش های کلیدی در هوش مصنوعی زاینده





دانشگاه علوم فنی های میان رشته ای
دانشکده سازمان های به شنبه



پرده آتی مدل های بزرگ(تر) و چندوجهی

هوش مصنوعی زاینده: محدودیت‌ها و ریسک‌ها

درباره هادی ویسی استاد دانشگاه تهران برای توضیح بده



Show drafts ▾

دکتر هادی ویسی، دانشیار گروه علوم سیاسی دانشکده علوم اجتماعی و سیاسی دانشگاه تهران هستند. ایشان در حوزه‌های رژیولینتیک، رئوکونومیک، منطقه‌گرایی، گردشگری، دولت محلی و حکمرانی شهری تخصص دارند.

دکتر ویسی ساقمه‌ی پژوهشی و آموزشی درخشنده دارد.

سوابق پژوهشی:

• مقالات:

- بیش از 42 مقاله در کنفرانس‌های معتبر داخلی و خارجی ارائه کرده‌اند.
- 25 مقاله در مجلات علمی پژوهشی به چاپ رسانده‌اند.

• کتاب:

◦ تالیف 5 کتاب در زمینه‌های مرتبط با تخصصشان

• مدارک تحصیلی:

- دکتری علوم سیاسی از دانشگاه تربیت مدرس
- کارشناسی ارشد علوم سیاسی از دانشگاه علامه طباطبائی
- کارشناسی علوم سیاسی از دانشگاه باقر العلوم

سوابق آموزشی:

- تدریس دروس مختلف در مقطع کارشناسی و کارشناسی ارشد در دانشگاه تهران
- راهنمایی و مشاوره دانشجویان تحصیلات تکمیلی
- ارائه سخنرانی در همایش‌ها و کارگاه‌های آموزشی

□ اطلاعات نادرست و توهمند (Hallucinations)

◀ تولید پاسخ با محتواهای مصنوعی و غیرواقعی

□ بایاس در یادگیری (Bias)

◀ یادگیری از روی داده دارای جهت‌گیری

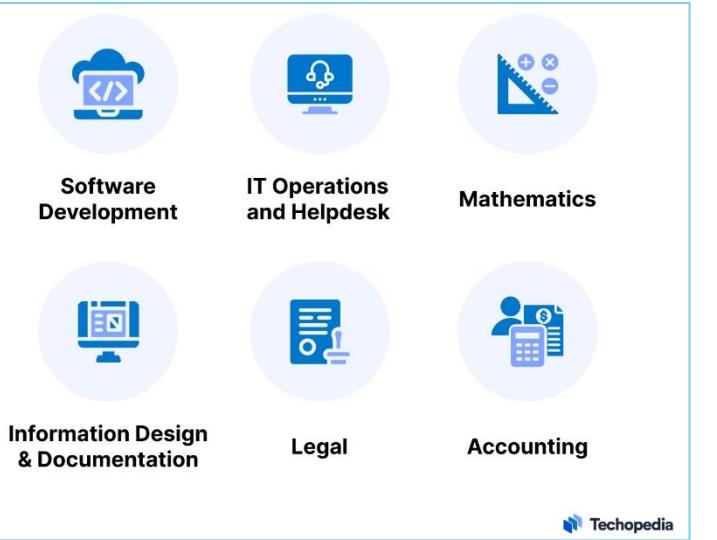
□ محرومگی و امنیت

◀ مانند چتبات‌های مبتنی بر ChatGPT

<https://zigap.ir/mag/chatgpt-ai-racism-study-african-american/>

■ Hadi Veisi (h.veisi@ut.ac.ir)

هوش مصنوعی زاینده: محدودیت‌ها و ریسک‌ها



□ جایگزینی شغل‌ها

◀ مانند جایگزین شدن چتبات در مراکز تماس

□ اثرات محیط زیستی (مثال: آموزش GPT-4)

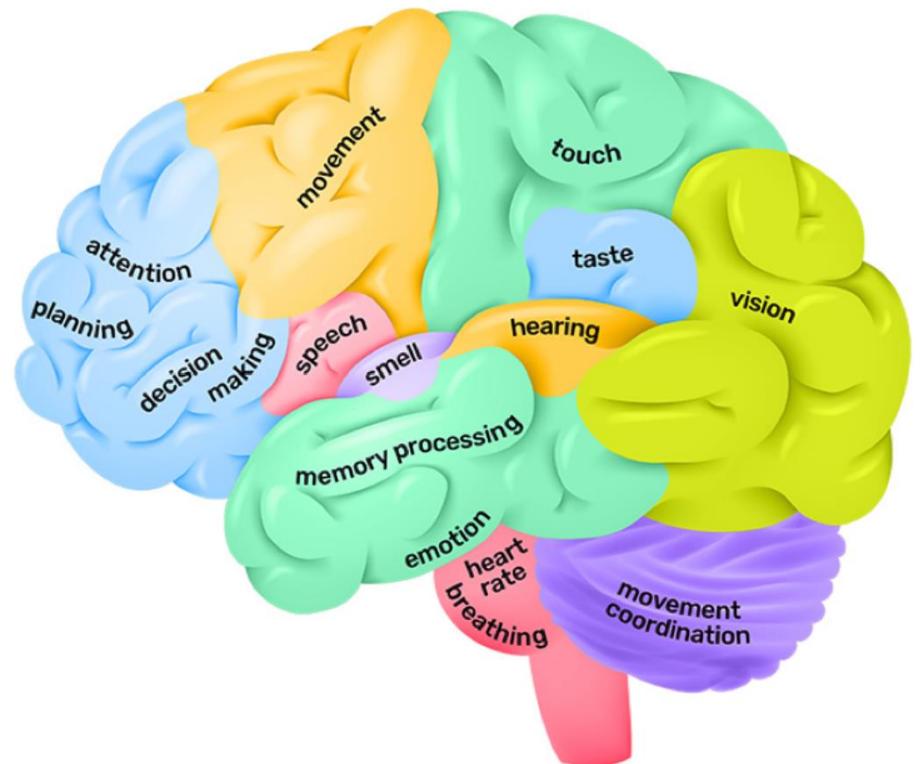
◀ ۲۵۰۰۰ جیپی یو A100 * ۴۰۰ وات (هر کدام) * ۱۰۰ روز

$$6,912,000 \text{ kilograms CO}_2 = 0.24 \text{ kg CO}_2 \text{ per kWh} = 24,000,000 \text{ kW/h}$$

◀ رانندگی با خودروی بنزینی به مسافت ۳۹ میلیون کیلومتر (۷۶ برابر فاصله زمین تا ماه) یا برقرسانی به خانه در یک سال

□ نقض مالکیت فکری

در مسیر تسخیر مغز

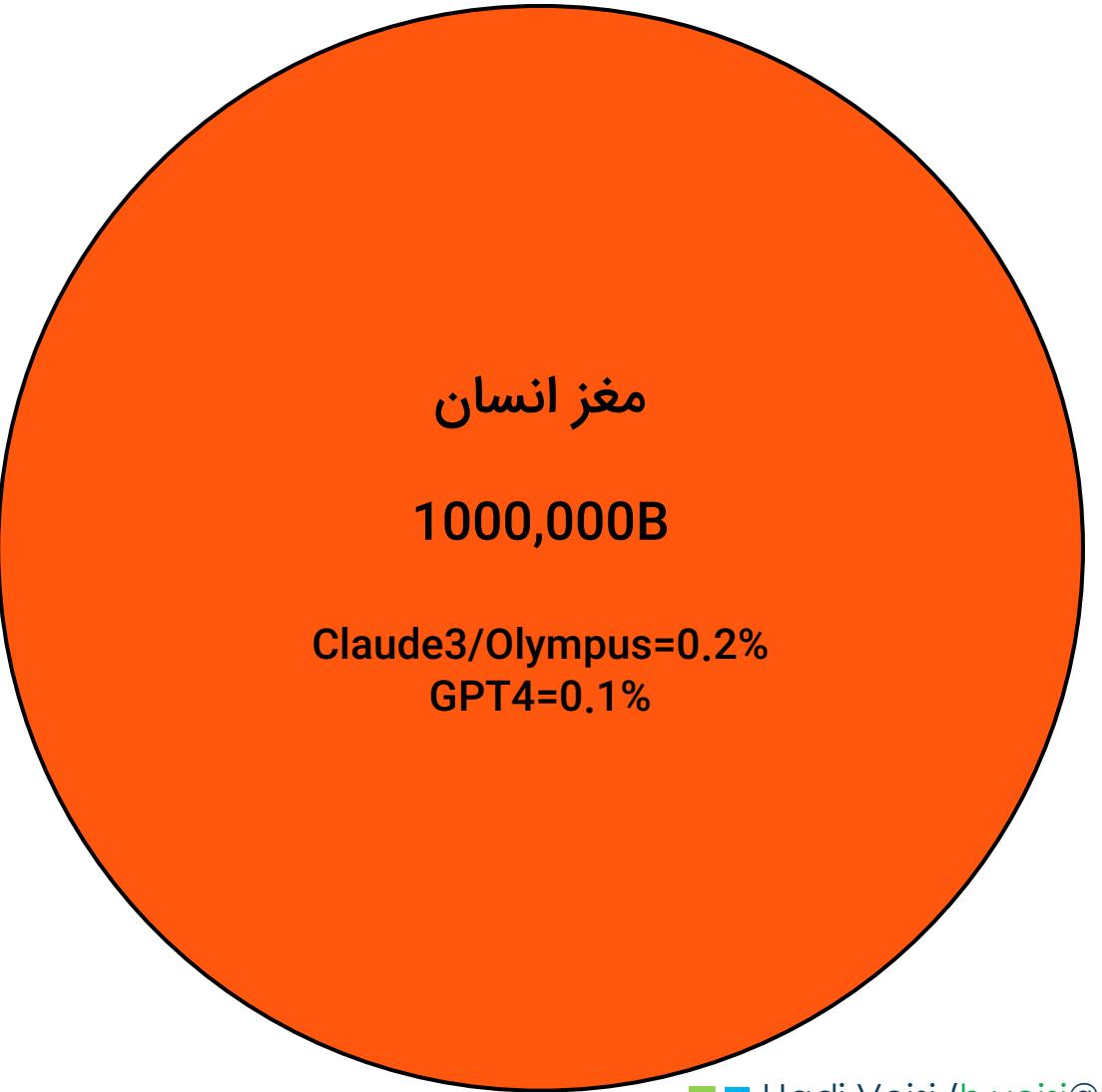
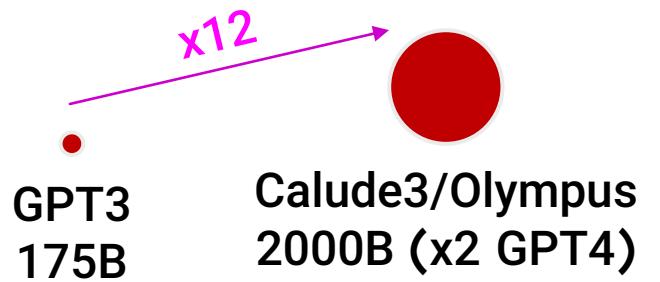


Region	Function	Tech	Progress	Verb
<i>Frontal lobe</i>	Memory processing, reasoning, regulating emotions...	GPT, BlenderBot 2.0...		PROCESS
<i>Parietal lobe</i>	Planning, attention, problem solving, decisions, morality...	MuZero, GPT-3, GPT-J...		DECIDE
<i>Parietal lobe</i>	Reading and comprehension...	GPT-3, GPT-J...		READ
<i>Temporal lobe</i>	Sense of time/space, taste, touch, link between functions of other lobes...	Tasty, Gastrograph AI...		SENSE
<i>Temporal lobe</i>	Hearing, speech...	General speech AI (Otter, Dragon, Synthesia...)		LISTEN/ SPEAK
<i>Temporal lobe</i>	Understanding, language, memory, learning...	BlenderBot 2.0, LaMDA...		LEARN
<i>Occipital lobe</i>	Vision and integrating visual information: colour, shape, distance...	CLIP, Wudao 2.0...		SEE
<i>Cerebellum</i>	Movement coordination, balance...	Robotics (Toyota, Boston Dynamics...)		MOVE
<i>Brain stem</i>	Body functions, heart rate, breathing, temperature...	General monitoring		MONITOR

Alan D. Thompson. August 2021. <https://lifearchitect.ai/>

General learning methods that can **scale** with computation are ultimately the most effective, Richard Sutton, The Bitter Lesson

مقیاس و دیگر هیچ!



خیز قهرمانی: چندوجهی پیش نیاز هوش مصنوعی عمومی



دانشگاه علوم فناوری های میان رشته ای
دانشکده سالانه های بومی

Company	Total funding	Mode	Multimodal features
OpenAI	\$14B	GPT-4V Text, image → text	
ANTHROPIC	\$3.4B	Claude-2 Text only	
Inflection	\$1.5B	Inflection-2 Text → text, image	
ALEPH ALPHA	\$642M	Magma Image → text	
cohere	\$469M	Text only	
AI21 labs	\$317M	Text only	
XAI	\$135M	Grok Text only	

مدل GPT5: بزرگ تر و چندوجهی



❑ پیش به سوی هوش مصنوعی عمومی
Artificial general intelligence (AGI) ◀

❑ در حال آموزش!

خودتان ببینید!

❑ نحوه تولید متن در روش های هوش مصنوعی زاینده

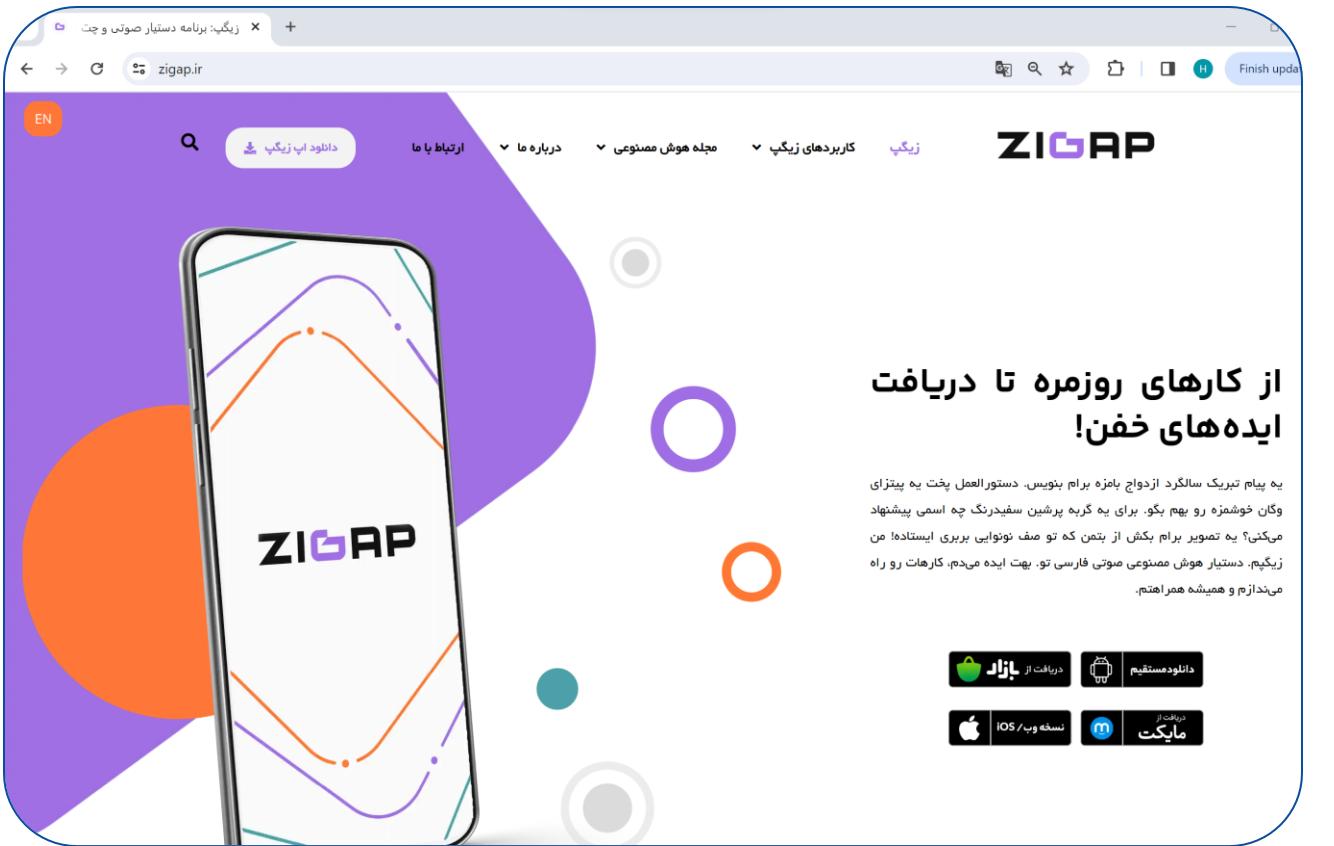
<https://zigap.ir/generative-ai/> ◀



با هم ببینیم

❑ تولید متن فارسی با ChatGPT در دستیار صوتی زیگپ

Zigap.ir ◀



■ Hadi Veisi (h.veisi@ut.ac.ir)