GAIA
დაიჯესტი

5 აპრილი, 2024
N5

გაეცანით AI-ის მთავარ სიახლეებს ამ დაიჯესტში და გაიგეთ მეტი

Grok-1.5 და ილონ მასკი OpenAI-ის წინააღმდეგ

Benchmark	Grok-1	Grok-1.5	Mistral Large	Claude 2	Claude 3 Sonnet	Gemini Pro 1.5	GPT-4	Claude 3 Opus
MMLU	73% 5-shot	81.3% 5-shot	81.2% 5-shot	75% 5-shot	79% 5-shot	83.7% 5-shot	86.4% 5-shot	86.8 5-shot
MATH	23.9% 4-shot	50.6% 4-shot	—	—	40.5% 4-shot	58.5% 4-shot	52.9% 4-shot	61% 4-shot
GSM8K	62.9 8-shot	90% 8-shot	81% 5-shot	88% 0-shot CoT	92.3% 0-shot CoT	91.7% 11-shot	92% 5-shot	95% 0-shot CoT
HumanEval	63.2% 0-shot	74.1% 0-shot	45.1% 0-shot	70% 0-shot	73% 0-shot	71.9% 0-shot	67% 0-shot	84.9% 0-shot

ილონ მასკისა და OpenAI-ის დაპირისპირება ბოლო პერიოდის აქტუალური საკითხია. თუმცა, ამ დაპირისპირების გარდა, ხელოვნური ინტელექტის სტარტაპ კომპანია xAI, რომელიც 2023 წელს ილონ მასკმა დააფუძნა, უკვე OpenAI-ის კონკურენტია. xAI-მ მარტის ბოლოს **xAI-მ თავისი ენის დიდი მოდელი Grok-1.5 გამოუშვა.**

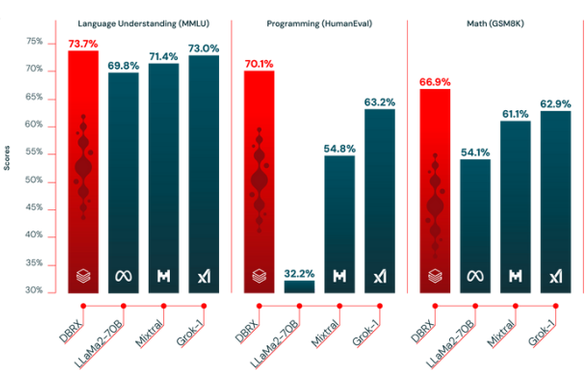
Grok ღია კოდის მქონე AI მოდელია - ჩატბოტი, რომლის მუშობის პრინციპი იგივეა როგორიც ChatGPT-ის, Gemini-ის ან Claude-ის, თუმცა **კონკურენტებისგან განსხვავებით მას აქვს წვდომა X-ის მონაცემებზეც**. Grok-1.5-ის ყველაზე მნიშვნელოვანი გაუმჯობესება არის მისი შესაძლებლობები და შედეგები **პროგრამირებასა და მათემატიკასთან დაკავშირებულ ამოცანებში.**

Grok-ის ტესტირებისას, Grok-1.5-მა მიაღწია 50.6%-იან შედეგს MATH ბენჩმარკზე და 90%-იან შედეგს GSM8K ბენჩმარკზე - ორივე მათემატიკური ბენჩმარკია, სადაც Grok მცირედით ჩამორჩება ისეთ ცნობილ ჩატბოტებს, როგორებიცაა: ChatGPT 4, Gemini, Claude 3 და ა.შ. დამატებით, **მოდელმა 74.1%-იანი შედეგი აჩვენა HumanEval ბენჩმარკზე,** რომელიც აფასებს კოდის გენერირებისა და პრობლემების გადაჭრის უნარს. აღნიშნულით კი **Grok გაცილებით უკეთესია, ვიდრე სხვა AI მოდელები, გარდა Claude 3 Opus-ისა**. შესაბამისად, Grok ახალი მოთამაშეა, რომელიც მომხმარებლებს, განსაკუთრებით დეველოპერებს სამუშაო პროცესში გაცილებით უკეთეს შედეგებს სთავაზობს.

Grok-1.5 სიახლეა ასევე მსჯელობის და კონტექსტის გაგების გაუმჯობესებული შესაძლებლობებიც. კერძოდ, მოდელს შეუძლია **128 ათას ტოკენამდე მონაცემთა ერთობლიობის** - დიდი მოცულობის დოკუმენტების დამუშავება ძალიან მცირე დროში.

Grok-1.5 ტესტირებისთვის მალე გახდება ხელმისაწყვდომი X-ზე. ნახეთ მეტი

DBRX: კიდევ ერთი ზოგადი AI მოდელი



Databricks-მა ახალი ზოგადი დანიშნულების ღია წვდომის AI მოდელი - **BBRX** წარადგინა, რომლის მუშაობის შედეგებიც **აღემატება** ისეთ მოდელებს, როგორებიცაა **GPT-3.5** და **LLaMA2**. DBRX განსაკუთრებულად ძლიერია პროგრამირებისა და მათემატიკური ამოცანების გადაჭრაში, მაგალითად **HumanEval** და **GSM8k** ბენჩმარკებზე აჭარბებს სპეციალიზებულ CodeLLaMA-70B მოდელსაც კი.

DBRX იყენებს **Mixture of Experts - MoE არქიტექტურას** და ეფუძნება **12 ტრილიონ ტოკენზე** დამუშავებულ მონაცემებს, რაც მას უფრო ეფექტურს ხდის როგორც დამუშავების სისწრაფის, ასევე ფუნქციონალების პარამეტრების რაოდენობის მიმართულებით.

შედეგად, მოდელის სიჩქარე 2-ჯერ აღემატება LLaMA2-70B-ის, ხოლო პარამეტრების რაოდენობა 40%-ით ნაკლებია ვიდრე Grok-1 მოდელის.

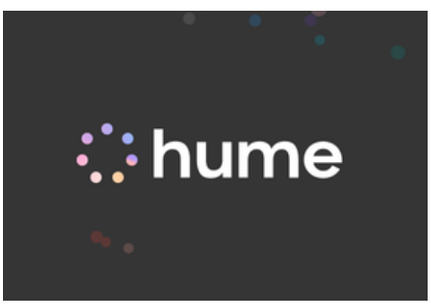
BBRX-ის მიზანია, განვითარებასთან ერთად, შექმნას მომხმარებლებისთვის ზოგადი დანიშნულების AI მოდელი, რომელიც გაამარტივებს სხვადასხვა სირთულის პრობლემის გადაჭრას, შესაძლებელს გახდის მრავალფეროვანი ამოცანების ძალიან სწრაფად და ეფექტურად შესრულებას, ვიდრე ეს დღეს არის შესაძლებელი. **DBRX ხელმისაწვდომია HuggingFace-ზე.** გარდა ამისა, Databricks-ის მომხმარებლებს შეუძლიათ მისი გამოყენება API-ების საშუალებით. ნახეთ მეტი

Hume - პირველი AI ჩატბოტი ემოციური ინტელექტით

Hume AI ჩატბოტია, რომელსაც ადამიანთან საუბარი **რეალურ დროში** შეუძლია, თუმცა მხოლოდ ეს არ არის. Hume პირველი AI მოდელია **ემოციური ინტელექტით**. ეს კი ნიშნავს, რომ მას ადამიანთან საუბრის პარალელურად პასუხებში ადამიანის **ემოციების ამოცნობა შეუძლია**. Hume მოსაუბრის ემოციური მდგომარეობისა და საუბრის თემის კონტექსტის გააზრებით შესაბამისი პასუხი აბრუნებს.

Hume **ემპათიური AI** ჩატბოტია, რომელსაც ადამიანის ემოციური მდგომარეობის ამოცნობის, გაგების და შესაბამისი პასუხის დაბრუნების უნარი აქვს, რასაც Empathic Voice Interface (EVI)-ის საშუალებით ახერხებს. კერძოდ, EVI-ს ესმის მომხმარებლის ხმის ტემბრი, რომელიც მნიშვნელობას აძლევს თითოეულ სიტყვას და მოდელი მომხმარებლის ხმოვან სიგნალებს იყენებს საკუთარი ენისა და მეტყველების მოსარგებად. **თქვენ შეგიძლიათ ესაუბროთ მას, როგორც ადამიანს და ისიც პასუხს გიბრუნებთ შესაბამისი ემოციებით, კონტექსტის გააზრებით და მაქსიმალურად ბუნებრივი სასაუბრო ტონით. Hume რეაგირებს უკეთესად, სწრაფად და ბუნებრივად, ვიდრე ნებისმიერი სხვა ხმოვანი AI ჩატბოტი.**

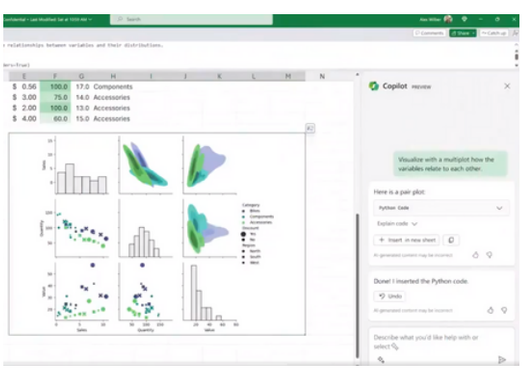
Hume-ის გამოყენება ძალიან ბევრ სფეროში შეიძლება, არამარტო როგორც პირადი AI ასისტენტის ფუნქციით, ასევე მაგალითად, **დეველოპერებს** შეუძლიათ გამოიყენონ EVI ნებისმიერი აპლიკაციისთვის მომხმარებელთან კომუნიკაციის პროცესში - გაბრაზებული მომხმარებლების დასამშვიდებლად ქოლ ცენტრში თუ ონლაინ ფსიქოლოგიური კონსულტაციისთვის.



ემპათიური AI ჩატბოტი დეტალურად ამუშავებს მომხმარებლის ხმოვან მონაცემებს და შეუძლია ემოციების ამოცნობა. მათ შორის, **შეუძლია ყურადღების, განწყობის, მენტალური ჯანმრთელობის და სხვა გარემოებების პროგნოზირება.** ამას კი გამოყენება შეიძლება ჰქონდეს, მაგალითად, სწავლებაში, მენტალურ მედიცინაში დიაგნოსტიკისა და პრევენციისთვის.

Hume საჯაროდ ხელმისაწვდომია, თუმცა EVI 2024 წლის აპრილში გახდება ხელმისაწვდომი. [ნახეთ მეტი]

>>> AI in Action: გამოყენების ქეისები



Microsoft-მა **Copilot Pro** წარადგინა. კოპაილოტი Microsoft-ის **მთელ ეკოსისტემაშია ინტეგრირებული** და ბევრ ახალ, საინტერესო შესაძლებლობებს აძლევს მომხმარებელს. **მთავარი სიახლეები:**

1. Copilot ახლა **iOS**-ზე და **Android**-ზეც მუშაობს

2. **ვორდში** დოკუმენტების **შეჯამება**, **შემოკლება** და **მარტივი Prompt-ებით ან სხვა ფაილების საფუძველზე ახალი ტექსტების შექმნა** შეგიძლიათ - მაგალითად პროდუქტის FAQ-ის მიხედვით ონბორდინგის გაიდლაინის შექმნა

3. შეგიძლიათ ექსელში **მონაცემთა ვიზუალიზაციის, მონაცემების დამუშავების (Transform) და ML ამოცანებისთვის** პითონის კოდის დაგენერირება **პროგრამირების** ცოდნის **გარეშე**, მხოლოდ Prompt-ით (სურ.)

4. **ფაილებიდან PPT პრეზენტაციების შექმნა** და გრძელი სლაიდების რამდენიმე მთავარ სლაიდად შეჯამებაც შესაძლებელი გახდა - როცა მთელი პრეზენტაციის წაკითხვა გვეზარება

5. **კორპორაციული GPT-ის შექმნა კოდის გარეშე**, Plugin-ებით და ჩაშენებული ინსტრუმენტებით ასევე შესაძლებელია - მაგალითად, თანამშრომლებისთვის კომპანიის ბიუჯეტის შესახებ ჩატბოტის შექმნა ფინანსური დოკუმენტების საფუძველზე

Copilot Pro $20 ღირს და მისი გამოყენება ოფისის პროგრამებში Microsoft 365-ის მომხმარებლებს შეუძლიათ. თუმცა, თუ Microsoft-ის ეკოსისტემას ძალიან აქტიურად არ იყენებთ ან სხვა ასისტენტის პრემიუმ ვერსიას უკვე კარგად იყენებთ, შეიძლება Copilot Pro-ს გამოწერა **არც ღირდეს** - [ნახეთ მეტი და თავად შეადარეთ]

>>> დაგეგმილი ღონისძიებები

Digital Design Festival Touch RGB

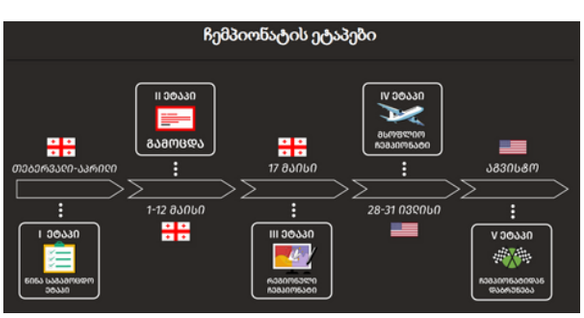
Touch RGB პირველი საერთაშორისო ციფრული დიზაინის ფესტივალია, სადაც საერთაშორისო მომხსენებლები პრეზენტაციების, პანელური დისკუსიებისა და ვორქშოპების ფორმატში ისაუბრებენ დიზაინში გამოყენებად ახალ ციფრულ ინსტრუმენტებზე, ხედვებსა და შემოქმედებითი პროცესების გამარტივების გზებზე. ფესტივალის მიზანია ადგილობრივი დიზაინერების პროფესიული უნარების განვითარებისა და თანამშრომლობის ხელშეწყობა. ღონისძიების მთავარი თემებია: UI/UX დიზაინი, გენერაციული დიზაინი (კრეატიული, AI, ახალი მედია) და იმერსიული დიზაინი.

თარიღი: 11 აპრილი
ადგილი: აღმაშენებლის გამზირი 127, ვასო აბაშიძის სახელობის ახალი თეატრი.

ღონისძიებაზე დასასწრებად საჭიროა [ბილეთის შეძენა].

>>> აღმოაჩინე შესაძლებლობა

Adobe მსოფლიო ჩემპიონატი დიზაინში



Adobe-ს მსოფლიო ჩემპიონატისთვის შესარჩევ კონკურსში რეგისტრაცია დაიწყო.
ვისთვის: 13-იდან 23 წლამდე ნებისმიერ ადამიანს, ვისაც აინტერესებს გრაფიკული დიზაინი
ეტაპები: შესარჩევი ეტაპი, Adobe-ის სერტიფიკატის აღება, რეგიონული ჩემპიონატი და საქართველო მსოფლიოზე აშშ-ში.
რამდენ ხანში: რეგიონულ ჩემპიონატზე მონაწილეებს ექნებათ 8 საათი ბრიფზე პრინტ ბანერების შესაქმნელად.

შესარჩევ ეტაპზე მონაწილეობისთვის საჭიროა [რეგისტრაცია]. **ნახეთ მეტი**