Gemini 4 Carbon tuli välja: kui tahad parimat tehisaru, pead rohkem maksma

14 min tagasi - 10.11.2026 Autor: AM

Tehnoloogiahiiud toovad uusi juturoboteid turule juba tihedamini kui pagarid sooje saiakesi. Enamasti lubatakse igal esitlusel maad ja ilmad kokku ning mõistagi on uus parem kui klunagi varem - loogiline. Kuid reaalsuses piirdub uuendus vaid pisut kiirema vastuse või mõne uue nipiga. Nüüd on aga Google lauale toonud uue mudeli, mida ei raatsita tavelise 30-dollarise kuiutasu eest kõigile jagada. Pärast eelmise kuu lõpus toimunud esmast eelvaadet on uusim tehisaru nimega Gemini 4 Carbon lõpuks jõudnud kasutajate ja arendajate käeulatusse.

Uus mudel ei tule siiski tühjale kohale ega asenda kõike eelnevat, vaid võtab koha sisse senise Gemini 4 Argoni kõrval. Kui teed Gemini äpis mudelivaliku lahti, vaatab sealt vastu selge tööjaotus: Argon on mõeldud koodi ja analüüsi jaoks ("*Code & analysis*"), samas kui Carboni taha on kirjutatud ei vähem ega rohkem kui "Kõige keerukam töö" ("*Most complex work*"). Ühtlasi tähendab uue raskekaallase saabumine seda, et vanem 3.1 Pro mudel saadeti äpist lõplikult pensionile.

Kui tavakasutaja jaoks on ligipääs seotud krõbeda 99-dollarilise kuutasuga AI Ultra paketis (mis avab Carboni nii Gemini äpis, otsingu tehisaru-režiimis kui ka Antigravity keskkonnas), siis arendajatele pakutakse Google AI Studio ja ärikliendi liidese kaudu sissejuhatavat hinda, mis on identne Argoniga: kaks dollarit miljoni sisend- ja kümme dollarit miljoni väljundsõna ehk *token*'i eest (alates detsembrist kerkib hind kahekordseks).

Maratonijooksja

Miks on Google'il üldse vaja kahte nii sarnase nimega mudelit korraga töös hoida? Google'i ametlik seisukoht tõmbab siin selge piiri:

> "Argon jääb ideaalseks valikuks enamiku tarkvaraarenduse ja teadmistepõhise töö jaoks, samas kui Carbon on loodud ülesanneteks, mis panevad proovile mudeli intelligentsuse piirid."

Ettevõtte sõnul kuuluvad selliste pähklite hulka näiteks "keerukas, mitmeastmeline programmeerimine ja vigade otsimine ning avatud lõpuga

teaduslik uurimistöö ja avastused".

Lihtsamalt öeldes: kui Argon on nagu terane ja kiire kontoritöötaja, kes teeb sinu süleris igapäevased tabelid ja koodijupid viuhti valmis, siis Carbon on nagu vaikne teadlasest maratoonar, kes lukustab end päevadeks kabinetti, närib end läbi tuhandetest lehekülgedest ja tuleb välja alles siis, kui on iga viimse kui detaili kolm korda üle kontrollinud. Carbon kulutab ühe vastuse peale märgatavalt rohkem mõtlemisaega ja ressursse, et teemat sügavuti uurida ning teha pikki enesekontrolli tsükleid.

Selleks, et sellist mõttejõudu suurte projektide juures päriselt ära kasutada, tutvustas Google oma Antigravity arenduskeskkonnas uut funktsiooni nimega */swarm* (parv). See muudab üheainsa sinu antud käskluse terveks koordineeritud tehisaru-agentide meeskonnaks, mille toimetamist saad ekraanil reaalajas jälgida. Arendajad saavad määrata Carboni kõige suuremat ajumahtu nõudvatesse rollidesse – näiteks süsteemi arhitektuuri loojaks, tööde koordineerijaks või audiitoriks, samal ajal kui kiiremad ja kergemad mudelid teevad paralleelsetes tööruumides ära musta töö.

Ühe muljetavaldava näitena suutis selline Carboni agentide parv luua vaid ühe viiba põhjal ajaloolistest fotodest ja inseneerijoonistest interaktiivse ruumilise *Three.js*-näituse Eiffeli torni ehitamisest aastatel 1887–1889.

Kas maailma targim või lihtsalt väga osav?

Turundusjutt on üks asi ja sõltumatud jõudlustestid hoopis teine. Kui vaatame suurt pilti, on Gemini 4 Carbon tõepoolest tõusnud tippmudelite edetabelite absoluutsesse tippu, kuid päris üksinda siiski seal troonil ei istu.

Sõltumatu platvormi Artificial Analysis laiapõhjalises intelligentsusindeksis (Intelligence Index v5) haaras Carbon esikoha 59 punktiga, edestades napilt oma suurimaid rivaale Claude Opus 5.5 (58 punkti) ja GPT-6 Astrat (57 punkti).

Carboni tõeline muskel tuleb välja just loogilises arutluskäigus, teabe otsimises, teadusülesannetes ja töövoogude automatiseerimises. Samuti särab mudel finants- ja õigusvaldkonnas: Balyasny üldise finantsintelligentsuse testis edestas see kõiki 23 tippkonfiguratsiooni. Google lõi koos Artificial Analysisiga ka uue multimodaalse teadmusagendi testi (MKA), mis hindab, kuidas tehisaru suudab pureda tihedaid faile, mis kubisevad graafikutest, joonistest ja tabelitest. Seal saavutas Carbon Fable 5.2 järel teise koha, kuid tegi seda võitjast enam kui kuus korda odavama ülesandepõhise kuluga.

Sarnast rahakotisõbralikkust näitas mudel ka CursorBench 4.0, Perplexity WANDR-i ja Zapieri AutomationBenchi testides, kus võeti esikoht enam kui poole väiksema kuluga kui järgmisel parimal välisel mudelil.

Kui sa aga otsid universaalset võluvitsa igaks tööks, tasub hoog maha võtta. Nimelt ei ole Carbon kaugeltki igas kategoorias võitmatu. Puhtalt programmeerimisagentide mõõduvõttudes (nagu Coding Agent Index, Terminal-Bench 4.0 ja FrontierCode v1.1) peab Google'i uus lipulaev tunnistama Anthropic'u Claude Opus 5.5 paremust. Samuti jääb see tarkvaraarenduse testis FrontierSWE 2 alla GPT-6 Astrale ning Vals AI täpsusindeksis napsab esikoha hoopis Claude Sonnet 5.5 (70,85% vs Carboni 70,36%).

Seega on aus öelda, et Carbon kuulub küll vaieldamatult maailma kahe parima tehisaru hulka, kuid koodikirjutamise spetsiifilistes nurkades on konkurentidel endiselt kerge edumaa.

Leiab turvaaugu paari tunniga

Üks valdkond, kus Gemini 4 Carbon aga tõeliselt hirmuäratavat võimekust näitab, on küberturvalisus.

Juba kaks nädalat enne avalikku esitlust anti mudel Fairwindi programmi kaudu valitud küberturbeekspertide käsutusse, ja tulemused räägivad enda eest. Carbon juhib edetabeleid nii turvanõrkuste leidmises (SEC-bench Pro), nende ärakasutamise simuleerimises (ExploitGym) kui ka lappimises (CWE-bench v1).

Google'i enda autonoomses turvasüsteemis Project Fortify saavutas Carbon vigade skaneerimisel 100-protsendilise leidlikkuse ja ohuhinnangute andmisel 95-protsendilise täpsuse. Avatud lähtekoodiga tarkvara testimisel parandas uus mudel senise Argoni niigi tugevat tulemust enam kui veerandi võrra, tuues päevavalgele varem märkamatuks jäänud loogikavead.

Näiteks suutis Google'i pilveturbe uurimismeeskond (CVR) Carboni abiga tõestada, et kaks pealtnäha tavalist viga ühes populaarses vabavaralises programmis olid tegelikult kriitilised ja ärakasutatavad turvaaugud, mis olid oma ohtlikkuselt võrreldavad kurikuulsa PHP2Shelli ründega.

Veelgi kõnekam juhtum pärineb programmist Wiz Scan for Good, kus Carbon pandi proovile avaliku digitaalse taristu kaitsmisel. Kahe suure pilveteenuse (SaaS) pakkuja kontrollimisel leidis tehisaru senitundmatud ehk nullpäeva turvaaugud vähem kui kahe tunniga. Töö, mis varem nõudis inimestest turvaekspertidelt nädalatepikkust ränka kaevamist, suruti kokku üheks

pärastlõunaks.

PLUSSID

- Edetabelite liider üldises intelligentsuses, teaduslikes ülesannetes, finantsanalüüsis ja töövoogude automatiseerimises.
- Suudab iseseisvalt leida ja lappida kriitilisi turvaauke tundidega, mitte nädalatega.
- Soodne ülesandepõhine kulu arendajale: pakub paljudes testides tipptulemust konkurentidest kaks kuni kuus korda madalama hinnaga ühe ülesande kohta.
- Suudab Antigravity keskkonnas juhtida tehisaru-agentide parve ja luua terviklikke 3D- ning veebilahendusi.

MIINUSED

- Gemini äpis ja otsingus kasutamiseks on vajalik 99-dollariline AI Ultra kuupakett.
- Jääb mitmes spetsiifilises programmeerimisagentide testis alla peamisele konkurendile Claude Opus 5.5-le.
- Põhjalikum mõtlemisprotsess tähendab, et mudel kulutab Argoni-sarnaste kiiremate mudelitega võrreldes rohkem sõnaühikuid (*token*'eid) ja aega.

Google Gemini 4 Carbon (gemini-4-carbon)

Kättesaadavus:	Gemini äpp, Google Search AI Mode, Google Antigravity, Google AI Studio, Vertex AI, Gemini for Business
Hind tavakasutajale:	99 USD kuus (Google AI Ultra pakett)
API sissejuhatav hind (kuni 30.11.2026):	2 USD / 1 miljon sisendsõna (input tokens), 10 USD / 1 miljon väljundsõna (output tokens)
API tavahind (alates 01.12.2026):	4 USD / 1 miljon sisendsõna, 20 USD / 1 miljon väljundsõna

Google Gemini 4 Carbon (gemini-4-carbon)

Kontekstiaken ja väljund:	Kuni 1 miljon token'it (ühtne sisend, mõtlemine ja väljund)
Artificial Analysis Intelligence Index v5:	1. koht (59 punkti; võrdluseks Claude Opus 5.5 sai 58 ja GPT-6 Astra 57 punkti)
Artificial Analysis Coding Agent Index:	2. koht (59 punkti; liider on Claude Opus 5.5 62 punktiga)
Terminal-Bench 4.0 ja FrontierCode v1.1:	2. koht (vastavalt 64% ja 51%, mõlemas Claude Opus 5.5 järel)
CursorBench 4.0, Perplexity WANDR, Zapier AutomationBench:	1. koht (enam kui 2x madalama ülesandepõhise kuluga kui järgmine parim väline mudel)
Multimodal Knowledge Agent (MKA) test:	2. koht (Fable 5.2 järel, kuid üle 6x odavama ülesandepõhise kuluga)
Vals AI täpsusindeks:	2. koht (70,36%; jääb napilt alla Claude Sonnet 5.5-le, kuid edestab Claude Opus 5.5-t)

- Uudised
- Tehisintellekt

Pilt

