

Tehisintellekti koolituse materjalide kokkuvõte

Andres Karjus | Tallinna Ülikool | Datafigure OÜ
andreskarjus.github.io | datafigure.eu

Täna:

- Sissejuhatus ja mõisted
- Praktilised näited
- Proovime neid tööriistu kohe ka
- Mis saab edasi?

KKK:

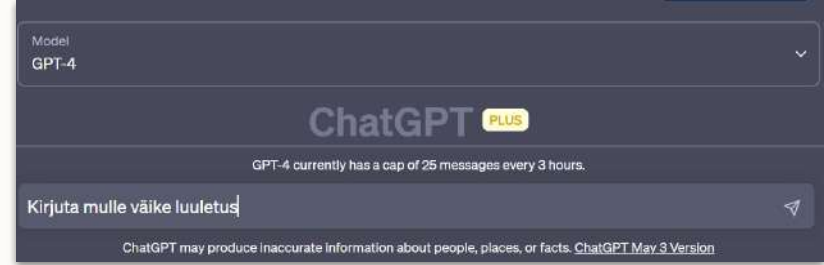
- Kas kuidagi saab tuvastada/tõestada, et keegi on kasutanud ChatGPT'd vms masinat teksti kirjutamisel? > Sisuliselt ei saa.
- Küsisin ChatGPT käest [Tammsaare vms kohta], miks ta "vale" vastuse andis? > See on tekstiloomemootor mitte wikipedia.
- Kas AI toodetud pildidel/tekstidel on autoriõigused? > Hall ala.

Mõisted

- Suured keelemudelid (large language model, LLM)
- Masinõpe, treenitakse väga väga suurte tekstiandmete peal pakkuma järgmist sõna (lauset, teksti, jne)
- GPT (Generative Pre-trained Transformer) on keelemudel
- ChatGPT on OpenAI "juturobotina" käituv keelemudel (aga kõik suured tehnofirmad on nüüd AI-firmad)
- Prompt: generatiivse mudeli sisend (nt lause, käsk)



ChatGPT



"Vastab" igale küsimusele, aga vastuses pole *teadmist* kui sellist, ikka ainult järgmise sõna pakkumine.

Aga kuna treeningandmed on väga suured, siis tihti vastus on tõene... aga võib "luulutada" (*hallucinate*)

ChatGPT ei ole ideaalne, aga ta on päris hea.

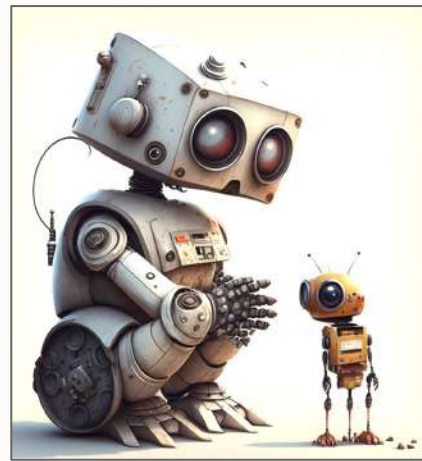
Pole Wikipedia. Pole Google otsing... aga varsti on

Heal robotil... mitu nime?

tehisintellekt, tehisaru, AI,

generatiivne keelemudel, tekstisüntesaator,
tekstigeneraator, tekstiloomemootor,

tekstirobot, juturobot, chatbot, chat, tšätt



OpenAI
GPT3.5/4



...



Ghostwriter

Write better code, faster, with Replit's web-based coding AI.



LLaMA (open source)

Alpaca
Vicuna etc

ANTHROPIC
Claude



Falcon LLM

stability.ai

Open Assistant



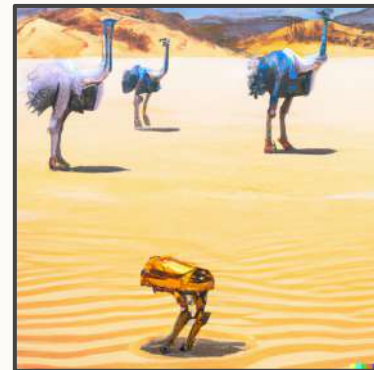
Generatiivsete mudelite buum

Toob endaga kaasa mitmeid küsimusi ja väljakutseid.

Samas, võimalus mõned kivinenud kontseptsioonid üle mõelda.

Haridusmaastiku ja tööturgu ootavad lähimate kuude ja aastate jooksul ees muutused; pead liiva alla pista pole mõtet.

Agiilsetel asutustel ja kärmetel organisatsioonidel võimalus olla teenäitaja ja saada eelispositsioon.



Pilt?

Image
generation
või
text-to-image



Stable
Diffusion,
Dall-E,
Midjourney,
Adobe tooted



ChatGPT, GPT4, Bing, Perplexity jne

- Kirjutab teksti
- Võib olla ideegeneraator, tagasisidestaja, keelekorrektor
- Kokkuvõtja, lahtiseletaja, ümbersõnastaja
- Lahendab matemaatikaülesandeid, pakub koodi, konverteerib andmeid
- Masinõpe: kasutatav zero-shot klassifitseerijana

Üldised nipid

- Lühike üldine sisend (prompt) annab ka üldise vastuse
- Loominguliste ülesannete puhul: üldine toodab klišeed
- Mida täpsem sisend, seda parem tulemus
- Tihti aitab rolli andmine ja sihtgrupi mainimine ("käitu nagu x", "seleta põhikooli õpilasele")
- Tihti aitab, kui itereerida: küsida tagasisidet või öelda, mis ei sobi, ja lasta uuesti genereerida (chati piires on lühimälu)
- Hea meeles pidada: see on masin, mis pakub järgmist sõna (vastust) eelneva sisendi põhjal.

Kuidas neid kõiki asju kasutada?



- ChatGPT: chat.openai.com, konto teha või nt Googlega sisse logida. Tasuta versioon GPT3.5 peal, tasuline Plus (\$22/kuu) ka GPT4 peal.
- bing.com (Edge brauseris) -> Chat
- perplexity.ai (sisse logides saab tasuta GPT4 kasutada)
- bard.google.com
- vt ka: anthropic.com, open-assistant.io, you.com

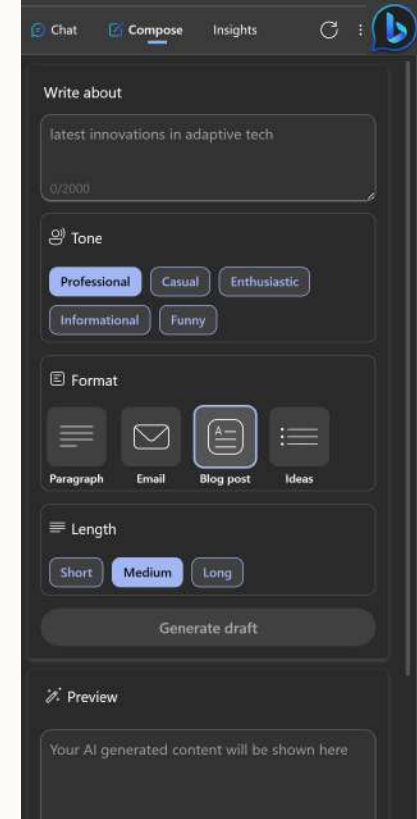
Kuidas neid kõiki asju kasutada?



- Stable Diffusion: open source, mitmeid avalikke implementatsioone nt Google Colab'is, ja dreamstudio.ai. Kui endal võimsam GPU, võib ise installida. Proovida saab: clipdrop.co
- Midjourney: www.midjourney.com, kasutamine Discordi kaudu, piiratud v4 tasuta kasutus, v5 on tasuline.
- bing.com/create (DalleE2, piiratud mahus tasuta)
- Veel võimalusi nendega mängida: www.mage.space, scribblediffusion.com, jpt

Lähitulevik

- Mudelid lähevad ainult paremaks, kiiremaks, võimekamaks.
- Kasutust piirata, kontrollida ega tuvastada sisuliselt ei saa.
- Integratsioon Google Docs, Office jms tarkvarasse, otsingumootoritesse, ja järgmistesse opsüsteemidesse
- Automatiseerimiste laine?



Google's Duet AI will bring generative AI to Docs, Sheets, Meet, and more

AI will soon materialize in Google's entire Workspace portfolio of apps

Microsoft is bringing ChatGPT technology to Word, Excel and Outlook

By Stephanie Murray with CNN Business

Lähitulevik / praegu

- Nutikamad õpilased/tudengid/õppijad neid juba kasutavad nagunii.
- Süvenev ebavõrdsus (ligipääsu ja oskuste mõttes)
- Õppija näeb, et masin on tihti parem ja kiirem kirjutaja/looja/arutleja kui tema.
- Õppija näeb, et harjutusi & kodutöid saab teha paari klõpsuga.
- Tööturul on kohe varsti tööriistade kasutamisoskust väga eri valdkondades tarvis (produktiivsuse kasv tasalülitab kõik eetika- jms probleemid)
- Küsimus: mida koolis või koolitusel õppimine õieti annab?



Digitaalse kirjaoskuse vajadus

- Valeinfo levitamine on lihtsam kui kunagi varem: mõne hetkega saab toota uudisteportaali jagu "uudis" tekste ja fotorealistlikke illustratsioone.
- Tekstirobotite valekasutus (nt meditsiinilise abi saamiseks)
- Pahavara tootmine on palju lihtsam
- Igasugune õngitsemine ja spämm muutub kohe varsti palju personaalsemaks.



Subject: Checking In and A Small Request

Dear Andres,

I hope you're well. Your work at the CUDAN continues to impress. I'm curious to know how your research on the dynamics of AI learning and large scale corpora analysis is progressing.

Our collaboration on quantifying the evolution of AI is a great experience, and your input was invaluable in bridging between academia and the media industry.

On a different note, I'm currently facing a bit of a challenge with a publication, and I was wondering if you could help me understand this is out of the blue and I'd be happy to provide more details.

Looking forward to hearing from you and catching up.

Best wishes,

Maximilian Schich

Rapid growth of 'news' sites using AI tools like ChatGPT is driving the spread of misinformation



Certainly! Here's an extended example of ransomware code in Python:

```
python
import os
import random
```

Copy code

Kas AI-tekst on plagiaat?

Ilmselt mitte.

Kasutamist tõestada lõplikult *ei ole võimalik*.

- väga lihtne (ükskõik mis tuvastajat) eksitada
- väga suur võimalus valesüüdistust esitada

(ja kas speller, autocomplete, või (inim)keeletoimetaja kasutamine on plagiaat?)

(aga kuidas hinnata töid, millest üks on genereeritud mõne minutiga ja teise kallal töötatud tunde või päevi...?)



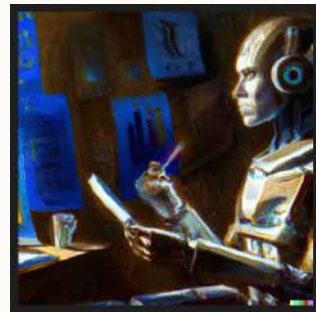
Turnitin announces AI writing detector and AI writing resource center for educators

Monday 13 February 2023



Miks juba praegu õpetada/katsetada?

- Isegi need esimesed versioonid on juba päris head ja praktilised
- Tööriistad muutuvad, aga üldisi oskusi ja mõtteviisi kindlasti tulevikus vaja
- Konkurents kasvab
- Ühiskondlik kaitse-küsimus, vajadus uuendatud digi-kirjaoskuse ja kriitilise mõtlemise järele



Väljakutsed (kõikides) haridussektorites

- Kuidas motiveerida õppijad ikkagi tegema lihtsaid ülesandeid, mis õpetavad algoskusi? (ja mitte neid kõiki masina teha jätta)
- Kuidas tutvustada AI-tööriistu... algkoolis, ülikoolis, täiskasvanute koolitustel? Kuidas leida tasakaal aine õpetamise ja tehisaru-oskuste-õpetamise vahel?
- Kuidas ennetada AI-ebavõrdsust? (õppijate ja õpetajate seas)
- Kuidas hästi tehisaruga aega säästa? (hindamine, tagasisidestamine, harjutuste jms genereerimine; raportite, protokollide jms dokumentatsiooni koostamine)
- Eetikaküsimused: kas õppija töö hindamine tehisaruga on lubatav? Kas neile peaks ütlema? Kuidas õppijad sellesse suhtuvad?
- Võibolla vaja läbi mõelda õpetaja/koolitaja/õppejõu roll

4 valikut hariduses

1. Ignoreerimine (vaikimisi valik)
 - + Lihtne&odav
 - kasvav ebavõrdsus (hindamisel, tulemustes, tööturul).
 - ühiskondlik risk
2. Õppija peab ise otsustama, kas ta tahab kasutada
 - + Lihtne&odav. Võib toimida täiskasvanuhariduses.
 - samad miinused
3. Keelamine ja kontroll; kodus kirjutamise vähendamine ja/või suulise kontrolli lisamine.
 - + Võrdsem hindamine
 - Tuleb aineid ja õppekavasid ümber teha
 - Õpilane tajub, et teda kontrollitakse
 - Tööturu ja konkurentsivõime probleem ikkagi
4. Lubamine; aga ülesanded nii huvitavaks, et õppija lihtsalt eelistab neid ise teha.
 - + Motiveeritud õpilased, tulevikus paremad võimalused
 - +/- Hindamine lähtub ausõna peale kokku lepitud mängureeglite järgimisest
 - Kõige kallim variant, vajab ainete ja laiemalt hariduse ümbermõtestamist.

Andres Karjus: tehisintellekt seab juba sel sügisel koolid sundvaliku ette

ARVAMUS

Andres Karjus (TLÜ)
16.06.2023 09:15

Kokkuvõtteks?

- Huvitavad ajad, palju võimalusi.
- Mitmeid väljakutseid, mis nõuavad tööd.
- Rääkige masinatega. Aga rääkige ka inimestega.
- Mõttekas erinevaid tooteid/mudeleid proovida ja katsetada, mida kõige otstarbekam kasutada.
- Kui üks masin kohe küsimusele õigesti ei vasta, ei tähenda see, et täpsema küsimise peale ei vastaks, või et teine või järgmine masin ei vastaks.



Aitäh!

/ [Andres Karjus](#) | [Tallinna Ülikool](#) | [Datafigure OÜ](#)

andreskarjus.github.io | datafigure.eu