Kā mākslīgais intelekts ietekmē uzņēmumu finanses? 1/19/24

Ģeneratīvais mākslīgais intelekts (“GenAI”) ir kļuvis par būtisku uzņēmējdarbības instrumentu, palīdzot uzņēmumiem optimizēt procesus, uzlabot efektivitāti un mazināt izmaksas. Tomēr, lai pilnībā izprastu GenAI ietekmi uz finansēm, ir svarīgi apskatīt šī instrumenta izmaksas no dažādiem aspektiem.

Katru reizi, kad GenAI modelis saņem vaicājumu (*prompt*), tas darbina lielos valodas modeļus, skaitļošanai izmantojot grafiskās procesoru vienības. Process, kurā ievadītos datus apstrādā, ir pazīstams kā dabiskās valodas apstrāde, un datus mēra žetonos (*tokens*). Katrs žetons atbilst apmēram četrām angļu valodas rakstzīmēm, un ar 1000 žetoniem var apstrādāt aptuveni 750 vārdus. Teksta ģenerēšanai GPT-4 modelim 1000 žetoni teksta ievadei izmaksā ap 0,03 USD un 1000 žetoni teksta izvadei izmaksā ap 0,06 USD. Piemēram, ja uzņēmumā ir 1000 darbinieku un katrs no tiem dienā veic desmit vaicājumus, katru 300 vārdu apmērā, ģenerējot 100 vārdu rezultātu, tad teksta ģenerēšanas izmaksas dienā būtu 198 USD – gadā ap 45 000 USD.

Izmaksu lielāko daļu rada GenAI modeļa pielāgošana (*fine-tuning*) konkrētajam uzdevumam vai domēnam, lai modelis ģenerētu vēlamo rezultātu. Piemēram, GenAI modelis, kas apmācīts veikt finanšu analīzi no gada pārskata datiem, ir jāpielāgo nodokļu piemērošanas izmaiņu dēļ. Modeļa pielāgošana ietver gan modeļa parametru pielāgošanu, gan jaunas datu kopas izmantošanu, lai padarītu modeļa izvades precīzākas attiecīgās valsts nodokļu kontekstā. Piemēram, OpenAI izmanto formulu, lai aprēķinātu kopējās modeļa pielāgošanas izmaksas.

Kopējās pielāgošanas izmaksas ir vienādas ar pamatmaksu par 1000 žetoniem reiz žetonu skaits ievades failā reiz modeļa pielāgošanai nepieciešamo iterāciju skaits. AI terminoloģijā iterācijas definē kā epohas (*epochs*). Piemēram, ja jāveic desmit iterācijas ar 50 000 žetoniem katrā un 1000 žetonu cena ir 0,008 USD (*gpt-3.5-turbo* modeļa apmācīšanas maksa), tad kopējās pielāgošanas izmaksas būtu:

10 * 50 000 / 1000 * 0,008 = 400 USD.

Lielo valodas modeļu infrastruktūras izbūve prasa ievērojamus ieguldījumus, tāpēc vairums uzņēmumu izvēlas izmantot mākoņpakalpojumus, lai uz tiem darbinātu GenAI modeļus. Mākoņpakalpojumu izmaksas iedalās vairākās pozīcijās: maksa par izmantošanas laiku, krātuves maksa, vispārējas nozīmes instances maksa, optimizētas skaitļošanas instances maksa, cena par gada abonēšanu. Lielākie publisko mākoņpakalpojumu sniedzēji šobrīd ir Amazon Web Services, Azure, Google Cloud Platform un Oracle, un katram no tiem cenas konkrētajā izmaksu pozīcijā atšķiras.

Izvēloties mākoņpakalpojumu arhitektūru GenAI modelim, galvenais apsvērums ir, vai tas tiks darbināts uz publiskā mākoņa, privātā mākoņa vai daudzmākoņu risinājumiem. Piemēram, medicīnas iestādēm ir svarīgi nodrošināt pacientu datu aizsardzību, tāpēc šķietami piemērotākas būtu privātās mākoņu krātuves, kas sākumā var izskatīties izdevīgākas par publiskajām, taču ilgtermiņā var novest pie papildu izmaksām to uzturēšanai un modeļu atjaunošanai.

Kopējās GenAI izmantošanas un pielāgošanas izmaksas var atšķirties atkarībā no vairākiem faktoriem, tostarp izmantotajiem pakalpojumiem un modeļa specifikācijām. Būtiski uzsvērt, ka viens no galvenajiem faktoriem GenAI modeļu veidošanā ir datu kvalitāte, jo bez augstas kvalitātes datiem GenAI modelis var prasīt vairākkārtēju pielāgošanu, tādējādi palielinot izmaksas un resursu patēriņu. Lai padarītu aprēķinus precīzākus, jāņem vērā arī darbaspēka apmācīšana GenAI lietošanā, datu drošības pasākumi un ētikas apsvērumi. Protams, svarīgi ir vienmēr konsultēties ar nozares ekspertiem, lai izsvērtu efektīvu un

uzņēmumam izdevīgu GenAI izmantošanu.

Lai uzzinātu vairāk par GenAI tehnoloģiju, šā gada 23. maijā aicinām apmeklēt *PwC's Academy* bezmaksas vebināru "Mākslīgā intelekta (GenAI) izmantošana praksē".