1.3 Atbild?ga MI izmantošana

Galven? doma

- **Atbildība ir dizaina uzdevums, nevis kontroles uzdevums.** MI atklāšana pienācīgi nedarbojas.
- **Un MI atklāšana ir netaisnīga, kad tā kļūdās.** Pastāv risks kļūdaini atzīmēt atbildes kā MI ģenerētas vai otrādi. Vairākas Eiropas valstis šos rīkus jau ir izslēgušas no lietošanas.
- **Pārliecinoša izdomāšana un mantotie aizspriedumi izriet no tā, kā šīs sistēmas ir veidotas un vērtētas.** Sekas jāizdara šādas: mācīt par tiem, nevis gaidīt labojumu.
- **Darbības princips ir: vispirms padomā, tad MI.** Liekot cilvēkiem apņemties savu atbildi, tas pārspēj skaidrojumu par mašīnas darbību; ierobežojumu iebūvēšana programmatūrā pārspēj tā aizliegšanu.
- **Vienas uzvednes vien nepietiek.** Tērzēšanas robots, kam uzdots darboties kā pasniedzējam, mēdz atgriezties pie tiešu atbilžu sniegšanas, kas padara uzticamus ierobežojumus par iegādes, nevis uzvednes veidošanas jautājumu.
- **Pastiprināšana ir noklusējuma stāvoklis; izlīdzināšana jāveido kopā no pieejas, dizaina un atbalsta** — un no šiem trim retākais ir apmācīts skolotājs.{{rev}}
- **Iekļaušana ir vispārliecinošākais ieguvums**, un tieši šeit paši skolotāji saskata visspēcīgāko argumentu.

> Ja no šīs lapas atceraties tikai vienu teikumu: **Dambji plīst. Iemāciet peldēt.**

Las?šanas dienasgr?mata, otrs skatiens

Atgriezīsimies pie mūsu devītās klases skolēna un viņa aizdomīgi izslīpētās lasīšanas dienasgrāmatas. Instinkts ir gandrīz universāls, un tas ir laba skolotāja instinkts: noskaidrot, vai viņš izmantoja MI. Uzbūvēt labāku dambi.

Šī nodaļa ir par to, kāpēc šis instinkts pats par sevi nekur neaizved. Mēs nonāksim pie tā, kas patiešām darbojas.

Sāksim ar neērto daļu. Skolotāji nespēj uzticami atšķirt MI radītu tekstu no skolēna rakstīta teksta. Vācijas pētījumā 289 dažādas pieredzes skolotājiem tika dots skolēnu un tērzēšanas robota rakstītu tekstu maisījums un lūgts to sašķirot. Nevienai grupai tas neizdevās. Taču abas grupas bija pārliecinātas, ka spēj to izdarīt. Pieredze gandrīz neko nemainīja. Plašākā mērogā situācija ir vēl sliktāka. Vienā universitātē pētnieki caur parastajiem kanāliem iesniedza tērzēšanas robota

ģenerētas atbildes reālajā eksāmenu sistēmā, nebrīdinot vērtētājus. 94% netika pat apšaubītas. Turklāt vidēji tās tika novērtētas augstāk nekā reālo skolēnu darbi. Kādas tad ir sekas? Mēs ķeramies pie mašīnas! Un tieši šeit atklājums pārstāj būt neveikls un kļūst par ētisku problēmu!

Detektors, kas k??daini sod? nepareizos skol?nus

Pētnieki septiņus no visplašāk izmantotajiem MI detektoriem pārbaudīja uz TOEFL esejām, ko rakstījuši skolēni, kuriem angļu valoda ir svešvaloda. Līdzās tam viņi izmantoja esejas, ko rakstījuši angliski runājoši astotās klases amerikāņu skolēni. Dzimtās valodas runātāju darbi tika klasificēti pareizi gandrīz vienmēr. No svešvalodā rakstītajām esejām vairāk nekā sešas no desmit tika kļūdaini atzīmētas kā mašīnas rakstītas.

Mehānisms nav noslēpumains. Šie rīki neatklāj lielo valodas modeļu darbu — tie atklāj paredzamību. Teksts, kas izmanto ierastus vārdus paredzamā secībā, tiek novērtēts kā mākslīgs. Ikviens, kurš raksta otrajā valodā, ar mazāku vārdu krājumu un drošākām teikumu konstrukcijām, rada tieši šādu signālu. Rīka kļūdas sistemātiski skar tos skolēnus, kuriem jau tā ir vissmagāk. Iestādes to ir izrēķinājušas. Vanderbilta universitāte norādīja, ka pat tad, ja šāds rīks kļūdītos tikai 1% gadījumu, tas joprojām nozīmētu aptuveni 750 nepatiesi apsūdzētus skolēnus gadā no tās 75 000 iesniegtajiem darbiem. Tāpēc tā šo rīku izslēdza. Vācijas federālās izglītības ministrijas norādījumi ir viennozīmīgi: tehnoloģija "nav pietiekami uzticama, lai nodrošinātu juridiski drošu pierādījumu".

Liela daļa Eiropas par šo vairs nediskutē. Francija, Nīderlande un Šveice iesaka izvairīties no detektoriem vai tos aizliedz. Apvienotās Karalistes eksāmenu regulators apzināti izvēlējies cilvēka spriedumu, nevis programmatūru. Nīderlandes arguments iet vēl soli tālāk un ir vērts atcerēties: skolēna darba ievietošana detektorā pati par sevi ir viņa personas datu apstrādes darbība, un to darīt bez atļaujas pārkāpj datu aizsardzības tiesību aktus. → 1.4 Tiesību aktiem atbilstoša MI izmantošana

Šķiet, ka krāpšanās panika ir trauslāka nekā ap to valdošais troksnis. Pētnieki aptaujāja skolēnus trijās vidusskolās pirms ChatGPT pastāvēšanas un vēlreiz — pēc tam. Krāpšanās apjoms kopumā palika tāds pats. Lielākā daļa skolēnu noraidīja domu, ka tērzēšanas robotam vienkārši būtu jāizveido viņu darbs, taču uzskatīja par godīgu to izmantot, lai sāktu darbu vai saņemtu skaidrojumu.

Patiesais jautājums nekad nav bijis "Kā es viņu pieķeršu?". Tas ir: ko mājasdarba uzdevums joprojām mēra, ja tā rezultātu var uzģenerēt pēc pieprasījuma?

Divas lietas, ko nevar salabot, t?p?c t?s j?m?ca

Pārliecinoša izdomāšana ir šo sistēmu iebūvēta īpašība. Tās rada ticamu teksta turpinājumu, kas nav tas pats, kas patiess turpinājums. Pētnieki tagad ir formāli parādījuši, no kurienes tas rodas:

standarta testi, ko izmanto lielo valodas modeļu vērtēšanai, atalgo pārliecinošu minēšanu vairāk nekā nenoteiktības atzīšanu, tāpēc, kamēr tie tiek vērtēti šādi, saglabāsies noteikts raiti izteiktu, labi noformulētu nepatiesību īpatsvars. Tā ir šo sistēmu uzbūves un vērtēšanas īpašība, nevis kļūda, kas gaida labojumu.

Arī apmācības materiālā ietvertie aizspriedumi ir iebūvēti. Turklāt tie nesamazinās, tehnoloģijai uzlabojoties. Eiropas pētījums par dzimumu stereotipiem dažādās valodās atklāja, ka lielāki modeļi — pat pēc papildu apmācības, kas paredzēta, lai tos padarītu drošākus un taisnīgākus — dažkārt stereotipizē vairāk, nevis mazāk. Ja kādu īpašību nevar tehniski novērst, tā kļūst par mācību vielu. Un to var iemācīt: vienā somu programmā to skolēnu īpatsvars, kuri spēja identificēt aizspriedumus MI sistēmas rezultātos, pieauga no 7% līdz 44%. Secinājums mācīšanai ir tāds, ka jāapmāca kļūdu analīze, nevis kļūdu novēršana. Citiem vārdiem — mašīnas kļūdas kļūst par mācību stundu.

Kas darbojas: dizains, nevis kontrole

Pierādījumi norāda uz vienu vienkāršu un viegli iegaumējamu principu: vispirms padomā, tad MI.

- **Liekiet cilvēkam vispirms apņemties savu atbildi.** Vairākos eksperimentos cilvēki, kuriem bija jāpieraksta sava atbilde, pirms redzēt mašīnas ieteikumu, pamanīja daudz vairāk tās kļūdu nekā tie, kuriem tika parādīti detalizēti skaidrojumi par to, kā mašīna spriedusi. MI darbības skaidrošana nepalīdzēja atrisināt šo problēmu, taču spriešanas piespiešana palīdzēja. Ar vienu godīgu atrunu: versijas, kas darbojās vislabāk, bija tās, kas dalībniekiem patika vismazāk, un tās visvairāk palīdzēja tiem, kam jau patīk piepūles pilna domāšana. Tātad tā ir arī parādība, kas pazīstama kā Mateja efekts: tiem, kam ir daudz, tiks dots vēl vairāk. Tieši viņi no šīs pieejas gūst vislielāko labumu.
- **Iebūvējiet ierobežojumus.** Tērzēšanas robots no 1.3. nodaļas, kas atbildes neatklāja un vietā uzdeva jautājumus, ir tā pati ideja programmatūrā: skolēni, kuri to izmantoja, eksāmenā nezaudēja neko, savukārt tie, kas izmantoja neierobežoto versiju, zaudēja krietni.
- Padariet redzamu procesu, nevis medījiet gala rezultātu. Melnraksti, darba piezīmes, īsa saruna par padarīto. Tie parāda, ko skolēns spēj izdarīt, nevienam nav jāpierāda, ko viņš patiešām ir darījis.

Dānija izmēģina šo principu: iesaistītajās skolās skolēni var izmantot MI, gatavojoties angļu valodas mutiskajam eksāmenam, savukārt rakstiskajā eksāmenā ir iekļauta ar roku rakstīta, bez palīglīdzekļiem pildāma daļa.

Pirms paļauties uz šo pieeju, vērts zināt divus ierobežojumus. Pirmais: populārais ieteikums, ka skolotājiem vienkārši jāuzdod tērzēšanas robotam darboties kā Sokrata metodes pasniedzējam, nepiepildās ilgākā sarunā. Vismaz pašlaik sistēma mēdz atgriezties pie tiešu atbilžu sniegšanas. Uzticami ierobežojumi ir jāiebūvē sistēmā, kas padara tos par iegādes lēmumu, nevis uzvednes veidošanas prasmi. Otrais: šķiet acīmredzami, ka MI, kas izskaidro savu spriešanu, ir drošības garants, taču skaidrojumi neuzticami novērš pārmērīgu uzticēšanos. Vienā pētījumā tie nemainīja cilvēku gatavību pieņemt nepareizu padomu, un pārāk tehniski vai pārāk vienkāršoti skaidrojumi var pat palielināt nepamatotu pārliecību. Skaidrojums ir priekšnoteikums, nevis risinājums.

Attieksme: divi neveiksmes veidi, nevis viens

Mēs uztraucamies, ka skolotāji pārāk uzticas MI. Taču mums vienlīdz jāraizējas arī par pretējo. Pat pieredzējušus cilvēkus var pievilt pārmērīga uzticēšanās. Vienā pētījumā skolotāji pārskatīja MI ģenerētas atzīmes. Viņi 42% no mašīnas atgriezeniskās saites atzina par neskaidru vai nepareizu, taču izmainīja tikai 9% no tās. Kļūdas pamanīšana un tās labošana ir divi atšķirīgi procesi.

Pretējais gals ir slēpšana. Skolotāji slēpj savu MI izmantošanu no kolēģiem un skolēniem, aizliedzot tiem darīt to, ko viņi paši dara: 77% privāti izmanto MI, kamēr 57% saka, ka skolēniem nekad nevajadzētu to izmantot ideju radīšanai. Tikmēr pētījums par akadēmiski spējīgiem 6. līdz 8. klases skolēniem atklāja, ka no trešdaļas līdz divām piektdaļām jau uzskata, ka MI zina vairāk nekā viņu skolotāji. Profesija, kas slēpj savu praksi, nevar būt paraugs atbildīgai praksei, un caurskatāmība nav tikai pieklājības jautājums. Tas ir mehānisms, ar kura palīdzību darbojas mācīšana ar piemēru.

Vai MI izl?dzina sp?les laukumu vai to sag?ž?

Ar darbu saistītos uzdevumos MI izlīdzina atšķirības. Vienā eksperimentā dalībnieki ar augstāku izglītības līmeni uzrādīja ievērojami labākus rezultātus nekā tie ar zemāku izglītības līmeni. Kad katram tika dots MI asistents, šīs atšķirības gandrīz pilnībā izzuda. Tad, atņemot asistentu, liela daļa atšķirību atgriezās. MI novērsa atšķirību rezultātā, nevis atšķirību spējās.

Mācīšanās uzdevumos notiek pretējais. Vienā pētījumā vidējā ieguvuma nebija vispār, taču palielinājās plaisa starp skolēniem, kuri jau daudz zināja, un tiem, kuri nezināja. Citā pētījumā tika aplūkoti skolēni, kuri rakstīja svešvalodā: vājākie rakstītāji no MI saņēma vairāk labojumu un veiksmīgi izmantoja mazāk no tiem, un vairāki uzskatīja atgriezeniskās saites plūdus par nomācošiem. Atgriezeniskā saite bez spējas to sakārtot nepalīdz — tā apbēdē.

Šī plaisa pastāv arī pirms jebkura rīka pieskāriena. Vācijā 80% skolēnu no vislabvēlīgākajām mājsaimniecībām uzskata MI par iespēju, salīdzinot ar 55% no vismazāk labvēlīgajām. Lielbritānijā šī plaisa iet cauri skolotāju istabai, nevis ierīču skapim: 45% privāto skolu skolotāju ir saņēmuši formālu apmācību MI izmantošanā, salīdzinot ar 21% valsts skolās. Valsts skolu sektorā turīgākas skolas atkal apsteidz nabadzīgākās. Aptauja par bērniem 20 Eiropas valstīs atklāj to pašu sociālo šķelšanos — gan cik daudz, gan cik daudzveidīgi viņi izmanto MI. Viena atziņa aptver visu problēmu labāk nekā jebkurš skaitlis: "turīgajiem ir pieejama gan tehnoloģija, gan cilvēki, kas palīdz to izmantot, savukārt nabadzīgajiem ir pieejama tikai tehnoloģija." Tas ir spēcīgākais pieejamais arguments par to, ka tieši skolotāji ir izšķirošie šīs plaisas mazināšanā.

Tātad izšķirošais faktors ir tas, vai skolēnam ir skolotājs, kurš ir apmācīts. MI kompensē atšķirības tad, kad kopā pastāv trīs lietas: droša piekļuve, mācīšanās vajadzībām veidots rīks (nevis vienkārši atbilžu sniegšanai), un skolotājs, kas atbalsta tā izmantošanu. Ja kāda no šīm trim lietām trūkst, MI atšķirības nevis mazina, bet pastiprina. Tā kā visas trīs parasti ir sastopamas vienās un tajās pašās skolās un iztrūkst tajās pašās skolās, pastiprināšana ir noklusējuma stāvoklis, un izlīdzināšana ir kaut kas, kas jāizveido apzināti.

Revision #12
Created 2026-09-01 09:14:31 UTC by Alexander
Updated 2026-09-03 13:50:29 UTC by Alexander