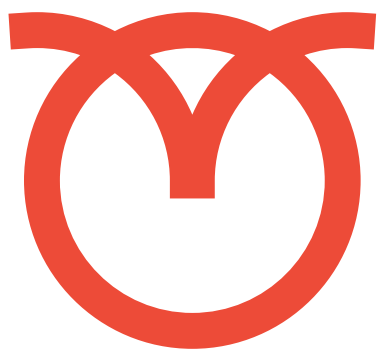


REZÜMÉ



OTDK

37. ORSZÁGOS
TUDOMÁNYOS DIÁKKÖRI
KONFERENCIA 2025



Orvos- és
Egészségtudományi
Szekció

PÉCS, 2025. ÁPRILIS 22-25.

Pécsi Tudományegyetem, Általános Orvostudományi Kar
Pécsi Tudományegyetem, Egészségtudományi Kar
Pécsi Tudományegyetem, Gyógyszerésztudományi Kar



OTDK'70

KOVÁCS FANNI

Általános orvos, Osztatlan
Semmelweis Egyetem Általános Orvostudományi Kar

TÓTH DALMA

Általános orvos, Osztatlan
Semmelweis Egyetem Általános Orvostudományi Kar

Témavezetők:

Baksa Barnabás PhD hallgató, SE ÁOK
Maurovich Horvát Pál egyetemi tanár, SE ÁOK

A ChatGPT-4o megbízhatóságának vizsgálata orvosszakmai szempontból koronária-CT angiográfia (CCTA) leletek esetén

Bevezetés: A mesterséges intelligencián (MI) alapuló nagy nyelvi modulok (LLM) megjelenése új lehetőségeket nyit a betegek számára az orvosi leletek értelmezésében. Azonban az LLM megbízhatósága és alkalmazhatósága az egészségügyben jelenleg nem tisztázott.

Célkitűzések: A ChatGPT-4o megbízhatóságának vizsgálata orvosszakmai szempontból koronária-CT angiográfia (CCTA) vizsgálatok leleteinek esetén.

Módszer: Retrospektív kutatásunkban CCTA leleteket elemeztünk. Az anonizált és metaadatot nem tartalmazó leleteket az OpenAI API Platformján keresztül a ChatGPT-4o modellbe tápláltuk. Elemzéseink során egyedi programkód alkalmazása mellett az alábbi utasításokat adtuk: 1. Parancs: Készíts összefoglalót. 2. Parancs: Tedd a beteg számára is érthetővé. 3. Parancs: Mit javasolsz a betegnek? Az eredeti leletek minőségét és a válaszok szakmai megbízhatóságát 2 orvos és 2 medikus értékelte, meghatározott kritériumrendszerű ötfokozatú skálán. Rögzítettük a dokumentumok szószámát és értékeltük a nyelvhelyességét is. ICC és t-próbákat alkalmaztunk.

Eredmény: Összesen 100 beteg leletét dolgoztuk fel és elemeztük a megadott Parancsok szerint. A leletek a konzervatívnak megfelelő arányban tartalmaztak különböző típusú és súlyosságú eltéréseket. A vizsgálok között kiváló ICC értékeket mértünk. Az eredeti lelet (4.87 ± 0.39), az 1. Parancs ($4.62 \pm 0.$), a 2. Parancs (4.54 ± 0.93) és a 3. Parancs 3 átlagpontszámai szignifikánsan nem különböztek (minden $p > 0.05$). A 2. Parancs nyelvi minőségét 4.92 ± 0.28 pontra értékeltük. Az MI jelentősen csökkentette a szószámot az 1. Parancs (-49%), 2. Parancs (-45%) és a 3. Parancs (-63%) esetén is az eredeti lelethez képest. MI hallucinációt nem találtunk. Az MI teljesítménye a graft alcsoportban a Parancs 1. és 2. esetében volt a legrosszabb minden értékelőnél (2.8 ± 1.25 , 2.8 ± 1.25 és 2.8 ± 0.84 , 2.5 ± 0.5 orvosi és medikusi értékelőknél ebben a sorrendben).

Következtetés: A ChatGPT-4o képes a betegek számára érthető módon, kiváló nyelvhelyességgel, lényeges szakmai hibák nélkül összefoglalót írni és releváns javaslatot tenni az orvosi lelet alapján. Az MI szakmailag káros vagy valótlan állítást nem tett, de teljesítménye a graft beültetésen átesett betegek leletei esetében rosszabb volt.

KOVÁCS KINGA JUDIT

Általános orvos, Osztatlan
Szegedi Tudományegyetem
Szent-Györgyi Albert Orvostudományi Kar

SZAKÁCS PÉTER

Általános orvos, Osztatlan
Szegedi Tudományegyetem
Szent-Györgyi Albert Orvostudományi Kar

Témavezetők:

Dr. Fodor Gergely egyetemi adjunktus, SZTE ÁOK
Dr. Peták Ferenc egyetemi tanár, SZTE ÁOK

Mesterséges intelligencián alapuló nagy nyelvi modellek alkalmazása az orvoképzés hatékonyságának javítására

Bevezetés: A mesterséges intelligencia rendszerek, mint a nagy nyelvi modellek (LLM) használata egyre nagyobb teret hódít mindennapjainkban, felhasználásuk egyre jelentősebb. Mivel egyre több hallgató is használ már ilyen eszközöket, így nem kerülhető meg tudatos alkalmazásuk az orvoképzésben sem. Tanulmányunkban azt vizsgáltuk, hogy a LLM-alapú chatbotok használata javíthatja-e az orvoképzés hatékonyságát.

Módszer: Vizsgálatunkban egy ChatGPT-alapú chatbotot készítettünk, melybe az „Introduction to medical informatics” tantárgy követelményeit vittük fel. Megvizsgáltuk a kurzus teljesítéséhez kapcsolódó kulcsinformációk megtalálásának hatékonyságát hagyományos forrásokkal (a tantárgy követelményeit Coospace-re feltöltve), illetve a chatbot segítségével. Tanulmányunkba 44 angol nyelvű képzésben tanuló elsőéves orvostanhallgatók két csoportját vontunk be: 21 fő a chatbot használatával, 23 fő hagyományos források használatával dolgozott. A vizsgálat során a tantárgykövetelményekre vonatkozó válaszok pontosságát és gyorsaságát Coospace-teszt segítségével értékeltük ki.

Eredmények: Chatbot használatával a hallgatók 32,4%-kal magasabb pontszámot értek el ($p < 0,01$). A válaszok helyességében 2 kérdés esetén volt szignifikáns különbség ($p < 0,01$). Ellenben a teszt kitöltésére fordított időben és az egyes kérdésekre való visszatérések számában nem találtunk szignifikáns eltérést ($p = 0,62$ és $0,77$).

Megbeszélés: Chatbot alkalmazásával hatékonyabb információszerezést észleltünk, melynek lehetséges magyarázata, hogy a nem anyanyelvükön tanuló, az egyetemmel még csak ismerkedő hallgatók számára a követelmények hagyományos szöveg alapú értelmezése kihívást jelenthet. A chatbottal a célzott kérdésekre célzott választ kapnak, melyekre pontosítás céljából visszakérdezhetnek. Eredményeink alapján javasolt a mesterséges intelligencia alkalmazása a modern orvoképzésben, mivel hatékonyan támogatja a személyre szabott tanulást, figyelembe véve a hallgatók egyéni igényeit. Ennek részeként javasolt olyan chatbotok fejlesztése és használata, melyek kifejezetten oktatási célokat szolgálnak, például a kikérdezés folyamatában