

Magyar páciensek narratív tapasztalatainak elemzése BERT témamodellezéssel és szentimentelemzéssel

Osváth Mátyás¹, Yang Zijian Győző², Kósa Karolina¹

¹ Magatartástudományi Intézet, Általános Orvostudományi Kar, Debreceni Egyetem,
osvath.matyas@med.unideb.hu

² Nyelvtudományi Kutatóközpont, Sightspot Network Kft.,
yang.zijian.gyozo@nytud.hu

¹ Magatartástudományi Intézet, Általános Orvostudományi Kar, Debreceni Egyetem,
kosa.karolina@med.unideb.hu

Kivonat: A pácienseknek a betegellátás során szerzett tapasztalatai fontos információkkal szolgálnak az ellátás problémáiról, amelyek elemzésével és felhasználásával az egészségügyi ellátás minősége javítható. A páciens-élményeket vizsgáló kutatások száma folyamatosan növekszik, azonban hazánkban kevés kutatás használ erre természetesnyelv-feldolgozáson alapuló módszereket. A jelen vizsgálat célja a hazai betegellátásban szerzett tapasztalatokról szóló írásokban megjelenő témák azonosítása és az írások érzelmi polaritásának meghatározása volt. Egy hazai online fórumon 2009 és 2020 közt megjelent, Web Scaper-rel letöltött 1663 blogbejegyzést elemeztünk témamodellezéssel és szentimentelemzéssel, előtanított huBERT és egy finomhangolt HIL-SBERT transzformer modell segítségével. A modell 326 és 200 témát azonosított, amelyeket manuális elemzéssel kategóriákká vontuk össze. A modell az írások 94,4%-át negatívan osztályozta, hasonló tendenciát mutatva az éves bontásra is. A szóbeágyazási BERT modellen alapuló témamodellezéssel eredményesen kinyerhetők a betegellátásban szerzett tapasztalatokat leíró szövegekben megjelenő mintázatok. Az eredmények minőségét azonban befolyásolja a korpusz nagysága, azaz a blogbejegyzések száma és terjedelme.

1. Bevezetés

A betegellátás teljesítményének egyik indikátora az ellátás minősége. A beteget közép-pontba állító, magas minőségű ellátás az egészségpolitikai ágazat és a betegellátást végző intézmények számára is fontos prioritás, melynek mérésére több, különféle indikátort tartalmazó elméleti keretrendszert fejlesztettek ki. Avedis Donabedian 1966-ban megjelent úttörő közleménye az egészségügyi ellátás minőségét a struktúra, folyamat és kimenet dimenziókba sorolható indikátorokkal javasolta jellemezni (Donabedian, 2005). A Gazdasági Együttműködési és Fejlesztési Szervezet (OECD) 2006-ban tette közzé az egészségügyi rendszer teljesítményének mérésére javasolt keretrendszerét többdimenziós mátrix formájában, amely az ellátás Donabedian-i minőségét továbbfejlesztve tartalmazza (Carinci és mtsai., 2015; Mattke és mtsai., 2006). A mátrix egyik fontos dimenzióját a betegtapasztalatok alkotják, amelyek a páciensekkel az ellátás során egyénekenként történő események eredményeként alakulnak ki, jelentősen

befolyásolva a pácienseknek az ellátásról alkotott véleményét. E tapasztalatok közé tartozik egyebek közt a személyzet és beteg közötti kommunikáció, az ellátáshoz való hozzáférés és várakozási idő. A páciensek tapasztalataitól el kell különíteni a szakirodalomban sokszor szinonimaként megjelenő, de nem azonos indikátort, a betegelégedettséget, amely a beteg előzetes elvárásai és a ténylegesen kapott ellátás közötti diszkrépancia mértékét mutatja (Jenkinson és mtsai., 2002).

A betegtapasztalatok nemzetközi felmérése és összehasonlítása érdekében az OECD összesen 11 konkrét, úgynevezett „Patient-reported experience” mutatót (PREM) határozott meg. Ezek olyan indikátorok, amelyek kiemelkedő jelentőségűek a páciens tapasztalata és az ellátás kimenete szempontjából egyaránt, mint például az orvosnak a pácienssel töltött elegendő ideje, a beteg megfelelő tájékoztatása az orvos által, a beteg bevonása a döntéshozásba, vagy a felírt, de anyagi okok miatt ki nem váltott gyógyszerek aránya (Brito Fernandes és mtsai., 2019). A konkrétan meghatározott indikátorok révén a betegellátás minőségi dimenziói nemcsak országokon belül, hanem országok közt is összehasonlíthatóvá válnak.

Az indikátorokhoz kvantitatív, például zárt kérdéseket tartalmazó kérdőívekkel, és kvalitatív módon, például szövegelemzéssel is lehet adatokat gyűjteni. A kutatások és rendszeres felmérések legnagyobb része kvantitatív módszereket alkalmaz (pl. az USA-ban a Hospital Consumer Assessment of Healthcare Providers and Systems, HCAHPS, (Giordano és mtsai., 2010), azonban az utóbbi időben egyre nagyobb az érdeklődés a páciensek narratív visszajelzései iránt is. Az USA-ban több szervezet működtet olyan honlapokat (RateMD, Healthgrades, Yelp), amelyeken páciensek szöveges visszajelzést adhatnak egy meghatározott intézményben kapott ellátásról, kvantitatív és kvalitatív formában. A brit Nemzeti Egészségügyi Szolgálat (NHS) online portálján is lehetőség van szöveges vélemény írására (Lagu és mtsai., 2013).

Nemzetközi kutatások szerint a betegeknek az orvosokról online írt értékelései és az azokra vonatkozó hozzászólások befolyással vannak az orvosválasztásra. A páciensek 59%-a tartja fontosnak az orvosokról írt értékeléseket (Burkle & Keegan, 2015; Hanauer és mtsai., 2014). Az online térben elérhető vélemények, értékelések és hozzászólások egyre nagyobb jelentőséggel bírnak, mivel számuk növekszik, és egyre többen használják ezeket az internetes forrásokat és értékeléseket a döntéshozáshoz.

Hazánkban több kutatás irányult páciens tapasztalatok felmérésére kérdőíves módszerekkel (Brito Fernandes és mtsai., 2019; Fernandes és mtsai., 2020; Lucevic és mtsai., 2019). Hazai és nemzetközi eredmények szerint is összefüggés van a páciensek pozitív tapasztalatai és az alap- és járóbeteg ellátás használata között (pl. kórházi felvételek száma és kórházba való visszatérések száma) (Doyle és mtsai., 2013; Price és mtsai., 2014). A magasabb betegelégedettséggel jellemzett kórházakban nagyobb mértékű a betegek adherenciája, kisebb a halálozás, és rövidebb időt töltenek az intézményben a betegek (Bidmon és mtsai., 2020; Nawab és mtsai., 2020).

A kvantitatív módszerek alkalmazása dominánsnak tekinthető a betegek tapasztalatainak és élményeinek értékelésére. Ezek a módszerek azonban az információvesztésből fakadóan szükségszerűen korlátozottak, és érthetőségük is alacsonyabb mértékű a dolgozók számára. A szöveges (narratív) visszajelzések jobban érthetőek az érintetteknek, de ezek elemzése sokkal bonyolultabb, az intézmény(ek) egészére vonatkozó következtetéshez pedig nagy számú narratívára van szükség. Ugyanakkor az ellátottaktól érkező szöveges visszajelzések részletes és alapos elemzése iránt egyre nagyobb az igény ezek információgazdagsága miatt (Khanbhai és mtsai., 2021) (Emmert és mtsai.,

2014; López és mtsai., 2012; Zhang és mtsai., 2018). A nagy mennyiségű, strukturálatlan szövegek azonban megnehezítik, sőt egy bizonyos volumenen túl lehetetlenné teszi a manuális elemzést. A probléma megoldását a természetesnyelv-feldolgozás mint a mesterséges intelligencia egyik részterületének robbanásszerű fejlődése hozta el, amely lehetővé tette nagy mennyiségű szövegek gépi elemzését.

1.1. Kapcsolódó irodalom

Greaves és munkatársai a páciensek hozzászólásait szentimentelemzéssel vizsgálták, valamint gépi tanuláson alapuló modelleket alkalmaztak különböző kategóriák előrejelzésére, mint például a higiénia, tisztaság és a beteggel szembeni tisztelet (Greaves és mtsai., 2013). Doing-Harris és munkatársai gépi tanuláson alapuló megközelítést használtak betegeknek az egészségügyre vonatkozó hozzászólásaiban rejlő témák azonosítására. Az elemzés során a kórházi időpont elérhetősége, az empátia és az orvos által nyújtott érthető magyarázat bukkantak fel, mint a leggyakoribb kategóriák (Doing-Harris és mtsai., 2016). Li és munkatársai több mint 122 ezer, kínai orvosokat véleményező szöveget kívántak rendszerezni irodalmi áttekintést, manuális annotációt és természetesnyelv-feldolgozást alkalmazva. Az elemzés eredményeként olyan rendszert hoztak létre, amely három fő (orvossal, beteggel, rendszerrel kapcsolatos témák) és 9 alkategóriát tartalmazott (Li és mtsai., 2018). 2021-ben jelent meg az első olyan rendszeres áttekintés az orvosi szakirodalomban, amely betegek által írott szabad szövegeket elemző, az elmúlt 20 évben megjelent, természetesnyelv-feldolgozást és gépi tanulást alkalmazó, összesen 19 tudományos közleményt foglalt össze (Khanbhai és mtsai., 2021). Ezek között hat olyan közlemény volt, amelyben nem felügyelt gépi tanuláson alapuló modelleket, specifikusan témamodellezést használtak a szövegben és dokumentumokban megjelenő témák és kategóriák automatikus azonosítására. Tudomásunk szerint egy kutatás alkalmazta a legkorszerűbb word2vec és kontextuális BERT modelleken alapuló módszereket a páciensek szövegeinek feldolgozására és az információk kinyerésére, mint például a betegség súlyossága, időbeli hossza és fennállása (Saha és mtsai., 2020).

Hazánkban tudomásunk szerint eddig nem történt olyan vizsgálat, amelynek fókuszában szöveges páciens-élmények elemzése állt volna természetesnyelv-feldolgozással.

1.2. Célkitűzés

A jelen tanulmány célja az volt, hogy a hazai betegellátásban szerzett tapasztalatokat bemutató spontán, nem strukturált írásokban rejlő jellemző információkat természetesnyelv-feldolgozás segítségével kinyerjük. A betegtapasztalatok alapján meghatározható mintázatok és témák automatizált kinyerésével a magyar betegellátás minőségének fejlesztéséhez szeretnénk hozzájárulni. Két kontextuális BERT modell került felhasználásra a szövegek enkódolására, azután nem felügyelt gépi tanulási algoritmussal klaszterelemzést végeztünk. További célunk volt finomhangolt BERT modellek segítségével szentimentelemzést végezni az írások érzelmi polaritásának (pozitív, negatív, semleges) meghatározására.

2. Módszerek

2.1. Adatgyűjtés

A betegek vagy hozzátartozóik által spontán írott szövegek egy magyar online egészség-fórumról kerültek letöltésre (<https://praxis.blog.hu/>). A magánkézben lévő platform 2009 óta működik, napi nézettsége 5000-7000 látogató között van.

A hazai betegellátásban szerzett tapasztalatokat leíró, spontán írott szövegeket a páciensek vagy hozzátartozóik a platform üzemeltetőinek küldik meg emailben, melyek cenzúrázás és szerkesztés nélkül kerülnek feltöltésre, év és hét szerint rendezve. A posztokhoz bárki fűzhet kommentárt. Az írások terjedelme nincs korlátozva. Az adatok nyilvánosan elérhetőek és szabadon hozzáférhetőek, ezért nem szükséges beleegyező nyilatkozat és etikai engedély a felhasználásukhoz. A jelenlegi kutatás keretében a portálon 2009 és 2020 közt feltöltött, páciensek tapasztalatait narratívan ismertető blogbejegyzéseket elemeztük. A blogbejegyzéseket egy keresőrobot segítségével töltöttük le. Minden, személyes azonosításra szolgáló szentitív információt eltávolítottunk. A megtisztított nyers szöveg 1660 blogbejegyzésből, 54,843 mondatból és 873,442 szóból állt, blogbejegyzésenként átlagosan 532,6 szóval.

2.2. Adatelemzés

Az írások numerikus adatokká való transzformálására szóbeágyazási modelleket alkalmaztunk. Ezek a kontextuális hierarchia megtartásával a szavakat vektorrepresentációkká konvertálják, amelyekből távolsági mutatókat számíthatók (pl. koszinusz távolság). Korpusz elemzése érdekében gyakran kerülnek felhasználásra mélytanulási modellek - mint például a word2vec - a jellemzők kinyerésére és kontextus figyelembevételére, azonban a nagyon specifikus kontextusban problémák merülhetnek fel. Ennek elkerülésére a „Bidirectional Encoder Representations from Transformer” (BERT) modellt használtuk, ami kétirányú nyelvmoddellen alapuló, ún. kontextualizált szóbeágyazás, az egyes szavakhoz kontextusfüggő jellemzővektort rendelve (Devlin és mtsai., 2019). A BERT modell előtanítása és finomhangolása nagy méretű korpuszt igényel, azért, hogy különböző kontextusokban is érzékeny legyen.

A jelen tanulmány Angelov (2020) korábbi munkájára épít, amely a legkorszerűbb transzformer modellen alapuló szóbeágyazásokat alkalmazza a témamodellezéshez. A szövegekből kinyert mondatok kódolására kétféle BERT modellt használtunk. Az első, huBERT modellhez – a tokenizálást leszámítva – nincs szükség a bemeneti szövegek előfeldolgozására, mivel a modellt a 9 milliárdos tokenizált Magyar Webkorpusz 2.0 változatán tanították elő (Nemeskey, 2020), teljesítményben felülmúlva más magyar nyelvi tudással rendelkező BERT modelleket több nyelvtechnológiai feladatban (Nemeskey, 2019). A huBERT modell architektúrája ugyanaz, mint a BERT-base modell. 12 kódoló réteggel és 12 figyelmi fejjel rendelkezik, és 768 rejtett feldolgozó elemmel rendelkezik (Ács és mtsai., 2021). HIL-SBERT modellt is alkalmaztunk a

nyelvi jellemzők kinyerésére¹, aminek célja a magas minőségű mondatbeágyazások elérése (Reimers & Gurevych, 2019).

A témamodellezésre használt, nem felügyelt klaszterelemzésre alkalmas algoritmus eredményeinek javítása érdekében elsőként dimenzióredukcióra került sor, amihez az „Uniform Manifold Approximation and Projection for Dimension Reduction” (UMAP) algoritmust használtuk (McInnes és mtsai., 2018). Következő lépésben egy sűrűség alapú térbeli klaszterezési algoritmust (HDBSCAN) alkalmaztunk az alacsony dimenziójú vektorreprezentációkra, hogy az adatstruktúrát felhasználva klasztereket, majd azokról témákat alkossunk. A HDBSCAN egy sűrűség alapú, hierarchikus klaszterelemzésre alkalmas nem parametrikus algoritmus, amely optimális megoldást kínál a futásidőre, valamint a klaszterek számát nem szükséges előre meghatározni (McInnes és mtsai., 2017). A modellt 15 szó per klaszterre állítottuk. A klaszterekben lévő szavakat lemmatizáltuk (spaCy 0.3.1 magyar modell² segítségével) és az NLTK³ magyar stopwords-listában szereplő szavakat töröltük, valamint a duplikált témákat és a null súllyal rendelkező témákat eltávolítottuk a könnyebb értelmezhetőség érdekében. A klaszter eredményeinek értelmezéséhez, illetve a témák azonosításához első körben c-TF-IDF⁴ (class-based TF-IDF) automatikus módszerrel végeztünk témakiválasztást, majd ezt követően humán interpretáció volt szükséges. Egyes hasonló jelentésű témákat kategóriákká vontuk össze.

A blogbejegyzésekből kinyert mondatok emocionális töltetének (pozitív, negatív és semleges) meghatározására szentimentelemzést alkalmaztunk. Az előtanított huBERT modell finomhangolására a Magyar Twitter Szentiment Korpuszt használtuk⁵. A nemzetközi irányelveknek megfelelően két részkorpuszt készítettünk (Wang és mtsai., 2018), amelyek a szövegek emocionális tartalmának prediktálására szolgáltak.

- Bináris (kétosztályos) korpusz: Az ötfokozatú skálából az 1 és 2 értékeket negatív, a 4 és 5 értékeket pozitív kategóriába konvertáltuk, a középső (3) értéket figyelmen kívül hagytuk. Tanító anyag: 2,468 szegmens. Teszt anyag: 269 szegmens.
- Háromosztályos korpusz: Az ötfokozatú skálából az 1 és 2 értékeket negatív, a 3-as értéket semleges, a 4 és 5 értékeket pedig pozitív kategóriába soroltuk. Tanító anyag: 3,600 szegmens. Teszt anyag: 400 szegmens.
- Ötosztályos korpusz: az eredeti ötfokozatú skálán jelölt korpusz, ahol az 1 – nagy negatív, 2: negatív, 3: semleges, 4: pozitív, 5: nagyon pozitív. Tanító anyag: 3,600 szegmens. Teszt anyag: 400 szegmens.

A finomhangolást az alábbi módosított hiperparaméterekkel végeztük: 8-as batch méret / GPU (4 darab GeForce GTX 1080Ti); tanulási ráta: $2e-5$; maximális szekvenciahossz: 128; epoch 15 (legjobb eredményt elérő modelleket alkalmaztuk). A modellek pontossága 85,92% -os volt a bináris osztályozásnál, 72,18% a háromosztályos és 68,50% az ötosztályos korpusznál.

¹ <https://hilanco.github.io/models/sbert.html>

² <https://github.com/spacy-hu/spacy-hungarian-models>

³ <https://www.nltk.org/>

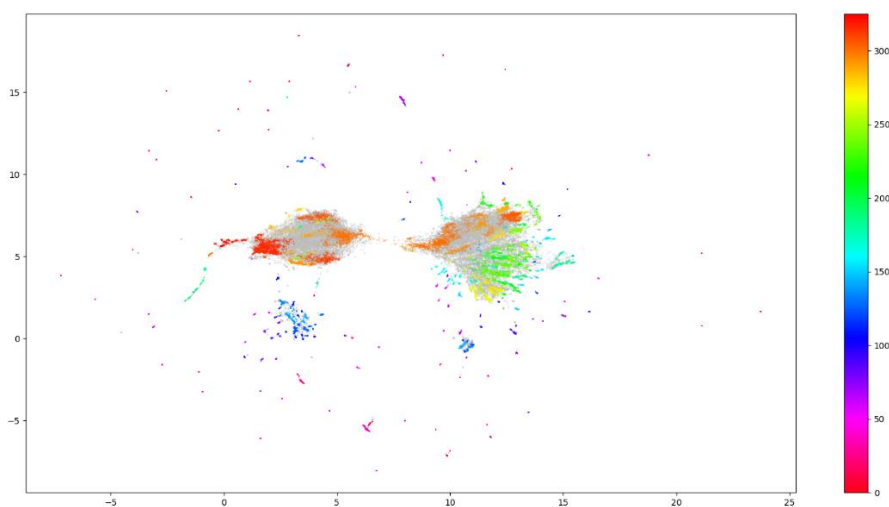
⁴ <https://github.com/MaartenGr/cTFIDF>

⁵ Magyar Twitter Szentiment Korpusz. Letöltve: 2021. november 3. <http://opendata.hu/dataset/hungarian-twitter-sentiment-corpus>.

3. Eredmények

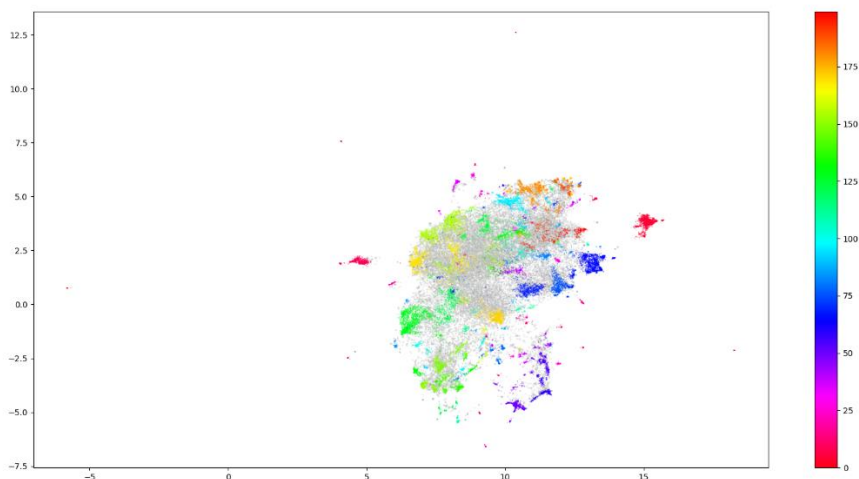
3.1. Témamodellezés

A klaszterelemzés során a finomhangolt BERT modell szemantikailag értelmezhető témákat eredményezett. A huBERT és HIL-SBERT által feldolgozott mondatok alapján 326 és 200 téma került meghatározásra. Érdekes módon a huBERT modell csak 20,503 mondatot, míg a HIL-SBERT modell 20,502 mondatot használt fel az elemzés során, a maradékot zajként (huBERT: 33,744, HIL-SBERT: 34,745) értelmezve. A klaszterek vizualizációjára dimenziócsökkentést alkalmaztunk, a mondatrepresentációkat vektortérben ábrázolva. A színes területek a HDBSCAN algoritmus által azonosított mondatok sűrűségét mutatják, míg a zaj szürke színeként van ábrázolva (1. és 2. ábra).



1. ábra: UMAP-redukált mondatvektorok a huBERT modellre.

Kvantitatív metrikát nem találtunk a két modell kiértékeléséhez, ezért manuálisan vizsgáltuk meg a két modell által kiadott eredményeket. Azt tapasztaltuk, hogy a huBERT modell kimenetét könnyebb volt értelmezni, ezért a további elemzéshez a huBERT modell által azonosított témákat használtuk fel. A modell eredményeként meghatározott 326 téma közül kizártuk azokat, amelyek az elemzés szempontjából nem relevánsak. A megmaradt témák nagy száma miatt humán interpretáció volt szükséges ahhoz, hogy az összetartozó témákat kategóriákká csoportosítsuk. A kategóriák a szerzők közötti iteratív megbeszélések révén kerültek kiválasztásra a torzítás csökkentése érdekében. A mondatok nagymértékben függenek a betegellátásban megélt konkrét helyzettől, valamint a blogbejegyzésekben található további témáktól, mivel egy blogbejegyzés számos témát említhet. A témák nagy száma miatt a továbbiakban csak az öt legjelentősebb kategóriát mutatjuk be (1. táblázat), illetve a modell által fontosnak jelölt, 2. táblázatban látható 18 témát.



2. ábra: UMAP-redukált mondatvektorok a HIL-SBERT modellre.

Diagnózis, tünetek és betegségek

Ez a kategória a páciensek vagy hozzátartozóik által észlelt tüneteket, a betegségeket és a kapott diagnózisokat tartalmazza. A tünetek akut és krónikus betegségekhez és diagnózishoz kapcsolódtak, az influenzától a gyomordaganatig terjedve. A kategóriához számos téma tartozott, mivel a modell több betegségkategóriát ismert fel és különített el. Ebbe a kategóriába tartoztak például a fogászati, gerinc- és szemészeti problémák, nemi betegségek, valamint a plasztikai beavatkozások. A témák között többször megjelenő fájdalom kifejezés a betegséggel való nehéz állapotra és megküzdésre utalhat. Például, az egyik páciens így érzett: „..., leírhatatlan volt a fájdalom, hisz nem érzéstelenítettek.”, illetve egy másik az alábbi módon fogalmazott: „A diagnózis már itt gennyes agyhártyagyulladás gyanúja volt, valószínűleg helyesen...”.

Nőgyógyászat, szülés és születés

A nőgyógyászattal, szüléssel és születéssel kapcsolatos élményeket és eseményeket a BERT modell markánsan differenciálta. A témák visszatérő előfordulása indokoltá tette a témák kategóriává való összevonását. Ez a kategória az alábbi témákat foglalta magában: *nőgyógyászati vizsgálatok, szülés, terhesség, abortusz, kisbaba, újszülött, ol-tás és védettség*. A modell olyan témát is azonosított, mint a szülés utáni lehangoltság és postpartum depresszió. Olyan mondatok tartoztak ide, mint például *“A szülés utáni időszak az anyák életébe paradox módon nem hoz feltétlenül boldogságot.”* és *“A szülés maga simán ment, az orvos korrekt volt.”*

Család és gyermek

A kategória a családok életében nagy valószínűséggel megjelenő témákat foglalta magában, úgymint a *gyerek hospitalizációja, aggodás a gyermekért, családi és szülői szerepek, gyermek védőoltása, és egyéb hozzátartozók állapota*. Megemlítendő, hogy több blogbejegyzés számolt be az iskolában és iskolaorvossal kapcsolatos eseményekről is. A kategória előbbtől való szétválasztását indokolja, hogy míg az előbbi az anyára fókuszál, addig az utóbbi témákban többször jelennek meg a gyermekkel és más hozzátartozókkal kapcsolatos szavak. Egy hozzátartozó így fogalmazott: *„Egy gyermekkel, főleg, ha beteg, türelmesen és odafigyeléssel kell bánni.”*

Struktúra

A páciensek több olyan témát tárgyaltak, amelyek besorolhatók az ellátás szervezeti alapját képező struktúra kategóriába, összhangban a Donabedian által korábban megalkotott dimenzióval. A modell által azonosított témákban megjelennek szakmai problémák, a humán erőforrás hiánya, a betegellátás és a politika kapcsolata, valamint az ellátórendszer nem megfelelő működése. Több utalás történt a kórházakban és osztályokon tapasztalt nem megfelelő higiéniai körülményekre is. Egy páciens következőképp fogalmazott: *„Azért gondoltam ezt megosztani veletek, mert most tényleg mindenhol káosz, zűrzavar, csőd, orvoshiány van.”*

Ellátás kimenete

Ebbe a kategóriába soroltuk a kedvező és a kedvezőtlen kimeneteket egyaránt. A kedvezőtlen kimenethez olyan témák kapcsolódtak, mint a haldoklás, halál, gyász, a gyógyulás bizonytalansága. A modell több kedvezőtlen kimenetű témát azonosított. Például egy hozzátartozó ezt írta: *„A család úgy véli, a kórházban mulasztás történt, és gyógyszer-tuladagolás vezetett szerettük halálához.”* A kedvező kimenet a *gyógyulás és hála* témákban megjelenő kifejezéseket foglalta magában, valamint utalásokat a személyzet által nyújtott minőségi ellátásra. Egy beteg az alábbi módon fejezte ki: *„A többieknek, akik minden nehézség ellenére emberségesek, ezer hála.”*

#	Kategória	Gyakori szavak
1	Diagnózis, tünetek és betegségek	gerinc, derék, fogfájás, herpesz, fájdalom, porckorongsérv, infúzió, panasz, szklerózis, zsibbadás, kettős látás, izomfájdalom
2	Struktúra	nővérhiány, szervezatlenség, orvoshiány, megbánt, ellátórendszer, igazságtalan, szakmai, fejetlenség, nagyfokú,
3	Ellátás kimenete	munka, osztály, hálás, szeret, nővér, dolgozó, kötelék, szívvel, gyász, ember, gyászmunka, halál, halott, folyamat
4	Nőgyógyászat és szülés	anyuka, nőgyógyász, gyermekágy, szülőszoba, szülés, abortusz, ribanc, oltás, védettség
5	Család és gyerek	kisgyermek, ágy, intelligencia, igény, fejlődés, gyermek, apa, szerep, anya, nyelv, szülő, hordozó

1. táblázat. Páciens tapasztalatok jelentőségteljes kategóriái.

#	Téma	Kulcsszavak
1	Orvosi eszköz	Eszköz, hippokratészi, ha, orvosi, méltó, belátás, rész, elismerés, teljesít
2	Gyógyszerek	Frontin, hormon, escitilt, hatóanyag, acetyl, torokfájás, ibuprofen, tableta, paracetamol, szer, szopogatótableta
3	Védőoltás	Oltás, oltópont, elleni, influenza, védettség, védőoltás, patika, beadás
4	Vér	Vér, véradás, véradó, hív, potenciális, vérnyomás
5	Alternatív gyógyászat	Magnezium, hosszú, hatás, kezdeti, alternatív, romlás, gyógyászat, különféle, váratlan, javulás
6	Isten és tudomány	Isten, ezó, tudomány, szélsőséges, orvostudomány, ezotéria, piedesztál,
7	Plasztikai beavatkozások	Plasztikai, borzalmas, műtét, mellő, részletez, hét, rend, cső, cm, pár, sebész, mellnagyobbító
8	Egyéb egészségügyi ellátás	Megnyugtató, diszpécser, mentős, mosoly, türelmet, beteg, kocsi, nyugtat, ápol, szakszerű
9	Ünnepek és vallás	Koszorú, adventi, gyertya, tűzhely, keresztény, hangulat, mitológia, díszít, fenyőfa
10	Egészséges életmód	Étrend, étel, diéta, változatos, hétféle, táplálkozástudományi, magnézium, stressz, munkahelyi, szervezet, magnéziumhiány, b6, stresszes
11	Edzés	Edzés, mozgás, amplifitness, megfelelő, edzésterv, fitnesscenter, izomcsoport, testbresztő, készülék
12	Mellékhatás(ok)	Mellékhatás, truxal, generalizált, ruid, pszichotikus, paranoid, szorongó
13	Dohányzás	Dohányzás, jár, tapasztalat, bárki, orvos, sokan, visszazokik, szenvedélybetegség
14	Állapotközlés	Orvos, állapot, közöl, ct, testvér, hány, nap
15	Beutaló	Időpont, vizsgálat, kér, lap, beutaló, ad, tud, doktornő
16	Tapasztalattapasztalás	Történet, szeret, megoszt, tapasztalat, mesél, blog, kórház, ír, leír
17	Pénz és anyagiak	Tb, fizet, ft, adó, rokkantjára, segély, fizetés, támogat, pénz
18	Várakozási idő és időpont	Három, hét, nap, múlva, óra, időpont, kap, év, műtét, vizsgálat, is, kerül, vár, sor, múlva, fél, év, várakozás

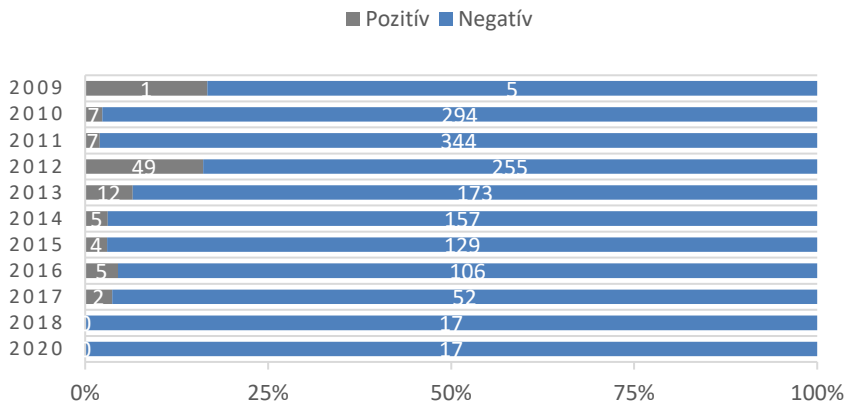
2. táblázat. Azonosított témák és kulcsszavak.

A további témák a 2. táblázatban láthatóak. A páciensek és hozzátartozóik számos egyéb témáról számoltak be a betegellátással kapcsolatban, mint vér, védőoltás, gyógyszerek és mellékhatásaik, alternatív gyógyászatok, vallás és tudomány, dohányzás, életmód és edzés. A megjelenő témák egyedisége miatt további kérdések merülnek fel,

amelyek további vizsgálatokkal válaszolhatók meg, amelyek például az alternatív gyógymódokra, az egészséges életmódra, valamint a vallás és Isten szerepére vonatkozik az orvostudományban és egészségügyi ellátásban.

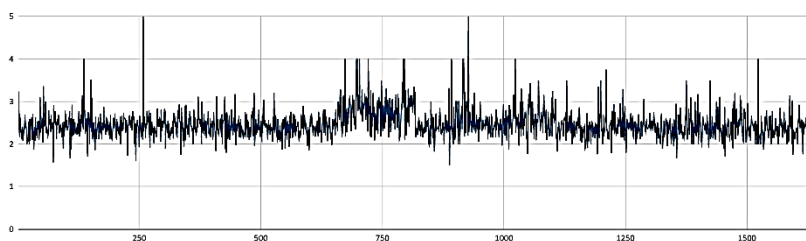
3.2. Szentimentelemzés

A páciens tapasztalatok osztályozása során a finomhangolt huBERT modell a 10 éves periódus során az ötosztályos változat esetében a tapasztalatok 94,4%-át negatívan osztályozta. Hasonló mintázat mutatkozott a másik két változattal is. Az ötosztályos modellben a vágóértéket a 3-as, vagyis a semleges értéknél határoztuk meg a polaritás eldöntéséhez. Egy páciens tapasztalatban több emocionális kifejezés is megjelenhetett, mivel az ellátás különböző aspektusai és folyamata során a páciens az eseményeket különbözőképpen észlelheti, illetve ugyanaz a kórházi ellátás az egyik páciensnek pozitív, míg egy másiknak negatív tapasztalattal végződhet. A páciens tapasztalatokat 10 éves periódusra osztva az ötosztályos modell alapján az íráskor valenciája hasonló tendenciát mutat, többségében negatív szentiment felé tolódva (3. ábra).

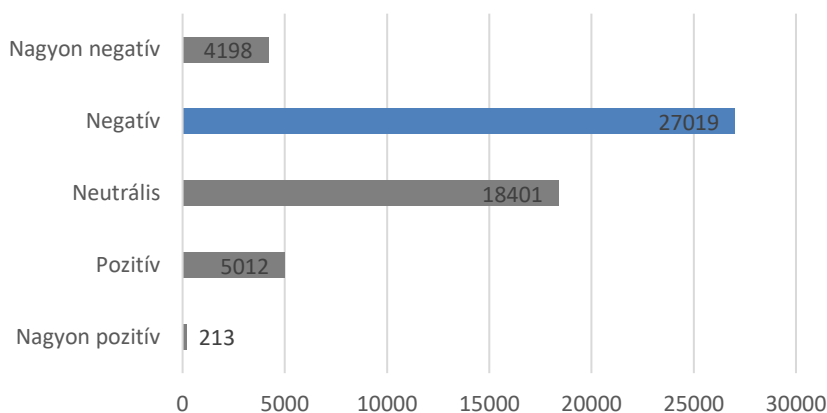


3. ábra. Szentimentelemzés a 10 éves periódusra.

A blogbejegyzések emocionális mintázatának évek során történő változására trendelemzést alkalmaztunk. Az osztályozott mondatok szentiment értékeit egy blogbejegyzésre átlagolva és évek szerint sorba rendezve ábráztuk. Amint a 4. ábrán látható, a blogbejegyzések legnagyobb része negatív szentimentet tükröz. Hasonló mintázat látható a 5. ábrán, amelyen az osztályozott mondatok érzelmi töltetének összesített eloszlását ötfokozatú Likert skálán ábráztuk.



4. ábra: Az évek szerint sorbarendezett blogbejegyzések megoszlása egy átlagos szentiment érték szerint.



5. ábra: Az osztályozott mondatok összesített eloszlása.

Az eredmények azt mutatják, hogy a fórumra beküldött páciens tapasztalatok túlnyomó többsége negatív. Ez valószínűleg részben a negativitási hibával hozható összefüggésbe, vagyis azzal a pszichológiai jelenséggel, hogy a negatív események érzelmi töltése és emlékképei intenzívebbek, mint a pozitív eseményeké, és ezek hosszabb ideig tartó hatást fejtenek ki (Ito és mtsai., 1998).

A negatív szentimentek dominanciája arra is utalhat, hogy ez a nyilvános fórum azokat a betegeket vonzza, akik nem tudnak vagy nem akarnak a betegellátásban létező legális úton, a betegjogi képviselőknél panaszt tenni. Fontos megemlíteni, hogy az elemzésre használt adatok forrása miatt az eredmények nem alkalmasak arra, hogy azokból bármilyen általános következtetést lehessen levonni a hazai betegellátás minőségére. Sőt, azt figyelembe véve, hogy évente több mint egymillió ápolási nap kerül elszámolásra az összes hazai fekvőbeteg-intézményben⁶, a 10 év alatt beküldött 1663, zömmel negatív betegtapasztalat akár kevésnek is tartható.

⁶ Nemzeti Egészségbiztosítási Alapkezelő - Fekvőbeteg Statisztikák. Letöltve: 2021. november 21
http://www.oep.hu/akadalymentes/virtualis_rovat/altfin_virt_dok2/besorolo/fekvo_stat?pagenum=2.

A narratívákban azonosított témák és kategóriák a betegellátásnak a páciensek által fontosnak vélt, gyakran előforduló problémáit mutatják. A panaszokban azonosított témák közötti differenciálás nehéz, mivel egy beteg több problémát említhet, valamint egy probléma leírása több mondaton keresztül is történhet.

Kutatásunk fontossága abban rejlik, hogy ráirányítja a figyelmet a betegek élményeiről spontán írt szövegek elemzésének fontosságára, mivel az lehetőséget ad a jellemzően legnagyobb mértékű frusztrációt és elégedetlenséget okozó konfliktushelyzetek és témák feltárására. A betegek tapasztalatainak elemzése – különösen nagyobb volumenben – fontos adalékkal szolgálhatna az ellátás minőségének javításához, különösen ha az elemzés nevekre és lokalizációkra is kiterjeszhető lenne.

A jelen tanulmányban ismertetett kutatás kezdeti eredményeket mutat be, amelyek a BERT témamodellezés alkalmazhatóságát támasztják alá páciensek narratíváinak automatikus elemzésére. További kutatások szükségesek a módszertani részletek kidolgozására, így például a témák megfelelő számának megállapítására vagy a témák közötti differenciálásra a pontosabb eredmények érdekében.

4. Következtetés

A páciensek tapasztalatainak feltárása és megértése kulcsfontosságú az egészségügyi ellátás javításához, mivel azok egyrészt tükrözik az ellátás minőségét, másrészt befolyással vannak más jelenlegi és leendő betegek attitűdjeire és véleményeire, valamint az ellátással kapcsolatos döntéseikre.

Kutatásunk során egy évtized alatt keletkezett beteg-narratívák automatizált feldolgozását és elemzését végeztük el. Témamodellezést alkalmaztunk BERT alapú modellekkel a nagy gyakorisággal előforduló témák és kategóriák azonosítására. Az írásokból öt jelentőségteljes kategóriát azonosítottunk: 1) diagnózis, tünetek és betegségek; 2) struktúra; 3) ellátás kimenete; 4) család és gyerek; és a 5) nőgyógyászat és szülés. A kategóriák közül kettő, a struktúra és az ellátás kimenete megegyezik a Donabedian-féle modell két dimenziójával, ezért is alkalmaztunk azonos elnevezést. Emberi interpretációt használtunk a modell teljesítményének értékelésére, valamint a témák kategóriákká való összevonására. Szentimentelemzést alkalmaztunk a mondatok, egyben a blogbejegyzések polaritásának meghatározására. Az írások döntő részét a modell negatívan osztályozta, utalva a negatív páciens tapasztalatokra. A betegellátás problematikus témáinak feltárására és az orvos-beteg találkozások mélységi megértésére hatékonyan használhatók a természetesnyelv-feldolgozás eszközei. Az alkalmazott módszereket azonban feltétlenül finomítani kell annak érdekében, hogy az eredmények értelmezése során minél kevesebb manuális közreműködésre legyen szükség.

Hivatkozások

- Ács, J., Lévai, D., Nemeskey, D. M., & Kornai, A. (2021). *Evaluating Contextualized Language Models for Hungarian*. <http://arxiv.org/abs/2102.10848>
- Angelov, D. (2020). Top2Vec: Distributed Representations of Topics. *arXiv*. <http://arxiv.org/abs/2008.09470>

- Bidmon, S., Elshiewy, O., Terlutter, R., & Boztug, Y. (2020). What Patients Value in Physicians: Analyzing Drivers of Patient Satisfaction Using Physician-Rating Website Data. *Journal of medical Internet research*, 22(2), e13830. <https://doi.org/10.2196/13830>
- Brito Fernandes, Ó., Baji, P., Kringos, D., Klazinga, N., Gulácsi, L., Lucevic, A., Boncz, I., & Péntek, M. (2019). Patient experiences with outpatient care in Hungary: results of an online population survey. *European Journal of Health Economics*, 20(Suppl 1), 79–90. <https://doi.org/10.1007/s10198-019-01064-z>
- Burkle, C. M., & Keegan, M. T. (2015). Popularity of internet physician rating sites and their apparent influence on patients' choices of physicians. *BMC Health Services Research*, 15(1), 1–7. <https://doi.org/10.1186/s12913-015-1099-2>
- Carinci, F., Van Gool, K., Mainz, J., Veillard, J., Pichora, E. C., Januel, J. M., Arispe, I., Kim, S. M., & Klazinga, N. S. (2015). Towards actionable international comparisons of health system performance: Expert revision of the OECD framework and quality indicators. *International Journal for Quality in Health Care*, 27(2), 137–146. <https://doi.org/10.1093/intqhc/mzv004>
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference, 1(Mlm)*, 4171–4186.
- Doing-Harris, K., Mowery, D. L., Daniels, C., Chapman, W. W., & Conway, M. (2016). Understanding patient satisfaction with received healthcare services: A natural language processing approach. *AMIA ... Annual Symposium proceedings. AMIA Symposium, 2016*, 524–533. <https://europepmc.org/articles/PMC5333198>
- Donabedian, A. (2005). Evaluating the quality of medical care. In *Milbank Quarterly* (Köt. 83, Szám 4, o. 691–729). Blackwell Publishing Inc. <https://doi.org/10.1111/j.1468-0009.2005.00397.x>
- Doyle, C., Lennox, L., & Bell, D. (2013). A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. In *BMJ Open* (Köt. 3, Szám 1, o. 1570). British Medical Journal Publishing Group. <https://doi.org/10.1136/bmjopen-2012-001570>
- Emmert, M., Meier, F., Heider, A. K., Dürr, C., & Sander, U. (2014). What do patients say about their physicians? An analysis of 3000 narrative comments posted on a German physician rating website. *Health Policy*, 118(1), 66–73. <https://doi.org/10.1016/j.healthpol.2014.04.015>
- Fernandes, Ó. B., Péntek, M., Kringos, D., Klazinga, N., Gulácsi, L., & Baji, P. (2020). Eliciting preferences for outpatient care experiences in Hungary: A discrete choice experiment with a national representative sample. *PLoS ONE*, 15(7 July). <https://doi.org/10.1371/journal.pone.0235165>
- Giordano, L. A., Elliott, M. N., Goldstein, E., Lehrman, W. G., & Spencer, P. A. (2010). Development, implementation, and public reporting of the HCAHPS survey. In *Medical Care Research and Review* (Köt. 67, Szám 1, o. 27–37). Med Care Res Rev. <https://doi.org/10.1177/1077558709341065>
- Greaves, F., Ramirez-Cano, D., Millett, C., Darzi, A., & Donaldson, L. (2013). Use of sentiment analysis for capturing patient experience from free-text comments posted online. *Journal of Medical Internet Research*, 15(11), 1–9. <https://doi.org/10.2196/jmir.2721>
- Hanauer, D. A., Zheng, K., Singer, D. C., Gebremariam, A., & Davis, M. M. (2014). Public awareness, perception, and use of online physician rating sites. *JAMA - Journal of the American Medical Association*, 311(7), 734–735. <https://doi.org/10.1001/jama.2013.283194>
- Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. *Journal of personality and social psychology*, 75(4), 887–900. <https://doi.org/10.1037//0022-3514.75.4.887>
- Jenkinson, C., Coulter, A., Bruster, S., Richards, N., & Chandola, T. (2002). Patients' experiences and satisfaction with health care: Results of a questionnaire study of specific aspects of care. *Quality and Safety in Health Care*, 11(4), 335–339.

- <https://doi.org/10.1136/qhc.11.4.335>
- Khanbhai, M., Anyadi, P., Symons, J., Flott, K., Darzi, A., & Mayer, E. (2021). Applying natural language processing and machine learning techniques to patient experience feedback: A systematic review. In *BMJ Health and Care Informatics* (Köt. 28, Szám 1, o. 100262). BMJ Publishing Group. <https://doi.org/10.1136/bmjhci-2020-100262>
- Lagu, T., Goff, S. L., Hannon, N. S., Shatz, A., & Lindenauer, P. K. (2013). A mixed-methods analysis of patient reviews of hospital care in England: Implications for public reporting of health care quality data in the United States. In *Joint Commission Journal on Quality and Patient Safety* (Köt. 39, Szám 1, o. 7–15). Joint Commission Resources, Inc. [https://doi.org/10.1016/s1553-7250\(13\)39003-5](https://doi.org/10.1016/s1553-7250(13)39003-5)
- Li, J., Liu, M., Li, X., Liu, X., & Liu, J. (2018). Developing embedded taxonomy and mining patients' interests from web-based physician reviews: Mixed-methods approach. *Journal of Medical Internet Research*, 20(8). <https://doi.org/10.2196/jmir.8868>
- López, A., Detz, A., Ratanawongsa, N., & Sarkar, U. (2012). What patients say about their doctors online: A qualitative content analysis. *Journal of General Internal Medicine*, 27(6), 685–692. <https://doi.org/10.1007/s11606-011-1958-4>
- Lucevic, A., Péntek, M., Kringos, D., Klazinga, N., Gulácsi, L., Brito Fernandes, Ó., Boncz, I., & Baji, P. (2019). Unmet medical needs in ambulatory care in Hungary: forgone visits and medications from a representative population survey. *European Journal of Health Economics*, 20(1), 71–78. <https://doi.org/10.1007/s10198-019-01063-0>
- Mattke, S., Epstein, A. M., & Leatherman, S. (2006). The OECD Health Care Quality Indicators Project: History and background. *International Journal for Quality in Health Care*, 18(SUPPL. 1), 1–4. <https://doi.org/10.1093/intqhc/mzl019>
- McInnes, L., Healy, J., & Astels, S. (2017). hdbscan: Hierarchical density based clustering. *The Journal of Open Source Software*, 2(11), 205. <https://doi.org/10.21105/joss.00205>
- McInnes, L., Healy, J., & Melville, J. (2018). *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction*. <http://arxiv.org/abs/1802.03426>
- Nawab, K., Ramsey, G., & Schreiber, R. (2020). Natural Language Processing to Extract Meaningful Information from Patient Experience Feedback. *Applied Clinical Informatics*, 11(2), 242–252. <https://doi.org/10.1055/s-0040-1708049>
- Nemeskey, D. M. (2020). Natural Language Processing Methods for Language Modeling. In *Eötvös Loránd University*.
- Nemeskey, D. M. (2019). Introducing huBERT. *XVII. Magyar Számítógépes Nyelvészeti Konferencia*, 3–14.
- Price, R. A., Elliott, M. N., Zaslavsky, A. M., Hays, R. D., Lehrman, W. G., Rybowski, L., Edgman-Levitan, S., & Cleary, P. D. (2014). Examining the role of patient experience surveys in measuring health care quality. In *Medical Care Research and Review* (Köt. 71, Szám 5, o. 522–554). SAGE Publications Inc. <https://doi.org/10.1177/1077558714541480>
- Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference*, 3982–3992. <http://arxiv.org/abs/1908.10084>
- Saha, B., Lisboa, S., & Ghosh, S. (2020). Understanding patient complaint characteristics using contextual clinical BERT embeddings. *42nd IEEE Annual Conference of Engineering in Medicine and Biology Society*.
- Wang, A., Singh, A., Michael, J., Hill, F., Levy, O., & Bowman, S. R. (2018). GLUE: A Multi-Task Benchmark and Analysis Platform for Natural Language Understanding. *7th International Conference on Learning Representations, ICLR 2019*. <http://arxiv.org/abs/1804.07461>
- Zhang, W., Deng, Z., Hong, Z., Evans, R., Ma, J., & Zhang, H. (2018). Unhappy Patients Are Not Alike: Content Analysis of the Negative Comments from China's Good Doctor Website. *Journal of Medical Internet Research*, 20(1), e35. <https://doi.org/10.2196/jmir.8223>