



A depresszió keretezési módjainak vizsgálata online fórumokon természetesnyelv-feldolgozással

MÁTÉ FANNI¹

ABSZTRAKT

A depresszió jelenségének vizsgálata nem újkeletű a szociológiában, de egyre szélesebb társadalmi problémává válása folytán ma is aktuális. A depresszióhoz kapcsolódó biomedikális és pszichológiai aspektus mellett egyre figyelemre méltóbb a szociológiai nézőpont beemelése a depresszió lehetséges okairól zajló párbeszédbe. A depresszióról szóló diskurzus kutatásában az online elérhető szövegek számos új lehetőséget nyújtanak, hiszen a fórumok anonimitása, egyszerű elérése népszerűvé teszi az online segítségkérést az érintettek számára. Kutatásomban természetesnyelv-feldolgozást, azon belül logisztikus regressziót használtam, hogy feltérképezem a depresszió értelmezési kereteinek laikus diskurzusokban való megjelenési mintáztatának gépi tanuló algoritmussal való feltérképezhetőségét. A gépi tanulás által ugyanis lehetővé vált, hogy 67857 hozzászólást elemezzek, amelynek emberi erőforrással való feldolgozása nehezen valósulhatott volna meg. Az elemzés során angol nyelvű, depresszió témájú online fórumok bejegyzéseinek kategorizálást végeztem el a tudományos diskurzus-kategóriák mentén. Első eredményeim azt mutatják, hogy a logisztikus regresszió mérsékelten jobb eredményt nyújtott, mint a véletlenszerű besorolás, egyben az emberi besoroláshoz hasonló eredményt ért el. Annak ellenére, hogy a kutatás angol nyelvű fórumokat elemzett, eredményeim hasznosak lehetnek bárki számára, aki absztrakt szociológiai fogalmak felhasználók által írt online szövegekben való megjelenésével foglalkozik.

KULCSZAVAK: depresszió, természetesnyelv-feldolgozás, mentális egészség, online fórumok

ABSTRACT

Classification of depression-related online forums using Natural Language Processing

The study of the phenomenon of depression is not new in sociology, but since the depression is becoming a wider social problem, it is still a relevant issue today. In addition to the biomedical and psychological aspects of depression, the sociological perspective is becoming more noteworthy in the discourse about the causes of depression. In the research of the discourse on depression, the online texts offer many new possibilities, as the forum's anonymity and accessibility make the online seeking for help popular. In this research, natural language

¹ Eötvös Loránd Tudományegyetem, Társadalomtudományi Kar, email: mate.fanni@tatk.elte.hu



TEMATIKUS TANULMÁNYOK – Digitális szociológia

processing (logistic regression) was applied to find patterns in the definition of depression in lay discourses. These methods make it possible to analyze a large amount of text - which would have been difficult to process with human resources. During the analysis, 67 857 posts of English-speaking online forums were categorized along the categories of the scientific discourse about depression. This study presents the first results, which shows logistic regression classifier performs like the annotators. . Although the research has analyzed English-speaking forums, my findings may be useful to anyone observing abstract sociological concepts in online texts written by users.

KEYWORDS: *depression, Natural Language Processing mental health, online forums*

Bevezetés

A depresszió online szövegekre alapuló kutatása új távlatokat nyit meg a szociológiában. Egyfelől amiatt, mert a platformok és azokon keresztül a hasonló tapasztalatokkal rendelkező személyek könnyen, anonim módon elérhetőek, így népszerűek a depresszióval, szorongással élők egyre növekvő csoportjában. Másfelől az oldalakon keletkező szövegek nem kutatási célból álltak elő, amely kétségtelen előnye, ahogy Dessewffy és Láng is kiemeli (2015), hogy a „hagyományos” kutatások kérdezési szituációjában olyan véleményeket ismerünk meg, amelyeket alakított a kérdezettek elvárásoknak való megfelelésének igénye (lásd: megfigyelői paradoxon (Labov 1972: 61)), sőt a vélemények kérdezés nélkül talán meg sem születtek volna. Ezzel szemben az interneten elérhető viselkedések ténylegesek, nem torzítja őket a kutatói igényeknek való megfelelés. A szerzők ugyan elsősorban a strukturált, számszerűsíthető digitális lábnyom kapcsán említik ezt az előnyt, de a megnyilatkozásokra szintén igaz, hogy a véleményeket, hozzászólásokat nem a kutatási szituáció hozta létre, azt külső tényezőként legfeljebb az adott tér normái, szabályai alakítják, de azok is organikus részét képezik a diskurzus terének.

Jelen tanulmányban a depresszió értelmezési módjait, annak okairól és lehetséges kezelési módjairól folyó globális, laikus diskurzust vizsgálom két aspektusból. A kutatás elsődleges célja, hogy hozzájáruljon a természetesnyelv-feldolgozás társadalomtudományi alkalmazhatóságának felméréséhez laikus diskurzusok elemzése nyomán. A vizsgálat abból a feltételezésből indul ki, hogy a szakértői diskurzusban megtalálható keretelési módok a laikus diskurzusokban is megjelennek. A kutatás implicit módon ezt a feltételezést is vizsgálja, amennyiben az osztályozás sikeressége azon is múlik, párhuzamba állítható-e a laikus és a szakértői diskurzus. Annak megértése, hogy az érintettek miként vélekednek a depresszióról, betegségként, lelki, avagy társadalmi problémaként tekintenek inkább rá, a hatékony segítségnyújtás formáit is meghatározhatja. Az online fórumokon zajló párbeszéd tartalmának ismerete amiatt is értékes a szakemberek számára, mert az online fórumok könnyen elérhető információforrásként szolgálnak az érintetteknek.



TEMATIKUS TANULMÁNYOK – Digitális szociológia

Tanulmányomban először bemutatom a depresszió fogalmát és különböző (biomedikális, pszichológiai és szociológiai) keretelési módjait, majd depresszióval kapcsolatos szövegeket vizsgáló kutatásokat ismertetek. Kitérek hagyományos kvalitatív tartalomelemzést, és újabb, gépi tanulást alkalmazó kutatásokra is. Az elméleti bevezető után a használt adatbázist, annak előfeldolgozását és az alkalmazott algoritmusokat mutatom be. A tanulmány utolsó részében az eredményeimről és azok lehetséges magyarázatáról, értelmezéséről írok – visszacsatolva ezeket fentiekben leírt kutatási kérdésekhez.

A depresszióról

Az Amerikai Pszichiátriai Társaság (APA) 2013-ban kiadott szakkönyve² szerint akkor állítható fel a depresszió diagnóza, ha az alábbi³ kilenc jellemző közül az első vagy második és további legalább négy jellemző legalább két héten keresztül megfigyelhető:

- levertség
- érdeklődés, örömeire való képesség elvesztése
- étvágyban vagy testsúlyban bekövetkező, nagymértékű növekedés vagy csökkenés
- megnövekedett vagy csökkent alvásigény, álmatlanság
- pszichomotoros agitáció (nyugtalanosság) vagy gátoltság (meglassultság)
- fáradtság, energiahány
- csökkent önértékelés, értéktelenség érzése, inadekvát büntudat érzése
- gondolkodásra való képtelenség, döntésképtelenség, koncentrációhiány
- halálra vonatkozó gondolatok gyakorisága, öngyilkossági gondolatok.

Az Egészségügyi Világszervezet (WHO) adatai alapján 2017-ben világszerte több mint 300 millió ember szenvedett depressziótól, s világszinten 2015-ben az összes haláleset 1,5 százaléka öngyilkosság volt, amely a depresszió egyik súlyos következménye lehet (World Health Organization 2017). Az előbbi adat természetesen csak iránymutatás, hiszen az egyes régiók, országok között széles skálán mozog a depresszió diagnosztizálásnak gyakorlata. Az adatokat továbbá azért is óvatosan kell kezelnünk, mert a depresszióhoz és a mentális betegségekhez erős társadalmi stigma tartozik (Rogers – Pilgrim 2005), amely befolyásolhatja az orvoshoz, pszichológushoz fordulás valószínűségét, így a diagnózis felállítását is. Ezen tényezők azonban csak tovább növelnék a megfigyelt előfordulások számát, így a fenti eset-számokra leginkább alsó becslésként tekinthetünk.

² Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition

³ A magyar fordításhoz Torzsa és munkatársai tanulmányában (2009) szereplő kifejezéseket használtam.



TEMATIKUS TANULMÁNYOK – Digitális szociológia

A depresszió keretezési módjai

A „depresszió keretezése” alatt a depresszió értelmezésének egyének (a szakértők és a laikusok) vagy intézmények általi kijelölését értem, azt a fogalmi keretet, amely mentén a depresszióra tekintenek.⁴ A tudományos diskurzusban (Busfield 2000, Comer 2013, Rogers – Pilgrim 2005) három keretezés különíthető el egymástól: (1) a biomedikális, amely a depressziót elsősorban szomatikus eredetűnek írja le, s azt gyógyszeres úton kezeli; (2) a pszichológiai, amikor lelki eredetű problémaként jelenik meg a depresszió és végül (3) a szociológiai, amely a két értelmezés mellett a társadalom szerepét hangsúlyozza a mentális zavarok kialakulásában. Annak, hogy milyen módon tekintünk a depresszióra, melyik a domináns keretezési mód, társadalmi szinten is nagy tétje van. A keretezés ugyanis kijelöli, milyen okoknak tulajdonítjuk a depressziót, így azt is, milyen kezelési, megküzdési módokat tartunk célravezetőknek. Mindez pedig hatással van arra, hogy a depresszióban szenvedő egyének felmerik-e vállalni problémájukat. A depresszió értelmezésének következményei társadalmi és egyéni szinten egyaránt megjelennek, például olyan kérdésekben, minthogy támogassa-e a társadalombiztosítás az antidepresszánsokat vagy a pszichológiai kezelést; hogy az érintett személy hatékony terápiát választ-e; de akár abban is, hogy a depressziós egyénre a társadalom miként tekint. Az alábbiakban a fenti három fő keretezési módot mutatom be röviden.

BIOMEDIKÁLIS KERETEZÉS

A szakemberek előtt még nem ismert, hogy pontosan milyen biológiai tényezők okozzák depressziót, csupán az, milyen különbségek figyelhetők meg a depresszióval érintettek testi folyamataiban az azzal nem érintettekhez képest. Egyes szerzők szerint (Busfield 2000, Comer 2013) ez a keretezés a jelenlegi uralkodó diskurzus a tudományos életben, főként amiatt, hogy a természettudományok jellegükből fakadóan objektíven mérhető eredményekkel tudják alátámasztani a hipotéziseiket.

A területen megjelenő különböző nézetek a genetika, illetve a biokémiai folyamatok szerepét hangsúlyozzák. A genetika szerepét támasztják alá azok a kutatások, amelyek szülők és gyerekek depresszióját vizsgálták olyan esetekben is, amikor a gyermek nem a vér szerinti szüleivel élt (Kamali – McInnis 2011). E szerint az iskola szerint tehát a depresszió, az arra való hajlam öröklődhet. A biokémiai nézőponthoz tartozik egyfelől az a megfigyelés, hogy két neurotranszmitter, a noradrenalin és a szerotonin aktivitása erősen kötődik a depresszióhoz, másfelől a depressziós

⁴ A használt keretezés fogalom részben eltér Lakoff (2004., o. xv) keretezés fogalmától, mely szerint a keretek azok a „mentális struktúrák, amelyek formálják azt, miként látjuk a világot”. Az általam használt keretezési fogalom ennél szűkebb, arra vonatkozik, a vizsgált jelenséget mely diszciplína szerint értelmezzük.



TEMATIKUS TANULMÁNYOK – Digitális szociológia

személyekben megfigyelt magas kortizol szint (Gao – Bao 2011), amely a stressz kiváltásáért felelős hormon. A képpalkotó rendszerek segítségével napjainkban az is ismert, hogy az agy különböző részeiben (például az amygdalában vagy a prefrontális kéregben (Ressler – Mayberg 2007)) eltérő az aktivitás szintje a depresszióban szenvedő és az abban nem szenvedő egyéneknél. Azonban sem a biokémiai faktorok esetén, sem az agyi területek aktivitása esetén nem sikerült (még) kimutatni az okság irányát, azaz nem tudhatjuk, az agyi, biokémiai elváltozások okozzák-e a depressziót, vagy hogy éppen a depresszió hatására történnek változások ezekben.

A biomedikális keretezéshez kapcsolódó kezelési módok a megváltozott biokémiai folyamatokra hatnak, így jellemzően antidepresszánsok szedésével kezelik a depressziót (Goldstein et al. 2011).

PSZICHOLÓGIAI KERETEZÉS

A pszichológiai gyakorlatban a depresszió lelki eredetű problémaként definiált, azonban ezen belül is meg kell különböztetnünk többféle irányzatot: a pszichoanalitikus, a kognitív és a behaviorista nézetet. A pszichoanalízis irányzata szerint a depresszió a gyászhoz hasonló folyamat, a gyászhoz képest hosszabb távon fennálló, tényleges vagy szimbolikus veszteség váltja ki (Freud 1985, Sik 2018).

A behaviorista nézet szerint a depressziót a jutalmazás, megerősítés csökkenése, a negatív megerősítés növekedése idézi elő (Lewinsohn 1985). Súlyosbítja a helyzetet, hogy a depressziós személy társadalmi jutalmainak csökkenése a konstruktív viselkedést is csökkenti, amely a további jutalmak elmaradását okozhatja, így eredményezve depressziós spirált (Martell et al. 2013).

A kognitív nézet két jellemző, depressziót célzó magyarázata a negatív gondolkodás és a tanult tehetetlenség elmélete. A negatív gondolkodás nézete szerint a depresszió fő okozója az egyént körülvevő világ negatív interpretációja (Beck 1979). A tanult tehetetlenség (Seligman 1972) pedig az a jelenség, amikor az egyén korábbi negatív tapasztalatai alapján úgy gondolja, nem ő irányítja az életét, a jutalmakra és büntetésekre nincs hatása, így egy idő után nem is próbál javítani az életkörülményein (Comer 2013).

A depresszió kezelése a pszichológiában terápiával történik, de az egyes iskolák ebben is különböző terápiás módszereket alkalmaznak. A korábban domináló pszichoanalízisben a depressziót kiváltó veszteséget hívták elő és tudatosították a személyekben, de a kutatások nem minden esetben támasztották alá ennek a módszernek a hatékonyságát (Comer 2013). A behaviorista terápia arra helyezi a hangsúlyt, hogy az egyén olyan tevékenységeket végezzen, amelyek örömet okoznak számára, és amelyek pozitív megerősítést eredményeznek. A kognitív terápiában pedig tevékenyebbé teszik az egyént, napi célokat, feladatokat tűznek ki elé, majd az au-



TEMATIKUS TANULMÁNYOK – Digitális szociológia

tomatikus negatív gondolatok automatikusságára és a gondolatok negatív mivoltára ébresztik rá, végül pedig az elsődleges, maladaptív attitűdjeit változtatják meg.

SZOCIOLÓGIAI VAGY SZOCIOKULTURÁLIS KERETEZÉS

A szociológiai keretezés az eddig tárgyalt két keretezési módot egészíti ki egyfelől azzal, hogy az egyént nem különálló tényezőként, hanem a társadalom szerves részeként tekinti, amely eredője lehet a depresszióhoz köthető pszichológiai problémáknak is; másfelől pedig abban az értelemben, hogy a biomedikális diskurzushoz képest, ahol az oksági kapcsolat iránya nem ismert, (a pszichológiai keretezéshez hasonlóan) a depresszió kiváltó okait keresi. Akárcsak a másik két értelmezési mód esetén, a szociológia különböző irányzatai is eltérő tényezőket emelnek ki a depresszió vagy általában a mentális zavarok vizsgálatakor.

A multikulturális megközelítés⁵ a depresszió társadalmon belüli eloszlását, szabályszerűségeit vizsgálja, magyarázatként akár olyan egyéni jellemzőket keresve, amelyek a társadalom struktúrájából fakadnak és így kapcsolódnak a depresszióhoz (Comer 2013).

Sik (2018) arra hívja fel a figyelmet a depresszió szociológiai aspektusát vizsgálva, hogy az örömtelenség, amely a depresszió egyik tünete, a társadalomból, annak működéséből fakadhat – például a különböző erőforrások hiányából vagy a társadalmi normák szigorú szabályaiból. Emellett a cselekvésre való képesség leépülése is összefügg a társadalom működésével, például a társadalmi elismerés hiánya vagy a munkanélküliség problémái kapcsán. Átfogó meghatározása (Sik 2018: 8) szerint ez alapján a depresszió tekinthető az eltorzult társadalom, a társadalmi zavarok interiorizálásaként is.

A depresszió szociológiai értelmezésének hermeneutikai vonatkozása is van, amelyben a szociológia szintén a két másik keretezési módot egészíti ki: eszerint a mentális zavarok, így ezen belül a depresszió meghatározása, vagyis a normális és patológikus elhatárolása társadalmi konstrukciók függvénye (Busfield 2000, Rogers – Pilgrim 2005), amelyek térben és időben is változnak.

A depresszió szociológiai értelmezése különösen fontos lehet a megfelelő kezeléshez: ha ugyanis csak medikális vagy terápiás úton történik a kezelés, azonban az egyén társadalmi helyzete nem változik meg (azaz a depressziót kiváltó probléma nem oldódik meg), a kezelés lényegében csak tüneti lesz. Természetesen a társadalom rendszerét, szabályszerűségeit nem lehet egyéni szinten megváltoztatni: ahhoz összehangolt, társadalmi szintű változások, (szociál)politika szükséges, egyéni szinten azonban hatékonyak lehetnek azok a terápiák, amelyek érzékenyek az egyén

⁵ Rogers és Pilgrim (2005) ezt a nézőpontot oksági magyarázatként nevezi, hiszen a diagnózist elfogadva ez a nézőpont az okokat keresi.



TEMATIKUS TANULMÁNYOK – Digitális szociológia

társadalmi helyzetére és környezetére, a terápia során ezen tényezőkre is figyelnek (Comas-Díaz 2011).

Korábbi kutatások a témában

„HAGYOMÁNYOS” MÓDSZEREKET HASZNÁLÓ KUTATÁSOK

Az alábbiakban olyan, hagyományos eszköztárral készült kutatásokat mutatok be, amelyek a depresszió keretezési módjait vizsgálják laikus diskurzusokban. Kokanovic munkatársaival (2012) az érintettek diskurzusát interjúk alapján kutatta. Az interjúk fókuszában a diagnózis felállítása, annak esetleges tagadása és az életvitelre történő hatása állt. Az Ausztráliában 2008 és 2011 között felvett interjúk eredményei azt mutatják, hogy az érintettek (orvosi diagnózissal vagy öndiagnózissal rendelkezők) hol elfogadják, hol kiegészítik a medikális diskurzust, ingáznak a medikális és nem-medikális elemek között. Sokan gondolják például úgy, hogy az antidepresszánsok segíthetnek a tünetek enyhítésében, mindazonáltal nem elegendők a probléma kezeléséhez, hiszen a depresszió komplexebb annál, mintsem, hogy gyógyszerrel orvosolható legyen (amely rímel a szociológiai keretezésben megjelenő kezelés-problematikára). Érdemes kiemelni, hogy az orvosi diagnosztikai gyakorlatot nem bírálták az interjúalanyok, miközben ennek bírálata a tudományos életben megjelenik (Busfield 2000, Kokanovic et al. 2012). Sőt elmondásuk szerint egyéb segítség híján az orvossal beszélnek meg a depresszióhoz köthető mindennapi, nem egészségügyi problémáikat is az interjúalanyok – akkor is, ha tudják, hogy azok nem az orvos kompetenciájába tartoznak. A tanulmány a szociológiai aspektust is említi, amikor arról ír, hogy világunkban a vidámság és elégedettség a cél, így patológikus viselkedésnek hat, ha valaki elégedetlen vagy lehangolt. Az interjúalanyok közül többen is úgy látták, a lehangoltságuk nem igényel orvosi beavatkozást, nem biológiai eredetű, hanem adott élethelyzetükből adódik (Kokanovic et al. 2012). E tanulmány nagyon jól mutatja, hogy az egyéni gondolkodásban nem feltétlenül van csak egyféle, egyértelmű értelmezési keret, sokkal inkább azok keveredése jellemző.

Az előbbieken bemutatott tanulmányban tehát kutatási célból készített interjúkat elemeztek. A következőkben bemutatott kutatások adatait már elemzésem szempontjából relevánsabb, online elérhető közösségi média hozzászólások képezik, feldolgozási módszerük azonban a jelen tanulmányéhoz képest eltérő, hagyományos, kutatói kvalitatív kódolás. A depresszió nyilvános értelmezését vizsgálja Pan munkatársaival (2018) egy kínai közösségi médium, a Twitterhez hasonló mikroblog, a Sina Weibo egyes hozzászólásainak tartalomelemzésével. Kvalitatív tartalomelemzéssel vizsgáltak 902 hozzászólást egy többszempontú kódrendszer alapján. Az egyik ilyen vizsgált szempont az volt, hogy a depresszió milyen tényezők miatt alakul ki, és milyen lehetőség van a megoldására. A szerzők négy kategóriát



TEMATIKUS TANULMÁNYOK – Digitális szociológia

különítenek el: a biológiai és medikális, a személyes (például önszabályozás), a szülőkkel vagy más közeli hozzátartozókkal való kapcsolathoz kötődő és végül a strukturális szintű okok és megoldások kategóriáját (Pan et al. 2018: 776). E négy kategória lényegében megfeleltethető a kutatásom alapját képező kategóriáknak – bár a kapcsolatokat mind a pszichológiai, mind a szociológiai kategóriába besorolhatjuk. E kategóriákkal kapcsolatban eredményeik azt mutatják, hogy ha strukturális vagy biológiai okok merülnek fel, akkor kisebb a hozzászólásokban megjelenő stigmatizáció, mintha a felhasználók egyéni tényezőket társítanak a depresszióhoz. Kutatásuk során a szerzők nem csak személyek, hanem intézményi fiókok hozzászólásait is elemezték, s összehasonlították, hogy a különböző típusú szereplők jellemzően milyen keretezési módokat használnak. Eredményeik szerint az állami média, a médiaszereplők és (az elsősorban abból tájékozódó) civil felhasználók hajlamosabbak az egyéni felelősséget hangsúlyozni; a tudományos élethez vagy az egészségügyhöz tartozó szereplők a biológiai tényezők fontosságát emelik ki; a strukturális tényezők pedig jellemzően a „civil újságírók” (citizen journalist) vagy különböző társaságok hozzászólásaiban jelennek meg. A korábban ismertetett tanulmányhoz hasonlóan e tanulmányban is megjelenik a biológiai és a nem-biológiai keretezés ellentéte.

A tanulmányban leírt vizsgálatnak azonban vannak korlátai. Csupán 902 hozzászólást elemeztek, holott hároméves intervallumban 2014 és 2016 között vizsgáldták, amikor (az adatgyűjtés leírásából következtethetően) nagyjából 36000 hozzászólás született az oldalon. A kis elemszám lehetővé teszi a kvalitatív kódolást, azonban a rendelkezésre álló adatoknak így csak kis részét elemzik. Ezen kutatás tehát viszonylag kevés adatra épül, amely mintázatok keresésére kevésbé alkalmas és jobban kitett az adatok esetlegességének. Továbbá az alacsony elemszám hátránya, hogy nem feltétlenül találkozunk minden, kutatási kérdésnek megfelelő kategóriával, ha annak számossága alacsony a vizsgált térben.

Ezzel szemben nagyobb mennyiségű szöveg gépi tanulási módszerekkel történő feldolgozása kevésbé hordozza magán a kutatói besorolás hátrányait, hiszen nagyobb mennyiségű adat feldolgozását teszi lehetővé kevésbé szubjektív, reprodukálható módon. Az alábbiakban ilyen, gépi tanulási módszert alkalmazó kutatások eredményeit mutatom be.

GÉPI TANULÁST ALKALMAZÓ KUTATÁSOK

A legtöbb természetesnyelv-feldolgozást használó depresszió témájú kutatás azt célozza, hogy a gép ismerje fel a depresszióval küzdők által írt szövegeket és ezáltal online szövegek alapján azonosítsa a depressziót, sőt, akár a potenciális öngyilkossági szándékot is.

Nguyen munkatársaival (2014) ennek érdekében egy általános, különböző témájú fórumoknak helyet adó oldal speciálisan depresszió, bipoláris zavar, szeparáció,



TEMATIKUS TANULMÁNYOK – Digitális szociológia

öngyilkosság és önsértés, önkárosítás témájú csoportjainak hozzászólásait vetette össze általános (például étel, állatok, divat) témájú csoportjainak hozzászólásaival. Előbbiből több mint 30 000, utóbbiból közel 230 000 hozzászólást gyűjtöttek le. A két csoportba sorolás nyelvi és tematikus jegyek alapján történt. A gépi tanulóval létrehozott tematikus csoportokon (topik modell eredményein) alapuló modell magas, 93 százalékos pontossággal működött. A témacsoportok, amelyek a besorolást segítették, olyan általános és egyértelműen a depresszióhoz sorolható kifejezések mellett, mint *suicide*, *mental health*, *cut*, *depression*, vagy *bipolar* olyan szavakat is tartalmaztak, mint a *meds*, *heart*, *pain*, *away*, *mother*, *father*, vagy *normal*, amelyek a biomedikális keretezésre és a társas kapcsolatokra, így szociológiai keretezés megjelenésére engednek következtetni a vizsgált oldalon zajló párbeszédekben is.

Thaul Lehrman és munkatársai a fórumhozzászólásokban megjelenő káros feszültség, vagyis distressz azonosíthatóságát vizsgálták (2012). Négy kategóriát határoztak meg: magas distressz, alacsony distressz, válaszhozzászólás, és boldog hozzászólás. A kutatás során kényelmi mintát alkalmaztak, a 200 hozzászólást különböző mentális egészséggel foglalkozó fórumokról úgy választották ki, hogy a hozzászólások fele distressz, másik fele pedig egyenlő arányban boldog és válaszhozzászólás legyen. A distressz csoportba tartozó hozzászólásokat ezt követően tovább bontották, elkülönítették az alacsony és a magas distressz szintű hozzászólásokat. A felderítő kutatás során különböző osztályozási módszereket (döntési fákat, Naiv Bayes, és maximum entrópia modelleket) használtak annak érdekében, hogy képet kapjanak a modellek teljesítményéről, amelyeket a későbbiekben nagy adatbázison terveztek alkalmazni. Eredményeik szerint a gépi modellek 20 százalékkal nyújtottak jobb teljesítményt, mint a véletlen allokáció.

O'Dea és munkatársai (2015) öngyilkosságra utaló tweeteket osztályoztak három kategóriába aszerint, hogy mennyire látszik öngyilkossági szándék a tweetek mögött. A három kategória a következő volt: (1) az öngyilkossági szándék erősen feltételezhető („*strongly concerning*”); az (2) öngyilkossági szándék előfordulhat („*possibly concerning*”); és (3) biztonsággal figyelmen kívül hagyható a hozzászólás („*safe to ignore*”). Összesen 14 701 tweetet töltöttek le 2014 februárja és áprilisa között, s egy abból vett 2000 elemű véletlen mintát kódolt be két kódoló a három kategória valamelyikébe. A modellezés során ennek a 2000 elemű bekódolt mintának egy részét biztosították a gép számára „tanulás” céljából és a maradékon tesztelték a tanulás utáni besorolás hatékonyságát. A klasszifikáció során logisztikus regressziót és support vector machine-t alkalmaztak különböző modellspecifikációkkal. Eredményeik szerint elkülöníthető az öngyilkossági fenyegetettség mértéke humán kódolók munkájára épített gépi tanulási modellekkel, bár a modellek teljesítménye még javítandó egy igazán megbízható kategorizáció érdekében. (A modellek 55-76 százalékos átlagos pontosságot mutattak attól függően, milyen algoritmust és milyen specifikációkat alkalmaztak.)



TEMATIKUS TANULMÁNYOK – Digitális szociológia

A GÉPI TANULÁS ELŐNYEI ÉS HÁTRÁNYAI

A gépi tanulással történő elemzés elsődleges előnye a rendelkezésre álló rengeteg adat feldolgozhatósága. Az online elérhető adatok hagyományos kvalitatív szövegelemzéssel történő feldolgozása magán hordozza a kvalitatív eszköztár olyan hátrányait, mint a nehéz reprodukálhatóság vagy a korlátozott általánosíthatóság, amely a relatíve kis elemszámból adódik. A természetesnyelv-feldolgozás alkalmazásának segítségével jelentősen nagyobb mennyiségű szöveget reprodukálható módon lehet feldolgozni, amely humán erőforrással amúgy képtelenség lenne. A rendelkezésre álló teljes adatmennyiség (például egy fórum összes hozzászólása, vagy egy adott kulcsszót tartalmazó összes hozzászólás) így hiánytalanul feldolgozható, az elemzett szövegek egyértelműen körülhatárolhatók ezáltal jobban azonosíthatóak a kutatás megfigyelésének határai. Szintén a gépi feldolgozás előnye az adatok feldolgozásának és így az eredmények jobb reprodukálhatósága és visszakövethetősége, amely a programozott lépéseknek köszönhető.

A gépi tanulással végzett kutatások hátránya a magyar kutatóközösség számára, hogy egyes témákban, akár csak jelen vizsgálat témája, nem áll rendelkezésre elegendő adat ahhoz, hogy a gépi tanuló algoritmusok megfelelően tudjanak működni. Amiatt, hogy a szövegelemzés nyelvspecifikus, nagy mértékben függ az osztályozás sikeressége attól, az adott nyelvre alkalmazható, rendelkezésre álló programok milyen minőségűek, mennyire megbízhatók. Szintén megnehezíti a szociológiai kutatásokban való alkalmazhatóságát, hogy jelentős programozási ismereteket igényel, emiatt kevésbé gyorsan tanulható, mint egy felhasználóbarát elemzőprogram (például az ATLAS.ti vagy az NVivo). Ugyan a gépi tanulás elsődlegesen ott válik a szociológia előnyére, hogy a nagy mennyiségű adat válik általa feldolgozhatóvá, ez egyben hátránya is, ugyanis a szövegek osztályozását rábízuk a gépre, így nem ellenőrizhető minden szöveg esetén, hogy valóban megfelelő osztályba sorolta-e a gép.

Az online elérhető, felhasználók által írt szövegek elemzésének hátránya, hogy széles közönség számára elérhetőek a honlapok, fórumok, és a kutatóknak nincs pontos ismerete arról, kik írják ezeket a szövegeket, tehát nincs egy jól definiálható populáció, emiatt az adatokból nem is vonhatunk le általánosítható következtetéseket, azokat csak a megfigyelt honlapok diskurzusaira (és nem a felhasználókra) tudjuk megtenni.

Témám érzékeny aspektusa szempontjából kiemelendő a szövegek automatikus kategorizálásának egy másik előnye is, miszerint nem embereknek kell feldolgoznia több tízezer, depressziós személyek által írt szöveget. A gépi feldolgozáshoz elengedhetetlen bizonyos mennyiségű szöveg elolvasása, de ezen szám alacsonyan tartásával a lelki teher is kisebb lehet, amely a kutatásban közreműködőket érheti.



TEMATIKUS TANULMÁNYOK – Digitális szociológia

Adatok és módszerek

A kutatás során használt, angol nyelvű adatbázist az RC²S² kutatócsoport⁶ keretében végzett kutatás során állítottuk elő. A kutatásban az angol mellett magyar nyelvű szövegek elemzését is célul tűztük ki, magyar nyelven azonban nem volt elérhető megfelelő mennyiségű fórumhozzászólás. A depresszióval foglalkozó online fórumokat a „depression forum”, „depression online” kifejezésekre történő kereséssel találtuk meg. A keresés célja annak modellezése volt, ahogyan egy érintett a gyakorlatban rátalálhat ezekre a fórumokra. A keresés eredményeit a következő elvek alapján szűkítettük: csak olyan fórumok hozzászólásait gyűjtöttük le, amelyek az utóbbi három évben is aktívak voltak és amelyek hozzászólásait regisztráció nélkül lehet olvasni. Az így előálló 19 fórum között szerepelnek szűken a depresszióval foglalkozó honlapok fórumai,⁷ általánosan (mentális) egészséggel foglalkozó honlapok fórumai,⁸ illetve a felhasználók által akármilyen témában indított fórumokat tartalmazó honlap depresszió témájú fórumai is.⁹ Ez utóbbiak annak érdekében kerültek az adatbázisba, hogy ne csak tematikus, orvosi vagy pszichológiai oldalak fórumai, hanem általános fórum is szerepeljen, így lehetőség szerint csökkentve a honlapok tematikájából eredő torzító hatást.

A honlapokról ezt követően legyűjtöttük a 2016. februárja és 2019. februárja közötti összes hozzászólást a SentiOne webes szöveganalitikai platform adatgyűjtési szolgáltatásával. Ezt követően a nyers adatbázison az alábbi, kétlépcsős szűrést alkalmaztuk a kutatócsoportban hozott döntés alapján:

(1) Csak azokat a beszédflowamokat¹⁰ (úgynevezett threadeket) választottuk ki, amelyekben előfordul a „depression” vagy „depressed” szó.

(2) A kiválasztott beszédflowamokból pedig csak azokat a hozzászólásokat tartottuk meg, amelynél a beszédflowam címében, a beszédflowam linkjében vagy magában a hozzászólásban szerepel a depresszió szó, vagy annak valamilyen szinonimája.¹¹

⁶ A tanulmány alapját képező, kiinduló kutatást Németh Renátával, Sik Domonkossal és Katona Eszterrel hajtottuk végre. Az itt bemutatott eredmények a kutatás annotált adatbázisának általam történő újbóli előfeldolgozására és a szintén általam készített új modellekre alapulnak.

⁷ például depression-understood.com

⁸ például mentalhealth.com, healtunlocked.com

⁹ például reddit.com

¹⁰ Beszédflowamok alatt azokat a fórumtopikokat értem, amelyeket egy-egy felhasználó, bizonyos témában, kérdésben indít el.

¹¹ A szinonimák: depression, depressed, bummer, desolation, desperation, moody, upset, gloom, hopelessness, depressant, melancholia, sorrow, unhappiness, feeling blue, depressive, depressive disorder, unipolar depression, bipolar, bipolar depression, major depression, mdd, persistent depressive disorder, pdd, cyclothymia, mood disorder, adjustment disorder, chronic fatigue syndrome, cfs, premenstrual dysphoric disorder. A szinonimák kiválasztásánál ügyeltünk arra, hogy amennyire lehetséges, a találatok keretezési módjait ne befolyásolják a kiválasztott szavak, így köznapi kifejezéseket és szakszavakat is használtunk. Az esetleges torzításokat enyhítik a beszédflowam címében vagy a linkjében előforduló szinonimák miatt bekerülő teljes beszélgetések.



TEMATIKUS TANULMÁNYOK – Digitális szociológia

A szűrést több tényező is indokolta. Technikai ok volt, hogy bizonyos honlapokról (például reddit.com) a honlapstruktúra miatt nem lehetett csak a depresszióval foglalkozó fórumokat leválogatni az adatgyűjtés során, így a teljes fórum hozzászólásai-ból a letöltés után kellett kiválogatni a kutatás szempontjából relevánsakat. Továbbá a legújított hozzászólások olvasása során azt tapasztaltuk, hogy egy általánosan depresszióval foglalkozó fórumon belül sem csak ilyen témájú beszédfolyamok fordulnak elő (például gyakori a fórumszabályokkal kapcsolatos beszédfolyam, de előfordultak filmekkel, zenéssel kapcsolatos beszélgetések is). Emellett azt találtuk, hogy a depresszióval foglalkozó beszédfolyamokon belüli hozzászólások is szerteágazóak lehetnek, így nagyon magas lehet a téma szempontjából irreleváns hozzászólások aránya. Ezen megfigyelésünk indukálta a szinonimák előfordulásának kritériumát. Ha azonban csak olyan beszélgetésfolyamokat elemeztünk volna, amelyek címében előfordul a depresszió szó, kizártunk volna olyan beszélgetésfolyamokat, amelyek például gyógyszerek vagy kezelések köré szerveződnek, de magát a depresszió szót nem tartalmazzák nevükben. Éppen ezért azon hozzászólásokat is beemeltük az elemzésbe, amelyek bár olyan beszélgetésfolyamokban fordultak elő, amelyeknek címében nem szerepelt a „depresszió” szó, azonban a hozzászólásban magában igen.

Az adatbázis tisztítása során elsőként az adatgyűjtés fázisában keletkező hibákat kezeltük: megszüntettük a duplikátumokat és olyan esetekben, amikor a válaszként érkező hozzászólásban szerepelt az eredeti hozzászólás szövege is, ez utóbbit kitöröltük. Eltávolítottuk a 20-nál kevesebb szót tartalmazó hozzászólásokat, hiszen tapasztalataink alapján azok kevésbé hordoznak elemezhető információt: inkább csak üdvözlések, megerősítések, így a depresszió értelmezése, keretezése szempontjából kevésbé nyújtanak információt, ellenben nagy mértékben megnövelik az adatbázis méretét, amely hosszabb futásidőt eredményez. A hozzászólásokból ezt követően el kellett távolítani a különböző linkeket, e-mail címeket is, amelyek szintén irrelevánsak a kutatás tárgyát illetően, ám a későbbi elemzési fázisban hátráltatták volna a modellezési eljárást.¹² A fentiek alapján leválogatott és kitisztított adatbázisban, vagyis a végleges szövegkorpuszban 67 857 hozzászólás szerepelt.

A gépi tanulás során a számítógép számára elsőként egy olyan adatbázist kellett biztosítani, amin „megtanulhatja” az általunk elvárt kategorizációt. Ehhez egy, a teljes adatbázishoz képest kisebb adatmennyiségen kézi, humán erővel be kell sorolnunk a hozzászólásokat és utána ezen kisebb adatbázis egy részét átadni a számítógépnek, hogy a besorolt hozzászólásokban mintázatokat keressen, amikről a kategóriák felismerhetővé válnak. (Ezen „tanult” mintázatok alapján fogja a gép elvégezni a teljes, nagy adatbázisban meglévő hozzászólások kategorizációját később.) A kisebb adatbázison végrehajtott, emberi erőforrással történő besorolást nevezzük annotálásnak.

¹² Annak ellenére, hogy egy link hordozhat információt arról, a hozzászólás írója miként értelmezi a depressziót, ezen linkek tartalmának felkeresése túlmutat a jelenleg használt modellek működésén.



TEMATIKUS TANULMÁNYOK – Digitális szociológia

Az annotálás folyamata

Ahhoz tehát, hogy a számítógép képes legyen felismerni azokat a mintázatokat, amelyek az egyes hozzászólásokat adott csoportba tartozóvá teszik, a kategóriákat egy kisebb adatbázison először emberi erővel szükséges megadni. Ezt a feladatot gyakran különböző online platformokon toborzott munkavállalók végzik (például Amazon Mechanical Turk, Figure Eight), azonban a hozzászólások osztályozásához szükséges elméleti ismeretek miatt kutatásunkban szociológiai tudással rendelkező, (egy fő kivételével) mesterszakos hallgatók végezték a hozzászólások osztályozását. A tíz egyetemi hallgató 4500, az adatbázisból egyszerű véletlen módszerrel kiválasztott hozzászólást osztályozott. (Az egyszerű véletlen kiválasztás nagy valószínűséggel adja a teljes adatbázis egy reprezentatív mintáját.) A besoroláshoz ötféle címke állt rendelkezésre: a három, már említett keretezés mellett bevezettük a (1) transzcendentális kategóriát,¹³ azt, hogy (2) a hozzászólás nem besorolható egyik keretezési módba sem, de a depresszióról szól, továbbá azt, hogy (3) a hozzászólás nem a depresszióról szól. Mivel nem feltételezhettük, hogy minden hozzászólás csak egy értelmezési módban íródott, ezért lehetősége volt az annotátoroknak a hozzászóláshoz második címkét (kategóriát) is adni, ha úgy látták, második keretezési mód is megjelenik a hozzászólásban.

Az annotálást követően egységesítettük a hozzászólásokhoz tartozó címkéket (vagyis a kategóriát jelölő értéket). Ezt úgy tettük meg, hogyha csak egy olyan címke volt, amely mindkét annotátornál előfordult, akkor az vált az elsődleges címkévé. Ha csak az egyik annotátor adott a három főkategóriába tartozó címkét, akkor az vált a címkévé. Ha mind az első, mind a második címke megegyezett, akkor azok, helyzetüknek megfelelően kerültek az első és a második címke helyére. Ha a két kategória sorrendje nem volt azonos, akkor a szakértői kódolás alapján történt a besorolás, akárcsak abban az esetben, ha csak eltérő kategóriába tartozó címkék szerepeltek.

Az egységesítés után előálló kategóriák eloszlásából (1. táblázat) látható, hogy az esetek túlnyomó többségében nem szerepelt második címke és az általunk vizsgált hozzászólásokban a pszichológiai keretezési mód a leggyakoribb első címke (32,7%), majd pedig a biomedikális (28,6%), a szociológiai keretezési mód pedig viszonylag ritka első kategóriaként (8,6%). Azokban az esetekben, ahol két címkét is kaptak a hozzászólások, elsősorban a pszichológiai első címkéhez tartozik szociológiai második címke, illetve a biomedikális első címkéhez pszichológiai második. Tehát az adatbázisban leghangsúlyosabb pszichológiai értelmezés mellett feltűnik a szociológiai keretezési mód is, ahogyan az elsődlegesen biomedikális nézőpont mellett is megjelenik a pszichológiai az általunk vizsgált hozzászólásokban. A modellezés során csak az első címkét vettem figyelembe, mivel a két címkével rendelkező

¹³ A transzcendentális kategóriába tartoztak olyan hozzászólások, amelyek a depresszió okaként vagy kezelési módjaként is természetfeletti, vallási vagy esetleg áltudományos tényezőket említettek.



TEMATIKUS TANULMÁNYOK – Digitális szociológia

hozzászólások esetén is az első címke a hangsúlyos, annak megfelelősége megbízhatóbb.

1. táblázat. Az első és második címkék kategóriánkénti együttes előfordulása

Első címke	Második címke					
	Nincs második címke	Bio-medikális	Pszichológiai	Szociológiai	Egyéb	Összesen
Biomedikális	858	0	327	88	11	1284
Pszichológiai	915	196	0	333	23	1467
Szociológiai	209	25	150	0	2	386
Egyéb	172	8	18	6	0	204
Nem depr.	1144	0	0	0	0	1144
Összesen	3298	229	495	427	36	4485

Forrás: Saját szerkesztés

Az adatbázis előkészítése

A szövegeket bizonyos előfeldolgozási lépéseknek kell alávetni annak érdekében, hogy egységesebb, a számítógép által könnyebben értelmezhető szöveget kapjunk, majd ezeket speciális formátumú adatbázisba kell rendezni annak érdekében, hogy az algoritmusok fel tudják azt dolgozni. Az előfeldolgozásba tartozik például a számok és a betűformátumok kezelése, a különböző módon ragozott szavak azonos alakra hozása (lemmatizálás), a több szóból álló kifejezések (szignifikáns bigramok) együtt kezelése és a valódi jelentéssel nem bíró elsősorban mondatösszetartó szerepű szavak eltávolítása (stopszavazás).

Ahhoz pedig, hogy a szöveget algoritmikus, számítógép általi feldolgozásnak vessük alá, számokká kell alakítani azt. Egy ilyen lehetséges folyamat a vektorizálás, amelynek során egy dokumentumon belül a mondat szerkezeti sajátosságokat nem vesszük figyelembe, csak azt, hogy milyen szavak szerepelnek az adott szövegben. Ennek során egy hatalmas, dokumentum-szó mátrix keletkezik, amelynek minden sora egy dokumentumot jelöl, minden oszlopa pedig egy adott szót vagy más dokumentum-jellemzőt (például azt, milyen honlapról származik a hozzászólás, vagy milyen hosszú). A mátrix cellájában pedig olyan számok szerepelnek, amelyek a szavak adott dokumentumbeli fontosságát jelzik.



TEMATIKUS TANULMÁNYOK – Digitális szociológia

Felügyelt osztályozó algoritmusok

Ahogy arra már korábban utaltam, egy ilyen, általam is alkalmazott felügyelt szöveg-osztályozási feladat esetén a modellnek először egy emberek által felcímkezett kisebb adatbázisból kiindulva meg kell „tanulnia”, fel kell ismernie azokat a hozzászólásokban lévő mintázatokat, amelyek az egyes kategóriákhoz tartoznak, s amelyek alapján meg lehet különböztetni azokat. Ezen kisebb adatbázison tanulva, a megismert mintázatok segítségével a gépnek új, a kisebb, tanuló adatbázisban nem szereplő, „még nem látott” szövegeket az előre definiált osztályok valamelyikébe kell sorolnia. Így téve lehetővé nagymennyiségű adatok automatizált osztályba sorolását. Az ilyen algoritmusok hétköznapi használata például, ha egy ügyfélszolgálatnak írt levelet annak szövege alapján az algoritmus az adott területtel foglalkozó osztályra irányít. A szövegosztályozó algoritmusok tehát a szöveget alkotó, előfeldolgozás utáni szavak, illetve egyéb, kutatók által megadott jellemzőkből kiindulva döntenek el, hogy az egyes dokumentumok melyik osztályba sorolhatók.

Az osztályozó algoritmusok új adatok besorolásában nyújtott teljesítményének megítéléséhez szükségünk van legalább még egy olyan, emberek által már felcímkezett adatbázisra, amelynek adatai (tehát az általa tartalmazott hozzászólások) elkülönülnek azoktól az adatoktól, amelyeken „tanult” a modell, tehát amelyben már megismerte a hozzászólások kategóriáját. Ennek érdekében az emberek által már bekódolt, ismert kategóriákkal rendelkező kisebb adatbázist véletlenszerűen két részre: tanító és teszt adatokra bontjuk. Előbbiekben tanítjuk a modellt, utóbbiakon pedig mérjük az alkalmazott modell osztályozási teljesítményét. A tanító adatokon tehát mind a kutatók, mind a gép ismeri a helyes besorolást, a teszt adatoknál azonban csak a kutatók ismerik a helyes besorolást, és nem bocsátják azt a gép rendelkezésére. A teszt adatokon éppen azt vizsgálják ugyanis, hogy a gép mennyire tudja eltalálni a helyes kategóriákat. A tanulási folyamat során olyan modell kifejlesztése a célunk, amely nem csak a tanító adatokra illeszkedik jól, de a teszt adatokon is jó teljesítményt nyújt, hiszen az jelzi a modell általánosíthatóságát.

Annak érdekében, hogy a nem ismert adatok besorolását a legjobb modellen hajtsuk végre, a különböző modellspecifikációkat validálni szükséges. A kutatás során ennek érdekében keresztvalidálást alkalmaztam, mely során a tanító adatokat öt egyenlő, diszjunkt részre bontottam, ahol négy rész szolgált tanító adatbázisként, egy rész pedig tesztadatként. A tanítást és tesztelést ebből következően ötször végeztem el, mindig egy egységet hagyva tesztalumnak. Ekkor a modell teljesítménye az ötszöri futtatás átlagos eredményeként írható le.

Alkalmazott modell: Logisztikus regresszió

A társadalomkutatásban jól ismert logisztikus regressziós modell a szöveganalitikában is használatos a jelen esetben is fennálló klasszifikációs problémák esetén. Szö-



TEMATIKUS TANULMÁNYOK – Digitális szociológia

vegelemzésben osztályozási problémára alkalmazott logisztikus regressziónál az egyes szavak vagy egyéb, osztályozáshoz használt tényezők értékei veszik fel a „független változók” szerepét, míg a kategóriába tartozás lesz a „függő változó”.

Logisztikus regresszió során a becslés tehát arra vonatkozik, hogy az adott dokumentum melyik kategóriába tartozik a benne szereplő szavak és hozzá tartozó egyéb információk alapján. Akárcsak a szociológiában elterjedt használata esetén, szövegelemzésnél is megkülönböztetünk kétféle modellt attól függően, hogy csak két vagy több osztályba kell az eseteket besorolni. Bináris logisztikus regresszió esetén csak két kategória van, tehát azt kell eldönteni, adott kategóriába tartozik-e a dokumentum vagy sem. Multinomiális logisztikus regresszió esetén, amikor a kimeneti változónk nem csak két értéket vehet fel, egy referenciakategóriához viszonyítva mindegyik kategóriára elvégezzük a számításokat, majd abba a kategóriába soroljuk az esetet, amelyik kategóriának legnagyobb a valószínűsége (a referenciakategóriához képest).

A logisztikus regresszió előnye, hogy a szavakhoz és egyéb jellemzőkhöz tartozó együttthatók segítik a modellek és a kategóriák interpretálhatóságát, hiszen az értékeken keresztül megérthetjük, hogy mely szavak (vagy egyéb, a modellbe beépített jellemzők) befolyásolták jelentősen az adott kategóriába sorolást (magas pozitív együtttható) és mely szavak előfordulása hatott az adott csoportba sorolás ellen (nagy negatív együtttható). Így tehát „beleláthatunk” a modell működésébe, megkapjuk azokat a szavakat vagy egyéb használt jellemzőket, amelyek a besorolás során fontosnak bizonyultak.

Fontos kiemelni, hogy a szövegelemzésben klasszifikációs problémáknál használt logisztikus regresszió esetén a modell jóságának megítélését nem a társadalomtudományból ismert magyarázómodellek esetén használatos mérőszámokkal (például AIC, vagy pszeudo R^2) mérik, hanem a tesztadatokra a modell által becsült és a valós kategória közötti eltéréssel. Tehát az alapján, hogy a tesztadatokon a modell mennyire találta el a tényleges, helyes kategóriát. Az alábbiakban bemutatom ezen mérőszámokat, mivel eredményeim elemzése során is részben ezekre fogok hivatkozni. Ilyen mérőszám az *accuracy* (pontosság), a *precision* (precizitás)¹⁴, a *recall* (visszaidézés) vagy az F1-score. A pontosság a jól besorolt esetek arányát mutatja az összes eseten belül, míg a precizitás osztályonként adja meg azon esetek arányát, amelyek az osztályba sorolt esetek közül valóban jó helyre kerültek. A visszaidézés számításánál szintén az egyes kategóriáknál számítjuk ki, hogy a kategóriába tartozó esetek mekkora arányát sorolta be valóban abba az osztályba a modell. Az F1-score a precizitás és visszaidézés harmonikus átlagaként számolandó.

¹⁴ A magyar szakirodalomban a *precision*t jellemzően pontosságnak fordítják (Sebők et al. 2016, Tikk 2007), én azonban az *accuracy* szótól való megkülönböztetés miatt a precizitás fogalmát használom (ld még: (Tan et al. 2012). Ez utóbbinak nincs bevett magyar fordítása, hol hitelességnek (Sebők et al. 2016), hol szabatoságnak (Tikk 2007) fordítják.



TEMATIKUS TANULMÁNYOK – Digitális szociológia

Eredmények

KÜLÖNBÖZŐ MODELLSPECIFIKÁCIÓK HATÁSA AZ OSZTÁLYOZÁSRA

A modellezés során különböző modellspecifikációkat alkalmaztam annak érdekében, hogy megtaláljam a legalkalmasabb modellt, továbbá azért, hogy a modellek robusztusságát is teszteljem. A különböző specifikációk eredményeit a 2. táblázat tartalmazza. A fentiekben bemutatott mérőszámok mellett minden modell esetén közöltem a Cohen's kappa értékét is, amely azt mutatja meg, a tökéletes besorolás és a véletlen besorolás különbségének mekkora részét képezi az, amennyivel a modell jobban teljesített a véletlenszerű besoroláshoz képest. A mérőszám értéke 1, ha a modell tökéletesen sorolja be az eseteket és 0, ha a véletlenszerű egyezésnél nem jobb a modell besorolása. A köztes értékek interpretálásához a Landis és Koch (1977) által leírt értelmezés szolgál ökölszabályként.¹⁵

Kezdetben a legalapvetőbb modellt használtam, amelyben csak a hozzászólások szövege szerepel egyszer stopszavazva (lásd az előfeldolgozási lépéseknél), másszor pedig stopszavazás nélkül. A két modell között minimális különbség van a stopszavazott szöveg javára (1-2. modell). Ez a különbség minden egyéb modellspecifikációnál is megfigyelhető volt, így a továbbiakban csak a stopszavazott modell eredményeit közlöm.

Mivel a modell besorolásának jósága csak mérsékelten volt jobb, mintha véletlenszerűen soroltuk volna be a hozzászólásokat (a Cohen-kappa értéke 0,45), a hozzászólásokban szereplő szavakon túl olyan változókat szándékoztam bevonni a modellbe, amelyek a szöveget kiegészítő információt hordoznak, így segíthetik a kategóriákba sorolás döntési folyamatait. Az első információ, amit beépítettem a modellbe, az a hozzászólást tartalmazó beszélgetésfolyam címe volt (3. modell). Ezt először szöveggként (bináris vektorizálással) építettem be a modellbe – így kívántam megragadni azt, hogy a kategóriába sorolásánál a beszélgetésfolyam témáját is figyelembe vegye a modell. Az eredmények ellentmondtak várakozásaimnak, ugyanis a modell rosszabb teljesítményt mutatott, mint a pusztán szövegeket tartalmazó modell. Ha azonban a beszélgetésfolyamoknak nem a címét, hanem csak azt emeltem be a modellbe, hogy melyik beszélgetésfolyamhoz tartozik a hozzászólás (4. modell), számmal jelölve az egyes beszélgetésfolyamokat, akkor a csak szöveget tartalmazó (1.) modellhez hasonló, bár szintén nem jobb eredmények tapasztalhatók.

A korábbiakhoz képest jobb besorolási teljesítmény tapasztalható azonban, ha a modell az osztályozás során figyelembe veszi azt, hogy az adott hozzászólás beszélgetésfolyama pszichológiai, egészségügyi vagy általános tematikájú fórumról származik (5. modell).

¹⁵ Eszerint 0,00-0,20 kis mértékű a javulás; 0,21-0,40-ig elfogadható; 0,41-0,60-ig mérsékelt; 0,61-től 0,80-ig jelentős és 0,81-től 1,00-ig közel tökéletes az osztályozás (Landis – Koch 1977: 165).



TEMATIKUS TANULMÁNYOK – Digitális szociológia

Nem csak a hozzászólások lelőhelyére (fórumára), hanem a hozzászólások tartalmára vonatkozó információkat is építettem a modellbe, nevezetesen azt, előfordul-e gyógyszernev a hozzászólásban. Ez az információ elsősorban a biomedikális kategória felismerését szolgálja, hiszen a gyógyszerekről zajló párbeszéd a biomedikális nézőpontot jelzi. Beépítésével a modell teljesítménye a csak szöveget tartalmazó modellhez képest néhány ezreddel jobb eredményeket mutat minden mérőszámban (6. modell).

A fent írt jellemzők kombinációjával hasonló teljesítményt nyújt a modell (7–8. modell). A gyógyszernev és beszélgetésfolyam száma, és ezek mellett a fórum csoportját tartalmazó modell is 61 százalékos pontosság és 0,45-ös Cohen-kappa értéket mutatott.

2. táblázat. Különböző modellspecifikációk eredményei

	Modell	Accuracy/ Pontosság	Precision/ Precizitás	F1-score	Cohen- kappa
1.	Stopszavazott szöveg	0,608	0,613	0,579	0,445
2.	Stopszavazás nélküli szöveg	0,605	0,611	0,575	0,441
3.	Stopszavazott szöveg, és a beszélgetésfolyam neve	0,559	0,548	0,539	0,381
4.	Stopszavazott szöveg, és a beszélgetésfolyam száma	0,605	0,607	0,576	0,441
5.	Stopszavazott szöveg, és a fórum csoportja	0,609	0,611	0,581	0,447
6.	Stopszavazott szöveg és előfordul-e gyógyszernev a szövegben	0,609	0,617	0,582	0,448
7.	Stopszavazott szöveg, és előfordul-e gyógyszernev a szövegben és beszélgetésfolyam száma	0,610	0,610	0,582	0,449
8.	Stopszavazott szöveg, és fórum csoportja, beszélgetésfolyam száma, előfordul-e gyógyszernev a szövegben	0,609	0,615	0,582	0,448

Forrás: Saját szerkesztés

A fenti eredmények alapján látható, hogy a gyógyszernev előfordulásának és a fórum csoportjának meghatározása minimálisan javította csak a modell teljesítményt a pusztán a szöveget tartalmazó modellhez képest, de egy kivételével mindegyik modell 0,61-es pontosságot és 0,45-ös kappa értéket mutatott.

A csak stopszavazott, alapvető modell tévesztési mátrixát (*confusion matrix*) vizsgálva (lásd 3. táblázat) látható, hogy a rossz besorolás elsődleges oka, hogy a modell nem ismeri fel sem a szociológiai, sem az egyéb kategóriákat, ezen belül is



TEMATIKUS TANULMÁNYOK – Digitális szociológia

a szociológiai kategóriát legnagyobb arányban a pszichológiai kategóriába sorolja. Nagyobb arányú tévesztés látható továbbá a nem depressziós kategória pszichológiába történő besorolásban, illetve az egyéb kategóriába tartozó hozzászólások közel negyedét nem depressziósnak címkézi a modell.

3. táblázat. Tévesztési mátrix a legalapvetőbb modellre

Becsült	Nem depr.	Biomedikális	Pszichológiai	Szociológiai	Egyéb	Összesen
Tényleges						
Nem depr.	126	27	66	1	1	221
	46%	10%	24%	0%	0%	
Bio-medikális	26	194	41	0	0	261
	8%	60%	13%	0%	0%	
Pszichológiai	50	30	224	2	0	306
	13%	8%	60%	1%	0%	
Szociológiai	20	3	36	5	0	64
	23%	3%	42%	6%	0%	
Egyéb	14	4	22	0	5	45
	23%	6%	35%	0%	8%	
Összesen	236	258	389	8	6	897

Forrás: Saját szerkesztés

Megjegyzés: A mátrix a tényleges és a modell által becsült kategóriák eloszlását mutatja be. A cellákban található százalék a tényleges kategóriákon belüli, vagyis sorszázalék

KATEGÓRIÁK CSÖKKENTÉSÉNEK HATÁSA AZ OSZTÁLYOZÁSRA

Már az annotálás során felmerült, hogy a pszichológiai és szociológiai keretezési mód nehezen megkülönböztethető, a társadalom és egyén közti határok a hozzászólásokban gyakran elmosódnak. Az annotált adatokban is kis számú a szociológiai kategóriába besorolt eset, amely eredményeként, mivel csak kevés eseten tanulhatja meg a számítógép a kategóriához tartozó mintázatokat, nehezebb feladat a szociológiai szövegek gyakori jellemzőit megtalálnia. Korábbi eredményeink alapján az a felvetés is megfogalmazódott kutatótársaimban (Németh et al. 2020), hogy a szociológiai keretezés kiforratlansága, intézményrendszerének hiánya nem teszi lehetővé, hogy jól felismerhető fogalomkészlet tartozzon hozzá.

Ezen hipotézis alapján megfigyeltem az osztályozó modell teljesítményét akkor is, amikor a szociológiai és a depresszió témájú, de nem besorolható (egyéb) hozzászólásokat egy csoporttá vontam össze. Vagyis a szakirodalomban megfigyelhető



TEMATIKUS TANULMÁNYOK – Digitális szociológia

biomedikális – pszichológiai – szociológiai bontás helyett elsősorban a biomedikális és a pszichológiai keretezési módot kellett elkülönítenie az osztályozónak. Összesen tehát négy kategória közül kellett a modellnek döntenie: nem depresszióról szól a hozzászólás, biomedikális keretezés, pszichológiai keretezés jelenik meg benne vagy pedig depresszió témájú, de egyéb keretezésű a szöveg. Ezt a kategorizációt már csak a korábban jobb teljesítményt nyújtó modelleken teszteltem. Az eredményeket a 4. táblázat tartalmazza, melyből látható, hogy a fentebb ismertetett modellekhez képest mindegyik új modell jobb teljesítményű, egymáshoz képest azonban továbbra is minimális eltéréseket találunk csak a modellek teljesítményében.

4. táblázat. A különböző modellek specifikációi a szociológiai és egyéb kategóriákat összevontan tartalmazó osztályokra

	Modell	Accuracy/ Pontosság	Precision/ Precizitás	F1-score	Cohen-kappa
1.	Stopszavazott szöveg	0,613	0,617	0,593	0,454
2.	Stopszavazott szöveg, és előfordul-e gyógyszernév a szövegben és beszélgetés- folyam száma	0,616	0,621	0,596	0,458
3.	Stopszavazott szöveg, és fórum csoportja, beszélgetés- folyam száma, előfordul-e gyógyszernév a szövegben	0,613	0,618	0,595	0,455

Forrás: Saját szerkesztés

Mivel a biomedikális kategória a korábbi eredmények alapján jól felismerhető és az ahhoz tartozó szókészlet is specifikus (például gyógyszernevek vagy különböző szakszók), tovább egyszerűsítettem a kategorizációs problémát három kategóriára: nem depresszió témájú hozzászólás, biomedikális keretezésű és nem biomedikális keretezésű. Ekkor a modell teljesítménye számottevően javul (5. táblázat): mindegyik modell 69 százalékos pontosságot és precizitást mutat. A Cohen-kappa értékek is 5 századot javultak. Mindazonáltal a különböző jellemzőket tartalmazó modellek között itt sem figyelhető meg nagy eltérés.



TEMATIKUS TANULMÁNYOK – Digitális szociológia

5. táblázat. A különböző modellek specifikációi a nem depressziós, biomedikális és nem biomedikális osztályokra

	Modell	Accuracy/ Pontosság	Precision/ Precizitás	F1-score	Cohen-kappa
1.	Stopszavazott szöveg	0,690	0,688	0,679	0,500
2.	Stopszavazott szöveg, és előfordul-e gyógyszernév a szövegben és beszélgetés-folyam száma	0,690	0,689	0,679	0,499
3.	Stopszavazott szöveg, és fórum csoportja, beszélgetés-folyam száma, előfordul-e gyógyszernév a szövegben	0,692	0,692	0,681	0,503

Forrás: Saját szerkesztés

Ahogy korábban is említettem, a logisztikus regresszió „független változóiként” szereplő szavak és egyéb tényezők együttthatói megmutatják, hogy melyek azok a szavak és tényezők, amelyek jelentős szerepet játszottak a kategorizáció során. A magas pozitív együttthatók azokat a szavakat jellemzik, amelyek jelenléte egy hozzászólást az adott kategóriába tartozóvá tesz, a nagy negatív értékek pedig éppen a kategóriába sorolás ellen szólnak. Az 5. táblázatban található utolsó, mindhárom tényezőt tartalmazó modellből a *nem-depressziós* kategória esetén azokat a szavakat érdemes megfigyelni, amelyek a kategória *ellen* szólnak, hiszen ezek azok, amelyekkel a modell igazából a depresszió témájú hozzászólásokat ismeri fel. A nem-depressziós kategóriánál a *depression*, *anxiety*, *depressed* és *help* szavakat találjuk legnagyobb negatív (tehát a kategória ellen szóló) együttthatóval. A biomedikális kategóriához kerülnek jellemzően azok a hozzászólások, amelyek a *medication*, *med*, *antidepressant*, *doctor*, *drug* szavakat tartalmazzák, s a kategória ellen szólnak a *friend*, *want*, *love*, *therapist*, *like*, *write* kifejezések. A nem biomedikális osztálynál magas együttthatókat találunk a *depression*, *therapist*, *psychologist*, *anxiety*, *thought* és *step* szavak esetén, míg alacsonyokat az egészségügyi kifejezéseknél, mint a *med*, *medication*, *doctor*, *antidepressant*, *drug* és *sleep*. A fontos tényezőknél tehát jól elkülönülnek a fizikai és a lelki egészséghez kötődő szavak és kifejezések.

Ha megfigyeljük ezen, legjobb modell tévesztési mátrixát (lásd 6. táblázat), akkor látható, hogy elsősorban a nem depressziós kategóriában figyelhető meg a bizonytalanság, méghozzá a nem depressziós hozzászólások nagy arányának nem biomedikális kategóriába sorolása miatt.



TEMATIKUS TANULMÁNYOK – Digitális szociológia

6. táblázat. Tévesztési mátrix az 5. táblázatban található 3. modellre

Becsült	Nem depr.	Biomedikális	Nem biomedikális	Összesen
Tényleges				
Nem depr.	94	23	104	221
	43%	10%	47%	
Biomedikális	23	185	53	261
	9%	71%	20%	
Nem biomedikális	43	23	349	415
	10%	6%	84%	
Összesen	160	231	506	897

Forrás: Saját szerkesztés

Megjegyzés: A mátrix a tényleges és a modell által becsült kategóriák eloszlását mutatja be. A cellákban található százalék a tényleges kategóriákon belüli, vagyis sorszázalék

A szociológiai és nem besorolható kategória alacsony előfordulása kézenfekvővé tette e két kategória összevonását és a biomedikális és pszichológiai típus mellett harmadik, egyéb kategóriaként történő szerepeltetését (lásd a 4. táblázat modelljeit). Azonban az irodalmak (Billings – Moos 1982) és az annotálás tapasztalatai alapján is két lehetséges kritikus pont merült fel. Az egyik, hogy a depressziós, de nem besorolható és a nem depressziós kategóriák között nem minden esetben volt egyértelmű a választás. A másik pedig, hogy a pszichológiai és szociológiai keretezés közötti határvonal sokszor nem volt egyértelmű. Éppen ezért úgy véltem, hogy megalapozott lehet egy olyan kategorizáció is, amelyben a nem besorolható kategóriát a nem-depressziós kategóriával, a szociológiai kategóriát pedig a pszichológiaival vontam egybe. Így a modellnek végül három kategóriába kellett sorolnia a hozzászólásokat: nem kategorizálható, biomedikális, pszichoszociális. A modellek eredményei a 7. táblázatban láthatók. Az eredeti (1. táblázatban látható) modellekhez képest ezek a modellek is jelentősen jobb teljesítményre voltak képesek. Ugyan a hatékonyságuk elmaradt a korábbi háromkategóriás (nem depressziós, biomedikális, nem biomedikális) modellek teljesítményétől, azonban a jelenlegi kategorizáció nem egyszerűsíti túl a feladatot és jobban megfelel a depresszió keretezési módjainak foglalkozó elméleteknek is.

Ennek a kategorizációnak a modellezés szempontjából az is előnye, hogy ez felosztás eredményezte a legkiegyensúlyozottabb tanító adatbázist: a nem depressziós esetek az adatok 29 százalékát, a biomedikális esetek a 30 százalékát, a pszichoszociális esetek pedig az adatok 41 százalékát adják, így nincs olyan kategória, amely túlzottan domináns lenne.



TEMATIKUS TANULMÁNYOK – Digitális szociológia

7. táblázat. A különböző modellek specifikációi a nem depressziós vagy nem azonosítható, biológiai és pszichoszociális kategóriákra

	Modell	Accuracy/ Pontosság	Precision/ Precizitás	F1-score	Cohen- kappa
1	Stopszavazott szöveg	0,673	0,674	0,669	0,494
2	Stopszavazott szöveg, előfordul-e gyógyszernév a szövegben és beszélgetésfolyam száma	0,676	0,678	0,672	0,499
3	Stopszavazott szöveg, és fórum csoportja, beszélgetésfolyam száma, előfordul-e gyógyszernév a szövegben	0,675	0,677	0,672	0,498

Forrás: Saját szerkesztés

Ezen koncepció 3. modelljének alkalmasságát alátámasztják az osztályba sorolás során magas pozitív vagy alacsony negatív együtthatóval rendelkező szavak, amelyek részben természetesen azonosak a korábbiakban bemutatottakkal. A gyógyszernév előfordulását és a beszélgetésfolyam számát tartalmazó modell esetén a nem-depresszió kategóriához magas együtthatókkal tartoznak például a *god, friend, guy, jesus* szavak, míg alacsonnyal a *depression, anxiety, help, mental_health*. A vallásos szavak megjelenésének az az oka, hogy a transzcendentális kategóriát, vagyis amikor a depresszióhoz vallásos, természetfeletti keretezés társult, kezdetben külön kategóriaként kezeltük, de alacsony előfordulása miatt az egyéb kategóriába került. A biomedikális diskurzust a *medication, med, doctor, antidepressant, drug, depression, dose, side_effect, 00mg,¹⁶ take* szavak segítettek felismerni, a pszichoszociálist pedig a *therapist, psychologist, depression, anxiety, step, therapy, thought, hard, write*. Ez utóbbi szavak alapján is látható, hogy a pszichológiai aspektus a domináns. A biomedikális keretezés ellen hatottak a *friend, want, love, therapist, write, like, life* vagy *others* szavak, melyek között az érzelmekhez kapcsolódó kifejezések mellett a *write* is megjelenik, amely a naplórészhez, blogíráshoz is kapcsolódhat, amely az egyik, nem medikális jellegű megküzdési módszer és a *life* és *others*, ami inkább a szociológiai keretezéshez kapcsolható. A pszichoszociális kategória esetén nagy negatív együtthatókkal bír a *medication, med, god, antidepressant, drug, doctor* szavak, de a tíz legalacsonyabb együttható között az *eat* is megtalálható.

A részletesebben ismertetett modell tévesztési mátrixa (8. táblázat) az együttműködőkhöz hasonlóan azt mutatja, hogy a modell könnyebben felismerte a nem-depressziós hozzászólásokat a korábbi legjobb modellhez képest. Ennek oka vélhetően az, hogy a transzcendentális témák fogódzót nyújtottak ehhez. Megjegyzendő azon-

¹⁶ Az előfeldolgozás során a különböző számokat mind 0-ra cseréltem.



TEMATIKUS TANULMÁNYOK – Digitális szociológia

ban, hogy a nem depressziós kategóriában továbbra is bizonytalan a modell és nagy arányban sorolja ezeket a hozzászólásokat a pszichoszociális kategóriába.

8. táblázat. Tévesztési mátrix a 2. modellre

Becsült Tényleges	Nem depr.	Biomedikális	Pszichoszoc.	Összesen
Nem depr.	145 55%	26 10%	95 36%	266
Biomedikális	33 13%	187 72%	41 16%	261
Pszichoszoc.	69 19%	22 6%	279 75%	370
Összesen	247	235	415	897

Forrás: Saját szerkesztés

Megjegyzés: A mátrix a tényleges és a modell által adott kategóriák eloszlását mutatja be. A cellákban található százalék a tényleges kategóriákon belüli, vagyis sorszázalék

Összegzés

Kutatásomban azt vizsgáltam, alkalmazható-e a természetesnyelv-feldolgozás automatikus osztályozása annak felmérésére, hogy a depresszió szakértői keretezése megjelenik-e laikus diskurzusokban. A dolgozatban először a depresszió három, szakirodalomban előforduló keretezési módjának automatikus osztályozását végeztem el logisztikus regresszió segítségével. A kapott eredmények alapján az eredeti kategorizálást többféleképpen módosítottam, amely eredményeként a modell teljesítménye jelentősen javult. Ezek a módosítások nem csak a modellek hatékonyságának növelését szolgálták, hanem azt is, hogy megértsük, egy ilyen komplex osztályozási feladat során milyen attribútumok azok, amelyek segítik és melyek azok, amelyek nehezítik a besorolási feladatot. Mindezen eredmények pedig a kutatott téma és a gépi tanuláson alapuló automatikus osztályozás mélyebb megértéséhez is hozzájárulnak.

Az eredmények azt mutatják, az alkalmazott modellspecifikációktól függetlenül, az osztályozási feladat egyszerűsítésével is 60 és 70 százalék közötti pontosság és 40-50 százalék közötti Cohen-kappa értéket érhetünk el. Annak megítélésében, hogy ez az eredmény megfelelően megbízható-e vagy sem, többféle tényezőt kell figyelembe vennünk. Az annotátorok besorolását alapul véve azt mondhatjuk, az automatikus osztályozás is hasonló teljesítményt nyújtott, mint az emberi besorolás, vagyis az annotátorok (0,45 és 0,42 az eredeti osztályozási feladatban).



TEMATIKUS TANULMÁNYOK – Digitális szociológia

Ennek egyik oka az eredeti modellben származhat abból, hogy az osztályok eloszlása nagyon egyenlőtlen volt a tanuló adatbázisban: az egyéb és a szociológiai kategória mind számában, mind arányában annyira alacsony volt, hogy a gépnek nem volt elég lehetősége mintázatokat találni az ezekhez tartozó hozzászólásokban. Amely ilyen, többféle témát tartalmazó, ezért szerteágazó szókészlettel bíró kategóriák esetén még nehezebb feladat, hiszen nincsen sok hívószó, amelyek jelenléte segítené a kategóriába sorolást.

Annak megítéléséhez, hogy az általam alkalmazott modellek teljesítménye jó-e, érdemes megvizsgálni a témához köthető, depresszióval, öngyilkossággal kapcsolatos, laikusok által írt online szövegek automatikus felismerésével foglalkozó kutatások eredményeit. A fent ismertetett depresszió témájú kutatások pontossága 60-93 százalék közötti, (O’Dea et al. 2015, Thaul Lehrman et al. 2012) Cohen-kappa értékek pedig 20 és 55 százalék közöttiek. Ezen modellek azonban nagyobb különbségekkel bíró nyelvi struktúrákon állnak: korábbi kutatások alapján ismert, hogy eltérő szóhasználat jellemzi a depresszióval küzdőket (Losada – Crestani 2016). Az általam vizsgált kérdés nem a depressziós-nem depressziós elkülönítést kívánja vizsgálni, hanem azt, a különböző értelmezési módok felismerhetők-e az automatikus osztályozó számára. A kutatások leírása alapján a jelen kutatásban használt adatbázis variábilisabb volt, ugyanis a depresszió különböző szinonimái voltak maguk a kulcsszavak, amelyeknek a hozzászólásokban elő kellett fordulnia – s nem az a depresszióhoz, öngyilkossághoz szorosabban kapcsolódó kifejezések. Ez alapján a kapott eredményeim nem mutatnak jóval alacsonyabb teljesítményű modelleket, mint a hasonló kutatások, egyszersmind az idézett kutatások is kiemelik a feladat bonyolultságát, mind az emberi kódolás, mind a gépi besorolás során. Az alkalmazott modellek érvényességét támasztja alá az is, hogy sem a szövegek részben eltérő feldolgozása, sem az osztályozást segítő egyéb jellemzők bevonása nem eredményezett jelentős változást a modellek teljesítményében.

Az eredeti osztályozási problémához képest a szociológia kategória egyéb kategóriával történő összevonása némiképpen javította a modell teljesítményét, de az eredmények nagyobb mértékű javulása igazán a szociológiai és pszichológiai kategória összevonásával történt meg. biomedikális biomedikális Ezt támasztják alá az annotálás eredményei és az irodalmak (Kokanovic et al. 2012) is a szociológiai és a pszichológiai keretezési mód a laikus beszélgetésekben nem különül el egymástól, ugyanaz az egyén mindkettőt egyszerre tarthatja érvényesnek azokat saját vagy mások helyzetére.

Az eredmények általánosíthatósága korlátozott, nem állítható, hogy a depresszió keretezésétől eltérő kérdések esetén vagy magyar nyelven is ilyen megbízhatósággal működne a modellek, ez azonban nem is célja a kutatásnak. A vizsgálat az új technológiák nyújtotta lehetőségek, a gépi tanulás szociológiai alkalmazhatóságának felméréséhez kíván hozzájárulni azáltal, hogy megvizsgálja, egy hermeneutikailag bonyolultabb feladat esetén hogyan használható, miként teljesít egy gépi tanuló



TEMATIKUS TANULMÁNYOK – Digitális szociológia

algoritmus. Kutatásom eredményei alapján azonban elmondható, hogy az általam vizsgált fórumon az alkalmazott gépi tanuló algoritmus (logisztikus regresszió) az emberi annotáláshoz hasonló teljesítményt nyújtott. A modell teljesítményét a modellspecifikációk módosítása érdemben nem befolyásolta, így robusztus modellről beszélhetünk.

Irodalom

- Aggarwal, C. C. (2018): *Machine Learning for Text*. Springer Publishing Company, Incorporated. Yorktown Heights USA
<https://doi.org/10.1007/978-3-319-73531-3>
- Beck, A. T. (1979): *Cognitive therapy and the emotional disorders*. International Universities Press. Boston <https://doi.org/10.1192/S0007125000214918>
- Billings, A. C. – Moos, R. H. (1982): Psychosocial theory and research on depression: An integrative framework and review. *Clinical Psychology Review*, 2(2): 213–237. [https://doi.org/10.1016/0272-7358\(82\)90013-7](https://doi.org/10.1016/0272-7358(82)90013-7)
- Busfield, J. (2000): Introduction: Rethinking the sociology of mental health. *Sociology of Health & Illness*, 22(5): 543–558.
- Comas-Díaz, L. (2011). Multicultural approaches to psychotherapy. In: Norcross, J. C. – VandenBos, G. R. – Freedheim, D. K. (eds.): *History of psychotherapy: Continuity and change*. American Psychological Association, 243–267.
<https://doi.org/10.1037/12353-008>
- Comer, R. J. (2013): *Abnormal Psychology*. Worth Publishers, New York
- Dessewffy T. – Láng L. (2015): Big Data és a társadalomtudományok véletlen találkozása a műtőasztalon. *Replika (92-93)*: 157–170.
- Freud, S. (1985): Mourning an Melancholia. In: Coyne, J. C. (ed.): *Essential Papers on Depression*. New York University Press, New York, 48–63.
- Gao, S.-F. – Bao, A.-M. (2011): Corticotropin-Releasing Hormone, Glutamate, and γ -Aminobutyric Acid in Depression. *The Neuroscientist*, 1(17): 124–144.
- Goldstein, D. J. – Potter, W. Z. – Ciraulo, D. A. – Shader, R. I. (2011): Biological theories of depression and implications for current and new treatments. In: Ciraulo, D. A. – Shader, R. I. (eds.): *Pharmacotherapy of depression*. Humana Press, Totowa, NJ, 1–32. https://doi.org/10.1007/978-1-60327-435-7_1
- Kamali, M. – McInnis, M. G. (2011): Genetics of mood disorders: General principles and potential applications for treatment resistant depression. In: Greden, J. F. – Riba, M. B. – McInnis, M. G. (eds.): *Treatment resistant depression: A roadmap for effective care*. American Psychiatric Publishing, Washington, DC



TEMATIKUS TANULMÁNYOK – Digitális szociológia

- Kokanovic, R. – Bendelow, G. – Philip, B. (2012): Depression: The ambivalence of diagnosis. *Sociology of health & illness*, 35(3): 377-390
<https://doi.org/10.1111/j.1467-9566.2012.01486.x>
- Labov, W. (1972): *Sociolinguistic Patterns*. University of Pennsylvania Press, Philadelphia
- Lakoff, G. (2004): *Don't think of an elephant!: know your values and frame the debate: the essential guide for progressives*. Chelsea Green Publishing
- Landis, J. R. – Koch, G. G. (1977): The measurement of observer agreement for categorical data. *Biometrics*, 33(1): 159–174. <https://doi.org/10.2307/2529310>
- Lewinsohn, P. M. (1985): *A Behavioral Approach to Depression*. In: Coyne, J. C. (ed.): *Essential Papers on Depression*. New York University Press, 150–180.
<https://doi.org/10.1016/j.jbtep.2010.10.001>
- Losada, D. E. – Crestani, F. (2016): A test collection for research on depression and language use. *International Conference of the Cross-Language Evaluation Forum for European Languages*: 28–39.
https://doi.org/10.1007/978-3-319-44564-9_3
- Martell, C. R. – Dimidjian, S. – Herman-Dunn, R. (2013): *Behavioral activation for depression: A clinician's guide*. Guilford Press, New York, NY
- Németh R. – Sik D. – Máté F. (2020): Machine Learning of Concepts Hard Even for Humans: The Case of Online Depression Forums. *International Journal of Qualitative Methods*, (19): <https://doi.org/10.1177/1609406920949338>
- Nguyen, T. – Phung, D. – Dao, B. – Venkatesh, S. – Berk, M. (2014): Affective and Content Analysis of Online Depression Communities. *IEEE Transactions on Affective Computing*, (5): 217–226. <https://doi.org/10.1109/TAFFC.2014.2315623>
- O'Dea, B. – Wan, S. – Batterham, P. J. – Cleave, A. L. – Paris, C. – Christensen, H. (2015): Detecting suicidality on Twitter. *Internet Interventions*. 2(2): 183–188.
<https://doi.org/10.1016/j.invent.2015.03.005>
- Pan, J. – Liu, B. – Kreps, G. L. (2018): A content analysis of depression-related discourses on Sina Weibo: Attribution, efficacy, and information sources. *BMC Public Health*, 18(1): 772–782. <https://doi.org/10.1186/s12889-018-5701-5>
- Ressler, K. J. – Mayberg, H. S. (2007): Targeting abnormal neural circuits in mood and anxiety disorders: From the laboratory to the clinic. *Nature neuroscience*, 10(9): 1116–1124.
- Rogers, A. – Pilgrim, D. (2005): *A sociology of mental health and illness 5th Edition*. Open University Press, Berkshire
- Sebők M. (Szerk.) (2016): *Kvantitatív szövegelemzés és szövegbányászat a politikatudományban*. L'Harmattan, Budapest



TEMATIKUS TANULMÁNYOK – Digitális szociológia

- Seligman, M. E. (1972): Learned helplessness. *Annual review of medicine*, 23(1): 407–412. <https://doi.org/10.1146/annurev.me.23.020172.002203>
- Sik D. (2018): From mental disorders to social suffering: Making sense of depression for critical theories. *European Journal of Social Theory*, 22(4): 477–496. <https://doi.org/10.1177/1368431018760947>
- Tan, P.-N. – Kumar, V. – Steinbach, M. (2012): Bevezetés az adatbányászatba – Alapvetés. Panem Kft., Budapest
- Thaul Lehrman, M. – Ovesdotter Alm, C. – Proaño, R. A. (2012): Detecting Distressed and Non-distressed Affect States in Short Forum Texts. *Proceedings of the Second Workshop on Language in Social Media*: 9–18. <https://www.aclweb.org/anthology/W12-2102>
- Tikk D. (Szerk.) (2007): Szövegbányászat. Typotex, Budapest
- Torzsa P. – Szeifert L. – Dunai K. – Kalabay L. – Novák M. (2009): A depresszió diagnosztikája és kezelése a családorvosi gyakorlatban. *Orvosi Hetilap*, 150(36): 1684–1693. <https://doi.org/10.1556/OH.2009.28675>
- World Health Organization (2017): Depression and other common mental disorders: Global health estimates. World Health Organization