Eötvös Loránd Tudományegyetem

Társadalomtudományi Kar

MESTERKÉPZÉS

Multi-label szövegklasszifikáció online fórumbejegyzéseken

Konzulens:

Németh Renáta

Készítette:

Kerekes Norbert

OA2ZIC

survey statisztika szak

2020. november

Tartalomjegyzék

„The most common type of modeling tasks no one talks about."

– Derrick Higgins
Senior Director of Data Science
(Blue Cross and Blue Shield)

1. Bevezető

Napjainkban az információs forradalom egyre nagyobb mértékű kiteljesedésével online tartalmak elképesztő méretű tárháza vált elérhetővé. Ami viszont talán még lényegesebb: ezeket az online tartalmakat nagy részben már nem „hozzáértők" hozzák létre, hanem maguk a felhasználók. Blogok, vlogok, fórumok, podcastek, bejegyzések a közösségi médián, valamint az ezekre érkező kommentek, illetve a kibontakozó eszmecserére egyre jobban alkalmas kommentszekciók rengeteg információt tárolnak, méghozzá mindenféle kérdezői közvetítés nélkül. Az online térben végbemenő diskurzusok mindinkább hasonlítanak a mindennapi élet során végbemenő személyek közti interakciókhoz: kérdéseket, ötleteket vethetünk fel, ezekre mások válaszolnak, elmondják a véleményüket, megjegyzést fűznek hozzá, vagy épp kíméletlen kritikát. De az internet megad még valamit, amit személyes kommunikáció nem: az anonimitást. Érzékeny témákról tudunk sokkal könnyebben beszélni, olyan emberekkel, akik akár sokezer kilométerre vannak tőlünk, arc nélkül, név nélkül, kiszabadulva a mindenféle megbélyegzés lehetősége alól. Problémáink megosztását sokkal kevésbé korlátozzák kulturális normák és tabuk, mint korábban.

A társadalomkutatás mind kvantitatív, mind kvalitatív ága régóta foglalkozik az érzékeny információk minél hatékonyabb felderítésével. Ezek az érzékeny információk sokfélék lehetnek, az adott társadalomra jellemző tabuk határozzák meg, mit nem osztunk meg szívesen, akár kvantifikálható adatról van szó, akár nem (pl. fizetés, vagyoni helyzet, egészségügyi állapot, alkoholfogyasztás, droghasználat, szexuális témák). Már a kvantitatív adatfelvételek terén is kitűnik a számítógépes és online lekérdezési módok és a nyitott kérdések pozitív hatása a válaszmegtagadásra az érzékeny témák tekintetében: ha nem egy élő személynek kell felelnünk, vagy egy kérdezőbiztos jelenlétében sorszámozott kérdőíven karikáznunk, nagyobb biztonságban érezzük magunkat (McNeeley, 2012). Viszont egy jó survey design sem tudja igazán jól kiküszöbölni azt a három tényezőt, ami miatt egy szenzitív témára nem válaszolunk szívesen őszintén (Tourangeau & Yan, 2007): (1) tolakodónak érezzük a kérdést (2) félünk valamiféle lelepleződéstől (3) tudjuk, hogy az őszinte válasz nem felel meg az elfogadott társadalmi normáknak. Ezen felül természetesen a kérdés- és válaszsorrend, illetve ezek megfogalmazása, az adatfelvétel megrendelője, az adatfelvétel célja mind tovább torzítja az érzékeny információkra kérdező kérdésekre adott válaszok minőségét. A szenzitív témák mentén definiálni tudunk szenzitív populációkat (McNeely, 2012), akiknek a

lekérdezése különös gondossággal összeállított módszertant igényelne. A mentális betegségekkel és zavarokkal kapcsolatos társadalmi stigmák miatt az ebben érintett egyének is szenzitív populációnak számítanak (Seeman et al., 2015).

Holtz és munkatársai (2012) szerint, miután németországi online fórumokon vizsgálta a neonáci diskurzust illetve a muszlimok életét, az internetes adatok az érzékeny témákban könnyeben elérhetőek, természetesebb közegben keletkeznek, ezáltal autentikusabbak – napjainkban is számos, kérdőívekkel vagy interjúkkal nehezen vizsgálható témában jelentenek fontos alternatívát (Latkovikj & Popovska, 2020, Canoy & Topacino, 2020, Boursier et al., 2019, Leitão, 2019, Noack-Lundberg et al., 2019, Halls et al., 2018).

Szakdolgozatomban egy már korábban említett szenzitív populációt, a depresszióban és szorongásban szenvedőket fogom megvizsgálni online fórumbejegyzéseken és -hozzászólásokon keresztül. A cél egy machine learning modell létrehozása, amely a szöveg alapján képes eldönteni, hogy az adott bejegyzés milyen diskurzusban jeleníti meg a depressziót: biomedikális, pszichológiai vagy társadalmi megközelítést használ-e hozzá. Mivel ezek a kategóriák nem kölcsönösen kizáróak, ezért egy multi-label jellegű probléma megoldása a feladat – egy olyan klasszifikációs probléma, melyben az esetek nem csak egy esethez tartozhatnak, de egyszerre többhöz is.

A multi-label klasszifikáció a machine learning óriási területén egy kevésbé népszerű terület – legalábbis kevésbé, mint amennyire a probléma hétköznapisága (egy eset egyszerre több kategóriába tartozik) indokolná. Klasszikus alkalmazási területei a szövegklasszifikáció mellett a zene-, videó- és képfelismerés, illetve az biostatisztika, főként a gén- és fehérjekutatás. Akárcsak a gépi tanulás többi területén, a multi-label klasszifikáció esetében is sokféle algoritmusból (sőt, algoritmuscsaládból) válogathatunk. Célom ezeknek az algoritmusoknak az összehasonlító vizsgálata a fent említett klasszifikációs probléma tükrében.

A szakdolgozatom második, szakirodalmi részében áttekintem a probléma pszichológiai oldalának (a depressziónak és a szorongásnak) a főbb aspektusait, majd a multi-label klasszifikációs feladat, illetve az annak megoldására használatos algoritmusok releváns irodalmát. Megvizsgálom a használható illeszkedésmutatókat, illetve ezeknek a probléma multi-label jellegéhez való viszonyát. A harmadik fejezetben röviden bemutatom az adatbázist, mely szerencsémre tisztítva, címkézve, azon számos szövegelemzési előkészületet elvégezve kaptam meg. Végül pedig Pythonban, az Scikit-multilearn library segítségével

alkalmazom a korábban bemutatott klasszifikációs módszereket a problémára, melyet az említett illeszkedésmutatók segítségével fogok összehasonlítani.

2. Szakirodalmi áttekintés

2.1. Depresszió

2.1.1. Fogalmi tisztázás

Ahhoz, hogy vizsgálhassunk egy témát, törekednünk kell a minél nagyobb fokú megértésére. A depresszió és a szorongás két jól ismert és gyakran használt fogalom, a hétköznapi életünk során is számtalanszor emlegetjük. A mindennapokban általában szabatosan használjuk a kifejezést, és inkább egy hangulat, egy érzelmi állapot leírására használjuk őket, de mindkét szó értelmezhető mélyebb, tudományosabb keretek közt. A kutatás során szintén hús-vér emberek bejegyzéseit, kommentjeit vizsgáljuk, ezért fontos tisztában lennünk mindkét értelmezéssel.

A *depresszió, mint hangulatállapot*, egy negatív érzelmekkel telt időszak, amely hatással van a gondolatainkra, az érzelmeinkre, a motivációnkra, a viselkedésünkre, a szubjektív jól-létünkre. A depressziós emberek gyakran éreznek levertséget, reménytelenséget, súlyosabb esetben öngyilkos gondolatok kerülgetik őket. A depresszió lehet rövid- vagy hosszútávú (de Zwart et al., 2019). Fő tünete az örömtelenség, az érdeklődés elvesztése olyan aktivitások iránt, amik általában örömmel töltik el az embert (Gilbert, 2007). A depresszió, mint hangulat, maga is tünet, súlyosabb mentális zavart is jelezhet, ilyen például a súlyos depressziós zavar (major depressive disorder, vagy MDD). A hétköznapi nyelvben a depresszió, mint hangulat, és a depresszió, mint mentális zavar, nincs elkülönítve: mindkét jelenségre ugyanazt a kifejezést használjuk. A depressziós zavart fontos elkülöníteni a depresszív hangulattól, mivel utóbbi általában a szomorúság velejárója, e módon a hétköznapi élet természetes része.

A *depresszió, mint mentális zavar*, sokféle tünettel rendelkezik, melyek a depresszió típusától és az érintett személytől függően változhatnak. Vannak, akik konstans szenvednek a tünetektől, de gyakran depresszív periódusokról beszélünk, amik közt „normális” időszakok vannak (American Psychiatric Association, 2013). Életkort tekintve a legérintettebbek a húszas, harmincas éveikben járók, nemet tekintve pedig a nők esetében kétszer nagyobb az előfordulás. 2017-ben világszerte csaknem 163 millió ember szenvedett depresszióban – és ők csak azok, akiket ténylegesen diagnosztizáltak is (Kessler & Bromet, 2013, Disease and Injury Incidence and Prevalence Collaborators, 2018).

A tünetek alapvetően az emberi működés öt nagyobb területére hatnak ki (Comer, 2015). *Érzelmi* tünetek közé soroljuk a szomorúságot, az elutasítottság, szánalmasság, üresség, megalázottság érzését. Egyesek tapasztalhatnak szorongást, dühöt, idegességet. *Motivációs* tünetként az depressziós emberek gyakran elvesztik a vágyat a korábban örömöt okozó tevékenységek iránt. Az akaraterő, a spontaneitás, a munkakedv, az étvágy, a szexuális vágyak csökkenése mind motivációs tünet. *Viselkedésbeli* tünet az aktivitás és a produktivitás lecsökkenése, az egyedüllét iránti vágy, egyes esetekben a beszéd lelassulása. *Kognitív* tünet a határozott negatív énkép, kisebbségérzés, önhibáztatás, súlyos pesszimizmus, reménytelenség érzése. Depresszióban szenvedők gyakran beszámolnak arról is, hogy az intellektuális képességeik lecsökkennek. Súlyos kognitív tünetnek számít az öngyilkos gondolatok megjelenése, ami a tehetetlenségből fakad. *Fizikai* tünetek közé pedig tartozhat a fejfájás, emésztési zavarok, szédülés, alvászavar vagy általános fájdalom. A fizikai tünetek miatt a depressziót gyakran félrediagnosztizálják és orvosi problémával próbálják magyarázni azokat (Parker & Hyett, 2010).

A depresszió fajtája befolyásolhatja ezeknek a tüneteknek a súlyosságát, megjelenését. A *melankolikus depresszió* tünetei közt megtaláljuk a nagymértékű étvágytalanságot, ebből fakadóan a súlyvesztést, míg az *atipikus depresszió* pont megnövekedett étvágyat okoz. Ez utóbbi aluszékonysággal (hiperszomniával) is jár, ellenben a *szorongásos depesszió* alvászavarokat okozhat (American Psychiatric Association. 2000, Yang et al., 2014). Nem minden típusú depresszió érinti a teljes populációt: a *terhesség utáni vagy közbeni depresszió* a nőket sújtja szülés előtt vagy után, incidenciahányadosa a becslések szerint 10-15% körüli a kismamák körében (Nonacs, 2019).

A depresszió egy rendkívül szerteágazó és összetett jelenség, a tünetek sokfélék, mégis gyakran hasonlóak, egyik mentális zavar jelezheti a másik jelenlétét. Gyakoribb fajtái elterjedtek, a ténylegesen diagnosztizált esetek száma (mely valószínűleg elmarad az igazságtól) magas. A témával foglalkozni tehát korántsem csak elvont pszichologizálás, embermilliók mindennapi életére komoly hatást gyakorló jelenségekről van szó. Az okok sokfélék lehetnek, fontosságukkal kapcsolatban nincs kialakult konszenzus, de abban többé-kevésbé megegyezés van, hogy az okokat egy hármas modell szerint kategorizálhatjuk.

2.1.2. Tudományos megközelítések

A depresszió létrejöttéért felelős tényezőket egy hármas rendszer, a bio-pszicho-szociális modell alapján kategorizálhatjuk. A modell hetvenes évekbeli bemutatása (Engel, 1977) óta a három kategória folyamatosan tágul, és az különösen a társadalmi pillérre igaz: olyan absztrakt, tág fogalmakat is magába foglalt, mint a vallás és a kultúra. Hagyományosan biológiainak tekintjük a fizikai egészséget, a genetikai múltat valamint gyógyszerek hatásait, pszichológiainak a megküzdési készségeket, szociális képességeket, az énképet és a mentális egészséget. Társadalmi okok közé tartoznak a családi körülmények, családon belüli kapcsolatok, párkapcsolatok és társas kapcsolatok (Borrel-Carrió et al., 2004). A modell a depresszió keletkezésére nem ad egyértelmű, megdönthetetlen választ, de széles körben elfogadottá tette, hogy a három aspektus kölcsönhatásban van egymással.

A különböző tudományos megközelítések nemcsak a kiváltó okok, hanem a kezelési módok és lehetőségek területén is elkülönülnek. Noha az esetek 20%-ában a tünetek kezelés nélkül is mérséklődni kezdenek egy idő után (Posternak & Miller, 2001), kezelés nélkül nagy a visszaesés kockázata, egy élethosszra az átlag négy epizód (Fava et al., 2006, Limosin et al., 2007). Kezelés nélkül a tünetek krónikussá is válhatnak (Culpepper et al., 2015), de randomizált kontrollteszt kimutatta, hogy a tünetek elmúlását követően még három hónapig fenntartva a kezelést, a visszaesés valószínűsége mintegy 70%-kal lecsökkenthető (Geddes et al., 2003).

2.1.2.1. Biológiai megközelítés

A biológiai, esetleg más néven orvosi vagy biomedikális megközelítés létjogosultságák orvosi tanulmányok sora támasztja alá. A területen azonban nem uralkodik abszolút konszenzus – szervezetünk sokféle funkciójának zavara lehet az orvosi magyarázat kulcsa, azonban az, hogy ezek hogyan hatnak egymásra, vagy milyen együttes előfordulás váltja ki a zavarokat, még nem véglegesen tisztázott.

A depresszió *genetikai* magyarázatára sok és sokféle kísérlet és tanulmány készült, ami alapján valószínű, hogy a depresszióra való hajlam örökölhető (Elder & Mosack, 2011). A molekuláris biológiai vizsgálatok közelmúltbeli szaporodása esélyt adott arra is, hogy azonosíthassuk azokat a konkrét kromoszómákat, amik felelősök lehetnek a hajlam

örökítésében (Kamali & McInnis, 2011), azonban genomszinten egyelőre még nem állnak rendelkezésre szignifikáns bizonyítékok (Levinson & Nichols 2018).

Biokémiai faktorok is magyarázhatják a depressziót, külön elmélet tárgyalja a neurotranszmitterek szerepét. Ezek olyan ingerületátvivő anyagok, amik kapcsolatban állnak a hangulatszabályozással, az alvásritmussal, emésztéssel, a fight or flight („üss vagy fuss") reakcióval, motiváció érzetével. Két neurotranszmitter, a norepinefrin és a szerotonin alulműködése erősen kapcsolódik a depresszióhoz (Goldstein et al., 2011). Más kutatások az endokrin rendszer szerepét vizsgálják. Az endokrin rendszer részei, a mirigyek hormonokat termelnek, amik közvetlenül a véráramba kerülnek. A depresszióban szenvedőknek általában abnormálisan magas a kortizol szintje, ami egy mellékvese által stressz hatására termelt hormon (Gao & Bao, 2011, Veen et al., 2011).

Érdemes megemlíteni a *neurológiai* megközelítést, ami az agyi képalkotó eljárások fejlődésével párhuzamosan kimutattak bizonyos mentális zavarok és az agyunk neurális hálózatának különböző pályái közti kapcsolatokat. A szorongásos zavar, a pánikbetegség és a kényszerbetegség (OCD) után a depresszió területén is születtek eredmények, amik a prefrontális kéreg, a hippokampusz és az amigdala szerepét vetik fel a depresszió kialakulásában (Brockmann et al., 2011).

A fent említett, főbb megközelítések mellett természetesen számtalan más magyarázó tényező is felmerült az orvostudomány területén a depresszióval kapcsolatban. Vizsgálják az immunrendszer (Blume et al., 2011), a különböző tápanyaghiányok (Anglin et al., 2012), a cirkadián ciklus (Carlson, 2013) szerepét.

A biomedikális megközelítés nem korlátozódik a depresszió magyarázatára, annak kezelésében is jelen van. A legelterjedtebb a gyógyszeres kezelés, melynek az első fajtája, a monoamin-oxidáz gátlók tulajdonképpen egy véletlennek köszönheti létét: a tébécés betegek kezelésére használt iproniazid nem várt mellékhatásként a depressziós tüneteket is csökkentette (Sandler, 1990, Kline 1958), a hatóanyag azonban csak szigorú diéta betartásával használható. Ehhez hasonlatosan egy másik fajta gyógyszer, a triciklikus antidepresszánsok is véletlenül kerültek tudományos fókuszba: hatóanyagukkal először a skizofréniát akarták kezelni, amiben hatástalannak bizonyult, ellentétben viszont a depresszió tüneteit pár nap szedés után hatékonyan csökkentette (Kuhn, 1958). A második generációs antidepresszánsok, melyek többsége szelektív szerotonon-visszavétel gátlókként is ismertek, csak a közelmúltban kezdtek igazán teret nyerni, de hatásosnak bizonyultak a depresszió kezelésében: a legtöbb

ma ismert antidepresszáns (Zoloft, Prozac, Lexapro) ezek közé tartozik (Ciraulo et al 2011). A kezelés meghatározásakor figyelembe veszik a kórtörténetet, a gyógyszeres kezelést pedig napjainkra sok helyen már csak a legszükségesebb (súlyos) esetekben rendelik el, mivel a kockázat-haszon rátája alacsony (NICE, 2009). Az antidepresszánsok pozitív hatása érdekében és a mellékhatások minimalizálása végett a gyógyszeres kezelés során gondos finomhangolás, nem ritkán különböző gyógyszerek együttes alkalmazása szükséges (Fochtmann & Gelenberg, 2005).

Az elektrokonvulzív terápia (korábbi, jobban ismert nevén elektrosokk terápia) abból az elképzelésből indult ki, hogy az epilepszia és más mentális betegségek egymást kizáróak, tehát előbbi előidézésével utóbbiak kezelhetők (Berrios, 1997). Napjainkban már csak engedéllyel, végső esetben használható módszer (Beloucif, 2013).

2.1.2.2. Pszichológiai megközelítés

A depresszió megközelítései közül talán a pszichológiai értelmezések a leginkább támogatottak, mind tudományos, mind hétköznapi értelemben. Ahogy a biológiai megközelítések tekintetében, úgy itt sem beszélhetünk mindent átfogó, konszenzuális értelmezésről, a különböző iskolák eltérő okokat feltételeznek a depresszió hátterében, ilyenformán kezelések tekintetében is eltérőek a módszerek.

A *pszichodinamikus* felfogás, amit Freud és Abraham neve fémjelez, a depressziós állapot és a gyász közötti párhuzamból indul ki. A koncepció szerint a depresszió hátterében mindig valamilyen valós vagy szimbolikus veszteség áll (Homans, 1999), de a későbbiekben a gyermekkori, szülőkkel való kapcsolat is felmerült: sok depresszió által sújtott beteg gyerekkorára jellemző az alacsony szintű törődés, de magas szintű szigorúság szülők irányából (Martín-Balnco et al., 2004). A pszichodinamikus megközelítésnek viszont megvannak a maga korlátai, sok inkonzisztens eredmény született különböző kutatások során (Parker, 1993), illetve a depresszió-veszteség kapcsolat univerzalitása is megkérdőjelezhető (Bonanno, 2004, Paykel & Cooper, 1992).

A pszichodinamikus terápiás kezelési módok az elméletnek megfelelően a szabad asszociációra épülnek, melynek során a terapeuták segítenek a pácienseknek olyan múltbéli eseményeket és érzéseket felidézni, melyek okozhatják a jelenlegi állapotukat (Busch et al.,

2004). A cél a tudatosság megteremtése, a veszteségek és a depressziós állapot közti ok-okozati összefüggés megértése, ezáltal irányítása.

A *behaviorista* elméletek (nevükhöz hűen) leginkább a maladaptív cselekvések szerepeit vizsgálják a depresszió kialakulásában és fenntartásában. Egy funkcionalista megközelítése (behaviorial activation) szerint a depressziós emberek hajlamosak olyan módon viselkedni, amivel fenntartják állapotukat. Ezek a viselkedési módok leginkább elkerülő mechanizmusok, amik elállják a kontrol visszaszerzésének és a pozitív megerősítéseknek az útját (Jacobson et al., 2001). Egy másik megközelítés szociálisabb nézőpontot kínál: eszerint a depresszióban szenvedő emberek kevésbé kommunikálnak másokkal, a kevés alkalmakkor pedig cselekvéseik diszfunkcionálisak. Ez szociális izolációhoz vezethet, ami tovább fokozza a negatív énképet és a magányosságot (Alloy et al., 1998, Prkachin et al., 1977). Ezzel rokon a „boldogságkeresés társadalmi normájának" elmélete, amely szerint az emberek azért keresnek kapcsolatot egymással, hogy pozitív élményeket és tapasztalatokat szerezzenek, de a depressziós emberek megszegik ezt a normát. A lehetségesen előforduló szótlanság, kedvtelenség, vagy épp hevesebb érzelmi reakció, bár általános tünete a depressziónak és egyéb mentális zavaroknak, de társas szituációkban legtöbbször mégis félreértelmezik, idegenkedésnek, zárkózottságnak, esetleg beképzeltségnek fogják fel (Alloy et al., 1998, Prkachin et al., 1977).

A behaviorista elméletek szerint megalkotott viselkedésterápia során ezért a pácienseknek segítenek újra felfedezni az örömöt okozó aktivitásokat, különválasztani a depresszív és nem depresszív viselkedési jegyeket, és feljeszteni a szociális skilljeiket (Farmer & Chapman, 2008).

A *kognitív* megközelítések inkább az alanyok gondolataira, semmint viselkedésére támaszkodnak, ez alapján alakult ki a két legbefolyásosabb elmélet: a negatív gondolatok és a tanult tehetetlenség elmélete. A *negatív gondolatok elmélete* szerint a depresszió igazából a gondolkodás negatív irányba való szisztematikus eltolódása. Beck (2002) kognitív triádja szerint állandó, körforgó kapcsolat áll fenn az én, a környezet és a jövő között. Eszerint a negatív eltolódás okait a kognitív torzítások között kell keresni (Gross, 2015), amilyen például a minden/semmi szemlélet, a túláltalánosítás, a negatív szűrés, az elhamarkodott következtetés, a jövendőmondás, az érzelmi logika, a „kell" és „kellene" felfogás, vagy a perszonalizáció. A *tanult tehetetlenség elmélete* az élet során néha megtapasztalható kontrollvesztés érzésére alapul. Seligman (1975) szerint az emberek akkor lesznek depressziósak, ha elvesztik az irányítást az életük jutalom-büntetés rendszere felett, ami miatt

saját magukat okolják. A tanult tehetetlenséget a közelmúltban felülvizsgálta a tudományos közeg, ebből jött létre az *attribúciós elmélet*, mi szerint fontos, hogy az egyén a kontrollvesztést egy olyan belső okhoz kösse, ami általános és állandó (Taube-Schiff & Lau 2008, Abramson et al 2002).

A kognitív terápiás módszer szintén Beck nevéhez kötődik, célja pedig, hogy a páciensek felismerjék és megváltoztassák negatív kognitív folyamataikat (Beck & Weisharr, 2011). Szokták kognitív viselkedésterápiának is nevezni, mivel sokban merít a behaviorista megközelítésből. A kezelés általában gyorsan eredményes (kevesebb, mint húsz alkalom szükséges), és négy fázisból áll: (1) örömteli aktivitások keresése és növelése (2) az automatikus gondolkodási sémák felismerése, okainak beazonosítása (3) a kognitív torzítások felismerése (4) a maladaptív attitűdök, jellegek megváltoztatása.

2.1.2.3. Szociológiai megközelítés

Sajnálatos, de helyénvaló kérdés, amit Ajit Avasthi (2006) feltesz a *szociológiai megközelítés*sel kapcsolatban: maradt-e relevanciájuk ezeknek az elméleteknek a modern pszichiátriai gyakorlatban? A múlt században ezek a teóriák forradalmasították a mentális zavarokról alkotott elképzeléseket, az utóbbi időben viszont előretörni látszanak a biomedikális, genetikai, neuroimaging és pszichofarmakológiai irányzatok – a pszichiátriai gyakorlat viszont még mindig őrzi multipragmatikus tulajdonságait.

Ha a szociológiai megközelítéseket akarjuk számba venni, érdemes elkülöníteni a társadalmi faktorokat, illetve a szociológiai elméleteket, mivel előbbiek sokkal inkább vitán felül állnak. Longitudinális vizsgálat (Fryers et al., 2005) szignifikáns összefüggést talált az esélyhányadosok között, ha az iskolai végzettség, munkanélküliség, vagyoni helyzetet függvényében vizsgáljuk a depresszió előfordulását.

Társadalmi faktorként számba vehetjük a családot, mint szocializációs színteret: a családi környezet és történések által okozott depresszió generációkon átívelő zavar lehet (Warner et al., 1999), amit okozhat a szülők maladaptív kognitív stílusa (Alloy et al., 2001), a szűkebb és tágabb családi környezetben uralkodó hangulat, konfliktusok (Kane & Garber, 2004, Marmorstein & Iacono, 2004), illetve a felfokozott, stresszes élethelyzetek közben a biztonság érzetének a hiánya, a szülők érzelmi elérhetetlensége (Shirk et al., 2005). Különösen fontos az ilyetén „örökség” vizsgálata azokban a családokban, amelyek pár generációra visszamenőleg

átélhettek olyan társadalmi változásokat, azok által kiváltott kollektív szenvedést és traumát, melynek utóhatásai szabadon áramlanak a generációk közt, gyakran beazonosíthatatlanul a szenvedő sokadik leszármazott által (Orvos-Tóth, 2018). A depresszió behaviorista magyarázatának része a *családi-társas perspektíva*, ami az elmélet szociális kapcsolatokkal foglalkozó részéhez kötődik: eszerint egy erőteljes, támogató családi közeg hiánya szorosan kapcsolódik a depresszió kialakulásához – házastársi konfliktusok, válás, bántalmazó kapcsolat fokozottan kitetté teszi az egyént a depressziónak (Shultz, 2007, Weissman et al., 1991).

Egy másik szociológiai megközelítés, a *multikulturális perspektíva* a depresszió jelenségét olyan inherens emberi tulajdonságokhoz köti, mint a nem, a származás vagy a társadalmi helyzet. Ahogy korábban említettem, diagnosztizált esetek alapján a depresszió előfordulása kétszeres a nők körében a férfiakhoz képest, ennek a magyarázatára pedig számos elmélet született. Az *artifact theory* szerint igazából nincs is különbség a nemek között a depresszió előfordulásának tekintetében, hanem kulturális okok miatt nehezebben ismerik azt fel a férfiak esetében, akik kevésbé hajlamosak segítséget kérni, szakértőhöz fordulni, illetve az érzelmi tüneteket kimutatni (Emmon, 2010, Brommelhoff et al., 2004). A *life stress theory* szerint a különbséget az magyarázza, hogy a mindennapi élet során a nőket több stressz éri, jobban fenyegeti őket a szegénység, megalázó munkák elvégzése és a diszkrimináció (Astbury, 2010, Keyes & Goodman, 2006). A *lack-of-control theory* erősen kapcsolódik a kognitív magyarázatok közül a tanult tehetetlenséghez: ez alapján a nők hajlamosabbak erre, mivel sok szituációban sokkal könnyebben válik belőlük áldozat, mint a férfiakból (Le Unes et al., 1980, Nolen-Hoeksema, 2002). A sokféle magyarázat sejteti, hogy ezen a területen is hiányzik a tudományos konszenzus, azonban a nemi különbségek szerepe a depresszióban egy sokat tárgyalt témává vált az utóbbi időben.

Vizsgálhatjuk a depresszió egy funkcionalista megközelítését, mely levezethető Durkheim anómia-elméletéből (Durkheim, 2003). Durkheim munkájának tárgya az öngyilkosság, de az általa kínált értelmezés a depressziót kutatva is hasznosnak bizonyulhat: gyors társadalmi változások, közmegegyezések felborulása, a kollektív reprezentáció megtörése számottevő faktor lehet a mentális zavarok kialakulásában.

Becker címkézés-elmélete (Becker, 1973) nyomán fontos megvizsgálni a stigma-jelenséget, tehát, hogy miként tekint a többségi társadalom a depresszióra. A fentebbiekben összefoglaltam a depresszió kialakulásának lehetséges társadalmi vonatkozásait, de legalább ugyanennyire fontosak a depresszió fennmaradásában szerepet játszó aspektusok. Philips és

munkatársai (2002) szerint a hangadó társadalmi csoportok viszonyulása a depresszióhoz átalakulóban van, de a kezdeti elutasítás és elítélés szerepét leginkább az ignorancia kezdi átvenni.

2.1.3. Társadalmi keretezés

Ám ha mégis beszélünk a témáról, fontos áttekinteni, hogyan is beszélünk róla, mert ez erősen befolyásolja, hogyan is konceptualizáljuk azt (Bergen, 2012, Landau et al., 2010). Entman (1993) szerint a *keretezés*, mint folyamat, nem más, mint egy koncepció bizonyos elemeinek kiválasztása és kiemelése a kommunikációban. Reali és munkatársai (2006) arra jutottak, hogy a depresszió biomedikális és genetikai magyarázata, ezzel összefüggésben pedig a gyógyszeres kezelés egyrészről csökkenteni tudja a stigma erejét: a fizikai betegségekhez hasonlóan leveszi a felelősséget a depresszióban szenvedők válláról. Másrészről viszont újratermeli a kialakult sztereotípiákat a mentális zavarokkal küzdőkről, és gerjeszti a társadalmi távolságtartást is, ilyen módon a páciensekre visszahullva akadályozza a depresszióból való kitörésüket. Ezzel szemben a pszichoszociális megközelítések segítenek rombolni a stigmát, csökkentik az előítéleteket és a félelmet.

A depressziónak viszont nemcsak a tudományos megközelítése játszik szerepet a mindennapi kereteződésében. Az utóbbi években (sőt évtizedekben) a média (és a közösségi média) nemcsak a közbeszéd részévé tette a mentális zavarokat, de narratívát, keretezést is adott neki. Ezt vizsgálva következtetéseket vonhatunk le arra nézve, hogy számolni kell a társadalmanként is kimutatható depresszió-keretezések között is. Egy tanulmány szerint (Zhang & Jin, 2017), ami tizenkét évnyi depresszióval kapcsolatos sajtómegjelenést vizsgált, a nyugati (individualista) társadalmakban a depresszió leírása epizodikus módon, egyéni történeteket és emberi vonatkozásokat bemutatva történik, a keleti (kollektivista) társadalmakban viszont nagyobb hangsúlyt kap tematikus keretezés, a big-picture kiemelése, az orvosi szemlélet, a kezelési lehetőségek. Az okok között kereshetjük az eltérő társadalmi értékek rendszerét, de érdemes megfontolni azt is, hogy a nyugati, nagyobb mértékben piacosodott sajtóipar jobban rá van kényszerülve az olvasók „becsalogatására”, és az emberi, könnyebben befogadható történetekkel jobban eladhatóvá tehető egy téma.

Végül pedig, nem szabad figyelmen kívül hagyni, hogy maguk a depresszió által sújtottak hogyan nyilatkoznak saját állapotukról – ez a szakdolgozatban bemutatott kutatás

szempontjából is hangsúlyosan releváns. Barnes mélyinterjús tanulmánya (2003) szerint a depresszióban szenvedők számára a depresszió magány, izoláció, reménytelenség, boldogtalanság. De ezek mellett fontos családi tabu, ami szégyent hozna a rokonságra. A szubjektív okok változatosak, köztük van a szervezet egyensúlyának megbomlása, stressz, családi légkör, a munka miatti nyomás, az iskolai elvárások, a bullying.

Korábban látszott, hogy már a pszichológiai elméletek nagy része sem tudta teljesen nélkülözni a társadalomnak, mint fontos faktornak a bevonását. A depressziónak nyilvánvalóan van társadalmi vonatkozása, de ezek a hatások gyakran közvetettek, hosszú távon hatnak. Maga a társadalmi keretezés is változhat (és változik is) időben és térben egyaránt. Ezek a megközelítések és keretezési módok könnyen normává alakulhatnak, meghatározó véleménnyé, amely hatással van arra is, hogy maguk a depressziós emberek hogyan vélekednek saját állapotukról.

Ez a három, az előző alfejezetekben áttekintett keretezési mód (a biomedikális, a pszichológiai és a társadalmi) adja a szakdolgozatban vizsgált multi-label feladat alapját, hiszen a narratívák között nincs, nem is lehet éles elválasztó határ, azok gyakran átfedésben vannak egymással.

2.2. Multi-label klasszifikáció

A multi-label klasszifikáció egy olyan besorolási probléma, amiben az esetek egyszerre több osztályhoz is tartozhatnak. Tekinthető a multiclass klasszifikáció egy speciális alesetének, ugyanis utóbbiban több mint két osztály létezik, de az esetek csak pontosan egybe tartozhatnak – a multi-label klasszifikációnál nincs meghatározva a felvehető címkék száma. A multi-label klasszifikáció egy jellemző területe a szövegelemzés, de rengeteg egyéb területen használják, mint például fehérje- és génosztályozás, zene-, kép- és videóklasszifikálás (Elisseeff & Weston, 2002, Boutel et al., 2004). A multi-label klasszifikációra legjobban hasonlító módszer a ranking (Tsakoumas & Katakis, 2009), amely során a címkékhez valószínűségeket rendelünk, és a legnagyobb valószínűséggel rendelkező kimenetet fogadjuk el.

Multi-label (továbbiakban: ML) feladatról beszélünk, ha egy D adatbázisunk a következő szerint épül fel:

$$D = \{(\boldsymbol{x_1}, \boldsymbol{y_1}), (\boldsymbol{x_2}, \boldsymbol{y_2}), \ldots, (\boldsymbol{x_n}, \boldsymbol{y_n})\},$$

ahol $\boldsymbol{x_i}$-k a bemeneti X tér vektorai, $\boldsymbol{y_i}$-k pedig a label setek, melynek λ elemeit a L nem üres label-halmaz tartalmazza. Ha X bemeneti, euklidészi tér p dimenziós, akkor

$$D = (\boldsymbol{X}, \boldsymbol{Y}),$$

ahol $\boldsymbol{X} = (\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_n})$ $és$ $\boldsymbol{Y} = (\boldsymbol{y_1}, \boldsymbol{y_2}, \ldots, \boldsymbol{y_n})$. $\boldsymbol{Y}$ egy eleme y_{ij}, ami 1, ha a *j*. label releváns az *i*. bemeneten. A feladat egy olyan h függvény megtalálása, ami

$$h : X \rightarrow Y = P(L) = \{0, 1\}^L$$

módon transzformál X-ből Y-ba. *P(L)* az esethez tartozó labelek power setje.

A ML klasszifikációnak több altípusa is létezik, ilyen

- a hierarchikus ML klasszifikáció, amikor a címkék valamilyen struktúrába rendezhetők (Clare & King, 2001).
- ML problémák: mikor egy esethez több kimenet van hozzárendelve, de ezek közül csak egy bizonyos számú kimenet a jó (Jin & Ghahramani, 2002).
- multiple instance probléma: mikor a feladat egy *koncepció* megtanulása, amely során a többféle felvett kimenetből csak egy szükséges a feladat megoldásához, de az eset definiálásához egyértelműen szükség van az összes kiementre (Maron & Lozano-Perez, 1997). (Babenko (2008) kulcstartó-szitucáiója.)

A ML klasszifikációs módszerek csoportokba sorolását először Tsoumakas és Katakis (2009) tette meg, két fő kategóriát meghatározva: a probléma transzformálásnak módszere, illetve

az adaptált algoritmusok módszere. Az alábbiakban ebből a két csoportból fogom bemutatni azokat a módszereket, melyeket az adatbázisok elemzése közben használtam. Fontos kiemelni, hogy a ML klasszifikációs technikák köre folyamatosan bővül, a meglévő módszerek pedig finomodnak, illetve mind újabb és újabb verziók kerülnek bemutatásra: a téma teljeskörűen átfogó ismertetése bőven meghaladja ennek a szakdolgozatnak a kereteit.

2.2.1. Probléma transzformálása

A módszer arra az elképzelésre alapul, hogy ha a ML klasszifikációs problémát single-label problémává transzformáljuk át, akkor már rendelkezésre áll számos jól ismert könnyen használható algoritmus, hogy megoldjuk a feladatot (Tsakoumas & Katakis, 2009).

Létezik két módszer, amit probléma transzformálásnak szoktak tekinteni, mivel valóban single-label klasszifikációs feladatot hoz létre, noha mindkettő jelentős információvesztéssel jár. Az egyik a felvett labelek közül random módon választ (mely választáshoz inicializálhatunk súlyokat), a másik pedig szimplán törlni az adatbázisból azokat az eseteket, amelyekhez egynél több label tartozik (Boutell et al., 2004).

X	λ_1	λ_2	λ_3
1	+	+	
2		+	
3	+		+
4	+		
5		+	+
6	+		+

→

X	λ_1	λ_2	λ_3
1	+		
2		+	
3	+		
4	+		
5			+
6			+

1. ábra: random választás a labelek közül

X	λ_1	λ_2	λ_3
1	+	+	
2		+	
3	+		+
4	+		
5		+	+
6	+		+

→

X	λ_1	λ_2	λ_3
2		+	

2. ábra: valódi ML esetek törlése

2.2.1.1. Bináris relevancia

A bináris relevancia (BR) módszere a legintuitívabb azok közül, melyek már valóban ML klasszifikáció kivitelezése használhatók. Lényege, hogy minden címkére önálló klasszifikációs algoritmust hozunk létre (címkénként akár eltérőket), ezek pedig egymással párhuzamosan és egymásra való tekintet nélkül működnek.

A megfigyeléseink egy ismeretlen P eloszlást követnek $X \times Y$-on, ahol X egy p dimenziós euklidészi tér és $L = \{ \lambda_1, \lambda_2, \dots, \lambda_q \}$ a labelek tere és q a lehetséges labelek száma. $D = \{ (\boldsymbol{x^i}, \boldsymbol{y^i}) \mid 1 \leq i \leq m \}$ a multi-label tanulóhalmaz, ahol $\boldsymbol{x^i} \in X$ egy p dimenziós feature vektor: $\left(x_1^i, x_2^i, \dots, x_p^i\right)^T$ és $\boldsymbol{y^i} \in \{-1, +1\}^q$ egy q elemű bináris vektor: $\left(y_1^i, x_2^i, \dots, x_q^i\right)^T$, ahol $y_j^i = +1\ (-1)$ jelöli, hogy a *j.* y releváns-e az *i.* bemeneten. Egy új $\boldsymbol{x}^* \in X$ eset releváns label setje legyen $\boldsymbol{Y}^*$, ami az $\boldsymbol{Y}^* = h(\boldsymbol{x}^*) \subseteq Y$ módon adható meg. A bináris relevancia módszere tehát a ML feladatot q darab bináris problémára bontja szét, amely során mindegyik probléma egy adott label megtanulásának felel meg (Zhang et al., 2018).

X	λ_1	λ_2	λ_3
1	+	+	
2		+	
3	+		+
4	+		
5		+	+
6	+		+

→

X	λ_1
1	1
2	0
3	1
4	1
5	0
6	1

X	λ_2
1	1
2	1
3	0
4	0
5	1
6	0

X	λ_3
1	0
2	0
3	1
4	0
5	1
6	1

3. ábra: a bináris relevancia működési elve

A módszer fontos tulajdonsága a konceptuális egyszerűség, hogy minden labelre külön tudja javítani a teljesítményt, hiányzó címkék esetén is könnyen implementálható. Gyengesége viszont, hogy figyelmen kívül hagyja a labelek közti korrelációt (Zhang et al., 2014), ami a legtöbb esetben csökkenti a teljesítőképességét.

2.2.1.2. Klasszifikáló láncok

A klasszifikáló láncok (classifier chains, CC) egy megoldás a bináris relevancia „figyelmetlenségére” a korrelációk felé. Az ötlet az, hogy állítsunk fel egy τ sorrendet a címkék között az alábbi injektív függvénnyel:

$$\tau: \{1, \dots, m\} \rightarrow \{1, \dots, m\},$$

ami megadja a lánc pontjainak kapcsolódási sorrendjét:

$$z_{\tau(1)} \rightarrow z_{\tau(2)} \rightarrow \cdots \rightarrow z_{\tau(m)}.$$

Legyen minden λ_k labelhez egy C_k klasszifikáló algoritmus, ami $C_k = X \times \{0,1\}^{k-1} \rightarrow \{0,1\}$, ahol $\{0,1\}^0 = \emptyset$, így

$$\hat{y}_1^{(l)} = C_1\left(x^{(l)}\right),$$

$$\hat{y}_2^{(l)} = C_2\left(x^{(l)}, \hat{y}_1^{(l)}\right),$$

$$\hat{y}_3^{(l)} = C_3\left(x^{(l)}, \hat{y}_1^{(l)}, \hat{y}_2^{(l)}\right),$$

$$\vdots$$

$$\hat{y}_m^{(l)} = C_m\left(x^{(l)}, \hat{y}_1^{(l)}, \hat{y}_2^{(l)}, \dots, \hat{y}_{m-1}^{(l)}\right).$$

X	λ_1	λ_2	λ_3
1	+	+	
2		+	
3	+		+
4	+		
5		+	+
6	+		+

→

X	λ_1
1	1
2	0
3	1
4	1
5	0
6	1

X	$\widehat{\lambda_1}$	λ_2
1	1	1
2	0	1
3	0	0
4	1	0
5	0	1
6	1	0

X	$\widehat{\lambda_1}$	$\widehat{\lambda_2}$	λ_3
1	1	1	0
2	0	1	0
3	0	1	1
4	1	0	0
5	0	1	1
6	1	0	1

4. ábra: a klasszifikáló láncok működési elve

A klasszifikáló láncok módszerének nagy előnye, hogy még magában hordozza a bináris relevancia intuitivitását, emellett viszont az egyik legjobban teljesítő ML klasszifikációs módszerként ismert (Madjarov et al., 2012). Viszont a labelek (így az algoritmusok) sorrendjének a kiválasztása lényeges kérdés. Alaphelyzetben ez random módon definiált, viszont így könnyű belefutni olyan szerencsétlen helyzetbe, mikor a lánc első szemének az alacsony accuracy-ja magával rántja az egész modellt (Dembczyński et al., 2010). Javasolt megoldás az ECC módszer, a klasszifikáló láncok ensemble-je (Read et al., 2011), illetve a GACC (genetic algorithm for classifier chains) (Gonçalves et al., 2013), amely különböző, illeszkedés jóságát mérő mutatószámok „egybegyúrásával” versenyezteti a láncokat. Ez viszont inkább akkor hasznos és fontos, ha a label-halmazunk számossága magas, hiszen egy tíz, vagy száz szemű klasszifikáló láncnál már nagyon sok múlik azon, hogy a könnyen felismerhető labeleket a lánc elejére helyezzük, és így ezek segítségével, az általuk jelenlévő korrelációs információval becsüljük a többi labelt. Ugyanezen esetben ütközik csak ki a klasszifikáló láncok

számolásikapacitás-igénye, mely nagyméretű label setek esetén nehezíti a lefutást (Read et al., 2011).

2.2.1.3. Label power-set

A label power-set (LP) megközelítés különbözik az előzőktől abban, hogy nem egy bizonyos darabszánú single-class klasszifikációs feladattá transzformálja a problémát, hanem egy darab multiclass problémává. Ezt úgy teszi meg, hogy külön kezel minden lehetséges label-kombinációt, majd ezeket a kombinációkat kezeli új labelként. Ha azt mondjuk, hogy a bináris relevancia (és egy kicsit a klasszifikáló láncok) módszere nem tudja jól kezelni a labelek egymástól való függőségét, akkor a label power-set a másik véglet, mert minden kombinációt teljesen egyediként és kezel, amikben mintha nem lenne semmi közös (Higgins, 2018). Szemléletesen: ha van egy [1, 2] és egy [1, 3] kategóriám, akkor mindkettőben benne van az információ, hogy az a két label együtt áll, de maga a power set algoritmus „nem tudja", hogy a két kombinációban az [1] közös.

Ha a lehetséges labelek halmaza L, ahol $L = \{\lambda_1, \lambda_2, \dots, \lambda_m\}$, akkor legyen L' egy olyan halmaz, melynek elemei L minden lehetséges részhalmaza. Ekkor $|L'| = min\left(D, 2^{|L|}\right)$. Labelezzük újra eszerint a D adatbázisunkat, ezután a megoldandó feladat egy szokásos klasszifikáció h klasszifikáló függvénnyel:

$$h(x) = arg\,\max_i h_i(x).$$

X	λ_1	λ_2	λ_3
1	+	+	
2		+	
3	+		+
4	+		
5		+	+
6	+		+

→

X	λ_1	λ_2	λ_3	$\lambda_1\lambda_2$	$\lambda_1\lambda_3$	$\lambda_2\lambda_3$	$\lambda_1\lambda_2\lambda_3$
1				+			
2		+					
3					+		
4	+						
5						+	
6					+		

5. ábra: a label power-set működési elve

2.2.1.4. RAKEL

A RAKEL, vagyis random k-labelsets módszere (Tsoumakas et al., 2011) a label powerset módszer egy továbbfejlesztése, aminek a célja, hogy az LP-módszer fent említett gyengeségét kiküszöbölje. Ennek érdekében LP-algoritmusok egy ensemble-jét állítja elő, melyből

mindegyik egyesével a lehetséges labelek egy random halmazán (*k* darab labelen) alapul. A módszer outputja a labelenkénti bináris klasszifikáló algoritmusok eredményének az átlaga.

L továbbra is a lehetséges labelek halmaza, így $L = \{\lambda_1, \lambda_2, \dots, \lambda_m\}$. Legyen $Y \subseteq L$, úgy, hogy $k = |Y|$ a k-labelset. L^k jelöli az összes különféle labelsetet L-en. Ekkor $|L^k| = \binom{|L|}{k}$. A RAKEL iteratívan állítja elő *m* darab LP-klasszifikáló algoritmus ensemble-jét. Minden *i.* iterációban ($i = 1, 2, \dots, m$) random módon, visszatevés nélkül kiválaszt egy *k*-labelsetet, Y_i -t, amire megtanul egy klasszifikációt: $h_i: X \to P(Y_i)$ (Tsoumakas & Vlahavas, 2007). Minden új esetre minden h_i model bináris döntést hoz minden λ_j labelre, ami benne van a megfelelő Y_i -ben. Ezután a RAKEL kiszámolja a döntések átlagát a λ_j-kre, ami pozitív outputot ad, ha az átlag nagyobb egy *t* küszöbértéknél.

Belátható, hogy ha *k = 1* és *m = |L|*, tehát egy label van a kiválasztott halmazban, és az iterációk száma megegyezik az eredeti labelek számával, akkor a bináris relevancia módszerét valósította meg az algoritmus, ha viszont *k = |L|* és *m = 1*, vagyis minden labelt felhasználunk a részhalmazok képzéséhez és csak egy iteráció van, akkor az LP-módszert kapjuk vissza. Logikus, hogy ez alapján *m* és *k* paraméterek megadhatók, $m = [1, |L^k|]$ és $k = [2, |L| - 1]$ intervallumokon.

X	λ_1	λ_2	λ_3	λ_4
1	+	+		+
2		+		+
3	+		+	
4	+			
5		+	+	+
6	+		+	

→

X	λ
1	1101
2	0101
3	1010
4	1000
5	0111
6	1010

→

X	$Y_1 \in 2^k$
1	110
2	010
3	101
4	100
5	011
6	101

X	$Y_2 \in 2^k$
1	101
2	101
3	010
4	000
5	111
6	010

…

X	$Y_m \in 2^k$
1	111
2	001
3	110
4	100
5	011
6	100

6. ábra: a random k-labelset működési elve

2.2.2. Adaptált algoritmusok

Az adaptált algoritmusok (vagy más néven algoritmus-adaptáció) felőli megközelítés alternatívát teremt a probléma transzformálásának módszerével szemben: nem a multi-label feladatot transzformálja single-label feladattá, hogy azokat már korábban kifejlesztett módszerek sorával képesek legyünk megoldani, hanem eredetileg single-label problémák megoldására létrejött algoritmusokat módosít úgy, hogy azok képesek legyenek kezelni multi-label problémákat is.

Az adaptált algoritmusok a 2000-es évek elejétől kezdtek a multi-label klasszifikációval foglalkozók figyelmébe kerülni. Ekkor fejlesztette ki Clare és King (2001) a C4.5 módszer multi-label verzióját, ami egy információs entrópiát használó döntési fa-algoritmusra alapuló eljárás (Quinlan, 1993). A multi-label-kompatibilitáshoz az entrópia képletét kellett módosítani:

$$entropy = \sum_{i=1}^{N} \big(p(c_i)logp(c_i) + q(c_i)logq(c_i)\big),$$

ahol $p(c_i)$ a c_i osztály relatív gyakorisága, és $q(c_i) = 1 - p(c_i)$.

Két másik korai adaptált algoritmus az AdaBoost.MH és az AdaBoost.MR, amely az AdaBoost (Freund & Schapire, 1997) két módosítása. Az AdaBoost.MH-ban, ha a gyenge klasszifikálók outputja poztív egy új x esetre az l labelen, akkor azt mondjuk, hogy x felveszi az l-t. Az AdaBoost.MR-ben a gyengén teljesítő algoritmusok outputja alapján minden L-ben szereplő labelre kapunk egy sorrendet (Schapire & Singer, 2000).

A felsorolt algoritmusok már észrevehetően adaptáltak a multi-label problémára, de a mélyükre nézve, probléma transzformálásból indulnak ki, mert minden (x, Y) esetük fel van bontva $|L|$ esetre (Tsoumakas & Katakis, 2009).

2.2.2.1. ML-kNN

A Zhang és Zhou által 2007-ben bemutatott ML-kNN (vagyis multi-label k-nearest neighbor) az első volt, ami lazy learning-megközelítést tett lehetővé a témában. A módszer nyilvánvalóan a kNN-algoritmusra (Cover & Hart, 1967) alapul, azzal a különbséggel, hogy amíg a kNN egy új esethez a legtöbb *k* legközelebbi szomszéd által felvett osztályt rendeli, addig az ML-kNN a leggyakrabban előforduló labeleket.

Vegyünk egy x esetet, és a hozzá tartozó $Y \subseteq L$ label-setet. Legyen $\boldsymbol{y_x}$ az x kategória-vektora, melynek az l-edik komponense 1, ha $l \in Y$, és 0 máskülönben. Emellett jelölje $N(x)$ az x eset k legközelebbi szomszédját. Ezalapján meghatározhatunk egy számláló vektort a következőképpen:

$$\boldsymbol{C_x}(l) = \sum_{a \in \mathrm{N(x)}} \boldsymbol{y_a}(l), \qquad l \in \mathrm{Y},$$

ahol $\boldsymbol{C_x}(l)$ számlálja azokat az szomszédokat x közelében, amik felveszik az l-edik labelt. Minden t teszthalmazba tartozó esetre az algoritmus definiálja az $N(t)$ halmazt. Jelölje H_1^l azt, ha t felveszi l-t, és H_0^l azt, ha nem. Ezen kívül, jelölje E_j^l $(j \in \{0, 1, 2, \ldots, k\})$ azt az esetet,

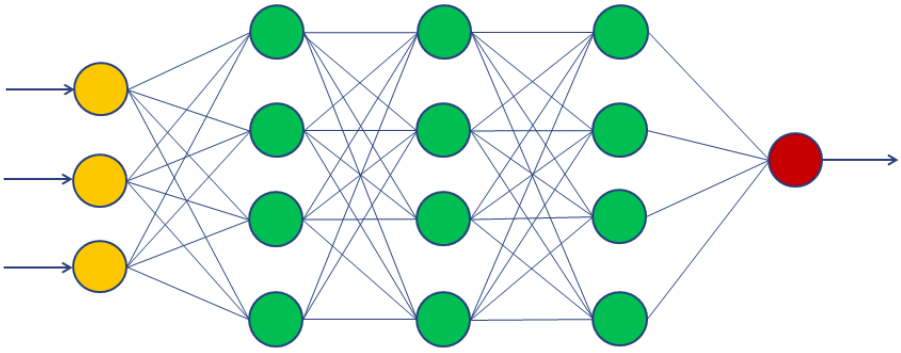
amikor t-nek a k darab legközelebbi szomszédjából pontosan j darab veszi fel az l-t. Ekkor $\boldsymbol{y_t}(l)$ megkapható maximum a posteriori (MAP) becsléssel:

$$\boldsymbol{y_t}(l) = arg \max_{b\in[0,1]} P\left(H_b^l \middle| E^l_{\boldsymbol{C_t}(l)}\right) = arg \max_{b\in[0,1]} P\left(H_b^l\right) P\left(E^l_{\boldsymbol{C_t}(l)} \middle| H_b^l\right)$$

Látható, hogy a kategória vektor becsléséhez az a priori és az a posteriori valószínűségek kellenek, amiket a tanulóhalmazból közvetlenül megkaphatunk (Zhang & Zhou, 2007).

2.2.2.2. Neurális hálók

Manapság már aligha beszélhetünk gépi tanulásról, klasszifikációról vagy modellezésről anélkül, hogy legalább érintőlegesen ne említenénk meg a neurális hálókat. A neurális hálók olyan gráf alapú modellek, melyek univerzális függvényaproximátorként használhatók, felépítésük és működésük pedig a biológiai neurális hálózatokéhoz hasonlatos (Chen et al., 2019). A rétegekbe szervezett neuronok valójában számítási egységek, melyek több bemenettel és egy kimenettel rendelkeznek (mint az igazi neuronok dendritjei és axonja). A bemenetek súlyokkal és egy eltolással kombinálva, majd ennek az eredményét egy nemlineáris aktivációs függvényen keresztülvíve előállítható a neuron kimeneti állapota, mely bináris: „tüzel” (tehát a kimenet 1), vagy „nem tüzel” (a kimenet 0). A neurális hálók tanításának a kulcsa igazából a hibavisszaterjesztés (*backpropagation of errors*), mely a gradiensereszkedés segítségével lehetővé teszi a differenciálható aktivációs függvénnyel rendelkező neuronok tanítását.



7. ábra: egy sűrűn kapcsolt neurális háló vázlatos rajza

Pont a hibavisszaterjesztés módszere az, ami a legerősebb hívószó a mesterséges neurális hálók elnevezését kritizálók számára: a biológiai neuronok működési elvében nincs ehhez hasonló folyamat (Francis, 1989). Másik fő kritika, hogy az eljárás erősen fekete-doboz jellegű, ami úgy is kivitelezhető, hogy sem a problémáról, sem a hálók működésének mélyebb elveiről

nincsenek ismereteink: Alexander Dewdney szerint: „ […] nincs benne az emberi kéz (vagy elme), a megoldások szinte mágia-szerűen képződnek, és ezért úgy tűnik, igazából senki nem tanult meg semmit.”, illetve „[…] a rengeteg szám, ami megragadná a lényegét, csak egy érthetetlen, olvashatatlan halmaz… tudományos forrásként értéktelen.” (Dewdney, 1997).

Bár a fenti kritikának van némi igazságtartalma, azt kár lenne tagadni, hogy a közelmúltban fellángoló adatforradalom idején a neurális hálók egyfajta deep-learning-svájci-bicskaként teszik a dolgukat, és ezen az eredményorientált piacon valószínűleg kevés embert foglalkoztatnak a tudományos aggályok. A neurális hálók működésének az alapjai mindazonáltal egyszerűen hozzáférhetők és könnyen emészthetők.

A neurális hálókban (pontosabban a feedforward neurális hálókban) az információ a neuronokból álló rétegekben egy irányba, „előre felé” áramlik: az első rejtett réteg neuronjai a bemeneti réteg által tartalmazott nyers információnak a lineáris kombinációját veszik, majd egy ez alapján képzett outputot adnak tovább az utánuk következő rétegnek. Legyen a bemenetünk feature-vektorok összessége, $\boldsymbol{x_1}, \boldsymbol{x_2}, \dots, \boldsymbol{x_n}$. Tekintsük a bemeneti réteget (layer) az első rétegnek, $l = 1$. Az első rejtett réteg ($l = 2$) neuronjai (i) az eredeti adatokat kapják meg, és ezeknek $\boldsymbol{w}$ súlyokkal és $\boldsymbol{b}$ eltolással vett lineáris kombinációját lépzik:

$$z_i^{[2]} = w_i^{\boldsymbol{T}} \cdot \boldsymbol{x}^{[1]} + \boldsymbol{b_i}$$

Ezután a kapott z eredményt egy $g(\,)$ nemlineáris aktivációs függvénybe helyezik:

$$a^{[2]} = g^{[2]}\left(z_i^{[2]}\right)$$

Az következő réteg neuronjai ezt az $\boldsymbol{a}$ outputot kapják meg inputként és viszik tovább. Az információ áramlása a következőképpen formalizálható:

$$z_i^{[l]} = w_i^T \cdot g^{[l]}\left(z^{[l-1]}\right) + b_i$$

A rétegek neuronjainak száma variálható, hiszen a $\boldsymbol{z}$ vektor a súlymátrixszal való szorzás által befogadhatóvá válik a következő réteg neuronjai számára.

A $g(\,)$ aktivációs függvény az egyik fontos pont a neurális hálókban, hiszen enélkül a háló csak lineáris függvények tömege lenne, így maga is csak lineáris problémák megoldására lenne képes, a neuronok tüzel / nem tüzel típusú válasza pedig nem tenné hasznosabbá, mint egy egyszerű logisztikus regresszió. Az aktivációs függvények nemlineáris volta adja a hálók flexibilitását. A manapság leggyakrabban használt aktivációs függvény a ReLU:

$$g(x) = max(0, x),$$

illetve ennek egy kis módosítása, a leaky ReLU, ami segít abban, hogy negatív tartomány esetén se ragadjon be a gradiens:

$$g(x) = \begin{cases} x & ha\ x > 0 \\ 0.01x & különben \end{cases}$$

Ha 0 és 1 közötti outputokat szeretnénk (mert például valószínűségként szeretnénk őket interpretálni), akkor pedig a sigmoid:

$$g(x) = \frac{e^x}{e^x + 1}$$

A másik fontos pont a neurális hálók tanításában a hibavisszaterjesztés. Ehhez kell egy differenciálható veszteségfüggvény. A veszteségfüggvény ($L = (\hat{y}, y)$) megmutatja, mennyire vagyunk messze az ideális eredménytől. A veszteségfüggvény értéke alapján kell módosítani a $\boldsymbol{W}$ és $\boldsymbol{b}$ paramétereket, hogy megtalálhassuk a függvény minimumát a gradiensereszkedés módszerével (*gradient descent*). A szélsőértékkeresési probléma során a veszteségfüggvény paraméterek szerinti parciális deriváltját „küldjük vissza" a hálón, pont fordított irányban, ahogy az információ áramlik, a paramétereket pedig egy hangolható ráta segítségével igazítjuk:

$$\boldsymbol{W}^{[l]} = \boldsymbol{W}^{[l]} - \alpha \boldsymbol{\partial W}^{[l]}$$

$$\boldsymbol{b}^{[l]} = \boldsymbol{b}^{[l]} - \alpha \boldsymbol{\partial b}^{[l]},$$

ahol $\boldsymbol{\partial W}$-t és $\boldsymbol{\partial d}$-t a lánc-szabály adja meg, α hiperparaméter pedig a tanulóráta, amivel az igazítás mértéke szabályozható. Ha túl kicsi, akkor a háló lassan tanul, ha pedig túl nagy, akkor könnyen átugarhatja a függvényminimumot.

A neurális hálók használata a multi-label feladatokban egy teljesértékű, különálló témaként is megállná a helyét – a szakirodalmak alapján meg is állja. A neurális hálók családja egy folyamatosan fejlődő és bővülő algoritmusgyűjtemény, a deep learning témaköre már jócskán meghaladta az egyszerűbb, sűrűn kapcsolt feedforward hálókat. Ez a multi-label problémákra fókuszálva sincs másképp: alkalmazhatók local networkök (Cerri et al., 2014), neurális hálók ensemble-je (Lenc & Král, 2017), a képfelismerésekben alapvető konvolúciós hálók (Vallet & Sakamoto, 2015), recurrent networkök (Yazici et al., 2020).

A neurális hálók olyan algoritmusok, amik natív módon alkalmassá tehetők multi-label feladat megoldására, ennek a kulcsa pedig a veszteségfüggvényben, illetve a kimeneti réteg (output layer) aktivációs függvényében van. Míg multiclass problémák esetén a kimeneti réteg neuronjai (melyeknek száma a lehetséges osztályok számával egyezik) által adott

valószínűségek összegződnek, addig multi-label problémák kezelésekor ez nem hasznos, így a gyakran alkalmazott softmax aktivációs függvény helyett sigmoid függvényt szokás használni.

Softmax aktivációs függvény: $g(x) = \frac{e^{x_i}}{\sum_{j=1}^{K} e^{x_j}}$ Sigmoid aktivációs függvény: $g(x) = \frac{e^x}{e^x+1}$

Ahogy az a függvények képletéből is látható, a sigmoid függvény külön-külön „teljes értékű" valószínűségeket ad outputként, amiket labelenként külön tudunk interpretálni és thresholdot vezetni be rájuk.

A veszteségfüggvény tekintetében szerteágazó a szakirodalom, az utóbbi években több tanulmány is megjelent, melyben egyedi, multi-label feladatokra konstruált veszteségfüggvényeket mutattak be (Du et al., 2019, Grodzicki et al., 2008, Zhang & Zhou, 2006). Egy könnyen interpretálható és implementálható (ezért gyakran alkalmazott) veszteségfüggvény a bináris kereszt-entrópia (Krakowski, 2018).

$$L = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot log\big(p(y_i)\big) + (1 - y_i) \cdot log\big(1 - p(y_i)\big),$$

ahol y a label és $p(y_i)$ annak a valószínűsége, hogy az adott label releváns az adott kimeneten. Az entrópia itt a valószínűségi eloszlásban lévő bizonytalanságot jelenti, jelen esetben pedig az általunk közelíteni kívánt ideális eloszlás és a valódi eloszlás közti Kullback-Leibler távolságot használjuk fel a veszteség megállapítására. Minél jobban tudjuk közelíteni az ideális eloszlást, annál kisebb a veszteség (Ramos et al., 2018).

2.2.3. Illeszkedésvizsgálat

Egy modell illeszkedésének a vizsgálata már akkor sem egyértelműen meghatározott, ha csak single-label (például multiclass) problémával állunk szemben. Nincs olyan mérőszám, amely egyszerre minden oldalról jól meg tudja fogni a klasszifikáció sikerességének a mértékét, ezért célszerű többféle metrikát alkalmazni, nem elhanyagolva, hogy melyik mit is jelent pontosan. Multi-label problémák tekintetében ez a kérdés még egy aspektussal tovább „osztódik", ez pedig egyik maga a feladat neve: egy esethez többféle label tartozik, vagy esetleg egy sem tartozik hozzá – mi jelzi jól a sikeres besorolást? Ha minden labelt pontosan eltaláltunk? Vagy legyünk megengedőbbek, és ha egy részüket eltaláltuk, azt tekintsük félsikernek?

Wu és Zhou (2017) tanulmánya nem kevesebb, mint tizenegyféle illeszkedésvizsgálatra alkalmas mutatót mutat be multi-label modellekre. A célom az, hogy könnyen interpretálható, jól megfogható mutatószámokat alkalmazzak az eredmények összehasonlítására. Mivel multi-label feladatok körében egy viszonylag egyszerűbb esettel állunk szemben (három osztály, nincs hierarchia, a címkék megfelelő számú reprezentációja az adatbázisban), ezért a bonyolultabb, sokféle adatbázis-jellegzetességet (pl. labelek hierarchizáltsága, szélsőségesen egyenlőtlen label-eloszlás, *imbalance*) is számba vevő mutatókat nem használtam fel.

$$MR = \frac{1}{n}\sum_{i=1}^{n} I(Y_i = Z_i)$$

A pontos egyezés (*exact match ratio)* egy szigorú mérőszám, ami azt mutatja meg, hogy az egyes esetekhez tartozó label-setek mekkora részét sikerült pontosan eltalálni.

A következő mérőszámokat szokás *konfúziós mérőszámok*nak is nevezni, mivel a valós és becsült értékekből konstruálható konfúziós mátrixon alapul a logikájuk.

		Becslés	
		+	–
Valóság	+	TP (true positive)	FN (false negative)
	–	FP (false positive)	TN (true negative)

Az alábbi négy konfúziós mérőszám közül egyik sem a legjobb, mivel az, hogy melyiket tekintjük mérvadónak, mindig azon múlik, hogy mire is vagyunk kíváncsiak.

$$A = \frac{TN + TP}{TN + FN + TP + FP}$$

Az *accuracy* minden esetre veszi a jól eltalált labelek arányát a becsült és valós labelek uniójához képest, majd ezeket átlagolja. Talán ez a legintuitívabb mérőszám, egyszerűen azt mondja meg, hányszor becsült jól a modell.

$$P = \frac{TP}{TP + FP}$$

A *precision* hasonló az *accuracy*hez, de a jól eltalált labeleket csak a valós labelekhez viszonyítja.

$$R = \frac{TP}{TP + FN}$$

A *recall* (más néven *szenzitivitás* vagy *true positive rate*) ennek mintegy párja, a jól eltalált labeleket a becsült labelekhez viszonyítja.

$$IR = \frac{TN}{TN + FP}$$

$$FPR = 1 - IR = \frac{FP}{TN + FP}$$

Az *inverse recall* (más néven *specificitás* vagy *true negative rate*) a recall ellentéteként is értelmezhető – önmagában nem sokszor használatos, hiszen azt az esetet kezeli, amikor az esetek nem veszik fel az adott labelt.

$$F_1 = \frac{2 \cdot R \cdot P}{R + P}$$

Az *F_1-score* a *precision* és a *recall* harmonikus átlaga, ezáltal kissé nehezebben interpretálható, mint a fentebbi mérőszámok. Ettől függetlenül hasznos mutató, főleg, ha a labelek eloszlása nem egyenletes a mintán. Ha viszont a becslésben nagy a hamis pozitívok és a hamis negatívok aránya közti eltérés, kevésbé jó, mert amelyik sokkal nagyobb szám, „magához rántja" az *F_1-score*t. Ilyenkor érdemes külön nézni a *recall*t és a *precision*t.

Az *inverse recall* kissé kilóg a mérőszámok közül hasznosságát tekintve, viszont definiálása elengedhetetlen egy sokkal fontosabb mérőeszköz, a ROC-görbe használatához, amely vizuális segítséget nyújt a különböző modellek összehasonlításában. A ROC-görbe nem más, mint a *true positive rate* és a *false positive rate* összevető ábrázolása különböző thresholdok függvényében. Mivel mindkét mérőszám nullától egyig terjedhet, ezért az ideális pont a bal felső sarok, a (0, 1), ahol mindent pozitív eset igaz, ezt a pontot próbáljuk közelíteni a (0,0) és (1,1) pontok közti görbével. A modellek közt a görbe alatti terület alapján (*AUC, Area Under Curve*) válogathatunk – minél nagyobb ez a terület, annál közelebb vagyunk az ideális ponthoz, annál jobb a modell (Powers, 2007, Hanley & McNeil, 1982).

A fenti mérőszámok jól működnek, viszont egyelőre csak különálló label-setekre, vagy egy bináris feladatra. A multiclass és multi-label feladatokhoz hozzátartozik az, hogy ezeket a mérőszámokat kiterjesszük a többféle felvehető kategóriára, és ezeknek a kategóriáknak a különálló eredményeit összefogjuk. Erre többféle módszer is a rendelkezésünkre áll, a négy leggyakrabban használt a *micro averaging,* a *macro averaging,* a *sample averaging* és a *weighting.* Mindegyik konfúziós mérőszámnak létrehozhatjuk az aggregált változatát, és ahogy maguk a mérőszámok, úgy az aggregálás különböző módszerei sem egyértelműen

hasznosak vagy kevésbé hasznosak – ugyanannak a kérdésnek különböző megközelítéseit kínálják.

A macro-averaged mérőszámok a legintuitívabbak: a kategóriák (labelek) különálló metrikáinak számtani átlagát veszi. Ha *C* egy illeszkedést mutató mérőszám, akkor a macro-verziója:

$$C_{macro} = \frac{1}{|L|} \sum_{j=1}^{|L|} C_{\lambda_j}$$

Ezzel a módszerrel tulajdonképpen ugyanolyan súlyt adunk minden osztálynak, így nem vagyunk tekintettel a különböző labelek gyakoriságának a különbségeire. Erre megoldás a weighted-mérőszámok létrehozása, amikor a labelekre kapott mutatók értékét aggregáláskor besúlyozzuk az általuk jelzett osztály gyakoriságával:

$$C_{weighted} = \frac{1}{|L|} \sum_{j=1}^{|L|} \frac{|\lambda_j|}{|\lambda|} C_{\lambda_j}$$

Érdemes megjegyezni, hogy a súlyozással egyidejűleg változnak a mérőszámok tulajdonságai: az F_1-score már nem biztos, hogy a recall és a precision között lesz.

A macro-módszer egy módosításának is tekinthető a sample-averaging, ami nem a labelek halmazán átlagol, hanem a teszthalmazba került eseteken, ezért nem label-szinten számolja a metrikákat, hanem eset-szinten:

$$C_{samples} = \frac{1}{N} \sum_{i=1}^{N} C_i$$

A micro-averaged mérőszámok liberálisabban kezelik az osztályokat: ilyenkor a labelek becsült értékeit „egy kalap alá vesszük”, nem teszünk különbséget aközött, hogy melyik érték hova tartozik. Mivel ez nem a korábban labelenként megkapott mérőszámokra alapul, nem lehet könnyen formalizálni. Ha például a precisionnél érzékeltetjük, hogy a konfúziós mátrixot csak egy adott labelen értelmezünk: $P_{\lambda_j} = \frac{TP_{\lambda_j}}{TP_{\lambda_j} + FP_{\lambda_j}}$, akkor $P_{micro} = P_L = \frac{TP_L}{TP_L + FP_L}$. Ilyenformán például a micro-averaged F_1: $F_{1micro} = 2\frac{R_{micro}P_{micro}}{R_{micro}+P_{micro}}$

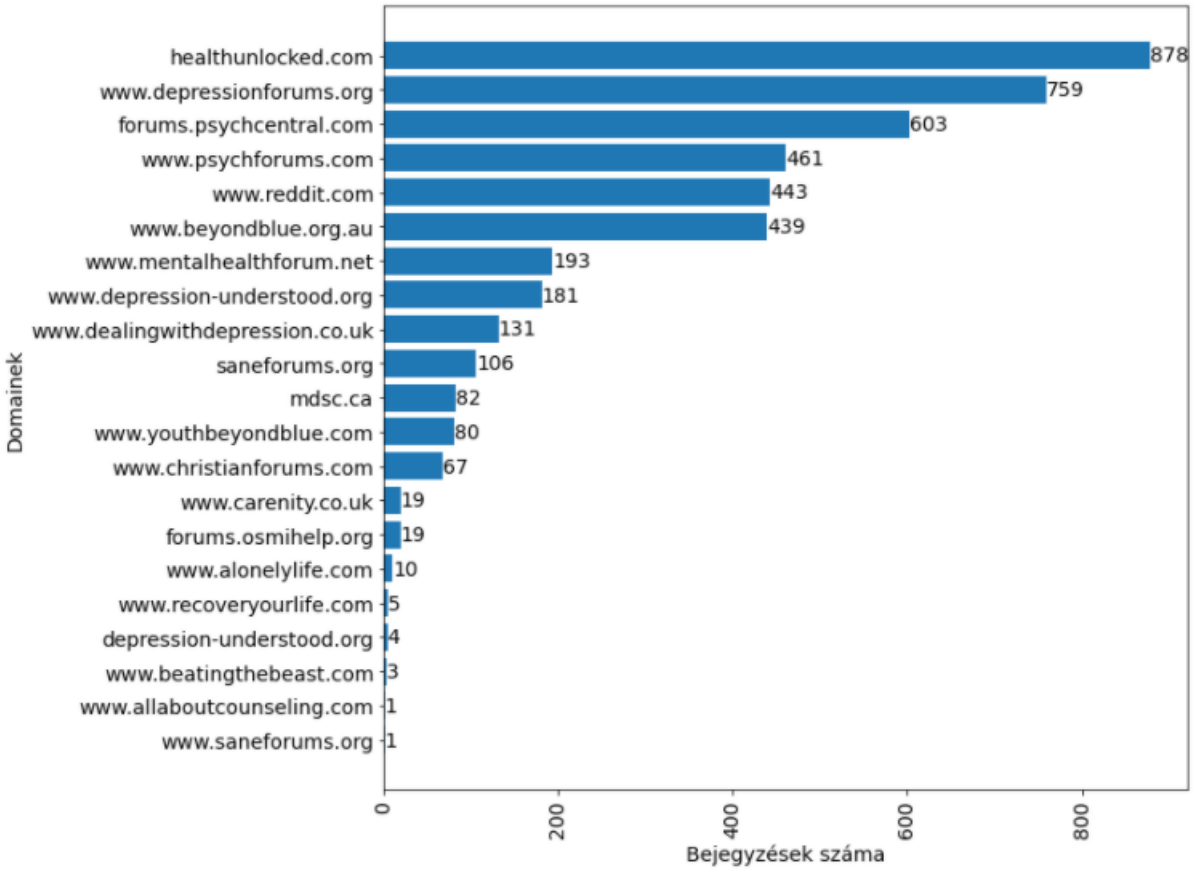
Azt mondhatjuk:

$$C_{micro} = f_C(\boldsymbol{\theta_L}),$$

ahol f_c az adott mérőszámot megadó függvény, $\boldsymbol{\theta_L}$ pedig a teljes label-halmazon számolt paraméterek értékei. (Sorower, 2010).

3. Adatbázis

A szakdolgozatomban depresszióval és szorongással kapcsolatos fórumbejegyzéseket és -hozzászólásokat elemzek, melyek mentális egészséggel foglalkozó népszerű internetes fórumokról lettek begyűjtve. A legtöbb bejegyzés forrása a *healthunlocked.com,* a *depressionforums.org*, a *forums.psychcentral.com*, a *psychforums.com*, a *reddit.com* és a *beyondblue.org.au* volt. Ezek regisztráció nélkül is olvasható, anonim felületek, ezért az adatgyűjtés az adatkezelési szabályoknak, a GDPR-nak megfelelt. A bejegyzések kétkörös filterezésen estek át: az első körben kiválasztásra kerültek azok a threadek, amiknek a címében vagy az indító bejegyzésében szerepelt a *depression* vagy a *depressed* szó, majd ez alól azok a bejegyzések kerültek be az adatbázisba, amelyek tartalmaztak legalább egyet egy meghatározott, depresszióval kapcsolatos szócsoportból[1]. A bejegyzések 2016 február 15. és 2019 február 15. között születtek. A megszűrt korpusz 79 889 bejegyzést tartalmazott, ebből 4 485 került annotálásra.

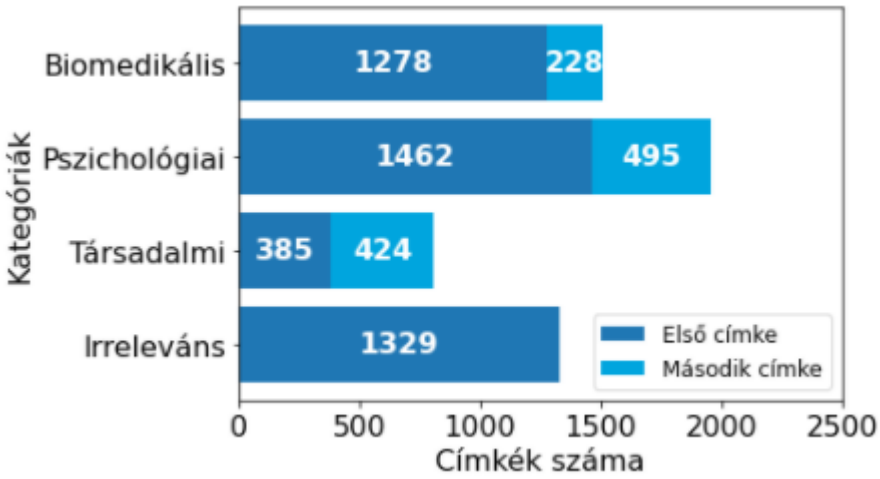


8. ábra: Annotált fórumbejegyzések száma oldalanként

[1] Az adatgyűjtés részletesebb leírása: http://ijqm.rc2s2.eu/

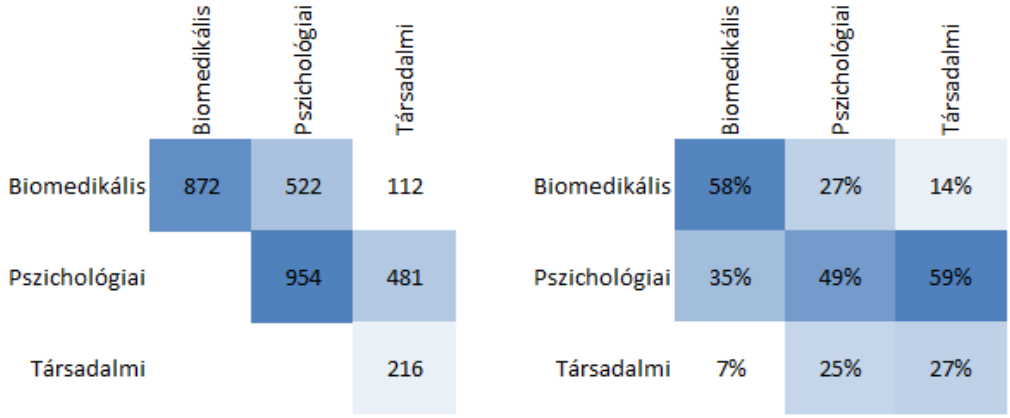
Az annotálást (mely a programozói zsargonban körülbelül címkézést jelent) társadalomtudományi területről érkező hallgatók végezték, mivel a feladat nagyobb rálátást igényelt annál, hogy nem szakértői annotálással megoldható legyen. Az annotátorok a címkézés megkezdése előtt tréninget kaptak, az ennek alapjául szolgáló anyag pedig a munka ideje alatt folyamatosan frissítésre került, hogy lehetőség szerint mindenki ugyanazokat a nézőpontokat alkalmazza egy adott bejegyzés címkézése közben. Minden besorolt bejegyzést két annotátor kódolt, akik egy elsődleges, és opcionálisan egy másodlagos címkét adtak. A végső címke részleges egyezés esetén többségi döntés alapján, teljes különbözőség esetén szakértői döntés alapján született (Németh et al., 2020).

A végső adatbázisban az alábbi labelek (depressziós keretezési módok) lehetségesek: *biomedikális, pszichológiai, társadalmi* és *irreleváns.* Érdemes észrevenni a társadalmi keretezés megjelenését: a másik két kategóriával ellentétben az elsődleges és másodlagos címkék aránya sokkal kiegyensúlyozottabb, vagyis ez az artikuláció vélhetően sokkal inkább kiegészítő jellegű, gyakorisága alapján jóval esetlegesebb.



9. ábra: Kategóriák megjelenése az annotált adatbázisban.

A dolgozat témája szempontjából nagyon fontos kérdés, hogy mit tudunk mondani a változók együttállásáról. Noha a multi-label feladatokban egy hierarchizáltság bevezetésével kezelhetővé tehető a szó szoros értelmében vett *elsődleges* és *másodlagos* jellege a címkéknek, én ezzel most nem éltem, ugyanis így a probléma matematikailag könnyebben kezelhetővé vált, a bemutatott machine learning módszerek pedig alapvető tulajdonságukban lettek bemutathatók.



10. ábra: A végleges címke-kombinációk darabszámai és arányai

A bal oldali táblázat a darabszámokat ábrázolja, a bal felső sarokból induló átlóban az egy címkével rendelkező bejegyzések darabszáma olvasható. A jobb oldali táblázat a labelek egymáshoz viszonyított arányát ábrázolják, az oszlopokban található a címke, ami*hez* viszonyítunk, a sorokban pedig ami*t*. Látható, hogy a biomedikális és a társadalmi keretezés az egymástól legtávolabb álló párosítás. A biomedikális keretezésű bejegyzések mindössze 7%-ában fordult elő szociológiai keretezés, a szociológiai keretezésűeknek pedig 14%-a tartalmazott biomedikális keretezést. A biomedikális és a pszichológiai keretezésű bejegyzések közel fele egy címkét kapott, a társadalmi keretezésű bejegyzések tekintetében viszont kevesebb, mint egyharmaduknál volt ennyire biztos a besorolás.

4485 eset összesen 3157 olyan címkét kapott, ami alapján az adott bejegyzés releváns és megjelenik benne legalább az egyik narratíva. A multi-label adathalmazokat a címkesűrűség alapján két mérőszámmal jellemezhetjük: a kardinalitással (*cardinality, Card*) és a sűrűséggel (*density, Dens*).

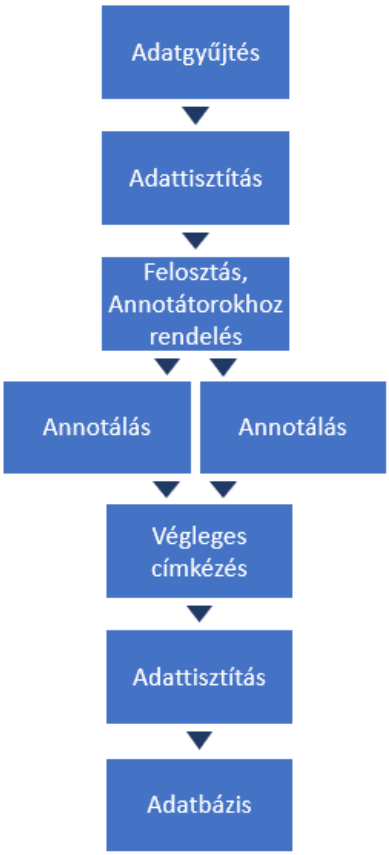
$$Card = \frac{1}{N}\sum_{i=1}^{N}|Y_i|$$

$$Dens = \frac{Card}{|L|} = \frac{1}{N}\sum_{i=1}^{N}\frac{|Y_i|}{|L|}$$

A kardinalitás az egy esetre jutó átlagos címkeszámot adja meg, a sűrűség pedig a label-halmaz számosságával normált címkeszámoknak az átlagát. Az általam használt adatbázisban a kardinalitás 0,96, a sűrűség pedig 0,32. A számok mutatják, hogy sok eset (ahogy a 9. ábrán is látszik, a bejegyzések majdnem harmada) címke nélküli. Azonban ez is lényeges információ

lehet gépi tanulás esetén, hiszen fontos, hogy az algoritmus ne próbáljon mindenképp minden, korábban nem látott bejegyzést „belepréselni” egy kategóriába, ha annak nincs megalapozottsága. Ha csak azokat az eseteket vesszük, ahol legalább egy érvényes label megtalálható, akkor a kardinalitás 1,34, a sűrűség pedig 0,45 – látszik, hogy az adatbázis multi-label jellege nem annyira domináns, ami érthető, hiszen az annotálás során a második címke opcionálisan adható volt, olyan eset pedig nincs, amihez három címke is tartozna. A kardinalitás és a sűrűség nemcsak puszta leíró statisztika, előzetes információt adhat a multi-label tanulás várható sikerességéről. Minél magasabb a kardinalitás és hozzá képest alacsonyabb a sűrűség, annál nehezebb dolga van egy algoritmusnak. Emellett a két mutatóhoz köthetünk illeszkedésvizsgálati mérőszámok által mutatott érzékenységet is: az accuracy jobban, az F_1-score és a ROC kevésbé reagál ezeknek a változására (Bernardini et al., 2014).

Az elemzéshez már egy előtisztított, címkézett, mondhatni „konyhakész” adatbázis állt rendelkezésemre[2].



11. ábra: Az adatbázis létrejötte

[2] Az adattisztításról bővebben: http://ijqm.rc2s2.eu/

4. Adatelemzés és eredmények

A szakdolgozatom célja az volt, hogy áttekintést adjak a multi-label klasszifikáció problémájáról, valamint a lehetséges megoldásokról. Ebben a fejezetben a korábban bemutatott megoldási technikákat implementálom és ismertetem az eredményeket. Nem volt célom state-of-the-art modellek létrehozása, a valódi információ ugyanis a különböző megoldási módok összehasonlításából fakad, hiszen mindegyiket ugyanarra a problémára, ugyanazon az adatbázison alkalmazom.

Az adatok előkészítését és a modellek illesztését Pythonban végeztem[3] a Scikit-multilearn library (Szymański & Kajdanowicz, 2019) segítségével, ami a jól ismert Scikit-learn (Pedregosa et al., 2011) egy részleges kiterjesztése multi-label problémák kezelésére. A két könyvtár között kompatibilitás van, az illeszkedésvizsgálat általam kiválasztott mérőszámait multi-label megoldásokra is képes kezelni a Scikit-learn.

4.1. Klasszifikációs algoritmusok

A multi-label klasszifikáció problématranszformációs megközelítése önmagában nem ad megoldást, csak egy módot arra, hogy már ismert klasszifikációs algoritmusoknak hogyan tudjuk „odaadni" a multi-label problémát, hogy azok képesek legyenek azt megoldani. Három algoritmust választottam ki: a *Naïve Bayes*t, a *Support Vector Machine*t és a *Logisztikus Regresszó*t. Fontosnak tartottam, hogy ne csak egy modellt próbáljak ki, hiszen a probléma transzformálásának pont ez az egyik előnye, hogy a segítségével bármilyen, single-label klasszifikációra alkalmas algoritmussal folytathatjuk a feladat megoldását.

4.1.1. Naïve Bayes

A Naïve Bayes egy nagyon intuitív és egyszerű klasszifikációs modell algoritmus, bár a neve kissé megtévesztő, mivel nem (vagy nem feltétlenül) egy bayesiánus statisztikai módszer, hanem a Bayes-tétel alkalmazása (Hand & Yu, 2001):

[3]A kód elérhető: https://gist.github.com/nk94-th/917fb7c36b53b38d1ec213cecf76318b

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Egyik alapvetése, hogy a független változók között függetlenséget feltételez. Előnye, hogy nagy adathalmazon is gyorsan lefut, ellenben viszonylag kis tanítóhalmazzal is elboldogul. Ha bemenetünk feature-vektorok összessége $(\boldsymbol{x_1}, \boldsymbol{x_2}, \dots, \boldsymbol{x_n})$, a kimenetünk pedig $\boldsymbol{y}$, akkor azt mondhatjuk, hogy

$$P(\boldsymbol{y}|x_1, \dots, x_n) = \frac{P(x_1, \dots, x_n|y)P(y)}{P(x_1, \dots x_n)}.$$

Ezután „naivan" feltételezzük (erre utal a név), hogy a független változóink között nincs összefüggés, ami miatt

$$P(y|x_1, \dots x_n) = \frac{P(y)\prod_{i=1}^{n} P(x_i|y)}{P(x_1, \dots, x_n)}.$$

Mivel a nevezőben szereplő kifejezés adott problémára nézve konstans (hiszen az input adatokon nem tudunk változtatni), a becslést megadhatjuk úgy, mint

$$\hat{y} = \arg\max_y P(y) \prod_{i=1}^{n} P(x_i|y).$$

Naïve Bayes klasszifikáció megvalósításakor az adatunknak megfelelően feltételeznünk kell egy eloszlást, és aszerint dolgozni. Szövegklasszifikációs modellek tekintetében szokásosan a Bernoulli eloszlás a használatos (McCallum & Nigam, 1998) a multinomiális mellett, ilyenkor a független változóink bináris jellegűek. Ekkor annak a valószínűsége, hogy az x esetünk a C_k osztályba tartozik:

$$p(x|C_k) = \prod_{i=1}^{n} p_{ki}^{x_i}(1 - p_{ki})^{(1-x_i)}.$$

A modellnek van egy kiküszöbölhető, de fontos gyengesége, ami abból fakad, hogy valószínűségek produktumát veszi. Ha olyan kategóriával találkozik a teszthalmazban, ami nem volt benne a tanítóhalmazban, akkor ahhoz automatikusan nulla valószínűséget rendel, ami a szorzások miatt az egész egyenletre kihat. Ezen egy hiperparaméter, α segítségével tudunk segíteni, amivel igazából azt erősítjük meg, hogy az adathalmazunk véges, és nulla megfigyelt valószínűség esetén sem mondhatjuk azt, hogy az adott kategória tényleges valószínűsége is valóban nulla.

4.1.2. Support Vector Machine

A Support Vector Machine egy robusztus nemvalószínűségi lineáris klasszifikáló algoritmus. A modell az adathalmazt egy p-dimenziós térben kezeli, a feladat pedig egy $(p-1)$-dimenziós hipersíkot találni, mely képes az adatot két részre elkülöníteni (klasszifikálni). A feltételezés az, hogy ha sok ilyen hipersík is van, akkor azt válasszuk, amely a legszélesebb margóval tudja elválasztani az eseteket – magyarán, amelyik a legjobban működik, hiszen minél szélesebb a margó, annál kevésbé valószínű, hogy egy új, ismeretlen eset rossz osztályba sorolódik be. Ezt nevezzük maximális margójú klasszifikációnak. A hipersík egyenlete a következőképpen írható fel:

$$y = w_0 + w_1x_1 + w_2x_2 + \cdots + w_px_p \; = \\ w_0 + \boldsymbol{w}^T\boldsymbol{X},$$

ahol $\boldsymbol{w}$ a hipersík normálvektora, w_0 a torzítás. A hípersíkhoz legközelebb eső elemek a tartópontok, azok határozzák meg a margó szélességét, ezáltal azok „tartják" a döntési határt (innen ered a módszer neve). Mivel szeretnénk minél szélesebb margót tartani, így a hipersík tartóvektorainak a normáját alacsonyan kell tartani.

Természetesen az adatok általában nem tökéletesen szeparábilisak, ezért a margó maximalizálása mellett kezelni kell a rosszul besorolást is. Erre bevezethetünk egy ξmérőszámot. Mivel a hipersík két részre (egy pozitív és egy negatív altérre) bontja a p dimenziós teret, ezért tudjuk, hogy

$$w_0 + \boldsymbol{w}^T\boldsymbol{X} \; = \; \pm\mathbf{1},$$

attól függően, hogy a tér melyik felén vagyunk. Így

$$y_i(w_0 + \boldsymbol{w}^T\boldsymbol{x_i}) \geq \mathbf{1}$$

jelöli a helyes besorolást, hiszen ha az egyenlet értéke kisebb egynél, akkor az y_i elem „belelóg" a margóba, ha pedig nullánál is kisebb, akkor rossz oldalra sorolta a klasszifikáció. Használjuk a ξ-t a következőképpen:

$$y_i(w_0 + \boldsymbol{w}^T\boldsymbol{x_i}) \geq \mathbf{1} - \boldsymbol{\xi_i}$$

Ha $\xi_i > 0$, akkor hibás a klasszifikáció, és hogy mennyire hibás, azt ξ_i értéke árulja el. A minimalizálási feladat tehát a következő:

$$\min_{\boldsymbol{w},w_0} \frac{1}{2}\left|\left|w\right|\right|^2 + \sum_{i=1}^{n} \xi_i,$$

hiszen a support vektorok hosszának minimalizálásával a margót szélesítve növeljük a besorolás biztosságát, míg a hibát alacsonyan tartva a helyes besorolási arányt tartjuk magasan.

Ami az SVM-et igazán előnyössé teszi, mint klasszifikációs módszer, az a kernelizálhatósága: a kernel-trükk alkalmazása ugyanis képessé teszi lineáris klasszifikációval nem megoldható feladatok elvégzésére is. A kernel függvények az adatok skaláris szorzatait használják fel arra, hogy a lineárisan nem megoldható problémát egy megfelelően magas dimenziójú jellemzőtérbe helyezzék át, ahol már lineárisan megoldhatóvá válik.

A szövegklasszifikálási feladatokhoz érdemes lineáris kernelt választani. Egyrészt azért, mert a szöveges adathalmazokra általában jellemző, hogy lineárisan szeparálhatók (Joachimns, 1998), másrészt pedig mert a lineáris kernel jól működik, ha sok feature van az adatban, ilyenkor ugyanis egy nemlineáris (például RBF) kernel nem változtat sokat a döntési határon, egy sokkal magasabb dimenzióba való áttranszformálás nem nyújt lényeges mértékű segítséget (Hsu et al., 2003), így elhagyásával számolási kapacitást spórolhatunk meg.

Az SVM illesztése egy jellemzően idő- és memóriaigényes folyamat, ezért nagy adathalmazokon az algoritmus lefutása lassú.

4.1.3. Logisztikus regresszió

A logisztikus regresszió egy széles körben használt bináris klasszifikációs eljárás, ami annak a feltételezésére alapul, hogy az *x* független változók, valamint az *y* bináris független változó bekövetkezésének az esélyének a logaritmusa között lineáris kapcsolat van, ami a következőképpen formalizálható:

$$\ell = \log \frac{p}{1-p} = \beta_0 + \beta_i x_i,$$

ahol p annak a valószínűségét jelöli, hogy $y = 1$. A logisztikus regresszió tanítása során a cél az, hogy megtaláljuk a legoptimálisabb β paramétereket, amik minimalizálják a veszteséget. A veszteségfüggvény szélsőértékének (globális minimumának) a keresését a modell illesztésekor gradiensmódszerrel végeztem, „saga” solverrel (Defazio et al., 2014), mely nagyméretű adat kezelésekor ideálisabb választás, mert mintát vesz a korábbi gradiensekből, azzal dolgozik tovább, így memóriatakarékos.

4.2. Eredmények

		Accurcay	Precision	Recall	F_1-score	AUC
Bináris Relevancia						
Naïve Bayes	Biomedikális	0,84	0,81	0,68	0,74	0,90
	Pszichológiai	0,68	0,66	0,53	0,59	0,74
	Szociológiai	0,75	0,34	0,39	0,36	0,72
	Micro-average	0,76	0,63	0,56	0,59	0,77
	Macro-average	0,75	0,60	0,53	0,56	0,79
	Weighted average	0,75	0,65	0,56	0,60	0,79
	Sample average	0,76	0,83	0,69	0,58	0,79
Log. regresszió	Biomedikális	0,82	0,80	0,64	0,71	0,88
	Pszichológiai	0,70	0,68	0,59	0,63	0,75
	Szociológiai	0,82	0,52	0,28	0,37	0,71
	Micro-average	0,78	0,70	0,55	0,62	0,81
	Macro-average	0,78	0,66	0,51	0,57	0,78
	Weighted average	0,77	0,69	0,55	0,61	0,79
	Sample average	0,78	0,84	0,69	0,62	0,78
SVM	Biomedikális	0,80	0,71	0,67	0,69	0,83
	Pszichológiai	0,64	0,59	0,58	0,59	0,67
	Szociológiai	0,78	0,39	0,36	0,37	0,64
	Micro-average	0,74	0,59	0,57	0,58	0,77
	Macro-average	0,74	0,56	0,54	0,55	0,71
	Weighted average	0,72	0,60	0,57	0,58	0,72
	Sample average	0,74	0,74	0,7	0,56	0,71
Klasszifikáló láncok						
Naïve Bayes	Biomedikális	0,84	0,81	0,68	0,74	0,90
	Pszichológiai	0,68	0,66	0,54	059	0,74
	Szociológiai	0,75	0,33	0,38	0,36	0,73
	Micro-average	0,76	0,63	0,56	0,59	0,77
	Macro-average	0,75	0,60	0,53	0,56	0,79
	Weighted average	0,75	0,65	0,56	0,60	0,79
	Sample average	0,76	0,83	0,69	0,59	0,79
Log. regresszió	Biomedikális	0,82	0,78	0,65	0,74	0,87
	Pszichológiai	0,70	0,67	0,62	0,64	0,74
	Szociológiai	0,80	0,42	0,31	0,36	0,70
	Micro-average	0,77	0,67	0,57	0,62	0,80
	Macro-average	0,77	0,62	0,53	0,57	0,77
	Weighted average	0,76	0,66	0,57	0,61	0,78
	Sample average	0,77	0,80	0,71	0,61	0,77
SVM	Biomedikális	0,82	0,78	0,65	0,71	0,87
	Pszichológiai	0,70	0,67	0,62	0,64	0,74
	Szociológiai	0,80	0,42	0,31	0,36	0,70
	Micro-average	0,77	0,67	0,57	0,62	0,80
	Macro-average	0,77	0,62	0,53	0,57	0,77
	Weighted average	0,76	0,66	0,57	0,61	0,78
	Sample average	0,77	0,80	0,71	0,61	0,77
Label Powerset						
Naïve Bayes	Biomedikális	0,83	0,83	0,61	0,70	0,89
	Pszichológiai	0,66	0,66	0,50	0,56	0,75
	Szociológiai	0,79	0,38	0,26	0,31	0,73
	Micro-average	0,76	0,67	0,49	0,57	0,79
	Macro-average	0,76	0,62	0,46	0,53	0,79
	Weighted average	0,75	0,67	0,49	0,57	0,80
	Sample average	0,78	0,85	0,65	0,57	0,79
Log. regresszió	Biomedikális	0,83	0,79	0,66	0,72	0,87
	Pszichológiai	0,70	0,72	0,53	0,61	0,77
	Szociológiai	0,82	0,52	0,21	0,30	0,74
	Micro-average	0,78	0,73	0,51	0,60	0,82
	Macro-average	0,78	0,68	0,47	0,54	0,79
	Weighted average	0,77	0,71	0,51	0,59	0,80
	Sample average	0,76	0,87	0,66	0,61	0,79
SVM	Biomedikális	0,81	0,73	0,69	0,71	0,87
	Pszichológiai	0,66	0,62	0,59	0,61	0,75
	Szociológiai	0,80	0,42	0,34	0,38	0,76
	Micro-average	0,76	0,63	0,58	0,60	0,81
	Macro-average	0,76	0,59	0,54	0,56	0,80
	Weighted average	0,74	0,62	0,58	0,60	0,80
	Sample average	0,76	0,77	0,71	0,59	0,80

A táblázat folytatódik a következő oldalon.

		Accuracy	Precision	Recall	F_1-score	AUC
RAKEL						
Naïve Bayes	Biomedikális	0,83	0,83	0,61	0,70	0,89
	Pszichológiai	0,66	0,66	0,50	0,56	0,75
	Szociológiai	0,80	0,38	0,26	0,31	0,73
	Micro-average	0,76	0,67	0,49	0,57	0,79
	Macro-average	0,76	0,62	0,46	0,53	0,79
	Weighted average	0,74	0,67	0,49	0,57	0,80
	Sample average	0,76	0,85	0,65	0,57	0,79
Log. regresszió	Biomedikális	0,83	0,77	0,69	0,73	0,87
	Pszichológiai	0,69	0,67	0,60	0,63	0,76
	Szociológiai	0,79	0,52	0,32	0,40	0,70
	Micro-average	0,78	0,69	0,58	0,63	0,81
	Macro-average	0,78	0,65	0,54	0,59	0,78
	Weighted average	0,77	0,68	0,58	0,62	0,79
	Sample average	0,78	0,82	0,71	0,62	0,78
SVM	Biomedikális	0,80	0,72	0,69	0,70	0,87
	Pszichológiai	0,67	0,63	0,59	0,61	0,76
	Szociológiai	0,82	0,44	0,35	0,39	0,76
	Micro-average	0,76	0,63	0,58	0,60	0,82
	Macro-average	0,76	0,60	0,54	0,57	0,80
	Weighted average	0,74	0,63	0,58	0,60	0,80
	Sample average	0,76	0,77	0,71	0,59	0,80
Neurális háló						
	Biomedikális	0,80	0,74	0,65	0,69	0,87
	Pszichológiai	0,66	0,62	0,61	0,61	0,74
	Szociológiai	0,81	0,46	0,23	0,31	0,73
	Micro-average	0,76	0,64	0,55	0,59	0,81
	Macro-average	0,76	0,60	0,50	0,54	0,78
	Weighted average	0,74	0,63	0,55	0,58	0,78
	Sample average	0,76	0,77	0,69	0,58	0,78

12. ábra: Eredmények

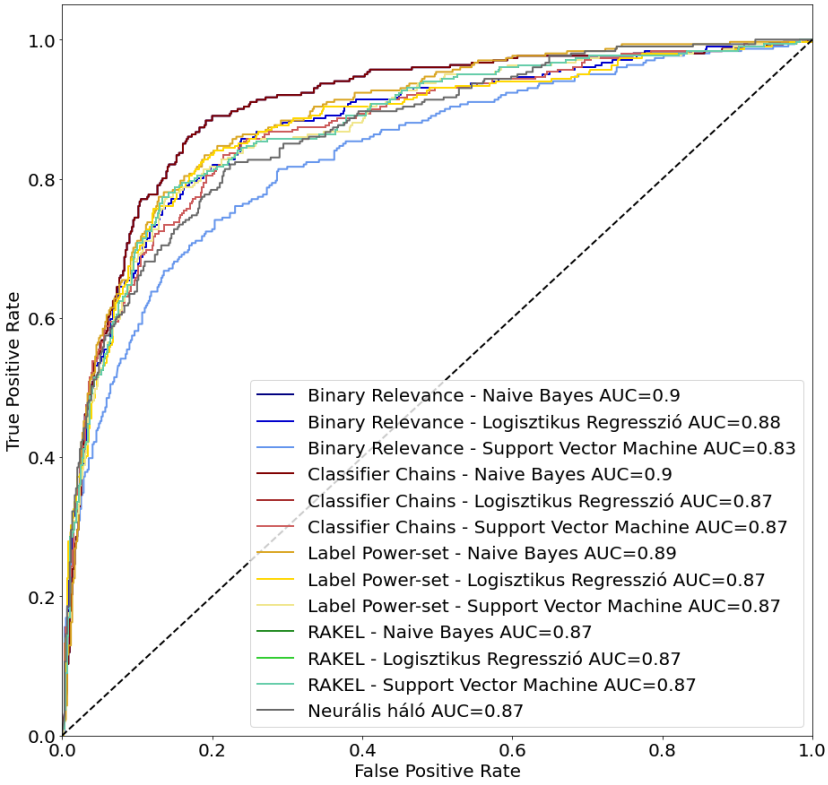
Az eredmények között elsőként egy módszer, a multi-label k-neares neighbors hiányára kell magyarázatot adnom. Az algoritmus sajnos a többszöri iterációs paraméterkeresés után sem adott értékelhető outputot a címkékre, aminek a magyarázata többrétű lehet. Egyrészt a módszer jellemző tulajdonsága, hogy a leggyakoribb címke hajlamos elhúzni az eredményeket (Tan, 2005). Érdemes emlékezni, hogy a közel négyezer-ötszáz eset a három felvehető címkét bináris kódolással hordozta (vagy épp nem hordozta) a modellekben, és még a leggyakoribb (szám szerint legtöbbször előforduló) pszichológiai narratívát jelölő címke is csak közel kétezerszer fordult elő, ezerháromszáz esetnél pedig egyik címke sem volt releváns. Így tulajdonképpen címkénként nézve, és összességében is az esetek inkább gyakrabban *nem vettek fel* címkét, mint igen. Ez klasszifikációnál a 0 kimenetet egyértelműbbé teszi, hiszen az a leggyakoribb. Ez a többi algoritmusnál is hasonlóan hathat a becslési folyamatra, a kNN által használt többségi szavazás (*majority voting*) gyaníthatóan ráerősített erre a jellegre (Syaliman et al., 2018). Másrészt, talán még jobban köthető ehhez a specifikus problémához, hogy a kNN algoritmusa érzékeny a magas dimenziójú, kis címkesűrűségű adatokra. Ahogy a független változók száma, és így a dimenzionalitás is nő, úgy tágul folyamatosan az adattér, amely folyamatot ideális esetben az adatnak követnie kellene – de nem követi. Az általam használt

adatbázisban a négyezer-ötszáz esetre közel huszonnégyezer független változó jutott (ennyi szóból állt a bejegyzések összessége által létrehozott tisztított szótár). Az algoritmus számára az lenne ideális, ha az adattérben a függőváltozók által létrehozott újabb és újabb tengelyek megjelenésekor ezekhez a tengelyekhez az adat egy megfelelő mértékű reprezentációja társulna. Viszont minél több dimenziós a tér, minél több tengely van, annál nehezebb két pontnak megfelelően közel lennie egymáshoz megfelelően sok tengely mentén – a szakirodalom ezt nevezi a „dimenzionalitás átka”-ként (Grus, 2015).

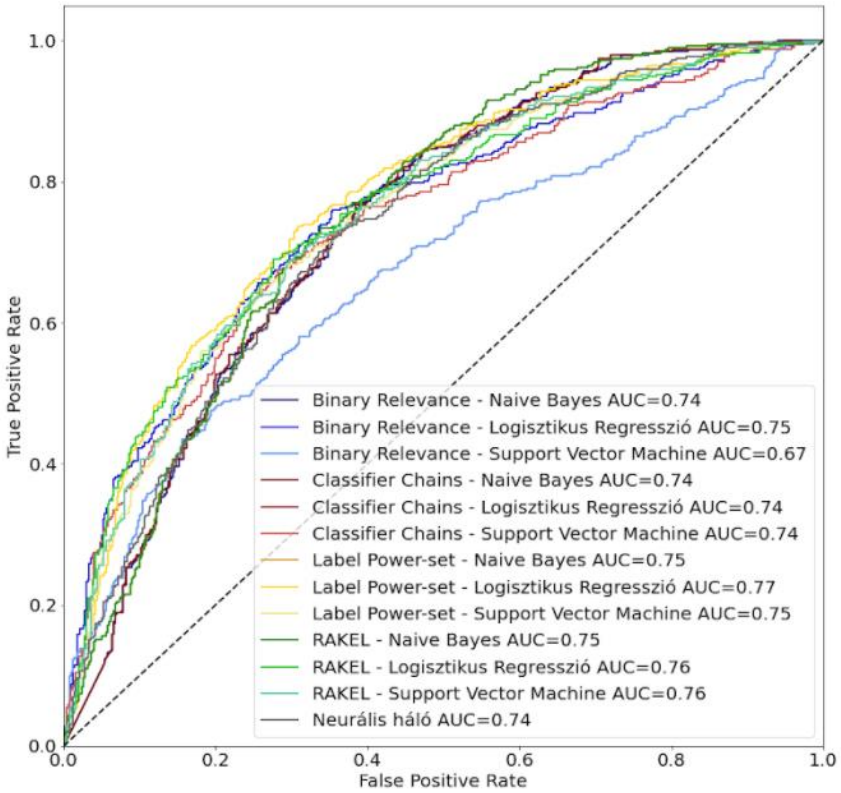
A fentebb említettek, valamint a táblázatok alapján elsőként szembetűnő kép, miszerint a modellek közel hasonló eredményt hoztak, ugyanabba az irányba mutat: az adatokon, azoknak előkészítésén, tisztításán, a változószelekción legalább annyira múlik (ha nem jobban) az illeszkedés jósága, mint a helyes modell és a helyes paraméterek megtalálásán. Ennek az elemzésnek a folyamán nem tértem ki részletesen erre, mivel a modelleket a tisztított, de más egyéb statisztikai módszerekkel (például korrelációs vizsgálatok, főkomponens elemzés, feltáró faktoranalízis, klaszterelemzés) általam nem szelektált adatokon akartam illeszteni. Ezeknek az alkalmazásai, és (lehetségesen pozitív) hatásai a klasszifikációs folyamatra alighanem külön témaként is relevánsak lennének.

Szintén fontosnak tartom kiemelni, hogy a használt modellek nagy része (mint például a support vector machine, a neurális hálók, maga a multi-label k-nearest neighbors is) nagy számítási kapacitást kívánó, memóriaigényes algoritmus, így a lehetőségeim viszonylag korlátozottak voltak többek közt a paraméterkeresésben, illetve azok keresztvalidálása során. Nem kizárt, hogy komolyabb technikai háttérrel nekiindulva tovább lehetne finomítani a modelleken, és javítani azok tulajdonságain.

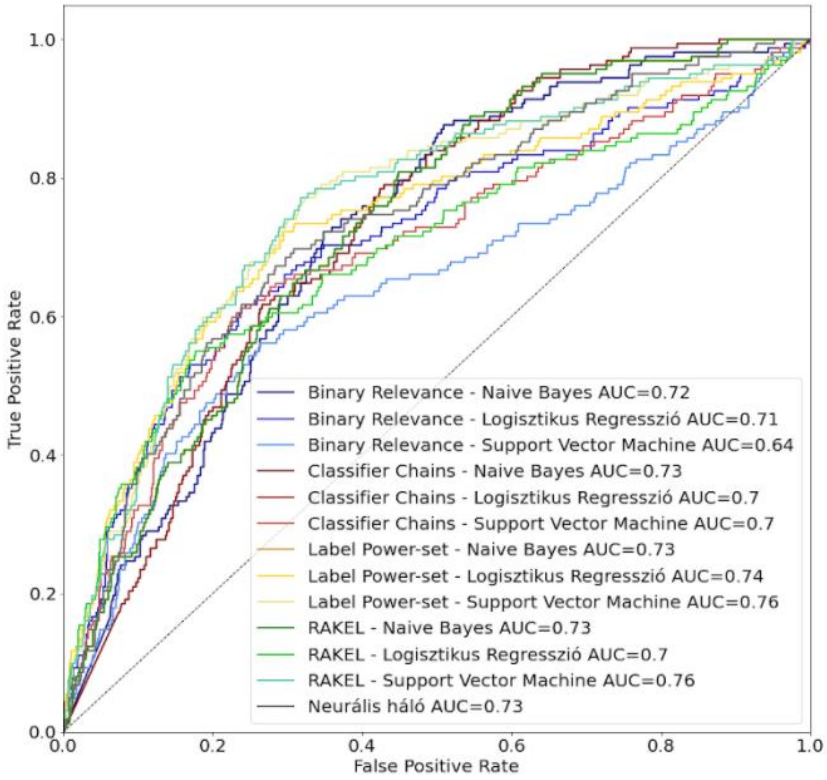
A modellek, bár az illeszkedésvizsgálati mutatókra nézve egész hasonló eredményeket hoztak, egy sűrű táblázat böngészése nem célszerű ennek a vizsgálatára. A vizuális összehasonlításra a különböző modellek által generált labelenkénti ROC-görbéket használtam, hiszen végső soron, bár az aggregált mutatók is fontos információkat tartalmaznak, mégiscsak az eredeti címkék helyes becslése a cél. A táblázatban négyféleképpen létrehozott aggregált mutatószámot mutattam be, ezek között helyenként nagy eltérés figyelhető meg. Ez újból megerősíti az általam korábban leírtakat, miszerint modellek illeszkedésvizsgálatakor (és ez különösen fontos a klasszifikációnál) nem hagyatkozhatunk egy- vagy kétféle mutatószámra, hiszen azzal gyakorlatilag a saját látószögünket szűkítjük be.



13. ábra: A biomedikális címke ROC-görbéi



14. ábra: A pszichológiai címke ROC-görbéi



15. ábra: A társadalmi címke ROC-görbéi

A ROC-görbék tanulmányozásakor rögtön látszik, hogy a modelleknek a biomedikális címkével volt a legkönnyebb dolga. Ezt magyarázhatja az orvosi narratívához gyakran kötődő szakszóhasználat, a gyógyszerek, antidepresszánsok nevei, azok szedésével kapcsolatos kifejezések, mennyiségek, mértékegységek, a klinikai állapotokat leíró terminusok. Ezzel kontrasztban a pszichológiai és társadalmi címkék már nehezebb feladatnak bizonyultak a modellek számára, amit többek közt az is okozhatott, hogy ehhez a kettőhöz kevésbé kötődtek a mindennapi szóhasználatból megkülönböztető jelleggel kiugró kifejezések.

Szintén érdemes megemlíteni ennek a vonatkozásában az eredmények szórását. A görbéket nézve jól látható, hogy a biomedikális és a pszichológiai címkék esetében a modellek közt többnyire „erős egyetértés volt”, a társadalmi címkénél viszont sokkal hangsúlyosabb a modellek közti különbség. Már a szakirodalmi összefoglaló társadalmi keretezésre vonatkozó fejezeténél is kiemeltem, hogy ez a fajta narratíva általában sokkal kevésbé nyilvánvalóan, sokkal látensebb módon jelenik meg a diskurzusban, ez összekapcsolható a fent látható teljesítménykülönbséggel: ha egy címkét könnyű, vagy kevésbé nehéz megjósolni (mint a biomedikálisat vagy a pszichológiait), akkor arra a modellek hasonló módon lehetnek képesek, ellenben ha nehéz a feladat, akkor nagyobb különbségek is kiütközhetnek.

Érdemes ezt a különbséget a modellek komplexitásának a szempontjából is vizsgálni. Látható, hogy a biomedikális címkék esetén az egyszerűbb modellek (bináris relevancia és klasszifikációs láncok) jól tudtak teljesíteni, de a társadalmi keretezés esetében a címkék közti korrelációt jobban figyelembe vevő modellek (label power-set és RAKEL) teljesítménye volt a jobb. Rögtön párhuzamot lehet vonni a 9. ábrával: a biomedikális címkével rendelkező bejegyzések majdnem kétharmada egyértelműen ebbe a narratívába tartozott, míg a társadalmi keretezésű bejegyzések kevesebb, mint harmadáról volt elmondható ugyanez. Nem csoda tehát, ha a társadalmi narratíva jobban megfoghatóvá vált a másik két címke előfordulását is számításba véve. Kiemelhető a neurális háló teljesítményének különbsége a különböző címkék esetén: a biomedikális és a pszichológiai kategóriákhoz képest a társadalmiban jobb eredményt tudott elérni, ez is utalhat arra, hogy ennek a narratívának a vizsgálatakor mélyebb struktúrákra van szükség.

A multi-label k-nearest neighbors módszernél említett jelenség, miszerint a 0 kimenet evidensebb lehet egy modell számára, szimplán a túlsúlya miatt, a recallal és a precisionnel kapcsolatban köszön vissza a többi modellnél: a recall mindig alacsonyabb, mint a precision, ami alapján, ha megnézzük a mutatószámok képletét, azt mondhatjuk, hogy a modellek inkább gazdaságosak, mint érzékenyek, hajlamosabbak egy létező címke esetén negatív irányba dönteni (így növelve a fals negatívok számát), mint fordítva, nem létező címke esetén azt relevánsnak jósolni.

Az eredmények természetesen szinte maguk vetik fel a további lehetőségeket a témával kapcsolatban. A már tárgyalt változószelekció és megfelelőbb technikai háttér mellett érdemes lehet levonni a tanulságot a fentiekből, és a továbbiakban felhasználni: a különböző címkékre különbözőképpen működő klasszifikációs algoritmusok a problématranszformálás módszerével egybevetve egyedi megoldásokra is lehetőséget adhatnak: egyedileg, külön címkékhez rendelt algoritmusokkal is dolgozhatunk. Emellett a klasszifikációhoz kapcsolódóan az optimális threshold megkeresése is érdekes feladat lehet – ismét csak, ezt akár címkékhez külön is meghatározhatjuk. Ezek a lehetőségek összekombinálva rengeteg utat nyithatnak meg a további kutatásokhoz.

5. Konklúzió

Szakdolgozatomban a machine learning témakörében kissé negligáltnak tekintett multi-label klasszifikációval foglalkoztam, illetve annak gyakorlati alkalmazásával egy multi-label problémát magában hordozó, depressziók különböző (biomedikális, pszichológiai, társadalmi) narratíváit vizsgáló, online fórumbejegyzésekből álló adatbázison. A bevezetésben igyekeztem kiemelni a szenzitív témákkal kapcsolatos online tartalmak egyre bővülő vizsgálatának pozitív hozadékait, amelyek kárpótolhatják a kutatókat a hagyományos adatgyűjtésekhez képest jóval szegényesebb metainformációkért. A szakirodalmi áttekintést két fő részre különítettem el: először megvizsgáltam a depressziónak, mint témának a különböző tudományos megközelítéseit. A szociológiai nézőpont mellett kiemeltem a téma társadalmi konceptualizációjának fontosságát. A második részben a multi-label klasszifikáció főbb módszereit, azoknak a matematikai és statisztikai alapjait tekintettem át, a szakirodalmak által használt felosztást követve: a probléma transzformálásának, illetve az adaptált algoritmusoknak a csoportját. Végül bemutattam a problémához alkalmazható illeszkedésmutatókat, valamint azoknak a fontosabb tulajdonságait. A harmadik fejezetben az adatbázist mutattam be, röviden tárgyalva annak létrejöttét, valamint bemutatva a multi-label probléma jellegét, illetve a területen használatos leíró statisztikák értékeit, azoknak esetleges előrejelző tulajdonságait. A negyedik fejezetben bemutattam az multi-label módszerek által alkalmazott klasszifikáló algoritmusokat, azoknak matematikai alapjait, végül pedig az illesztett modellek által adott eredményeket, a korábban ismertetett mérőszámok segítségével. Ezután igyekeztem minél mélyebben ismertetni az eredményeket, az azokban megfigyelhető tendenciákat és különbségeket a korábban említettekhez kapcsolni.

Szakdolgozatommal egy már lezajlott kutatást (Németh et al., 2020) követtem, annak az adatbázisához igyekeztem egy újfajta módon közelíteni. Célom az volt, hogy a multi-label klasszifikációs problémák megoldási lehetőségeit összehasonlító szándékkal bemutassam, azoknak működését egy valós adatbázison szemléltessem. Egy szövegklasszifikációs feladat komplex folyamatából a különböző modelltípusok összevetését emeltem ki. Ennek megfelelően az eredmények összehasonlításakor az említett modellek működési jellegzetességeire támaszkodtam.

Az eredmények alapján megállapítható, hogy a biomedikális narratíva messze a legkönnyebben felismerhető a tanuló algoritmusok számára, míg a pszichológiai, és főleg a

társadalmi keretezés felderítése nehezebb feladat. Utóbbinál különösen igaz ez, hiszen a társadalmi keretezés az adatok szintjén is ritkán állt önmagában – így jobb megoldást kínálnak a címkék közti összefüggést figyelembe vevő algoritmusok. Az adatbázisra alacsony relatív címkesűrűség jellemző, amit a mérőszámok viselkedése is tükrözött: a modellek hajlamosabbak voltak a nullás kimenet prediktálására. Az alacsony címkesűrűség komolyan érintette az egyik adaptált algoritmus, a multi-label k-nearest neighbors alkalmazásának lehetőségeit.

Az eredmények az adatok által megalapozott elvárásokat jellemzően alátámasztják, de legalább ugyanannyi kérdést és lehetőséget is felvetnek. A további vizsgálódás útjai két nagyobb csoportra oszthatók: az első az adatoknak a különböző statisztikai módszerekkel való további tisztítása, a célzott változószelekció, mely a dimenziócsökkentés révén jobban teljesítő és robusztusabb modelleket eredményezhet; a második pedig az általam ismertetett eredmények felhasználásával egy akár bonyolultabb, de az adatbázis jellegzetességeihez jobban illeszkedő modellstruktúra kidolgozása, melybe az algoritmusok címkénként eltérő viselkedése, mint információ, beépíthető.

Véleményem szerint a szakdolgozatban bemutatott eredmények igazolják, hogy a multi-label klasszifikáció jogosan tölthet be igen fontos szerepet a machine learning témakörén belül – az adatforradalom küszöbén állva ugyanis nem hagyhatók figyelmen kívül azok a modellstruktúrák, melyek a sokféle kimeneti lehetőséget, kategóriát megfelelően képesek egymás mellett, egyidejűleg kezelni. Komplexebb, nehezebben kvantifikálható feladatoknál, mint például egy téma társadalmi narratívájának a vizsgálata, láthatóan fontos szempont olyan módszereket használni, melyek képesek a sokféle kimenet egymással való kapcsolatát, együttes előfordulását vizsgálni – hiszen így kevésbé jól megfogható, absztraktabb módon artikulálódó nézőpontok is jobban megközelíthetővé válhatnak. Különösen fontos ez egy olyan érzékeny, még mindig sok tabuval, elfojtással vagy épp megvetéssel övezett, mégis széles köröket érintő témánál, mint a mentális betegségek, zavarok. Társadalomtudományos felelősségnek is tekinthető, hogy azzal párhuzamosan, hogy az emberek egyre többen, többször és többet beszélnek erről és ehhez hasonló témákról, az elérhető adatok segítségével mi is lépést tarthassunk ezzel a közbeszédben megjelenő paradigmaváltással – egyes témákat övező tabuk leomlása talán az egyik legfontosabb társadalmi mechanizmus, nincs hát ok rá, hogy ne próbáljuk meg minél jobban megérteni azt.

6. Irodalomjegyzék

Abramson, L.Y., Alloy, L. B., Hankin, B. L., Haeffel, G. J., MacCoon, D. G., & Gibb, B. E. (2002). *Cognitive vulnerability – Stress models of depression in a self-regulatory and psychobiological context*. In Gotlib, I. H. & Hammen, C. L. (szerk.). *Handbook of depression.* Guilford Press, New York, NY.

Alloy, L., Fedderly, S., Kennedy-Moore, E., & Cohan, C. (1998). Dysphoria and social interaction: An integration of behavioral and confirmation and interpersonal perspectives. *Journal of Personality and Social Psychology.* 74(6): 1566-1579.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5. kiadás). American Psychiatric Publishing, Arlington, VA.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4. kiadás). American Psychiatric Publishing, Washington, DC.

Andrews, P. W., & Thompson, J. A. (2009). The bright side of being blue: depression as an adaptation for analyzing complex problems. *Psychological Review*. 116(3): 620-654.

Anglin, R. E., Tarnopolsky, M. A., Mazurek, M. F., & Rosebush, P. I. (2012). The Psychiatric Presentation of Mitochondrial Disorders in Adults. *The Journal of Neuropsychiatry and Clinical Neurosciences.* 24(4): 394-409.

Appleton, K. M., Sallis, H. M., Perry, R., Ness, A. R., & Churchill, R. (2015). Omega-3 fatty acids for depression in adults. *The Cochrane Database of Systematic Reviews.* (11): CD004692.

Avasthi, A. (2006). Are social theories still relevant in current psychiatric practice. *Indian Journal of Social Psychiatry.* 32(1): 3-9.

Babenko, B. (2008). *Multiple Instance Learning: Algorithms and Applications.* http://ailab.jbnu.ac.kr/seminar_board/pds1_files/bbabenko_re.pdf (letöltve: 2020. 10. 03.)

Barnes, E. (2003). *Young People and Depression: A Sociological Perspective.* Masters thesis, National University of Ireland Maynooth.

Beck, A. T. (2002). Cognitive Models of Depression. In: Leahy, R. (szerk.). *Clinical Advances in Cognitive Psychotherapy: Theory and Application.* Springer Publishing Company, New York, NY.

Beck, A.T., & Weishaar, M. E. (2011). Cognitive Therapy. In Corsini, R. J. & Wedding, D.

(szerk.). *Current psychotherapies* (9. kiadás). Brooks/Cole, Belmont, CA.

Beloucif, S. (2013). Informed consent for special procedures: electroconvulsive therapy and psychosurgery. *Current Opinion in Anesthesiology*. 26(2): 182-185.

Bergen, B. K. (2012). *Louder Than Words: The New Science of How the Mind Makes Meaning.* Basic Books, New York, NY.

Bernardini, F. C., da Silva, R. B., Rodovalho, R. M., & Meza, E. B. M. (2014). Cardinality and Density Measures and Their Influence to Multi-Label Learning Methods. *Learning and Nonlinear Models.* 12(1): 53-71.

Berrios, G. E. (1997). The scientific origins of electroconvulsive therapy. *History of Psychiatry*. 8(29, pt. 1): 105-119.

Blume, J., Douglas, S. D., Evans, & Evans, D. L. (2011). Immune Suppression and Immune Activation in Depression. *Brain, Behavior and Immunity.* 25(2): 221-229.

Bonanno, G. A. (2004). Loss, trauma, and human resilience. *American Psychologist.* 59(1): 20–28.

Borrell-Carrió, F., Suchman, A., & Epstein, R. (2004). The Biopsychosocial Model 25 Years Later: Principles, Practice, and Scientific Inquiry. *Annals of Family Medicine*. 2(6): 576-582.

Boursier, V., Gioia, F., Coppola, F., & Schimmenti, A. (2019). Digital storytellers: Parents facing with children's autism in an Italian web forum. *Mediterranean Journal of Clinical Psychology,* 7(3). https://cab.unime.it/journals/index.php/MJCP/article/view/2104/pdf (letöltve: 2020. 10. 01.)

Boutell, M. R., Luo, J., Shen, X., & Brown, C. M. (2004). Learning multi-label scene classification. *Pattern Recognition*. 37(9): 1757-1771.

Bridle, C., Spanjers, K., Patel, S., Atherton, N. M., & Lamb, S. E. (2012). Effect of exercise on depression severity in older people: systematic review and meta-analysis of randomised controlled trials. *The British Journal of Psychiatry.* 201(3): 180-185.

Brockmann, H., Zobal, A., Schumacher, A., Daamen, M., Joe, A., Biermann, K., & Biecker, H. (2011). Influence of 5-HTTLPR polymorphism on resting state perfusion in patients with major depression. *Journal of Psychiatric Research*. 45(4): 442–451.

Brommelhoff, J. A., Conway, K., Merikangas, K., & Levy, B. R. (2004). Higher rates of depression in women: Role of gender bias within the family. *Journal of Women's Health*. 13(1): 69–76.

Busch, F. N., Rudden, M. G., & Shapiro, T. (2004). *Psychodynamic treatment of depression.* American Psychiatric Publishing, Washington, DC.

Canoy, N., & Topacino, A. M. D. C. (2020). Unhearing Online Suicide Talk: Becoming-Voice through the Use of Maddening Poetic Conversations. *Journal of Constructivist Psychology,* DOI: 10.1080/10720537.2020.1727392

Carlson, N. R. (2013). *Psychology of behavior.* Perason, Boston, MA.

Cerri, R., Barros, R. C., & de Carvalho, A. C. P. L. F. (2014). Hierarchical multi-label classification using local neural networks. Journal of Computer and System Sciences. 80(1): 39-56.

Chen, Y-Y., Lin, Y-H., Kung, C-C., Chung, M-H., & Yen, I-H. (2019). Design and Implementation of Cloud Analytics-Assisted Smart Power Meters Considering Advanced Artificial Intelligence as Edge Analytics in Demand-Side Management for Smart Homes. *Sensors*. 19 (9): 2047.

Ciraulo, D. A., Evans, J. A., Qiu, W. Q., Shader, R. I., & Salzman, C. (2011). Antidepressant treatment of geriatric depression. In: Ciraulo, D. A., & Shader, R. I. Pharmacotherapy for depression. (2. kiadás). Springer Science & Business Media, New York, NY. pp. 125-183.

Clare A., & King R.D. (2001). Knowledge Discovery in Multi-label Phenotype Data. In: De Raedt L., & Siebes A. (szerk.). *Principles of Data Mining and Knowledge Discovery. PKDD 2001.* Springer, Berlin.

Comer, R. J. (2015). *Abnormal Psychology.* Worth Publishers, New York, NY.

Cooney, G. M., Dwan, K., Greig, C. A., Lawlor, D. A., Rimer, J., Waugh, F. R., McMurdo, M., & Mead, G. E. (2013). Exercise for depression. *The Cochrane Database of Systematic Reviews.* 9(9): CD004366.

Cover, T. M., & Hart, P. E. (1967). Nearest neighbor pattern classification. *IEEE Transactions of Information Theory.* 13(1): 21-27.

Culpepper, L., Muskin, P. R., & Stahl, S. M. (2015). Major Depressive Disorder: Understanding the Significance of Residual Symptoms and Balancing Efficacy with Tolerability. *The American Journal of Medicine*. 128 (Suppl. 9): S1–S15.

de Zwart, P. L., Jeronimus, B. F., & de Jorge, P. (2019). Empirical evidence for definitions of episode, remission, recovery, relapse and recurrence in depression: a systematic review. *Epidemiology and Psychiatric Sciences*. 28 (5): 544–562

Defazio, A., Bach, F., & Lacoste-Julien, S. (2014). *SAGA: A Fast Incremental Gradient Method With Support for Non-Strongly Convex Composite Objectives.* https://arxiv.org/abs/1407.0202 (letöltve: 2020. 11. 05.)

Dembczyński, K., Waegeman, W., Cheng, W., & Hüllermeier, E. (2012). On label dependence and loss minimization in multi-label classification. *Machine Learning.* 88: 5-45.

Dewdney, A. K. (1997). *Yes, We Have No Neutrons: An Eye-Opening Tour Through the Twists and Turns of Bad Science.* Wiley, New York, NY. pp. 82.

Disease and Injury Incidence and Prevalence Collaborators (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 392 (10159): 1789–1858.

Du, J., Chen, Q., Peng, Y., Xiang, Y., Tao, C., & Lu, Z. (2019). ML-Net: multi-label classification of biomedical texts with deep neural networks. *Journal of the American Medical Informatics Association*. 26(11): 1279–1285.

Durkheim, É. (2003). *Az öngyilkosság.* Budapest, Osiris.

Elder, B. L., & Mosack, V. (2011). Genetics of Depression: An Overview of the Current Science. *Mental Health Nursing.* 32(4): 192-202.

Elisseeff, A., & Weston, J. (2002, december 3-8). *A kernel method for multi-labelled classification* [Paper presentation]. Advances in Neural Information Processing Systems 14, Vancouver.

Emmons, K. K. (2010*). Black dogs and blue words: Depression and gender in the age of self-care.* Rutgers University Press, Piscataway, NJ.

Engel, G. L. (1977). The need for a new medical model: a challenge for biomedicien. *Science.* 196(4286): 129-136.

Entman, R. M. (1993). Framing: towards clarification of a fractured paradigm. *Journal of Communication.* 43(4): 51-58.

Farmer, R. F., & Chapman, A. L. (2008). *Behavioral interventions in cognitive behavior therapy: Practical guidance for putting theory into action.* American Psychological Association, Washington, DC.

Fava, G. A., Park, S. K., & Sonino, N. (2006). Treatment of recurrent depression. *Expert Review of Neurotherapeutics.* 6(11): 1735-1740.

Fochtmann, L. J., & Gelenberg, A. J. (2005). Guideline Watch: Practice Guideline for the Treatment of Patients With Major Depressive Disorder, 2nd Edition. *Focus.* 3(1): 34-42.

Francis, C. (1989). The recent excitement about neural networks. *Nature*. 337(6203): 129-132.

Freund, Y. & Schapire, R.E. (1997), A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*. 55(1): 119-139.

Fryers, T., Melzer, D., Jenkins, R., Brugha, T. (2005). The distribution of the common mental disorders: social inequalities in Europe. *Clinical Practice and Epidemiology in Mental health.* 1(14).

Gao, S-F., & Bao, A-M. (2011). Corticotropin-releasing hormone, glutamate, and γ-aminobutyric acid in depression. *Neuroscientist.* 17(1): 124-144.

Geddes, J. R., Carney, S. M., Davies, C., Furukawa, T. A., Kupfer, D. J., Frank, E., & Goodwin, G. M. (2003). Relapse prevention with antidepressant drug treatment in depressive disorders: a systematic review. *Lancet*. 361(9358): 653–661.

Gross, R. (2015). *Psychology: The Science of Mind and Behaviour.* (7. kiadás). Hodder Education, London.

Gilbert, P. (2007). *Psychotherapy and counselling for depression.* SAGE, Los Angeles.

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. (szerk.). *Pharmacotherapy of depression* (2. kiadás). Springer Science & Business Media, New York, NY. pp. 1-32.

Gonçalves, E.C., Plastino, A., & Freitas, A.A. (2013). A genetic algo-rithm for optimizing the label ordering in multi-label classifier chains. In: *2013 IEEE 25th International Conference on Tools with ArtificialIntelligence, IEEE.* pp. 469–476

Grodzicki, R., Mańdziuk, J., & Wang, L. (2008). Improved Multilabel Classification with Neural Networks. In: Rudolph, G., Jansen, T., Beume, N., Lucas, S., & Poloni, C. (szerk.) *Parallel Problem Solving from Nature – PPSN X.* Springer, Berlin. pp. 409-416.

Grus, J. (2015). *Data Science from Scrach: First Principles with Python.* O'Reilly, Sebastopol, CA. pp. 226-232.

Halls, A., Muller, I., & Angier, E. (2018). 'Hope you find your 'eureka' moment soon': a qualitative study of parents/carers' online discussions around allergy, allergy tests and eczema. *BMJ Open,* 8(11). https://bmjopen.bmj.com/content/8/11/e022861 (letöltve: 2020. 10. 01.)

Han, D. J., & Yu, K. (2001). Idiot's Bayes — not so stupid after all? *International Statistical Review*. 69(3): 385-399.

Hanley, J. A., & McNeil, B. J. (1982). The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. *Radiology*. 143: 29-36.

Higgins, D. (2018 május 30-31). *Classify all the Things (with multiple labels): The most common type of modeling task no one talks about* [conference presentation]. Data Science Leaders Summit. New York, NY.

Holtz, P., Kronberger, N., & Wagner, W. (2012). *Journal of Media Psychology Theories Methods and Applications*, 24(2): 55-66.

Homans, P. (1999). We (Not so) Happy Few: Symbolic Loss and Mourning in Freud's Psychoanalytic Movement and the History of Psychoanalysis. *Psychoanalysis and History.* 1(1): 69-86.

Hsu, C-W., Chih-Chung, C., & Chih-Jen, L. (2003). A Practical Guide to Support Vector Classification. http://www.datascienceassn.org/sites/default/files/Practical%20Guide%20to%20Support%20Vector%20Classification.pdf (letöltve: 2020. 11. 05.)

Jacobson, N., Martell, C., & Dimidjian, S. (2001). Behavioral activation treatment for depression: Returning to contextual roots. *Clinical Psychology: Science and Practice.* 8(3): 255-270.

Jin, R., & Ghahramani, Z. (2002). Learning with multiple labels.

Joachims, T. (1998). Text categorization with Support Vector Machines: Learning with many relevant features. In: Nédellec, C., & Rouveirol, C. (szerk.) *Machine Learning: ECML-98.* Springer, Berlin, Heidelberg. pp. 137-142.

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. (szerk.). *Treatment resistant depression: A roadmap for effective care*. American Psychiatric Publishing, Arlington, VA. pp. 293–308.

Kessler, R. C., & Bromet, E. J. (2013). The epidemiology of depression across cultures. *Annual Review of Public Health.* 34: 119–38.

Kline, N. S. (1958). Clinical experience with iproniazid (Marsilid). *Journal of Clinical and Experimental Psychopathology.* 19(1, Suppl.): 72–78.

Krakowski, M. (2018). Approaches for the Improvement of the Multilabel Multiclass Classification with a Huge Number of Classes. In: Klassen, G., & Conrad, S. (szerk.). *Proceedings of the 30th GI-Workshop Grundlagen von Datenbanken.* pp. 65-70. http://ceur-ws.org/Vol-2126/ (letöltve: 2020. 10. 26).

Kuhn, R. (1958). The treatment of depressive states with G-22355 (imipramine hydrochloride). *American Journal of Psychiatry.* 115: 459–464.

Landau, M. J., Meier, B., & Keefer, L. (2010) A metaphor enriched social cognition. *Psychological Bulletin.* 136(6): 1045-1067.

Latkovikj, M. T., & Popovska, M. B. (2020). Online research about online research: advantages and disadvantages. *E-methodology*, 6(6): 44-56.

Leitão, R. (2019). Technology-Facilitated Intimate Partner Abuse: a qualitative analysis of data from online domestic abuse forums. *Human-Computer Interaction (online).* DOI: 10.1080/07370024.2019.1685883

Lenc, L., & Král, P. (2017). Ensemble of Neural Networks for Multi-label Document Classification. *ITAT 2007.* CreateSpace Independent Publishing Platform, Scotts Valley, CA. pp. 186-192.

Levinson, D. F., & Nichols, W. E. (2018). Genetics of Depression. In: Chanrey, D. S., Sklar, P., Buxbaum, J. D., & Nestler, E. J. (szerk.) *Charney & Nestlers Neurobiology of Mental Illness.* (5. kiadás). Oxford University Press, New York, NY.

Levinson, D. F., & Nichols, W. E. (2014). Major depression and genetics. Stanford, School of Medicine, Stanford, CA.

In Charney, Dennis S.; Sklar, Pamela; Buxbaum, Joseph D.; Nestler, Eric J. (eds.). Charney & Nestlers Neurobiology of Mental Illness (5th ed.). New York: Oxford University Press. p. 310.

Limosin, F., Mekaoui, L., & Hautecouverture, S. (2007). Prophylactic treatment for recurrent major depression. *Presse Médicale.* 36(11, pt. 2): 1627-1633.

Madjarov, G., Kocev, D., Gjorgjevikj, D., & Saso, D. (2012). An Extensive Experimental Comparison of Methods for Multi-Label Learning. *Pattern Recognition.* 45(9): 3084-3104.

Maron, O., & Lozano-Perez, T. (1997). A framework for multiple-instance learning. Proceedings of Neural In-formation Processing Systems, 10. https://papers.nips.cc/paper/1346-a-framework-for-multiple-instance-learning (letöltve: 2020. 10. 03.)

Martín-Blanco, A., Soler, J., Villalta, L., Feliu-Soler, A., Elices, M., Pérez,V., & Pascual, J. C. (2014). Exploring the interaction between childhood maltreatment and temperamental traits on the severity of borderlinepersonality disorder. *Comprehensive Psychiatry*. 55(2): 311–318.

McCallum, A., & Nigam, K. (1998). A comparison of event models for Naive Bayes text classification. *AAAI-98 workshop on learning for text categorization.* 41-48.

McNeely, S. (2012). Sensitive Issues in Surveys: Reducing Refusals While Increasing Reliability and Quality of Responses to Sensitive Survey Items. In: L. Gideon (Szerk.), *Handbook of Survey Methodology for the Social Sciences* (pp.377-396). Springer, New York.

National Institute for Health and Care Excellence. (2009, október 28). *Depression in adults: recognition and management.* https://www.nice.org.uk/guidance/cg90 (letöltve: 2020. 10. 03.)

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: 1-18.

Noack-Lundberg, K., Liamputtong, P., Marjadi, B., & Ussher, J. (2019). Sexual violence and safety: the narratives of transwomen in online forums. *Culture, Health & Sexuality,* 22(2): 1-14.

Nonacs, R. M. (2019, október 11.). *Postpartum depression.* https://reference.medscape.com/article/271662-overview (letöltve: 2020. 10. 03.)

Orvos-Tóth, N. (2018). *Örökölt sors. Családi sebek és a gyógyulás útjai.* Kulcslyuk, Budapest.

Parker, G. (1993). Parental rearing style: examining for links with personality vulnerability factors for depression. *Psychiatry and Psychiatric Epidemiology*. 59(1): 20-28.

Parker, G., & Hyett, M. (2010). Screening for depression in medical settings: Are specific scales useful? In Mitchell, A. J., & Coyne, J. C. (szerk.). *Screening for depression in clinical practice: An evidence-based guide.* Oxford University Press, New York, NY. pp. 191–201.

Parker, G. B., Brotchie, H., & Graham, R. K. (2017). Vitamin D and depression. *Journal of Affective Disorders.* 208: 56-61.

Paykel, E. S., & Cooper, Z. (1992). Life events and social stress. In Paykel, E. S. (szerk.). *Handbook of affective disorders.* Guilford Press, New York, NY.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research.* 12: 2825-2830.

Posternak, M. A., & Miller, I. (2001). Untreated short-term course of major depression: a meta-analysis of outcomes from studies using wait-list control groups. *Journal of Affective Disorders*. 66 (2–3): 139-46.

Powers, D. M. W. (2007). Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation. *Journal of Machine Learning Technologies.* 2(1): 37-63.

Prkachin, K., Craig, K., Papageorgis, D., & Reith, G. (1977). Nonverbal communication deficits and response to performance feedback in depression. *Journal of Abnormal Psychology.* 86(3): 224-234.

Quinlan, R. (1993). *C4.5: Programs for Machine Learning.* Morgan Kaufmann Publishers, San Mateo, CA.

Ramos, D., Franco-Pedroso, J., Lozano-Diez, A., & Gonzalez-Rodriguez, J. (2018). Deconstructing Cross-Entropy for Probabilistic Binary Classifiers. *Entropy*. 20(3): e20030208.

Read, J., Pfahringer, B., Holmes, G., & Frank, E. (2011). Classifier Chains for Multi-Label Classification. *Machine Learning*. 85(3): 333-359.

Reali, F., Soriano, T., & Rodríguez, D. (2016). How we think about depression: The role of linguistic framing. *Revista Latinoamericana de Psicología*. 48(2): 127-136.

Sandler, M. (1990). *Monoamine oxidase inhibitors in depression: History and mythology.* Journal of Psychopharmacology. 4(3): 136-139.

Schapire, R. E., & Singer, Y. (2000), Boostexter: a boosting-based system for text categorization, *Machine Learning.* 39(2/3): 135-168.

Schultz, G. (2007, May 24). *Marital breakdown and divorce increases rates of depression, StatCan study finds.* https://www.lifesitenews.com/news/marital-breakdown-and-divorce-increases-rates-of-depression-statcan-study-f (letöltve: 2020. 11. 01.)

Seeman, N., Tang, S., Brown, A. D., & Ing, A. (2015). World survey of mental illness stigma. *Journal of Affective Disorders*, 190: 115-121.

Seligman, M. E. P. (1975). Helplessness. Freeman, San Francisco, CA.

Sorower, M. S. (2010). A Literature Survey on Algorithms for Multi-label Learning. *Oregon State University.* 18: 1-25.

Syaliman, K. U., Nabahan, E. B., & Sitompul, O. S. (2017). Improving the accuracy of k-nearest neighbor using local mean based and distance weight. *Journal of Physics: Conference Series.* 978: 012047.

Szymański, P., Kajdanowicz, T. (2019). Scikit-multilearn: A Python library for Multi-Label Classification. *Journal of Machine Learning Research.* 20(6): 1-22.

Tan, S. (2005). Neighbor-weighted K-nearest neighbor for unbalanced text corpus. *Expert Systems with Applications.* 28: 667-671.

Taube-Schiff, M., & Lau, M. A. (2008). Major depressive disorder. In Hersen, M. & Rosqvist, J. (szerk.). Handbook of psychological assessment, case conceptualization, and treatment, Vol. 1: Adults. John Wiley & Sons, Hoboken, NJ. pp. 319-351.

Tourangeau, R., & Yan, T. (2007). Sensitive Questions in Surveys. *Psychological Bulletin*, 133(5): 859-883.

Tsoumakas, G., & Katakis, J. (2009). Multi-Label Classification: An Overview. *International Journal of Data Warehousing and Mining*. 3(3): 1-13.

Tsoumakas, G., Katakis, I., & Vlahavas, I. (2011). Random k-labelsets for multi-label classification. *IEEE Transactions on Knowledge and Data Engineering*. 23(7): 1079-1089.

Tsoumakas, G., & Vlahavas, V. (2007). Random k-labelsets: An ensemble method for multilabel classification. In: Kok, J. N., Koronacki, J., Lopez de Mantaras, R., Matwin, S., & Mladenic, S. (szerk.) *Machine Learning: ECML 2007: The 18th European conference on machine learning.* Springer, Heidelberg. pp. 406-417.

Vallet, A., & Sakamoto, H. (2015). Multi-Label Convolutional Neural Network forAutomatic Image Annotation. *Journal of Information Processing*. 23(6): 767-776.

Veen, G., van Vliet, I. M., DeRijk, R. H, Giltay, E. J., van Pelt, J., & Zitman, F. G. (2011). Basal cortisol levels in relation to dimensions and DSM-IV categories of depression and anxiety. *Psychiatric Research.* 185(1-2): 121-128.

Weissman, M. M., Livingston, B. M., Leaf, P. J., Florio, L. P., & Holzer, C., III. (1991). Affective disorders. In Robins, L. N. & Regier, D. A. (szerk.). *Psychiatric disorders in America: The Epidemiologic Catchment Area Study.* Free Press, New York, NY.

Wu, X-Z., & Zhou, Z-H. (2017). A Unified View of Multi-Label Performance Measures. *ICML 2017: Proceedings of the 34th International Conference on Machine Learning.* 70: 3780-3788.

Yazici, V. O., Gonzalez-Garcia, A., Rimasa, A., Twardowski, B., & de Weijer, J. (2019). Orderless Recurrent Models for Multi-label Classification. https://arxiv.org/abs/1911.09996v3 (letöltve: 2020. 10. 24.)

Zhang, M-L., & Zhou, Z-H. (2014). A review on multi-label learning algorithms. *IEEE Transactions on Knowledge and Data Engineering.* 26(8): 1819-1837.

Zhang, M-L., & Zhou, Z-H. (2006). Multilabel Neural Networks with Applications to Functional Genomics and Text Categorization. *IEEE Transactions on Knowledge and Data Engineering.* 18(10): 1338-1351.

Zhang, M-L., Zhou, Z-H. (2007). ML-KNN:A lazy learning approach to multi-label learning. *Pattern Recognition.* 40(2007): 2038 – 2048.

Zhang, M-L., Li, Y., Liu, X., & Geng, Y. (2018). Binary relevance for multi-label learning: an overview. *Frontier of Computer Science*. 12(2): 191-202.

Zhang, Y., & Jin, Y. (2017). Thematic and Episodic Framing of Depression: How Chinese and American Newspapers Framed a Major Public Health Threat. *Athens Journal of Mass Media and Communications*. 3(2): 91-106.

Zimmerman, M., Martin, J., McGonigal, P., Harris, L., Kerr, S., Balling, C., Keifer, R., Stanton, K., & Dalrymple, K. (2018). Validity of the dsm-5 anxious distress specifier for major depressive disorder. *Depression and Anxiety*, 36(1): 31-38.