

37. Országos Tudományos Diákköri Konferencia  
Had- és Rendészettudományi Szekció  
Budapest, április 22–24  
Nemzeti Közszolgálati Egyetem  
Rendészettudományi Kar

---

## **A mesterséges intelligencia és a gépi tanulás hatása a dezinformációra**

Jelige: A megtévesztés korszaka

Pályamunka lezárásának időpontja:

2024. május

# Tartalomjegyzék

|                                                                                                              |    |
|--------------------------------------------------------------------------------------------------------------|----|
| Bevezetés.....                                                                                               | 4  |
| Téma aktualitás .....                                                                                        | 4  |
| Tudományos probléma megfogalmazása .....                                                                     | 5  |
| Kutatási célkitűzések .....                                                                                  | 5  |
| Hipotézisek .....                                                                                            | 6  |
| Kutatási módszertan.....                                                                                     | 6  |
| 1. A gépi tanulás és a mesterséges intelligencia a dezinformációban .....                                    | 7  |
| 1.1 A mesterséges intelligencia és a gépi tanulás fogalma .....                                              | 7  |
| 1.2 A mesterséges intelligencia tanításához szükséges adatgyűjtési fázis.....                                | 10 |
| 1.3 Szentiment analízis és álláspont vizsgálat (Stance detection) .....                                      | 11 |
| 1.4 Generatív ellenséges hálózatok (generative adversarial networks/GANs) használata a dezinformációban..... | 14 |
| 1.5 Tartalomkészítés természetes nyelvi generálással (NLG) .....                                             | 17 |
| 1.6 Közösségi botok (social bots) használata dezinformálásra .....                                           | 19 |
| 1.7 A dezinformáció automatizálása chatbotok segítségével .....                                              | 20 |
| 1.8 Következtetések .....                                                                                    | 23 |
| 2 A gépi tanulás és a mesterséges intelligencia szerepe a dezinformáció elleni küzdelemben .....             | 24 |
| 2.1 A dezinformáció szövegelemzés alapú felismerése .....                                                    | 24 |
| 2.2 Deepfake felismerése gépi tanulási modellek által.....                                                   | 27 |
| 2.3 Közösségi botok detektálása a gépi tanulás segítségével .....                                            | 30 |
| 2.4 A dezinformáció gráf és hálózatalapú felismerése.....                                                    | 33 |
| 2.5 Következtetések .....                                                                                    | 35 |
| 3 Az empirikus kutatás bemutatása .....                                                                      | 36 |
| 3.1 A kérdőív kiértékelésének és összeállításának a módszere .....                                           | 36 |
| 3.2 Empirikus kutatásom elemzése .....                                                                       | 38 |
| 3.3 Következtetések .....                                                                                    | 45 |
| Összegzett következtetések .....                                                                             | 45 |
| Eredmények.....                                                                                              | 46 |

|                                                          |    |
|----------------------------------------------------------|----|
| Felhasznált irodalom .....                               | 47 |
| 4. Függelék.....                                         | 50 |
| 4.1 Táblázat jegyzék .....                               | 50 |
| 4.2 Ábrák jegyzéke.....                                  | 50 |
| 4.3 1.számú melléklet (Kérdőíves felmérés kérdései)..... | 51 |
| 4.4 2.számú melléklet (Kérdőívek kiértékelése).....      | 55 |
| Rövidítések jegyzéke .....                               | 59 |

## Bevezetés

### Téma aktualitás

Manapság az egyre jobban digitalizálódó világban, folyamatosan még inkább releváns kérdéssé válik, hogy honnan szerezzük be napi híreinket. A múltban és a jelenben is az információ képviselte és képviseli az egyik legnagyobb értéket. Napjainkban egyre nehezebbé válik megkülönböztetni a hamis vagy félrevezető információkat, híreket a valósaktól. Ezért hatalmas probléma a dezinformáció jelensége a 21. században.<sup>1</sup> A dezinformáció olyan hamis információ, amelyet szándékosan terjesztenek a közvélemény befolyásolása vagy az igazság elfedése érdekében.<sup>2</sup>

Ezeknek a folyamatoknak a súlyosbodása nem csak a digitalizálódásnak és a közösségi média által nyújtott platformoknak köszönhető, hanem az egyre nagyobb technológiai fejlődésnek is. A mesterséges intelligencia és a gépi tanulás robbanásszerű fejlődése lehetővé teszi a dezinformáció könnyebb terjesztését. Azonban ezeket a lehetőségeket nem csak rossz célokra lehet felhasználni, hanem az álhírek és dezinformáció elleni küzdelemre és észlelésére is. Több esemény is bizonyítja, hogy a mesterséges intelligencia és különböző ágai dezinformációs kampányokban is felhasználhatók. Az egyik kiemelkedő eset a 2016-os amerikai elnökválasztás, amelyben közösségi botokat használtak a választás menetének befolyásolására.<sup>3</sup> Brazíliában a 2018-as elnökválasztás során chatbotokat használtak a WhatsApp alkalmazáson a jelöltekről szóló dezinformáció terjesztésére.<sup>4</sup> Az Egyesült Királyságban pedig a brexiti népszavazás alatt közösségi botokat alkalmaztak a Twitter nevű platformon.<sup>5</sup> Véleményem szerint ezek a példák jól mutatják, hogy a mesterséges intelligenciát és különböző ágait egyre többször veszik használatba a dezinformáció terjesztése érdekében és ily módon jelentős kockázatokat hordozva. Ezért az állami szereplőknek és a közösségi média platformoknak fontos szerepet kell játszaniuk az álhírek és dezinformáció elleni küzdelemben.

---

<sup>1</sup> D.M.J. Lazer és mtsai., „The Science of Fake News: Addressing Fake News Requires a Multidisciplinary Effort”, *Science* 359, sz. 6380 (2018): 1056–57, <https://doi.org/10.1126/science.aao2998>.

<sup>2</sup> Don Fallis, „What Is Disinformation?”, *Library Trends* 63, sz. 3 (2015): 401–26, <https://doi.org/10.1353/lib.2015.0014>.

<sup>3</sup> Adam Badawy, Emilio Ferrara, és Kristina Lerman, „Analyzing the Digital Traces of Political Manipulation: The 2016 Russian Interference Twitter Campaign”, in *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 2018, 258–65, <https://doi.org/10.1109/ASONAM.2018.8508646>.

<sup>4</sup> Marco Aurélio Ruediger és mtsai., „Bots and Brazil’s Electoral Legal System: 2018 Elections”, 2019. január, <http://bibliotecadigital.fgv.br:80/dspace/handle/10438/26229>.

<sup>5</sup> Philip N. Howard és Bence Kollanyi, „Bots, #StrongerIn, and #Brexit: Computational Propaganda during the UK-EU Referendum” (arXiv, 2016. június 20.), <https://doi.org/10.48550/arXiv.1606.06356>.

## **Tudományos probléma megfogalmazása**

A mesterséges intelligencia manapság egy rohamosan fejlődő ága a tudománynak. Ez sok problémát felvet, mivel a technológia fejlesztése nem csak a kutatók és átlag felhasználók számára hasznos, hanem a bűnözők vagy fenyegető szereplők számára egyaránt. Az innováció és a kutatás és fejlesztés gyors fejlődésének érdekében ezeket a technológiákat nem minden esetben felügyelik és korlátozzák. Erre egy különösen jó példa a chatbotok politikai radikalizálódása, ha hozzáféréssel rendelkeznek az internethez. A dezinformációs tér manapság hatalmas probléma. Olyannyira, hogy egy Észak-atlanti Szerződés Szervezete (NATO) által támogatott kutatásban felmerült, hogy a kognitív hadviselés legyen a hatodik hadviselési domén.<sup>6</sup> Kutatásomban ennek a két területnek az összefonódását vizsgálom, a fenyegető szereplő és a dezinformációt észlelő szakember szempontjából.

## **Kutatási célkitűzések**

Kutatásommal fő célkitűzésem, hogy megvizsgáljam a mesterséges intelligencia technológiai fejlődésének hatásait a dezinformáció terjesztésére és annak megakadályozására. Az első fejezetben a fenyegető szereplő szemszögéből vizsgálom a mesterséges intelligencia és gépi tanulás alapú programok nyújtotta lehetőségeket. Ezt azért vizsgálom elsőként, mivel úgy vélem, a problémát először a fenyegető szemszögéből kell megérteni és utána tudunk hatékonyan választ adni rá. Kutatásom második fejezetében azt vizsgálom, hogy kutatók és a közösségi média vállalatok, hogyan próbálják ezeket az álhíreket vagy dezinformációt kiszűrni és észlelni. Itt van olyan ága a mesterséges intelligenciának - mint például a természetes nyelvi feldolgozás -, melyet a terjesztés és észlelési fázisban is felhasználhatnak.

Kutatásom utolsó részében pedig kérdőívem kitöltési eredményei alapján fogom vizsgálni, hogy van-e összefüggés az életkor és a közösségi oldalak, a dezinformáció és a mesterséges intelligencia percepcionálása között. Itt megvizsgálom, mennyire tartják a kitöltők a dezinformációt és álhíreket problémásnak. Továbbá, hogy az életkor milyen erős tényező a közösségi oldalakon látottak befolyásoló erejének és a mesterséges intelligencia fejlődésének megítélésében.

---

<sup>6</sup> Hervé Le Guyader. „Cognitive Domain: A Sixth Domain of Operations.” Bernard Claverie; Baptiste Prébot; Norbou Buchler; François du Cluzel. *Cognitive Warfare: The Future of Cognitive Dominance*, NATO Collaboration Support Office, pp.3, 1-5, 2022, 978-92-837-2392-9. (hal-03635898)

## **Hipotézisek**

Kutatásomhoz a következő hipotéziseket fogalmaztam meg:

H1: A fiatalabb generáció hajlamosabb függővé válni a közösségi oldalak használatától, mint az idősebb korosztály.

H2: A mesterséges intelligencia fejlesztése felé irányuló optimizmus magasabb a fiatalabb generációnál, mint az idősebb korosztálynál.

H3: Az idősebb korosztályhoz képest, a fiatalabb generáció nagyobb valószínűséggel érzi, hogy a közösségi média megkönnyítette más emberek számára, hogy a magánéletükbe behatoljanak.

## **Kutatási módszertan**

Kutatásom alatt nemzetközileg elismert szakirodalomra támaszkodtam, melyet a Scopus bibliográfiai adatbázisának segítségével gyűjtöttem össze. A felhasznált kutatások és folyóiratok hitelességét a Scimago Journal Rank által vizsgáltam meg. Kutatásom végén a hipotéziseim alátámasztása érdekében kérdőívet készítettem és az IBM SPSS Statistics 27 szoftverprogrammal vizsgáltam kérdőívem eredményeit. Itt megvizsgáltam, hogy különböző faktorok, mint végzettség, nem és életkor mennyire befolyásolja az általam feltett kérdésekre a válaszokat. A kérdőívem kérdéseit a „Technology Adoption Propensity” (TAP), vagyis a technológia elfogadási hajlandósági modell által dolgoztam ki. A TAP modell segített abban, hogy kérdéseimet négy kategória alapján alkossam meg, mely később segítségemre volt a specifikus kérdések összevetésében.

Kutatásomban a hivatkozásokat a Zotero nevű hivatkozáskezelő program segítségével készítettem a Chicago Manual of Style of 17th Edition hivatkozási stílusban.

## 1. A gépi tanulás és a mesterséges intelligencia a dezinformációban

A technológiai fejlődés alapvetően nagyon sok képességet megerősít a dezinformációban és megkönnyíti a dezinformáló szereplők munkáját, de nem teljesen automatizálja azt. Ma már számtalan nyíltforrásból elérhető eszköz van, melyek dezinformációra alkalmazhatóak és bárki számára elérhetőek, legyen az fizetős vagy ingyenes szolgáltatás. A következő fejezetben a generatív ellenséges hálózatokat fogom bemutatni, amelyek a deepfake készítésében és szintetikus hangok generálásában játszanak fő szerepet. A neuro-lingvisztikus programozáson (Neuro-Linguistic Programming) belül a természetes nyelvi feldolgozás (Natural Language Processing) és a természetes nyelvi generálás (Natural Language Generation) eszközei segítettek abban, hogy a gépi tanulási rendszerek képesek legyenek szöveget olvasni, írni és értelmezni, a közösségi botok (social bots) pedig segítettek a hamis információk elterjesztésében. Ezeknek a terminológiáknak a fogalmát kifejtem a rájuk szánt fejezeteken belül. Végül bemutatom a chatbotok hatását a dezinformációs környezetre. A kutatásomban egy Center for Security and Emerging Technology tanulmány által lefektetett dezinformációs fázisokat fogom alkalmazni.<sup>7</sup> Ezeket a fázisokat a 2. ábrán tüntettem fel.

### 1.1 A mesterséges intelligencia és a gépi tanulás fogalma

Nagyon fontos megérteni, hogy mit is értünk mesterséges intelligencia és gépi tanulás alatt. A mesterséges intelligencia az emberi intelligencia szimulációját jelenti olyan gépekben, amelyeket oly módon programoztak, hogy úgy gondolkodjanak és tanuljanak, mint az emberek.<sup>8</sup> A gépi tanulás, pedig a mesterséges intelligencián belül található.

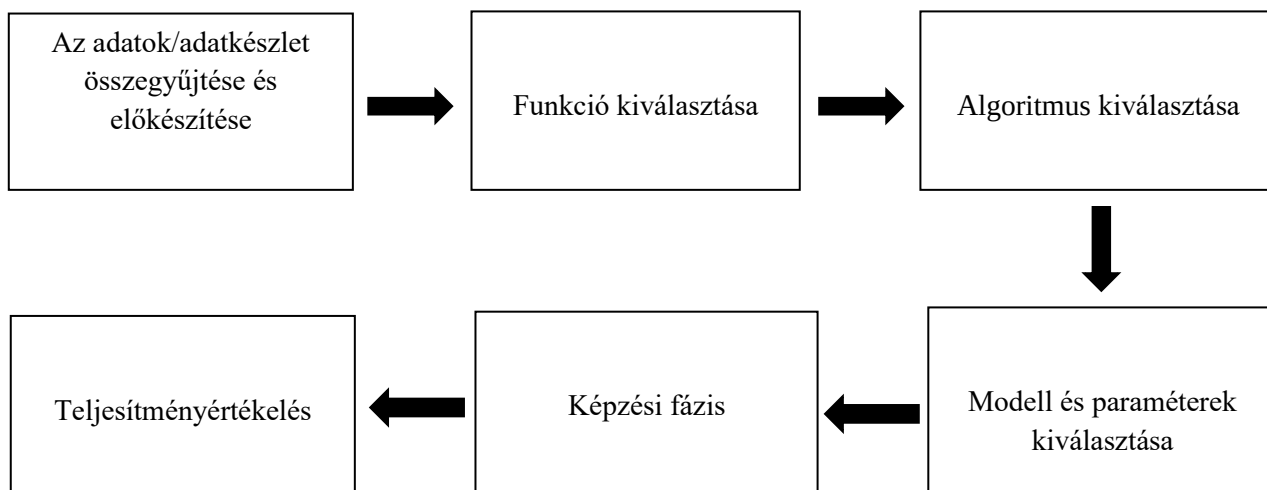
A gépi tanulási algoritmus egy olyan számítási folyamat, amely a bemeneti adatokat felhasználja egy tervezett feladat elvégzésére anélkül, hogy kifejezetten erre programozták volna, vagyis nem „kemény kódolva” van. Mivel ezek az algoritmusok az ismétlés, vagyis tapasztalat eredményeként automatikusan módosítják vagy adaptálják a tervüket, azaz "puhán kódoltak" abban az értelemben, hogy egyre jobban és jobban teljesítik a kitűzött célt. A képzési folyamat során a bemeneti adatok mintáit a kívánt eredményekkel együtt adjuk meg. Az algoritmus ezután a legjobb képességei szerint konfigurálja magát, hogy képes legyen általánosítani, hogy a képzési bemeneti adatok megadásakor a kívánt eredmény előállítása

---

<sup>7</sup> Katerina Sedova, Christine McNeill, Aurora Johnson, Aditi Joshi, and Ido Wulkan, „AI and the Future of Disinformation Campaigns”, é. n., <https://doi.org/10.51593/2021CA011>.

<sup>8</sup> Jerry Kaplan, *Artificial Intelligence: What Everyone Needs to Know* (Oxford University Press, 2016), 1.

mellett a kívánt kimenetet is létrehozza a friss, korábban fel nem tárt adatokból. Ezt a szakaszt nevezzük „tanulási” folyamatnak.<sup>9</sup>

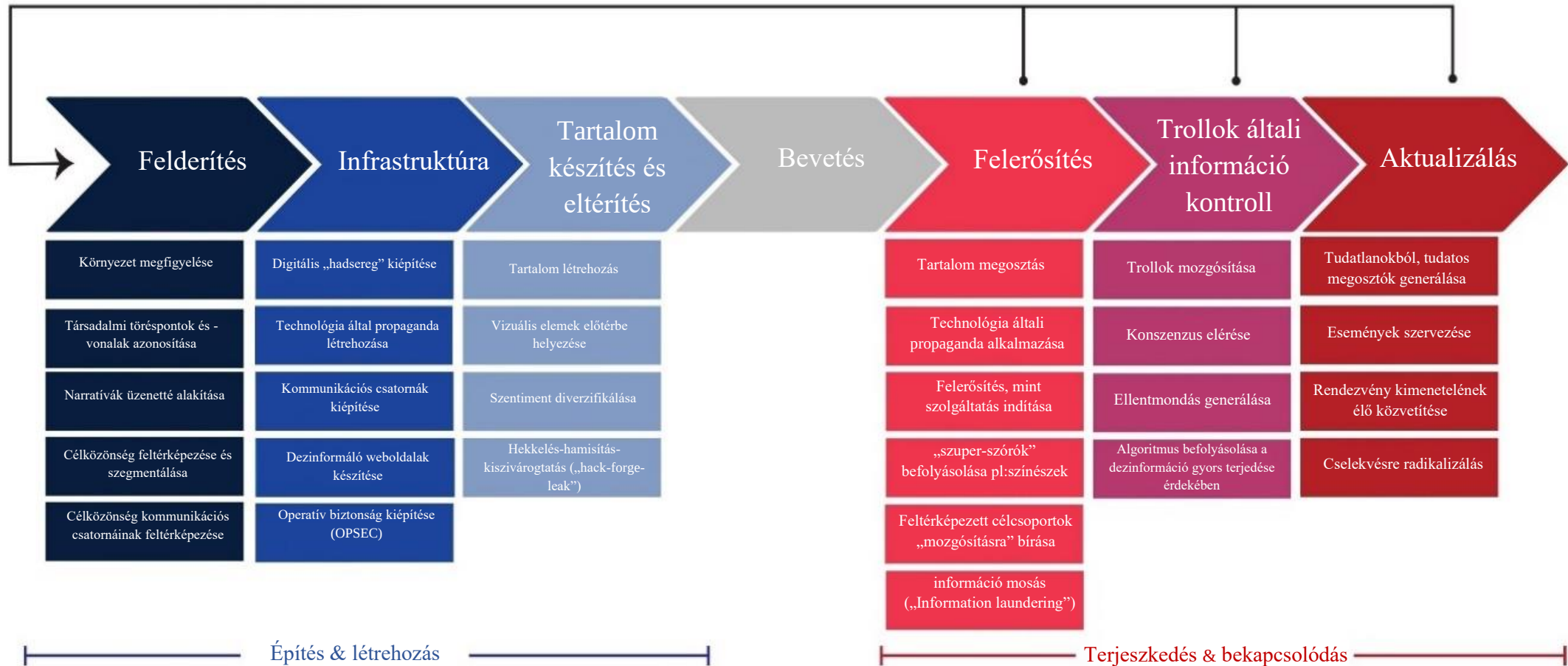


1. ábra Egy általános gépi tanulási modell elemei

(Saját szerkesztés, forrás: Jafar Alzubi, Anand Nayyar, és Akshi Kumar, „Machine Learning from Theory to Algorithms: An Overview”, Journal of Physics: Conference Series 1142, sz. 1 (2018): 012012, <https://doi.org/10.1088/1742-6596/1142/1/012012>.)

<sup>9</sup> Issam El Naqa és Martin J. Murphy, „What Is Machine Learning?”, in *Machine Learning in Radiation Oncology: Theory and Applications*, szerk. Issam El Naqa, Ruijiang Li, és Martin J. Murphy (Cham: Springer International Publishing, 2015), 4, [https://doi.org/10.1007/978-3-319-18305-3\\_1](https://doi.org/10.1007/978-3-319-18305-3_1).

## visszacsatolási kör



2. ábra A dezinformáció fázisai

(Saját szerkesztés; forrás: Katerina Sedova, Christine McNeill, Aurora Johnson, Aditi Joshi, and Ido Wulkan, „AI and the Future of Disinformation Campaigns”, é. n., <https://doi.org/10.51593/2021CA011>.)

## 1.2 A mesterséges intelligencia tanításához szükséges adatgyűjtési fázis

Mielőtt rátérnék kutatásomban a mesterséges intelligencia különböző fenyegető módszereinek felhasználására fontos kiemelni az adatgyűjtési fázist, mivel ez egy nélkülözhetetlen előkészület, ha valaki célzottan akar gépi tanulással működő programokat készíteni. A dezinformatív tartalmakat előállítók legtöbbször a közösségi médiából vagy már meglévő nyíltforrásokból szerzik meg a szükséges adatokat. A közösségi médiaelemzés folyamatának négy lépése van, amelyet a dezinformatív szereplők vagy maguk hajtanak végre, vagy kiszervezik cégek számára. Ez a négy az adatok feltárása, gyűjtése, előkészítése és végül elemzése.<sup>10</sup>

Ahhoz, hogy a gépi tanulással működő programok képesek legyenek jól ellátni munkájukat, hatalmas mennyiségű adatra van szükségük. Ez fáradságos munka lehet, hiszen a „big data” új kihívások elé állítja a dezinformatív tartalmakat előállító személyeket. Négy tényező van, amely ezeket a munkákat megnehezíti. Először is a szükség van tároló helyre, központi adattárra, amely képes a különböző forrásokból származó és különböző formátumú adatokat összegyűjteni. Erre felhőalapú tárolási módokat szoktak javasolni. Második hatalmas problémát az adatok keletkezésének sebessége adja, ráadásul minél naprakészebb az elemzés, annál jobban felhasználható, amely folyamatos munkát igényel. Harmadik az adatok struktúratlansága és változatos formája. Utolsó probléma a bizonytalanság az adott adat minőségében, hiszen a tanulási procedúrában nem csak az adat mennyiség számít, hanem a minőség is nagy szerepet játszik.<sup>11</sup>

A dezinformációt tartalmazó tartalmak előállítói vagy trollok, még nyílt forráskódú eszközök használatával vagy saját eszközök létrehozásával gyűjthetnek adatokat. Az olyan adathalász eszközök, mint a Twint egy nyílt forráskódú eszköz, amely megkerüli a platform azon törekvését, hogy korlátozza az adatokhoz való hozzáférést.<sup>12</sup> A tweetek letöltésére, a letöltés mennyiségére és sebességére vonatkozó korlátozások megkerülésére használható. Ma az adatok

---

<sup>10</sup> Stefan Stieglitz és mtsai., „Social Media Analytics – Challenges in Topic Discovery, Data Collection, and Data Preparation”, *International Journal of Information Management* 39 (2018. április 1.): 156, <https://doi.org/10.1016/j.ijinfomgt.2017.12.002>.

<sup>11</sup> Stieglitz és mtsai., 158.

<sup>12</sup> „TWINT - Twitter Intelligence Tool”, Python (2017; repr., TWINT Project, 2023. április 19.), <https://github.com/twintproject/twint>.

hasznosítása sem egy egyszerű dolog, amely szakembereket igényel, mint például adatelemző tudósokat és a szoftvermérnököket.<sup>13</sup>

Láthatjuk, hogy ez egy időigényes és költséges eljárás, ezért egyes szereplők kiszervezik ezt a folyamatot olyan vállalatoknak, amelyek előfeldolgozási szolgáltatásokat nyújtanak a „big data” számára. A dezinformációs szereplők a következő szintre lépve felhasználhatják ezeket az adatokat a programok képzésére és a dezinformációs kampány támogatására, miután azokat bevitték és feldolgozták.

### **1.3 Szentiment analízis és álláspont vizsgálat (Stance detection)**

Az adatok összegyűjtése után a következő lépés annak eldöntése, hogy milyen programmal fogjuk felhasználni ezeket az adatokat. Először a szentiment analízist fogom szemléltetni, mely a természetes nyelvi feldolgozáson belül található kutatási terület, és azután az álláspont vizsgálatot.<sup>14</sup> A természetes nyelvi feldolgozás és a természetes nyelvi generálás lehetővé teszi a gépi tanulási rendszerek számára a szövegek olvasását, írását és értelmezését, és lehetővé teszi ezáltal szentiment analízist is.<sup>15</sup>

A szentiment analízis az emberek egy entitással kapcsolatos véleményének, attitűdjének és érzelmeinek számítógépes tanulmányozása. Az entitás jelenthet személyeket, eseményeket vagy témákat. Egyes gépi tanulási modellek 94-96%-os pontossággal képesek a felhasználói bejegyzésekből és kommentekből következtetni a polarizáció szintjére. Bár a korai változatok hatékonyságát illetően még mindig létezik némi kritika, a szentiment analízis egyre jobba válik. Legtöbbször marketing célokra használják, mely azt szemlélteti, hogy látható profit szerezhető ennek a technológiának továbbfejlesztésében. A közösségi médiában használt szlenget, hangulatjeleket, emojiakat és rövidítéseket képesek jobban megérteni a fejlettebb gépi tanulási algoritmusok, amelyeket emberi nyelvészek által kiválasztott adathalmazokon képeznek ki. A képességek, amelyet ez a technológiai fejlesztés nyújt a dezinformáció és közösségi média területén megkönnyíti a dezinformációs szereplők munkáját. A szentiment analízis ugyanis

---

<sup>13</sup> „Juršėnas, A., Karlauskas, K., Ledinauskas, E., Maskeliūnas, G., Ruseckas, J., The Double-Edged Sword of AI: Enabler of Disinformation (2021). Riga: NATO Strategic Communications Centre of Excellence.”, é. n., 12.

<sup>14</sup> Walaa Medhat, Ahmed Hassan, és Hoda Korashy, „Sentiment Analysis Algorithms and Applications: A Survey”, *Ain Shams Engineering Journal* 5, sz. 4 (2014. december): 1109, <https://doi.org/10.1016/j.asej.2014.04.011>.

<sup>15</sup> „What Is Natural Language Processing? | IBM”, elérés 2023. január 4., <https://www.ibm.com/topics/natural-language-processing>.

segíthet meghatározni az emberek beállítottságát, például politikailag az általuk követett emberek és közzétett posztjaik alapján.<sup>16</sup>

A szentiment analízisnek három fő osztályozási szintje van:

1. A dokumentumszintű szentiment analízis egy véleményt közlő dokumentumot igyekszik kedvező vagy kedvezőtlen véleményt vagy hozzáállást kifejező kategóriába sorolni. A teljes anyagot egyetlen alapvető információs egységnek tekinti (egy témáról beszél).<sup>17</sup>
2. A mondatszintű szentiment analízis az egyes mondatokban megjelenített hangulatot igyekszik kategorizálni. Meghatározza, hogy a mondat eleve szubjektív vagy objektív. A mondatszintű szentiment analízis megállapítja, hogy a mondat pozitív vagy negatív véleményt képvisel-e, ha szubjektív.<sup>18</sup>
3. Az aspektusszintű szentiment analízis az érzelmek kategorizálására törekszik bizonyos entitások attribútumainak fényében. Az entitások és összetevőik azonosítása a kezdeti szakasz. Ugyanannak a dolognak különböző aspektusairól a véleményt alkotók különböző nézőpontokat fejezhetnek ki. Amikor megjegyezzük, hogy "Ez egy gyors és biztonságos autó, de az üzemanyag-fogyasztása meglehetősen elavult", ez egy illusztráció erre.<sup>19</sup>

Az álláspont vizsgálat (Stance detection) arra törekszik, hogy meghatározza a szöveg szerzőjének hozzáállását egy olyan entitáshoz vagy véleményhez, amelyet a szöveg közvetlenül megszólít vagy sugall. Az álláspont vizsgálat folyamatát a lejjebb látható ábrán szemléltetem. Általában a szentiment analízis egyik részproblémájának tekintik. Ez egy osztályozási probléma, ahol a szöveg szerzőjének álláspontját keressük egy kategóriabesorolás formájában ebből a halmazból: „pártol, ellenez, egyik sem”. Alkalmanként a „semleges” kategóriabesorolás is hozzáadódik az álláspontkategóriák halmazához.<sup>20</sup> Maga az álláspont egy olyan nyilvános

---

<sup>16</sup> Katerina Sedova, Christine McNeill, Aurora Johnson, Aditi Joshi, and Ido Wulkan, „AI and the Future of Disinformation Campaigns”, 17.

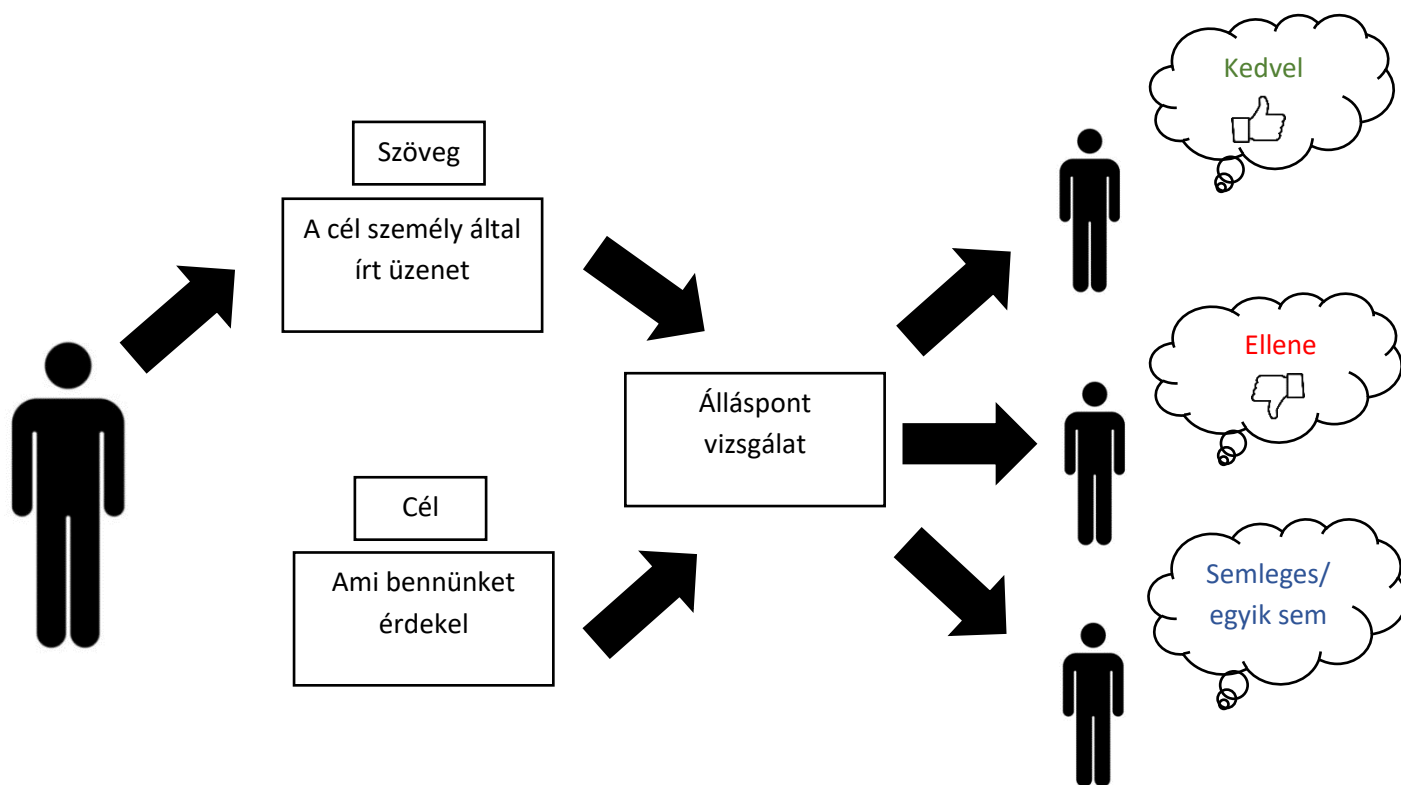
<sup>17</sup> Medhat, Hassan, és Korashy, „Sentiment Analysis Algorithms and Applications”, 1093.

<sup>18</sup> Medhat, Hassan, és Korashy, 1093.

<sup>19</sup> Medhat, Hassan, és Korashy, 1094.

<sup>20</sup> Dilek Küçük és Fazli Can, „Stance Detection: A Survey”, *ACM Computing Surveys* 53, sz. 1 (2021. január 31.): 2, <https://doi.org/10.1145/3369026>.

cselekvés, amelyet dialógikusan, nyílt kommunikációs módszerekkel hajtanak végre adott entitások értékelésére, de ezek az értékelések igazodnak a szociokulturális mező bármely kiemelkedő dimenzióján belül szubjektumokhoz.<sup>21</sup>



3. ábra Álláspont vizsgálat (Stance detection) leegyszerűsített menete

(Saját szerkesztés; forrás: Dilek Küçük és Fazli Can, „Stance Detection: A Survey”, *ACM Computing Surveys* 53, sz. 1 (2021. január 31.): 1–37, <https://doi.org/10.1145/3369026>.)

Az álláspont vizsgálatra legtöbbször a felügyelettel végrehajtott tanulási módszert a support vector machine-t (továbbiakban SVM) alkalmazzák. Az álláspont vizsgálati tanulmány - amiből dolgoztam - 40 olyan tanulmányra hivatkozik, melyben a legjobb pontot elérő megközelítés között volt az SVM.<sup>22</sup> A SVM egy olyan számítógépes algoritmus, amely példák alapján tanulja meg a besorolás hozzárendelését az objektumokhoz. Ezzel a móddal az arcok felismerésétől a hitelkártya csalásokig betaníthatóak a programok.<sup>23</sup> Az álláspont vizsgálatával kapcsolatos kezdeti kutatások célja a résztvevők ideológiai vagy vitás kérdésekkel kapcsolatos álláspontjának azonosítása online vitafórumokon. A rosszindulatú szereplők az elemzések ezen formáinak segítségével azonosíthatják célcsoportjaikat, hogy dezinformálásukkal csak ezeket a

<sup>21</sup> John Du Bois, „The stance triangle”, 2007, 169, <https://doi.org/10.1075/pbns.164.07du>.

<sup>22</sup> Küçük és Can, „Stance Detection”, 17.

<sup>23</sup> William S Noble, „What Is a Support Vector Machine?”, *Nature Biotechnology* 24, sz. 12 (2006. december): 1, <https://doi.org/10.1038/nbt1206-1565>.

személyeket célzhatnak meg. Ilyen vitatott kérdések lehetnek a feminizmus, ateizmus, abortusz vagy éppen politikai párt/személy.<sup>24</sup>

#### **1.4 Generatív ellenséges hálózatok (generative adversarial networks/GANs) használata a dezinformációban**

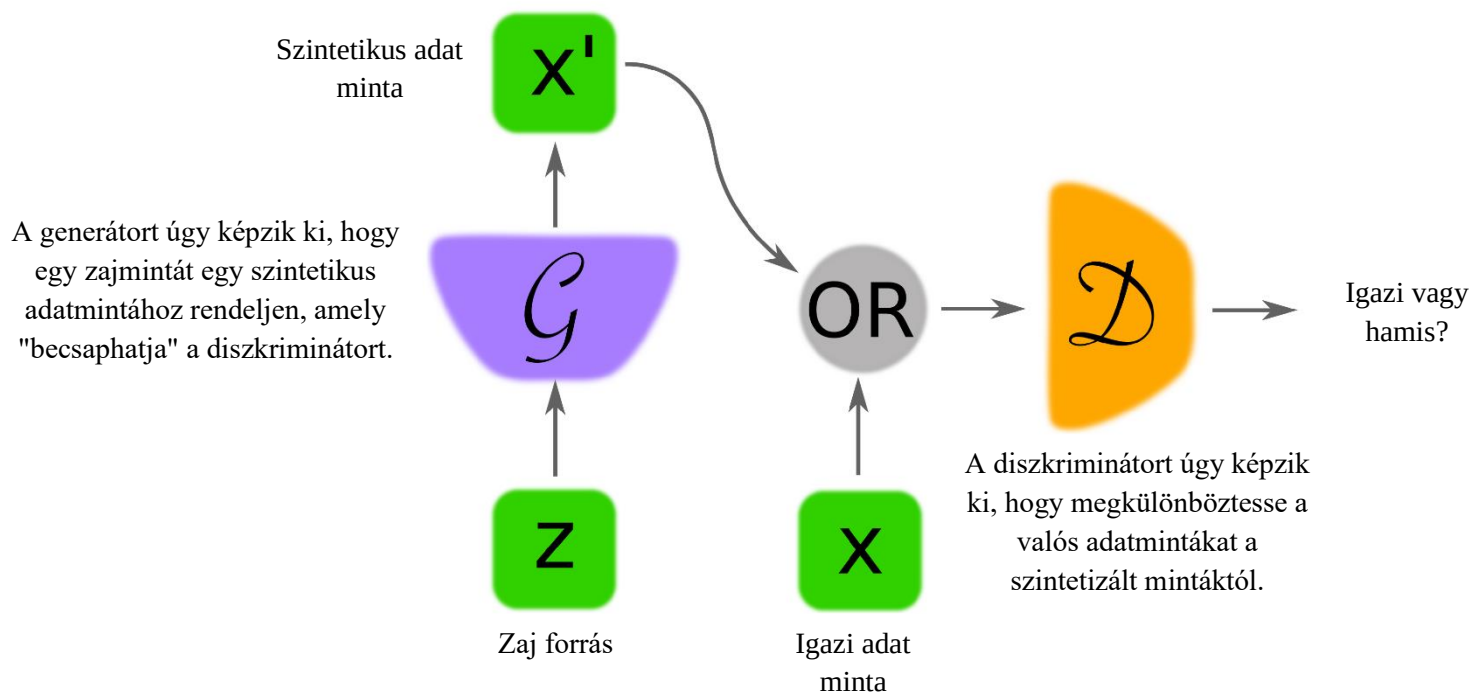
A generatív ellenséges hálózatok (továbbiakban GAN) olyan típusú neurális hálózatok, amelyek két neurális hálózati modellből állnak, egy generátorból és egy diszkriminátorból, amelyeket a GAN arra használ, hogy megkülönböztesse a valódi mintákat a generált mintáktól. Ahogy a neurális hálózatok versenyeznek, tanulnak egymástól, és a versenyfolyamat egyre valóságosabb kimenetet eredményez kép, videó és hang formátumban. Fontos, hogy a generátor csak a diszkriminátorral való interakció révén tanul és nincs közvetlen hozzáférése a valódi képekhez. A szintetikus minták és a valódi képek halmazából vett minták egyaránt a diszkriminátor rendelkezésére állnak. Amikor a diszkriminátor kielemezett egy képet egy mérési alap (ground truth) alapján tudja, hogy a kép valós, vagy a generátortól érkezett. Ez betanításra tökéletes, mivel a kettes számrendszeren alapszik. Valós kép lehet 1-es, a kreált pedig 0-s. A diszkriminátoron keresztül a generátor betanítható, hogy még valóságosabb és jobb minőségű hamisítványokat készítsen.<sup>25</sup> A neurális hálózat egy olyan módszer, amellyel olyan számítógépes programokat hozhatunk létre, amelyek képesek tanulni az információkból. Nagyon lazán arra épül, ahogyan az emberi agy szerintünk működik. Először is, egymáshoz kapcsolódó szoftver "neuronok" hálózatát építjük fel, és lehetővé tesszük, hogy kommunikáljanak egymással. Ezután a hálózatot arra kérjük, hogy oldjon meg egy problémát, amit ismételtén megpróbál, miközben a sikerhez vezető kapcsolatokat erősíti, a kudarchoz vezető kapcsolatokat pedig gyengíti.<sup>26</sup>

---

<sup>24</sup> Küçük és Can, „Stance Detection”, 5.

<sup>25</sup> Ian Goodfellow és mtsai., „Generative adversarial networks”, *Communications of the ACM* 63, sz. 11 (2020. 02. 22.): 2–3, <https://doi.org/10.1145/3422622>.

<sup>26</sup> Su-Hyun Han és mtsai., „Artificial Neural Network: Understanding the Basic Concepts without Mathematics”, *Dementia and Neurocognitive Disorders* 17, sz. 3 (2018. szeptember): 3–4, <https://doi.org/10.12779/dnd.2018.17.3.83>.



4. ábra A GAN képzési folyamata során tanuló két modell a diszkriminátor "D" és a generátor "G"

(Saját szerkesztés; forrás Antonia Creswell és mtsai., „Generative Adversarial Networks: An Overview”, *IEEE Signal Processing Magazine* 35, sz. 1 (2018. január): 53–65, <https://doi.org/10.1109/MSP.2017.2765202..>)

A GAN-ok legfőképpen a szintetikus hanganyagok előállításában, deepfake és videók készítésében kiemelkedőek. Azért kapta a „deepfake” nevet ez a technológia, mivel mélytanulás alapú modell az alapja az ilyen technológiáknak és módszereknek. A mélytanulás a gépi tanulás egy részterülete, mely több rétegű neurális hálózatok képzését foglalja magában. A deepfake arra is használható, hogy arcot cseréljenek egy népszerű vagy ellentmondásos médiaszereplővel vagy politikai szereplővel. Az arccserét általában szintetikus hanggal párosítják, amely utánozza a valódi szereplő hangját, így a dezinformatív tartalmakat előállító személy a deepfake technológia segítségével valósnak tűnő videókat hozhat létre. Ez az eszköz megtévesztésre, félreinformálásra és botrány okozására tökéletes. Az emberek szemében ez növelheti a szkepticizmust és megnehezítheti eldönteni, hogy hamis vagy igaz hírt és videót látnak vagy hallanak.<sup>27</sup> A GAN-okat nem csak a tartalom készítési fázisban lehet használni, hanem az infrastruktúra kidolgozásában is. A dezinformatív üzenetek hatékony terjesztése érdekében tömérdek hamis profilt állítanak elő. Ezekhez a profilokhoz profil képekre van szükség, hogy még hitelesebbek legyenek. Ilyen hihető képet a gépi tanulás segítségével könnyedséggel előállíthatunk. Erre egy jó példa a *thispersondoesnotexist* nevű oldal. Minden

<sup>27</sup> Katerina Sedova, Christine McNeill, Aurora Johnson, Aditi Joshi, and Ido Wulkan, „AI and the Future of Disinformation Campaigns”, 28.

egy alkalommal, amikor a weboldalt frissítjük, egy új arckép jelenik meg előttünk, amelyet a GAN technológia hozott létre.<sup>28</sup> Egy ma már sokkal fejlettebb weboldal, mely mesterséges intelligenciát használ a kép generáláshoz az a Midjourney. Midjourney kulcsszavak megadása alapján kreál valóság-hű képeket. Egy kitűnő példa a hamis hangok generálására az ElevenLabs által nyújtott hang klónozási szolgáltatás. A weboldal ingyenesen elérhető felületén egészen 2500 karakterig adhatjuk meg saját szövegünket és az mp3 hangfájlok feltöltése után a megadott hangon fogja az általunk betáplált szöveget elismételni.<sup>29</sup> Az itt felsorolt példák mind ingyen használatba vehetőek.

Deepfake alkalmazását a „hack-forgé-leak” (meghekkelés-hamisítás-kiszivárogtatás) nevezetű támadásokban hasznosíthatják a támadók. A kiszivárogtatott videóknak vagy hanganyagoknak nem is kell igaznak lenniük vagy éppen a meghekkelt e-mailből származniuk, hanem csak a média társaságokkal és a közönséggel kell elhitetni, hogy az a forrás. Ilyenkor a médiában ambivalens érzések keletkezhetnek, hogy közzé tegyék-e a hírt vagy sem. Az ilyen fajta hírek, - mivel általában figyelmet keltenek és botrányosak - profitot és nézettséget hozhatnak a média cégeknek, de ha később kiderül, hogy hamis hírt adtak le hírnevük nagy mértékben romolhat. Az ilyen fajta hírek választásoknál jelenthetnek nagy problémát, mivel, ha a hírt már egyszer közzétették és még ha a reagálás gyors is, akkor ez már eljutott az emberekhez és befolyásolhatta őket döntéseikben, állásfoglalásaikban.<sup>30</sup>

A deepfake terjedése, mint úgyszintén a dezinformáció állandóvá válása, az általános bizalom erodálását hozhatja a hírekben, amely olyan általános dolgok megkérdőjelezéséhez vezethet, amelyek veszélyesek lehetnek magára a szkeptikus emberre és más embertársaikra. Erre egy jó példa a COVID-19 járvány alatti oltásellenesség.

---

<sup>28</sup> Katerina Sedova, Christine McNeill, Aurora Johnson, Aditi Joshi, and Ido Wulkan, 20.

<sup>29</sup> „ElevenLabs - Prime AI Text to Speech | Voice Cloning”, elérés 2023. április 19., <https://beta.elevenlabs.io/>.

<sup>30</sup> Katerina Sedova, Christine McNeill, Aurora Johnson, Aditi Joshi, and Ido Wulkan, „AI and the Future of Disinformation Campaigns”, 57.

## 1.5 Tartalomkészítés természetes nyelvi generálással (NLG)

A természetes nyelvi generálás (továbbiakban NLG) alatt a szöveg vagy beszéd nem nyelvi bemenetéből történő generálásának feladatát értjük. Az NLG a nyelvi bemenetből előállított szövegre vagy beszédre is vonatkozhat, mint például összegzés vagy szövegegyszerűsítés.<sup>31</sup> Ennek a szöveg generálási módszernek a 6 lépését a következő táblázatban szemléltetem.

A tartalomkészítési fázison belül ez hozta a legnagyobb változást a dezinformációs kampányok számára. Ez a technológia képes meghatározott paraméterek alapján szövegeket létrehozni, melyek hihetőnek és egyedinek hangzanak, így megkönnyítve a dezinformáció írásának folyamatát kampányok számára. Ez a technológia képes meghatározott paraméterek alapján szövegeket létrehozni, melyek hihetőnek és egyedinek hangzanak, így megkönnyítve a dezinformáció írásának folyamatát a kampányok számára.

1. táblázat A természetes nyelvi generálás 6 lépése

(Saját szerkesztés; forrás: „What Is Natural Language Generation?”, Enterprise AI, elérés 2023. április 14., <https://www.techtarget.com/searchenterpriseai/definition/natural-language-generation-NLG.>)

| Tartalomelemzés                                     | Adatok megértése                                                     | Dokumentumok strukturálása                       | Mondatok összesítése                                          | Nyelvtani strukturálás                                                              | Nyelv ismertetése                     |
|-----------------------------------------------------|----------------------------------------------------------------------|--------------------------------------------------|---------------------------------------------------------------|-------------------------------------------------------------------------------------|---------------------------------------|
| az adatok szűrése és a fontos tartalmak azonosítása | az adatok értelmezése, a minták azonosítása és kontextusba helyezése | dokumentumterv és narratív struktúra létrehozása | releváns mondatok kombinálása a téma összefoglalása érdekében | nyelvtani szabályok alkalmazása a természetes hangzású szöveg létrehozása érdekében | a végső input formázása és bemutatása |

Egy „Journal of Experimental Political Science” folyóiratban publikált kutatásban a kutatók a nyelvi modellek társadalmi hatásainak két különböző kísérleti módon történő vizsgálatát végezték el. A kutatás során mesterséges intelligencia által írott híreket használtak arra, hogy megvizsgálják azok hitelességét és azt, hogy képesek-e megváltoztatni a külpolitikával kapcsolatos sztereotípiákat és előítéleteket. Ezt a kutatást az OpenAI által fejlesztett GPT-2

<sup>31</sup> Chris van der Lee és mtsai., „Human Evaluation of Automatically Generated Text: Current Trends and Best Practice Guidelines”, *Computer Speech & Language* 67 (2021. május 1.): 2, <https://doi.org/10.1016/j.csl.2020.101151>.

(Generative Pre-trained Transformer) három különböző verziójával készítették, amelyek a közepes nagyságú 355 milliós, a 724 milliós nagy és a 1,5 milliárdos extra-nagy paraméterekre épülő modellek voltak. A GPT modell a természetes nyelv feldolgozásban használt neurális hálózat architektúrájának egyik típusa. Ez azt jelenti, hogy nagy mennyiségű szöveges adaton, felügyelet nélküli tanítási technikát alkalmaznak. Itt fontos megemlíteni, hogy az OpenAI ma mára GPT-5-öt fejleszt, ha csak a GPT-2 és GPT-3 közti paraméter szám különbségét nézzük, akkor nagy mértékű fejlődést tapasztalhatunk. Hiszen a GPT-2 csak 1,5 milliárd paraméterrel rendelkezik, de a GPT-3 már 175 milliárddal ez 116 szorosa a GPT-2 kapacitásának. Ezek a modellek hatalmas hasznára válhatnak az emberiség számára, de veszélyeket is hordozhatnak magukban. A kutatók az első kísérletben a mesterséges intelligencia által generált szövegek emberi hitelesség érzékelésének felső határát vizsgálták, valós hírekkel összehasonlítva. A Második kísérlet párhovatarozás hatását vizsgálta a mesterséges intelligencia által generált hírek hitelességére. Ez vitákat váltott ki arról, hogy az emberek, akik politikailag egy párt mellett állnak nagyobb valószínűséggel hisznek-e a politikailag kedvező híreknek, és nagyobb valószínűséggel ragaszkodnak-e ezekhez a véleményekhez, még úgyis, ha fel van tüntetve, hogy az a szöveg szintetikus lett generálva. A harmadik kísérletben három különböző méretű mesterséges intelligencia szöveggeneráló modell képességét vizsgálták, hogy ezek mennyire képesek emberi segítség nélkül automatikusan híreket létrehozni és azok mennyire hitelesek.<sup>32</sup>

Az első kísérlet azt az eredményt hozta, hogy az olvasók a mesterséges intelligencia által generált, de emberek által szerkesztett szöveget ugyanolyan hitelesnek vélték, mint egy ugyanarról az eseményről szóló, eredeti és emberi kézzel írott cikket. A második kísérletben arra jutottak a kutatók, hogy egyértelműen a pártállás jelentős hatással volt arra, hogy az emberek mennyire ítélik hitelesnek az adott hírt. Ez arra a következtetésre juttatta a kutatókat, hogy ez nagy veszélyt jelenthet, ha specifikusan célozzák ezeket a közegeket a dezinformálók. Utoljára véleményem szerint a legjelentősebb eredmény, amire a kutatók jutottak, az a harmadik kísérletből származik. A szöveggeneráló modellek képesek voltak emberi beavatkozás nélkül is hitelesnek tűnő hírcikket alkotni, de ezek nyilván nem voltak olyan sikeresek, mint az emberek által utólag szerkesztett változatok.<sup>33</sup>

---

<sup>32</sup> Sarah Kreps, R. Miles McCain, és Miles Brundage, „All the News That’s Fit to Fabricate: AI-Generated Text as a Tool of Media Misinformation”, *Journal of Experimental Political Science* 9, sz. 1 (2022. ed): 3–4, <https://doi.org/10.1017/XPS.2020.37>.

<sup>33</sup> Kreps, McCain, és Brundage, 4.

Ezen szöveggeneráló rendszerek képességei lehetővé tehetik a célzott és pontos tartalomfejlesztést, mivel a generatív nyelvi rendszerek példák alapján képesek írásmódokat utánozni. Ez a már előbb említett hack-forgé-leak támadásnál előny lehet. Hiteles hamis leveleket, vagy e-maileket gyárthat, mely kompromittálja a célzott közszereplőt. Ezeket a hamisítványokat a meghekkelt e-mailek közé feltölthetik és közzétehetik, hogy minél hihetőbb legyen a lejárató kampány.<sup>34</sup>

## 1.6 Közösségi botok (social bots) használata dezinformálásra

A közösségi bot egy olyan számítógépes program, amely automatikusan készít tartalmakat, és kapcsolatba lép a felhasználókkal a közösségi médiában, hogy utánozza és esetleg megváltoztassa a viselkedésüket.<sup>35</sup> Ezeknek a közösségi botoknak a célja lehet egyes profilok, csoportok közönségének mesterséges növelése, illetve egy vállalat vagy egy nyilvánosan ismert személyiség hírnevének rombolása.

A múltban a botok főként egyfajta tevékenységet végeztek, mégpedig a tartalmak automatikus közzétételét. Ezek a korai botok nem voltak annyira fejlettek, így a kiszűrésük sokkal kevésbé volt bonyolult feladat, mint napjainkban. Manapság a közösségi botok sokkal kifinomultabbak. A botok képesek információt keresni a weben, amivel feltölthetik a profiljukat. Emellett képesek úgy is megosztani tartalmakat, hogy ne keltsenek gyanút a közösségi botokat észlelő rendszerekben, például előre meghatározott időpontokban posztolnak, így utánozva a tartalom előállításának emberi időbeli jellegzetességeit. Ezek a botok képesek beszivárogni a vitákba is, és természetes nyelvi algoritmusok segítségével a témához illő válaszokat produkálni. Ezekkel a válaszokkal a bot profilja új követőkre tehet szert, vagy tartalmait valódi emberek is megoszthatják. A botok a korábban említett szentiment analízissel elemezhetik a szöveget, ezzel segítve a válaszok létrehozását.<sup>36</sup>

Egy kiváló ingyen elérhető példa a Prank Me Not weboldal.<sup>37</sup> Ezen a weboldalon hamis tweetek vagy facebook csevegések állíthatók elő, amelyeket a közösségi botok felhasználhatnak hamis információk terjesztésére. Egy általam kreált példát illesztettem a bekezdés alá, amely az

---

<sup>34</sup> Katerina Sedova, Christine McNeill, Aurora Johnson, Aditi Joshi, and Ido Wulkan, „AI and the Future of Disinformation Campaigns”, 25.

<sup>35</sup> Emilio Ferrara és mtsai., „The Rise of Social Bots”, *Communications of the ACM* 59, sz. 7 (2016. június 24.): 96, <https://doi.org/10.1145/2818717>.

<sup>36</sup> Ferrara és mtsai., 99.

<sup>37</sup> „Fake Tweet, Chat & Facebook Status Generator | Prank Me Not”, elérés 2023. április 19., <https://www.prankmenot.com/>.

Amerikai Egyesült Államok elnökének tweetjét mutatja, melyben az amerikai szárazföldi egységek bevetését jelenti Ukrajnába. A weboldal használatához minimális angol tudás szükséges, de más szakértelmet nem igényel különösebben.



5. ábra Prank Me Not weboldal segítségével készített hamis tweet (Saját szerkesztés)

A clearneten és a darkneten vannak olyan piacok, ahol az internetezők igénybe vehetik a közösségi botok által nyújtott szolgáltatásokat egy megszabott ár fejében. A clearnet alatt a nyilvánosan elérhető szolgáltatásokat értjük. Darknet alatt pedig azt értjük mikor egy specifikus keresőmotort használ a felhasználó és a személyazonossága rejtve marad. Ezek a szolgáltatások legáltalánosabb esetben kedveléseket, megosztásokat és kommenteléseket biztosítanak. Ezek a közösségi botok túlterheléses támadásokban és phishingben is előszeretettel használt eszközök. Végző soron a közösségi botok jelentős szerepet játszanak a dezinformációs kampányokban, mivel a már rendelkezésre álló építőelemeket, például a szövegelemzéshez szükséges hangulelemzést és a GAN-rendszert felhasználhatják egy hihető profil kialakításához.<sup>38</sup>

## 1.7 A dezinformáció automatizálása chatbotok segítségével

A chatbot egy olyan számítógépes szoftver, amely utánozza és értelmezi az emberi (beszélt vagy gépelt) interakciót, lehetővé téve a felhasználók számára, hogy úgy beszélgessenek a digitális eszközökkel, mintha valódi emberekkel beszélgetnének. A chatbotok lehetnek olyan egyszerűek, mint kérdésekre válaszoló egysoros programok, de lehetnek olyan összetettek is,

<sup>38</sup> Dennis Assenmacher és mtsai., „Demystifying Social Bots: On the Intelligence of Automated Social Media Actors”, *Social Media + Society* 6, sz. 3 (2020. július): 5, <https://doi.org/10.1177/2056305120939264>.

mint a digitális asszisztensek, amelyek idővel tanulnak és fejlődnek, hogy egyre személyre szabottabb szolgáltatást nyújtsanak, ahogy egyre több adatot gyűjtenek és elemeznek. A fejlettebb botok már személyre szabhatóak és beállíthatóak, hogy hogyan és miként közöljék a válaszukat.<sup>39</sup> A chatbotok lehengerlő fejlesztésére jó példák lehetnek a nemrégiben OpenAI által kifejlesztett és megjelentetett ChatGPT vagy a Google által létrehozott Bard.

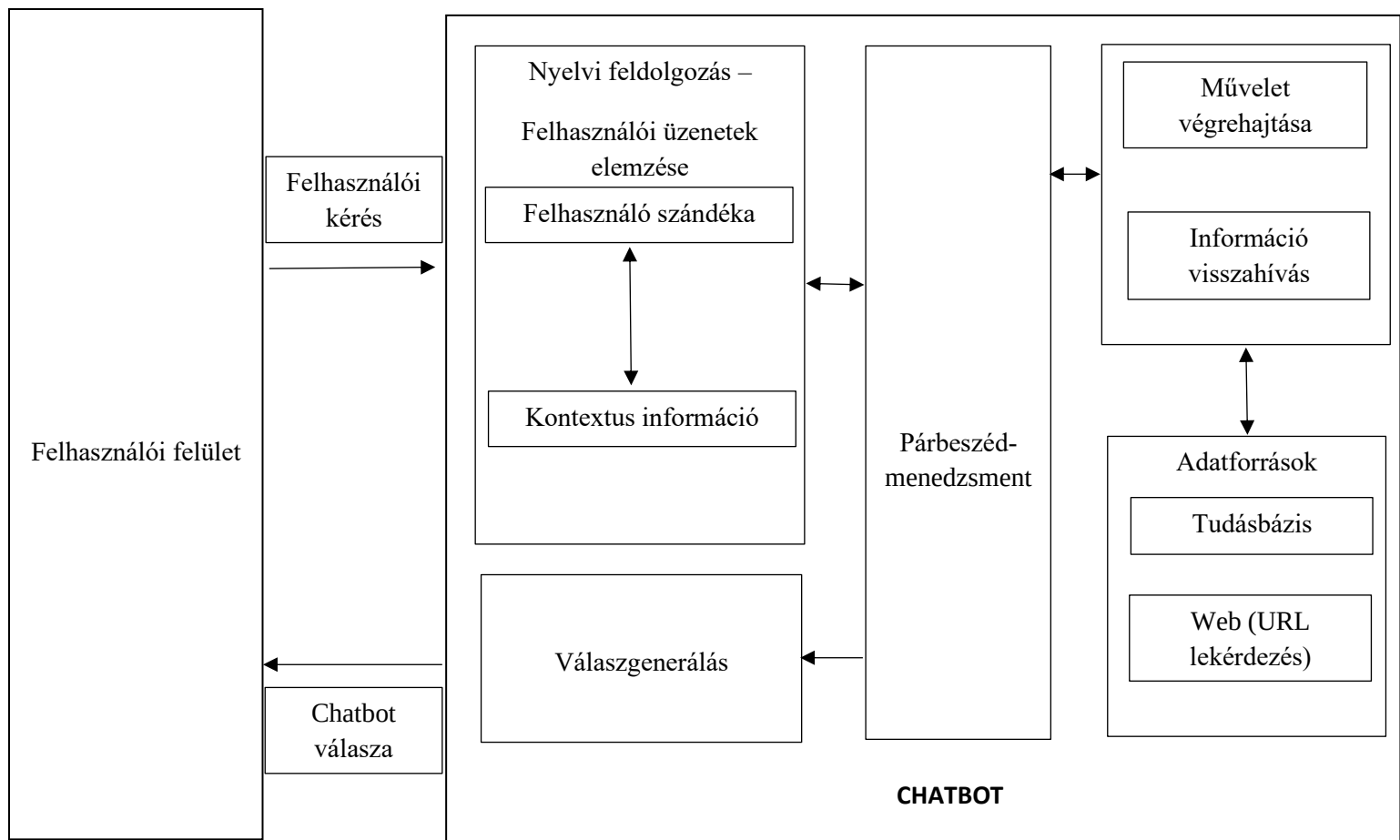
A dezinformációban többnyire az adatvezérelt és a prediktív chatbotokat használják, mert ezek fejlettebb dolgokra képesek, mint egy feladatorientált. Ezek a chatbotok kontextustudatosak, és a természetes nyelvi megértést (NLU), a természetes nyelvi feldolgozást (továbbiakban NLP) és a gépi tanulást kihasználva tanulnak menet közben. Prediktív intelligenciát és analitikát alkalmaznak, hogy lehetővé tegyék a felhasználói profilok és a korábbi felhasználói viselkedés alapján történő személyre szabást. Ezeknek a képességeknek a fejlesztése valamilyen módon könnyebbé vált, mivel nyílt forráskódúak. Példa erre a Facebook (ma már Meta) és általuk 2020-ban kiadott BlenderBot teljes modellje, kódja és kiértékelési beállításai nyílt forrásúvá tétele.<sup>40</sup> A nyílt forráskódúvá tétel mögött az innováció és a kutatás továbbfejlesztése is áll. Mivel ezek a dolgok könnyen elérhetővé válnak, a rosszindulatú szereplők ugyanolyan könnyen bevonhatják ezeket a képességeket a saját kampányaikba. A chatbotok többfordulós beszélgetéseket folytathatnának a hozzászólások alatti kommentekben, miközben a természetes nyelvi generáló rendszerek logikus, egy érvet előremutató bekezdéseket írnak. Eközben a gépi tanulással működő szentiment analízis értékeli az üzeneteik hatását, és optimalizálja a hozzászólásokhoz való bevetést. Még ha a rosszindulatú szereplők nem is automatizálják teljesen a folyamatot, egy ember-gép csapat csökkentheti a munkaigényes folyamatot, és jobban segíthet a hiteles felhasználó megszemélyesítésében.<sup>41</sup>

---

<sup>39</sup> „What Is a Chatbot?”, elérés 2023. január 4., <https://www.oracle.com/chatbots/what-is-a-chatbot/>. NEM A VÉGLEGES

<sup>40</sup> Katerina Sedova, Christine McNeill, Aurora Johnson, Aditi Joshi, and Ido Wulkan, „AI and the Future of Disinformation Campaigns”, 32.

<sup>41</sup> Katerina Sedova, Christine McNeill, Aurora Johnson, Aditi Joshi, and Ido Wulkan, 38.



6. ábra Egy általános chatbot felépítése

(Saját szerkesztés; forrás: Eleni Adamopoulou és Lefteris Moussiades, „An Overview of Chatbot Technology”, *Artificial Intelligence Applications and Innovations* 584 (2020. május 6.): 373–83, [https://doi.org/10.1007/978-3-030-49186-4\\_31](https://doi.org/10.1007/978-3-030-49186-4_31)).

Egy kifinomult, nagy nyelvi modelleket használó chatbot több platformon is meggyőző személyiség lehet, kevés emberi segítséggel. A dezinformáló szereplők finomhangolással megtaníthatnak egy chatbotot bizonyos trollkodási módszerekre, például provokációra, bosszantásra vagy, hogy social engineeringre specializálódjon. A chatbotok válaszok széles skáláját nyújthatnák, amelyek közül az emberi kezelő választhatna, válogathatna a kimenetből és közzétehetné a választ, ha a gépi tanulási rendszerek képesek lennének "megérteni" az online interakciókat és természetes módon tudnának válaszolni. Ez megkönnyítené egy emberi, különösen egy más nyelvet beszélő operátor számára a kommunikációs folyamatok kezelését. A chatbot-trollokat egyre nehezebb lesz azonosítani a valódi emberektől, függetlenül attól, hogy interakciót próbálnak-e terjeszteni, egy bizonyos álláspontot képviselnek-e, vagy az ellenkező véleményt valló személyeket veszik célba. Az exponenciális méretnövekedés révén az ember-gép együttműködés csökkentheti az egyetlen emberi kezelő munkaigényes eljárását,

segíthet a valódi felhasználók pontosabb megszemélyesítésében, és több embert vonhat be az online interakciókba.<sup>42</sup>

## 1.8 Következtetések

Kutatásom első fejezetében jól láthatóvá vált, hogy a gépi tanulás és a mesterséges intelligencia számos fejlesztését és eszközét használhatják dezinformációs kampányokra. Ezek nem is főként a teljes automatizálását hozzák el a dezinformáció terjesztésére, de - a 2. ábrán látható fázisok közül - mindegyikében segítő hatással lehet a mesterséges intelligencia és a gépi tanulás. Sokszor azt is leszűrhetjük, hogy ezekhez a képességekhez nem is kell hatalmas tudással rendelkezni, mivel sokszor a már megalkotott eszközök nyíltforrásból beszerezhetők és csak a használatukat kell elsajátítani.

Összességében a mesterséges intelligencia az egy folyamatosan fejlődő tudományterület, és még mindig nem tudjuk, hová vezethet minket a jövőben. Ezért kell figyelemmel lenni, hogyha ezeket az eszközöket rosszindulattal használják, akkor legyen egy mesterséges intelligencia által vezérelt alternatíva, ami ezeket a folyamatokat képes kivédeni. Kutatásom következő fejezetében ezt kívánom kifejteni.

---

<sup>42</sup> Katerina Sedova, Christine McNeill, Aurora Johnson, Aditi Joshi, and Ido Wulkan, 39–40.

## 2 A gépi tanulás és a mesterséges intelligencia szerepe a dezinformáció elleni küzdelemben

A gépi tanulást és a mesterséges intelligencia fejlődését legjobban talán egy kétélű kardként lehetne leírni. Az első fejezetben a „káros” hatásairól írtam, azonban a második fejezetben azt fogom kifejteni, hogy hogyan használható a gépi tanulás és a mesterséges intelligencia a dezinformáció felismerésére és terjesztésének gátlására vagy fékezésére. Az információk megbízhatóságának a vizsgálata az mindig is egy nehéz és komplikált feladat volt, melyben az emberek nem minden esetben voltak sikeresek. Ez arra készítette a tudósokat, hogy ezt a procedúrát automatizálják gépi tanulással, neurális hálózatokkal és mély tanulási rendszerekkel.

Ebben a fejezetben főként a hamis információt tartalmazó szövegek elemzését és felismerését, a deepfake felismerését hang és videó formátumban és a közösségi botok detektálását vizsgálom, valamint, hogy ebben a gépi tanulási modellek mennyit tudnak segíteni.

### 2.1 A dezinformáció szövegelemzés alapú felismerése

Talán a legalapvetőbb módszere annak, hogy a dezinformációt felismerjük, ha a szövegelemzést, az NLP-t használjuk fel. A jellemzőknek és adatoknak közvetlen a vizsgált cikk tartalmából történő kinyerése szolgál a mintafelismerő rendszer felépítéséhez. A szöveg nyelvi kontextus nélküli elemzésében a kutatók leggyakrabban a szóhalmazt (bag-of-words) és N-gramm modelleket használnak. A szóhalmaz módszer a kifejezések előfordulását számolja a szövegben, az N-grammok pedig megpróbálják előre jelezni a következő szót. Itt hangsúlyozni kell, hogy az N-grammok száma természetesen függ a dokumentum hosszától is, ezért a tanulási folyamaton belül ezt is figyelembe kell venni.<sup>43</sup>

A szövegelemzés metódusokon belül a szövegosztályzásról is beszélhetünk, amely egy mesterséges intelligenciamodell természetes nyelvi feldolgozás segítségével, betanítás után osztályozza a szövegeket, kiszámítva annak a valószínűségét, hogy az adott szöveg dezinformáció-e vagy sem. Ezt úgy teszi, hogy egy előre meghatározott kategóriába helyezi a szövegrészeket. A mesterséges intelligencia a szövegelemzésben akadályokba ütközhet, mivel

---

<sup>43</sup> Michał Choraś és mtsai., „Advanced Machine Learning Techniques for Fake News (Online Disinformation) Detection: A Systematic Mapping Study”, *Applied Soft Computing* 101 (2021. március 1.): 6, <https://doi.org/10.1016/j.asoc.2020.107050>.

eredendően a szövegbe keveredhet az igazságot elhomályosító félrevezetés vagy hibás adat, így megnehezítve a döntést a numerikus két számrendszerben gondolkodó gép számára, hogy állírn-e az adott cikk vagy sem.<sup>44</sup> Egy másik nehézség, ha a hírek tartalmának, vagyis a szövegben szereplő információk, valamint a videók, fényképek és egyéb, a hírt alkotó elemek pontosságának ellenőrzésénél csak a szöveget használják fel a dezinformáció felismeréséhez.<sup>45</sup>

Az első ilyen szövegosztályzási módszer, amiről beszélhetünk, az a dezinformáció érvelésen alapuló felderítése. Ez azt jelenti, hogy az ismert tényekkel összehasonlítja a szöveget, és ha ezzel ellentmondásos információt talál, jelzi ezt a kutatónak. Tekintettel arra, hogy a most elérhető szöveg generáló mesterséges intelligenciamodellek még nem teljesen tökéletesek hosszú szövegeket alkotni ellentmondásosságok kreálása nélkül, ezért ezeket a szövegeket megvizsgálhatjuk következtetési algoritmusok segítségével. Az ilyen fajta rendszerek számára komplikációt jelent a megfelelő adatbázisok létrehozása. Ez a folyamat könnyíthető, ha a vizsgált területet leszűkítjük egy adott politikai párt, vagy akár orvosi területre, mint a COVID-19.<sup>46</sup>

A COVID-19-cel kapcsolatos tanulmányában Adrian Groza egy érvelésen alapuló dezinformáció-felismerő rendszert hozott létre a természetes nyelvi feldolgozás és ontológiai következtetés segítségével. A természetes nyelvi feldolgozás segít a rendszernek a szövegeket a gép által ismert logikai kijelentésekké formálni. Ennek a rendszernek a lényege, hogy megkülönböztesse a megbízható és a nem megbízható orvosi forrásokat.<sup>47</sup> Az ilyen logikai érvelésen alapuló rendszer egy másik problémája az új, kitalált eseményekkel magyarázott dezinformáció. A vitatott állításokat a rendszer már korábbi megerősített tényekre alapozva dönti el, hogy az állítás logikailag helyes-e vagy sem. Ez a rendszer így önmagában nem lenne képes például új, kitalált események vizsgálatára. Amire viszont képes, például az előbb említett COVID-19 alapú dezinformáció felismerése, mivel ez egy új vírus, de a róla szóló

---

<sup>44</sup> „Juršėnas, A., Karlauskas, K., Ledinauskas, E., Maskeliūnas, G., Rondonas, D., Ruseckas, J. The Role of AI in the Battle Against Disinformation (2022). Riga: NATO Strategic Communications Centre of Excellence”, é. n., 9.

<sup>45</sup> Uğur Mertoğlu, Burak Genç, és Hayri Sever, „Text-Based Fake News Detection via Machine Learning”, in *Trends in Data Engineering Methods for Intelligent Systems*, szerk. Jude Hemanth és mtsai., Lecture Notes on Data Engineering and Communications Technologies (Cham: Springer International Publishing, 2021), 122, [https://doi.org/10.1007/978-3-030-79357-9\\_12](https://doi.org/10.1007/978-3-030-79357-9_12).

<sup>46</sup> „Juršėnas, A., Karlauskas, K., Ledinauskas, E., Maskeliūnas, G., Rondonas, D., Ruseckas, J. The Role of AI in the Battle Against Disinformation (2022). Riga: NATO Strategic Communications Centre of Excellence”, 10.

<sup>47</sup> Adrian Groza, „Detecting fake news for the new coronavirus by reasoning on the Covid-19 ontology” (arXiv, 2020. április 26.), 1, <http://arxiv.org/abs/2004.12330>.

állításoknak össze kell függeniük és igaznak kell lenniük, amit a vírusokról általában tudunk már.<sup>48</sup>

A szövegelemzésen belül egy másik módszer a tudásbázis alapú dezinformáció felismerés. Ez a módszer megerősített cikkek és információk tudásbázisán alapszik, és az álláspont vizsgálat képességeit is felhasználja.

Az "Association for Information Science and Technology" folyóirat egyik kutatásában az álhírek detektálását egy osztályozó rendszer segítségével vizsgálták a kutatók. Ehhez már meglévő nyíltforrású adathalmazokat használtak fel, amelyek hamis és valódi híreket tartalmaztak. A kutatásban két almodulon volt a hangsúly, az egyik az hitelesség felismerő almodul (veracity detection submodule) és a másik, a stílusalapú (style based) almodul. A hitelesség felismerő almodul arra szolgál, hogy meghatározza az információ igaz vagy hamis voltát. A stílus alapú almodul viszont a szöveg stílusára és hangnemére koncentrál. Hasonló technikákat alkalmaz, mint a szentiment analízis, iróniát és szarkazmust is fel tud ismerni. Ezeket használja szöveg hangnemének kiértékelésére. A megállapításokat súlyozott átlagot alkalmazó szavazási módszerrel kombinálja, ahol a súlyokat szintén egy neurális hálózat tanulja meg. A főmodul további almodulokkal bővíthető, melyek metaadatokból állhatnak. Az álláspont vizsgálatnál itt három besorolást kaphatnak a dokumentumok: hamis, gyanús vagy igazi. A kutatók azt is megfigyelték, hogy az álhírek szókincse és hangvétele általában agresszívebb, főként az erős érzelmek kiváltására alkalmasak, valamint előítéleteket közvetítenek. A kutatásban a stílusalapú almodell volt sikeresebb az álhírek felismerésében. A modellnek 80%-os sikeressége volt.<sup>49</sup>

---

<sup>48</sup> „Juršėnas, A., Karlauskas, K., Ledinauskas, E., Maskeliūnas, G., Rondonas, D., Ruseckas, J. The Role of AI in the Battle Against Disinformation (2022). Riga: NATO Strategic Communications Centre of Excellence”, 11.

<sup>49</sup> „Towards automatic fake news classification - Ghosh - 2018 - Proceedings of the Association for Information Science and Technology - Wiley Online Library”, 1–2, elérés 2023. január 24., <https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/pra2.2018.14505501125>.

## 2.2 Deepfake felismerése gépi tanulási modellek által

Ezen a kategórián belül két lényeges részből lehet szót ejteni, ezek szintetikus hanganyagok és a deepfake technológiával manipulált videók. Ezt a kettőt kombinálni szokták, hogy a dezinformátorok még nagyobb hatást érjenek el az átadott üzenetükkel.

Először is a szintetikus hanganyagok észlelésére térek ki. A „hangcserénél” hangkonverziót használnak, amely lehetővé teszi a beszélő hangjának módosítását, hogy úgy hangozzon, mint a célszemély, miközben megőrzi az eredeti beszéd tartalmát. A legegyszerűbb módja az ilyen hanganyagok kiszűrésére, hogy a hangfelvételt egy spektogrammá alakítjuk a különböző hangfrekvenciák időbeli eloszlásának kiszámításával. A hangfelvételtől így egy Mel-spektogrammnak nevezett 2D-s grafikon készül, amely az egyik tengelyen az időt, a másikon pedig a hangfrekvenciát ábrázolja. Míg a fül számára a valódi és szintetikus hang szinte hasonló, a spektogrammokon különbözően néznek ki ezek a hangok. Ezáltal a hangfelismerési probléma átalakul egy képosztályozási problémává, amelyben a mesterséges intelligencia kiemelkedő képességekkel rendelkezik és ezt a problémát egy mély tanulási modell megoldhatja.<sup>50</sup>

Az egyik legnagyobb kihívás az ilyen hanganyag felismerésével az az általánosítási (generalization) képesség. Ez azt jelenti, hogy a rendszer képes-e a betanított példaktól eltérő anyagokat is észlelni. Ez azért is fontos terület, hiszen a támadók folyamatosan új módszereket fejlesztenek ki a felismerés kijátszására. Ezt a problémát a kutatók az adatkészletek továbbfejlesztésével, mély tanulási rendszerek bevonásával és transzfer tanulással korrigálják.<sup>51</sup> A transzfer tanulás egy olyan technika, ahol egy feladatra betanított modellt egy második, kapcsolódó feladat modelljének kiindulópontjaként használnak. A transzfer tanulás lényege, hogy az egyik feladatra betanított modell gyorsabban és jobb teljesítménnyel használható egy másik, kapcsolódó feladat megtanulására, mintha a modellt a nulláról kellene betanítani. Például egy nem hamisított példákból álló nagy adathalmazon betanított modell kiindulópontként használható egy hamisítás-felismerő modellhez. Az előre betanított modell ezután finomhangolható egy kisebb, hamisított példákból álló adathalmazon. Ez lehetővé teheti,

---

<sup>50</sup> „Juršėnas, A., Karlauskas, K., Ledinauskas, E., Maskeliūnas, G., Rondonas, D., Ruseckas, J. The Role of AI in the Battle Against Disinformation (2022). Riga: NATO Strategic Communications Centre of Excellence”, 12–13.

<sup>51</sup> (Müller és mtsai., 2022, o. 133)

hogy a hamisítás-felismerő modell gyorsabban és jobb teljesítménnyel tanuljon, mintha a nulláról tanítanák be.<sup>52</sup>

Másodjára a szintetikus vagy manipulált videó felismerésben használt mesterséges intelligenciát fejtem ki. A deepfake technológia fejlődését talán egy macska-egér játékhoz lehetne hasonlítani, hiszen a deepfake készítő és felismerő programok is folyamatosan fejlődnek. A mostanában használt legelterjedtebb módja a deepfake felismerésének gépi tanuláson és deepfake adathalmazokon alapuló programok. Ezzel a módszerrel is tudjuk automatizálni ezt, az amúgy emberek számára kihívásokkal teli procedúrát.<sup>53</sup> Az arcok manipulálási módját elkülöníthetjük két típusra. Először az arckifejezés-manipuláció a célszemély valós arcát hagyja a testen, de a támadó tudja az arckifejezéseit irányítani és az ajkait szinkronizálni. Másodjára az arcazonosság-manipuláció a célszemély arcvonásait megváltoztatja, így a személy fizikai megjelenését változtatja meg. A rendszer általában a pislogási, szín vagy textúra hibákat figyeli meg és próbálja felismerni, hogy a kép eredeti-e vagy sem. Fontos hangsúlyozni ennél a résznél, hogy az adat módosítással járó dolgok, mint például a tömörítés, eltávolíthatja az adatmódosításra utaló jeleket.<sup>54</sup> Egyik példája a tömörített videóval járó problémának az, amikor a tömörítés következtében a videó képzaja erősen degradálódik így nem lehet alkalmazni rajta mikroszkópikus elemzést.<sup>55</sup>

Egy másik nehézség lehet az ilyen hamis videók felismerésében, ha a deepfake által manipulált arc csak rövid ideig látható. Az „Institute of Electrical and Electronics Engineers” (IEEE) által kiadott egyik kutatásban egy olyan rekurrens neurális hálózati rendszert hoztak létre, amely a videókat képkockáknak elemzi. A rendszer a „Long Short Term Memory” (továbbiakban LSTM) elnevezésű rendszert használja. Ez lehetővé tette, hogy akár két másodperces videónál 97%-os sikerrel felismerje a rendszer, ha deepfake arcot lát. A rekurrens neurális háló és az LSTM hálózatok eltérnek abban a többi neurális hálótól, hogy egy korábbi esemény figyelembevételével tudnak egy predikciót tenni, mivel az információkat hosszú ideig

---

<sup>52</sup> „Transfer Learning: Yang, Qiang, Zhang, Yu, Dai, Wenyan, Pan, Sinno Jialin: 9781107016903: Amazon.com: Books”, 8, elérés 2023. január 24., <https://www.amazon.com/Transfer-Learning-Qiang-Yang/dp/1107016908>.

<sup>53</sup> Akash Chintla és mtsai., „Recurrent Convolutional Structures for Audio Spoof and Video Deepfake Detection”, *IEEE Journal of Selected Topics in Signal Processing* 14, sz. 5 (2020): 1025, <https://doi.org/10.1109/JSTSP.2020.2999185>.

<sup>54</sup> Andreas Rössler és mtsai., „FaceForensics++: Learning to Detect Manipulated Facial Images” (arXiv, 2019. augusztus 26.), 3, <https://doi.org/10.48550/arXiv.1901.08971>.

<sup>55</sup> Darius Afchar és mtsai., „MesoNet: a Compact Facial Video Forgery Detection Network”, in *2018 IEEE International Workshop on Information Forensics and Security (WIFS)*, 2018, 3, <https://doi.org/10.1109/WIFS.2018.8630761>.

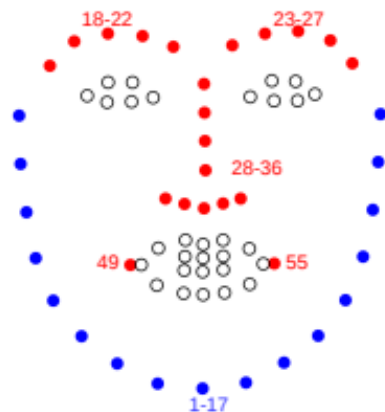
megjegyzik. Ez azt jelenti, hogy a korábbi képkockákat felhasználja az elemzéshez és ezáltal predikciót tud alkotni a rendszer, hogy hamis videóról van-e szó vagy sem.<sup>56</sup>



7. ábra 4 korszerű deepfake modell

(Saját szerkesztés; forrás: Andreas Rössler és mtsai., „FaceForensics++: Learning to Detect Manipulated Facial Images” (arXiv, 2019. augusztus 26.), <https://doi.org/10.48550/arXiv.1901.08971>.)

Egy további módszer a deepfake felismerésére az arc 3D orientációjának eltérésein alapul. Amikor egy arcot rekonstruálunk a sajátunkéra az arcon kiemelkedő pontok, mint például az ajkak sarkai vagy az orr hegye nem mindig marad meg a helyes pozíciójában. Ezek általában az emberi szemmel észrevehetetlen elváltozások. Egy másik fontos pont, hogy a deepfake a központi arctartományban cserél csak arcot. Ez ahhoz vezet, hogy a központi és a teljes arc tájékozódási pontokból kiszámított 3D fejpózok ellentmondásosak lesznek, mivel az arc középső és külső kontúrján eltérések keletkeznek.<sup>57</sup>



8. ábra Az ábrán egy arc látható tájékozódási pontokból generálva. A piros pontok a központi arctartományt jelölik. A kék pontok a külső kontúrt adják. Az üres köröket nem vesszük figyelembe a tájékozódási pontok számításánál

Forrás: Xin Yang, Yuezun Li, és Siwei Lyu, „Exposing Deep Fakes Using Inconsistent Head Poses” (arXiv, 2018. november 13.), 8262–63, <https://doi.org/10.48550/arXiv.1811.00661>.

<sup>56</sup> David Güera és Edward J. Delp, „Deepfake Video Detection Using Recurrent Neural Networks”, in *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 2018, 5, <https://doi.org/10.1109/AVSS.2018.8639163>.

<sup>57</sup> Xin Yang, Yuezun Li, és Siwei Lyu, „Exposing Deep Fakes Using Inconsistent Head Poses” (arXiv, 2018. november 13.), 8262–63, <https://doi.org/10.48550/arXiv.1811.00661>.

### 2.3 Közösségi botok detektálása a gépi tanulás segítségével

A gépi tanulás alapú bot-felderítési technikákat először is aszerint osztályozzák, hogy felügyelt, félig-felügyelt vagy felügyelet nélküli gépi tanulási módszerről van-e szó. Azután tovább osztják tartalomalapúra vagy viselkedésalapúra. A 9. ábrán szerepel egy nem gépi tanuláshoz köthető bot észlelés. Ilyen a „crowdsourcing” és az ember által készített struktúra elemzés. A nem gép által végzett emberi munkát általában adathalmazok gyűjtésére és kutatásokban való részvételre használják tesztalanyként.<sup>58</sup>

A vizsgálati skála szerint 2 csoportra oszthatjuk a technikákat egyéni fiók alapú detektorokra és csoportokon alapuló detektorokra. Az egyéni fiók alapúak alatt azokat a fiókokat értjük, amelyek egyetlen felhasználó által vannak irányítva és automatizálva vannak feladatok, mint a tartalom megosztás vagy más felhasználók követése. Az ilyen feladatokra több online elérhető bot detektor is van. Ilyen bot detektor például a Botometer, mely annak a valószínűségét értékeli, hogy az általunk megadott Twitter felhasználó bot-e vagy sem. Ez az alkalmazás a fiók különböző jellemzőit elemzi, ebbe beleértve a fiók tweetjeinek tartalmát, a fiók követőit és a fiók aktivitását. A rendszer egy bináris pontozás alapján dönt, amely azt jelenti, hogy a 0 pont emberi fiók és az 1 pont pedig botfiók. Manapság már az olyan botok, amelyek csak követésre vannak kifejlesztve, könnyedén felismerhetők egy gépi tanulási rendszer által, mivel nem generálnak tartalmat vagy nem beszélgetnek mással. A detektorok létrehozásában jelentkeztek kihívások is. Az első ilyen kihívás, hogy nincs egy olyan mérési alap (ground truth) adatkészlet, melyet a képzési fázisban felhasználhatnánk. Legtöbbször nincs mérési alap, és a besorolásokat (labels) a kutatók adják meg manuálisan. Ez problémás lehet, mivel az emberek elfoglaltságokkal rendelkeznek, így nehezen ismerik fel a különbségeket botok és emberek által használt fiókok között. Manapság a közösségi botok automatizált és az ember által vezérelt tevékenységek kombinációját mutatják, amelyeket nem lehet csak egyszerű bináris kategorizálásokkal felismerni.<sup>59</sup>

---

<sup>58</sup> Mariam Orabi és mtsai., „Detection of Bots in Social Media: A Systematic Review”, *Information Processing & Management* 57, sz. 4 (2020. július 1.): 8–9, <https://doi.org/10.1016/j.ipm.2020.102250>.

<sup>59</sup> Stefano Cresci, „A Decade of Social Bot Detection”, *Communications of the ACM* 63, sz. 10 (2020. szeptember 23.): 76–77, <https://doi.org/10.1145/3409116>.



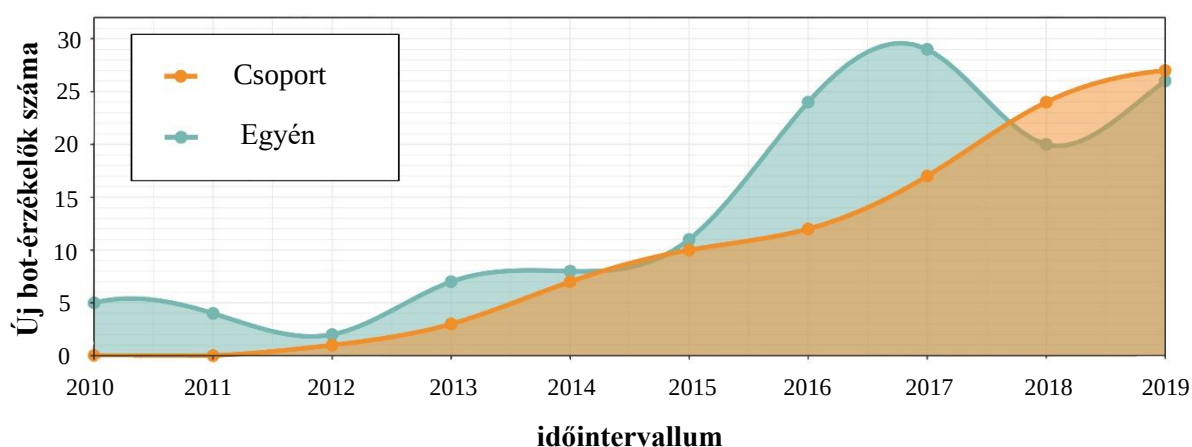
9. ábra Közösségi bot észlelési technikák

(Saját szerkesztés, forrás: „Juršėnas, A., Karlauskas, K., Ledinauskas, E., Maskeliūnas, G., Rondonas, D., Ruseckas, J. The Role of AI in the Battle Against Disinformation (2022). Riga: NATO Strategic Communications Centre of Excellence”.)

A csoportos bot detektor viszont több olyan fiók koordinációját figyeli meg, amelyek közös cél elérése érdekében működnek együtt. Ilyenek általában az üzenet felerősítése, megosztása vagy az online beszélgetések manipulálása azt a látszatot keltve, hogy az üzenettel több felhasználó is egyetért. Ez a módszer időt és energiát takaríthat meg a felderítési fázisban, mivel a hamis bot fiókok általában együttműködnek egymással és botneteket hoznak létre. A botnetek többnyire egy entitás irányában rendelkeznek célokkal. Ennek hatására több bizonyítékot hagynak maguk után automatizáltságukról, mint az egyetlen fiók alapú közösségi botok. A detektorok csoportos bot fiókok vizsgálatával még több információhoz is hozzáférnek ezáltal több adattal tudják ellátni a mesterséges intelligencia algoritmusokat. Ennek köszönhetően csoportos felismerésnél a felügyelet nélküli és a félig felügyelt tanítási módszerre támaszkodnak inkább a kutatók.<sup>60</sup>

<sup>60</sup> Cresci, 78.

Egy további módszer az „anómália” alapú bot felismerés. A felismerés alapja itt az, hogy a különösnek vélt online viselkedésből következteti ki a modell, hogy bot vagy valódi felhasználó-e a fiók. Ezt a módszert legtöbbször a botok észlelésének első szakaszában használják fel. Ennek segítségével javítva a jövőbeli észlelési technikákat alapigazságsadatkészlet létrehozásával. Az ilyen rendellenes viselkedések kiszűrésének egy módszere a kutatók általi „honey potok”-kal való interakciója a botoknak. A honey potokat arra használják a kutatók vagy biztonsági szakértők, hogy a támadók tevékenységét megfigyelhessék és információkat gyűjtsenek a támadók taktikáiról. Egy „Information Processing & Management” folyóiratban publikált kutatásban 60 ilyen honey potot hoztak létre a Twitteren, amelyek tartalomszennyezőként szolgáltak. Ezekkel a honey potokkal kapcsolatba lépő fiókokat botoknak bélyegezték, azt feltételezve, hogy egy valódi felhasználó nem lépne kapcsolatba vagy nem követne ilyen fiókokat. A botok pedig azért lépnek ilyen fiókokkal kapcsolatba, hogy saját kommunikációs hálójukat kiterjesszék.<sup>61</sup>



10. ábra 2010-óta közzétett botdetektor osztályozása az idő függvényében

(Saját szerkesztés, forrás: Cresci, „A Decade of Social Bot Detection”.)

<sup>61</sup> Orabi és mtsai., „Detection of Bots in Social Media”, 13.

## 2.4 A dezinformáció gráf és hálózatalapú felismerése

A matematikai gráfok több ponton is segítségre vannak a dezinformáció felismerésében. A közösségi hálók szerkezetét gyakran ábrázolják gráfokkal, élek és pontok segítségével. Ezek a profilok közötti kölcsönhatásokra deríthetnek fényt, amelyek leginkább segítséget nyújthatnak a csoportos botfelismerésben.<sup>62</sup>

A gráf alapú felismerésre két rendszer példát fogok felhozni. Az egyik rendszer egy „áldozati osztályzás”-i rendszert használ arra, hogy a valódi felhasználókat magasabb rangsorba helyezze, mint az általa botnak vélt fiókokat. Ez a rendszer felügyelt gépi tanulás mellett a gráfokat úgy használja fel, hogy a potenciális áldozatokhoz, vagyis valódi felhasználókhoz alacsony súlyokat rendel. Egy megbízható felhasználói fióktól pedig elkezdi vizsgálni a gráfokon random éleken elindulva, hogy a többi pont valódi felhasználó-e vagy sem. A másik detektálási módszer az az időbeli aktivitás megfigyelése. A hasonló időbeli viselkedéssel rendelkező, erősen összekapcsolt fiókok nagy valószínűséggel botok. Ez a módszer a csoportos bot felismerésnél játszhat nagy szerepet.<sup>63</sup>

A hálózati jellemzők több módon is felhasználhatóak az álhírek és dezinformáció felismerésében. Először a visszhangkamrákat (echo chambers) segít azonosítani, amely egy olyan környezetet jelent, ahol egy személy csak olyan információkkal vagy nézőpontokkal találkozhat, amelyek a sajátját tükrözik. Ez megnehezíti az egyén számára a kritikus gondolkodást és az alternatív nézőpontok mérlegelését, és fokozhatja a társadalmi polarizációt. Másodszor feltárja az egyes felhasználók különböző szerepeit az álhírek terjesztésében.<sup>64</sup>

A hálózatok két fajtája homogén és a heterogén. A homogén hálózatok azonos csomópont- és kapcsolattípusokon osztoznak. A heterogén hálózatok pedig különböző csomópont- és kapcsolat típusokkal rendelkeznek. A heterogén hálózatok előnye, hogy az adatokat és a kapcsolatokat sokféle szempontból képesek reprezentálni és kódolni. A felhasználók, a közösségi médiában közzétett bejegyzések, a tényleges hír és más, különféle entitások mind részt vesznek a hírterjesztési folyamatokban.<sup>65</sup>

---

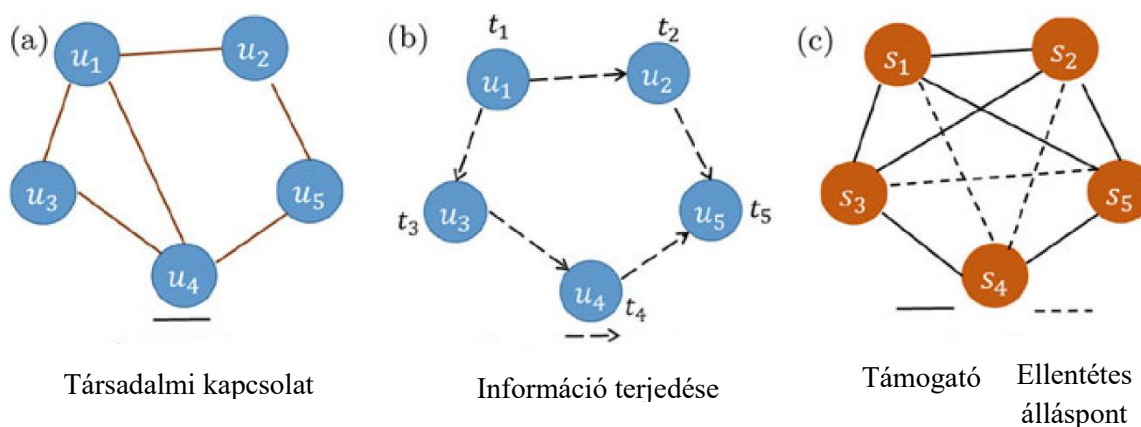
<sup>62</sup> Orabi és mtsai., 8.

<sup>63</sup> Orabi és mtsai., 9.

<sup>64</sup> Kai Shu, H. Russell Bernard, és Huan Liu, „Studying Fake News via Network Analysis: Detection and Mitigation”, in *Emerging Research Challenges and Opportunities in Computational Social Network Analysis and Mining*, szerk. Nitin Agarwal, Nima Dokuohaki, és Serpil Tokdemir, Lecture Notes in Social Networks (Cham: Springer International Publishing, 2019), 45, [https://doi.org/10.1007/978-3-319-94105-9\\_3](https://doi.org/10.1007/978-3-319-94105-9_3).

<sup>65</sup> Shu, Bernard, és Liu, 47–48.

A homogén hálózatokon belül itt hármat fogok bemutatni. A hitelességi-, diffúziós- és barátsági hálózatokat. A barátság hálózatok, ahogy a neve is sugallja a felhasználók közötti barátságot szemlélteti élekkel és csomópontokkal. Ez azért is fontos, mivel a hírek terjesztésének elsődleges módszere az ismerősökön keresztül történik. A barátsághálózat a hírek terjedésének alapvető útvonalát tárhatja fel. A diffúziós hálózat segítséget nyújthat abban, hogy az időbeli tendenciákat megértsük, és hogy az álhírek a kiindulópontját megállapíthassuk. A felhasználók közötti terjedési útvonalakat megállapíthatjuk ennek a hálózatnak a segítségével. Utoljára a hitelességi hálózat a hírek általános valószínűségét értékeli. Ezt a felhasználók által nyújtott vélemények segítik elő. A támogató hozzászólások növelik a hitelességi pontszámot, az ellentétes álláspontok pedig csökkentik azt.<sup>66</sup>



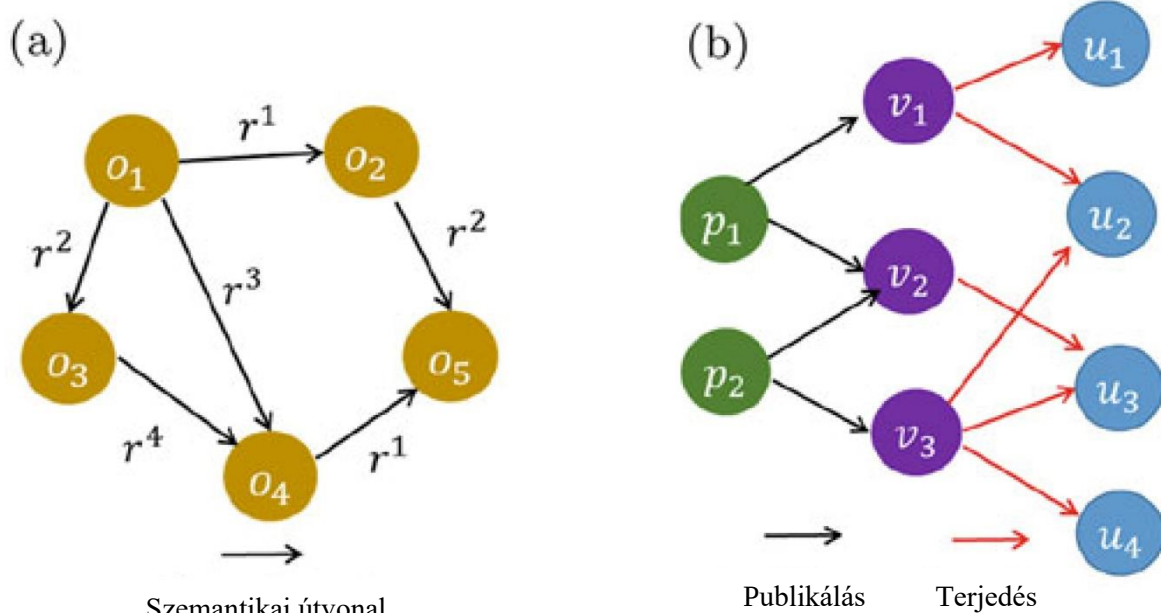
11. ábra Homogén hálózatok. A homogén hálózatok három típusa: a) baráti hálózat, b) diffúziós hálózat c) hitelességi hálózat. Az "u" csomópont egy felhasználót jelöl, az "s" pedig egy közösségi média posztot.

(Saját szerkesztés, forrás: Shu, Bernard, és Liu, „Studying Fake News via Network Analysis”.)

A heterogén hálózatoknál főként tudáshálózatokról és interakciós hálózatokról beszélhetünk. A tudáshálózatok tényellenőrzést végeznek az általuk összekapcsolt nyílt adatok segítségével. Ha egy híranyagban szereplő állítás levezethető a tudáshálózatban már meglévő tényekből, akkor az átment a tudáshálózat tényellenőrzési vizsgálatán, így valós a hír. Az interakciós hálózat a különféle entitások, köztük a kiadók, a hírek és a közösségi médiában megjelenő posztok közötti kapcsolatok hálózatként való ábrázolását végzi. Itt a kiadók és a felhasználók tulajdonságai, valamint a kiadók, a hírek és a felhasználók közötti korreláció képesek megkülönböztetni a valódi híreket az álhírektől.<sup>67</sup>

<sup>66</sup> Shu, Bernard, és Liu, 47–48.

<sup>67</sup> Shu, Bernard, és Liu, 48–49.



12. ábra Heterogén hálózatok. A két típust szemlélteti a) tudáshálózat b) interakciós hálózat. Az „o” csomópont egy tudás egységet, „v” egy hírt, „p” egy hírkiadót jelent, az „u” csomópont pedig továbbra is felhasználót jelöl.

(Saját szerkesztés, forrás: Shu, Bernard, és Liu, „Studying Fake News via Network Analysis”.)

## 2.5 Következtetések

Kutatásom második fejezetében a kard másik élet mutattam be, mely talán egy kis reményt adhat számunkra. Ez egy lényeges terület, számos kutatás folyik a dezinformáció, álhírek témáján belül, főként azoknak felismeréséről és megakadályozásáról. Itt nem tudom eléggé hangsúlyozni, hogy a mesterséges intelligencia ezen szemszögéből egy macska-egér játék bontakozik ki, amely arra készíti a kutatókat, hogy minél naprakészebb kutatásokkal és megoldásokkal szolgáljanak.

Véleményem szerint a szakembereknek, akik a felismeréssel és ezeknek a gépi tanulás alapú modellek kidolgozásával foglalkoznak sokkal ravaszabbnak kell lenniük, mint azoknak, akik a dezinformációs kampányokban tevékenykednek. Az álhírek gyors terjedése a dezinformációs szereplők malmára hajtja a vizet és emellett az ilyen aktorokat az állam is támogathatja, így emberi erőforrás és pénzügyi támogatási többlettel rendelkezhetnek.

### 3 Az empirikus kutatás bemutatása

Ebben a fejezetben ismertetem kérdőíves felmérésemnek az összeállítási és kiértékelési módszerét. Az SPSS program által kiadott eredményekre és korrelációkra pedig következtetéseket nyújtok, melyeket alábbi az összefüggésekkel támasztok alá.

#### 3.1 A kérdőív kiértékelésének és összeállításának a módszere

A kérdőívemben főként kutatásom fő témáit dolgoztam fel, melyek a dezinformáció, álhírek a közösségi oldalakon és a mesterséges intelligencia voltak. A tanulmányomat kiegészítő kérdőív segítségével adatokat akartam szerezni, melyek alátámasztják hipotéziseimet és segítenek egy átfogóbb képet alkotni az emberek dezinformációról és mesterséges intelligenciáról alkotott elképzeléseiről.

A kérdőívem kérdéseit két fő csoportba lehet sorolni. Az első a mesterséges intelligenciára vonatkozik és az arról alkotott nézetéről a kitöltőknek. A második csoport pedig a közösségi médiáról és annak hatásairól szólt, például: függőségről, befolyásolhatóságról vagy az álhírek észleléséről és szabályozásáról. Itt fontos kiemelni, hogy a közösségi oldalak kérdéskörének azért szenteltem ekkora szerepet kérdéseim között, mivel legtöbbször az álhírek és dezinformáció itt jelenik meg és terjed el először.

Kutatásom kérdéseit a „technológia elfogadási hajlandósági modell” alapján alkottam (továbbiakban TAP - “Technology Adoption Propensity”). Ez az index a fogyasztók hajlandóságát jelzi a különböző új csúcstechnológiai termékek és szolgáltatások befogadásával kapcsolatban. Ez a modell a mesterséges intelligenciával kapcsolatosan is hasznos volt kutatásom során. A TAP modellt a „technológia elfogadása modell” (továbbiakban TAM – „Technology Acceptance Model”) alapján dolgozták ki a kutatásban, melyből dolgoztam. A TAM modellben szereplő hasznossági és használhatósági kritériumok két alapvető kérdést tükröznek, amelyeket a felhasználók feltehetnek maguknak, amikor véleményt alkotnak egy bizonyos technológiáról:

1. Milyen előnyökre számíthat a fogyasztó az új technológia alkalmazásakor?
2. Mekkora kihívást jelentene a fogyasztó számára a technológia helyes használatának elsajátítása?

A fogyasztók, ha az első kérdést szemlélik felismerhetik, hogy a technológia használatának előnyei és hátrányai egyaránt vannak. A fogyasztók olyan előnyöket tapasztalhatnak, mint a

nagyobb társadalmi beilleszkedés, alkalmazkodóképesség és hatékonyság. A hátrányok közé tartozhat a függőség kialakulása és a társadalmi elszigetelődés. A fogyasztó ezeket a tényezőket mind számításba veszi, mielőtt egy új technológiát használna.<sup>68</sup>

A különbség a TAP és TAM modell között, amíg a TAM modell kiemeli az észlelt hasznosság és a könnyű használat jelentőségét, mint a technológia felhasználói elfogadottságának jelentős tényezőit. Addig a TAP modell a technológia és a felhasználó közötti illeszkedésre összpontosít.

A TAP modell által olyan kérdéseket tettem föl, mely négy kategóriába sorolható. Először az emberek mesterséges intelligencia és közösségi oldalak felé mutatott optimizmusuk felől érdeklődtem. Erre a jelzőre példa mondat: „A mesterséges intelligencia kényelmesebbé teszi a munkámat, mivel automatizál sok tevékenységet.”. Másodjára a jártasság kérdéskörében tettem fel kérdéseket. Erre egy példa: „A mesterséges intelligencia fejlesztési területén a kiemelkedő újításokat és fejlesztéseket figyelemmel kísérem.”. Harmadjára a közösségi oldalak által kialakult függőséget vizsgáltam. Erre egy példa: „Ha nem követem a közösségi oldalon, hogy mi történik a világban attól félek, hogy kimaradok valamiből.” Végül azt is megvizsgáltam, hogy az emberek mennyire érzik magukat veszélyeztetettnek az online befolyásolással szemben, illetve, hogy milyen aggodalmaik vannak a mesterséges intelligencia fejlesztéseivel kapcsolatban. Ez a vizsgálat a résztvevők sebezhetőségére fókuszált. Ilyen kérdés volt például: „A közösségi média túlságosan megkönnyítette a vállalatok és más emberek számára, hogy behatoljanak a magánéletembe.”.

A fő pont melyeket a kérdéseken kívül összevetettem, az a korral való összefüggés volt. Ezzel azt vizsgáltam, hogy különböző korú kitöltők válaszai mennyire térnek el egy másik korcsoporttól. Az adatok elemzéséhez az IBM SPSS Statistics 27 szoftverprogramot használtam és azon belül is a Kruskal-Wallis tesztet alkalmaztam. Az SPSS program elsajátításához és megértéséhez a Sajtos László és Mitev Ariel SPSS kutatási és adatelemzési könyvet alkalmaztam.<sup>69</sup>

Kruskal-Wallis tesztet alkalmaztam az adatok elemzéséhez. A Kruskal-Wallis teszt egy nem paraméteres statisztikai teszt, amelyet két vagy több független csoport mediánjainak összehasonlítására használnak. A Kruskal-Wallis teszt úgy működik, hogy az összes értéket

---

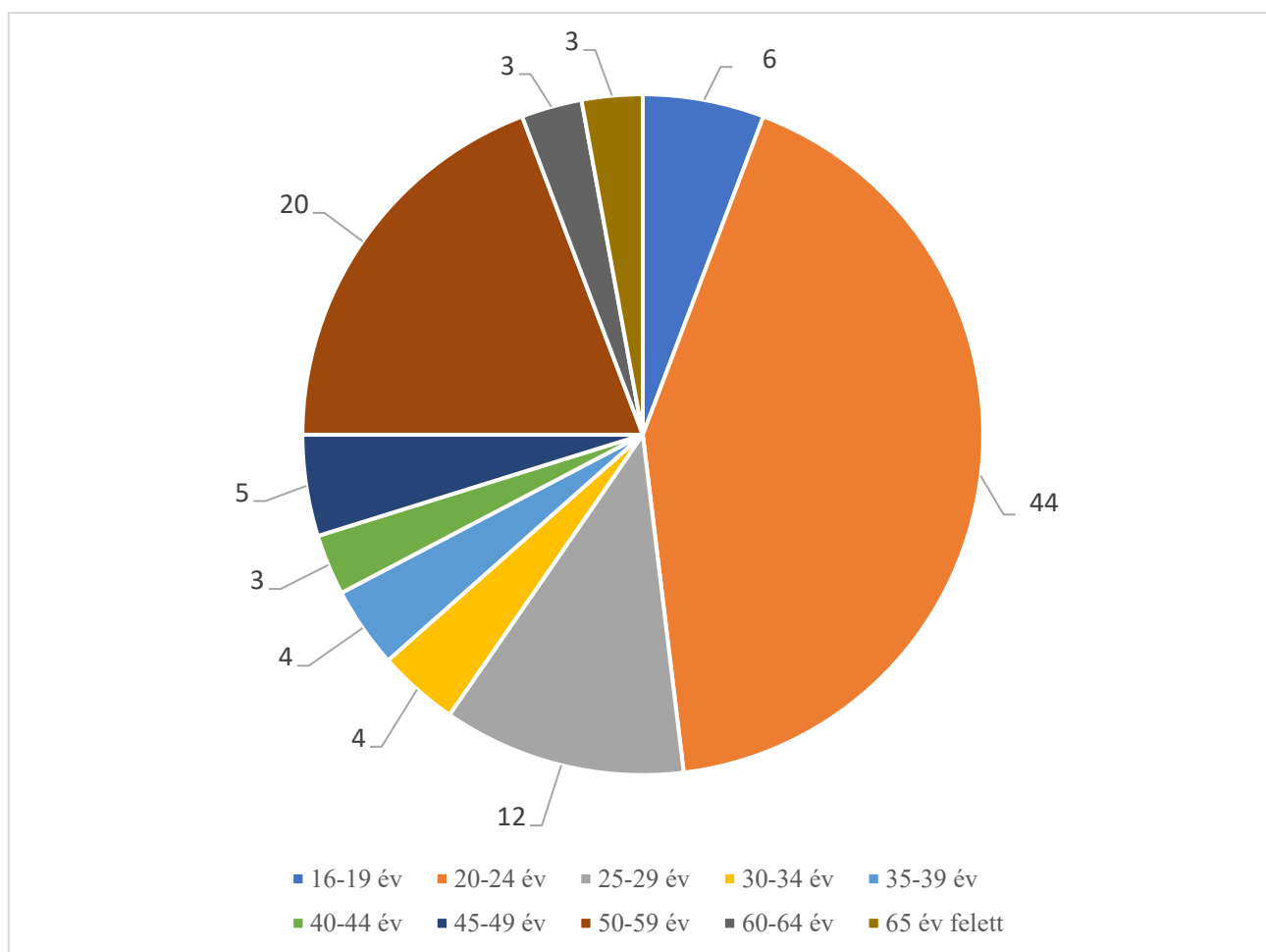
<sup>68</sup> Mark Ratchford és Michelle Barnhart, „Development and Validation of the Technology Adoption Propensity (TAP) Index”, *Journal of Business Research* 65, sz. 8 (2012. augusztus 1.): 1210–11, <https://doi.org/10.1016/j.jbusres.2011.07.001>.

<sup>69</sup> Ariel Mitev és Laszlo Sajtos, *SPSS kutatási és adatelemzési kézikönyv*, 2007.

rangsorolja az csoportok között, majd kiszámítja a rangok összegét. A módszer lényege, hogy ezeket a rangsorokat összehasonlítja, hogy kiderüljön, vannak-e közöttük különbségek. A tesztstatisztika a rangösszegek közötti különbségeken alapul, és egy olyan khi-négyzet eloszlással hasonlítja össze, amelynek szabadságfoka egyenlő a meghatározott tartomány mínusz egy. Ha a teszt p-értéke vagyis az aszimptotikus szignifikanciája kisebb, mint 0,05, akkor a kérdéseknél elutasítjuk a nullhipotézist, miszerint a csoportok között nincs különbség, és arra következtetünk, hogy legalább egy csoport átlaga (M) eltér a többitől.

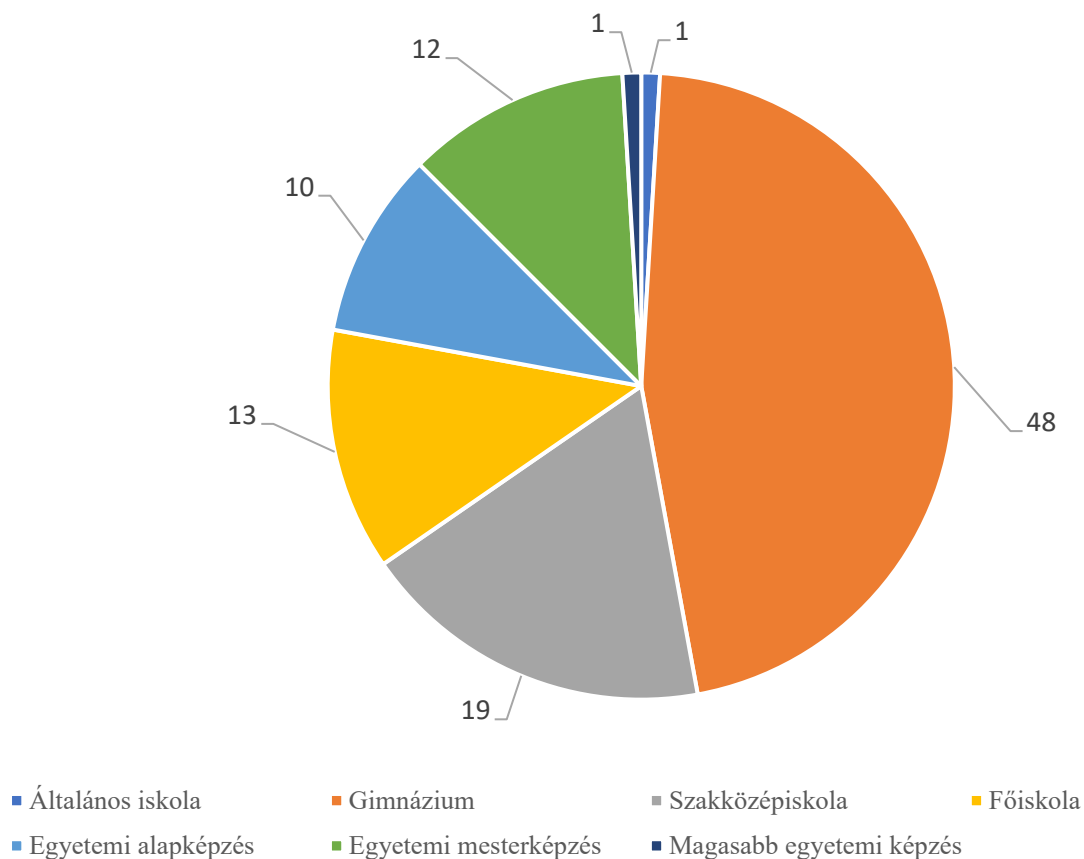
### 3.2 Empirikus kutatásom elemzése

Kérdőívemet 104-en (n=104) töltötték ki, ezen belül ebből 56,7% nő és 43,3% férfi volt. A legtöbb kitöltést a 20-24 éves korcsoporttól mely 42,3% volt, és a legkevesebb kitöltést a 60-65 év feletti korcsoporttól (n=3) tudtam begyűjteni.



13. ábra A kérdőívet kitöltők megoszlása korcsoportok alapján (Saját szerkesztés)

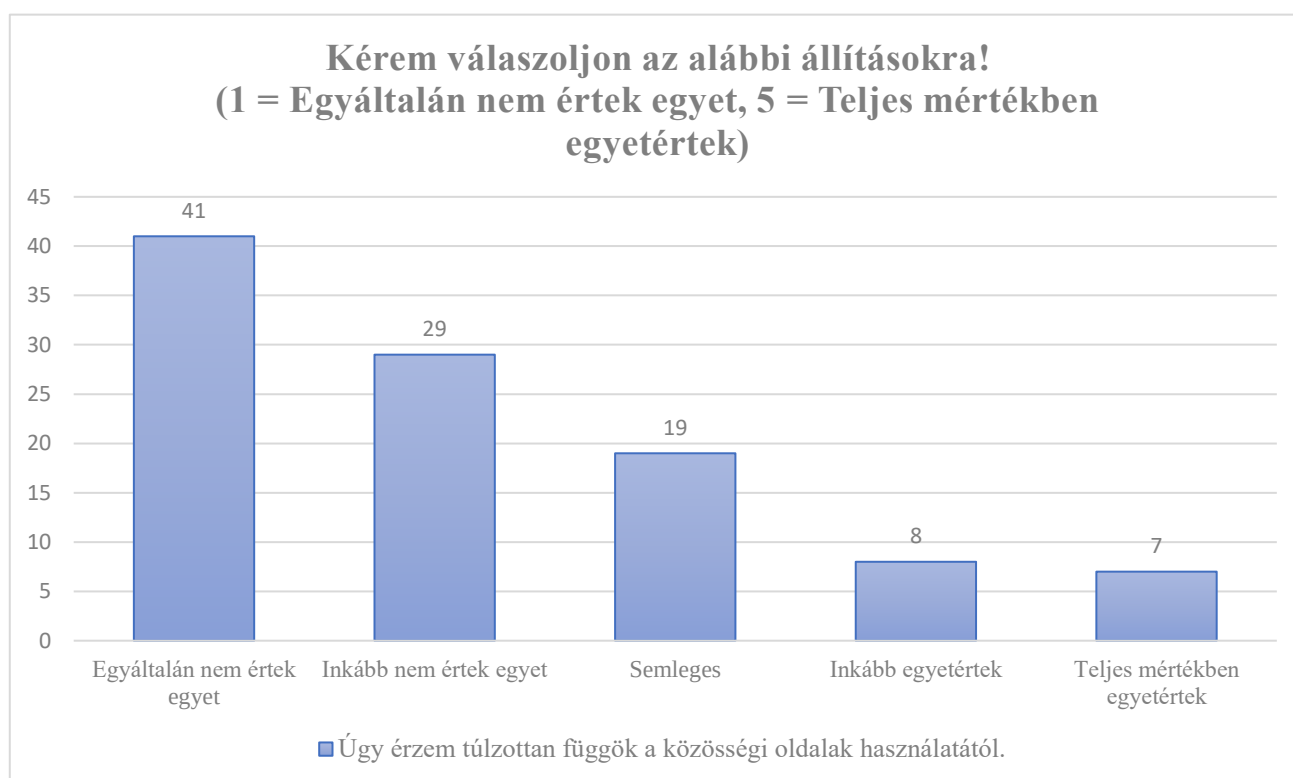
Végzettség szempontjából pedig megállapítható, hogy egy ember kivételével (n=103) mindenki rendelkezik érettségi bizonyítvánnyal. Egyetemi alapképzettek száma 10 (9,6%), mesterképzést végzettek száma 12 (11,5%) és magasabb egyetemi képzést végzett személy csak egy volt (1%).



14. ábra A kérdőívet kitöltők megoszlása iskolai végzettségük alapján (Saját szerkesztés)

A Kruskal-Wallis tesztel főként az életkor befolyását elemeztem különböző kérdések megválaszolásakor, mivel az életkor 10 definiált tartományként (16-19 évtől egészen 65 év felettig) vittem fel az SPSS-be. Először a közösségi média hatását elemeztem és a végén tértem ki a mesterséges intelligencia hatására.

Az első állítás „Úgy érzem túlzottan függök a közösségi oldalak használatától.”, melyet egy 1-től 5-ös skálán lehetett megjelölni. Az 1-es „egyáltalán nem értek egyet” és az 5-ös pedig „teljes mértékben egyetértek”. A további állításoknál, melyeket Kruskal-Wallis tesztel vizsgáltam, mindegyikénél ez az 1-5-ös skálát alkalmaztam így később nem említem meg. Itt a válaszok meglepő módon 1-től 5-ig lefelé ívelő tendenciát mutattak. Az 1-est 41-en, 2-est 29-en, 3-ast 19-en, 4-est 8-an és az 5-öst pedig csak 7-en jelölték meg. Ebből az látszik, hogy a kitöltők önpercepciója azt sugallja, hogy nem függenek a közösségi oldalak használatától.



15. ábra Kérdőívemben állítás: „Úgy érzem túlzottan függök a közösségi oldalak használatától.” (Saját szerkesztés)

A Kruskal-Wallis teszt eredményei alapján szignifikáns különbséget találtunk az emberek különböző korcsoportjai között és a közösségi média túlzott függőségének érzékelésében, Kruskal Wallis  $H = 18557$ ,  $df = 9$ ,  $p = 0.029$ . A különbséget az aszimptotikus szignifikanciánál figyelhetjük meg, mely a  $p = 0.029$  volt, mivel ez az érték kisebb, mint  $p = 0.05$ , ezért mutat az eredmény szignifikáns különbséget. A Kruskal Wallis  $H$  értéke alapján tudjuk még meghatározni, hogy a csoportok közötti különbségek statisztikailag szignifikánsak-e vagy sem. A Kruskal Wallis  $H = 18557$  is egy viszonylag erős bizonyíték mutat arra, hogy a csoportok közötti különbség magas. A minta 104 résztvevőből állt, változó mintanagyságokkal az egyes életkorcsoportokban, beleértve 6 résztvevőt a 16-19 éves korcsoportban, 44 résztvevőt a 20-24 éves korcsoportban, 12 résztvevőt a 25-29 éves korcsoportban, 4 résztvevőt a 30-34 és 35-39 éves korcsoportokban, 3 résztvevőt a 40-44 és 60-64 éves korcsoportokban, 5 résztvevőt a 45-49 éves korcsoportban, 20 résztvevőt a 50-59 éves korcsoportban, és 3 résztvevőt a 65 év feletti korcsoportban. A többi állításnál is ez az eloszlás játszik szerepet.

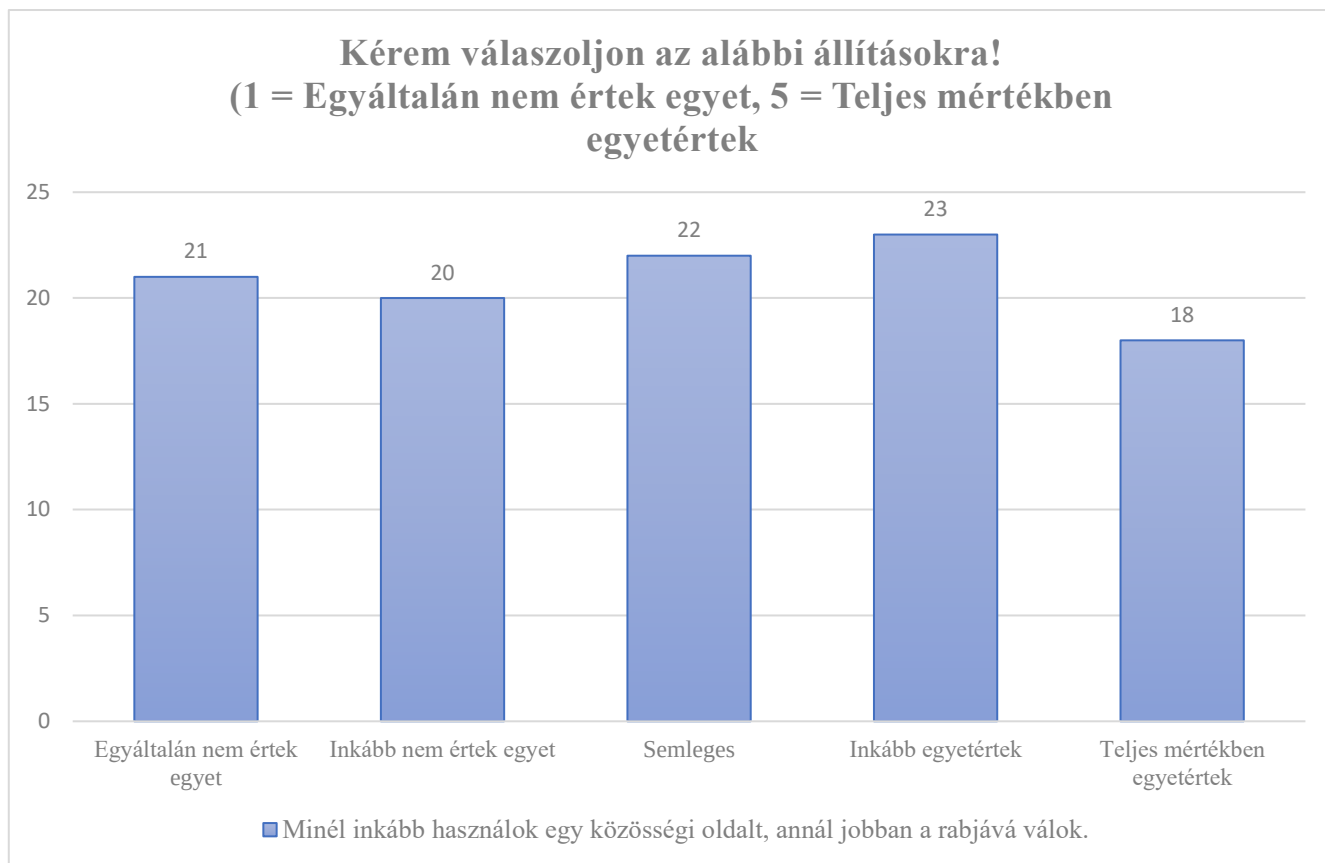
A következő magyarázatban az  $M$  az átlagot jelenti. Az egyes korcsoportokhoz tartozó átlag rangok a következők voltak: 16-19 éveseknél ( $n = 6$ )  $M = 61,67$ , 20-24 éveseknél ( $n = 44$ )  $M = 61,86$ , 25-29 éveseknél ( $n = 12$ )  $M = 49,17$ , 30-34 éveseknél ( $n = 4$ )  $M = 50,50$ , 35-39

éveseknél ( $n = 4$ )  $M = 35,75$ , 40-44 éveseknél ( $n = 3$ )  $M = 84,50$ , 45-49 éveseknél ( $n = 5$ )  $M = 35,00$ , 50-59 éveseknél ( $n = 20$ )  $M = 37,48$ , 60-64 éveseknél ( $n = 3$ )  $M = 44,33$  és 65 év felettiiek esetén ( $n = 3$ )  $M = 40,67$ .

Az átlag rangok alapján megállapíthatjuk, hogy a fiatalabb korosztályok (16-19 és 20-24) magasabb átlag rangokkal rendelkeznek, ami arra utal, hogy nagyobb valószínűséggel érzékelik a közösségi média függőségét erősebbnek. Az idősebb korosztályok 45-49 évesektől a 65 év felettiig alacsonyabb átlag rangokkal rendelkeznek, ami arra utal, hogy kevésbé valószínű, hogy erősnek érzékelik a közösségi médiától való függőséget.

Egy hasonló állítást is feltettem a kérdőívemben mely a „Minél inkább használok egy közösségi oldalt, annál jobban a rabjává válok.” volt. Itt a Kruskal-Wallis teszt eredményei alapján a következők jöttek ki. A Kruskal-Wallis  $H = 17017$ ,  $df = 9$  és a  $p = 0,048$ . Ennél az állításnál ki kell hangsúlyozni, hogy a 0,048 aszimptotikus szignifikancia a küszöbérték 0,05-nél nem sokkal van lejjebb, ezért az eredményt fenntartásokkal kezeljük. Az átlagos rangsorok az egyes korcsoportokban a következők voltak 16-19 évesek ( $n = 6$ ) esetén  $M = 70,92$ , 20-24 éveseknél ( $n = 44$ )  $M = 60,76$ , 25-29 éveseknél ( $n = 12$ )  $M = 39,04$ , 30-34 éveseknél ( $n = 4$ )  $M = 53,25$ , 35-39 éveseknél ( $n = 4$ )  $M = 58,50$ , 40-44 éveseknél ( $n = 3$ )  $M = 60,00$ , 45-49 éveseknél ( $n = 5$ )  $M = 19,20$ , 50-59 éveseknél ( $n = 20$ )  $M = 44,73$ , 60-64 éveseknél ( $n = 3$ )  $M = 53,00$  és 65 év felettiiek esetén ( $n = 3$ )  $M = 38,67$ . Ebben az esetben az aszimptotikus szignifikancia  $p = 0,048$  volt, amely, bár nem jelent szignifikáns különbséget az előző állítás értékéhez képest, azonban még mindig kisebb, mint a  $p = 0,05$ , ezért vizsgálhatjuk tovább a korosztályok között a különbséget.

Az átlag rangok alapján megállapítható, hogy az fiatalabb korcsoportok (16-19 és 20-24 évesek) magasabb rangot kaptak, ami azt jelzi, hogy nagyobb valószínűséggel értenek egyet azzal a kijelentéssel, hogy minél többet használnak egy közösségi oldalt, annál inkább a rabjává válnak tőle. Az idősebb korosztályok (45-49 és 65 év felettiiek) sokkal alacsonyabb rangot kaptak, ami arra utal, hogy kevésbé értenek egyet a kijelentéssel.



16. ábra Kérdőívemben állítás: „Minél inkább használok egy közösségi oldalt, annál jobban a rabjává válok.” (Saját szerkesztés)

Az előző két tesztben hasonlóan megfogalmazott kérdésekre hasonló eredmények jöttek ki, mely arra utal, hogy statisztikailag alá lehet támasztani megfigyelésem. Ebből az szűrhető le, hogy a fiatalabb generáció könnyebben válhat függővé a közösségi oldalak által. Ez több veszéllyel is járhat, mivel a különféle álhírek és dezinformáció a közösségi oldalakon üti fel először a fejét és ott terjednek előszeretettel.

A Kruskal-Wallis tesztet alkalmaztam az utolsó közösségi oldal kapcsolatos állításnál is, mely a „A közösségi média túlságosan megkönnyítette a vállalatok és más emberek számára, hogy behatoljanak a magánéletembe.” volt. Ennél az állításnál is a kort vettem csoportosító változónak.

A Kruskal-Wallis-teszt eredményei azt mutatják, hogy a különböző korcsoportok között szignifikáns eltérés van az állítás megítélésében, Kruskal-Wallis  $H = 19111$   $df = 9$  és  $p = 0,024$ . A minta 104 résztvevőből állt, korcsoportonként eltérő mintamérettel. Az egyes korcsoportok átlagos rangsorai a következők voltak: 16-19 ( $n = 6$ )  $M = 68,83$ , 20-24 ( $n = 44$ )  $M = 57,78$ , 25-29 ( $n = 12$ )  $M = 63,96$ , 30-34 ( $n = 4$ )  $M = 62,88$ , 35-39 ( $n = 4$ )  $M = 38,25$ , 40-44 ( $n = 3$ )  $M = 33,50$ , ( $n = 5$ ) 45-49  $M = 43,30$ , 50-59 ( $n = 20$ )  $M = 45,78$ , 60-64 ( $n = 3$ )  $M = 9,17$  és 65 év felett ( $n = 3$ ) ( $M = 24,17$ ).

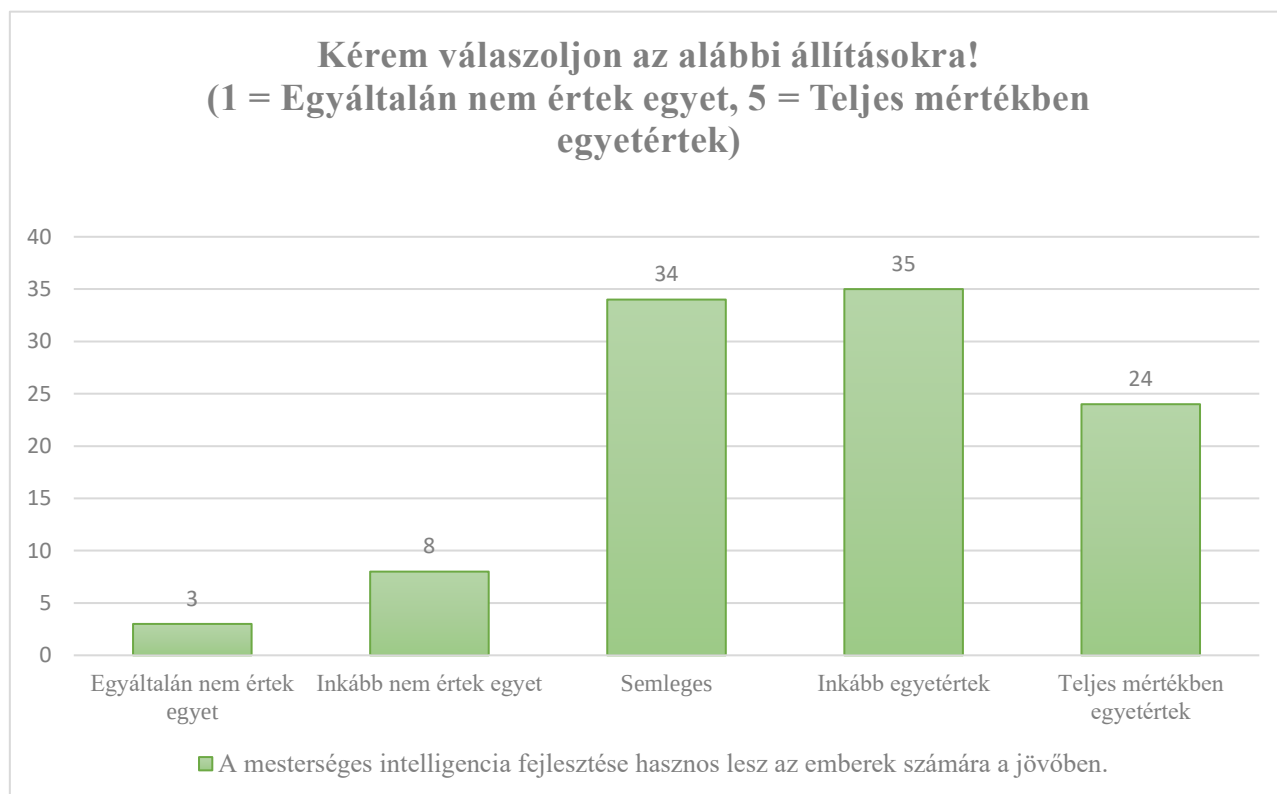
Ezekből az eredményekből leszűrhetjük, hogy egészen a 16-19 éves korosztálytól a 30-34 évesekig úgy értékelték, hogy más emberek a közösségi média által nagyobb hatással lehetnek magánéletükre és jobb belátást nyújt arra. Míg a 35-39 éves korcsoporttól, már jelentős az átlag értékek csökkenése. Ebből arra következtethetünk, hogy a fiatalabb korosztály többször használja a közösségi oldalakat, vagy esetleg több dolgot tesz ki adott oldalakra.



17. ábra Kérdőívemben állítás: „A közösségi média túlságosan megkönnyítette a vállalatok és más emberek számára, hogy behatoljanak a magánéletembe.” (Saját szerkesztés)

Egy másik állítás, melyet a Kruskal-Wallis módszerrel vizsgáltam a „A mesterséges intelligencia fejlesztése hasznos lesz az emberek számára a jövőben.”, melyet az életkorral, mint csoportosító változóval egyeztettem. Az eredmények alapján arra a következtetésre jutottam, hogy a mesterséges intelligencia hasznosságának megítélése különböző korosztályok között szignifikánsan eltér. A Kruskal-Wallis H érték 17125 volt és a  $df = 9$ . Emellett az aszimptotikus szignifikancia értéke, ami  $p = 0,047$ , amely okot ad arra, hogy a megfigyelésemet tovább vizsgáljam külön korcsoportok szintjén, mivel szignifikáns különbség figyelhető meg köztük.

A 16-19 évesek ( $n=6$ )  $M = 47,17$ , a 20-24 évesek ( $n=44$ )  $M = 62,47$ , a 25-29 évesek ( $n=12$ )  $M = 59,29$ , a 30-34 évesek ( $n=4$ )  $M = 61,75$ , a 35-39 évesek ( $n=4$ )  $M = 25,25$ , a 40-44 évesek ( $n=3$ )  $M = 40,00$ , a 45-49 évesek ( $n=5$ )  $M = 45,60$ , az 50-59 évesek ( $n=20$ )  $M = 38,38$ , a 60-64 évesek ( $n=3$ )  $M = 51,50$ , a 65 év felettiek ( $n=3$ )  $M = 33,00$  volt.



18. ábra Kérdőívemben állítás: „A mesterséges intelligencia fejlesztése hasznos lesz az emberek számára a jövőben.”  
 (Saját szerkesztés)

A Kruskal-Wallis teszt átlagos rangjaiból látható, hogy a fiatalabb korosztályok (16-19 és 20-24 évesek) magasabb átlagos rangokkal rendelkeztek, ami arra utal, hogy valószínűbb, hogy egyetértenek abban, hogy a mesterséges intelligencia fejlesztése hasznos lesz az emberek számára a jövőben. A 25-29, 30-34 és 40-44 éves korosztályoknak is viszonylag magas átlagos rangjaik voltak, ami arra utal, hogy valamennyire egyetértenek. Másrésről azonban a 35-39,

45-49, 50-59 és 65 év feletti korosztályok alacsonyabb átlagos rangokkal rendelkeztek, ami azt sugallja, hogy kevésbé értenek egyet a kijelentéssel. Fontos azonban megjegyezni, hogy az 0,047 aszimptotikus szignifikancia érték csak enyhén van a 0,05 szignifikancia küszöbérték alatt, így ezeket az eredményeket óvatosan kell értelmezni.

### **3.3 Következtetések**

A kérdőívem kiértékelése segítségével sikerült alátámasztani hipotéziseimet, melyek arra világítottak rá, hogy a mesterséges intelligenciával szemben sokkal pozitívabb a hozzáállása a fiatalabb generációnak, mint az idősebbnek. Ezek mellett egy negatív szempontot is megfigyelhettünk, mely a TAP modell által felállított függőség és veszélyeztetettség érzékelése a fiatalabb generációnál magasabb volt, mint az idősebbeknél. Ez annak is köszönhető lehet, hogy az idősebb generáció másként viszonyul a technológiához és magánélethez, ezért sem érzik magukat annyira veszélynek kitéve. A függőség az idősebb generációnál láthatóan kisebb probléma. Ez számtalan tényezőnek betudható, például a digitális készségek meg nem léte, közösségi médián kívüli szociális kapcsolatok preferálása, vagy csak szimplán a fiatalabb generációtól eltérő érdeklődési körök és életszemlélet. Egy fő tényező lehet még, hogy az idősebbek saját magukon kevésbé veszik észre a függőség jeleit, így kevésbé értékelik magukat függésre hajlamosnak.

### **Összegzett következtetések**

Kutatásomban a mesterséges intelligencia és a gépi tanulás hatását vizsgáltam a dezinformációs kampányokra és az álhírek terjedésére. Ezt a témát megvizsgáltam, dezinformáló és azt akadályozó és észlelő szemszögből is. Ez azért volt fontos, hogy ne csak az egyik oldalt ismerjük meg, hanem a dezinformációt terjesztők cselekedeteiből is tudunk tanulni.

Kutatásom első fejezetében azt dolgoztam fel, hogy mesterséges intelligencia hogyan képes a dezinformációs munkálatokat megkönnyíteni. Itt több példán keresztül bemutattam, hogy mely fázisban, mely technológia képes hasznossá válni. Számomra az egyik legfontosabb következtetés az a technológia fejlődésével kapcsolatos volt. Egy technológiát sokkal könnyebb megalkotni, mint utána szabályozni azt, hogy ne tudják a fenyegető szereplők azokat felhasználni. Második fejezetben a szakértők vagy kutatók munkáját vizsgáltam, akik a dezinformációt próbálják felismerni különböző gépi tanulási programokkal. Ennél a

legfontosabb megfigyelésem a technológia folytonos fejlődésének hatása volt. Hiszen egy konstans verseny alakul ki a dezinformációt terjesztők és a szakértők vagy azt észlelők között.

Utolsó sorban kérdőívem segítségével mértem fel a mesterséges intelligencia felé mutatott attitűdöt, a közösségi oldalak hatását, és hogy az álhíreket mennyire vélik a kitöltők problémásnak. Egyik fő kérdése a kérdőívemnek, hogy a kitöltők találkoztak-e már közösségi oldalon álhírrrel, amelyre egy kiemelkedő 94%-os igen volt a válasz. Itt vizsgáltam még, hogy az álhírek ellen kormányzati szabályzás kéne, vagy inkább a közösségi oldalaknak kéne azokat szabályozni. A közösségi oldalak saját maguk szabályozását a kitöltők helyesebbnek vélték, mivel 87%-uk igennel válaszolt, míg a kormányzati szabályozásnál csak 64% volt az igenek aránya. Ez azt mutatja, hogy az emberek jobban tartják felelősnek a felületek szolgáltatóit ilyen jellegű problémák orvoslására, mint az államot.

## **Eredmények**

A kérdőív statisztikai vizsgálata után a hipotéziseimet alá tudtam támasztani. Erre szerettem volna általam vélt magyarázatokkal szolgálni. Itt ki kell emelni, hogy a kérdőív eredményei és abból levont eredmények csak a kitöltők körére igaz. Ezáltal nem lehet általános állítást alkotni, amely minden emberre igaz.

T1: A fiatalabb generáció, mely kutatási eredményeim alapján 16-24 évesig definiálhatunk, ténylegesen hajlamosabbak a közösségi média használatától függeni. A fiatalabb generációkat jobban érdekelheti az online szocializáció és a digitális jelenlét fenntartása. Ez a közösségi média fokozott használatához és a függőség nagyobb kockázatához vezethet.

T2: A fiatalabb generáció 16-24 éves korig láthatóan az eredményekből optimistábban viszonyul a mesterséges intelligencia fejlesztéséhez. Ennek oka az lehet, hogy ez a generáció technológiával együtt nőtt fel, és már fiatal kortól kezdve találkozott vele, míg az idősebb generációnak esetleg később kellett alkalmazkodni a technológiához. Emellett a fiatalabb generáció az oktatás és a média révén jobban megismerheti a mesterséges intelligenciát, ami befolyásolhatja a fejlődéssel kapcsolatos hozzáállást.

T3: Nem csak a fiatalabb generáció érzi a magánéletbe való behatolást közösségi médián keresztül, hanem egészen 34 éves korcsoportig bezárólag. Ez azon is alapulhat, mivel a fiatalabb generáció aktívabb a közösségi média oldalakon és több személyes információt osztanak meg a külvilággal. Emellett a fiatalabb generációt jobban aggaszthatja a magánélet és

az adatbiztonság, mivel olyan digitális korban nőtt fel, ahol gyakoriak a személyazonossággal való visszaélések.

## Felhasznált irodalom

- Afchar, Darius, Vincent Nozick, Junichi Yamagishi, és Isao Echizen. „MesoNet: a Compact Facial Video Forgery Detection Network”. In *2018 IEEE International Workshop on Information Forensics and Security (WIFS)*, 1–7, 2018. <https://doi.org/10.1109/WIFS.2018.8630761>.
- Assenmacher, Dennis, Lena Clever, Lena Frischlich, Thorsten Quandt, Heike Trautmann, és Christian Grimme. „Demystifying Social Bots: On the Intelligence of Automated Social Media Actors”. *Social Media + Society* 6, sz. 3 (2020. július): 205630512093926. <https://doi.org/10.1177/2056305120939264>.
- Badawy, Adam, Emilio Ferrara, és Kristina Lerman. „Analyzing the Digital Traces of Political Manipulation: The 2016 Russian Interference Twitter Campaign”. In *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 258–65, 2018. <https://doi.org/10.1109/ASONAM.2018.8508646>.
- Chintha, Akash, Bao Thai, Saniat Javid Sohrawardi, Kartavya Bhatt, Andrea Hickerson, Matthew Wright, és Raymond Ptucha. „Recurrent Convolutional Structures for Audio Spoof and Video Deepfake Detection”. *IEEE Journal of Selected Topics in Signal Processing* 14, sz. 5 (2020): 1024–37. <https://doi.org/10.1109/JSTSP.2020.2999185>.
- Choraś, Michał, Konstantinos Demestichas, Agata Giełczyk, Álvaro Herrero, Paweł Ksieniewicz, Konstantina Remoundou, Daniel Urda, és Michał Woźniak. „Advanced Machine Learning Techniques for Fake News (Online Disinformation) Detection: A Systematic Mapping Study”. *Applied Soft Computing* 101 (2021. március 1.): 107050. <https://doi.org/10.1016/j.asoc.2020.107050>.
- Cresci, Stefano. „A Decade of Social Bot Detection”. *Communications of the ACM* 63, sz. 10 (2020. szeptember 23.): 72–83. <https://doi.org/10.1145/3409116>.
- Du Bois, John. „The stance triangle”, 139–82, 2007. <https://doi.org/10.1075/pbns.164.07du>.
- El Naqa, Issam, és Martin J. Murphy. „What Is Machine Learning?” In *Machine Learning in Radiation Oncology: Theory and Applications*, szerkesztette Issam El Naqa, Ruijiang Li, és Martin J. Murphy, 3–11. Cham: Springer International Publishing, 2015. [https://doi.org/10.1007/978-3-319-18305-3\\_1](https://doi.org/10.1007/978-3-319-18305-3_1).
- „ElevenLabs - Prime AI Text to Speech | Voice Cloning”. Elérés 2023. április 19. <https://beta.elevenlabs.io/>.
- „Fake Tweet, Chat & Facebook Status Generator | Prank Me Not”. Elérés 2023. április 19. <https://www.prankmenot.com/>.
- Ferrara, Emilio, Onur Varol, Clayton Davis, Filippo Menczer, és Alessandro Flammini. „The Rise of Social Bots”. *Communications of the ACM* 59, sz. 7 (2016. június 24.): 96–104. <https://doi.org/10.1145/2818717>.
- Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, és Yoshua Bengio. „Generative adversarial networks”. *Communications of the ACM* 63, sz. 11 (2020. 0 22.): 139–44. <https://doi.org/10.1145/3422622>.
- Groza, Adrian. „Detecting fake news for the new coronavirus by reasoning on the Covid-19 ontology”. arXiv, 2020. április 26. <http://arxiv.org/abs/2004.12330>.
- Güera, David, és Edward J. Delp. „Deepfake Video Detection Using Recurrent Neural Networks”. In *2018 15th IEEE International Conference on Advanced Video and*

- Signal Based Surveillance (AVSS)*, 1–6, 2018.  
<https://doi.org/10.1109/AVSS.2018.8639163>.
- Han, Su-Hyun, Ko Woon Kim, SangYun Kim, és Young Chul Youn. „Artificial Neural Network: Understanding the Basic Concepts without Mathematics”. *Dementia and Neurocognitive Disorders* 17, sz. 3 (2018. szeptember): 83–89.  
<https://doi.org/10.12779/dnd.2018.17.3.83>.
- Howard, Philip N., és Bence Kollanyi. „Bots, #StrongerIn, and #Brexit: Computational Propaganda during the UK-EU Referendum”. arXiv, 2016. június 20.  
<https://doi.org/10.48550/arXiv.1606.06356>.
- „Juršėnas, A., Karlauskas, K., Ledinauskas, E., Maskeliūnas, G., Rondonanskas, D., Ruseckas, J. The Role of AI in the Battle Against Disinformation (2022). Riga: NATO Strategic Communications Centre of Excellence”, é. n.
- „Juršėnas, A., Karlauskas, K., Ledinauskas, E., Maskeliūnas, G., Ruseckas, J., The Double-Edged Sword of AI: Enabler of Disinformation (2021). Riga: NATO Strategic Communications Centre of Excellence.”, é. n.
- Kaplan, Jerry. *Artificial Intelligence: What Everyone Needs to Know*R. Oxford University Press, 2016.
- Katerina Sedova, Christine McNeill, Aurora Johnson, Aditi Joshi, and Ido Wulkan,. „AI and the Future of Disinformation Campaigns”, é. n. <https://doi.org/10.51593/2021CA011>.
- Kreps, Sarah, R. Miles McCain, és Miles Brundage. „All the News That’s Fit to Fabricate: AI-Generated Text as a Tool of Media Misinformation”. *Journal of Experimental Political Science* 9, sz. 1 (2022. ed): 104–17. <https://doi.org/10.1017/XPS.2020.37>.
- Küçük, Dilek, és Fazli Can. „Stance Detection: A Survey”. *ACM Computing Surveys* 53, sz. 1 (2021. január 31.): 1–37. <https://doi.org/10.1145/3369026>.
- Lazer, D.M.J., M.A. Baum, Y. Benkler, A.J. Berinsky, K.M. Greenhill, F. Menczer, M.J. Metzger, és mtsai. „The Science of Fake News: Addressing Fake News Requires a Multidisciplinary Effort”. *Science* 359, sz. 6380 (2018): 1094–96.  
<https://doi.org/10.1126/science.aao2998>.
- Lee, Chris van der, Albert Gatt, Emiel van Miltenburg, és Emiel Krahmer. „Human Evaluation of Automatically Generated Text: Current Trends and Best Practice Guidelines”. *Computer Speech & Language* 67 (2021. május 1.): 101151.  
<https://doi.org/10.1016/j.csl.2020.101151>.
- Medhat, Walaa, Ahmed Hassan, és Hoda Korashy. „Sentiment Analysis Algorithms and Applications: A Survey”. *Ain Shams Engineering Journal* 5, sz. 4 (2014. december): 1093–1113. <https://doi.org/10.1016/j.asej.2014.04.011>.
- Mertoğlu, Uğur, Burkay Genç, és Hayri Sever. „Text-Based Fake News Detection via Machine Learning”. In *Trends in Data Engineering Methods for Intelligent Systems*, szerkesztette Jude Hemanth, Tuncay Yigit, Bogdan Patrut, és Anastassia Angelopoulou, 113–24. Lecture Notes on Data Engineering and Communications Technologies. Cham: Springer International Publishing, 2021.  
[https://doi.org/10.1007/978-3-030-79357-9\\_12](https://doi.org/10.1007/978-3-030-79357-9_12).
- Mitev, Ariel, és Laszlo Sajtos. *SPSS kutatási és adatelemzési kézikönyv*, 2007.
- Müller, Nicolas M., Pavel Czempin, Franziska Dieckmann, Adam Froggyar, és Konstantin Böttinger. „Does Audio Deepfake Detection Generalize?” arXiv, 2022. április 21.  
<https://doi.org/10.48550/arXiv.2203.16263>.
- Noble, William S. „What Is a Support Vector Machine?” *Nature Biotechnology* 24, sz. 12 (2006. december): 1565–67. <https://doi.org/10.1038/nbt1206-1565>.
- Orabi, Mariam, Djedjiga Mouheb, Zaher Al Aghbari, és Ibrahim Kamel. „Detection of Bots in Social Media: A Systematic Review”. *Information Processing & Management* 57, sz. 4 (2020. július 1.): 102250. <https://doi.org/10.1016/j.ipm.2020.102250>.

- Ratchford, Mark, és Michelle Barnhart. „Development and Validation of the Technology Adoption Propensity (TAP) Index”. *Journal of Business Research* 65, sz. 8 (2012. augusztus 1.): 1209–15. <https://doi.org/10.1016/j.jbusres.2011.07.001>.
- Rössler, Andreas, Davide Cozzolino, Luisa Verdoliva, Christian Riess, Justus Thies, és Matthias Nießner. „FaceForensics++: Learning to Detect Manipulated Facial Images”. arXiv, 2019. augusztus 26. <https://doi.org/10.48550/arXiv.1901.08971>.
- Ruediger, Marco Aurélio, Carlos Augusto Liguori Filho, Ezequiel Fajreldines dos Santos, Guilherme Kenzo Santos, João Pedro Favaretto Salvador, Rodrigo Moura Karolczak, Tatiane Guimarães, Theófilo Miguel de Aquino, és Victor Doering Silveira. „Bots and Brazil’s Electoral Legal System: 2018 Elections”, 2019. január. <http://bibliotecadigital.fgv.br:80/dspace/handle/10438/26229>.
- Shu, Kai, H. Russell Bernard, és Huan Liu. „Studying Fake News via Network Analysis: Detection and Mitigation”. In *Emerging Research Challenges and Opportunities in Computational Social Network Analysis and Mining*, szerkesztette Nitin Agarwal, Nima Dokoochaki, és Serpil Tokdemir, 43–65. Lecture Notes in Social Networks. Cham: Springer International Publishing, 2019. [https://doi.org/10.1007/978-3-319-94105-9\\_3](https://doi.org/10.1007/978-3-319-94105-9_3).
- Stieglitz, Stefan, Milad Mirbabaie, Björn Ross, és Christoph Neuberger. „Social Media Analytics – Challenges in Topic Discovery, Data Collection, and Data Preparation”. *International Journal of Information Management* 39 (2018. április 1.): 156–68. <https://doi.org/10.1016/j.ijinfomgt.2017.12.002>.
- „Towards automatic fake news classification - Ghosh - 2018 - Proceedings of the Association for Information Science and Technology - Wiley Online Library”. Elérés 2023. január 24. <https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/pra2.2018.14505501125>.
- „Transfer Learning: Yang, Qiang, Zhang, Yu, Dai, Wenyan, Pan, Sinno Jialin: 9781107016903: Amazon.com: Books”. Elérés 2023. január 24. <https://www.amazon.com/Transfer-Learning-Qiang-Yang/dp/1107016908>.
- „TWINT - Twitter Intelligence Tool”. Python. 2017. Reprint, TWINT Project, 2023. április 19. <https://github.com/twintproject/twint>.
- „What Is a Chatbot?” Elérés 2023. január 4. <https://www.oracle.com/chatbots/what-is-a-chatbot/>.
- „What Is Natural Language Processing? | IBM”. Elérés 2023. január 4. <https://www.ibm.com/topics/natural-language-processing>.
- Yang, Xin, Yuezun Li, és Siwei Lyu. „Exposing Deep Fakes Using Inconsistent Head Poses”. arXiv, 2018. november 13. <https://doi.org/10.48550/arXiv.1811.00661>.

## 4. Függelék

### 4.1 Táblázat jegyzék

|                                                           |    |
|-----------------------------------------------------------|----|
| 1. táblázat A természetes nyelvi generálás 6 lépése ..... | 17 |
|-----------------------------------------------------------|----|

### 4.2 Ábrák jegyzéke

|                                                                                                                                                                                                                                         |    |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 1. ábra Egy általános gépi tanulási modell elemei.....                                                                                                                                                                                  | 8  |
| 2. ábra Egy általános gépi tanulási modell elemei.....                                                                                                                                                                                  | 8  |
| 3. ábra A dezinformáció fázisai.....                                                                                                                                                                                                    | 9  |
| 4. ábra A GAN képzési folyamata során tanuló két modell a diszkriminátor "D" és a generátor "G"..                                                                                                                                       | 15 |
| 5. ábra Prank Me Not weboldal segítségével készített hamis tweet (Saját szerkesztés) .....                                                                                                                                              | 20 |
| 6. ábra Egy általános chatbot felépítése .....                                                                                                                                                                                          | 22 |
| 7. ábra 4 korszerű deepfake modell.....                                                                                                                                                                                                 | 29 |
| 8. ábra Az ábrán egy arc látható tájékozódási pontokból generálva. A piros pontok a központi arctartományt jelölik. A kék pontok a külső kontúrt adják. Az üres köröket nem vesszük figyelembe a tájékozódási pontok számításánál.....  | 29 |
| 9. ábra Közösségi bot észlelési technikák .....                                                                                                                                                                                         | 31 |
| 10. ábra 2010-óta közzétett botdetektor osztályozása az idő függvényében.....                                                                                                                                                           | 32 |
| 11. ábra Homogén hálózatok. A homogén hálózatok három típusa: a) baráti hálózat, b) diffúziós hálózat c) hitelességi hálózat. Az "u" csomópont egy felhasználót jelöl, az "s" pedig egy közösségi média posztot. ....                   | 34 |
| 12. ábra Heterogén hálózatok. A két típust szemlélteti a) tudáshálózat b) interakciós hálózat. Az „o” csomópont egy tudás egységet, „v” egy hírt, „p” egy hírkiadót jelent, az "u" csomópont pedig továbbra is felhasználót jelöl. .... | 35 |
| 13. ábra A kérdőívet kitöltők megoszlása korcsoportok alapján (Saját szerkesztés) .....                                                                                                                                                 | 38 |
| 14. ábra A kérdőívet kitöltők megoszlása iskolai végzettségük alapján (Saját szerkesztés).....                                                                                                                                          | 39 |
| 15. ábra Kérdőívemben állítás: „Úgy érzem túlzottan függök a közösségi oldalak használatától.” (Saját szerkesztés).....                                                                                                                 | 40 |
| 16. ábra Kérdőívemben állítás: „Minél inkább használok egy közösségi oldalt, annál jobban a rabjává válok.” (Saját szerkesztés).....                                                                                                    | 42 |
| 17. ábra Kérdőívemben állítás: „A közösségi média túlságosan megkönnyítette a vállalatok és más emberek számára, hogy behatoljanak a magánéletembe.” (Saját szerkesztés).....                                                           | 43 |
| 18. ábra Kérdőívemben állítás: „A mesterséges intelligencia fejlesztése hasznos lesz az emberek számára a jövőben.” (Saját szerkesztés).....                                                                                            | 43 |

### 4.3 1.számú melléklet (Kérdőíves felmérés kérdései)

1. Kérem adja meg a nemét!
  - a) férfi
  - b) nő
  
2. Kérem adja meg az életkorát!
  - a) 16-19 év
  - b) 20-24 év
  - c) 25-29 év
  - d) 30-34 év
  - e) 35-39 év
  - f) 40-44 év
  - g) 45-49 év
  - h) 50-59 év
  - i) 60-64 év
  - j) 65 év felett
  
3. Kérem adja meg a legmagasabb iskolai végzettségét!
  - a) Általános iskola
  - b) Gimnázium
  - c) Szakközépiskola
  - d) Főiskola
  - e) Egyetemi alapképzés
  - f) Egyetemi mesterképzés
  - g) Magasabb egyetemi képzés
  
4. Előfordult már, hogy anélkül osztott meg egy hírt a közösségi médián, hogy ellenőrizte volna annak hitelességét, és később kiderült volna, hogy az hamis vagy félrevezető?
  - a) Igen
  - b) Nem
  - c) Nem tudom
  
5. Ön, hogyan értékeli az álhírek terjedését?
  - a) Nagyon problémásnak értékelem
  - b) Problémásnak értékelem
  - c) Semlegesnek értékelem
  - d) Nem értékelem problémásnak
  - e) Egyáltalán nem értékelem problémásnak
6. Honnan szerzi elsősorban a híreket?
  - a) Közösségi médiáról
  - b) Online hírportálokról
  - c) Nyomtatott sajtóból
  - d) Nem szokott hírek után nézni

7. Átlagosan mennyi időt tölt naponta a hírek olvasásával?
- a) Nem szoktam híreket olvasni
  - b) 1-30 perc
  - c) 30-60 perc
  - d) 1-2 óra
  - e) 2 óránál több
8. Milyen eszközt használ elsősorban, hogy tájékozódjon hírekről?
- a) Televízió
  - b) Okostelefon
  - c) Rádió
  - d) Számítógép
  - e) Egyiket sem
9. Találkozott már álhírekkel vagy dezinformációval a közösségi média platformokon, (például a Facebookon, Twitteren, Instagramon vagy a TikTokon)?
- a) Igen
  - b) Nem
  - c) Nem tudom
10. Ön szerint van szükség kormányzati szabályozásra az álhírek terjedése ellen?
- a) Igen
  - b) Nem
  - c) Nem tudom
11. Ön szerint kell szabályozni a közösségi média felületeknek az álhírek terjedését?
- a) Igen
  - b) Nem
  - c) Nem tudom
12. Utána szoktam nézni különböző forrásokból, hogy az adott cikk vagy adat igaz-e.  
1 - 5 skála

Kérem válaszoljon az alábbi állításokra!

(1 = Egyáltalán nem bízom meg, 5= Teljes mértékben megbízom)

13. Megbízom a közösségi média platformokról származó hírekben és információkban, melyeket ismerőseim osztanak meg.
14. Megbízom a közösségi média platformokról származó hírekben és információkban, melyeket ismeretlen személyek osztanak meg.
- 1 - 5 skála

Kérem válaszoljon az alábbi állításokra!

(1 = Egyáltalán nem aggaszt, 5= Nagyon aggaszt)

15. Aggaszt a mesterséges intelligencia további fejlesztése.

16. Aggaszt a mesterséges intelligencia programok által generált tartalmak (például. deepfake) lehetséges hatása az álhírek terjedésére.  
1 – 5 skála

Kérem válaszoljon az alábbi állításokra!

(1 = Egyáltalán nem értek egyet, 5= Teljes mértékben egyetértek)

17. A mesterséges intelligencia fejlesztésének köszönhetően a mindennapi életem könnyebbé vált. (például: ChatGPT)  
18. A mesterséges intelligencia fejlesztése hasznos lesz az emberek számára a jövőben.  
19. A mesterséges intelligencia kényelmesebbé teszi a munkámat, mivel automatizál sok tevékenységet.  
20. A mesterséges intelligencia fejlesztési területén a kiemelkedő újításokat és fejlesztéseket figyelemmel kísérem.  
1 – 5 skála

Kérem válaszoljon az alábbi állításokra!

(1 = Egyáltalán nem értek egyet, 5= Teljes mértékben egyetértek)

21. Élvezem, ha az újabb mesterséges intelligencia fejlesztéseket ki tudom használni.  
22. A mesterséges intelligencia megkönnyíti az életemet.  
23. A mesterséges intelligencia sebezhetőbbé tesz bűnözők számára.  
24. Úgy gondolom, hogy a mesterséges intelligenciát a közösségi médiaplatformok ellenőrzésére és szabályozására kellene használni az álhírek terjedésének megakadályozása érdekében.  
1 – 5 skála

Kérem válaszoljon az alábbi állításokra!

(1 = Egyáltalán nem értek egyet, 5= Teljes mértékben egyetértek)

25. Minél inkább használok egy közösségi oldalt, annál jobban a rabjává válok.  
26. Ha nem követem a közösségi oldalon, hogy mi történik a világban attól félek, hogy kimaradok valamiből.  
27. Bízom abban, hogy képes vagyok különbséget tenni a valós és hamis információk között az interneten.  
28. Annyira sok új információ és impulzus ér az interneten különösképp a közösségi oldalakon, hogy nehéz kiszűrni a félrevezető bejegyzéseket a valósaktól.  
1 – 5 skála

Kérem válaszoljon az alábbi állításokra!

(1 = Egyáltalán nem értek egyet, 5= Teljes mértékben egyetértek)

- 29. A közösségi oldalakon láttak valami újat tudtak tanítani nekem.
- 30. A közösségi oldalakon megjelenő posztok és bejegyzések befolyásolják a gondolkodásom.
- 31. Úgy érzem túlzottan függök a közösségi oldalak használatától.
- 32. A közösségi média túlságosan megkönnyítette a vállalatok és más emberek számára, hogy behatoljanak a magánéletembe.

1 – 5 skála

#### 4.4 2.számú melléklet (Kérdőívek kiértékelése)

##### NPar Tests

##### Descriptive Statistics

|                                                               | N   | Mean | Std. Deviation | Minimum | Maximum |
|---------------------------------------------------------------|-----|------|----------------|---------|---------|
| Úgy érzem túlzottan függök a közösségi oldalak használatától. | 104 | 2,14 | 1,218          | 1       | 5       |
| Kérem adja meg az életkorát!                                  | 104 | 4,19 | 2,791          | 1       | 10      |

##### Kruskal-Wallis Test

##### Ranks

|                                                               | Kérem adja meg az életkorát! | N   | Mean Rank |
|---------------------------------------------------------------|------------------------------|-----|-----------|
| Úgy érzem túlzottan függök a közösségi oldalak használatától. | 16-19 év                     | 6   | 61,67     |
|                                                               | 20-24 év                     | 44  | 61,86     |
|                                                               | 25-29 év                     | 12  | 49,17     |
|                                                               | 30-34 év                     | 4   | 50,50     |
|                                                               | 35-39 év                     | 4   | 35,75     |
|                                                               | 40-44 év                     | 3   | 84,50     |
|                                                               | 45-49 év                     | 5   | 35,00     |
|                                                               | 50-59 év                     | 20  | 37,48     |
|                                                               | 60-64 év                     | 3   | 44,33     |
|                                                               | 65 év felett                 | 3   | 40,67     |
|                                                               | Total                        | 104 |           |

##### Test Statistics<sup>a,b</sup>

Úgy érzem  
túlzottan függök  
a közösségi  
oldalak  
használatától.

|                  |        |
|------------------|--------|
| Kruskal-Wallis H | 18,557 |
| df               | 9      |
| Asymp. Sig.      | ,029   |

a. Kruskal Wallis Test

b. Grouping Variable: Kérem adja meg az életkorát!

## NPar Tests

### Descriptive Statistics

|                                                                            | N   | Mean | Std. Deviation | Minimum | Maximum |
|----------------------------------------------------------------------------|-----|------|----------------|---------|---------|
| Minél inkább használok egy közösségi oldalt, annál jobban a rabjává válok. | 104 | 2,97 | 1,390          | 1       | 5       |
| Kérem adja meg az életkorát!                                               | 104 | 4,19 | 2,791          | 1       | 10      |

## Kruskal-Wallis Test

### Ranks

|                                                                            | Kérem adja meg az életkorát! | N   | Mean Rank |
|----------------------------------------------------------------------------|------------------------------|-----|-----------|
| Minél inkább használok egy közösségi oldalt, annál jobban a rabjává válok. | 16-19 év                     | 6   | 70,92     |
|                                                                            | 20-24 év                     | 44  | 60,76     |
|                                                                            | 25-29 év                     | 12  | 39,04     |
|                                                                            | 30-34 év                     | 4   | 53,25     |
|                                                                            | 35-39 év                     | 4   | 58,50     |
|                                                                            | 40-44 év                     | 3   | 60,00     |
|                                                                            | 45-49 év                     | 5   | 19,20     |
|                                                                            | 50-59 év                     | 20  | 44,73     |
|                                                                            | 60-64 év                     | 3   | 53,00     |
|                                                                            | 65 év felett                 | 3   | 38,67     |
|                                                                            | Total                        | 104 |           |

### Test Statistics<sup>a,b</sup>

|                                                                            |        |
|----------------------------------------------------------------------------|--------|
| Minél inkább használok egy közösségi oldalt, annál jobban a rabjává válok. |        |
| Kruskal-Wallis H                                                           | 17,017 |
| df                                                                         | 9      |
| Asymp. Sig.                                                                | ,048   |

a. Kruskal Wallis Test

b. Grouping Variable: Kérem adja meg az életkorát!

## NPar Tests

### Descriptive Statistics

|                                                                                                                     | N   | Mean | Std. Deviation | Minimum | Maximum |
|---------------------------------------------------------------------------------------------------------------------|-----|------|----------------|---------|---------|
| A közösségi média túlságosan megkönnyítette a vállalatok és más emberek számára, hogy behatoljanak a magánéletembe. | 104 | 3,66 | 1,146          | 1       | 5       |
| Kérem adja meg az életkorát!                                                                                        | 104 | 4,19 | 2,791          | 1       | 10      |

### Kruskal-Wallis Test

#### Ranks

|                                                                                                                     | Kérem adja meg az életkorát! | N   | Mean Rank |
|---------------------------------------------------------------------------------------------------------------------|------------------------------|-----|-----------|
| A közösségi média túlságosan megkönnyítette a vállalatok és más emberek számára, hogy behatoljanak a magánéletembe. | 16-19 év                     | 6   | 68,83     |
|                                                                                                                     | 20-24 év                     | 44  | 57,78     |
|                                                                                                                     | 25-29 év                     | 12  | 63,96     |
|                                                                                                                     | 30-34 év                     | 4   | 62,88     |
|                                                                                                                     | 35-39 év                     | 4   | 38,25     |
|                                                                                                                     | 40-44 év                     | 3   | 33,50     |
|                                                                                                                     | 45-49 év                     | 5   | 43,30     |
|                                                                                                                     | 50-59 év                     | 20  | 45,78     |
|                                                                                                                     | 60-64 év                     | 3   | 9,17      |
|                                                                                                                     | 65 év felett                 | 3   | 24,17     |
|                                                                                                                     | Total                        | 104 |           |

### Test Statistics<sup>a,b</sup>

|                                                                                                                     |        |
|---------------------------------------------------------------------------------------------------------------------|--------|
| A közösségi média túlságosan megkönnyítette a vállalatok és más emberek számára, hogy behatoljanak a magánéletembe. |        |
| Kruskal-Wallis H                                                                                                    | 19,111 |
| df                                                                                                                  | 9      |
| Asymp. Sig.                                                                                                         | ,024   |

a. Kruskal Wallis Test

b. Grouping Variable: Kérem adja meg az életkorát!

## NPar Tests

### Descriptive Statistics

|                                                                                    | N   | Mean | Std. Deviation | Minimum | Maximum |
|------------------------------------------------------------------------------------|-----|------|----------------|---------|---------|
| A mesterséges intelligencia fejlesztése hasznos lesz az emberek számára a jövőben. | 104 | 3,66 | 1,011          | 1       | 5       |
| Kérem adja meg az életkorát!                                                       | 104 | 4,19 | 2,791          | 1       | 10      |

## Kruskal-Wallis Test

### Ranks

|                                                                                    | Kérem adja meg az életkorát! | N   | Mean Rank |
|------------------------------------------------------------------------------------|------------------------------|-----|-----------|
| A mesterséges intelligencia fejlesztése hasznos lesz az emberek számára a jövőben. | 16-19 év                     | 6   | 47,17     |
|                                                                                    | 20-24 év                     | 44  | 62,47     |
|                                                                                    | 25-29 év                     | 12  | 59,29     |
|                                                                                    | 30-34 év                     | 4   | 61,75     |
|                                                                                    | 35-39 év                     | 4   | 25,25     |
|                                                                                    | 40-44 év                     | 3   | 40,00     |
|                                                                                    | 45-49 év                     | 5   | 45,60     |
|                                                                                    | 50-59 év                     | 20  | 38,38     |
|                                                                                    | 60-64 év                     | 3   | 51,50     |
|                                                                                    | 65 év felett                 | 3   | 33,00     |
|                                                                                    | Total                        | 104 |           |

### Test Statistics<sup>a,b</sup>

A mesterséges  
intelligencia  
fejlesztése  
hasznos lesz az  
emberek  
számára a  
jövőben.

|                  |        |
|------------------|--------|
| Kruskal-Wallis H | 17,125 |
| df               | 9      |
| Asymp. Sig.      | ,047   |

a. Kruskal Wallis Test

b. Grouping Variable: Kérem adja meg az életkorát!

## Rövidítések jegyzéke

| Rövidítés | Magyar megnevezés                          | Idegen nyelvű megnevezés                          |
|-----------|--------------------------------------------|---------------------------------------------------|
| NATO      | Észak-atlanti Szerződés Szervezete         | The North Atlantic Treaty Organization            |
| TAP       | Technológia elfogadási hajlandósági modell | Technology Adoption Propensity                    |
| TAM       | technológia elfogadása modell              | Technology Acceptance Model                       |
| NLP       | Természetes nyelvi feldolgozás             | Natural Language Processing                       |
| NLG       | Természetes nyelvi generálás               | Natural Language Generation                       |
| NLU       | Természetes nyelvi megértés                | Natural Language Understanding                    |
| SVM       | Támogató vektor gép                        | Support Vector Machine                            |
| GAN       | Generatív ellenséges hálózatok             | Generative Adversarial Networks                   |
| GPT       | Általános előtanított transzformátor       | Generative Pre-trained Transformer                |
| LSTM      | Hosszú rövidtávú memória                   | Long Short Term Memory                            |
| df        | Szabadsági fok                             | Degrees of Freedom                                |
| IEEE      | -                                          | Institute of Electrical and Electronics Engineers |