

Csonka Viktor

MESTERSÉGES INTELLIGENCIA TANULÁSÁNAK MEGFIGYELÉSE ÉS BEFOLYÁSOLÁSA

Témavezető: Dr. Guzsvinecz Tibor

Műszaki Informatikai Kar, Informatikai Rendszerek és Alkalmazásai Tanszék

Pannon Egyetem
Intézményi Tudományos Diákköri Konferencia

2023



Mesterséges intelligencia tanulásának megfigyelése és befolyásolása

Készítette: Csonka Viktor

Programtervező informatikus BSc

Műszaki Informatikai Kar, Informatikai Rendszerek és Alkalmazásai Tanszék

Témavezető: Dr. Guzsvinecz Tibor

Napjainkban nagyon elterjedt témává vált a mesterséges intelligencia. Ezt mi sem bizonyítja jobban mintsem, hogy a szintén máig népszerű Netflixnek kettő és fél évbe telt míg elérte az egymillió regisztrált felhasználót. Míg azonban az OpenAI csevegő robotjának ez csupán öt napba került. A mesterséges intelligencia terjedését mi sem indokolja jobban mint, hogy képes komplex feladatokat megoldani, általunk emberek helyett.

Az elterjedésének hála olyan modellek, algoritmusok kerülnek kifejlesztésre melyek segítik a hatalmas mennyiségű adat (Big Data) feldolgozását és ezek alapján képesek döntések meghozatalára. Olyan területek alkalmaznak mesterséges intelligenciát, melyekre elsöre nem is gondolnánk: egészségügy [1], autonóm rendszerek [2], és kiberbiztonság [3]. Ezen területek korábban ember által végzett rendkívül megterhelő és időigényes feladatokat igényeltek.

Ahhoz, hogy a mesterséges intelligencia akár összetett feladatokat is el tudjon látni, tanulnia kell. Eközben pedig külső környezeti ingerek befolyásolhatják és zavarhatják meg az adatfeldolgozást. Ebben a kutatásban arra szeretnék rávilágítani, hogyan befolyásolhatja a gépi tanulást a felhasználó általi inger egy feladat elvégzésében. Ennek megfigyelésére a Unity rendszerében egy háromdimenziós virtuális környezetet implementáltam, amelyben a karakter építkezni tanul. Az eredmények azt mutatják, hogy a külső ingerek lassítják a tanulási folyamatot, de nem akadályozzák meg a végleges cél elérésében.

[1] Davenport, T. Kalakota, R., The potential for artificial intelligence in healthcare, *Future Healthc. J.* **2019**, 6, (2) 94–98, doi: 10.7861/futurehosp.6-2-94.

[2] Grigorescu, S., Trasnea, B., GA survey of deep learning techniques for autonomous driving. *J. Field Robotics* **2019**, 37, (3), 362–386.

[3] Li, J.-H., Cyber security meets artificial intelligence: a survey, *Front. Inf. Technol. Electron. Eng.* **2018**, 19, (12), 1462–1474, doi: 10.1631/fitee.1800573.

Tartalomjegyzék

Ábrajegyzék	4
Táblázatjegyzék	6
Rövidítésjegyzék	7
Köszönetnyilvánítás	8
1. Bevezetés	9
2. Szakirodalmi áttekintés	11
2.1. Felügyeletlen tanulás	11
2.2. Felügyelt tanulás	12
2.3. Ajánló rendszerek	13
2.4. Megerősítéses tanulás	14
2.5. Kutatási kérdés és hipotézisek	14
3. Felhasznált módszerek	16
3.1. A szakirodalmi áttekintés módszere	16
3.2. A virtuális környezet fejlesztése	16
3.2.1. Az alkalmazás struktúrája	17
3.2.2. ML-Agents	18
3.2.3. Környezeti változók	18
3.2.4. A virtuális környezet mögötti logika	19
3.3. Az elkészült virtuális környezet	21
3.4. Adatgyűjtés	22
3.4.1. A mérések	23
3.4.2. Beavatkozások a tanulási folyamatba	23
3.5. Adatelemzés	24
4. Eredmények	25
4.1. Jutalmak	25

4.2. Építkezés	27
4.3. Kiemelkedő epizódok	35
5. Átbeszélés	38
6. Összefoglalás	40
7. Irodalomjegyzék	41

Ábrajegyzék

1. ábra: A felügyelt tanulás folyamata. A [29]-ben található ábra alapján készült.	12
2. ábra: A felhasználói értékelések mátrixa. Minden cella $(r_{u,i})$ megfelel egy u felhasználó által adott, i termék értékelésének. Az algoritmusnak meg kell állapítania, hogy a hiányzó $r_{u,i}$ cellában milyen értékelést fog adni az u felhasználó. A [41]-ben található ábra alapján készült.	13
3. ábra: A megerősítéses tanulás folyamata. A [46]-ben található ábra alapján készült.	14
4. ábra: Az alkalmazás osztálydiagramja.	17
5. ábra: A környezeti változókat szabályzó grafikai elemek.	18
6. ábra: A virtuális környezet mögötti logika.	20
7. ábra: A virtuális környezet és a felhasználói felület.	21
8. ábra: A láthatatlan falak a virtuális térben.	22
9. ábra: Az adatok formátuma.	23
10. ábra: A jutalmak hisztogramjai, csoportosítva a különböző felhasználói beavatkozások alapján: Nincs beavatkozás (a); a szél bekapcsolásra került (b); a földrengés bekapcsolásra került (c); mindkettő bekapcsolásra került (d).	26
11. ábra: Az építkezés tanulásának hisztogramjai, csoportosítva a különböző felhasználói beavatkozások alapján: Nincs beavatkozás (a); a szél bekapcsolásra került (b); a földrengés bekapcsolásra került (c); mindkettő bekapcsolásra került (d).	29
12. ábra: Az ágens építkezést tanuló szokásai felhasználói beavatkozás nélkül.	30
13. ábra: Az ágens építkezést tanuló szokásai szeles környezetben.	31
14. ábra: Az ágens építkezést tanuló szokásai földrengést tartalmazó környezetben.	32
15. ábra: Az ágens építkezést tanuló szokásai szelet és földrengést tartalmazó virtuális környezetben.	32
16. ábra: A különbségek a szél ereje és a tanulási szokások között. Az ábrán a p értékek szerepelnek.	33

17. ábra: A különbségek a földrengés ereje és a tanulási szokások között. Az ábrán a p értékek szerepelnek.	34
18. ábra: A különbségek a szél ereje, a földrengés ereje és a tanulási szokások között. Az ábrán a p értékek szerepelnek.	35
19. ábra: Egy torony, aminek az alapja két gömb.	36
20. ábra: Az ágens a gyenge szél ellenére is felmászott a toronyra.	36
21. ábra: Az építkezés során a falnak lökték az ágenst az objektumok.	37
22. ábra: A földrengés miatt összedőlt a torony.	37

Táblázatjegyzék

1. táblázat: A jutalmak leíró statisztikái a különböző futások során.	25
2. táblázat: Azon cselekedetek leíró statisztikái, amikor az ágens építkezni tanult.	28
3. táblázat: Azon cselekedetek leíró statisztikái, amikor az ágens nem tanult építkezni.	28

Rövidítésjegyzék

American Psychological Association	APA
Artificial intelligence	AI
Hipotézis	H
Mesterséges intelligencia	MI
Natural language processing	NLP
Non-playable character	NPC
Objektumorientált programozás	OOP
Unity Machine Learning Agents Toolkit	ML-Agents

Köszönetnyilvánítás

Ezen sorokkal szeretném kifejezni mély hálámat Dr. Guzsvinecz Tibornak, témavezetőmnek, aki kivételes hozzáértésével és szakmai segítségnyújtással járult hozzá a Tudományos Diákköri Konferenciára benyújtott kutatásom sikeres megvalósításához. Konzulensem remek szakmai tudása és hasznos észrevételei nélkül a kutatás elkészítése elképzelhetetlen lenne.

Az általa nyújtott iránymutatás és támogatás nemcsak segített a kutatás minőségének javításában, hanem személyes és szakmai fejlődésemet is elősegítette.

Szeretném megköszönni továbbá a szakmai segítséget és bátorítást Czeglédi Konrádnak, Fogl Barnának és Sági Biankának akik ezzel támogatták a dolgozat elkészültét.

1. Bevezetés

Az 1940-es és 50-es években számos kutató kezdett el foglalkozni azzal a gondolattal, hogy létrehozzanak egy mesterséges agyat. Neuropszichológusok munkájával létre is jött az első neurális hálózat modell, ami az „M-P model” névre hallgatott [1]. Ez volt az első matematikai modell, ami az emberi agyban működő neuronok hálóját és működését utánozta. Ez egyben az első mesterséges neurális hálónak is tekinthető. Ezt az említett neurális hálót az 1. egyenlet írja le:

$$g(x_1, x_2, x_3, \dots, x_n) = g(x) = \sum_{i=1}^n x_i$$
$$y = f(g(x)) = 1, \text{ ha } g(x) \geq \theta$$
$$y = f(g(x)) = 0, \text{ ha } g(x) < \theta,$$

ahol g a bemenet és szimbolizálja az emberi agyban található dendriteket. Ez a g összegzést végez el, θ pedig egy küszöbérték paraméter. Végül f dönti el, hogy mi a kimenet értéke g és θ alapján. Ezt a részét az 1. egyenletnek küszöbérték logikának is nevezzük.

A következő mérföldkő az volt, amikor Hebb ajánlott egy neuropszichológián alapuló tanuló algoritmust 1949-ben [2]. Ez az algoritmus egy felügyeletlen tanulási szabályon alapszik, ami tanító halmazokból nyeri ki a statisztikai információkat és adathasonlóság alapján osztályoz. 1952-ben Arthur Samuel fejlesztett egy dámapjáték programot, ami képes volt a következő lépések megtételére [3]. 1956-ban John McCarthy volt az első, aki használta a mesterséges intelligencia (MI, angolul: artificial intelligence, AI) kifejezést [4]. Ő tekinthető emiatt az MI atyjának.

Napjainkban számos feladathoz használunk MI-t. Az MI célja, hogy az emberi intelligenciához hasonló, vagy jobb képességekkel rendelkezzen. Az MI képes olyan összetett feladatok megoldására, amelyeket korábban csak az emberi intelligencia tudott elvégezni. Ehhez pedig olyan algoritmusok, modellek és rendszerek kerülnek kifejlesztésre, amelyek lehetővé teszik a gépeknek az adatok feldolgozását, tanulását és döntéshozatalát. Emiatt az MI alkalmazása széleskörű és sokféle területen megtalálható, beleértve az ipart [5], beszédfelismerést [6, 7],

egészségügyet [8–10], autonóm rendszereket [11, 12], oktatást [13], pénzügyi elemzést [14], kiberbiztonságot [15], képfeldolgozást [16, 17], és még sok mást.

Az MI-nek viszont tanulnia szükséges a feladatok elvégzése előtt. Simon a következőképpen definiálta a tanulást [18]: *„A tanulás olyan változásokat jelent a rendszerben, amelyek abban az értelemben adaptívak, hogy lehetővé teszik a rendszer számára, hogy a következő alkalommal ugyanazokat a feladatokat vagy az azonos populációból származó feladatokat hatékonyabban és eredményesebben végezze el”*. Azonban, tanulási módszerek kiválasztása előtt két kérdést kell megválaszolni: Mit tanuljunk? Hogy tanuljunk? Ezekre a kérdésekre pontos választ kell adni, hiszen a tanulási folyamat tekinthető a problémamegoldás egy fajtájának is [19, 20].

Tekintve, hogy tanulásról és problémamegoldásról van szó, nagyon fontos az adathalmaz megfelelő felépítése és óvatos kiválasztása. A nem megfelelő, vagy hibás adathalmaz, valamint a külső ingerek hatása más és más eredményt produkálhat. Ez annak köszönhető, hogy ezek a változók befolyásolhatják a tanulási folyamatot.

Kutatásomban a külső ingerek hatására fókuszálok, mivel a célom az, hogy megvizsgáljam ezeknek az előbb említett hatását, valamint erejét a tanulási fázisra mérve. Ennek érdekében elkészítettem egy háromdimenziós virtuális környezetet, amiben egy MI-vel felruházott karakter építkezni próbál. Tehát, az említett karakter célja nem más, mint, hogy egy tornyot építsen a virtuális világban található objektumokból, majd végül megmássza azt. Mindeközben a felhasználó valós időben változtathatja a virtuális világ paramétereit, valamint olyan hatásokat hozhat létre, ami befolyásolja a tanulást, és ezzel együtt az építkezést.

A dolgozatom a következő struktúrával rendelkezik. A 2. fejezetben egy szakirodalmi elemzést mutatok be az MI tanulási módszereiről, valamint a szakirodalom elemzés során létrejött kutatási kérdésemet és hipotéziseimet definiálom. A 3. fejezetben bemutatom a kutatás során felhasznált módszereket. A 4. fejezetben az eredmények kerülnek részletezésre, miközben az 5. fejezetben átbeszélésre kerülnek. A 6. fejezetben pedig a konklúziókat vonom le.

2. Szakirodalmi áttekintés

Az MI-nek négy különböző tanulási módszerét különböztetjük meg. Az első a felügyeletlen tanulás (unsupervised learning), a másik a felügyelt tanulás (supervised learning), a harmadik az ajánló rendszerek (recommender systems), a negyedik pedig a megerősítéses tanulás (reinforcement learning). Az irodalmi áttekintés során kitérek mind a négy tanulási módszerre. A 2.1-es szekcióban a felügyeletlen tanulást vizsgálom, a 2.2-es szekcióban pedig a felügyelt tanulást. A 2.3-as szekcióban az ajánló rendszereket mutatom be, míg a 2.4-esben a megerősítéses tanulás módszert tekintem át, miközben pár szóval kitérek az imitációs tanulás fajtájára is. A szakirodalmi áttekintés után pedig vázoló a kutatási kérdésemet és hipotéziseimet a 2.5-ös szekcióban.

2.1. Felügyeletlen tanulás

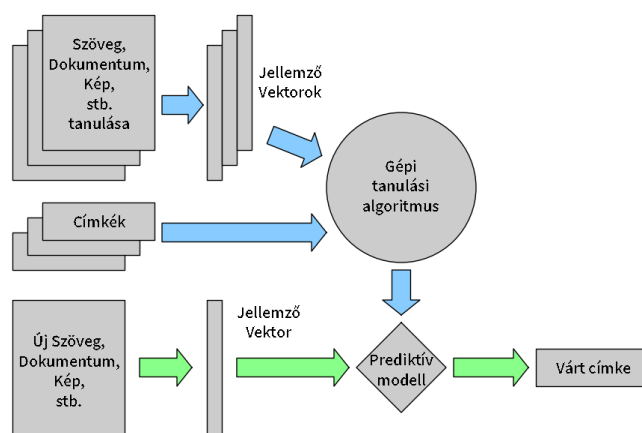
Ez a tanulási folyamat a számított kimenet és a valós kimenet összehasonlításán alapul, azaz a tanulás a hiba kiszámítását és a hiba beállítását jelenti a várt kimenet elérése érdekében. Más szóval, a felügyeletlen tanulás célja, hogy megpróbáljunk rejtett struktúrát találni címkézetlen adatokban. Mivel a tanulónak adott példák nem rendelkeznek címkével, ezért nincs hiba- vagy jutalomjelzés a lehetséges megoldás értékeléséhez. Például, adott lóerejű személygépkocsik adatsorát adjuk meg a tényleges árakkal, majd a felügyelt algoritmus célja, hogy több helyes választ állítson elő annak érdekében, hogy új személygépkocsi vásárlása esetén mi lenne annak az ára.

A felügyeletlen gépi tanulási algoritmusok megvizsgálhatják például cégek ügyfeleinek adatbázisát, és automatikusan felfedezhetik a piaci szegmenseket. Ebből adódóan automatikusan csoportosíthatják az ügyfeleket ezekbe a szegmensekbe. Következtetésképpen, a cég automatikusan és hatékonyabban tudja értékesíteni a különböző piaci szegmenseket. Ebben az esetben nem ismerjük előre, hogy melyek az adott piaci szegmensek, vagy hogy melyik ügyfél melyik szegmenshez tartozik. Hasonlóképpen, a felügyeletlen gépi tanulási algoritmusok automatikusan azonosíthatják több weboldal (pl. Facebook, Twitter) felhasználói körét, illetve azok ismerőseit. Valamint azt is, hogy mely embercsoportok ismerhetik egymást [21].

Manapság, a digitalizáció során az elektronikus kórlapok is elterjedtté váltak, így ha az orvosi feljegyzéseket adattá alakítjuk, a betegségeket jobban megérthetjük felügyeletlen tanulás használatával [22, 23]. A bioinformatika tudomány területén is alkalmazzák, hiszen a biológusok rengeteg adatot gyűjtenek, többek között a génszekvenciákról, DNS-szekvenciákról, amiknek elemzésében is nagy segítséget nyújt [24–27].

2.2. Felügyelt tanulás

A felügyelt tanulás az egy olyan gépi tanulási feladat, ahol a bemeneti objektumok és a kívánt kimeneti érték egy modellt képeznek. Más szóval, a felügyelt tanulásban minden példa egy pár, amely egy bemeneti objektumból (általában egy vektorból) és egy kívánt kimeneti értékből áll. Egy felügyelt tanulási algoritmus elemzi a tanítási adatokat, és egy függvényt állít elő, amely felhasználható új példák esetén is a leképezésére [28]. Az 1. ábrán szemléltetem ezt a folyamatot.



1. ábra: A felügyelt tanulás folyamata. A [29]-ben található ábra alapján készült.

A felügyelt tanulási módszert nagyon sok helyen alkalmazzák. Gyakran e-mailek kezelésénél fordul elő. Képes személyre szabott, automatikus válaszadásra; e-mailek mappákba való rendezésére; valamint levélszemét (spam) észlelésére is [30, 31]. Ezeken felül kéz-, arc-, és beszédfelismerésre is alkalmazható [32, 33]. Szintén gyakran használt tanulási módszer különböző jelek feldolgozásánál [34]. A természetes nyelvfeldolgozó módszerek (natural language processing, NLP) is ezt a tanulási módot használják [31, 35, 36]. A legfontosabb felhasználási területe az adatok elemzésénél van, hiszen dimenziócsökkentésre is alkalmazzák [37–40]. Következésképpen a felügyelt tanulás segítségével egyszerűbbé válhat az adatok kiértékelése, és a konklúziók levonása.

2.3. Ajánló rendszerek

Az ajánló rendszerek célja, hogy jelentőségteljes ajánlásokat adjon egy rendszer felhasználói körének olyan termékekről, amelyek érdekelhetik őket. [41] Az utóbbi időben az ajánló rendszerek fejlesztésére mesterséges intelligenciát használnak. Azonban megfelelő tanuló algoritmust találni a feladatra kihívásokkal teli, mivel számtalan fellelhető a szakmában. Ahhoz, hogy a lehetőségek tárházából a legjobbat tudjuk kiválasztani, először fel kell ismerni trendeket. Mindezek mellett azonosítani és beépíteni kell a nyitott kérdéseket. Illetve, valamelyképpen segíteni az új felhasználókat abban, hogy tájékozódjanak a rendszerben. Az ajánló rendszereket tartalom alapú, kollaboratív vagy hibrid szűrés alapján lehet csoportosítani [42]. Ajánló rendszert használ például a Netflix is, amely a “Tetszik” és “Nem tetszik” visszajelzések segítségével más hasonló érdeklődési körben mozgó felhasználók számára ad olyan ajánlásokat, melyeket például mi nézünk. A 2. ábrán látható egy példa, hogy hogyan működik egy ajánlórendszer.

		Termékek					
		1	2	...	i	...	m
Felhasználók	1	5	3		1	2	
	2		2				4
	:			5			
	u	3	4		2	1	
	:					4	
	n			3	2		
		a	3	5		?	1

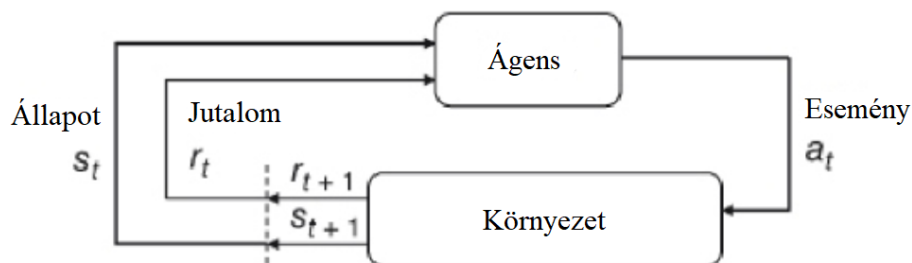
2. ábra: A felhasználói értékelések mátrixa. Minden cella $(r_{u,i})$ megfelel egy u felhasználó által adott, i termék értékelésének. Az algoritmusnak meg kell állapítania, hogy a hiányzó $r_{a,i}$ cellában milyen értékelést fog adni az a felhasználó. A [41]-ben található ábra alapján készült.

Ahogy sejthető a korábbiakból, egyik leggyakoribb felhasználási területe a személyre szabott reklámok ajánlása [43]. Ezt többféleképpen teheti meg. Figyelheti azt is, hogy egy felhasználó mit ír be egy keresőmotorba, hogy milyen weboldalt olvas a felhasználó, vagy akár azt is, hogy milyen videót néz a felhasználó. Ezekon felül, ha például a felhasználó szövegesen értékeli egy általa megvásárolt terméket, akkor az érzelmeket is kinyerheti belőle [44]. A kinyert érzelmek alapján pedig több terméket tud ajánlani. Ha például egy vásárlónak nem tetszett egy tojásfőző, akkor a jövőben megpróbál nem ajánlani hasonló terméket és/vagy márkát. Érdeemes

lehet megjegyezni, hogy olyan tanuló módszerek is léteznek, amik figyelik a felhasználó internetes kattintási szokásait [45].

2.4. Megerősítéssel tanulás

A megerősítéssel tanulás módszerét a viselkedéssel foglalkozó pszichológia inspirálta. A megerősítéssel tanulás egyedi módszere nem más, mint hogy az ágensek jutalmat (vagy büntetést) kapnak bizonyos tettek után. Az ágensek célja, hogy maximalizálják a jutalmukat. A megerősítéssel tanulás folyamata a 3. ábrán látható.



3. ábra: A megerősítéssel tanulás folyamata. A [46]-ben található ábra alapján készült.

Egyik gyakori példája a megerősítéssel tanulásnak a számítógépes játékok területén fordul elő. Vannak olyan emberek, akik az MI ellen szeretnek játszani, és nem valódi játékosok ellen. Ilyen célra olyan MI-t kell alkotniuk a játékfejlesztőknek, ami leköti a játékosokat [47]. De nem csak számítógépes játékok területén használják ezt a típusú tanulást: A személygépjármű forgalmat is próbálják jósolni vele [48]. Sőt, a tőzsdén is alkalmazzák [49].

A megerősítéssel tanulás egyik fajtája az imitációs tanulás. Az imitációs tanulás során az ágens az emberi viselkedést igyekszik utánozni [50]. A megfigyelésekből készít egy leképezést a tettek felé az ágens. A megfigyelés is két különböző típusú lehet: Az első, hogy az ágens megfigyel egy oktató karaktert, míg a második, hogy a saját hibáiból tanul. Az imitációs tanulást a robotika egyik kulcsfontosságú részének lehet tekinteni [51]. Elmondható, hogy az ágens interakcióba lép a környezetével [52]. Ezt a fajta tanulást gyakran használják a séta tanítására [53], szimulációk alatt [54], robotok kézmozgásainak kalibrálására [55, 56], földi járművek tanítására [57, 58], repülő eszközök tanítására [59], és számítógépes játékok esetén is [60].

2.5. Kutatási kérdés és hipotézisek

A szakirodalomból látható, hogy több típusú gépi tanulási módszert tudunk alkalmazni, mindegyiket egy egyedi célra. Éppen ezért mindegyiknek megvan az előnye és a hátránya. Kutatásom során az imitációs tanulóval foglalkozom, ezért ennek a vizsgálatát és befolyásolását

tűztem ki célként. Mivel ez a fajta tanulás nagy gyakorisággal használható valós környezetben és szimulációban is, ezért hasznosnak érzem a kivizsgálását. Ebből adódóan a következő kutatási kérdést tűztem ki: *Mennyire tudják külső tényezők befolyásolni a gépi tanulás ezen formáját?*

A kutatási kérdés megválaszolása érdekében kettő nullhipotézist (H) fogalmaztam meg. Ezek a következők:

- H1: *A felhasználói beavatkozások nem befolyásolják az ágens által elért jutalom mennyiségét.*
- H2: *A felhasználói beavatkozások nem befolyásolják az ágens építkezést tanuló szokásait.*

A két nullhipotézis elfogadása vagy elutasítása érdekében létrehoztam egy háromdimenziós virtuális környezetet, ahol egy imitációs tanulást alkalmazó ágenst befolyásolhat különböző környezeti erőkkal a felhasználó. A kutatáshoz felhasznált módszerek a 3. fejezetben olvashatók.

3. Felhasznált módszerek

Ebben a fejezetben bemutatom a kutatás során felhasznált módszereket. A 3.1-es szekcióban a korábban olvasható szakirodalmi áttekintés módszereit mutatom be. A 3.2-es szekcióban pedig a korábban említett virtuális környezet fejlesztését fejtém ki. Az elkészült környezet a 3.3-as szekcióban kerül bemutatásra. A 3.4-es szekcióban az adatgyűjtési módszereket írom le. Végül, a 3.5-ös szekcióban pedig az adatelemzéshez használt módszereket mutatom be.

3.1. A szakirodalmi áttekintés módszere

Ahogy korábban olvasható volt, első lépésként egy szakirodalmi áttekintést valósítottam meg. A vizsgálat során három fő tudományos adatbázist használtam. Ez a három a Web of Science, IEEE Xplore, illetve a Google Scholar volt. Segítségükkel megvizsgáltam a témához tartozó releváns, nemzetközi publikációkat. Az alábbi kulcsszavakat, és azok kombinációit használtam: „artificial intelligence”, „learning”, „behavior”, „adaptability”, „adaptation”, valamint „virtual environment”. A szakirodalmi kutatáshoz a PRISMA 2020 elveit használtam. Ez a következőt jelenti:

A Web of Science 4874230 publikációt eredményezett ezekre a kulcsszavakra, míg az IEEE Xplore 934302 találatot adott. A Google Scholar pedig nagyjából 7700000 találatot eredményezett. Cím és kulcsszavak alapján mindegyik adatbázisból megvizsgáltam az első 200-at, tehát összesen 600 publikációt néztem át. Ezeket kimentettem egy Excel fájlba, ahol cím szerint rendeztem őket. Így könnyen megtaláltam a duplikációkat. Következőnek 182 duplikációt töröltem az említett fájlból. A maradék 418 publikáció absztraktját megvizsgáltam, azonban ezekből 341 nem illett a dolgozatomba. Végül, 77 közlemény került áttanulmányozásra. Ezekből pedig 60 publikációra való hivatkozás került be a dolgozatom végleges verziójába. Ezekre a publikációkra épül a szakirodalmi elemzésem, valamint a dolgozatom bevezető fejezete. A hivatkozott publikációk mennyiségébe nem számolom bele a weboldal típusú hivatkozásokat.

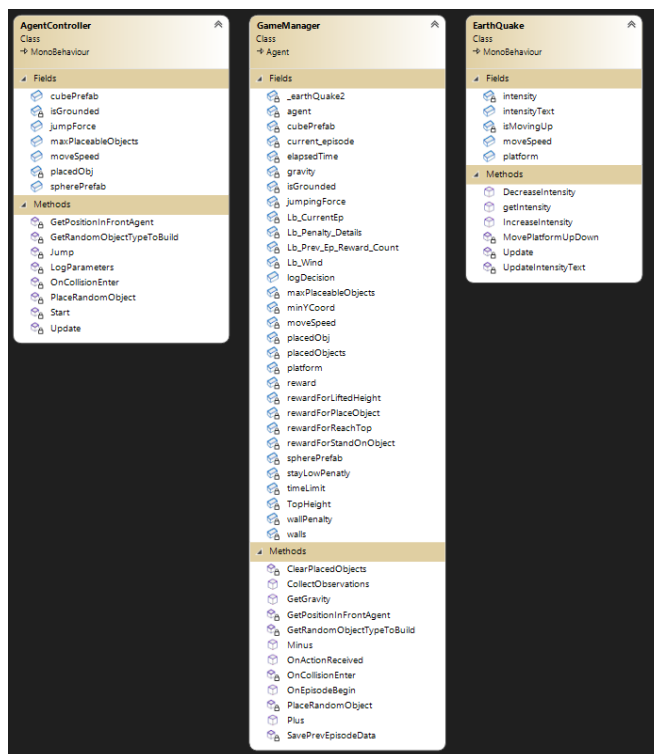
3.2. A virtuális környezet fejlesztése

A virtuális környezetet a Unity 3D (2021.3.16f1) játékfejlesztő motor segítségével valósítottam meg. A kódot C# nyelven írtam a Visual Studio integrált fejlesztői környezet segítségével. Az implementációhoz az objektumorientált programozási (OOP) paradigmát használtam. Ebből adódóan a létrejött osztályok az adatokat és a hozzájuk tartozó műveleteket egyesítik, így lehetőséget adnak a hatékony, moduláris, és az újra felhasználható kód írására. Végeredményképpen létrejött egy egyszerű, és érthető kód, ami könnyen karbantartható és bővíthető.

Az implementációs fázis elején különböző környezeti változókat hoztam létre. + A környezeti változók mellett a Unity Machine Learning Agents Toolkit (ML-Agents) nevű bővítményt használtam, hogy létrehozzak intelligens ágenseket [61]. Ezen felül elkészítettem a Unity környezetében két különböző előregyártott objektumot, egy kockát és egy gömböt. Ezeket használhatja az ágens az építkezéshez.

3.2.1. Az alkalmazás struktúrája

Ahogy korábban említettem, a kódban mindent OOP szemlélet alapján készítettem. Külön metódusba raktam például az ugrást, objektumok törlését és elhelyezését is, mivel ezeket többször is felhasználom a kódban. Az OOP szemlélet egyik előnye, hogy az alkalmazás struktúrája könnyen ábrázolható. Ezt a 4. ábrán szemléltetem egy osztálydiagram segítségével.



4. ábra: Az alkalmazás osztálydiagramja.

Ahogy a 4. ábrán is látható, angol neveket adtam a változóknak és a metódusoknak is. 26 különböző metódust használtam a virtuális környezet implementálása során. Ezekből három fontosabbat szeretnék megemlíteni. Ez a három az “OnEpisodeBegin”, “CollectObservations”, valamint az “OnActionReceived” nevet kapta.

Az “OnEpisodeBegin” metódus minden új epizód elindulásakor fut le. Kitöröl minden előre létrehozott építményt, és véletlenszerűen pozicionálja az ágens karakterét. A véletlenszerű elhelyezésre azért van szükség, mert, ha az ágens rájön egy nyerő stratégiára, akkor állandó ciklusba kerülhet és ugyanazt a mozgást végezheti. A kutatási céлом miatt kiküszöböltem az ágens ezen stílusú viselkedését. A “CollectObservations” metódusban azt definiálom, hogy az ágens milyen információk tudatában legyen egy epizód kezdésekor. Itt a saját pozícióját, a falak és az egyéb objektumok helyzetét kapja meg. Az “OnActionReceived” metódusban megadom azokat az értékeket, amelyek segítségével tudja a mesterséges intelligencia befolyásolni a virtuális környezetben mozgó karakterét. Emellett itt definiálom a jutalom rendszert, és figyelek arra, hogy csak a megengedett ideig fusson a szimuláció.

3.2.2. ML-Agents

Az ML-Agents egy nyílt forráskódú bővítmény, amely lehetővé teszi, hogy intelligens ágensek jöjjenek létre Unity környezetben a PyTorch felhasználásával. Ebből adódóan, ezek az ágensek játékok és egyéb szimulációk alatt használhatók, megfigyelhetők. Mindemellett elég robusztus, hiszen 2D, 3D, valamint virtuális- és kiterjesztett valóság alapú környezetekben is használható. Mind játékfejlesztők, mind kutatók is használják a bővítmény által biztosított Python API-t ágensek létrehozásához. Ezek az ágensek használhatnak megerősítő tanulást, imitációs tanulást, neuroevolúciót, vagy akár más gépi tanulási módszereket.

Ezeket a képzett ágenseket több célra használhatják, beleértve a non-playable character (NPC) típusú karakterek tanítását. Továbbá, használható viselkedés irányítására (különböző beállításokban, például több ágenszt tartalmazó, versenyszerű környezetekben), a játékok készítésének automatizált tesztelésére és különböző játéktervi döntések kiértékelésére a megjelenés előtt. Az ML-Agents Toolkit egy központi platformot biztosít, ahol az AI fejlesztések elemezhetők a Unity környezetében, majd hozzáférhetővé tehetők a kutatói és játékfejlesztői közösség számára. A kutatás során használt ágens karaktere a 3D Balls sablon projektből származott.

3.2.3. Környezeti változók

Az alkalmazásban megtalálhatók a virtuális környezetet befolyásoló függvények is, amelyek a szél és a földrengés erejének módosítását végzik el. Ezeket a felhasználó állíthatja be

valós időben a felületen elhelyezett grafikai elemek segítségével. Az 5. ábrán szemléltetem ezeket a grafikai elemeket.



5. ábra: A környezeti változókat szabályzó grafikai elemek.

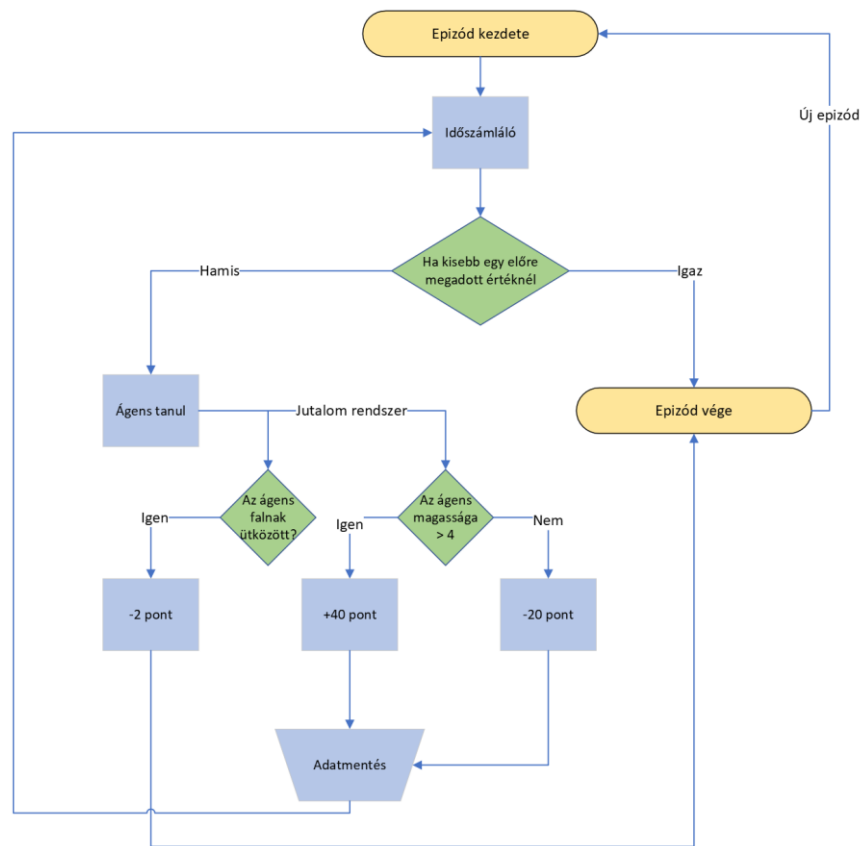
Ahogy az 5. ábrán is látható, négy különböző gomb került elhelyezésre. Ezen gombok funkciói a következők:

- Az első gomb megnyomása növeli 0,1 egységgel a szél erejét,
- A második gomb megnyomása csökkenti 0,1 egységgel a szél erejét,
- A harmadik gomb megnyomása növeli 0,01 egységgel a földrengés erejét,
- A negyedik gomb megnyomása csökkenti 0,01 egységgel a földrengés erejét.

A szél a következőképpen működik a virtuális környezetben. Logikáját tekintve egy oldalirányú gravitációról beszélünk. Az első két gomb megnyomására változik az x tengelyen lévő gravitáció ereje. A maradék két gomb megnyomására pedig elkezd fel-le mozogni a platform a beállított lépésközzel. A platform mozgásának gyorsaságán és a fel-le mozgás között eltelt időt is lehet állítani. A mérések alatt ez 0,01 és 1 volt.

3.2.4. A virtuális környezet mögötti logika

A korábbiakban már említettem, hogy a kutatásomban a megerősítéses tanulás elvét használom. Ezt az elvet a feladat összetettsége miatt kiegészítettem egy imitációs tanulás technikával. Ez alapján létrehoztam a virtuális környezet mögött rejlő rendszert, ami a 6. ábrán látható.



6. ábra: A virtuális környezet mögötti logika.

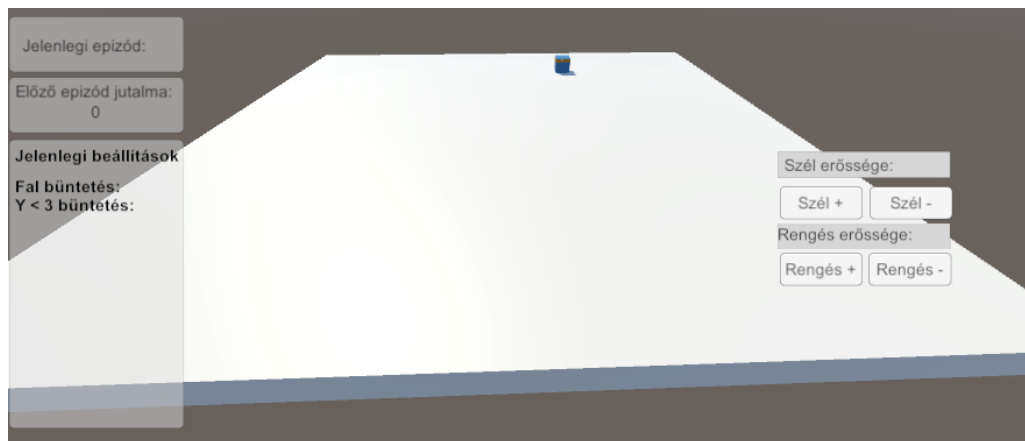
Ahogy a 6. ábrán is látható, az első lépés mindig az epizód kezdete. Ez az a pont, ahol megtörténik az adott epizód kalibrációja. Ezután pedig az alkalmazás háttérében elkezdem számolni az epizód elkezdése óta eltelt időt. A számlálás egészen 30 másodpercig tart. Ha az eltelt idő elérte ezt az értéket, az epizód véget ér, és egy új kezdődik el. Ellenkező esetben az ágens folyamatosan tanul a virtuális környezetben. Itt jön képbe a korábban megemlített jutalom rendszer. Vegyük az ágens magasságát a virtuális környezetben. Ha a magassága négy egységnél kisebb, akkor azt állapítjuk meg, hogy ágens nem jutott magasabbra. Ezt a korlátot azért állítottam négy egységre, mert így belekalkulálom, hogy esetlegesen ugrik egyet az ágens az epizód végén. Ebben az esetben az ágens jutalma -20 pont. Ha tanulás közben valamely láthatatlan falhoz ér, szintén -20 pontot kap. Ilyenkor szigorúbb a büntetés, mivel az epizód is véget ér. Ha az ágens

több, mint négy egység magasra jutott, akkor 40 pont jutalmat kap. Ha egy építő objektumnak a tetejéhez ért, az eredménye mellé további 200 pontot kap. Ezzel az epizód végi jutalommal motiválom a cselekvésre.

Az ágens a következő féleképpen építkezhet: Két típusú objektum közül fog véletlenszerűen kapni egyet maga mellé (az x tengelyen elcsúsztatva). A két objektum típus a gömb és a kocka, utóbbit kétharmados valószínűséggel kapja meg. Azt is meghatároztam, hogy legfeljebb mennyi objektumot tud lerakni a játéktérre. A tesztekben lerakható objektumok mennyisége 10 volt.

3.3. Az elkészült virtuális környezet

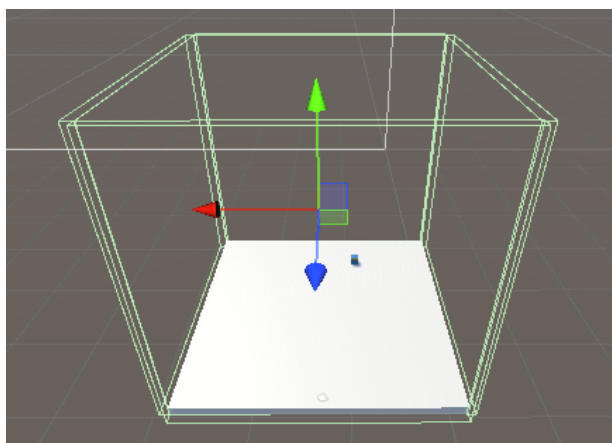
A kutatás céljából adódóan a virtuális környezet felépítése egyszerű. Található benne egy 30x30 egységnyi platform. Ezen a platformon helyezkedik az ágens karaktere, aki mozogni és építkezni tanul a különböző epizódok lefolyása alatt. Az ágens karaktere egy kék kockaként van reprezentálva a virtuális környezetben. Ahogy korábban is elhangzott, egy felhasználói felülettel is rendelkezik ez a virtuális tér. A 7. ábrán látható az elkészült virtuális környezet, illetve a felhasználói felület. Bal felül látható a jelenlegi epizód száma, valamint az előző epizódban szerzett jutalom. Alatta a jelenlegi beállítások láthatók, ahol az ágenst érő büntetések értékei szerepelnek. Jobb oldalt pedig a korábban említett környezeti változókat befolyásoló felületi elemek láthatók.



7. ábra: A virtuális környezet és a felhasználói felület.

A tesztelési fázis során gyakran tapasztaltam, hogy az ágens lefutott a platformról és ennek következtében leesett. A probléma kiküszöbölésére több láthatatlan, 30 egység magas falat helyeztem el a platform körül. Ezzel a megoldással, valamint a korábban említett negatív

jutalommal az ágens nagyobb fókuszot helyezett az építkezésre. A láthatatlan falak a 8. ábrán vannak szemléltetve.



8. ábra: A láthatatlan falak a virtuális térben.

3.4. Adatgyűjtés

A mérésekhez egy laptopot használtam, ami a következő konfigurációval rendelkezett. Ez a laptop egy kilencedik generációs processzort tartalmazott (Intel Core i7 9750H) egy NVIDIA GTX 1650-es videokártya mellett. Az említett számítógép 8 GB memóriával rendelkezett.

Az adatokat a virtuális környezet automatikusan elmenti és kiírja egy fájlba. Ennek a módszere a következő: Első lépésként minden döntés után hozzáfűzi az aktuális állapotot egy *string* típusú változóhoz. Ez kiíratásra kerül minden epizód kezdetén egy, szöveget tartalmazó fájlba (.txt). Annak érdekében, hogy az adatokat egyszerűen lehessen kezelni, tabulátorok használatával választottam el. Ezt használva, könnyen beilleszthetővé vált a Google Sheets alkalmazásba. A 9. ábrán szemléltetem a kimentett adatok formátumát.

	A	B	C	D	E	F	G
1	Epizód	Szél	Földrengés	Mozgás	Ugrás	Építés	Jutalom
2	1	0	0	2	1	1	
3	1	0	0	2	1	1	
4	1	0	0	2	1	1	
5	1	0	0	2	1	1	
6	1	0	0	2	1	1	
7	1	0	0	4	1	1	
8	1	0	0	4	1	1	
9	1	0	0	4	1	1	
10	1	0	0	4	1	1	
82	1	0	0	2	0	0	
83							
84							-20

9. ábra: Az adatok formátuma.

A 9. ábrán látszik, hogy milyen adatagok kerülnek rögzítésre egy futás alatt. Ezek a következők: Epizód száma, szél és földrengés erőssége, mozgás, ugrás, építés, valamint a jutalom. Látható, hogy egy epizód több sorból épül fel, hiszen a meghatározott idő alatt több dolgot is tehet az ágens. Az ágens által elért jutalom mennyisége a többi adattal ellentétben csak egyszer jelenik meg a legutolsó döntési sorozat után.

3.4.1. A mérések

A programot 26 alkalommal futtattam le az előre elkészített imitációs fájl segítségével. Ebből 12 alkalommal történt adatmentés. Ezekből a kimentett adatokból került kiválasztásra a négy elemzett adathalmaz. A fájlok hossza eltérő, mivel 79 és 120 közötti döntés lehetséges epizódonként a megadott idő alatt. Példának okáért, az alapbeállításokkal futó mérés adatmentése 128094 sor lett a Google Sheets alkalmazásban. A földrengéseket tartalmazó fájl viszont ennél hosszabb. Ennek okán nem is tudta feldolgozni a Google Sheets a sorokat érintő limitációi miatt. A földrengések esetében csak az első 69 futás fért el a Google Sheets alkalmazásában.

3.4.2. Beavatkozások a tanulási folyamatba

A négy kiválasztott adathalmazból csak egy esetében nem történt felhasználó általi beavatkozás. A maradék három adathalmaz esetében viszont igen. Egyik alkalommal megváltoztattam a szél erősségét 0,1-re a 15. epizód közben. A szél erősségét később növeltem 0,4-re az 56. epizód alatt. A földrengés esetében már az első epizód alkalmával bekapcsoltam a nehezítést 0,01-re, és nem változtattam meg ezt az értéket. Végül, az utolsó futás alkalmával a szelet és a földrengést is bekapcsoltam az első epizód alatt, majd pedig figyeltem az ágens viselkedését.

3.5. Adatelemzés

Az adatelemzés az R statisztikai programcsomag használatával készült RStudio grafikus környezetben. Az adatok importálását a *rio* nevű csomagban található függvényekkel végeztem el. Az importálás befejezése után az ágens viselkedését kétféleképpen elemeztem: Jutalmak szerint, valamint az alapján, hogy tanult-e építkezni. A jutalmak elemzése egyszerűbben történt, mivel epizódonként csak egyet kap az ágens. Az építkezési szokások vizsgálatához pedig megszámláltam epizódonként, hogy tanult-e építkezni az ágens, vagy sem. Az R nyelvben található *table* függvény segítségével kontingenciamátrixokat hoztam létre, amik az összehasonlítás alapjául szolgáltak.

Az elemzéshez, illetve a hipotézisek vizsgálatához $\alpha = ,05$ értéket választottam. Minden vizsgálat előtt elemeztem az adathalmazok eloszlásait hisztogramokkal, valamint a Shapiro-Wilk normalitás próbával. Ahogy a következő fejezetben is olvasható, az adathalmazok nem követték a Gauss eloszlást, ezért vagy nemparametrikus próbákat alkalmaztam, vagy olyanokat, amelyek kevésbé érzékenyek az adathalmazok eloszlására.

Az adathalmazok összehasonlításához a Wilcoxon rangösszeg próbát alkalmaztam. Következőnek megvizsgáltam az ágens által elért jutalom mennyiségére, valamint az építkezést tanuló szokásaira mért felhasználói beavatkozások hatását. Ehhez a Kruskal-Wallis próbát alkalmaztam az adathalmazok eloszlásai miatt. Hogy alaposabban megértsem a Kruskal-Wallis próba eredményét, megállapítottam, hogy mely párok között van szignifikáns különbség. Ehhez a Wilcoxon rangösszeg próba többszörös összehasonlítási verzióját használtam Holm korrekcióval.

Az ágens építkezést tanuló szokásainak vizsgálatára Pearson χ^2 próbáját, valamint a Cochran-Mantel-Haenszel χ^2 próbáját is használtam. Előbbi módszerrel megállapítottam az adathalmazok arányai közt lévő különbségeket, aminek erejét Cramér V értékével mondtam meg. A Cochran-Mantel-Haenszel χ^2 próbát a felhasználói beavatkozások megléte esetén alkalmaztam, mivel ekkor háromdimenziós kontingenciamátrixokkal dolgoztam. Továbbá, az is megállapítható vele, hogy létezik-e szignifikáns összefüggés egy bináris változó (építkezni tanul-e az ágens), valamint a vizsgált felhasználói beavatkozások között.

A hisztogramok és a vonaldiagramok elkészítéséhez a *ggplot2* csomagot használtam az RStudio környezetben. A *p* értékeket tartalmazó mátrixok kirajzolásához a *corrplot* csomagot alkalmaztam. Az adatok közléséhez az American Psychological Association (APA) szabványát használtam alapul.

4. Eredmények

Ahogy korábban említettem, az eredményeket kétféleképpen értékeltem ki. A kiértékelési módszerekből adódóan több szekcióra bontottam ezt a fejezetet. A 4.1-es szekcióban a jutalmak szerinti kiértékelést mutatom be, míg a 4.2-es szekcióban pedig azt írom le, hogy hányszor tanult az epizódok alatt építkezni az ágens. A 4.3-as szekcióban pedig megemlítek pár kiemelkedő, valamint számomra érdekesnek vélt epizódot.

4.1. Jutalmak

A fő célkitűzés az volt, hogy az ágens karaktere egy tornyot építsen és arra felmásszon. A megfigyelések azt mutatják, hogy képes volt tornyot építeni 100 futáson belül. Azonban, legtöbb esetben arra nem jött rá, hogy fel is kell rá másznia. Őt futtatás alatt kettő esetén tanulta meg a felmászást. Az egyik ilyen eset a környezeti változók bekapcsolása mellett történt. Mindemellett, az ágens tornya vegyesen állt kockákból és gömbökből, amelyek a földrengés hatására kigurultak, így az hamar össze is dőlt.

Érdeemes megemlíteni, hogy amikor az ágens elkezdte első epizódjait, megpróbálta felfedezni a pályát. Ennek eredményeképpen el is találta a láthatatlan falak valamelyikét. Így azonnal megszakadt az epizód, és az ágens -20 pontot kapott. Átlagosan 30-33 epizód kellett ahhoz, hogy az ágens ne akarjon nekimenni a láthatatlan falaknak. Végző soron elmondható, hogy a láthatatlan falak miatt kapta a legkevesebb jutalmat az ágens. A legnagyobb mennyiségű jutalom pedig a felhasználói beavatkozásoktól függött. Ezeket a jutalmakat az 1. táblázatban szemléltetem.

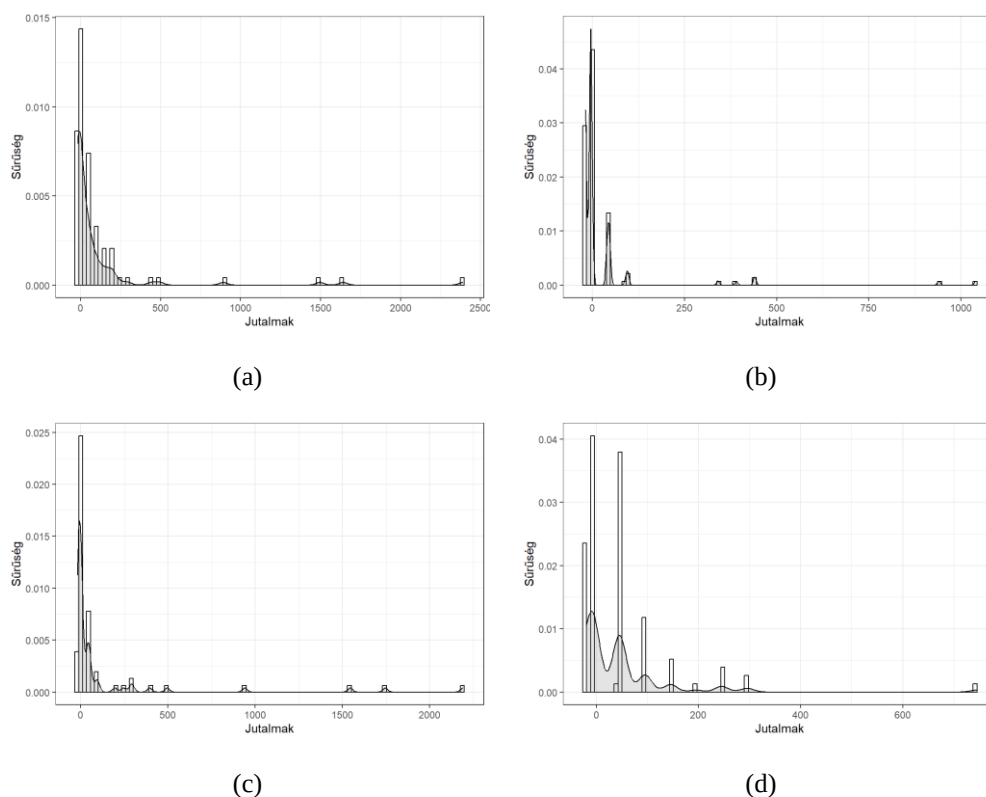
1. táblázat: A jutalmak leíró statisztikái a különböző futások során.

Beavatkozás	Minimum	Maximum	Medián	Átlag	Szórás
Nincs beavatkozás	-20	2390	-5	105,55	338,03
Szél	-20	1040	-5	27,29	140,60
Földrengés	-20	2190	-5	127,90	396,77
Mindkettő	-20	745	40	45,76	99,66

Az 1. táblázatban látható, hogy a legtöbb jutalmat akkor szerezte az ágens, amikor nem történt felhasználói beavatkozás. A legkevesebb jutalmat pedig akkor, amikor mindkét típusú

beavatkozás megtörtént. Az is látható az 1. táblázatban, hogy a szél gátolta meg a legjobban abban, hogy jutalmat szerezzen. Az is elmondható, hogy földrengés esetén átlagosan több jutalmat szerzett, mint amikor nem történt semmiféle felhasználói beavatkozás.

A következő lépésként összehasonlítottam a jutalmak mennyiségét a különböző felhasználói beavatkozások között. Azonban az összehasonlítás előtt az adathalmaz eloszlását is megvizsgáltam. Ez hisztogramokkal, valamint a Shapiro-Wilk normalitás próbával történt. A különböző adathalmazok eloszlásai a 10. ábrán láthatók.



10. ábra: A jutalmak hisztogramjai, csoportosítva a különböző felhasználói beavatkozások alapján: Nincs beavatkozás (a); a szél bekapcsolásra került (b); a földrengés bekapcsolásra került (c); mindkettő bekapcsolásra került (d).

Ahogy a 10. ábrán észrevehető, az adathalmazok nem követték a Gauss eloszlást. Ezt a megállapítást a Shapiro-Wilk normalitás próba is megerősítette az összes adathalmazra ($W_{nincs} = ,368$; $p_{nincs} < ,001$; $W_{szél} = ,318$; $p_{szél} < ,001$; $W_{földrengés} = ,387$; $p_{földrengés} < ,001$; $W_{mindkettő} = ,588$; $p_{mindkettő} < ,001$). Emiatt a jutalmak a Wilcoxon rangösszeg próbával kerültek összehasonlításra. Az eredmények a következőket mutatják:

- A beavatkozás nélküli és a szelet tartalmazó futásokban szerzett jutalmak között szignifikáns különbség van ($W = 8418,5$; $p < ,001$),
- A beavatkozás nélküli és a földrengést tartalmazó futásokban szerzett jutalmak között nincs szignifikáns különbség ($W = 3441,5$; $p = ,978$),
- A beavatkozás nélküli és a mindkét beavatkozást tartalmazó futásokban szerzett jutalmak között nincs szignifikáns különbség ($W = 4867$; $p = ,834$),
- A szelet és földrengést tartalmazó futásokban szerzett jutalmak között szignifikáns különbség van ($W = 3232$; $p < ,001$),
- A szelet és a mindkét beavatkozást tartalmazó futásokban szerzett jutalmak között szignifikáns különbség van ($W = 4480$; $p < ,001$),
- A földrengést és a mindkét beavatkozást tartalmazó futásokban szerzett jutalmak között nincs szignifikáns különbség ($W = 3259$; $p = ,599$).

A kutatás következő lépéseként a különböző felhasználói beavatkozások hatásait vizsgáltam az ágens által elnyert jutalom mennyiségére. Ehhez a vizsgálathoz a korábban megállapított eloszlások miatt a Kruskal-Wallis próbát alkalmaztam. Az említett próba eredményei a következőt mutatják:

- A szélnek van szignifikáns hatása az ágens által elnyert jutalom mennyiségére, $H(4) = 10,863$; $p = ,028$,
- A földrengésnek szintén van szignifikáns hatása az ágens által elnyert jutalom mennyiségére, $H(2) = 16,531$; $p < ,001$,
- A kettő együttesen is szignifikáns hatással rendelkezik az ágens által elnyert jutalom mennyiségére, $H(8) = 52,734$; $p < ,001$.

4.2. Építkezés

Következő lépésként azt vizsgáltam a jutalmak kiértékelése után, hogy hányszor tanult, és hányszor nem tanult építkezni az ágens a különböző felhasználói beavatkozások alatt. A korábbiakhoz hasonlóan, ebben az esetben is négy különböző felhasználói beavatkozási típust különböztetek meg. A 2. táblázatban szemléltetem azt, hogy hányszor tanult építkezni az ágens. A 3. táblázatban pedig azt mutatom be, hogy hányszor nem tanult építkezni a virtuális környezetben.

2. táblázat: Azon cselekedetek leíró statisztikái, amikor az ágens építkezni tanult.

Beavatkozás	Minimum	Maximum	Medián	Átlag	Szórás
Nincs beavatkozás	0	235	5	22,43	44,96
Szél	0	545	0	8,25	41,76
Földrengés	0	215	0	7,68	29,23
Mindkettő	0	240	0	2,27	15,85

A 2. táblázatban látható, hogy volt olyan epizód, ahol egyszer sem akart építkezni tanulni az ágens. Ez mindegyik beavatkozási típusban is előfordult. Az tapasztalható, hogy átlagosan akkor tanult építkezni a legtöbbet az ágens, amikor nem történt felhasználói beavatkozás. Mindkét beavatkozás esetén akart a legkevésbé építkezni. Az is megemlíthető, hogy volt olyan epizód a széles virtuális környezet során, amikor a legtöbbet akart építkezni tanulni. Ettől függetlenül, átlagosan csak 8,25 alkalommal tanult építkezni a különböző epizódok lefolyása alatt.

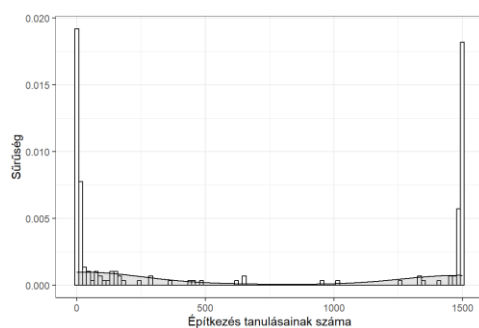
3. táblázat: Azon cselekedetek leíró statisztikái, amikor az ágens nem tanult építkezni.

Beavatkozás	Minimum	Maximum	Medián	Átlag	Szórás
Nincs beavatkozás	51	1500	1495	1283,00	446,13
Szél	0	1523	0	285,10	576,48
Földrengés	0	1500	70	703,90	734,93
Mindkettő	0	1500	0	168,20	461,08

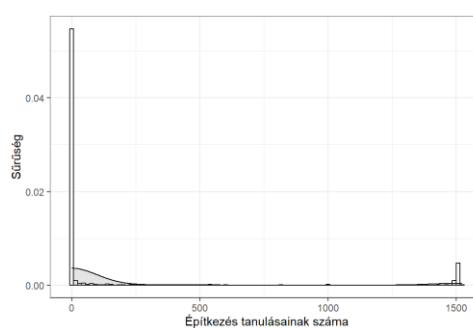
A 3. táblázatban megfigyelhető, hogy átlagosan akkor nem akart a legtöbbször tanulni az ágens, amikor nem történt felhasználói beavatkozás. Észrevehető, hogy átlagos értékek lecsökkentek bármely felhasználói beavatkozás esetén. Szél esetén 77,77%-os csökkenés tapasztalható, míg földrengés esetén ez a mérték csak 45,13%-os. Mindkét típusú beavatkozás esetén pedig 86,89%-os csökkenésről beszélhetünk. A 2. és 3. táblázatokban található adatok

alapján elmondható, hogy felhasználói beavatkozás nélkül akart átlagosan legtöbbször tanulni és nem tanulni. Mindkét esetben a szél bekapcsolása nagyobb hatással volt rá, mint a földrengés. Értelemszerűen mindkettő bekapcsolása esetén csökkentek a legnagyobb mértékben ezek a számok.

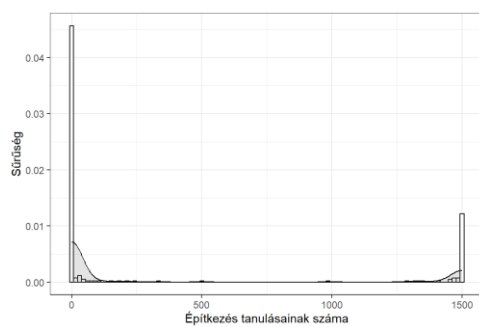
Következő lépésként összehasonlítottam az ágens tanulási szokásait a különböző felhasználói beavatkozások között. Mint ahogy korábban tettem, itt is megvizsgáltam először az adathalmazok eloszlását. Az adathalmazok eloszlásait hisztogramokkal szemléltetem a 11. ábrán.



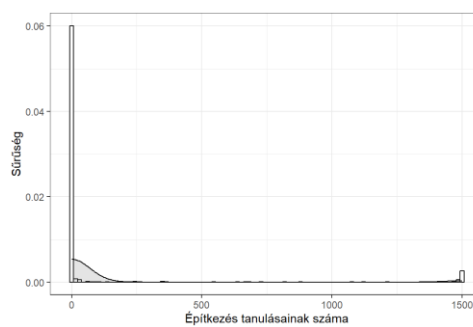
(a)



(b)



(c)



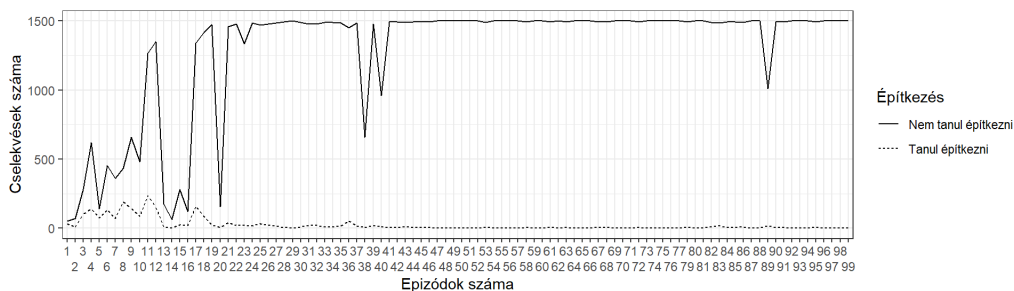
(d)

11. ábra: Az építkezés tanulásának hisztogramjai, csoportosítva a különböző felhasználói beavatkozások alapján: Nincs beavatkozás (a); a szél bekapcsolásra került (b); a földrengés bekapcsolásra került (c); mindkettő bekapcsolásra került (d).

Ahogy a 11. ábrán is észrevehető, a különböző adathalmazok nem követték a Gauss eloszlást. A Shapiro-Wilk normalitás próbával is megvizsgáltam őket. Az említett próba alkalmazása is megerősítette ezt a megállapítást ($W_{nincs} = ,688$; $p_{nincs} < ,001$; $W_{szél} = ,359$; $p_{szél} < ,001$; $W_{földrengés} = ,548$; $p_{földrengés} < ,001$; $W_{mindkettő} = ,257$; $p_{mindkettő} < ,001$). Az eloszlás miatt a Wilcoxon rangösszeg próbát alkalmaztam az ágens tanulási szokásainak összehasonlítására. Az összehasonlítás eredményei a következők:

- A beavatkozás nélküli és a szelet tartalmazó futásokban lévő építkezést tanító szokások között szignifikáns különbség van ($W = 156748$; $p < ,001$),
- A beavatkozás nélküli és a földrengést tartalmazó futásokban lévő építkezést tanító szokások között szignifikáns különbség van ($W = 37911$; $p < ,001$),
- A beavatkozás nélküli és a mindkét beavatkozást tartalmazó futásokban lévő építkezést tanító szokások között szignifikáns különbség van ($W = 262169$; $p < ,001$),
- A szelet és földrengést tartalmazó futásokban lévő építkezést tanító szokások között szignifikáns különbség van ($W = 112261$; $p < ,001$),
- A szelet és a mindkét beavatkozást tartalmazó futásokban lévő építkezést tanító szokások között szignifikáns különbség van ($W = 828414$; $p < ,001$),
- A földrengést és a mindkét beavatkozást tartalmazó futásokban lévő építkezést tanító szokások között szignifikáns különbség van ($W = 270284$; $p < ,001$).

Következőnek azt vizsgáltam, hogy a különböző felhasználói beavatkozások hatása melyik epizódban érvényesül. Először azt az esetet néztem meg, amikor semmilyen beavatkozás nem történt. Ebben az esetben érvényesülő építkezést tanuló szokásokat a 12. ábrán szemléltetem.

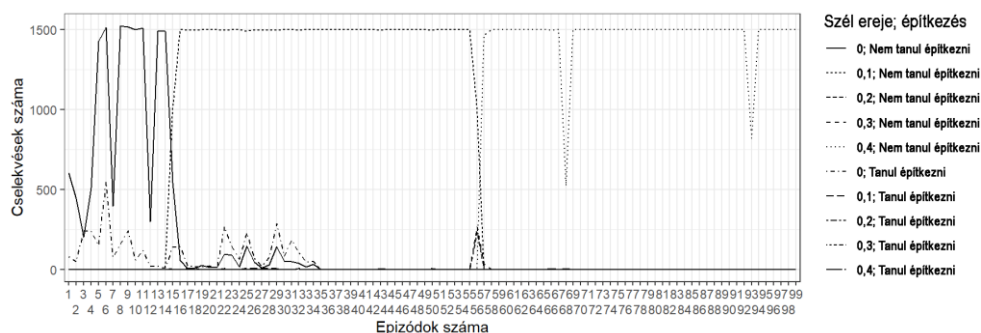


12. ábra: Az ágens építkezést tanuló szokásai felhasználói beavatkozás nélkül.

Ahogy a 12. ábrán is látható, beavatkozás nélkül az első pár epizódban sokkal aktívabban tanul az ágens építkezni. Az idő múlásával pedig elkezdi nem tanulni. χ^2 próbával megvizsgáltam a nem tanulási és a tanulási arányok közti hasonlóságot. Az eredmények azt mutatják, hogy a

kettő arányai között közepes erősségű szignifikáns különbség van, $\chi^2(97, N = 127895) = 16418; p < ,001; V = ,358$.

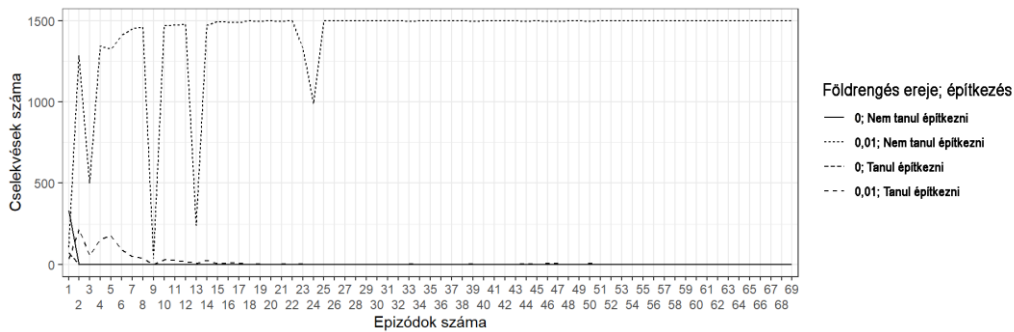
A beavatkozás nélküli adathalmaz elemzése után a következő vizsgálatom a szél befolyásoló erejére terjedt ki. A szeles virtuális környezetben érvényesülő építkezést tanuló szokásokat pedig a 13. ábrán mutatom be.



13. ábra: Az ágens építkezést tanuló szokásai szeles környezetben.

A 13. ábrán látható, hogy az ágens még akkor volt aktívabb, amikor nem tartalmazott szelet a virtuális környezet. Látható, hogy ahol először növeltem a szél erejét, drasztikusan csökkent az ágens aktivitási szintje. Gyakrabban döntött úgy, hogy inkább nem akar tanulni építkezni. Az is látható, hogy a további szélerősség növelésnél megpróbált újra építkezni tanulni, de nem túl gyakran döntött így. A Cochran-Mantel-Haenszel χ^2 próba segítségével megállapítottam, hogy létezik összefüggés az építkezést tanító szokások és a szél ereje között, $\chi^2_{MH}(392, N = 145204) = 311626; p < ,001$. Továbbá, ebben az esetben is meghatároztam a nem tanulási és a tanulási arányok közti hasonlóságot. A vizsgálat eredményéből kiderül, hogy létezik szignifikáns, valamint közepes erejű különbség a kettő kategória aránya között, $\chi^2(98, N = 145204) = 19696; p < ,001; V = ,368$.

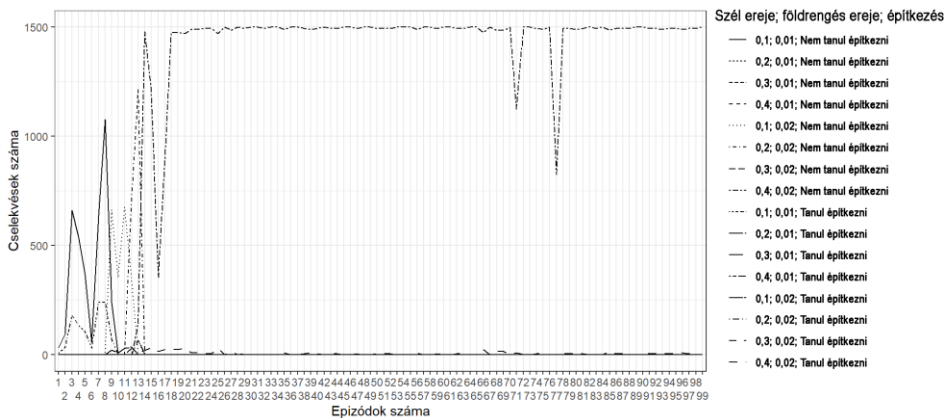
A következő vizsgálatom során a földrengés befolyásoló erejét határoztam meg. A földrengést tartalmazó virtuális környezetben érvényesülő építkezést tanuló szokásokat a 14. ábrán szemléltetem.



14. ábra: Az ágens építkezést tanuló szokásai földrengést tartalmazó környezetben.

Ahogy a 14. ábrán látható, az ágens ebben az esetben is aktívabban kezdte a tanulást, de szintén hamar csökkent az aktivitási szintje. Megállapítottam, hogy létezik összefüggés az építkezést tanító szokások és a földrengés ereje között, $\chi^2_{MH}(68, N = 145204) = 22240; p < ,001$. Földrengés alapú beavatkozás esetében is megvizsgáltam a nem tanulási és a tanulási arányok közti hasonlóságot. Az eredményből kiderül, hogy létezik szignifikáns, közepes erejű különbség a kettő kategória aránya között, $\chi^2(68, N = 98199) = 8608,7; p < ,001; V = ,296$.

Végül, a szél és a földrengés együttes hatását vizsgáltam meg. A két típusú beavatkozást tartalmazó virtuális környezetben érvényesülő tanulási szokásokat a 15. ábrán szemléltetem.



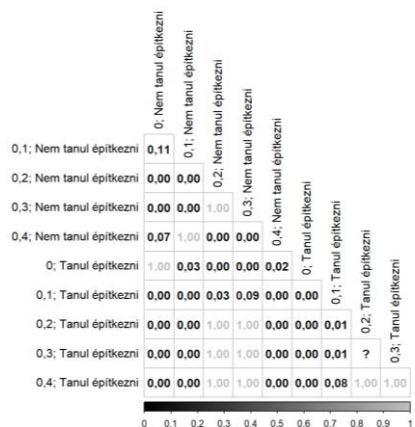
15. ábra: Az ágens építkezést tanuló szokásai szél és földrengést tartalmazó virtuális környezetben.

A 15. ábrán látható, hogy a korábbiakhoz hasonlóan az ágens aktívabb volt a mérések elején. A nyolcadik epizódig gyakran tanult építkezni. Utána csökkent a tanulási hajlandósága. Az is látható, hogy a mérés végéig próbálkozott továbbra is építkezni tanulni. Szél és földrengés együttes használata esetén is megvizsgáltam a nem tanulási és a tanulási arányok közti hasonlóságot. Az eredményből kiderül, hogy létezik szignifikáns, közepes erejű különbség a kettő aránya között, $\chi^2(97, N = 133616) = 16753; p < ,001; V = ,354$.

Következőnek az ágens építkezést tanuló szokásaira mért felhasználói hatásokat vizsgáltam. Az adathalmazok eloszlása miatt a Kruskal-Wallis próbát alkalmaztam. A próba eredményei segítségével a következőket állapítottam meg:

- A szélnek van szignifikáns hatása az ágens építkezést tanuló szokásaira, $H(4) = 4319,7; p < ,001$,
- A földrengésnek szintén van szignifikáns hatása az ágens építkezést tanuló szokásaira, $H(2) = 1535,3; p < ,001$,
- A kettő együttesen is szignifikáns hatással rendelkezik az ágens építkezést tanuló szokásaira, $H(10) = 19857; p < ,001$.

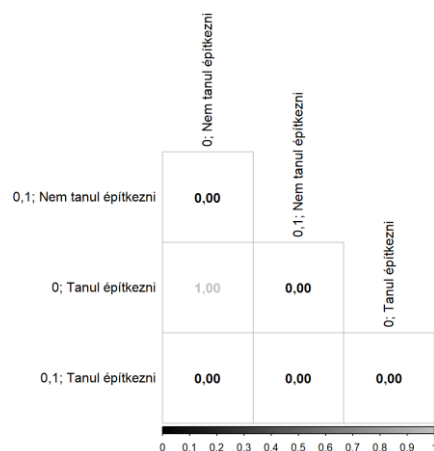
Tekintve, hogy a felhasználói beavatkozások rendelkeznek szignifikáns befolyásoló erővel, ezért megvizsgáltam, hogy mely esetekben léteznek ezek a hatások. Először a szél esetén határoztam meg ezt. A vizsgálathoz a Wilcoxon rangösszeg próba többszörös összehasonlítási verzióját használtam Holm korrekcióval. Az eredmények a 16. ábrán láthatók.



16. ábra: A különbségek a szél ereje és a tanulási szokások között. Az ábrán a p értékek szerepelnek.

Ahogy megállapítható a 16. ábrán szereplő adatokból, a legtöbb pár között észrevehető szignifikáns különbség. Az is látható, hogy a p értékek csak kettő tizedesjegyre lettek kirajzolva az RStudio által. Ez természetesen nem azt jelenti, hogy ezek az értékek egyenlők a nullával, hanem csak azt, hogy meglehetősen kicsik. A “0,3-as erejű szél, tanul építkezni” és a “0,2-es erejű szél, tanul építkezni” párok között pedig nem lehet értelmezni a különbség szignifikanciaszintjét a kevés esetszámból adódóan.

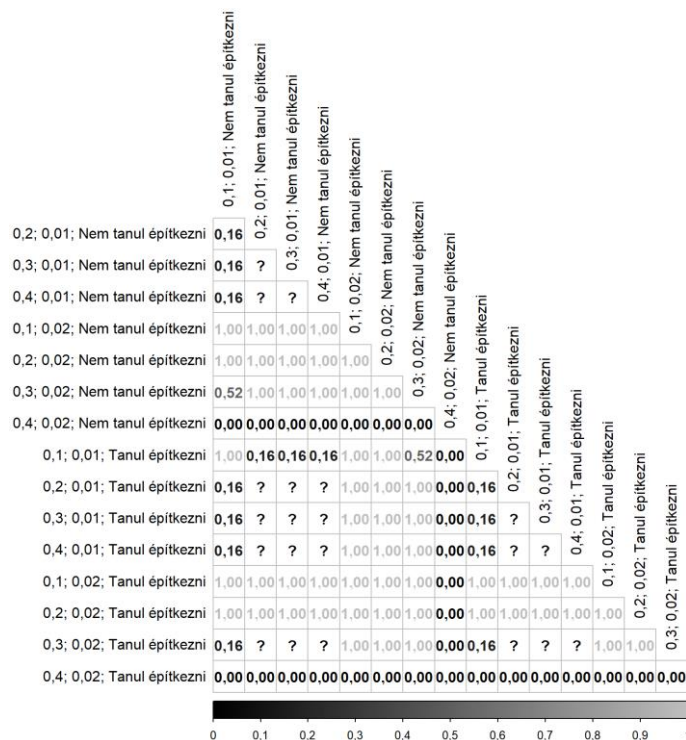
Következőnek egy hasonló vizsgálatot végeztem a földrengés hatásának megfigyelésére. Ebben az esetben is a Wilcoxon rangösszeg próba többszörös összehasonlítási verzióját használtam Holm korrekcióval. Az eredményeket a 17. ábrán szemléltetem.



17. ábra: A különbségek a földrengés ereje és a tanulási szokások között. Az ábrán a p értékek szerepelnek.

Ahogy a 17. ábrán is látható, egy pár kivételével szignifikáns különbségek találhatók az adathalmazban. Hasonlóan a 16. ábrához, itt is csak kettő tizedesjegyre lettek kirajzolva a p értékek. Az említett kivétel pedig a “0-ás erejű földrengés, tanul építkezni” és a “0-ás erejű földrengés, nem tanul építkezni” párok esetén állapítható meg.

Végül, egy hasonló vizsgálatot alkalmaztam a szél és a földrengés együttes hatásának megfigyelésére. Szintén a Wilcoxon rangösszeg próba többszörös összehasonlítási verzióját használtam Holm korrekcióval. A vizsgálat eredményeit a 18. ábrán mutatom be.



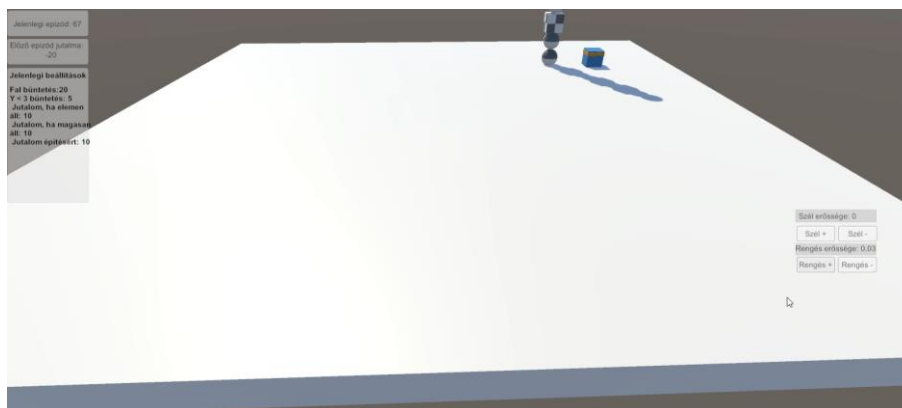
18. ábra: A különbségek a szél ereje, a földrengés ereje és a tanulási szokások között. Az ábrán a p értékek szerepelnek.

Hasonlóan a 16. és 17. ábrához, a 18. ábra esetében is csak kettő tizedesjegyre lettek kirajzolva a p értékek. A vizsgálat eredménye a következőt mutatja meg: Ha egy pár bármely tagjaként szerepel a 0,4-es erejű szél és a 0,02-es erejű földrengés, szignifikáns különbség található köztük és egy másik pár között. Továbbá, ez független attól, hogy tanult-e építkezni az ágens, vagy sem. Más pár között semmilyen esetben sem található szignifikáns különbség az adathalmazban. Az elemzésből az is kiderül, hogy gyakran előfordult nem értelmezhető szignifikanciaszint is a párok között a kevés esetszámból adódóan.

4.3. Kiemelkedő epizódok

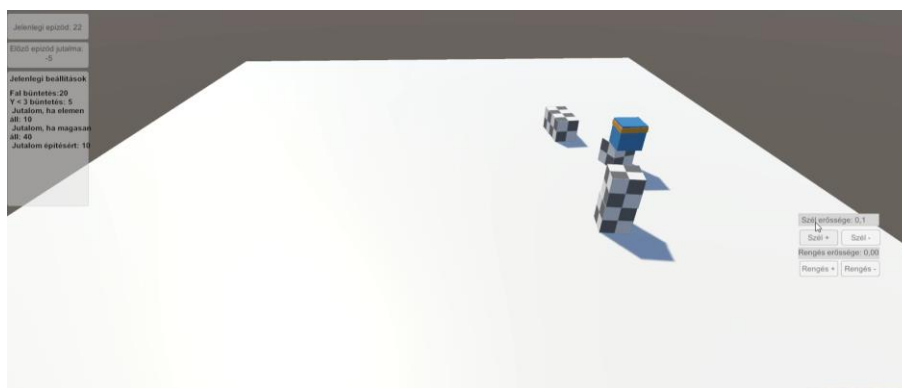
Ahogy említettem a fejezet elején, ebben a szekcióban bemutatok pár számomra érdekes epizódot. Ezek az epizódok egy kisebb rálátást nyújthatnak az ágens tanulási szokásaira. Példának okáért, létezett olyan epizód is, ahol az ágens karaktere két gömbből épített egy tornyot. Az ágens

azonban nem állt meg csak két gömb felhasználásánál. Következőnek egy kockát is elhelyezett rajta, ami az utolsó epizódig nem dőlt össze. Ezt a tornyot a 19. ábrán szemléltetem.



19. ábra: Egy torony, aminek az alapja két gömb.

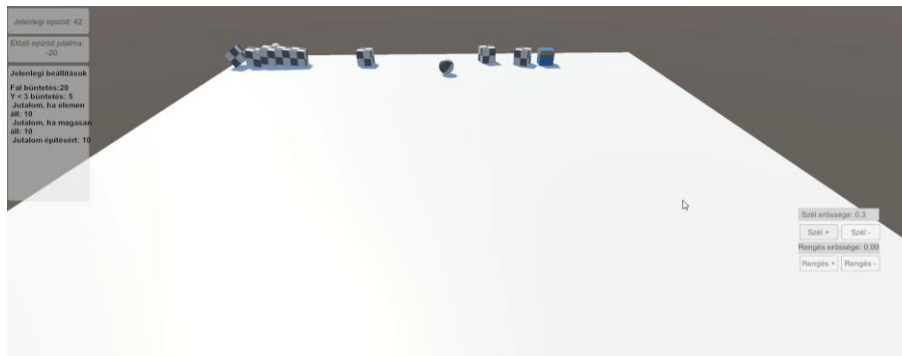
A mérések során előfordult az is, hogy a gyenge szél nem tántorította el az ágens karakterét a tornyok építésében, valamint azok megmászásában. A 20. ábrán egy ilyen példát szemléltetek. Ebben az esetben az ágens meg is mászta az általa épített tornyot, és 45 ponttal zárta az adott epizód végét.



20. ábra: Az ágens a gyenge szél ellenére is felmászott a toronyra.

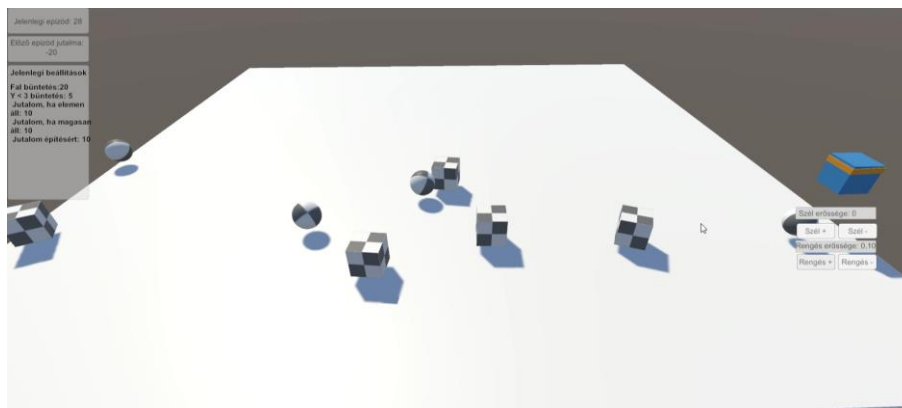
Építkezés közben gyakran előfordult az is, hogy az ágens egyszerre akarta az összes elemét lerakni. Ez sajnos ahhoz vezetett, hogy neki lőttek az építőkockák a láthatatlan falak

valamelyikének. Ebből adódóan az epizód azonnal véget is ért, és az ágens -20 pontot kapott. Ez sajnos csökkentette a tanulási kedvét, mert az objektumok lerakását és ezzel együtt az építkezést negatív jutalommal kapcsolta össze. Erre látható egy példa a 21. ábrán.



21. ábra: Az építkezés során a falnak lökték az ágenst az objektumok.

Következőnek egy olyan példát hoztam, ahol a földrengés miatt összedőlt a torony. Sikerült egy olyan pillanatot megörökítenem, amikor a torony már összedőlt, de az ágens még zuhan a platform felé. Ez a 22. ábrán látható.



22. ábra: A földrengés miatt összedőlt a torony.

5. Átbeszélés

A tanulmányban bemutatott eredmények alapján elmondható, hogy választ kaptam a kutatási kérdésekre. Következő lépésként, mindkét nullhipotézist elutasítottam, és a két alternatív hipotézist fogadtam el. Következtetésképpen ez azt jelenti, hogy a felhasználói beavatkozások hatással vannak az ágensre. Mind a jutalom mértéke, mind az építkezést tanító szokások is változnak tőlük.

Az alkalmazott imitációs tanulási módszer során megfigyeltem, hogy a jutalom mértéke kulcsfontosságú tényező az MI tanulásában. A szél, a földrengés, vagy akár mindkettő szignifikáns hatással rendelkezik és befolyásolja a tanulási szokásokat. Ezáltal az MI más mennyiségű jutalmat érhet el: A jutalom mértéke csökkent a környezeti hatások miatt. Ezen környezeti változások nem csak a konkrét tanulási feladatok eredményeire gyakorolnak hatást, hanem kihatnak az MI általános tanulási képességeire is. Az építkezés tanulása is mutatott szignifikáns csökkenést a környezeti tényezők miatt. A jutalom mértékének csökkenése és az építkezés tanulásának visszaesése megmutatja, hogy az MI rendszerek fejlesztése során különös figyelmet kell fordítani a környezeti változásokhoz való alkalmazkodó képességük optimalizálására. Ettől függetlenül, az MI valamennyire alkalmazkodik környezeti változásokhoz, mivel megváltoztatja a tanulási szokásait. Emiatt elmondható, hogy az MI dinamikusan fejlődik.

Ez a fejlődés lehetővé teszi az MI számára, hogy hatékonyabban alkalmazkodjon a változó környezeti feltételekhez. Ezáltal nemcsak a felhasználói beavatkozásokra tud gyorsan reagálni, hanem hosszú távon is képes optimalizálni a tanulási folyamatokat a környezeti változásoknak megfelelően. A kutatásom eredménye nem csak elméleti, hanem gyakorlati alkalmazhatósággal is rendelkezik. A tanulási módszerek további kutatása és fejlesztése ígéretes lehetőségeket kínál az MI területén, főként olyan helyeken, ahol folyamatosan alkalmazkodniuk kell a környezeti változásokhoz.

Mint a legtöbb kutatás, az én kutatásom is rendelkezik korlátokkal. Az elsődleges korlátot a mintavételi epizódok száma határozta meg. A kijelölt 100 epizód mérete már önmagában is jelentős kihívást jelentett a feldolgozandó adatmennyiség szempontjából. Emiatt az adatok kezelése és elemzése is hosszadalmasra nyúlt. Mindemellett a rendelkezésre álló hardver miatt egyszerre csak egy virtuális környezet tudott futni. A felvett videó is akad néhány ponton, ami nehezítette a kulcsfontosságú jelenetek felkutatását. További korlátot jelentett az online térbe való adatfelvitel is, mely gyakran nem zárult sikerrel a Google Sheets limitációi miatt. Előfordult, hogy emiatt többször újra kellett próbálni az adatfelvitelt.

A jövőbeli terveim között szerepel a kutatás időtartamának meghosszabbítása. A jelenlegi kutatásban a 100 epizód mérete korlátozta az adatgyűjtést. A hosszabb időtartamú kutatás lehetővé tenné a minták számának növelését. Ezáltal a kapott eredmények megbízhatóbbak lehetnek, és a kutatás hozzájárulhat az MI tanulásának további, alaposabb megértéséhez. A jövőbeli kutatásokban érdemes lehet kipróbálni alternatív mintavételi módszereket, vagy egyszerűen növelni a meglévő módszerek számát a pontosabb eredmény érdekében. Nem csak alternatív mintavételi módszereket lehetne alkalmazni a jövőben, hanem más tanulási módszerrel is meg lehetne ismételni a kutatást és összevetni az eredményeket. További jövőbeli kutatási lehetőség lehet, hogy egyszerre több ágens is versenyezzen egymással a virtuális környezetben. Ezáltal egy versenyhelyzet is teremthető számukra. Érdemes lehet megfigyelni, hogy ilyen versenyhelyzetben hogyan viselkednek és tanulnak akár egymástól a különböző ágensek. Figyelembe vehető továbbá az is, hogy éppen hány ágens versenyzik egymással a legjobb eredmény vagy jutalom elérése érdekében.

6. Összefoglalás

A kutatásom során az MI tanulási szokásait vizsgáltam imitációs tanulási módszert alkalmazva. Céлом az volt, hogy megfigyeljem, hogy egy dinamikus környezetben hogyan tanul egy ágens, miközben a felhasználói interakciók befolyásolhatják. Ehhez létrehoztam egy virtuális környezetet a Unity 3D játékfejlesztő motor alkalmazásával, valamint az ML-Agents bővítmény segítségével az imitációs tanulási módszert is implementáltam. Miközben az ágens tanul a virtuális környezetben, a felhasználó különböző erejű szelet és földrengést idézhet elő számára valós időben. Az ágens karakterének célja, hogy maximalizálja a jutalmát építkezéssel és toronymászással. Tornyt építhet kockákból és gömbökből. Jutalmat kap, miután megmássza az általa épített tornyt.

Négy különböző adathalmazt használtam az MI tanulási szokásainak megfigyelése céljából. Egy adathalmazban nem történt semmilyen felhasználói interakció, míg a maradékban a szelet, a földrengést, vagy mindkettőt befolyásoltam. A földrengést tartalmazó adathalmaz kivételével mindegyik 100 epizódot tartalmazott. Az eredményeket az RStudio segítségével értékeltem ki.

Az eredmények azt mutatják, hogy a maximális jutalom nagyobb volt a felhasználói beavatkozás nélküli esetben. Azt is megfigyeltem, hogy ágens nem mindig hozott megfelelő és hatékony döntéseket. A kutatás jól reprezentálja a jutalom fontosságát ennél a tanulási modellnél, hiszen a jutalom rendszer optimalizálása során különböző irányból megfigyelt jutalmazás láthatóan más és más eredményt tud produkálni. Emellett megállapítottam, hogy a felhasználók befolyásolhatják az építkezés tanulását is a különböző környezeti hatásokkal.

Összességében elmondható, hogy az MI valamilyen szinten alkalmazkodott a környezeti változásokhoz, ami hosszú távon kulcsfontosságú a rendszerek hatékonyságának és alkalmazkodóképességének fenntartásában. A kutatásom eredménye így nemcsak elméleti, hanem gyakorlati alkalmazhatósággal is rendelkezik, hiszen értékes információt kínál az MI továbbfejlesztése terén, főként a változó és dinamikus környezetekben.

7. Irodalomjegyzék

- [1] Zhang, L., Zhang, B. A geometrical representation of McCulloch-Pitts neural model and its applications. *IEEE Trans. Neural Netw.* **1999**, 10, (4), 925–929, doi: 10.1109/72.774263.
- [2] Kuriscak, E., Marsalek, P., Stroffek, J., Toth, P. G. Biological context of Hebb learning in artificial neural networks, a review. *Neurocomputing* **2015**, 152, 27–35, 2015, doi: 10.1016/j.neucom.2014.11.022.
- [3] Samuel, A. L. Some studies in machine learning using the game of checkers. *IBM J. Res. Dev.* **2000**, 44, (1.2), 206–226, doi: 10.1147/rd.441.0206.
- [4] McCarthy, J., Minsky, M. L., Rochester, N., Shannon, C. E. A proposal for the Dartmouth summer research project on artificial intelligence. Stanford: AI Magazine, 1995.
- [5] Peres, R. S., Jia, X., Lee, J., Sun, K., Colombo, A. W., Barata, J. Industrial artificial intelligence in industry 4.0 - systematic review, challenges and outlook. *IEEE Access* **2020**, 8, 220121–220139, doi: 10.1109/access.2020.3042874.
- [6] Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A. -r., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Process. Mag.* **2012**, 29, (6), 82–97, doi: 10.1109/msp.2012.2205597.
- [7] Graves, A., Mohamed, A.-R., Hinton, G., Speech recognition with deep recurrent neural networks. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, 2013, 6645–6649, doi: 10.1109/ICASSP.2013.6638947.
- [8] Davenport, T. Kalakota, R., The potential for artificial intelligence in healthcare, *Future Healthc. J.* **2019**, 6, (2) 94–98, doi: 10.7861/futurehosp.6-2-94.
- [9] Yu, K.-H., Beam, A. L., Kohane, I. S., Artificial intelligence in healthcare, *Nat. Biomed. Eng.* **2018**, 2, (10), 719–731, doi: 10.1038/s41551-018-0305-z.
- [10] Elbadawi, M., McCoubrey, L. E., Gavins, F. K. H., Ong, J. J., Goyanes, A., Gaisford, S., Basit, A. W., Harnessing artificial intelligence for the next generation of 3D printed medicines. *Adv. Drug Deliv. Rev.* **2021**, 175, (113805), 113805, doi: 10.1016/j.addr.2021.05.015.
- [11] Menouar, H., Guvenc, I., Akkaya, K., Uluagac, A. S., Kadri, A., Tuncer, A., UAV-enabled intelligent transportation systems for the smart city: Applications and challenges. *IEEE Commun. Mag.* **2017**, 55, (3), 22–28, doi: 10.1109/mcom.2017.1600238cm.

- [12] Grigorescu, S., Trasnea, B., GA survey of deep learning techniques for autonomous driving. *J. Field Robotics* **2019**, 37, (3), 362–386.
- [13] Borges, A. F. S., Laurindo, F. J. B., Spínola, M. M., Gonçalves, R. F., Mattos, C. A., The strategic use of artificial intelligence in the digital era: Systematic literature review and future research directions. *Int. J. Inf. Manage.* **2021**, 57, (102225), 102225, doi: 10.1016/j.ijinfomgt.2020.102225.
- [14] Jaksic, M., Marinc, M., Relationship banking and information technology: The role of artificial intelligence and fintech. *SSRN Electron. J.* **2017**, doi: 10.2139/ssrn.3059426.
- [15] Li, J.-H., Cyber security meets artificial intelligence: a survey, *Front. Inf. Technol. Electron. Eng.* **2018**, 19, (12), 1462–1474, doi: 10.1631/fitee.1800573.
- [16] Chollet, F., Xception: Deep learning with depthwise separable convolutions. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2017, pp. 1800–1807, doi: 10.1109/CVPR.2017.195.
- [17] He, K., Zhang, X., Ren, S., Sun, J., Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.
- [18] Simon, H. A., Why should machines learn? In *Machine Learning*, Elsevier, 1983, pp. 25–37.
- [19] Langley, P., Simon, H., Bradshaw, G., Zytkow, J. Scientific Discovery: Computational Explorations of the Creative Processes. Cambridge, MA: MIT Press, 1987.
- [20] Shaw, M. J., Whinston, A. B., Learning and adaptation in distributed artificial intelligence systems. In *Distributed Artificial Intelligence*, Elsevier, 1989, pp. 413–429, doi: 10.1016/B978-1-55860-092-8.50020-4.
- [21] Haider, P., Chiarandini, L., Brefeld, U. Discriminative clustering for market segmentation. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2012, pp. 417–425, doi: 10.1145/2339530.2339600.
- [22] Kononenko, I. (2001). Machine learning for medical diagnosis: history, state of the art and perspective. *Artificial Intelligence in Medicine*, **2001**, 23, (1), 89–109, doi:10.1016/s0933-3657(01)00077-x.

- [23] Magoulas, G. D., Prentza, A. Machine learning in medical applications. In *Advanced course on artificial intelligence*, 1999, pp. 300–307. Berlin, Heidelberg: Springer Berlin Heidelberg, doi: 10.1007/3-540-44673-7_19.
- [24] Caragea, C., Honavar, V. Machine Learning in Computational Biology. *Encyclopedia of Database Systems*, **2009**, 1663–1667.
- [25] Guyon, I., Elisseeff, A. An introduction to variable and feature selection. *J. Mach. Learn. Res.*, **2003**, 3, 1157–1182.
- [26] Tarca, A. L., Carey, V. J., Chen, X.-W., Romero, R., Drăghici, S. Machine learning and its applications to biology. *PLoS Computational Biology*, **2007**, 3, (6), e116, doi: 10.1371/journal.pcbi.0030116
- [27] Sajda, P. (2006). Machine learning for detection and diagnosis of disease. *Annu. Rev. Biomed. Eng.*, **2006**, 8, (1), 537–565, doi: 10.1146/annurev.bioeng.8.061505.095802
- [28] Mohri, M., Rostamizadeh, A., Talwalkar, A. *Foundations of Machine Learning*, The MIT Press, 2012.
- [29] Rehurek, R. Practical Data Science in Python, https://radimrehurek.com/data_science_python/ (Elérhető online, megnyitva: 2023.11.09.)
- [30] Tzanis, G., Katakis, I., Partalas, I., Vlahavas, I. *Modern applications of machine learning*. In *proceedings of the 1st annual SEERC doctoral student conference–DSC*, 2006, pp. 1–10.
- [31] Bratko, A., Filipič, B., Cormack, G. V., Lynam, T. R., & Zupan, B. Spam filtering using statistical data compression models. *J. Mach. Learn. Res.*, **2006**, 7, 2673–2698.
- [32] Valenti, R., Sebe, N., Gevers, T., & Cohen, I. Machine learning techniques for face analysis. In *Machine Learning Techniques for Multimedia: Case Studies on Organization and Retrieval*, 2008, pp. 159–187. Berlin, Heidelberg: Springer Berlin Heidelberg, doi: 10.1007/978-3-540-75171-7_7.
- [33] Al-Hmouz, A. An adaptive framework to provide personalisation for mobile learners, 2012. Doktori értekezés, University of Wollongong.
- [34] Makeig, S., Kothe, C., Mullen, T., Bigdely-Shamlo, N., Zhang, Z., Kreutz-Delgado, K. Evolving signal processing for brain–computer interfaces. In *Proceedings of the IEEE, 100 (Special Centennial Issue)*, 2012, 1567–1584, doi: 10.1109/JPROC.2012.2185009.

- [35] Sebastiani, F. Machine learning in automated text categorization. *ACM Comput. Surv.*, **2002**, 34, (1), 1–47, doi: 10.1145/505282.505283.
- [36] Xiong, L., Poczos, B., Connolly, A., Schneider, J. Anomaly detection for astronomical data, 2010. Carnegie Mellon University, doi: 10.1184/R1/6475475.v1.
- [37] Hou, S., Qiu, R. C., Chen, Z., Hu, Z. SVM and dimensionality reduction in cognitive radio with experimental validation, 2011.. arXiv preprint arXiv:1106.2325.
- [38] Tsagkaris, K., Katidiotis, A., Demestichas, P. Neural network-based learning schemes for cognitive radio systems. *Comput. Commun.*, **2008**, 31, (14), 3394–3404, doi: 10.1016/j.comcom.2008.05.040.
- [39] Tabaković, Ž. A survey of cognitive radio systems. Post and Electronic Communications Agency, *Jurišićeva*, **2011**, 13.
- [40] Hosey, N., Bergin, S., Macaluso, I., O'Donoghue, D. Q-learning for cognitive radios. In *Proceedings of the China-Ireland Information and Communications Technology Conference (CICT 2009)*, 2009. National University of Ireland Maynooth.
- [41] Melville, P., Sindhvani, V. Recommender systems. *Enc. Mach. Learn.*, **2010**, 1, 829–838.
- [42] Portugal, I., Alencar, P., Cowan, D. The use of machine learning algorithms in recommender systems: A systematic review. *Expert Syst. Appl.*, **2018**, 97, 205–227, doi: 10.1016/j.eswa.2017.12.020.
- [43] Broder, A. Z. Computational advertising and recommender systems. In *Proceedings of the 2008 ACM conference on Recommender systems*, 2008, pp. 1–2, doi: 10.1145/1454008.1454009.
- [44] Pang, B., Lee, L., Vaithyanathan, S. Thumbs up? Sentiment classification using machine learning techniques, 2002. arXiv preprint cs/0205070.
- [45] Clarke, B., Fokoue, E., Zhang, H. H. Principles and theory for data mining and machine learning, 2009. New York: Springer, doi: 10.1007/978-0-387-98135-2.
- [46] Schwartz, H. M. Multi-agent machine learning: A reinforcement approach. John Wiley & Sons, 2014.
- [47] Machines, T. G. P. Machine Learning Applications in Computer Games. ICML, 2008.

- [48] Horvitz, E. J., Apacible, J., Sarin, R., Liao, L. Prediction, expectation, and surprise: Methods, designs, and study of a deployed traffic forecasting service, 2012. arXiv preprint arXiv:1207.1352.
- [49] Shen, S., Jiang, H., Zhang, T. Stock market forecasting using machine learning algorithms. Department of Electrical Engineering, Stanford University, Stanford, CA, 2012, 1–5.
- [50] Hussein, A., Gaber, M. M., Elyan, E., Jayne, C. Imitation learning: A survey of learning methods. *ACM Comput. Surv.*, **2017**, 50, (2), 1–35, doi: 10.1145/3054912.
- [51] Schaal, S. Is imitation learning the route to humanoid robots? *Trends Cogn. Sci.*, **1999**, 3, (6), 233–242, doi: 10.1016/S1364-6613(99)01327-3.
- [52] Russell, S. J., Norvig, P. Artificial intelligence a modern approach, 2010. London.
- [53] Berger, E., Amor, H. B., Vogt, D., Jung, B. Towards a simulator for imitation learning with kinesthetic bootstrapping. In *Workshop Proceedings of Intl. Conf. on Simulation, Modeling and Programming for Autonomous Robots (SIMPAN)*, 2008, pp. 167–173.
- [54] Argall, B., Browning, B., Veloso, M. Learning by demonstration with critique from a human teacher. In *Proceedings of the ACM/IEEE international conference on Human-robot interaction*, 2007, pp. 57–64, doi: 10.1145/1228716.1228725.
- [55] Kober, J., Peters, J. Policy search for motor primitives in robotics. *Adv. Neur. In.*, **2008**, 21.
- [56] Mülling, K., Kober, J., Kroemer, O., Peters, J. Learning to select and generalize striking movements in robot table tennis. *Int. J. Robot. Res.*, **2013**, 32, (3), 263–279, doi: 10.1177/0278364912472380.
- [57] Silver, D., Bagnell, J., Stentz, A. High performance outdoor navigation from overhead data using imitation learning. Robotics: Science and Systems IV, Zurich, Switzerland, 2008, 1.
- [58] Ollis, M., Huang, W. H., Happold, M. A bayesian approach to imitation learning for robot navigation. In *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2007, pp. 709–714, doi: 10.1109/IROS.2007.4399220.
- [59] Morales, E. F., Sammut, C. Learning to fly by combining reinforcement learning with behavioural cloning. In *Proceedings of the twenty-first international conference on Machine learning*, 2004, p. 76., doi: 10.1145/1015330.1015384.

- [60] Thureau, C., Bauckhage, C., Sagerer, G. Learning human-like movement behavior for computer games. In *Proc. Int. Conf. on the Simulation of Adaptive Behavior*, 2004, pp. 315–323.
- [61] Juliani, A., Berges, V.-P., Teng, E., Cohen, A., Harper, J., Elion, C., Goy, C., Gao, Y., Henry, H., Mattar, M., Lange, D. Unity: A general platform for intelligent agents, 2018. arXiv preprint arXiv:1809.02627.