Tudományos Diákköri Konferencia  
Egyéb tudományok: Informatika szekció  
Budapest, 2020. november 27.  
Kodolányi János Egyetem

***PLA-felmérés adatainak***

***mesterséges intelligencia alapú elemzése***

***Jelige: kakukmarci***

***2020. november 27.***

# Bevezetés

A TDK-dolgozatban szó lesz a kiindulási helyzetről, a feladatról magáról, a motivációról, a célcsoportokról, a feladat hasznosságáról:

## Kiindulási helyzet

Amikor elkezdtem a témával foglalkozni, még nem volt korábbi tapasztalatom kutatómunkával kapcsolatban. Eredetileg a BPROF-képzés elkezdéséhez szükség volt egy PLA- (vagyis egy Prior Learning Assesment) felmérés elkészítésére.

## Motiváció

Amikor felvetődött egy TDK-dolgozat írásának lehetősége, abban a pillanatban kissé korainak éreztem, de aztán úgy gondoltam semmi módon nem fogok tudni elkerülni hasonló helyzeteket a jövőben. Erre való tekintettel úgy döntöttem, akkor most jött el annak az ideje, hogy elkezdjek gyakorolni, megismerkedjek ennek a világnak a „formáival és színeivel”, valamint megtanuljam, hogy hogyan kell jól, magabiztosan előadni egy témát. A téma választás az adatok rendszerezésére és elemzésére esett, egyrészt azért, mert korábbról volt tapasztalatom ezen területen, másrészt azért, mert ha pl. azonnal egy valódi hasznosságot elváró programozási feladatról lett volna szó, akkor ahhoz több előfeltételnek kellett volna adva lennie.

## A feladat ismertetése

A világ fejlődésével exponenciálisan nő a gyűjtött és tárol adatok mennyisége, és mind ezeknek már csak a tárolása és rendszerezése, oly mértékű feladat, amit majdnem, hogy senki sem vállalna el. Ezért egy menekülési útvonal maradt: az automatizálás. Tehát a feladatom a felmérés elvégzése, annak adatainak összegyűjtése. Majd egy általam választott programban ezek tárolása rendszerezése, elemzése robot által, hiszen ez ember nem volna képes, vagy csak hosszadalmas idő után tudná átlátni és következményeit levonni a sok számérték pókhálójából. Végül a robot elemzéséből levonni a konklúziót, és azt érthető formában tálalnom.

Az IT-aspektusok mellett a mindenkori adattartalom szakmai értelmezése a lényegi kihívás: a PLA-felmérések adatai alapján személyiség-profilírozás és személyiségek csoportosítása végezhető el automatikusan. A személyiség-profil kapcsán megállapításra kerül: mennyire konzisztens egy személyiség. A csoportképzés kapcsán levezetésre kerül: ki kivel összekapcsolva eredményez konzisztens párt (nagyobb csoportot).

## Célcsoport

Elsősorban bárki hasznát veheti az általam végzett munkának, aki fel szeretné mérni például a munkára jelentkező emberek tudását. Vagy bárkinek, aki bajba jutott nagy adathalmazok rendszerezésében vagy elemzésében. De lehet a kedves Olvasó akár csak egy érdeklődő, aki bele akar kóstolni a folyamat menetébe.

Az érdemi célcsoport egy BPROF = üzemmérnök informatikus Hallgató esetében, akinek tantervében az oktatási üzemmel való hatásos és hatékony együttműködés szerepel, értelemszerűen maga az oktatási intézmény (bármely oktatási intézmény). A PLA-felmérések mellett pl. a hallgatói elégedettségmérés rendszere keretében összegyűlt adatok alapján is lehet anonim személyiség-profilokat és összeférhetetlenségeket levezetni.

## Hasznosság

Elképzelésem szerint minden munkáltatónak fejtörést okoz, melyik jelentkezőt vegyék fel, és azok után is megmarad a kérés, hogy a megfelelő ember tölti-e be azt a pozíciót? Egy ilyen felméréssel meg lehet könnyíteni a döntést és quasi biztosra menni, hogy olyan valaki kerül a betöltendő helyre, aki megfelel a pozíció követelményeinek. A személyiség profilírozás és/vagy a csoportképzés automatizálása így haszontermelésre képes a konfliktusok minimalizálásán, az erőforrás-hatékonyság növelésén keresztül.

## A dolgozat szerkezetéről

A dolgozat lényegében klasszikus tagolást használ, hogy könnyebben el lehessen különíteni a egyes részeket és, meg lehessen találni amit az Olvasó keres. Először egy nyitó fejezet van, ahol megismerjük többek között a feladatot, valamint a dolgozat célját. Ezután maga a feladat megvalósításának fejezetei következnek, itt lépésről lépésre végig lehet követni a kidolgozás folyamatát. Végül zár egy összefoglalással és jegyzékekkel. A TDK kapcsán felvállalt feladathoz érdemi szakirodalom azért nem tartozik, mert az ismert megoldások nem érik el a knuth-i elv belépő szintjét, ahol tudás az, ami forráskódba átírható – így a dolgozat látszólag kevés szakirodalommal dolgozik, de a lehetőségekhez képest ez lényegében a maximum is egyben.

# Szakirodalmi háttér

Itt bemutatásra kerül a PLA és a mesterséges intelligencia, mint keretek, amik között, és segítségével létrejött a dolgozat. Ahogy azt az előző alfejezet jelezte, a szakirodalmi alapvetés csak a jelenségkörre (PLA) és az automatizálást lehetővé tevő technológiai alapvetésre tud és kell, hogy koncentráljon:

## A PLA – a keretrendszer

A Prior Learning Assesment eredete messzebbre nyúlik vissza, mint ahogy valaki gondolná. Először még a régi céhek idején alkalmazták a mester és a tanítvány között (<https://en.wikipedia.org/wiki/Recognition_of_prior_learning>). Ennek az volt a lényege, hogy a mester ellenőrzi a tanítványa munkáját, és az elkészült munka minőségét, biztosítva annak versenyképességét és megfelelését az akkori normáknak. Az 1980-as években a szakoktatás szerves részévé vált világszerte (<https://en.wikipedia.org/wiki/Recognition_of_prior_learning>), tehát ha valaki egy bizonyos szakmát szeretne megtanulni, abban specializálódni, akkor számíthat egy PLA-felmérés kitöltésére. Ez a fajta felmérés sok előny hordoz magával. Elsősorban a folyamat egyszerű és könnyen kivitelezhető. Az önbevallásra alapuló felmérés nem különbözteti meg a résztvevőket az alapján, hogy ki, hol, mikor és milyen módon szerezte az ember a tudását, dolgozott-e már korábban vagy sem? Továbbá nem méri fel a résztvevő életkorát, nemét vagy fizikai tulajdonságait. És végül, nem formál ítéletet valakinek a korábbi képesítései, vagy eredményei alapján.

## A mesterséges intelligencia szerepe és potenciálja

Egy hozzá nem értő embernek először egy kézzel fogható, látható robot juthat eszébe a mesterséges intelligencia szó hallatán. Azonban a valóság kissé más. Minden olyan intelligenciát mesterséges intelligenciának nevezhetünk, amit egy gép, program vagy egy mesterségesen létrehozott tudat alkotott meg. Maga koncepció egészen a 17. századig nyúlik vissza. Itt még csak digitális számológépekről volt szó. A 19. századra már eljutottak a programozhatóság létrejöttéig. A 20. század második felében megjelent a mesterséges intelligencia fogalma és az első primitív mesterséges intelligenciák, míg napjainkig már igen kifejlett példányok is születtek. A mi esetünkben az igazság valahol a kettő között van; önmagában nem tesz semmit, viszont, ha „megkérjük”, hogy elemezzen nekünk egy adathalmazt, azt elvégzi és az ember számára is könnyen értelmezhető formában mutatja meg, hogy mire jutott ott, ahol a naiv (nem optimalizált) emberi gondolatok semmire nem jutnak (vö. 10. ábra - korreláció = 0.21 ott, ahol az MI hibátlan egyéni és csoportos mintázatokat talált)…

A dolgozat hátterében alkalmazott MI a hasonlóságelemzés: <https://miau.my-x.hu/myx-free/> - mely a nemzetközi összehasonlító vizsgálatok alapján triviálisan alkalmas a vállalt feladatok elvégzésére (vö. <https://miau.my-x.hu/miau/196/My-X%20Team_A5%20fuzet_HU_jav.pdf>)

# Az elemzés menete

Ebben a fejezetben arról lesz szó, mely úton kell végig haladni a 0-ról az 1-re, vagyis a felmérés elkészítésétől az adatok megfelelő kezelésén át a konklúzióig:

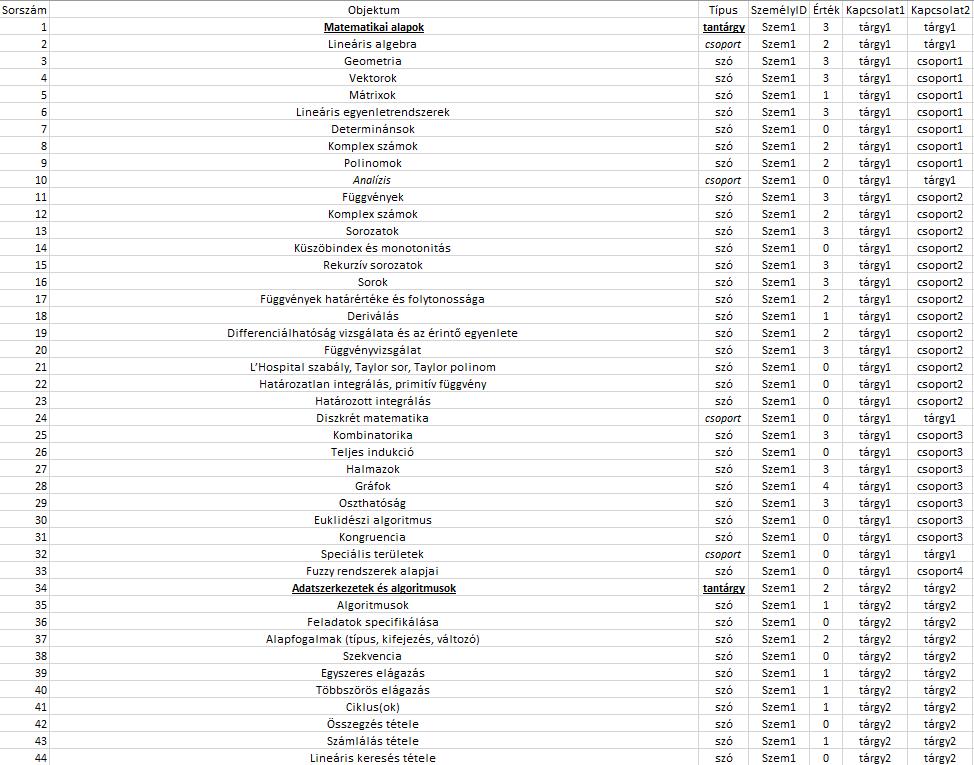
## A felmérés elkészítése

Az első feladatom a felmérés elkészítése majd annak adatainak rendszerezése és kezelése volt. Pár héttel ezután felvetődött a TDK, mely kapcsán kézenfekvő volt, hogy a PLA-felméréssel és a mesterséges intelligencia erre való hatásával fog foglalkozni a dolgozat. Az általam végzett felmérésnek két alanya volt. A felmérésen a BPROF képzés 3 éve (6 szemesztere) alatt tanulandó tantárgyak, a tantárgyakba tartozó kulcsszócsoportok, és azon belül kulcsszavak szerepeltek – a BPROF-képzés akkreditációs anyagainak kivonataként. A felmérésen résztvevőknek ezeket kellet értékelni egy előre definiált értékskála alapján. Ezen skálának hat eleme volt, ami a -1 és a + 5 között mozogtak. Azt gondolná az ember, hogy ezek megfelelnek a közéletben használt iskolai értékeléseknek, de ezúttal kicsit mást jelentenek, bár a számok növekedése továbbá is a tudásszint növekvő mértékét jelzik. A „-1” az adott objektum (tárgy, kulcsszócsoport, kulcsszó) szubjektív feleslegességére utal, a „0” az abszolút semmilyen tudás jele, az „1” az éppen csak annyi, hogy nem idegen a fogalom szintjét szimbolizálja, és 2-től 5-ig pedig az adott objektumról az alany megfelelő tudása és az objektum magabiztos tanítása másoknak szintek között mozog. Tehát megvan minden a felmérés kitöltéséhez és értelmezéséhez. Miután ez elkészült, jöhetett az adatok rendszerezése.

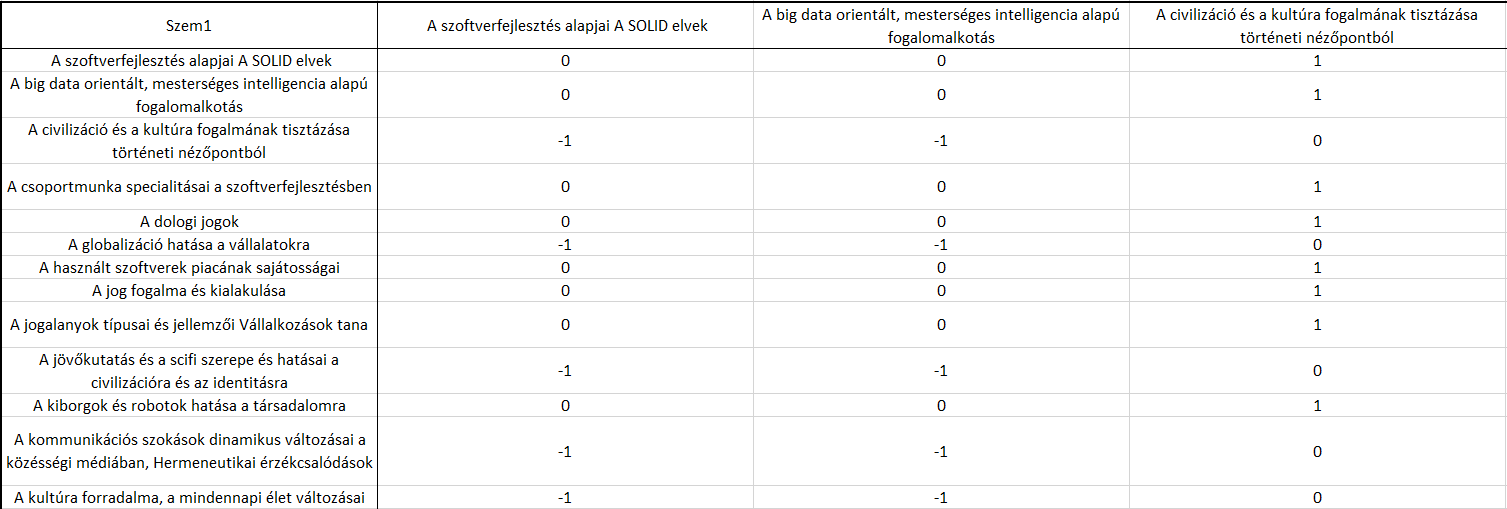
## Az adatok feldolgozása

Kezdetben két program merült fel az adatfeldolgozás támogatására: az első a Microsoft Access, a másik az Excel. Végül az utóbbit választottam, mert ugyanazt meg lehet csinálni benne, amit Access-ben, de ez még többet is tud, ha tudja az ember, mikor hova nyúljon. Mire a megfelelő változatig eljutottunk tizenhárom különböző változaton ment keresztül a fejlesztési folyamat. Míg az elsőkben még a táblázat adatai sem voltak megfelelőek, vagy pontosak, addig a 13. verzió már teljesen alkalmas volt a cél eléréséhez. A fejlesztés oktatásmódszertani szempontból számos típushibára és az állandó és mindenre kiterjedő önellenőrzés kényszerű fontosságára is rámutatott. Az Excel fájl 10+3 munkalapot tartalmaz (URL = <https://miau.my-x.hu/bprof/?C=M;O=D>, ill. <https://miau.my-x.hu/bprof/oam_pla_13%20(1).xlsx>). Az első 10 tartalmazz a tényleges adatokat, számításokat, eredményeket. A másik 3 pedig egy régi eredményt - összehasonlítás céljából. Az első munkalap a PLA nevet viseli. Itt található a teljes adatbázis (1.ábra).

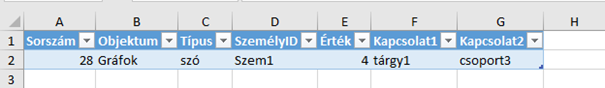
Megjelennek az objektumok, azok sorszáma minden személynél, az objektum típusa (tantárgy, csoport vagy szó) az alanyok által adott értékelés, a személyid, hogy melyik személyre vonatkozik az értékelés és rögzítve vannak az objektumok egymáshoz való kapcsolata (egy szó tartozhat egy csoporthoz és egy tantárgyhoz is). Ez az egy kézzel megírt része az adatbázisnak. A többi adatot és számolást az Excel kimutatás-varázslójával számítottam ki. A második munkalapon (Trianguláris mátrix) három darab trianguláris mátrix található (2.ábra). Ebben három 281\*281 cellányi táblázat található az egyik későbbi eredmény meghatározásához. A harmadik és negyedik munkalapon (maxid1 és maxid2) (3.ábra, 4.ábra) egy-egy kimutatás van, melyek megmondják, hogy az egyes személyek melyik objektumot értékelték a legjobbra, és azt is, hogy mik az objektum egyéb adatai.



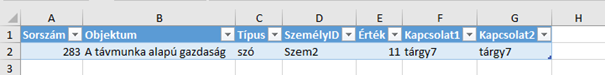
1.ábra - 2\*282+1fejléc sorból álló táblázat egy részlete (Forrás: saját ábrázolás, ahol az értékoszlop mértékegysége: pontszám/kód)



2.ábra – a trianguláris mátrix bal felső részlete (Forrás: saját ábrázolás, ahol az értékek előjeles pontszám-különbségek arra utal, hogy ha – előjel akkor az adott objektum kevésbe ismert a másiknál, a 0 az ugyan akkora tudás jele, míg + előjel az adott objektum nagyobb ismeretére utal a másiknál)



3.ábra – szemid1 legjobban értékelt objektumának ellenőrzése (Forrás: saját ábrázolás – a 4-es érték mértékegysége = A tudásszint, a -1 és 5 intervallumban)



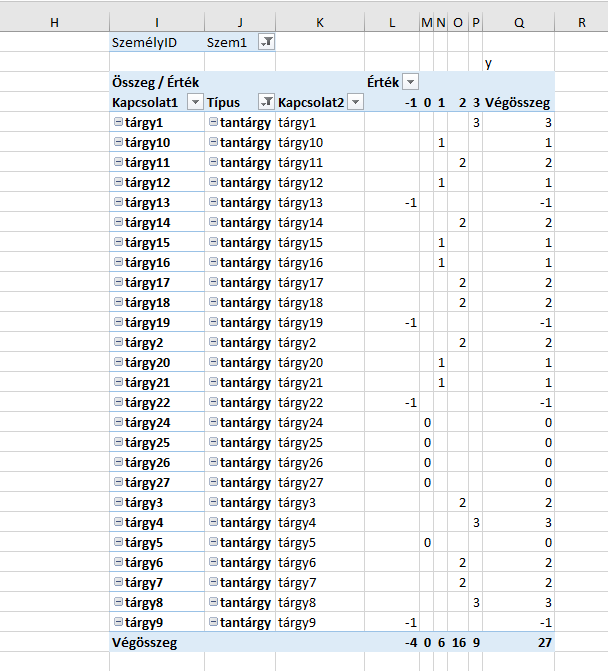
4.ábra – szemid2 legjobban értékelt objektumának ellenőrzése) (Forrás: saját ábrázolás – a 11-es érték mértékegysége = A tudásszint, a -1 és 5 intervallumban)

## A nyers adatoktól a kimutatásokon át az OAM-ig

Az ötödik és hatodik munkalap (szemid1 és szemid2), személyenként lebontja, és kimutatásokon keresztül segít előállítani az OAM-ot (Objektum Attribútum Mátrix). A munkalap elején ismét látható az adatbázis, amit az első munkalapról hivatkozunk át ide is. Az egyelőre elemzendő OAM-ba csak a tantárgyaknak kell bekerülnie ezért a kimutatások sem foglalkoznak mással csak azokkal annak érdekében, hogy gyorsan eljuthassunk a kívánt eredményhez.

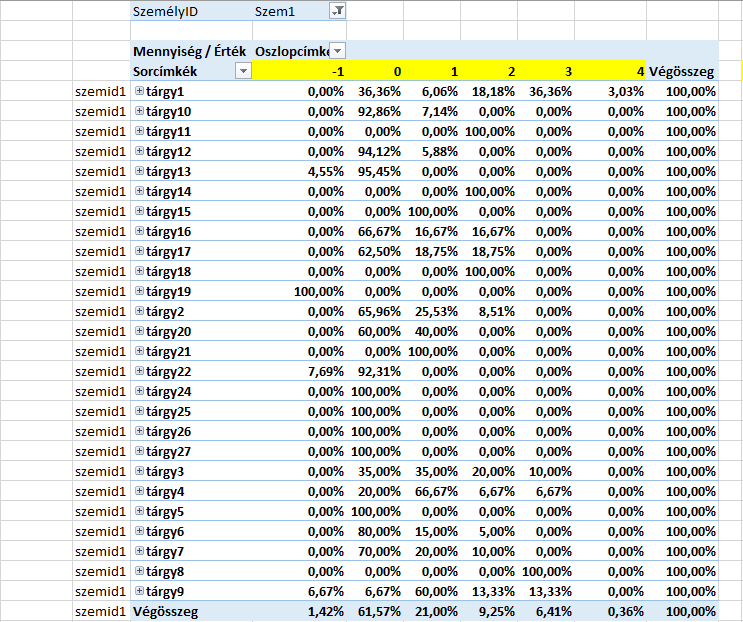
### Az első kimutatás

Meg lehet nézni, hogy az kiválasztott alany melyik tantárgyat milyen tudásszint-értékkel látta el (5. ábra):

5.ábra – szemid1 tantárgyakra adott értékelésének kimutatása (Forrás: saját ábrázolás, ahol a cellák értékeinek mértékegysége az adott oszlopban a fejlécen látható érték/pontszám/kód)

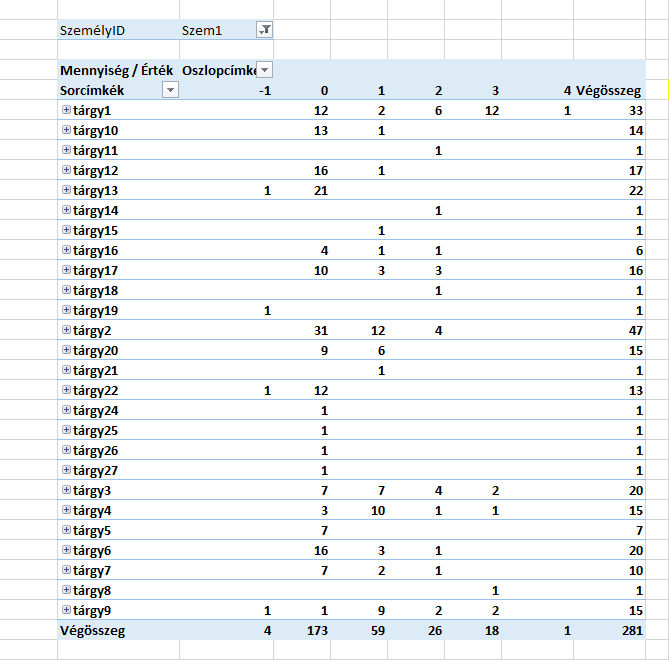
### A második kimutatás

Már nem a tényleges értékeket mutatja, hanem azt, hogy arra a tantárgyra és a tantárgyhoz köthető többi objektumra, hány darab -1, 0, 1, 2, 3, 4 vagy 5-ös értékelés érkezett sor-százalék-nézetben (6.ábra):

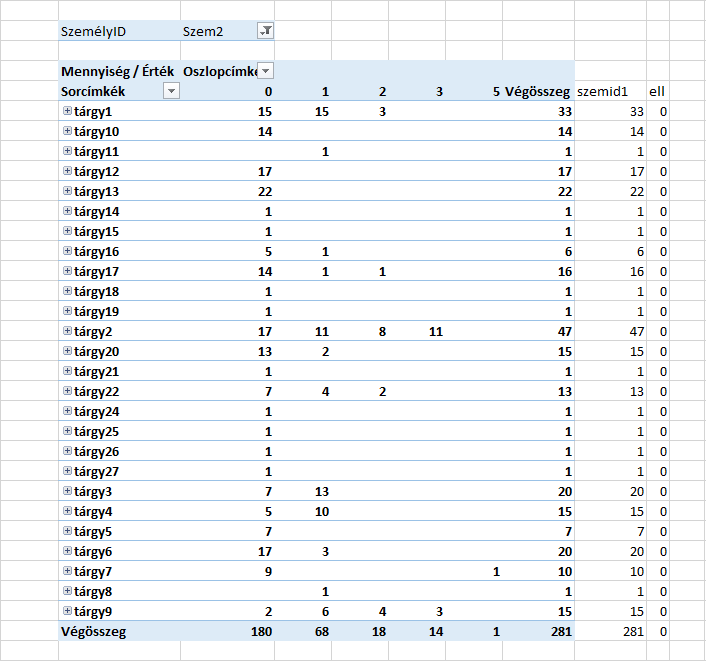


6.ábra – szemid1 tantárgyaként adott érékeléseinek darabszáma százalékosan soronként - (Forrás: saját ábrázolás)

### További kimutatások

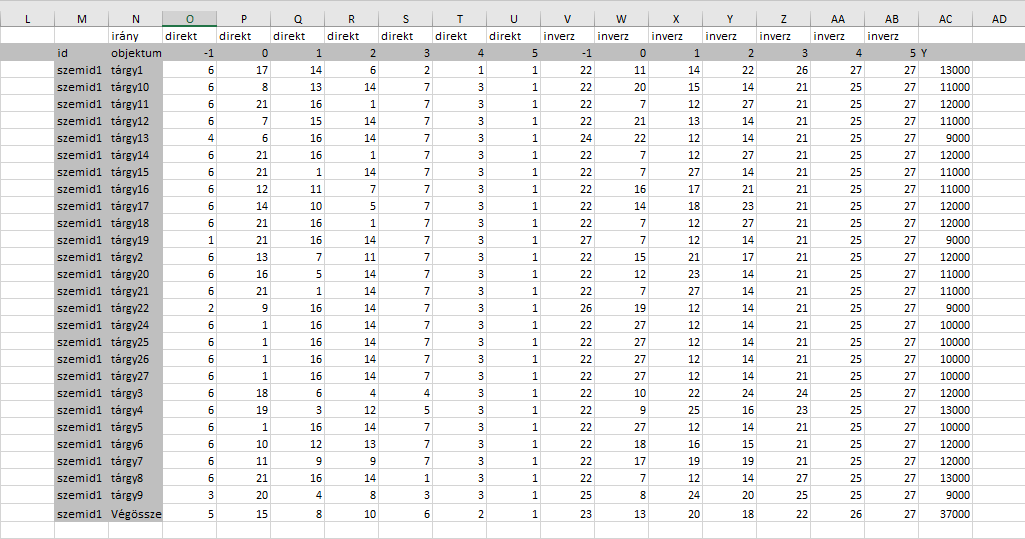
Ellenőrzés céljából született még meg két kimutatás ugyan ezeken a munkalapokon, ahol a két alany által mondott értékelések darabszámát hasonlítjuk össze, könnyen láttatva, ha valahol eltérés van, ami pl. mutatja, ha valaki nem töltötte ki az összes érékelendő mezőt. 

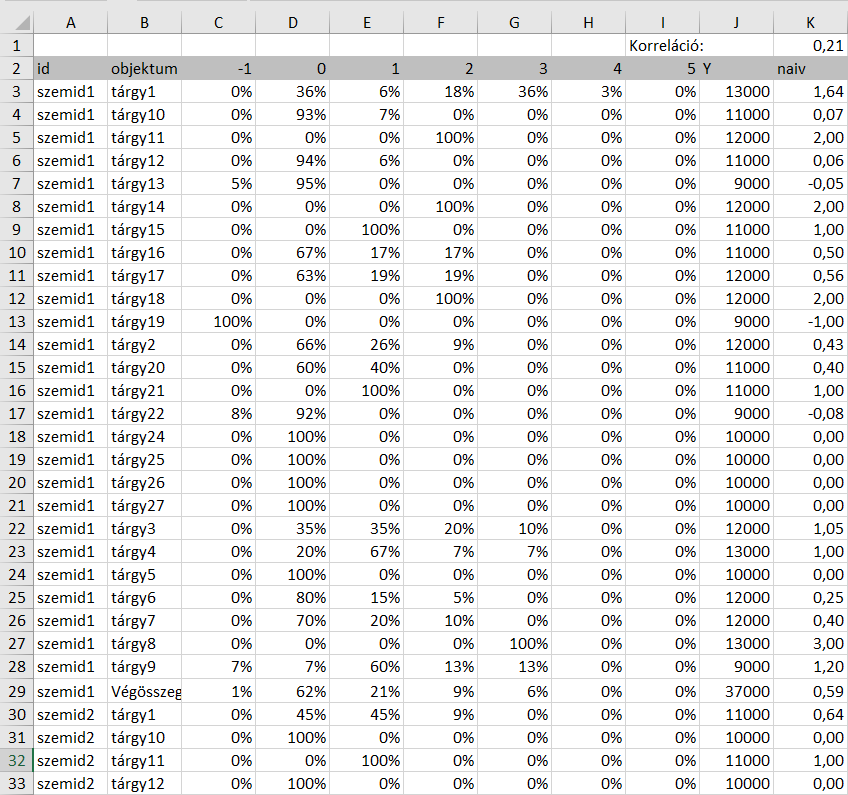
7.ábra – szemid1 tantárgyakként adott értékeléseinek darabszáma (Forrás: saját ábrázolás, ahol a mértékegység a tantárgyan belüli, oszlopfejlécen látható értékek darabszáma – vagyis 281 db önálló értékelés áll rendelkezésre minden alany esetén, melyek tantárgyakhoz eltérő mértékben kötődnek – lehet tehát olyan tantárgy is, mely önmagában értékelendő – pl. tárgy8/11/14/15/18/19/21/24/25/26/27)

8.ábra – szemid2 tantárgyakként adott értékeléseinek darabszáma (Forrás: saját ábrázolás, ahol a mértékegység a tantárgyan belüli, oszlopfejlécen látható értékek darabszáma)

## Az OAM

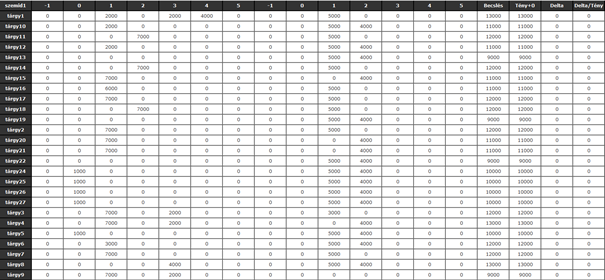
Ezek után a hetedik munkalapon (OAM) létrejött az OAM, amit majd át tudunk majd adni a mesterséges intelligenciának elemzésre (9.ábra). Az OAM a másodikként elkészített kimutatás adatait fogja majd átvenni, tehát egy százalékos táblázatot látunk, amiben az adott érékek darabszáma van összeszámolva (10.ábra). Ezek sor számozva lesznek, hozzákerül egy Y módosítottoszlop, ami úgy jön létre, hogy egy-egy tantárgynál az arra mondott nyers értékelést megszorozzuk 1000-rel majd hozzáadunk 10000-ret annak érdekében, hogy a robot által készített modellt könnyen és gyorsan értelmezni lehessen. Így végül lesz egy táblázatunk, ahol látjuk a 28 tantárgynak (objektumnak) a sorszámozott értékelését (Y\_nyers) direkt úton, ugye itt minél nagyobb a sorszám annál hangsúlyosabb annál a tantárgynál az az érték, vagyis minél nagyobb annál erősebben befolyásolja a személyről kialakuló tudásképet. Egy inverz irányú sorszámozást is kell készíteni itt logikusan minél kisebb egy sorszám annál hangsúlyosabb, vagyis annál nagyobb mértékben befolyásolja az alanyról kialakuló tudásképet.

9.ábra - szemid1 OAM-ja (Forrás: saját ábrázolás – mértékegység: rangsorszámok oszloponként)

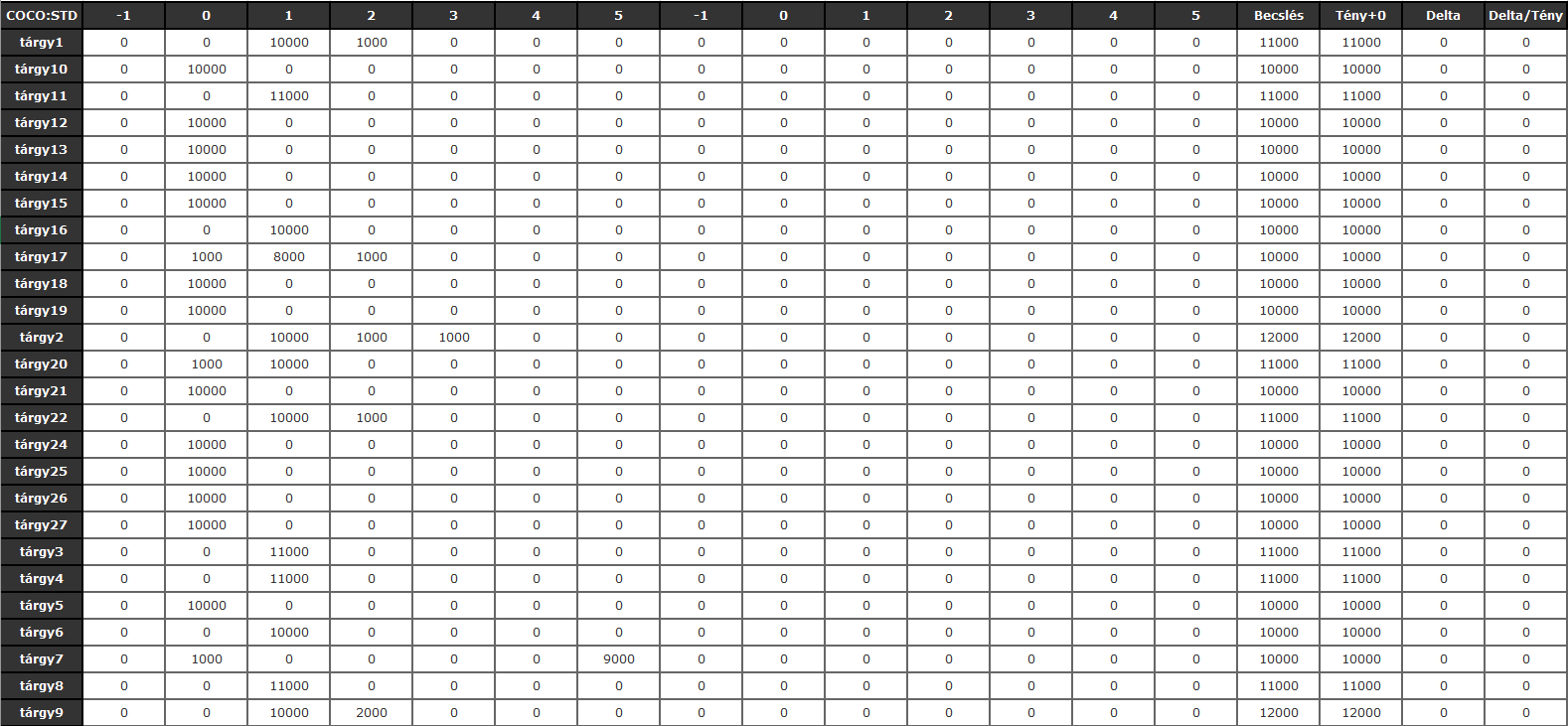
10.ábra – a végleges OAM előkészítését szolgáló táblázat (Forrás: saját ábrázolás)

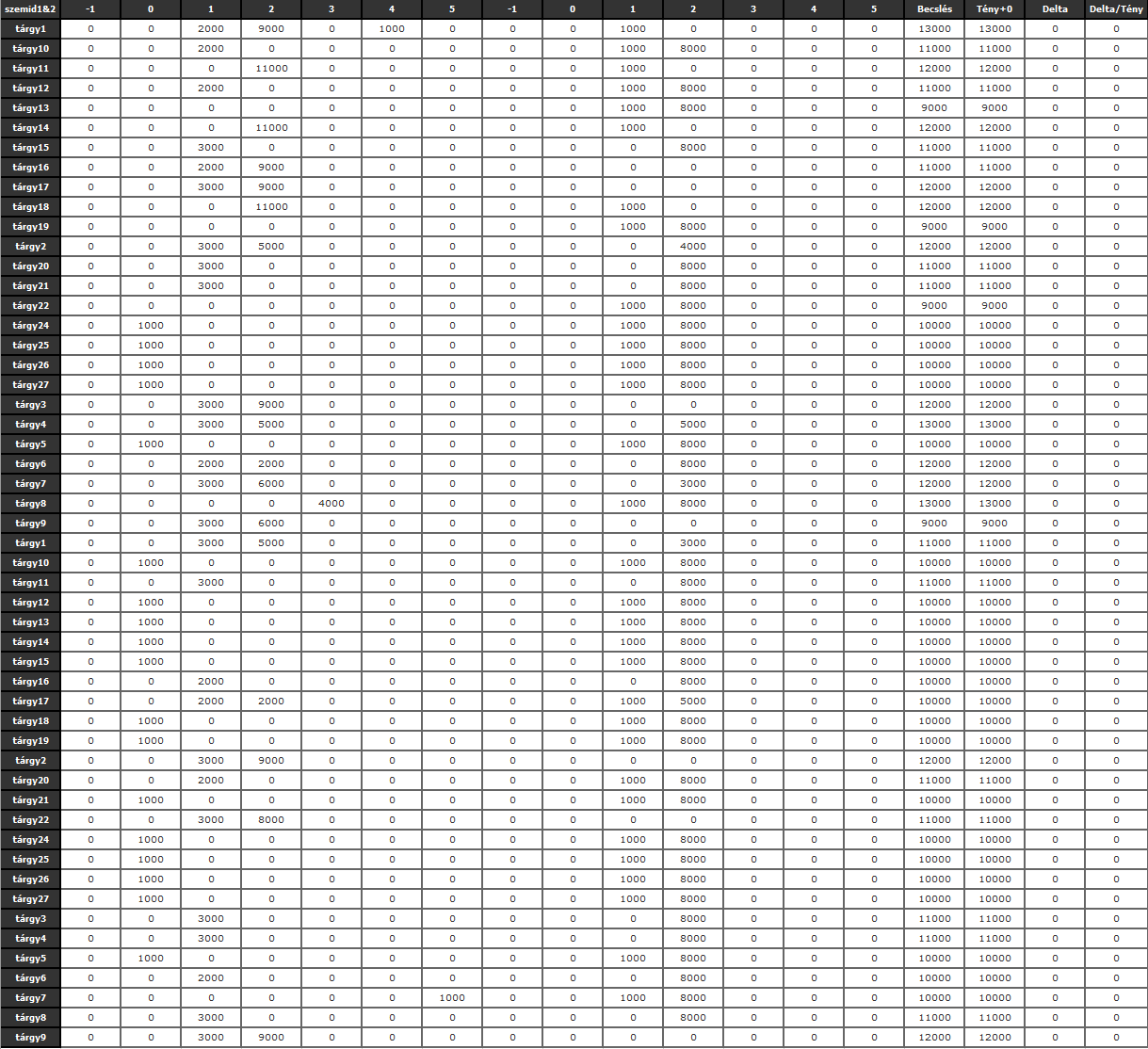
## A mesterséges intelligencia modellje

Az ember számára ezen számok halmaza nem mondana sokat, s ezt a naiv értékelések, korreláció-számítások félreérthetetlenül jelzik is (10.ábra), de ha átadjuk egy context free, optimalizáló robotnak az adatokat, a robot(szem) létrehoz egy olyan modellt, amin már mi emberként is láthatjuk, hogy mi a lényeg. Három modell jött létre, az első kettő a két alany külön-külön való elemzése (11.ábra, 12.ábra), és a harmadikon pedig a kettő együttes összevetése (13.ábra). Ezekben a modellekben láthatók a tantárgyak sorokként és a direkt és inverz irányú -1től 5ig tartó skála is, valamint egy becslés, tény és egy delta oszlop. A becslés és a tény igazából mind kettő az Y oszlop értékét (vagyis -1 és 5 közötti érték\*1000+10000) értéket mutatja. Ami fontos számunkra az a delta oszlop, ami az eltérést mutatja. Ebben az oszlopban minél nagyobb szám szerepel annál jobban ellentmondásos a tantárgy értékelése és az abba a tantárgyba tartozó többi objektum értékelése. A harmadik modell, ami a két alanyt veti össze, ugyan úgy épül fel, mint az előző kettő csak it a sorok száma az egy-személyes modellek duplája, hiszen két alany értékei szerepelnek benne azonos attribútumok mellett. Hasonlóan a delta oszlopot kell néznünk és itt is, akárcsak korábban, mindenütt 0 érték szerepel. Vagyis a robot szerint a két személy értékelési egymáshoz képest is rendben vannak, teljes mértékben lehetségesek.



11. ábra – szemid1 robot által készített modellje (Forrás: saját ábrázolás, ahol a baloldali -1 és 5 intervallum a direkt iránynak felel meg, míg a jobboldali -1 és 5 intervallum az inverz iránynak felel meg – vö. 9. ábra oszlopfejléce.)

12.ábra – szemid2 robot által készített modellje (Forrás: saját ábrázolás, ahol a baloldali  
-1 és 5 intervallum a direkt iránynak felel meg, míg a jobboldali -1 és 5 intervallum az inverz iránynak felel meg – vö. 9. ábra oszlopfejléce.)



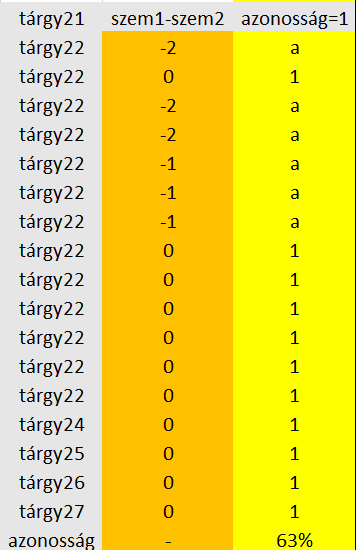
13.ábra – szemid1 és 2 egyszerre való vizsgálatakor, robot által készített modell (Forrás: saját ábrázolás, ahol a baloldali -1 és 5 intervallum a direkt iránynak felel meg, míg a jobboldali -1 és 5 intervallum az inverz iránynak felel meg – vö. 9. ábra oszlopfejléce.)

A 11-12-13. ábrák értelmezése:

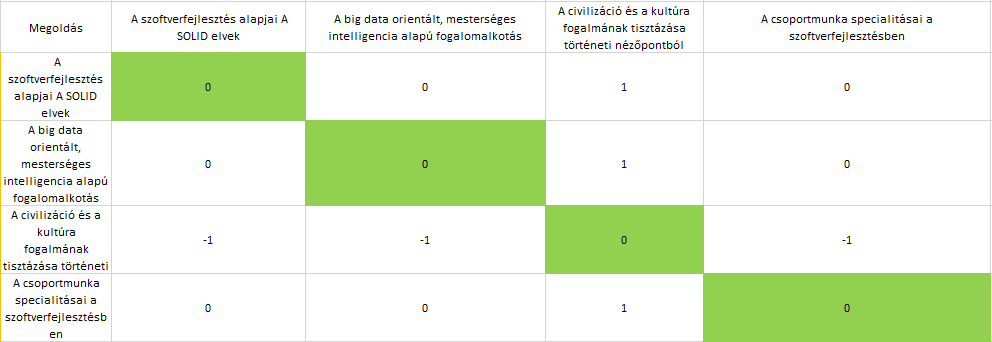
* A 11. ábra alanya nem csak a direkt oldalon mutat fel erőtereket, hanem az indirekt oldalon is.
* A 12. ábra alanya csak direkt erőterekre támaszkodva is értelmezhető.
* A két alany egymáshoz képest racionálisnak tekinthető, amit a tény-becslés eltérések monoton nullaértékei jeleznek.

# Eredmények

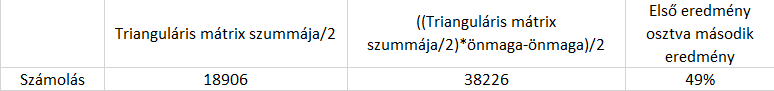
A felmérés célja az volt, hogy megvizsgáljuk két, vagy akár több személy tudásának egyenértékét. Amit az lehetett látni a mesterséges intelligencia szerint ez az egymáshoz és önmagukon belüli konzisztencia adottnak tételezhető fel, tehát a két személy egymáshoz képest ugyan úgy logikusak, konzisztensek a kulcsszavak értékelései és az aggregált tantárgyi értékelés egymásból való levezethetősége tekintetében. Ezen felül még min. kétféle statisztikai módon lehet a megegyezéseket vizsgálni. Az első mód a numerikus egyezés (14.ábra) ami annyit tesz, hogy hányszor mondták mindketten ugyan arra az objektumra ugyan azt az értéket, ez egy egyszerű számolással, személy1 értékeiből kivonjuk a személy2 értékeit és ahol 0 van ott ugyan azt mondták, összeszámoljuk az 0-kat (a sárga oszlopban 1-es jelenik meg ahol az eredmény 0 volt és egy „a”, ha bármi más) és ki is jön, hogy a mi esetünkben (62.8) ~ 63%-ban egyeznek meg. A másik mód a tendencia szerinti azonosságok. (15.ábra) Ez annyit tesz, hogy hányszor értékelték az egyik objektumot jobbra, mint a másikat. Itt jön vissza a már említett trianguláris mátrix. Személyenként létre kell hozni egy táblázatot, aminek a sorai és oszlopai is az objektumok, ugyan abban a sorrendben, s ez megmutatja az adott objektum viszonyát a másikhoz képest. Ebben az esetben az előjeleket kell vizsgálnunk, vagyis azt, hogy egy-egy személy két objektum kapcsán melyiket értékelte nagyobb tudásszinttel kecsegtetőnek. A harmadik mátrix, ami a két személy mátrixaiból létrehozza az eredményt, tehát megadja, hány ponton volt azonos a két személy azonos objektumpárra vonatkozó értékelésének „tendenciája” (16.ábra). Ez úgy történik, hogy megnézi, hogy az egyik, ahova 0 került oda itt is 0 fog, ha nem akkor megnézi, hogy az érték kisebb-e, mint 0 ha igen akkor ide is 0 kerül, ha nem akkor 1, majd ezt a táblázatot úm. „szummázzuk”, és annak eredményét elosztjuk 2-vel majd a kapott összeget megszorozzuk önmagával, kivonjuk belőle önmagát és elosztjuk ismét 2-vel. Végül ezt a két összeget egymással elosztjuk és kijön, hogy a tendencia szerű megegyezés 49%-ban történt meg. Emellett még az is kiderül a kutatásból, mennyire jó egy tanterv, ill. mennyire gondos az akkreditáció az alanyok által reprezentált célcsoportra vonatkozóan. Ezt úgy állapíthatjuk meg, hogy megvizsgáljuk a tantárgyakra, kulcsszavakra adott átlagos értékeléseket félévekre lebontva. Egy jó tanterv esetében, ha az értékeket grafikonon nézzük meg (17.ábra), akkor egy csökkenő tendenciát mutató grafikont kell látnunk, amennyiben az X-tengelyen a szemeszterek szerepelnek 1-től 6-ig, hiszen akkor jó egy tanterv, ha a tanulás kezdetén rendelkezik az egyed a legnagyobb bemeneti tudással az akkor tanult tantárgyakról és a legkevesebbel tudással rendelkezik a képzés végére tervezett tantárgyak esetében.



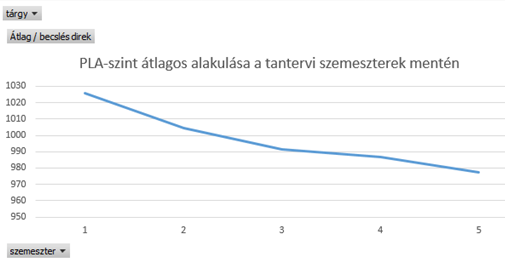
14.ábra – numerikus megegyezést kiszámító táblázat alsó része – narancságában a szemid1 érékeléséből kivont szemid2 értékeléseinek eredménye, sárgában segédszámolás: ha a mellette a narancsban 0 van akkor 1-est ír, ha más akkor egy a (tetszőleges nem szám karakter) - (Forrás: saját ábrázolás, ahol a mértékegységek = narancssárga: szemid1 objektumra adott értékelése – szemid2 azonos objektumra adott értékelése, citromsárga: nincs mértékegység, ez egy segédszámolás, bármilyen tetszőleges számmal és nem szám karakterrel lehet dolgozni)



15.ábra – példa egy tendencia szerű megegyezés kiszámolásához alkalmas táblázatra (Forrás: saját ábrázolás, ahol a mértékegységek = ahol az értékek előjeles pontszám-különbségek arra utal, hogy ha – előjel akkor az adott objektum kevésbe ismert a másiknál, a 0 az ugyan akkora tudás jele, míg + előjel az adott objektum nagyobb ismeretére utal a másiknál)



16.ábra – a harmadik trianguláris mátrix aljánál található számítások a tendenciaszerinti megegyezés meghatározásához (Forrás: saját ábrázolás)



17.ábra – a tudásszint grafikonja szemeszterenként (Forrás: saját ábrázolás, ahol az Y-tengely mértékegysége: jóságpont, az X-tengely a szemeszterek sorrendjének jele)

**A dolgozat tehát teljesítette a profilírozás és a csoportképzés kapcsán vállalt feladatokat, sőt, az oktatástervezés és akkreditáció kapcsán is hasznos ellenőrző riportot/adatvizualizációt volt képes levezetni.**

# Összefoglalás

## Rezümé – HU

Kezdésképpen érdemes megismerni az alapkérdést, ami mentén kialakult az alábbi dolgozat. A kérdés így hangzott: Vajon lehet-e kettő vagy több tanuló másképpen, de egyforma? Ennek kiderítésnek érdekében végeztünk egy anonim PLA-felmérést (PLA=prior learning assesment), vagyis megkérdeztünk Hallgatókat, vajon a BPROF-tanterv egyes tantárgyai, témakörei és kulcsszavai kapcsán „-1” és „5” között hogyan értékelik saját tudásszintjüket, ahol a „-1” az adott objektum feleslegességére utalt, míg „0” volt a nem tudás, nem ismeret jele, s 1-től 5-ig a további értékek az egyre növekvő tudásszintre utalnak (1<5). A kapott adatokat egy Excel táblázatba foglalva (ahol az egyes kulcsszavak témakörökhöz és tantárgyakhoz való kapcsolata is megadásra került), valamint kimutatásokat használva elkészítettük az OAM-ot (OAM=objektum-attribútum-mátrix). Majd ezt az OAM-ot átadtuk a context free, optimalizáló mesterséges intelligenciának, és megnéztük, hogy a „robotszem” mit lát ezekben az adathalmazokban. A felmérésben kísérleti jelleggel két személy szerepelt. Kétféle módon adtuk át az adatokat a robotnak. Az első rétegben a személyeket külön-külön vizsgáltuk. Míg a második rétegben a részvevőket egyszerre értelmeztük. Eredményként azt kaptuk, hogy az azonos szak iránt érdeklődő alanyok bizonyos dolgokról hasonlóan gondolkodnak - az elvárásoknak megfelelően - vö. numerikus (x/n=62.8%) és tendenciaszerű azonosságok (y/(n\*n-n/2)=49%). Ennek ellenére a robot azt mutatja, hogy a két alany jól képes az együttműködésre nagyrészt, és az modellezett értékeik alapján mindketten másként egyformán logikusnak, konzisztensnek, értékelhetők (önmagukban és egymáshoz képest is) az alacsonyabb aggregációs szint (kulcsszavak) és a magasabb aggregációs szintek (pl. tantárgyak) kapcsolatait tekintve. Azonban szembetűnő egy kevés eltérés is, míg az egyik alany csak a direkt világban volt aktív. Addig a másik az inverz világban is feltűnik. Mindezen eredmények igazak voltak kétféle modellezés eljárás esetén is (antidiszkriminatív modell vs. standardmodell).

## Abstract – EN

To begin with it would be wise to know the question along which the following dissertation was written. The question was: Could two or more students be different, but still the same in a way. To find this out, we conducted an anonymous PLA survey (PLA = prior learning assessment) that is, we asked students to rate their level of knowledge between „-1” and „5” for each subject, topic and keyword in the BPROF curriculum, where „-1” referred to the redundancy of the object, while 0 was the no knowledge level and further values from 1 to 5 indicate the increasing level of knowledge (1<5) The survey’s data was stored in an Excel chart (where the relationship of each keyword to topics and subjects were fixed) where an OAM was made (OAM = object attribute matrix) with the help of statements. Then, we gave this OAM to the a context-free, optimising artificial intelligence to see what the „robot eye” make of these values. Two people too part in the survey on an experimental basis, and we gave the data to the robot in two ways. The first way was when the two participants were examined separately. While in second way they were examined simultaneously. As a result, we found out that the subjects were interested in a similar way towards certain things, as expected (numerical x / n = 62.8% and trend like y / (n \* n - n / 2) = 49%). Nevertheless, the robot shows that the two subjects are well able to cooperate for the most part, and based on their modelled values, they can both be evaluated as equally logical, consistent, in their own right, and in relation to each other, in terms of the relationship between lower levels of aggregation (keywords) and higher levels of aggregation (e.g., subjects). However, a small difference is also worth noting, while one subject was active only in the direct world, the other also appeared in the inverse world. All of these results were also true for two types of modeling procedures (anti-discriminatory model vs. standard model).

# Források

<https://miau.my-x.hu/miau2009/index.php3> (utoljára használva: 2020.11.13.)

<https://miau.my-x.hu/bprof/?C=M;O=D> + <https://miau.my-x.hu/bprof/oam_pla_13%20(1).xlsx> (utoljára használva: 2020.11.13.)

<https://en.wikipedia.org/wiki/Recognition_of_prior_learning> (utoljára használva: 2020.11.13.)

<https://hu.wikipedia.org/wiki/Mesters%C3%A9ges_intelligencia> (utoljára használva: 2020.11.13.)

https://miau.my-x.hu/myx-free/ (utoljára használva: 2020.11.13.)

<https://miau.my-x.hu/miau/196/My-X%20Team_A5%20fuzet_HU_jav.pdf> (utoljára használva: 2020.11.13.)

# Mellékletek

## Rövidítések jegyzéke

* BPROF – Bachelor of Profession (felsőoktatási alapképzés egy formája)
* EN – English (angol)
* HU – Hungaryan (magyar)
* IT – Information Technology (információs technológia)
* MI – Mesterséges Intelligencia
* OAM – Objektum Attribútum Mátrix (mátrix fajta)
* PLA – Prior Learning Assesment (egy felmérési forma)
* Szemid1 – Személy Indentifier 1 (anonim megjelölése egy személynek/alanynak)
* Szemid2 – Személy Indentifier 2 (anonim megjelölése egy személynek/alanynak)
* TDK – Tudományos Diákköri Konferencia

## Ábrák jegyzéke

1. 2\*282+1fejléc sorból álló táblázat egy részlete (Forrás: saját ábrázolás, ahol az értékoszlop mértékegysége: pontszám/kód) 7. oldal
2. a trianguláris mátrix bal felső részlete (Forrás: saját ábrázolás, ahol az értékek előjeles pontszám-különbségek arra utal, hogy ha – előjel akkor az adott objektum kevésbe ismert a másiknál, a 0 az ugyan akkora tudás jele, míg + előjel az adott objektum nagyobb ismeretére utal a másiknál) 8. oldal
3. szemid1 legjobban értékelt objektumának ellenőrzése (Forrás: saját ábrázolás – a 4-es érték mértékegysége = A tudásszint, a -1 és 5 intervallumban) 8. oldal
4. 4.ábra – szemid2 legjobban értékelt objektumának ellenőrzése) (Forrás: saját ábrázolás – a 11-es érték mértékegysége = A tudásszint, a -1 és 5 intervallumban)
5. szemid1 tantárgyakra adott értékelésének kimutatása (Forrás: saját ábrázolás, ahol a cellák értékeinek mértékegysége az adott oszlopban a fejlécen látható érték/pontszám/kód) 9. oldal
6. szemid1 tantárgyaként adott érékeléseinek darabszáma százalékosan soronként - (Forrás: saját ábrázolás) 10. oldal
7. szemid1 tantárgyakként adott értékeléseinek darabszáma (Forrás: saját ábrázolás, ahol a mértékegység a tantárgyan belüli, oszlopfejlécen látható értékek darabszáma – vagyis 281 db önálló értékelés áll rendelkezésre minden alany esetén, melyek tantárgyakhoz eltérő mértékben kötődnek – lehet tehát olyan tantárgy is, mely önmagában értékelendő – pl. tárgy8/11/14/15/18/19/21/24/25/26/27) 11. oldal
8. szemid2 tantárgyakként adott értékeléseinek darabszáma (Forrás: saját ábrázolás, ahol a mértékegység a tantárgyan belüli, oszlopfejlécen látható értékek darabszáma) 12. oldal
9. szemid1 OAM-ja (Forrás: saját ábrázolás – mértékegység: rangsorszámok oszloponként) 13. oldal
10. a végleges OAM előkészítését szolgáló táblázat (Forrás: saját ábrázolás) 14. oldal
11. szemid1 robot által készített modellje (Forrás: saját ábrázolás, ahol a baloldali -1 és 5 intervallum a direkt iránynak felel meg, míg a jobboldali -1 és 5 intervallum az inverz iránynak felel meg – vö. 9. ábra oszlopfejléce.) 15. oldal
12. szemid2 robot által készített modellje (Forrás: saját ábrázolás, ahol a baloldali  
    -1 és 5 intervallum a direkt iránynak felel meg, míg a jobboldali -1 és 5 intervallum az inverz iránynak felel meg – vö. 9. ábra oszlopfejléce.) 15. oldal
13. szemid1 és 2 egyszerre való vizsgálatakor, robot által készített modell (Forrás: saját ábrázolás, ahol a baloldali -1 és 5 intervallum a direkt iránynak felel meg, míg a jobboldali -1 és 5 intervallum az inverz iránynak felel meg – vö. 9. ábra oszlopfejléce.) 16. oldal
14. numerikus megegyezést kiszámító táblázat alsó része – narancságában a szemid1 érékeléséből kivont szemid2 értékeléseinek eredménye, sárgában segédszámolás: ha a mellette a narancsban 0 van akkor 1-est ír, ha más akkor egy a (tetszőleges nem szám karakter) - (Forrás: saját ábrázolás, ahol a mértékegységek = narancssárga: szemid1 objektumra adott értékelése – szemid2 azonos objektumra adott értékelése, citromsárga: nincs mértékegység, ez egy segédszámolás, bármilyen tetszőleges számmal és nem szám karakterrel lehet dolgozni) 18. oldal
15. példa egy tendencia szerű megegyezés kiszámolásához alkalmas táblázatra (Forrás: saját ábrázolás, ahol a mértékegységek = ahol az értékek előjeles pontszám-különbségek arra utal, hogy ha – előjel akkor az adott objektum kevésbe ismert a másiknál, a 0 az ugyan akkora tudás jele, míg + előjel az adott objektum nagyobb ismeretére utal a másiknál) 18. oldal
16. a harmadik trianguláris mátrix aljánál található számítások a tendenciaszerinti megegyezés meghatározásához (Forrás: saját ábrázolás) 19. oldal
17. a tudásszint grafikonja szemeszterenként (Forrás: saját ábrázolás, ahol az Y-tengely mértékegysége: jóságpont, az X-tengely a szemeszterek sorrendjének jele) 19. oldal

# Tartalomjegyzék

[Bevezetés 2](#_Toc56432468)

[Kiindulási helyzet 2](#_Toc56432469)

[Motiváció 2](#_Toc56432470)

[A feladat ismertetése 2](#_Toc56432471)

[Célcsoport 3](#_Toc56432472)

[Hasznosság 3](#_Toc56432473)

[A dolgozat szerkezetéről 3](#_Toc56432474)

[Szakirodalmi háttér 4](#_Toc56432475)

[A PLA – a keretrendszer 4](#_Toc56432476)

[A mesterséges intelligencia szerepe és potenciálja 4](#_Toc56432477)

[Az elemzés menete 6](#_Toc56432478)

[A felmérés elkészítése 6](#_Toc56432479)

[Az adatok feldolgozása 6](#_Toc56432480)

[A nyers adatoktól a kimutatásokon át az OAM-ig 8](#_Toc56432481)

[Az első kimutatás 9](#_Toc56432482)

[A második kimutatás 10](#_Toc56432483)

[További kimutatások 11](#_Toc56432484)

[Az OAM 12](#_Toc56432485)

[A mesterséges intelligencia modellje 14](#_Toc56432486)

[Eredmények 17](#_Toc56432487)

[Összefoglalás 20](#_Toc56432488)

[Rezümé – HU 20](#_Toc56432489)

[Abstract – EN 21](#_Toc56432490)

[Források 22](#_Toc56432491)

[Mellékletek 23](#_Toc56432492)

[Rövidítések jegyzéke 23](#_Toc56432493)

[Ábrák jegyzéke 24](#_Toc56432494)

[Tartalomjegyzék 26](#_Toc56432495)