The background features a complex, abstract network of glowing blue lines and nodes, forming various geometric shapes like triangles and polygons. The lines are thin and light blue, while the nodes are small, bright blue dots. The overall effect is a sense of digital connectivity and data flow.

**AZ OKTATÁS, A KUTATÁS ÉS
A KÖZGYŰJTEMÉNYEK DIGITÁLIS
TRANSZFORMÁCIÓJA FELSŐFOKON**

**NETWORKSHOP 2024
33. Országos Informatikai Konferencia**

**2024. április 3–5.
Eszterházy Károly Katolikus Egyetem, Eger**

AZ OKTATÁS, A KUTATÁS ÉS A KÖZGYŰJTEMÉNYEK DIGITÁLIS TRANSZFORMÁCIÓJA FELSŐFOKON

NETWORKSHOP 2024
33. Országos Informatikai Konferencia

2024. április 3–5.
Eszterházy Károly Katolikus Egyetem, Eger

Szerkesztette: Tick József, Kokas Károly, Holl András

HUNGARNET Egyesület
Budapest, 2024



HUN-REN
Magyar Kutatási Hálózat

NETWORKSHOP

Szerkesztette: Tick József, Kokas Károly, Holl András

Tipográfia és tördelés: Vas Viktória

Korrektúra: Danyi Melinda

Angol nyelvi lektor: Cseresnyés Dorottya

Networkshop 2024 konferencia előadásainak közleményei

Eszterházy Károly Egyetem, Eger

2024. április 3–5.

ISBN 978-615-82243-2-1

DOI: <https://doi.org/10.31915/NWS.2024>

Kiadja a HUNGARNET Egyesület
az MTA Könyvtár és Információs Központ közreműködésével

Budapest

2024

Borítókép: [freepik.com](https://www.freepik.com)

TARTALOMJEGYZÉK

Előszó	5
Ungváry Rudolf A MARC21 formátum kettős szerkezete és a formátum felhasználói szintjének fordításai	7
Holl András, Andódy Katalin Adatbányászati gyakorlatok repositóriumra és MTMT-re.....	16
Simon András Mesterséges intelligenciával támogatott adatgazdagítás a Nemzeti Levéltárban.....	22
Soós Gábor, Rövid András, Ormos Pál V2X – A járművek közötti kommunikáció kihívásai	29
Csernai Zoltán Egy online tanulást támogató portál kurzusának vizsgálata Big Data adatelemző módszerekkel.....	36
T. Nagy László, Németh Áron A mesterséges intelligencia (MI) teológiai kompetenciái	45
Mészáros Erika Kodaktól a jövőig – Egy könyvtári digitalizálás szinterei.....	55
Frankó Máté, Sándor Ákos Adatvizualizáció a könyvtári menedzsmentben: fejlesztések az SZTE Klebelsberg Könyvtár döntéstámogató rendszerében.....	63
Hernek István A felhasználóképzés szintjei az SZTE Klebelsberg Könyvtárban: az elsőévesektől a kutatókig	72
Némethi-Takács Margit, Borbély Mária Bibliográfiai kapcsolatok az általános megjegyzés adatmezőben.....	78
Dobás Kata, Tüskés Anna A magyar irodalomtörténet bibliográfiájának migrációja az ITIdata szemantikus adatbázisba	87
Horváth Péter A kanonikus magyar költészet versformakeresője.....	96
Sebestyén Ádám, Sárközi-Lindner Zsófia Történeti források szemantikus feldolgozása – Az ELTEdata adatbázis új gyűjteményei	105

Bolya Mátyás	
Lyukkártya és népdalrendezés – Egy mechanikus népzenei adatbázis digitális rekonstrukciójának lehetőségei.....	112
Kovácsházy Tamás	
Az idő, mint alapvető infrastruktúra, az idő szerepe az adatközpontban.....	121
Albert Ágota Katalin	
A mesterséges intelligencia használatának követelményei az oktatási szektorban, különös tekintettel a mesterséges intelligencia használatáról szóló rendeleltre.....	129
Varga Emese	
Digitális szövegszerkesztés a dHUpla keretrendszerében	135
Nemoda Zsuzsanna, Héjja Balázs, Nagy Andor, Tóth Máté	
A Pest Megyei Digitális Könyvtár fejlesztése	141
Nagy Dóra, Sándor Ákos	
Voice2text: a hanganyagátírás lehetőségei MI segítségével.....	149
Kalcsó Gyula	
Képek és metaadatok gyűjteményezése scrapingtechnológiával közösségi képmegosztó oldalról	157
Péter Róbert, Szántó Zsolt, Biacsi Zoltán, Kocsis Zoltán, Berend Gábor, Bilicki Vilmos	
Az AVOBMAT (Analysis and Visualization of Bibliographic Metadata and Texts) többnyelvű kutatási eszköz munkafolyamata és új funkciói	163
Máray Tamás	
Kvantum-számítástechnika: ez már a „jövő”?.....	171
Fellegi Zsófia	
Digitális kiadások migrációja: gépi és emberi intelligencia együttműködése.....	177
Palkó Gábor	
Posztmodern intertextualitás és digitális szövegkiadás	184
Antal Dániel	
A szlovák adatkicserélési tér magyarországi föderációjának lehetőségei.....	192
Vass Johanna	
Kutatási adatok megosztása a gyakorlatban – Adatrepozitóriumok használata az Ökológiai Kutatóközpont publikációiban	199
Mihály Eszter, Micsik András, Nagy Kadosa	
Irodalmi levélváltások nyomában TEI-vel és térképpel.....	208

Egy online tanulást támogató portál kurzusának vizsgálata Big Data adatelemző módszerekkel¹

Investigating the course of an online learning portal using Big Data analytics²

Csernai Zoltán

Eszterházy Károly Katolikus Egyetem, Digitális Technológia Intézet,

Humáninformatika Tanszék

csernai.zoltan@uni-eszterhazy.hu

Absztrakt

Napjainkban a Big Data technológia egyre népszerűbb az oktatásban, mivel lehetőséget teremt a tanulási folyamatok hatékonyabbá tételére, a tanulók személyre szabott támogatására és az intézményi döntéshozatal javítására. (Bamiah et al., 2018).

A Big Data technológiák ezen kívül javítják és támogatják az online tanulást támogató rendszerek által generált adatok gyűjtését, tárolását, feldolgozását, elemzését, biztonságát és virtualizációját (Otoo-Arthur & van Zyl, 2020).

A cikk célja a Big Data technológiához kapcsolódó két kulcsfontosságú oktatási terület, az oktatási adatbányászat és a tanulási analitika folyamatának, módszereinek és alkalmazási lehetőségeinek bemutatása. A kutatás egy online tanulást támogató portál kurzusának tevékenységeit vizsgálja Big Data adatelemző módszerek és eszközök segítségével, egy online platformon keresztül.

A cikk célkitűzése a tanulási folyamatok mélyebb megértése és a tanulók jövőbeli eredményeinek előrejelzése Big Data elemzési módszerekkel. A tanulmány olyan elemzési megközelítéseket kínál, amelyek segíthetik az oktatókat a személyre szabott támogatás nyújtásában és a tanulói siker előmozdításában.

„A Kulturális és Innovációs Minisztérium ÚNKP-23-3 kódszámú Új Nemzeti Kiválóság Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával készült.”

Kulcsszavak: Big Data, oktatási adatbányászat, tanulási analitika, Moodle, online tanulást támogató portál.

Abstract

Nowadays, Big Data technology is becoming increasingly popular in education as it offers the potential to make learning processes more efficient, provide personalised support to learners and improve institutional decision-making (Bamiah et al., 2018).

1 „A Kulturális és Innovációs Minisztérium ÚNKP-23-3 kódszámú Új Nemzeti Kiválóság Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával készült.”

2 „Supported by the ÚNKP-23-3 New National Excellence Program of the Ministry for Culture and Innovation from the source of the National Research, Development and Innovation Fund.”

In addition, Big Data technologies improve and support the collection, storage, processing, analysis, security and virtualisation of data generated by e-learning systems (Otoo-Arthur & van Zyl, 2020).

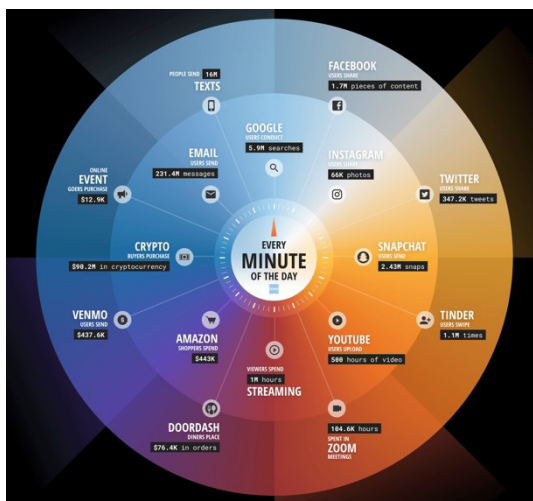
The aim of this paper is to describe the processes, methods and applications of two key educational areas related to Big Data technology: educational data mining and learning analytics. The research investigates the activities of a course in an online learning support portal using Big Data data analytics methods and tools through an online platform.

The objective of the paper is to gain a deeper understanding of learning processes and to predict learners' future performance using Big Data analytics methods. It offers analytical approaches that can help educators provide personalized support and promote student success.

Keywords: Big Data, educational data mining, learning analytics, Moodle, online learning support portal.

1. Bevezetés

Az internet rohamos fejlődésével párhuzamosan exponenciálisan növekszik a digitális adatok mennyisége. A 10. éves „Data Never Sleeps” infografika (lásd. 1. ábra) szemléletesen bemutatja ezt a folyamatot, kiemelve az adatok mennyiségének és sokféleségének gyorsulását.



1. ábra: Data Never Sleeps 10.0 infografika

A kép forrása:

<https://www.domo.com/data-never-sleeps>

Az 1. ábra szemlélteti a digitális adatok exponenciális növekedését, amely jellemzően a Big Data három alapvető jellemzőjében ragadható meg: mennyiség, sebesség és változatosság. A percnként keletkező hatalmas mennyiségű adat (pl. Google-keresések, Facebook-bejegyzések, YouTube-videók) rávilágít a jelenség mértékére, míg az adatok keletkezésének sebessége és formátumbeli sokfélesége a feldolgozás és elemzés kihívásaira hívja fel a figyelmet.

Mivel a „Big Data” kifejezésnek jelenleg nincs egyértelműen meghatározott, általánosan elfogadott definíciója, a következőkben áttekintjük a szakirodalomban található legfontosabb definíciós kísérleteket. Ez az áttekintés elengedhetetlen a téma megfelelő kontextusba helyezéséhez és a kutatás elméleti kereteinek meghatározásához.

2. Big Data definíciós kísérletek

A „Big Data” fogalmának nincs egyértelműen meghatározott definíciója. A szakirodalomban található megközelítések (pl. 3V-, 4V-, 5V-modell) különböző szempontokat (mennyiség, sebesség, változatosság, megbízhatóság, érték) emelnek ki. Kutatásunkban az 5V-modellt tekintjük kiindulópontnak, mivel az oktatási környezetben az adatok megbízhatósága és értéke kiemelt jelentőséggel bír. A tanulók tanulmányi előmenetelének elemzéséhez például elengedhetetlen, hogy az adatok megbízhatóak legyenek, és lehetővé tegyék a tanulási folyamatok mélyebb megértését.

„A Big Data négy jellemzője a mennyiség, a sebesség, a változatosság és az érték.” (Dijcks, 2011)

„A Big Data együttese magába foglalja a korábban soha nem látott mértékű és változatos forrásból érkező adatok rögzítését, feldolgozását, elemzését, megosztását, illetve az eredmények vizualizálását.” (Szűts & Yoo, 2016)

„Big Data alatt olyan rendkívül nagy adathalmazokat értünk, amelyek az emberi viselkedéssel és interakciókkal kapcsolatos mintákat, trendeket és összefüggéseket tárnak fel.” (Yeoman, 2019)

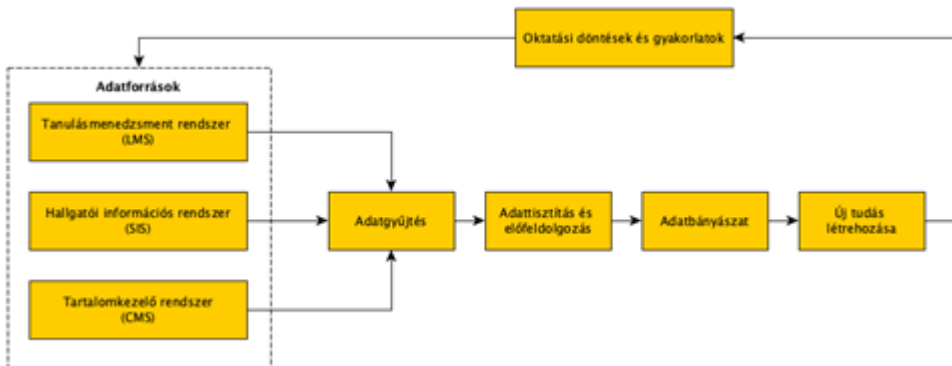
A „Big Data” kifejezés szakirodalmi vizsgálata során kutatásom fókuszpontjába két növekvő terület került: az oktatási adatbányászat és a tanulási analitika. Mindkét terület a Big Data oktatási alkalmazására fókuszál, célja a tanulási folyamatok mélyebb megértése és a tanulók személyre szabott támogatása. Az oktatási adatbányászat a tanulási folyamatokból származó adatok elemzésére összpontosít, míg a tanulási analitika a tanulási környezet és a tanulók viselkedésének elemzésével foglalkozik. Ezen területek integrálása lehetővé teszi a tanulási folyamatok hatékonyabbá tételét, a tanulók személyre szabott támogatását és az intézményi döntéshozatal javítását.

3. Oktatási adatbányászat

Az oktatási adatbányászat definíciója szerzőnként eltérő lehet, de általánosságban úgy fogalmazhatjuk meg, mint az oktatási környezetből származó adatok elemzését a tanulási folyamatok javítása érdekében. Az International Educational Data Mining Society (2018) szerint az oktatási adatbányászat célja a tanulók és a tanulási környezetük jobb megértése. Romero és Ventura (2020) definíciója kiegészíti ezt a tanulói teljesítmény előrejelzésével és modellek létrehozásával. Mindkét definíció hangsúlyozza az adatok elemzésének fontosságát az oktatás területén.

Az oktatási adatbányászat folyamata több lépésből áll (lásd. 2. ábra): először adatokat gyűjtünk oktatási adatforrásokból, majd megtisztítjuk és előfeldolgozzuk őket. Ezután

adatbányászati módszerekkel mintákat és kapcsolatokat keresünk az adatokban. Végül az eredményeket értelmezzük, és új ismereteket generálunk az oktatásról.



2. ábra: Az oktatási adatbányászat folyamatábrája
A kép angol nyelvű forrása: Ashrafimoghari, 2022

Néhány példa az oktatási adatbányászatra:

- a tanulók mentális modelljeinek következtetése a kérdéseik alapján;
- a plágium felderítése a diákok által benyújtott munkák összehasonlításával az online forrásokkal és más tanulók beadványaival;
- modellek készítése a hallgatók jövőbeli teljesítményének előrejelzésére;
- a veszélyeztetett tanulók beazonosítása a viselkedési minták alapján.

4. Tanulási analitika

A tanulási analitika a tanulási folyamatok adatalapú elemzését jelenti, célja a tanulás és a tanulási környezet megértése és javítása. Clow (2013) szerint a tanulási analitika a tanulási gyakorlatok megértésére és javítására fókuszál. A Society for Learning Analytics Research (SoLAR) (2024) definíciója kiegészíti ezt a tanulás és a tanulási környezet optimalizálásával. Mindkét definíció hangsúlyozza az adatok elemzésének fontosságát a tanulási folyamatok fejlesztésében.

A tanulási analitika célja a Big Data elemzése, hogy olyan visszajelzéseket generáljon, amelyek közvetlenül befolyásolják a tanulókat, oktatókat és a tanulási folyamat részleteit. Egy online tanulást támogató portál kurzusában a tanulók tevékenységeinek, például a fórumon folytatott kommunikációjuk elemzése lehetővé teszi személyre szabott tanulási ajánlások nyújtását a tanulók számára, vagy javaslatok megfogalmazását az oktatók részére a tananyag módosításához.

A tanulási analitika folyamata több lépésből áll (lásd. 3. ábra): először adatokat gyűjtünk a tanulási környezetből, majd elemezzük őket. Ezután jelentéseket készítünk, és vizualizáljuk az adatokat. A negyedik lépés a döntéshozatal, amit a végrehajtás követ. Végül a folyamat a finomítással zárul.



3. ábra: A tanulási analitika folyamatábrája
 A kép angol nyelvű forrása: Ashrafimoghari, 2022

Néhány példa a tanulási analitikára:

- a hallgatók elkötelezettségének nyomon követése az online tanulást támogató portál kurzusában található tartalmakkal kapcsolatban;
- a tanulmányi teljesítmény előrejelzése közösségi média adatok alapján;
- a különböző típusú visszajelzések hatásának mérése a tanulók teljesítményére;
- az adatok felhasználása annak eldöntéséhez, hogy mely beavatkozási stratégiák a leghatékonyabbak a nehézségekkel küzdő tanulók számára.

5. Legnépszerűbb módszerek és eszközök az oktatási adatbányászatban és a tanulási analitikában

Az oktatási adatbányászat és a tanulási analitika számos közös alapot oszt meg mind a kutatási célok, mind az alkalmazott módszerek és technikák terén. Bár sokféle módszer és technika létezik, a legnépszerűbbek közé tartozik az osztályozás és előrejelzés, a klaszterezés, valamint a kapcsolatbányászat. Ezeknek a módszereknek a leírását és alkalmazási példáit az 1. táblázat tartalmazza.

Név	Leírás	Alkalmazási példa
Osztályozás és előrejelzés	Az adatok múltbeli értékei alapján kapcsolatot keres és azonosít a független változó és a függő változók között.	A tanulók teljesítményének előrejelzése az online környezetben tanúsított viselkedésük alapján.
Klaszterezés	A teljes adathalmaz különböző kategóriájú kis adathalmazokra osztása annak eldöntésére, hogy az adatpontok mennyire hasonlóak.	A hasonló tananyagok vagy a tanulók csoportosítására a tanulási és interakciós mintáik alapján.
Kapcsolatbányászat	Az adathalmazban lévő változók közötti kapcsolatok felfedezése és szabályok létrehozása a későbbi felhasználás céljából.	A tanulók gyenge teljesítménye és a tanulási folyamat során tanúsított viselkedésük közötti összefüggések azonosítása.

1. táblázat: Az oktatási adatbányászat és a tanulási analitika legnépszerűbb módszerei és technikái

A 2. táblázat bemutatja az oktatási adatbányászat és a tanulási analitika legnépszerűbb eszközeit és platformjait. A megfelelő eszköz kiválasztásakor a szervezeteknek figyelembe kell venniük a költségvetést, az adatok hozzáférhetőségét és a platform által kínált funkciók sokféleségét. Ezek az eszközök segíthetnek nyomon követni a tanulók elkötelezettségét, teszteredményeit és a kurzusok befejezési arányait egy online tanulást támogató portálon.

Név	Leírás	Webcím
Viselkedéselemzés (Behaviour Analytics)	A hallgatók tevékenységei alapján viselkedésminták előállítását teszi lehetővé.	https://moodle.org/plugins/block_behaviour
Analitikai grafikonok (Analytics graphs)	Összegyűjti a hallgatók aktivitási adatait és vizuális formában megjeleníti.	https://moodle.org/plugins/block_analytics_graphs

2. táblázat: Az oktatási adatbányászat és a tanulási analitika legnépszerűbb eszközei és platformjai

6. A Big Data adatelemző módszerek alkalmazási lehetőségei egy online tanulási portálon

A Big Data adatelemző módszerek széleskörű alkalmazási lehetőségeket kínálnak az online tanulási portálokon. Ezek a módszerek segíthetnek a tanulási folyamatok mélyebb megértésében, a tanulók személyre szabott támogatásában és az oktatási erőforrások hatékonyabb felhasználásában. A következőkben egy esettanulmány keretében bemutatjuk a viselkedéselemzés (Behaviour Analytics) és az analitikai grafikonok (Analytics Graphs) eszközök alkalmazását az Eszterházy Károly Katolikus Egyetem Moodle rendszerében. A kutatás a 2023/24-es tanév második félévében zajlott, és a „Webplapszerkesztés I. – Csernai Zoltán” kurzus adatait elemezte. A gyűjtött adatok többek között a tanulók bejelentkezési adatait, a fórumhozzászólásaikat és a teszteredményeiket tartalmazták. Az adatok elemzéséhez a Moodle beépített viselkedéselemzési és analitikai grafikon eszközeit használtuk.

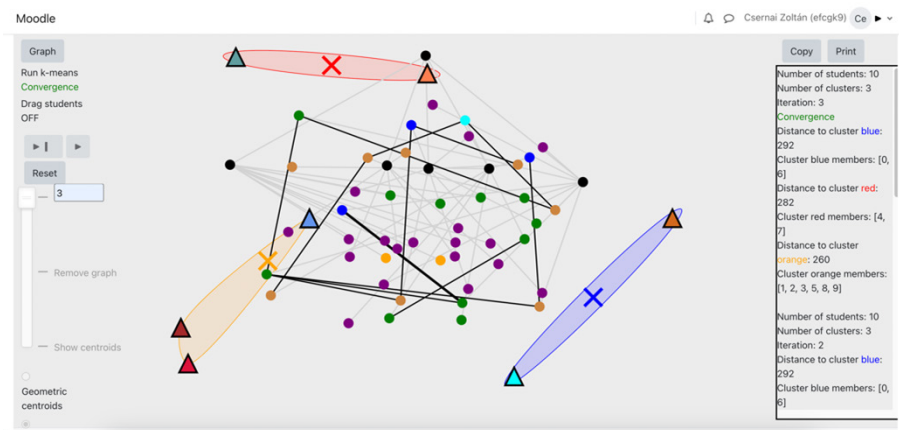
6.1. Viselkedéselemzés (Behaviour Analytics)

A viselkedéselemzés (Behavior Analytics) egy olyan Moodle bővítmény, amely a kurzus-hoz való hozzáférési naplókából elemzi a hallgatók szekvenciális viselkedési mintáit.

Az összes tevékenységet grafikon csomópontként jeleníti meg a kurzusoldalon, ahol a fekete a témát, a narancssárga a fórumot, a zöld a fájlt, a lila az URL-t, a piros a kvízt, és a sárga a külső eszközöket jelöli.

A csomópontok közötti kapcsolatok a hallgatók tevékenységei közötti hozzáféréseket ábrázolják. Amikor egy kiválasztott diák nevére kattintunk, egy vonalhálózat jelenik meg, amely megmutatja, hogyan mozognak a hallgatók a csomópontok által jelölt különböző tanulási tevékenységek között.

A klaszterezés módszerének célja, hogy azokat a tételeket csoportosítsa, amelyek bizonyos szempontok szerint hasonlóságot mutatnak. A klaszterelemzés egyik fontos lépése annak meghatározása, hogy az adatokat hány különálló klaszterbe kell rendezni. Az alábbi ábrán (lásd. 4. ábra) látható, hogy a tíz hallgató három különálló klaszterbe lett csoportosítva.



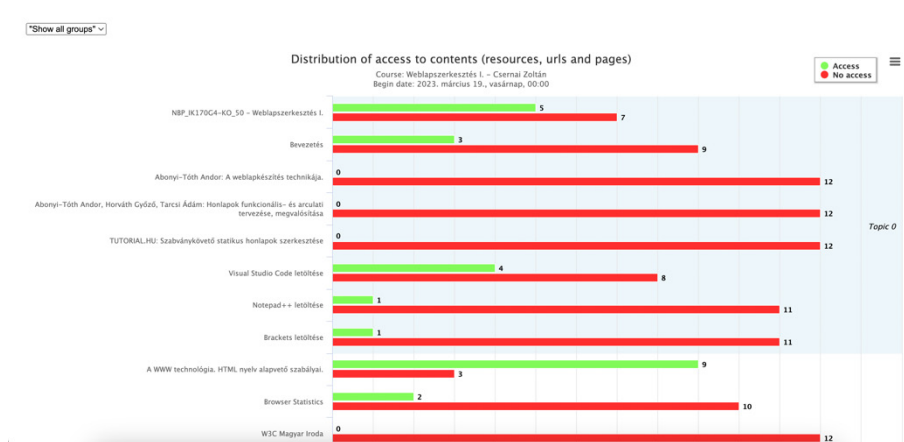
4. ábra: Klaszterezés

Minden hallgatónak van egy középpontja, amely a tevékenységekhez való hozzáférésükből és a csomópontok koordinátáiból származik. Ezeket a hallgatói centrális pontokat klaszterekbe rendezhetjük, hogy csoportosítsuk a hallgatókat és felismerjük a közös hozzáférési mintákat, ezáltal lehetővé téve a tanulók viselkedésének csoportosítását oktatási környezetben.

6.2. Analytics Graphs

Az analitikai grafikonok (Analytics Graphs) egy olyan Moodle bővítmény, amely összegyűjti az online tanulást támogató portálon a tanulói aktivitási adatokat, és különféle grafikonokon jeleníti meg azokat. Az alábbi ábra a hallgatók viselkedését mutatja a tanár által megosztott erőforrásokhoz való hozzáférés szempontjából. A zöld sávok azon hall-

gatók számát jelzik, akik hozzáfértek a jelzett tevékenységhez, míg a piros sávok azokat a hallgatókat mutatják, akik nem fértek hozzá (lásd. 5. ábra).



5. ábra: A tevékenység teljesítményelemzés eredményeit bemutató grafikonok

7. Összegzés és a kutatás folytatása

Ebben a tanulmányban áttekintettük a Big Data különböző definíciós megközelítéseit, különös tekintettel a „V”-modellekre, amelyek a Big Data legfontosabb jellemzőit írják le. Részletesen bemutattuk az oktatási adatbányászat és a tanulási analitika technikáit és alkalmazásait, mint például az osztályozás és előrejelzés, a klaszterezés és a kapcsolatbányászat. Tárgyaltuk továbbá az olyan eszközök, mint a viselkedélemzés (Behaviour Analytics) és az analitikai grafikonok (Analytics Graphs) hasznosságát, amelyek segíthetnek az oktatóknak, kutatóknak, hallgatóknak és adminisztrátoroknak a tanulók viselkedésének és teljesítményének elemzésében.

Úgy véljük, hogy az online tanulást támogató portálok által generált «Big Data» hatékony elemzéséhez az oktatási adatbányászat és a tanulási analitika kiváló eszközöket és módszereket kínál. Ezek a módszerek nemcsak a tanulási folyamatok mélyebb megértését teszik lehetővé, hanem hozzájárulhatnak a tanulók jövőbeli eredményeinek előrejelzéséhez is, ami alapvető fontosságú a személyre szabott támogatás és a tanulói siker előmozdítása szempontjából.

Irodalomjegyzék

- Ashrafimoghari, V. (2022). *Big Data and Education: Using big data analytics in language learning*. <https://doi.org/10.13140/RG.2.2.21946.06080>
- Bamiah, M., Brohi, S., & Bashari Rad, B. (2018). Big data technology in education: Advantages, implementations, and challenges. *Journal of Engineering Science and Technology*.
- Clow, D. (2013). An overview of learning analytics. *Teaching in Higher Education*, 18(6), 683–695. <https://doi.org/10.1080/13562517.2013.827653>
- Data Never Sleeps 10.0 | Domo*. (é. n.). Elérés 2024. június 16., forrás: <https://www.domo.com/data-never-sleeps>

- Dijcks, J.-P. (2011). *Oracle: Big data for the enterprise* [White paper]. Oracle Corporation. <https://www.oracle.com/technetwork/database/bi-datawarehousing/wp-big-data-with-oracle-521209.pdf>
- International Educational Data Mining Society. (2018). Educationaldatamining.Org. <https://educationaldatamining.org/>
- Otoo-Arthur, D., & van Zyl, T. L. (2020). A scalable heterogeneous big data framework for e-learning systems. 2020 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems, icABCD 2020 – Proceedings. Scopus. <https://doi.org/10.1109/icABCD49160.2020.9183863>
- Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *WIREs Data Mining and Knowledge Discovery*, 10(3), e1355. <https://doi.org/10.1002/widm.1355>
- Society for Learning Analytics Research (SoLAR). (2024). Society for Learning Analytics Research (SoLAR). <https://www.solaresearch.org/>
- Szűts Z., & Yoo J. (2016). Big Data, az információs társadalom új paradigmája. *Információs Társadalom*, 16(1), 8. <https://doi.org/10.22503/inftars.XVI.2016.1.1>
- Yeoman, I. (2019). Big Data. *Journal of Revenue and Pricing Management*, 18(1), 1–1. <https://doi.org/10.1057/s41272-019-00191-9>