

Oktatási adattudomány

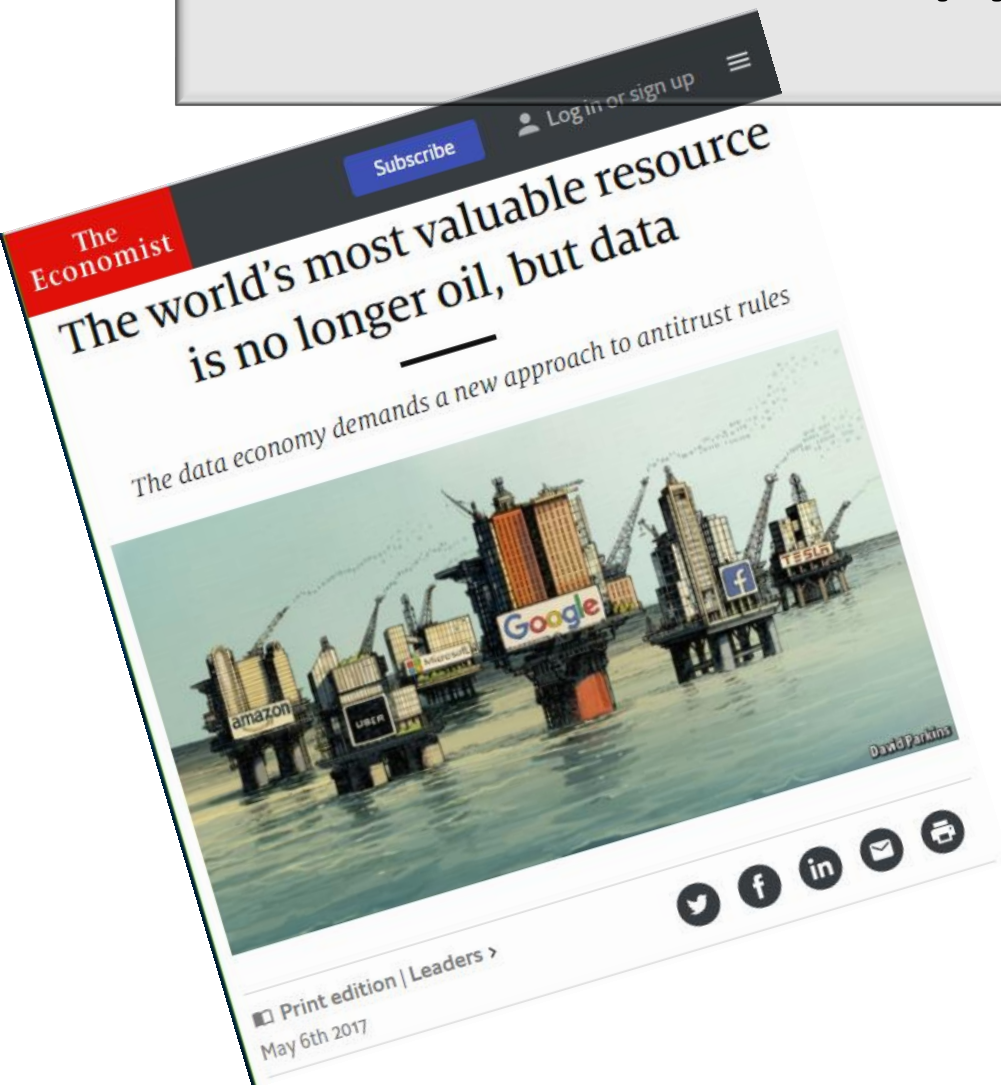
Molontay Roland,
BME, Matematika Intézet,
Sztochasztika Tanszék



M Ű E G Y E T E M 1 7 8 2



Adattudomány



Nagy mennyiségű adatokban rejlő érdekes, értelmes, értékes információk kinyerése különféle (elsősorban gépi tanulási) algoritmusok segítségével

index

TUDOMÁNY

A gépi tanulás
múlik

hvg.hu

Friss hírek Itthon Világ

2013. március 03. 15:00 SZT

A jövő üz

szerző:
hvg.hu

4

Ajánlom



BOLCSÓ DÁNIEL

KÖVETÉS

2017.11.20. 09:40

A héten Budapesten tartott előadást Margriet Gro
Watson adatplatform fejlesztője. A szakterülete a kl
elemzéssel vizsgálja nem várt környezeti események
adatvizualizációs eszközöket is fejleszt. A Bud
rendezvény két progr

index

TECH

A fél Szilícium-völgynek egy magyar cég bányász
adatokat



GÖTZ ATTILA

2017.01.23. 21:16

Egyre több olyan adat keletkezik a világban, ami Budapesten nyer értelmet.
Nemzetközi nagyvállalatok, köztük Fortune 500-as óriások egy hazai IT-
tanácsadó cég segítségével próbálnak megbirkózni a hatalmas
márciómennységgel, ami üzleti folyamataik során keletkezik. A magyar
értékes exportcikké vált, a budapesti csapat ügyfelei között
mint a Facebook, az Apple, a Tesla, a Netflix, a Disney
amokban 110-

masba szöktek,
killekkel

rében a
kat a
nára
zö

idő
resetebb

amokban 110-

Új korszak kezdődött a tudományban

SIMON TAMÁS | 2013.10.15. 17:59

Ez a cikk 4 éve frissült utoljára. A benne szereplő információk mára elavultak lehetnek.

Irdatlan mennyiségben laborokban, részecskegyártó távcsöveknél és számítástechnikai társadalomtudományok segítségével teszteli, hogy a tudósok korábban feltevéseire bukkanjanak-e összefüggésekre.

Breadth of Knowledge

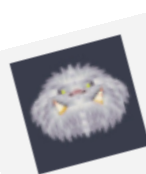
Classic researcher:
T-shaped

Domain
Specialty

Modern researcher:
Pi-shaped

Domain
Specialty
Stats/
Computing

Depth of Knowledge



Uj Péter
erdelyip

Sarkadi Zsolt

2016. január 6., szerda 15:31

- Óriási változás megy végbe a világ tudományi életében, a **Johns Hopkins Egyetem** pr...
- Minden évben annyi tudományos publikáció jelenik meg, mint az összesen az emberi történelem során.
- Ma az adatok tárolása és feldolgozása a tudomány alapja, és ez vezetett át Szalayt a kozmológiától az adattudomány területére.
- Az új adattudomány minden természettudomány alapja, és beépül az oktatásba is.

Adattudományi kutatás- fejlesztési projektek

- Statisztikai és adattudományi kutatás-
fejlesztések vállalati és akadémiai
partnerek részére



Oktatási adattudomány



- Nagy potenciál van a rengeteg rendelkezésre álló oktatási adatban is
 - Online tanulási szoftverek adatai, különféle Moodle-típusú rendszerek, teszteredmények, egyetemi tanulmányi rendszerek adatai
- Külföldön ez egy létező „tudományág”: külön konferenciával, külön folyóirattal stb.
 - Educational data mining/science, learning analytics
- Magyarországon több intézményben és az Oktatási Hivatalban nagyjából a miénkkel egy időben indultak ilyen irányú kezdeményezések

BME - Neptunos adatok bányászata

- Matematika Intézet + Központi Tanulmányi Hivatal együttműködése 2016 szeptembere óta
- Cél: a BME (döntően a Neptunban rendelkezésre álló) rendkívül nagy adatkincséből matematikailag korrekt eszközök felhasználásával oktatáspolitikai, oktatás-szervezési következtetések levonása
- Hozzáférés a neptunos adatkincshez
 - Szigorú titoktartási, adatkezelési nyilatkozat aláírása után



Budapesti Műszaki és Gazdaságtudományi Egyetem (BME)
Központi Tanulmányi Hivatal



Részvevők - KTH

Központi Tanulmányi Hivatal



**Dr. Szabó Mihály,
Igazgató**



Csabay Bálint



Bognár István

Részvevők – Matematika Intézet

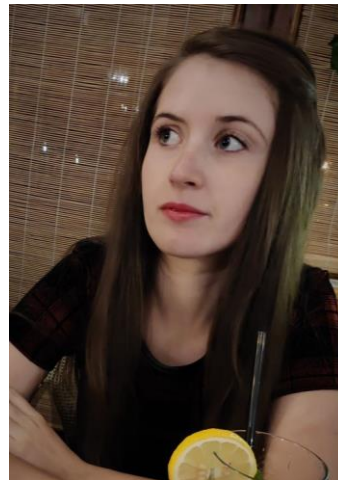
Matematika Intézet



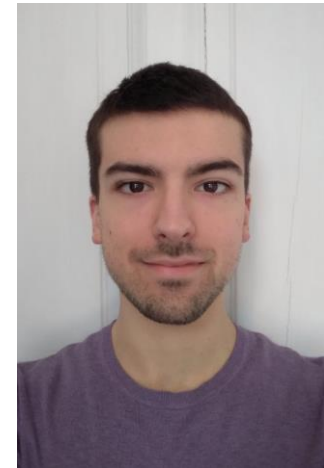
Molontay Roland



Nagy Marcell



Horváth Noémi



Baranyi Máté

Résztevők – MSc hallgatók

**17 hallgató kapcsolódott be a kutatásokba
Témalabor tárgy keretein belül**



Hallgatókkal közös eredmények

- TDK dolgozatok
 - Két OTDK I. díj
 - Három TDK I. díj és egy TDK II. díj
 - Egy Pro Progressio Különdíj és egy Rektori Különdíj
- Hallgatók által tartott előadások nemzetközi konferenciákon
 - Spanyolország, Olaszország, Csehország, Szlovákia
- Nemzetközi konferenciacikkek
- Hazai és nemzetközi folyóiratcikkek
- **Több működő webalkalmazás!**



Dashboard fejlesztése, adatvizualizációs megoldások

- Interaktív adatvizualizáció
- Döntéstámogatást elősegítő kimutatások
- Webes alkalmazások



Hallgatói csoportok teljesítményének alapstatisztikái

946

HALLGATÓK SZÁMA

34.84

FELVETT KREDIT
ÁTLAG

24.86

MEGSZERZETT
KREDIT ÁTLAG

3.39

TANULMÁNYI ÁTLAG

2.94

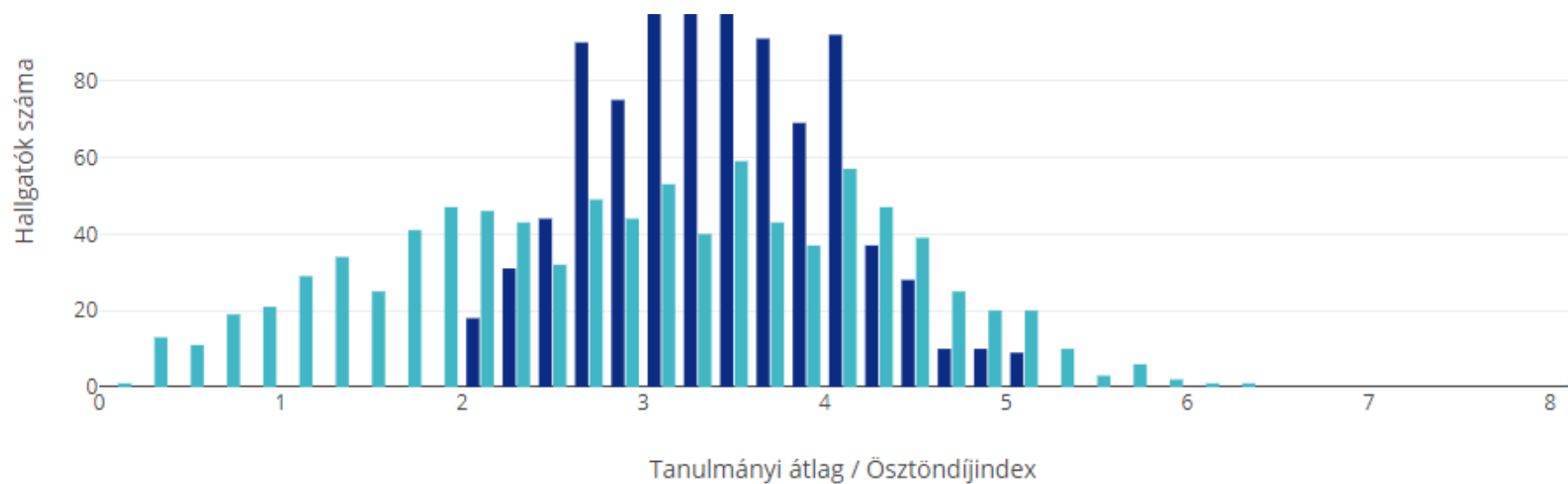
ÖSZTÖNDÍJINDEX

3.32

KUM. TANULMÁNYI
ÁTLAG

131.72

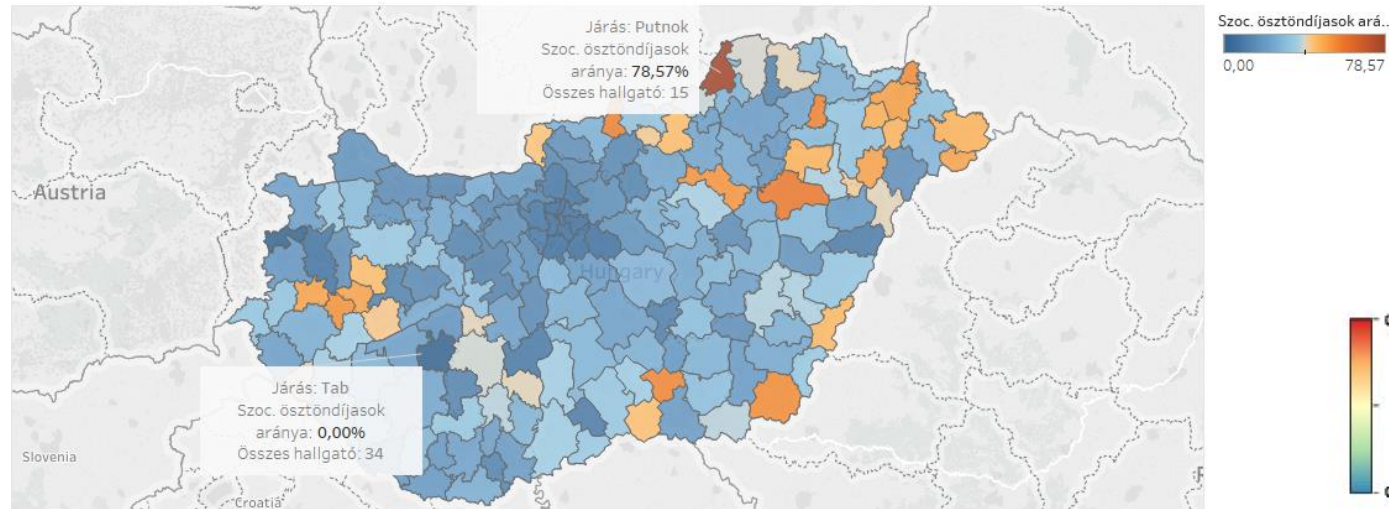
KUM. MEGSZERZETT
KREDIT



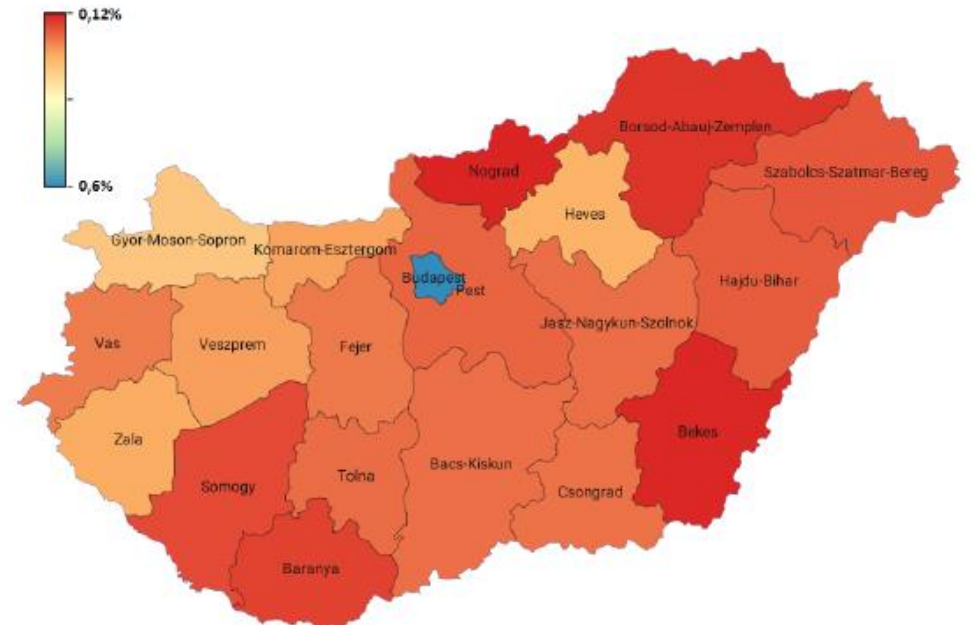
Horváth Noémi

Térképes adatvizualizáció

Szociális ösztöndíj

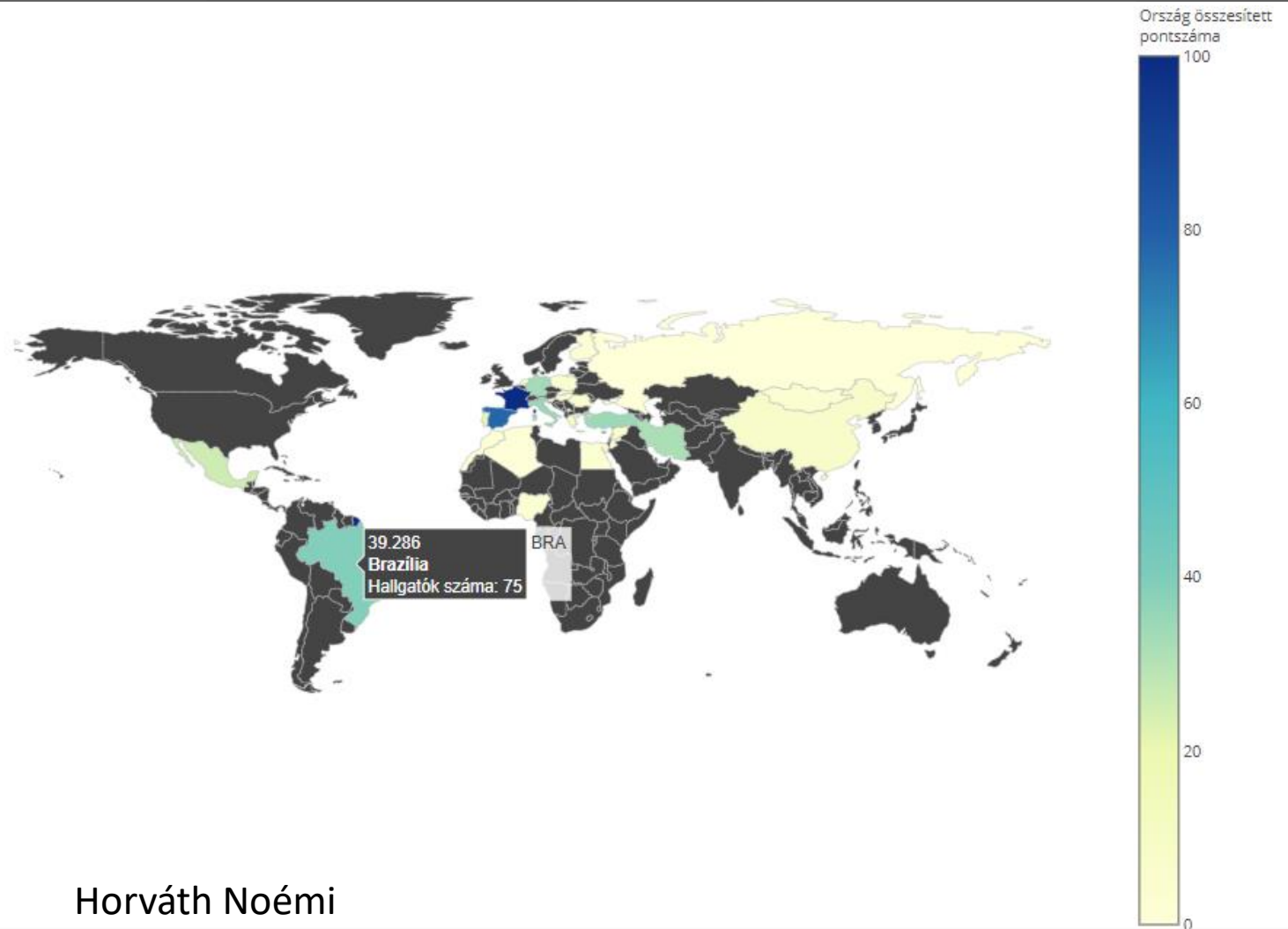


Map based on Longitude (generated) and Latitude (generated). Color shows $AVG([Szocos]) * 100$. Details are shown for Járás (group).



Nagy Marcell

Külföldről érkező hallgatók teljesítménye



Mennyibe kerül egy diploma?

P
O
Z
I
T
Í
V



2,236,452 Ft

ÁLLAMI KÖLTSÉG



31,200 Ft

HALLGATÓI ÖNERŐ



590

VÉGZETT HALLGATÓK



4.29

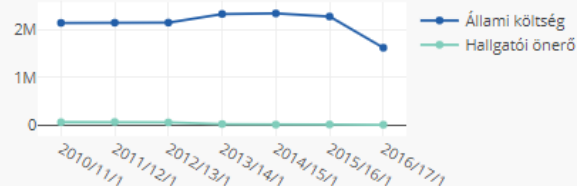
DIPLOMAÁTLAG



7.3

AKTÍV FÉLÉVEK

ÁLLAMI KÖLTSÉG ÉS HALLGATÓI ÖNERŐ A FELVÉTEL FÉLÉVE ALAPJÁN



VÉGZETT HALLGATÓK ARÁNYA



EGY DIPLOMA ÁRA

2,533,384 Ft

183,485 Ft

73,328 Ft



1,022,682 Ft

ÁLLAMI KÖLTSÉG



192,677 Ft

HALLGATÓI ÖNERŐ



129

ELBOCSÁTOTT HALLGATÓK



1.6

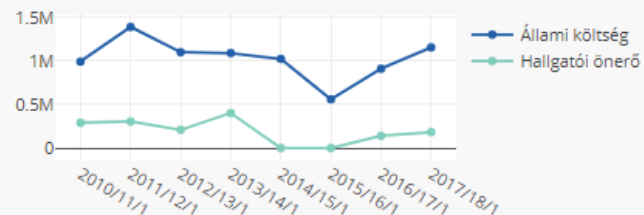
PASSZÍV FÉLÉVEK



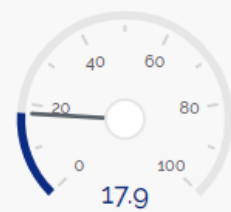
3.4

AKTÍV FÉLÉVEK

ÁLLAMI KÖLTSÉG ÉS HALLGATÓI ÖNERŐ A FELVÉTEL FÉLÉVE ALAPJÁN



KIBUKOTT HALLGATÓK ARÁNYA



HALLGATÓNKÉNTI VESZTESÉG

218,054 Ft

183,485 Ft

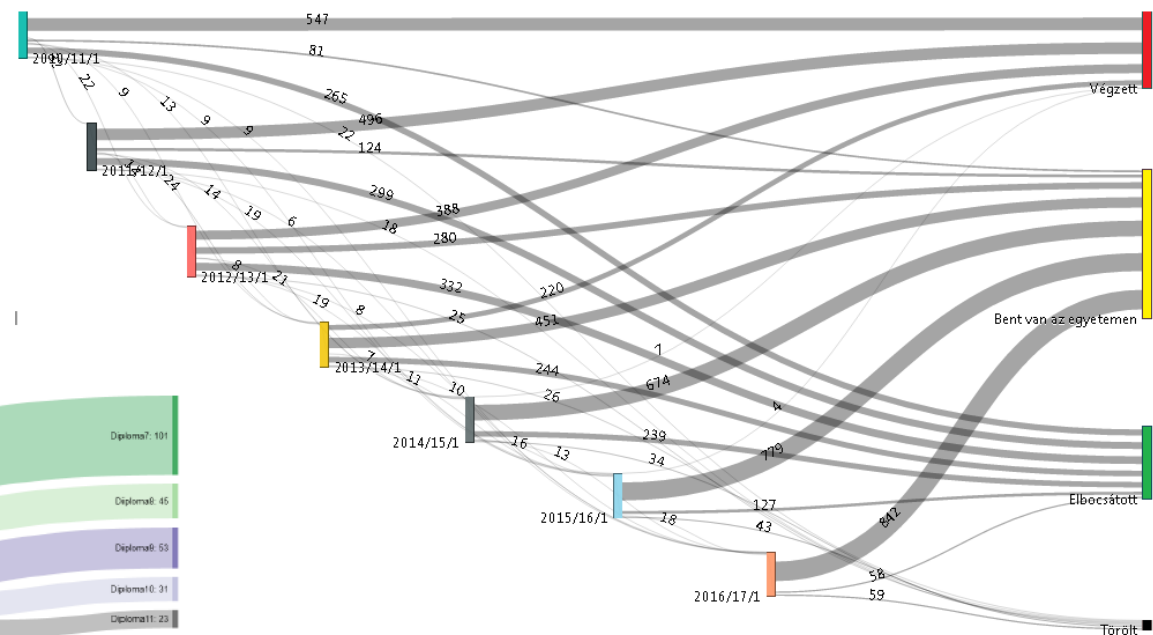
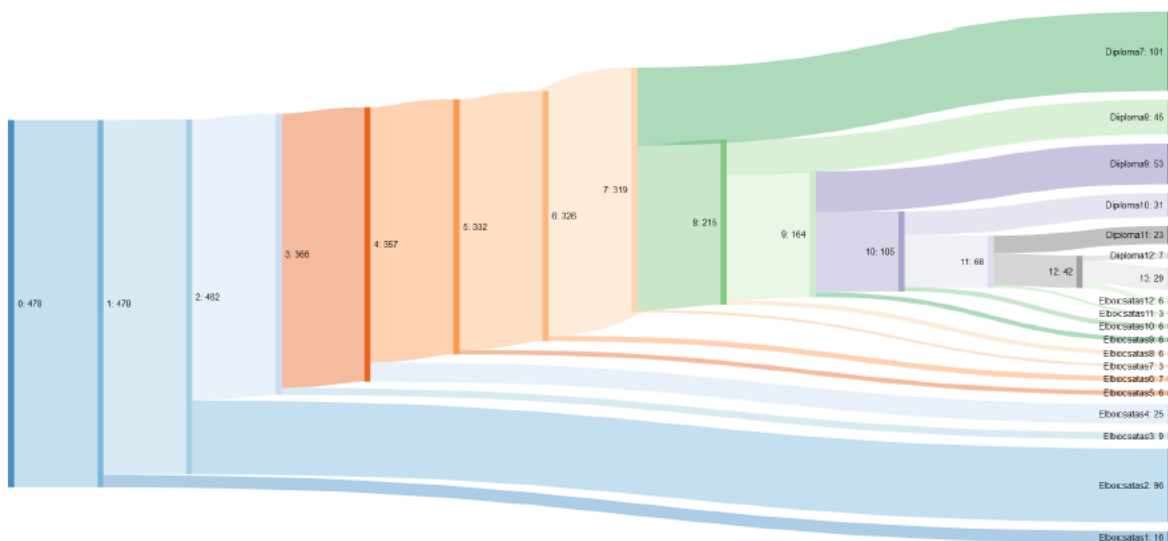
34,569 Ft

- Mennyibe kerül (államnak/hallgatónak) egy diploma?
- Mekkora össze megy „veszendőbe” a lemorzsolódás miatt?
- Mennyibe kerül egy „ elveszett” kredit?

N
E
G
A
T
Í
V

Hallgatói áramok vizualizációja

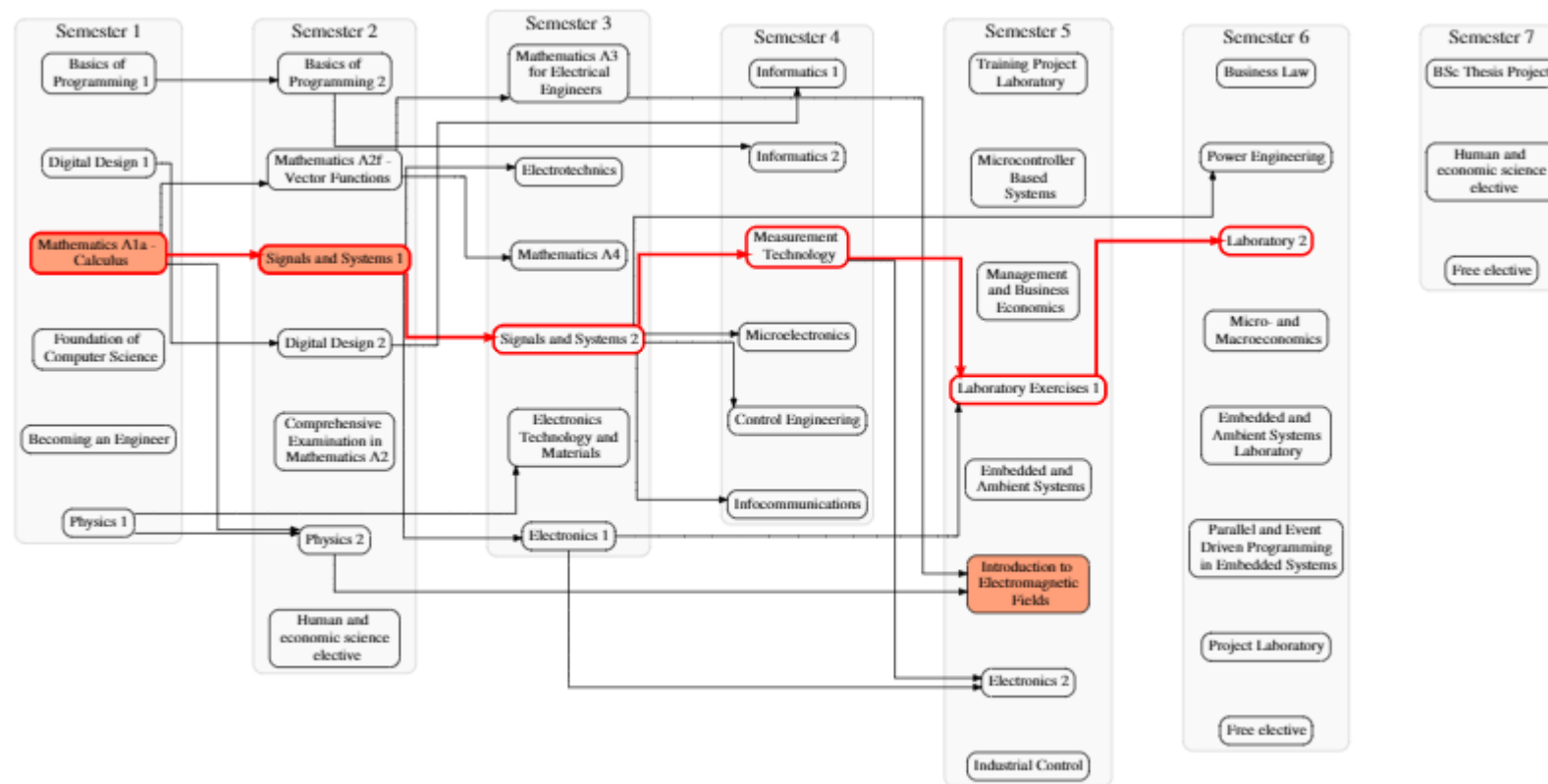
- Hallgatói áramok félévenkénti vizsgálata Sankey-diagrammal



Benyhe Ágota, Horváth Dániel, Horváth Noémi

Tantervi hálók elemzése

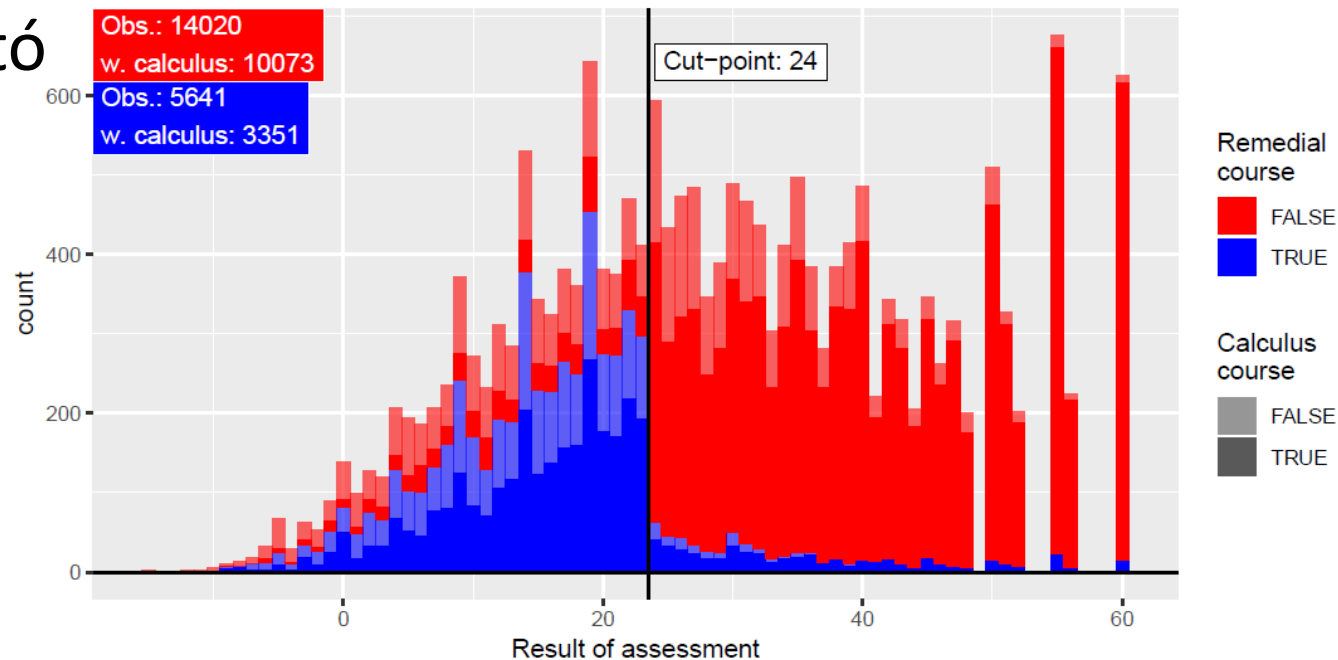
- Tantervi hálók struktúrája milyen hatással van a végzési időre
- Mely tárgyak teljesítési arányának változása van a legnagyobb hatása a végzési időre
- Lemorzsolódásra, késleltetett végzésre leginkább befolyással bíró beavatkozási pontok azonosítása



Horváth Noémi, Bergmann Júlia, Szekrényes Dóra

A felzárkóztatás sikerességének vizsgálata

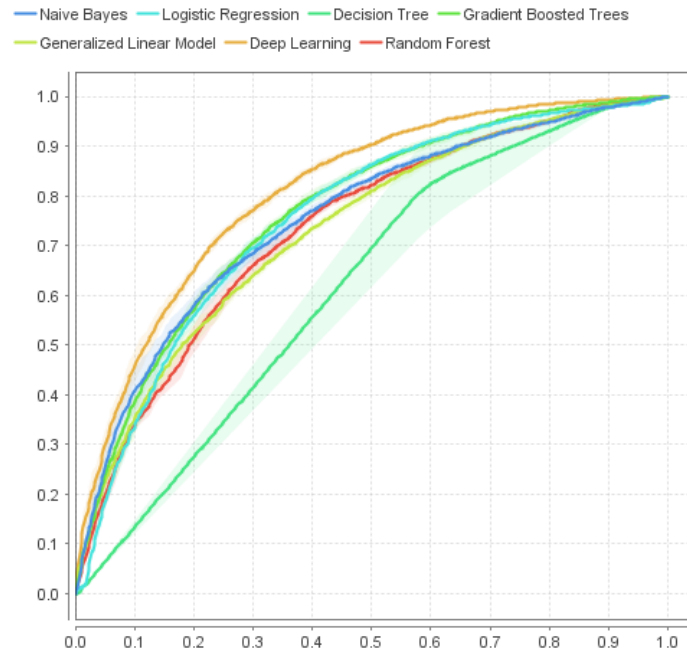
- A matematika felzárkóztató tárgy hatása a későbbi teljesítményre
 - Regression Discontinuity Design ökonometriai módszertan alkalmazásával
- Eredmények
 - Kicsi, de statisztikailag szignifikáns pozitív hatást mértünk



Baranyi Máté

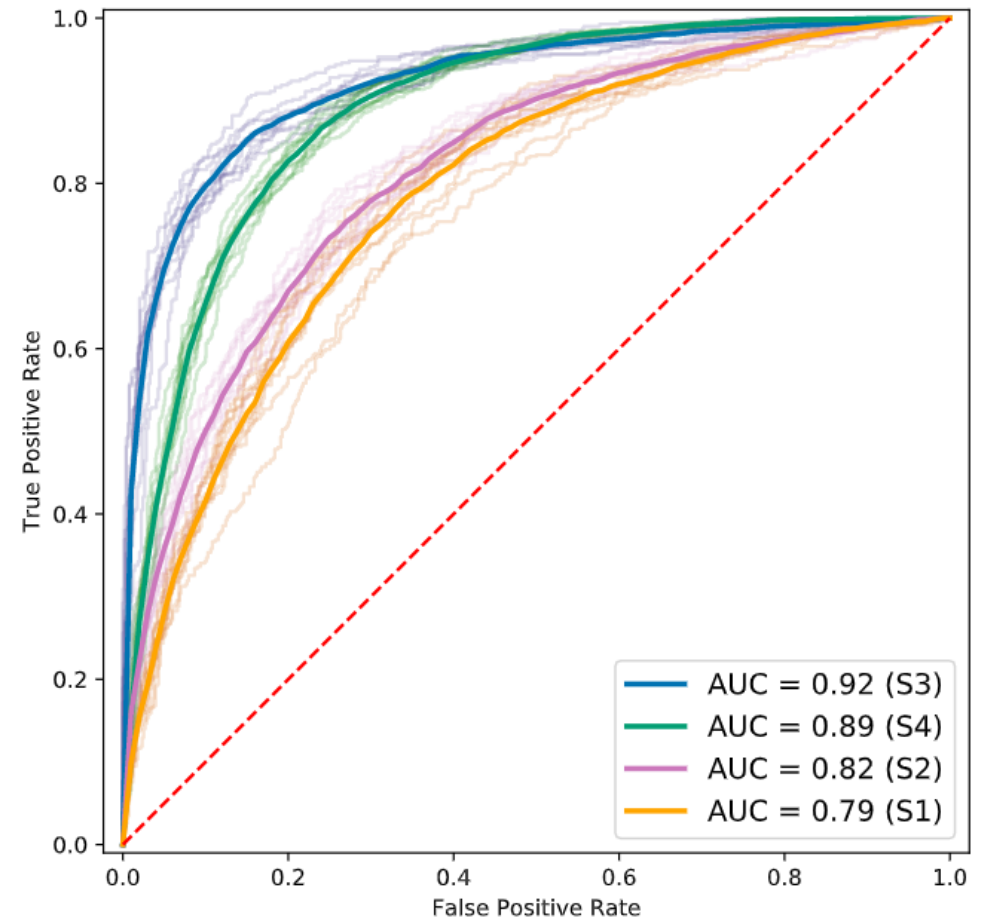
Lemorzsolódó hallgatók korai szűrése prediktív analitikai módszerekkel

- Historikus adatok alapján prediktív analitikai módszerekkel vizsgáljuk, mennyire jelezhető elő a beiratkozás pillanatában, hogy a hallgató sikeresen el fogja-e végezni az egyetemet
 - Így a lemorzsolódásban veszélyeztetett hallgatókat tudjuk beazonosítani
- Jól teljesítenek a modellek



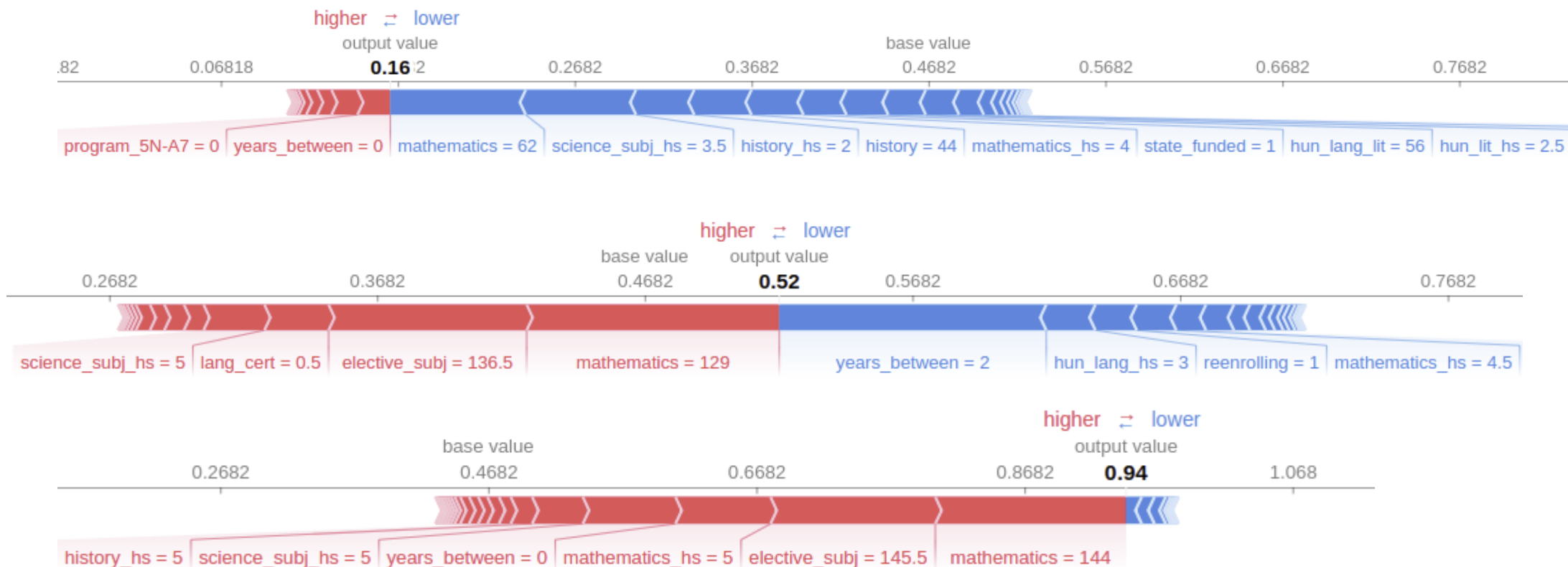
Lemorzsolódás előrejelzése az első féléves adatok ismeretében

- Mennyivel növeli a prediktív erőt, ha a beiratkozás pillanatában meglévő adatokhoz hozzávesszük az félév tanulmányi adatait is?
- Van-e inkrementális prediktív ereje a középiskolai tanulmányoknak az első félév eredményeihez képest?



Az attribútumok hatása a predikcióra vonatkozóan – értelmezhető gépi tanulás

- SHAP érték segítségével



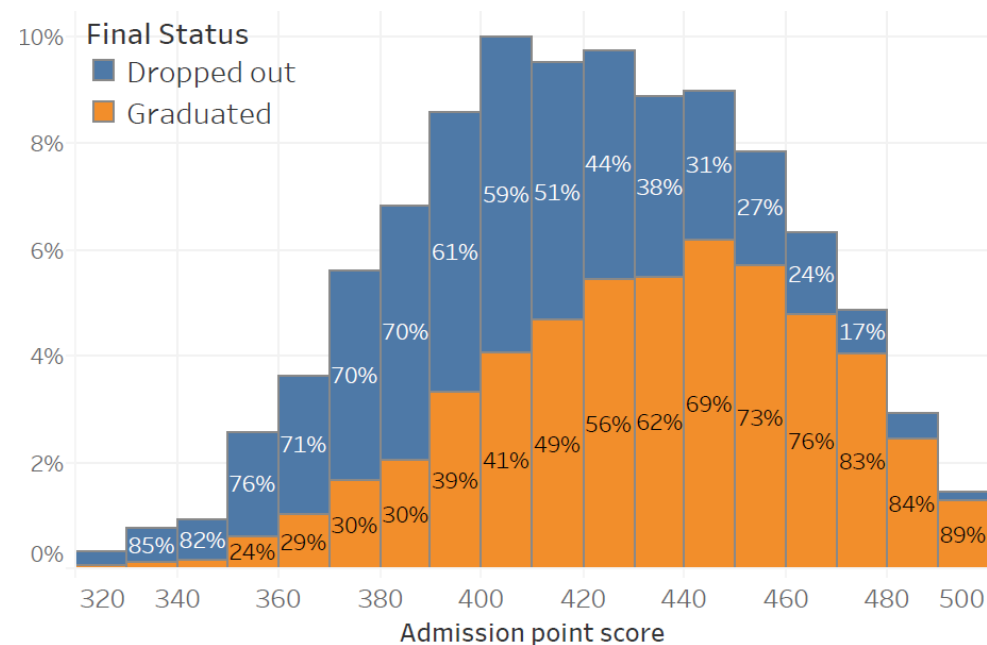
Hogyan használjuk az eredményeket?

- Etikai kérdések
- Szabad-e közölni a hallgatónak, hogy az algoritmus azt jósolja, ki fog bukni az egyetemről?
- Milyen jó/rossz felhasználási módja lehet az eredményeknek?



Felvételi pontszám prediktív ereje

- A felvételi pontszám mennyire bír jó prediktív erővel a későbbi egyetemi teljesítményre vonatkozóan?
 - Elég jó prediktív erővel bír
 - A prediktív ereje képzési területenként eltérő
 - A humán tárgyak mérnöki területeken is elég nagy prediktív erővel bírnak
 - A felvételi pontszám kis módosításaival még nagyobb prediktív erő érhető el



Nagy Marcell

Középiskolai rangsor készítése

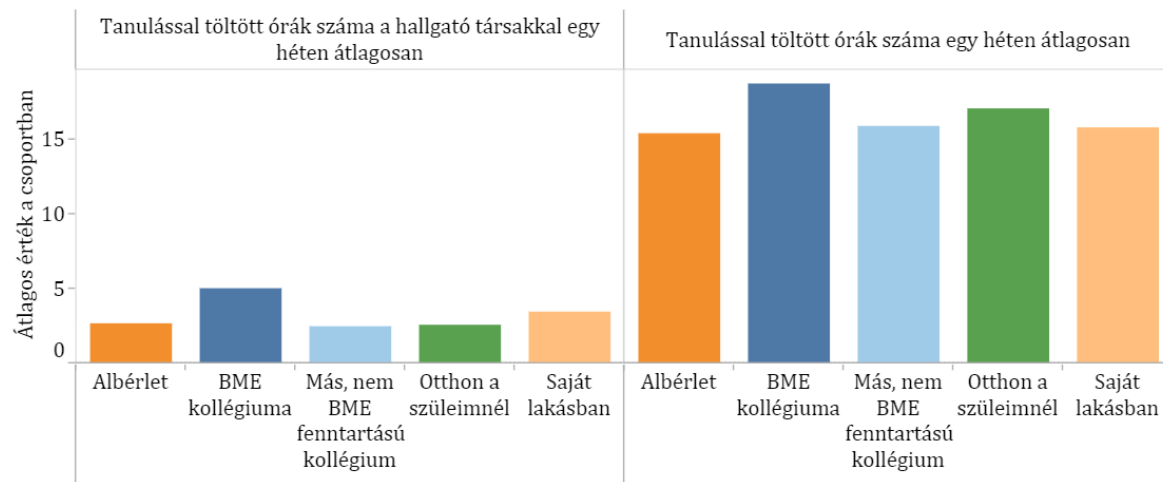
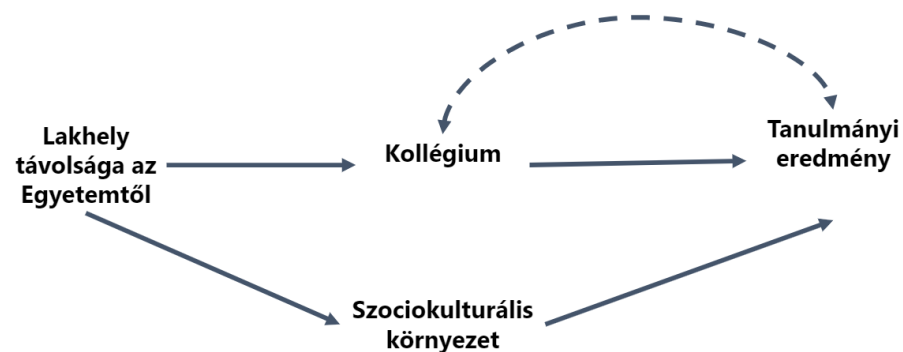
- Hazai középiskolák újszerű módszertanon alapuló rangsorának elkészítése
- A szokásos mutatókon túl figyelembe veszi az adott középiskolákból érkező egyetemi teljesítményét is
- A BME számára kiemelt jelentőségű középiskolák beazonosítása
- Terület-specifikus alrangsorok létrehozása

High school	Our rank	HVG rank	legjobbiskola.hu
Eötvös József Gimnázium, Budapest	1	3	3
Teleki Blanka Gimnázium, Székesfehérvár	2	60	269
Fazekas Mihály Fővárosi Gyakorló Ált. Isk. és Gimnázium, Budapest	3	4	10
Veres Péter Gimnázium, Szakközépiskola és Szakmunkásképző, Budapest	4	10	37
Janus Pannonius Gimnázium és Szakközépiskola, Pécs	5	43	161
Lovassy László Gimnázium, Veszprém	6	8	1
Belvárosi Általános Iskola és Gimnázium, Békéscsaba	7	100	171
Radnóti Miklós Kísérleti Gimnázium és Általános Iskola, Szeged	8	15	4
Apáczai Csere János Gyakorló Gimnázium (ELTE), Budapest	9	6	30
Városmajori Gimnázium, Budapest	10	11	20

Horváth Noémi

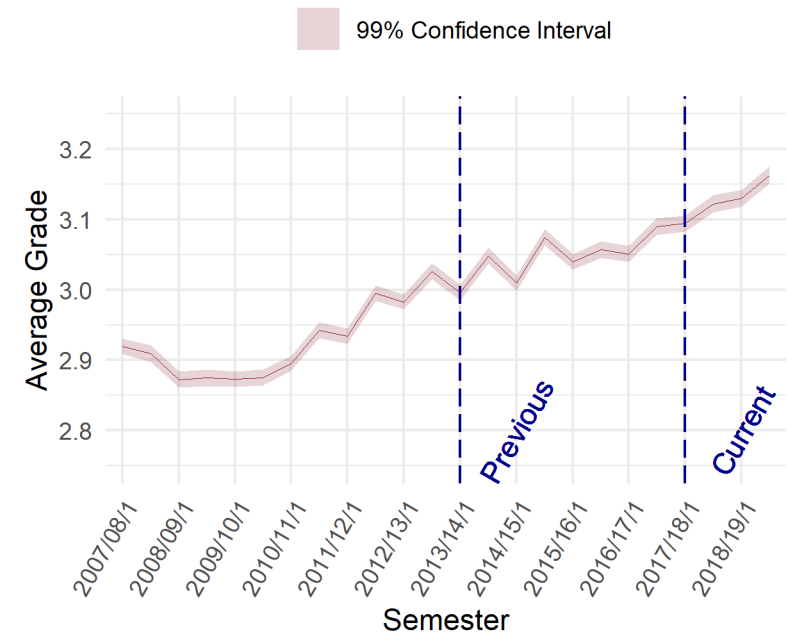
A kollégiumi lét hatása az egyetemi teljesítményre

- A kollégiumi lét milyen hatással van a hallgatók egyetemi teljesítményére
 - Endogenitás kezelése
 - Szignifikáns pozitív hatást mértünk többféle modellspecifikációval
 - Kérdőíves adatgyűjtéssel a hatás csatornáinak beazonosítása

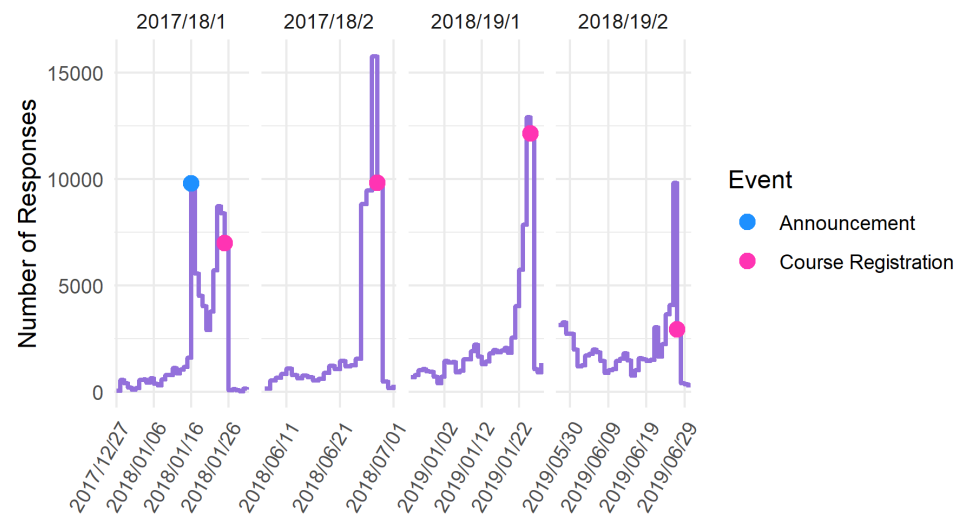
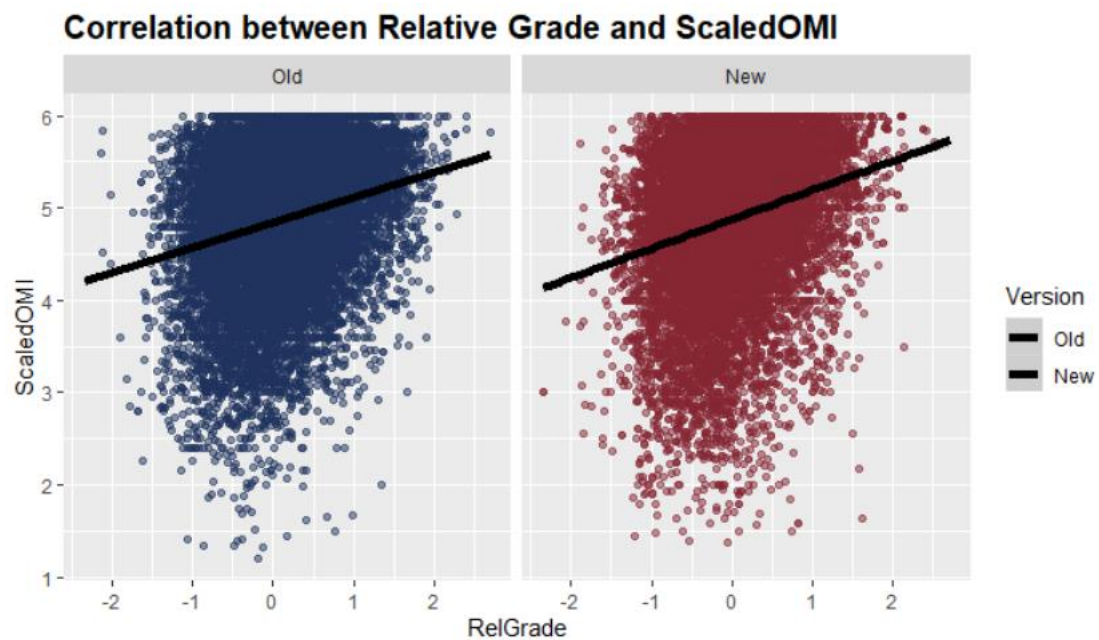


Jegyinfláció és az OHV értékelések kapcsolatának vizsgálata

- Van-e összefüggés a tantárgyi jegyeloszlás és az adott kurzus oktatójának OHV eredményei között
 - Szofisztikált ökonometriai módszerekkel azt mértük, hogy átlagosan, ha az osztályzat egy jeggyel jobb, akkor a hallgatók 0,25-tel jobb értékelést adnak az oktatónak
 - OLS regresszió, fixhatás becslés, TSLS regresszió

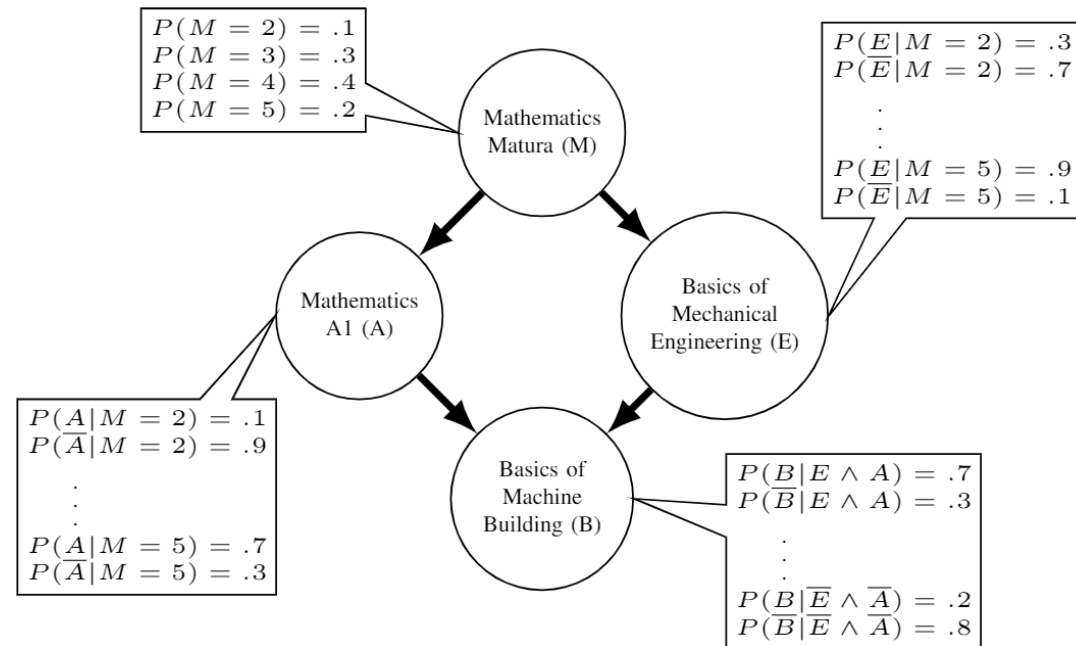


Az OHV reform hatásainak statisztikai elemzése



Feltételes összefüggőségek feltárása a hallgatók különböző teljesítménymutatói között

- Bayes-hálókkal felderíteni a személyes adatok, a középiskolai eredmények és az egyetemi teljesítmény közti összefüggéseket



Gál Kristóf

További irányok

- A középiskola típusa (szakközépiskola/gimnázium) és fenntartója (állami/egyházi/alapítványi) hogyan befolyásolja a későbbi egyetemi teljesítményt
 - Dornai Zsófia
 - Kovács Szilvia
- Gimnazisták és szakgimnáziumból érkező hallgatók teljesítményének összehasonlítása az első félévekben tantárgyi lebontásban
 - Szőke Márton
- Konkrét tantárgyak sikeres teljesítése milyen tényezőktől függ leginkább
 - Borgula Annamária
 - Major Levente
- Tanulmányi átlagok előrejelzése (regresszió)
 - Török Roland
- Szociális helyzet és egyetemi teljesítmény összefüggései
 - Barabás Zoltán



Csoportunk publikációi

- Dániel Horváth, Roland Molontay, Mihály Szabó: Visualizing student flows to track retention and graduation rates, 22nd IEEE International Conference on Information Visualisation, pp. 338-343, 2018
- Máté Baranyi, Roland Molontay: Effect of mathematics remediation on academic achievement – a regression discontinuity approach, IEEE International Symposium on Educational Technology (ISET), pp. 29-33, 2019
- Marcell Nagy, Roland Molontay: Predicting Dropout in Higher Education based on Secondary School Performance, 22nd IEEE International Conference on Intelligent Engineering Systems, pp. 389-394, 2018
- Marcell Nagy, Roland Molontay, Mihály Szabó: A web application for predicting academic performance and identifying the contributing factors, SEFI 47th Annual Conference, 1794-1806, 2019
- Kiss Botond, Nagy Marcell, Molontay Roland, Szabó Mihály: Predicting Dropout Using High School and First-semester Academic Achievement Measures, 17th International Conference on Emerging eLearning Technologies and Applications, 2019
- Máté Baranyi, Kristóf Gál, Roland Molontay, Bálint Csabay: Modeling Students' Academic Performance Using Bayesian Networks, 17th International Conference on Emerging eLearning Technologies and Applications, 2019
- Noémi Horváth, Roland Molontay, Mihály Szabó: Who are the most important „suppliers” for universities? - Ranking secondary schools based on their students' university performance, 2nd Danube Conference for Higher Education Management: In search of excellence in higher education, pp 133-143, 2019
- Zombor Berezvai, Gergely Lukáts, Roland Molontay: A pénzügyi ösztönzők hatása az egyetemi oktatók osztályozási gyakorlatára, 66/7, pp 733-750, 2019
- Bergmann Júlia, Molontay Roland, Szabó Mihály, Szekrényes Dóra: Kreditrendszerű képzések mintatanterveinek és előtanulmányi hálóinak elemzése a hazai matematika alapszakok példáján, Alkalmazott Matematikai Lapok (megjelenés alatt)
- Roland Molontay, Noémi Horváth, Júlia Bergmann, Dóra Szekrényes, Mihály Szabó: Characterizing curriculum prerequisite networks by a student flow approach, IEEE Transactions on Learning Technologies (megjelenés alatt)
- Zombor Berezvai, Gergely Dániel Lukáts, Roland Molontay: Can Professors Buy Better Evaluation for More Lenient Grading? - The Effect of Grade Inflation on Student Evaluations of Teaching (előkészületben)
- Marcell Nagy, Roland Molontay: Predictive Validity of the Hungarian University Entrance Score and its Modifications (előkészületben)

Érdekel az oktatási adattudomány?

Szívesen csatlakoznál?



Keressetek bátran!
molontay@math.bme.hu



Kérdések?

Köszönöm szépen a figyelmet!