

Mintatanterv a hallgatói teljesítések tükrében

Curriculum reflected by student achievements

Dr. TERDIK György

Debreceni Egyetem Informatikai Kar
4028 Debrecen, Kassai út 26., Magyarország,
<http://www.inf.unideb.hu>, Terdik.Gyorgy@inf.unideb.hu

Abstract

One of the key focuses of university education is high-quality undergraduate teaching. There are well-established, standardized methodologies for evaluating educational quality. In this paper, we adopt an unconventional approach to highlight the effectiveness of education through the relationship between the curriculum and student performance. By conducting a statistical analysis of student grades, we employ a graphical modeling approach to infer potential dependency structures among academic subjects.

Keywords: Graphical model, Conditional independence, Partial correlation, Curriculum, Student performance.

Kivonat

Az egyetemi oktatás egyik súlypontja a színvonalas alapképzés. A minőség megítélésének vannak bevált jól működő szabványos módszerei. Jelen közleményünkben egy nem szokványos megközelítést választunk, hogy rávilágítsunk a mintatanterv és a hallgatói teljesítmény kapcsolatán keresztül az oktatás hatékonyságára. A hallgatói érdemjegyek statisztikai vizsgálata során egy grafikus modell segítségével következtetünk a tárgyak közötti lehetséges függőségi kapcsolatrendszerre.

Kulcsszavak: Grafikus modell, feltételes függetlenség, parciális korreláció, tanrend, hallgatói teljesítmény.

1. BEVEZETÉS

1.1. A mintatanterv

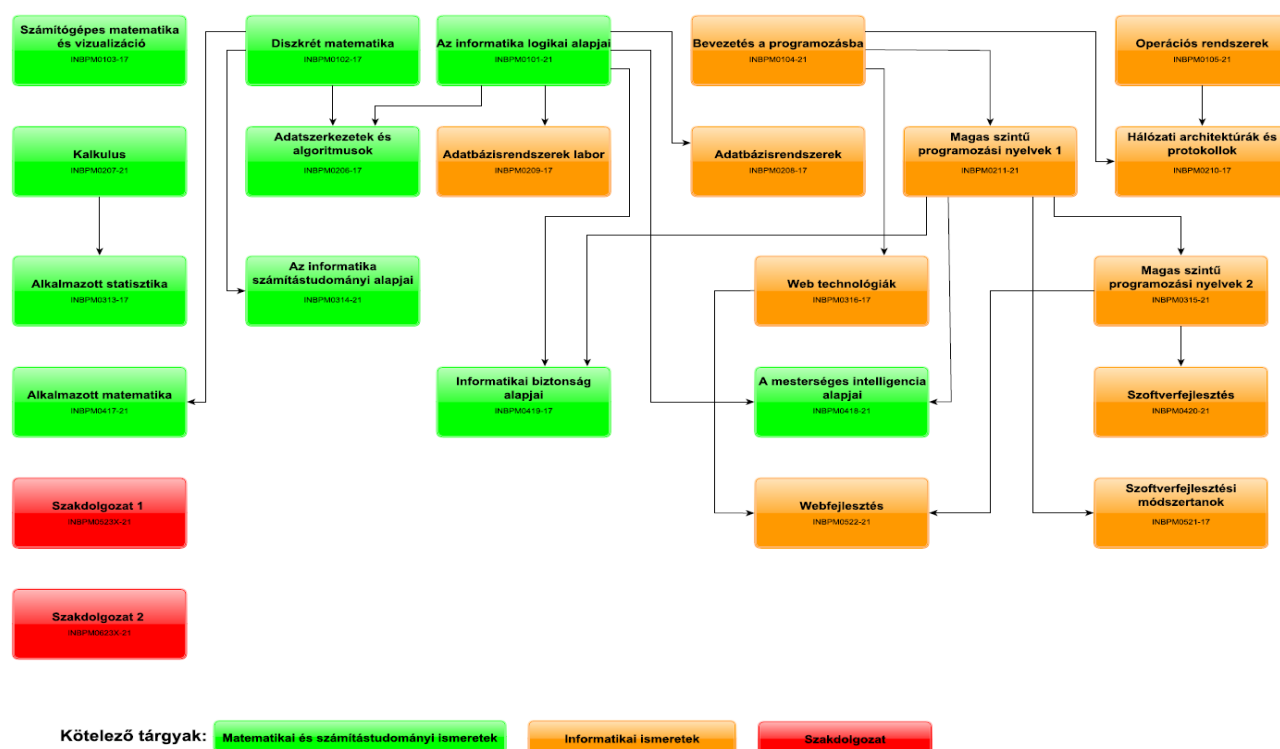
A programtervező informatikus alapszak az egyik meghatározó képzés a Debreceni Egyetem Informatikai Karán, lsd. [1]. Az angol és magyar nyelvű képzés párhuzamosan van meghirdetve, mi csak a magyar nyelvű képzésre szorítkozunk és annak is csak a kötelező tantárgyaira. A hallgatók számára is hasznos képzési gráf, 1. ábra, mutatja hogyan épülnek egymásra a meghirdetett tárgyak. A kötelező tárgyak alapvetően két csoportra oszlanak: Matematikai és számítástudományi ismeretek (zöld) és Informatikai ismeretek (barna). A nyílak mutatják a feltételeket az illető tárgy felvehetőségének, illetve az előzetes ismeretek kötelező elsajátításának. A sorok az ajánlott tárgyakat tartalmazzák félévenként.

1.2 Adatok

Az adatok az egyetem tanulmányi rendszeréből származnak. Az utolsó három tanévben, azaz hat félévben álltak rendelkezésünkre a hallgatók által szerzett érdemjegyek, ami nálunk 1-től 5-ig terjedő osztályzatot jelent, a 'nem teljesítés' bejegyzést 1-nek tekintettük. A minta elemszáma, azaz az összes osztályzat 13919 volt. Érdemes megjegyezni, hogy az osztályzatok nem egyenletesen oszlanak el a tantárgyak között. Köszönhető ez a tantárgyfelvétel rugalmasságának, illetve a hallgatói létszám változásának stb. Az 1. táblázat mutatja az egyes tárgyak teljesítésének adatait százalékosan. Az adatok feldolgozása során az egyes tantárgyakhoz tartozó osztályzatokat ötös csoportokba rendeztük, és ezek átlagával számoltunk. Ennek elsősorban kényelmi oka volt mert így nem diszkrét hanem folytonos változókkal számoltunk. Továbbá a normális eloszlás is feltételezhető lett. A csoportokból minden esetben azonos számút vettünk 86-ot, ami elegendő a statisztikai vizsgálathoz.

Az 1. táblázat önmagában is tartalmaz hasznos információt a képzésre vonatkozóan, de a jelen munkának ez nem tárgya.

Programtervező informatikus BSc szak



1. ábra PTI BSc, tantervi háló

2. STATISZTIKAI MODELL A FÜGGŐSÉG FELDERÍTÉSÉRE

A tantervi háló (1. ábra) mutatja, hogy a tantárgyak, ha csak a felvehetőségüket tekintjük is, függenek egymástól. A kérdés, hogy a tantárgyak teljesítése során ez a függőség megmutatkozik-e, illetve milyen egyéb információk fedezhetők fel a tárgyak kölcsönös kapcsolatából?

A szokásos jól ismert mérőszám két véletlen például X_1 és X_2 mennyiség függetlenségére a korreláció

$$\rho_{X_1, X_2} = \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var } X_1 \text{ Var } X_2}}.$$

Példaként mutatjuk a 'Matematikai és számítástudományi ismeretek' csoport korrelációs mátrixát, 2. táblázat. Amiből kiolvasható az összes lehetséges párosítás viszonya. A teljes kapcsolat rendszer ennél természetesen sokkal bonyolultabb hiszen mindezen közben figyelembe kell venni a többi tárgy teljesítését is. Ez azt jelenti, hogy az összes tárgyat egy vektor változónak tekintjük: $\mathbf{X}$, és az X_p , X_q koordinátákat a $\mathbf{X}_{\setminus(p,q)}$ maradék változók rögzítése mellett vizsgáljuk. A feltételes korreláció

$$\rho_{X_p, X_q | \mathbf{X}_{\setminus(p,q)}} = \frac{C_{X_p, X_q | \mathbf{X}_{\setminus(p,q)}}}{\sigma_{X_p | \mathbf{X}_{\setminus(p,q)}} \sigma_{X_q | \mathbf{X}_{\setminus(p,q)}}}.$$

mutatja a két koordináta, jelen esetben a két tárgy feltételes függőségét miközben a többi tárgy is jelen van. Ha $\rho_{X_p, X_q | \mathbf{X}_{\setminus(p,q)}} = 0$ akkor az X_p , X_q változókat feltételesen korrelálatlannak nevezzük. Itt jegyezzük meg, hogy míg a korrelációk mátrixát a $\Sigma_{\mathbf{X}}$ kovariancia mátrixból számíthatjuk, hasonlóan a feltételes korreláció mátrixot is az Ω feltételes kovariancia mátrixból számítjuk. Általában a feltételes kovariancia számítása elméletileg és statisztikailag is bonyolult feladat. Azonban, ha feltételezzük, hogy $\mathbf{X}$ többváltozós normális eloszlású akkor az *Inverz Szórás Lemma*, [2] 141 old., egy egyszerű számítást állít: $\Omega = \Sigma_{\mathbf{X}}^{-1}$.

Nevezetesen a feltételes korreláció mátrixa két lépésben megkapható: 1. kovariancia mátrix invertálásával 2. az így kapott mátrix skálázásával. Annyi kiegészítést érdemes tenni, hogy az így kapott mátrix (p,q) eleme a $\rho_{X_1, X_2|X_{\setminus(p,q)}}$ feltételes korreláció negatívja. Természetesen a legérdekesebb eset a feltételes függetlenség $\rho_{X_1, X_2|X_{\setminus(p,q)}} = 0$, ami a normalitásból következik.

A Ω feltételes kovariancia/korrelációs mátrixot precíziós (koncentrációs) mátrixnak szokás nevezni. Érdemes megjegyezni, hogy eredetileg $\rho_{X_1, X_2|X_{\setminus(p,q)}}$ -t parciális korrelációnak is nevezik, [3], és a normális eloszlás ill. a lineáris modellek esetén ma is így hivatkozunk rá.

1. táblázat – A tárgyak teljesítése

		1	2	3	4	5
MestI	A mesterséges intelligencia alapjai	37.53	25.89	17.12	14.25	5.21
ABázR	Adatbázisrendszerek	46.94	21.7	16.1	10.49	4.78
ABázL	Adatbázisrendszerek labor	19.02	29.45	22.85	15.64	13.04
Aszerk	Adatszerkezetek és algoritmusok	42.19	15.52	16.67	7.19	18.44
AMat	Alkalmazott matematika	23.4	26.43	20.54	13.97	15.66
AStat	Alkalmazott statisztika	33.74	42.78	10.26	8.23	4.99
ILog	Az informatika logikai alapjai	12.72	14.94	15.26	27.19	29.89
ISzAI	Az informatika számítástudományi alapjai	8.32	13.93	18.96	33.27	25.53
BevPr	Bevezetés a programozásba	11.3	8.22	16.1	20.55	43.84
DMat	Diszkrét matematika	12.98	31.7	18.38	15.51	21.42
HálóA	Hálózati architektúrák és protokollok	24.82	17.28	24.37	16.84	16.69
IBizt	Informatikai biztonság alapjai	36.22	20.7	22.8	13.99	6.29
Kalk	Kalkulus	24.35	21.73	22.65	16.8	14.48
MSzP1	Magas szintű programozási nyelvek 1	39.9	7.06	8.15	16.79	28.1
MSzP2	Magas szintű programozási nyelvek 2	15.35	19.88	24.8	19.88	20.08
OpRen	Operációs rendszerek	21.48	12.63	18.87	17.85	29.17
SzgViz	Számítógépes matematika és vizualizáció	14.35	17.1	20.65	20	27.9
SzFej	Szoftverfejlesztés	23.55	27.13	27.65	15.02	6.66
SzMód	Szoftverfejlesztési módszertanok	4.81	1.14	3.66	30.21	60.18
Wtech	Web technológiák	44.79	31.36	12.47	6.54	4.84
WebF	Webfejlesztés	11.94	6.79	33.02	18.03	30.21

2. táblázat – Korrelációs mátrix, 'Matematikai és számítástudományi ismeretek'

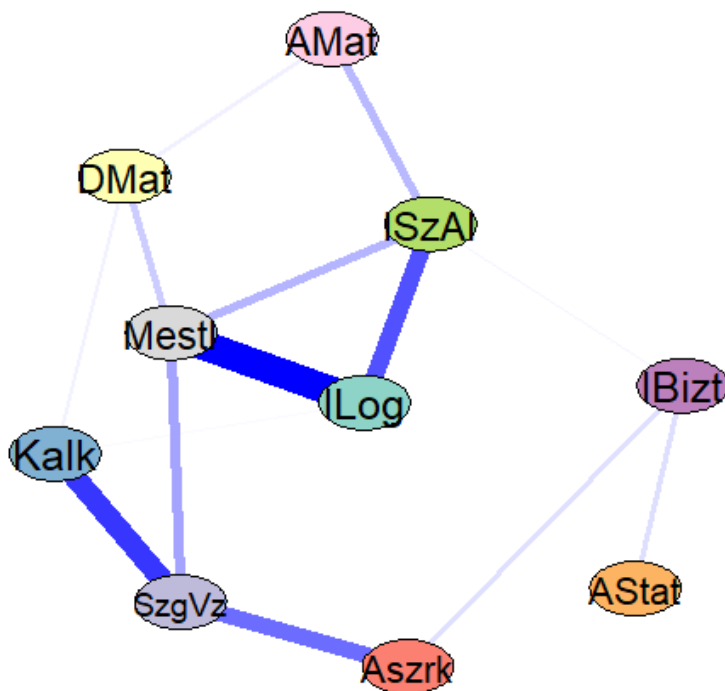
	ILog	DMat	SzgViz	Aszerk	Kalk	AStat	ISzAI	AMat	MestI	IBizt
ILog	1	0.036	0.132	0.134	-0.14	-0.111	-0.225	0.115	-0.265	-0.057
DMat	0.036	1	0.046	0.063	-0.145	-0.076	-0.091	-0.146	-0.163	-0.09
SzgViz	0.132	0.046	1	0.211	-0.238	-0.091	0.069	0.038	-0.184	-0.054
Aszerk	0.134	0.063	0.211	1	-0.097	-0.099	0.008	-0.009	0.014	0.154
Kalk	-0.14	-0.145	-0.238	-0.097	1	0.07	0.13	-0.083	0.117	-0.003
AStat	-0.111	-0.076	-0.091	-0.099	0.07	1	-0.016	-0.024	-0.024	-0.154
ISzAI	-0.225	-0.091	0.069	0.008	0.13	-0.016	1	-0.173	0.175	-0.143
AMat	0.115	-0.146	0.038	-0.009	-0.083	-0.024	-0.173	1	0.087	0.042
MestI	-0.265	-0.163	-0.184	0.014	0.117	-0.024	0.175	0.087	1	-0.012
IBizt	-0.057	-0.09	-0.054	0.154	-0.003	-0.154	-0.143	0.042	-0.012	1

3. táblázat: Ω precíziós mátrix, 'Matematikai és számítástudományi ismeretek'

	ILog	DMat	SzgViz	Aszerk	Kalk	AStat	ISzAI	AMat	Mestl	IBizt
ILog	1	0.048	-0.039	-0.142	0.058	0.128	0.19	-0.099	0.242	0.138
DMat	0.048	1	0.021	-0.074	0.125	0.098	0.104	0.153	0.133	0.131
SzgViz	-0.039	0.021	1	-0.201	0.208	0.071	-0.13	-0.057	0.184	0.083
Aszerk	-0.142	-0.074	-0.201	1	0.035	0.026	-0.038	0.037	-0.101	-0.185
Kalk	0.058	0.125	0.208	0.035	1	-0.032	-0.1	0.072	-0.031	0.001
AStat	0.128	0.098	0.071	0.026	-0.032	1	0.065	0.015	0.077	0.177
ISzAI	0.19	0.104	-0.13	-0.038	-0.1	0.065	1	0.174	-0.13	0.167
AMat	-0.099	0.153	-0.057	0.037	0.072	0.015	0.174	1	-0.135	-0.016
Mestl	0.242	0.133	0.184	-0.101	-0.031	0.077	-0.13	-0.135	1	0.062
IBizt	0.138	0.131	0.083	-0.185	0.001	0.177	0.167	-0.016	0.062	1

Ami a feltételes korreláció elnevezést esetünkben indokolja az a precíziós mátrix grafikai megjelenítése. Ezt a 'Matematikai és számítástudományi ismeretek' csoport precíziós mátrixán keresztül (3. Táblázat) mutatjuk meg. A precíziós mátrixhoz tartozó gráf az un. Markov modell elvén alapul. A gráf csúcspontjai a tantárgyak. Ha két csúcs között nincs él akkor feltételesen függetlenek az összes többire nézve. Viszont feltételnek elegendő szabni csak a két csúcsot összekötő legrövidebb útvonalon lévő csúcsokat. Ilyen módon például az Alkalmazott statisztika feltételesen független az Adatszerkezetek és algoritmusoktól, ha az Informatikai biztonság alapjai adott. Hasonlóképpen következtetés vonható le a tantárgyak csoportjairól is, például a Számítógépes matematika és vizualizáció két csoportra osztja a tárgyakat így ezek a csoportok feltételesen függetlenek, ha a Számítógépes matematika és vizualizáció értéke adott. Az élek vastagsága arányos a feltételes függőség mértékével. A gráf jelentősen csökkenti a tárgyak közötti lehetséges bonyolult viszonyrendszert. Fontos megjegyezni, hogy a gráf szerkesztése során a feltételes függetlenség, azaz $\rho_{X_1, X_2 | X_{\setminus \{p, q\}}} = 0$ megállapításához tartozó p érték nem a statisztikai próba szigorú szabályai szerint lett megállapítva, hanem egy olyan küszöbérték *megadásával*, ami könnyebb érthetőséget biztosít az elemzéshez. A Ω precíziós mátrix becslésére van direkt statisztikai módszer GLASSO (Graphical Least Absolute Selection and Shrinkage Operator) [4].

Feltételes Függőség Gráfja - Alapozó Tantárgyak



2. ábra. A precíziós mátrix megjelenítése

3. ÖSSZEFOGLALÁS

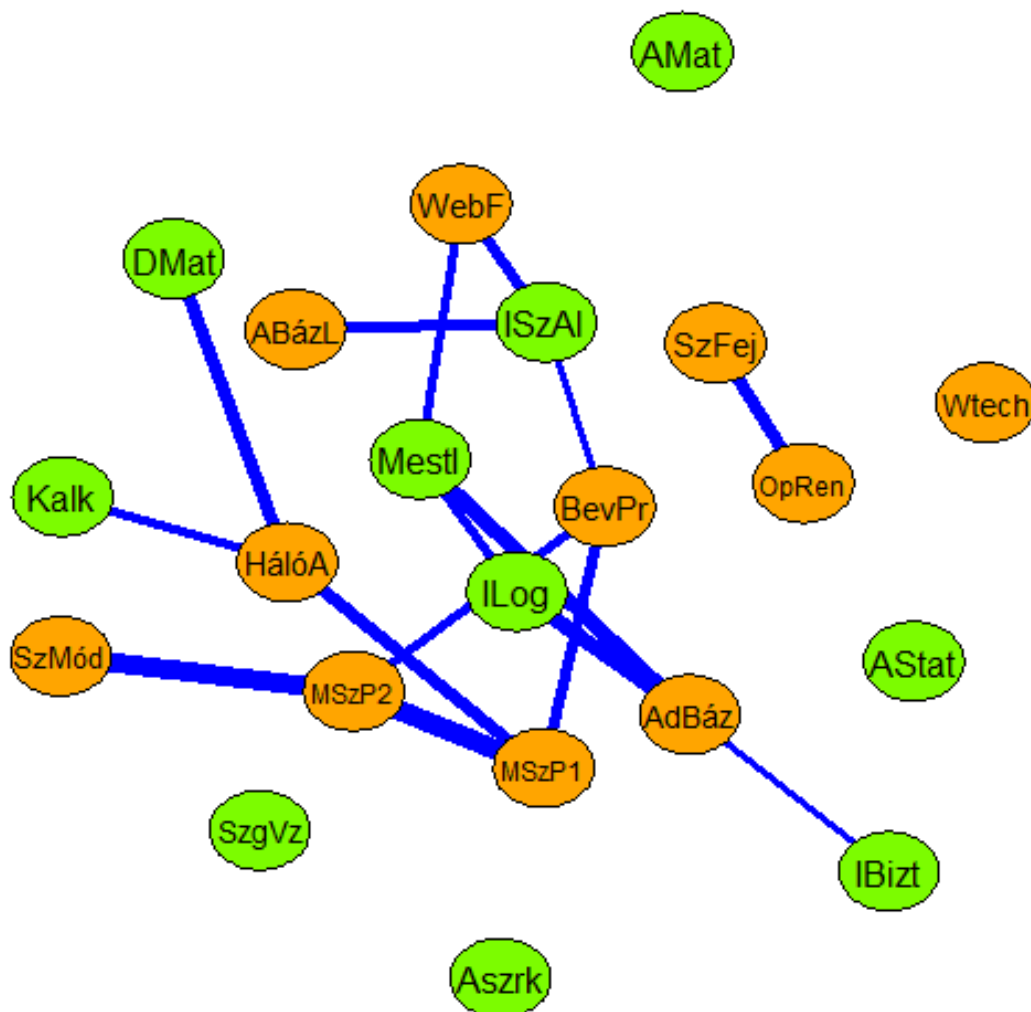
A PTI BSc, tantervi hálózhoz, 1. ábra, tartozó tantárgyak grafikai megjelenítése, 3. ábra, számos következtetésre ad lehetőséget. A teljesség és kritikai szándék nélkül fűzünk ehhez az ábrához néhány megjegyzést.

1. Négy alaptantárgy: 'Alkalmazott matematika', 'Alkalmazott statisztika', 'Adatszerkezetek és algoritmusok', 'Számítógépes matematika és vizualizáció' és egy informatikai tárgy a 'Web technológiák' szeparálódnak a többitől, azaz feltételesen nem függenek tőlük. Hasonlóképpen viselkedik a két tárgyból, 'Operációs rendszerek' és 'Szoftvertechnológia', álló csoport.

2. A 'Diszkrét matematika' és a 'Kalkulus' csak a 'Hálózati architektúrák' és protokollok-on keresztül kapcsolódik a nagyobb csoporthoz. Illetve az 'Informatikai biztonság alapjai' az 'Adatbázisrendszerek'-en keresztül.

3. 'A mesterséges intelligencia alapjai', 'Az informatika számítástudományi alapjai' és 'Az informatika logikai alapjai' tárgyak szervesen illeszkednek az informatikai tárgyak közé.

Feltételes Függettség Gráfja



3. ábra. *PTI BSc, tantervi gráfja*

KÖSZÖNETNYILVÁNÍTÁS

Köszönettel tartozom Márton Ágnesnek az Informatikai Kar Tanulmányi Osztály vezetőjének az adatok biztosításáért és Abdaljbbar B. A. Dawod PhD hallgatónak az R kódhoz nyújtott segítségért.

Ezt a munkát a Debreceni Egyetem QoS-HPC-IoT Laboratórium és a TKP2021-NKTA-34 projekt támogatta. A TKP2021-NKTA-34 számú projekt a Kulturális és Innovációs Minisztérium Nemzeti Kutatási Fejlesztési és Innovációs Alapból nyújtott támogatásával, a TKP2021-NKTA pályázati program finanszírozásában valósult meg.

IRODALMI HIVATKOZÁSOK

- [1] Programtervező informatikus (BSc) szak, mintatanterv, https://inf.unideb.hu/sites/default/files/inline-files/PTI%20BSc%202021%20tanterv_2025_1.pdf
- [2] Kendall, M. and Stuart, A. *The advanced theory of statistics*, London: Griffin, 4th ed. 1977.
- [3] Whittaker, J. *Graphical Models in Applied Multivariate Statistics*. Wiley, 1st edition, 2009
- [4] Tibshirani R. Regression shrinkage and selection via the Lasso, *Journal of the Royal Statistical Society: Series B (Methodological)*, 58 (1996), 267–288.