

OTDK-dolgozat

2025

**KODOLÁNYI JÁNOS EGYETEM
ÚJMÉDIA ÉS KREATÍVIPARI KAR**

**TÖBBKRITÉRIUMÚ DÖNTÉSTÁMOGATÓ KONCEPCIÓRENDSZER
– FEJLESZTÉS A TDK-S TÉMAVÁLASZTÁS OPTIMALIZÁLÁSÁRA**

**CRITERION-BASED DECISION SUPPORT SYSTEM FOR
OPTIMIZING TOPIC SELECTION OF THE SCIENTIFIC STUDENT
ASSOCIATIONS CONFERENCE**

Dénes Attila Norbert

Konzulens: Pitlik László

Kézirat lezárásának dátuma: 2025. 11. 21.

Bevezetés

Ez egy rendhagyó tudományos dolgozat, amely kísérletként megpróbálja az előírásuk tudatos áttekintését és kikerülését, annak minden várható pozitív és negatív következményeit felvállalva. A Tudományos Diákköri Konferencián való indulás egyik legegységesebb és legmeghatározóbb lépése a kutatási téma kiválasztása. Azonban ez különösen nehéznek bizonyulhat egy frissen a tudományos világ tengerébe csöppenő fiatal számára. Tovább növeli e nehézség mértékét az időszűkösség, az egyetemi profilkapcsolódási korlátozottságok, és az a kritikus tényező is, hogy nincs egy objektív rendszer, amely segítene rangsorolni a különböző opciókat, lehetőségeket. Ezen kutatás célja utóbbi kidolgozása, egy olyan rendszer megalkotása, amely objektív mérések által segít a valóban releváns, innovatív, nagy hatékonysággal és lehetőségrendszerrel bíró TDK-téma kiválasztásában és megvalósításában. Ennek megoldására szóelemzést fogok alkalmazni, olyan kulcsszavakat kiválasztva, amelyek a már említett célkitűzések mentén kerültek kialakításra.

Szakirodalmi alapok

Szakirodalmak tekintetében az Magyar Internetes Agrár / Alkalmazott Informatikai Újság megjelent rezümé és TDK objektivitását, tudományos célszerűségeit tartalmazó írásainak olvasására fordítottam kiemelt figyelmet, s ezekből tanulva igyekeztem egy objektív és innovációközpontú tízelemes kulcsszócsomót megalkotni, amelyek illeszkedését aztán elemezhetem a különböző interneten fellelhető TDK és OTDK dolgozatokon. Kulcsfontosságú volt, hogy először a kulcsszavakat állítsam össze, s utána kezdjem megnézni a rezüméket, mivel így elkerültem, hogy tudatos tapasztalatokból és befolyásolódásokból szubjektív módon, egyes elemeket esetlegesen önkaraton kívül is előtrébe helyezzek.

A rezümék kiválasztása

A rezümék kiválasztásához az internet eszköztárait hívtam segítségül. Főként az online fellelhető REZÜMÉKÖTET 2023. április 20-22 fájlából válogattam, illetve a <https://miau.my-x.hu/miau2009/index.php3> nyilvánosan elérhető eszköztárából. A kiválasztásnál több AI rendszert is segítségül hívtam, hogy a választott elemek területi eloszlása között legyen egyfajta arányosság.

Kulcsszavak és elemzések

A kulcsszavak megjelenésének mértékét szintén igyekeztem AI eszközökkel megvizsgálni. Azonban több mint 5 eszközt kipróbálva változatos promptolásokkal, szembesülnöm kellett azzal, hogy erre még nem alkalmasak a rendszerek. Így, kissé talán favágó módszerrel, de maradt a CTRL+F billentyűkóddal való elemzés. Az effektív munka érdekében egy dokumentumba gyűjtöttem az elemeket, s ott futtattam le a kereséseket, majd egy táblázatba foglaltam össze az egyesével is leellenőrzött adatokat. A szavakra igyekeztem olyan módon keresni, hogy alaktól független megjelenítődjenek (pl. innov., ez magában foglalja innováció, innovatív stb.). A rezümé címeit nem vettem bele az elemzésbe, csak a tartalmukat

Látszólagos eredmények

Ezek után kialakult egy eredménylista-féleség. Ez az a pont ahol természetesen még egyáltalán nem lehet megállni, mivel fontos kérdés, hogy az adott rezümé hny szót tartalmaz. A következő lépés tehát a szószám és kulcsszavak előfordulásának arányának a vizsgálata volt. Itt megint manuálisan végeztem el a szavak számának ellenőrzését, a rezümé (- cím) tartalmának kijelölésével és Word-os szószám megállapításával. Ezek után jött a százalék-számítás. Ezt szintén próbáltam a mesterséges intelligenciákat felhasználva és összevetve elvégeztetni. A következő AI-kat használtam: Grok, ChatGPT, Kimi, Gemini, Qwen. Örömmel láttam, hogy ezt a feladatot az összes AI megegyező eredményekkel végezte el, csupán a tizedesvesszőben szerepeltek különbségek. Így sikerült kialakítanom a következő táblázatokat:

[illegible]

20 0% 0.5917% 0% 0.5917% 0.2959% 0% 0% 0% 0% 0.5917%

Ezek alapján pedig a következő táblázat százalékarányának csökkenő sorrendbe helyezésével már ki is lehetne alakítani egy TDK-s dobogót:

ID	szószám	kulcsszavak_db	kulcsszavak_%
1	231	8	3.4632%
2	232	2	0.8621%
3	237	9	3.7975%
4	235	4	1.7021%
5	154	11	7.1429%
6	232	4	1.7241%
7	156	6	3.8462%
8	156	1	0.6410%
9	211	8	3.7915%
10	205	9	4.3902%
11	297	8	2.6936%
12	305	10	3.2787%
13	203	9	4.4335%
14	250	12	4.8000%
15	132	9	6.8182%
16	334	5	1.4970%
17	206	6	2.9126%
18	195	2	1.0256%
19	173	0	0.0000%
20	338	7	2.0710%

Azonban e művelet elvégzése csupán egy látszólagos és súlyosan felületes eredményt adna, mivel nem minden kulcsszó bír ugyanolyan erővel.

Pontosítások, túl a naív becsléseken a matematika világáig

Az előző lista alapján szinte egyetlen kulcsszó ismételt használatával már nyerhetővé is válhatna a verseny. Ezért érdemes több matematikai műveletet is elvégezni, elemezve az elsődleges, rendszerint mindenkor látszólagos eredmény valósághűségét. Ennek eredményéül, különböző matematikai műveletek alkalmazásával a következő táblázatok jöttek létre:

ID	bizonyí	haték	szigni	modell	elemz	eredm	piac	alkalm	innov	gyakorl	Y0 (idealitás index)	naiv aggrega cio2	naiv sorrend2
1	4	7	3	4	11	4	3	5	6	4	1000000	5,1	4
2	4	7	3	8	6	13	8	5	6	4	1000000	6,4	12
3	4	1	3	8	7	1	8	5	6	4	1000000	4,7	1
4	3	5	3	8	13	10	8	5	6	4	1000000	6,5	14
5	4	2	3	1	17	7	8	5	1	4	1000000	5,2	5
6	4	7	3	8	12	13	4	5	6	3	1000000	6,5	14
7	4	7	1	8	5	8	8	5	2	4	1000000	5,2	5
8	4	7	3	8	8	13	8	5	6	4	1000000	6,6	17
9	4	7	3	8	4	9	8	3	6	2	1000000	5,4	8
10	4	4	3	5	10	2	8	2	6	4	1000000	4,8	2
11	2	7	3	3	14	13	8	5	6	4	1000000	6,5	14
12	4	6	3	8	15	6	5	1	4	4	1000000	5,6	10
13	4	7	3	2	17	3	1	5	6	4	1000000	5,2	5
14	4	7	2	6	2	11	8	4	6	4	1000000	5,4	8
15	1	7	3	8	1	5	8	5	6	4	1000000	4,8	2
16	4	7	3	8	9	12	7	5	5	4	1000000	6,4	12
17	4	7	3	8	3	13	6	5	3	4	1000000	5,6	10
18	4	7	3	8	17	13	2	5	6	4	1000000	6,9	19
19	4	7	3	8	17	13	8	5	6	4	1000000	7,5	20
20	4	3	3	7	16	13	8	5	6	1	1000000	6,6	17

inverz	inverz	inverz	inverz	inverz	inverz	inverz	inverz	inverz	inverz	inverz	validitás	objektív becslés	objektív sorrend	naiv sorrend2	naiv sorrend1	delta2	delta1	
17	14	18	17	10	17	18	16	15	17	1000000	1	1000007	4	4	9	0	-5	
17	14	18	13	15	8	13	16	15	17	1000000	1	999992	15	12	18	3	-3	
17	20	18	13	14	20	13	16	15	17	1000000	1	1000009	1	1	7	0	-6	
18	16	18	13	8	11	13	16	15	17	1000000	1	999991	16	14	15	2	1	
17	19	18	20	4	14	13	16	20	17	1000000	1	1000007	5	5	1	0	4	
17	14	18	13	9	8	17	16	15	18	1000000	1	999996	14	14	14	0	0	
17	14	20	13	16	13	13	16	19	17	1000000	1	1000007	5	5	6	0	-1	
17	14	18	13	13	8	13	16	15	17	1000000	1	999990	18	17	19	1	-1	
17	14	18	13	17	12	13	18	15	19	1000000	1	1000006	8	8	8	0	0	
17	17	18	16	11	19	13	19	15	17	1000000	1	1000009	2	2	5	0	-3	
19	14	18	18	7	8	13	16	15	17	1000000	1	999991	16	14	12	2	4	
17	15	18	13	6	15	16	20	17	17	1000000	1	1000005	9	10	10	-1	-1	
17	14	18	19	4	18	20	16	15	17	1000000	1	1000007	5	5	4	0	1	
17	14	19	15	19	10	13	17	15	17	1000000	1	1000003	11	8	3	3	8	
20	14	18	13	20	16	13	16	15	17	1000000	1	1000009	2	2	2	0	0	
17	14	18	13	12	9	14	16	16	17	1000000	1	999997	13	12	16	1	-3	
17	14	18	13	18	8	15	16	18	17	1000000	1	1000005	9	10	11	-1	-2	
17	14	18	13	4	8	19	16	15	17	1000000	1	999990	19	19	17	0	2	
17	14	18	13	4	8	13	16	15	17	1000000	1	999981	20	20	20	0	0	
17	18	18	14	5	8	13	16	15	20	1000000	1	1000000	12	17	13	-5	-1	
																max	3	8
																min	-5	-6

Ezekből kimutatható, hogy egy 20-as elemlistán 14!! helykülönbségnyi változás tapasztalható, ha túllépve a látszólagos eredményen az objektív és matematikai logikát is segítségül hívjuk.

Végleges variáció

Utóbbiak mentén sikerült a következő táblázat kialakítása:

COCO:Y0	X(A1)	X(A2)	X(A3)	X(A4)	X(A5)	X(A6)	X(A7)	X(A8)	X(A9)	X(A10)	Becsülés	Tény+0	Delta	Delta/Tény
O1	3,5	499974,9	499971,9	3	10	3	2,5	4	7,5	12,5	999992,8	1000000	7,2	0
O2	3,5	499974,9	499971,9	7	5	12	9,5	4	7,5	12,5	1000007,8	1000000	-7,8	0
O3	3,5	499968,9	499971,9	7	6	0	9,5	4	7,5	12,5	999990,8	1000000	9,2	0
O4	2	499972,9	499971,9	7	12	9	9,5	4	7,5	12,5	1000008,3	1000000	-8,3	0
O5	3,5	499969,9	499971,9	0	16	6	9,5	4	0	12,5	999993,3	1000000	6,7	0
O6	3,5	499974,9	499971,9	7	11	12	3,5	4	7,5	8	1000003,3	1000000	-3,3	0
O7	3,5	499974,9	499969,9	7	4	7	9,5	4	1	12,5	999993,3	1000000	6,7	0
O8	3,5	499974,9	499971,9	7	7	12	9,5	4	7,5	12,5	1000009,8	1000000	-9,8	0
O9	3,5	499974,9	499971,9	7	3	8	9,5	2	7,5	7	999994,3	1000000	5,7	0
O10	3,5	499971,9	499971,9	4	9	1	9,5	1	7,5	12,5	999991,8	1000000	8,2	0
O11	1	499974,9	499971,9	2	13	12	9,5	4	7,5	12,5	1000008,3	1000000	-8,3	0
O12	3,5	499973,9	499971,9	7	14	5	4,5	0	3	12,5	999995,3	1000000	4,7	0
O13	3,5	499974,9	499971,9	1	16	2	0	4	7,5	12,5	999993,3	1000000	6,7	0
O14	3,5	499974,9	499970,9	5	1	10	9,5	3	7,5	12,5	999997,8	1000000	2,2	0
O15	0	499974,9	499971,9	7	0	4	9,5	4	7,5	12,5	999991,3	1000000	8,7	0
O16	3,5	499974,9	499971,9	7	8	11	6,5	4	4	12,5	1000003,3	1000000	-3,3	0
O17	3,5	499974,9	499971,9	7	2	12	5,5	4	2	12,5	999995,3	1000000	4,7	0
O18	3,5	499974,9	499971,9	7	16	12	1	4	7,5	12,5	1000010,3	1000000	-10,3	0
O19	3,5	499974,9	499971,9	7	16	12	9,5	4	7,5	12,5	1000018,8	1000000	-18,8	0
O20	3,5	499970,9	499971,9	6	15	12	9,5	4	7,5	0	1000000,3	1000000	-0,3	0

Ezek tovább erősítik, hogy nagyon fontos az elsődleges eredmények utólagos átgondolása, átvizsgálása matematikai logika alapján. Ugyanakkor ezekből kiolvasható az is, hogy releváns és mérhető a TDK dolgozatok minősége a szóhasználati kifejezéselemek vizsgálatával, s így kialakítható egy olyan objektív rendszer, amely a TDK-tervek rangsorolása révén segíthet a diákoknak a valódi tudományos munkákat bemutató elindulásban.

Összegzés

Dolgozatomban szóbányászati módszerekkel igyekeztem objektív módon TDK-rezümék kulcsszó alapú értékelésével kialakítani egy módszert, amely segíthet a TDK témaválasztások tudományos megvalósításában. Elemzésem hosszú utak folyamán rávilágított számos hibalehetőségre, a modern eszközök korlátaira, az emberi érzékelés vakfoltjaira és a matematikai elemzések fontosságára. Végeredményben elmondható, hogy kezdeti célom sikerült, elkészült egy olyan kritériumrendszer, amely objektív különbséget tud formálni a dolgozatok között. Ráadásul, maga az ezen vezető út is bemutatta ennek kiemelt szükségét a látszólagos elemek totális megkérdőjelezésének prezentációjával. A dolgozat további hasznosítása lehet ennek alapján épülő AI elemzőrendszer megalkotása, illetve a módszer finomhangolása szótőcsalád, szinonimák és egyéb pontosító elemzésekkel, melyek jelen dolgozatom limitációi közé tartoznak, de megnyitják az utat a további tudományos pontosságot taglaló elemzések számára. A kísérlet reprodukálhatóságának és további magyarázatának elemei itt találhatóak: <https://miau.my-x.hu/miau/328/tdk/> .