

TERMÉKPÁROSÍTÁS MACHINE LEARNING ESZKÖZÖKKEL

Seprenyi Péter

Virgo Systems Kft.

seprenyi.peter@virgo.hu

Gordon Ákos

Virgo Systems Kft

gordon.akos@virgo.hu

Tarczali Tünde

Pannon Egyetem, Rendszer- és Számítástudományi Tanszék

tarczali.tunde@mik.uni-pannon.hu

Absztrakt

Az e-kereskedelmi rendszerek egyik központi kihívása a különböző forrásokból származó termékleírások automatikus összevetése és azonosítása. Jelen közlemény egy *RePrice* nevű hibrid termékpárosító rendszert mutat be, amely szöveges jellemzők, strukturált attribútumok és képi információk együttes feldolgozásával végzi az egyezésvizsgálatot. A módszert különböző e-kereskedelmi platformokon végzett kísérletek során értékeltük, amelyek azt mutatják, hogy a kombinált megközelítés jelentősen javítja a párosítási pontosságot a kizárólag szövegalapú vagy képalapú modellekhez képest. Az eredmények azt jelzik, hogy a *RePrice* rendszer hatékonyan alkalmazható dinamikus árfigyelő és versenytársелеző rendszerek fejlesztésében.

Kulcsszavak: termék párosítás, árazási intelligencia, gépi tanulás, CLIP (Contrastive Language-Image Pre-Training) modell, e-kereskedelmi adat integráció

Abstract

Product Matching with Machine Learning Tools

One of the central challenges in e-commerce is the automatic identification of whether different product listings from various sources refer to the same real-world item. This paper presents *RePrice*, a hybrid product matching system that integrates textual features, structured attributes, and visual data for comprehensive similarity analysis. The proposed method was evaluated through experiments conducted on multiple e-commerce platforms. Results demonstrate that the hybrid approach significantly outperforms models relying solely on textual or image-based information. These findings suggest that the

RePrice system can be effectively utilized in the development of dynamic price monitoring and competitor intelligence applications.

Keywords: product matching, price intelligence, machine learning, CLIP (Contrastive Language-Image Pre-Training) model, e-commerce data integration

Bevezetés

Az e-kereskedelmi árfigyelő rendszerek egyik alapvető funkciója annak meghatározása, hogy két különböző terméklistázás ugyanarra a valós termékre vonatkozik-e. Ez az egyezésfelismerés kulcsszerepet játszik a duplikált bejegyzések konszolidálásában, az árak összehasonlíthatóságában és a készletmenedzsmentben [1], így pontos megoldása alapvető fontosságú az ár-összehasonlítás, a készletkövetés és a termékszintű elemzés szempontjából is. A feladatot ugyanakkor jelentősen nehezíti az adatok heterogenitása és hiányossága: a globális azonosítók gyakran hiányoznak, az elnevezések pedig nem egységesek.

A hagyományos szövegalapú technikák, mint például a fuzzy string matching, korlátozott hatékonysággal működnek, mivel nem képesek megbízhatóan kezelni a szemantikai különbségeket. A modern, transzformer-alapú nyelvi modellek – mint a BERT – kontextusérzékeny szövegreprezentációt alkalmaznak, amely jelentősen javítja az egyezés pontosságát [2]. Kiegészítő információforrásként strukturált attribútumok (pl. márka, modell, műszaki jellemzők) is rendelkezésre állnak, bár ezek megbízhatósága és teljessége nem garantált [3].

A vizuális jellemzőkön alapuló összevetés különösen hasznos az olyan kategóriákban, ahol a termékek megjelenése alapján jól elkülöníthetők. A CLIP-modellhez hasonló, a több modalitást – képi és szöveges adatokat – egyesítő architektúrák lehetővé teszik ezen információk hatékony kihasználását [4]. A legjobb teljesítmény ugyanakkor multimodális rendszerekkel érhető el, amelyek képesek integrálni a különböző típusú adatokat, miközben tolerálják az adatminőséggel kapcsolatos problémákat [5].

Jelen cikk célja egy ilyen hibrid megközelítésű rendszer, a *RePrice* bemutatása, amelyet valós piaci környezetre optimalizáltunk.

Módszerek és eredmények

A termékazonosítás gyakorlati megvalósítása összetett feladat, amely többféle adatforrást és elemzési szempontot igényel [6]. Ebben a fejezetben részletesen bemutatjuk a *RePrice* rendszerben alkalmazott hibrid megközelítést, amely különféle egyezési technikákat integrál egyetlen, több lépcsőből álló feldolgozási láncba.

A termékazonosítási folyamat első lépése strukturált adatokra épül. Elméletileg a globális termékazonosítók – például a GTIN-, EAN- vagy UPC-kódok – egyértelmű párosítást tennének lehetővé [3], gyakorlati alkalmazhatóságuk azonban korlátozott, mivel ezek az

azonosítók a nyilvánosan elérhető termékoldalakon jellemzően hiányoznak. Alternatív megoldásként a gyártó- és modellnévmezők információtartalma hasznosítható, különösen a műszaki cikkek esetében, ahol ezek az attribútumok viszonylag gyakran elérhetők [7]. Ugyanakkor más termék kategóriákban – például a ruházati vagy lakásfelszerelési termékeknél – ezen mezők hiányossága gyakori. Az ilyen típusú strukturált mezők alapján végzett egyezésvizsgálat találati aránya általában alacsony, jellemzően 1–5% közötti értéket mutat.

A feldolgozás következő lépéseként, a rendszer skálázhatóságának növelése érdekében névalapú egyezésvizsgálatot alkalmazunk. Mivel a terméknevek minden listázásban rendelkezésre állnak, először normalizáljuk ezeket – többek között kisbetűsre alakítással és a felesleges írásjelek eltávolításával –, majd fuzzy string matching eljárásokkal – például Levenshtein-távolság vagy token-alapú metrikák segítségével – mérjük az egyezést.

Azonban a különböző weboldalakon szereplő terméknevek jelentős eltérést mutathatnak, még az azonos termékek esetén is, ezért elengedhetetlen a kategóriaspecifikus szótárak, referenciatáblák és távolságküszöbök kialakítása. E módszer önmagában körülbelül 35–40%-os egyezési arányt képes biztosítani, ami meglehetősen jó eredménynek számít ilyen adatminőség mellett.

A pontosabb egyezés elérése érdekében további strukturált attribútumokat is bevonunk az elemzésbe. Ilyen jellemzők például a szín, méret, digitális termékek esetében a memória, valamint az ár, amelyeket kiegészítő információként használunk a névalapú hasonlóság finomítására [8]. Ugyanakkor ezen jellemzők bevonása növelheti a hamis pozitív egyezések kockázatát – különösen olyan esetekben, amikor eltérő termékek hasonló ár kategóriába esnek, vagy részleges átfedést mutatnak a specifikációkban. Ennek ellenére a szöveges és attribútum-alapú információk integrálása már 45–50%-os találati arányt is eredményezhet.

A nyilvánosan elérhető webes adatok nemcsak hiányosak, hanem gyakran strukturálatlanok és zajosak is, ezért a párosítási pontosság további javítása érdekében képalapú összehasonlítást is beépítünk a rendszerbe. Ennek során minden termékhez tartozó képet letöltünk, egységes méretre és formátumra normalizálunk, majd a CLIP-modell segítségével vektortérbeli reprezentációt (beágyazást) képzünk [10]. A CLIP-modell előnye, hogy a képi és szöveges információkat közös szemantikus térben ábrázolja, így lehetővé teszi a vizuális hasonlóság kvantitatív értékelését.

A képpárok közötti hasonlóságot koszinusz-távolság alapján számítjuk, és az alábbi háromlépcsős döntési szabályrendszert alkalmazzuk:

- 0,85 feletti hasonlósági érték esetén az egyezést automatikusan elfogadjuk;
- 0,70–0,85 közötti tartományban kézi ellenőrzésre jelöljük ki;
- 0,70 alatti értékeknél az egyezést elutasítjuk.

Ez a képalapú módszer kategóriafüggetlen 54–67%-os párosítási arányt eredményezett, különösen hatékonyak bizonyulva a divatcikk és elektronikai termékek esetében.

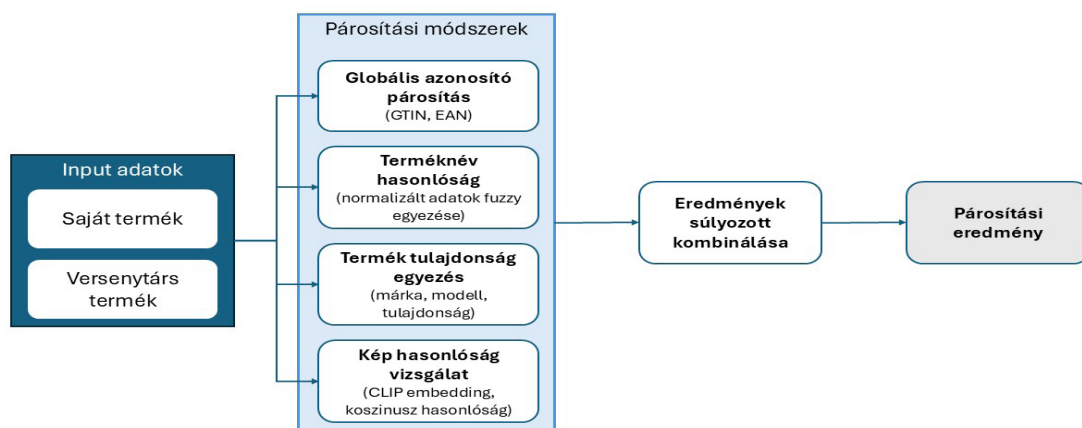
A fentiek alapján egyértelmű, hogy egyik modalitás sem képes önmagában megfelelő megbízhatóságot biztosítani. A vizsgálatok eredménye alapján megoldásunkban a végső egyezés meghatározása multimodális, súlyozott modell alapján történik. Ez a megközelítés lehetővé teszi a különböző típusú jelek mérlegelését: a determinisztikus tényezők (például a globális termékazonosítók megléte) biztos támpontot nyújtanak, míg a valószínűségi jellegű komponensek (például beágyazásokból származó hasonlósági pontszámok) a bizonytalanabb, de informatív összefüggések felismerését segítik. A kettő kombinációja így megbízhatóbb és robusztusabb döntéshozatalt tesz lehetővé [9][11].

A *RePrice* rendszer integrációs modulja ezek kombinációját alkalmazza egy előre definiált súlyozási séma szerint, amely az egyes jellemzők informatív értékén alapul. A rendszerben alkalmazott modulokat az 1. ábra reprezentálja.

Az alkalmazott súlyarányokat az alábbiak szerint állítottuk be:

- Globális azonosítók egyezése (GTIN, EAN): 0,40
- Névalapú fuzzy egyezés: 0,20
- Strukturált attribútumok (pl. márka, modell): 0,10
- CLIP-alapú vizuális hasonlóság: 0,20
- Árközeliség: 0,10

A modell együttes pontszámot képez az egyes forrásokból származó részpontszámok súlyozott összegeként, és ennek alapján végzi el az egyezés eldöntését egy küszöbérték szerint. A súlyok értékét empirikusan kalibráltuk, valós adathalmazokon végzett szisztematikus tesztek során.



1. ábra: A rendszer sematikus architektúrája

A különböző modalitások integrálásán alapuló kombinált megközelítés jelentős mértékben javította a rendszer általános teljesítményét. A modell értékelését egy 29 000 termékleírást tartalmazó tesztalmazon végeztük, amely 23 különböző e-kereskedelmi platformról származott, és olyan termékkategóriákat foglalt magában, mint az elektronikai eszközök, ruházat, gyorsan forgó fogyasztási cikkek (FMCG), valamint háztartási termékek. A vizsgálati mintában 5200 előre validált, egyező termékpár szerepelt.

A kizárólag névalapú egyezésvizsgálat alkalmazása esetén a rendszer precizitása 0,37, a recall értéke 0,49, az így számított F1-mutató pedig 0,42 volt. Ezzel szemben a teljes, multimodális megközelítés alkalmazásával a precizitás 0,86-ra, a recall 0,66-ra, az F1-érték pedig 0,75-re emelkedett. Az F1-mutató – mint a precizitás és a recall harmonikus átlaga – jól jellemzi az osztályozási rendszer általános hatékonyságát.

Az alábbi 1. táblázat néhány termékkategóriára kibontva jeleníti meg a vizsgálati eredményeket.

Termékkategória	Termékek száma	Képek száma	Név alapú egyezés			Súlyozott kombinált egyezés		
			Precizitás	Recall	F1	Precizitás	Recall	F1
Szórakoztató elektronika	2365	12342	0,41	0,49	0,45	0,83	0,63	0,72
Fashion	1681	27931	0,32	0,53	0,40	0,74	0,66	0,70
Barkács	658	5223	0,45	0,58	0,51	0,88	0,66	0,75
Háztartás	871	3776	0,42	0,47	0,44	0,87	0,71	0,78
FMCG	342	3889	0,31	0,53	0,39	0,87	0,68	0,76

1. táblázat: A RePrice rendszer által adott eredmények összefoglalása különböző termékkategóriák esetén

Az eredmények azt mutatják, hogy a különböző típusú adatok együttes használata számottevő előnyt biztosít: a rendszer képes tolerálni az egyes modalitásokban előforduló hiányosságokat vagy zajokat, miközben a rendelkezésre álló további jellemzők révén megerősíti az egyezések megbízhatóságát. A képalapú beágyazások különösen hatékonynak bizonyultak azokban a termékkategóriákban – elsősorban az elektronikai cikkeknel és divatáruknál –, ahol a vizuális megjelenés domináns szerepet játszik a termékek megkülönböztetésében.

Következtetések és további irányok

Jelen cikk az online árfigyelési rendszerekben kulcsszerepet játszó termékazonosítás problémáját vizsgálta. A cél annak meghatározása volt, hogy különböző forrásokból származó, eltérő szerkezetű és minőségű adatok alapján megállapítható-e, hogy két terméklistázás ugyanarra a valós árucikkre vonatkozik. Bár a feladat jelentős kihívást jelentett az adatok heterogenitása és hiányosságai miatt, az e-kereskedelemben alapvető fontosságú az ár-összehasonlítás, készletmonitorozás és termékszintű elemzés lehetővé tételéhez.

A vizsgálat során három fő adatmodalitás – szöveges leírások, strukturált attribútumok és képi megjelenítés – szerepét elemeztük a termékpárosítási feladatban. A szövegalapú megközelítések magas precizitást kínálnak, de érzékenyek a nem szabványos, hiányos vagy zajos elnevezésekre. Az attribútumalapú azonosítás akkor lehet hatékony, ha az adatok teljessége és konzisztenciája biztosított. A képalapú egyezés – különösen a CLIP-modellek által előállított beágyazások alkalmazásával – a vizuálisan könnyen elkülöníthető termékcsoportoknál, például a divatárúk és elektronikai eszközök esetén bizonyult kiemelkedően hatékornak.

Az említett modalitások integrálásával kialakított hibrid döntési modell egyértelműen a legjobb teljesítményt nyújtotta. Eredményeink azt mutatják, hogy a multimodális információfeldolgozás képes kiegyenlíteni az egyes adatforrások hiányosságait, és robusztus, megbízható teljesítményt biztosít gyakorlati környezetben is. A *RePrice* rendszer, amely ezen elvek mentén került megvalósításra, ipari környezetben is bizonyította alkalmazhatóságát, lehetővé téve a többplatformos termék- és ármonitorozást.

Kísérleteink megerősítik, hogy még zajos, strukturálatlan webes környezetben is megvalósítható a pontos termékazonosítás egy jól felépített, mesterséges intelligenciát alkalmazó hibrid megközelítéssel. A bemutatott rendszer megbízható alapként szolgálhat fejlett árfigyelő megoldások, ajánlórendszerek, valamint termékanalitikán alapuló döntéstámogató rendszerek számára.

Jelenlegi megoldásunkban az egyes adatmodalitásokhoz rögzített, heurisztikusan meghatározott súlyokat rendeltünk. Ugyanakkor valószínűsíthető, hogy a rendszer teljesítménye tovább javítható lenne, ha a súlyok optimalizálását felügyelt tanulás révén végeznénk. Ha rendelkezésre állna egy címkézett tanulókészlet ismert egyező termékpárokkal, akkor például logisztikus regresszióval vagy gradiens alapú boosting technikákkal meghatározhatók lennének az optimális súlyparaméterek.

Jövőbeli munkánkban olyan tanulható integrációs rétegek fejlesztését tervezzük, amelyek akár teljesen kiválthatják, vagy dinamikusan finomíthatják a jelenlegi súlybeállításokat. Ez új lehetőségeket nyithat meg az intelligens e-kereskedelmi alkalmazások – például személyre szabott árfigyelő rendszerek, termékajánlók és dinamikus kategória-rendszerezések – fejlesztése előtt.

Köszönetnyilvánítás

A kutatást az „Intelligens, gépi tanuláson alapuló termékár- és népszerűség-elemző alkalmazás tervezése és fejlesztése” című projekt támogatta, projektazonosító: 2019-1.1.1-PIA-CI-KFI-2019-00254.

Irodalomjegyzék

- [1] A. de la Comble, A. Dutt, P. Montalvo, and A. Salah, “Multi-modal attribute extraction for e-commerce,” arXiv preprint [arXiv:2203.03441](https://arxiv.org/abs/2203.03441), 2022.
- [2] S. Gheewala, S. Xu, S. Yeom, and S. Maqsood, “Exploiting deep transformer models in textual review-based recommender systems,” *Expert Systems with Applications*, vol. 235, p. 121120, 2024.
- [3] K. Gupte, L. Pang, H. Vuyyuri, and S. Pasumarty, “Multimodal product matching and category mapping: Text+Image based deep neural network,” in *Proc. 2021 IEEE Int. Conf. on Big Data (Big Data)*, Orlando, FL, USA, 2021, pp. 4500–4505.
- [4] M. Hovsepyan, “Deep learning models for product mapping,” M.S. thesis, Faculty of Mathematics and Physics, Charles Univ., Prague, Czech Republic, 2024.
- [5] M. Jia, L. Shen, L. A. Tuan, M. Chen, and J. Xu, “MuJo-SF: Multimodal joint slot filling for attribute value prediction of e-commerce commodities,” *IEEE Transactions on Multimedia*, vol. 26, pp. 10354–10366, 2024.
- [6] Y. Jiang, K. Liao, S. Lin, H. Qiao, K. Yu, C. Yang, and Y. Chen, “Self-supervised multimodal representation learning for product identification and retrieval,” in B. Luo, L. Cheng, Z. G. Wu, H. Li, and C. Li, Eds., *Neural Information Processing (CCIS, vol. 1965)*. Singapore: Springer, 2024, pp. 579–594.
- [7] S. Mallapragada, Y. Xie, V. R. Chawan, Z. Hailat, and Y. Wang, “Multi-modality transformer for e-commerce: Inferring user purchase intention to bridge the query-product gap,” in *Proc. 2024 IEEE Int. Conf. on Big Data (BigData)*, Washington, DC, USA, 2024, pp. 589–598.
- [8] A. Nawaz, “ASIF: Attention-based sentiment inquiry framework for profound product recommendations,” *Multimedia Tools and Applications*, vol. 84, pp. 9689–9712, 2025.
- [9] B. Zhang, C. Xu, H. Tang, C. Wei, and Y. Jin, “ACE-BERT: Adversarial cross-modal enhanced BERT for e-commerce retrieval,” in X. Song, R. Feng, Y. Chen, J. Li, and G. Min, Eds., *APWeb-WAIM 2023 (LNCS, vol. 14334)*. Singapore: Springer, 2024, pp. 294–311.
- [10] S. Tóth, S. Wilson, A. Tsoukara, E. Moreu, A. Masalovich, and L. Roemheld, “End-to-end multi-modal product matching in fashion e-commerce,” arXiv preprint [arXiv:2403.11593](https://arxiv.org/abs/2403.11593), 2024.
- [11] J. Pollack, H. Köpcke, and E. Rahm, “Intermediate fusion for multimodal product matching,” in *Proc. of the Workshop on AI for Fashion at AAI 2024 (CEUR Workshop Proc., vol. 3710)*, 2024, pp. 131–137.