

EGÉSZSÉGÜGYI ÁLHÍRKERESŐ ELJÁRÁSOK INFORMATIKAI MEGOLDÁSAI

COMPUTATIONAL TOOLS FOR HEALTH RELATED FAKE NEWS DETECTION

Csendes Tibor

MTA doktora, Szegedi Tudományegyetem, Informatikai Intézet, Szeged
csendes@inf.szte.hu

Berend Gábor

PhD, Szegedi Tudományegyetem, Informatikai Intézet, Szeged
berendg@inf.szte.hu

Németh T. Enikő

MTA levelező tagja, Szegedi Tudományegyetem,
Magyar Nyelvi és Irodalmi Intézet, Szeged
nemethen@hung.u-szeged.hu

ÖSSZEFOGLALÁS

Az MTA Tudomány a Magyar Nyelvért Nemzeti Program *Álhírek, áltudományos nézetek nyelvészeti azonosítása* című projektje keretében eljárásokat fejlesztünk egészségügyi álhírek automatikus felismerésére. A jelen cikkkel kettős célunk van. Egyrészt röviden bemutatjuk az álhírkereső eljárások fejlesztése terén elért első eredményeinket, azon belül is az informatikai megoldásainkat. Másrészt szeretnénk a már használható számítógépes eszközeink példáival az ilyen munkák iránt érdeklődőket megnyerni együttműködőknek.

ABSTRACT

In the frame of the Science for the Hungarian Language National Programme in the Linguistic Identification of Fake News and Pseudo-Scientific Views project we develop computational procedures to automatically recognize health care related fake news. We have dual aims with the present article. First, we want to introduce our first results in fake news recognition,

especially our computational tools. Second, we intend to attract interested colleagues for possible collaborations by the examples of our computational tools.

Kulcsszavak: automatikus áhírfelismerés, áhír, egészségügy, pragmatika, reguláris kifejezés

Keywords: automatic fake news recognition, fake news, health care, pragmatics, regular expression

Előzmények

A Szegedi Tudományegyetem Általános Nyelvészeti Tanszéke és a hozzá kötődő Pragmatika Hallgatói Kutatóműhely évek óta foglalkozott a manipuláció és az áhírek nyelvészeti eszközökkel felismerhető jeleinek, stratégiáinak a keresésével. 2022 nyarán az SZTE Informatikai Intézetében Gáspár Tamás mesterszakos hallgató elkészített egy olyan okostelefonos alkalmazást, amely az áhírek nyelvészeti tulajdonságainak az ellenőrzésére volt képes. Ugyanazon év őszén az MTA Tudomány a Magyar Nyelvért Nemzeti Programja keretében az MTA-SZTE-DE Elméleti Nyelvészeti és Informatikai Kutatócsoport kapott támogatást négy évre az *Áhírek, áltudományos nézetek nyelvészeti azonosítása* című projektje megvalósítására.

2023 novemberében zártuk az első évet, az informatikai eszközeink megjelenése kialakult, lehet őket tesztelni. Ez megfelel az ún. „zöld banán” fejlesztési elvnek, illetve a felhasználói élmény (UX) korai kialakításának. Az érdemi algoritmusok finomítása, továbbfejlesztése valószínűleg kitölti majd a projekt teljes idejét, de a most elérhető eljárásaink is jók és használhatók már.

Első eredmények

Okostelefonos alkalmazás

Továbbfejlesztettük a fentebb említett működőképes okostelefonos alkalmazást, amely képes a nyelvészcsoporthunk által összeállított jellemzők automatikus felismerésére. Az alkalmazás a háttérben futva, a képernyőképen megjelenő tetszőleges szövegen végzi a detektálást, megjelöli a gyanús részeket, rövid magyarázatot ad az okra, és összegzi a teljes látható szövegről való benyomását az áhírnyanú valószínűségének megadásával. Az alkalmazás egy online adatbázis segítségével dolgozik, amelyet a nyelvészek egy egyszerű kezelőfelületen tudnak bővíteni vagy javítani. Az összegzett értékelés egyelőre ad hoc jellegű, nincs visszamérve, illetve a valós helyzethez igazítva. Az 1. ábra segít elképzelni az alkalmazás eredményét, amelyről egy bemutató videó is elérhető az [al hirdetektor.hu](https://www.youtube.com/watch?v=ybg5NU-UwD4&t=36s) oldalról, vagy közvetlenül a

<https://www.youtube.com/watch?v=ybg5NU-UwD4&t=36s>

címről is. A későbbiekben tervezzük ennek összekötését a mesterséges intelligencia alapú álhírfelismerő programmal. Az alkalmazás maga is letölthető az alhirdetektor.hu vendégoldalról. Android 8 és későbbi rendszerekben működik.



1. Ábra. Az álhírfelismerő okostelefonos alkalmazás működés közben. A program csak a kijelzőn megjelenő szöveget használja. Ezen megjelöli a gyanúra okot adó jeleket, és a jobb felső sarokban színnel és százalékos értékkel is jelzi az álhírnyír mértékét.

Mesterséges intelligencia alapú felismerő motor

Az álhírfelismerésre egy mesterséges neuronháló alapú módszert is fejlesztettünk 5000 egészségügyi álhírral és ugyanennyi korrekt egészségügyi szöveggel való betanítással. Ez egy magyar nyelvi modell segítségével dolgozik, a 110 millió paraméteres HuBERT (Nemeskey, 2020) mesterséges neuronháló finomhangoláson átesett változata hozza meg a döntést. A HuBERT modell használatán túl, a kompaktabb HuBERTUSz modellcsalád (Ficsor, Berend, 2023) használatával is kísérleteztünk, amellyel hasonlóan jó eredményekre sikerült jutnunk. Álhírdetektáló modelljeink az ún. transzformer neurális háló architektúrát alkalmazó nyelvi modellre támaszkodnak. Ennek az a lényege, hogy a program a szöveget alkalmas rövid jelsorozatokra bontja, és az érdemi elemzés ezeken történik. A transzformer modellek az utóbbi időben sok természetesnyelv-feldolgozási (natural language processing; NLP) probléma megoldásával bizonyították sikeres alkalmazhatóságukat, mivel ez a fajta architektúra lehetővé teszi a szövegekben rejlő belső összefüggések, kontextusok és nyelvi jelenségek feltárását, amelyek rendkívül hasznosak lehetnek változatos, a szövegek feldolgozásával kapcsolatos feladat elvégzésében, így az álhírek azonosításában is.

Tesztelésünk szerint a modell kb. 90%-os pontossággal dolgozik, jelenleg egy bekezdésnyi szövegen hatékony. Ez a pontosság az átlagos emberi megítélésnél kicsit jobb az általunk vizsgált korpuszon. A pontos kiértékelést az 1. Táblázat mutatja a tanítási és tesztelési adathalmazunk alapján. A táblázatban szereplő F1 mutató a pontosság és a fedés harmonikus

közepe. A pontosság itt a rendszer által helyesen azonosított álhírek aránya az összes álhírként azonosítotthoz képest. A fedés pedig a helyesen azonosított álhírek aránya az összes álhírhez képest.

	HuBERTUSz-tiny	HuBERTUSz-small	HuBERTUSz-medium	HuBERT
Paraméterek száma	4,5 millió	29,5 millió	42 millió	110 millió
címekre	0.9399	0.9423	0.9423	0.9549
tartalomra	0.9196	0.9272	0.9322	0.9422

1. Táblázat. A kifejlesztett mesterséges neuronháló teljesítménye az F1 mutató szerint címekre, illetve tartalomra használva, 4 különböző nagyságrendű kapacitással rendelkező magyar nyelvi modell szerint.

A modellhez tartozik egy visszajelzést támogató üzenetküldő funkció is. Az ezzel szerzett tapasztalattal tudjuk javítani a program eredményességét. A program általában jól működik, akár a végleges megoldásunk is lehet.

A projektünk hátralévő három évében leginkább ennek az eszköznek a továbbfejlesztését szeretnénk végezni. A többi informatikai eszköz vagy kész van, vagy épp a mesterséges intelligencián alapuló fejlesztésekre vár. Szeretnénk tisztázni a mesterséges neuronhálónk döntése interpretációját, kipróbálni a tisztán a nyelvészeti szabályokkal elérhető osztályozást, illetve ezek hibridjét.

Az említett neurális hálót használó modell kipróbálható a <https://www.inf.u-szeged.hu/~gencsi/alhir-felismero.html> címen, illetve a kutatócsoport vendégoldalán <https://www.alhirdetektor.hu/Links>. A működését a 2. ábra illusztrálja.

Egészségügyi álhírfelismerő

+Használati útmutató **Elo olvasom**

+Egészségügyi álhírfelismerő alkalmazás

Tudnivalók:

- A felismerő magyar nyelvű egészségügyi szövegen működik.
- A rendszer 20 és 500 szó közötti hosszúságú szövegekre ad megbízható eredményt.
- A hosszabb szövegeknek csak az elejét fogja feldolgozni a rendszer.
- Az ellenőrzendő szövegen a felismerő minimális technikai szerkesztést hajt végre.
- A rendelkezésre álló felismerők közül a rendszer véletlenül választ ki egyet.
- Ha a felismerő által adott megállapítás az Ön megítélése szerint nem megfelelő, akkor kérjük segítse a munkánkat azzal, hogy visszajelzést küld.

Egy saját fejlesztésű termék sejtjeink védelmének érdekében. Építsd be mindennapjaidba ezt a 18-féle sejtvédelmet és regenerációt támogató antioxidáns hatású terméket. A dobozban található vitamin, ásványi anyag és növényi hatóanyag formulákat praktikus, napi pakkok formájában szedheted.

Felismerés

Eredmény: **Alhír**

+Feldolgozott szöveg

↳ Visszajelzés küldés a felismerőről

© 2023 Copyright: Mihály Gencsi
Látogatók száma: 7

2. Ábra. Az egészségügyi álhírfelismerő mesterséges intelligencia alapú alkalmazás működés közben. A program a kijelölt helyre másolt szöveget vizsgálja. Pár másodperc alatt

döntésre jut, és piros vagy zöld színnel kiemelten mutatja az eredményt, de egy kék színű sáv is jelzi az álhírgyanú százalékos mértékét.

Kereső robot

Dolgozunk egy automatikus, egészségügyi álhíreket kereső roboton is. Ez már autonóm módon, emberi beavatkozást nem igénylően működőképes. Az interneten kulcsszavak segítségével keres egészségügyi szövegeket, és az előző pontban leírt algoritmussal álhíreket tud azonosítani. Ilyen eszköz fejlesztése felmerült a Nemzeti Népegészségügyi és Gyógyszerészeti Központtal (NNGYK, korábbi nevén Országos Gyógyszerészeti- és Élelmiszeregészségügyi Intézet, OGYÉI) tervezett együttműködésben is. A hatóság ugyan továbbra is érdekelt az együttműködésben, de az átszervezés miatt türelmünket kérték a részletek megbeszélése céljából. A keresés hatékonysága javítása végett megkerestük az Országos Széchenyi Könyvtárt (OSZK), és az SZTE-vel létrejött egy olyan együttműködési szerződés, amelynek a keretében az OSZK és az SZTE a webes aratás, illetve az ehhez kapcsolódó mesterséges intelligencia alapú alkalmazások kifejlesztésére vonatkozó szándéknyilatkozatot fogalmaztak meg. A program rendben működik, a tesztelések a várakozásoknak megfelelő eredményeket adtak.

Keresőprogramba beépülő bővítmény

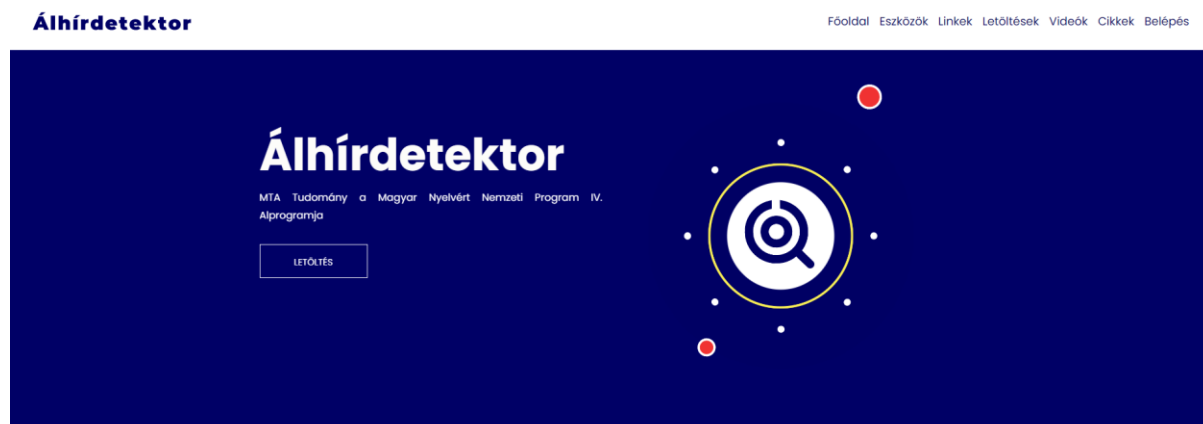
Készítettünk egy keresőprogramba beépülő álhírdetektáló alkalmazást is. Ennek továbbfejlesztésén még dolgozunk, de az első működőképes változatot bemutató videó elérhető a <https://1drv.ms/v/s!AsNC4EZvD52RgUW3j81BxBsXoZHy> címen, illetve a <https://www.alhirdetektor.hu/Videos> vendégoldalról. A 3. ábra illusztrálja a működését. Lényegében hasonló a tudása, mint az okostelefonos alkalmazásunké. Ennek a megoldásnak a kidolgozottsága egyelőre a legkisebb.



3. Ábra. A keresőprogramba beépülő bővítmény alkalmazás működés közben. Lényegében az okostelefonos alkalmazáshoz hasonló a döntése, de a teljes, az oldalon található szöveg

alapján dönt. Az osztályozás összegzett eredményét színnel is jelzi, valamint utal a gyanús részletek problémáira.

Folyik egy egységes vendégoldal fejlesztése is. Ezen a kutatócsoport eredményeit tesszük közzé hírekkel, letöltési lehetőséggel és bemutató anyagokkal együtt. Ennek fejlesztés alatt lévő változata már elérhető a <http://www.alhirdetektor.hu/> címen, a szerkeszthető hálózati környezet beüzemelése folyik még. A főoldalt mutatja a 4. ábra.



AZ ÁLHÍREKET KISZŰRŐ APPLIKÁCIÓ MÖGÖTT IS GIGANTIKUS

ALAPKUTATÁS ÁLL

Létrehozta: Vörös Richárd Dátum: 2024-01-30

Közel egy évvel ezelőtt indult el egy kutatói program, amiben közösen dolgoznak nyelvészek és informatikusok, időnként gyógyszerészek tanácsaira támaszkodva. A bölcészek a kutatásaik eredményeit megosztják az informatikusokkal, akik ezeket felhasználva fejlesztenek álhírfelismerő eszközöket, illetve „tanítják” a mesterséges intelligenciát. A végeredmény: megbízható appok, amikkel biztonságosan kiszűrhetjük az álhíreket. A program, aminek az első fázisa

4. Ábra. Az MTA-SZTE-DE Elméleti Nyelvészeti és Informatikai Kutatócsoport álhírdetektálási munkái gyűjtőoldala.

Terveink

Együttműködési megállapodást kötöttünk a PTE gyógyszerész kutatóival a gépi tanulás alkalmazása az online gyógyszerforgalmazói felületek webes tartalmának automatikus kategorizálását célzó kutatás közös megvalósítására. Az érdemi munka nemrég kezdődött meg. Nagy nyelvi modellen alapuló, többnyelvű mesterséges neuronhálót fejlesztünk. Mivel a probléma megoldása az Európai Unió hatósági feladatai között szerepel, ezért keressük majd a lehetőségét, hogy a megfelelő hatóság munkáját segítsük.

Jól kiegészítjük egymást a HDMO projekttel: ők többek között tényellenőrzést végeznek és módszertani ajánlásokat tesznek (URL2). A mi, gyors becslést kínáló eszközeink hasznosak lehetnek számukra az előszűrésben, illetve a megkeresések egy részének kezelésére. A meglévő együttműködési szerződés alapján kölcsönös előnyök kihasználására törekszünk.

Hasonló közös munkára számítunk a Nemzeti Népegészségügyi és Gyógyszerészeti Központtal (NNGYK, korábbi nevén Országos Gyógyszerészeti- és Élelmiszerégszségügyi Intézet, OGYÉI). Az első megbeszélésekre már sor került, de az együttműködési szerződés megkötése előtt vagyunk még. Várhatóan az interneten elérhető egészségügyi tartalmú szövegek gyors

előszűrése, és keresőrobot segítségével való korai figyelmeztetés lesz majd a cél. De felmerült a határozatok összeállításában való mesterséges intelligencia alapú algoritmikus előkészítés is.

Keressük a külföldi kutatócsoportokkal való együttműködést is, akiket hasonló kérdések foglalkoztatnak. Felvettük a kapcsolatot például a FakeSpeak projekttel (URL3). A külföldi kapcsolatok felvetik a többnyelvű rendszerek fejlesztési igényét. Aktívan szerepelünk nemzetközi konferenciákon, hogy az elért eredményeink ismertetésével bővítsük az együttműködők körét.

Köszönetnyilvánítás

Köszönjük az MTA támogatását a Tudomány a Magyar Nyelvért Nemzeti Program keretében, valamint a Szegedi Tudományegyetem és a Debreceni Egyetem segítségét a projekt végrehajtása helyi feltételeinek a biztosításában. Hálásak vagyunk tanácsadóinknak, akik rendszeresen támogatják a munkánkat: Bari Ferencnek, Boldogkői Zsoltnak, Csupor Dezsőnek és Halasi Attilának, valamint az informatikai megoldások kidolgozóinak: Ficsor Tamásnak, Gencsi Mihálynak és Vörös Richárdnak. Sok segítséget kaptunk az SZTE Innovációs Igazgatóságától.

IRODALOM

URL1: www.alhirdetektor.hu

URL2: www.detekto.hu, <https://www.hdmo.eu/section/about-us/who-we-are/?lang=hu>

URL3: <https://www.hf.uio.no/ilos/english/research/projects/fakespeak/index.html>

Berend Gábor, Csendes Tibor, Ficsor Tamás, Gencsi Mihály, Szécsényi Tibor, Vörös Richárd és Németh T. Enikő: Álhírek nyelvi modellek segítségével történő felismerése. *Nyelv és Tudomány*, nyest.hu, közlésre benyújtva, 2024.

Ficsor, Tamás, and Gábor Berend. "HuBERTUSz: Alacsony paraméterszámú transzformer modellek létrehozása és kiértékelése magyar nyelvre." XIX. Magyar Számítógépes Nyelvészeti Konferencia, MSZNY-2023. Szegedi Tudományegyetem, Szeged, (2023): 417-432.

Nemeskey, Dávid Márk. "Natural language processing methods for language modeling." PhD Értekezés, ELTE, Budapest (2020).