

Záverečná karta projektu

Názov projektu Evidenčné číslo projektu **SK-IL-RD-18-0004**

Detekcia falošných informácií a dezinformácií v doméne zdravotníctva

Zodpovedný riešiteľ **doc. Ing. Marián Šimko, PhD.**

Príjemca **Slovenská technická univerzita v Bratislave - Fakulta informatiky a informačných technológií**

Názov pracoviska, na ktorom bol projekt riešený

Slovenská technická univerzita v Bratislave, Fakulta informatiky a informačných technológií

Názov a štát zahraničného pracoviska, ktoré spolupracovalo pri riešení

Bar-Ilan University, Ramat-Gan, 5290002 Israel

Udelené patenty/podané patentové prihlášky, vynálezy alebo úžitkové vzory, ktoré sú výsledkami projektu

Výsledkami projektu neboli žiadne patenty, alebo podané patentové prihlášky, vynálezy alebo úžitkové vzory.

Najvýznamnejšie publikácie (knihy, články, prednášky, správy a pod.) zhrňujúce výsledky projektu – uveďte aj publikácie prijaté do tlače

AFC05 PIKULIAK, Matúš - ŠIMKO, Marián. Combining cross-lingual and cross-task supervision for zero-shot learning. In Text, Speech, and Dialogue : proceedings of 23rd International Conference, TSD 2020, Brno, Czech Republic, September 8–11, 2020. 1. vyd. Cham : Springer Nature, 2020, S. 162-170. ISSN 0302-9743 (print). ISBN 978-3-030-58322-4 (print). V databáze: DOI: 10.1007/978-3-030-58323-1_17 ; SCOPUS: 2-s2.0-85091185995.

AFC06 PIKULIAK, Matúš - ŠIMKO, Marián. Exploring parameter sharing techniques for cross-lingual and cross-task supervision. In Statistical Language and Speech Processing : 8th International Conference on Statistical Language and Speech Processing (SLSP 2020), Cardiff, 14 -16 October 2020. Cham : Springer Nature Switzerland, 2020, S. 97-108. ISSN 0302-9743. ISBN 978-3-030-59429-9 (2019: 0.427 - SJR, Q2 - SJR Best Q). V databáze: SCOPUS: 2-s2.0-85092138617 ; DOI: 10.1007/978-3-030-59430-5_8.

AFC07 ŠTEFANCOVÁ, Elena - SRBA, Ivan. POI Recommendation Based on Locality-Specific Seasonality and Long-Term Trends. In SOFSEM 2020: Theory and Practice of Computer Science, 46th International Conference on Current Trends in Theory and Practice of Informatics, SOFSEM 2020, Limassol, Cyprus, January 20–24, 2020, Proceedings. 1. vyd. Cham : Springer, 2020, S. 338-349. ISBN 978-3-030-38918-5. V databáze: SCOPUS: 2-s2.0-85079104182.

AFC08 VASKO, Dominik - PECÁR, Samuel - ŠIMKO, Marián. Automatic Text Generation in Slovak Language. In SOFSEM 2020: Theory and Practice of Computer Science, 46th International Conference on Current Trends in Theory and Practice of Informatics, SOFSEM

Uplatnenie výsledkov projektu

Uplatnenie výsledkov projektu vidíme v doméne zdravotníctva v oblasti narábania s informáciami a vyvracania falošných informácií a dezinformácií. Výsledky projektu prispeli k vývoju Monant platformy pre monitoring, detekciu a zmierňovanie dopadov antisociálneho správania.

Vzhľadom na odchod väčšiny riešiteľov projektu, vrátane zodpovedného riešiteľa doc. Ing. Mariána Šimka, PhD. z pracoviska, berúc do úvahy neochotu spolupracovať pri tvorbe záverečnej správy až na jednu riešiteľku, nie je v súčasnosti možné kvalifikovane zodpovedať v plnej šírke otázku miery uplatniteľnosti výsledkov projektu vrátane s tým spojeného naplnenia cieľov. Vzhľadom na čerpanie finančných prostriedkov predpokladáme, že ich výška plne zodpovedá stanoveným cieľom projektu. Považujeme za potrebné poukázať tiež na budúce napĺňanie monitorovacích údajov vyplývajúcich z projektového zámeru.

Súhrn výsledkov riešenia projektu a naplnenia cieľov projektu v slovenskom jazyku (max. 20 riadkov)

V sledovanom období bola vyplatená časť personálnych výdavkov na riešiteľov projektu a prebehla prezentácia výstupov projektu na 2 konferenciách SOFTWARE SEMINAR 2020 (46th International Conference on Current Trends in Theory and Practice of Computer Science) a DIONE 2020 (EAI International Conference on Data and Information in Online Environments). Konferenčné príspevky sú súčasťou zborníkov.

Predpokladáme, že riešiteľský kolektív sa riadil harmonogramu daného projektovým zámerom:

V marci 2020 skončila druhá fáza projektu, v rámci ktorej bolo plánované mať dokončenú väčšinu modelov a detekčných metód, a začala posledná fáza, ktorá bola zameraná na dokončenie sprístupnenie výsledkov metód, ich vysvetľovanie (prezentovanie vysvetlení používateľom) a zapojenie expertov do procesu odpovedania na otázky ohľadom medicínskych dezinformácií. Prebehli 3 projektové stretnutia riešiteľov projektu na jar 2020. Harmonogram projektu 2020:

Zväčšovanie anotovanej dátovej sady medicínskych informácií - Predpokladáme, že v rámci platformy Monant bolo pokračované v zbere medicínskych novinových článkov, blogov ako aj článkov verifikujúcich medicínske tvrdenia.

Klasifikácia stanoviska (postoja) článku k danému tvrdeniu - Predpokladáme, že boli zrealizované obe navrhnuté metódy strojového učenia. Prvá z nich využívala iba malý počet viet najpodobnejších s daným tvrdením, a ich okolie. Druhá všetky vety, pričom bol využitý mechanizmus pozornosti (angl. attention mechanism).

Analýza kredibility vybraných medicínskych zdrojov odbornými hodnotiteľmi - Heuristická evaluácia odborníkmi nebola realizovaná. Kredibilita medicínskych zdrojov sa rieši v rámci bakalárskych prac analýzou existujúcich metrik a modelov, kde je skúmaná kredibilita z pohľadu analýzy zdrojov medicínskych dokumentov.

Vývoj a nasadenie webového portálu FireAnt - K nasadeniu portálu neprišlo, „chat-botí“ neboli zapojení.

Metóda skorej detekcie falošných správ s veľkým dopadom - Predpokladáme, že v prípade detekcie medicínskych falošných správ boli preskúmané možnosti kombinácie viacerých modelov strojového učenia (angl. ensemble), ktorá boli použité pre spojenie predikcie dopadu a detekcie falošných správ.

Anotačná aplikácia pre dátovú sadu medicínskych dezinformácií - Anotačná sada bola vytvorená <https://annotation-app.monant.fiit.stuba.sk/login?next=%2F>.

Vysvetľovanie detekčných a klasifikačných metód - Predpokladáme, že finálna verifikácia navrhnutých metód a následná prezentácia vhodnej argumentácie pre expertov a dátových vedcov, ako aj laických používateľov sa stihli uskutočniť len čiastočne, formou účasti na konferenciách.

Súhrn výsledkov riešenia projektu a naplnenia cieľov projektu v anglickom jazyku (max. 20 riadkov)

A part of personnel expenses was paid during the monitored period and the project outputs were presented at 2 conferences SOFTware SEMinar 2020 (46th International Conference on Current Trends in Theory and Practice of Computer Science) and DIONE 2020 (EAI International Conference on Data and Information in Online Environments).

We assume, (since most of the project team members have left the project institution before finalization of the project), the research team has followed the schedule as presented by the project plan:

The second phase of the project has been completed in March 2020, when most of the models and detection methods were to be finalized. Afterwards, the last project phase started, when the method results were supposed to be available and explained to users, also, experts were to be involved in the process of answering questions regarding the medical misinformation. 3 project meetings were organized in spring of 2020.

Project schedule 2020:

Extension of an annotated medical information dataset - We assume that the collection of medical newspaper articles, blogs as well as articles verifying medical claims continued within the Monant platform.

Classification of the article position on a given statement - We assume both proposed machine learning methods have been implemented. The first one used only a small number of sentences, which were the most similar to the given statement, and its surroundings. The second has used all sentences, using the "attention mechanism".

Credibility analysis of selected medical sources by professional evaluators - Heuristic evaluation by experts was not performed. The credibility of medical resources has been dealt with within the student's theses, where existing metrics and models were analysed. Credibility has been examined from the point of view of the medical document sources analysis.

Development and deployment of the FireAnt web portal - The portal was not deployed, "chat-bots" were not connected.

An early detection method of high-impact fake messages - We assume, in the case of the medical false information detection, a variety of possibilities were explored by combining several machine learning models (ensemble). These models were to combine an impact prediction with false information detection.

Annotation application for a dataset of medical misinformation - has been created:

<https://annotation-app.monant.fiit.stuba.sk/login?next=%2F>.

Detection and classification methods explanation - We assume, the final verification of the proposed methods and the subsequent presentation of appropriate argumentation for experts and data scientists, as well as normal users, were fulfilled only partially through the conference participation.