

Decodificarea limbajului dezinformării: O analiză comparativă a știrilor false și reale COVID-19

Andrei Luchici

Associate Professor, Department of Computer Science for Business Management

Director, Center for Research in Artificial Intelligence (CRAI)

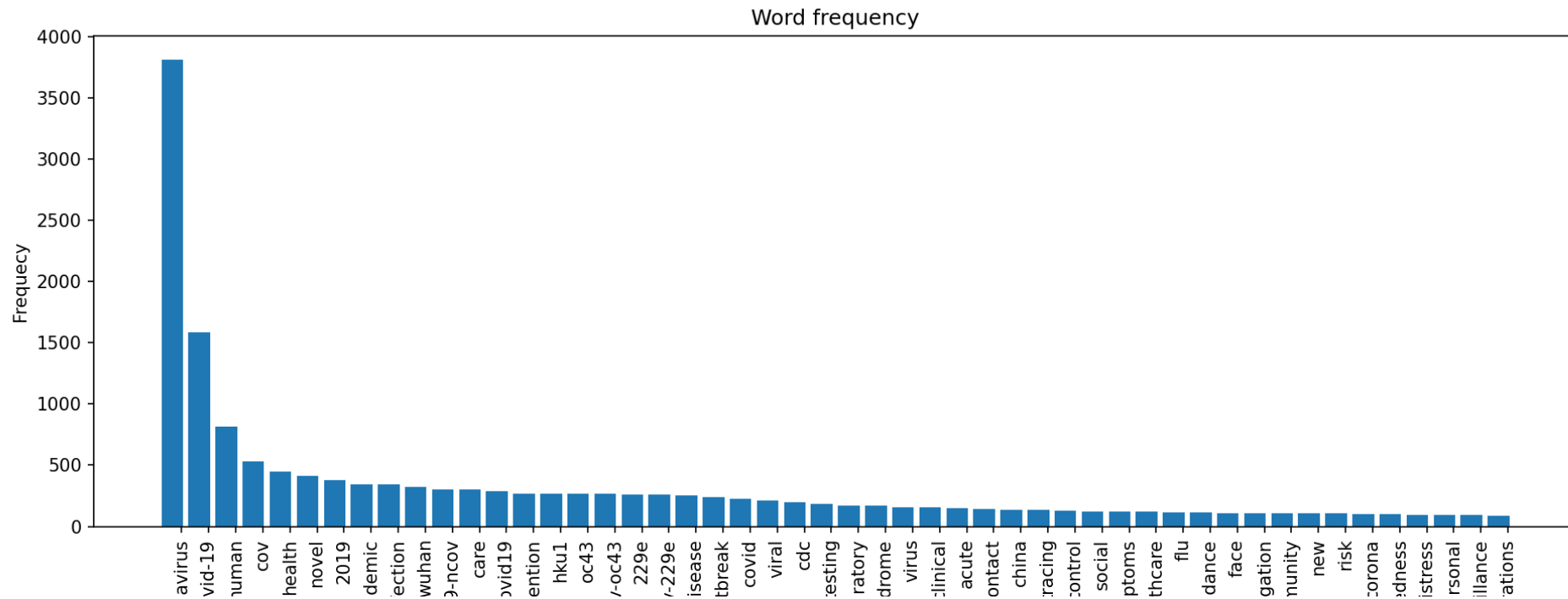
Romanian-American University

24 October 2023

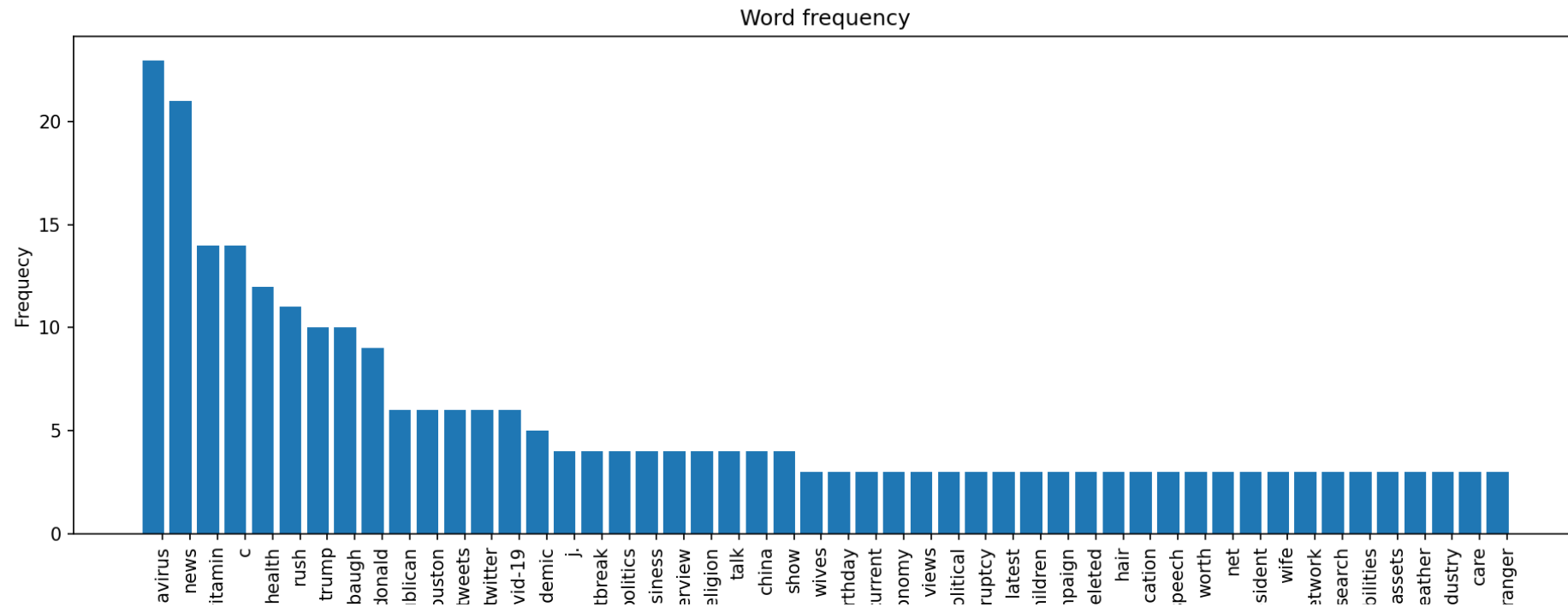
Informații preliminare

- Știri despre COVID-19, adnotate de experți
- Datele inițiale: 894 știri false; 3610 știri reale
- După filtrarea și curățarea datelor: 378 știri false; 3176 știri reale
- O parte din experții care au adnotat datele:
 - healthfeedback.org
 - factcheck.afp.com
 - mythdetector.ge
 - leadstories.com

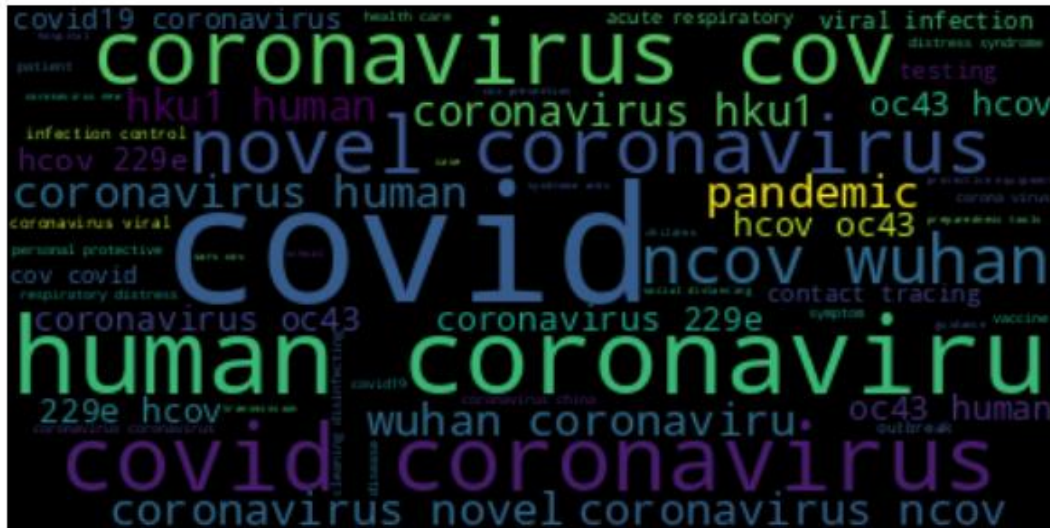
Cuvinte cheie pentru știrile reale



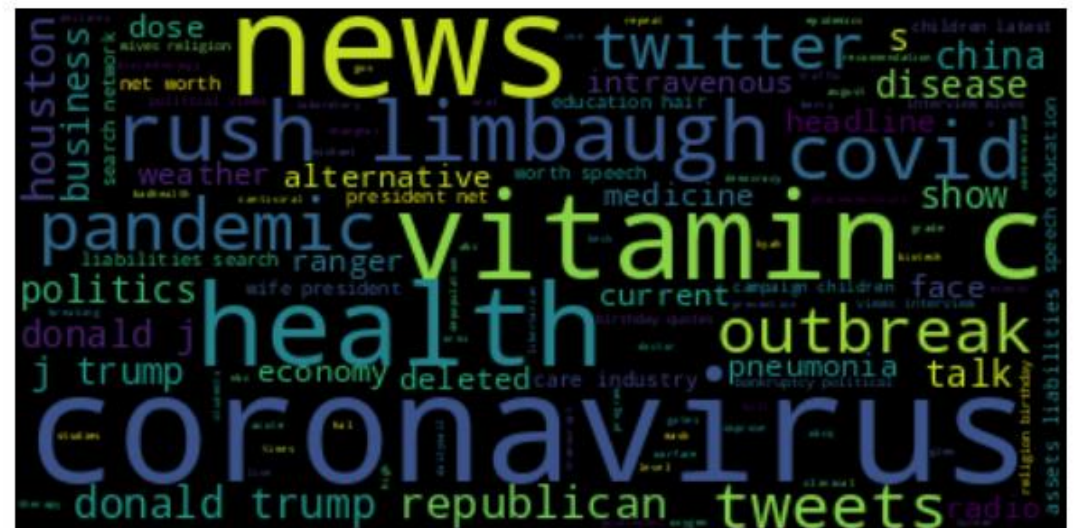
Cuvinte cheie pentru știrile false



Știrile false raportează mai mult opiniile persoanelor publice

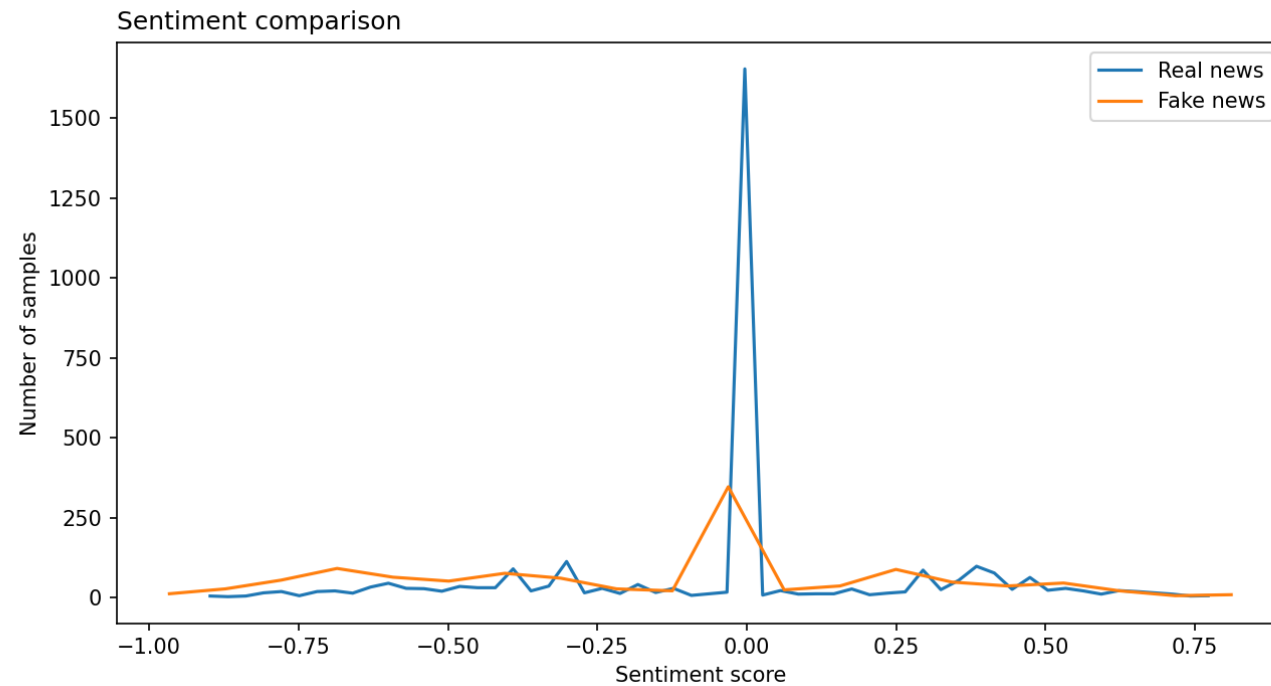


Știri reale

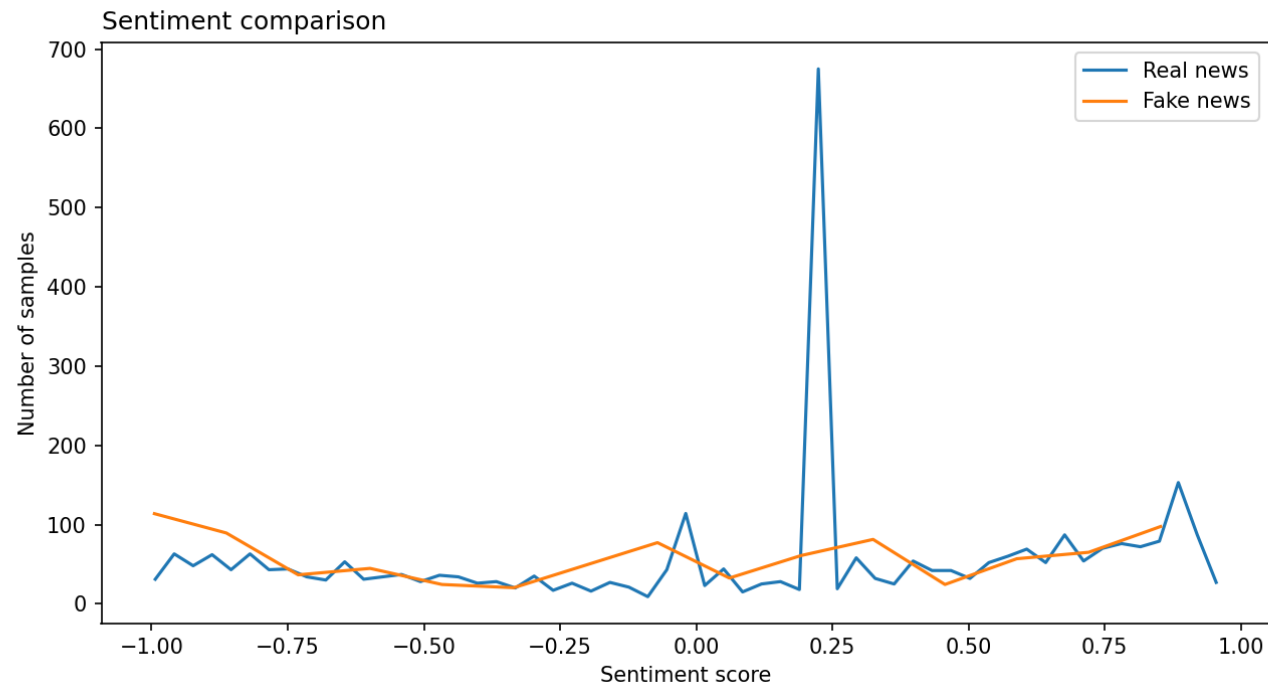


Știri false

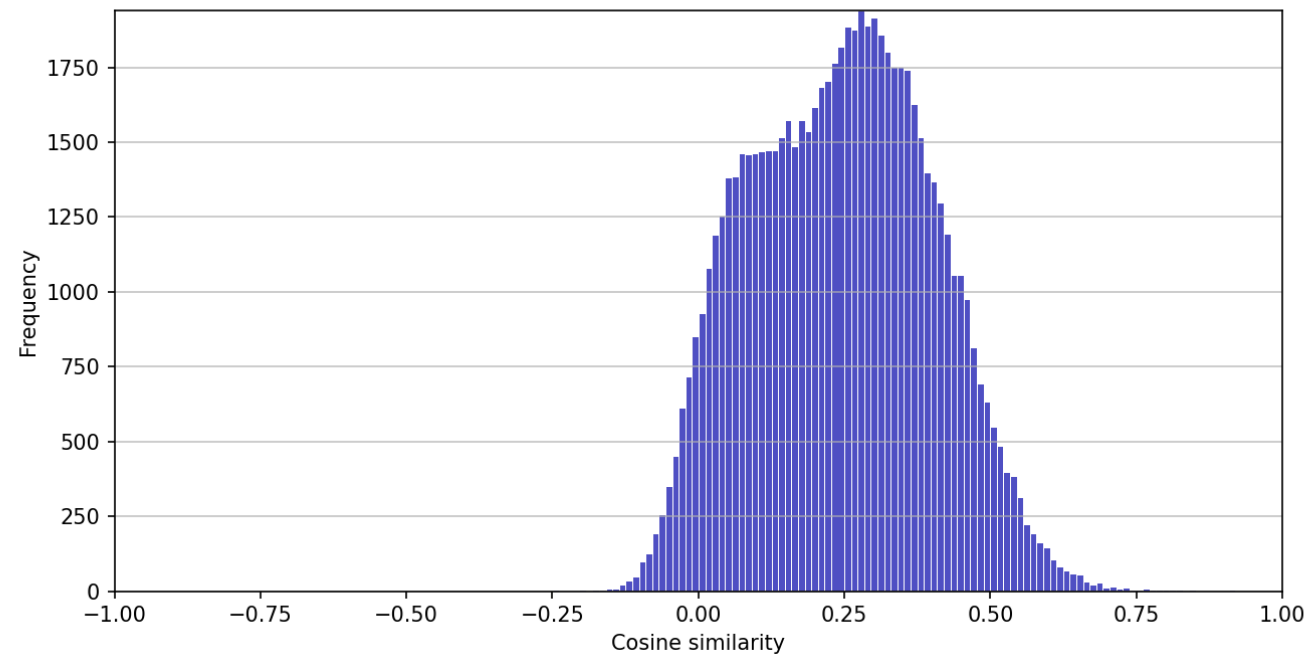
Titlurile știrilor reale sunt preponderent neutre



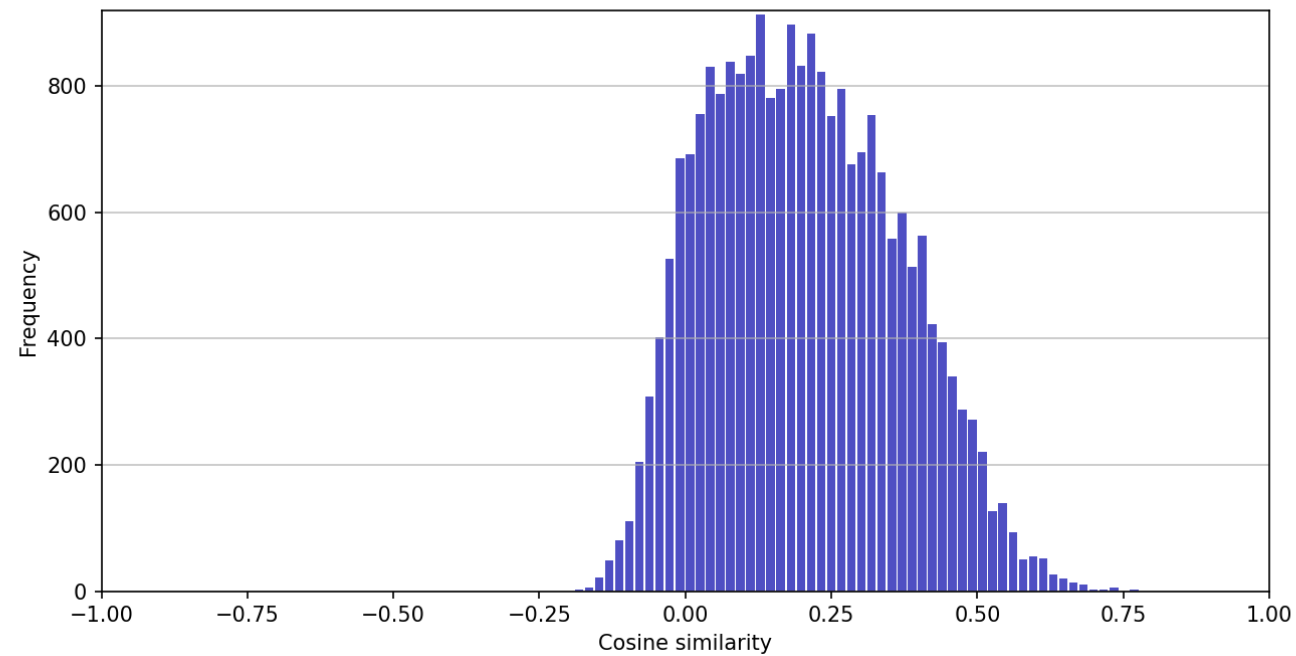
Conținutul știrilor reale este relativ pozitiv



Titlurile știrilor false încearcă să imite limbajul folosit în știrile reale



Știrile false încearcă să imite limbajul folosit în știrile reale



Concluzii și întrebări

- Autorii știrilor false se folosesc preponderant de “povești personale” nu fapte
- Sentimentul știrilor false este mai negativ decât cel al știrilor reale
- O analiză statistică a limbajului folosit nu este suficientă pentru a determina dacă o știre este reală sau falsă
- O analiză a temelor prezentate poate contribui mai mult pentru a determina valoarea de adevăr a unei știri?