



POLITEHNICA  
Bucharest



Faculty of  
Automatic Control  
and Computers



Computer  
Science and  
Engineering  
Department

# SCAN-NEWS: Sistem inteligent pentru detectia și atenuarea răspândirii dezinformării și știrilor false în rețelele sociale

Director proiect: Elena Simona Apostol

Membru: Alexandru Predescu



*Sesiunea Științifică AOȘR – 05 Iulie 2024*



# Scurtă prezentare director proiect

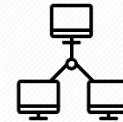


Conf.Dr.Ing. Elena-Simona Apostol  
[elena.apostol@upb.ro](mailto:elena.apostol@upb.ro)

- Conferențiar la departamentul de Calculatoare, Facultatea de Automatică și Calculatoare, POLITEHNICA București
- Membru al echipei de cercetare **MobyLab** (distributed and pervasive systems lab)
- Subiecte de cercetare:



Machine Learning



Distributed Computing



Big Data Analytics



Time Series Analysis



Data Management



Internet of Things



# Contextul proiectului și importanța temei

- Tendințele actuale și avantajele oferite de tehnologiile digitale au exacerbat practicile de răspândire a știrilor false
- Consecințe grave asupra comunităților:



**Alegeri Prezidențiale**



**Atacuri teroriste**



**Război Ucraina**



**COVID-19**



# Obiectivul Principal

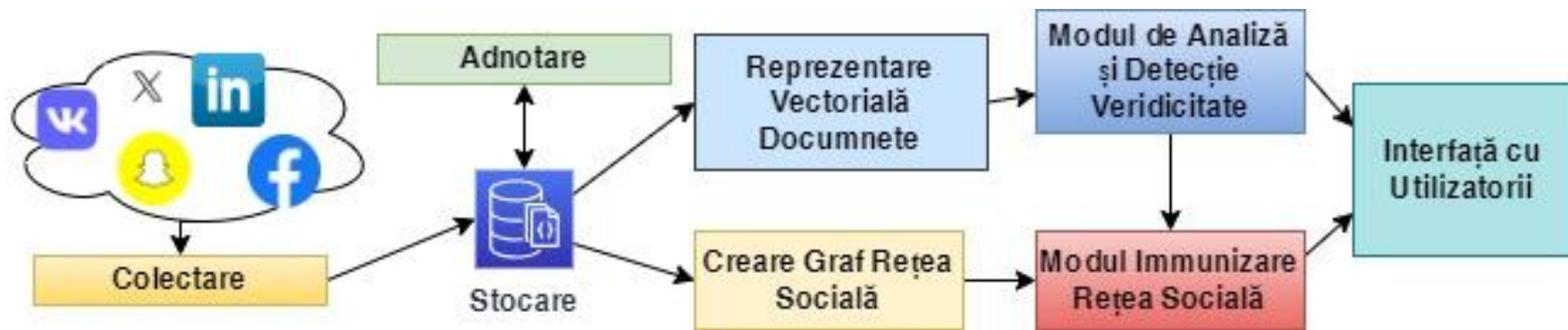
Cercetare și dezvoltarea de:

- noi **modele de detecție** de tip ansamblu care să considere în timpul antrenării pe lângă contextul cuvintelor extras cu transformare și informații despre polaritatea și tematica știrilor cât și graful de difuzie al știrii în rețeaua socială
- noi **strategii de imunizare** a rețelelor sociale care au ca obiectiv principal oprirea răspândirii online a informațiilor false





# Arhitectura generală SCAN-NEWS





# Activități de cercetare

- **Analiza modelelor existente** pentru detecția veridicității știrilor pe platformele sociale
- Începerea activităților de **colectare, stocare și adnotare** a datelor
- **Modelarea vectorială** a documentelor
- **Antrenare a modelelor** de tip ansamblu pentru detectarea veridicității știrilor publicate pe platformele sociale
- Antrenarea unor modele bazate pe **vectorizări de tip tranformer și vectorizări ale platformei sociale** modelate ca un graf pentru detectarea știrilor false
- Propunerea mai multor **modele de învățare automată** bazată pe rețele neuronale adânci
- Implementarea unui nou algoritm de **imunizare a rețelei** bazat pe **detecția comunităților** în mediile sociale





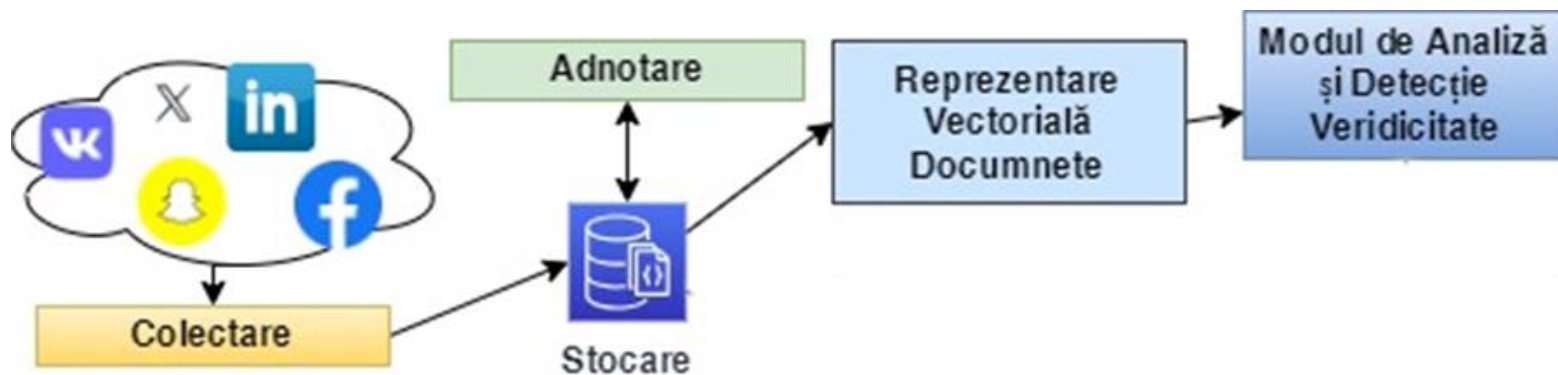
# Diseminare: Publicații raportate în cadrul proiectului

- **Elena-Simona Apostol**, Özgür Coban, Ciprian-Octavian Truică. *CONTAIN: A community-based algorithm for network immunization*. Engineering Science and Technology, an International Journal. Elsevier. 55:1-10(101728), ISSN 2215-0986, Iulie 2024. DOI: [10.1016/j.jestch.2024.101728](https://doi.org/10.1016/j.jestch.2024.101728) (Jurnal cotatec ISI Q1, F.I. = 5.1 conform JCR 2023)
- **Elena-Simona Apostol**, Adrian-Cosmin Cojocaru, Ciprian-Octavian Truică. *Large-Scale Graphs Community Detection using Spark GraphFrames*. The 23rd International Symposium on Parallel and Distributed Computing (ISPDC 2024). Iulie 2024 (Conferință cotatec C în CORE)
- Alexandru Petrescu, Ciprian-Octavian Truică, **Elena-Simona Apostol**. *Language-based Mixture of Transformers for EXIST2024*. Conference and Labs of the Evaluation Forum (CLEF2024). Publicată în Lecture Notes in Computer Science (LNCS). Septembrie 2024





# Modul de Analiză și Detecție Veridicitate: Arhitectură







## Sub-Modulul de Preprocesare și Reprezentare Vectorială

- Curățarea textului de: semne de punctuație, URL-uri, Emoji, cuvinte oprite (Punctuation Marks, URLs, emoji, cuvinte frecvente (Stop Words))
- Lematizare
- Padding /Truncare
  
- Modelarea vectorială a documentelor (Embedding): BERT, DeBERTa, RoBERTa, XML-RoBERTa



## Detecția – soluții bazate pe BiLSTM, BiGRU, CNN & LLMs

Propunerea unei noi arhitecturi Deep Learning bazate pe:

- Bidirectional Long Short-Term Memory,
- Convolutional Neural Network,
- Gated Recurrent Units și
- LLM-uri



# Experimente și rezultate

Statistic	Twitter15	Twitter16
source tweets	1490	818
false	370	205
non-rumor	374	205
true	372	207
unverified	374	201
users	276,663	173,847
tweets	331,612	204,820

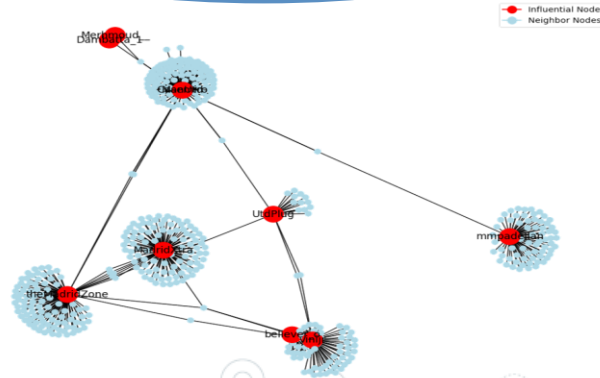
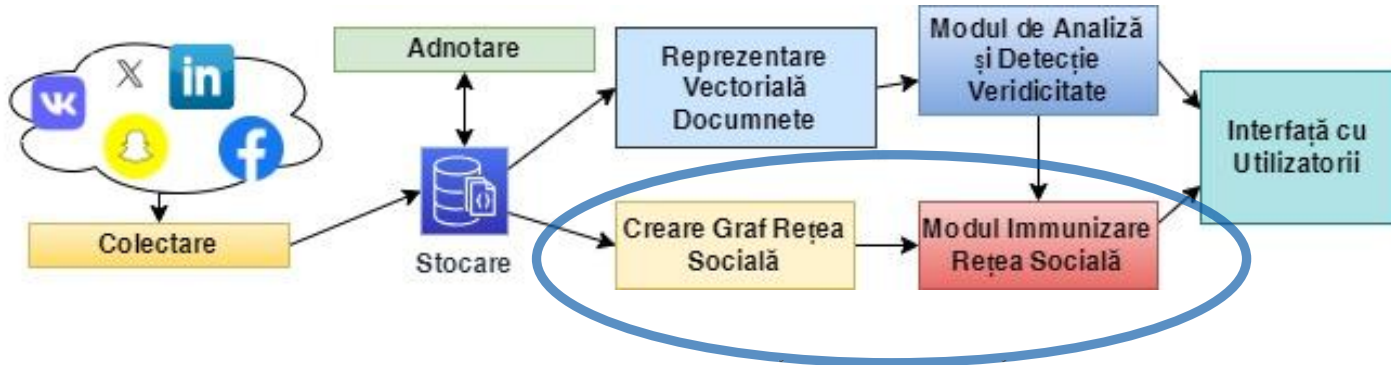
## EXIST 2024 Dataset

Engleză	4,727
Spaniolă	5,307

ModelName	Epoch	F1	Loss	Train(s)	Eval(s)
twitter-roberta	4	0.7789	0.4863	1463	5
twitter-xlm-roberta-base-sentiment-multilingual	3	0.7665	0.4670	7039	159
twitter-xlm-roberta-base-sentiment	3	0.7482	0.4902	6373	146
bert-toxic-comment-classification	4	0.7463	0.5211	6969	117
distilbert-uncased-english	4	0.7406	0.5348	2918	3
distilbert-base-multilingual-cased-sentiments	4	0.7379	0.5123	4407	3
MiniLM-L12-H384	5	0.7338	0.5059	296	3
xlm-roberta	4	0.7327	0.5520	1834	9
roberta-hate-speech-dynabench-r4	3	0.7126	0.5220	6080	154



# Modulul de Imunizare Rețea Socială





# CONTAIN: A community-based algorithm for network immunization



---

**Algorithm 1: CONTAIN: Community-based Algorithm for Network Immunization**

---

**Input** :  $G$  - the graph for immunizing  
 $S$  - the list of harmful nodes, i.e., seeds  
 $k$  - the immunization budget  
 $\gamma$  - the resolution for communities' size  
 $\Delta\gamma$  - the step value to increase the resolution

**Output:**  $r$  - the ranked list of harmful communities

---

```
1  $N(s) \leftarrow \text{ExtractNeighbours}(s \in S)$  // extract the subgraph formed by each seed  $s$ 
2  $G' \leftarrow \text{Compose}(N(s))$  // create the composed subgraph  $G'$ 
3 Loop
4  $C \leftarrow \text{Louvain}(G, \gamma)$  // extract communities using Louvain
5 if  $|C \cap G'| \geq k$  // verify if the budget  $k$  constraint is meet
6 then
7  $r \leftarrow \text{Ranked}(C \cap G')$  // rank harmful communities
8 return  $r$  // return the ranked list
9 else
10  $\gamma \leftarrow \gamma + \Delta\gamma$  // increase the resolution
```

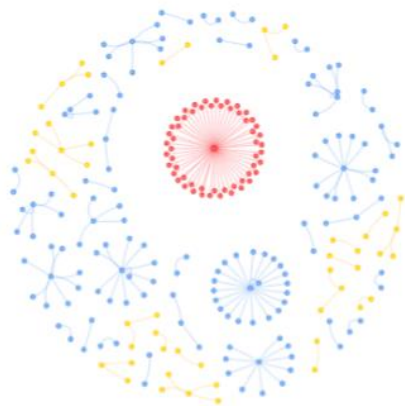
---



# Detecția de comunități



- Propus un framework de detectare a comunității în medii distribuite dezvoltat folosind GraphFrames, Apache Hadoop și Apache Spark
- Algoritmi implementați: **Louvain**, Fast Greedy, K-Cliques

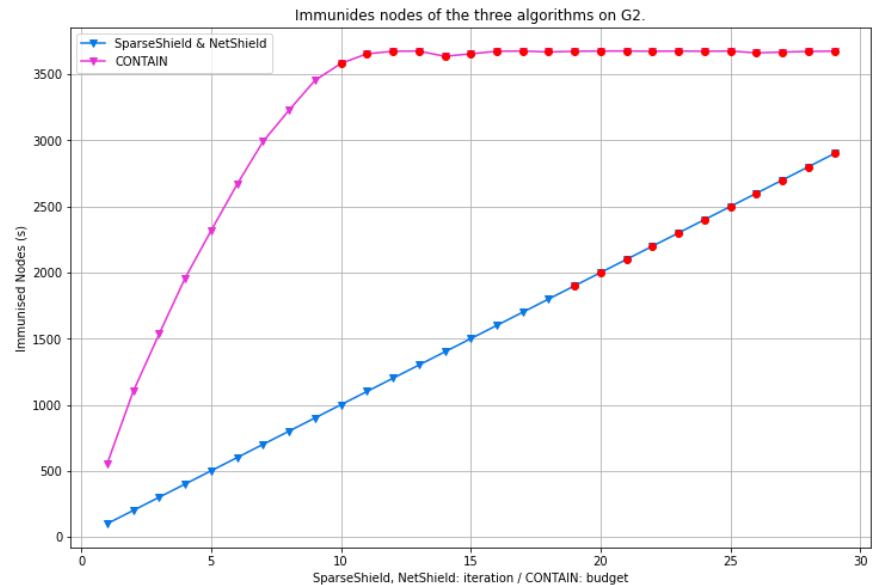
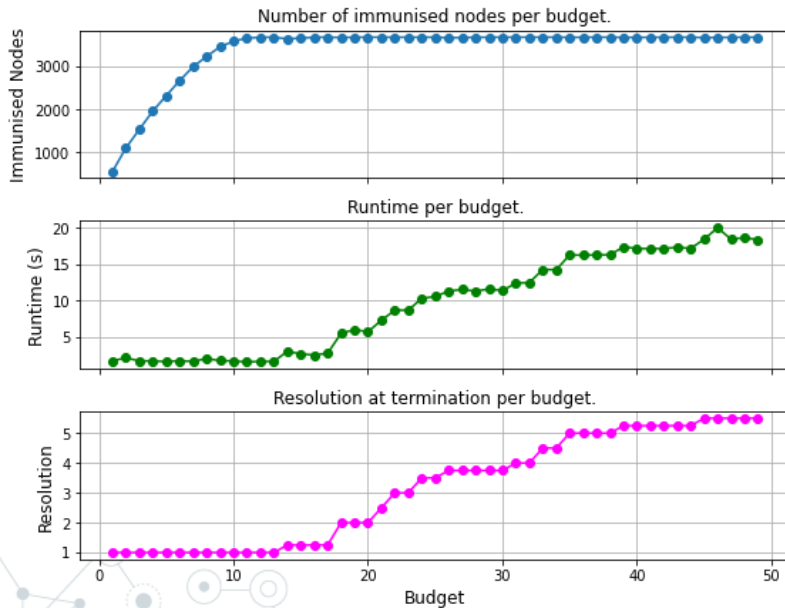




# Rezultate pentru CONTAIN

Seturile de date:

- FakeNewsNet
- Facebook Social Circles





## Concluzii

- Propus un **model de detecție** a informației false din rețele sociale care ține cont de context și conținut
- Propus un nou **algoritm de imunizare** a rețelei bazată pe comunități
- Propus o nouă soluție de distribuție a algoritmilor de **detecție comunități** în mediile sociale
- Diseminat rezultatele într-un articol de jurnal Q1 și două articole de conferință



# Vă mulțumesc !



Elena-Simona Apostol  
[elena.apostol@upb.ro](mailto:elena.apostol@upb.ro)

