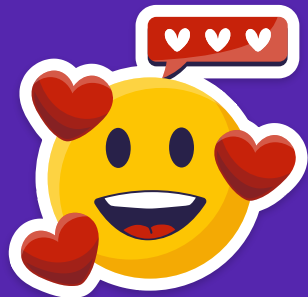




DETECȚIA IRONIEI CU AJUTORUL



EMOTICOANELOR



Lăpăduș Daria 333

Radu Raluca 333

Trandafir Alexandru 333





INTRODUCERE



- ★ În era digitală, comunicarea este rapidă și vizuală – emoji-urile joacă un rol cheie în transmiterea emoțiilor (ex: sarcasm, ironie, frustrare).
- ★ Modelele NLP tradiționale ignoră adesea semnificația emoji-urilor.
- ★ Articolul „**What a Sunny Day: Toward Emoji-Sensitive Irony Detection**” oferă o abordare practică și inovatoare.
- ★ Obiectivul nostru: **dezvoltarea unui model capabil să detecteze ironia influențată de emoji-uri**, cu aplicabilitate în social media monitoring, analiza opiniei și detectarea conținutului toxic.



LUCRĂRI ASEMĂNĂTOARE



Benchmark: SemEval-2018 a oferit date standard pentru detectarea ironiei.

Contribuții importante:

- ★ *Imoji Dataset* → evidențiază rolul emoji-urilor.
- ★ *DeepMoji* → folosește emoji-uri pentru învățarea emoțiilor.
- ★ *BERTweet & RoBERTa + emoji* → îmbunătățesc performanța detectării ironiei.

DESCRIEREA METODEI



1. **Preprocesarea datelor** – normalizarea emoji-urilor, curățarea textului, tokenizare.
2. **Vectorizarea** – folosirea de embeddings pre-antrenate (GloVe pentru text și embeddings speciali pentru emoji-uri).
3. **Modelul** – rețea biLSTM cu atenție, urmată de un strat de clasificare softmax.
4. **Antrenarea și evaluarea** – modelul a fost antrenat pe două subseturi: original și emoji-removed.





COMPARAREA **METODELOR**



Metodă	Avantaje	Dezavantaje
Reguli lingvistice	Explicabil, rapid	Performanță slabă
DeepMoji	Emoții bine reprezentate	Nu e specific pentru ironie
BERTweet	Adaptat la Twitter	Nu include emoji-uri explicit
biLSTM + Attention	Sensibil la polaritatea emoji	Necesită date suplimentare





DATASET

Am utilizat inițial **datasetul de la SemEval 2018 Task 3**.

Am extins acest dataset, construind trei subseturi:

- ★ **original** – tweet-urile așa cum apar în setul SemEval;
- ★ **emoji-removed** – tweet-uri ironice din care au fost eliminate emoji-urile pentru a vedea impactul asupra clasificării.



PREPROCESARE

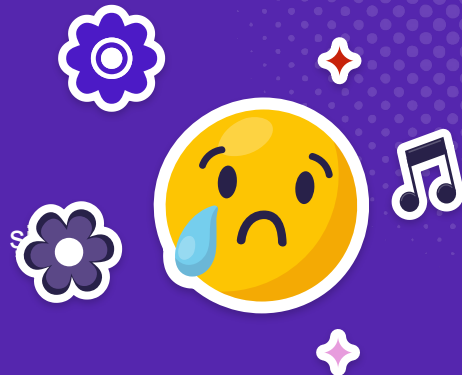
- ★ Eliminarea mențiunilor de utilizator (@username)
- ★ Eliminarea link-urilor
- ★ Conversia la litere mici (lowercasing)
- ★ Eliminarea caracterelor speciale și a spațiilor multiple
- ★ Tokenizarea textului
- ★ Normalizarea emoji-urilor
- ★ Padding și tăiere (padding/truncating)



MODEL

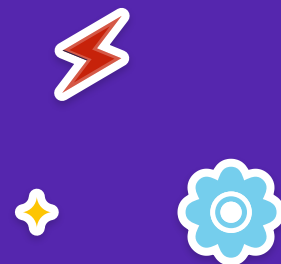
A. Model utilizat

- ★ biLSTM (Bidirectional Long Short-Term Memory) + attention → captează context și relevanță semnatică



B. Parametri

- ★ Dropout: 0.4
- ★ Optimizator: Adam (lr=0.001)
- ★ Dim. embedding: 200 (text), 50 (emoji)
- ★ Batch size: 64
- ★ Epoci: 15





C. Experimente

Model	Set date	Accuracy	F1 (Irony)
biLSTM	original	68.4%	62.3%
biLSTM + attention	emoji-removed	69.5%	64.1%








D. Interpretabilitate

- ★ Atenția se concentrează pe emoji-uri și cuvinte-cheie → înțelegere mai bună a ironiei.



LIMITĂRI



- ★  *Lingvistică*: Modelul funcționează doar pentru limba engleză.
- ★  *Generalizare*: Exemplele sintetice pot fi nerealiste.
- ★  *Resurse*: Modele avansate necesită GPU-uri puternice.
- ★  *Scalabilitate*: Procesarea emoji-urilor nu este complet automatizată.
- ★  *Ambiguitate semantică*: Emoji-urile pot avea multiple interpretări în funcție de context.





CONCLUZII



- ★ Emoji-urile oferă indicii afective utile în detectarea ironiei online.
- ★ Modelele NLP tradiționale, lipsite de sensibilitate vizuală, sunt insuficiente în contextul actual.
- ★ Adăugarea emoji-urilor ca tokeni semantici a îmbunătățit performanța modelului.
- ★ Proiectul a oferit o înțelegere mai profundă a interacțiunii dintre text și simboluri vizual





„What a Sunny Day: Toward Emoji-Sensitive Irony Detection” – Chaudhary et al., 2019

- Lucrarea principală pe care se bazează proiectul.
- Introduce datasetul *Imoji* și o abordare cu biLSTM + attention pentru ironie sensibilă la emoji-uri.

„DeepMoji: Using Millions of Emoji Occurrences to Learn Any Domain” – Felbo et al., 2017

- Demonstrează că emoji-urile pot servi ca proxy pentru emoție, generând embedding-uri utile în analiză afectivă.

„BERTweet: A pre-trained language model for English Tweets” – Nguyen et al., 2020

- Model pre-antrenat pe tweet-uri, utilizat ca bază pentru modele mai complexe de detecție a ironiei.

„SemEval-2018 Task 3: Irony detection in English tweets” – Van Hee et al., 2018

- Competiție NLP ce oferă un benchmark recunoscut pentru detecția ironiei în tweet-uri.

