

Analiza emoțiilor în text utilizând NLP

Octavian Măcelaru

octavian-andrei.macelar@unibuc.ro

Lorena Vătău

lorena.vatau@unibuc.ro

Nicoleta Caramaliu

nicoleta.caramaliu@unibuc.ro

Abstract

Acest proiect prezintă o aplicație de analiză emoțională a textului, bazată pe procesarea limbajului natural (NLP). Sunt integrate mai multe modele pentru detectarea emoțiilor: VADER (pentru engleză), DistilRoBERTa (pentru engleză și română) și un model personalizat LSTM, Nicupiticu, antrenat pe un set de date românesc. Aplicația permite clasificarea multi-etichetă a emoțiilor, vizualizarea distribuției emoționale și compararea rezultatelor. Platforma, dezvoltată cu Streamlit, oferă o interfață interactivă și contribuie la extinderea resurselor NLP pentru limba română.

1 Introducere

Analiza emoțională, cunoscută și sub denumirea de analiză afectivă, reprezintă un subdomeniu esențial al procesării limbajului natural (NLP – Natural Language Processing), ce are ca scop identificarea și interpretarea emoțiilor exprimate în texte. Aceasta joacă un rol din ce în ce mai important într-o varietate largă de domenii, precum jurnalism digital, psihologie computațională, interfețe om–calculator, marketing, asistenți conversaționali și monitorizarea opiniei publice pe rețelele sociale. Identificarea automată a emoțiilor permite înțelegerea profundă a conținutului afectiv din texte, facilitând dezvoltarea de aplicații sensibile la contextul emoțional.

Cu toate acestea, majoritatea cercetărilor și aplicațiilor din acest domeniu sunt centrate în jurul limbii engleze sau a altor limbi cu resurse bogate, în timp ce limbile cu resurse limitate, cum este româna, beneficiază de un suport mult mai redus. Lipsa dataset-urilor emoționale bine etichetate, a modelelor pre-antrenate specializate și a instrumentelor de vizualizare interpretabilă reprezintă obstacole semnificative în dezvoltarea unor soluții viabile pentru limba română.

Lucrarea de față prezintă o platformă completă de analiză emoțională care integrează mai multe

modele de învățare automată: două modele pre-antrenate larg utilizate — VADER (bazat pe reguli, pentru engleză) și DistilRoBERTa (bazat pe Transformers, pentru ambele limbi), precum și un model neural dezvoltat și antrenat local, dedicat în mod special limbii române, denumit *Nicupiticu*. Acest model este antrenat de la zero pe un set de date românesc, utilizând o arhitectură bidirecțională LSTM și un vocabular construit manual din corpusul folosit.

Motivația principală a proiectului constă în completarea lacunelor existente în infrastructura NLP pentru limba română, în special în contextul analizei emoționale. În plus, ne propunem să oferim o soluție interpretabilă și ușor de utilizat, care permite compararea rezultatelor obținute de diverse modele, precum și vizualizarea emoțiilor în mod intuitiv (e.g., prin roata emoțiilor, diagrame radar, grafice comparative).

Contribuțiile principale ale lucrării sunt:

- dezvoltarea unui model personalizat pentru română, capabil să identifice 12 emoții distincte, incluzând emoții precum optimism, pesimism, încredere și anticipare;
- integrarea și compararea a trei abordări complementare de analiză emoțională: bazată pe reguli (VADER), bazată pe Transformers (DistilRoBERTa) și model neural local (Nicupiticu);
- realizarea unei platforme interactive, cu suport bilingv, pentru testarea și vizualizarea rezultatelor emoționale;
- crearea și preprocesarea unui set de date etichetat cu emoții multiple, adaptat limbii române.

Prin această lucrare, ne propunem să susținem eforturile de extindere a tehnologiilor NLP pentru

limbi cu resurse limitate și să oferim un cadru extensibil și reproductibil pentru cercetări viitoare în analiza afectivă.

2 Lucrări Asemănătoare

Inițial, metodele erau bazate pe lexicoane de sentimente (de exemplu, VADER (Hutto and Gilbert, 2014)), care etichetau cuvintele individuale ca pozitive, negative sau neutre. Aceste abordări oferă interpretabilitate ridicată, dar sunt limitate în captarea contextului semantic.

Odată cu avansul rețelelor neuronale și al transformerelor, modelele de tip BERT și DistilRoBERTa (Vaswani et al., 2017) au devenit standard în analiza emoțională, oferind performanțe superioare în înțelegerea nuanțelor afective. Modele precum *GoEmotions* (Demszky et al., 2020) au fost antrenate pe date anotate cu 27 de emoții fine-grained, fiind frecvent utilizate ca bază pentru transfer learning.

Pentru limba română, resursele sunt mai puține. Există inițiative precum *RoBERT* (Dumitrescu et al., 2020), dar puține modele sunt optimizate pentru detecția emoțiilor. Contribuția noastră constă în integrarea unui model personalizat, Nicupiticu, antrenat pe date românești, și compararea acestuia cu modele pre-antrenate, demonstrând astfel importanța adaptării locale în NLP-ul afectiv.

- Am consultat surse precum Google Scholar, ACL Anthology și PapersWithCode pentru identificarea celor mai recente lucrări în domeniul detecției emoțiilor și clasificării multi-etichetă.
- Cele mai relevante direcții includ transfer learning pentru limbi cu resurse reduse, prelucrarea contextuală a emoțiilor și vizualizarea distribuției afective.

3 Metodologie

Această secțiune descrie arhitecturile utilizate, strategiile de preprocesare și modul în care au fost integrate diferitele modele pentru analiza emoțiilor. Ne-am concentrat pe comparabilitatea rezultatelor și pe evaluarea complementară a abordărilor statistice, neuronale și bazate pe transformers, în contextul unui set de date adaptat limbii române.

3.1 Setul de date

Setul de date utilizat pentru antrenarea modelului Nicupiticu conține exemple de propoziții în limba

română etichetate cu una dintre cele 12 emoții definite în roata emoțiilor a lui Plutchik: bucurie, tristețe, furie, frică, surpriză, neutru, iubire, dezgust, optimism, pesimism, încredere și anticipare. Etichetarea a fost realizată manual, iar distribuția este aproximativ echilibrată, deși anumite emoții precum „anticipare” sau „dezgust” apar mai rar.

Pentru evaluare, am folosit și texte de test din surse externe, inclusiv exemple generate manual pentru a observa generalizarea modelelor pe date realiste și variate.

3.2 Preprocesare

Preprocesarea a constat în:

- normalizarea textului: conversie la litere mici, eliminarea semnelor de punctuație și a spațiilor multiple;
- tokenizare și mapare în secvențe numerice folosind un vocabular construit pe întregul set de date;
- completarea secvențelor la o lungime fixă (padding) pentru a fi compatibile cu rețeaua neuronală LSTM;
- etichetarea emoțiilor folosind un cod numeric prin LabelEncoder pentru a putea fi folosite în CrossEntropyLoss.

Pentru modelele bazate pe transformer, am folosit tokenizer-ul nativ pre-antrenat (BERT tokenizer), iar pentru VADER, textul a fost lăsat aproape intact, întrucât modelul este sensibil la punctuație și majuscule.

3.3 Modele

Am testat următoarele arhitecturi:

- **VADER** – un model lexicon-based utilizat pentru text în limba engleză. Returnează scoruri pentru polaritate: pozitiv, negativ, neutru și un scor compus.
- **DistilRoBERTa** – folosit în două variante: pre-antrenat pentru engleză și pentru română, adaptat pentru clasificare multi-etichetă.
- **Nicupiticu** – un model bidirecțional LSTM cu două straturi ascunse, 200 de dimensiuni pentru embedding, dropout 0.5 și clasificator dense cu funcție de activare ReLU. Modelul a fost antrenat timp de 50 de epoci cu AdamW și ReduceLROnPlateau ca scheduler.

Am comparat acuratețea de validare a modelului personalizat (Nicupiticu) cu celelalte două. Nicupiticu a obținut o acuratețe generală de aproximativ 84%, fiind competitiv cu DistilRoBERTa, care a înregistrat scoruri ușor mai mari doar pentru emoțiile frecvente. VADER, fiind optimizat pentru engleză, a funcționat slab pe date în română, confirmând importanța alegerii limbii în modelare.

Interpretabilitate: rezultatele pot fi interpretate atât numeric, cât și vizual (roata emoțiilor, scoruri detaliate, gauge chart), iar emoțiile dominante sunt ușor de identificat. Modelul VADER oferă scoruri intuitive, dar limitate la polaritate, în timp ce Nicupiticu și DistilRoBERTa oferă distribuții mai fine, utile în analiza afectivă complexă.

4 Lucrări Viitoare

Pentru dezvoltări viitoare, ne propunem extinderea setului de date, introducerea de date din alte domenii (psihoterapie, social media), testarea transferului cross-lingvistic și implementarea unei interfețe interactive care să permită etichetarea participativă. De asemenea, vom explora integrarea emoțiilor ca trăsături în sarcini conexe, cum ar fi detectarea discursului instigator la ură sau analiza intenției.

5 Concluzie

Această lucrare a prezentat dezvoltarea unei platforme avansate pentru analiza emoțiilor în texte scrise, cu un accent special pe limba română – o limbă cu resurse semnificativ mai reduse față de engleză. Am propus și validat un model neural propriu, denumit *Nicupiticu*, antrenat pe un set de date adaptat, capabil să recunoască 12 emoții diferite. Rezultatele experimentale au arătat că acest model este competitiv în raport cu alternativele consacrate, cum ar fi DistilRoBERTa și VADER, oferind în același timp un grad sporit de interpretabilitate, crucial pentru aplicații sensibile, cum ar fi psihologia computațională sau analiza discursului public.

Pe lângă validarea modelului, platforma noastră a demonstrat valoarea integrării mai multor tipuri de modele (bazate pe reguli, pe transformers și pe rețele neuronale clasice), fiecare aducând perspective diferite asupra textului analizat. Un alt rezultat important al proiectului este modul intuitiv de prezentare a rezultatelor, prin vizualizări interactive și clare, care facilitează înțelegerea distribuției emoțiilor în text.

Dezvoltarea sistemului a evidențiat și importanța

unor componente adesea subestimate, precum selecția riguroasă a datelor de antrenament, preprocesarea lingvistică specifică limbii române și echilibrarea distribuției emoțiilor în setul de date. Experiența acumulată ne-a permis să înțelegem în profunzime provocările antrenării de la zero a unui model într-un domeniu complex, precum cel afectiv, și să construim o soluție scalabilă și reutilizabilă în contexte viitoare.

În ansamblu, proiectul ne-a oferit o experiență completă de cercetare aplicată, pornind de la problematizare și documentare, trecând prin dezvoltarea și testarea arhitecturilor de rețea, până la evaluarea comparativă și integrarea într-o aplicație interactivă. Aceste etape ne-au consolidat abilitățile atât în procesarea limbajului natural, cât și în design-ul de sisteme software orientate spre utilizator.

6 Limitări

Modelul Nicupiticu este antrenat exclusiv pe date în limba română, deci nu este aplicabil cross-lingvistic fără adaptare. Antrenarea este sensibilă la dezechilibre în date, iar interpretarea emoțiilor ambigue poate genera predicții incerte. Platforma necesită resurse moderate (GPU), iar scorurile pot depinde de lungimea textului.

7 Declarație Etică

Analiza emoțională automată ridică riscuri etice legate de interpretarea greșită a stărilor psihologice și de potențiala utilizare abuzivă în contextul supravegherii sau manipulării. Pentru a evita aceste riscuri, am folosit doar date publice și anonime, iar modelul nostru nu ia decizii automatizate fără supraveghere umană.

Conștientizăm existența unor posibile biasuri în datele de antrenare, motiv pentru care recomandăm utilizarea responsabilă și transparentă a sistemului. Modelul trebuie folosit exclusiv ca instrument de sprijin, nu ca substitut pentru evaluarea umană.

References

- Dorottya Demszky, Dana Movshovitz-Attias, Jeongwoo Ko, Alan Cowen, Gaurav Nemade, and Sujith Ravi. 2020. Goemotions: A dataset of fine-grained emotions. *arXiv preprint arXiv:2005.00547*.
- Stefan Daniel Dumitrescu, Andrei M. Avram, Alina Maria Ciobanu, and Liviu P. Dinu. 2020. Introducing the romanian bert language model. In *Findings of the*

Association for Computational Linguistics: EMNLP 2020, pages 432–441.

C.J. Hutto and Eric Gilbert. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Eighth International Conference on Weblogs and Social Media (ICWSM-14)*.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems*.

A Anexă

Aceasta este secțiunea unde se pot adăuga tabele, formule sau rezultate suplimentare — elemente care oferă o înțelegere detaliată a cercetării, dar nu sunt esențiale pentru prezentarea metodei în sine.