

Proiect de cercetare:

” Îmbunătățirea prognozelor ratei inflației în România folosind analiza de sentiment și machine learning (învățare automată)”



Director de proiect: Prof. univ. dr. habil. CS I Mihaela Simionescu

Membru în proiect: dr. CS III Alexandru-Sabin Nicula

Prezentul raport intermediar include un studiu corelat cu tema proiectului împreună cu articolul corespunzător publicat într-o revistă WoS, autorii având afilierea AOSR.

1. Introducere

Context

- România – cea mai mare inflație din UE la începutul anului 2024 (7,3%);
- Creștere semnificativă în anul 2024: 7,3% (ianuarie) vs. 3,1% (media UE);
- Scădere ușoară prognozată de BNR: 4,7% (sfârșitul anului 2024), 3,5% (sfârșitul anului 2025).

Importanța prognozelor cu acuratețe mare

- controlul inflației: esențial pentru convergența către zona euro;
- luarea deciziilor bine fundamentate: agenți economici și factori de decizie politică.

Impactul inflației asupra agenților economici

- anticiparea schimbărilor pieței: ajustarea prețurilor, bugetarea cheltuielilor.
- gestionarea riscurilor: costuri mai mari ale intrărilor, modele de preț ajustate.
- protejarea profiturilor: decizii strategice bazate pe prognoze inflației.

Impactul inflației asupra băncilor centrale

- politica monetară: stabilizarea prețurilor și creșterea economică.
- ratele dobânzii și alte instrumente: menținerea inflației sub control.
- mediu economic sănătos: stabilitate macroeconomică.

Impactul inflației asupra așteptărilor pieței

- influențarea deciziilor economice: investiții, negocieri salariale, cheltuieli.
- perspectivele economiștilor și băncilor centrale: baza pentru luarea deciziilor.

2. Analiza de sentiment folosită în elaborarea prognozelor economice

2.1. Definiția sentimentului

Definiția sentimentului

- Sentimentul ca parametru sau variabilă în modelarea econometrică;
- Sentometrie: o zonă de cercetare nouă cu potențial în previziune.

- Abordări din diverse domenii:
 - lingvistică: punct de vedere subiectiv (pozitiv/negativ);
 - sociologie: stare umană legată de contextul spațio-temporal;
 - psihologie: expresie a sentimentelor, credințelor și dorințelor;
 - economie: tonul știrilor asociate unui subiect (incertitudine, pozitivitate, negativitate).

Tipuri de sentiment

- sentiment explicit: transmite stări subiective private;
- sentiment implicit: transmite informații factuale;
- valență: gradul de pozitivitate/negativitate;
- stimulare: nivelul de intensitate.

Aplicații ale sentimentului în econometrie

- contabilitate: evaluarea riscului și performanței;
- finanțe: previziunea prețurilor acțiunilor și a rentabilității investițiilor;
- politică: analiza opiniei publice și a sentimentului electoral;
- marketing: înțelegerea preferințelor clienților și a tendințelor pieței;
- psihologie: studierea comportamentului și a emoțiilor umane;
- informatică: dezvoltarea de sisteme de procesare a limbajului natural.

Potențialul sentimentului în previziune

- așteptările pieței influențate de sentiment;
- sentimentul ca indicator al șocurilor economice;
- îmbunătățirea modelelor econometrice prin incorporarea sentimentului.

2.2. Analiza sentimentului în prognoze

- sentimentul – rol important în luarea deciziilor economice.
- analiza sentimentului – îmbunătățirea acurateții previziunilor economice.

Importanța sentimentului în economie

- economiștii – influența sentimentului asupra piețelor financiare;
- analiza sentimentului poate identifica tendințe și schimba comportamentul;
- sentimentul – previziuni ale indicatorilor economici.

Metode de analiză a sentimentului

- analiza de sentiment – tehnici matematice și date textuale;
- metode principale de măsurare a sentimentului:
 - indici de sentiment bazați pe conversia datelor calitative;
 - indici bazați pe date din chestionare și sondaje;
 - indici determinați de comunicări oficiale.

Integrarea sentimentului în prognozele economice

- studii -> sentimentul poate îmbunătăți acuratețea previziunilor;
 - sentimentul integrat în modele econometrice;
 - analiza sentimentului pentru a predicta indicatorii economici, cum ar fi inflația.
-
- acest studiu -> utilitatea analizei sentimentului în prognoza inflației;
 - indici de sentiment derivați din rapoartele BNR.

- Clements și Reade (2020) – abordare bazată pe dicționar pentru a calcula indicii sentimentului, concentrându-se pe incremente și decremente;
- Indici de sentiment folosind texte din ziare:
 - Garcia (2013) – articole din New York Times pentru sentimentul pieței financiare;
 - Baker et al. (2016)– zece ziare pentru a crea un indice care indică incertitudinea politicii economice;
 - Fraiberger (2016) – articole din Reuters pentru a prezice randamentele și riscurile pieței bursiere.
 - Nyman et al. (2021) – prognoze ale dificultăților sistemului financiar folosind date din ziare;
 - Thorsrud (2020) – îmbunătățirea prognozelor actuale ale PIB folosind sentimentul din ziarele de afaceri din Norvegia;
 - Bortoli et al. (2018) – îmbunătățirea prognozelor pe termen scurt ale PIB folosind analiza de sentiment pe baza textelor din ziarul francez *Le Monde*.

3. Metodologia de cercetare

Procesarea limbajului natural (NLP)

- un domeniu al inteligenței artificiale;
- obiectiv: interacțiunea dintre computere și limbajul uman;
- utilizări: analiza sentimentelor, rezumatul textului, traducerea automată și recunoașterea entității de nume.

Analiza sentimentelor

- o tehnică NLP;
- obiectiv: determinarea polarității unui text;
- identifică dacă un text este pozitiv, negativ sau neutru;
- utilizări: analiza opiniilor clienților, monitorizarea rețelelor sociale și urmărirea sentimentului pieței.

Principalele abordări

- **Abordare bazată pe lexicon:**

- dicționar de cuvinte pre-marcate ca fiind pozitive, negative sau neutre;
- analiza textului prin căutarea cuvintelor din dicționar și atribuirea unui scor în funcție de polaritatea cuvântului.

- **Abordare bazată pe învățarea automată:**

- model de învățare automată instruit pe un set de date de text etichetat cu polaritatea;
- model utilizat apoi pentru a prezice polaritatea textului nou.

Calculul indicelui de sentiment

Indicele de sentiment= valoare numerică care reprezintă polaritatea generală a unui text.

- calcul indice de sentiment: însumarea scorurilor individuale ale cuvintelor din text;
- calcul scoruri cuvinte: o abordare bazată pe lexicon sau o abordare bazată pe învățarea automata;
- indicele de sentiment– normalizat la o valoare cuprinsă între 0 și 1 (0– text negativ și 1– text pozitiv).

Indicele de sentiment calculat (*is*) joacă rolul de variabilă explicativă în diverse reprezentări ale modelelor autoregresive cu decalaj distribuit (ARDL), care sunt utilizate ca metodă de prognoză a inflației.

$$inflatie_t = \alpha_1 + \beta_1 \cdot inflatie_{t-1} + \gamma_1 \cdot is_t + \varepsilon_{1t} \quad (1)$$

$$\ln(inflatie_t) = \alpha_2 + \beta_2 \cdot \ln(inflatie_{t-1}) + \delta_1 \cdot \ln(somaj_t) + \varepsilon_{2t} \quad (2)$$

$$\ln(inflatie_t) = \alpha_3 + \beta_3 \cdot \ln(inflatie_{t-1}) + \gamma_2 \cdot is_t + \delta_2 \cdot \ln(somaj_t) + \varepsilon_{3t} \quad (3)$$

inflatie- rata inflației trimestriale înregistrată în perioada actuală *t*;

is- indice de sentiment;

u- rata șomajului;

$\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \gamma_1, \gamma_2, \delta_1, \delta_2$ - parametrii;

$\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t}$ - erori;

t- indice de timp.

Indicatori utilizați pentru a evalua acuratețea previziunilor inflației:

- Eroarea medie (ME);
- Eroarea medie absolută (MAE);
- Rădăcina pătrată a erorii medii pătrate (RMSE);
- Coeficientul U1 al lui Theil;
- Coeficientul U2 al lui Theil;
- Eroarea medie procentuală absolută (MAPE).

4. Rezultate

- metodă de previziune a inflației utilizând indicatori de sentiment;
 - modele ARDL (Autoregressive Distributed Lag) pentru a construi previziuni;
 - variabile în modele: rata inflației înregistrată, indicele de sentiment și rata șomajului.
-
- valorile indicilor de sentiment calculate cu ajutorul IntelliDocker.
 - se verifică staționaritatea seriilor de timp înainte de a construi modelele ARDL;
 - modele ARDL pentru a explica rata inflației înregistrată pe baza indicelui de sentiment;
 - serii de timp din perioada 2006: Q1 - 2023: Q4.
-
- statisticile descriptive ale datelor (Tabelul 1).
 - modelele ARDL, inclusiv coeficienții modelului și performanța previziunilor.

Tabelul 1. Statistici descriptive

Statistici	Rata inflației	Rata șomajului
Media	4.92%	6.73%
Mediana	4.82%	6.10%
Valoare maximă	16.37%	9.3%
Valoare minimă)	-1.70%	3.70%
Deviația standard	3.78%	1.78%
Coeficientul de asimetrie	1.00	0.05
Coeficientul de aplatizare	4.29	1.51
Testul Jarque-Bera (valoarea p în paranteze)	17.03 (0.0002)	6.67 (0.03)
Numărul de observații	72	72

Sursa: calcule proprii.

Rata maximă a inflației (16,37%) din România – în ultimul trimestru al anului 2022.

Cauze hiperinflație :

- războiul din Ucraina: creșterea prețurilor la energie și alimente la nivel global;
- perturbările lanțului de aprovizionare: Pandemia a continuat să afecteze aprovizionarea cu diverse bunuri.
- cererea globală ridicată: presiune suplimentară pe prețuri.

Cea mai mare dezinflație din România– în trimestrul al treilea din 2019.

Cauze dezinflație: epuizarea cererii excedentare a consumatorilor.

Rata șomajului:

- rata trimestrială a șomajului maximă – în primul trimestru al anului 2010, în timpul recesiunii economice;
- rata trimestrială a șomajului minimă – în al doilea trimestru al anului 2008, în timpul boom-ului economic.

Factori de influență pentru rata șomajului:

- creșterea economică: o creștere economică rapidă -> noi locuri de muncă -> reducerea șomajului;
- investițiile: investițiile străine și interne -> crearea de noi firme și extinderea celor existente-> crearea de locuri de muncă;
- politicile guvernamentale: politicile care sprijină afacerile -> țară mai atractivă pentru investiții > crearea de locuri de muncă.

Testul Jarque-Bera

- ipoteza distribuției normale nesuținută pentru majoritatea seriilor de date la un nivel de semnificație de 5%;
- la un nivel de semnificație de 1%, distribuție normală pentru seria de date privind rata șomajului.

Staționaritatea seriilor de timp

- serii de timp ajustate sezonier;
- testul Augmented Dickey-Fuller (ADF): seria pentru indicele de sentiment staționară, restul seriilor $I(1)$ (Tabelul 2).

Tabelul 2. Rezultatele testului ADF

Variabila	Seriile de date	Ecuția	stat. ADF	Valoarea p	Concluzia
Rata inflației	Seria în dublă diferență	fără termen fără termen liber	-6,801	<0,01	fără rădăcină unitate
	Seria în primă diferență		-7,593	<0,01	fără rădăcină unitate
	Seria în nivel		-1,608	0,101	rădăcină unitate
Rata șomajului	Seria în dublă diferență	fără termen fără termen liber	-8,572	<0,01	fără rădăcină unitate
	Seria în primă diferență		-7,125	<0,01	fără rădăcină unitate
	Seria în nivel		-0,420	0,5280	rădăcină unitate
Indice de sentiment	Seria în dublă diferență	fără termen fără termen liber	-6,620	<0,01	fără rădăcină unitate
	Seria în primă diferență		-14,305	<0,01	fără rădăcină unitate
	Seria în nivel	termen liber	-7,670	<0,01	fără rădăcină unitate
ln(inflație)	Seria în dublă diferență	fără termen fără termen liber	-6,862	<0,01	fără rădăcină unitate
	Seria în primă diferență		-5,882	<0,01	fără rădăcină unitate
	Seria în nivel		-0,614	0,4471	rădăcină unitate
ln(șomaj)	Seria în dublă diferență	fără termen fără termen liber	-12,414	<0,01	fără rădăcină unitate
	Seria în primă diferență		-6,866	<0,01	fără rădăcină unitate
	Seria în nivel		-0,306	0,5718	rădăcină unitate

- rezultatele estimărilor în Tabelul 3;
- problema multicoliniarității atenuată în modelul care include rata șomajului prin transformarea datelor în logaritm natural;
- corelație nesemnificativă între rata înregistrată a inflației și indicele de sentiment.

Tabelul 3. Rezultatele estimărilor pentru modelele ARDL

Variabila	Coeficienți (valorile p în paranteze)		
Model	M1	M2	M3
rata inflației în perioada anterioară	0.796 (<0.01)	-	-
indice de sentiment	-5.220 (0.0315)	-	-
ln(rata inflației în perioada anterioară)	-	0.822 (<0.01)	0.768 (<0.01)
ln(rata șomajului)	-	0.114 (0.0758)	-0.038 (0.0865)
ln(indice de sentiment)	-	-	-0.793 (0.0304)
Constanta	4.138 (0.0105)	-	-0.041 (0.936)
Testul White: stat. (valoarea p între paranteze)	1.668 (0.8928)	0.392 (0.531)	0.956 (0.3281)
Testul Breusch-Godfrey pentru lag=1: stat. (valoarea p între paranteze)	0.338 (0.5608)	0.350 (0.5540)	2.171 (0.1406)
Testul Breusch-Godfrey pentru lag=2: stat. (valoarea p între paranteze)	0.618 (0.7341)	1.445 (0.4855)	2.228 (0.3282)
Testul Jarque-Bera: stat. (valoarea p între paranteze)	31.599 (<0.01)	28.792 (<0.01)	51.088 (<0.01)
Testul Ramsey Reset (valoarea p între paranteze)	1.249 (0.2677)	1.269 (0.2641)	0.761 (0.3863)

Sursa: calcule proprii.

- previziuni pentru rata inflației pe baza modelelor ARDL;
- acuratețea previziunilor comparată cu cea a previziunilor BNR;
- predicții combinate între așteptările numerice ale BNR și celelalte previziuni folosind scheme de combinare tradiționale.

Acuratețea previziunilor (Tabelul 4):

- previziunile de sentiment bazate pe modelul ARDL (indicele de sentiment și rata șomajului ca variabile explicative) – acuratețe maximă;
 - previziuni bazate doar pe indicele de sentiment – mai bune decât previziunile numerice ale BNR.
 - predicțiile combinate nu au îmbunătățit previziunile BNR sau cele de sentiment;
 - predicțiile combinate bazate pe modelul ARDL (indicele de sentiment și rata șomajului ca variabile explicative) și predicțiile BNR (scheme cu ponderi optime și inverse)- superioare celorlalte predicții după MAPE.
-
- indicii de sentiment bazați pe așteptările experților– rol important în realizarea unor previziuni cu acuratețe mare pentru rata trimestrială a inflației în România.
 - metoda propusă, care a determinat cele mai bune previziuni– utilizată pentru a realiza următoarele previziuni trimestriale ale inflației pentru anii 2024 și 2025.

Tabelul 4. Acuratețea previziunilor privind rata trimestrială a inflației pentru România (2006: T1-2023: T4)

Previziuni	ME	MAE	RMSE	U1	U2	MAPE
Previziuni BNR	0,005	0,725	1,954	0,114	0,29	0,196
Previziuni de sentiment bazate pe modelul ARDL și indici de sentiment (sf1)	0,004	0,343	1,924	0,064	0,23	0,157
Previziuni de sentiment bazate pe modelul ARDL, indici de sentiment și șomaj (sf2)	-0,004	0,219	1,147	0,022	0,174	-0,052
Previziuni bazate pe modelul ARDL	-0,943	1,1	2,145	0,568	0,375	-1,033
Previziuni combinate bazate pe previziunile BNR și sf1, cu ponderi optime	0,011	1,354	2,933	0,164	0,93	0,352
Previziuni combinate bazate pe previziunile BNR și sf1, cu ponderi inverse	0,0089	0,983	2,414	0,128	0,595	0,246
Previziuni combinate bazate pe previziunile BNR și sf1, cu ponderi egale	0,008	0,957	2,344	0,127	0,571	0,24
Previziuni combinate bazate pe previziunile BNR și sf2, cu ponderi optime	0,085	0,301	0,447	0,233	0,426	0,047
Previziuni combinate bazate pe previziunile BNR și sf2, cu ponderi inverse	0,062	0,266	0,398	0,131	0,718	0,027
Previziuni combinate bazate pe previziunile BNR și sf2, cu ponderi egale	0,056	0,258	0,389	0,200	0,348	0,020

Sursa: calcule proprii.

5. Concluzii

- previziuni ale ratei inflației-> luarea unor decizii economice informate de către firme și factorii de decizie politică;
- controlul inflației-> stabilitatea financiară și bunăstarea economică a României.
- previziuni precise ale inflației -> gestionare eficientă a acestei probleme actuale ale economiei românești.

- abordare inovatoare a prognozei inflației prin integrarea indicilor de sentiment;
- previziunile de sentiment bazate pe ARDL (indicele de sentiment și rata șomajului ca variabile explicative) - superioare previziunilor numerice ale BNR.

Limitele cercetării

- o singură metodă de construire a indicilor de sentiment;
- numai scheme tradiționale de combinare a previziunilor.

Direcții viitoare de cercetare

- metode alternative de construire a indicilor de sentiment;
- tehnici avansate de combinare a previziunilor;
- extinderea analizei la alte țări.

Bibliografie

- Angeletos, G.M., Collard, F., Dellas, H., 2018. Quantifying confidence. *Econometrica* 86 (5), 1689–1726.
- Angeletos, G.M., La'O, J., 2013. Sentiments. *Econometrica* 81 (2), 739–779.
- Ardia, D., Bluteau, K., Borms, S., Boudt, K., 2020. Econometrics meets sentiment: an overview of methodology and applications. *J. Econ. Surv.* 34 (3), 512–547.
- Ardia, D., Bluteau, K., Boudt, K., 2021. The R package sentometrics to compute, aggregate and predict with textual sentiment. *J. Stat. Softw.* 99 (2), 1–40.
- Ardia, D., Bluteau, K., Boudt, K., 2019. Questioning the news about economic growth: Sparse forecasting using thousands of news-based sentiment values. *International Journal of Forecasting*, 35(4), 1370-1386.
- Baciu, I.C., 2015. Stochastic models for forecasting inflation rate. Empirical evidence from Romania. *Procedia Econ. Financ.* 20, 44–52.
- Bajo, E. and Raimondo, C., 2017. Media sentiment and IPO underpricing. *Journal of Corporate Finance*, 46, 139-153.
- Baker, S.R., Bloom, N. and Davis, S.J., 2016. Measuring economic policy uncertainty. *The quarterly journal of economics*, 131(4), 1593-1636.
- Bortoli, C., Combes, S., Renault, T., 2018. Nowcasting GDP growth by Reading newspapers. *Econ. Stat.* 505 (1), 17–33.
- Cambria, E., Das, D., Bandyopadhyay, S., Feraco, A., 2017. Affective computing and sentiment analysis. In: *A Practical Guide to Sentiment Analysis*. Springer, Cham.
- Casey, G.P. and Owen, A.L., 2013. Good news, bad news, and consumer confidence. *Social Science Quarterly*, 94(1), 292-315.
- Castle, J.L., Hendry, D.F. and Martinez, A.B., 2017. Evaluating forecasts, narratives and policy using a test of invariance. *Econometrics*, 5(3), 39.
- Clements, M.P. and Reade, J.J., 2020. Forecasting and forecast narratives: The Bank of England inflation reports. *International Journal of Forecasting*, 36(4), 1488-1500.
- Di Fatta, G., Reade, J.J., Jaworska, S. and Nanda, A., 2015, December. Big social data and political sentiment: the tweet stream during the UK general election 2015 campaign. In *2015 IEEE international conference on smart city/socialcom/sustaincom (smartcity)* (pp. 293-298). IEEE.
- Ericsson, N.R., 2016. Eliciting GDP forecasts from the FOMC's minutes around the financial crisis. *Int. J. Forecast.* 32 (2), 571–583.

- Eugster, P. and Uhl, M.W., 2024. Forecasting inflation using sentiment. *Economics Letters*, 111575.
- Evans, J.A. and Aceves, P., 2016. Machine translation: Mining text for social theory. *Annual review of sociology*, 42, 21-50.
- Feldman, R., Govindaraj, S., Livnat, J. and Segal, B., 2010. Management's tone change, post earnings announcement drift and accruals. *Review of Accounting Studies* 15(4), 915–953.
- Fraiberger, S., 2016. November. News sentiment and cross-country fluctuations. In: *Proceedings of the First Workshop on NLP and Computational Social Science*, 125–131.
- Garcia, D., 2013. Sentiment during recessions. *J. Financ.* 68 (3), 1267–1300.
- Grimmer, J. and Stewart, B.M., 2013. Text as data: the promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis* 21(3), 267–297.
- Hansen, S., McMahon, M., 2016. Shocking language: Understanding the macroeconomic effects of central bank communication. *J. Int. Econ.* 99, S114–S133.
- Henry, E., 2008. Are investors influenced by how earnings press releases are written? *Journal of Business Communication* 45(4), 363–407.
- Hussein, D.M.E.-D.M. 2018. A survey on sentiment analysis challenges. *J. King Saud Univ.-Eng. Sci.* 30, 330–338.
- Jegadeesh, N. and Wu, D., 2013. Word power: A new approach for content analysis. *Journal of financial economics*, 110(3), 712-729.
- Jones, J.T., Sinclair, T.M. and Stekler, H.O., 2020. A textual analysis of Bank of England growth forecasts. *International Journal of Forecasting*, 36(4), 1478-1487.
- Kearney, C. and Liu, S., 2014. Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis*, 33, 171-185.
- Kräussl, R. and Mirgorodskaya, E., 2017. Media, sentiment and market performance in the long run. *The European Journal of Finance*, 23(11), 1059-1082.
- Liang, B., Su, H., Gui, L., Cambria, E. and Xu, R., 2022. Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks. *Knowledge-Based Systems*, 235, 107643.
- Ludvigson, S.C., 2004. Consumer confidence and consumer spending. *J. Econ. Perspect.* 18 (2), 29–50.
- Lukauskas, M., Pilinkienė, V., Bruneckienė, J., Stundžienė, A., Grybauskas, A. and Ruzgas, T., 2022. Economic activity forecasting based on the sentiment analysis of news. *Mathematics*, 10(19), 3461.

- Munezero, M., Montero, C.S., Sutinen, E. and Pajunen, J., 2014. Are they different? Affect, feeling, emotion, sentiment, and opinion detection in text. *IEEE transactions on affective computing*, 5(2), 101-111.
- Nyman, R., Kapadia, S. and Tuckett, D., 2021. News and narratives in financial systems: exploiting big data for systemic risk assessment. *Journal of Economic Dynamics and Control*, 127, 104119.
- Romer, C.D., Romer, D.H., 2008. The FOMC versus the staff: where can monetary policymakers add value? *Am. Econ. Rev.* 98 (2), 230–235.
- Shapiro, A.H., Sudhof, M. and Wilson, D., 2018. Measuring news sentiment. Technical Report 2017-01, Federal Reserve Bank of San Francisco.
- Shapiro, A. H., Sudhof, M., & Wilson, D. J. 2022. Measuring news sentiment. *Journal of econometrics*, 228(2), 221-243.
- Sharpe, S.A., Sinha, N.R. and Hollrah, C.A., 2023. The power of narrative sentiment in economic forecasts. *International Journal of Forecasting*, 39(3), 1097-1121.
- Simionescu, M., 2020. Bayesian combined forecasts and Monte Carlo simulations to improve inflation rate predictions in Romania. *Rese. Pap. Econ. Financ.* 4 (1), 7–20.
- Simionescu, M., 2022. Econometrics of sentiments-sentometrics and machine learning: the improvement of inflation predictions in Romania using sentiment analysis. *Technological Forecasting and Social Change*, 182, p.121867.
- Stekler, H. and Symington, H., 2016. Evaluating qualitative forecasts: The FOMC minutes, 2006–2010. *International Journal of Forecasting*, 32(2), 559-570.
- Susskind, Z., Arden, B., John, L.K., Stockton, P., John, E.B., 2021. Neuro-Symbolic AI: An Emerging Class of AI Workloads and Their Characterization. arXiv preprint arXiv: 2109.06133.
- Taboada, M., 2016. Sentiment analysis: An overview from linguistics. *Annual Review of Linguistics*, 2, 325-347.
- Tetlock, P.C., 2007. Giving content to investor sentiment: The role of media in the stock market. *The Journal of finance*, 62(3), 1139-1168.

- Thorsrud, L.A., 2018. Words are the new numbers: a newsy coincident index of the business cycle. *Journal of Business & Economic Statistics* 38(2), 393–409.
- Thorsrud, L.A., 2020. Words are the new numbers: A newsy coincident index of the business cycle. *Journal of Business & Economic Statistics*, 38(2), 393-409.
- Van de Kauter, M., Desmet, B. and Hoste, V., 2015. The good, the bad and the implicit: a comprehensive approach to annotating explicit and implicit sentiment. *Language Resources & Evaluation* 49(3), 685–720.
- Wang, W., Guo, L., Wu, Y.J., 2022. The merits of a sentiment analysis of antecedent comments for the prediction of online fundraising outcomes. *Technol. Forecast. Soc. Chang.* 174, 121070.
- Wang, Y., Song, W., Tao, W., Liotta, A., Yang, D., Li, X., Zhang, W., 2022. A Systematic Review on Affective Computing: Emotion Models, Databases, and Recent Advances. *Information Fusion*.
- Zhang, L., Wang, S., Liu, B. 2018. *Deep learning for sentiment analysis: A survey*. Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 8, e1253.

6. Diseminarea rezultatelor cercetării



Marketing and Management of Innovations, 15(2), 2024



Sentiment Analysis as an Innovation in Inflation Forecasting in Romania

Mihaela Simionescu^{1,2,3,*} , Alexandru-Sabin Nicula^{2,4} 

- 1 Faculty of Business and Administration, University of Bucharest, Bucharest, Romania
 - 2 Academy of Romanian Scientists, Bucharest, Romania
 - 3 Institute for Economic Forecasting, Romanian Academy, Bucharest, Romania
 - 4 National Institute for Economic Research "Costin C. Kirițescu", Romanian Academy, Bucharest, Romania
- * Corresponding author: Mihaela Simionescu, mihaela.simionescu@unibuc.ro

Type of manuscript: Research paper

Cite as: Simionescu, M., & Nicula, A. S. (2024). Sentiment Analysis as Innovation in the Inflation Forecasting in Romania. *Marketing and Management of Innovations*, 15(2), 13–25. <https://doi.org/10.21272/mmi.2024.2-02>

Received: 10 January 2024
Revised: 13 May 2024
Accepted: 14 June 2024

Publisher & Founder: Sumy State University



Copyright: © 2024 by the authors. For open-access publication within the terms and conditions of the Creative Commons Attribution (CC BY) licence (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Romania faced the highest inflation rate in the European Union at the beginning of 2024, but progress has been made compared to that in 2023 due to the increasing interest rate. This inflation stemmed from a combination of global and domestic factors (global factors such as the Russia-Ukraine war, supply chain disruptions caused by the COVID-19 pandemic and war, rising commodity prices, domestic factors such as wage and pension increases, tax and charge hikes, and a strategy of gradual increase in the monetary policy interest rate). The National Bank of Romania (NBR) uses a combination of monetary policy instruments to target inflation and provides quarterly forecasts. However, under uncertain conditions, numerical forecasts are less reliable, and the inclusion of sentiment analysis in forecasts might lead to innovation in the field by improving the prediction accuracy. Sentiment analysis has become increasingly important in the field of economics, offering valuable insights and potentially improving economic forecasting and decision-making due to rapid technological progress. Sentiment analysis can identify potential changes in consumer behaviour and business decisions before they are translated into actual economic data, providing an early warning system for economic trends and potential crises. The methodological background relies on natural language processing to extract sentiment indices for large amounts of texts in Inflation Reports provided by NBR. Moreover, the sentiment indices calculated by IntelliDocker are incorporated into autoregressive distributed lag (ARDL) models to provide quarterly inflation forecasts. This type of econometric model has the advantage of addressing endogeneity. Moreover, the unemployment rate is considered an inflation predictor since tensions in the labour market might impact inflation. This paper contributes to empirical forecasting by proposing sentiment forecasts that are more accurate than NBR numerical forecasts corresponding to the 2006: Q1-2023: Q4 horizon. The new forecasting method might be used to make inflation predictions for the next quarters. More accurate forecasts would be valuable for businesses, the central bank, policymakers, and the general public. However, while sentiment analysis offers valuable insights, it is important to remember that human judgment and expertise remain essential for interpreting the data and making informed economic decisions.

Keywords: inflation; forecasts; sentiment index; natural language processing.

VĂ MULTUMESC !