

Lista de lucrări publicate

Maria Nuțu*

Publicații în cadrul conferințelor internaționale

- [1] **Maria Nuțu**. *Deep Learning Approach for Automatic Romanian Lemmatization*. In *2021 25th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES2021)*, Procedia Computer Science, Elsevier Publisher, vol. 192, pp. 49-58.
<https://www.sciencedirect.com/science/article/pii/S1877050921014939>
Rank B, 4 puncte.
- [2] Adriana Mihaela Coroiu, Alina Delia Călin, and **Maria Nuțu**. Communication Style - An Analysis from the Perspective of Automated Learning. In *Artificial Neural Networks and Machine Learning (ICANN)*, Cham Springer International Publishing, pp. 589-597, 2018 ISBN: 978-3-030-01418-6
https://link.springer.com/chapter/10.1007/978-3-030-01418-6_58
Rank B, 4 puncte.
- [3] Adriana Mihaela Coroiu, Alina Delia Călin, and **Maria Nuțu**. Topic Modeling in Medical Data Analysis. Case Study Based on Medical Records Analysis. In *2019 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, pp. 1-5, 2019.
<https://ieeexplore.ieee.org/document/8903900>
Rank B, 4 puncte.
- [4] Beáta Lőrincz, **Maria Nuțu**, Adriana Stan and Mircea Giurgiu. An Evaluation of Post-filtering for Deep Learning Based Speech Synthesis with Limited Data. In: *2020 IEEE 10th International Conference on Intelligent Systems (IS)*, pp. 437-442, 2020,
DOI:10.1109/IS48319.2020.9199932.
<https://ieeexplore.ieee.org/document/9199932>
Rank C, 1 punct.
- [5] **Maria Nuțu**, Beáta Lőrincz and Adriana Stan Deep Learning for Automatic Diacritics Restoration in Romanian. In *2019 IEEE 15th International Conference on Intelligent Computer Communication and Processing (ICCP)*, IEEE Computer Society, pp. 235-240, 2019.
<https://ieeexplore.ieee.org/document/8959557>
Rank C, 2 puncte.
- [6] Beáta Lőrincz, **Maria Nuțu**, and Adriana Stan “Romanian Part of Speech Tagging using LSTM Networks”. In *2019 IEEE 15th International Conference on Intelligent Computer*

*Clasificare articole folosind Computing Research and Education Association of Australasia, CORE Inc. website: <http://portal.core.edu.au/conf-ranks/>, valabilă în anul publicării fiecărui articol

Communication and Processing (ICCP), IEEE Computer Society, pp. 223–228, 2019.
<https://ieeexplore.ieee.org/document/8959730>

Rank C, 2 puncte.

7. [7] Adriana Stan, Beáta Lőrincz, **Maria Nuțu** and Mircea Giurgiu, "The MARA corpus: Expressivity in end-to-end TTS systems using synthesised speech data," 2021 International Conference on Speech Technology and Human-Computer Dialogue (SpeD), Bucharest, Romania, 2021, pp. 85-90
<https://ieeexplore.ieee.org/document/9587438>

Rank D, 0.5 puncte.

Punctaj total: 17.5 puncte

Citările lucrărilor științifice (sursa: Google Scholar)

- [1] **Maria Nuțu** Deep Learning Approach for Automatic Romanian Lemmatization. In *2021 25th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES2021)*, Procedia Computer Science, Elsevier Publisher, vol. 192, pp. 49-58.

Citări

1. [8] Pratama, Angga, Raksaka Indra Alhaqq, and Yova Ruldeviyani. "Sentiment Analysis Of The Covid-19 Booster Vaccination Program as a Requirement for Homecoming During Eid Fitr In Indonesia." *Journal Of Theoretical And Applied Information Technology* , vol.101, No.1, ISSN: 1817-3195 (2023).
 2. [9] Mouyassir, Kawtar, Abderrahmane Fathi, and Nouredine Assad. "Elevating Aspect-Based Sentiment Analysis in the Moroccan Cosmetics Industry with Transformer-based Models." *International Journal of Advanced Computer Science & Applications* 15.6 (2024).
 3. [10] Yoon, Liu Jun, et al. "A Comparative Study of Lemmatization Approaches for Rojak Language." *The International Conference on Data Science and Emerging Technologies*. Springer Nature Singapore, 2023.
 4. [11] Kawtar, Mouyassir, et al. "Hierarchical Spatiotemporal Aspect-Based Sentiment Analysis for Chain Restaurants using Machine Learning." *International Journal of Advanced Computer Science & Applications* 15.3 (2024).
- [7] Adriana Stan, Beáta Lőrincz, **Maria Nuțu** and Mircea Giurgiu, "The MARA corpus: Expressivity in end-to-end TTS systems using synthesised speech data," 2021 International Conference on Speech Technology and Human-Computer Dialogue (SpeD), Bucharest, Romania, 2021, pp. 85-90,

Citări

1. [12] Ungureanu, D., Badeanu, M., Marica, G. C., Dascalu, M., and Tufis, D. I. (2021, October). Establishing a Baseline of Romanian Speech-to-Text Models. In *2021 International Conference on Speech Technology and Human-Computer Dialogue (SpeD)* (pp. 132-138). IEEE.
2. [13] Beáta Lőrincz, Elena Irimia, Adriana Stan, and Verginica Barbu Mititelu. "RoLEX: The development of an extended Romanian lexical dataset and its evaluation at predicting concurrent lexical information." *Natural Language Engineering* (2022): 1-26.

3. [14] Gasan, Carol-Luca, and Păis, Vasile. "Investigation of Romanian speech recognition improvement by incorporating Italian speech data." *Linguistic resources and tools for natural language processing* (2023): 235.
4. [15] Stan, Adriana, and Johannah O'Mahony. "An analysis on the effects of speaker embedding choice in non auto-regressive TTS." arXiv preprint arXiv:2307.09898 (2023).
- [4] Beáta Lőrincz, **Maria Nuțu**, Adriana Stan and Mircea Giurgiu. An Evaluation of Post-filtering for Deep Learning Based Speech Synthesis with Limited Data. In: *2020 IEEE 10th International Conference on Intelligent Systems (IS)*, pp. 437–442, 202, DOI:10.1109/IS48319.2020.9199932

Citări:

1. [16] Eren, Eray, and Cenk Demiroglu. *Deep learning-based speaker-adaptive postfiltering with limited adaptation data for embedded text-to-speech synthesis systems*. *Computer Speech & Language* (2023): 101520.
2. [17] Beáta Lőrincz. Contributions to neural speech synthesis using limited data enhanced with lexical features. In *Proc. 2021 ISCA Symposium on Security and Privacy in Speech Communication* (pp. 83-85).
3. [18] Anas Fahad Khan et. al. *When Linguistics Meets Web Technologies. Recent advances in Modelling Linguistic Linked Open Data*. In *The journal Semantic Web – Interoperability, Usability, Applicability*, publisher IOS Press, ISSN: 1570-0844,
- [5] **Maria Nuțu**, Beáta Lőrincz and Adriana Stan. Deep Learning for Automatic Diacritics Restoration in Romanian. In *2019 IEEE 15th International Conference on Intelligent Computer Communication and Processing (ICCP)*, IEEE Computer Society, pp. 235–240, 2019.

Citări:

1. [19] Stankevičius, L., Lukoševičius, M., Kapočiūtė-Dzikienė, J., Briedienė, M., & Krilavičius, T. (2022). Correcting diacritics and typos with a ByT5 transformer model. *Applied Sciences*, 12(5), 2636.
2. [20] Pakalniškis, L. (2022). *Giliujoju mokymusi grįstas diakritinių ženklų atstatymas lietuvių kalbai* (Doctoral dissertation, Kauno technologijos universitetas).
3. [21] Stan, A., & Lőrincz, B. (2021). Generating the Voice of the Interactive Virtual Assistant. In *Virtual Assistant*. IntechOpen.
4. [22] Hifny, Y. (2021, June). Recent Advances in Arabic Syntactic Diacritics Restoration. In *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 7768-7772). IEEE.
5. [23] Náplava, J., Straka, M., & Straková, J. (2021). Diacritics Restoration using BERT with Analysis on Czech language.
6. [24] Esmail, S., Bar, K., & Dershowitz, N. (2021). How Much Does Lookahead Matter for Disambiguation? Partial Arabic Diacritization Case Study. (Master thesis) Tel Aviv University, Blavatnik School of Computer Science)
7. [25] Scott, K. M., Ashby, S., & Cibin, R. (2020, September). Implementing text-to-speech tools for community radio in remote regions of Romania. In *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers* (pp. 123-126).

8. [26] Al-Thubaity, A., Alkhalifa, A., Almuhareb, A., & Alsanie, W. (2020). Arabic diacritization using bidirectional long short-term memory neural networks with conditional random fields. *IEEE Access*, 8, 154984-154996.
 9. [27] Iordache, F., Georgescu, L., Oneată, D., & Cucu, H. (2019). Romanian Automatic Diacritics Restoration Challenge. In *Proceedings of the 14th international conference Linguistic resources and tools for natural language processing* (pp. 64-74).
 10. [28] Ogheneruemu, Kingsley Lucky Ogheneovo. Development of Yoruba Diacritic Restoration for Under Dot and Diacritic Mark for Yoruba Text Using Deep Learning Model. MS thesis. Kwara State University (Nigeria), 2022.
 11. [29] Özge, Asiye Tuba, Özge Bozal, and Umut Özge. "Diacritics correction in Turkish with context-aware sequence to sequence modeling." *Turkish Journal of Electrical Engineering and Computer Sciences* 30.6 (2022): 2433-2445.
 12. [30] Esmail, Saeed, Kfir Bar, and Nachum Dershowitz. "How much does lookahead matter for disambiguation? partial arabic diacritization case study." *Computational Linguistics* 48.4 (2022): 1103-1123.
 13. [22] Hifny, Yasser. "Recent advances in Arabic syntactic diacritics restoration." *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021.
- [6] Beáta Lőrincz, **Maria Nuțu**, and Adriana Stan "Romanian Part of Speech Tagging using LSTM Networks". In *2019 IEEE 15th International Conference on Intelligent Computer Communication and Processing (ICCP)*, IEEE Computer Society, pp. 223–228, 2019.
 1. [31] Shafahat Sardarov. *Development and Design of Deep Learning-based Parts-of-Speech Tagging System for Azerbaijani language*, Thesis for Master of Science in Engineering in Computer Science, 2022, Khazar University, Azerbaijan
 2. [32] Josipa Juričić. *Označavanje vrsta riječi pomoću neuronskih mreža*. Master thesis, University of Split, Faculty of Science. Department of Informatics, 2022.
 3. [33] Harjanto, Shadifa Auliatama, and Ade Romadhony. "Question Template Extraction Using Sequence Labeling Approach." *2024 International Conference on Data Science and Its Applications (ICoDSA)*. IEEE, 2024.
 4. [34] Aydinov, Farhad, et al. "Investigation of automatic part-of-speech tagging using CRF, HMM and LSTM on misspelled and edited texts." *Proceedings of the 2022 5th artificial intelligence and cloud computing conference*. 2022.
 5. [35] Juričić, Josipa, and Branko Žitko. "POS-Only Tagging Using RNN for Croatian Language." *International Conference on Digital Transformation in Education and Artificial Intelligence Application*. Springer Nature Switzerland, 2023.
 - [3] Adriana Mihaela Coroiu, Alina Delia Călin, and **Maria Nuțu**. Topic Modeling in Medical Data Analysis. Case Study Based on Medical Records Analysis. In *2019 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, pp. 1–5, 2019.
- Citări:**
1. [36] Gupta, Aditi, and Hoor Fatima. "Topic Modeling in Healthcare: A Survey Study." *NEUROQUANTOLOGY* 20.11 (2022): 6214-6221.

2. [37] Kenei, J., Opiyo, E., & Machii, J. Modeling and visualization of clinical texts to enhance meaningful and user-friendly information re-trieval. In *Med. Sci. Forum*, Vol. 1, February, 2022.

Bibliografie

- [1] M. Nuțu, “Deep learning approach for automatic romanian lemmatization,” *Procedia Computer Science*, vol. 192, pp. 49–58, 2021, Knowledge-Based and Intelligent Information and Engineering Systems: Proceedings of the 25th International Conference KES2021, ISSN: 1877-0509. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1877050921014939>.
- [2] A. M. Coroiu, A. D. Călin, and M. Nuțu, “Communication style - an analysis from the perspective of automated learning,” in *Artificial Neural Networks and Machine Learning – ICANN 2018*, V. Kůrková, Y. Manolopoulos, B. Hammer, L. Iliadis, and I. Maglogiannis, Eds., Cham: Springer International Publishing, 2018, pp. 589–597, ISBN: 978-3-030-01418-6.
- [3] A. M. Coroiu, A. D. Călin, and M. Nuțu, “Topic modeling in medical data analysis. case study based on medical records analysis,” in *2019 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, 2019, pp. 1–5. DOI: 10.23919/SOFTCOM.2019.8903900.
- [4] B. Lőrincz, M. Nuțu, A. Stan, and M. Giurgiu, “An evaluation of postfiltering for deep learning based speech synthesis with limited data,” in *2020 IEEE 10th International Conference on Intelligent Systems (IS)*, 2020, pp. 437–442. DOI: 10.1109/IS48319.2020.9199932.
- [5] M. Nuțu, B. Lőrincz, and A. Stan, “Deep learning for automatic diacritics restoration in romanian,” in *2019 IEEE 15th International Conference on Intelligent Computer Communication and Processing (ICCP)*, 2019, pp. 235–240. DOI: 10.1109/ICCP48234.2019.8959557.
- [6] B. Lőrincz, M. Nuțu, and A. Stan, “Romanian part of speech tagging using lstm networks,” in *2019 IEEE 15th International Conference on Intelligent Computer Communication and Processing (ICCP)*, IEEE, 2019, pp. 223–228.
- [7] A. Stan, B. Lőrincz, M. Nuțu, and M. Giurgiu, “The mara corpus: Expressivity in end-to-end tts systems using synthesised speech data,” in *2021 International Conference on Speech Technology and Human-Computer Dialogue (SpeD)*, 2021, pp. 85–90. DOI: 10.1109/SpeD53181.2021.9587438.
- [8] A. Pratama, R. I. Alhaqq, and Y. Ruldeviyani, “Sentiment analysis of the covid-19 booster vaccination program as a requirement for homecoming during eid fitr in indonesia,” *Journal of Theoretical and Applied Information Technology*, vol. 101, no. 1, 2023.
- [9] K. Mouyassir, A. Fathi, and N. Assad, “Elevating aspect-based sentiment analysis in the moroccan cosmetics industry with transformer-based models,” *International Journal of Advanced Computer Science & Applications*, vol. 15, no. 6, 2024.
- [10] L. J. Yoon, X. Y. Tan, K. Y. Lim, C. W. Tan, L. E. Cheng, and J. Tan, “A comparative study of lemmatization approaches for rojak language,” in *The International Conference on Data Science and Emerging Technologies*, Springer, 2023, pp. 3–16.
- [11] M. Kawtar, A. Fathi, N. Assad, and A. Kartit, “Hierarchical spatiotemporal aspect-based sentiment analysis for chain restaurants using machine learning,” *International Journal of Advanced Computer Science & Applications*, vol. 15, no. 3, 2024.

- [12] D. Ungureanu, M. Badeanu, G.-C. Marica, M. Dascalu, and D. I. Tufis, “Establishing a baseline of romanian speech-to-text models,” in *2021 International Conference on Speech Technology and Human-Computer Dialogue (SpeD)*, IEEE, 2021, pp. 132–138.
- [13] B. Lőrincz, E. Irimia, A. Stan, and V. B. Mititelu, “Rolex: The development of an extended romanian lexical dataset and its evaluation at predicting concurrent lexical information,” *Natural Language Engineering*, pp. 1–26, 2022.
- [14] C.-L. Gasan and V. PĂIȘ, “Investigation of romanian speech recognition improvement by incorporating italian speech data,” *LINGUISTIC RESOURCES AND TOOLS FOR NATURAL LANGUAGE PROCESSING*, p. 235, 2023.
- [15] A. Stan and J. O’Mahony, “An analysis on the effects of speaker embedding choice in non auto-regressive tts,” *arXiv preprint arXiv:2307.09898*, 2023.
- [16] E. Eren and C. Demiroglu, “Deep learning-based speaker-adaptive postfiltering with limited adaptation data for embedded text-to-speech synthesis systems,” *Computer Speech and Language*, p. 101 520, 2023.
- [17] B. Lorincz, “Contributions to neural speech synthesis using limited data enhanced with lexical features,” in *Proc. 2021 ISCA Symposium on Security and Privacy in Speech Communication*, 2021, pp. 83–85.
- [18] A. F. Khan, C. Chiarcos, T. Declerck, D. Gifu, E. G.-B. Garcia, J. Gracia, M. Ionov, P. Labropoulou, F. Mambrini, J. P. McCrae, *et al.*, “When linguistics meets web technologies. recent advances in modelling linguistic linked open data,”
- [19] L. Stankevičius, M. Lukoševičius, J. Kapočiūtė-Dzikienė, M. Briedienė, and T. Krilavičius, “Correcting diacritics and typos with a byt5 transformer model,” *Applied Sciences*, vol. 12, no. 5, p. 2636, 2022.
- [20] L. Pakalniškis, “Giliujoju mokymusi grįstas diakritinių ženklų atstatymas lietuvių kalbai,” Ph.D. dissertation, Kauno technologijos universitetas, 2022.
- [21] A. Stan and B. Lőrincz, “Generating the voice of the interactive virtual assistant,” in *Virtual Assistant*, IntechOpen, 2021.
- [22] Y. Hifny, “Recent advances in arabic syntactic diacritics restoration,” in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2021, pp. 7768–7772.
- [23] J. Náplava, M. Straka, and J. Straková, “Diacritics restoration using bert with analysis on czech language,” *arXiv preprint arXiv:2105.11408*, 2021.
- [24] S. Esmail, K. Bar, and N. Dershowitz, “How much does lookahead matter for disambiguation? partial arabic diacritization case study,” 2021.
- [25] K. M. Scott, S. Ashby, and R. Cibin, “Implementing text-to-speech tools for community radio in remote regions of romania,” in *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers*, 2020, pp. 123–126.
- [26] A. Al-Thubaity, A. Alkhalifa, A. Almuhareb, and W. Alsanie, “Arabic diacritization using bidirectional long short-term memory neural networks with conditional random fields,” *IEEE Access*, vol. 8, pp. 154 984–154 996, 2020.
- [27] F. IORDACHE, L. GEORGESCU, D. ONEAȚĂ, and H. CUCU, “Romanian automatic diacritics restoration challenge,” in *Proceedings of the 14th international conference “linguistic resources and tools for natural language processing*, 2019, pp. 64–74.

- [28] K. L. O. Ogheneruemu, “Development of yoruba diacritic restoration for under dot and diacritic mark for yoruba text using deep learning model,” M.S. thesis, Kwara State University (Nigeria), 2022.
- [29] A. T. Özge, Ö. Bozal, and U. Özge, “Diacritics correction in turkish with context-aware sequence to sequence modeling,” *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 30, no. 6, pp. 2433–2445, 2022.
- [30] S. Esmail, K. Bar, and N. Dershowitz, “How much does lookahead matter for disambiguation? partial arabic diacritization case study,” *Computational Linguistics*, vol. 48, no. 4, pp. 1103–1123, 2022.
- [31] S. Sardarov, “Development and design of deep learning-based parts-of-speech tagging system for azerbaijani language,” Ph.D. dissertation, Khazar University, Azerbaijan, 2022.
- [32] J. Juričić, “Označavanje vrsta riječi pomoću neuronskih mreža,” Ph.D. dissertation, University of Split. Faculty of Science. Department of Informatics, 2022.
- [33] S. A. Harjanto and A. Romadhony, “Question template extraction using sequence labeling approach,” in *2024 International Conference on Data Science and Its Applications (ICoDSA)*, IEEE, 2024, pp. 242–247.
- [34] F. Aydinov, I. Huseynov, S. Sayadzada, and S. Rustamov, “Investigation of automatic part-of-speech tagging using crf, hmm and lstm on misspelled and edited texts,” in *Proceedings of the 2022 5th artificial intelligence and cloud computing conference*, 2022, pp. 21–28.
- [35] J. Juričić and B. Žitko, “Pos-only tagging using rnn for croatian language,” in *International Conference on Digital Transformation in Education and Artificial Intelligence Application*, Springer, 2023, pp. 45–62.
- [36] A. Gupta and H. Fatima, “Topic modeling in healthcare: A survey study,” *NEUROQUANTOLOGY*, vol. 20, no. 11, pp. 6214–6221, 2022.
- [37] J. Kenei, E. Opiyo, and J. Machii, “Modeling and visualization of clinical texts to enhance meaningful and user-friendly information re-trieval,” in *Med. Sci. Forum*, The 2nd International Electronic Conference on Healthcare, vol. 1, 2022.