

Named Entity Recognition based Resume Parser and Summarizer

Narendra G O and Hashwanth S

School of Computer Science and Engineering

Vellore Institute of Technology, Chennai, Tamil Nadu, India

Abstract: Resumes are unstructured documents and the language used in the resume will be ambiguous and variable. Screening documents in large amounts is a difficult operation, and recruiters and hiring managers waste a lot of time going through each and every CV of the candidates. The extraction of valuable information from a resume is difficult due to this variability. It necessitates an urgent need to comprehend the context in which words are used. Using spaCy, a powerful natural language processing (NLP) package, this study suggests a machine learning method to phrase matching in resumes, focused on the extraction of specific entities. Like a human recruiter, it can examine and extract comprehensive information from resumes. While parsing, it maintains track of the terms in order to classify people. Specific Important entities are extracted from the resume document and then it is stored for later classification. Document ranking is processed and the final rankings are depicted in descending order with respect to the score of the resume in order to make it easy for the recruiters to shortlist the resumes quickly and efficiently. Recruiters may pick the needed applicants based on the scores rather than sifting through stacks of resumes from unqualified prospects.

Keywords: NER, Resume, Ranking, spaCy

REFERENCES

- [1]. Xu, M., Jiang, H., & Watcharawittayakul, S. (2017). A Local Detection Approach for Named Entity Recognition and Mention Detection. *ACL*.
- [2]. Crichton, G., Pyysalo, S., Chiu, B. et al. A neural network multi-task learning approach to biomedical named entity recognition. *BMC Bioinformatics* 18, 368 (2017).
- [3]. E. S. Chifu, V. R. Chifu, I. Popa and I. Salomie, "A system for detecting professional skills from resumes written in natural language," 2017 13th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP), 2017, pp. 189-196, doi: 10.1109/ICCP.2017.8117003.
- [4]. Mourad Gridach, Character-level neural network for biomedical named entity recognition, *Journal of Biomedical Informatics*, Volume 70, 2017, Pages 85-91, ISSN 1532-0464,
- [5]. Maryam Habibi, Leon Weber, Mariana Neves, David Luis Wiegandt, Ulf Leser, Deep learning with word embeddings improves biomedical named entity recognition, *Bioinformatics*, Volume 33, Issue 14, 15 July 2017, Pages i37–i48, <https://doi.org/10.1093/bioinformatics/btx228>
- [6]. Lin, B., Xu, F.F., Luo, Z., & Zhu, K.Q. (2017). Multi-channel BiLSTM-CRF Model for Emerging Named Entity Recognition in Social Media. *NUT@EMNLP*.
- [7]. Dernoncourt, Franck & Lee, Ji & Szolovits, Peter. (2017). NeuroNER: an easy-to-use program for named-entity recognition based on neural networks.
- [8]. Sanyal, Satyaki & Hazra, Souvik & Ghosh, Neelanjan & Adhikary, Soumyashree. (2017). Resume Parser with Natural Language Processing. [10.13140/RG.2.2.11709.05607](https://doi.org/10.13140/RG.2.2.11709.05607).
- [9]. Abd, Maan & Mohd, Masnizah. (2018). A comparative study of word representation methods with conditional random fields and maximum entropy markov for bio-named entity recognition. *Malaysian Journal of Computer Science*. 31. 15-30. [10.22452/mjcs.sp2018no1.2](https://doi.org/10.22452/mjcs.sp2018no1.2).
- [10]. Riedl, M., & Padó, S. (2018). A Named Entity Recognition Shootout for German. *ACL*.

- [11]. G. Popovski, B. K. Seljak and T. Eftimov, "A Survey of Named-Entity Recognition Methods for Food Information Extraction," in IEEE Access, vol. 8, pp. 31586-31594, 2020, doi: 10.1109/ACCESS.2020.2973502.
- [12]. Perera, N., Dehmer, M., & Emmert-Streib, F. (2020). Named Entity Recognition and Relation Detection for Biomedical Information Extraction. *Frontiers in cell and developmental biology*, 8, 673. <https://doi.org/10.3389/fcell.2020.00673>
- [13]. Jain, Rachna & Sharma, Abhishek & Mishra, Gouri & Nand, Parma & Chakraborty, Sudeshna. (2020). Named Entity Recognition in English Text. *Journal of Physics: Conference Series*. 1712. 012013. 10.1088/1742-6596/1712/1/012013.
- [14]. Sarath, A., & Sreedevi, D. M. An Effective Resume Data Extraction with Advance Algorithms.
- [15]. Zhang, Q., Fu, J., Liu, X., & Huang, X. (2018). Adaptive Co-attention Network for Named Entity Recognition in Tweets. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1). Retrieved from <https://ojs.aaai.org/index.php/AAAI/article/view/11962>