References

1. Abhisek Chakrabarty Onkar Arun Pandit and Utpal Garain. 2017. Context Sensitive Lemmatization Using Two Successive Bidirectional Gated Recurrent Networks. Conference: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1).

2. Abhisek Chakrabarty, Akshay Chaturvedi and Utpal Garain. 2019. CNN-based Context Sensitive Lemmatization. Publication Proceedings of the ACM India Joint International Conference on Data Science and Management of Data. Pages 334–337.

3. Adamson, G.W., Boreham, J. 1974. The use of an association measure based on character structure to identify semantically related pairs of words and document titles. Information Storage and Retrieval. Vol 10, pp 253-260

4. Anjali G. Jivani. 2011. A Comparative Study of Stemming Algorithm**s.** IJCTA | ISSN: 2229-6093.

5. Anjali Ganesh Jivani, Rupam Gupta. 2014. Empirical Analysis of Affix Removal Stemmers: et al, Int.J.Computer Technology &Applications, Vol 5 (2), 393-399 IJCTA.

6. Bart Jongejan, Hercules Dalianis. 2-7 August 2009. Automatic training of lemmatization rules that handle morphological Proceedings of the 47th Annual Meeting of the ACL and the 4th IJCNLP of the AFNLP, pages 145–153, Suntec, Singapore.

7. Benno Stein and Martin Potthast. 2007. Putting Successor Variety Stemming to Work. Advances in Data Analysis pp 367–374.

8. C. Douglas Johnson. 1972. Formal Aspects of Phonological Description. Mouton.

9. C.D. Manning and H. Schütze. 1999. Foundations of Statistical Natural Language Processing. The MIT, Press.

10. Carmen Galvez, Fe´lix de Moya-Anego´n and Vı´ctor H. Solana. 2005. In stemming and lemmatization processes, two term conflation approaches: non-linguistic and linguistic: Department of Information Science, University of Granada, Granada, Spain.Journal of Documentation.Vol. 61 No. 4. pp. 520-547.

11. Chapin, Paul. 1967. On The Syntax of word-derivation in English, MIT Press.

12. Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, David McClosky. June, 2014. Linguistics & Computer Science: Stanford University, The Stanford CoreNLP Natural Language Processing Toolkit

13. D.R. Morrison. October 1968. "PATRICIA-Practical Algorithm To Retrieve Information Coded in Alphanumeric," J. ACM, vol. 15, no. 4, pp. 514

14. Daniel Jurafsky & James H. Martin. Copyright c 2007. WORDS & TRANSDUCERS: Speech and Language Processing: An introduction to speech recognition, natural language processing, and computational linguistics.

15. Dawson, J. 1974. "Suffix Removal and Word Conflation." ALLC Bulletin, Michelmas, 33-46.

16. Dimitrios P. Lyras, Kyriakos N. Sgarbas, Nikolaos D. Fakotakis. 2007. Using the Levenshtein Edit Distance for Automatic Lemmatization: A Case Study for Modern Greek and English: 19th IEEE International Conference on Tools with Artificial Intelligence(ICTAI 2007), Vol-2. Pp.428-435.

17. Douglas W. Oard, Gina-Anne Levow, and Clara. Cabezas. 2001. CLEF Experiments at Maryland: Statistical Stemming and Backoff Translation. Print ISBN: 978-3-540-42446-8.

18. Eric Gaussier. 1999. Unsupervised learning of derivational morphology from inflectional lexicons. Xerox research Center Europe 6, chemin de Maupertuis 38240 Meylan F.

19. Fazli Can & Seyit Kocberber & Erman Balcik & Cihan Kaynak & H. Cagdas Ocalan & Onur M. Vursavas, February 2008. Information Retrieval on Turkish Texts: 'Information retrieval on Turkish texts.' Journal of the American Society for for Information Science and Technology. Vol. 59, No. 3, pp. 407-421.

20. Frakes William B. 2003. Strength and similarity of affix removal stemming algorithms. ACM SIGIR Forum, Volume 37, No. 1. 26-30.

21. Francis Katamba. 1994. English Words: by Routledge: ISBN 0-415-10468-8 (pbk) .

22. Funchun Peng, Nawaaz Ahmed, Xin Li and Yumao Lu. 2007. Context sensitive stemming for web search. Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval., 639-646.

23. Haibin Liu, Tom Christiansen. 2012. BioLemmatizer: a lemmatization tool for morphological processing of biomedical text. Published in J. Biomedical Semantics. Computer Science, Medicine Journal of Biomedical Semantics, DOI: 10.1186/2041-1480-3-3. Corpus ID: 15618342.

24. Harman Donna. 1991. "How effective is suffixing?" Journal of the American Society for Information Science. 42, 7-15 7.

25. Harris Z.S. 1955. Distributional Structure. In: Papers in Structural and Transformational Linguistics. Formal Linguistics Series. Springer, ISBN978-94-017-6059-1.

26. Hull D. A. and Grefenstette,. January 31, 1996. A detailed analysis of English Stemming Algorithms, XEROX Technical Report, http://www.xrce.xerox.

27. J. B. Lovins. 1968. Development of a stemming algorithm," Mechanical Translation and Computer Linguistic., vol.11, no.1/2, pp. 22-31.

28. Jiaul H. Paik, Mandar Mitra, and Swapan K. Parui. . 2011. GRAS: An Effective and Efficient Stemming Algorithm for Information Retrieval. ACM Transactions on Information Systems, Vol. 29, No. 4, Article 19.

29. Jiaul H. Paik, Swapan K. Parui (ISI, Kolkata). July 2011. A Novel Corpus-Based Stemming Algorithm using Co-occurrence Statistics: Publication: SIGIR '11: Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval Pages : 863–87

30. John Goldsmith. 2001. Linguistica 5: Unsupervised Learning of Linguistic Structure.

31. John Goldsmith. 2001. Unsupervised Learning of the Morphology of a Natural Language: Study Report : Computational Linguistic: Issue 2, pp153-198.

32. Jorge Morato, Miguel Ángel Marzal, Juan Lloréns, and José Moreiro. 2004. WordNet Applications: Petr Sojka, Karel Pala, Pavel Smrž, Christiane Fellbaum, Piek Vossen (Eds.): GWC, Proceedings, pp. 270–278.

33. Juan-Manuel Torres-Moreno: 2012. Beyond Stemming and Lemmatization. Ultra-stemming to Improve Automatic Text Summarization,: Laboratoire Informatique d'Avignon, BP 91228 84911, Avignon, Cedex 09,

France.

34. Kimmo Koskenniemi. 1984. Two-level morphology: A general computational model for word-form recognition and production: Publications/Department of General Linguistics, University of Helsinki, ISBN-10: 9514532015.

35. Krister Lindén. 2008. A Probabilistic Model for Guessing Base Forms of New Words by Analogy. Springer, Berlin, Heidelberg. Online ISBN: 978-3-540-78135-6.

36. Krovetz Robert. July 1993 "Viewing morphology as an inference process". Proceedings of the 16th annual international ACM SIGIR conference on Research.

37. LAURI KARTTUNEN AND KENNETH R. BEESLEY. July 13, 2005. Twenty-five years of finite-state morphology: CSLI Publications, pp.71-83.

38. M. Hafer And Stephen F. Weiss. 1974. Word Segmentation by Letter Successor Varieties:Pergamon Press. Printed in Great Britain: Vol. 10, pp. 371-385.

39. M.P. Schützenberger. 1961. A remark on finite transducers. Information and Control, 4:185–187.

40. Mark Aronoff and Kirsten Fudeman. 25 October 2010. What is Morphology? ISBN: 978-1-405-19467- 9.

41. Matjaž Juršič, Igor Mozetič, Tomaž Erjavec, Nada Lavrač. 1214, 2010. LemmaGen: Multilingual Lemmatisation with Induced Ripple-Down Rules: Journal of Universal Computer Science, vol. 16, no. 9, 1190-   .

42. Michela Bacchin *, Nicola Ferro, Massimo Melucci. 2005. A probabilistic model for stemmer generation: Department of Information Engineering, University of Padova, Italy, Information Processing and Management 41 121–137 123

43. Milan Straka and Jana Strakova´.2017. Tokenizing, POS Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe, Charles University, Faculty of Mathematics and Physics, Institute of Formal and Applied Linguistics, CoNLL.

44. Morris Halle. 1973. Prolegomena to a Theory of Word Formation. Linguistic Inquiry Volume 4

45. NOAM CHOMSKY, MORRIS HALLE. Winter, I973.Harper & Row, Publishers, New York, Evanston, and London. Published by: The MIT Press Stable URL: ttp://www.jstor.org/stable/4177749 Accessed: 14-04-2018 01:41 UTC

46. Paice Chris D. 1990. "Another stemmer". ACM SIGIR Forum, Volume 24, No. 3. 56-61.

47. Paice Chris D. 1994. An evaluation method for stemming algorithms". Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval 42-50.

48. Porter M.F. 1980. "An algorithm for suffix stripping". Program: Electronic Library and Information System. ISSN: 0033-0337, Vol. 14, 130-137.

49. Porter M.F. 2001. "Snowball: A language for stemming algorithms". Computer Science.

50. Prasenjit Majumder, Mandar Mitra, Kalyan Kumar Datt. 2006. Statistical vs. Rule-Based Stemming for Monolingual French Retrieval Adhoc Monolingual Information Retrieval.

51. Prasenjit Majumder, Mandar Mitra, Swapan K. Parui, and Govinda Kole. 2007. YASS: Yet another suffix stripper. Article in ACM Transactions on Information Systems.

52. R. Baeza-Yates and B. Ribeiro-Neto. 1999. Modern Information Retrieval. Addison Wesley.

53. Richard Wicentowski. 2002. Modeling and Learning Multilingual Inflectional Morphology in a Minimally Supervised Framework. Baltimore, Maryland.

54. Ronald M. Kaplan* & Martin Kay. 1980. Regular Models of Phonological Rule Systems. Xerox Palo Alto Research Center and Stanford University.

55. Rupam Gupta, and Anjali G. Jivani. 2020. A Lemmatizer: International Journal on Emerging Technologies. ISSN No. (Print): 0975-8364, 11(2): 817-824.

56. Rupam Gupta, Anjali G. Jivani. 2017. Analyzing the Stemming Paradigm: Information and Communication Technology for Intelligent Systems (ICTIS 2017) - Volume 2 pp 333-342.

57. Sylvain Neuvel. 2002. Whole Word Morphologizer: Expanding the WordBased Lexicon: University of Chicago. A Nonstochastic Computational. Brain and Language 81, 454–463.

58. Toms Bergmanis and Sharon Goldwater. 2018. Context Sensitive Neural Lemmatization with Lematus. Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers) P: 1391–1400

59. W. B. Frakes. October 2010. CHAPTER 8: STEMMING ALGORITHMS: Software Engineering Guild, Sterling, VA 22170, Wiley-Blackwell.

60. Xu Jinxi and Croft Bruce W. 1998. "Corpus-based stemming using co-occurrence of word variants". ACM Transactions on Information Systems. Volume 16, Issue 1. 1998, 61-81.

61. Y.-L. Yeong, T.-P. Tan, and S. K. Mohammad. 2016. Using dictionary and lemmatizer to improve low resource English-Malay statistical machine translation system, Pro- cedia. Comput. Sci., Vol. 81, pp. 243–249.

URLs

1. http://morphadorner.northwestern.edu/morphadorner
2. http://textanalysisonline.com/spacy-word-lemmatize
3. http://ufal.mff.cuni.cz/udpipe
4. http://www-nlpir.nist.gov/projects/duc/data.html
5. https://cst.dk/tools/index.php?lang=en
6. https://devopedia.org/stemming
7. https://duc.nist.gov/data.html
8. https://en.wikipedia.org/wiki/Christopher_D._Paice
9. https://en.wikipedia.org/wiki/Finite-state_transducer
10. https://en.wikipedia.org/wiki/Information_extraction
11. https://en.wikipedia.org/wiki/Information_retrieval
12. https://en.wikipedia.org/wiki/Julie_Beth_Lovins

13. https://en.wikipedia.org/wiki/Morphological_derivation

14. https://en.wikipedia.org/wiki/Natural_language_processing

15. https://github.com/hugomailhot/MorphoLex-en/blob/master/MorphoLEX_en.xlsx

16. https://nlp.stanford.edu/IR-book/html/htmledition/stemming-and-lemmatization-1.html

17. https://usefulenglish.ru/writing/list-of-derivative-adjectives

18. https://usefulenglish.ru/writing/list-of-nouns-with-suffixes

19. https://web.stanford.edu/~jurafsky/slp3/slides/Chapter18_introandsimilarity.pdf

20. https://web.stanford.edu/~jurafsky/slp3/slides/Chapter18_introandsimilarity.pdf

21. https://web.stanford.edu/~laurik/publications/ciaa-2000/fst-in-nlp/fst-in-nlp.html

22. https://wordnet.princeton.edu/

23. https://www.kaggle.com/datasets/nltkdata/brown-corpus

24. https://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html

25. https://www.merriam-webster.com/browse/thesaurus/

26. https://www.oxfordlearnersdictionaries.com/wordlist/american_english/oxford3000/

_____