

## ПЕРЕЛІК ПОСИЛАНЬ

1. Олійник Ю.О. Аналіз методів визначення ваг ознак текстових документів. [Текст] / Ю.О. Олійник, Д.О. Катющенко // Журнал науковий огляд № 3(46), 2018 – С. 112–123.
2. Катющенко Д. Дослідження Методу Максимуму Ентропії Для Класифікації Текстів На Прикладі Apache OpenNlp Doccat [Текст]: / Д. Катющенко, Ю. Олійник // Математичне та імітаційне моделювання систем. МОДС '2017" – 2017.
3. Катющенко Д. Дослідження методів попередньої обробки та векторизації текстових документів [Текст]: / Д. Катющенко, Ю. Олійник //Матеріали наукової конференції студентів, магістрантів та аспірантів «Інформатика та обчислювальна техніка» – ІОТ-2018, 24– м. Київ: С. 193-195
4. Ханько Г. В. Огляд алгоритмів текстового аналізу даних[Текст]: / Г.В. Ханько // “Інформатика та обчислювальна техніка” – ІОТ-2017– Київ: 2017. С. 17 – 20.
5. Батура Т.В. Основы обработки текстовой информации [Текст]: / Т.В.Батура, М.В. Чаринцева // Учебное пособие –Институт систем информатики им. А.П. Ершова СО РАН– Новосибирск: 2016– С.45.
6. Козлов П.Ю. Способы представления текстовой информации при автоматизированном рубрицировании коротких текстовых документов [Текст]: / П.Ю. Козлов // Программные продукты и системы / Software&SystemsТ.30. №4, 2017 – С. 678–683.
7. Файловый архив студентов StudFiles [Електронний ресурс] // Режим доступу: <https://studfiles.net/preview/4533558/page:12/>

8. Joachims T. “Learning to Classify Text Using Support Vector Machines: Methods, Theory and Algorithms” [Текст]: / Т. Joachims // Kluwer Academic Publishers Norwell, MA, USA: 2002.

9. Батура Т.В. Методы автоматической классификации текстов [Текст]: / Т.В. Батуреждународный научно-практический журнал программные продукты и системы №1 : 2017. С. 85 – 99.

10. Воронцов К.В. Лекции по линейным алгоритмам классификации [Электронный ресурс] // Режим доступа: <http://www.machinelearning.ru/wiki/images/6/68/voron-ML-Lin.pdf>

11. Оценка классификатора (точность, полнота, F-мера) [Электронный ресурс] / Суровая реальность – 2012. – Режим доступа: <http://bazhenov.me/blog/2012/07/21/classification-performance-evaluation.html>

12. Ballard D.H. Computer Vision [Текст]: / D.H. Ballard, C.M. Brown C.M. // Prentice Hall Inc., 1982. – 539p.

13. Mikolov T. Distributed Representations of Words and Phrases and their Compositionality [Текст]: / Т. Mikolov, I. Sutskever, K. Chen, G. Corrado, J. Dean // Proceeding NIPS 13 Proceedings of the 26th International Conference on Neural Information Processing Systems – Volume 2. – 2013. – pp. 3111-3119.

14. Mikolov T. Linguistic Regularities in Continuous Space Word Representations [Текст]: / Т. Mikolov, W. Yih, G. Zweig // Proceedings of NAACL-HLT 2013 – Atlanta, Georgia – 2013. pp.746–751.

15. Zhang W. A comparative study of TF\*IDF, LSI and multi-words for text classification [Текст]: / W. Zhang, T. Yoshida, X. Tang // Expert Systems with Applications - Volume 38 Issue 3 - 2011, pp. 2758-2765.

16. Нугуманова А.Б. Обогащение модели BAG-OF-WORDS семантическими связями для повышения качества классификации текстов предметной области [Текст] : / А.Б. Нугуманова, И.А. Бессмертный, П.

Пецина, Е.М. Байбурин // “Програмные продукты и системы”. 2016. №2 (114). С. 89 – 99.

17. Sahlgren M. Using bag-of-concept to improve the performance of support vector machines in text categorization [Текст]: / M. Sahlgren, R. Coster // Proc. of the 20th Int. Conf. on Computational Linguistics, Association for Computational Linguistics. 2004. p. 487.

18. Huang A. Clustering documents using a Wikipedia-based concept representation [Текст]: / A. Huang // Advances in Knowledge Discovery and Data Mining. SpringerBerlin: Heidelbergpubl. 2009. pp 628-636.

19. Shehata S. An efficient concept-based mining model for enhancing text clustering. [Текст]: / S. Shehata, F. Karray, M. Kamel // IEEE Transactions on Knowledge and Data Engineering. 2010, vol. 22, #10. pp.1360-1371.

20. Egozi O. Concept-based indexing text information retrieval [Текст]: / O. Egozi, S. Markovitch, E. Gabrilovich // ACM Transactions on Information System (TOIS). 2011, vol.29 #2. p 8.

21. Boubekour F. Concept-based indexing in text information retrieval [Текст]: / F. Boubekour, W. Azzoug // Int. J. Comput. Sci. Inf. Technol. 2013, vol. 5 pp. 119-136.

22. Huang L. Learning a concept-based document similarity measure. [Текст]: / L. Huang // Journ. of the American Society for Information Science and Technology. 2012, vol. 63, #8. pp.1593-1608.

23. Wang P. Building semantic kernels for text classification using Wikipedia [Текст]: / P. Wang, C. Domeniconi // Proc. of the 14th ACM SIGKDD Intern. Conf. on Knowledge discovery and data mining: ACM Publ., 2008, pp.713-721.

24. Turdakov D.Y. Word sense disambiguation methods. [Текст]: / D.Y. Turdakov // Programming and Computer Software. 2010, vol. 36, #6, pp.309-326.

25. Dinh D. Combining global and local semantic contexts for improving biomedical information retrieval. [Текст]: / D. Dihn, L. Tamine // Advances in Information Retrieval – Berlin: Heidelberg Publ. – 2011. pp. 375-386.

26. Carrillo M. Concept Based Representations as complement of Bag Of Words in Information Retrieval [Текст]: / M. Carrillo, A. Lopez-Lopez // Artificial Intelligence Applications and Innovations – Berlin: Heidelberg Publ. – 2010. pp. 154-161.

27. Deerwester S.C. Indexing by latent semantic analysis [Текст]: / S.C. Deerwester, S.T. Dumais, T.K. Landauer, G.W. Furnas, R.A. Harshman // JAsls - 1990, no.1, pp 91 - 102.

28. Ju R. An Efficient Method for Document Categorization Based on Word2vec and Latent Semantic Analysis. [Текст]: / R. Ju // — 2015 IEEE Int. Conf. on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Auto-nomic and Secure Computing; Pervasive Intelligence and Computing. Liverpool, UK, 2015, pp. 2276–2283.

29. Agirre E. A study on similarity and relatedness using distributional and WordNet-based approaches. [Текст]: / E. Agirre // Proc. of Human Language Technologies: the 2009 Annual Conf. of the North American Chapter of the Association for Computational Linguistics. – ACL Publ. – 2009. pp. 19 – 27.

30. Mashechkin I. Automatic text summarization using latent semantic analysis. [Текст]: / I. Mashechkin, M. Petrovsky, D. Popov, D. Tsarev // Programming and Computer Software. – 2011, vol. 37, no. 6, pp. 299-305.

31. Mashechkin I. Methods of text fragment relevance estimation based on the topic model analysis in the text summarization problem. [Текст]: / I. Mashechkin, M. Petrovsky, D. Tsarev // Vychislitelnye metody i programmirovaniye [Computing Methods and programming]. - 2013, vol. 14, no. 1, pp. 91-102.

32. Xiang Zhang Character-level convolutional networks for text classification. [Текст]: / Junbo Zhao, Yann LeCun. // Proc. Neural Inform. Processing Systems Conf. (NIPS 2015). Montreal, Canada - 2015.

33. Victor S. Miller Variations on a theme by Ziv and Lempel, Combinatorial algorithms on words. [Текст]: / Victor S. Miller, Mark N. Wegman – Springer Berlin Heidelberg, 1985. – pp. 131-140.

34. Гасфилд Д. Строки, деревья и последовательности в алгоритмах: Информатика и вычислительная биология [Текст]: / Д. Гасфилд Пер. с англ. И. В. Романовского. — 2-е изд. — СПб.: Невский Диалект, 2003. — 654 с.

35. Окулов С. М. Алгоритмы обработки строк. — М.: Бином, 2013. — 255с.

36. Ricardo Baeza-Yates Modern information retrieval [Текст]: / Ricardo Baeza-Yates, Berthier Ribeiro-Neto. — Reading, Massachusetts: Addison-Wesley Longman, 1999. — 192 с.

37. Попков М.И. Автоматическая система классификации текстов для базы знаний предприятия: магистерская диссертация — М: 2014. - 57 с.

38. Victor S. Miller Variations on a theme by Ziv and Lempel, Combinatorial algorithms on words. [Текст]: / Victor S. Miller, Mark N. Wegman – Springer Berlin Heidelberg, 1985. – pp. 131-140.

39. Corder G. W. Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach. [Текст]: / G. W. Corder, D. I. Foreman// — New York: Wiley, 2009.

40. Yiming Yang A comparative study of of feature selection in text categorization [Текст]: / Yiming Yang Jan O. Pedersen // ICML '97 Proceedings of the fourteenth international conference on machine learning - San Francisco, CA, USA, 1997 — pp. 412-420.

41. Zhang C. Automatic Keyword Extraction from Documents Using Conditional Random Fields [Текст]: / C. Zhang, H. Wang, Y. Liu, D. Wu, Y. Liao, B. Wang // *Journal of Computational Information Systems* 4:3. 2008. pp. 1169-1180.

42. Sharma A. R. Literature survey of statistical, deep and reinforcement learning in natural language processing Sign In or Purchase [Текст]: / A. R. Sharma, P. Kaushik // *2017 International Conference on Computing, Communication and Automation (ICCCA)*, Greater Noida, 2017, pp. 350-354.

43. Ghahramani Z. An introduction to hidden Markov models and Bayesian networks. [Текст]: / Z. Ghahramani // *International Journal of Pattern Recognition and Artificial Intelligence*, 15(01). 2001. pp. 9–42.

44. Manning R.R. Introduction. [Текст]: R.R. Manning (eds) *Retrieving the Radical Tillich. Radical Theologies*. Palgrave Macmillan— NewYork— 2015. pp. 1-19.

45. Suthaharan S. Support Vector Machine. In: *Machine Learning Models and Algorithms for Big Data Classification*. [Текст]: / S. Suthaharan // *Integrated Series in Information Systems*, vol 36. Springer — Boston, MA — 2016. pp. 207-235.

46. Гришанов К.М. Метод классификации  $k$  - nn и его применение в распознавании символов [Текст]: / К.М. Гришанов, Ю.С. Белов // *Фундаментальные проблемы науки*. — Тюмень: НИЦ АЭТЕРНА — 2016. 30-33 с.

47. Manning D. Christopher Introduction to information retrieval [Текст]: / Christopher D. Manning, Prabhakar Raghavan // Cambridge University Press - 2008.

48. Maimon Oded Z. *Data Mining With Decision Trees: Theory And Applications (2nd Edition)* [Текст]: / Maimon Oded Z, RokachLior — 2014. 304 p.

49. Смирнова О.С. Выбор топологии нейронных сетей и их применение для классификации коротких текстов [Текст]:/ О.С. Смирнова, В.В. Шишков. International Journal of Open Information Technologies, vol. 4.//– 2016, pp. 50–54.

50. Гавриленко О. В. Огляд та аналіз алгоритмів TEXT MINING / О.В. Гавриленко, Ю.О. Олійник, Г. В. Ханько. // Управління проектами, системний аналіз і логістика. – К.: НТУ, 2017. – Вип.

51. Apache™ OpenNLP [Електронний ресурс] // Режим доступу: <http://opennlp.apache.org/>

52. uClassify [Електронний ресурс] // Режим доступу: <https://www.uclassify.com/>

53. Ленкин А.В. Исследование систем для TextMining [Текст]: / А.В. Ленкин, Р.И. Баженов // ISSN 2414-4487. Постулат. №1. – 2017 –с. 326-340

54. Joachims T. “Learning to Classify Text Using Support Vector Machines: Methods, Theory and Algorithms” [Текст]: / Т. Joachims // Kluwer Academic Publishers Norwell, MA, USA. –2002.

55. Рысаков С.В. Статистические методы снятия омонимии [Текст]: / С.В. Рысаков, С.Э. Клышинский // Новые информационные технологии в автоматизированных системах. – 2015. – С. 555 – 563.

56. Зеленков Ю.Г. Вероятностная модель снятия морфологической омонимии на основе нормализующих подстановок и позиций соседних слов [Електронний ресурс] // режим доступу: [http://www.dialog-21.ru/media/2444/zelenkov\\_segalovich.pdf](http://www.dialog-21.ru/media/2444/zelenkov_segalovich.pdf)

57. Моніторинговий проект “ОКО” [Електронний ресурс] / Режим доступу: <http://www.ukroko.org>

58. Електронне врядування «Розумне місто» [Електронний ресурс] / Режим доступу: <https://rozzumnemisto.org>

59. Мова програмування Java [Електронний ресурс] / Режим доступу:  
<https://java.com/ru/>

60. Платформа для створення графічного інтерфейсу додатків JavaFX [Електронний ресурс] / Режим доступу:  
<http://www.oracle.com/technetwork/java/javafx/>

61. Model View Controller (MVC) [Електронний ресурс]/ Режим доступу:  
<https://ru.wikipedia.org/wiki/Model-View-Controller>

62. Інтегроване середовище розробки Eclipse [Електронний ресурс] /  
Режим доступу: <http://www.eclipse.org/>

63. JavaFX SceneBuilder [Електронний ресурс] / Режим  
доступу:<https://o7planning.org/en/10621/install-javafx-scene-builder-into-eclipse>