

uses the KNN algorithm which used to classify objects based on the basis of their similarities or distance metric. KNN classifiers are based on learning by analogy. On the basis of stylistics features they find out the result in the analytical manner. The dichotomy data for the Stylometry authentication experiments contained 1770 records for each subset of six subjects. Each subset was run against the other yielding 76.72% and 66.72% accuracy. The author [14] faces some difficulties and their future work is to extend the authentication task to identify patterns in frequently used misspelled and misused words.

5. Conclusion And Future Work

Through the overall discussion the paper we discussed first the basic behind the stylometry then later in the discussion move to the literature review where we have discussed about stylometry that can be used for the identification and authentication of the author in different fields like Author identification; detection of hoaxes, frauds, and deception in writing styles; gender identification from emails, plagiarism detection etc.. We have also analyse the result on the basis of the stylometric features for the scientific articles and email author identification. So in this manner, we just see that the stylometry can be used in many broad areas. A still lot of research has to be done in field of author identification but we have chosen to implement it for the security of email by identifying the author and with this the security of the email system will be improved.

6. References

- [1] David I. Holmes, "*The Evolution of Stylometry in Humanities Scholarship*," Literary and Linguistic Computing 13/3: Pages: 111-117, 1998.
- [2] T. C. Mendenhall, "*The Characteristic Curves of Composition*," Science 214, Pages : 237-246, 1887.
- [3] C. Mascol, "*Curves of Pauline and Pseudo-Pauline Style I*," Unitarian Review, 30 Pages: 452-60, 1888; C. Mascol, "*Curves of Pauline and Pseudo-Pauline Style II*," Unitarian Review, 30 Pages: 539-46, 1888.
- [4] L. A. Sherman, "*Analytics of Literature: A Manual for the Objective Study of English Prose and Poetry*", Boston: Ginn, 1893.
- [5] Stefan Gruber, and Stuart Noven, "*Tool support for plagiarism detection in text documents*", Symposium on Applied Computing archive Proceedings of the 2005 ACM , Pages: 776 – 781, 2005.
- [6] D. Pavelec, L. S. Oliveira, E. Justino, F. D. Nobre Neto, and L. V. Bastista, "*Compression and Stylometry for Author Identification*", Proceedings of International Joint Conference on Neural Networks, 2009.
- [7] Burrows, J. F., "Computers and the Study of Literature",. In: C. S. Butler (ed.): Computers and Written Texts. Oxford: Blackwell, Pages: 167-204, 1992.
- [8] H. Maurer, F. K., "*Plagiarism - a survey*", Journal of Universal Computer Science, vol. 12, no. 8 , Pages: 1050-1084, 2006.
- [9] Shane Bergsma, Matt Post, David Yarowsky, "*Stylometric Analysis of Scientific Articles*", 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies , Pages: 327-337, 2012.
- [10] Daniel Pavelec, Edson Justino, and Luis S.Oliveira, "*Author Identification using Stylometric Features*", Inteligencia Artificial, Revista Iberoamericana de Inteligencia Artificial. Vol 11, No 36, Pages: 59-65, 2007.
- [11] Na Cheng, Xiaoling Chen, R. Chandramouli, K. P. Subbalakshmi, "*Gender Identification from E-mails*", Computational Intelligence and Data Mining , Pages: 154-158, 2009.
- [12] Michael Brennan, Rachel Greenstadt, "*Practical Attacks Against Authorship Recognition Techniques*", Innovative Applications of Artificial Intelligence (IAAI), 2009.
- [13] Sadia Afroz, Michal Brennan and Rachel Greenstadt, "*Detecting Hoaxes, Frauds, and Deception in Writing Style Online*", IEEE Symposium on Security and Privacy, 2012 .
- [14] K. Calix, M. Connors, D. Levy, H. Manzar, G. McCabe, and S. Westcott, "*Stylometry for E-mail Author Identification and Authentication*" Proceedings of CSIS Research Day, Pace University, May 2008.