

DAFTAR PUSTAKA

- Ahuja, Y. & Yadav, K.S., 2012. Multiclass Classification and Support Vector Machine. *Global Journal of Computer Science and Technology Interdisciplinary*, 12(11), pp.14–19. Available at: https://globaljournals.org/GJCST_Volume12/2-Multiclass-Classification-and.pdf.
- Ali, S. & Smith, K.A., 2003. Automatic Parameter Selection For Polynomial Kernel. *Proceedings of the 2003 IEEE International Conference on Information Reuse and Integration, IRI 2003*, pp.243–249.
- Anderson, E., 1936. The Species Problem in Iris. *Annals of the Missouri Botanical Garden*, 23(3), p.457. Available at: <http://www.jstor.org/stable/2394164?origin=crossref>.
- Anton, H. & Rorres, C., 2010. *Elementary Linear Algebra; Applications* Version. 10th ed., New York : John Wiley & Sons Inc.
- Banerjee, A., Merugu, S., Dhillon, I. S. & Ghosh, J., 2005. Clustering with Bregman Divergences. *Journal of Machine Learning Research*, 6, pp.1705–1749.
- Bapat, R.B., 2000. *Linear Algebra and Linear Models*, Second Edition. New York: Springer.
- Bellman, R.E., 1954. The Theory of Dynamic Programming. *Bulletin of the American Mathematical Society*, 60(6), pp.503–515.
- Ben-David, S., Simon, H.U. & Pal, D., 2007. Stability of K-Means Clustering. *Conference on Learning Theory*, pp.20–34.
- Ben-Hur, A., Horn, D., Siegelmann, H.T. & Vapnik, V., 2001. Support Vector Clustering. *Journal of Machine Learning Research* 2, 2, pp.125–137.
- Ben-Hur, A., Elisseeff, A. & Guyon, I., 2002. A Stability Based Method For Discovering Structure In Clustered Data. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, 17(February 2002), pp.6–17.
- Ben-Israel, A. & Levin, Y., 2006. The Geometry Of Linear Separability In Data Sets. *Linear Algebra and Its Applications*, 416(1), pp.75–87.
- Bezdek, J.C., 1981. *Pattern Recognition with Fuzzy Objective Function Algorithm*. New York : Plenum Press.
- Bezdek, J.C., Ehrlich, R. & Full, W., 1984. FCM: The Fuzzy C-Means Clustering Algorithm. *Computers & Geosciences*, 10(2–3), pp.191–203. Available at: <https://www.sciencedirect.com/science/article/pii/0098300484900207?via%3Dihub>
- Bhosale, D. & Ade, R., 2014. Feature Selection Based Classification Using Naive Bayes, J48 and Support Vector Machine. *International Journal of Computer Applications* (0975-8887). Volume 99–No.16, August 2014, 99(16), pp.14–18.
- Borovinskiy, V., 2009. Classification of Iris Data Set. Available at : http://lab.fs.uni-lj.si/lasin/wp/IMIT_files/neural/doc/seminar8.pdf
- Bronson, R. & Costa, G.B., 2009. *Matrix Methods; Applied Linear Algebra* 3rd ed., New York : Elsevier Inc.

- Brown, J.D., 2009. Questions and Answers About Language Testing Statistics : Choosing the Right Number of Components or Factors in PCA and EFA Choosing the Number of Components or Factors to Include in a PCA or EFA Examples Illustrating the Five Stopping Rules. *JALT Testing & Evaluation SIG*, 13(May), pp.19–23.
- Chakraborty, R.C., 2010. Fundamentals of Neural Networks Soft Computing. Available at: http://www.myreaders.info/02-Fundamentals_of_Neural_Network.pdf
- Chang, Q., Chen, Q. & Wang, X., 2005. Scaling Gaussian RBF Kernel Width To Improve SVM Classification. *2005 International Conference on Neural Networks and Brain*, 1, pp.19–22. Available at: http://ieeexplore.ieee.org.gate.lib.buffalo.edu/xpls/abs_all.jsp?arnumber=1614559%5Cnhttp://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1614559.
- Cherney, D., Denton, T. & Waldron, A., 2012. *Linear Algebra*, California: Davis.
- Chiang, M.M.T. & Mirkin, B., 2010. Intelligent Choice Of The Number Of Clusters In K-Means Clustering: An Experimental Study With Different Cluster Spreads. *Journal of Classification*, 27(1), pp.3–40.
- Cover, T.M., 1965. Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition. *IEEE Transactions on Electronic Computers*, EC-14(3), pp.326–334.
- Cristianini, N., 2001. Support Vector and Kernel Machines. *Technologies, Biowulf*, p.85.
- Davies, D.L. & Bouldin, D.W., 1979. A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-1(2), pp.224–227.
- Deborah, J. L., Baskaran, R. & Kannan, A., 2010. A Survey on Internal Validity Measure for Cluster Validation. *International Journal of Computer Science & Engineering Survey*, 1(2), pp.85–102.
- Deepa, M. & Revathy, P., 2012. Validation of Document Clustering Based On Purity And Entropy Measures. *International Journal of Advanced Research in Computer and Communication Engineering*, 1(3), pp.147–152.
- Dhillon, I.S., 2004. Kernel K-Means, Spectral Clustering and Normalized Cuts. *Compute*, CI(78712), pp.551–556.
- Dillon, D.R. & Matthew, G., 1984. *Multivariate Analysis; Methods and Applications*, New York: John Wiley & Sons Inc.
- Dunn, J.C., 1974. A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters. *Journal of Cybernetics*, 3(3), pp.32–57.
- Fisher, R.A., 1936. The Use Of Multiple Measurements In Taxonomic Problems. *Annual Eugenics*, 7, Part II, pp.179–188.
- Gentle, J.E., 2007. *Matrix Algebra: Theory, Computations, And Applications In Statistics*, Available at: http://books.google.com/books?hl=en&lr=&id=rQg2sgPo2gQC&oi=fnd&pg=PR7&dq=Springer+exts+in+Statistics&ots=23Txmqw8eo&sig=TYcfX23MnVw8VTddpT_XCf0_8VA.

- Guestrin, C., 2007. Linear separability Addressing Non-Linearly Separable Data-Option 1, Non-Linear Features, pp.1–26. Available at : <https://pdfs.semanticscholar.org/presentation/0866/da2957b61b3881e26373cb0f34e684ef0385.pdf>
- Halkidi, M., Batistakis, Y. & Vazirgiannis, M., 2001. On Clustering Validation Techniques. *Journal of Intelligent Information Systems*, 17(2–3), pp.107–145.
- Han, J. & Kamber, M., 2006. *Data Mining, Concept and Techniques* 2nd ed., San Francisco: Morgan Kaufmann Publishers.
- Hand, D.J. & Yu, K., 2001. Idiot’s Bayes--Not So Stupid After All? *International Statistical Review*, 69(3), pp.385–398.
- Hautamaki, V., Cherednichenko, S., Karkkainen, I., Kinnunen, T. & Franti, P., 2005. Improving K-Means by Outlier Removal. *Scia 2005, Lncs 3540*, pp.978–987.
- Haykin, S., 1999. *Neural Networks - A Comprehensive Foundation - Simon Haykin*. 2nd ed., Singapore : Pearson Education, Inc.
- Hoffmann, H., 2007. Kernel PCA For Novelty Detection. *Patt Recogn*, 40(3), pp.863–874.
- Hubert, L. & Arabie, P., 1985. Comparing Partition. *Journal of classification*, 218, pp.193–194.
- Jain, a. K., Murty, M.N. & Flynn, P.J., 1999. Data Clustering: A Review. *ACM Computing Surveys*, 31(3), pp.264–323. Available at: <http://portal.acm.org/citation.cfm?doid=331499.331504>.
- Jain, A.K., 2009. Data Clustering 50 Years Beyond K-Means. Michigan State University. *19th International Conference in Pattern Recognition (ICPR)*, pp.651–666.
- Johnson, R.A. & Wichern, D.W., 2007. *Applied Multivariate Statistical Analysis*. New York : Pearson Education, Inc.
- Jolliffe, I.T., 2002. Principal Component Analysis, Second Edition. *Encyclopedia of Statistics in Behavioral Science*, 30(3), p.487. Available at: <http://onlinelibrary.wiley.com/doi/10.1002/0470013192.bsa501/full>.
- Kreyszig, E., 1978. *Introductory Functional Analysis With Applications*, New York: John Wiley & Sons.
- Kullback, S. & Leibler, R.A., 1951. On Information and Sufficiency Author (s) : S . Kullback and R . A . Leibler Source : The Annals of Mathematical Statistics , Vol . 22 , No . 1 (Mar ., 1951), pp . 79-86 Published by : Institute of Mathematical Statistics Stable URL : <http://www.jstor.> , 22(1), pp.79–86.
- Lang, S., 1986. *Introduction to Linear Algebra* 2nd ed., New York: Springer.
- Larsen, B. & Aone, C., 1999. Fast And Effective Text Mining Using Linear-Time Document Clustering. *Proceedings of the Fifth ACM SIGKDD International Conference On Knowledge Discovery and Data Mining-KDD '99*, pp.16–22.
- Lee, S. H. & Daniels, K., 2005. Gaussian Kernel Width Generator for Support Vector Clustering. *Advances in Bioinformatics and Its Applications- Proceedings of the International Conference*, pp.151–162. Available at: http://eproceedings.worldscinet.com/9789812702098/9789812702098_0014.html.

- Lee, S. & Lee, W., 2013. A Study of Initial Seeds Methods Comparison for K-Means Algorithm, *Proceedings, The 4th International Conference on Security-enriched Urban Computing and Smart Grid*.ASTL.Vol.26, pp.68-71
- Lee, Y. & Choi, S., 2007. Minimum Entropy, K-Means, Spectral Clustering. *Biometrics Technology Research*,(1).
- Lee, Y., Lin, Y. & Wahba, G., 2004. Multicategory Support Vector Machines: Theory And Application To The Classification Of Microarray Data And Satellite Radiance Data. *Journal of the American Statistical Association*, 99(465), pp.67–81.
- Li, T. & Ding, C., 2008. Weighted Consensus Clustering. *Proceedings of the Eighth SIAM International Conference on Data Mining*, pp.798–809. Available at: https://users.cs.fiu.edu/~taoli/tenure/dm08_72_li.pdf.
- Liu, T., Liu, S., Chen, Z., & Ma, W., 2003. An Evaluation On Feature Selection For Text Clustering. *Icml*, pp.488–495. Available at: <http://www.aaai.org/Papers/ICML/2003/ICML03-065.pdf>.
- Liu, Y., Li, Z., Xiong, H., Gao, X. & Wu, J. 2010. Understanding Of Internal Clustering Validation Measures. *IEEE Internatinal Conference on Data mining*, pp.911–916.
- Liu, Z., Chen, D. & Bensmail, H., 2005. Gene Expression Data Classification With Kernel Principal Component Analysis. *Journal of Biomedicine and Biotechnology*, 2005(2), pp.155–159.
- Luxburg, U. Von, 2010. Clustering Stability: An Overview. *Foundations and Trends in Machine Learning*, 2(3), pp.235–274. Available at: https://books.google.com/books?hl=de&lr=&id=Qwf_YAmuyFUC&pgis=1.
- MacQueen, J.B., 1967. K-Means Some Methods for Classification and Analysis of Multivariate Observations. *5th Berkeley Symposium on Mathematical Statistics and Probability 1967*, 1(233), pp.281–297. Available at: <http://projecteuclid.org/euclid.bsmsp/1200512992>.
- Mardia, K. V., Kent, J.T. & Bibby, J.M., 1979. *Multivariate Analysis*, London: Academic Press Limited.
- Margineantu, D.D. & Dietterich, T.G., 2003. Learning Decision Trees For Loss Minimization In Multi-Class Problems. *Citeseer*, 99-30-03, pp.1–14. Available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.7.1088&rep=rep1&type=pdf>.
- McLachlan, G.J. & Khan, N., 2004. On A Resampling Approach For Tests On The Number Of Clusters With Mixture Model-Based Clustering Of Tissue Samples. *Journal of Multivariate Analysis*, 90(1 SPEC. ISS.), pp.90–105.
- Melnykov, I. & Melnykov, V., 2014. On K-Means Algorithm With The Use Of Mahalanobis Distances. *Statistics and Probability Letters*, 84(1), pp.88–95. Available at: <http://dx.doi.org/10.1016/j.spl.2013.09.026>.
- Mewada, A., Gedam, P., Khan, S. & Reddy, M.U., 2010. Network Intrusion Detection using Semi Supervised Support Vector Machine,1(2,3,4), *Special Issue of IJCTT Vol.1 Issue 2,3,4; 2010 for International Conference [ACCTA-2010]*, 3-5 August 2010.pp.172–175.
- Mirkin, B., 2005. *Clustering for A Data Recovery Approach*, Boca Raton: Chapman & Hall/CRC.

- Mohamad, I. Bin & Usman, D., 2013. Standardization and its Effects on K-Means Clustering Algorithm. *Research Journal of Applied Sciences, Engineering and Technology*, 6(17), pp.3299–3303.
- Mufti, G.B., Bertrand, P. & Moubarki, L. El, 2005. Determining The Number Of Groups From Measures Of Cluster Stability. *Proceedings of International Symposium on Applied Stochastic Models and Data Analysis*, pp.404–413. Available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.98.4941&rep=rep1&type=pdf>.
- Nazeer, K.A.A. & Sebastian, M.P., 2009. Improving the Accuracy and Efficiency of the K-Means Clustering Algorithm, I, *Proceedings of the World Congress on Engineering 2009 Vol I.WCE 2009*, July 1-3, 2009, London, U.K. pp.1–5.
- Ott, R.L. & Longnecker, M., 2010. *An Introduction To Statistical Methods And Data Analysis - sixth edition*. New York : Duxbury, Thomson Learning.
- Park, H., Lee, J. & Jun, C., 2008. A K-Means-like Algorithm for K-Medoids Clustering and Its Performance. *System*, (C), pp.2005–2008. Available at: <https://pdfs.semanticscholar.org/11fb/cfedfdc701e563a8888df6cd8de8c7652b69.pdf>
- Parrado-Hernández, E., Mora-Jimenez, J., Arenas-Garcia, J., Figueiras-Vidal, A.R. & Navia_Vazquez, A., 2003. Growing Support Vector Classifiers With Controlled Complexity. *Pattern Recognition*, 36(7), pp.1479–1488.
- Pitoy, C., 2017. Empirical Study On Bandwidth Optimization For Kernel PCA In The K-Means Clustering Of Non-Linearly Separable Data, *Journal of Theoretical and Applied Information Technology (JATIT)*. 95(20), pp.5464–5472.
- Raftery, A., 1986. A Note On Bayes Factors For Log-Linear Contingency Table Models With Vague Prior Information. *Journal of the Royal Statistical Society, Series B*, 48(2), pp.249–250.
- Rai, P., 2011a. Data Clustering : K-Means and Hierarchical Clustering What is Data Clustering ? *Machine Learning*, 2011, pp.1–24.
- Rai, P., 2011b. Kernel Methods and Nonlinear Classification slides, 2011. Available at: <https://www.cs.utah.edu/~piyush/teaching/15-9-slides.pdf>.
- Rand, W., 1971. Objective Criteria for the Evaluation of Clustering Methods Author (s): William M . Rand Source : Journal of the American Statistical Association , Vol . 66 , No . 336 (Dec ., 1971), pp . 846- Published by : American Statistical Association Stable URL. , 66(336), pp.846–850.
- Rencher, A.C., 2002. *Methods of Multivariate Analysis*. New York : John Wiley & Sons, Inc. Publication.
- Rendón, E., Abundez, I., Alejandra, A. & Quiroz. E.M., 2011. Internal Versus External Cluster Validation Indexes. *International Journal of Computers and Communications*, 5(1), pp.27–34. Available at: <http://w.naun.org/multimedia/UPress/cc/20-463.pdf>.
- Ross, S., 2009. *Introduction To Probability And Statistics For Engineers And Scientists* 5 th., San Diego : Elsevier Academic Press.
- Russell, S. & Norvig, P., 2010. *Artificial Intelligence: A Modern Approach* 3rd ed. New Jersey : Prentice Hall Series.
- Rynne, B.P. & Youngson, M.A., 2008. *Linear Functional Analysis*, London: Springer.

- Sakthi, M. & Thanamani, A.S., 2011. An Effective Determination of Initial Centroids in K-Means Clustering Using Kernel PCA, 2(3), pp.955–959. Available at: http://www.ijcsit.com/docs/Volume_2/vol2issue3/ijcsit2011020303.pdf.
- Salman, R., Kecman, K., Li, Q., Strack, R. & Test. E., 2011. Fast K-Means Algorithm Clustering, 3(4), *International Journal of Computer Networks & Communications (IJCNC)* Vol.3, No.4, July 2011pp.17–31.
- Santini, 2016. Advantages & Disadvantages of K-Means and Hierarchical Clustering (Unsupervised Learning), Available at: http://stp.lingfil.uu.se/~santinim/ml/2016/Lect_10/10c_UnsupervisedMethods.pdf.pp.1–5.
- Sathiya, G. & Kavitha, P., 2014. An Efficient Enhanced K-Means Approach With Improved Initial Cluster Centers. *Middle - East Journal of Scientific Research*, 20(4), pp.485–491.
- Scholkopf, B., Sung, K., Burges, C., Girosi, F., Niyogi, P., Poggio, T. & Vapnik, V., 1997. Comparing Support Vector Machines With Gaussian Kernels To Radial Basis Function Classifiers. *IEEE Transactions on Signal Processing*, 45(11), pp.2758–2765.
- Scholkopf, B., Williamson, R., Smola, A., Shawe-Taylor, J., & Platt, J., 1999. Support Vector Method for Novelty Detection. MIT Press (2000). *Nips*, 12, pp.582–588.
- Scholkopf, B., Smola, a J. & Muller, K.R., 2012. Kernel Principal Component Analysis. *Computer Vision And Mathematical Methods In Medical And Biomedical Image Analysis*, 1327, pp.583–588. Available at: http://www.tribesandclimatechange.org/docs/tribes_450.pdf%5Cnpapers2://publication/uuid/D79004C6-EBB3-42CE-9F15-120270D985BE.
- Shannon, C.E., 1948. A Mathematical Theory Of Communication. *The Bell System Technical Journal*, 27(July 1928), pp.379–423.
- Shawe-Taylor, J. & Cristianini, N., 2005. *Kernel Methods for Pattern Analysis*. New York: Cambridge University Press.
- Shimodaira, H., 2015. Clustering and Visualisation of Data. , pp.3–6. Available At : <http://www.inf.ed.ac.uk/teaching/courses/inf2b/learnnotes/inf2b-learn-note03-2up.pdf>
- Silverman, B., 1986. Density Estimation For Statistics And Data Analysis. *Chapman and Hall*, 37(1), pp.1–22. Available at: http://books.google.com/books?hl=en&lr=&id=e-xsrjsL7WkC&oi=fnd&pg=PR9&dq=Density+Estimation+for+Statistics+and+Data+Analysis&ots=ivSonp7D_q&sig=XfZFIEyzmSO4nm54dgq22EiW9iA.
- Singh, N. & Singh, D., 2013. The Improved K-Means with Particle Swarm Optimization. *Journal of Information Engineering and Applications*, 3(11), pp.1–8.
- Smith, L.I., 2002. A Tutorial On Principal Components Analysis Introduction. *Statistics*, 51, p.52.
- Soliman, M.A.K.M.A. & Abo-Bakr, R.M., 2011. Linearly And Quadratically Separable Classifiers Using Adaptive Approach. *Journal of Computer Science and Technology*, 26(5), pp.908–918.

- Souza, C., 2010. Kernel Functions for Machine Learning Applications. Available at: <http://crsouza.com/2010/03/kernel-functions-for-machine-learning-applications/>.
- Sripada, S.C., 2011. Comparison of Purity and Entropy of K-Means Clustering and Fuzzy C Means Clustering. *Indian Journal of Computer Science and Engineering*, 2(3), pp.343–346. Available at: <http://www.ijcse.com/docs/IJCSE11-02-03-105.pdf>.
- Steinbach, M., Karypis, G. & Kumar, V., 2000. A Comparison of Document Clustering Techniques. *KDD Workshop On Text Mining*, 400(X), pp.1–2. Available at: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4721382>.
- Strehl, A. & Ghosh, J., 2002. Cluster Ensembles – A Knowledge Reuse Framework for Combining Multiple Partitions. *Journal of Machine Learning Research*, 3, pp.583–617.
- Tan, I.N.G., Steinbach, M. & Kumar, V., 2005. *Introduction to Data Mining*, Boston : PearsonEducation, Inc.
- Theodoridis, S. & Koutroumbas, K., 2003. *Pattern Recognition 2nded.*, New York: Elsevier.
- Thomas, M., De Brabanter, K. & De Moor, B., 2014. New Bandwidth Selection Criterion for Kernel PCA: Approach to Dimensionality Reduction and Classification Problems. *BMC Bioinformatics*, 15, p.137. Available at: <http://www.biomedcentral.com/1471-2105/15/137>.
- Thorstensen, N., Segonne, F. & Kerivan, R., 2008. Normalization And Preimage Problem In Gaussian Kernel PCA. Available at: <https://hal-enpc.archives-ouvertes.fr/hal-00835106/document>
- Tzortzis, G. & Likas, A., 2014. The MinMax k-Means Clustering Algorithm. *Pattern Recognition*, 47(7), pp.2505–2516.
- UCI Machine Learning Repository: Iris Data Set. 1988. Available At : <https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.names>
- Vapnik, V., 1995. *The Nature of Statistical Learning*. New York: Springer Verlag.
- Vimal, A., Valluri, S.R. & Karlapalem, K., 2008. An Experiment with Distance Measures for Clustering *. *Matrix*, pp.5–8.
- Weiss, G.M. & Hirsh, H., 1998. Learning to Predict Rare Events in Event Sequences. *Proceedings of the 4th International Conference on Knowledge Discovery and Data Mining*, (January 1998), pp.359–363.
- Wu, J., 2012. Advances in K-Means Clustering. *Advances in K-Means Clustering: A Data Mining Thinking*, pp.155–175. Available at: <http://link.springer.com/book/10.1007/978-3-642-29807-3>.
- Wu, J., Xiong, H. & Chen, J., 2009. Adapting the Right Measures for K-Means Clustering. *Proceedings of the 15th ACM SIGKDD International Conference On Knowledge Discovery And Data Mining - KDD '09*, p.877. Available at: <http://portal.acm.org/citation.cfm?doid=1557019.1557115>.
- Wu, X. & Kumar, V. (Ed), 2008. *Top Ten Algorithms in Data Mining*. New York: Taylor & Francis Group, LLC
- Yudi Agusta, 2007. K-Means – Penerapan, Permasalahan dan Metode Terkait. *Jurnal Sistem dan Informatika*, 3(Pebruari), pp.47–60.



- Zhang, B., Kleyner, G. & Hsu, M., 1999. A Local Search Approach To K-Clustering. *Hp Laboratories Technical Report Hpl*, (119). Available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.26.7606&rep=rep1&type=pdf>.
- Zhang, C. & Xia, S., 2009. K-Means Clustering Algorithm With Improved Initial Center. *2009 Second International Workshop on Knowledge Discovery and Data Mining*, 1(2), pp.790–792.
- Zhao, Q., 2012. *Cluster Validity in Clustering Methods*, Academic Dissertation. Available at: <http://cs.joensuu.fi/sipu/pub/qinpei-thesis.pdf>
- Zhao, Y. & Karypis, G., 2001. Technical Report Criterion Functions for Document Clustering: Experiments and Analysis, pp.1–30. . Available at: <https://pdfs.semanticscholar.org/a40f/670c985e45e973337284fcb4fd44fe5858fb.pdf>.
- Zhao, Y. & Karypis, G., 2002. Evaluation Of Hierarchical Clustering Algorithms For Document Datasets. *Proceedings Of The Eleventh International Conference On Information And Knowledge Management*, pp.515–524. Available at: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Evaluation+of+hierarchical+clustering+algorithms+for+document+datasets#0>.