

BAB VI

KESIMPULAN DAN SARAN

6.1 Kesimpulan

Berdasarkan hasil eksperimen yang meliputi pengujian preprocessing dengan klasifikasi serta analisis hasil klasifikasi yang telah dilakukan, maka dapat disimpulkan beberapa hal sebagai berikut :

- a. *Preprocessing* pada dataset wine, iris, wisconsin, pima dan haberman dapat dilakukan dengan metode *iterative partitioning filter*. Dalam melakukan *preprocessing* berbagai data ini, data diolah menggunakan software keel. Selanjutnya untuk melakukan validasi pada preprocessing data ini digunakan klasifikasi menggunakan *Backpropagation* pada software weka. Hasil eksperimen menunjukkan bahwa skema filter majority lebih meningkatkan performansi akurasi hasil klasifikasi dibandingkan skema konsensus.
- b. Analisis hasil *preprocessing* dapat dilakukan dengan melihat pengaruh perubahan parameter pada metode *iterative partitioning filter* terhadap hasil akurasi klasifikasi, *Kappa static*, *MAE*, *Root Mean Squared error* dan *confusion matrix*. Analisis juga dapat dilakukan dengan melihat grafik-grafik akurasi hasil klasifikasi pada masing-masing tabel, dan juga grafik sebelum data di *preprocessing* dengan *Iterative Partitioning Filter* dan sesudah di *preprocessing* dengan *Iterative Partitioning Filter*. Hasil akurasi pada dataset wine sebelum di

preprocessing sebesar 98.361% sedangkan setelah *dipreprocessing* dengan *Iterative Partitioning Filter* hasil akurasi menjadi 100%, untuk akurasi pada dataset iris sebelum di *preprocessing* sebesar 96.078% sedangkan setelah *dipreprocessing* dengan *Iterative Partitioning Filter* hasil akurasi menjadi 100%, lalu akurasi pada dataset wisconsin sebelum di *preprocessing* sebesar 94.118% sedangkan setelah *dipreprocessing* dengan *Iterative Partitioning Filter* hasil akurasi menjadi 99.524%, kemudian untuk akurasi pada dataset pima sebelum di *preprocessing* sebesar 79.694% sedangkan setelah *dipreprocessing* dengan *Iterative Partitioning Filter* hasil akurasi menjadi 95.313%, dan untuk akurasi pada dataset haberman sebelum di *preprocessing* sebesar 72.115% sedangkan setelah *dipreprocessing* dengan *Iterative Partitioning Filter* hasil akurasi menjadi 100%.

6.2 Saran

Pada penelitian selanjutnya, peneliti dapat menambahkan nilai number partitions dan nilai confidence yang lebih bervariasi agar dapat mengetahui seberapa jauh kinerja iterative partitioning filter dengan semakin banyak pemotongan data dan nilai confidence. Pada proses validasi peneliti juga menggunakan metode klasifikasi yang lain agar dapat dilihat kinerja iterative partitioning filter dengan metode klasifikasi tersebut.

DAFTAR PUSTAKA

- Anagnostopoulos, I. et al., 2005. The Wisconsin Breast Cancer Problem: Diagnosis and DFS time prognosis using probabilistic and generalised regression neural classifiers. *Oncology Reports*.
- Anand, R., Kirar, V.P.S. & Burse, K., 2013. K-Fold Cross Validation and Classification Accuracy of PIMA Indian Diabetes Data Set Using Higher Order Neural Network and PCA. *International Journal of Soft Computing and Engineering (IJSCE)*, II(6), pp.436-38.
- Anand, R., Kirar, V.P.S. & Kavita, B., 2012. Data Pre-processing and Neural Network Algorithms for Diagnosis of Type II Diabetes: A Survey. *International Journal of Engineering and Advanced Technology (IJEAT)*, II(1), pp.49-52.
- Bargeron, B., 2003. *Essential of Knowledge Management*. New Jersey: John Wiley & Sons, Inc.
- Br Ginting, S.L., Zarman, W. & Hamidah, I., 2014. ANALISIS DAN PENERAPAN ALGORITMA C4.5 DALAM DATA MINING UNTUK MEMPREDIKSI MASA STUDI MAHASISWA BERDASARKAN DATA NILAI AKADEMIK. *Prosiding Semianar Nasional Aplikasi Sains & Teknologi (SNAST)*, pp.263-72.
- Fahlman, S.E. & Hinton, G.E., 1987. Connectionist Architectures for Artificial Intelligence. *IEEE Computer*, 20(1), pp.100-09.
- Han, J. & Kamber, M., 2006. *Data Mining Concepts and Techniques*. 2nd ed. San Fransisco, CA, USA: Morgan Kaufmannn.
- Haryati, D.F., Abdillah, G. & Hadiana, A.I., 2016. KLASIFIKASI JENIS BATUBARA MENGGUNAKAN JARINGAN SYARAF TIRUAN DENGAN ALGORITMA BACKPROPAGATION. *Seminar Nasional Teknologi Informasi (SENTIKA)*, pp.557-62.
- Ho, T.K. & Basu, M., 2002. Complexity Measures of Supervised Classification Problems. *IEEE*

Transactions on Pattern Analysis and Machine Intelligence, pp.1-20.

J.Roiger, R. & W.Geatz, M., 2003. *Data Mining A Tutorial-Based Primer*. International Edition ed. Canada: Pearson Education, Inc.

Junaedi, H., Herman, B., Maryati, I. & Melani, Y., 2011. Data Transformation Pada Data Mining. *IDEaTech*, pp.93-99.

Karthick, R. & Malathi, D.A., 2015. Preprocessing of Various Data Sets Using Different Classification Algorithms for Evolutionary Programming. *International Journal of Science and Research (IJSR)*, IV(4), pp.2730-33.

Khoshgoftaar, T.M. & Rebours, P., 2007. Improving Software Quality Prediction by Noise Filtering Techniques. *JOURNAL OF COMPUTER AND TECHNOLOGY*, 22(3), pp.387-96.

Kotsiantis, S.B., Kanellopoulos, D. & Pintelas, P.E., 2006. Data Preprocessing for Supervised Learning. *INTERNATIONAL JOURNAL OF COMPUTER SCIENCE*, I(2), pp.111-17.

Mabrur, A.G. & Lubis, R., 2012. PENERAPAN DATA MINING UNTUK MEMPREDIKSI. *KRITERIA NASABAH KREDI*, I(1), pp.53-57.

Maharani, W., 2009. KLASIFIKASI DATA MENGGUNAKAN JST BACKPROPAGATION MOMENTUM DENGAN ADAPTIVE LEARNING RATE. *Seminar Nasional Informatika (semnasIF)*, pp.25-31.

Mansilla, E.B. & Tin, K.H., 2005. Domain of competence of XCS classifier system in complexity measurement space. *IEEE Transactions on Evolutionary Computation*, 9(1), pp.82-104.

Nawi, N.M., Atomi, W.H. & Rehman, M.Z., 2013. The Effect of Data Pre-Processing on Optimized Training of Artificial Neural Networks. *Procedia Technology*, 8(C), pp.33-40.

- Quinlan, J.R., 1993. *C4.5 : Programs for Machine Learning*. San Mateo: Morgan Kaufmann, CA.
- Raviya, K.H. & Gajjar, B., 2013. Performance Evaluation of Different Data Mining Classification Algorithm Using WEKA. *Indian Journal of Research*, II(1).
- Ronsen, P., 2012. Data Mining : Masa Lalu, Sekarang dan Masa Mendatang. *JSM STMIK Mikroskil*, 13(1), p.41.
- Saez, J.A., Galar, M., Luengo, J. & Herrera, F., 2016. An iterative class noise filter based on the fusion of classifiers with noise sensitivity control. *Information Fusion*, 27, pp.19-32.
- Saucha, D. & Setiawan, N.A., 2015. PERBANDINGAN ALGORITME J48 DAN NBTREE UNTUK KLASIFIKASI DIAGNOSA PENYAKIT PADA SOYBEAN. *Seminar Nasional Teknologi Informasi dan Komunikasi (SENTIKA)*, pp.205-12.
- Singhal, S. & Jena, M., 2013. A Study on WEKA Tool for Data Preprocessing, Classification and Clustering. *international Journal of Innovative Technology and Exploring Engineering*, 2(6), pp.250-53.
- Sobran, N.M.M., Ahmad, A. & Ibrahim, Z., 2013. CLASSIFICATION OF IMBALANCED DATASET USING CONVENTIONAL NAÏVE BAYES CLASSIFIER. *Proceeding of the International Conference on Artificial Intelligence in Computer Science and ICT*, pp.35-42.
- Swain, M., Dash, S.K., Dash, S. & Mohapatra, A., 2012. AN APPROACH FOR IRIS PLANT CLASSIFICATION USING NEURAL NETWORK. *International Journal on Soft Computing (IJSC)*, III(1), pp.79-89.
- Swastina, L., 2013. Penarapan Algoritma C4.5 Untuk Penentuan Jurusan Mahasiswa. *GEMA AKTUALITA*, II(1), pp.93-98.
- Tong, D.L., 2011. A Simpler Method Of Preprocessing MALDI-TOF MS Data For Differential Biomarker Analysis : Sistem Cell and Melanoma Cancer Studies. *Clinical Proteomics*, 8(14), pp.1-18.

Witten, I.H., Frank, E. & Hall, M.A., 2011. *Data Mining : Practical Machine Learning Tools and Techniques*. 3rd ed. Elsevier.

Zhao, Y. & Zhang, Y., 2008. Comparison of decision tree methods for finding active objects. *Advances on Space Research*, 41(12), pp.1955-59.

Zhu, X., Wu, X. & Chen, Q., August 2003. Eliminating class noise in large datasets. *Proc. the 20th Int. Conf. Machine Learning, Washington DC*, pp.920-27.

