

BAB VI

KESIMPULAN DAN SARAN

6.1 Pendahuluan

Pada bab sebelumnya penulis telah melakukan analisis terhadap kinerja dari algoritma *k-Nearest Neighbor* dan *Modified k-Nearest Neighbor* untuk pengklasifikasian dataset *website phishing*. Pada bab ini kesimpulan dan saran akan penulis buat dengan harapan agar dapat digunakan sebagai acuan bagi penelitian selanjutnya pada bidang yang linier dengan penelitian yang penulis lakukan.

6.2 Kesimpulan

Pada penelitian yang telah penulis lakukan, berfokus pada kinerja dari algoritma klasifikasi *k-Nearest Neighbor* dan *Modified k-Nearest Neighbor* untuk pengklasifikasian data, sehingga dapat ditarik beberapa kesimpulan, yaitu sebagai berikut :

1. Algoritma *k- Nearest Neighbor* dan *Modified k-Nearest Neighbor* dapat digunakan untuk pengklasifikasian data menggunakan dataset *website phishing*. Berdasarkan hasil yang telah peneliti lakukan ternyata algoritma *Modified k- Nearest Neighbor* tidak mampu meningkatkan kualitas klasifikasi secara signifikan dan menambah waktu komputasi yang diperlukan.
2. Algoritma *Modified k- Nearest Neighbor* berdasarkan hasil penelitian yang telah dilakukan terjadi peningkatan terhadap nilai rata-rata dari *precision*, *recall* dan *accuracy* pada proses klasifikasi menggunakan dataset *website phishing*, dengan penambahan proses perhitungan nilai validitas dan *wight*

voting sehingga menghasilkan peningkatan nilai rata-rata dari *precision*, *recall* dan *accuracy* dibandingkan menggunakan algoritma *k – Nearest Neighbour*. Dimana algoritma *Modified k – Nearest Neighbour* menghasilkan nilai akurasi terbaik yang dapat dicapai sebesar 95.60 %. Sedangkan hasil klasifikasi menggunakan algoritma *k – Nearest Neighbour* menghasilkan nilai rata – rata akurasi terbaik sebesar 95.35%. Dalam menentukan nilai *k* pada *k – Nearest Neighbour* dan *Modified k – Nearest Neighbour* berdasarkan ruang solusi yang dilakukan peneliti nilai *k = 5* memiliki pengaruh terhadap nilai *precision*, *recall* dan *accuracy* terbaik untuk kedua algoritma. Untuk waktu komputasi algoritma *Modified k – Nearest Neighbour* memiliki waktu komputasi yang lebih besar dibandingkan algoritma *k – Nearest Neighbour*. Sedangkan untuk memori yang digunakan pada algoritma *Modified k – Nearest Neighbour* lebih kecil dibandingkan algoritma *k – Nearest Neighbour*.

6.3 Saran

Berdasarkan percobaan yang telah dilakukan pada bab sebelumnya terdapat beberapa saran dari peneliti, diharapkan:

1. Memperbaiki kelemahan algoritma *Modified k – Nearest Neighbour* yang memiliki waktu komputasi lebih lama dibandingkan algoritma *k – Nearest Neighbour*.
2. Lakukan perbandingan kinerja antara algoritma *k-Nearest Neighbor* dan *Modified k-Nearest Neighbor* untuk *dataset* yang berbeda.
3. Lakukan pengujian menggunakan algoritma klasifikasi lainnya.

DAFTAR PUSTAKA

- Bhatia, N. (2010). Survey of nearest neighbor techniques. *arXiv preprint arXiv:1007.0085*.
- Eyupoglu, C. (2018). Breast cancer classification using k-nearest neighbors algorithm. *The Online Journal of Science and Technology-July*, 8(3).
- Frank, E., & Hall, M. A. (2011). *Data mining: practical machine learning tools and techniques*: Morgan Kaufmann.
- Gazalba, I., & Reza, N. G. I. (2017). *Comparative analysis of k-nearest neighbor and modified k-nearest neighbor algorithm for data classification*. Paper presented at the Information Technology, Information Systems and Electrical Engineering (ICITISEE), 2017 2nd International conferences on.
- Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*: Elsevier.
- Hassanat, A. B., Abbadi, M. A., Altarawneh, G. A., & Alhasanat, A. A. (2014). Solving the problem of the K parameter in the KNN classifier using an ensemble learning approach. *International Journal of Computer Science and Information Security*, 12.
- Islam, M. J., Wu, Q. J., Ahmadi, M., & Sid-Ahmed, M. A. (2007). *Investigating the performance of naive-bayes classifiers and k-nearest neighbor classifiers*. Paper presented at the 2007 International Conference on Convergence Information Technology (ICCIT 2007).
- James, L. (2006). Phishing exposed. Tech target article sponsored by: Sunbelt software. *available at: searchexchange.com*.
- Karegowda, A. G., Jayaram, M., & Manjunath, A. (2012). Cascading k-means clustering and k-nearest neighbor classifier for categorization of diabetic patients.
- Mohammad, R. M., Thabtah, F., & McCluskey, L. (2015). Phishing websites features *School of Computing and Engineering*: University of Huddersfield.
- Olson, D. L., & Delen, D. (2008). *Advanced data mining techniques*: Springer Science & Business Media.
- Parvin, H., Alizadeh, H., & Minaei-Bidgoli, B. (2010). *MKNN: Modified k-nearest neighbor*. *Global Journal of Computer Science and Technology*.

- Pratiwi, S. N. D., & Ulama, B. S. S. (2016). Klasifikasi Email Spam dengan Menggunakan Metode Support Vector Machine dan k-Nearest Neighbor. *Jurnal Sains dan Seni ITS*, 5(2).
- Purwiantono, F. E., & Tjahyanto, A. (2017). Model Klasifikasi Untuk Deteksi Situs Phising Di Indonesia: no.
- Refaeilzadeh, P., Tang, L., & Liu, H. (2009). Cross-validation. *Encyclopedia of database systems*, 532-538.
- Wafiyah, F., Hidayat, N., & Perdana, R. S. (2017). Implementasi Algoritma Modified K-Nearest Neighbor (MKNN) untuk Klasifikasi Penyakit Demam. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN*, 2548, 964X.
- Yusa, M., Utami, E., & Luthfi, E. T. (2016). Analisis Komparatif Evaluasi Performa Algoritma Klasifikasi pada Readmisi Pasien Diabetes. *Jurnal Buana Informatika*, 7(4).
- Yusra, Y. (2016). Perbandingan Klasifikasi Tugas Akhir Mahasiswa Jurusan Teknik Informatika Menggunakan Metode Naïve Bayes Classifier dan K-Nearest Neighbor. *Jurnal Sains dan Teknologi Industri*, 14(1), 79-85.
- Zaki, M. J., Meira Jr, W., & Meira, W. (2014). *Data mining and analysis: fundamental concepts and algorithms*: Cambridge University Press.