

BAB 6

KESIMPULAN DAN SARAN

Bab ini membahas tentang kesimpulan dari penelitian yang sudah dilakukan dan saran-saran yang dapat digunakan untuk penelitian berikutnya.

6.1 Kesimpulan

Berdasarkan hasil eksperimen dan analisis yang sudah dilakukan, terdapat beberapa kesimpulan:

1. Faktor-faktor yang berhubungan dengan stroke dapat ditentukan dengan melihat sebaran atribut stroke dengan atribut lain, menggunakan metode *chi square*, dan menghitung nilai *information gain*. Dari langkah-langkah yang sudah dilakukan, terdapat 21 faktor yang berhubungan dengan stroke yaitu kesulitan berjalan, serangan jantung, kondisi kesehatan, penyakit jantung koroner, kesulitan melakukan tugas sendirian, kesulitan berpakaian, usia, penyakit paru-paru, penyakit ginjal, kesulitan berkonsentrasi, radang sendi, diabetes, kanker lain, kanker kulit, asma, depresi, peminum berat, status merokok, pendapatan, ras, dan kategori BMI. Faktor-faktor ini digunakan sebagai atribut untuk model klasifikasi. Jumlah atribut terbaik untuk algoritma *Decision Tree* dan *Random Forest* adalah 7 atribut pertama dari 21 faktor di atas dan untuk algoritma *Naïve Bayes* adalah 20 atribut pertama dari 21 faktor di atas yang ditentukan dengan metode *cross-validation*.
2. Algoritma klasifikasi yang tepat untuk mengolah data survei BRFSS adalah menggunakan algoritma *Decision Tree*, *Categorical Naïve Bayes*, dan *Random Forest* karena fitur yang digunakan bertipe kategorik.
3. Model terbaik dilihat dari nilai akurasi, presisi, *recall*, dan *f1-score*, tetapi nilai yang lebih diutamakan adalah nilai *recall* karena akan sangat besar risikonya jika model menyatakan penderita stroke sebagai orang yang sehat. Model terbaik yang diambil dari setiap algoritma adalah *Decision Tree* menggunakan 7 atribut dengan nilai akurasi 75%, presisi 74%, *recall* 79%, dan *f1-score* 76%, *Categorical Naïve Bayes* menggunakan 20 atribut dengan nilai akurasi 74%, presisi 76%, *recall* 69%, dan *f1-score* 72%, dan *Random Forest* menggunakan 7 atribut dengan nilai akurasi 75%, presisi 73%, *recall* 79%, dan *f1-score* 76%.
4. Perangkat lunak dirancang agar pengguna dapat memilih algoritma yang ingin digunakan untuk memprediksi stroke berdasarkan model terbaik. Atribut-atribut dari model setiap algoritma diimplementasikan ke dalam bentuk pertanyaan yang harus dijawab pengguna. Perangkat lunak berhasil memprediksi penderita stroke dan bukan berdasarkan pertanyaan yang sudah dijawab.

6.2 Saran

Berikut merupakan saran-saran yang dapat dilakukan untuk penelitian berikutnya:

1. Melakukan pembuatan model klasifikasi menggunakan algoritma lain yang dapat menangani fitur-fitur bertipe kategorik seperti *Support Vector Machine*.
2. Melakukan lebih banyak kombinasi atribut untuk eksperimen agar mendapatkan nilai evaluasi model yang lebih baik.

DAFTAR REFERENSI

- [1] Ramadany, A. F., Pujarini, L. A., dan Candrasari, A. (2013) Hubungan diabetes melitus dengan kejadian stroke iskemik di rsud dr. moewardi surakarta tahun 2010. *Algorithmica*, **66**, 595–614.
- [2] Rachmawati, D., Ningsih, D. K., dan Andarini, S. (2020) Factors affecting knowledge towards stroke risks and early symptoms in emergency department. *Malang Neurology Journal*, **6**, 11–19.
- [3] Han, J., Kamber, M., dan Pei, J. (2011) *Data Mining Concepts and Techniques*, 3rd edition. Elsevier, Massachusetts.
- [4] Mendenhall, W., Beaver, R. J., dan Beaver, B. M. (2009) *Introduction to Probability and Statistics*, 13th edition. Cengage Learning, Canada.
- [5] Chawla, N. V., Bowyer, K. W., Hall, L. O., dan Kegelmeyer, W. P. (2002) Smote: Synthetic minority over-sampling technique. *Journal Of Artificial Intelligence Research*, **16**, 321–357.