

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, maka dapat disimpulkan bahwa:

1. Model terbaik untuk memprediksi tingkat pemulihan cacat merupakan model regresi yang dibangun oleh *log loss function* dan dievaluasi performanya berdasarkan RMSE pada pohon ke-146 yang dibangun berdasarkan pohon sebelumnya. Variabel yang digunakan adalah 71,03589% dari keseluruhan variabel serta data yang digunakan pada model ini adalah hanya 85,48898% dari total keseluruhan data. Model ini membuat 146 buah pohon yang memiliki kedalaman sebesar 10 dengan jumlah dari turunan kedua dari *loss function* adalah sebesar 37 dan menjalankan 5 tugas yang dilakukan dalam satu waktu. Parameter regularisasi yang digunakan pada model ini adalah η sebesar 0,1809017, λ sebesar 19,87699, dan γ sebesar 1,58336.
2. Variabel yang paling penting dalam memengaruhi akurasi dari tingkat pemulihan cacat adalah durasi pemberian manfaat kepada tertanggung setelah periode eliminasi, diikuti oleh *own to any occupation transition: OwnOther*, *age band*, *own to any occupation transition: Own+1*, *disability category: injury other than back*, dan *disability category: other musculoskeletal*.
3. Hasil prediksi tingkat pemulihan cacat dengan menggunakan metode XGBoost menghasilkan prediksi yang lebih akurat dibandingkan dengan menggunakan metode GBM. Hal ini terlihat dari nilai RMSE dari metode XGBoost lebih rendah dibandingkan dengan metode GBM. RMSE dari metode XGBoost adalah 0,1066381, sedangkan RMSE dari metode GBM adalah 0,1070939.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, beberapa saran yang dapat direkomendasikan adalah:

1. Penelitian selanjutnya dapat menerapkan metode XGBoost untuk membuat model prediksi pada data tingkat pemulihan asuransi grup cacat di Indonesia.
2. Penelitian selanjutnya dapat mengimplementasikan metode *machine learning* lainnya, seperti *dropouts meet multiple additive regression trees* (DART), *light gradient-boosting machine* (LightGBM), dan *categorical boosting* (CatBoost) untuk memprediksi tingkat pemulihan cacat pada data asuransi grup cacat.

DAFTAR REFERENSI

- [1] Maleh, J. dan Bosley, T. (2020) *Disability and Death Probability Tables for Insured Workers Born in 2000*. Baltimore: Social Security Administration, 1.
- [2] Kopinsky, M. (2017) *Predicting Group Long Term Disability Recovery and Mortality Rates Using Tree Model*. Society of Actuary, 3rd edition. Springer-Verlag, USA.
- [3] Vivian (2020) Prediksi Tingkat Pemulihan Kehamilan Peserta Asuransi Cacat Berkelompok Jangka Panjang Menggunakan Metode *Random Forest*. Skripsi. Universitas Katolik Parahyangan, Indonesia.
- [4] Mangaratua, L. (2020) Aplikasi Metode *Random Forest* dalam Memprediksi Tingkat Mortalita Asuransi Grup Cacat Jangka Panjang. Skripsi. Universitas Katolik Parahyangan, Indonesia.
- [5] Mercyana, A. (2020) Pemodelan Tingkat Pemulihan Asuransi Grup Cacat Jangka Panjang Menggunakan *Gradient Boosting Machine* (GBM). Skripsi. Universitas Katolik Parahyangan, Indonesia.
- [6] Reinhart, R. (2022) Prediksi Tingkat Mortalita Asuransi Grup Cacat Jangka Panjang Menggunakan *Neural Network*. Skripsi. Universitas Katolik Parahyangan, Indonesia.
- [7] Gunawan, V. N. (2022) Prediksi Tingkat Pemulihan Kehamilan pada Asuransi Cacat Berkelompok Jangka Panjang Menggunakan Metode *Neural Network*. Skripsi. Universitas Katolik Parahyangan, Indonesia.
- [8] Budiana, S. (2022) Aplikasi *Bayesian Additive Regression Tree* (BART) untuk Memprediksi Tingkat Pemulihan Kehamilan Asuransi Disabilitas Kelompok Jangka Panjang. Skripsi. Universitas Katolik Parahyangan, Indonesia.
- [9] Thamara, F. I. (2022) Pemodelan Tingkat Pemulihan Asuransi Grup Cacat Jangka Panjang Menggunakan *Neural Network*. Skripsi. Universitas Katolik Parahyangan, Indonesia.
- [10] Chen, T. dan Guestrin, C. (2016) *XGBoost: A Scalable Tree Boosting System*. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, NY, USA, August KDD '16, pp. 785–794. ACM.
- [11] Wade, C. (2020) *Hands-On Gradient Boosting with XGBoost and Scikit-Learn: Perform Accessible Machine Learning and Extreme Gradient Boosting with Python*. Packt Publishing Ltd, United Kingdom.
- [12] Juwanto, H., Bungsu, A., Pasaribu, T., Demus, N., dan Iman, I. (2022) Tabel Morbiditas Indonesia. Technical Report 1. Asosiasi Asuransi Jiwa Indonesia, Indonesia.
- [13] Office of The Chief Actuary (2018) *The Long-range Disability Assumptions for The 2018 Trustees Report*. Technical Report 1. Social Security Administration, USA.
- [14] Russel, S. (2021) *Artificial Intelligence: A Modern Approach*, 4th edition. Pearson, Hoboken.
- [15] Schmuller, J. (2013) *Statistical Analysis with Excel for Dummies*. John Wiley & Sons, USA.

-
- [16] Sá, A., Almeida, A., Rocha, B., Mota, M., Souza, J., dan Dentel, L. (2011) *Lightning Forecast Using Data Mining Techniques on Hourly Evolution of The Convective Available Potential Energy*. *Brazilian Congress on Computational Intelligence, Fortaleza, November, Brazil, November*, pp. 8–11. Brazilian Society on Computational Intelligence (SBIC).
- [17] Breiman, L., Friedman, J., Stone, C., dan Olshen, R. (1984) *Classification and Regression Trees*, 1st edition. Chapman & Hall/CRC, New York.
- [18] James, G., Witten, D., Hastie, T., dan Tibshirani, R. (2013) *An Introduction to Statistical Learning*. Springer, California.
- [19] Zhou, Z.-H. (2012) *Ensemble Methods: Foundations and Algorithms*. CRC press.
- [20] Friedman, J. H. (2002) Stochastic Gradient Boosting. *Computational Statistics & Data Analysis*, **38**, 367–378.
- [21] Sammut, C. dan Webb, G. I. (ed.) (2010) *Mean Squared Error*. Bagian dari Sammut, C. dan Webb, G. I. (ed.), *Encyclopedia of Machine Learning*. Springer US, Boston, MA.
- [22] Varberg, D. E., Purcell, E. J., dan Rigdon, S. E. (2007) *Calculus*. Pearson Educación, USA.
- [23] Ekman, M. (2021) *Learning Deep Learning: Theory and Practice of Neural Networks, Computer Vision, NLP, and Transformers Using TensorFlow*. Addison-Wesley Professional.
- [24] Yan, M. (2012) *Extension of Convex Function*. *arXiv preprint arXiv:1207.0944*, **1**.
- [25] Shen, Y. (2005) *Loss Functions for Binary Classification and Class Probability Estimation*. University of Pennsylvania.
- [26] Chen, T., He, T., Benesty, M., dan Khotilovich, V. (2019) *Package ‘XGBoost’*. *R version*, **90**.