

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan pembahasan pada bab-bab sebelumnya, diperoleh kesimpulan sebagai berikut.

1. Dari hasil eksperimen, diperoleh 3 model *Cox*, yaitu model *Cox* tanpa penalti, model *Cox* dengan penalti *L2*, dan model *Cox* dengan penalti *L1*. Disimpulkan bahwa model *Cox* dengan penalti *L2* merupakan model yang terbaik dalam memprediksi probabilitas hidup pasien kanker paru-paru ras Asia maupun keseluruhan ras. Model *Cox* dengan penalti *L2* merupakan model terbaik karena model ini memiliki nilai *concordance* tertinggi dan bebas dari *overfitting* maupun *underfitting*.
2. Model *Cox* dengan penalti *L2* juga dapat digunakan untuk memodelkan pasien kanker paru-paru dari *data set* selain yang terdapat dalam skripsi ini, namun dengan syarat variabel bebasnya sama. Hal ini dapat dilakukan karena model *Cox* dengan penalti *L2* telah diregularisasi dan divalidasi silang, sehingga prediksi probabilitas yang diperoleh tidak terbatas untuk *data set* pada skripsi ini.
3. Pada bulan ke-0 hingga 100, pasien kanker paru-paru dengan ras Asia memiliki probabilitas hidup yang lebih tinggi daripada keseluruhan pasien kanker paru-paru. Setelah 100 bulan, keduanya memiliki probabilitas hidup yang hampir identik.

5.2 Saran

Saran untuk pengembangan skripsi ini diberikan sebagai berikut.

1. Melakukan uji statistik untuk mengetahui signifikansi dari model *Cox*. Hal ini perlu dilakukan karena penilaian model *Cox* terbaik hanya menggunakan skor kebaikan model sehingga signifikansi dari model *Cox* tidak diketahui.
2. Pada skripsi ini, banyaknya variabel bebas yang digunakan adalah 11, padahal terdapat lebih dari 50 variabel bebas pada sumber tempat *data set* pasien kanker diperoleh. Variabel bebas tersebut banyak yang tidak digunakan karena keterbatasan waktu untuk memahami seluruh definisi dari variabel bebas. Dengan demikian, diharapkan pada penelitian selanjutnya, variabel bebas tersebut dapat dipahami definisinya dan digunakan pada saat estimasi parameter.

DAFTAR REFERENSI

- [1] Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., dan Bray, F. (2021) Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, **71**, 209–249.
- [2] Kaplan, E. L. dan Meier, P. (1958) Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, **53**, 457–481.
- [3] Cox, D. R. (1972) Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, **34**, 187–202.
- [4] George, B., Seals, S., dan Aban, I. (2014) Survival analysis and regression models. *Journal of nuclear cardiology*, **21**, 686–694.
- [5] Klugman, S. A., Panjer, H. H., dan Willmot, G. E. (2012) *Loss models: from data to decisions*. John Wiley & Sons.
- [6] London, D. (1997) *Survival models and their estimation*. ACTEX publications.
- [7] Han, J., Pei, J., dan Kamber, M. (2011) *Data mining: concepts and techniques*. Elsevier.
- [8] Mittal, S., Madigan, D., Burd, R. S., dan Suchard, M. A. (2014) High-dimensional, massive sample-size cox proportional hazards regression for survival analysis. *Biostatistics*, **15**, 207–221.
- [9] Mittal, S., Madigan, D., Cheng, J. Q., dan Burd, R. S. (2013) Large-scale parametric survival analysis. *Statistics in medicine*, **32**, 3955–3971.
- [10] Genkin, A., Lewis, D. D., dan Madigan, D. (2007) Large-scale bayesian logistic regression for text categorization. *technometrics*, **49**, 291–304.
- [11] Lin, D. (2007) On the breslow estimator. *Lifetime data analysis*, **13**, 471–480.
- [12] Moore, D. F. (2016) *Applied survival analysis using R*. Springer.
- [13] Tay, K., Simon, N., Friedman, J., Hastie, T., Tibshirani, R., dan Narasimhan, B. (2021) Regularized cox regression.