

BAB 6

KESIMPULAN DAN SARAN

6.1 Kesimpulan

Berdasarkan analisis dan eksperimen yang telah dilakukan pada tugas akhir ini, kesimpulan yang dapat dibuat adalah

1. Fitur yang berpengaruh terhadap IPK, lama studi, dan jalur masuk UNPAR ditentukan menggunakan teknik statistika, yaitu korelasi *Pearson* dan korelasi *Chi-Square*. Fitur yang memiliki pengaruh untuk IPK adalah SEMESTER TEMPUH, PROGRAM STUDI, dan JALUR. Fitur yang memiliki pengaruh untuk lama studi adalah IPK, PROGRAM STUDI, Provinsi asal SMA, dan JALUR.
2. Hubungan antara IPK atau lama studi, dan jalur masuk UNPAR dianalisis menggunakan teknik statistika yaitu metode visualisasi dan metode korelasi. Hasil yang diperoleh adalah untuk setiap jalur masuk, IPK dan lama studi memiliki arah hubungan negatif, yaitu jika nilai IPK mengalami kenaikan maka lama studi akan turun. Sebaliknya, jika nilai IPK mengalami penurunan maka lama studi akan naik. Jalur masuk Seleksi Khusus dan PMDK memiliki rata-rata IPK yang lebih tinggi dibandingkan dengan jalur USM 1, USM 2, dan USM 3. Lama studi untuk jalur Seleksi Khusus dan PMDK memiliki rata-rata yang lebih rendah dibandingkan dengan jalur USM 1, USM 2, dan USM 3. Ini menunjukkan bahwa jalur masuk memiliki pengaruh terhadap nilai IPK dan lama studi.
3. Model yang dibangun untuk analisis hubungan antara IPK atau lama studi dan jalur masuk UNPAR menggunakan algoritma klasterisasi. Algoritma yang menghasilkan model terbaik adalah algoritma *K-Means*. Hasil pengelompokan yang diperoleh adalah jalur masuk PMDK merupakan anggota terbanyak pada *Cluster 0* dan *Cluster 3*. *Cluster 0* memiliki rata-rata IPK 3.568 dan rata-rata lama studi 7 semester, *Cluster 3* memiliki rata-rata IPK 3.284 dan rata-rata lama studi 8 semester.
4. Model analisis prediktif untuk memprediksi nilai IPK dan lama studi diluncurkan menggunakan GUI. Model terbaik untuk melakukan prediksi IPK adalah model dengan algoritma *Decision Tree*, model terbaik untuk melakukan prediksi lama studi adalah model dengan algoritma *Naive Bayes*. GUI yang dibuat memiliki tiga halaman. Halaman pertama berisi deskripsi *dataset*, rata-rata IPK dan lama studi untuk setiap program studi dan jalur masuk. Halaman kedua berisi peluncuran model prediksi IPK, halaman ketiga berisi peluncuran model prediksi lama studi. Berdasarkan pengujian fungsional yang sudah dilakukan, model berhasil diluncurkan dengan baik.

6.2 Saran

Analisis statistika yang telah dilakukan hanya membandingkan dua atribut saja, yaitu menggunakan *Pearson* dan *Chi-Square*. Berdasarkan hasil analisis dan eksperimen yang telah dilakukan pada tugas akhir ini, saran yang dapat dilakukan untuk meningkatkan analisis dan eksperimen adalah menggunakan teknik analisis statistika lanjutan seperti ANOVA untuk membandingkan rata-rata dari tiga atribut atau lebih.

DAFTAR REFERENSI

- [1] Han, J., Kamber, M., dan Pei, J. (2012) *Data Mining. Concepts and Techniques, 3rd Edition*, 3rd edition. Morgan Kaufmann.
- [2] Wes McKinney (2010) Data Structures for Statistical Computing in Python. , ?, 56–61.
- [3] Hunter, J. D. (2007) Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, **9**, 90–95.
- [4] Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., dan Oliphant, T. E. (2020) Array programming with NumPy. *Nature*, **585**, 357–362.
- [5] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., dan Duchesnay, E. (2011) Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, **12**, 2825–2830.
- [6] Foundation, P. S. (2024) tkinter — python interface to tcl/tk. <https://docs.python.org/3/library/pickle.html>. 10 Februari 2024.
- [7] Faraglia, D. (2014) Welcome to faker’s documentation! <https://faker.readthedocs.io/en/master/>. 10 Februari 2024.
- [8] Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., dan SciPy 1.0 Contributors (2020) SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, **17**, 261–272.
- [9] Foundation, P. S. (2024) tkinter — python interface to tcl/tk. <https://docs.python.org/3/library/tkinter.html>. 22 Maret 2024.
- [10] Philips, N. (2018) Yarr! the pirate’s guide to r. *The Observer*, -.
- [11] Budiati, D., Wilandari, Y., dan Suparti, M. J. (2014) Analisis hubungan antara lama studi, jalur masuk dan indeks prestasi kumulatif (ipk) menggunakan model log linier. *JURNAL GAUSSIAN*, **1**, 41–50.
- [12] Yuan, C. dan Yang, H. (2019) Research on k-value selection method of k-means clustering algorithm. *J*, **2**, 226–235.
- [13] Müllner, D. (2011) Modern hierarchical, agglomerative clustering algorithms. *arXiv*, **1**, 1–29.

- [14] Raschka dan Sebastian (2018) Model evaluation, model selection, and algorithm selection in machine learning. *arXiv*, **3**.
- [15] Wang, Z. dan Bovik, A. C. (2009) Mean squared error: Love it or leave it? a new look at signal fidelity measures. *IEEE Signal Processing Magazine*, **26**, 98–117.
- [16] Hodson, T. O. (2022) Root mean square error (rmse) or mean absolute error (mae): When to use them or not. *Geoscientific Model Development Discussions*, **2022**, 1–10.
- [17] Podrzaj dan Primoz (2019) A brief demonstration of some python gui libraries. *Proceedings of the 8th International Conference on Informatics and Applications ICIA2019*, Wilmington, New Castle, DE 19801, USA, 2-4 August, pp. 1–6. SDIWC.