

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian dalam Skripsi ini telah memperoleh nilai R^2 untuk kedua model *Principal Component Analysis* (PCA) dan *Partial Least Square* (PLS) untuk mengatasi masalah multikolinieritas pada kasus sektor mobilitas yang paling mempengaruhi jumlah kasus COVID-19. Berdasarkan analisis dan pembahasan pada bab sebelumnya, kesimpulan yang dapat diambil adalah sebagai berikut.

1. Dalam menangani multikolinieritas, PCA memanfaatkan multikolinieritas dengan menggabungkan variabel-variabel yang berkorelasi tinggi menjadi satu set variabel yang tidak berkorelasi. Sedangkan PLS membuat model dengan mengidentifikasi komponen yang merupakan kombinasi linier dari variabel prediktor asli, kemudian membangunnya untuk memperhitungkan sebagian besar varian dalam variabel respons sambil meminimalkan korelasi antara variabel tersebut.
2. Tabel 5.1 menjelaskan sektor-sektor mobilitas yang memiliki hubungan asosiatif dengan jumlah kasus COVID-19 serta *error* dari setiap model PLS dan PCA (ada di halaman selanjutnya).
3. Metode *Principal Component Analysis* dapat menangani multikolinieritas lebih baik daripada metode *Partial Least Square*. Dalam penelitian ini, semua model yang mengalami multikolinieritas berhasil diatasi dengan kedua metode *Principal Component Analysis* (PCA) dan *Partial Least Square* (PLS).

Tabel 5.1: Sektor-Sektor Mobilitas yang Mempengaruhi Jumlah Kasus COVID-19 di Setiap Negara serta nilai RMSE dari setiap model PCA dan PLS

	Retail	Grocery	Parks	Stations	Workplaces	Residential	RMSE PCA Train	RMSE PCA Test	RMSE PLS Train	RMSE PLS Test
Australia 1				✓	✓	✓	125,67755	151,8068	123,5301	152,3642
Australia 2		✓	✓	✓			61,99788	85,74780	549,0174	550,3078
Australia 3	✓	✓	✓		✓		65,74101	69,99074	18224,76	26285,75
Kanada 1	✓	✓	✓	✓	✓		491,1352	884,1940	1326,492	1040,138
Kanada 2	✓	✓	✓	✓			379,4949	780,5637	2027,251	1963,784
Kanada 3	✓	✓				✓	419,8364	810,8091	7225,520	6843,554
Indonesia 1	✓	✓	✓				94,28556	120,1457	1610,440	1637,269
Indonesia 2	✓	✓	✓	✓	✓	✓	80,35740	95,40816	5686,887	5058,544
Indonesia 3	✓	✓	✓	✓			138,5329	160,9648	10421,33	11210,17
India 1	✓		✓		✓	✓	289,3258	351,9327	16920,00	15669,36
India 2	✓		✓	✓		✓	184,9081	169,9720	43080,74	34959,20
India 3	✓	✓		✓	✓	✓	103,8540	183,7177	39657,34	52591,47
Amerika Serikat 1	✓		✓	✓		✓	167,4930	237,0205	45748,50	47302,42
Amerika Serikat 2	✓	✓	✓				104,0467	187,7547	53004,52	65490,08
Amerika Serikat 3	✓	✓			✓	✓	92,43321	127,9358	147425,5	148925,5

Dalam Tabel, terlihat bahwa variabel mobilitas *retail* mendominasi sebagai variabel mobilitas yang memiliki hubungan asosiatif yang kuat dengan jumlah kasus COVID-19, dengan menyusul variabel-variabel mobilitas lainnya. Dalam hasil nilai RMSE, terlihat bahwa nilai RMSE pada model PLS negara Australia periode I lebih rendah dibandingkan nilai RMSE pada model PCA negara Australia periode I, namun untuk negara-negara dan periode-periode lainnya, nilai RMSE model PCA lebih rendah dibandingkan dengan nilai RMSE pada model PLS.

5.2 Saran

Metode yang digunakan dalam penelitian ini masih dapat dikembangkan agar mendapatkan hasil yang lebih akurat dan memberikan gambaran yang lebih jelas mengenai faktor-faktor yang menentukan performa kedua metode. Pengembangan terhadap penelitian ini meliputi hal-hal sebagai berikut.

1. Peneliti selanjutnya dapat menggunakan faktor-faktor selain jumlah kasus harian, seperti jumlah kasus mingguan, jumlah kematian atau jumlah orang yang sudah divaksinasi. *Dataset* juga dapat dibersihkan secara lebih baik untuk menghasilkan model yang lebih akurat.
2. Penelitian selanjutnya dapat mencoba metode lain yang dapat mengatasi masalah multikolinieritas seperti regresi LASSO atau regresi Ridge. Dalam regresi LASSO, metode ini menyusutkan koefisien taksiran sampai mendekati angka nol dan menyeleksi variabel-variabel independen sehingga menghasilkan model dengan variabel terbaik. Regresi Ridge juga menyusutkan koefisien taksiran namun tidak sampai mendekati angka nol, kemudian menyeleksi variabel-variabel independen sehingga menghasilkan model dengan variabel terbaik.
3. Penelitian selanjutnya dapat menggunakan negara lain, terutama yang memiliki jumlah kasus harian yang tinggi dan tidak konsisten. Biasanya, negara dengan jumlah kasus harian yang tinggi dan tidak konsisten rentan terhadap masalah multikolinieritas.
4. Penelitian selanjutnya dapat menghapus *outlier* dalam *dataset*. Adanya *outlier* dalam sebuah *dataset* dapat disebabkan karena adanya anomali seperti kesalahan pengukuran. Menghapus *outlier* dapat membantu menghasilkan model yang lebih akurat.

5. Penelitian selanjutnya dapat menambahkan *lag* pada data. Menambahkan *lag* dalam kumpulan data melibatkan pembuatan variabel baru yang mewakili nilai variabel sebelumnya.

