

BAB I

PENDAHULUAN

1.1 Latar Belakang

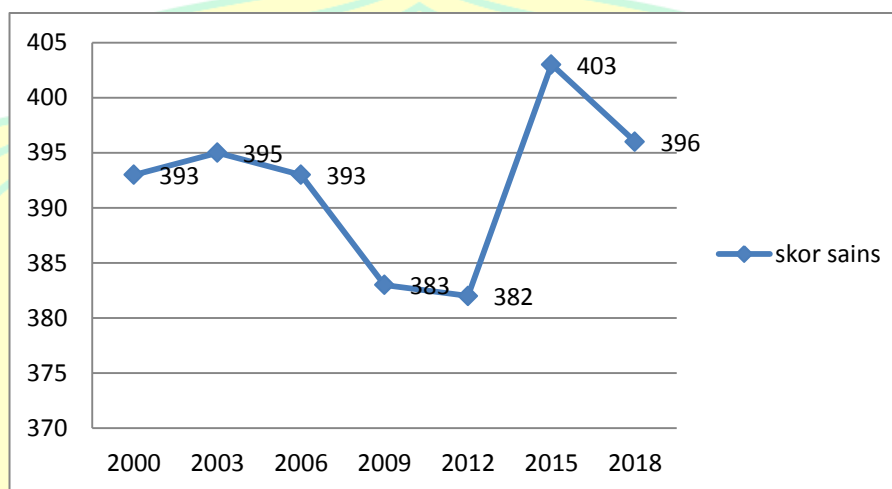
Teknologi, informasi, dan komunikasi saat ini berkembang dengan sangat pesatnya. Hal tersebut memberikan pengaruh dalam kehidupan manusia di berbagai aspek, tidak terkecuali bidang pendidikan (Wiyono, 2013). Anil (2019) menyatakan bahwa pendidikan dapat menyediakan hal-hal yang diperlukan untuk dapat berkontribusi di masyarakat dan menjadikan pendidikan sebagai fondasi utama yang harus dimiliki baik dalam bentuk keterampilan maupun pengetahuan. Salah satu pengetahuan yang menjadi kebutuhan utama untuk seorang pelajar adalah literasi sains (Deming et al., 2007).

Literasi sains menurut OECD (2019b) merupakan kemampuan seseorang menerapkan pengetahuan yang dimilikinya untuk mengidentifikasi masalah, mengkonstruksi pengetahuan baru, memberikan penjelasan secara ilmiah, mengambil kesimpulan, dan kemampuan mengembangkan pola pikir secara reflektif sehingga dapat berpartisipasi dalam mengatasi isu-isu dan gagasan terkait sains. Selain itu, literasi sains juga menawarkan pemenuhan kebutuhan personal dan kegembiraan yang dapat dibagikan dengan siapapun (Zuriyani, 2013). Hal tersebut menjadi salah satu landasan mengapa literasi sains merupakan suatu kebutuhan utama bagi seorang pelajar. Oleh karena itu, pengukuran dan evaluasi terkait literasi sains pada siswa Indonesia sangat diperlukan.

Programme for International Students Assessment (PISA) merupakan penilaian pendidikan secara internasional yang diselenggarakan oleh *Organization for Economic Cooperation and Development* (OECD) setiap 3 tahun sekali. PISA mengukur dan menilai literasi sains, membaca dan matematika. Tujuan diselenggarakannya PISA, yaitu untuk mengetahui sejauh mana siswa dapat berpikir melintasi batas disiplin mata pelajaran dan menerapkan pengetahuan mereka secara kreatif. Peserta PISA merupakan

pelajar berumur 15 tahun yang terdaftar dalam lembaga pendidikan di kelas 7 atau lebih tinggi yang diambil secara acak (Schleicher, 2018).

Indonesia telah mengikuti *Programme for International Students Assessment* (PISA) sejak tahun 2000. Pencapaian skor sains PISA dari tahun 2000 hingga tahun 2018 dapat dilihat pada Gambar 1.1 di bawah ini.



Gambar 1.1 Grafik Skor Sains PISA Indonesia Tahun 2000 - 2018

Awal mula keikutsertaan Indonesia dalam PISA yaitu tahun 2000, menempatkan Indonesia di peringkat ke-38 dari 41 partisipan dengan skor 393. Pencapaian tersebut tidak jauh berbeda dengan tahun – tahun berikutnya. Tahun 2009 dan 2012 Indonesia mengalami penurunan dengan skor 383 dan 382. Selanjutnya tahun 2015, Indonesia mengalami kenaikan dengan skor 403 dan menempatkan Indonesia pada peringkat ke-62 dari 69 negara. Namun demikian, tahun 2018 Indonesia kembali mengalami penurunan, dimana Indonesia mendapatkan peringkat ke-71 dari 79 negara dengan skor 396 (Hewi & Shaleh, 2020). Tentunya hal tersebut menjadi sebuah catatan penting baik untuk pemerintah, guru, pelajar, maupun orang tua. Melakukan evaluasi dan mengetahui faktor-faktor apa saja yang memengaruhi hasil PISA sangat diperlukan sebagai tindakan preventif untuk mendapatkan skor yang lebih baik kedepannya

Survey diagnostik yang diukur dalam PISA mengandung banyak peubah kategorik yang mengukur latar belakang dan karakteristik siswa. Mempertimbangkan penggunaan model yang tepat dan sederhana untuk

melihat hubungan antara latar belakang dan karakteristik siswa dengan hasil skor literasi sains PISA 2018 sangat diperlukan. Salah satu metode untuk mengetahui hubungan antar beberapa peubah dan melihat pengaruh suatu peubah terhadap peubah lainnya, yaitu dengan analisis regresi linear. Akan tetapi, metode ini kurang tepat digunakan pada penelitian ini dikarenakan beberapa peubah berkorelasi kuat antara satu dengan yang lainnya atau yang dikenal dengan multikolinearitas. Terjadinya Multikolinearitas dalam regresi linear menjadi suatu masalah dikarenakan pengaruh peubah penjelas terhadap peubah respon sulit untuk dilihat secara terpisah (Gujarati, 1992). Selain itu, ragam yang dihasilkan akan semakin besar seiring dengan bertambahnya multikolinearitas. Hal tersebut juga menyebabkan selang kepercayaan untuk parameter regresi semakin melebar dan menghasilkan inferensi yang buruk.

Robert Tibshirani (1996) memperkenalkan sebuah metode statistika yang dapat menangani masalah multikolinearitas dan dapat digunakan pada data dengan jumlah peubah penjelas yang sangat banyak, metode tersebut adalah *Least Absolute Shrinkage and Selection Operator* (LASSO). LASSO merupakan salah satu metode regresi terpinalti yang sama dengan regresi ridge, dimana keduanya melakukan pendugaan koefisien regresi dengan meminimalkan jumlah kuadrat galat ditambahkan suatu fungsi pinalti. Meskipun bentuk fungsi untuk pendugaan koefisien regresi pada keduanya terlihat sama, LASSO dan regresi ridge memiliki perbedaan yang cukup jelas. Fungsi pinalti yang digunakan pada regresi ridge hanya dapat menyusutkan koefisien regresi mendekati nol, sedangkan pada LASSO dapat menyusutkan nilai koefisien regresi sama dengan nol.

LASSO kurang tepat diterapkan pada data kategorik karena LASSO sangat bergantung pada pilihan strategi pengkodean (Susongko & Afrizal, 2018). Ketika strategi pengkodean yang dilakukan berbeda, maka pemilihan peubah dan model fit yang dihasilkan akan berbeda pula. Hal tersebut dapat diatasi dengan metode *group* LASSO yang diperkenalkan oleh Bakin (1999) dan kemudian dikembangkan oleh Yuan dan Lin (2006). *Group* LASSO merupakan generalisasi dari LASSO untuk melakukan penyeleksian peubah

berdasarkan kelompok. Algoritme *Group LASSO* memungkinkan kelompok peubah penjelas yang telah ditentukan sebelumnya dipilih masuk atau keluar dari model secara bersama-sama.

Penelitian terkait dengan regresi *LASSO* maupun *group LASSO* sudah banyak dilakukan. Penelitian sebelumnya yang dilakukan Soleh (2013) memberikan kesimpulan bahwa *LASSO* merupakan alternatif bagi pendugaan koefisien regresi dan penyeleksian peubah pada kondisi peubah mengalami multikolinearitas. Selain itu, *LASSO* dianggap lebih baik dibandingkan dengan regresi ridge dalam hal kekonsistenan hasil pendugaan parameter. Penelitian yang dilakukan oleh Chen (2020) menyatakan bahwa *group LASSO* lebih baik daripada *LASSO* dalam mengidentifikasi faktor-faktor yang berhubungan dengan tuberculosis di Jawa Barat berdasarkan nilai MSE yang lebih kecil. Penggunaan *Lasso* untuk data PISA sendiri telah dilakukan oleh Santi et al. (2019) menggunakan peubah respon skor Matematika. Penelitian ini menggunakan 200 variabel dan menghasilkan 11 variabel yang signifikan memengaruhi skor Matematika dengan metode *LASSO* berdasarkan nilai RMSE dan BIC (*Bayesian Information Criteria*).

Berdasarkan pada hal-hal yang telah ditulis sebelumnya, penyeleksian kelompok peubah menggunakan *group LASSO* pada skor sains data PISA Indonesia dipilih sebagai topik dalam Tugas Akhir ini. Penelitian ini diharapkan dapat membantu pihak-pihak terkait untuk mengetahui peubah-peubah yang memengaruhi skor sains data PISA sehingga dapat menentukan strategi yang lebih baik dalam penilaian berikutnya.

1.2 Perumusan Masalah

Rumusan masalah dalam penelitian ini diantaranya:

1. Bagaimana penerapan metode *group LASSO* dalam menyeleksi kelompok peubah yang memengaruhi skor sains data PISA Indonesia?
2. Kelompok peubah apa saja yang memengaruhi skor sains data PISA Indonesia secara signifikan berdasarkan metode *group LASSO*?

1.3 Tujuan Penelitian

Tujuan yang ingin di capai dalam penelitian ini diantaranya:

1. Mengetahui penerapan metode *group* LASSO dalam menyeleksi kelompok peubah yang memengaruhi skor sains data PISA Indonesia.
2. Mengetahui kelompok peubah apa saja yang memengaruhi skor sains data PISA secara signifikan berdasarkan metode *group* LASSO.

1.4 Manfaat Penelitian

Manfaat yang akan diperoleh dalam penelitian ini diantaranya:

1. Menambah wawasan tentang pengolahan data PISA terutama dengan metode *group* LASSO.
2. Dapat mengetahui kelompok peubah yang memengaruhi skor sains data PISA secara signifikan berdasarkan metode *group* LASSO.

