

ABSTRAK

Analisis regresi multivariat merupakan suatu analisis yang digunakan untuk mengetahui hubungan kausal antara beberapa variabel dependen dengan beberapa variabel independen. Masalah akan terjadi apabila data yang digunakan dalam analisis regresi multivariat mengandung data hilang. Data hilang dapat diabaikan apabila data yang hilang sedikit, namun apabila data yang hilang berjumlah besar maka data hilang tersebut tidak dapat diabaikan. Metode Algoritma EM merupakan suatu metode yang dapat digunakan dalam penanganan data hilang. Algoritma EM memiliki dua tahap yaitu langkah E (*Expectation*) dan langkah M (*Maximization*). Langkah E menentukan ekspektasi dari fungsi log likelihood dari data lengkap yang bergantung pada data yang teramati dan θ . Langkah M memaksimumkan fungsi log likelihood yang diestimasi dengan memperhatikan θ . Langkah E dan M akan terus berulang sampai estimasi dari parameter konvergen. Setelah data hilang tersebut diestimasi, analisis regresi multivariat dapat dilakukan dengan menggunakan data lengkap hasil estimasi dari algoritma EM.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Data merupakan keterangan tentang sesuatu. Keterangan tersebut dapat berupa angka (bilangan) dan dapat pula berupa kalimat. Secara statistik data dapat dibagi menjadi dua yaitu data metrik dan data non metrik. Data metrik merupakan data yang didapat dengan jalan mengukur dan bisa mempunyai nilai desimal, data metrik dikategorikan sebagai data interval atau data rasio, sedangkan data non metrik merupakan data yang didapat dengan jalan menghitung, tidak mempunyai desimal serta dilakukan dengan katagorisasi. Data non metrik dapat dikategorikan sebagai data nominal atau data ordinal

Uji data multivariat dilakukan untuk memastikan bahwa berbagai metode multivariat dapat digunakan pada suatu data tertentu, sehingga hasil proses multivariat dapat diinterpretasikan dengan tepat. Jika uji data diabaikan maka kesimpulan yang diambil atau bahkan metode multivariat yang digunakan untuk pengolahan data tersebut salah atau data tidak bisa diproses. Salah satu cara untuk melakukan uji data multivariat adalah dengan pengujian adanya data hilang, yakni menguji apakah data yang tidak lengkap akan mempengaruhi data secara keseluruhan. Salah satu alasan perlunya uji data multivariat dilakukan adalah jika data (yang terdiri dari banyak variabel) mempunyai banyak data hilang atau tidak ada isinya, jika data tersebut dipaksakan untuk tetap diproses maka kesimpulan yang dihasilkan akan sangat berbeda jika dibandingkan dengan data yang tidak

terdapat data hilang (data lengkap), bahkan pada data yang sangat banyak mengandung data hilang, proses multivariat tidak bisa dilakukan. Data hilang terjadi karena informasi untuk suatu objek tidak diberikan, sulit dicari atau memang informasi tersebut tidak ada.

Kendala yang dihadapi dengan adanya data hilang adalah banyaknya waktu dan tenaga yang terbuang untuk melakukan proses uji data. Sedangkan jika data ternyata tidak layak untuk diproses lebih lanjut, tidak ada jaminan untuk “pengobatan” terhadap data tersebut sehingga menjadi layak digunakan.

Secara prinsip multivariat merupakan perluasan dari univariat atau bivariat, dimana jika univariat atau bivariat hanya menghitung maksimal dua variabel, multivariat menghitung lebih dari dua variabel, dalam multivariat semua variabel dianalisis secara simultan atau bersamaan. Sebagai contoh tingkat pertumbuhan sebuah tanaman yang melibatkan seperti jumlah pupuk yang diberikan, curah hujan, tingkat keasaman tanah, intensitas sinar matahari dan sebagainya tidak bisa diselesaikan dengan menggunakan analisis univariat atau bivariat, karena keterbatasan kedua analisis tersebut.

Analisis regresi multivariat merupakan perluasan dari analisis regresi linier yang digunakan untuk mengetahui hubungan antar variabel-variabel independen dengan variabel-variabel dependen, untuk selanjutnya dibentuk p model persamaan regresi.

Regresi multivariat merupakan salah satu teknik multivariat yang mempelajari tentang regresi antara beberapa variabel dependen dan beberapa variabel independen, yang dilakukan secara bersama-sama.