

## Pemilihan Model Regresi Terbaik Dengan BIC (*Bayesian Information Criteria*)

Oleh:  
Curie Putri Hijrihani  
05305141019

### ABSTRAK

Analisis regresi adalah teknik statistika untuk membuat model dan menyelidiki hubungan antara dua variabel atau lebih. Hubungan antara variabel bebas dan variabel tak bebas dinyatakan dengan persamaan estimasi yang disebut persamaan regresi. Untuk menentukan persamaan regresi terbaik belum ada suatu prosedur statistik yang baku. Terdapat beberapa kriteria yang digunakan sebagai tolak ukur dalam menilai suatu kecocokan model dengan data, salah satunya yaitu dengan kriteria BIC. Tujuan penulisan skripsi ini adalah untuk menjelaskan mengenai langkah-langkah pemilihan model regresi terbaik dengan BIC dan menjelaskan penerapan pemilihan model regresi terbaik dengan BIC.

Persamaan model regresi linier dapat dituliskan sebagai berikut :  $Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon_i$  dengan  $\beta_0, \beta_1, \beta_2, \dots, \beta_k$  adalah parameter,  $X_1, X_2, \dots, X_k$  adalah variabel bebas pada amatan ke- $i$  dan  $\varepsilon_i$  adalah galat yang diasumsikan saling bebas. Metode yang digunakan untuk pemilihan model regresi terbaik pada penulisan ini adalah BIC atau *Bayesian Information Criteria*. Dalam *Bayesian Information Criteria* ( BIC) ini menyatakan bahwa semakin kecil nilai BIC maka semakin baik pula modelnya dan jika diperoleh nilai BIC yang negatif maka digunakan nilai absolutnya.

Dalam pemilihan model regresi terbaik dengan BIC terdapat langkah-langkah yang harus dilakukan yaitu: membuat plot antara variabel tak bebas dengan masing-masing variabel bebasnya, menentukan persamaan dugaan, pemeriksaan asumsi-asumsi regresi linier, melakukan transformasi jika asumsi dalam analisis regresi linier tidak terpenuhi, uji koefisien regresi, uji koefisien korelasi atau pemeriksaan kolinearitas, seleksi model regresi, dan menafsirkan model regresi terbaik. Untuk pemilihan model regresi terbaik ini dapat diterapkan dalam berbagai bidang, salah satunya adalah di bidang pendidikan, yaitu mengenai hubungan *Test of English Proficiency* atau TOEP untuk memprediksi skor TOEFL (Y) siswa di SMA. TOEP terdiri atas *Listening* (TOEP\_LIST yang dimisalkan menjadi variabel  $X_1$ ) dan *Reading* (TOEP\_READ yang dimisalkan menjadi variabel  $X_2$ ). Setelah dilakukan analisis diperoleh persamaan regresi terbaiknya adalah  $\hat{Y} = 254,2444 + 1,9168X_1 + 1,6925X_2$  yang berarti bahwa setiap penambahan skor TOEP\_LIST sebesar 1, maka akan menambah skor TOEFL sebesar 1,9168 dan setiap penambahan skor TOEP\_READ sebesar 1, maka akan menambah skor TOEFL sebesar 1,6925.