



ALLVISION/ISTOCK/THINKSTOCK

BAGIAN 1

Tentang Kriteria Pemilihan Model

Memilih suatu model tidaklah mudah. Setidaknya ada dua kriteria yang harus dipertimbangkan, yakni *goodness of fit* dan kesederhanaan.

PENULIS



BAGUS SARTONO
Dosen di Departemen Statistika
Institut Pertanian Bogor



ALFIAN FUTUHUL HADI
Dosen di Jurusan Matematika
Universitas Jember

BEBERAPA PEKAN yang lalu ada salah seorang teman yang menyampaikan pertanyaan tentang bagaimana memilih model. Dia mengatakan bahwa saat ini dia memiliki tiga buah kandidat model untuk dipilih dan tim di perusahaannya tidak sepakat untuk menentukan mana yang akan digunakan. Selanjutnya penulis menyarankan

untuk mempertimbangkan tidak memilih salah satu yang dianggap paling baik, tetapi menggunakan ketiganya untuk melakukan prediksi dengan teknik *ensemble* (lihat *InfoKomputer* edisi Desember 2016). Saran tersebut tidak sepenuhnya dapat diterima oleh beliau dengan pertimbangan kemudahan, dan tetap pada permasalahan semula yaitu

memilih satu dari beberapa model yang ada.

Model statistika pada prinsipnya adalah suatu upaya dari kita untuk mendekati proses dari suatu kejadian atau fenomena yang terjadi di sekitar kita. Munculnya kejadian tertentu diyakini atau diasumsikan dipengaruhi oleh hal-hal atau kondisi-kondisi tertentu.

Dalam pemodelan selanjutnya kejadian yang menjadi perhatian tersebut dijadikan sebagai target atau *output*, sedangkan kondisi-kondisi yang berpengaruh terhadap kejadian itu dikenal sebagai input. Padanan istilah *output-input* yang banyak digunakan adalah *respons-predictor* atau *dependent-independent*.

Model yang dibangun tentu saja diharapkan sangat mendekati proses yang sesungguhnya (yang sayangnya dalam banyak hal kita tidak pernah mengetahui dengan pasti). Karena ingin sangat



Secara naluriah, manusia lebih menyukai hal-hal yang sederhana dan tidak rumit, sehingga muncul kriteria kedua yaitu kesederhanaan.

mendekati yang sesungguhnya, jelas bahwa kriteria pertama yang dijadikan pertimbangan untuk menentukan model yang baik adalah ketepatan prediksi *output* yang dikenal dengan istilah *goodness of fit*.

**Kriteria Pertama:
Goodness of Fit**

Besaran ukuran *goodness of fit* berkenaan dengan seberapa mirip prediksi kejadian *output* yang dihasilkan oleh model dengan kejadian yang sebenarnya. Pada model-model dengan *output* numerik seperti regresi linear dikenal beberapa ukuran seperti R-squared (dikenal dengan nama koefisien determinasi), MAPE (*mean absolute percentage error*) dan MAD (*mean absolute deviation*).

Ukuran-ukuran tersebut pada dasarnya adalah selisih antara nilai aktual dengan nilai prediksi dari model. Makin kecil selisihnya maka model dikatakan makin baik. Pendekatan pemodelan yang menerapkan kriteria ini umumnya akan menggunakan metode yang didasarkan pada jarak antara nilai data dengan dugaan model, seperti metode *least squares* dengan berbagai macam variasinya. Model dengan R-squared besar, MAPE dan MAD kecil adalah model yang lebih disukai.

Sementara itu kriteria *goodness of fit* pada model-model dengan *output* berupa kelas (*output* yang bersifat kategorik) memiliki ukuran seperti akurasi, sensitivitas, *specificity*, *area under the ROC-curve*, serta beberapa uku-

ran lainnya. Semuanya didasarkan pada kesamaan antara kelas hasil prediksi dengan kelas yang sebenarnya. Makin besar persentase kesamaan antar keduanya maka modelnya disebut makin baik. Terdapat pula ukuran lain yang tidak berpikir hanya sama atau tidak sama antara prediksi dan aktual, tetapi didasarkan pada kejadian aktual dengan peluang terjadinya seperti ukuran *log loss*.

Pada pemodelan statistika atau *statistical learning* yang menggunakan pendekatan *maximum likelihood* juga dikenal ukuran seperti nilai dari *likelihood function*. Untuk kemudahan komputasi, sering yang digunakan adalah logaritma dari *likelihood function*. Ukuran ini pada prinsipnya ingin melakukan pencocokan distribusi sebaik mungkin dari model terhadap data yang dimiliki.

Secara umum (meskipun tidak selalu terjadi), model-model yang makin kompleks atau makin rumit bentuknya akan memberikan *goodness of fit* yang makin baik. Model polinomial dengan derajat tinggi dan nonlinear akan terlihat memiliki ukuran ketepatan prediksi yang lebih baik dibandingkan model linear.

Model yang melibatkan ratusan variabel *predictor* umumnya akan memberikan prediksi yang lebih baik dibandingkan model yang hanya menggunakan 5 atau 6 buah variabel *predictor*. Namun, secara naluriah, manusia lebih menyukai hal sederhana dan tidak rumit sehingga muncul kriteria kedua yaitu kesederhanaan.

**Kriteria Kedua:
Kesederhanaan**

Jika bisa menggunakan model yang sederhana, buat apa yang rumit? Bukankah kalau sederhana biasanya lebih mudah dipahami? Kalau melibatkan sedikit variabel *predictor* kan lebih mudah dan lebih murah dalam mengumpulkan data?

Pertanyaan-pertanyaan di atas adalah pertanyaan yang menyiratkan bahwa secara naluriah manusia lebih menyukai yang sederhana (*parsimonious*). Sehingga kemudian muncul ukuran-ukuran seperti banyaknya parameter model, derajat bebas *error*, banyaknya *node* pada *tree*, dan sebagainya. Beberapa orang berpandangan tidak masalah kalau ukuran ketepatan prediksi berkurang sedikit, asalkan modelnya bisa jauh lebih sederhana.

Berbicara tentang prinsip kesederhanaan, banyak orang merujuk pada apa yang disebut *Occam's razor* yang dikemukakan oleh filsuf William of Ockham (1287–1347). Prinsipnya kira-kira berbunyi bahwa dari sekian banyak hipotesis yang harus diambil adalah yang memiliki paling sedikit asumsi. Dalam beberapa tulisan terdapat kutipannya "*Numquam ponenda est pluralitas sine necessitate*" dengan terjemahan bebas penulis adalah untuk apa menambah rumit jika tidak banyak membantu dan diperlukan.

Dengan dasar pemikiran dari kedua kriteria utama yang disebutkan di atas, kemudian berkembanglah kriteria-kriteria yang kalau ditelisik tidak lain adalah gabungan dari kedua kriteria yang telah disebutkan, yaitu memiliki ketepatan prediksi yang tinggi dengan tetap memperhatikan kesederhanaan bentuk model. 