

Classification

Bab 3 berfokus pada konsep dasar klasifikasi, menggunakan dataset MNIST, yang merupakan kumpulan gambar digit tulisan tangan berukuran 28 x 28 piksel¹. Setelah memuat data dan membaginya menjadi *training* dan *test set*², sebuah *classifier* sederhana seperti Stochastic Gradient Descent (SGD) digunakan sebagai model awal untuk memprediksi digit berdasarkan pola piksel³.



Figure 3-1. A few digits from the MNIST dataset

Awalnya, model dibangun sebagai *binary classifier*, dengan tujuan membedakan apakah sebuah gambar adalah digit "5" atau bukan⁴. Pada tahap ini, disoroti bahwa metrik akurasi saja tidak cukup untuk evaluasi karena data yang tidak seimbang⁵. Oleh karena itu, diperkenalkan metrik yang lebih tepat, yang disajikan di dalam *confusion matrix*⁶.

Metrik Evaluasi Kunci

- *Precision*: Ukuran keakuratan prediksi positif⁷⁷⁷.
- *Recall*: Ukuran kemampuan model untuk menemukan semua kasus positif⁸⁸⁸.
- *F1-score*: Rata-rata harmonik dari *precision* dan *recall*⁹.

Bab ini menjelaskan trade-off yang terjadi antara *precision* dan *recall*, dan menunjukkan bagaimana mengubah *threshold* keputusan dapat memengaruhi kedua metrik ini¹⁰¹⁰¹⁰¹⁰.

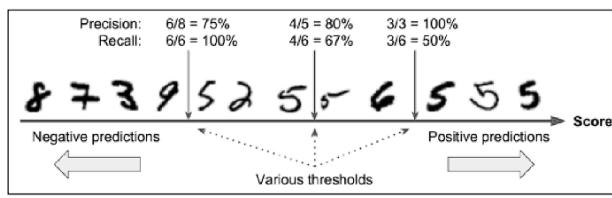


Figure 3-3. Decision threshold and precision/recall tradeoff

Setelah pemahaman tentang *binary classification*, model diperluas menjadi *multiclass classifier* yang mampu mengenali ke-10 digit sekaligus¹¹. Algoritma seperti OvA (One-vs-All) dan OvO (One-vs-One), serta *classifier* seperti Random Forest, digunakan dan dibandingkan performanya¹². Evaluasi model *multiclass* dilakukan menggunakan *confusion matrix* yang diperbesar, di mana pola kesalahan dapat

diidentifikasi, terutama dengan menormalisasi matriks untuk melihat kelas mana yang paling sering salah diklasifikasikan¹³.

Akhirnya, bab ini memperkenalkan *multioutput classification* sebagai bentuk generalisasi dari *multilabel* dan regresi multivariat¹⁴. Contohnya adalah memperbaiki gambar MNIST yang rusak, di mana model dilatih untuk memprediksi 28 x 28 piksel bersih sekaligus dari *input* gambar yang berisik¹⁵. Secara keseluruhan, Bab 3 menekankan pentingnya evaluasi yang tepat menggunakan metrik yang sesuai dengan konteks masalah dan perlunya menggali pola kesalahan melalui *confusion matrix* untuk peningkatan model di masa depan¹⁶.