

# Implementasi Bayes Decision Rule berbasis Maximum Posterior Probability and Minimum Risk

## Deskripsi Masalah

Suatu pabrik pengalengan ikan menggunakan mesin untuk mengalengkan ikan. Mesin membedakan ikan ke dalam dua jenis, yaitu salmon dan seabass. Proses klasifikasi ikan dilakukan dengan mencocokkan fitur *lightness* pada suatu ikan pada suatu basis data yang berisi penggolongan ikan berdasarkan *lightness*. Karena adanya ketidakpastian dalam kategorisasi ikan berdasarkan *lightness* maka klasifikasi dilakukan dengan cara mencari peluang ikan masuk dalam kategori salmon atau seabass berdasarkan nilai *lightness* yang dimiliki. Selain itu perusahaan juga menggolongkan ikan berdasarkan resiko kerugian yang ditimbulkan dari kesalahan klasifikasi ikan. Pada cara ini, ikan digolongkan sebagai salmon atau seabass berdasarkan nilai kerugian yang dapat terjadi jika terjadi kesalahan penggolongan.

## Deskripsi Global Implementasi

Perhitungan *maximum posterior probability* dilakukan dengan rumus berikut.  $P(kelas | lightness) \cong P(lightness | kelas) \cdot P(kelas)$ . Ikan digolongkan ke dalam kelas yang memiliki peluang paling tinggi. Perhitungan  $P(kelas = Salmon)$  dan  $P(kelas = Seabass)$  dilakukan dengan cara menghitung jumlah salmon dan seabass dari semua data set. Perhitungan  $P(lightness | kelas)$  dilakukan dengan cara mencari *range* tempat *lightness* tersebut berada dan dari *range* tersebut dapat diperoleh jumlah salmon (*countSalmon*) dan seabass (*countSeabass*) yang memiliki *range* tersebut dan dibagi dengan jumlah masing-masing kelas.

Klasifikasi berdasarkan *minimum risk* dilakukan dengan cara menghitung kerugian jika suatu ikan diklasifikasikan ke dalam kelas tertentu. Nilai resiko ditentukan oleh fungsi *loss*. Resiko (*risk*) ini dihitung dengan cara  $risk_{total} = risk_{klasifikasi\ benar} + risk_{klasifikasi\ salah}$ . Perhitungan *risk* dilakukan dengan cara  $risk_{klasifikasi} = loss_{klasifikasi} \cdot peluang_{klasifikasi}$ . Fungsi  $loss_{klasifikasi}$  bernilai 0 jika ikan diklasifikasikan dengan benar dan 1 jika dimasukkan ke dalam kategori yang salah.

Implementasi dilakukan dalam bentuk program yang dibuat dengan Bahasa Python. Program ini terdiri dari dua bagian, yaitu data set dan program itu sendiri. Data set adalah berkas CSV yang memiliki tiga kolom, yaitu *range*, *countSalmon*, dan *countSeabass*. Dari file tersebut dapat dihitung  $P(kelas = Salmon)$  dan  $P(kelas = Seabass)$  secara umum. Dari file tersebut juga dapat dihitung  $P(lightness | kelas)$ .

Program diuji dengan menggunakan data yang sama dengan data *range lightness* untuk ikan seabass dan ikan salmon. Hasil uji menampilkan bahwa untuk *range lightness* 1 sampai dengan 5.5, ikan cenderung diklasifikasikan sebagai ikan salmon. Untuk *range* yang lain, ikan digolongkan dalam ikan seabass.

## Kode dan Video Demo

Video Demo dapat diakses di [https://youtu.be/x2q2ihGc\\_wo](https://youtu.be/x2q2ihGc_wo)

Kode program dapat diakses di <https://github.com/anwari32/Pattern-Recognition-HW-1>