

## Abstrak

Menurut data dari *World Health Organization* (WHO), kanker merupakan salah satu penyakit penyumbang utama kematian didunia. Sekitar 8,2 juta orang meninggal karena kanker. Oleh karena berbahayanya penyakit ini berbagai cara dilakukan untuk melakukan pencegahan maupun pendeteksian secara dini. Namun, pendeteksian penyakit kanker sangat rentan terjadinya kesalahan yang dilakukan manusia, seperti yang dimuat pada jurnal kesehatan *British Medical Journal* (BMJ), kesalahan medis menyebabkan 251.454 kematian setiap tahunnya di Amerika Serikat. Guna menanggulangi masalah tersebut digunakanlah teknologi *gene expression* dengan bantuan teknologi *microarray*. Namun pengolahan data *microarray* dirasa kurang memungkinkan, hal ini dikarenakan besarnya dimensi yang dimiliki. Apabila dimensi dikurangi secara tiba-tiba tentu akan merusak informasi yang dimiliki dan akan berakibat data tidak dapat diklasifikasikan. Karenanya dibutuhkan sistem yang mampu mengklasifikasi data *microarray* tanpa menghilangkan informasi penting. Pada penelitian ini, sistem dibangun dengan menggunakan pendekatan *machine learning* yaitu dengan metode *Naive Bayes*. Selain itu guna menyelesaikan permasalahan dalam hal reduksi dimensi, sistem dibangun dengan pendekatan *feature selection*, yaitu dengan metode *Mutual Information*. Untuk mengukur performansi sistem yang dibangun, digunakan *F1-score*. Sistem yang dibangun mampu mengklasifikasi kanker pada data *microarray*, dengan *F1-score* rata-rata sebesar 0.89.

**Kata kunci:** Kanker, *Microarray*, *Naive Bayes*, *Mutual Information*.