Abstrak

Kanker merupakan salah satu dari sekian banyak penyakit mematikan. Namun kanker bukanlah sumber penyakit sesungguhnya karena faktanya setiap orang dewasa setiap hari pasti akan ada 30 miliar sel yang mati dan diantaranya 1 sampai 10 persennya merupakan sel kanker, intinya setiap orang pasti punya sel kanker. Kanker tumbuh di bagian tubuh manapun, salah satunya adalah paru-paru. Berdasarkan data dari *World Health Organization* (WHO) kanker paru-paru adalah salah satu jenis kanker yang banyak menyerang pria, hal itu diperkuat oleh *International Agency for Research on Cancer* (IARC) bahwasanya pada tahun 2012 di Indonesia terdapat 25.322 kasus kanker paru-paru pada pria dan 9374 kasus pada wanita. Pada penelitian ini dicoba untuk mengklasifikasikan data ekspresi gen dari kanker paru-paru. Data ekspresi gen pada penelitian ini tergolong kompleks oleh karena itu perlu dilakukan adanya perekdusian dimensi atau seleksi fitur dan metode yang dipakai adalah *Principal Components Analysis* (PCA). Untuk algoritma untuk klasifikasinya dipakai *Naive Bayes* dan juga *Support Vector Machine* (SVM). Pada penelitian ini kedua algoritma klasifikasi tersebut bekerja secara bersamaan dan akan diteliti seberapa optimalnya hal tersebut. Hasil akurasi dari penelitian menggunakan klasifikasi dengan *Naive Bayes* dan SVM ini adalah 100%.