

# Implementasi algoritma kernel K-Means based Co-clustering untuk memprediksi penyakit kanker paru-paru = :Implementation of K-Means-based K-Means Kernel algorithm to predict lung cancer

Sari Gita Fitri, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20514760&lokasi=lokal>

---

## Abstrak

Kanker adalah penyakit yang disebabkan oleh ketidakaturan perjalanan hormon yang mengakibatkan tumbuhnya daging pada jaringan tubuh yang normal atau sering dikenal sebagai tumor ganas. Kanker terjadi saat sel-sel dalam tubuh membelah diri diluar kendali. Sel-sel abnormal ini kemudian menyerang jaringan terdekat. Salah satu kanker yang paling umum terjadi adalah kanker paru-paru. Kanker paru-paru adalah kanker yang dimulai di paru-paru dan paling sering terjadi pada orang yang merokok. Paru-paru kanan memiliki 3 bagian, yang disebut dengan lobus, sedangkan paru-paru kiri memiliki 2 lobus. Kanker paru-paru merupakan penyebab utama kematian terkait kanker di seluruh dunia dengan 30%-40% terjadi di negara berkembang. Untuk memprediksi apakah seseorang menderita kanker paru-paru atau tidak dapat dilihat dari terdapatnya tumor ganas pada paru-paru yang dapat dilakukan melalui CT scan. Namun, hasil CT scan tidak cukup dalam mendeteksi atau mendiagnosis secara dini terdapatnya tumor ganas di dalam paru-paru. Untuk itu, dapat digunakan machine learning dalam mendeteksi secara dini adanya tumor ganas di dalam paru-paru. Dalam penelitian ini, penulis menggunakan Kernel K-Means based Co-clustering yang merupakan pengembangan dari K-Means based Co-clustering. K-Means mengelompokkan data menggunakan jarak Euclidean. Akan tetapi, jika data yang dipisahkan adalah data nonlinear, maka konvergensi dari data yang dipisahkan tersebut akan kecil dan membutuhkan waktu yang lama, sehingga masalah ini dapat diselesaikan dengan menggunakan fungsi kernel untuk menggantikan jarak Euclidean. Co-clustering mempartisi baris dan kolom dari suatu matriks data secara simultan, sehingga blok yang diinduksi oleh partisi adalah klaster yang baik. Metode Kernel K-Means based Co-clustering memasukkan banyak titik untuk mewakili masing-masing pusat klaster, sehingga titik-titik di dalam klaster saling berdekatan, akan tetapi jauh dari titik yang mewakili klaster lain. Data yang digunakan adalah data kanker paru-paru yang diperoleh dari laboratorium radiologi RSUPN Cipto Mangunkusumo, Jakarta. Hasil akurasi yang diperoleh untuk memprediksi penyakit kanker paru-paru dengan menggunakan metode Kernel K-Means based Co-clustering adalah 94,5%.

.....Cancer is a disease caused by an irregular course of hormones that results in the growth of flesh in normal body tissues or often known as malignant tumors. Cancer occurs when cells in the body divide out of control. These abnormal cells then attack nearby tissues. One of the most common cancers is lung cancer. Lung cancer is cancer that starts in the lungs and most often occurs in people who smoke. The right lung has 3 parts, which are called lobes, while the left lung has 2 lobes. Lung cancer is the leading cause of cancer-related deaths worldwide with 30%-40% occurring in developing countries. To predict whether someone has lung cancer or can not be seen from the presence of malignant tumors in the lungs that can be done through a CT scan. However, CT scan results are not enough to detect or diagnose the presence of malignant tumors early in the lungs. For this reason, machine learning can be used to detect malignant tumors early in the lungs. In this research, the writer uses Kernel K-Means based Co-clustering which is the development of K-Means-based Co-clustering. K-Means groups data using Euclidean distances. However, if the separated data is non-linear data, the convergence will be small and take a long time, so this problem can be solved by

using the kernel function to replace the Euclidean distance. Co-clustering partitioned rows and columns of a data matrix simultaneously, so the blocks induced by partitions are good clusters. Kernel K-Meansbased Co-clustering method includes many points to represent each cluster center, so that the points within the cluster are close together, but far from the points representing other clusters. The data used are lung cancer data obtained from the radiology laboratory of Cipto Mangunkusumo General Hospital, Jakarta. Accuracy results obtained to predict lung cancer by using the Kernel K-Meansbased Co-clustering method are 94.5%.