

# Reduksi Dimensi Menggunakan Autoencoder untuk Deteksi Kanker pada Data *Microarray*

Vina Fadriani Effendi<sup>1</sup>, Adiwijaya<sup>2</sup>, Untari Novia Wisesty<sup>3</sup>

<sup>1,2,3</sup>Fakultas Informatika, Universitas Telkom, Bandung

<sup>1</sup>effendivina@students.telkomuniversity.ac.id, <sup>2</sup>adiwijaya@telkomuniversity.ac.id,

<sup>3</sup>untari@telkomuniversity.ac.id

---

## Abstrak

Kanker merupakan salah satu penyakit dengan jumlah kematian yang tinggi [2]. Beberapa tahun terakhir, *microarray* merupakan teknologi yang digunakan untuk menganalisis dan mendeteksi kanker. *Microarray* digunakan sebagai representasi ekspresi gen pada DNA manusia. Namun, permasalahan utama untuk mendeteksi kanker dengan menggunakan data *microarray* adalah jumlah sampel data yang sedikit dan dimensi yang tinggi. Maka dari itu, data *microarray* harus diproses terlebih dahulu dengan salah satu pendekatan reduksi dimensi sebelum diproses dengan metode klasifikasi. Pada penelitian ini, pendekatan reduksi dimensi yang digunakan adalah *autoencoder*. *Autoencoder* telah berhasil diterapkan dalam rekonstruksi data citra dibandingkan dengan beberapa metode reduksi dimensi lainnya. Arsitektur *autoencoder* akan menjadi pembahasan utama dalam penelitian ini. Selanjutnya, algoritma *Support Vector Machine* (SVM) dan *Multilayer Perceptron* (MLP) digunakan untuk proses klasifikasi. Data *microarray* yang digunakan pada penelitian ini adalah data *leukemia*, *colon cancer*, *lung cancer*, dan *ovarian cancer*. Metode SVM mencapai nilai rata-rata akurasi sebesar 93,81%, sementara nilai rata-rata akurasi dengan metode MLP sebesar 80,95%.

**Kata kunci:** reduksi dimensi, *microarray*, *autoencoder*, SVM, MLP

---

## Abstract

Cancer is one of the disease with high mortality rate [2]. In the last few years, *microarray* is a technology that used to analyze and detect cancer. *Microarray* is used as a representation of gene expression in human DNA. However, the major problems to detect cancer using *microarray* data are the data has very little number of sample data and the high data dimensionality. Therefore, *microarray* data must be processed by one of the approach of dimensionality reduction before it will be processed by classification method. In this research, the approach of dimensionality reduction that used is *autoencoder*. *Autoencoder* has been applied in image reconstruction successfully compared to the other dimensionality reduction method. *Autoencoder* architecture will be the main topic in this research. Next, *Support Vector Machine* (SVM) and *Multilayer Perceptron* (MLP) is used for classification process. *Microarray* datas that used in this research are *Leukemia*, *Colon Cancer*, *Lung Cancer*, and *Ovarian Cancer* datasets. SVM method achieves the average value of 93.81% accuracy, while the average value using MLP method is 80.95%.

**Keywords:** dimensionality reduction, *microarray*, *autoencoder*, SVM, MLP

---