

An enhanced feature selection and cancer classification for microarray data using relaxed Lasso and support vector machine

Aina Umairah Mazlan¹, Noor Azida Sahabudin^{1,*}, Muhammad Akmal Remli^{2,3}, Nor Syahidatul Nadiyah Ismail¹ and Kayode I. Adenuga⁴

¹Faculty of Computing, Universiti Malaysia Pahang, Pekan, Pahang, Malaysia; ²Institute for Artificial Intelligence and Big Data, Universiti Malaysia Kelantan, Kota Bharu, Kelantan, Malaysia; ³Data Science Department, Universiti Malaysia Kelantan, Kota Bharu, Kelantan, Malaysia;

⁴Faculty of Enterprise, Creative and Professional Studies, Farnborough College of Technology, Hampshire, United Kingdom

*Corresponding author: E-mail: azida@ump.edu.my

Chapter outline

1. Introduction	193	3.2.4. Support vector machine	196
2. Related work	194	3.2.5. Generalized multiclass SVM (GenSVM)	196
3. Materials and methods	195	3.3. Evaluation indicator and model development	196
3.1. Datasets	195	4. Results and discussion	198
3.2. Methods	196	5. Conclusion	198
3.2.1. Relaxed Lasso (rL)	196	Acknowledgment	199
3.2.2. Kruskal–Wallis rank sum test (KW)	196	References	199
3.2.3. Sparse group Lasso (SGL)	196		

1. Introduction

Cancer, also known as malignant tumors in medical terms, is the biggest killer disease in a number of countries. The number of deaths and new cancer cases is arising year by year. It is estimated that the cancer burden worldwide will increase by an approximate 22.2 million by 2030 (Mahdavifar et al., 2016). This deadly disease may affect everyone at all ages, but the risk increases with age (Lin et al., 2019). According to the Yu et al., (2018), cancer is the term for diseases triggered by the unregulated production of irregular body cells. Cancer cells may also invade other tissues via the blood and lymph systems to certain areas of the body.

Cancer can be classified into diverse types that are generally named based on the type of cell where the cancer originates (Yu et al., 2018). Lung cancer, breast cancer, prostate cancer, and ovary cancer are some examples of the common cancer types. Lung cancer and breast cancer had the highest mortality rates in men and women, respectively (Bray et al., 2018; Lin et al., 2019). The high morbidity and mortality of cancer mostly is caused from diagnosed disease at a late stage (Dogan, 2019).

As discussed in this section, it is proven that cancer is a serious burden disease in health problems. Therefore, much research has been done to find out better ways to prevent, diagnose, and treat this disease to reduce number of deaths. This