

## 1. Pendahuluan

Cancer is a various diseases characterized by abnormal cell growth that spreads to different parts of the body. These cells divide uncontrollably and in many cases can lead to death [1]. In 2014, cancer prevalence increased sharply, with around 1,665,540 individuals diagnosed and 585,720 fatalities. Currently, there are 1,958,310 new cancer cases and 609,820 related deaths. The incidence of cancer increased by 3% annually from 2014 to 2019, following two decades of decline, leading to an additional 99,000 new cases. [2][3].

Deaths caused by cancer are often due to metastasis and recurrence due to acquired drug resistance. Conventional chemotherapy, including some agents such as alkylating agents, mitotic inhibitors, and antimetabolites alone are not sufficient to treat most cancers, and often cause cancers to become drug-resistant and metastatic to become more aggressive [4]. The drawback of these conventional chemotherapy methods is that apart from being costly, they can damage healthy cells, organs and tissues. Although chemotherapy has reduced morbidity and mortality, almost all chemotherapeutic agents damage rapidly dividing and growing cells [2][5]. This stimulates exploration of novel compounds targeting newly identified mechanisms. One such target relevant in cancer drug development is polo-like kinase 1 (PLK1), which may be utilized either alone or in combination with current chemotherapy treatments [6].

PLK1 regulates mitotic entry and G2/M checks, coordinates centrosome and cell cycles, controls spindle assembly and chromosome separation, has various functions in the mid-spindle zone and during abscission, facilitates DNA replication, and plays a role in cytokinesis and meiosis [6][7]. One alternative for cancer drug development is using in silico techniques, which apply machine learning to PLK1 inhibitors to predict the bioactivity of PLK1 and its derivatives [1][6].

Several studies have been conducted related to the implementation of machine learning to patients in the prediction of cancer and its derivatives. In 2023, Garima and his colleagues conducted QSAR research on tetrahydropteridine derivatives as PLK1 inhibitors using molecular docking and dynamics approaches, with an accuracy of 0.8213 [6]. In 2018, Cai Huang and his colleagues conducted research related to machine learning in predicting the individual response of cancer patients to therapeutic drugs with high accuracy using data on 175 cancer patients with the SVM algorithm and got an accuracy of 82.6% [8].

Based on the literature survey, the accuracy results obtained from the model produce a fairly good value, but these results can be improved. One way to improve the accuracy of the model is to perform a feature selection process and in previous studies there have not been many that use feature selection to determine the derived genes of cancer inhibitors. Methods that can be used in this feature selection process are heuristic methods such as the Gravitational Search Algorithm.

This study aims to predict the bioactivity of PLK1 inhibitor derivatives using the Gravitational Search Algorithm-Support Vector Machine. The Gravitational Search Algorithm method is used for feature selection, while the Support Vector Machine method is used for the prediction process. The Gravitational Search Algorithm, inspired by the physics concept of gravitational force to solve optimization problems, can be used for training and feature selection in models. Meanwhile, the Support Vector Machine, which works by finding the best hyperplane to separate two classes in the feature space, can be used for classification, regression, image processing, and bioinformatics.