1. Introduction

Risk of lung cancer in among 10 years quite variative between the participant inside the Computed Tomography (CT) study cases [1]. From 15% of 68 years old that smokes for 2 packs per day and continuously smoking, to 0.8% for 51 years-old women that smokes 1 packs per day during 28 years before quited for smoking 9 years before [2]. Tobacco smoking is more most dangerous with 80%-90% than the cigarrets smokers. The symptoms are can be continuously if the disease uncontrolled with therapy [3]. However, it will get another symptoms [3]. The most reported symptom of lung cancer is counghing, and follows with the respiratory symptomps which is chest pain, dyspnea, and hemptysis [3]. Patients with lung cancer usually present with many symptoms that respiratory and constitutional includes. There are several recent treatment for the lung cancer, one of them is Chemoterapic [3]. Currently, some big part of chemoterapic agents in use empirically developed [13]. With chemoterapic, it will be have a reason for subset patient that can be surgerical medically using T1 N0 Tumor peripheral [13]. Example for the Small Cell Lung Cancer (SCLC), currently that the recent detection one of them are using Chemoteraphy combined with Chest Radiation, if its only Chemoterapy alone, bad local control from center of the tumors in the cest morbidity will significantly increased (airway, obstruction, dyspnea, cough, superior vena cava syndrome) when relapse [4].

Then, the consequence of Chemoteraphy that mostly and difficult to manage is nauseous vomit (Suh, 2021) [4] . The condition that makes the patient becomes nauseous vomit can cause unbalance of electrolyte, dehydration, anocorrection, and weight decreases (Genc, Can, & Aydiner, 2013) [4]. The things that can reduce the symptom of the nauseous vomit there was in two ways such as in pharmacology and non-pharmacology [4]. For the non-pharmacology there was in several ways such as music therapy, sports, relax, hypnosis, massage, yoga, accupuntur and acupressure [4]. Non-pharmacology ways to control the nauseous vomit can be easy to learn and less cost than the pharmacology [4]. Therefore, for the alternative lung cancer detection is using Machine Learning Gene Expression Data [5].

A study conducted by Jayadeep Pati about Gene Expression Analysis for Lung Cancer Predicition using Machine learning techniques [6]. The study show that machine learning techniques can be uses for Lung cancer detection in a Huge Gene Expression data [6]. Another study about Evaluation of Machine Learning Classifier for Early Stage Prediction of Lung Cancer conducted by Muhammad Imran Faisal and coworkers [7]. It results that the best model classifier are multi layer perceptron, gradient boosting, and support vector machine [7]. All models achieve an accuracy score 88.57, precision score 84.44, recall score 76.57, and f-measure score 80.31 [7]. In a typical microarray experiment, hundreds or even thousands of genes are individually tested for statistical significance. The process of testing multiple hypotheses leads to false-positive results, a phenomenon that must be controlled without significantly affecting the strength of the research. Other statistical challenges, such as determining sample size, overfitting with data, and unstable gene lists, must also be considered. For facing the challenges, we using meta-heurestic feature selection, such as Particle Swarm Optimization. Thus, Study of Implementation metaheuristic feature selection in micro-array data for the smoker detection case is verry rare

In this study, we aim to build an implementation of Particle Swarm Optimization – Support Vector Machine on Gene Expression data for lung cancer identification in smoker person. we use Particle Swarm Optimization for feature selection and Support Vector Machine for model building. The reason why this study using Particle Swarm Optimization because its quite effective on optimization technique and swarm intelligence are behind it, and it successfully applied in many practical engineering problem. The reason we use Support Vector Machine for model building because it able to learn classification data patterns with great accuracy and reproducibility.

Research Problem

The research problem that must be proofen in this research as the following :

1. How to perform feature selection using Particle Swarm Optimization?
2. How to build Lung Cancer Identification in Smoker Person predictive model using Support Vector Machine method?
3. Performance result for applying Particle Swarm Optimization -Support Vector Machine for Lung Cancer Identification in Smoker Person for making the prediction.

Purposes

Based on the preliminary study, so author can get the aim of the research as follows:

1. Apply Particle Swarm Optimization method for making feature selection
2. Build Lung Cancer Identification in Smoker Person predictive model using Support Vector Machine method
3. Know the performance result of applied Particle Swarm Optimization – Support Vector Machine for lung cancer identification in smokers person