1. INTRODUCTION

Human Immunodeficiency Virus (HIV) is a virus that can attack cells and causes a person more susceptible to infections and other diseases[1], and also causes Acquired Immune Deficiency Syndrome (AIDS). This virus has two main species, namely HIV-1 and HIV-2. The HIV-1 was firstly found in chimpanzees and gorillas that lived in West Africa, while the HIV-2 was firstly found in mangabey primates that also lived in West Africa[2]. WHO reported that there were around 770 thousand deaths in 2018 caused by HIV[3]. HIV spread through contact with people via fluid media, such as sharing injecting drug equipment.

Since the spread of the HIV epidemic, several efforts have been made to develop therapies that use HIV-1 antiretroviralsas the target. The knowledge about the role of various components in the HIV-1 life cycle can help researchers to develop new drug candidates. One of the active targets used in this development is HIV-1 protease enzyme[4]. This enzyme is an essential enzyme needed in the assembly and maturation of virions[5]. Aspartic proteinase from HIV-1 is commonly used as a target for AIDS treatment. Many drug candidates are derived by use aspartic proteases as the target. Several available licensed drugs have been used as HIV-1 protease inhibitors, such as saquinavir, indinavir, and ritonavir[4]. Unfortunately, most of the inhibitors are accompanied by side effects if got used in long-term treatment. The most common side effects are HIV protease inhibitor-induced metabolic syndromes, such as dyslipi-demia, insulin-resistance, and lipodystrophy/lipoatrophy, as well as cardiovascular and cerebrovascular diseases[4].

The widespread provision of Anti Retroviral Therapy (ART) also has an impact on the mutation of the HIV virus that can cause resistance. Resistance to HIV can occur due to several factors. one of the factors is due to the very high speed of viral replication that causes the formation of new HIV subtypes that have the possibility of resistance to certain ARVs[6], [7].

Regarding the resistance problem, the activity of HIV-1 protease inhibitors is necessary to be investigated in a laboratory so we can get a new HIV-1 protease inhibitors. However, the investigation of drug activity takes a long time and costly[8]. To overcome this problem, an alternative method is needed to be able to predict the activity before being tested in the laboratory. One of the method that can be used is Quantitative Structure-Activity Relationship

(QSAR) method. The QSAR method is applied to establish a relationship between the activity of a compound and its molecular structure[9]. By using a set of the molecular descriptor as an input, QSAR can be used to predict the activity of HIV-1 protease inhibitors as a drug.

In this study, we aim to build a QSAR model to predict the activity of HIV-1 protease inhibitors. The QSAR model is developed in two stages, i.e. feature selection and prediction model development. The feature selection was conducted by using statistical analysis and gravitational search algorithm (GSA). The GSA was chosen because of the ability to improve the prediction accuracy by selecting a set of appropriate descriptors[18]. Then, the prediction model was developed by using artificial neural network (ANN) that has been widely used in QSAR studies[19]–[21].

Formulation of the problem

Based on the description above, several main problems in this Final Project can be formulated, including:

1. How to do feature selection using the gravitational search algorithm?
2. How to build a prediction model using the gravitational search algorithm - artificial neural network method?
3. How is the performance of artificial neural network methods in predicting the activity of HIV-1 protease inhibitors?

Purpose

The objectives to be achieved in this thesis are as follows:

1. Use the gravitational search algorithm to perform feature selection.
2. Build a prediction model using the gravitational search algorithm - artificial neural network method.
3. Use artificial neural network methods to predict the activity of HIV-1 protease inhibitors.

Scope of problem

In this Final Project, the object of research is limited to the following scope:

1. HIV-1 protease as a target is used to look for inhibitors.
2. The descriptors used are two-dimensional (2D) descriptors.