

**HIV/AIDS severity prediction system using Artificial Neural Network**

**Murhula Mwambali Bienfait-099271**

**SUPERVISOR:** Edmond Odhiambo

A proposal submitted in partial fulfillment of the requirement for the degree of Bachelor of Science in Informatics and Computer Science

Date of Submission: June 2020

# **Abstract**

The Human Immunodeficiency Virus (HIV) continues to be a major global health issue. Since its discovery in 1981, HIV has claimed more than 32 million lives. Due to its dynamic nature of morphing into new forms after a very short period, it has no cure as for now. If left untreated HIV would end up developing into its most severe stage the Acquired Immune Deficiency Syndrome (AIDS). AIDS is a syndrome that weakens a patient's immune system and leaves him vulnerable to other infections.

Antiretroviral Therapy (ART) has been used as a combative mechanism to slow down the virus progression in the human body. ART consists of a combination of antiretroviral drugs. The progression of the virus is tracked by counting Cluster Differentiation 4 (CD4) positive cells, and the amount of the virus in the blood (viral load) every 6 months.

This research introduces the use of Artificial Neural Network (ANN) to predict HIV/AIDS viral load levels over a given period. The ANN will be accessed via a web application created for demonstration purposes. This becomes far more efficient in comparison to the biannual viral load test; since it takes into consideration patient lifestyle, diet, alcohol consumption, and the limited resources of the hospital.

**Keywords: Artificial Neural Network, CD4+, Viral Load**

# **Contents**

[**Abstract** 3](#_Toc52255102)

[**Contents** 4](#_Toc52255103)

[**List of Tables** 6](#_Toc52255104)

[**List of Figures** 7](#_Toc52255105)

[**Chapter 1: Introduction** 1](#_Toc52255106)

[1.1 Background 1](#_Toc52255107)

[1.2 Problem Statement 2](#_Toc52255108)

[1.3 Aim 2](#_Toc52255109)

[1.4 Research Objectives 3](#_Toc52255110)

[1.5 Research Questions 3](#_Toc52255111)

[1.6 Justification 3](#_Toc52255112)

[1.7 Scope and Limitation 3](#_Toc52255113)

[**Chapter 2: Literature Review** 4](#_Toc52255114)

[2.1 Introduction 4](#_Toc52255115)

[1. 2.2 Methods for measuring HIV/AIDS Progression 4](#_Toc52255116)

[**2.2.1. Viral load** 4](#_Toc52255117)

[**2.2.2. The cluster of Differentiation 4** 4](#_Toc52255118)

[2.3. Current HIV/AIDS Predictions System 5](#_Toc52255119)

[**2.3.1. Random Forest Algorithm** 5](#_Toc52255120)

[**2.3.2. Support Vector Machine** 7](#_Toc52255121)

[**2.3.3. Prediction of HIV/AIDS Status using Artificial Neural Networks** 8](#_Toc52255122)

[**2.3.4. K-Nearest Neighbor** 9](#_Toc52255123)

[2.4. Conceptual Design 10](#_Toc52255124)

[**Chapter 3: Research Methodology** 12](#_Toc52255125)

[3.1 Introduction 12](#_Toc52255126)

[3.2 CRISP-DM Development Methodology 12](#_Toc52255127)

[3.3 Research Design 13](#_Toc52255128)

[**3.3.1 Business Understanding** 13](#_Toc52255129)

[**3.3.2 Data Understanding** 13](#_Toc52255130)

[**3.3.2 Data Preparation** 13](#_Toc52255131)

[**3.3.3 Modelling** 13](#_Toc52255132)

[**3.3.4 Evaluation** 13](#_Toc52255133)

[**3.3.5 Deployment** 13](#_Toc52255134)

[**References** 16](#_Toc52255135)

# **List of Tables**

[Table 1: RNA viral load 10](#_heading=h.1y810tw)

# **List of Figures**

[Figure 1ROC curves for the RF models with and without genotype 6](#_Toc52255040)

[Figure 2Support Vector Machine Example 7](#_Toc52255041)

[Figure 3Neural Network Structure(Adapted from(Steven walczack,2004)) 8](#_Toc52255042)

[Figure 4KNN example (Adapted from(Jose,2018)) 9](#_Toc52255043)

[Figure 5Conceptual Framework 11](#_Toc52255044)

[Figure 6 CRISP-DM Development Methodology 12](#_Toc52255045)

# **Chapter 1: Introduction**

## Background

According to the World Health Organization (WHO), HIV remains a major global public health concern. The virus has claimed more than 32 million lives since its discovery in 1981. HIV is a human immunodeficiency virus. (Organisation, n.d.) . HIV causes immunodeficiency syndrome or AIDS syndrome to develop if it is not treated. HIV is a retrovirus and, due to its dynamic nature of morphing into news forms after a very short time, it has no cure.

Sub-Saharan Africa alone represents an estimated 69% of all HIV-positive diagnoses, and 70% of all AIDS-related deaths in 2011. (Ayesha B.M. Kharsany and Quarraisha A. Karim, 2016). HIV is the leading cause of death in sub-Saharan Africa. Rural Areas, where there is limited or no access to health and welfare services, carries the major weight of the subcontinent's HIV/AIDS burden.

Kenya is the third largest HIV epidemic worldwide, with an estimated 1.6 million people living with HIV in 2018. (avert., 2018) The first occurrence of the virus in Kenya was in 1984. By 1990, HIV was one of the leading causes of mortality in Kenya. (avert., 2018). Kenya's HIV outbreak affects 4.7% of its general population, but groups of men who have sex with men, women, sex workers, and injection drug users are still more susceptible to infection. (avert., 2018). The main factor contributing to the high incidence of HIV/AIDS in Kenya has been attributed to the high level of poverty among Kenyans, where over 50 percent of the population lives on less than $1 per day. (Programme, 2014)

There exist two major phenotypes of HIV; namely, HIV-1 and HIV-2. HIV-1 is the most prevalent type of virus worldwide. HIV is further subdivided into three strains. It comprises group M (main), O (outlier), and N (Non-M, or Non-O). Relatively uncommon HIV-2 is mostly spread in West Africa and is becoming more common in India, although still in small numbers. It is subdivided into two groups, labeled as HIV-2 A and HIV-2 B. (D S Callaway, n.d.)

So far HIV has been managed through several therapies. Antiretroviral the. ART combines several antiretroviral (ARV) drugs to maximally suppress the progression of the virus. ART aims to slow the virus progression from the Acute HIV infection to the clinical latency, to the Acquired immunodeficiency syndrome (AIDS) stage (Organization, 2013). Acute primary infection is nearly one to four weeks after getting HIV. At this stage, the patient has a large amount of the virus in the blood and is very infectious. Clinical latency, often referred to as asymptomatic HIV infection, at this stage of the viral infection, is still active but morphs at a low rate. Without medication, an HIV positive patient at the clinical latency stage may last about a decade. AIDS is the most severe phase of virus infection. The common symptoms of AIDS include chills, fever, sweats, swollen lymph glands, weakness, and weight loss. At this stage, the CD4 cell count has dropped below 200 cells/mm. CD4 cell count is a test used to measure how many CD4 cells (white blood cells) that you have in your blood. AIDS-infected patients have a high viral load and are highly infectious. (Prevention, 2019)

Currently, resources are being allocated to sponsoring activities to predict an improvement in a patient's viral load. These activities focus on using the nucleotide sequences of the Reverse Transcriptase (RT), the viral load, the CD4 count cell, and the Protease (PR) to predict HIV short term progression. The above variable has shown a high correlation to the amount of HIV within the patient's body(viral load). (Murell, 2020)

## Problem Statement

Currently, patients are to visit the hospital at least twice a year for viral load tests. The test is conducted clinically through the use of various clinical markers within the hospitals’ laboratories. This approach is so far not efficient since it fails to appreciate three factors such as the unavailability of the patient to commit to this biannual appointment. Secondly, the patient's lifestyle (diet and alcohol consumption) affects the immune system in the short term. Lastly, the limited resources to conduct these tests within a larger populated HIV infected population.

CD4 count changes are significantly correlated to a group-level viral load but are limited in the identification of immunological failure of the patient. CD4 cell count is an inadequate solution to the measurement of viral load for early diagnosis of virological failure. (Motasim Badri, 2008)

## Aim

This study aims to predict the likelihood that the infection of an HIV patient will become less severe.

## Research Objectives

1. To investigate how human expert carry out the HIV laboratory test
2. To review existing predictive models for viral diseases
3. To investigate the use of Artificial Neural Network in the medical field
4. To design an Artificial Neural Network and to develop a viral load predictive system
5. To test the proposed solution on a web application to prove the concept.

## Research Questions

1. How is the HIV/AIDS progression measured?
2. What are the factors influencing the progression of HIV among HIV positive patients?
3. What are the prediction models currently used to predict viral diseases?
4. How can Artificial Neural Network be used to predict HIV/AIDS viral load levels?

## Justification

The benefits of artificial intelligence in the medical field can range from simplifying the tasks of the health care professional, energy, and time, to detecting patterns that would otherwise have been missed by the physician. With this, artificial intelligence can help the physician make necessary changes to the patient’s HIV therapy regimen at any given time.

It is worth noting that the use of predictive analytics in the medical field can lead to an improvement in diagnostic medical strategies. This would automatically lead to a better understanding of the virus progression and ultimately to better outcomes for the patients.

In the medical field, therefore, a system that combines the power of artificial intelligence and predictive analysis can be used to improve HIV patient care, along with the skills and input of physicians.

## Scope and Limitation

The focus of the project is to implement and test a machine learning algorithm with a web application that will be used to demonstrate a proof of concept. We will also focus on the analysis of already existing datasets to predict HIV viral load count as opposed to developing our dataset.

Because of the sensitive nature of the data under consideration, getting primary data is very challenging as most of them are very confidential. The data was obtained from Stanford HIV/DB. Due to this, the model is more likely to perform better in the United State of America as opposed to Africa.

# **Chapter 2: Literature Review**

## 2.1 Introduction

This chapter examines the methods for measuring HIV / AIDS progression in the human body. We will also evaluate the existing systems that are used to predict the viral load level in HIV positive patients and highlight the results of the research.

## 2.2 Methods for measuring HIV/AIDS Progression

HIV/AIDS progression can be measured in two ways, by measuring the CD4+ count and the viral load levels in the human body

### **2.2.1. Viral load**

Viral load is a term that describes the amount of virus in an organism, this in terms of virus particles per milliliter (mL). Virus load is one of the best measures to assess the progression of HIV disease and to compute how well ART medication works. Viral loads may be expressed either as copies of Ribonucleic acid ( **RNA**) per milliliter of blood (copies/ml) or as logs. The viral load is considered to be undetectable when the viral load is too small to be assessed. However, being an undetectable viral load does not mean that HIV has been eradicated. HIV may not be detectable in the blood, it may be found in the female genital secretions, tissues, lymph nodes, and cerebrospinal fluid.

The viral load test is instrumental in determining the necessity of a change in the ART regimen for an HIV positive patient. This prevents unnecessary ART regimen changes and can encourage the patient to adhere to the ART regimen given.

### **2.2.2. The cluster of Differentiation 4**

The Cluster of Differentiation 4 (CD4+) is a glycoprotein found on the surface of the T-cells. CD4 count is a blood test to check the amount of CD4 cells per milliliter (mL) of blood. The test is to estimate the number of white blood cells per milliliter of blood. HIV attacks the CD4 cells in the blood. This causes the number of CD4 cells in the body to drop, making it challenging to fight the infection.

CD4 cell count predicts clinical progression and survival in a brief period. (Claris Shoko, 2019). The downside in using CD4+ count as a measure of HIV severity in HIV positive patients is that their detection comes after resistance to the drug has been developed. This necessitates a switch of the regimen to a second or third-line regimen, which is more expensive. (AIDS, 2016)

## 2.3. Current HIV/AIDS Predictions System

Now, we will look at the various existing prediction systems, analyzing their strengths and weaknesses.

### **2.3.1. Random Forest Algorithm**

Random Forest is a supervised ensemble learning algorithm that is utilized for both classifications as well as regression tasks. (Chauhan, 2020). As a classifier combination, it utilizes the L tree-structure based classifier {h(X, Ѳn), N=1, 2, 3…L}, where X defines the input data and {Ѳn} is a set of identical and dependent distributed random vectors. Each decision tree is created through a random selection (bagging) on the datasets. Then get the prediction from each of the trees then select the best solution utilizing voting. (Chauhan, 2020)

The Random Forest algorithm has been used in the prediction of HIV positive patient's severity following the change in therapy. (Susan Kamal, 08 Apr 2020) Used seven input variables for training and prediction. These inputs variables included: CD4 count, nadir CD4 count, duration of ART, CD4 count, and baseline viral load.

The second study (Andrew D Revell, 2012) Model was developed with 3188 treatment change episodes (TCE) from North America, Western Europe, Australia, and tested with 100 TCEs. Two models were developed; one considering the genotype and the other ignoring the genotype. Without the genotype, the model resulted in the area under the curve (AUC) of 0.88 while that with the genotype had an AUC result of 0.86. With the TCE originating from Romania, the AUC result was 0.60. The result suggested that this approach might be generalizable and the model useful for treatment-experienced patients in a limited-resource setting where genotyping may not be readily available.

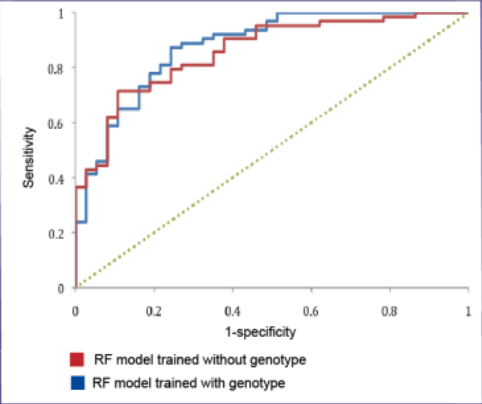
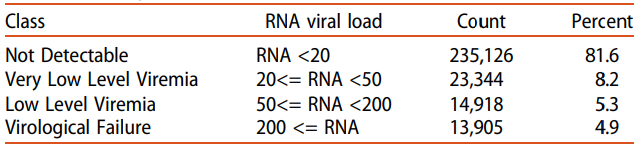


Figure 1:ROC curves for the RF models with and without genotype

Lastly, (Sezerman, 2016) incorporates the sequence and structural features for predicting HIV resistance with a random forest classifier. Tested on the mutants of HIV-1 protease and reverse transcriptase; the model showed accuracy measures ranging between 98-99.2%. The focus of the research was to identify the best attributes for the determination of HIV resistance. Unlike the linear sequence representation, this research had combined both sequence and structure features. The features that were taken into consideration include, hydrophobicity measure, evolutionary conservation, flexibility measure, disordered proteins, and amino acid volume information, for the sequence features, whereas the structural features considered the 2D and 3D representation combined with the contact energies

Table 1: RNA viral load



### **2.3.2. Support Vector Machine**

Support Vector Machine (SVM) is a statistical technique that works by embedding data into a high dimensional vector. The SVM algorithm aims at finding a hyperplane in an N-dimensional space that separately classifies the data points.

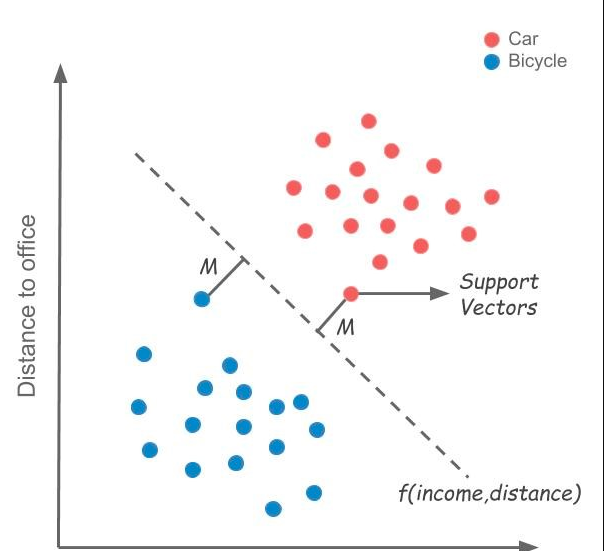


Figure 2Support Vector Machine Example

(Sezerman, 2016) Implemented a support vector machine classifier with both sequence and structural features for drug resistance. The study obtained accuracy measures of 95-96%. To obtain the best possible classification, the parameter was tweaked. To overcome SVM sensitivity regarding class imbalance, (Sezerman, 2016), applied sampling techniques to the datasets, this to avoid misclassification.

### **2.3.3. Prediction of HIV/AIDS Status using Artificial Neural Networks**

Artificial Neural Networks (ANN) are biologically inspired computational models that tried to mimic the human brain. ANN might have up to thousands of processor units. Processing elements (also known as neurode or perceptron) are arranged in layer and connected to other neurons containing the activation function. The output of one layer serves as the input to the next layer and possibly to other layers.

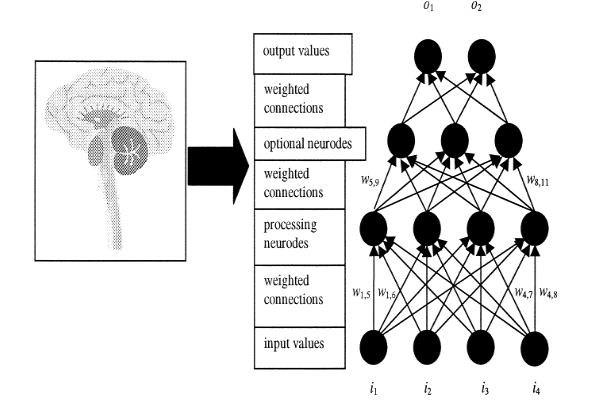


Figure 3Neural Network Structure(Adapted from(Steven walczack,2004))

(Emuoyibofarhe, 2016), introduced a test for computing the accuracy of HIV/AIDS diagnosis; using Artificial Neural Network. Enzyme-Linked ImmunoSorbent Assay (ELISA) experiments were used as inputs to the neural network. ELISA is an immunological test used for screening antibodies, antigens, proteins, and glycoproteins in blood samples. The research has successfully developed neural networks for predicting and hence the diagnosis of HIV using the sample results of the ELISA tests.

The accuracy of the model was 94% based on 9 epochs. This successfully confirmed the need for an Artificial Neural Network to diagnose HIV based on ELISA test result and is to be used to assist medical practitioners to give a more accurate diagnosis of HIV.

(Lee, n.d.) Applied ANN to classify and predict the health status of HIV/AIDS-related patients. (Lee, n.d.) The research utilizes the AIDS Cost and Services Utilization Survey (ACSUS), a longitudinal study of persons with HIV-related disease in which a combination of personal interviews and abstraction of medical records is used. 9 variables were selected after controlling them by the variance inflation factor (VIF) method for detecting multicollinearity between the input variable. (Lee, n.d.)

Two indices were to evaluate the reliability of the test: positive predictive power (PPP) and negative predictive power (NPP). The PPP is the true positive probability. The PPP indicates that the probability of the well-health status being present given that health status is predicted as well. Where the NPP is the true negative probability. NPP indicates that of not-well-health status being present given that health status is predicted as not-well. The PPP correct classification varies between 86.4% and 88.7% where NPP varies from 77.0% to 79.6%. (Lee, n.d.).

### **2.3.4. K-Nearest Neighbor**

K-Nearest Neighbor (KNN) is a supervised machine learning that can be used for both supervised classification problems as well as for regression problems. The neighbors are found by the use of distance measures. K-value determines the number of neighbors, which are to be considered for the case, based on their distance. The obtained distances are ranked, and the K smallest distances are as neighbors.

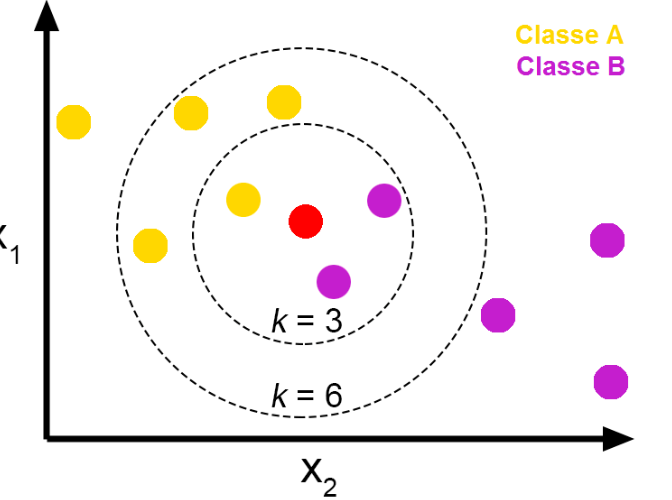


Figure 4KNN example (Adapted from(Jose,2018))

In (Dandan Tang, 2018) ‘s research the k nearest algorithm was used. The goal of the study was of using a data mining algorithm to establish the identification model of HIV infection. A confusion matrix and other indicators including accuracy, sensitivity, specification, precision, and recall of the receiver operating characteristic (ROC) curve were used to appraise the performance of the model. A 10 fold cross-validation was applied to KNN validation. AUC scores for the KNNW was 0.9747 and classification of 0.91528, a precision of 0.89513 and a recall of 0.855135

(ChenHsiang Shen, 2016), also applied the KNN algorithm. To train the learning model the k-value was set to 6, with 210 dimension vectors of training samples paired with phenotypic data. KNN algorithms are faster, but the validation of the model is very slow when working with a large number of cases. The **R2** values were obtained ranging from 0.719 to 0.9188 across 5 –folds, with better precision accuracy of the model when **R2** was set to 1.

## 2.4. Conceptual Design

The data needs to be cleaned for use in the training of the ANN system. With this only specific features are to be extracted; including the viral load, the CD4+ count for a specific patient, and the phenotype. The ANN will consist of three layers of nodes. The input layer consists of input variables given to generate the prediction. Hidden since does not interact directly with other layers, but works on data being transformed by the previous. It serves as a link to the output layer that displays the final prediction. After training of the neural network, the system will provide an estimated prediction of the patient viral load over a given time.



Figure 5Conceptual Framework

# **Chapter 3: Research Methodology**

## 3.1 Introduction

This chapter outlines the research methodology to be used. This include the sample population, the sampling method and the research design selected. The data collection and analysis techniques of data are also explained.

## 3.2 CRISP-DM Development Methodology

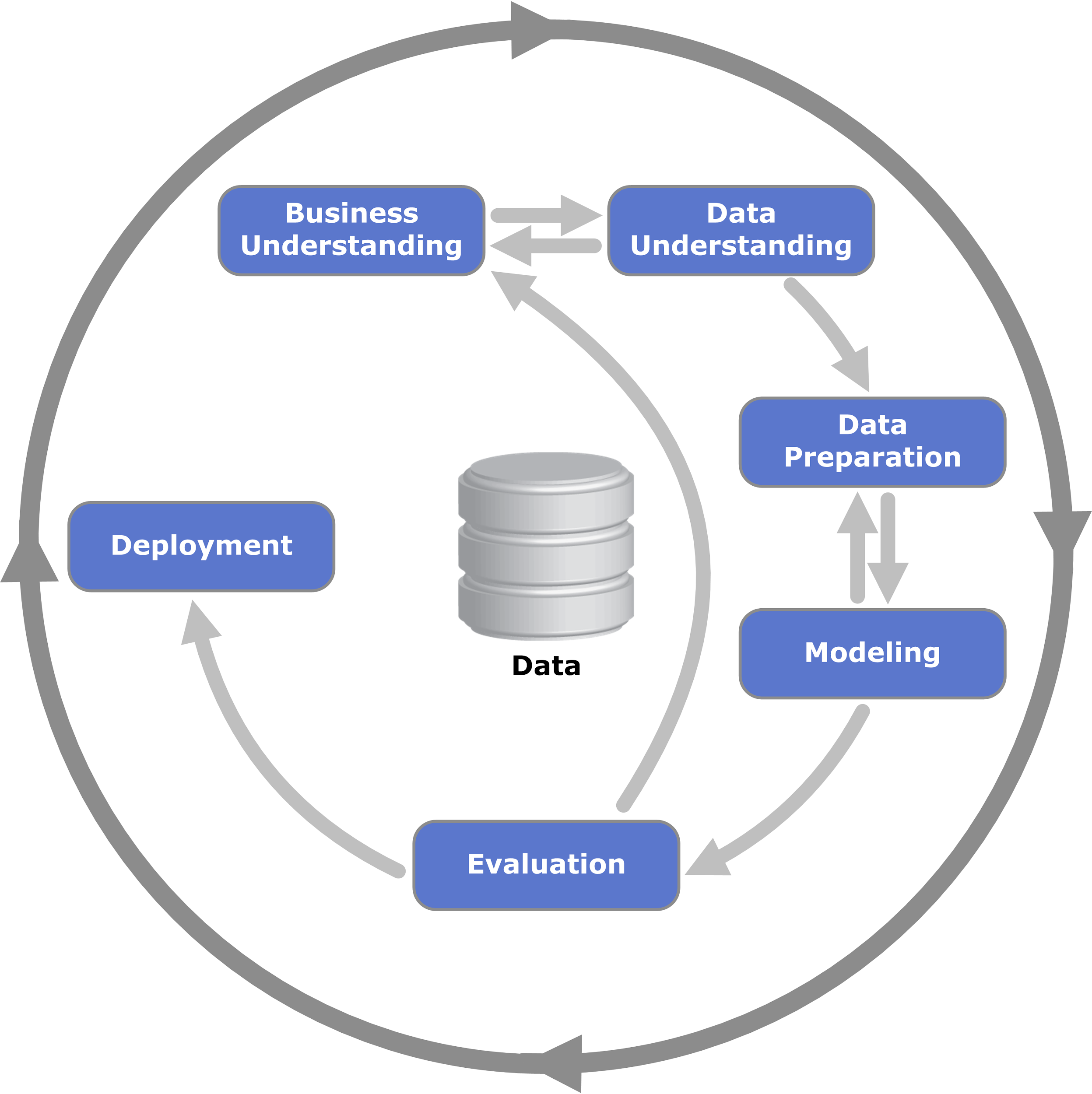


Figure 6 CRISP-DM Development Methodology

CRISP-DM stands for Cross Industry Standard Process for Data Mining. It is a methodology created in 1996 to shape data mining projects. CRISP-DM involves a series of steps; six to be precise.

The first step of the CRISP-DM methodology is the business understanding. This step focuses on the understanding of the project objectives, requirements and defines the scope of its activities. The second step is the Data Understanding; with a start on initial data collection. Its objective is to get familiar with the data, to identify the data quality problem and lastly to discover insight into the data. The third step is the data preparation, it involves cleaning the data, generating new features and merging and aggregating data. The fourth step is modelling. Modelling techniques are selected and applied. The fifth step is the evaluation. Here is where it is up to verify that the results are correct. This would be the time to evaluate model performance. The sixth and final step is the deployment; which consists of presenting the results in a useful and understandable manner.

## 3.3 Research Design

Searching for potential solutions on a defined problem is an integral part of design science research . This research followers a mixture of the quantitative and qualitative research methods. The qualitative research methods ought to get understanding of data; and the quantitative research methods will be applied to get enough information for the research in order to feed our model.

This research is prescriptive research, and seeks to offer potential solutions to problems.

### **3.3.1 Business Understanding**

This is the first stage of the CRISP-DM software methodology. At this stage the objectives of the research with the algorithm requirements are defined, jointly with the scope and the application requirements. These are explained in chapter one when discussing the objectives.

### **3.3.2 Data Understanding**

This stage starts with the actual data collection. The data that is to be used in building the neural network regressor; was obtained from HIV Stanford database. The data is in an XML format, and consists of 1527 XML files, with their individual TCEs.

We also check the quality of the data; such as data completeness and values distribution. This is a critical part of the project because it defines how viable can be the final result.

### **3.3.2 Data Preparation**

This stage involved cleaning of data, generating new features, feature selection, merging and aggregating data. The TCE data was first extracted from the XML file using a python script developed. The duplicates were removed, and 15 different Data Frame created, were merged together into one CSV file; that will be used for training, testing and model validation.

### **3.3.3 Modelling**

The neural network regressor model will be developed using scikit-learn library, a free software machine learning library for the Python programming language.

### **3.3.4 Evaluation**

### **3.3.5 Deployment**

The deployment of our model will consist of presenting the results in an useful and understandable manner. This will be done through developing a web application; using of Flask, a micro web framework written in Python, and Streamlit’s open-source app framework

# **References**

(National AIDS and STI Control Programme, 2. & UNAIDS, 2017. s.l.: s.n.

AIDS, U. N., 2016. *The need for routine viral load testing,* s.l.: United Nation AIDS.

Andrew D Revell, 1. E. D. W. M. Y. P., 2 (Alan R. Hevner, 2004)012. The use of computational models to predict response to HIV therapy for clinical cases in Romania. *GERMS 2012.*

avert., 2018. *Avert.org.* [Online]   
Available at: https://www.avert.org/professionals/hiv-around-world/sub-saharan-africa/kenya  
[Accessed 20 june 2020].

Ayesha B.M. Kharsany and Quarraisha A. Karim, 2016. HIV Infection and AIDS in Sub-Saharan Africa: Current Status, Challenges and Opportunities. *Bentham Open.*

Chauhan, N. S., 2020. *medium..* [Online]   
Available at: https://towardsdatascience.com/random-forest-a-powerful-ensemble-learning-algorithm-2bf132ba639d  
[Accessed 15 june 2020].

ChenHsiang Shen, X. Y. R. W. H. &. I. T. W., 2016. Automated prediction of HIV drug resistance from genotype data. *BMC Bioinformatics,* Volume 278, p. 17.

ChenHsiang Shen, X. Y. R. W. H. &. I. T. W., 2016. Automated prediction of HIV drug resistance from genotype data. *BMC Bioinformatics ,* Volume 278, p. 17.

Claris Shoko, D. C., 2019. A superiority of viral load over CD4 cell count when predicting mortality in HIV patients on therapy. *bmc infectious diseases.*

D S Callaway, R. M. R. a. M. A. N., n.d. Virus phenotype switching and disease progression in HIV-1 infection..

Dandan Tang, 1. M. Z. J. X. X. Z. Y. L. L. F. K. W.. Y. Z., 2018. Application of Data Mining Technology on Surveillance Report Data of HIV/AIDS High-Risk Group in Urumqi from 2009 to 2015. *Panayiotis Vlamos,* Volume 2018(10 Dec 2018), p. 17.

Emuoyibofarhe, O. J. O. J. B. O. I. O. P. O. P. &. A. J., 2016. *Prediction of HIV/AIDS Status using Artificial Neural Network.,* Lagos: University of Lagos.

https://aidsinfo.nih.gov/understanding-hiv-aids/fact-sheets/19/46/the-stages-of-hiv-infection, n.d. [Online].

https://www.unaids.org/en/resources/documents/2019/2019-UNAIDS-data, n.d. s.l.: s.n.

https://www.who.int/hiv/pub/guidelines/clinicalstaging.pdf, n.d. s.l.: s.n.

Lee, N. K. K. a. C., n.d. A Neural Network Application to Classification of. *Journal of Medical Systems,* Volume Vol. 21, No. 2.

Motasim Badri, c. a. S. D. L. a. R. W., 2008. Utility of CD4 cell counts for early prediction of virological failure during antiretroviral therapy in a resource-limited setting. *BMC Infectious Diseases.*

Murell, D., 2020. *healthline.* [Online]   
Available at: https://www.healthline.com/health/hiv-aids/cd4-viral-count#relationship-between-cd-count-and viral-load  
[Accessed 18 june 2020].

Organisation, W. H., n.d. *World Health Organisation.* [Online]   
Available at: https://www.who.int/gho/hiv/hiv\_text/en/

Organization, W. H., 2013. *who.int.* [Online]   
Available at: https://www.who.int/hiv/topics/treatment/art/en/  
[Accessed 18 june 2020].

Prevention, C. f. D. C. a., 2019. *Center for Disease Control and Prevention.* [Online]   
Available at: https://www.cdc.gov/hiv/basics/whatishiv.html  
[Accessed 18 june 2020].

Programme, N. A. a. S. C., 2014. *KENYA HIV ESTIMATES,* Nairobi: National AIDS and STI Control Programme.

Sezerman, Z. K. a. O. U., 2016. Prediction of HIV Drug Resistance by Combining. *IEEE/ACM Transactions on Computational Biology and Bioinformatics,* Volume 15.

Susan Kamal, J. U. M. C. H. L. R. K., 08 Apr 2020. Random forest machine learning algorithm predicts virologic outcomes among HIV infected adults in Lausanne, Switzerland using electronically monitored combined antiretroviral treatment adherence. *AIDS Care.*