Title: Prediction of malaria Diseases' High Prevalence Time and place using

Machine Learning Algorithm in Wolaita Zone, Ethiopia

Submitted by: Fetenech Meskele

Abstract

Malaria is one of the leading causes of death in Ethiopia. Plasmodium falciparum malaria still

poses one of the greatest threats to human life with over 200 million cases globally leading to

Half-million deaths annually. Of these, 90% of cases and of the mortality occurs in sub-Saharan

Africa, mostly among children. Though there are many efforts to control malaria, the complexity

of the problems is still very severe. So there is a need to investigate in detail the synergic effect

Of risk factors with temperature, altitude, type of visit and malaria type and their causes of death.

Hence in this research an attempt is made to determine the hierarchical importance of different

Risk factors and their patterns on malaria death occurrence. The main aim of this paper is to develop

And validate a computational system that exploits the predictive power of current Machine

Learning approaches on 3-years of prospective data from the high-transmission holoendemic

Malaria Ruler-densely-population of wolita Zone. Our dataset of >9x104

Screened study participants attending our clinical and community services from 2020 to 2022

Contains monthly prevalence, temporal, environmental and host features. Our Locality-specific

Elastic-Net based Malaria Prediction System (LEMPS) achieves good generalization

Performance, both in magnitude and direction of the prediction, when tasked to predict monthly

Prevalence on previously unseen validation data (MAE<=6x10-2, MSE<=7x10-3) within a range

of (+0.1 to -0.05) error-tolerance which is relevant and usable for aiding decision-support in a

holoendemic setting. LEMPS is well-suited for malaria prediction, where there are multiple Features which are correlated with one another, and trading-off between regularization-strength L1-norm and L2-norm allows the system to retain stability. Data-driven systems are critical for regionally-adaptable surveillance, management of control strategies and resource allocation across stretched healthcare systems.

Background

Malaria remains one of the most serious public health problems in the world. In 2015, there were an estimated 429,000 malaria-related deaths (range 235,000–639,000) worldwide. Most of these deaths occurred in Africa (92%), Southeast Asia, and the Eastern Mediterranean region]. In Ethiopia, malaria is endemic across three-quarters of the landmass, and an estimated 68% of the Population lives in these affected areas. Over the last decade, the burden of the disease has Declined significantly, which may have resulted from the improved coverage of high impact? Interventions including prompt treatment of cases using artemisinin-based combination therapy (ACT), prevention and control of malaria among pregnant women using intermittent preventive therapy (IPT), and use of vector control methods, i.e. insecticide-treated bed nets (ITNs) and Indoor residual. Malaria is one of the main public health challenges in Ethiopia that hinder the Productivity and development of the country. In 2021, Ethiopia is on track to minimize the Incidence of malaria by 40% as per its 2020 malaria reduction strategy. Currently, Ethiopia is Working to eliminate malaria in 2030 by extending the 2020 strategy. Hence, this study aimed to

Analyze the trend of malaria prevalence in Humbo Woreda, Abelal Abaya and Dugna Fango in Wolita Zone Ethiopia, 2020 to 2022.

General Objective

The general objective of this research is to investigate the potential applicability of Machine Learning technologies to discover hidden knowledge in malaria data and develop a model that can help in predicting trend of malaria causes and deaths in order to address malaria health related issue. It focuses on determining the most hierarchical risk factors and their patterns that affect malaria occurrence of deaths and type of cases identification.

Specific Objectives

In order to achieve the above general objective properly, the research attempt to achieve the following specific objectives.

To review different literatures that can support the study. These are used to:

• Indicate how malaria is a critical public health issue.

•Show how Machine Learning technologies easily solve and give support to malaria intervention and control program especially by build a model/rules

•To select data mining techniques suitable for investigation and show the characteristics of disease pattern which is important for the prevention and control interventions of malaria in the country.

•To pre-process the data set for training and testing in order to adjust inconsistent data encoding, accounting for missing values, and deriving missing fields from existing ones.

• To build a prototype (predictive model) using Machine Learning software so as to uncover the knowledge in the training datasets.

• To evaluate the performance of the predictive model built using test dataset and propose further research direction.

Methodology

Methods Study area this study was carried out at Wolita Zone in three ruler Woredas, in the

Southern part of Ethiopia. Malaria is the most prevalent seasonal disease in the area, which peaks in May to December. Study design A retrospective malaria data review was conducted to determine the

Three years (2020–2022) malaria prevalence trends from January to December 2020 at the

November Health Centre. Data collection three-year malaria prevalence data were obtained from

the three woreda Health Centre for the period 2010–2021. Throughout the reviewed period,

Microscopy was used as the gold standard to confirm Plasmodium parasite presence by

Examination of peripheral smears of stained blood films, as per the WHO protocol. In Ethiopia,

All hospitals and health centers follow a standard operating procedure (SOP) for blood smear

Preparation, staining and blood film examination for malaria parasite detection throughout the

country. The woreda's Health Centre has a very consistent laboratory data recording system.

Data were checked for its completeness before analysis; in this regard a very small number of

Records which were incomplete. Also, the laboratory technologists/technicians engaged in

Malaria microscopy was competent and well trained in malaria microscopy.
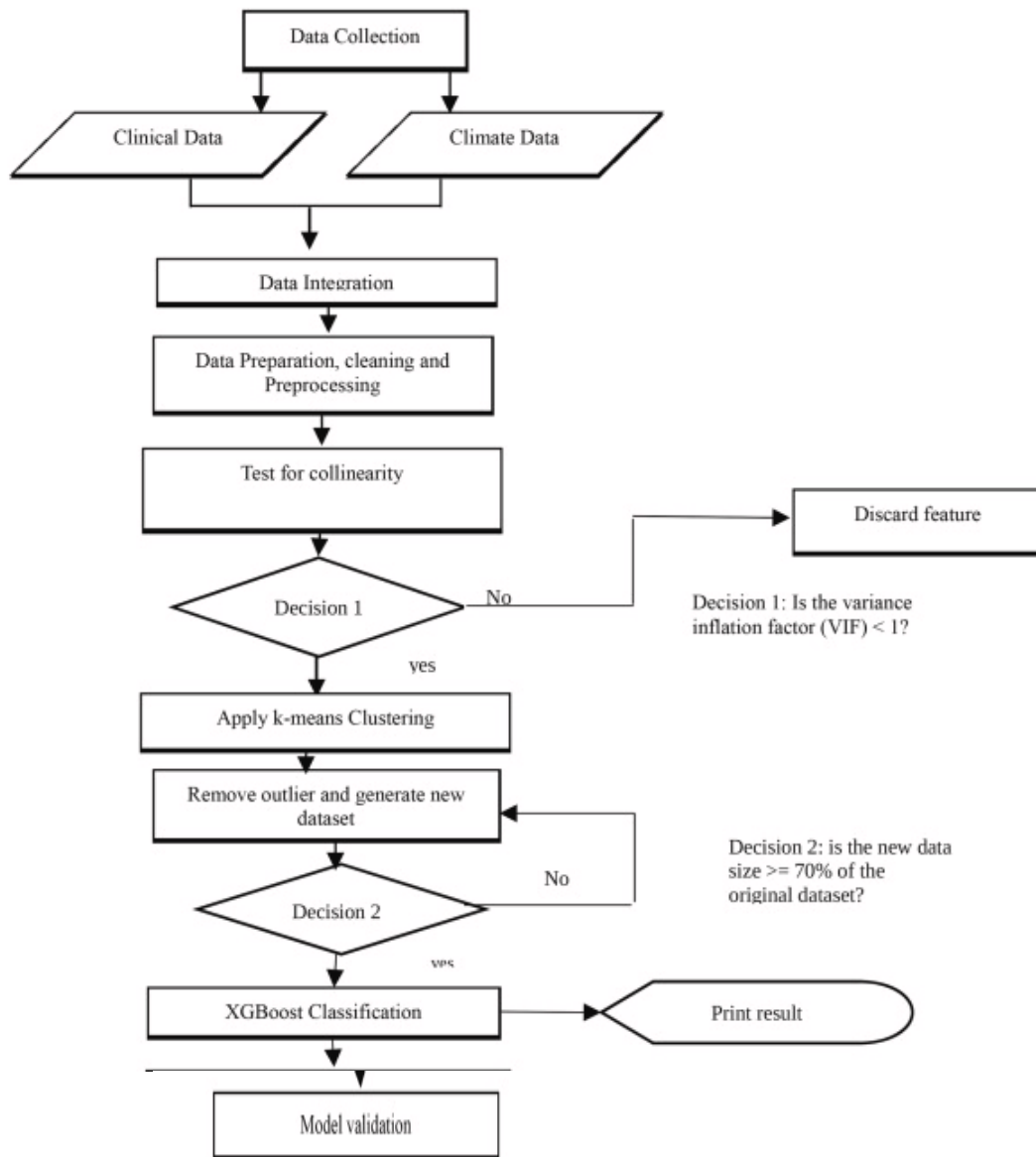
Results

From 2020 to 2022, blood films were prepared and examined from 6780 malaria-suspected

Patients at the three woreda's Centre. Of these, 2066 (4.52%) were microscopically confirmed

Malaria cases (annual mean = 158). The minimum (n = 43) and maximum (n = 122) number of

Annual malaria cases were reported in 2018 and 2020, respectively the highest (n = 1768) and

Lowest (n = 558) number of malaria-suspected patients were examined in 2019 and 2020,

Respectively. Throughout the reviewed period, malaria was more common in the urban

Population than the rural population. In both urban and rural people, higher malaria cases were

Observed in 2010 .Plasmodium falciparum was the most frequently reported species and

Reference

1. hagaw Anagew (2002), "Application of Data Mining Technology to Predict

Child Mortality Pattern: The Case of Butajira Rural Health Project", MSC Thesis,

Addis Ababa University, Ethiopia.

2. World Health Organization. World malaria report 2018. Geneva: World Health Organization;

2018. https://www.who.int/malaria/publications/world-malaria-report-2018/en/ Accessed 30 July

2019.

3. World Health Organization. Malaria: Factsheet. Geneva: World Health Organization; 2019.

https://www.who.int/news-room/fact-sheets/detail/malaria Accessed 30 July 2019.