# **PROJECT PROPOSAL ON**

**EARLY CLASSIFICATION AND IDENTIFICATION OF HEADACHES USING MACHINE LEARNING**

**BY:**

**AMODU DANIEL OJONUGWA**

**(SCI17CSC043)**

**BEING SUBMITTED TO DEPARTMENT OF COMPUTER SCIENCE, FACULTY OF SCIENCE. FEDERAL UNIVERSITY LOKOKA**

**INTRODUCTION**

In recent times, headaches have caused the death of many individuals in the world as less care and attention is given to individuals with headaches, as it is considered a minor sickness. Due to the nature of this sickness, identifying the type of headache an individual is suffering from becomes difficult and wrong medication may be prescribed.

Headache is a painful sensation in any part of the head, ranging from sharp to dull, that may occur with other symptoms. Headaches can have causes that are not due to underlying diseases. Examples, lack of sleep, an incorrect eyeglass prescription, stress loud noise exposure, or tight headwear. Headaches are divided into several types such as; Primary headaches, cluster headaches, secondary headaches, allergy or sinus headaches, hormone headaches, caffeine, post-traumatic, etc. thus the identification of a particular headache is very difficult as most of them operate almost in the same way.

Headaches can have causes that are not due to underlying disease. Examples include lack of sleep, an incorrect eyeglass prescription, stress, loud noise exposure, or tight headwear.

A working population is a carrier and the backbone of every society. This population is reproductively an active part of any society as well. Having these aspects in mind, it can be said that they are the most important group in society. Nevertheless, there is still a discrepancy between society’s expectations and investments when it comes to the working population. The common practice is to pay attention to the health of this population when it comes to risky occupations and perform regular medical checkups. They care about employees who suffer from a primary headache that begins only when they go to see a doctor. Headache disorders are the most prevalent of all neurological conditions and they are among the most frequent medical complaints seen in general practice. Half of the general population experiences a headache during any given year, and more than 90% report a lifetime history of head pain. (Krawczyk *et al*. 2013)

Despite its high prevalence, headache diagnosis is still mainly based on clinical interviews, patient diaries, and physical examinations. More advanced diagnostic methods are therefore desired for both clinical and research purposes and can potentially aid in early diagnosis and assessment of disease progression. Given the lack of consistent structural abnormalities, clinical neurophysiology methods are particularly suited to study the pathophysiology of headaches.

**1.1 STATEMENT OF THE PROBLEM**

Given the sensitive nature and high rate of headaches amongst individuals, early diagnosis can be done primarily by observing the lifestyle choices and social habits of the individuals. Since most people do not like being monitored nor do they give an accurate account of some of their lifestyle and health condition to someone else, the development of this system aims to address that issue, and provide a method of early classification and identification of headaches.

**1.2 AIM AND OBJECTIVES**

This project aims to design a model for the classification of headaches and identification of various types of headaches using patient symptoms.

The objectives are as follows:

a) To study and understand the dataset for this project work.

b) To adopt the use of machine learning algorithms K-Nearest Neighbor (KNN) and Random Forest (RF) to classify the various types of headaches.

c) To build a model using the above-mentioned algorithms and evaluate the results.

d) To implement an application for the classification and identification of different types of headaches.

**1.3 SIGNIFICANCE OF THE STUDY**

Through the development of this system model, one can determine the type of headache a patient is suffering from. This will also reduce the death rate or any complication among patients having this headache, as the right medication will be prescribed by the health practitioners to the affected patients effectively and on time.

**1.4 RESEARCH METHODOLOGY**

To effectively carry out this project, the data required need to be sourced from local and international medical facilities. Information such as the different types of headaches and their symptoms. Will be gathered in electronic format, and prepared in columns and rows on a Microsoft Excel spreadsheet, to allow for easy training and manipulation of data.

Furthermore, a programming language (python) capable of carrying out machine learning algorithms and training data sets will be used to analyze and classify the data to achieve the set goal and aim of this project. Here, the Random Forest (RF) algorithm will be used for the classification and identification of the types of headaches.

Finally, a MySQL database is used for the storage of the preprocessed and processed data. The SQL database will be used to normalize the data during analysis. Hence various columns and rows will be added and dropped during the processing stages to get the most accurate results and the most functional model with the least amount of faults.

**1.5 EXPECTED RESULT**

At the end of this project, we are expecting to build a model that comfortably identifies the different types of headaches early enough before the situation becomes critical