**CHAPTER THREE**

**Research Methodology**

**Materials and Methods**

This research work incorporated the use of a wide range of tools. The algorithm and implemented majorly with MATLAB programming language. An excel dataset in .xls format containing one thousand (1000) different cases with eight (8) features and one output was synthesized. The features consist of a cross section of important subjects offered in senior secondary schools for each case the output contains a human counselor's course recommendation based on the students performances in the eight subjects, in the implementation this recommendations are represented by an arbitrary number system,

'ENGINEERING'= 1;

'MEDCINE/SURGERY'= 2 ;

'PHARMACY'= 3;

'LAW'= 4;

'ACCOUNTING'= 5;

'ARCHITECTURE'= 6;

'QUANTITY SURVEYING'= 7;

'MASS COMMUNICATION'= 8;

'TOO POOR'= 9;

'GOOD AT ALL'= 10; without any loss of generality.

The data set was then normalized using the min-max normalization technique represented by the formula; , where

This normalization produced values between 0 and 1.

Using this dataset as a case-base, a GUI- based matlab application is developed to receive the user’s performance in the various courses into a row vector, which is also normalized following the min-max normalization technique.

The Euclidean distance between the new vector and each in the case-base is then computed Given an m-by-n data matrix X, which is treated as m(1-by-n) row vectors x, x, ..., x , the various distances between the vector xs and xt are defined as follows:

Euclidean distance:

These distances are sorted in ascending order and the 10 nearest vectors in the case-base to the new vector are selected the output recommendation for each of the neighbors in the case is considered the vote of that neighbor. The votes of each of the 10 nearest neighbors is determined and the majority vote is outputted as the recommendation for the new case .