Accuracy Measurements and Decision Making by Naïve Bayes and Forward Chaining Method to Identify the Malnutrition Causes and Symptoms

Muhammad Ibtasam
*Department of Computer Science, Faculty of Sciences, University of Agriculture Faisalabad, Pakistan
Email: Ibyasam76@gmail.com

Abstract.
Malnutrition is characterized as muscle weakening and cognitive disparity caused by social, dietary, political, food security issues. It appears as many underlying symptoms like fatigue, weakness, micronutrient deficiencies, weight loss to apparent symptoms of muscle mass reduction. Every 1 in 5 children is malnourished in developing countries.

Purpose: Policies and program formulation require prevalence facts to classify the most prevalent cause. Diagnostic tools and computer modeling have revolutionized the world of health sciences. Much algorithmic formulation can help to predict the prognosis of diseases based on the previous fact sheets.

Methods/Study design/approach: Naïve Bayes provides the posterior probability value that gives an analysis of the member with the whole sample set. Forward chaining gives the logistic conclusion with IF and THEN approach

Result/Findings: Naïve Bayes provided high accuracy of 88% as compared to 85% forward chaining.

Novelty/Originality/Value: In this study, the Naïve Bayes algorithm approach is coupled with the forward chaining system to provide a highly accurate measurement of the cause of malnutrition.

Keywords: Malnutrition, Naïve Bayes, Forward Chaining Method

INTRODUCTION
Malnutrition is a well-addressed issue among international organizations World Health Organization (WHO), World Food Program (WFP), and many other United Nations (UN) initiatives [1]. It is more prevalent in developing countries of South Asia due to social, political, and food insecurity [2]. It is mainly characterized as a clinical nutritional deficiency that emerged due to much socio-economic depravity, dietary causes, overweight, obesity, food insecurity, poor national policies, underlined health disease conditions, and unawareness of maternal and child nutrition [3]. It leads to mass muscle destruction, loss of cognitive functions, and a high risk of disease in severe cases. Diseases person is also at risk of malnutrition due to low appetite and poor absorption [4].

Stunting, low birth weight, undernutrition, weakness, low body mass, sarcopenia, loss of appetite, underlie disease causes and risks are the main symptoms of malnutrition. It is most prevalent among children and the maternal population [5].

Children under five years of age are at risk. The most prevalent symptoms in children are stunting, muscle wasting, hair follicle loss, cognitive impairment followed by fatigue and weakness. Maternal malnutrition gives birth to malnourished infants or low birth weight (LBW) infants with poor developmental growth in
the fetus [6]. The proper management of the issue requires policies, governance structure at the national level, and international surveillance of the effectiveness of the implemented policies [6][7].

Computer modeling and statistical modeling are used to quantify the extent of the problem. Naïve Bayes algorithm has the ability to identify malnutrition in children under 5 years with accurate probability [8][9]. Forward chaining method logistically concludes the probability of a case with a naïve bayes system. Naïve Bayes and Forward Chaining were used by [10] to determine childbirth process. Research done by [11][12] show that Naïve Bayes has good performance.

Research gap:
Global and regional malnutrition surveys and estimations by UNICEF, WHO, and World Bank show that the world is still far away from getting rid of malnutrition [13]. Malnutrition is a worldwide problem affecting both developed and underdeveloped countries. It is required to identify the main causative agents responsible for the prognosis of malnutrition at such a drastic level. It could be helpful in marking the most prevalent cause to be targeted in the policy-making.

METHODS
Decision-making methods were used to logically classify the main symptom of malnutrition on basis of the causes of malnutrition i.e., food security or dietary choices. The naïve Bayes method classifies the whole data set after being labeled, into parameters for the application of the model. The Naïve Bayes model works on the Bayesian algorithm which classified the data by probability and statistical modeling. It tests the probability of the specific attribute of the set dividing it into a subset and providing the result of probability by using the Bayes equation i.e.

\[
P(Y|X) = \frac{P(X|H)P(H)}{P(X)}
\]  

\(X = \) unknown class data  
\(H = \) Hypothesis of data  
\(P(H|X) = \) Posterior Probability of hypothesis (H) based on attribute X  
\(P(H) = \) Hypothesis probability  
\(P(X|H) = \) Probability of hypothesis based on attribute X  
\(P(X) = \) Probability of X

RESULT AND DISCUSSION
Naïve Bayes
This paper included 3 step calculation of the probability of causes of malnutrition on basis of food security and dietary choices and symptoms of malnutrition. The steps are Prior probability, Likelihood probability, and Posterior probability.

Prior probability
It is obtained by comparing the member of a class with the collected data of the whole sample. The following equation is used

\[
P = X / A
\]  

\(P = \) prior value  
\(X = \) total data present per class  
\(A = \) total data

Collected data from parameters of Sample 1 patients for Naïve Bays calculations:
1. Patients weight: 45 kg  
2. Other disease histories: none  
3. Hair loss: No  
4. Stunting: No  
5. Micronutrient deficiency symptoms: No  
6. Fatigue: No  
7. Weakness:  
8. Muscle wasting: No
Application of the Naïve Bays above-mentioned equation:

\[
P(R_{01}) = \frac{150}{269} = 0.55 \\
P(R_{02}) = \frac{119}{269} = 0.44
\]

**Likelihood probability**

Calculation of each attribute or parameter probability value to its class. The following equation is followed

\[
L = \frac{F}{B}
\]  

\( L = \text{Value of Likelihood} \)

\( F = \text{Number of attributes or parameters classified per class} \)

\( B = \text{Entire data per class} \)

Putting values in the above-mentioned Equation 3

\[
P(G_{01} = NO|R_{01}) = \frac{110}{212} = 0.51 \\
P(G_{02} = NO|R_{02}) = \frac{102}{212} = 0.48
\]

Likelihood values of sample 1 is shown in Table 1.

<table>
<thead>
<tr>
<th>Sr No.</th>
<th>Code</th>
<th>Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R01</td>
<td>0.51</td>
</tr>
<tr>
<td>2</td>
<td>R02</td>
<td>0.48</td>
</tr>
</tbody>
</table>

**Posterior probability**

Result of the obtained likelihood probability in the form of featured probability to the entire class. It is used to identify characteristics in the sample that should be prescribed as the primary symptom. The following equation is used to compute the final probability value

\[
P = \{H|E\} = P\{H\} \times P\{E|H\}
\]  

\( P = \text{Posterior probability} \)

\( H = \text{Extracted feature} \)

\( E = \text{Data per class} \)

Based on sample 1, the posterior probability of Naïve Bayes is as follows computed by using equation 4

Posterior P1: \(0.51 \times 0.47 \times 0.41 \times 0.36 \times 0.45 \times 0.55 \times 0.49 \times 0.42 = 0.0016\)

The posterior probability value obtained from P2 is high so in sample 1.

**Forward chaining method:**

Transfer of knowledge has been described in two categories i.e., Malnutrition causes and malnutrition symptoms. Codes for malnutrition causes and symptoms are shown in Table 2 and Table 3.

**Table 2. Malnutrition Causes**

<table>
<thead>
<tr>
<th>Sr No</th>
<th>Malnutrition Causes</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Dietary Choices</td>
<td>R01</td>
</tr>
<tr>
<td>2</td>
<td>Food Security</td>
<td>R02</td>
</tr>
</tbody>
</table>

**Table 3. Malnutrition Symptoms**

<table>
<thead>
<tr>
<th>Sr No</th>
<th>Symptoms</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Weight</td>
<td>G01</td>
</tr>
<tr>
<td>2</td>
<td>Disease history</td>
<td>G02</td>
</tr>
<tr>
<td>3</td>
<td>Hair loss</td>
<td>G03</td>
</tr>
<tr>
<td>4</td>
<td>Stunting</td>
<td>G04</td>
</tr>
<tr>
<td>5</td>
<td>Micronutrient deficits</td>
<td>G05</td>
</tr>
<tr>
<td>6</td>
<td>Fatigue</td>
<td>G06</td>
</tr>
<tr>
<td>7</td>
<td>Weakness</td>
<td>G07</td>
</tr>
<tr>
<td>8</td>
<td>Muscle wasting</td>
<td>G08</td>
</tr>
</tbody>
</table>
Rule base forward chaining
Production rules are used to investigate the forward chaining logistic method. The decision tables store the data about the malnutrition causes and symptoms that provide a rule base and arrangement of the relationship of each attribute.

<table>
<thead>
<tr>
<th>Rule</th>
<th>If</th>
<th>THEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>G01=NO^G02=NO^ R01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G03=NO^G04=NO^</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G05=NO^G06=NO^</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G07=NO^G08=NO^</td>
<td></td>
</tr>
</tbody>
</table>

Table 4 concludes that If selected subjects’ symptoms are not in accordance with R01 i.e., Dietary choices and the problem lies in R02 i.e., food security. The decision making can take two courses either detected the main problem or inference in the conclusion or not. In case the final result finds more than 1 conclusion the research is not concluded to its final and continues until it gets the final remarks. On applying the forward chaining method to sample 1 following conclusion are derived.

<table>
<thead>
<tr>
<th>Sr No.</th>
<th>Symptoms</th>
<th>Causes</th>
<th>Direction searches</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>G01=No</td>
<td>R01, R02</td>
<td>G02</td>
<td>Continue</td>
</tr>
<tr>
<td>2</td>
<td>G02=No</td>
<td>R01, R02</td>
<td>G03</td>
<td>Continue</td>
</tr>
<tr>
<td>3</td>
<td>G03=No</td>
<td>R01, R02</td>
<td>G04</td>
<td>Continue</td>
</tr>
<tr>
<td>4</td>
<td>G04=No</td>
<td>R01, R02</td>
<td>G05</td>
<td>Continue</td>
</tr>
<tr>
<td>5</td>
<td>G05=No</td>
<td>R01, R02</td>
<td>G06</td>
<td>Continue</td>
</tr>
<tr>
<td>6</td>
<td>G06=No</td>
<td>R01, R02</td>
<td>G07</td>
<td>Continue</td>
</tr>
<tr>
<td>7</td>
<td>G07=No</td>
<td>R01, R02</td>
<td>G08</td>
<td>Continue</td>
</tr>
<tr>
<td>8</td>
<td>G08=No</td>
<td>R02</td>
<td></td>
<td>Finalized</td>
</tr>
</tbody>
</table>

Table 5 states that the main cause of malnutrition lies in food security and muscle wasting as the main symptom leading to immediate malnutrition.

Testing of the Data
Corresponding and un-corresponding data are presented in Table 6.

<table>
<thead>
<tr>
<th>Corresponding data</th>
<th>Naïve Bayes</th>
<th>269</th>
<th>Forward chaining</th>
<th>212</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Corresponding data</td>
<td>32</td>
<td>30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

![Figure 1. System Diagnosis](image-url)
Accuracy values can be obtained by Equation 5

\[ \text{Accuracy value} = \frac{\text{amount of accurate data}}{\text{total data}} \times 100 \]  

In Naïve Bayes algorithm can be calculated as follow:
Total data = 269
Corresponding data = 237
Un-corresponding data = 32
Accuracy value = 237/269 * 100 = 88%

In Forward chaining can be calculated as follow:
Total Data = 269
Corresponding data = 239
Un-corresponding data = 30
Accuracy value = 239/269 * 100 = 85%

CONCLUSION
Naïve Bayes and the forward chaining method are efficient in logical concluding results. Naïve Bayes is more efficient and accurate in providing the results from the collected data. The naïve Bayes system provides the efficient identification of the cause of malnutrition i.e., muscle wasting and factor causing the malnutrition i.e., food insecurity. The efficiency of the forward chaining method is 85% and that of the naïve Bayes system is 88%.

REFERENCES