Classification Using Naïve Bayes and Decision Tree on Food Addiction

Adriyendi

IAIN Batusangkar & UPI YPTK Padang, Indonesia
elektronikpos@gmail.com

Abstract

In food consumption, food addiction is behavioral and biological overlaps have been observed between eating and addictive disorders. Food addiction influence to healthy life. Food addiction caused by overeating, bingo eating, eating disorder, eating addiction, mindless eating, craving, chocaholic, and emotional eating. Determination between addiction and normal condition in food consumption, need to classification. Classification is very important in determine signs of food addiction. Classification using Naïve Bayes Algorithm and Decision Tree Algorithm. Class target is Class Normal and Class Addiction. Classification using Naïve Bayes Algorithm by criterion is Calorie Dense Food, Fatty Food, Sweet Food, Diet and Stress. This criterion as causal factor of food addiction. Classification using Decision Tree Algorithm by criterion is Stress, Fatty Food and Calorie Dense Food. This criterion as causal factor of food addiction. The experimental result is a Classification Model. This model became data source for national policy in public health.

Keywords: *Classification, Food Addiction, Naïve Bayes Algorithm, Decision Tree Algorithm*

1. Introduction

People can be addicted to food consumption caused by overeating. Based on analysis in the paper [1], an addictive perspective suggests an interactive effect between an individual's predisposition for developing an addiction and an addiction agent. From analysis on paper [2], food addiction based on factor neurochemistry like alteration dopamine and endogenous opioids. Also, alteration on limbic system and self-medication behaviors. Food may be addictive is sweets, carbohydrates, fats, sweet/fat combinations and/or high salt foods. Eating topography has been identified in response to some food caused by eating addiction [3]. Response to some food like persistent desire to eat certain foods or cravings. Also, need to eat more of the food to get same felling. Negative effect of behaviors when eating like emotional eating and mindless eating. Assumption that addictive were 'bad', people felt that 'addiction' should be reduced in food consumption [4]. Based on research in the paper [5] shown that food addiction is factor contributing to over-eating then to obesity.

Determination between addiction and normal condition in food consumption, need to classification. Classification became very important in determine signs of food addiction. Food addiction influence healthy life. Result of classification on food addiction as data source for national policy in public health. Classification is a form of data analysis that extracts models describing important data classes. Effective and scalable methods heave developed in for Decision Tree Induction, Naïve Bayes Classification, rule-based classification and many other classification methods [6].

Specification of Naïve Bayes is a statistical classifier (performs probabilistic prediction), foundation based on Bayes Theorem, performance (comparable performance with decision tree and selected neural network classifier), incremental (each training

example can incrementally increase/decrease the probability that a hypothesis is correct-prior knowledge can be combined with observed data), and standard (even when Bayesian methods are computationally interact-able, they can provide a standard of optimal decision making against which other methods can be measured). Advantages is easy to implement and good results obtained in most of the cases.

Specification of Decision Tree is speed (relatively faster learning speed/than other classification methods), easy (convertible to simple and easy to understand classification rules), access (can use SQL queries for accessing databases), and accuracy (comparable classification accuracy with other methods). Advantages is produce a set of rules that are transparent, easy to understand, minimal error and considering most important attributes first.

2. Related Work

In the paper [7], twenty well known data mining classification methods are applied on ten UCI machine learning medical datasets, and the performance of various classification methods are empirically compared while varying the number categorical and numeric attributes. The types of attributes and the number of instances in datasets. Out of this twenty classification methods Bayes Net., Naïve Bayes, Classification via Regression, Logistic Regression and Random Forest Classification methods are the best. For numeric attribute datasets Classification via Regression, NB-Tree and Multiclass Classifier methods are the best. For categorical attribute datasets NB-Tree, Classification via Regression and Bayes Net methods are the best. Out of these above five rule based classification methods PART and Decision Tree methods are the best.

In the paper [8], presents a case study on predicting performance of students at the end of a university degree at an early stage of the degree program. Naïve Bayes has given an accuracy of 83.65% on Dataset II. The results show that it is possible to predict the graduation performance in 4th year at university using only pre-university marks and marks of 1st and 2nd year courses, no socio-economic or demographic features, with a reasonable accuracy.

In the paper [9], this research aims to compare C4.5, Naive Bayes, and Neural Network algorithm to determine which one is the most accurate in predicting heart disease. The test result of the three algorithms shows that the C4.5 algorithm has an accuracy value of 96.15%, while the Naive Bayes and Neural Network algorithm each has an accuracy value respectively of 86.15% and 94.23%. Thus, C4.5 algorithm can predict heart disease better than Naive Bayes and Neural Network algorithm.

In the paper [10], C5.0 algorithm is used to create a decision tree based on the experts, while fuzzy logic used to categorize the type of hypertensive disease that suffered by the patient and increase the level of accuracy of the diagnosing system. The accuracy of the combination between c5.0 algorithm and fuzzy logic is about 97.19%. In the paper [11], treats an evaluation of the analyzed results of classification algorithms selected for better prediction based on the characteristics of data from the data set with liver disease, by Naïve Bayes, Decision Tree, Multi-Layer Perceptron and k-NN used in a previous study, which developed our data set, and additionally Random forest, Logistic which proposed by us. Those algorithms were compared in several kinds of evaluating criteria like precision, recall, sensitivity, specificity, and so on. Through the experiments, we could know that in view of precision, Naïve Bayes is preferable than others, but in other criteria such as Recall and Sensitivity, Logistic and Random Forest took precedence over other algorithms in the performance of prediction test as considering the algorithmic characteristics to liver patient data set.

3. Framework

3.1. Classification Model

Classification is the processing of finding a set of model that distinguish data class for the purposes of being able to use the model to predict the class of objects whose class label is unknown. Classification is the process to construct a model based on the training set and uses it to classify new data or test set. It is a supervised learning as observations; measurements are accompanied by known class labels in a large amount of training set and new data is classified based on training set. Classification is one of main activities in the distillation of knowledge by learning supervised approach [12]. Classification on model can be depicted in a framework, shown in Figure 1.
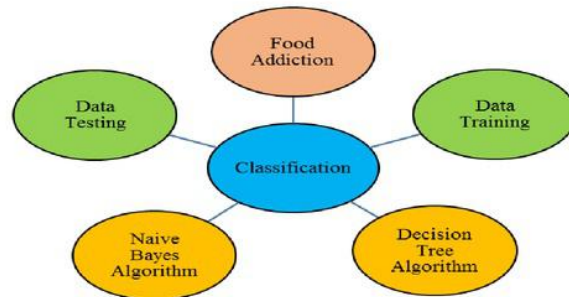

Figure 1. Classification Model

Figure 1 shown classification model using Naïve Bayes Algorithm and Decision Tree Algorithm. Data Training used to data analysis in classification. Data Testing used to extract data classes. Food Addiction as source and sample data processing in classification. Food Addiction Model shown on Figure 2.

Figure 2. Food Addiction Model

Figure 2 shown that food addiction caused by eating disorder, eating addiction, over eating, bingo eating, emotional eating, and mindless eating. Craving so make people food addiction. Factor chocaholic can be seen as addictive to some food in certain quantity.

3.2. Naïve Bayes Model

Naïve Bayes Classifier is classification with the method probability and statistic, namely predict opportunities in future based on experience so time formerly known as Bayes Theorem [13]. The theorem combined with Naïve where assumed conditions attribute between on and another is free. Probabilistic approach as Naïve Bayes Classifier, having several advantages among others, simple, quick and high accuracy, display on Figure 3.
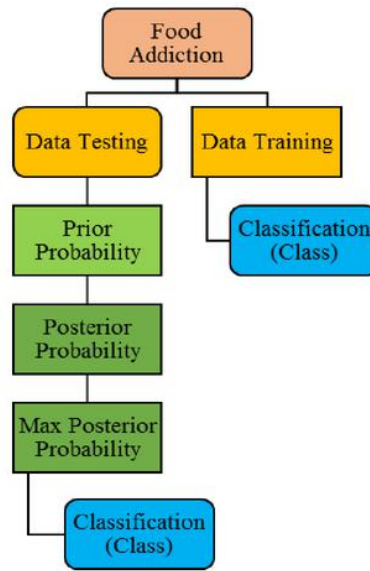

Figure 3. Naïve Bayes Model

Figure 3 shown Naïve Bayes Model by several step. Food Addiction as data source. Data Testing used to prior probability, posterior probability, max value of posterior probability, and Class as target in classification. Data Training as data source based on Food Addiction. Classification based Bayes Theorem: Probability (B given A) = Probability (A and B) / Probability (A). To calculate the probability of B given A, the algorithm counts the number of cases, A and B occur together, and divides it by the number of case where A occurs alone. Let X be a data tuple. In Bayesian terms, X is considered 'evidence'. Let H be some hypothesis, that the tuple X belongs class C. P (H | X) is the posterior probability of H conditioned on X. In contrast, P (H) is the prior probability of H. Bayes Theorem formula:

$$P(H \mid X) = P(X \mid H)P(H) / P(X) \tag{1}$$

Similarly, P (X | H) is the posterior probability of X conditioned on H. P (X) is the prior probability of X.

$$Posterior = (likelihood * prior) / evidence \tag{2}$$

It use a very intuitive technique. Bayes classifier, do not several free parameters that must be set. This greatly simplifies the design process. Since the classifier returns probabilities, it is simpler to apply these results to a wide variety of tasks than if an arbitrary scale was used. It does not require large amounts of data before learning can begin. Naïve Bayes Classifier are computationally fast when making decisions [14].

3.3. Decision Tree Model

A decision tree is a predictive modeling technique from the fields of machine learning and statistics that builds a simple tree-like structure to classify the data according their categories. A decision tree is a flowchart like tree structure, where each internal node denotes a test on an attribute [15]. Each branch represents an outcome of the test. Least nodes represent classes or class distributions, shown Figure 4.
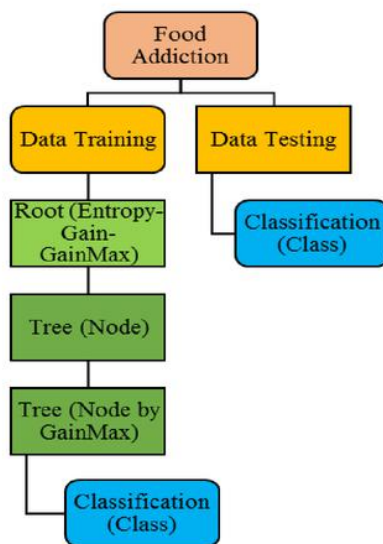

Figure 4. Decision Tree Model

Figure 4 shown Decision Tree Model step by step. Food Addiction as data source. Data Training used to determine entropy an information gain. Then, selected max value gain. Choice node by gain maximum. Target is Class. Decision Tree based on Decision Tree Induction. Step 1, tree is constructed in a top-down recursive divide-and-conquer manner. Step 2, at start, all the training examples are at the root. Step 3, attributes are categorical (if continuous-value, they are discretized in advance). Step 4, examples are partitioned recursively based on selected attributes. Step 5, test attributes are selected on the basis of a heuristic or statistical measure. Conditions for stopping partitioning when all samples or a given node belong to the same class. There are no remaining attributes for further partitioning-majority voting is employed for classifying the leaf, and there are no samples left. Entropy is a measure of uncertainty associated with a random variable. Interpretation, if higher entropy then higher uncertainty and if lower entropy then lower uncertainty. Entropy using formula:

$$Entropy \; (S) = \sum_{i=1}^{n} - \, pi * \log_2 \, pi \qquad (3)$$

Where S is dataset (data testing), n is number of partition and pi is proportion Si on S.

$$Gain \; (S, A) = Entropy \; (S) - \sum_{i=1}^{n} \frac{|S_i|}{|S|} * Entropy \; (S_i) \qquad (4)$$

Where S is dataset (data testing), n is number of partition attribute (A), Si is number of dataset on partition. Attribute selection measure using Information Gain (ID3/C4.5). Select the attribute by the highest information gain [16].

Classification using Naïve Bayes Algorithm and Decision Tree Algorithm. Step of Naïve Bayes Algorithm and Decision Tree Algorithm can be seen in Table 1.

Table 1. Step Algorithm

Step	Naïve Bayes Algorithm	Decision Tree Algorithm
1	Read data training	Prepare data training
2	Calculate number of class	Select attribute as root
3	Calculate the same case equal the same class	Make branch for each value
4	Multiply all value of result by data X to be classified	Repeat process 4 for each branch until all the case at the branch the same class

4. Result and Discussion

4.1. Classification Model

Classification Model using Naïve Bayes and Decision Tree can be seen on Figure 5.
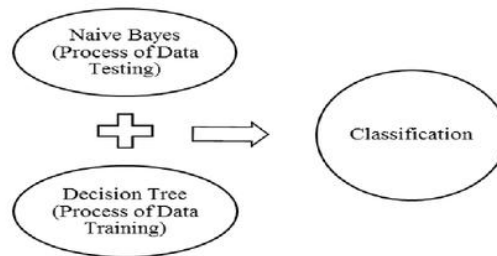
Figure 5. Process of Classification

Figure 5 shown class as target is class normal and class addiction. Calculation using Naïve Bayes Algorithm is process of data testing. Calculation using Decision Tree Algorithm is process of data training. Next, result it will be compared to determine class.

4.2. Naïve Bayes Model

Data training to classification model using Naïve Bayes Algorithm, shown on Table 2.

Table 2. Data Training for Naïve Bayes Algorithm

Id	Calorie Dense Food	Fatty Food	Sweet Food	Diet	Stress	Class
1	Low	High	High	No	No	Normal
2	Middle	Low	High	No	Yes	Addiction
3	High	Middle	Low	Yes	No	Normal
4	Low	High	Middle	Yes	Yes	Addiction
5	High	Low	Middle	Yes	No	Addiction
6	Low	Middle	High	No	No	Normal
7	Middle	High	Low	No	Yes	Addiction
8	High	Middle	High	Yes	No	Addiction
9	Middle	Low	Middle	No	Yes	Normal
10	Low	High	Low	Yes	Yes	Addiction
11	High	Low	High	Yes	Yes	Addiction
12	Middle	Middle	High	No	Yes	Normal

Table 2 shown data training, criteria and class. Next, data testing to classification, shown on Table 3.

Table 3. Data Testing for Naïve Bayes Algorithm

Id	Calorie Dense Food	Fatty Food	Sweet Food	Diet	Stress	Class
X	High	High	Middle	No	Yes	?

Table 3 shown criteria is Calorie Dense Food (High), Fatty Food (High), Sweet Food (Middle), Diet (No), and Stress (Yes). Class as target is: X…?. Naïve Bayes Classification, shown in Table 4.

Table 4. Naïve Bayes Classification

Data Testing	
Class	C1: Class = 'Normal' and C2 : Class = 'Addiction'
Prior Probability	
Data to be classified	X = (Calorie Dense Food = High, Fatty Food = High, Sweet Food = Middle, Diet = No, Stress = Yes) X → Class =? P(C$_i$) : P(Class = "Normal") = 5/12 = 0.41667 and P(Class = "Addiction") = 7/12 = 0.58333
Compute P(X|C$_i$) for each class	P(Calorie Dense Food = "High" | Class = "Normal") = 1/5 = 0.20000 P(Calorie Dense Food = "High" | Class = "Addiction") = 3/7 = 0.42857 P(Fatty Food = "High" | Class = "Normal") = 1/5 = 0.20000 P(Fatty Food = "High" | Class = "Addiction") = 3/7 = 0.42857 P(Sweet Food = "Middle" | Class = "Normal") = 1/5 = 0.20000 P(Sweet Food = "Middle" | Class = "Addiction") = 2/7 = 0. 28571 P(Diet = "Yes" | Class = "Normal") = 4/5 = 0.80000 P(Diet = "Yes" | Class = "Addiction") = 2/7 = 0. 28571 P(Stress = "Yes" | Class = "Normal") = 2/5 = 0.40000 P(Stress = "Yes" | Class = "Addiction") = 5/7 = 0.71429
Max Posterior Probability	
Compare Posterior Probability	X = (Calorie Dense Food = High, Fatty Food = High, Sweet Food = Middle, Diet = No, Stress = Yes) P(X|C$_i$) : P(X| Class = "Normal") = 0.20000 x 0.20000 x 0.20000 x 0.80000 x 0.40000 = 0.00256. P(X| Class = "Addiction") = 0.42857 x 0.42857 x 0. 28571 x 0. 28571 x 0.71429 = 0.01071. P(X|C$_i$)*P(C$_i$) : P(X| Class = "Normal") * P(Class = "Normal") = 0.41667 x 0.00256 = 0.00107P(X| Class = "Addiction") * P(Class = "Addiction") = 0.58333 x 0.01071 = 0.00625
Classification	
Class	X belongs to class → 0. 00625 > 0. 00107 (Class = "Addiction")

Table 4 shown result is class target = addiction based on data testing.

4.3. Decision Tree Model

Data testing classification model using Decision Tree Algorithm, shown on Table 5.

Table 5. Data Testing for Decision Tree Algorithm

Id	Calorie Dense Food	Fatty Food	Sweet Food	Diet	Stress	Class
1	Low	High	High	No	No	Normal
2	Middle	Low	High	No	Yes	Addiction
3	High	Middle	Low	Yes	No	Normal
4	Low	High	Middle	Yes	Yes	Addiction
5	High	Low	Middle	Yes	No	Addiction
6	Low	Middle	High	No	No	Normal
7	Middle	High	Low	No	Yes	Addiction
8	High	Middle	High	Yes	No	Addiction
9	Middle	Low	Middle	No	Yes	Normal

10	Low	High	Low	Yes	Yes	Addiction
11	High	Low	High	Yes	Yes	Addiction
12	Middle	Middle	High	No	Yes	Normal

Table 5 shown data (Id1, Id2, until Id12). Criteria is Calorie Dense Food (Low, Middle, High), Fatty Food (Low, Middle, High), Sweet Food (Low, Middle, High), Diet (Yes, No), and Stress (Yes, No). Class as target is: Normal and Addiction.

Table 6. Entropy Root

Entropy	Calculation
$E_{(Total)}$ $= E(T)$	$= E_{(Total)} = ([5+, 7-]) = - (5/12) \log_2 (5/12) - (7/12) \log_2 (7/12) = - (0.41667) ((\log (5/12)) / \log 2) - (0.58333) ((\log (7/12)) / \log 2) = - (0.41667) (-0.38021) / (0.30103) - (0.58333) (-0.23408) / (0.30103) = - (0.41667) (-1.26303) - (0.58333) (-0.77759) = (0.52627) + (0.45359) = 0.97986$
$E_{(Calorie}$ $_{Dense\ Food)}$ $=E(CDF,H)$	$= E_{(High)} = ([1+, 3-]) = - (1/4) \log_2 (1/4) - (3/4) \log_2 (3/4) = - (0.25) ((\log (1/4)) / \log 2) - (0.75) ((\log (3/4)) / \log 2) = - (0.25) (- 0.60206 / (0.30103) - (0.75) (- 0.12494) / (0.30103) = - (0.25) (-2) - (0.75) (-0.41504) = (0.5) + (0.31128) = 0.81128$
$E(CDF,M)$	$= E_{(Middle)} = ([2+, 2-]) = - (2/4) \log_2 (2/4) - (2/4) \log_2 (2/4) = - (0.5) ((\log (2/4)) / \log 2) - (0.5) ((\log (2/4)) / \log 2) = - (0.5) (-0.30103 / (0.30103) - (0.5) (-0.30103) / (0.30103) = - (0.5) (-1) - (0.5) (-1) = (0.5) + (0.5) = 1$
$E(CDF,L)$	$= E_{(Low)} = ([2+, 2-]) = - (2/4) \log_2 (2/4) - (2/4) \log_2 (2/4) = - (0.5) ((\log (2/4)) / \log 2) - (0.5) ((\log (2/4)) / \log 2) = - (0.5) (-0.30103 / (0.30103) - (0.5) (-0.30103) / (0.30103) = - (0.5) (-1) - (0.5) (-1) = (0.5) + (0.5) = 1$
$E_{(Fatty\ Food)}$ $= E(FF,H)$	$= E_{(High)} = ([1+, 3-]) = - (1/4) \log_2 (1/4) - (3/4) \log_2 (3/4) = - (0.25) ((\log (1/4)) / \log 2) - (0.75) ((\log (3/4)) / \log 2) = - (0.25) (- 0.60206 / (0.30103) - (0.75) (- 0.12494) / (0.30103) = - (0.25) (-2) - (0.75) (-0.41504) = (0.5) + (0.31128) = 0.81128$
$E(FF,M)$	$= E_{(Middle)} = ([3+, 1-]) = - (3/4) \log_2 (3/4) - (1/4) \log_2 (1/4) = - (0.75) ((\log (3/4)) / \log 2) - (0.25) ((\log (1/4)) / \log 2) = - (0.75) (-0.12494/ (0.30103) - (0.25) (- 0.60206) / (0.30103) = - (0.75) (-0.41504) - (0.25) (-2) = (0.31128) + (0.5) = 0.81128$
$E(FF,L)$	$= E_{(Low)} = ([1+, 3-]) = - (1/4) \log_2 (1/4) - (3/4) \log_2 (3/4) = - (0.25) ((\log (1/4)) / \log 2) - (0.75) ((\log (3/4)) / \log 2) = - (0.25) (- 0.60206 / (0.30103) - (0.75) (- 0.12494) / (0.30103) = - (0.25) (-2) - (0.75) (-0.41504) = (0.5) + (0.31128) = 0.81128$
$E_{(Sweet\ Food)}$ $= E(SF,H)$	$= E_{(High)} = ([3+, 3-]) = - (3/6) \log_2 (3/6 - (3/6) \log_2 (3/6) = - (0.5) ((\log (3/6)) / \log 2) - (0.5) ((\log (3/6)) / \log 2) = - (0.5) (- 0.30103 / (0.30103) - (0.5) (-0.30103) / (0.30103) = - (0.5) (-1) - (0.5) (-1) = (0.5) + (0.5) = 1$
$E(SF,M)$	$= E_{(Middle)} = ([1+, 2-]) = - (1/3) \log_2 (1/3) - (2/3) \log_2 (2/3) = - (0.33333) ((\log (1/3)) / \log 2) - (0.66667) ((\log (2/3)) / \log 2) = - (0.33333) (-0.47712/ (0.30103) - (0.66667) (-0.17609) / (0.30103) = - (0.33333) (-1.58496) - (0.66667) (-0.58496) = (0.52831) + (0.38998) = 0.91829$
$E(SF,L)$	$= E_{(Low)} = ([1+, 2-]) = - (1/3) \log_2 (1/3) - (2/3) \log_2 (2/3) = - (0.33333) ((\log (1/3)) / \log 2) - (0.66667) ((\log (2/3)) / \log 2) = - (0.33333) (-0.47712/ (0.30103) - (0.66667) (-0.17609) / (0.30103) = - (0.33333) (-1.58496) - (0.66667) (-0.58496) = (0.52831) + (0.38998) = 0.91829$
$E_{(Diet)}$ $= E(D,Y)$	$= E_{(Yes)} = ([1+, 5-]) = - (1/6) \log_2 (1/6) - (5/6) \log_2 (5/6) = - (0.16667) ((\log (1/6)) / \log 2) - (0.83333) ((\log (5/6)) / \log 2) = - (0.16667) (-0.77815) / (0.30103) - (0.83333) (-0.07918) / (0.30103) = - (0.16667) (-2.58496) - (0.83333) (- 0.26303) = (0.43084) + (0.21919) = 0.65003$
$E(D,N)$	$= E_{(No)} = ([4+, 2-]) = - (4/6) \log_2 (4/6) - (2/6) \log_2 (2/6) = - (0.66667) ((\log (4/6)) / \log 2) - (0.33333) ((\log (2/6)) / \log 2) = - (0.66667) (-0.17609) / (0.30103) - (0.33333) (-0.47712) / (0.30103) = - (0.66667) (-0.58496) - (0.33333) (-1.58496) = (0.38998) + (0.52831) = 0.91829$
$E_{(Stress)}$ $= E(S,Y)$	$= E_{(Yes)} = ([2+, 5-]) = - (2/7) \log_2 (2/7) - (5/7) \log_2 (5/7) = - (0.28571) ((\log (2/7)) / \log 2) - (0.71429) ((\log (5/7)) / \log 2) = - (0.28571) (-0.28571) / (0.30103) - (0.71429) (-0.14613) / (0.30103) = - (0.28571) (-0.94911) - (0.71429) (- 0.48543) = (0.27117) + (0.34674) = 0.61791$

E(S,N)	= $E_{(No)}$ = ([3+, 2-]) = - (3/5) $\log_2$ (3/5) - (2/5) $\log_2$ (2/5) = - (0.6) ((log (3/5)) / log 2) – (0.4) ((log (2/5)) / log 2) = - (0.6) (-0.22185) / (0.30103) – (0.4) (-0.39794) / (0.30103) = - (0.6) (-0.73697) – (0.4) (-1.32193) = (0.44218) + (0.52877) = 0.97095

    Table 6 shown process calculate Entropy Total (value of comparison class Normal and Addiction) for each criteria.

Table 7. Gain Root

Gain	Calculation
Gain(S, CDF) $S_{High\,[1+,\,3-]}$ $S_{Middle\,[2+,\,2-]}$ $S_{Low\,[2+,\,2-]}$	= E(S)–(4/12)E(S_{High})–(4/12)E(S_{Middle}) – (4/12) E (S_{Low}) = 0.97986 – (4/12) 0.81128 – (4/12) 1 (4/12) 1 = 0.97986 – (0.33333) 0.81128 (0.33333) 1 – (0.33333)1 = 0.97986 – 0.27042 - 0.33333 - 0.33333 = 0.04278
Gain(S, FF) $S_{High\,[1+,\,3-]}$ $S_{Middle\,[3+,\,1-]}$ $S_{Low\,[1+,\,3-]}$	= E(S)–(4/12)E(S_{High})–(4/12)E(S_{Middle}) (4/12)E(S_{Low}) = 0.97986 – (4/12) 0. 81128 – (4/12) 0. 81128 –(4/12) 0. 81128 = 0.97986 – (0.33333) 0.81128 – (0.33333) 0. 81128 – (0.33333) 0. 81128 = 0.97986 – 0.27042 - 0.27042 -0.27042 = 0.16860
Gain(S, SF) $S_{High\,[3+,\,3-]}$ $S_{Middle\,[1+,\,2-]}$ $S_{Low\,[1+,\,2-]}$	= E(S)–(6/12)E(S_{High})–(3/12)E(S_{Middle}) –(3/12) E (S_{Low}) = 0.97986 – (6/12) 1 – (3/12) 0.91829 – (3/12) 0.91829 = 0.97986 – (0.5) 1 – (0.25) 0.91829 – (0.25) 0.91829 = 0.97986 – 0.5 - 0.22957 - 0.22957 = 0.02072
Gain(S, D) $S_{Yes\,[1+,\,5-]}$ $S_{No\,[4+,\,2-]}$	= E(S)–(6/12)E(S_{Yes})–(6/12)E(S_{No}) = 0.97986 –(6/12) 0.65003 – (6/12) 0.91829 = 0.97986 – (0.5) 0. 65003 – (0.5) 0.91829 = 0.97986 – 0.32502 – 0.45915 = 0.19569
Gain(S, S) $S_{Yes\,[2+,\,5-]}$ $S_{No\,[3+,\,2-]}$	= E(S)–(7/12)E(S_{Yes})–(5/12)E(S_{No}) = 0.97986 – (7/12) 0.61791 – (5/12) 0.97095 = 0.97986 – (0.58333) 0.61791 – (0.41666) 0.97095 = 0.97986 – 0.36045 – 0.40456 = 0.21485

    Table 7 shown information gain and calculation, where higher gain based on entropy for each class.

Table 8. Attribute Entropy and Gain Root

Attribute		Total (S)	Normal (S_1)	Addiction (S_2)	Entropy	Gain
Total		12	5	7	0.97986	
CDF						0.04278
	H	4	1	3	0.81128	
	M	4	2	2	1	
	L	4	2	2	1	
FF						0.16860
	H	4	1	3	0.81128	
	M	4	3	1	0.81128	
	L	4	1	3	0.81128	
SF						0.02072
	H	6	3	3	1	
	M	3	1	2	0.91829	
	L	3	1	2	0.91829	
D						0.19569
	Y	6	1	5	0.65003	
	N	6	4	2	0.91829	
S						0.21485
	Y	7	2	5	0.61791	
	N	5	3	2	0.97095	

Table 8 shown value of attributes, entropy and gain. Gain selected is higher gain. Result is criteria Stress = 0.21485. Criteria Stress became next node in decision tree.

Table 9. Entropy Branch for Stress = Yes

Entropy	Calculation
$E_{(Total)} = E(T)$	$= E_{(Total)} = ([2+, 5-]) = - (2/7) \log_2 (2/7) - (5/7) \log_2 (5/7) = - (0.28571) ((\log (2/7)) / \log 2) - (0.71429) ((\log (5/7)) / \log 2) = - (0.28571) (-0.54407) / (0.30103) - (0.71429) (-0.14613) / (0.30103) = - (0.28571) (-1.80736) - (0.71429) (-0.48543) = (0.51638) + (0.34674) = 0.86312$
$E_{(CalorieDenseFood)} = E(CDF,H)$	$= E_{(High)} = ([0+, 1-]) = - (0/1) \log_2 (0/1) - (1/1) \log_2 (1/1) = - (0) ((\log (0/1)) / \log 2) - (1) ((\log (1/1)) / \log 2) = - (0) (0 / (0.30103) - (1) (0) / (0.30103) = - (0) (0) - (1) (0) = (0) + (0) = 0$
$E(CDF,M)$	$= E_{(Middle)} = ([2+, 2-]) = - (2/4) \log_2 (2/4) - (2/4) \log_2 (2/4) = - (0.5) ((\log (2/4)) / \log 2) - (0.5) ((\log (2/4)) / \log 2) = - (0.5) (-0.30103 / (0.30103) - (0.5) (-0.30103) / (0.30103) = -(0.5) (-1) - (0.5) (-1) = (0.5) + (0.5) = 1$
$E(CDF,L)$	$= E_{(Low)} = ([0+, 2-]) = - (0/2) \log_2 (0/2) - (2/2) \log_2 (2/2) = - (0) ((\log (0/2)) / \log 2) - (1) ((\log (2/2)) / \log 2) = - (0) (0 / (0.30103) - (1) (0) / (0.30103) = - (0) (0) - (1) (0) = (0) + (0) = 0$
$E_{(FattyFood)} = E(FF,H)$	$= E_{(High)} = ([0+, 3-]) = - (0/3) \log_2 (0/3) - (3/3) \log_2 (3/3) = - (0) ((\log (0/3)) / \log 2) - (1) ((\log (3/3)) / \log 2) = - (0) (0) / (0.30103) - (1) (0) / (0.30103) = - (0) (0) - (1) (0) = (0) + (0) = 0$
$E(FF,M)$	$= E_{(Middle)} = ([1+, 0-]) = - (1/1) \log_2 (1/1) - (0/1) \log_2 (0/1) = - (1) ((\log (1/1)) / \log 2) - (0) ((\log (0/1)) / \log 2) = - (1) (0) / (0.30103) (0) (0) / (0.30103) = - (1) (0) - (0) (0) = (0) + (0) = 0$
$E(FF,L)$	$= E_{(Low)} = ([1+, 2-]) = - (1/3) \log_2 (1/3) - (2/3) \log_2 (2/3) = - (0.33333) ((\log (1/3)) / \log 2) - (0.66667) ((\log (2/3)) / \log 2) = - (0.33333) (- 0.47712) / (0.30103) - (0.66667) (- 0.17609) / (0.30103) = - (0.33333) (-1.58496) - (0.66667) (- 0.58496) = (0.52831) + (0.38998) = 0.91829$
$E_{(SweetFood)} = E(SF,H)$	$= E_{(High)} = ([1+, 2-]) = - (1/3) \log_2 (1/3) - (2/3) \log_2 (2/3) = - (0.33333) ((\log (1/3)) / \log 2) - (0.66667) ((\log (2/3)) / \log 2) = - (0.33333) (- 0.47712) / (0.30103) - (0.66667) (- 0.17609) / (0.30103) = - (0.33333) (-1.58496) - (0.66667) (- 0.58496) = (0.52831) + (0.38998) = 0.91829$
$E(SF,M)$	$= E_{(Middle)} = ([1+, 1-]) = - (1/2) \log_2 (1/2) - (1/2) \log_2 (1/2) = - (0.5) ((\log (1/2)) / \log 2) - (0.5) ((\log (1/2)) / \log 2) = - (0.5) (-0.30103 / (0.30103) - (0.5) (-0.30103) / (0.30103) = - (0.5) (-1) - (0.5) (-1) = (0.5) + (0.5) = 1$
$E(SF,L)$	$= E_{(Low)} = ([0+, 2-]) = - (0/2) \log_2 (0/2) - (2/2) \log_2 (2/2) = - (0) ((\log (0/2)) / \log 2) - (1) ((\log (2/2)) / \log 2) = - (0) (0 / (0.30103) - (1) (0) / (0.30103) = - (0) (0) - (1) (0) = (0) + (0) = 0$
$E_{(Diet)} = E(D,Y)$	$= E_{(Yes)} = ([0+, 3-]) = - (0/3) \log_2 (0/3) - (3/3) \log_2 (3/3) = - (0) ((\log (0/3)) / \log 2) - (1) ((\log (3/3)) / \log 2) = - (0) (0) / (0.30103) - (1) (0) / (0.30103) = - (0) (0) - (1) (0) = (0) + (0) = 0$
$E(D, N)$	$= E_{(No)} = ([2+, 2-]) = - (2/4) \log_2 (2/4) - (2/4) \log_2 (2/4) = - (0.5) ((\log (2/4)) / \log 2) - (0.5) ((\log (2/4)) / \log 2) = - (0.5) (-0.30103 / (0.30103) - (0.5) (-0.30103) / (0.30103) = - (0.5) (-1) - (0.5) (-1) = (0.5) + (0.5) = 1$

Table 10. Gain for Stress = Yes

Gain	Calculation
Gain(S,CDF) $S_{High [0+, 1-]}$ $S_{Middle [2+, 2-]}$ $S_{Low [0+, 2-]}$	$= E(S) - (1/7)E(S_{High}) - (4/7)E(S_{Middle}) - (2/7)E(S_{Low}) = 0.86312 - (1/7) 0 - (4/7) 1 - (2/7) 0 = 0.86312 (0.14286) 0 - (0. 57143) 1 - (0.28571) 0 = 0.86312 - 0 - 0. 57143 - 0 = 0.29169$
Gain(S, FF) $S_{High [0+, 3-]}$ $S_{Middle [1+, 0-]}$ $S_{Low [1+, 2-]}$	$= E(S) - (3/7)E(S_{High}) - (1/7)E(S_{Middle}) - (3/7)E(S_{Low}) = 0.86312 - (3/7) 0 - (1/7) 0 - (3/7) 0.91829 = 0.86312 - (0.42857) 0 - (0.14286) 0 - (0.42857) 0.91829 = 0.86312 - 0 - 0 - 0.39355 = 0.46957$
Gain(S, SF) $S_{High [1+, 2-]}$ $S_{Middle [1+, 1-]}$ $S_{Low [0+, 2-]}$	$= E(S) - (3/7)E(S_{High}) - (2/7)E(S_{Middle}) - (2/7)E(S_{Low}) = 0.86312 - (3/7) 0.91829 - (2/7) 1 - (2/7) 0 = 0.86312 - (0.42857) 0.91829 - (0.28571) 1 - (0.28571) 0 = 0.86312 - 0.39355 - 0.28571 - 0 = 0.18386$

Gain(S, D) $S_{Yes\ [0+,\ 3-]}$ $S_{No\ [2+,\ 2-]}$	= E(S)–(3/7)E(S_{Yes})–(4/7)E(S_{No}) = 0.86312 – (3/7) 0 – (4/7) 1= 0.86312 – (0.42857) 0 – (0. 57143) 1= 0.86312 – 0 – 0. 57143 = 0.29169

Table 9 shown entropy branch. Table 10 shown higher gain is Fatty Food = 0.46957.

Table 11. Attribute Entropy and Gain for Stress = Yes

Attribute		Total (S)	Normal (S₁)	Addiction (S₂)	Entropy	Gain
Stress = Yes		7	2	5	0.86312	
CDF						0.29169
	H	1	0	1	0	
	M	4	2	2	1	
	L	2	0	2	0	
FF						0.46957
	H	3	0	3	0	
	M	1	1	0	0	
	L	3	1	2	0.91829	
SF						0.18386
	H	3	1	2	0.91829	
	M	2	1	1	1	
	L	2	0	2	0	
D						0.29169
	Y	3	0	3	0	
	N	4	2	2	1	

Table 11 shown value of attributes, entropy and gain. Gain selected is higher gain. Result is criteria Fatty Food (FF) = 0.46957. Fatty Food became next node in decision tree.

Table 12 (a). Entropy Branch for Stress = No 1

Entropy	Calculation
E($_{Total}$)=E(T)	= E($_{Total}$) = ([3+, 2-]) = - (3/5) $\log_2$ (3/5) – (2/5) $\log_2$ (2/5) = - (0.6) ((log (3/5)) / log 2) – (0.4) ((log (2/5)) / log 2) = - (0.6) (-0.22185) / (0.30103) – (0.4) (-0.39794) / (0.30103) = - (0.6) (-0.73697) – (0.4) (-1.32193) = (0.44218) + (0.52877) = 0.97095
E($_{Calorie Dense Food}$) =E(CDF,H)	= E($_{High}$) = ([1+, 2-]) = - (1/3) $\log_2$ (1/3) - (2/3) $\log_2$ (2/3) = - (0.33333) ((log (1/3)) / log 2) – (0.66667) ((log (2/3)) / log 2)= - (0.33333) (-0.47712) / (0.30103) – (0.66667) (-0.17609) / (0.30103) = - (0.33333) (-1.58496) – (0.66667) (- 0.58496) = (0.52831) + (0.38996) = 0.9182 4
E(CDF,M)	= E($_{Middle}$) = ([0+, 0-]) = - (0/0) $\log_2$ (0/0) - (0/0) $\log_2$ (0/0) =- (0) ((log (0/0)) / log 2) – (0) ((log (0/0)) / log 2) = - (0) (0) / (0.30103) – (0) (0) / (0.30103) =- (0) (0) – (0) (0) = (0)+(0) = 0
E(CDF,L)	= E($_{Low}$) = ([2+, 0-]) = - (2/0) $\log_2$ (2/0) - (0/2) $\log_2$ (0/2) = - (0) ((log (2/0)) / log 2) – (0) ((log (0/2)) / log 2) = - (0) (0) / (0.30103)) – (0) (0) / (0.30103) = - (0) (0) – (0) (0) = (0) + (0) = 0
E($_{Fatty Food}$) =E(FF,H)	= E($_{High}$)= ([1+, 0-]) = - (1/1) $\log_2$ (1/1) - (0/1) $\log_2$ (0/1)= - (1) ((log (1/1)) / log 2) – (0) ((log (0/1)) / log 2) = - (1) (0) / (0.30103) – (0) (0) / (0.30103) = - (1) (0) – (0) (0) = (0) + (0) = 0
E(FF,M)	= E($_{Middle}$) = ([2+, 1-]) = - (2/3) $\log_2$ (2/3) - (1/3) $\log_2$ (1/3) = - (0.6667) ((log (2/3)) / log 2) – (0.33333) ((log (1/3)) / log 2) = - (0.6667) (-0.17609) / (0.30103) – (0. 33333) (-0.47712) / (0.30103) = - (0.6667) (-0.58496) – (0. 33333) (-1.58496) = (0.38998) + (0.52831) = 0.91829
E(FF,L)	= E($_{Low}$) = ([0+, 1-]) = - (0/1) $\log_2$ (0/1) - (1/1) $\log_2$ (1/1) = - (0) ((log (0/1)) / log 2) – (1) ((log (1/1)) / log 2) = - (0) (0) / (0.30103) – (1) (0) / (0.30103) = - (0) (0) – (1) - (0) = (0) + (0) = 0

Table 12 (a) and (b) shown value of Entropy Total, Calorie Dense Food (H, M, L) and Fatty Food (H, M, L) to criteria Stress. Entropy branch to criteria Stress = No.

Table 12 (b). Entropy Branch for Stress = No 1

Entropy	Calculation
$E_{(Sweet\ Food)}$ =E(SF,H)	= $E_{(High)}$ = ([2+, 1-]) = - (2/3) $\log_2$ (2/3) - (1/3) $\log_2$ (1/3) = - (0.6667) ((log (2/3)) / log 2) − (0.33333) ((log (1/3)) / log 2) = - (0.6667) (-0.17609) / (0.30103) − (0.33333) (-0.47712) / (0.30103) = - (0.6667) (-0.58496) − (0. 33333) (-1.58496) = (0.38998) + (0.52831) = 0.91829
E(SF,M)	= $E_{(Middle)}$ = ([0+, 1-]) = - (0/1) $\log_2$ (0/1) - (1/1) $\log_2$ (1/1) = - (0) ((log (0/1)) / log 2) − (1) ((log (1/1)) / log 2) = - (0) (0) / (0.30103) − (1) (0) / (0.30103) = - (0) (0) −(1)-(0) =(0)+(0) = 0
E(SF,L)	= $E_{(Low)}$= ([1+, 0-]) = - (1/1) $\log_2$ (1/1) - (0/1) $\log_2$ (0/1) = - (1) ((log (1/1)) / log 2) − (0) ((log (0/1)) / log 2) = - (1) (0) / (0.30103) − (0) (0) / (0.30103) = - (1) (0) − (0) (0) = (0) + (0) = 0
$E_{(Diet)}$ =E(D,Y)	= $E_{(Yes)}$ = ([1+, 2-]) = - (1/3) $\log_2$ (1/3) - (2/3) $\log_2$ (2/3) = - (0.33333) ((log (1/3)) / log 2) − (0.66667) ((log (2/3)) / log 2) = - (0.33333) (-0.47712) / (0.30103) − (0.66667) (-0.17609) / (0.30103) = - (0.33333) (-1.58496) − (0.66667) (- 0.58496) = (0.52831) + (0.38996) = 0.91827
E(D, N)	= $E_{(Low)}$ = ([2+, 0-]) = - (2/0) $\log_2$ (2/0) - (0/2) $\log_2$ (0/2) = - (0) ((log (2/0)) / log 2) − (0) ((log (0/2)) / log 2) = - (0) (0) / (0.30103)) − (0) (0) / (0.30103) = - (0) (0) − (0) (0) = (0) + (0) = 0

Table 13. Entropy Branch for Stress = No 2

Entropy	Calculation
$E_{(Fatty\ Food)}$ =E(FF,H)	= $E_{(High)}$= ([1+, 0-]) = - (1/1) $\log_2$ (1/1) - (0/1) $\log_2$ (0/1) = - (1) ((log (1/1)) / log 2) − (0) ((log (0/1)) / log 2) = - (1) (0) / (0.30103) − (0) (0) / (0.30103) = - (1) (0) − (0) (0) = (0) + (0) = 0
E(FF,M)	= $E_{(Middle)}$ = ([2+, 1-]) = - (2/3) $\log_2$ (2/3) - (1/3) $\log_2$ (1/3) = - (0.6667) ((log (2/3)) / log 2) − (0.33333) ((log (1/3)) / log 2) = - (0.6667) (-0.17609) / (0.30103) − (0.33333) (-0.47712) / (0.30103) = - (0.6667) (-0.58496) − (0. 33333) (-1.58496) = (0.38998) + (0.52831) = 0.91829
E(FF,L)	= $E_{(Low)}$ = ([0+, 1-]) = - (0/1) $\log_2$ (0/1) - (1/1) $\log_2$ (1/1) = - (0) ((log (0/1)) / log 2) − (1) ((log (1/1)) / log 2) = - (0) (0) / (0.30103) − (1) (0) / (0.30103) = - (0) (0) − (1) - (0) = (0) + (0) = 0
$E_{(Sweet\ Food)}$ =E(SF,H)	= $E_{(High)}$ = ([2+, 1-]) = - (2/3) $\log_2$ (2/3) - (1/3) $\log_2$ (1/3) = - (0.6667) ((log (2/3)) / log 2) − (0.33333) ((log (1/3)) / log 2) = - (0.6667) (-0.17609) / (0.30103) − (0.33333) (-0.47712) / (0.30103) = - (0.6667) (-0.58496) − (0. 33333) (-1.58496) = (0.38998) + (0.52831) = 0.91829
E(SF,M)	= $E_{(Middle)}$ = ([0+, 1-]) = - (0/1) $\log_2$ (0/1) - (1/1) $\log_2$ (1/1) = - (0) ((log (0/1)) / log 2) − (1) ((log (1/1)) / log 2) = - (0) (0) / (0.30103) − (1) (0) / (0.30103) = - (0) (0) − (1) - (0) = (0) + (0) = 0
E(SF,L)	= $E_{(Low)}$ = ([1+, 0-]) = - (1/1) $\log_2$ (1/1) - (0/1) $\log_2$ (0/1) = - (1) ((log (1/1)) / log 2) − (0) ((log (0/1)) / log 2) = - (1) (0) / (0.30103) − (0) (0) / (0.30103) = - (1) (0) − (0) (0) = (0) + (0) = 0
$E_{(Diet)}$ =E(D,Y)	= $E_{(Yes)}$ = ([1+, 2-]) = - (1/3) $\log_2$ (1/3) - (2/3) $\log_2$ (2/3) = - (0.33333) ((log (1/3)) / log 2) − (0.66667) ((log (2/3)) / log 2) = - (0.33333) (-0.47712) / (0.30103) − (0.66667) (-0.17609) / (0.30103) = - (0.33333) (-1.58496) − (0.66667) (- 0.58496) = (0.52831) + (0.38996) = 0.91827
E(D, N)	= $E_{(Low)}$ = ([2+, 0-]) = - (2/0) $\log_2$ (2/0) - (0/2) $\log_2$ (0/2) = - (0) ((log (2/0)) / log 2) − (0) ((log (0/2)) / log 2) = - (0) (0) / (0.30103)) − (0) (0) / (0.30103) = - (0) (0) − (0) (0) = (0) + (0) = 0

Table 13 shown value of entropy branch. Entropy Fatty Food (High, Middle, Low) to criteria Stress. Entropy Sweet Food (High, Middle, Low) to criteria Stress. Entropy Diet (Yes, No) to criteria Stress = No.

Table 14. Gain for Stress = No

Gain	Calculation
Gain(S, CDF) $S_{High\,[1+,\,2-]}$ $S_{Middle\,[0+,\,0-]}$ $S_{Low\,[2+,\,0-]}$	$= E(S)-(3/5)E(S_{High})-(0/5)E(S_{Middle})-(2/5)E(S_{Low})$ $= 0.97095 - (3/5)\,0.91827 - (0/5)\,0 - (2/5)\,0$ $= 0.97095 - (0.6)\,0.91827 - (0)\,0 - (0.4)\,0$ $= 0.97095 - 0.55096 - 0 - 0 = 0.41999$
Gain(S, FF) $S_{High\,[1+,\,0-]}$ $S_{Middle\,[2+,\,1-]}$ $S_{Low\,[0+,\,1-]}$	$= E(S)-(1/5)E(S_{High})-(3/5)E(S_{Middle})-(1/5)E(S_{Low})$ $= 0.97095 - (1/5)\,0 - (3/5)\,0.91829 - (1/5)\,0$ $= 0.97095 - (0.2)\,0 - (0.6)\,0.91829 - (0.2)\,0$ $= 0.97095 - 0 - 0.55097 - 0 = 0.41998$
Gain(S, SF) $S_{High\,[2+,\,1-]}$ $S_{Middle\,[0+,\,1-]}$ $S_{Low\,[1+,\,0-]}$	$= E(S)-(3/5)E(S_{High})-(1/5)E(S_{Middle})-(1/5)E(S_{Low})$ $= 0.97095 - (3/5)\,0.91829 - (1/5)\,0 - (1/5)\,0$ $= 0.97095 - (0.6)\,0.91829 - (0.2)\,0 - (0.2)\,0$ $= 0.97095 - 0.55097 - 0 - 0 = 0.41998$

Table 14 shown information gain for each class in branch criteria Stress = No. Result it all of value gain is equal. Conditions for stopping partitioning when all samples for given node belong to the same class.

Table 15. Attribute Entropy and Gain for Stress = No

Attribute	Total (S)	Normal (S_1)	Addiction (S_2)	Entropy	Gain
Stress = No	5	3	2	0.97095	
CDF					0.41999
H	3	1	2	0.91827	
M	0	0	0	0	
L	2	2	0	0	
FF					0.41998
H	1	1	0	0	
M	3	2	1	0.91829	
L	1	0	1	0	
SF					0.41998
H	3		H	3	
M	1		M	1	
L	1		L	1	
D					0.41999
Y	3		Y	3	
N	2		N	2	

Table 15 shown gain for each class is equal. Conditions for stopping partitioning when all samples for given node belong to the same class.

Table 16. Data Testing for Fatty Food = High

Id	Calorie Dense Food	Fatty Food	Sweet Food	Diet	Stress	Class
1	Low	High	High	No	No	Normal
4	Low	High	Middle	Yes	Yes	Addiction
7	Middle	High	Low	No	Yes	Addiction
10	Low	High	Low	Yes	Yes	Addiction

Table 16 shown data testing (12 point data from data source) filtering became 4 point data based on criteria Fatty Food = High to each class.

Table 17. Entropy for Fatty Food = High

Entropy	Calculation
$E_{(Total)} = E(T)$	$= E_{(Total)} = ([1+, 3-]) = -(1/4) \log_2 (1/4) - (3/4) \log_2 (3/4) = -(0.25) ((\log (1/4)) / \log 2) - (0.75) ((\log (3/4)) / \log 2) = -(0.25) (-0.60206) / (0.30103) - (0.75) (-0.12494) / (0.30103) = -(0.25) (-2) - (0.75) (-0.41504) = (0.5) + (0.31128) = 0.81128$
$E_{(CalorieDense\ Food)} = E(CDF,H)$	$= E_{(High)} = ([0+, 0-]) = -(0/0) \log_2 (0/0) - (0/0) \log_2 (0/0) = -(0/0) ((\log (0/0)) / \log 2) - (0/0) ((\log (0/0)) / \log 2) = -(0/0)(0/0)/(0.30103)-(0/0)(0/0)/(0.30103) = -(0/0) (0.30103) - (0/0) (0.30103) = (0) + (0) = 0$
$E(CDF,M)$	$= E_{(Middle)} = ([0+, 1-]) = -(0/0) \log_2 (0/0) - (1/1) \log_2 (1/1) = -(0) ((\log (0/0)) / \log 2) - (1) ((\log (1/1)) / \log 2) = -(0) (0) / (0.30103) - (1) (0) / (0.30103) = -(0) (0) - (1) (0) = (0) + (0) = 0$
$E(CDF,L)$	$= E_{(Low)} = ([1+, 2-]) = -(1/3) \log_2 (1/3) - (2/3) \log_2 (2/3) = -(0.33333) ((\log (1/3)) / \log 2) - (0.66667) ((\log (2/3)) / \log 2) = -(0.33333) (-0.47712) / (0.30103)) - (0.66667) (-0.17609) / (0.30103) = -(0.33333) (-0.58496) - (0.66667) (-0.58496) = (0.19498) + (0.38998) = 0.58496$
$E_{(SweetFood)} = E(SF,H)$	$= E_{(High)} = ([1+, 0-]) = -(1/1) \log_2 (1/1) - (0/1) \log_2 (0/1) = -(1) ((\log (1/1)) / \log 2) - (0) ((\log (0/1)) / \log 2) = -(1) (0) / (0.30103) - (0) (0) / (0.30103) = -(1) (0) - (0) (0) = (0) + (0) = 0$
$E(SF,M)$	$= E_{(Middle)} = ([0+, 1-]) = -(0/1) \log_2 (0/1) - (1/1) \log_2 (1/1) = -(0) ((\log (0/1)) / \log 2) - (1) ((\log (1/1)) / \log 2) = -(0) (0) / (0.30103) - (1) (0) / (0.30103) = -(0) (0) - (1) - (0) = (0) + (0) = 0$
$E(SF,L)$	$= E_{(Low)} = ([0+, 2-]) = -(0/2) \log_2 (0/2) - (2/2) \log_2 (2/2) = -(0) ((\log (0/2)) / \log 2) - (1) ((\log (2/2)) / \log 2) = -(0) (0) / (0.30103) - (1) (0) / (0.30103) = -(0) (0) - (1) (0) = (0) + (0) = 0$
$E_{(Diet)} = E(D,Y)$	$= E_{(Yes)} = ([0+, 2-]) = -(0/2) \log_2 (0/2) - (2/2) \log_2 (2/2) = -(0) ((\log (0/2)) / \log 2) - (1) ((\log (2/2)) / \log 2) = -(0) (0) / (0.30103) - (1) (0) / (0.30103) = -(0) (0) - (1) (0) = (0) + (0) = 0$
$E(D, N)$	$= E_{(No)} = ([0+, 1-]) = -(0/1) \log_2 (0/1) - (1/1) \log_2 (1/1) = -(0) ((\log (0/1)) / \log 2) - (1) ((\log (1/1)) / \log 2) = -(0) (0) / (0.30103) - (1) (0) / (0.30103) = -(1) - (0) = (0) + (0) = 0$
$E_{(Stress)} = E(S,Y)$	$= E_{(Yes)} = ([0+,3-]) = -(0/3) \log_2 (0/3) - (3/3) \log_2 (3/3) = -(0) ((\log (0/3)) / \log 2) - (1) ((\log (3/3)) / \log 2) = -(0) (0) / (0.30103) - (1) (0) / (0.30103) = -(0) (0) - (1) (0) = (0) + (0) = 0$
$E(S,N)$	$= E_{(No)} = ([1+, 0-]) = -(1/1) \log_2 (1/1) - (0/1) \log_2 (0/1) = -(1) ((\log (1/1)) / \log 2) - (0) ((\log (0/1)) / \log 2) = -(1) (0) / (0.30103) - (0) (0) / (0.30103) = -(1) (0) - (0) (0) = (0) + (0) = 0$

Table 17 shown entropy for Fatty Food = High to each class based on previous node. Entropy Calorie Dense Food = Low by value 0.58496. Others entropy result is 0.

Table 18. Gain for Fatty Food = High

Gain	Calculation
Gain(S, CDF) $S_{High\ [0+,\ 0-]}$ $S_{Middle\ [0+,\ 1-]}$ $S_{Low\ [1+,\ 2-]}$	$= E(S) - (0/4)E(S_{High}) - (1/4)E(S_{Middle}) - (3/4)E(S_{Low}) = 0.81128 - (0/4) 0 - (1/4) 0 - (3/4) 0.58496 = 0.81128 - (0) 0 - (0.25) 0 - (0.75) 0.58496 = 0.81128 - 0 - 0 - 0.43872 = 0.37256$
Gain(S, SF) $S_{High\ [1+,\ 0-]}$ $S_{Middle\ [0+,\ 1-]}$ $S_{Low\ [0+,\ 2-]}$	$= E(S) - (1/4)E(S_{High}) - (1/4)E(S_{Middle}) - (2/4)E(S_{Low}) = 0.81128 - (1/4) 0 - (1/4) 0 - (2/4)0 = 0.81128 - (0.25) 0 - (0.25) 0 - (0.5) 0 = 0.81128 - 0 - 0 - 0 = 0.81128$
Gain(S, D) $S_{Yes\ [0+,\ 2-]}$ $S_{No\ [0+,\ 1-]}$	$= E(S) - (2/4)E(S_{Yes}) - (2/4)E(S_{No}) = 0.81128 - (2/4) 0 - (2/4) 0 = 0.81128 - (0.5) 0 - (0.5)0 = 0.81128 - 0 - 0 = 0.81128$
Gain(S, S) $S_{Yes\ [0+,\ 3-]}$ $S_{No\ [1+,\ 0-]}$	$= E(S) - (3/4)E(S_{Yes}) - (1/4)E(S_{No}) = 0.81128 - (3/4) 0 - (1/4) 0 = 0.81128 - (0.75) 0 - (0.25) 0 = 0.81128 - 0 - 0 = 0.81128$

Table 18 shown information gain for each class equal. Conditions for stopping partitioning when all samples for given node belong to the same class.

Table 19. Attribute Entropy and Gain for Fatty Food = High

Attribute		Total (S)	Normal (S₁)	Addiction (S₂)	Entropy	Gain
Fatty Food = High		4	1	3	0.81128	
CDF						0.37256
	H	0	0	0	0	
	M	1	0	1	0	
	L	3	1	2	0.58496	
SF						0. 81128
	H	1	1	0	0	
	M	1	0	1	0	
	L	2	0	2	0	
D						0. 81128
	Y	2	0	2	0	
	N	1	0	1	0	
S						0. 81128
	Y	3	0	3	0	
	N	1	1	0	0	

Table 19 shown Entropy Calorie Dense Food = Low by value .58496. Others entropy is equal by value 0. Gain to Entropy Calorie Dense Food = .37256 and others entropy is equal by value 0.81128. Entropy (SF = Sweet Food, D = Diet, S = Stress) have value is equal, then partitioning is stopped. Choice one of them (Sweet Food).

Table 20. Data Testing for Fatty Food = Middle

Id	Calorie Dense Food	Fatty Food	Sweet Food	Diet	Stress	Class
3	High	Middle	Low	Yes	No	Normal
6	Low	Middle	High	No	No	Normal
8	High	Middle	High	Yes	No	Addiction
12	Middle	Middle	High	No	Yes	Normal

Table 20 shown data testing (12 point data) filtering became 4 point data based on Fatty Food = Middle.

Table 21 (a). Entropy for Fatty Food = Middle

Entropy	Calculation
$E_{(Total)}$ =E(T)	$= E_{(Total)} = ([3+, 1-]) = -(3/4) \log_2 (3/4) - (1/4) \log_2 (1/4) = -(0.75) ((\log (3/4)) / \log 2) - (0.25) ((\log (1/4)) / \log 2) = -(0.75) (-0.12494) / (0.30103) - (0.25) (-0.60206) / (0.30103) = -(0.75) (-0.41504) - (0.25) (-2) = (0.31128) + (0.5)$ $= 0.81128$
$E_{(Calorie\ Dense\ Food)}$ =E(CDF,H)	$= E_{(High)} = ([1+, 1-]) = -(1/2) \log_2 (1/2) - (1/2) \log_2 (1/2) = -(0.5) ((\log (1/2)) / \log 2) - (0.5) ((\log (1/2)) / \log 2) = -0.5 (-0.30103) / (0.30103) - (0.5) (-0.30103) / (0.30103) = -(0.5) (-1) - (0.5) (-1) = (0.5) + (0.5)$ $= 1$
E(CDF,M)	$= E_{(Middle)} = ([1+, 0-]) = -(1/1) \log_2 (1/1) - (0/1) \log_2 (0/1) = -(1) ((\log (1/1)) / \log 2) - (0) ((\log (0/1)) / \log 2) = -(1) (0) / (0.30103) - (0) (0) / (0.30103)$ $= -(1) (0) - (0) (0) = (0) + (0)$ $= 0$
E(CDF,L)	$= E_{(Low)} = ([1+, 0-]) = -(1/1) \log_2 (1/1) - (0/1) \log_2 (0/1) = -(1) ((\log (1/1)) / \log 2) - (0) ((\log (0/1)) / \log 2) = -(1) (0) / (0.30103) - (0) (0) / (0.30103)$ $= -(1) (0) - (0) (0) = (0) + (0)$ $= 0$

Table 21 (b). Entropy for Fatty Food = Middle

Entropy	Calculation
$E_{(SweetFood)}$ =E(SF,H)	= $E_{(High)}$ = ([2+, 1-]) = - (2/3) $\log_2$ (2/3) - (1/3) $\log_2$ (1/3) = - (0.66667) ((log (2/3)) / log 2) – (0.) ((log (1/3)) / log 2) = - (0.66667) (-0.17609) / (0.30103) – (0.33333) (-0.47712) / (0.30103) = - (0.66667) (-0.58496) – (0.33333) (-1.58496) = (0.38998) + (0.52831) = 0.91829
E(SF,M)	= $E_{(Middle)}$ = ([0+, 0-]) = - (0/0) $\log_2$ (0/0) - (0/0) $\log_2$ (0/0) = - (0) ((log (0/0)) / log 2) – (0) ((log (0/0)) / log 2) = - (0) (0) / (0.30103) – (0) (0) / (0.30103) = - (0) (0) – (0) - (0) = (0) + (0) = 0
E(SF,L)	= $E_{(Low)}$= ([1+, 0-]) = - (1/1) $\log_2$ (1/1) - (0/1) $\log_2$ (0/1) = - (1) ((log (1/1)) / log 2) – (0) ((log (0/1)) / log 2) = - (1) (0) / (0.30103) – (0) (0) / (0.30103) = - (1) (0) – (0) (0) = (0) + (0) = 0
$E_{(Diet)}$ =E(D,Y)	= $E_{(Yes)}$ = ([1+, 1-]) = - (1/2) $\log_2$ (1/2) - (1/2) $\log_2$ (1/2) = - (0.5) ((log (1/2)) / log 2) – (0.5) ((log (1/2)) / log 2) = - (0.5 -0.30103) / (0.30103) – (0.5) (-0.30103) / (0.30103) = - (0.5) (-1) – (0.5) (-1) = (0.5) + (0.5) = 1
E(D, N)	= $E_{(No)}$ = ([2+, 0-]) = - (2/2) $\log_2$ (2/2) - (0/0) $\log_2$ (0/0) = - (1) ((log (2/2)) / log 2) – (0) ((log (0/)) / log 2)= - (1) (0) / (0.30103) – (0) (0) / (0.30103) = - (1) (0) – (0) - (0) = (0) + (0) = 0
$E_{(Stress)}$ =E(S,Y)	= $E_{(Yes)}$ = ([1+,0-]) = - (1/1) $\log_2$ (1/1) - (0/1) $\log_2$ (0/1) = - (1) ((log (1/1)) / log 2) – (0) ((log (0/1)) / log 2) = - (1) (0) / (0.30103) – (0) (0) / (0.30103) = - (1) (0) – (0) (0) = (0) + (0) = 0
E(S,N)	= $E_{(No)}$ = ([2+, 1-]) = - (2/3) $\log_2$ (2/3) - (1/3) $\log_2$ (1/3) = - (0.66667) ((log (2/3)) / log 2) – (0.) ((log (1/3)) / log 2) = - (0.66667) (-0.17609) / (0.30103) – (0.33333) (-0.47712) / (0.30103) = - (0.66667) (-0.58496) – (0.33333) (-1.58496) = (0.38998) + (0.52831) = 0.91829

Table 21 (a) and (b) shown Entropy Total, Entropy Calorie Dense Food, Entropy Sweet, Entropy Diet, and Entropy Stress to criteria Fatty Food = Middle.

Table 22. Gain for Fatty Food = Middle

Gain	Calculation
Gain(S, CDF) $S_{High\,[1+,\,1-]}$ $S_{Middle\,[1+,\,0-]}S_{Low\,[1+,\,0-]}$	= E(S)–(2/4)E(S_{High})–(1/4)E(S_{Middle})–(1/4)E(S_{Low}) = 0.81128 – (2/4) 1 – (1/4) 0 – (1/4) 0 = 0.81128 – (0.5) 1 – (0.25) 0 – (0.25) 0 = 0.81128 – 0.5 - 0 – 0 = 0.31128
Gain(S, SF) $S_{High\,[2+,\,1-]}$ $S_{Middle\,[0+,\,0-]}S_{Low\,[1+,\,0-]}$	= E(S)–(3/4)E(S_{High})–(0/4)E(S_{Middle})–(1/4)E(S_{Low}) = 0.81128 – (3/4) 0.91829 – (0/4) 0 – (1/4) 0 = 0.81128 – (0.75) 0.91829 – (0) 0 – (0.25) 0 = 0.81128 – 0.68872 - 0 – 0 = 0. 12256
Gain(S, D) $S_{Yes\,[1+,\,1-]}$ $S_{No\,[2+,\,0-]}$	= E(S)–(2/4)E(S_{Yes})–(2/4)E(S_{No}) = 0.81128 – (2/4) 1 – (2/4) 0 = 0.81128 – (0.5) 1 – (0. 5) 0 = 0.81128 – 0.5 – 0 = 0.31128
Gain(S, S) $S_{Yes\,[1+,\,0-]}$ $S_{No\,[2+,\,1-]}$	= E(S)–(1/4)E(S_{Yes})–3/4)E(S_{No}) = 0.81128 – (1/4) 0 – (3/4) 0.91829 = 0.81128 – (0.25) 0 – (0.75) 0.91829 = 0.81128 – 0 – 0.68872 = 0.12256

Table 22 shown information gain (S, CDF) and gain (S, D) is equal. Gain (S, SF) and gain (S, S) is equal.

Table 23. Attribute Entropy and Gain for Fatty Food = Middle

Attribute		Total (S)	Normal (S₁)	Addiction (S₂)	Entropy	Gain
Fatty Food = Middle		4	3	1	0. 81128	
CDF						0.31128
	H	2	1	1	1	
	M	1	1	0	0	
	L	1	1	0	0	
SF						0. 12256
	H	3	2	1	0.91829	
	M	0	0	0	0	
	L	1	1	0	0	
D						0.31128
	Y	2	1	1	1	
	N	2	2	0	0	
S						0.12256
	Y	1	1	0	0	
	N	3	2	1	0.91829	

Table 23 shown information gain (Sweet Food = SF) and gain (Stress = S) is equal. Gain (Calorie Dense Food = CDF) and (Diet = D) is equal. It is meaning that 2 couple entropy have same value, then partitioning is stopped. Compare and choice higher gain (CDF and D). Choice one of them (CDF).

Table 24. Data Testing for Fatty Food = Low

Id	Calorie Dense Food	Fatty Food	Sweet Food	Diet	Stress	Class
3	High	Low	Low	Yes	No	Normal
6	Low	Low	High	No	No	Normal
8	High	Low	High	Yes	No	Addiction
12	Middle	Low	High	No	Yes	Normal

Table 24 shown data testing (12 point data) filtering became 4 point data based on Fatty Food = Low.

Table 25 (a). Entropy for Fatty Food = Low

Entropy	Calculation
$E_{(Total)}$ =E(T)	= $E_{(Total)}$ = ([1+, 3-]) = - (1/4) log₂ (1/4) – (3/4) log₂ (3/4) = - (0.25) ((log (1/4)) / log 2) – (0.75) ((log (3/4)) / log 2) = - (0.25) (-0. 60206) / (0.30103) – (0.75) (-0. 12494) / (0.30103) = - (0.25) (-2) – (0.75) (-0. 41504) = (0.5) + (0. 31128) = 0. 81128
$E_{(Calorie\ Dense\ Food)}$ =E(CDF,H)	= $E_{(High)}$ = ([0+, 2-]) = - (0/2) log₂ (0/2) - (2/2) log₂ (2/2) = - (0) ((log (0/2)) / log 2) – (1) ((log (2/2)) / log 2) = - (0 (0) / (0.30103) – (1) (0) / (0.30103) = - (0) (0) – (1) (0) = (0) + (0) = 0
E(CDF,M)	= $E_{(Middle)}$ = ([1+, 1-]) = - (1/2) log₂ (1/2) - (1/2) log₂ (1/2) = - (0.5) ((log (1/2)) / log 2) – (0.5) ((log (1/2)) / log 2) = - (0.5 (-0.30103) / (0.30103) – (0.5) (-0.30103) / (0.30103) = - (0.5) (-1) – (0.5) (-1) = (0.5)+(0.5) = 1
E(CDF,L)	= $E_{(Low)}$ = ([0+, 0-]) = - (0/0) log₂ (0/0) - (0/0) log₂ (0/0) = - (0) ((log (0/0)) / log 2) – (0) ((log (0/0)) / log 2) = - (0) (0) / (0.30103) – (0) (0) / (0.30103) = - (0) (0) – (0) - (0) = (0) + (0) = 0

Table 25 (b). Entropy for Fatty Food = Low

Entropy	Calculation
$E_{(SweetFood)}$ =E(SF,H)	= $E_{(High)}$ = ([0+, 2-]) = - (0/2) log$_2$ (0/2) - (2/2) log$_2$ (2/2) = - (0) ((log (0/2)) / log 2) – (1) ((log (2/2)) / log 2) = - (0 (0) / (0.30103) – (1) (0) / (0.30103) = - (0) (0) – (1) (0) = (0) + (0) = 0
E(SF,M)	= $E_{(Middle)}$ = ([1+, 1-]) = - (1/2) log$_2$ (1/2) - (1/2) log$_2$ (1/2) = - (0.5) ((log (1/2)) / log 2) – (0.5) ((log (1/2)) / log 2) = - (0.5 (-0.30103) / (0.30103) – (0.5) (-0.30103) / (0.30103) = - (0.5) (-1) – (0.5) (-1) = (0.5) + (0.5) = 1
E(SF,L)	= $E_{(Low)}$= ([0+, 0-]) = - (0/0) log$_2$ (0/0) - (0/0) log$_2$ (0/0) = - (0) ((log (0/0)) / log 2) – (0) ((log (0/0)) / log 2) = - (0) (0) / (0.30103) – (0) (0) / (0.30103) = - (0) (0) – (0) = (0) + (0) = 0
$E_{(Diet)}$ = E(D,Y)	= $E_{(Yes)}$ = ([0+, 2-]) = - (0/2) log$_2$ (0/2) - (2/2) log$_2$ (2/2) = - (0) ((log (0/2)) / log 2) – (1) ((log (2/2)) / log 2) = - (0 (0) / (0.30103) – (1) (0) / (0.30103) = - (0) (0) – (1) (0) = (0) + (0) = 0
E(D, N)	= $E_{(No)}$ = ([1+, 1-]) = - (1/2) log$_2$ (1/2) - (1/2) log$_2$ (1/2) = - (0.5) ((log (1/2)) / log 2) – (0.5) ((log (1/2)) / log 2) = - (0.5 (-0.30103) / (0.30103) – (0.5) (-0.30103) / (0.30103) = - (0.5) (-1) – (0.5) (-1) = (0.5) + (0.5) = 1
$E_{(Stress)}$ = E(S, Y)	= $E_{(Yes)}$ = ([1+,2-]) = - (1/3) log$_2$ (1/3) - (2/3) log$_2$ (2/3) = - (0.33333) ((log (1/3)) / log 2) – (0.66667) ((log (2/3)) / log 2) = - (0.33333) (-0.47712) / (0.30103) – (0.66667) (-0.17609) / (0.30103) = - (0.33333) (-1.58496) – (0.66667) (-0.58496) = 52831) + (0.38998) = 0.91829
E(S,N)	= $E_{(No)}$ = ([0+, 1-]) = - (0/1) log$_2$ (0/1) - (1/1) log$_2$ (1/1) = - (0) ((log (0/1)) / log 2) – (1) ((log (1/1)) / log 2) = - (0) (0) / (0.30103) – (1) (0) / (0.30103) = - (0) (0) – (1) (0) = (0) + (0) = 0

Table 26. Gain for Fatty Food = Low

Gain	Calculation
Gain(S, CDF) $_{High [0+, 2-]}$ $S_{Middle [1+, 1-]}$ $S_{Low [0+, 0-]}$	= E(S)–(2/2)E(S$_{High}$)–(2/2)E(S$_{Middle}$)–(0/0)E(S$_{Low}$) = 0.81128 – (2/2) 0 – (2/2) 1 – (0/2) 0 = 0.81128 – (1) 0 – (1) 1 – (0) 0 = 0.81128 – 0 - 1– 0 = - 0.18872
Gain(S, SF) $S_{High [0+, 2-]}$ $S_{Middle [1+, 1-]}$ $S_{Low [0+, 0-]}$	= E(S)–(2/2)E(S$_{High}$)–(2/2)E(S$_{Middle}$)–(0/0)E(S$_{Low}$) = 0.81128 – (2/2) 0 – (2/2) 1 – (0/2) 0 = 0.81128 – (1) 0 – (1) 1 – (0) 0 = 0.81128 – 0 - 1– 0 = - 0.18872
Gain(S, D) $S_{Yes [0+, 2-]}$ $S_{No [1+, 1-]}$	= E(S)–(2/2)E(S$_{Yes}$)–(2/2)E(S$_{No}$) = 0.81128 – (0/2) 0 – (2/2) 1 = 0.81128 – (0) 0 – (1) 1 = 0.81128 – 0 – 1 = -0.18872
Gain(S, S) $S_{Yes [1+, 1-]}$ $S_{No [0+, 1-]}$	= E(S)–(2/2)E(S$_{Yes}$)–1/2)E(S$_{No}$) = 0.81128 – (2/2) 1 – (1/2) 0 = 0.81128 – (1) 1 – (0.5) 0 = 0.81128 – 1 – 0 = -0.18872

Table 27. Attribute Entropy and Gain for Fatty Food = Low

Attribute		Total (S)	Normal (S$_1$)	Addiction (S$_2$)	Entropy	Gain
Fatty Food = Low		4	1	3	0. 81128	
CDF						-0.18872
	H	2	0	2	0	
	M	2	1	1	1	
	L	0	0	0	0	
SF						-0.18872
	H	2	0	2	0	
	M	2	1	1	1	
	L	0	0	0	0	
D						-0.18872
	Y	2	0	2	0	
	N	2	1	1	1	
S						-0.18872
	Y	3	1	2	0.91829	
	N	1	0	1	0	
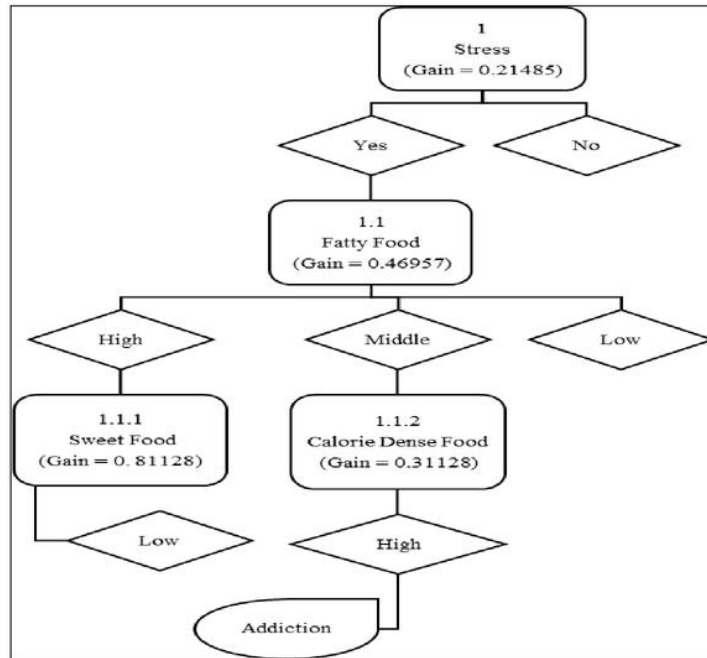
Figure 6. Decision Tree Classification

Table 28. Comparison

Methods	Algorithm
Naïve Bayes Algorithm	If [Calorie Dense Food] = "High", and if [Fatty Food] = "High", and if [Sweet Food] = "Middle", and if [Diet] = "No", and if [Stress] = "Yes", then Class = "Addiction"
Decision Tree Algorithm	If [Stress] = "Yes", and if [Fatty Food] = "Middle", and if [Calorie Dense Food] = "High", then Class = "Addiction"

Table 25 (a) and (b) show Entropy Fatty Food = Low to Entropy Total in this class. Table 26 shown information gain (S, CDF), gain (S, SF), gain (S, D) and gain (S, S). Table 27 shown attributes, entropy and gain. Information gain is equal. Conditions for stopping partitioning when all samples for a given nod belong to the same class. Figure 6 shown Decision Tree to making decision. Decision Tree by 3 Node. Node 1 = Stress, Node 1.1 = Fatty Food, Node 1.1.1 = Sweet Food, Node 1.1.2 = Calorie Dense Food. Final decision is Class = Addiction. Table 28 shown comparison between Naïve Bayes Algorithm and Decision Tree Algorithm in classification.

5. Conclusion

Classification using Naïve Bayes by criteria Calorie Dense Food (High), Fatty Food (High), Sweet Food (Middle), Diet (No), and Stress (Yes). Its factor influence Food

Addiction. Classification using Decision Tree by criteria Stress (Yes), Fatty Food (Middle), and Calorie Dense Food (High). Its factor influence Food Addiction.

Calorie Dense Food, Fatty Food, and Stress is the same factor on both algorithm in classification. Level criteria influence each class, depend on major or minor effect in each level criteria. Future work, classification on Food Addiction using Artificial Neural Network.

References

[1]   E. M. Schulte, M. A. Joyner, M. N. Potenza, C. M. Grilo and A. N. Gearhardt, "Current considerations regarding food addiction", Curr Psychiatry Rep, **(2015)**, pp. 17-19.
[2]   A. Muele and A. Gearhardt, "Five years of the yale food addiction scale: taking stock and moving forward", Current Addiction Reports, vol. 1, no. 3, **(2014)**, pp. 193-205.
[3]   J. Hebebrand, Ö. Albayrak, R. Adan, J. Antel, C. Dieguez, J. Jong, G. Leng, J. Menzies, J. G. Mercer, M. Murphy, G. Plasse and S. L. Dickson, "Eating addiction", rather than "food addiction", better captures addictive-like eating behavior", Neuroscience and Bio-behavioral Reviews: Elsevier, vol. 47, **(2014)**, pp. 295-306.
[4]   B. Kaptan and S. Kayısoglu, "Consumers' attitude towards food additives", American Journal of Food Science and Nutrition Research, vol. 2, no. 2, **(2015)**, pp. 21-25.
[5]   J. M. L. Cabrera, F. Carvajal and P. L. Legarrea, "Food addiction as a new piece of the obesity framework", Nutrition Journal, BioMed Central, **(2016)**, vol. 15, no. 5.
[6]   A. B. D. Ahmed and I. S. Elaraby, "Data mining: a prediction for student's performance using classification method", World Journal of Computer Application and Technology, vol. 2, no. 2, **(2014)**, pp. 43-47.
[7]   S. Saha and D. Nandi, "Data classification based on decision tree, rule generation, bayes and statistical methods: an empirical comparison", International Journal of Computer Applications, vol. 129, no.7, **(2015)**, pp. 36-41.
[8]   R. Asif, A. Merceron and M. K. Pathan, "Predicting student academic performance at degree level: a case study", I. J. Intelligent Systems and Applications, MECS Publisher, no. 01, **(2015)**, pp. 49-61.
[9]   M. Wahyudi, and S. N. N. Alfisahrin, "Comparison between c4.5 algorithm, naïve bayes, and neural network for prediction of cardiovascular", Journal Technology of Information and Communication, vol. 2, no. 2, **(2014)**, pp. 99-106.
[10] I. M. N. A. Yasa, I. K. G. D. Putra and N. M. I. M. Mandenni, "Hypertension expert system with c5.0 algorithm and fuzzy logic", Telkomnika Indonesian Journal of Electrical Engineering, vol. 12, no. 7, **(2014)**, pp. 5669-5677.
[11] H. Jin, S. Kim and J. Kim, "Decision factors on effective liver patient data prediction", International Journal of Bio-Science and Bio-Technology, vol. 6, no. 4, **(2014)**, pp. 167-178.
[12] M. A. Manjiri, "A new architecture for making moral agents based on c4.5 decision tree algorithm", I. J. Information Technology and Computer Science, MECS Publisher, no. 05, **(2014)**, pp. 50-57.
[13] B. F. Z. Bayaty and S. Joshi, "Comparative analysis between naïve bayes algorithm and decision tree to solve wsd using empirical approach", Lecture Notes on Software Engineering, vol. 4, no. 1, **(2016)**, pp. 82-86.
[14] R. Gupta and P. K. Shukla, "Performance analysis of anti-phishing tools and study of classification data mining algorithms for a novel anti-phishing system", I. J. Computer Network and Information Security, MECS Publisher, no. 12, **(2015)**, pp. 70-77.
[15] S. Khazaee and A. Bozorgmehr, "A new hybrid classification method for condensing of large datasets: a case study in the field of intrusion detection", I. J. Modern Education and Computer Science, MECS Publisher, no. 4, **(2015)**, pp. 32-41.
[16] T. Choudhury, V. Kumar and D. Nigam, "Intelligent classification & clustering of lung & oral cancer through decision tree & genetic algorithm", International Journal of Advanced Research in Computer Science and Software Engineering, vol. 5, no. 12, **(2015)**, pp. 501-510.

Artikel_Jurnal_2_-_Food_Addiction.pdf