NIPES

Journal of Science and Technology Research

Journal homepage: www.nipesjournals.org.ng

The Mutual Information Approach for Determining the Strength of Associations Between Features of Clinical Depression

Augusta Aghaulor
Department of Computer Science, Delta State University, Abraka, Nigeria
Email: augustacindy@gmail.com Phone: 08068765174

Article Info

Received 11 June 2021
Revised 03 July 2021
Accepted 06 July 2021
Available online 31 August 2021

***Keywords**:*
Clinical depression, mutual information, ict, alcohol

https://doi.org/10.37933/nipes/3.3.2021.7

https://nipesjournals.org.ng

Abstract

This study addresses a crucial challenge relating to key predictors and their association with human depression using information theory, focusing on mutual information. Mutual information is a well-known technique for determining the strength of statistical relationships between variables in healthcare and many other research fields. Finding mutual information using unbalanced and limited dataset data set is a demanding task. The results from the mutual information and information gain indicate high mutual relationship between "depression" and "alcohol or other drug consumption"; "depression" and family support and availability of accommodation". But a low mutual relationship between "depression" and "cigarette smoking". The results also indicate significant mutual relationship between "depression" and a synergy of "impaired function and alcohol and other drug consumption". Given the challenges posed by depression, it is hoped that the findings from the study will be among the current universal study for the inclusion of ICT model in the identification of the predictors of depression.

1.0 Introduction

In recent times, depression has been recognized as a widespread emotional discomfort in many areas of clinical literature. This health condition affects the entire society irrespective of sex, status, size, religious affiliation, cultural background, race, skin colour, or age. About 350,000 million people with no exception of any age group are estimated to be suffering from depressive disorder globally, from which estimated 800,000 people are lost yearly to suicide, a condition resulting from unidentified and inadequate management of depression [1]. According to the American Psychiatric Association [2], characteristics of depression include: feelings of guilt, hypersomnia, sad mood, loss of pleasure or interest, loss of energy, worthlessness, insomnia, inability to concentrate, withdrawal from social activities, disturbed appetite, and in worse situations suicidal ideation, diagnosing a patient of this condition requires at least two weeks consistent presence of a number of these symptoms. Consequently, resulting in impairment in the routine activities and/or conspicuous relationship challenges with people [3]. To diagnose that a patient is suffering from depression, a minimum of five from the nine listed criteria must be satisfied (Diagnostic and statistical manual of mental disorders, American Psychiatric Association [2].

Although depression is a widespread health problem, world health organization (WHO) in 2012, stated that third world countries are at a higher risk of been affected by depression due to limited mental health professionals and lack of state-of-the-art medical equipment used in managing

depression. Studies have shown that in Nigeria, for example, depression is almost totally neglected [3] and this gives more than 30% of all the outpatient attendance; more than 20% of all admitted patients in majority of the hospitals in Nigeria [4]; also, more than 35% of all death within dedicated psychiatric hospitals in the country is as a result of depression. According to WHO [1], in every 10 persons suffering from mental health in Nigeria, 9 do not get any medical care. This high negligence according to them, is increasingly threatening public health.
As socially advanced as Nigeria is, survey by many scholars showed that higher institution students here are at a higher rate of being depressed compared to students of other countries [5]–[7]. According to Afolabi et al [8], Adewuya et al [9] and Peltzer et al [10], depression rate among university students is significantly higher than other population within the country, this is as a result of poor academic performance, high alcohol consumption, inadequate family support, lack of accommodation, and cigarette smoking. Although some options are publicly available for detecting and treating depression, it is often misdiagnosed and wrongly-treated [3]. WHO [1] stated that this disease has greatly affected the social and economic situation of Nigeria and many developing countries. However, Gureje et al [11] suggested that the impact of this disease can be minimized with clear knowledge of the predictors and the association between the predictors of depression. A number of modern computing and statistical tools have been used to reveal the association and dependency among variables in the medical literature. This includes Cohen's kappa coefficient and odds ratio [12]. This study focuses on establishing each feature's (symptoms) of depression strength and the relationships that exist between the symptoms and the class target (depression). Given the effectiveness and the simplicity of mutual information in bioinformatics [13], text categorisation Gao et al [14], and many other AI solutions [15], this work uses mutual information an information theoretic approach, in finding the significance of the predictors (symptoms, in this case) and calculating the dependence relationships between the symptoms and depression. Mutual information has been applied in several studies to compare multiple healthcare survey instruments [16]–[20]. For instance, Hernández & Samengo [20] proposed a novel estimator for mutual information of discrete variables X and Y, which was adequate when X had a much larger number of effective states than Y.

1.2 Information Gain (IG)

An information gain refers to the method through which the relevancy of individual symptoms of depression are scored and ranked. This defines the amount of information for each attribute with regards to its classification target [21] by looking at individual feature in isolation. It further compute the gain, determine how crucial and significant it is to the class label based on mutual information and likely communication amongst features [22]. In other to achieve this, only the features with reasonably large information gain are selected. Given that entropy is the general measure for the information [23], computation for IG begins with the calculation of entropy for the target class. According to [23] Entropy is defined by:

$$info(D) = -\sum_{i=1}^{m} p_i log_2(p_i) \qquad 1$$

Where:
$info(D)$ denoted the information required to classify a tuple in D,this is as well referred to entropy of D
D refers to dataset sample size
p_i defines the fraction of D with regards to class target
m denotes the total number of likely outcome.
The extreme entropy values for $info(D)_{max}$ are 0 (perfectly classified) the deterministic value and 1 (totally random).

The next step is to calculate the expected information needed in classifying a tuple from D based on the partitioning of attribute A.
The expression is:

$$info_A(D) = \sum_{j=1}^{v} (|\ D_j\ |/|\ D\ |) * info(D_j) \qquad 2$$

where D_j is the subset of D containing distinct value of A and v is the number of distinct values in A
To measure information gain, the relative difference between posterior entropy and prior entropy of class is calculated.

$$Gain(A) = info(D) - info_A(D) \qquad 3$$

Doquire and Verleysen [24] opined that mutual information attempt to calculate the amount of information that is averagely communicated in one random variable about another. In other words, mutual information states how much to which the joint probability of the predictor (in this case, symptom), and the target (depression) deviates from what it would be if the predictor was independent of the target [25-26]. For instance, if Y represents the rolling of a fair 6-sided die, and X represents if the roll is odd (0 if even, 1 if odd). Without doubt, the value of X gives information about the value of Y and vice versa. This therefore means that mutual information is shared among these variables. In this study, mutual information helps in the accurate measurement of the amount of information that individual symptom of depression has over each other or depression itself, as such minimizing the uncertainty of each symptom against each other. Consequently, the value of mutual information of symptoms, and the total contribution of each symptom to the target nodes (depression) were computed. : Equation 4 to 6 shows the mutual information between a predictor E and a target C:

$$\mathrm{MI}(C, \mathrm{E}) = \mathrm{H}(C) - \mathrm{H}(\mathrm{C/E}) \qquad 4$$

$$\approx \sum_{\mathrm{c} \in \mathrm{C}} \sum_{\mathrm{e} \in \mathrm{E}} \mathrm{P(c,e)} \log_2 \frac{\mathrm{P(c,e)}}{\mathrm{P}(c)\mathrm{P(e)}} \qquad 5$$

$$\approx \sum_{\mathrm{e} \in \mathrm{E}} \mathrm{P(e)} \sum_{\mathrm{c} \in \mathrm{C}} \mathrm{P(c/e)} \log_2 \frac{\mathrm{P(c/e)}}{\mathrm{P}(c)} \qquad 6$$

Where:
$\mathrm{H}(C)$ defines the entropy of class variable in the training set.
$\mathrm{H(C/E)}$ defines entropy of class variable if C feature is given.
According to Conrady & Jouffe [26], the equation helps in the computation of the mutual information between any possible predictors and classification target. Consequently, both the predictor which gives predictive importance and overall information gain are found easily. Remarkable success has been recorded by many authors from the application of mutual information in the determination of relevancy between symptom features and target class.
Below is the general algorithm used for implementing information gain:

Function IG $\mathrm{C/E}$ feature ranking-based entropy
initialisation:
$\mathrm{S} = 0$;
C ← domain of a class label;
E← domain of an attribute values;
For each $c_i \epsilon$ C do:
    calculate $\mathrm{P(c[i])}$;
    $\mathrm{H_c = S + P(c[i)) * \log_2 P(c[i]))}$;
    $\mathrm{S \leftarrow H_c}$:
End For

```
For each e_i ∈ E:
        calculate P(e[j])
        Sum = S + P(e[j]) * log_2 P(e[j]));
        C ← Sum;
End For
For each c_i do:
        For each e_j do:
                calculate P(c[i] | e[j])
                M = S + P( c[i] | e[j] ) * log_2 P( c[i] | e[j] );
                S← M;
        End For
End For
H(C/E) = (−1) * Sum * (−1) * M;
IG = H_c − H(C/E)
return IG
End function
```

2. Methodology

As mentioned earlier, the study focuses on determining the statistical relationships between the symptoms of depression and their contributions towards depression using mutual information technique. The dataset was analysed using a software tool, BaysiaLab 6, a universal analytics platform, which provides scientists, researchers and practitioners with a comprehensive "lab" environment for machine learning, knowledge modeling, diagnosis, analysis, simulation, and optimization [26].

In the study, below is the procedure proposed for the computation of information gain:

1. Prepare the depression dataset
2. Find the entropy for the target class (that is, depression)
3. Determine the mutual information values between the symptoms and depression in the dataset. Probability density estimation method was used to extract the linear and nonlinear relationship between depression symptoms.
4. Determine the overall contribution of each symptom to the target class

In the information-theoretic sense, mutual information attempt to measure the amount of information the absence or presents of a symptom can contribute in making accurate classification decision on depression. The above-stated four-step procedure was followed for information again and original depression dataset in BayesiaLab 6 was used for testing [26].

Just as it was described by Korb and Nicholson [25] that mutual information is symmetric, it therefore means that equal amount of mutual information is generally reported by both predictor on the target as well as target on the predictor.

Put mathematically,

$\mathrm{MI}(C{:}\,\mathrm{E}) = \mathrm{MI}(\mathrm{E}{:}\,\mathrm{C}) = 0$ iff C *and* E are independent statistically.

Given the above concept, also the data items involved; depression and its symptoms, Table 3 represents the ranking of each symptom's overall mutual information contribution to depression using a relevant statistical measure.

2.1 Data collection and description

The researcher started by firstly conducting a semi-structured interview with depression patients. Thereafter, records of previously diagnosed mental health patients of the mental health section of University of Benin Teaching Hospital (UBTH) and one other primary healthcare centre (PHC) were extracted. The dataset comprises a total of 1789 data instances of which 1020 of them were male and 778 were female. The age range is maximum 92 and minimum 12, the standard deviation is 13.92 while the mean age is 42.55. Initially, the dataset contained one class attribute (depression) and main attributes of a total of 27. The data set attributes value (the degree of depression), codes, and the types of the data are shown in table 1. The depression class attributes are classified into four(4) categories: severe depression (3), moderate depression (2), mild depression (1), and No depression (0). In the selection process, efforts were made to ensure that the clinical criteria for the selection of depression cases as stipulated by International classification of diseases-10th edition (ICD-10) [27] and diagnostic and statistical manual-fifth edition (DSM-5) [2] were religiously followed.

Table 1: Initial dataset and attribute value

S/N	Attributes	Code	Values		Data type
			Absent	Present	
1	Sad mood	SM	0	1	Integer
2	Thought of suicide	SU	0	1	Integer
3	Loss of pleasure	LP	0	1	Integer
4	insomnia	IN	0	1	Integer
5	Hypersomnia	HY	0	1	Integer
6	Loss of appetite	LA	0	1	Integer
7	Psychomotor agitation	PA	0	1	Integer
8	Psychomotor retardation	PR	0	1	Integer
9	Loss of energy	LE	0	1	Integer
10	Feeling of worthlessness	FW	0	1	Integer
11	Lack of thinking	LT	0	1	Integer
12	Indecisiveness	ID	0	1	Integer
13	Recurrent thought of death	TD	0	1	Integer
14	Impaired function	IF	0	1	Integer
15	Weight gain	WG	0	1	Integer
16	Weight loss	WL	0	1	Integer
17	Employment status	ES	0	1	Integer
18	Depression in family	DF	0	1	Integer
19	Stressful life events	SL	0	1	Integer
20	Financial pressure	FP	0	1	Integer
21	Level of family support	LS	None: 0; low: 1; medium: 2, High: 3		Integer
22	Availability of Accommodation	AA	None: 0; low: 1; medium: 2, High: 3		Integer
23	Alcohol consumption	AC	None: 0; low: 1; medium: 2, High: 3		Integer
24	Other drug consumption	DC	None: 0; low: 1; medium: 2, High: 3		Integer
25	Job satisfaction	JS	None: 0; low: 1; medium: 2, High: 3		Integer
26	Academic performance	AP	None: 0; low: 1; medium: 2, High: 3		Integer
27	Cigarette smoking	CS	None: 0; low: 1; medium: 2, High: 3		Integer
	Depression diagnosis	**DG**	**None: 0; mild: 1; Moderate: 2, severe: 3**		**Nominal**

2.2. Data validation

In other to authenticate the dataset integrity, eight healthcare professionals from six different areas of specialization were employed. While three are Psychiatrists, two specialize in Clinical Psychologists, one is a Nurse Psychotherapists, and another one is a Clinical Social Worker, then the last, a Child/Adolescent Psychiatrist. For the control group, a total of 372 no depression cases were also collected. These two datasets were combined, and their attributes were also merged by these healthcare experts for the purpose of compatibility. On some occasions were the attributes did not match, the PHC team was consulted by the mental healthcare experts for clarification. After careful considerations by the mental healthcare experts, certain attributes in the dataset were merged. For example, accommodation availability and level of family support were merged, alcohol and consumption of other drugs were merged, and academic performance and job satisfaction were also merged. Consequently, the predictor variables of Table 2 from serial number 1-24 and the diagnosis which is in boldface were selected.

Table 2 Selected attribute set

S/N	Attributes	Code	Values		Data type
			Absent	Present	
1	Sad mood	SM	0	1	Integer
2	Thought of suicide	SU	0	1	Integer
3	Loss of pleasure	LP	0	1	Integer
4	insomnia	IN	0	1	Integer
5	Hypersomnia	HY	0	1	Integer
6	Loss of appetite	LA	0	1	Integer
7	Psychomotor agitation	PA	0	1	Integer
8	Psychomotor retardation	PR	0	1	Integer
9	Loss of energy	LE	0	1	Integer
10	Feeling of worthlessness	FW	0	1	Integer
11	Lack of thinking	LT	0	1	Integer
12	Indecisiveness	ID	0	1	Integer
13	Recurrent thought of death	TD	0	1	Integer
14	Impaired function	IF	0	1	Integer
15	Weight gain	WG	0	1	Integer
16	Weight loss	WL	0	1	Integer
17	Employment status	ES	0	1	Integer
18	Depression in family	DF	0	1	Integer
19	Stressful life events	SL	0	1	Integer
20	Financial pressure	FP	0	1	Integer
21	Level of family support availability of accommodation	LS	None: 0; low: 1; medium: 2, High: 3		Integer
22	Alcohol or drug consumption	AC	None: 0; low: 1; medium: 2, High: 3		Integer
23	Job satisfaction/Academic performance	JS	None: 0; low: 1; medium: 2, High: 3		Integer
24	Cigarette smoking	CS	None: 0; low: 1; medium: 2, High: 3		Integer
	Depression diagnosis	**DG**	**None: 0; mild: 1; Moderate: 2, severe: 3**		**Nominal**

Some pre-processing activities were carried out on the dataset both on instance level and structurally before it could be used. These activities are, dataset transposition into row and columns, value conversion and attribute normalization, data conversion into tables form, features merging, and data type's conversions. Table two is the result from the pre-processing of the dataset. The dataset trained was represented using Boolean format which showed depression presence in two major levels: absence (0) and presence (1) for symptoms within serial number 1 – 20), while the symptoms degree of presence was categorized within serial number 1-24. The essence of the pre-processing activities was to satisfy the machine learning analysis tool requirements and quality of data element as specified by Cai and Zhu [28] and Murphy [29].

3. Results and discussion

This research work, mutual information between symptom of depression and depression was carried out in other to ascertain symptoms strength if combined with respect to the class target and also determine significance of individual symptom relative to the classification target. The Table 3 below shows MI with its class level total contributions in BayesiaLab (version 6) from the depression dataset collection to processing [26].

Table 3: Mutual information and Symptoms contribution to depression

S/N	Parent/ Target/depression	Symptom	Mutual information	Overall contribution (%)
1	Depression	Sad mood	0.0863	2.3532
2	Depression	Lack of thinking	0.3101	8.4521
3	Depression	Loss of appetite	0.2022	5.5093
4	Depression	Impaired function	0.6309	17.1944
5	Depression	Loss of energy	0.5621	15.318
6	Depression	Indecisiveness	0.3069	8.3627
7	Depression	Suicide attempt	0.0680	1.8526
8	Depression	Alcohol or other drug consumption	0.3098	8.3152
9	Depression	Loss of pleasure	0.1618	4.4101
10	Depression	Stressful life events	0.0202	0.5503
11	Depression	Insomnia	0.1179	3.2135
12	Depression	Job satisfaction or academic performance	0.1310	3.8041
13	Depression	Recurrent thought of death	0.1898	5.1730
14	Depression	Weight loss	0.3445	9.3879
15	Depression	Hypersomnia	0.0214	0.5825
16	Depression	Weight gain	0.0352	0.9602
7	Depression	Worthlessness	0.3650	9.9484
18	Depression	Psychomotor retardation	0.1896	5.1674
19	Depression	Depression in family	0.0277	0.7540
20	Depression	Family support and availability of accommodation	0.1774	5.1689
21	Depression	Psychomotor agitation	0.0171	0.4654
22	Depression	Financial pressure	0.032	0.0873
23	Depression	Cigarette smoking	0.0152	0.0550
24	Depression	Employment status	0.0097	0.0257

As shown in Table 3, ‘impaired function’ exhibited highest mutual information of 0.6309 with also the highest overall contribution of 17.1944% towards depression. Loss of energy was the second highest with 0.5621 and 15.318% overall mutual information contribution towards depression. The third major contributor is Worthlessness with mutual information of 0.3650 and 9.9484 contribution to depression. Surprisingly, alcohol and consumption of other drugs, showed 0.3098 mutual information with 8.3152% contributor to depression making it the 6th highest contributor. The 10th highest contributor is ‘accommodation availability’ and level of family support with 0.1774 mutual information and a total of 5.1689% to depression. In position 20 is ‘Stressful life events’ with 0.0202 mutual information and 0.5503 contribution to depression. Noticeably is financial pressure with 0.032 is mutual information and 0.0873 total contribution. The least is ‘Employment status’ with 0.0097 mutual information and also the lowest overall contributor of 0.0257% to depression.

Table 4 shows the synergistic combination of symptoms, though symptoms with positive contributions were selected. The result showed that ‘sad mood’ and ‘loss of appetite’ topped the table with a mean synergy of 6.9968. Followed closely is ‘weight loss’ and ‘sad mood’ combination with a total of 6.7256 mean synergy. In the middle of the table is synergy of Indecisiveness and Psychomotor agitation with 0.1434 mean value. Lastly, ‘weight gain’, recurrent thought of death’ and ‘suicide attempt’ synergy exhibited the lowest contribution to depression with 0.0003 mean synergy.

Table 4 Synergistic combination of depression symptoms to depression

s/n	Symptom	Symptom	Mean synergy (%)	S/N	Symptom	Symptom	Mean synergy (%)
1	Sad mood	Loss of appetite	6.9968	23	Weight loss	Lack of thinking	0.8029
2	Weight loss	Sad mood	6.7256	24	Depression in family	Impaired function	0.7392
3	Insomnia	Loss of pleasure	4.5318	25	Job satisfaction or academic performance	Sad mood	0.6845
4	Loss of appetite	Loss of pleasure	4.3210	26	Impaired function	Cigarette smoking	0.6705
5	Alcohol or other drug consumption	Impaired function	4.1431	27	Weight loss	Recurrent thought of death	0.6430
6	Loss of energy	Insomnia	4.0504	28	Weight loss	Psychomotor retardation	0.6206
7	Sad mood	Impaired function	3.6677	29	Psychomotor agitation	Impaired function	0.5298
8	Loss of energy	Loss of pleasure	3.2593	30	Depression in family	Indecisiveness	0.5031
9	Loss of appetite	Insomnia	3.1744	31	Lack of thinking	Depression in family	0.4873
10	Loss of energy	Loss of appetite	2.5662	32	Suicide attempt	Impaired function	0.4427
11	Worthlessness	Impaired function	2.1636	33	Worthlessness	Depression in family	0.4084
12	Indecisiveness	Impaired function	1.7308	34	Indecisiveness	Job satisfaction or academic performance	0.3392
13	Impaired function	Lack of thinking	1.6686	35	Cigarette smoking	Weight loss	0.3210
14	Impaired function	Recurrent thought of death	1.4990	36	Depression in family	Psychomotor retardation	0.2645
15	Job satisfaction or academic performance	Family support and availability of accommodation	1.3994	37	Depression in family	Weight loss	0.2423
16	Impaired function	Psychomotor retardation	1.3568	38	Weight loss	Suicide attempt	0.2318
17	Alcohol or other drug consumption	Sad mood	1.2224	39	Recurrent thought of death	Depression in family	0.2210
18	Worthlessness	Weight loss	1.0756	40	Psychomotor agitation	Worthlessness	0.2161
19	Weight loss	Impaired function	1.0288	41	Stressful life events	Loss of pleasure	0.2081
20	Sad mood	Insomnia	0.8742	42	Weight loss	Psychomotor agitation	0.2027

21	Weight loss	Indecisiveness	0.8229	43	Financial pressure	Alcohol or other drug consumption	0.1786
22	Family support and availability of accommodation	Weight gain	0.8125	44	Weight gain	Hypersomnia	0.1684

S/N	Symptom	Symptom	Mean synergy (%)	S/N	Symptom	Symptom	Mean synergy (%)
45	Lack of thinking	Psychomotor agitation	0.1622	68	Stressful life events	Loss of appetite	0.0611
46	Depression in family	Psychomotor agitation	0.1608	69	Hypersomnia	Indecisiveness	0.0569
47	Loss of pleasure	Job satisfaction or academic performance	0.1600	70	Psychomotor agitation	Suicide attempt	0.0541
48	Cigarette smoking	Psychomotor agitation	0.1573	71	Recurrent thought of death	Family support and availability of accommodation	0.0517
49	Alcohol or other drug consumption	Family support or availability of accommodation	0.1552	72	Weight loss	Hypersomnia	0.0510
50	Family support and availability	Weight loss	0.1541	73	Lack of thinking	Job satisfaction or academic performance	0.05024
51	Hypersomnia	Impaired function	0.1532	74	Weight gain	Lack of thinking	0.0490
52	Psychomotor agitation	Recurrent thought of death	0.1446	75	Suicide attempt	Psychomotor agitation	0.0482
53	Indecisiveness	Psychomotor agitation	0.1434	76	Alcohol or other drug consumption	Insomnia	0.0444
54	Family support and availability of accommodation	Impaired function	0.14008	77	Family support and availability of accommodation	Sad mood	0.0442
55	Alcohol or other drug consumption	Suicide attempt	0.13414	78	Weight loss	Weight gain	0.0441
56	Impaired function	Weight gain	0.1271	79	Worthlessness	Financial pressure	0.0440
57	Depression in family	Cigarette smoking	0.1252	80	Financial pressure	Impaired function	0.0417
58	Worthlessness	Hypersomnia	0.1199	81	Weight gain	Psychomotor retardation	0.0407
59	Hypersomnia	Lack of thinking	0.1180	82	Psychomotor agitation	Weight gain	0.0378
60	Hypersomnia	Psychomotor retardation	0.1178	83	Financial pressure	Recurrent thought of death	0.0328
61	Depression in family	Hypersomnia	0.1090	84	Alcohol or other drug consumption	Weight loss	0.0315

62	Hypersomnia	Recurrent thought of death	0.1025	85	Financial pressure	Lack of thinking	0.0309
63	Stressful life events	Insomnia	0.1022	86	Hypersomnia	Psychomotor agitation	0.0307
64	Weight gain	Indecisiveness	0.0885	87	Indecisiveness	Financial pressure	0.0295
65	Weight gain	Depression in family	0.0851	88	Psychomotor retardation	Financial pressure	0.0260
66	Loss of energy	Stressful life events	0.0781	89	Psychomotor agitation	Psychomotor retardation	0.0259
67	Suicide attempt	Depression in family	0.0731	90	Weight gain	worthlessness	0.0258

S/N	**Symptom**	**Symptom**	**Mean synergy (%)**	**S/N**	**Symptom**	**Symptom**	**Mean synergy (%)**
91	Weight gain	Recurrent thought of death	0.0252	99	Financial pressure	Hypersomnia	0.0039
92	Financial pressure	Cigarette smoking	0.0225	100	Weight gain	Financial pressure	0.0023
93	Suicide attempt	Family support and availability of accommodation	0.02147	101	Suicide attempt	Hypersomnia	0.0020
94	Weight loss	Financial pressure	0.0187	102	Cigarette smoking	Hypersomnia	0.0014
95	Suicide attempt	Financial pressure	0.01402	103	Employment status	Job satisfaction or academic performance	0.010
96	Depression in family	Financial pressure	0.0134	104	Weight gain	Cigarette smoking	0.0009
97	Psychomotor agitation	Financial pressure	0.0090	105	Family support and availability of accommodation	Hypersomnia	0.0006
98	Job satisfaction or academic performance	Lack of energy	0.0077	106	Weight gain	Suicide attempt	0.0003

4. Conclusion

Given the increasing rate of depression among the Nigerian university students, it will be highly inappropriate and detrimental if nothing is urgently done. As such, it is essential that multiple solutions that help reduce the detrimental effects of depression be sought. Establishing the statistical associations between the features of depression and their contributions towards depression with ICT tools is a promising step. The study focused on the use of mutual information to measure the strength of association between the various symptoms of depression. The sampled data was collected from two mental health divisions: University of Benin Teaching Hospital, Nigeria and primary healthcare centre, and was validated by eight experienced mental health professionals. The results reveal the several number of unexplored possibilities for building clinical decision support systems. In particular, the results showed the rich capability of ICT-aided support, in relation to association analysis of the symptoms of depression.

References

[1] WHO, “Depression: A global public health concern. Available at http://www.who.int/mental_health/management/depression/who_paper_depression_wfmh_2012.pdf,” WHO Department of Mental Health and Substance Abuse, 2012. [Online]. Available: http://www.who.int/mental_health/management/depression/who_paper_depression_wfmh_2012.pdf.

[2] American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*, Fift Editi. Washington, DC: American Psychiatric Association, 2013.

[3] M. Abas, G.-C. Ali, E. Nakimuli-Mpundu, and D. Chabanda, “Depression in people living with HIV in Subsaharan Africa,” *Trop. Med. Int.*, vol. 19, no. 12, pp. 1392–1396, 2014, doi: 10.1111/tmi.12382.

[4] O. Gureje, R. Uwakwe, B. Oladeji, V. O. Makanjuola, and O. Esan, “Depression in adult Nigerians: Results from the Nigerian Survey of Mental Health and Well-being,” *J. Affect. Disord.*, vol. 120, no. 1–3, pp. 158–164, 2010, doi: 10.1016/j.jad.2009.04.030.

[5] C. J. Othieno, R. O. Okoth, K. Peltzer, S. Pengpid, and L. O. Malla, “Depression among university students in Kenya: Prevalence and sociodemographic correlates,” *J. Affect. Disord.*, vol. 165, pp. 120–125, 2014, doi: 10.1016/j.jad.2014.04.070.

[6] A. Ibrahim, S. Kelly, C. Adams, and C. Glazebrook, “A systematic review of studies of depression prevalence in university students,” *Journal of Psychiatric Research*, vol. 47, no. 3. pp. 391–400, 2013, doi: 10.1016/j.jpsychires.2012.11.015.

[7] D. Eisenberg, S. E. Gollust, E. Golberstein, and J. L. Hefner, “Prevalence and correlates of depression, anxiety, and suicidality among university students.,” *Am. J. Orthopsychiatry*, vol. 77, no. 4, pp. 534–542, 2007, doi: 10.1037/0002-9432.77.4.534.

[8] M. O. Afolabi, E. a Abioye-Kuteyi, F. O. Fatoye, I. S. Bello, and A. O. Adewuya, “Pattern of Depression among Patients in a Nigerian Family Practice Population,” *South African Fam. Pract.*, vol. 50, no. 2, 2008, [Online]. Available: http://www.safpj.co.za/index.php/safpj/article/view/858.

[9] A. O. Adewuya, B. A. Ola, O. O. Aloba, B. M. Mapayi, and O. O. Oginni, “Depression amongst Nigerian university students: Prevalence and sociodemographic correlates,” *Soc. Psychiatry Psychiatr. Epidemiol.*, vol. 41, no. 8, pp. 674–678, 2006, doi: 10.1007/s00127-006-0068-9.

[10] K. Peltzer, S. Pengpid, S. Olowu, and M. Olasupo, “Depression and Associated Factors Among University Students in Western Nigeria,” *J. Psychol. Africa*, vol. 23, no. 3, pp. 459–466, 2013, doi: 1433-0237.

[11] O. Gureje, J. Abdulmalik, L. Kola, E. Musa, M. T. Yasamy, and K. Adebayo, “Integrating mental health into primary care in Nigeria: report of a demonstration project using the mental health gap action programme intervention guide.,” *BMC Health Serv. Res.*, vol. 15, no. 1, p. 242, 2015, doi: 10.1186/s12913-015-0911-3.

[12] Y. Kang, M. R. Steis, A. M. Kolanowski, D. Fick, and V. V. Prabhu, “Measuring agreement between healthcare survey instruments using mutual information,” *BMC Med. Inform. Decis. Mak.*, vol. 16, no. 1, pp. 1–11, 2016, doi: 10.1186/s12911-016-0335-y.

[13] L. Wang, Y. Wang, and Q. Chang, “Feature selection methods for big data bioinformatics: A survey from the search perspective,” *Methods*. 2016, doi: 10.1016/j.ymeth.2016.08.014.

[14] Z. Gao, Y. Xu, F. Meng, F. Qi, and Z. Lin, “Improved information gain-based feature selection for text categorization,” 2014, doi: 10.1109/VITAE.2014.6934421.

[15] G. Chandrashekar and F. Sahin, “A survey on feature selection methods,” *Comput. Electr. Eng.*, 2014, doi: 10.1016/j.compeleceng.2013.11.024.

[16] P. S. Russell, M. Basker, S. Russell, P. D. Moses, M. K. C. Nair, and K. A. Minju, “Comparison of a self-rated and a clinician-rated measure for identifying depression among adolescents in a primary-care setting,” in *Indian Journal of Pediatrics*, Jan. 2012, vol. 79, no. SUPPL. 1, pp. 45–51, doi: 10.1007/s12098-011-0438-5.

[17] J. A. Seago, “A comparison of two patient classification instruments in an acute care hospital,” *J. Nurs. Adm.*, vol. 32, no. 5, pp. 243–249, 2002, doi: 10.1097/00005110-200205000-00004.

[18] P. N. Tan, V. Kumar, and J. Srivastava, “Selecting the right objective measure for association analysis,” in *Information Systems*, Jun. 2004, vol. 29, no. 4, pp. 293–313, doi: 10.1016/S0306-4379(03)00072-3.

[19] [A. E. Nicholson and N. Jitnah, “Using mutual information to determine relevance in Bayesian networks,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 1531, no. October, pp. 399–410, 1998, doi: 10.1007/BFb0095287.

[20] D. G. Hernández and I. Samengo, “Estimating the mutual information between two discrete, asymmetric variables with limited samples,” *Entropy*, vol. 21, no. 6, 2019, doi: 10.3390/e21060623.

[21] T. A. Alhaj, M. M. Siraj, A. Zainal, H. T. Elshoush, and F. Elhaj, “Feature selection using information gain for improved structural-based alert correlation,” *PLoS One*, 2016, doi: 10.1371/journal.pone.0166017.

[22] M. Bennasar, Y. Hicks, and R. Setchi, “Feature selection using Joint Mutual Information Maximisation,” *Expert Syst. Appl.*, 2015, doi: 10.1016/j.eswa.2015.07.007.

[23] C. E. Shannon, “A mathematical theory of communication,” *Bell Syst. Tech. J.*, 1948, doi: 10.1145/584091.584093.
[24] G. Doquire and M. Verleysen, “Mutual information-based feature selection for multilabel classification,” *Neurocomputing*, vol. 122, pp. 148–155, 2013, doi: 10.1016/j.neucom.2013.06.035.
[25] K. Korb and A. Nicholson, *Bayesian Artificial Intelligence*, Second edi., vol. 2, no. 1. Chapman & Hall, 2004.
[26] S. Conrady and L. Jouffe, *Bayesian Networks & BayesiaLab: A Practical Introduction for Researchers*, vol. 1, no. 1. Bayesia USA, 2015.
[27] World Health Organization, *The ICD-10 Classification of Mental and Behavioural Disorders*, vol. 10. 1992.
[28] L. Cai and Y. Zhu, “The Challenges of Data Quality and Data Quality Assessment in the Big Data Era,” *Data Sci. J.*, vol. 14, no. 0, p. 2, 2015, doi: 10.5334/dsj-2015-002.
[29] K. Murphy, *Machine Learning: A Probabilistic Perspective*. The MIT Press, Cambridge, 2012.