Probabilistic Expert Systems for Reasoning in Clinical Depressive Disorders

Blessing Ojeme[1], Audrey Mbogho[2], Thomas Meyer[3]
[1,3]CAIR, Department of Computer Science, and CSIR, University of Cape Town, Rondebosch 7701, Cape Town, South Africa
[1]bojeme@cs.uct.ac.za, [3]tmeyer@cs.uct.ac.za
[2]Department of Mathematics and Physics, Pwani University, Kilifi, Kenya
[2]a.mbogho@pu.ac.ke

***Abstract*—Like other real-world problems, reasoning in clinical depression presents cognitive challenges for clinicians. This is due to the presence of co-occuring diseases, incomplete data, uncertain knowledge, and the vast amount of data to be analysed. Current approaches rely heavily on the experience, knowledge, and subjective opinions of clinicians, creating scalability issues. Automating this process requires a good knowledge representation technique to capture the knowledge of the domain experts, and multidimensional inferential reasoning approaches that can utilise a few bits and pieces of information for efficient reasoning. This study presents knowledge-based system with variants of Bayesian network models for efficient inferential reasoning, translating from available fragmented depression data to the desired information in a visually interpretable and transparent manner. Mutual information, a Conditional independence test-based method was used to learn the classifiers.**

Keywords—Bayesian networks; probability; expert systems; depressive disorders; inferential reasoning; graphical models.

I. INTRODUCTION

First developed in the mid-1960s [1] as an important applied subfield of artificial intelligence (AI), expert systems (ES) attempt to solve complex problems in a particular domain by mimicking human experts' problem-solving methodology. ES have two main objectives: task-shifting from human experts to machine or non-experts [2] and sharing knowledge and know-how (as a way of reducing the gap between the have and have-nots [3] in the form of information products. By mimicking the problem-solving methods of human expert, ES help human experts to get a second opinion in decision-making and also act as advice to non-experts in a particular domain. For ES to perform at this human-expert level, knowledge is provided by a human expert, which it integrates and makes available in readable and understandable formats. Variants of ES (see Fig.1) include rule-based ES (RBES), frame-based ES (FBES), fuzzy logic ES (FLES), neural-based ES (NBES) and probability-based ES (PBES) [4]. Knowledge representation techniques for these various ES are production rules, frames, fuzzy rules, a combination of production rules and neural networks, and Bayesian networks (BN), respectively [4].

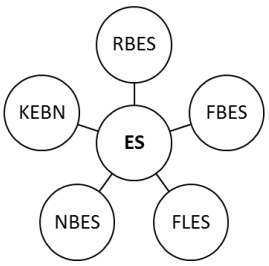

Fig. 1. Variants of expert systems

PBES, also called knowledge-based systems or knowledge engineering with Bayesian networks (KEBN) [5], which is the primary focus of this paper, uses BN for exact and approximate modeling of physical and biological systems [6]. The decision to use KEBN as the modeling platform was influenced by its strength for simplifying conditionalization and for handling uncertainty using probabilistic representation [6]. With respect to the objective of this study, the limitations of the other ES were another influencing factor. For instance, besides the inability to reason omni-directionally, the traditional RBES, based upon Buchanan and Shortliffe's MYCIN [7] for inference and decision-making has proved to be both brittle and cumbersome where problems are not well defined [5], [8]. Again, the single-disorder assumption of the RBES [9], that a patient only suffers from one disorder at a time, renders it ineffective for reasoning in the depression domain since depression has high comorbidity with other mental and/or physical illnesses [10], [11]. FBES, although they provide a natural way of describing the features and properties of objects in slots, may not be effective for knowledge representation in a complex domain such as depression because of similar drawbacks as in RBES [4]. Even though FLES are faster than RBES and FBES because of the smaller number of rules they require, Negnevitsky [4] highlights some major drawbacks, which make them unsuitable for building medical diagnostic decision support systems. These include huge computational overhead and

allocation of equal importance to all symptoms that are combined in the diagnostic process. Another drawback is difficulties of tuning in fuzzy systems, and the rules of combining membership functions, known as the min-max rules for conjunctive (AND) and disjunctive (OR) reasoning that, do not fit the human-expert reasoning process in disease diagnosis. Though the proficiency of the NBES in building medical diagnosis support systems has been well explored [12], Ahmed et al [13] noted that it is sensitive to data formats and requires large datasets to produce reliable results. On the other hand, the difficulty of access to mental health datasets has been well discussed by Doherty et al [14] and in most cases comes as a mixture of symbolic, textual and numeric data, but neural networks works best only when the data is in numeric formats.

As in other real-world problems, reasoning in clinical depression problems presents enormous cognitive challenges for clinicians of all categories. Adding uncertainty regarding the structure of the domain itself and the high comorbidity with physical and/or mental disorders with depression compounds the challenges. Omni-directional inferencing capability (reason from cause to effect, or from effect to cause), inherent ability to explicitly model uncertainty [5], and a combination of principles from numerous disciplines (see Fig. 2) [15], make KEBN suitable for direct representations of many complex problem domains such as depression.

The main contributions of this paper are the design and description of probabilistic (graphical) system models for addressing problems in representation, inference and knowledge engineering within the decision-making process of clinical depression.

The rest of the paper is structured as follows: Section II discusses the advantages of learning graphical KEBN models. Section III presents the step-by-step methodology taken to achieve the objective of the study. Section IV presents the experimental results and their analysis. Section V discusses the evaluation metrics used to measure the performance of our models. Section VI presents the works that are closely related to our study. Section VII concludes the study and sets out an agenda for future work.

II. LEARNING GRAPHICAL KEBN MODELS

A KEBN is a graphical model or probabilistic networks <N, A, Θ> [16], which is a directed acyclic graph (DAG) with nodes N, and directed arcs A, between the nodes, augmented by a conditional probability table (CPT) for each node, collectively represented by Θ. The network DAG and CPT present one convenient way of representing assumptions of conditional independence (CI). Each node n ϵ N in the graph represents a random variable X (an attribute in a dataset), and has a value corresponding to the probability of the random variable, P(X) [17].

Each arc a ϵ A between the nodes represents a probabilistic dependency, for instance, a direct arc from node X to node

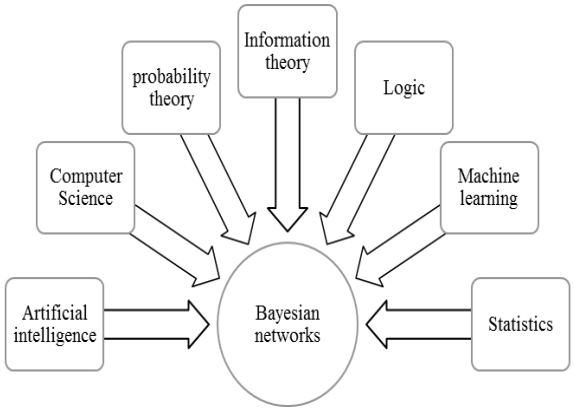

Fig. 2. Disciplines of KEBN

Y shows that node X has a direct influence on node Y, written as P(Y|X). The nodes and the arcs define the structure of the network, and the conditional probabilities are the parameters given the structure. Learning a graphical model has two parts [5]: The first is the learning of parameters given a structure. The second, more difficult and interesting part, is to learn the graph structure. Though much effort is required to learn the structure from depression data, the result is worth the effort as it helps to avoid the potentially serious consequences of a diagnostic error that would have arisen if left to the mercy of intuition and subjective judgment of clinicians [5]. Through its omnidirectional inference system and inherent ability to explicitly model uncertainty, KEBN is suitable for direct representations of many real-world problems, such as reasoning about depression [5].

III. METHODOLOGY

To achieve the objective of the study, which is to develop and describe graphical KEBN models for reasoning in depression, the following steps were taken:

a. Collection of depression data from the mental health unit of University of Benin Teaching Hospital (UBTH) and primary care centres in Nigeria.

b. Data Preparation by discretisation, that is, transforming data into qualitative data.

c. Presentation of dataset as an $N_i * M_j$ matrix ('i' varies from 1 to 1090 and 'j' varies from 1 to 21.

d. Extraction of significant features by reducing the number of symptoms (dimensionality reduction) using an unsupervised technique, the principal component analysis (PCA). This is similar to the way physicians extract the most significant symptoms of an illness during medical diagnosis.

e. Presentation of new dataset with the significant symptoms extracted in (d) above (N_i * M_j matrix of the principal components).

f. Generation of graphical KEBN model

g. Specification of conditional probability distribution for each node to quantity the relationship between connected nodes.

A. *Data Processing*

Dataset collected from the hospital were 1090 data instances, 23 attributes (21 features and 2 class (target) attributes. The features are: age, sex, Sad mood, suicide attempt, loss of pleasure, insomnia, hypersomnia, loss of appetite, psychomotor agitation, psychomotor retardation, loss of energy, feeling of worthlessness, lack of thinking indecisiveness, recurrent thought of death, weight gain, weight loss, stressful life events, financial pressure, depression in family, employment status, depression diagnosis, and depression comorbidity. A smaller part of this dataset was first used in a study by Ojeme and Mbogho [18], [19] to test the predictive strength of BN in detecting depression. This study achieves the same objective, but beyond merely producing predictions like the previous ones, performed dimensionality reduction with PCA and used graphical knowledge-based system models to precisely quantify the importance of individual symptoms, and identified the most efficient path towards the target nodes. Data discretisation was performed with the Waikato environment for knowledge analysis (Weka) [20] on the default settings. Next was the extraction of significant features by reducing the number of features (dimensionality reduction) using the principal component analysis (PCA), also performed in Weka. This was necessary in order to remove redundant features and outliers from the dataset, which would have degraded the classification accuracy [21]. The step-by step PCA approach is summarized as follows: 1) take the whole dataset consisting of d-dimensional samples ignoring the class label 2) compute the d-dimensional mean vector (i.e., the mean for every dimension of the whole dataset 3) compute the covariance of the whole dataset 4) compute the eigenvectors $e_1, e_2, \ldots\ldots, e_n$ and the corresponding eigenvalues $\lambda_1, \lambda_2, \ldots\ldots, \lambda_n$ 5) sort the eigenvectors by decreasing eigenvalues and choose m eigenvectors with the largest eigenvalues to form an n * m dimensional matrix (where every column represents an eigenvector) 6) Use this n * m eigenvector matrix to transform the samples onto the new dataset.

The initial input vector (eigenvector) having 21 features and their corresponding eigenvalues (marked in boldface for visualisation purpose) are shown in Table 1. Eigenvectors

Table 1. Eigenvectors and eigenvalues in PCA

	Features	**Eigen-value**		**Features**	**Eigen-value**
1	Sad mood	**1.6597**	12	Psychomotor agitation	**0.9070**
2	Loss of pleasure	**1.2290**	13	Psychomotor ret.	**0.8165**
3	Insomnia	**1.1959**	14	lack of thinking	**0.7921**
4	Worthlessness	**1.1583**	15	financial pressure	**0.7314**
5	Impaired function	**1.1470**	16	Hypersomnia	**0.6512**
6	suicide attempt.	**1.1358**	17	Thought of death	**0.6232**
7	Employment status	**1.1254**	18	Depression in family	**0.4943**
8	Indecisiveness	**1.1161**	19	Loss of appetite	**0.4692**
9	Loss of energy	**1.1094**	20	Age	**0.4512**
10	Weight loss	**0.9840**	21	Sex	**0.4091**
11	Weight gain	**0.9772**			

have been sorted by decreasing eigenvalues and those with the largest eigenvalues are the principal components (PC) [21]. In line with the concept of PCA, the PC for this study are the features with eigenvalues greater or equal to 1 (That is, features number 1 to 9 in Table 1).

IV. EXPERIMENTAL RESULTS AND ANALYSIS

Although KEBN provides an intuitive medium for knowledge acquisition and inference in many systems, graphical tools are required for the creation and manipulation of any nontrivial network structure. A number of software packages exist for KEBN modeling and inference. These include BayesiaLab [5], (Weka) [20], Multi-label extension to Weka (Meka) [22], Analytical, Bayes Net Toolbox, GeNIe, Hugin, JavaBayes, MSBNx, and Netical [6]. Built on the foundation of the BN formalism with a sophisticated GUI, BayesiaLab (Evaluation Version, 6.0.2), a powerful desktop application (Windows/Mac/Unix), was used in this study to automatically generate structural models from depression data and describe the probabilistic relationships between variables. BayesiaLab contains: (1) a graph editor, (2) a probability editor, and (3) a numerical engine for rapid prototyping and implementation of KEBN [5].

Using a 10-fold cross-validation, the structure of the PCA-transformed depression dataset was learnt using the Tree Augmented Naïve Bayes (TAN), an improved extension of the Naïve Bayes (NB) algorithm, which relaxes the problematic assumption that all attributes are independent of each other given the class.

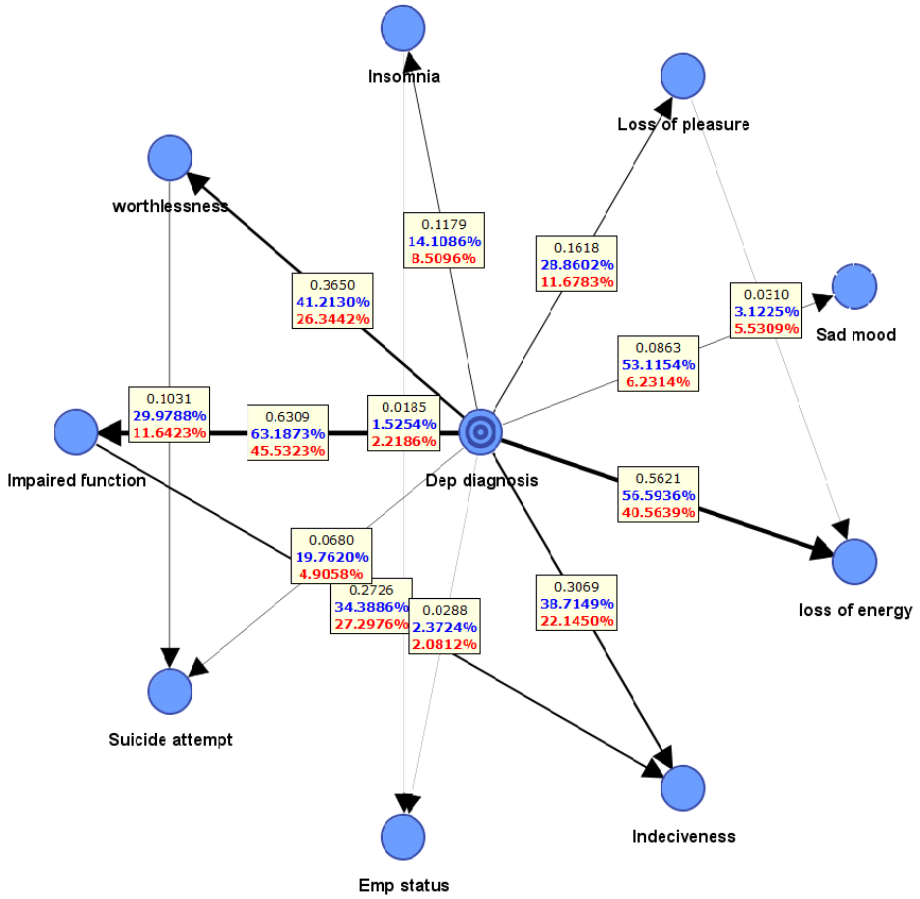

Fig. 3. Tree augmented Bayesian network graphical model of depression diagnosis

TAN, as discussed in Cheng and Greiner [23] has a high performance that is comparable to state-of-the art classifiers like decision trees despite its robust nature and computational simplicity. Its learning procedure is: 1) take the training set X = (x_1, x_2,…,x_n/c) as input (where X is the feature node and c is the classification node); 2) replace every mutual information test I(x_i, x_j) with a conditional mutual information test I(x_i, x_j /{c}); 3) add c as a parent of every x_i where $I \leq i \leq n$.; 4) learn the parameters and output the TAN. Development of the KEBN graphical model with the PCA-transformed dataset was performed using the default settings of BayesiaLab. The experimental result is shown in Fig. 3 and analysed in Table 2.

The top numbers in the yellow boxes (Fig. 3) show the mutual information (MI) value. Mutual information reports how much one variable, say a child, tells about another variable, say a parent [6]. The feature, impaired function, which has the highest MI of 0.6309, is the most important predictive variable with regard to depression diagnosis while the least important predictive feature with regard to depression diagnosis is employment status with MI of 0.0288. The middle blue numbers show the relative mutual information with regard to the child nodes while the bottom red numbers show the relative mutual information with regard to the parent nodes. The results confirm a good performance of this model in terms of the considered metrics but given that a classifier has its own inductive bias, we found it helpful testing out a variety of other supervised learning classifiers and selecting the best model.

Table 3 presents a summary of the results from TAN and other variants of BN classifiers including Naïve Bayes (NB), Augmented Naïve Bayes (ANB), Tree Augmented Naïve Bayes (TAN), Markov Blanket (MB), Augmented Markov Blanket (AMB) and the Minimal Augmented Markov Blanket (MAMB) [24]. Though the results from the table show all the classifiers as reasonably good (none had less than 88% in terms of ROC and Precision, TAN had the best overall performance while NB had the least performance in the experiment.

V. EVALUATION

In order to ensure that nothing is neglected and that results from our experiments are consistent with expectations, we evaluate the performance of our models using a 10-fold cross-validation and the following six model evaluation metrics discussed in Chai and Draxler [25]: 1) The root-mean-square error (RMSE) measures the differences between values predicted by a model and the values actually observed.

Table 2: Relationship analysis with target node

Parent/ target	child	Mutual information	Effect on target	Overall contribution (%)
Depression diagnosis	Loss of energy	0.5621	0.9654	30.9540
Depression diagnosis	Impaired function	0.6309	0.9590	22.7671
Depression diagnosis	worthlessness	0.3650	0.8777	13.1727
Depression diagnosis	Loss of pleasure	0.1618	0.9495	5.8394
Depression diagnosis	insomnia	0.1179	0.5978	4.2550
Depression diagnosis	Sad mood	0.0863	1.9125	3.1158
Depression diagnosis	indecisiveness	0.3069	0.8466	1.9826
Depression diagnosis	Employment status	0.0288	-1147	1.9751
Depression diagnosis	Suicide attempt	0.0680	0.6871	0.2655
Loss of pleasure	Loss of energy	0.0310		11.7903
insomnia	Employment status	0.0185		1.6035
worthlessness	Suicide attempt	0.1031		1.5337
Impaired function	indecisiveness	0.2726		0.7452

Mathematically,

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i} - X_{model,i})^2}{n}} \quad (1)$$

where X_{obs} is observed values and X_{model} is modelled values at time/place i and n is the number of observations. The values of RMSE range from 0 to infinity with 0 indicating a perfect model performance. 2) The normalized root mean square error (NRMSE) is the non-dimensional form of the RMSE (to the range of the observed data) used for comparing RMSE with different units. NRMSE is expressed as a percentage, where lower values indicate less residual variance and better performance. Mathematically,

$$NRMSE = \frac{RMSE}{X_{obs,max} - X_{obs,min}} \quad (2)$$

3) The Pearson correlation coefficient (r) shows the strength and direction of a linear relationship between two variables, X (model output) Y (observed values). It is obtained by dividing the covariance of the two variables by the product of their standard deviations, given a value between -1 and +1. A correlation coefficient of +1 shows a total correlation, 0 is no correlation and -1 is a total negative correlation. Mathematically,

$$R = \frac{\sum_{i=1}^{n}(x_i - \bar{x}).(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2.\sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad (3)$$

where x_i = actual number; y_i= predicted number; $\bar{x}$ and $\bar{y}$ are average numbers for actual and predicted, respectively. 4) The square of the Pearson correlation coefficient (R^2), measures the power of correlation between predicted and actual number of faults. Like R, this metric's value should be near to 1 if the model is to be acceptable. 5) Receiver operating characteristics (ROC) provides the area under the curve (AUC) of the plot of the true positive rate (y-axis) against the false positive rate (x-axis). An excellent classifier will have ROC area values between 0.9 and 1.0 (90 and 100%) while a poor classifier will have ROC area values between 0.6 and 0.7 (60 and 70%) [26]. 6) A precision of 91 to 94% was achieved by all the classifiers indicating correct predictions among the positive predictions. Precision = TP/(TP + FP) where TP and FP refer to true positive and false positive, respectively.

Table 3: Results from performance of variants of BN classifiers

	TAN	NB	ANB	MB	AMB	MAMB
RMSE	0.2148	0.3152	0.2141	0.3009	0.2262	0.2278
NRMSE (%)	7.1603	10.5078	7.1358	10.0311	7.5412	7.5942
R	0.9532	0.8967	0.9536	0.9064	0.9481	0.9472
R^2	0.9087	0.8041	0.9093	0.8216	0.8988	0.8973
ROC (%)	89.9676	88.7414	89.9047	89.5604	89.9676	89.9047
Precision (%)	94.8624	91.9266	94.9541	93.2110	94.4954	94.4954

VI. RELATED WORK

Studies show several attempts by KEBN researchers to design diagnostic tools for medical conditions. We highlight some of these studies that are closely related to the use of graphical models for reasoning in mental health.

For the identification of factors that affect diseases and their correlation, Curiac et al [27] developed a BN model for the analysis of psychiatric data from a Romanian specialised clinic. The study found that the probability of patients diagnosed with specified psychiatric diseases fluctuates for mixed dementia paranoid schizophrenia but drops by about 50% for simple schizophrenia. Seixas et al [28] proposed a BN decision model for assisting clinicians in the detection of dementia diseases (Alzheimer's Disease (AD) and Mild Cognitive Impairment (MCI)) using a combination of data and knowledge from clinicians while the network parameters were estimated using a supervised learning algorithm from a dataset of real clinical cases.

VII. CONCLUSIONS AND FUTURE WORK

Today's data-driven world requires researchers and clinicians to be able to explore and visualize data very quickly for informed decision-making. In this paper, we have demonstrated the strength of graphical KEBN in making reasoning in depression, not only appealing but convenient, by bringing out hidden structures in data. The model computed the predictive importance of various symptoms with regard to depression and showed that methods based on CI tests, such as mutual information, are suitable for BN classifier learning. The study demonstrates that more tools can be developed to help improve research data visualisation. However, the study can be extended and improved in several ways. We can use a scoring-based test that searches through possible network structure for a best scored network for the purpose of finding a graph and parameters that maximize the likelihood. Secondly, in increasing the role of computationally managed knowledge in healthcare, we will look at social, economic, legal and ethical issues that tend to create a wide gap between research and healthcare [29]. Lastly, we will investigate the performance of these methods in other contexts.

REFERENCES

[1] E. Feigenbaum, "Expert Systems: Principles and Practice," *The Encyclopedia of Computer Science and Engneering. Harlow, England Addison-Wesley*, vol. 581. pp. 1–12, 1992.

[2] S-H Liao, "Expert System Methodologies and Applications—A Decade Review from 1995 to 2004," *Expert Syst. Appl.*, vol. 28, no. 1, pp. 93–103, 2005.

[3] C. Smith, B. McGuire, T. Huang, and G. Yang, "The History of Artificial Intelligence," *University of Washington*. p. 27, 2006.

[4] M. Negnevitsky, *Artificial intelligence: A Guide to Intelligent Systems*, Second edi. Addison Wesley: Pearson Education, 2005.

[5] S. Conrady and L. Jouffe, *Bayesian Networks & BayesiaLab: A Practical Introduction for Researchers*, vol. 1, no. 1. Bayesia USA, 2015.

[6] K. Korb and A. Nicholson, *Bayesian Artificial Intelligence*, Second edi., vol. 2, no. 1. Chapman & Hall, 2004.

[7] B. Buchanan and E. Shortliffe, *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Addison-Wesley, Reading, MA, 1984.

[8] G. Luger and C. Chakrabarti, "Knowledge-Based Probabilistic Reasoning from Expert Systems to Graphical Models," *Department of Computer Science University of New Mexico https://www.cs.unm.edu/~treport/tr/09-11/luger.pdf*. 2008.

[9] A. Onisko, P. Lucas, and M. J. Druzdzel, "Comparison of Rule-Based and Bayesian Network Approaches in Medical Diagnostic Systems," in *S. Quaglini, P. Barahona, and S. Andreassen (Eds.): Proceedings of the 8th Conference on Artificial Intelligence in Medicine in Europe, AIME 2001, Springer,-Verlag, Berlin, 2001*, 2001, pp. 283–292.

[10] V. K. Sharma and J. R. Copeland, "Detecting Mental Disorders in Primary Care," *Ment. Health Fam. Med.*, vol. 6, no. 1, pp. 11–13, 2009.

[11] A. J. Ferrari, F. J. Charlson, R. E. Norman, S. B. Patten, G. Freedman, C. J. L. Murray, T. Vos, and H. Whiteford, "Burden of Depressive Disorders by Country, Sex, Age, and Year: Findings from the Global Burden of Disease Study 2010," *PLoS Med.*, vol. 10, no. 11, pp. 1–12, 2013.

[12] F. Amato, A. López, E. M. Peña-Méndez, P. Vaňhara, A. Hampl, and J. Havel, "Artificial Neural Networks in Medical Diagnosis," *J. Appl. Biomed.*, vol. 11, no. 2, pp. 47–58, 2013.

[13] M. U. Ahmed, S. Begum, and P. Funk, "Case Studies on the Clinical Applications Using Case-Based Reasoning," in *proceedings of the Federated Conference on Computer Science and Information Systems*, 2012, pp. 3–10.

[14] G. Doherty, D. Coyle, and M. Matthews, "Design and Evaluation Guidelines for Mental Health Technologies," *Interact. Comput.*, vol. 22, no. 4, pp. 243–252, 2010.

[15] I. Ben-Gal, "Bayesian Networks," *In Ruggeri F., Faltin F. & Kenett R: Encyclopedia of Statistics in Quality & Reliability*, vol. 1, no. 1. Wiley & Sons, p. 6, 2007.

[16] D. Nikovski, "Constructing Bayesian Networks for Medical Diagnosis from Incomplete and Partially Correct Statistics," *IEEE Trans. Knowl. Data Eng.*, vol. 12, no. 4, pp. 1–18, 2000.

[17] Alpaydın Ethem, *Introduction to Machine Learning*, Second Edi. The MIT Press Cambridge, Massachusetts London, England, 2010.

[18] B. Ojeme and A. Mbogho, "Predictive Strength of Bayesian Networks for Diagnosis of Depressive Disorders," in *Proceedings of the 8th KES International Conference on Intelligent Decision Technologies (KES-IDT 2016)-Tenerife, Spain, June 15-17*, 2016, pp. 373–382.

[19] B. Ojeme and A. Mbogho, "Selecting Learning Algorithms for Simultaneous Identification of Depression and Comorbid Disorders," in *Proceedings of the 20th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES-2016). York, UK, September 5-7, Vol. 96*, 2016, pp. 1294–1303.

[20] R. R. Bouckaert, E. Frank, M. Hall, R. Kirkby, P. Reutemann, A. Seewald, and D. Scuse, *WEKA Manual for Version 3-7-12*. University of Waikato, Hamilton, New Zealand, 2014.

[21] H. Jeong, C. Ziemkiewicz, W. Ribarsky, and R. Chang, "Understanding Principal Component Analysis Using a Visual Analytics Tool. Charl," *Charlotte Visualization Center, UNC Charlotte*. pp. 1–10, 2009.

[22] J. Read, P. Reutemann, B. Pfahringer, and G. Holmes, "MEKA: A Multi-label / Multi-target Extension to WEKA," *J. Mach. Learn. Reasearch. Available http//jmlr.org/papers/volume17/12-164/12-164.pdf*, vol. 17, no. 21, pp. 1–5, 2016.

[23] J. Cheng and R. Greiner, "Comparing Bayesian Network Classifiers," in *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence*, 1999, pp. 101–108.

[24] R. Bouckaert, "Bayesian Network Classifiers in Weka for Version 3-5-7," Department pf Computer Science, University of Waikato, New Zealand, 2008.

[25] T. Chai and R. R. Draxler, "Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)? -Arguments Against Avoiding RMSE in the Literature," *Geosci. Model Dev.*, vol. 7, no. 3, pp. 1247–1250, 2014.

[26] T. Saito and M. Rehmsmeier, "The Precision-Recall Plot is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets.," *PLoS One*, vol. 10, no. 3, p. e0118432, 2015.

[27] D.-I. Curiac, G. Vasile, O. Banias, C. Volosencu, and A. Albu, "Bayesian Network Model for Diagnosis of Psychiatric Diseases," in *Proceedings of the ITI 2009 31st International Conference on Information Technology Interfaces*, 2009, pp. 61–66.

[28] F. L. Seixas, B. Zadrozny, J. Laks, A. Conci, and D. C. Muchaluat Saade, "A Bayesian Network Decision Model for Supporting the Diagnosis of Dementia, Alzheimer's Disease and Mild Cognitive Impairment.," *Comput. Biol. Med.*, vol. 51, pp. 140–58, 2014.

[29] T. Beck, S. Gollapudi, S. Brunak, N. Graf, H. U. Lemke, D. Dash, I. Buchan, C. D??iaz, F. Sanz, and A. J. Brookes, "Knowledge Engineering for Health: A New Discipline Required to Bridge the 'ICT Gap' Between Research and Healthcare," *Hum. Mutat.*, vol. 33, no. 5, pp. 797–802, 2012.