Predicting Depression Treatment Using Bayesian Networks Based on Mental Health Factors

Altaaf Ally
2424551
Probabilistic Graphical Modelling
COMS4062A

I. INTRODUCTION

Depression is a prevalent mental health disorder that affects millions of people worldwide. Early detection and appropriate treatment are crucial for managing depression and improving the quality of life for affected individuals. In recent years, machine learning techniques, particularly Probabilistic Graphical Models (PGMs), have shown promise in predicting depression treatment outcomes based on various mental health factors. This study aims to develop a Bayesian network model to predict whether a person needs treatment for depression using a dataset of relevant mental health factors.

Numerous factors can impact the treatment outcome for individuals with depression, including health issues, sleep disturbances, traumatic brain injury, stress, chronic illness, substance abuse, and others. A key challenge lies in evaluating these factors to determine whether an individual requires treatment for depression. As discussed above, predicting whether a person requires depression treatment presents a significant challenge. Regression and simple statistical models are commonly used approaches. However, these models may not fully account for the complex interdependencies between variables. To address this predictive challenge, this study will explore the use of Bayesian Networks.

Mathematically, Bayesian Networks represent a joint distribution over a set of variables by factoring it into a product of conditional probabilities. They capture conditional dependencies between variables, allowing for a nuanced understanding of the interplay of factors affecting treatment outcomes. This probabilistic graphical model provides a robust framework for predicting the need for depression treatment.

The significance of this problem lies in its potential to assist mental health professionals in making informed decisions about treatment recommendations. By accurately predicting the need for depression treatment, the model can help prioritize resources and ensure timely interventions for individuals who are most likely to benefit. Moreover, this study contributes to the growing body of research on the application of PGMs in mental health, showcasing their effectiveness in modeling complex relationships between variables.

II. DATA HANDLING AND ETHICAL CONSIDERATIONS

This section details the data collection and preprocessing and ethical considerations.

A. Data Handling and Preprocessing

The dataset used in this study was obtained from Kaggle and created by Jikadara (n.d.) [Data set]. It contains sensitive information related to individuals' mental health. To ensure dataset balance, the model was split into a training set of 60,000 samples, with validation and test datasets each containing 20,000 samples. The model is trained on a Bayesian Network and evaluated using Accuracy, Precision, Recall, F1-score, and ROC AUC. The following variables were captured:

TABLE I
DESCRIPTION OF VARIABLES

Variable	Description
Timestamp	Time of data entry
Gender	Gender of the individual
Country	Country of residence
Occupation	Occupation of the individual
self_employed	Whether the individual is self-employed
family_history	History of mental illness in the family
Days_Indoors	Number of days spent indoors
Growing_Stress	Level of growing stress
Changes_Habits	Changes in habits due to mental health
Mental_Health_History	Personal history of mental health issues
Mood_Swings	Presence of mood swings
Coping_Struggles	Struggles in coping with mental health
Work_Interest	Interest in work
Social_Weakness	Weakness in social interactions
mental_health_interview	Experience of mental health interviews
care_options	Available care options

B. Data Collection and Preprocessing

To prepare the data for analysis, several preprocessing steps were performed:

- Missing values in categorical variables were handled by filling them with the mode of the respective column,

ensuring dataset completeness while preserving the most frequent category.

- Categorical variables were encoded using label encoding to convert them into numerical representations suitable for the Bayesian network model.
- Randomized the data, since the data was ordered by gender, after encoding to prevent the model from being trained exclusively on one gender, ensuring a balanced representation.

C. Ethical Consideration

Throughout the data handling process, I prioritized respectful use of the dataset, acknowledging both its sensitive nature and the anonymity of the 300,000 individuals who contributed their information. I acknowledge the dataset creator and limit its use strictly to this specific research study. The results and findings of this study are presented in an aggregated manner, ensuring that no individual can be identified based on the reported information.

III. PGM SELECTION AND APPLICATION

A. Choice of Bayesian Network Model

For this study, a Bayesian network model was chosen to predict whether a person needs treatment for depression based on mental health factors. Bayesian networks are particularly suitable for this task due to their ability to represent complex relationships between variables and incorporate domain knowledge in the form of prior probabilities. In the context of mental health, Bayesian networks can capture the intricate dependencies between various factors and provide probabilistic predictions for depression treatment.

B. Bayesian Network Representation

The Bayesian network in this study is represented as a Directed Acyclic Graph (DAG), where nodes represent mental health factors and edges represent conditional dependencies between them. This graphical representation allows for a clear visualization of the relationships between variables and facilitates the understanding of the model structure as seen in Figure 1

1) (treatment, Gender)Treatment is linked to gender due to variations in treatment response and side effects between males and females.
2) (Social_Weakness, Gender)Social weakness is linked to gender due to societal expectations and gender roles affecting social interaction and support.
3) (Occupation, Gender)Occupation is linked to gender due to gender-based occupational segregation and disparities in workplace conditions and opportunities.
4) (Coping_Struggles, Gender)Coping struggles are linked to gender due to differences in coping mechanisms and societal pressures experienced by different genders.
5) (Mood_Swings, Occupation)Mood swings are linked to occupation due to the impact of work-related stressors and job satisfaction on mental health.

6) (Mental_Health_History, Occupation)Mental health history is linked to occupation due to occupational stressors and the influence of work environment on mental well-being.
7) (Days_Indoors, Occupation)Days indoors are linked to occupation due to variations in work schedules and indoor exposure to environmental factors affecting health.
8) (Changes_Habits, Occupation)Changes in habits are linked to occupation due to the influence of work-related stress and lifestyle factors on health behaviors.
9) (Growing_Stress, Occupation)Growing stress is linked to occupation due to occupational demands and pressures affecting stress levels and mental health.
10) (self_employed, care_options)Being self-employed is linked to care options due to differences in access to healthcare benefits and support services.
11) (Country, self_employed)Country is linked to being self-employed due to variations in entrepreneurship rates and regulatory environments across countries.
12) (family_history, self_employed)Family history is linked to being self-employed due to familial influences on career choices and entrepreneurial tendencies.
13) (mental_health_interview, self_employed)Mental health interview is linked to being self-employed due to mental health stigma and disclosure concerns in entrepreneurial settings.
14) (Timestamp, family_history)Timestamp is linked to family history due to temporal associations and changes in health status over time within families.
15) (Work_Interest, Coping_Struggles)Work interest is linked to coping struggles due to the impact of job satisfaction and engagement on stress management.
16) (care_options, None)Care options are not linked to any other variable in the dataset.
17) (self_employed, family_history)Being self-employed is linked to family history due to differences in work-related stressors and lifestyle factors affecting health outcomes.
18) (self_employed, mental_health_interview)Being self-employed is linked to mental health interview due to the potential impact of entrepreneurial stressors on mental well-being and help-seeking behavior.
19) (self_employed, care_options)Being self-employed is linked to care options due to differences in access to healthcare benefits and support services.
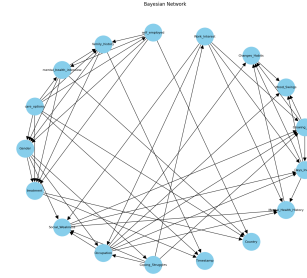

Fig. 1. Bayesian Network for Health Factors

C. Structure Learning

To learn the structure of the Bayesian network from the data, the Hill Climb Search algorithm from the pgmpy library was employed. The Hill Climb Search is a heuristic search algorithm that navigates the search space to find the optimal structure that best fits the data. It is particularly useful for large datasets like the one used in this study, as it efficiently explores the space of possible network structures.

D. Model Evaluation

To evaluate the goodness of fit of the learned Bayesian network structure, the Bayesian Information Criterion (BIC) was used in conjunction with structure learning. BIC assesses how well the Bayesian network fits the mental health data by considering the likelihood of the data given the model while penalizing model complexity. A higher BIC score indicates a better fit of the model to the data, striking a balance between model complexity and explanatory power.

E. Parameter Estimation

Once the Bayesian network structure was learned, the model parameters were estimated using the Maximum Likelihood Estimator (MLE). MLE determines the most likely parameter values given the observed data, maximizing the likelihood function. In this study, MLE was used to estimate the conditional probability distributions for each node in the network, capturing the probabilistic relationships between mental health factors.

F. Model Complexity Management

To manage model complexity, the number of variables and edges in the Bayesian network was carefully considered. The selection of relevant mental health factors was based on domain knowledge and statistical significance. The Hill Climb Search algorithm with a maximum in-degree constraint was used to control the complexity of the network structure, preventing overfitting and ensuring interpretability.

G. Experiment Design

The experiment design involved splitting the dataset into training, validation, and test sets. The training set was used to learn the Bayesian network structure and estimate the parameters. The validation set was employed to tune hyperparameters and assess the model's performance during development. Finally, the test set was used to evaluate the model's predictive accuracy on unseen data.

H. Model Evaluation Metrics

To evaluate the trained Bayesian network model, several metrics were utilized, including accuracy, precision, recall, F1-score, and ROC AUC. These metrics provide a comprehensive assessment of the model's performance in predicting depression treatment. The model's predictions were compared against the actual treatment labels, and the evaluation metrics were calculated both with and without evidence to examine the impact of additional information on prediction accuracy.

I. Incorporation of Evidence

One of the key strengths of Bayesian networks is their ability to incorporate evidence and update the probabilities of the target variable accordingly. In this study, evidence was incorporated by setting the observed values of certain mental health factors and then inferring the probability of needing depression treatment given that evidence. This allows for a more personalized and context-specific prediction, as the model takes into account the individual's specific mental health profile. By comparing the model's performance with and without evidence, we can assess the impact of incorporating additional information on the accuracy and reliability of the predictions.

IV. INFERENCE AND RESULTS ANALYSIS

A. Inference using the Bayesian Network

Inference in the Bayesian network model was performed using the Variable Elimination algorithm. This exact inference technique efficiently computes the posterior probabilities of the target variable (the need for depression treatment) given observed evidence. Evidence includes demographic information and specific symptoms related to depression.

The Variable Elimination algorithm exploits the conditional independence assumptions in the Bayesian network's structure to streamline computations. It works by eliminating variables one-by-one, summing out their contributions to arrive at the desired marginal probability distribution. This process creates intermediate probability distributions ("factors") over subsets of variables, which are combined and marginalized until only the target variable remains.

The resulting posterior probability offers a personalized prediction of whether treatment is needed, aiding mental health professionals in their decision-making. Variable Elimination is exact and computationally efficient for smaller Bayesian networks. However, large and complex networks may necessitate approximate inference techniques (like sampling) to balance accuracy and computational cost

B. Evaluation Metrics

1) Accuracy

Accuracy was used to assess the overall performance of the Bayesian network model in correctly predicting whether a person needs treatment for depression based on their mental health factors. The model's predictions were compared against the actual treatment labels in the test set, and the accuracy was calculated using the formula:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (1)$$

where TP represents the number of correctly predicted positive instances (individuals who need treatment and were correctly

identified), TN represents the number of correctly predicted negative instances (individuals who do not need treatment and were correctly identified), FP represents the number of incorrectly predicted positive instances, and FN represents the number of incorrectly predicted negative instances.

2) Precision

Precision was used to evaluate the model's ability to correctly identify individuals who need treatment for depression among all the positive predictions made by the model. It was calculated using the formula:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

Precision is particularly relevant in this context because falsely identifying individuals as needing treatment (false positives) can lead to unnecessary interventions and resource allocation.

3) Recall

Recall was used to assess the model's ability to identify all the individuals who actually need treatment for depression among the total number of individuals who need treatment. It was calculated using the formula:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

Recall is important in this study because failing to identify individuals who need treatment (false negatives) can have serious consequences, as they may not receive the necessary support and interventions.

4) F1 Score

The F1 score was used to provide a balanced measure of the model's performance, considering both precision and recall. It was calculated using the formula:

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

The F1 score is particularly useful in this study because it provides a single metric that balances the trade-off between precision and recall, ensuring that the model is effective in identifying individuals who need treatment while minimizing false positives and false negatives.

5) ROC AUC

The ROC AUC was used to evaluate the Bayesian network model's ability to discriminate between individuals who need treatment for depression and those who do not, considering various classification thresholds. The ROC curve was plotted by calculating the true positive rate (recall) and the false positive rate at different thresholds, and the area under the curve was computed to obtain the ROC AUC score. The ROC AUC provides a comprehensive measure of the model's predictive power, with higher values indicating better discrimination between the positive and negative instances. It helps assess the model's performance across different operating points and

is particularly useful when the decision threshold needs to be adjusted based on the specific requirements of the application.

6) Comparison of Prediction Methods

The comparison of prediction methods, as shown in Figure 2, demonstrates the impact of incorporating evidence on the accuracy of the Bayesian network model. When evidence is taken into account, the model achieves significantly higher accuracy compared to making predictions without evidence. For the validation set, the accuracy with evidence is 0.765,
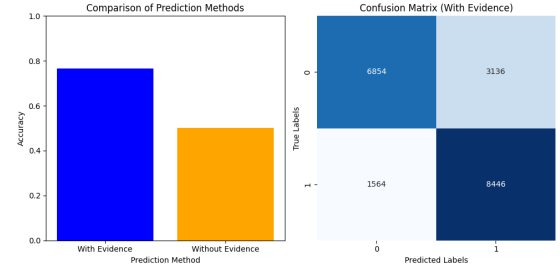

Fig. 2. Comparison of Prediction Methods

while the accuracy without evidence is 0.5005. This indicates that by incorporating relevant mental health factors as evidence, the model's ability to correctly predict the need for depression treatment improves substantially. The number of correct predictions with evidence (15,300) is notably higher than the number of correct predictions without evidence (10,010). Similarly, for the test set, as shown in Figure 3, the accuracy with evidence is 0.74, surpassing the accuracy without evidence of 0.50125. The model correctly predicts the need for depression treatment for 14,800 instances when evidence is considered, compared to 10,025 correct predictions without evidence. These results highlight the importance of
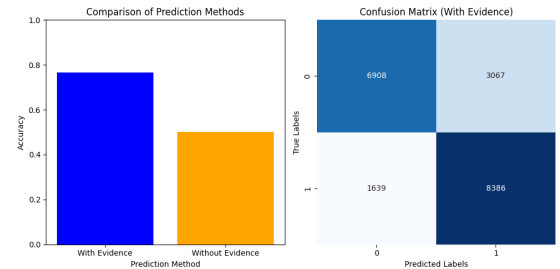

Fig. 3. Comparison of Prediction Methods (Test Set)

leveraging available information and incorporating it as evidence in the Bayesian network model. By doing so, the model can make more accurate predictions and provide better support for clinical decision-making. Table II summarizes the test and validation sets.

7) Evaluation Metrics Comparison

The comparison of evaluation metrics, as depicted in Figure 4 and Figure 5, provides a comprehensive assessment of the model's performance using precision, recall, F1-score, and

Dataset	Evidence	Accuracy
Validation Set	With Evidence	0.765
	Without Evidence	0.5005
Test Set	With Evidence	0.74
	Without Evidence	0.5012

TABLE II
SUMMARY OF RESULTS FOR TESTING AND VALIDATION SETS

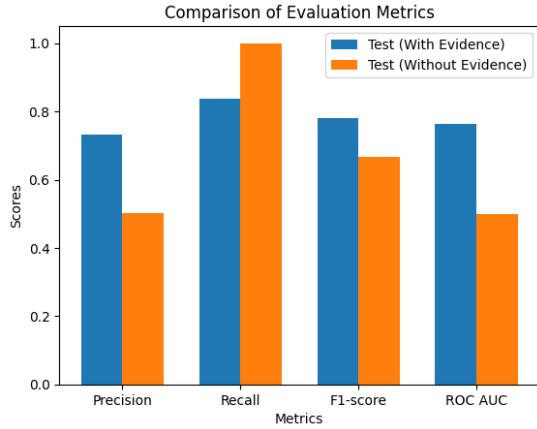ROC AUC. For the test set (Figure 4), the model achieves

Fig. 4. Comparison of Evaluation Metrics (Test Set)

higher scores across all metrics when evidence is incorporated. The precision, recall, and F1-score are notably higher with evidence compared to without evidence. This indicates that the model is more effective in correctly identifying individuals who need treatment (precision) and capturing a larger proportion of those who actually require treatment (recall) when evidence is considered. The higher F1-score suggests a better balance between precision and recall. Similarly, for the
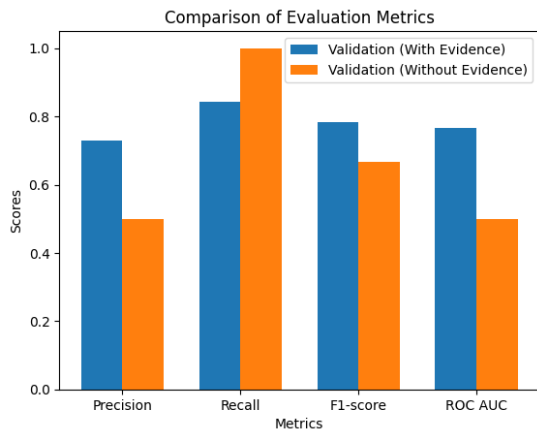

Fig. 5. Comparison of Evaluation Metrics (Validation Set)

validation set (Figure 5), the model demonstrates improved performance across all metrics when evidence is taken into account. The precision, recall, and F1-score are consistently higher with evidence compared to without evidence. This fur-

ther reinforces the model's ability to make accurate predictions and effectively identify individuals who need depression treatment when relevant mental health factors are considered. The ROC AUC scores for both the test and validation sets are also higher when evidence is incorporated. This indicates that the model has a better ability to discriminate between individuals who need treatment and those who do not when evidence is taken into account. Overall, the evaluation metrics provide strong support for the effectiveness of the Bayesian network model in predicting the need for depression treatment when relevant mental health factors are considered as evidence.

V. CONCLUSION

The Bayesian network model's ability to handle missing data, incorporate prior knowledge, and update predictions based on evidence makes it a powerful tool for predicting depression treatment based on mental health factors. The model can capture complex relationships between variables and provide probabilistic predictions to assist mental health professionals in making informed treatment decisions.

However, it is important to acknowledge the limitations of this approach. The current model does not directly account for certain life events or external circumstances that may significantly impact an individual's mental health. Moreover, mental health conditions often involve complex feedback loops and self-reinforcing dynamics, which the current model does not explicitly account for, potentially limiting its ability to capture the full complexity of an individual's mental health journey.

Despite these limitations, the current study demonstrates the potential of Bayesian networks in predicting the need for depression treatment. The model's predictions can serve as a valuable supportive aid for mental health professionals, but should be considered in conjunction with clinical expertise and individual patient assessments.

In conclusion, while the Bayesian network model has shown promising results, further research is needed to address the complex dynamics of mental health. By continually refining and expanding the model, incorporating a wider range of factors, and exploring advanced modeling techniques, we can work towards developing more comprehensive and accurate tools to support the identification and treatment of individuals with depression.

REFERENCES

Jikadara, B. (n.d.). Mental Health Dataset [Data set]. Kaggle. https://www.kaggle.com/datasets/bhavikjikadara/mental-health-dataset