The International Journal of Indian Psychology
ISSN 2348-5396 (e) | ISSN: 2349-3429 (p)
Volume 7, Issue 3, DIP: 18.01.031/20190703
DOI: 10.25215/0703.031
http://www.ijip.in | July - September, 2019

Research Paper

Modeling the Detection of Internet Addiction Disorder Using Bayesian Networks

Anju Singh[1]*, Sakshi Babbar[2]

ABSTRACT

Today's life revolves around online activities and has made internet a necessity. The unregulated excessive use of internet interferes in daily life and may lead to addiction to it. In literature, Internet Addiction Disorder (IAD) is defined as compulsive online behaviour which dominates and becomes organizing principle of one's life impairing physical, emotional and social well-being. The common conventional assessment tools available to detect IAD are Internet Addiction Test (IAT), Compulsive Internet Use Scale (CIUS), Problematic Internet Use Questionnaire (PIUQ). Recently researchers are adopting Machine Learning (ML) to automate the detection of IAD. This study proposes to develop graphical interface IAD models by harnessing the significant features of Bayesian Networks (BN), a powerful ML technique. The models are constructed with real data sets collected online through IAT, CIUS and PIUQ questionnaires as domain knowledge and we propose these models as IAT-BN, CIUS-BN and PIUQ-BN respectively. The graphical presentations of models provide an effective visualization of causes and symptoms of IAD and facilitate in the better explanation of its occurrence. The promising experimental results show superior performance of the IAT-BN model with 100% accuracy followed by CIUS-BN as 95% and PIUQ-BN model with 90% accuracy. The study also highlights the important factors which when controlled may substantially reduce the risk of IAD. The models can be an efficient psychiatric decision support system in predicting the occurrence of IAD, risk assessment and management.

Keywords: *Internet Addiction Disorder, Bayesian Networks, Predictive Model, Sensitivity.*

Internet has made life easier and convenient but the world has started to see the problems with over-dependence on the internet. The internet penetration is on increase and as per the recent report of January 2019 by Statista presented in Figure 1, almost 4.4 billion people worldwide are active internet users (Statista, 2019).

[1]PhD Scholar, G D Goenka University, Gurugram, Haryana, India
[2]Associate Professor, G D Goenka University, Gurugram, Haryana, India
*Responding Author

Received: June 28, 2019; Revision Received: August 16, 2019; Accepted: September 25, 2019

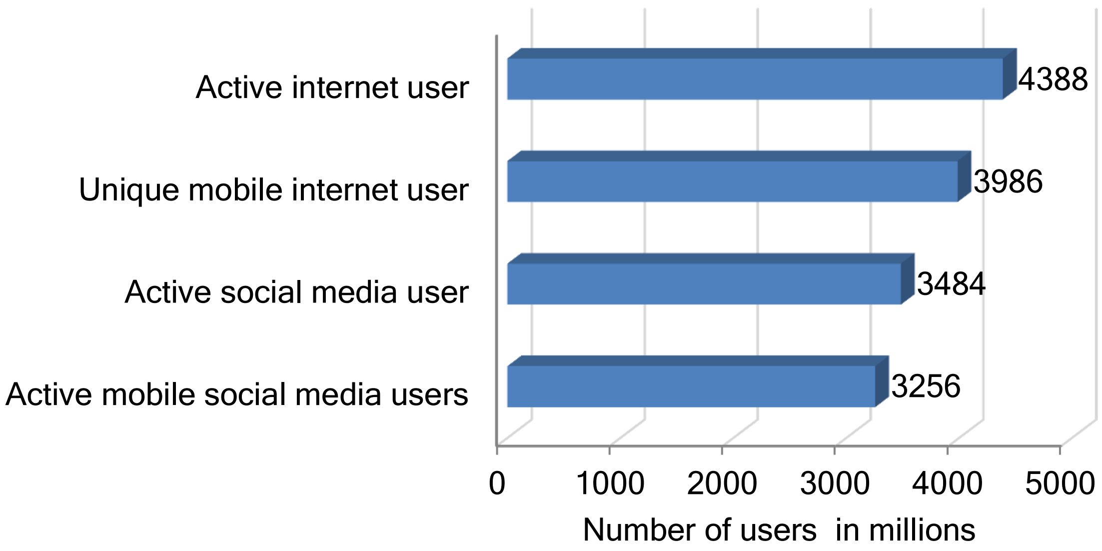

Figure 1 Global digital population as of Jan, 2019 (in millions) from Statista portal*. Active internet users are almost 4.4 billion which makes 58 percent of the global population. *Source: We Are Social; Data Reportal; Hootsuite.

The pervasiveness of internet in the life of people has escalated the risk of addiction to it and unhealthy use of internet may be detrimental to physical, social as well as mental health of the internet users. Recently, researchers are studying various types of mental disorders existing in society, to upkeep the mental health of the people (Auerbach et al., 2018, Musser et al., 2016). This study aims to improve the health of online society by studying the key parameters that contribute to IAD. American Psychiatric Association characterizes IAD as, lack of control of individuals over internet, resulting in deterioration of social, occupational and academic performance. Internet addiction is an umbrella term encompassing following areas of internet usage (Gregory, 2017): Net Compulsions: too much time spent online; Cyber Relationship Addiction: Excessive use of social networking sites to build up virtual friends and relationships; Cybersex Addiction: Too much engaged in online sex and porn sites; Information Overload: Always engrossed in seeking information and searching net.

The popular tools available for diagnosis of IAD are IAT (IAT Manual, 2015), CIUS (Meerkerk et al, 2009), and PIUQ (Demetrovics et al, 2 008) questionnaires, which have been validated by researchers in various studies (Chaudhari, 2015; Demetrovics , 2016; kaltiala-heinoa , 2004; Shao, 2018). Due to the rise of awareness regarding the health hazards associated with IAD, recently researchers are attracted towards studying the underlying factors associated with it. In order to gain actionable insight into the domain knowledge researchers are exploring ML techniques to detect the prevalence of IAD. In few recent studies, researchers used ML techniques like Naive Bayes, Random forest, K-nearest neighbour, Support Vector Machine (SVM), decision tree etc. (Ioannidis K et al., 2016; Akhter, 2017; Chaudhury & Tripathy, 2018).

This study focuses on the use of Bayesian Networks in assisting internet addiction disorder diagnosis. BN are a probabilistic presentation of the problem domain and has significant capabilities that can be used for medical predictions and diagnosis. The study uses the domain knowledge of standard assessment tools of IAD (IAT, CIUS, PIUQ) to develop three IAD models proposed to be named as IAT-BN, PIUQ-BN,CIUS-BN respectively.

Following are the key contributions of this paper:

1. The models depict the basic characteristics of IAD in a simplified graphical manner, can be used as a complementary tool to facilitate the decision-making process of detection and prediction of occurrence of IAD by the counsellors, teachers, psychiatrists and health practitioners.
2. The paper explores the power of BN, captures the key parameters of IAD and can help in propensity assessment of IAD for the internet users, thus prompting the individuals to take corrective actions before falling prey to IAD.
3. The testing of models with real cases highlighted the robustness of the models and thus we may use the models for the real-world application.
4. The sensitivity analysis of the models can interpret the sensitivity of IAD towards the key factors which plays a major role in the occurrence of IAD. Thus the paper reveals that if these key factors are controlled then the incidence of IAD can be minimized to a large extent.

The remainder of the paper is organized as follows: capabilities of ML techniques in real-world applications is presented followed by a Literature review of studies using traditional assessment tools with statistical data analysis approach and few recent studies using ML techniques. The powerful features of BN are discussed in detail. The steps in formalization and construction of proposed BN models are presented, followed by results wherein we describe the experiment setup including data collection, pre-processing of data, and implementation of models with results. Lastly, evaluation of models is illustrated using prediction and sensitivity analysis and testing of models with real cases. Finally, the discussion is summarized.

Common Abbreviations followed in the paper are illustrated in Table 1.

Table 1 Common Abbreviations followed in the Paper

ABBREVIATIONS	FULL FORM
IAD	Internet Addiction Disorder
BN	Bayesian Networks
ML	Machine Learning
IAT	Internet Addiction Test
CIUS	Compulsive Internet Use Scale
PIUQ	Problematic Internet Use Questionnaire
IAT-BN	Internet Addiction Test-Bayesian Network
CIUS-BN	Compulsive Internet Use Scale-Bayesian Network
PIUQ-BN	Problematic Internet Use Questionnaire-Bayesian Network

MACHINE LEARNING

Today's data-rich world is driven by generation of large amount of data. Researchers are developing human-friendly automated systems by harnessing the data, learning the relevant patterns from the data to make better decisions. This process of moving from data to actionable insight is accomplished by Machine Learning (Kelleher, Namee & DArey, 2015).

ML is an art of building an automated model that learns from the historical data (domain/expert knowledge) to achieve the prediction capabilities and find out interesting new patterns in data for a better understanding of the problem area. There are four types of ML methods: Supervised, Unsupervised, Semi-supervised and Reinforcement learning. In

Supervised learning method the historical data that consists of expert knowledge in the form of inputs and corresponding outputs with labels, is used to train the models and based on the patterns identified, the model predicts the future events. The supervised learning methods are implemented by using the following algorithm: decision tree, Naive Bayes, Artificial Neural Networks, BN, SVM. Unsupervised ML method works with unknown outputs i.e. unlabeled data and explores data to find the structure in the data and then find attributes that separate the data in segments. Nearest-neighbour mappings, k-means clustering are popular techniques of unsupervised learning. Semi-supervised learning uses both types of data: label data with known outputs and unlabeled data with unknown outputs and can be implemented with methods such as classification, regression, prediction. Reinforcement learning is used for gaming and navigation environments and uses trial and error methods to discover the best relevant pattern in the data (David, 2017).

In this study, we have employed Bayesian Networks which can be used both in a supervised and unsupervised environment. In this study since the objective is to detect IAD from the historical data, we chose the supervised learning method of learning in order to gain insight and describe the reasons for the occurrence of internet addiction disorder. Any Supervised learning model has a general framework of training and testing environment, consisting of different phases as represented in Figure 2. It is a technique of inferring a relationship among the factors of the domain area from the training data. The training data is a set of examples of the domain consisting of inputs and the desired output value. A supervised learning algorithm analyzes the training data, produces an inferred model that can be used for mapping new unseen instances. The model is then tested with test data to ascertain its accuracy (Mcnulty, 2015). When all these phases are combined together we get prediction capabilities from a supervised ML model.

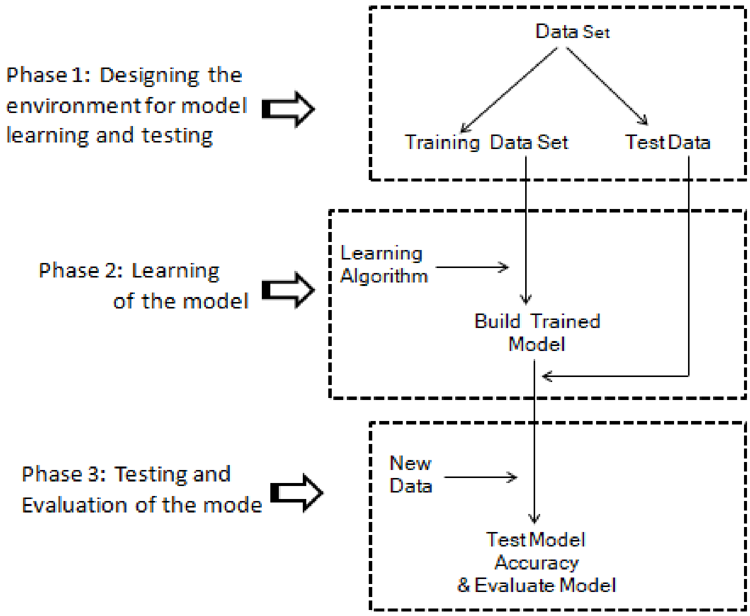

Figure 2 Phases in Supervised Learning technique of Machine Learning. These Phases are Combined together to get Prediction Capabilities from ML Models.

Limitations of traditional statistical data analysis and how ML tools are important for the development of real-world models is discussed in the following points:

1. The conventional methods need the intervention of domain experts to provide the result. Whereas ML methods integrate computer-based learning systems in the problem domain, providing opportunities to facilitate and enhance the work of experts with novel outcomes and ultimately improve problem management.
2. The ML techniques automatically detect relationships between the independent and the dependent variables and can identify interactions among the variables. Like in the case of detection of IAD the independent variables are the symptoms of IAD and dependent variable is the incident of IAD.
3. As regards to this study on IAD, traditional methods are based on theoretical subscales, involves self-report mode and have some inherent drawbacks like participants may misinterpret items in the questionnaire, gives dishonest answers and miss some items in the questionnaire or conceal their excessive usage. The diagnosis of internet addiction disorder becomes difficult as the internet users tend to become sceptical about their online activities and do not disclose their true internet usage habits. ML models offer techniques where missing values can be interpreted to match the true population. As ML works well with wide number of attributes, high number of observations and can handle missing data efficiently, it best fits our approach.

LITERATURE REVIEW

In literature, there are a number of studies on the diagnosis of IAD using conventional statistical analysis but very few research works are done using ML techniques. We present a few recent studies based on conventional methods and ML methods for diagnosis of IAD in internet users.

In a recent cross-sectional study conducted by Gedam SR et al (2017), 846 students of various faculties from Deemed University were classified into normal and internet addicted using IAT and Mental Health Inventory. SPSS (Statistical Package for Social Sciences) was employed as a tool to perform the data analysis using chi-square test and binary logistic regression model. The study found 19.85% internet addiction prevalence with 19.5% moderate and 0.4 % severe addiction. Researchers also concluded that gender, computer usability, lower emotional, behavioural and psychological well-being are important indicators of internet addiction.

Samha AA et al. (2018) investigated the psychometric properties of IAT with Lebanese University medical students, ages ranged from 18 to 29 years. The participants were given score using IAT questionnaire, with the maximum score being 100. Statistical analysis was performed using SPSS and cronbach's alpha established the internal consistency of each factor. The result showed 38.2% of the participants scored between 31 and 49 on the total Internet Addiction Scale signifying mild addiction level, 28.9% scored between 50 and 79 that is a sign of a moderate level of internet addiction, while 1.2% of the participants scored above 80, indicating severe reliance upon the internet.

The 14-items Japanese version of CIUS was validated by Yong RKF et al (2017), on a sample of 623 effective respondents stratified by age and sex. The study conducted by the researchers employed CIUS as a measure of IAD which has a five-point Likert scale (0 = never; 1 = seldom; 2 = sometimes; 3 = often; 4 = very often) giving a total score in the range 0–56 with a cut-off score of 28. Data were analyzed and evaluated by statistical analysis tools SPSS and AMOS. The study did not find any significant differences in the level of IAD in terms of demographic factors and result showed that 3% of the tested sample had a high CIUS score.

In recent years, researchers are exploring the idea of detecting IAD using ML techniques. Due to the excessive penetration of internet in the lives of people, it is essential to explore and extract patterns pertaining to the factors and causes of IAD in order to predict the probable cases of IAD in the internet users to benefit the internet population and minimize the risk of IAD. Although there are not many works that explored the use of ML techniques to predict the occurrence of IAD, we present a few recent studies that explored the ML techniques to detect IAD.

Ioannidis K et al (2016) used ML prediction model which included logistic regression, random forests and Naive Bayes techniques to evaluate whether Problematic Internet Use (PIU) can be predicted from recognized impulsive, compulsive traits and symptomatology. The online sample evaluated consisted of 2006 cases, out of which 181(9.0%) had moderate/severe PIU. The model robust result showed the prediction of PIU was possible and ML can be useful in the Psychiatry process.

In another study, Akhter SA (2017) predicted potential online gambling addicts using ML. The researcher used SVM and Naive Bayes classifier and find that SVM over performed Naive Bayes classifier on an average. The system can be used as an indicator for finding the potential gambling addicts.

Zeinab et al (2018) developed prediction models for diagnosis of Internet addiction users. The study used two heuristics algorithms: Particle Swarm optimization and Binary Fierfly algorithm to eliminate the insignificant features in IAT used for diagnosis of IAD. The study also employed supervised learning K-nearest neighbour algorithm for classification. The study concluded that the model developed by considering the important risk factors can identify the status of the user's internet addiction with 99.75% accuracy.

Literature suggests that the craze for adopting ML techniques to model real-life applications is increasing among researchers due to ML liberal, inductive, accurate and efficient performance over statistical data analysis. The recent studies employed ML techniques like Naive Bayes, decision tree, random forest, SVM for the purpose of exploring the existence of IAD in internet users. From literature, we also found that Naive Bayes and SVM are used more to detect IAD in the taken data sets, than other ML techniques. Naive Bayes has a disadvantage it assumes that all features of the problem domain are independent to each other so it is not a good choice for domains where features are correlated, like in this study the features of IAD coexist and are correlated. SVM cannot handle missing data, in fact it eliminates the missing data from the data set before the analysis. Also, SVM models are more like a black box, where how the reasoning occurs is not explained.

Our proposed ML models are based on BN, which provides a pictorial understanding of the key aspects of IAD and can be easily used by health practitioners as an efficient tool for reasoning in a dynamic and complex online health environment of internet users. The focus of this study is on causal reasoning among IAD factors, accurate prediction even if data is missing, and inference based system for answering queries related to the risk of IAD and its management. For example: 1) if a person is identified with a severe level of one of the factors of IAD, then what is its impact on other factors and what will be the tendency of the person towards IAD? 2) provide accurate prediction even if a person is not able to provide complete information about his/her online behaviour, 3) if a person shows a moderate and severe level of factors of IAD then what is the likelihood of a person to be addicted to internet?.

BAYESIAN NETWORKS

Bayesian Networks belong to the probabilistic graphical model framework- an automated system that is trained from available data and discovers patterns for making decisions. These models are explicitly used for uncertainties that exist in most of the real-world applications. They develop probabilistic relations among a set of random variables and can be used to predict the possible outcome from the input data. The probabilistic approach models are useful in diagnostic inference (from effect to cause), Predictive inference (from causes to effect) and combining evidence (combined effect of variables on other variables) (Zhang, 2008).

The probabilistic graphical model are based on Bayes theorem that is also referred to the theorem of Probability (Equation 1). It plays a pivotal role in updating the probabilities of unobserved events. Let us consider that there is a prior probability for the unobserved event. So if a related event occurs then by using the Bayes rule we can update the prior probability to get the posterior probability of the event. The Bayes theorem can be stated as:

$$p(A|B) = \frac{\boldsymbol{p(B|A) \times p(A)}}{\boldsymbol{p(B)}} \qquad \text{(Equation 1)}$$

where p(A|B) is the conditional probability of A, when event B has already happened, p(B|A) is the conditional probability of event B, when A has previously occurred, p(A) is the prior or probability of event A and p(B) is the probability that event B has occurred. The Bayes theorem forms the basis of computation in BN and permits to update the probabilities in a network if any instance or information changes.

Mathematically BN can be defined as B=(G, p) where G(V, E) is a directed acyclic graph with V as a set of nodes $X_1, X_2, X_3 ...... X_n$ and E as a set of directed edges. p is the joint probability over the set of variables V and n is the number of nodes in BN, depicted in Equation 2 (Estabragh et al., 2013**)**.

$$p(X_1, X_2, X_3 ..... X_n) = \prod_{i=1}^{n} (p(X_i / parent(X_i)) \qquad \text{(Equation 2)}$$

Bayesian Networks are regarded as powerful probabilistic models having two components:

- Qualitative component is a Directed Acyclic Graph having nodes representing variables and directed edges (arrows) denoting direct dependencies among the variables. The node from which the edge starts is the Parent node and the node to which the edge arrow points to is the child node. Child nodes are conditionally dependent on their parent nodes and depict the causal relationship between the parent and child nodes.
- Quantitative component forms the basis of the prior learning of the BN from the historical data set. It constitutes of Conditional Probability Table (CPT) which provides the probability at each node depending on the values for each combination of values of nodes on which the node depends.

Let us consider an example illustrated in Figure 3 to explain the components of BN. suppose we are interested to diagnose cancer in the patients who visit the clinic. X represents the probability of patient having normal/abnormal X-Ray, Y represents the probability that patient is a smoker or not and Z represents the cancer state as Yes/NO. Variables X and Y are parent nodes and variable Z is the child node, arrows depict the dependency of Z node on the X and Y nodes. On the basis of historical data, the CPT shows the probability of cancer node, in accordance with the configuration of its parent states in X-Ray and Smoking node. Using Bayes rule we can find predictions such as p (Cancer=Yes/No | X-Ray=Abnormal, Smoking=Yes).

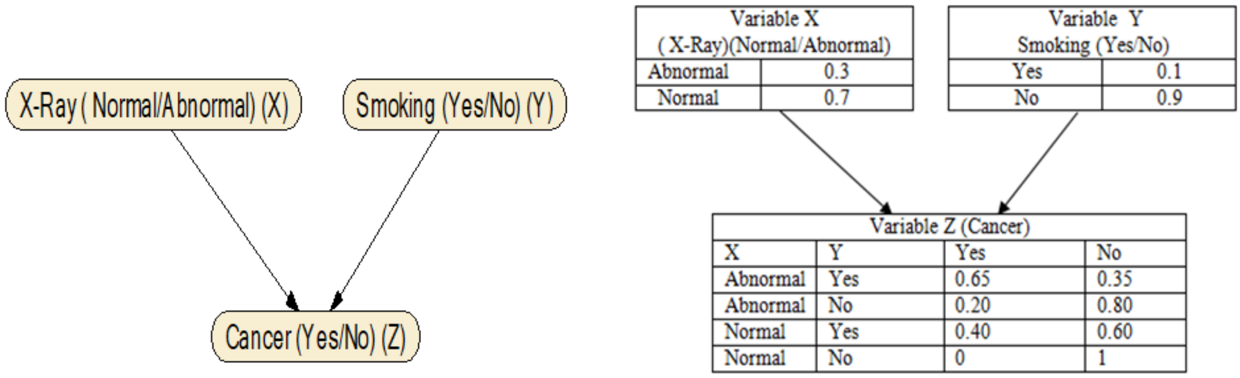

Figure 3. An Example of a Bayesian Graphical Network showing Qualitative component (Directed Acyclic Graph- Nodes with directed Arrows) and Quantitative component (Conditional Probability Table).

BN was chosen to model the detection of IAD due to its following remarkable features:

1. Bayesian Networks belongs to the probabilistic graphical model which combine qualitative analysis and quantitative analysis of data and encodes the variables of the problem domain, so missing data can be handled effectively
2. The probabilistic presentation of BN allows assessment of risks and uncertainties, thus support in decision-making process.
3. Causal relationships among variables can be obtained through BN, as a result it assists in better understanding of the problem domain and can be used to anticipate the consequences of an intervention.
4. BN versatility shows good prediction accuracy even with smaller and incomplete data sets, it also can be updated when new information becomes available.
5. BN models are analytically constructed with the conditional probability distribution available for each variable, therefore they provide fast responses to queries and instances.

METHODOLOGY

The objective of this paper is to develop BN models for analyzing IAD, based on the IAT, CIUS and PIUQ predictors, symptoms and prior known causal relationships, with a key objective of helping counsellors, teachers and health professionals in their decisional process.

As mentioned earlier, the two components of BN involves the process of identification of variables of IAD and establishing the cause and effect relationship among them (Qualitative analysis) and then construct the conditional probability tables (Quantitative analysis) to achieve the prior learning of the model. The models constructed are tested for accuracy before the models can be used for real-world applications.

The IAD models are built by employing three assessment tools of IAD: IAT, CIUS, PIUQ as data collection tools. The automation process of IAD models was achieved by exploiting the significant features of BN, so the three IAD models were named as IAT-BN, CIUS-BN, PIUQ-BN.

Internet Addiction Test (IAT)

IAT questionnaire is a widely used 20-item questionnaire developed to measure the presence and severity of IAD. Each item is based upon a 5-point Likert scale, where zero denotes not applicable, one indicates rare online behaviour, two refers to occasional engagement with internet, three as frequent usage, four as often involvement with internet and five denotes always online behaviour. The maximum score 100 points represent the levels of Internet addiction as 0 to 30 points considered as normal; 31 to 49 indicates the presence of mild level of IAD; 50-79 reflects moderate level of IAD and 80 to 100 points indicates severe level of IAD. (See questions in the third column of Table 2)

Table 2 Psychometric Factors of IAT with Measuring Criteria

Observation & Target Variables	**Variable description**	**Questions**	**5 point Likert scale (0 to 5)**	**Total Score with classification (normal,mild, moderate,severe)**
Salience (Node A) Observation variable 5 items	Most reliable factor of IAD, where internet dominates the user's thoughts and behaviour	1)How often do you block out disturbing thoughts about your life with soothing thoughts of the Internet? 2)How often do you fear that life without the Internet would be boring, empty, and joyless? 3)How often do you snap, yell, or act annoyed if someone bothers you while you are online? 4)How often do you feel preoccupied with the Internet when off-line, or fantasize about being online? 5)How often do you choose to spend more time online over going out with others?	0: Does not Apply; 1: Rarely; 2:Occasionally; 3:Frequently; 4: Often; 5: Always	Total score -25 0 to 6 normal; 6 to 12 mild; 12to 18 moderate; 18 to 25 severe.
Excessive Use (Node B) Observation Variable 5 items	Factor which shows excessive online engagement and indicate compulsive online behaviour of the internet users	1)How often do you find that you stay online longer than you intended? 2)How often do you neglect household chores to spend more time online? 3)How often do you lose sleep due to late-night log-ins? 4)How often do you try to hide how long you've been online? 5)How often do you feel depressed, moody or nervous when you are off-line, which goes away	0: Does not Apply; 1: Rarely; 2:Occasionally; 3: Frequently; 4: Often; 5: Always	Total score -25 0 to 6 normal; 6 to 12 mild; 12 to 18 moderate; 18 to 25 severe.

Observation & Target Variables	Variable description	Questions	5 point Likert scale (0 to 5)	Total Score with classification (normal,mild, moderate,severe)
		once you are back online?		
Neglect work (Node C) Observation variable 3 items	Major factor which hampers the performance and productivity of the internet users	1)How often do your grades or school work suffers because of the amount of time you spend online? 2)How often does your job performance or productivity suffer because of the Internet? 3)How often do you become defensive or secretive when anyone asks you what you do online?	0: Does not Apply; 1: Rarely; 2: Occasionally; 3: Frequently; 4: Often; 5: Always	Total score-15 0 to 3 normal; 3 to 7 mild; 7 to 11 moderate; 11 to 15 severe.
Anticipation (Node D) Observation variable 2 items	This refers to dreaming about being online and keeps the internet users preoccupied	1)How often do you check your e-mail before something else that you need to do? 2)How often do you find yourself anticipating when you will go online again?	0: Does not Apply; 1: Rarely; 2: Occasionally; 3: Frequently; 4: Often; 5: Always	Total score-10 0 to 2 normal; 2 to 5 mild; 5 to 8 moderate; 8 to 10 severe.
Lack of control (Node E) Observation variable 3 items	Factor which is identified by staying online longer than intended	1)How often do others in your life complain to you about the amount of time you spend online? 2)How often do you find yourself saying “just a few more minutes” when online? 3)How often do you try to cut down the amount of time you spend online and fail?	0: Does not Apply; 1: Rarely; 2: Occasionally, 3: Frequently; 4: Often; 5: Always	Total score-15 0 to 3 normal; 3 to 7 mild; 7 to 11 moderate; 11 to 15 severe.
Neglect Social Life (Node F) Observation variable 2 items	Social relationship deficit factor	1)How often do you prefer the excitement of the Internet to intimacy with your partner? 2)How often do you form new relationships with fellow online users?	0: Does not Apply; 1: Rarely; 2: Occasionally; 3: Frequently; 4: Often; 5: Always	Total score-10 0 to 2 normal; 2 to 5 mild; 4 to 8 moderate; 8 to 10 severe
Internet Addiction Disorder (Node H) Target Node	Unobserved node which represent the absence and presence of IAD in internet users			Total score-100 0 to Below 50: IAD Absent; 50 to 100: IAD Present.

Note: Conducted Pilot study for Psychometric Factors of IAT (Reference: DOI https://doi.org/10.1007/978-981-10-8527-7_8)

Compulsive Internet Use Scale (CIUS)

CIUS questionnaire consists of 14 items, each item is measured using a 5-point Likert scale where zero denotes no engagement with internet; one means seldom usage; two denotes sometimes; three indicates internet is used often; four signify very often usage of internet. Total score point sums up to 56 points. A cut-off score of 28 was considered and the scores were divided into three equal-width tiers (tier 1 = 0–18; tier 2 = 19–37; tier 3 = 38–56) to test the level of Internet addiction. (See questions in the third column of Table 3)

Table 3 Psychometric Factors of CIUS with Measuring Criteria

Observation & Target Variables	Variable description	Questions	5 point Likert scale (0 to 4)	Total Score with classification (normal,mild, moderate,severe)
Loss of Control (Node A) Observation variable 4 items	This factor refers to urge for being online and user is unable to stop using internet	1) How often do you find it difficult to stop using the internet when you are online? 2) How often do you continue to use the internet despite your intention to stop? 3) How often are you short of sleep because of the internet? 4) How often have you unsuccessfully tried to spend less time on the internet?	0: Never; 1: Seldom; 2: Sometimes; 3: Often; 4: Very often	Total score - 16 0 to 4 normal; 4 to 8 mild; 8 to 12 moderate; 12 to 16 severe.
Preoccupation (Node B) Observation Variable 3 items	Factor includes mental and behavioural preoccupation. It shows excessive online engagement	1) How often do you prefer to use the internet instead of spending time with others (e.g. partner, children, parents, friends)? 2) How often do you think about the internet, even when not online? 3) How often do you look forward to your next internet session?	0: Never; 1: Seldom; 2: Sometimes; 3: Often; 4: Very often	Total score -12 0 to 3 normal; 3 to 6 mild; 6 to 9 moderate; 9 to 12 severe.
Withdrawal (Node C) Observation variable 1 item	Factor which depicts a negative mood when offline	1) How often do you feel restless, frustrated, or irritated when you cannot use the internet?	0: Never; 1: Seldom; 2: Sometimes; 3: Often; 4: Very often	Total score-4 0 to 1 normal; 1 to 2 mild; 2 to 3 moderate; 3 to 4 severe.
Coping/Mood modifications (Node D) Observation variable 2 items	This refers to enhancement of mood by going online	1) How often do you go on the internet when you are feeling down? 2) How often do you use the internet to escape from your sorrows or get relief from negative feelings?	0: Never; 1: Seldom; 2: Sometimes; 3: Often; 4: Very often	Total score-8 0 to 2 normal; 2 to 4 mild; 4 to 6 moderate; 6 to 8 severe.

Observation & Target Variables	Variable description	Questions	5 point Likert scale (0 to 4)	Total Score with classification (normal,mild, moderate,severe)
Conlict (Node E) Observation variable 4 items	Factor which includes inter and intrapersonal conflict. It is identified by staying online longer and neglecting the daily routine activities.	1) How often do others (e.g. partner, children, parents, friends) say you should use the internet less? 2) How often do you think you should use the internet less often? 3) How often do you rush through your (home) work in order to go on the internet? 4) How often do you neglect your daily obligations (work, school or family life) because you prefer to go on the internet?	0: Never; 1: Seldom; 2: Sometimes; 3: Often; 4: Very often	Total score - 16 0 to 4 normal; 4 to 8 mild; 8 to 12 moderate; 12 to 16 severe.
Internet Addiction Disorder (Node F) Target variable	Unobserved node which represent the absence and presence of IAD in the internet users			Total score- 56 0 to Below 28: IAD Absent; 28 to 56: IAD Present.

Problematic Internet Use Questionnaire (PIUQ)

PIUQ is comprised of 18 items and each item is measured on a 5-point Likert scale where one denotes that person has never used internet for the particular situation; two indicates person is rarely using internet; three refers to sometimes usage of internet; four signifies often use of internet and five indicates always online behaviour. The 18 item points sum up to 90 and the level of internet addiction ranges from 18 to 90. (See questions in the third column of Table 4)

Table 4 Psychometric Properties of PIUQ with Measuring Criteria

Observation & Target Variables	Variable description	Items	5 point Likert scale (1 to 5)	Total Score with classification normal,mild, moderate, severe
Obsession (Node A) Observation variable 6 items	Factor which indicates being obsessed with Internet activities	1)How often do you fantasize about the Internet, or think about what it would be like to be online when you are not on the Internet? 2) How often do you daydream about the Internet? 3) How often do you feel tense, irritated, or stressed if you cannot use the Internet for as long as you want to?	1: Never; 2: Rarely; 3: Sometimes; 4: Often; 5: Always	Score Range: 6-30 6 to 12 normal; 12 to 18 mild; 18 to 24 moderate; 24 to 30 severe.

Observation & Target Variables	**Variable description**	**Items**	**5 point Likert scale (1 to 5)**	**Total Score with classification normal,mild, moderate, severe**
		4) How often do you feel tense, irritated, or stressed if you cannot use the Internet for several days? 5) How often does it happen to you that you feel depressed, moody, or nervous when you are not on the Internet and these feelings stop once you are back online? 6) How often do you dream about the Internet?		
Neglect (Node B) Observation variable 6 items	Factor which depicts ignoring non-internet activities	1) How often do you neglect household chores to spend more time online? 2) How often do you spend time online when you'd rather sleep? 3) How often do you choose the Internet rather than being with your partner? 4) How often does the use of Internet impair your work or your efficacy? 5) How often do people in your life complain about spending too much time online? 6) How often do you choose the Internet rather than going out with somebody to have some fun?	1: Never; 2: Rarely; 3: Sometimes; 4: Often; 5: Always	Score Range: 6-30 6 to 12 normal; 12 to 18 mild; 18 to 24 moderate; 24 to 30 severe
Control Disorder (Node C) Observation variable 6 items	Factor which indicates being unable to stop the use of the internet	1) How often do you feel that you should decrease the amount of time spent online? 2) How often does it happen to you that you wish to decrease the amount of time spent online but you do not succeed? 3) How often do you try to conceal the amount of time spent online? 4) How often do you feel that your Internet usage causes problems for you? 5) How often do you realize saying when you are online, "just a couple of more minutes and I will stop"? 6) How often do you think that you should ask for help in relation to your Internet use?	1: Never; 2: Rarely; 3: Sometimes; 4: Often; 5: Always	Score Range: 6-30 6 to 12 normal; 12 to 18 mild; 18 to 24 moderate; 24 to 30 severe.

Observation & Target Variables	Variable description	Items	5 point Likert scale (1 to 5)	Total Score with classification normal,mild, moderate, severe
Internet Addiction Disorder (Node D) Target variable	Unobserved node which represent the absence and presence of IAD in the internet users			Score Range: 18-90 18 to Below 54 IAD Absent; 54 to 90 IAD Present.

The questionnaires form the domain knowledge for IAD models and in the following sections of the paper construction and implementation of the BN models are explained by discussing the Qualitative and Quantitative analysis of models in detail.

Qualitative Analysis of IAD Models

The structure of models is realized by describing the psychometric properties (Observed and Target Variables) of the assessment tools of IAD and presenting them in a graphical manner.

IAT-BN Model: In order to achieve the qualitative component of the IAT-BN model, the IAT items are grouped in 6 factors: Salience with 5 items; Excessive Use comprised of 5 items; Neglect work formed by 3 items; Anticipation has 2 items; Lack of control and Neglect social life with 3 and 2 items respectively. These six factors of IAD become the nodes in IAT-BN and to gain the CPT at each node we divided the measurement of each factor in four tiers (indicating normal, mild, moderate and severe online behaviour) as shown in the fifth column of Table 2. Figure 4 portrays the graphical structure of the IAT-BN model showing the causal relationship depicted by arrows among the identified variables. A pilot study of the IAT-BN model was conducted and model was evaluated using randomly generated data set (Ref: DOI: 10.1007/978-981-10-8527-7_8).

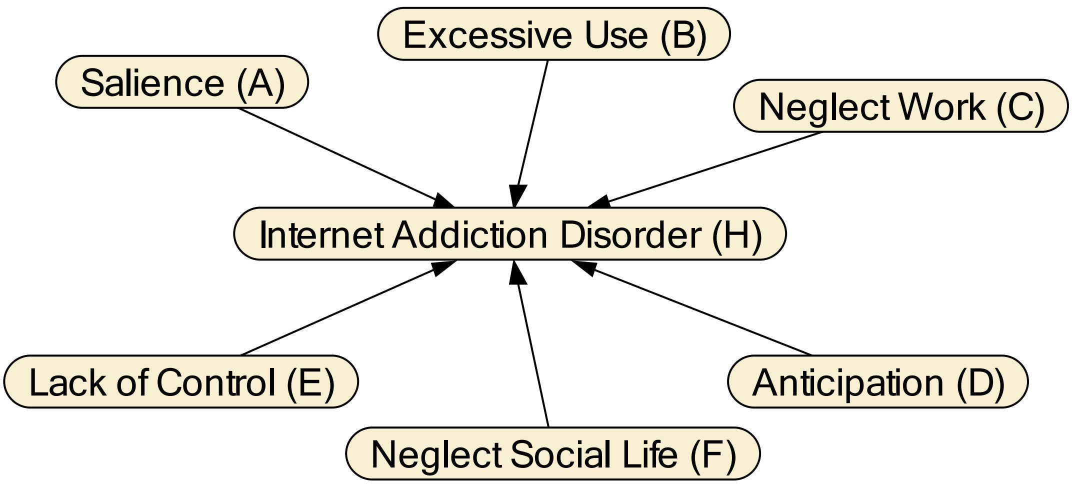

Figure 4. The established Bayesian Network structure of IAT-BN Model. Nodes-A, B, C, D, E, F (Observation variables). Node-H (Target variable).

CIUS-BN Model: The CIUS-BN model based on the psychometric properties of CIUS was formulated by following the same procedure as of IAT-BN model. Table 3 describes the factors of IAD with their measuring criteria and Figure 5 portrays the graphical structure of the CIUS-BN model. The CIUS-BN model in Figure 5, illustrates five observed variables as nodes A,B,C,D,E (Loss of control, preoccupation, Withdrawal, Coping/mood modifications, Conflict) and target variable IAD as node F. The arrows depict the influence of observed variables on the target variable.

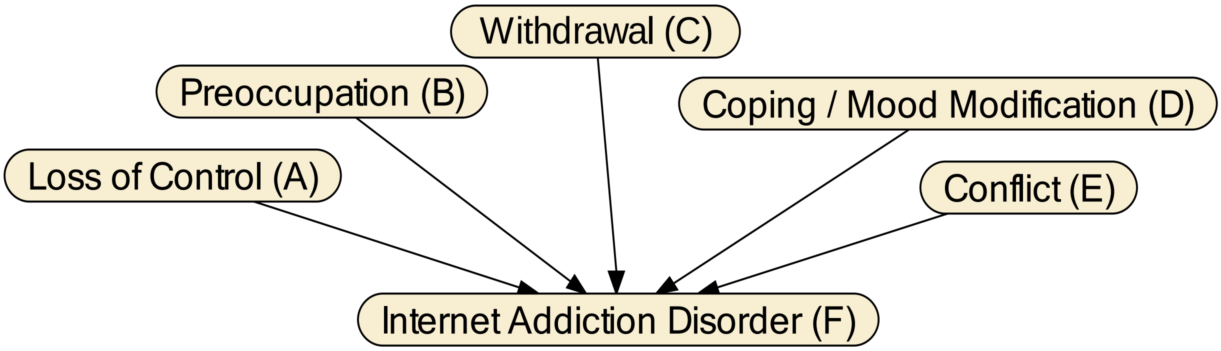

Figure 5. The established Bayesian Network structure of CIUS-BN Model. Nodes-A, B, C, D, E (Observation variables). Node-F (Target variable).

PIUQ-BN Model: The psychometric properties associated with the PIUQ questionnaire were grouped in three factors and their measurement criteria are summarised in Table 4. The established BN structure of the PIUQ-BN model is displayed in Figure 6, where three observed variables are nodes A, B, C (Obsession, Neglect, Control disorder) and target variable is node D (IAD). The edges signify the effect of observed variables on the target variable.

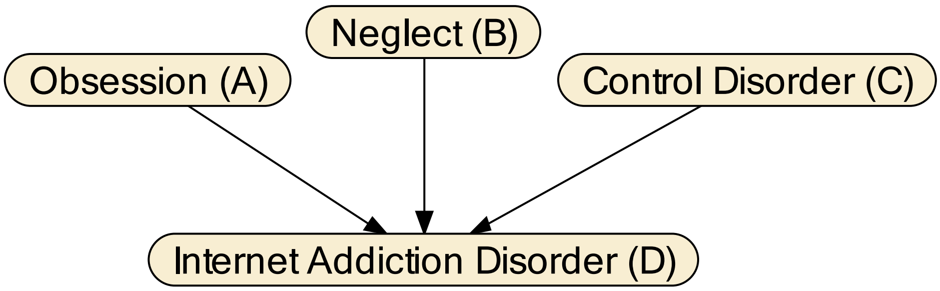

Figure 6. The established Bayesian Network structure of PIUQ-BN Model. Nodes-A, B and C (Observation variables). Node-D (Target variable).

Quantitative Analysis of Models

The domain knowledge i.e. the description of the factors responsible for the occurrence of IAD with their measure criteria becomes the qualitative parameter of the BN models. In order to attain the quantitative component of the models, we used reliable and high power BN development software, Netica from Norsys (Norsys Corporation, http://www.norsys.com**)**. We created the IAD graphical network in Netica and encoded the network that forms the conditional probability tables at each node. The features of Netica were exploited for quantitative analysis of the models i.e. prior learning environment was achieved through the real data and CPT was assigned to the nodes in the models. The three IAD models IAT-BN, CIUS-BN, PIUQ-BN generated after prior learning with the historical data, are depicted in Figure 7, Figure 8 and Figure 9 respectively.

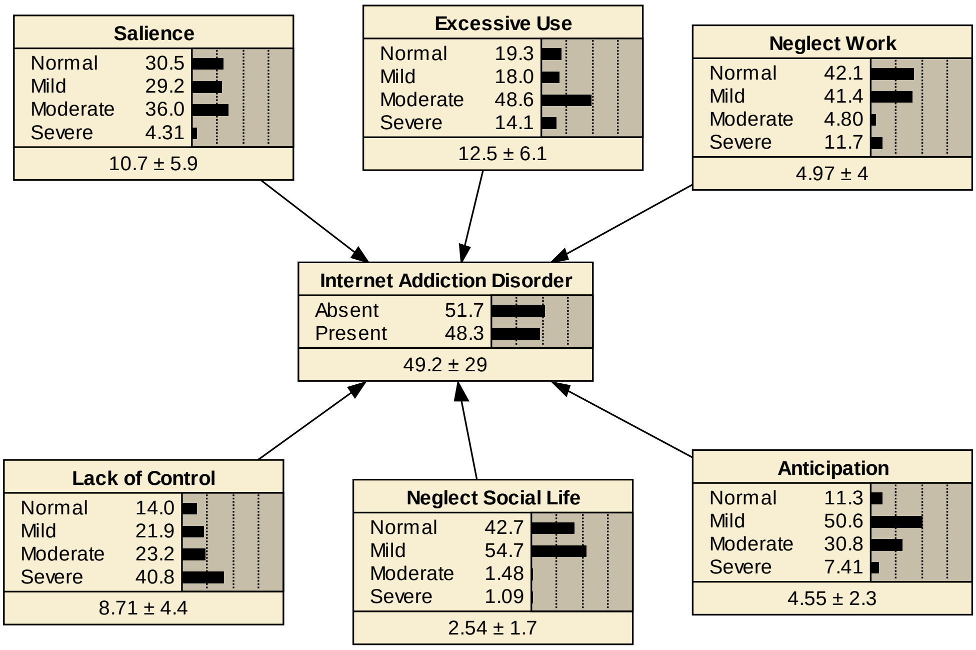

Figure 7. IAT-BN Model after prior learning - Structured to analyze the occurrence of IAD.

The IAT-BN model in Figure 7 shows that the prior probability of IAD absent in the data set is 51.7% and the prior probability of IAD present in the historical data set is 48.3%.

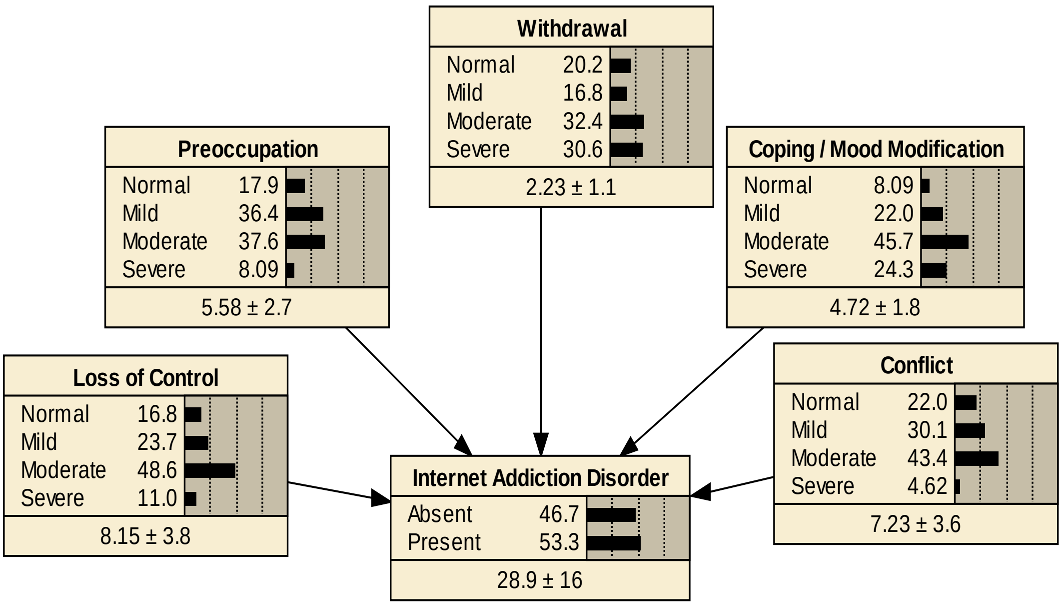

Figure 8. CIUS-BN Model after prior learning - Structured to analyze the occurrence of IAD.

The prior probability of IAD absent and IAD present in the data set was found to be 46.7 % and 53.3% respectively, in the CIUS-BN model, see Figure 8. The PIUQ-BN model portrayed in Figure 9, was achieved by prior learning through historical real data and showed IAD absent and IAD present prior probability as 45% and 55% respectively.

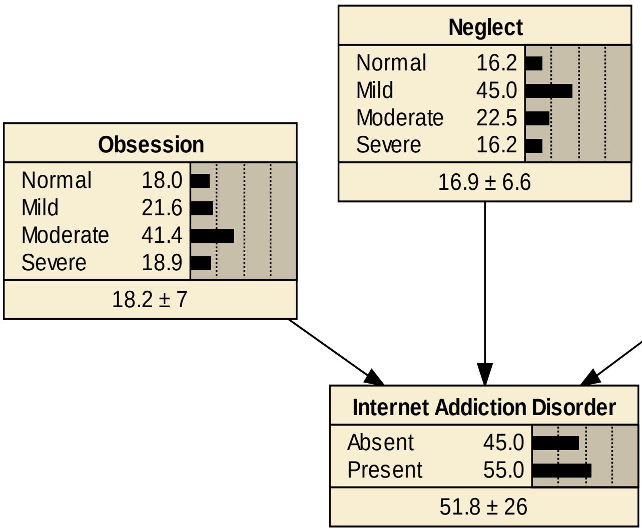

Figure 9. PIUQ-BN Model after prior learning - Structured to analyze the occurrence of IAD.

EXPERIMENT AND RESULTS

Following section of the paper explains the setup of the experiments for the models including data set preparation and summarises the results obtained.

Experiment setup

Data collection and data pre-processing: Respondents were approached through facebook and whatsapp, to fill online three sets of questionnaires, based on IAT, CIUS, PIUQ. We present the details of the three data sets in Figure 10.

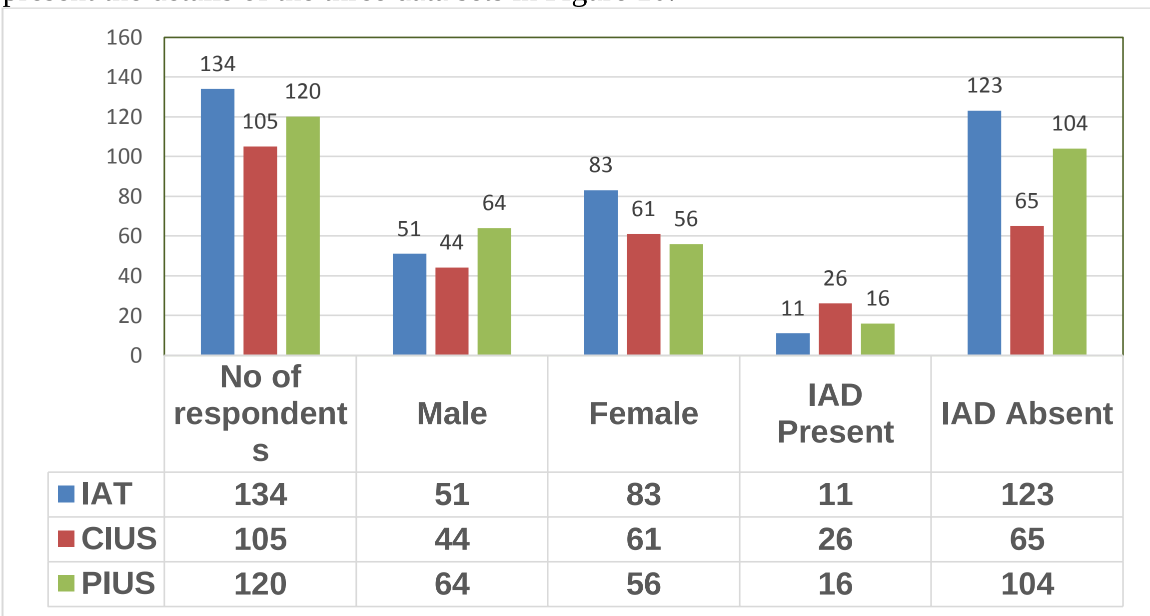

Figure 10. Description of three Real Data Sets collected online using IAT, CIUS and PIUQ Questionnaires.

After analyzing the statistical data, we observed that data sets were with imbalanced class distributions, which is quite common in many real applications (Chawla et al., 2002). The nature of the problem of internet addiction predicates dealing with highly imbalanced classes, as respondents do not necessarily disclose their true behaviour of online activities. We observed that more than 80% of respondents were not addicted to internet in the IAT and PIUQ data set. In CIUS data set more than 60% of respondents did not fall in the addicted category. The construction of models using the real data sets yielded a highly biased classifier which predicted the majority class most number of times. In order to diminish the imbalanced class problem, we first eliminated the response of the respondents who were apprehensive in revealing their true internet usage activities and responded mostly in “Does not apply”, “Never” or “Rarely”. Secondly, to mitigate the effects of imbalance, we applied the SMOTE (Synthetic Minority Oversampling Technique) approach to create a balanced data set (Chawla et al., 2002**)**. Finally, 100 respondents per data set were considered for IAD analysis.

IMPLEMENTATION OF MODELS

To implement and analyze the performance of the models, we employed the Hold-out technique of model validation. Once the models are constructed it is important to measure their performance. In the Hold-out method of validation the data is split into two parts: training set and test set (either 50-50 or two-thirds and one-third). The training set forms the historical data through which the models learn and gain prior knowledge. The test set is used to measure the performance of the model to examine how well it performs on unseen data (Tan et al., 2006). In this study, each set of data was split into two sets: 80% data for training/learning of models and 20% data for testing of models. The graphical interface feature of Netica was utilized to train the models with the training data set. The three IAD models IAT-BN, CIUS-BN, PIUQ-BN obtained after the prior learning of the parameters with training data set, are depicted in Figure 7, Figure 8 and Figure 9 respectively.

EXPERIMENT RESULTS

The models were tested with 20% test data (20 real cases) and results obtained are depicted in Table 5, which denotes the confusion matrix often used to describe the performance of the classification models. True Positives (TP) stands for the number of positive cases predicted correctly i.e. the cases predicted by the model as IAD present and in actual they do have IAD, False Positives (FP) corresponds to the number of negative cases wrongly predicted by the model i.e. the model predicted IAD present in the cases, but they do not actually have IAD. False Negatives (FN) corresponds to the number of positive cases predicted as negative i.e. the cases where the model predicted IAD absent, but they actually have IAD. True Negatives (TN) is the number of negative cases predicted correctly by the models i.e. cases where the model predicted IAD absent and actually also they do not have IAD. Error rate represents the number of all incorrect predictions divided by the total number of cases in the dataset (the best error rate is 0.0 whereas the worst is 1.0). The Area under receiver operating curve (AUC) is an important measure that can be used to ascertain the performance of the classification models (A perfect test is represented by an area of 1 and an area of 0.5 represents a worthless test) (Narkhede, 2018).

Table 5 Result of Experiments performed on IAD Models (Confusion Matrix)

IAD Models	**True Positives (TP)**	**False Positives (FP)**	**False Negatives FN)**	**True Negatives (TN)**	**Error Rate**	**Area Under Curve (AUC)**
IAT-BN	18	0	0	2	0%	1
CIUS-BN	14	0	1	5	5%	0.987
PIUQ-BN	14	0	2	4	10%	0.968

The performance of the models was also evaluated by measuring Accuracy, Precision and Recall of the models, the result is shown in Table 6. Accuracy is the ratio of correctly predicted cases to the total cases and it denotes how often the model is correct. The Precision determines the fraction of cases that turn out as positive and model has also declared them as positive cases i.e. when the model predicts IAD present, then how often it is correct. Recall measures the fraction of positive cases correctly predicted by the model i.e. when it is actually IAD present, how often model predicts IAD present. Analysis of test results in Table 5 and Table 6 shows that experiment results were promising and the IAT-BN model is the best fit predictive model among the three, which can detect the instances of IAD with 100% accuracy.

Table 6 Performance Comparison of IAD Models

IAD Models	Accuracy	Precision	Recall
IAT-BN	100%	1	1
CIUS-BN	95%	1	0.933
PIUQ-BN	90%	1	0.875

EVALUATION OF MODELS

Before the application of models in real-life instances, the robustness of the models was assessed by using prediction and sensitivity analysis techniques.

Prediction Analysis

The quality of models was established by performing probabilistic inference i.e. causal reasoning (Heaton, 2013). The causal reasoning in Bayesian Network demonstrates the influence of varying the values of observation nodes on the prospect of occurrence of IAD. Let us consider a hypothetical instance/query and compare it with the IAT-BN inference result. In the model, the evidence can be entered by clicking on the state of the variable and this procedure will condition the probabilities of the node on the clicked nodes. With this, we can explore and infer how the model reasons and represent knowledge. Due to space limitations, we present one instance for the IAT-BN model only.

Instance 1 for IAT-BN model: In case of Severe level of *Salience*, Moderate level of *Excessive Use* and Severe *Lack of Control*, one should consider IAD present.

Inference: We click *Salience*=Severe, *Excessive Use*=Moderate and *Lack of Control*=Severe and observe the conditional probabilities of the IAD node. We see in Figure 11, that in particular IAD absent probability comes down from 51.7 % to 37.1%;

IAD present probability goes up from 48.3% to 62.9%. So it can be inferred that the person has a high probability to be diagnosed as addicted to the internet.

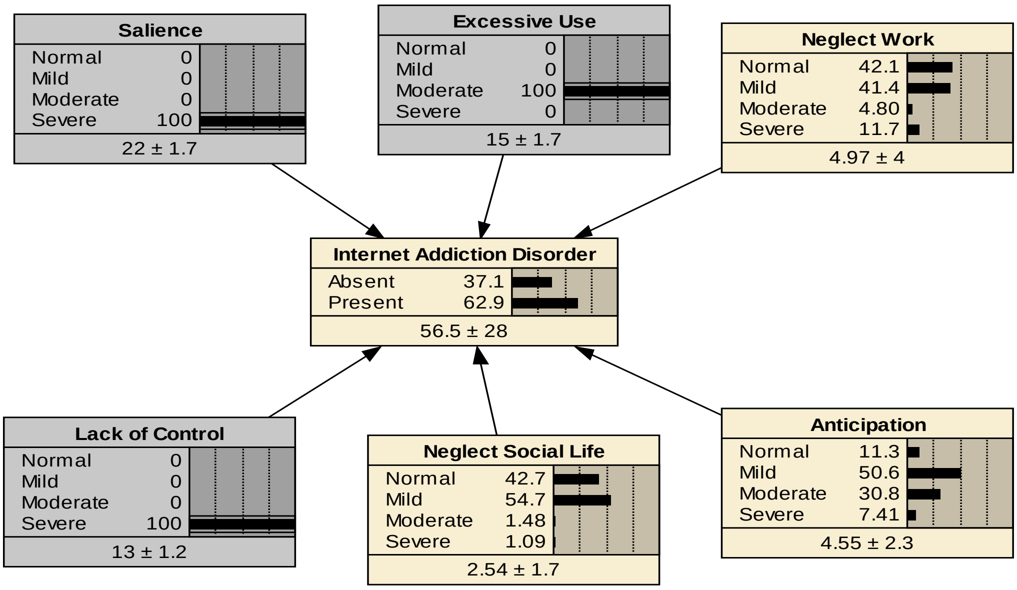

Figure 11. IAT-BN Model in the state representing a Hypothetical Instance. Highlighted nodes are evidence clicked.

In order to gain insight into the performance of the Models we also performed prediction analysis, by matching the predicted outcome of the models with the actual results (Mohammadfam et al., 2017). Few real cases were tested with IAT-BN, CIUS-BN and PIUQ-BN models and the result is elaborated in Table 7.

Table 7 Matching Result of Actual outcome with Outcome from IAD Models for Real Cases

IAD Models	Real Cases	Actual result	Predicted outcome from model
IAT-BN	Case: *Salience* - Severe,*Excessive Use* - Severe, *Neglect Work* - Severe,*Anticipation* - Moderate, *Lack of Control* - Moderate, *Neglect Social Life* - Moderate	Present	Present
CIUS-BN	Case: *Withdrawal* - Normal,*Preoccupation* -Mild, *Loss of Control* - Mild, *Conflict* - Mild *Coping/Mood modifications* - Mild	Absent	Absent
PIUQ-BN	Case: *Obsession* - Moderate, *Neglect* - Severe, *Control Disorder* - Severe	Present	Present

Sensitivity Analysis

Bayesian Network software Netica has a significant feature Sensitivity to findings, which can be used to determine which observed variables/factors most affect the unobserved variable of interest (Lucas et al., 2004). We utilized this feature to identify the factors which when restricted may substantially reduce the probability of a person to be diagnosed as addicted to Internet. Results for sensitivity findings at various nodes of the IAD models are presented in Figure 12.

The sensitivity analysis result is based on the historical data and we observe that among the six parameters, the most sensitivity of the IAD node in the IAT-BN model is related to the node of Lack of Control followed by the salience factor in the IAT-BN model. The most sensitivity of the IAD node is associated to Loss of Control factor followed by Conflict factor, among the five underlying factors causing IAD in the CIUS-BN model. The most sensitivity of the IAD node is linked to the Obsession factor followed by Control Disorder factor in the PIUQ-BN model.

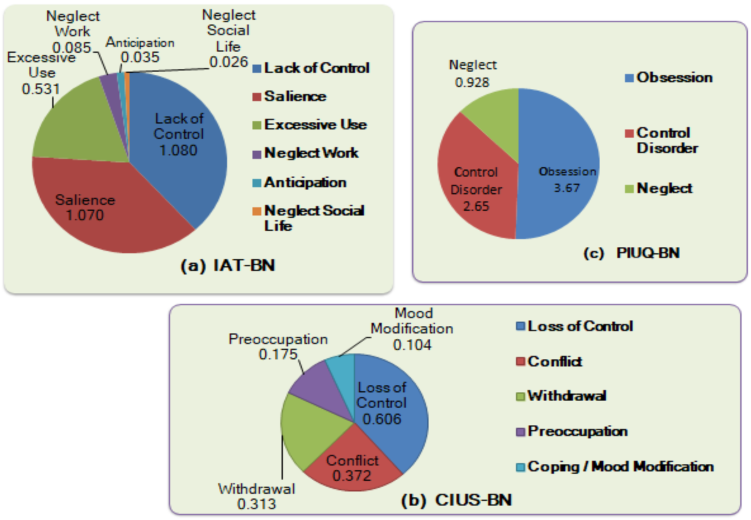

Figure 12. Sensitivity Analysis results of the three IAD Models (a) IAT-BN, (b) CIUS-BN and (c) PIUQ-BN.

DISCUSSION

In this paper, we present a comprehensive study of the detection and description of Internet Addiction Disorder in internet users. We proposed and developed three graphical interface predictive models, using supervised machine learning technique, Bayesian Networks. Bayesian Networks was chosen for this study as it reflects the symptoms of the IAD in a graphical manner and describes how the symptoms are related by probabilities. Thus providing an effective visualization of IAD, hiding the intricate calculations underlying in its detection.

We employed Internet Addiction Test, Compulsive Internet Use Scale, Problematic Internet Use questionnaire as domain knowledge, to design helpful and efficient predictive models. The models were constructed using real data sets and their robustness was evaluated on real cases. However, the historical data sets collected online did not consider any particular age group of the respondents.

The experimental results showed the superior performance of the IAT-BN model with 100% accuracy, over CIUS-BN model with 95% and the PIUQ-BN model with 90%. We see a variation in the performance of the models, the key reason can be that the respondents actively participated in filling the first questionnaire based on IAT, in comparison to the CIUS and PIUQ questionnaires which were filled in sequence. Also, the respondents found few items as repetition and may not have filled the other two questionnaires with interest. However, if the real data sets have been collected from clinics or health organizations, the proposed models may have performed better.

We also performed prediction and sensitivity analysis and presented the influence of certain factors on the probability of occurrence of IAD and summarized the identification of underlying factors when controlled, can reduce the probability of a person being addicted to internet. The paper result vividly demonstrates how the use of BN modelling can help to analyze the factors for IAD and how their interpretation helps to manage the health of the online population. The study also illustrated the rich knowledge base of IAD models by comparing a hypothetical instance with the BN inference result. The models thus created can be implemented by professional health care-takers to assist in the diagnosis of the IAD in the internet users.

In future, the historical data sets will be replaced by exploring the posts or messages of active users of social networking sites like Facebook, Twitter, Instagram etc. and we propose to ascertain their tendency towards IAD by integrating machine learning techniques.

REFERENCES

Akhter, S.A. (2017). Using machine learning to predict potential online gambling Addicts. ResearchGate. doi: 10.13140/RG.2.2.22661.19685

American Psychiatric Association (2014). New Research Press Briefing: Internet Addiction:Review of Neuroimaging Studies. Retrieved from https://www.psychiatry.org/newsroom/news-releases/internet-addiction-review-of-neuroimaging-studies.

Auerbach, R. P., Mortier, P., Bruffaerts, R. et al.,(2018). WHO World Mental Health Surveys International College Student Project: Prevalence and Distribution of Mental Disorders,Journal of Abnormal Psychology, American Psychological Association, 127,623–638. http://dx.doi.org/10.1037/abn0000362.

Chaudhari, B., Menon, P., Saldanha, D., Tewari, A., Bhattacharya, L.(2015). Internet addiction and its determinants among medical students. Ind Psychiatry J ,24,158-62.

Chaudhury, P., Tripathy, H.K. (2018). A Study on impact of smartphone Addiction on academic performance, International Journal of Engineering & Technology, 7 (2.6), 50-53. doi: 10.14419/ijet.v7i2.6.10066.

Chawla, N. V., Bowyer, K. W., Hall, L. O., Kegelmeyer, W.P. (2002).SMOTE:Synthetic Minority Over Sampling Technique. Journal of Artificial Intelligence Research, 16,28-331.

David, F. (2017).Types of Machine Learning Algorithms You Should Know; Towards Data Science. Retrieved from https://towardsdatascience.com/types-of-machine-learning-algorithms-you-should-know-953a08248861.

Demetrovics, Z., Király, O., Koronczai, B., Griffiths, M.D., Nagygyörgy, K., Elekes, Z., et al.(2016). Psychometric Properties of the PIUQ Short-Form (PIUQ-SF-6) in a Nationally Representative Sample of Adolescents. PLoS ONE 1(8), e0159409.doi:10.1371/journal. pone.0159409.

Demetrovics, Z., Szeredi, B., Rozsa, S. (2008). The three-factor model of Internet addiction: the development of the PIUQ. Behavior Research Methods, 40(2),563–74.

Estabragh, Z. S., Riahi Kashani, M.M., Jeddi Moghaddam F. et al.(2013). Netw Model Anal Health Inform Bioinforma, 2, 257. https://doi.org/10.1007/s13721-013-0042-x.

Gedam, S. R., Ghosh, S., Modi, L., Goyal, A., Mansharamani, H.(2017). Study of internet addiction: Prevalence, pattern, and psychopathology among health professional undergraduates. Indian Journal of Social Psychiatry,33,(4),305-31. doi: 10.4103/ijsp.ijsp_70_16.

Gregory, C.(2017). PSYCOM Internet addiction disorder Signs, Symptoms, and Treatments. Retrieved from https://www.psycom.net/iadcriteria.html.

Heaton, J. (20130. BN for Predictive Modeling,Article from: Forecasting &Futurism, 7. Retrieved from https://www.soa.org/Library/Newsletters/Forecasting.../july/ffn-2013-iss7-heaton.aspx.

IAT-Manual(2015).Retrieved from http:\\\netaddiction.com/wp-content/uploads/2015/11/IAT-Manual.doc.

Ioannidis, K., Chamberlain, S.R., Treder, M.S., Kiraly, F. et al.(2016). Problematic internet use(PIU): Associations with the impulsive-compulsive spectrum. An application of machine learning in psychiatry , 83:94-102. doi: 10.1016/j.jpsychires.2016.08.010.

kaltiala-heinoa, R., Lintonena, T., Rimpela, A. (2004). A internet addiction? potentially problematic use of the internet in a population of 12–18 year-old adolescents. Addiction research and theory, 12, 1, 89–96. https://doi.org/10.1080/1606635031000098796.

Kelleher, J. D.,Namee, B.M., DÁrey, A. (2015). Fundamentals of Machine Learning for

Lucas, P.J., Van der Gaag, L.C., Abu-Hanna, A. (2004). A Bayesian networks in biomedicine and health-care, In Proceedings of Artificial Intelligence in Medicine. Artif Intell Med,30(3), 201-14. doi:10.1016/j.artmed.2003.11.001.

Mcnulty, E. (2015). Dataconamy. Retrieved from https://dataconomy.com/2015/01/whats-the-difference-between-supervised-and-unsupervised-learning/.

Meerkerk, G.J., Van Den Eijnden, R.J., Vermulst, A.A., Garretsen, H.F. (2009). The Compulsive Internet Use Scale: some psychometric properties. CyberPsychology & Behavior, 12(1),1–6. doi: 10.1089/cpb.2008.0181.

Mohammadfam, I., Ghasemi, F., Kalatpour, O., Moghimbeigi, A. (2017). Constructing a Bayesian network model for improving safety behavior of employees at workplaces. Elsevier Applied Ergonomics, 58, 35-47. doi: 10.1016/j.apergo.2016.05.006.

Musser, E. D., Karalunas,S.L., Dieckmann, N., Peris,T,S, Nigg,J.T. (2016). Attention-Deficit/Hyperactivity Disorder Developmental Trajectories Related to Parental Expressed Emotion. Journal of Abnormal Psychology, American Psychological Association,125, No. 2, 182–195, http://dx.doi.org/10.1037/abn0000097.

Narkhede, S. (2018). Understanding Confusion Matrix. Retrieved from https://towardsdatascience.com/understanding-confusion-matrix.Norsys Software Corporation .Application for belief networks and influence diagrams.. http://www.norsys.com.

Predictive Data Analytics, MIT Press,40-43,312. Retrieved from www.amazon.com/Fundamentals-Machine-Learning-Predictive-Analytics-ebook/dp/B013FHC8CM.

Samaha, A. A., Fawaz, M., Yahfoufi, N.E., Gebbawi, M., Abdallah, H., Baydoun, S. A., Ghaddar, A., & Eid, A.H. (2018). Assessing the Psychometric Properties of the IAT (IAT) Among Lebanese College Students. Front Public Health. https://doi.org/10.3389/fpubh.2018.00365.

Shao, Y., Zheng, T.,Wang, Y., Liu, L., Chen, Y., Yao, Y. (2018). Internet addiction detection rate among college students in the People's Republic of China: a meta-analysis. Child adolescent psychiatry ment health,12: 25. doi: 10.1186/s13034-018-0231-6.

Statista internet Demographic & Use Worldwide digital population (2019). Retrieved from https://www.statista.com/statistics/617136/digital-population-worldwide/

Tan, P.N., Steinbach, M., Kumar, V. (2006). Introduction to Data Mining, Pearson Education In., 186-188,294-301.

Yong, R.K.F.,Inoue, A., Kawakami, N. (2017). The validity and psychometric properties of the Japanese version of the CIUS. BMC Psychiatry. https://doi.org/10.1186/s12888-017-1364-5.

Zeinab, H., Pakzad, A., Asghari, A. (2018). An Artificial Predictive Modeling Framework for Automatically Detecting Problematic Use of Internet. InternationalJournal of Computer Applications, 79, (0975 – 8887).

Zhang, N.L.(2008). Introduction to Bayesian Networks. Retrieved from https://www.cse.ust.hk/bnbook/pdf/l02.h.pdf.

Acknowledgements

The author appreciates all those who participated in the study and helped to facilitate the research process.

Conflict of Interest

The author declared no conflict of interests.

How to cite this article: A. Singh, & S. Babbar (2019). Modeling the Detection of Internet Addiction Disorder Using Bayesian Networks. *International Journal of Indian Psychology*, *7*(3), 253-276. DIP:18.01.031/20190703, DOI:10.25215/0703.031