Journal of Namibian Studies, 34(2023): 3124-3140 ISSN: 2197-5523 (online)

Examine The Internet Addiction Scale Of Students In Turkey And Iraq Comparatively With The Multivariate Adaptive Regression Splines (MARS) Method

Hewa Hassan[1], Dr. Murat Kayri [2], Dr. Hikmet Şevgin[3]

[1] Department of Accounting, Gasha Technical Institute, Iraq.
[2] Dept. of Computer and Instructional Technology Education, Faculty of Education, Van Yüzüncü Yıl University, Turkey.
[3] Department of Educational Sciences, Department of Measurement and Evaluation in Education, Faculty of Education, Van Yüzüncü Yıl University, Turkey.

Abstract

In recent years, the Internet has altered how modern life is lived. Internet addiction is considered a current condition that could affect mental health. The aim of the study is to develop a scale for internet addiction and examine the factors that can affect the addiction status of individuals. The literature uses the novel non-parametric method known as MARS. MARS has the ability to provide accurate parametric forecasts as well as readable model curves. The sample of the study is at the international level, and cosmopolitan universities from each country (Turkey and Iraq), and the target audience was determined as university students.

The data for the study came from a questionnaire designated, which was meant to evaluate a person's level of addiction. The Internet addiction scale is a 5-point Likert scale, with the lowest score that each participant can get from the scale is 35 and the highest score being 175. The population of the study consists of 2235; 1220 students (427 males and 793 females) from the Turkish sample and 1015 students (461 males and 550 females) from the Iraqi sample were assigned by random sampling method. For MARS data mining analysis, the default values of the SPM 8.2 program were taken as a basis, and operations were carried out considering the entire data set. MARS obtained nine base functions of the model for Turkey and twenty base functions of the model for Iraq. As a result of the MARS analysis method, it was proven that the predictors of Daily Internet Usage may be seen as the primary initiating factor for Internet dependence for Turkish and Iraqi students following the research model identified by MARS for both datasets. These results have been supported in the literature, which eventually resulted in the conclusion that using the Internet leads to Internet dependence.

Keywords: Internet use, Addiction, MARS, Classification.

Introduction

There are various types of regression methods, and below will be outlined some. The most important are linear, non-linear and mixed. It is clear that with increasing numbers of variables, models become correspondingly more complex, and thus analyzing relationships between observed effects and their corresponding causes is more complicated. Axel W. et al (2004) If there is a situation where parameter estimations might have become biased, it is necessary for regression methods to have different functions. For the three kinds of regression function, having the ability to use parametric methods assumes a de facto initial linear correlation. The other side of the coin is that non parametric approaches apply to situations where there is non-linear correlation among the variables. Because nothing in life is ever truly black or white, the same applies to data analysis: additive algorithm processes are employed to obtain parameter estimations that are non-biased Kayri M, Zırhlıoğlu G (2009), and result in readable curves if initial data correlation turns out as too rough and there are interpretation issues Fabio Trojani (2006).

One of the many non-parametric methods for data analysis now available, MARS – or Multivariate Adaptive Regression Spline approach – has the advantage of generating strong algorithms on which multiple variables can be incorporated and unbiased, independent parameter estimations obtained. Hébert-Losier, K., & Beaven, C. M. (2014) stated that MARS is, at its core the use of step-wise linear regression for a given set, whereby data performance can be improved. The way it does so is by generating, for each linear region of the model, a novel regression equation. This is referred to a knot. There are two approaches to MARS – forwards and backward. With a forwards approach, all the predictive variables are used to generate an effective regression analysis; while with the opposite approach, the least squares approach is used to eliminate those functions which are of little use or effectiveness. These two approaches use the concept of “knots” to eliminate as far as possible the issue of multidimensionality. According to Nalcaci, G., Özmen, A. & Weber, G.W. (2019) The main advantage of MARS compared to approaches using a single regression approach is that the research can be structured such that it includes multiple predicted variables: nominal, ordinal, or continuous, for example.

Friedman, J. H. (1991) pioneered the concept of MARS, during a time when the need to generate more complex models was becoming apparent, and different approaches were needed. It attempts to smoothen out curves derived from “rough” data points without compromising on linearity. This can happen because each section, or “knot” is analyzed separately, with its own kind of equation particular to it. The greater the number of knots means a more significant number of separate equations and, thus a much greater deal of finesse in the curve and what it shows in the relationship between variables. The Basis Function, or BF, is the specific regression equation for each knot, while MARS is like a conglomerate of all BF. Equation 1 below is an illustration of how predictive variables and their subsequent effect on the application of MARS can be calculated by Maciej et al. (2019).

$$Y = \sum_{k=1}^{M} \beta_k BF_k(X) + \varepsilon \qquad [1]$$

In the above equation, the set of predictive variables is denoted by X, and the predicted variable is denoted by Y. BF_k for each linear knot has k as the number of its associated Basis Function. This, therefore, means that a regression equation for each linear section is generated, thus transforming a non-linear correlation structure into one which is linear. The total number of BF is denoted by M, while b_k relates to an estimation of the lowered Mean Square Error, or MSE.

Selection of an appropriate model, under the MARS approach, is carried out through the use of GCV, or Generalised Cross-Validation (9) as shown in Equation 2 below.

$$GCV = \frac{1}{N} \frac{\sum_{i}^{N}[Y_i - f_M(X_i)]^2}{\left[1 - \frac{C(M)}{N}\right]^2} \qquad [2]$$

In the above equation, the specific observation number is denoted by N for the particular data set in question. The Basis Function contains a number of squared residuals in the final model, M, whose sum is $f_M(X_i)$. These squared residuals are associated with a lack of fit, the sum of which is $\sum_{i}^{N}[Y_i - f_M(X_i)]^2$. It remains finally to denote the penalty term $\left[1 - \frac{C(M)}{N}\right]^{-2}$ which is applied to M number of Basis Functions. The reason for the introduction and application of a penalty term is to reduce the volume of Basis Functions, because a greater amount can affect model efficiency in certain scenarios. By calculating the lowest possible value for GCV, it is possible to generate the optimum model of MARS.

There are certain limitations to non-parametric methods, namely how to generate regression curves that are easily readable, as well as the generation of unbiased estimations based on both a solution approach and a split method Elith, J., Leathwick, J.R. and Hastie, T. (2008),. Accurate and easy interpretation, together with speed of calculation mean that

MARS is preferred even if it is dependent on the variable structure Xiaochen et al. (2022).

According to several studies Li-Yen Chang, Hsiu-Wen Wang (2006), and Diniz, M.A. (2022), it has been shown that MARS is of greater efficiency than other types of statistical analysis when looking at predicted variables and their associated predictive ones. It is, however, the case that MARS requires a data set of sufficient size for it to work to any degree of accuracy Kim, BS., Lee, YB. & Choi, DH. (2009).

It was hoped within this study to demonstrate how the application of MARS through a non-parametric approach could model variables where the structure was mixed and thus present clearly the extent to which Internet Addiction is an issue for both Iraq and Turkey. In taking this somewhat novel approach, it attempts to illustrate another way in which variables and their associated predictive value can ascertain levels of Internet dependency within the two cohorts. A comparative discussion, through the incorporation of classification techniques, will follow below.

Materials and Methods

Sample

The sample of the study is at the international level, and cosmopolitan universities from each country (Turkey and Iraq) and a university from mentioned countries were selected, which is Van Yüzüncü Yıl University from Turkey and Soran University from Iraq. The target audience was determined as university students. The fact that the sample was composed of adolescents is due to the fact that internet addiction is mainly seen in adolescents, and individuals in this period are open to all kinds of influences. The number of individuals in the sample is 2235; 1220 students from Van Yüzüncü Yıl University and 1015 from Soran University were assigned by random sampling method. The age range of the sample ranged from 11-67, and the mean age was found to be 21.66.

Data Collection Tool

The Internet Dependency Scale (IDS) is used as the source of the data set to which the MARS approach is applied, which is designed by Günüç and Kayri (2016) and the researcher with a total of 54 items to ask. It was semi- rather than fully structured in form. The idea was to look at the extent to which Internet addiction was a factor for the students at these universities. The study was divided into three separate sections; the first, it looked at the classic social and demographic backgrounds of the participants including the usual questions relating to socioeconomic status, age and gender. This comprised 5 closed-ended questions. The second addressed the effects of the Covid 19 pandemic; more specifically, how it had changed their use of the Internet by asking about length of time spent on it, how has use changed, whether they had it at home, on what sort of device they accessed it, and so on. This comprised 10 questions both closed and open in style. The third part incorporated a scale for Internet addiction, where four diagnostic sub-dimensions were explored.

Responses required grading according to a five-point Likert scale and the following attitudes are present: “I totally agree”, “I agree”, “I’m undecided”, “I don’t agree”, “I strongly disagree”. The scale items are scored from 5 to 1, with 5 points corresponding to the "I totally agree" degree and 1 point to the "I strongly disagree" degree. All items in the scale are oriented towards Internet dependency. The dependency level measurement tool has 35 elements, with a minimum score of 35 and a maximum score of 175 for each. Higher IDS test results indicate a stronger dependency. In addition, A two-step clustering approach was used to group different levels of addiction.

The Cronbach's alpha (α) internal consistency coefficient of the scale was found to be 0.957 for the Turkey sample and 00.936 for the Iraq sample. Exploratory factor analysis obtained by Günüç and Kayri (2010) regarding the construct validity of the scale, the scale consists of four sub-dimensions. In order to show that the results and structure of this study are valid, confirmatory factor analysis was applied to the obtained data. These four sub-dimensions are; It was named as “Deprivation”, “Control Difficulty”, “Distortion in Functioning” and “Social Isolation”.

Test for normality has been applied to the dataset by using Kolmogorov - Smirnov text, the test for both sample exhibit $P>0.05$, meaning that a normal distribution does not apply, a parametric approach to data analysis is inappropriate given the significance values in both data sets. Thus, a non-parametric approach needs to be taken. It should be noted, however that there is no requirement to test data set assumptions when employing a MARS approach.

The Cronbach's alpha (α) reliability coefficients for the four sub-dimensions were found to be 0.887 for Deprivation, 0.898 for Control Difficulty, 0.908 for Distortion in Functioning, and 0.878 for Social Isolation for the Turkish sample, whereas in other hands, 0.847 for Deprivation, 0.859 for Control Difficulty, 0.853 for Distortion in Functioning, and 0.813 for Social Isolation, for Iraq sample. IDS is asserted to be perfectly valid and reliable in the context of the provided data.

Data Analysis

According to Kayri (2007), generally, more cohesive and consistent results can be obtained where similarities in variables and participants are clustered accordingly. This helps in mitigating against the case where inconvenient outliers and bias in the data occur, where the sample is non-homogeneous. A clustering analysis consisting of two steps, combined with a logarithmic probability approach, can be most useful in cluster sampling, where it is possible to measure distance and similarity to a good degree of accuracy. Two-stage clustering analysis allows the data set to be divided into homogeneous subgroups. In other words, it aims to divide the heterogeneous data set into homogeneous subclasses or clusters. It is reported in the literature that statistical studies obtained from the clusters thus formed have healthier results (Kayri, 2007). After the two-stage cluster analysis, 2 clusters with a cut-off score of 93 points (including 93) for the Turkey sample and 88 points (including 88) for the Iraq sample were obtained. These clusters are named as “non-dependent” and “dependent”.

In this study, besides the descriptive statistics on some demographic variables, the relationships between these variables and internet addiction status were also examined, as result; there is a relationship between country and all dependent variables (Internet Addiction and its components; Deprivation, Difficulty in control, Distortion in functionality, Social isolation). In addition, Stage, Siblings Number, Smoke, Being infected with Covid19, Quality of online education during Covid19 do not have significant correlation with dependent variables. Moreover, gender has significant correlation with the variables Internet Addiction, Deprivation, Distortion in functionality, Social isolation. Furthermore, Age has significant correlation with the variables Internet Addiction, Deprivation, Difficulty in control, Distortion in functionality. Also, Mother and Father Education do not have significant correlation with dependent variables. In other hand, Father Job has significant correlation with the variables Internet Addiction, Deprivation, Distortion in functionality, Social isolation. Also, Mother Job and Family Income has significant correlation with Deprivation. As well as, Having Internet at Home has significant correlation with the variables Internet Addiction, Deprivation, Difficulty in control, Social isolation. In other respect, Purpose of using Internet has significant correlation with the variables Internet Addiction, Deprivation, Difficulty in control, Distortion in functionality, and also, Hours of using the Internet has a significant correlation with the variables Internet Addiction, Deprivation, Difficulty in control, Distortion in functionality, Social isolation. Moreover, Force of Using Internet During Covid19 has significant correlation with the variables Internet Addiction, Deprivation, Difficulty in control, Distortion in functionality. Furthermore, Electronic Device has significant correlation with the variables Internet Addiction, Difficulty in control, Distortion in functionality, Social isolation. Also, Hours of Using Internet During Covid19 has significant correlation with the variables Internet Addiction, Deprivation, Difficulty in control, Distortion in functionality, Social isolation. Finally, Hours of Using Internet before Covid19 has significant correlation with the variables Internet Addiction, Deprivation, Difficulty in control, Distortion in functionality, Social isolation.

As a result of the analysis made with MARS data mining method for the Iraq and Turkey datasets which has been shown in the confusion matrix table 1 and table 2, when students are classified as “addicted/ not-addicted” in the name of the Internet addiction, the number of students whom have been classified as dependent are 209 and 317 students for Turkey and Iraq respectively, in addition, the number of students whom have been classified as Not-dependent are 614 and 405 students for Turkey and Iraq respectively. As a consequence of these results, it can be state that the students from Iraq are more addicted to the Internet.

	Independent	Dependent	Total Number of Students
Independent	614	279	893
Dependent	118	209	327
Total Number of Students	732	488	1220

Table 1: Confusion Matrix Obtained from MARS Analysis (Turkey Sample)

	Non-Dependent	Dependent	Total Number of Students
Non-Dependent	405	137	543
Dependent	156	317	473
Total Number of Students	561	454	1015

Table 2: Confusion Matrix Obtained from MARS Analysis (Iraq Sample)

As a result of using the MARS analysis approach, model classification rates for accuracy, specificity, sensitivity, and precision, as well as F1-statistic values and AUC values of the ROC curve, are provided in table 3.

Criteria	Turkey	Iraq
Accurate Classification Rate	67.46%	71.13%
Specificity Ratio/ Specificity	63.91%	67.02%
Sensitivity Rate	68.76%	74.72%
Precision Rate	83.88%	72.19%
F1 Statistic	75.57%	73.44%
Area Under the ROC Curve (AUC)	71.28%	80.23%

Table 3, MARS Analysis Result Classification Performance Rates in Turkey and Iraq Sample

In terms of correct classification rate, MARS analysis method differs among both countries, while the correct classification rate obtained with the MARS analysis method for Turkey and Iraq is 67.46% and 71.13% respectively, it is seen that the sample from Iraq has a higher correct classification rate than the Turkey sample. Moreover, in terms of Specificity Rate, MARS analysis method differs among both countries, while the specificity rate obtained with the MARS analysis method for Turkey and Iraq is 63.91% and 67.02% respectively, it is seen that the sample from Iraq has a higher specificity rate than the Turkey sample. In addition, in terms of sensitivity rate, MARS analysis method differs among both countries, while the sensitivity rate obtained with the MARS analysis method for Turkey and Iraq is 68.76% and 74.72% respectively, it is seen that the sample from Iraq sample has a higher sensitivity rate than the Turkey sample. In terms of precision, MARS analysis method differs among both countries, while the precision rate obtained with the MARS analysis method for Turkey and Iraq is 83.88% and 72.19% respectively, it is seen that the sample from Turkey sample has a higher precision rate than the Iraq sample. In terms of F1–Statistic, the F1-statistic is a statistical measure obtained as a result of the harmonic average of precision and precision measures. MARS analysis method differs among both countries, while the F1-statistic obtained with the MARS analysis method for Turkey and Iraq is 75.57% and 73.44% respectively, it is seen that the sample from Turkey has a higher F1–Statistic rate than the Iraq sample. In other words, the F-1 statistical result for Turkey sample, which is the harmonic mean of MARS analysis method sensitivity and precision, showed a higher classification success. Furthermore, the MARS analysis method for Turkey sample was higher than the MARS analysis method for Iraq sample in identifying dependent and distinguishing independent category. In regard to the area under the ROC curve, MARS analysis method differs among both countries, while the area under the ROC curve obtained with the MARS analysis method for Turkey and Iraq is 71.28% and 80.23% respectively, it is seen that the sample from Iraq sample has a higher specificity rate than the Turkey sample. In other words, the MARS analysis method for Iraq sample classified dependent and independent category with fewer errors than the MARS analysis method Turkey Sample. As can be seen in Table 3, it was observed that there was no significant difference in the percentage levels considered.

Finding

Figures 1 and 2 illustrate data derived from the Generalized Cross Validation (GCV) model. As has been explained previously, the MARS process incorporates the use of algorithms that compute in both directions: backwards and forwards. This means that it is a reliable means of extracting data to the minimal degree of error possible, depending on the circumstances in which they were obtained.

Looking at the data presented in these figures, it can be seen that the basis function has been expressed according to a variable structure as well as taking into consideration the interactive and independent variables. With 20 regression equations for the Iraq data and 9 for the Turkish data, we can see that these exist at the point where there is the least error in the GCV value. Tables T and G show the results of the number of functions produced during data analysis.

As it has been shown le the Table 4, and Table 5, as a result, 0.179 and 0.205 are the lowest GCV values for Turkey and Iraq, which determined that 38 and 37 are the maximum number of basic functions for both datasets, Turkey and Iraq respectively.

Maximum Number of Basic Functions	GCV value	Maximum Number of Basic Functions	GCV value
35	0.18015	42	0.18015
36	0.18002	43	0.18022
37	0.18012	44	0.18027
38	0.17999	45	0.18033
39	0.18009	46	0.18025
40	0.18017	47	0.18030
41	0.18006	48	0.18032

Table 4: GCV Values for Determining the Maximum Number of Basic Functions

Maximum Number of Basic Functions	GCV value	Maximum Number of Basic Functions	GCV value
35	0.20480	42	0.20527
36	0.20510	43	0.20518
37	0.20467	44	0.20550
38	0.20496	45	0.20480
39	0.20499	46	0.20501
40	0.20536	47	0.20512
41	0.20559	48	0.20516

Table 5: GCV Values for Determining the Maximum Number of Basic Functions for Iraq sample

The MARS analysis method, that starts model building with 0 basic functions, reaches its most complex level with 36 basic functions. Then, the basic functions that did not contribute to the model were removed from the model with the backward pruning process, and the model using 8 predictors with 9 basic functions and GCV value of 0.17999 for Turkey, and 14 predictors with 20 basic functions and GCV value of 0.20467 for Iraq.

As a result of the variance analysis performed to define the relative contributions of each independent variable and the interactions between the variables, it was observed that the most appropriate model consisted of 8 and 15 functions for Turkey and Iraq correspondingly which it has been shown le figure 1 and 2.

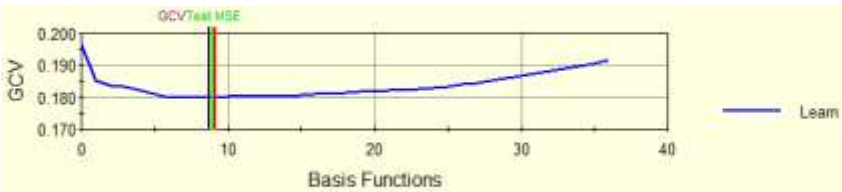

Figure 1. The number of basic functions indicated by the GCV value obtained as a result of the MARS analysis of Turkey sample.

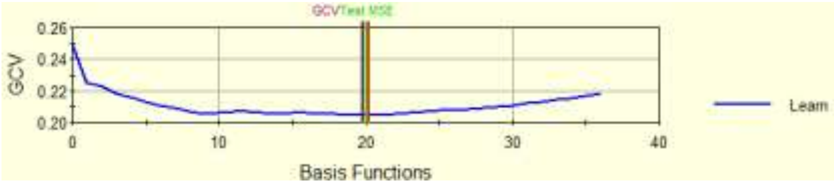

Figure 1. The number of basic functions indicated by the GCV value obtained as a result of the MARS analysis of Iraq Sample

The contribution of the independent variables in the model of the variance analysis graph created for the most appropriate model was calculated by estimating the adjusted R2 of the model as a result of excluding the ANOVA function at each step. The variance analysis information regarding these graphically illustrated functions is given in Table 6 and 7 for Turkey and Iraq sample.

Function	Standard Deviation	Cost of omission	No of basis Functions	No of Effective Parameters	Variables
1	0.14136	0.18308	2	5,333	NUMBER OF SIBLINGS, DAILY INTERNET USE
2	0.06553	0.18210	1	2,667	MOTHER EDUCATION, DAILY INTERNET USE
3	0.09935	0.18177	1	2,667	AGE, DAILY INTERNET USE
4	0.05405	0.18210	1	2,667	MOTHER EDUCATION, SMOKING
5	0.10707	0.18168	1	2,667	DAILY INTERNET USE, DAILY INTERNET USE IN COVID19 QUARANTINE
6	0.03038	0.18001	1	2,667	DAILY INTERNET USE, QUALITY OF ONLINE EDUCATION IN TH COVID19 EPIDEMIC
7	0.03011	0.18011	1	2,667	DAILY INTERNET USE, DEVICE PREFERENCE FOR ONLINE IN QUARANTINE
8	0.04381	0.18113	1	2,667	SMOKING, THE QUALITY OF ONLINE EDUCATION IN THE COVII EPIDEMIC

Table 6. Variance Analysis for the Optimal Model of the Turkey sample.

Function	Standard Deviation	Cost of duty	No of basis Functions	No of Effective Parameters	Variables
1	0.08903	0.20538	1	2,750	DAILYINTERNETUSAGE
2	0.06707	0.20693	1	2,750	FAMILYINCOME, DAILYINTERNETUSAGE
3	0.05464	0.20584	1	2,750	INTERNETUSAGEPURPOSE, DAILYINTERNETUSAGE
4	0.1503	0.20693	2	5,500	DAILYINTERNETUSAGE, DAILYINTERNETUSAGEINCOVID19QUARANTI
5	0.12145	0.20921	2	5,500	INTERNETUSAGECOVID19PANDEMIC, QUALITYONLINEEDUCATIONCOVID19PANDEMIC
6	0.05282	0.20631	1	2,750	AGE, DAILYINTERNETUSAGE
7	0.07625	0.20544	2	5,500	MOTHER'SEDUCATION, INTERNETUSAGECOVID19PANDEMIC
8	0.03528	0.20483	1	2,750	DAILYINTERNETUSAGE, DEVICEPREFERENCEFORONLINEINQUARANT
9	0.08515	0.20734	2	5,500	COVID19PREQUARANTINEDAILYINTERNETUSAGE, QUALITYONLINEEDUCATIONCOVID19PANDEMIC
10	0.03457	0.20483	1	2,750	NUMBEROFSIBLINGS, COVID19PREQUARANTINEDAILYINTERNETUSA
11	0.04430	0.20538	1	2,750	INTERNETHOME, INTERNETUSAGECOVID19PANDEMIC
12	0.06437	0.20503	1	2,750	INTERNETUSAGEPURPOSE, INTERNETUSAGECOVID19PANDEMIC
13	0.07758	0.20531	1	2,750	FATHERSJOB, DAILYINTERNETUSAGE
14	0.16633	0.20765	1	2,750	MOTHERSJOB, DAILYINTERNETUSAGE
15	0.1193	0.20541	2	5,500	NUMBEROFSIBLINGS, DAILYINTERNETUSAGE

Table 7. Variance Analysis for the Optimal Model of the Iraq sample.

In both tables, the predictor variables in the functions and the number of basic functions and the values indicating the loss that will occur if the predictor variables are removed from the model are given. Moreover, as can be seen from the table, it is seen that some predictor variables are included in the model with more than one basic function alone and combined with more than one variable and included in the model with a single basic function. The most suitable model obtained as a result of the forward step and backward step applications of the MARS analysis method is called the final model. Table 8,9 shows the table information about the final model.

Basis Function	coefficients	Variable	Sign in	Parent Sign	parent	knot
0	0.1103					
4	-0.0350	Number of siblings	+	-	DAILY INTERNET USAGE	11.00
5	-0.0062	Number of siblings	+	-	DAILY INTERNET USAGE	11.00
6	0.0116	MOTHER EDUCATION	-	-	DAILY INTERNET USAGE	3.000
9	0.0042	AGE	+	+	DAILY INTERNET USAGE	30,00
11th	0.0767	MOTHER EDUCATION	-	+	SMOKING	5,000
12	0.0085	DAILY INTERNET USAGE DURING COVID19 QUARANTINE	+	+	DAILY INTERNET USAGE	1,000
13	0.0062	THE QUALITY OF ONLINEEDUCATION IN THE COVID19 OUTPUT	-	+	DAILY INTERNET USAGE	6.000
15	-0.0453	IN QUARANTINEONLINETO THE DEVICE CHOICE	+	+	DAILY INTERNET USAGE	3.000
18	0.0565	THE QUALITY OF ONLINEEDUCATION IN THE COVID19 OUTPUT	-	+	SMOKING	4.000

Table 8, Final Model Created for the Most Appropriate Model Turkey Sample

Basis Function	coefficients	Variable	Sign in	Parent Sign	parent	knots
0	0.2830					
1	0.0365	DAILYINTERNETUSAGE	+			3.0000
4	0.2705	FAMILYINCOME	-	-	DAILYINTERNETUSAGE	2,0000
7	-0.0167	INTERNETUSAGEPURPOSE	+	+	DAILYINTERNETUSAGE	3.0000
8	0.0176	DAILYINTERNETUSAGEINCOVID19QUARANTINE	-	+	DAILYINTERNETUSAGE	4.0000
9	0.0359	DAILYINTERNETUSAGEINCOVID19QUARANTINE	+	-	DAILYINTERNETUSAGE	4.0000
10	-0.0573	QUALITYONLINEEDUCATIONCOVID19PANDEMIC	+	+	INTERNETUSAGECOVID19PANDEMIC	2,0000
11th	-0.3137	QUALITYONLINEEDUCATIONCOVID19PANDEMIC	-	+	INTERNETUSAGECOVID19PANDEMIC	2,0000
12	0.0157	AGE	+	+	DAILYINTERNETUSAGE	22,0000
14	0.8877	MOTHER'SEDUCATION	-	+	INTERNETUSAGECOVID19PANDEMIC	6.0000
15	0.0362	MOTHER'SEDUCATION	+	+	INTERNETUSAGECOVID19PANDEMIC	6.0000
16	-0.0451	DEVICEPREFERENCEFORONLINEINQUARANTINE	-	+	DAILYINTERNETUSAGE	3.0000
19	0.0097	QUALITYONLINEEDUCATIONCOVID19PANDEMIC	+	+	COVID19PREQUARANTINEDAILYINTERNETUSAGE	2,0000
20	0.0689	QUALITYONLINEEDUCATIONCOVID19PANDEMIC	-	+	COVID19PREQUARANTINEDAILYINTERNETUSAGE	2,0000
23	0.0521	NUMBEROFSIBLINGS	+	+	COVID19PREQUARANTINEDAILYINTERNETUSAGE	10,0000
25	0.2395	INTERNETHOME	-	+	INTERNETUSAGECOVID19PANDEMIC	1,0000
27	-0.0360	INTERNETUSAGEPURPOSE	+	+	INTERNETUSAGECOVID19PANDEMIC	6.0000
28	-0.0669	FATHERSJOB	-	+	DAILYINTERNETUSAGE	1,0000
29	0.0575	MOTHERSJOB	+	-	DAILYINTERNETUSAGE	1,0000
30	0.2636	NUMBEROFSIBLINGS	+	+	DAILYINTERNETUSAGE	5,0000
32	-0.2517	NUMBEROFSIBLINGS	-	+	DAILYINTERNETUSAGE	4.0000

Table 9, Final Model Created for the Most Appropriate Model of the Iraq sample

According to fundamental equations of functions for the optimal model which has been shown in the table 10 and 11

```
BASIC FUNCTION
BF1 = max (0, DAILY INTERNET USAGE - 1);
BF2 = max (0, SMOKING - 2);
BF3 = max (0, 2 - SMOKING);
BF4 = max (0, NUMBER - 11) * BF1;
BF5 = max (0, 11 - NUMBER OF SISTER ) * BF1;
BF6 = max (0, MOTHEREDUCATION - 3) * BF1;
BF9 = max (0, 30 - AGE) * BF1;
BF11 = max (0.5 - MOTHEREDUCATION) * BF2;
BF12 = max (0, DAILY INTERNET USAGE DURING COVID19 QUARANTINE - 1) * BF1;
BF13 = max (0, QUALITY OF ONLINEEDUCATION IN THE COVID19 OUTPUT - 6) * BF1;
BF15 = max (0, DEPARTMENT CHOICE FOR ONLINE IN QUARANTINE - 3) * BF1;
BF18 = max (0.4 - QUALITY OF ONLINE EDUCATION IN THE COVID19 OUTPUT) * BF3;
```

Table 10. Fundamental Equations of Functions for the Optimal Model of Turkey sample

```
BASIC FUNCTION
BF1 = max (0, DAILYINTERNETUSAGE - 3);
BF2 = max (0, 3 - DAILYINTERNETUSAGE);
BF4 = max (0, 2 - FAMILY INCOME) * BF2;
BF5 = max (0, INTERNETUSAGECOVID19PANDEMIC - 1);
BF7 = max (0, 3 - INTERNETUSAGEPURPOSE) * BF1;
BF8 = max (0, DAILYINTERNETUSAGEINCOVID19QUARANTINE - 4) * BF1;
BF9 = max (0.4 - DAILYINTERNETUSAGEINCOVID19QUARANTINE) * BF1;
BF10 = max (0, QUALITYONLINEEDUCATIONCOVID19PANDEMIC - 2) * BF5;
BF11 = max (0, 2 - QUALITYONLINEEDUCATIONCOVID19PANDEMIC) * BF5;
BF12 = max (0, AGE - 22) * BF1;
BF14 = max (0, MOTHER'SEDUCATION - 6) * BF5;
BF15 = max (0.6 - MOTHER'SEDUCATION) * BF5;
BF16 = max (0, DEVICEPREFERENCEFORONLINEINQUARANTINE - 3) * BF1;
BF18 = max (0, COVID19PREQUARANTINEDAILYINTERNETUSAGE - 1);
BF19 = max (0, QUALITYONLINEEDUCATIONCOVID19PANDEMIC - 2) * BF18;
BF20 = max (0, 2 - QUALITYONLINEEDUCATIONCOVID19PANDEMIC) * BF18;
BF23 = max (0, NUMBEROFSIBLINGS - 10) * BF18;
BF25 = max (0, INTERNETHOME - 1) * BF5;
BF27 = max (0.6 - INTERNETUSAGEPURPOSE) * BF5;
BF28 = max (0, FATHERSJOB - 1) * BF2;
BF29 = max (0, MOTHERSJOB - 1) * BF2;
BF30 = max (0, NUMBEROFSIBLINGS - 5) * BF2;
BF32 = max (0, NUMBEROFSIBLINGS - 4) * BF2;
```

Table 11. Fundamental Equations of Functions for the Optimal Model of Iraq sample

the maximum number of basic functions determined based on the GCV value in the MARS analysis method was determined as 18. Then, 9 of the 18 basic functions were used to establish the most appropriate model for Turkey sample, Table 12. Conversely, the maximum number of basic functions of Iraq sample determined based on the GCV value in the MARS analysis method was determined as 32. Then, 20 of the 32 basic functions were used to establish the most appropriate model for Iraq sample, Table 13. The basic functions BF1, BF2 and BF3 from Turkey sample, and the basic functions BF2, BF5 and BF18 from Iraq sample, were used in the creation of other basic functions and were not included in the final model.

```
MODEL INTERNETADDICTIONCATEGORY = BF4 BF5 BF6 BF9 BF11 BF12

BF13 BF15 BF18;
```

Table 12, Internet Addiction Model Table for the Most Appropriate Model of Turkey sample

```
MODEL INTERNETADDICTIONCATEGORIC = BF1 BF4 BF7 BF8 BF9 BF10 BF11 BF12 BF14 BF15 BF16
BF19 BF20 BF23 BF25 BF27 BF28 BF29 BF30  BF32;
```

Table 13, Internet Addiction Model Table for the Most Appropriate Model of Iraq Sample

Regression equation for optimal model table 14 and 15, the F and p-values calculated as F=18.29641 and F=17.98916 for Turkey and Iraq, the p-value was calculated for both of them as p = 0.0000, this indicate that the models are meaningful.

```
Y = 0.11026 - 0.0350244 * BF4 - 0.00620594 * BF5
+ 0.0116034 * BF6 + 0.00422424 * BF9
+ 0.0766942 * BF11 + 0.00848087 * BF12
+ 0.0061938 * BF13 - 0.0453293 * BF15
+ 0.0564568 * BF18;
```

Table 14: Regression Equation for Optimal Model of Turkey sample

```
Y = 0.11026 - 0.0350244 * BF4 - 0.00620594 * BF5
+ 0.0116034 * BF6 + 0.00422424 * BF9
+ 0.0766942 * BF11 + 0.00848087 * BF12
+ 0.0061938 * BF13 - 0.0453293 * BF15
+ 0.0564568 * BF18;
```

Table 15: Regression Equation for Optimal Model of Turkey sample

The data obtained with the Internet Addiction Scale for Turkey and Iraq samples were analyzed with the MARS analysis method and the most important predictors of Internet Addiction category were determined. The predictor variables included in the analysis and the significance levels of these variables on the variable predicted within themselves in the established model

The most important predictors obtained on Turkey sample as a result of the analysis made with the MARS analysis method are given in Table 16. Accordingly, the most important predictors of Internet Addiction Scale for Turkey sample are respectively; Daily Internet Usage, Mothers' Education, Number of Siblings, Smoking, Age, Daily Internet Usage During Covid-19 Quarantine, The Quality of Online education in The Covid-19 Pandemic, the device used, The variables associated with the dependent variable in the MARS analysis method were ranked according to their importance, starting from 100 points. The variable with the highest relationship with internet addiction scale was the self-efficacy perception variable, while the variable with the lowest relationship was the type of device to use the Internet. Variables that have little or no relationship with internet addiction scale and are not included in the analysis are respectively; Daily use of the Internet before COVID-19 pandemic, Fatherhood, Mother's profession, Gender, Fathers' education, Family income, Covid-19 pandemic and internet use, have you effected by Covid-19, do you have Internet at home, Internet use purpose, Mothers' Profession, and Stage of study were obtained as insignificant.

Variable	score	
DAILY INTERNET USAGE	100.00	
MOTHEREDUCATION	50.82	
Number of siblings	47.58	
SMOKING	45.52	
AGE	36.07	
DAILY INTERNET USAGE DURING COVID19 QUARANTINE	35.18	
THE QUALITY OF ONLINEEDUCATION IN THE COVID19 OUTPUT	25.74	
IN QUARANTINEONLINETO THE DEVICE CHOICE	9.47	
DAILY USE OF THE INTERNET BEFORE COVID19QUARANT	0.00	
FATHERHOOD	0.00	
MOTHER'S PROFESSION	0.00	
GENDER	0.00	
FATHEREDUCATION	0.00	
FAMILY INCOME	0.00	
COVID19PANDEMIDEINTERNETUSE	0.00	
Have you had COVID19	0.00	
HOME INTERNET	0.00	
INTERNETUSEPURPOSE	0.00	
MOTHER'S PROFESSION_mis	0.00	
Class	0.00	

Table 16. MARS Analysis Method Table of Significance Levels of Variables of Turkey Sample

The most important predictors obtained on Iraq sample as a result of the analysis made with the MARS analysis method for Iraq sample are given in Table 17. Accordingly, the most important predictors of Internet Addiction Scale for Iraq sample are respectively; Daily Internet Usage, Internet Usage Covid-19 Pandemic, Quality Online Education Covid-19 Pandemic, Covid-19 Pre-Quarantine Daily Internet Usage, Mothers' Job, internet Usage Purpose, Family Income, Daily Internet Usage in Covid-19 quarantine, Age, Number of Siblings, Mothers' Eduation, Internet Home, Fathers' Job, Device Preference for Online in Quarantine. The variables associated with the dependent variable in the MARS analysis method were ranked according to their importance, starting from 100 points. The variable with the highest relationship with internet addiction scale was the self-efficacy perception variable, while the variable with the lowest relationship was Device Preference for Online in Quarantine. Variables that have little or no relationship with internet addiction scale and are not included in the analysis are respectively; Gender, Fathers' Education, Have You Had Covid-19, Do You Smoke and Stage of study were obtained as insignificant.

Variable	score	
DAILY INTERNET USAGE	100.00	
MOTHEREDUCATION	50.82	
Number of siblings	47.58	
SMOKING	45.52	
AGE	36.07	
DAILY INTERNET USAGE DURING COVID19 QUARANTINE	35.18	
THE QUALITY OF ONLINEEDUCATION IN THE COVID19 OUTPUT	25.74	
IN QUARANTINEONLINETO THE DEVICE CHOICE	9.47	
DAILY USE OF THE INTERNET BEFORE COVID19QUARANT	0.00	
FATHERHOOD	0.00	
MOTHER'S PROFESSION	0.00	
GENDER	0.00	
FATHEREDUCATION	0.00	
FAMILY INCOME	0.00	
COVID19PANDEMIDEINTERNETUSE	0.00	
Have you had COVID19	0.00	
HOME INTERNET	0.00	
INTERNETUSEPURPOSE	0.00	
MOTHER'S PROFESSION_mis	0.00	
Class	0.00	

Table 17. MARS Analysis Method Table of Significance Levels of Variables of Iraq Sample

Discussion

Daily Internet Usage may be seen as the primary initiating factor for Internet dependence for Turkish and Iraqi students following the research model identified by MARS for both datasets. These results have been supported in the literature, which eventually resulted in the conclusion that using the Internet leads to Internet dependence. In addition, the current study, according to Kayri (2010), demonstrated that both the average amount of time spent online and the anticipated usage of the Internet were significant indicators of internet dependence.

Because MARS takes the novel approach in that discrete basis functions are obtained in respect of a particular knot – or linear area – and thus examines these regions accordingly, it is not a traditional "parametric" statistical regression tool. The advantage is therefore that all basis functions can be included, forming part of the outcome as a whole, and is thus not a "generalization" but a much more accurate means of analyzing data.

With the present model of MARS drawing on nineteen predicted variables forming the scale of Internet Dependency, as well as the dependent variable, one aim of this study was to demonstrate how MARS can be employed for such studies as well as epidemiology as a whole. The predicted variable of $P<0.5$ is a key takeaway and shows how it is affected by the various predictors such as Daily Internet Usage, Mother education, Number of Sibilings, Age, Daily internet Usage During COVID-19. It is shown, too, that magnitude of each variable may impact the model, but is nevertheless a prime advantage of MARS as against other methods.

Recommendations arising from the present study are that MARS is a tool of great significance for work in the field of epidemiology, as has been shown by the positive outcomes in terms of statistical analysis. Because it is able to generate predictions relatively free of bias due to the lack of need for a continuous/discrete scale, one can say that MARS does so by drawing on and analyzing data at a much higher level of granularity than other statistical methods.

Findings from this study have important clinical implications in general and university students specially. Ministry of higher education, mental health authorities, should be aware of these high rates of internet addiction among the university students in both countries, and should create awareness among students regarding internet addiction and its potential harms; this could be included in foundation course of curriculum implementation support program for students. We Initiative should be taken to provide ample opportunities for students to involve in extracurricular activities and interact with friends. There should be provision of counsellor for emotional and mental support of students as they are overburden with studies and long posting schedules. Measures should be undertaken to prevent further increase in rates and manage the possible cases.

Reference:

Axel W, Hardle W, Müler M, Sperlich S (2004). Non-parametric and Semiparametric Models. Springer Series in Statistics: Springer Press, Germany.

Diniz, M.A. (2022) Statistical methods for validation of predictive models. J. Nucl. Cardiol. 29, 3248–3255. https://doi.org/10.1007/s12350-022-02994-7

Elith, J., Leathwick, J.R. and Hastie, T. (2008), A working guide to boosted regression trees. Journal of Animal Ecology, 77: 802-813. https://doi.org/10.1111/j.1365-2656.2008.01390.x

Fabio Trojani (2006) Semiparametric Regression for the Applied Econometrician, Journal of the American Statistical Association, 101:473, 397-398, DOI: 10.1198/jasa.2006.s78.

Friedman, J. H. (1991). Multivariate Adaptive Regression Splines. The Annals of Statistics, 19(1), 1–67. http://www.jstor.org/stable/2241837.

Günüç, S., Kayri, M., (2016). Comparing Internet Addiction in Students with High and Low Socioeconomic Status Levels.

Hébert-Losier, K., & Beaven, C. M. (2014). The MARS for squat, countermovement, and standing long jump performance analyses: are measures reproducible?. Journal of strength and conditioning research, 28(7), 1849–1857. https://doi.org/10.1519/JSC.0000000000000343.

Kayri M (2007). Two-step clustering analysis in researches: A case study. Eurasian Journal of Educational Research, 7(28): 89- 99.

Kayri M, Zırhlıoğlu G (2009). Kernel smoothing function and choosing bandwidth for non-parametric regression methods. Ozean Journal of Applied Science, 2: 49-54.

Kim, BS., Lee, YB. & Choi, DH. (2009) Comparison study on the accuracy of metamodeling technique for non-convex functions. J Mech Sci Technol 23, 1175–1181. https://doi.org/10.1007/s12206-008-1201-3

Li-Yen Chang, Hsiu-Wen Wang (2006) Analysis of traffic injury severity: An application of non-parametric classification tree techniques, Graduate Institute of Transportation and Logistics, National Chia-Yi University, 300 University Road, Chia-Yi 60004, Taiwan.

Maciej Przybyłek, Tomasz Jeliński, Julia Słabuszewska, Dorota Ziółkowska, Karina Mroczyńska, and Piotr Cysewski (2019) Application of Multivariate Adaptive Regression Splines (MARSplines) Methodology for Screening of Dicarboxylic Acid Cocrystal Using 1D and 2D Molecular Descriptors. Crystal Growth & Design. 19 (7), 3876-3887 DOI: 10.1021/acs.cgd.9b00318.

Nalcaci, G., Özmen, A. & Weber, G.W. (2019) Long-term load forecasting: models based on MARS, ANN and LR methods. Cent Eur J Oper Res 27, 1033–1049. https://doi.org/10.1007/s10100-018-0531-1.

Xiaochen Zhang, Kuangnan Fang & Qingzhao Zhang (2022) Multivariate functional generalized additive models, Journal of Statistical Computation and Simulation, 92:4, 875-893, DOI: 10.1080/00949655.2021.1979550