applied sciences

MDPI

Article

Predicting Phubbing Through Machine Learning: A Study of Internet Usage and Health Risks

Ayşen Yalman [1], Mehmet Arif Arık [2,*], Mehmet Kayakuş [3,*], Murad Karaduman [2], Sibel Karaduman [4], Fatma Yiğit Açıkgöz [5], Tuba Livberber [2] and Fahrettin Kayan [5]

[1] Department of Marketing and Advertising, Serik Gulsun-Suleyman Sural Vocational School, Akdeniz University, Antalya 07500, Türkiye; aysenyalman@akdeniz.edu.tr
[2] Department of Journalism, Faculty of Communication, Akdeniz University, Antalya 07058, Türkiye; mkaraduman@akdeniz.edu.tr (M.K.); tubalivberber@akdeniz.edu.tr (T.L.)
[3] Department of Management Information Systems, Faculty of Social and Human Sciences, Akdeniz University, Antalya 07600, Türkiye
[4] Department of Radio, Tv and Cinema, Faculty of Communication, Akdeniz University, Antalya 07058, Türkiye; sibelkaraduman@akdeniz.edu.tr
[5] Department of Marketing and Advertising, Social Sciences Vocational School, Akdeniz University, Antalya 07030, Türkiye; fatmayigit@akdeniz.edu.tr (F.Y.A.); kayan@akdeniz.edu.tr (F.K.)
* Correspondence: m.arif.arik@gmail.com (M.A.A.); mehmetkayakus@akdeniz.edu.tr (M.K.)

Abstract: Phubbing, defined as the disruption of social relationships and interactions due to excessive cell phone use, is becoming an increasing concern in modern society. Since one of the primary motivations for cell phone use is internet access, it is crucial to assess the time that individuals spend online to measure the prevalence of phubbing as a social behaviour disorder. This study aimed to better understand and evaluate the phubbing phenomenon by predicting future trends in internet usage using machine learning techniques. Four machine learning models—an artificial neural network (ANN), support vector regression (SVR), random forest (RF) regression, and time series—were employed to predict the average internet usage. Data from 2014 to 2024 were obtained from the World Bank, and cross-validation was used to enhance the reliability and accuracy of the models. All four models were successful in predicting internet usage, with the ANN showing the highest accuracy, followed by SVR, RF, and the time series. According to the data, the average daily time spent online increased from 277 min in 2014 to 417 min in 2024. Projections based on these machine learning models estimate that this figure will rise to 507 min by 2030 and 603 min by 2035. These findings provide valuable insights into the potential risks of increased phubbing behaviours on social interactions and offer a foundation for the exploration of the long-term health implications of excessive internet use. Future research could further examine the effects of phubbing on mental health and develop strategies to mitigate this social behaviour disorder.

Keywords: phubbing; health; internet use; addiction; machine learning; communication

check for updates

Academic Editors: Diego R. Faria and Jing Jin

Received: 14 October 2024
Revised: 27 December 2024
Accepted: 22 January 2025
Published: 24 January 2025

Citation: Yalman, A.; Arık, M.A.; Kayakuş, M.; Karaduman, M.; Karaduman, S.; Yiğit Açıkgöz, F.; Livberber, T.; Kayan, F. Predicting Phubbing Through Machine Learning: A Study of Internet Usage and Health Risks. *Appl. Sci.* **2025**, *15*, 1157. https://doi.org/10.3390/app15031157

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The concept of phubbing first emerged in the early 2010s and has rapidly become widespread with digitalisation in modern society. In particular, the fact that mobile devices are at the centre of our daily lives makes it important to understand the effects of this behaviour on individuals and communities. However, there are not enough studies on the long-term social and psychological effects of phubbing in the existing literature. This study aims to fill this gap in the literature, to predict the potential future risks of phubbing, and

to raise public awareness. The findings of this study will enable the development of new strategic approaches for both academic circles and policymakers.

Phubbing refers to individuals neglecting their social ties by focussing on their mobile phones instead of face-to-face interaction with the people around them. This behaviour, which is becoming increasingly common today, can have negative effects on both psychological and physical health. Phubbing can lead to feelings of insecurity, alienation, and loneliness in interpersonal relationships, while increasing stress and anxiety levels due to a lack of social support. In addition, physical health problems such as sleep disorders, eye strain, and posture disorders may also occur due to constant phone use. Therefore, it is clear that phubbing negatively affects both social relationships and the general health of individuals; therefore, it is an important issue that needs attention in modern society.

In parallel with the changes in technology, communication tools and forms have also changed, and the internet and digital-based communication tools have become a constant part of daily and social life [1–4]. In particular, portable phones, which provide instant interaction independently of space and time and are constantly connected to the online network, have a strong impact on people's social relationships and daily lives [5,6]. The main reason for this impact is that individuals increasingly use the internet to communicate with friends, family, or colleagues or meet strangers [7], to have fun, to engage in hobbies and play games [8], or to obtain information, learn, and research [9].

Several studies have utilised advanced machine learning techniques for time series forecasting [10,11]. However, few have applied these methods to understand the socio-behavioural implications of internet usage, as addressed in this study. Our research builds on these works by leveraging machine learning models to explore the long-term effects of digital behaviours like phubbing.

Individuals tend to prefer smartphones to computers for internet use [12]. In China, mobile phones are a prevalent tool for internet access, with over 97% of online users relying on them to connect to the web. On the other hand, in Turkey, 73.5% of internet traffic is generated through mobile devices [13]. Although these high usage rates are mainly due to the obvious benefits of smartphones, such as comfort and convenience, especially in recent years, researchers and scientists have begun to examine the harm that they cause to mental and physical health. In addition to this harm, there has been growing concern about the potential negative effects of smartphones on the quality of social interactions and relationships [14–17]. One of the negative effects that arises due to intensive smartphone use is the concept of "phubbing", the act of focusing on the smartphone instead of the person being interacted with in a disrespectful and condescending manner during social interaction. Phubbing is a term derived by combining the terms "phone" and "snubbing", while "being phubbed" means being ignored and belittled by someone who uses their cell phone while in the company of others [18]. The negative consequences for social relationships are broad and deep, including romantic partners [18] or parents [17]. In this respect, it is of great importance to consider the concept of phubbing as an important addiction and to consider the extent of the negative effects that it may cause in the coming years.

This research aims to reveal the dimensions of the danger of "phubbing" by estimating the duration spent on the internet in 2030 and 2035, considering the correlation between internet use and smartphone use.

In the Introduction of this study, the concept of "phubbing", related to excessive phone use, is discussed. The Section 2 of the study reviews similar works in the literature and provides an analysis of these studies. The Section 3, which is the Materials and Methods, describes the dataset and analysis techniques used. In the Section 4, the Findings and Discussion, the parameters used in the methods, the error and success metrics, and the results from the analysis,

are presented. Finally, in the Section 6, the findings are interpreted, highlighting the unique, superior, and improvable aspects of the study, and its impact on the literature is discussed.

This study provides a novel contribution by integrating phubbing analysis with machine learning-based time series forecasting. Unlike prior studies that predominantly used survey or traditional statistical methods, our work employs AI-based prediction models to estimate future internet usage patterns, directly addressing the digital behaviours that contribute to phubbing. The inclusion of diverse independent variables (e.g., urban–rural distribution, literacy rates, mobile traffic share) enriches the analysis, offering a comprehensive perspective on this social phenomenon.

2. Literature

Phone addiction, which has become a "disease" in the modern world, has given rise to the concept of phubbing. The term "phubbing", or "phone preoccupation", has become a subject of increasing concern in recent years and is a phenomenon that can have serious effects on our social relationships and psychological well-being. Phubbing refers to when a person constantly checks or uses their phone while interacting with others. This action leads to a lack of attention to the other person and communication breakdowns. Recent studies reveal that phubbing is an important problem in relationships and social connections [12].

Research shows that the addictive act of looking at the phone screen has negative effects on individuals' social, physical, and psychological conditions. In addition to research revealing that phubbing reduces trust between people [18], there are studies indicating that the quality of communication decreases in such an environment [19–23]. Therefore, communication not only becomes mundane but the quality of the interaction is potentially reduced, threatening the structure of the relationship [24]. When the research studies were examined, it was found that these studies mostly focus on adults, namely university students and adolescents [25]. Although the number of studies in this field seems to be quantitatively high in the literature, the studies conducted in this field are traditional studies that specifically address the impact of phubbing on individuals [26,27].

On the other hand, in addition to these traditional studies, there are also studies using machine learning methods. For example, the study by Rahman et al. [9] adopts an approach based on data mining and machine learning to identify patterns related to phishing behaviour. In their study, various models were developed based on the collected data. The study found that the developed addiction measures failed to accurately predict phishing. Moreover, as highlighted by the much higher explained variance, the machine learning approach is considered more suitable for predicting pubbing than traditional linear statistical methods. This study also emphasises that phubbing cannot be solely associated with addiction dynamics (such as age, social anxiety, and loneliness) and that modelling procedures that can account for non-linearity are needed to accurately assess users' phubbing levels.

In Mahmud et al. [28], a machine learning model is introduced, offering a departure from conventional analysis techniques. This research explores the potential conflict arising from self-phubbing and partner phubbing by constructing a model that connects these behaviours. The aim is to understand how self- and partner phubbing influence relationship satisfaction and, in turn, impact an individual's overall life satisfaction.

Chotpitayasunondh and Douglas [12] investigated the factors that contribute to phubbing behaviour and delved into how individuals reshape social communication norms in response to changes in their own behaviour and the behaviour of others around them. Their study investigated how smartphone addiction can lead to phubbing behaviour and the potential effects of gender and various variables, and they developed a data collection and prediction model through Qualtrics, a software program. Thus, at the end of this artificial intelligence-supported study, it was found that phubbing is a direct consequence of problem-

atic smartphone use. However, it is noted that artificial intelligence was not completely used in the analysis in this study. In addition to this, the quantitative method of the questionnaire was applied in this study. Most of the data were obtained through a questionnaire. Therefore, it is not possible to state that this study was conducted entirely using artificial intelligence.

In Liyanage [11], the effect of a fear of missing out (FOMO) as a determinant of phubbing behaviour is investigated by machine learning methods. This study aimed at determining the FOMO effect, which was selected as a physiological variable in modelling problematic social media and cell phone use. The study provides an in-depth and comprehensive analysis of the severity of problematic social media use, using a wider set of prediction factors than previously used, as well as more sophisticated methods utilising machine learning. The study concludes that the FOMO effect is identified as the most important contributor among other factors.

Nadeem [22] developed a machine learning model that can accurately identify potential cases of depression, feelings of loneliness, and various negative conditions underlying problematic social media use.

The research introduces a machine learning framework called Social Network Mental Disorder Detection, which leverages features derived from social network data. Tested on 3126 social media users, the model is reported to be able to identify stressed users on social media platforms.

Although phubbing has become an area of increasing interest in recent years, research on the social antecedents and effects of phubbing is extremely limited. In addition, there is a limited number of studies that include a research model supported by artificial intelligence [10]. However, AI-based studies allow the analysis of larger datasets and the discovery of more complex relationships. Through these methods, the factors underlying phubbing behaviour can be understood more deeply, and predictions can be made about future changes.

In the articles on phubbing, it is seen that the number of studies using machine learning methods is quite low. In addition to this situation, artificial intelligence techniques are extremely important in the field of social sciences in terms of working with large datasets [28].

3. Materials and Methods

This study employed machine learning techniques to explore the phenomenon of phubbing, setting it apart from traditional research methods. By analysing large datasets, machine learning methods can help us to understand how phubbing behaviour is associated with different variables and how this behaviour may develop in the future. There are 5.35 billion internet users worldwide, and 96.5% of this number use mobile phones to access the internet. Most of the total internet traffic is also carried out by smartphones [11]. In this direction, it would be meaningful to use the correlation between internet usage and mobile phones to reveal the future status of the phishing concept, which is a phenomenon caused by excessive phone usage.

This study aimed to apply machine learning-based analyses, offering a fresh perspective on academic research regarding phubbing and facilitating a more comprehensive examination of the phenomenon.

Besides the four models used in this study (ANN, RF, SVR, and time series), comparisons with more advanced time series forecasting methods, such as LSTM and XGBoost, which are frequently suggested in the literature, have not been made. This shortcoming is recognised as a limitation regarding a more comprehensive evaluation of the results. Future studies can improve the accuracy and reliability of time series analyses by including these methods in model comparisons.

3.1. Dataset

This study utilised machine learning models to predict the average time spent online, measured in minutes, by employing sixteen independent variables. The dependent variable was the average internet usage time for the period from 2014 to 2024, as sourced from the World Bank. The analysis incorporated the following independent variables to explore these relationships:

- Total population;
- Number of active internet users;
- Number of active social media users;
- Cellular mobile network users;
- Proportion of women by population;
- Proportion of men by population;
- Proportion of population living in rural areas;
- Proportion of population living in urban areas;
- Average (median) age of the population;
- Average mobile internet connection speed;
- Average literacy rate of the population;
- Female literacy rate;
- Male literacy rate;
- Mobile (cell) phone ownership rate of the population;
- Proportion of the population owning a smartphone or mobile phone;
- Share of mobile phones in total web traffic.

Table 1 presents the statistical metrics, including the mean, minimum, maximum, and standard deviations of the variables utilised in the dataset for this study.

Table 1. The statistical properties of the variables.

	Average	Minimum	Maximum	SD
Number of active internet users	56.81	35.40	74.41	12.90
Number of active social media users	52.17	38	68.90	9.302
Cellular mobile network users	69.44	51.45	77.39	7.920
Total population	82.08	76.70	85.59	2.734
Proportion of women by population	50.28	49.80	51.00	0.004
Proportion of men by population	49.72	49.00	50.20	0.004
Proportion of population living in rural areas	24.68	22.40	26.92	0.014
Proportion of population living in urban areas	75.34	73.08	77.60	0.014
Average (median) age of the population	31.52	30.70	32.20	0.397
Average mobile internet connection speed	31.89	25.47	34.90	2.949
Average literacy rate of the population	96.04	95.00	96.70	0.005
Female literacy rate	93.42	92.00	94.70	0.008
Male literacy rate	98.79	98.00	99.10	0.003
Mobile (cell) phone ownership rate of the population	95.42	86.00	98.90	0.041
Proportion of the population owning a smartphone or mobile phone	76.37	34.00	98.80	0.222
Share of mobile phones in total web traffic	6.38	4.14	8	1.343
Average time spent online	62.59	29.43	74.80	0.135

Based on the statistical analysis of the dependent and independent variables presented in Table 1, the data show an even distribution with no missing values, outliers, or noise. Consequently, the entire dataset was incorporated into the model.

3.2. Artificial Neural Networks

Artificial neural networks (ANNs) are computational frameworks designed to emulate the cognitive functions of the human brain. The development of AI stems from efforts to mathematically model the brain's mechanisms of processing information, thereby contributing to advancements in artificial intelligence applications [29].

Artificial nerve cells closely resemble their biological counterparts. These artificial neurons connect to create an ANN, functioning similarly to biological neurons. They have components dedicated to receiving input signals, processing them, and then transmitting outputs. An artificial neural cell comprises five essential parts: inputs, weights, a summation function, an activation function, and an output. The inputs symbolise values or signals obtained from other neural cells or the external environment. Weight values are coefficients that determine the effect of the inputs received by the neural cell on the neural cell. Each input has its own weight value. The summation function computes the total input to a neural cell by adding together the inputs, each multiplied by its corresponding weight. The activation function then processes this net input, deciding on the output that the cell will generate in response [30,31]. The output of each neuron is determined using the following equation:

$$f(x) = f\left(\sum_{i=1} w_i x_i + b\right) \quad (1)$$

In the equation, x_i and w_i represent the input values of the neuron and their weights, respectively. f is the non-linear activation function used throughout the network. The bias value b represents the activation threshold of the neuron.

ANNs can be categorised into feedforward and feedback networks based on the configuration of the interconnections among the neurons. In feedforward networks, neurons are systematically organised in successive layers, beginning with the input and culminating at the output layer. Conversely, feedback neural networks exhibit a more complex interconnection scheme, where the output of a neuron can serve as input not only to neurons in subsequent layers but also to neurons within the same layer or in preceding layers [32,33].

3.3. Random Forest Regression

The random forest (RF) algorithm is a prominent method within the realm of supervised classification algorithms; it finds application in both regression and classification tasks. Its primary objective is to enhance the accuracy of classification outcomes by generating multiple decision trees. This algorithm functions by aggregating numerous decision trees, each operating independently, and subsequently selecting the most optimum result from among them [34]. As the quantity of trees escalates, the likelihood of achieving an accurate outcome simultaneously rises. A principal distinction between the decision tree algorithm and the RF algorithm lies in the latter's utilisation of randomness during the selection of the root node and the division of the nodes. The RF algorithm incorporates an added layer of randomness into the model when cultivating trees. Rather than identifying the most significant feature for node division, it evaluates the best feature from a randomly chosen subset of features. Thus, this algorithm considers a limited, randomly selected set of features for node splitting, fostering significant diversity that typically enhances model performance. Furthermore, the RF algorithm necessitates minimal data preparation, streamlining its implementation process [35].

A significant issue with traditional decision trees is their tendency to overlearn or, in other words, overfit the data. To address this, the RF model employs a strategy of randomly selecting numerous subsets from both the dataset and the feature set for training. This approach generates hundreds of decision trees, each of which makes individual predictions [36,37].

3.4. Support Vector Regression

Support vector machines (SVMs) are a supervised learning approach used to address both classification and regression challenges. They are particularly effective in dealing with complex but relatively small to medium-sized datasets. SVMs function by drawing a line, or hyperplane, that separates the data points on a plane. This line is designed to maximise the distance between points of different classes. The points that lie closest to this separating line are referred to as support vectors [38].

SVR consists of hyperplanes, support vectors, and margins. Hyperplanes are points in space; the support vector is the samples obtained by passing through these points; and the margin is the distance that helps to determine the correct hyperplane to maximise the distances between the data points belonging to the nearest group and the hyperplane. SVR focusses on capturing samples by drawing another hyperplane with the largest possible margin between the hyperplane and the support vector (samples). It maximises this margin to obtain the best decision limit [39].

SVR is suitable for solving linear and non-linear problems. Linear SVR classifies the data with a hyperplane, easily separating the data groups with lines. In non-linear SVR, data groups may not always be separated by lines (curves), and, if the data cannot be classified by a hyperplane, the kernel function method is used [40]. Kernelisation (kernel trick—kernel method) is the process of dimensioning low-dimensional, unanalysable data and processing them with additional operations. There are many kernels for this process. The polynomial kernel method is an analysis model in which the dimensioning process is performed with a transition to a higher subspace in the space under study. Upon applying the radial basis function (RBF) kernel method, a non-linear interval is delineated [41].

3.5. Time Series Method

A time series is a dataset consisting of values of a given variable recorded at regular intervals over time. This method of analysis is used to identify trends, cycles, seasonal effects, and random fluctuations from past data and to use these patterns to predict future behaviour. Time series analysis is widely used in a wide range of disciplines, such as economics, finance, energy, environmental sciences, and social sciences [42].

The main objective of time series analysis is to reveal the structural features of past values of a variable to predict its future trends and thus support decision-making processes. This analysis process is usually based on three main components: trends (long-term trends), seasonality (regular cyclical changes), and random fluctuations (noise). For example, a long-term upward trend in the energy demand may be combined with seasonal effects, resulting in a higher demand during the summer months [43].

The methods used in time series analysis can be divided into two main groups: parametric and non-parametric approaches. Parametric methods assume that the data fit a model and estimate the parameters of this model. For example, the Autoregressive Integrated Moving Average (ARIMA) model is one of the most widely used methods in time series analysis and provides effective results, especially in stationary series. The Seasonal ARIMA (SARIMA) model is a more appropriate approach in series with seasonal effects. In addition, the Holt–Winters method is another popular method that considers both trend and seasonal effects [44].

Non-parametric methods do not rely on model assumptions and provide more flexible analysis. Machine learning and deep learning approaches have become widespread in time series analysis in recent years. Deep learning methods such as long short-term memory (LSTM) are highly effective in learning complex and non-linear relationships [45].

Time series analysis plays a critical role not only in forecasting but also in anomaly detection and policy development. For example, strategic communication policies can be developed by analysing the changes in the perceptions of electric vehicles on social media over time. Therefore, time series analysis has an important place in both theoretical and applied studies [46].

4. Results and Discussion

This study utilised data from the World Bank, spanning 2014 to 2024. Within the model developed for the study, the average time spent on the internet served as the dependent variable. Four distinct machine learning techniques were employed: an ANN, RF regression, SVR, and time series. The model's performance and error rate were evaluated using statistical metrics such as the R^2, MSE, MAE, and MAPE. Figure 1 shows the fundamental working structure of the study.

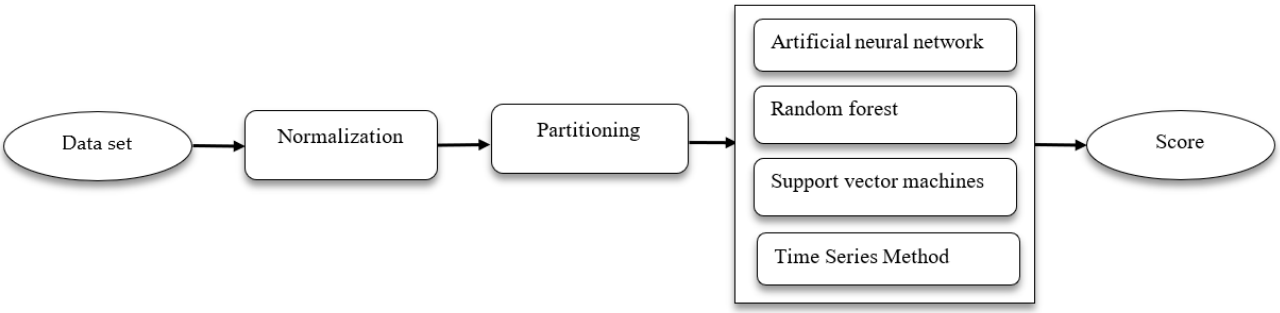


Figure 1. Structure of the study.

Within the domain of data science, an attribute refers to each column within a dataset, serving as a crucial element in producing the desired model output. Attribute selection, or feature selection, encompasses the process of identifying and selecting the most valuable attributes from the dataset. This step is pivotal in influencing the efficacy and efficiency of a machine learning model. Including extraneous attributes can lead to prolonged training times and may compromise the model's simplicity and transparency. Furthermore, an excess of attributes can diminish the interpretability of the model, complicating its analysis and understanding [47]. Overfitting can lead to a model performing well on the training dataset but poorly on a test dataset. This discrepancy arises when the records in the test dataset differ significantly from those in the training set, resulting in a high error rate. Feature selection methods are typically categorised into three main types: filter methods, wrapper methods, and embedded methods. Filter methods evaluate the importance of features by analysing their relationship with the target variable. This process filters the dataset, producing a subset that includes only the most relevant features [48].

In this study, we employed the Pearson correlation method, a type of filter method, to assess the degree of relationship between variables that had a linear connection. Within the scope of the study, the number of cellular mobile network users, the average literacy rate of the population, the female literacy rate, the male literacy rate, and the mobile (cell) phone ownership rate of the population were excluded from the independent variables because they were below the threshold value determined according to the Pearson correlation method.

Normalisation is a technique employed to rectify inconsistencies that arise when data are collected using varying scales or coding systems, thereby ensuring the preservation of the data's integrity [49]. In this study, the dataset underwent normalisation using the min-max method. This technique involves applying a linear transformation to the original data, typically confining them within the range of [0–1]. In this context, the minimum value represents the lowest possible value that a data point can assume, while the maximum value indicates the highest. During the min-max normalisation process, the method evaluates the extent to which each data point exceeds the minimum value and ranks these differences accordingly [50]. Equation (2) shows min-max normalisation:

$$x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2)$$

Cross-validation is a popular validation technique employed in data analysis methods like machine learning and statistical modelling. Its primary aim is to ensure that the dataset is utilised effectively when assessing a model's performance and its ability to generalise to new data [51]. K-fold cross-validation is a widely utilised technique in machine learning that involves partitioning the training dataset into K equal segments. In each iteration, one segment serves as the validation set, while the remaining segments collectively form the training set. This iterative process is repeated K times, ensuring that each segment is used as the validation set exactly once. Consequently, this method yields K distinct accuracy values, from which the overall performance of the model is assessed by calculating their average. Typically, the values of K most frequently adopted are 5 and 10, as they strike a balance between computational efficiency and the reliability of performance estimates [52]. As the dataset grows larger, it is advisable to increase the K value. For this study, a K value of 10 was selected.

The study results were evaluated, analysed, and interpreted using several methods, including the coefficient of determination (R^2), mean squared error (MSE), mean absolute error (MAE), and mean absolute percentage error (MAPE).

The R^2 serves as an indicator of the concordance between the actual values and the predictions generated by a model. It is computed by taking the ratio of the variance explained by the predicted values to the total variance present in the actual values. This statistical measure effectively conveys the effectiveness of the model's fit, providing insights into its predictive accuracy. The R-squared metric takes a value between 0% and 1% after the mathematical process. Here, 0% indicates that the model explains nothing, and the model fit is not good [53]. This outcome suggests that the current data are not appropriate for the model, necessitating a change in the model. On the other hand, a score of 1% signifies that the model offers a perfect explanation of the data, indicating an excellent fit [54]. The MSE serves as an evaluative statistical metric that quantifies the performance of predictive models by calculating the average of the squares of the deviations between the observed actual values and their corresponding predicted counterparts. This metric is instrumental in determining the predictive accuracy of a model. Conversely, the MAE assesses model accuracy by averaging the absolute discrepancies between the actual and predicted values. This metric prioritises the actual magnitude of the prediction errors, thereby providing direct insights into the model's performance. Furthermore, the MAPE offers a comparative measure of forecast inaccuracy by expressing it in percentage terms relative to the actual demand. This renders the MAPE particularly intuitive and user-friendly for interpretive purposes. Importantly, the MAPE metric circumvents the challenges associated with comparing models evaluated across different units of measurement, offering a standardised comparative framework [55].

Among these performance measures, the R^2 serves as the coefficient of determination to assess the model's accuracy. A high value for this coefficient indicates a good prediction relationship. The MSE, MAE, and MAPE, on the other hand, are error measures, and low results indicate high performance, being inversely proportional to the performance [56,57].

In this study, four distinct machine learning methodologies were employed: an ANN, RF regression, SVR, and time series. For the ANN approach, a feedback model was designed, comprising eleven input neurons and a single output neuron. Through iterative experimentation, it became evident that a configuration featuring four hidden layers, each containing four neurons, yielded more effective outcomes. The architecture of the model developed is depicted in Figure 2.

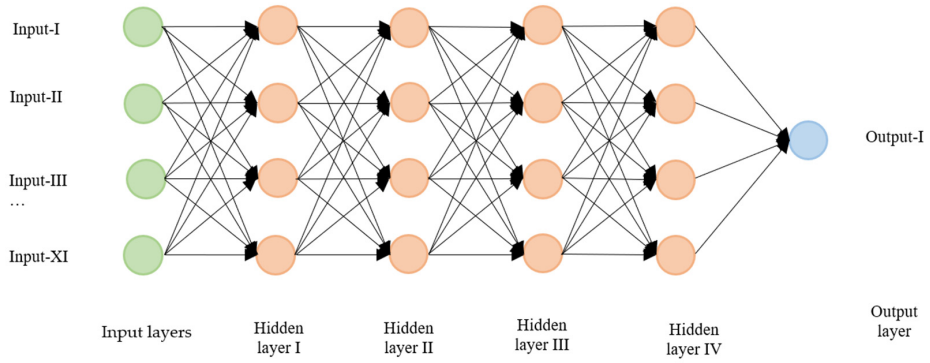


Figure 2. ANN model.

Following an extensive evaluation of several potential activation functions, the sigmoid function emerged as the most effective option due to its superior performance. The sigmoid activation function is characterised by its continuous and differentiable nature, yielding output values constrained between 0 and 1 for any input. In the context of the learning method, a feedback learning algorithm was employed, serving as a critical parameter influencing the model's success in ANNs. In a feedback ANN, at least one cell generates an output that is subsequently fed back as an input either to itself or other cells. This feedback mechanism may occur among cells within the same layer or across different layers [58]. In the model, 500 iterations were performed to determine the best result. The Gini index is preferred as the division criterion in the RF method. The tree depth value is set to 10, and the model is configured. In this study, non-linear SVR was utilised, specifically testing radial basis function (RBF) and polynomial kernels. The RBF kernel was selected as it provided the most successful results. This kernel assesses the input features of data points using a Gaussian measure to determine the similarity or distance between pairs. The model's parameters included an overlapping penalty value of 200 and an RBF sigma value of 0.5. In this study, the past five time steps were used for time series forecasting. That is, for each forecast, the model considers data from the previous five time periods. The data are reshaped to relate the past time steps of the independent variables (X) to the current values of the dependent variable (y).

Table 2 presents a comparison of the models' evaluations using the ANN, RF regression, SVR, and time series methods.

Table 2. Evaluation of models.

	ANN	RF	SVR	Time Series
R^2	0.925	0.878	0.905	0.863
MSE	0.001	0.015	0.003	0.017
MAE	0.016	0.098	0.025	0.103
MAPE	0.043	0.137	0.078	0.144

Figure 3 shows the R^2 analysis. The R^2 is a statistical measure used to assess a model's performance in regression analysis. It indicates how closely the data align with the regression line. In other words, it shows how much the dependent variable will change, on average, in terms of its own unit of measurement, when the independent variable changes by one unit in its own unit of measurement [59]. The study's findings revealed that the R^2 metric was 92.5% for the ANN, 87.8% for RF regression, 90.5% for SVR, and 86.3% for the time series. The R^2 metric, ranging from 0 to 1, measures the fit quality of a regression model, with values approaching 1 indicating a superior fit. Although the ANN emerged as the most successful model in this analysis, all four models demonstrated efficacy in their predictive capacities.

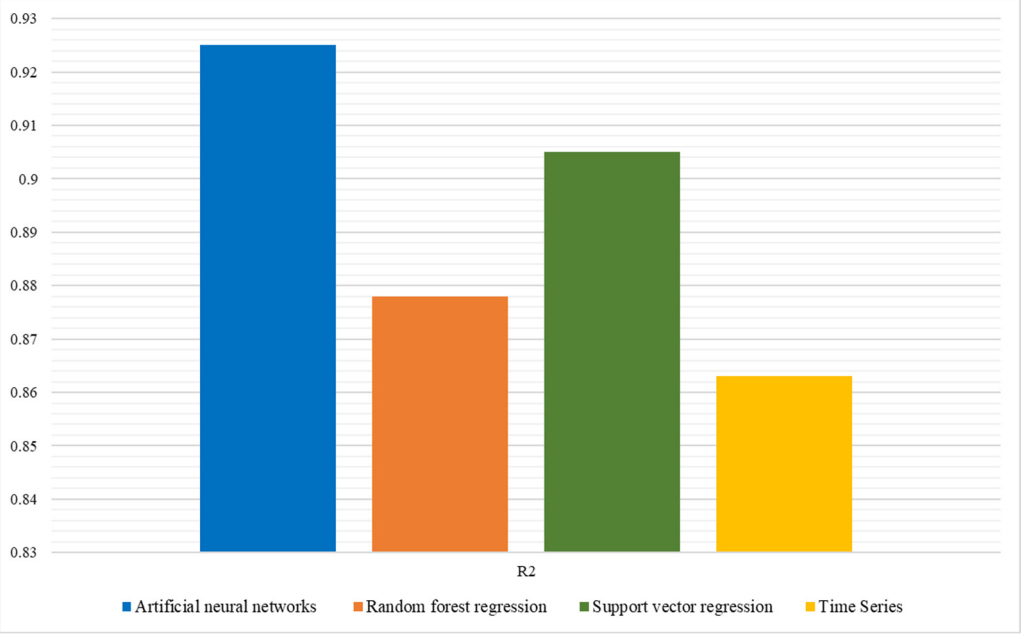


Figure 3. R^2 analysis.

Figure 4 shows the MSE analysis. The MSE serves as a metric to gauge the accuracy of a regression model's predictions compared to actual values. It is determined by finding the average of the squared differences between the predicted and true values. A lower MSE signifies better model performance [60]. This study found that the MSE metric for the ANN was 0.001; for RF regression, it was 0.015; for SVR, it was 0.003; and for time series, it was 0.017. Although the model with the fewest errors in this study was the ANN, it was found that the error rates of all four models were low and acceptable.

Figure 5 shows the MAE analysis. The MAE metric calculates the sum of the absolute differences between the predicted values and actual values. Essentially, it quantifies how far away each prediction is from the true value [61]. This study revealed values of 0.016 for the ANN, 0.098 for RF regression, 0.025 for SVR, and 0.025 for the time series. The ideal value of the MAE is zero because it means that the predictions match the actual values exactly. Lower MAE values represent better model performance. This study found that the MAE value in all four models was remarkably close to the ideal one.

Figure 6 shows the MAPE analysis. The MAPE metric shows the incorrectness of each prediction as a percentage compared to the true value [60]. This study revealed values of 0.043 for the ANN, 0.137 for RF regression, 0.078 for SVR, and 0.144 for time series. It is widely recognised that the nearer the value of the MAPE is to zero, the more accurate the predictions are.

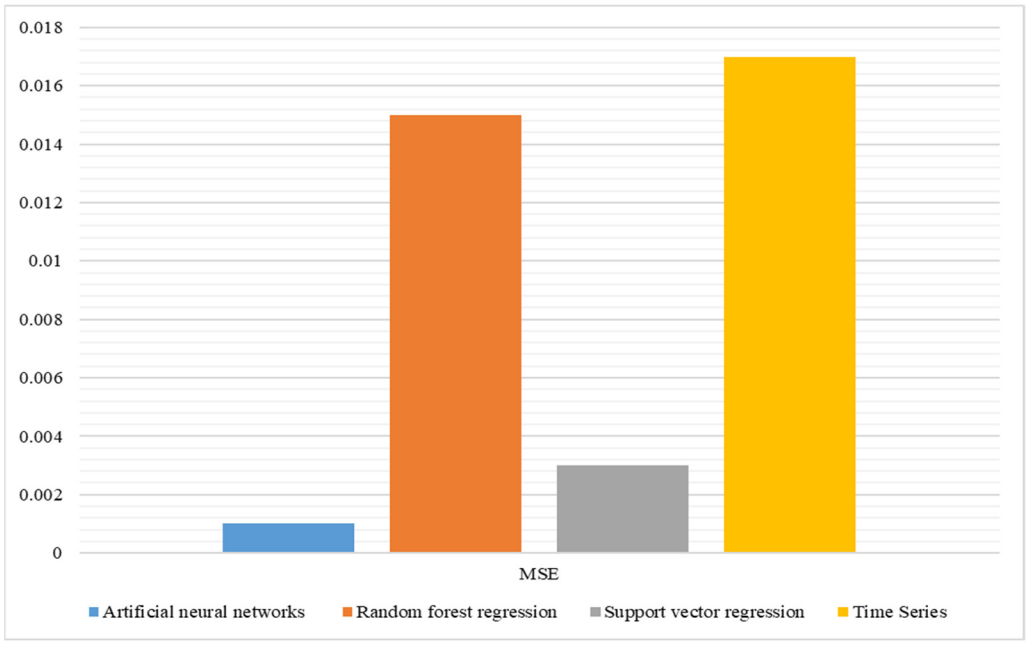


Figure 4. MSE analysis.

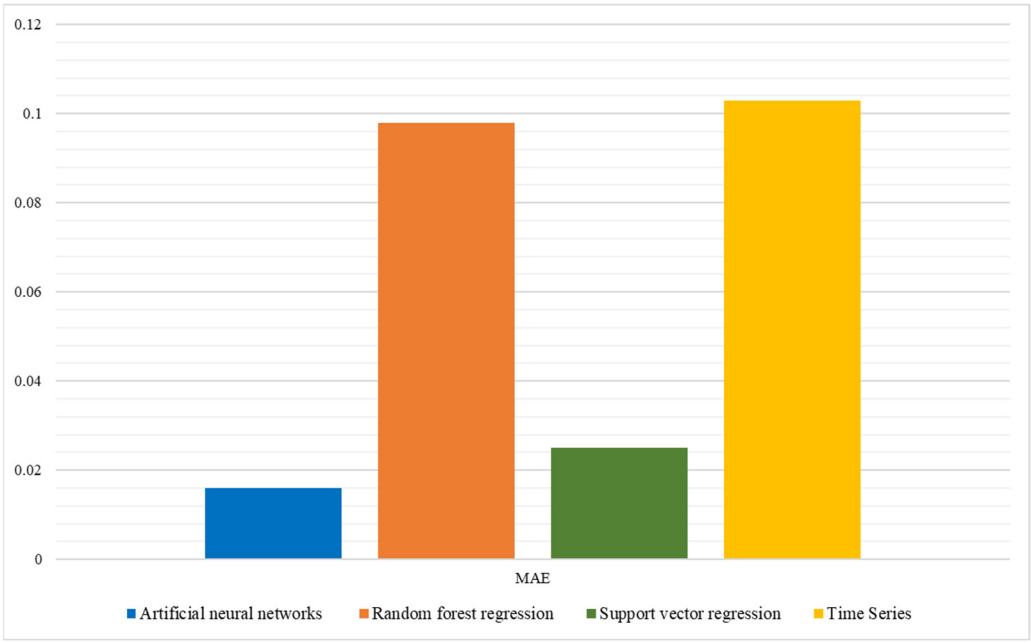


Figure 5. MAE analysis.

This study found that the MAPE metric across all four models closely approached the ideal and acceptable benchmark. Upon analysing the error and success metrics, the models were ranked in terms of performance, with the ANN achieving the highest success and lowest error rates, followed by SVR, RF regression, and subsequently time series.

The study results were made easy to understand using data visualisation techniques. Data visualisation helps to render the features in the dataset more visible and make the dataset more understandable. This study explored the relationship between the predicted and actual data by employing scatter and line plots.

A scatter plot is used to explore how two values change in relation to each other across two dimensions. With this diagram, although the reason for the relationship cannot be understood, the presence and strength of a relationship between two variables can be

determined, facilitating classification into categories such as strong positive, weak positive, strong negative, weak negative, or no relationship [62]. Figure 7 shows the scatter plot curves for the ANN, RF regression, SVR, and time series methods.

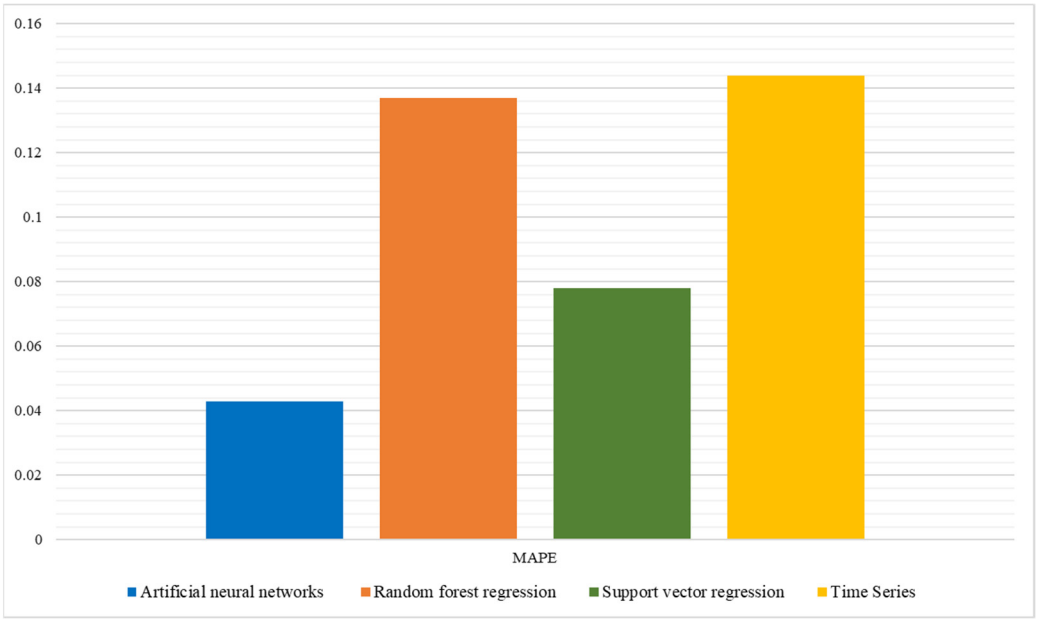


Figure 6. MAPE analysis.

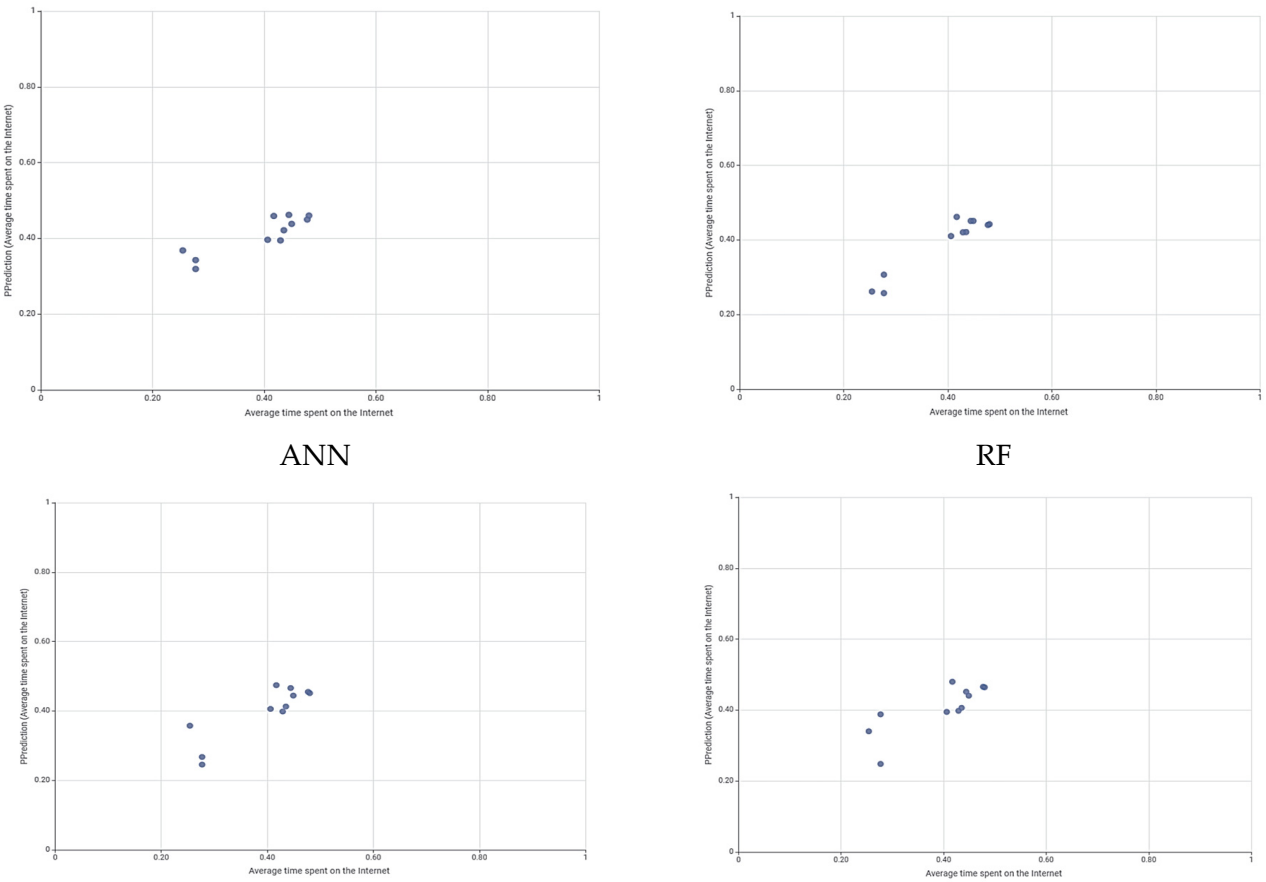


Figure 7. Scatter plots of methods.

Figure 7 shows scatter plot graphs comparing the actual average time spent on the internet with the predicted values generated by the machine learning models. On these plots, the x-axis represents the actual values, while the y-axis represents the predicted values. Each point on the graph corresponds to a specific data sample. Scatter plots serve as an excellent tool for the analysis of the relationship between variables, as well as detecting patterns, highlighting outliers, and assessing the data distribution. The figures indicate a strong positive correlation between the predicted and actual values; as one variable increases, so does the other, resulting in points that largely cluster around the line.

A line chart serves as an effective tool in illustrating temporal trends. In this type of visualisation, the dimension, such as time, is invariably plotted along the x-axis, while the calculated values occupy the y-axis. To construct a line on the chart, the dataset must include a minimum of two data points. In contrast, a dataset containing only a single value is depicted merely as a point. They are graphs used to compare the mean values obtained from continuous variables of different categorical variables [63]. Figure 8 shows the line plots for the ANN, RF regression, SVR, time series methods.

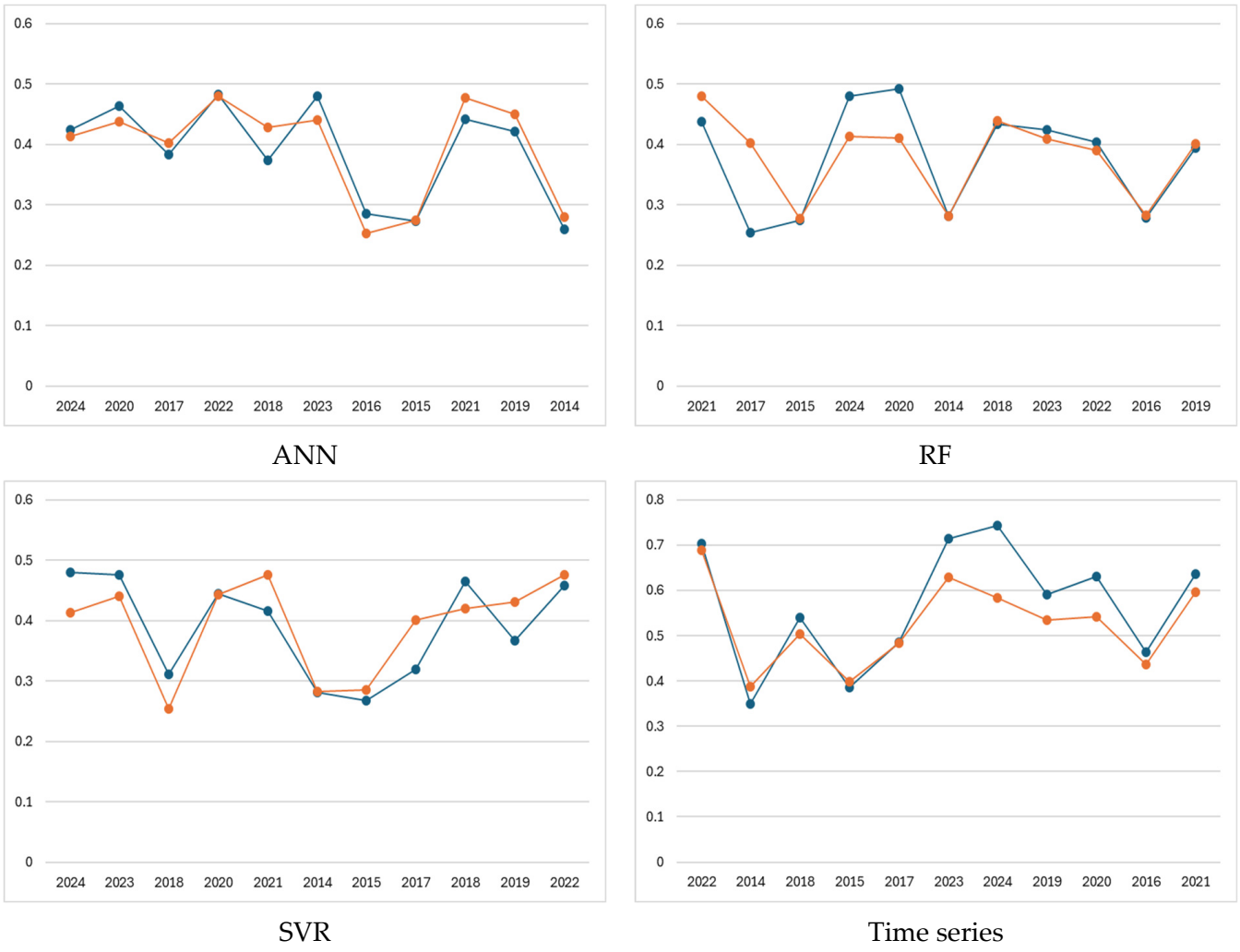


Figure 8. Line plots of methods.

Figure 8 shows the comparative analysis between the actual and predicted values concerning the average time spent on the internet, evaluated using several machine learning models. The x-axis of these line plots signifies the time interval, while the y-axis represents the normalised average time spent online. In these visual representations, the red lines denote the actual measurements, whereas the blue lines depict the predicted estimates.

The figures collectively suggest that the methodologies employed exhibit a high degree of predictive accuracy. Among these, it is evident that the ANN offers the most precise alignment with the actual data, followed sequentially by SVR, RF, and the time series. This observation underscores that the ANN approach demonstrates a superior fit and predictive capabilities compared to its counterparts.

In 2014, the average daily duration spent on the internet amounted to 277 min, increasing to 417 min by 2024. This study utilised machine learning models to estimate the average time devoted to internet use. According to the results of this model, it is calculated that it will reach 507 min in 2030 and 603 min in 2035.

The results showed that the artificial neural network (ANN) model provided the highest accuracy rate and was superior to the other models in predicting internet usage times. The ANN model was characterised by low error rates due to its ability to work with complex data structures. Although random forest (RF), support vector regression (SVR), and time series also provided high accuracy rates, the superiority of the ANN in terms of the accuracy and error metrics made a significant difference. These findings suggest that the ANN is a model that should be particularly favoured in the prediction of time series data.

5. Discussion

This research employed four distinct machine learning techniques to estimate the average duration of internet usage, utilising data spanning from 2014 to 2024, sourced from the World Bank. To enhance the model's reliability and precision, a cross-validation approach was implemented. The findings revealed that all four models effectively predicted the average time spent online. Evaluation metrics, including the R^2, MSE, MAE, and MAPE, were calculated to assess the models' performance. Notably, the ANN emerged as the most successful model, achieving an R^2 of 92.5%, MSE of 0.001, MAE of 0.016, and MAPE of 0.043. The ANN model outperformed the SVR, RF, and time series methods. Consistent with the existing literature, the study demonstrated that the ANN was the most effective technique. Furthermore, the results indicated that the model provided predictions with greater accuracy and reduced errors compared to previous studies in the literature.

The relationship between internet usage and phubbing is well documented in prior studies [10]. Increased time spent online correlates with reduced face-to-face interaction quality, thereby escalating phubbing behaviours. This study bridges this gap by forecasting future internet usage patterns and discussing their potential impact on phubbing's prevalence. Our findings highlight the importance of addressing phubbing through predictive modelling to mitigate its social and psychological consequences.

Although the artificial neural network (ANN) provided the highest accuracy rate among the models used in this study, the advantages and disadvantages of other models for different scenarios should be discussed. For example, the random forest (RF) model may be faster and more effective on small datasets with its less complex structure, while support vector regression (SVR) may give better results on non-linear data. The success of a time series model depends on the characteristics, seasonality, and non-linearity of the data, so the advantages of different models should be assessed for specific datasets. In future studies, it is suggested to evaluate the performance of more advanced models (e.g., LSTM and XGBoost) on complex datasets. These comparisons will be useful to determine which model is more appropriate in different contexts.

Technological developments in smartphones increase phubbing behaviours by offering individuals factors such as the internet, social media, and games together. This situation causes people to focus on their smartphones instead of communicating with their surroundings in social environments and negatively affects their social communication. There are

many studies indicating that internet use has a negative impact on individuals' mental health [64–69]. Therefore, considering that the duration of internet usage will increase by almost 1.5 times, the physical and social problems that may be experienced in the future will be directly related to the increasing duration of internet usage.

The results of this study are consistent with prior research on the predictive capabilities of machine learning models for internet usage. For instance, the ANN's superior performance in terms of accuracy aligns with findings from similar studies [10,22]. However, our study differs by focusing on the social implications of these predictions, particularly their relevance to phubbing behaviours, a topic rarely explored in the existing literature.

The phubbing phenomenon should be addressed not only with quantitative analyses but also in a qualitative context. For example, interview or survey studies can be conducted to understand individuals' experiences and perceptions of this phenomenon. Furthermore, case studies examining the effects of phubbing on family relationships, work environments, and social interactions may allow for a broader perspective on this phenomenon. Such qualitative studies would deepen the meaning of existing findings and contribute to providing more comprehensive solutions.

While this study employed basic machine learning models due to their interpretability and ease of replication, future research could benefit from incorporating advanced models such as LSTM or XGBoost. These models are well suited for capturing non-linear and temporal dependencies in large datasets, which could provide deeper insights into the behavioural patterns associated with phubbing.

This study focused on the quantitative analysis of phubbing behaviour. However, qualitative insights into the social and relational effects of phubbing were not addressed in this study. To develop a more holistic perspective, future interviews or case studies with individuals may contribute to a deeper understanding of the social and relational effects of this phenomenon.

This study focused on internet use, but psychological and behavioural variables (e.g., social anxiety, loneliness) were not included in the model. This can be considered a factor that limits the scope of the findings. Future studies may contribute to a more in-depth examination of phubbing behaviour by integrating such behavioural variables into the model.

The dataset used in this study is based on global averages and does not include geographical diversity. This is considered a factor that limits the generalisability of the findings. In future research, it is recommended to create a more diverse dataset by collecting data from different cultural and regional contexts. This approach may allow a better understanding of the local effects of phubbing behaviour.

6. Conclusions and Future Work

This study examined how future increases in internet usage durations may affect phubbing behaviours. The findings suggest that internet use may have potential negative effects on social relationships. Among the machine learning models used in the study, the ANN has shown superior performance compared to other methods by providing the highest accuracy rate.

By addressing the relationship between internet usage, phubbing behaviours, and their broader social implications, this study aims to contribute to the interdisciplinary discourse on technology's role in shaping human interactions. Our findings provide a basis for future AI-driven behavioural studies, offering actionable insights for both academic researchers and policymakers.

Phubbing, namely engaging with a phone or other digital device during face-to-face contact with a person, is a behaviour that has become common in modern society.

From a health perspective, phubbing can have negative effects on both individuals and relationships. Constant digital distraction can weaken interpersonal bonds, reduce empathy, and increase feelings of loneliness. Furthermore, this behaviour can raise stress levels and negatively impact mental health. Research shows that phubbing can also be associated with health problems such as sleep disorders, depression, and anxiety. Therefore, having a balanced relationship with digital devices is important in maintaining both physical and mental health.

The average daily time spent on the internet reached 277 min in 2014 and 417 min in 2024. In this study, the average time spent on the internet was estimated according to machine learning models. According to the results of this model, it is calculated that it will reach 507 min in 2030 and 603 min in 2035. These predictions may indicate serious health risks. Prolonged internet use can encourage sedentary lifestyles, leading to chronic conditions such as obesity, heart disease, and diabetes. Furthermore, screen addiction can impair eye health, negatively impact sleep patterns, and trigger mental health problems, particularly anxiety and depression. To manage this increase, digital detox practices should be encouraged, individuals should limit their screen time, and regular physical activity and social interactions should be increased. Healthy internet usage habits should be promoted through education and awareness-raising campaigns.

The findings of this study reveal the long-term social effects of the phubbing phenomenon. However, it is critical to develop strategies to reduce the negative effects of this phenomenon. Suggested interventions may include digital detox programs, promoting mobile applications that limit individuals' screen time, and organising awareness campaigns. In addition, policies that limit the use of digital communication tools can be implemented in social settings such as educational institutions and workplaces. In future research, it is recommended to conduct studies on the effectiveness of such interventions.

The social, economic, and ecological impacts of the phubbing phenomenon are quite wide-ranging. In the social context, a decrease in the frequency of individuals' face-to-face communication could lead to consequences such as loneliness, insecurity, and relationship dissatisfaction. Economically, the need for constant digital interaction can reduce individuals' productivity and cause a loss of attention in the business world. In financial terms, it has been observed that technology dependency affects individuals' spending habits and encourages the consumption of more technological devices. In the ecological context, increased internet usage and device production may contribute to the growth of the carbon footprint by increasing energy consumption. For these reasons, it is suggested that individuals should implement digital detox strategies and adopt sustainable technology policies to reduce the effects of phubbing.

The dataset used in this study, which included global averages for 2014–2024, provided by the World Bank, highlights that technology use and social norms exhibit regional and cultural differences. Future studies could explore these local effects by collecting data from various geographical regions and cultural contexts, such as comparing internet usage habits in developed and developing countries and examining how these differences influence phubbing. With the proliferation of 5G/6G technologies and edge–fog–cloud infrastructures, the expected surge in digital data will necessitate the use of more complex and large datasets for the analysis of social phenomena like phubbing. In this context, the applicability of artificial intelligence and machine learning techniques in edge environments should be prioritised, integrating real-time data analytics and prediction methods, particularly with data collected from user devices like smartwatches. Real-time data analysis for stress management and behaviour prediction offers a promising approach to better understand and mitigate phubbing. Future research should focus on larger datasets, edge–fog–cloud architectures, and advanced hybrid AI models to process data in real time and predict digital

behaviours more effectively. Given phubbing's potential to cause more significant negative effects than other types of addiction, as highlighted in prior research [70], its dangers are expected to increase with rising internet usage. Therefore, it is critical to develop more robust AI-based models, utilise additional datasets and independent variables, and devise innovative methods to predict and address the social impacts of phubbing.

Author Contributions: Conceptualisation, M.A.A., M.K. (Murad Karaduman) and T.L.; data curation, S.K. and F.K.; formal analysis, M.K. (Murad Karaduman); funding acquisition, A.Y.; investigation, M.K. (Mehmet Kayakuş), F.Y.A. and T.L.; methodology, M.A.A. and M.K. (Mehmet Kayakuş); project administration, A.Y.; resources, A.Y., M.K. (Murad Karaduman) and S.K.; software, M.K. (Mehmet Kayakuş); validation, M.K. (Murad Karaduman), F.Y.A., T.L. and F.K.; writing—original draft, A.Y., M.K. (Mehmet Kayakuş), S.K., F.Y.A. and T.L.; writing—review and editing, M.A.A., F.Y.A. and F.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset is available from the authors upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of the data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Atteh, E.; Assan-Donkoh, I.; Mensah, Y.A.; Boadi, A.; Badzi, S.C.; Lawer, V.T. A thoughtful overview of social media usage among students and its impact on their academic work. *Asian J. Adv. Res. Rep.* **2020**, *8*, 30–39. [CrossRef]
2. Jones, T. Cell phone use while walking across campus: An observation and survey. *Elon J. Undergrad. Res. Commun.* **2014**, *5*, 1–2.
3. Roberts, J.; Yaya, L.; Manolis, C. The invisible addiction: Cell-phone activities and addiction among male and female college students. *J. Behav. Addict.* **2014**, *3*, 254–265. [CrossRef] [PubMed]
4. Yıldırım, M.; Öztürk, A.; Solmaz, F. Fear of COVID-19 and sleep problems in Turkish young adults: Mediating roles of happiness and problematic social networking sites use. *Psihologija* **2023**, *56*, 497–515. [CrossRef]
5. Boase, J.; Kobayashi, T. Kei-Tying teens: Using mobile phone e-mail to bond, bridge, and break with social ties—A study of Japanese adolescents. *Int. J. Hum.-Comput. Stud.* **2008**, *66*, 930–943. [CrossRef]
6. Burchell, K. Tasking the everyday: Where mobile and online communication take time. *Mob. Media Commun.* **2015**, *3*, 36–52. [CrossRef]
7. Schou Andreassen, C.; Pallesen, S. Social network site addiction-an overview. *Curr. Pharm. Des.* **2014**, *20*, 4053–4061. [CrossRef]
8. Cheok, A.D.; Sreekumar, A.; Lei, C.; Thang, L.N. Capture the flag: Mixed-reality social gaming with smart phones. *IEEE Pervasive Comput.* **2006**, *5*, 62–69. [CrossRef]
9. Zhang, K.Z.; Chen, C.; Lee, M.K. Understanding the Role of Motives in Smartphone Addiction. In Proceedings of the The Pacific Asia Conference on Information Systems, Chengdu, China, 24–28 June 2014.
10. Rahman, M.A.; Duradoni, M.; Guazzini, A. Identification and prediction of phubbing behavior: A data-driven approach. *Neural Comput. Appl.* **2022**, *34*, 3885–3894. [CrossRef]
11. Liyanage, L.R.M.C.; Sulaiman, M.F.; Ismail, R.; Gunaratne, G.P.; Dharmakeerthi, R.S.; Rupasinghe, M.G.; Mayakaduwa, A.P.; Hanafi, M.M. Carbon mineralization dynamics of organic materials and their usage in the restoration of degraded tropical tea-growing soil. *Agronomy* **2021**, *11*, 1191. [CrossRef]
12. Chotpitayasunondh, V.; Douglas, K.M. How "phubbing" becomes the norm: The antecedents and consequences of snubbing via smartphone. *Comput. Hum. Behav.* **2016**, *63*, 9–18. [CrossRef]
13. Datareportal, Digital 2023: Global Overview Report. 2024. Available online: https://datareportal.com/reports/digital-2023-global-overview-report (accessed on 12 June 2024).
14. Baron, N.S.; Campbell, E.M. Gender and mobile phones in cross-national context. *Lang. Sci.* **2012**, *34*, 13–27. [CrossRef]
15. Campbell, S.W.; Kwak, N. Mobile communication and civic life: Linking patterns of use to civic and political engagement. *J. Commun.* **2010**, *60*, 536–555. [CrossRef]
16. Lee, Y.-K. The dark side of smartphone usage: Psychological traits, compulsive behavior and technostress. *Comput. Hum. Behav.* **2014**, *31*, 373–383. [CrossRef]

17. Xie, X.; Chen, W.; Zhu, X.; He, D. Parents' phubbing increases Adolescents' Mobile phone addiction: Roles of parent-child attachment, deviant peers, and gender. *Child. Youth Serv. Rev.* **2019**, *105*, 104426. [CrossRef]
18. Roberts, J.A.; David, M.E. My life has become a major distraction from my cell phone: Partner phubbing and relationship satisfaction among romantic partners. *Comput. Hum. Behav.* **2016**, *54*, 134–141. [CrossRef]
19. Chotpitayasunondh, V.; Douglas, K.M. The effects of "phubbing" on social interaction. *J. Appl. Soc. Psychol.* **2018**, *48*, 304–316. [CrossRef]
20. Mishra, A.R.; Singh, R.K.; Motwani, D. Intuitionistic fuzzy divergence measure-based ELECTRE method for performance of cellular mobile telephone service providers. *Neural Comput. Appl.* **2020**, *32*, 3901–3921. [CrossRef]
21. Przybylski, A.K.; Weinstein, N. Digital screen time limits and young children's psychological well-being: Evidence from a population-based study. *Child Dev.* **2019**, *90*, e56–e65. [CrossRef] [PubMed]
22. Vanden Abeele, M.M.; Postma-Nilsenova, M. More than just gaze: An experimental vignette study examining how phone-gazing and newspaper-gazing and phubbing-while-speaking and phubbing-while-listening compare in their effect on affiliation. *Commun. Res. Rep.* **2018**, *35*, 303–313. [CrossRef]
23. Zhang, N.; Jia, W.; Lei, H.; Wang, P.; Zhao, P.; Guo, Y.; Dung, C.H.; Bu, Z.; Xue, P.; Xie, J.; et al. Effects of human behavior changes during the coronavirus disease 2019 (COVID-19) pandemic on influenza spread in Hong Kong. *Clin. Infect. Dis.* **2021**, *73*, e1142–e1150. [CrossRef]
24. Miller, D.; Loftus, A.M.; O'Boyle, P.J.; McCloskey, M.; O'Kelly, J.; Mace, D.; McKeon, N.; Ewan, S.L.; Moore, L.; Abbott, A.; et al. Impact of a telephone-first consultation system in general practice. *Postgrad. Med. J.* **2019**, *95*, 590–595. [CrossRef] [PubMed]
25. Masood, A.; Zhang, Q.; Ali, M.; Cappiello, G.; Dhir, A. Linking enterprise social media use, trust and knowledge sharing: Paradoxical roles of communication transparency and personal blogging. *J. Knowl. Manag.* **2023**, *27*, 1056–1085. [CrossRef]
26. Balta, S.; Emirtekin, E.; Kircaburun, K.; Griffiths, M.D. Neuroticism, trait fear of missing out, and phubbing: The mediating role of state fear of missing out and problematic Instagram use. *Int. J. Ment. Health Addict.* **2020**, *18*, 628–639. [CrossRef]
27. Franchina, V.; Vanden Abeele, M.; Van Rooij, A.J.; Lo Coco, G.; De Marez, L. Fear of missing out as a predictor of problematic social media use and phubbing behavior among Flemish adolescents. *Int. J. Environ. Res. Public Health* **2018**, *15*, 2319. [CrossRef] [PubMed]
28. Mahmud, I.; Supramaniam, K.; Jahan, N.; Begum, A.; Masud, A. Dark Side of Mobile Phone Technology: Assessing the Impact of Self-Phubbing and Partner-Phubbing on Life Satisfaction. *Interdiscip. J. Inf. Knowl. Manag.* **2024**, *18*, 893–914. [CrossRef] [PubMed]
29. Kumar, K.; Thakur, G.S.M. Advanced applications of neural networks and artificial intelligence: A review. *Int. J. Inf. Technol. Comput. Sci.* **2012**, *4*, 57. [CrossRef]
30. Kayakuş, M.; Erdoğan, D.; Terzioğlu, M. Predicting the share of tourism revenues in total exports. *Alphanumer. J.* **2023**, *11*, 17–30. [CrossRef]
31. Sharma, S.; Sharma, S.; Athaiya, A. Activation functions in neural networks. *Towards Data Sci.* **2017**, *6*, 310–316. [CrossRef]
32. Dongare, A.; Kharde, R.; Kachare, A.D. Introduction to artificial neural network. *Int. J. Eng. Innov. Technol. (IJEIT)* **2012**, *2*, 189–194.
33. Kayakus, M.; Tutcu, B.; Terzioglu, M.; Talaş, H.; Ünal Uyar, G.F. ROA and ROE forecasting in iron and steel industry using machine learning techniques for sustainable profitability. *Sustainability* **2023**, *15*, 7389. [CrossRef]
34. Osisanwo, F.Y.; Akinsola, J.E.; Awodele, O.; Hinmikaiye, J.O.; Olakanmi, O.; Akinjobi, J. Supervised machine learning algorithms: Classification and comparison. *Int. J. Comput. Trends Technol. (IJCTT)* **2017**, *48*, 128–138.
35. Svetnik, V.; Liaw, A.; Tong, C.; Culberson, J.C.; Sheridan, R.P.; Feuston, B.P. Random forest: A classification and regression tool for compound classification and QSAR modeling. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1947–1958. [CrossRef] [PubMed]
36. Amro, A.; Al-Akhras, M.; Hindi, K.E.; Habib, M.; Shawar, B.A. Instance reduction for avoiding overfitting in decision trees. *J. Intell. Syst.* **2021**, *30*, 438–459. [CrossRef]
37. Archer, K.J.; Kimes, R.V. Empirical characterization of random forest variable importance measures. *Comput. Stat. Data Anal.* **2008**, *52*, 2249–2260. [CrossRef]
38. Kayakuş, M.; Açıkgöz, F.Y. Classification of news texts by categories using machine learning methods. *Alphanumer. J.* **2022**, *10*, 155–166. [CrossRef]
39. Suganyadevi, M.; Babulal, C. Support vector regression model for the prediction of loadability margin of a power system. *Appl. Soft Comput.* **2014**, *24*, 304–315. [CrossRef]
40. Brereton, R.G.; Lloyd, G.R. Support vector machines for classification and regression. *Analyst* **2010**, *135*, 230–267. [CrossRef]
41. Basak, D.; Pal, S.; Patranabis, D.C. Support vector regression. *Neural Inf. Process.-Lett. Rev.* **2007**, *11*, 203–224.
42. Munoz-Tuduri, M.; García-Moro, C.; Walker, P.L. Time series analysis of the epidemiological transition in Minorca 1634–1997. *Hum. Biol.* **2006**, *78*, 619–634. [CrossRef] [PubMed]
43. Lim, B.; Zohren, S. Time-series forecasting with deep learning: A survey. *Philos. Trans. R. Soc. A* **2021**, *379*, 20200209. [CrossRef] [PubMed]
44. Chen, G.; Abraham, B.; Bennett, G.W. Parametric and non-parametric modelling of time series—An empirical study. *Environmetr. Off. J. Int. Environmetr. Soc.* **1997**, *8*, 63–74. [CrossRef]

45. Storlie, C.B.; Swiler, L.P.; Helton, J.C.; Sallaberry, C.J. Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models. *Reliab. Eng. Syst. Saf.* **2009**, *94*, 1735–1763. [CrossRef]
46. Cook, A.A.; Mısırlı, G.; Fan, Z. Anomaly detection for IoT time-series data: A survey. *IEEE Internet Things J.* **2019**, *7*, 6481–6494. [CrossRef]
47. Dash, M.; Liu, H. Feature selection for classification. *Intell. Data Anal.* **1997**, *1*, 131–156. [CrossRef]
48. Venkatesh, B.; Anuradha, J. A review of feature selection and its methods. *Cybern. Inf. Technol.* **2019**, *19*, 3–26. [CrossRef]
49. Nayak, S.; Misra, B.B.; Behera, H.S. Impact of data normalization on stock index forecasting. *Int. J. Comput. Inf. Syst. Ind. Manag. Appl.* **2014**, *6*, 13.
50. Saif, A.S.; Garba, A.G.; Awwalu, J.; Arshad, H.; Zakaria, L.Q. Performance comparison of min-max normalisation on frontal face detection using haar classifiers. *Pertanika J. Sci. Technol.* **2017**, *25*, 163–171.
51. Krstajic, D.; Buturovic, L.J.; Leahy, D.E.; Thomas, S. Cross-validation pitfalls when selecting and assessing regression and classification models. *J. Cheminform.* **2014**, *6*, 1–15. [CrossRef]
52. Nti, I.K.; Nyarko-Boateng, O.; Aning, J. Performance of machine learning algorithms with different K values in K-fold cross-validation. *Int. J. Inf. Technol. Comput. Sci.* **2021**, *13*, 61–71.
53. Piepho, H.P. A coefficient of determination (R2) for generalized linear mixed models. *Biom. J.* **2019**, *61*, 860–872. [CrossRef] [PubMed]
54. Saunders, L.J.; Russell, R.A.; Crabb, D.P. The coefficient of determination: What determines a useful R2 statistic? *Investig. Ophthalmol. Vis. Sci.* **2012**, *53*, 6830–6832. [CrossRef]
55. Chicco, D.; Warrens, M.J.; Jurman, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *Peerj Comput. Sci.* **2021**, *7*, e623. [CrossRef]
56. Kayakuş, M.; Terzioğlu, M.; Yetiz, F. Forecasting housing prices in Turkey by machine learning methods. *Aestimum* **2022**, *80*, 33–44. [CrossRef]
57. Mueller, T.; Pierce, F. Soil carbon maps: Enhancing spatial estimates with simple terrain attributes at multiple scales. *Soil Sci. Soc. Am. J.* **2003**, *67*, 258–267. [CrossRef]
58. Rezaeian Zadeh, M.; Amin, S.; Khalili, D.; Singh, V.P. Daily outflow prediction by multi layer perceptron with logistic sigmoid and tangent sigmoid activation functions. *Water Resour. Manag.* **2010**, *24*, 2673–2688. [CrossRef]
59. Lüdecke, D.; Ben-Shachar, M.S.; Patil, I.; Waggoner, P.; Makowski, D. performance: An R package for assessment, comparison and testing of statistical models. *J. Open Source Softw.* **2021**, *6*, 31–39. [CrossRef]
60. Tofallis, C. A better measure of relative prediction accuracy for model selection and model estimation. *J. Oper. Res. Soc.* **2015**, *66*, 1352–1362. [CrossRef]
61. Hodson, T.O. Root mean square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geosci. Model Dev.* **2022**, *15*, 5481–5487. [CrossRef]
62. Kahng, S.; Iwata, B.A.; Fischer, S.M.; Page, T.J.; Treadwell, K.R.H.; Williams, D.E.; Smith, R.G. Temporal distributions of problem behavior based on scatter plot analysis. *J. Appl. Behav. Anal.* **1998**, *31*, 593–604. [CrossRef] [PubMed]
63. Larson–Hall, J. Moving beyond the bar plot and the line graph to create informative and attractive graphics 1. *Mod. Lang. J.* **2017**, *101*, 244–270. [CrossRef]
64. Ergün, N.; Özkan, Z.; Griffiths, M.D. Social media addiction and poor mental health: Examining the mediating roles of internet addiction and phubbing. *Psychol. Rep.* **2023**, 1–21. [CrossRef] [PubMed]
65. Hadlington, L. Human factors in cybersecurity; examining the link between Internet addiction, impulsivity, attitudes towards cybersecurity, and risky cybersecurity behaviours. *Heliyon* **2017**, *3*, e00346. [CrossRef] [PubMed]
66. Kuss, D.; D Griffiths, M.; Karila, L.; Billieux, J. Internet addiction: A systematic review of epidemiological research for the last decade. *Curr. Pharm. Des.* **2014**, *20*, 4026–4052. [CrossRef] [PubMed]
67. van Den Eijnden, R.J.; Spijkerman, R.; Vermulst, A.A.; van Rooij, T.J.; Engels, R.C. Compulsive Internet use among adolescents: Bidirectional parent–child relationships. *J. Abnorm. Child Psychol.* **2010**, *38*, 77–89. [CrossRef] [PubMed]
68. Chern, K.C.; Huang, J.H. Internet addiction: Associated with lower health-related quality of life among college students in Taiwan, and in what aspects? *Comput. Hum. Behav.* **2018**, *84*, 460–466. [CrossRef]
69. Huang, C. Time spent on social network sites and psychological well-being: A meta-analysis. *Cyberpsychol. Behav. Soc. Netw.* **2017**, *20*, 346–354. [CrossRef]
70. Krasnova, H.; Abramova, O.; Notter, I.; Baumann, A. Why phubbing is toxic for your relationship: Understanding the role of smartphone jealousy among "Generation Y" users. In Proceedings of the Twenty-Fourth European Conference on Information Systems (ECIS), İstanbul, Turkey, 12–15 June 2016.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.