engineering proceedings

Proceeding Paper

Decision-Tree-Computing-Based Usage Intention Prediction of School Social Media †

Huai-Te Huang [1] and Hao-En Chueh [2,*]

1 Department of Industrial Education, National Taiwan Normal University, Taipei 106209, Taiwan; whiter0819@gmail.com
2 Department of Information Management, Chung Yuan Christian University, Taoyuan 320314, Taiwan
* Correspondence: hechueh@cycu.edu.tw

† Presented at the 3rd IEEE International Conference on Electronic Communications, Internet of Things and Big Data Conference 2023, Taichung, Taiwan, 14–16 April 2023.

Abstract: Social media is the main channel that teenagers use to exchange information. The purpose of this study is to construct a prediction model of school social media usage intention. To collect training data for modeling, we conducted a questionnaire survey on students of a senior high school in Taoyuan City, Taiwan. The training data processing method was decision tree computing. In this study, the decision tree computing software, Weka, was used to analyze the training data to extract the key factors affecting high school students' intentions to use school social media. The research results showed that perceived usefulness was the most important factor affecting school social media usage intentions, and trust was the second most important factor affecting school social media usage intention. The prediction model was proposed in this study to predict students' intentions to use school social media. It serves as a guide for schools to use social media as a channel for distributing important information.

Keywords: social media; decision tree computing; prediction model; usage intention

check for updates

Citation: Huang, H.-T.; Chueh, H.-E. Decision-Tree-Computing-Based Usage Intention Prediction of School Social Media. *Eng. Proc.* **2023**, *38*, 16. https://doi.org/10.3390/engproc2023038016

Academic Editors: Teen-Hang Meen, Hsin-Hung Lin and Cheng-Fu Yang

Published: 20 June 2023

1. Introduction

Following the advancement and innovation of information technology (IT), more social media platforms [1], such as Facebook (FB), Instagram (IG), and Twitter, have been launched. Users not only obtain information, receive comments, and share information via these social media, but they also promote interpersonal relationships through them.

Among all social media, FB is the most used and its users are of all ages. The emergence of FB was due to its use by young people, but gradually more and more middle-aged people are using it, as extensive news and content is available on FB regardless of the correctness of the information. This is convenient for businesspeople and office workers who need to monitor the news every day. Although the use time of young people is decreasing, many of them still use FB to read the news or receive messages. Thus, FB is still the most prominent social media platform [2,3].

IG is a popular platform for the younger generation, and almost everyone in the younger generation has an IG account. In other words, the usage time that is slowly decreasing on FB is being transferred to the IG platform. This generation of young people uses social media to build their styles, share their lives, and allow more people to know and see them. IG is the type of platform that best allows young people to express their characteristics [4,5].

Twitter is also a platform used mostly by young people, especially in the United States and Japan. Unlike IG, which contains mainly pictures, Twitter is a text-based platform, and much of the content is not about sharing knowledge, but simply expressing emotions. Hence, many politicians and entertainers like to use the platform to express their opinions.

Compared to IG, which is mainly used for personal image management, Twitter is more expressive of the user's real personality [6].

According to analysis of the three most popular social media, it is evident that the main functions of FB are providing information and news. IG is becoming more commercialized every year as the number of young users increases. Twitter has more text-based communication than the above two, but it is also free in terms of usage, leading to difficulties in controlling the dimensions and quality of content [7].

In recent years, the issue of digital transformation has been widely discussed, and every industry, including schools, continues to invest in IT hardware and software to reform education and strengthen students' knowledge in response to the digital trend. In addition, schools are using various IT platforms to collect and analyze student data and actively communicate with parents and students to provide student learning profile data. The latter enables parents to understand important information on students and schools. Such data not only enhance parents' and students' recognition of the school but also serve as an important resource for school development and marketing [8].

Because of the functions and features of FB, many schools are using it as a channel to disseminate important information to parents and students to inform them of the latest policies, events, and school-related educational data. If parents and students can access important school information on FB in real time, the information asymmetry between schools and parents and students reduces.

Information asymmetry means that the parties involved in a transaction do not have the same information that affects the transaction, and the party with information advantage uses inappropriate means to gain more benefits, thereby impairing the interests of the party with an information disadvantage. Even if the party with an information advantage does not intend to hide important information, the party with an information disadvantage may not have enough information to make the best decision for itself. When information asymmetry occurs among schools, parents, and students, it affects the responsibility for the educational performance of schools and the right of choice of parents and students [9].

Although social media is an important channel for adolescents to obtain and transmit information, it is not known whether they are willing to use this channel to obtain important information from schools. Therefore, it is necessary to construct a prediction model of school social media usage intention to extract the key factors that influence high school students' use of school social media as a channel to obtain important information.

To construct the prediction model of school social media usage intention, the decision tree computing software, Weka, was used to extract the key factors affecting the usage intention of school social media. The rest of the paper is organized as follows: In Section 2, the concepts of the usage intention model and decision tree computing are introduced. In Section 3, the research method is presented. In Section 4, the computing results and the discussions are displayed. The conclusions are given in Section 5.

2. Literature Review

2.1. Usage Intention Model

The basic theoretical model of the Technology Acceptance Model (TAM) was developed by Davis et al. [10] based on the Theory of Reasoned Action (TRA), which analyzes users' intention to use new information technology (IT) based on users' perceived usefulness (PU) and perceived ease-of-use (PEOU) of IT. The TAM is used as the theoretical basis for many empirical studies, and a considerable amount of empirical support has accumulated for it. It is also used to evaluate and predict users' acceptance of new IT systems, and therefore the model is widely used in IT-related research.

PU refers to the user's psychological perception of whether a particular system can make their work more efficient. The higher the PU, the higher the intention to use the system. PEOU refers to the user's perception of the ease of use of a particular system, and the higher the PEOU, the higher the intention to use the system. Attitude refers to the positive or negative evaluation of an individual's performance of a specific behavior.

According to TAM, attitudes are influenced by both PU and PEOU. When users perceive the system to have higher PU and PEOU, their attitudes toward the system tend to be positive. Usage Intention is the degree to which users are willing to use a particular system, and according to TAM, usage intention is influenced by both attitude and PU at the same time.

The Theory of Planned Behavior (TPB) [11] was proposed by Ajzen to explain how people change their behavioral patterns. TPB assumes that human behavior is the result of deliberate planning, and the variables in its theoretical model include behavioral attitudes, subjective norm (SN), perceived behavioral control (PBC), and behavior intention (BI). SN refers to the perceived positive or negative perceptions by significant others that an individual performs a particular behavior. PBC refers to the degree of control or mastery that an individual expects to have when adopting a particular behavior.

Trust is a key issue in the relationship between people and technology because it is one of the most important factors influencing users' use of any information technology, especially in the use of social media. When users use social media, they unknowingly reveal personal information, and social media platforms also collect this information from users all the time. Therefore, whether social media can gain the trust of users and effectively protect their information affects their intention to use social media [12].

Based on the above, we suggest that the five factors of PU, PEOU, SN, PBC, and trust have a significant impact on students' intentions to use school social media as an information access channel.

2.2. Decision Tree Computing

Machine learning (ML) is an important branch of artificial intelligence (AI) that focuses on building systems that can learn or improve performance based on the data they use. ML algorithms look for patterns and associations in large amounts of data and make optimal decisions or predictions based on the patterns and associations identified. The algorithms continue to improve as they learn, and the more data they use, the more accurate they become. ML consists of different types of learning models, and the user can decide to use supervised or unsupervised learning models depending on the nature of the data and the desired outcome [13].

The main techniques of ML include classification analysis, prediction analysis, cluster analysis, association rule analysis, and sequential pattern analysis [14,15]. The main purpose of classification analysis is to categorize a new paradigm, which does not have an explicit category, into a predefined category. This technique has been widely used to solve problems in various fields, such as identifying the key attributes of each type of user from their past usage records to find out what attributes make them loyal users and potential churners.

The main purpose of prediction analysis is to use information or conditions recorded in the past or at the present stage to determine possible future outcomes. Most of the methods used for classification analysis can also be used for prediction analysis. While classification analysis is used to make judgments about the present, prediction analysis is used to make judgments about the future. Cluster analysis is the analysis of the similarity of a large number of records to produce a combination of clusters. The combination of clusters produced by cluster analysis is characterized by the high similarity between records within a cluster and the low similarity between records in different clusters. Cluster analysis is mainly used in situations where it is unclear how to classify the data, so it is suitable for analyzing medical images and social networks, or for finding anomalies.

Association rule analysis is mainly used to identify possible relationships between attributes in a database. It is commonly used in the analysis of sales databases. Hence, it is referred to as shopping basket analysis in the field of marketing data science, which focuses on identifying combinations of items in sales databases that are frequently purchased by customers at the same time. Sequential pattern analysis is a technique similar to association rule analysis, which emphasizes the temporal order relationship of the set of items. Therefore, the main purpose of sequential pattern analysis is to identify the sequential relationship between the sets of items that frequently appear in the database during a specific time interval.

This study aimed to predict the intention of high school students to use school social media as an important channel to obtain school information. Therefore, the questionnaire data of high school students' usage intention on school social media was analyzed by using the decision tree technique to build a decision-tree-based usage intention model of school social media.

The decision tree technique is a supervised learning method in ML that is mainly used to deal with classification problems. Decision trees use the hierarchical conception of an inverted tree structure to express the classification process. Starting from the root node at the top level, a research variable is selected at each node, and the data are divided into subsets based on the possible values of the research variable to form the next level of the hierarchy. This process stops when there are no more suitable research variables to be selected or the value of the target variable can be decided. Common decision tree techniques include ID3, C4.5, and CART. ID3 is the first proposed decision tree algorithm, which uses information gain to determine the selected research variable at each node of the tree structure. C4.5 is an improved version of ID3, which does not use the information gain directly but uses the information gain ratio as the basis for research variable selection. CART can be used for both classification and regression problems with the Gini coefficient instead of the information entropy model. In this study, C4.5 was used to construct the prediction model of the school social media usage intention for high school students.

3. Research Method

3.1. Training Data for Modeling

To construct the prediction model of school social media usage intention, we conducted a questionnaire survey of the students in a senior high school in Taoyuan, Taiwan. The questionnaire was divided into two parts. The first part was used to understand the background information, including gender, age, frequency of visiting FB, awareness of the school's official FB fan page, and frequency of visiting the school's official FB fan page. The second part of the questionnaire consisted of six items on the six variables of this study (perceived usefulness, perceived ease-of-use, subjective norm, perceived behavioral control, trust, and usage intention). Each questionnaire item was measured on a five-point Likert scale with the numbers 5 to 1, with 5 indicating strongly agree and 1 indicating strongly disagree.

Before the questionnaire survey, each participant was asked to browse the school's official FB fan page on their smartphones for one minute and then to fill out the questionnaire. The questionnaire was administered from 1 June to 5 July 2022, and completed questionnaires were obtained from 223 participants.

3.2. Decision Tree Computing Software

Weka was used in this study to analyze the collected questionnaire data. Weka is a free software that provides various ML technologies, including data pre-processing, classification analysis, cluster analysis, and association rule analysis, and visualizes the data. Weka was developed by the University of Waikato, and is distributed under the GNU License in addition to open source [16].

There are two ways to use Weka: one is to analyze the data directly with the GUI of the Weka software, and the other is to analyze the data by calling the library provided by Weka in the program code. This study used the GUI of Weka to analyze the data directly (Figure 1). To do this, in the GUI of Weka, the Explorer option is chosen to enter the preprocess tab page (Figure 2), then the Open file option is selected in the Preprocess tab page to enter the questionnaire data to be analyzed. After entering the questionnaire data to be analyzed (Figure 3), the user proceeds to the decision tree analysis.

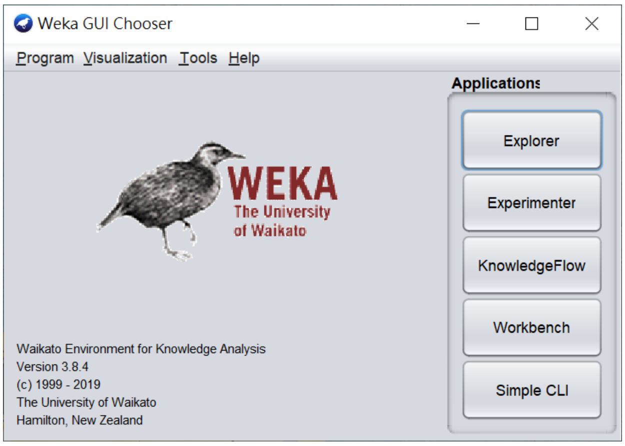

Figure 1. GUI of Weka.

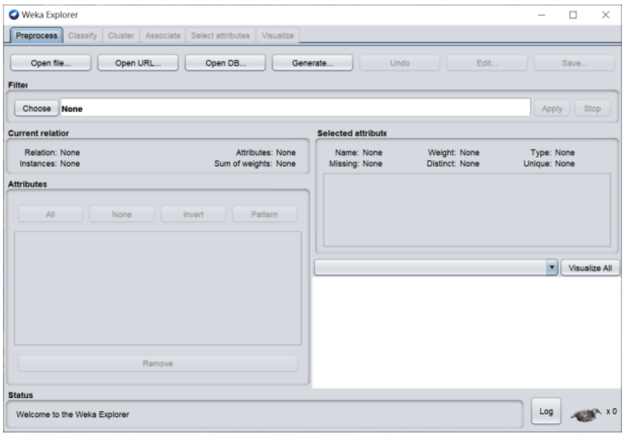

Figure 2. Preprocess tab page.

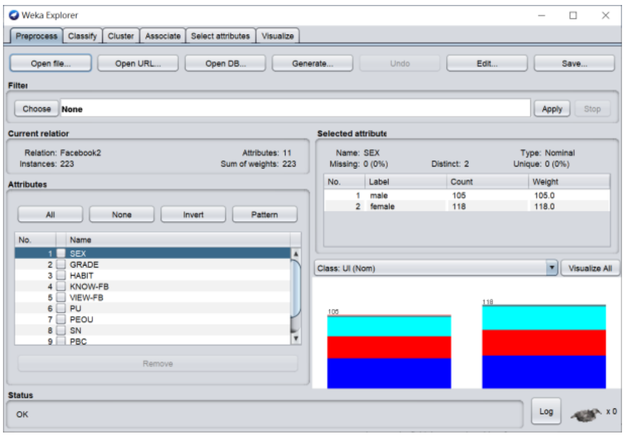

Figure 3. Preprocess tab page after data input.

In Weka, the decision tree analysis is in the Classify tab page (Figure 4), and the decision tree algorithm can be selected by clicking the Choose option on the Classify tab page. In the options of the decision tree algorithm of Weka, the J48 algorithm is the C4.5 decision tree algorithm. After selecting the J48 algorithm, the Test options are set to use the training set, and the target attribute is selected as usage intention (UI), then the Start option is selected to start the decision tree analysis (Figure 5).

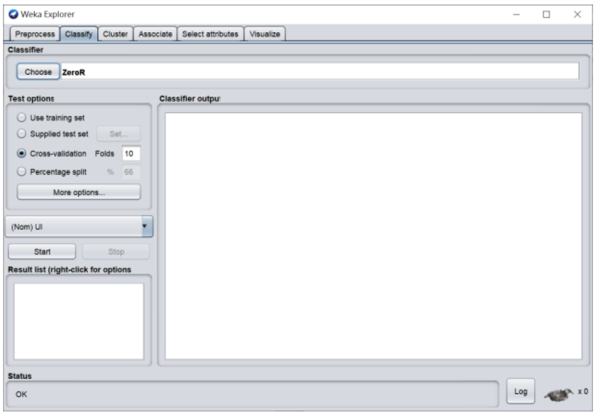

Figure 4. Classify tab page.

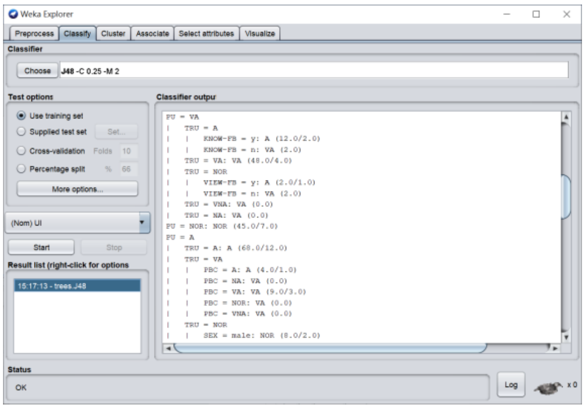

Figure 5. Classify tab page after executing J48.

4. Computing Results and Discussions

4.1. Description of Training Data

The training data for modeling are shown in Tables 1 and 2. According to Table 1, the percentage of students who frequently visited FB did not reach 50%, but most students were aware that the school had an official FB fan page, and most of them had visited the school's official FB fan page.

Table 1. Background information of participants.

Question	Option	Frequency
Gender	Male	105
	Female	118
Age	15 years old	102
	16 years old	71
	17 years old	5
	18 years old	45
Do you visit Facebook frequently?	Strongly disagree	9
	Disagree	16
	Normal	102
	Agree	46
	Strongly agree	50
Do you know that your school has an official Facebook fan page?	Yes	199
	No	24
Have you visited the school's official Facebook fan page?	Yes	170
	No	53

Table 2. Usage intentions of participants.

Question	Option	Frequency
Perceived usefulness (PU): I think it is useful to get important information about the school from school social media.	Strongly disagree	3
	Disagree	4
	Normal	54
	Agree	105
	Strongly agree	66
Perceived ease of use (PEOU): I think it is easy to use school social media to get important information about the school.	Strongly disagree	1
	Disagree	8
	Normal	63
	Agree	84
	Strongly agree	67
Subjective norm (SN): Friends and family around me think I should use school social media to get important school information.	Strongly disagree	3
	Disagree	3
	Normal	66
	Agree	92
	Strongly agree	59
Perceived behavioral control (PBC): I think I can get important school information from school social media.	Strongly disagree	2
	Disagree	4
	Normal	55
	Agree	86
	Strongly agree	76
Trust (TRU): I think it is trustworthy to get important information about the school from school social media.	Strongly disagree	1
	Disagree	4
	Normal	63
	Agree	91
	Strongly agree	64
Usage Intention (UI): I would like to receive important school information from school social media.	Strongly disagree	2
	Disagree	5
	Normal	67
	Agree	89
	Strongly agree	60

According to Table 2, the proportion of students who strongly agreed with PBC was the highest, followed by the proportion of students who strongly agreed with PEOU. This showed that students generally had no difficulty in using school social media to access important information about the school, either in terms of using the system or searching for information. A possible reason is that nowadays, high school students are generally digital natives, who have grown up in the era of the rapid development of IT. Their information skills are better than those of the previous generations of X and Y. Therefore, most high school students think they can quickly learn to use any information platform.

4.2. Decision-Tree-Computing-Based Usage Intention Prediction

The prediction model of school social media usage intention based on decision tree computing is shown in Figure 6. According to Figure 6, five decision rules can be obtained.

Rule 1: if perceived usefulness is "strongly agree" and trust is "strongly agree", then the usage intention is mostly "strongly agree".

Rule 2: if perceived usefulness is "strongly agree" and trust is "agree", then the usage intention is mostly "agree".

Rule 3: if perceived usefulness is "agree", trust was "strongly agree", then the usage intention is mostly "strongly agree" or "agree".

Rule 4: if perceived usefulness is "agree", trust is "agree", then the usage intention is mostly "agree".

Rule 5: if perceived usefulness is "agree", trust is "agree", and gender is female, and PBC is "strongly agree" or "agree", then the usage intention is mostly "agree".

```
PU = VA
|   TRU = VA: VA (48.0/4.0)
|   TRU = A: A (12.0/4.0)
|   TRU = NOR
|   |   VIEW-FB = y: A (2.0/1.0)
|   |   VIEW-FB = n: VA (2.0)
PU = A
|   TRU = VA
|   |   PBC = VA: VA (9.0/3.0)
|   |   PBC = A: A (4.0/1.0)
|   TRU = A: A (68.0/12.0)
|   TRU = NOR
|   |   SEX = male: NOR (8.0/2.0)
|   |   SEX = female
|   |   |   PBC = VA: A (1.0)
|   |   |   PBC = A: A (6.0/1.0)
|   |   |   PBC = NOR: NOR (5.0/1.0)
|   |   |   PBC = NA: NOR (1.0)
|   TRU = NA: NOR (3.0/1.0)
PU = NOR: NOR (45.0/7.0)
PU = NA: NOR (4.0)
PU = VNA: NOR (3.0/2.0)
```

Figure 6. Prediction model of school social media usage intention.

The above results show that perceived usefulness is the key factor that influenced high school students' usage intention of school social media, while trust is the secondary factor. Therefore, if schools want to use social media as the main channel to announce important information, they need to take measures to ensure that students think that it is useful to learn important information from school social media. In addition, schools must protect students' personal information and not allow social media to become a channel for the disclosure of students' personal information.

5. Conclusions

With the rapid advancement of technology and the greatly reduced costs of computing and storage, diversified data, such as images, text, sound, and maps, can be provided to parents in compliance with regulations. This allows schools to market themselves or provide parents and the public with a clear overview of the school to increase their recognition of the school and build satisfaction and loyalty to the school brand. In the past, when there was no low birth rate and information was relatively unavailable, most schools focused only on the quality of teaching and learning, attaching importance to student growth and teacher professionalism. However, with the professionalization of education, many schools are researching school administration, emphasizing data, and using the empirical evidence provided by data and information for decision-making and management. With the rapid development of new media, smartphones have become an ideal carrier for schools to deliver valuable information to parents and students through social media in a fast and low-cost way. We constructed a prediction model of school social media usage intention and produced five decision rules. The findings of this study can be used to predict students' intentions to use school social media, and can also be used as a guide for schools to use social media as an important channel for disseminating information.

Author Contributions: Conceptualization, H.-T.H. and H.-E.C.; methodology, H.-T.H.; software, H.-E.C.; investigation, H.-T.H.; writing—original draft preparation, H.-T.H. and H.-E.C.; writing—

review and editing, H.-T.H. and H.-E.C.; visualization, H.-E.C. 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: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Carr, C.T.; Hayes, R.A. Social media: Defining, developing, and divining. *Atl. J. Commun.* **2015**, *23*, 46–65. [CrossRef]
2. Levy, R.E. Social media, news consumption, and polarization: Evidence from a field experiment. *Am. Econ. Rev.* **2021**, *111*, 831–870. [CrossRef]
3. Lamot, K. What the metrics say. The softening of news on the Facebook pages of mainstream media outlets. *Digit. J.* **2022**, *10*, 517–536. [CrossRef]
4. Wahid, R.M.; Gunarto, M. Factors driving social media engagement on Instagram: Evidence from an emerging market. *J. Glob. Mark.* **2022**, *35*, 169–191. [CrossRef]
5. Fagundes, L.S.; Marot, T.A.; Natividade, J.C. Use of Instagram, social comparison, and personality as predictors of self-esteem. *Psico-USF* **2021**, *25*, 711–724. [CrossRef]
6. Foroughi, B.; Griffiths, M.D.; Iranmanesh, M.; Salamzadeh, Y. Associations between Instagram addiction, academic performance, social anxiety, depression, and life satisfaction among university students. *Int. J. Ment. Health Addict.* **2021**, *20*, 2221–2242. [CrossRef]
7. Rhee, L.; Bayer, J.B.; Lee, D.S.; Kuru, O. Social by definition: How users define social platforms and why it matters. *Telemat. Inform.* **2021**, *59*, 101538. [CrossRef]
8. Dhiman, D. The Practice of Media Education and Media Research: A Review on Five Asian Countries. The Practice of Media Education and Media Research: A Review on Five Asian Countries. *Glob. Media J.* **2021**, *19*, 1–7.
9. Goldstein, A.; Eaton, C. Asymmetry by design? Identity obfuscation, reputational pressure, and consumer predation in US for-profit higher education. *Am. Sociol. Rev.* **2021**, *86*, 896–933. [CrossRef]
10. Davis, F.D.; Bagozzi, R.P.; Warshaw, P.R. User acceptance of computer technology: A comparison of two theoretical models. *Manag. Sci.* **1989**, *35*, 982–1003. [CrossRef]
11. Ajzen, I. The theory of planned behavior. *Organ. Behav. Hum. Decis. Process.* **1991**, *50*, 179–211. [CrossRef]
12. De Meo, P.; Messina, F.; Rosaci, D.; Sarné, G.M. Combining trust and skills evaluation to form e-Learning classes in online social networks. *Inf. Sci.* **2017**, *405*, 107–122. [CrossRef]
13. Badillo, S.; Banfai, B.; Birzele, F.; Davydov, I.I.; Hutchinson, L.; Kam-Thong, T.; Siebourg-Polster, J.; Steiert, B.; Zhang, J.D. An introduction to machine learning. *Clin. Pharmacol. Ther.* **2020**, *107*, 871–885. [CrossRef]
14. Mahesh, B. Machine learning algorithms—A review. *Int. J. Sci. Res.* **2020**, *9*, 381–386.
15. Ray, S. A quick review of machine learning algorithms. In Proceedings of the 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing, Faridabad, India, 14–16 February 2019.
16. Attwal, K.P.S.; Dhiman, A.S. Exploring data mining tool-Weka and using Weka to build and evaluate predictive models. *Adv. Appl. Math. Sci.* **2020**, *19*, 451–469.