Article

Usability and Security Testing of Online Links: A Framework for Click-Through Rate Prediction Using Deep Learning

Robertas Damaševičius [1,*] and Ligita Zailskaitė-Jakštė [2,3]

1   Department of Applied Informatics, Vytautas Magnus University, LT-44404 Kaunas, Lithuania
2   Department of Multimedia Engineering, Kaunas University of Technology, LT-51386 Kaunas, Lithuania; ligita.zailskaite@ktu.lt
3   Institute of Business and Economics, Mykolas Romeris University, LT-08303 Vilnius, Lithuania
*   Correspondence: robertas.damasevicius@vdu.lt

Abstract: The user, usage, and usability (3U's) are three principal constituents for cyber security. The effective analysis of the 3U data using artificial intelligence (AI) techniques allows to deduce valuable observations, which allow domain experts to design practical strategies to alleviate cyber-attacks and ensure decision support. Many internet applications, such as internet advertising and recommendation systems, rely on click-through rate (CTR) prediction to anticipate the possibility that a user would click on an ad or product, which is key for understanding human online behaviour. However, online systems are prone to click on fraud attacks. We propose a Human-Centric Cyber Security (HCCS) model that additionally includes AI techniques targeted at the key elements of user, usage, and usability. As a case study, we analyse a CTR prediction task, using deep learning methods (factorization machines) to predict online fraud through clickbait. The results of experiments on a real-world benchmark Avazu dataset show that the proposed approach outpaces (AUC is 0.8062) other CTR forecasting approaches, demonstrating the viability of the proposed framework.

Keywords: human-centric cyber security; CTR prediction; factorization machine; artificial intelligence; deep learning

1. Introduction

Novel human–computer interaction systems are being continuously developed, and new communication possibilities are emerging. The digitization of diverse information is becoming a usual pattern, and with technological advancements, human commands or statements can be processed in computer systems. Automatic chatbots, question–answer systems, and other software solutions that facilitate human–machine interactions (HMI) require adequate methods for ensuring the security of the system [1]. The creation of user interfaces for mobile devices is difficult because it addresses specialized mobile device usage—users demand quick replies to their activities while limiting the quantity of information submitted. Because of their nature, mobile devices are subject to several limits, such as restricted display size and motoric limitations on information entry. Good usability is a key need for interface design and is essential for attracting and maintaining users, particularly in commercial applications such as mobile e-commerce platforms [2].

The greater emphasis on a human-centric design [3] is the spotlight of current cyber security research. The emphasis is shifting towards integrating human behaviour and cognitive reasoning to achieve human-centric cyber security that can secure humans and institutions from the malignant aftereffects of cyberattacks, whilst also being in sync with human cognitive and behavioural patterns. Human-centric cyber security (HCCS) is a nebulous notion that is difficult to describe due to the intrinsic link between humans and technology, as well as individuals and security systems; however, the emphasis is on the link between online user behaviour and its implications. Therefore, the methods

for predicting the behaviour of users when interacting with internet systems, expressed through a variety of quantitative metrics such as click-through rate (CTR), are relevant [4].

Usability means how easily a product or service can be accessed, how user-friendly it is, how well or efficiently it performs its function, and how prone it is to errors (international standard ISO 9241-11:2018). A website with a high level of usability produces a positive attitude and greater satisfaction than one with less usability, information content, customer service, and perceived security [5]. Usable design must play a part in creating the effective business use of social media. The popularity of social media posts and materials can be increased by using relevant keywords, attractive image colours [6], and other search engine optimization (SEO) techniques. Usability is also related to credibility. It has been shown that highly credible websites reached a significantly higher CTR on sponsored content, as their credibility facilitated online visitors to explore the website for longer [7].

Usability and security issues are closely related [8,9]. Usability is also closely related to the security of the safe use of apps and browsing on the internet. For example, phishing warnings designed with usability in mind have significantly reduced CTR compared to other types of warning. This shows that usability-aware warning placement has a significant impact on phishing warning adherence [10]. The user should also be able to make security-aware decisions, and assess the intention of the sender and the possible cyber security implications if they continue with clicking a link on a website or social media. Moreover, as underscored by Franz et al. [11], any anti-malware or anti-phishing solution must have high usability to be used by the users effectively. The attractiveness and usability of apps or websites can be increased by introducing motivation-enhancement and supporting measures such as gamification [12]. The understanding of user experience (UX) integrates into real-world usable privacy and security research studies that are aimed at changing risk-taking behaviour, thus increasing the security of critical applications [13].

Forecasting the likelihood of users selecting (i.e., clicking on) ads or products (also known as predicting click-through rates) is an important issue for many web applications such as online advertising and recommendation systems [14]. The effectiveness of the forecast has a direct impact on the end revenues of service providers. It is generating an increasing interest in both academia and commercial circles. Machine learning (ML) plays a defining role in predicting CTR, with item attributes and user profiles, and as input functions. Understanding human online behaviour based on contextual variables requires accurate CTR prediction. The purpose of CTR prediction is to estimate the likelihood of people clicking ads or objects, which is important for many internet applications such as online advertising and recommendation systems. There are numerous factors that influence CTR, but because certain factors may have less impact than others, it is crucial to determine the factors that are important for boosting click-through rate estimation. The success of CTR prediction is dependent on the modelling of complex feature relationships [15]. It is challenging to learn the complicated feature relationships that underpin user behaviour. The experts' past expertise can be used to develop low-order feature interactions. The typical CTR prediction approach is heavily reliant on feature design. Data features are arbitrarily chosen and analysed. The data have a complicated mapping connection, particularly for important data, and it is critical to also consider feature interactions. Many effective techniques rely heavily on manually manufacturing combinatorial features [16], that is, building new features by mixing numerous predictor variables, often called cross-features. The importance of such features may incur a great cost because designing effective features involves extensive technical work and domain understanding [17–19].

In this paper, we follow the vision of Grobler et al. [20] by acknowledging user, usage, and usability (3U's) as the three principal elements of cyber security. In the cyber security domain, the effective analysis of the data using artificial intelligence techniques allows to deduce high value insights that allow domain experts to design practical strategies to mitigate cyberattacks and ensure decision support [21,22]. Many internet applications, such as online advertising and recommendation systems, rely on CTR prediction to anticipate the possibility that a user would click on an ad or product. This problem is extremely

difficult because: (1) input functions (e.g., user age and item category, etc.) are typically sparse; and (2) effective inference relies on high-order functions, which are difficult to process manually by domain experts and are not enumerable.

In this paper we present a novel artificial intelligence-based model form modelling security and usability interactions in web applications. The proposed model, called the Human-Centric Cyber Security (HCCS) model, extends the user, usage, and usability (3U's) model [20] with a deep learning methodology that can be used for CTR prediction. The model is evaluated on real-world CTR data using the Avazu benchmark dataset.

The remaining parts of the paper are outlined in the following. Section 2 discusses the relevant work. Section 3 introduces and explains our methodology. Section 4 discusses the results. Finally, Section 5 formulates the conclusions and overviews the future efforts.

2. Related Work

2.1. Machine Learning Techniques for CTR Prediction

Many models have been proposed by researchers for CTR prediction, most of which employ the ML techniques. A logistic regression (LR) model was adopted to forecast CTR on Google Ads [23]. The online sparse learning technique was introduced to train the model using numerous aspects of characteristics such as user data, query keywords, advertising information, and metadata with ads as the input of the model. Chapelle [24] introduced an ML framework based on LR with the goal of predicting Yahoo's CTR. The absence of data, especially for new ads, is one of the most significant issues in CTR prediction. The previous display data are insufficient to offer a click-through model with a pre-calculated benchmark. Graepel et al. [25] analysed the sponsored search application scenario, the critical importance of CTR prediction in general, and the task-specific limitations, such as accuracy, calibration, scalability, dynamics, and exploration. They introduced the adPredictor algorithm, a new Bayesian online learning system for binary prediction. The algorithm uses a generalised linear model and a factorising Gaussian belief distribution on feature weights to calculate the approximate posterior, resulting in the automatic learning rate adaptation. Guo et al. [26] proposed a numerical feature embedding learning framework (AutoDis) for CTR prediction with unique representation attributes retained. Discretization, meta-embeddings, and aggregation are the three major components of AutoDis. For each numerical field, they offer meta-embeddings to learn knowledge from the perspective of a field with a relatively small number of parameters. Then, soft discretization is performed, capturing the correlations between numerical characteristics and meta-embeddings.

2.2. Deep Neural Networks for CTR Prediction

High-order feature interactions might be harder to grasp. Deep neural networks (DNNs) can learn these deep-feature connections. Recently, DNNs have gained popularity because of their ability to learn high-order features. DNNs are promising in user response prediction because they can investigate high-order hidden patterns and automatically develop more informative feature representations to deliver higher prediction performance. Zhang et al., for example, investigated feature mappings and developed the FM-supported neutral network (FNN) [27]. To choose extremely important features, An et al. [28] employ extreme gradient boosting (XGBoost) for the feature engineering phase. In actual mobile advertising circumstances, the selected characteristics are mobile contextual factors such as time contextual, geographical contextual, and other contextual attributes. The XGBoost deep FM-supported neutral network (XGBDeepFM) combines the capabilities of XGBoost for feature selection, FM for cross-feature interaction, and DNN for feature learning. Zhang et al. [29] developed an embedding approach based on the pre-trained FM. Multi-layer perceptrons (MLPs) were created to investigate higher-order features based on the concatenated embedding vectors. The FM, on the other hand, severely limits the quality of embedding initialization. Product-based neural network (PNN) was introduced by Qu et al. [30] to capture high-order feature interactions by inserting a product layer after the embedding layer but before the fully-connected layer. Furthermore, PNN, in the

same manner as other deep models, captures low-order features, which are critical for CTR prediction. DeepFM [31] imposed an FM unit in the Wide and Deep architecture [32], with no pre-training or feature engineering, and highlighted the interplay of low- and high-order features. The combination of the factor decomposition machine and the deep neural network is a characteristic of this model. DeepIntent used a word attention technique for query and ad representations to increase the quality of the learnt semantic match and to capture query intent [33]. Deep Crossing presented a deep architecture to estimate click likelihood by learning elements of an impression such as query text, ad text, or keywords automatically [34]. Deep Character-level Model [35] and Match-Tensor [36] presented very deep dual architectures for search query, and embeddings with a matching layer to learn the deep features for CTR forecasting.

Zhou et al. created a unique technique, termed the Deep Interest Network (DIN), by constructing a local activation unit, which uses the attention mechanism to learn significant sections of past actions and which uses weighted sum pooling to produce a mapping of a user's interest [37]. Higher active weights are assigned to behaviours that are more relevant to the candidate item in terms of representing a user's interest. Gligorijevic et al. used attention models to learn feature mappings that concentrate on important parts of queries and online ads, and introduce a supervised model that learns semantic embeddings, as well as their corresponding CTR [38]. Chen et al. [39] offer a deep belief nets (DBN)-based mobile advertising CTR estimation fusion model. The approach combines the strength of DBN's data mapping and feature extraction capabilities with LR models. To tackle the fundamental difficulty of acquiring the optimal feature expression in ML, our algorithm discovers deep features instead of basic original characteristics and then feeds them into an LR model to forecast the CTR.

Gan et al. [40] propose a recent-recurrent neural network (R-RNN), to divide a user's click-through (CT) behaviours into the global CT sequence and recent CT sequence, and then use an RNN to analyse the user's recent CT sequence, which roughly reflects the user's current interests. The RNN model also employs long short-term memory (LSTM) to address the issue of long-term reliance in the time series data of CT behaviours. In the work of Jiang et al. [41], the stacked autoencoder (SAE) is constructed using the unsupervised layer-by-layer pre-training to extract the features of input data, after which the CTR of the user can be obtained. Li et al. [42] proposed an attentive deep interest-based network model. The interest sequence is captured in the interest extractor layer, and the auxiliary losses are used to construct the interest state with deep supervision. They use a bidirectional gated recurrent unit (Bi-GRU) to describe the reliance between behaviours. Then they extract the target's interest evolving process and suggest an interest evolving layer. Simultaneously, the attention mechanism is incorporated in the sequential structure. The model then learns very non-linear feature interactions using stack autoencoders for CTR prediction.

Lu et al. [43] present a hybrid model for CTR prediction that combines the deep interest network (DIN) with the eXtreme Deep FM (xDeepFM). DIN employs an adaptive local activation unit that integrates an attention unit to learn user interest from past actions associated with certain adverts. A fundamental component of xDeepFM is a compressed interactions network (CIN), which aims to automatically produce feature interactions at a vector-wise level. xDeepFM is formed by combining a CIN, a simple DNN, and a linear component into a single unified model. The end-to-end hybrid model is an MLP-based parallel ensemble of models.

To create multi-scale features based on distinct receptive fields, Qiang et al. [44] propose using multi-scale stacking pooling (MSSP). By creating several observers with diverse angles and fields of vision, the structure stacks multi-scale characteristics bi-directionally from depth and breadth angles, assuring the variety of retrieved features. Furthermore, factorization is used for learning parameters, while the architecture ensures that high-order features are learnt well in sparse data. The authors next merge the MSSP with a traditional DNN to create DeepFM, a unified model.

Qin et al. [45] suggested a framework for user behaviour retrieval for CTR prediction (UBR4CTR). Using a learnable search strategy, the most informative user behaviours will be extracted from the whole user history sequence in UBR4CTR. The goal is to anticipate the CTR between a target user and a target item in a certain environment; all three pieces of information are merged to generate a prediction target. The prediction target comprises the target user features, target object, and context. To construct the final forecast, the collected behaviours are fed into a deep attention-based model.

Song et al. [46] present the dual-view attention network (DVAN), which enhances both the user-view and the item-view to acquire numerous interaction connections for CTR prediction. The cross-domain attention module extracts coarse-level correlations, while a local homogeneous-domain attention module adapts to extract fine-level correlations. A unique selection technique is developed to connect the cross-domain attention unit with the locally homogeneous attention unit, gradually building representation from the coarse to the fine level. To depict the nonlinear related relationship of features, Wang et al. [47] employ a stack autoencoder for higher-order feature interactions and enhanced FM for lower-order feature interactions. Xu et al. [48] propose an ideally connected deep belief net (OCDBN) for click prediction using a rotation code whitening technique based on optimum mean removal. OCDBN removes optimum mean to increase the accuracy of CTR prediction. OCDBN can extract global main characteristics from input instances with a variety of components, which may be used for both single and sequential ad impressions. Yang et al. [49] propose an embedding approach for learning feature mappings in user CTR forecasting systems, called operation-aware embedding, and build a new deep learning model, called operation-aware neural networks (ONN). In contrast to the standard feature embedding approach, which learns a single mapping for all operations, operation-aware embedding may learn many mappings for diverse operations.

2.3. Factorization Machines

Two-order feature interactions have been studied using factorization machines (FMs). Factorization machines (FMs) [50,51] are supervised ML approaches that embed features in a latent manifold and model feature interactions using the inner product of their embedding vectors. Factorization machines (FM), combining polynomial regression models with factorization methods, were designed to model interactions of features, and have proven their effectiveness for various problems [52,53]. The CTR prediction models based on 2nd polynomial degree mapping and FMs are commonly employed.

Following that, other variations of factorization machines were proposed. Juan et al. created the factorization-based prediction field-aware FMs [54]. Zhang et al. [55] use classic feature combination methods and deep neural networks (DNNs) to automate feature combinatorics to increase CTR forecast accuracy. They propose the field-aware neural FM (FNFM), in which a DNN is employed for higher-order feature learning. The model outperforms existing deep feature learning models such as DCN, DeepFM, and NFM in terms of expression ability. The importance of diverse interactions of second-order features was examined by GBFM [56] and AFM [57]. All of these techniques, however, are cantered on simulating the interactions of low-order objects. Some recent work has been conducted to imitate higher-order object interactions. Neural factorization machines (NFM) [58], for example, modelled higher-order objects by stacking DNNs using the output of second-order object interactions. Higher-order factorization machines (HOFM) [59] have suggested efficient learning algorithms for higher order FMs. Zhou et al. [60] propose DGRU, a novel hybrid neural network that combines factorization-machine-based neural network (DeepFM) with gated recurrent unit neural network (GRUNN). The DeepFM component oversees combining autonomic features, while GRU oversees modelling user interests and their evolution. The GRU is given a series of 1 and 0 values that reflect user click activities. It includes data on what consumers like and hate.

Summarizing, there are some research gaps in the previous studies as follows: the lack of context (the CTR prediction problem is usually treated out of context as a simple data

mining problem, thus ignoring the complex relationships with security and usability of web applications) and feature interaction (the modelling of feature interactions is still under-researched: when a neural network learns high-order features, it is unable to identify the importance and association of combined features, limiting the model's prediction accuracy even further).

3. Background and Methodology

3.1. Problem Description

Let us assume having n instances and m observations, $D = \{(x_i, y_i) | i = 1, \dots, n\}$, in which $X^{\mathrm{T}} = \left(x^{(1)^T}, \dots, x^{(m)^T} \right)$ and $y_i \in \{0, 1\}$ is the label of the i-th instance. For the CTR problem, 0 is the label of no click, and 1 is the label of click. The task is to find a mapping $f : R^{I_1} \times \dots \times R^{I_m} \to \{0, 1\}$. If $y_i \in R$, it can be solved as a regression problem.

3.2. Model Description

The term "user-centred security" is used to describe security models, procedures, systems, and software with usability as a main motivator or goal and was presented in [61]. At first, security was the main aspect of every system, but later, the instrumental approach has changed, seeking instead to involve the aesthetic elements which are associated with the quality of the technology. Almost all human–computer interaction (HCI) research has focused on achieving user behavioural goals [62]. As mentioned earlier, we use Grobler et al.'s [20] proposed components of cyber security 3Us (user, usage, and usability), which we have extended with the Artificial Intelligence (AI) elements, as follows:

User. The human who interacts with cyber systems for authorized objectives is considered by user components. Usually users are defined by demographic, psychographic, and behavioural characteristics offline. These characteristics are possible to use on-line as well. When there are not enough demographic characteristics, it is appropriate to use other characteristics such as user typology, Big Five, or consumer roles based on consumer actions on social media [63]. The user component of cyber security considers the varied spectrum of these users, each with varying cyber security awareness because of personal preferences, social characteristics, and past experiences, and user reaction toward cyber security hazards. The user characteristics are presented in Table 1.

Usage. The functional features of technological and non-technological safeguards put in place to secure users against explicit security threats are the focus of usage components [64]. Usage focuses on how a cyber-security system and procedures, such as antivirus products, spamming detection algorithms, business rules, and cyber security regulations, are intended to be used.

Usability. The usability of a system is determined by how well it can be evaluated by the end user. It gives an overview of how humans interact with technology. According to Voorveld et al. [65], the user experience in each internet platform is different because of the functionality of the differing internet pages and the user interface, etc. The usability combines technological and non-technological aspects. Gaver and Martin [66] proposed that technology should be used to serve a variety of non-instrumental demands, such as surprise, diversion, and closeness.

AI. The intelligence of the system is ensured by machine learning or deep learning methods such as deep neural networks. AI aids users to secure their systems during use while ensuring usability is maintained.

Table 1. Summary of the components of the Human-Centric Cyber Security model.

User Type	Definition	Reference
Legitimate user	The legitimate users' goals are to defend or use the system.	[20]
	Legitimate users know the system's files and where they can be found well.	[67]
Malicious user or attacker	Attacker gains access to data by exploiting legitimate users.	[20,68]
	Attackers exploit holes in IoT systems for personal gain.	[69]
Expert user	If a cyber danger is detected, it would use a variety of informative sources to alert legitimate users and sustain its continuity under attack.	[20]
	Experts know a lot about a subject and are familiar with how the discipline is structured. This involves the ability to comprehend and contribute to the discipline's vocabulary and technique.	[70]
Non-expert user	A non-expert user is unaware of the underlying functional, usability, and security principles, and can use the system only with a set of instructions given.	[71]

3.3. Click-Through Rate (CTR)

We demonstrate the role of AI techniques in online cyber security in relation to usability through the CTR prediction task.

Definition (prediction of CTR), let x R n represent the combination of user functions u and element characteristics v, where categorical attributes are encoded via one-time coding, and n is the size of the concatenated functions. The task of the likelihood that the user u will click on the element v is based on the feature vector x.

CTR is the ratio of the number of clicks on a specific link or an element of interface to the number of times people were exposed to the link or element. Prediction of CTR is critical in current online personalized services. To create an effective CTR prediction model, it is crucial to capture users' drifting interests by modelling successive user activities [45]. The user behaviour, especially for marketing specialists, is an object of careful analysis, because this issue influences business success, increased incomes, and reached companies goals. One of the essential user behaviour metrics for advertising revenue is the click-through rate (CTR). It is a short term, non-financial consumer behaviour measurement [6], which possibly links to financial results, and is also equally vital for increasing incomes. Many online applications, such as web search, recommender systems [14], and display advertising, rely on click-through rate (CTR) prediction. As a result, the ranking method can be changed to CTRbid across all options, where "bid" refers to the profit received by the system if a user clicks on an item. It is obvious that the purpose is to accurately estimate the CTR.

3.4. Factorization Machines

The method starts with a sparse feature vector x, which is then fed to a nesting layer that maps features (categorical and numeric) into the lower-dimensional manifold. The nesting of all fields is then introduced into a new interacting layer that is designed as a multipurpose (multilayer) self-tensioned neural network. Higher-order objects are combined using the attention mechanism for each interacting layer, and various types of combinations can be evaluated using multilayer mechanisms that map objects to different subspaces. We model different orders of combinatorial objects by superimposing multiple interacting layers (Figure 1). The final layer produces a low-dimensional mapping of the

input object that models high-order features and is then used to estimate the CTR with a sigmoidal function.
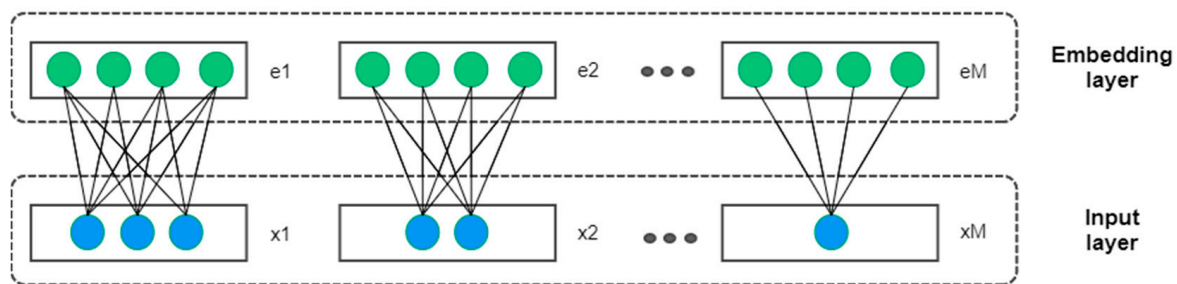
Figure 1. The conceptual structure of the factorization machine.

First, we present the user profiles and attributes of an element as a sparse vector, which aggregates all fields. Because the mappings of categorical features are usually very sparse, the generally accepted method is to map them in low-dimensional manifolds (for example, word embeddings). Often, categorical features can be multivalued. For compatibility with multivalued input data, we represent the multivalued feature field as the arithmetic mean of the feature embedding vectors.

To facilitate interactions between categorical and numeric attributes, we present numeric attributes in the low-dimensional attribute space as categorical attributes. As a result, the embedding layer's output will be a concatenation of several embedding vectors. After transferring categorical and numerical features to the same low-dimensional space, we proceed to modelling high-order features in this space.

The loss function is a loss log, which is defined as follows: where y are valid data on user CTR, and N is the count of training samples.

$$y = w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n} \sum_{j=i+1}^{n} \langle v_i, v_j \rangle x_i x_j, \tag{1}$$

This formula has one difference from the previous formula—the weights w_{ij} for multiplicative effects, features i and j are replaced by the dot product of the vectors corresponding to them. These vectors have dimension k, i.e., $v_i = (u_{i1}, u_{i2} \ldots u_{ik})$ and the larger the dimension is, the deeper the interactions will be learned by the model, but it will also take more time and is more likely to be retrained. Otherwise, this is the usual regression model. One question remains to be solved—how to select these vectors in linear time, because it is unclear from the above equation.

$$\frac{1}{2} \sum_{l=1}^{k} \left(\left(\sum_{i=1}^{d} v_{i,l} x_i \right)^2 - \sum_{i=1}^{d} v_{i,l}^2 x_i^2 \right), \tag{2}$$

The parameters are optimized using stochastic gradient descent (SGD) as follows.

$$\frac{\partial}{\partial \theta} \hat{y}(x) = \begin{cases} 1, & \text{if } \theta \text{ is } w_0 \\ x_i, & \text{if } \theta \text{ is } w_i \\ x_i \sum_{j=1}^{n} v_{j,f} x_j - v_{i,f} x_i^2, & \text{if } \theta \text{ is } v_{i,f} \end{cases}, \tag{3}$$

We use the cross-entropy (a.k.a. LogLoss) metric to calculate the loss per task. When a network output may be interpreted as an independent hypothesis, the final activation relates to the probability that the hypothesis is correct. It calculates the difference between the expected and empirical probability distributions. In addition, cross-entropy addresses the vanishing gradient problem, which is a concern for deep neural networks. It is calculated as follows for binary classification where the targets are either 0 or 1.

$$LogLoss = -\frac{1}{N} \sum_{i=1}^{N} (y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)), \tag{4}$$

where y_i and $\hat{y}_i$ are the ground truth and estimated interest of the user (candidate or corporation), respectively, and N is the total count of training examples.

During training, the model parameters are updated by using gradient descent to minimize the weighted LogLoss.

3.5. Neural Network Architecture

The neural network architecture used in this study is based on DeepFM [72]. The neural network model consists of two principal units: the factorization machine (FM), and the deep neural network (DNN), which share the same input, while the output of the embedding layer is fed into the DNN. The embedding layer is used to compress the input vector to a dense real-valued vector that is fed downstream to the first hidden layer. The FM unit can model both linear (1st order) and 2nd order interactions between features. The DNN unit is a feed-forward neural network (FFNN), which learns high-order feature combinations. A small part of the model showing some input, embedding layers, and FM and DNN units, is visualized in Figure 2. The model has 427,089,562 trainable parameters.
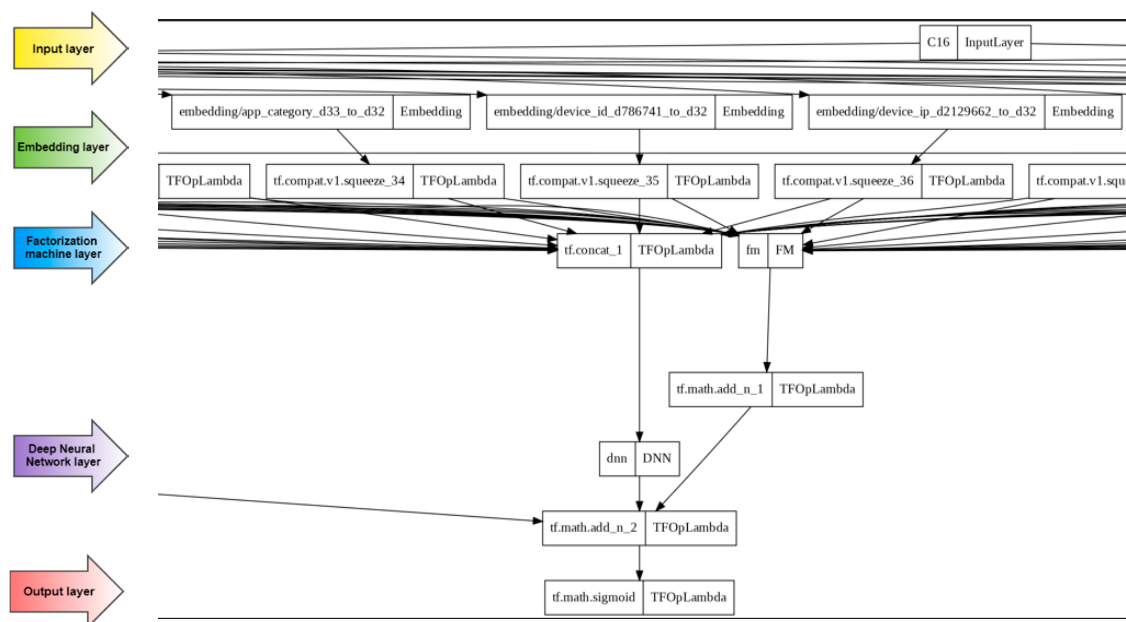

Figure 2. Part of the model's architecture.

The working of the model is summarized as Algorithm 1.

Algorithm 1 Pseudocode of factorization neural network

Input: Sparse feature vector.
Output: Predicted CTR value.
1. **Input layer** to input sparse feature vector into model.
2. **Embedding layer** to compress the input vector to a dense real-valued vector.
3. **Factorization machine** to compute hidden (latent) features.
4. A feed-forward **neural network** that learns high-order feature combinations.
5. **Output layer** with sigmoid function to predict the CTR value.

3.6. Evaluation

As in many other studies, we randomly partitioned the dataset into three parts: training (70%), validation (10%), and test (20%). The training set is utilized for model training, the validation set is used for hyper-parameter tweaking, and the test set is used for performance reporting. To assess the efficacy of various models, we employed two prominent measures that have been used in earlier studies.

LogLoss, also known as binary cross-entropy, measures the divergence of forecast probabilities from actual labels. LogLoss has a value between 0 and 1, and the label has a binary value.

Area under curve (AUC) expresses the likelihood that a binary classifier would assign a higher value to a randomly selected positive test case than to a randomly selected negative test instance. A higher AUC value denotes superior performance. When the output of the model is a number reflecting the likelihood of a particular data record belonging to a certain class, AUC and LogLoss are more useful than accuracy and recall. To compute precision and recall, probabilities are translated to class labels based on a user-defined threshold, and the threshold chosen might have a significant impact on the findings. AUC and LogLoss, on the other hand, can circumvent such thresholds.

4. Results and Discussion

4.1. Dataset and Settings

All methods are implemented in Python language using TensorFlow library on a desktop computer with Intel® Core i5-8265U 64 bit processor with 8 GB RAM and Windows 10 operating system.

Here we use the Avazu dataset for the experimental validation of the proposed methodology. The Avazu dataset is generated from Avazu's 2015 CTR prediction contest. The dataset is composed of 24 columns, most of which are categorical variables and a few of which are numeric identifiers. It has 40,428,967 samples (of which 33,563,901 are positive, and 6,886,566 are negative), 23 fields, and 1,544,488 features. The dataset is large, and, therefore, following [72], we use only 10% of the dataset for training and testing.

4.2. Results

The training process is illustrated in Figure 3. For optimization, we use the Adam optimizer, and for loss function, the binary cross-entropy. We stop training after 10 epochs, as the loss was decreasing very slowly.
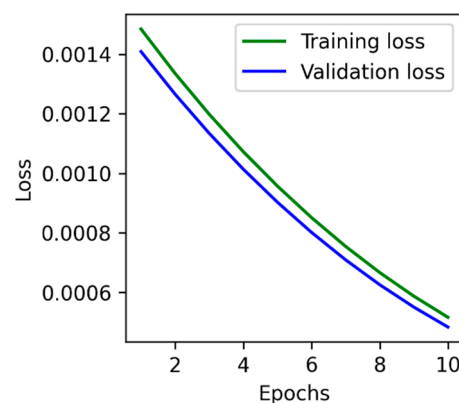

Figure 3. Training of the model.

The confusion matrix of the classification results is presented in Figure 4. The overall accuracy reached is 88.82%

4.3. Comparison with Related Work

We contrast our method with machine-based factorization methods that consider second-order features, and methods that can capture interactions of higher-order objects. The methods and their performance are summarized in Table 2 and explained below. FM [50] uses factorization methods to model the interaction of second-order functions. AFM [57] is one of the most modern models that captures interactions of second-order functions. It extends the FM, using the attention mechanism to categorize the importance of second-order combinatorial properties. DeepCrossing [16] uses deep, fully connected, resid-

ual neural networks to study nonlinear feature relationships in an implicit way. NFM [58] stacks neural networks on top of the second-order object interaction layer. The interactions of high-order features are indirectly fixed by the nonlinearity of neural networks. In CrossNet [34], the cross-net, which is the core of the Deep and Cross model, uses the external product of a concatenated feature vector at the bit level to explicitly model feature relationships. HOFM [59] offers kernel-based algorithms for learning higher order FMs. They build a third-order FM using a public implementation. AutoInt [73] maps both numerical and category characteristics into the same low-dimensional space. Then, to explicitly represent feature interactions in this space, a multi-head self-attentive neural network (SANN) with residual connections is developed. Multiple orders of feature combinations of input characteristics may be represented using different layers of multi-head SANNs. The compressed interaction network (CIN), which is the core of the xDeepFM model [74], takes the external product of the folded feature matrix at the vector level.
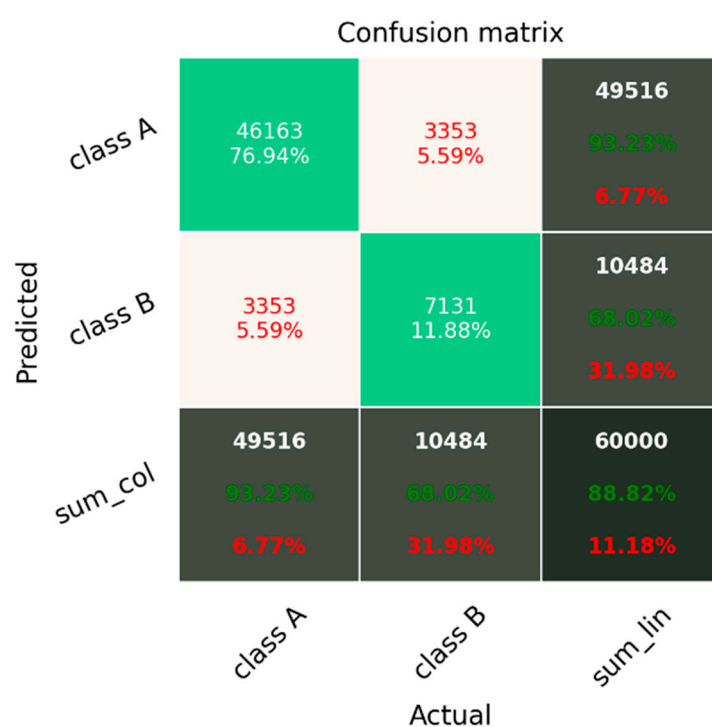
Figure 4. Confusion matrix of the classification results.

Table 2. Comparison of performance on the Avazu dataset. Best value is in bold.

Method	LogLoss	AUC
FM [50]	0.3856	0.7706
AFM [57]	0.3854	0.7718
NFM [58]	0.3864	0.7708
HOFM [59]	0.3854	0.7701
CrossNet [34]	0.3868	0.7667
DeepCrossing [16]	0.3889	0.7643
AutoInt [73]	0.3824	0.7752
Proposed (this paper)	**0.3562**	**0.8062**

5. Conclusions

In this study, we suggested applying artificial intelligence (AI) methodologies as an additional part of a cyber security-oriented user, use, and usability (3U) framework. The effective analysis of the 3U data using the AI approaches enables domain specialists to build viable tactics for mitigating cyberattacks and providing decision assistance. Many

internet applications, such as internet advertising and recommendation systems, rely on click-through rate (CTR) prediction to forecast the likelihood of a user clicking on an ad or product, which is critical for understanding human online behaviour. Web systems, on the other hand, are vulnerable to click fraud attacks. We demonstrate the approach by a case study in CTR prediction that intends to forecast online fraud using deep learning approaches such as factorization machines.

The experimental results on a real-world Avazu dataset show that our suggested approach not only outpaces known modern CTR forecasting approaches, but also demonstrates the viability of the proposed framework.

The main limitation of the proposed model is as follows: the model does not account for the concept drift when the underlying features change over time in unpredictable ways.

Future work will involve capturing sequential features of click-throughs during feature engineering and creating deeper neural network architectures to account for higher-order feature interactions. Moreover, we will consider the adoption of the model for CTR prediction using streamed data.

Author Contributions: Conceptualization, R.D.; Formal analysis, R.D. and L.Z.-J.; Funding acquisition, R.D.; Investigation, R.D. and L.Z.-J.; Methodology, R.D.; Software, R.D.; Supervision, R.D.; Validation, R.D.; Writing—original draft, R.D. and L.Z.-J.; Writing—review & editing, R.D. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Dataset used in this study is available from: https://www.kaggle.com/c/avazu-ctr-prediction (accessed on 5 January 2022).

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

References

1. Woźniak, M.; Połap, D.; Damaševičius, R.; Wei, W. Design of computational intelligence-based language interface for human-machine secure interaction. *J. Univ. Comput. Sci.* **2018**, *24*, 537–553.
2. Paskevicius, P.; Damasevicius, R. Design of usable interface for a mobile e-commerce system. *CEUR Workshop Proc.* **2016**, *1712*, 31–35.
3. Grundy, J.; Khalajzadeh, H.; McIntosh, J. Towards human-centric model-driven software engineering. In Proceedings of the 15th International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE 2020), Prague, Czech Republic, 5–6 May 2020; pp. 229–238.
4. Carmichael, D.; Cleave, D. How effective is social media advertising? A study of facebook social advertisements. In Proceedings of the 2012 International Conference for Internet Technology and Secured Transactions (ICITST 2012), London, UK, 10–12 December 2012; pp. 226–229.
5. Vila, N.; Kuster, I. Consumer feelings and behaviours towards well designed websites. *Inf. Manag.* **2011**, *48*, 166–177. [CrossRef]
6. Zailskaite-Jakste, L.; Ostreika, A.; Jakstas, A.; Staneviciene, E.; Damasevicius, R. Brand communication in social media: The use of image colours in popular posts. In Proceedings of the 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO 2017), Opatija, Croatia, 22–26 May 2017; pp. 1373–1378. [CrossRef]
7. Colbert, M.; Oliver, A.; Oikonomou, E. The effect of credibility of host site upon click rate through sponsored content. In Proceedings of the 9th International Conference on Persuasive Technology (PERSUASIVE 2014), Padua, Italy, 21–23 May 2014; Springer: Berlin/Heidelberg, Germany, 2014; pp. 56–67. [CrossRef]
8. Ullah, A.; Xiao, H.; Barker, T. A study into the usability and security implications of text and image based challenge questions in the context of online examination. *Educ. Inf. Technol.* **2019**, *24*, 13–39. [CrossRef]
9. Ali, S.; Anwar, S.; Solehria, S. User interaction based framework for protecting user privacy in online social networks. In Proceedings of the ICISO 2013—The 14th International Conference on Informatics and Semiotics in Organisations, IFIP WG8.1 Working Conference, Riga, Latvia, 6–7 November 2013; pp. 82–90.
10. Petelka, J.; Zou, Y.; Schaub, F. Put your warning where your link is: Improving and evaluating email phishing warnings. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, UK, 4–9 May 2019. [CrossRef]
11. Franz, A.; Zimmermann, V.; Albrecht, G.; Hartwig, K.; Reuter, C.; Benlian, A.; Vogt, J. SoK: Still plenty of phish in the sea—A taxonomy of user-oriented phishing interventions and avenues for future research. In Proceedings of the 17th Symposium on Usable Privacy and Security (SOUPS 2021), Vancouver, BC, Canada, 8–10 August 2021; pp. 339–357.
12. Petrykina, Y.; Schwartz-Chassidim, H.; Toch, E. Nudging users towards online safety using gamified environments. *Comput. Secur.* **2021**, *108*, 102270. [CrossRef]

13. Distler, V.; Lenzini, G.; Lallemand, C.; Koenig, V. The framework of security-enhancing friction: How UX can help users behave more securely. In Proceedings of the Pervasive Health: Pervasive Computing Technologies for Healthcare, Atlanta, GA, USA, 18–20 May 2020; pp. 45–58. [CrossRef]

14. Ji, Z.; Pi, H.; Wei, W.; Xiong, B.; Wozniak, M.; Damasevicius, R. Recommendation based on review texts and social communities: A hybrid model. *IEEE Access* **2019**, *7*, 40416–40427. [CrossRef]

15. Li, Z.; Cui, Z.; Wu, S.; Zhang, X.; Wang, L. Fi-gnn: Modeling feature interactions via graph neural networks for ctr prediction. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management, Beijing, China, 3–7 November 2019; pp. 539–548.

16. Shan, Y.; Hoens, T.R.; Jiao, J.; Wang, H.; Yu, D.; Mao, J. Deep crossing: Web-scale modeling without manually crafted combinatorial features. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 255–262.

17. Shubham, S.; Jain, N.; Gupta, V.; Mohan, S.; Ariffin, M.M.; Ahmadian, A. Identify glomeruli in human kidney tissue images using a deep learning approach. *Soft Comput.* **2021**, 1–12. [CrossRef]

18. Yazdinejad, A.; Parizi, R.; Dehghantanha, A.; Srivastava, G.; Mohan, S.; Rababah, A. Cost optimization of secure routing with untrusted devices in software defined networking. *J. Parallel Distrib. Comput.* **2020**, *143*, 36–46. [CrossRef]

19. Vallathan, G.; John, A.; Thirumalai, C.; Mohan, S.; Srivastava, G.; Lin, J.C.-W. Suspicious activity detection using deep learning in secure assisted living IoT environments. *J. Supercomput.* **2020**, *77*, 3242–3260. [CrossRef]

20. Grobler, M.; Gaire, R.; Nepal, S. User, Usage and Usability: Redefining Human Centric Cyber Security. *Front. Big Data* **2021**, *4*, 583723. [CrossRef]

21. Damaševičius, R.; Toldinas, J.; Venčkauskas, A.; Grigaliūnas, Š.; Morkevičius, N.; Jukavičius, V. Visual analytics for cyber security domain: State-of-the-art and challenges. In Proceedings of the International Conference on Information and Software Technologies (ICIST 2019), Jingde, China, 19–22 November 2019; pp. 256–270. [CrossRef]

22. Damasevicius, R.; Toldinas, J.; Venckauskas, A.; Grigaliunas, S.; Morkevicius, N. Technical threat intelligence analytics: What and how to visualize for analytic process. In Proceedings of the 24th International Conference Electronics (ELECTRONICS 2020), Palanga, Lithuania, 15–17 June 2020. [CrossRef]

23. McMahan, H.B.; Holt, G.; Sculley, D.; Young, M.; Ebner, D.; Grady, J.; Nie, L.; Phillips, T.; Davydov, E.; Golovin, D.; et al. Ad click prediction: A view from the trenches. In Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Chicago, IL, USA, 11–14 August 2013; pp. 1222–1230.

24. Chapelle, O. Modeling delayed feedback in display advertising. In Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 24–27 August 2014; pp. 1097–1105.

25. Graepel, T.; Borchert, T.; Candela, J.Q.; Herbrich, R. Web-scale Bayesian click-through rate prediction for sponsored search advertising in Microsoft's bing search engine. In Proceedings of the 27th International Conference on International Conference on Machine Learning, Madison, WI, USA, 21–24 June 2010; pp. 13–20. [CrossRef]

26. Guo, H.; Chen, B.; Tang, R.; Zhang, W.; Li, Z.; He, X. An Embedding Learning Framework for Numerical Features in CTR Prediction. *arXiv* **2020**, arXiv:2012.08986.

27. Zhang, Y.; Dai, H.; Xu, C.; Feng, J.; Wang, T.; Bian, J.; Wang, B.; Liu, T.Y. Sequential Click Prediction for Sponsored Search with Recurrent Neural networks. In Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, Québec City, QC, Canada, 27–31 July 2014.

28. An, H.; Ren, J.; Wu, S. XGBDeepFM for CTR predictions in mobile advertising benefits from ad context. *Math. Probl. Eng.* **2020**, *2020*, 1747315. [CrossRef]

29. Zhang, W.; Du, T.; Wang, J. Deep learning over multi-field categorical data. In *European Conference on Information Retrieval*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 45–57.

30. Qu, Y.; Fang, B.; Zhang, W.; Tang, R.; Niu, M.; Guo, H.; Yu, Y.; He, X. Product-based neural networks for user response prediction. In Proceedings of the 2016 IEEE 16th International Conference on Data Mining, Barcelona, Spain, 12–15 December 2016; pp. 1149–1154.

31. Guo, H.; Tang, R.; Ye, Y.; Li, Z.; He, X. DeepFM: A factorization-machine based neural network for CTR prediction. In Proceedings of the 26th International Joint Conference on Artificial Intelligence, Melbourne, Australia, 19–25 August 2017; pp. 1725–1731.

32. Cheng, H.-T.; Koc, L.; Harmsen, J.; Shaked, T.; Chandra, T.; Aradhye, H.; Anderson, G.; Corrado, G.; Chai, W.; Ispir, M.; et al. Wide & deep learning for recommender systems. In Proceedings of the 1st Workshop on Deep Learning for Recommender Systems, Boston, MA, USA, 15 September 2016; pp. 7–10.

33. Zhai, S.; Chang, K.; Zhang, R.; Zhang, Z.M. DeepIntent: Learning attentions for online advertising with recurrent neural networks. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 1295–1304.

34. Wang, R.; Fu, B.; Fu, G.; Wang, M. Deep Cross Network for Ad Click Predictions. In Proceedings of the ADKDD'17, New York, NY, USA, 14 August 2017.

35. Edizel, B.; Mantrach, A.; Bai, X. Deep character-level click-through rate prediction for sponsored search. In Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, Tokyo, Japan, 7–11 August 2017; pp. 305–314.

36. Jaech, A.; Kamisetty, H.; Ringger, E.; Clarke, C. Match-Tensor: A Deep Relevance Model for Search. *arXiv* **2017**, arXiv:1701.07795.

37. Zhou, G.; Mou, N.; Fan, Y.; Pi, Q.; Bian, W.; Zhou, C.; Zhu, X.; Gai, K. Deep Interest Evolution Network for Click-Through Rate Prediction. In Proceedings of the AAAI conference on artificial intelligence, Honolulu, HI, USA, 27 January–1 February 2019; Volume 33, pp. 5941–5948. [CrossRef]

38. Gligorijevic, J.; Gligorijevic, D.; Stojkovic, I.; Bai, X.; Goyal, A.; Obradovic, Z. Deeply supervised model for click-through rate prediction in sponsored search. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, New York, NY, USA, 4–8 August 2019; pp. 1–22.

39. Chen, J.; Zhao, Z.; Shi, J.; Zhao, C. A new approach for mobile advertising click-through rate estimation based on deep belief nets. *Comput. Intell. Neurosci.* **2017**, *2017*, 7259762. [CrossRef]

40. Gan, M.; Xiao, K. R-RNN: Extracting user recent behavior sequence for click-through rate prediction. *IEEE Access* **2019**, *7*, 111767–111777. [CrossRef]

41. Jiang, Z.; Gao, S.; Dai, W. A CTR prediction approach for text advertising based on the SAE-LR deep neural network. *J. Inf. Process. Syst.* **2017**, *13*, 1052–1070. [CrossRef]

42. Li, H.; Duan, H.; Zheng, Y.; Wang, Q.; Wang, Y. A CTR prediction model based on user interest via attention mechanism. *Appl. Intell.* **2020**, *50*, 1192–1203. [CrossRef]

43. Lu, Q.; Li, S.; Yang, T.; Xu, C. An adaptive hybrid XdeepFM based deep interest network model for click-through rate prediction system. *PeerJ Comput. Sci.* **2021**, *7*, e716. [CrossRef]

44. Qiang, B.; Lu, Y.; Yang, M.; Chen, X.; Chen, J.; Cao, Y. SDeepFM: Multi-scale stacking feature interactions for click-through rate prediction. *Electronics* **2020**, *9*, 350. [CrossRef]

45. Qin, J.; Zhang, W.; Wu, X.; Jin, J.; Fang, Y.; Yu, Y. User Behavior Retrieval for Click-Through Rate Prediction. In Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, New York, NY, USA, 25–30 July 2020. [CrossRef]

46. Song, K.; Huang, Q.; Zhang, F.; Lu, J. Coarse-to-fine: A dual-view attention network for click-through rate prediction. *Knowl.-Based Syst.* **2021**, *216*, 106767. [CrossRef]

47. Wang, Q.; Liu, F.; Xing, S.; Zhao, X. A new approach for advertising ctr prediction based on deep neural network via attention mechanism. *Comput. Math. Methods Med.* **2018**, *2018*, 8056541. [CrossRef] [PubMed]

48. Xu, R.; Wang, M.; Xie, Y. Optimally connected deep belief net for click through rate prediction in online advertising. *IEEE Access* **2018**, *6*, 43009–43020. [CrossRef]

49. Yang, Y.; Xu, B.; Shen, S.; Shen, F.; Zhao, J. Operation-aware neural networks for user response prediction. *Neural Netw.* **2020**, *121*, 161–168. [CrossRef] [PubMed]

50. Rendle, S. Factorization machines. In Proceedings of the 2010 IEEE 10th International Conference on Data Mining (ICDM), Sydney, Australia, 13–17 December 2010; pp. 995–1000.

51. Rendle, S. Factorization machines with LIBFM. *ACM Trans. Intell. Syst. Technol.* **2012**, *3*, 57. [CrossRef]

52. Lang, F.; Liang, L.; Huang, K.; Chen, T.; Zhu, S. Movie recommendation system for educational purposes based on field-aware factorization machine. *Mob. Netw. Appl.* **2021**, *26*, 2199–2205. [CrossRef]

53. Wen, P.; Yuan, W.; Qin, Q.; Sang, S.; Zhang, Z. Neural attention model for recommendation based on factorization machines. *Appl. Intell.* **2021**, *51*, 1829–1844. [CrossRef]

54. Juan, Y.; Zhuang, Y.; Chin, W.-S.; Lin, C.-J. Field-aware factorization machines for CTR prediction. In Proceedings of the 10th ACM Conference on Recommender Systems, Boston, MA, USA, 15–19 September 2016.

55. Zhang, L.; Shen, W.; Huang, J.; Li, S.; Pan, G. Field-aware neural factorization machine for click-through rate prediction. *IEEE Access* **2019**, *7*, 75032–75040. [CrossRef]

56. Cheng, C.; Xia, F.; Zhang, T.; King, I.; Lyu, M.R. Gradient boosting factorization machines. In Proceedings of the 8th ACM Conference on Recommender systems, New York, NY, USA, 6–10 October 2014; pp. 265–272.

57. Xiao, J.; Ye, H.; He, X.; Zhang, H.; Wu, F.; Chua, T.S. Attentional factorization machines: Learning the weight of feature interactions via attention networks. In Proceedings of the 26th International Joint Conference on Artificial Intelligence, Melbourne, Australia, 19–25 August 2017; pp. 3119–3125.

58. He, X.; Chua, T.S. Neural factorization machines for sparse predictive analytics. In Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, Tokyo, Japan, 7–11 August 2017; pp. 355–364.

59. Blondel, M.; Fujino, A.; Ueda, N.; Ishihata, M. Higher-order factorization machines. In Proceedings of the Advances in Neural Information Processing Systems 29 (NIPS 2016), Barcelona, Spain, 5–10 December 2016; pp. 3351–3359.

60. Zhou, R.; Liu, C.; Wan, J.; Fan, Q.; Ren, Y.; Zhang, J.; Xiong, N. A Hybrid Neural Network Architecture to Predict Online Advertising Click-Through Rate Behaviors in Social Networks. *IEEE Trans. Netw. Sci. Eng.* **2021**, *8*, 3061–3072. [CrossRef]

61. Zurko, M.E.; Simon, R.T. User-centered security. In Proceedings of the 1996 workshop on New security paradigms, Lake Arrowhead, CA, USA, 17–20 September 1996; pp. 27–33.

62. Hassenzahl, M.; Tractinsky, N. User experience—A research agenda. In *Behaviour & Information Technology*; Informa UK Limited: London, UK, 2006; Volume 25, pp. 91–97. [CrossRef]

63. Muntinga, D.G.; Moorman, M.; Smit, E.G. Introducing COBRAs: Exploring motivations for brand-related social media use. *Int. J. Advert.* **2011**, *30*, 13–46. [CrossRef]

64. Kraus, L.; Wechsung, I.; Möller, S. Psychological needs as motivators for security and privacy actions on smartphones. *J. Inf. Secur. Appl.* **2017**, *34*, 34–45. [CrossRef]

65. Voorveld, H.A.M.; van Noort, G.; Muntinga, D.G.; Bronner, F. Engagement with Social Media and Social Media Advertising: The Differentiating Role of Platform Type. *J. Advert.* **2018**, *47*, 38–54. [CrossRef]

66. Gaver, B.; Martin, H. Alternatives. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*; ACM Press: New York, NY, USA, 2000. [CrossRef]

67. Stolfo, S.J.; Salem, M.B.; Keromytis, A.D. Fog computing: Mitigating insider data theft attacks in the cloud. In Proceedings of the 2012 IEEE symposium on security and privacy workshops, San Francisco, CA, USA, 20–23 May 2012; pp. 125–128.

68. Namasudra, S.; Devi, D.; Kadry, S.; Sundarasekar, R.; Shanthini, A. Towards DNA based data security in the cloud computing environment. *Comput. Commun.* **2020**, *151*, 539–547. [CrossRef]

69. Ande, R.; Adebisi, B.; Hammoudeh, M.; Saleem, J. Internet of Things: Evolution and technologies from a security perspective. *Sustain. Cities Soc.* **2020**, *54*, 101728. [CrossRef]

70. Chen, Q.; Bakare, S.; Verma, A.; Casasanta, C.; White, C.; Cotoranu, A.; Leider, A. Analyzing Expert Cybersecurity Twitter Accounts by Using Thesaurus Methods for Text Analytics. In Proceedings of the Student-Faculty Research Day, CSIS, Pace University, New York, NY, USA, 4 May 2018.

71. Kabil, A.; Duval, T.; Cuppens, N. Alert Characterization by Non-Expert Users in a Cybersecurity Virtual Environment: A Usability Study. In Proceedings of the 7th International Conference on Augmented Reality, Virtual Reality and Computer Graphics, Lecce, Italy, 8–11 June 2020. [CrossRef]

72. Qi, G.; Li, P.; Carta, S. Deep Field-Aware Interaction Machine for Click-Through Rate Prediction. *Mob. Inf. Syst.* **2021**, *2021*, 5575249. [CrossRef]

73. Song, W.; Shi, C.; Xiao, Z.; Duan, Z.; Xu, Y.; Zhang, M.; Tang, J. AutoInt: Automatic Feature Interaction Learning via Self-Attentive Neural Networks. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM 2019), Beijing, China, 3–7 November 2019; pp. 1161–1170.

74. Lian, J.; Zhou, X.; Zhang, F.; Chen, Z.; Xie, X.; Sun, G. xdeepfm: Combining explicit and implicit feature interactions for recommender systems. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, New York, NY, USA, 19–23 August 2018; pp. 1754–1763.