INTERNATIONAL JOURNAL OF SOFTWARE ENGINEERING & COMPUTER SYSTEMS (IJSECS)
ISSN: 2289-8522 e-ISSN: 2180-0650
VOL. 8, ISSUE 2, 33 – 42
DOI: https://doi.org/10.15282/ijsecs.8.2.2022.4.0101

ORIGINAL ARTICLE

Deep Neural Network for Click-Through Rate Prediction

Erni[1], D. Riana[1]

[1]Faculty of Information Technology, University of Nusa Mandiri, Jakarta, Indonesia.

ABSTRACT – Predicted click-through rate is one of the most frequently used criteria to determine the effectiveness of an ads. In advertising production, click-through predictions are very influential for the company that places the ads. Click-through rates need to be predicted accurately because accurate prediction results determine whether the click-through rate is exactly clicked or not by the viewing consumer. Predicted click-through can be done on advertising and social network datasets. The use of these two datasets is intended to make the comparison results more convincing from the proposed method. The purpose of this study is to compare two advertising and social network datasets, by proposing the application of the Deep Neural Network (DNN) model by testing hyperparameter variations to find a better architecture in predicting whether or not users click on an advertisement. The hyperparameter variations include 3 variations of the hidden layer, 2 variations of the activation function, namely ReLu and Sigmoid, 3 variations of the optimization (RMSprop, Adam, and Adagrad), and 3 variations of the learning rate (0.1, 0.01, and 0.001). The results of experiments conducted with the advertising parameter dataset with hidden layer of 3, learning rate of 0.01, and Adam optimization with an accuracy value of 99.90%, AUC of 99.90% and Precision-Recall of 99.89% while the data for social network ads parameters with hidden layer of 5, learning rate of 0.1 and Adam optimization with accuracy of 92.25%, AUC of 92.72%, and Precision-Recall of 89.70%.

ARTICLE HISTORY
Received: 17 Feb 2022
Revised: 22 June 2022
Accepted: 13 July 2022

KEYWORDS
Click-Through Rate
Deep Neural Network
Preprocessing
Advertising
Social Network

INTRODUCTION

Click-through rate prediction is very important and is considered one of the most profitable stories in the machine learning domain. Click-through rates need to be predicted accurately because accurate prediction results determine whether the click-through rate is exactly clicked or not by the viewing consumer. In previous research, there were no more optimal predictions in the Deep Neural Networks method for the Advertising dataset [1] and Social Network dataset [2] (click-through rate), especially for consumers who only click on ads or don't click on ads [3][4][5]. Click-through rate prediction has been used historically in every type of ads format like search engine, textual and contextual ads, video ads, and others. Rapid development of Ads network and click-through rate prediction require huge data analysis [6].

Dataset Advertising [1], known as internet advertising or online marketing is one of the fastest growing and most profitable businesses. One of the datasets containing an advertising dataset can be obtained from Kaggle which contains 1,000 samples and ten features consisting of a "timestamp" column which is expressed as a dataset which consists of month, day, and time [3]. In the dataset [1], the researcher carried out the pre-processing stage, namely by converting the "country" and "city" columns into unique integer values, performing feature selection by deleting the Ad Topic Line column based on previous research references [7], and performing Extracting date by separating month, day, and time in the "timestamp" column. Therefore, there are differences in the features of the dataset that we used in this study when compared to research [3]. This allows different results from previous studies. The dataset of Social Network ads [2] is an advertising dataset retrieved from the Kaggle dataset repository and is a standard dataset [5]. In dataset [2], we also carried out the pre-processing stage by converting the "gender" column into a unique integer value. Therefore, there are differences in the features of the dataset that we used in this study when compared to research [5].

In advertising production, click-through predictions are very influential for the company that places the ads. Companies must also know whether the ads are actually seen by the user or just clicked on it. In addition to predicting the click-through rate of an ads, the use of the model or algorithm used is also very important in analyzing the click-through rate that occurs.

Machine learning approaches have the ability to overcome these limitations. This approach has been carried out in several previous studies by comparing several methods such as support vector machine, logistic regression, random forest, and decision tree [7] [6]. Previous studies have also compared several datasets and methods [4][3].

The Deep Neural Network approach has the ability to solve this problem by using a variety of hyperparameters used including 2 forms of layer decoder and encoder, 3 variations of hidden layers (3 hidden layers, 4 hidden layers, and 5 hidden layers), 2 variations of the activation function namely ReLu and Sigmoid, 3 variants of the optimization

*CORRESPONDING AUTHOR | D. Riana | ✉ dwiza@nusamandiri.ac.id

(RMSprop, Adam, and AdaGrad), and 3 variations of learning rate (0.1, 0.01, 0.001) as carried out in previous studies [8].

Deep Neural Networks are part of the broad field of AI, namely the science and engineering of creating intelligent machines that have the ability to achieve goals like humans do [9]. DNN has more than 3 layers (input layer, hidden layer, and output layer), in other words Multilayer perceptron with more layers. Because the layers are relatively many, it is called Deep. The learning process at DNN is known as Deep Learning [10]. Each layer has multiple nodes and nodes from adjacent layers are fully connected and have a bigger data processing capacity that is stronger than ANN [11][12][13].

AdaGrad optimization is more sophisticated and sets the inversely proportional scaling of the learning rate to the square root of the cumulative squared gradient. AdaGrad is not effective for all DNN training. Since the change in learning rate is a function of the historical gradient, AdaGrad becomes prone to convergence. The RMSProp algorithm is a modification of the AdaGrad algorithm to make it effective in nonconvex problem spaces. RMSPod replaces the summation of the squared gradients in AdaGrad with an exponentially decreasing average of the gradient, effectively dropping the impact of the historical gradient. Adam demonstrated that adaptive moment estimation is the latest evolution of an adaptive learning algorithm that integrates ideas from AdaGrad, RMSProp, and momentum. Just like AdaGrad and RMSPod, Adam provides individual learning levels for each parameter. Adam benefits from both the previous methods doing a better job of handling non-stationary goals and noisy and infrequent gradient problems. Adam uses the first moment (i.e., the mean as used in RMSProp) as well as the second moment of the gradient (uncentralized variance). The exponential moving average of the squared gradient [8].

This research is very important to determine the effectiveness of an advertisement on a mobile application from the side of the application user and the party doing the advertising. The specifics of this research are important because the prediction of click-through rate is one of the most frequently used criteria to determine advertising effectiveness. In ad production, click-through predictions are very influential for the company that places the ad. This relates to the profits earned by advertisers. Therefore, it is very important to predict the click-through rate. Click-through predictions can be used by models or algorithms. It is necessary to determine the best model or algorithm. Therefore, this study aims to get a deep learning model that is able to predict click-through rates with good accuracy. For this purpose, this study uses two data sets, namely advertising and social network dataset [2] and proposes the application of a Deep Neural Network (DNN) model by testing hyperparameter variations to find an architecture that is better at predicting click-through rates. The hyperparameter variations include 3 variations of the optimizer (RMSprop, Adam, and Adagrad), 2 variations of the activation function, namely ReLu and Sigmoid, 3 variations of the learning rate (0.1, 0.01, and 0.001), and 3 variations of the hidden layer.

In this study, the authors continue the previous research [14] by proposing a prediction of click-through rate using the Deep Learning algorithm for comparison of two datasets in predicting click-through rate by doing different pre-processing from the previous studies in order to get more optimal results

RELATED WORK

Research related to the prediction of click-through rate has been done by many previous researchers. Dan [3] conducted research on the prediction of click-through rate with the proposed algorithms including SVM, DNN, LR, FM, DSL, AND, GAN, and MTF. The results showed that the MTF method obtained the highest score of 0.69 with Avazu data, 0.68 with Avito data, 0.67 with Talking data, 0.91 with Kad data, and 0.65 with Corpon data, but the accuracy results produced can still be improved. Research conducted by Ying [3] shows the prediction of click-through rate by testing several data and methods. The method used from the research results shows that the RTILKE method produces significant predictions with the acquisition of AUC 0.86 with Display data of 0.87 with Avazu data, 0.89 with Avito data, 0.83 with Talking data, and 0.89 with Kad data while the results with the DNN method with AUC 0.69 with data Display of 0.67 with Avazu data, 0.57 with Avito data, 0.65 with Talking data and 0.66 with Card data, but the accuracy results produced can still be improved.

Similar research was also conducted by Thejas [15] with the aim of increasing site revenue or managing revenue from advertising by combining several datasets and two learning models for the transformation and classification of features with the results of 90.4% RF accuracy, 93.2% SVM, 94.2% GBDT, 94.6% NB, 96% ETCF and 96.4% CFXGB.

Researcher by author's name Saraswathi [6] conducted a study on the precision level of the custom model being observed to be 96% which is much higher than the standard model. This file is used in an iOS application, to read input component and give results 0 or 1. Along with the information of this framework built, accuracy of click prediction framework click-through rate which can anticipate if user click promotion increases up to 96% where previously this model had speed precision of only 92%.

Research on click-through prediction [7] conducted an empirical study of the association of various web exploitation techniques to predict whether a click-through rate will be clicked or not. The results obtained show that the SVM algorithm is better than other algorithms with an accuracy of 97.5%.

Furthermore, Poulomi [5] conducted a study to investigate six types of Machine Learning, Classifier algorithms (i.e. Logistics Regression, SVM, Naive Bayes, KNN, Decision Tree, and Random Forest). In this study, they demonstrated their comparative analysis and predicted whether someone will buy a certain product immediately after being launched on the market. The results of the research show the accuracy level of Support Vector Machine of 93%, Naïve Bayes of 90%, Logistic Regression of 89%, K-Nearest Neighbour of 93%, Decision Tree of 91%, and Random Forest of 92%.

Similar research was also conducted by Tulin [16] which conducted a Prediction of Advertisement Clicks-based research with the aim of increasing the click-through rate of digital media in the most efficient way to increase the market share. The results show that XGBoost gives the highest R-Squared value of 81% in the prediction of all metrics used in this study. Research conducted by Hary [17] predicted a false job vacancy click-through rate. The evaluation and validation model used Naïve Bayes as the baseline model and Stochastic Gradient Descent as the final model. For the Naïve Bayes model, the accuracy value is 0.971 and the F1 score is 0.743 while the Stochastic Gradient Descent obtained an accuracy value of 0.977 and an F1 score of 0.81. This final result shows that Stochastic Gradient Descent has slightly better performance than Naïve Bayes.

The latest research related to click-through rate prediction was carried out by Oscar [18] regarding bank promotion with the DNN algorithm using multi-layer interconnected input and output layers where each hidden layer has 3 variations including 2 architectures, 3 learning variations, and 3 optimization variations with an accuracy of 90% using the DNN method. Related work can be seen in Table 1.

Table 1. Related Research

Authors	Dataset	Accuracy
Dan Jiang, Rongbin Xu, Xin Xu, Ying Xie, (2021) [3]	Avazu (40 million), Avito (20 million), Gifts (1,000), Talking (18.7 million), and Coupon-purcase (2.8 million)	The DNN method produces AUC 0.69% for Avazu data, 0.66% for Avito data, 0.57% for Talking data, 0.65% for gift data and 0.64% for Corpon data. The highest MTF 0.91% Card data.
Xie, Y., Jiang, D., Wang, X., & Xu, R. (2019). [4]	Avazu (40 million), Avito (20 million), Gifts (1,000), Talking (18.7 million), and Display (46 million).	DNN method AUC 0.69% Display data, 0.67% Avazu data, 0.57% Avito data, 0.65% Talking data and the highest is RTILKE 0.89% Kad data.
Thejas (2020) [15]	Double click scam	Method RF 90.4%, SVM 93.2%, GBDT 94.2%, NB 94.6%, ETCF 96% and CFXGB 96.4%.
Saraswati (2019) [6]	Advertising (1000)	RL accuracy is 92%, SVM accuracy is 92% and DT accuracy is 92% and RF accuracy is 96%.
Mahindra, Jitendra (2020) [7]	Advertising (1000)	SVM method with 97.5% accuracy results.
Poulomi (2019) [5]	Social_Network_Ads (400)	SVM 93%, NB 90%, LR 89%, KNN 93%, DT 91%, and RF 92%.
Tulin (2019) [19]	OTA Dashboard (800,237K)	The XGBoost method has the highest R-Squared value of 81%
Hary, sabita (2021) [17]	Job vacancies (17.880K)	Naïve Bayes accuracy 0.971% and F1 score 0.743%. Stochastic Gradient Descent accuracy 0.977% and F1 score 0.81%
Orcar, Maulidiah (2021) [18]	Telemarketing	3 hidden layers, learning rate and 10 epochs produce the highest accuracy of 90% with the DNN method

PROPOSED METHOD

An overview of the method is the flow of the experiments carried out, among others, starting from the preparation stage of the advertising and social network dataset [2] taken from the official Kaggle website. The next stage is conducting pre-processing data which aims to obtain data that is ready to be used for testing the proposed model. The next stage evaluates the performance of the model or algorithm by doing cross validation and the last step is testing the model based on the pre-processing data stage model to get optimal accuracy results in predicting click-through rate. The following is a click-through rate research method that can be seen in Figure 1.

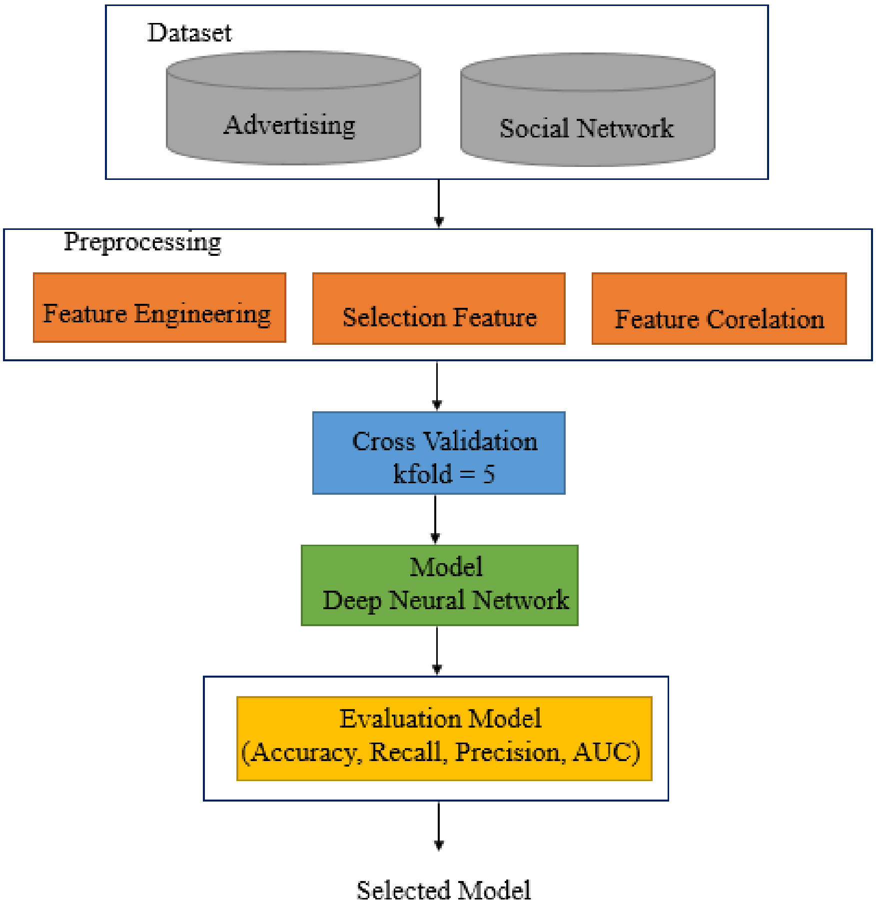

Figure 1. Proposed Method

Data Pre-processing

Data Pre-processing is a technique for preparing data in such cases as cleaning data from noise or changing data formats [20].

a. Engineering Features
Extracting date is one of the Feature Engineering used in this research phase. Extracting date is done by separating the month, day, and time contained in the timestamp feature. Therefore, 3 new attributes appear, namely month, day, and time. This attribute will explain when the user clicks on the click-through rate on advertising data.

b. Selection Features
Deletion of several features in the dataset based on previous research references. Removed features are Ad Topic Line and Timestamp.

c. Featured Correlation
Feature Correlation is used to see the correlation between attributes [21] [22]. The more positive the correlation value, the better the classification results will be obtained. If the resulting correlation is negative, then it can be considered for deleting attributes because it will affect the model results to be obtained

Cross Validation

After preparing the pre-processing data, the next step is Cross-validation which is used as a test data for model validation or evaluation. In this study, cross-validation was used 5 times.

Dataset Test

The proposed DL algorithm model is the DNN method. The model was then experimented by comparing two datasets that predict click-through rates. Furthermore, optimization was carried out by setting the model parameters. Optimal parameters can produce a good model so that the author conducts several experiments on the model as follows:
a. Comparing the number of hidden layers.
b. Comparing the number of epochs on a pre-determined model.
c. Comparing the use of optimization.
The parameters used can be seen in Table 2.

Table 2. Parameters of Deep Neural Network

Parameter	Parameter Value
Activation Input Layer	ReLU (Rectified Linear Unit)
Activation Hidden Layer	ReLU (Rectified Linear Unit)
Activation Output	Sigmoid
Optimization	Adam , RMSprop, Adagrad
Learning Rate	0.1, 0.01, 0.001
Epoch	50
Batch Size	50

Table 2 describes the parameters used in the DNN architecture. The tests carried out vary such as the value of learning rate, epoch, and batch size using the activation function on the type of layer and the type of optimization that will be used in this study.

Description of Dataset

The data used was obtained from public data on Kaggle under the name of the advertising dataset [1] in Table 3 and the social network dataset [2] in Table 4.

Table 3 shows a collection of Advertising dataset [1] which has ten variables in the form of Daily time spent on site, age, income area, daily internet usage, ads topic line, city, male, count, timestamp, and clicked-on-ads with a total of 1000 datasets. The use determination of a sample of 1000 datasets refers to research references [3] and [2] [5].

Table 3. Advertising Dataset

Daily Time Spent on Site	Age	Area Income	Daily Internet Usage	Ad Topic Line	City	Male	Country	Timestamp	Clicked on Ad
68.95	35	61833.9	256.09	Cloned 5thgeneration orchestration	Wrightburgh	0	Tunisia	27/03/2016 00:53	0
80.23	31	68441.85	193.77	Monitored national standardization	West Jodi	1	Nauru	04/04/2016 01:39	0
69.47	26	59785.94	236.5	Organic bottom-line service-desk	Davidton	0	San Marino	13/03/2016 20:35	0
74.15	29	54806.18	245.89	Triple-buffered reciprocal time-frame	West Terrifurt	1	Italy	10/01/2016 02:31	0
68.37	35	73889.99	225.58	Robust logistical utilization	South Manuel	0	Iceland	03/06/2016 03:36	0
59.99	23	59761.56	226.74	Sharable client-driven software	Jamieberg	1	Norway	19/05/2016 14:30	0
88.91	33	53852.85	208.36	Enhanced dedicated support	Brandonstad	0	Myanmar	28/01/2016 20:59	0
66.0	48	24593.33	131.76	Reactive local challenge	Port Jefferybury	1	Australia	07/03/2016 01:40	1
74.53	30	68862.0	221.51	Configurable coherent function	West Colin	1	Grenada	18/04/2016 09:33	0

Table 4 shows a collection of social network dataset [2] that has five variables in the form of user id, gender, age, estimated salary, and purchased with a total of 400 data.

Table 4. Social Network Dataset

User ID	Gender	Age	EstimatedSalary	Purchased
15624510	Male	19	19000	0
15810944	Male	35	20000	0
15668575	Female	26	43000	0
15603246	Female	27	57000	0
15804002	Male	19	76000	0
15728773	Male	27	58000	0
15598044	Female	27	84000	0
15694829	Female	32	150000	1
15600575	Male	25	33000	0

Advertising dataset [1] consisting of 1000 datasets has two class labels, namely not clicking click-through rate (0) and clicking click-through rate (1) of which 500 datasets are not clicking click-through rate and 500 datasets are clicking click-through rate. The class distribution statistics can be seen in Figure 2.

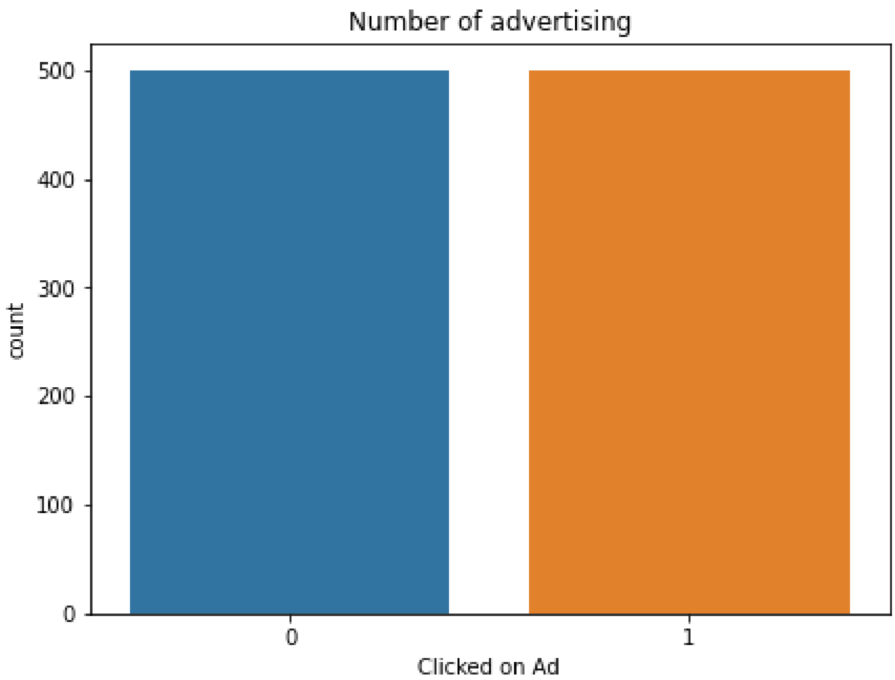

Figure 2. Research Methodology Dataset Advertising [1]

Social network dataset [2] consisting of 400 datasets has two class labels, namely not clicking click-through rate (0) and clicking click-through rate (1) of which 257 datasets are not clicking click-through rate and 143 datasets are clicking click-through rate. The class distribution statistics can be seen in Figure 3.

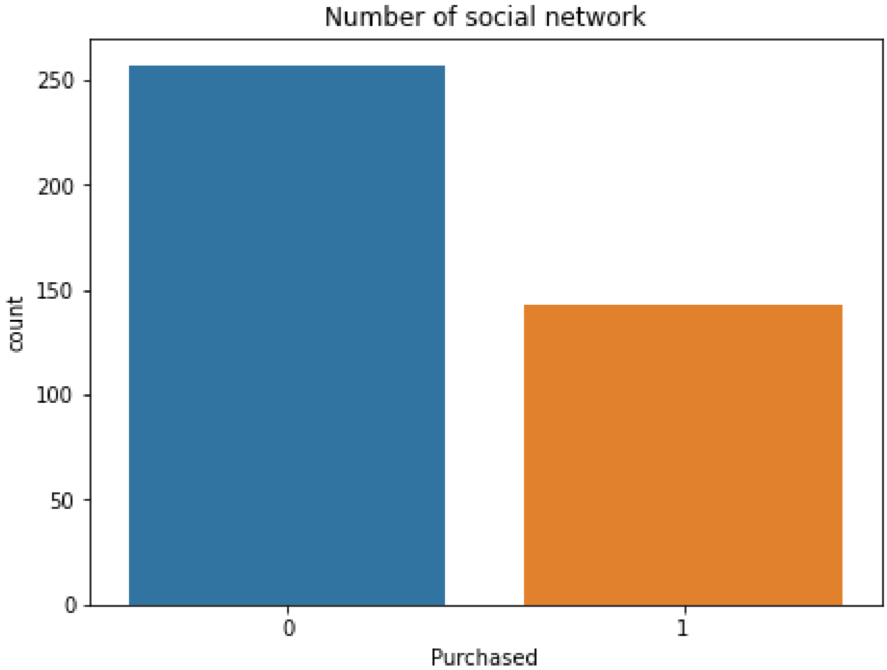

Figure 3. Research Methodology Dataset Social Network

RESULTS AND DISCUSSION

Pre-processing

The pre-processing carried out in the study has several stages including removing unnecessary attributes, adding new attributes, feature engineering, feature correlation, and checking for missing values with the aim of producing clean data so that it can be used in model testing.

a. Engineering Features

Feature Engineering used in this research phase, namely extracting date was done by separating the month, day, and time contained in the timestamp feature. Therefore, 3 new attributes appear, namely month, day, and time. This attribute will describe the time where the user clicked the click-through rate. The timestamp features before separation can be seen in Table 5.

Table 5. Features of Extracting Timestamp

Before	After		
Timestamp	Month	Day	Time
3/27/2016 0:53	3	27	0
04/04/2016 01:39	4	4	1
3/13/2016 20:35	3	13	20
01/10/2016 02:31	1	10	2
06/03/2016 03:36	6	3	3

The advertising dataset [1] transformation was carried out on the city and country features while the social network dataset [2] collection was also performed on gender features. An example of data before and after data transformation can be seen in Table 6.

Table 6. Transformation Features of City, Country, and Gender

Dataset Advertising				Dataset Social Network	
Before		After		Before	After
City	Country	City Codes	Country Codes	Gender	Gender
Wrightburgh	Tunisia	961	215	male	1
West Jodi	Nauru	903	147	male	1
Davidton	San Marino	111	184	female	0
West Terrifurt	Italy	939	103	female	0
South Manuel	Iceland	805	96	male	1

b. Selection Features

At this stage, feature selection was carried out by deleting several features in the dataset based on previous research references [7]. Removed features are Ad Topic Line and Timestamp. Therefore, the dataset that is ready to be processed is based on Table 7 as follows:

Table 7. Dataset after Pre-processing

Daily Time	Age	Area Income	Daily Internet Usage	City Codes	Country Codes	Male	Month	Day	Time	Clicked on Ad
68.95	35	61833.9	256.09	961	215	0	3	27	0	0
80.23	31	68441.85	193.77	903	147	1	4	4	1	0
69.47	26	59785.94	236.5	111	184	0	3	13	20	0
74.15	29	54806.18	245.89	939	103	1	1	10	2	0
68.37	35	73889.99	225.58	805	96	0	6	3	3	0
59.99	23	59761.56	226.74	282	158	1	5	19	14	0

Based on Table 7 above, the features in the advertising dataset [1] that will be used in model testing are Daily Time, Age, Area Income, Daily Internet Usage, City Codes, Country Codes, Male, Month, Day, Time, and Clicked-on-Add as data classes.

Table 8. Dataset after Pre-processing

User ID	Gender	Age	Estimated Salary	Purchased
15624510	1	19	19000	0
15810944	1	35	20000	0
15668575	0	26	43000	0
15603246	0	27	57000	0
15804002	1	19	76000	0
15728773	1	27	58000	0

Based on Table 8 above, the features in the social network dataset [2] that will be used in model testing are User id, gender, age, estimated salary, and purchased as data classes.

Model Experiment Results

Experiments were carried out on one of the deep learning methods, namely the Deep Neural Network. There are three stages carried out on the DNN model, namely by optimizing RMSprop and Adam, and by continuing with AdaGrad optimization by activating based on data that has been pre-processed previously, with 5 times cross-validation on both predictions of Click-through rate data.

The following are the most optimal results from the performance of the DNN model with RMSprop optimization, Adam, and AdaGrad optimization can be seen in Table 9.

Table 9. DNN Model Performance Results

Dataset	Optimization	Model Performance Results					
		Epoch/ Batch Size	Learning Rate	Hidden Layer	Accuracy	AUC	Precision-Recall
Advertising	RMSprop	50	0.001	3	99.30%		99.29%
	Adam		0.01	3	99.90%		99.89%
	Adagrad		0.1	3	99.50%		99.49%
Social Network	RMSprop		0.01	5	91.25%	90.24%	87.63%
	Adam		0.1	5	92.25%	92.72%	89.70%
	Adagrad		0.1	5	90.25%	89.31%	86.31%

The results of the Advertising dataset [1] research from the results of the optimized model performance of all the models that have been tried show that the DNN model with Adam optimization gets the best model performance results with accuracy of 99.90%, AUC of 99.90%, and Precision-Recall of 99.89%. Meanwhile, the results of the research on social network datasets [2] show that the DNN model with Adam optimization gets the best model performance with an accuracy of 92.25%, AUC of 92.72%, and Precision-Recall of 89.70%. The results of the best model performance will then be used as a proposed predictive model.

Selected Parameters

After experimenting with the DNN model comparing the three optimizations of RMSprop, Adam, and AdaGrad and then comparing the use of the three selected optimizations, the DNN parameters selected as the proposed predictive model were obtained and can be seen in the Table 10.

Table 10. DNN Model Performance Results

Dataset Advertising		Dataset Social Network	
Parameter	Parameter Value	Parameter	Parameter Value
Input Layer	10	Input Layer	4
Hidden Layer	3	Hidden Layer	5
Activation Input	ReLU	Activation Input	ReLU
Activation Output	Sigmoid	Activation Output	Sigmoid
Optimization	Adam (Lr=0.01)	Optimization	Adam (Lr=0.1)
Epoch	50	Epoch	50
Batch Size	50	Batch Size	50

Selected parameters from the advertising dataset [1] are input layer of 10, hidden layer of 3, ReLu input activation, sigmoid output activation, Adam optimization with Learning Rate of 0.01, epoch of 50, and batch size of 50 while the selected parameters of social network dataset are input layer of 4, hidden layer of 5, ReLu input activation, sigmoid output activation, optimization, Adam with Learning Rate of 0.1, epoch of 50, and batch size 50.

Comparison of the results of the proposed research with previous research

At the comparison stage, the accuracy displays the results of the DNN model with a comparison of the model with previous studies. This stage is carried out to see the novelty or novelty of the experimental results from previous studies using the same data set. The following is a comparison of results can be seen in Table 11 as follows.

Table 11. Comparison of Research Results

Dataset	Algorithms	AUC/ROC	Accuracy	Precision-recall
Advertising	DNN Xie [3]	66.00%	-	-
	Model DNN	99.90%	99.90%	99.89%
Social Network	SVM Poulomi [5]	-	93.00%	89.00%
	Model DNN	92.72%	92.25%	89.70%

Comparison of the results of research that has been done with previous research with advertising datasets [1] using the DNN method of previous studies [3] with AUC result of 66.00%. The findings of this study indicate that the AUC value in the advertising dataset [1] increased by 33.90% with an accuracy value of 99.90% and a precision-recall of 99.89%. The social network dataset of the previous research [2] reached 92.72% where the previous research [5] was not included. Another finding of this study resulted in a higher precision-recall value compared to [5] which was 89.70% even though the accuracy is 0.75% lower.

CONCLUSION

The DNN model can be used as a prediction of the best click impression ratio to estimate the possibility of users clicking on ads or products on the advertising dataset [1] that has been tried. After pre-processing, the results were more optimal than those that were not pre-processed in previous studies. From the accuracy comparison of the two datasets, the most optimal advertising parameter data is with hidden layer 3, 0.01 learning rate and with Adam optimization with an accuracy value of 99.90%, AUC 99.90% and Precision-Recall 99.89%, while social network ads parameter data with a hidden layer of 5, learning rate of 0.1 with Adam optimization with the accuracy of 92.25%, AUC of 92.72%, and Precision-Recall of 89.70%.

REFERENCES

[1] T. BYRNES, “Advertising,” [Online]. Available: https://www.kaggle.com/datasets/tbyrnes/advertising.

[2] “Social Network Ads,” [Online]. Available: https://www.kaggle.com/datasets/sash1563/social-network-ads.

[3] Y. Xie, D. Jiang, X. Wang, and R. Xu, “Robust transfer integrated locally kernel embedding for click-through rate prediction,” *Inf. Sci. (Ny).*, vol. 491, pp. 190–203, 2019, doi: 10.1016/j.ins.2019.04.006.

[4] D. Jiang, R. Xu, X. Xu, and Y. Xie, “Multi-view feature transfer for click-through rate prediction,” *Inf. Sci. (Ny).*, vol. 546, pp. 961–976, 2021, doi: 10.1016/j.ins.2020.09.005.

[5] P. Saha, “Performance analysis of the Machine Learning Classifiers to predict the behaviour of the customers,

when a new product is launched in the market," vol. 5, no. 3, pp. 1907–1911, 2019.

[6] S. Saraswathi, V. Krishnamurthy, D. Venkata Vara Prasad, R. K. Tarun, S. Abhinav, and D. Rushitaa, "Machine learning based ad-click prediction system," *Int. J. Eng. Adv. Technol.*, vol. 8, no. 6, pp. 3646–3648, 2019, doi: 10.35940/ijeat.F9366.088619.

[7] J. Mahindra Bagul and T. B. Kute, "Ad-Click Prediction using Prediction Algorithm: Machine Learning Approach," *Int. Res. J. Eng. Technol.*, 2020, [Online]. Available: www.irjet.net.

[8] A. Shrestha and A. Mahmood, "Review of deep learning algorithms and architectures," *IEEE Access*, vol. 7, pp. 53040–53065, 2019, doi: 10.1109/ACCESS.2019.2912200.

[9] S. M. Ieee, F. Ieee, V. Sze, Y.-H. Chen, T.-J. Yang, and J. S. Emer, "Efficient Processing of Deep Neural Networks: A Tutorial and," *Proc. IEEE*, vol. 105, no. 12, pp. 2295–2329, 2017, [Online]. Available: http://ieeexplore.ieee.org/document/8114708/.

[10] A. Noviar and S. Mukti, "Klasifikasi Arritmia pada Sinyal EKG menggunakan Deep Neural Network," *JUPITER*, vol. 13, pp. 29–38, 2021.

[11] C. Y. Low, J. Park, and A. B. J. Teoh, "Stacking-Based Deep Neural Network: Deep Analytic Network for Pattern Classification," *IEEE Trans. Cybern.*, vol. 50, no. 12, pp. 5021–5034, 2020, doi: 10.1109/TCYB.2019.2908387.

[12] G. Li, F. Hu, Y. Zhao, and N. Chi, "Enhanced performance of a phosphorescent white LED CAP 64QAM VLC system utilizing deep neural network (DNN) post equalization," *2019 IEEE/CIC Int. Conf. Commun. China, ICCC 2019*, no. Iccc, pp. 173–176, 2019, doi: 10.1109/ICCChina.2019.8855926.

[13] D. Xianzhi, "Research on Camera Calibration Technology Based on Deep Neural Network in Mine Environment," *Proc. - 2020 Int. Conf. Comput. Vision, Image Deep Learn. CVIDL 2020*, no. Cvidl, pp. 375–379, 2020, doi: 10.1109/CVIDL51233.2020.00-68.

[14] B. LAILIAH, "AD CLICK PREDICTION DENGAN MACHINE LEARNING," *Tesis*, 2021.

[15] T. G.S., S. Dheeshjith, S. S. Iyengar, N. R. Sunitha, and P. Badrinath, "A hybrid and effective learning approach for Click Fraud detection," *Mach. Learn. with Appl.*, vol. 3, p. 100016, 2021, doi: 10.1016/j.mlwa.2020.100016.

[16] T. Çakmak, A. T. Tekin, Ç. Şenel, T. Çoban, Z. E. Uran, and C. Okan Sakar, "Accurate prediction of advertisement clicks based on impression and click-through rate using extreme gradient boosting," *ICPRAM 2019 - Proc. 8th Int. Conf. Pattern Recognit. Appl. Methods*, no. Icpram, pp. 621–629, 2019, doi: 10.5220/0007394306210629.

[17] H. Sabita, F. Fitria, and R. Herwanto, "Analisa Dan Prediksi Iklan Lowongan Kerja Palsu Dengan Metode Natural Language Programing Dan Machine Learning," *J. Inform.*, vol. 21, no. 1, pp. 14–22, 2021, doi: 10.30873/ji.v21i1.2865.

[18] Oscar, N. Maulidiah, A. Purnamawati, D. Putri, and H. F. Pardede, "Prediksi Tingkat Kesuksesan Promosi Bank Dengan Algoritma DNN." 2021.

[19] T. Çakmak, A. T. Tekin, Ç. Şenel, T. Çoban, Z. E. Uran, and C. Okan Sakar, "Accurate prediction of advertisement clicks based on impression and click-through rate using extreme gradient boosting," *ICPRAM 2019 - Proc. 8th Int. Conf. Pattern Recognit. Appl. Methods*, no. February, pp. 621–629, 2019, doi: 10.5220/0007394306210629.

[20] D. Gunawan, "Evaluasi Performa Pemecahan Database dengan Metode Klasifikasi Pada Data Preprocessing Data mining," *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 2, no. 1, p. 10, 2016, doi: 10.23917/khif.v2i1.1749.

[21] M. T. Leleuly and P. H. Gunawan, "Analysis of Feature Correlation for Music Genre Classification," *2020 8th Int. Conf. Inf. Commun. Technol. ICoICT 2020*, pp. 5–8, 2020, doi: 10.1109/ICoICT49345.2020.9166333.

[22] J. K. Kim and S. Kang, "Neural Network-Based Coronary Heart Disease Risk Prediction Using Feature Correlation Analysis," *J. Healthc. Eng.*, vol. 2017, 2017, doi: 10.1155/2017/2780501.