energies

MDPI

Article

Assessing China's Investment Risk of the Maritime Silk Road: A Model Based on Multiple Machine Learning Methods

Jing Xu [1], Ren Zhang [1,2,*], Yangjun Wang [1], Hengqian Yan [1], Quanhong Liu [1], Yutong Guo [1] and Yongcun Ren [1]

[1] Institute of Meteorology and Oceanology, National University of Defense Technology, Changsha 410073, China
[2] Collaborative Innovation Center on Meteorological Disaster Forecast, Warning and Assessment, Nanjing University of Information Science and Engineering, Nanjing 210044, China
* Correspondence: zhang_ren17@nudt.edu.cn

Abstract: The maritime silk road policy of China brings opportunities to companies relating to overseas investment. Despite the investment potentials, the risks cannot be ignored and have still not been well assessed. Considering the fact that ICRG comprehensive risk has certain subjectivity, it is not completely applicable to China's overseas investment. Therefore, based on the data of the China Statistical Yearbook and International Statistical Yearbook, a new indictor is adopted to better capture the Chinese investment risk and to make our prediction more objective. In order to acquire the ability to predict the investment risk in the future which is essential to stakeholders, machine learning techniques are applied by training the ICRG data of the previous year and Outward Foreign Direct Investment (OFDI) data of the next year together. Finally, a relative reliable link has been built between the OFDI indicator in the next year and the left ICRG indicators in the last year with both the best precision score of 86% and recall score of 86% (KNN method). Additionally, the KNN method has a better performance than the other algorithms even for high-level risk, which is more concerning for stakeholders. The selected model cannot only be used to predict an objective and reasonable investment risk level, but can also be used to provide investment risk predictions and suggestions for stakeholders.

Keywords: investment risk prediction and assessment; machine learning; deep learning; international country risk guide; K-nearest neighbor

Citation: Xu, J.; Zhang, R.; Wang, Y.; Yan, H.; Liu, Q.; Guo, Y.; Ren, Y. Assessing China's Investment Risk of the Maritime Silk Road: A Model Based on Multiple Machine Learning Methods. *Energies* **2022**, *15*, 5780. https://doi.org/10.3390/en15165780

Academic Editor: Francesco Bellotti

Received: 2 June 2022
Accepted: 15 July 2022
Published: 9 August 2022

1. Introduction

The Chinese government has long been devoted to opening up to the rest of the world, and actively guides and encourages businesses to "go global". This will facilitate the rapid transformation and upgrading of China's economy and improve the ecological environment significantly. In addition, it is of great significance for China to establish a new diplomatic situation, shape the international perception of China's peaceful growth, and increase its global discourse power.

When investing abroad, investors must frequently examine a variety of social, cultural, economic, legal, diplomatic, racial, religious, and war-related aspects. Middle Eastern nations, for instance, are abundant in fossil fuels and offer substantial investment possibilities. However, they often have less stable political environments and a greater chance of violence than other regions of the globe. This potential risk might disrupt China's investment initiatives.

In 2011, the Libyan civil war broke out. China implemented a highly successful evacuation program. The outcome of this, however, was that numerous investment projects in Libya were shelved or abandoned, affecting 75 firms and 50 significant pro-jects, with contract losses totaling $18.8 billion. Furthermore, after Sirisena was unexpectedly elected president of Sri Lanka in 2014, he altered China's policy significantly. He halted the majority of Chinese loans and investment projects, which dealt a severe blow to Chinese

businesses in Sri Lanka. The aforementioned problems occurred because government functional departments lacked investment risk assessment and loss investigation processes. The absence of appropriate methods for mitigating potential investment hazards leads to blind offshore investments being made, potentially leading to substantial losses. Therefore, it is of both scientific and practical importance to conduct an adequate investment security assessment prior to foreign investment, as this will assist businesses in gaining benefits and the nation in avoiding hazards. Currently, the study of overseas investment focuses mostly on multiple perspective analyses of foreign investment. First postulated by Hall, the cultural risk of foreign investment is that cultural differences have a substantial impact on investment activities [1]. Schinas and others analyzed the economic and financial risks of foreign investment security, noting that the "going out" strategy has a substantial impact on the global economy and trade and may increase trade efficiency, reduce energy consumption, and provide investment advantages [2]. Busse remarked that shifts in the political climate and policy discontinuities could have a big impact on the investment. This political unpredictability underlines the growing significance of political risk in international investment [3]. Zhang discussed the problem of legal risk management for international investments [4]. Colin Flint offered a new theory of outbound investment and geopolitics, proposing that China should fully address geopolitical risks in outbound investment, effectively communicate China's geopolitical narrative, and investigate and construct geopolitics with Chinese characteristics [5].

Some international organizations evaluate the investment risk of each nation based on macro quantitative analysis, with the international country risk guide being the most representative. The ICRG evaluates political, economic, and financial elements to produce an overall risk assessment. Based on the ICRG data set, Ukwueze examined the influence of political factors on international aid and proposed that donors consider political issues when making gifts [6]. Based on the ICRG data set, Javaid examined the effect of political factors on climate change and concluded that improving political quality can effectively lower climate risk [7]. The preceding research demonstrates that the data set is credible and can be used in this investigation. Nonetheless, the existing ICRG Composite Risk Rating is not fully utilized for Chinese ventures abroad. Moreover, the building procedure of ICRG indicators contains a degree of subjectivity. In contrast, machine learning and deep learning approaches may effectively mine data relationships and have been successfully implemented in remote sensing, finance, and other domains [8–10].

According to the aforementioned literature, in order to conduct an exhaustive analysis of China's foreign investment, multiple elements must be evaluated simultaneously. The conclusions of the ICRG International Comprehensive Risk Ratings are not derived from a Chinese perspective, but rather from a global one. Therefore, the results of the complete risk rating cannot be used to assess China's foreign investment risk.

Based on the actual situation of China's overseas investment over the years, this study used the data mining capabilities of machine learning and deep learning technology to overcome the subjectivity of the traditional international comprehensive risk rating model and constructed a model suitable for China's overseas investment risk assessment, providing a framework for Chinese firms' international investment decisions. The following are the primary contributions of this paper:

1. As far as we know, this study is the first to use ICRG data combined with the machine learning methods to predict China's investment risks in the Maritime Silk Road region. In the prediction process, China's foreign investment data was used to replace the weighted risk results from ICRG data, improving the assessment results' objectivity and effectiveness.
2. Machine learning and deep learning technologies were applied to the prediction model, and multi-source information of the current year was used to predict the investment risk of China in the Maritime Silk Road region in the next year, with an accuracy rate of 86%.

The remainder of the paper is structured as follows: Section 2 presents data sources and data preprocessing, Section 3 presents the method and design of comparative experiments, Section 4 presents and discusses the results of comparative experiments, and Section 5 Summarizes this study.

2. Data

2.1. ICRG

Despite the difficulties in gauging country risk exactly, it is possible to make comparisons between states by utilizing the country risk ratings provided by relevant international agencies. This article explores the association between the country risk ratings given by the American PRS Group's International Country Risk Guide (ICRG) and systemic risk. The International Credit Rating Group (ICRG) is the most professional organization that analyzes country risk; it produces monthly rating results based on those of existing rating agencies.

Political risk (PR), economic risk (ER), and financial risk (FR) are the three components of country risk according to ICRG's classification. C composite risk (CR) is also considered. Political risk is weighed by 12 indicators, economic risk by 5, and financial risk by 5. The greater the value is, the lower the risk.

Indicators such as government stability, socioeconomic conditions, investment profiles, internal conflict, external conflict, corruption, religious tensions, law and order, ethnic conflict, and bureaucratic quality between 2002 and 2019 are covered by the data used in this study. The ICRG includes both descriptive and economic data in its country analyses. The ICRG model used for forecasting financial, economic, and political risk was created in 1980 by the editors of International Reports, a widely respected weekly international financial and economic report. In response to a client's request for a comprehensive research examination of the potential dangers associated with foreign business activities, the editors developed a statistical model to evaluate the risk. They supported it with analyses, interpretation of figures, and verification of what the numbers did not indicate. The outcome was a complete system capable of measuring and comparing various types of risk across countries. One of the benefits of ICRG is that it enables users to conduct risk assessments based on ICRG data in accordance with their individual needs. ICRG models are utilized by institutional investors, banks, multinational enterprises, importers, exporters, foreign exchange dealers, and shipping companies, among others, to estimate the impact of financial, economic, and political risks on their existing and future businesses and investments.

ICRG's nation risk assessment is widely utilized in several country risk-related studies and carries a high level of credibility [6,11,12]. In this research, multiple ICRG indicators are used as surrogate variables for economic, political, and other indicators in order to model and fit the available data to forecast foreign investment risks.

2.2. OFDI

This study made use of the "China Statistical Yearbook" and "International Statistical Yearbook" from the National Bureau of Statistics. Since the statistics of China's foreign direct investment are only available for 2003 onwards, the sample period of this study was determined to be 2003–2019 and included data from the foreign direct investment stock of 214 countries and regions in the world.

2.3. Historical Situation Analysis

Based on the historical data provided by ICRG and the calculated risk values, this study drew risk zoning maps for 2004, 2009, 2014, and 2019. It can be seen from the figure that the combined risk based on ICRG data changed over time. For example, Iraq was a high-risk area in 2004 due to the second Gulf War, which took place in 2003. As the intensity of the war decreased before it then stopped, the risk in Iraq decreased but remained high for a long time.

Based on OFDI data, OFDI zoning maps of 2004, 2009, 2014, and 2019 were drawn in this study. As can be seen from the figure, the stock of OFDI changed over time. As a result of the second Gulf War in 2003, it was evident that the stock of investment in Iraq was rapidly and substantially reduced. The change seen was consistent with the overall change in risk reflected in the ICRG data.

In addition, it can be seen from Figures 1 and 2 that the stock of outbound investment is not entirely consistent with the risk assessment results of ICRG, with some differences being evident. For example, Canada's ICRG risk from 2004 to 2019 was consistently lower than that of the United States. According to the actual data of OFDI, the stock of outbound investment in the United States is always higher than that in Canada. Therefore, it is not reasonable for investors to make foreign investments solely based on the value at risk provided by the ICRG.

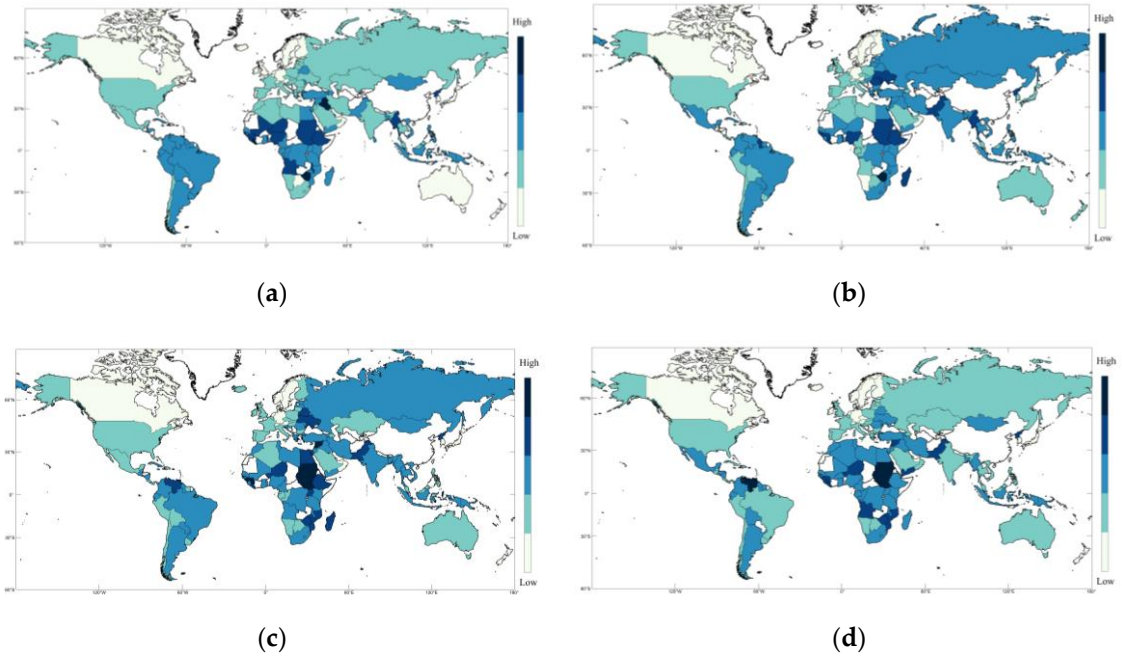


Figure 1. ICRG's Perennial Change Risk Zoning Map. (**a**–**d**) represent ICRG composite risk in 2004, 2009, 2014, 2019.

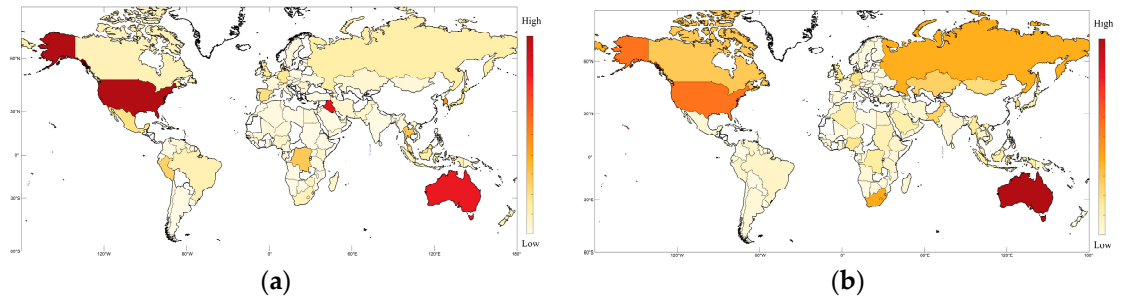
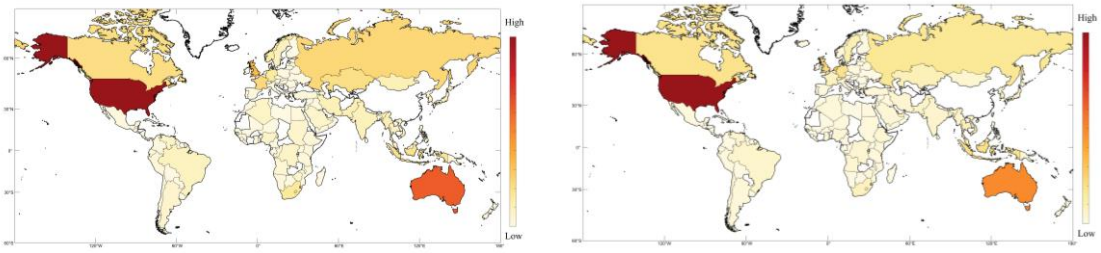


Figure 2. *Cont.*



(**c**) (**d**)

Figure 2. Foreign Direct Investment Stock's Perennial Change Zoning Map. (**a**–**d**) represent Foreign Direct Investment Stock risk in 2004, 2009, 2014, 2019.

2.4. Data Preprocessing

Twenty-five factors were considered in this study, and the data scale was not large. The existing equipment could be used for all kinds of model calculations. Therefore, this problem does not need to reduce the amount of computation required by the model, nor does it need to reduce the data dimension, and all the data can be fully used to train the model.

In this study, ICRG data from 2002 to 2017 and OFDI stock data from 2003 to 2018 were used as training sets for risk prediction model training. The Figure 3 below shows that ICRG data in 2018 and OFDI stock data in 2019 were used as test sets to test the prediction effect of the model.

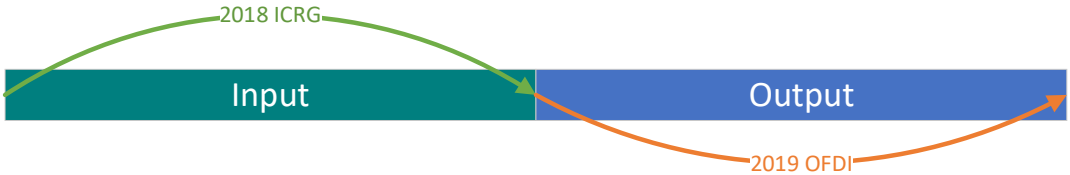


Figure 3. Schematic diagram of model input and output.

All data were normalized first, and the risk level was calibrated.

3. Methods

Based on machine learning and deep learning methods, this experiment applies these methods to predict the risk of foreign investment. In addition to comparing the SVM, XGB, LightGBM, Random Forest, and KNN machine learning models [13], this study also compares the prediction effect between the machine learning model and the deep learning model [14,15].

3.1. Machine Learning

3.1.1. SVM

The support vector machine (SVM) is a machine learning technique that combines structural risk minimization into the classification procedure. At the same time, by introducing a kernel function, the optimal classification hyperplane can be obtained while the classification error is minimized and the generalization ability of the model is improved. This method is suitable for nonlinear classification problems and can effectively overcome high-dimensional problems [16]. Due to its excellent performance, SVM is widely used in pattern recognition, time series prediction, remote sensing, image recognition, medical diagnosis, and many other fields [17,18].

3.1.2. XGB

In machine learning, the use of the XGB technique for optimizing gradient-boosting decision trees is gaining popularity. Since the modification of the GBDT loss function, XGB has become less resource-intensive, more adaptable, and more effective; as a result, it has

found widespread application in data mining, system evaluation, and other fields [19–22]. Unlike the random forest, the XGB model is additive and based on the principle of boosting integration learning. Individual learners have a strong dependence on each other. Training involves an algorithm called gradient descent, whereas greedy training applies forward distribution. Each iteration learns a tree to fit the residuals between the predicted results of the previous T-1 tree and the true value of the training sample [23]. The objective function of XGB is:

$$Obj^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t)}) + \sum_{i=1}^{t} \Omega(f_i) \quad (1)$$

where, l is the error function, which represents the deviation between the true value and the predicted value. The regularization term in Eq. can penalize $\Omega(f_i)$ complex models.

The algorithm uses the second-order Taylor formula to derive optimization loss function and leaf node grouping, and the final objective function is:

$$Obj^{(t)} = \sum_{j=1}^{T} \left[G_j \omega_j + \frac{1}{2}(H_j + \lambda)\omega_j^2 \right] + \gamma T \quad (2)$$

This algorithm can effectively reduce the occurrence of over-fitting problems, and the blocks storage structure is convenient for parallel computing, which has prominent advantages in the processing of classification regression problems [24–27].

3.1.3. LightGBM

LightGBM is an evolutionary algorithm for gradient lifting decision trees. LightGBM can further optimizes the gradient lifting decision tree through histogram algorithm, unilateral gradient sampling, mutually exclusive feature binding, leaf-wise decision tree growth strategy, and other schemes, thus accelerating the training speed of the model and improving its generalization ability while maintaining a high algorithm accuracy [28]. LightGBM is applied extensively due to its quick speed and excellent performance in multiple applications, including classification, regression, sorting, and others. As a tool for fault discovery, financial analysis, and medical diagnostics, it shows great potential [29,30].

3.1.4. Random Forest

Random forest is a classical model of ensemble learning based on the idea of a weakly supervised model decision tree. Random forest generates multiple decision trees, and each decision tree performs law fitting and merging independently to make predictions. Finally, the bagging algorithm combines and summarizes multiple decision tree classifiers to complete the final prediction [31]. In essence, random forest is an improvement of the decision tree algorithm. The generalization error of the algorithm depends on the strength of each decision tree in the random forest and the correlation between the decision trees. Dividing the features of each node by random selection can produce a lower error rate and be more robust than noise [32].

Random forest can do regression analysis on nonlinear variables. In the decision-making process, the bagging algorithm guides the aggregation, effectively reducing the over-fitting problem. The model obtained after training can classify and regression the random forest. If some trees use a particular variable during training and others do not, the importance of the variable can be obtained by comparing information values. Since bagging and selection processes do not need to satisfy linear constraints, these ideas also apply to nonlinear regression. Due to its universality, the algorithm has been widely used in biology, finance, transportation, remote sensing, computer recognition, and other fields [33,34].

Random forests have the following advantages. The current algorithm has higher accuracy. It can run efficiently on large data sets and process high-dimensional feature data without dimensionality reduction. In the random forest model, two critical parameters need to be adjusted: the number of decision trees (*n_estimators*), and the maximum number of features per tree (*max_features*).

3.1.5. KNN

KNN is a mature machine learning algorithm of nonparametric statistical methods, which has a wide range of applications in classification and regression problems. Based on the VSM model, KNN realizes the classification of samples by calculating the distance between unknown samples and all known samples according to voting rules [35].

When choosing the similarity of two instances, the Euclidean distance is generally used, that is:

$$L_p(x_i, x_j) = \left(\sum_{l=1}^{n} \left| x_i^{(l)} - x_j^{(l)} \right|^p\right)^{\frac{1}{p}} \quad (3)$$

where p is a variable parameter,$x_i \in R^n$, $x_j \in R^n$, where L_∞ is defined as:

$$L_\infty(x_i, x_j) = \max_l \left| x_i^{(l)} - x_j^{(l)} \right| \quad (4)$$

The KNN method is user-friendly, efficient, low in complexity, and effective with large data sets. Due to the advantages mentioned earlier, the KNN approach has achieved successful classification results in multiple fields, including text classification, visual recognition, and medical diagnosis [36–41].

3.1.6. Logistic Regression

Logistic regression is a generalized linear regression analysis model that belongs to supervised learning in machine learning. Its derivation and calculation are similar to the regression process but it is mainly used to solve the classification problem. The model is trained by the given N sets of data (training set), and the given one or more sets of data (test set) are classified after the training [13].

3.2. Deep Learning

DNN

Deep neural network (DNN) is a biological heuristic computing and learning model. The network takes weighted inputs from neurons or the environment, passing through processing units at each level to produce discrete or continuous outputs. The development of DNN can be traced back to the proposal of perceptron, but the single-layer perceptron has defects and cannot deal with XOR problems [42,43]. DNN solves the defect of perceptron with its multi-hidden layer structure, and the proposed gradient descent algorithm makes the parameter training of the DNN network feasible [44]. For a single neuron, the output of all neurons in the upper layer is the input of this neuron. A weighted summation of multiple inputs can obtain the neuron's output [45]. Finally, the output of all neurons in this layer is transferred to the next layer after mapping the nonlinear mapping function, which is the activation function [46].

DNN networks are mainly used to deal with classification problems, such as the risk level prediction, in this study. This network is suitable for processing lattice data. That is, each sample is from independent observation, and there is no spatial correlation [47,48].

3.3. Research Flow

3.3.1. Machine Learning

The Figure 4 below shows a flow chart for predicting risks based on machine learning methods. The specific process is as follows:

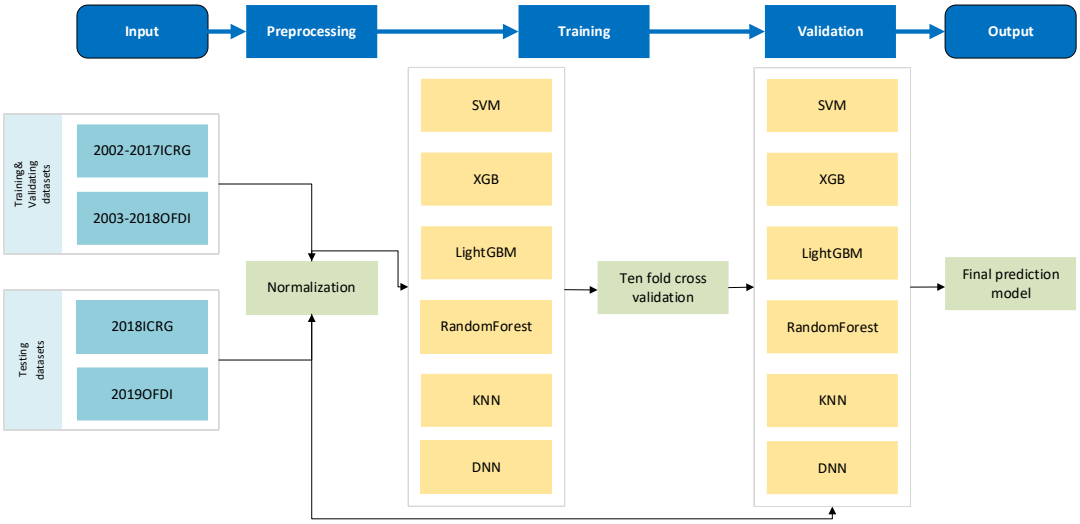


Figure 4. Flowchart of Outbound Investment Risk Prediction.

Step one. Data preprocessing. The ICRG data from 2002 to 2018 were matched with the OFDI stock data from 2003 to 2019 so that the ICRG data of the previous year could be used to predict the OFDI stock situation of the following year. The matched ICRG data in 2018 and the OFDI stock data in 2019 were taken as the test data, and the other matched data were training data. There were 2063 training data and 136 test data. Twenty-five ICRG variables were selected as the input data of the prediction model. The output was the stock of outward direct investment.

Due to the significant difference in output data, the normalization process was complex. Therefore, before training, the output data with a value of less than 782 were set to the minimum value of 0, the output data with a value greater than 113500 were set to the maximum value of 1, and the remaining data were normalized to 0-1. Input data were normalized directly in the 0–1 range.

The output data normalized from 0 to 1 were divided into five risk levels, with 1 representing the highest risk and 5 representing the lowest risk.

Step two. Training a predictive model to ensure the best results for each model classification. The most critical parameters of SVM are C and gamma. The most critical parameters of XGB are *max_depth* and *min_child_weight*. The most critical parameters of LightGBM are *num_leaves*, *min_data_in_leaf*, and *max_depth*. The most critical parameters of random forest are *n_estimators* and *max_depth*. The most critical parameters of KNN are *n_neighbors*, *leaf_size*, and *p*.

Here, the GridSearchCV method was used to obtain the optimal model hyperparameters of each machine learning model. The Figure 5 below shows the optimal parameters of each machine learning model were selected through 10-fold cross-validation.

Step three. Prediction with the trained model. Feeding the 2018 ICRG data into the model predicts the stock of OFDI. The model effect was evaluated according to the accuracy, F1, precision, and recall of the predicted foreign direct investment stock and actual data.

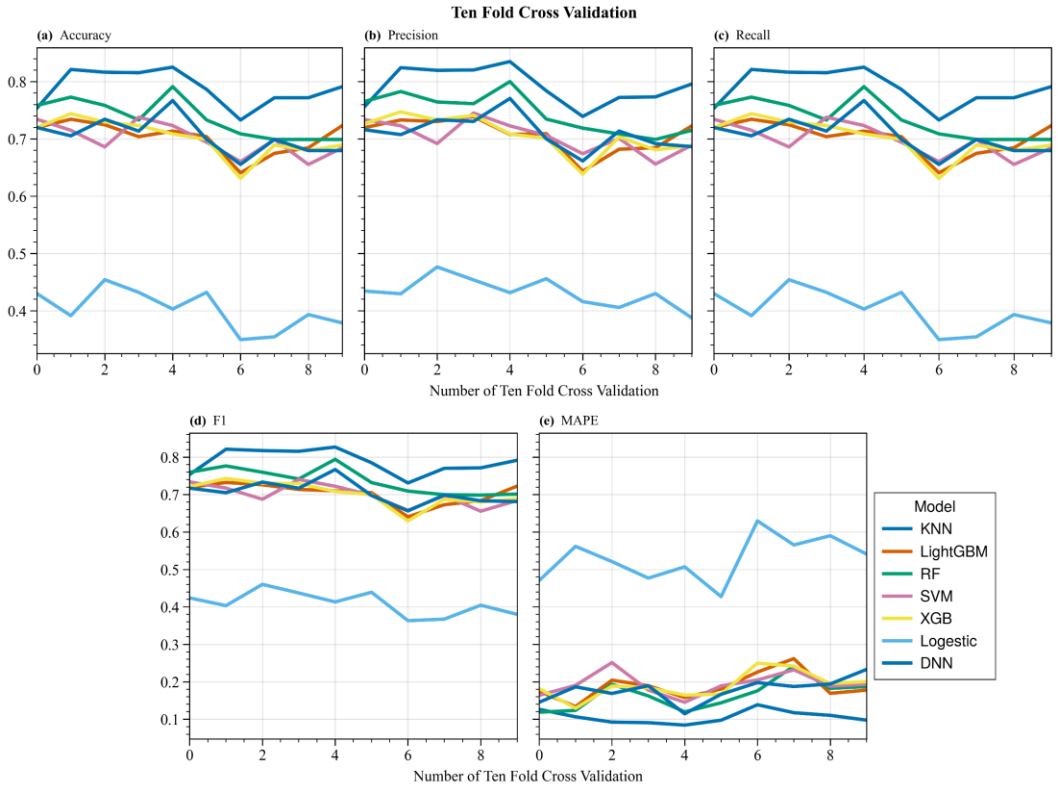


Figure 5. The results of Ten-Folder Cross Validation from panel (**a**) to panel (**e**): Accuracy, Precidsion, Recall, F1, MAPE.

3.3.2. Deep Learning

The model effect was evaluated according to the Accuracy, F1, Precision, and Recall of the predicted foreign direct investment stock and actual data.

Based on the processing of the machine learning data, after the input data of the DNN model were one-hot encoded, the risk level of the output result was determined by calculating the probability of data belonging to a particular class. The training set of DNN was 2002–2017 ICRG data and 2003–2018 OFDI data, and the test set was 2018 ICRG data and 2019 OFDI data. The DNN in this paper was built on the pytorch platform and divided into two layers, three layers, and four layers according to the number of network layers for comparative experiments.

The hyperparameters of DNN mainly include the number of neurons and activation functions in the different network layers. For networks with different layers, the input layer and hidden layer of the network used the ReLU activation function, and the output layer used the Softmax activation function. Using the ReLU activation function can improve the computational efficiency while avoiding the gradient explosion phenomenon as far as possible. The Softmax activation function was used with one-hot encoding to convert the prediction results of the model into probability representations, which facilitated the loss calculation in subsequent model training and improved the model's prediction accuracy. The number of neurons and network parameters in different layers of the model are shown in Table 1 below:

Table 1. Network Parameters of Models with Different Layers.

Layers	Number of Neurons	Output Category	Batch Size	Learning Rate	Optimizer
2	(256, 128)	5	64	0.01	Adam
3	(256, 128, 64)	5	64	0.01	Adam
4	(256, 128, 64, 64)	5	64	0.01	Adam

After the output layer of the model passes through the activation function of Softmax, the output of the model was the probability of multiple attributions, which had the following form:

$$y = (y_1, y_2, \ldots, y_c) \tag{5}$$

$$\sum_{i=1}^{C} y_i = 1 \tag{6}$$

That is, the sum of the above probabilities is 1. The loss function suitable for this kind of probability calculation is the multi-class cross-entropy loss function, and the calculation formula of this function is as follows:

$$L_k = -\sum_{i=1}^{C} t_i \log(y_i) \tag{7}$$

Therefore, the model needs to reduce the size of the above loss function as much as possible through the gradient descent algorithm during parameter training. The DNN model in this experiment used an Adam optimizer during training, which is an effective stochastic optimization method. This method combines the advantages of two optimization algorithms, AdaGrad and RMSProp, which require only one-order gradients, are computationally efficient, and require only a tiny amount of memory [49]. The batch size during training was set to 64, and the number of epochs was 600. The initial learning rate was set to 0.001. This decays with the number of iterations and the effect (if the accuracy does not improve after 50 iterations, the learning rate decays to 60% of the previous, and the minimum decay is 0.0001).

3.4. Evaluation Indicators

In order to compare the prediction effect of each model, in this study, we used the four indicators of accuracy, precision, recall, F1, and MAPE (mean absolute percentage error) value between the predicted value and the test value for the evaluation. These five indicators could evaluate the effect of the predictive model from multiple perspectives. They were calculated as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

$$precision = \frac{TP}{TP + FP} \tag{9}$$

$$recall = \frac{TP}{TP + FN} \tag{10}$$

$$F1 = \frac{precision * recall * 2}{precision + recall} \tag{11}$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{\hat{y}_i - y_i}{y_i} \right|$$

In the above formula, TP represents the True Positive class if an instance is a positive class and is predicted to be a positive class. FN represents a False Negative class if an instance is a positive class but is predicted to be a negative class. FP represents a False Positive class if an instance is a negative class but is predicted to be a positive class. TN represents the True Negative class if an instance is a negative class and is predicted to be a negative class.

4. Results and Discussion

4.1. Accuracy

Although the above models can predict investment risk for the next year during training, their effectiveness is unknown. Therefore, to test the prediction accuracy, this study compared the Accuracy, F1-score, Precision, and Recall predicted by each model. The results are shown in Table 2 below:

Table 2. Evaluation Results of Each Prediction Model.

Indicator	SVM	XGB	LightGBM	RF	KNN	Logistic	DNN
Accuracy	0.75	0.70	0.71	0.77	**0.86**	0.42	0.71
F1	0.75	0.71	0.71	0.78	**0.86**	0.42	0.71
Precision	0.78	0.72	0.73	0.80	**0.86**	0.44	0.73
Recall	0.75	0.70	0.71	0.77	**0.86**	0.42	0.71
MAPE	9.1%	20.3%	18.9%	19.1%	**4.5%**	38.5%	13.7%

As observed in the preceding graph, the KNN prediction model is the model with the best performance among multiple machine learning prediction models, whereas the XGB prediction model is the model with the worst performance.

In order to determine whether the prediction effect of machine learning is superior to that of deep learning, we created a deep neural network prediction model with multiple layers based on the data processing described above and compared its prediction results to those of various machine learning models.

The prediction model based on a four-layer deep neural network achieved the highest value for each index. The prediction model based on a three-layer deep neural network scored the lowest value across all indicators. In other words, the four-layer deep neural network prediction model had the best performance of these deep neural network prediction models, and the three-layer deep neural network prediction model had the worst performance.

Compared to machine learning predictive models, deep learning predictive models are less predictive than machine learning predictive models. The optimal model in deep learning, the four-layer deep neural network prediction model, had a little greater prediction impact than the worst prediction model in machine learning, the XGB prediction model, and the LightGBM prediction model had equivalent prediction ability. Both the two-layer and three-layer deep neural network prediction models were less accurate than all other machine learning prediction models.

In this study, the poor performance of the prediction ability of the deep learning prediction model was attributed to the fact that the deep neural network is very dependent on data and requires a big quantity of data to achieve the prediction model through learning. However, the amount of data used in the study was relatively small for deep neural networks, and so deep learning networks cannot effectively exhibit their learning capabilities. Consequently, the predictive effect was not optimal. Moreover, deep learning requires increased computational power. Under identical equipment settings, the deep learning prediction model had a longer calculation time and a higher processing cost than the machine learning prediction model.

4.2. Local Prediction Effect

To test the ability of different models to predict high-risk and low-risk, this study considered the high-risk part and the low-risk part separately, and analyzed the predictive ability of each prediction model, namely the four indicators of Accuracy, F1, precision, and recall.

It can be seen from the figure that when the SVM prediction model and the KNN prediction model faced high risks, all the prediction indicators had achieved optimal values. However, the prediction indicators of the XGB prediction model and the LightGBM

prediction model were not ideal. This shows that both the SVM and KNN models had good prediction effects in predicting high risk. The specific results are shown in Table 3 below:

Table 3. High Risk Predictive Model Evaluation Results.

Indicator	SVM	XGB	LightGBM	RF	KNN
Accuracy	**0.88**	0.5	0.5	0.62	**0.88**
F1	**0.93**	0.67	0.67	0.77	**0.93**
Precision	**1**	1	1	1	**1**
Recall	**0.88**	0.5	0.5	0.62	**0.88**

It can be seen from the figure that when the KNN prediction model faced low risk, all the prediction indicators achieved optimal values. However, the prediction indicators of the SVM prediction model were not ideal. This shows that the KNN prediction model had a good effect on predicting low risk. Overall, the models performed better at predicting low risk. The results are shown in Table 4 below:

Table 4. Low Risk Predictive Model Evaluation Results.

Indicator	SVM	XGB	LightGBM	RF	KNN
Accuracy	0.74	0.77	0.79	0.82	**0.91**
F1	0.85	0.87	0.88	0.9	**0.95**
Precision	1	1	1	1	**1**
Recall	0.74	0.77	0.79	0.82	**0.91**

4.3. Discussion

In the aforementioned experiments, the prediction effect of the deep learning DNN risk prediction model was typically inferior to that of the majority of machine learning prediction models. This may be due to the fact that the DNN model is a sort of deep learning model, and training deep learning models requires a large amount of data [50–52].

The amount of data employed in this study may have been insufficient for deep learning, hence limiting deep learning's capacity for model fitting. In the future, we will be able to use additional data to construct a risk prediction model based on DNN and other deep learning techniques in order to improve its predictive abilities.

Although the risk prediction model based on KNN has the best prediction effect among several models, this model also has certain limitations. For example, in terms of data, there is an imbalance in the samples, which may affect the prediction effect of the KNN model. Therefore, follow-up work may aim to optimize the KNN model to improve its predictive ability further.

This study uses ICRG data from each year to predict the probability of OFDI for the following year, with a one-year forecast horizon. In practice, foreign investment is long-term, and these projects frequently take several years. Taking simply one year's forecasted results into account cannot cover all investing conditions. As a result, the follow-up work may examine the use of previous ICRG data to estimate the risk of foreign direct investment in the next few years in order to remind relevant businesses to implement risk aversion measures quickly and attempt to limit the risk.

In addition, if adequate data are available, this study's prediction model can be applied to other specific industries of foreign investment, such as overseas oil investment, port investment, and construction. This research aims to provide more precise risk assessment results for associated firms' overseas investments, thereby assisting Chinese enterprises in "going global" from every viewpoint. It is necessary to develop and support global cooperative development with all nations.

5. Conclusions

As China continues to expand its scale of foreign investment, it also faces many risks related to the significant amount of foreign investment involving a wide range of countries and many types of projects. It is therefore necessary to predict investment risks effectively.

A widely used indicator which can comprehensively describe the investment risk of each country in ICRG was obtained by weighting other indicators subjectively. It was proven that this indictor could not predict the investment risk of China well.

More importantly, ICRG data are commonly used to rate the investment risk in the previous year and do not have the ability to predict the investment risk in the future, which stakeholders are eager to know.

In this paper, a new indictor named OFDI was adopted to better capture the Chinese investment risk and to make our prediction more objective. As the new indicator had no direct link to the other indictors of ICRG, some new techniques (e.g., machine learning algorithms and deep learning methods) were adopted to obtain the prediction ability by training the ICRG data of the previous year and OFDI data of the next year together.

All models were well trained based on 10-fold cross validation, and different experiment schemes were designed to test the model robustness and prediction accuracy in terms of multiple aspects.

Finally, we built a relative reliable link between the OFDI indicator in the next year and the left ICRG indicators in the last year with both the best precision score of 86% and recall score of 86% (KNN method). Additionally, the KNN method had a better performance than the other algorithms even for high-level risk, being more concerning for stakeholders.

The selected model cannot only be used to predict an objective and reasonable investment risk level, but can also be used to provide investment risk predictions and suggestions for stakeholders.

Author Contributions: Conceptualization, J.X.; Data curation, J.X.; Formal analysis, J.X.; Funding acquisition, R.Z.; Methodology, J.X.; Project administration, R.Z. and Y.W.; Software, J.X., H.Y. and Q.L.; Supervision, R.Z., Q.L., Y.G. and Y.R.; Validation, Q.L.; Visualization, J.X. and H.Y.; Writing—original draft, J.X.; Writing—review & editing, J.X. and Y.W. All authors have read and agreed to the published version of the manuscript.

Funding: This study is supported by the Chinese National Natural Science Fund (No. 41976188).

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

References

1. Hall, S.; Du Gay, P. *Questions of Cultural Identity*; Sage: Thousand Oaks, CA, USA, 1996.
2. Schinas, O.; von Westarp, A.G. Assessing the Impact of the Maritime Silk Road. *J. Ocean. Eng. Sci.* **2017**, *2*, 186–195. [CrossRef]
3. Busse, M.; Hefeker, C. Political Risk, Institutions and Foreign Direct Investment. *Eur. J. Polit. Econ.* **2007**, *23*, 397–415. [CrossRef]
4. Zhang, X. Interpret the Legal Risk Management of Overseas Investment under the New Situation of "One Belt and One Road". *Int. Eng. Labor* **2015**.
5. Flint, C.; Zhu, C. The Geopolitics of Connectivity, Cooperation, and Hegemonic Competition: The Belt and Road Initiative. *Geoforum* **2019**, *99*, 95–101. [CrossRef]
6. Ukwueze, E.R.; Ugwu, U.C.; Okafor, O.A. Impact of Institutional Quality on Multilateral Aid in Nigeria. *J. Econ. Sci. Res.* **2021**, *4*, 3116. [CrossRef]
7. Javaid, A.; Arshed, N.; Munir, M.; Zakaria, Z.A.; Alamri, F.S.; Abd El-Wahed Khalifa, H.; Hanif, U. Econometric Assessment of Institutional Quality in Mitigating Global Climate-Change Risk. *Sustainability* **2022**, *14*, 699. [CrossRef]
8. Mullainathan, S.; Spiess, J. Machine Learning: An Applied Econometric Approach. *J. Econ. Perspect.* **2017**, *31*, 87–106. [CrossRef]
9. Lecun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
10. Kamilaris, A.; Prenafeta-Boldú, F.X. Deep Learning in Agriculture: A Survey. *Comput. Electron. Agric.* **2018**, *147*, 70–90. [CrossRef]
11. Brown, C.L.; Cavusgil, S.T.; Lord, A.W. Country-Risk Measurement and Analysis: A New Conceptualization and Managerial Tool. *Int. Bus. Rev.* **2015**, *24*, 246–265. [CrossRef]
12. Gezikol, B.; Tunahan, H. The Econometric Analysis of the Relationship between Perceived Corruption, Foreign Trade and Foreign Direct Investment in the Context of International Indices. *Alphanumeric J.* **2018**, *6*, 117. [CrossRef]
13. Rawson, A.; Brito, M. A Survey of the Opportunities and Challenges of Supervised Machine Learning in Maritime Risk Analysis. *Transp. Rev.* **2022**, 1–23, *in press*. [CrossRef]

14. Xiong, Y.; Zhu, M.; Li, Y.; Huang, K.; Chen, Y.; Liao, J. Recognition of Geothermal Surface Manifestations: A Comparison of Machine Learning and Deep Learning. *Energies* **2022**, *15*, 2913. [CrossRef]
15. Akyuz, E.; Cicek, K.; Celik, M. A Comparative Research of Machine Learning Impact to Future of Maritime Transportation. *Procedia Comput. Sci.* **2019**, *158*, 275–280. [CrossRef]
16. Pisner, D.A.; Schnyer, D.M. Support Vector Machine. In *Machine Learning: Methods and Applications to Brain Disorders*; Mechelli, A., Vieira, S., Eds.; Elsevier: Amsterdam, The Netherlands, 2019; pp. 101–121. [CrossRef]
17. Christopher, J.C. Burges A Tutorial on Support Vector Machines for Pattern Recognition. *Data Min. Knowl. Discov.* **1998**, *2*, 121–167.
18. Ali, W.; Tian, W.; Din, S.U.; Iradukunda, D.; Khan, A.A. Classical and Modern Face Recognition Approaches: A Complete Review. *Multimed. Tools Appl.* **2021**, *80*, 4825–4880. [CrossRef]
19. Pan, B. Application of XGBoost Algorithm in Hourly PM2.5 Concentration Prediction. *IOP Conf. Ser. Earth Environ. Sci.* **2018**, *113*, 12127. [CrossRef]
20. Nobre, J.; Neves, R.F. Combining Principal Component Analysis, Discrete Wavelet Transform and XGBoost to Trade in the Financial Markets. *Expert Syst. Appl.* **2019**, *125*, 181–194. [CrossRef]
21. Gumus, M.; Kiran, M.S. Crude Oil Price Forecasting Using XGBoost. In Proceedings of the 2017 International Conference on Computer Science and Engineering (UBMK), Antalya, Turkey, 5–8 October 2017; pp. 1100–1103. [CrossRef]
22. Zhang, D.; Qian, L.; Mao, B.; Huang, C.; Huang, B.; Si, Y. A Data-Driven Design for Fault Detection of Wind Turbines Using Random Forests and XGboost. *IEEE Access* **2018**, *6*, 21020–21031. [CrossRef]
23. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference, San Francisco, CA, USA, 13–17 August 2016.
24. Shi, X.; Wong, Y.D.; Li, M.Z.F.; Palanisamy, C.; Chai, C. A Feature Learning Approach Based on XGBoost for Driving Assessment and Risk Prediction. *Accid. Anal. Prev.* **2019**, *129*, 170–179. [CrossRef]
25. Ren, X.; Guo, H.; Li, S.; Wang, S. A Novel Image Classi Fi Cation Method with CNN-XGBoost Model. In Proceedings of the 16th International Workshop, IWDW 2017, Magdeburg, Germany, 23–25 August 2017; pp. 378–390. [CrossRef]
26. Ogunleye, A.; Wang, Q.G. XGBoost Model for Chronic Kidney Disease Diagnosis. *IEEE ACM Trans. Comput. Biol. Bioinform.* **2020**, *17*, 2131–2140. [CrossRef] [PubMed]
27. Fan, J.; Wang, X.; Wu, L.; Zhou, H.; Zhang, F.; Yu, X.; Lu, X.; Xiang, Y. Comparison of Support Vector Machine and Extreme Gradient Boosting for Predicting Daily Global Solar Radiation Using Temperature and Precipitation in Humid Subtropical Climates: A Case Study in China. *Energy Convers. Manag.* **2018**, *164*, 102–111. [CrossRef]
28. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.Y. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017; pp. 3147–3155.
29. Tang, M.; Zhao, Q.; Ding, S.X.; Wu, H.; Li, L.; Long, W.; Huang, B. An Improved LightGBM Algorithm for Online Fault Detection of Wind Turbine Gearboxes. *Energies* **2020**, *13*, 807. [CrossRef]
30. Wang, D.; Zhang, Y.; Zhao, Y. LightGBM: An Effective MiRNA Classification Method in Breast Cancer Patients. In Proceedings of the 2017 International Conference on Computational Biology and Bioinformatics, Newark, NJ, USA, 18–20 October 2017; pp. 7–11. [CrossRef]
31. Javier, P.; Camino, G.; Javier, M. Electricity Price Forecasting with Dynamic Trees: A Benchmark Against the Random Forest Approach. *Energies* **2018**, *11*, 1588.
32. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
33. Liaw, A.; Wiener, M. Classification and Regression by Random Forest. *R News* **2002**, *2/3*, 18–22.
34. Karballaeezadeh, N.; Ghasemzadeh Tehrani, H.; Mohammadzadeh Shadmehri, D.; Shamshirband, S. Estimation of Flexible Pavement Structural Capacity Using Machine Learning Techniques. *Front. Struct. Civ. Eng.* **2020**, *14*, 1083–1096. [CrossRef]
35. Guo, G.; Wang, H.; Bell, D.; Bi, Y.; Greer, K. KNN Model-Based Approach in Classification. In Proceedings of the On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE, Sicily, Italy, 3–7 November 2003; pp. 986–996. [CrossRef]
36. Zhang, M.L.; Zhou, Z.H. ML-KNN: A Lazy Learning Approach to Multi-Label Learning. *Pattern Recognit.* **2007**, *40*, 2038–2048. [CrossRef]
37. Zhang, S.; Li, X.; Zong, M.; Zhu, X.; Wang, R. Efficient KNN Classification with Different Numbers of Nearest Neighbors. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 1774–1785. [CrossRef]
38. Weinberger, K.Q.; Saul, L.K. Distance Metric Learning for Large Margin Nearest Neighbor Classification. *J. Mach. Learn. Res.* **2009**, *10*, 207–244.
39. Zhang, H.; Berg, A.; Maire, M.; Malik, J. SVM-KNN: Discriminative Nearest Neighbor Classification for Visual Category Recognition. In Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, NY, USA, 17–22 June 2006; pp. 2126–2136.
40. Trstenjak, B.; Mikac, S.; Donko, D. KNN with TF-IDF Based Framework for Text Categorization. *Procedia Eng.* **2014**, *69*, 1356–1364. [CrossRef]
41. Pratama, B.Y.; Sarno, R. Personality Classification Based on Twitter Text Using Naive Bayes, KNN and SVM. In Proceedings of the 2015 International Conference on Data and Software Engineering (ICoDSE), Yogyakarta, Indonesia, 25–26 November 2016; pp. 170–174. [CrossRef]

42. Nguyen, A.; Yosinski, J.; Clune, J. Deep Neural Networks Are Easily Fooled: High Confidence Predictions for Unrecognizable Images. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015.
43. Moosavi-Dezfooli, S.M.; Fawzi, A.; Frossard, P. DeepFool: A Simple and Accurate Method to Fool Deep Neural Networks. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016.
44. Seltzer, M.L.; Dong, Y.; Wang, Y. An Investigation of Deep Neural Networks for Noise Robust Speech Recognition. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013.
45. Lee, S.; Cho, S.; Kim, S.H.; Kim, J.; Chae, S.; Jeong, H.; Kim, T.; Sciubba, E. Deep Neural Network Approach for Prediction of Heating Energy Consumption in Old Houses. *Energies* **2020**, *14*, 122. [CrossRef]
46. Sun, Y.; Liang, D.; Wang, X.; Tang, X. DeepID3: Face Recognition with Very Deep Neural Networks. *Comput. Sci.* **2015**. [CrossRef]
47. Montavon, G.; Samek, W.; Müller, K.R. Methods for Interpreting and Understanding Deep Neural Networks. *Digit. Signal Process.* **2018**, *73*, 1–15. [CrossRef]
48. Veselý, K.; Ghoshal, A.; Burget, L.; Povey, D. Sequence-Discriminative Training of Deep Neural Networks. *Proc. Interspeech* **2013**, *2013*, 2345–2349.
49. Kingma, D.; Ba, J. Adam: A Method for Stochastic Optimization. *Comput. Sci.* **2014**. [CrossRef]
50. Kaplan, H.; Tehrani, K.; Jamshidi, M. A Fault Diagnosis Design Based on Deep Learning Approach for Electric Vehicle Applications. *Energies* **2021**, *14*, 6599. [CrossRef]
51. Kaplan, H.; Tehrani, K.; Jamshidi, M. Fault Diagnosis of Smart Grids Based on Deep Learning Approach. In Proceedings of the 2021 World Automation Congress (WAC), Taipei, Taiwan, 1–5 August 2021; pp. 164–169. [CrossRef]
52. Fontana, V.; Blasco, J.M.D.; Cavallini, A.; Lorusso, N.; Scremin, A.; Romeo, A. Artificial Intelligence Technologies for Maritime Surveillance Applications. In Proceedings of the 2020 21st IEEE International Conference on Mobile Data Management (MDM), Versailles, France, 30 June–3 July 2020; pp. 299–303. [CrossRef]