Published by
NIGERIAN SOCIETY OF PHYSICAL SCIENCES
Available online @ https://asr.nsps.org.ng/index.php/asr

African Scientific Reports 3 (2024) 192

African Scientific Reports

Predicting long-term deposit customers using convolutional neural network and data conversion technique

Adebayo Abdulganiyu Keji[a,*], Oluwafemi Fakeye[a], Nneka N. Onochie[b], Olumide Sangotoki[a]

[a]Department of Computer Science, Faculty of Computing Studies, Nile University of Nigeria, Abuja FCT, 900001 Nigeria
[b]Department of Electronic Engineering, Faculty of Engineering, University of Nigeria, Nsukka, 410001 Enugu State, Nigeria

Abstract

The banking industry is the foundation of any nation's economy, and bank deposits are its primary source of profitability. Bank deposits play a significant role in determining a nation's saving rate. Globalization has resulted in substantial technological changes, business strategy, and customer service across many industries, including the financial services sector. This study proposes a deep learning model using Residual Network architectural design and transfer learning on a Portuguese banking institution containing 40,811 training data, with 36,202 belonging to label 0 and 4639 belonging to label 1. Clearly, this shows a significant level of bias between the two labels. Hence, a SMOTE method of class balancing was applied. This dataset, in comma-separated value (CSV), was converted into images coupled with the weight transfer from the residual network trained on ImageNet; our fully connected layer was built and trained with the image files. The highest performance reached by the conventional machine learning models, Random Forest (RF), is 90.78% for accuracy, 59.37% precision, 96.78% recall, and 85.28% F1 score, tested on 412 test samples. However, our proposed methodology achieves an outstanding result with an accuracy of 93.00%, 97.00% precision, 90.00% recall, 93.00% F1 score, and 94.00% ROC, with test samples of size 5601. Since long-term deposits are necessary for the banking system to fund the individual, corporate, and industrial loans needed for the country's growth and development, these results will provide an effective and reliable marketing technique required to determine the target population.

DOI:10.46481/asr.2024.3.3.192

Keywords: Machine learning, Banking, CNN, Deep learning, Marketing

Article History :
Received: 03 March 2024
Received in revised form: 21 July 2024
Accepted for publication: 28 August 2024
Published: 12 September 2024

1. Introduction

The banking industry is the foundation of any nation's economy, and bank deposits are its primary source of profitability. Bank deposits play a significant role in determining a nation's saving rate [1]. Mobilizing deposits is a direct banking action that boosts their effectiveness [2]. The majority of businesses mostly rely on bank loans for funding, and deposits have a big impact on how much money is available for lending. Consequently, the significance of deposits cannot be overstated [3]. Due to the growth in banking and financial crises over the previous 25 years, there has been a great deal of research conducted on the causes of crises and how they affect the real economy [4]. There are vast numbers of notable factors that affect clients' belief in bank savings, which

*Corresponding author: Tel. No.: +234-706-623-1665.
Email address: Gleebayour@gmail.com (Adebayo Abdulganiyu Keji)

are but are not limited to constant interest rates, country-specific wage rates, total annual bank loss, and regulatory and institutional strength, among others [5].

All businesses, regardless of size of operation, have been impacted by the globalization of economies. Globalization has resulted in significant technological changes, business strategy, and customer service across many industries, including the financial services sector. These shifts have become a source of competitive advantage for industries like education, healthcare, finance, and insurance [6]. The banking sector has been significantly impacted by globalization, which has led to a revolution in how it offers cutting-edge technologies, customer service, increased productivity, and enhanced profitability worldwide [7]. The banking industry is facing increased competition in adopting new technologies. As a result of increased competition, banking companies began utilizing the cutting-edge, technology-enabled marketing strategy known as telemarketing [8]. A company's lifeblood is its marketing, which comes in various forms.

One of the common advertising strategies used by retail banks to advertise their goods is telemarketing. A direct marketing strategy is employed to reach consumers and achieve particular goals. Managing the promotion activities section will be more accessible by concentrating on client connection through remote communication in a contact center [9]. But in order to maximize client value and build a more meaningful, long-lasting relationship that aligns with business goals, a telemarketing business plan must be developed. Classifying the traits or criteria of potential clients is necessary to accomplish this; otherwise, the technique may need to be more effective [10]. A couple of techniques have been used, including machine learning.

Computer science subfields that collectively focus on classification, prediction, and clustering or grouping tasks are called machine learning. Machine learning (ML) is a field that creates algorithms designed to be applied to datasets. Machine learning can be classified into two primary categories: supervised and unsupervised [11]. ML has gained momentum in almost every application, ranging from the healthcare system, the agricultural industry, the banking sector, speech recognition, and to management. However, we are proposing the application of deep learning, a subfield of machine learning to predict customers who can afford long term deposits in the banking industry based on their past transaction records. This would enable the financial industries to have a customer focus which will in turn reduce cost of marketing and improve the performance of the banking industries.

2. Related work

An overview of the relevant literature and key concepts for the suggested investigation is provided below. In addition, we examine recent research on predicting the possibility of customers to subscribe for long-term deposits that use conventional and artificial neural network techniques.

Kim *et al*. [12] developed a deep neural network with an initial weight of 0.05 and convolutional layer varied from 1 to 3. The data set employed in this study was gathered for 30 months with an instance of 45,211. The proposed deep neural network obtained an accuracy peaking at 76.70%. Grosicki [13] presented an artificial neural network that was applied to the campaign prediction of telemarketing in the banking ecosystem. This study employed two types of model training, which are supervised and unsupervised learning. However, 89.8% accuracy was obtained. However, there is a need in the improvement of the developed model performance. Farooqi and Iqbal [14] employed data mining techniques to predict customer subscriptions for long-term deposits. Five data mining techniques were employed in the study and they were evaluated using 14 parameter classifiers. The dataset used in this study was analyzed using the Waikato environment for knowledge analysis. The decision tree achieved the best performance compared to other employed models with an accuracy of 91.2%, error rate of 8.8%, specificity of 95.9% and sensitivity of 53.8%. This study does not consider parameter selection and data balancing, which may introduce bias to the minority class of the developed model. Patwary [15] created a group innovative technique to forecast long-term deposits subscriber. The dataset which included 45,211 entries with 16 features was obtained for the study. The naïve bayes, neural network, and support vector machine were adopted for their experiment. 96.62%, 99.08%, and 97.17% accuracy, specificity, and sensitivity respectively was obtained for the best ensemble model. This research work doesn't give priority to feature selection, in the sense that little or no data preprocessing was done on the dataset. Muslim *et al*. [16] emphasis is on managing bank resources by effectively telemarketing their process of delivering products.

This study contends that the absence of data-balancing strategies throughout development is the primary cause of the significantly low accuracy of the previously established models. However, they used the SMOTE data balancing technique with LightGBM, logistic regression, and random forest model to increase the accuracy of the current models. Using training and testing samples of roughly 32,950 and 6,590 instances, respectively, the random forest model yields an accuracy of 90.34%, the logistic regression yields an accuracy of 88.89%, and the LightGBM yields an accuracy of 90.63%. Appropriate feature selection received little to no attention in this work, which may have a direct effect on how well the created model performs. Rony [17] focused on learning the peculiarities of his customers in order to increase sales of banking products. Principal component analysis, factor analysis, correlation analysis, and exploratory data analysis are some of the statistical techniques being used to preprocess the dataset employed in this study. With 90.64 and 99.05% accuracy and sensitivity, respectively, the LR model performs the best out of all of these. But instead of using a contact-based dataset like this study did, it was proposed that real-world or client-based data be investigated for the case study's prediction. Also, there is room for improvement in the obtained accuracy of their work. Borugadda [8] used machine learning techniques to incorporate telemarketing into the banking industry to encourage customers to subscribe for long-term deposits. For this work, five distinct classic machine learning techniques—DT, RF, GNB, LR, and SVM—were used using the

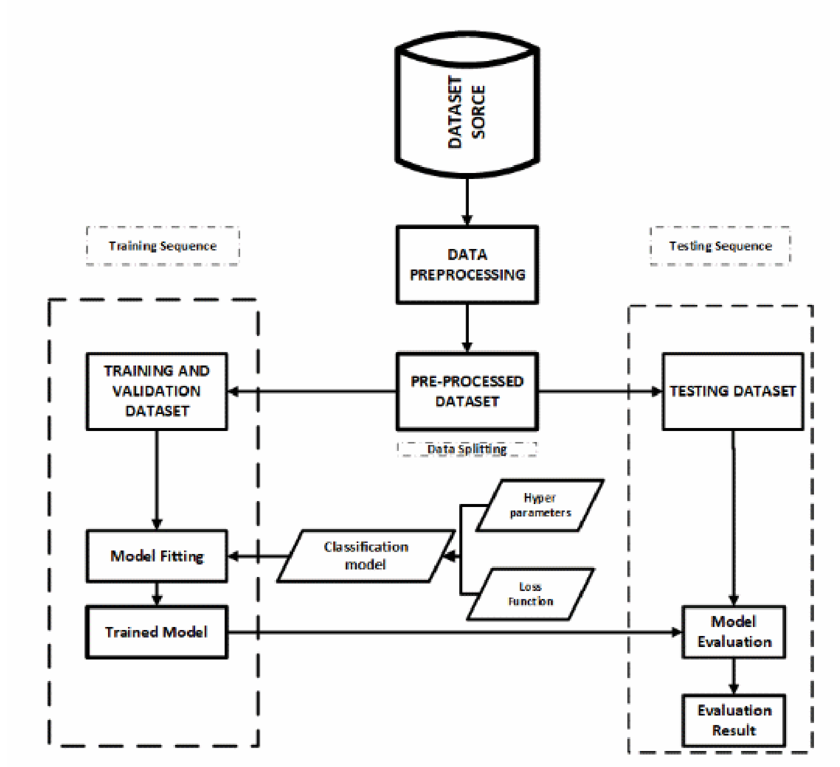

Figure 1: Design framework.

Portuguese bank dataset. The results of the experiment showed that logistic regression outperform other models with an accuracy of 92.48%. However, as feature extraction and data balancing strategies can greatly improve model performance, they are not considered in this study. However, there are notable gaps in the predictions of potential clients for long-term deposits currently. Research from the studied literatures shows that while most researchers employed shallow machine learning approaches, none of them considered the application of deep neural network and transfer learning on a dataset with a plentiful number of instances, which may act as an ideal source domain. We also observed that only a small number of the examined publications looked at ways to address the class imbalance in the dataset that was being used. In this study, we suggest applying deep neural network to forecast whether a consumer will sign up for long-term deposits in the banking sector. The synopsis of the literatures reviewed in this research study are summarized in Table 1.

Many studies have been conducted in this field, however the literatures only include traditional machine-learning methods. Nonetheless, this study is innovative since it employs deep neural networks to address the performance issue in this field of study. The study's other driving force is the requirement for a noteworthy and trustworthy method to pinpoint target consumers for the banking sector's efficient long-term deposit marketing in order to finance business, industry, and individual loans that are essential for the country's expansion and development.

3. Methodology

The direct marketing dataset information run by a Portuguese bank was adopted for this study. This dataset contains 45211 instances and 17 columns. This dataset was preprocessed by transforming the categorical columns into numerical variables to enable us fit the dataset into the model. Furthermore, important features were selected and we applied the SMOTE over sampling data balance technique to eradicate bias that could be introduced as a result of the minority class present in the dataset. The novelty of the study lies in the conversion of the CSV dataset format to image to enable us apply Convolutional Neural Network for classification of long-term deposit prediction. Moving forward, we however split the dataset into validation, training and test set. This set of datasets were employed in validating, training, and testing of our developed model. To measure the performance of the developed model the accuracy, precision, F1 score and recall metrics were employed. The design framework of this research study is shown in Figure 1.

3.1. Dataset description

The data employed to carry out this study was obtained from the UCL repository. The direct marketing dataset information run by a Portuguese bank was adopted for this study. This marketing was based on conversation via phone. They were determining if

Table 1: Summary of related works.

Research paper	Model	Strength	Weakness	Performance metrics
[18]	The model used should have been mentioned.	dataset balancing technique with a correlation-based feature selection	The model wasn't mentioned; thus, the study needs more transparency.	True positive and ROC
[19]	The model name wasn't stated	The inclusion of customer lifetime value (LTV)	Dataset class imbalance	ROC and ALIFT
[12]	Deep neural network (1-3) and six traditional machine learning	Binarization of the nominal and categorical input data.	76.7% accuracy	Accuracy
[20]	14 different models	the smooth-threshold estimating equation, smoothly-clipped absolute deviation, elastic net, and lasso for feature selection	Dataset class imbalance	AUC, mean, and standard deviation
[13]	Autoencoder denoising and regularized neural network	Supervised and non-supervised model training.	Dataset class imbalance and lack of feature selection	Accuracy
[14]	NB, J48 DT, NN, KNN, and sequential minimal optimization	Evaluated using 14 different parameter classifiers	Dataset class imbalance and lack of feature selection	Specificity, sensitivity, Error rate, and accuracy
[21]	Random under-sampling, bagging, and cluster-based under-sampling methods.	Ensemble learning approach	Selective feature extraction	Accuracy, Area under the curve, TPR, and BCR
[8]	DT, RF, GNB, SVM and LR	The model achieves 92.48% accuracy.	Imbalance data set, lack of selection of important feature.	Accuracy
[17]	KNN, SVM, RF, and LR	factor analysis, EDA, Principal component analysis.	Imbalance dataset.	Accuracy and Sensitivity.
[16]	LightGBM, LR, and random forest	Application of data balancing technique (SMOTE).	Sufficient feature selection received little or no attention.	Accuracy
[15]	Naïve bayes, Neural network and SVM	Ensemble Technique, sensitivity of 97.17%, Accuracy of 96.62 and Specificity of 99.08%.	Imbalance dataset.	Specificity, sensitivity, and accuracy.
[22]	HLVQ, SOM, SMO, SVM and RBNF	Relief-F attribute, one-R attribute, and filtered attributes evaluator for dataset preprocessing.	Dataset class imbalance	Accuracy

the client would opt ('yes') or not ('no'). Table 2 shows the description of the dataset used for training, validation and testing of the predictions of potential clients for long-term deposits model. Table 3 identifies the name and link to the dataset used in this study.

3.2. Data analysis

The data analysis of the features present in the dataset is conducted to enhance our understanding of the dataset we are working with. There are 11 occupations named in the dataset, and other occupations were used to represent other possible occupations. Figure 2a shows the distribution of clients to their various occupations as present in the dataset. According to the dataset, Blue-collar jobs have the highest percentage of workers. Figure 2b show three different marital statuses present in the dataset: married, single, and divorced. The percentage of married customers is higher than the other two statuses. The status of a customer can affect his/her decision to subscribe for a long-term deposit. Figure 2c shows how a client's education level can affect his/her decision to subscribe

Table 2: Data description.

Features	Data type
Age	Numeric
Job	Categorical: "admin.", "blue-collar", "entrepreneur", "housekeeper", "management", "retired", "self-employed", "services", "student", "technician", "unemployed", and "unknown"
Marital	Categorical: 'divorced', 'married', 'single', 'unknown'; note: 'divorced' means divorced or widowed
Education	Categorical: 'basic.4y', 'basic.6y', 'basic.9y', 'high.school', 'illiterate', 'professional.course', 'university.degree', 'unknown'
Default	Categorical: 'no', 'yes', 'unknown.'
Loan	Categorical: 'no', 'yes', 'unknown')
Contact	Categorical: 'cellular', 'telephone'
Month	Categorical: 'jan', 'feb', 'mar', ..., 'nov', 'dec'
Day	Categorical: 'mon', 'tue', 'wed', 'thu', 'fri')
Duration	Numeric
campaign	Number of contacts performed during this campaign and for this client (numeric, includes the last contact)
pdays	Number of days that passed by after the client was last contacted from a previous campaign (numeric; 999 means clients were not previously contacted)

Table 3: Data description.

Data	Link
Bank marketing	http://archive.ics.uci.edu/dataset/222/bank+marketing [23]

for a long-term deposit. The customers in their secondary education have the highest percentage, followed by customers in tertiary education. The housing feature tells us more about customers who currently have accommodation. This feature can also influence the decision to subscribe for long-term deposits. As seen in Figure 2d, the customers with housing have more percentages in terms of numbers. The loan feature tells us more about customers' loan status. This feature can also influence the decision to subscribe for long-term deposits. As seen in Figure 2e, the customers without an active loan subscription have a higher percentage in numbers. The response binary represents the customers' decision after the conversation with the Portuguese bank official. This can be referred to as a label in the machine learning model. 88.6% of the customer population choose not to subscribe for long-term deposits, while 11.4% prefer to subscribe, as seen in Figure 2.

3.3. *Data preprocessing*

The actions made to improve the data gathered make them far more valuable for training the machine learning algorithm. Before our data is fed into our model, several prior transformations were made to it, which include data mapping, data balancing and transformation of data. The categorical features were mapped to numerical values. Also, we employed the Synthetic Minority Oversampling Technique (SMOTE) to balance our dataset and the Comma Separated Value (CSV) dataset was converted to images to make it possible for us to apply deep learning coupled with transfer learning to achieve our aim and objectives. Figure 3 shows the pictorial representation of the converted dataset.

3.4. *Classification algorithm*

The model employed in this research study is the Residual Network (ResNet). The transfer of learning from ResNet trained on the ImageNet dataset was used to optimize the ResNet model. This expedites the training process and enhances the model's functionality. The base network's weight was frozen to accomplish transfer learning, and the fully connected layer—the "head"—was eliminated. This allowed us to add a series of fully connected layers that would be trained using our data collection. Figure 4 shows the Residual Network architectural network.

Training a very deep neural network is a challenging task but the advantage of their depth in terms of model performance is astonishing. Exploding gradient is one of the shortcomings of increased depth which in turn disturb the convergence of the models. Thus, the novel deep residual learning framework by Ref. [24] to address the varnishing problem of the gradients of deep networks.

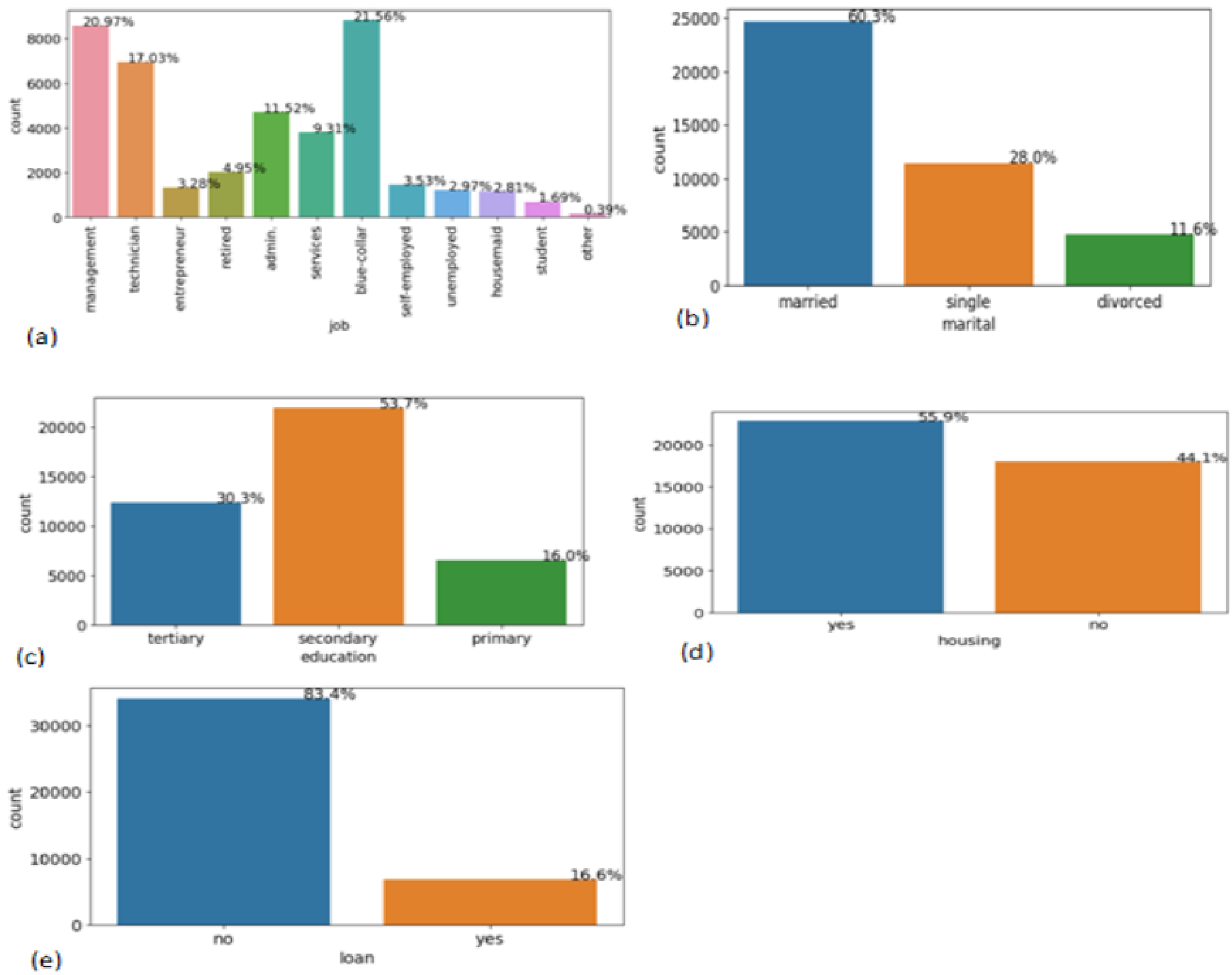

Figure 2: Occupation data analysis (b) Analysis of marital status feature (c) Analysis of educational feature (d) Analysis of housing feature) (e) Analysis of the customer response.

Representing an input to a layer with x and we have several layers stacked together to fit a mapping $\mathcal{M}(x)$. These stacked layers are then approximated to a residual function in equation (1) instead of the underlying mapping $\mathcal{M}(x)$ [24].

$$F(x) := M(x) - x. \tag{1}$$

This function thus become:

$$M(x) = F(x) + x. \tag{2}$$

Hence, the weight of the multiple layers can be drive to zero by solver with the residual learning reformation if the identity mapping are optimal [24]. Defining a building block as:

$$y = A + x, \tag{3}$$

where x is the input to the layer, y is the output from the layer and A denotes the residual mapping that is to be learn [24].

$$A = F(x, w_i). \tag{4}$$

For the layers in Figure 5,

$$F = W_2\sigma(W_1 x), \tag{5}$$

where σ denotes the activation function (ReLu). However, multiple convolutional layers can be represented by the function A and thus obtaining a residual mapping F depending on the number of layers without increase in computational cost.

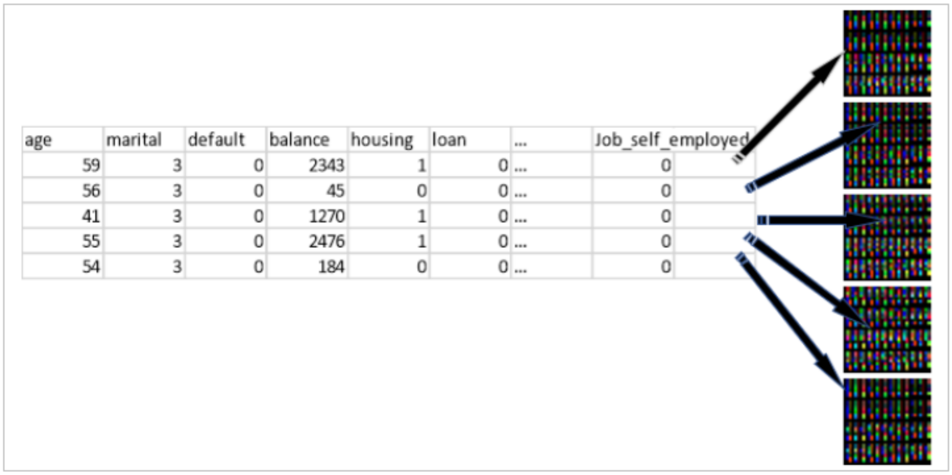

Figure 3: Conversion of CSV data format to images.

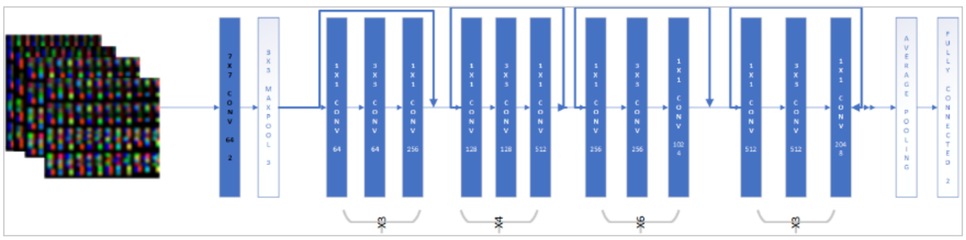

Figure 4: Residual Network (ResNet) architecture.

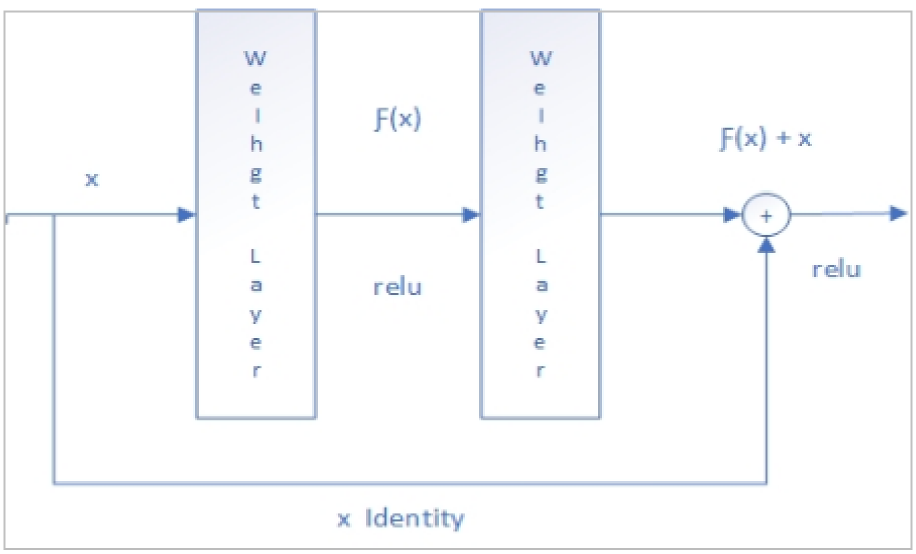

Figure 5: Building block.

4. Results and discussion

The Portuguese banking dataset was collected from an online dataset repository and was preprocessed by encoding the categorical features using the ordinal encoding method and converted to images using the OpenCV python library. Moving forward, both the traditional machine learning and deep neural network models were implemented and trained using the preprocessed dataset (images). We implement, evaluate, and visualize the results of our model using the open-source Python libraries; TensorFlow, Keras, Scikit-

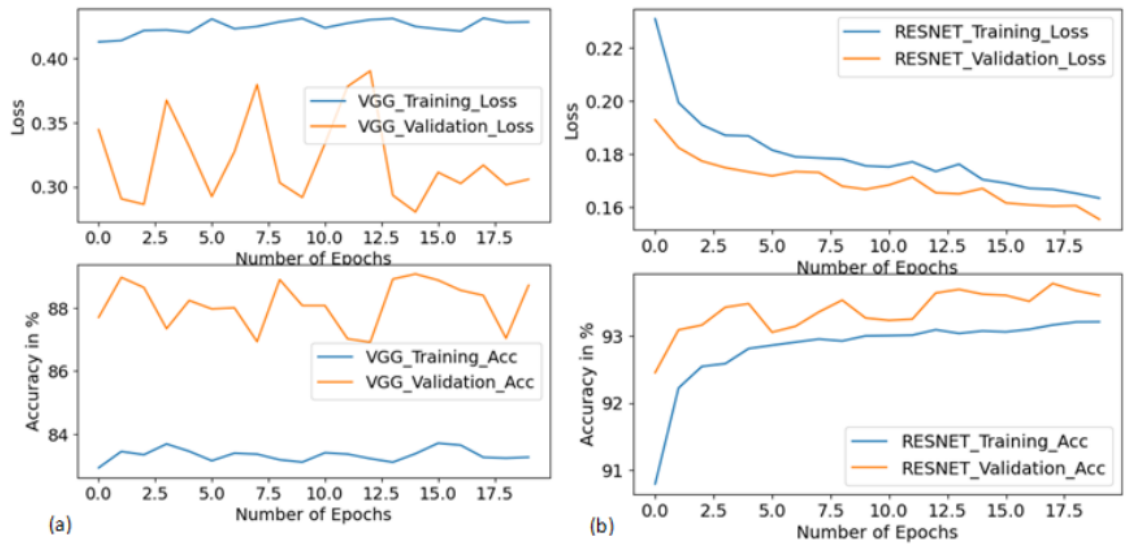

Figure 6: (a)Training and validation loss and accuracy plot for VGG model (b) Training and validation accuracy plot and accuracy plot for ResNet model.

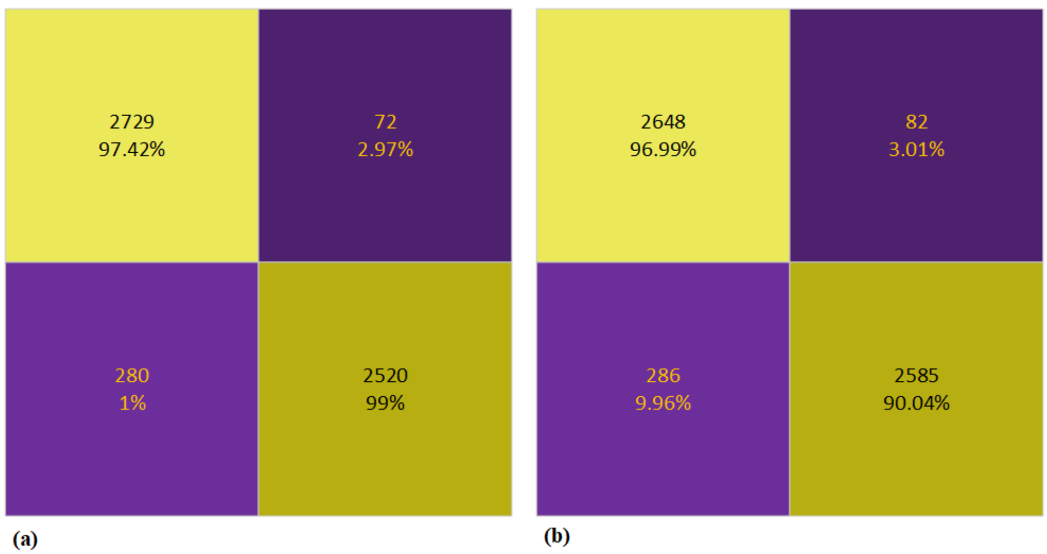

Figure 7: (a) Confusion matrix for VGG Model (b) Confusion matrix for ResNet model.

Classification Report :

	precision	recall	f1-score	support
0	0.83	0.97	0.90	2853
1	0.97	0.80	0.87	2748
accuracy			0.89	5601
macro avg	0.90	0.89	0.89	5601
weighted avg	0.90	0.89	0.89	5601

(a)

Classification Report :

	precision	recall	f1-score	support
0	0.90	0.97	0.94	2730
1	0.97	0.90	0.93	2871
accuracy			0.93	5601
macro avg	0.94	0.94	0.93	5601
weighted avg	0.94	0.93	0.93	5601

(b)

Figure 8: (a) Classification report for VGG model (b) Classification report for ResNet model.

learn, NumPy, pandas and matplotlib. The GPU package was employed since the study involves thousands of images and need to perform a large number of computations, from feature extraction through loss reduction to the bare minimum by guaranteeing the convergence of the model's weight values. However, we make use of the Jupiter notebook of Google online training platform known as Colab to train our model. Figures 5a and 5b display the training and validation loss and accuracy for the Visual Geometry Group

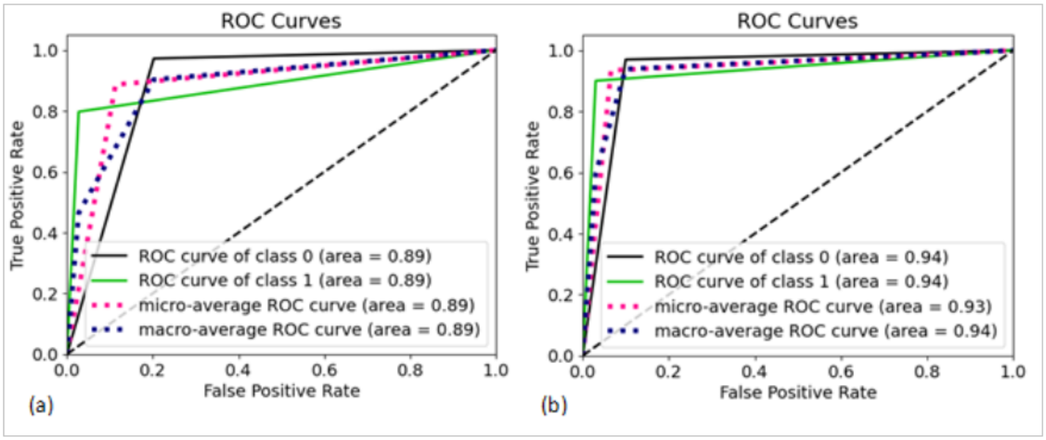

Figure 9: (a) Receiver operating characteristics curve for VGG Model (b) Receiver operating characteristics curve for ResNet model.

Table 4: Comparison of model results.

S/N	Paper	Model	Best performing model	Accuracy	ROC
1	[9]	Logistic regression, Support vector machine, Naïve Bayes, K-nearest neighbor, Decision tree, Random forest and Neural Network.	Support vector machine	91.07%	92.5%
2	[14]	K-nearest neighbor, Naïve Bayes, J48-Decisison tree, Sequential minimal Optimization,	J48 - Decision tree	91.20%	88.40%
3	[22]	Radial basis function network, Sequential minimal optimization, Support vector machine, Self-organizing map, Multilayer-perceptron, Hierarchical learning vector quantization.	Multilayer perceptron	90.01%	-
4	[24]	Logistic regression, random forest, multilayer perceptron, and decision tree.	Random forest	86.80%	92.7%
5	Proposed method	VGG and ResNet	ResNet	93.00%	94.00%

(VGG) and the Residual Network (ResNet) model. Figure 6a shows the accuracy rising as the number of iterations increases and loss decreasing as the number of iterations increases but in zig-zag pattern while Figure 6b shows a smooth reduction in loss value and increment in accuracy value as number of iterations increases. This suggests an important training procedure.

Testing our algorithm to identify potential clients who would sign up for a long-term deposit in the banking industry yielded the confusion matrix. From Figure 7a it can be seen that a total of 2775 samples fall into the true positive segment, 78 samples fall into the false negative segment, 556 samples fall into the false positive segment, and 2192 samples fall into the true negative segment of the confusion matrix for the VGG model while Figure 7b shows a total of 2648 samples fall into the true positive segment, 82 samples fall into the false negative segment, 286 samples fall into the false positive segment, and 2585 samples fall into the true negative segment of the confusion matrix for the ResNet model. The classification report explains the four key performance metrics of the machine learning classification models: accuracy, recall, precision, and F1 score. Figures 8a and 8b show the classification report of our trained model. Also, an efficient method for assessing the performance of the created model is to plot the true positive rate (y coordinate) against the false positive rate (x coordinate) the higher this rate is the better the model performance as it can be seen in Figure 9a and 9b, the area under the Receiver Operating Characteristics curve is 89% for the VGG and 94% for ResNet.

4.1. Performance comparison of our developed model with other research studies

Table 4 compares our developed model with existing models from the previous literature. Four different studies were compared with our work and it could be seen that our methodology achieve a better performance which shows the effect of adequate data processing and the capacity of deep neural network towards the prediction of long-term deposit subscribers in the banking industry.

5. Conclusion

In this study, we developed a deep learning model using ResNet architectural design and transfer learning on a Portuguese banking institution that initially contained 40,811 training data, with 36,202 belonging to label 0 and 4639 belonging to label 1. Clearly, this shows a significant level of bias between the two labels. However, a SMOTE method of class balancing was applied to make both labels have the same number of samples. Moving forward, this dataset in comma-separated value (CSV) was converted to images to allow us to conveniently apply a deep neural network to extract the notable features in the images. Thus, with weight transfer from the residual network trained on ImageNet, our fully connected layer was built and trained with the image files. We, therefore, evaluate our proposed model using notable evaluation metrics like accuracy, receiver operating cure, precision, F1 score and recall. Our Residual Network (ResNet) achieves the best result with an accuracy of 93.00%, 97.00% precision, 90.00% recall, 93.00%, F1 score, and 94.00% ROC with a test sample of size 5601 for model evaluation. I would recommend the implementation of this study by utilizing other deep convolutional neural network types with fewer training parameters that maintain the desired result in the future work as this could counterbalance the tradeoff between performance and time complexity that could be noticed in our work. A fresh dataset with sufficient samples for clients who would sign up for long-term deposits should also be sought out for since this can reduce the bias brought about by applying the SMOTE data balancing technique.

References

[1] S. Mushtaq & D. A. Siddiqui, "Effect of interest rate on bank deposits: Evidences from Islamic and non-Islamic economies", Futur. Bus. J. **3** (2017) 1. http://dx.doi.org/10.1016/j.fbj.2017.01.002.

[2] I. N. Yakubu & A. H. Abokor, "Factors determining bank deposit growth in Turkey: an empirical analysis", Rajagiri Manag. J. **14** (2020) 121. http://dx.doi.org/10.1108/RAMJ-05-2020-0017.

[3] Y. A. Ünvan & I. N. Yakubu, "Do bank-specific factors drive bank deposits in Ghana?", J. Comput. Appl. Math. **376** (2020) 112827. http://dx.doi.org/10.1016/j.cam.2020.112827.

[4] T. Jokipii & P. Monnin, "The impact of banking sector stability on the real economy", J. Int. Money Financ. **32** (2013) 1. https://doi.org/10.1016/j.jimonfin.2012.02.008.

[5] O. Paul & O. Omosefe, "The impact of interest rate on bank deposit: evidence from the Nigerian banking sector", Mediterranean Journal of Social Sciences **5** (2013) 232. http://dx.doi.org/10.5901/mjss.2014.v5n16p232.

[6] S. Raghunandan, G. Lavina & S. Jose, "Electronic customer relationship management an effective tool in the banking sector", Asian Journal of Management **9** (2018) 913. https://doi.org/10.5958/2321-5763.2018.00144.0.

[7] L. S. Goldberg, "Understanding banking sector globalization", IMF Staff Pap. **56** (2009) 171. https://econpapers.repec.org/RePEc:pal:imfstp:v:56:y:2009:i:1:p:171-197.

[8] B. Premkumar, D. P. Nandru & D. C. Madhavaiah, "Predicting the success of bank telemarketing for selling long-term deposits: an application of machine learning algorithms", St. Theresa J. Humanit. Soc. Sci. **7** (2021) 91. https://www.researchgate.net/publication/352755139_Predicting_the_Success_of_Bank_Telemarketing_for_Selling_Long-term_Deposits_An_Application_of_Machine_Learning_Algorithms?enrichId=rgreq-fb47eb7ab7b8daaf3d14c037b2f50066-XXX&enrichSource=Y292ZXJQYWdlOzM1Mjc1NTEzOTtBUzoxMDM4NzcwMTE5OTA1MjgwQDE2MjQ2NzM1MDQ3MzA%3D&el=1_x_2&_esc=publicationCoverPdf.

[9] A. Ilham, L. Khikmah, Indra, Ulumuddin & I. B. A. I. Iswara, "Long-term deposits prediction: A comparative framework of classification model for predict the success of bank telemarketing", J. Phys. Conf. Ser. **1175** (2019) 102035. https://iopscience.iop.org/article/10.1088/1742-6596/1175/1/012035.

[10] B. S. Neysiani, N. Soltani, & S. Ghezelbash, *A framework for improving find best marketing targets using a hybrid genetic algorithm and neural networks*, Conf. Proc. 2015 2nd Int. Conf. Knowledge-Based Eng. Innov. KBEI 2015, 2016, pp. 733-738. http://dx.doi.org/10.1109/KBEI.2015.7436136.

[11] S. Athey, "The impact of machine learning on economics" in *Economics of Artificial Intelligence: An agenda*, Mara Lederman (Ed.), University of Chicago press, Chicago, USA, 2019, pp. 507-547. https://www.nber.org/books-and-chapters/economics-artificial-intelligence-agenda.

[12] K. H. Kim, C. S. Lee, S. M. Jo, & S. B. Cho, *Predicting the success of bank telemarketing using deep convolutional neural network*, Proc. 2015 7th Int. Conf. Soft Comput. Pattern Recognition, SoCPaR 2015, Fukuoka, Japan, 2016, pp. 314–317. https://doi.org/10.1109/SOCPAR.2015.7492828.

[13] M. Grosicki, "Application of artificial neural networks to prediction of success of bank telemarketing campaign", Proceeding of Transcom European conference of young researcher and scientist, Slovak republic, 2015, pp. 25-314. https://doi.org/10.1109/SOCPAR.2015.7492828.

[14] M. Rashid-Farooqi & N. Iqbal, "Performance evaluation for competency of bank telemarketing prediction using data mining techniques", Int. J. Recent Technol. Eng. **8** (2019) 5666. http://dx.doi.org/10.35940/ijrte.A1269.078219.

[15] M. J. A. Patwary, S. Akter, M. S. Bin Alam & A. N. M. Rezaul Karim, "Bank deposit prediction using ensemble learning", Artif. Intell. Evol. **2** (2021) 42. https://doi.org/10.37256/aie.222021880.

[16] M. A. Muslim, Y. Dasril, A. Alamsyah & T. Mustaqim, "Bank predictions for prospective long-term deposit investors using machine learning LightGBM and SMOTE", J. Phys. Conf. Ser. **1918** (2021) 042143. http://dx.doi.org/10.1088/1742-6596/1918/4/042143.

[17] M. A. T. Rony, M. M. Hassan, E. Ahmed, A. Karim, S. Azam & D. S. A. A. Reza, *Identifying long-term deposit customers: a machine learning approach*, 2nd Int. Informatics Softw. Eng. Conf. IISEC 2021, Ankara, Turkey, 2021, pp. 1-6. http://dx.doi.org/10.1109/IISEC54230.2021.9672452.

[18] C. Vajiramedhin & A. Suebsing, "Feature selection with data balancing for prediction of bank telemarketing", Appl. Math. Sci. **8** (2014) 113. https://api.semanticscholar.org/CorpusID:18039108.

[19] S. Moro, P. Cortez & P. Rita, "Using customer lifetime value and neural networks to improve the prediction of bank deposit subscription in telemarketing campaigns", Neural Comput. & Applic. **26** (2015) 351. http://dx.doi.org/10.1007/s00521-014-1703-0.

[20] Y. Kawasaki & M. Ueki, "Sparse predictive modeling for bank telemarketing success using smooth-threshold estimating equations", *J. Japanese Soc. Comput. Stat.* **28** (2015) 53. http://dx.doi.org/10.5183/jjscs.1502003_217.

[21] M. Amini, J. Rezaeenour & E. Hadavandi, "A cluster-based data balancing ensemble classifier for response modeling in bank direct marketing", Int. J. Comput. Intell. Appl. **14** (2015) 1. https://doi.org/10.1142/S1469026815500224.

[22] A. Panigrahi & M. C. Patnaik, "Customer deposit prediction using neural network techniques", Int. J. Appl. Eng. Res. **15** (2020) 253. https://www.ripublication.com/ijaer20/ijaerv15n3_9.pdf.

[23] P. Moro, S. Rita & P. Cortez, "Bank marketing," 2012. [Online]. Available: http://archive.ics.uci.edu/dataset/222/bank+marketing.

[24] J. Asare-Frempong & M. Jayabalan, "Predicting customer response to bank direct telemarketing campaign," *Int. Conf. Eng. Technol. Technopreneurship, ICE2T 2017*, 2017, pp. 1–4. http://dx.doi.org/10.1109/ICE2T.2017.8215961.