NIPES

Journal of Science and Technology Research

Journal homepage: www.nipesjournals.org.ng

Science and Technology Research

Application of Iterative Machine Learning in Predicting Fake Documents in Job Applications

Aru, Okereke Eze, Adimora, Kyrian Chinemeze, and Umunnakwe ,Franklin Ugochukwu
Department of Computer Engineering, Michael Okpara University of Agriculture Umudike, PMB 7267, Umuahia Abia State, Nigeria
Email: okezearu@gmail.com, kyrimanjero@gmail.com, and franklin2noble@gmail.com

Article Info

***Keywords*:**
Fake documents, Iterative machine learning, Convergence threshold, Job-recruitment, and Datasets.

Received 11 October 2022
Revised 24 October y 2022
Accepted 6 November 2022
Available online 13 March 2023

https://doi.org/10.5281/zenodo.7728808

Abstract

The increase in the rate of industry liquidation and unproductiveness resulting from false persons' recruitment through fake document assessment during the job application is a great concern. Most companies have gone bankrupt whereas some have shut down untimely due to unqualified staff recruited with fake documents. This paper proposed a machine learning (ML) model that predicts fake documents in job applications. The machine learning model was designed with iterative learning algorithms based on a comparative analysis of applicants' data and online information from the alma mater discussed in this article. An iterative similarity convergence threshold of 0.5 was set to ensure the prediction accuracy of the model. Compared documents that are less than the defined threshold are considered fake and not original. The iteration results of several trained datasets presented in Tables 4 and 5, filtered fake documents. These datasets were trained, tested, and validated using the MATLAB application. Significant results achieved, revealed the efficient performance of the model in filtering fake documents in job applications.

1.0 Introduction

The algorithms in machine learning are classified as either iterative or non-iterative algorithms. The machine learning algorithm accurately predicts future data based on the supplied data and interaction with past experience. Besides, reinforcement and semi-supervised learning algorithms supervised and unsupervised learning is widely used in today's research that is categorized as either iterative or non-iterative. In this paper, an iterative type of machine learning is adopted in the design of the proposed machine learning analytical model to predict fake documents in job applications. An iterative machine learning algorithm is chosen because of the document, authentication, verification, and validation are repeated processes. The goal of iterative algorithm is to find an optimum solution from the training dataset. A stopping criterion known as an iterative convergence threshold of 0.5 is taken for this paper's research as future iterations may result in minor changes to the data. This is to ensure accuracy in the prediction model. The proposed model iteratively increases the accuracy of the prediction until the similarity convergence threshold of 0.5 is reached. In order to meet this convergence criterion, the new model learns from the applicant's supplied data and is repeated iteratively. In this section, several relevant works of literature were reviewed with the gaps identified and improved by the proposed system.[1] conducted research that employs a neural network for person-job fit through the candidates' profiles and related recruitment history. Its shortcoming is that the research did not specify how the issue of fake documents can be detected in job applications.

Chowdhury *et al.*[2] carried out a systematic review of multi-disciplinary literature required to develop AI capability in human resource management (HRM). It is a review work and did not provide a solution to the issue of the fake document on job applications. [3] Researched on AI-enabled job characteristics and possible substitution crises resulting from state-of-art technology. The work did not consider the issue of the fake documents and how they can be detected in the job recruitment process. The critical components of online recruitment include tracking the status of candidates, employer's website, job portals, online testing, and social networking [4]. The advantages of e-recruitment include effectiveness, high value, easiness, and efficiency [5]. The job recruitment process is a very tedious task for the human resource management unit of every organization. This is because the time meant for selecting the best candidate applicants for the advertised job most times is used in critical evaluation of applicants' documents to get rid of fake documents. In most cases after these rigorous or critical evaluations on the applicants' submitted documents, organizations are still experiencing the great number of false people being recruited. This research gap is what led to the writing of this research paper. According to Nitu *et al.* [6], who conducted a research on the recruitment process that employed data mining for initial recruitment process of applicants' resume using back propagation neural network. The research focused only on the applicants' resume information for fake document assessment. Muhammad *et al.* [7] conducted research with the application of deep learning using Convolutional neural network for determination of ink age and forgery documents. According to the research work, ink mismatch detection is a key step in document forgery detection. However, the work only focused on a hyper spectral document image for forgery detection and may not be applicable in detection fake digital document. [8] Research paper that employed application filter for company's CVs selection. The research study made use of independent classifiers and Naïve Bayes methodology in its design. Chloe *et al*. [9] research work focused on fraud detection that focused on detecting documents containing at least one forgery in a flow of documents and spotting and localizing these forgeries within documents.
Omar *et al.* [10] research work on a Machine Learning (ML) based classification model to classify Bengali text into non-suspicious and suspicious categories based on its original contents.
Heejung *et al.* [11] research on data driven automatic fake news detection methods using Bidirectional Encoder Representations. Bandar [12] research model that detects fraud in the online recruitment environments. The research employed ensemble approach based on Random Forest Classifier to detect Online Recruitment Fraud (ORF). Ahmed *et al.* [13] carried out research for fake news detection among politicians that employed n-gram analysis and machine learning techniques. The research also made use of a linear support vector machine as a classifier. The work focused only on fake news and may not be applicable for detecting fake documents in a job application.
[14] Conducted thesis research that aimed at improving the human resource management (HRM) recruitment process using artificial intelligence. The work is restricted to improving the recruitment process without consideration of fake document detection in a job application. [15] Carried out research to distinguish documents produced by laser printers, inkjet printers and electronic copiers, based on features extracted from the characters in the documents. The research emphasis is mainly on printed hard copy document without consideration of digital document.
According to Bogle [16], research conducted on the use of mobile intelligent agent, and cryptography instead of human agents to perform the job search using fuzzy preference to validate applicant identity and accuracy. The work focused on the applicant's identity verification and not to detect fake document in job recruitment.
In all the reviewed literature, none were able to address the issue of fake documents in job recruitment using iterative machine learning algorithms.

2.0 Materials and Methods

The work has been implemented in the working platform of MATLAB.

2.1 These system specifications are used.

I. The Visual Studio IDE for project code
II. The MATLAB IDE for result testing and validation
III. OS: Windows 8
IV RAM: 8GB
V Processor: Intel Core i5

2.2 Data Acquisition and Processing

2.2.1 Validation and Verification Algorithm Approach

The algorithm 1 defines the candidate's parameters by comparing the applicant's submitted document, online information and update from the alma mater.

Algorithm 1: Document Comparison Test (Wide Range Approach)

Input to the algorithm:

Applicant's Details:

Name:
Age:
Qualification:
Year of Graduation:
Institution:
Profession:
Work Experience:
State of Origin:
State of Residence:
Date of Birth:
Sex:
Degree Status:
Marital Status:

Step 1: Read the applicant's data
Step 2: Compare Submitted Data and Online Information
What is the Match?
If less than 50 %
Consider the submitted document fake
Else
Accept the document as not fake
Step 3: Print candidates information
Step 4: End.

The algorithm1 analyzes and compares the information on the applicant's submitted document with update information from the alma mater. Similarity convergence less than 50 percent is a confirmation that such a document is fake and not original. It is a wide range approach because many parameters are involved.

Algorithm 2: Document Comparison Test (Narrow Approach)

The approach here, is to measure the similarity of document belonging to an applicant. The method is divided into four major groups, Status degree-based (Sd_b), Institutional Certificate-based (Ic_b), Knowledge-based (K_b), Professional-based (P_b), and Graduation year-based (Gy_b) as shown in Figure 1. This approach is called narrow because few parameters are considered in the algorithm and will be detailed in the next subsections.

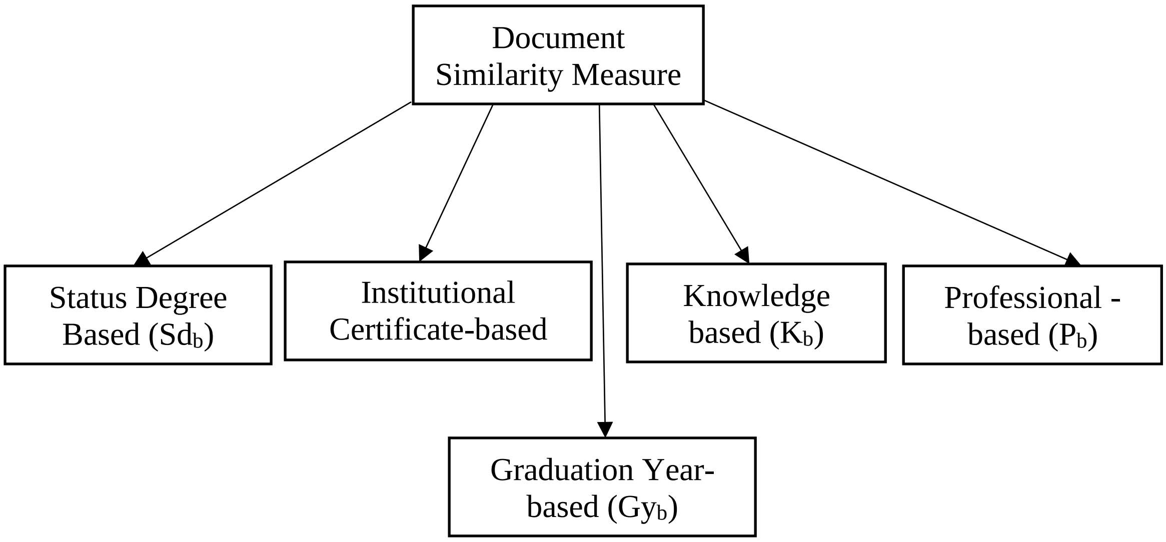

Figure 1: Five major groups of document similarity convergence acceptance

2.2.1. Status degree based (sd_b)

The status degree based similarity is the measurement approach that validates the status of degree claimed by the applicant in job application. This approach extracts information from the applicant's alma mater and compares it with the information supplied by the applicants. Then compared, if the similarity convergence is less than 50 percent (0.5) the document is considered fake. These are the status degree parameters used in this algorithm.

I. First class (1^{st} class degree honors)
II. Distinctions honors
III. Second class degree honors upper division
IV. Second class degree honors lower division
V. Third class degree honors
VI. Upper credit
VII. Lower credit
VIII. Pass

The two main types of status degree similarity functions are character-based similarity functions (Found worthy in character) and learning based similarity functions (Found worthy in learning) always printed on the degree certificate. The character-based similarity function is also called applicant's acceptance attribute behavioral measurement. It takes eight stings of characters as itemized and then calculates the applicant's acceptance attribute (including insertion, deletion, and substitution) between them. Character-based quantifies character similarity between eight strings to quantify the similarity, for instance evaluates the applicant's minimum status degree for the job application acceptance and compare records from the claimed institution and what is presented. If the similarity convergence function is less than 0.5 the document is fake and rejected.

2.2.2. Institutional Certificate-based (Ic_b)

Institutional-based similarity convergence uses a semantic approach. This similarity approach determines the relationship between the information on the applicant's submitted documents and update from online and alma mater. The aim is to validate if the applicants was really the product of the claimed institution. The approach allows the human resource management or the recruitment center to verify if the applicant was a graduate of the claimed institution. If the candidate's attribute is not found on the institutional record the document is termed fake and not original.

2.2.3. Knowledge based (K_b)

Knowledge based similarity convergence approach is aimed at confirming the claimed acquired skills of the applicant and job experiences as stated in the submitted documents for job consideration. The approach enforces the applicant to demonstrate all the claimed skills presented on paper for job consideration. The new system operates to evaluate the skill on paper and in practice using intelligent logical comparison schema that considers the document fake when it is 50 % less in similarity. The schema of knowledge representation generally includes the rules of conclusions, logical propositions, and network semantics such as taxonomy and ontology.

2.2.4. Profession based (P_b)

This approach for similarity check assesses the applicants on the basis of claimed profession to ascertain if the candidate's profile portrays such identity using calculated values of Newton R. iteration integrated into the system. Third iteration with the value less than 50 percent confirms the document to be fake.

2.2.5. Graduate Year based (Gy_b)

The graduate year based approach of similarity convergence deals with applicant's year of graduation. It is the approach that determines the candidate's academic record year of graduation and compares it with the document presented. Notable variation in the comparative analysis of the graduation year proves the document fake. The experimental evaluation result of each group is shown in Table 1 in section 3.

Algorithm 3: Applicant's Alma mater Document Verification

Step 1: Employer Initiate communication with the applicant's Alma mater for an applicant
Step 2: Alma mater checks if this name is in your record?
Step 3: If yes?
Step 4: Help confirm applicant's data signature and education details in format of
    Full Name:
    Discipline:
    Year of graduate:
    Class of degree obtained or awarded
Step 5: If applicant's data signature and education details correspond to what is in the employer's disposition, then
Step 6: Accept the document as original
    Else
    Document is taken as fake and won't be processed further
Step 7: Print result
Step 8: End Communication.

Algorithm 4: Machine learning algorithms model design and feature selection

Step 1: Begin
Step 2: Input original labeled data as input to auto analyze which is labeled sample matrix i.e.
    $D = \{i_1, i_2, \dots, i_n\}^T \in A^{nxd}$
Step 3: Analysis function that performs the encoding of document feature i.e.
    $f(D) = \sigma_1(DW^{(1)})$;
Step 4: Generated features set after serious of transforms;
Step 5: Analysis function that performs decoding of document feature i.e.
    $\hat{D} = g(f(i)) = \sigma_2(f(D)W^{(2)})$;
Step 6: The output of decoding function is equal to original document feature set.
    Else
    Document is fake
Step 7: End

Here, D is the applicants' documents, A is the analyzer, i_1, i_2, to i_n is the applicants' identity, T is the document transformation, n is the number variable required to be verified, f(D) is the function of the analysis, W is the interpretation function, f(i) is the applicants' identity function.
The proposed system framework undergoes the seven steps as stated in algorithm 4 for verifying the candidate's document. The framework is presented in Figure 2.

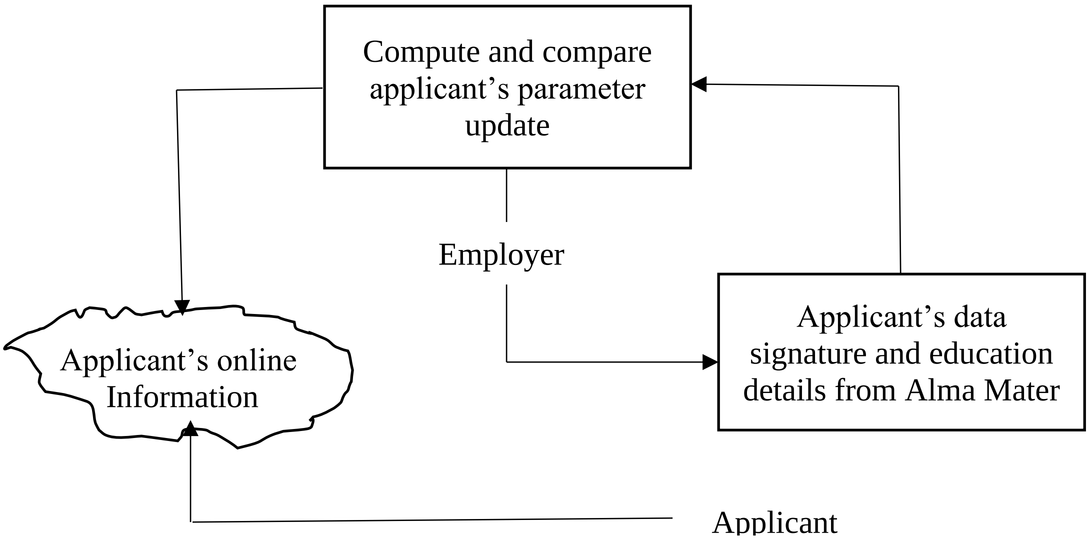

Figure 2: Document validation model

2.2.6 Data Set Description

The total datasets sample collection experiment from three algorithms were summarized in Table1. Here, we selected 1200 samples from algorithms 1 and 3 as training sets and 400 samples from algorithm 2 as a testing dataset. Approximately 5% of training data as summarized in Table 2 was used as validation to predict fake documents in job applications. Table 3 compares the proposed system to the existing system.

Table 1: Datasets sample collection

Parameters	**Algorithm 1**	**Algorithm 2**	**Algorithm 2**
Applicants' data	300	200	400
Institution data	300	200	400
Total	600	400	800
Grand Total	= 600 + 400 + 800 = 1800 samples		

Table 2: Datasets training, validation, and testing

Condition	**Training**	**Validation**	**Testing**
Applicants' Data	85%	8%	20%
Institution Data	80%	11%	18%

Table 3: Proposed system and existing system comparison

	Training Parameter	Testing Accuracy
Proposed system	1800	97%
Existing system	3000	82%

The number of the trained parameters processed in the proposed system is 1800 while the existing system has 3000 parameters processed. The proposed system testing accuracy is preferred to that of the existing system 97% against 82%.

2.3 Document Divergence and Convergence Test

In this section, we have two series of positive terms and the terms of one of the series is always larger than the terms of the other series. Then if the larger series is convergent the smaller series must also be convergent. Likewise, if the smaller series is divergent then the larger series must also be divergent. Note as well that in order to apply this test we need both series to start at the same place. These series are used to compare applicant's documents to validates its originality or determine whether it is fake. In this research, a document is said to be divergent when the comparative test is less than 50 percent ($Ct < 50\%$) and convergent when the comparison test is 50 percent or greater ($Ct \geq 50\%$).
Having two series (documents a_n and b_n) such that:

$$\sum a_n \; and \; \sum b_n \; with \; a_n, b_n \geq 0 \; for \; all \; n \; and \; a_n \leq b_n \; for \; all \; n. \; Then.$$

1. $if \; \sum b_n \; is \; convergent \; then \; so \; is \; \sum a_n$
2. $if \; \sum a_n \; is \; divergent \; then \; so \; is \; \sum b_n$

Note as long as the requirement that $a_n, b_n \geq 0 \; and \; a_n \leq b_n$ really only need to be true for the comparative test to work. Then the conditions of the test are only true for $n \geq N + 1 \; and \; for$ $k \leq n \leq N$ at least one of the conditions is not true. If we then look at $\sum a_n$ (the same thing could be done for$\sum b_n$) we get,

$$\sum_{n=k}^{\infty} a_n = \sum_{n=k}^{N} a_n + \sum_{n=N+1}^{\infty} a_n \quad (1)$$

The first series is nothing more than a finite sum (no matter how large NN is) of finite terms and will be finite. The original series will be convergent/divergent only if the second infinite series on the right is convergent/divergent and the test can be done on the second series as it satisfies the conditions of the test. Mathematically document convergent and divergent status are stated in equation (2), (3), (4), and (5) respectively.

$$S_n = \sum_{i=1}^{n} a_i \quad (2)$$

$$t_n = \sum_{i=1}^{n} b_n \quad (3)$$

$$S_n \leq S_n + a_n + 1 \quad (4a)$$

$$\sum_{i=1}^{n} a_i + a_n + 1 = \sum_{i=1}^{n+1} a_i = S_n + 1 \quad (4b)$$

$$\Rightarrow S_n \leq S_n + 1 \quad (4c)$$

$$t_n \leq t_n + b_n + 1 \quad (5a)$$

$$\sum_{i=1}^{n} b_i + b_n + 1 = \sum_{i=1}^{n+1} b_i = t_n + 1 \quad (5b)$$

$$t_n \leq t_n + 1 \quad (5c)$$

Figure 3 is the architecture of the machine learning model that predicts fake documents in Job applications.

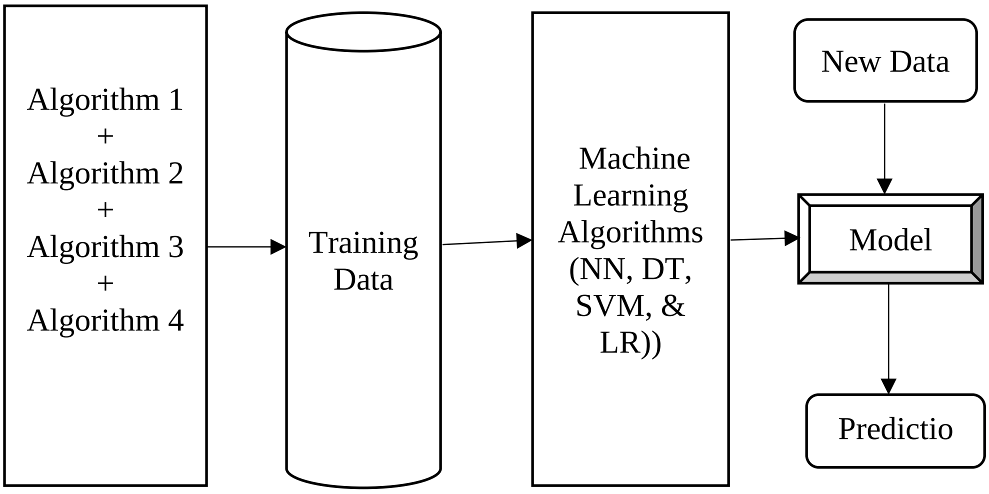

Figure 3: Architecture of the new model

Algorithm1, contains the wide range of applicant's document comparison test data, Algorithm 2, contains the narrow range of applicant's document comparison test data, Algorithm 3, and contains the applicant's Alma mater document verification data. The consortium datasets generated by these algorithms are processed by Algorithm 4 and made available for training which in turn constructs the machine learning analytical model to predict fake documents in job applications.

2.3.1. Similarity Convergence Algorithm

The mathematical model for the similarity convergence algorithm is shown below. In this research we use a convergence threshold of $\varepsilon = 0.5$ for acceptance of document similarity.

When a sequence of elements gets closer to a single value it is said to converge. In this research paper, we adopt iterative convergence principle for analyzing our document similarity using a converged numerical value at acceptable threshold of 0.5. We also made use of sequence of real numbers $x_1, x_2, \dots, x_n$ representing a good number of documents for verification. In our design, x_n converges to a given number N if in every positive error, there is an x_m such that every element x_n that comes after $x_m$differs from N by less than that error.

Clearly if a root of $\varphi f(n_1, n_2)$ exists, this root is a fixed point of iteration, that is, a point for which $f_{k+1}(n_1, n_1) = f_k(n_1, n_1)$. That is with$f_0(n_1, n_1) = 0$.

$$f_{k+1}(n_1, n_1) = f_k(n_1, n_1) + \beta\varphi\big(f_k(n_1, n_1)\big) = \beta g(n_1, n_1) + \big(\delta(n_1, n_1) - \beta d(n_1, n_1)\big) \qquad (6)$$

Where $f_k(n_1, n_1)$ denotes the acceptance degree of acceptance at the k-th iteration step, $\delta(n_1, n_1)$ the discrete delta function, and β the relaxation parameter which controls the convergence, as well as the rate of convergence of the iteration. Iteration (6) is the basis of a large number of iterative similarity algorithms, and therefore analyzed as follows:

$$F_{K+1}(u, v) = \beta G(u, v) + (1 - \beta DS(u, v))F_K(u, v) \qquad (7)$$

Where u is the information on the submitted document, v is the information from the alma mater and online, and $F_K(u, v)$, *G (u, v)*, and *D (u, v)* represent, $f_k(n_1, n_1)$, *g (n*1,*n*2*)*, and *d(n*1,*n*2*)* respectively. Next we express $F_K(u, v)$ in terms of $F_0(u, v)$.

$$F_1(u, v) = \beta G(u, v) - n\emptyset \qquad (8)$$

$$F_2(u,v) = \beta G(u,v) - n\emptyset + \left(1-\beta D(u,v)\right)F_K(u,v) + Hsin\,\theta, \quad (9a)$$

$$= \sum_x^k \left((1-\beta D(u,v) + Hsin\,\theta)\right)^x \beta G(u,v),$$

(9b)$F_K(u,v) = \sum_{x=0}^{k-1}\left((1-\beta D(u,v) + Hsin\,\theta)\right)^x \beta G(u,v),$ (9c)

$$= H_K(u,v)G(u,v), \quad (9d)$$

We, therefore, see that the acceptance similarity convergence at the K-th iteration step is given by

$$H_K(u,v) = \beta \sum_{x=0}^{k-1}\left((1-\beta D(u,v) + Hsin\,\theta)\right)^x \quad (10)$$

The sufficient condition for similarity convergence is satisfied after several iterations, the value converges to the original signal as denoted by C in Figure 4. This is also the inverse solution obtained directly from the iterative convergence of equation (6).

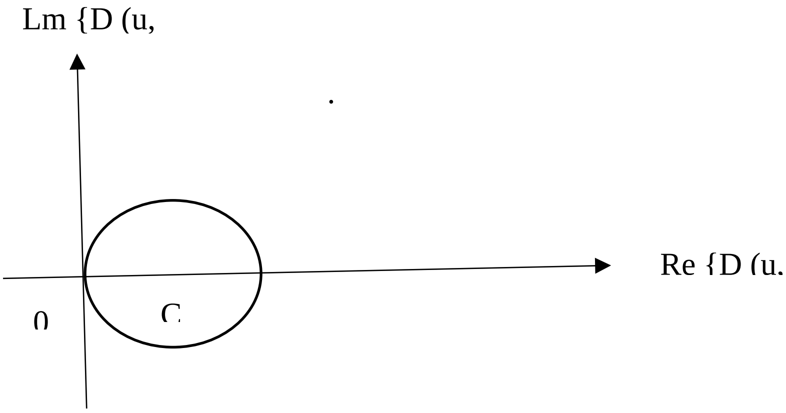

Figure 4: Sufficient Condition for Convergence

$$Where\ C = \left({}^{1}/_{\beta},0\right) \quad (11)$$

Figure 5 is an overview of the iterative machine algorithm and its convergence criteria set for iterating the datasets and outputting fake documents in job applications.

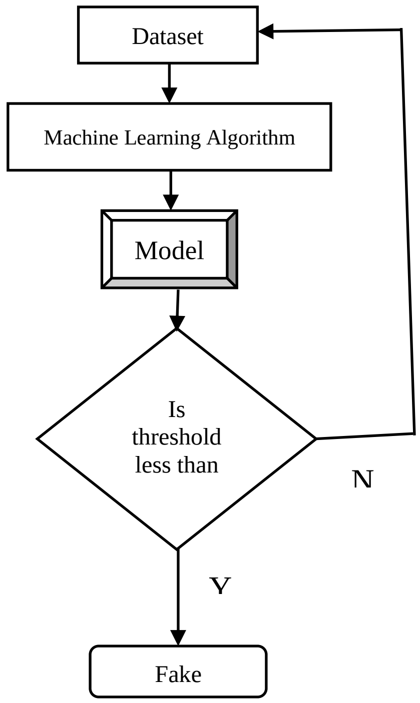

Figure 5: Operational Overview of the new model

3.0 Results and Discussion

This section provides a description of implementation results and performance evaluation of the proposed machine learning model and a comparison section to predict fake documents in job applications. Table 4 depicts the machine learning algorithms dataset training and testing.

Table 4: Machine learning dataset training and testing

Iteration Parameter	Dataset	Neural Network		Decision Tree		Support Vector Machine		Logistic Regression	
		Test	Train	Test	Train	Test	Train	Test	Train
FD	41.2	27.8	22.2	37.5	29.9	38.6	35.5	28.2	23.7
OD	38.9	25.1	20.5	36.2	28.4	36.3	32.6	24.9	20.1
AD			5.1		7.7		5.5		4.7
FD	35.0	20.2	18.7	27.2	23.4	32.5	30.4	19.9	17.3
OD	39.5	23.6	19.2	28.1	24.3	31.7	30.5	19.4	18.8
AD			5.9		3.8		1.7		1.6
FD	48.1	32.3	28.5	34.9	30.6	35.4	34.1	27.1	23.9
OD	42.7	33.1	27.2	36.7	28.1	36.2	33.0	24.2	22.5
AD			4.9		6.5		2.3		2.5
FD	31.2	21.0	15.2	28.1	24.2	30.0	29.7	18.9	14.2
OD	33.1	19.3	14.5	25.2	22.1	31.8	28.4	15.1	13.3
AD			5.3		3.5		1.9		3.3

Description of the variables in Table 4. FD is Fake Document, OD is Original Document, and AD is Average Deviation. All parameters are in percentage (%). During the first iteration, the average deviation in a neural network is 5.1, a decision tree is 5.5, the support vector machine is 5.5, and logistic regression is 4.7. In the second iteration, the average deviation in the neural network is 5.9, the decision tree is 3.8, the support vector machine is 1.7, and the logistic regression is 1.6. In the third iteration, the average deviation is 5.3, a decision tree is 3.5, the support vector machine is 1.9, and logistic regression is 3.3. Table 5 is the iterative convergence result that confirms fake and original documents of the applicants through the similarity acceptance computation of the proposed system.

Table 5: Documents similarity results

N/S	Iterative Algorithm	D1 (Pair 1)	D2 (Pair 2)	D3 (Pair 3)	Remarks
1.	Model	0.5334	0.8716	0.9253	D. Good
2.	Micro	0.1032	0.1509	0.1845	D. Fake
3.	Macro	0.3217	0.0345	0.4651	D. Fake
4.	Meta	0.2731	0.0256	0.3451	D. Fake
5.	Human	0.0014	0.0035	0.0026	D. Fake
6.	Support Vector Machine	0.6721	0.8921	0.5672	D. Good
7.	Gradient Descent	0.5320	0.9321	0.7145	D. Good
8.	K-fold	0.9451	0.6523	0.5498	D. Good
9.	Jacobi	-0.0234	0.0034	-0.0674	D. Fake
10.	Gauss-Seidel	0.0356	0.0873	0.4213	D. Fake
11.	Boltzmann Machine	0.7823	0.9134	0.6534	D. Good
12.	Naïve Bayes	0.6342	0.8253	0.7451	D. Good
13.	Logistic Regression	0.2118	0.0754	0.0063	D. Fake
14.	Newton R	0.0042	0.0011	0.0047	D. Fake

The values in Table 1 reflect good document when the machine learning iterative result of document similarity convergence is greater than 50 percent in all the three pairs. However, when the document (D1, D2, and D3)) similarity convergence is less than 50 percent in all the pairs, it implies the document detected is fake and not original. The results showed that the approaches and algorithms applied are efficient and acceptable in detecting applicant's fake document in job consideration.

3.1 Iterative Convergence Regression

We now test the documents similarities using iterative convergence regression on the fitted lines. The testing parameters are extracted from the series of algorithms defined in section 2 of this paper. Figure 6 depicts the results of fake and original documents based on the updated information assessment from applicant's employer and alma mater.

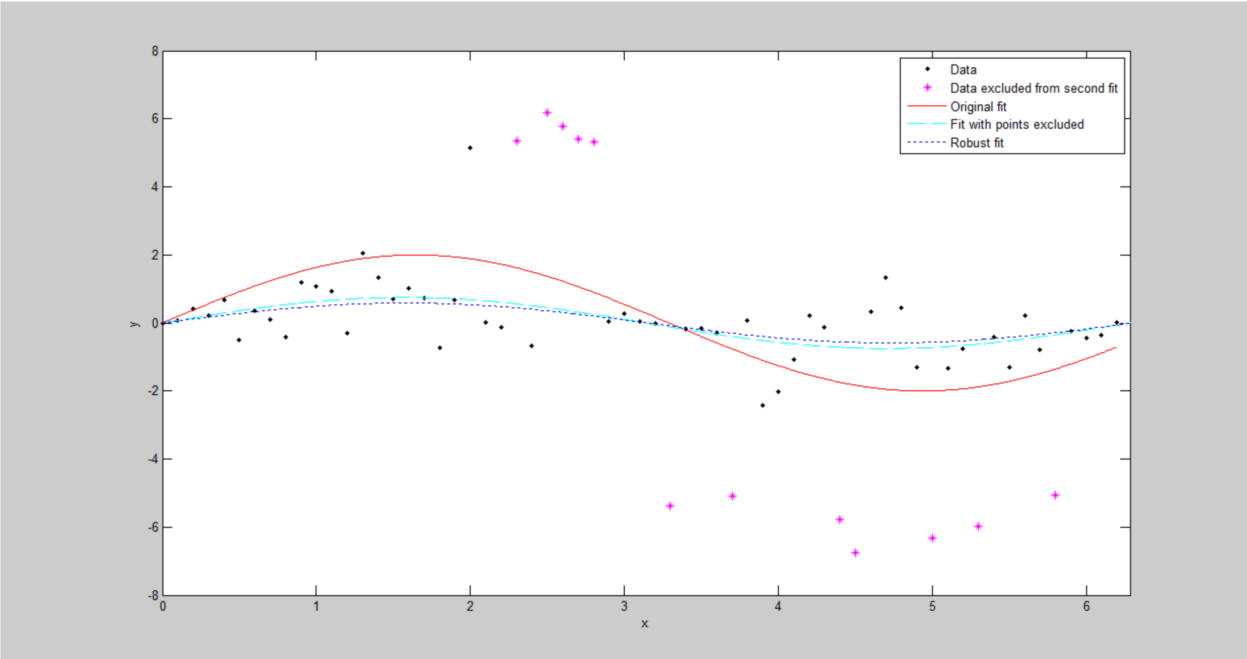

Figure 6: Document assessment result

The dots (data) on the original fit with the red color line represent the confirmed original documents of the applicants after convergence test assessment from the alma mater. Also the stars (data excluded from second fit) with pink color represent the confirmed fake documents of the applicants after convergence test assessment from the applicant's alma mater. The sky blue dash lines fit points denote fake applicants' documents. Therefore, the study results revealed that the machine learning model iterative approach is effective in detecting applicants' fake documents in job applications.

4.0. Conclusion and Suggestions

Fake documents have watered the recruitment process for job consideration. This has affected many organizations adversely because wrong persons are employed with fake documents and most times this brings about the dignity and integrity loss of a company. However, it is now imperative that a system is developed to detect fake applicants' documents in job applications. In this paper, we developed an iterative machine learning model specifically for detecting fake documents in a job application. Its test accuracy has shown that it outperforms the existing system. We suggest that the future direction should include a hybrid of data mining and deep learning in filtering fake documents in job recruitment.

References

[1] Wang Z., Wei W., Xu C., Xu J., Mao X. (2022). "Person-Job Fit Estimation from Candidate Profile and Related Recruitment History with Co-attention Neural Networks",ScienceDirect of Elsevier, doi: http://doi.org/10.1016/j.neac.com.2022.06.012

[2] Chowdhury S., Dey P., Joel-Edgar S., Bhattacharya S., Rodriguez-Espindola O., Abadie A., and Truong L. (2022). "Unlocking the value of Artificial Intelligence in Human resource management through AI Capability Framework", ScienceDirect of Elsevier. https://doi.org/10.1016/j.hrmr.2022.100899

[3] Verma S. and Singh V. (2022). "Impact of Artificial intelligence-enabled job characteristics and perceived substitution crisis on innovative work behavior of employees from high-tech firms", ScienceDirect of Elsevier. https://doi.org/10.1016/j.chb.2022.107215

[4] Hada, B. and Gairola, S. (2015) Opportunities and Challenges of E-Recruitment. *Journal ofManagement Engineering and Information Technology*, 2, pp. 1-4.

[5] Kaur, P. (2015) E-Recruitment: A Conceptual Study. *International Journal of Applied Research*, 10(3), pp. 78-82.

[6] Nitu M. Ashish B., Svit I., and Savits I. (2012). Applicability of back propagation neural network for recruitment data mining. *International journal of Engineering and technology (IJERT)*. 1(3), pp. 1-7.

[7] Muhammad J.K., Adeel Y., Asad A., and Khurram K. (2018). Deep Learning for Automated Forgery Detection Hyper Spectral Document Images. *Journal of Electronic Imaging. DOI: 10.1117/1JEI.27.5.05300*, Pp. 1-23.

[8] Anca A., Evelyn S., and Magnolia T., (2017). Machine Learning Based Application for Staff Recruitment. Acta Technica Naocensis Electronics and Telecommunications, Vol. 58(4), pp.1-7.

[9] Chloe A., Nicolas S., Antoine D., and Jean-marc O. (2018). Fraud Detection Contest Report. 24th *International Conference on Pattern Recognition (ICPR 2018),* Aug. 2018, Beijing, China. pp. 13 -18. Hal-02316399.

[10] Omar S. Mohammed M.H., A.S.M. Kayes, Raza N., and Iqbal H. Sarker (2020). Detecting Suspicious Texts Using Machine Learning Techniques. *Applied Sciences MDPI*, pp. 1- 23.

[11] Heejung J., Dongsuk O., Jang M.K., and Heuiseok Lim (2019). exBake: Automatic Fake News Detection Model based on Bidirectional Encoder Representations from Transformers (BERT). *Applied Sciences MDPI*, pp. 1-9.

[12] Bandar Alghamdi, and Fahad Alharby (2019). An Intelligent Model for Online RecruitmentFraud Detection. *Journal of Information Security*, http://www.scirp.org/journal/jis

[13] [Ahmed H., Traore I., and Saad S. (2017). Detection of online fake news using n-gram analysis and machine learning techniques. Springer International publishing, ISDDC 2017, LNCS 10618, pp. 127-138.

[14] Johansson J., Senja H., and McCauley B. (2019). The application of artificial intelligence (AI) in human resource management (Bachelor Thesis) Jonkoping University, pp. 1-61.

[15] Shang, S.,Memon, N., and Kong, X. (2014). Detecting documents forged by printing and copying. *EURASIP Journal on Advanced in Signal Processing*, pp.1-13.

[16] Bogle, S., and Sankaranarayanan S. (2012). Biometric application of intelligent agents in fakedocument detection of job applicants. *International Journal of Computer Science, Engineering and Information Technology (IJCSEIT)*, 2(1), pp. 75-94.