iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** III **Month of publication:** March 2026

DOI: https://doi.org/10.22214/ijraset.2026.77916

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

Detecting Fake Job Postings Using BI-LSTM and Attention Mechanism

Shilpa K
Department of Computer Science, Avinashilingam Institute of Home Science and Higher Education, Coimbatore

Abstract: ***With the advent of online medias, fakeness in job have been increased. The rapid growth of job sites that offering numerous jobs multiplied the amount of fakeness. Through fraudulent job postings many innocent people have huge financial losses and which in turn leads to identity theft also. Therefore we have to devise effective methodologies to avoid its impact. Various machine learning and Natural Language Processing techniques have been used in various domains such as fake news detection, e-mail spam detection etc. Using Recurrent Neural Networks and its variants can capture the patterns in the text. In this paper a new idea is presenting in which fake jobs notification in online medias can be detected with BI-LSTM and attention mechanism. Both methods can capture the dependencies between text more efficiently and helps to detect the fakeness in job sites more accurately.***

Keywords: fake jobs, dependencies between text, Identity theft, spam detection

I. INTRODUCTION

In today's technology era, the job notifications can be accessed through online media instantly. As job opportunities has been increased, the amount of fakeness in job notifications also multiplied. With the help of this paper, we can identify the fakeness in job sites and make people keeping safe from these traps.

In the proposed methodology we are using Bi-LSTM(Bi-directional Long Short Term Memory) with attention mechanism. Bi-LSTM works well with complex data and try to capture the data is genuine or fake. Together with Bi-LSTM we are using attention mechanism which helps to identify relevant and meaningful data and capture the relationships between the data. It can identify long range dependencies between the data by selectively attending features or contexts.

Due to this specific characteristic ,we can easily identify the amount of fakeness in job postings sites. By incorporating both forward and backward context, Bi- LSTM models can capture the subtleties of job descriptions, requirements, and other textual features that may signal fraudulent behaviour.

By incorporating the attention mechanism, our model is able to concentrate on critical features, such as suspicious phrases or abnormal job requirements, that may indicate fraud. The integration of both text and structured data, such as job location, employment type, and salary range, further strengthens the model's capability to sort out between real and fake job advertisements.

This paper explores the use of Bi-LSTM for detecting fake job postings, leveraging both text and numerical features to build a robust detection model. By applying advanced word embedding techniques and integrating structured data, this study aims to enhance the correctness and accuracy of fraud detection systems.

II. LITERATURE REVIEW

Fake Job Postings using several algorithm have been used in past years. But many algorithms failed to achieve the relative dependencies between the text and context of text.

In the paper "Identifying fake or real posting-Machine learning Approach", [Aug 2021] some et al Naive Bayes algorithm is used which is based on the conditional probability of the values which is used for classification of data .It uses relatively small dataset which predicts the given job is fake or authentic. In this, dependencies between text is not considered.

In another research paper "Real and Fake classification using NLP and machine learning techniques" published on [3rd JAN 2024], it uses a comparision of several machine learning and NLP techniques Random forest classifier achieved greater accuracy but bidirectionally dependencies between text doesn't mentioned in this also.

Other work on" Detecting fake job posting using Bidirectional LSTM"[3rd march 2023] ,which helps to capture complex patterns and contextual relationships between text.Long term dependencies is not mentioned here.

Some related works in this area and their research gaps are listed in table

Author	Datas et	Classifers Used	Accura cy	Resear ch gaps
Lee, D., Kim, H.	EMSCA D	Neural networks	97.9%	Scalability issues
Dutta & Bandyopadh yay (2020)	EMSCA D (17,880)	Naive Bayes, Decision Tree, KNN,	98.27%	Deep learning models not explored.
Habiba et al. (2021)	EMSCA D (18,000)	KNN, Decision Tree, SVM, Naive Bayes, Random Forest, DNN	97.7%	No use of text features in DNN models.
Verma, S., Gupta, R.	Kaggle Real/Fa ke Job	Machine leraning approaches	98.2%	No use of deep learning.
Brown, J., Davis, A.	Custom Dataset		98.6%	High computati onal cost with BERT
Pillai (2023)	Kaggle (17,880)	Bi-LSTM, Random Forest, LightGB M	98.71%	Class imbalance, generalizat ion challenges
Nasser & Alzaanin (2020)	Kaggle (17,880)	Naive Bayes, SVM, Decision tree, Random forest, KNN	98.2%	Limited deep learning model exploratio n
Gupta, A., Kumar, R.	Kaggle	SVM	97.2%	The class imbalance issue
Zhang, Y., Li, K.	EMSC AD	CNN	97.9%	Does not Capture
				long term dependenc ies.
Chen, L., Wu, P.	EMSC AD	Logistic regression	95.4%	It struggles with unstructed Data
Anderson, B., Huang, C.	Kaggle dataset	XGBOOS T	98.3%	No time series analysis
Li, X., Zhang, M.	EMSC AD	Random forest	96.5%	No hybrid approach

Patel, R., Singh, P.	Kaggle Real/Fa ke Job Posting	LSTM	97.5%	Missing attention mechanis m
Khan, S., Aurangzeb, Z.	EMSC AD		96.3%	Heavy reliance on traditional text mining Only.
Wang, Y., Li, F.	Kaggle	Neural networks	97.8%	Lacks domain specific embedding S

By taking into account all the research gaps, we are proposing a system which can detect fakeness even more accurately. In the proposed methodology, we have to identify fake job posting using BILSTM and together with attention mechanism, we help to capture the contextual relationship between the text as well as used to predict the long term dependencies and filter out relevant and important words in sentences using this attention mechanism.

III. METHODOLOGY

Detection of fake job posting in online media using BI-LSTM and attention mechanism.

The BI-LSTM model captures both past and future context in job descriptions, while the attention mechanism specifically focuses on the most important parts of the job description, such as suspicious phrases, job titles, and abnormal requirements. By integrating both text and structured data, this study seeks to create a more accurate and an interpretable model for detecting fraudulent job advertisements. The goal is to create a system that overcomes the limitations of traditional models, offering higher accuracy, better focus on relevant patterns, and enhanced real-time fraud detection capabilities in the online job market.

To predict the fakeness in job postings using BI-LSTM and attention mechanism.

1) Enhance BI-LSTM to capture contextual information: With the help of Bi-LSTM to effectively process job descriptions by capturing both past (previous words) and future (subsequent words) context within the textual data, which is crucial for identifying subtle fraudulent patterns in job postings.

2) Incorporate an Attention Mechanism to enhance focus: Integrate an attention mechanism that effectively focuses on important features of the job description, such as keywords, suspicious phrases, or anomalies in job requirements and benefits, helps the model's capability to sort out between real and fake job posts.

3) Enhance accuracy and robustness: Improves the amount of accuracy in detecting fake job postings in assessing with traditional machine learning models, develop a model providing better results by calculating precision, recall, and F1-score.

4) Handle complex, unstructured textual data: Manages structured(Job type,job location) and unstructured data(Job desc,companyprofile).

5) Enhance interpretability: Through the attention mechanism, provide greater interpretability by revealing which parts of the job postings contribute most to the fraud detection decision, making the model's predictions more transparent and actionable.

6) Develop a scalable and real-time detection solution: Creating a scalable model which can esily deployed in real-world applications.

The proposed method focus on BI-LSTM with attention mechanism .BI- LSTM captures the context of words in both forward and backward direction.In attention mechanism it will focus on the most revelant information. Together with BI-LSTM and attention mechanism it will help to collect the revelant pieces of information easily .So it can definitely help to detect fakeness in the fake job advertisements.

IV. METHODOLOGY DIAGRAM AND PSEUDOCODE OF PROPOSED MODIFICATION

Steps of fake detection using Bi-LSTM with Attention Mechanism to detect fake job postings

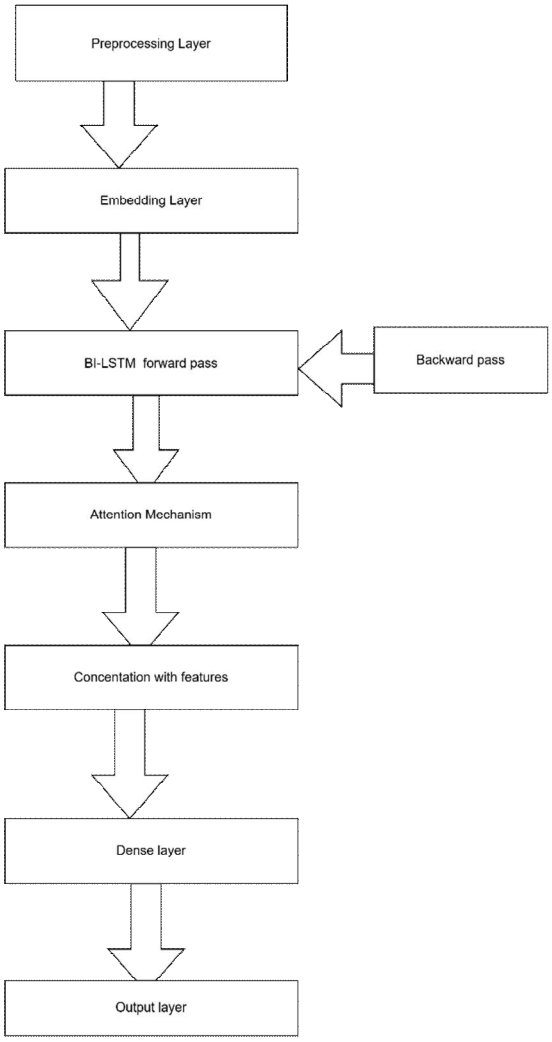


Fig 1.Proposed methodology flowchart

1) Step 1: Preprocess the data (Tokenize, Clean, Pad sequences) Pad the sequences to ensure uniform input size.
2) Step 2: Define the BI-LSTM model with Attention

- Bi-LSTM Layer: Capture sequential information in both forward and backward directions
- Attention Mechanism: Calculate attention scores for each word in the sequence
- Dense Layer: Learn from the combined output
- Output Layer: Predict if "the job posting is fake (1) or real (0) "

3) Step 3: Define the Attention Layer
4) Step 4: Train the proposed model
5) Step 5: Evaluate the proposed model

V. EXPERIMENTAL SETUP

Hyperparameter	Description	Value/Range
Embedding Dimension	The size of the word embedding vectors (e.g., using GloVe or Word2Vec).	100
LSTM Units	Number of units in the BI-LSTM layer.	128
Attention Units	Number of units in the attention mechanism for calculating attention weights.	1 (per time step)
Batch Size	Number of units processed before changing the model's weights.	25
Epochs	Number of complete passes through which training dataset goes.	2
Max.Sequence Length	Maximum length of the input sequence (job description) after padding/truncating.	100-150 Tokens

VI. RESULTS

The experimental results shows that this method of bilstm with attention mechanism acquire 99% accuracy in detecting fake job postings which is higher than any than earlier models.

The results are given in the table which is showing the fake value accuracy is 99%. Precision,Recall,F1-score and Support values using BI-LSTM and attention mechanism

.

Model	Accuracy	Precision	Recall	F1-score
Bidirectional LSTM and attention mechanism	99.38%	0.98	0.99	0.99

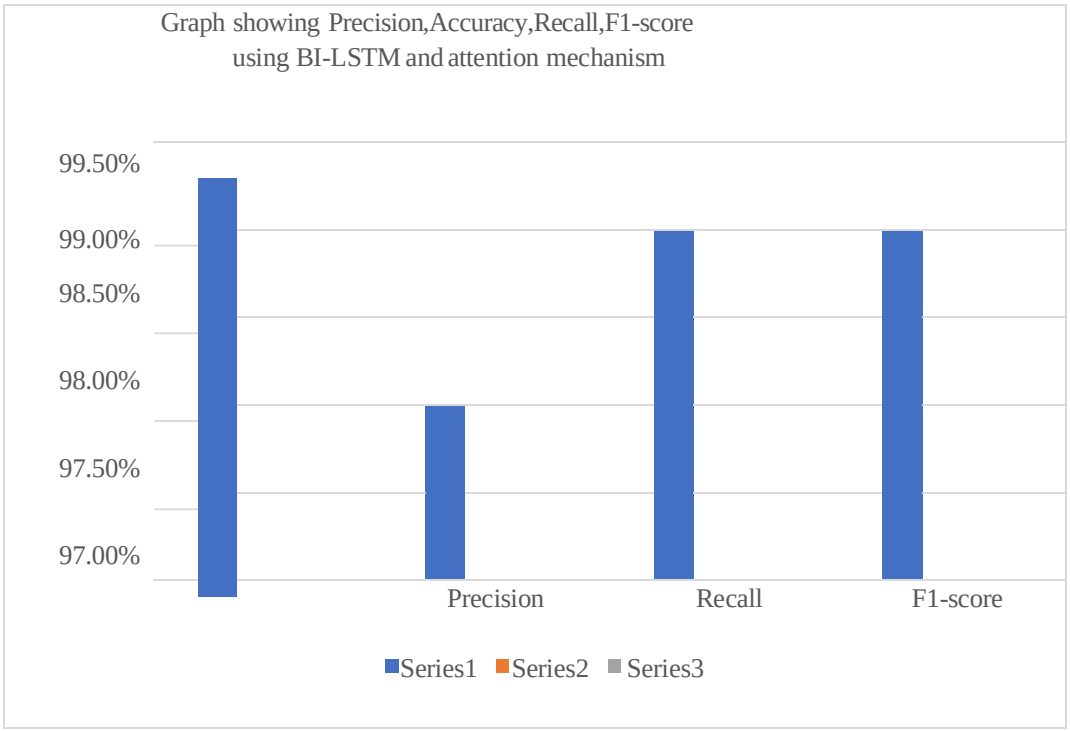


Fig 2:Graph showing Precision,Accuracy,Recall,F1-score using BI-LSTM and attention mechanism

In the previous works they focus only on the BI-LSTM model and got the results as follows:

Model	Accuracy	Precision	Recall	F1-score
Bidirectional LSTM	98.71 %	0.89	0.83	0.86

The performance metrics achieved by ensemble models is as follows:

Model	Accuracy	Precision	Recall	F1-score
Light GBM	98.18%	0.94	0.66	0.77
Random Forest	97.04%	0.97	0.40	0.57
Gradient Boosting	96.67%	0.86	0.37	0.62

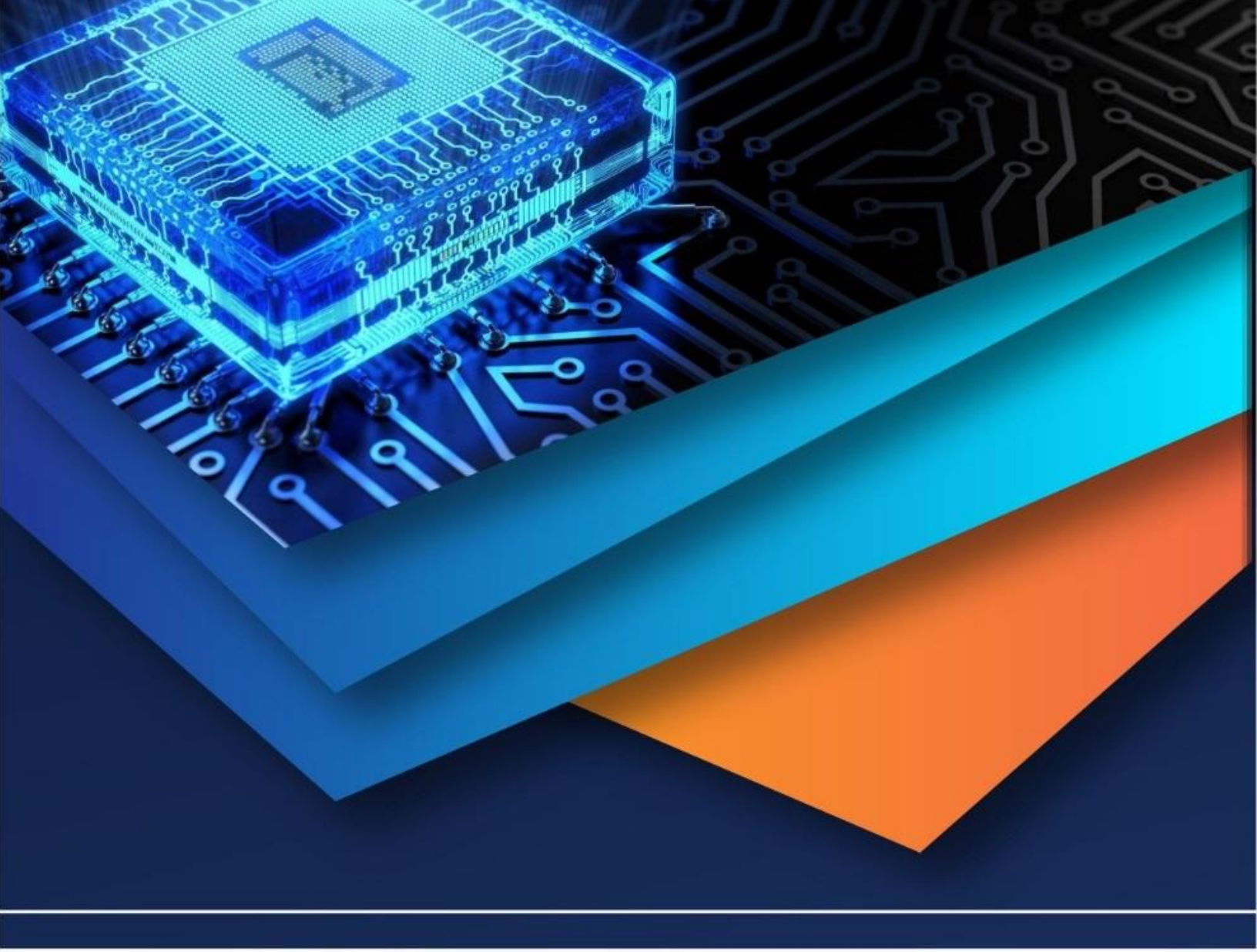
By these looking these results we can say that BI-LSTM with attention mechanism model got more accurate results of fake job postings than any other model.

VII. CONCLUSION

In this article we presented an approach to detect fake job posting using BILSTM and attention mechanism and from the experimental results we got this model acquired 99% accuracy. So from this study, we can conclude that this model can detect fake jobs very efficiently. Therefore this model is a boon for job seekers from getting trap into fake job sites and they can make them safe from themselves from financial and mental tortures.

REFERENCES

[1] K. Sridevi, G. Likitha , Real or Fake Job posting detection,2024
[2] Maddi Sravya Reddy, Maddikera Hemanth Lal, Fake Job post detection using machine learning ,2024.
[3] V Anbarasu, Dr. S. Selvakani, Mrs. K. Vasumathi, Fake Job Prediction Using Machine Learning,2024.
[4] Sherina Sara Jaison, Mallikarjuna Kodabagi , Identifying real and fake job posting using machine learning ,2023.
[5] Mrinal Kumari, NSK Satya kala, Nandini R, Fake job posting prediction using machine learning approach ,2023.
[6] Hina Afsal,Furqan Rustam,Identifying fake job posting using selective features and resampling techniques,2023.
[7] Aravind Sasidharan Pillai,Detecting fake job posting using Bidirectional LSTM,2023.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY