

# Deep Learning Based Sentiment Analysis Regarding Job Sectors in Bangladesh Using Glassdoor Data

Sovon Chakraborty  
Computer Science and Engineering  
BRAC University  
Dhaka Bangladesh  
sovon.chakraborty@bracu.ac.bd

Annajiat Alim Rasel  
Senior Lecturer  
Computer Science and Engineering  
BRAC University  
Dhaka Bangladesh  
annajiat@bracu.ac.bd

Muhammed Yaseen Morshed Adib  
Computer Science and Engineering  
BRAC University  
Dhaka Bangladesh  
yaseen.morshed.adib@g.bracu.ac.bd

Munim Bin Muquith  
Computer Science and Engineering  
BRAC University  
Dhaka Bangladesh  
munim.bin.muquith@g.bracu.ac.bd

**Abstract— will be updated**

**Keywords—sentiment analysis, deep learning, Bi-LSTM, Bi-GRU (key words)**

## I. INTRODUCTION

The growth of computer and internet-based technologies is a factor in the quick rise in the volume and accessibility of data globally. Larger amounts of data are now easier to obtain than ever before, enabling new information processing or decision-making tools to be developed through analytical formulas and rule-development opportunities (data-processing algorithms).

A potent quantitative technique that extracts insights from industrial data has most recently been developed as a result of the proliferation of clever machine learning algorithms in the field of computer science. Political science [1], transportation [2], biology and medical research [3], and many more fields use supervised machine learning techniques. Researchers investigated a variety of machine learning techniques to enhance HR (human resources) management outcomes in response to developments in information technology [4-6].

Many data analytics projects now go by the name "big data." The phrase "Big Data" first came into use to describe a technological advancement that made it possible to collect a lot of data [7]. Depending on the context in which big data has been referenced since, the term has expanded to describe various features of analysis in many disciplines. The phrase is now used to describe data qualities and processing capabilities, and it covers both technical and commercial aspects of data collection activities [8].

Gathering proper knowledge regarding University job employees is the main idea here. A specialized dataset will be built in order to analyze the sentiments of university employees. The dataset will further be used for understanding the satisfactions and objections of University job employees.

Machine learning and Deep learning models are performing outstandingly regarding text analysis. In this research, several deep learning architectures namely Bi- directional LSTM (Bi-LSTM), Bi-directional Gated recurrent unit (Bi-

GRU) and 1-D Convolutional Neural network (CNN) have been studied with the gathered dataset. Though, these architectures have been widely used but in the case of Bangladeshi University employees job satisfaction these models have been applied barely. Data is visualized and analyzed precisely so that this research can be used in future also.

This paper's primary contribution can be summed up as follows:

1. Building a specialized dataset that has more than 10,000 comments scrapped from Glassdoor. After a critical review, the comments are annotated with the appropriate polarity value.
2. Understanding the way in which most university employees currently respond to the job situations after the Post Covid-19 period and finding out a proper pattern related to this field.
3. Three deep learning structures are fine tuned for achieving maximum utilization. The models are also validated properly.

## II. LITERATURE REVIEW

Many researchers have become interested in sentiment analysis in recent years. Despite the abundance of sentiment analysis works in several languages in the literature, the contribution regarding university employee job satisfaction is still insufficient. The prior research produced satisfactory results by utilizing several machine learning and deep learning techniques. Convolutional neural networks (CNN), transformers, artificial neural networks (ANN), and other deep learning architectures are well known for their use in sentiment analysis applications.

Luo et al. [16] conducted a sentiment study of the tourism industry based on consumer reviews. The researchers used Bi-LSTM designs, which only take into account positive and negative polarity; neutral sentiment was not taken into account. Based on data from tweets made by people during the Covid-19 pandemic, analysts evaluated public reaction toward Moderna, AstraZeneca, and Pfizer. [17]. No prior Machine learning or deep learning architectures have been suggested for this research. The findings [18] employs ELM

(Emitted Light Modulation) for aspect-sentiment embedding. The authors initially collected data from Glassdoor.com, after which they used the NLTK Tokenize Package to tokenize the reviews into sentences. The unaltered text is then changed into embedding's of review-level summaries. According [19] a semi-open question seems to be useful for determining work satisfaction in the paper. But the credibility of these questions are not properly given. In [20] authors have predicted company ratings based on comments that are available on Glassdoor. For this research they have used Bi-LSTM and Naive Bayes classifiers. The architectures were not fine-tuned as a result relatively less accuracy has been achieved here. Till date, CNN and Bi-GRU have been used scarcely in the field of Sentiment analysis.

In this research, the prime purpose is to understand the satisfaction of University employees with a number of deep learning based architectures. All the architectures are fine-tuned. To compare between these models several performance metrics have been used. In the related work section, a minimal amount of research has been conducted in order to detect the sentiments of University employees.

#### REFERENCES

- [1] Bajpai, R., Hazarika, D., Singh, K., Gorantla, S., Cambria, E., & Zimmerman, R. (2019). Aspect-sentiment embeddings for company profiling and employee opinion mining. arXiv preprint arXiv:1902.08342.
- [2] Chinazzo, G. (2021). Investigating the indoor environmental quality of different workplaces through web-scraping and text-mining of Glassdoor reviews. *Building Research & Information*, 49(6), 695-713.
- [3] Luo, N., Zhou, Y., & Shon, J. (2016). Employee satisfaction and corporate performance: Mining employee reviews on glassdoor. Com.
- [4] Uchida, E., & Kino, Y. (2021). Study on the Relationship Between Employee Satisfaction and Corporate Performance in Japan via Text Mining. *Procedia Computer Science*, 192, 1730-1739.
- [5] Dina, N. Z., & Juniarta, N. (2020). Aspect-based Sentiment Analysis of Employee's Review Experience. *Journal of Information Systems Engineering and Business Intelligence*, 6(1), 79-88.
- [6] Das Swain, V., Saha, K., Reddy, M. D., Rajvanshy, H., Abowd, G. D., & De Choudhury, M. (2020, April). Modeling organizational culture with workplace experiences shared on glassdoor. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1-15).
- [7] Kashive, N., Khanna, V. T., & Bharthi, M. N. (2020). Employer branding through crowdsourcing: understanding the sentiments of employees. *Journal of Indian Business Research*.
- [8] Feng, S. F. (2020). Job Satisfaction, Management Sentiment, and Financial Performance: text Analysis with Job Reviews from Indeed. com. *Authorea Preprints*.
- [9] Wijngaards, I., Burger, M., & van Exel, J. (2020). Unpacking the quantifying and qualifying potential of semi-open job satisfaction questions through computer-aided sentiment analysis. *Journal of Well-Being Assessment*, 4(3), 391-417.
- [10] Dube, S., & Zhu, C. (2021). The Disciplinary Effect of Social Media: Evidence from Firms' Responses to Glassdoor Reviews. *Journal of Accounting Research*, 59(5), 1783-1825.
- [11] Jung, Y., & Suh, Y. (2019). Mining the voice of employees: A text mining approach to identifying and analyzing job satisfaction factors from online employee reviews. *Decision Support Systems*, 123, 113074.
- [12] Leah-Martin, V. (2017). Relative compensation and employee satisfaction. Available at SSRN 2896268.
- [13] Kibriya, A. M., Frank, E., Pfahringer, B., & Holmes, G. (2004, December). Multinomial naive bayes for text categorization revisited. In *Australasian Joint Conference on Artificial Intelligence* (pp. 488-499). Springer, Berlin, Heidelberg.
- [14] Su, J., Shirab, J. S., & Matwin, S. (2011, January). Large scale text classification using semi-supervised multinomial naive bayes. In *ICML*.
- [15] Singh, G., Kumar, B., Gaur, L., & Tyagi, A. (2019, April). Comparison between multinomial and Bernoulli naïve Bayes for text classification. In *2019 International Conference on Automation, Computational and Technology Management (ICACTM)* (pp. 593-596). IEEE.
- [16] Singh, M., Bhatt, M. W., Bedi, H. S., & Mishra, U. (2020). Performance of bernoulli's naïve bayes classifier in the detection of fake news. *Materials Today: Proceedings*.
- [17] Yoo, S. H., Geng, H., Chiu, T. L., Yu, S. K., Cho, D. C., Heo, J., ... & Lee, H. (2020). Deep learning-based decision-tree classifier for COVID-19 diagnosis from chest X-ray imaging. *Frontiers in medicine*, 7, 427.
- [18] Rani, S., & Kumar, P. (2019). Deep learning based sentiment analysis using convolution neural networks. *Arabian Journal for Science and Engineering*, 44(4), 3305-3314.
- [19] Dube, S., & Zhu, C. (2021). The Disciplinary Effect of Social Media: Evidence from Firms' Responses to Glassdoor Reviews. *Journal of Accounting Research*, 59(5), 1783-1825.