totox

by D H

Submission date: 16-Jul-2023 01:32AM (UTC-0400)
Submission ID: 2131722429
File name: 6372-17076-1-RV.docx (187.63K)
Word count: 3686
Character count: 21570

JIKO (Jurnal Informatika dan Komputer) Accredited KEMENDIKBUD RISTEK, No.105/E/KPT/2022
Vol. x, No. x, April 2023, hlm. x-x p-ISSN: 2614-8897
DOI: 10.33387/jiko e-ISSN: 2656-1948

EXPLORING THE EFFECTIVENESS OF DEEP LEARNING IN ANALYZING REVIEW SENTIMENT

Mariyanto [1]*, Hilman F. Pardede [2]

[1,2] Ilmu Komputer, Universitas Nusa Mandiri, Jakarta, Indonesia
*Email: [1]14210250@nusamandiri.ac.id, [2]hilm003@brin.go.id

(Received: dd mmm yyyy, Revised: dd mmm yyyy, Accepted: dd mmm yyyy)

Abstract

This research focuses on the analysis of sentiments in office product reviews using LSTM and CNN methods, with the addition of a novel approach called GRU combined with Multihead Attention. This study aimed to compare the performance of these methods and provide insights into their effectiveness in sentiment analysis tasks. The analysis involved training and evaluating the models using a dataset of office product reviews. Various evaluation metrics, such as the accuracy and loss, were employed to assess the performance of the models.
The results show that the LSTM and CNN methods play significant roles in the sentiment analysis of office product reviews. In addition, the newly introduced GRU + Multihead Attention method shows promising results, indicating its potential for improving the accuracy of sentiment analysis. The findings of this study provide valuable insights for researchers and practitioners in the field of sentiment analysis, particularly in the context of office product reviews. This study serves as a foundation for further research and development of more advanced sentiment analysis methods, leveraging the benefits of both GRU and Multihead Attention in the analysis of office product sentiment.

Keywords: *word embedding, CNN, LSTM, GRU*

**Corresponding Author: Author1*

1. INTRODUCTION

Studying product reviews from customers holds immense importance in the business world. Customer reviews provide valuable insights into the perceptions, opinions, and experiences of consumers with specific products or services. This information is highly beneficial for companies in several ways.

Firstly, analyzing product reviews allows businesses to understand the strengths and weaknesses of their offerings. By studying customer feedback, companies can identify the aspects of their products that are well-received and appreciated, enabling them to further enhance and highlight these positive features. On the other hand, negative reviews provide valuable feedback on areas that need improvement, allowing companies to address any issues and make necessary adjustments to enhance customer satisfaction.

Research by Li et al. (2019) [1] emphasizes the significance of analyzing customer reviews for improving product quality. They conducted a study on sentiment analysis of customer reviews for a range of products and found that companies can identify specific areas of improvement based on customer feedback. This approach enables companies to make data-driven decisions and enhance product development strategies.

Secondly, customer reviews serve as a powerful tool for reputation management. Positive reviews act as testimonials, showcasing the satisfaction and positive experiences of customers, which can significantly influence the purchasing decisions of potential buyers. Conversely, negative reviews highlight areas of concern and give companies an opportunity to respond and rectify any issues, demonstrating their commitment to customer service and resolving problems.

Research by Zhang and Duan (2020) [2] highlights the importance of managing online reputation through effective handling of customer reviews. They found that companies that actively engage with customers and respond to their feedback, both positive and negative, can improve their online reputation and build trust with consumers.

Furthermore, studying product reviews enables businesses to gain a competitive edge in the market. By analyzing customer feedback, companies can identify emerging trends, preferences, and demands, allowing them to tailor their products and services to meet the evolving needs of their target audience. This customer-centric approach can help companies stay ahead of the competition, attract new customers, and retain existing ones.

In conclusion, studying product reviews from customers is crucial for businesses to gain insights, improve their offerings, manage their reputation, and stay competitive. By actively analyzing and leveraging customer feedback, companies can make informed decisions, enhance customer satisfaction, and ultimately drive business growth.

This study focuses on the analysis of a dataset consisting of reviews of office products from Amazon annotated with sentiment labels. The dataset undergoes pre-processing steps, such as text cleansing, tokenization, and stopword removal. To convert textual data into numeric vectors, word embedding, which captures the semantic meaning of the word, is used.

The core of this research is the CNN-LSTM model in an office product review. In addition to the CNN-LSTM hybrid model, the authors introduced the GRU+Multihead method. By combining a Gated Recurrent Unit (GRU) architecture with Multihead Attention, this new method enhances the model's ability to capture sequential information and pay attention to relevant features in office product review text.

The experimental results show that the CNN-LSTM model is effective for sentiment analysis of office product reviews. In addition, the GRU method showed better performance, outperforming the hybrid CNN-LSTM model and the baseline model in terms of accuracy and other evaluation metrics.

2. RELATED WORK

Rehman et.al. [3] presented a new approach to sentiment analysis of film reviews by combining the strengths of convolutional neural networks (CNN) and short-term memory (LSTM) networks. The main objective of this study is to improve the accuracy of sentiment analysis in the context of film reviews. The study began by collecting a dataset of film reviews, including positive and negative sentiment labels. The weights and biases of the CNN-LSTM hybrid model were optimized using backpropagation and gradient descent to minimize the loss function and increase the accuracy of sentiment prediction. The experimental results show that the CNN-LSTM hybrid model achieves higher accuracy and better sentiment analysis performance than the traditional approach. The integration of CNN and LSTM enabled the model to effectively capture local and sequential features, resulting in more accurate sentiment predictions for film reviews.

Suryana [4] introduces an approach to extract contextual insights from product reviews using LSTM networks and word insertion. The aim of this study is to harness the power of deep learning techniques to gain a deeper understanding of the context and sentiments expressed in product reviews. Word embeddings such as Word2Vec or GloVe are used to represent words as numeric vectors that capture semantic information. The experimental results showed that the in-depth LSTM model, along with word embedding, effectively generated contextual insights from product reviews. The model achieves high accuracy in sentiment prediction and provides valuable contextual information that helps understand review nuances.

3. RESEARCH METHOD

By expressing words as dense numerical vectors and using machine learning, word embedding revolutionizes the analysis of product reviews. This is done by allowing computers to comprehend the context and sentiments represented in the text. Businesses can extract specific product features customers talk about and positive and negative feelings from customer feedback. This is done using techniques like sentiment analysis and aspect-based sentiment analysis. Companies may automate and expand their review research and make data-driven decisions to increase customer satisfaction and their products and services by utilizing word embedding and machine learning.

3.1 Data Source

The "Office product reviews" dataset, available at the Stanford Network Analysis Project (SNAP) website, contains 800,357 reviews of office products from Amazon. It includes information about the reviewers, products, ratings, helpfulness, and timestamps. This dataset is valuable for studying customer feedback and sentiment towards office products, and it can be used for tasks such as sentiment analysis and understanding customer preferences in the office product market.

3.2 Analysis of Variance (ANOVA)

ANOVA, or Analysis of Variance, is a statistical technique used to compare means across two or more groups and determine if there are significant differences between them. It is widely employed in data science and various research fields to analyze the impact of categorical independent variables on a continuous dependent variable.

In ANOVA [5][11], the total variation in the data is partitioned into two components: variation within each group and variation between the groups. The analysis aims to determine whether the observed between-group differences are statistically significant or simply due to random variation. ANOVA calculates an F statistic, which is used to assess the significance of these differences.

ANOVA offers several advantages in data science analysis. It allows for the simultaneous comparison of multiple groups, providing a more efficient and comprehensive analysis compared to conducting pairwise t-tests. ANOVA also provides insights into the magnitude of the observed differences and can account for potential interaction effects between independent variables.

3.2 Machine Learning

Deep learning is a potent branch of machine learning that focuses on building multi-layered artificial neural networks that analyze and comprehend complex data representations. Deep learning algorithms can extract high-level characteristics from enormous volumes of data and learn complex patterns by utilizing these powerful neural networks. This enables machines to execute tasks like speech recognition, natural language processing, and image recognition with remarkable accuracy. Deep learning has significantly advanced many disciplines, including computer vision, natural language processing, and data analytics. This has paved the way for creative applications and game-changing developments in artificial intelligence. Deep learning can automatically learn from data and adapt.

Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs)[12] are a class of deep learning models widely used for various computer vision tasks, including image classification and object recognition. CNNs have revolutionized the field of computer vision due to their ability to automatically learn meaningful representations from raw pixel data. The core concept behind CNNs is the convolutional layer, which applies a set of learnable filters or kernels to the input image, extracting local features while preserving spatial relationships.
The formula for the convolution operation in a CNN can be expressed as follows:

$$C(i,j) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} I\,(i+m, j+n).\ K(m,n)$$

where $C(i,j)$ represents the output at position (i,j), $I(i+m, j+n)$ is the input pixel value at position $(i+m, j+n)$, $K(m,n)$ denotes the kernel weight at position (m,n), and M and N represent the size of the kernel.

One seminal work that contributed to the development of CNNs is the paper by LeCun et al. (1998) [6]. The authors introduced the LeNet-5 architecture, which featured a convolutional layer followed by pooling layers and fully connected layers. This pioneering work demonstrated the effectiveness of CNNs in handwritten digit recognition and laid the foundation for subsequent advancements in the field.

Since then, numerous variations and improvements to CNN architectures have been proposed. Notably, the VGG network proposed by Simonyan and Zisserman (2014) [7] achieved state-of-the-art performance on the ImageNet dataset by utilizing deeper architectures with smaller convolutional filters. This work highlighted the importance of depth in CNNs and inspired the development of deeper networks, such as the popular ResNet and Inception models.

In conclusion, CNNs have revolutionized computer vision by automatically learning meaningful features from raw pixel data. They have become a fundamental tool for various tasks, and their effectiveness has been demonstrated through groundbreaking works like LeNet-5 by LeCun et al. (1998) [6] and the VGG network by Simonyan and Zisserman (2014) [7].

Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) architecture that is particularly effective in modeling sequential data, such as natural language processing and time series analysis. LSTMs address the vanishing gradient problem associated with traditional RNNs by incorporating memory cells and gates that control the flow of information. This enables LSTMs to capture long-term dependencies and learn complex patterns over extended sequences.

The key components of an LSTM cell include an input gate (i_t), a forget gate (f_t), an output gate (O_t) a memory cell (c_t) and a hidden state (h_t). The formulas governing the operations of an LSTM cell are as follows:

$$i_t = \sigma(W_{ix}.x_t + W_{ih}.h_{(t-i)} + b_i$$
$$f_t = \sigma(W_{fx}.x_t + W_{fh}.h_{(t-i)} + b_f$$
$$o_t = \sigma(W_{ox}.x_t + W_{oh}.h_{(t-i)} + b_o$$
$$g_t = tanh(W_{gx}.x_t + W_{gh}.h_{(t-i)} + b_g$$
$$c_t = f_x.c_t + i_t.g_t$$
$$h_t = o_t.\tanh(c_t)$$

where x_t represents the input at time step t, σ denotes the sigmoid activation function, W and b are the weight matrices and bias vectors, and $\cdot$ indicates matrix multiplication.

The LSTM architecture was first introduced by Hochreiter and Schmidhuber (1997) [8]. Their groundbreaking paper proposed the LSTM model as a solution to the vanishing gradient problem in RNNs. They demonstrated the effectiveness of LSTM networks in various tasks, including speech recognition and language modeling, highlighting their ability to capture long-term dependencies.

Since its inception, LSTM has become one of the most popular and widely adopted RNN architectures. Its effectiveness has been demonstrated in numerous applications, including machine translation, sentiment analysis, and speech recognition, among others.
In conclusion, LSTM networks have revolutionized the modeling of sequential data by addressing the

vanishing gradient problem and capturing long-term dependencies [13]. The LSTM architecture, proposed by Hochreiter and Schmidhuber (1997) [8], has become a fundamental tool in various fields and continues to drive advancements in sequence modeling.

Gated Recurrent Unit (GRU)

Gated Recurrent Unit (GRU) is a variant of the recurrent neural network (RNN) architecture that addresses vanishing gradient problem and offers an efficient alternative to Long Short-Term Memory (LSTM) networks. GRUs utilize gating mechanisms to control the flow of information through the network, enabling them to capture and propagate relevant information over long sequences [14].

Similar to LSTM, GRU cells consist of an update gate z_t, a reset gate r_t, and a candidate activation h'_t. The formulas governing the operations of a GRU cell are as follows:

$$z_t = \sigma(W_{zx}.x_t + W_{zh}.h_{(t-i)} + b_z$$
$$r_t = \sigma(W_{rx}.x_t + W_{rh}.h_{(t-i)} + b_r$$
$$h'_t = \sigma(W_{hx}.x_t + W_{hh}.(r_t \odot h_{t-1} + b_h)$$
$$h_t = (1 - z_t) \odot h_{(t-1)} + z_t \odot h'_t$$

where x_t represents the input at time step t, denotes the sigmoid activation function, $\odot$ indicates element-wise multiplication, and W and b are the weight matrices and bias vectors.

The GRU architecture was introduced by Cho et al. (2014) [1]. Their seminal paper presented GRUs as a simplified variant of LSTM, demonstrating comparable performance on various sequential tasks while requiring fewer parameters. GRUs were shown to effectively capture long-term dependencies and exhibit better training efficiency than LSTM.

Since its introduction GRU has gained widespread popularity and has been successfully applied to various domains, including natural language processing, speech recognition, and machine translation. Its efficiency, simplicity, and competitive performance have made GRU a popular choice for modeling sequential data [15].

In summary, GRU networks offer an efficient and effective solution for modeling sequential data. Introduced by Cho et al. (2014) [9], GRUs utilize gating mechanisms to capture long-term dependencies while maintaining computational efficiency. Their simplicity and competitive performance have established GRUs as a valuable tool in various applications.

4. RESULT AND DISCUSSION

The Figure 1 illustrates the data distribution for the "Overall" class in the dataset used in this study. This data distribution provides an overview of the sample distribution across the different "Overall" classes present in the dataset. In sentiment analysis and product review classification, the "Overall" class represents the overall sentiment or rating of a product. This data distribution provides valuable insights into the balance or imbalance of sample sizes within each "Overall" class.

In Figure 1, the x-axis represents the available "Overall" classes, while the y-axis indicates the number of samples in each class. Data distribution reflects a balanced or imbalanced distribution. If there is a significant difference in the number of samples across each "Overall" class, the data distribution shows varying bar heights on the graph. Understanding the data distribution for the "Overall" class is crucial for comprehending the relationship between the sample sizes within each class and the classification outcomes generated by the model. In the case of imbalanced data distribution, the classification results in the majority class tend to have a greater impact on the overall model evaluation.

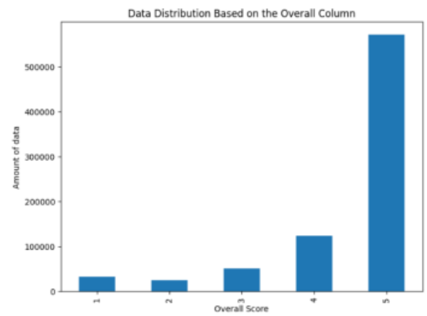

Figure 1: Data distribution for the Review class

The results of the modeling using the ANOVA method are shown in Figure 2. This provides a clear visualization of the features that significantly influence the target and aims to evaluate the relative importance of each feature to the target variable to be predicted. If the ANOVA significance value is high, this indicates that the feature has a significant influence on the target. The visualization presented in Figure 2 shows that the "summary" feature has the greatest influence on the target (overall). Furthermore, the data used in this study only uses the "summary" and "overall" features for word embedding processing.

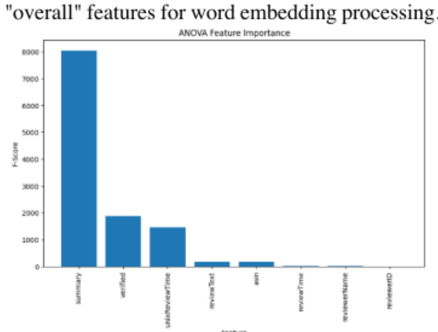

Figure 2. ANOVA Test Results

Figure 3 displays a visual representation of the test results obtained from the analysis of office product reviews using the LSTM, CNN, and GRU methods. The purpose of this test is to compare the performance of the three methods in the sentiment analysis of office product reviews. The results depicted in Figure 3 provide an overview of the relative performance of LSTM, CNN, and GRU in terms of accuracy. Through this visualization, it is easy to observe and compare the test results of the three methods.

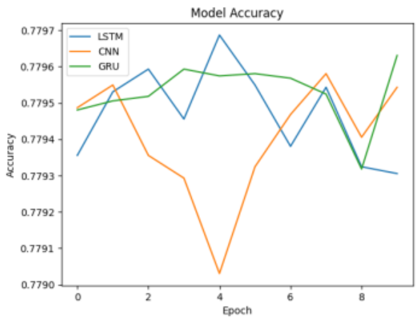

Figure 3. Results of machine learning model accuracy

Based on Figure 3, several conclusions can be drawn regarding the performance of this method. For example, if the LSTM demonstrates higher accuracy than CNN and GRU, it may be more effective at performing sentiment analysis of office product reviews. On the other hand, if CNN or GRU perform better in a particular evaluation metric, it indicates the superiority of that method in certain aspects of sentiment analysis.

In Figure 4, the x-axis indicates the number of epochs or iterations performed during training, and the y-axis indicates the model loss value. The model loss change curve in Figure 4 provides information about the extent to which the model experiences a decrease in error as the training progresses.

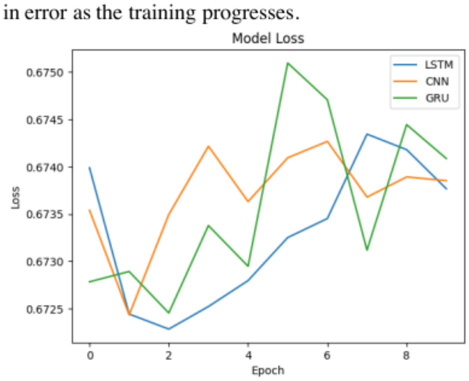

Figure 4. Loss generated by Machine Learning

When the model loss curve increases as the training progresses, this indicates that the model cannot reduce errors or errors during the learning process. This can be caused by several factors, such as the complex model architecture, dataset size that is too small, or training conditions that are not optimal.

An increase in model loss may indicate that the model is unable to learn the patterns in the dataset properly or may experience overfitting, where the model "memorizes" the training data too much, so that it cannot be generalized properly to new data.

In the process of training a neural network, we can compare the time taken for an epoch, which represents a complete training cycle, across the three different types of architectures. Let us now examine the average time estimates for an epoch in each architecture. LSTM networks typically require approximately 90 s per epoch. LSTMs are generally used to process sequential data, and are known for their ability to handle long-term dependencies. The CNN architecture typically completed an epoch of approximately 59 s. The GRU network, similar to LSTM, showed an average epoch of approximately 84 s. The GRU model is often used for sequential data modeling and has the advantage of maintaining short-term information.

5. CONCLUSION

The test results show that three methods, LSTM, CNN, and GRU, play an important role in the sentiment analysis of office product reviews. Each method has its strengths and weaknesses when dealing with this task. LSTM had the highest accuracy with the longest epoch cycle time. The accuracy generated by the CNN is the lowest but requires the shortest time. GRU has moderate accuracy with moderate epoch times.

The selection of an appropriate method depends on the specific context of the sentiment analysis and the characteristics of the dataset. Based on the test results, it is suggested that LSTM, CNN, and GRU be considered for sentiment analysis of office product reviews, choosing the method that best suits the research objectives and requirements.

This study serves as an initial foundation for the further development of sentiment analysis of office product reviews. Future research may involve exploring other methods, optimizing models, incorporating additional features, or adapting to more specific contexts and types of product review.

REFERENCES

REFERENCE:

[1] J. Panthati, J. Bhaskar, T. K. Ranga, and M. R. Challa, “Sentiment Analysis of Product Reviews using Deep Learning,” in *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2018, pp. 2408–2414. doi: 10.1109/ICACCI.2018.8554551.

[2] R. Thakur, “Customer engagement and online reviews,” *J. Retail. Consum. Serv.*, vol. 41, pp. 48–59, 2018, doi: https://doi.org/10.1016/j.jretconser.2017.11.002.

[3] A. U. Rehman, A. K. Malik, B. Raza, and W. Ali, “A Hybrid CNN-LSTM Model for Improving Accuracy of Movie Reviews Sentiment Analysis,” *Multimed. Tools Appl.*,

vol. 78, no. 18, pp. 26597–26613, 2019, doi: 10.1007/s11042-019-07788-7.

[4] N. Suryana, A. S. H. Basari, and others, "Generate Contextual Insight of Product Review Using Deep LSTM and Word Embedding," in *Journal of Physics: Conference Series*, 2020, p. 12006.

[5] X. Liang and J. Li, "Optimize Statistical Analysis via Propensity Score Matching and Repeated-Measures Analysis of Variance," *Cardiovasc. Interv.*, vol. 16, no. 3, pp. 361–362, 2023.

[6] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998, doi: 10.1109/5.726791.

[7] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *3rd International Conference on Learning Representations (ICLR 2015)*, Computational and Biological Learning Society, 2015, pp. 1–14.

[8] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.

[9] K. Cho *et al.*, "Learning Phrase Representations using {RNN} Encoder{--}Decoder for Statistical Machine Translation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing ({EMNLP})*, Doha, Qatar: Association for Computational Linguistics, Oct. 2014, pp. 1724–1734. doi: 10.3115/v1/D14-1179.

[10] A. Vaswani *et al.*, "Attention is All you Need," in *Advances in Neural Information Processing Systems*, I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds., Curran Associates, Inc., 2017. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

[11] Yu, Z., Guindani, M., Grieco, S.F., Chen, L., Holmes, T.C. and Xu, X., 2022. Beyond t test and ANOVA: applications of mixed-effects models for more rigorous statistical analysis in neuroscience research. *Neuron*, *110*(1), pp.21-35.

[12] Ding, X., Zhang, X., Han, J. and Ding, G., 2022. Scaling up your kernels to 31x31: Revisiting large kernel design in cnns. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 11963-11975).

[13] Zha, W., Liu, Y., Wan, Y., Luo, R., Li, D., Yang, S. and Xu, Y., 2022. Forecasting monthly gas field production based on the CNN-LSTM model. *Energy*, p.124889.

[14] Zhang, W., Li, H., Tang, L., Gu, X., Wang, L., & Wang, L. (2022). Displacement prediction of Jiuxianping landslide using gated recurrent unit (GRU) networks. *Acta Geotechnica*, *17*(4), 1367-1382.

[15] Zhang J, Jiang Y, Wu S, Li X, Luo H, Yin S. Prediction of remaining useful life based on bidirectional gated recurrent unit with temporal self-attention mechanism. Reliability Engineering & System Safety. 2022 May 1;221:108297.

totox

ORIGINALITY REPORT

11%	7%	6%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to University of London External System Student Paper	1%
2	ejournal.unkhair.ac.id Internet Source	1%
3	www.ijert.org Internet Source	1%
4	doctorpenguin.com Internet Source	1%
5	"Web and Big Data", Springer Science and Business Media LLC, 2020 Publication	1%
6	Ambuj Mehrish, Navonil Majumder, Rishabh Bharadwaj, Rada Mihalcea, Soujanya Poria. "A review of deep learning techniques for speech processing", Information Fusion, 2023 Publication	1%
7	www.mikevestil.com Internet Source	1%
	Submitted to City University of Hong Kong	

8	Student Paper	1%
9	Mohammad Shahin, F. Frank Chen, Ali Hosseinzadeh, Neda Zand. "Using Machine Learning and Deep Learning Algorithms for Downtime Minimization in Manufacturing Systems: An Early Failure Detection Diagnostic Service", Research Square Platform LLC, 2023 Publication	1%
10	Submitted to University of Northumbria at Newcastle Student Paper	1%
11	Submitted to University of Portsmouth Student Paper	1%
12	Wenshu Zha, Yuping Liu, Yujin Wan, Ruilan Luo, Daolun Li, Shan Yang, Yanmei Xu. "Forecasting monthly gas field production based on the CNN-LSTM model", Energy, 2022 Publication	1%
13	mdpi-res.com Internet Source	1%
14	Submitted to British University In Dubai Student Paper	1%
15	Ruihao Cao, Guanjun Liu, Yu Xie, Changjun Jiang. "Two-Level Attention Model of	1%

Representation Learning for Fraud Detection", IEEE Transactions on Computational Social Systems, 2021
Publication

16	webthesis.biblio.polito.it Internet Source	1 %

Exclude quotes	On	Exclude matches	< 1%
Exclude bibliography	On		

totox