GENERATIVE OF ORIGIN MODEL DISTRIBUTION MASKED WITH EMOTIONS AND TOPICS DISTRIBUTION IN HYBRID METHOD

Anonymous authors
Paper under double-blind review

ABSTRACT

There is a vast amount of data in the form of natural signals in the world, and difficult expression processors are required to analyze such data. Traditional embedding methods are susceptible to generalization failure. In this study, we developed a classification model that creates and approximates an origin hypothesis model using limited emotions and topics. To solve the hypothesis, the proposed model utilizes dynamic learner modules. Using this mechanism, a text-based origin distribution representation learning model was designed. In order to simulate and generalize, we analyzed the experimental evaluation results via various natural language data sets and measured the corresponding performance. Thus, we demonstrated that the machine achieves the classification task more effectively by integrating learning distribution and multiple learning methods.

1 INTRODUCTION

1.1 EXISTING CHALLENGE

In the problem of classifying videos, images, audio, and text information, establishing a connection with the network is critical to express data. Depending on the internal mechanism of the model, there are several approaches to establish the connection. For example, there are improved versions of convolutional neural networks (CNNs) used with regions of interest exhibiting enhanced accuracy for object detection (Xie et al. (2021);Qiao et al. (2021)). Additionally, merging models have also been proposed to accomplish multiple tasks by using big data sets. Examples in this regard include long-short term memory, CNNs, and support vector machines (He et al. (2021);Sharma et al. (2021)). Deep neural networks have also been proposed as a solution to nonlinear problems, along with generative adversarial nets using appropriate information injection of lost data. Internally, these have been the challenges pertaining to effectively expressing learning. In addition, attempts have been made to establish language learning using bidirectional encoder representations from transformers without a predictive sample; however, they have not worked well in certain fields. An enhanced version has also been proposed for this purpose. For instance, a knowledge-based hybrid model (Wang et al. (2021)) has also been used in various applications such as sentiment analysis (Obied et al. (2021)).

1.2 THE CHALLENGE

Combining Distributions and Approximation in Learning. In this study, both supervised learning and unsupervised learning, having limitations in labeling large amounts of data, were utilized to complement each other. Moreover, we hypothesized from the topics and emotions from the original data to construct the generative initialization module. The intention was to break out of the biased distribution and shape it into a normal distribution. To this end, we employed three learning modules with effectively selectable selectors. First, we attempted to solve the scarcity more robustly to correct the hidden distribution via a simple and predictable recursive learning method. Classification performance was thus demonstrated. Main contributions of this study In this study, marginal random distribution was used considering the overall sparsity and appropriate weight bias. This is the first such attempt, and it is expected that the model will effectively reduce the memory requirement in the future. The contributions are summarized as follows:

- A novel method to solve the scarcity that is different from the traditional method was developed.
- An integrated structure with multiple learning methods that combine the advantages of hybrid supervised learning and unsupervised learning was realized.
- A mechanism that complements each other via modal analysis in text analysis was devised.
- Experiments with the new structure reveal that the attained classification performance was higher than the existing classification performance, and the previously known neural.

2 BACKGROUND

In this section, we will look at existing technology trends for information presentation.

2.1 INFORMATION SELECTION

The weights were selected after the model was trained independently on the COVID-19 document. Several bagging methods were adopted, and when the bagging parameter was between 2 and 4, the Meta-Ensemble method enabled deep expression and exhibited the best performance Rakotoson et al. (2021). It performed sentiment classification using Glove, information gain, wrapper-based, and evolutionary algorithms. Using the speech and SST data of the president of the World Health Organization related to COVID-19, the evaluation results exhibited good performance in terms of time, operation, and accuracy. Specifically, the logistic regression model is 0.845 in SST, which is an increase of approximately 0.08 as compared with the existing technology. Additionally, in the data related to COVID-19, this value was 0.864, an increase of approximately 0.01 compared with the existing technology. By function, it was reduced by approximately 0.78 (Deniz et al. (2021)). It is a classification of emotions to facilitate effective feedback between students and professors. For this purpose, machine learning and deep learning technologies were employed. The performance of the proposed model was evaluated using the data sets collected from Coursera and MOOC and was found to be 99.43% in RF. Furthermore, approximately 99% of the performance was achieved for positive, negative, and neutral cases (Edalati et al. (2021)).

2.2 HEURISTIC SYMBOLIC

It is a type of forensics, dividing two categories in the process of author verification, and devised POsNoise technology that effectively captures the subject if a bias problem occurs when verification is performed via explicit and implicit definitions. Accordingly, only the subject text in the author's writing style is tagged. Compared to the conventional AV method, there is an improvement in performance. However, issues pertaining to English and time and completeness should be considered in the future (Halvani & Graner (2021)). Papaloukas et al. (2021) introduces 47,000 legal texts in Greek. Metadata was pre-processed and transformed using the Python library, and heuristics and regular expression patterns were used to handle many typos and whitespace. In addition, evaluation was conducted using various deep learning methods. BERT performed best. As a result of evaluating precision, recall, and f1 based on frequency in 2285 classes, 83.0, 85.5, and 84.2 were recorded, respectively.

2.3 DEEP ARCHITECTURE

It proposes and describes various deep learning-based methods to detect malware in the security area. Popular detection techniques include L-BFGS, GAN, DL models, and Hot/Cold. Each model has a different area where it excels. Generally, the optimal number of neurons in each component constituting deep learning is described in terms of the size of the data. As a hybrid architecture, DN reduces the dimensionality before the application of classifiers, making it suitable for large and complex data. Meanwhile, conventional classifiers are suitable for small amounts of data. Dropout is often used to prevent overtraining of the data set. However, in this study, it was found that the use of deep learning is not necessary in any domain. In other words, it is a little simpler to explain, and other models may perform better (Mahdavifar & Ghorbani (2019)).

3 GENERATIVE OF ORIGIN MODEL DISTRIBUTION MASKED WITH EMOTIONS AND TOPICS DISTRIBUTION IN HYBRID

In this section, we introduce which extant model is adopted by the proposed model, its structure, which model generates training, and what data set it is applied to.

3.1 MOTIVATION

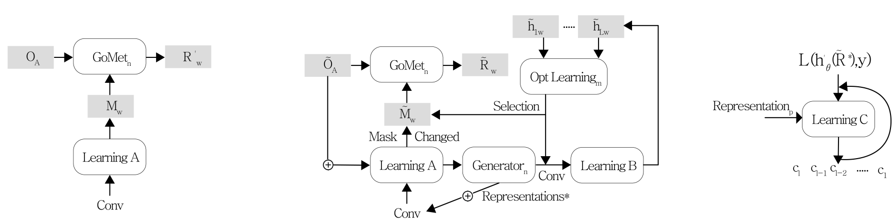
Several existing models have already used the adversarial learning method to learn the distribution by referring to information that hints at effectively presenting the learning structure for the missing data, whereas some models have utilized topic analysis (Yoon et al. (2018);Lee et al. (2021);Park et al. (2022a)). Previously, methods that complement the traditional weights, such as the following have been widely used (Tian & Tong (2010)).

$$IDF_i = \log_2\left(\frac{N}{n_j}\right) + 1 = \log_2(N) - \log_2(n_j) + 1 \tag{1}$$

In this paper, we propose an origin guessing model that uses a limit value selection window, focusing on the multiple extraction method for representing the information of existing natural language and the learning structure of deep learning. This is one of the newest methods that has not been proposed as a complementary method of learning using random distribution with topic and sentiment analysis. For information learning, we analyzed a malicious spam dataset that is commonly encountered and articles in the field of reviews and healthcare that contain user sentiment.

3.2 REPRESENTATION DISTRIBUTION SELECTOR VIA MULTI-LEARNING

The following illustrates the overall schematic diagram.



(a) GoMet Architecture (b) Generator of Representation (c) Class Estimation

Figure 1: (a) Structural diagram of the generator of the original model with emotion and topic in mask, depicting initialization. The gray-brown box represents the variable $\in R^k$. White boxes indicate crucial modules for each action. The direction of the arrow symbolizes the direction and content of argument information transfer. (b) m as expression constructor X, Hypothesis h_0 is set for the matrix and h is implemented via the learner. (C) illustrates class classification by predicting and learning the expression via the estimator h.

3.3 DEFINITION 1

The proposed origin creation model comprises several modules. First, in the input module, the original matrix for the informal data X is received as input. It is referred to as O(m × n). Second, the generative of original model (GOM) module is one that contains the core algorithm that analyzes the input and provides an output. Here, R represents the solution via data analysis and learning. The variables and modules that support the core algorithms are considered in detail. Accordingly, the heuristic variable is the element h_1, and the error of the hypothesis is obtained using the mean square error, M represents the mask variable, and w represents the corresponding weight. Moreover, Conv represents a function variable to be converted and enables convolution operation. During conversion, the value generated by the generator is fed back to be used for learning after embedding.

3.4 DEFINITION 2

The following learning module is composed of three types according to the learning support method and the selection method. It comprises Learning A, Learning B, and Learning C stages. Following that is the module representing the output layer. The expression of Learning A is as follows:

$$\mathrm{L_a}\left(\mathrm{R_{mn}}, \tilde{\mathrm{h}}\left(\mathrm{E_F}\right), \mathrm{y}\right) \tag{2}$$

In the original A matrix, a random matrix is generated by randomly masking topic A generation, data analysis, and Sentiment A generation. Subsequently, a transform operation is performed to approximate the masked TA, masked data, and masked SA and O. This is again used for error calculation for selecting learning for dimensional learning and distribution learning, and also for selecting by performing error calculation for mask selectors. The expression as a mask and learner selection function is as follows:

$$\mathrm{O}_{\text{Selector}}\left(\tilde{\mathrm{h}}\left(\mathrm{E_F}\right), \mathrm{y}\right) \tag{3}$$

where E represents the original embedding before receiving the feed and implies the combination of S_A, T_A, and D_x for learning. E_F denotes the process of vectorizing and loading a masked random distribution for actual operation after its generation. Moreover, E_A represents a transition from O. The equation for Learning B is as follows:

$$\tilde{\mathrm{L}}_{\mathrm{b}}\left(\hat{\mathrm{h}}\left(\mathrm{R}^*\right), \mathrm{y}'\right) \tag{4}$$

As a result of the first learning, the representations generated by the n_{th} generator and the ones passed through the selector in optimization learning are computed. This is calculated as mean square error or alternating least square, before proceeding to convolution and affects later. Learning C obtains h_θ for optimal R and Y_c by reasoning. $Estimator_l$ is estimated after receiving the best representation and starting learning about theta. In total, 100 iterations were performed to compute and estimate the optimal class. While learning close to the original O, representation learning was performed based on the extracted probability distribution, while making it approximately normal. Classify the class following iterative training using the estimator using the mean square error. When performing a convolution operation, the A matrices participate in learning and learn this with a random distribution.

$$\text{Conv}_{\,i} = \begin{cases} E_F'(T, S, D) \oplus \text{ Generator }_i \\ R_A\left(T', S', D'\right) \oplus E_A\left(O'\right) \end{cases} \tag{5}$$

Hybrid analysis and learning: The following figure illustrates initialization that expresses data distribution in supervised learning and unsupervised learning via topic and emotion clustering analysis on original data.

Topic A indicates that the topic is analyzed when the original data are received, and the topic extraction method used is as follows (Jelodar et al. (2019)).

$$p(\theta, \mathbf{z}, \mathbf{w} \mid \alpha, \beta) = p(\theta \mid \alpha) \prod_{n=1}^{N} p(z \mid \theta) p\left(w \mid z_n, \beta\right) \tag{6}$$

For example, in the case of malicious spam documents, they are in the form of supervised learning. Thus, the semantic makes it associatively constant. Data k is a module that analyzes data for basic statistics. Additionally, Sentiment A generates emotion clustering in conjunction with the original data to form unsupervised learning S_a = E(a) = S_{w1}, S_{w2}, S_{w3}, S_{w4},...,S_{wn} (Park et al. (2021)). For instance, it is assumed that words that affect user emotions are clustered among similar classes and then converted into vectors. Mask arbitrarily partially masks the analyzed topic, data, and emotion matrix to cover it. Select one of the analysis modules for O (j=1,2,3..).

There are two types of learning selectors. It is important to learn how to choose and where to apply the mask. First, M_j(h, L(w)) is utilized to determine how and how (row units) and how many (0 and 1) are there to hide. Vectorized D represents a single document in the case of a malicious document, and T represents a list of each grouping subject. S represents each spam, and represents a value weighted according to irony. To compensate for the large loss, L was calculated for the weak part in row units, and the corresponding distribution outlier was reduced by estimation. The matrices are masked before randomization and convolution. Multi-Mask and Generator: The mask technique has

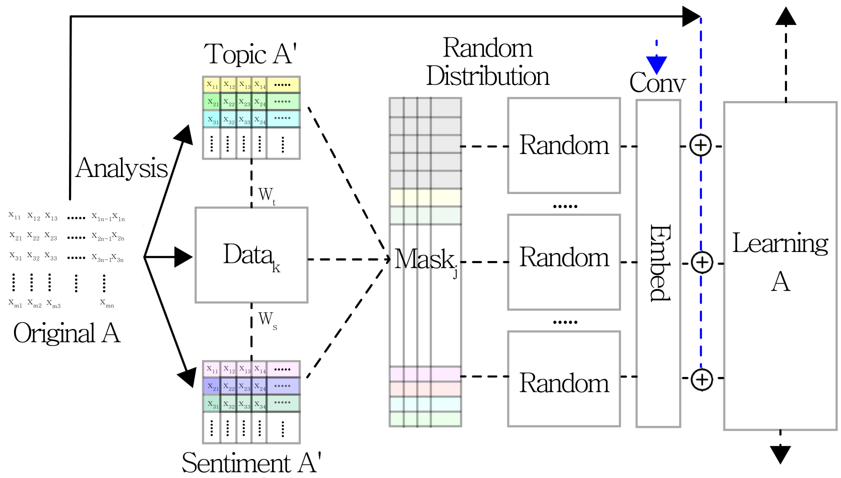


Figure 2: As data initialization analysis, topic analysis A and sentiment analysis A for the original matrix m × n, and data embedding and transformation into random masks and learning mechanisms are learned via the convolution operation.

the meaning of a mask weight for w and undergoes learning in Learning A and Learning Optimization modules. To vectorize the original and masked topic A, sentiment A, and Data A, a random distribution is created and vectorized. The loss in the process of learning the difference is calculated and the finally generated expression is calculated. This is called Generator 1 (repeated n times). Following Learning A module and optimization learning, the mask is selected by combining it with the selection information. It is related to determining the loss when training and deciding where to embed the mask.

3.5 LEARNING+M

Moreover, there is a type of learning that performs optimization by mapping h:=$h^{\sim}$. The learning module selects effective information by decomposition, approximation distribution, and multi-learning, and passes the information to the route connected to it to update the mask (refer to Fig. 1 (b)). For instance, in optimization learning, dimensional decomposition learning and distribution learning are performed depending on whether there are m_{th} types and tag = 0 or 1. To optimize the transformed matrix, alternating square, stochastic gradient descent, and mean square error are used as m pieces of error information received between the original matrix and transformed matrix.

3.6 ALGORITHM OF THE PROPOSED MODEL+M

We initialized the generator of the original model using the input of dataset X. The parameters were as follows: sentiment S_a, topic Ta, limit value k, number of topics n_T, decomposition order r, antonym a, synonym s, and neutral n. We analyzed the clustering technique Gibbs sampling and latent dichstra topic A on the original data A. We analyzed sentiment A based on the weight value of the clustering input O into synonyms and antonyms. Based on the probability statistics embedding technique, w_s and w_t were exchanged and interlocked to express the data. Random distribution was performed for each mask. After randomization was performed on each analysis, it was converted to a convolution operation, then embedding and training A were performed. The training was performed on randomly distributed R, T, D, and O (masked and randomized) and convolutional feeds (training and mask selectors). There were two types of operation for obtaining the value: using the original embedding () and using the fed embedding (), which is fed while repeating the training. To optimize, we selected the type of optimization based on the loss between the existing hypothesis and the correct answer and used a selector. Subsequently, the representation generated in the generator was

Algorithm 1: Algorithm of the proposed model

Input: X$\{x_1, x_2, x_3, ...x_n\}$
Initialization:
1. parameter $\{S_a, T_a, k, n_T, r, a, s, n\}$
2. Preprocessing

Procedure:
for $m, n \leftarrow 0$ **to** $k-1$ **do**
    $W_T \leftarrow Gibbssampling\ using\ LDA\ and\ O_A$
    $W_S \leftarrow sentimentdataanalysisfromO_A$
    $D_k \leftarrow statisticalinformationextraction$
    T'_A, D', S'_A
  for *each in* T'_A, D', S'_A **do**
      $M_T \leftarrow threshold\ calculation\ and\ randomly$
      $D" \leftarrow threshold\ calculation\ and\ randomly$
      $M'_A \leftarrow threshold\ calculation\ and\ randomly$
  end for
  Randomly Distribution
    $Embedding()_n \leftarrow output\ using\ Eq.\ (5)$
    $Generation_n \leftarrow output\ using\ learningA()$
    Distribution and optimization in learning+m
    $Selector_m \leftarrow output\ using\ Eq.\ (4)$
    $Selector_l \leftarrow output\ using\ learningA()$
    $Selector_m \leftarrow output\ using\ learningA()$
    Loss calculation   Feed and combined computation
    $S_{w'_1}, S_{w'_2}, ..., S_{w'_n}$
    $T_{w'_1}, T_{w'_2}, ..., T_{w'_n}$
    Loss calculation using MSE
    $Mask_j for\ random\ distribution$
    $Update_n$
end for
  Generation of estimation
for $c \leftarrow 0$ **to** $l-1$ **do**
  Repeat
    2. Find $h_\theta \leftarrow\ learningc()$
    3. Training
    4. Solve
    5. Until convergence
end for

re-calculated after combined computation and was combined with conversion module again. It is a structure that goes into the convolution of O with a random distribution generated from T', S', and D'. Thus, it strengthens the embedding convolution. The estimator was estimated after receiving the best representation and learning about theta. In total, 100 iterations were performed to compute and estimate the optimal class. It was trained approximately 100 times or more and learned until convergence. Therefore, the default learning rate was set to 0.01. While learning close to the original O, representation learning was performed based on the extracted probability distribution to make it approximately normal. After inputting the optimally expressed matrix R and y correct values and a hypothetical h model representing L = Loss, a classifier was generated (c_1, c_2,...), and the class was then classified. Similar to traditional machine learning methods, it performed better than SVM, DT, and KNN.

Table 1: The following table shows the accuracy. A comparison was performed with existing models that achieved the best performance based on traditional methods for three different types of data, such as security and healthcare.

ACC	D=1	D=2	D=3
B=1	0.876	0.978	0.837
B=2	0.873	0.981	0.849
B=3	0.505	0.944	0.843
SM	0.878	0.982	0.873

4 EXPERIMENTS AND EVALUATION

4.1 EXPERIMENTAL SETUP

We used a ratio of 0.5 on the D1-popcorn dataset (Sadeghian & Sharafat (2015);Mhatre et al. (2017);Sholanbayev (2016);Qiao et al. (2021)), the D2-spam dataset (Hidalgo et al. (2012);Sartaj & Mollah (2021);Sousa et al. (2021);Nandhini & KS (2020)), D3-pubmed article dataset (Abdirizak et al. (2019);Harvey et al. (2019);Park et al. (2022b)). GeForce RTX to implement training methods for 2060 GPU, and Windows 10 Home 64-bit wi32 core Intel(R) Core(TM) i9-9980HK CPU @ 2.40GHz was used. A classification evaluation was performed using Python, TensorFlow, and Keras. We learned approximately 100 times and set the same basic learning rate to 0.01 based on S_a, T_a, k, and n_T, which represent emotions, topics, threshold values, and the number of topics, respectively.

4.2 COMPARISON WITH EXISTING WORK

We performed comparative evaluations of four versions of the architecture of the underlying traditional model. Base 1 utilized the subject model and performed a decomposition-learning task (Park et al., 2021a). Base 2 performed an approximation of the subject distribution task (Park et al., 2022b). Base 3 performed weighted decomposition learning for emotion clustering (Park et al., 2021b). A table evaluating the model (SM) proposed in this study is shown in Table 1.

4.3 ANALYSIS

Existing studies have proven that the models achieved higher performance than traditional statistical embedding, and among them, those that had the highest performance were compared. For D1, B = 1 achieved approximately 87.6%, B = 2 achieved 87.3%, and the next basic model achieved 50.5%. However, SM achieved the highest performance (87.8%). The area under the curve performance was also the highest at 0.95. Among the existing models, Base 1 achieved the best performance (87.6%), followed by B = 2. B = 3 performed poorly. For D2, we sequentially achieved a performance of approximately 97.8% at B = 1, 98.1% at B = 2, and 94.4% at B = 3. However, in SM, 98.2% showed that it was similar to the existing deep learning performance, which indicates that approximating it using multiple learning using random masked emotions and subjects for the optimal R expression method is highly effective. The area under the curve was the highest at 0.992. In addition, for D3, we achieved 87.3% by enabling multi-learning through optimal representation and estimator despite the small dataset. A previous model showed a high performance in the order B = 2, B = 3 and B = 1. This was possible because it was improved through the error function learning method and multi-analysis in the study that presented the existing models, and it has achieved great success in the past.

Figure 3 shows how accuracy improves with a pure number of times when training D2 according to dynamic parameters (decomposition order, topic parameter, synonym parameter, antonym parameter not conducive to malicious documents, and medium parameter not affecting spam) indicates the extent to which it continued to increase and then rose to approximately 0.982 and showed no further increase.

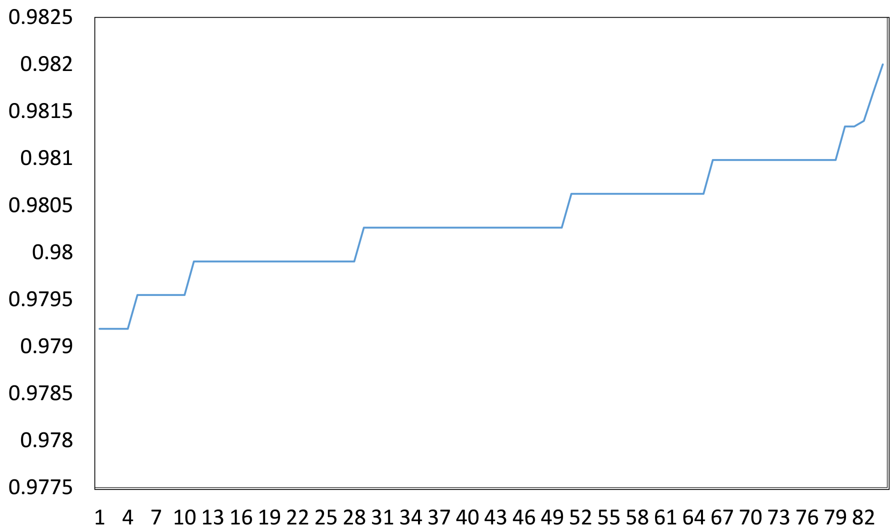


Figure 3: The improvement in the accuracy based on the training progress of the proposed GoMet.

5 CONCLUSION

In this paper, a generative original model distribution masked with emotions and topic distribution in a hybrid model is presented. This is a hybrid model of supervised and unsupervised learning, and expression learning was performed to effectively embed and classify sparse information in the text. Specifically, multi-learning was conducted with the problem of data scarcity, multi-analysis of emotions and themes of the original data, and masked random distribution. The advantage of this approach is that multiple analyses complement each other by tuning multiple parameters. It was found that this not only performed well in expression learning on more than 10,000 existing data but also exhibited more effective performance on small amounts of data. As a result of performing a classification evaluation of the proposed GoMet model, it increased by approximately 4%. The proposed model is a type of multi-task modal that has been in the spotlight these days and is studied and developed to support various expression creation and classification tasks in the future.

REFERENCES

Fatima Abdirizak, Rayleen Lewis, and Gerardo Chowell. Evaluating the potential impact of targeted vaccination strategies against severe acute respiratory syndrome coronavirus (sars-cov) and middle east respiratory syndrome coronavirus (mers-cov) outbreaks in the healthcare setting. *Theoretical Biology and Medical Modelling*, 16(1):1–8, 2019.

Ayça Deniz, Merih Angin, and Pelin Angin. Evolutionary multiobjective feature selection for sentiment analysis. *IEEE Access*, 9:142982–142996, 2021.

Maryam Edalati, Ali Shariq Imran, Zenun Kastrati, and Sher Muhammad Daudpota. The potential of machine learning algorithms for sentiment classification of students' feedback on mooc. In *Proceedings of SAI Intelligent Systems Conference*, pp. 11–22. Springer, 2021.

Oren Halvani and Lukas Graner. Posnoise: An effective countermeasure against topic biases in authorship analysis. In *The 16th International Conference on Availability, Reliability and Security*, pp. 1–12, 2021.

Ruth Harvey, Giada Mattiuzzo, Mark Hassall, Andrea Sieberg, Marcel A Müller, Christian Drosten, Peter Rigsby, Christopher J Oxenford, and study participants. Comparison of serologic assays for middle east respiratory syndrome coronavirus. *Emerging Infectious Diseases*, 25(10):1878–1883, 2019.

Kaijian He, Lei Ji, Chi Wai Don Wu, and Kwok Fai Geoffrey Tso. Using sarima–cnn–lstm approach to forecast daily tourism demand. *Journal of Hospitality and Tourism Management*, 49:25–33, 2021.

José María Gómez Hidalgo, Tiago A Almeida, and Akebo Yamakami. On the validity of a new sms spam collection. In *2012 11th International Conference on Machine Learning and Applications*, volume 2, pp. 240–245. IEEE, 2012.

Hamed Jelodar, Yongli Wang, Chi Yuan, Xia Feng, Xiahui Jiang, Yanchao Li, and Liang Zhao. Latent dirichlet allocation (lda) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, 78(11):15169–15211, 2019.

Junglim Lee, Youngji Kim, Eunju Kwak, and Seungmi Park. A study on research trends for gestational diabetes mellitus and breastfeeding: focusing on text network analysis and topic modeling. *The Journal of Korean Academic Society of Nursing Education*, 27(2):175–185, 2021.

Samaneh Mahdavifar and Ali A Ghorbani. Application of deep learning to cybersecurity: A survey. *Neurocomputing*, 347:149–176, 2019.

Mayuri Mhatre, Dakshata Phondekar, Pranali Kadam, Anushka Chawathe, and Kranti Ghag. Dimensionality reduction for sentiment analysis using pre-processing techniques. In *2017 International Conference on Computing Methodologies and Communication (ICCMC)*, pp. 16–21. IEEE, 2017.

S Nandhini and Jeen Marseline KS. Performance evaluation of machine learning algorithms for email spam detection. In *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)*, pp. 1–4. IEEE, 2020.

Zeinab Obied, Aiman Solyman, Atta Ullah, Ahmed Fat'hAlalim, and Alhag Alsayed. Bert multilingual and capsule network for arabic sentiment analysis. In *2020 International Conference On Computer, Control, Electrical, And Electronics Engineering (ICCCEEE)*, pp. 1–6. IEEE, 2021.

Christos Papaloukas, Ilias Chalkidis, Konstantinos Athinaios, Despina-Athanasia Pantazi, and Manolis Koubarakis. Multi-granular legal topic classification on greek legislation. *arXiv preprint arXiv:2109.15298*, 2021.

W Park, Nawab Muhammad Faseeh Qureshi, and Dong Ryeol Shin. Pseudo nlp joint spam classification technique for big data cluster. *Computers, Materials and Continua*, 71(1):517–535, 2022a.

Woo Hyun Park, Dong Ryeol Shin, and Nawab Muhammad Faseeh Qureshi. Effective emotion recognition technique in nlp task over nonlinear big data cluster. *Wireless Communications and Mobile Computing*, 2021, 2021.

Woo Hyun Park, Isma Farah Siddiqui, Dong Ryeol Shin, and Nawab Muhammad Faseeh Qureshi. Nlp-based subject with emotions joint analytics for epidemic articles. *CMC-COMPUTERS MATERIALS & CONTINUA*, 73(2):2985–3001, 2022b.

Limeng Qiao, Yuxuan Zhao, Zhiyuan Li, Xi Qiu, Jianan Wu, and Chi Zhang. Defrcn: Decoupled faster r-cnn for few-shot object detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 8681–8690, 2021.

Loïc Rakotoson, Charles Letaillieur, Sylvain Massip, and Fréjus Laleye. Bagbert: Bert-based bagging-stacking for multi-topic classification. *arXiv preprint arXiv:2111.05808*, 2021.

Amir Sadeghian and Ali Reza Sharafat. Bag of words meets bags of popcorn. *CS224N Proj*, pp. 4–9, 2015.

Sahil Sartaj and Ayatullah Faruk Mollah. An intelligent system for spam message detection. In *Intelligent Systems*, pp. 387–395. Springer, 2021.

Rajesh Sharma, Akey Sungheetha, et al. An efficient dimension reduction based fusion of cnn and svm model for detection of abnormal incident in video surveillance. *Journal of Soft Computing Paradigm (JSCP)*, 3(02):55–69, 2021.

Uan Sholanbayev. Sentiment analysis on movie reviews, 2016.

Gustavo Sousa, Daniel Carlos Guimarães Pedronette, João Paulo Papa, and Ivan Rizzo Guilherme. Sms spam detection through skip-gram embeddings and shallow networks. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pp. 4193–4201, 2021.

Xia Tian and Wang Tong. An improvement to tf: Term distribution based term weight algorithm. In *2010 Second International Conference on Networks Security, Wireless Communications and Trusted Computing*, volume 1, pp. 252–255. IEEE, 2010.

Bo Wang, Tao Shen, Guodong Long, Tianyi Zhou, Ying Wang, and Yi Chang. Structure-augmented text representation learning for efficient knowledge graph completion. In *Proceedings of the Web Conference 2021*, pp. 1737–1748, 2021.

Xingxing Xie, Gong Cheng, Jiabao Wang, Xiwen Yao, and Junwei Han. Oriented r-cnn for object detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 3520–3529, 2021.

Jinsung Yoon, James Jordon, and Mihaela Schaar. Gain: Missing data imputation using generative adversarial nets. In *International conference on machine learning*, pp. 5689–5698. PMLR, 2018.