Jurnal Sistem Informasi Bisnis 02(2024)

On-line: http://ejournal.undip.ac.id/index.php/jsinbis/article/view/59586

Classification of SWOT Statements Employing BERT Pre-Trained Model Embedding

Husni Thamrin[a*], Dewi Oktafiani[b], Ibrahim Ihsan Rasyid[c], Irfan Miftahul Fauzi[d]

[a] Dept of Informatics, Universitas Muhammadiyah Surakarta
[b] Dept. of Information Technology, STMIK Amikom Surakarta
[c] Dept. of Informatics, Institut Teknologi Bandung
[d] Dept. of Bio-enterpreneurship, Universitas Muhammadiyah Madiun

Submitted: February 2nd, 2024; accepted: March 9th, 2024
DOI: 10.21456/vol14iss2pp143-152

Abstract

SWOT analysis is a highly effective method for organizations to develop strategic planning and gain widespread adoption by various institutions, industries, and businesses. The importance of SWOT analysis lies in its ability to provide a comprehensive assessment of an organization's internal and external factors. Despite its advantages, there are several challenges in its implementation, such as the challenge to identify the four elements of SWOT and to put statements into their correct position as strength, weakness, opportunity, or threat. This study aims to determine the best SWOT statement classification from a combination of using BERT models as feature extraction technique and compare it with traditional method of TF-IDF. The SWOT statement is input to the model to get a vector as a sentence representation. More similar vector representations indicate the closer meaning of the sentences. The similarity is the basis for the classifier to determine whether a sentence falls into the domain S, W, O, or T. We examined two classification algorithms, namely Support Vector Machine (SVM) and Naïve Bayes Classifier (NBC). Data consists of 635 SWOT statements from study programs of a higher education institution. Five combinations of feature extraction techniques and classification algorithms were tested. The study finds that SBERT model embedding in conjunction with support vector machine classification yield the best performance with an accuracy of 0.73 and an F1-score of 0.738. It outperforms the more traditional method of feature extraction of TF-IDF and other combinations using the Naive Bayes Classifier.

Keywords: SWOT analysis; strategic planning; classification; deep learning; model embedding;

1. Introduction

SWOT analysis is a strategic planning tool that helps organizations assess their current position and plan for the future (Achmad et al., 2021). The acronym represents Strengths, Weaknesses, Opportunities, and Threats. The importance of SWOT analysis lies in its ability to provide a comprehensive overview of an organization's internal and external factors (Alalie et al., 2019; Granulo and Tanović, 2020; Teimoori and Alinezhad, 2019). By identifying its strengths and weaknesses, a business is able to capitalize on its advantages and resolve its shortcomings. Simultaneously, by recognizing opportunities and threats in the external environment, it assists in making informed decisions and developing strategies that align with the organization's goals and mitigate potential risks. The SWOT analysis facilitates improved decision-making, enhances strategic planning, and supports businesses to adapt to a dynamic business environment. It's a versatile and valuable tool for businesses of all sizes and across various industries, including business or education institution (Achmad et al., 2021; Fafurida et al, 2020; Fuadi et al, 2019). The Indonesian National Accreditation Board for Higher Education (Badan Akreditasi Nasional Perguruan Tinggi) encourages the adoption of SWOT analysis in self-evaluation documents (Aurachman and Putri, 2020; Thamrin and Pamungkas, 2017). Such documents should have SWOT statements that align with the written narratives for each criterion within the nine accreditation criteria.

Aside from its advantages, there are difficulties that people may encounter when applying SWOT analysis. The drawbacks include oversimplification, the long list of SWOT statements, and the reluctance to admit weaknesses. Oversimplification is the reduction of complex factors to overly simplistic terms, missing essential nuances and interrelationships. It can lead to superficial insights and ineffective strategic planning, failing to capture the full scope of an organization's situation (Namugenyi et al., 2019; Tweheyo and Mugarura, 2020). A lengthy list of SWOT statements can be overwhelming and reduce the strategic value of the analysis. It distracts from critical and prioritized factors that truly impact the organization. While reluctance to admit weaknesses can impede strategic planning, obstructs growth and risk mitigation, which in turn delays improvement, and long-term success

*) Corresponding author: husni.thamrin@ums.ac.id

(Lohrke, Mazzei et al., 2022; Vlados and Chatzinikolaou, 2019).

Another limitation of using SWOT analysis is the difficulty in identifying the four elements of SWOT. It may result in an incomplete analysis, leaving out essential aspects of the organization's situation, which could potentially lead to biased or incomplete strategies. This difficulty can also impair the process of decision-making, causing inefficiencies in resource allocation. The difficulty arises from several factors such as lack of clarity, subjectivity, and inexperience. The lack of clarity make it hard to distinguish between a strength and an opportunity or between a weakness and a threat. Subjective interpretations from different stakeholders may result in varying perspectives and the lack of understanding may cause confusion in distinguishing between SWOT statements. Inexperienced or hasty analyses can lead to overlooking critical factors, hence complicating the identification process (Taherdoost and Madanchian, 2021; Togayev and Zayniddinova, 2023).

To overcome these obstacles, organizations must ensure comprehensive data, objectivity, and a thorough understanding of their internal and external conditions. Technical tools such as a SWOT analysis template is common and can be very helpful to ensure that all relevant factors are considered (Etokidem et al., 2020; Mateos, 2020). On the other hand, machine learning algorithms are invaluable in reducing subjectivity and extracting valuable insights from vast and scattered information. By applying machine learning techniques, organizations can gain a better understanding of complex data and identify SWOT patterns that can help to make data-driven decisions with greater accuracy (Makmun and Thamrin, 2018).

This study aims to evaluate the performance of machine learning algorithms in identifying textual statements into the four elements of SWOT. The process applies when the organization's team has reached a consensus on the SWOT statements that accurately align with the organization's profile. At this point, a text classification algorithm can be employed to categorize each statement as either a strength, weakness, opportunity, or threat. By automating this categorization process, the algorithm helps accelerate the SWOT analysis, reduce subjectivity and ensure a more objective assessment.

Various research groups have carried out similar investigations. Abu-Alaish et al. (2021) successfully compiled and extracted the most significant expressions from collected SWOT statements analyzed for a university. Their effort was to utilize machine learning through the use of the Weka application, employing methods such as classification, clustering, and association rules. In addition, Akella (2021) attempted to employ machine learning to make decisions by selecting strategies based on SWOT statements. They assigned weights to each statements and made decisions based on linear and quadratic regression. On the other hand, Lima et al., (2021) conducted a study where they developed an application to collect both internal and external organizational conditions, constructed linguistic word definitions, and assess the association of each input variable. The application automated the information collection and SWOT matrix creation by utilizing the fuzzy inference method.

To the extent of our knowledge, there have been no attempt to apply BERT for the classification of SWOT statements or to assist in the formulation of SWOT matrices, although BERT has found a wide application in text classifications such as those conducted by (Bilal and Almazroi, 2023; Imaduddin et al., 2023; Munikar et al., 2019; Prabhu et al., 2021). BERT is a deep learning model built on the transformer architecture and can be applied to large language modeling, similar to ChatGPT. BERT and its derivatives, such as SBERT, outperform Word2Vec or GloVe techniques for many classification tasks because Word2Vec and Glove generate vector representations that disregard context (Kale et al., 2023; Shen and Liu, 2021).

The contribution of this paper is to prove that one of the deep learning models, SBERT, is capable of generating vectors that effectively reflect SWOT assertions. Observations have been carried out on real datasets from an educational institution. However, the SBERT model has been trained with Indonesian sentences not limited to educational context so the results of this research may apply to SWOT analysis for all business lines, not just educational areas.

2. Theoretical Background

2.1. BERT and SBERT

BERT is an acronym for Bidirectional Encoder Representations from Transformers. It is a state-of-the-art natural language processing (NLP) model developed by Google in 2018. It is a deep learning model based on the Transformer architecture, designed for understanding the context and meaning of words in a sentence or text by considering the surrounding words both to the left and the right (Devlin et al., 2018). BERT has achieved significant advancement in various NLP tasks, including text classification, sentiment analysis, named entity recognition, and question answering (González-Carvajal and Garrido-Merchán, 2020; Hao et al., 2019; Wang et al., 2019).

The distinguishing factor of BERT is in its pre-training on a massive volume of textual data, which allows it to learn contextual word representations. As a result, it excels at capturing the nuances and details of language, making it one of the most powerful models for a wide range of natural language tasks. BERT, as a variant of the Transformer model, is characterized by its bidirectional and unsupervised pre-training approach (Devlin et al., 2018). The key

On-line: http://ejournal.undip.ac.id/index.php/jsinbis/article/view/59586

components of BERT's architecture includes a multi-head self-attention mechanism, a transformer encoder stack, multiple embeddings, pre-training and fine-tuning stages, bidirectional context, masked language modeling, and next sentence prediction. BERT model can be used to tackle various NLP applications by adapting the input and output layer. For text classification tasks, each input texts are treated as a "single sentence" where the words are tokenized to generate the respective word embeddings. A process of training involves thousands of steps to tune the neural network parameter, while processing a huge number of masked sentences. The output layer contains the class label and output tokens.

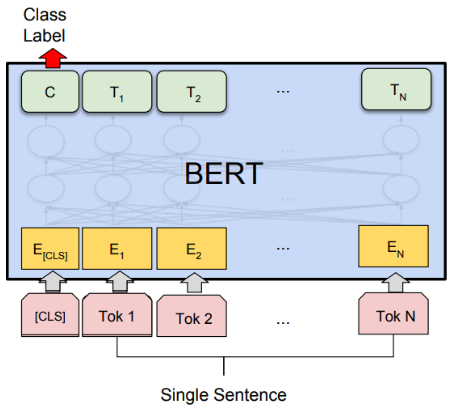

Figure 1. BERT architecture (Devlin et al., 2018)

Scientific enthusiasts have constructed various new models and architectures inspired or based on BERT models (Azizah et al., 2023). To name a few, there are ALBERT, RoBERTa, and DistilBert. This study investigates a variant, called SBERT or Siamese BERT. SBERT uses two identical BERT models during training and execution (see Figure 2.). The outputs of both BERT models u and v with the difference |u – v| are concatenated and processed with a Softmax function classifier.

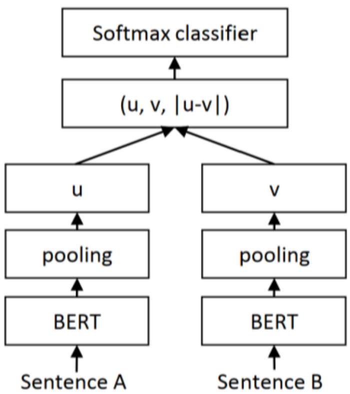

Figure 2. SBERT architecture with classification objective function (Reimers and Gurevych, 2019)

SBERT focuses on learning sentence embeddings that capture semantic similarity between sentences, encouraging the model to pull together representations of similar sentences and push apart representations of dissimilar sentences. It is very good in generating sentence embeddings or representations that capture the semantic meaning of texts (Milios and BehnamGhader, 2022; Reimers and Gurevych, 2019).

Pre-trained SBERT model using database in Bahasa Indonesia has been available. Two works are notable here. One work is by Firqa, (2022) (SBERT-F) who trained SBERT model using 500 MB data and the other is by Denaya, (2023) (SBERT-D) who trained the model using 1.5 GB data.

2.2. *TF-IDF*

TF-IDF (Term Frequency-Inverse Document Frequency) is one of the statistical methods to determine the weight or importance of a word (term) to a document or article in a collection or corpus (Cahyawijaya et al., 2022). The applied weighting technique combines term frequency and inverse document frequency. Term frequency represents the number of occurrences of a term within a document, where the more frequent the term, the higher the weight assigned to it. Meanwhile, inverse document frequency measures how important a term to identify a document. The weight of a term can be calculated as follows.

$$tf_{d,t} = 0.5 + 0.5 \times \frac{f_{d,t}}{\max(f_{d,t})} \quad (1)$$

$$idf_t = \log\left(\frac{D}{df_t}\right) \quad (2)$$

$$W = tf_{d,t} \times idf_t \quad (3)$$

In equations (1) to (3) tf is the term frequency, f is the raw count of a term in a document, max(f) is the highest count of a term in all documents, idf is the inverse document frequency, D is the total number of documents, df is the number of documents containing the term, and W is the TF-IDF. In the equations, subscript d denotes the document d, and subscript t denotes the term t.

2.3. *Support Vector Machine*

SVM (Support Vector Machine) is a machine learning method that employs a linear function hypothesis in a high-dimensional feature space (Cervantes et al., 2020; Pisner and Schnyer, 2020). It employs learning algorithms and is based on optimization theory. SVM algorithm was developed by Broser, Guyon, Vapnik. It was first introduced in 1992 at the Annual Workshop on Computational Learning Theory.

The utilization of the SVM method has several advantages, including generalization, addressing the curve of dimensionality, and feasibility. However, it may be challenging to apply this method to large-scale problems or samples. Figure 3. illustrates the Support Vector Machine model.

On-line: http://ejournal.undip.ac.id/index.php/jsinbis/article/view/59586

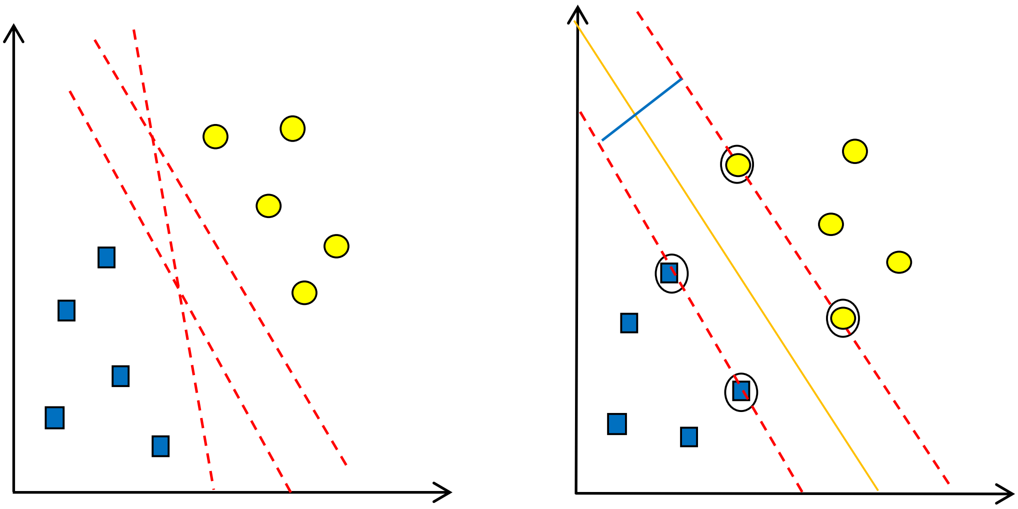

Figure 3. Support vector machine works by defining a hyperplane separating two classes of data

The concept of SVM can be described conveniently as an attempt to find the best hyperplane that serves as a separator between two classes. The data can be represented as x_i, and the labels for each data point are represented as y_i, for i = 1, 2, ... l, where l denotes the total number of data. The support vector machine (SVM) operates under the assumption that the two classes, say 1 and 0, can be perfectly separated by a hyperplane in a d-dimensional space. A hyperplane can be written as the set of points x that satisfies equation (4).

$$\vec{w}.\vec{x} + b = 0 \tag{4}$$

where w is the normal vector the hyperplane. Two parallel hyperplanes can be selected so that they separate the two classes. All data for class 1 meet the inequalities of equation (5).

$$\vec{w}.\vec{x} + b \geq 1 \tag{5}$$

Meanwhile, the data that belong to class 0 satisfy the following equation (6)

$$\vec{w}.\vec{x} + b \leq 0 \tag{6}$$

The distance between the two hyperplanes are called the margin, which needs to be maximized. The margin may be expressed as equation (7).

$$\frac{1}{\|\vec{w}\|} \tag{7}$$

The problem may be approached as quadratic programming problem to find the minimum of equation (8),

$$\min_{w} \tau(w) = \tfrac{1}{2}\|\vec{w}\|^2 \tag{8}$$

considering the constraints expressed in equation (9)

$$y_i(\vec{w}.\vec{x}_i) - 1 \geq 0, \forall i \tag{9}$$

The solution to the problem can be obtained using computational methods, such as the Lagrange multiplier, which is expressed in equation (10).

$$L(\vec{w}, b, \alpha) = \frac{1}{2}\| \vec{w}^2 - \sum_{i=1}^{l} \alpha^i (y^i((\vec{w}.\vec{x} + b) - 1)) \tag{10}$$
$$i = 1,2,....l$$

α_i is the Lagrange multipliers, which can be zero or positive, $\alpha_i \geq 0$. The optimal value for equation (11) and equation (12) may be modified as the problem of maximization which contain only the variable α_i.

$$\sum_{i=1}^{l} \alpha_i - \tfrac{1}{2}\sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j \vec{x}_i \vec{x}_j \tag{11}$$

$$\alpha_i \geq 0 (i = 1,2,...l)$$
$$\sum_{i=1}^{l} \alpha_i y_i = 0 \tag{12}$$

The best model in SVM classification largely depends on specific datasets and evaluation criteria. No single approach guarantees to reach the best model for any features, regardless of using OvA (One-vs-All) or OvO (One-vs-One) strategy. However, there are good practices that give a chance to obtain one. In the training phase, explorations should start from a simple architecture, such as a linear kernel, to tune hyperparameters. Cross validation helps one get a robust solution. Complex kernels exploration becomes necessary when linear kernel is not sufficient. Model training with both OvA and OvO strategy is a good practice while comparing the performance metrics like accuracy, precision, recall, and F1-score. Once a model establishes, its improvement may proceed with analysis of misclassified instances and data patterns.

3. Method

The research method employed in this study involves several stepwise activities as depicted in Figure 4. The activities start from data collection and

On-line: http://ejournal.undip.ac.id/index.php/jsinbis/article/view/59586

pre-processing, document feature extraction, classification, and end with evaluation. This study employs two deep learning pre-trained models to obtain vectors that represent the SWOT statements contained in the text documents and utilize the more traditional TF-IDF algorithm to create such vectors. The vectors, which are sequences of floating-point values, become the basis of language processing. Similarity of two textual statements, for example, can be computed by taking the dot product of two vectors that represent the two corresponding texts.

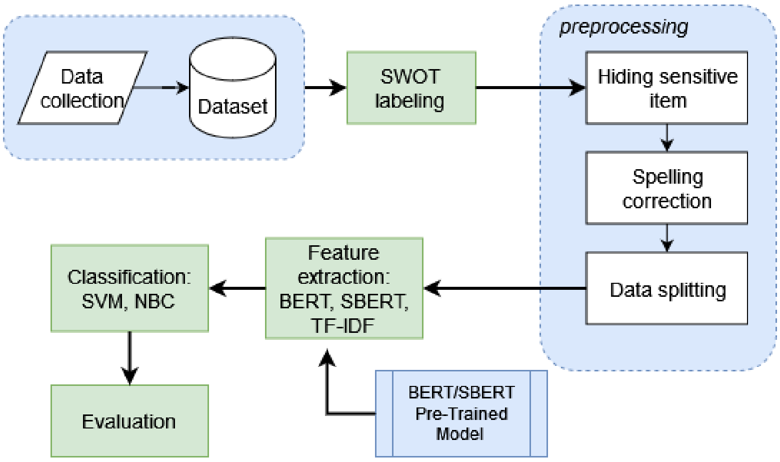

Figure 4. Research flow

3.1. Data collection and pre-processing

Data collection involves obtaining SWOT statements from a higher education institution. Such data is generated when the institution formulates strategic planning and proposes operational activities. The institution employs SWOT analysis to gain insights into factors that influence teaching and learning process, research related activities, and institutional development programs.

The collected textual statements are in Bahasa Indonesia and have previously been labeled with S, W, O, and T, respectively. However, the authors conduct a thorough reassessment of the statements to ensure their accurate placement into S, W, O, and T quadrants. Corrections take place as needed.

In order to protect the privacy of individuals or organizations and comply with ethical standards, sensitive items, particularly those containing personally identifiable information, are concealed by converting named entities into their general holonymy. This modification guarantees the preservation of data anonymity while retaining the meaningful content for analysis purposes.

Furthermore, to enhance the quality of the dataset, manual spelling correction is performed to rectify any typographical errors or inconsistencies in the collected data. This step aims to improve the integrity of data and ensure that the analysis is based on accurate information.

Unlike many typical text analysis approaches, our method does not involve stemming or stop word removal. Stemming, which reduces words to their root form, is omitted to maintain the originality and context of terms. Similarly, stop word removal is not performed to preserve even commonly used words, as they might contain important information in the context of the study.

The dataset is then split into two subsets: a training set comprising 70% of the data and a testing set containing the remaining 30%. All algorithms under investigation utilize identical training and testing set to ensure that they face similar scenarios and tackle similar problems.

3.2. Feature Extraction

Feature extraction, or text vectorization, is the process of converting textual data into their representative numbers. Various methods can be used to perform this task, such as TF-IDF, bag-of-words, n-gram, and word2vec. The aforementioned methods work mostly on word statistics and hence ignore the contextual meaning of words that construct sentences or texts. In this study we investigate the use of deep learning BERT models in the text feature extraction task. BERT models have many components including word and sentence embeddings, which is the dense vector representations for words and sentences. These embeddings possess the capture of semantic relationships among words and sentences; hence their use can expectedly improve the result of natural language processing tasks including text classification.

Below is an example Python codes to get sentence embeddings from a group of sentences. First step is to prepare for the Python modules including transformer and torch. Second step is to make a function to calculate the average vectors from the model output.

```
# !pip install transformers
from transformers import AutoTokenizer,
AutoModel
import torch
```

```
# Mean pooling - take attention mask into
account for correct averaging
def mean_pooling(model_output,
attention_mask):
    token_embeddings = model_output[0]
    input_mask_expanded = attention_mask.\
        unsqueeze(-
1).expand(token_embeddings.size()).float()
    return torch.sum(token_embeddings *
input_mask_expanded, 1) /
torch.clamp(input_mask_expanded.sum(1),
min=1e-9)
```

The third step is to load a pre-trained model. We need to create tokenizer and model object, stating the name of the data as the parameter.

```
# Load pre-trained model
tokenizer =
AutoTokenizer.from_pretrained('cahya/bert-
base-indonesian-522M')
model =
AutoModel.from_pretrained('cahya/bert-base-
indonesian-522M')
```

Considering that the our sentences are in Bahasa, we use BERT model that has been pre-trained using dataset in Bahasa (Azizah et al., 2023). The only pre-trained model that is available at the time of the study is bert-base-indonesian-522M, which was trained using 522 MB textual data.

Table 1. Sample of data

Sample SWOT statements	English translation
Minat masyarakat untuk menyekolahkan putra-putrinya di prodi menunjukkan kecenderungan meningkat	The desire of parents to enroll their children in the study program shows a growing trend
Calon mahasiswa berminat kuliah di program studi karena visi, misi dan tujuan yang spesifik	Prospective students are attracted to the study program because of its distinct vision, mission, and aims.
Model tracer alumni yang dilakukan belum efektif dan efisien	The current model of the alumni tracer study lacks effectiveness and efficiency

For example, we wish to find vectors of three sentences in table 1. The fourth step is to make the list of texts, and calculate their sentence embeddings as follows.

```
sentences = ["Minat masyarakat untuk
menyekolahkan putra-putrinya di prodi
menunjukkan kecenderungan meningkat", "Calon
mahasiswa berminat kuliah di program studi
karena visi, misi dan tujuan yang spesifik",
"Model tracer alumni yang dilakukan belum
efektif dan efisien"]

# Tokenize sentences and compute token
embeddings
encoded_input = tokenizer(sentences,
padding=True, truncation=True,
return_tensors='pt')
with torch.no_grad():
    model_output = model(**encoded_input)
embeddings = mean_pooling(model_output,
encoded_input['attention_mask'])
```

The variable `embeddings` in the code above stores a list of 3 vectors that represent the three input sentences, respectively. The vectors are the sentence feature that can be fed into classification algorithms for the task of text classification.

Three deep learning models are tested in this study. First model is cahya/bert-base-Indonesian-522M, shich is a BERT model pre-trained using 522MB data in Bahasa Indonesia. Second is firqaaa/indo-sentence-bert-base, which is an SBERT model pre-trained using 500MB data by Firqa, (2022) (SBERT-F). Third is denaya/indoSBERT-large, which is an SBERT model pre-trained using 1.5 GB data by Denaya, (2023) (SBERT-D). BERT and SBERT-F produces sentence embeddings of 768 floating-point values in size, while SBERT-D results in 384 dimensional vectors. As a comparison, we complete the study with classification using the more traditional technique of TF-IDF for feature extraction. TF-IDF feature represents word frequency and word importance within the document corpus and its dimension varies depending on the text under processing. For our SWOT data, TF-IDF produces vectors with the size of 1293 floating-point values.

3.3. Classification

Classification in this context is the technique to assign S, W, O, T labels to open ended texts. The common procedure in examining classification performance is to firstly train the classifier with labeled training data and then let the classifier label texts in the testing data into predefined categories. We use SVM algorithm as the main classifier, and also examine the use of NBC to compare the performance.

In total, we investigate five combinations of text feature extraction and classification techniques, as follows:

1) BERT – SVM
2) SBERT-F – SVM
3) SBERT-D – SVM
4) TF-IDF – SVM
5) TF-IDF – NBC

3.4. Evaluation

We use accuracy and F1 score as the metric to compare the performance of each model/algorithm. Accuracy describes how close a given set of measurements (observations or readings) are to their true value. In this context, it describes how many statements are classified into their correct SWOT label compared to the number of all statements. F1-score on the other hand represents the harmonic mean of precision and recall, which takes into account of imbalances in the numbers of data among categories.

4. Results and Discussion

4.4. Result

We tested 635 SWOT statements which have been manually labeled. They are split into training data consisting of 444 statements (70%) and testing data

On-line: http://ejournal.undip.ac.id/index.php/jsinbis/article/view/59586

containing 191 statements (30%). The data are not equally distributed among SWOT categories, but the imbalance is only mild. The majority class is the Strength with 217 data or 34% support, while the minority class is the Threat with 109 data or 17% support. Table 2. display the data distribution among classes.

Table 2. Data distribution among categories

Label	Support	Percent
S	217	34
W	188	30
O	121	19
T	109	17

Classification using SVM algorithm on BERT model embeddings results in a confusion matrix as depicted in Figure 5. In most cases, the algorithm are able to predict the statements into the correct categories. The accuracy is at 0.58, which is fairly accurate. Table 3 displays the precision, recall and F1-score for each category. F1-Score varies between 0.45 and 0.64, and the weighted average, which takes into account the number of supports, is 0.579.

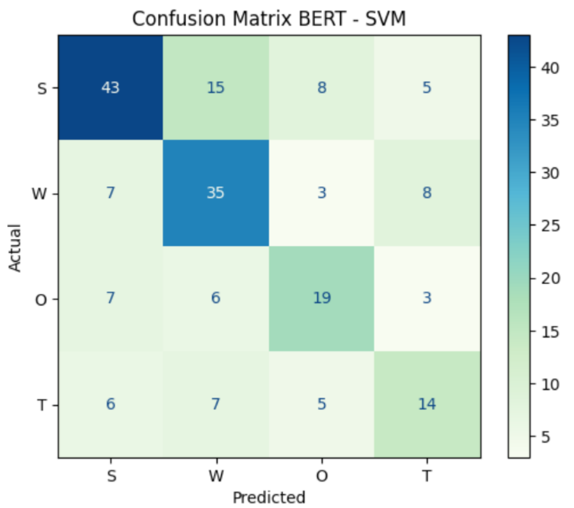

Figure 5. Confusion matrix of classification using BERT feature extraction and support vector machine

Table 3. Performance metrics for classification using BERT feature extraction and support vector machine

Label	Precision	Recall	F1-Score	Support
S	0.68	0.61	0.64	71
W	0.56	0.66	0.60	53
O	0.54	0.54	0.54	35
T	0.47	0.44	0.45	32

Classification on the same data using different algorithms and feature extractions provide different accuracy and F1-score. Table 4. displays values of the metric for various combination of classification algorithms and feature extractions. The table suggests that SVM algorithm on SBERT model embeddings that was pre-trained by Denaya achieve the highest classification accuracy and F1-score. On the other hand, the least accuracy is obtained when the classification is conducted using SVM algorithm on BERT model embeddings.

Table 4. Performance of various combinations of feature extraction and classification algorithms

Algorithm	Accuracy	F1-Score
BERT – SVM	0.58	0.579
SBERT-F – SVM	0.69	0.695
SBERT-D – SVM	0.73	0.738
TF-IDF – SVM	0.66	0.666
TF-IDF – NBC	0.64	0.608

Figure 6. depicts confusion matrix produced by the best observed combination of classification algorithm and feature extraction, which uses SVM for classification and SBERT-D model embeddings as feature extraction. By comparing it with confusion matrix in Figure 5., we can obviously see that the best model can predict more data accurately whereby higher values are on the diagonal cells. For example, the best model correctly predicts 51 statements to be in class S among 71 statements. In contrast, combination of BERT-SVM can only predict 43 statements correctly (see Figure 5.).

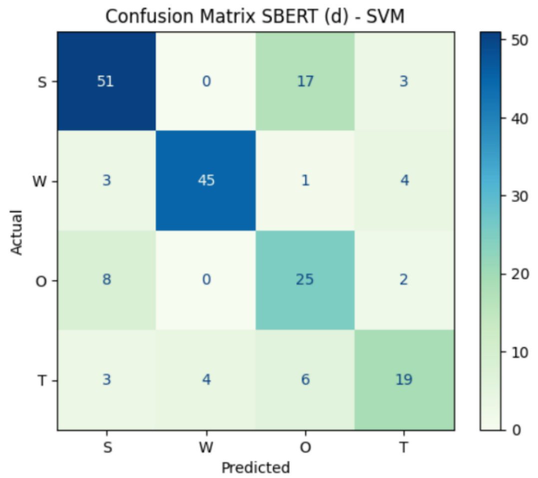

Figure 6. Confusion matrix of classification using SBERT-D feature extraction and support vector machine

4.5. *Discussion*

In this section, we discuss the results obtained from our experiment on classification using various combination of feature extraction and classification techniques. Our findings reveal that SBERT stands out as the most effective feature extraction method, achieving an accuracy of 0.73 and an F1-Score of 0.738. This underscores the significance of contextual sentence embeddings to enhance the performance of text classification tasks especially if the data is at sentence level.

The label SBERT-D in Table 4 denotes denaya/indoSBERT-large, which is an SBERT model that is pre-trained by Denaya, (2023) using a database

Copyright ©2024, JSINBIS, p-ISSN: 2502-2377, e-ISSN: 2088-3587
On-line: http://ejournal.undip.ac.id/index.php/jsinbis/article/view/59586

of Bahasa Indonesia of approximately 1.5 GB in size. While label SBERT-F denotes firqaa/indo-sentence-bert-base, which is an SBERT model pre-trained by Firqa, (2022) using a database at about 500 MB in size. The classification using later SBERT model is less accurate than using the former. This fact suggests that a larger pre-training database improves the accuracy of SBERT vector representation of a sentence, which leads to an overall better classification or NLP task performance.

This study shows that SBERT models outperform BERT model in the classification task of SWOT statements. The fact applies for SBERT-F despite the fact that SBERT-F and BERT were pre-trained using approximately the same dataset size. BERT is designed for contextualized word embeddings and a wide range of NLP tasks, whereas SBERT produces semantically meaningful sentence embeddings to handle tasks involving semantic similarity between sentences. SWOT statements are mostly in the form of short sentences or even phrases so SBERT works better against them rather than BERT.

The results indicate that TF-IDF feature extraction supports SWOT classification surprisingly better than BERT model embeddings. The results contrast the finding by several other investigators that state BERT model outperform other traditional feature extraction methods including TF-IDF (Gomes et al., 2023; González-Carvajal and Garrido-Merchán, 2020). Apparently, the domain of SWOT classification favors TF-IDF over BERT. This outcome might be attributed to the fact that the BERT model was pre-trained on a relatively small database of 522 MB. The small number of dataset may limit the ability of the model to capture the specificity of the investigated domain. The fact highlights the crucial role of pretraining data size in the performance of BERT-based models (Pérez-Mayos et al., 2021). In fact, the previous paragraph has shown that SBERT model with larger pre-trained dataset performs better.

We find that classification using the Support Vector Machine (SVM) exhibits superior performance compared to the Naive Bayes Classifier. SVM's ability to capture complex decision boundaries and handle non-linear data distribution within our dataset contributes to its enhanced classification accuracy, reinforcing its suitability for similar classification tasks.

We have repeated our experiments using several data split. Admittedly, the experimental results show variations but in overall the trend is similar. This highlights the importance of the size and variation of training data to improve the classification performance.

We have also made an additional investigation to employ a majority vote among the five classification methods to decide the class of a statement. The investigation demonstrates that employing a majority vote does not yield an improvement in the result, where accuracy attains a value of 0.7 and weighted average of F1-Score is at a level of 0.731. Table 5 displays the precision, recall and F1-Score for each SWOT category. This suggests that the individual models perform adequately on their own and the fusion of their works does not significantly enhance the overall performance.

Table 5. Performance of majority vote

Label	Precision	Recall	F1-Score	Support
S	0.78	0.69	0.73	71
W	0.82	0.87	0.84	53
O	0.71	0.63	0.67	35
T	0.80	0.50	0.62	32

We believe that the crucial point in the SWOT statement classification process is determining the domain of the sentence as external and internal. Determining the domain is essential to decide whether a statement falls into SW or OT categories. Distinguishing positive and negative tone of a statement affects the decision, but deep learning algorithms have shown a reputation in sentiment analysis showing high accuracy (Pamungkas et al., 2023). This fact suggests the necessity to strive for a more accurate technique to determine the domain of the statement.

5. Conclusion

This paper concludes that SBERT model embeddings are the best sentence feature among the investigated models for the SWOT statement classification task. Since SWOT statements are typically short sentences, SBERT may be the most appropriate BERT-base model that can capture the contextual meaning of the statements. SWOT classification using SBERT model embeddings pre-trained with 1.5 GB dataset achieves an accuracy of 0.73 and F1-Score of 0.738. SBERT outperforms the traditional TF-IDF method to produce representative sentence vectors, even for a model that is pre-trained by a smaller dataset.

Reference

Abu-Alaish, A., Jaradat, G., Al-Shugran, M., Alodat, I., 2021. Automating SWOT Analysis Using Machine Learning Methods. *International Journal of Advances in Soft Computing and its Applications*, 13(2), 140-161.

Achmad, N., Sholahuddin, M., Murwanti, S., Nugroho, S.P., 2021. Determining SWOT of Culinary Tourism in South Square of Surakarta Palace. *International Journal of Business, Economics and Management*, *4*(2), 427–432. https://doi.org/10.31295/ijbem.v4n2.1729

Akella, S., 2021. Applying SWOT for B2B Decisions, Extension to larger data with Machine Learning

Regression. *Journal of Physics: Conference Series*, 1998. 10.1088/1742-6596/1998/1/012005

Alalie, H.M., Harada, Y., Noor, I.M., 2019. Impact of Strength, Weakness, Opportunities, Threats (SWOT) Analysis on Realizing Sustainable Competitive Advantage in Banking Industry Sector in Iraq. *International Journal of Scientific and Research Publications*, *9*(3), 49–52. http://dx.doi.org/10.29322/IJSRP.9.03.2019.p8708

Aurachman, R., Putri, E.M., 2020. University Strategic System Engineering based on BAN PT Accreditation Criteria One using SysML and Semantic Approach. *Journal of Physics: Conference Series*, 1477, 1-6. 10.1088/1742-6596/1477/5/052022

Azizah, S.F.N., Cahyono, H.D., Sihwi, S.W., Widiarto, W., 2023. Performance Analysis of Transformer Based Models (BERT, ALBERT and RoBERTa) in Fake News Detection. *ArXiv Preprint ArXiv:2308.04950*. https://doi.org/10.48550/arXiv.2308.04950

Bilal, M., Almazroi, A.A., 2023. Effectiveness of fine-tuned BERT model in classification of helpful and unhelpful online customer reviews. *Electronic Commerce Research*, 23(4), 2737–2757. http://dx.doi.org/10.1007/s10660-022-09560-w

Cahyawijaya, S., Lovenia, H., Aji, A.F., Winata, G.I., Wilie, B., Mahendra, R., … others, 2022. NusaCrowd: Open Source Initiative for Indonesian NLP Resources. *ArXiv Preprint ArXiv:2212.09648*. https://doi.org/10.48550/arXiv.2212.09648

Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., Lopez, A., 2020. A Comprehensive Survey on Support Vector Machine Classification: Applications, Challenges and Trends. *Neurocomputing*, 408, 189–215. https://doi.org/10.1016/j.neucom.2019.10.118

Denaya, D., 2023. IndoSBERT: Indonesian SBERT for Semantic Textual Similarity tasks. Retrieved November 9, 2023. https://huggingface.co/denaya/indoSBERT-large

Devlin, J., Chang, M.-W., Lee, K., Toutanova, K., 2018. BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. *ArXiv Preprint ArXiv:1810.04805*. https://doi.org/10.48550/arXiv.1810.04805

Etokidem, A.J., Nkpoyen, F., Ekanem, C., Mpama, E., Isika, A., 2020. Strengths, Weaknesses, Opportunities and Threats (SWOT) Analysis of Civil Society Organisations for Immunisation in a Nigerian Rural Community. *West African Journal of Medicine*, 37(6), 650–655.

Fafurida, Oktavilia, S., Karsinah, Fauziah, S., 2020. Development of Potential Culinary and Shopping Tourism in Central Java Province. *Jurnal Ekonomi Pembangunan: Kajian Masalah Ekonomi dan Pembangunan*, 21(2), 107–117. https://doi.org/10.23917/jep.v21i2.10590

Firqa, A., 2022. Indo-Sentence-BERT: Indonesian Sentence BERT for Semantic Similarity. Retrieved November 9, 2023. https://huggingface.co/firqaaa/indo-sentence-bert-base

Fuadi, D., Suwandi, J., Sutama, Maryadi, 2019. Developing Learning Model of Life Skills-Based Courses in University. In *Proceedings of the 4th Progressive and Fun Education International Conference, Profunedu 2019, 6-8 August 2019, Makassar, Indonesia*. http://dx.doi.org/10.4108/eai.7-8-2019.2288438

Gomes, L., da Silva Torres, R., Côrtes, M.L., 2023. BERT-and TF-IDF-Based Feature Extraction for Long-Lived Bug Prediction in FLOSS: a Comparative Study. *Information and Software Technology*, 160, 107217. https://doi.org/10.1016/j.infsof.2023.107217

González-Carvajal, S., Garrido-Merchán, E.C., 2020. Comparing BERT Against Traditional Machine Learning Text Classification. *ArXiv Preprint ArXiv:2005.13012*. https://doi.org/10.48550/arXiv.2005.13012

Granulo, A., Tanović, A., 2020. The advantage of using SWOT analysis for companies with implemented ITIL framework processes. *2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*, 1656–1661. https://doi.org/10.23919/MIPRO48935.2020.9245393

Hao, Y., Dong, L., Wei, F., Xu, K., 2019. Visualizing and understanding the effectiveness of BERT. *ArXiv Preprint ArXiv:1908.05620*. https://doi.org/10.48550/arXiv.1908.05620

Imaduddin, H., A'la, F.Y., Nugroho, Y.S., 2023. Sentiment Analysis in Indonesian Healthcare Applications using IndoBERT Approach. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 14(8), 113-117. https://dx.doi.org/10.14569/IJACSA.2023.0140813

Kale, A.S., Pandya, V., Troia, F.D., Stamp, M., 2023. Malware Classification with Word2Vec, Hmm2Vec, BERT, and ELMo. *Journal of Computer Virology and Hacking Techniques*, 19(1), 1–16. http://dx.doi.org/10.1007/s11416-022-00424-3

Lima, E., Nascimento, M.H.R., Alencar, D.B.D., Nascimento, M.R., Júnior, J.R.L.P., Silva, A.L.F.D., 2021. Swot Analysis Implemented With Fuzzy Inference to Support Decision Making. *International Journal for Innovation Education and Research*. https://doi.org/10.31686/ijier.vol9.iss9.3368

Lohrke, F.T., Mazzei, M.J., Frownfelter-Lohrke, C., 2022. Should it Stay or Should it Go? Developing an Enhanced SWOT Framework for Teaching Strategy Formulation. *Journal of Management Education*, 46(2), 345–382. https://doi.org/10.1177/10525629211021143

Makmun, A., Thamrin, H., 2018. Performance of Similarity Algorithms for Statement Mapping in a SWOT Analysis Application. *AIP Conference Proceedings*, 1977. https://doi.org/10.1063/1.5042902

Mateos, M.C., 2020. Hitting the Target. *Quality Progress*, *53*(4), 34–41.

Milios, A., BehnamGhader, P., 2022. An Analysis of Social Biases Present in BERT Variants Across Multiple Languages. *ArXiv Preprint ArXiv:2211.14402*. https://doi.org/10.48550/arXiv.2211.14402

Munikar, M., Shakya, S., Shrestha, A., 2019. Fine-Grained Sentiment Classification using BERT. *2019 Artificial Intelligence for Transforming Business and Society (AITB)*, 1, 1-5. https://doi.org/10.1109/AITB48515.2019.8947435

Namugenyi, C., Nimmagadda, S.L., Reiners, T., 2019. Design of a SWOT Analysis Model and Its Evaluation in Diverse Digital Business Ecosystem Contexts. *Procedia Computer Science*, 159, 1145-1154. https://doi.org/10.1016/j.procs.2019.09.283

Pamungkas, E.W., Putri, D.G.P., Fatmawati, A., 2023. Hate Speech Detection in Bahasa Indonesia: Challenges and Opportunities. *International Journal of Advanced Computer Science and Applications*, 14(6). https://dx.doi.org/10.14569/IJACSA.2023.01406125

Pérez-Mayos, L., Ballesteros, M., Wanner, L., 2021. How Much Pretraining Data do Language Models Need to Learn Syntax?. *ArXiv Preprint ArXiv:2109.03160*. https://doi.org/10.48550/arXiv.2109.03160

Pisner, D.A., Schnyer, D.M., 2020. Support Vector Machine. *Machine learning*, 101-121. https://doi.org/10.1016/B978-0-12-815739-8.00006-7

Prabhu, S., Mohamed, M., Misra, H., 2021. Multi-Class Text Classification using BERT-Based Active Learning. *ArXiv Preprint ArXiv:2104.14289*. https://doi.org/10.48550/arXiv.2104.14289

Reimers, N., Gurevych, I., 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, 3982-3992. https://doi.org/10.18653/v1/D19-1410

Shen, Y., Liu, J., 2021. Comparison of Text Sentiment Analysis Based on BERT and word2vec. *2021 IEEE 3rd International Conference on Frontiers Technology of Information and Computer (ICFTIC)*, 144-147. https://doi.org/10.1109/ICFTIC54370.2021.9647258

Taherdoost, H., Madanchian, M., 2021. Determination of Business Strategies Using SWOT Analysis; Planning and Managing the Organizational Resources to Enhance Growth and Profitability. *Macro Management \& Public Policies*, 3(1), 19-22. https://doi.org/10.30564/mmpp.v3i1.2748

Teimoori, D., Alinezhad, A., 2019. Organizational Sustainable Competitive Advantage using ORESTE, TRIZ, SWOT Approaches in Gray Conditions. *Iranian Journal of Optimization*, 11(2), 85–96. https://dorl.net/dor/20.1001.1.25885723.2019.11.2.2.3

Thamrin, H., Pamungkas, E.W., 2017. A Rule Based SWOT Analysis Application: A Case Study for Indonesian Higher Education Institution. *Procedia Computer Science*, 116, 144–150. https://doi.org/10.1016/j.procs.2017.10.056

Togayev, S., Zayniddinova, D., 2023. SWOT-The Essence and Methodological Problems of Analysis. *Educational Research in Universal Sciences*, 2(4), 1003–1008. https://doi.org/10.5281/zenodo.7922139

Tweheyo, G., Mugarura, A., 2020. Relevance of SWOT Analysis in Universal Secondary Education. *International Journal of Education and Social Science Research*, 3(1), 242–257. http://dx.doi.org/10.37500/IJESSR.2020.30121

Vlados, C., Chatzinikolaou, D., 2019. Towards a restructuration of the conventional SWOT analysis. *Business and Management Studies*, 5(2), 76-84. https://doi.org/10.11114/bms.v5i2.4233

Wang, Z., Ng, P., Ma, X., Nallapati, R., Xiang, B., 2019. Multi-Passage BERT: A Globally Normalized BERT Model for Open-Domain Question Answering. *ArXiv Preprint ArXiv:1908.08167*. https://doi.org/10.48550/arXiv.1908.08167