Received: December 16, 2021. Revised: March 7, 2022.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Aspect-Based Sentiment Analysis of Financial Headlines and Microblogs Using Semantic Similarity and Bidirectional Long Short-Term Memory

Agus Tri Haryono[1] **Riyanarto Sarno[1]*** **Rachmad Abdullah[1]**

Department of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
* Corresponding authors Email: riyanarto@if.its.ac.id

Abstract: Financial headlines and microblogs usually have sentiment of finance and research further about aspect of sentiment analysis still needed. From dataset FiQA 2018 challenges 4 aspect represent about financial, which is corporate, economy, stock and market. This research proposes method to determine financial headlines and microblogs to financial aspect. A sentence form dataset pre-processed into keyword. Aspect categorization using semantic similarity calculate similarity measurement on word vector of sentence with term list of each aspect, the highest similarity value determines the aspect. Neural network is used for sentiment classification of headlines and microblogs, the method used word embedding and bidirectional long short-term memory (BLSTM). The sentiment classification results are combined with the aspect categorization results to determine which aspects have the highest positive sentiment. The highest aspect categorization performance is obtained combined semantic similarity and bidirectional long short-term memory which reach 88.1% and semantic classification accuracy reach 77.0%. The stock aspect has received more positive sentiment compared to the sentiment of other aspect, this can be used for started investment in stock.

Keywords: Financial analysis, Aspect categorization, Semantic similarity, Word embedding, Bidirectional LSTM.

1. Introduction

Choosing a good company to invest in the stock market will be a difficult and high-risk task, because the unstable natural environment of the stock market and changes in stock prices are affected by many aspects. So, a strong analysis is needed to evaluate aspects of the company to reduce the risk of investment failure.

In the research that has been done, changes in stock prices can be influenced by news [1]. News about companies listed on the stock exchange have varied information and complex structures, coupled with information from social media, increasing the diversity of data to be processed.

With the high diversity of data from news and microblogging social media, natural language processing (NLP) techniques can be used to extract and identify sentiments from a given text. The sentiment analysis method used to classify text [2]. Sentiment analysis can be a meaningful process of classifying categories to find patterns in words or sentences [3].

However, sometimes only part of the story to shown as result of sentiment analysis, for illustration "the company has a very good finances but the market share is declining". Multiple aspect can present from a given aspect and different sentiment on each of them, positive sentiment on corporate financial and negative sentiment on market share.

Aspect based sentiment analysis (ABSA) introduce in 2010 [4], this method is used for extract aspect from given text and sentiment associated. ABSA has several sub-tasks to identifying relevant aspect and determining sentiment of each aspect. Dataset from FiQA 2018 challenge on sub-task 1 is used on this research. The dataset supplied consist of headline news and tweet on financial subject. The feature of dataset is score of sentiment, relevant snippets of sentence and some of level aspect of each sentence. To predict aspect of each sentence, this experiment will be using two method such as neural network model and semantic similarity, whereas determine sentiment score using neural network classification.

In the challenge that has been implemented and written in a research journal [5], the previous researcher uses a neural network method, namely for aspect extraction using bidirectional long short-term memory recurrent neural network (BLSTM RNN) and word embedding Google-News-Word2Vec, for sentiment analysis using multi-channel convolutional neural network (CNN) and word embedding with combination of pre-trained model Godin and Word2Vec-Google-News. Result of previous research for the measure performance of aspect model reach 0.69 and for model of sentiment classification has R squared score 0.288 with mean squared error (MSE) 0.112. Another study showed that using the a lite bidirectional encoder representations from transformers (ALBERT) model as a deep learning method and the semeval-2014 dataset task 4, subtask 1 aspect extraction resulted an accuracy of 83.04 % [6]. The self-attention gated convolutional neural network (CNN) method, which used in previous research with the SemEval restaurant 2016 dataset, produced a performance of 81.40 % [7].

In addition to using a neural network in determining aspect, another method that can be used is semantic similarity. Semantic similarity is measurement distance of semantic meaning from each document or term [8]. Semantic similarity from previous research has good performance on determine aspects, such as on research on restaurant reviews has performance 84 % [9] and on hotel reviews has performance 84 % [10].

Deep learning method to predict imbalanced data produce low accuracy or semantic similarity with imbalance term list produce high false positive and false negative, deep learning method and semantic similarity can be combined to improve accuracy and reduce false results.

Semantic similarity, bidirectional LSTM and proposed technique combined semantic similarity and bidirectional LSTM is used as compared method on determine aspect in this research. The proposed technique is expected to improve the performance of supervised learning with Bidirectional LSTM method by utilizing semantic similarity.

In this paper, to determine aspect categorization is uses semantic similarity method with Table 1 as glossary of terms, bidirectional LSTM as deep learning method and combined semantic similarity and Bidirectional LSTM as proposed method. Each method's performance will be evaluated using accuracy and the F1 score. Furthermore, bidirectional LSTM, will be used to determine the sentiment

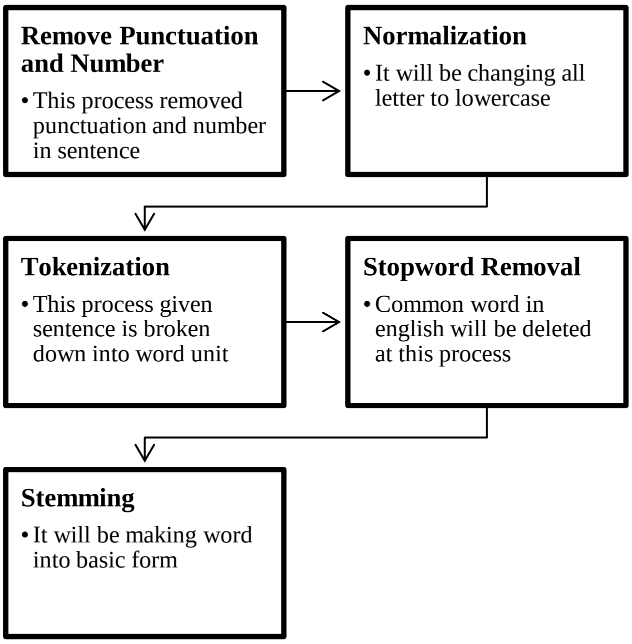

Figure. 1 Pre-processing technique

classification on each document or term. The accuracy, recall, precision, and F1-score are used to evaluate performance results and determine which aspect sentiment has the least error.

2. Related theory

A number of theories related to the research will be explained in this section.

2.1 Pre-processing

The pre-processing technique often used in natural language processing consist of the following technique shown Fig. 1.

2.2 Aspect keyword term list

Corporate, stock, economy and market is aspect in this experiment. These aspects are determined from dataset. In addition to the four aspects, the research also looking on the sentiment of sentence is positive or negative to represent financial condition.

The sentence will be classified into predetermined aspects after reviewed. Keyword aspect determined using keyword extract Wikipedia page related to each aspect and keyword from dataset, data in Table 1.

2.3 Semantic similarity

Method used for measurement distance of semantic meaning each document or term is semantic similarity [11]. For similarity calculation have two types, which are based on the existing synonym

Received: December 16, 2021. Revised: March 7, 2022.

Table 1. Listicle of keyword aspect vocab

Aspect	Keyword Aspect Vocabulary
Corporate	credit, court, acquisition, legality, product, claim, factor, lending, pay-out, shareholder, reputation, risk, ridiculous, standardization, rate, economic, equity, environmental, export, member, position, successfully, investment, accounting, cost, reference, criminal, negativity, transaction, anxiety, strategy, image, legal, information, reputable, brand, status, appointive, finance, trust, companies, stock, regulate, consolidation, theory, general, policy, company, regulation, expectation, acquirer, legislation, marketing, stereotype, fact, regulatory, strategic, appointment, debt, operating, joke, manufacturing, communicate, need, lawful, management, control, merger, growth, dividend, judicial, federal, quarterly, tax, pattern, reputational, quality, corporation, business, civil, production, industry, income, quarter, law, problem, appoint, constitutional, customer, illegality, regulator, designate, press, relationship, financial, initiative, exercisable, communication, banking, corporate, special, probability, identity, plan, planning, sale, rise, sell, client
Stock	indicate, confidential, support, value, ratio, sentiment, day, statistical, selling, turnover, signalling, buyside, buying, knowledge, investing, analysis, action, additional, testing, fundamentally, average, illegal, option, broker, level, insider, technical, sell, investment, point, fundamental, offering, initial, consider, firm, measure, price, offer, indicator, backtesting, private, forecasting, growth, mathematical, theory, trading, investor, illegality, trend, open, empirical, available, uptrending, stock, uptrend, method, outstanding, selloff, signal, pricing, inventory, activity, alternative, examination, shares
Economy	oil, bank, debt, capital, depress, healthcare
Market	inflation, business, forex, trading, timing, industry, sentiment, currencies, generally, annualize, monetary, exchange, demand, extremely, result, risk, period, devaluation, fx, pricing, crisis, economy, price, global, stock, euro, consumer, volatile, derivative, currency, dollar, volatility, market, extreme

resource, and based on collection of words in a corpus [12].

$$Similarity(X_a, Y_b) = \frac{\sum_{a=1}^{n} X_a Y_b}{\sqrt{\sum_{a=1}^{n} X_a^2}\sqrt{\sum_{a=1}^{n} Y_b^2}} \qquad (1)$$

The first (X_a) and second (Y_b) words will measured the proximity between the words that produce the similarity of the two words. $\sum_{a=1}^{n} X_a Y_b$ is the number of repetitions on the addition n vector words. Similarity distance with score 1 is the maximum value, that defines extremely the same meaning. If score 0 is the minimum value of similarity distance, which explained excessively different meaning. Accordingly, 0 and 1 is ranged value of similarity meaning.

2.4 Bidirectional LSTM + word embedding

Long-short term memory (LSTM) [13] is a type of recurrent neural network architecture (RNN) that is used to model the intercourse among phrase with long lapse. The output of the LSTM categorization depends on the training data provided. Based on the interests assigned to the information, keep or remove information input (e.g., whether to close the gate or otherwise) is determine by long-short term memory (LSTM) layers. Algorithms have also deliberated the use of weights in assigning notable relationship. That means the long-short term memory (LSTM) layers learns which information is important and which is not [14].

Bidirectional LSTM network [15] is architecture to utilize both earlier context and later context of particular word, architecture illustrated in Fig. 2. It has of two hidden layers, first layer computes the forward hidden thread $\overrightarrow{h_t}$, subsequently the second hidden layer compute backward thread $\overleftarrow{h_t}$, after that both layers combine $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ to generate output y_t. BLSTM is implemented by the ensuing functions:

$$\overrightarrow{h_t} = H\left(W_{x\overrightarrow{h}} x_t + W_{\overrightarrow{h}\overrightarrow{h}} \overrightarrow{h}_{t-1} + b_{\overrightarrow{h}}\right) \qquad (2)$$

$$\overleftarrow{h_t} = H\left(W_{x\overleftarrow{h}} x_t + W_{\overleftarrow{h}\overleftarrow{h}} \overleftarrow{h}_{t-1} + b_{\overleftarrow{h}}\right) \qquad (3)$$

$$y_t = W_{\overrightarrow{h}y} \overrightarrow{h_t} + W_{\overleftarrow{h}y} \overleftarrow{h_t} + b_y \qquad (4)$$

In this experiment, bidirectional long-short term memory (BLSTM) method is used for aspect categorization and sentiment classification, as input, sentence from dataset to extract using word embedding become corpus and convert into word vector feature. Word embedding used is word2vec

Received: December 16, 2021. Revised: March 7, 2022.

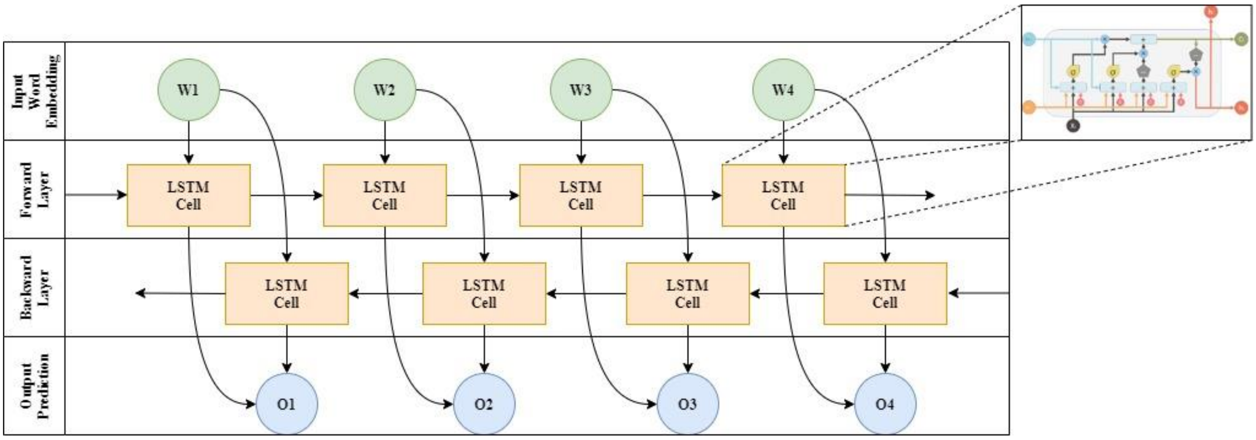

Figure. 2 Bidirectional LSTM architecture

Table 2. Deputation of the document in the first form before reformatted.

Index	1
sentence	Royal Mail chairman Donald Brydon set to step down
info	[{'snippets': "['set to step down']", 'target': 'Royal Mail', 'sentiment_score': '-0.374', 'aspects': "['Corporate/Appointment']"}]

from pre-trained google news model. Google news model contains word vector with vocabulary of 3 million words and phrases that trained on approximately 100 billion words from Google News dataset. The feature length of word vector is 300. So, Google new model has good accuracy to be implantation, because google news model can estimate with good representation of the word [16]. Aspect and sentiment classification will be using this theory for determined aspect and sentiment of each sentence.

2.5 Evaluation

In this work, we propose an accuracy measurement approach for the prediction of aspects and the precision, recall, and F-1 score for sentiment analysis. This measurement approach is thought to be the most dependable for assessing the suggested method's performance [17]. The performance of data testing with ground truth value represented by true positive (BP), false positive (SP), true negative (BN) and false negative (SN) determines the recall and precision calculating procedure.

$$P = \frac{BP}{(BP+SP)} \tag{5}$$

Precision (P) is the exactness value between the ground truth information and the response decided by the system.

$$R = \frac{BP}{(BP+SN)} \tag{6}$$

Recall (R) is the success rate to regain information by the system.

$$F = 2 \times \frac{P \times R}{(P+R)} \tag{7}$$

3. Research method

3.1 Dataset

Dataset from FiQA 2018 [5] provided two files training with json format, 653 snippets tweets and 404 snippets headlines news with financial subject. Each file has sentence and info feature. In info feature has snippets, target, sentiment score and aspects. The data need to be reformatted before take step pre-processing. Table 2 shows an example of a data block for a single sentence.

Each sentence's aspect can have up to 6 levels, but we will only use aspects up to level 1 for this experiment with 4 aspect. Fig. 3 lists the names and frequency of each aspect. In Fig. 3 there is a significant class imbalance. As a result, classification certain aspects are difficult.

Range in sentiment scores are [−1, 1]. −1 is

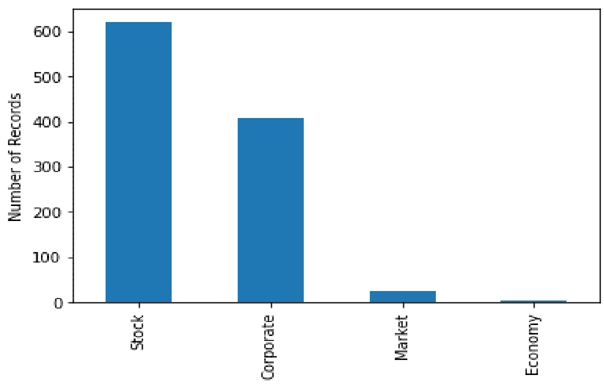

Figure. 3 Density classification of each aspect in dataset

Received: December 16, 2021. Revised: March 7, 2022.

Start

1. Take a sentence from dataset
2. Perform pre-processing in table 1 with sequence: Remove punctuation → Remove number → Normalization to lowercase.
3. Perform pre-processing using Python Natural Language Toolkit (NLTK) and Spacy Model with sequence: Tokenization → Stop word Removal → Stemming
4. Save the result for next step.

Stop

Figure. 4 Pseudocode pre-processing

Table 3. Result of the pre-processing

Before	After
Royal Mail chairman Donald Brydon set to step down	'royal', 'mail', 'chairman', 'donald', 'set', 'step'

incredibly negative and 1 is incredibly positive. In this experiment sentiment score will be normalize to binary number, with [-1, 0] is 0 and [>0, 1] is 1.

3.2 Pre-processing

This stage involves the steps shown Fig. 4. Pre-processing pseudocode result into corpus in Table 3.

3.3 Aspect categorization

Two experiments will be carried out to determined the best aspect categorization performance. The first method, hereinafter referred to as AC1, the aspect term keywords and the pre-processed keywords is computed the distance by the semantic similarity. Another method called AC2 uses the BLSTM neural network to determine the aspect of each sentence. The last method, as proposed called as AC3, used combine technique semantic similarity and the BLSTM to improve accuracy in identifying the aspect of each sentence.

The aspect categorization experiment description are follows:

3.3.1. Aspect categorization using semantic similarity (AC1)

The AC1 will carry out the process shown Fig. 5. Data from Table 3 has extracted to keyword list using pre-processing will be calculated for similarity score with list of term in Table 1 to categorize the data to the pre-established 4 aspects. The rate of similarity score ranged between 0 and 1, value close to 0 is representing dissimilar sentence to aspect. If score near to 1, the keyword from sentence is identical and potential able to indicate the aspect. So, the highest value of similarity value of each aspect will be used for determine aspect of sentence.

Start

1. Take list keyword from previous stage and aspect term list
2. Using python gensim and google news model word2vec controlling word embedding to expand every word into vector
3. Calculate semantic similarity between keyword and aspect term using Eq. (1)
4. Find highest similarity value from all aspects for determine aspect of sentence.

Stop

Figure. 5 Pseudocode of AC1

Table 4. Result of semantic similarity (AC1)

Term List	Aspect 1	Aspect 2	Aspect 3	Aspect 4
	Corporate	Stock	Economy	Market
'royal', 'mail', 'chairman', 'donald', 'set', 'step'	**0.4040**	0.2761	0.2116	0.2097

Start

1. Take list keyword from pre-processing stage
2. Using python gensim and google news model word2vec controlling word embedding to expand every word into vector
3. Convert aspects into binary multi class
4. Classifying the aspect using the BLSTM Eq. (2) to Eq. (4).

Stop

Figure. 6 Pseudocode of AC2

3.3.2. Aspect categorization using BLSTM (AC2)

The subsequent processes were carried for the AC2 shown Fig. 6.

The example of the data in Table 3 is cultivate for aspect-based classification. The keyword from Table 3 would be extended in term of the term list using word embedding into vector feature. Word2Vec from Google News model pre-trained was the word embedding used. The vector feature obtained of word is used by bidirectional long short-term memory

Received: December 16, 2021. Revised: March 7, 2022.

Table 5. Bidirectional LSTM (AC1) result

Term List	Aspect Prediction
'royal', 'mail', 'chairman', 'donald', 'set', 'step'	**Corporate**

Start

1. Take list result form AC1
2. Calculate average score similarity of each aspect.
3. Compare average aspect score similarity and score similarity of each sentence.
4. If score similarity above average aspect score, result of aspect not changed.
5. If score similarity below average aspect score, used model from AC2 to identifying aspect.

Stop

Figure. 7 Pseudocode of AC3

(BLSTM) to build classification model.

3.3.3. Aspect categorization using semantic similarity and BLSTM (AC3)

The procedure will be carried out by the AC3 shown Fig. 7.

The aspect average score was used as the limit value to identify the error in determining the aspect in AC1. If the similarity score is less than the aspect average score, the model created in AC2 was used to determine the aspect.

3.4 Sentiment classification

The process of sentiment classification almost looks like to aspect categorization using BLSTM process, different just label in use and type classification binary single class. Label use is sentiment score with change value [-1, 1] into binary number with rule [-1, 0] is 0 and [>0, 1] is 1. The sentiment classification results are expressed as a score between 0 and 1, with a score close to 0 representing a negative sentiment and a score close to 1 representing a positive sentiment.

3.5 Evaluation

Aspect categorization and sentiment classification performance was evaluated by comparing a few results. Accuracy, precision, recall, and F1 score would be used to evaluate each performance.

Table 6. Sentiment classification result

Term List	AC1	AC2	Sentiment	
			Binary	Score
'royal', 'mail', 'chairman', 'donald', 'set', 'step'	**Corporate**	**Corporate**	Negative	0.0054

Table 7. Aspect categorization approach

Description	Approach
Semantic similarity is used to categorized the term list of result pre-processing into 4 aspects that have been determined. The proximity between the aspect in Table 1 to the term list will calculate by semantic similarity.	**AC1**
Create a trained model with Bidirectional LSTM and Word Embedding to classify the pre-processing results into each aspect.	**AC2**
The aspect determination method used combination of semantic similarity and the Bidirectional LSTM model, with a limit value for the average score of aspects.	**AC3**

Table 8. Aspect categorization performance

Approach	Data Test	Accuracy	Weighted Average F1 Score
AC1	1057	0.801	**0.803**
AC2	212	0.811	**0.790**
AC3	1057	0.881	**0.881**

4. Result

Aspect categorization and sentiment analysis result will be explained.

4.1 Aspect categorization approach

Experiments were carried out to calculate the performance of the method for determining aspects and evaluating the accomplishment for aspect distribution using two approaches aspect categorization. The AC1 technique uses 1057 data, which is the entire dataset, whereas the AC2 approach uses 212 data, which is 20 % of the dataset, with the remaining 80 % used as training data in modelling.

Principle in Eq. (5) as F1-measure and accuracy is used to evaluated each aspect categorization and sentiment classification. Table 8 shows the evaluation based on aspect categorization using semantic similarity and bidirectional LSTM.

According to Table 8, approach aspect categorization 1 (AC1) evaluated using term list in Table 1 by semantic similarity and aspect categorization 2 (AC1) evaluated using bidirectional

Received: December 16, 2021. Revised: March 7, 2022.

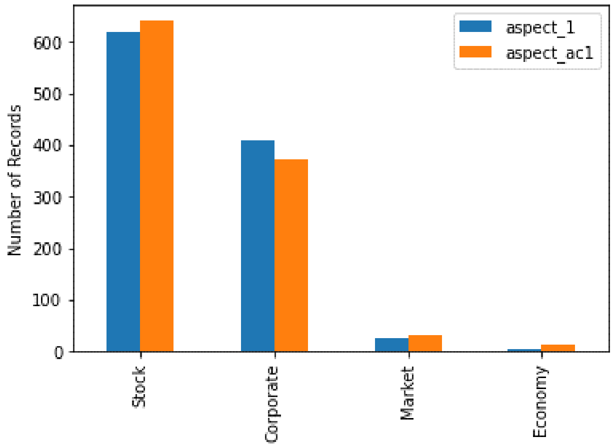

Figure. 8 Frequency distribution AC1

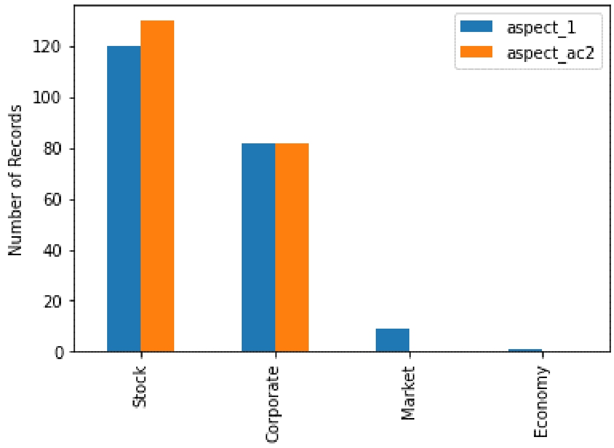

Figure. 9 Frequency distribution AC2

Table 9. Frequency distribution dataset and AC1

Aspect	Dataset	AC1
Stock	620	641
Corporate	409	372
Market	25	31
Economy	3	13
Total	**1057**	**1057**

Table 10. Frequency distribution data test and AC2

Aspect	Data test	AC2
Stock	120	138
Corporate	82	74
Market	9	0
Economy	1	0
Total	**212**	**212**

LSTM plus word embedding. The performance approach of aspect categorization 1 is similar to the performance approach of aspect categorization 2, with AC1 producing application performance up to 0.803 and AC2 producing application performance up to 0.790. With a difference of 1.3 %, AC1 outperforms AC2.

According to Fig. 8 and Fig. 9, AC1 performance better than AC2 because AC1 better to determine aspect Economy and Market. Shown in Table 9, all dataset can be determined into 4 aspect with AC1 approach, although the result still have false positive and false negative. With AC2 approach, aspect Economy and Market has zero result from data testing, shown in Table 10.

AC1 generated 1057 data score sentiment from the result, and AC3 computed the average score sentiment for the limit value, as showed in table 11. If the score sentiment is less than a certain threshold, the model from AC2 was used for the determined aspect; otherwise, the score sentiment and aspect from AC1 are still used.

Fig. 10 showed that AC3 outperformed AC1.

Table 11. Average aspect score

Aspect	Average Score
Stock	0.457
Corporate	0.457
Market	0.528
Economy	0.449

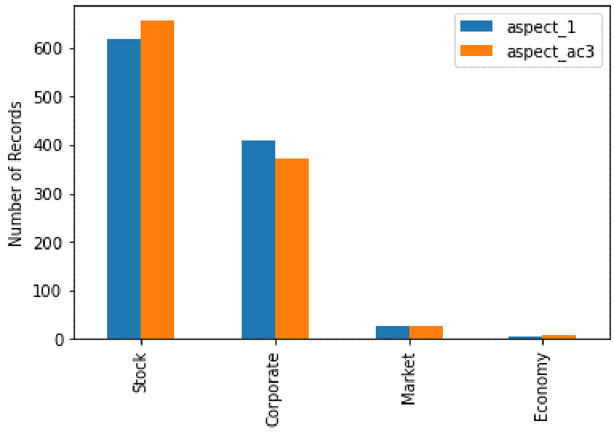

Figure. 10 Frequency distribution AC3

Table 12. Frequency distribution Dataset and AC3

Aspect	Dataset	AC3
Stock	620	656
Corporate	409	370
Market	25	25
Economy	3	6
Total	**1057**	**1057**

AC3 success reduced false positive and false negative from semantic similarity technique using bidirectional LSTM model, but did not eliminate semantic similarity's ability to detect aspects with limited data. As showed in Table 12, the frequency distribution of each aspect from AC3 better than AC1 and AC2, particularly on the market and economy aspects. According to Table 8, there was significant improvement on AC3 with a difference F1 score 7.8 % higher than AC1.

Table 11. Sentiment classification performance

Sentiment Analysis Performance	
Accuracy	0.77
Precision	0.74
Recall	0.75
F1-Score	0.74

Table 12. Frequency distribution dataset and result of sentiment classification

Aspect	Sentiment	Dataset (in Percent)	Sentiment Result (in Percent)
Corporate	Negative	14.00	12.20
	Positive	24.69	22.80
Stock	Negative	18.54	19.11
	Positive	40.11	42.95
Market	Negative	1.51	1.32
	Positive	0.85	1.04
Economy	Negative	0.19	0.57
	Positive	0.09	0.00
Total		100	100

4.2 Sentiment classification approach

After aspect categorization, next step is determined sentiment of each sentence from dataset. The approach sentiment classification processing using similar step with aspect categorization using BLSTM, which is from 1057 dataset split into 845 as data training and 212 as data testing.

According to Table 11, performance from sentiment classification using bidirectional LSTM and word embedding is 77 % in accuracy and 74 % in precision, 75 % recall and 74 % f1-score. calculation evaluation performance using formula showed in Eqs. (5), (6), and (7). With previous trained model sentiment classification, researcher analysed sentiment of each aspect from AC3 result and will be compared with ground truth from dataset.

From Table 12, it can be concluded distribution of sentiment financial headlines and microblogs. As the ground truth of dataset, sentiment negative had 362 data and 695 as sentiment positive. Compared with result of sentiment classification sentiment negative had 351 data and sentiment positive had 706 data. Aspect stock with sentiment positive had difference result 2.84 %, which is the biggest difference from all aspect sentiment. The result of the smallest difference is owned by aspect Market with sentiment negative with difference result 0.18 %.

5. Conclusion

The proposed method proved can classify dataset into 4 aspect financial and can reduce false positive and false negative by semantic similarity using bidirectional LSTM model, hence enhancing the accuracy of the method. The highest aspect categorization performance obtained by combined Semantic Similarity and bidirectional LSTM model which reach 88.1 % and sentiment classification accuracy using bidirectional LSTM and Word Embedding reach 77 %. The result of aspect categorization outperforms previous research with same data, which reached 0.69 % [5]. The result sentiment analysed of each aspect has biggest difference result on aspect Stock with sentiment positive which is 2.84 % and smallest difference result which is 0.18 % on aspect market with sentiment negative.

This study still has a flaw that can be addressed. Future research ideas include optimizing the term list for each aspect, adding polysemy recognition capability to increase semantic similarity, and improving imbalance data using under or over sampling. Following that, financial headlines from portal news and tweets from microblogging social media can be used in research as real-time data.

Acknowledgments

This research was funded by Lembaga Pengelola Dana Pendidikan (LPDP) under Riset Inovatif-Produktif (RISPRO) Invitation Program managed, the Indonesian Ministry of Education and Culture under Penelitian Terapan Unggulan Perguruan Tinggi (PTUPT) Program, and Institut Teknologi Sepuluh Nopember (ITS) under project scheme of the Publication Writing and IPR Incentive Program (PPHKI).

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

This research can work well and successfully because of the following research contributions: Conceptualization by Agus Tri Haryono and Prof. Riyanarto Sarno; methodology and software by Agus Tri Haryono; validation, formal analysis, investigation, resources by Prof. Riyanarto Sarno, Rachmad Abdullah, and Agus Tri Haryono; data curation, writing-original draft preparation, and writing-review and editing by Agus Tri Haryono; visualization by Rachmad Abdullah; supervision by Prof Riyanarto Sarno.

References

[1] W. Wang, K. Y. Ho, W. M. Liu, and K. Wang, "The relation between news events and stock

price jump: An analysis based on neural network", In: *Proc. - 20th Int. Congr. Model. Simulation, MODSIM 2013*, No. December, pp. 1406–1411, 2013, doi: 10.36334/modsim.2013.f8.wang.

[2] B. Saberi and S. Saad, "Sentiment analysis or opinion mining: A review", *Int. J. Adv. Sci. Eng. Inf. Technol.*, Vol. 7, No. 5, pp. 1660–1666, 2017, doi: 10.18517/ijaseit.7.5.2137.

[3] B. Pang and L. Lee, "Opinion Mining and Sentiment Analysis", *Foundations and Trends in Information Retrieval*, Vol. 2, No. 1, 2008, pp. 1–2.

[4] T. T. Thet, J. C. Na, and C. S. G. Khoo, "Aspect-based sentiment analysis of movie reviews on discussion boards", *J. Inf. Sci.*, Vol. 36, No. 6, pp. 823–848, 2010, doi: 10.1177/0165551510388123.

[5] Hitkul, S. Shahid, S. Singhal, D. Mahata, P. Kumaraguru, and R. R. Shah, "Aspect-Based Sentiment Analysis of Financial Headlines and Microblogs", *Deep Learning-Based Approaches for Sentiment Analysis Algorithms for Intelligent Systems*, No. January, Springer, Singapore, 2020, pp. 111–137.

[6] E. S. Alamoudi and N. S. Alghamdi, "Sentiment classification and aspect-based sentiment analysis on yelp reviews using deep learning and word embeddings", *J. Decis. Syst.*, Vol. 30, No. 2–3, pp. 259–281, 2021, doi: 10.1080/12460125.2020.1864106.

[7] J. Yang and J. Yang, "Aspect Based Sentiment Analysis with Self-Attention and Gated Convolutional Networks", In: *Proc. IEEE Int. Conf. Softw. Eng. Serv. Sci. ICSESS*, Vol. 2020-Octob, pp. 146–149, 2020, doi: 10.1109/ICSESS49938.2020.9237640.

[8] R. Abdullah, S. Suhariyanto, and R. Sarno, "Aspect Based Sentiment Analysis for Explicit and Implicit Aspects in Restaurant Review using Grammatical Rules, Hybrid Approach, and SentiCircle", *Int. J. Intell. Eng. Syst.*, Vol. 14, No. 5, pp. 294–305, 2021, doi: 10.22266/ijies2021.1031.27.

[9] A. Firmanto and R. Sarno, "Aspect-based sentiment analysis using grammatical rules, word similarity and SentiCircle", *Int. J. Intell. Eng. Syst.*, Vol. 12, No. 5, pp. 190–201, 2019, doi: 10.22266/ijies2019.1031.19.

[10] D. A. K. Khotimah and R. Sarno, "Sentiment analysis of hotel aspect using probabilistic latent semantic analysis, word embedding and LSTM", *Int. J. Intell. Eng. Syst.*, Vol. 12, No. 4, pp. 275–290, 2019, doi: 10.22266/ijies2019.0831.26.

[11] E. W. Pamungkas, R. Sarno, and A. Munif, "B-BabelNet: Business-specific lexical database for improving semantic analysis of business process models", *Telkomnika (Telecommunication Comput. Electron. Control.*, Vol. 15, No. 1, pp. 407–414, 2017, doi: 10.12928/TELKOMNIKA.v15i1.3176.

[12] E. Iosif and A. Potamianos, "Similarity computation using semantic networks created from web-harvested data", *Nat. Lang. Eng.*, Vol. 21, No. 1, pp. 49–79, 2015, doi: 10.1017/S1351324913000144.

[13] R. A. Priyantina and R. Sarno, "Sentiment analysis of hotel reviews using Latent Dirichlet Allocation, semantic similarity and LSTM", *Int. J. Intell. Eng. Syst.*, Vol. 12, No. 4, pp. 142–155, 2019, doi: 10.22266/ijies2019.0831.14.

[14] J. Wang, B. Peng, and X. Zhang, "Using a stacked residual LSTM model for sentiment intensity prediction", *Neurocomputing*, Vol. 322, pp. 93–101, 2018, doi: 10.1016/j.neucom.2018.09.049.

[15] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures", *Neural Networks*, Vol. 18, No. 5–6, pp. 602–610, 2005, doi: 10.1016/j.neunet.2005.06.042.

[16] S. M. Rezaeinia, R. Rahmani, A. Ghodsi, and H. Veisi, "Sentiment analysis based on improved pre-trained word embeddings", *Expert Syst. Appl.*, Vol. 117, pp. 139–147, 2019, doi: 10.1016/j.eswa.2018.08.044.

[17] Y. Jiao and P. Du, "Performance measures in evaluating machine learning based bioinformatics predictors for classifications", *Quant. Biol.*, Vol. 4, No. 4, pp. 320–330, 2016, doi: 10.1007/s40484-016-0081-2.