applied sciences

MDPI

Article

DPMS: Data-Driven Promotional Management System of Universities Using Deep Learning on Social Media

Mohamed Emran Hossain [1], Nuruzzaman Faruqui [2], Imran Mahmud [2], Tony Jan [3], Md Whaiduzzaman [3,4,*] and Alistair Barros [4]

[1] Department of Development Studies, Daffodil International University, Dhaka 1341, Bangladesh; emran@daffodilvarsity.edu.bd
[2] Department of Software Engineering, Daffodil International University, Dhaka 1342, Bangladesh; faruqui.swe@diu.edu.bd (N.F.); imranmahmud@daffodilvarsity.edu.bd (I.M.)
[3] Center for Artificial Intelligence Optimisation Research, Torrens University, Melbourne, VIC 3000, Australia; tony.jan@torrens.edu.au
[4] School of Information Systems, Queensland University of Technology, Brisbane, QLD 4000, Australia; alistair.barros@qut.edu.au
* Correspondence: md.whaiduzzaman@qut.edu.au

Abstract: SocialMedia Marketing (SMM) has become a mainstream promotional scheme. Almost every business promotes itself through social media, and an educational institution is no different. The users' responses to social media posts are crucial to a successful promotional campaign. An adverse reaction leaves a long-term negative impact on the audience, and the conversion rate falls. This is why selecting the content to share on social media is one of the most effective decisions behind the success of a campaign. This paper proposes a Data-Driven Promotional Management System (DPMS) for universities to guide the selection of appropriate content to promote on social media, which is more likely to obtain positive user reactions. The main objective of DPMS is to make effective decisions for Social Media Marketing (SMM). The novel DPMS uses a well-engineered and optimized BiLSTM network, classifying users' sentiments about different university divisions, with a stunning accuracy of 98.66%. The average precision, recall, specificity, and F1-score of the DPMS are 98.12%, 98.24%, 99.39%, and 98.18%, respectively. This innovative Promotional Management System (PMS) increases the positive impression by 68.75%, reduces the adverse reaction by 31.25%, and increases the conversion rate by 18%. In a nutshell, the proposed DPMS is the first promotional management system for universities. It demonstrates significant potential for improving the brand value of universities and for increasing the intake rate.

Keywords: deep learning; BiLSTM network; promotional management system; social media marketing; decision support system; data-driven decision

check for updates

Citation: Hossain, M.E.; Faruqui, N.; Mahmud, I.; Jan, T.; Whaiduzzaman, M.; Barros, A. DPMS: Data-Driven Promotional Management System of Universities Using Deep Learning on Social Media. *Appl. Sci.* **2023**, *13*, 12300. https://doi.org/10.3390/app132212300

Academic Editor: Chilukuri K. Mohan

Received: 18 September 2023
Revised: 4 November 2023
Accepted: 9 November 2023
Published: 14 November 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Data-driven decision-making is more effective than decision-making from experience or intuition [1]. Despite its effectiveness, its application is missing from the promotional management system of Higher Educational Institutions (HEIs). Choosing the right content to promote plays a crucial role in the overall success of a Social Media Marketing (SMM) campaign [2]. This paper proposes a Deep Learning (DL)-based Data-Driven Promotional Management System (DPMS) to automatically choose the most effective content to use for social media marketing, which reduces the rate of adverse reaction. The DPMS uses a sentiment analyzer to identify the strengths and weaknesses of Higher Educational Institutions (HEIs). It guides the marketing team by predicting the aspects of HEIs to highlight in order to obtain more positive reactions and comments on social media. As a result, it increases the success rates of SMM campaigns for HEIs.

The proposed Data-Driven Promotional Management System (DPMS) is trained to utilize stakeholders' feedback on various university divisions, including faculties, the library, the admissions office, finance and accounts, human resource management, and the registrar's office. This feedback is collected and refined through the application of Natural Language Processing (NLP) techniques that are well-suited for analyzing such data [3]. Subsequent to the collection process, this study undertakes extensive feature engineering to ascertain the most productive feature extraction methods [4]. Given the nature of the dataset features, a Bidirectional Long Short-Term Memory (BiLSTM) network is selected for crafting the data-centric promotional management system tailored for university settings. The BiLSTM network is meticulously architected to categorize users' feedback into four distinct classes, facilitating the extraction of valuable insights into stakeholders' experiences across the various university divisions. Based on these insights, promotional posts are crafted to accentuate the divisions that elicit positive experiences from users.

This novel DPMS has the potential to revolutionize universities' promotional management systems. It is a unique application of the BiLSTM network in Natural Language Processing (NLP)-based social media marketing. The novelties and outstanding contributions of this system are listed below.

- Novel Concept: The concept of applying the BiLSTM network to develop a data-driven decision-making system for universities on social media is the first of its kind.
- Effective Network Design: The BiLSTM network to design the DPMS has been carefully engineered and optimized to maximize the stakeholders' feedback dataset performance.
- Outstanding Performance: The DPMS demonstrates outstanding performance, with an average validation accuracy of 98.66%. Its precision, recall, specificity, and F1-score are 98.12%, 98.24%, 99.39%, and 98.18%, respectively.
- Improving Social Media Marketing Impression: This exceptional promotional management system demonstrates an increase in positive impressions by 68.75% and reduces the negative responses by 31.25% in real-world settings.

This paper presents a novel application of the BiLSTM network in NLP by developing a data-driven promotional management system for universities. The remainder of the article is organized into six different sections. The Section 2 discusses the relevant literature and compares it with our approach. The Section 3 is the proposed methodology for the design of the DPMS. The experimental results and evaluation have been discussed in Section 4. The practical implementation and analysis in real-world settings have been presented in Section 5. The limitations and corresponding future scope have been explained in the sixth section. Finally, the paper is concluded in Section 7.

2. Literature Review

The underlying technology behind the proposed DPMS is a human sentiment classifier developed using a BiLSTM network. E. A. Vetrova et al. [5], R. Pizarro Milian et al. [6], E. Nedbalova [7], and A. I. M. Elfeky et al. [8] researched a university promotional management system. However, none of them use a Deep Learning (DL)-based automatic approach. It is a significant research gap that the proposed DPMS has filled.

Although there is no data-driven promotional management system for HEIs, plenty of research has been conducted on sentiment analysis and the BiLSTM network, which are the underlying technology of the proposed DPMS. S. Murugaiyan et al. [9] developed an aspect-based sentiment classifier using a Deep Convolutional Neural Network (DCNN) and the BiLSTM network. It achieved an accuracy of 93.28%. The combination of BiLSTM and Bidirectional Encoder Representations from Transformers (BERT) developed by M. Wankhade et al. [10] classifies sentiment from SemEval2014 datasets [11] with an accuracy of 80.78%. A. S. Talaat et al. [12] analyzed the performance of sentiment classification on a small Apple dataset using a hybrid BERT model and achieved an accuracy of 91.72%. The proposed DPMS's sentiment classification accuracy is 98.12%, which is much better than these DL methods.

The application of DMPS is Social Media Marketing (SMM). B. S. Arasu et al. [13] developed a Machine Learning (ML)-based approach to enhance social media marketing, which achieves an average accuracy of 86.6%. E. Kongar et al. [14] applied ML to target customers through social media post mining. This approach reached an accuracy of up to 82%. A reinforcement learning-based approach developed by P. Eklund et al. [15] analyzes the effects of advertisements. The review of the recent literature ensures that the proposed DPMS is unique and that none of the concurrent approaches is similar.

The proposed DPMS is a BiLSTM network-based sentiment classification-centric approach. S. Aslan et al. [16] used a BiLSTM-based sentiment classifier using people's tweets on the Ukraine–Russia war. This approach obtains an accuracy of 91.79%. Z. Gou et al. [17] focuses on emotional classification using sentiment analysis with the BERT and BiLSTM network combination. It achieved an accuracy of 85.44%. M. Lestandy et al. [18] worked on the same objective—emotion classification. Their CNN-BiLSTM-based approach achieved an accuracy of 92.85%. Although the proposed DPMS's technology is similar to these approaches, the application is unique. To our best knowledge, this is the first approach of applying the BiLSTM network to identify stakeholders' sentiments and to make social media promotion more effective.

Research Gap Analysis

The literature review and summary are in Table 1 demonstrate that sentiment analysis using the BiLSTM network and machine learning models is a well-developed field of research. It is also noticeable that the university's promotional management strategy is an active research field. Applying a sentiment classifier built using the BiLSTM network in developing a data-driven promotional management system is an innovative approach. However, it is a significant research gap that has not been addressed yet. The proposed DPMS has been developed to fill this research gap.

Table 1. A summary of the literature review with comparable features.

Authors	Objective	Method	Dataset	Accuracy
S. Murugaiyan et al. [9]	Aspect-based sentiment analysis	CNN and BiLSTM	Customer Speech Data	93.28%
M. Wankhade et al. [10]	Sentiment classification from sentences	BiLSTM-BERT	SemEval2014 datasets	80.78%
A. S. Talaat et al. [12]	Sentiment classification performance analysis	Hybrid BERT	Apple	91.72%
B. S. Arasu et al. [13]	Socia media data analytics using machine learning	Multiple Models	Private Dataset	86.60%
E. Kongar et al. [14]	Performance analysis using machine learning	AutoML	Private Dataset	82%
S. Asla [16]	Sentiment analysis from tweets	MF-CNN-BiLSTM	Tweet Dataset	91.79%
Z. Gou et al. [17]	Emotion analysis of dialog	BERT-BiLSTM	Real-world Dialog	85.44%
M. Lestandy et al. [18]	Emotion classification and Word2Vec Weighting	CNN-BiLSTM	Emotion Dataset	92.85%

Social Media Marketing (SMM) is now a mainstream marketing strategy. There are multiple machine-learning approaches that have been applied in SMM. However, the core focus of these methods is on data analytics and consumer behavior prediction. There is a research gap in increasing the impression using positive reactions and responses on social media with a data-driven approach. The proposed DPMS guides the social media marketing team in choosing effective content to promote, which minimizes negative reactions, increases positive responses, and contributes to better impressions.

The literature review summary in Table 1 shows that the highest classification accuracy is 93.28%, and this has been achieved by S. Murugaiyan et al. [9]. The success of the proposed DPMS depends on correct prediction from the BiLSTM network. As it is a data-driven approach, incorrect data lead to devastating effects. There is a research scope for optimizing the BiLSTM network architecture to enhance the classification performance. This scope has been utilized in this paper. The well-engineered, properly optimized BiLSTM network of the DPMS achieves an accuracy of 98.66%.

3. Methodology

The proposed methodology illustrated in Figure 1 starts with the data preparation. In this phase, the dataset is collected and then the text data are processed. After that, the processed data are split for training, testing, and validation. Feature engineering is performed next to analyze the appropriate method to vectorize the processed text. After that, an appropriate network is selected through network characteristics analysis. Finally, the network architecture is designed and optimized for training.

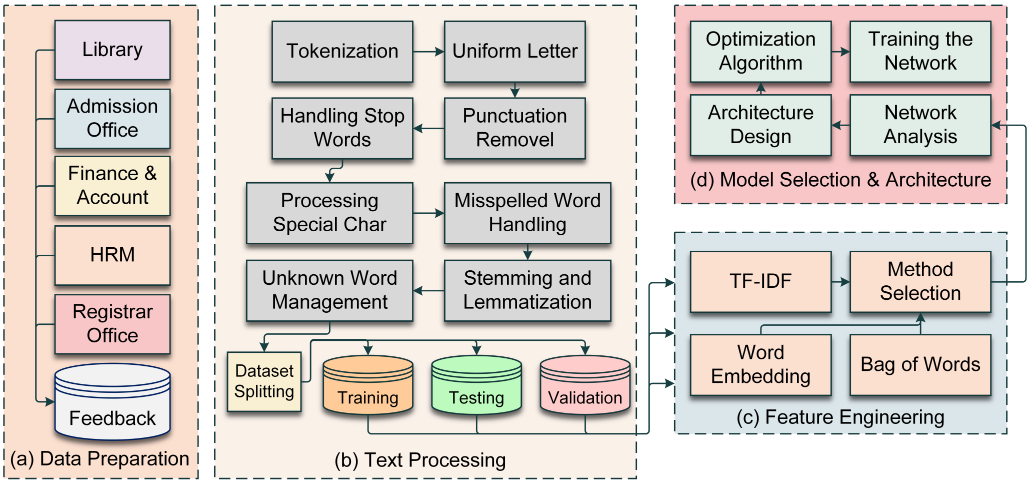

Figure 1. The overview of the methodology developed in this paper.

3.1. Dataset Preparation

One of the crucial steps of the Deep Learning (DL) approach is dataset preparation [19]. A well-engineered DL model collapses if not trained and tested with the appropriate dataset. This experiment uses a dataset that contains university stakeholders' feedback. This section discusses the data collection strategy and processing methods to develop the Data-driven Promotional Management System (DPMS) for higher educational institutions.

3.1.1. Dataset Collection

The data collection method of the proposed DPMS is a unique approach that involves the stakeholders' feedback only. It consists of collection strings that contain opinions, appreciation, concerns, complaints, and other types of feedback, which is the sentiment of the stakeholders. Usually, the core components of a university are faculties, departments, library, admission office, the finance and account division, the Human Resource Management (HRM) division, registrar office, and more [20]. The stakeholders' feedback about the quality of services in these university divisions is the dataset of the proposed DMPS. It is an array of variable strings defined by Equation (1), which contains the sentiments of the stakeholders. These feedback are collected using a traditional Opinion Box hanging outside each university section and Google Form.

$$M[S_{array}] = \sum_{i=0,j=0}^{m,n} (UID, D_{ijk}) \quad (1)$$

In Equation (1), the $M[S_{array}]$ is the memory location of S_{array}, which is the dataset. The elements of this array are ordered pair (UID, D_{ijk}), where UID is the unique identification key for every instance of the dataset and D_{ijk} is the string. Here, i is the university's division, j represents different individuals, and k is the length of the string, which varies according to the number of characters in the string. Each instance is labeled through human inspection, which is defined by Equation (2), which is a set of ordered pairs. Here, $H(S_{array})$ is the human inspection and L is the label of the data D.

$$\{(D, L)\} = \{(S_{array}, H(S_{array}))\} \tag{2}$$

The labeled dataset prepared according to Equation (2) contains a 'string, label' pair. This labeled dataset is used to train the proposed DPMS. The label of the dataset is a set defined by Equation (3) [21].

$$\begin{aligned} &L = \{xx \in \text{Sentiment}\} \\ &\text{where Sentiment = \{Good, Average, Neutral, Poor\}} \end{aligned} \tag{3}$$

The labeled dataset is the raw dataset for the proposed data-driven promotional management system for higher educational institutions. This dataset needs further processing to feed it into the Deep Learning model, to train it and to predict the sentiments of the stakeholders about a particular division of the university [22]. Making promotional decisions based on the predicted sentiment about a particular decision increases the impact of social media promotional posts.

3.1.2. Text Processing

Tokenization

The string of the dataset has been split into multiple tokens, which is the first step of the text processing in this paper. This process has been defined by Equation (4), where $f(w_i)$ is the function responsible for generating tokens. Only unique tokens are allowed after tokenization, and duplicate tokens are removed [23].

$$f(w_i) = w_i \; \forall \; w_i \in T \tag{4}$$

The manual inspection performed on the dataset shows multiple improper stakeholder responses. There are responses with meaningless sentences. In multiple instances, there are spelling mistakes and emojis. Moreover, there is some ambiguous feedback that contains no meaningful information. Furthermore, there are special characters such as punctuation marks, exclamatory symbols, numerical lists, and improper capitalization. It is essential to process this dataset to remove instances that are not useful for training the DL model.

Uniform Letter-Casing

The first step of the text processing of the proposed DPMS is to lowercase the letters. According to the English language standard, capitalizing is meaningful to human readers. However, it carries no information variation for computer-aided systems. Moreover, having lowercase and uppercase letter combinations makes the vector space complicated. That is why at the beginning of text processing, the letters have been made lowercase by following the mathematical model explained in Equation (5) [24].

$$low(D_i) = \begin{cases} \text{lower}(D_i), & \text{if } D_i \in Cap \\ \text{otherwise} \end{cases} \tag{5}$$

In the uniform letter-casing defined in Equation (5), the $low(D_i)$ represents the function where D_i is the dataset, which is the stakeholders' feedback. If D_i belongs to capital letters, the function converts to a lowercase letter.

Punctuation Removal

The punctuation makes the contents more readable. However, it is applicable to human readers. Punctuation marks do not contain any useful features while training DL models and predicting the sentiments through Natural Language Processing (NLP). The punctuation marks removal process is defined by Equation (6) [25].

$$R_p(D_i) = \begin{cases} D_i, & \text{if } D_i \notin P \text{ remove,} \\ \text{otherwise} & \end{cases} \tag{6}$$

In Equation (6), the R_p represents the punctuation removal function, which takes D_i as the input. If the character is not a punctuation mark, the input remains unchanged. However, if it belongs to punctuation marks P, the character is removed.

Handling Stop Words

The stop words (for example, articles, prepositions, conjunctions, auxiliary verbs, etc.), do not carry enough information for Natural Language Processing (NLP). However, their presence in the dataset adds additional vectors in the word vector and increases the complexities of the vector space without making it more effective. That is why removing stop words makes the sentiment classification process used in the DPMS more effective. It is performed using Equation (7) [26].

$$H_s(D_i) = \begin{cases} D_i, & \text{if } D_i \notin S \text{ remove,} \\ \text{otherwise} & \end{cases} \tag{7}$$

$$S_w = \{xx \in \text{Stop-Words}\} \tag{8}$$

The $H_s(D_i)$ represents the filtering function that checks the stop words available in the D_i. The stop word has been listed as a set defined by Equation (8). If any word is found in the set of stop words, it is removed from the D_i string.

Processing Special Characters and Numbers

The dataset used in this experiment contains multiple special characters and numeral lists. These special characters and numbers are not useful in NLP. In this paper, these special characters and numbers have been removed to reduce the vector-space complexity and to make the sentiment classifier faster. The process of handling special characters and numbers is defined by Equation (9). Here, the function $C_{sn}()$ takes the D_i and removes the special characters if $D_i \in A$, where A is the set of special characters and numbers [27].

$$C_{sn}(D_i) = \begin{cases} D_i, & \text{if } D_i \in A \text{ remove,} \\ \text{otherwise} & \end{cases} \tag{9}$$

Handling Misspelled Words

The dataset for this paper is prepared from stakeholders' feedback where there are multiple spelling mistakes. Training the model with misspelled words hampers the overall quality of prediction. The misspelled words have been corrected in the DPMS using Equation (10), where the correct word is c. The purpose is to maximize the probability $P(cw)$. Here, w represents the misspelled word. The mathematical model in Equation (10) follows the mathematical structure of Bayes' theorem [28].

$$\hat{c} = \arg\max_{c \in C} P(cw) = \arg\max_{c \in C} \frac{P(wc)P(c)}{P(w)} \tag{10}$$

Stemming and Lemmatization

The suffixes and prefixes are removed in Natural Language Processing (NLP) practice. It has been maintained in this paper, which is expressed by Equation (11). It is defined as a function, $s : W \to W'$. The set of experimenting words in an instance is w, and w' is the corresponding stemmed word. In Equation (11), $s(w)$ is the stemming function [29].

$$s(w) = \text{stem}(w) \tag{11}$$

The term $s(w)$ is a heuristic function. That means that it does not guarantee the base form of every word. Obtaining inaccurate words after stemming is common, and lemmatization is used to overcome this issue. It is an advanced method that covers the words into their base form. This base form is known as the lemma. It is performed through a morphological process. This part of speech is also considered in lemmatization. The mathematical background of lemmatization is expressed in Equation (12). Because it is a heuristic, stemming may not always give you the best representation of a word's root. Sometimes, inaccurate answers are produced because stemming ignores the context and the part of speech of the word. That is why lemmatization is performed as well. It is a more sophisticated approach to text normalization that reduces words to their base or dictionary form, called a lemma. It involves morphological analysis and considers the context and part of speech of the word in the text. The lemmatization process used in this research is expressed in Equation (12).

$$l(w) = \text{lemma}(w, \text{POS}(w)) \tag{12}$$

In Equation (12), the lemma of the word w is $\text{lemma}(w, \text{POS}(w))$. And the part of speech is denoted by $\text{POS}(w)$.

Unknown Word Management

The dataset contains some unknown words that are written using English letters but that carry meaning in different languages. Some of these words have been replaced where the English meaning is traceable. However, untraceable words have been removed from the dataset. It involves re-sequencing the original letters into new sequences that are meaningful. The process involves the mathematical representation of Equation (13). This function, defined as, $f : S \to S'$, maps the existing character sequence, which is unknown according to the dictionary, and is re-sequenced to S', which is meaningful. If no meaningful sequence is discovered, the existing sequence S is removed [30].

$$f(w_i) = \begin{cases} w_i, & \text{if } w_i \notin R \\ g(w_i), & \text{if } w_i \in R \end{cases} \tag{13}$$

3.1.3. Dataset Splitting

After processing, there are a total of 8810 instances in the dataset. This dataset has been split into the training, testing, and validation datasets, with a ratio of 70:15:15. At this ratio, there are 6167 instances in the training dataset. The training and validation datasets contains 1320 and 1321 instances, respectively. The training dataset has been used to train the BiLSTM network. The validation dataset has been used during the training to validate the learning progress. The testing dataset has been used to test the performance of the network.

3.2. Feature Engineering

The feature engineering of the proposed promotional management system involves extracting features from the processed text data. According to the literature review, Bag of Words (BoW), Word Embedding, and Term Frequency-Inverse Document Frequency (TF-IDF) are widely used feature extraction methods [31]. All three of these methods have been explored in this paper, and the most appropriate method has been adopted in this paper.

Term Frequency-Inverse Document Frequency (TF-IDF)

One of the common NLP feature extraction approaches is Term Frequency-Inverse Document Frequency (TF-IDF), which has been defined by Equation (14). Here, i is the string and j is the word. The term frequency, $x_{i,j}$, represents the weight of a particular word j from the i feedback. The inverse term frequency is $\mathrm{idf}(w_j)$. The mathematical model of TF-IDF is defined in Equation (14) and expresses the feedback from the stakeholders as a vector of term frequencies. The inverse part carries the information related to the importance of the word in the document [32].

$$x_{i,j} = \mathrm{tf}(w_j, d_i) \times \mathrm{idf}(w_j) \tag{14}$$

Bag of Words (BoW)

The Bag of Words (BoW) is another frequently used feature extraction method in NLP. It represents the document as the vector of word frequencies. A vector dimension represents each word in the vocabulary, whereas the value of the vector is the word frequency in the experimenting feedback. The BoW follows the mathematical operation defined in Equation (15) [33].

$$x_{i,j} = \mathrm{freq}(w_j, d_i) \tag{15}$$

In Equation (15), $x_{i,j}$ represents the frequency of words in the stakeholders' feedback. Here, i is the number of frequencies and j is the document. The word is expressed as w_j in $\mathrm{freq}(w_j, d_i)$ and d_i means the ith document.

Word Embedding (WE)

The Word Embedding (WE) method used in this paper has been defined by Equation (16) as a Word2Vec approach. It has been used to map words or phrases to real numbers. These vectors represent the semantic meaning of the words. It also allows for mathematical operation among the words. The embedding function in Equation (16) converts the input word w_j into a vector. Eventually, it converts every word into vector $x_{i,j}$, where i is the vector of the i document [34].

$$x_{i,j} = \mathrm{emb}(w_j) \tag{16}$$

Feature Engineering Method Selection

In this study, the Word2Vec method has been selected as the feature extractor. The DL model used in this paper is a Bidirectional Long Short-Term Memory (BiLSTM) network [35]. The Word2Vec generates vector spaces and aligns the semantically similar words. That means it is a semantic-aware feature extraction method. As a result, the BiLSTM network gains context-aware classification capabilities when the Word2Vec method is used [36]. Moreover, the vector space formed by Word2Vec is richer in information than the one-hot-encoded vectors. It reduces the dimensionality, making it computationally more efficient, and the BiLSTM network learns faster. Additionally, the Word2Vec approach is a continuous representation of the vectors, which makes mathematical operation among words easier. It becomes beneficial when the BiLSTM network is coupled with Word2Vec, because the gates and memory cells of this network rely on mathematical operations [37]. Finally, the BiLSTM network captures patterns from both the forward and backward directions. When combined with the contextual embeddings from Word2Vec, it becomes potent in understanding the nuances in the data. Considering the advantages of using Word2Vec with the BiLSTM network, it has been used as a feature extractor from the dataset to train the DL model.

3.3. Model Selection and Architecture

3.3.1. Selection Process

There is a BiLSTM network at the heart of the proposed DPMS. It is trained to analyze and to classify the sentiments of the stakeholders. The labeled dataset used in this paper has four classes, defined as $C = \{Good, Average, Neutral, Poor\}$. The network classifies the stakeholders' feedback into one of these four classes. Based on this classification, a promotional management system is designed. The users are allowed to express their opinions without any word limitations. That is why the input to the network does not have a fixed length. The BiLSTM network is an ideal choice to work on inputs of variable lengths. Moreover, the user feedback are sequential data, where exploring both the forward and backward directions reveals valuable features for classifying the opinions accurately [38]. Considering these features and necessities, the BiLSTM network has been selected to design the proposed data-driven promotional management system.

3.3.2. Network Architecture

The BiLSTM network used in this paper is illustrated in Figure 2. It is a branch of the Recurrent Neural Network (RNN) that has the capability of processing data in both the forward and backward directions [39]. The network has been designed to handle input strings with a maximum of 2000 nodes. After the input layer, there is an embedding layer. This layer is responsible for converting the input into a feature vector. The vectors then enter the bi-directional layer, where the features are extracted from both the forward and backward directions. These features enter into a global max pooling layer. There is a dense layer after it with 128 nodes. The signals from the dense layer go to the output layer. The output layer has four nodes, each of which is processed through the Softmax function to calculate the probability of the likelihood of a certain class.

The BiLSTM layers have input, forget, and output gates, which are defined by Equations (17), (18), and (19), respectively. Each of these gates is time-dependent. In these equations, x_t is the input to the BiLSTM layer at time t. And the hidden state is defined as h_t at time t. And the output is expressed by o_t at time t.

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \tag{17}$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \tag{18}$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \tag{19}$$

The BiLSTM network stores information in the candidate cell. It is an essential element of the network, which has been calculated using Equation (20). Each candidate cell changes state when it receives input. The state of the candidate cells has been calculated using Equation (21). The weight matrix of the cell is w and the bias is b.

$$\tilde{c}t = \tanh(Wxcx_t + W_{hc}h_{t-1} + b_c) \tag{20}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{21}$$

The output is calculated using Equation (22). The element-wise multiplication operation is performed to obtain the output, where σ is the Sigmoid function and $\odot$ means the element-wise multiplications.

$$h_t = o_t \odot \tanh(c_t) \tag{22}$$

One of the reasons behind using the BiLSTM network is to utilize its capability of both forward and backward propagation to extract learnable features. The forward process is the sequence in a progressive direction, while the backward process does the opposite. The working principle is the same except for their direction, which is defined by Equation (23).

Here, $h_t^{(f)}$ is the output produced by the forward processing, and $h_t^{(b)}$ is the output for backward processing.

$$y_t = [h_t^{(f)}, h_t^{(b)}] \tag{23}$$

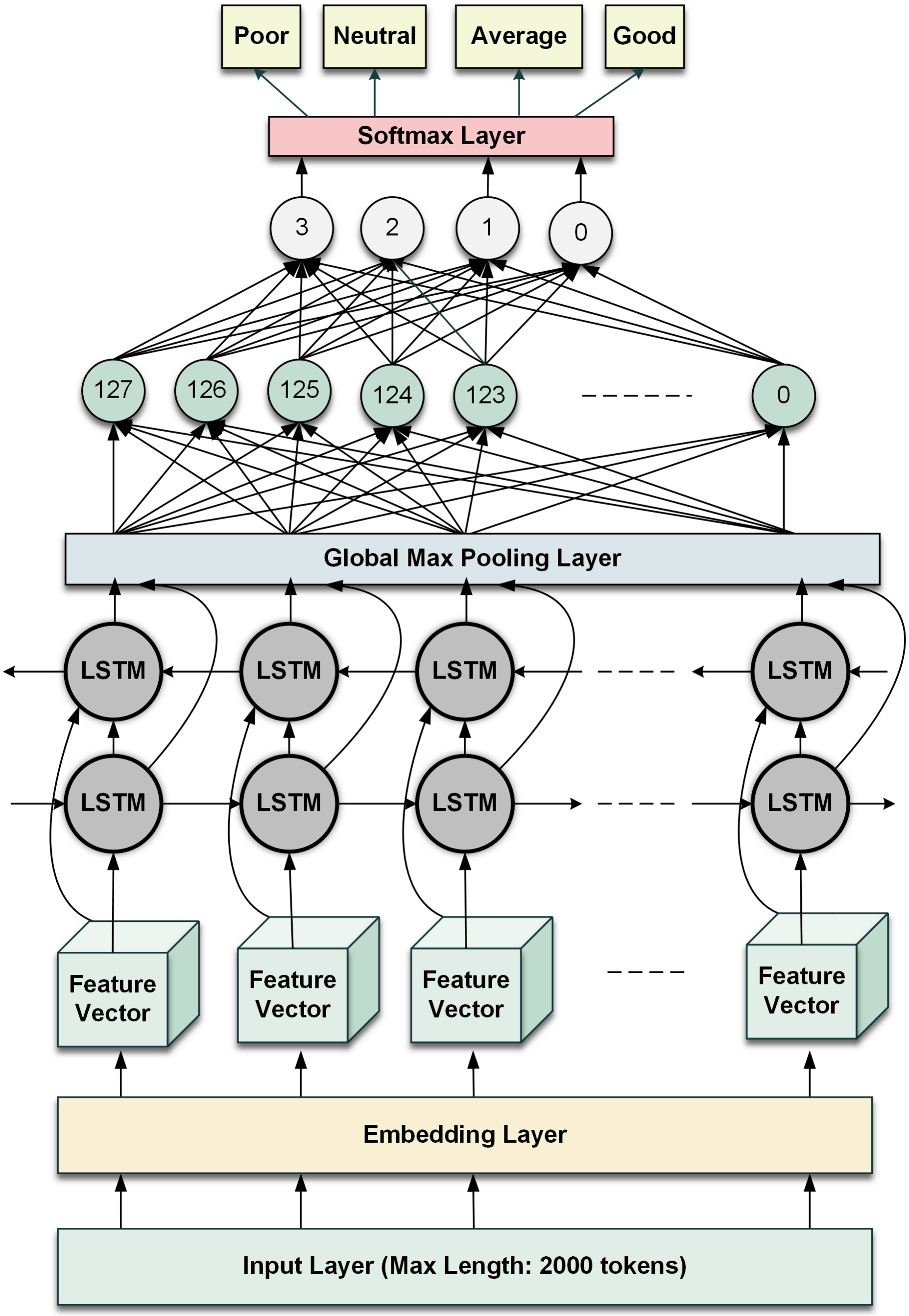

Figure 2. The BiLSTM network architecture.

3.3.3. Optimizer Algorithm

The Adaptive Moment Estimation (ADAM) has been used to optimally train the proposed BiLSTM network to design the data-driven promotional management system for universities. It has been used to utilize the advantages of both momentum and adaptive learning rate-based approaches. In this experiment, the moments have been calculated first using Equations (24) and (25), respectively [40].

$$m_i^{t+1} = \beta_1 m_i^t + (1 - \beta_1) \nabla_{w_i} J(w) \tag{24}$$

$$m_i^{t+1} = \beta_1 m_i^t + (1 - \beta_1) \nabla_{w_i} J(w) \tag{25}$$

The purpose of calculating the moments is to use them to correct the biases so that the BiLSTM network learns the effective features. The process involves two mathematical operations defined by Equations (25) and (26).

$$\hat{m}_i^{t+1} = \frac{m_i^{t+1}}{1 - \beta_1^{t+1}} \tag{26}$$

$$\hat{v}_i^{t+1} = \frac{v_i^{t+1}}{1 - \beta_2^{t+1}} \tag{27}$$

After having the momentum and the corrected biases using that momentum, the optimizer algorithm is ready to update the weights of the dense layer optimally. The better the weight optimization, the better the performance of the network. This important and final phase has been developed using the mathematical structure expressed in Equation (28)

$$w_i^{t+1} = w_i^t - \frac{\eta}{\sqrt{\hat{v}_i^{t+1}} + \epsilon} \hat{m}_i^{t+1} \tag{28}$$

4. Experimental Results and Evaluation

The proposed data-driven promotional management system for universities using Deep Learning on social media has been designed to promote the target university on social media with minimum negative and maximum positive reactions. Every institution has both strengths and weaknesses. When the good side of a university is publicized over social media, the tendency of the audience to give a positive response is higher. The proposed DPMS classifies the sentiment of the stakeholders' feedback and identifies what the stakeholders like or dislike. Based on these findings, the decision to promote content is made.

4.1. Evaluation Metrics

It has been observed from the literature review that Machine Learning and Deep Learning approaches are evaluated using four evaluation metrics [41]. These are: accuracy, precision, recall, specificity, and F1-score [42]. It has been further observed that recall and sensitivity have been used interchangeably [43]. These evaluation metrics are defined by Equations (29), (30), (31), (32), and (33), respectively.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{29}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{30}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{31}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{32}$$

$$\text{F1 Score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}} \tag{33}$$

The accuracy, precision, recall, specificity, and F1-score calculations are dependent on the values of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) [44]. A confusion matrix has been generated by using the test dataset on the BiLSTM network. The TP, TN, FP, and FN have been calculated from the confusion matrix [45].

4.2. Confusion Matrix Analysis

We have generated a confusion matrix from the prediction made by the BiLSTM network on testing data, which is illustrated in Figure 3. The variables in the testing dataset are 'Good', 'Average', 'Neutral', and 'Poor'. Each variable has 350 instances. The confusion matrix analysis for these four classes shows that the overall accuracy of the system is 98.66%.

The analysis of the confusion matrix for the categories Good, Average, Neutral, and Poor reveals an impressively high-performing model, as indicated by a global accuracy of 98.66%. The values of different evaluation metrics are listed in Table 2. The individual class metrics further underscore this performance. The precision, which measures how many of the items are identified as belonging to a particular class actually do belong to that class, is highest for 'Average', at approximately 98.86%, followed closely by 'Poor', at 98.33%. 'Good' and 'Neutral' are also notably high, at approximately 97.49% and 97.79%, respectively. Recall, which assesses how many items of a particular class were correctly identified, is similarly high across all categories, peaking for 'Poor' at 99.04%. Specificity, indicating how well the model correctly identifies negative cases for each class, also maintains high rates—being above 99% for all classes. The F1-score, a balanced metric that considers both precision and recall, remains consistently high, ranging from approximately 97.66% for 'Good' to 98.68% for 'Poor'. These statistics collectively signify a robust model that accurately distinguishes between the Good, Average, Neutral, and Poor classes.

Figure 3. The confusion matrix analysis to find TP, TN, FP, and FN.

Table 2. Performance evaluation metrics values for each class.

Class	Precision	Recall/Sensitivity	Specificity	F1-Score
Good	0.9749	0.9784	0.9933	0.9766
Average	0.9886	0.983	0.9958	0.9858
Neutral	0.9779	0.9779	0.9943	0.9779
Poor	0.9833	0.9904	0.9922	0.9868

The analytical results of individual classes have been illustrated in Figure 4. From the performance visualization, it is evident that none of the values in the evaluation metrics are less than 97%. This proves the outstanding classification performance of the BiLSTM network developed in this research.

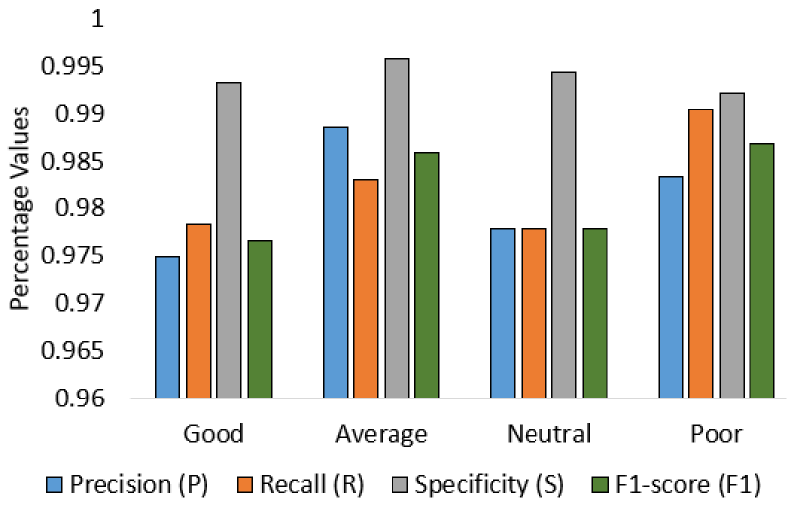

Figure 4. The analysis of individual classes of the confusion matrix.

4.3. AUC-ROC Analysis

The True Positive Rate (TPR) and False Positive Rate (FPR) values offer additional insights into the model's high-performance characteristics, which have been calculated using Equations (34) and (35). The values of TPR and FPR for all classes are listed in Table 3. The TPR, also known as sensitivity or recall, signifies the proportion of actual positives correctly identified by the model. For all four classes—'Good', 'Average', 'Neutral', and 'Poor'—the TPR values are notably high, ranging from approximately 97.79% for 'Neutral' to as high as 99.04% for 'Poor'. These high TPRs indicate that the model has an excellent ability to identify cases within each class correctly.

$$\text{True Positive Rate (TPR)} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{34}$$

$$\text{False Positive Rate (FPR)} = \frac{\text{FP}}{\text{FP} + \text{TN}} \tag{35}$$

Table 3. True Positive Rate (TPR) and False Positive Rate (FPR).

Metrics	Good	Average	Neutral	Poor
TPR	0.9784	0.983	0.9779	0.9904
FPR	0.0067	0.0041	0.0057	0.0078

Conversely, the FPR, which measures the proportion of negative cases incorrectly classified as positive, remains notably low across all classes. The FPR ranges from around 0.41% for 'Average' to about 0.78% for 'Poor', signifying that the model has a minimal tendency to generate false alarms. The graphical representations of TPR and FPR have been illustrated in Figure 5. These extremely low FPR values and high TPR values substantiate the model's accuracy in distinguishing between the classes while minimizing the risk of misclassification. Overall, the TPR and FPR metrics affirm that the model performs exceedingly well across all categories.

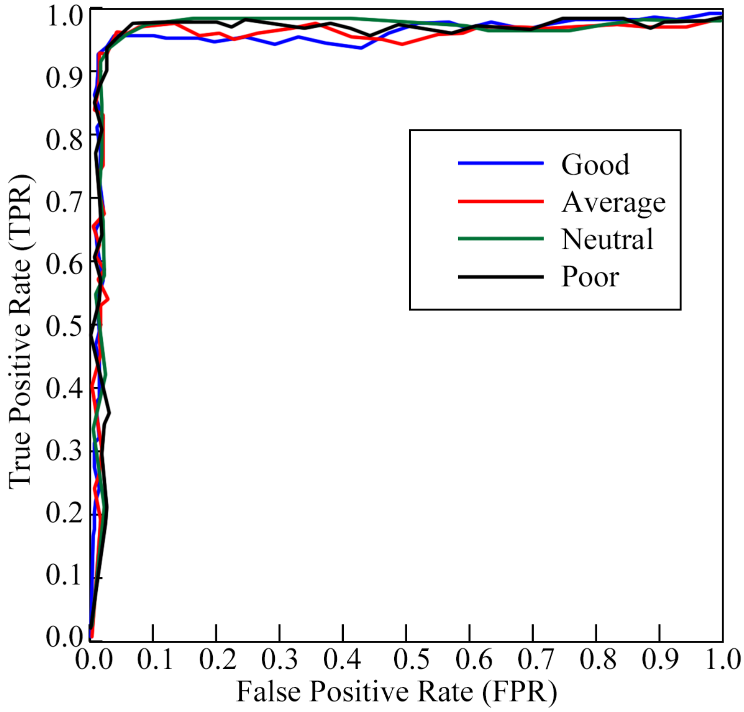

Figure 5. AUC-ROC Analysis.

5. DPMS Implementation and Analysis

The proposed DPMS illustrated in Figure 6, has been implemented in a Virtual Private Server (VPS) with 4 GB primary memory, four CPU cores, 120 GB SSD RAID storage, and 3000 GB bandwidth. The BiLSTM network has been developed in a desktop computer with 16 GB primary memory, 4 GB video memory, and a Core i3-9100 CPU with four cores having a maximum 3.60 GHz clock speed. The PyCharm Community Edition has been used to write the code. The BiLSTM network has been implemented using the TensorFlow library in Python program language.

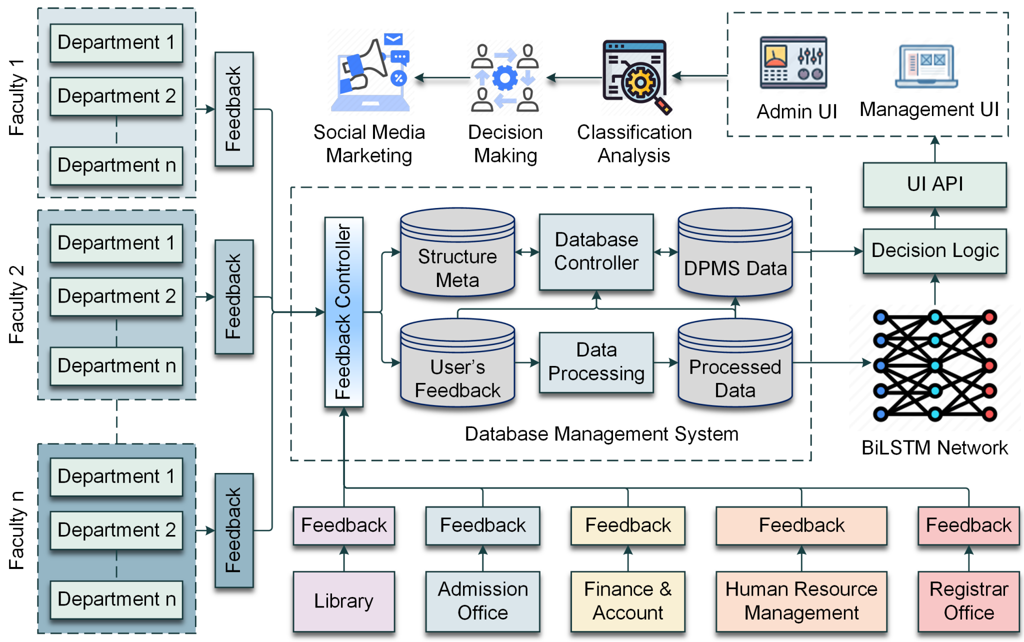

Figure 6. The implementation overview of the proposed HEI PMS system.

5.1. System Component and Workflow

The major components illustrated in Figure 6 of the proposed DPMS systems are the university's faculties, library, admission office, finance and account, human resource management, and registrar office. These divisions have been considered as individual nodes in the development phase. Each node is connected to its separate feedback, which has been created using Google Forms. Initially, the data are stored in multiple Google Sheets, each of which is uniquely identifiable by their nodes' names. A module Feedback Controller has been designed to transfer the data from the Google Sheets to the MySQL database running in the VPS. The Feedback Controller stores the metadata in the Structure Meta table. The raw user's feedback is stored in the User's Feedback table. A processing module that follows the mathematical structure is explained in Section 3.1.2. The processed data are stored in the Processed Data table. The predictions from the BiLSTM network are stored in the DPMS Data table as historical data. The tables are accessed and controlled by the Database Controller that has been specially designed for the DPMS. A BiLSTM network receives the input from the Processed Data table and makes a prediction. The prediction is processed through the decision logic. There are a User Interface (UI) and Application Programming Interface (API) connected with the Admin UI and Management UI. According to the command from the Management UI, the classification is analyzed and the decision is made to promote a particular content on social media.

5.2. Machine vs. Human Analysis

Obtaining the correct prediction from the BiLSTM network is a mandatory requirement for the effectiveness of the proposed DPMS. To evaluate the prediction from the BiLSTM network, the classifications from the network and from human inspection have been compared in this section, side by side. In this analysis, 100 random instances have been used to classify them into one of the four classes, ten times. Concurrently, the feedback has been analyzed by humans to check if the classification from the BiLSTM network is accurate or not. Table 4 lists the comparison between the class identification of the BiLSTM network and human inspection.

Table 4. The comparison between human inspection and classification from BiLSTM network.

	BiLSTM Classification				**Human Inspection**			
Attempt	**Good**	**Average**	**Neutral**	**Poor**	**Good**	**Average**	**Neutral**	**Poor**
1	28	35	22	15	29	34	21	16
2	30	27	25	18	30	28	25	17
3	34	30	15	21	35	29	15	21
4	36	30	18	16	36	29	19	16
5	32	34	19	15	31	35	19	15
6	29	25	26	20	29	25	25	21
7	31	28	26	15	30	29	26	15
8	30	28	19	23	30	28	20	22
9	33	30	17	20	33	30	17	20
10	29	31	25	15	30	30	25	15

Table 4 compares human inspection and BiLSTM network classification across the 'Good', 'Average', 'Neutral', and 'Poor' categories, revealing a high degree of alignment between the two methods. Across ten attempts, the deviations are remarkably minimal, often within a one- to two-unit range, underscoring the model's reliability. For instance, in categories like 'Good' and 'Average', the BiLSTM closely mirrors human assessments, diverging by, at most, one unit. The 'Neutral' and 'Poor' categories show similar minor variances. However, a noticeable discrepancy exists in the eighth attempt for the 'Poor' category, where the BiLSTM model categorizes one more than the human inspection. Despite this, the overarching consistency between the human and machine evaluations affirms the BiLSTM model's robustness and high reliability. Its ability to mimic human-level categorization for various sentiments makes it an invaluable tool in sentiment analysis tasks. The comparison has been illustrated in Figure 7. The visual representation also supports the similarity between BiLSTM classification and human inspection.

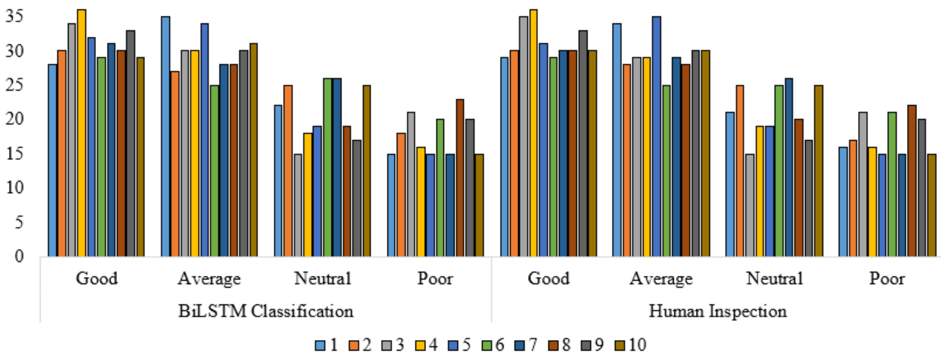

Figure 7. Comparison between the BiLSTM prediction and human observation.

5.3. Real-Life Social Media Response

According to the classification of the BiLSTM network, the experience of the stakeholders is better at the library. And it is poor in the accounts and finance division. The admission office belongs to the average category. And the registrar's office is neutral, according to the feedback. Four social media promotional posts have been created in this experiment and shared, with a difference of seven days. The reactions on social media are listed in Table 5.

Table 5. Reaction in social media toward four types of posts classified by BiLSTM network.

	Positive				Negative		
Classification	**Like**	**Love**	**Care**	**Wow**	**Haha**	**Sad**	**Angry**
Good	119	115	32	25	3	0	5
Average	63	60	56	42	18	7	19
Neutral	24	15	10	5	7	10	5
Poor	15	11	6	1	52	18	95

Table 5 provides an in-depth analysis of social media reactions to four types of posts—Good, Average, Neutral, and Poor—as classified by a BiLSTM network. For the 'Good' posts, most reactions are positive, with 119 'Likes' and 115 'Loves', signifying high levels of audience approval and emotional engagement. Negative reactions are almost negligible, suggesting an effective level of identification for high-quality content. 'Average' posts garner a balanced mix of positive and negative reactions, with 63 'Likes' and 60 'Loves', but also 19 'Angry' and 18 'Haha' reactions, indicating a moderate level of resonance with the audience. 'Neutral' posts elicit relatively muted reactions, with 'Like' leading at 24, and minimal instances of 'Angry' and 'Sad'. However, 'Poor' posts show a distinct pattern, attracting many negative reactions—95 'Angry' and 52 'Haha', juxtaposed against only 15 'Likes' and 11 'Loves'. This strongly suggests that the BiLSTM network's classification correlates well with public sentiment, effectively distinguishing high-quality posts from those that are poorly received. See Figure 8.

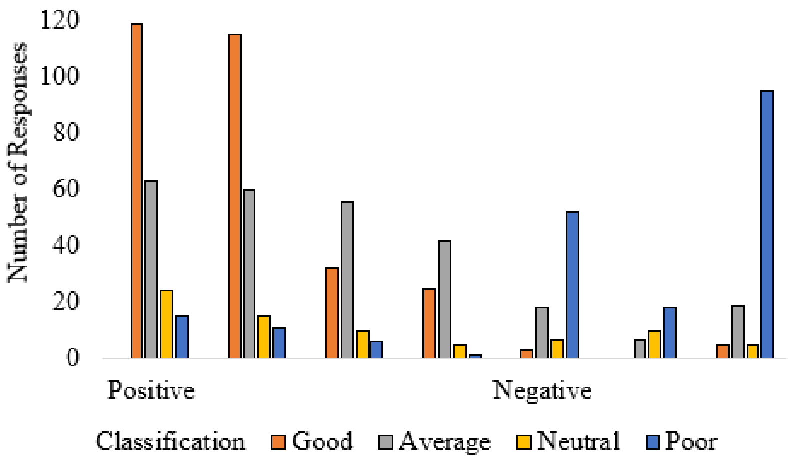

Figure 8. Social media response on four posts classified as Good, Average, Neutral, and Poor by the BiLSTM network.

The graphical illustration in Figure 8 demonstrates that the content classified as 'Good' by the proposed DPMS receives a higher positive response, which results in a higher engagement in impression. It involves the brand value of the university. On the other hand, content classified as 'Poor' by the DPMS receives a higher number of negative reactions and marginal positive reactions. It causes less engagement and impression. It also negatively impacts on the brand value, and students become demotivated to submit admissions to the promoted university. The data presented in Table 5 and the graphical representation in Figure 8 strongly suggest that DPMS-guided promotional activities are more effective than randomly sharing content on social media. The average negative reaction reduction calculated by Equation (36) is 31.25%.

$$N_{avg} = \frac{\frac{1}{k}\sum_{i=1}^{k} Pos_i - \sum_{j=1}^{k} Neg_j}{\sum_{i=1}^{k} Pos_i} \times 100 \tag{36}$$

6. Limitations and Future Scope

The remarkable performance of a data-driven promotional management system does not come without any harmful side effects. Like everything else in this world, it has both strengths and weaknesses. However, instead of considering the weaknesses as limitations for this paper, we have taken them as an opportunity for further development. In this section, the limitations and corresponding future scopes of this paper are discussed.

6.1. Institution-Specific Approach

The DPMS trained for a particular institution may not be applicable to another institution because the organizational structures and quality of different divisions are different. From this context, the DPMS is not robust. For every new institution, the DPMS needs to be trained with the data of that institution. From this context, the DPMS lacks generalization [46]. However, training the proposed BiLSTM network using a dataset constructed from different institutions is a potential solution for overcoming this limitation, which will be analyzed in the future scope of this paper.

6.2. Adversarial Machine Learning (AML) Attack

A successful Adversarial Machine Learning (AML) attack can cause incorrect predictions, resulting in unexpected results from promotional activities [47]. The proposed DPMS has no defense against AML attacks. The success of the DPMS depends on the correct classification from the BiLSTM network. Any manipulation through an AML attack can impact the overall integrity of the system. Although the current version of DPMS is defenseless against AML attacks, the subsequent version will have a defense mechanism against it.

6.3. Dependence on the Quality of the Feedback

The Natural Language Processing (NLP) performance depends on the quality of the training data [48]. In the proposed data-driven promotional management system, the feedback comes directly from the stakeholders. This means that the dataset's quality depends on the quality of the feedback. It makes the entire system vulnerable if an adequate volume of high-quality feedback is unavailable. An ongoing research effort involves figuring out how to improve the quality of users' feedback without altering meaning. This will be published in the future scope of this research.

6.4. Observational Period

The response to the social media posts suggested by DPMS has been observed over a period of 8 weeks. Although the performance during these eight weeks is satisfactory, this observational period is not long enough. The seasonal impact on the promotional posts is still unexplored, and it requires at least a one-year observational period. The researchers of this paper are collecting data, and in the future scope of this paper, the impact of DPMS for the longer observational period will be published.

This paper presents the first version of the proposed DPMS. The researchers behind this experiment are working on the subsequent versions where the limitations identified in this paper will be addressed. Thus, an improved DPMS will be obtained for more effective promotion of the activities of universities.

7. Discussion and Conclusions

Promotional activities are essential for most businesses. Business growth, brand value, revenue flow, and many other important business factors directly relate to promotional activities. A business can suffer from a significant loss and even go bankrupt if there is weakness or major mistakes in promotional activity [49]. The university is no different. Promoting universities to maintain good student intake rates, improve brand value, and to gain popularity is essential [50]. Instead of an intuition or experience-based approach, a data-driven promotional management system is more effective. A BiLSTM-based DPMS has been presented in this paper. Network design and implementation are simple tasks. However, designing the dataset structure and collecting the appropriate data are challenging, and these have been overcome in this paper. The data collection strategy presented in this paper collects feedback from the stakeholders from different university divisions and processes them using the standard Natural Language Processing (NLP) standard. And finally, the dataset is converted into vectors. These vectors are used to train the BiLSTM network, which classifies the feedback into Good, Average, Neutral, and Poor categories. Based on this classification, the content to promote on social media is selected.

The methodology used in this paper achieves an accuracy of 98.66%. The precision, recall, specificity, and F1-score obtained in this paper are 98.12%, 98.24%, 99.39%, and 98.18%, respectively. These are the performances of the BiLSTM network. The stunning performance is revealed when the proposed DPMS is applied to real-world settings. The experiment in real-world settings shows that it increases the impression by 68.75%. After making the decision from the DPMS's classification, the negative reaction to the social media posts was reduced by 31.25%. This remarkable result proves the outstanding performance of the proposed data-driven promotional management system for universities on social media. However, this system suffers from several limitations. The DPMS designed for one institution does not guarantee that it will be effective for others. This is because different institutions have different types of strengths and weaknesses. That is why it is less likely to obtain similar feedback. The observational period of the DPMS experiment is short; it cannot defend against AML attack, and the quality of feedback is uncontrollable. These are a few of the other weaknesses of this approach. This limitation paves the path toward conducting subsequent research to develop a better version of the DPMS that will have a more positive impact on successful social media marketing strategies.

Author Contributions: Conceptualization, M.E.H.; Methodology, M.E.H.; Validation, M.W.; Formal analysis, I.M.; Writing—original draft, N.F.; Writing—review & editing, T.J.; Visualization, A.B.; Supervision, M.W. and A.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the Australian Research Council Discovery Project, Re-Engineering Enterprise Systems for Microservices in the Cloud, under Grant DP190100314.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to organizational privacy.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Colombari, R.; Geuna, A.; Helper, S.; Martins, R.; Paolucci, E.; Ricci, R.; Seamans, R. The interplay between data-driven decision-making and digitalization: A firm-level survey of the Italian and US automotive industries. *Int. J. Prod. Econ.* **2023**, *255*, 108718. [CrossRef]
2. Jami Pour, M.; Hosseinzadeh, M.; Amoozad Mahdiraji, H. Exploring and evaluating success factors of social media marketing strategy: A multi-dimensional-multi-criteria framework. *Foresight* **2021**, *23*, 655–678. [CrossRef]
3. Kang, Y.; Cai, Z.; Tan, C.W.; Huang, Q.; Liu, H. Natural language processing (NLP) in management research: A literature review. *J. Manag. Anal.* **2020**, *7*, 139–172. [CrossRef]
4. Tabassum, A.; Patil, R.R. A survey on text pre-processing & feature extraction techniques in natural language processing. *Int. Res. J. Eng. Technol. (IRJET)* **2020**, *7*, 4864–4867.
5. Vetrova, E.A.; Kabanova, E.E.; Medvedeva, N.V.; Jukova, E.E. Management of Educational Services Promotion in the Field of Higher Education (The Example of "Russian State Social University"). *Eur. J. Contemp. Educ.* **2019**, *8*, 370–377.
6. Pizarro Milian, R. What's for sale at Canadian universities? A mixed-methods analysis of promotional strategies. *High. Educ. Q.* **2017**, *71*, 53–74. [CrossRef]
7. Nedbalova, E. Understanding the Interaction between a University and Promotional Services—A Case Study. Ph.D. Thesis, University of Southampton, Iskandar Puteri, Malaysia, 2015.
8. Elfeky, A.I.M.; Masadeh, T.S.Y.; Elbyaly, M.Y.H. Advance organizers in flipped classroom via e-learning management system and the promotion of integrated science process skills. *Think. Ski. Creat.* **2020**, *35*, 100622. [CrossRef]
9. Murugaiyan, S.; Uyyala, S.R. Aspect-Based Sentiment Analysis of Customer Speech Data Using Deep Convolutional Neural Network and BiLSTM. *Cogn. Comput.* **2023**, *15*, 914–931. [CrossRef]
10. Wankhade, M.; Annavarapu, C.S.R.; Abraham, A. MAPA BiLSTM-BERT: Multi-aspects position aware attention for aspect level sentiment analysis. *J. Supercomput.* **2023**, *79*, 11452–11477. [CrossRef]
11. Agirre, E.; Banea, C.; Cardie, C.; Cer, D.; Diab, M.; Gonzalez-Agirre, A.; Guo, W.; Mihalcea, R.; Rigau, G.; Wiebe, J. Semeval-2014 task 10: Multilingual semantic textual similarity. In Proceedings of the Proceedings of the 8th international workshop on semantic evaluation (SemEval 2014), Dublin, Ireland, 23–24 August 2014; pp. 81–91.
12. Talaat, A.S. Sentiment analysis classification system using hybrid BERT models. *J. Big Data* **2023**, *10*, 110. [CrossRef]
13. Arasu, B.S.; Seelan, B.J.B.; Thamaraiselvan, N. A machine learning-based approach to enhancing social media marketing. *Comput. Electr. Eng.* **2020**, *86*, 106723. [CrossRef]
14. Kongar, E.; Adebayo, O. Impact of social media marketing on business performance: A hybrid performance measurement approach using data analytics and machine learning. *IEEE Eng. Manag. Rev.* **2021**, *49*, 133–147. [CrossRef]
15. Eklund, P. Reinforcement Learning in Social Media Marketing. In *Research Anthology on Applying Social Networking Strategies to Classrooms and Libraries*; IGI Global: Hershey, PA, USA, 2023; pp. 836–853.
16. Aslan, S. A deep learning-based sentiment analysis approach (MF-CNN-BILSTM) and topic modeling of tweets related to the Ukraine–Russia conflict. *Appl. Soft Comput.* **2023**, *143*, 110404. [CrossRef]
17. Gou, Z.; Li, Y. Integrating BERT Embeddings and BiLSTM for Emotion Analysis of Dialogue. *Comput. Intell. Neurosci.* **2023**, *2023*, 6618452. [CrossRef] [PubMed]
18. Lestandy, M.; Abdurrahim, A. Effect of Word2Vec Weighting with CNN-BiLSTM Model on Emotion Classification. *J. Nas. Pendidik. Tek. Inform. JANAPATI* **2023**, *12*, 99–107. [CrossRef]
19. Munappy, A.; Bosch, J.; Olsson, H.H.; Arpteg, A.; Brinne, B. Data management challenges for deep learning. In Proceedings of the 2019 45th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), Kallithea, Greece, 28–30 August 2019; IEEE: Piscateville, NJ, USA, 2019; pp. 140–147.

20. Biglan, A. Relationships between subject matter characteristics and the structure and output of university departments. *J. Appl. Psychol.* **1973**, *57*, 204. [CrossRef]
21. Huang, X.; Jin, G.; Ruan, W. Machine Learning Basics. In *Machine Learning Safety*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 3–13.
22. Shorten, C.; Khoshgoftaar, T.M.; Furht, B. Text data augmentation for deep learning. *J. Big Data* **2021**, *8*, 101. [CrossRef]
23. Webster, J.J.; Kit, C. Tokenization as the initial phase in NLP. In Proceedings of the COLING 1992 Volume 4: The 14th International Conference on Computational Linguistics, Nantes, France, 23–28 August 1992.
24. Zhang, H.; Cheng, Y.C.; Kumar, S.; Huang, W.R.; Chen, M.; Mathews, R. Capitalization normalization for language modeling with an accurate and efficient hierarchical RNN model. In Proceedings of the ICASSP 2022—2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Singapore, 22–27 May 2022; IEEE: Piscateville, NJ, USA, 2022; pp. 6097–6101.
25. Raina, V.; Krishnamurthy, S.; Raina, V.; Krishnamurthy, S. Natural language processing. *Building an Effective Data Science Practice: A Framework to Bootstrap and Manage a Successful Data Science Practice*; Apress: BerNeley, CA, USA, 2022; pp. 63–73.
26. Nothman, J.; Qin, H.; Yurchak, R. Stop word lists in free open-source software packages. In Proceedings of the Workshop for NLP Open Source Software (NLP-OSS), Melbourne, Australia, 20 July 2018; pp. 7–12.
27. Nawab, K.; Ramsey, G.; Schreiber, R. Natural language processing to extract meaningful information from patient experience feedback. *Appl. Clin. Inform.* **2020**, *11*, 242–252. [CrossRef] [PubMed]
28. Kwon, O.; Kim, D.; Lee, S.R.; Choi, J.; Lee, S. Handling out-of-vocabulary problem in hangeul word embeddings. In Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume, Online, 19–23 April 2021; pp. 3213–3221.
29. Khyani, D.; Siddhartha, B.; Niveditha, N.; Divya, B. An interpretation of lemmatization and stemming in natural language processing. *J. Univ. Shanghai Sci. Technol.* **2021**, *22*, 350–357.
30. Tapsai, C.; Rakbumrung, W. Solving Unknown Word Problems in Natural Language Processing. In Proceedings of the International Academic Multidisciplinary Research Conference, Amsterdam, The Netherlands, 8–10 May 2019; p. 204.
31. Wambsganss, T.; Engel, C.; Fromm, H. Improving explainability and accuracy through feature engineering: A taxonomy of features in NLP-based machine learning. In Proceedings of the Forty-Second International Conference on Information Systems, Austin, TX, USA, 12–15 December 2021.
32. Christian, H.; Agus, M.P.; Suhartono, D. Single document automatic text summarization using term frequency-inverse document frequency (TF-IDF). *ComTech Comput. Math. Eng. Appl.* **2016**, *7*, 285–294. [CrossRef]
33. Soumya George, K.; Joseph, S. Text classification by augmenting bag of words (BOW) representation with co-occurrence feature. *IOSR J. Comput. Eng.* **2014**, *16*, 34–38. [CrossRef]
34. Wang, B.; Wang, A.; Chen, F.; Wang, Y.; Kuo, C.C.J. Evaluating word embedding models: Methods and experimental results. *APSIPA Trans. Signal Inf. Process.* **2019**, *8*, e19. [CrossRef]
35. Li, Y.; Dong, H. Text sentiment analysis based on feature fusion of convolution neural network and bidirectional long short-term memory network. *J. Comput. Appl.* **2018**, *38*, 3075.
36. Wu, P.; Li, X.; Ling, C.; Ding, S.; Shen, S. Sentiment classification using attention mechanism and bidirectional long short-term memory network. *Appl. Soft Comput.* **2021**, *112*, 107792. [CrossRef]
37. Yue, W.; Li, L. Sentiment analysis using Word2vec-CNN-BiLSTM classification. In Proceedings of the 2020 Seventh International Conference on Social Networks Analysis, Management and Security (SNAMS), Paris, France, 14–16 December 2020; IEEE: Piscateville, NJ, USA, 2020; pp. 1–5.
38. Marapelli, B.; Carie, A.; Islam, S.M. RNN-CNN model: A bi-directional long short-term memory deep learning network for story point estimation. In Proceedings of the 2020 5th International Conference on Innovative Technologies in Intelligent Systems and Industrial Applications (CITISIA), Sydney, Australia, 25–27 November 2020; IEEE: Piscateville, NJ, USA, 2020; pp. 1–7.
39. Li, H.; Lin, Z.; An, Z.; Zuo, S.; Zhu, W.; Zhang, Z.; Mu, Y.; Cao, L.; Garcia, J.D.P. Automatic electrocardiogram detection and classification using bidirectional long short-term memory network improved by Bayesian optimization. *Biomed. Signal Process. Control* **2022**, *73*, 103424. [CrossRef]
40. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
41. Trivedi, S.; Patel, N.; Faruqui, N. A Novel Lightweight Lung Cancer Classifier Through Hybridization of DNN and Comparative Feature Optimizer. In *Hybrid Intelligent Systems, Proceedings of the 22nd International Conference on Hybrid Intelligent Systems, Online, 13–15 December 2022*; Springer: Cham, Switzerland, 2022; pp. 188–197.
42. Achar, S.; Faruqui, N.; Bodepudi, A.; Reddy, M. Confimizer: A Novel Algorithm to Optimize Cloud Resource by Confidentiality-Cost Trade-off using BiLSTM Network. *IEEE Access* **2023**, *11*, 89205–89217. [CrossRef]
43. Faruqui, N.; Yousuf, M.A.; Whaiduzzaman, M.; Azad, A.; Alyami, S.A.; Liò, P.; Kabir, M.A.; Moni, M.A. SafetyMed: A Novel IoMT Intrusion Detection System Using CNN-LSTM Hybridization. *Electronics* **2023**, *12*, 3541. [CrossRef]
44. Vujović, Ž. Classification model evaluation metrics. *Int. J. Adv. Comput. Sci. Appl.* **2021**, *12*, 599–606. [CrossRef]
45. Patel, N.; Trivedi, S.; Faruqui, N. An Innovative Deep Neural Network for Stress Classification in Workplace. In Proceedings of the 2023 International Conference on Smart Computing and Application (ICSCA), Hail, Saudi Arabia, 5–6 February 2023; IEEE: Piscateville, NJ, USA, 2023; pp. 1–5.

46. Suciu, O.; Marginean, R.; Kaya, Y.; Daume III, H.; Dumitras, T. When does machine learning FAIL? generalized transferability for evasion and poisoning attacks. In Proceedings of the 27th USENIX Security Symposium (USENIX Security 18), Baltimore, MD, USA, 15–17 August 2018; pp. 1299–1316.
47. Trivedi, S.; Tran, T.A.; Faruqui, N.; Hassan, M.M. An Exploratory Analysis of Effect of Adversarial Machine Learning Attack on IoT-enabled Industrial Control Systems. In Proceedings of the 2023 International Conference on Smart Computing and Application (ICSCA), Hail, Saudi Arabia, 5–6 February 2023; IEEE: Piscateville, NJ, USA, 2023; pp. 1–8.
48. Dima, A.; Lukens, S.; Hodkiewicz, M.; Sexton, T.; Brundage, M.P. Adapting natural language processing for technical text. *Appl. AI Lett.* **2021**, *2*, e33. [CrossRef]
49. Hayran, C.; Ceylan, M. Impact of social media brand blunders on brand trust and brand liking. *Int. J. Mark. Res.* **2023**, *65*, 466–483. [CrossRef]
50. Mathur, M.; Lawrence, D.; Chakravarty, A. Leveraging consumer personality and social media marketing to improve a brand's social media equity. *Int. J. Consum. Stud.* **2023**, *47*, 1076–1094. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.