

Article

Using the AraBERT Model for Customer Satisfaction Classification of Telecom Sectors in Saudi Arabia

Sulaiman Aftan ^{1,*}  and Habib Shah ² ¹ Department of Computer Science, Texas Tech University, Lubbock, TX 79709, USA² Department of Computer Science, College of Computer Science, King Khalid University, Abha 62529, Saudi Arabia

* Correspondence: sulaiman.aftan@ttu.edu

Abstract: Customer satisfaction and loyalty are essential for every business. Feedback prediction and social media classification are crucial and play a key role in accurately identifying customer satisfaction. This paper presents sentiment analysis-based customer feedback prediction based on Twitter Arabic datasets of telecommunications companies in Saudi Arabia. The human brain, which contains billions of neurons, provides feedback based on the current and past experience provided by the services and other related stakeholders. Artificial Intelligent (AI) based methods, parallel to human brain processing methods such as Deep Learning (DL) algorithms, are famous for classifying and analyzing such datasets. Comparing the Arabic Dataset to English, it is pretty challenging for typical methods to outperform in the classification or prediction tasks. Therefore, the Arabic Bidirectional Encoder Representations from Transformers (AraBERT) model was used and analyzed with various parameters such as activation functions and topologies and simulated customer satisfaction prediction takes using Arabic Twitter datasets. The prediction results were compared with two famous DL algorithms: Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN). Results show that these methods have been successfully applied and obtained highly accurate classification results. AraBERT achieved the best prediction accuracy among the three ML methods, especially with Mobily and STC datasets.

Keywords: BERT; AraBERT; deep learning; customer satisfaction classification



Citation: Aftan, S.; Shah, H. Using the AraBERT Model for Customer Satisfaction Classification of Telecom Sectors in Saudi Arabia. *Brain Sci.* **2023**, *13*, 147. <https://doi.org/10.3390/brainsci13010147>

Academic Editor:
Muthuraman Muthuraman

Received: 12 December 2022
Revised: 6 January 2023
Accepted: 11 January 2023
Published: 14 January 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

Customers play an essential role in the business of various small, medium, and large enterprises (SMEs). These businesses or industries can reach the top or bottom based on customer relations, loyalty, trust, support, feedback, opinions, surveys, and other comments, individually or combined. It is very significant to understand customers' needs and comfort levels in commercial and individual-based industries, i.e., customer satisfaction during or after service utilization. Researchers have used different customer feedback acquisition and prediction methods, such as social platforms, electronic surveys, calls, emails, online mobile applications, and websites [1]. Based on sharing customer feedback, comments, advice, suggestions, recommendations, and views, the quality and quantity of services can be improved and extended [2].

Telecommunication is one of the crucial global fields that can play a significant role in each sector, such as business, defense, investment, production, and individual. The fast, reliable, secure, and accurate service can increase the service quality of the corresponding communication companies. Therefore, the prediction of customer feedback is essential to the country's development. Various computer science, theoretical and mathematical, and statistical techniques have been proposed and simulated to accurately predict customer satisfaction so that service quality can be improved effectively according to customer needs and expectations [3,4]. In the Saudi Arabian region, customer feedback is taken very

seriously in companies' government and private sectors. Different departments have been established to supervise and resolve customer complaints in various sectors using different strategies and channels, which has built an extremely competitive telecom market [5]. The Saudi telecom industry is rapidly changing in technological developments, service delivery, competitive landscape, and telecommunications services expansion in the non-traditional telecom services sector. These include managed infrastructure, data center/colocation, and cloud services. Saudi Telecom Company (STC, Saudi Arabia), Integrated Telecom Company (ITC, Saudi Arabia), Saudi Mobily Company (Etisalat, Saudi Arabia), Zain, Virgin, and Go Telecom are the most famous [6]. Saudi Arabia nations are among the most populous nations in the Gulf Cooperation Council (GCC) region, where most of the population comprises young people. These young nations believe in and utilize advanced technology for education, research, business, production, and other sectors. The high-speed 5G network and COVID-19 have increased the future impact of uncertain business situations [7]. The STC, Mobily, and Zain won the 5G awards in 2021. Also, according to the source, around 11 million people are using the Twitter platform on smartphones and computers, while the percentage of users is increasing rapidly due to the population and interest [8].

These three companies are very serious about taking customer feedback and comments on various channels such as Twitter, Facebook, and websites. Suggestions, words, issues resolution, and company stock value prediction are significant in classifying customer loyalty. Previously, various techniques have been used to predict customer feedback on social media platforms [9]. Twitter in Saudi Arabia is also widely used as a source of information to predict the financial market, movements in stock markets, and others [1]. Various researchers have utilized Arabic Twitter data to forecast the outcomes of the corresponding company and provide the best analysis to customers through the emerging sentiment analysis approaches based on advancements in Natural Language Processing (NLP) and text analytics techniques to identify and evaluate the opinions of users expressed in their tweets [10].

According to the Global Competitiveness Report of the World Economic Forum, people in Saudi Arabia are included in the top 36 competitive nations out of 140 countries. Therefore, their business market, production, customer relationship, quality and honesty, and demand are among the top attractive business models in the Middle East and an open market [11]. Apart from computer science and engineering-based problems, the various novel ML and DL techniques have played an essential role in the world economy, business, customer relations, satisfaction, prediction of the trend of customers over the industry concerning time and quality, and quantity [12–14]. However, it is difficult to predict accurate values due to the challenges of analyzing the customer views and understanding the real meaning and classification of this feedback.

Physiological researchers have simulated the various neural network models extracted from the human brain processing systems to create intelligent machines that can perform tasks that usually require human intelligence, such as visual perception, speech recognition, decision-making, and language translation [15]. Machine learning techniques are often used in conjunction with neural networks, and both are inspired by how the brain processes information and learns from experience [15,16]. Overall, the study of the brain and how it works has significantly influenced the development of AI, and the structure and function of the brain inspire many of the techniques used in AI.

Apart from others, the most widely used technique is Sentiment analysis in various product domains using the customer online feedback dataset to identify the positive and negative suggestions, complaints, and others about the product, company, and others. In computer science, classification such as speeches, videos, customer feedback, images, and text are the most common and complex tasks in ML and other typical approaches [17,18]. Like NLP, which can analyze the various types of data for multiple tasks such as sentiment analysis, cognitive assistant, spam filtering, lexical (structure) analysis, parsing, semantic analysis, discourse integration, pragmatic analysis, detecting fake news, detection of false

income, and various real-time types of language translation [19–21]. According to [22], different ML techniques have been used and simulations to predict telecommunication companies' customer satisfaction based on Arabic tweets.

Generally, various methods based on DL, such as Convolution Neural Networks (CNN) [23], Recurrent Neural Networks (RNN) [24], Hierarchical Attention Networks (HAN) [25], Support Vector Machine (SVM) [26], Residual Learning with Simplified CNN Extractor [27], distant, subjective supervision [28], adaptive recursive neural network [29], Random Forest (RF), Decision Tree (DT) [30], Bidirectional Long Short-Term Memory (Bi-LSTM), a hybrid of CNN and Bi-LSTM, Naive Bayes (NB) [31], Emotion Tokens, BiGRU-CNN model [32], Improved Negation Handling, and other effective intelligent methods for classification of the Turkish, Chinese, Thai, Covid, business, and medical-based Twitter datasets for sentiment analysis [33,34]. For Arabic language tweets, in the datasets analysis for various tasks such as classification or prediction, the researchers have used such Deep Attentional Bidirectional LSTM, Chi-Square and K-Nearest Neighbor, Convolutional Neural Networks, Narrow Convolutional Neural Networks (NCNN), CNN and RNN, Bidirectional LSTM, SVM, KNN, Decision Trees, NB, and others for Arabic Sentiment Analysis using the Twitter dataset for solving different tasks [35,36].

Predicting the Saudi stock exchange market, including STC datasets, VM, KNN, and Naive Bayes methods, has been simulated successfully with higher accuracy of 97.10 and 95.71% using SVM Precision and Recall phases, respectively [37]. It has been found that DL outperforms typical ML in various performance measures [38]. However, the classification/prediction accuracy can be increased using the emerging DL-based model such as BERT, which will be discussed next.

DL methods have also shown promising results in classification, prediction, and sentiment analysis in the last decades [39]. Especially in sentiment analysis, DL methods outperform the bag-of-words approach in feature generation. A few ways with limited datasets were found to get accurate simulation results in Saudi Telecommunications Companies, which are not enough for future trends, classifications, and service improvement based on customer tweets. For this purpose, deep learning-based methods such as CNN, RNN, and the recent AraBERT techniques have been proposed to predict the accurate results of these critical companies named STC, Zain, and Mobily. The AraBERT model has been chosen because it has been trained on a large corpus of Arabic text and can be used for various NLP tasks such as language translation, text classification, and question answering. It can also be fine-tuned for specific tasks, such as sentiment analysis or named entity recognition [40].

The rest of the paper is organized into five sections, namely: related works that will summarize the previous work in the corresponding area; the three Deep Learning methods (CNN, RNN, and the recent AraBERT) with a short intro and parameters; proposed simulation methodology; simulation results and discussion; and conclusion to finish.

2. Deep Learning Methods

Sentiment analysis (SA) is one of the complex tasks of computationally identifying and categorizing the opinion of various parties expressed in different text formats. It has various contributions in multiple areas, such as forecasting market movements, quality prediction, and improvements based on sentiment in various platforms such as blogs, news, social media posts, comments, and ratings [41]. Furthermore, the complexities increase when these comments are in the form of local language, feelings, or emotions.

Based on sentiment analysis, the identifications of various customer satisfaction and dissatisfactions can be easily obtained with multiple classes and lead to recommender systems for other customers. For large global businesses, the number of customers is increasing rapidly, along with their assessments, comments, and suggestions. The conventional approaches cannot manage and identify the future risk of such significant stakeholders' views; therefore, computer tools are needed for such complex analyses. Various computational and mathematical tools have been proposed to simulate the identifications of accurate

feedback and forecasting values with these social and non-social sentiments. These methods are Naïve Bayes (NB) classifier, Long Short-Term Memory (LSTM), CNN, and other Deep Learning techniques that have established new success histories for solving various complex computer science-related problems [42–46]. Some of the famous types of ML algorithms have been mentioned in Figure 1.

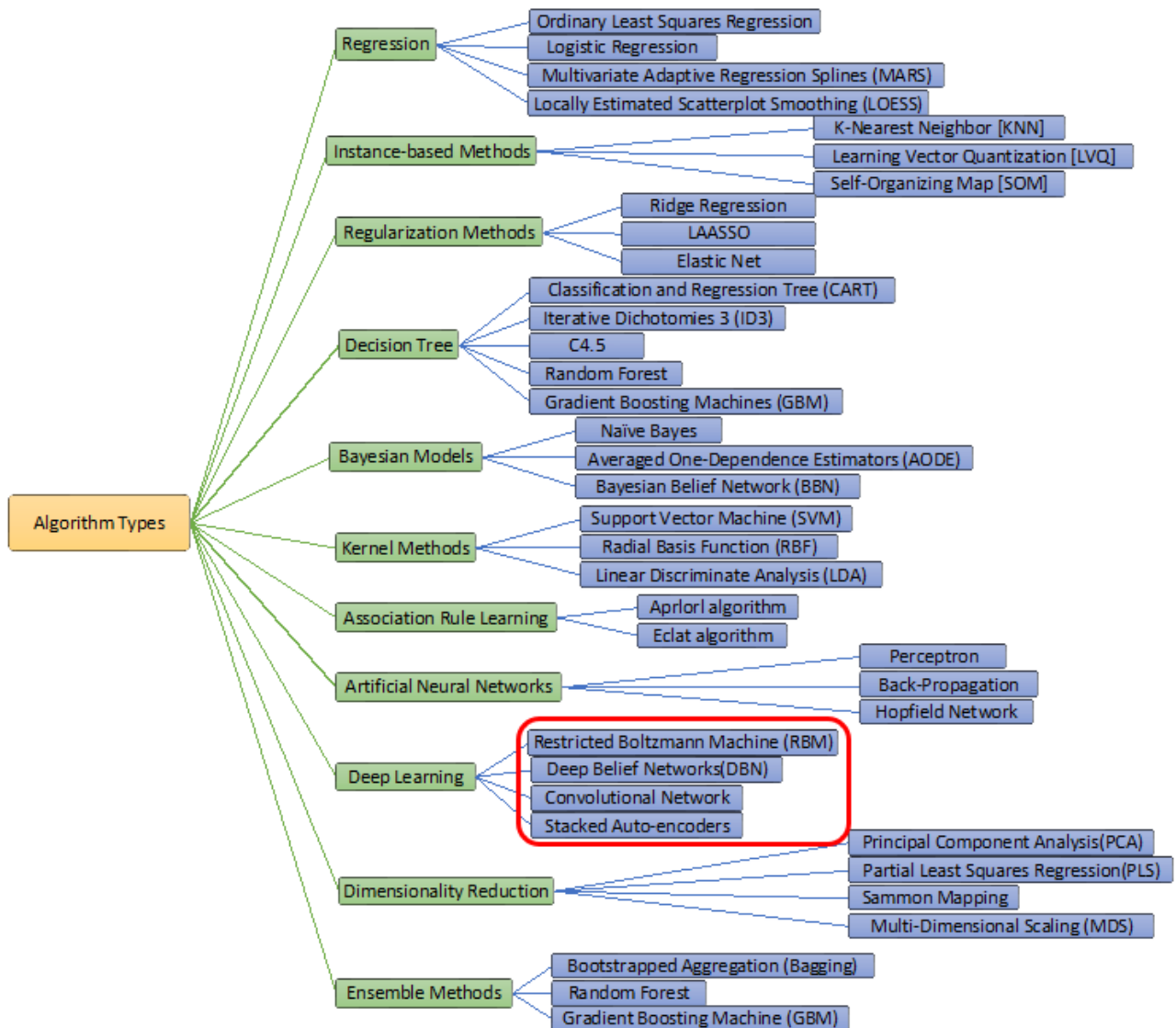


Figure 1. The various types of ML algorithms.

The automated and parallel way of human processing methodologies of DL has increased the effectiveness of these methods and the motivation of researchers from various backgrounds. Many new, improved, and hybrid DL algorithms have been introduced, used, and published in high-quality journals with an outstanding performance from typical methods. Also, the researchers used various DL techniques to indicate the relationship between multiple companies and their customers based on the feedback, quality, comment, and surveys in different areas. Along with, text classification analysis in NLP these three methods have gotten more attention than others based on highly accurate results, which have played an essential role in various domains such as business and customer relations, social impacts on future trends, etc. [19,47,48]. The process includes a pretreatment, feature extraction, selection of sentiment classification algorithm, and sentiment classification performance measures in the mentioned sentiment classification. The three classification algorithms, CNN, RNN, and BERT, have been explained in the following sections.

2.1. Convolutional Neural Network

Convolutional Neural Network (CNN), which was initially developed in the Neural Network Image Processing Community (NNIPC), is a famous kind of Feed-forward Neural Network (FNN) with a profound structure and has shown outstanding simulation results in various tasks, particularly in NLP tasks, such as sentence analysis of various languages in different applications. Multiple types of CNN, including the typical model, have been an important focus of research as they can be applied to complex problems involving time-varying patterns [49–51]. The standard CNN involves two operations, which feature extractors, convolution and pooling; the obtained output is then associated with the following [23]. CNN can interpret spatial data through the convolution layers (CL). A CL has various filters or kernels, which it learns to extract specific features from the corresponding dataset. The kernel is a 2D window that has sided over the convolution operation's input data. We use temporal convolution in our experiments to analyze sequential data like tweets. The typical CNN architecture is used to simulate the Arabic tweets datasets of STC, Zain, and Mobily for classification purposes, as given in the following Figure 2.

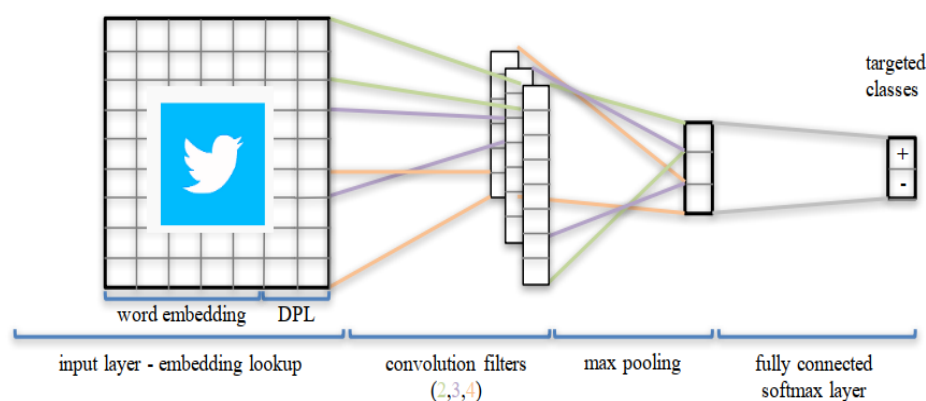


Figure 2. Typical CNN architecture adopted for the Twitter dataset classification purpose.

The different layers of CNN are: Convolutional Layer (the first layer used to extract the features from the given dataset using mathematical operation), the Pooling Layer (using pooling operations to decrease the computational costs), the Fully Connected Layer (consists of the weights and biases values and neurons for connection to others layers), Dropout (to overcome the overfitting problem, some neurons can be dropped), and Activation Functions (to produce the output in the desired template with the suitable functions), which are considered in the typical model.

Previously, CNN and its various versions have been successfully used to classify and predict different business models based on customer feedback and reviews. In 2021, the CNN model was used, along with RNN and RoBERT, to predict the rating given to the products [14]. The simulation results obtained by CNN for Tweet sentiment analysis, which contains 377,616 geotagged tweets, achieved the highest accuracy of 66.0%, based on the one text feature. In contrast, the highest accuracy of 78.0% is achieved using a combination of text and count of nearby location categories features. This accuracy has been increased to 74 and 78 % by using the CNN using a pre-trained 6B GloVe model. It is noted that the CNN using the pre-trained 27B GloVe model achieved the highest accuracy, 83.9% and 94%, respectively, for the same datasets [52]. The hybrid version of CNN and LSTM, called the ensemble model, also successfully used for the classification of the Arabic Twitter dataset by [53], achieves an F1-score of 64.46%, which outperforms the state-of-the-art deep learning model's F1-score of 53.6%, which is higher than CNN and LSTM versions.

Another hybrid version of the Convolutional Neural Network and Differential Evolution Algorithm called DE-CNN is proposed and simulated on various Arabic twitter datasets, achieving high accuracy and being less time-consuming than the state-of-the-art algorithms [54]. Using five types of the Twitter dataset, Deep CNN has been simulated

successfully and obtained outstanding results for sentiment classification [55]. Using several techniques to classify churn prediction in the telecommunication industry, the CNN algorithm showed higher precision with a value of 97.78% [47]. In many famous significant airlines worldwide, CNN outperformed in analyzing tweets extracted based on customers' experiences [12].

On the other hand, a study used deep learning technology to evaluate the torsional strength of Reinforced Concrete (RC) beams. The data-driven model is based on a 2D convolutional neural network (CNN) and uses information such as the beam width, concrete compressive strength, etc., in the model inputs. An improved bird swarm algorithm (IBSA) was used to optimize the hyperparameters of CNN, which was then tested using a dataset of 268 groups of lab tests of RC beams. The results showed that the proposed 2D CNN outperforms other machine learning models, building codes and empirical formulas in evaluation metrics [56]. Diagnosing surface cracks of concrete structures is essential for assessing the safety of a structure. However, traditional methods for doing this are often time-consuming and not very accurate. This paper proposes a new way of identifying surface conditions of concrete structures using a computer vision-based automated method. This method uses different convolutional neural networks (CNNs), which are pre-trained and can be used to make predictions. A modified Dempster-Shafer algorithm combines different CNN results to get more accurate predictions and creates a more reliable result. This method is checked with different types and noise levels and has been tested in real-world scenarios, which shows that it is potentially very accurate in identifying surface cracks of concrete structures [57].

2.2. Recurrent Neural Networks

The Recurrent Neural Network (RNN), developed in 1990, is a practical simulation resulting in modeling, classification, and other complex tasks in science, engineering, medical, and industrial areas. Aside from the others, RNN is successfully used in NLP applications such as text classifications and analysis based on various data acquisitions. The First order RNN uses context units to store the output of the state neurons from the computation of previous time steps [58]. One of the famous RNN architectures among NLP researchers is Long-Short Term Memory (LSTM). Figure 3 is proposed to simulate the three mentioned datasets for the classification task. The LSTM is simulated here because it is outstanding in solving problems such as tagging, sequence-to-sequence predictions, language modeling, and other complex computer science and engineering issues [59].

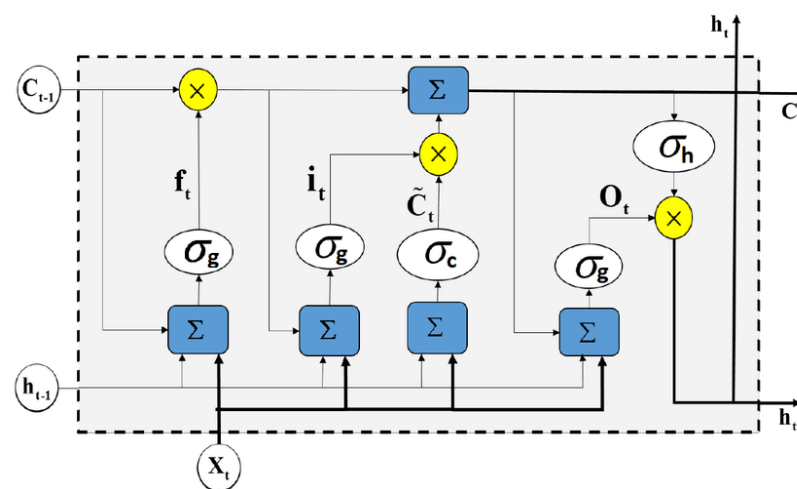


Figure 3. LSTM architecture cell.

Each LSTM has a memory cell, input gate (i_t), output gate (O_t), a forget gate (f_t), and hidden state (h_t) in a classical recurrent neural network (RNN), where the typical equation is:

$$S_t = \tanh(U_{x_t} + W_{s_{t-1}}) \quad (1)$$

$$Y = \text{Softmax}(V_{s_t}) \quad (2)$$

This paper will simulate the RNN with LSTM, expressed in the following step-by-step equations from the input to output gate: below.

Equation of Input Gate of RNN (LSTM):

$$i_t = \sigma(W_i h_{t-1} + U_i x_t + b_i) \quad (3)$$

$$\tilde{C}_t = \tanh(W_h h_{t-1} + U_h x_t + b) \quad (4)$$

Equation of Forget Gate of RNN (LSTM):

$$f_t = \sigma(W_f h_{t-1} + U_f x_t + b_f) \quad (5)$$

Equation of Memory State of RNN (LSTM):

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (6)$$

Equation of Output Gate of RNN (LSTM):

$$O_t = \sigma(W_o h_{t-1} + U_o x_t + b_o) \quad (7)$$

Using seven models on 12 different forecasting problems, RNN (LSTM) obtained the most accurate forecasts [60]. In Customer Lifetime Value (CLV), RNN was used for churn prediction and performed better than other algorithms [61]. Moreover, the recurrent neural network results improved the accuracy of the financial and social network Stock Twits [62]. The various RNN models have been applied to client loyalty number (CLN) applications. Furthermore, we have obtained state-of-the-art results in customer prediction, recency, frequency, and monetary (RFM) variables. Here, the RNN will be used to classify the Twitter dataset of Saudi telecommunications companies along with the CNN and BERT models.

2.3. BERT Model

The latest model, BERT, adopts the structure of a transformer, which includes multiple encoded layers, proposed in 2018, and has shown its advantages on many NLP tasks, like inference and semantic understanding, classification, etc. [63]. BERT is a bidirectional DL Model that looks to text from the two sides, left and right, rather than one side. It is a pre-trained model for various tasks, feature-based, fine-tuning, etc. The performance of BERT depends on the nature of the datasets, tasks, and Encoder (self-attention and Decoder) transfer parts.

Transformer structure is fast to process words simultaneously, and the context of words is better learned as they can understand the context from both directions simultaneously. The transformer includes two major components: the encoder and decoder layers. The Google AI developer, back in 2018, took advantage of transformer components by working on the encoder layers (self-attention and feed-forward neural network (FFNN), and proposed BERT: Bidirectional Encoder Representations from Transformers. BERT is intended to jointly condition both the left and right context in all layers to pre-train deep bidirectional representations from an unlabeled text [63].

Training BERT should be in two phases. The first phase is pre-training, where the model understands language and context, and the second phase is fine-tuning, where the model learns the language and how to solve the problems. BERT can solve many issues, such as Neural Machine Translation, Question Answering, Sentiment Analysis, Text

summarization, etc. Also, BERT achieved state-of-the-art outcomes in more than 11 NLP tasks [63].

The first phase is pre-training unsupervised datasets simultaneously by two techniques. The first is masking out some percentage of the words in the input and then conditioning each word bidirectionally to predict the masked words with Masked Language Modeling (MLM). In this stage, they used WordPiece embedding, which is subword tokenization that enables the model to process the unknown words by decomposing them into known subwords [63].

As shown in Figure 4, every embedding token has a particular classification token at the beginning of every sentence [CLS] and uses [SEP] to separate them. Also, to help the model differentiate among the different sentences, they add a learned embedding indicating sentence A or sentence B is added to each token, which is segment embedding. In this process, E has been marked as input embedding at the first hidden vector of [CLS] token as $C \in \mathbb{R}^H$. Furthermore, at the end of the final hidden subscriber, the i th input token can be $T_i \in \mathbb{R}^H$. [2]

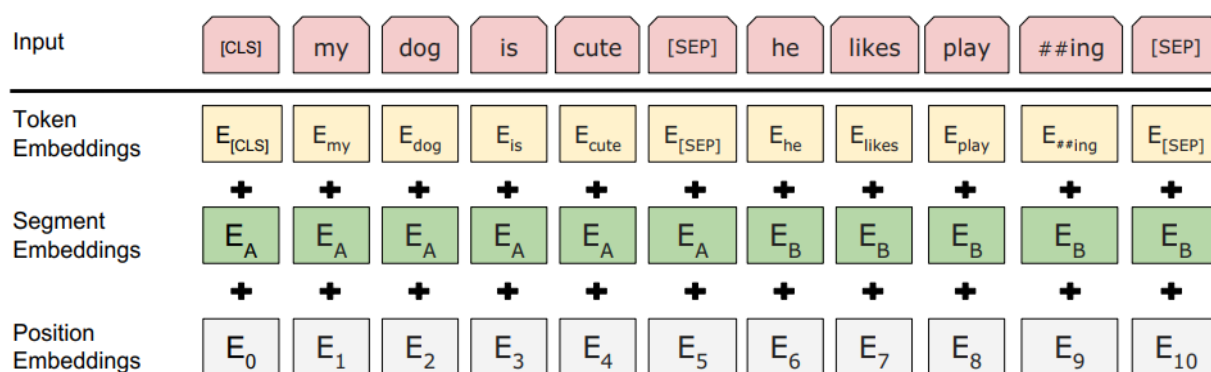


Figure 4. BERT input representation. The input embeddings are the sum of the token embeddings, the segmentation embeddings, and the position embeddings.

The second technique is understanding the relationship between sentences, which is essential in this model on Question Answering (QA) and Natural Language Inference (NLI), by applying the Next Sentence Prediction (NSP) classification task to predict whether sentence B immediately follows sentence A. For each pre-training example, the sentences A and B are chosen randomly from the corpus, with 50% of the time B being the sentence that follows A (labeled as IsNext) and 50% being a random sentence (labeled as NotNext) [63].

The capability of the transformer's self-attention mechanism to simulate many downstream tasks, whether they entail single texts or pairs of texts, by simply swapping out the appropriate inputs and outputs makes fine-tuning simple. As encoding a concatenated text pair with self-attention effectively involves bidirectional cross-attention between two sentences, BERT uses this method to combine both stages. Feed the inputs and outputs particular to each task into BERT and fine-tune all the parameters end-to-end [63].

The second phase is fine-tuning the model, where the pre-trained parameters are used to initialize the BERT model and labeled data from the downstream tasks is used to fine-tune each parameter. Figure 5 represents the overall pre-training and fine-tuning procedures for BERT. In both pre-training and fine-tuning, the same architectures are used as given.

By pre-training the model, we have to minimize the loss; word vectors T_i have the same size and are generated simultaneously. So, we have to pass these word vectors into a fully connected layered output with the same number of neurons equal to the number of tokens in the vocabulary and apply a SoftMax activation. In this way, the word vector is converted to distribution, and the actual label of this distribution would be one hot encoded vector for the exact word. So, we compare these two distributions and then train the network using the cross-entropy loss. On the other hand, the [MASK] token does not

appear during the fine-tuning to prevent a conflict between pre-training and fine-tuning the model [63].

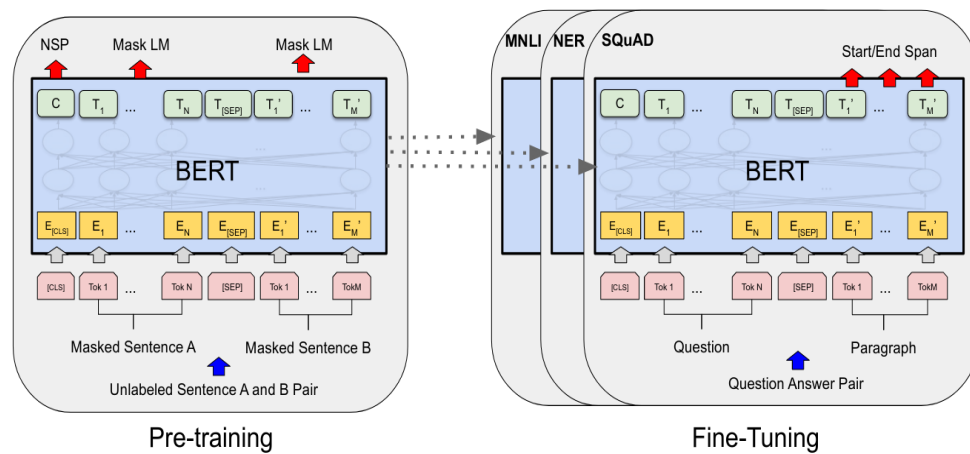


Figure 5. The general pre-training and fine-tuning procedures for BERT.

Despite being initialized with the same pre-trained parameters, each downstream task has fine-tuned parameters. For example, during the fine-tuning of classification, a weights layer W add where $W \in \mathbb{R}^{K \times H}$ K is the number of labels. Performance is essential in the BERT model to achieve higher accuracies, so they presented two sizes of the model: BERT_{BASE}, which is “(L = 12, H = 768, A = 12, Total Parameters = 110 M) and BERT_{LARGE} (L = 24, H = 1024, A = 16, Total Parameters = 340 M)”, where L is the number of layers, H is the hidden size, and A is the number of self-attention heads [63].

AraBERT Model

The latest effective model, successfully used for various BERT tasks, has been improved and extended by researchers with multiple parameters and methods [3]. AraBERT is a pre-trained language model for Arabic language processing tasks. Some potential merits of using AraBERT have been trained on a large dataset of Arabic text, which means it has a strong understanding of the language and can perform tasks such as language translation, text classification, and sentiment analysis with high accuracy. Because AraBERT has already been pre-trained on a large dataset, it can be fine-tuned for specific tasks with relatively little additional training data, reducing the time and resources required to develop and deploy natural language processing models. AraBERT is specifically designed for Arabic language processing, which makes it a valuable tool for tasks that involve Arabic text. AraBERT is a pre-trained model that can be easily integrated into existing natural language processing pipelines, making it easy to use and adapt for various tasks. The ability to fine-tune pre-trained models such as AraBERT allows for transfer learning, where knowledge learned from one task can be applied to another related task, further improving performance and reducing the training data [40,64,65]. Here, the same Saudi telecommunications datasets with the standard preprocessing method, as given in the next section, have been used by the AraBERT model. Also, the AraBERT model was configured with the same topology as the typical BERT-Model, which has 12 encoder blocks, 768 hidden dimensions, 12 attention heads, 512 maximum sequence lengths, and ~110 M parameters.

Figure 6 shows the general AraBERT model using the Twitter dataset for classification purposes to know the fundamental values of customer satisfaction. The standard AraBERT model will be used to simulate the mentioned telecommunication companies' Twitter datasets for accurate classification and CNN and RNN DL models, as given in Figure 7.

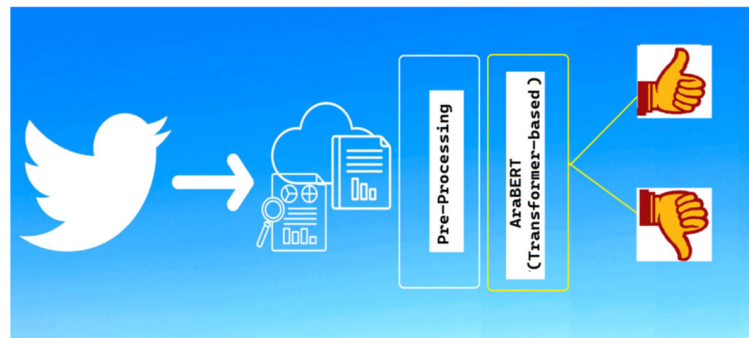


Figure 6. The proposed AraBERT overview.

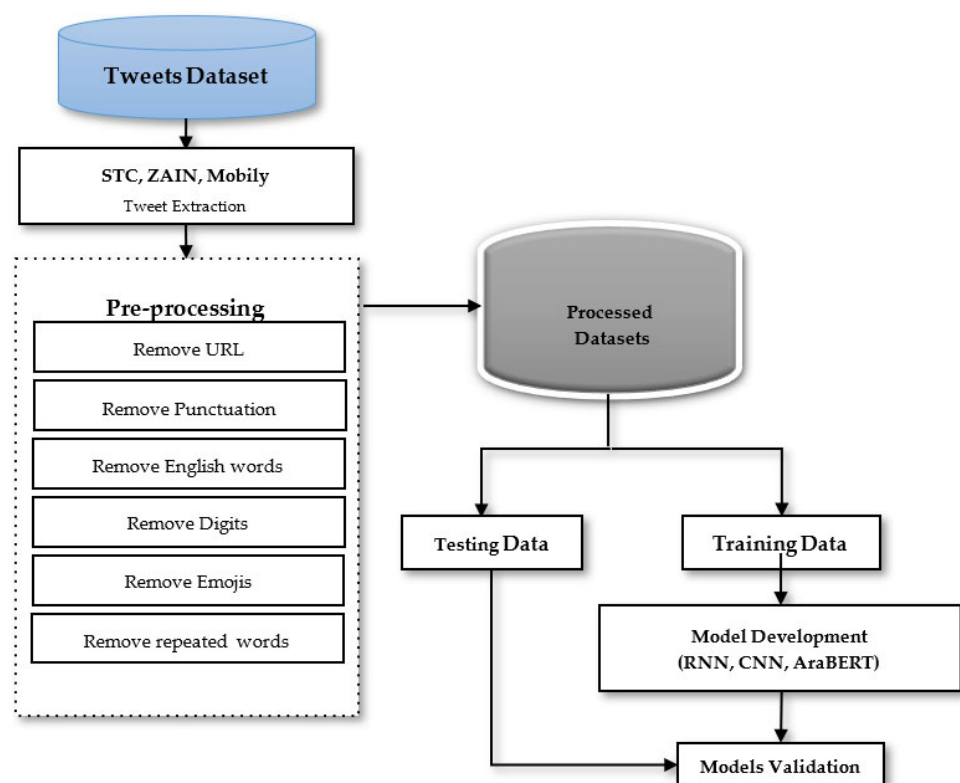


Figure 7. The proposed research methodology steps for classification tasks.

3. Research Methodology

The AraBERT model, successfully used in many natural and artificial-based applications with outstanding performance, especially in NLP, has never been used to predict or classify the telecommunication sector's customer satisfaction in the Saudi Arabian region. Therefore, the AraBERT model will predict Saudi telecommunication companies' (government and private) customer feedback quality, production, and planning. For the actual characteristics, feedback and company profile, the following models will be used for analysis with AraBERT. The RNN with Long Short-Term Memory (LSTMs) and Convolutional Neural Networks (CNN) algorithms are used as DL algorithms. The collected dataset will be preprocessed, trained, and tested with various topologies to predict customer satisfaction in these companies. The results of these DL models with different executing criteria will be measured through relevant performance metrics like accuracy, training, testing errors, etc. The preprocessed steps will be explained in Section 3.1.

In the steps trained on the LSTM model, by embedding each word in a tweet as a vector to generate word vectors embedding with 300 dimensions per word, we split the datasets

into 80% training and 20% testing data. From that, we fed this model embedding words using a 64-dimensional hidden state, applied a fractional of 0.1 dropout rate over the batch of sequences, then fed the model another LSTM layer with a 64-dimensional hidden state that returns one hidden state and a single unit dense layer was applied followed by Sigmoid activation. Also, the learning rate by Adam optimizer is ($lr = 10^{-3}$) with 10 epochs, and the Keras-TensorFlow library is implemented. On the CNN model, afterword embedding and splitting the datasets, the model is fed with several layers of convolutional layer (Conv1D with filters of 32 and kernel size of 8), and a pooling layer (MaxPooling1D with a pool size of 2) and a single unit dense layer was applied, followed by Sigmoid and ReLU activations. Also, this was with the same learning rate and epochs as the LSTM model.

In our approach, we used AraBERT based on the BERT model. There are six versions of the AraBERT model. We used AraBERTv02-large with a size of 1.38 G MB, 371 million parameters, no pre-segmentation needed, trained on 200 million sentences with a data size of 77 GB, and 8.6 billion words. Also, it contains 12 transformer blocks, 768 hidden sizes, and a self-attention of 12. In the preprocessing stage, we cleaned the data, as shown in Section 3.1. Then we pre-trained and fine-tuned each dataset separately and merged all the datasets into one dataset to come up with the result in Section 3.3. We applied randomly splitting of 80–20% for train-test datasets. We applied the experiment settings: maximum length = 128, patch size = 16, epochs = 2, adam epsilon = 10^{-8} , learning rate = 2×10^{-5} , and GELU activation function along with Transformer and Scikit-Learn libraries.

3.1. Twitter Dataset

The raw dataset for this study comprised customer reviews on Saudi communications companies named AraCust, which contain: STC, Zain, and Mobily, which consisted of 20,000 observations; STC had 7590, Mobily 6450, and Zain 5950 tweets in mixed Arabic and English languages with two output positive and negative ranking [66]. The data-processing is a crucial step applied to any collected raw data before embedding it with a sentiment extraction approach with higher quality data to obtain highly accurate simulation results. The stages include stemming, cleaning datasets by removing irrelevant data, blank space, etc., and tokenization. Furthermore, all this raw data was converted into numeric conversion so that the DL algorithms could understand and deal with it. Tokenization breaks the text stream into words, phrases, symbols, or other meaningful words, and a text-to-sequences tokenizer has been used. The following steps were used to clean the data for accurate simulation classification results. The preprocessing steps for all datasets have been mentioned in the following table.

Figure 7 shows the big picture of this research, where the Twitter datasets have been simulated after preprocessing (as mentioned in Table 1) by CNN, RNN, and AraBERT for the given task.

3.2. Performance Evaluation Metrics

We evaluate the performance of models through the classification report that presents the main evaluation metrics of a classification-based machine learning model. It shows the model's accuracy, recall, precision, F1 score, and support. Accuracy is the ratio of the total number of correctly predicted examples and the number of examples in the test set. The precision is the ratio of genuinely positive classifieds and the sum of true and false positive examples by the test set's model. The recall talks about the accurate prediction of positive examples. Moreover, the F1 score is the weighted harmonic mean of precision and recall of the model.

Table 1. Pre-processing steps of the datasets.

Preprocessing Steps		Example before Processing	Example after Processing	Translate
Step 1	Removing Emojis	@STCcare شكراً يا السمي ♥	@STCcare بالشمي شكراً	Thanks my friend
Step 2	Removing English Words	STCcare لك يعطيك العافيه وشكراً لك	يعطيك العافيه وشكراً لك	Bless you and thank you
Step 3	Removing English Symbols	@ الله يعينا ياخوك طحنا بعصابه	الله يعينا ياخوك طحنا بعصابه	May Allah help us brother We fell into this gang
Step 4	Removing Mobile No/ID/numbers	رقم الجوال ****05008 المشكلة ان كانت عندي مشكلة الهوية 4436*** سحبت من عندي تقريبا ريال 100 قبل اسبوع انها	رقم الجوال ورقم الهوية المشكلة ان كانت عندي مشكلة قبل اسبوع انها سحبت من عندي تقريبا ريال	Mobile number and identity number The problem is that I had a problem a week ago because it withdrew from me almost 100 riyals
Step 5	Removing stopping words	مشكور...	مشكور	Thanks
Step 6	Removing website link	متى ياموبايلي https://t.co/34jDalwW4o (accessed on 10 January 2017)	متى ياموبايلي	When Mobile
Step 7	Removing repeated Arabic words	الله، شركة، اشكركم		God, a company, thank you

For accurate classification analysis of the companies' datasets, the typical performance measure matrices were used for all three methods: AraBERT, CNN, and RNN. A classification report is generated after every classification process. Mathematical models of the classification accuracy, Precision, Recall, and F-measure are given in Equations (8)–(11).

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (8)$$

where TP—true positive; TN—true negative; FP—false positive; FN—false negative

$$\text{Precision} = \frac{TP}{TP + FP} \quad (9)$$

Recall—the proportion of actual positive cases that are correctly identified (Equation (10))

$$\text{Recall} = \frac{TP}{TP + FN} \quad (10)$$

where TP—true positive; FN—false negative

The F1-score is mathematically defined by Equation (11)

$$\text{F1 - score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (11)$$

3.3. Simulation Results on Twitter Sentiment Analysis

The three models, as mentioned above, have been simulated with different topologies and parameters for the classification purpose using three datasets: STC, Zain, and Mobily. COLAB has successfully simulated each method with the python platform. Each method's training and testing simulation results have been evaluated based on the training and testing error and accuracy.

Table 2 shows training error results measured during the training experiments on the RNN, CNN, and AraBERT models. The CNN model has the lowest training error for all the training data sets, i.e., Zain, Mobily, and STC. While AraBERT has the highest training errors for these data sets. The AraBERT is not converging in training experiments due to

its more comprehensive scope of classifying emojis and Arabic text; thus, it has significant training errors. This property makes it capable of generalizing for testing/validation data sets. In this way, it avoids overfitting.

Table 2. Training Error Results obtained by CNN, RNN and AraBERT.

Model\Dataset	Zain	Mobily	STC
RNN	0.0060	0.0229	0.0039
CNN	0.0058	0.0066	0.0032
AraBERT	0.02	0.04	0.03

Figure 8 shows the training accuracies obtained by the RNN, CNN, and AraBERT models. The training accuracies of RNN and CNN are higher than that of the AraBERT, leading the RNN and CNN to over-fit and thus may lead to significant errors for unknown datasets (e.g., validation datasets). The simulation results of the training experiments are further presented in the form of a confusion matrix for the three models mentioned earlier (i.e., CNN, RNN, and AraBERT).

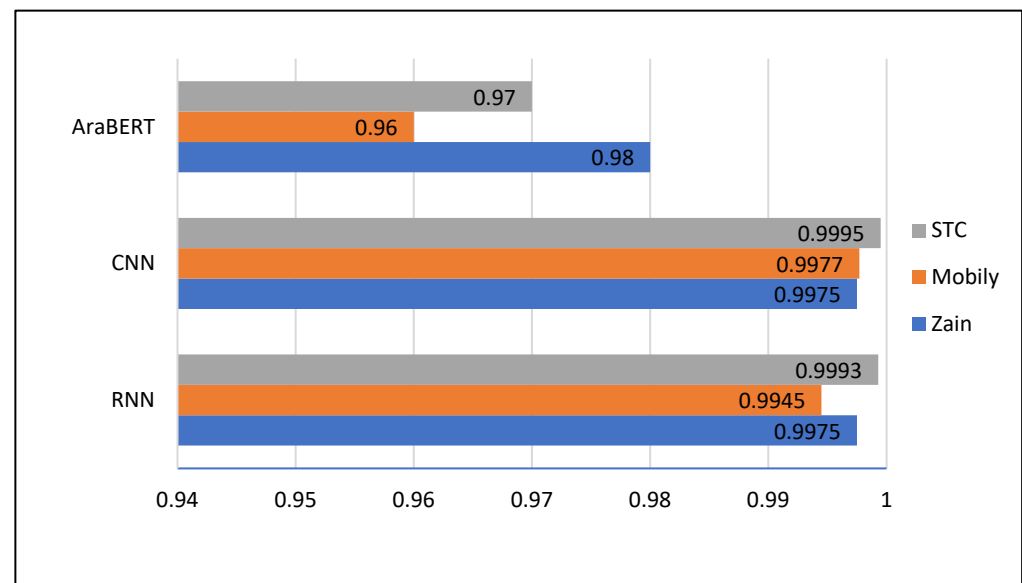


Figure 8. Training accuracies obtained by RNN, CNN, and AraBERT.

Table 3 tabulates the confusion matrix for the CNN on the STC dataset. It precisely classifies the negative examples while giving an 8% error in the case of positive examples.

Table 3. Confusion matrix obtained by CNN on STC dataset.

CNN_STC	Precision	Recall	F1-Score
Negative	1.00	0.99	1.00
Positive	0.92	1.00	0.96
Accuracy	-	-	0.99
Macro average	0.96	1.00	0.98
Weighted average	0.99	0.99	0.99

Using the CNN for the classification purpose, the results in the confusion matrix (which represents the summary of prediction results on a classification problem), obtained from the STC, ZAIN, and Mobily datasets have been presented in Tables 3–5, respectively. The average best result obtained by the CNN on the STC dataset where the F1 values are

very close to 1, significantly, the weight reached 0.99, which shows that the CNN model has achieved the state of art results from other datasets (Mobily and Zain).

Table 4. Confusion Matrix obtained by CNN on the Zain dataset.

CNN_ZAIN	Precision	Recall	F1-Score
Negative	1.00	0.96	0.98
Positive	0.00	0.00	0.00
Accuracy	-	-	0.96
Macro average	0.50	0.48	0.49
Weighted average	1.00	0.96	0.98

Table 5. Confusion Matrix obtained by CNN on the Mobily dataset.

CNN_Mobily	Precision	Recall	F1-Score
Negative	0.61	0.99	0.75
Positive	0.99	0.48	0.65
Accuracy	-	-	0.71
Macro average	0.80	0.74	0.70
Weighted average	0.82	0.71	0.70

Furthermore, the RNN model has been successfully simulated to obtain the results in the confusion matrix of the Twitter dataset of STC, ZAIN, and Mobily companies for classification tasks. The confusion matrix values after the simulation have been presented in Tables 6–8. The results shown in Tables 6 and 8 show that the RNN has achieved outstanding results on the STC and Zain dataset, where the weighted average reached 0.99.

Table 6. Confusion matrix obtained by RNN on the STC dataset.

RNN_STC	Precision	Recall	F1-Score
Negative	1.00	0.99	1.00
Positive	0.92	1.00	0.96
Accuracy	-	-	0.99
Macro average	0.96	1.00	0.98
Weighted average	0.99	0.99	0.99

Table 7. Confusion matrix obtained by RNN on the Mobily dataset.

RNN_Mobily	Precision	Recall	F1-Score
Negative	0.61	1.00	0.76
Positive	1.00	0.48	0.65
Accuracy	-	-	0.71
Macro average	0.80	0.74	0.70
Weighted average	0.82	0.71	0.70

Table 8. Confusion matrix obtained by RNN on the Zain dataset.

RNN_Zain	Precision	Recall	F1-Score
Negative	0.98	1.00	0.99
Positive	1.00	0.98	0.99
Accuracy	-	-	0.99
Macro average	0.99	0.99	0.99
Weighted average	0.99	0.99	0.99

The proposed AraBERT model has been simulated with the three types of Twitter datasets to evaluate the classification results in the confusion matrix. The results presented

in Figures 9–11 show that AraBERT also achieved the state of the results on all datasets, where the results are more accurate and stable than CNN and RNN models.

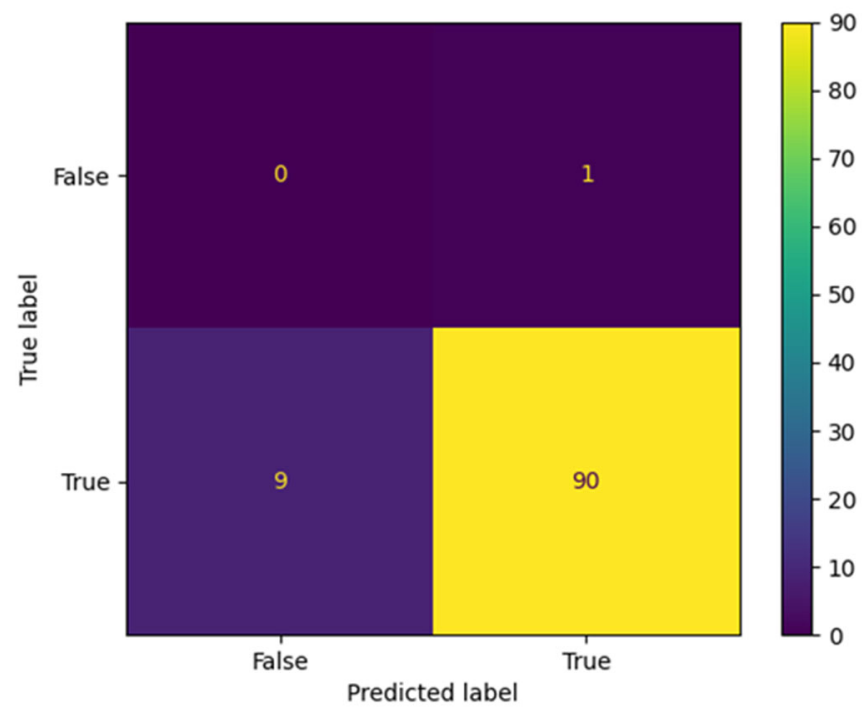


Figure 9. Confusion matrix obtained by AraBERT on the STC dataset.

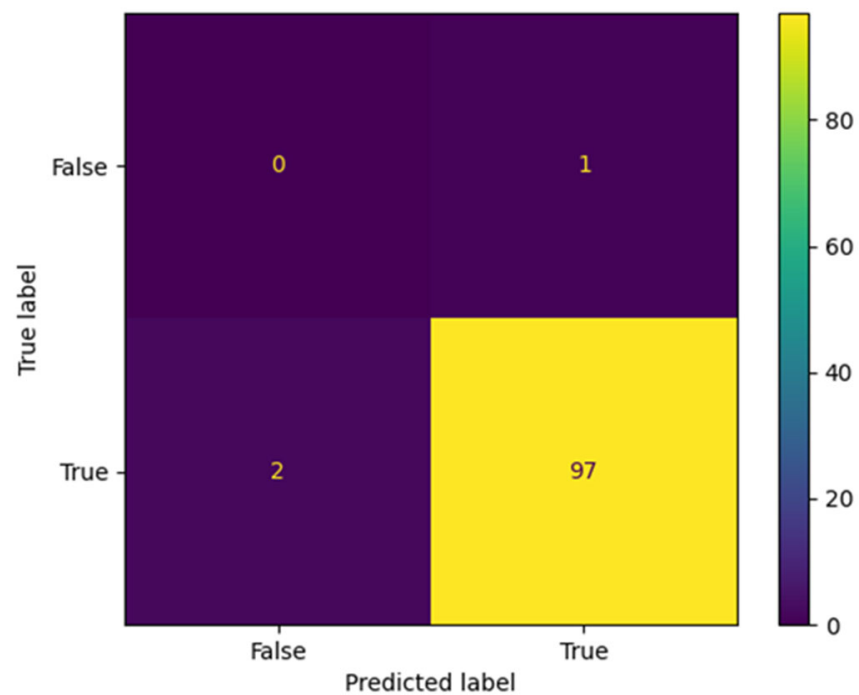


Figure 10. Confusion matrix obtained by AraBERT on the Zain dataset.

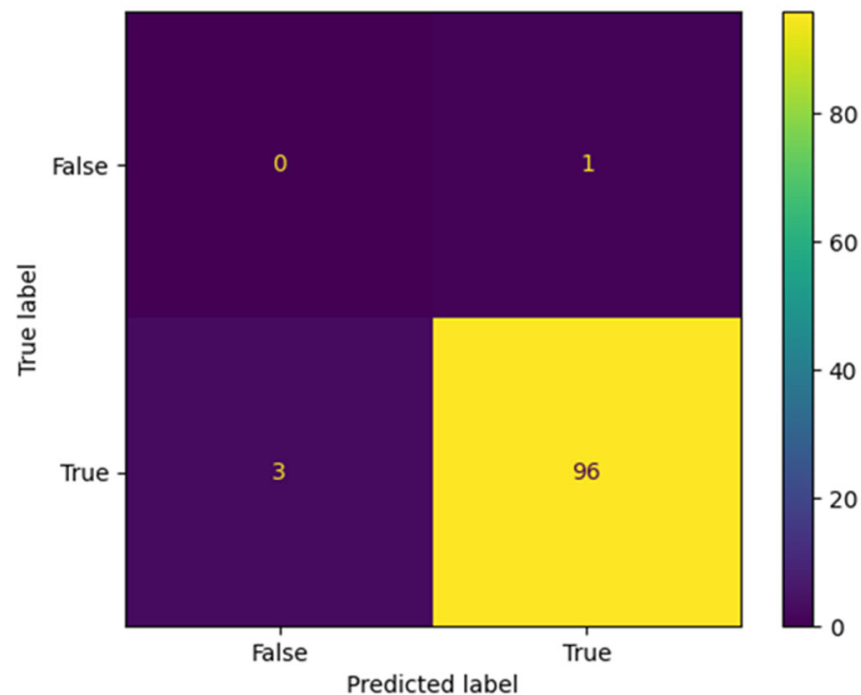


Figure 11. Confusion matrix obtained by AraBERT on the Mobily dataset.

CNN, RNN, and AraBERT obtained the above simulation results on various datasets with different simulation structures and parameters.

Table 9 shows validation loss results measured after each validation experiment on the validation datasets. RNN has the lowest validation loss for the STC dataset in the validation experiments. While for Zain and Mobily, the RNN validation loss is higher than that of the AraBERT model. On average, AraBERT has the lowest validation loss for all three datasets.

Table 9. Validation loss results obtained by CNN, RNN, and AraBERT.

Model\Dataset	Zain	Mobily	STC
RNN	0.0412	0.2141	0.004
CNN	0.0504	0.2895	0.0099
AraBERT	0.04	0.10	0.03

Table 10 shows validation accuracy results measured after each validation experiment on the datasets. The results tabulated in Table 10 clearly show that the AraBERT has the highest average validation accuracy compared to the CNN and RNN for all the datasets.

Table 10. Validation accuracy obtained by CNN, RNN, and AraBERT.

Model\Dataset	Zain	Mobily	STC	Average
RNN	0.9588	0.7859	0.9960	0.9135
CNN	0.9496	0.7105	0.9901	0.8834
AraBERT	0.96	0.90	0.97	0.9433

4. Conclusions

This research has used Arabic Sentiment Analysis to measure customer satisfaction in Saudi Arabia Telecom Companies based on tweets. This customer satisfaction has been tested and evaluated by three methods: CNN, RNN, and AraBERT, which is based on various performance metrics, including the confusion matrix. It was found that the

AraBERT model, which we simulated for the first time for the given dataset, obtained highly accurate and stable simulation results compared to other models that monitor a customer's satisfaction on social media. The highly accurate results can be successfully used to predict customer satisfaction, providing companies with the best knowledge and early warnings.

5. Future Work

The authors would like to extend their current research model to multiple applications such as NLP analysis, time series dataset prediction, and classifications of various datasets using AraBERT and other DL methods. The AraBERT will be improved based on hybridization with bio-inspired and typical methods and will be simulated in a numerical time-series dataset. Also, the typical model of AraBERT will be improved based on the psychological features of customer feedback, like motivations such as the desire for convenience, quality, or value [67]. Customer feedback can reveal a person's values, such as their priorities or beliefs about what is essential in a product or service [68], in addition to customers' attitudes toward a company, brand, or product, including their level of loyalty or favorability [69].

Author Contributions: Conceptualization and writing—original draft preparation, S.A.; writing—review and editing, H.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The datasets has been taken from this link <https://peerj.com/articles/cs-510/>.

Acknowledgments: I would like to express my sincere gratitude to my supervisor, Yu Zhuang, for his invaluable comments and suggestions on an early draft of this research paper. His expertise and guidance were crucial in helping me to refine my ideas and approach to the study. I am also grateful to my co-author, Habib Shah, for contributing to this project. Their insights and expertise were invaluable in the development of my research. I would like to thank both of them for their time and effort in helping to make this research paper a success. Also, I am grateful for the support and commitment from Texas Tech University to my research. The author extends their appreciation to the Deanship of Scientific Research at King Khalid University through the Large Group under Grant RGP 2/212/1443.

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

References

1. Alalwan, A.A.; Rana, N.P.; Dwivedi, Y.K.; Algharabat, R. Social media in marketing: A review and analysis of the existing literature. *Telemat. Inform.* **2017**, *34*, 1177–1190. [CrossRef]
2. Susanti, C.E. The Effect of Product Quality and Service Quality Towards Customer Satisfaction and Customer Loyalty in Traditional Restaurants in East Java. In Proceedings of the International Conference on Managing the Asian Century, Singapore, 11–13 July 2013; Springer: Singapore, 2013; pp. 383–393.
3. Abiodun, R. Development of mathematical models for predicting customers satisfaction in the banking system with a queuing model using regression method. *Am. J. Op. Manag. Inf. Syst.* **2017**, *2*, 86–91.
4. Mugion, R.G.; Musella, F. Customer satisfaction and statistical techniques for the implementation of benchmarking in the public sector. *Total. Qual. Manag. Bus. Excel.* **2013**, *24*, 619–640. [CrossRef]
5. Al-Ghamdi, S.M.; Sohail, M.S.; Al-Khaldi, A. Measuring consumer satisfaction with consumer protection agencies: Some insights from Saudi Arabia. *J. Consum. Mark.* **2007**, *24*, 71–79. [CrossRef]
6. The Communication and Information Technology Commission. Annual Report of (CITC). Available online: https://www.cst.gov.sa/en/mediacenter/reports/Documents/PR_REP_013Eng.pdf (accessed on 11 December 2022).
7. Hassounah, M.; Raheel, H.; Alhefzi, M. Digital response during the COVID-19 pandemic in Saudi Arabia. *J. Med. Internet Res.* **2020**, *22*, e19338. [CrossRef]

8. Digital 2019 Saudi Arabia. Available online: <https://www.slideshare.net/DataReportal/digital-2019-saudi-arabia-january-2019-v01> (accessed on 18 June 2022).
9. Bhatia, S.; Li, J.; Peng, W.; Sun, T. Monitoring and analyzing customer feedback through social media platforms for identifying and remedying customer problems. In Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Association for Computing Machinery, New York, NY, USA; 2013; pp. 1147–1154.
10. Duwairi, R.M.; Marji, R.; Sha'ban, N.; Rushaidat, S. Sentiment analysis in arabic tweets. In Proceedings of the 2014 5th international conference on information and communication systems (ICICS), Irbid, Jordan, 1–3 April 2014; pp. 1–6.
11. Alshammari, T.S.; Ismail, M.T.; AL-WADI, S.; Saleh, M.H.; Jaber, J.J.; Alshammari, J.J.J.; Tariq, S.; Ismail, M.T.; AL-WADI, S.; Saleh, M.H. Modeling and Forecasting Saudi Stock Market Volatility Using Wavelet Methods. *J. Asian Financ. Econ. Bus.* **2020**, *7*, 83–93. [\[CrossRef\]](#)
12. Kumar, S.; Zymbler, M. A machine learning approach to analyze customer satisfaction from airline tweets. *J. Big Data* **2019**, *6*, 62. [\[CrossRef\]](#)
13. Considine, E.; Cormican, K. Self-service Technology Adoption: An Analysis of Customer to Technology Interactions. *Procedia Comput. Sci.* **2016**, *100*, 103–109. [\[CrossRef\]](#)
14. Hossain, M.S.; Muhammad, G.; Amin, S.U. Improving consumer satisfaction in smart cities using edge computing and caching: A case study of date fruits classification. *Futur. Gener. Comput. Syst.* **2018**, *88*, 333–341. [\[CrossRef\]](#)
15. Copeland, B.J. "Artificial Intelligence". Encyclopedia Britannica, 11 November 2022. Available online: <https://www.britannica.com/technology/artificial-intelligence> (accessed on 11 December 2022).
16. Hosch, W.L. "Machine Learning". Encyclopedia Britannica, 13 December 2022. Available online: <https://www.britannica.com/technology/machine-learning> (accessed on 11 December 2022).
17. Webber, B.L. *Natural Language Processing: A Survey* BT-On Knowledge Base Management Systems: Integrating Artificial Intelligence and Database Technologies; Brodie, M.L., Mylopoulos, J., Eds.; Springer: New York, NY, USA, 1986; pp. 353–363. [\[CrossRef\]](#)
18. Waheeb, W.; Ghazali, R.; Herawan, T. Ridge Polynomial Neural Network with Error Feedback for Time Series Forecasting. *PLoS ONE* **2016**, *11*, e0167248. [\[CrossRef\]](#)
19. Zhang, A.; Li, B.; Wang, W.; Wan, S.; Chen, W. MII: A Novel Text Classification Model Combining Deep Active Learning with BERT. *Comput. Mater. Contin.* **2020**, *63*, 1499–1514. [\[CrossRef\]](#)
20. Jiang, K.; Lu, X. Natural Language Processing and Its Applications in Machine Translation: A Diachronic Review. In Proceedings of the 2020 IEEE 3rd International Conference of Safe Production and Informatization (IICSPI), Chongqing, China, 28–30 November 2020; pp. 210–214. [\[CrossRef\]](#)
21. KM, A.K.; Abawajy, J. Detection of False Income Level Claims Using Machine Learning. *Int. J. Mod. Educ. Comput. Sci.* **2022**, *14*, 65–77.
22. Almuqren, L.; Cristea, A.I. Twitter Analysis to Predict the Satisfaction of Telecom Company Customers. In Proceedings of the HT '16: 27th ACM Conference on Hypertext and Social Media, Halifax, NS, Canada, 10–13 July 2016.
23. Albawi, S.; Mohammed, T.A.; Al-Zawi, S. Understanding of a convolutional neural network. In Proceedings of the 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 21–23 August 2017; pp. 1–6. [\[CrossRef\]](#)
24. Jain, L.C.; Medsker, L.R. *Recurrent Neural Networks: Design and Applications*; CRC Press, Inc.: Boca Raton, FL, USA, 1999.
25. Yang, Z.; Yang, D.; Dyer, C.; He, X.; Smola, A.; Hovy, E. Hierarchical attention networks for document classification. In Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: Human language technologies; 2016; pp. 1480–1489.
26. Chang, C.C.; Lin, C.J. LIBSVM: A Library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2013**, *2*, 1–39. [\[CrossRef\]](#)
27. Thinh, N.K.; Nga, C.H.; Lee, Y.; Wu, M.; Chang, P.; Wang, J. Sentiment Analysis Using Residual Learning with Simplified CNN Extractor. In Proceedings of the 2019 IEEE International Symposium on Multimedia (ISM), San Diego, CA, USA, 09–11 December 2019; pp. 335–3353. [\[CrossRef\]](#)
28. Sahni, T.; Chandak, C.; Chedeti, N.R.; Singh, M. Efficient Twitter sentiment classification using subjective distant supervision. In Proceedings of the 2017 9th International Conference on Communication Systems and Networks (COMSNETS), Bengaluru, India, 04–08 January 2017; pp. 548–553. [\[CrossRef\]](#)
29. Dong, L.; Wei, F.; Tan, C.; Tang, D.; Zhou, M.; Xu, K. Adaptive Recursive Neural Network for target-dependent Twitter sentiment classification. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, Baltimore, MD, USA, 23–25 June 2014; Volume 2. [\[CrossRef\]](#)
30. Singh, J.; Tripathi, P. Sentiment analysis of Twitter data by making use of SVM, Random Forest and Decision Tree algorithm. In Proceedings of the 2021 10th IEEE International Conference on Communication Systems and Network Technologies (CSNT), Bhopal, India, 18–19 June 2021; pp. 193–198. [\[CrossRef\]](#)
31. Tunyan, E.V.; Cao, T.A.; Ock, C.Y. Improving Subjective Bias Detection Using Bidirectional Encoder Representations from Transformers and Bidirectional Long Short-Term Memory. *Int. J. Cognit. Lang. Sci.* **2021**, *15*, 329–333.
32. Yang, L.; Li, Y.; Wang, J.; Sherratt, R.S. Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning. *IEEE Access* **2020**, *8*, 23522–23530. [\[CrossRef\]](#)
33. Shehu, H.A.; Sharif, M.H.; Sharif, M.H.U.; Datta, R.; Tokat, S.; Uyaver, S.; Kusetogullari, H.; Ramadan, R.A. Deep Sentiment Analysis: A Case Study on Stemmed Turkish Twitter Data. *IEEE Access* **2021**, *9*, 56836–56854. [\[CrossRef\]](#)

34. Sindhu, C.; Som, B.; Singh, S.P. Aspect-Oriented Sentiment Classification using BiGRU-CNN model. In Proceedings of the 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 8–10 April 2021; pp. 984–989. [\[CrossRef\]](#)
35. Elfaik, H.; Nfaoui, E.H. Deep Attentional Bidirectional LSTM for Arabic Sentiment Analysis In Twitter. In Proceedings of the 2021 1st International Conference on Emerging Smart Technologies and Applications (eSmarTA), Sana'a, Yemen, 10–12 August 2021; pp. 1–8. [\[CrossRef\]](#)
36. Alzyout, M.; Bashabsheh, E.A.L.; Najadat, H.; Alaiad, A. Sentiment Analysis of Arabic Tweets about Violence Against Women using Machine Learning. In Proceedings of the 2021 12th International Conference on Information and Communication Systems (ICICS), Valencia, Spain, 24–26 May 2021; pp. 171–176. [\[CrossRef\]](#)
37. AL-Rubaiee, H.; Qiu, R.; Li, D. Analysis of the relationship between Saudi twitter posts and the Saudi stock market. In Proceedings of the 2015 IEEE Seventh International Conference on Intelligent Computing and Information Systems (ICICIS), Cairo, Egypt, 12–14 December 2015; pp. 660–665. [\[CrossRef\]](#)
38. Dong, S.; Wang, P.; Abbas, K. A survey on deep learning and its applications. *Comput. Sci. Rev.* **2021**, *40*, 100379. [\[CrossRef\]](#)
39. Medhat, W.; Hassan, A.; Korashy, H. Sentiment analysis algorithms and applications: A survey. *Ain Shams Eng. J.* **2014**, *5*, 1093–1113. [\[CrossRef\]](#)
40. Antoun, W.; Baly, F.; Hajj, H. Arabert: Transformer-based model for arabic language understanding. *arXiv* **2020**, arXiv:2003.00104.
41. Muhammad, A.N.; Aseere, A.M.; Chiroma, H.; Shah, H.; Gital, A.Y.; Hashem, I.A.T. Deep learning application in smart cities: Recent development, taxonomy, challenges and research prospects. *Neural Comput. Appl.* **2020**, *33*, 2973–3009. [\[CrossRef\]](#)
42. Teoh, K.H.; Ismail, R.C.; Naziri, S.Z.M.; Hussin, R.; Isa, M.N.M.; Basir, M. Face Recognition and Identification using Deep Learning Approach. *J. Phys. Conf. Ser.* **2021**, *1755*, 012006. [\[CrossRef\]](#)
43. Qiu, J.; Wu, Q.; Ding, G.; Xu, Y.; Feng, S. A survey of machine learning for big data processing. *EURASIP J. Adv. Signal Process.* **2016**, *2016*, 1–16. [\[CrossRef\]](#)
44. Trinh, C.; Meimaroglou, D.; Hoppe, S. Machine Learning in Chemical Product Engineering: The State of the Art and a Guide for Newcomers. *Processes* **2021**, *9*, 1456. [\[CrossRef\]](#)
45. Yang, J.; Li, S.; Wang, Z.; Dong, H.; Wang, J.; Tang, S. Using deep learning to detect defects in manufacturing: A comprehensive survey and current challenges. *Materials* **2020**, *13*, 5755. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Bello, A.A.; Chiroma, H.; Gital, A.Y.; Gabralla, L.A.; Abdulhamid, S.M.; Shuib, L. Machine learning algorithms for improving security on touch screen devices: A survey, challenges and new perspectives. *Neural Comput. Appl.* **2020**, *32*, 13651–13678. [\[CrossRef\]](#)
47. Gabhane, M.D.; Suriya, D.S.B.A. Churn Prediction in Telecommunication Business using CNN and ANN. *J. Posit. Sch. Psychol.* **2022**, *6*, 4672–4680.
48. DiPietro, R.; Hager, G.D. Deep learning: RNNs and LSTM. In *Handbook of Medical Image Computing and Computer Assisted Intervention*; Zhou, S.K., Rueckert, D., Fichtinger, C.A.I., Eds.; The Elsevier and MICCAI Society Book Series; Academic Press: Cambridge, MA, USA, 2020; pp. 503–519. [\[CrossRef\]](#)
49. El Kader, I.A.; Xu, G.; Shuai, Z.; Saminu, S.; Javaid, I.; Ahmad, I.S. Differential Deep Convolutional Neural Network Model for Brain Tumor Classification. *Brain Sci.* **2021**, *11*, 352. [\[CrossRef\]](#)
50. Hussain, I.; Ahmad, R.; Muhammad, S.; Ullah, K.; Shah, H.; Namoun, A. PHTI: Pashto Handwritten Text Imagebase for Deep Learning Applications. *IEEE Access* **2022**, *10*, 113149–113157. [\[CrossRef\]](#)
51. Shah, H. Using new artificial bee colony as probabilistic neural network for breast cancer data classification. *Front. Eng. Built Environ.* **2021**, *1*, 133–145. [\[CrossRef\]](#)
52. Lim, W.L.; Ho, C.C.; Ting, C.Y. Tweet sentiment analysis using deep learning with nearby locations as features. In *Lecture Notes in Electrical Engineering*; Springer: Singapore, 2020; Volume 603, pp. 291–299. [\[CrossRef\]](#)
53. Heikal, M.; Torki, M.; El-Makky, N. Sentiment Analysis of Arabic Tweets using Deep Learning. *Procedia Comput. Sci.* **2018**, *142*, 114–122. [\[CrossRef\]](#)
54. Dahou, A.; Elaziz, M.A.; Zhou, J.; Xiong, S. Arabic Sentiment Classification Using Convolutional Neural Network and Differential Evolution Algorithm. *Comput. Intell. Neurosci.* **2019**, *2019*, 2537689. [\[CrossRef\]](#)
55. Jianqiang, Z.; Xiaolin, G.; Xuejun, Z. Deep Convolution Neural Networks for Twitter Sentiment Analysis. *IEEE Access* **2018**, *6*, 23253–23260. [\[CrossRef\]](#)
56. Yu, Y.; Liang, S.; Samali, B.; Nguyen, T.N.; Zhai, C.; Li, J.; Xie, X. Torsional capacity evaluation of RC beams using an improved bird swarm algorithm optimised 2D convolutional neural network. *Eng. Struct.* **2022**, *273*, 115066. [\[CrossRef\]](#)
57. Yu, Y.; Samali, B.; Rashidi, M.; Mohammadi, M.; Nguyen, T.N.; Zhang, G. Vision-based concrete crack detection using a hybrid framework considering noise effect. *J. Build. Eng.* **2022**, *61*, 105246. [\[CrossRef\]](#)
58. Elman, J.L. Finding Structure in Time. *Cogn. Sci.* **1990**, *14*, 179–211. [\[CrossRef\]](#)
59. Naeem, M.; Mashwani, W.K.; ABIAD, M.; Shah, H.; Khan, Z.; Aamir, M. Soft computing techniques for forecasting of COVID-19 in Pakistan. *Alex. Eng. J.* **2022**, *63*, 45–56. [\[CrossRef\]](#)
60. Lara-Benítez, P.; Carranza-García, M.; Riquelme, J.C. An experimental review on deep learning architectures for time series forecasting. *Int. J. Neural Syst.* **2021**, *31*, 2130001. [\[CrossRef\]](#)
61. Ljunghed, J. Predicting Customer Churn Using Recurrent Neural Networks. Master's Thesis, School of Computer Science and Communication, KTH, Stockholm, Sweden, 2017.

62. Sohangir, S.; Wang, D.; Pomeranets, A.; Khoshgoftaar, T.M. Big Data: Deep Learning for financial sentiment analysis. *J. Big Data* **2018**, *5*, 3. [[CrossRef](#)]
63. Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv* **2018**, arXiv:1810.04805.
64. Nada, A.M.A.; Alajrami, E.; Al-Saqa, A.A.; Abu-Naser, S.S. Arabic text summarization using arabert model using extractive text summarization approach. *Int. J. Acad. Inf. Syst. Res.* **2020**, *4*, 6–9.
65. Faraj, D.; Abdullah, M. Sarcasmdet at sarcasm detection task 2021 in arabic using arabert pretrained model. In Proceedings of the Sixth Arabic Natural Language Processing Workshop, Kyiv, Ukraine (Virtual), 19 April 2021; pp. 345–350.
66. Almuqren, L.; Cristea, A. AraCust: A Saudi Telecom Tweets corpus for sentiment analysis. *PeerJ Comput. Sci.* **2021**, *7*, e510. [[CrossRef](#)] [[PubMed](#)]
67. Grönroos, C. From Marketing Mix to Relationship Marketing: Towards a Paradigm Shift in Marketing. *Manag. Decis.* **1994**, *32*, 4–20. [[CrossRef](#)]
68. Maio, G.R. Chapter One: Mental representations of social values. *Adv. Exp. Soc. Psychol.* **2010**, *42*, 1–43. [[CrossRef](#)]
69. Oliver, R.L. *Satisfaction: A Behavioral Perspective on the Consumer*, 2nd ed.; Routledge: New York, NY, USA, 2010.

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.