International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Architectural Design of aChatbotused For Artificial Intelligence with NLP Classification using Deep Learning

Ria Kohli[1], Deepak K Sinha[2], Dr. G. Nanda Kishor Kumar[3], Badria Sulaiman Alfurhood[4], Rajesh Gupta[5], Sudhir Kumar Chaturvedi[6]

Submitted: 20/09/2022 **Accepted:** 24/12/2022

Abstract:A chatbot is a piece of technology that uses natural language to mimic human behaviour. In order to improve customer service and happiness, there are many types of chatbots that can be employed as conversational agents in different business fields. This research proposed novel technique in chatbot data based NLP classification utilizing DL structures. Here the information has been gathered from chatbot and handled for clamor expulsion and standardization. The handled information has been highlight extricated with ordered utilizing Bi-LSTM based Intermittent brain organizations. The trial examination has been completed regarding exactness, accuracy, F-1 score.The trials on two data sets of articles revealed that employing natural language processing and the suggested technique established its viability for creating an automatic categorization system of articles with an accuracy of above 91%.

Keywords: *Chatbot, NLP, Classification, Deep Learning, Normalization, Feature Extraction*

1. Introduction

The expression "chatbot" alludes to programming specialists that repeat human correspondence through message or voice messages. They are likewise alluded to as "chat robots."One of the first and essential targets of chatbots was to seem shrewd and mislead others about their actual nature.Their use has extraordinarily expanded as extra chatbots with various plans and capacities have been created [1].These conversational specialists are exceptionally restricted in their capacity to extend their insight bases continuously, however they can fool clients into thinking they are addressing humans.Thechatbot utilizes man-made consciousness and profound learning methods to understand client info and proposition an important response. Furthermore, they speak with individuals involving normal language in an assortment of chatbot applications, including call focuses and clinical chatbots [2].Doctors, medical caretakers, patients, or their families could profit from a chatbot. Chatbots might have the option to step in and let some free from the strain on clinical staff by better getting sorted out quiet information, overseeing drugs, aiding emergencies or with medical aid, and giving an answer for minor clinical challenges.With the merging of NLP and deep learning, such machine-human interaction has substantially improved in recent years. Severalmethods for executing chatbots utilizing profound learning calculations are accessible. Seq2Seq is a famousdeeplearning procedure that is easy to use into a chatbot [3].

The contribution of this research is as follows:

1. To propose novel technique in chatbot data based NLP classification using deep learning architectures
2. The processed data has been feature extracted with classified using Bi-LSTM based Recurrent neural networks.

2. Related Works

A few text-based human-computer interaction systems, such as ELIZA [5], which mimics a psychiatrist, and PARRY [4], which proposes the thoughts of a paranoid patient, have been developed.In two different trials, [5] compared interactions between a virtual person and a real human in the context of a medical consultation. Their findings indicate a correlation between virtual and actual encounters.The work of A. [6]A medical recommendation system particularly created to engage with users and take on the role of a doctor is presented by [7].Pharmabot is a paediatric generic medicine consultant chatbot that [8]

[1] *Computer Science, Research Scholar, Bhagwantuniversity Ajmer, Rajasthan, India. Email: kohliria7337@gmail.com*
[2] *Professor, Department of CSE, Jain(Deemed-to-be University), Bangalore, India. Email: sk.deepak@jainuniversity.ac.in https://orcid.org/0000-0003-0317-8624*
[3] *Professor, Computer Science and Engineering, Malla Reddy University, Telangana, India. Email: dr.gnandakishor@gmail.com 0000-0001-6395-3587*
[4] *Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Saudi Arabia. Email: bsalfurhood@pnu.edu.sa*
[5] *Pro Chancellor, Department of Management, Sanskriti University, Mathura, Uttar Pradesh, India. Email: prochancellor@sanskriti.edu.in 0000-0002-2118-8474*
[6] *Associate Professor, School of Engineering, UPES Dehradun, India Email: sudhir.chaturvedi@ddn.upes.ac.in*

offer as a tool for prescribing and providing helpful knowledge about generic medications for children.Because they collect contextual information, the model, known as ELMo [9], performs significantly better than word2vec.2018 will see the launch of OpenAI's embedding model development leveraging Google's newNN architecture, Transformer.Transformer is entirely attention-based, which significantly boosts the effectiveness of extensive model training on TPU.Their original design, known as GPT, is now frequently utilised for text generating tasks. In the same year, Google creates BERT, a bidirectional transformer-based system. The most sophisticated embedding model currently available is called BERT, which has 340M parameters and was trained on 3.3 billion words.More training data and larger models are still in style [10].

3. System Model

This segment examine novel method in chatbot information based NLP grouping utilizing profound learning models. Here the information has been gathered from chatbot and handled for commotion evacuation and standardization. The handled information has been include removed with characterized utilizing Bi-LSTM based Intermittent brain organizations.

ChatBot based AI:

Figure 1 depicts the strata of the campus's architecture. The first layer is in charge of gathering data that prioritises the use of IoTdevices.The Internet of Things (IoT) devices adapt to the needs of each group and gather crucial data regarding all of their activities.In the specific example of the students, various IoT devices and sensor systems gather data from them, including the precise amount of time each student spends at the university, the locations he frequents, as well as details about the tasks he completes and his credentials.This knowledge is crucial when the AI needs to make a judgement call regarding an event it is researching.

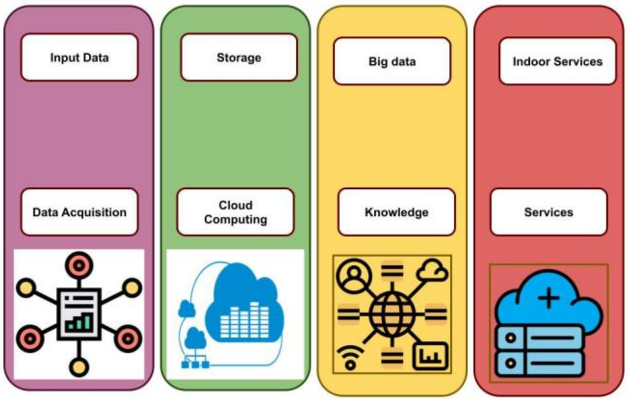

Fig. 1. Layers of the architecture of a Chatbot that uses artificial intelligence

Bi-LSTM based Recurrent neural networks:

Financial time series that continuously fluctuate are categorised into M movements: After document d is released, there is a price movement called fd in d-FinLDA. The number of categories, M, is dependent on thresholds; for example, if the threshold is a 0.5% change in both rise and reduction, then there are only two movements. For all Dirichlet distributions in the model that are conjugate to multinomial distributions, we also utilise symmetric Dirichlet priors. As a result, the hyper parameters,, and each have η, γ and α, single value. Following its probabilistic graphical model, the assumption of d-FinLDA is described by the following generative process.

1 For each topic $k \in \{1, \dots, K\}$:
   a) Produce$\beta_k \mid \eta = (\beta_{k,1}, \dots, \beta_{k,V}) \sim \text{Dir}\ (\eta)$
   b) Produce$\delta_k \mid \gamma = (\delta_{k,1}, \dots, \delta_{k,M}) \sim \text{Dir}\ (\gamma)$

2 For every document $\in \{1, \dots, D\}$:
   a) Produce$\theta_d \mid \alpha = (\theta_{d,1}, \dots, \theta_{d,K}) \sim \text{Dir}\ (\alpha)$
   b) For every word token $\in \{1, \dots, N_d\}$:
      i) Produce$z_{d,n} \mid \theta_d \in \{1, \dots, K\} \sim \text{Mult}\ (\boldsymbol{\theta}_d)$
      ii) Produce$w_{d,n} \mid \boldsymbol{\beta}, z_{d,n} \in \{1, \dots, V\} \sim \text{Mult}\ (\boldsymbol{\beta}_{z_{d,n}})$
   c) Produce$f_d \mid \delta, z_d \in \{1, \dots, M\} \sim \text{Mult}\ (\mu_d)$

For d-FinLDA, the joint distribution of the hidden and observed variables is:

$$P(\beta, \delta, \theta, z, w, f \mid \alpha, \eta, \gamma)$$
$$= \prod_{k=1}^{K} P(\boldsymbol{\beta}_k \mid \eta) \prod_{k=1}^{K} P(\boldsymbol{\delta}_k \mid \gamma) \prod_{d=1}^{D} (P(\boldsymbol{\theta}_d \mid \alpha)$$
$$h_t = \tan h\ (W_{xh}x_t + W_{hh}h_{t-1} + b_h)$$

$y_t = W_{hy}h_t + b_y$

where W is a weight lattice, b is an inclination vector, tanh is a secret layer initiation capability, and ht is the secret state yield. Wxh is a weight interfacing the info (x) to the secret layer (h).The result of the secret state, which gets the results of the past stage, is determined utilizing condition (1).In request to utilize LSTM, the accompanying conditions should be utilized.$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i)$

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

$$c_t = f_t c_{t-1} + i_t \tanh\ (W_{xc}x_t + W_{hc}h_{t-1} + b_c)$$
$$h_t^b = \tanh\ (W_{xh}^b x_t + W_{hh}^b h_{t+1}^b + b_h^b)$$
$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o)$$

$h_t = o_t \tanh\ (c_t)$

In any case, LSTM actually has a downside in that it can consider the information's past setting and is consequently unfit to consider any ensuing (i.e., future) context.The input is considered by the forward secret layer h f t in climbing request, i.e., t = 1, 2, 3,..., T.The in reverse secret layer, then again, considers the contribution to diminishing request, i.e., t = T,...,3,2,1. At long last, yield yt is created by joining h f t and h b t. These conditions are utilized to execute the BiLSTM model:

$$h_t^f = \tanh\left(W_{xh}^f x_t + W_{hh}^f h_{t-1}^f + b_h^f\right)$$
$$h_t^b = \tanh\left(W_{xh}^b x_t + W_{hh}^b h_{t-1}^b + b_h^b\right)$$

$$y_t = W_{hy}^f h_t^f + W_{hy}^b h_t^b + b_y$$

Our model's training goal is to minimise the cross-entropy loss given a set of annotated training instances and an L2 regularisation term, which is provided by

$$L(\theta) = -\sum_i \log p_{g_i} + \frac{\beta}{2} \parallel \theta \parallel_{2,}^2$$

where θ stands for each model parameter. PGI denotes the model's prediction of the gold relation type for the i-th training case. The regularisation parameter is calledβ.

4. Experimental Analysis

To create word vector representations, we employed the GloVe word embedding layer, pre-trained word vectors, and an unsupervised learning approach.Data from samples examined within the first 90 days of the Chatbot's adoption indicate a significant improvement in student performance. data on activity compliance in regard to tasks sent, finished, and delivered, as well as the level of development during the relevant time period.The time frame taken into account is 90 days after implementation; however, the percentage of production before the Chatbot implementation is presented as a reference in the second column.

Table-1 Comparative analysis between proposed and existing technique

Parameters	PARRY	BERT	ADC_AI_NLP_DL
Accuracy	83	89	93
Precision	85	88	91
Recall	75	79	85
F1_Score	65	72	80

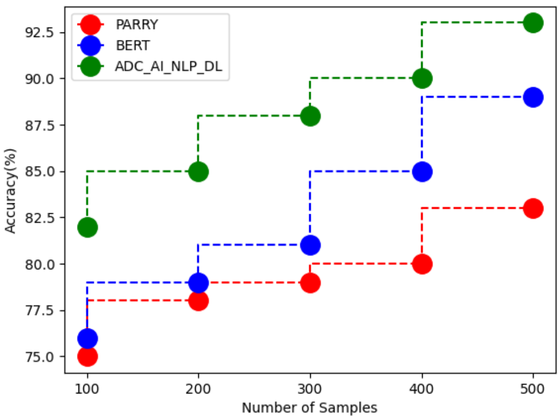

Fig.-2 Comparison of accuracy

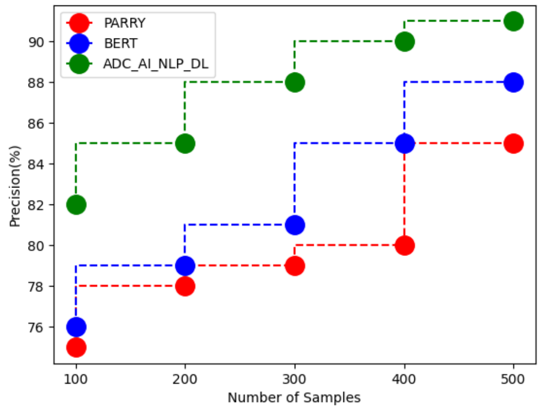

Fig.-3 Comparison of precision

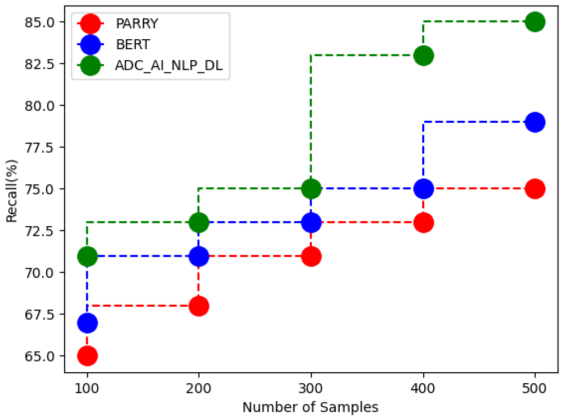

Fig.-4 Comparison of recall

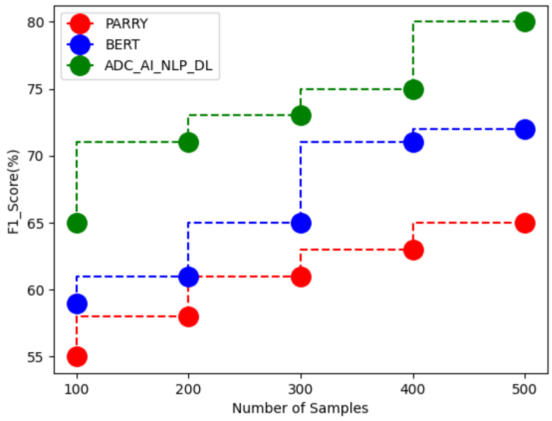

Fig.-5 Comparison of F-1 score

From above figure 2-5 the comparative analysis between proposed and existing technique has been shown as far as exactness, accuracy, review and F_1 score. Here the proposed method attainedaccuracy of 93%, accuracy of 91%, review of 85% and F_1 score of 80%. One specification for assessing classification methods isaccuracy. Official description of accuracy is as follows: Total number of accurate guesses isequal to total number of accurate guesses. By dividing number of accurate predictions byoverall sample size, we may determine accuracy. Our model was 44 percent accurate on thismulticlass problem, according to the outcome. One indicator of the model's performance isprecision, or the nature of a fruitful expectation. Complete number of exact positivepredictions is separated by absolute number of genuine encouraging points to decide accuracy. Review literally refers to how many right hits were also discovered, or how many genuine positives wereremembered. Precision is the percentage of returning hits that were true positive, or correct hits. The recall is determined by comparing the proportion of correctly labelled Positivesamples to all Positive samples. Model ability to make distinctions Recall is used tomeasure positive samples. As more positive samples are identified, the recall rises. The F1score is computed using harmonic mean of recall as well as precision. Recall that harmonicmean serves as a replacement for the arithmetic mean, which is utilised more frequently. Itfrequently helps when figuring up an average rate. We figure out the F1 score meanprecision and recall.

5. Conclusion

This research propose novel technique in chatbot data based NLP classification using deep learning architectures. The processed chatbox data has been feature extracted with classified using Bi-LSTM based Recurrent neural networks. Thepre-processing of train information utilizes NLP, a strategy used to comprehend PC informationand oversee human cooperations. Also, the text remarks that are given to the model arebeing pre-handled. The two snippets of data are communicated to the feeling library, where thepre-handled information highlights are removed. The exploratory examination has been completed as far as exactness, accuracy, review, F-1 score.the proposed procedure attainedaccuracy of 93%, accuracy of 91%, review of 85% and F_1 score of 80%.

References

Rahman, M. M., Watanobe, Y., & Nakamura, K. (2021). A bidirectional LSTM language model for code evaluation and repair. *Symmetry*, *13*(2), 247.

Kanungsukkasem, N., &Leelanupab, T. (2019). Financial latent Dirichlet allocation (FinLDA): Feature extraction in text and data mining for financial time series prediction. *IEEE Access*, *7*, 71645-71664.

Suman, S., & Kumar, J. (2022). Interactive Agricultural Chatbot Based on Deep Learning. In *Intelligent Data Communication Technologies and Internet of Things* (pp. 965-973). Springer, Singapore.

Rahat, F. R., Mim, M., Mahmud, A., Islam, I., Mahbub, M. J. A., Mehedi, M. H. K., &Rasel, A. A. (2022). Data Analysis Using NLP to Sense Human Emotions ThroughChatbot. In *International Conference on Advances in Computing and Data Sciences* (pp. 64-75). Springer, Cham.

Soufyane, A., Abdelhakim, B. A., & Ahmed, M. B. (2021). An intelligent chatbot using NLP and TF-IDF algorithm for text understanding applied to the medical field. In *Emerging Trends in ICT for Sustainable Development* (pp. 3-10). Springer, Cham.

Ni, P., Okhrati, R., Guan, S., & Chang, V. (2022). Knowledge Graph and Deep Learning-based Text-to-GQL Model for Intelligent Medical Consultation Chatbot. *Information Systems Frontiers*, 1-19.

Mittal, M., Battineni, G., Singh, D., Nagarwal, T., & Yadav, P. (2021). Web-based chatbot for frequently asked queries (faq) in hospitals. *Journal of Taibah University Medical Sciences*, *16*(5), 740-746.

Kurup, G., & Shetty, S. D. (2022). AI Conversational Chatbot for Primary Healthcare Diagnosis Using Natural Language Processing and Deep Learning. In *Computational Intelligence in Pattern Recognition* (pp. 259-272). Springer, Singapore.

Anki, P., Bustamam, A., &Buyung, R. A. (2021). Comparative Analysis of Performance between Multimodal Implementation of Chatbot Based on News Classification Data Using Categories. *Electronics*, *10*(21), 2696.

Bhagchandani, A., &Nayak, A. (2022). Deep Learning Based Chatbot Framework for Mental Health Therapy. In *Advances in Data and Information Sciences* (pp. 271-281). Springer, Singapore.