# Automated Shop Assistant using NLP and Mask Detection

**Sudiksha Acharya[1], Snehasish Dawn[1]**

*[1]SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India*

---

**Abstract -** *In recent times, due to the widespread of COVID, where keeping one's distance from others in public is mandatory, it would be more efficient if we could limit human interaction in public places like stores. A well-trained machine which is capable of doing a wide range of tasks can reduce human engagement. A chatbot is an application that replicates human interaction through NLP, they are constructed in such a way that users cannot identify whether they are receiving a response from a machine or a real human. This research paper introduces an Artificial Intelligence-based Shop Assistant which can be implemented in the shops as an interactive kiosk, which detect whether the person is wearing a mask as an authentication and then gives access to a chatbot for product search. It takes input from a customer in form of speech and text, supporting multiple languages and provides with enquired information about the product - stock, prices, discounts, etc. Kiosk can redirect the customer to the designated selves where the enquired products are placed. This system will even alert the shop manager when product availability is reducing. At the same time, it increases customer engagement as the customer can access information about a wider range of products than physically searching for it which reduces human dependency. This system can also be used for direct and indirect marketing to promote products. Data analysis of the searched data provides insights into the best and worst-performing products, which aids in sales forecasting and inventory management.*

***Key Words***: CNN, Mask Detection, NLP, Exploratory Data Analysis, Kiosk, Business Automation

## 1.INTRODUCTION

COVID pandemic started in late 2019 and within a year it has affected more than 180 million and led to the death of over 3.8 million people worldwide. The disease's spread is still ongoing in numerous parts of the world as of June 2021. Controlling future epidemics will include social distancing initiatives [1]. Maintaining social distance is not only a choice but also a necessity as a result of an unanticipated shift caused by COVID [2]. Because the human footprint is higher in public settings, the viruses are more likely to propagate there. To reduce human intervention, an Artificial Intelligence-based Kiosk system can be implemented that does not require any form of human-to-human interaction to break the chain. In this research article, we presented a web-based kiosk system with relevant Graphical User Interface (GUI) which can be deployed in shops and utilized as an automated shop assistant, with numerous functions that will lessen dependency on human assistance. The suggested system's features are explained further below.

Mask detection functionality works as an authenticating system for this kiosk, which will let the customer enquire about the product only if the customer is wearing a mask else the system would not let the user use the kiosk. Computer vision is used to train the mask detection model, which is used to discern between items, categorize them, and order them based on the trained data points. Previously, Computer Vision was solely intended to replicate human visual systems until we discovered how an Artificial Intelligent model can be used to augment the functionality of a system [3]. Classifiers from Convolutional neural networks (CNN) are also employed in the proposed model, resulting in a model with considerably greater accuracy with the same quantity of test and train data [3]. Regardless of the layout of the visual recorded by the computer vision, the model can track the face.

The kiosk is trained with a natural language processing model which can take input from customers in the form of text and speech in various supported languages. A database is created by the system admin which contains all the products with their corresponding amount and aisle number (location) where the products are stocked up. The kiosk will be able to analyze and process the user's request and fetch the required data from the table. The entity is extracted from the user's request which is then matched with the database and then the data is retrieved and produced to the customer.

The queries that are entered into the system are saved in our database, which can then be retrieved and utilized to analyze product sales and trends. The dataset is subjected to exploratory data analysis in order to better understand the pattern of the data and discover the data points. In this system, we have analyzed the data we collected to determine the product's trend over a specific period of time, as well as the word cloud, which makes it easier to analyze on-demand products and determine the time period when users are most active, which can be used for product marketing and endorsement. To increase the reach time of the ads, the ads should be placed at the time when the active time of the system is high [4].

## 2. OVERVIEW OF THE PROPOSED SYSTEM

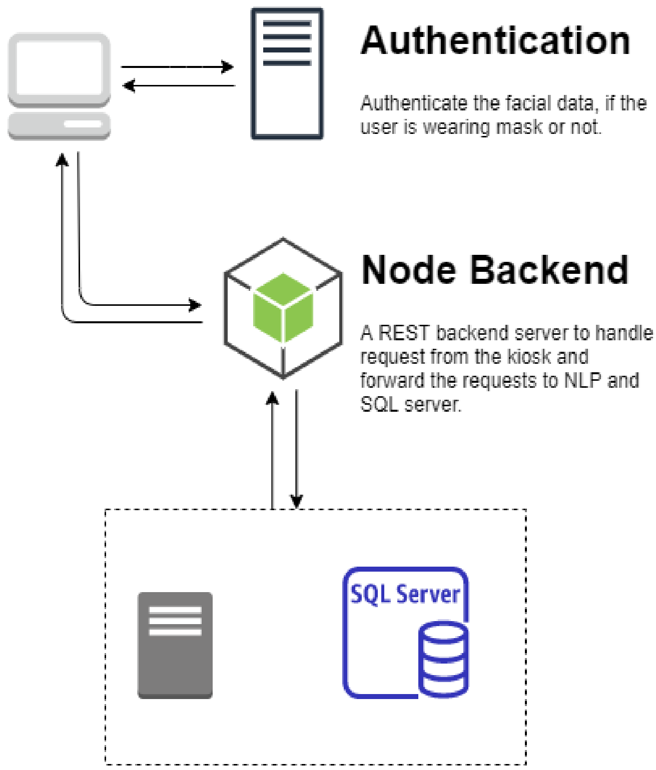



**Fig - 1**: Overview of Kiosk Model

The user will approach the kiosk for help, which will be augmented by our Artificial Intelligence Chatbot. When a user approaches the Kiosk, a front-facing camera is used to determine if the customer is wearing the mask by computer vision. The user will be denied service if they are not wearing a mask. This is accomplished through the use of a CNN-based face detection model. When it's confirmed that the user was wearing a mask, they'll be directed to a chatbot screen where they can ask questions in the form of voice. The frontend system, which also serves as a kiosk, will transform voice queries to text using speech-to-text techniques, and then send the query to the backend server, where natural language processing technologies will be used to determine the intent and entities. For instance, the user asks "Where can I find some Breakfast?", here the Intent will be to "BUY" and the entity will be "BREAKFAST". This intent and entities will invoke a REST API call to a NodeJS server which acts as a layer between the database and Kiosk [5]. The database has all the information regarding a particular product, when the NodeJS server queries about a product it receives all the information regarding that particular product and generates a response to be sent to the kiosk which the user can interpret.

## 3. METHODOLOGY

### 3.1 Mask Detection for user authentication

The proposed Artificial Intelligent based system includes face mask screening to prevent customers from entering the shop without a mask. This works as the authentication system for our kiosk model, a customer can not access the chatbot until computer vision system detects the customer is wearing a mask

#### 3.1.1 Data collection and processing

The sources [6][7] provide image datasets divided as Mask and Without Mask, which are then used to train and test the algorithm. The mask dataset has 2602 pictures, whereas the non-masked dataset has 2604 images. Image augmentation [8] is a technique for modifying pictures in terms of magnification, rotation range, and other parameters, resulting in a greater number of data points and hence improved algorithm accuracy. The images are transformed to a certain size(50x50) and color scale(greyscale) as the extra details in the image captured by computer vision are unnecessary and will increase the processing time of the model. This transformation will ensure consistency across all pictures in the collection, which speeds up the on-camera live recognition.

#### 3.1.2 CNN used for Face Detection

A Convolutional Neural Network is a Deep Learning method that takes an input image and assigns priority to various aspects in the image based on learnable weights, which are often provided by the developer in the form of training and testing data allowing it to distinguish one from the other [9]. In the proposed model, there are two convolutional layers and the pre-processed greyscale images of 50x50 are passed into the first convolutional layer which consists of 200 kernels of 3x3 image matrix which is fed into the second convolutional layer which consists of 100 kernels of 3x3 image matrix, both the layers are followed by Relu (Rectifier unit) and Pooling Layers. Max pooling layer of size 2x2 image matrix is used in this model which is used to extract the maximum dominant features that are present in the kernels, this layer also works as a noise suppressant which is important for reducing the extra and irrelevant features present in the kernels [10]. After the image is passed from both the convolutional neural network layers, it is further passed to the classifier stage. In the classifier layer, the convolutions are flattened which is connected to a dense layer of 50 neurons. A dropout layer is added in the intermediate stage to reduce overfitting in our model. The convolutions are further classified into two neurons for 'Mask' and 'No Mask' which are also known as label classes. The SoftMax technique is used for the classification. Categorical cross-entropy is used to find out the loss of the model while training and testing the datasets as we have two categorical outputs and lastly Adam Optimizer is used for gradient descent, this is a high accuracy optimization method with the usage of low memory.

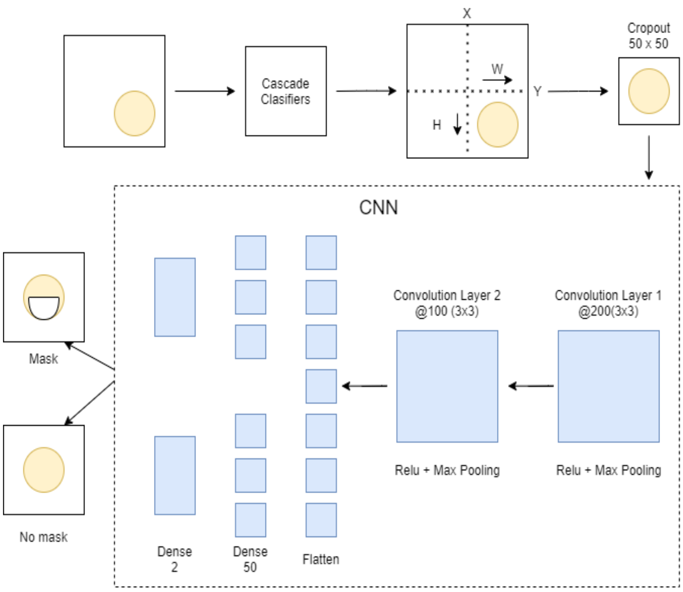



**Fig – 2**: Overview Mask Detection

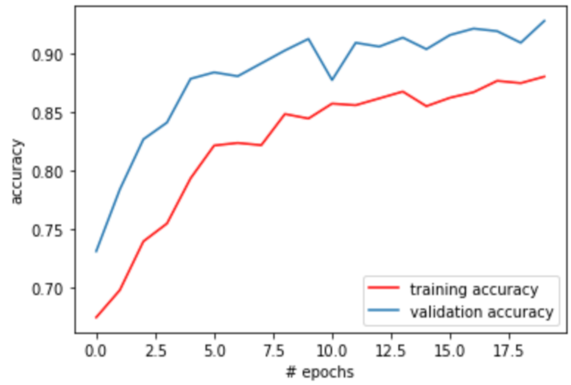
### 3.1.3 Training and testing of the model

The model was trained using training data for the first 80% of the data points and test data for the remaining 20%. Both with mask and without mask data are included in the data points. The constructed algorithm is put to the test throughout 20 epochs (iterations) and the results are present in Table 1.

**Table - 1:** Loss and accuracy percentage produced

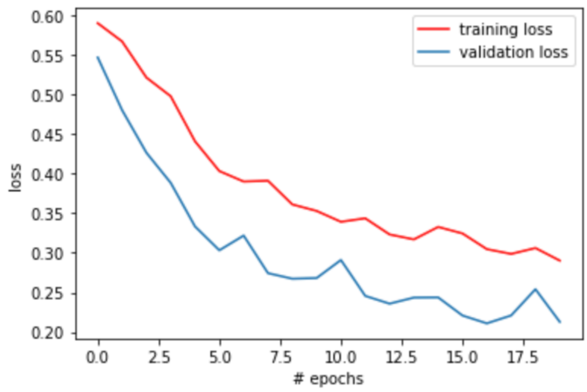
| Epoch | Loss | Accuracy | Val_Loss | Val_Accuracy |
|---|---|---|---|---|
| 1 | 0.5900 | 0.6751 | 0.5467 | 0.7316 |
| 2 | 0.5670 | 0.6983 | 0.4805 | 0.7844 |
| 3 | 0.5214 | 0.7400 | 0.4564 | 0.8273 |
| 4 | 0.4978 | 0.7553 | 0.3884 | 0.8416 |
| 5 | 0.4408 | 0.7939 | 0.3333 | 0.8790 |
| 6 | 0.4032 | 0.8219 | 0.3032 | 0.8845 |
| 7 | 0.3900 | 0.8242 | 0.3218 | 0.8812 |
| 8 | 0.3910 | 0.8222 | 0.2742 | 0.8922 |
| 9 | 0.3611 | 0.8489 | 0.2674 | 0.9032 |
| 10 | 0.3529 | 0.8480 | 0.2681 | 0.9131 |
| 11 | 0.3392 | 0.8478 | 0.2752 | 0.8779 |
| 12 | 0.3436 | 0.8564 | 0.2456 | 0.9098 |
| 13 | 0.3232 | 0.8622 | 0.2358 | 0.9065 |
| 14 | 0.3169 | 0.8681 | 0.2432 | 0.9142 |
| 15 | 0.3327 | 0.8556 | 0.2437 | 0.9043 |
| 16 | 0.3245 | 0.8628 | 0.2208 | 0.9219 |
| 17 | 0.3046 | 0.8675 | 0.2109 | 0.9197 |
| 18 | 0.2986 | 0.8772 | 0.2208 | 0.9098 |
| 19 | 0.2992 | 0.8753 | 0.2342 | 0.9123 |
| 20 | 0.2902 | 0.8808 | 0.2128 | 0.9285 |

Graph 1 depicts the accuracy model which shows the data accuracy curves for testing and training datasets. Both the datasets accuracy are nearly comparable, showing that the model can almost accurately predict live data with minimal overfitting




**Graph - 1**: Accuracy model

Graph 2 depicts the data loss model. After each epoch, data loss decreases, and the training and testing data sets produce equivalent findings which is an ideal situation [11].




**Graph - 2**: Loss model

The model's accuracy (Fig 3) was determined to be **92.85**

```
print(model.evaluate(test_data,test_target))
29/29 [==============================] - 13s 445ms/step - loss: 0.2128 - accuracy: 0.9285
[0.21276439726352692, 0.9284928441047668]
```

**Fig - 3**: Output of accuracy of model

### 3.1.4 Implementation and results

A hardware webcam is used to video capture in real-time. Cascade classifier, which is a pre-trained model provided by OpenCV [12] is used to extract the region of interest of the image and crop it out and transformed it into a 50x50 image size and that live image is passed through the trained CNN model which labels the input into two classifications. Following successful implementation, we can see that when the individual is wearing a mask, the face in real-time will be labeled as "Mask," and when the individual is not wearing a mask, the face is labeled "No Mask."

## 4.1 NLP for Interactive Chatbot

### 4.1.1 Naive Bayesian Classifiers

Naive Bayes' classification, a probabilistic classifier which is derived from Bayes' theorem which is used for classifying the data into multiple segments. For example, for the dataset used here, the data are classified into categories like 'Buy' and 'Return'. This classifier is based on Bayes' theorem (Fig 4) which is to find the probability of Ck given that event X is true evidence. P(Ck ) is the event occurring when there is a finite amount of outcomes that can occur at equal probability and P(x|Ck ) is the probability after the evidence is seen[13].

$$p(C_k \mid \mathbf{x}) = \frac{p(C_k)\, p(\mathbf{x} \mid C_k)}{p(\mathbf{x})}$$

**Fig - 4** : Bayes Theorem

### 4.1.2 Training and Testing of datasets

A dataset of input data has been created manually which is further divided into training and testing data. 80% of the data in the database is used for training the model and 20% is used for the test the same. Supervised learning is used to train the data, which refers to labeling input and desired output together. Classification technique is used to create this model which efficiently maps data into its target category. It can recognize the intents of the sentences passed by the user and can categorize it into the target category. This algorithm asserts the training data and a function is created, which is required to map into new sentences [14]. An accurate model will allow the user to determine proper results for new data which are not passed in training datasets. Table 2 shows a few examples of data passed into the training dataset.

**Table - 2:** Training model example

| Source request | Decision |
|---|---|
| Will you accept the merchandise as a return | Return |
| How do I give back this sanitizer? | Return |
| Where can I find the milk? | Buy |
| What is the best way for me to return this sanitizer? | Return |
| I need some milk | Buy |
| Do you have return policy for this product? | Return |
| Do you have eggs available? | Buy |
| I cannot find bread here. | Buy |

### 4.1.3 Implementation and output

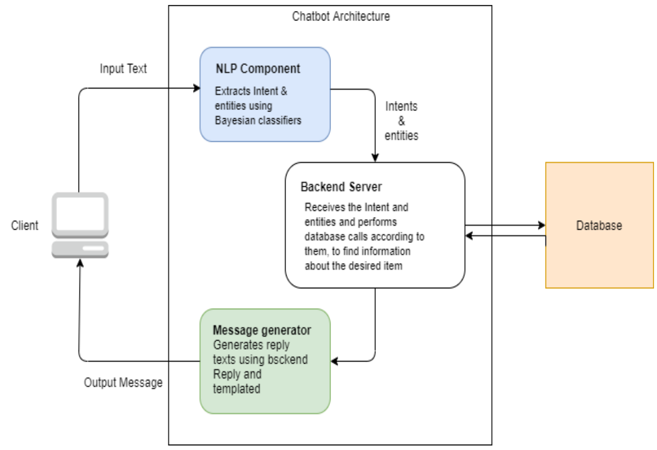



**Fig - 5**: Overview of Chatbot

In this system, a Bayesian algorithm is used to train the dataset to understand the user's intents. The intents are primarily words like 'where', 'how', 'when' etc which are used to determine the requirement of the user[15]. These intents are further used to extract information from the user request like 'shampoo', 'milk' etc which are known as entities. The entities and intent along with each other enable the machine to classify the data into components like whether the customer wants to 'return 'or 'buy' the entity product [16][17].

## 5.1 Sales Analysis and Marketing using Exploratory data analysis

### 5.1.1 Data collection and Pre-processing

The backend database stores the customer's inquiries about any product in columns such as Text and

Timestamp. For cleansing, the data is retrieved from a table in a dataset [18]. This procedure is followed to maintain consistency across the entire dataset and to remove any unnecessary or extraneous data. Extra words like 'I', 'Where', 'You', 'there' and many more are stored in the list and are removed together from the whole dataset using a loop. The words in the list are replaced with whitespace and then splitting the sentence using split and re-joining the whole sentence.

### 5.1.2 Trend Analysis using WordCloud

It is critical to comprehend customer behavior in order to comprehend the pattern of a specific period of time. Consumers' purchasing habits are intertwined. From the dataset which is created after the preprocessing of the input data, the frequency of each product over a given time period is assessed, and a word cloud is constructed to help visualize the data [19]. Figure 6 shows that a few texts are displayed in a larger font than the others. This is because the more often searched data is displayed in larger font size, while the less frequently searched data is displayed in a smaller font size, which clearly displays the trend for that period of time.

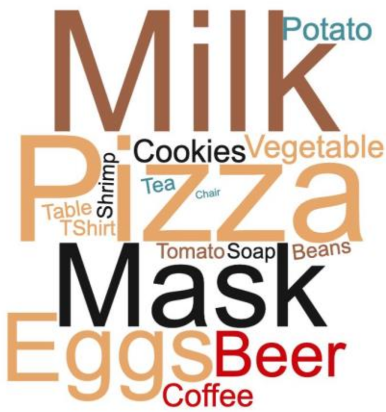
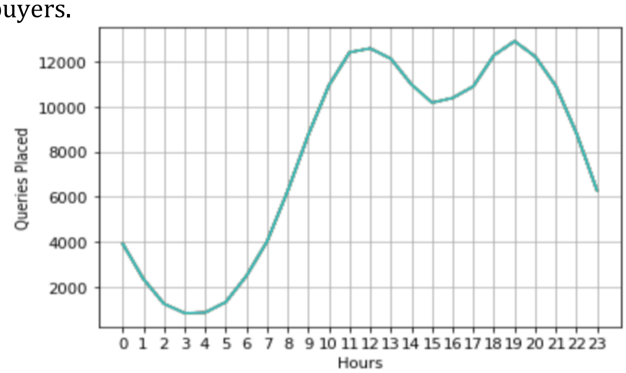



**Fig - 6**: WordCloud for Trend Analysis

### 5.1.3 Data Analysis for Marketing

The customer data that is stored in our internal system can be used for direct and in-direct marketing of the products. For example, in the dataset created after the data is cleaned, we will have columns such as keywords (which will primarily consist of products and reviews, if any) and the timestamp in which the product is searched. We can determine when the kiosk system is most frequently utilized and use the kiosk display banner for any type of product branding. For example, based on the data collected by the system, it can be inferred (Graph 3) that the major number of activities is between 11:00 hours and 19:00 hours. Whereas, 1:00 hours to 5:00 being the lowest for the kiosk activity. So, within the high active time period, we can run an advertisement that will be effective in increasing product interaction among a large number of buyers.




**Graph - 3**: Kiosk activity time Graph

## 6. CONCLUSIONS

In this study, a business automation process is proposed that reduces any kind of human interaction inside a closed space, which has been described as the new normal since the outbreak of the pandemic and also lowers the shop's operating costs by reducing the shop's reliance on human shop assistants. This kiosk is trained to detect whether the customer entering the shop is wearing a mask. This system uses NLP to understand customer's queries and gives feedback according to the query. By training the model with a variety of different sorts of data and enhancing the model's accuracy, the system's computer vision can be improved. In the chatbot, a log file will be maintained for all the requests which the kiosk is unable to process which will be later accessed by the admins to train the model in a more efficient way. A constant update patch will be required to handle various queries from the user. The data which are stored can also be further analyzed to maintain the inventory [20] of the store in a more efficient way which will make sure the on-demand products are in stock while reducing overstocking of items that have fewer sales.

## REFERENCES

[1] Khataee, H., Scheuring, I., Czirok, A. et al. Effects of social distancing on the spreading of COVID-19 inferred from mobile phone data. Sci Rep 11, 1661 (2021). https://doi.org/10.1038/s41598-021-81308-2

[2] Howard, Jeremy & Huang, Austin & Li, Zhiyuan & Tufekci, Zeynep & Ždímal, Vladimír & Westhuizen, Helene-Mari & Delft, Arne & Price, Amy & Fridman, Lex & Tang, Li-Han & Tang, Viola & Watson, Gregory & Bax, Christina & Shaikh, Reshama & Questier, Frederik & Hernandez, Danny & Chu, Larry & Ramirez, Christina & Rimoin, Anne. (2020). Face Masks Against COVID-19: An Evidence Review. 10.20944/preprints202004.0203.v1.

[3] A. Das, M. Wasif Ansari and R. Basak, "Covid-19 Face Mask Detection Using TensorFlow, Keras and OpenCV," 2020 IEEE 17th India Council International Conference (INDICON), 2020, pp. 1-5, doi: 10.1109/INDICON49873.2020.9342585.

[4] Iacobucci, Dawn & Petrescu, Maria & Krishen, Anjala & Bendixen, Michael. (2019). The state of marketing analytics in research and practice. Journal of Marketing Analytics. 7. 10.1057/s41270-019-00059-2.

[5] Kulkarni, Chaitanya & Takalikar, Mukta. (2020). CSEIT183535 | Analysis of REST API Implementation.

[6] Omkar Gaurav. (2020). Face Mask Detection Dataset. [Datafile]. Retrieved from https://www.kaggle.com/omkargurav/face-mask-dataset

[7] BalajiSrinivas. (2020). Face Mask Detection. [Datafile]. Retrieved from https://github.com/balajisrinivas/Face-Mask-Detection/tree/master/dataset

[8] A. Mikołajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," 2018 International Interdisciplinary PhD Workshop (IIPhDW), 2018, pp. 117-122, doi: 10.1109/IIPHDW.2018.8388338.

[9] O'Shea, Keiron & Nash, Ryan. (2015). An Introduction to Convolutional Neural Networks. ArXiv e-prints.

[10] Yamashita, R., Nishio, M., Do, R.K.G. *et al.* Convolutional neural networks: an overview and application in radiology. Insights Imaging 9, 611–629 (2018). https://doi.org/10.1007/s13244-018-0639-9

[11] M. M. Rahman, M. M. H. Manik, M. M. Islam, S. Mahmud and J. -H. Kim, "An Automated System to Limit COVID-19 Using Facial Mask Detection in Smart City Network," 2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), 2020, pp. 1-5, doi: 10.1109/IEMTRONICS51293.2020.9216386.

[12] S. Nie, Z. Jiang, H. Zhang, B. Cai and Y. Yao, "Inshore Ship Detection Based on Mask R-CNN," IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, 2018, pp. 693-696, doi: 10.1109/IGARSS.2018.8519123.

[13] Kaviani, Pouria & Dhotre, Sunita. (2017). Short Survey on Naive Bayes Algorithm. International Journal of Advance Research in Computer Science and Management. 04.

[14] Akinsola, J E T. (2017). Supervised Machine Learning Algorithms: Classification and Comparison. International Journal of Computer Trends and Technology (IJCTT). 48. 128 - 138. 10.14445/22312803/IJCTT-V48P126.

[15] Jiao, Anran. (2020). An Intelligent Chatbot System Based on Entity Extraction Using RASA NLU and Neural Network. Journal of Physics: Conference Series. 1487. 012014. 10.1088/1742-6596/1487/1/012014.

[16] Lalwani, Tarun and Bhalotia, Shashank and Pal, Ashish and Rathod, Vasundhara and Bisen, Shreya, Implementation of a Chatbot System using AI and NLP (May 31, 2018). International Journal of Innovative Research in Computer Science & Technology (IJIRCST) Volume-6, Issue-3, May-2018, Available at SSRN: https://ssrn.com/abstract=3531782 or http://dx.doi.org/10.2139/ssrn.3531782

[17] Deshmukh, Ratnadeep & Wangikar, Vaishali. (2011). Data Cleaning: Current Approaches and Issues.

[18] Ayanouz, Soufyane & Anouar Abdelhakim, Boudhir & Benhmed, Mohammed. (2020). A Smart Chatbot Architecture based NLP and Machine Learning for Health Care Assistance. 10.1145/3386723.3387897.

[19] Depaolo, Concetta & Wilkinson, Kelly. (2014). Get Your Head into the Clouds: Using Word Clouds for Analyzing Qualitative Assessment Data. TechTrends. 58. 38-44. 10.1007/s11528-014-0750-9.

[20] Seyedan, M., Mafakheri, F. Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities. J Big Data 7, 53 (2020). https://doi.org/10.1186/s40537-020-00329-2