iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.42099

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

Virtual Shopping Assistant for Online Fashion Store

Prof. Anushree Deshmukh[1], Smit Shah[2], Heena Puthran[3], Naisargi Shah[4]
Department of Information Technology, Mumbai University, India
Manjara Charitable Trust's Rajiv Gandhi Institute of Technology, Mumbai, India.

Abstract: ***A new way for individuals to interact with computer systems will be done through chatbots or conversational interfaces. Historically, introducing a matter answered by a software package involves employing a program or filling out a type. The technology at the core of the increase of the chatbot is NLP i.e., Natural Language Processing.***
Sequence to Sequence (often abbreviated to seq2seq) fashions is a specific type of Recurrent Neural Network architectures that we commonly use (but no longer restricted) to clear up complicated Language issues like Machine Translation, Question Answering, growing Chatbots, Text Summarization, and so forth.
Recent advances in machine learning have greatly improved the accuracy and effectiveness of NLP, creating chatbots a viable choice for several organizations like e-commerce, Customer service, Conversational apps, social media, Sales/Marketing/Branding, as Voice modules, Travel industry, Medicine, Hospitality, Human Resources etc. An NLP primarily based chatbot is a pc software or synthetic brain that communicates with a patron by means of textual or sound strategies This improvement in NLP is firing a great deal of additional analysis and research which should lead to continued growth in the effectiveness of chatbots in the years to come.
Stochastic gradient descent (often abbreviated SGD) is an iterative approach for optimizing an goal characteristic with appropriate smoothness homes (e.g. differentiable or sub differentiable). Usage of Chatbots can also prove to be beneficial in ways like economically offering 24/7 service, improving customer satisfaction, reaching a younger demographic, reducing costs, increasing revenue and much more.
Keywords: Chatbots, Natural Language Processing (NLP), Stochastic Gradient Descent (SGD), Sequential Model, Machine Learning.

I. INTRODUCTION

Today most websites depend on a menu primarily based on navigation and a search bar to grant data to the user. However, websites with a giant quantity of content material and poorly structured navigation can make it challenging for a person to locate the facts effortlessly and quickly. On occasion, if you reflect on consideration of an online purchasing portal, it has a massive catalog of products. Browsing through the merchandise can be difficult and time-eating given the range of aspects a product can have. To get an extra element of a product, humans are probably to ask alternatively than study the description about it. Virtual Assistants for consumer providers have turned out to be greater and greater famous with customer-oriented businesses. Most of them are constructed from human conversations in the past, which is easy however confronted with troubles of information scale and privacy. Most of the time, clients want to wait in line to get an aid body of a worker's answer, which is much less advantageous and tough to scale up. Meanwhile, clients might also have privacy issues about the conversations, as a result, conversations with clients can't be effortlessly leveraged to teach a chatbot. FAQs are in most cases furnished on the company's internet site to inform their carrier and product. It's simply that the FAQ is normally much less interactive and provides too much data that is much less practical. A chatbot can be used as a choice in supplying FAQs. This is a promising choice in contrast to the usage of search and kind-based tools. It makes it less complicated for the person to locate information. The chatbot has pre-programmed responses, however, it can work with dynamic data from a person's message to make an applicable dialog and endorse applicable information. Given a precise product web page and a purchaser question, the chatbot selects a satisfactory reply from current facts sources. It carries a set of today's NLP and computing devices gaining knowledge of techniques. The two foremost elements in this challenge are the internet site and the chatbot. They are built-in seamlessly to grant a precise personal experience. The product database is impartial to the saved responses; more modern products underneath the respective class can be without problems brought and eliminated and require no amendment of the saved chatbot responses.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

II. RELATED WORKS

The chatbot developed in this project is a First-party chatbot. First-party chatbots refer to conversation engines developed by large enterprises for their own business to improve customer service quality and reduce overall customer service budget like the 'Super Agent'. It uses sub-engines to filter queries and if the probability of the correctness is maximum, that response is generated to the user. [1] Another use of chatbot developed using ML is to use it as a digital organizer to provide variety of services to its master. It expands its digital abilities to organize events, order food, play music, travelling guide, game prediction etc. [3] When a query is input, the Chatbot calls API of AIML interpreter where the query is processed. AIML Interpreter exams the textual content of entry based totally on understanding in AIML documents uploaded into AIML Interpreter. Every time customers chat with a bot, it calls Pandorabots API. [5] There are a lot of options to choose from, on the internet. Thus, having a chatbot that personalises products and caters to one's distinct needs can prove to be quite beneficial. This can mean that the bot acts as a personalised assistant. [2] To familiarise civil crowd with one's features and services, large organisations say, Banks may also put-up FAQs on their websites. This can become tiresome to read and can be enhanced by the use of virtual answering devices. Thus, as suggested by authors in [4] NLP is used so that the system can understand user queries in the form of natural language. Cosine similarity algorithm finds similarities between queries and patterns in the knowledge base. Patterns with the highest cosine values are considered to be most similar to user queries so they can be used as a response to user queries. Parse tree is used to check sentence structure of user query; the meaning of the sentence is deduced here.

III. PROPOSED WORK

A. *Methodology*

1) *SGD Algorithm:* The term "stochastic" refers to a system or process that has a random probability associated with it. As a result, instead of selecting the entire data set for each iteration in Stochastic Gradient Descent, a few samples are chosen at random. The term "batch" is used in Gradient Descent to refer to the total number of samples from a dataset that are utilized to calculate the gradient for each iteration. The batch is taken to be the entire dataset in standard Gradient Descent optimization, such as Batch Gradient Descent. While using the entire dataset is really valuable for arriving to the minima in a less noisy and random manner, the challenge occurs when our datasets become large. If one's dataset contains a million samples, then using a standard Gradient Descent optimization technique, one must use all one million samples to complete one iteration of the Gradient Descent, and this must be done for each iteration until the minima are achieved. The use of SGD Classifier(loss='log') outcomes in logistic regression, i.e. a mannequin equal to Logistic Regression which is geared up with the aid of SGD alternatively of being geared up via one of the different solvers in Logistic Regression. Similarly, SGDRegressor(loss='squared_error', penalty='l2') & Ridge clear up the identical optimization problem, by way of unique means.

a) *Natural Language Processing (NLP):* NLP stands for Natural Language Processing. Using NLP technology, you can assist a computer recognize human speech and spoken words. NLP combines computational linguistics that is the rule-based modelling of the human spoken language with smart algorithms such as statistical, machine, and deep mastering algorithms. There are a variety of human errors, differences, and extraordinary intonations that human beings use each and every day in their speech. NLP science permits the computer to understand, process, and reply to massive volumes of textual content unexpectedly in real-time. In day-to-day life, you have encountered NLP tech in voice-guided GPS apps, digital assistants, speech-to-text word advent apps, and different app aid chatbots. This tech has located sizeable use instances in the enterprise sphere the place it's used to streamline processes, display worker productivity, and extend income and after-sales efficiency.

b) *Classification:* The classification SGD Classifier implements an undeniable stochastic gradient descent getting to know pursuits which helps one of a kind loss features and penalties for classification. Below is the choice boundary of a SGD Classifier educated with the hinge loss, equal to a linear SVM.

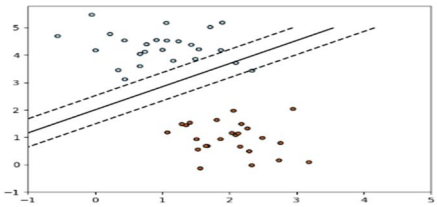


Fig.1: Decision Boundary of SGD Classifier

©IJRASET: All Rights are Reserved

c) *Regression:* The type SGDRegressor implements a simple stochastic gradient descent gaining knowledge of events which helps distinct loss features and penalties to in shape linear regression models. SGDRegressor is nicely suitable for regression troubles with a giant variety of coaching samples (> 10.000), for different troubles we advise Ridge, Lasso, or Elastic Net.

2) *IDE:* A built-in improvement surroundings (IDE) is a software program software that offers full software program improvement facets to laptop programmers. It was once employed in the provided undertaking and aided in the writing of code, deployment, and execution of the underlying model, as nicely as reaching the critical results.

a) *Visual Studio Code:* Microsoft's Visual Studio Code is some built-in improvement surroundings for Windows, Linux, and Mac OS X. Debugging, syntax highlighting, sensible code completion, snippets, code refactoring, and embedded Git are amongst the features. Users can personalize the theme, keyboard shortcuts, and preferences, as nicely as deploy extensions that provide new features.

3) *Libraries Used*

a) *Pandas:* Pandas is a Python bundle for statistics manipulation and analysis. It is primarily based on the Matplotlib and NumPy libraries. As a result, it has a huge vary of aspects for each managing and visualizing data.

b) *NumPy:* NumPy is a Python programming language. NumPy is a library that consists of multidimensional array objects and a series of features for manipulating them. NumPy lets in you to behavior mathematical and logical operations on arrays.

c) *Keras:* Keras is a free open-source Python library for constructing and analyzing deep-learning models that is both powerful and simple to use. It covers Theano and TensorFlow, two efficient numerical computation frameworks, and allows you to create and train neural network models with just a few lines of code.

d) *TensorFlow:* This is a program that allows you to create tensor. TensorFlow is a machine learning software library that is free and open-source. It can be used for a variety of applications, but it focuses on deep neural network training and inference. TensorFlow supports both CPUs and GPUs, which is one of the biggest advantages for tensorflow. It also compiles faster than other deep learning libraries such as Keras and Torch.

4) *Framework Used*

a) *Django:* Django is a high-level Python internet framework that permits speedy improvement of tightly closed and maintainable websites. Built by means of skilled developers, Django takes care of plenty of the problem of net development, so you can focal point on writing your app barring desiring to reinvent the wheel. It is free and open source, has a thriving and energetic community, terrific documentation, and many selections for free and paid-for support.

B. *System Architecture*

From the user's perspective, the internet site has a chat overlay the usage of which the consumer can chat with the bot. Any records the chatbot requires, the person immediately enters into the message window. The chatbots take this enter and suit it with the programmed responses. It then affords facts in its responses and in the structure of hyperlinks to the appropriate products.

C. *Process Flow*

The discern proven beneath illustrates the usual working float of the system. To put quickly into words, it can be stated that consumer inputs his question in the chatbot and the question is processed at some stage in which chatbot appears at already existing records strings. If question enter and question database match, then chatbot offers the fabulous response from its database.

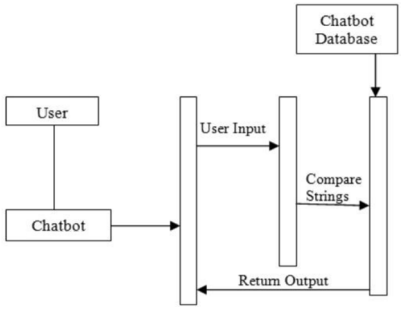


Fig.2: Process flow diagram

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

D. Data Set

For current purposes, the hand-build dataset in the form of "Intents" has been integrated in the system. Readymade datasets are easily available on websites like GitHub, Kaggle but for testing purposes of accuracy, that has not been used at this point. Although, it can be a future scope for the same.

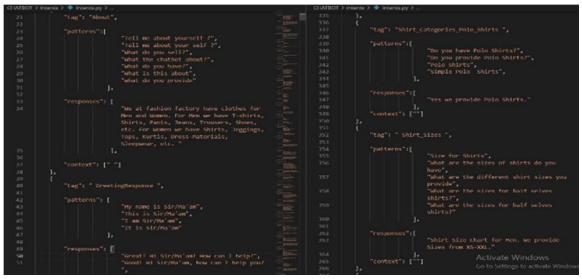


Fig.3: Intents

IV. EXPERIMENTAL ANALYSIS

A. Model Implementation

The two foremost aspects are the internet site and the chatbot. They are built-in seamlessly to supply a suitable person experience.

- *Website:* The internet site is coded in HTML/CSS with PHP used for scripting. The internet site has a MySQL database that shops the product small print and inventory. The internet site will be built-in with Chatbot the usage of Django Framework.
- *Chatbot:* The Chatbot implementation will be achieved with the aid of enforcing a Sequential Model. The Sequential Model is defined below. Our task does now not comprise any dataset. The dataset will be the Intents that we have created. Intents are quick messages which will comprise a "tag, pattern, and response". The chatterbot makes use of the Sequential model, to fetch responses based totally on person input. Sequential Model is written in python and is an easy scripting language for giving Genius to chatbots and different conversational entities. It's a simple text, line-based scripting language to is easy to learn, speedy to type, and convenient to examine and maintain. We are the use of NumPy, Pandas, Keras, and TensorFlow libraries. The algorithm which we will be the use of is the SGD (Stochastic Gradient Descent) algorithm.

1) Working Of Chatbot

a) Step 1 – Intents

Intents can be termed as brief messages.

{
tag: 'T1'
pattern: ['p1', 'p2,…..]
response: ['A', 'B','C',….]
},
Example:
tag = 'greeting'
pattern: ['Hi', 'How are you']
response: ['hello, how may I help you']
{
tag: 'T2'
pattern: ['p3', 'p4', …..]
response: ['X', 'Y','Z', ….]
}, and so on.

Tag	Pattern
T1	P1
T1	P2
T2	P3
T2	P4

Tag	Pattern
greetings	Hi
greetings	Hello
greetings	Have a nice day
greetings	How are you etc.

b) Step 2 – Separating the words from the pattern

Example:

Pattern is – "greetings, how may I help you"

Now the words, punctuation, special characters will be separated. It will be stored in a Word Dictionary. Everything will be given an index. Here any duplicate entry will be removed and the prefix & suffix will be neglected.

Index	Word
0	W1
1	W2
2	W3
3	W4
4	W5

Index	Word
0	,
1	?
2	I
3	how
10	may
25	greet
38	help
46	you

Fig.4: Word Dictionary

c) Step 3 – Encoding

All the possible pattern values will be separated and added to the dictionary. Now let's say the size of the dictionary is 38, so accordingly, to the indexes values every time the particular word is present 1 will be assigned, or else 0 will be assigned.

Assuming the size of Dictionary = 46

For Encoding, an array with (size of Dictionary + 1) is formed.

[_ , _ , _ , ... _ , _ , _ , _]

Index → 0th 1st 2nd 43rd 44th 45th 46th

So, according to the Table drawn in the previous section, the value '1' shall be assigned to words present at the index in array and '0' shall be assigned to those not present in array.

[1 , 1 , 1 , 1 ..... 1 , ... 1 , ... 1 , ... 1]

Index → 0th 1st 2nd 3rd 10th 25th 38th 46th

Fig.5: Encoding

Now, this encoding is passed on to the Train part.

Train

[1 1 1 1 0 . . . 0 1 1]

[1 1 0 1 . 0 . . 0 1 1]

[0 0 1 1 0 . . . 0 0 1]

This Train will be the input to our Sequential Model.

B. Model Output

EPOCHS	LOSS	ACCURACY
90	0.6203	0.8173
100	0.4526	0.8635
130	0.4402	0.8804
170	0.6406	0.8239
200	0.8954	0.8763

Table.1. Epochs vs loss and accuracy

So we are getting the highest accuracy of 0.8804 at 130 epochs and the loss we are getting is 0.4402.

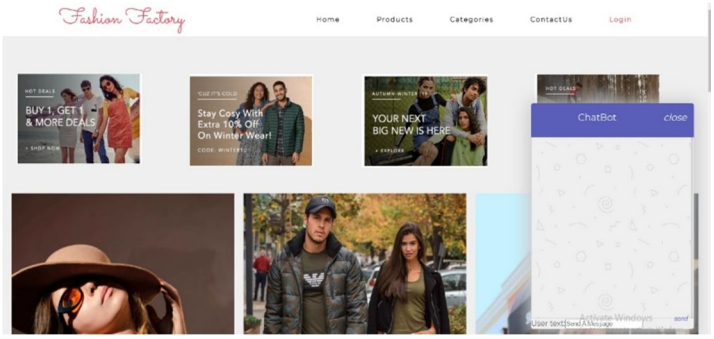


Fig.6: Website snippet

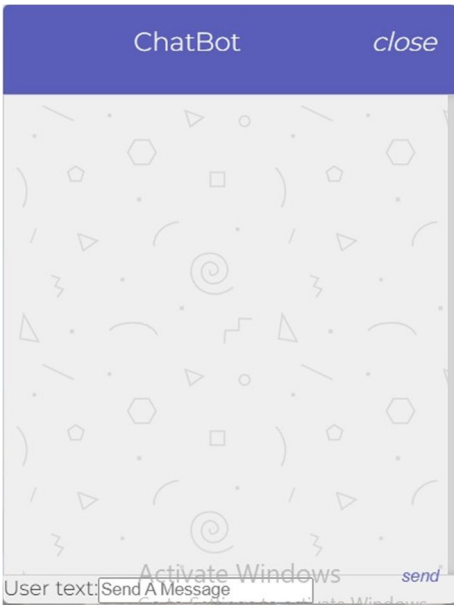


Fig.8: Chatbot Interface

The following graphs give insights on the loss and accuracy of the model. The model has achieved highest accuracy of 0.8804 at 130 epochs. At this accuracy, we will also be able to predict the accuracy of answers responded back to user by the chatbot.

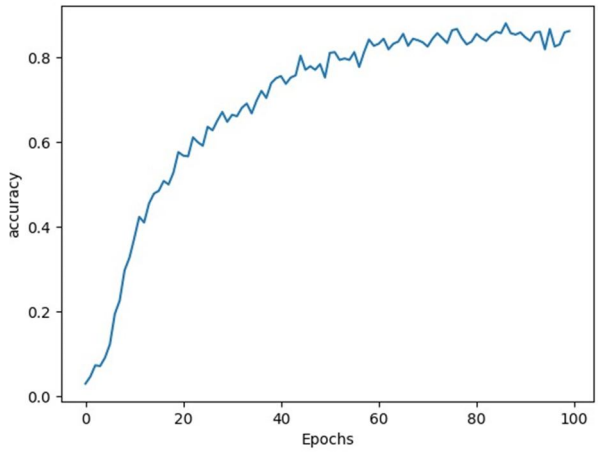


Fig.9: Epochs vs Accuracy graph

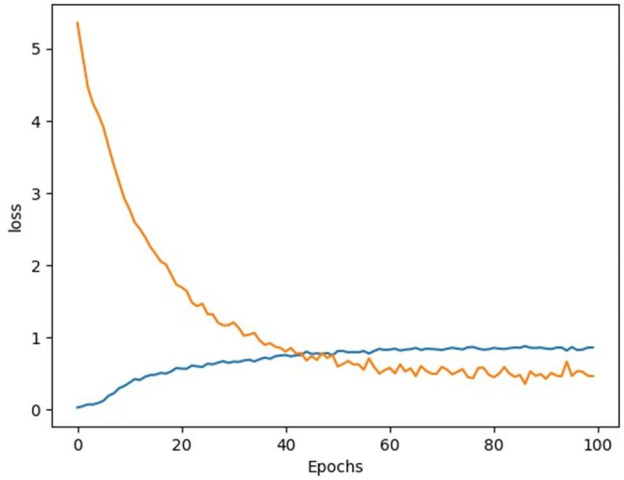


Fig.10: Epochs vs Loss graph

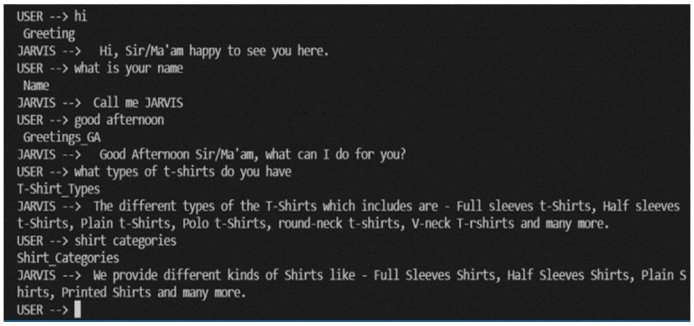


Fig.11: Working of Chatbot on Terminal

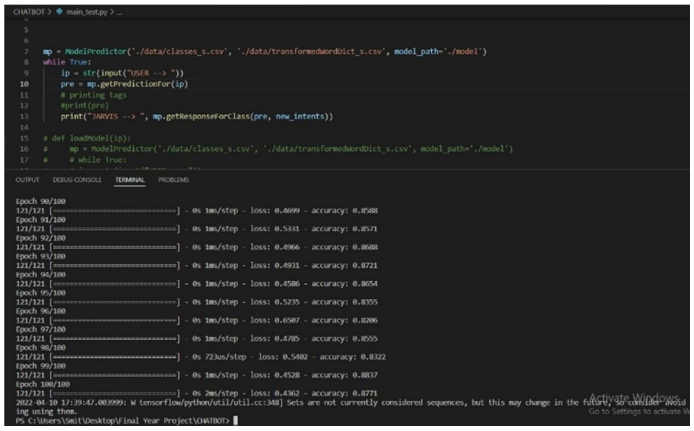


Fig.12: Accuracy vs Epochs on Terminal

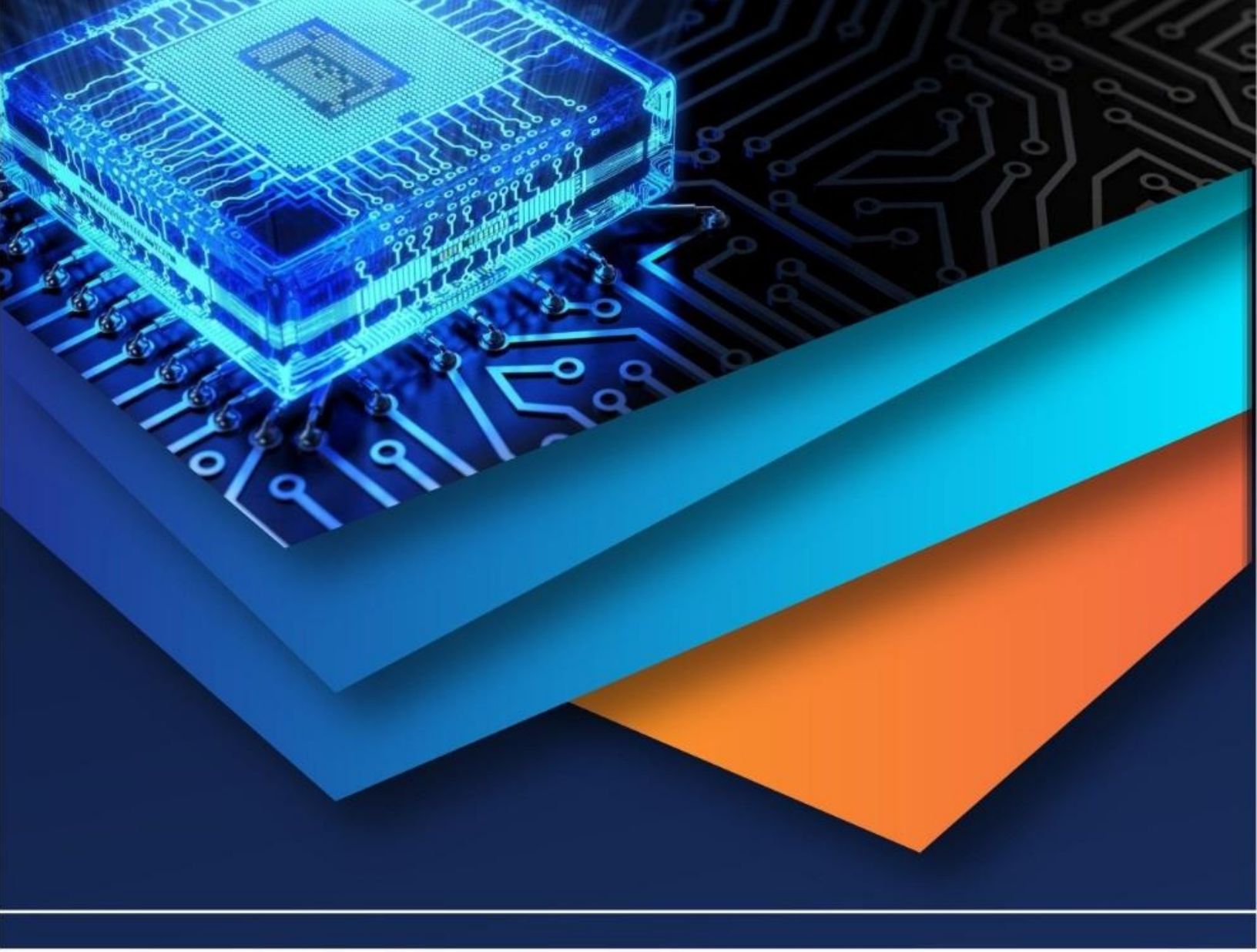
V. CONCLUSION

Tech-savvy shoppers of nowadays are continually on the lookout for the fantastic and most customized patron experiences. It can appear a not possible undertaking to fulfil the onslaught of ever-evolving demands. However, there is one answer primed to fulfil the contemporary customer, and that is a chatbot. With a chatbot, your company can effortlessly provide tremendous assist and hostilities decision any time of day, and for a giant volume of clients simultaneously. According to Microsoft, 90% of customers assume an on line portal for purchaser service. As a vast component of enterprise evolution, the want for AI-powered chatbots will solely proceed to rise. Now is the time to deploy a chatbot solution so that your company doesn't get left behind.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

REFERENCES

[1] "Super-Agent: A Customer Service Chatbot for E-commerce Websites" - Lei Cui∗, Shaohan Huang∗, Furu Wei, Chuanqi Tan, Chaoqun Duan, and Ming Zhou Microsoft Research Asia {lecu, shaohanh, fuwei, v-chutan, v-chadu, mingzhou} @microsoft.com
[2] "An E-Commerce Website based Chatbot" - Siddharth Gupta, Deep Borkar, Chevelyn De Mello, Saurabh Patil Department of Computer Engineering Xavier institute of Engineering, Mumbai, India, Vol. 6(2)
[3] "ML Based Virtual Personal Assistant" - V. Lalitha, A. Dinesh, L. Parameswaran, S. Dinesh Kumar, Department of CSE, Sri Sairam Engineering College, Chennai, Vol 6, Issue 7, July 2019
[4] "Chatbot using NLP to Provide Banking Information" - Abidah Elcholiqi, Aina Musdholifah, Master Program of Computer Science; FMIPA UGM, Yogyakarta, Indonesia Department of Computer Science and Electronics, FMIPA UGM, Yogyakarta, Indonesia Vol.14, No.1, January 2020
[5] "Implementing a Chatbot for E-Commerce Site Using Artificial Intelligence Markup Language (AIML)" - Stefanus Ardhito Prasetya, Alva Erwin, Maulahikmah Galinium, Faculty of Engineering and Information Technology, Swiss German University, 2018.
[6] Abbafati, L. 2017. "Disruptive Innovation in the Luxury E---commerce: The Case of Farfetch. Dissertation", Academic Year 2016---2017, Department of Business and Management, LUISS University
[7] Allied Market Research 2017. "Online Clothing Rental Market by End-to-End-Users (Women, Men, and Kids) and Clothing Style (Ethnic, Western, and Others) --- Global Opportunity Analysis and Industry Forecast", 2017---2023. Mar 2017 https://www.alliedmarketresearch.com/online---clothing---rental---market consulted on 06/08/2019
[8] Anderson, C., 2006. "The Long Tail: Why the Future of Business is Selling More for Less".
[9] BAGAAR, 2018. "Digital leading the way to more sustainable textile & fashion industries" (blo https://www.bagaar.be/insights/digital---leading---the---way---to---a---more--- sustainable---textile---fashion---industry consulted on 09/08/2019
[10] Barthes, R., 1967. "The Fashion System, trans. Ward M., Howard R. New York: Random House"



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)