

# Financial Literacy Chatbot

Yashraj Mulwani

dept. of Computer Engineering  
V.E.S. Institute of Technology  
Mumbai, Chembur

[2020.yashraj.mulwani@ves.ac.in](mailto:2020.yashraj.mulwani@ves.ac.in)

Gauri Nagral

dept. of Computer Engineering  
V.E.S. Institute of Technology  
Mumbai, Chembur

[2020.gauri.nagral@ves.ac.in](mailto:2020.gauri.nagral@ves.ac.in)

Kritika Yadav

dept. of Computer Engineering  
V.E.S. Institute of Technology  
Mumbai, Chembur

[d2020.kritika.yadav@ves.ac.in](mailto:d2020.kritika.yadav@ves.ac.in)

Bhavesh Bhatia

dept. of Computer Engineering  
V.E.S. Institute of Technology  
Mumbai, Chembur

[2020.bhavesh.bhatia@ves.ac.in](mailto:2020.bhavesh.bhatia@ves.ac.in)

Mrs. Abha Tewari

dept. of Computer Engineering  
V.E.S. Institute of Technology  
Mumbai, Chembur

[abha.tewari@ves.ac.in](mailto:abha.tewari@ves.ac.in)

**Abstract—** A Financial Literacy chatbot based on Artificial Intelligence, Machine Learning, NLP (Natural Language Processing) technologies, and many modern frameworks, will help to provide a user-friendly and accurate solution to the problem and queries. Being a very interactive chatbot, more data would be collected from the user. With proper processing using the technologies mentioned above, the user can get awareness and guidance on specific Government related schemes and how to opt for them. Using NLP, the chatbot also seems to provide a beginner-friendly summary of the documents related to the schemes and help the user to get clear guidance. This project also aims at implementing voice commands as input and output and aims at making a successful multi-lingual prospect. This project provides the latest news about government schemes.

**Keywords—** Financial Chatbot, Financial Transaction, Artificial Intelligence, Natural Language Processing (NLP), Awareness, Guidance, Machine Learning

## I. INTRODUCTION

Managing finances through financial literacy is one of the most undervalued talents that might impact a nation's future in a world that is changing quickly. Governments have professional analysts to manage and forecast the financial future of the nations while researching the shifting geopolitical linkages at a macro level. The need to include financial literacy in the micro-levels of the economy is emerging, nevertheless.

The COVID-19 epidemic has increased the demand for financial education even more to boost the expansion of the economy. In a couple of months, the epidemic rendered millions of people unemployed, leaving them with mounting debt. Although the economy is gradually improving and investments in the retail sector are rising, the impending danger of a lack of financial literacy might have long-term effects.

Personal economic affairs, calculating, and investing are just any of the fiscal capabilities that fall under the umbrella term of "commercial knowledge". An individual's friendship accompanying money is a byproduct of the groundwork of monetary education. Experts frequently hold that economic instruction is a life process that is to say best started as early as doable. Financial properties and duties have currently become commonplace between folk of all financial classes. Financially

savvy community are less likely to enhance martyrs of commercial trickery.

Data suggests that financial literacy among Indians is about 24%. With a literacy rate of close to 80%, India is home to over one-fifth of the world's population. Unfortunately, just 24% of the country's population is financially literate. For a nation that depends on the economy for its prosperity, the fact that three-fourths of the population in India does not know about or appreciate the urgent need for good financial management is concerning. Only 15% of the country's population was able to manage their money and savings in 2013. However, during the past eight years, the proportion has significantly improved.

An ordinary person finds it difficult to keep track of their money outlays. It might be quite demanding to keep track of your earnings, expenses, debts, taxes, and investments. For a sustainable present and future, financial planning is necessary. Financial advisors provide advice to their customers on how to budget their spending so that it doesn't exceed their income and how to invest in various markets to increase their income. The markets are always changing, therefore the advisors must be alert at all times to keep up. They demand a high price for their services since they must keep up with developments and counsel customers on the best investments.

They demand a high price for their services since they must keep up with developments and counsel customers on the best investments. Even the financial sector has grown less dependent on humans as a result of technological improvements. Machine learning algorithms-based web apps have begun to take off in the market. However, their inaccurate predictions and suggestions made them not very trustworthy. However, the introduction of artificial intelligence (AI) radically altered the situation. Neural networks enable AI to make predictions more quickly and accurately than anybody else. The field will become more lucrative as the usage of AI for personal financial advisors grows in popularity and algorithm accuracy increases.

Now, the financial sector has grown less dependent on humans as a result of technological improvements. Machine learning algorithms-based web apps have begun to take off in the market. The introduction of artificial intelligence (AI) radically altered the situation. Neural networks enable AI to make predictions more quickly and accurately than anybody else. The field will become more lucrative as the usage of AI for

personal financial advisors grows in popularity and algorithm accuracy increases.

A Financial Literacy chatbot based on Artificial Intelligence, Machine Learning, NLP (Natural Language Processing) technologies, and many modern frameworks, will help to provide a user-friendly and accurate solution to the problem and queries. Being a very interactive chatbot, more data would be collected from the user, and with proper processing using the technologies mentioned above, the user can get awareness as well as guidance to specific Government related schemes.

**Contributions:** Contributions: The following are the main contributions of the paper:

- A chatbot application software to simplify user-software interaction.
- To suggest safe and advisable government schemes.
- Initial user data inputs, storage, and proper processing using ML models with the assistance of AI.
- To achieve maximum efficiency of the ML model in the predictions made.
- To enable voice-based input-output for easier interaction.
- To support input and output in multiple languages.
- To educate users on basic financial aspects.

## II. BACKGROUND

There is a brief overview of chatbots in (Section II-A), artificial intelligence in (Section II-B), machine learning in (Section II-C), and natural language processing (in Section II-D).

### A. ChatBot

The same or similar repetitive tasks can be efficiently completed by chatbots, which are interactive systems. They are popular because they free up time for humans to focus on undiscovered, intellectual and mindful ideas and actions that require human attention.

**Classification:** Bots are divided into two types:

- **Keyword recognition-based:** Chatbots use keyword recognition to respond to users using programmable keywords and an AI tool called Natural Language Processing (NLP). When asked a lot of similar questions, NLP chatbots struggle due to keyword overlaps. Examples of chatbots combine menu/button-based and keyword recognition-based functionality are common.
- **Voice-bots:** Businesses are increasingly using voice-based chatbots or voice bots to make conversational interfaces more vernacular. Voice bots have been on the rise for the last couple of years due to their convenience, making it easier for customers to speak rather than type. A voice-activated chatbot brings frictionless experiences directly to the end customer.

**Mechanisms:** The User Interface (UI), via which a user interacts with a text chatbot and asks questions or makes choices, is its front-facing element. Voice-activated chatbots use the microphone of the associated devices to take commands and input from the user. Option-based text chatbots are the simplest to create, while dynamic textual and voice chatbots require extra parts and sophisticated algorithms. NLP techniques such as Sentiment Analysis and Part-of-Speech Tagging are used to comprehend queries and deliver relevant output.

**Financial chatbots:** Financial chatbots are making common investment easier for consumers by simplifying economic affairs. More banks are implementing bots to work as in essence helpers to clients, adding adeptness to movements. Chatbots are powerful enough to warn public on their commercial health and help ruling class take a reliable decision concerning an expense or conditional plan. AI-powered monetary aids chatbots are able to offer reliable pieces of advice established the analysis of consumer's giving habits and report following. We focus particularly on offering commercial recommendation inside the scope concerning this paper.

### B. Artificial Intelligence

Artificial Intelligence is the process of making calculating copy the way persons believe, appropriating traits from human thinking and applying bureaucracy as algorithms in a mathematical habit. It is the endeavor of building systems gifted accompanying the intellectual processes characteristic of persons, such as reason, find message, statement, and gain experience. Despite current incident in estimating facilities, there are now no structures that can counterpart human adaptability across a best range of endeavors or those lacking a solid amount of background information. History implies that calculating is capable of being efficiently set up to act troublesome jobs, in the way that improving traditional, fogged pictures or artwork into a clear, new figure.

### C. Machine Learning

Machine Learning is an arm of Artificial Intelligence that uses dossier and algorithms to determine about new belongings. Regression and categorization algorithms are ultimate established, while reversion algorithms are used to forecast effects. Machine learning algorithms have two classifications: directed and alone. Unsupervised algorithms depend dossier to "determine" on their own, while directed algorithms demand a preparation dataset. Subsets of Machine Learning involve neural networks, deep education, and support knowledge.

### D. Natural Language Processing(NLP)

Natural language processing (NLP) is a arm of machine intelligence that handles comprehending and altering uttered and human language. A field of AI that has been around for a short period but has currently acquire recognition as machine intelligence and deep learning have progressive. NLP is secondhand in requests to a degree Google Translate, Siri,

Alexa, and added personal helpers to define and put oneself in the place of another human terminology. Text inputted into these requests is treated and elucidated utilizing NLP, and it will touch grow in style and veracity accompanying the incident of artificial intelligence.

### III. PROPOSED SYSTEM

The initiative to develop an AI Chatbot that can identify and providing financial advice and information on current, and prospective government initiatives in light of all the problems that underprivileged communities in India experience provides a unique perspective on the applications of AI. According to the research paper authored by Jon T.S Quah and Y.W Chua (Chatbot Assisted Marketing in the Financial Service Industry) Chatbots are playing a significant role in Singapore's financial services.

Their research indicates that an unbiased and concise interaction that a Chatbot provides helps people retain and use the financial information more effectively. Choosing a Chatbot to promote financial literacy is the right call for following reasons:

- As a result, consumer vindication is uncompromised, but the bank saves plenty opportunity and services on help. 24/7 Support Banking chatbots support 24/7 support.
- Enhanced Productivity of Bank Personnel Banking chatbots can help banks humiliate their operators' opportunity on bureaucratic tasks and ordinary query judgment.
- Personalized Client Experience Banking chatbots can help banks embody their consumer happening by including news about each consumer's weaknesses and experiences into the conversation.
- Using chatbots, banks can have significant chats on friendly publishing to accumulate concerns and transfer ruling class to the worried areas – it can go at a great distance in construction the association's connected to the internet honor.

This Chatbot will use a talk plan to uphold dialogue accompanying a consumer. Once a consumer inputs a question or a affidavit, the Natural Language Processing Layer will extract keywords from the recommendation, it will tokenize, grammatize, and stem the recommendation to programming language.

The Data Store file that will hold the passage of management blueprints and other contribution alternatives will be treated and distinguished against the consumer recommendation. Multinational Naive Bayes treasure for NLP and passage categorization will be used to find discussion matches for likely sentences each class.

The categorization score recognizes the class with the topmost term couples. The production maybe translated to different vocabularies in addition to visual and audio entertainment transmitted via radio waves output. A consumer will should corroborate every occasion he/she needs to approach computer network use. The initial sign-up will need the consumer to complete a form in consideration of accumulate dossier for further analysis.

Once sufficient data has been collected, the system will also be able to deploy a machine learning-based predictor of the financial plans opted by a customer base comparable to the user. This web application will be designed and tuned considering the technical abilities of the user base

### IV. ARCHITECTURE & IMPLEMENTATION

Here, the first module is a user interface hosted on a website. The user interface will consist of a dialog box where users can say or ask anything. The Chatbot will reply accordingly. For awareness, the user will have to type Awareness in the dialog box. For advice, the user will have to type Advice in the dialog box. In the Advice section the user has to answer to all the relevant questions asked by the chatbot to get accurate results.

The second module consists of various system features such as Text Translation and Voice to Text functionality which will enable greater accessibility for people also consists of a new panel which is integrated using google news which will provide users all the latest news related to all government finance schemes all over India.

The third module is the framework for processing user input and giving appropriate responses. This will be handled by Natural Language Processing. And data will be researched and put in a file which will be accessed by the Chatbot. A UI module consisting of user authentication, form filling and sign in may be required for the purposes of data collection and analysis but it is not necessary.

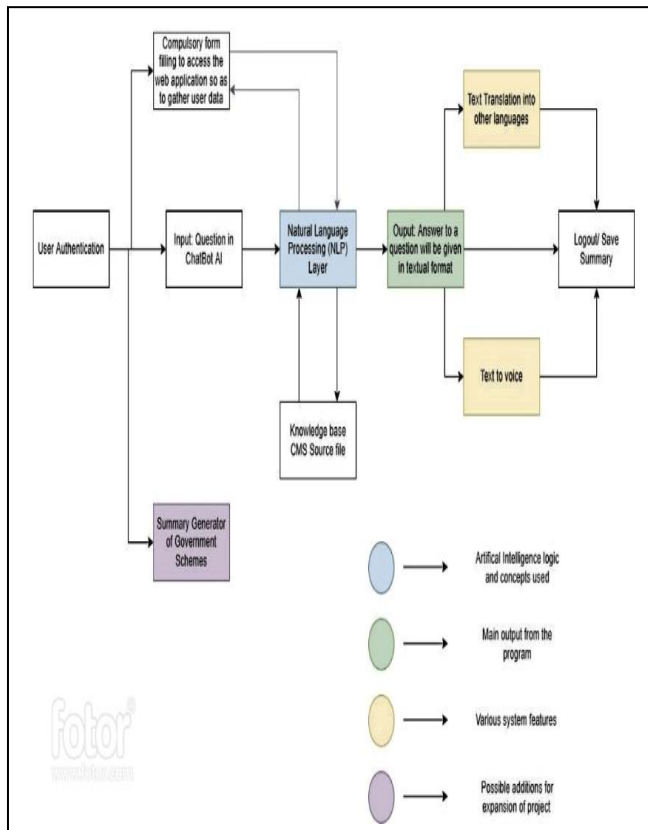


Fig. 1. Framework of Website.

## V. ALGORITHM AND PROCESS DESIGN

```
import nltk
intents = json.load(file)
data= intents
all_words=[ ]
for i in intents
    tokenize(data)
    stem(data)
    bg_data=bag_of_words(data)
    all_words.append(bg_data)
tag = [ ]
for tag in intents{intents}
    tag.append(tags)

training_set_X.layer1.train(all_words)
training_set_X.layer2.train(tags)
save_data = train_set_X.pth

user_input = sentence
word=nlp(sentence)
p=calc_class_probability(word)
c= class
If p>0.75
    print(c.random.response)
else
    print("I did not understand")
```

Fig. 2. Psuedo code for training the model.

The Nave Bayes algorithm attempts to classify text into specific categories so that the chatbot can identify the user's intent, thereby narrowing the possible range of responses.

The invention depends commonality, that basically means that sure conversation should have more burden for classifications based on the commonness of their presentations in that classification. The most easy way of experiment this invention is by using k-fold cross confirmation. This includes preparation the chatbot with sure inputs and their equivalent classifications, then utilizing a test fight assess in what way or manner frequently the chatbot can correctly categorize a likely input. Confusion molds, veracity, precision, and recall will be secondhand for determining the performance of the invention.

## VI. METHODOLOGY APPLIED

The main intelligence behind the Dialog connect is the JSON file. JSON file exists of miscellaneous aims accompanying a type of patterns and appropriate answers. Manually constructed the JSON file utilizing CATHOVEN and QUILLBOT to accumulate differing question patterns. This JSON will before acquire information on Pytorch Machine Learning 3-tier model to reckon the chance appropriate for a likely class that is, a particular tag.

**Knowledge modeling:** This phase determines how knowledge is represented and stored in the knowledge base.

**Conversation flow:** Both the lexicon used by the tutor and the order in which ideas are presented should be defined in this phase.

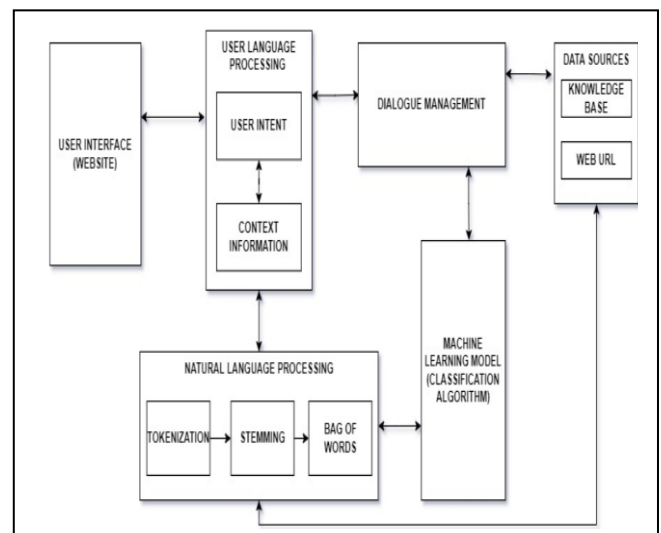


Fig. 3. Block Diagram .

Although many dossier buildings exist to store the computerized data in system, most chat-bot discourse buildings are established trees. Each node in the tree represents a unique response, from a simple greeting to detailed information about previous queries. It is also important to note that in order for the conversation service to determine which

Fig. 4. JSON file consisting of Government Schemes.

response the user is looking for, the similarity between the user input and all known queries must be calculated. This process is usually done by machine learning algorithms using similarity measures between sentences in which each word or character may represent a single dimension, and its accuracy is refined by providing thousands of correctly labeled examples of user inputs.

Programmers can build a complex neural network with ease using PyTorch as it has a core data structure, Tensor, multi-dimensional array like Numpy arrays. PyTorch use is increasing in current industries and in the research community as it is flexible, faster, easy to get the project up and running, due to which PyTorch is one of the top deep learning tools. Thus we have made extensive use of PyTorch in our project.

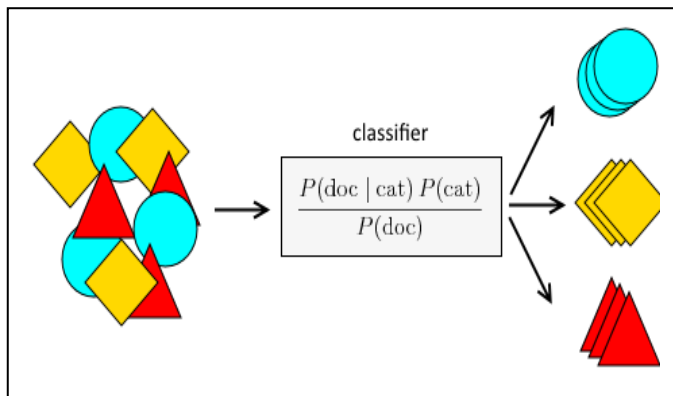


Fig. 5. Naïve Bayes Classifier.

The central methodology is inputting a question or a statement from the user consisting of keywords. The keywords will then be analyzed by the NLP layer to generate an appropriate response.

- Tokenization: breaks down text into smaller semantic units or single clauses
- Part-of-speech-tagging: marking up words as nouns, verbs, adjectives, adverbs, pronouns, etc
- Stemming and lemmatization: standardizing words by reducing them to their root forms
- Stop word removal: filtering out common words that add little or no unique information, for example, prepositions and articles (at, to, a, the).

Keywords will be detected from a CMS source file consisting of data of various government schemes for the NLP algorithm

to compare to rule based approaches relying on hand-crafted grammatical rules need to be implemented.

## VII. HARDWARE & SOFTWARE SPECIFICATIONS

Python Libraries: These are some to the Python Libraries used:

a) *Natural Language Toolkit*: The Natural Language

```
"intents": [
  {
    "tag": "atalyोजना_what",
    "patterns": [
      "What is Atal Pension Yojana?",
      "Atal Pension Yojana",
      "Describe the Atal Pension Yojana.",
      "Tell me about Atal Pension Yojana",
      "What is APY"
    ],
    "responses": [
      "It is the universal social security scheme for all Indians who",
      "It is the national social security programme for all non-taxpa",
      "Atal Pension Yojana is a pension scheme that focuses on helpin",
      "The Atal Pension Yojana is a pension system that aims to provi"
    ]
  }
],
```

```
import torch
import torch.nn as nn

class NeuralNet(nn.Module):
    def __init__(self, input_size, hidden_size, num_classes):
        super(NeuralNet, self).__init__()
        self.l1=nn.Linear(input_size, hidden_size)
        self.l2=nn.Linear(input_size, hidden_size)
        self.l3=nn.Linear(input_size, num_classes)
        self.relu=nn.ReLU()

    def forward(self, x):
        out=self.l1(x)
        out=self.relu(out)
        out=self.l2(x)
        out=self.relu(out)
        out=self.l3(x)
        #no activation and no softmax
        return out
```

Toolkit (NLTK) is the preferred API for NLP (Natural Language Processing) in Python. It is a very powerful tool for pre-processing text data for further analysis, such as ML models. It helps convert text to numbers, which can then be easily consumed by the model.

b) **PUNKT**: Punkt Sentence Tokenizer This tokenizer divides a text into a list of sentences by using an unsupervised algorithm to build a model for abbreviation words, collocations, and words that start sentences. It must be trained on a large collection of plaintext in the target language before it can be used.

```
C:\Users\gauri\abcbot>C:\Users\gauri\Newfolder\Scripts\activate
(base) C:\Users\gauri\abcbot>conda activate base
(base) C:\Users\gauri\abcbot>python -u "c:\Users\gauri\abcbot\train.py"
epoch 100/1000,loss=0.8861
epoch 200/1000,loss=0.8634
epoch 300/1000,loss=0.8055
epoch 400/1000,loss=0.1279
epoch 500/1000,loss=0.8067
epoch 600/1000,loss=0.0002
epoch 700/1000,loss=0.0141
epoch 800/1000,loss=0.0023
epoch 900/1000,loss=0.0904
epoch 1000/1000,loss=0.0000
final loss=0.0000
training complete. file saved to data.pth
(base) C:\Users\gauri\abcbot>
```

Fig. 6. Feed Forward Neural Network Initialization.

c) **TORCH**: Manage the spinned words as you want.. PyTorch is a device getting to know framework primarily based at the Torch library, used for programs consisting of computer vision and herbal language processing, firstly developed through Meta AI and now a part of the Linux basis umbrella. it's far loose and open-source software program released underneath the modified BSD license.

TABLE I. HARDWARE REQUIREMENTS

TABLE II. SOFTWARE REQUIREMENTS

d) **NUMPY**: NumPy may be used to carry out a extensive ype of mathematical operations on arrays. It provides powerful information structures to Python that guarantee efficient calculations with arrays and matrices and it resources an significant library of excessive-level mathematical functions that perform on those arrays and matrices.

e) **PANDAS**: Pandas are honestly powerful. They offer you with a huge set of essential commands and functions which are used to effortlessly examine your facts. we can use Pandas to perform various tasks like filtering your information in keeping with certain conditions, or segmenting and segregating the facts in step with preference, and many others.

## VIII.IMPLEMENTATION & RESULTS

Fig. 7. Training the model.

| Tools                | Language                                   |
|----------------------|--------------------------------------------|
| IDE                  | Jupyter notebook, IDLE, Visual Studio Code |
| Front-End Tech Stack | HTML, CSS, JS, Typescript                  |
| Backend Tech Stack   | Python, NodeJS, PyTorch                    |

An epoch refers to at least one complete bypass thru the entire training dataset during the education method of a neural community. at some point of an epoch, the set of rules methods each schooling example as soon as and updates the version's parameters based at the predictions and mistakes made at the education facts.

The wide variety of epochs is an important hyperparameter that determines how in many instances the model will see the entire education dataset. usually, growing the number of epochs can improve the accuracy of the model, however if the variety of epochs is just too high, the version may overfit to the schooling statistics and perform poorly on new, unseen records. consequently, the most advantageous range of epochs is generally determined through experimentation and validation.

According to the scale of dataset of the government Schemes intents, which has been manually accrued, the

| Resource   | Minimum Requirement                                                                                      | Recommended             |
|------------|----------------------------------------------------------------------------------------------------------|-------------------------|
| CPU        | 2 x 1.8GHz 32-bit (x86)                                                                                  | 4 x 2.4GHz 64-bit (x64) |
| RAM        | 4GB                                                                                                      | 8GB                     |
| Disk Space | 3.5 GB for new installations, 5 GB for upgrades (including temporary files required during installation) | N/A                     |

experimentation resulted to 10 epochs to be the most effective range for the ChatBot. Having experimented with various metric estimation strategies, the nice perfect became Holdout method for splitting the data into train set and test set. The train set is then divided into batches to perform batch

clever schooling.

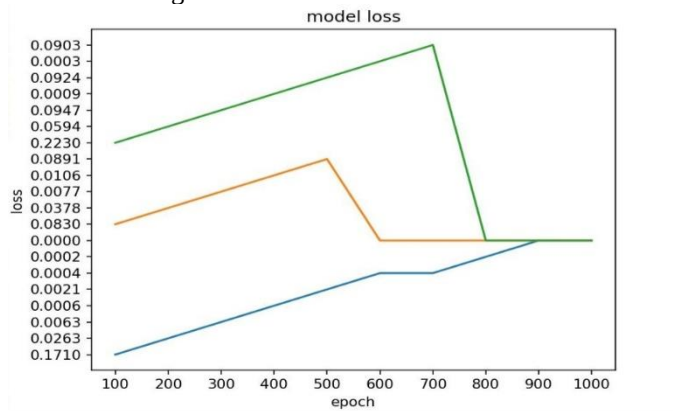


Fig. 8. Graph to show model precision.

The above graph shows 3 iterations of model training parameters. The loss is compared with the number of epochs. As presented in the graph, loss eventually converges to 0 in all the iterations around 10 epochs, thus cementing the best number to be 10.

## IX. CONCLUSION

In this paper, we have presented Financial Literacy chatBot, a Machine learning empowered chatbot for financial transactions. At first, we have formulated a set of requirements based on a common person using this bot for his/her financial related queries. The architecture of Chatbot has been designed to satisfy the formulated requirements and to satisfy the users need. Furthermore, we have evaluated its performance and analysed its accuracy, advantages, and limitations. Using this application, one can execute financial transactions within a chatbot.

## REFERENCES

- [1] Tewari, Abha and Chhabria, Amit and Khalsa, Ajay Singh and Chaudhary, Sanket and Kanal, Harshita, A Survey of Mental Health Chatbots using NLP (April 25, 2021).
- [2] Rukmini Murugesan, V.Manohar, "AI in Financial Sector – A Driver to Financial Literacy", Volume: 7, Issue: 3, Month: July, Year: 2019, ISSN: 2320-4168.
- [3] Annamaria Lusardi, "Financial literacy and the need for financial education: evidence and implications", DOI: 10.1186/s41937-019-0027-5.
- [4] Jon T. S. Quah and Y. W. Chua, "Chatbot Assisted Marketing in Financial Service Industry", DOI: 10.1007/978-3-030-23554-3\_8.
- [5] Ashish Shah, Pratik Raj, Pushpam Kumar, Supriya P and Asha H V, "FinAID, A Financial Advisor Application using AI", IRJTE, ISSN: 2277-3878 (Online), Volume-9 Issue-1, May 2020.