CAN THO UNIVERSITY

COLLEGE OF INFORMATION AND COMMUNICATION TECHNOLOGY

🙞 🕮 🙜

Logo

Description automatically generated

GRADUATION THESIS

BACHELOR OF ENGINEERING IN

INFORMATION TECHNOLOGY

(HIGH-QUALITY PROGRAM)

BUILD A CHATBOT APPLICATION

Student: Nguyen Duy Khang

Student ID: B1910652

Class: 2019-2023 (K45)

Student: Tran Ba Phuong

Student ID: B1908407

Class: 2019-2023 (K45)

Advisor: Dr. Tran Cong An

Can Tho, 12/2022

CAN THO UNIVERSITY

COLLEGE OF INFORMATION AND COMMUNICATION TECHNOLOGY

DEPARTMENT OF INFORMATION TECHNOLOGY

🙞 🕮 🙜

Logo

Description automatically generated

GRADUATION THESIS

BACHELOR OF ENGINEERING IN

INFORMATION TECHNOLOGY

(HIGH-QUALITY PROGRAM)

BUILD A CHATBOT APPLICATION

Student: Nguyen Duy Khang

Student ID: B1910652

Class: 2019-2023 (K45)

Student: Tran Ba Phuong

Student ID: B1908407

Class: 2019-2023 (K45)

Advisor: Dr. Tran Cong An

Can Tho, 12/2022

# ACKOWLEDGEMENT

To complete this report, first, we would like to thank the teachers and teachers of the Faculty of Information Technology, School of Technology and Communication, who believe in us, teach us, and equip us with knowledge. valuable in the learning process and help us better understand the research areas to complete the assigned funding topic.

We would like to express our deepest gratitude to Dr. Tran Cong An, who has guided and mentored wholeheartedly in the process of completing the thesis.

Due to the lack of experience in doing the subject as well as the limitations of knowledge, in the basic annual report, it is inevitable that there will be shortcomings. We look forward to receiving your comments, suggestions, and criticisms to improve the basic yearbook.

Wishing your health, happiness, and success on your career path Teaching.

We sincerely thank you!

Can Tho, December 1, 2022

# ABSTRACT

The use of chatbots evolved rapidly in numerous fields in recent years, including Marketing, Supporting Systems, Education, Health Care, Cultural Heritage, and Entertainment. In this paper, we first present a historical overview of the evolution of the international community’s interest in chatbots. Next, we discuss the motivations that drive the use of chatbots, and we clarify chatbots’ usefulness in a variety of areas. Moreover, we highlight the impact of social stereotypes on chatbots design. After clarifying necessary technological concepts, we move on to a chatbot classification based on various criteria, such as the area of knowledge they refer to, the need they serve and others. Furthermore, we present the general architecture of modern chatbots while also mentioning the main platforms for their creation. Our engagement with the subject so far, reassures us of the prospects of chatbots and encourages us to study them in greater extent and depth.

# LIST OF FIGURES

[Figure 1 Websocket protocol illustration 7](#_Toc121337713)

[Figure 2 Illustrate the preprocessing of the dataset step 1 and 2 10](#_Toc121337714)

[Figure 3 Illustrate the preprocessing of the dataset step 3 10](#_Toc121337715)

[Figure 4 Illustrating the process of training 7 epochs on Google Colab 12](#_Toc121337716)

[Figure 5 System design overview 14](#_Toc121337717)

[Figure 6 System files and folders 15](#_Toc121337718)

[Figure 7 Server running illustration 16](#_Toc121337719)

[Figure 8 Interface running illustration 16](#_Toc121337720)

[Figure 9 Website interface illustration 1 17](#_Toc121337721)

[Figure 10 Website interface illustration 2 17](#_Toc121337722)

[Figure 11 Website interface illustration 3 18](#_Toc121337723)

[Figure 12 Illustration of MetaMask extension on chrome 18](#_Toc121337724)

[Figure 13 Illustration of chatbot website interface 19](#_Toc121337725)

# LIST OF TABLES

[Table 1 Work assignment 2](#_Toc121337726)

[Table 2 Example of a dialogue in the Family Guy Dialogue dataset 9](#_Toc121337727)

[Table 3 Parameters of the model 11](#_Toc121337728)

[Table 4 Model evaluation results 13](#_Toc121337729)

TABLE OF CONTENTS

[ACKOWLEDGEMENT i](#_Toc121337730)

[ABSTRACT ii](#_Toc121337731)

[LIST OF FIGURES iii](#_Toc121337732)

[LIST OF TABLES iii](#_Toc121337733)

[CHAPTER 1: INTRODUCTION 1](#_Toc121337734)

[1. The purpose of the study 1](#_Toc121337735)

[2. The problem of the study 1](#_Toc121337736)

[3. Related work – citing the previous work 1](#_Toc121337737)

[4. Research Contents 2](#_Toc121337738)

[5. Outline 2](#_Toc121337739)

[6. Work assignment 2](#_Toc121337740)

[CHAPTER 2: LITERATURE REVIEW 3](#_Toc121337741)

[1. Machine learning 3](#_Toc121337742)

[2. Chatbot 3](#_Toc121337743)

[2.1. Define 3](#_Toc121337744)

[2.2. Use-cases 3](#_Toc121337745)

[2.3. Classification 4](#_Toc121337746)

[2.4. Mechanisms 4](#_Toc121337747)

[2.5. Financial chatbots 4](#_Toc121337748)

[2.6. Security and Privacy issues 4](#_Toc121337749)

[3. Blockchain 5](#_Toc121337750)

[3.1. Define 5](#_Toc121337751)

[3.2. Classification 5](#_Toc121337752)

[4. DialoGPT 6](#_Toc121337753)

[5. System analysis and design 6](#_Toc121337754)

[5.1. Websocket 6](#_Toc121337755)

[5.2. ReactJS 7](#_Toc121337756)

[5.3. ViteJs 8](#_Toc121337757)

[CHAPTER 3: CHATBOT TRAINING 9](#_Toc121337758)

[1. Domain 9](#_Toc121337759)

[2. Dataset 9](#_Toc121337760)

[3. Dataset preprocessing 9](#_Toc121337761)

[4. Model training 10](#_Toc121337762)

[5. Result evaluation 12](#_Toc121337763)

[6. System design overview 13](#_Toc121337764)

[CHAPTER 4: INSTALL SOLUTION 15](#_Toc121337765)

[1. Install the program 15](#_Toc121337766)

[2. Program specification 17](#_Toc121337767)

[CHAPTER 5: TEST REVIEW 20](#_Toc121337768)

[1. Software testing and evaluation 20](#_Toc121337769)

[2. Install 20](#_Toc121337770)

[3. Test 20](#_Toc121337771)

[4. Rate the results 20](#_Toc121337772)

[CHAPTER 6: CONCLUSION 21](#_Toc121337773)

[1. Result achieved 21](#_Toc121337774)

[2. Development 21](#_Toc121337775)

[REFERENCES 22](#_Toc121337776)

# CHAPTER 1: INTRODUCTION

In this chapter, we will present an overview of the research situation of the chatbot topic and the reason why we chose this topic along with the objectives, objects and research scope of the topic. Finally, we list the contents of the study by period and some related studies.

## The purpose of the study

In our country, answering questions from the customer service department via online messages is popular. However, this is also done manually and faces many difficulties such as time consuming and costly to pay staff just to answer the same simple and similar questions. Therefore, there is an urgent need for an intelligent and automatic control system to bring higher efficiency and chatbot is a perfect choice.

## The problem of the study

A Chatbot is a popular platform to enable users to interact with a software or website to gather information or execute actions in an automated fashion. In recent years, chatbots are being used for executing financial transactions, however, there are several security issues, such as secure authentication, data integrity, system availability and transparency, that must be carefully handled for their wide-scale adoption. Recently, the blockchain technology, with several security advantages, has emerged as one of the foundational technologies with the potential to disrupt several application domains, particularly in the financial sector. In this paper, we forward the idea of integrating a chatbot with blockchain technology in the view to improve the security issues and learn new technologies.

## Related work – citing the previous work

Alan Turing in 1950 proposed the Turing Test (“Can machines think?”), and it was at that time that the idea of a chatbot was popularized [1]. The first known chatbot was Eliza, developed in 1966, whose purpose was to act as a psychotherapist returning the user utterances in a question form [2]. It used simple pattern matching [3] and a template-based response mechanism. Its conversational ability was not good, but it was enough to confuse people at a time when they were not used to interacting with computers and give them the impetus to start developing other chatbots [4]. An improvement over ELIZA was a chatbot with a personality named PARRY developed in 1972 [5]. In 1995, the chatbot ALICE was developed which won the Loebner Prize, an annual Turing Test, in years 2000, 2001, and 2004. It was the first computer to gain the rank of the “most human computer” [6]. ALICE relies on a simple pattern-matching algorithm with the underlying intelligence based on the Artificial Intelligence Markup Language (AIML) [7], which makes it possible for developers to define the building blocks of the chatbot knowledge [6]. Chatbots, like SmarterChild [8] in 2001, were developed and became available through messenger applications. The next step was the creation of virtual personal assistants like Apple Siri [9], Microsoft Cortana [10], Amazon Alexa [11], Google Assistant [12] and IBM Watson [13].

## Research Contents

The research content is divided into 5 main phases including:

Stage 1: Theoretical research

The content of theoretical research focuses on understanding the following documents:

* Basic knowledge of machine learning, chatbot, blockchain.
* Research on GODEL.
* Researching methods and processes for developing app websites

Stage 2: Studying the training data: DSTC7 End-to-End Conversation Modeling.

Stage 3: Building a model chatbot.

Stage 4: Experiment and evaluate the results.

Stage 5: Build a chatbot website that integrates payments with blockchain.

## Outline

CHAPTER 1: INTRODUCTION: Introduce and give directions to research the chatbot problem.

CHAPTER 2: LITERATURE REVIEW: An overview of chatbots and blockchain.

CHAPTER 3: SOLUTION DESIGN: overview of the system architecture to build.

CHAPTER 4: INSTALL SOLUTION: Build a chat bot application combined with blockchain on web platform.

CHAPTER 5: TEST REVIEW: Description of test objectives, test scenarios and test results.

CHAPTER 6: CONCLUSION: Provide research results, assessment, and direction.

## Work assignment

|  |  |
| --- | --- |
| **Work** | **Performer** |
| Build source | Tran Ba Phuong |
| Model training | Tran Ba Phuong |
| Code server (back-end) | Nguyen Duy Khang |
| Code ui | Tran Ba Phuong |
| Blockchain integration | Tran Ba Phuong |
| Write a report | Nguyen Duy Khang |

Table 1 Work assignment

# CHAPTER 2: LITERATURE REVIEW

In this chapter, we present the theory of machine learning, chatbots, blockchains, model training datasets, and technologies used to build systems.

## Machine learning

Currently, there are many problems that can be applied with machine learning technology to solve such as human face recognition, natural language processing, intelligent robot development, etc. The essence of machine learning technology is training. Train the computer to become smarter in the direction that people want it by building computational models capable of self-improving parameters based on the data it encounters. To get a machine learning model that can be applied to real-world problems with high accuracy, we need to expose it and process through a lot of data. Doing the above is often referred to as training a machine learning model.

The process of training machine learning models can be divided into two forms: supervised learning and unsupervised learning.

* Supervised learning is a form of model training used in problems where people already know the answers, then using data sets with answers to put them into the model for processing. The parameters of the model will be automatically adjusted so that the output data processing results match the given answer.
* Unsupervised learning is often used in problems where the answer is unknown. A specific example is the classification problem, the model will be trained with a data set with no given answer, the goal of the problem is to classify the objects of that data set into classes based on the object characteristics.

## Chatbot

### Define

Chatbot (or a bot in short) is an application program that can make auditory or textual conversations in real time with users [14]. This is a smart implementation of AI providing a user-friendly conversational experience for users via multiple channels. It is the upcoming leading technology for vast potential for sales, customer service and marketing. In the next section, we explore several aspects of a chatbot.

### Use-cases

Chatbots are increasingly being used as personal assistants for users, enabling people to converse with a chatbot, ask questions and get things done such as call someone, pay bills, set up a meeting and carry on many other activities that a personal assistant is supposed to do. On March 24, 2017, a 4 years old child Roman even saved his mother’s life using Siri, a chatbot from Apple [15], [10]. Other popular such chatbots are Google Assistant [12] and Amazon’s Alexa [11]. Chatbots are also being used at call centres enabling customers to query regarding their products and receive instant replies 24/7.

### Classification

Text-based: A user interacts with a text-based chatbot with texts only. Users will query with texts and get answers with texts also. Such chatbots can be of two types. One is a bot providing fixed options and users need to select an option to interact with. The other is a dynamic chatbot where the bot, on taking random queries from a user, provides a dynamic answer to the user.

Voice-activated: This is the most sophisticated class of chatbots in which users interact with the bot using voice.

### Mechanisms

Here, we provide a simple working mechanism of a chatbot. A chatbot consists of a number of compo- nents. The front-facing component for a text chatbot is the User Interface (UI) using which a user interacts and submits queries or selects options. A voice-activated chatbot utilises the microphone of the corresponding devices to receive in- structions/inputs from the user. An option-based text chatbot is the easiest to develop as it just needs to be equipped to handle a limited number of pre-selected options. Dynamic textual and voice chatbots, on the other hand, need to utilise a number of additional components and advanced algorithms, such as voice translation and Speech To Text Reporter (STTR), to function properly. These chatbots also need to apply other Natural Language Processing mechanisms, such as Part-Of-Speech [16] Tagging and Sentiment Analysis [17] to understand the query and to produce a a suitable output.

### Financial chatbots

A financial chatbot is a specific type of chatbot which is used in financial domains with a wide- range of use-cases, such as allowing users to execute financial transactions, providing financial advises, preventing financial frauds and so on [18], [19]. In the scope of this paper, In the scope of this paper, In the scope of this paper, we only focus on implementing toll when using chatbot by combining blockchain.

### Security and Privacy issues

Because of their wide usages in different applications domains, chatbots often need to handle sensitive data. Therefore, the security and privacy issues are of great importance for chatbots. Here, we highlight a few of such issues, mostly applicable to financial chatbots, such as secure authentication, data confidentiality and integrity, system availability, accountability and transparency [20], [21], [22]. Only authenticated users should be allowed to interact with a chatbot so that they can submit queries/transact for their respective bank account. Data confidentiality and integrity will guarantee that the submitted transaction is accessible by an authorised entity and is secure against any corruption. System availability will ensure uninterrupted access while accountability and transparency of the system will help to increase the trustworthiness of the system. The principal data privacy issues mostly arise from the lack of control and consent over any submitted transaction.

## Blockchain

### Define

Bitcoin is regarded as the first widely-used decentralised digital currency that does not rely on a central entity, such as a central bank, for its creation and circulation [23]. Its main technological breakthrough is due to its underlying mechanism called blockchain, an example of a distributed ledger shared among a group of Peer-to-Peer (P2P) nodes [24]. The ledger is an ordered data structure consisting of many blocks chained together by cryptographic mechanisms. Each block contains some transactions where each transaction enables a user to transact a certain amount of bitcoin to another user/users. Each block refers to its previous block using a cryptographic hash, which refers to its previous block and so on, hence forming a chain and colloquially known as blockchain.

### Classification

Evolving from the Bitcoin blockchain, a new type of blockchain system has emerged, facilitating the deployment and autonomous execution of computer programs, known as smart-contracts, on top of the respective ledger [12]. Being part of the ledger makes smart-contracts and their executions immutable and irreversible, a sought-after property having a wide range of applications in different domains. Besides, a smart-contract supporting blockchain system has some other advantages, such as distributed data control, data persistence, data provenance, accountability, and transparency. Based on who can access a ledger in a blockchain system, there are generally two types of blockchain:

* Public blockchain: In a public blockchain, also known as the permissionless blockchain, anyone can join and participate in the network for blockchain governance and transaction creation at any time. Examples of public blockchain systems are Bitcoin [25], Ethereum [26], Litecoin [27], Monero [28] and so on.
* Private blockchain: In a private blockchain, also known as permissioned blockchain, only authorised and trusted entities are allowed to participate supporting different levels of permissions and privacy. Examples of private blockchain systems are Hyperledger Platforms [29], Quo-ram [30], and others.

## DialoGPT

DialoGPT [31] contains the source code and trained model for a large-scale pretrained dialogue response generation model. The [human evaluation results](https://github.com/microsoft/DialoGPT?fbclid=IwAR2bIgT_XpJ7FBzQVu1p_YVyDAuNWXytwF2pHlW6laZv5Wup8K1xwD-cfJM" \l "human_eval) indicate that the response generated from DialoGPT is comparable to human response quality under a single-turn conversation Turing test.

The repository is based on [huggingface pytorch-transformer](https://github.com/huggingface/transfer-learning-conv-ai) and [OpenAI GPT-2](https://github.com/openai/gpt-2), containing data extraction script, model training code and pretrained small (117M) medium (345M) and large (762M) model checkpoint.

The model is trained on 147M multi-turn dialogue from Reddit discussion thread. The largest model can be trained in several hours on a 8 V100 machines (however this is not required), with distributed training and FP16 option.

The include script can be used to reproduce the results of DSTC-7 grounded dialogue generation challenge and a 6k multi-reference dataset created from Reddit data.

## System analysis and design

### Websocket

WebSocket API [32] is an advanced technology that allows open two-way interactive communication session between the user's browser and the server. With this API we can send messages to the server and get event driven responses without having to poll the server for an answer.

Diagram

Description automatically generated

Figure 1 Websocket protocol illustration

Real-time web application uses WebSocket to display data at the client end, which is continuously sent by the server. In WebSocket, data is continuously pushed/passed into the same connection that was opened and that's also why WebSocket is faster and improves performance.

### ReactJS

To develop web-based interfaces easily, the topic chooses the ReactJS [33] library because of its popularity and convenience. It can be said that Meta's success in creating ReactJS has helped a lot in developing web interfaces. ReactJS is an open-source JavaScript library created to improve the smoothness and speed of web pages when users interact. Besides, ReactJS also brings high extensibility and is extremely simple. This library helps developers work on individual parts of the site, making teamwork easier than ever.

To increase efficiency in teamwork, the topic decided to integrate TypeScript language into ReactJS. TypeScript is an open-source project developed by Microsoft; it can be seen as an upgraded version of JavaScript. TypeScript is added with many new features and is supported by programming applications that make it easier for the institute to detect and correct errors, and intelligently suggested code snippets also help speed up application development.

### ViteJs

Vite [34] (French word for “quick”) is a build tool that aims to provide a faster and leaner development experience for modern web projects. It consists of two major parts:

* A dev server that provides rich feature enhancements over native ES modules, for example extremely fast Hot Module Replacement (HMR)
* A build command that bundles your code with [Rollup](https://rollupjs.org" \t "_blank), pre-configured to output highly optimized static assets for production.

Vite is opinionated and comes with sensible defaults out of the box, but is also highly extensible via its Plugin API and JavaScript API with full typing support.

# CHAPTER 3: CHATBOT TRAINING

In this chapter, we present the training process of a chatbot model and an overview of the design of a chatbot website system that integrates blockchain payments.

## Domain

Open domain, users can create conversations in any field. It is not necessary to define goals or intentions in advance. Conversations on social networking sites like twitter and reddit are good examples of open domains. They can go through different topics. There are countless topics with a lot of insight required to generate a reasonable answer, which is a difficult problem. We have trained the chatbot in the direction of open domain with the general topic of basic communication sentences. The data set is taken from conversations in a family cartoon, and it fits our requirements quite well.

## Dataset

The dataset used to train the chatbot model is Family Guy Dialogue [35]. This dataset consists of dialogue from the family guy the fun cartoon series and there are three columns in it. First is a character column that shows the name of the character and followed by dialogue which is used by the character on the show and finally seasons of the show. We chose the character with the most dialogue in this dataset as Peter for training, accounting for 29% of the total.

|  |  |  |
| --- | --- | --- |
| **character** | **dialog** | **season** |
| Peter | Brian, teach me how to be a gentleman. | season 1 |
| Brian | Well, Peter, its not really that hard. Lets start with a polite conversation. For example, its a pl... | season 1 |
| Peter | Its a pleasure to see you again. After Hogans Heroes, Bob Crane got his skull crushed in by a friend... | season 1 |
| Brian | Wow. Perfect. My work is done. But just for the heck of it, lets try it again. | season 1 |

Table 2 Example of a dialogue in the Family Guy Dialogue dataset

## Dataset preprocessing

Step 1: remove session field in dataset.

Step 2: rename character field to name and dialog field to line.

1. Graphical user interface, text, application, email

   Description automatically generated

Figure 2 Illustrate the preprocessing of the dataset step 1 and 2

Step 3: create a model-fit dataset consisting of a response field of and contexts.

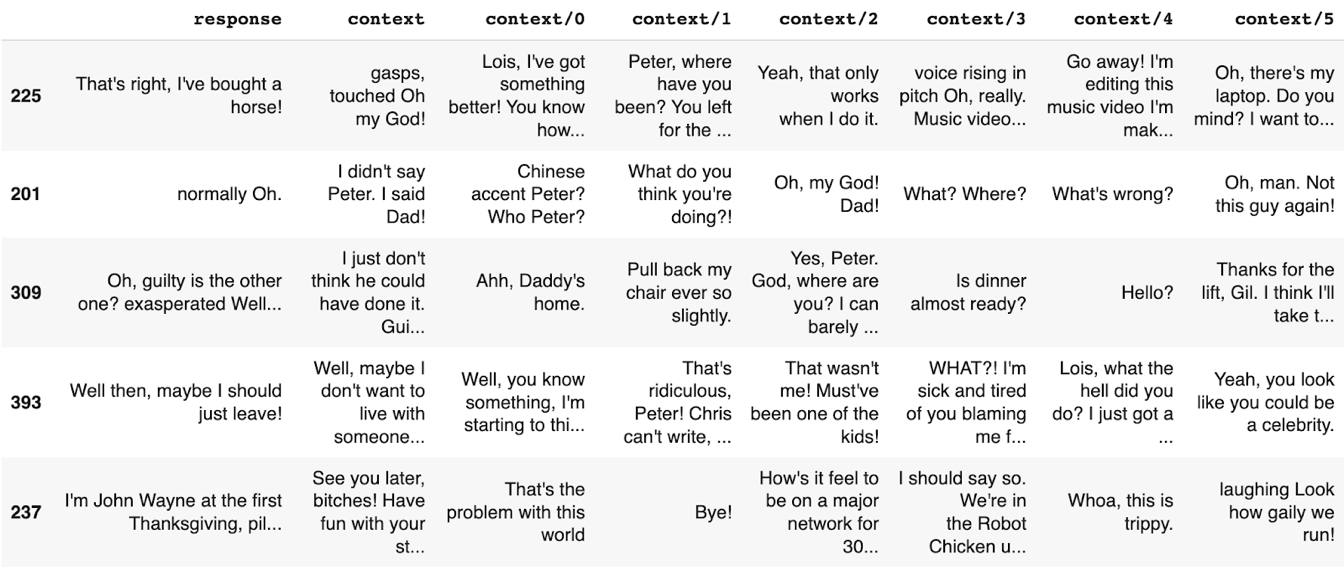
1. 

Figure 3 Illustrate the preprocessing of the dataset step 3

A total of 394 responses will be divided between the training is 354 and 40 for validation.

## Model training

We have used Microsoft/DialoGPT-small Model [31] for this ChatBot. The training script accept several arguments to tweak the training presented in Table 3.

|  |  |  |  |
| --- | --- | --- | --- |
| **Argument** | **Type** | **Default value** | **Description** |
| fp16 | boolean | f | Whether to use 16-bits floating point for model training. |
| gradient\_accumulation\_steps | int | 1 | Accumulate gradients on several steps |
| learning\_rate | float | 5e-5 | Learning rate |
| continue\_from |  |  | Resuming the training after a specified number of steps |
| num\_train\_epochs | int | 5, 7, 12 | Hyper parameters that determine how many times we put all the data from the training dataset in |
| per\_gpu\_train\_batch\_size | int | 4 | Batch size for training |
| per\_gpu\_eval\_batch\_size | int | 4 | Batch size for validation |
| gradient\_accumulation\_steps | int | 1 | Accumulate gradients on several steps |
| max\_seq\_length | int | 200 | Maximum number of tokens for each training instance. |

Table 3 Parameters of the model

To train the model, we use Google Colab connected to a Google Drive account containing the source code we have uploaded. Then change the path to the source code directory. Google Colab provides the ability to execute a python file using the “! python file\_name.py" command. All the printed messages will be displayed below the execution line, from which we can observe and check the progress of the training. To make the training process as fast as possible, we will change the hardware of Google Colab from CPU to GPU.

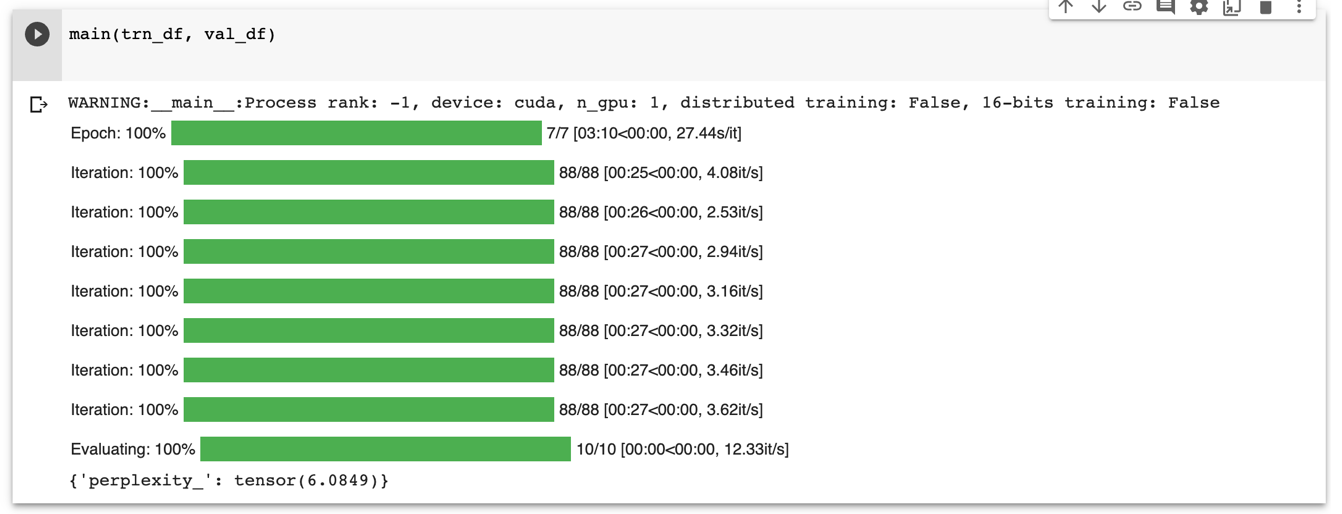


Figure 4 Illustrating the process of training 7 epochs on Google Colab

## Result evaluation

To evaluate the model, we chose one of the popular methods, perplexity [36], instead of BLEU score because because the reasonableness of the answer may contain different words or phrases.

Suppose we have a set of sentences to test (held-out: not in the training set).

|  |  |
| --- | --- |
|  | (1) |

Every sentence have is a sequence of words , in there is the length of the second sentence .

We calculate the probability for each through the language model that has just been trained. Then, the quality of this language model will be calculated as formula (2).

|  |  |
| --- | --- |
|  | (2) |

The higher the value obtained from the above calculation, the better the quality of the model for the data that is not seen in the training set.

Perplexity is defined as formula (3).

|  |  |
| --- | --- |
|  | (3) |

In there:

|  |  |
| --- | --- |
|  | (4) |

According to the above formula (4), the smaller the value of perplexity, the better the language model can be built.

|  |  |  |
| --- | --- | --- |
| **Number of epoch** | **Perplexity** | **Test dialogue** |
| 5 epochs | 6.7179 | Text  Description automatically generated |
| 7 epochs | 6.0849 | Text  Description automatically generated |
| 12 epochs | 3.1306 | Text  Description automatically generated |

Table Model evaluation results

As evaluated in Table 4, the model after training 12 epochs gives lower perplexity results than the model training 7 or 5 epochs. We also tried to train with a higher number of epochs, but the results were not significantly different from the 12 epochs training model, so we decided to choose this model for the system to use.

The results show that the accuracy of the model when answering questions is not too high and the scope is not wide because the data set is quite small compared to the requirements of an open domain chatbot.

## System design overview

After having the chatbot model, we proceed to build a model website system that is placed on the server side and we use MetaMask wallet to authenticate payments through Ethereum.

1. Diagram

   Description automatically generated

Figure 5 System design overview

# CHAPTER 4: INSTALL SOLUTION

In this chapter, we show how to install and use the chatbot website system with integrated blockchain payment. In addition, we also present the program specification.

## Install the program

We have two main folders in the project, ui (contains the website's interface) and server (contains the website's server).

Graphical user interface, application

Description automatically generated

Figure 6 System files and folders

To be able to run the website we will start both parts at the same time

* Server part:

$ cd server: open the server at the command line

$ pip install requirements.txt: install all necessary packages

$ uvicorn main:app –reload: Start the server at port 8000

After successfully running the command line interface will notify Uvicorn running on <http://127.0.0.1:8000> and Application startup complete.

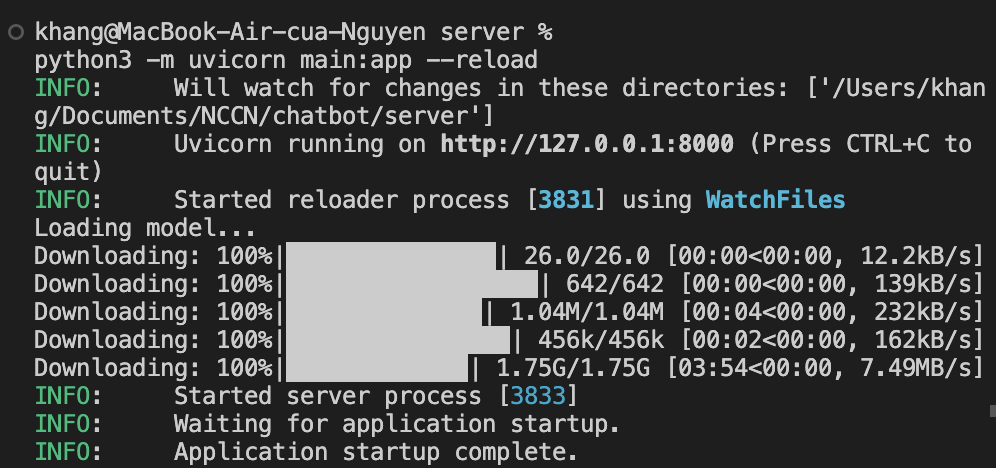


Figure 7 Server running illustration

* UI part

$ cd ui: open ui at command line

$ yarn: install all necessary libraries in package.json

$ yarn vite: Start the interface Vite

After successfully running the command line interface will notify Vite running on <http://localhost:5173/>

Text

Description automatically generated

Figure 8 Interface running illustration

## Program specification

The home page interface includes information about discoverable benefits. To access the chatbot site the user must pay a certain fee through the Metamask wallet (blockchain) the currency used is GoerliETH.

1. Graphical user interface

   Description automatically generated with medium confidence
2. Figure 9 Website interface illustration 1
3. Graphical user interface, application

   Description automatically generated
4. Figure 10 Website interface illustration 2
5. Graphical user interface, application

   Description automatically generated
6. Figure 11 Website interface illustration 3
7. Graphical user interface, application

   Description automatically generated
8. Figure 12 Illustration of MetaMask extension on chrome

After the user completes the payment, you can go to the page to chat with the bot.

1. Graphical user interface, text, application, chat or text message

   Description automatically generated

Figure 13 Illustration of chatbot website interface

# CHAPTER 5: TEST REVIEW

Testing is crucial because it identifies flaws and errors before the website is delivered to the client, ensuring the website quality. Besides, Blockchain testing is the methodical assessment of the functioning elements of the blockchain, it is our smart contracts. Blockchain testing, in contrast to standard software testing, contains several components, including blocks, mining, transactions, wallets, and so forth, all of which need specialized tools to be tested.

## Software testing and evaluation

The blockchain-powered chatbot website is basically done and testing done at the local level

## Install

Server Requirement: An Internet connection is required for the server to load the model.

## Test

Test your website on a local level.

## Rate the results

Basically, the Website has met the requirements of a chatbot combined with blockchain website.

# CHAPTER 6: CONCLUSION

In this chapter, we summarize the results we have achieved in the process of researching this topic. At the same time, we also make some suggestions to develop and improve this research in the future.

## Result achieved

After the implementation of the project, the program has completed the following results:

* Learn and understand chatbot knowledge
* Learn and understand blockchain knowledge
* Understand how chatbots integrate blockchain
* Learn and understand the tools to build websites such as: ReactJS, Vite
* Perform website testing at local level.

## Development

Replace DialoGPT with what GODEL has higher accuracy.

Maybe converting the chatbot to sounding like a virtual assistant would be more relevant for the time being.

Convert the system to a plugin so that it can be integrated into other systems.

# REFERENCES

|  |  |
| --- | --- |
| [1] | Turing, Alan, Computing machinery and intelligence in “Mind”, vol, 433-460, 1950. |
| [2] | Weizenbaum, Joseph, ELIZA—a computer program for the study of natural language communication between man and machine, Communications of the ACM, 9(1), 36-45, 1966. |
| [3] | Brandtzaeg, Petter Bae, and Asbjørn Følstad, Why people use chatbots, International conference on internet science (pp. 377-392), 2017. |
| [4] | Klopfenstein, L. C., Delpriori, S., Malatini, S., & Bogliolo, A, The rise of bots: A survey of conversational interfaces, patterns, and paradigms, Proceedings of the 2017 conference on designing interactive systems (pp. 555-565), 2017, June. |
| [5] | Colby, Kenneth Mark, Sylvia Weber, and Franklin Dennis Hilf, Artificial paranoia, Artificial Intelligence, 2(1), 1-25, 1971. |
| [6] | Wallace, Richard S, The anatomy of ALICE, Parsing the turing test (pp. 181-210), 2009. |
| [7] | Marietto, M. D. G. B., de Aguiar, R. V., Barbosa, G. D. O., Botelho, W. T., Pimentel, E., França, R. D. S., & da Silva, V. L, Artificial intelligence markup language: a brief tutorial, arXiv preprint arXiv:1307.3091, 2013. |
| [8] | Molnár, György, and Zoltán Szüts, The role of chatbots in formal education, 2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY) (pp. 000197-000202), 2018. |
| [9] | [Online]. Available: https://www.apple.com/siri/. |
| [10] | [Online]. Available: https://www.microsoft.com/en-us/cortana. |
| [11] | [Online]. Available: https://www.digitaltrends.com/home/what-is-amazons-alexa-and-what-can-it-do/. |
| [12] | [Online]. Available: https://assistant.google.com. |
| [13] | [Online]. Available: https://www.ibm.com/watson. |
| [14] | Abdul-Kader, Sameera A., and John C. Woods, Survey on chatbot design techniques in speech conversation systems, International Journal of Advanced Computer Science and Applications 6.7 , 2015. |
| [15] | [Online]. Available: https://www.cnet.com/culture/child-saves-mother-iphone-siri-uk/. |
| [16] | Brill, Eric, "Transformation-based error-driven learning and natural language processing: A case study in part-of-speech tagging," in *Computational linguistics, 21(4), 543-565*, 1995. |
| [17] | Bakshi, R. K., Kaur, N., Kaur, R., & Kaur, G, "Opinion mining and sentiment analysis," in *2016 3rd international conference on computing for sustainable global development (INDIACom) (pp. 452-455)*, 2016. |
| [18] | Okuda, Takuma, and Sanae Shoda, "AI-based chatbot service for financial industry," in *Fujitsu Scientific and Technical Journal, 54(2), 4-8*, 2018. |
| [19] | [Online]. Available: https://www.artificial-solutions.com/blog/chatbots-financial-services-benefits-use-cases. |
| [20] | Bozic, Josip, and Franz Wotawa, "Security testing for chatbots," in *IFIP International Conference on Testing Software and Systems (pp. 33-38)*, 2018. |
| [21] | Lai, Sen-Tarng, Fang-Yie Leu, and Jeng-Wei Lin, "A banking chatbot security control procedure for protecting user data security and privacy," in *International conference on broadband and wireless computing, communication and applications*, 2018. |
| [22] | Yan, F., Xu, M., Qiao, T., Wu, T., Yang, X., Zheng, N., & Choo, K. K. R, "Identifying wechat red packets and fund transfers via analyzing encrypted network traffic," in *2018 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/12th IEEE International Conference on Big Data Science and Engineering (TrustCom/BigDataSE) (pp. 1426-1432)*, 2018. |
| [23] | Nakamoto, Satoshi, "Bitcoin: A peer-to-peer electronic cash system," in *Decentralized Business Review, 21260*, 2008. |
| [24] | Chowdhury, M. J. M., Ferdous, M. S., Biswas, K., Chowdhury, N., Kayes, A. S. M., Alazab, M., & Watters, P, "A comparative analysis of distributed ledger technology platforms," in *EEE Access, 7, 167930-167943*, 2019. |
| [25] | [Online]. Available: https://www.bitcoin.org. |
| [26] | [Online]. Available: https://www.ethereum.org. |
| [27] | [Online]. Available: https://litecoin.org. |
| [28] | [Online]. Available: https://www.getmonero.org. |
| [29] | [Online]. Available: https://www.hyperledger.org. |
| [30] | [Online]. Available: https://consensys.net/quorum/. |
| [31] | Zhang, Y., Sun, S., Galley, M., Chen, Y. C., Brockett, C., Gao, X., ... & Dolan, B, "Dialogpt: Large-scale generative pre-training for conversational response generation," in *arXiv preprint arXiv:1911.00536*, 2019. |
| [32] | [Online]. Available: https://tenten.vn/tin-tuc/websocket/. |
| [33] | [Online]. Available: https://reactjs.org. |
| [34] | [Online]. Available: https://vitejs.dev. |
| [35] | [Online]. Available: https://www.kaggle.com/datasets/gokulrajkmv/family-guy-dialog. |
| [36] | [Online]. Available: http://www.cs.columbia.edu/~mcollins/courses/nlp2011/notes/lm.pdf. |
| [37] | Chopra, S., Auli, M., & Rush, A. M., "Abstractive sentence summarization with attentive recurrent neural networks," in *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, San Diego, California, 2016. |
| [38] | Peng, B., Galley, M., He, P., Brockett, C., Liden, L., Nouri, E., ... & Gao, J, "Godel: Large-scale pre-training for goal-directed dialog," in *arXiv preprint arXiv:2206.11309*, 2022. |
| [39] | Galley, M., Brockett, C., Gao, X., Dolan, B., & Gao, J, "End-to-end conversation modeling: Dstc7 task 2 description," in *DSTC7 workshop (forthcoming)*, 2018. |
| [40] | Khashabi, Daniel, Yeganeh Kordi, and Hannaneh Hajishirzi, "Unifiedqa-v2: Stronger generalization via broader cross-format training," in *arXiv preprint arXiv:2202.12359*, 2022. |
| [41] | Bajaj, P., Campos, D., Craswell, N., Deng, L., Gao, J., Liu, X., ... & Wang, T, "Ms marco: A human generated machine reading comprehension dataset," in *arXiv preprint arXiv:1611.09268*, 2016. |
| [42] | Mosig, Johannes EM, Shikib Mehri, and Thomas Kober, "Star: A schema-guided dialog dataset for transfer learning," in *arXiv preprint arXiv:2010.11853*, 2020. |
| [43] | [Online]. Available: https://files.pushshift.io/reddit. |