

PROJECT REPORT

**CHATBOT AND MACHINE LEARNING FOR DATA CLASSIFICATION
FOR JUZZ MARKETING**

**MASTER OF TECHNOLOGY IN INTELLIGENT SYSTEM
MODULE: INTELLIGENT REASONING SYSTEM**

**BY: LAU XIANGMING
STUDENT ID: A0267144A**

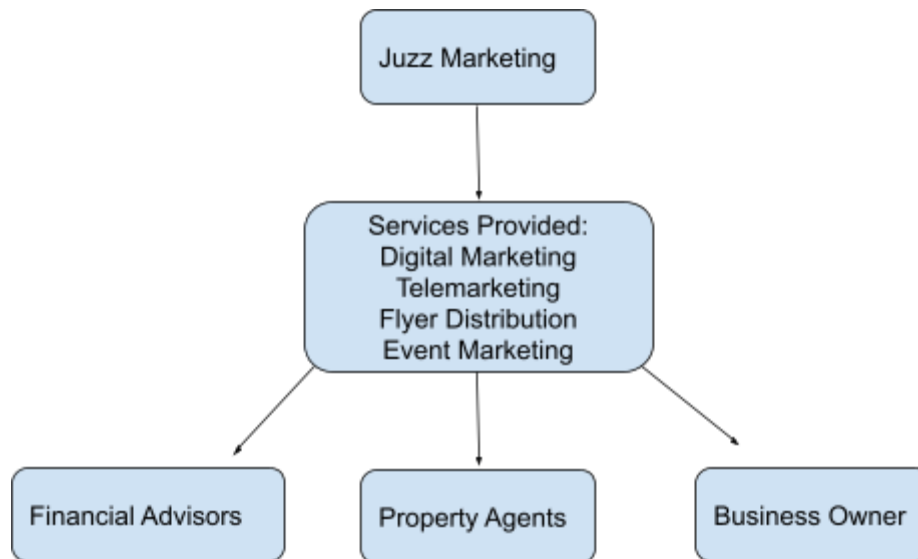
Table of Contents

Executive Summary	<u>3</u>
Introduction	<u>5</u>
Project Planning & Design	<u>6</u>
Picture Flow Diagram	<u>6</u>
WhatsApp Chatbot	<u>6</u>
Sales Enquiry Handling	<u>7</u>
Complaint Handling	<u>7</u>
Database Classification	<u>7</u>
Overall Project Scope	<u>8</u>
Project Technicality	<u>8</u>
Juzz Marketing Website Link Scraping	<u>8</u>
Question Database	<u>9</u>
DialogueFlow	<u>9</u>
User Category	<u>9</u>
Testing of the Jazz Chatbot Workflow	<u>16</u>
Classify Database's Gender	<u>21</u>
Limitations of Jazz WhatsApp Chatbot	<u>22</u>
Limitations of Gender Classification	<u>22</u>
Future Works	<u>22</u>
Rerefences	<u>23</u>
Appendices	<u>24</u>

Executive Summary

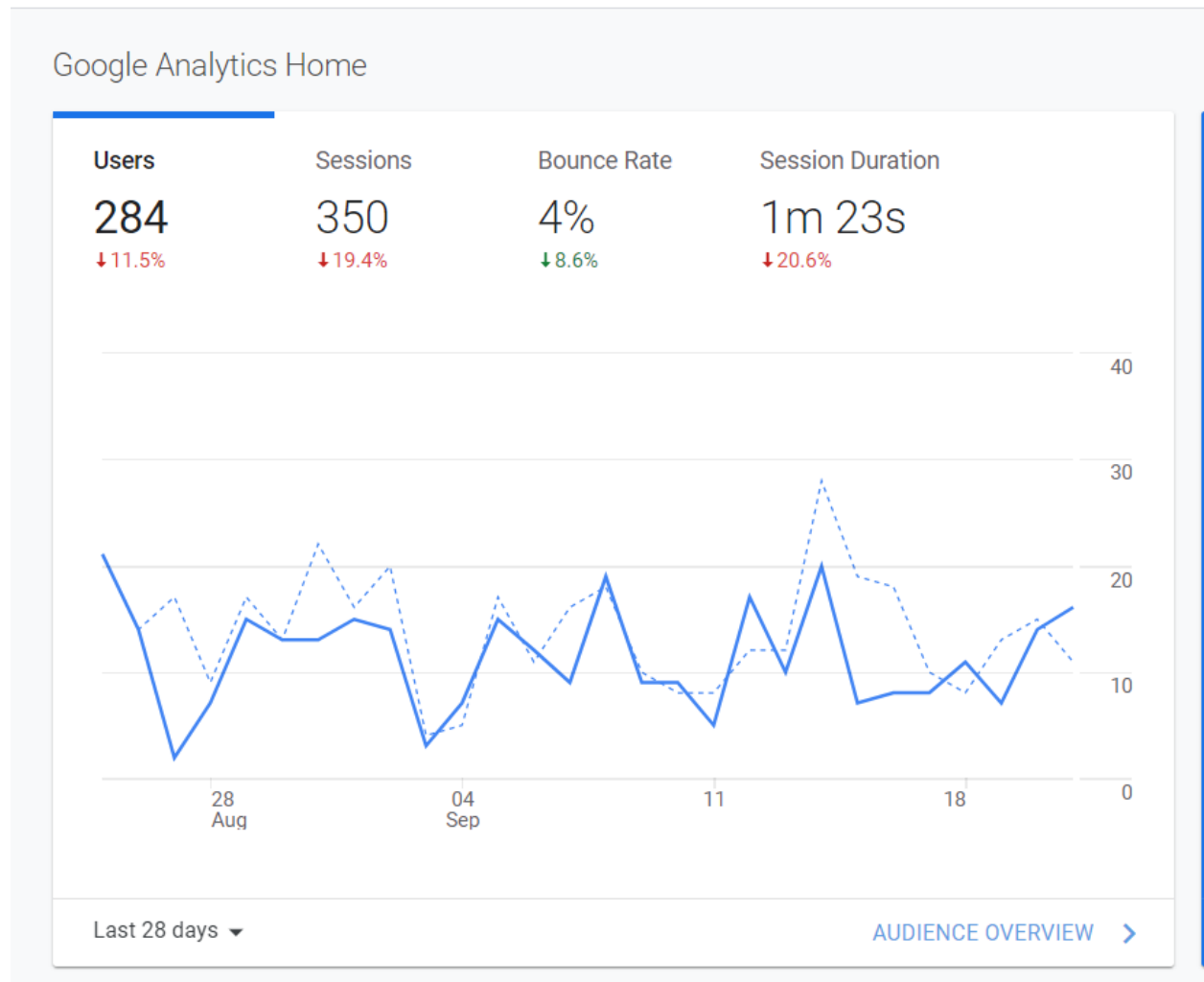
Juzz Marketing is a marketing agency that provides lead generation services to business owners, property agents and financial advisors. The website has multiple web pages that provide a lot of information about their lead generation services. However when a user visits the website, they could get confused and have no idea what to look for. Based on the google analytics as shown in picture 1, the website has a 4% bounce rate which means that for every 25 visitors who visited the website, 1 of them will exit the website after viewing 1 page. Whenever a user tries to whats app Juzz for enquiry, it will take hours for them to reply due to manpower shortage. Juzz could be potentially losing customers due to this issue and we have to find possible solutions to fix this problem. Juzz also do not have a sitemap to assist user in the navigation.

Juzz also has a huge consumer database that consists of only name and mobile generated for its clients and wish to seek more meaningful information from this such as finding out their gender using artificial intelligence instead of manually input. There is a need to do gender classification because there may be some clients who like to target specific gender. Take for example, a facial spa that provides their services only to females may want to target all female customers only



The services that Jazz Marketing provides are Digital Marketing, Telemarketing, Flyers Distribution and Event Marketing and its customers can be classified into 3 main categories - Financial advisors, Property Agents and Business Owners.

Picture 1(screenshot on 22 Sep 2022)



Juzz Chatbot is created to help solve this problem. When someone visits the website, Juzz Chatbot will interact with them and start to ask them questions. Google Dialog flow will be used to integrate with whats app for the chatbot design.

Introduction

Objective of This Project

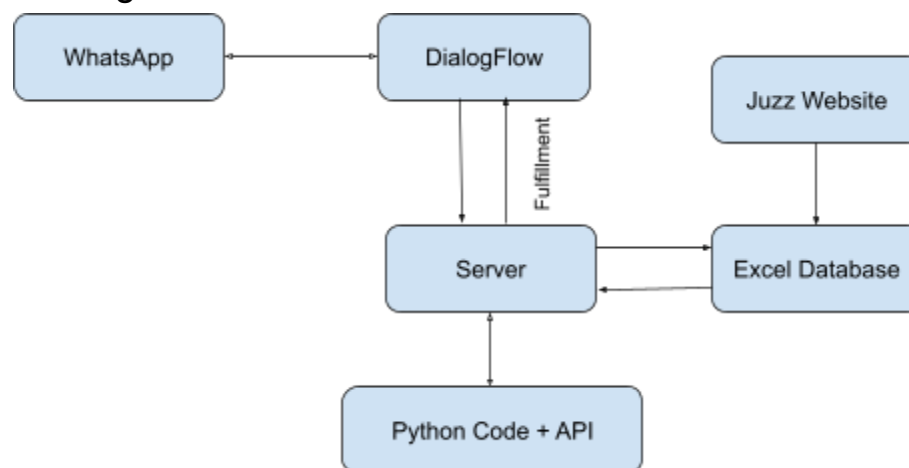
The objective of this project is to design a chatbot in English using Google Dialog Flow for Juzz Marketing to automate its whats app messages and

to use machine learning to help Juzz Marketing automate the gender classification of its consumer database that consists of only name and mobile which is currently done manually. The consumer database is generated from its lifestyle portal <https://upgradelifestyle.fun/> Juzz Marketing website URL is <https://juzzmarketingservices.com/> and it uses <https://upgradelifestyle.fun/> a lifestyle portal to generate leads for its clients. Juzz clients consist of business owners, property agents and financial advisors. Its consumer database will consist of members of the public from all walks of life who opt in to the lifestyle portal.

Project Planning & Design

This project aims to solve 2 problems. Firstly, to automate whats app replies by creating a chatbot using Google DialogFlow. Secondly, to apply machine learning techniques to help Juzz classify its database's gender into male and female that has only name and mobile.

Picture Flow Diagram



WhatsApp Chatbot

The project planning starts from a thought. “How can I apply what I have learned in NUS Intelligent Reasoning System modules to help Juzz Marketing solve its existing problems and use Artificial Intelligence to automate its workflow?”

Sales Enquiry Handling

Juzz Marketing sales enquiries usually come in from whats app and is currently manually replied by their internal staff. Due to their heavy workload, the whats app messages are usually replied only after a couple of hours or in the worst case scenario , the next working day. Juzz Chatbot is designed to help solve this problem- To help automate the whats app chats so that the messages are replied within seconds instead of hours.

Complaint Handling

Just like any lead generation companies, Juzz marketing is unable to guarantee that all the leads they generate are of good quality because when a lead opts in, it is out of their control whether the lead will become a potential customer. From time to time, Juzz Marketing will also receive complaints from customers. As humans contain emotions and cannot work 24/7, these complaints are not handled promptly and if Juzz employs a customer service officer who is hot tempered and starts a quarrel with clients, it can result in a loss of clients. As robots do not contain emotions and can work 24/7, it makes every sense for Juzz to automate this process and make complaint handling more effective.

Database Classification

The database that Juzz Marketing generates for its client generally consists of only name and number only. It does not consist of other details such as gender. A staff will look at the name of the prospects and try to gauge whether it is a male or female and manually key into the database and this takes up a lot of time. Machine Learning is applied to help Juzz automate this process to improve efficiency.

Overall Project Scope

1. The queries addressed by the chatbot will not be limited to just information accessible on Juzz website as it will consist of complaint handling standard operating procedures(SOP) which will not be appropriate to put inside the website.
2. The chatbot would be integrated to whats app
3. The database gender classification will be analyzed and classified using Natural Language Processing(NLP) techniques.

Project Technicality

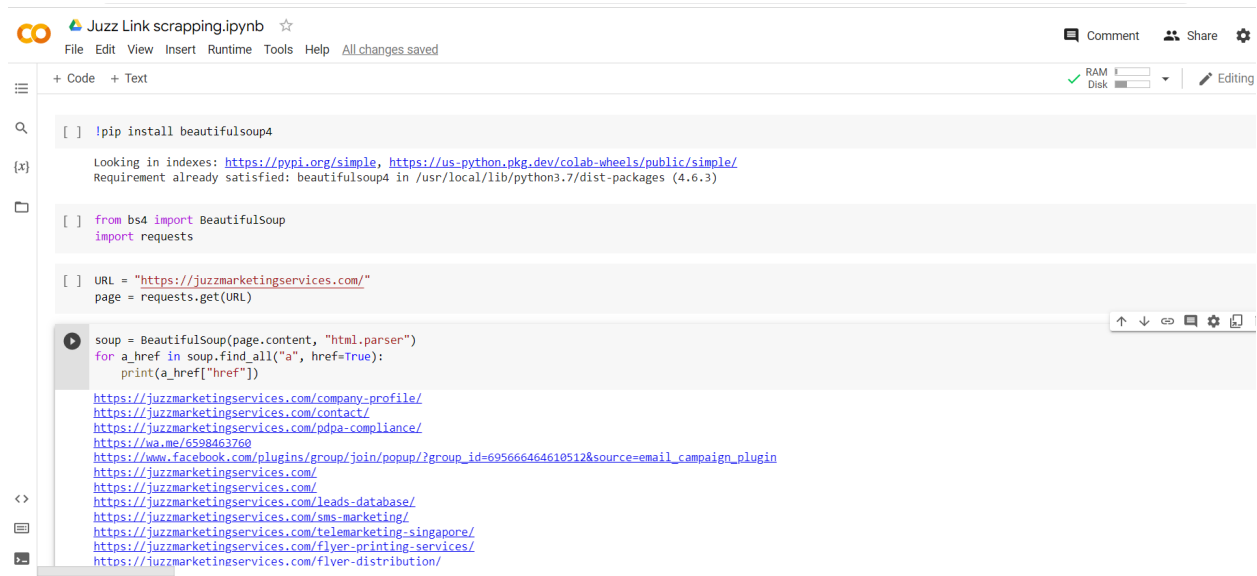
1. Beautiful soup is used to scrap the hyperlinks of Juzz Marketing in order to derive at the potential questions users will ask in whats app. A total of 32 hyperlinks has been scrapped.
2. The scrapped information is used to build into potential questionnaires that potential users will ask
3. Google Dialog Flow NLP is used to process user intents
4. Machine Learning using Python is use to classify database
5. Chat History with 50 clients has been studied and recorded down for the design of Juzz Chatbot so that we can consider different phrases and utterances to train the intents of Google Dialog Flow.

Juzz Marketing Website Link Scraping

Beautiful Soup is a Python package for parsing HTML and XML documents. Beautiful library has been imported into python to scrub the hyperlinks of Juzz Marketing website to ensure no links are missed out. The reason why beautiful soup is chosen for scraping of the website is because it is fast and handy to use and is suitable for smaller projects. The scrapper looked through the main website of Juzz Marketing

www.juzzmarketingservices.com and search for hyperlinks within. The process is reflected in picture 2 below.

Picture 2



```
[ ] | pip install beautifulsoup4

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: beautifulsoup4 in /usr/local/lib/python3.7/dist-packages (4.6.3)

[ ] | from bs4 import BeautifulSoup
import requests

[ ] | URL = "https://juzzmarketingservices.com/"
page = requests.get(URL)

[ ] | soup = BeautifulSoup(page.content, "html.parser")
for a_href in soup.find_all("a", href=True):
    print(a_href["href"])

https://juzzmarketingservices.com/company-profile/
https://juzzmarketingservices.com/contact/
https://juzzmarketingservices.com/odpa-compliance/
https://wa.me/6598463760
https://www.facebook.com/plugins/group/join/popup/?group_id=695666464610512&source=email_campaign_plugin
https://juzzmarketingservices.com/
https://juzzmarketingservices.com/
https://juzzmarketingservices.com/leads-database/
https://juzzmarketingservices.com/sms-marketing/
https://juzzmarketingservices.com/telemarketing-singapore/
https://juzzmarketingservices.com/flyer-printing-services/
https://juzzmarketingservices.com/flyer-distribution/
```

Question Database

After deriving the hyperlinks, a list of questions was then drafted to identity possible questions that users will potentially ask Juzz Chatbot. All the content of Juzz Marketing has been read through to understand what kind of information and data is available and for which of the 3 user types(Business Owners,Property agents or financial advisors) is the information meant for.

DialogueFlow

User category

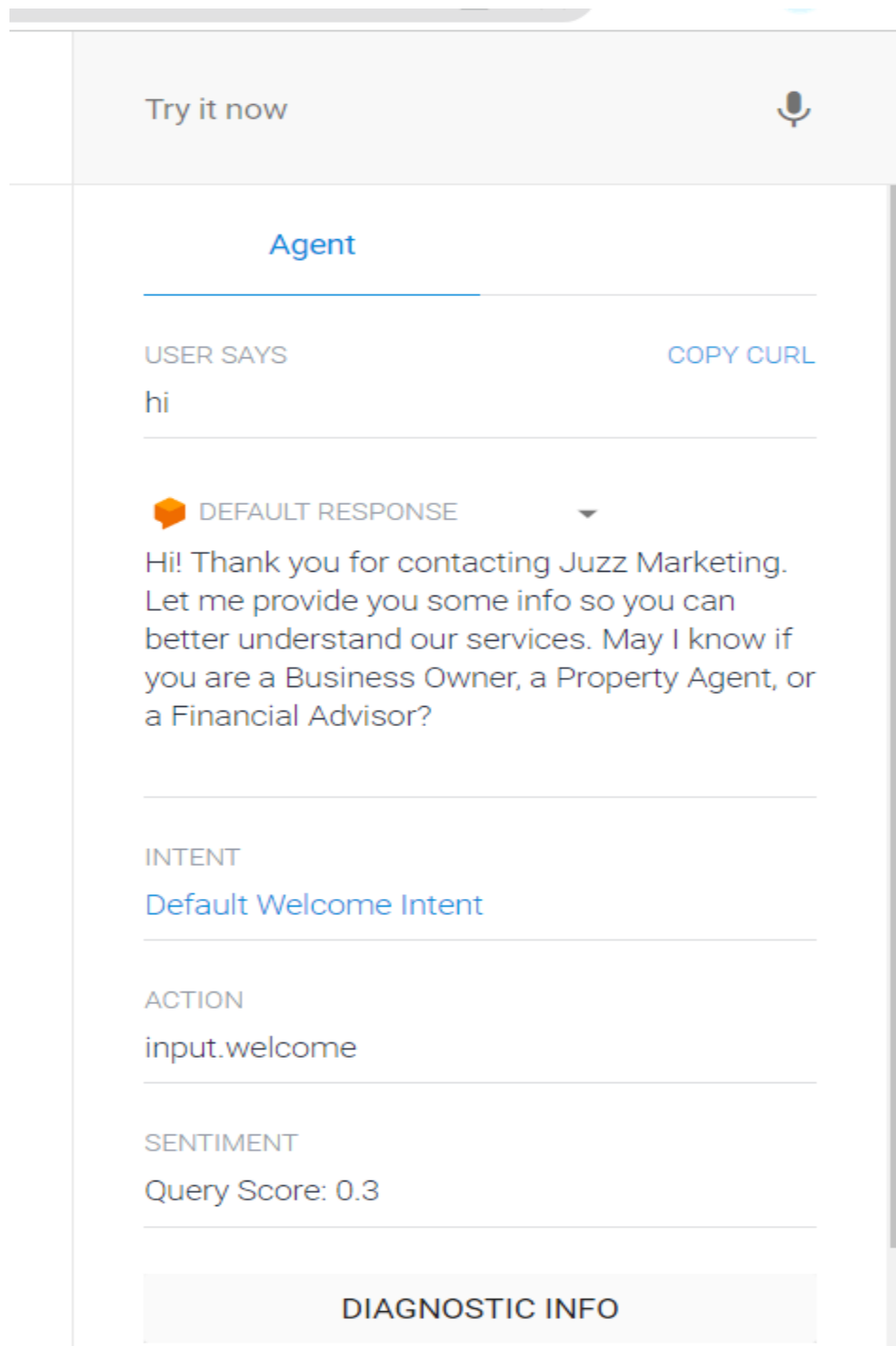
The user input(utterances) are processed by Google Dialog Flow and matched against the intent. “ Default Welcome Intent” and “Default Fallback Intent” are designed to handle the greeting utterance. The chatbot needs to

be able to identify the user type. The user types are basically classified into 3 main categories. They are businesses, property agent and financial advisors

- Businesses- This refers to all range of businesses from F&B to retail and to e-commerce except for property and insurance.
- Property agents- There around 30,000 property agents in Singapore who conduct selling and renting of both commercial and residential properties
- Financial advisors- They provide financial planning in terms of wealth protection and wealth accumulation to their clients

The welcome intent is shown in picture 1 below. When someone sends a greeting utterance such as “Hi”. Google Dialog flow will ask them whether are they business owners, property agents or financial advisors


Picture 3



After identifying the User type, Juzz Chatbot will reply accordingly. Take for example, in Picture 4 below, the user replied “ i am a financial advisor” Juzz Chatbot will respond and provide the services and rates to the user.

Picture 4


Try it now



USER SAYS

COPY CURL

i am a financial advisor

 DEFAULT RESPONSE


▼

Hi here is some info for you : Package (1) 500 targeted leads \$1000 (suitable for agents who are comfortable doing cold calling) Package (2) 50 appt settings(\$3000) + 50 opt in leads (\$1500) = \$4500 -Appt settings (We set appts for you with name, age, occupation, date, time, and location of appt for you) -Opt in leads (obtained through digital marketing by offering a free gift+wealth sharing session/financial review) Guarantee min 3 cases closed Target 10 cases closed If the unlikely event if you don't close at least 3 cases, we will send you more leads until you close 3 cases Because we guarantee a min of 3 cases closed while you target 10 cases closed, therefore it's totally risk-free for you while you enjoy so much upside potential. There is also a 20min video on our website at the bottom of the page that we shared about our services and there are FAQs too.

<https://juzzmarketingservices.com/leads-for-financial-advisors> Let us know if you have additional enquiries. Thank you!

If the user is a property agent, Jazz chatbot will reply accordingly. Take for example, in Picture 5 below, the user replied “property agent” Jazz Chatbot will respond and provide the services and rates to the user.

Picture 5


Try it now

Agent

USER SAYS

COPY CURL

property agent

 DEFAULT RESPONSE ▾

We provide the following services for property agents: 1) Flyer Printing and Distribution 2) Homeowner Leads 3) Appointment Settings 4) Digital Marketing & Management. Which service do you require?

INTENT

Property-agent-Intent

ACTION

Not available

SENTIMENT


Query Score: 0.1

DIAGNOSTIC INFO

When the user identifies himself as a property agent and keen on flyer distribution services, jazz Chatbot will reply with flyer distribution services details and rates to the user.

Picture 6

Try it now




Agent

USER SAYS

COPY CURL

flyer distribution

 DEFAULT RESPONSE

Here are the types of flyer distribution available. 1) HDB Letterbox \$39/1000 units 2) HDB Door-to-door \$59/1000 units 3) Landed/Industrial/Commercial/Shophouse Letterbox \$99/1000units 4) Carpark \$60/1000 units 5) Street/MRT/Bus Stop \$25/hr/150 flyers. This rate is for A4 size (157gsm) flyer & smaller with 5k minimum order quantity.

INTENT

flyer-distribution-intent

ACTION

Not available

SENTIMENT

Query Score: 0.0

A default intent has also been created to handle questions that users will ask but does not fall inside the intent already designed yet as shown in the picture below. Users will always ask question out of the box such as “Who are you” or some people could message us wrongly such as ‘ Where are you now?” Therefore a fallback intent is very important.

Try it now

Agent

USER SAYS

who are you

COPY CURL

DEFAULT RESPONSE

Sorry i didn't get that. I will ask our client servicing manager to contact you

CONTEXTS

RESET CONTEXTS

__system_counters__

INTENT

Default Fallback Intent

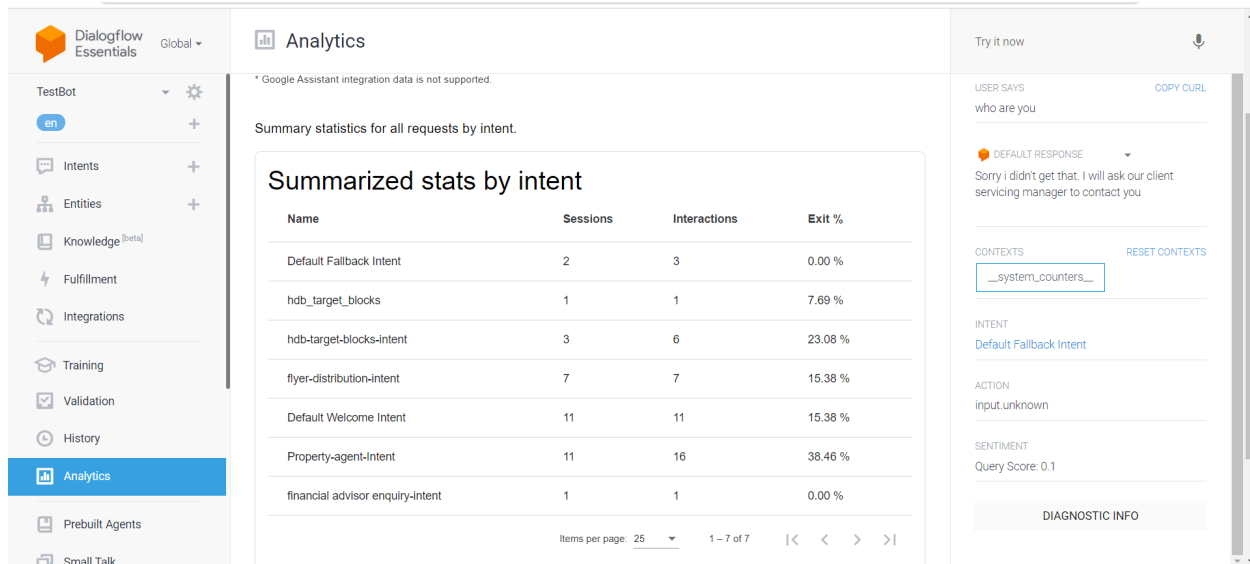
ACTION

input.unknown

SENTIMENT

Query Score: 0.1

Analytic will be done on a regular basis as shown in the picture below to study the fallback intent to improve Juzz chatbot. The goal is to have as little fallback intent as possible and let Juzz Chatbot be automated.



Testing of the Juzz Chatbot Workflow

User: Hi hello

Juzz Chatbot: Hi! Thank you for contacting Juzz Marketing. Let me provide you with some information so you can better understand our services. May I know if you are a Business Owner, a Property Agent, or a Financial Advisor?

User; Financial Advisor

Juzz Chatbot: Hi, here is some information for you:

Package 1: 500 targeted leads \$1000 (suitable for agents who are comfortable doing cold calling)

Package 2: 50 appt settings(\$3000) + 50 opt in leads (\$1500) = \$4500 (Appt settings; We set appts for you with name, age, occupation, date, time, and location of appt for you + Opt in leads; obtained through digital marketing by offering a free gift+wealth sharing session/financial review)

Both packages come with Guarantee min 3 cases closed. Target 10 cases closed. In the unlikely event if you don't close at least 3 cases, we will send you more leads until you close 3 cases. Because we guarantee a min of 3

cases closed while you target 10 cases closed, therefore it's totally risk-free for you while you enjoy so much upside potential.

Go to <https://juzzmarketingservices.com/leads-for-financial-advisors/> for additional information. Let us know if you have additional enquiries. Thank you!

User: Can zoom to share more?

Juzz Chatbot: Sure. May i know what's name?

User: Sandy


Juzz Chabot: Thanks, sandy! We are open for a 30-min zoom appointment every Tuesday and Thursday between 2pm and 6pm. You may book directly at our website <https://juzzmarketingservices.com/contact>

Agent

USER SAYS

COPY CURL

hi hello

 DEFAULT RESPONSE ▼

Hi! Thank you for contacting Juzz Marketing.
Let me provide you with some information so
you can better understand our services. May I
know if you are a Business Owner, a Property
Agent, or a Financial Advisor?

USER SAYS

COPY CURL

financial advisor



DEFAULT RESPONSE



Hi, here is some information for you:

Package 1: 500 targeted leads \$1000 (suitable for agents who are comfortable doing cold calling)

Package 2: 50 appt settings(\$3000) + 50 opt in leads (\$1500) = \$4500 (Appt settings; We set appts for you with name, age, occupation, date, time, and location of appt for you + Opt in leads; obtained through digital marketing by offering a free gift+wealth sharing session/financial review)

Guarantee min 3 cases closed. Target 10 cases closed. If the unlikely event if you don't close at least 3 cases, we will send you more leads until you close 3 cases Because we guarantee a min of 3 cases closed while you target 10 cases closed, therefore it's totally risk-free for you while you enjoy so much upside potential.

Go to
<https://juzzmarketingservices.com/leads-for-financial-advisors/> for additional information.
Let us know if you have additional enquiries.
Thank you!

Agent

USER SAYS

[COPY CURL](#)

can do a zoom to share more



DEFAULT RESPONSE



Hi! Could you tell me your name?

CONTEXTS

[RESET CONTEXTS](#)

08b55f01-d5e8-4fa0-ac23-8f7eb92b6d6
d_id_dialog_context

zoom_appointment_intent_dialog_context

zoom_appointment_intent_dialog_params_person_name

__system_counters__

INTENT

[Zoom Appointment Intent](#)

Agent

USER SAYS

[COPY CURL](#)

sandy



DEFAULT RESPONSE



Thanks, sandy! We are open for a 30-min zoom appointment every Tuesday and Thursday between 2pm and 6pm. You may book directly at our website <https://juzzmarketingservices.com/contact>

INTENT

[Zoom Appointment Intent](#)

ACTION

Not available

PARAMETER

VALUE

person_name

{ "name": "sandy" }

SENTIMENT

Query Score: 0.3

Classify Database's Gender

Machine Learning techniques have been used to classify Juzz Consumer Database into Male and Female that consists of only its names as shown in the picture below. Previously, it was done manually by one of Juzz employees. With Machine Learning, this can be done automatically within minutes while it can take up to hours for a human to do it. Count Vector has been used to transform the text into a vector on the basis of the frequency of each word that occurs in the entire text. Dict Vectorizer has also been used to transform lists of feature value mapping to the vectors. Naive Bayes Classifier has been used for the classification of gender. The reason why Naive Bayes Classifier was selected is because it is easy and fast and need lesser training dataset in the prediction. It also assumes that the dataset is independent which holds true in this case.

The mode has an accuracy of 60.60%



```
# Naive Bayes Classifier
[209] from sklearn.naive_bayes import MultinomialNB
      clf = MultinomialNB()
      clf.fit(X_train,y_train)
      clf.score(X_test,y_test)

0.6060606060606061

[210] # Accuracy of our model
      print("Accuracy of Model",clf.score(X_test,y_test)*100,"%")

Accuracy of Model 60.60606060606061 %

[211] #Accuracy of our model
      print("Accuracy of model", clf.score(X_train,y_train)*100,"%")

Accuracy of model 100.0 %

[216] #Sample1 Prediction
      sample_name = ["Mark"]
      vect =cv.transform(sample_name).toarray()

#Female is 0, Male is 1
      clf.predict(vect)

array([1])
```

Limitations of Juzz WhatsApp Chatbot

Some prospective customers may prefer to call in directly instead of sending a whats app text. Currently the Chatbot does not have any voice recognition feature and therefore unable to pick on phone call enquiries

Limitations of Gender Classification

The dataset that we are given are pretty small. Due to this small dataset, the machine may not have enough data to train and predict the model accurately. Currently the machine is also unable to predict names that are non Singaporean as the dataset uploaded for training comprises only of local Singaporean

Future Works of Juzz Whats App Chatbot

- Training Corpus can be future increased for intent detection.
- To integrate voice recognition so that both whats app text and phone enquiries can be attended.

Future Works of Juzz Gender Classification

More data can be added to the training dataset to help the machine better predict the gender.

References

How to classify Gender by Name

<https://www.geeksforgeeks.org/python-gender-identification-by-name-using-nltk/>

Google Dialog Flow

<https://cloud.google.com/dataflow/docs/tutorials>

How to scrub data using Python beautiful soup

<https://realpython.com/beautiful-soup-web-scraper-python/>

Appendices

Details of all defined Intents

1. Default Welcome Intent
 - a. O/P context: gotousercontext
2. GetGivenName
 - b. i/p context: gotousercontext
 - c. o/p context: awaiting servicetype
 - d. Entity: ServiceType (Financial Advisor, Property Agent, Business Owner)