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

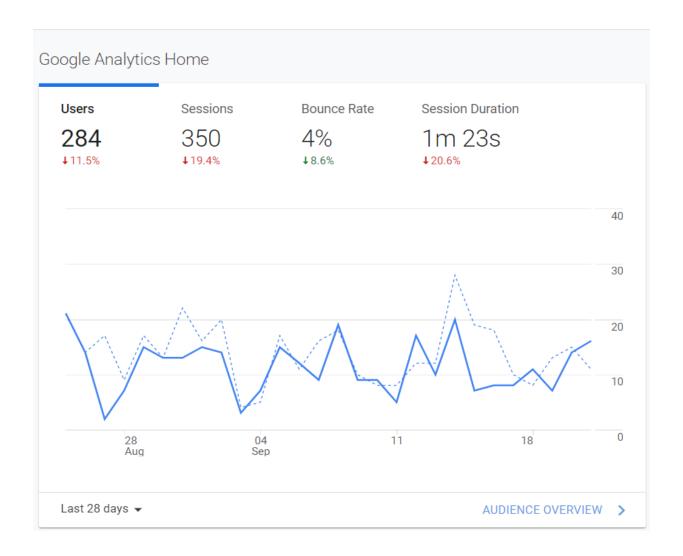
Executive Summary	4
Introduction	8
Project Planning & Design	8
Picture Flow Diagram	9
WhatsApp Chatbot	9
Sales Enquiry Handling	9
Complaint Handling	10
Database Classification	10
Overall Project Scope	11
Project Technicality	11
Juzz Marketing Website Link Scraping	11
Question Database	12
DialogFlow	13
User Category	13
Testing of the Juzz Chatbot Workflow	20
Integration with WhatsApp	26
Integrate with Google DialogueFlow	27
Advertisement User Journey	28
Classify Database's Gender	30

Monitoring of Bounce rate and Sales Conversion of Landing Page	32
Limitations of Juzz WhatsApp Chatbot	33
Limitations of Gender Classification	34
Future Works of Juzz WhatsApp Chatbot	34
Future Works of Juzz Gender Classfication	34
References	35
Appendices	36

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. Although there is a lot of information in the website, the website is unstructured and there is no sitemap to assist in user navigation and users usually struggles to find the queries they wanted and eventually would need to request for a phone call or zoom meeting which is time consuming and unproductive for Juzz Staffs.

Based on the google analytics as shown in picture below, 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 WhatsApp 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 does not have a sitemap to assist user in the navigation.

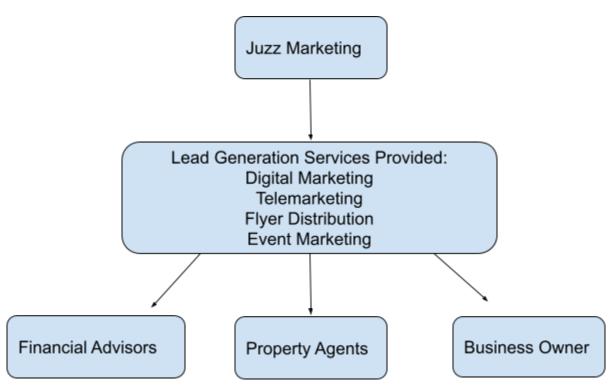


Juzz has a huge consumer database that consists of only name and mobile generated for its clients and wish to seek more meaningful information from this such as finding out their gender using artificial intelligence instead of manually input. There is a need to do gender classification because there 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 Juzz 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.

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 WhatsApp Business for the chatbot design.

Juzz Services Chart



Introduction

Objective of This Project

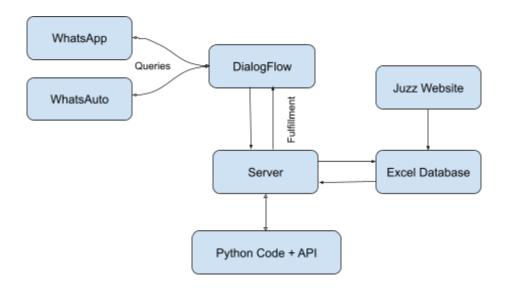
The objective of this project is to design a chatbot in English using Google Dialogflow for Juzz Marketing to automate its WhatsApp Business enguiry messages and to use machine learning to help Juzz Marketing automate the gender classification of its consumer database that consists of only name and mobile 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. And most of these consumers opt in their name and mobile without the gender. The reason why upgradelifestlye fun only collects name and mobile is because when less data is required, the cost of conversion perl lead will be lower. If consumers need to fill in a lot of info, less people will be willing to do so resulting in a higher cost per lead.

Project Planning & Design

This project aims to solve 2 problems. Firstly, to automate WhatsApp Business replies on service enquiries by creating a chatbot using Google

DialogFlow. Secondly, to apply machine learning techniques to help Juzz classify its consumer database's gender into male and female that currently has only names and mobiles.

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 WhatsApp and are currently manually replied by their internal staff. Due to their heavy workload, the WhatsApp 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 WhatsApp 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 WhatsApp Business
- 3. The database gender classification will be analyzed and classified using Natural Language Processing(NLP) techniques.

Project Technicality

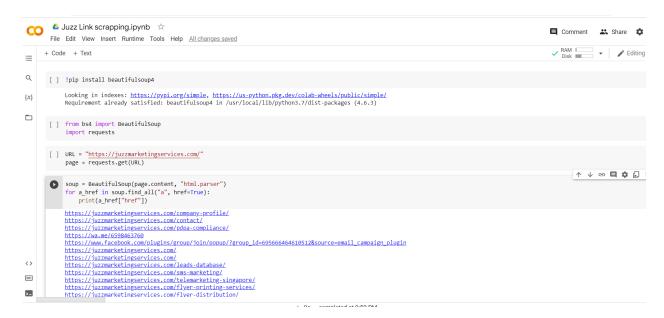
- Beautiful soup is used to scrap the hyperlinks of Juzz Marketing in order to derive at the potential questions users will ask in WhatsApp Business. A total of 32 hyperlinks has been scrapped.
- 2. The scrapped information is used to build into potential questionnaires that potential users will ask
- 3. Google Dialog Flow NLP is used to process user intents
- 4. Machine Learning using Python is use to classify database
- 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 has been chosen to do website scrapping. 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 searched for hyperlinks within. The process is reflected in picture 2 below.

Picture 2



Question Database

After deriving the hyperlinks, a list of questions was then drafted to identify 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. Past conversation that users had with Juzz that was replied manually by one of Juzz Staffs was studied and sorted out to help

the team to design possible intents that potential customers will ask during an enquiry. A total of sample questions was coded to google dialogflow to map out possible user intents to ensure that juzzchatbot is able to answer most of the user queries.

DialogFlow

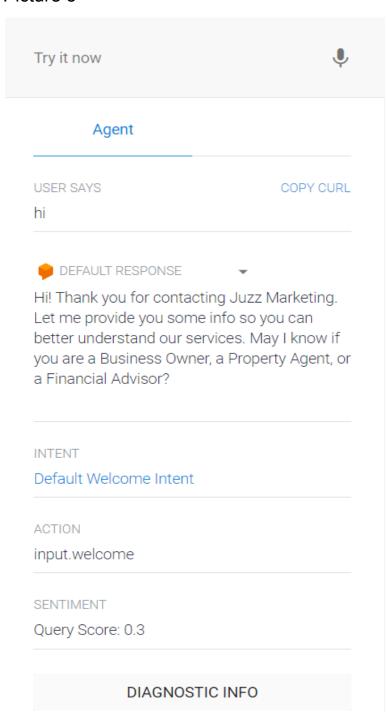
User Category

The user input(utterances) are processed by Google Dialog Flow and matched against the intent. "Default Welcome Intent" are designed to handle the greeting utterance and "Default Fallback Intent" are designed to capture user inputs(utterance) which do not get matched to any other category. 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. There are around 20,000 financial advisors in Singapore. They provide financial planning in terms of wealth protection and wealth accumulation and estate planning to their clients

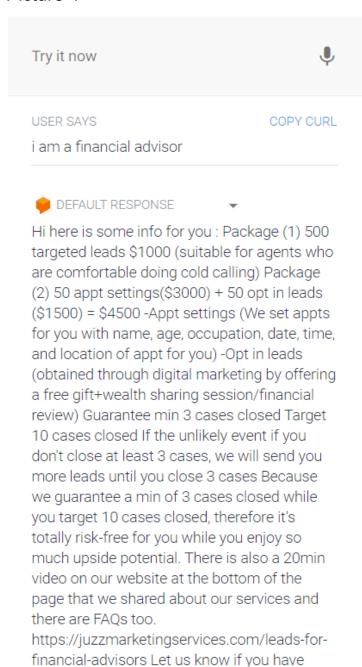
The welcome intent is shown in Picture 3 below. When someone sends a greeting utterance such as "Hi". Google Dialog flow will ask them whether are they business owners, property agents or financial advisors.

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.

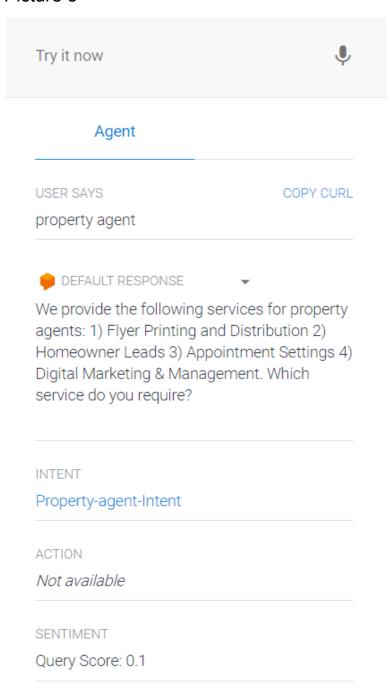
Picture 4



additional enquiries. Thank you!

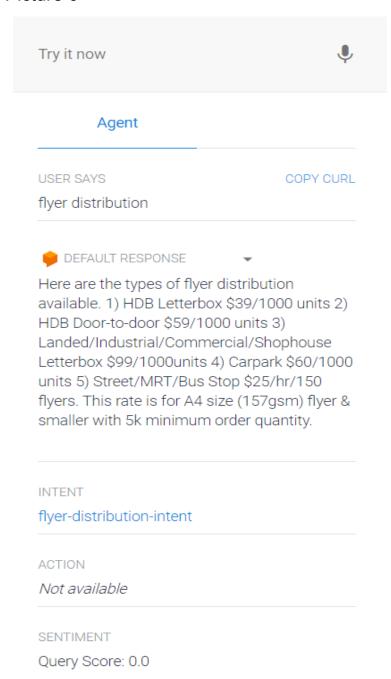
If the user is a property agent, Juzz chatbot will will reply accordingly. Take for example, in Picture 5 below, the user replied "property agent" Juzz Chatbot will responses and provide the services and rates to the user.

Picture 5

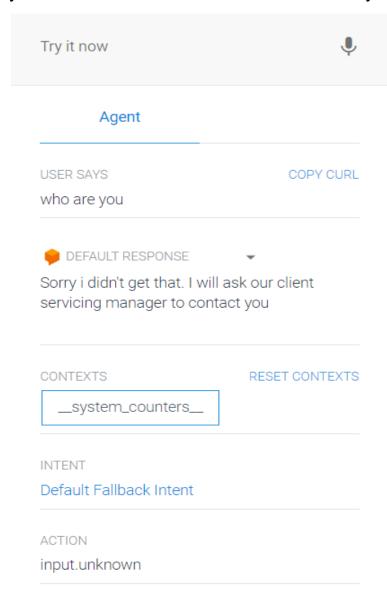


When the user identifies himself as a property agent and keen on flyer distribution services, juzz Chatbot will reply with flyer distribution services details and rates to the user.

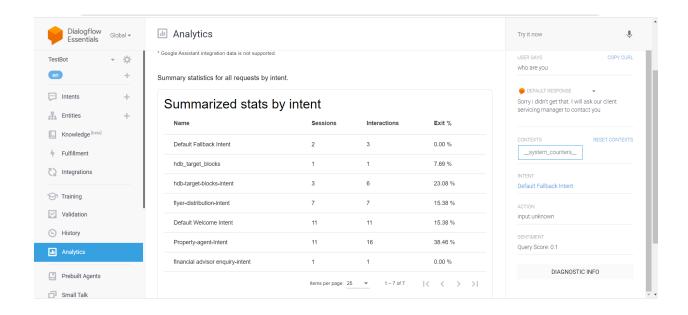
Picture 6



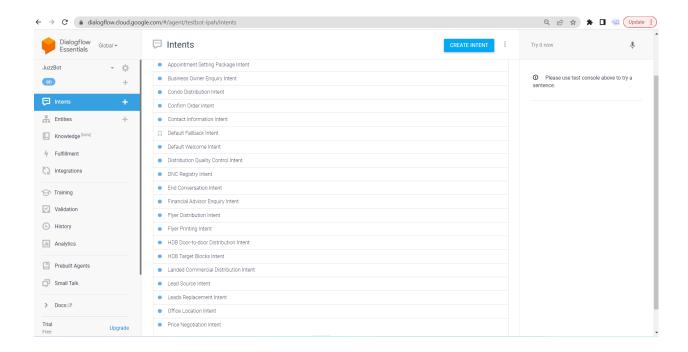
A default intent has also been created to handle questions that users will ask but does not fall inside the intent already designed yet as shown in the picture below. Users will always ask questions out of the box such as "Who are you" or some people could message us wrongly such as "Where are you now?"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



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.



Besides default and welcome intents the following intents were also programmed into google diagueflow as well as shown in the picture below.



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 Chatbbot: 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

21

Agent

USER SAYS

COPY CURL

hi hello



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?

financial advisor



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-forfinancial-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_contex

zoom_appointment_intent_dialog_param s_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

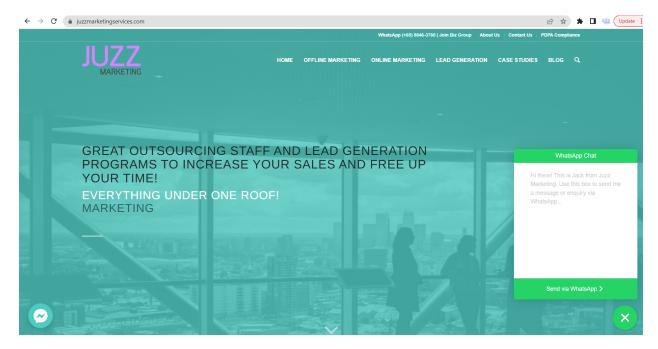
{ "name": "sandy" } person_name

SENTIMENT

Query Score: 0.3

Integration with WhatsApp

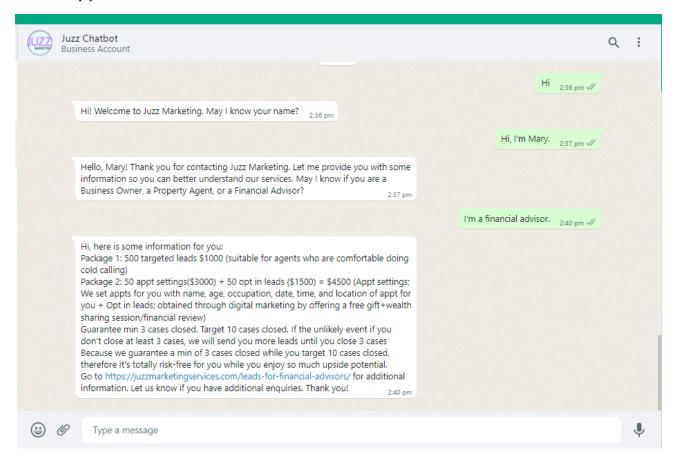
For integration with WhatsApp Business, user just need to contact Juzz WhatsApp business mobile number by clicking on the whats app icon as shown in the screenshot below.



Alternatively, users such as business owners, property agents and financial advisors can also save Juzz WhatsApp business mobile number +6598463760 to their phone and search the Juzz WhatsApp Business mobile number on their WhatsApp App.

Once a user clicks on the whats app icon, the user will be able to interact with the chatbot as shown in the screenshot below.

WhatsApp Web Screenshot



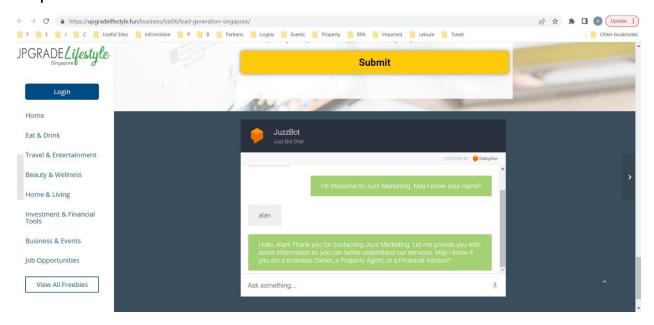
Just like any normal conversation in WhatsApp, user may now start chatting with Juzz Chatbot.

Integrate With Google Dialogue Flow

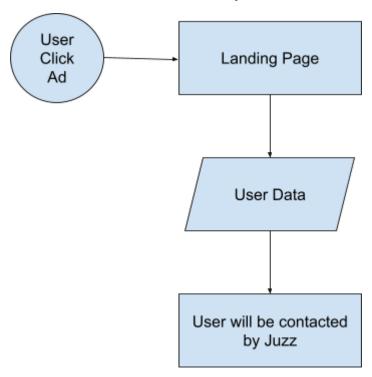
Juzz Marketing also uses <u>upgradelifestyle.fun</u> for lead generations to generate its potential clients such as financial advisors, property agents and business owners. It runs facebook ads to target business to business clients and uses

https://upgradelifestyle.fun/business/biz06/lead-generation-singapore/

as a landing page to generate leads for its business. Google Dialogue Flow has been integrated into this landing page too as shown in the screenshot below so that when users have some queries, google dialogue flow can help to answer them.



Advertisement User Journey



- 1. User clicks advertisement
- 2. User will be redirected to a landing page

https://upgradelifestyle.fun/business/biz06/lead-generation-singapore/

- 3. User will fill out the form and tick the privacy policy, disclaimer and terms & conditions
- Privacy Policy

https://upgradelifestyle.fun/privacy-policy/

Terms & Conditions

https://upgradelifestyle.fun/terms-conditions/

Disclaimer

https://upgradelifestyle.fun/disclaimer/

- 5. User proceed to submit the form
- 6. User will be contacted by Juzz Marketing

Sample Advertisement



Classify Database's Gender

The business goal is to classify Juzz's consumer database obtained from its lead generation website www.upgradelifestyle.fun and the only data feature that we have is the names. 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. The dataset has been preprocessed to check for column name consistency and to check missing values. All females have been replaced as 0 and all males have been replaced to 1 to make it machine readable. 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 it requires less training dataset in the prediction. It also assumes that the dataset is independent which holds true in this case.

```
[] # Sample2 Prediction
    sample_name1 = ["Peter"]
    vect2 = cv.transform(sample_name1).toarray()

[] #Female is 0, Male is 1
    clf.predict(vect2)
    array([1])

[] # Sample3 Prediction
    sample_name2 = ["Jessie"]
    vect3 = cv.transform(sample_name2).toarray()

#Female is 0, Male is 1
    clf.predict(vect3)
    array([0])
```

Based on the picture above, 3 random names were chosen to predict the gender. The first name chosen was "Mark" and it returns an array of 1 which means it has predicted that "Mark is male. This is accurate The second name that was chosen is "Peter" and it also returns an array of 1 which means it has predicted that "Peter" is a male. This is accurate too. The third name chosen was "Jessie" and it returns an array of 0 which means it has predicted that "Jessie is female. This is accurate too.

Moving forward, whenever Juzz has a name list that needs to predict gender. They do not have to do it manually anymore. Take for example, someone fills in the form at

https://upgradelifestyle.fun/investment/biz05/dear-home-owners/ which consists of only name, e mail and mobile and does not contain the person's gender, Juzz can use this model that is built to predict the name.

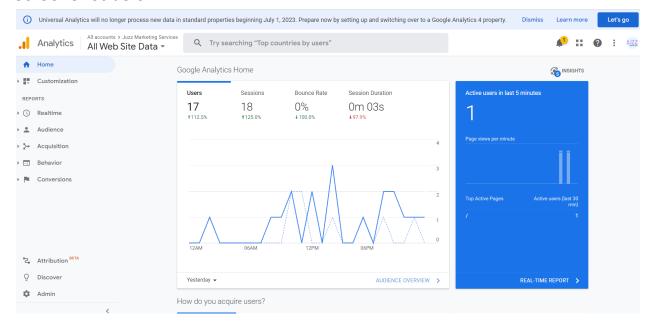
Your name			
Your email			
Mobile			

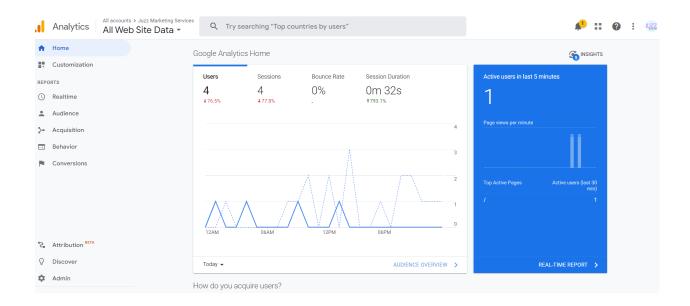
The model managed to achieve an accuracy of 60.60% on the testing dataset and 100% accuracy on the training dataset as shown in the picture below based on the 100 dataset provided. Previously, based on human manual prediction, it was only able to achieve a 50% accuracy. Therefore, using the model to predict the gender is still more accurate than using a

human to do manual prediction although there is room for improvement moving forward.

Monitoring of Bounce rate and Sales Conversion of Landing Page

After automating Juzz whats app chatbot, the website has managed to obtain a 0% whats app bounce rate for the past 2 days as shown in the 2 screenshot below





However, this does not mean that 0% bounce rate will last forever. The monitoring of the bounce rate of Juzz Marketing website https://juzzmarketingservices.com/ will be monitored on a monthly basis to ensure that the automation of whats app Juzz chatbot helps to reduce the bounce rate of the main website. The conversion of the landing page https://upgradelifestyle.fun/business/biz06/lead-generation-singapore/ will also help to increase the number of leads generated

Limitations of Juzz WhatsApp Chatbot

Some prospective customers may prefer to call in directly instead of sending a WhatsApp text. Currently the Chatbot does not have any voice recognition feature and therefore unable to pick up the phone call enquiries. Google dialog flow is only able to pick up the first word of non-trained inputs too.

Limitations of Gender Classfication

The dataset that Juzz gave to the project team is 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 WhatsApp Chatbot

- Training Corpus can be future increased for intent detection.
- To integrate voice recognition so that both WhatsApp text and phone enquiries can be attended.
- Integrate enquiries with Google Spreadsheets
- Integrate Google Calendar for appointment bookings

Future Works of Juzz Gender Classfication

More local Singaporean and foreigners data can be added to the training dataset to help the machine better predict the gender. Juzz also needs to include some foreign names as well because given Singapore's 5.6 million population, it has 1.3 Million foreigners. Some of the foreigners may opt in to upgradelifestyle.fun but currently their gender could not be classified accurately due to lack of foreign database. More training dataset needs to be fed into the model as well to increase the accuracy. As currently, its accuracy is only about 60% based on the test dataset.

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/

Compare different machine learning classifiers using Python

https://www.youtube.com/watch?v=QINjjSge65Y

Appendices

Details of all defined Intents

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

Business Owner)

- 3. Financial Advisor Enquiry Intent
 - e. i/p context: defined servicetype,
 - f. o/p context: defined servicetype,
 - g. Entity: none
- 4. Property Agent Enquiry Intent
 - h. i/p context: defined servicetype,
 - i. o/p context: defined servicetype,
 - j. Entity: none
- 5. Business Owner Enquiry Intent
 - k. i/p context: defined servicetype,
 - I. o/p context: defined servicetype,
 - m. Entity: none
- 6. Flyer Distribution Intent
 - n. i/p context: defined servicetype,
 - o. o/p context: defined servicetype,
 - p. Entity: none

7. Flyer Printing Intent

- q. i/p context: defined servicetype,
- r. o/p context: defined servicetype,
- s. Entity: none

8. Zoom Appointment Intent

- t. i/p context: defined servicetype,
- u. o/p context: defined servicetype,
- v. Entity: @sys.person person name

9. Confirm Order Intent

- t. i/p context: defined servicetype,
- u. o/p context: defined servicetype,
- v. Entity: @sys.person person-name, @any.package any-package,
- @sys.phone-number phone-number, @sys.email email

10. End Conversation Intent

- w. i/p context: defined servicetype,
- x. o/p context: defined servicetype,
- y. Entity: none