Introduction

## Background

The lack of expertise that people have on how to perform a good training session

has a series of issues. The two issues we encounter are that it leads to inefficiency and

bad practices that can lead people to quit or injure themselves. With the cost to have

a personal trainer or somebody to manage a person’s training session being very high,

there is an opportunity to improve the user experience with automated services that

provide relevant information and keeps track of the progress that is being made.

Artificial Intelligence has come a long way since the 1950s and today’s computing

capabilities have opened a whole world of possibilities, the evolution of tools in language

processing has seen an increasing amount of applications which mix human and machine

interaction like nobody has ever imagined.

The objective of this project is to explore the capabilities of Machine Learning with

Natural Language Processing by using the tools available and study it’s integration

with a real-world application such as a trainer chat bot. Over the years chat bots have

steadily been incorporated into society, the flexibility and cost chat bots have, makes a

very lucrative option for monotonous operations such as some provided by a personal trainer.

## Motivation

The motivation for doing this project comes from my own experience in gyms where I have seen people not knowing what exercises to do, and those that thought how to do them many times didn’t. So I bean investigating on the different patterns which people who go to the gym had in common, by looking at google trends we can see that there is a defined pattern that replicates year over year, where at new year’s eve everybody plans to start fresh the year, but as we can see it doesn’t last long. After that I decided to check with other people what they thought by doing a survey which was quite enlighttning.

* Over 55% would be motivated to go to the gym if they were assisted by a personal trainer.
* 35.1% would follow instructions if that personal trainer was a chatbot, this is normal as people still prefer a human touch, but the cost is also higher in that case.

Observing this and with my previous experience with chatbots I decided to create a proof of concept that could be used as a demonstration to see if further developing the chatbot would be wise.

## Objectives

The main objectives for this product are to reduce the amount of dropouts through out the year and improve the user’s health by keeping them motivated and accomplishing objectives in a way that doesn’t injure their muscles.

* Improve dropout rate in gyms
* Improve health

Context

## Socioeconomic Environment

* Social Impact

The positive impact of this product socially is the fact that when fully developed will

provide a way for connecting with new people and make it easier for people to train

without having to spend a large amount of money or time to learn how to do exercises

correctly or what routes are better to run.

Another social impact of this product if it catches the user’s attention and use it

regularly is that the health benefits of training will compensate indirectly with health

issues caused by obesity or heart problems.

* Economic Impact

There are two sides to the economic impact this project may have in society, first would be the increase in users using this platform will highly likely increase the

number of users going to gyms, with this the profits for these gyms will also increase,

creating not only more jobs to support the influx of users but maybe more opportunities

for gym openings. Also, by providing a cheaper tool for users, the cost of hiring a

personal trainer is reduced to the fee for using the service this product would provide.

The other side of the economic impact would be how this project affects personal

trainers, as can be verified in the query done 53,1% prefer the services of a professional rather than a chatbot, this means that there is still a market for personal trainers, still we think that the development of this project could impact negatively on the people working as personal trainers, but comparing the pros and cons we consider that the benefits in the economy compensates any negative impact on the sector.

Outside this the savings caused by chatbots can account for billions in expenses, those billions saved can be used to further invest on a company’s growth and by that impacting positively on the global economy.

## Regulatory Framework

* Ethical risks

----

The ethical impact of this product is something that must be analyzed as this type

of technology even though have been around for some time, at the pace it is evolving

can soon cause them to truly seem human like. This is a problem or at least something

that must be taken account as people don’t like to feel fooled, if we consider the impact

Google Duplex has had since it launched, people where amazed and scared of the fact

that the assistant used human behavior while talking like humming while it searched

for information.

Other things to consider as a social impact would be the other way around, the

vulnerability of chatbots, as it is a machine interacts with people it must be correctly

designed to stop users from accessing personal information or extracting information

from the system.

-----

Chatbots improve based on conversation data from users, this has several ethical

impacts that must be considered. People in conversations can distinguish when a person

is being racist or verbally abusive of other people. The problem is that chatbots

now don’t have those capabilities, so as the bot learns from conversations if the design

is not implemented to mitigate this, and it is built to generate its own responses it can

end up becoming racist or homophobic with other people. Most chatbots don’t have

this problem as the bot’s answers are mostly hard coded. But as chatbots evolve to

learn how to answer dynamically as a person, the design must be capable of cleaning

the phrases in order to remove any discriminating words.

If not you can find problems such as the chatbot built by Microsoft which used unsupervised training from users and became in less than 24 hours completely corrupted by what it

learnt.

* GDPR

General Data Protection Regulation, GDPR, is a legal framework that protects citizens

data in the European Union (EU), this framework means that companies must

comply with user privacy and facilitate EU consumers with guidelines about what information

is being acquired from the user as well as the option to delete it. Also, in

case of a data breach on the webpage it must notify the user if any personal data was

stolen.[6]

For chatbots there are a series of guidelines on how to be GDPR complaint.

• **Data transparency:** Know what data is going to be extracted from the user.

Notify it through the privacy policy.

* **Data storage:** Separate user data from the rest of the data and with any info

that can identify the user encrypt it.

• **Data deletion:** Offer the user the option to remove all the data if asked.

• **Data retrieval:** Allow the user to know what data is being stored and retrieve

it.

• **Privacy-first design:** Develop chatbots with privacy in mind in order to avoid

restructuring the bot afterwards. Ask for the users consent to acquire their data.

As chatbots are in constant dialog with users, this can be done easier by explaining

with a message what data is being extracted.[6]

* Future Regulations

AI is a fast-moving field, because of this it is important to be careful where regulations

are set, as a badly place regulation could negatively impact the innovation in this

field and reduce the economic growth generated by such innovations. For this reason,

regulatory entities should hold off in directly regulating AI until it stabilizes. This will

better provide insight into where it is better to add regulation in benefit of the consumer.

Another aspect of the chatbots is that people with personal issues prefer sharing

them with the bot to let out some steam, even comments about suicide instead of

looking for human help. This may be because they know the bot won’t judge as a

person would. The problem is that most bots aren’t designed to deal with this as it is

out of the scope of the bot’s core functionality. This may require future regulation or

standardization to make this kind of comments redirected to the corresponding entities

to provide the adequate assistance.

Solution Proposed

## Sam presentation

So, the chatbot that is set to be the best virtual personal trainer goes by the name of Sam, stands for Sports Assistant Machine, which was chosen as it is a name that is unisex and easy to remember. Sam has gone through two remakes, the first is a proprietary solution which focus was in control over Sam’s brains, here are the technologies used for this first design.

* Sam presentation

## Initial Design

* Technologies used

So from left to right we have:

* Telegram: Which was used as the user interface for the chatbot as it provides the capabilities to have a chatbot and also has a very well documented API to help when communicating the server with Telegram’s server. I also has the ability of using webhooks that in this design was used with scalability in mind.
* Raspberry Pi: On the server side I use a Raspberry Pi as Sam’s office, where all the different modules will be.
* In order for Telegram to communicate with our server I use Nginx Web Framework as it can be used as a Reverse Proxy, which is a way of using more ports which aren’t the predetermined ones without visually affecting the URL used to communicate with the server
* PostgreSQL was used as the relational database manager for Sam, this mainly was done as I had more experience using Postgres than other database managers.
* Python was also used as the programming language for the amount of support it has for machine learning libraries and Natural language processing tools and also because I have experience developing chatbots with this language.
* Structure

This is a simple schematics of how the chatbot will work functionally, the most important components are the Activity Manager and the Context Manager.

The most complex task was when developing the context manager as developing it to be conversational was quite challenging and ended being very time consuming. The process from Sam receiving the message and responding were the following.

1. When receiving the message if no context was detected it goes through the process of tokenizing, where the phrase is spliced into words, leaving out irrelevant punctuation signs.
2. Then it goes through the process of stemming the word where it limits the vocabulary Sam has to understand by removing conjugations such as exercising to exercising, as in the bot doesn’t need to learn both words, with one it can get enough understanding of what the user wants to use.
3. After that it translates this stems to number where each word has an identifier and it goes through a classifier. Sam will check what context is related to what the user is saying and it relates it to the user.
4. After knowing what the user is saying Sam tries to extract the relevant information to fully understand the data.
5. Then it will perform an action based on that information
6. Finally it sends the user a response.

* Time to change

So, the problem is that developing a functional context manager isn’t as simple and different factors limited the time to develop something interesting, so I had to make the choice of scrapping a big chunk of the work done up to then in order to use a quicker solution to develop Sam.

## Second Iteration

* Technologies used

So many of the components remained the same such as Telegram, Postrgres and Python.

But instead of creating a proprietary solution I decided to switch to Rasa which is a low level chatbot framework which provide several tools to create chatbots. It has to modules, the core module which is were the activities are launched and the initial communication happens, and the other module is the NLU that does it is used for understanding the user and also makes Sam conversational. The system Rasa uses to keep the context for the user is by the use of stories which are conversation flows that the interactions can have, it has it’s pros and cons. The pros is that it is easier to create conversation flows, the disadvantage is that it is hard to maintain and also it is very difficult to design robust conversation flows designed to direct the conversation with the user.

The next change had to do with rasa, which due to being pretty resource heavy it required of more powerful hardware, so in order to keep developing Sam I had to migrate it’s office to the cloud, for this I used Azure as they gave me 100 dollars in credit to use for being a user.

* Structure

The process the message goes through when going through rasa are the following:

1. First the message goes through the interpreter where Sam understands what the user is talking about,
2. Sam stores this information in a tracker to use it through the processing
3. After understanding what the user instead Sam tries to give context to what the user said and what needs to be done.
4. Based on that, Sam performs an action
5. Finally Sam sends a response to the user and also saves it in the tracker.

## Functionality

* Existing functionality
* **Exercise table creation:** One of the most time-consuming and boring things

to do is prepare a routine to follow, there are applications that help out with this,

but the idea for this project is to make the interaction more human-like, having

the bot automatically generate tables based on what muscles have been trained

and what muscles are not proportionate to the rest of the body. As this is a proof

of concept and due to the time constraint of this project, the functionality that

tracks what the user has exercised and the table creation will be limited.

* **Diets table creation:** Like what the above point talks about, the objective

here is to make it easier for the user to know what to eat in order to reach their

desired weight or to gain muscle mass. As some users may have different meal

preferences, for the final product the chatbot can take food preferences in to

account to provide different meals.

* **User tracking:** The objective is to track the user’s progress and see how close

they are of fulfilling their goals, this will provide data on each user’s performance and how to further adapt their routines to make it easier to follow.

## Demonstration

* Judges demonstration
* If wrong: Video presentation

Planning and Budget

## Planning

The original plan was to develop the whole project in five months in two different

stages and documenting the whole process, the first stage was programmed to las two

weeks designing and thinking on the objective of the application and making sure that

the idea is viable before writing a line of code with a form sent to possible users. The

next stage was the development step, first creating the tests that will confirm the application

works correctly, this stage was predicted to last 4 months.

After the five months of development, due to external factors, the project was seriously

delayed and a redesign had to be done as the complexity of the context manager

in the original design made it unviable to deliver for this year. This made it necessary

to rethink how to adapt the project to reach the target of September for delivery. In

order to reach the objective an extension on the different parts of the project have to

be done, this included two more weeks of planning and preparation, twelve more weeks

of development where the new configuration and implementation was completed, the

rest of the extension was filling out the documentation with the new design.

Even though this was a setback to the project, it allowed gaining experience in

how the perception of time in completing tasks has to be refined in order to avoid

overextending a project and having inefficiencies.

## Budget

For the hardware in the initial design the only cost was the raspberry Pi which was minimal.

For the software, the only software used was the virtual machine from azure. The

reason the cost for the Azure Machine is relative is that it has a pricing system based

on paying what you use, meaning that depending on how much you use the resources

the pricing varies. For this project the price of usage at the moment of writing this

memory has been 40 e as virtual machine was used only when required to minimize

costs.

The human resources used in the project were where most of the budget of the

project went to, the hourly rate is based on similar projects done and by the knowledge

of people in the industry. The developer works 4 hours a day in average on the project

for the length of the project considering only working days. For the project manager he

works only 5,5 days out of the 22 working days in a month, to control that the project

is on track and make adjustments to it.

Based on all these costs the total budget for the project is FORTY NINE THOUSAND

TWO HUNDRED NINETY EIGHT AND TWO CENTS EUROS.+

.

Future Work

## Next steps in development

* Complete existing functionality

In continuing Sams development I would continue developing it’s functionalities to make Sam a better personal trainer.

* Additional functionality

Also I would explore more advanced functionality to add to Sam like more activities or route planner.

* Voice assistant

With future innovations in voice technology the idea of integrating Sam with voice capabilities can open the door to big improvements such as assisting during a users training session.

Conclusions

## Market for product

* People would use it

Based on the surveys I did at the beginning I know that people would be open to using it, so there is a market there for Sam.

* Benefit to gyms

Also economically Sam would provide indirectly more users to gyms which will enjoy higher profits because of it.

If I were to further develop Sam I would consider returning to finishing my proprietary solution due to some limitations Rasa has.

That is all from me, if you have any questions about the project I will be more than happy to answer.