Chatbot for ManageEngine in Ministry of National Guard - Health Affairs

Prepared by:

- ❖ Nawaf Almutairi
- ❖ Surayyi Alqahtani
- ❖ Ahmad Alharthi
- ❖ Faisal Alsufyani
- ❖ Shoug Alkhathran

Instructor :
- ❖ Mejdal alqahtani

ISID
Information Systems
and Informatics Division

Project Outline

introduction

Department Overview

Data structure

NLP Preprocessing

Modeling

Chat bot

Future work

Conclusion

› **Introduction**

› The MNGHA consists of medical campuses located in Riyadh, Jeddah, Al-Ahsa, Dammam, and Medina. All these campuses are linked to one electronic medical record system (BestCare).

الشؤون الصحية بوزارة الحرس الوطني

أول منظومة صحية في العالم تحصل على الاعتماد خارج الولايات المتحدة

حصلت على **شهادة اعتماد المرحلة السابعة** في نموذج نضج أنظمة تحليل البيانات من الجمعية الدولية لنظم المعلومات وإدارة الرعاية الصحية (HIMSS)

نموذج نضج أنظمة تحليل البيانات

هو نموذج متكامل يرتكز على الجانب التقني وتكنولوجيا المعلومات إضافة إلى حوكمة البيانات وإدارتها وثقافة البيانات وإتاحتها وتمكين الأفراد بمختلف مستوياتهم للتعامل مع البيانات والاعتماد عليها

Department Overview

- **Information system and informatics division (ISID):**

- **ISD (information services department)** is a department for receiving some requests from ManageEngine which is the system that user can order any request that related to information systems issue.

- Provision of full support, help desk and asset management to end-users.

Project Goal

› The purpose of this project is creating a chatbot for the ME service desk system how can be interactive with the users and enters all the required information without the need for real Agent to save the time by decrease the load of the routine problems, in certain cases transfer the users to the real agent.

Data Structure

› There is no data , we create a templet data from crash.

› Until now we have :

  › 278 entity

  › 3 columns (tag – patterns – responses)

Data Structure

› The are details must fill in :

› Requester details
- Requester name
- Asset

› Service category
- Category
- Subcategory
- Item
- Select the software

› Asset information
- Isd name:

› Asset location
- location
- Building
- Floor number
- Room number

› contact information
- Extension , pager
- Badge number
- Email

Data structure

NLP Preprocessing

- Tokenization

- Remove stopping word and punctuations

- Steaming

- chucking (NER)

- Part of speech

- Count vectorizer

- Term Frequency — Inverse Document Frequency(TF-IDF)

NLP Preprocessing

n email I need an email I need a new email I want a new email Can you create a new mail for me Im a new employee and I dont have an email Im a new organizer and I dont have an email I have been moved from one department to another and I need a new email I have not received an email yet Department email user email email srevices My printer does not print papers I thank the lnk is empty My printer does not have a toner My printer does not have a lnkMy ink is empty My toner is emptyI want to order new lnkI want to order new tonerI want to order new black lnkI want to order new black tonerI want to request a new toner I want a black lnk I want a black toner the color is black I need all colors I need a red and blue I need a black and blue My ISD Name is My printer model is Nawaf AlmutairiAhmad AlharthiFaisal AlsufyaniSurayyi Alqahtani Shoug Alkhathran extension badge em building floor roomresponses Hellohow can I helpppp pphiwhat is your problem Have a nice day Your welcome bye Happy to help Any time My pleasure Sorry cant understand you Please give me more info Not sure I understand What kind of the software issue you havedo you need a medical application servicesdo you need mobile application servicesdo you need non medical services Ok I need your contact information please what is your name ok sure what is your justification for ordering the printer ok May I know your justification for requesting the printer ok no problem give me your information please what is your name ok Give me a description of your need to create a new mail please Fine What item do you need is it Department email or user email or email srevices ok what is your name Ok You want to order new toner or lnk What kind of color Sure i will do that You want to order new toner

NLP Preprocessing

Is these a steaming or lemmatization ?

```
greeting    :   greet
Thanks      :   thank
helping     :   help
information   :   inform
I   :   i
problem   :   problem
software    :   softwar
```

Yes , steaming

NLP Preprocessing

- Named Entity recognition

I need all colors, I need a red and blue, I need a black and blue My ISD Name **ORG** is ... My printer model is ..

Nawaf Almutairi **PERSON** , Ahmad Alharthi **PERSON** , Faisal Alsufyani **PERSON** , Surayyi Alqahtani **PERSON**

, Shoug Alkhathran **PERSON** extension badge em building floor room,,responses, Hello,how can I helpppppp?,hi,what

- Word cloud

NLP Preprocessing

Count Vectorizer

	access	account	address	ahmad	alharthi	alkhathran	all	almutairi	alqahtani	alsufyani	...	without	won	work	working	workplace	works	would
0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
273	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
274	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
275	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
276	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
277	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0

278 rows × 221 columns

TF-IDF

	about	about your	add	add request	allright	allright your	ask	ask you	badge	badge number	...	your building	your contact	your email	your extension	your information	your justification
0	0.000000	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.0	0.000000	0.0	0.0	0.0
1	0.000000	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.0	0.000000	0.0	0.0	0.0
2	0.000000	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.0	0.000000	0.0	0.0	0.0
3	0.000000	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.0	0.000000	0.0	0.0	0.0
4	0.000000	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.0	0.000000	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
273	0.000000	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.0	0.467702	0.0	0.0	0.0
274	0.238247	0.238247	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	...	0.238247	0.0	0.000000	0.0	0.0	0.0
275	0.000000	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.0	0.000000	0.0	0.0	0.0
276	0.000000	0.000000	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.0	0.000000	0.0	0.0	0.0
277	0.000000	0.000000	0.244568	0.244568	0.0	0.0	0.0	0.0	0.0	0.0	...	0.000000	0.0	0.000000	0.0	0.0	0.0

278 rows × 185 columns

Unsupervised Learning Model

- **We used two unsupervised learning models:**

- Latent Semantic Analysis (LSA).

- K-Means.
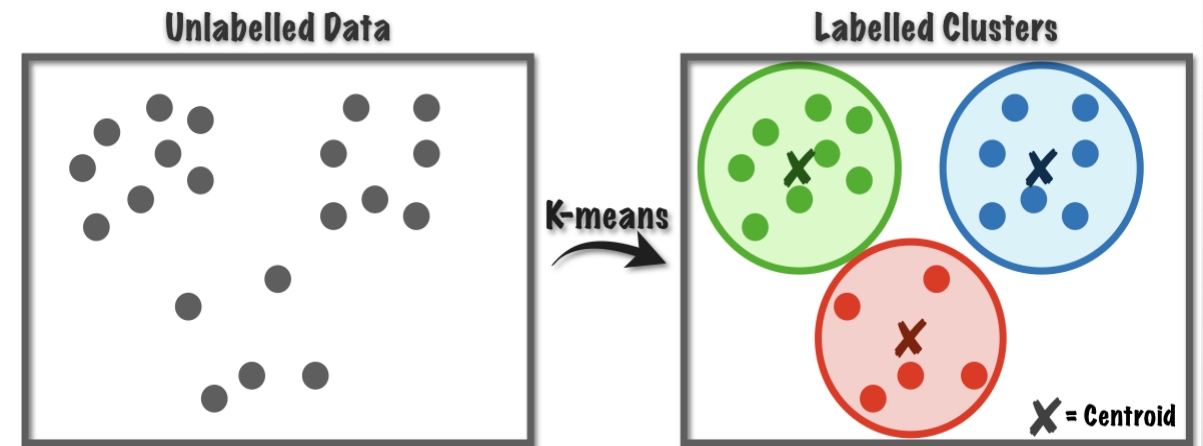
Latent Semantic Analysis (LSA)

- In LSA we got a two cluster:

|	access	account	address	ahmad	alharthi	alkhathran	all	almutairi	alqahtani	alsufyani	...	without	won	work	working	workplace	works
cluster_1	0.007	0.002	0.028	0.0	0.0	0.0	0.002	-0.0	-0.0	0.0	...	0.010	0.004	0.014	0.030	0.007	0.001
cluster_2	0.016	0.002	0.004	-0.0	-0.0	-0.0	0.006	-0.0	-0.0	-0.0	...	-0.004	0.020	0.017	0.084	0.001	0.004

2 rows × 221 columns

- Display topics:

```
Topic    0
printer, my, office, new, is, in, to, have, me, for

Topic    1
software, the, is, with, to, problem, me, application, there, working
```

K-Means

- First, we select range from 1 to 10 , to choose the No. of K

- The optimum number of clustering is 4.

- K-means result :
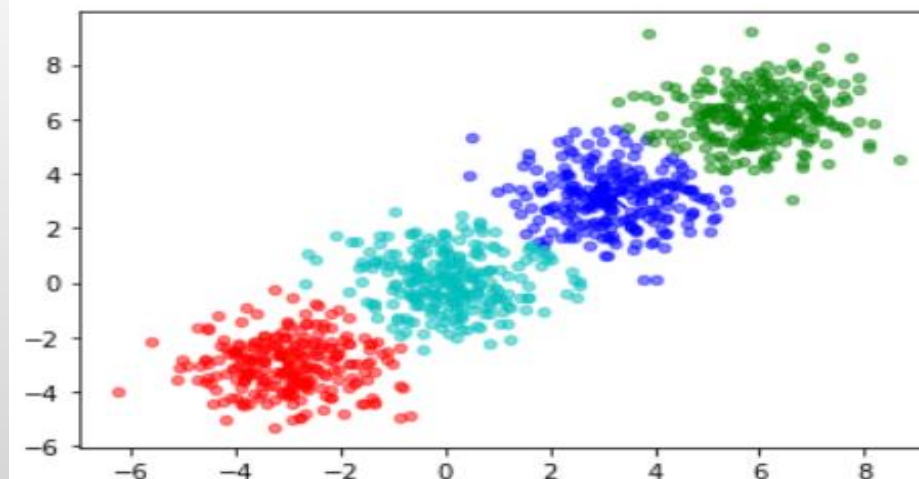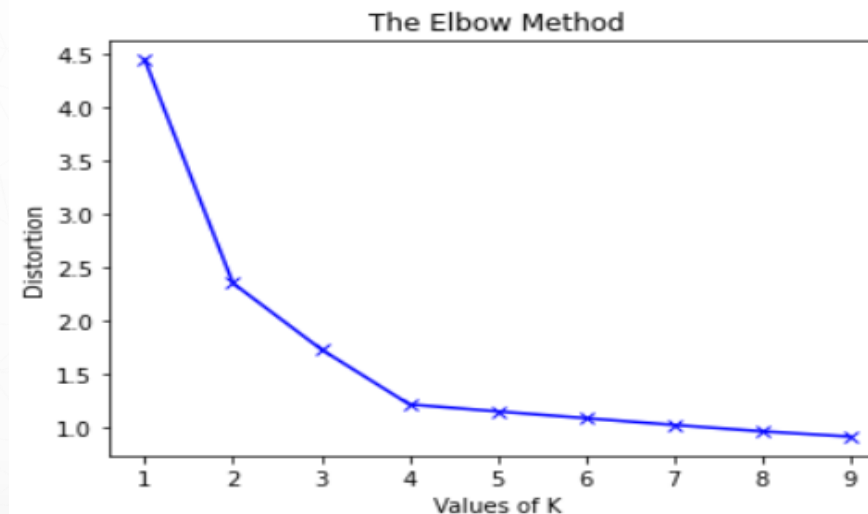
Live chatbot

Future Work

- Convert the ChatBot to CallBot (Interactive AI Call), How ?

- ✓ Transcribe speech to text.

- ✓ Decision making machine.

- ✓ Transcribe speech to speech.

Speech Recognition

Conclusion

- All these campuses in MNGHA are linked to one electronic medical record system (BestCare) which are under the ISD department.
- we create a templet data.
- We clean the text by Remove stopping word and punctuations
- We use a world cloud to see which word repeated.
- The topic models were LSA , K-means.
- The best model is K-means.
- The optimum number of clustering is 4.
- The future work is to build an interactive AI callbot