

## **Raport Lab 1-2**

### **Holobot descriere**

Aplicatia pe care o dezvoltam este o aplicatie de in care utilizatorul va putea interactiona cu un chatbot care este specializat in raspunsul la intrebarile utilizatorului legate de programare.

Problema pe care dorim sa o rezolvam este timpul pierdut de programator in cautarea raspunsului pe care il doreste la intrebarea pe care acesta o ridica in timp ce codeaza. Astfel acesta va putea intreba rapid chatbotul iar acesta va raspunde in timp util la intrebare salvand astfel programatorul de la intreruperea flowului de programare si pierderea timpului cu cautat. Totodata aplicatia poate fi folosita si ca un learning tool de catre studenti si elevi.

Ca si functionalitati aplicatia va permite punerea de intrebari si evaluarea raspunsului primit, astfel botul putand sa fie antrenat continuu si cu intrebari de la utilizator.

Partea legata de ai a aplicatiei va fi realizata in python. Va exista un server rest creat cu django care va avea endpointuri pentru primit raspunsuri si primit evaluari dar si un endpoint pentru primire de date de antrenament, va fi deployat in azure ca si web app si securizat cu jwt token sa nu poata fi utilizat de oricine si putand fi accesat doar de serverul de node.js care va fi intermediar intre frontendul aplicatiei si serviciul de python.

Modelul folosit fa vi chatterbot, o librarie de python care ofera un chat bot neantrenat care ca si algoritmi de baza are algoritmi de cautare si clasificare.

Datele vor fi stocate intr o baza de date nosql pentru a putea fi usor procesate.

Ca si dataseturi avem 60k intrebari de pe stack overflow :

<https://www.kaggle.com/datasets/imoore/60k-stack-overflow-questions-with-quality-rate/>

Si 10% din totalul de intrebari si raspunsuri de pe stack overflow:

<https://www.kaggle.com/datasets/stackoverflow/stacksample/data?select=Questions.csv>

### **Raport Lab 3**

Algoritmul inteligent este scris in python si are la baza biblioteca chatterbot. Aceasta biblioteca ajuta la generarea de raspunsuri la inputul utilizatorului. Chatbotul are si o serie de adaptori logici, filtrte si metode de comparare care ajuta la generarea cat mai buna a raspunsurilor la intrebarile de programare pe care le va avea utilizatorul.

Drumul unui raspuns la o intrebare incepe cu preprocesarea intrebari, cum ar fi curatarea spatiilor goale sau tokenizarea intrebarii. Ca si adaptori logici chatbotul nostru foloseste un adaptor premade din biblioteca chatterbot, best match cu un trashold acceptat de 0.9, si raspuns default pentru situati in care nu se intelege intrebarea.

Datele pe care le foloseste chatbotul sunt stocate intr o baza de date sqlite, chatbotul in momentul antrenari formand un graf de conversatii in baza de date, totodata si in momentul unei conversatii aceasta udateaza acest graf din baza de date.

Ca si antrenament chatbotul a fost antrenat cu conversatii sub forma de liste aceasta putand fi antrenat cu liste care reprezinta conversatiile sau cu corpus data. Listele care contin conversatiile sunt de forma intrebare raspuns si a fost antrenat cu 100000 de astfel de liste.

## **Raport Lab4**

Experimental chatbotul a fost testat prin 2 metode diferite de testare. Fiind un chatbot a fost foarte greu de cuantificat sub forma de numar performanta acestuia, in mod normal performanta unui chatbot fiind data de numarul de utilizatori a acestuia si de cat de multumiti sunt utilizatori cu acesta.

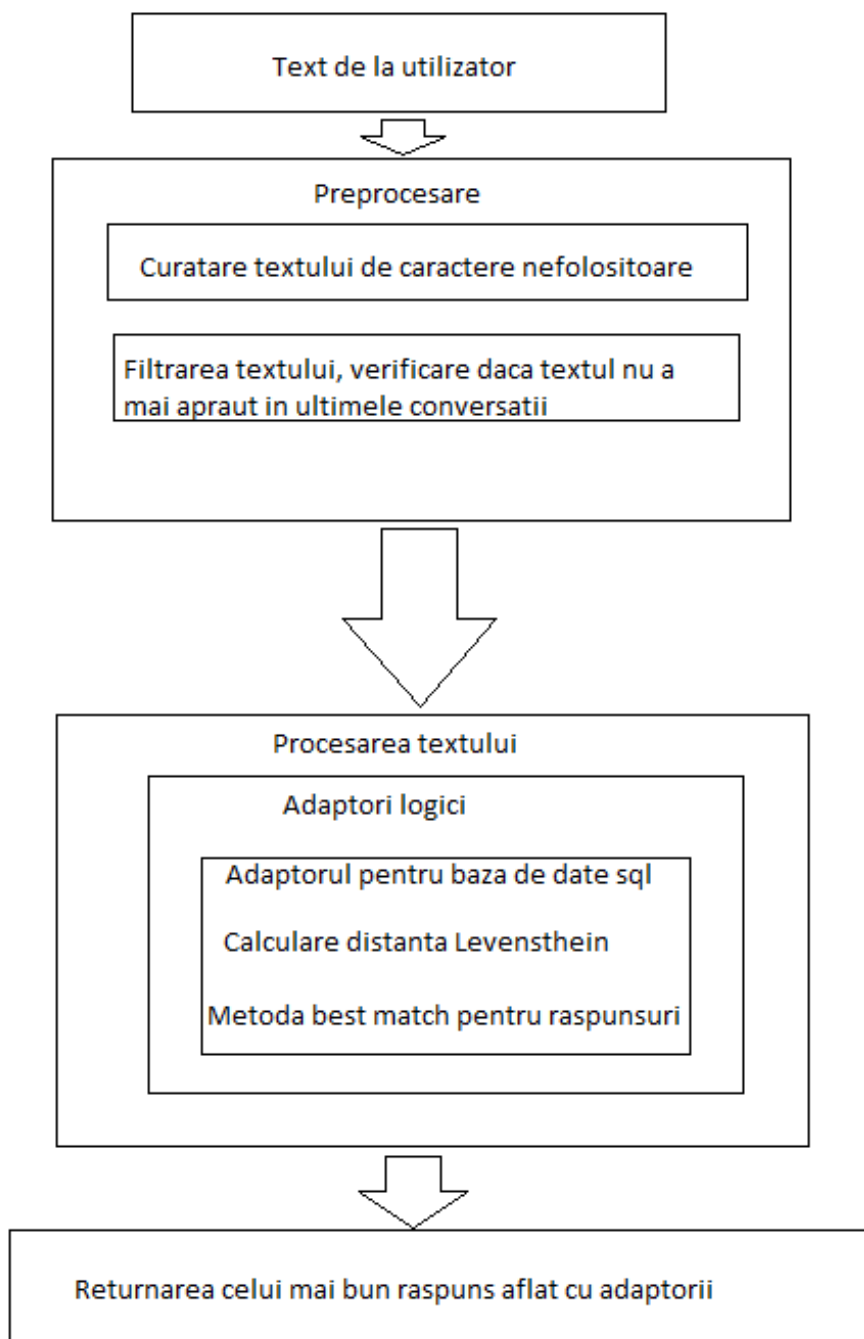
Partea aceasta de testare umana a fost realizata de echipa noastra de testeri in care testerul nostru a purtat conversatii reale cu chatbotul incercand sa rezolve diverse probleme ce pot aparea in mod uzual in cod, astfel ajungandu se la o concluzie ca chatbotul poate rezolva cateva probleme dar nu pe toate.

Cea de a doua metoda prin care s a testat performanta chatbotului a fost prin compararea similiaritai dintre raspunsurile reale la unele probleme si raspunsurile date de chatbot la aceleasi probleme, s a ales un lot de 500 de intrebari iar chatbotul a avut o similaritate de 30% a raspunsurilor lui cu cele reale.

Performanta buna pentru inceput care speram sa se imbunatateasca pe viitor.

Comparativ cu sota nu s au gasit date legat de benchmarkuri ale chatobtilor.

## Raport Lab5



## **Raport Lab6**

In acest raport se vor prezenta imbunatatirile aduse chatbotului de programare.

Una din imbunatatirile de baza aduse a fost introducerea de noi adaptorii logici, adaptorii pentru stocarea datelor dar si filtre pe inputuri.

Astfel chatbotul a primit un adaptor SQLStorage care sa il ajute sa isi salveze cat mai bine datele cu care este antrenat intr o baza de date interna sql in care are reprezentate grafuri conversationale.

Ca si filtre s a adaugat un filtru care pentru intrebari sau conversatii care se repeta in ultima vreme sa se dea acelasi raspuns, astfel usurand calculul raspunsurilor.

Adaptorii logici au suferit modificari majore, Adaptorul principal a ramas in continuare adaptorul BestMatch built in de la biblioteca chatterbot dar a fost imbunatatit alegandu se in threshold de 0.9 in raspunsuri. Functia de comparare a fost modificata si a fost astfel aleasa distanta Levenshtein ca si metoda de a determina similaritatea unei intrebari primite cu intrebarea la care e inregistratun raspuns. Metoda de selectie a fost imbunatatita la a alege primul raspuns care trece de thresholdul de 0.9 similaritate.

Totodata daca pentru o intrebare nu se gaseste nici o alta intrebare simialra in baza de date am adaugat un raspuns default pentru a gestiona cu usurinta aceste cazuri.

O alta imbunatatire semnificativa a fost antrenamentul pe care chatbotul l a primit. Acesta a fost invatat sa comunice fraze de baza in limba engleza astfel fiind posibila o conversatie mai fluenta cu acesta.

Totodata numarul de intrebari si raspunsuri din baza sa de date pe care le a invatat de pe stack overflow a crescut de la 10k la 500k.

Astfel cu aceste imbunatatiri chatbotul ca si performanta reuseste sa inteleaga cine e si ce scop are si reuseste sa raspunda la cat mai mukte intrebari astfel utilizatorii devenind din ce in ce mai satisfacuti.

