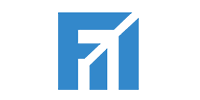
UNIVERSITATEA “ALEXANDRU IOAN CUZA” DIN IAȘI

**FACULTATEA DE INFORMATICĂ**

****

LUCRARE DE LICENȚĂ

**Învățare automată: chatbot-ul *Euterpe***

propusă de

**Alexandru Voaideș**

**Sesiunea:** *iulie, 2018*

**Coordonator științific**

**Colaborator Florin Olariu**

**UNIVERSITATEA ”ALEXANDRU IOAN CUZA” DIN IAȘI**

**FACULTATEA DE INFORMATICĂ**

**Învățare automată: chatbot-ul *Euterpe***

**Alexandru Voaideș**

**Sesiunea:** *iulie, 2018*

Coordonator științific

**Colaborator Florin Olariu**

Prin prezenta declar că Lucrarea de licenţă cu titlul „*Învățare automată: chatbot-ul Euterpe*” este scrisă de mine şi nu a mai fost prezentată niciodată la o altă facultate sau instituţie de învăţământ superior din ţară sau din străinătate. De asemenea, declar că toate sursele utilizate, inclusiv cele preluate de pe Internet, sunt indicate în lucrare, cu respectarea regulilor de evitare a plagiatului:

- toate fragmentele de text reproduse exact, chiar şi în traducere proprie din altă limbă, sunt scrise între ghilimele şi deţin referinţa precisă a sursei;

- reformularea în cuvinte proprii a textelor scrise de către alţi autori deţine referinţa precisă;

- codul sursă, imaginile etc. preluate din proiecte *open*-*source* sau alte surse sunt utilizate cu respectarea drepturilor de autor şi deţin referinţe precise;

- rezumarea ideilor altor autori precizează referinţa precisă la textul original.

Iași, *3 iulie 2018*

Absolvent *Alexandru Voaideș*

*\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_*

(semnătura în original)

Prin prezenta declar că sunt de acord ca Lucrarea de licență cu titlul „*Învățare automată: chatbot-ul Euterpe*”, codul sursă al programelor şi celelalte conţinuturi (grafice, multimedia, date de test etc.) care însoţesc această lucrare să fie utilizate în cadrul Facultăţii de Informatică.

De asemenea, sunt de acord ca Facultatea de Informatică de la Universitatea „Alexandru Ioan Cuza” din Iași, să utilizeze, modifice, reproducă şi să distribuie în scopuri necomerciale programele-calculator, format executabil şi sursă, realizate de mine în cadrul prezentei lucrări de licenţă.

Iași, *3 iulie 2018*

Absolvent *Alexandru Voaideș*

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(semnătura în original)

**Cuprins**

[Introducere 2](#_Toc517614347)

[**Capitolul 1.** Platforma de dezvoltare Android 3](#_Toc517614348)

[**Capitolul 2.** Tehnologii folosite 3](#_Toc517614350)

[**Capitolul 3.** Aplicația 3](#_Toc517614352)

[**Capitolul 4.** Componenta de inteligență artificială 3](#_Toc517614354)

[Bibliografie 3](#_Toc517614356)

# Introducere

La momentul actual, domeniul învățării automate se bucură de o mare popularitate și utilitate. Acesta a fost principalul motiv al alegerii temei de față. Un alt motiv a fost dorința de a retrezi interesul oamenilor pentru poezie, aflată astăzi într-un con de umbră. Pentru aplicație, am decis ca sistemul de operare să fie Android, având în vedere că oamenii se folosesc mai des de un smartphone pentru rezolvarea problemelor decât de un calculator personal. Această temă mi-a oferit și ocazia de a aprofunda domeniul învățării automate.

Lucrarea de față constituie documentația asupra unei aplicații pentru Android, bazată pe un chatbot românesc care vorbește în versuri. Există zeci de mii de chatbot-uri în lume, în aproximativ 60 de limbi (două în limba română) și pe o arie largă de specializări: de la îndrumător pentru gătit („Dinner Ideas”), până la specialist în tranzacții bancare („Neomy”) (Chatbots.org, 2018). Luând în seamă aceste statistici, aplicația nu reprezintă un lucru inedit. Însă, având în vedere numărul scăzut de chatbot-uri în limba română și abilitatea de a vorbi versificat, aceasta poate fi considerată o noutate.

Obiectivele principale ale lucrării sunt:

* Valorificarea oportunităților oferite de dezvoltarea domeniului învățării automate
* Cultivarea interesului pentru poezie
* Crearea unui instrument inteligent de divertisment

În realizarea chatbot-ului, am recurs la o serie de reprezentări a cuvintelor prin vectori de probabilități precalculați cu un algoritm Word2Vec. Algoritmul a fost aplicat pe corpusul CoRoLa, o mulțime de texte dintre care unele datează din 1989, „scopul creării acestuia fiind să ofere o imagine obiectivă a limbii române actuale scrise și vorbite” (Tufiș, et al., 2017). Numărul vectorilor preluați este de 250258, fiecare având dimensiunea 300. Am calculat cu datele din vectori utilizând similaritatea cosinus, pentru fiecare cuvânt, cele mai similare 10 cuvinte cu acesta, care au fost folosite ulterior la algoritmul de generare a versurilor.

Chatbot-ul este definit ca „an artificial living entity designed to have conversations with real human beings. This could encompass a text conversation via computers, a spoken conversation or even a non-verbal conversation” (Chatbots.org, 2018). Numele Euterpe dat chatbot-ului a fost preluat din mitologia greacă, fiind una dintre cele nouă fiice ale lui Zeus și considerată muza muzicii și a poeziei lirice.

În primul capitol se face o statistică asupra numărului de aplicații și al veniturilor dezvoltatorilor din acestea, o scurtă descriere a mediului de dezvoltare folosit, Android Studio, precum și o prezentare succintă a câtorva dintre cele mai populare chatbot-uri pe Android existente.

Capitolul doi descrie modalitățile prin care se poate realiza un chatbot și surprinde tehnologiile folosite pentru atingerea obiectivelor acestei lucrări.

Capitolul trei cuprinde o descriere completă asupra interfeței grafice și a funcționalităților aplicației.

Capitolul patru vorbește despre „creierul” chatbot-ului și algoritmul de generare a versurilor care se folosește de cunoștințele acestuia.

# Capitolul 1

## Platforma de dezvoltare Android

Aplicațiile se bucură la momentul actual de o popularitate extraordinară în rândul oamenilor, iar aceasta mulțumită facilităților și divertismentului pe care le oferă. Conform statisticilor, există 3.8 milioane de aplicații pe Google Play, 2 milioane pe Apple App Store, 669 de mii pe Windows Store, 430 de mii pe Amazon Appstore și 234 500 pe BlackBerry World (Statista, 2018).

Dezvoltatorii au profitat de nivelul de utilizare al aplicațiilor, care au devenit pentru ei o sursă de venit generoasă (YouTube, Facebook Messenger, Instagram ș.a.). Un dezvoltator Android câștigă anual 97.6 de mii de dolari, iar un dezvoltator mid-level iOS, nu cu o diferență prea mare, 96.6 mii de dolari (Charuza, 2018).

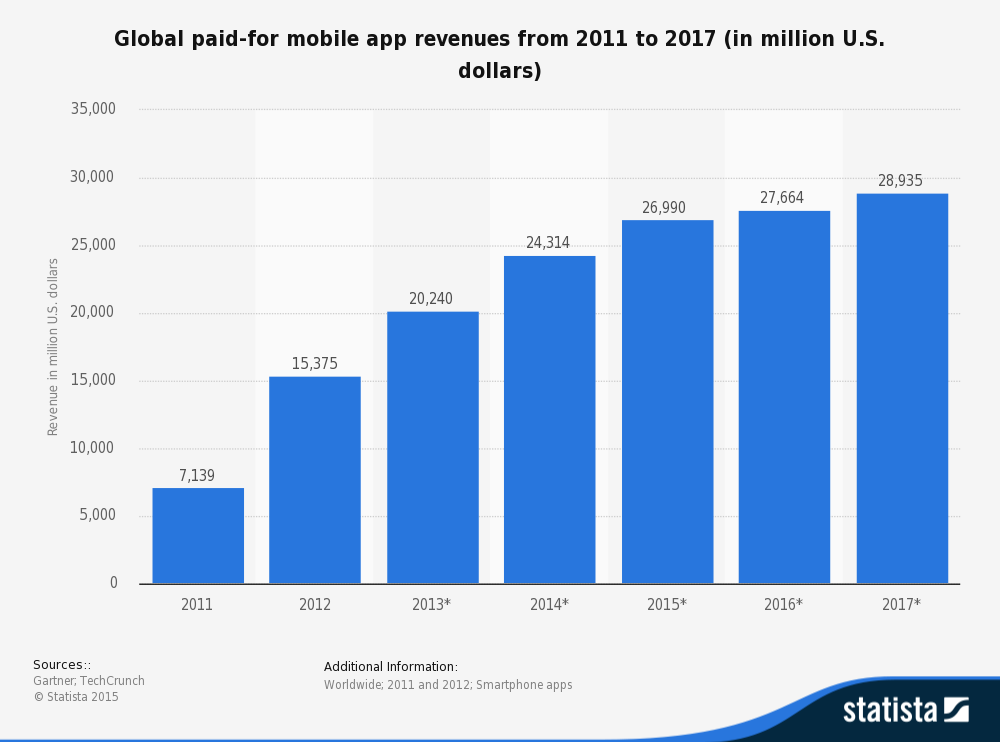


Fig.1 Creșterea veniturilor din aplicații la nivel global din anul 2011 până în 2017

În iunie 2014, Google lansează Android Studio, un IDE gratuit pentru dezvoltarea de aplicații pe sistemul de operare Android. IDE-ul vine la pachet cu un SDK și un debugger și prezintă următoarele instrumente:

* Un sistem de dezvoltare flexibil bazat pe Gradle;
* Un emulator rapid cu multe opțiuni;
* Un mediu de dezvoltare unic pentru toate dispozitivele cu Android;
* Instant Run ce permite adăugarea noilor modificări aduse aplicației fără necesitatea instalării unui nou APK;
* Șabloane de cod și integrare cu GitHub pentru a facilita construirea de caracteristici comune ale aplicațiilor și importarea codului simplu;
* Instrumente de testare extensive și framework-uri;
* Instrumente Lint pentru detecția performanței, uzabilității, versiunii compatibile și a altor probleme;
* Suport pentru C++ și NDK;
* Suport încorporat pentru Google Cloud Platform, făcând ușoară integrarea pentru Google Cloud Messaging și App Engine;

(AndroidDevelopers, 2014)

## Chatbot-uri pe Android

Așa cum s-a menționat în introducere, numărul chatbot-urilor a crescut foarte mult odată cu dezvoltarea domeniului inteligenței artificiale și al învățării automate, având o aplicabilitate foarte diversă: legislație, finanțe, sport, educație etc. Aceștia se specializează pe oferirea unei experiențe conversaționale autentice, în care interlocutorul uman nu ar trebui să poată afirma dacă poartă o discuție cu un om sau cu o mașină. Despre un astfel de chatbot se spune că a trecut testul Turing. Există companii care folosesc chatbot-uri pentru a furniza servicii mai bune clienților, dintre care menționez 10.

### Endurance: A Companion for Dementia Patients

# Capitolul 2

## Tehnologii folosite

# Capitolul 3

## Aplicația

# Capitolul 4

## Componenta de inteligență artificială

# Bibliografie

Charuza, P., 2018. *Much money can earn app.* [Interactiv]   
Available at: https://fueled.com/blog/much-money-can-earn-app/  
[Accesat 29 May 2018].

Chatbots.org, 2018. *About Us: Chatbots.org - Virtual assistants, virtual agents, chat bots.* [Interactiv]   
Available at: https://www.chatbots.org/community/static/about\_us  
[Accesat 1 5 2018].

Chatbots.org, 2018. *Chatbots.org - Virtual assistants, virtual agents, chat bots.* [Online]   
Available at: https://www.chatbots.org  
[Accessed 1 5 2018].

Developers, A., 2014. *Android Developers: Intro.* [Interactiv]   
Available at: https://developer.android.com/studio/intro/  
[Accesat 30 May 2018].

Gibson, A., Nicholson, C. & Patterson, J., 2017. *Deeplearning4j: Open-source, Distributed Deep Learning for the JVM.* [Interactiv]   
Available at: https://deeplearning4j.org/overview  
[Accesat 20 5 2018].

Statista, 2018. *Number of apps available in leading app stores: Statista.* [Interactiv]   
Available at: https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/  
[Accesat 29 May 2018].

Tufiș, D. și alții, 2017. *CoRoLa.* [Interactiv]   
Available at: http://corola.racai.ro/  
[Accesat 29 May 2018].