



**Universitatea
Transilvania
din Brașov**

**FACULTATEA DE MATEMATICĂ
ȘI INFORMATICĂ**

Lucrare Disertație

Autor: Hanganu Bogdan

Coordonator: Lect. Univ. Băicoianu Alexandra

Brașov
Iulie 2022



Universitatea
Transilvania
din Brașov

FACULTATEA DE MATEMATICĂ
ȘI INFORMATICĂ

Lucrare Disertație

Detectarea emoțiilor prin canale multiple

Autor: Hanganu Bogdan

Coordonator: Lect. Univ. Băicoianu Alexandra

Brașov
Iulie 2022

Cuprins

1	Abstract	1
2	Introducere	2
2.1	Interacțiunea om-mașină	2
2.2	3

1 Abstract

Această lucrare de disertație este elaborată în jurul problematicii identificării emoțiilor unei persoane. Utilizând metode de învățare automată, datele sunt înregistrate prin intermediul unor canale multiple (audio, video și text). Expresiile faciale, obținute prin intermediul canalului video, reflectă în mod intuitiv starea mentală a unei persoane, fiind una dintre cele mai bogate și importante forme de comunicare inter-umană. Tonalitatea vocii care se adresează în timpul comunicării ne poate oferi informații valoroase referitoare la starea de spirit. Mesajul care este transmis prin intermediul vocii, ne oferă informații referitoare la personalitatea individului, fiind de ajutor mai departe în procesul de analiză. Datele stocate sunt mai apoi prelucrate folosind tehnici de procesare specifice fiecărui canal. Pentru input-ul audio este folosită procesarea digitală de semnal, cu următoarele tehnici reprezentative: Transformata Fourier, Transformata Fourier pe termen scurt, Coeficienți Mels. Procesare de imagini pentru canalul video vine în adăugare cu: scalare de date, transformare imaginii în greyscale, iar pentru text este de menționat: tokenizare, lematizare. Fiecare bloc de date preprocesat în mod corespunzător canalului părinte, va trece mai departe prin pasul de recunoaștere cu ajutorul metodelor de învățare automată. În această lucrare au fost realizate o serie de experimente pentru: procesarea datelor, antrenarea modelelor de machine learning. Finalizarea acestor teste a avut ca urmare dezvoltarea aplicației "Multimodal Emotion Detection" care să vină în sprijinul procesului de interviuare.

2 Introducere

2.1 Interacțiunea om-mașină

Aplicația "Multimodal Emotion Detection" are ca audiență persoanele care doresc să facă o analiză a candidatului care a trecut printr-un proces de interviu. Fiind scrisă în limbajul de programare Python, permite o manevră concisă a datelor înregistrate, care pot fi mai apoi vizualizate de către utilizator prin intermediul framework-ului GUI (Graphical User Interface) QT (pronunțat "cute").

În cadrul lucrării, se propune recunoașterea emoțiilor utilizatorului într-un mod inteligent, utilizând tehnici și metode de Machine Learning și Deep Learning. Aceste două procedee sunt subcategorii ale domeniului numit inteligență artificială, domeniu care a început să se modeleze în funcție de nevoile oamenilor.

Dezvoltarea rapidă a inteligenței artificiale, "Big data science" și a tehnologiei "Block chain" a provocat multiple schimbări în structura socială umană. În majoritatea proceselor unde este nevoie de interacțiune umană, se implementează automatizări care să sporească eficiența, să folosească resursele umane, software și hardware în mod cât mai eficace. La nivel industrial, sistemele automatizate inteligente sunt deja folosite în uzine, fabrici, având rolul de a asigura în permanență buna funcționare a întregului ansamblu. Atât eficiența cât și performanțele acestor sisteme sunt motivate de către costul redus de mentenanță. La un nivel mai aproape de către utilizatori, putem realiza că inteligența artificială a început tot mai des să facă parte din viața de zi cu zi, ajungând în stadiul să devină indispensabil oamenilor.

În zilele noastre, interacțiunea dintre oameni și I.A. este în continuă creștere, ajungând să intre treptat în viața noastră de zi cu zi. De la asistenți virtuali (care au rolul de a sprijini utilizatorul prin intermediul interpretării comenzilor vocale), până la reclame personalizate, aceste sisteme inteligente interacționează din ce în ce mai mult cu ființe umane. Deoarece este un subiect în care interesul este unul foarte crescut, relația între om și mașinăria inteligentă poate să ajungă la un nivel mai înalt, prin integrarea cu emoțiile utilizatorului. Acesta este un domeniu crucial de cercetare, oferind diverse oportunități și aplicații pentru oameni.

Un aspect cheie în această sferă care să sporească parteneriatul între om și mașină este dat de înțelegerea emoțiilor umane. Emoția este un factor important atât în comunicarea verbală, cât și în comunicarea nonverbală (gesticulare, expresiile corpului). Identificarea stărilor unei persoane poate ajuta o mașină să înțeleagă intențiile utilizatorului, în scopul de a-i oferi o interacțiune mai potrivită. Printre primele studii care au fost făcute pentru integrarea sistemelor inteligente cu emoțiile umane, acestea s-au reflectat în identificarea emoțiilor prin voce, deoarece comunicarea verbală este una dintre cele mai rapide forme de socializare, cu cel mai mare impact în istorie.

Fiind una dintre cele mai consacrate aptitudini prin care o ființă inteligentă s-a putut diferenția și avansa în lanțul trofic, comunicarea verbală reprezintă un semnal complex, în care sunt transmise informații referitoare la mesaj, legate de emițător precum și de emoțiile transmise de acesta. Fiind o serie complexă, capacitatea sistemului care face identificarea emoțiilor trebuie să fie pe măsură, pentru a analiza cu acuratețe starea subiectului, oferindu-i astfel o experiență cât se poate de autentică. Nu numai atât, utilizând o astfel de recunoaștere, poate ajuta la crearea unor interfețe ușor navigabile ("Recunoașterea emoțiilor prin vorbire este deosebit de utilă pentru aplicațiile din domeniul interacțiunii om-mașină deoarece ajută la crearea unor interfețe ușor de utilizat"[1]).

Deoarece semnalul audio conține și alte informații precum emoția transmisă de către emițător, algoritmul de recunoaștere a emoțiilor umane prin extragerea trăsăturilor acustice capturate în vorbire, devine "baza pentru realizarea unei interacțiuni om-calculator mai armonioasă și mai eficientă, având o mare importanță în cercetare, precum și aplicativă"[2].

În cadrul aplicației, vocea utilizatorului (semnalul audio) este folosită în două contexte. Primul este dat de către identificarea emoțiilor candidatului pe baza a diferite trăsături acustice depistate din voce iar al doilea de către identificarea cuvintelor rostite, pentru a putea fi transformate în text. Deoarece în aplicație identificarea textului (speech-to-text) este utilizat prin intermediul unor apeluri de tip API Rest la funcționalitățile oferite de către Google, accentul se va pune pe prima utilizare în lucrare.

Seviciului oferit de către Google pentru recunoașterea textului este folosit cu scopul de a permite o captare a dialogului avut

2.2

Bibliography

- [1] K. Vimaladeví C. Vinola. A Survey on Human Emotion Recognition Approaches, Databases and Applications. Accessed: 16.05.2022 (cit. on p. 3).
- [2] Giannoukos I. Anagnostopoulos C N Iliou T. Features and classifiers for emotion recognition from speech: a survey from 2000 to 2011. Artificial Intelligence Review, 2015, 43(2):155-177. (in NETHERLANDS). Accessed: 16.05.2022 (cit. on p. 3).
- [3] Alexandra BAICOIANU Bogdan HANGANU Lucian Andrei RADU. Machine Learning For Condition Monitoring: Latest Trend And Review. [https : / / ibima . org/accepted-paper/machine-learning-for-condition-monitoring-latest-trend-and-review/](https://ibima.org/accepted-paper/machine-learning-for-condition-monitoring-latest-trend-and-review/). Accessed: 20.05.2020.

	Titlu trimitere	ID lucrare Turnitin	Trimis	Similaritate	
Vizualizare confirmare digitală	Utilizarea tehnicilor de procesare de semnal în domeniul învățării automate	1258270301	18/06/2020 22:44	2% 	Trimitre lucrare   --

Figura 1: Licență similaritate