

SOLUȚIE MACHINE LEARNING PENTRU RECUȚNOAȘTEREA EMOȚIEI ÎN VORBIRE

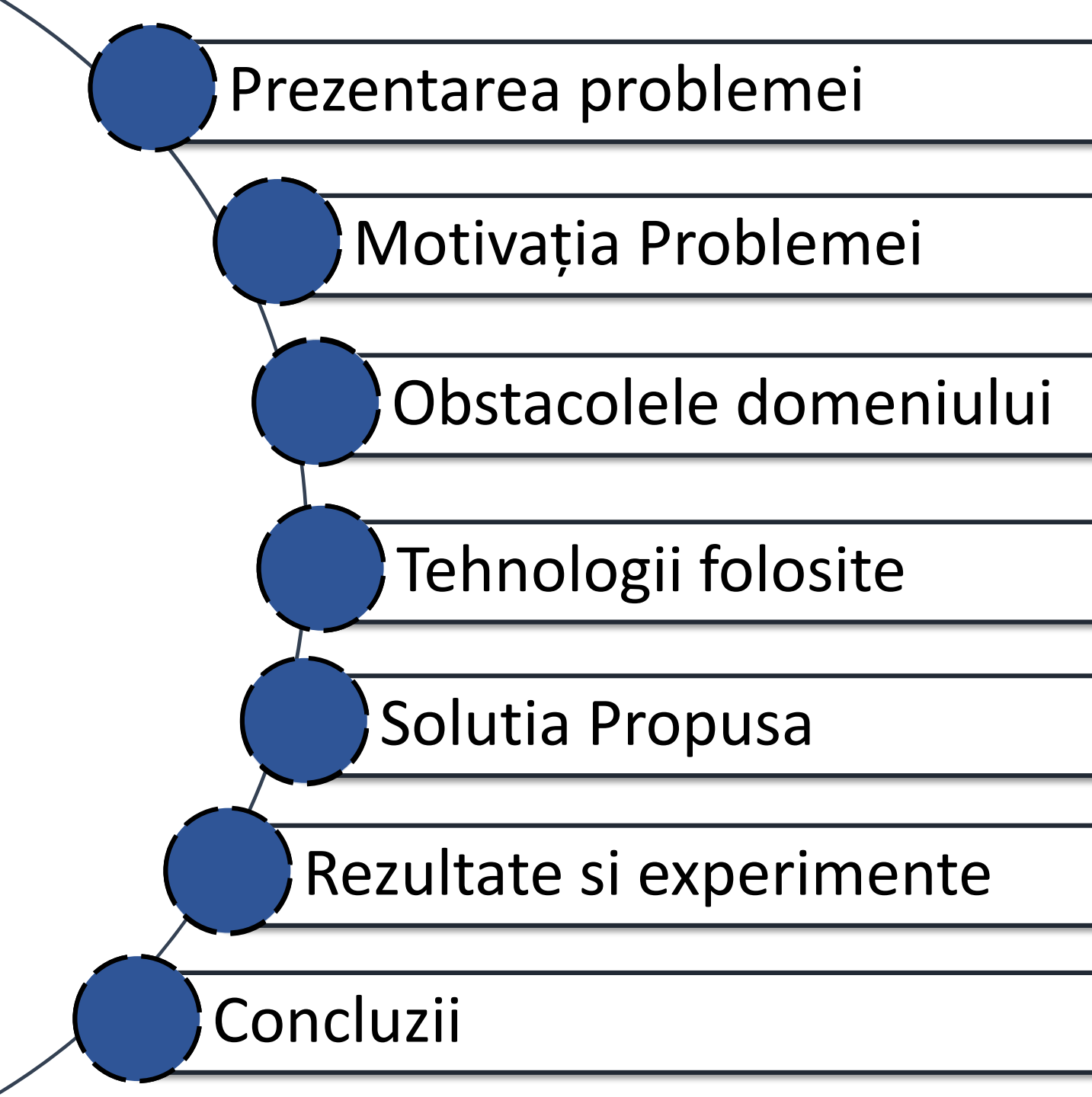
Coordonator

Prof. dr. Ing. Ștefan HOLBAN

Candidat

Steleac Raul-Dacian

IUNIE 2020

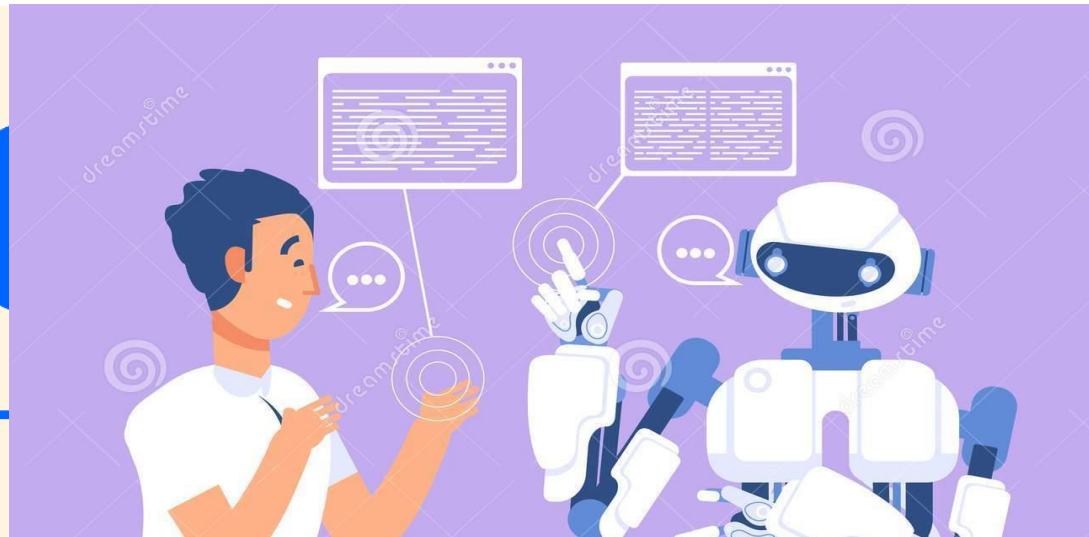


Recunoasterea emotiei in vorbire



Om - Om

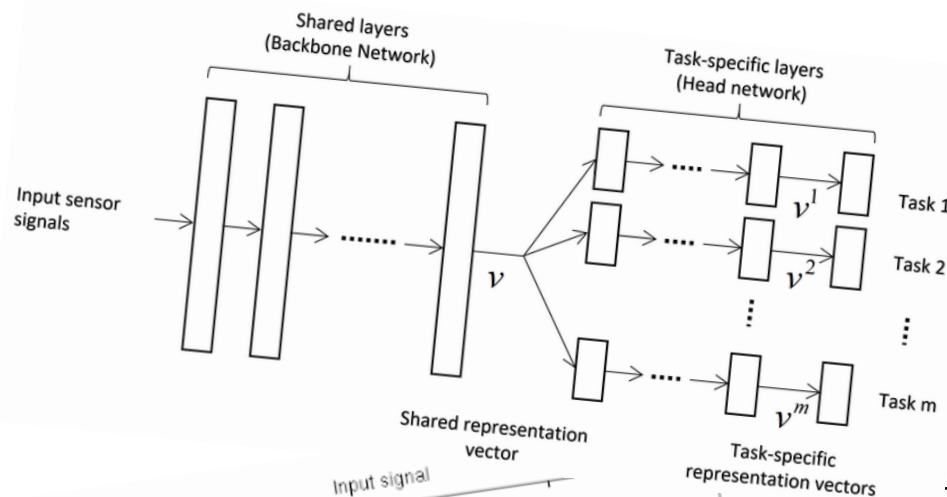
Informație lingvistică
Vs
Informație emoțională



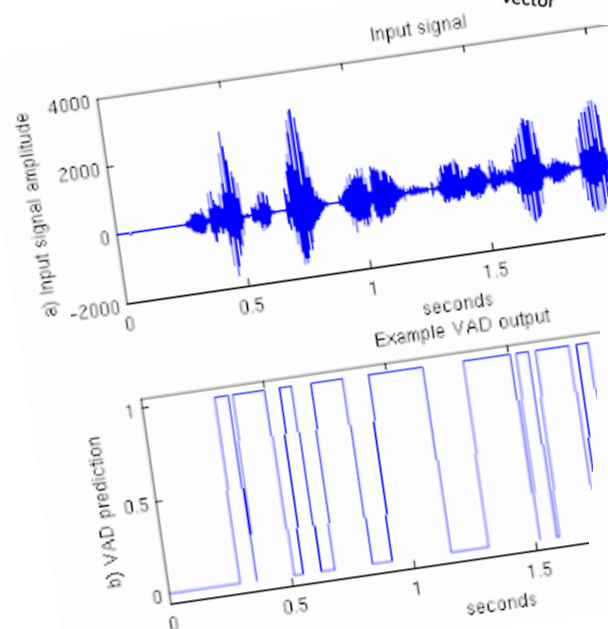
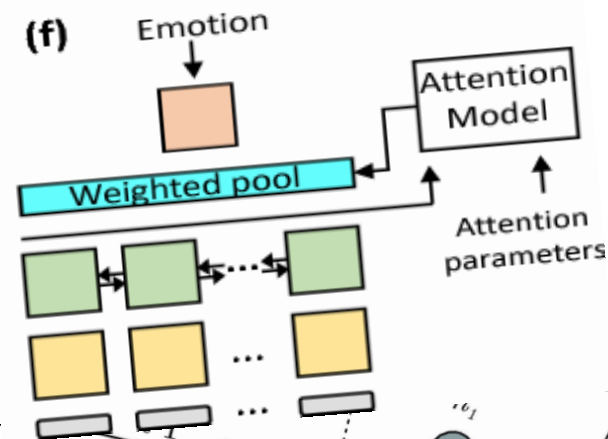
Om - Mașină

Noua generație de vorbitori

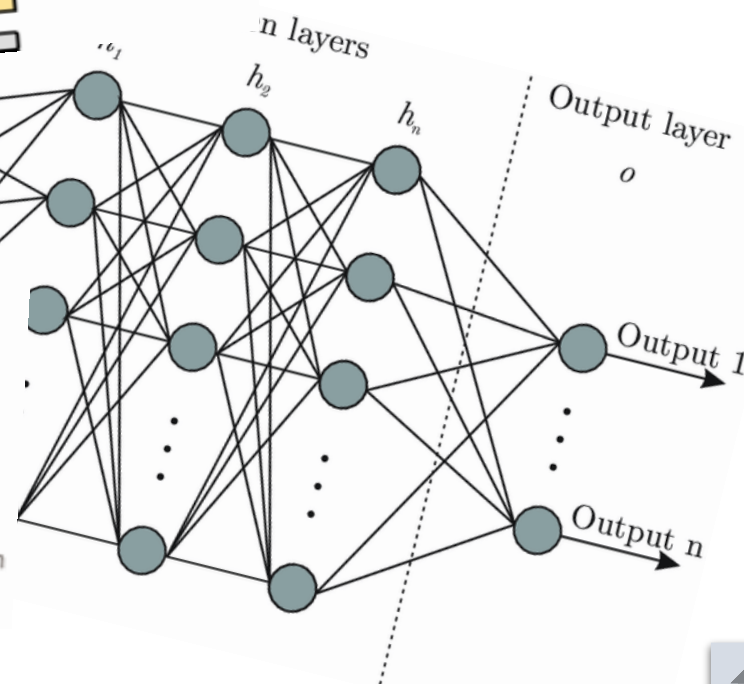
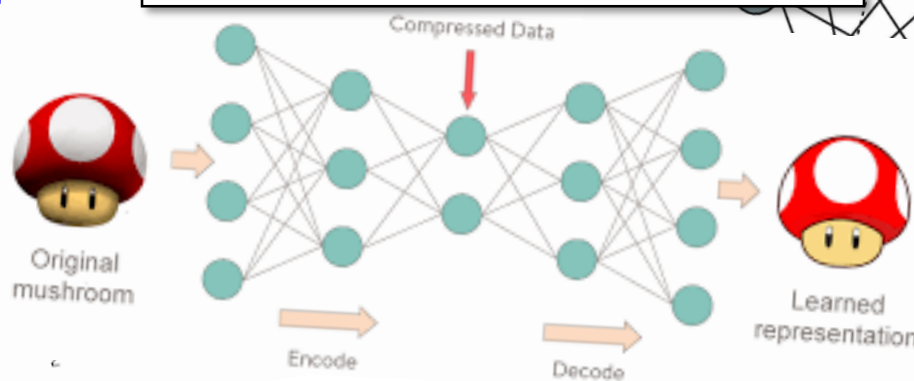
Domeniul recunoasterii emotiei in vorbire



1996 [1]
likelihood
neighbors



Încă nu sunt viabili pentru piață

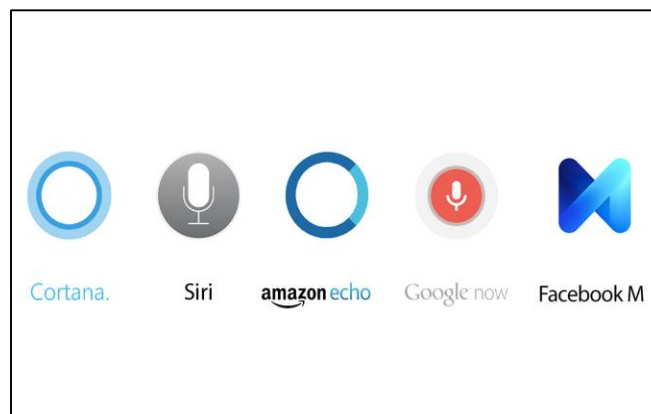


Motivații aplicative

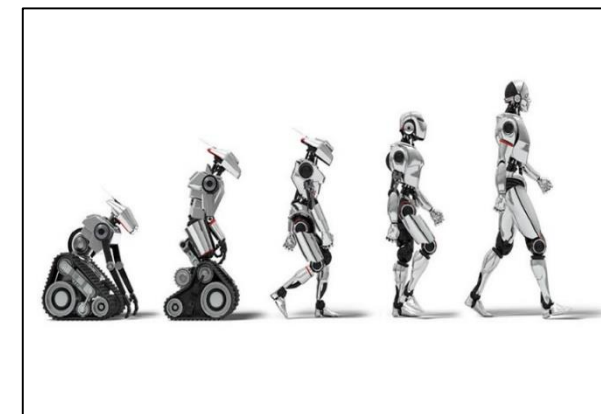
Aplicații care ar beneficia prin introducerea unui mecanism de recunoaștere a emoției în vorbire:



Mecanism de Feed-back



Asistenți inteligenți



Noua generație de roboți



Educație



Stații de call-center



Jocuri video



Inconsistența bazelor de date

- Numărul redus de exemple de antrenare din bazele de date.
- Incertitudinea din etichetarea emoțiilor bazelor de date.
- Înregistrări jucate.

Caracteristici de intrare nereprezentative

- Lipsa unui set de caracteristici de intrare reprezentative.
- “Hand-crafted”, necesitatea unor experți din domeniul audio.
- “End-to-end”, imposibilitatea determinării caracteristicilor obișnuite, “black-box”.



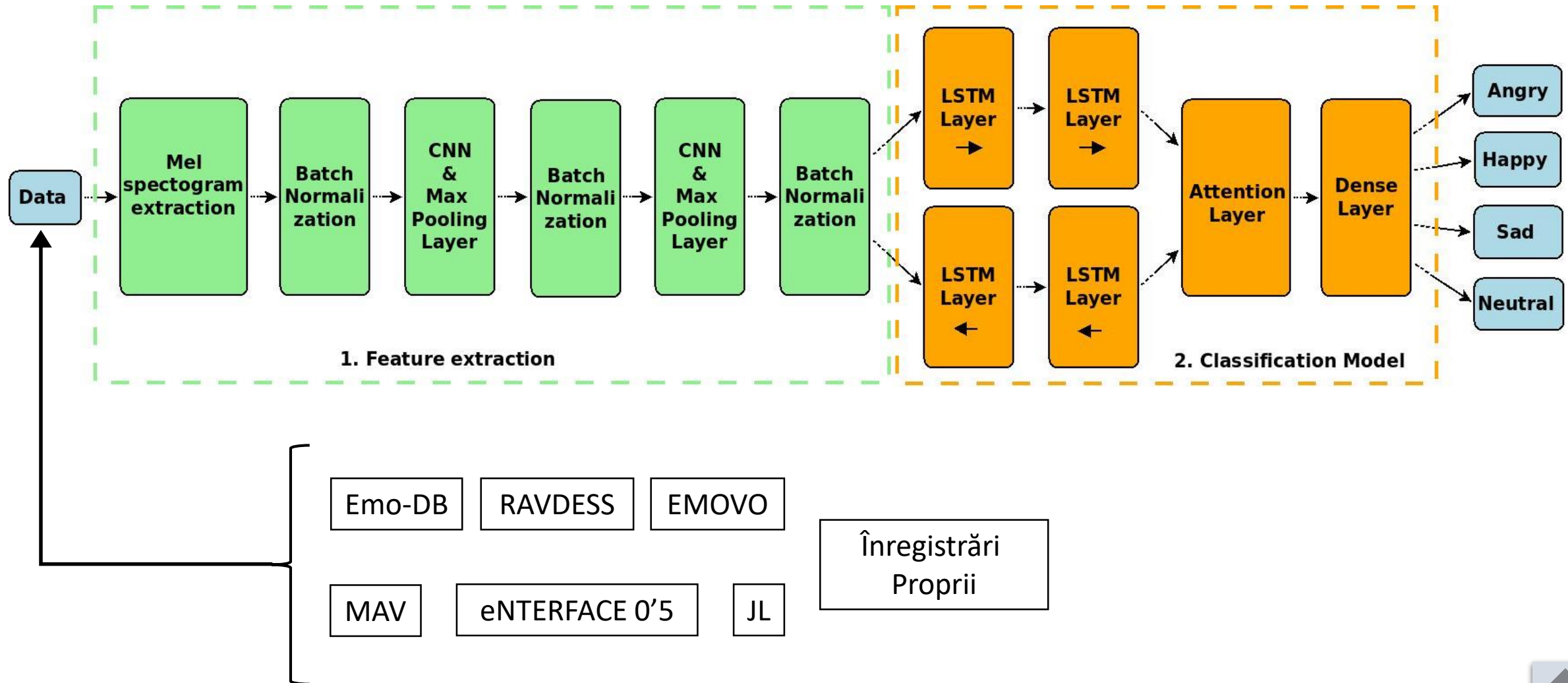
Tehnologii folosite



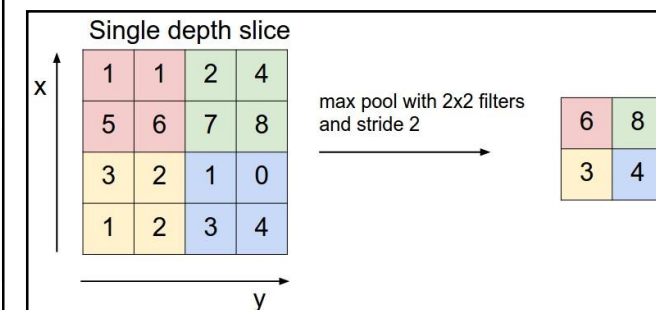
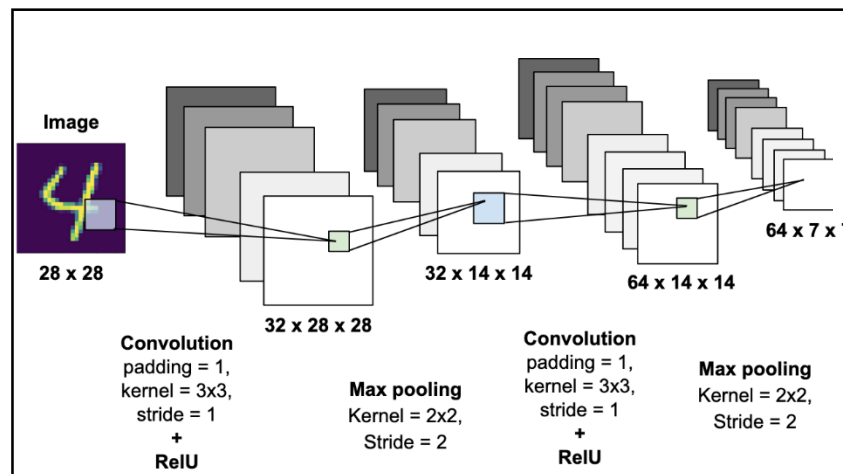
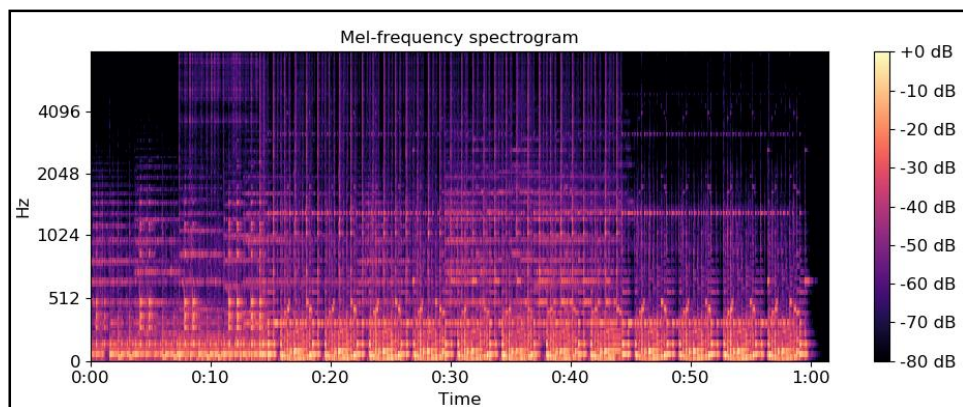
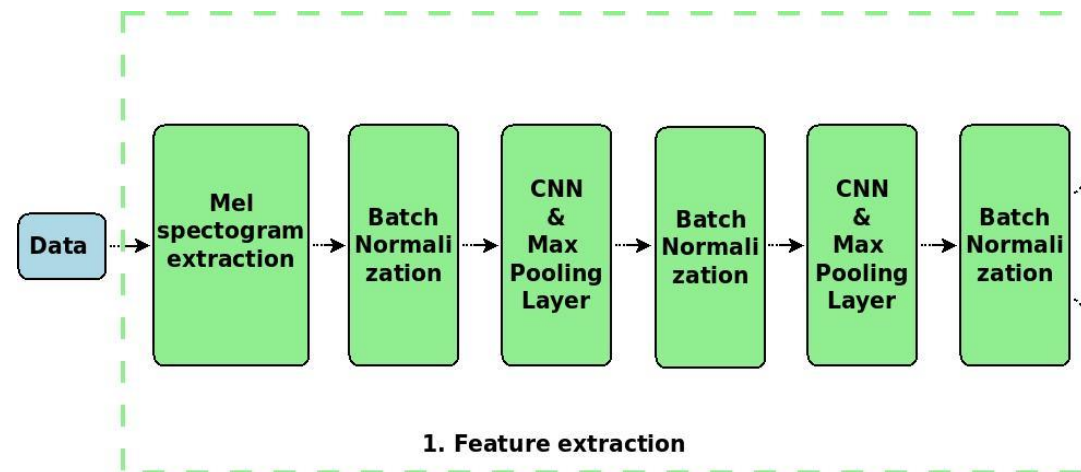
TensorFlow



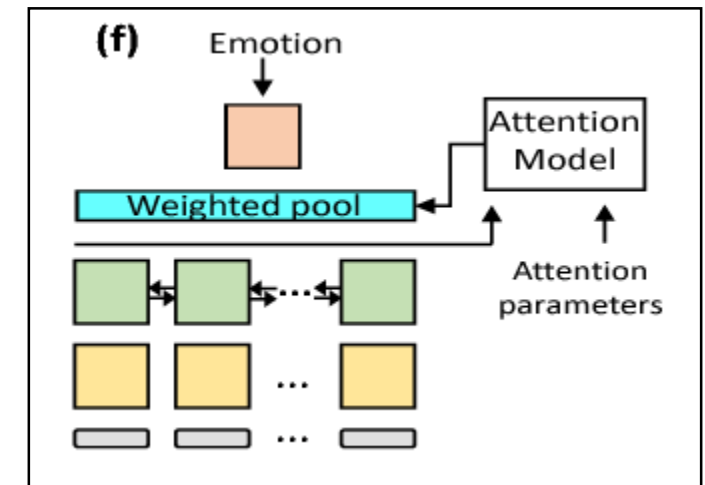
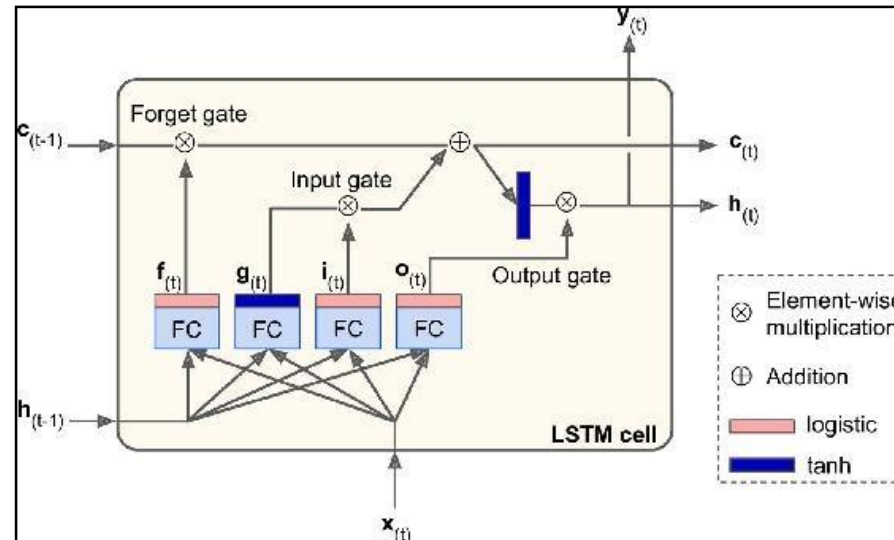
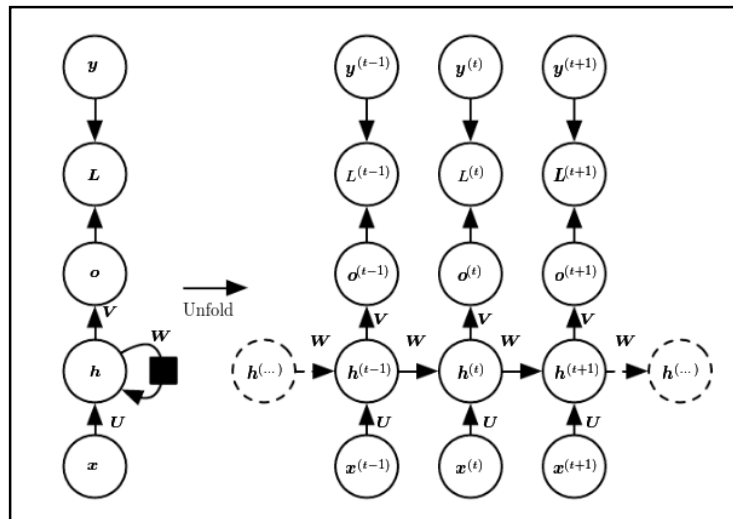
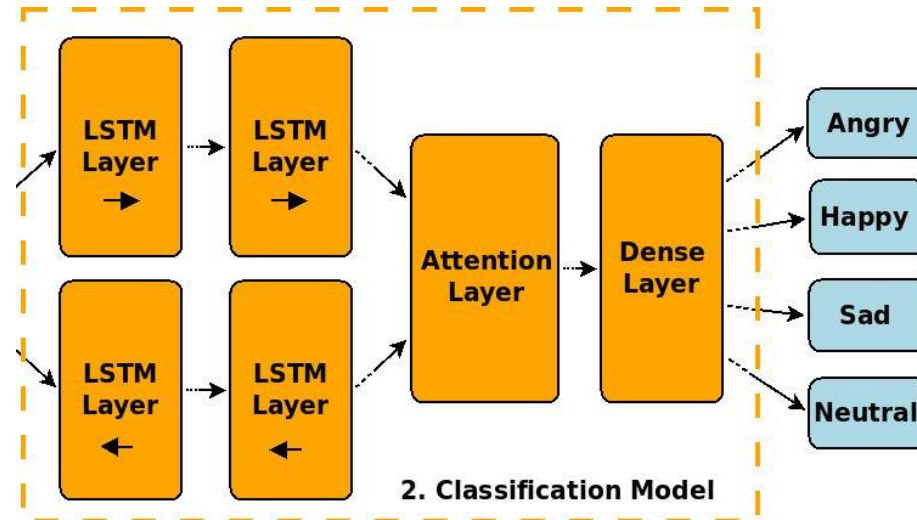
Soluția propusă



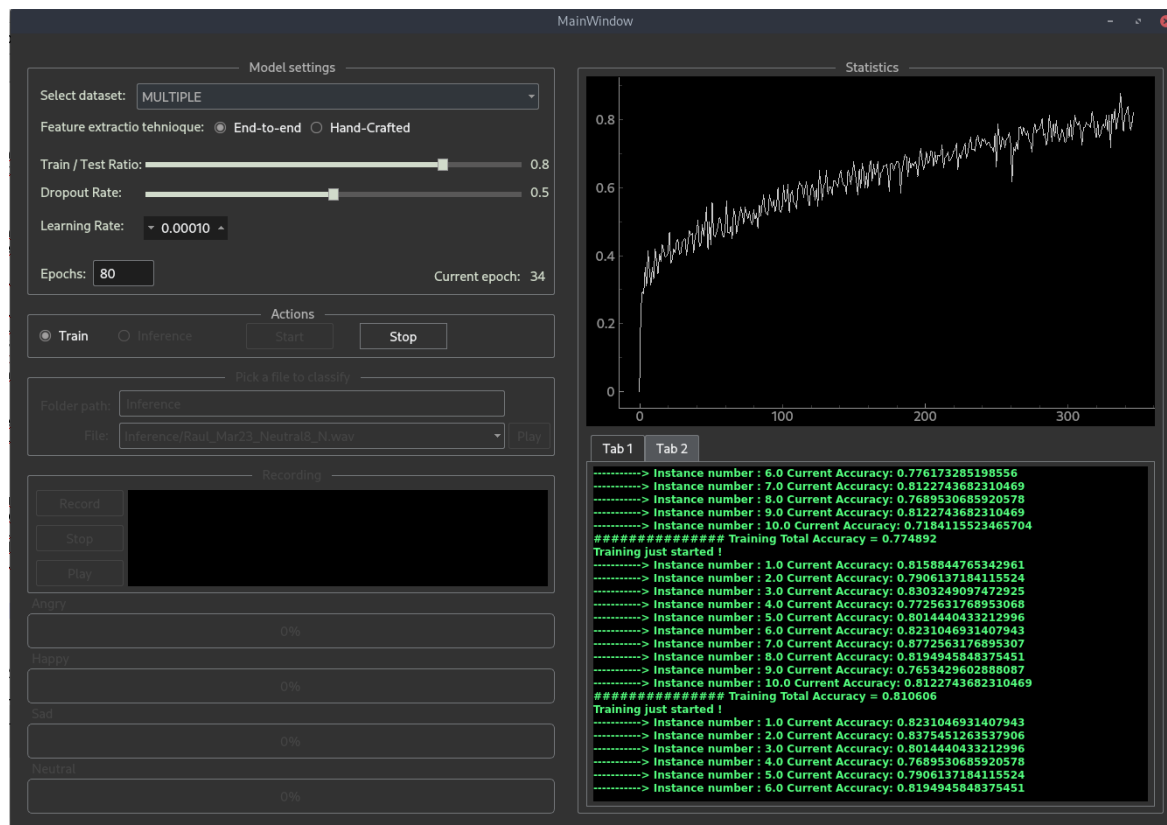
Extragerea caracteristicilor de intrare



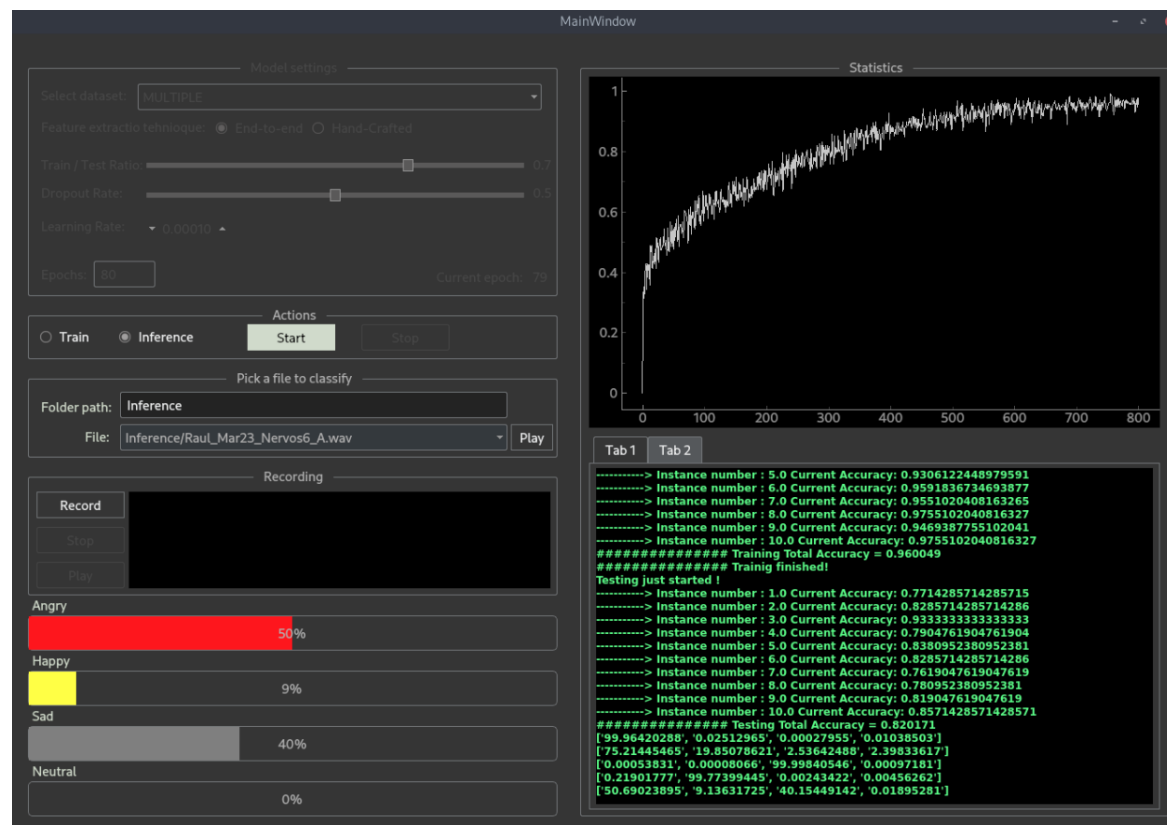
Modelul calificador



Interfața grafică



Interfața grafică în modul de antrenare

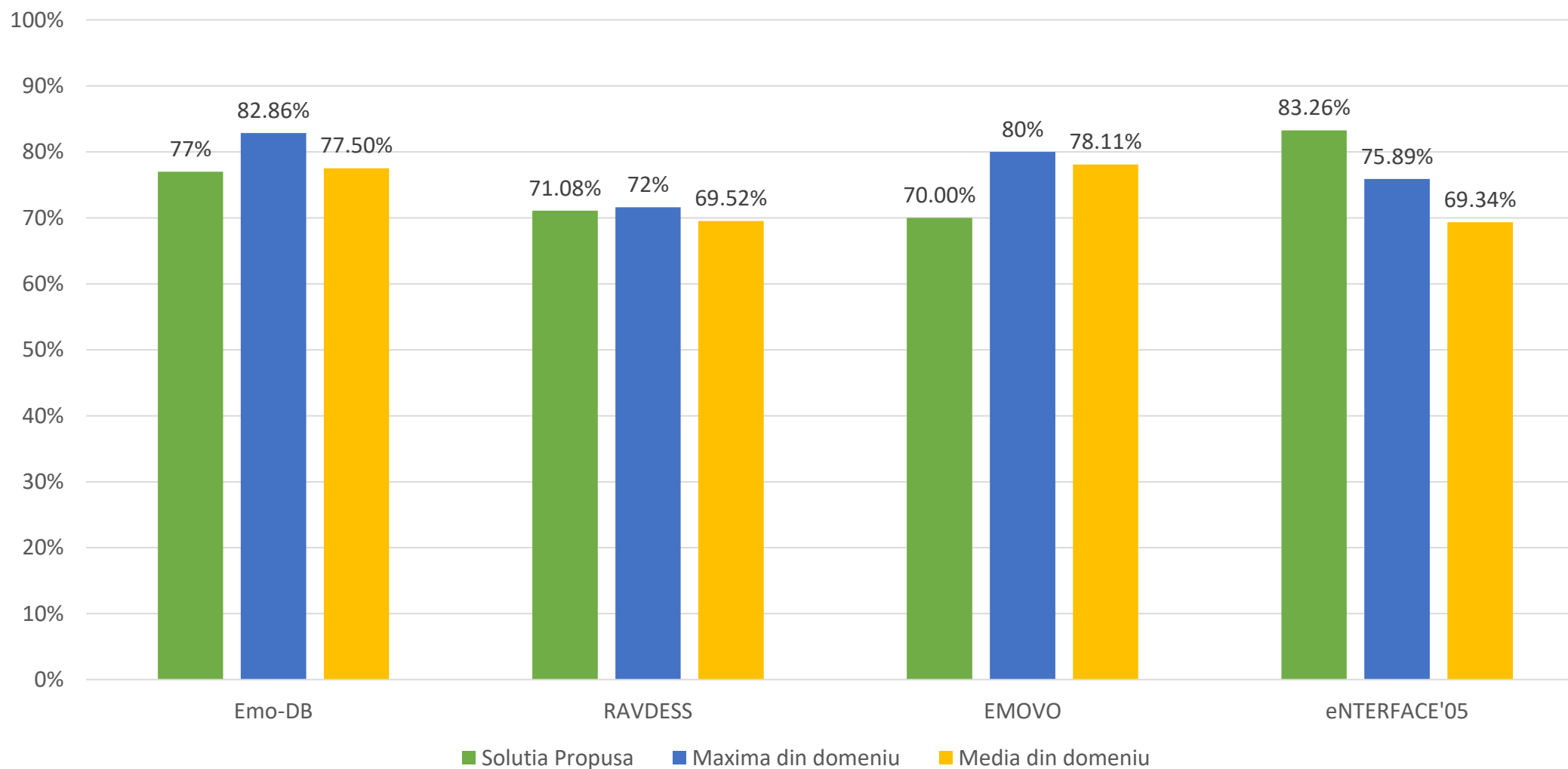


Interfața grafică în modul de inferență



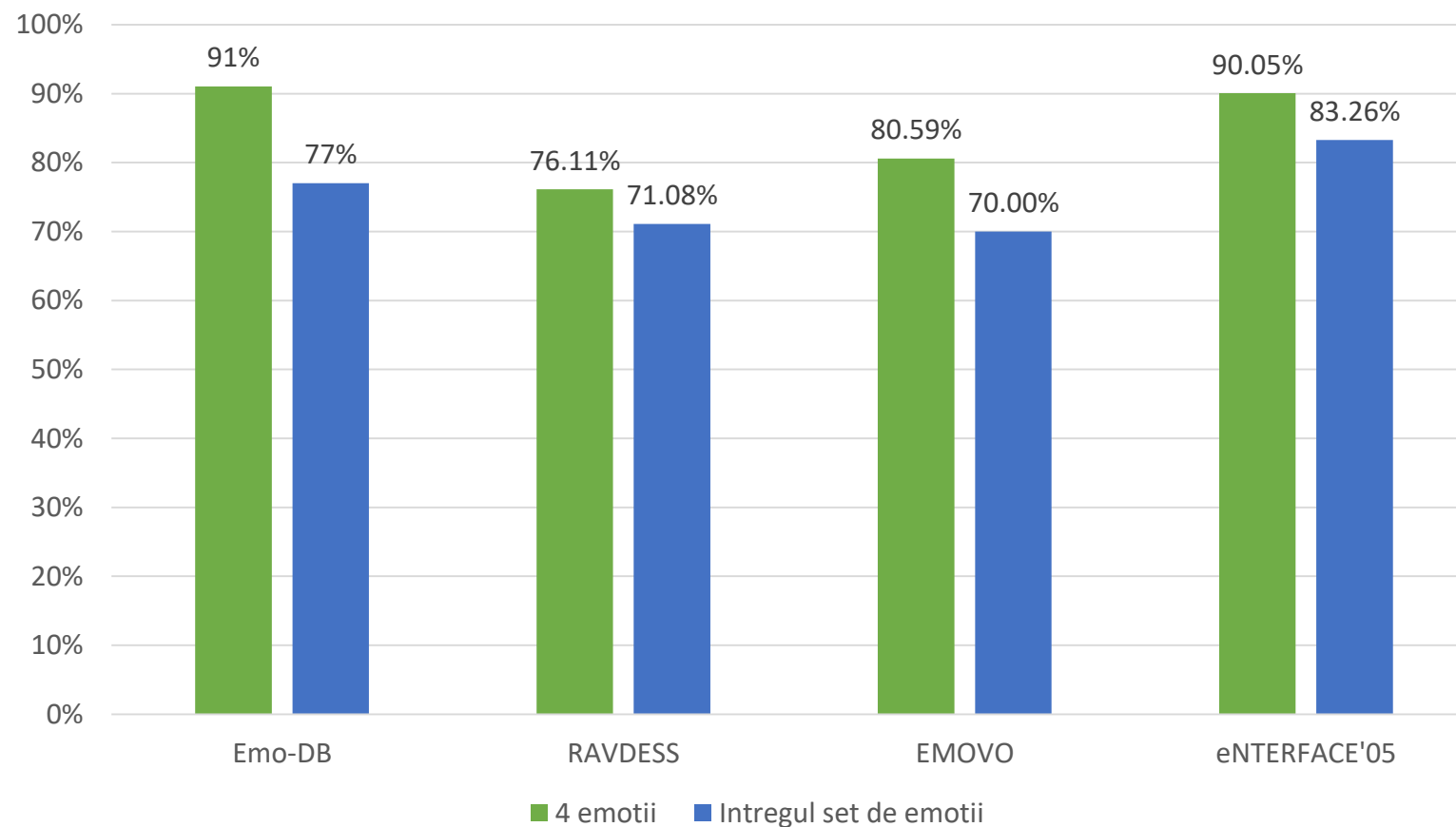
Rezultate si experimente

Rezultatele individuale pe bazele de date



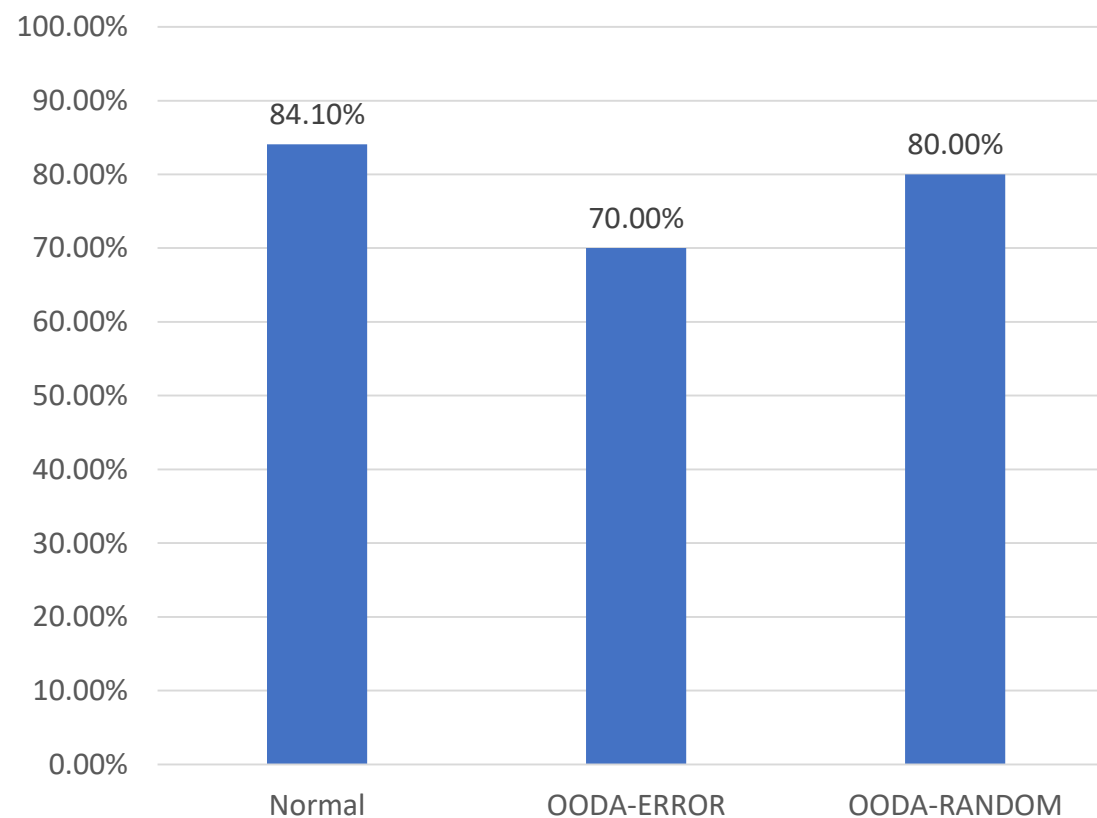
Rezultate si experimente

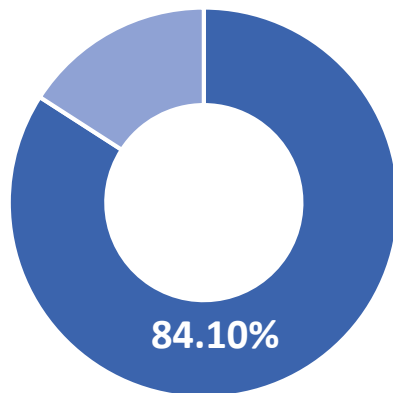
Rezultatele individuale pe bazele de date – 4 emoții



Rezultate si experimente

Rezultatele obținute pe intregul set de baze
de ate





Concluzii

- Rezultate asemanatoare, sau chiar mai bune, in cazurile individuale.
- Acuratetea modelului antrenat pe mai multe baze de date incurajatoare, dar inca nu destul de mare pentru a permite plasarea algoritmului pe piata.

Posibile îmbunătățiri

- Mărirea numărului de baze de date.
- Introducerea unui modul care să diminueze diferențele dintre înregistrările provenite din baze de date diferite.
- Antrenarea "multi-task".
- Combinarea soluției propuse cu un algoritm de detecție a emoției vizuale.

