

Music Emotion Classification

– LSTM i Transformer –

Julijana Jevtić Jelena Milošević

Matematički fakultet, Univerzitet u Beogradu

Septembar, 2025

1 Uvod

2 Modeli

- Long Short-Term Memory (LSTM)
- Transformer

3 Rezultati primene modela

- Loss and Accuracy
- Metrički rezultati modela
- F1 skor po klasama
- Normalizovana matrica konfuzije
- ROC krive po klasama
- Precision-recall krive

4 Zaključak

- Klasifikacija emocija iz muzičkih tekstova
- Cilj – prepoznavanje emocija (sreća, tuga, ljubav, ljutnja)
- Primena u muzičkim preporukama i personalizaciji
- Zadatak podrazumeva **analizu teksta** i **modelovanje sekvenci**

LSTM - Long Short-Term Memory

Sekvencijalni model za dugoročne zavisnosti.

Transformer

Model zasnovan na pažnji (attention), efikasniji u NLP zadacima.

Long Short-Term Memory (LSTM)

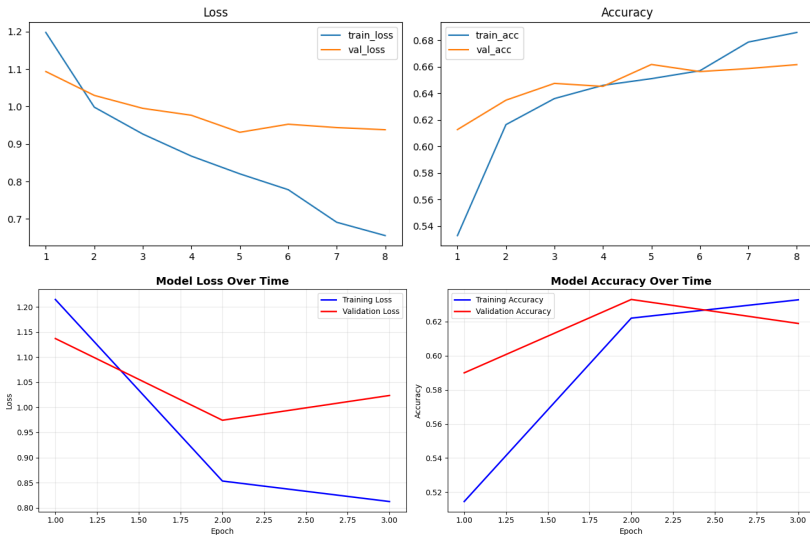
- Vrsta RNN-a koja rešava problem **nestajanja gradijenata**.
- Koristi **memorijske ćelije** i **gate mehanizme** (input, forget, output) za kontrolu informacija.
- Obrada sekvenci:
 - 1 Zaboravljanje irelevantnih podataka.
 - 2 Dodavanje novih informacija.
 - 3 Ažuriranje i prosleđivanje značajnih podataka.
- **Primene:** NLP (analiza, prevođenje, generisanje teksta), prepoznavanje govora, vremenske serije, biomedicina, računarska vizija, robotika, optimizacija resursa...
- **Prednosti:** stabilno učenje, dobar za dugoročne zavisnosti.
- **Ograničenja:** složeniji od RNN-a, osetljiv na inicijalizaciju.

Transformer

- Arhitektura zasnovana na **mehanizmu pažnje** (self-attention).
- Obrada celih sekvenci **paralelno**, za razliku od RNN-a.
- Ključne komponente:
 - ① **Self-attention** za hvatanje odnosa između svih reči.
 - ② **Multi-head pažnja** za učenje različitih tipova zavisnosti.
 - ③ **Pozicionalno kodiranje** za očuvanje redosleda.
 - ④ **Feedforward mreže i rezidualne veze**.
- **Primene:** mašinsko prevođenje, sažimanje teksta, generisanje teksta, odgovaranje na pitanja, obrada slika i govora.
- **Prednosti:** skalabilnost, hvatanje dugoročnih zavisnosti, osnova savremenih modela (BERT, GPT, T5...).
- **Ograničenja:** visoka složenost, velika potrošnja memorije i resursa.

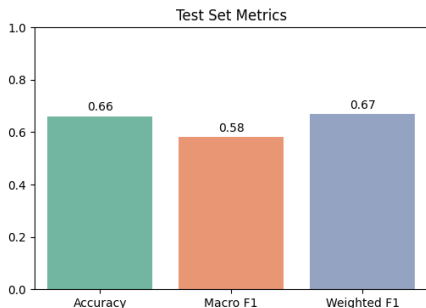
- U ovom poglavlju prikazani su rezultati evaluacije LSTM modela nad test skupom podataka.

Tok funkcije gubitka i tačnosti tokom epoha

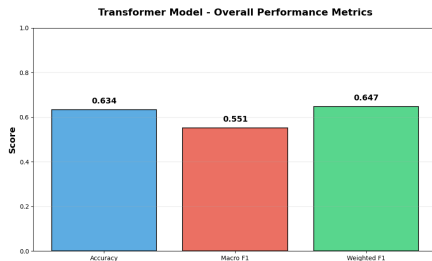


Slika: LSTM krive (gore) i Transformer krive (dole)

Metrički rezultati modela



Slika: LSTM metrike

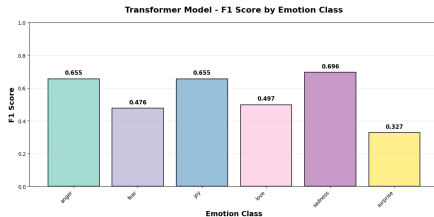


Slika: Transformer metrike

F1 skor po klasama

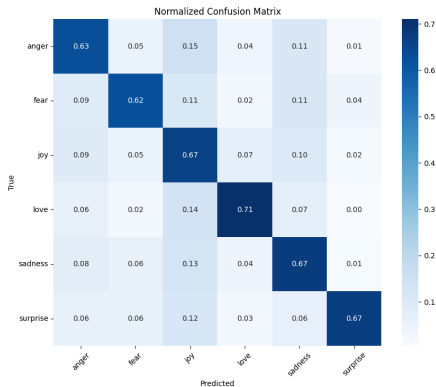


Slika: LSTM F1 Score

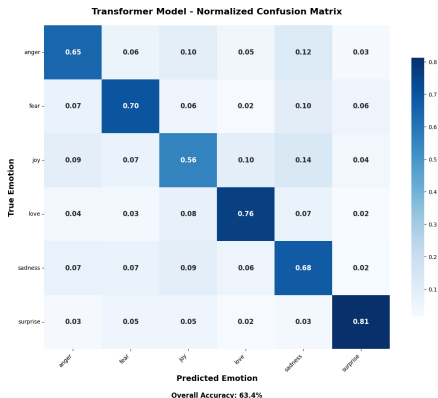


Slika: Transformer F1 Score

Normalizovana matrica konfuzije



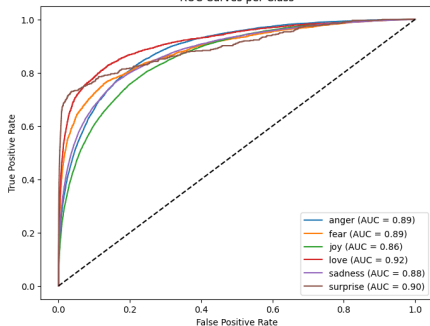
Slika: LSTM matrica



Slika: Transformer matrica

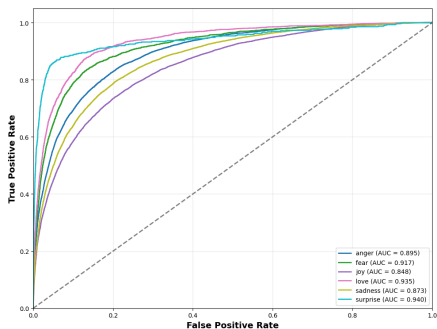
ROC krive po klasama

ROC Curves per Class



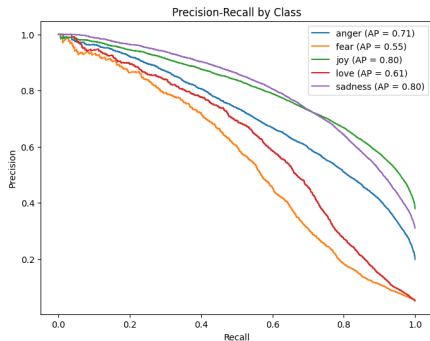
Slika: LSTM kriva

Transformer Model - ROC Curves per Emotion Class

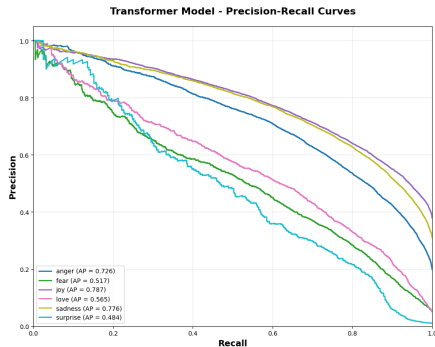


Slika: Transformer kriva

Precision-recall krive



Slika: LSTM kriva



Slika: Transformer kriva

Primeri predviđanja na osnovu teksta

| Text snippet | True | Pred | Prob |
|----------------------------------|-------|---------|-------|
| "i told you that i loved you..." | joy | love | 0.798 |
| "merry christmas have a very..." | joy | joy | 0.999 |
| "this ending is all but an..." | anger | sadness | 0.727 |
| "i catch a vibe when i am..." | joy | joy | 0.927 |
| "so many memories and so..." | joy | sadness | 0.614 |

Tabela: Primeri predikcija LSTM modela za emocije u pesmama (skraćeni tekstovi).

- Poređenje LSTM i Transformer modela dalo je **vrlo slične rezultate**.
- Uprkos **dominaciji** Transformer arhitektura u savremenim NLP zadacima, LSTM modeli su i dalje **relevantni i konkurentni** modeli.