

Recommending Appropriate Lyrics to Youngsters

Understanding the Presence of Inappropriate Content in Music Lyrics: Insights for Children’s Recommender Systems

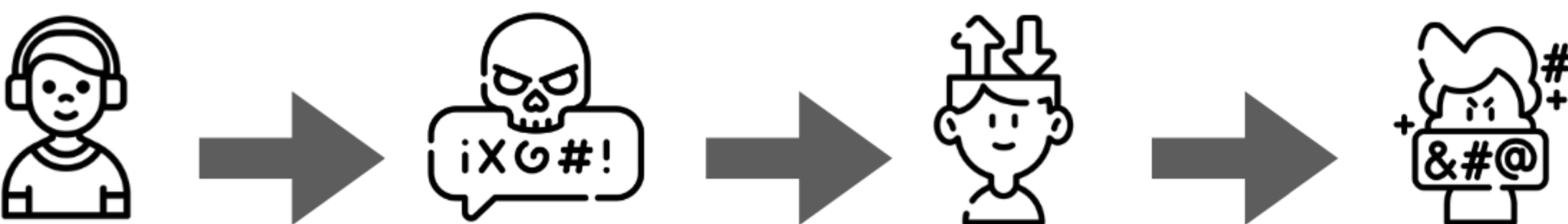
AUTHOR
Jasper Heijne

SUPERVISORS
Sole Pera & Robin Ungruh

PROBLEM

Children are easily influenced [1]; inappropriate music can negatively impact their behaviour [2]. Music recommender systems do not consider these influences well.

Main points missing in current research:
1. A large-scale overview of inappropriate content in music.
2. The relationship between inappropriate content and recommender systems.



RESEARCH QUESTION

What is the prevalence of inappropriate content in lyrics of modern Pop, Rock, Country, Rap, and R&B music and how does this influence music recommender systems for children?

DATA & MODELS

Figure 1 shows the distribution of 37,993 songs in genres from the Genius dataset.

We scan 4 facets of inappropriateness:

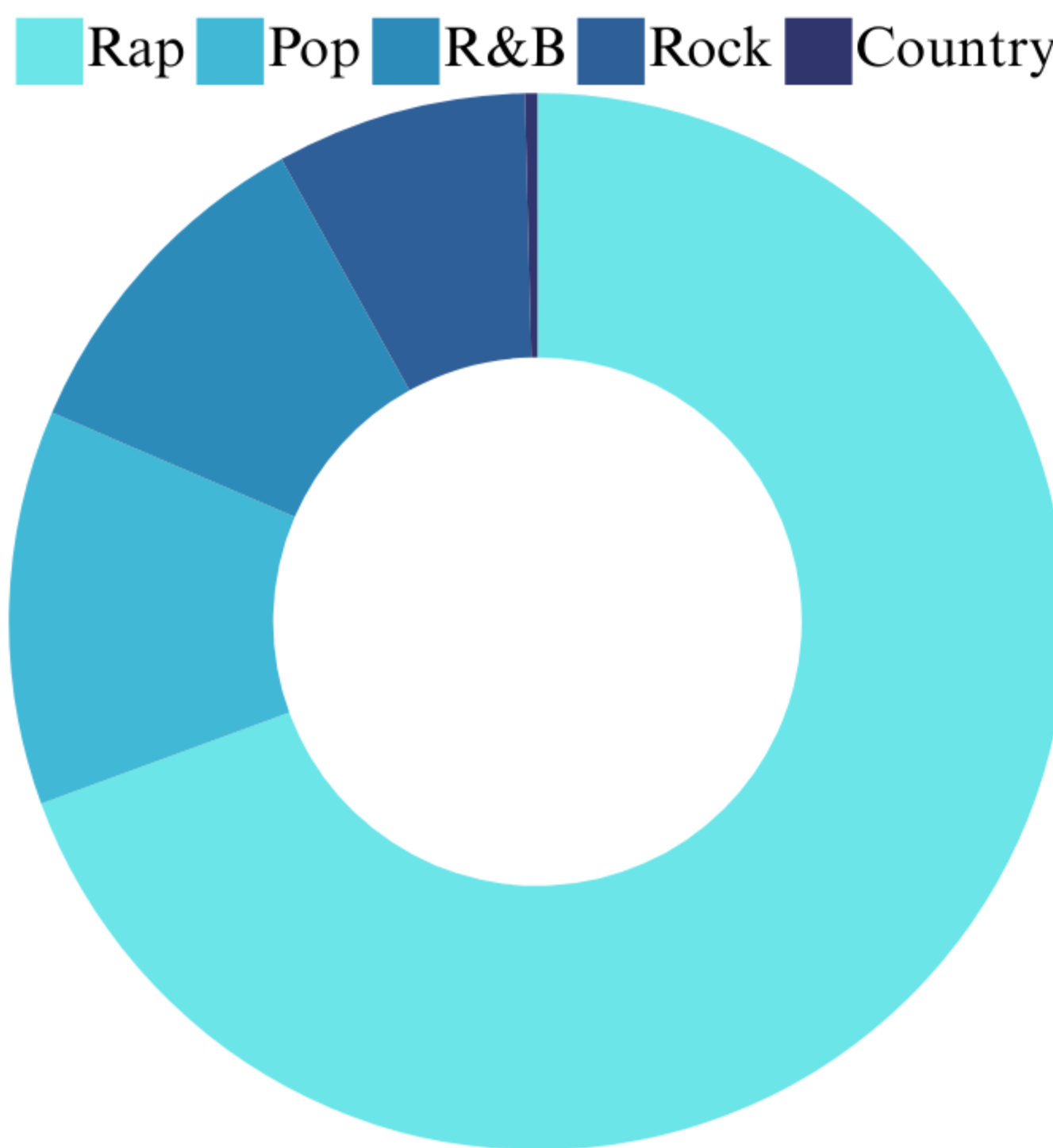
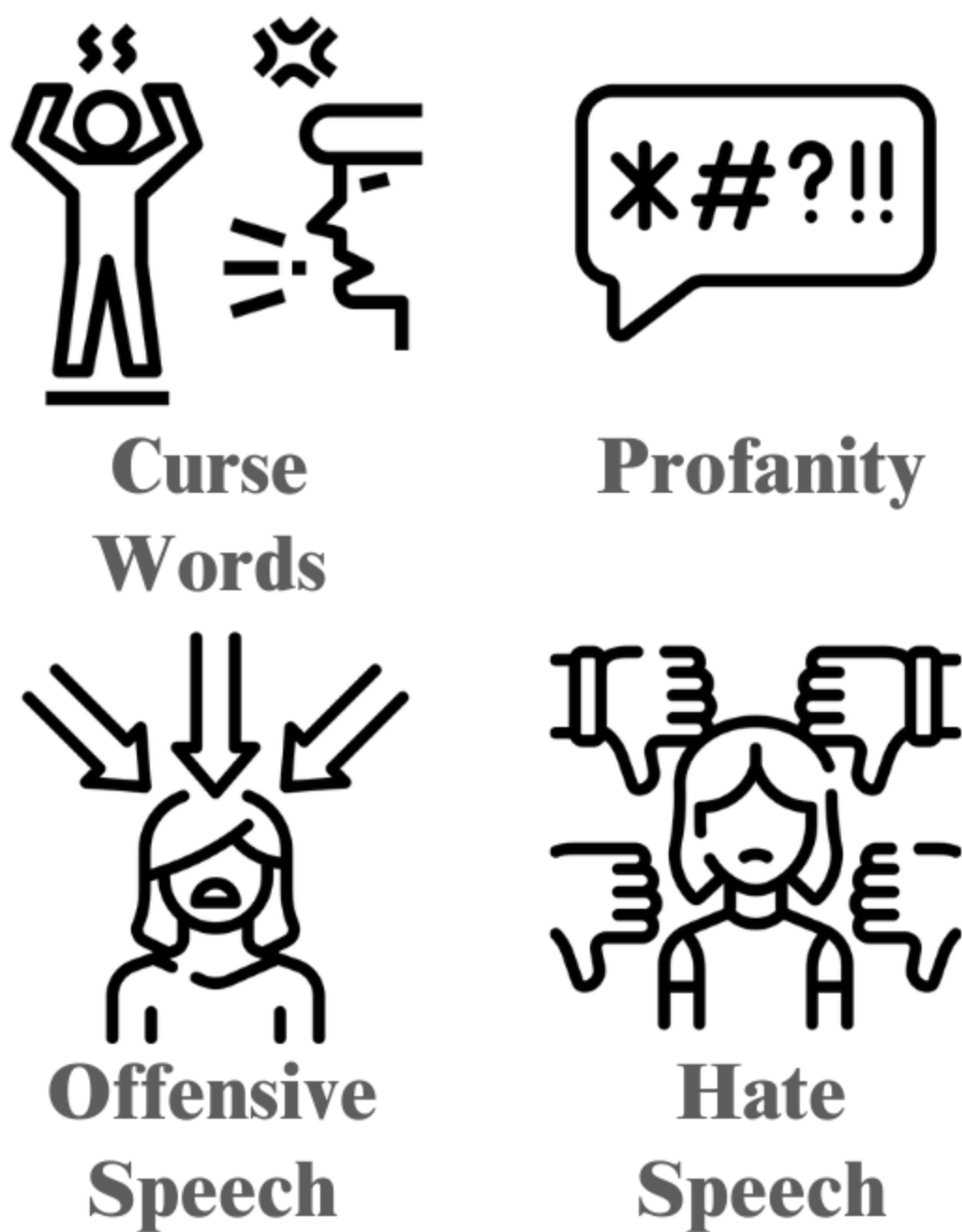


Fig 1: Distribution of music genres in the dataset

RESULTS

The lyrics of all songs are probed by the four models. Each model results in a percentage, or probability representing inappropriateness. For an overview of the distributions per facet of inappropriateness see Figures 2-5.

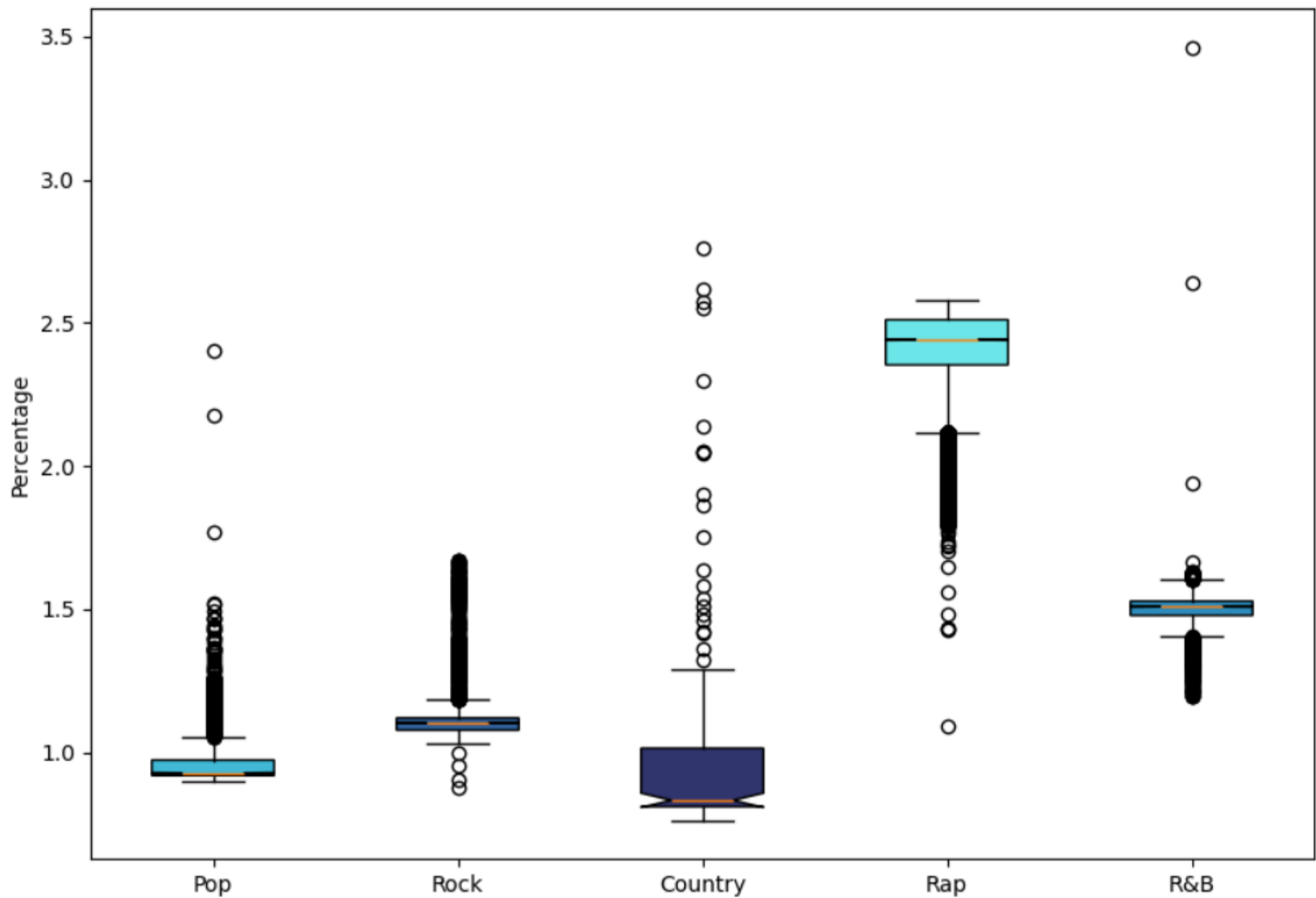


Fig 2: Distribution of percentage of curse words in lyrics for each genre.

- Rap scores highest in the percentage of curse words, profanity and offensive speech.
- R&B ranks second in inappropriateness based on the percentage of curse words, profanity and offensive speech.
- Rock scores highest in hate speech.

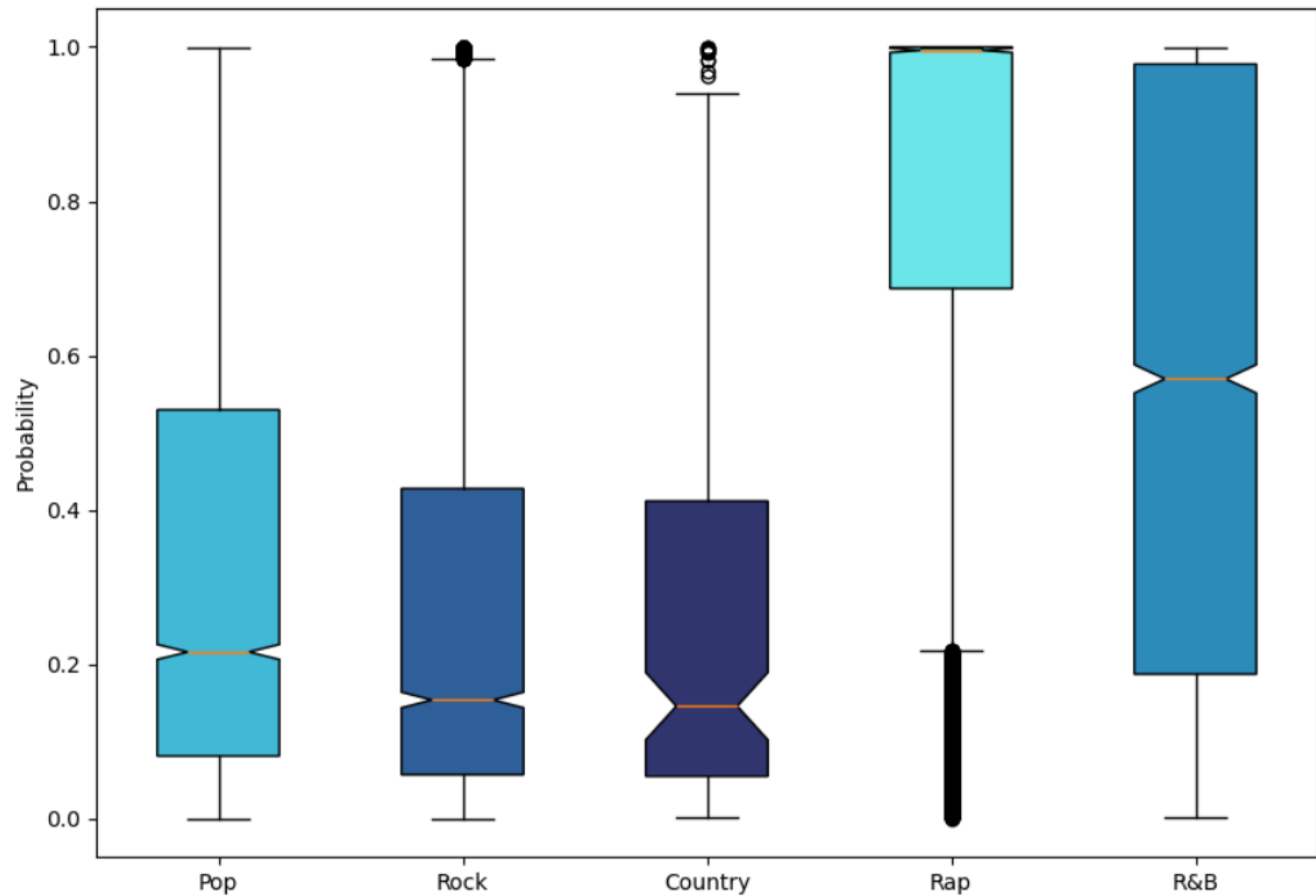


Fig 3: Probability distributions of songs to be profane per genre.

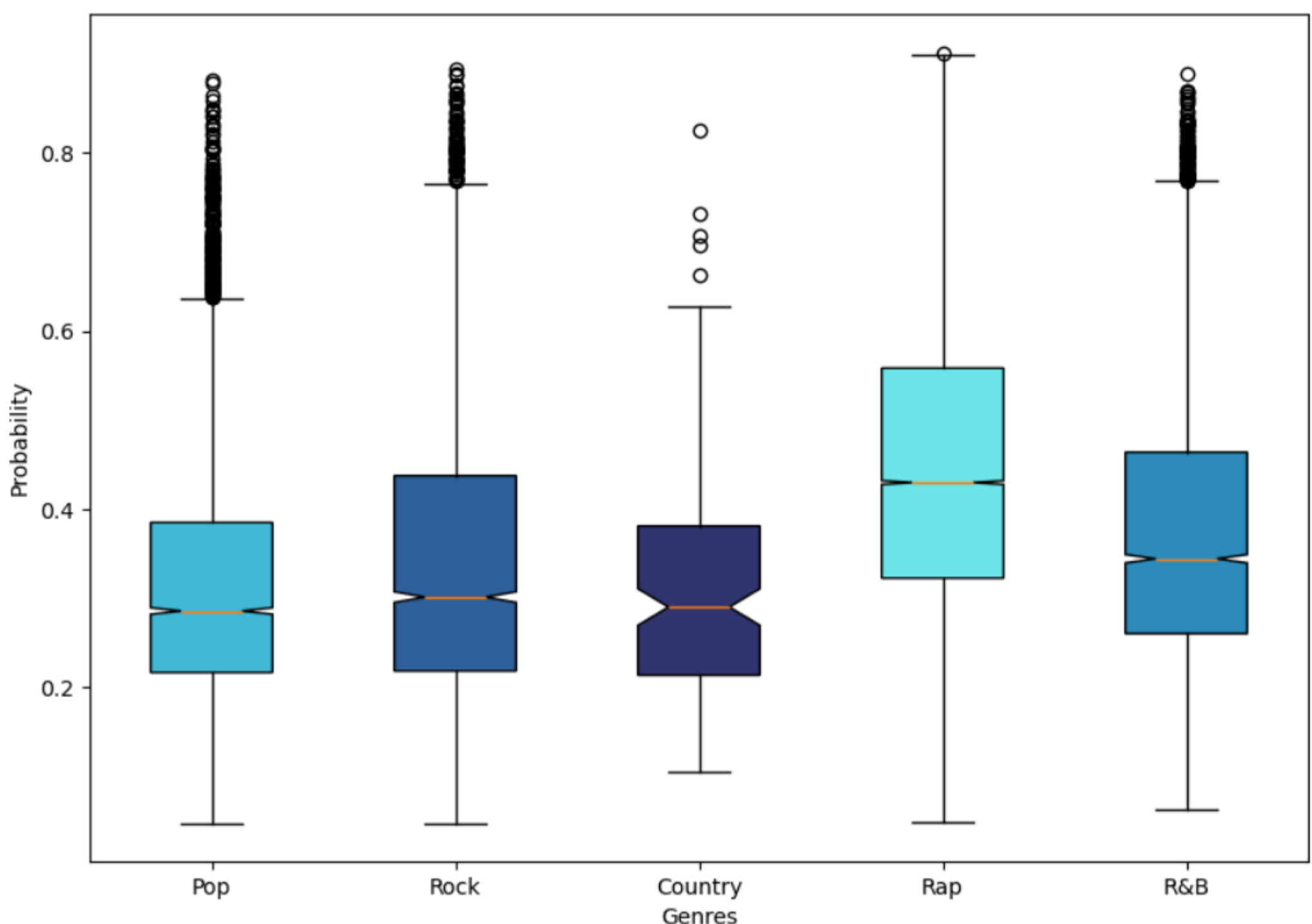


Fig 4: Probability distribution of songs to be offensive speech per genre

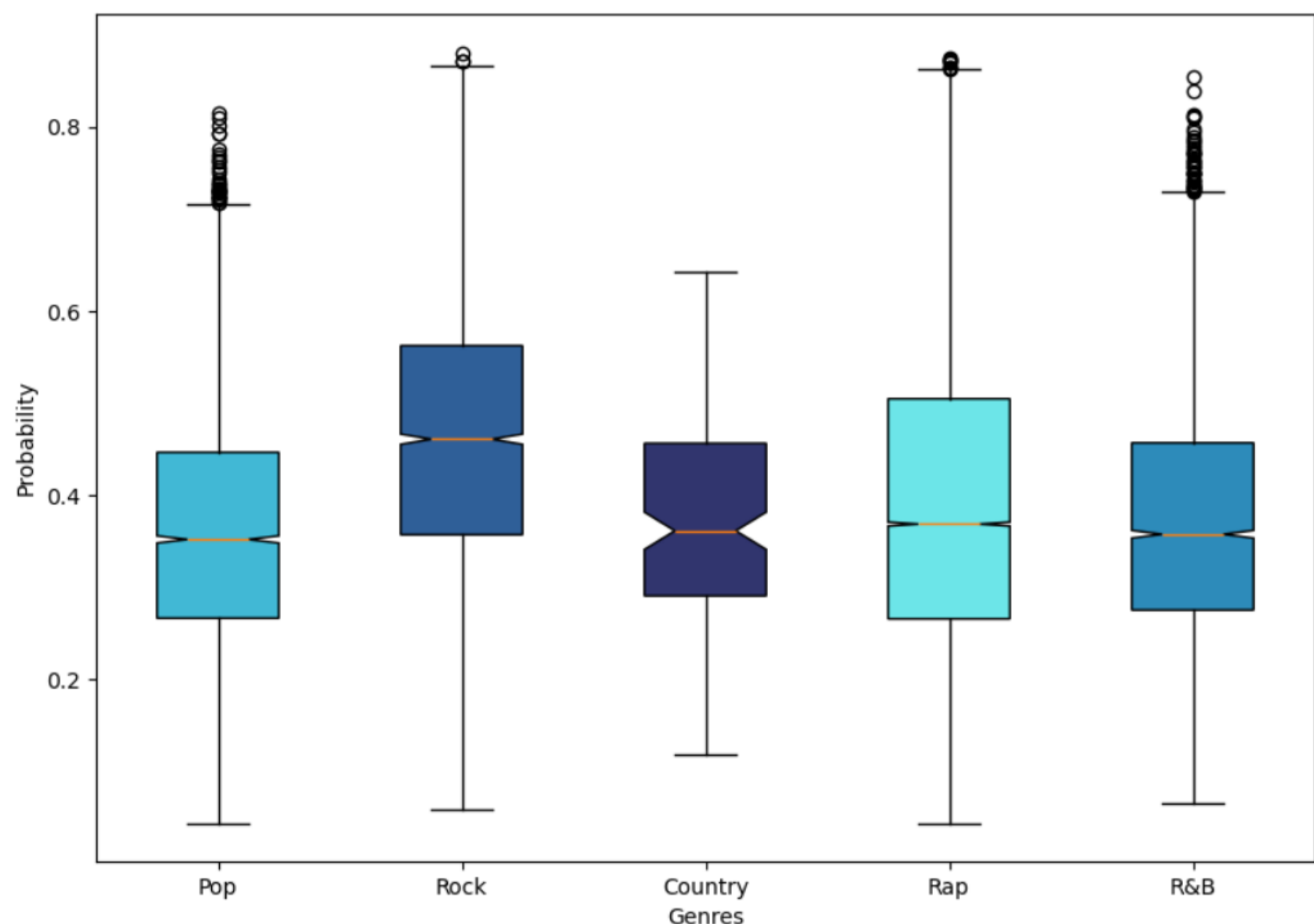


Fig 5: Probability distributions of songs to be hate speech per genre.

DISCUSSION

Results reveal that inappropriateness is not a binary label and is more nuanced. An in-depth overview of facets of inappropriate content is needed.

Three readability measurements reveal young children can understand the lyrics in the genres, showing the necessity of monitoring recommendations.

Overall we advise future children's music recommender systems to include a filter per song to filter on facets of inappropriateness in the lyrics.

CONCLUSION

1. All genres contain inappropriate content.
2. Rap and R&B are most inappropriate for children.
3. Music recommender systems should filter songs containing inappropriate lyrics.

