

Transformer-based approach towards Music Emotion Recognition from Lyrics

Yudhik Agrawal, Ramaguru Guru Ravi Shanker, Vinoo Alluri

Cognitive Science Lab, International Institute of Information Technology, Hyderabad



Find Our Work at Github

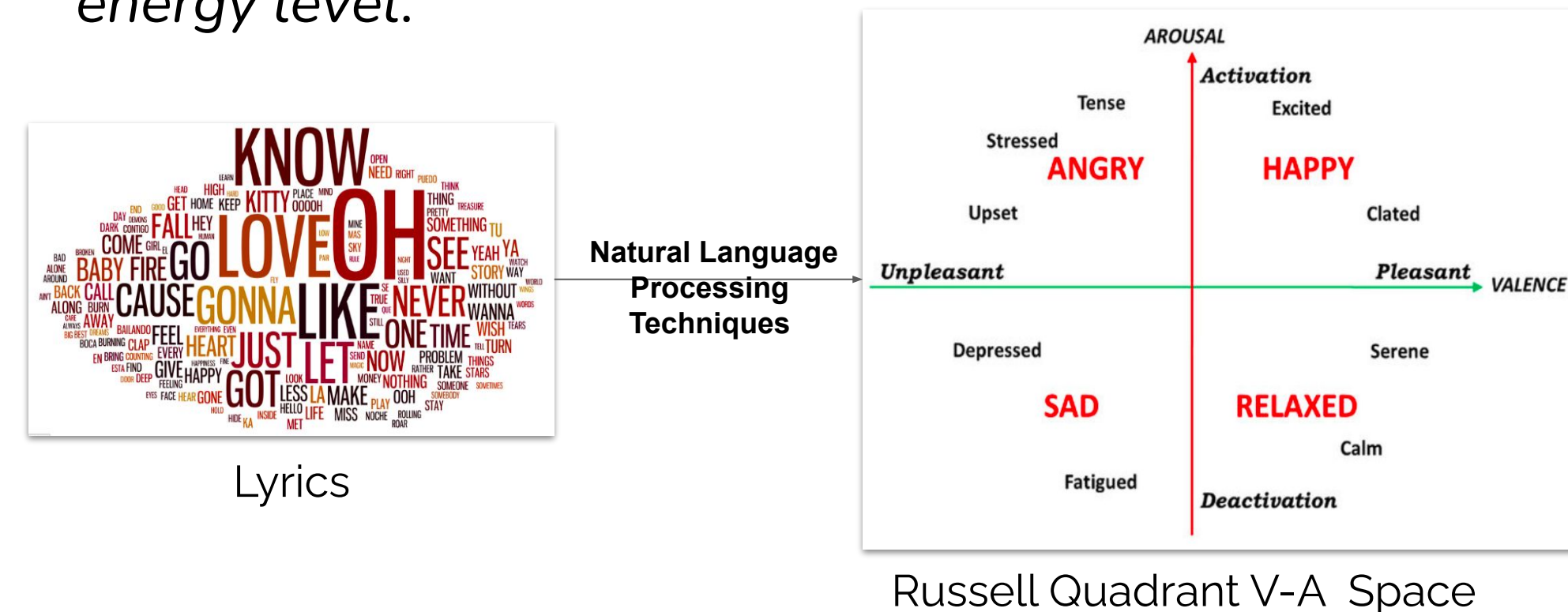


Aim & Motivation

- Individuals seek a varied range of emotional experiences via music.
- Identifying emotions from music helps in organization, retrieval, and recommendation of music to satisfy an individual's personal needs.
- Music Emotion Recognition has been limited to the usage of acoustic content, social tags, and metadata.



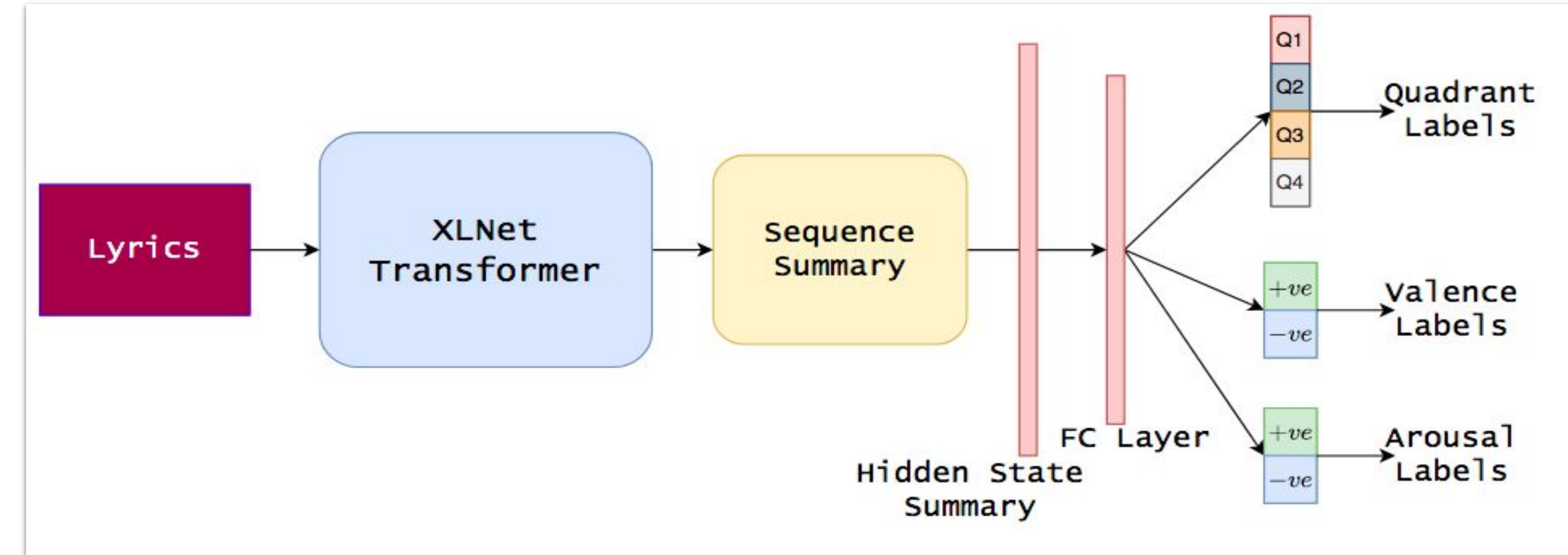
- Lyrics are a vital factor contributing to musical reward and play a crucial role in *eliciting emotions*. But have been largely neglected for Music Emotion Recognition tasks.
- Music Emotion Recognition via lyrics involves *identifying emotional connotations of lyrics* using NLP techniques and projecting them onto an emotion space (VA space) where *Valence* represents *pleasantness* and *Arousal* represents *energy level*.



Datasets

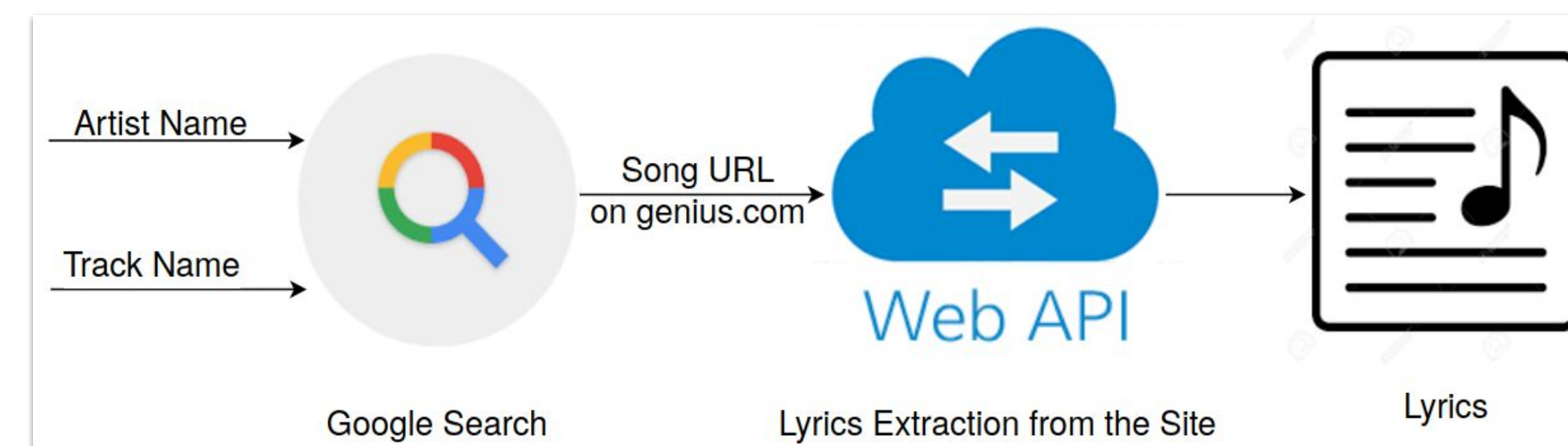
- MER Dataset**^[2]: contains 180 songs with manually assigned VA values uniformly distributed across 4 quadrants and solely based on lyrics (without audio).
- MoodyLyrics**^[3]: contains 2595 songs uniformly distributed across 4 emotion quadrants of the VA space. Valence and Arousal values were assigned based on lyrics using several lexicons like ANEW, WordNet.

Approach



- We propose a multi-task deep neural network architecture that, given the lyrics, outputs the classification of Emotion Quadrants, in addition to Valence and Arousal Hemispheres.
- The proposed network uses XLNet transformer as the base network.
- We also trained our network on individual tasks which involves training our model for a singular task i.e either Quadrant, Valence and Arousal.

Lyrics Extraction



- The first step for the task at hand, is to extract lyrics as the datasets do not provide lyrics due to copyright issues.
- All APIs require exact Artist and Track name for extracting lyrics, which are often misspelled in the datasets.
- We provide an improved approach for extracting lyrics using an added web-crawler which helps in getting the correct URI using Google search.
- Hirate improved from (60-80)% to ~99% in both the datasets.

References

- [1] Abdillan, J., Asror, I., Wibowo, Y.F.A., et al.: Emotion classification of song lyrics using bidirectional lstm method with glove word representation weighting. Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)4(4), 723–729 (2020)
- [2] Malheiro, R., Panda, R., Gomes, P., Paiva, R.P.: Emotionally-relevant features for classification and regression of music lyrics. IEEE Transactions on Affective Computing 9(2), 240–254 (2016)
- [3] Cano, E., Morisio, M.: Moodylyrics: A sentiment annotated lyrics dataset. In: Proceedings of the 2017 International Conference on Intelligent Systems, Meta-heuristics & Swarm Intelligence. pp. 118–124 (2017)

Results

Results of classification on MER dataset

Classification	Approach	Accuracy	Precision	Recall	\mathcal{F}_1 -score
Quadrant	Traditional NLP-based technique [2]	-	-	-	80.10%
Quadrant	Our Method	88.89%	90.83%	88.75%	88.60%
Valence	Traditional NLP-based technique [2]	-	-	-	90.00%
Valence	Our Method	94.44%	92.86%	95.83%	93.98%
Arousal	Traditional NLP-based technique [2]	-	-	-	88.30%
Arousal	Our Method	88.89%	90.00%	90.00%	88.89%

Results of classification by Quadrants on MoodyLyrics dataset

Approach	Accuracy	Precision	Recall	\mathcal{F}_1 -score
Naive Bayes [1]	83.00%	87.00%	81.00%	82.00%
BiLSTM + Glove [1]	91.00%	92.00%	90.00%	91.00%
Our Method	94.78%	94.77%	94.75%	94.77%

Ablation Study of Our Method on MoodyLyrics

Classification	Accuracy		\mathcal{F}_1 -score	
	Multi-Task	Single-Task	Multi-Task	Single-Task
Quadrant	94.78%	95.68%	94.77%	95.60%
Valence	95.73%	96.51%	95.67%	96.46%
Arousal	94.38%	94.38%	94.23%	94.35%

Conclusion

- First study to use a transformer model in identifying emotional connotations of lyrics.
- Used improved methodology to extract lyrics using a crawler.
- Model **outperforms existing state-of-the-art** methods on multiple datasets.

Future Work

- Helps Hybrid music recommendation systems by incorporating emotional connotations of lyrics for retrieval.
- This study can help us gain insight into the relationship between individual differences like personality traits and preferences for certain kinds of emotionally-laden lyrics.