RESULT ON MUSIC GENRE CLASSIFICATION USING DEEP LEARNING

Maulik Desai , Jay Rodge, Kapil Sahu, Advait Kulkarni , Prof. Bhavana Bahikar

SKNSITS, Lonavala

Abstract: *Music genre is evident cut names created by folks to categories bits of music. A musical style classification is depicted by the fundamental attributes shared by its people. These attributes usually area unit known with the instrumentation, harmonic content, and regular structure of the music. Genre is typically wont to structure the expansive accumulations of music accessible on the net. Presently, musical style labelling is performed manually. Automatic musical style classification arrangement will facilitate the people also as enterprises during this procedure and would be a valuable enlargement to music information retrieval frameworks and recommendation frameworks. Likewise, Automatic musical style classification provides a structure to making and assessing highlights for a substance based mostly examination of musical signals. During this paper, the automated musical style into a progressive system of musical style is explored mistreatment Long Short Term Memory (LSTM). LSTM may be a modification of perennial Neural Networks (RNNs), RNNs area unit specifically used for consecutive information adores Audio, Video and Text. The performance and relative importance of the planned options area unit investigated by coaching applied mathematics pattern recognition classifiers mistreatment GTZAN dataset. mistreatment the planned feature sets, classification of seventieth for 6 musical genres is achieved.*

Introduction:

A music may be a typical class that identifies some items of music as happiness to a shared tradition or set of conventions. it's to be distinguished from musical type and expressive style, though in apply these terms area unit generally used interchangeably

Existing System:

- No such commercial system
- Manual Classification

Propose System

- Classifier with better accuracy
- Using Long short term memory

Literature Survey:

1. **Paper Name: Musical Genre Classification of Audio Signals**

Author Name: George Tzanetakis

Description:

Musical genres are categorical labels created by humans to characterize items of music. A musical style is characterised by the common characteristics shared by its members. These characteristics usually are involving the instrumentation, rhythmic structure, and harmonic content of the music. Genre hierarchies are usually wont to structure the massive collections of music obtainable on the online. Presently musical style annotation is performed manually. Automatic musical style classification will assist or replace the human user during this method and would be a valuable addition to music data retrieval systems. Additionally, automatic musical style classification provides a framework for developing and evaluating options for any style of content- based analysis of musical signals. In this paper, the automated classification of audio signals into an hierarchy of musical genres is explored. Additional specifically, three feature sets for representing timbered texture, rhythmic content and pitch content are projected. The performance and relative importance of the projected options is investigated by training applied math pattern recognition classifiers exploitation real-world audio collections. Each whole file and period frame-based classification schemes are represented. Exploitation the projected feature sets, classification of sixty one for 10 musical genres is achieved. This result is similar to results reportable for human musical style classification.

2. **Paper Name: CONVOLUTIONAL RECURRENT NEURAL NETWORKS FOR MUSIC CLASSIFICATION**

Author name: Keunwoo Choi, Gyorgy Fazekas, Mark Sandler

Description: We introduce a convolution repeated neural network (CRNN) for music tagging. CRNNs benefit of convolution neural networks (CNNs) for native feature extraction and repeated neural networks for temporal report of the extracted options. We tend to compare CRNN with 3 CNN structures that are used for music tagging whereas dominant the number of parameters with relation to their performance and coaching time per sample. Overall, we tend to found that CRNNs show a robust performance with relation to the quantity of parameter and coaching time, indicating the effectiveness of its hybrid structure in music feature extraction and have summarisation.

3. **Paper Name: Deep Neural Networks: A Case Study for Music Genre Classification**

Author Name: Arjun Raj Rajanna

Description: Music classification could be a difficult downside with many applications in today's large-scale datasets with Gigabytes of music files and associated information and on-line streaming services. Recent success with deep neural network architectures on large-scale datasets has impressed varied studies within the machine learning community for numerous pattern recognition and classification tasks akin to automatic speech recognition, natural language process, audio classification and laptop vision. During this paper, we tend to explore a two-layer neural network with manifold learning techniques for style classification. We compare the classification accuracy rate of deep neural networks with a group of well-known learning models together with support vector machines (SVM and `1-SVM), supplying regression and `1- regression together with over sewn audio options for a genre classification task on a public dataset. Our experimental results show that neural networks square measure comparable classic learning models once the info is diagrammatical in an exceedingly made feature space.

4. **Paper Name: IMPROVED MUSIC FEATURE LEARNING WITH DEEP NEURAL NETWORKS**

Author Name: Siddharth Sigtia, Simon Dixon

Description: Recent advances in neural network coaching give the way to efficiently learn representations from information. Sensible representations are a very important demand for Music info Retrieval (MIR) tasks to be performed with success. However, a serious drawback with neural networks is that coaching time becomes preventative for terribly massive datasets and also the learning rule will mire in native minima for terribly deep and wide network architectures. during this paper we tend to examine three ways to boost feature learning for audio knowledge exploitation neural networks: one. Using corrected Linear Units (ReLUs) instead of standard sigmoid units; a pair of. Using a powerful regularisation technique referred to as Dropout; three. Using Hessian-Free (HF) optimization to improve coaching of sigmoid nets. we tend to show that these methods give important enhancements in coaching time and the options learnt area unit higher than state of the art handcrafted features, with a genre classification accuracy of eighty three ± 1.1% on the Tzanetakis (GTZAN) dataset. we tend to found that the rectifier networks learnt higher options than the sigmoid networks. we tend to additionally demonstrate the capability of the options to capture relevant info from audio knowledge by applying them to genre classification on the ISMIR 2004 dataset.

5. **Paper Name: Long Short-term Memory Recurrent Neural Network based Segment Features for Music Genre Classification**

Author Name: Jia Dai, Shan Liang, Wei Xue, Chongjia Ni

Description: In the standard frame feature primarily based music classification ways, the audio information is pictured by independent frames and also the consecutive nature of audio is completely ignored. If the consecutive information is well shapely and combined, the classification performance is considerably improved. The long short-run memory(LSTM) perennial neural network (RNN) that uses a collection of special memory cells to model for long-range feature sequence, has been successfully used for several sequence labelling and sequence prediction tasks. During this paper, we tend to propose the LSTM RNN based section options for music classification. The LSTM RNN is employed to find out the illustration of LSTM frame feature. The section options are the statistics of frame options in every section. What is more, the LSTM section feature is combined with the section illustration of initial frame feature to get the fissional section feature. The analysis on ISMIR info show that the LSTM section feature performs higher than the frame feature. Overall, the fissional segment feature achieves eighty nine.71% classification accuracy, about 4.19% improvement over the baseline model exploitation deep neural network (DNN). This vital improvement show the effectiveness of the planned section feature.

Architecture Design

System Design
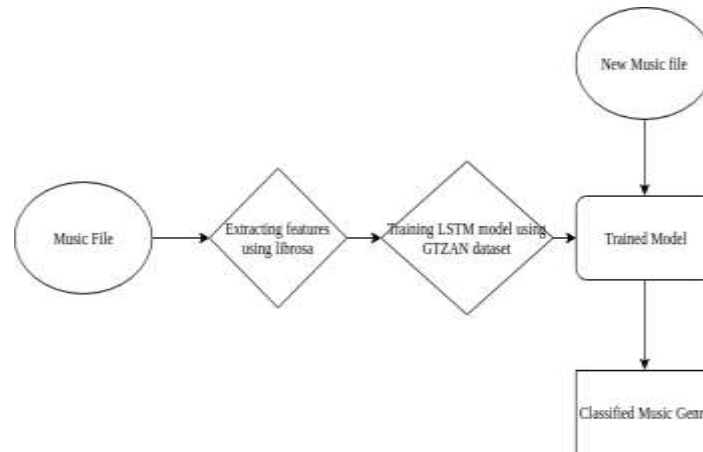
Future Scope:
- New dataset
- Accuracy
- Integration with other music services

Conclusion:

We propose the LSTM RNN primarily based phase options for music genre classification. LSTM RNN is sweet at modeling feature sequence with long-run context, which may higher represent the music knowledge. However the sequence coaching of LSTM RNN will cause that the incorrectly classified frames gather along, as well because the right classification frames. Then the fissional segment feature which mixes the LSTM RNN and scattering feature is projected to ameliorate this drawback. the many improvement of projected feature indicates the success of projected model.

References:

[1] A. Meng, P. Ahrendt, J. Larsen, and L. K. Hansen, "Temporal feature integration for music genre classification," Audio, Speech, and Language Processing, IEEE Transactions on, vol. 15, no. 5, pp. 1654–1664, 2007.

[2] Y. Song and C. Zhang, "Content-based information fusion for semi-supervised music genre classification," Multimedia, IEEE Transactions on, vol. 10, no. 1, pp. 145–152, 2008.

[3] I. H. Chung, T. N. Sainath, B. Ramabhadran, M. Picheny, J. Gunnels, V. Austel, U. Chauhari, and B. Kingsbury, "Parallel deep neural network training for big data on blue gene/q," in Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, 2015, pp. 745–753.

[4] Y. Zhao, D. P. Tao, S. Y. Zhang, and L. W. Jin, "Similar handwritten chinese character recognition based on deep neural networks with big data," Journal on Communications, 2014.

[5] G. E. Dahl, D. Yu, L. Deng, and A. Acero, "Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition," IEEE Transactions on Audio, Speech, and Language Processing, vol. 20, no. 1, pp. 30–42, 2012.

[6] James Bergstra, Olivier Breuleux, Fr´ed´eric Bastien, Pascal Lamblin, Razvan Pascanu, Guillaume esjardins, Joseph Turian, David Warde-Farley, and Yoshua Bengio. Theano: a CPU and GPU math expression compiler. In Proceedings of the Python for Scientific Computing Conference (SciPy), June 2010. Oral Presentation.

[7] James Bergstra, Norman Casagrande, Dumitru Erhan, Douglas Eck, and Bal´azs K´egl. Aggregate features and AdaBoost for music classification. Machine Learning, 65(2-3):473–484, 2006.

[8] James Bergstra, Michael I Mandel, and Douglas Eck. Scalable genre and tag prediction with spectral covariance. In ISMIR, pages 507–512, 2010.

[9] Christopher M Bishop. Neural networks for pattern recognition. Oxford university press, 1995.

[10] George E Dahl, Tara N Sainath, and Geoffrey E Hinton. Improving deep neural networks for LVCSR using rectified linear units and dropout. In Proc. ICASSP, 2013.