Neural Networks: Project Report.

Irvin Aloise, Mirco Colosi
Sapienza University of Rome

November 29, 2016

1 Introduction

MIDI databases are still very popular nowadays and the scientific community is always improving the way in which it is possible to categorize them. MIDI - which stands for *Musical Instrument Digital Interface* - is a technical standard that allows to connect and make them communicate properly several musical tools, e.g. instruments, digital equipment and computers [ref. https://en.wikipedia.org/wiki/MIDI]. It is composed by messages containing informations about *notation*, *pitch*, *velocity*, *control signals* - that describe parameters like vibrato or the volume - and *clock* signals in order to synchronize the tempo of all the instruments.

With the diffusion of machine learning approaches, those classification methods are always more powerful and precise, taking advantage of MIDI informations together with other audio features, in order to achieve categorization of the tracks by *author*, *genre*, *style* and so on.

In this paper, instead, it has been used a more general approach: it has been designed and developed a mechanism that measures the similarity between tracks and, given a new instance returns the author which mostly suits that instance - chosen between the authors available in the training set. To do that, we used a *universal* similarity metric based on **Kolmogorov complexity** [ref: LI, Ming, et al. The similarity metric. IEEE transactions on Information Theory, 2004, 50.12: 3250-3264.]. The peculiarity of this approach is that can be employed potentially on every kind of file with no modification and without the need of time expensive calculations for extracting audio features. Then, once that a similarity measure is retrieved, a simple *k-NN* has been employed to classify new instances.

The remaining of the document is organized as follows: in Section 2 a brief overview of the related approaches is given; Section 3 describes the methodology used in this project together with some results; finally in Section 4 are reported conclusions and possible future improvements.

2 Related works

Most of the works in the literature try to classify MIDI songs by *genre*. This because generally datasets are stored by author but recognize the style could be useful in automated platforms, for example to suggest a new song related to the one that we are listening.

For example, **Basili et al** [ref: Basili, Roberto, Alfredo Serafini, and Armando Stellato. "Classification of musical genre: a machine learning approach." ISMIR. 2004.] use some *coarse-grain features* in order to classify MIDI files by genre. Those features are basically provided directly by the MIDI file, in order to evaluate how good MIDI describe symbolic music; they are the following ones:

- Melodic intervals
- Instruments
- Instrument classes and drumkits
- Meter and time changes
- Note extension

Those features are given as input to several machine-learning algorithm - *Naive Bayes*, *VFI*, *J48*, *PART*, *NNge* and *JRip* - to extract a genre prediction for new MIDI instances.

Chet in his work *Classification of Musical Playing Styles using MIDI Information* [ref: Gnegy, Chet N. "Classification of Musical Playing Styles using MIDI Information."] instead tries to return a *style* classification of the instances, intended as a functional description of a specific instruments in a song. More precisely, the styles used are the following:

- **Bass**: the song has a predominant single line of deep tones and the other harmonic structure are just supporters.
- **Lead**: a melodic line, e.g. vocal performance or guitar solo.
- **Rhythm**: in general represented by repeated chords that give the rhythmic structure to the melody.
- **Acoustic**: this style is characterized by multiple individual notes - or melody lines - often performed on the same instruments.

A large variety of audio features is used in this work, spanning from simple statistics - *mean pitch*, *pitch standard derivation* and many others - to very complex ones - *coverage*, *liricality* and several others. Many different machine-learning algorithm are then used in order to assign a *style* to new instances, e.g. *decision tree*, *logistic regression*, *k-NN*, *QDA* and *SVM*.