Comparison of Musical Genre Classification Methods

Evan Gomolin   
*Faculty of Science*  
*Western University*London, Canada  
egomolin@uwo.ca

Ali Al-Tekriti  
*Faculty of Science*  
*Western University*London, Canada  
aaltekri@uwo.ca

David MacLean  
*Faculty of Science*  
*Western University*London, Canada  
[dmacle22@uwo.ca](mailto:dmacle22@uwo.ca)

Matvey Skripchenko  
*Faculty of Science*  
*Western University*London, Canada  
?????@uwo.ca

*Abstract*—This electronic document is a “live” template and already defines the components of your paper [title, text, heads, etc.] in its style sheet. *\*CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract*. (*Abstract*)

Keywords—sklearn, FMA, genre classification, model comparison, music analysis

# Introduction

This template, modified in MS Word 2007 and saved as a “Word 97-2003 Document” for the PC, provides authors with most of the formatting specifications needed for preparing electronic versions of their papers. All standard paper components have been specified for three reasons: (1) ease of use when formatting individual papers, (2) automatic compliance to electronic requirements that facilitate the concurrent or later production of electronic products, and (3) conformity of style throughout a conference proceedings. Margins, column widths, line spacing, and type styles are built-in; examples of the type styles are provided throughout this document and are identified in italic type, within parentheses, following the example. Some components, such as multi-leveled equations, graphics, and tables are not prescribed, although the various table text styles are provided. The formatter will need to create these components, incorporating the applicable criteria that follow.

# Related Work

## Selecting a Template (Heading 2)

First, confirm that you have the correct template for your paper size. This template has been tailored for output on the A4 paper size. If you are using US letter-sized paper, please close this file and download the Microsoft Word, Letter file.

## Maintaining the Integrity of the Specifications

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

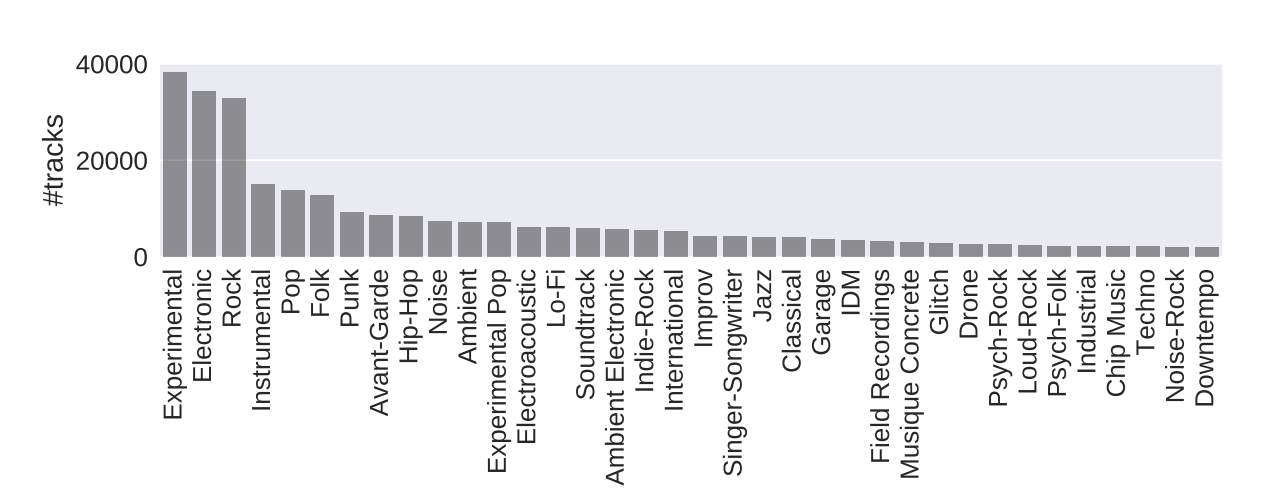
# Methods

Our purpose with this experiment is to compare the accuracy of different classification models in the sklearn library when classifying music into genres. Our methodology behind this experiment goes as follows:

1. Locate a sufficiently large and detailed dataset of music.
2. Preprocess the dataset such that it does not favour any single model.
3. Train many types of models from sklearn on identical preprocessed data
4. Test the models on reserved preprocessed data and compare their results.

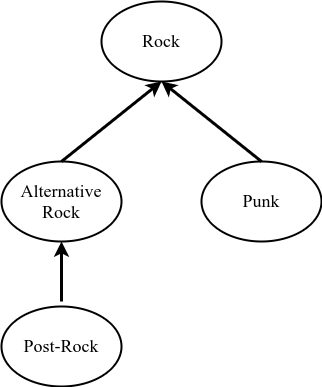
## Data

For the purposes of fair comparison, a dataset with a suffiently large number of songs and features must be selected. This will limit any bias produced by the stastically unsound nature of smaller datasets. For this experiment, we have selected the Free Music Archive (FMA) dataset created by Defferrard et al. [**citation needed]**. The FMA dataset includes 106,574 audio tracks of 30 seconds, from 16,341 artists, across 14854 albums. These tracks span 161 genres; however, as shown in Figure 1, the classes of genres are unbalanced.



1. Number of tracks in the 36 most populous genres in the FMA dataset. [**citation needed paper]**

In addition to the provided audio files, FMA provides hierarchical genre metadata. The metadata associated with a genre contains its parent genre, which indicates a more general genre that the genre in question is a subset of. For example, the genre “Post-Rock” has a parent genre of “Alternative Rock” . As seen in Figure 2, we will refer to genres at the top of this hierarchy, i.e. genres with reflexive parent references, as “root genres.”

1. Example of a possible genre hierarchy. Rock is the root genre of every genre in this tree.

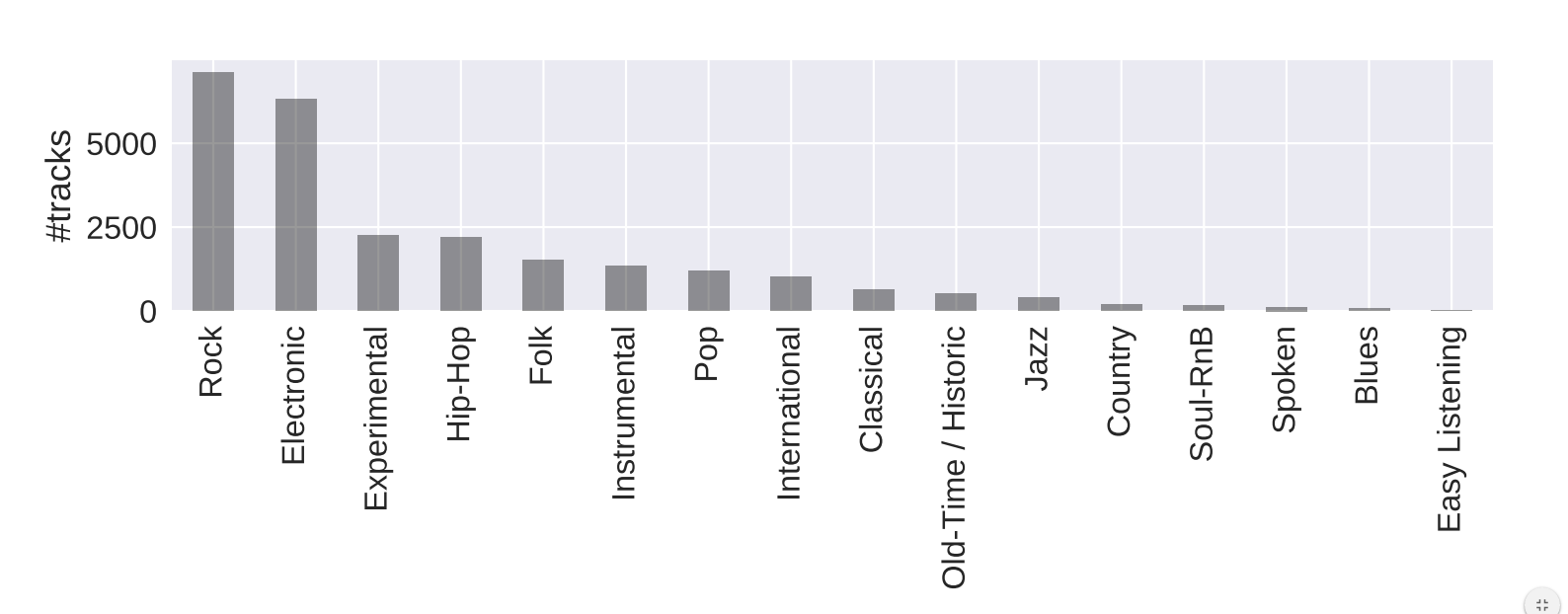
FMA also includes metadata describing each of the tracks. The metadata is split into 3 main categories: track information, librosa-features, and Echonest-features. Track information includes words describing the track, such as genre labels, artist names, location of recording, etc; as well as curated subsets and proposed splits. The subsets allow for less powerful computers to be able to process the dataset and have been curated to maximize the number of tracks with high quality metadata. Whereas the proposed splits indicate that 80% of the tracks are split into a training set, 10% into a validation set, and 10% into a test set. The librosa-features metadata includes information, like Mel-frequency cepstral coefficients, extracted from each track by the music processing library librosa. The Echonest-features metadata contains features extracted by the now dissolved company, Echonest; e.g., song tempo, or “danceability”.

## Data Preprocessing

In advance of training and comparing the various models, we developed a preprocessing stage to transform the FMA dataset into a consistent, usuable format. The goal with our preprocessing stage was to prepare the metadata in a generic way, such that the preprocessed metadata was equally useful to all the models. Additionally, the power of the computers which we had access to was considered when designing this stage.

Given the computational resources we had access to, we chose to only use the “medium” sized subset of the FMA dataset, as defined in the metadata. As shown in Figure 3, the resulting subset contained 25,000 tracks across 16 root genres. Furthermore, we decided not to use the FMA audio files due to similar resource limitations. Therefore, our dataset consists exclusively of metadata.

Next, each track was relabeled using the root genre for which it is a child of, instead of its specific (sub)genre. This allows us to limit the number of classes to 16, so that we had sufficiently many tracks in each class.



1. Number of tracks in “medium” subset of FMA across 16 root genres. [**citation needed paper]**

Once the tracks were relabeled, we organized the features associated with each of those tracks. First, we took the union of the librosa extracted features and the Echonest extracted features, because the source of the feature is irrelevant. Then we split the features and their associated labels into training, validation, and test sets in accordance with the FMA propositions. To prevent any unintended learning based on the natural order of the dataset, we randomly shuffled the order of the training set. It is important to note that every model will receive the same random shuffling. Finally, we centre all three sets about the training set’s mean, and make sure they are of unit variance using (1).

y(c) = (c – m) / s 1)

Where μ is the mean of the training set, and σ is the standard deviation of the training set. At this point all of our features are sufficiently preprocessed and ready to be passed on to the training stage.

## Model Training

Our goal when selecting models was to construct a set of of classifiers with high variety. This would allow us to measure how changes to the type of model affected the classification results. For this experiment we selected 10 models from the sklearn library.

1. Logistic Regression
2. Linear Support Vector Classification
3. Support Vector Classification with Radial Basis Function Kernel
4. k Nearest Neighbours
5. Multilayer Perceptron Classifier
6. Decision Tree Classifier
7. Gaussian Naive Bayes
8. Gradient Boosting Classifier
9. Linear Discriminant Analysis
10. Quadratic Dsicriminant Analysis

For each model, we trained it on several different subsets of the entire feature set. The purpose of this was to find the optimal performance of each model on the FMA dataset. Therefore, we would not be comparing models on a subset of features that favours certain models. Instead, every model can be evaluated by its best solution.

[INSERT FIGURE W/FEATURE SET COMPARISON]

## Model Testing

Given every model, trained on its optimal subset of features, each one predicts the genres of the test set features. For each model, a confusion matrix is contructed to characterize the performance of that model. Furthermore, the recall, and precision scores of each model operating on the test set are calculated with (2) and (3) respectively.

RecallScore = Tp / (Tp + Tn) 2)

PrecisionScore = Tp / (Tp + Fp) 3)

Where Tp is the number of true positive predictions, Tn is the number of true negative predictions, and Fp is the number of false positive predictions.

# Experimental Results

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

## Authors and Affiliations

**The template is designed for, but not limited to, six authors.** A minimum of one author is required for all conference articles. Author names should be listed starting from left to right and then moving down to the next line. This is the author sequence that will be used in future citations and by indexing services. Names should not be listed in columns nor group by affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization).

### For papers with more than six authors: Add author names horizontally, moving to a third row if needed for more than 8 authors.

### For papers with less than six authors: To change the default, adjust the template as follows.

#### Selection: Highlight all author and affiliation lines.

#### Change number of columns: Select the Columns icon from the MS Word Standard toolbar and then select the correct number of columns from the selection palette.

#### Deletion: Delete the author and affiliation lines for the extra authors.

## Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include Acknowledgments and References and, for these, the correct style to use is “Heading 5”. Use “figure caption” for your Figure captions, and “table head” for your table title. Run-in heads, such as “Abstract”, will require you to apply a style (in this case, italic) in addition to the style provided by the drop down menu to differentiate the head from the text.

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced. Styles named “Heading 1”, “Heading 2”, “Heading 3”, and “Heading 4” are prescribed.

## Figures and Tables

#### Positioning Figures and Tables: Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation “Fig. 1”, even at the beginning of a sentence.

1. Table Type Styles

| Table Head | Table Column Head | | |
| --- | --- | --- | --- |
| Table column subhead | Subhead | Subhead |
| copy | More table copya |  |  |

1. Sample of a Table footnote. (*Table footnote*)
2. Example of a figure caption. (*figure caption*)

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity “Magnetization”, or “Magnetization, M”, not just “M”. If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write “Magnetization (A/m)” or “Magnetization {A[m(1)]}”, not just “A/m”. Do not label axes with a ratio of quantities and units. For example, write “Temperature (K)”, not “Temperature/K”.

# Discussion

# References

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first ...”

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors’ names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [4]. Papers that have been accepted for publication should be cited as “in press” [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

1. G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. *(references)*
2. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
3. I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
4. K. Elissa, “Title of paper if known,” unpublished.
5. R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.
6. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
7. M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.

**IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.**