Comparison of learning models for music genre classification

Student[1], Student[2]

Abstract
Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Keywords
Keyword1 — Keyword2 — Keyword3

1
2

Contents

1. Introduction

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

and some mathematics $\cos\pi = -1$ and α in the text[1].

2. Data Analysis

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

$$\cos^3\theta = \frac{1}{4}\cos\theta + \frac{3}{4}\cos 3\theta \tag{1}$$

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

1. First item in a list
2. Second item in a list
3. Third item in a list

2.1 Subsection

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Paragraph Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

Paragraph Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetuer at, consectetuer sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

2.2 Subsection

Reference to Figure **??**.

3. Methods and Experiments

Fundamental intuition for choosing models: Multi-class classification being the underlying aspect of the problem statement, we turned to learning models which constitute this feature.

3.1 Logistic Regression

Motivation: Logistic regression though by nature the default for binary classification, the fact that it can be extended for multi-class models was the motivation for choosing it.

Our model predicted the probabilities of different possible genres (labels) corresponding to a given feature space (see the data analysis section). For each of the multi-classification modelling techniques [one-vs-one - *ovo*, one-vs-all - *ovr*], we empirically used various algorithms as solvers (like stochastic average gradient) and consecutively varied the following parameters.

- Maximum iterations for convergence $(100-900)$
- Initial class weights (balanced weighing, initializing to 1)
- Weight regularization strength $(0.1-0.9)$
- Tolerance of error across runs

in focus of improving on the accuracy of and reducing loss of the classification.

[1] And some mathematics $\cos\pi = -1$ and α in the text.

To optimize the hyper-parameters for modelling the logistic regression classifier we made use of the exhaustive Grid-search cross validation technique.

3.2 Support Vector Machines

Motivation: As the feature-space is of a relatively higher dimensionality, we decided to try out support vector machines to improve on the statistics achieved using the logistic model.

Initially, we modelled a linear support vector classifier which used set of hyperplanes for learning the training data. As we did not observe any drastic improvements in the statistics, we further analysed the data and decided to experiment with the non-linear modelling of the support vector machines.

While empirically using both *ovo* and the *ovr* techniques, we tested various kernels:

- Gaussian radial basis function
- Polynomial (degree 3)
- Sigmoid

For each of the kernels we varied the following parameters:

- Maximum iterations for convergence $(100-300)$
- Initial class weights (balanced weighing, initializing to 1)
- Weight regularization strength $(0.1-0.9)$
- Tolerance of error across runs

To optimize the hyper-parameters for modelling the support vector classifier we made use of the exhaustive Grid-search cross validation technique.

3.3 Ensemble Classifier

Motivation: After experimenting with the above two models for sometime, we figured that logistic regression classifiers were good with accuracy metric and SVMs were minimizing on the log-loss statistic. So we decided to ensemble them together to understand if we can get the better of both worlds.

We used the **voting classifier** strategy to combine the above classifiers using a majority vote to predict the genres. We used both hard and the average predicted probabilities voting (soft vote) methods and found that the hard voting produced better results.

3.4 Semi-Supervised Classifier

Motivation: After not being able to further deduce any meaningful direct relationship between the feature-space, we decided to try the semi-supervised learning methodology.

We chose the label-spreading strategy over label propagation as it is more robust to noise[1]. Modelling steps:

Step 1 Splitting the labelled (train) data as based on some percentage (we chose $\frac{1}{3}^{rd}$ of the it for validation)

Step 2 Merging the split of labelled data from [step-1] with the unlabelled data.

Step 3 Training the classifier/kernel chosen for label-spreading with respective parameters (listed below)

Step 4 Scoring/validating the trained model using the split test data from [step-1]

Step 5 Predicting the genres for unlabelled (actual test) data

We varied the following parameters over runs:

- Kernels (rbf, k-nearest neighbours). Varying the gamma parameter (b/w $0.1-00001$) and the number of neighbours attribute (b/w $50-300$) respectively.
- Maximum iterations for convergence $(100-300)$

Other models

Apart from the above mentioned classifiers, we experimented with more models that did not pass our benchmarks for both the accuracy and loss metrics.

- Naive Bayes'
- Decision trees
- Random forests

Though we initially considered reducing the dimensionality by performing principal component analysis, after some substantial scrutiny of the feature space we dropped it. The rhythm bands, pitch classes and the timbre coefficients we spread out across the statistics to discard.

Performance Metrics

Evaluation of all the models designed was done using the k-fold cross validation technique (with $k = 5$).

For splitting the data into folds we used stratified process of preserving the percentage of samples for each genre thereby enabling fair evaluation.

4. Results

Performance measures of the above experiments when evaluated with the test data.

4.1 Accuracy, Log-Loss statistics

- Logistic Regression Classifier across various values for its parameters. Results for the best 5 combinations:

Table 1. metrics on test data (LR)

modelling strategy	solver	accuracy	loss
ovr	sag	0.658	0.12
multinomial	sag	0.653	0.17
ovr	liblinear	0.650	0.21
ovr	lbfgs	0.646	0.26
ovr	newton-cg	0.646	0.26

- Support vector classifiers across various values for its parameters. Results for the best 5 combinations:

Table 2. metrics on test data (SVC)

modelling strategy	kernel	accuracy	loss
multinomial	rbf	0.636	0.07
ovr	rbf	0.636	0.07
multinomial	linear	0.607	0.20
ovr	linear	0.607	0.20
ovr	poly (degree 3)	0.567	0.36

- Ensemble classifier with best parametric combination of logistic regression, SVC and voting strategy produces

$$accuracy = 0.631, loss = 0.09$$

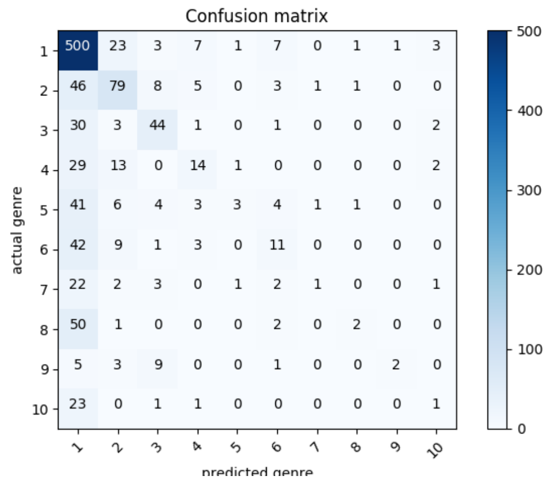
- Semi-supervised classifier with label spreading:

Table 3. metrics on test data (semi-supervised)

kernel	attribute	accuracy	loss
knn	neighbours(300)	0.533	0.14
rbf	$\gamma = 0.001$	0.612	0.29

4.2 Confusion matrix

The following confusion matrix represents the accuracy of classification. It is evident that the prediction is directly influenced by the frequency of respective genres in the training data.



The matrix entry i, j is the number of songs actually in genre i, but predicted to be genre j. Darker the blue, accurate the classification.

5. Discussion

6. Appendices

Thank you! Google!

References

[1] semi-supervised learning,