Music and AI: Music Genre Classification using Machine Learning Methods

John Mohoang and Mitchell Peterson
Intro to Artificial Intelligence - Susan Fox. December 2016

Introduction and Background

- **Motivation:** Our interest lies in the intersection between Artificial Intelligence and Music, specifically Machine Learning.
- **Why?** Music is such a subjective art form by nature, so what would happen if we applied completely objective algorithms to it? Would patterns emerge?
- **Our Goal:** Many music classification systems today recognize sounds by finding matches in a pre-existing database. We wanted to find a way to classify music that isn't in the database already.

Literature

- "Learning Features from Music Audio with Deep Belief Network" (2010) Hamel and Eck
- "A Machine Learning Approach to Automatic Music Genre Classification." (2008) Silla Jr et al.
- "Music Genre Classification of Audio Signals" (2002) Tzanetakis

Data Collection

- **Data Source**:
  - Tzanetakis Dataset
  - Built for his PhD thesis
  - 1000 songs split into 10 genres

python — We used Python and associated libraries for our analysis

- Utilized the Aubio library
  - Paul Brossier, 2006
  - Segment sound files before attack, perform pitch detection, detect tempo and beat changes, produce live midi streams and more.
- Extracted a frequency analysis (spectrogram), estimated average pitch and bpm for each song.

Data Cleaning

- Pre-trim: 667189000 data points.
- Averaged spectrogram data into 1 second groups and treated it as vectors to calculate the angle difference between each group
- Averaged pitch in 1 second intervals as well.
- Wrote a python script to extract, clean and export all 1000 songs worth of data into one .cvs file = 60991 data points.

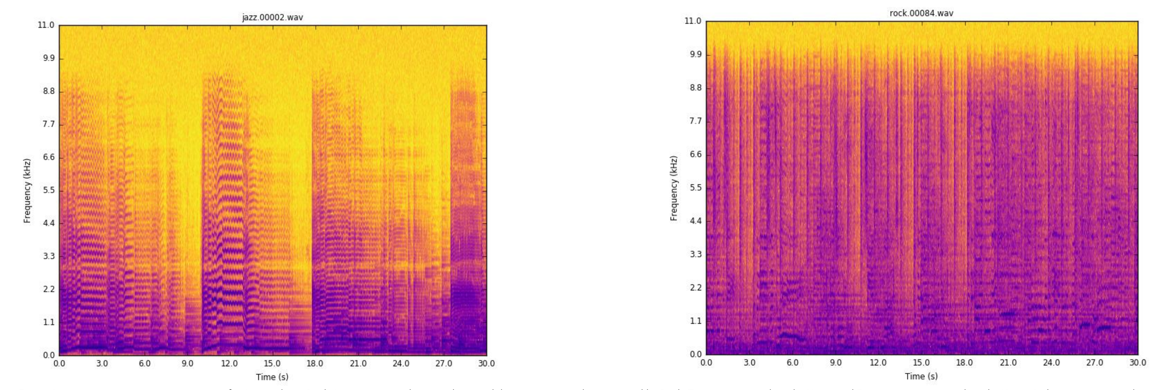

Figure 1. represents two spectrograms, one for Jazz and one Rock. It is easy to see how jazz has much less consistency than a genre like Rock. Time is measured on the X axis and Frequency is measured on the Y axis. Lightness corresponds to amplitude

Methods and Results

We used R Studio for our analysis.

Data was split 80% for training and 20% for testing.

Algorithms:

1). **Decision Tree**

- Predict that each observation belongs to the most commonly occurring class of training observations in the region to which it belongs.

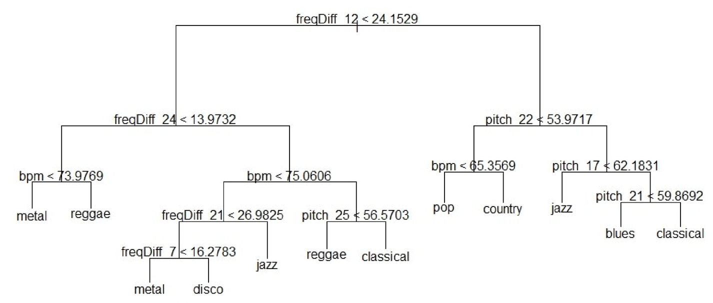

Figure 2. Our first, pre trimmed decision tree.

- Default classification tree with no pruning
- 81.05% misclassification rate

- Cross Validation to determine where to prune the tree

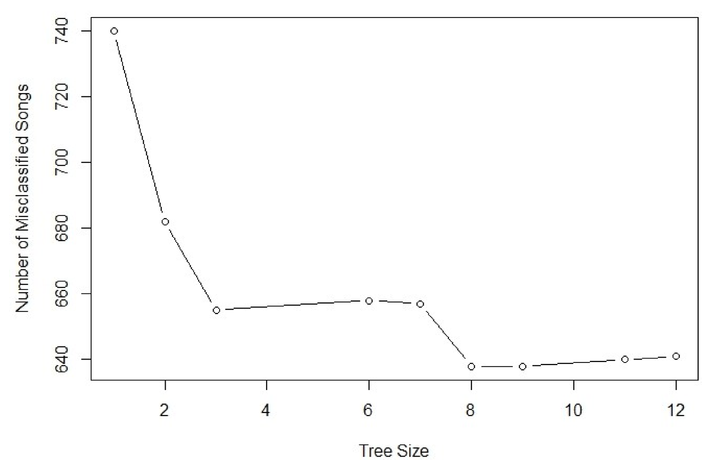

Figure 3. Graph of cross validation, used to find the most optimal tree size.

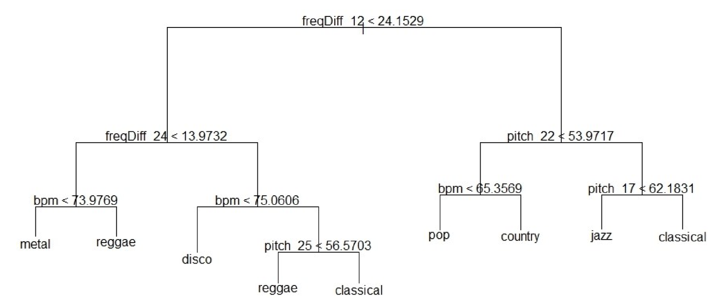

Figure 4. Pruned Decision Tree trimmed to size 9.

- Shortest, most optimal tree after pruning.
- 78.94% misclassification rate

Methods and Results Cont.

2). **Support Vector Machines (SVM)**

- Categorizes new data by using the optimal separating hyperplanes built using training data.
- 71.05% Misclassification Rate

3). **Linear Discriminant Analysis (LDA)**

- Creates classification by searching for a linear combination of predictors that best separate the data.
- 78.95% Misclassification Rate

4). **Quadratic Discriminant Analysis (QDA)**

- A more flexible modification of LDA that assumes each classification category has its own covariance matrix.
- 84.74% Misclassification Rate

Discussion

Summary of Misclassification Rates for all Models

	Decision Tree	SVM	LDA	QDA
Percentage Misclassification Rate	78.95	71.05	78.95	84.74

- Best Machine Learning algorithm was Support Vector Machine
- We hypothesize this is due to the non-linear nature of our collected data.
- Even though LDA and Decision trees are different, for this dataset they produce the same misclassification error rate.
- It makes sense for the error rate of QDA to be greater than that of LDA because QDA is a much more flexible algorithm.

Conclusion/Further Investigation

- Since our algorithms do not produce 100 percent error rate, this shows that they did in fact learn *something.*
- **Future Work**: classification using more complex algorithms that will be more accurate, extract more informative data from the dataset, find a better data trimming method.

References

1). Hamel, Philippe, and Douglas Eck. "Learning Features from Music Audio with Deep Belief Network." (2010): n. pag. International Society for Music Information Retrieval. Web.
2). Jr., Carlos N. Silla, Alessandro Koerich L., and Celso Kaestner A. A. "A Machine Learning Approach to Automatic Music Genre Classification." J. Braz. Comp. Soc. Journal of the Brazilian Computer Society 14.3 (2008): 7-18. Web.
3). Stamatatos, Efstathios, and Gerhard Widmer. "Automatic Identification of Music Performers with Learning Ensembles." *Artificial Intelligence* 165.1 (2005): 37-56. Web.
4). Tzanetakis, George, and Perry Cook. "Musical genre classification of audio signals." *IEEE Transactions on speech and audio processing* 10.5 (2002): 293-302.
5). James, Gareth, et al. *An introduction to statistical learning*. Vol. 6. New York: springer, 2013.

MACALESTER COLLEGE