Genre Classification in Digital Music Libraries

By: Jacob Grinstead

Division of Science and Mathematics
University of Minnesota, Morris
Morris, Minnesota

November 16, 2019

Introduction - How Genre in Music is Classified

When was the last time you searched for new music to listen to?

Introduction - How Genre in Music is Classified

When was the last time you searched for new music to listen to?

- Spotify
- Pandora
- iTunes

Introduction - How Genre in Music is Classified

When was the last time you searched for new music to listen to?

- Spotify
- Pandora
- iTunes

Problem: Manual genre classification is not practical for the size of music databases today

Introduction - How Genre in Music is Classified

When was the last time you searched for new music to listen to?

- Spotify
- Pandora
- iTunes

Problem: Manual genre classification is not practical for the size of music databases today

Solution: Automatic genre classification

Introduction - Past Difficulties in Classification

Challenges in genre classification come in two forms:

- Practical
- Technical

Introduction - Past Difficulties in Classification

Challenges in genre classification come in two forms:

- Practical
    - Genres can be subjective
    - The usefulness of using genres as classification
- Technical

Introduction - Past Difficulties in Classification

Challenges in genre classification come in two forms:

- Practical
    - Genres can be subjective
    - The usefulness of using genres as classification
- Technical
    - Machine learning training time
    - Unreliable or inefficient algorithms

Introduction - Importance of Genre

- People culturally identify with genres
- Users are already accustomed to searching for music via genre
- People use genres more than any other criteria when searching for music recommendations
- Manually entering genre metadata is less practical

Introduction - Diminishing the Technical Challenges

- Training efficiency has been increasing
- Music analysis has become more available and easy to use
- More efficient classification algorithms are being created

Outline

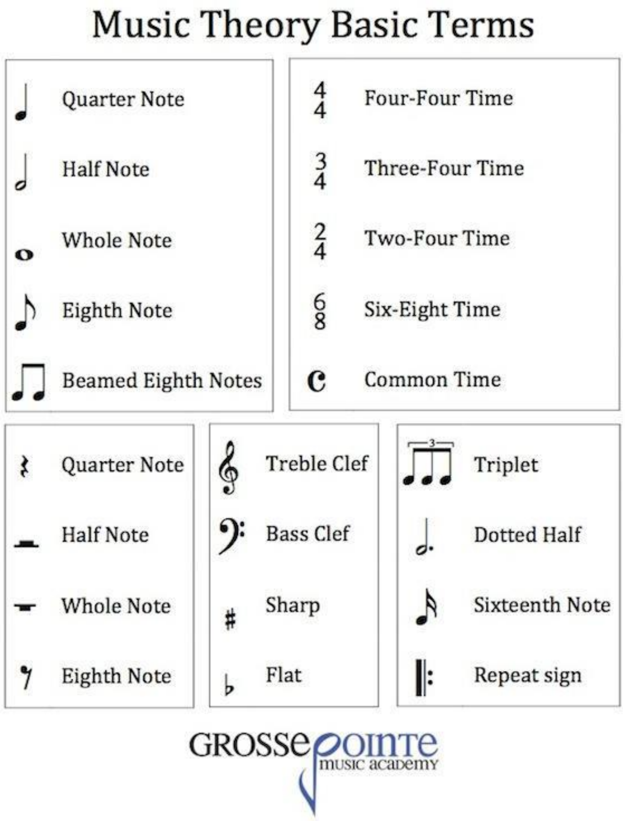
- Background
- Musical Encoding
- Classification Algorithms

Outline - Background

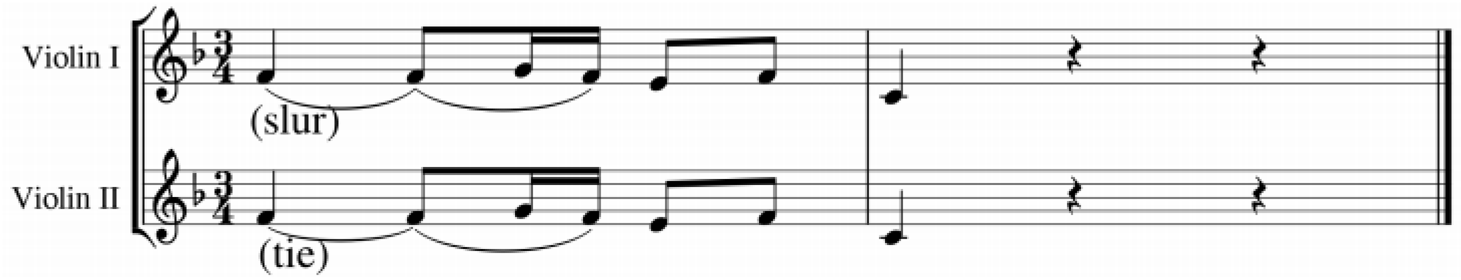
- **Background**
  - **Music Theory**
  - **Essentia**
  - **Machine Learning**
- Musical Encoding
- Classification Algorithms

Background - Music Theory

- Music is primarily made up of notes, ranging from A to G, and rests
- Notes and rests are written on music staffs
- Other symbols and notations are added to make it easier to read for musicians
- These symbols also indicate specific things about a piece



Background - Music Theory



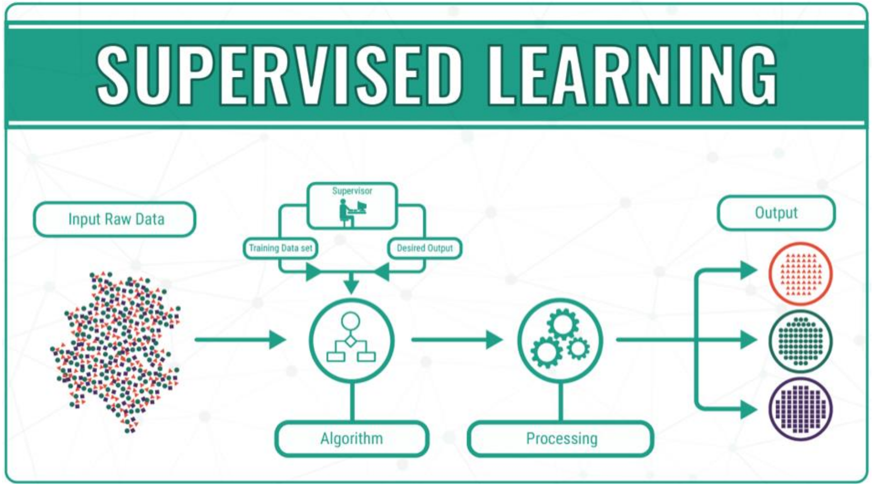
- Two measures from Beethoven’s Op. 18 No. 1
- The flat symbol shows the key of the piece is in F major
- Notice, there is a slight difference in appearance between a slur and a tie

Background - Essentia

- Open-source library for music analysis
- Able to extract different content-based features from a piece
- Optimized for speed

Background - Machine Learning

- Method of data analysis using patterns
- Machine learning algorithms use training data to make predictions
- Supervised learning - training data as input and prediction models as output
- The three classification algorithms later on are all machine learning algorithms



Outline - Musical Encoding

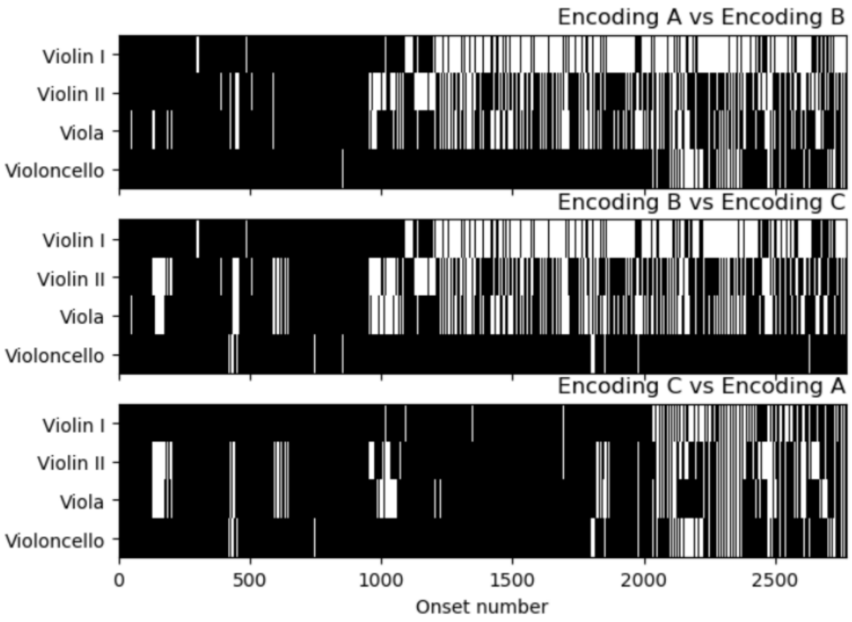
- Background
- **Musical Encoding**
  - **What is it?**
  - **Research - *Encoding Matters***
  - **Results of Research**
- Classification Algorithms

Musical Encoding

- Features of music is displayed via code
- Content-based and image-based representations
- Different file types encode things differently

Musical Encoding - *Encoding Matters*

- Research by Nestor Napoles, Gabriel Vigliensoni, and Ichiro Fujinaga
- Took the same piece from three different encodings
- Used matching note/rest onsets to measure discrepancies between the three pieces

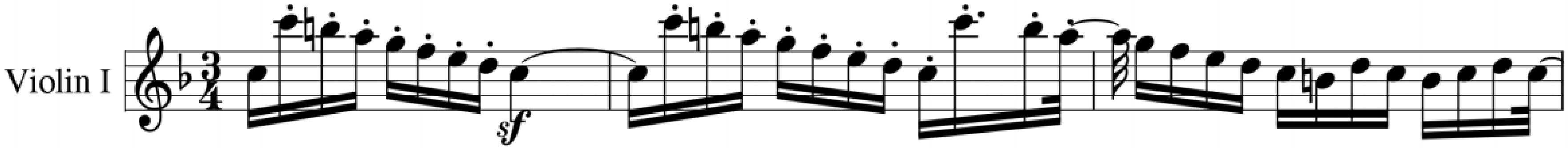


Black represents where note/rest onsets match each other and white represents where they do not match

Musical Encoding - Software Error

- The music notation software allows inconsistent encodings - overcrowded measures

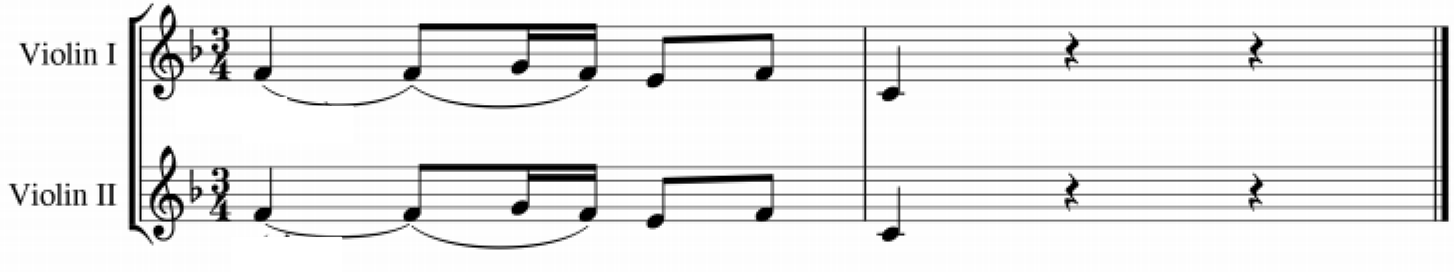
Encoding A - Overcrowded



Encoding B - Correct

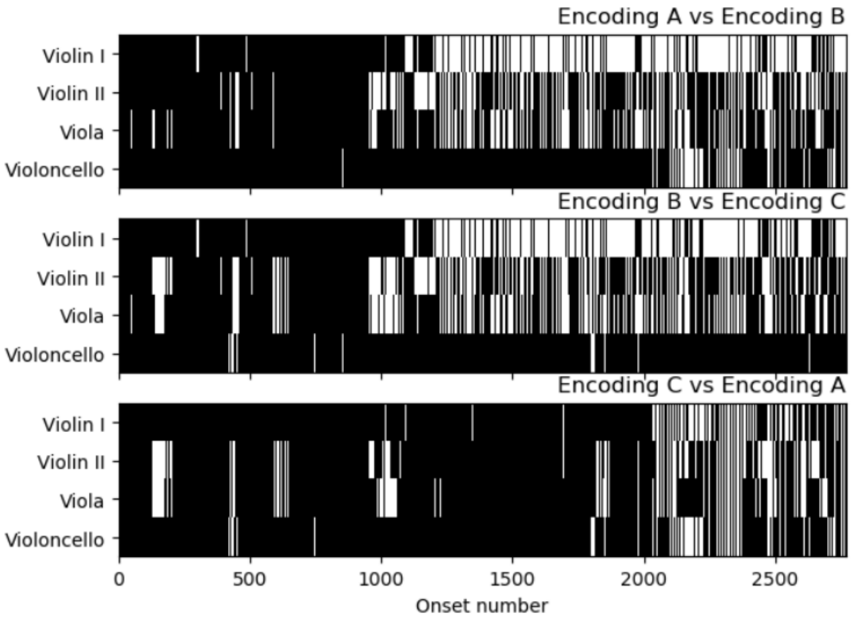
Musical Encoding - Human Error

- Difficulty of seeing the physical differences between pieces



Musical Encoding - Problems to Overcome

- The same piece in different encoded formats can be similar, but not the same
- Many different reasons for discrepancies of the same piece
- Potential for an interesting problem in genre classification



Outline - Classification Algorithms

- Background
- Musical Encoding
- **Classification Algorithms**
    - **Deep Neural Network (DNN)**
    - **ExtraTrees**
    - **XGBoost**
    - **Results**

Classification Algorithms - Background

Study done by Benjamin Murauer and Günther Specht

- Three machine learning classification algorithms
  - Deep Neural Network
  - Extra Trees
  - XGBoost
- Training data set of 25,000 pieces and testing data set of 35,000 pieces

genre	# of songs
Rock	7,103
Electronic	6,314
Experimental	2,251
Hip-Hop	2,201
Folk	1,519
Instrumental	1,350
Pop	1,186
International	1,018
Classical	619
Old-Time / Historic	510
Jazz	384
Country	178
Soul-RnB	154
Spoken	118
Blues	74
Easy Listening	21
total	**25,000**

Classification Algorithms - Background

- They used Essentia to extract features from the pieces of music

- Mean Log Loss Score, L

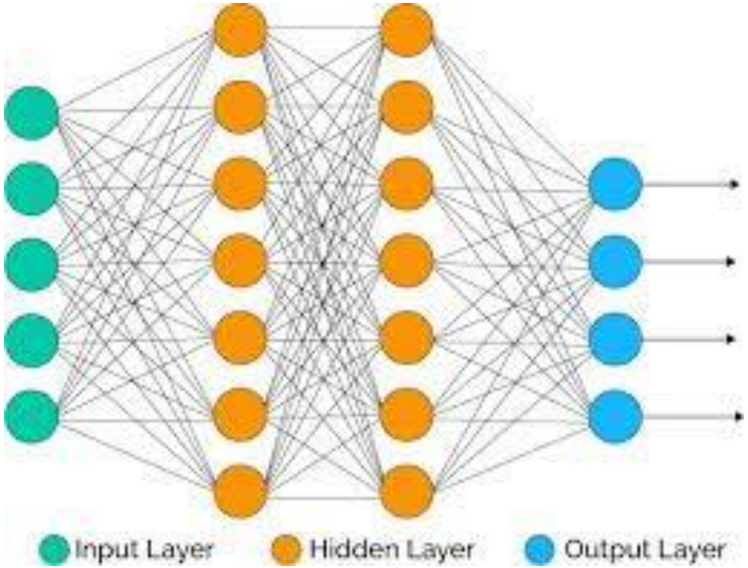
$$L = -\frac{1}{N}\sum_{n=1}^{N}\sum_{c=1}^{C} y_{nc}\ln(p_{nc})$$

feature name	exemplary value
low level average loudness	0.938
low level melbands skewness mean	2.246
low level spectral flux median	0.112
rhythm bpm	83.583
...	...
danceability	1.101
tonal key	'E'
tonal chord	'major'

Classification Algorithms - Neural Networks

Neural Networks are sets of algorithms designed to recognize patterns

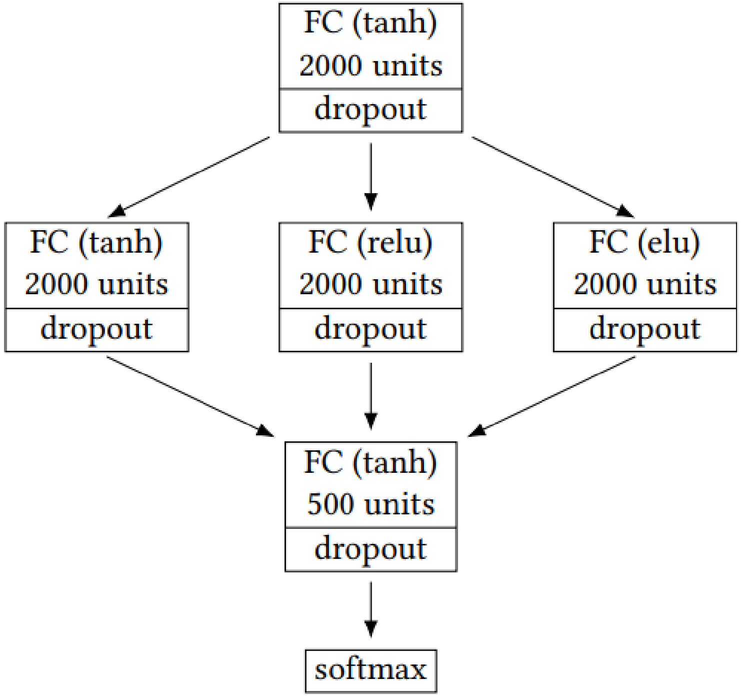
- Input Layer: Numerical data as vectors
- Hidden Layer: Activation functions are performed
- Output Layer: Numerical data



Classification Algorithms - Deep Neural Network (DNN)

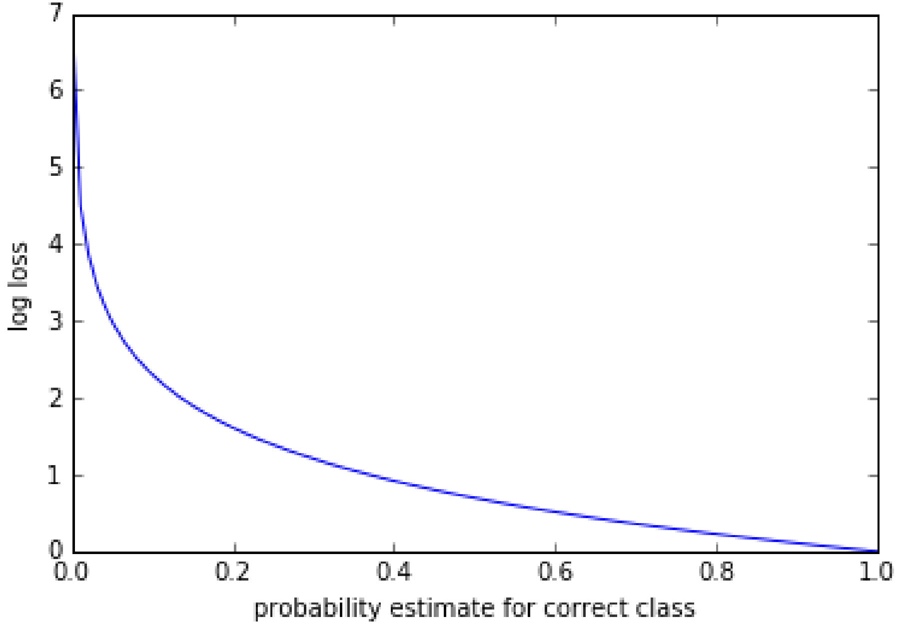
The main difference of a DNN is that they have more than one hidden layer

- Input Layer: Feature values from Essentia
- Hidden Layer: Activation functions
  - tanh
  - relu
  - elu
- Output Layer: Probabilities



Classification Algorithms - DNN Results

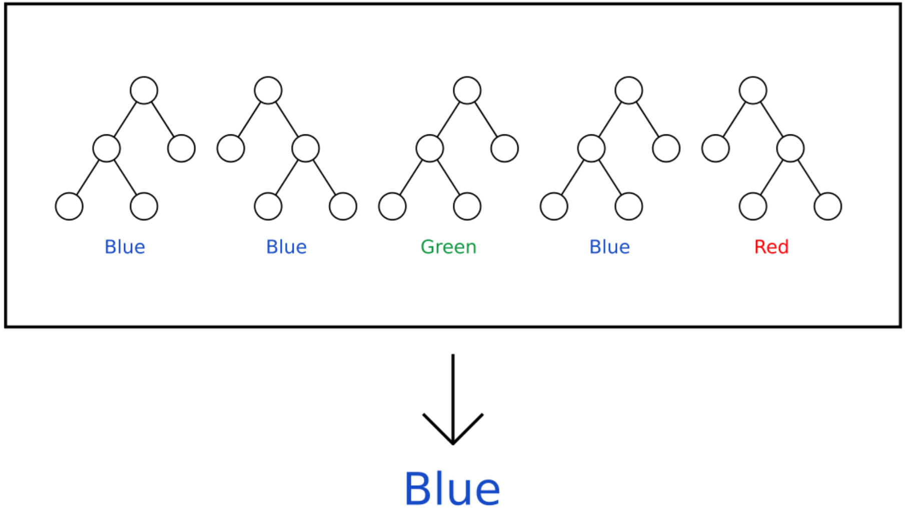
The DNN had a mean log loss score of 1.44



Classification Algorithms - ExtraTrees

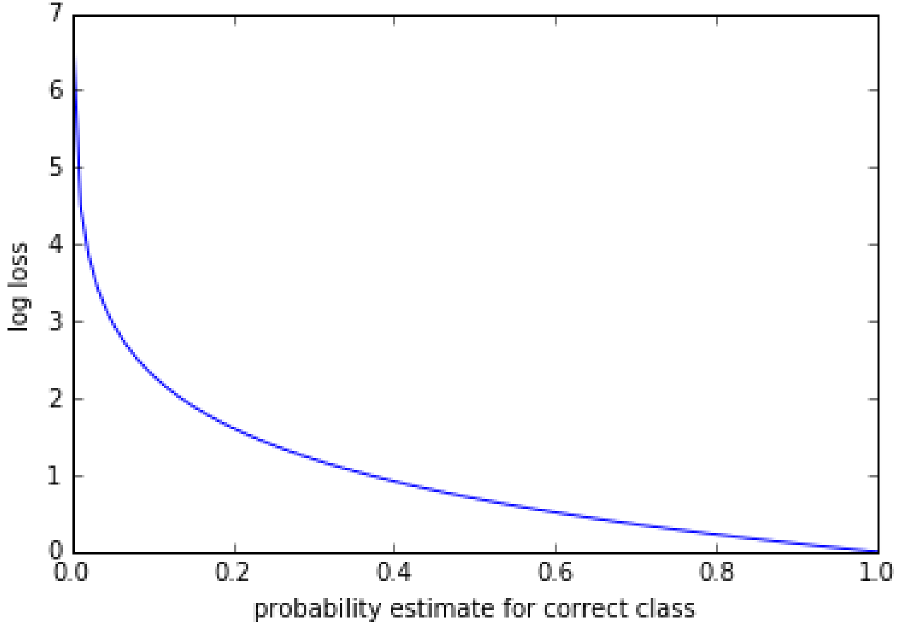
The ExtraTrees classifier algorithm is a variant of the random forest classifier

- Builds an ensemble of decision trees
- Nodes are split randomly
    - Decreased variance, increased bias
- Uses whole training data set to learn from



Classification Algorithms - ExtraTrees Results

The ExtraTrees classifier had a mean log loss score of .92



Classification Algorithms - XGBoost

The XGBoost classifier uses gradient boosting

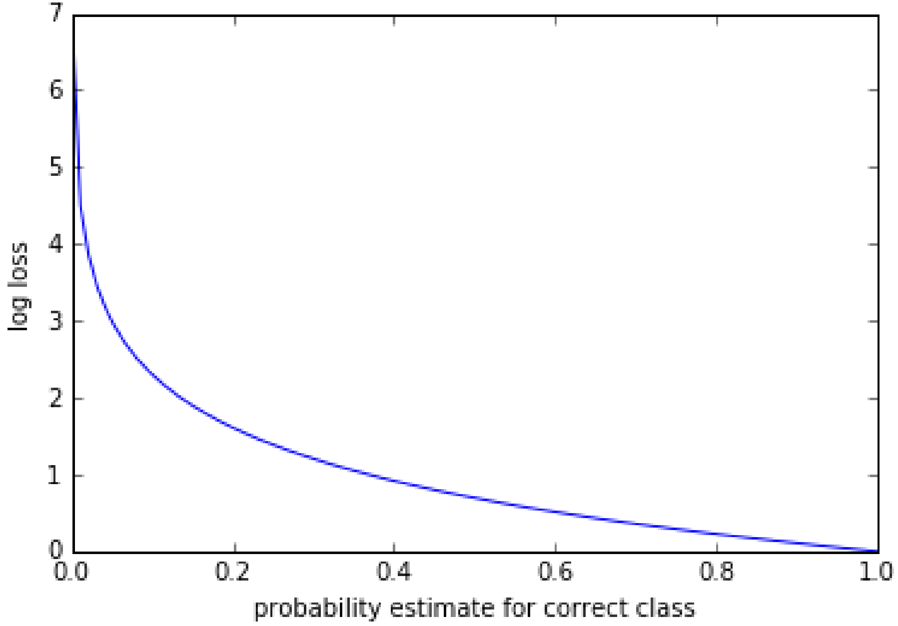
- It also creates an ensemble of decision trees as prediction models
- Aggregates them to create a final prediction

Gradient Boosting

- Uses a gradient descent algorithm
- Produces models that predict errors of previous models to better themselves
- Supports classification predictive modeling problems

Classification Algorithms - XGBoost Results

XGBoost had a mean log loss score of .82



Classification Algorithm - Results

- XGBoost has lowest mean log loss score
- Better than a DNN
- Potential for bias

genre	**# of songs**
Rock	7,103
Electronic	6,314
Experimental	2,251
Hip-Hop	2,201
Folk	1,519
Instrumental	1,350
Pop	1,186
International	1,018
Classical	619
Old-Time / Historic	510
Jazz	384
Country	178
Soul-RnB	154
Spoken	118
Blues	74
Easy Listening	21
total	**25,000**

Conclusion

Putting everything together:

- Different encodings of the same pieces could provide different log loss scores
- Only around a 50% chance of correctly guessing genres

Acknowledgements

- Kristin Lamberty
- Sam Score
- Paul Gans

References

https://medium.com/@chisoftware/supervised-vs-unsupervised-machine-learning-7f26118d5ee6

https://music-encoding.org/

https://medium.com/datadriveninvestor/when-not-to-use-neural-networks-89fb50622429

https://stackoverflow.com/questions/35013822/log-loss-output-is-greater-than-1

https://victorzhou.com/blog/intro-to-random-forests/

http://delivery.acm.org/10.1145/3200000/3191822/p1923-murauer.pdf?ip=146.57.93.60&id=3191822&acc=OPEN&key=70F2FDC0A279768C%2E1626CA105EEA6A29%2E4D4702B0C3E38B35%2E6D218144511F3437&__acm__=1573916539_c31c6787273041f19b1d26c32db8fa7a

http://delivery.acm.org/10.1145/3280000/3273027/p69-napoles.pdf?ip=146.57.93.60&id=3273027&acc=ACTIVE%20SERVICE&key=70F2FDC0A279768C%2E1626CA105EEA6A29%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&__acm__=1573917448_e54e9c33bf3054c436baf3c40e93c780