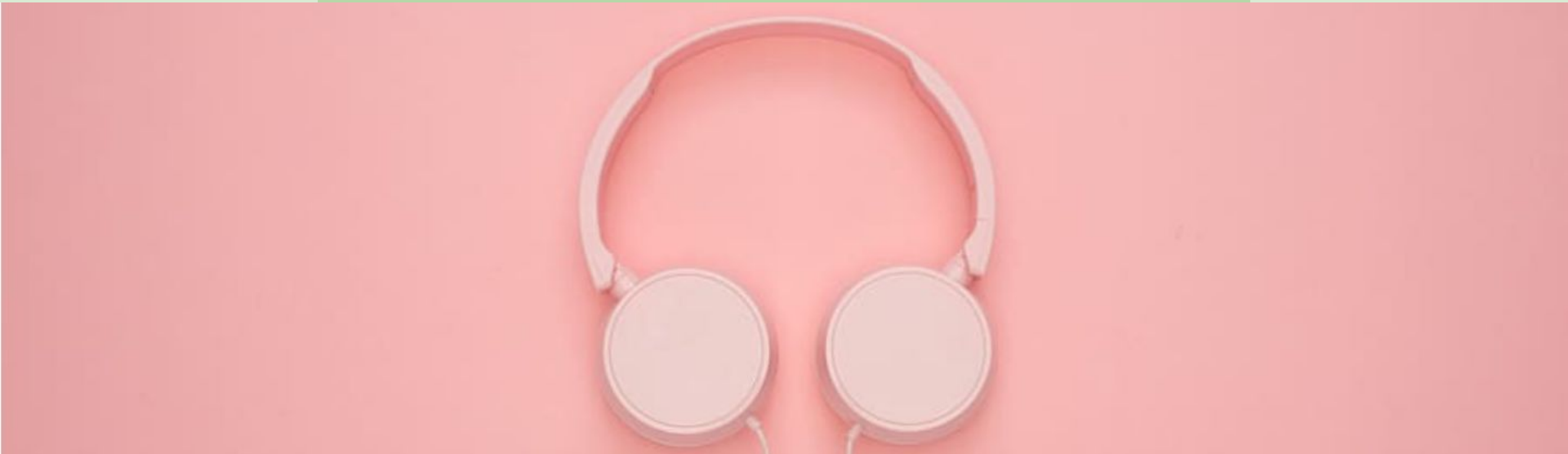
Anatomy of a song

Alex & Inès

Can lyrics define a music genre?



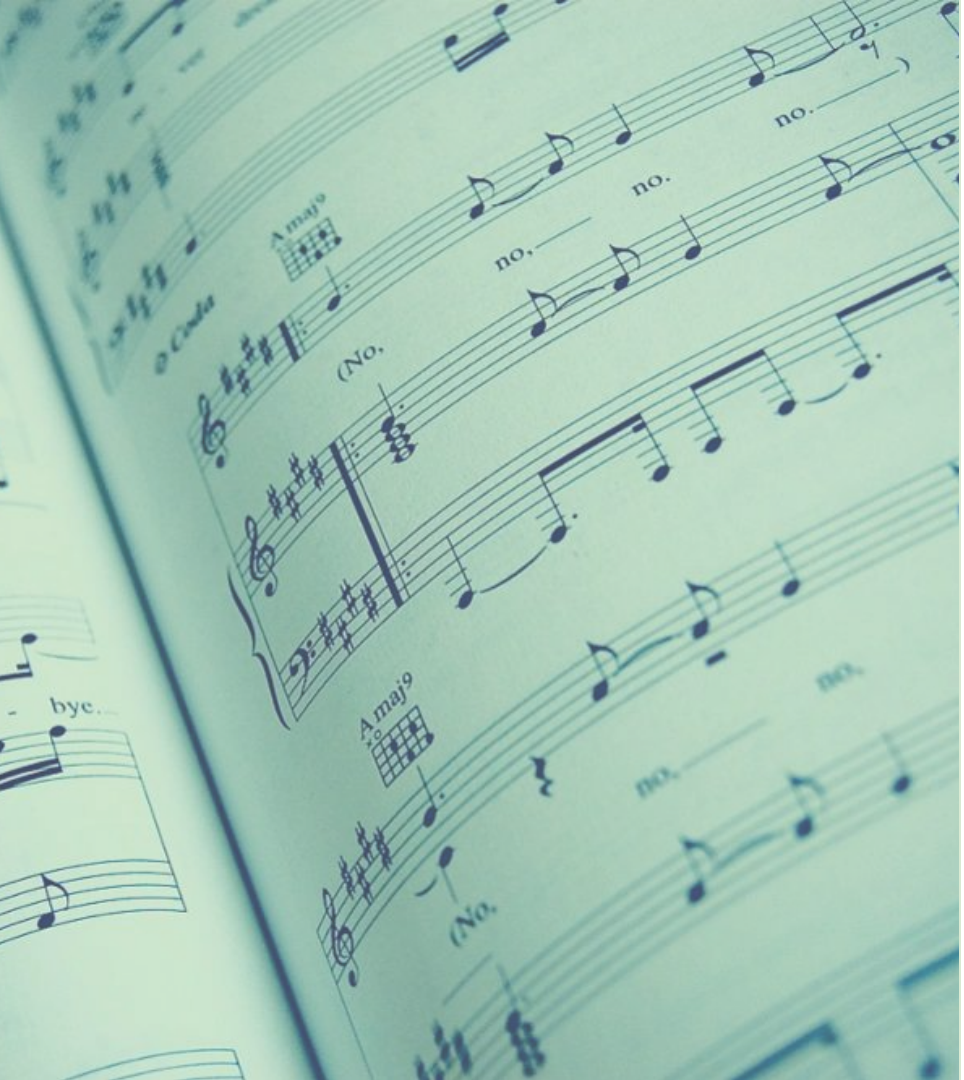


TABLE OF CONTENTS

01

Introduction

Musical Characteristics

Musical Characteristics

Hip Hop

- Heavy bassline
- Flow
- 85 - 115 BPM

Rock

- Strong beat
- Screaming or shouting
- 110 - 140 BPM

Pop

- Simple melody
- Rhythms made to dance
- 100 - 130 BPM

Lyrical Characteristics

Lyrical Characteristics

Hip Hop

- Poverty
- Violence
- Women

Rock

- Drug
- Intimacy
- Rebellion & Freedom

Pop

- Love
- Friendship
- Empowerment

02

Data Preparation

Data Set

Dataset

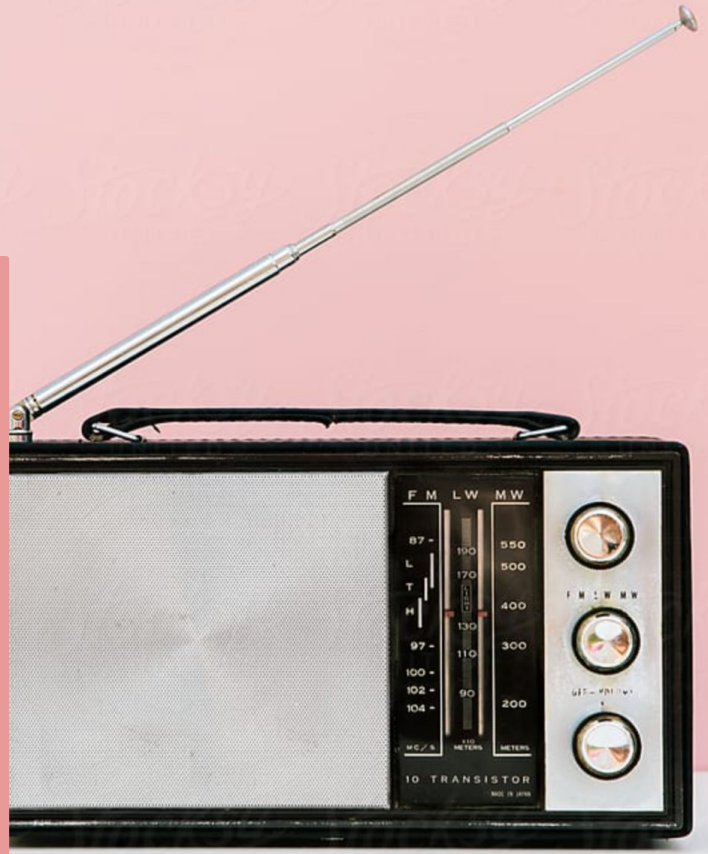
Where?

- Found on Kaggle
- Based off data scraped from a Brazilian music portal (Vagalume)

What?

- 6 music genres
- 2 main languages
- 3242 artists
- 209522 songs

Data Cleaning & Processing



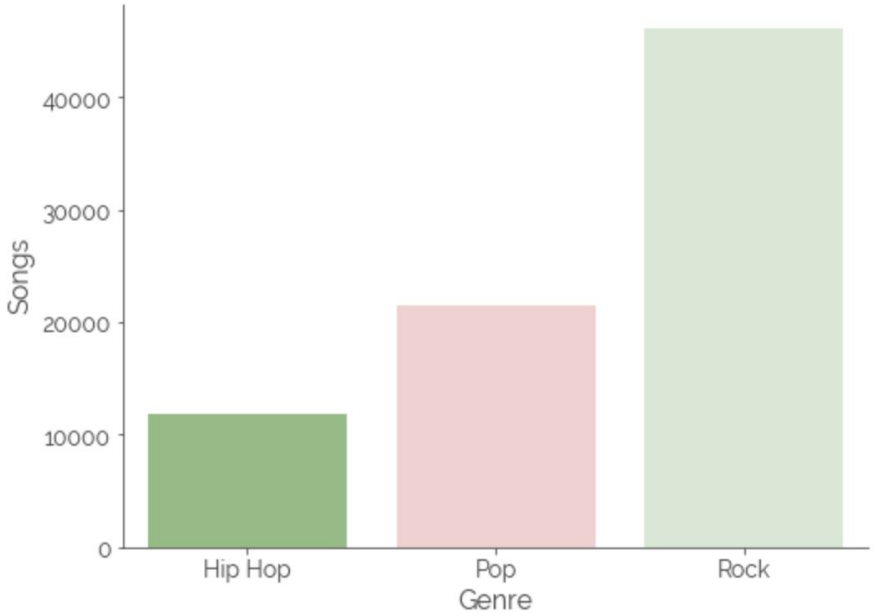
Data cleaning

- Songs in English
- 3 genres: Hip Hop, Pop and Rock
- Artists with a single main genre
- 79452 songs

Genre Repartition

- Hip Hop ~ 15 %
- Pop ~ 27 %
- Rock ~ 58 %

The dataset is unbalanced, in favour of Rock songs.



Natural Language Processing

- Remove and keep track of trailing apostrophes (eg. tryin')
- Remove any brackets and its content (eg. [Chorus])
- Remove multiple letters (eg. yeeeaaah –> yeah)
- Basic cleaning
- Stemming / Lemmatization
- Words vectorization

Hip Hop

- Most common words in Hip Hop music
- Does not include the top 200 words in Pop and Rock
- Top 10 words include: explicit words, hit, cuz'
- 8 of the top 10 words are explicit

Pop

- Most common words in Pop music
- Does not include the top 200 words in Hip Hop and Rock
- Top 10 words include: kiss, touch, beautiful, hurt

Rock

- Most common words in Rock music
- Does not include the top 200 words in Hip Hop and Pop
- Top 10 words include: dead, alive, woman, child

Average number of words

Music Genre	Average number of words
All	123
Hip Hop	233
Pop	124
Rock	94

In average, Hip Hop lyrics have:

- 190 % more words than Pop lyrics
- 250 % more words than Rock lyrics

Rock lyrics have 25 % fewer lyrics than average

Feature Enrichment

Top Ten Words

- 3 binary markers
- 1 for each genre
- Marker = 1 if any word from the 10 most common words for each genre is present in the lyrics.

Trailing Apostrophe

- Track how many there are per song
- In average, 5 times more trailing apostrophes in Hip Hop songs

03

Modelling

Basic models

Random Forest

- GridSearch to find best hyper parameters
- 99 % accuracy on the training test
- Only 75 % accuracy on the testing set
- Overfitting suspected

Logistic Regression - Model Selection

Stemmer

- Porter
- Snowball
- Lancaster

Solver

- Limited-memory BFGS (default)
- Newton CG
- SAGA

Penalizer

- L1
- L2 (default)
- Elastic Net

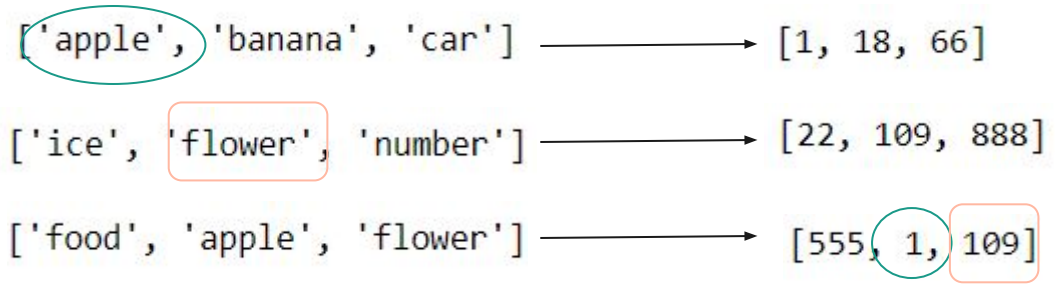
Logistic Regression - Results

- 76 % accuracy on the training set
- 74 % accuracy on the testing set
- 61% of the misclassified songs were Pop songs

Neural Networks

Sequences

What is a sequential approach to text transformation?



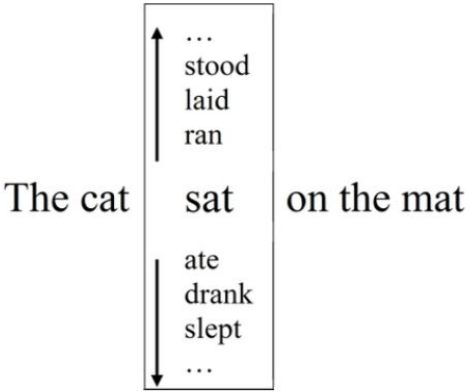
Advantages:

- Interpretability
- Count and evidence
- Streamline ease

RNN

What is RNN?

- Type of NN exhibiting dynamic, bidirectional behaviour where nodes can exchange information back and forth to reinforce learning.
- Requires maintaining logs of past epochs in order to travel back 'in time'.



Why RNN?

- Why not Word2Vec, N-grams, Skip-grams?

There is no *context* to this sentence

Multiple Layer Perceptron (MLP)

What is *SLP*? - Single Layer Perceptron Neural Network

- Simplest form of NN: One input, one output and one hidden layer.
- Implicit forward propagation in its basic format
- Limited choices for activation and other hyperparameters

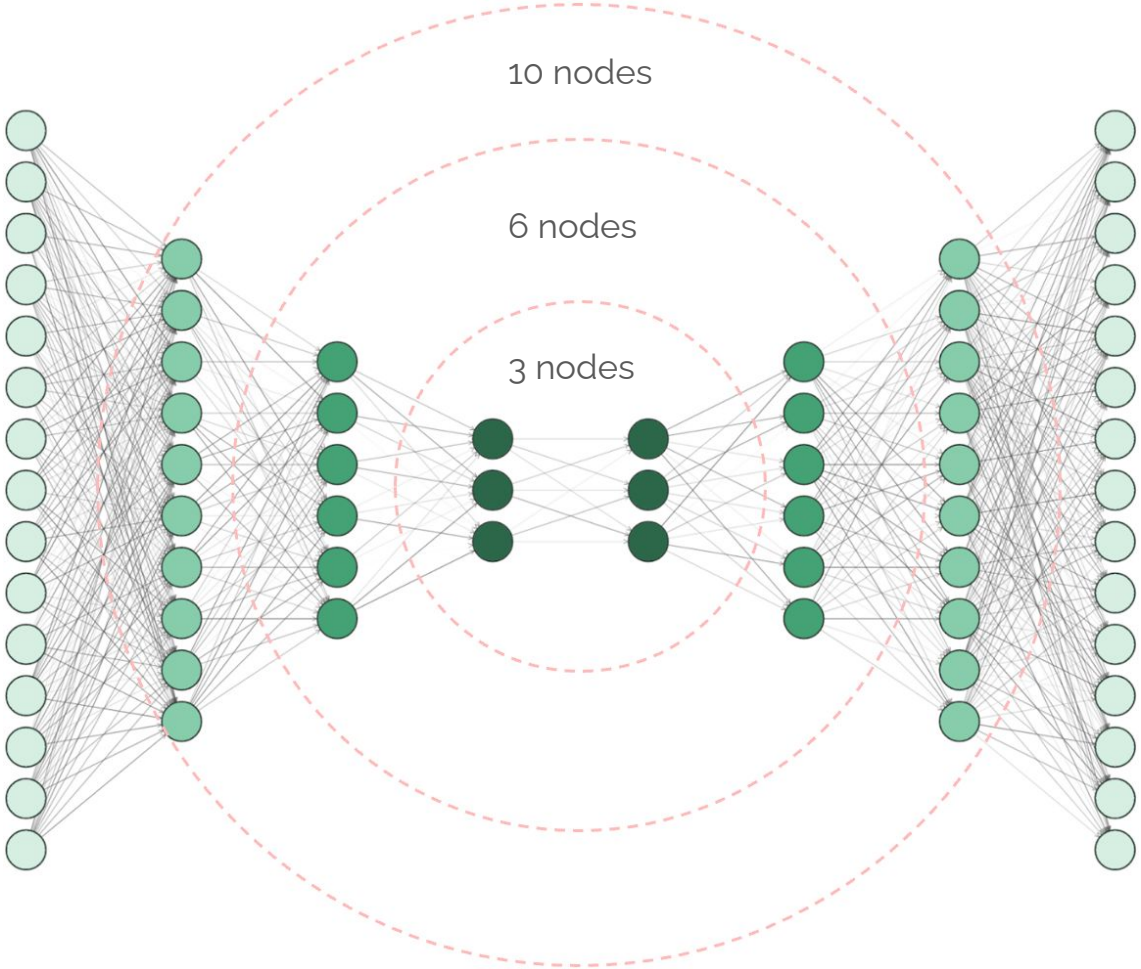
MLP = More layers!

- Free choice of architecture
- Multiple hyperparameter possibilities
- Different activations for each layer

MLP - parameters

Hyperparameter selection:

1. Activation function:
   - ReLu (Rectified Linear Unit) : f(x) = max(0,x)
   - Logistic: f(x) = 1/(1+exp(-x))
   - Tanh: f(x) = tanh(x)
2. Learning rate - Constant, adaptive, inverse
3. Solver - Adam, LBFGS
4. Architecture - symmetrical, inverse symmetrical (AE-like), stiff, diverse, randomized, …


10 nodes
6 nodes
3 nodes

Data formatting

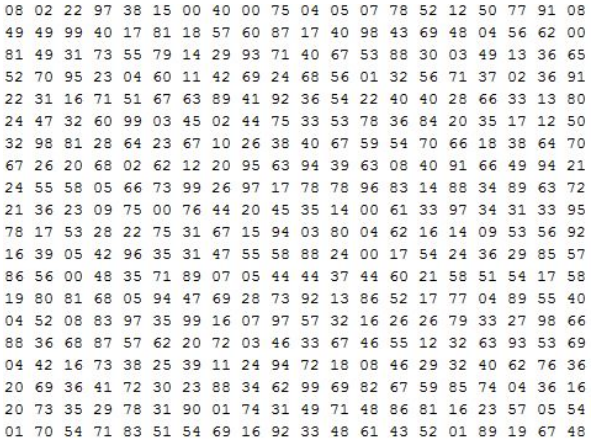
Format required: NxM array

Problem: Not all songs have the same number of words!

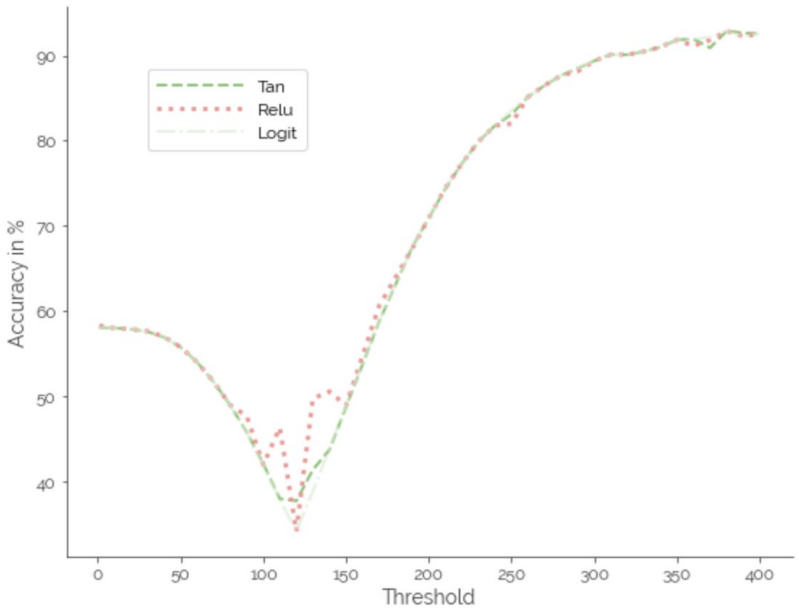
Solution:

- Use only songs with a number of words above a chosen threshold, M.
- For those songs, use only the first M words.

Result: We generate a performance spectrum based on the choice of thresholds.



Performance analysis across thresholds



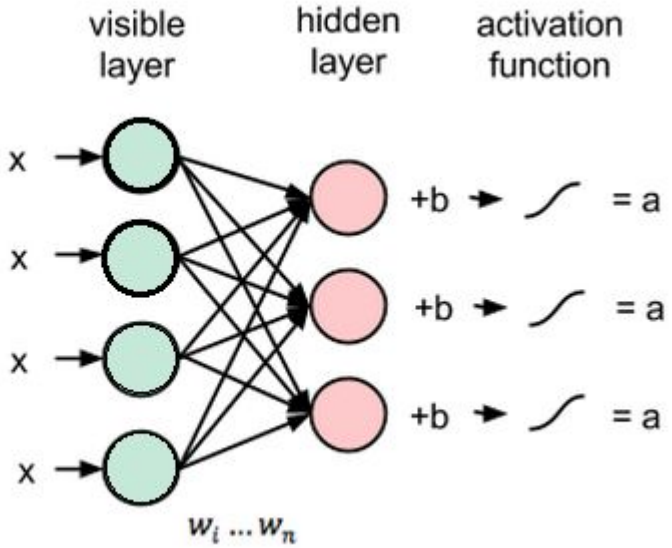
Behaviour:

- Constant ~58% until roughly M=7-8.
- Drop to minimum of ~36% accuracy at around M=120.
- Spike increase up to ~82% accuracy until the M=260 mark.
- Slower, continuous increase to maximum recorded of ~92% around M= 380 - 400

Why? More here.

Results

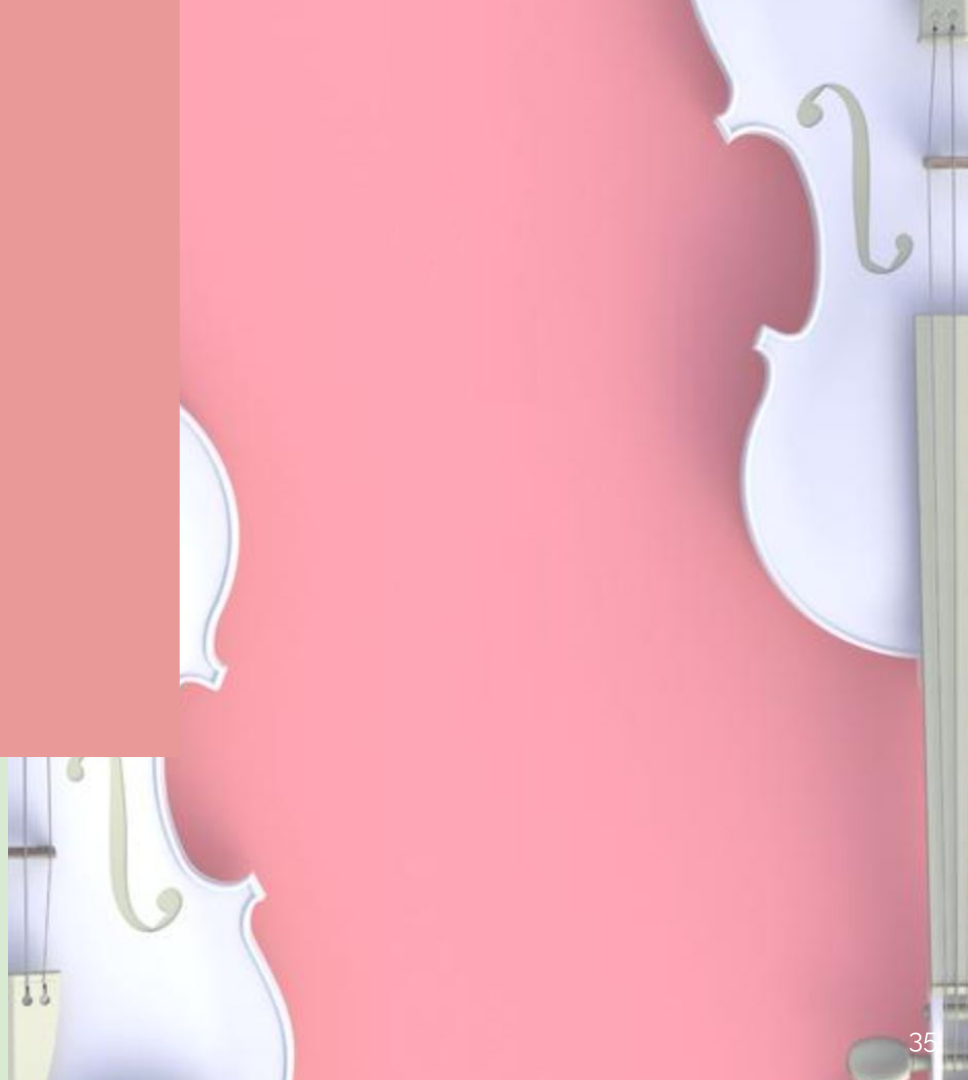
- MLP performs best on larger songs
- The model can be improved using *more data.*
- Other possible models:
  - RBM: Useful in both topic modelling and classification
  - Note: NO topic modelling.



04

Conclusion

Next Steps



Models Improvement

Feature Enrichment

- Check for same words repetition within lyrics

Logistic Regression

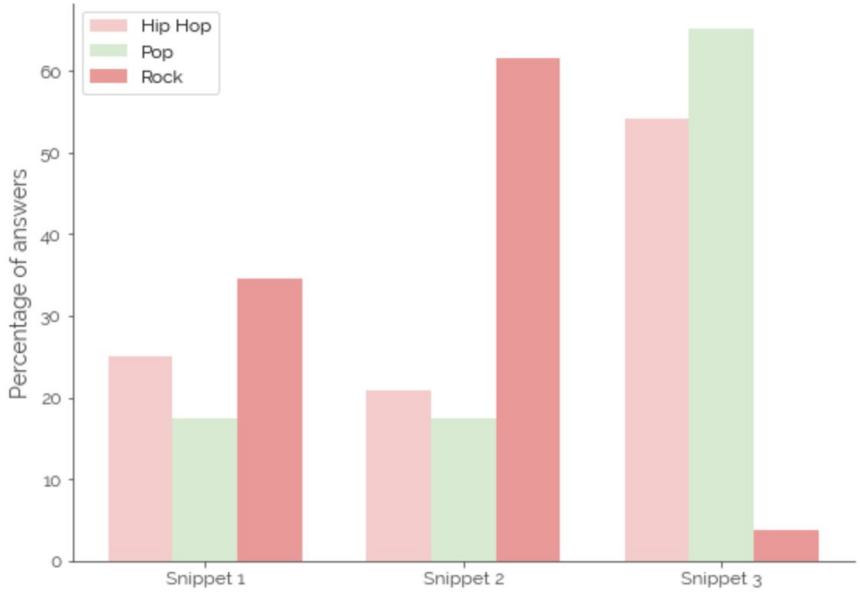
- Error analysis: 61% of the misclassified songs were pop songs

Neural Networks

- More data

Demonstration

Did you beat the machine?



Thank you