Folk Music as a Natural Language

Kevin Zecchini
March 31, 2016

Motivation:

Can we determine the musical influence of a song?

- Songs constantly build on each other over time
- Interested in predicting origin based on similarities in musical content

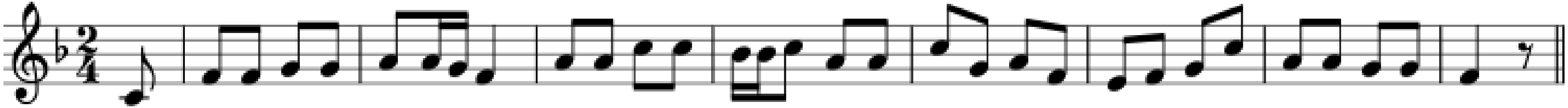
Approach:

- Treat sheet-music as a natural language
- Develop a model which can predict the origin of music
- Use n-gram and ‘bag-of-notes’ analysis as features

Data

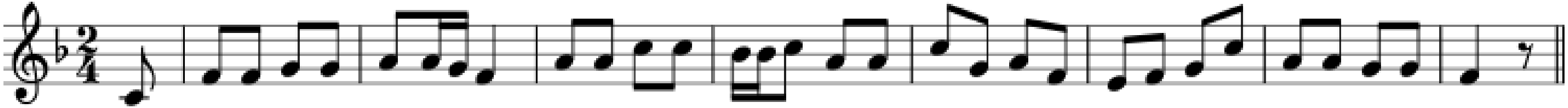
- Scraped approximately 10,000 folk songs stored in individual .krn files:
  - 5365 German (11 sub-regions)
  - 2276 Chinese (4 sub-regions)
  - 619 Luxembourg (4 sub-regions)
  - 339 French (2 sub-regions)
  - 153 UK
  - 373 Native American (3 sub-regions)
  - 152 Canadian
  - 136 USA
  - 848 Other European songs (Russian, Italian, Polish, Danish, Swiss, etc.)

NLP Features:

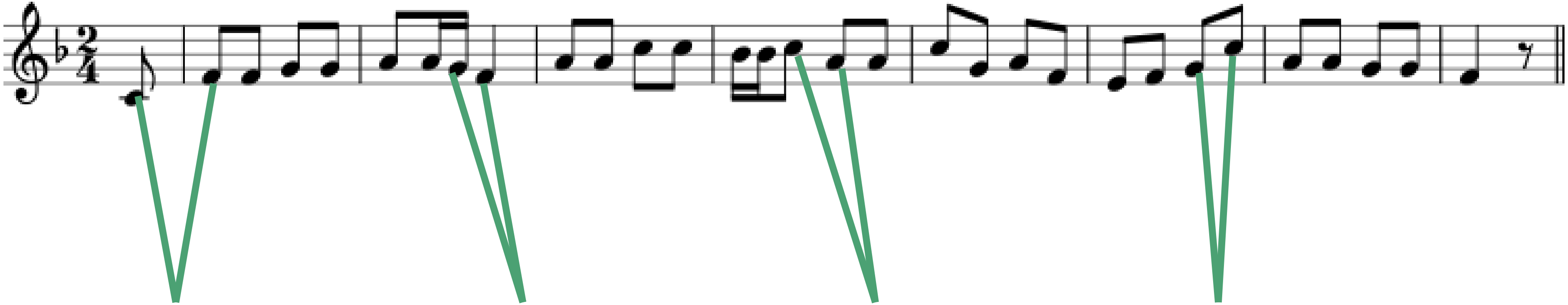


- Music can be transposed to various keys and played with various tempos - focus on relativity:
  - Relative pitch - intervals between notes
  - Relative duration - the equivalent quarter note value
- Remove ‘stopwords’ and ‘punctuation’:
  - Take away structure in the song
  - Remove measures/rests
  - Simplify chords to root note

NLP Features:



NLP Features:



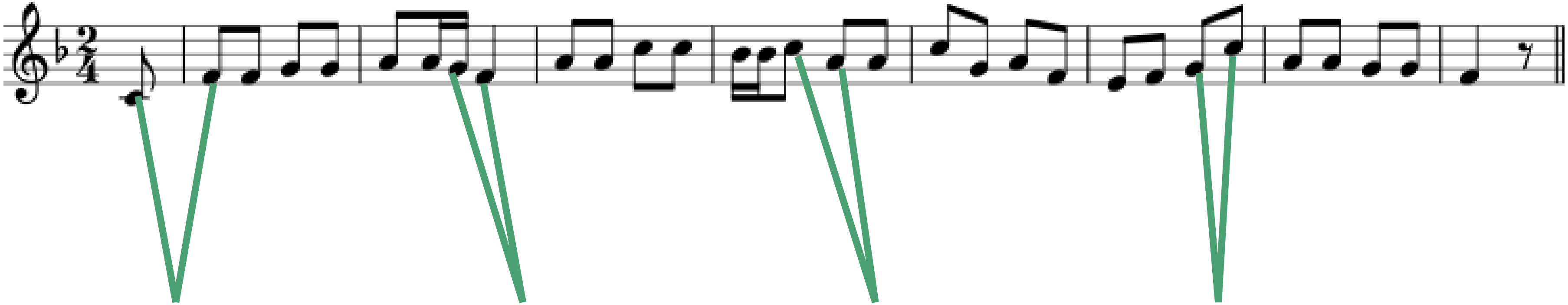
Interval: +5
Duration: (0.5, 0.5)

Interval: -2
Duration: (0.25, 1)

Interval: -3
Duration: (0.5, 0.5)

Interval: +5
Duration: (0.5, 0.5)

NLP Features:



Interval: +5
Duration: (0.5, 0.5)

Interval: -2
Duration: (0.25, 1)

Interval: -3
Duration: (0.5, 0.5)

Interval: +5
Duration: (0.5, 0.5)

Create three separate tf-idf feature sets with various n-gram length iterations:

- Only using intervals
- Only using duration
- Based both on interval and duration

Generating data - Markov Model:

- Highly unbalanced data
- Used Markov model to generate additional music of a given label
- Given n-gram number, will use that as the state to predict the next note
- Probability of next note determined by probability over the label’s corpus

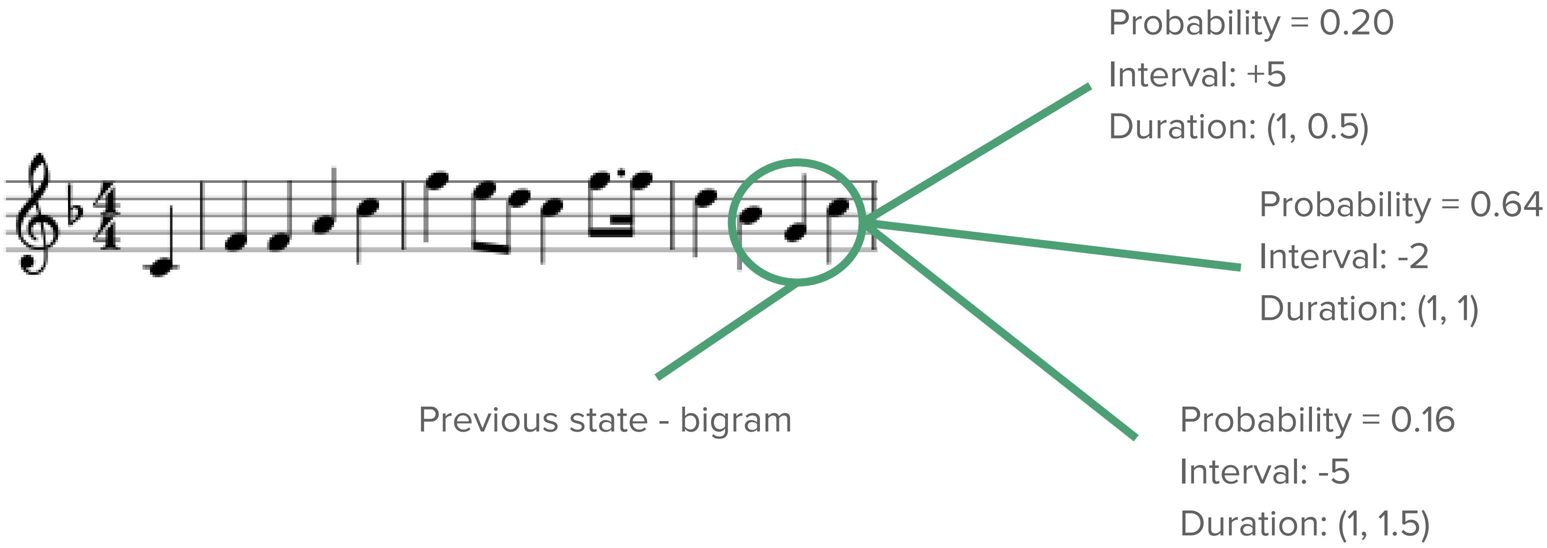
Generating data - Markov Model:

- Highly unbalanced data
- Used Markov model to generate additional music of a given label
- Given n-gram number, will use that as the state to predict the next note
- Probability of next note determined by probability over the label’s corpus

Previous state - bigram

Generating data - Markov Model:

- Highly unbalanced data
- Used Markov model to generate additional music of a given label
- Given n-gram number, will use that as the state to predict the next note
- Probability of next note determined by probability over the label's corpus



Final Model and Results:

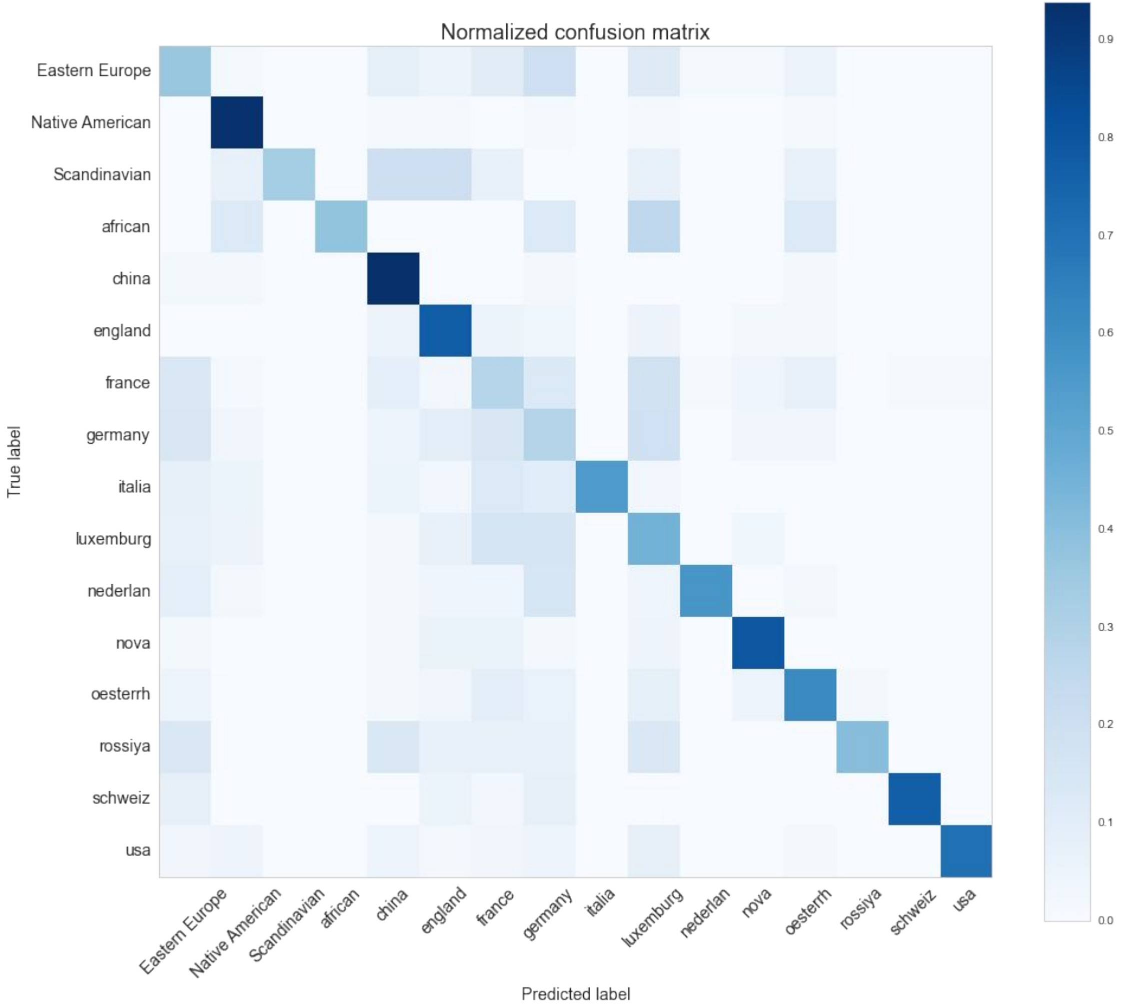
- Condensed labels into 16 representative regions
- Three distinct feature sets searched over n-gram range 1-6
- Best model takes into account both pitch and duration
- Logistic Regression - 61.4% accuracy based on interval and duration for n-gram range 3-4

Future Work:

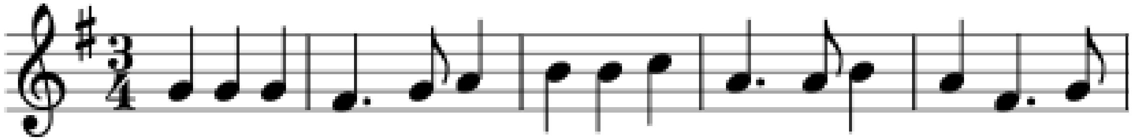
- Use a note2vec analysis in place of bag-of-notes approach
- Expand to music/scores with additional voicings
- Create similarity metric between songs:
    - Use this similarity score in a recommendation engine/search engine
- Use this approach to predict composer or additional characteristics of music
    - Complex sequences in classical music

Thank You!

kzecchini@gmail.com
github.com/kzecchini
linkedin.com/in/kzecchini


Normalized confusion matrix
True label
Predicted label
Eastern Europe
Native American
Scandinavian
african
china
england
france
germany
italia
luxemburg
nederlan
nova
oesterrh
rossiya
schweiz
usa
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0

Example data:



**kern
*M3/4
*k[f#]
*G:
=1
{4g
4g
4g
=2
4.f#
8g
4a}
=3
{4b
4b
4cc
=4
4.a
8a
4b}
=5
{4a
4.f#
8g

- Many aspects to music - key, time signature, melody, harmony, lyrics etc.
- Simplify to one voicing
- Construct nlp features based on some basic characteristics - pitch and duration