Generating Groove: Predicting Jazz Harmonization

Nicholas Bien[1] (nbien@stanford.edu) and Lincoln Valdez[2] (lincolnv@stanford.edu

[1] Department of Computer Science, Stanford University, [2] Department of Computer Science, Stanford University

Motivation

We aim to generate an appropriate jazz chord progression for a given melody. As musical minds, we wanted to see if it is possible to create a model that can make any melody sound jazzy. Classical theory is very predictable. Prior research has generated chord harmonizations for classical and pop music with high accuracy.[1] Jazz theory is a lot less formulaic, so predicting jazz progressions seemed a lot more challenging.

Dataset and Feature Mapping

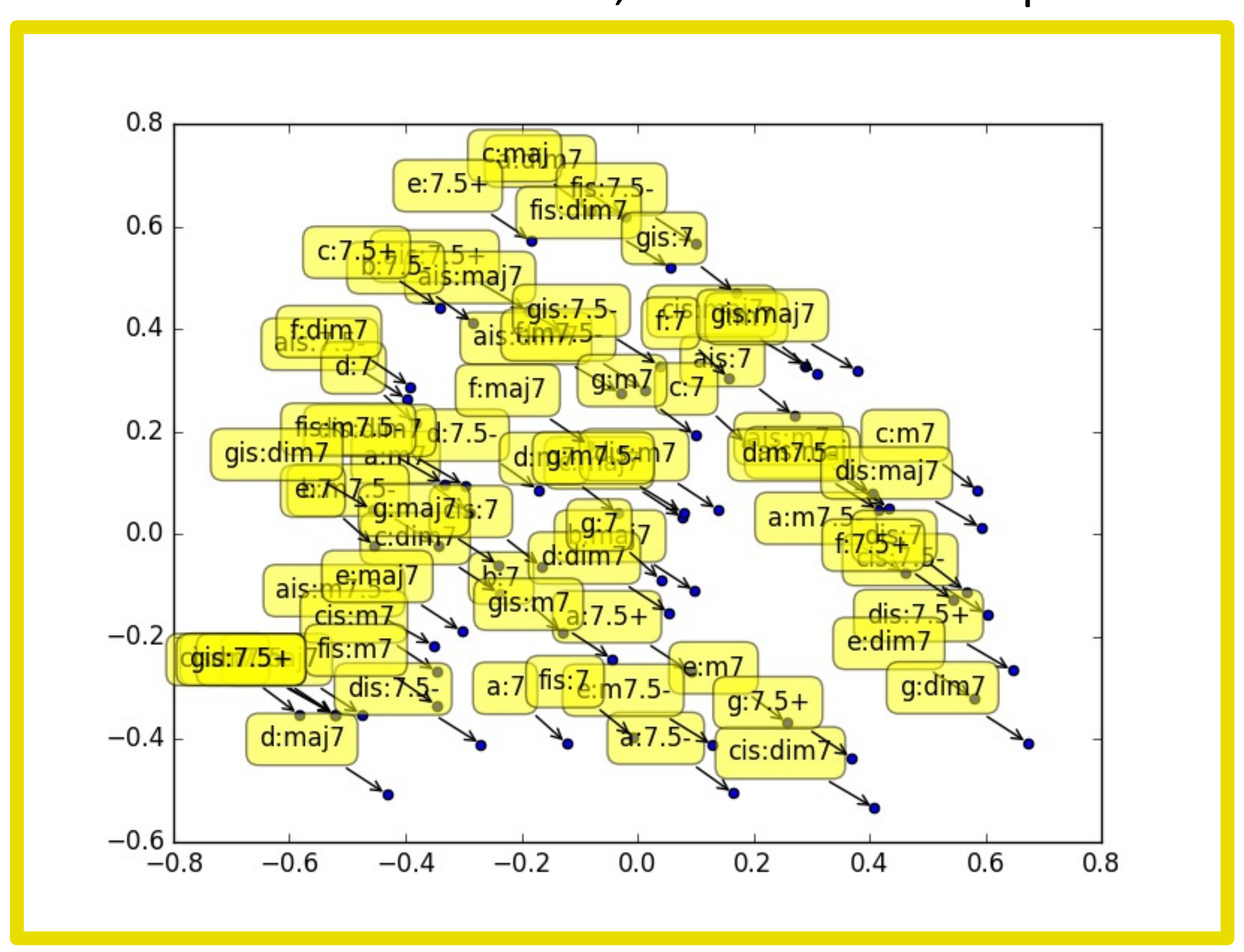
Our dataset was a set of 149 "Jazz Standards" encoded in Lilypond musical transcription format. Each song in the dataset consisted of a melody line with accompanying chord symbols. We transposed each song to the key of C major / A minor. For each chord symbol, we count the number of times each of the 12 notes in the chromatic scale occurs to obtain our features.

Clustering Algorithms

There were 246 chord types in our dataset. To facilitate classification, we explored methods for grouping chords into similar clusters.

K-means and Brown Clustering

We first tried using K-means clustering with the normalized note counts as features and seeing which chords centered around the initialized centroids. Our second method was Brown clustering, which is normally used for the semantics of words in sentences based on their context. Instead, we used it on sequences of chords.



PCA of aggregated note features after chord clustering

Clustering Results

Combining the results of the two algorithms, we observed that chords where notes were stacked on top of the 7th were clustered together, and that 7th chords had similar functions to their respective triads and 6ths. Our clustering efforts brought the number of chord classes down from 246 to 85.

Classification Algorithms

Naïve Bayes

For our first chord prediction algorithm, we implemented a one-off Laplace-smoothed Naïve Bayes classifier. The posterior likelihood of each class was calculated:

$$p(C_k \mid \mathbf{x}) = \frac{p(C_k)\, p(\mathbf{x} \mid C_k)}{p(\mathbf{x})}$$

The results were promising: though the accuracy wasn't that high, it was above baseline, and the actual chords predicted for a given melody were realistic.

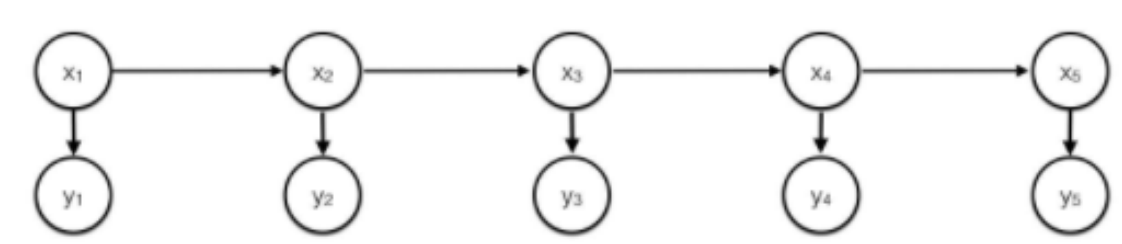
Support Vector Classifier

Our second method was a multi-class support vector classifier. Input vectors were 12-dimensional note counts. We used a radial bias function kernel:

$$K(\mathbf{x}, \mathbf{x}') = \exp(-\gamma \|\mathbf{x} - \mathbf{x}'\|^2)$$

The SVC outperformed Naïve Bayes and seemed to capture the harmonic content of the melody, especially when many notes were present in a feature vector.

Hidden Markov Model



Lastly, we tried incorporating a hidden Markov model with our SVC to capture relationships between chords.[2] For this task, we augmented the dataset with START and END tokens. Transition probabilities were calculated from bigrams in the training set. We used SVC outputs as noisy estimators of the ground truth chord to obtain emission probabilities. We then used the Viterbi algorithm to obtain the maximum likelihood sequence of hidden states (chords classes) given the observed states (SVC predictions).

Results

For testing, we used 80/20 cross validation (split by song). Depending on the run, this amounted to ~6200 chords in the training set and ~1500 chords in the test set.

Algorithm	% Accuracy
Baseline	17.6%
Naïve Bayes	26.2%
Support Vector Classifier	28.3%
Support Vector Classifier w/ HMM	27.4%

All algorithms easily outperformed the baseline. While the numbers could certainly be improved by reducing the number of classes even further, we were pleased with the aesthetic quality of the harmonizations produced. In particular, incorporating a Hidden Markov Model with the SVM lowered overall accuracy slightly but produced more canonical jazz progressions, even for non-jazz music.

Future Directions

What we learned from this project is that the problem we were trying to tackle is a lot more difficult than we initially thought. We likely could have achieved much higher accuracy with the same approach on classical music, since jazz music is far too expansive given the limited data set we had access to. Possible future approaches could be:

- Expanding research into the neural network realm[3]
- Obtaining/creating more data
- Trying the experiment first on classical theory, tuning, refining and then extrapolating to jazz

References

[1] Simon, Ian, Dan Morris, and Sumit Basu. "MySong: Automatic Accompaniment Generation for Vocal Melodies." CHI Proceedings. 2008.

[2] Cunha, Uraquitan Sidney and Geber Ramalho. "An Intelligent Hybrid Model for Chord Prediction." Organised Sound 4(2), pp. 115-119. 1999.

[3] Gang, Dan, Daniel Lehmann, and Naftali Wagner. "Tuning a Neural Network for Harmonizing Melodies in Real-Time." 1998.

[4] Paiement, Jean-François, Douglas Eck, and Samy Bengio. "Probabilistic Melodic Harmonization." LNAI 4013, pp. 218-229. 2006.