A CORRELATED TOPIC MODEL OF STRUCTURAL SCHEMATA IN POPULAR MUSIC

Dakota Killpack

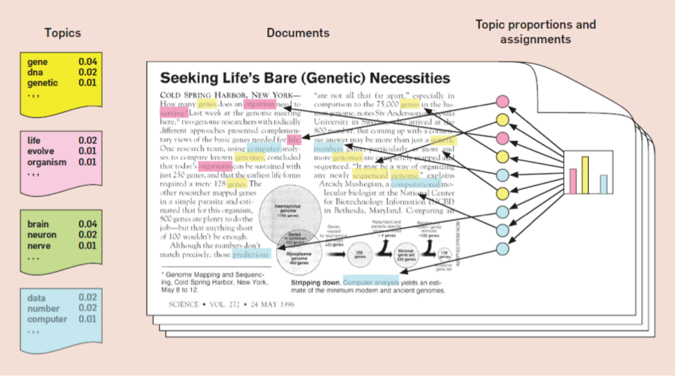
September 5, 2015

Princeton University

LATENT DIRICHLET ALLOCATION: AN OVERVIEW

LDA: Overview

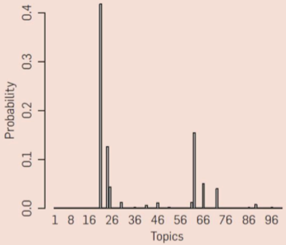
Figure 1. The intuitions behind latent Dirichlet allocation. We assume that some number of "topics," which are distributions over words, exist for the whole collection (far left). Each document is assumed to be generated as follows. First choose a distribution over the topics (the histogram at right); then, for each word, choose a topic assignment (the colored coins) and choose the word from the corresponding topic. The topics and topic assignments in this figure are illustrative—they are not fit from real data. See Figure 2 for topics fit from data.



[1]Blei, David M. "Probabilistic topic models." Communications of the ACM 55.4 (2012): 77-84

LDA: Overview

Figure 2. Real inference with LDA. We fit a 100-topic LDA model to 17,000 articles from the journal *Science*. At left are the inferred topic proportions for the example article in Figure 1. At right are the top 15 most frequent words from the most frequent topics found in this article.



"Genetics"	"Evolution"	"Disease"	"Computers"
human	evolution	disease	computer
genome	evolutionary	host	models
dna	species	bacteria	information
genetic	organisms	diseases	data
genes	life	resistance	computers
sequence	origin	bacterial	system
gene	biology	new	network
molecular	groups	strains	systems
sequencing	phylogenetic	control	model
map	living	infectious	parallel
information	diversity	malaria	methods
genetics	group	parasite	networks
mapping	new	parasites	software
project	two	united	new
sequences	common	tuberculosis	simulations

[1] Blei, David M. "Probabilistic topic models." Communications of the ACM 55.4 (2012): 77-84

STATISTICAL PROPERTIES OF MUSIC AND TEXT

From words to chords

In their attempts to codify patterns of harmony in pop/rock music, music theorists have devised a handful of systems that would be good candidates for ”topics”.

From words to chords

Walter Everett: Rock uses a continuum of tonal systems, spanning:

- ”Conservative” major/minor diatonic harmonies (e.g. Billy Joel, The Beatles)
- ”Conservative” with modal influences
- Blues-based harmonies relying on the pentatonic scale
- Major triads based on the minor pentatonic scale (e.g. heavy metal)

[1] Everett, Walter. “Making Sense of Rock’s Tonal Systems.” Music Theory Online 10.4 (2004): n. pag. Web. 4 Jan. 2010.

From words to chords

Nicole Biamonte: different ways of realizing underlying patterns

- Identifies several sets of chord pop/rock progressions that serve the same function as classical chord progressions, but using different musical materials
- Creates a typology of chord progression styles
- These styles, or 'microlanguages', are good topic candidates

[1] Biamonte, Nicole. "Triadic Modal and Pentatonic Patterns in Rock Music." Music theory spectrum: The journal of the Society for Music Theory 32.2 (2010): 95–110.

From words to chords

Corpus studies confirm the intuition that there are multiple 'microstyles' in rock, e.g. Temperley + De Clerq's corpus analysis:

- Two main harmonic profiles in rock
- A flat-side (blues/pentatonic) cluster of chords and a sharp-side cluster of chords (common practice/diatonic)
  - $\flat VI, \flat VII, \flat III$ vs. ii, VI, III
- Metal is almost a third cluster, and has a weak correlation between $\flat II$, $\sharp IV$, and a lack of V

[1] De Clercq, Trevor, and David Temperley. "A corpus analysis of rock harmony." Popular Music 30.01 (2011): 47-70.

From words to chords

Such a definition of topics has a cognitive analogue: listeners have distinct statistical profiles for harmonic expectation in different musical styles.

Bryn Hughes:

- When primed with classical music, trained musicians are more sensitive to out-of-key progressions
- Harmonic rhythm plays a significant role in expectations
- Listeners expect recurring chord patterns
- Listeners have different sets of expectations for rhythmic different positions in 12-bar blues schemata

[1] Hughes, Bryn, "Harmonic Expectation in Twelve-Bar Blues Progressions" (2011). Electronic Theses, Treatises and Dissertations. Paper 3663.

CORPUS

Billboard Annotations

McGill Billboard Corpus

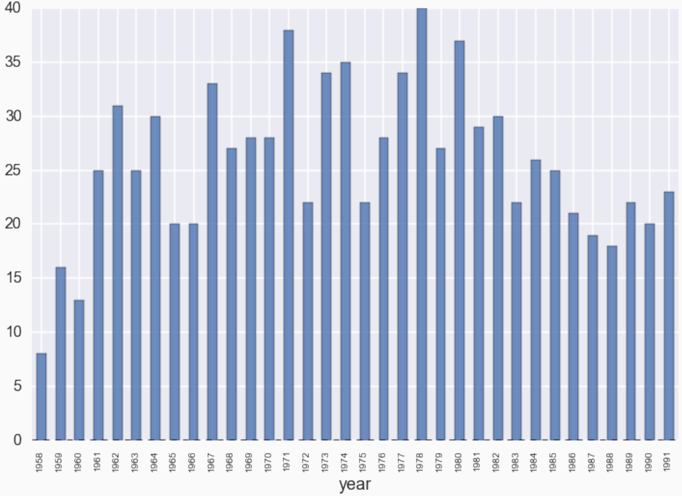
0.0 silence
0.15 C, intro, | A:min | A:min | A:min | A:min |
7.18 | A:min | A:min | A:min | A:min |
14.20 A, verse, | A:min | A:min | A:min | A:min |

[1] John Ashley Burgoyne, Jonathan Wild, and Ichiro Fujinaga, ‘An Expert Ground Truth Set for Audio Chord Recognition and Music Analysis’, in Proceedings of the 12th International Society for Music Information Retrieval Conference, ed. Anssi Klapuri and Colby Leider (Miami, FL, 2011), pp. 633–38

Billboard Annotations

Years are distributed approximately uniformly across the corpus, but the result approximates a normal distribution centered around 1977/1978.

Billboard Annotations



Billboard Annotations

Custom parser converts Billboard annotations to CSV and extracts metric information. All songs were transposed to C major.

```
0,4,C,min
0,3,A-,maj
3,1,B-,maj
0,4,C,min
```

- Starting beat
- Duration
- Transposed chord root
- Chord quality

Corpus Statistics

To simplify and regularize the corpus, a subset of songs was selected such that:

- Meter was $\frac{3}{4}$ or $\frac{4}{4}$
- There were no significant changes of meter

After selection, 876 songs remained.

DATA REPRESENTATION AND PREPROCESSING

From n-grams to bargrams

The amount of information conveyed by harmony is not necessarily constant.

This should be captured during tokenization.

From n-grams to bargrams

Bargram model:

1. Treat each measure as a token containing the aggregation of some number of chord tokens with labeled rhythmic positions
2. Perform collocation detection to identify chord progressions comprising multiple bars

Result:

- Long stretches of the tonic chord are identified as significant collocations, as are several 4-bar and 8-bar repeating circular chord progressions
- Each one of these tokens forms a building block for topics that describe 'style' rather than simply assembling musical primitives

Collocations are formed based on the probability of their constituents cooccurring, and are accepted if:

$$\frac{(\text{count}(w_i, w_j) - \delta) * N}{\text{count}(w_i) * \text{count}(w_j)} > \tau$$

where $\tau = 0.1$ is a threshold value, $\delta = 5$ is a minimum count value for tokens, and N is the vocabulary size.

After collocations are extracted, the corpus is thresholded, removing tokens that do not that appear in 3 or more ($\approx \frac{D}{300}$) documents, or 9 or more times overall ($\approx \frac{D}{100}$).

[1]Tomas Mikolov, Ilya Sutskever, Kai Chen, Gregory S. Corrado, Jeffrey Dean: Distributed Representations of Words and Phrases and their Compositionality. NIPS 2013: 3111-3119

From n-grams to bargrams

Table: Top 10 bargrams

Bargram	Corpus frequency
0,4,c,maj	598
0,4,f,maj	468
0,4,g,maj	393
0,4,g,7	251
0,4,a,min	234
0,4,b-,maj	159
0,4,c,7	133
0,4,c,min	125
0,4,d,min	119
0,4,d,min7	118

From n-grams to bargrams

Table: Top 10 bar-bigrams (___ indicates collocation)

Bar-bigram	Corpus frequency
0,4,c,maj___0,4,c,maj	388
0,4,c,maj	352
0,4,f,maj___0,4,f,maj	166
0,4,f,maj	152
0,4,g,maj___0,4,g,maj	147
0,4,c,maj___0,4,f,maj	142
0,4,g,maj	137
0,4,f,maj___0,4,c,maj	131
0,4,f,maj___0,4,g,maj	107
0,4,g,maj___0,4,c,maj	107

From n-grams to bargrams

Table: Bar-ngrams 100-110 (___ indicates collocation)

Bar-ngram	Corpus frequency
0,4,c,maj___0,4,c,maj___0,4,b-,maj___0,4,b-,maj	19
0,4,f,maj___0,4,g,maj___0,4,c,maj___0,4,c,maj	19
0,4,c,maj___0,4,f,maj___0,4,c,maj___0,4,c,maj	18
0,4,d,min___0,4,d,min	18
0,4,c,min___0,2,a-,maj—2,2,b-,maj (x2)	18
0,2,a-,maj—2,2,b-,maj	18
0,4,a,min___0,4,f,maj	18
0,4,a,maj	18
0,2,g,sus4—2,2,g,maj	18

From n-grams to bargrams

For the final models, two rounds of collocation detection were used, resulting in a final vocabulary size of 615 tokens.

PMI: THE FINAL STEP

PMI

Traditional methods of evaluating topic models, such as calculating perplexity or test data likelihood, suffer from a low correlation with human perception of semantic coherence. Pointwise Mutual Information, or PMI, alleviates this problem.

$$\mathrm{PMI}(w_i, w_j) = \log \frac{P(w_i, w_j)}{P(w_i)P(w_j)}$$

where $P(w_i, w_j)$ is based on a 3-term sliding window, and topics are scored based on the mean PMI score of each of the 8 most probable tokens in a topic with the other 7:

$$\text{Topic PMI} = \frac{1}{{}^{8}P_2} \sum_{i \neq j} \mathrm{PMI}(w_i, w_j), i, j \in 1 \ldots 8$$

PMI Regularized LDA

Newman et al. proposed a method for incorporating PMI information into into LDA via structured priors. This greatly improves the human-rated coherence of topics, especially for very small corpora (like Billboard).

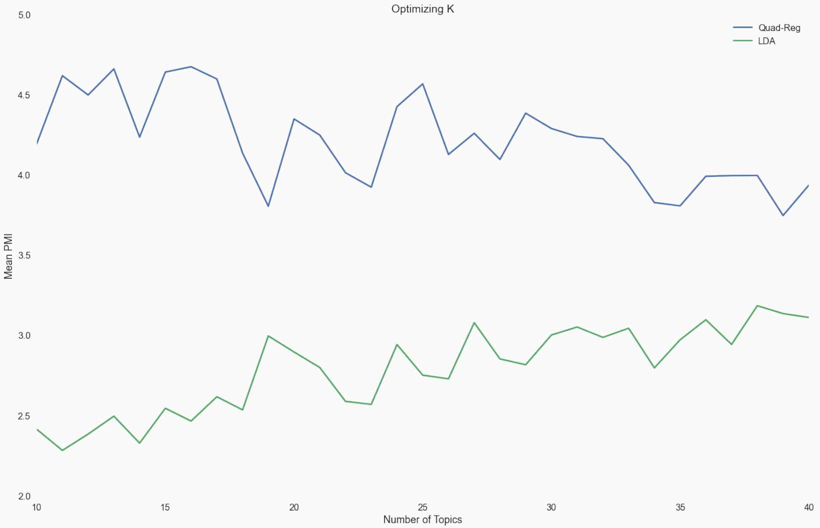
In their QUAD-REG model, prior knowledge is incorporated into the topic word-probability vectors ϕ_t via a quadratic form involving a sparse *WxW* matrix *C*, where entries are PMI scores between words.

$$p(\phi_t|C) \propto (\phi_t^\intercal C \phi_t)^\nu$$

[1] Newman, Bonilla, Buntine (2011). Improving Topic Coherence with Regularized Topic Models. In NIPS 2011

RESULTS

Finding the optimum number of topics



Topics

Topic 1

- 0,4,c,maj___0,4,c,maj(294)
- 0,4,c,maj___0,4,c,maj7/7___0,2,g,sus4(b7,9)—2,2,c,7___0,4,f,maj(15)
- 0,4,f,sus4(b7,9)___0,4,f,sus4(b7,9)(12)
- 0,4,f,maj/9___0,4,f,maj/9(20)
- 0,4,c,maj___0,2,d,min7—2,2,g,7(10)PM
- 0,2,f,maj7—2,2,f,maj6(13)
- 0,4,a,min9(10)
- 0,4,b-,maj(9)(13)

Mean PMI: 2.619989

Topics

Topic 2

- 0,4,c,5___0,4,c,5(45)
- 0,4,c,5(44)
- 0,4,c,5___0,4,c,5___0,4,c,5___0,4,c,5(51)
- 0,4,f,5___0,4,f,5(18)
- 0,4,c,5(b7)(11)
- 0,4,e-,5(20)
- 0,4,g,5(10)
- 0,2,a-,maj—2,2,b-,maj(37)

Mean PMI: 6.472619

Topics

Topic 3

- 0,4,g,maj___0,4,g,maj(138)
- 0,4,a,min___0,4,a,min(36)
- 0,4,g,maj___0,4,g,maj___0,4,g,maj___0,4,g,maj(40)
- 0,4,f,maj___0,4,g,maj(90)
- 0,4,d,maj___0,4,d,maj(27)
- 0,4,f,maj___0,4,c,maj/3(16)
- 0,4,g,maj___0,4,g,maj___0,4,f,maj___0,4,f,maj(36)
- 0,4,g,maj___0,4,f,maj(33)

Mean PMI: 4.165779

Topics

Topic 4

- 0,4,f,maj___0,4,f,maj___0,4,c,maj___0,4,c,maj(183)
- 0,4,c,maj___0,4,c,maj___0,4,a,min___0,4,a,min(104)
- 0,4,f,maj___0,4,f,maj___0,4,g,maj___0,4,g,maj(97)
- 0,4,f,maj___0,4,f,maj___0,4,f,maj___0,4,f,maj(29)
- 0,4,c,maj___0,4,c,maj___0,4,f,maj___0,4,f,maj(150)
- 0,4,a,min___0,4,a,min___0,4,f,maj___0,4,f,maj(34)
- 0,4,g,7___0,4,g,7___0,4,c,maj___0,4,c,maj(60)
- 0,4,f,maj___0,4,f,maj___0,4,g,7___0,4,g,7(13)

Mean PMI: 3.955108

Topics

Topic 5

- 0,3,c,maj___0,3,c,maj(27)
- 0,3,f,maj(19)
- 0,3,c,maj(29)
- 0,3,f,maj___0,3,f,maj(14)
- 0,3,g,maj___0,3,c,maj(12)
- 0,3,g,maj___0,3,g,maj(11)
- 0,3,a,min(14)
- 0,3,f,maj___0,3,g,maj(16)

Mean PMI: 11.213694

Topics

Topic 6

- 0,4,a-,maj(57)
- 0,4,b-,maj(79)
- 0,4,e-,maj(54)
- 0,4,g,min(26)
- 0,4,f,min7(67)
- 0,4,e-,maj___0,4,a-,maj(33)
- 0,4,a-,maj___0,4,b-,maj(48)
- 0,4,c,min7(49)

Mean PMI: 4.931101

Topics

Topic 7

- 0,4,g,11(26)
- 0,4,g,11___0,4,g,11(29)
- 0,4,c,maj___0,4,g,min7___0,4,c,maj___0,4,g,min7(9)
- 0,4,g,min7(43)
- 0,4,c,maj9___0,4,c,maj9(15)
- 0,2,f,maj—2,2,c,maj___0,4,g,maj(16)
- 0,2,c,min—2,2,b-,maj___0,4,a-,maj(16)
- 0,4,b-,maj/b7(15)

Mean PMI: 3.465998

Topic 8

- 0,4,c,7___0,4,c,7___0,4,c,7___0,4,c,7(234)
- 0,4,f,7___0,4,f,7___0,4,c,7___0,4,c,7(67)
- 0,4,g,7___0,4,f,7___0,4,c,7___0,4,c,7(29)
- 0,4,c,7___0,4,c,7___0,4,f,7___0,4,f,7(29)
- 0,4,c,7___0,4,c,7___0,4,c,7(17)
- 0,4,f,7___0,4,f,7(22)
- 0,4,g,7___0,4,f,7(15)
- 0,4,d,7___0,4,d,7___0,4,g,7___0,4,g,7(10)

Mean PMI: 7.060424

Topics

Topic 9

- 0,4,f,maj___0,4,c,maj(77)
- 0,4,c,maj/5___0,4,g,7(12)
- 0,4,f,maj___0,4,c,maj___0,4,c,maj(13)
- 0,4,g,7___0,4,c,maj___0,4,c,7___0,4,f,maj(15)
- 0,4,d,7___0,4,g,7___0,4,g,7___0,4,c,maj(15)
- 0,4,c,maj___0,4,g,min7(13)
- 0,2,g,maj—2,2,f,maj(15)
- 0,4,f,maj___0,4,c,maj___0,4,g,7___0,4,c,maj(9)

Mean PMI: 3.371266

Topics

Topic 10

- 0,4,c,maj___0,4,c,maj___0,4,c,maj___0,4,c,maj(562)
- 0,4,g,maj___0,4,c,maj___0,4,c,maj___0,4,c,maj(31)
- 0,4,c,maj___0,4,c,maj___0,4,c,maj___0,4,f,maj(23)
- 0,4,c,maj___0,4,c,maj___0,4,c,maj(65)
- 0,4,f,maj___0,4,c,maj___0,4,c,maj___0,4,c,maj(47)
- 0,4,c,maj___0,4,f,maj___0,4,f,maj___0,4,g,maj(10)
- 0,4,c,maj___0,4,c,maj___0,4,c,maj___0,4,g,maj(19)
- 0,4,c,7___0,4,f,maj___0,4,f,maj___0,4,c,maj(18)

Mean PMI: 5.373208

Topics

Topic 11

- 0,4,c,maj(521)
- 0,4,f,maj(241)
- 0,4,a,min7(65)
- 0,4,g,7(95)
- 0,4,d,min7(65)
- 0,4,g,maj(118)
- 0,4,d,min7___0,4,g,7(47)
- 0,4,f,maj7(50)

Mean PMI: 2.465027

Topics

Topic 12

- 0,4,g,maj___0,4,c,maj(82)
- 0,4,a,min___0,4,g,maj(45)
- 0,4,g,maj___0,4,c,maj___0,4,a,min___0,4,f,maj(12)
- 0,4,e,min___0,4,f,maj(14)
- 0,4,g,maj___0,4,a,min(30)
- 0,4,f,maj___0,4,d,min(24)
- 0,4,c,maj___0,4,a,min(44)
- 0,4,a,min___0,4,d,maj(26)

Mean PMI: 4.124833

Topics

Topic 13

- 0,4,c,min___0,4,c,min___0,4,c,min___0,4,c,min(319)
- 0,4,c,min___0,4,c,min(91)
- 0,4,c,min(122)
- 0,4,c,min___0,4,c,min___0,4,f,min___0,4,f,min(22)
- 0,4,c,min___0,4,c,min___0,4,c,min(13)
- 0,4,c,min___0,4,c,min___0,4,b-,maj___0,4,b-,maj(66)
- 0,4,b-,7___0,4,b-,7(14)
- 0,4,e-,maj___0,4,e-,maj___0,4,e-,maj___0,4,e-,maj(11)

Mean PMI: 4.321794

Topics

Topic 14

- 0,4,c,maj___0,4,c,maj___0,4,d,min___0,4,d,min(19)
- 0,2,d,min—2,2,g,maj___0,4,c,maj(14)
- 0,4,g,maj___0,4,g,maj___0,4,c,maj___0,4,c,maj(64)
- 0,4,e,min___0,4,e,min___0,4,f,maj___0,4,f,maj(18)
- 0,2,c,maj—2,1,f,maj/5—3,1,c,maj(18)
- 0,4,f,maj___0,4,f,maj___0,4,c,maj___0,4,g,maj(17)
- 0,4,f,maj___0,4,f,maj___0,4,a,min___0,4,a,min(14)
- 0,4,d,maj(12)

Mean PMI: 2.262797

Topics

Topic 15

- 0,4,c,maj___0,4,c,maj___0,4,c,maj___0,4,a,min(45)
- 0,4,g,7___0,4,c,maj___0,4,c,maj___0,4,c,maj(41)
- 0,4,a,min___0,4,a,min___0,4,a,min(10)
- 0,4,c,min7___0,4,c,min7___0,4,c,min7___0,4,c,min7(215)
- 0,4,g,maj___0,4,c,maj___0,4,c,maj___0,4,a,min(11)
- 0,4,f,maj___0,4,c,maj___0,4,c,7___0,4,f,maj(10)
- 0,2,f,maj—2,2,g,maj___0,4,c,maj___0,2,f,maj—2,2,g,maj___0,4,c,maj(34)
- 0,4,c,min11___0,4,c,min11(15)

Mean PMI: 2.091902

Topics

Topic 16

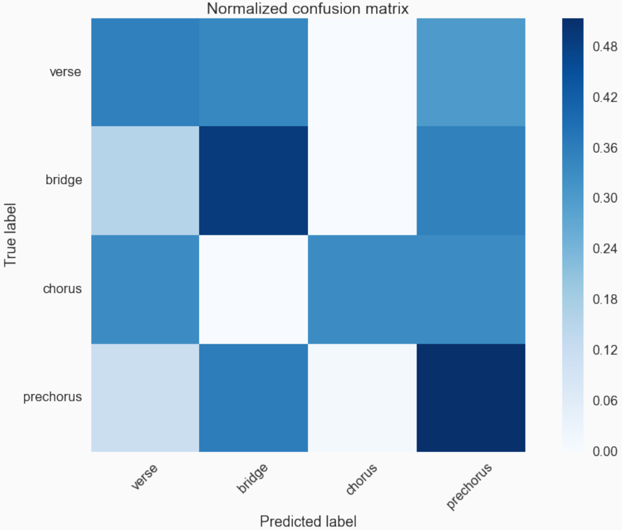
- 0,4,f,maj6(33)
- 0,4,a,min(114)
- 0,4,g,maj6(21)
- 0,4,a,min/b7(11)
- 0,4,c,min___0,4,c,min/b7___0,4,c,min/13___0,2,a-,maj7—2,2,g,7(108)
- 0,4,g,sus4(b7)___0,4,g,7(14)
- 0,4,a,7___0,4,d,min7(16)
- 0,4,c,maj7___0,4,c,maj7(30)

Mean PMI: 3.715426

Structural Schemata

- Random forest classifier with 100 trees
- Four labels: Verse, Chorus, Prechorus, Bridge
- Mean success rate over 10-fold cross-validation was 50.67% for the four-class case
- Bridge and Prechorus correctly identified in most cases, hinting at distinct topical content
- Classifier has most difficulty predicting Chorus label: least distinctive song section to maximize predictability for listeners?

Structural Schemata

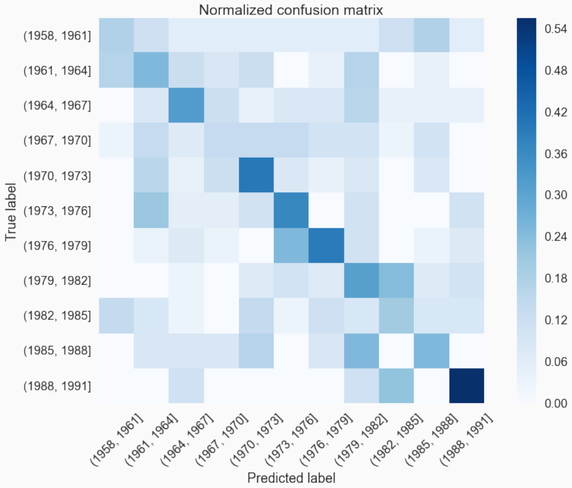


Year prediction

Year prediction experiment:

- Random forest classifier with 100 trees
- Corpus years sorted into 5 and 11 bins
- Mean success rate over 10-fold cross-validation was 27.59% for the eleven-class case and 40.51% for the five-class case

Year prediction



Questions?