



PREDICTING POETIC MOVEMENTS

BY JOSHUA SZYMANOWSKI

WHY?



A LOOK AT THE DATA



SOURCE

poetryfoundation.org

TOTAL POEMS

5,176 → 4,376

TARGET

genre

YEARS

1780 - present

WORDS

1,034,736 – 110,723 unique



TARGET

Pre-1900
(23.8%)



Modern
(29.2%)



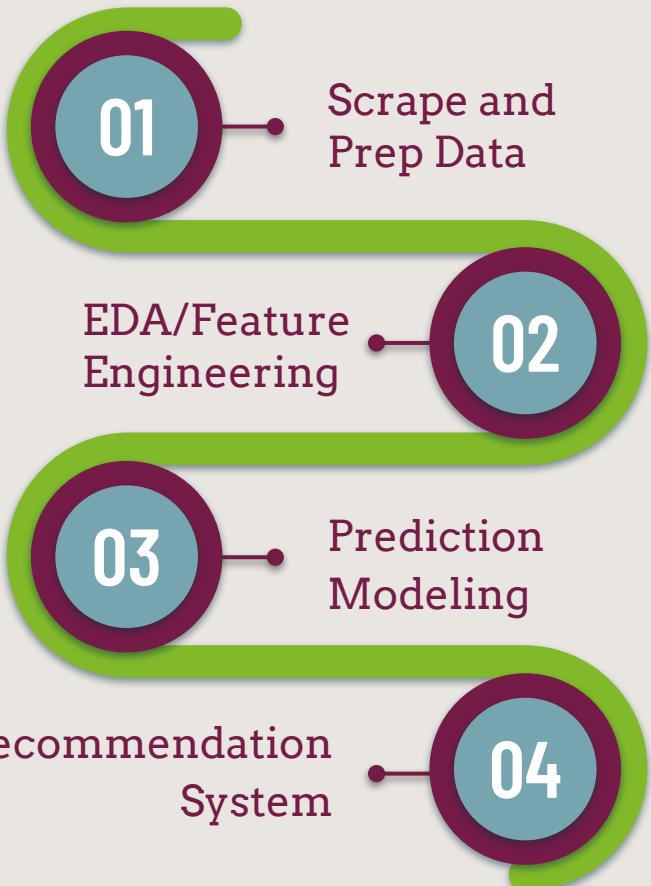
Metropolitan
(25.0%)



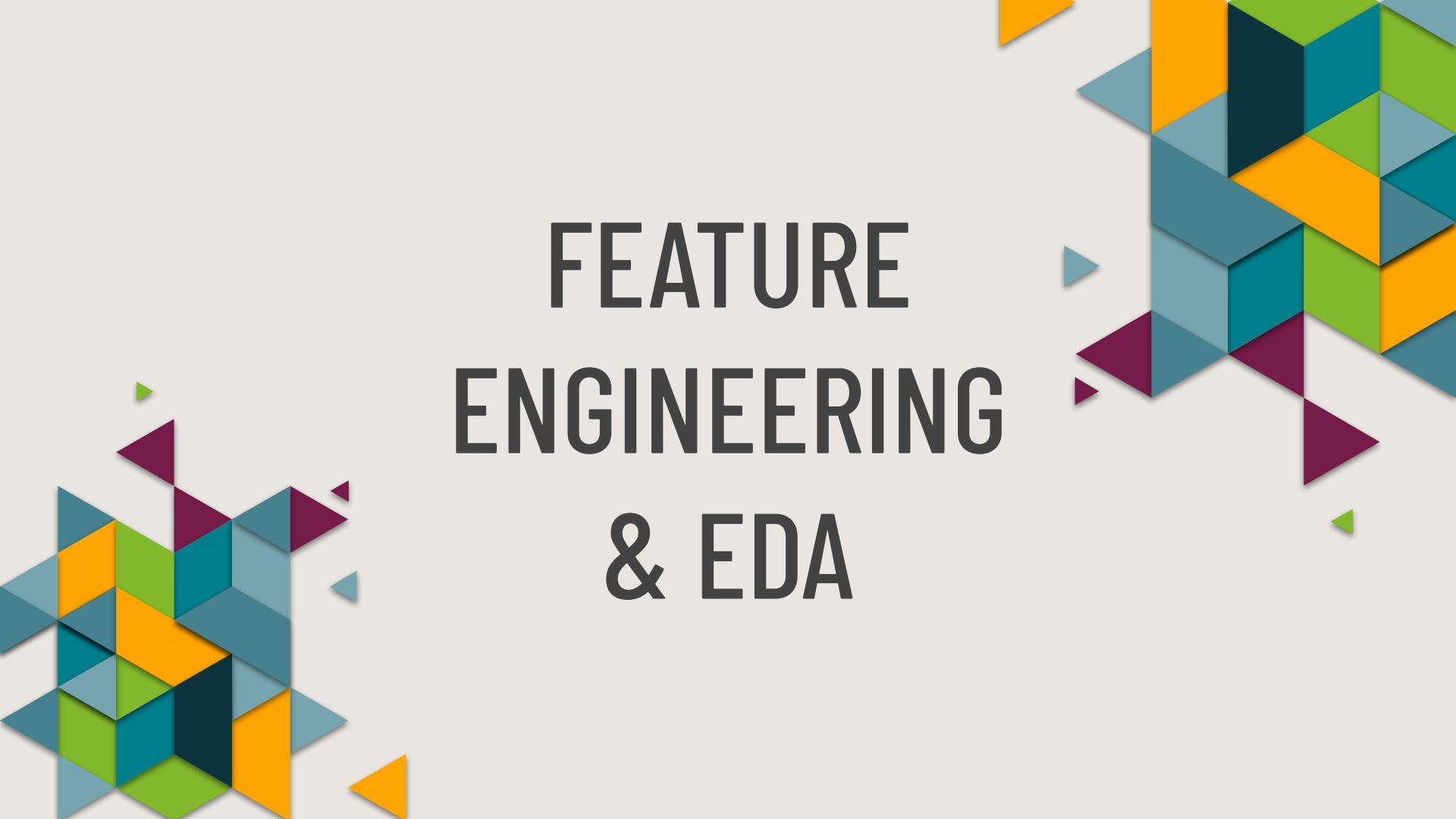
Avant-Garde
(22.0%)



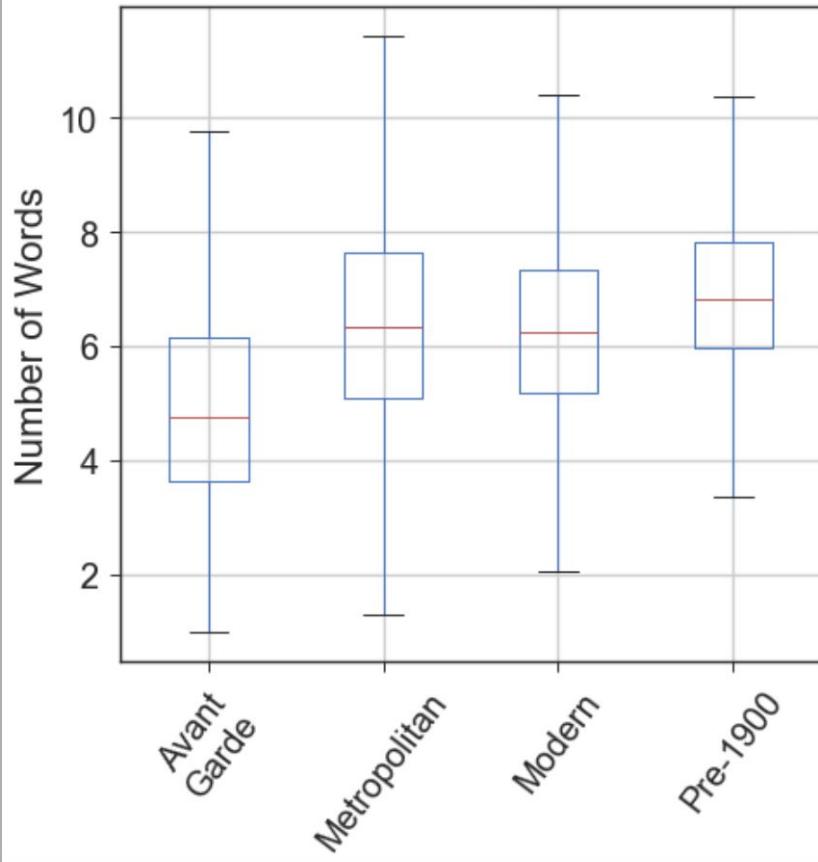
ROADMAP

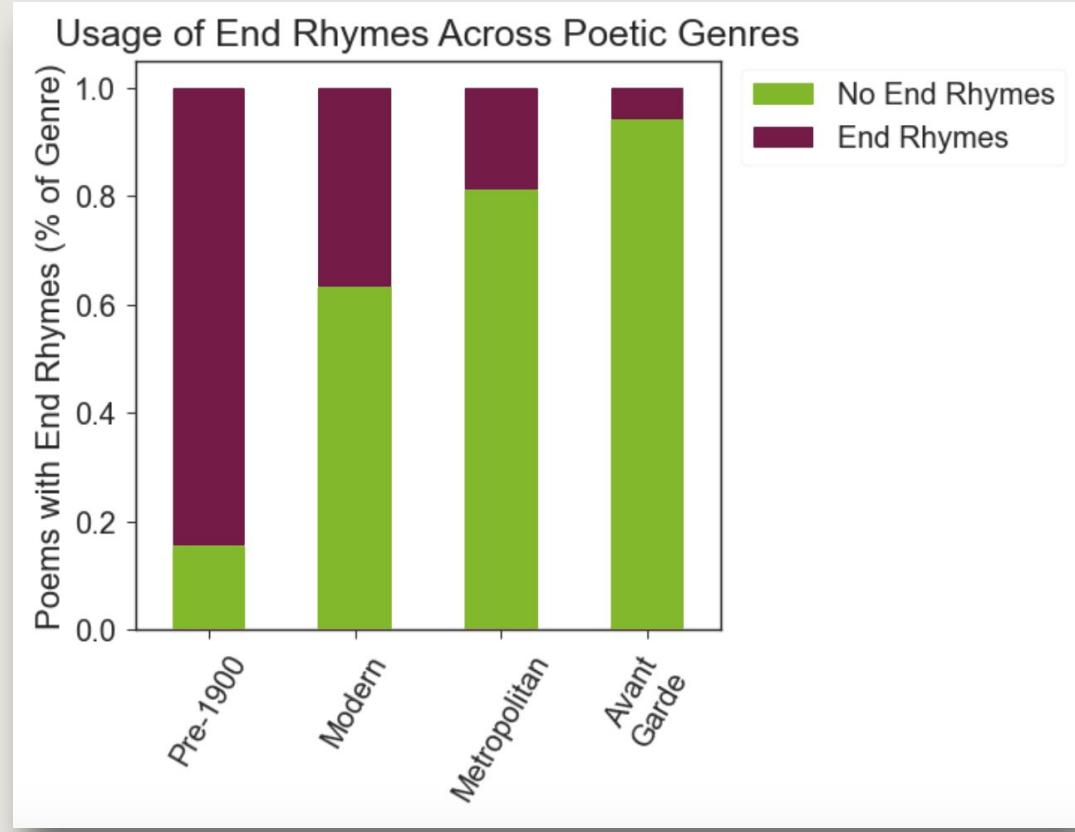


FEATURE ENGINEERING & EDA

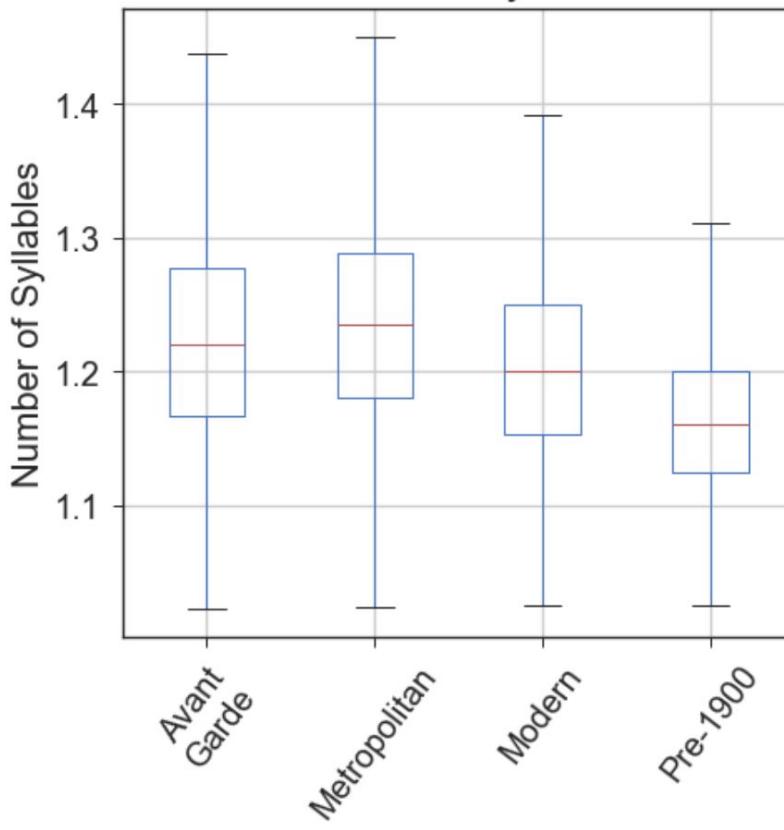


Average Line Length per Poem, Distribution by Genre

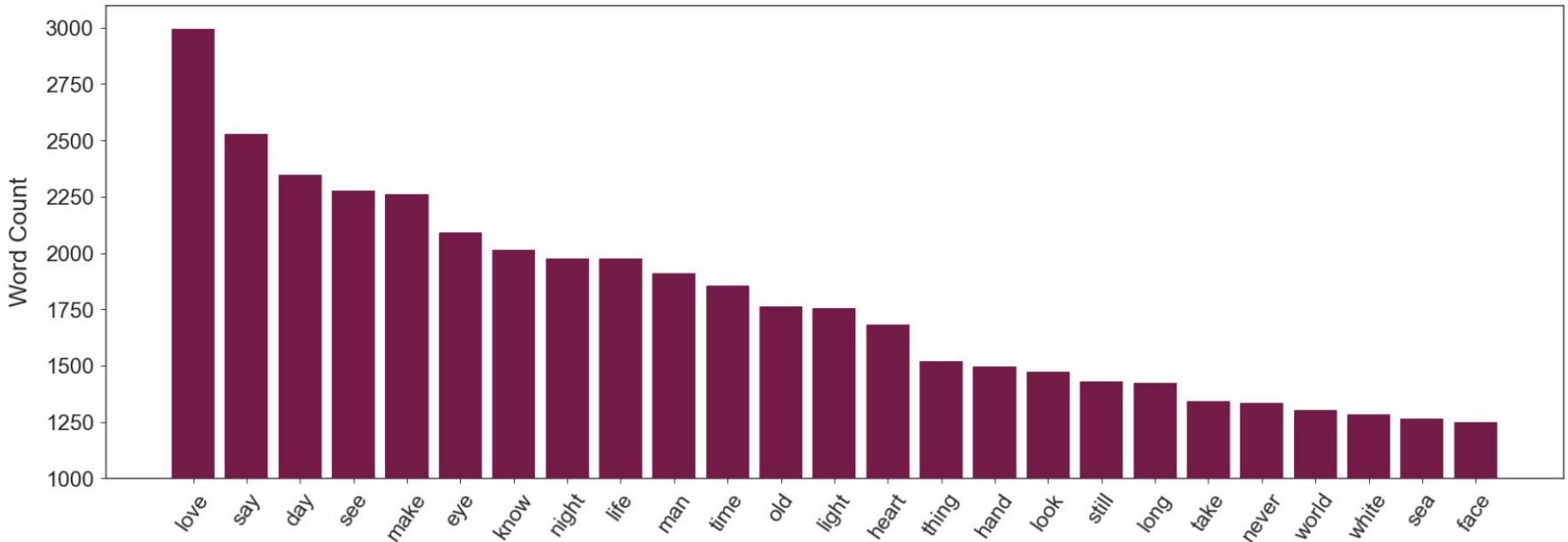




Average Number of Syllables Per Word, Distribution by Genre



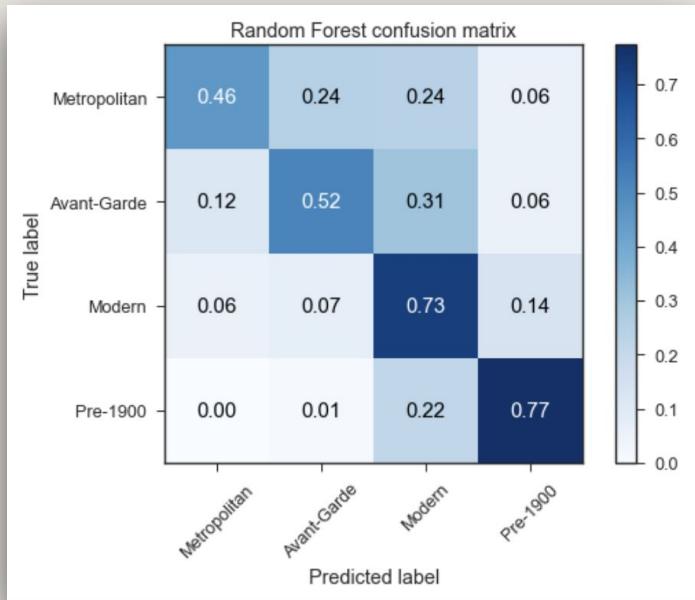
Top 25 Words



MODELS

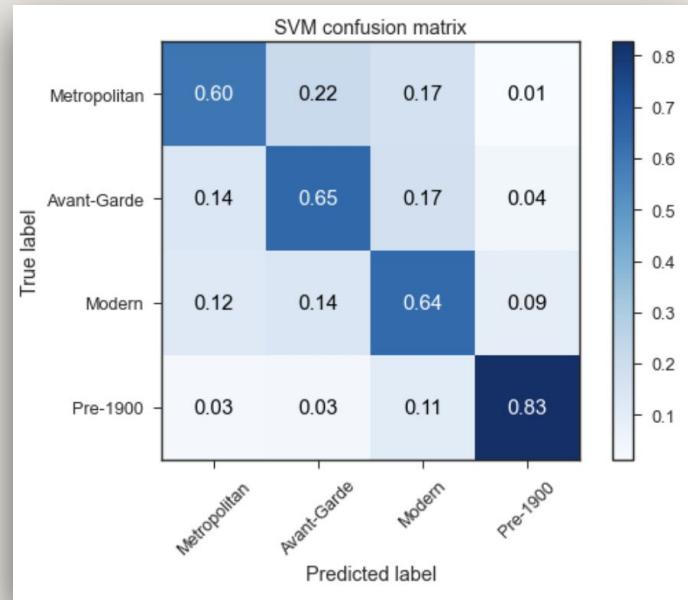


RANDOM FOREST MODEL



ACCURACY 62.6%
F1 SCORE 62.1%

SVM MODEL



ACCURACY 67.8%
F1 SCORE 67.8%

RANDOM FOREST

SVM

TOP FEATURES

number of end rhymes, length of line,
syllables per word, number of lines,
polarity score, sentiment score

crazy, among, nobody, form, look
length of line, dream, small, cry, star

RECOMMENDER APP DEMO

I am designed to recommend poetry based on certain parameters.

You can fiddle with my settings on the left of your screen.

Give me a word.

love

Would you like to use my filter?

no
 yes

Great news! I found 100 poems.

38.1% match

RETURNING TO ROOTS OF FIRST FEELING by Robert Duncan

GENRE: Black Mountain

34.8% match

FRAGMENT 5· WHOM SHOULDN'T CHOOSE FOR MY JUDGE? by Samuel Taylor Coleridge

poetryfoundation.org/poetrymagazine/poems/.../returning-to-roots-of-first-feeling



QUESTIONS?

Joshua Szymanowski

joshua.szymanowski@gmail.com

https://evening-cove-11137.herokuapp.com/

p-szymo

LinkedIn



TAKEAWAYS

- Pre-1900 poetry is easily recognized
 - End rhymes and length of line
 - Avant-garde is its polar opposite
 - Form is important for most prediction models
 - SVM relies mostly on vocabulary
 - Poetry is rarely neutral (and generally positive)
- 