



# PREDICTING POETIC MOVEMENTS

BY JOSHUA SZYMANOWSKI

# WHY?





## A LOOK AT THE DATA

SOURCE

[poetryfoundation.org](http://poetryfoundation.org)

TOTAL POEMS

5,176 → 4,376

TARGET

genre

YEARS

1780 - present

WORDS

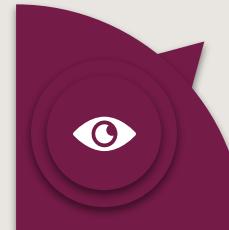
1,034,736 – 110,723 unique

# TARGET

Pre-1900  
(23.8%)



Modern  
(29.2%)



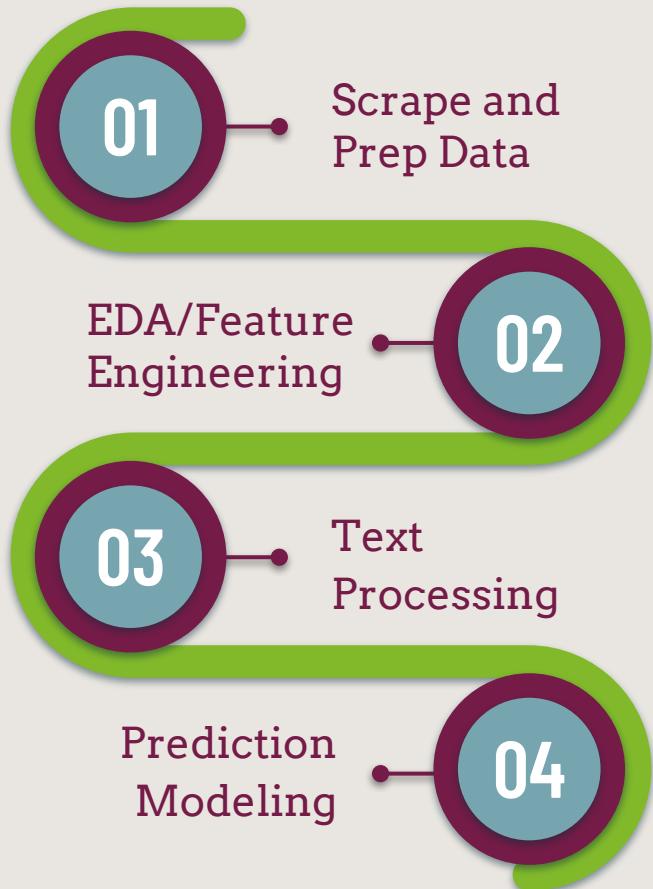
Metropolitan  
(25.0%)



Avant-Garde  
(22.0%)



# ROADMAP



# MODEL PREVIEW



ACCURACY

46.7%

F1 SCORE

42.6%



ACCURACY

47.2%

F1 SCORE

47.0%



ACCURACY

62.6%

F1 SCORE

62.1%



ACCURACY

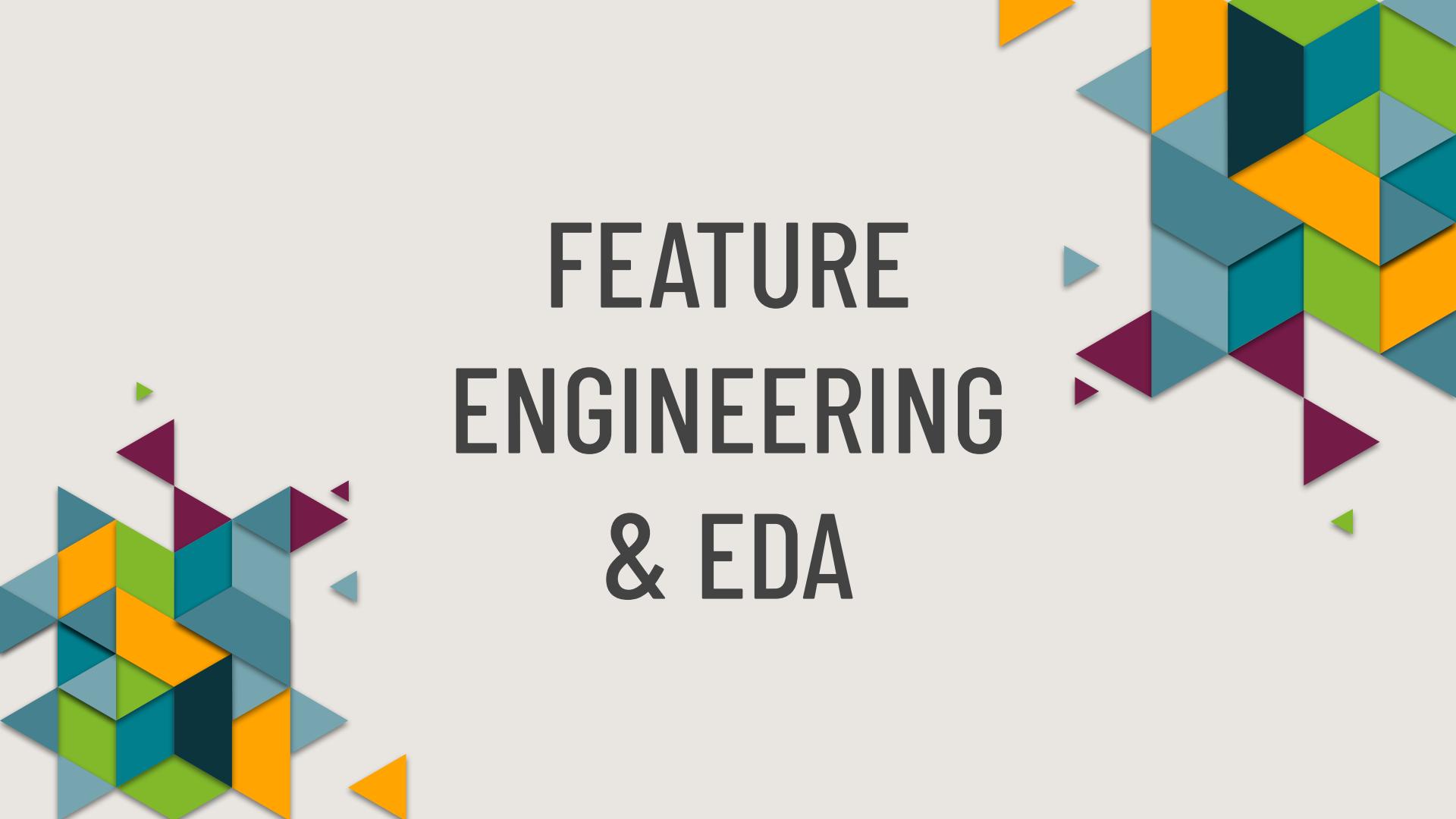
67.8%

F1 SCORE

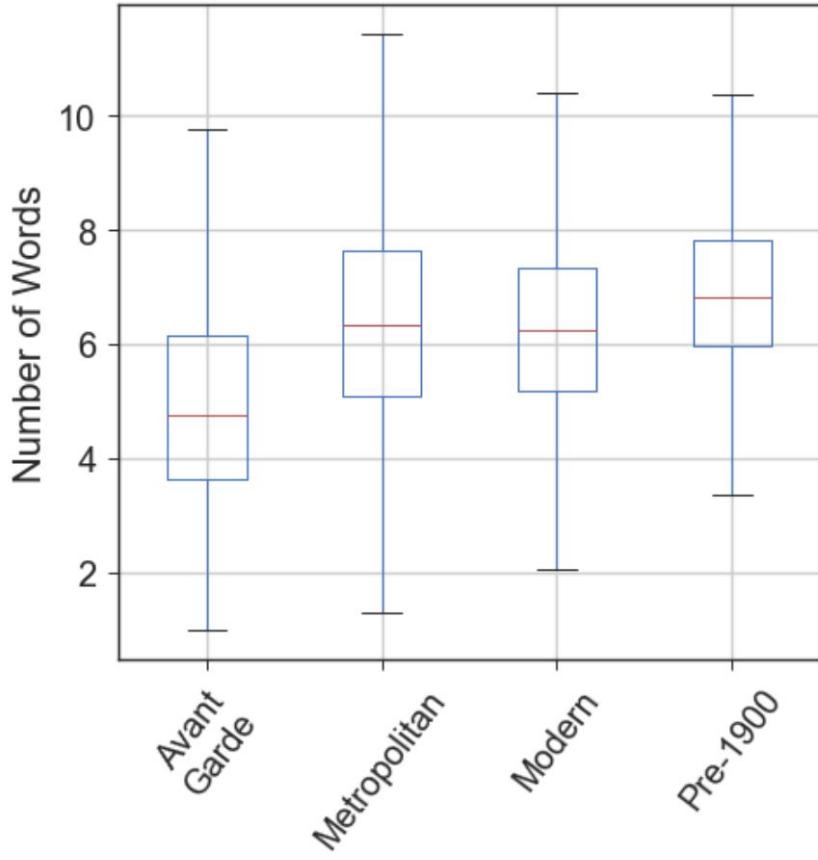
67.8%



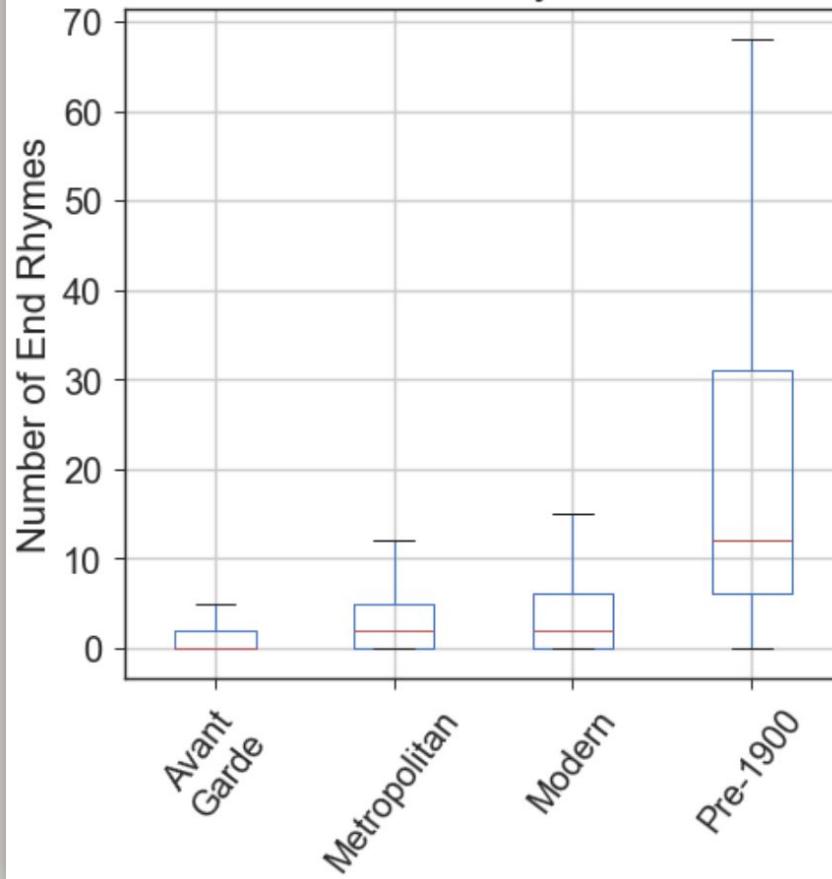
# FEATURE ENGINEERING & EDA



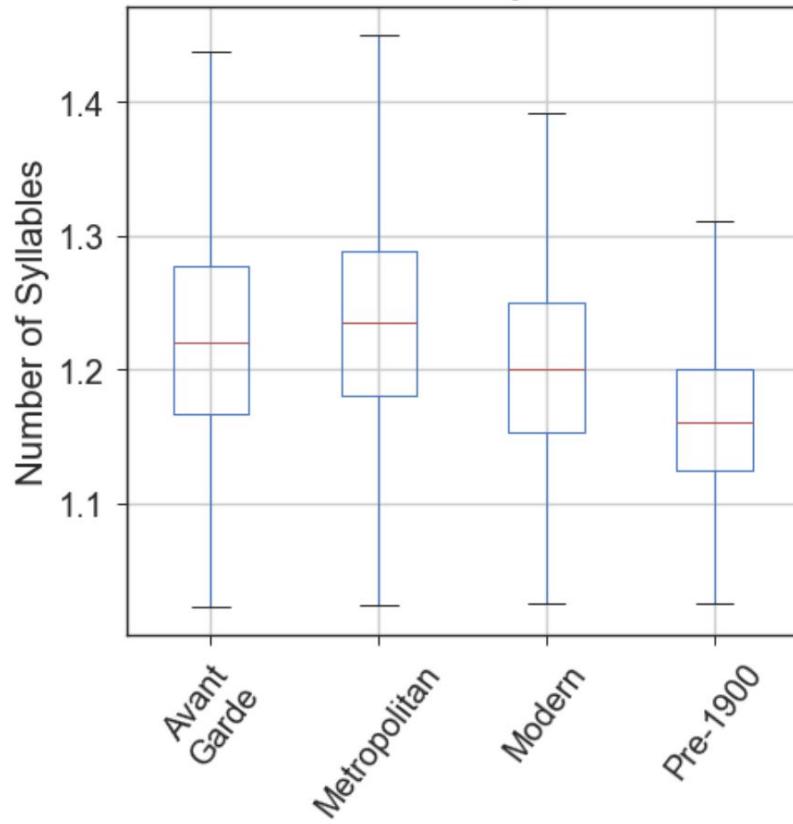
## Average Line Length per Poem, Distribution by Genre



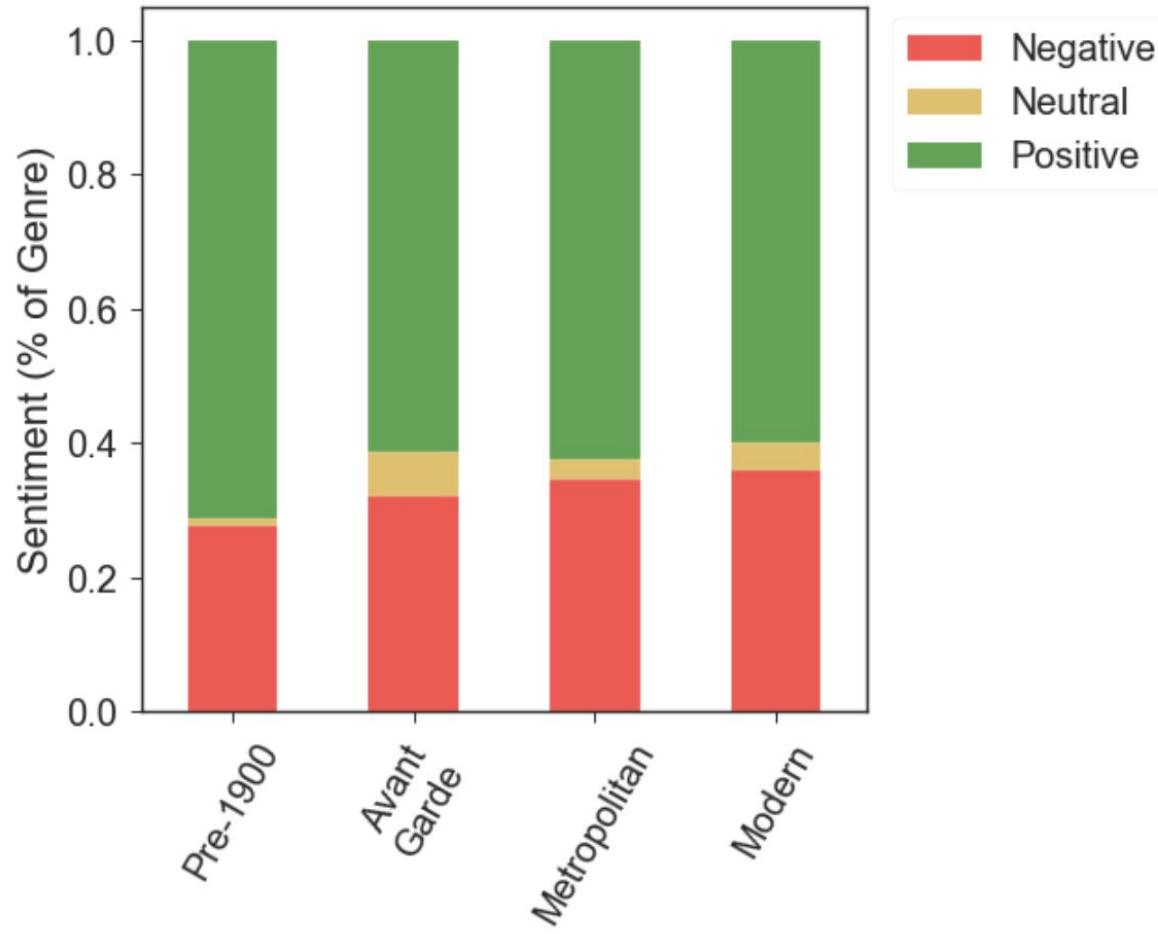
## Number of End Rhymes, Distribution by Genre



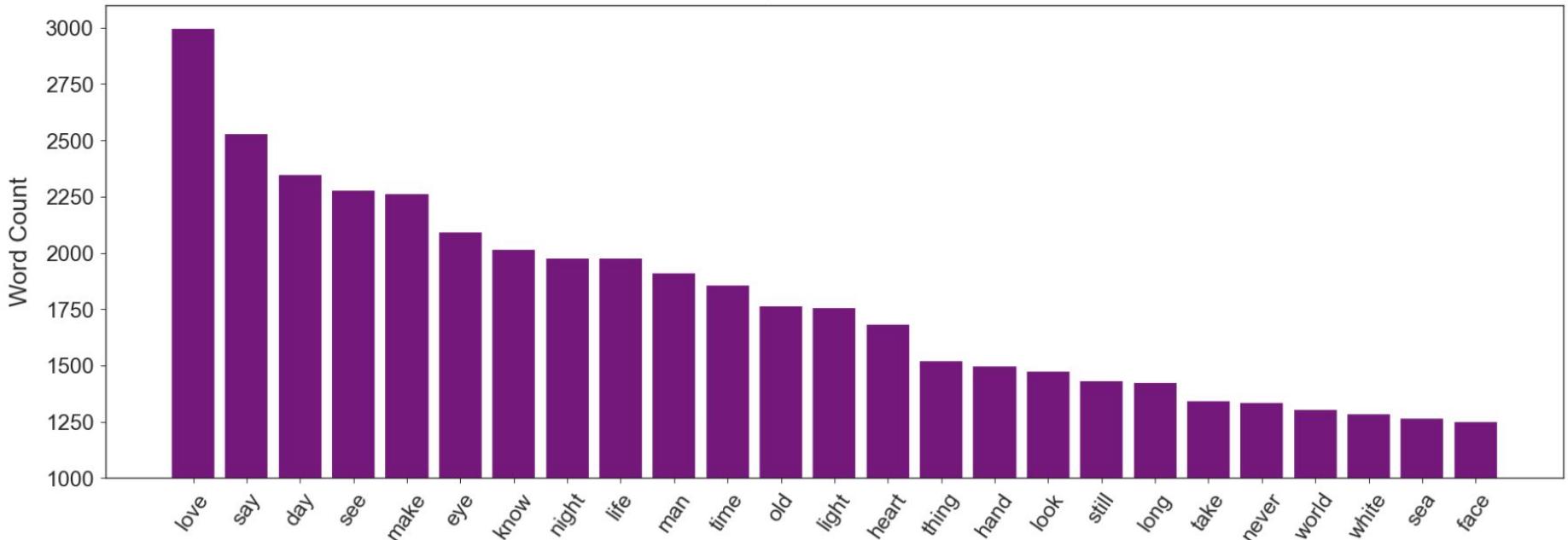
## Average Number of Syllables Per Word, Distribution by Genre



## The Polarity of Poetic Genres



## Top 25 Words



## Metropolitan

never new still even hear open boy in the ball stand first city name leaf earth  
american green sleep good empty star left  
father end sleep dream sun walk keep slide  
fall place right much color rain give feel way change  
death wait poem dark light blue call god thought  
soul leave summer begin river space shadow  
room child wind dead hand long move last  
beautiful everything dead hand long move last  
saw tree run house holy book bed  
let though something voice live done sweet fire talk  
sound rise remember around man take away hair  
head another remember around man take away hair  
work seem past mother past become sea  
blood tell always work seem past mother past become sea  
little write turn lie red wall body  
tell young air lose two girls morning wall  
hold sky turn sit up water morning wall  
want snow black sing full song well friend  
woman always write turn sit up water morning wall  
achos

## Modern

people light heart long grow mountain hold silver finger half verse  
black mind bird blood flower alone god yellow gold last  
hour shadow young color stone high set rain another  
away show left lie grass thought star stone  
moon without keep well sing green city  
music shadow till born touch cold street  
must first round foot soul men still dan song  
work find poem year red life world wing  
find spring sprout two sweet place end smile hill part  
fall water summer thing name strong land fire wave  
take move tree air back nothing turn and walk  
night far dust way even open sky word memory  
bright every laugh house sleep wall always night  
cry window whose cloud wait arm open dark  
full stand earth silence golden body friend  
day look over sun never think fly in play  
death sound door head earth change call hear heaven beauty  
old little many sea child great dream leaf

Avant-Garde

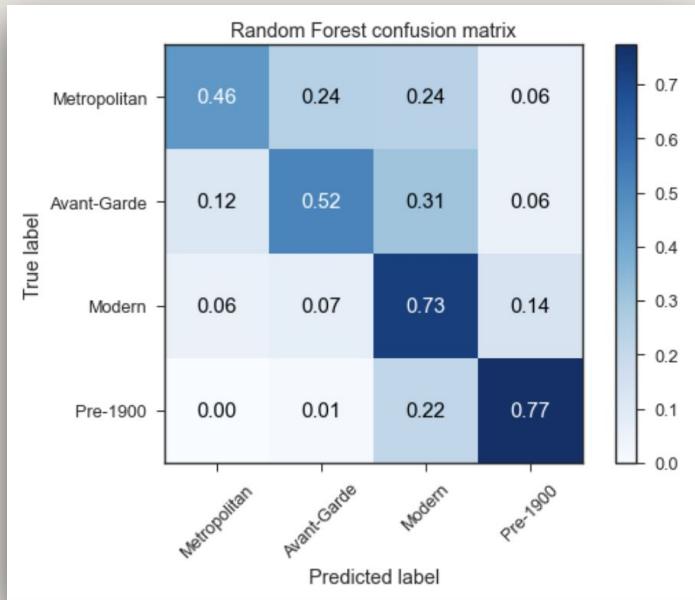
Pre-1900

golden  
man part  
sweet leaf-wood clear  
old poor cry voice star inside  
thought ever air must forth  
rise wing stood fill  
mountain smile Whose great gold  
day cold earth spring foot  
eye red song white  
far bed tell left mind sight  
thing night till though  
mineyly hand clear  
sea leave lie  
never pale good  
find set fear move  
wind land door  
dear field without  
World change art  
look bright nature blue lay  
take away keep  
tart green king  
bright shadow  
away hear  
green soft  
thus death  
live word  
watch ion  
time friend  
soul half  
time pa  
breast since  
grow young  
name saw  
black cloud light  
say deep snow  
still pas kiss  
light age  
child first

# MODELS

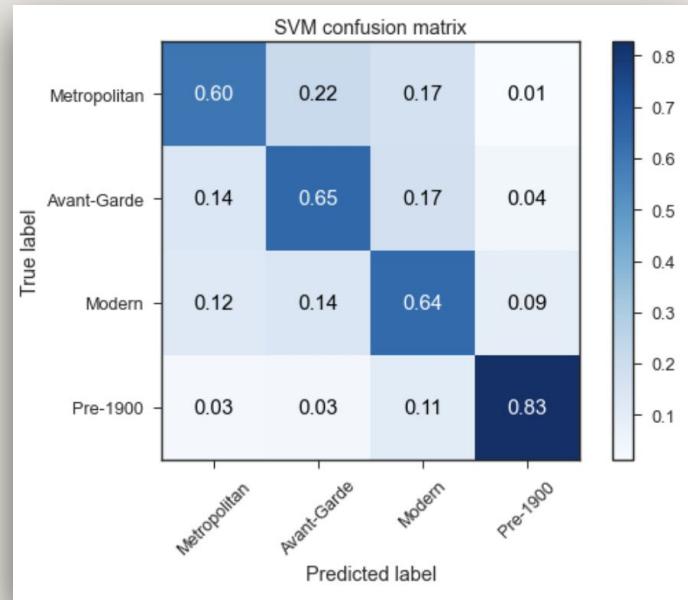


# RANDOM FOREST MODEL



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

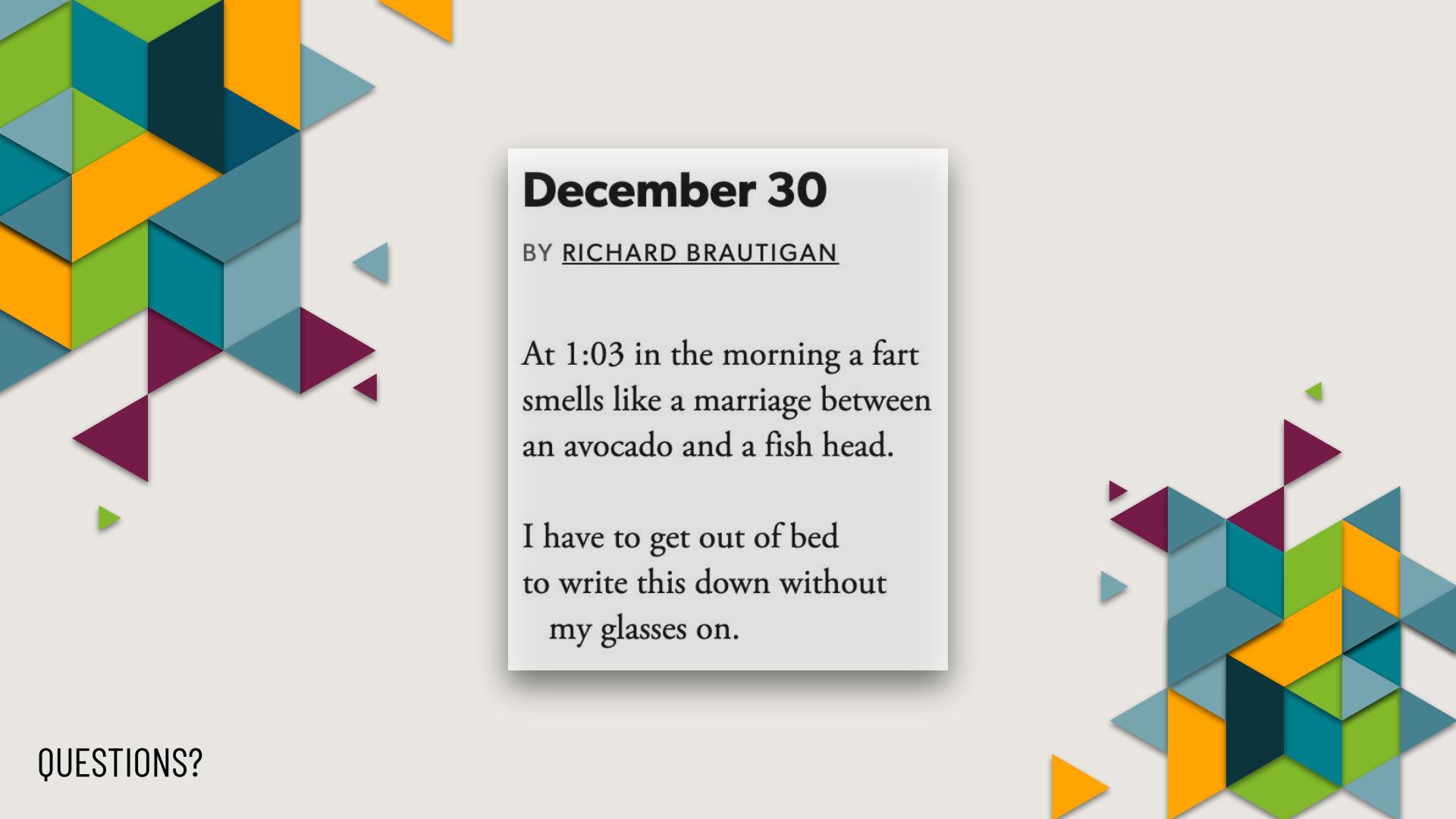
*crazy, among, nobody, form, look*  
*length of line, dream, small, cry, star*



# 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)
- 



# December 30

BY RICHARD BRAUTIGAN

At 1:03 in the morning a fart  
smells like a marriage between  
an avocado and a fish head.

I have to get out of bed  
to write this down without  
my glasses on.

QUESTIONS?

# Joshua Szymanowski

joshua.szymanowski@gmail.com

## CONTACT INFO

github

LinkedIn