To be, or not to be… Shakespeare?

Eva Mizer • 05.04.2023

The Problem

Client

- Literary historians
- Bardolators

Problem

- Mystery authors of Elizabethan/Jacobean era plays
- Who wrote what in collaborations

Solution

- Writing analysis

The Solution

Data

- Romantic comedies of three playwrights

Feature Extraction

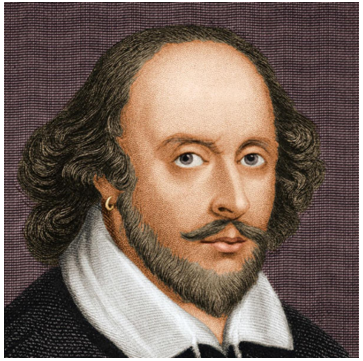
- Word choice
- Word combinations
- Punctuation
- Line length

Model Creation

- Process text and make prediction

The Process

Playwrights



William Shakespeare

- Much Ado About Nothing, 1598/99

Thomas Middleton

- A Trick to Catch the Old One, 1608

John Fletcher

- The Faithful Shepherdess, 1609

Features

Writing Style

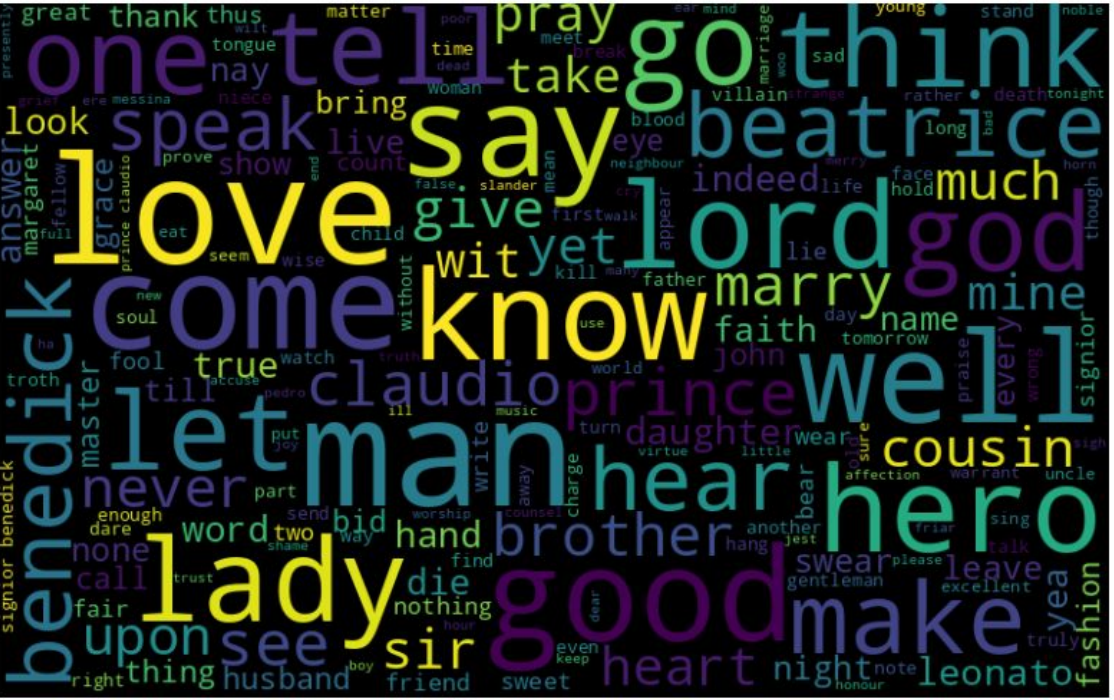
- Punctuation
- Line Length

Grams

- Unigrams - vocabulary
- Bigrams - word combos

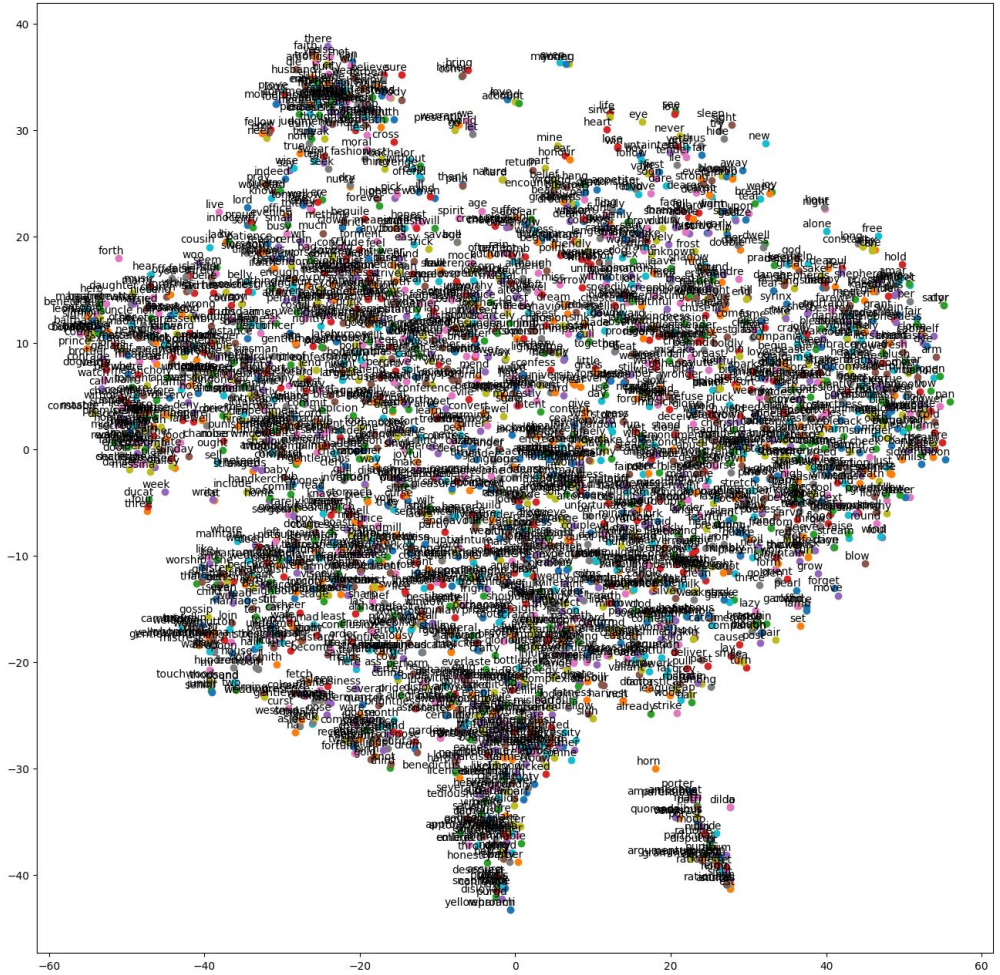
Word2Vec Embeddings

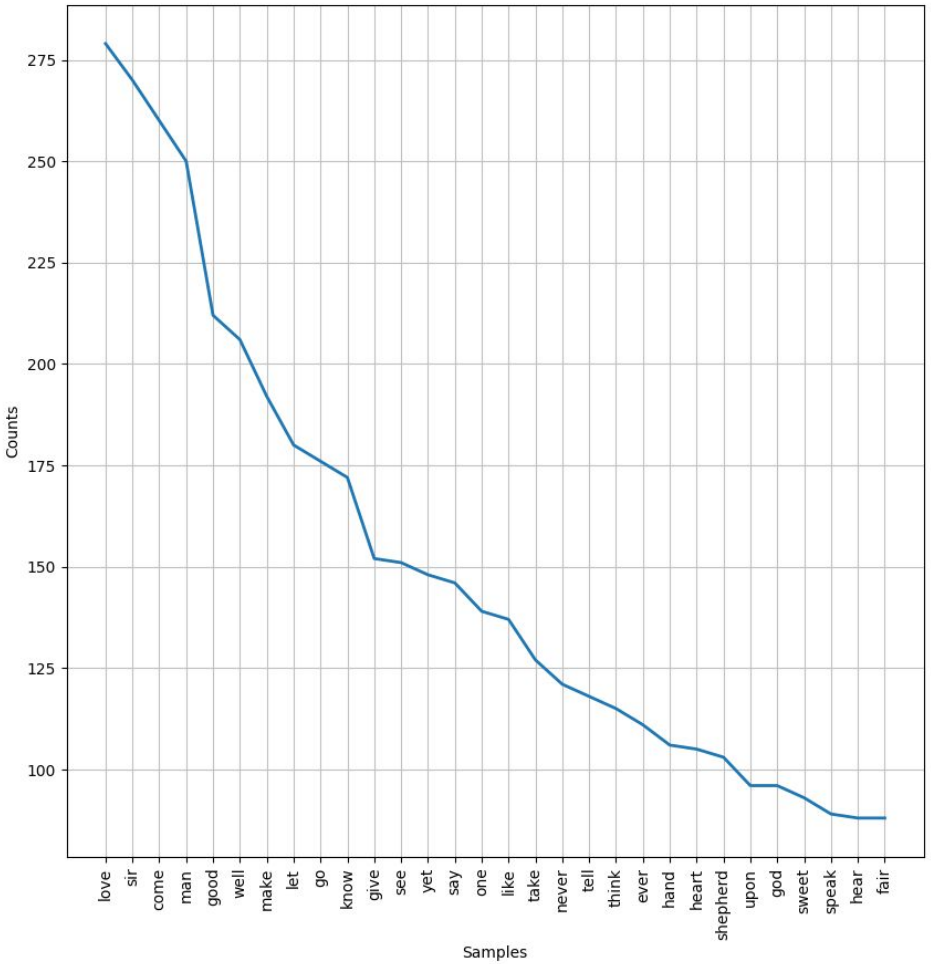
- Word proximity


one
tell
pray
go
think
say
speak
beatrice
love
lord
god
come
know
marry
mine
well
let
man
claudio
prince
hear
hero
cousin
never
lady
good
make
brother
hand
word
see
upon
sir
heart
leonato
fashion
benedick
signior benedick
great
thank
thus
matter
tongue
time
nay
bring
look
live
show
count
eye
blood
sad
villain
long
much
indeed
give
yet
wit
child
father
faith
name
tomorrow
every
true
fool
till
daughter
wear
cousin
affection
uncle
enough
dare
bid
hand
word
two
none
call
fair
die
nothing
swear
leave
gentleman
excellent
yea
truly
thing
husband
friend
sweet
night
note
honour
right
boy
even
keep
hour
dear
hang
jest
another
sing
talk
please
find
trust
way
shame
send
part
worship
counsel
charge
away
write
turn
virtue
little
sure
bear
master
fellow
answer
margaret
grace
prince claudio
full
eat
seem
wise
new
soul
ha
troth
watch
pedro
accuse
music
ill
put
joy
without
use
world
truth
kill
many
appear
firstwalk
false
slander
mean
neighbour
marry
face
hold
life
lie
day
tomorrow
praise
every
wrong
signior
sigh
though
horn
bad
long
rather
death
tonight
stand
noble
young
marriage
woo
meet
break
dead
poor
wilt
messina
niece
prove
end
count
woman
take
time
tongue
nay
great
thank
thus
matter
present

Vocabulary

Word Vectors





Word Frequency

Models

Random Forest Classifier

- Accuracy: 62%

Multinomial Naive Bayes

- Accuracy: 75%

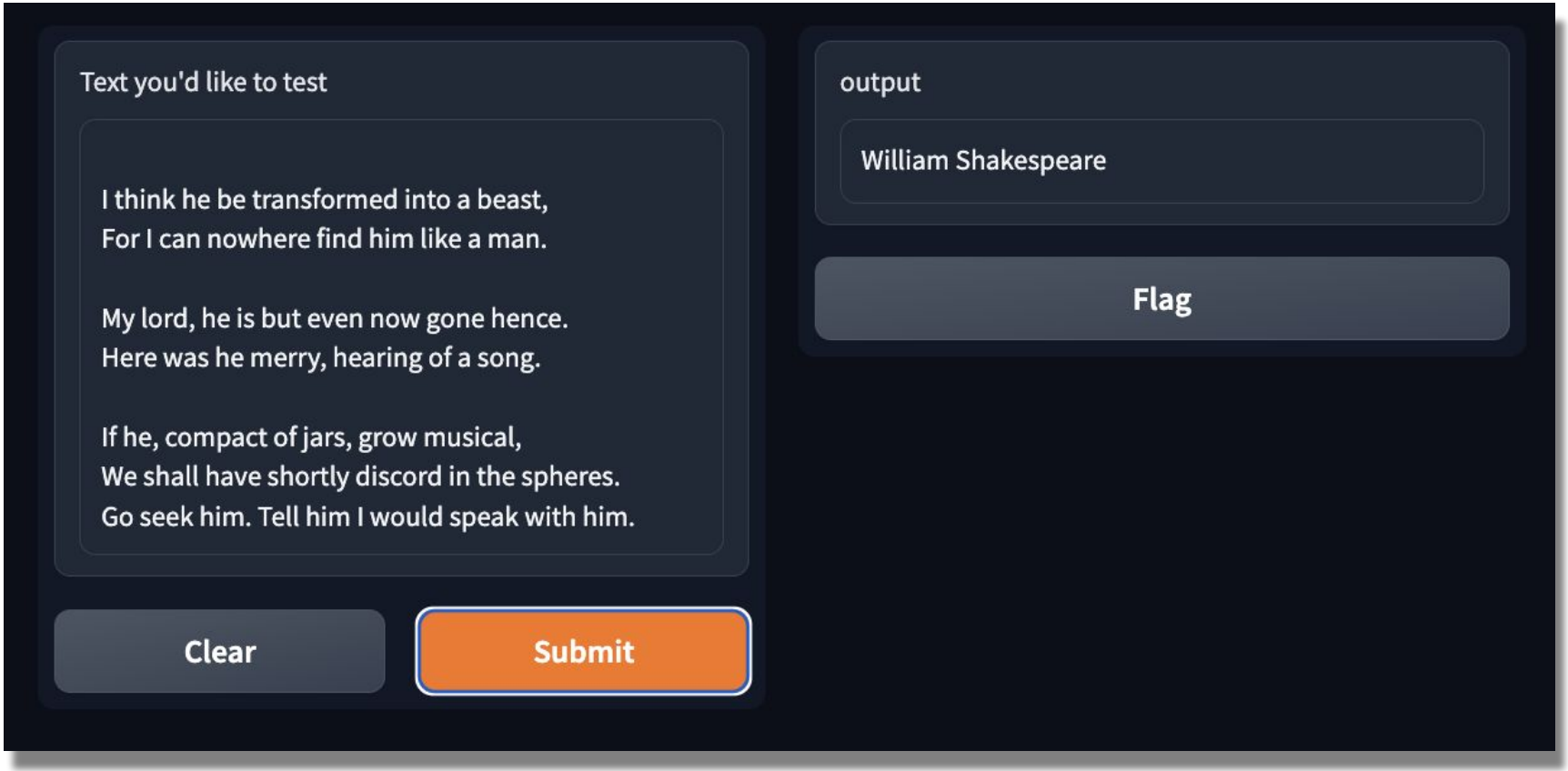
Support Vector Classifier (SVC)

- Accuracy: 69%

The Solution

User Friendly App via Gradio

Example from “As You Like It”



Text you'd like to test

His hands doth perspire, knees weak, arms art heavy, his sweater bearers a foul crimson mark, his mother's spaghetti

Clear | Submit

output

John Fletcher

Flag

His hands doth perspire, knees weak, arms art heavy, his sweater beareth a foul crimson mark; his mother's spaghetti

Recommendations

Base Use

- Input text of unknown author, get authorship guess

Tips

- Use as much text as possible
- Test scene, act, or whole play

Note:

- Model analyzes style, not fingerprints

Next steps

Expand Data

Add more plays to train, genre examination, etc.

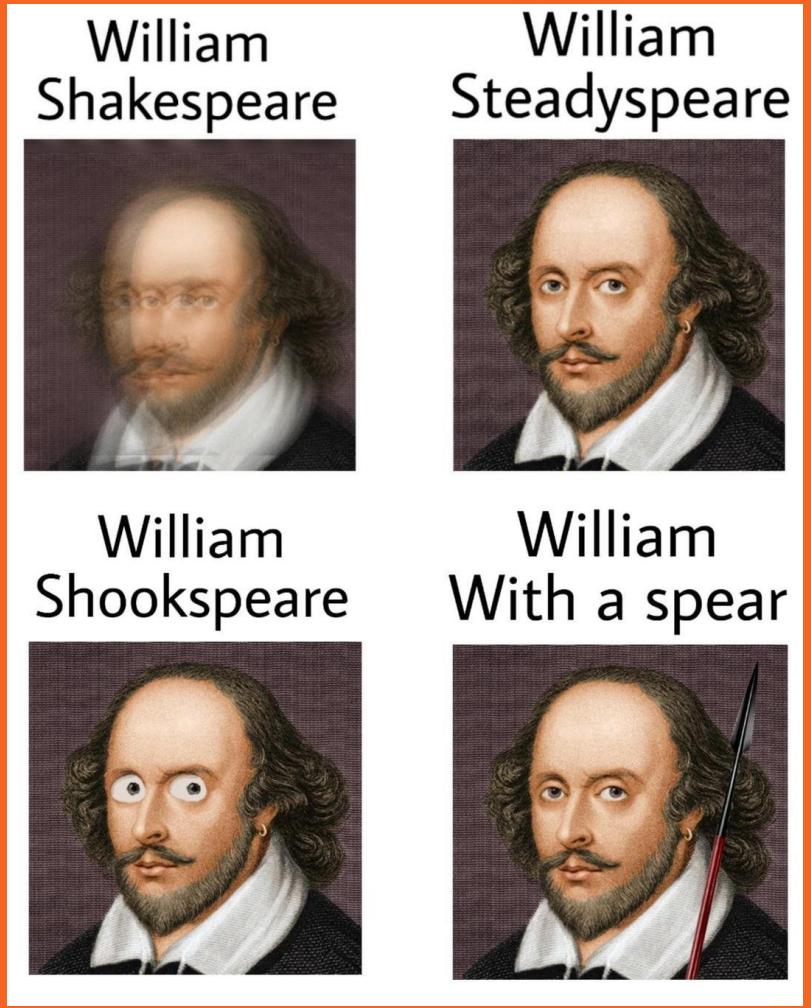
Link Model to Functions

Maximize functionality, automate analysis

Retrain Model

Experiment with genres, solve pen names, etc.

Thank you!


William Shakespeare
William Steadyspeare
William Shookspeare
William With a spear