

Identifying Shakespearean Authorship



Drew Holcombe

Data Scientist

Washington, DC

Agenda

- Problem and data understanding
 - Modeling process
 - Results and conclusions
 - Limitations and next steps



Understanding Questions of Authorship

- Frequent collaboration
 - John Fletcher, Thomas Middleton, George Wilkins
- Timon of Athens
 - Low quality
 - Stylistically unusual



Data

Shakespeare

Othello
King Lear
Antony and Cleopatra
Coriolanus

14,396 lines of dialogue

Middleton

A Trick to Catch the Old One
The Phoenix
The Puritan
The Revenger's Tragedy
Your Five Gallants

8,845 lines of dialogue

Modelling

Accuracy

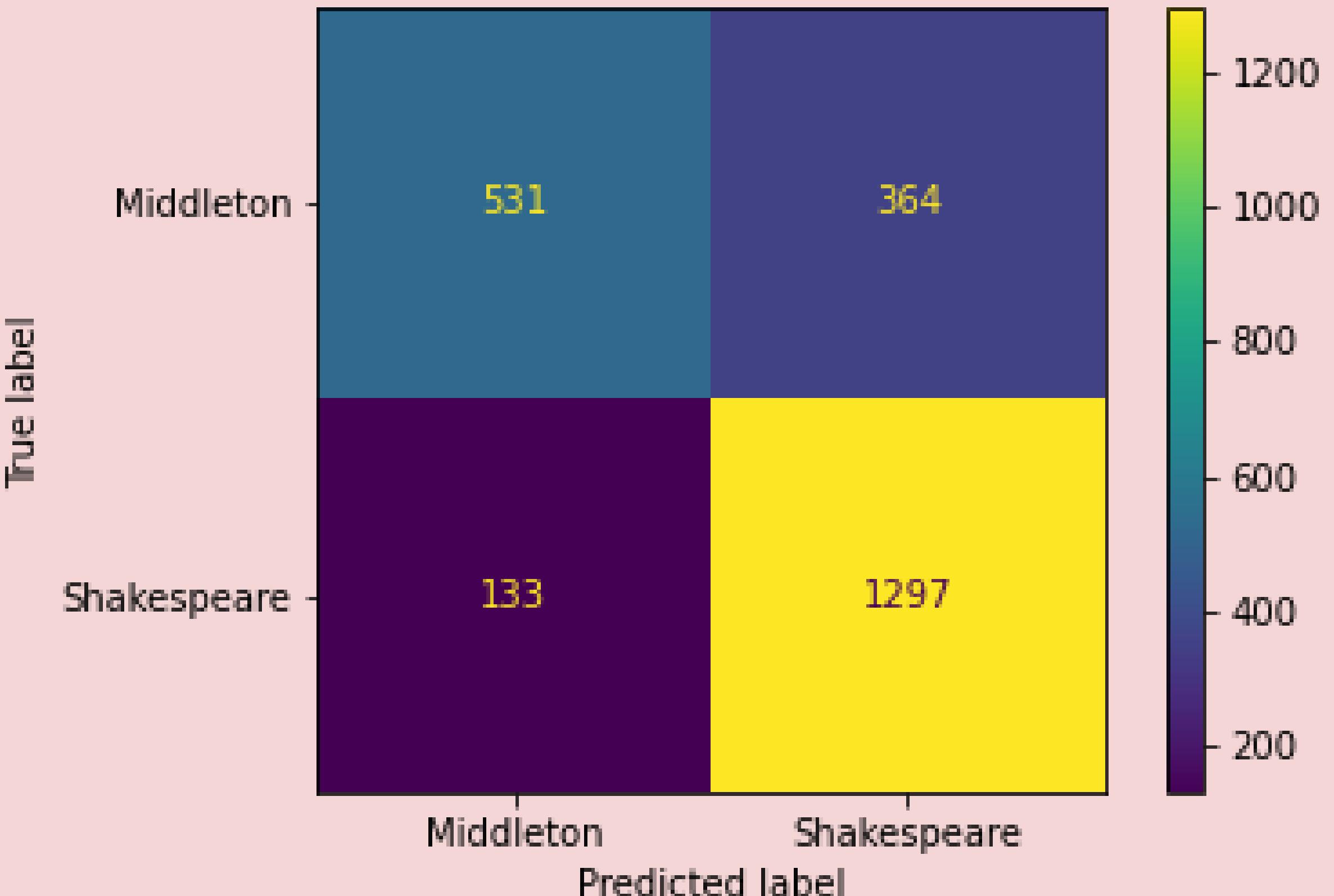
78.6%

Precision

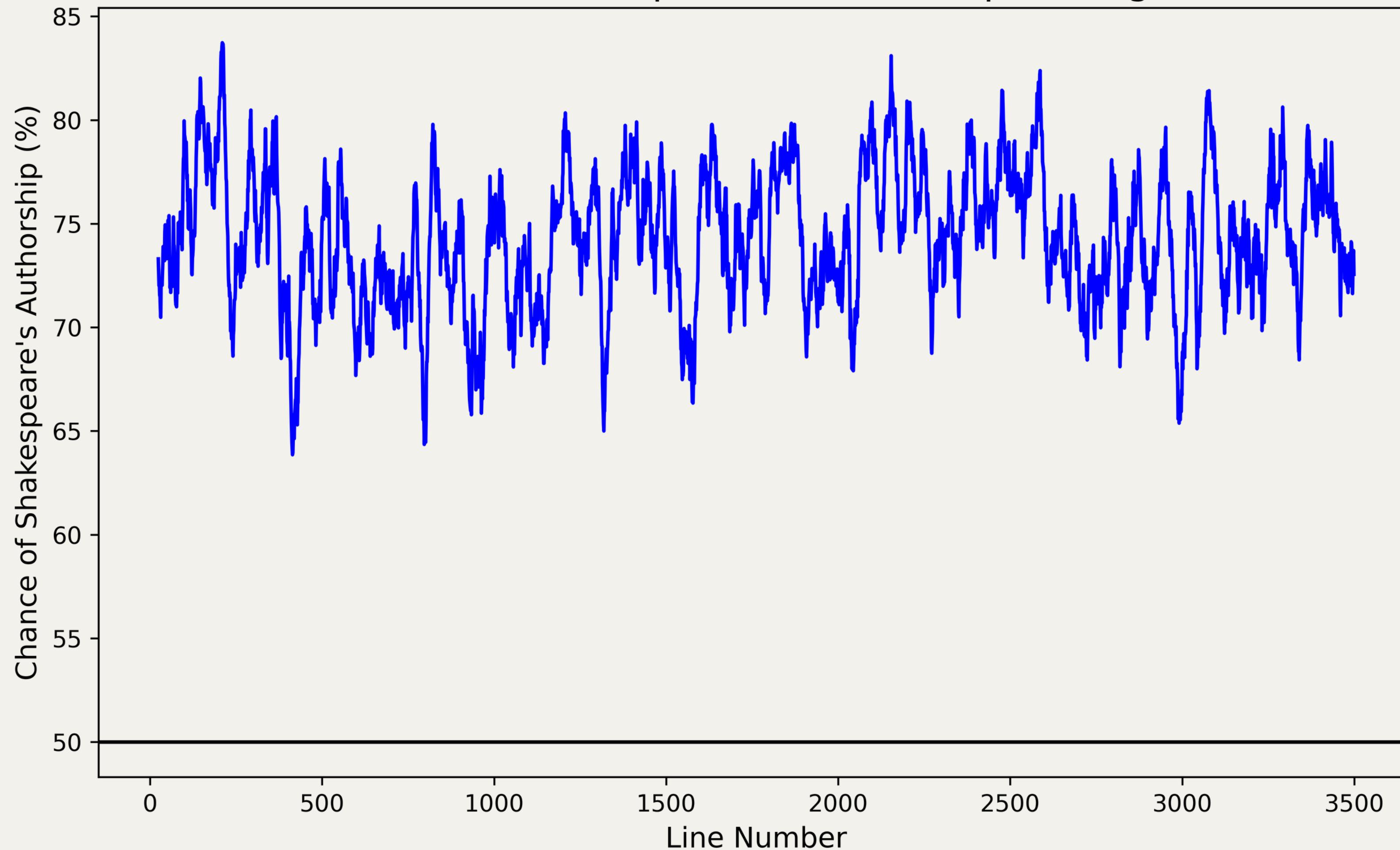
78.1%

Recall

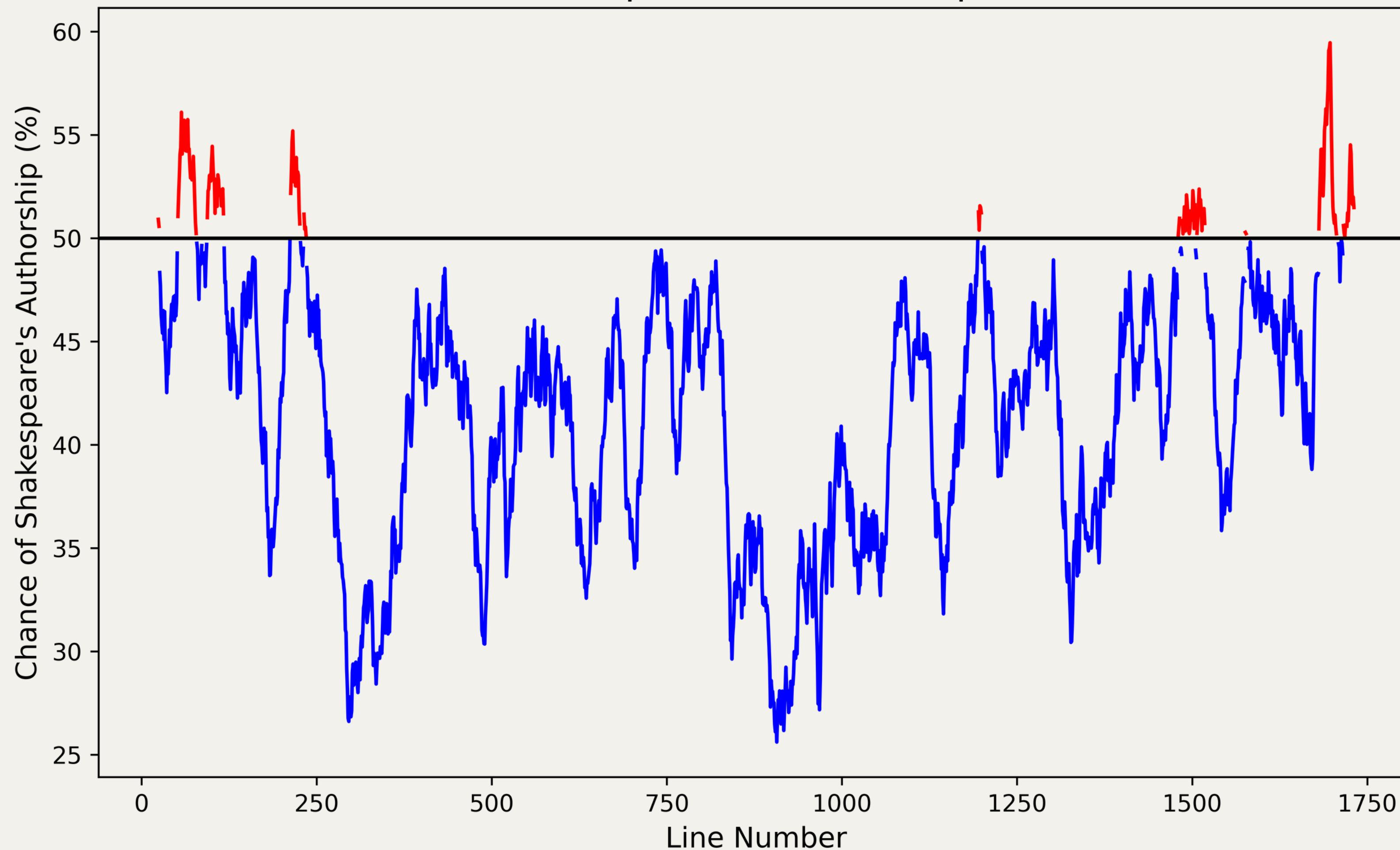
90.7%



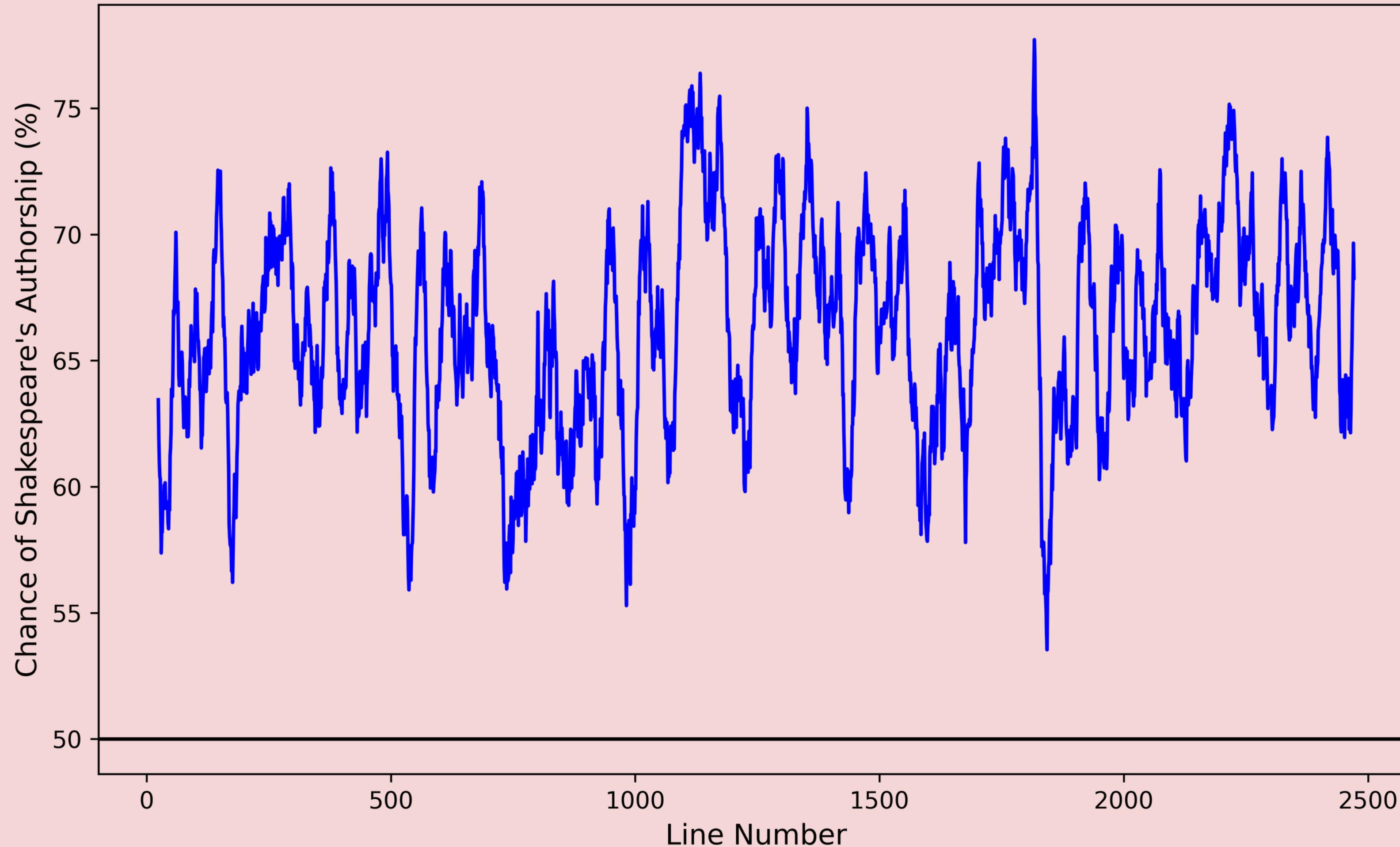
Likelihood of Shakespeare's Authorship of King Lear



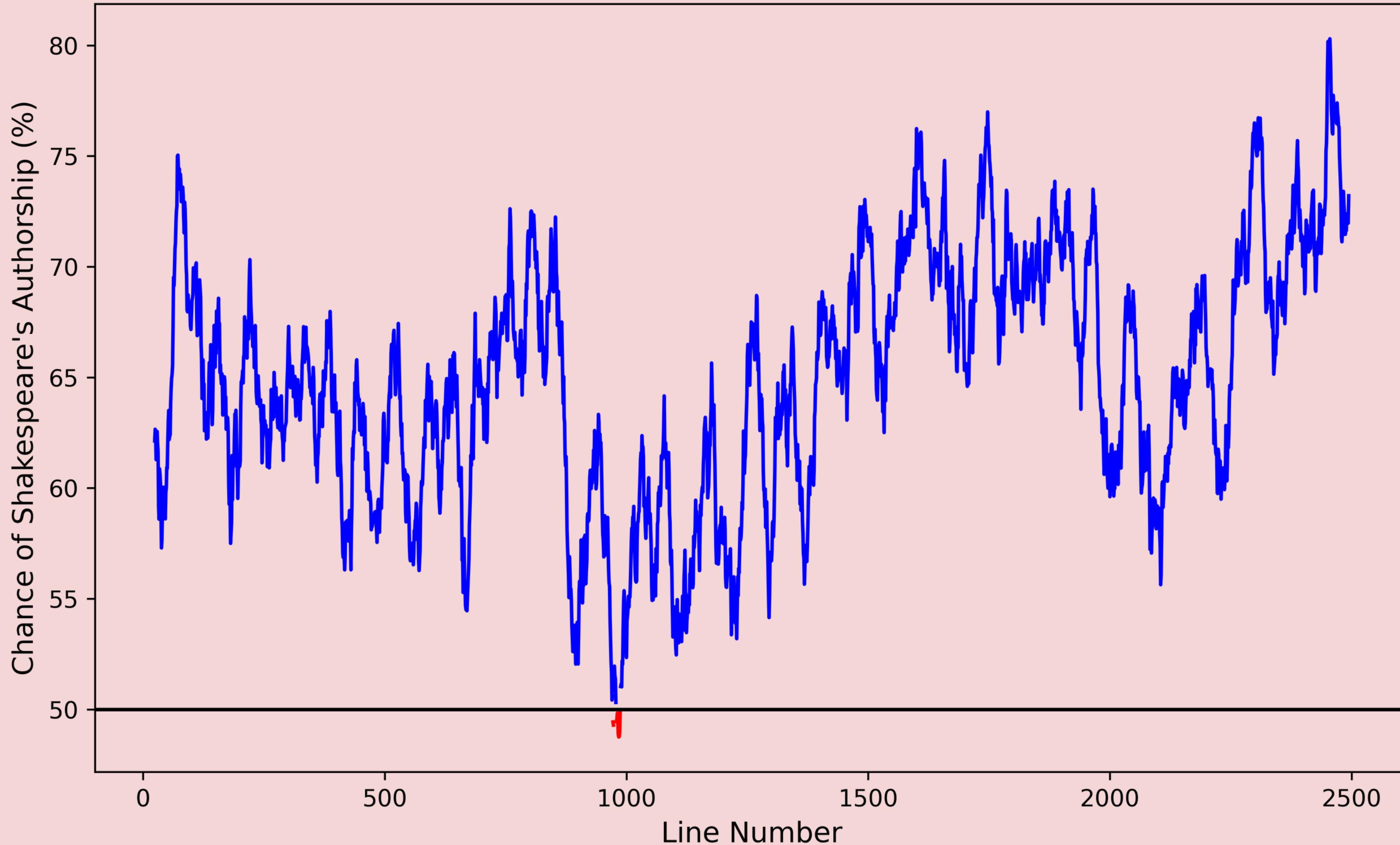
Likelihood of Shakespeare's Authorship of The Phoenix



Likelihood of Shakespeare's Authorship of Pericles, Prince of Tyre



Likelihood of Shakespeare's Authorship of Timon of Athens



Conclusions

Some evidence Middleton
contributed to
scenes 5 through 10

No ironclad findings

Next Steps

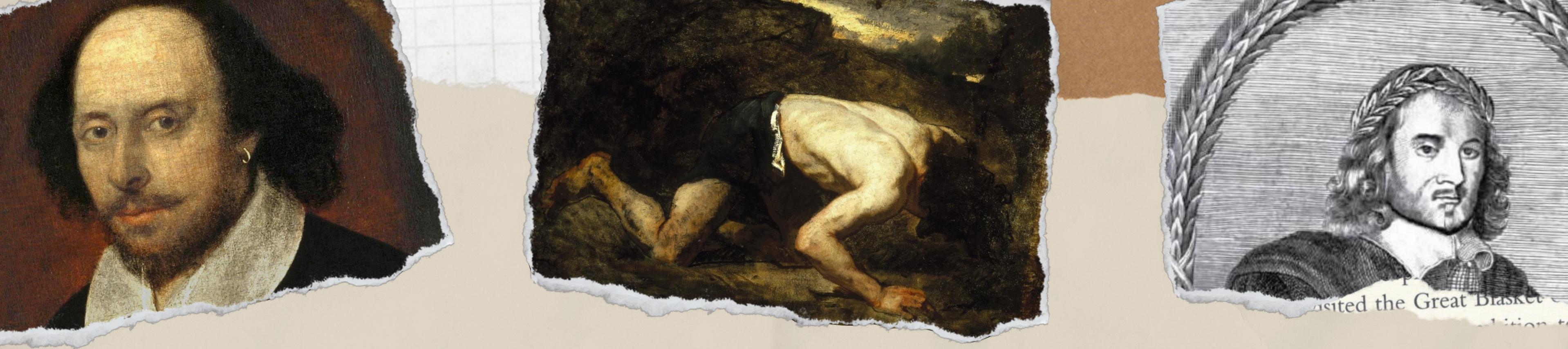
Train on additional plays by Middleton

Consider other collaborators

Continue iterative modeling process

Apply methods to other literary works
with contested authorship





Thank you!

drew.holcombe7@gmail.com

[linkedin.com/in/drewfholcombe/](https://www.linkedin.com/in/drewfholcombe/)

github.com/DrewFHolcombe

