* 1. The confusion matrix with Laplace Smoothing is when 80% of the data is training data and 20% test data.

[ 2. 1. 1. 15. 33. 9. 9.]

[1. 0. 0. 15. 27. 16. 17.]

[1. 0. 9. 11. 48. 7. 21.]

[1. 0. 0. 43. 70. 13. 35.]

[1. 0. 0. 14. 143. 22. 41.]

[1. 0. 1. 18. 55. 26. 45.]

[1. 0. 0. 23. 56. 24. 66.]

**Accuracy: 0.3067940552016985**

* 1. The confusion matrix with Laplace smoothing and using 80% test and 10% validation data.

[14. 6. 4. 0. 7. 1. 3.]

[ 3. 11. 0. 5. 8. 5. 6.]

[ 4. 8. 14. 3. 11. 1. 8.]

[ 3. 4. 2. 26. 28. 3. 15.]

[ 5. 2. 3. 6. 51. 21. 23.]

[ 3. 4. 2. 3. 17. 16. 28.]

[ 6. 5. 5. 12. 16. 9. 32.]

**Accuracy: 0.3474576271186441**

The NB classifier with the best accuracy on the validation was the one with a smoothing constant of 285.

[18. 3. 4. 2. 5. 2. 1.]

[5. 16. 4. 2. 2. 4. 5.]

[2. 2. 23. 3. 6. 4. 8.]

[3. 5. 0. 21. 25. 15. 12.]

[4. 6. 6. 6. 55. 11. 22.]

[6. 5. 3. 8. 18. 19. 14.]

[2. 6. 6. 11. 13. 15. 32.]

**Accuracy: 0.39148936170212767**

* 1. The NB classifier with the best accuracy on the test was the one with a smoothing constant of 285.

[14. 6. 4. 0. 7. 1. 3.]

[3. 12. 0. 5. 8. 5. 6.]

[4. 8. 13. 3. 11. 1. 8.]

[3. 4. 2. 26. 28. 3. 15.]

[5. 2. 3. 6. 51. 21. 23.]

[3. 4. 2. 3. 17. 16. 28.]

[6. 5. 5. 12. 16. 9. 32.]]

**Accuracy: 0.3474576271186441**

This indicates that the choice of smoothing value has a significant impact on the accuracy of the model.

* 1. The confusion matrix shows that books 5,6 and 7 are the most confused. This show that the JK Rowlings writing style has changed over time but books that where written closer together are more confused with each other.