Part I

Question 2. The information is challenging for the public to comprehend. (-2)

Question 3. Consider the overall trend of the word count over time. (-2)

Question 4. Because in the data, 86% of the changes consists of negative sentiment. (-1)

Part III

Section 0: Grounds?(-1)

Section 2: did you check the warning message? (-1)

Part IV. You have to try binary weight when testing precision/recall at first.

So what’s your answer? (-5)

The naïve bayes model does not work well on high dimensional data. You have to provide your intuition based on the metrics you found. (-5)

Part III

Section 1

Q2. What’s your understanding on test data?

Q4. Please read the question again. I was asking the top 10 tokens extracted from the review text you found in the question Q3.

Part III

Section 2

Did you check your code? You should have put the text file in `vectorizer.fit\_transform()`. Your `X\_train` includes other unnecessary columns so that the method does not work. Instead, why not trying `vectorizer.fit\_transform(X\_train.loc[‘text’])`

Why did you give up? I wish you had asked me for help.

Part I

Q1. They look for firm’s

Part I

2. So, why did the stock price not change? (-2)

Part IV. What insights can we derive from the metrics? When should we prioritize precision over recall? What factors contribute to Naive Bayes having the highest precision? Conversely, why does the hand-written rule logistic regression exhibit lower precision but higher recall? Ultimately, the crucial question is whether these findings are meaningful and coherent. (-5)