1. What's the difference between the three Naive Bayes classifiers?

- GaussianNB: assumes features are continuous and normally distributed. Suitable for real-valued data (can be positive or negative).
- BernoulliNB: assumes binary (0/1) features. With continuous inputs, it binarizes them at a threshold, treating values simply as present or absent.
- MultinomialNB: assumes features are non-negative counts (e.g., word frequencies). It does not allow negative values; in this task, scaling was applied to transform features into count-like integers.

2. Precision and Recall Comparison (Class 1 = Fraud)

Model	Accuracy	Precision (Class 1)	Recall (Class 1)
GaussianNB	0.976	0.381	0.861
BernoulliNB	0.995	0.929	0.759
MultinomialNB	0.995	0.962	0.737

GaussianNB achieves the highest recall (0.861), which means it detects most of the fraud cases. However, its precision is very low (0.381), so more than half of its fraud predictions are actually false alarms. This model favors catching fraud but would overwhelm investigators with too many false positives. BernoulliNB shows a good balance between recall (0.759) and precision (0.929). It still catches the majority of fraud cases while keeping false alarms relatively low. This balance makes it more practical for real-world fraud detection. MultinomialNB achieves the highest precision (0.962), meaning almost all flagged frauds are truly fraudulent. However, its recall is lower (0.737), so it misses more fraud cases compared to GaussianNB. This makes it suitable if false alarms are very costly, but less ideal if missing fraud is more dangerous.

3. Which model would you choose for fraud detection, and why?

Fraud detection requires a balance between recall (catching as many frauds as possible) and precision (avoiding too many false alarms). GaussianNB had the highest recall (0.861) but very low precision (0.381), which means it would generate many false positives. MultinomialNB showed the best precision (0.962) but lower recall (0.737), which risks missing fraud cases. BernoulliNB offered the best balance, with strong precision (0.929) and good recall (0.759), resulting in the highest overall F1 score. Therefore, I would choose BernoulliNB, because it achieves high accuracy and provides the most practical trade-off between catching fraud and avoiding false alarms without extra preprocessing.