1A) when we duplicate n feature into n+1 and retrain the model. It would lead to multicollinearity which can cause issues with stability of the coefficients. This could weigths of the two will be disturbed.

2A) 1. We have too little data to conclude that A is better or worse than any other template with 95% confidence.

This statement suggests indecision due to insufficient data. To determine this, we need to perform statistical tests.

E is better than A with over 95% confidence, B is worse than A with over 95% confidence. You need to run the test for longer to tell where C and D compare to A with 95% confidence.

• This statement is making specific claims about the templates' performance. We can check

these claims are supported by statistical tests. 3. Both D and E are better than A with 95% confidence. Both B and C are worse than A with over 95% confidence.

• Similar to statement 2, this statement makes specific claims about the templates' performance.

Now, let's perform pairwise comparisons:

Template A (control) vs. Template B

* p-value = 1-0.95 0.05 (assuming a one-tailed test)
If the p-value is less than 0.05, we reject the null hypothesis that B is not different from A.
Template A vs. Template C
p-value = 1-0.915 (for a two-tailed test, as it's not specified in the statement)
If the p-value is less than 0.05, we reject the null hypothesis that C is not different from A.
Template A vs. Template D
* p-value = 1-0.88
• If the p-value is less than 0.05, we reject the null hypothesis that D is not different from A.
• Template A vs. Template E
* p-value = 1-0.86
If the p-value is less than 0.05, we reject the null hypothesis that E is not different from A.
If we find that the p-values for comparisons with B and C are greater than 0.05, we can't confidently say that B is worse than A or that C is worse than A. Similarly, if the p-values for comparisons with D and E are less than 0.05, we can conclude that D and E are better than A.
So, let's evaluate the statements:
Statement 1: This could be true or false depending on the p-values

Statement 2: This statement could be partially true if the p-values for B and C are greater than 0.05, but it doesn't provide information about D and E.
Statement 3: This statement could be true if the p-values for D and E are less than 0.05, and the p-values for B and C are greater than 0.05.
To make a definitive conclusion, we would need to perform the statistical tests and compare the p-values.
3A) the computational cost is o(k)
4A)
Ranking based on potential impact on accuracy:
1. Method 1: This could have the highest impact on accuracy as it targets areas where the current model is uncertain.
2. Method 3: While it addresses cases where V1 is confidently wrong, it might not be as effective if the errors are due to uncertainty.
3. Method 2: It provides a general set of examples but may not be as targeted in addressing V1's weaknesses.

Overall:
• A combination of these methods might be beneficial. Incorporating challenging examples near the decision boundary (Method 1) and diverse examples (Method 2) could lead to a more robust model.
• It's important to validate the impact of these strategies through experimentation, as the actual effectiveness can depend on the nature of the data and the characteristics of misclassifications made by the initial model

5A) p MLE= (k/n)
P Bayesian =(k+1)/(n+2)
P Map = k/(n+1)