**Name: Manisha Koranga  
Roll No: BT20CSE033**

**College: IIIT Nagpur**

**Ans\_1-**When you copy n features by feature (n + 1) and retrain the new model using logistic regression, the repeated feature will have the same weight in the new model. The model assigns weights based on the contribution of features to the prediction, and if similar features provide comparable information, the model tends to distribute weight evenly across them. Therefore, in this case you can expect a high degree of correlation and redundancy in return characteristics, for example. However, note that the accuracy of the weights may vary depending on the specific dataset characteristics, optimization methods, and optimization methods used during the run.

**Ans\_2-**The result based on the data provided is:

2. If E is better than A, the confidence level exceeds 95%; If B is worse than A, the confidence level is 95%. To show with 95% confidence that C and D compare to A, you would need to run longer tests.

Template E has higher click-through rates. A and the difference is big. (14% vs 10%). The reliability of this result is quite high. Template B has a lower CTR than A (7% vs 10%). The reliability of this result is quite high. For samples C and D, it is recommended to run the test longer to determine how they compare to A with 95% confidence. This means that the data for C and D are currently uncertain and require further analysis before a reliable comparison can be made.

Therefore option 2 is the most accurate representation of the conclusions drawn from the given data.

**Ans\_3-** In logistic regression with sparse feature vectors, the computational value of each iteration of gradient descent can be estimated. The main steps include calculating the estimates, calculating the residuals, and calculating the slopes. For sparse data, the average number of non-zero entries in each training sample is k and k is less than n; The estimated value of each gradient descent is said to be O(k⋅m). Today's well-written packages (like scikit-learn in Python) are optimized for arbitrary data, further reducing computational costs.

**Ans\_4-** The third method seems to be the most effective way to improve the accuracy of V2 text in separating media into information and entertainment. The training data will be for complex problems by recording 1 million stories and selecting the subset where the V1 classifier is inaccurate and far from the decision limit. This strategy can improve products' performance on complex examples by providing more powerful models for different languages. The first approach focuses on examples close to the decision space that can give insight into potential areas for improvement, while the second approach provides a different tab but there are many types of notes and they are less accurate than the objective approach. that the effects of sex may be less significant. Therefore, to improve the accuracy of the V2 classifier, it seems best to address difficult problems where the original classifier got stuck.

**Ans\_5-** Given n tosses and the coin landing on heads, three different methods are used to estimate the probability of the coin landing on heads. The maximum likelihood estimate (MLE) is simply calculated as P\_mle = k / n, which represents the proportion of heads found for all tosses. For Bayesian estimation, assuming the prior is univariate, the posterior distribution is the Beta distribution with parameters (k + 1, n−k + 1) and the estimate is P\_bayesian = (k + 1) / (n + 2), which represents the value of the test. In the case of maximum a posteriori (MAP) estimation, the posterior beta distribution has the form P\_MAP = k/n and provides an estimate based on the most common value given observations and data first. Each method provides a unique view of estimating the probability of P, combining frequency bins, modified Bayesian, or searching for patterns in the presence of prior beliefs.