

CS 203: Software Tools & Techniques for AI
IIT Gandhinagar
Sem-II - 2024-25

Lab Assignment 10

Total: 10 Marks

Submission deadline: Wednesday, 16/04/2025 11:59:59 PM

Submission guidelines:

1. Implement the code in the jupyter notebook and share it during submission.
2. Add all screenshots of the results in a pdf file and share.
3. The deadline will not be extended at any cost.

Note: Submitting this assignment solution confirms that you will follow the IITGN's honor code. We shall strictly penalize the submissions containing plagiarized text/code.

23110049 - Aryan Solanki

23110237 - Parthiv Patel

Question 1

Z-score: 1.2130583676113753
P-value: 0.22510750571860028

Conclusion:

Z-score: 1.21 suggests a modest difference, but not one that's statistically relevant to us.

P-value: 0.225 indicates that the difference in click-through rates is not statistically significant.

What Does This Mean for our A/B Test?

Since the P-value is greater than 0.05, we do not have sufficient evidence to conclude that the ad position (Top vs. Side) significantly affects the click-through rate.

In other words, the difference observed in this test could easily be explained by chance, and both ad positions perform similarly in terms of click-through rate.

These results are **statistically insignificant** and do not support a strong preference for one ad position over the other.

Question 2

KS Test for test_1 vs train:

KS Statistic: 0.0190625, P-value: 0.9721940612395358

KS Test for test_2 vs train:

KS Statistic: 0.4075, P-value: 7.2019977111245e-96

No covariate shift detected in test_1.

Covariate shift detected in test_2.

We know, A covariate shift refers to a situation where the distribution of the input features (covariates) in the test data is different from the distribution in the training data, even though the relationship between the features and the target variable (NO2(GT) in our case) remains the same.

test_1 shows no covariate shift, meaning its data is similar to what the model saw during training.

test_2 exhibits covariate shift, meaning the distribution of NO2(GT) in test_2 is quite different from what was seen in the training data, which could impact model performance.

Hence, based on the results of the Kolmogorov-Smirnov (KS) test, we can ideally use test_1 for evaluating our model, since test_1 does not show covariate shift relative to the training dataset.