**Explainable ai for scheme recommendation model using Permutation Feature Importance Approach**

The objective of this analysis was to compute the feature importance of different features (specifically geography and stockist type) on the recommendations generated by a user-based recommendation model. The model was tested using an imbalanced dataset, where the features included geography columns (Geography\_East, Geography\_North, Geography\_South, Geography\_West) and stockist types (Stockist\_Type\_Distributor, Stockist\_Type\_Retailer, Stockist\_Type\_Wholesaler). The purpose was to evaluate how changing these features impacts the quality of the recommendation system and identify which features contribute the most to changes in recommendations.  
  
  
**2. Methodology**

**2.1. Permutation Feature Importance (PFI) Approach**

To assess feature importance, we employed the **Permutation Feature Importance (PFI)** approach. This methodology involves the following steps:

1. **Baseline Model Generation**:
   * The baseline recommendation system is generated by predicting user-product recommendations based on the current data.
   * We compute these recommendations using a user-based collaborative filtering approach, and the recommendations are based on user-specific preferences (e.g., Partner\_id, Product\_id, and their associated similarity scores).
2. **Perturbation**:
   * For each feature (geography and stockist type), we create a perturbed dataset. This is done by randomly shuffling the values of the chosen feature across the rows in the dataset. This randomization aims to disrupt the relationship between the feature and the target (recommendations).
3. **Retraining Model with Perturbed Data**:
   * After perturbing a feature, we retrain the recommendation model on the perturbed data and generate new recommendations. This helps us evaluate the impact of the feature perturbation on the recommendation results.
4. **Comparison of Recommendations**:
   * We compare the baseline recommendations with the perturbed recommendations to see how much they differ. If the recommendations change significantly, the perturbed feature is considered important.
5. **Importance Score Calculation**:
   * The importance score of each feature is calculated as the proportion of changes in the recommendations when the feature is shuffled. This is computed by counting the number of changes in the recommendations and dividing it by the total number of rows in the dataset.

**3.1. Feature Importance Insights:**

The results from the feature importance analysis indicate that both geography and stockist type features contribute to the variation in recommendations generated by the model. However, the magnitude of their impact may vary depending on the underlying data distribution.

* **Geography Features**:
  + Geography features (East, North, South, West) show significant impact on the recommendation output. This suggests that the geographical location of users influences their product preferences and, consequently, the product recommendations they receive.
* **Stockist Type Features**:
  + Similarly, stockist types (Distributor, Retailer, Wholesaler) also impact the recommendations. Different stockist types may represent different market segments with distinct purchasing behaviors, which affect the relevance of recommendations provided to the users.

**3.2 Assumptions, limitations**

* **Shuffling Assumption**:  
  The shuffling approach assumes that the feature values are randomly distributed when shuffled, and that any meaningful relationship between the feature and recommendations will be destroyed. After shuffling, if the feature is crucial for the recommendation, we expect the recommendations to change, and thus, the feature’s importance score will be higher.
* **Dropping Assumption**:  
  In the dropping approach, the assumption is that removing the feature will result in a loss of information that should affect the recommendations. This is because the feature is expected to contribute meaningfully to the recommendation system. If dropping the feature doesn’t change recommendations much, the feature is likely less important.

**Limitations**

* **Features with Low Variance**:  
  If a feature has low variance or few distinct values (e.g., a categorical variable with a small number of categories), shuffling or dropping that feature might not cause a large enough change in recommendations to accurately measure its importance. For example, a feature with a single value across most of the data may not affect recommendations even if it were shuffled or removed.

**Edge Cases:**

* **Missing or Sparse Data**: In cases where features have missing values or are sparse, the effect of perturbing those features may not be meaningful or may lead to poor model performance.
* **Outliers**: Outliers in feature values might distort the importance measurement, as these values could disproportionately affect the model’s output.
* **Feature Interactions**: If the features interact with each other (e.g., specific combinations of geography and stockist type are highly correlated with certain products), shuffling one feature may not fully capture the joint effect.