

The summary of our approach:

1. **Visualization** – The input data was visualized through spotfire with the sum of PriceUp values, and following was observed
 - a. Certain brands and colors had frequent PriceUps as compared to other
 - b. Wednesday had the highest number of PriceUp count
 - c. Some websites up the prices more frequently than others
 - d. Among the two categories mobiles and cameras, there was no priceUp observed in input data for cameras.
2. **Missing values** - A few variables chosen as predictors had missing values that were handled as below:
 - a. AvgRating – Replaced with mean
 - b. ReviewCount – Replaced with 0
 - c. Color – Replaced with NA
 - d. Shipping period – Replaced with 10+ days
3. **Categorization Scores**
 - a. Converted shipping period values in to the categorical scores grouping together the similar values in to single category score.
4. **Derived Variable** – Weekday variable was derived from date
5. **Creating Dummies** – The following variables were chosen as categorical variables based on the visualization observations and dummy variables were created
 - a. Brand – After converting the case
 - b. Color
 - c. SiteName
 - d. Weekday
 - e. Category
6. **Dimension Reduction**
 - a. Top 12 brands with highest priceUp count were chosen from 30 dummy variables for brand.
 - b. Top 6 colors with highest priceUp count were chosen from 20 dummy variables for color.
7. **Data Partition**
 - a. A part of training data set was used as validation set 6200 – 7765 records.
8. **Final Model**
 - a. **Logistic Regression classification Model** was generated including the categorical and transformed variables as discussed above, including variables inStock, freeShipping and timeNextPrice to arrive at the predicted class and success probability.
 - b. **Discriminant Analysis classification Model** was generated including the categorical and transformed variables as discussed above, including variables inStock, freeShipping and timeNextPrice to arrive at the predicted class and success probability.