* ***APPROACHES***

In total we have 15 columns, which are:

* uniq\_id
* crawl\_timestamp
* product\_url
* product\_name
* product\_category\_tree
* pid
* retail\_price
* discount
* image
* is\_FK\_Advantage
* description
* product\_rating
* overall\_rating
* brand
* product\_Specification

Clearly, out of these attributes; uniq\_id, crawl\_timestamp, pid, image, is\_FK\_Advantage, product\_rating, overall\_rating don’t affect the classification. Therefore, we will be considering other attributes.

In our approach:

1. Features = {*“description”*}

In this case we only have 1 Independent Variable [“description”] and 1 Dependent Variable [“Category”].

1. Features = {}

The following are the methods used:

* Decision Tree
* Random Forest
* Gradient Boosting
* ***RESULTS***