**Introduction**

Ratings given on a web-site can sometimes be biased. For example if a product is given a 5 star rating and description reads really bad review for the product then the rating is biased. The primary objective of this document is to consider various attributes of a rating and review and arrive at an overall score for the review added by the user.

**Attributes**

There are primarily 5 attributes considered to arrive at the algorithm:

* **verified\_purchase:** if a user writing the review has actually purchased the product. If an entity is verified purchase weight (w) is more.
* **helpfulness\_count:** how many people found the review helpful.
* **images\_count:** if images are uploaded, the review is more useful for the user.
* **words\_count:** total word count of the description.
  + w(helpfulness) > w(words count)
* **rating\_confidence:** This is a factor of rating and sentiment analysis.

**Algorithm:**

We give highest weightage if an entity is a verified purchase. And Helpfulness has more weight than words count.

These are the scores that are calculated:

1. **helpfulness\_score:**

If verified purchase, then weight is 5 for helpfulness\_count, else it is 1.

*If verified\_purchase:*

*helpfulness\_count / 100 \* w(5)*

*else*

*helpfulness\_count / 100 \* w(5)*

1. **images\_score:**

If verified purchase, then weight is 1 for images\_count, else it is 0.75.

*If verified\_purchase:*

images*\_count \* w(1)*

*else*

images*\_count \* w(0.75)*

1. **words\_score:**

If verified purchase, then weight is 2 for words\_count, else it is 0.5.

*If verified\_purchase:*

*words\_count / 100) \* w(2)*

*else*

*words\_count / 100) \* w(0.5)*

1. **rating\_confidence:**

Rating confidence is a deviation factor of ratings and sentiments. If rating is too far from sentiment, then we cannot trust the rating then rating confidence becomes 0 else it is a factor of rating/10 that needs to be added to overall score.

Example:

If rating is 5 but sentiment is 1 (very sad), it implies that even thought user has given high rating, the title and description have implied that user is sad.

***rating\_confidence = IF (ABS (rating – sentiment) ) < 2 THEN rating/10 ELSE 0***

1. **review\_score:**

This is an average of helpfulness\_score, words\_score and then addition of rating\_confidence and images\_score:

Review\_score = rating\_confidence + images\_score

+ (helpfulness\_score + words\_score) / 2

**Sample Data:**

**Table

Description automatically generated**

**Table

Description automatically generated**