Metrics

Learn about the metrics to capture the success of a feed-ranking system.

> **We'll cover the following** ∧

- User actions
- User engagement metrics
  - Selecting feed optimization metric
  - Negative engagement or counter metric
  - Weighted engagement

The feed-ranking system aims to maximize user engagement. So, let's start by looking at all the user actions on a Twitter feed.

User actions#

The following are some of the actions that the user will perform on their tweet, categorized as positive and negative actions.

Positive user actions

- Time spent viewing the tweet
- Liking a Tweet
- Retweeting
- Commenting on a Tweet
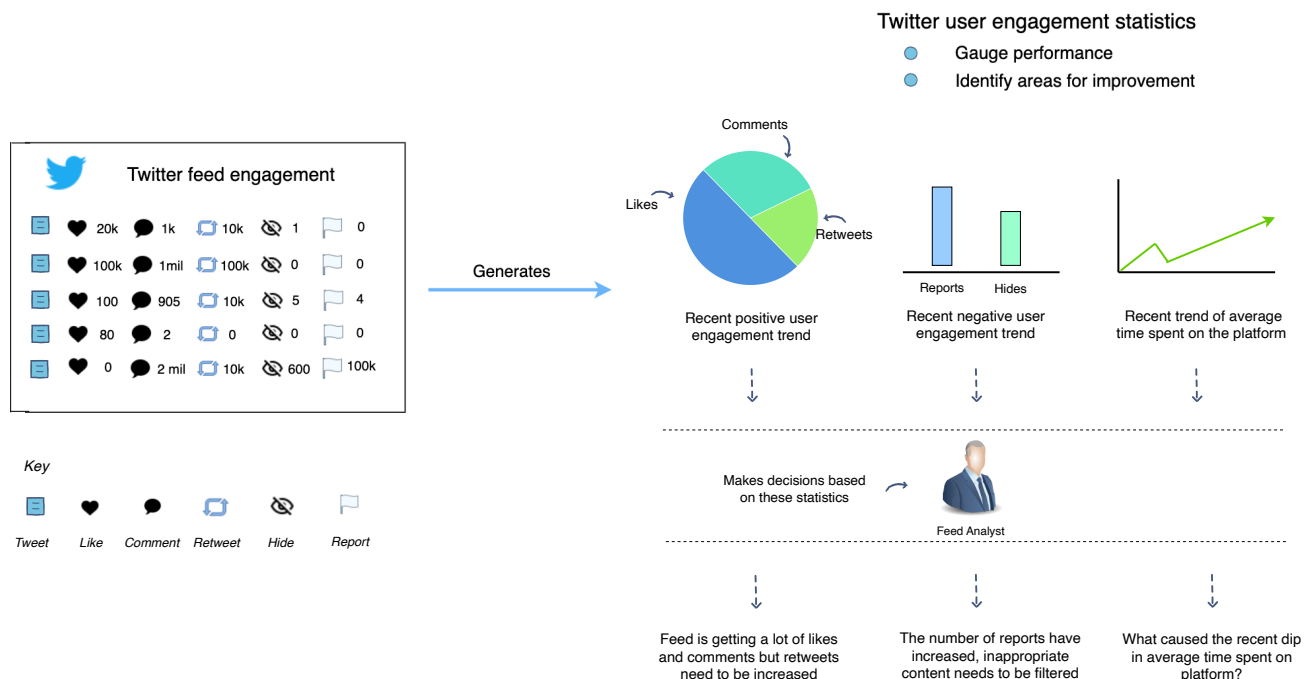
Negative user actions

- Hiding a Tweet

- Reporting Tweets as inappropriate

Now, you need to look at these different forms of engagements on the feed to see if your feed ranking system did a good job.

> 📝 All of the metrics we are going to discuss can be used to determine the user engagement of the feed generated by our feed ranking system.

User engagement metrics#

The following illustration shows how *feed engagement data* generates useful statistics for gauging user engagement.

User actions on the platform can help generate useful statistics

The above illustration shows different positive and negative engagements on a Twitter feed. Let's see which of these engagements would be a good one to target as your *overall system metric* to optimize for.

target as your overall system metric to optimize for.

Selecting feed optimization metric#

An important thing to understand in selecting a topline is that it's scientific as well as a business-driven decision.

The business might want to focus on one aspect of user engagement. For instance, Twitter can decide that the Twitter community needs to engage more actively in a dialogue. So, the topline metric would be to focus more on the **number of comments** on the Tweets. If the average number of comments per user increases over time, it means that the feed system is helping the business objective.

Similarly, Twitter might want to shift its focus to overall engagement. Then their objective will be to increase average overall engagement, i.e., **comments, likes, and retweets**. Alternatively, the business may require to optimize for the time spent on the application. In this case, **time spent on Twitter** will be the feed system metric.

Negative engagement or counter metric#

For any system, it's super important to think about counter metrics along with the key, topline ones. In a feed system, users may perform multiple negative actions such as reporting a Tweet as inappropriate, block a user, hide a Tweet, etc. Keeping track of these negative actions and having a metric such as **average negative action** per user is also crucial to measure and track.

Weighted engagement#

More often than not, all engagement actions are equally important.

However, some might become more important at a particular point in time

However, some might become more important at a particular point in time, based on changing business objectives. For example, to have [a] young audience, the number of comments might be more critical rather than just likes. As a result, we might want to have different weights for each action and then track the overall engagement progress based on that weighted sum. So, the metric would become a **weighted combination of these user actions**. The weighted combination metric can be thought of as a value model. It will summarize multiple impacts (of different forms of user engagements) into a single score. Let's see how this works.

In the following illustration, we are going to use the *weighted combination metric* to measure/score user engagement. Each user action will be assigned a weight according to the impact it should have towards the final score.

> 📝 These weights are assigned, keeping in mind the respective importance of different forms of engagement towards the business objectives.

Value model to measure user engagement
on Twitter feed

Engagement Aggregates for
the Twitter feed

Multiple impacts summarized into a single score
to measure user engagement

Tweet viewed	2000
Likes	70
Comments	80
Retweets	20
Reports	5

	Occurrence Aggregate	Weight	Weighted Impact
Likes	70	0.4	28
Comments	80	0.2	16
Retweets	20	0.3	6
Reports	5	-0.1	-0.5
Weighted Score			49.5

Negative
weight

$$\text{Normalized Score} = \frac{\text{Score}}{\text{Total number of active users}} = \frac{49.5}{1000} = 0.0495 = 4.95\%$$

Using the weighted metric to measure the performance of the feed, i.e., it's user engagement

The user engagements are aggregated across all users' feeds over a specific period of time. In the above diagram, two-thousand tweets were viewed in a day on Twitter. There were a total of seventy likes, eighty comments, twenty retweets, and five reports. The weighted impact of each of these user engagements is calculated by multiplying their occurrence aggregate by their weights. In this instance, Twitter is focusing on increasing "likes" the most. Therefore, "likes" have the highest weight. Note that the negative user action, i.e., "report", has a negative weight to cast its negative impact on the score.

The weighted impacts are then summed up to determine the score. The final step is to normalize the score with the total number of active users. This way, you obtain the engagement per active user, making the score comparable. To explain the importance of normalization, consider the following scenario.

The score calculated above is referred to as score A, in which we considered one-thousand active users. We now calculate the score over a different period where there are only five-hundred active users, referred to as score B. Assume that score B comes out to be less than score A. Now, score A and score B are not comparable. The reason is that the decrease in score B may just be the effect of less active users (i.e., five-hundred active users instead of one-thousand active users).

When it comes to interpretation, a higher score equates to higher user engagement.

The weights can be tweaked to find the desired balance of activity on the platform. For example, if we want to increase the focus on commenting, we can increase its weight. This would put us on the right track, i.e., we would be showing Tweets that get more comments. This will lead to an increase in the overall score, indicating good performance.

← **Back**

Problem Statement

Next →

Architectural Components

✅ Mark as Completed

⚠ Report an Issue