



# Diversity

Let's learn methods to reduce monotony on the user's Twitter feed.

We'll cover the following

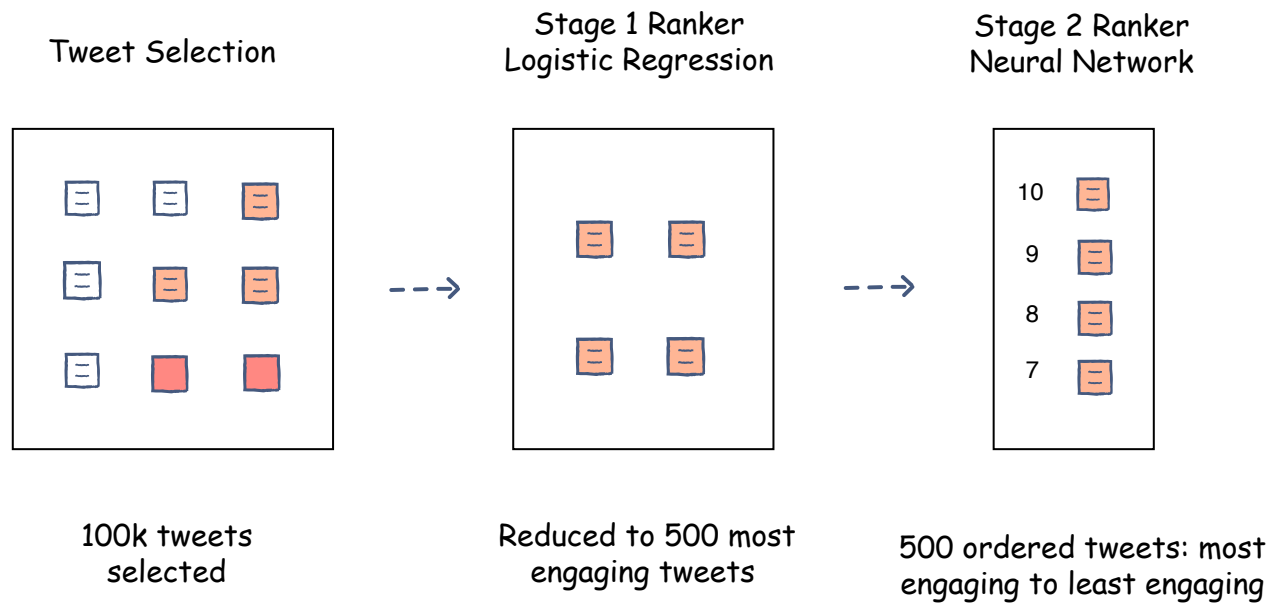


- Why do you need diverse Tweets?
  - Diversity in Tweets' authors
  - Diversity in tweets' content
- Introducing the repetition penalty

## Why do you need diverse Tweets?#

Let's assume that you adopted the following modelling option for your Twitter feed ranking system. The Tweet selection component will select one-hundred thousand Tweets for user A's Twitter feed. The stage one ranker will choose the top five-hundred most engaging Tweets for user A. The stage two ranker will then focus on assigning engagement probabilities to these Tweets with a higher degree of accuracy. Finally, the Tweets will be sorted according to the engagement probability scores and will be ready for display on the user's Twitter feed.





The layered model approach

Consider a scenario where the sorted list of Tweets has five consecutive posts by the same author. No, your ranking model hasn't gone bonkers! It has rightfully placed these tweets at the top because:

- The logged-in user and the Tweet's author have frequently interacted with each other's Tweets
- The logged-in user and the Tweet's author have a lot in common like hashtags followed and common followees
- The author is very influential, and their Tweets generally gain a lot of traction



This scenario remains the same for any modelling option.

## Diversity in Tweets' authors#

However, no matter how good of a friend the author is or how interesting



with regards to the Tweets’ author.

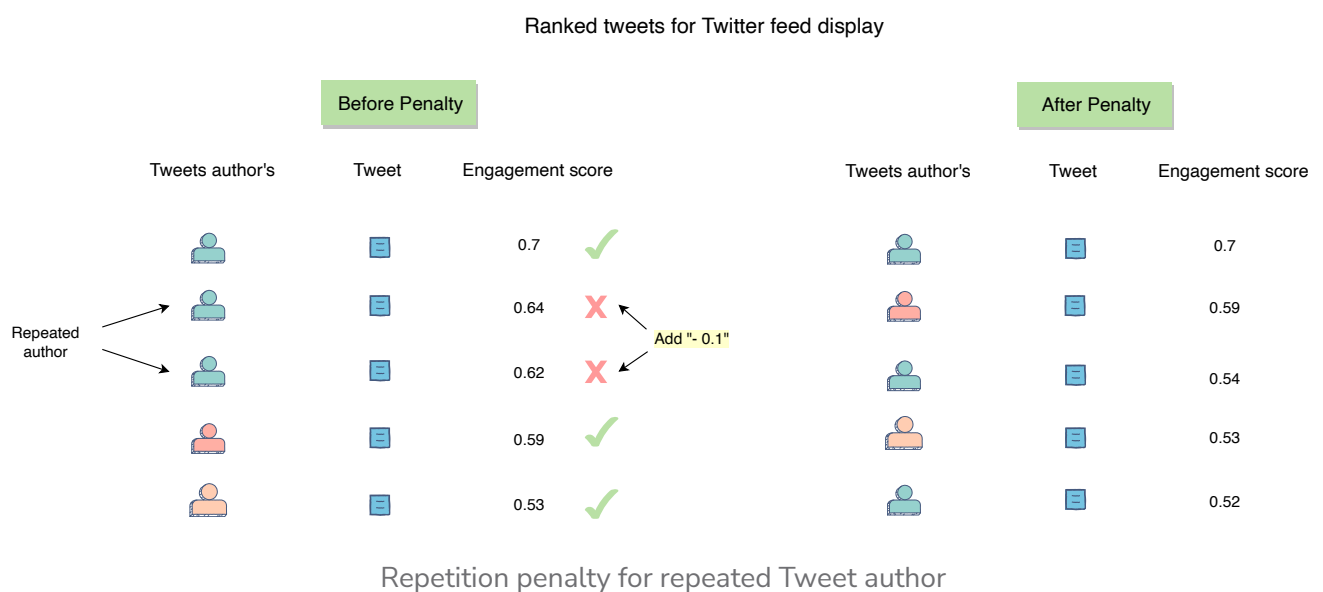
## Diversity in tweets’ content#

Another scenario where we might need to introduce diversity is the Tweet’s content. For instance, if your sorted list of Tweets has four consecutive tweets that have videos in them, the user might feel that their feed has too many videos.

## Introducing the repetition penalty#

To rid the Twitter feed from a monotonous and repetitive outlook, we will introduce a repetition penalty for repeated Tweet authors and media content in the Tweet.

One way to introduce a repetition penalty could be to add a negative weight to the Tweet’s score upon repetition. For instance, in the following diagram, whenever you see the author being repeated, you add a negative weight of  $-0.1$  to the Tweet’s score.



Another way to achieve the same effect is to bring the Tweet with repetition three steps down in the sorted list. For instance, in the following diagram,

when you observe that two consecutive Tweets have media content in them, you bring the latter down by three steps.

