**GOODREADS RECOMMENDATION SYSTEM**

Goodreads 10k dataset contains ratings for ten thousand popular books found on the internet. In this dataset there are 100 reviews for each book, although some have less - fewer - ratings. Ratings are given from 1 to 5.

Both book IDs and user IDs are contiguous. For books, it ranges from 1-10000, for users, 1-53424. All users have made at least two ratings. Median number of ratings per user is 8.

The main goal of this project is to suggest new books to the user that may interest them to increase the user engagement.

## **Collaborative filtering with model-based approach**

The idea of collaborative filtering is that the users who have agreed in the past tend to also agree in the future. If A and B read similar books, A is more likely to read a book that B has read a book which a random person has read.

### **Neural networks for recommendation**

We can estimate the ratings of the user for unread books by creating a neural network model for recommendation.

In this approach, instead of taking the dot product of user and book embeddings as in matrix factorization models, we concatenate the user and book embeddings to a new input feature to the neural network. This approach helps to learn complex non-linear relationships.