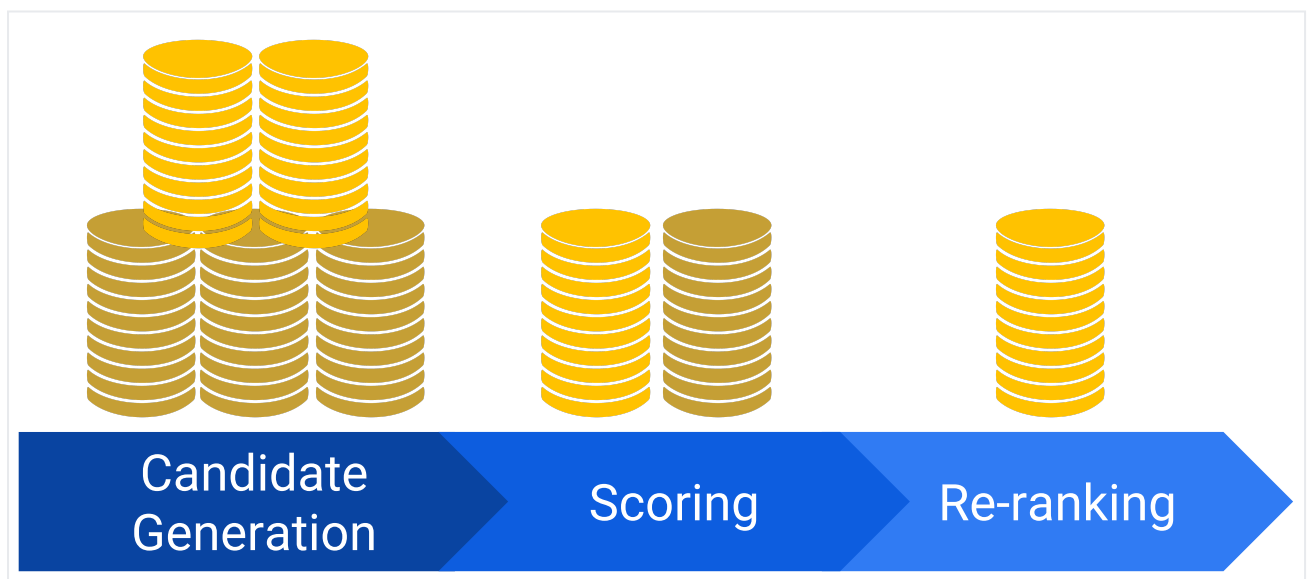


Recommendation Systems Overview

One common architecture for recommendation systems consists of the following components:

- candidate generation
- scoring
- re-ranking



Candidate Generation

In this first stage, the system starts from a potentially huge corpus and generates a much smaller subset of candidates. For example, the candidate generator in YouTube reduces billions of videos down to hundreds or thousands. The model needs to evaluate queries quickly given the enormous size of the corpus. A given model may provide multiple candidate generators, each nominating a different subset of candidates.

Scoring

Next, another model scores and ranks the candidates in order to select the set of items (on the order of 10) to display to the user. Since this model evaluates a relatively small subset of items, the system can use a more precise model relying on additional queries.

Re-ranking

Finally, the system must take into account additional constraints for the final ranking. For example, the system removes items that the user explicitly disliked or boosts the score of fresher content. Re-ranking can also help ensure diversity, freshness, and fairness.

We will discuss each of these stages over the course of the class and give examples from different recommendation systems, such as YouTube.

Resource: For a more comprehensive account of the technology, architecture, and models used in YouTube
[Jton et al., Deep Neural Networks for YouTube Recommendations](https://research.google.com/pubs/pub45530.html).
([://research.google.com/pubs/pub45530.html](https://research.google.com/pubs/pub45530.html))

[Previous](#)

← [Terminology](#) (/machine-learning/recommendation/overview/terminology)

[Next](#)

[Check Your Understanding](#) →

(/machine-learning/recommendation/overview/check-your-understanding)

Except as otherwise noted, the content of this page is licensed under the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) (<https://creativecommons.org/licenses/by/4.0/>), and code samples are licensed under the [Apache 2.0 License](https://www.apache.org/licenses/LICENSE-2.0) (<https://www.apache.org/licenses/LICENSE-2.0>). For details, see the [Google Developers Site Policies](https://developers.google.com/site-policies) (<https://developers.google.com/site-policies>). Java is a registered trademark of Oracle and/or its affiliates.

Last updated 2020-02-10 UTC.