

Online A/B tests!

↳ To tune your recommendation system
using your real customers, & measures how
they react to your recommendations.

Note:- offline Metrics such as Novelty,
~~operatior~~

Diversity, accuracy are used as
Indicators for developing Recommender Systems
offline.

→ But you shouldn't declare victory until
you have ~~succes~~ measured ~~your~~ real impact
on real users from your work.

↳ User behaviour is the ultimate test of
your work.

→ There is a real effect where often
accuracy metrics tell you that an algorithm
is great, only to have it to do horribly in a

Online test.

- ↳ Accurately predicted ratings don't necessarily make for good recommendations.
- ↳ YouTube said 'There is more art than science' in selecting 'Recommendation'
 - ↳ what they mean is that you can't always use accuracy as a surrogate for good recommendations.

Note:- End of the day, online A/B tests are the only evaluation that matters for your recommendation system.

Perceived Quality:-

- ↳ getting explicit feedback from users (e.g. reviews, like, dislike)