



Toloka

Improving Recommender Systems with Human-in-the-Loop

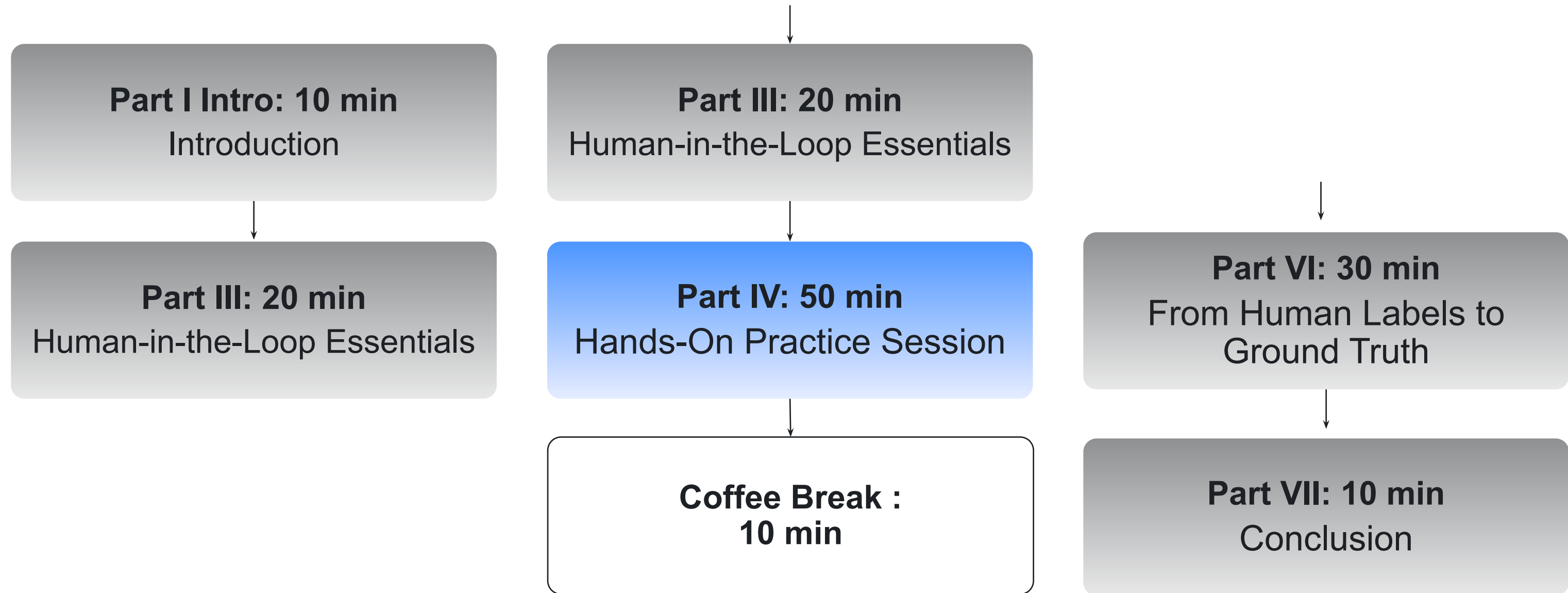
Dmitry Ustalov, Natalia Fedorova, Nikita Pavlichenko,
Maxim Kunakov, Fedor Zhdanov

Part IV

Practice Session

Natalia Fedorova / Maxim Kunakov,
Educational Project Manager / Solutions Engineer

Tutorial Schedule



Practice session: scope

Imagine that you develop a machine learning pipeline to help improve Recommender Systems

- ▶ You have a dataset of products and recommendations
- ▶ You need to get a ranked list of recommendations
- ▶ These collected data will further be used to train a Recommender System

Practice session: scope


Imagine that you develop a machine learning pipeline to help improve Recommender Systems

▶ You have a dataset of products and recommendations

▶ You need to get a ranked list of recommendations

▶ These collected data will further be used to train a Recommender System

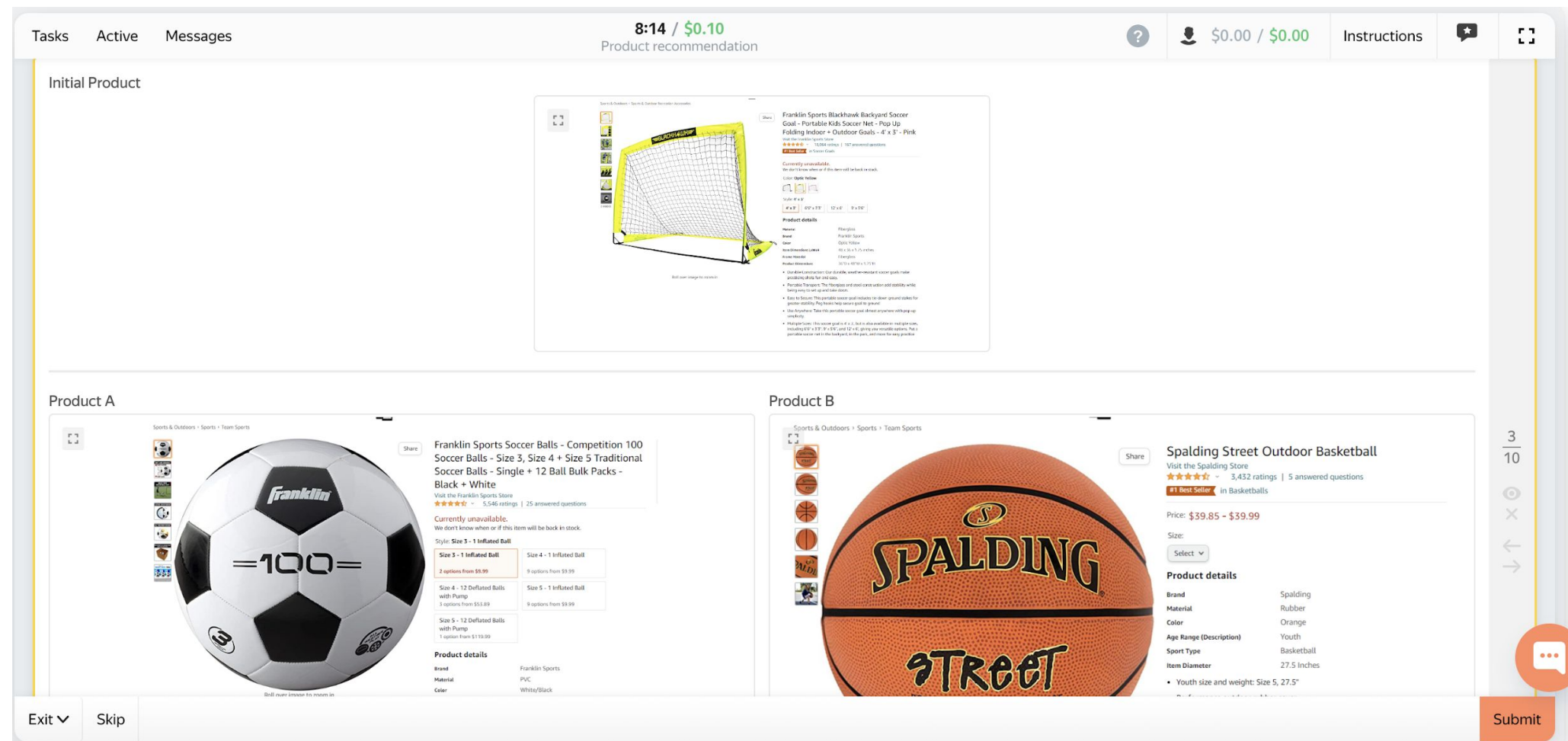
It is your goal during the practice session of our tutorial



What recommendation of a complementary product is better?

Tasks Active Messages 8:14 / \$0.10 Product recommendation ? \$0.00 / \$0.00 Instructions

Initial Product



The screenshot shows a product recommendation interface. At the top, there is a navigation bar with 'Tasks', 'Active', and 'Messages'. Below this, a status bar shows '8:14 / \$0.10 Product recommendation', a question mark icon, a balance of '\$0.00 / \$0.00', and 'Instructions'. The main content area is divided into three sections:

- Initial Product:** A product card for 'Franklin Sports Blackhawk Backyard Soccer Goal - Portable Kick Soccer Net - Pop-Up Folding Indoor + Outdoor Goals - 6 x 3' - Pink'. It includes an image of the goal and a 'Buy now' button.
- Product A:** A product card for 'Franklin Sports Soccer Balls - Competition 100 Soccer Balls - Size 3, Size 4 + Size 5 Traditional Soccer Balls - Single + 12 Ball Bulk Packs - Black + White'. It features an image of a soccer ball and a 'Buy now' button.
- Product B:** A product card for 'Spalding Street Outdoor Basketball'. It features an image of an orange basketball and a 'Buy now' button.

At the bottom of the interface, there are 'Exit' and 'Skip' buttons on the left, and a 'Submit' button on the right. A vertical sidebar on the right side of the interface contains navigation icons and a '3/10' indicator.

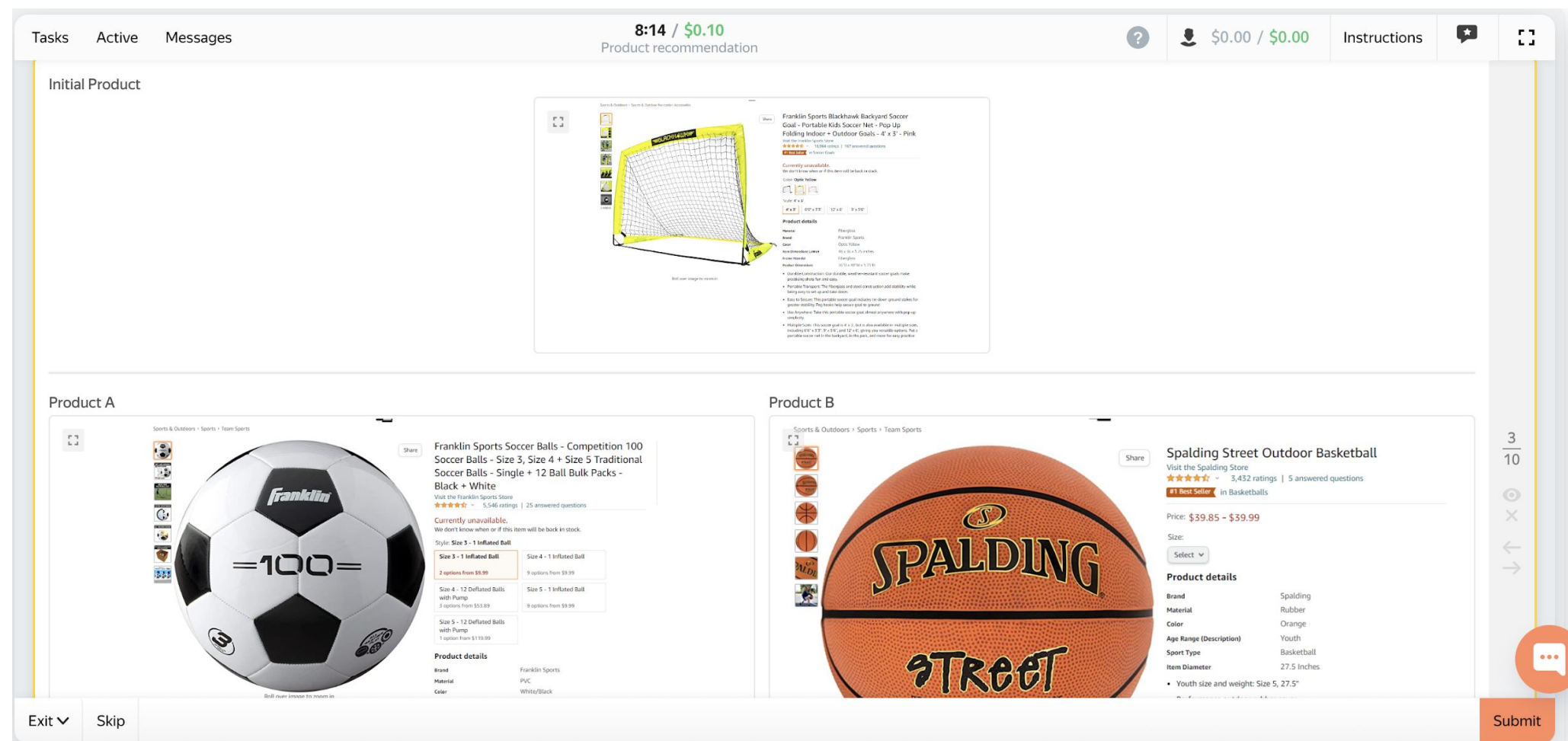
What recommendation of a complementary product is better?

Task

What recommendation of a complementary product is better?

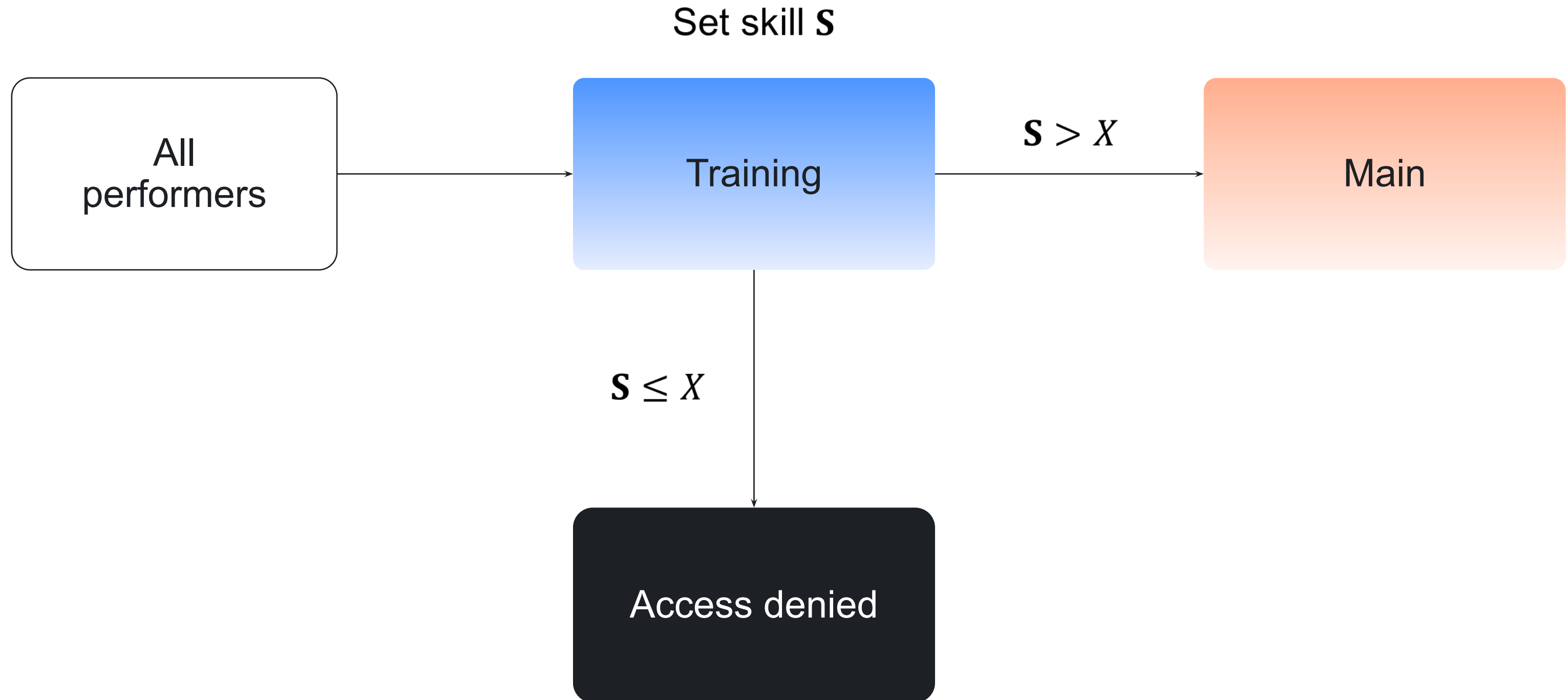
Settings

- ▶ Type: SbS
- ▶ Training: yes
- ▶ Quality control: golden set
- ▶ Pay: \$0.05 per 1 page with 10 tasks
- ▶ Time: 600 sec per 1 page



The screenshot displays a user interface for a product recommendation task. At the top, there is a navigation bar with 'Tasks', 'Active', and 'Messages' tabs. A timer shows '8:14 / \$0.10' and a balance of '\$0.00 / \$0.00'. Below the navigation bar, the 'Initial Product' is a Franklin Sports Blackhawk Backyard Soccer Goal. The main area shows two recommended products: Product A, Franklin Sports Soccer Balls, and Product B, Spalding Street Outdoor Basketball. The interface includes a 'Submit' button at the bottom right and 'Exit' and 'Skip' buttons at the bottom left.

Performer training



We are starting the practice session: 45 min

We share you links to:

- ▶ A step-by-step guide to configure and run data labeling
- ▶ A dataset with products and recommendations
- ▶ Promo code to use Toloka as a requester

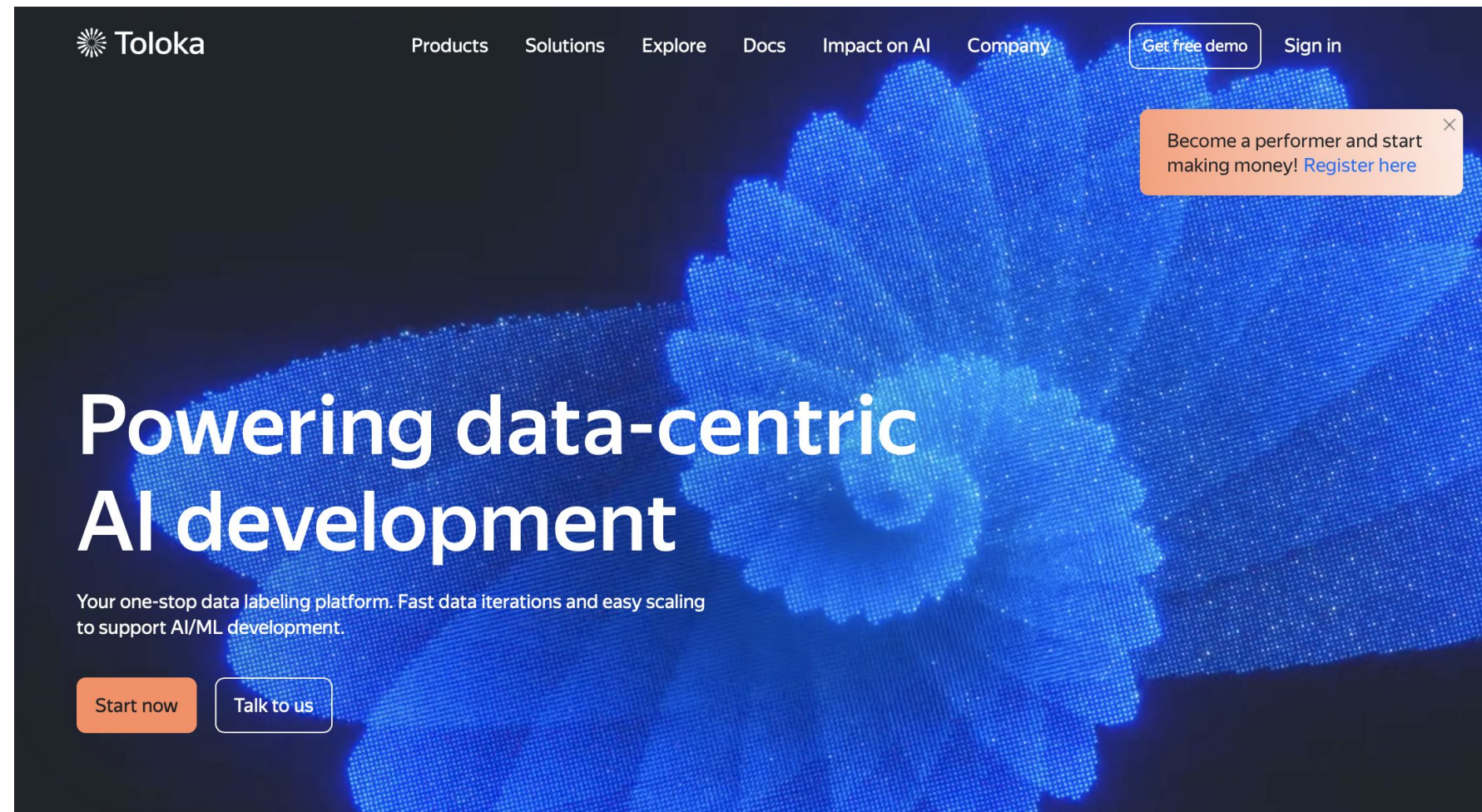
Step-by-step instruction:

bit.ly/3RVHt6S



Sign up on Toloka as a requester

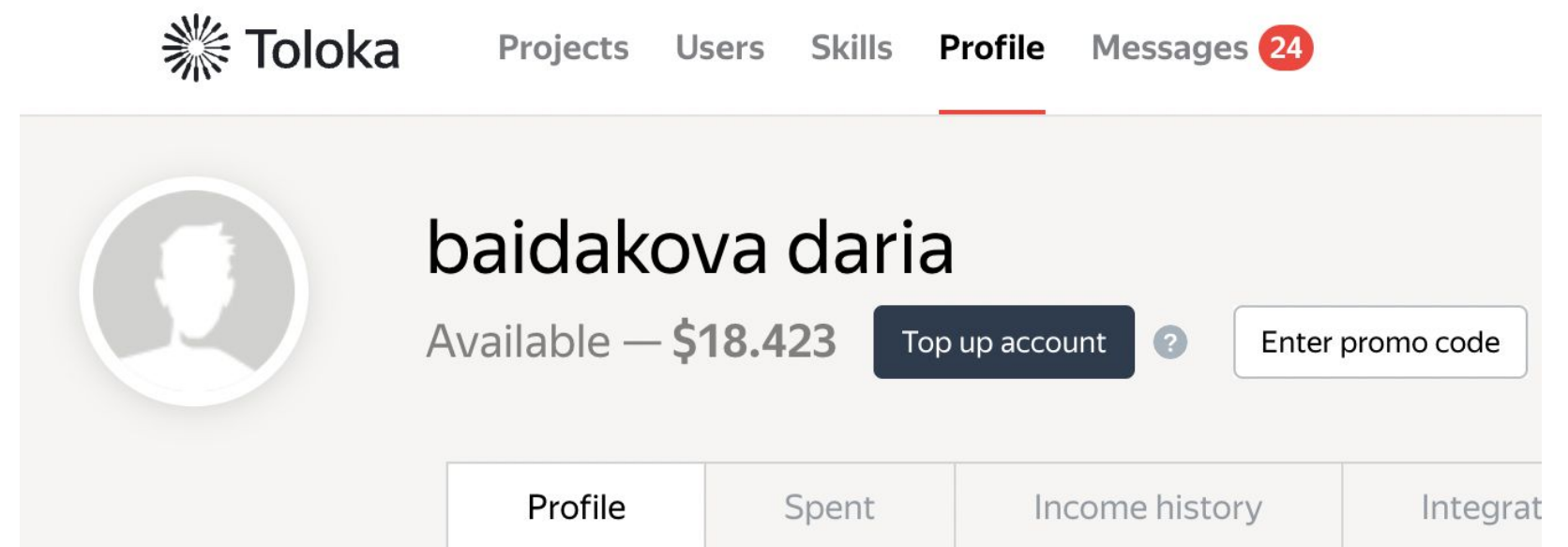
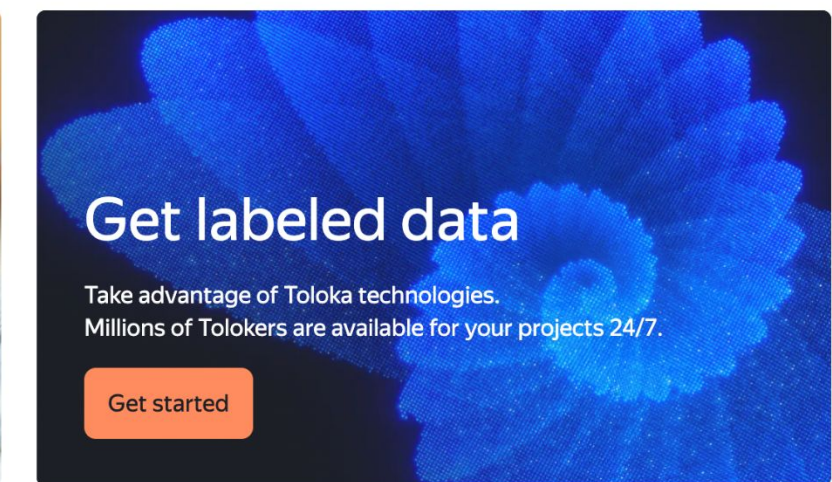
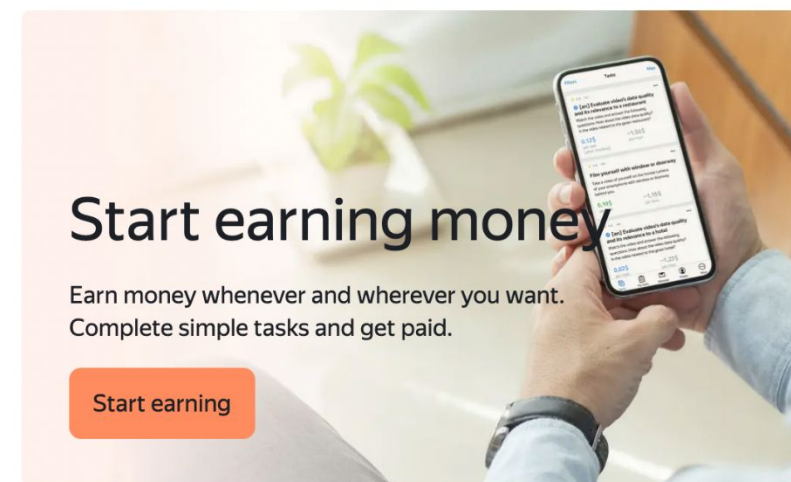
1. Go to toloka.ai
2. Click on “Start now”



Sign up on Toloka as a requester

3. Click on “Get labeled data”
4. Use your email and password
5. Go to Profile and use promocode

I want to...



PROMO CODE:

RecSys2022

Toloka demo

5 min

Join our Slack: **recsys_2022**

Natalia Fedorova

Educational Project Manager



natfedorova@toloka.ai



<https://toloka.ai/events/recsys-2022/>



<https://bit.ly/3eYIX2P>

Useful links

1. Manual bit.ly/3RVHt6S

2. Interface

<https://docs.google.com/document/d/12nauScNVSWpifDBYz30Zn03kLgrUPTndOfAB03ZQb-0/edit?usp=sharing>

3. Input data

<https://docs.google.com/document/d/1Bmlq-IRmDV3NtxOYRt0TPyLbj2cbfhZ-2oPTXVH6U9w/edit?usp=sharing>

4. Instructions

<https://docs.google.com/document/d/16PF7pFpnf5yvdjYT7UDYNx2MT3YPP8GwNMtbG3pqFtg/edit?usp=sharing>

5. Train dataset

<https://drive.google.com/file/d/1RoyvOcXCpQIFrqH9QM06Wd3n9H2gO9lf/view?usp=sharing>

6. Main dataset

<https://drive.google.com/file/d/1752D9fgc6xXBUO-bKy0KaJ0hcLLG984j/view?usp=sharing>