

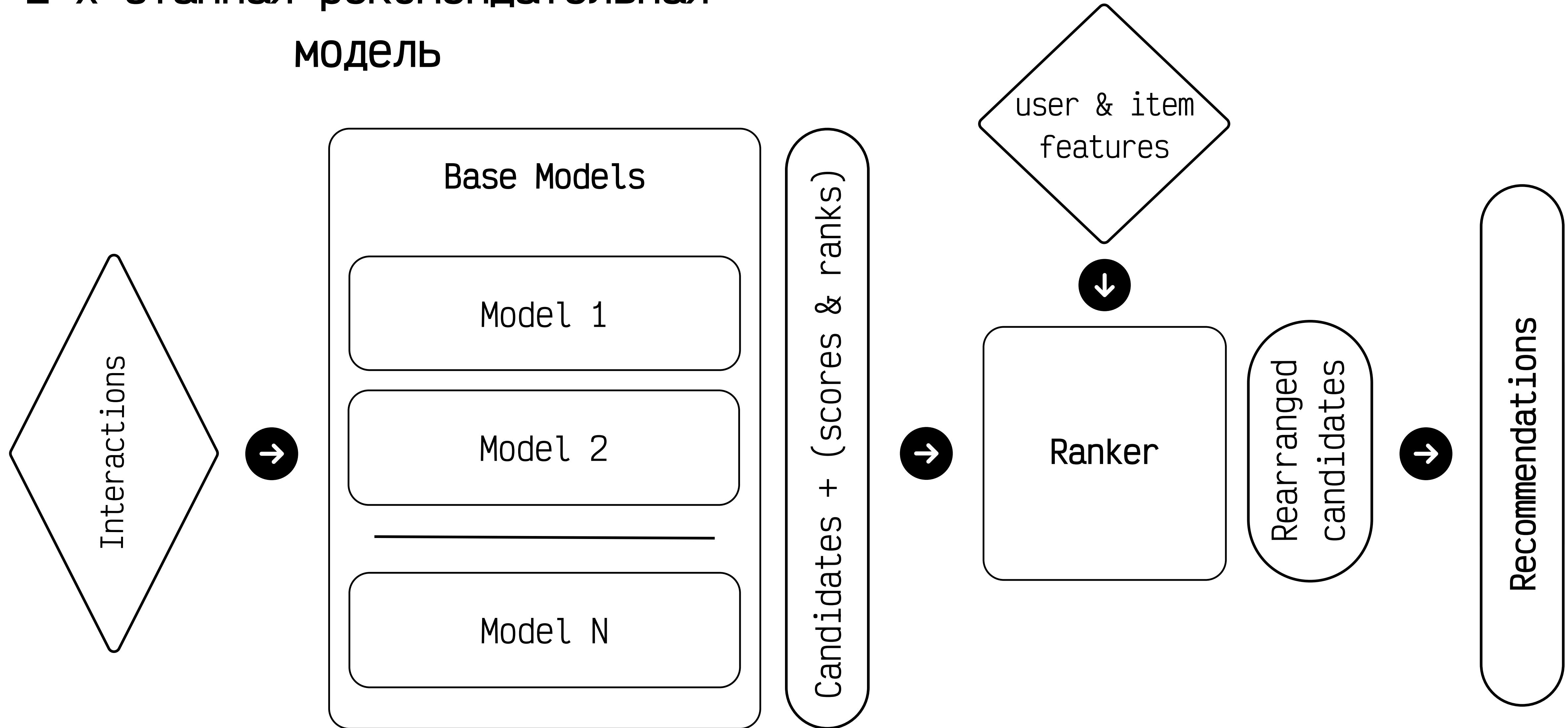
Рекомендательные системы

Модели

Виды моделей	Плюсы	Минусы
kNN	<ul style="list-style-type: none"> • скорость обучения и инференса • просто интерпретировать результат 	<ul style="list-style-type: none"> • необходимо пересчитывать матрицу при изменении количества юзеров или айтемов • холодный старт • ошибки и неточности при малом количестве взаимодействий
linear autoencoders (SLIM, EASE)	<ul style="list-style-type: none"> • аналогично kNN • хорошо работает с разреженными данными • MVP для многих задачи (быстро настроить и хороший результат) 	<ul style="list-style-type: none"> • матрица хранится в <code>dense</code> формате и занимает много памяти ⇒ не применимо на большом объеме данных
Matrix factorization (SVD, ALS)	<ul style="list-style-type: none"> • возможность использовать фитчи юзеров и айтемов • работа с разреженными данными • переиспользование скрытых представлений признаков 	<ul style="list-style-type: none"> • фитчи неявно участвуют в обучении векторов • возможно использовать фитчи только в <code>dense</code> формате (дорого и долго при большом количестве фитчей)
Factorization machines	<ul style="list-style-type: none"> • работа с разреженными данными • рекомендации для cold и warm юзеров и айтемов 	<ul style="list-style-type: none"> • риск переобучения(необходимо грамотно настраивать регуляризацию) • скорость инференса сильно зависит от выбранных гиперпараметров
DL (SASRec, Bert4Rec)	<ul style="list-style-type: none"> • выявление более сложных зависимостей • архитектура помогает учитывать различные данные 	<ul style="list-style-type: none"> • скорость и дороговизна обучения

Стратегия

2-х этапная рекомендательная модель



Модели первого уровня

Здесь модели должны быть максимально простыми и не учитывали фитчи айтемов, т.е. строились, основываясь на матрице взаимодействий. Также хотелось бы, чтобы модели были разными, т.е. основывались на различных методах.

В качестве моделей первого уровня можно рассмотреть:

- kNN (cosine & BM25)
- MF (ALS)
- FM (LightFM)
- EASE \ SLIM
- DL (SASRec \ Bert4Rec)
- Top Popular
- Last Viewed
- Рекомендации на основе изображений последних просмотренных товаров (<https://habr.com/ru/articles/675052/>)

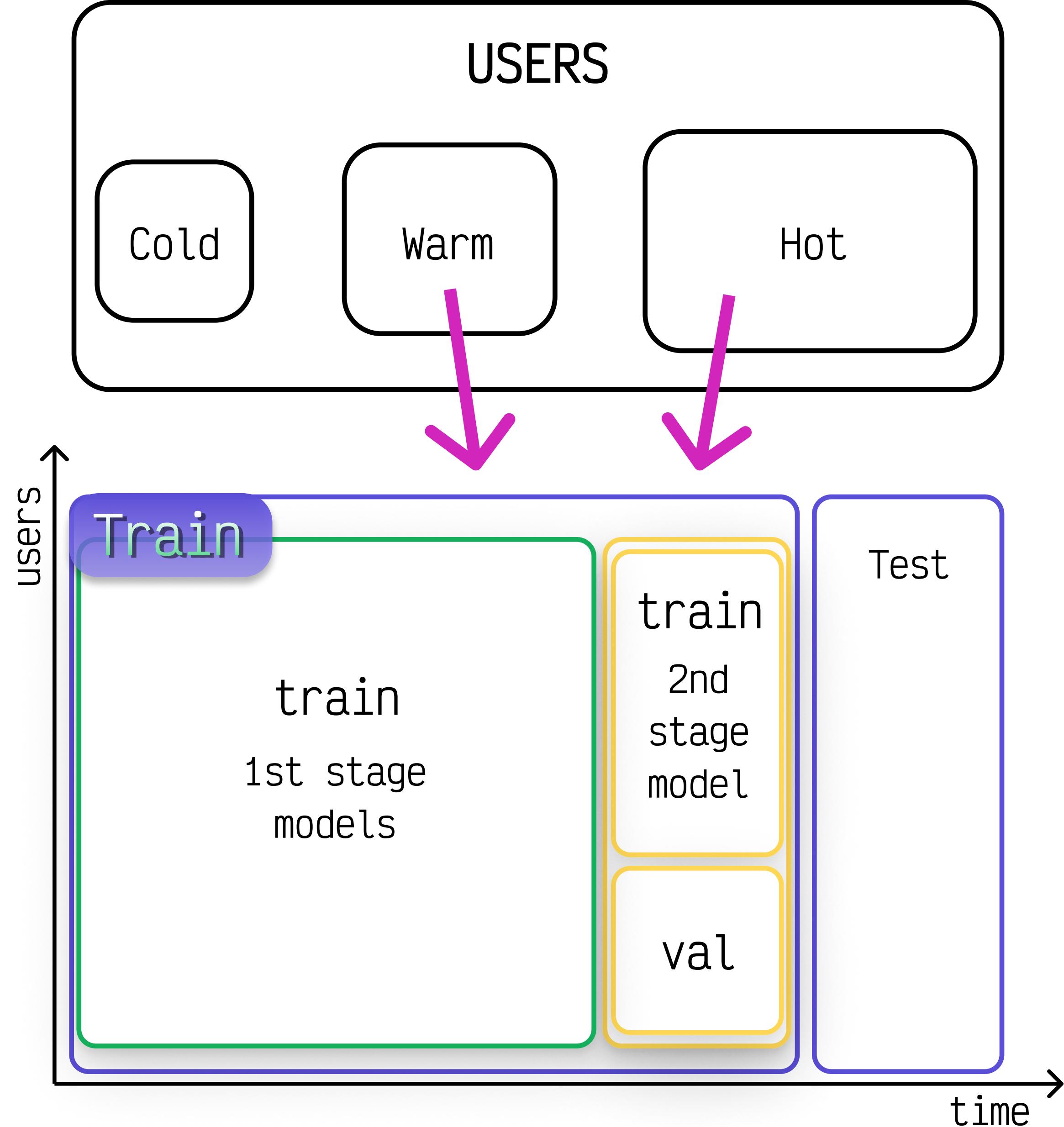
Модели второго уровня

Данная модель должна учитывать скоры посчитанных товаров, а так же фитчи товаров и пользователей и на основе этого выдавать свой ответ. В качестве ранкеров можно рассмотреть:

- LightGMB (LGBMRanker)
- CatBoost (CatBoostRanker)

*Модель второго уровня строиться на данных отличных от тех, что были использованы для построения моделей первого уровня

Валидация



При использовании Time-based валидации, естественно, что появятся Cold-users. Так как у нас отсутствует информация о пользователях, то для холодного старта будет использоваться модель TopPopular.

На оставшихся пользователях и будет обучаться наша модель.

Метрики

Будем использовать целый набор метрик:

- MAP@k
- Recall@k
- Serendipity
- NDCG
- Novelty

Вопросы