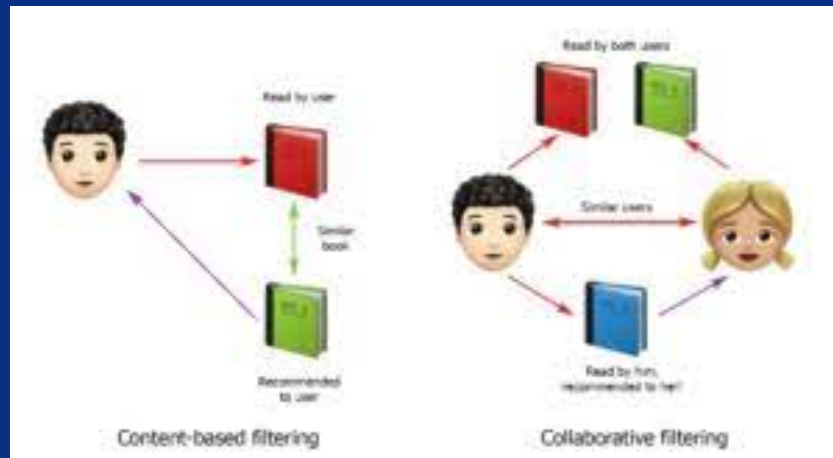




Рекомендательные системы для Рискмаркета

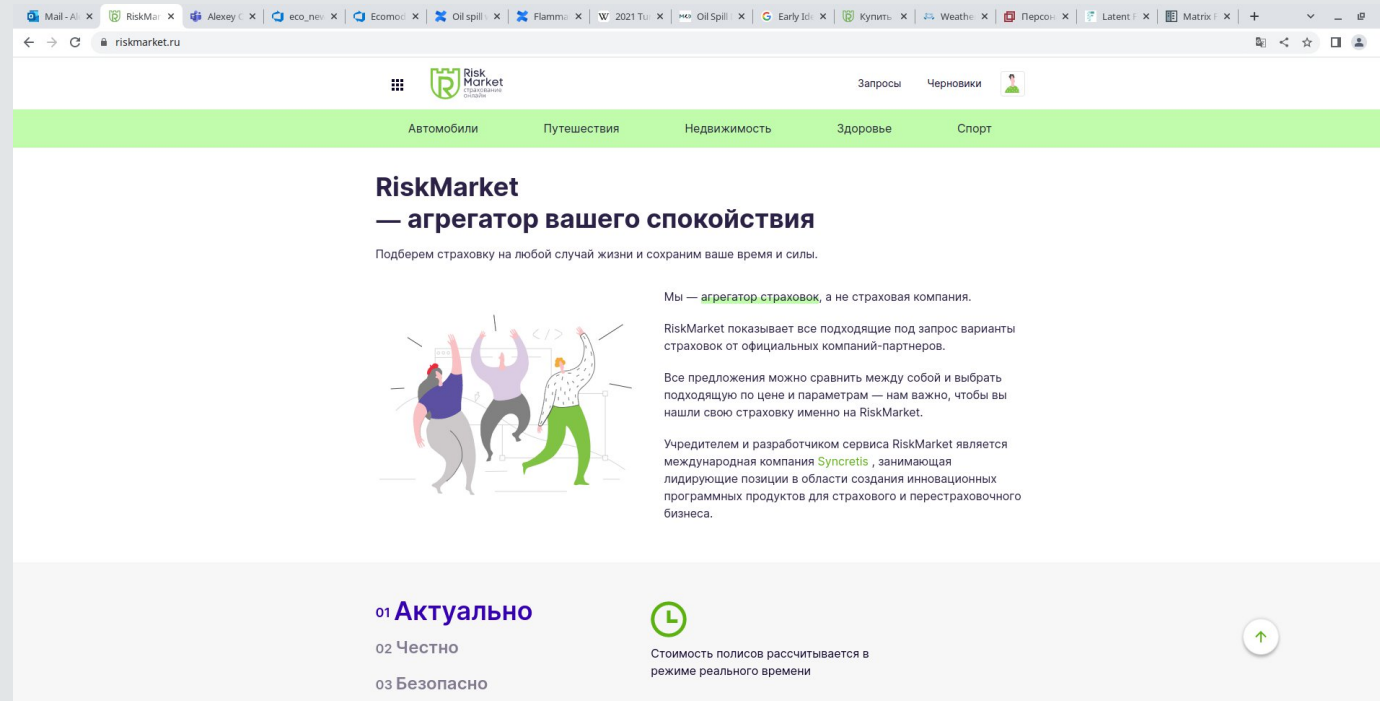


Контекст

Рекомендательные системы — это современная технология, способная повысить привлекательность любого сайта, занимающегося продажей.

Рискмаркет — это агрегатор страховок, показывающий предложения компаний-партнёров.

Целью данной презентации является обсуждение применимости рекомендательных систем в Рискмаркете.



Плюсы рекомендательных систем.

- Пользователю проще сделать выбор, что приводит к повышению продаж.
- Повышение заинтересованности пользователей.
- **Догадливость**, способность угадывать неожиданные предпочтения пользователей.
- Есть возможность как применять достаточно хорошие решения из коробки, так и улучшать их различными способами (учёт особенностей объектов и пользователей, улучшение методов).
- **Cross-selling**. С этим товаром также покупают.
- **Up-selling**. Дополнительные новые товары.



Почему именно Рискмаркет?

- По **CRISP-DM** мы должны начинать с понимания бизнеса и понимания данных. Поэтому выбор данных предшествует выбору модели.
- Большинство действующих рекомендательных систем привязаны к выбору данных.
- Универсальное решение не имеет особого смысла. Мы вряд ли сможем быстро его реализовать. Мы даже не знаем конкурентов. Кроме того, оно не будет привязано к Синкретису.
- Рискмаркет — это продукт в проде.
- Можно было бы использовать рекомендательные системы в рекомендациях резюме вакансиям или в рекомендациях контента на портале.



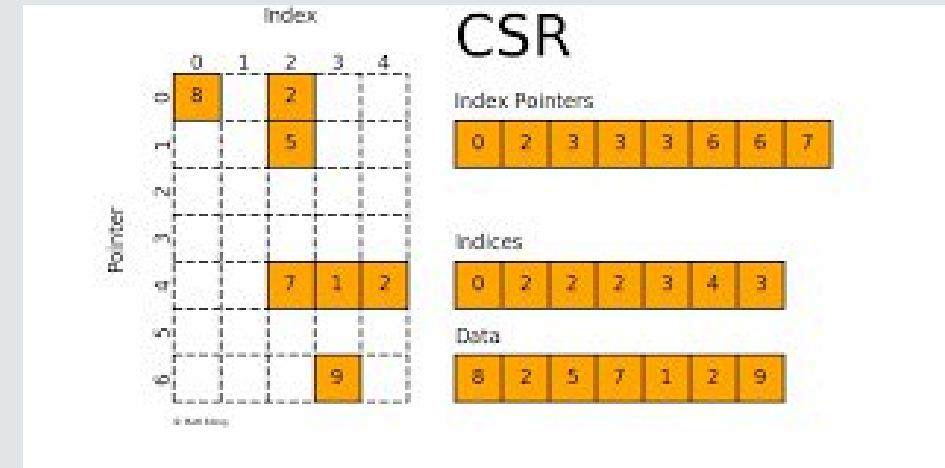
Терминология.

- **Пользователи.** Известна история их действий, другая информация минимальна.
- **Объекты.** Страховые полисы различных типов различных компаний. Есть дополнительные особенности полисов, такие как цена.
- **Задача.** Рекомендовать пользователям новые страховые полиса.
- **Явные предпочтения.** Покупка полиса данного типа данной компании.
- **Неявные предпочтения.** Просмотр условий полиса данного типа данной компании.



Структура данных.

- Разреженная матрица предпочтений с нулями и единицами.
- У каждого объекта есть информация об атрибутах.
- Хранение матрицы: **csr**-формат.
- Подобная матрица есть как для явных, так и для неявных предпочтений.
- Матрицы не противоречат друг другу.



Метрика.

RMSE, так как для простоты убрано ранжирование, т.е. порядок рекомендаций не важен.

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(\hat{y}_i - y_i)^2}{n}}$$

Основная модель.

- Неявные и явные предпочтения учитываются в минимизации:

$$\sum_{(u,i) \in D} c_{ui} \left(s_{ui} - \bar{s}_u - \bar{s}_i - \sum_{t \in T} p_{ut} q_{it} \right)^2 + \lambda \sum_{u \in U} \|p_u\|^2 + \mu \sum_{i \in I} \|q_i\|^2 \rightarrow \min_{P, Q}$$

- Явные предпочтения:

$$c_{ui} = 1 + \alpha r_{ui}:$$

- Начальные матрицы **p** и **q** берутся из SVD-разложения матрицы неявных предпочтений.

Вспомогательная модель.

Оценка рейтинга по схожим клиентам (User-Based CF):

$$\hat{r}_{ui} = \bar{r}_u + \frac{\sum_{v \in U(u)} S(u, v)(r_{vi} - \bar{r}_v)}{\sum_{v \in U(u)} S(u, v)}$$

- Близость понимается в смысле комбинации коэффициента Жаккара и кластеризации по домену.
- Модель делается для матрицы неявных предпочтений.

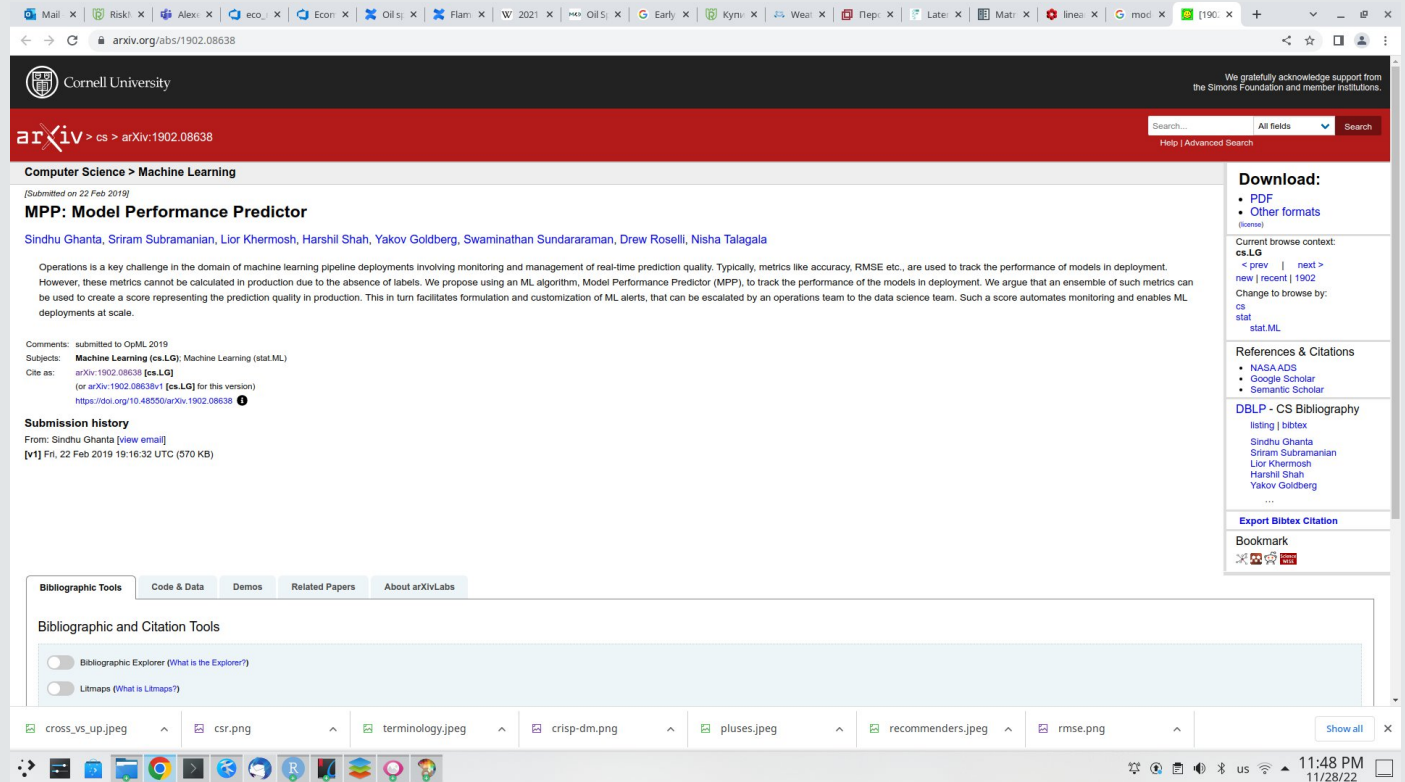
Комбинирование моделей.

- Рекомендации делаются по основной модели.
- Далее они уточняются по вспомогательной модели для схожих клиентов.
- Если ничего неизвестно, берутся самые частые объекты.
- **Cross-selling** и **up-selling** делаются по вспомогательной модели (только **item-based**, а не **user-based**, с грамотно подобранной метрикой близости).
- **Проблема холодного старта** (для пользователя или для объекта) решаются с помощью основной модели.



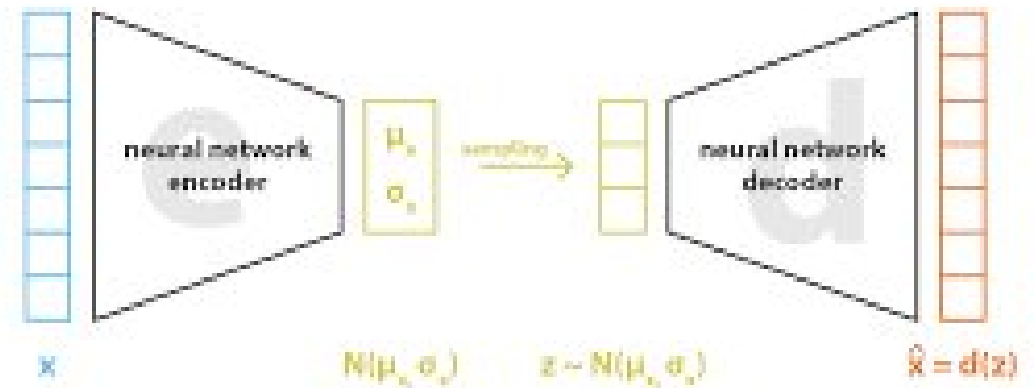
Мониторинг.

- Как понять что модель не работает?
- Сравнивать качество модели каждую неделю вручную.
- Либо сравнивать каждую неделю с помощью MPP (model performance predictor).
- Как интерпретировать результаты?
- Например, за счёт возмущения входных данных и переобучения.



Дальнейшие улучшения.

- Учёт атрибутов полисов, графовые модели.
- Использование нейронных сетей (encoder-decoder).
- Использование продвинутых архитектур (**Bert2Rec**)
- Добавление ранжирования.



$$\text{loss} = ||x - \hat{x}||^2 + \text{KL}[N(\mu_x, \sigma_x), N(0, I)] = ||x - d(z)||^2 + \text{KL}[N(\mu_x, \sigma_x), N(0, I)]$$