

Рекомендательные системы

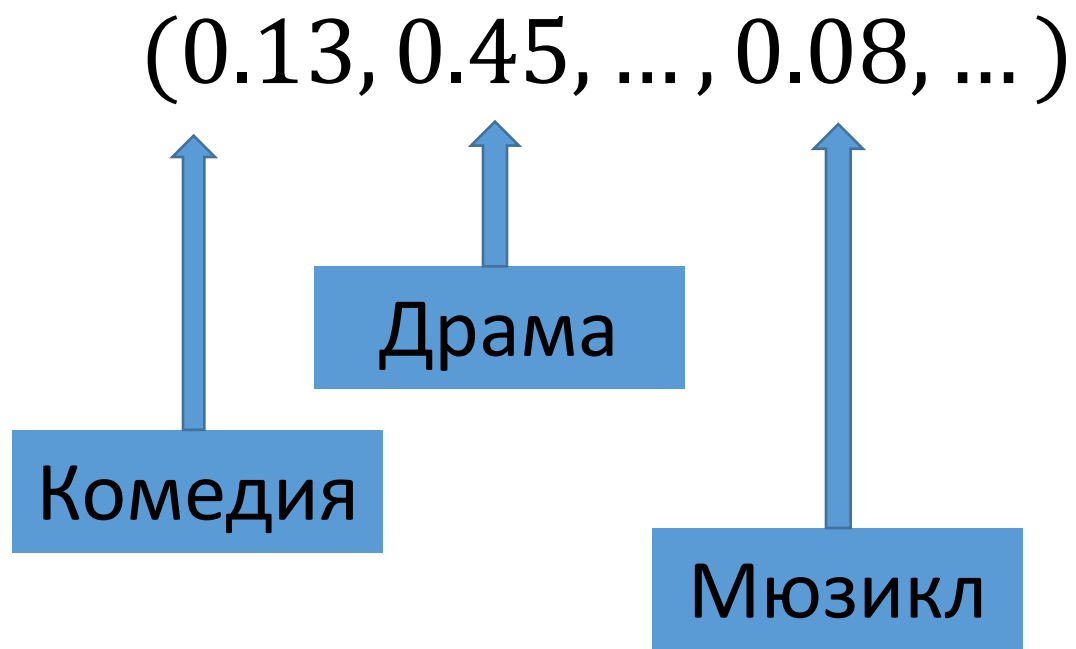
По материалам Евгения Соколова

Занятие 2

Модели со скрытыми
переменными

Векторы интересов

- Для пользователя — насколько он интересуется каждым жанром
- Для фильма — насколько он относится к каждому жанру



Рейтинг

- Предположение: заинтересованность определяется как скалярное произведение векторов пользователя и фильма

$$(0.1, 0.5, 0.01, 0.92) \times (0, 0, 0.1, 0.95) = 0.875$$

$$(0.1, 0.5, 0.01, 0.92) \times (0.9, 0, 0, 0.1) = 0.182$$

Пользователь

Фильм

Модели со скрытыми переменными

- Обучим вектор p_u для каждого пользователя u
- Обучим вектор q_i для каждого товара i
- Оценка приближается их скалярным произведением:

$$r_{ui} \approx \langle p_u, q_i \rangle$$

- Находим векторы только по известным оценкам
- После этого можем предсказать оценку для любой пары «пользователь-товар»

Модели со скрытыми переменными

- Оптимизационная задача:

$$\sum_{(u,i) \in R} (r_{ui} - \bar{r}_u - \bar{r}_i - \langle p_u, q_i \rangle)^2 \rightarrow \min_{P, Q}$$

- Решение: градиентный спуск, Alternating Least Squares (ALS) и другие методы

Модели со скрытыми переменными

2	5	
5		4
	1	
	2	5

Модели со скрытыми переменными

	(0.9, 0.05)	(0.02, 1.1)	(1.05, 0.01)
(2.1, 5)	2	5	
(4.6, 0)	5		4
(0, 1)		1	
(4.9, 0.9)		1	5

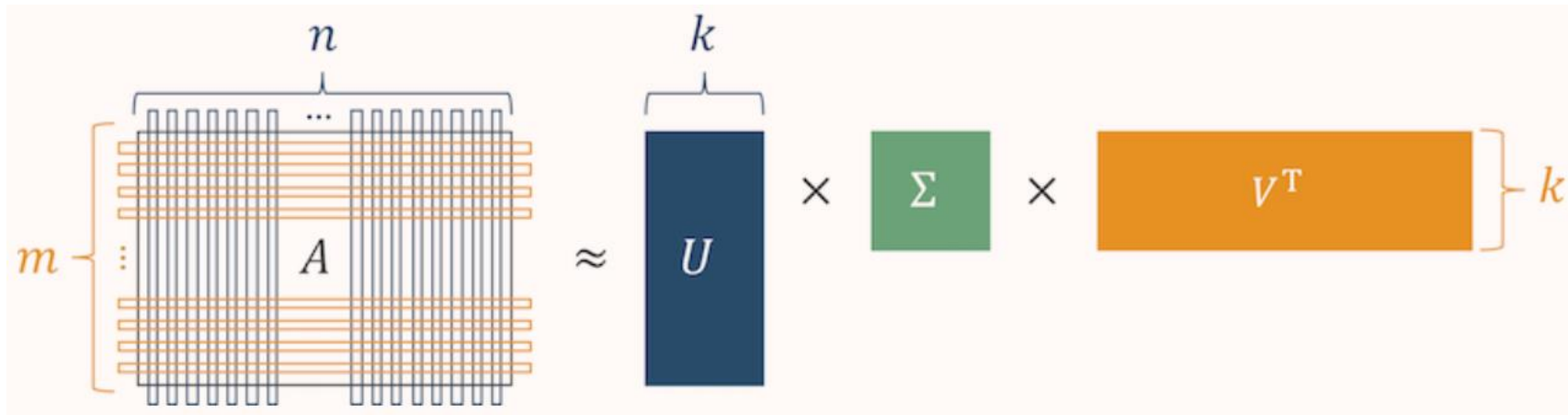
Singular Value Decomposition (SVD)

Известно, что любая матрица размера $n \times k$ (ранга k) представима в виде $X = VDU^T$, где

- 1) V – ортогональная матрица размера $n \times n$, ее столбцы – собственные векторы матрицы XX^T ;
- 2) D – матрица размера $n \times k$, $d_{ii} = \sqrt{\lambda_i}$, $d_{ij} = 0$, если $i \neq j$, где $\{\lambda_i\}_{i=1}^k$ – собственные числа матрицы $X^T X$ (и ненулевые собственные значения матрицы XX^T);
- 3) U – ортогональная матрица размера $k \times k$, её столбцы – собственные векторы матрицы $X^T X$.

SVD для построения рекомендаций

- Матрица товарных предпочтений (матрица, где строки это пользователи, а столбцы это продукты, с которыми пользователи взаимодействовали) представляется произведением трех матриц:

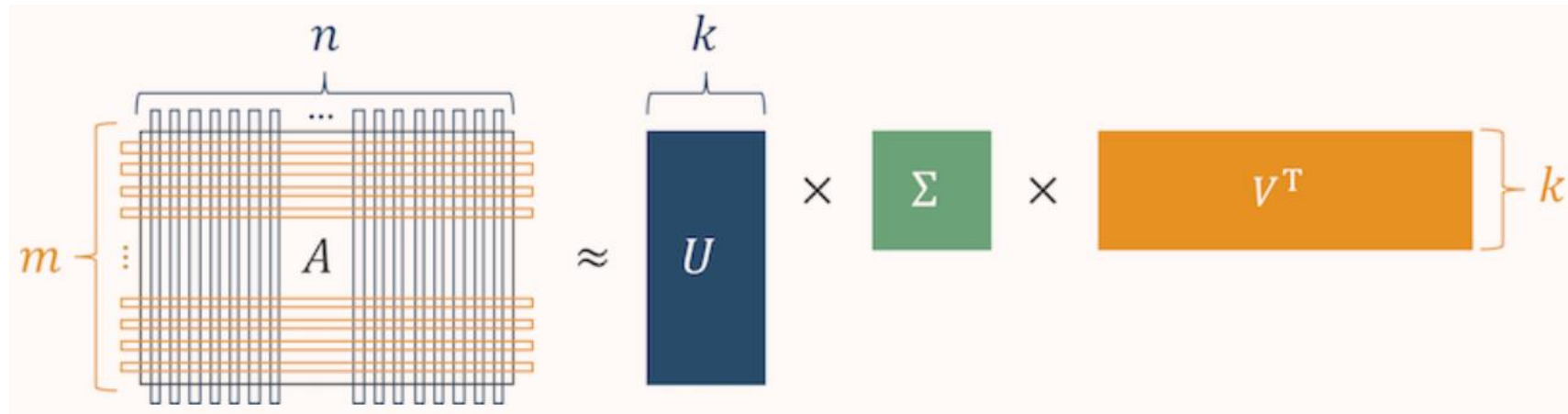


U – описание характеристик пользователя

V – описание характеристик продукта

SVD для построения рекомендаций

- Матрица товарных предпочтений (матрица, где строки это пользователи, а столбцы это продукты, с которыми пользователи взаимодействовали) представляется произведением трех матриц:



- После восстановления исходной матрицы, клетки, где у пользователя были нули, а появились «большие» числа, показывают степень латентного интереса к товару. Упорядочим эти цифры, и получим список товаров, релевантных для пользователя.

SVD для построения рекомендаций

- Матрица товарных предпочтений (матрица, где строки это пользователи, а столбцы это продукты, с которыми пользователи взаимодействовали) представляется произведением трех матриц:

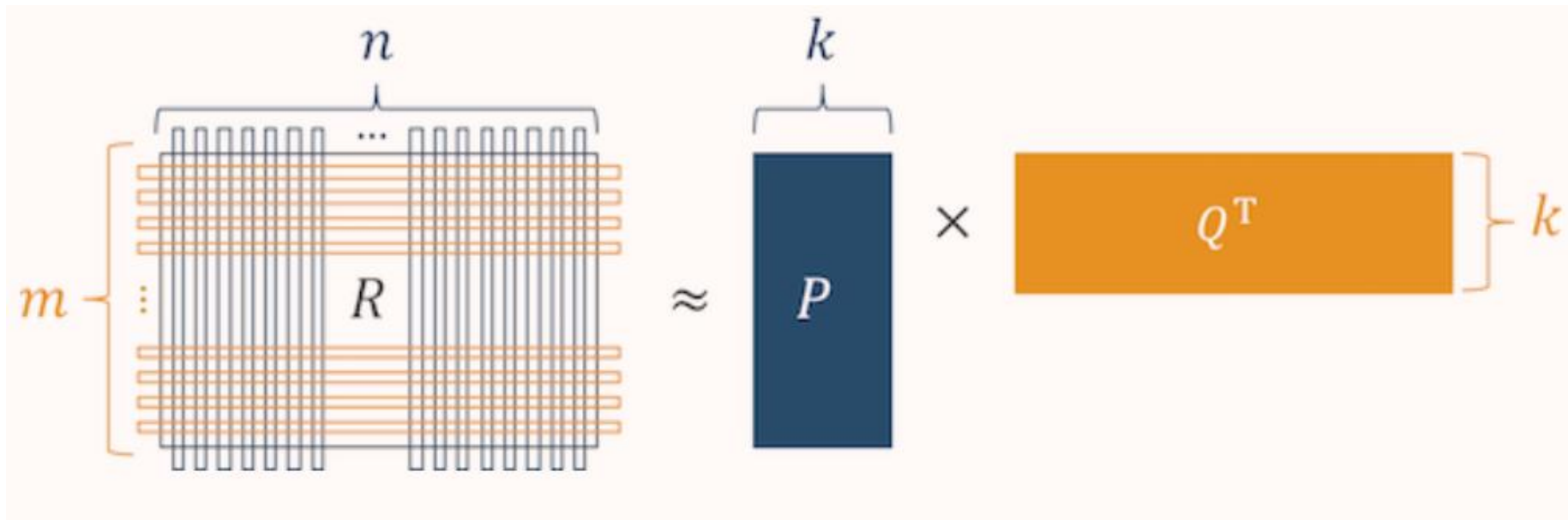
$$A \approx U \Sigma V^T$$

- После восстановления исходной матрицы, клетки, где у пользователя были нули, а появились «большие» числа, показывают степень латентного интереса к товару. Упорядочим эти цифры, и получим список товаров, релевантных для пользователя.
- При этой операции у пользователя и товара появляются «латентные» признаки. Это признаки, показывающие «скрытое» состояние пользователя и товара.

Общее семейство подобных алгоритмов называется ***NMF (non-negative matrix factorization)***. Как правило вычисление таких разложений весьма трудоемко, поэтому на практике часто прибегают к их приближенным итеративным вариантам.

Матричная факторизация (факторизационные машины) с помощью ALS

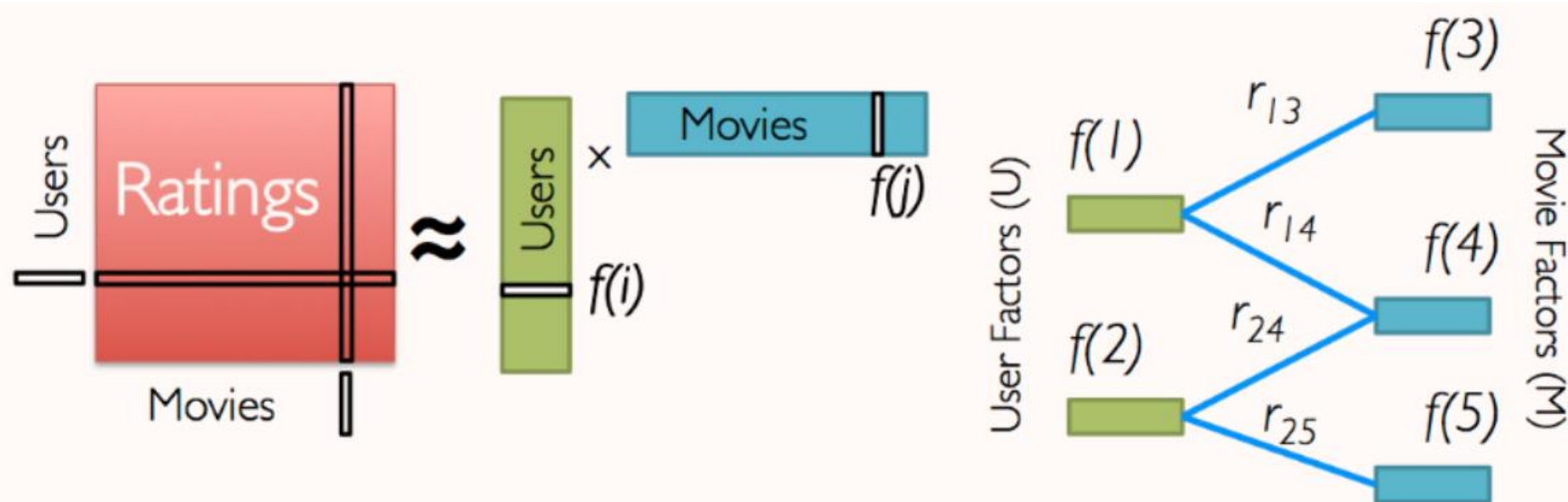
- **ALS (alternating least squares):** популярный *итеративный* алгоритм разложения матрицы предпочтений на произведение 2 матриц: факторов пользователей (U) и факторов товаров (Q).



Матричная факторизация с помощью ALS

- **ALS (alternating least squares):** популярный *итеративный алгоритм разложения матрицы предпочтений на произведение 2 матриц: факторов пользователей (U) и факторов товаров (I).*
- **Принцип работы:** *минимизация среднеквадратичной ошибки на предоставленных рейтингах.*
- **Оптимизация** происходит поочередно, сначала по факторам пользователей, потом по факторам товаров.
- **Для обхода переобучения** к среднеквадратичной ошибке добавляются регуляризационные коэффициенты.

Матричная факторизация



Iterate:

$$f[i] = \arg \min_{w \in \mathbb{R}^d} \sum_{j \in \text{Nbrs}(i)} (r_{ij} - w^T f[j])^2 + \lambda ||w||_2^2$$

Taken from the BerkeleyX Course Big Data Analysis with Spark

<https://habr.com/ru/company/lanit/blog/421401/>

Матричная факторизация

- Мы можем дополнить матрицу предпочтений новым измерением, содержащим информацию о пользователе или товаре. Таким образом, мы задействуем больше доступной информации и возможно получим более точную модель.

На практике именно факторизационные машины в большинстве кейсов дают наилучший результат!

Контентные методы

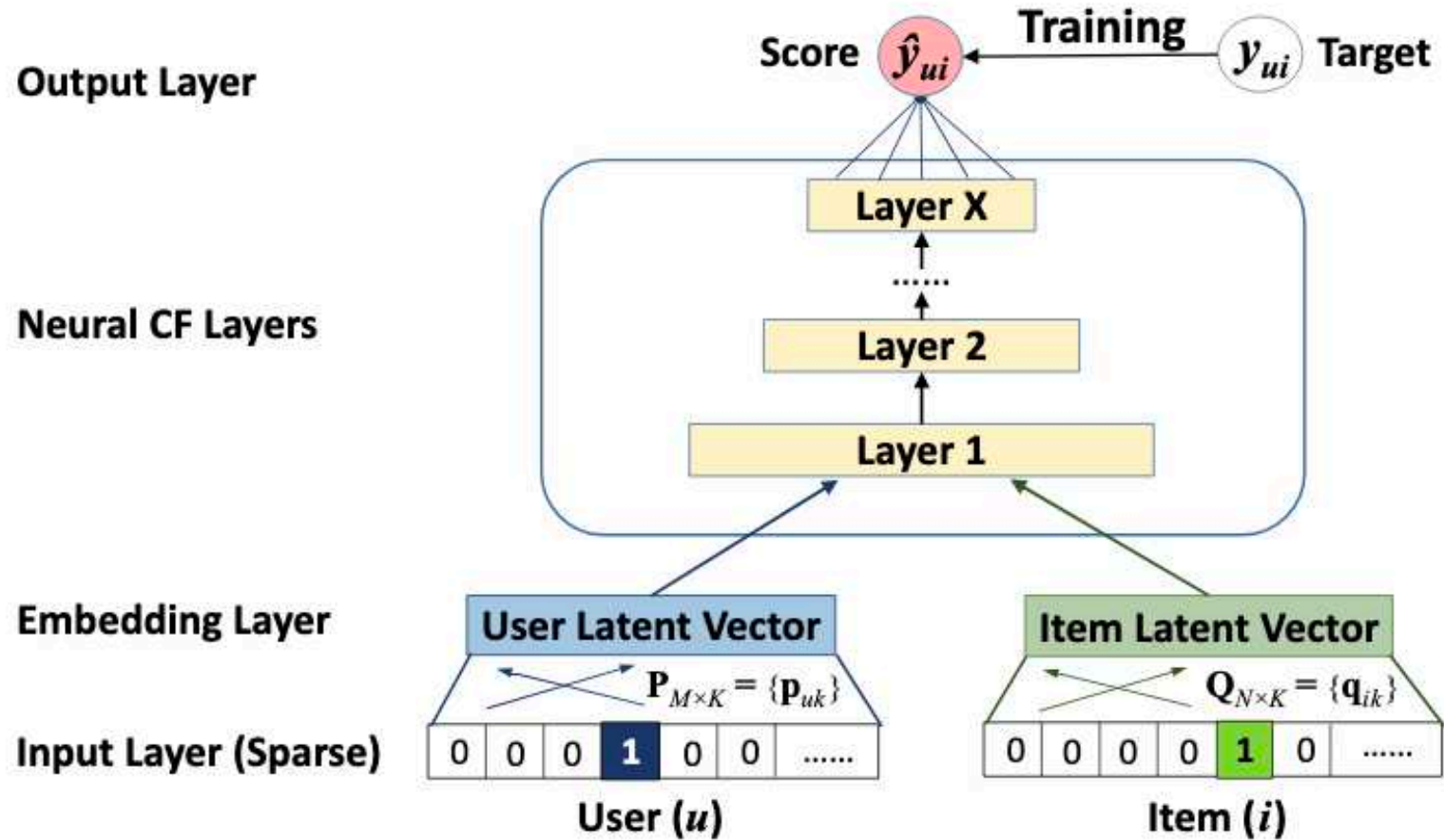
Контентные рекомендации

- Сведём задачу к обычному обучению с учителем
- Объект: пара «пользователь-товар» (u, i)
- Ответ: отклик пользователя
- Факторы: информация про пользователя и про товар
- Обучаем любую модель на этих данных
- Среди факторов могут быть и прогнозы коллаборативных моделей

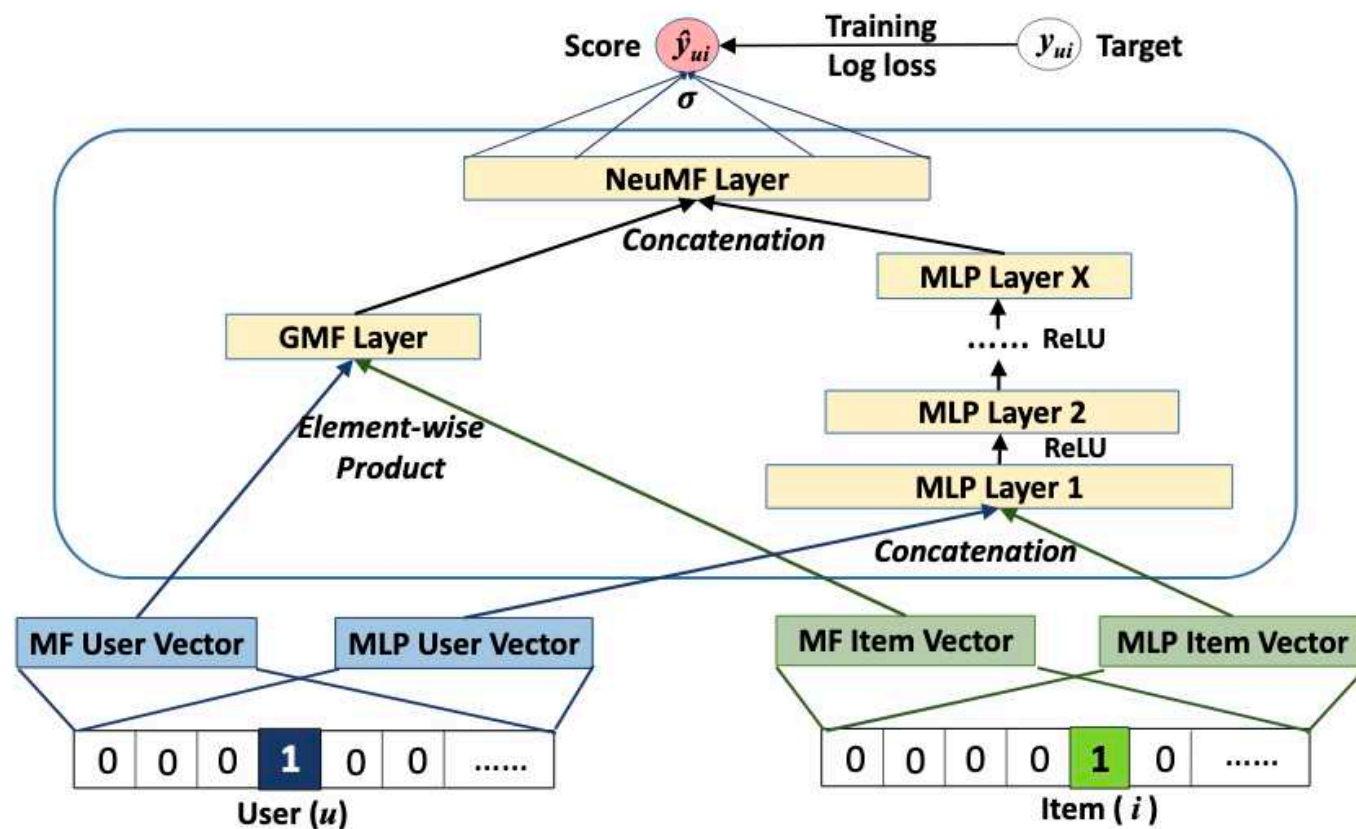
Нейросетевые методы

Neural Collaborative Filtering

Model #1



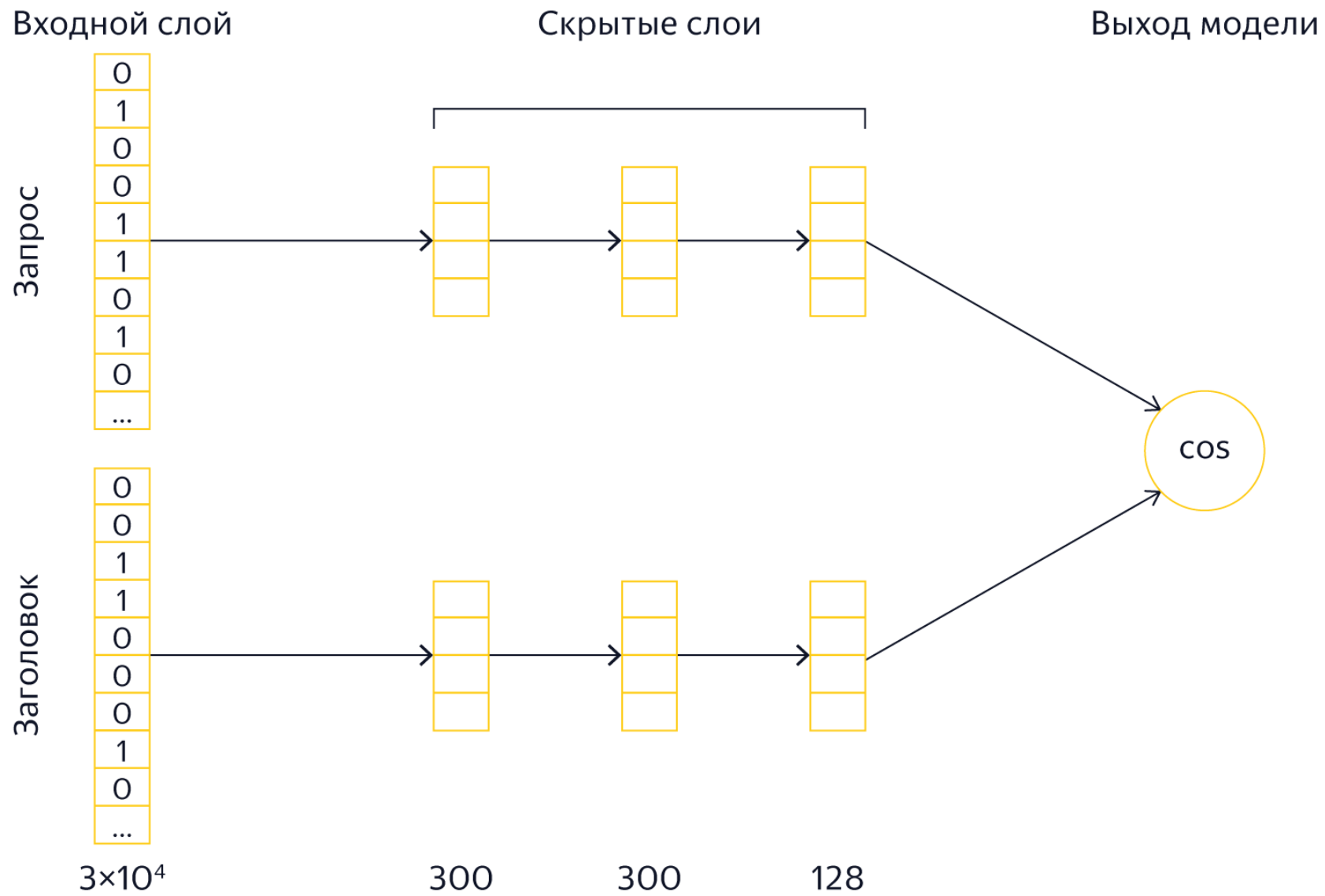
Model #2

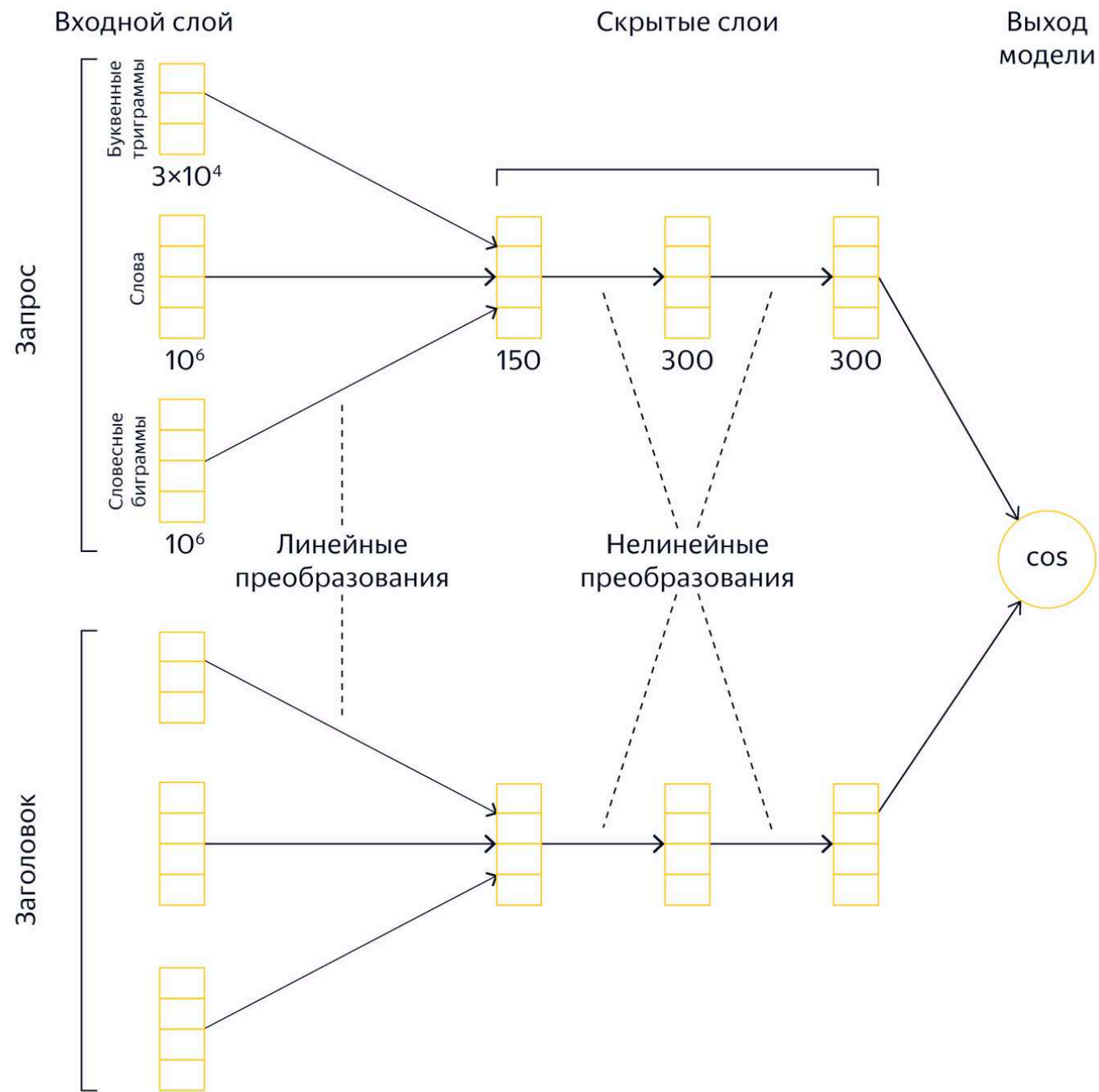


Deep Structured Semantic Model

Word-hashing

- К тексту добавляются маркеры начала и конца
- После чего он разбивается на буквенные триграммы
- Пример: [палех] → [па, але, лех, ех]





Метрики качества рекомендаций

Качество предсказаний

В зависимости от целевой переменной:

- MSE, MAE, R^2
- Accuracy, HitRate, precision/recall, AUC-ROC
- Метрики качества ранжирования (дальше в курсе)

Качество предсказаний

- Насколько хорошо мы предсказываем оценки r_{ui} ?
- Разделяем сессии пользователей на две части: обучаемся на первой, измеряем качество предсказания второй
- Оцениваем, насколько хорошо предсказываем поведение пользователя — но не факт, что нужно именно это
- Зачем рекомендовать то, что он и так купил бы?

Другие метрики

- Покрытие
 - Какая доля товаров рекомендовалась хотя бы раз?
 - Какой доле пользователей хотя бы раз показаны рекомендации?
- Новизна
 - Как много рекомендованных товаров пользователь встречал раньше?
- Прозорливость (serendipity)
 - Способность предлагать товары, которые отличаются от купленных ранее
- Разнообразие

Резюме

- Рекомендации — широкая задача с большим количеством коммерческих применений
- Модели: коллаборативная фильтрация, контентный подход
- Рекомендации товаров на основе сходства пользователей
- Модели со скрытыми переменными
- Активное развитие нейронных сетей для рекомендаций
- Обилие метрик качества

Почитать

- <https://www.benfrederickson.com/approximate-nearest-neighbours-for-recommender-systems/>
- <https://habr.com/ru/company/yandex/blog/314222/>
- <https://www.jefkine.com/recsys/2017/03/27/factorization-machines/>
- <https://www.csie.ntu.edu.tw/~b97053/paper/Rendle2010FM.pdf>
- <https://arxiv.org/abs/1708.05031>