

УДК 004.4

В. Яцишин к.т.н., доц., А. Давидов, Д. Подолян

Тернопільський національний технічний університет імені Івана Пулюя, Україна

МЕТОДИ КОЛАБОРАТИВНОЇ ФІЛЬТРАЦІЇ ПРИ ПОБУДОВІ СИСТЕМ КРОС-ПРОДАЖ

V. Yatsyshyn, Ph.D., Assoc. Prof., A. Davydov, D. Podolyan

METHODS OF COLLABORATIVE FILTERING IN CROSS-SELLING SYSTEMS BUILDING

Сучасні підходи до ведення бізнесу вимагають залучення інформаційних технологій, зокрема web-технологій, які покликані стимулювати та розширювати сферу послуг і товарів, і як результат підвищувати ефективність діяльності організацій.

Одним з підходів підвищення ефективності діяльності підприємств є впровадження “smart” програмного забезпечення, в основі якого лежать методи і засоби штучного інтелекту. Актуальним є впровадження такого програмного забезпечення в електронних магазинах, платформах з реалізації товарів і послуг та ін.

Одним з підходів щодо стимулювання продаж, який лежить в основі cross-selling, є побудова рекомендацій на основі колаборативної фільтрації. Підходи до побудови рекомендаційних систем та місце колаборативної фільтрації наведено на рис. 1.

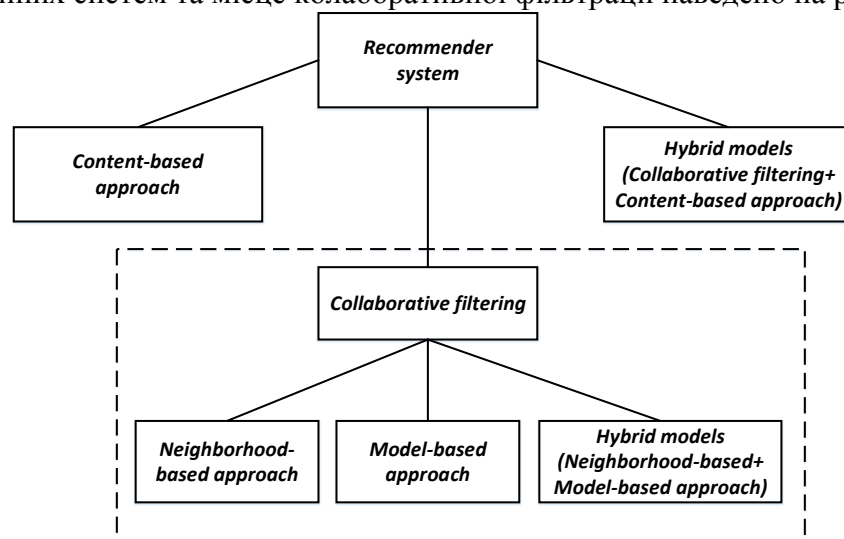


Рис. 1 – Місце колаборативної фільтрації при побудові рекомендаційних систем

Проаналізуємо принципи побудови рекомендаційних систем і метрики подібності товарів на основі підходу найближчих сусідів (neighborhood based approach) щодо стимулювання продаж (cross-selling).

Суть задач cross-selling (рис. 2), зводиться до рекомендації супутніх товарів і послуг для користувачів, які вже здійснювали певні покупки чи замовляли певні послуги. У результаті пропозиції товарів чи послуг зростають продажі товарів, а отже і прибуток організації. Основною проблемою при рекомендації cross-товарів є визначення подібних товарів або подібних покупців.



Рис. 2 – Задачі Cross-selling

Формалізація задачі колаборативної фільтрації для cross-selling можна представити, як наведено на рис. 3 та описано нижче.

	+	+		+		
	+		+		+	
	+		+	+	+	
		+				+
		+				+
			+	+		+

Рис. 3 – Матриця колаборативної фільтрації для задач cross-selling

Рядками матриці є користувачі, які придбали певні товари, а стовпцями – самі товари. Матриця може бути заповнена кількістю товарів, які придбав кожен покупець або позначкою, про придбання товарів. В такому випадку введемо позначення:

U – покупці електронного магазину;

I – товари (продукти харчування, книги, відео, музика та ін.);

$r_{ui} = [\text{покупець } u \text{ придбав товар } i]$ – матриця колаборативної фільтрації.

Задачі, які можна вирішувати на основі колаборативної фільтрації для електронних магазинів:

- оцінювання товару i для користувача u ;
- визначення списку рекомендованих товарів для користувача u ;
- рекомендації супутніх товарів (cross-selling);
- інформування користувачів про нові товари (up-selling);
- сегментація користувачів.

Для визначення подібності товарів чи користувачів можна використовувати метрики, зокрема: метрика Jaccard, метрики косинуса кута, метрика центрованого косинуса кута.

Метрика Jaccard формально представляється, як показано нижче

$$\text{sim}(A, B) = |r_A \cap r_B| / |r_A \cup r_B| \quad (1)$$

$\text{sim}(A, B)$ – подібність товарів або користувачів;

r_A і r_B – множина товарів, які купили відповідно користувачі А і В;

Недоліком метрики Jaccard є те, що ігноруються значення матриці. Метрика подібності на основі косинуса кута формально представляється, як показано нижче

$$\cos(\alpha) \equiv \frac{\bar{a} \cdot \bar{b}}{|\bar{a}| \cdot |\bar{b}|} \quad (2)$$

Метрика на основі центрованого косинуса кута представляється наступним чином:

$$\text{sim}(u, u') = \frac{\sum_{i \in I(u, u')} (r_{ui} - \bar{r}_u) \cdot (r_{u'i} - \bar{r}_{u'})}{\sqrt{\sum_{i \in I(u, u')} (r_{ui} - \bar{r}_u)^2 \cdot \sum_{i \in I(u, u')} (r_{u'i} - \bar{r}_{u'})^2}} \quad (3)$$

Таким чином, застосування метрик подібності дає змогу визначити подібні товари або користувачів, однак їх застосування не вирішує проблему «холодного старту» (у випадку появи нового користувача). Тому подальшого дослідження потребують методи колаборативної фільтрації та їх композиція з іншими методами штучного інтелекту, зокрема з методами на основі латентних моделей.