CS-4053 Recommender System

Spring 2023

Lecture 5: Naïve Bayes Collaborative Filtering

Course Instructor: Syed Zain Ul Hassan

National University of Computer and Emerging Sciences, Karachi

Email: zain.hassan@nu.edu.pk

Probabilistic Methods

- Given the user-item interaction matrix:
  - Find the probability that active user will like a given item
  - The rating is predicted based on probabilities
- Recommendations provided are more accurate

Naïve Bayes Classifier

- A supervised multi-class classification algorithm
- It is based on Bayes Theorem
- It has a naïve assumption that all pairs of variables are independent

Posterior, Likelihood, Prior

$$\underbrace{P(Y|X)}_{\text{Posterior}} = \frac{\overbrace{P(X|Y)}^{\text{Likelihood}} \times \overbrace{P(Y)}^{\text{Prior}}}{P(X)}$$

$$P(Y|X) = \frac{\prod_{i=1}^{d} P(X_i|Y) \times P(Y)}{P(X)}$$

Naïve Bayes Classifier

- It has a naïve assumption that all pairs of variables are independent

Posterior Likelihood Prior

$$P(Y|X) = \frac{P(X|Y) \times P(Y)}{P(X)}$$

$$P(Y|X) = \frac{\prod_{i=1}^{d} P(X_i|Y) \times P(Y)}{P(X)}$$

- Final classification is produced by the argument that maximizes ***y***

$$y = \mathrm{argmax}_{\mathrm{y}}\, \mathrm{P(y)} \prod_{i=1}^{n} P(xi \mid y)$$

Naïve Bayes Collaborative Filtering

- Calculation of rating probabilities based on Bayes rule
- Assumes ratings are independent
- This approach can both be user-based and item-based
- Proposed by *Priscila Valdiviezo-Diaz*[1]

Naïve Bayes Collaborative Filtering: User-based Approach

- We define **$P(r_i = y)$** as the prior probability that the **Item i** be rated by any user as **y**

$$P(r_i = y) = \frac{(\#\ \text{of users who rated item } i \text{ as } y) + \alpha}{(\#\ \text{of users who have rated item } i) + \beta}$$

- Where α is a hyper-parameter to avoid 0 probabilities and β is **R** x α (for R = no. of possible ratings)

Naïve Bayes Collaborative Filtering: User-based Approach

- We define **P(r_j = k | r_i = y)** as the likelihood that the **Item j** be rated **k** given that **Item i** is rated as **y**

$$\boldsymbol{P(r_j = k \mid ri = y) = \frac{(\#\ of\ users\ who\ rated\ item\ j\ as\ k\ and\ item\ \ i\ as\ y) +\ \alpha}{(\#\ of\ users\ who\ have\ rated\ item\ j\ and\ rated\ item\ i\ as\ y) +\ \beta}}$$

- Where $\boldsymbol{\alpha}$ is a hyper-parameter to avoid 0 probabilities and $\boldsymbol{\beta}$ is **R** x $\boldsymbol{\alpha}$ (for R = no. of possible ratings)

Naïve Bayes Collaborative Filtering:
User-based Approach Using an Example

- How probable is the rating 1 for Item 5 using Naïve Bayes approach

	Item1	Item2	Item3	Item4	Item5
User 1	1	3	3	2	?
User 2	2	4	2	2	4
User 3	1	3	3	5	1
User 4	4	5	2	3	3
User 5	1	1	5	2	1

Naïve Bayes Collaborative Filtering:
User-based Approach Using an Example

- How probable is the rating **1** for **Item 5** using Naïve Bayes approach

- Corresponds to conditional probability **P(Item 5 = 1 | X)**, where **X = User 1's previous ratings = (Item1=1, Item2=3, Item3= …)**

	Item1	Item2	Item3	Item4	Item5
User 1	1	3	3	2	?
User 2	2	4	2	2	4
User 3	1	3	3	5	1
User 4	4	5	2	3	3
User 5	1	1	5	2	1

Naïve Bayes Collaborative Filtering:
User-based Approach Using an Example

	Item1	Item2	Item3	Item4	Item5
User 1	1	3	3	2	?
User 2	2	4	2	2	4
User 3	1	3	3	5	1
User 4	4	5	2	3	3
User 5	1	1	5	2	1

$$\boldsymbol{P(X|\, r_{i5} = 1)} = P(r_{i1} = 1 \,|\, r_{i5} = 1) * P(r_{i2} = 3 \,|\, r_{i5} = 1) * P(r_{i3} = 3 \,|\, r_{i5} = 1) * P(r_{i4} = 2 \,|\, r_{i5} = 1)$$

$$= \frac{2+0.01}{2+0.05} * \frac{1+0.01}{2+0.05} * \frac{1+0.01}{2+0.05} * \frac{1+0.01}{2+0.05} = \mathbf{0.117}$$

X = our input features (existing ratings) e.g. r_{i1}, r_{i2},...

Naïve Bayes Collaborative Filtering:
User-based Approach Using an Example

	Item1	Item2	Item3	Item4	Item5
User 1	1	3	3	2	?
User 2	2	4	2	2	4
User 3	1	3	3	5	1
User 4	4	5	2	3	3
User 5	1	1	5	2	1

$$\boldsymbol{P(X|\, r_{i5} = 2)} = P(r_{i1} = 1\,|\, r_{i5} = 2) * P(r_{i2} = 3\,|\, r_{i5} = 2) * P(r_{i3} = 3\,|\, r_{i5} = 2) * P(r_{i4} = 2\,|\, r_{i5} = 2)$$

$$= \frac{0+0.01}{0+0.05} * \frac{0+0.01}{0+0.05} * \frac{0+0.01}{0+0.05} * \frac{0+0.01}{0+0.05} = \mathbf{0.0016}$$

X = our input features (existing ratings) e.g. r_{i1}, r_{i2},...

Naïve Bayes Collaborative Filtering: User-based Approach Using an Example

	Item1	Item2	Item3	Item4	Item5
User 1	1	3	3	2	?
User 2	2	4	2	2	4
User 3	1	3	3	5	1
User 4	4	5	2	3	3
User 5	1	1	5	2	1

$$\boldsymbol{P(X|\, r_{i5} = 3)} = P(r_{i1} = 1\,|\, r_{i5} = 3) * P(r_{i2} = 3\,|\, r_{i5} = 3) * P(r_{i3} = 3\,|\, r_{i5} = 3) * P(r_{i4} = 2\,|\, r_{i5} = 3)$$

$$= \frac{0+0.01}{1+0.05} * \frac{0+0.01}{1+0.05} * \frac{0+0.01}{1+0.05} * \frac{0+0.01}{1+0.05} = \mathbf{0.00000082}$$

X = our input features (existing ratings) e.g. r_{i1}, r_{i2},...

Naïve Bayes Collaborative Filtering:
User-based Approach Using an Example

	Item1	Item2	Item3	Item4	Item5
User 1	1	3	3	2	?
User 2	2	4	2	2	4
User 3	1	3	3	5	1
User 4	4	5	2	3	3
User 5	1	1	5	2	1

$$\boldsymbol{P(X|\, r_{i5} = 4)} = P(r_{i1} = 1 \,|\, r_{i5} = 4) * P(r_{i2} = 3 \,|\, r_{i5} = 4) * P(r_{i3} = 3 \,|\, r_{i5} = 4) * P(r_{i4} = 2 \,|\, r_{i5} = 4)$$

$$= \frac{0+0.01}{1+0.05} * \frac{0+0.01}{1+0.05} * \frac{0+0.01}{1+0.05} * \frac{1+0.01}{1+0.05} = \mathbf{0.00000083}$$

X = our input features (existing ratings) e.g. r_{i1}, r_{i2},...

Naïve Bayes Collaborative Filtering:
User-based Approach Using an Example

	Item1	Item2	Item3	Item4	Item5
User 1	1	3	3	2	?
User 2	2	4	2	2	4
User 3	1	3	3	5	1
User 4	4	5	2	3	3
User 5	1	1	5	2	1

$$\boldsymbol{P(X|\, r_{i5} = 5)} = P(r_{i1} = 1 \,|\, r_{i5} = 5) * P(r_{i2} = 3 \,|\, r_{i5} = 5) * P(r_{i3} = 3 \,|\, r_{i5} = 5) * P(r_{i4} = 2 \,|\, r_{i5} = 5)$$

$$= \frac{0+0.01}{0+0.05} * \frac{0+0.01}{0+0.05} * \frac{0+0.01}{0+0.05} * \frac{0+0.01}{0+0.05} = \mathbf{0.0016}$$

X = our input features (existing ratings) e.g. r_{i1}, r_{i2},...

Naïve Bayes Collaborative Filtering: User-based Approach Using an Example

	Item1	Item2	Item3	Item4	Item5
User 1	1	3	3	2	?
User 2	2	4	2	2	4
User 3	1	3	3	5	1
User 4	4	5	2	3	3
User 5	1	1	5	2	1

$$\boldsymbol{P(r_{i5} = 1) = \frac{2 + 0.01}{4 + 0.05} = 0.496}$$

Naïve Bayes Collaborative Filtering: User-based Approach Using an Example

	Item1	Item2	Item3	Item4	Item5
User 1	1	3	3	2	?
User 2	2	4	2	2	4
User 3	1	3	3	5	1
User 4	4	5	2	3	3
User 5	1	1	5	2	1

$$P(r_{i5} = 1 \mid X) = P(r_{i5} = 1) * P(X \mid ri_5 = 1)$$

$$P(r_{i5} = 1 \mid X) = (0.496) * (0.117) = 0.058$$

Naïve Bayes Collaborative Filtering:
User-based Approach Using an Example

	Item1	Item2	Item3	Item4	Item5
User 1	1	3	3	2	?
User 2	2	4	2	2	4
User 3	1	3	3	5	1
User 4	4	5	2	3	3
User 5	1	1	5	2	1

$$\boldsymbol{P(r_{i5} = 1 \mid X) = (0.496) * (0.117) = 0.058}$$

- Find all other posterior probabilities the same way and select the rating value that gives us the maximum posterior probability

Naïve Bayes Collaborative Filtering: Item-based Approach Using an Example

- Now we predict rating 1 for Item 5 using item-based Naïve Bayes approach

	Item1	Item2	Item3	Item4	Item5
User 1	1	3	3	2	?
User 2	2	4	2	2	4
User 3	1	3	3	5	1
User 4	4	5	2	3	3
User 5	1	1	5	2	1

Naïve Bayes Collaborative Filtering: Item-based Approach

- We define **P(r_u = y)** as the prior probability that the **active user** gives any item the rating **y**

$$P(r_{ui} = y) = \frac{(\#\ of\ items\ that\ user\ has\ given\ a\ rating\ y) + \alpha}{(\#\ of\ total\ items\ the\ user\ has\ rated) + \beta}$$

- Where α is a hyper-parameter to avoid 0 probabilities and β is **R** x α (for R = no. of possible ratings)

Naïve Bayes Collaborative Filtering: Item-based Approach

- We define $P(r_{uj} = y \mid r_{ui} = y)$ as the likelihood that the **user j** will give rating **y** given that **user i** has also given a rating **y**

$$P(r_{uj} = y \mid r_{ui} = y) =$$

$$\frac{(\#\ of\ items\ that\ both\ user\ j\ and\ user\ i\ has\ rated\ as\ y) + \alpha}{(\#\ of\ items\ that\ user\ j\ has\ given\ a\ rating\ and\ for\ which\ user\ i\ has\ given\ rating\ y) + \beta}$$

- Where α is a hyper-parameter to avoid 0 probabilities and β is **R** x α (for R = no. of possible ratings)

Naïve Bayes Collaborative Filtering: Item-based Approach Using an Example

	Item1	Item2	Item3	Item4	Item5
User 1	1	3	3	2	?
User 2	2	4	2	2	4
User 3	1	3	3	5	1
User 4	4	5	2	3	3
User 5	1	1	5	2	1

$$\boldsymbol{P(r_{uj} = 1 \mid r_{u1} = 1)} = P(r_{u2} = 1 \mid r_{u1} = 1) * P(r_{u3} = 1 \mid r_{u1} = 1) * P(r_{u4} = 1 \mid r_{u1} = 1) * P(r_{u5} = 1 \mid r_{u1} = 1)$$

$$\boldsymbol{P(r_{uj} = 1 \mid ru_1 = 1)} = \frac{0 + 0.01}{1 + 0.05} * \frac{1 + 0.01}{1 + 0.05} * \frac{0 + 0.01}{1 + 0.05} * \frac{1 + 0.01}{1 + 0.05} = \boldsymbol{0.0000839}$$

X = our input features (existing ratings) e.g. r_{u1}, r_{u2},...

Naïve Bayes Collaborative Filtering:
Item-based Approach Using an Example

	Item1	Item2	Item3	Item4	Item5
User 1	1	3	3	2	?
User 2	2	4	2	2	4
User 3	1	3	3	5	1
User 4	4	5	2	3	3
User 5	1	1	5	2	1

$$P(r_{u1} = 1) = \frac{1}{4} = 0.25$$

Naïve Bayes Collaborative Filtering:
Item-based Approach Using an Example

	Item1	Item2	Item3	Item4	Item5
User 1	1	3	3	2	?
User 2	2	4	2	2	4
User 3	1	3	3	5	1
User 4	4	5	2	3	3
User 5	1	1	5	2	1

$$\boldsymbol{P(r_{u1,\, i5} = 1 \mid X)} = \ (0.25) * (0.0000839) = \mathbf{0.0000209}$$

X = our input features (existing ratings) e.g. $r_{u1}, r_{u2}, \ldots$

Naïve Bayes Collaborative Filtering: Item-based Approach Using an Example

	Item1	Item2	Item3	Item4	Item5
User 1	1	3	3	2	?
User 2	2	4	2	2	4
User 3	1	3	3	5	1
User 4	4	5	2	3	3
User 5	1	1	5	2	1

$$\boldsymbol{P(r_{u1,\,i5} = 1 \mid X)} = (0.25) * (0.0000839) = \mathbf{0.0000209}$$

- Find all other posterior probabilities the same way and select the rating value that gives us the maximum posterior probability

Naïve Bayes Collaborative Filtering: Pros and Cons

Pros

- Provides more accurate recommendations in general
- Can provide ranking of predicted ratings
- Can provide confidence level for a prediction

Cons

- Can become computationally intractable
- Serendipity cannot be controlled
- Independence between ratings is required