

# IP items recommender

T.G. Hwang

October 15, 2024

## Keywords

- data mining
- linear regression
- tags scores
- preferences prediction

## Contents

<b>1</b>	<b>IPIRecModel</b>	<b>2</b>
1.1	Series1 . . . . .	2
1.2	Series2 . . . . .	3
1.3	Series3 . . . . .	4
<b>2</b>	<b>Estimators</b>	<b>5</b>
2.1	Series1 . . . . .	5
2.2	Series2 . . . . .	6
2.3	Series3 . . . . .	8
2.4	Series4 . . . . .	9

# 1 IPIRecModel

## 1.1 Series1

---

**Algorithm 1** IPIRecModelSeries1

---

**Require:** No. of top n decisioned tags  $NT(u)$ , threshold of co-occurred items

Freq.  $\theta_I$

PREPROCESS( $NT(u), \theta_I$ )

PROCESS( )

**function** PREPROCESS( $NT(u), \theta_I$ )

  top\_n\_decisioned\_tags()

  mean\_freq\_tags()

**function** PROCESS( )

  item\_based\_tags\_corr()

▷ in Equation (1)

  user\_based\_tags\_corr()

▷ in Equation (2)

  tags\_score()

▷ in Equation (3)

---

$$sim_U(x, y) = \frac{\sum_{u \in U(x) \cap U(y)} (r(u, x) - \bar{r}(x))(r(u, y) - \bar{r}(y))}{\sqrt{\sum_u (r(u, x) - \bar{r}(x))^2 \sum_u (r(u, y) - \bar{r}(y))^2}} \quad (1)$$

$$sim_I(x, y) = \frac{\sum_{i \in I(x) \cap I(y)} (r(i, x) - \bar{r}(x))(r(i, y) - \bar{r}(y))}{\sqrt{\sum_i (r(i, x) - \bar{r}(x))^2 \sum_i (r(i, y) - \bar{r}(y))^2}} \quad (2)$$

$$s(x, y) = \text{CoOccur}(x, y) \cdot \text{Pr}(x, y) \cdot \frac{1}{|\mathcal{S}|} \sum_{d \in \mathcal{S}} sim_d(x, y) \quad (3)$$

$$\text{CoOccur}(x, y) = \frac{\ln(\min(|I(x) \cap I(y)| + 1, \theta_I))}{\ln(\max_{x', y' \in T} |I(x') \cap I(y')|)}; \theta_I = 20 \text{ (optional)}$$

$$\text{Pr}(x, y) = |I(x)|^{-1} |I(x) \cap I(y)|$$

$$\bar{r}(x) = \frac{\sum_u |I(u) \cap I(V)|}{|U(x) \cap U(V)|} + \frac{\sum_i |U(i) \cap U(V)|}{|I(x) \cap I(V)|}$$

## 1.2 Series2

---

**Algorithm 2** IPIRecModelSeries2

---

**Require:** No. of top n decisioned tags  $NT(u)$ , threshold of co-occurred items

```

    Freq.  $\theta_I$ 
    PREPROCESS( $NT(u), \theta_I$ )
    PROCESS( )
    POSTPROCESS( )

function PREPROCESS( $NT(u), \theta_I$ )
    |   top_n_decisioned_tags()
    |   mean_freq_tags()

function PROCESS( )
    |   item_based_tags_corr()           ▷ in Equation (1)
    |   user_based_tags_corr()          ▷ in Equation (2)
    |   tags_score()                     ▷ in Equation (3)

function POSTPROCESS( )
    |   append_tendencies()              ▷ in Equation (4)

```

---

$$S \leftarrow S + \mathbb{B}(\mathbb{X}(V), S), \quad (4)$$

$$\bar{b}(\mathbb{X}(V), x, y) = \hat{b}^U(\mathbb{X}(V), x, y) + \hat{b}^I(\mathbb{X}(V), x, y) + \frac{b(\mathbb{X}(V), x) + b(\mathbb{X}(V), y)}{2}$$

$$\begin{aligned}
 b(\mathbb{X}(V), t) &= \frac{|R(t)|}{|R|} - \mu(\mathbb{X}(V)) \\
 \hat{b}^U(\mathbb{X}(V), x, y) &= \frac{\sum_{u \in U(x) \cup U(y)} b(u)}{|U(x) \cup U(y)|} \\
 b(\mathbb{X}(V), u) &= \frac{|I(u)|}{|I|} - \mu(\mathbb{X}(V)) \\
 \hat{b}^I(\mathbb{X}(V), x, y) &= \frac{\sum_{i \in I(x) \cup I(y)} b(i)}{|I(x) \cup I(y)|} \\
 b(\mathbb{X}(V), i) &= \frac{|U(i)|}{|U|} - \mu(\mathbb{X}(V))
 \end{aligned}$$

### 1.3 Series3

---

**Algorithm 3** IPIRecModelSeries3

---

**Require:** training set  $\mathbb{X}$ , test set  $\mathbb{Y}$

$\rho = \text{corr}(\mathbb{X}(V))$   $\triangleright$  in Equation (5)

$\mathcal{B} = \text{tendencies}(\mathbb{X}(V), \rho)$   $\triangleright$  in Equation (6)

$S = \rho + \mathcal{B}$   $\triangleright$  in Equation (7)

---

$$\rho(x, y) = J(x, y) \Pr(x, y) \Pi_{d \in \mathcal{S}} \text{sim}_d(x, y) \quad (5)$$

$$J(x, y) = \frac{|I(x) \cap I(y)|}{|I(x) \cup I(y)|}$$

$$\Pr(x, y) = \frac{|I(x) \cap I(y)|}{|I(x)|}$$

$$\text{sim}_{UB}(x, y) = \frac{\sum_{u \in U(x) \cup U(y)} |NT(u) \cap x| |NT(u) \cap y|}{\sqrt{\sum_u |NT(u) \cap x| \sum_u |NT(u) \cap y|}}$$

$$\text{sim}_{UF}(x, y) = \frac{\sum_{u \in U(x) \cup U(y)} |I(u) \cap I(x)| |I(u) \cap I(y)|}{\sqrt{\sum_u |I(u) \cap I(x)|^2 \sum_u |I(u) \cap I(y)|^2}}$$

$$\text{sim}_{IB}(x, y) = \frac{\sum_{i \in I(x) \cup I(y)} |T(i) \cap x| |T(i) \cap y|}{\sqrt{\sum_i |T(i) \cap x|^2 \sum_i |T(i) \cap y|^2}}$$

$$\text{sim}_{IF}(x, y) = \frac{\sum_{i \in I(x) \cup I(y)} \tau(x) \tau(y)}{\sqrt{\sum_i \tau(x)^2 \sum_i \tau(y)^2}}$$

$$\tau(t) = |I(t)|^{-1} \sum_{i \in I(t)} |T(i)|$$

$$\bar{b}(x, y) = \hat{b}^U(x, y) + \hat{b}^I(x, y) + \frac{b(x) + b(y)}{2} \quad (6)$$

$$b(t) = \frac{|R(t)|}{|R|} - \mu(\mathbb{X})$$

$$\hat{b}^U(x, y) = \frac{\sum_{u \in U(x) \cup U(y)} b(u)}{|U(x) \cup U(y)|}$$

$$b(u) = \frac{|I(u)|}{|I|} - \mu(\mathbb{X})$$

$$\hat{b}^I(x, y) = \frac{\sum_{i \in I(x) \cup I(y)} b(i)}{|I(x) \cup I(y)|}$$

$$b(i) = \frac{|U(i)|}{|U|} - \mu(\mathbb{X})$$

$$s(x, y) \leftarrow \rho(x, y) + \bar{b}(x, y) \quad (7)$$

## 2 Estimators

### 2.1 Series1

---

**Algorithm 4** IPIRecEstimatorSeries1

---

**Require:** *iter*

```

ADJUST_TAGS_SCORE( $\mathbb{X}(V)$ )
for  $a \in [V, L, P]$  do
    for  $\_$  in  $\text{range}(\text{iter})$  do
        TRAINING( $\mathbb{X}(a)$ )
function ADJUST_TAGS_SCORE( $\mathbb{X}$ )
     $W = \text{numpy.ones}(|T|, |T|)$ 
    for  $r(u, i) \in \mathbb{X}$  do
         $\hat{r}(u, i) = \text{estimate}(u, i)$   $\triangleright$  in Equation (8)
         $W = \text{feed\_back}(r(u, i), \hat{r}(u, i), W)$   $\triangleright$  in Equation (9)
     $S = SW$ 
function TRAINING( $\mathbb{X}$ )
    for  $r(u, i) \in \mathbb{X}$  do
         $\hat{r}(u, i) = \text{estimate}(u, i)$   $\triangleright$  in Equation (8)
         $\Omega = \text{feed\_back}(r(u, i), \hat{r}(u, i), \Omega)$   $\triangleright$  in Equation (10)

```

---

$$\hat{r}(u, i) = \frac{\sum_x \sum_y \omega(u, x, y) s(x, y) \nu(u, y)}{\sum_{x \in NT(u)} \sum_{y \in T(i)} |\omega(u, x, y)|} \quad (8)$$

$$\nu(u, y) = \begin{cases} NT(u) \cap y = \phi & \text{Default voting} \\ \text{otherwise} & 1 \end{cases}$$

$$\epsilon(u, i) = r(u, i) - \hat{r}(u, i)$$

$$w(x, y) = \frac{w + \eta(w(x, y) \epsilon(u, i) \lambda^S \gamma^S s(x, y))}{\sum_{z \in T(i)} |w(x, z) s(x, z)|} \quad (9)$$

$$\eta(x) = \begin{cases} x \geq 0 & 2^{-1}x \\ \text{otherwise} & 2x \end{cases}$$

$$\mathbb{W}(u) = \sum_{x \in NT(u)} \sum_{y \in T(i)} |\omega(u, x, y)|$$

$$\omega(u, x, y) = \omega(u, x, y) \frac{r(u, i) + \lambda^\Omega s(x, y)}{\hat{r}(u, i) + \gamma^\Omega \mathbb{W}(u)} \quad (10)$$

## 2.2 Series2

---

**Algorithm 5** IPIRecEstimatorSeries2

---

**Require:** *iter*

```

 $\Omega = \text{numpy.ones}(|U|, |T|, |T|)$ 
ADJUST_TAGS_SCORE( $\mathbb{X}(V)$ )
for  $a \in [V, L, P]$  do
    for  $\_$  in  $\text{range}(iter)$  do
        TRAINING( $\mathbb{X}(a)$ )
function ADJUST_TAGS_SCORE( $\mathbb{X}$ )
     $W = \text{numpy.ones}(|T|, |T|)$ 
    for  $r(u, i) \in \mathbb{X}$  do
        for  $x \in NT(u)$  do
            for  $y \in T(i)$  do
                 $\hat{r}(u, i) = \text{estimate}(u, i)$   $\triangleright$  in Equation (11)
                 $W = \text{feed\_back}(r(u, i), \hat{r}(u, i), W)$   $\triangleright$  in Equation (12)
     $S = SW$ 
function TRAINING( $\mathbb{X}$ )
    for  $r(u, i) \in \mathbb{X}$  do
         $\hat{r}(u, i) = \text{estimate}(u, i)$   $\triangleright$  in Equation (11)
         $\Omega = \text{fit\_weights}(r(u, i), \hat{r}(u, i), S, \Omega)$   $\triangleright$  in Equation (13)
         $S = \text{fit\_scores}(r(u, i), \hat{r}(u, i), S, \Omega)$   $\triangleright$  in Equation (14)

```

---

$$\hat{r}(u, i) = \frac{\sum_x \sum_y \omega(u, x, y) s(x, y) \nu(u, y)}{\sum_{x \in NT(u)} \sum_{y \in T(i)} |\omega(u, x, y)|} \quad (11)$$

$$\nu(u, y) = \begin{cases} NT(u) \cap y = \phi & DV \\ \text{otherwise} & 1 \end{cases}$$

$$\begin{aligned} \epsilon(u, i) &= r(u, i) - \hat{r}(u, i) \\ s(x, y) &= s(x, y) + \lambda^S \eta \left( \epsilon(u, i) \left( \frac{s(x, y)}{\|S\|_2} \right) \right) + \gamma^S \frac{\bar{s}(x) + \bar{s}(y)}{2} \end{aligned} \quad (12)$$

$$\eta(x) = \begin{cases} x \geq 0, 2^{-1}x \\ \text{otherwise}, 2x \end{cases}$$

$$\omega(u, x, y) = \omega(u, x, y) \frac{(r(u, i) + \alpha(\hat{r}(u, i))) + \lambda^\Omega s(x, y)}{(\hat{r}(u, i) + \alpha(\hat{r}(u, i))) + \gamma^\Omega \mathbb{W}(u, i)} \quad (13)$$

$$\alpha(x) = \begin{cases} x = 0 & 1 \\ \text{otherwise} & 0 \end{cases}$$

$$\mathbb{W}(u, i) = \sum_{x \in NT(u)} \sum_{y \in T(i)} |\omega(u, x, y)|$$

$$s(x, y) = s(x, y) \frac{(r(u, i) + \alpha(\hat{r}(u, i))) + \lambda^\Omega \omega(u, x, y)}{(\hat{r}(u, i) + \alpha(\hat{r}(u, i))) + \gamma^\Omega \mathbb{S}(u, i)} \quad (14)$$

$$\mathbb{S}(u, i) = \sum_{NT(u)} \sum_{y \in T(i)} |s(x, y)|$$

## 2.3 Series3

---

**Algorithm 6** IPIRecEstimatorSeries3

---

**Require:** Train.Set.  $\mathbb{X}$ , main Seq.  $\mathcal{A}^{\text{main}}$ , post Seq.  $\mathcal{A}^{\text{post}}$

$\Omega \leftarrow \text{np.ones(shape}=(|U|, |T|, |T|))$

$S = \text{ADJUST\_SCORES}(\mathbb{X}(V), S, \Omega)$

**for**  $a \in \mathcal{A}^{\text{main}}$  **do**

$\mathcal{L}_{\min} = \infty$

**while** True **do**

$\mathbb{L}^\Omega, S, \Omega = \text{FIT\_WEIGHTS}(\mathbb{X}, a, S, \Omega)$

$\mathbb{L}^S, S, \Omega = \text{FIT\_SCORES}(\mathbb{X}, a, S, \Omega)$

$\mathcal{L} = \frac{2\mathbb{L}^S \mathbb{L}^\Omega}{\mathbb{L}^S + \mathbb{L}^\Omega}$

$\triangleright$  in Equation (18)

**if**  $\mathcal{L}_{\min} \geq \mathcal{L}$  **then**

$\mathcal{L}_{\min} = \mathcal{L}$

$S_{\min} = S$

$\Omega_{\min} = \Omega$

**else**

$\mathcal{L} = \mathcal{L}_{\min}$

$S = S_{\min}$

$\Omega = \Omega_{\min}$

**break**

**for**  $a \in \mathcal{A}^{\text{post}}$  **do**

**if**  $a \neq V$  **then**

$\_, S, \Omega = \text{FIT\_SCORES}(\mathbb{X}(V), S, \Omega)$

$\_, S, \Omega = \text{FIT\_SCORES}(\mathbb{X}(a), S, \Omega)$

**function** ADJUST\_SCORES( $\mathbb{X}, S, \Omega$ )

$\_, S, \_ = \text{FIT\_SCORES}(\mathbb{X}, V, S, \Omega)$

$\mathcal{B} = \text{tendencies}(\mathbb{X}, S)$

$S = S + \mathcal{B}$

$\_, S, \_ = \text{FIT\_SCORES}(\mathbb{X}, V, S, \Omega)$

**return**  $S$

$\triangleright$  in Equation (6)

**function** FIT\_SCORES( $\mathbb{X}, a, S, \Omega$ )

**if**  $a \neq V$  **then**

$S, \Omega = \arg_{S', \Omega'} \min \mathbb{L}(\mathbb{X}(V), S, \Omega)$

**return**  $\arg_{\mathbb{L}', S', \Omega'} \min \mathbb{L}(\mathbb{X}(a), S, \Omega)$

**function** FIT\_WEIGHTS( $\mathbb{X}, a, S, \Omega$ )

$\mathbb{L}, S, \Omega = \arg_{\mathbb{L}', S', \Omega'} \min \mathbb{L}(\mathbb{X}(a), S, \Omega)$

$\_, S, \Omega = \text{FIT\_SCORES}(\mathbb{X}, a, S, \Omega)$

**return**  $\mathbb{L}, S, \Omega$

---



$$\hat{r}(u, i) = \frac{1}{|T(u)|} \sum_{x \in T(u)} \frac{\Pi_{y \in T(i)} \omega(u, x, y) s(x, y)}{\Pi_{y \in T(i)} |\omega(u, x, y)|} \quad (15)$$

$$\begin{aligned} \epsilon(u, i) &= r(u, i) - \hat{r}(u, i) \\ s(x, y) &= s(x, y) + \lambda^S \eta \left( \epsilon(u, i) \left( \frac{s(x, y)}{\|S\|_2} + \frac{\bar{s}(x) + \bar{s}(y)}{2} \right) \right) \end{aligned} \quad (16)$$

$$\eta(x) = \begin{cases} x \geq 0 & 2^{-1}x \\ \text{otherwise} & 2x \end{cases}$$

$$\omega(u, x, y) = \omega(u, x, y) \frac{r(u, i) + \lambda^\Omega s(x, y)}{\hat{r}(u, i) + \gamma^\Omega \mathbb{W}(u, i)}$$

$$\mathbb{W}(u, i) = \sum_{x \in NT(u)} \sum_{y \in T(i)} |\omega(u, x, y)|$$

$$s(x, y) = s(x, y) \frac{r(u, i) + \lambda^\Omega \omega(u, x, y)}{\hat{r}(u, i) + \gamma^\Omega \mathbb{S}(u, i)}$$

$$\mathbb{S}(u, i) = \sum_{NT(u)} \sum_{y \in T(i)} |s(x, y)|$$

$$\mathbb{L}(\mathbb{X}, S, \Omega) = \text{RMSE}(\mathbb{X}, S, \Omega) = \sqrt{\frac{1}{|\mathbb{X}|} \sum_{(u, i) \in \mathbb{X}} (r(u, i) - \hat{r}(u, i))^2} \quad (17)$$

$$\mathbb{L}^S(\mathbb{X}(a), S, \Omega) = \min \text{RMSE}^S(\mathbb{X}(a), S, \Omega)$$

$$\mathbb{L}^\Omega(\mathbb{X}, S, \Omega) = \min \text{RMSE}^\Omega(\mathbb{X}(a), S, \Omega)$$

$$\mathcal{L}(\mathbb{X}(a), S, \Omega) = \frac{2\mathbb{L}^S(\mathbb{X}(a), S, \Omega)\mathbb{L}^\Omega(\mathbb{X}(a), S, \Omega)}{\mathbb{L}^S(\mathbb{X}(a), S, \Omega) + \mathbb{L}^\Omega(\mathbb{X}(a), S, \Omega)} \quad (18)$$

$$\min \mathcal{L}(\mathbb{X}, S, \Omega)$$

## 2.4 Series4

식 (19)를 사용한 알고리즘 6으로 모델을 훈련함

$$\mathcal{F}(\mathbb{X}(a), S, \Omega) = \frac{|\mathbb{X}(a)|}{|U|^2|I|} \sum_{u \in U} \left( \frac{1}{|T(u)||I|} \sum_{x \in T(u)} \sum_{y \in T} (w(u, x, y) s(x, y))^2 \right)^{0.5}$$

$$\mathbb{L}(\mathbb{X}, S, \Omega) = \text{RMSE}(\mathbb{X}, S, \Omega) + \mathcal{F}(\mathbb{X}(a), S, \Omega) \quad (19)$$