

Práctica 8

Sistemas de recomendación

Javier Herrer Torres (NIP: 776609)



Aprendizaje automático
Grado en Ingeniería Informática



**Escuela de
Ingeniería y Arquitectura**
Universidad Zaragoza

Escuela de Ingeniería y Arquitectura
Universidad de Zaragoza
Curso 2020/2021

1. Objetivo

El objetivo es implementar un sistema de recomendación de películas basado en las valoraciones de los usuarios.

2. Estudio previo

- $r(i, j) = 1$ si el usuario j ha calificado el item i
- $y^{(i,j)}$: calificación de j para el item i (definido si $r(i, j) = 1$)
- $x^{(i)}$: atributos del item i
- $\theta^{(j)}$: parámetros del usuario j
- El usuario j calificará el item i como $(\theta^{(j)})^T x^{(i)}$
- $m^{(j)}$: número de items calificados por el usuario j

3. Metodología

3.1. Leer los datos

Se han cargado los siguientes datos:

- **X**: n.º de películas \times n.º de parámetros. La i -ésima fila de **X** corresponde a $x^{(i)}$ para la i -ésima película.
- **Theta**: n.º de usuarios \times n.º de parámetros. La j -ésima fila de **Theta** corresponde a $\theta^{(j)}$ para el j -ésimo usuario.
- **Y**: n.º de películas \times n.º de usuarios. Almacena la calificación $y^{(i,j)}$ de la i -ésima película para el usuario j .
- **R**: n.º de películas \times n.º de usuarios. Es una matriz binaria que almacena $r(i, j) = 1$ si el usuario j ha calificado la película i .

3.2. Implementar la función de coste

Una vez cargados los datos, se ha procedido a implementar la siguiente función de coste *sin regularización*.

$$J(x^{(1)}, \dots, x^{(n_m)}, \theta^{(1)}, \dots, \theta^{(n_m)}) = \frac{1}{2} \sum_{(i,j): r(i,j)=1} ((\theta^{(j)})^T x^{(i)} - y^{(i,j)})^2 \quad (1)$$

Con ella, se penaliza la diferencia entre el valor que se predice y el valor que el usuario ha dado de los items que ya conoce.

3.3. Añadir el gradiente

Después, se han añadido los gradientes tanto de $x_k^{(i)}$ como de $\theta_k^{(j)}$:

$$\frac{\partial J}{\partial x_k^{(i)}} = \sum_{j:r(i,j)=1} ((\theta^{(j)})^T x^{(i)} - y^{(i,j)}) \theta_k^{(j)} \quad (2)$$

$$\frac{\partial J}{\partial \theta_k^{(j)}} = \sum_{i:r(i,j)=1} ((\theta^{(j)})^T x^{(i)} - y^{(i,j)}) x_k^{(i)} \quad (3)$$

3.4. Añadir la regularización

Al añadirle regularización a la ecuación 1 se obtiene:

$$\begin{aligned} J(x^{(1)}, \dots, x^{(n_m)}, \theta^{(1)}, \dots, \theta^{(n_u)}) &= \frac{1}{2} \sum_{(i,j):r(i,j)=1} ((\theta^{(j)})^T x^{(i)} - y^{(i,j)})^2 \\ &+ \frac{\lambda}{2} \sum_{i=1}^{n_m} \sum_{k=1}^n (x_k^{(i)})^2 + \frac{\lambda}{2} \sum_{j=1}^{n_u} \sum_{k=1}^n (\theta_k^{(j)})^2 \end{aligned} \quad (4)$$

Esta función de distorsión tiene en cuenta tanto los atributos de cada ítem como los parámetros de cada ítem al mismo tiempo. Esta fórmula se construye penalizando valores grandes de θ y de x (últimas dos partes).

Añadiéndole regularización a las ecuaciones 2 y 3 se obtiene (para $k \neq 0$):

$$\frac{\partial J}{\partial x_k^{(i)}} = \sum_{j:r(i,j)=1} ((\theta^{(j)})^T x^{(i)} - y^{(i,j)}) \theta_k^{(j)} + \lambda x_k^{(i)} \quad (5)$$

$$\frac{\partial J}{\partial \theta_k^{(j)}} = \sum_{i:r(i,j)=1} ((\theta^{(j)})^T x^{(i)} - y^{(i,j)}) x_k^{(i)} + \lambda \theta_k^{(j)} \quad (6)$$

3.5. Entrenar el recomendador

Al minimizar la ecuación 4, el algoritmo de optimización debe encontrar el mejor valor de todos los $\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(n_u)}$ de forma que se minimice el promedio de los errores cuadráticos que se cometen, favoreciendo valores bajos para todos los θ .

Si se utilizan valores iniciales iguales a 0, en las ecuaciones 5 y 6, el descenso de gradiente siempre sería 0 por lo que nunca podría converger a un valor distinto del valor inicial (una única iteración). Por ello, se emplean valores iniciales aleatorios pequeños (intentando que sean distintos).

3.6. Predecir ratings y generar recomendaciones

Para predecir la calificación de la película i para el usuario j , se calcula:

$$(\theta^{(j)})^T x^{(i)} + \mu_i \quad (7)$$

A diferencia de la predicción especificada en la sección 2, al sumarle la media se obtienen calificaciones en un ratio del 1 al 10, en lugar del original del 1 al 5.

4. Resultados

Top recommendations for you:

Predicting rating 7.6 for movie Titanic (1997)
Predicting rating 7.6 for movie Schindler's List (1993)
Predicting rating 7.5 for movie Shawshank Redemption, The (1994)
Predicting rating 7.5 for movie Good Will Hunting (1997)
Predicting rating 7.4 for movie Star Wars (1977)
Predicting rating 7.2 for movie Raiders of the Lost Ark (1981)
Predicting rating 7.2 for movie Braveheart (1995)
Predicting rating 7.2 for movie Godfather, The (1972)
Predicting rating 7.2 for movie Usual Suspects, The (1995)
Predicting rating 7.1 for movie Silence of the Lambs, The (1991)

Original ratings provided:

Rated 4 for Toy Story (1995)
Rated 2 for Star Wars (1977)
Rated 5 for Forrest Gump (1994)
Rated 4 for Jurassic Park (1993)
Rated 3 for Home Alone (1990)
Rated 2 for Aladdin (1992)
Rated 4 for Willy Wonka and the Chocolate Factory (1971)
Rated 4 for Titanic (1997)
Rated 4 for Schindler's List (1993)
Rated 2 for Batman (1989)
Rated 2 for Pinocchio (1940)
Rated 2 for Cinderella (1950)
Rated 2 for Mary Poppins (1964)
Rated 2 for Alice in Wonderland (1951)

5. Conclusiones

Dado que las calificaciones introducidas corresponden con los gustos del autor de esta práctica, se ha procedido a comparar los resultados de la predicción

obtenida con las predicciones que realiza Netflix para la misma persona. Esto se ha realizado para las películas que no habían sido vistas previamente, y que aparecían en el catálogo de la plataforma Netflix:

- Para la película «Cadena perpetua» con un 7,5; Netflix la recomienda en un 83 %.
- Para la película «Los sospechosos de siempre» con un 7,2; Netflix la recomienda en un 69 %.