MAI4CAREU
Master programmes in Artificial Intelligence 4 Careers in Europe

Πανεπιστήμιο Κύπρου - Τεχνητή Νοημοσύνη

MAI612 - ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ

Διάλεξη 15: Recommender Systems

Βασίλης Βασιλειάδης, PhD
Χειμερινό Εξάμηνο 2022/23

CYENS
CENTRE OF EXCELLENCE

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Διάλεξη 15: Recommender Systems

Μαθησιακά αποτελέσματα

Θα κατανοήσετε:

1. Τι είναι τα recommender systems και το πρόβλημα που λύνουν
2. Ο αλγόριθμος Collaborative Filtering
3. Πώς να επεκτείνετε τον αλγόριθμο στην περίπτωση των δυαδικών ετικετών
4. Λεπτομέρειες εφαρμογής που θα κάνουν τον αλγόριθμο να λειτουργήσει καλύτερα στην πράξη
5. Φιλτράρισμα με βάση το περιεχόμενο
6. Πώς να βρείτε τα σχετικά στοιχεία μεταξύ ενός μεγάλου συνόλου

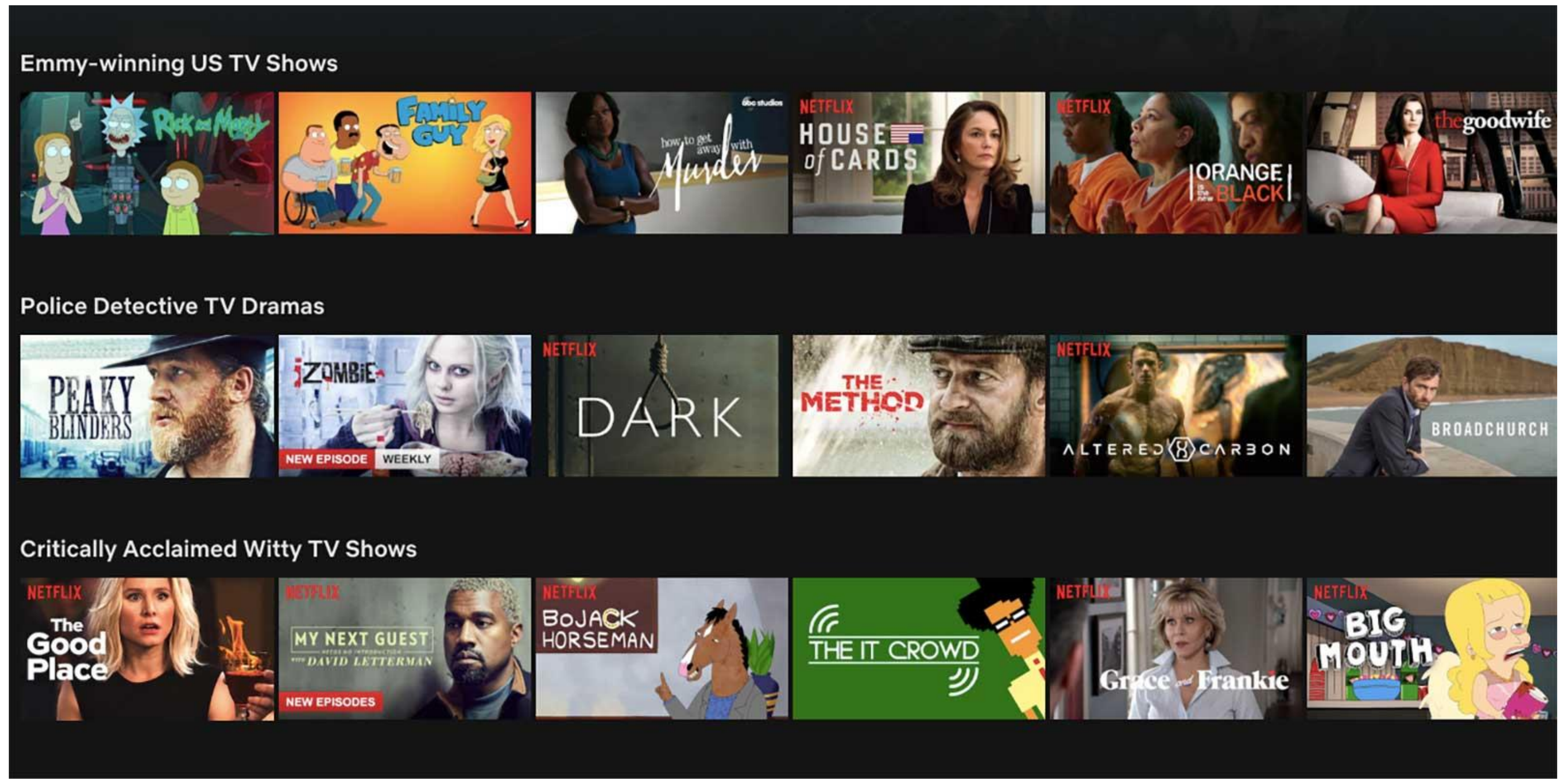
Co-financed by the European Union
Connecting Europe Facility

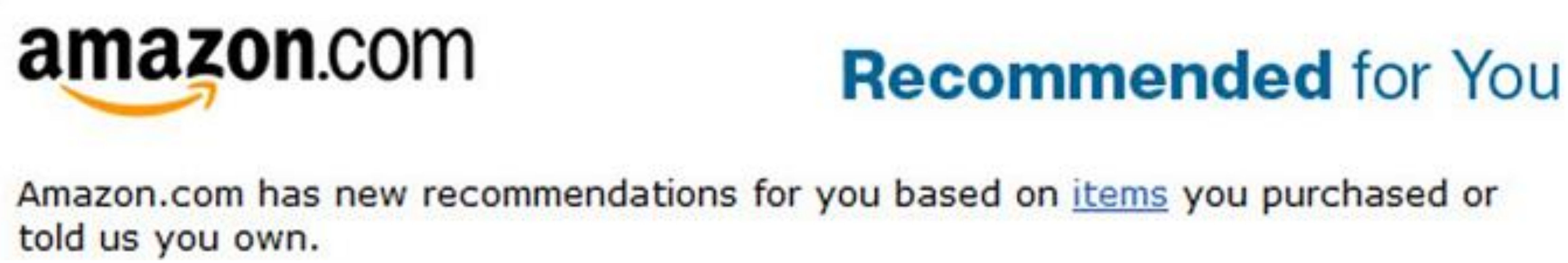
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Recommendations

Πολλές εταιρείες προωθούν την αξία τους κατανοώντας τις προτιμήσεις των χρηστών και παρέχοντας συστάσεις.

Οι συστάσεις μπορεί να επηρεάσουν τις ειδήσεις που διαβάζουμε, τα προϊόντα που αγοράζουμε και την ψυχαγωγία που καταναλώνουμε.





Recommender systems

Ένα Recommender system είναι ένα σύστημα που παρέχει προτάσεις για στοιχεία που είναι πιο σχετικά με ένα συγκεκριμένο χρήστη

Ιδιαίτερα χρήσιμο όταν ένας χρήστης πρέπει να επιλέξει ένα στοιχείο από έναν μεγάλο κατάλογο στοιχείων που μπορεί να προσφέρει μια υπηρεσία

Για παράδειγμα:

- Δημιουργία λίστας αναπαραγωγής για βίντεο και υπηρεσίες μουσικής
- Συστάσεις προϊόντων για ηλεκτρονικά καταστήματα
- Συστάσεις περιεχομένου για πλατφόρμες μέσων κοινωνικής δικτύωσης
- Ταξιδιωτικές συστάσεις
- Αποτελέσματα αναζήτησης
- ...

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Πρόβλεψη αξιολογήσεων ταινιών

Ο χρήστης βαθμολογεί ταινίες χρησιμοποιώντας 0 έως 5 αστέρια

Movie	Kate (1)	John (2)	Annie (3)	Steve (4)
Alone in the forest	5	5	0	0
Afraid of the dark	5	?	?	0
The black cave	?	4	0	?
Billy's night off	0	0	5	4
The animal whisperer	0	0	5	?

★				
★	★			
★	★	★		
★	★	★	★	
★	★	★	★	★

n_u = Αριθ. χρηστών
n_m = Αριθ. των ταινιών
$r(i,j) = 1$ εάν ο χρήστης j έχει βαθμολογήσει την ταινία i

$n_u = 4$
$n_m = 5$

$r(1,1) = 1$
$r(3,1) = 0$

$y^{(i,j)} = 4$

$y^{(i,j)}$ = βαθμολογία που δίνεται από τον χρήστη j στην ταινία i (καθορίζεται μόνο εάν $r(i,j) = 1$)

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Κι αν έχουμε χαρακτηριστικά των ταινιών;

$n_u = 4$
$n_m = 5$
$n = 2$

Movie	Kate (1)	John (2)	Annie (3)	Steve (4)	x_1 (θρίλερ)	x_2 (κωμωδία)
Alone in the forest	5	5	0	0	0.9	0
Afraid of the dark	5	?	?	0	1.0	0.01
The black cave	?	4	0	?	0.99	0
Billy's night off	0	0	5	4	0.1	1.0
The animal whisperer	0	0	5	?	0	0.9

$$\boldsymbol{x}^{(1)} = \begin{bmatrix} 1 \\ 0.9 \\ 0 \end{bmatrix}$$

$$\boldsymbol{x}^{(4)} = \begin{bmatrix} 1 \\ 0.1 \\ 1.0 \end{bmatrix}$$

Για τον χρήστη 1: Πρόβλεψη βαθμολογίας για την ταινία i ως: $\boldsymbol{\theta} \cdot \boldsymbol{x}^{(i)}$ ← γραμμική παλινδρόμηση

$\boldsymbol{\theta}^{(1)} = [0, 5, 0]^T$ $\quad \boldsymbol{x}^{(3)} = [1, 0.99, 0]^T$ $\quad \boldsymbol{\theta}^{(1)} \cdot \boldsymbol{x}^{(3)} = 4.95$

Διαφορετικό μοντέλο γραμμικής παλινδρόμησης για κάθε χρήστη

Για τον χρήστη j: Πρόβλεψη της βαθμολογίας j του χρήστη για ταινία i ως: $\boldsymbol{\theta}^{(j)} \cdot \boldsymbol{x}^{(i)}$

Συνάρτηση κόστους

Συμβολισμός:

$r(i,j) = 1$ εάν ο χρήστης j έχει βαθμολογήσει την ταινία i (0 διαφορετικά)
$y^{(i,j)}$ = βαθμολογία που δίνεται από τον χρήστη j στην ταινία i (εάν ορίζεται)
$\boldsymbol{\theta}^{(j)}$ =παράμετροι για το χρήστη j $\boldsymbol{\theta}^{(j)} \in \mathbb{R}^{n+1}$
$\boldsymbol{x}^{(i)}$ = χαρακτηριστικό διάνυσμα για την ταινία i

Για τον χρήστη j και την ταινία i, πρόβλεψη βαθμολογίας: $f_{\boldsymbol{\theta}^{(j)}}\left(\boldsymbol{x}^{(i)}\right) = \boldsymbol{\theta}^{(j)} \cdot \boldsymbol{x}^{(i)}$
$m^{(j)}$ = Αριθ. ταινιών που βαθμολογούνται από τον χρήστη j
Μάθε το $\boldsymbol{\theta}^{(j)}$

$m^{(j)}$ είναι απλά μια σταθερή, οπότε αν την αφαιρέσουμε θα καταλήξουμε στην ίδια λύση για $\boldsymbol{\theta}^{(j)}$

αριθμός χαρακτηριστικών

$$\min_{\boldsymbol{\theta}^{(j)}} L\left(\boldsymbol{\theta}^{(j)}\right) = \frac{1}{2\cancel{m^{(j)}}} \sum_{i:r(i,j)=1} \left(f_{\boldsymbol{\theta}^{(j)}}\left(\boldsymbol{x}^{(i)}\right) - y^{(i,j)}\right)^2 + \frac{\lambda}{2\cancel{m^{(j)}}} \sum_{k=1}^{n} \left(\theta_k^{(j)}\right)^2$$

Διαφάνειες βασισμένες στο Andrew Ng's Machine Learning course - Coursera

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Συνάρτηση κόστους

Για να μάθετε το $\boldsymbol{\theta}^{(j)}$(παράμετροι για το χρήστη j):

$$L(\boldsymbol{\theta}^{(j)}) = \frac{1}{2}\sum_{i:r(i,j)=1}\left(f_{\boldsymbol{\theta}^{(j)}}(\boldsymbol{x}^{(i)}) - y^{(i,j)}\right)^2 + \frac{\lambda}{2}\sum_{k=1}^{n}\left(\theta_k^{(j)}\right)^2$$

Μάθετε τις παραμέτρους $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \dots, \boldsymbol{\theta}^{(n_u)}$για όλους τους χρήστες:

$$L(\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \dots, \boldsymbol{\theta}^{(n_u)}) = \frac{1}{2}\sum_{j=1}^{n_u}\sum_{i:r(i,j)=1}\left(f_{\boldsymbol{\theta}^{(j)}}(\boldsymbol{x}^{(i)}) - y^{(i,j)}\right)^2 + \frac{\lambda}{2}\sum_{j=1}^{n_u}\sum_{k=1}^{n}\left(\theta_k^{(j)}\right)^2$$

Αλγόριθμος βελτιστοποίησης

$$\min_{\boldsymbol{\theta}^{(1)},\ldots,\boldsymbol{\theta}^{(n_u)}} \underbrace{\frac{1}{2}\sum_{j=1}^{n_u}\sum_{i:r(i,j)=1}\left(f_{\boldsymbol{\theta}^{(j)}}\left(\boldsymbol{x}^{(i)}\right)-y^{(i,j)}\right)^2+\frac{\lambda}{2}\sum_{j=1}^{n_u}\sum_{k=1}^{n}\left(\theta_k^{(j)}\right)^2}_{L\left(\boldsymbol{\theta}^{(1)},\ldots,\boldsymbol{\theta}^{(n_u)}\right)}$$

Ενημέρωση gradient descent:

$$\theta_k^{(j)} = \theta_k^{(j)} - a\left(\sum_{i:r(i,j)=1}\left(f_{\boldsymbol{\theta}^{(j)}}\left(\boldsymbol{x}^{(i)}\right)-y^{(i,j)}\right)\right) \qquad (\text{για } k=0)$$

$$\theta_k^{(j)} = \theta_k^{(j)} - a\underbrace{\left(\sum_{i:r(i,j)=1}\left(f_{\boldsymbol{\theta}^{(j)}}\left(\boldsymbol{x}^{(i)}\right)-y^{(i,j)}\right)x_k^{(i)}+\lambda\theta_k^{(j)}\right)}_{\frac{\partial}{\partial\theta_k^{(j)}}L\left(\boldsymbol{\theta}^{(1)},\ldots,\boldsymbol{\theta}^{(n_u)}\right)} \qquad (\text{για } k\neq 0)$$

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Collaborative Filtering

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Κι αν δεν ξέρουμε τα χαρακτηριστικά;

Movie	Kate (1)	John (2)	Annie (3)	Steve (4)	x_1 ~~(thriller)~~	x_2 ~~(comedy)~~
Alone in the forest	5	5	0	0	?	?
Afraid of the dark	5	?	?	0	?	?
The black cave	?	4	0	?	?	?
Billy’s night off	0	0	5	4	?	?
The animal whisperer	0	0	5	?	?	?

Υποθέστε ότι με κάποιο τρόπο μάθαμε τις παραμέτρους για τους χρήστες

Μπορούμε να υπολογίσουμε τις τιμές για τα χαρακτηριστικά από τα δεδομένα;

$$\boldsymbol{\theta}^{(1)} = \begin{bmatrix}0\\5\\0\end{bmatrix} \quad \boldsymbol{\theta}^{(2)} = \begin{bmatrix}0\\5\\0\end{bmatrix} \quad \boldsymbol{\theta}^{(3)} = \begin{bmatrix}0\\0\\5\end{bmatrix} \quad \boldsymbol{\theta}^{(4)} = \begin{bmatrix}0\\0\\5\end{bmatrix}$$

Χρήση $\boldsymbol{\theta}^{(j)} \cdot \boldsymbol{x}^{(i)}$

$$\left.\begin{aligned}\boldsymbol{\theta}^{(1)} \cdot \boldsymbol{x}^{(1)} &\approx 5\\ \boldsymbol{\theta}^{(2)} \cdot \boldsymbol{x}^{(1)} &\approx 5\\ \boldsymbol{\theta}^{(3)} \cdot \boldsymbol{x}^{(1)} &\approx 0\\ \boldsymbol{\theta}^{(4)} \cdot \boldsymbol{x}^{(1)} &\approx 0\end{aligned}\right\} \quad \boldsymbol{x}^{(1)} = \begin{bmatrix}1\\1.0\\0.0\end{bmatrix}$$

Διαφάνειες βασισμένες στο Andrew Ng’s Machine Learning course - Coursera

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Συνάρτηση κόστους

Δεδομένου $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots, \boldsymbol{\theta}^{(n_u)}$

Μάθε $\boldsymbol{x}^{(i)}$:

$$\min_{\boldsymbol{x}^{(i)}} \quad L\left(\boldsymbol{x}^{(i)}\right) = \frac{1}{2} \sum_{j:r(i,j)=1} \left(\boldsymbol{\theta}^{(j)} \cdot \boldsymbol{x}^{(i)} - y^{(i,j)}\right)^2 + \frac{\lambda}{2} \sum_{k=1}^{n} \left(x_k^{(i)}\right)^2$$

Μάθε $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(n_m)}$:

$$\min_{\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n_m)}} L\left(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(n_m)}\right) = \frac{1}{2} \sum_{i=1}^{n_m} \sum_{j:r(i,j)=1} \left(\boldsymbol{\theta}^{(j)} \cdot \boldsymbol{x}^{(i)} - y^{(i,j)}\right)^2 + \frac{\lambda}{2} \sum_{i=1}^{n_m} \sum_{k=1}^{n} \left(x_k^{(i)}\right)^2$$

Collaborative filtering

Επανέλαβε:

Δεδομένου $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \dots, \boldsymbol{x}^{(n_m)}$ και κινηματογραφικές βαθμολογίες $(r(i,j), y^{(i,j)}($

Εκτίμηση $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \dots, \boldsymbol{\theta}^{(n_u)}$

Δεδομένου $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \dots, \boldsymbol{\theta}^{(n_u)}$

Εκτίμηση $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \dots, \boldsymbol{x}^{(n_m)}$

Μάντεψε $\boldsymbol{\theta} \rightarrow \boldsymbol{x} \rightarrow \boldsymbol{\theta} \rightarrow \boldsymbol{x} \rightarrow \boldsymbol{\theta} \rightarrow \boldsymbol{x} \rightarrow \boldsymbol{\theta} \rightarrow \boldsymbol{x} \rightarrow \cdots$

Αυτό λειτουργεί, αλλά αντί να πηγαίνει μπρος-πίσω υπάρχει ένας πιο αποτελεσματικός τρόπος για να λύσει θ και x **ταυτόχρονα**

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Collaborative filtering

Συνάρτηση κόστους για να μάθουν $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots, \boldsymbol{\theta}^{(n_u)}$:

$$L\left(\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(n_u)}\right) = \frac{1}{2}\sum_{j=1}^{n_u}\sum_{i:r(i,j)=1}\left(\boldsymbol{\theta}^{(j)} \cdot \boldsymbol{x}^{(i)} - y^{(i,j)}\right)^2 + \frac{\lambda}{2}\sum_{j=1}^{n_u}\sum_{k=1}^{n}\left(\theta_k^{(j)}\right)^2$$

$\cancel{x_0 = 1}$

Λειτουργία κόστους για να μάθουν $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(n_m)}$:

$\boldsymbol{x} \in \mathbb{R}^n$

$\boldsymbol{\theta} \in \mathbb{R}^n$

$$L\left(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n_m)}\right) = \frac{1}{2}\sum_{i=1}^{n_m}\sum_{j:r(i,j)=1}\left(\boldsymbol{\theta}^{(j)} \cdot \boldsymbol{x}^{(i)} - y^{(i,j)}\right)^2 + \frac{\lambda}{2}\sum_{i=1}^{n_m}\sum_{k=1}^{n}\left(x_k^{(i)}\right)^2$$

Συνδυάστε τα (ελαχιστοποιήστε $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n_m)}$ και $\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(n_u)}$ ταυτόχρονα):

$$L\begin{pmatrix}\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n_m)}, \\ \boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(n_u)}\end{pmatrix} = \frac{1}{2}\sum_{(i,j):r(i,j)=1}\left(\boldsymbol{\theta}^{(j)} \cdot \boldsymbol{x}^{(i)} - y^{(i,j)}\right)^2 + \frac{\lambda}{2}\sum_{j=1}^{n_u}\sum_{k=1}^{n}\left(\theta_k^{(j)}\right)^2 + \frac{\lambda}{2}\sum_{i=1}^{n_m}\sum_{k=1}^{n}\left(x_k^{(i)}\right)^2$$

Αλγόριθμος collaborative filtering

1. Αρχικοποίηση $\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(n_m)}, \boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(n_u)}$ σε μικρές τυχαίες τιμές
2. Ελαχιστοποίηση $L\left(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(n_m)}, \boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(n_u)}\right)$ του gradient descemt

   Επανέλαβε:

   Για κάθε $i = 1, \dots, n_m, and\ j = 1, \dots, n_u$:

$$x_k^{(i)} = x_k^{(j)} - a \frac{\partial}{\partial x_k^{(i)}} L\left(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(n_m)}, \boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(n_u)}\right)$$

$$\theta_k^{(j)} = \theta_k^{(j)} - a \frac{\partial}{\partial \theta_k^{(j)}} L\left(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(n_m)}, \boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(n_u)}\right)$$

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Αλγόριθμος collaborative filtering

1. Αρχικοποίηση $\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(n_m)}, \boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(n_u)}$ σε μικρές τυχαίες τιμές
2. Ελαχιστοποίηση $L\left(\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(n_m)}, \boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(n_u)}\right)$ του gradient descent

   Επανέλαβε:

   Για κάθε $i = 1, \dots, n_m, and\ j = 1, \dots, n_u$:

$$x_k^{(i)} = x_k^{(i)} - a\left(\sum_{j:r(i,j)=1} \left(\boldsymbol{\theta}^{(j)} \cdot \boldsymbol{x}^{(i)} - y^{(i,j)}\right)\theta_k^{(j)} + \lambda x_k^{(i)}\right)$$

$$\theta_k^{(j)} = \theta_k^{(j)} - a\left(\sum_{i:r(i,j)=1} \left(\boldsymbol{\theta}^{(j)} \cdot \boldsymbol{x}^{(i)} - y^{(i,j)}\right)x_k^{(i)} + \lambda \theta_k^{(j)}\right)$$

3. Για έναν χρήστη με παραμέτρους $\boldsymbol{\theta}$ και μια ταινία με (μαθημένα) χαρακτηριστικά $\boldsymbol{x}$, προβλέψτε μια βαθμολογία $\boldsymbol{\theta}^T\boldsymbol{x}$

Κουίζ

Γιατί θεωρούμε το συνεργατικό φιλτράρισμα ως έναν μη εποπτευόμενο αλγόριθμο μάθησης;

Επειδή τα χαρακτηριστικά δεν παρέχονται από έναν δάσκαλο.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Δυαδικές ετικέτες

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Δυαδικές ετικέτες

Movie	Kate (1)	John (2)	Annie (3)	Steve (4)
Alone in the forest	1	1	0	0
Afraid of the dark	1	?	?	0
The black cave	?	1	0	?
Billy’s night off	0	0	1	1
The animal whisperer	0	0	1	?

- Πώς μπορούμε να επεκτείνουμε τον αλγόριθμο collaborative filtering για να δουλέψουμε για δυαδικές ετικέτες;
- Θέλουμε να προβλέψουμε αν στους χρήστες θα αρέσει ένα συγκεκριμένο στοιχείο που δεν έχουν ακόμη βαθμολογήσει.
- Θα θέλαμε να εκτιμήσουμε την πιθανότητα να αρέσει ένα προϊόν, προκειμένου να αποφασίσουμε πόσο θα προτείνουμε αυτά τα στοιχεία.

Διαφάνειες βασισμένες στο Andrew Ng’s Machine Learning course - Coursera

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Παραδειγματικές εφαρμογές

1. Ηλεκτρονικός ιστότοπος αγορών
   - Ο χρήστης j αγόρασε ένα προϊόν μετά την εμφάνιση του;
2. Τα social media
   - Άρεσε στον j χρήστη ένα αντικείμενο;
3. Online διαφήμιση
   - Ο χρήστης j έκανε κλικ σε ένα στοιχείο;
4. Χρησιμοποιήστε τη συμπεριφορά του χρήστη, αντί να ζητάτε ρητή αξιολόγηση
   - Ο χρήστης j ξόδεψε τουλάχιστον 30 δευτερόλεπτα με ένα στοιχείο;

Έννοια των αξιολογήσεων:

1: απασχολημένος μετά την εμφάνιση του στοιχείου

0: δεν συμμετείχε μετά την εμφάνιση του στοιχείου

? το στοιχείο δεν εμφανίζεται ακόμα

Από την παλινδρόμηση στη δυαδική ταξινόμηση

Στο παρελθόν:

Πρόβλεψη $y^{(i,j)}$ ως $\boldsymbol{\theta}^{(j)} \cdot \boldsymbol{x}^{(i)}$

Για τις δυαδικές ετικέτες:

Πρόβλεψε ότι η πιθανότητα $y^{(i,j)} = 1$
δίνεται από $g\left(\boldsymbol{\theta}^{(j)} \cdot \boldsymbol{x}^{(i)}\right)$

Μοντέλο logistic regression

που $g(z) = \frac{1}{1+e^{-z}}$

Λειτουργία κόστους για δυαδική εφαρμογή

Ανάγκη τροποποίησης της συνάρτησης κόστους από το τετραγωνικό σφάλμα στο δυαδικό σφάλμα διασταυρούμενης εντροπίας

Προηγούμενη συνάρτηση κόστους:

$$L\begin{pmatrix}\boldsymbol{x}^{(1)},\dots,\boldsymbol{x}^{(n_m)},\\ \boldsymbol{\theta}^{(1)},\dots,\boldsymbol{\theta}^{(n_u)}\end{pmatrix}=\frac{1}{2}\sum_{(i,j):r(i,j)=1}\Big(\overbrace{\boldsymbol{\theta}^{(j)}\cdot\boldsymbol{x}^{(i)}}^{\hat{y}^{(i,j)}}-y^{(i,j)}\Big)^2+\frac{\lambda}{2}\sum_{j=1}^{n_u}\sum_{k=1}^{n}\left(\theta_k^{(j)}\right)^2+\frac{\lambda}{2}\sum_{i=1}^{n_m}\sum_{k=1}^{n}\left(x_k^{(i)}\right)^2$$

$$\hat{y}^{(i,j)}=g\left(\boldsymbol{\theta}^{(j)}\cdot\boldsymbol{x}^{(i)}\right)$$

Απώλεια για δυαδικές ετικέτες:

$$loss\left(\hat{y}^{(i,j)},y^{(i,j)}\right)=-y^{(i,j)}\log\left(\hat{y}^{(i,j)}\right)-\left(1-y^{(i,j)}\right)\log\left(1-\hat{y}^{(i,j)}\right)$$

← Απώλεια για ένα παράδειγμα

$$L\begin{pmatrix}\boldsymbol{x}^{(1)},\dots,\boldsymbol{x}^{(n_m)},\\ \boldsymbol{\theta}^{(1)},\dots,\boldsymbol{\theta}^{(n_u)}\end{pmatrix}=\sum_{(i,j):r(i,j)=1}loss\left(\hat{y}^{(i,j)},y^{(i,j)}\right)+\frac{\lambda}{2}\sum_{j=1}^{n_u}\sum_{k=1}^{n}\left(\theta_k^{(j)}\right)^2+\frac{\lambda}{2}\sum_{i=1}^{n_m}\sum_{k=1}^{n}\left(x_k^{(i)}\right)^2$$

Λεπτομέρειες υλοποίησης

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Μέση κανονικοποίηση

Movie	Kate (1)	John (2)	Annie (3)	Steve (4)	Emily (5)
Alone in the forest	5	5	0	0	?
Afraid of the dark	5	?	?	0	?
The black cave	?	4	0	?	?
Billy’s night off	0	0	5	4	?
The animal whisperer	0	0	5	0	?

- Ας υποθέσουμε ότι ένας νέος χρήστης έρχεται που δεν έχει βαθμολογήσει καμία ταινία
- Ο στόχος βελτιστοποίησης πιθανότατα προβλέπει βαθμολογία 0 για όλες τις ταινίες, επειδή οι αξιολογήσεις του νέου χρήστη δεν έχουν καμία συμβολή στη λειτουργία τετραγωνικού σφάλματος.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Μέση κανονικοποίηση

$$\text{Βαθμολόγηση} = \begin{bmatrix} 5 & ? & ? & 0 & ? \\ ? & 4 & 0 & ? & ? \\ 0 & 0 & 5 & 4 & ? \\ 0 & 0 & 5 & 0 & ? \end{bmatrix} \qquad \mu = \begin{bmatrix} 2.5 \\ 2 \\ 2.25 \\ 1.25 \end{bmatrix}$$

$$\text{Κανονικοποιημένη Αξιολόγιση} = \begin{bmatrix} 2.5 & 2.5 & -2.5 & -2.5 & ? \\ 2.5 & ? & ? & -2.5 & ? \\ ? & 2 & -2 & ? & ? \\ -2.25 & -2.25 & 2.75 & 1.75 & ? \\ -1.25 & -1.25 & 3.75 & -1.25 & ? \end{bmatrix}$$

Αξιολογίσεις - μ

Για τους χρήστες j στην ταινία i προβλέπουν: $\boldsymbol{\theta}^{(j)} \cdot \boldsymbol{x}^{(i)} + \mu_i$

Οι προβλεπόμενες αξιολογήσεις του νέου χρήστη θα είναι ο μέσος όρος των αξιολογήσεων

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Μέση κανονικοποίηση

- Η ομαλοποίηση των σειρών είναι η πιο συνηθισμένη περίπτωση για αυτή την εφαρμογή: όταν ένας νέος χρήστης φτάνει, να του παρέχει λογικές βαθμολογίες
- Θα μπορούσαμε να ομαλοποιήσουμε τις στήλες, το οποίο είναι αυτό που θα κάναμε όταν μια νέα ταινία φτάσει χωρίς αξιολογήσεις
- Ωστόσο, τέτοιες ταινίες δεν θα πρέπει να συνιστώνται αρχικά σε πολλούς χρήστες, καθώς δεν γνωρίζουμε πολλά γι' αυτούς, οπότε είναι πιο λογικό να εξομαλύνουμε τις σειρές.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Διανυσματική εφαρμογή των προβλεπόμενων αξιολογήσεων

$$Y = \begin{bmatrix} 5 & ? & ? & 0 & ? \\ ? & 4 & 0 & ? & ? \\ 0 & 0 & 5 & 4 & ? \\ 0 & 0 & 5 & 0 & ? \end{bmatrix} \epsilon\ \mathbb{R}^{n_m \times n_u}$$

Προβλεπόμενες βαθμολογίες: $\begin{bmatrix} \boldsymbol{\theta}^{(1)} \cdot \boldsymbol{x}^{(1)} & \dots & \boldsymbol{\theta}^{(n_u)} \cdot \boldsymbol{x}^{(1)} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\theta}^{(1)} \cdot \boldsymbol{x}^{(n_m)} & \dots & \boldsymbol{\theta}^{(n_u)} \cdot \boldsymbol{x}^{(n_m)} \end{bmatrix}$

$$X\Theta\ \epsilon\ \mathbb{R}^{n_m \times n_u}$$

$$\boldsymbol{x}^{(i)} \epsilon\ \mathbb{R}^{n \times 1}$$

$$\boldsymbol{\theta}^{(j)} \epsilon\ \mathbb{R}^{n \times 1}$$

$$X = \begin{bmatrix} \left(\boldsymbol{x}^{(1)}\right)^T \\ \vdots \\ \left(\boldsymbol{x}^{(n_m)}\right)^T \end{bmatrix} \epsilon\ \mathbb{R}^{n_m \times n}$$

$$\Theta = \left[\boldsymbol{\theta}^{(1)} \quad \dots \quad \boldsymbol{\theta}^{(n_u)}\right] \epsilon\ \mathbb{R}^{n \times n_u}$$

Εύρεση σχετικών στοιχείων

- Κατά τη χρήση συνεργατικού φιλτραρίσματος, τα χαρακτηριστικά $\boldsymbol{x}^{(i)}$ του στοιχείου i είναι δύσκολο να ερμηνευτούν
  - Στο παράδειγμα, είχαμε χαρακτηριστικά όπως αν μια ταινία είναι θρίλερ ή κωμωδία, ωστόσο, δεν μπορούμε εύκολα να δούμε μεμονωμένα χαρακτηριστικά, π.χ., x_1 ή x_2 να καταλάβουμε τι είναι αυτά.
  - Παρ' όλα αυτά, συλλογικά τα χαρακτηριστικά μεταδίδουν κάποιο νόημα σχετικά με το στοιχείο

- Για να βρούμε άλλα στοιχεία που σχετίζονται με ένα στοιχείο ενδιαφέροντος, πρέπει να ταξινομήσουμε τα στοιχεία με βάση την απόστασή του $\boldsymbol{x}^{(i)}$.
  - Αυτό επιτυγχάνεται χρησιμοποιώντας τους k - κοντινότερους γείτονες του $\boldsymbol{x}^{(i)}$,όπου k δείχνει τον αριθμό των στοιχείων παρόμοια με αυτά που $\boldsymbol{x}^{(i)}$ θέλουμε να ανακτήσουμε

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Φιλτράρισμα με βάση το περιεχόμενο

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Collaborative filtering vs φιλτραρίσματος βάσει περιεχομένου

Collaborative filtering:

- Προτείνετε στοιχεία σε εσάς με βάση τις αξιολογήσεις των χρηστών που έδωσαν παρόμοιες αξιολογήσεις με εσάς

Φιλτράρισμα με βάση το περιεχόμενο:

- Προτείνετε στοιχεία σε σας με βάση τα χαρακτηριστικά του χρήστη και του στοιχείου για να βρείτε καλό ταίριασμα
- Μένουν κάποιες βαθμολογίες

Διαφάνειες βασισμένες στο Andrew Ng's Machine Learning course - Coursera

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Παραδείγματα χαρακτηριστικών χρήστη και στοιχείων

Χαρακτηριστικά του χρήστη:
- Ηλικία
- Φύλο
- Χώρα
- Ταινίες που παρακολούθησαν
- Μέση βαθμολογία ανά είδος
- ...

Χαρακτηριστικά ταινιών:
- Έτος
- Είδος(α)
- Κριτικές
- Μέση βαθμολογία
- ...

$\boldsymbol{x}_u^{(j)}$ για το χρήστη j

$\boldsymbol{x}_m^{(i)}$ για την ταινία

Το διανυσματικό μέγεθος θα μπορούσε να είναι διαφορετικό

Διαφάνειες βασισμένες στο Andrew Ng's Machine Learning course - Coursera

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Φιλτράρισμα με βάση το περιεχόμενο: μαθαίνοντας να ταιριάζετε

Υπολογισμός ενσωματώσεων από διανυσματικά χαρακτηριστικά:

$$\boldsymbol{v}_u^{(j)} = f_u\left(\boldsymbol{x}_u^{(j)}\right) \qquad \boldsymbol{v}_u^{(j)} \in \mathbb{R}^n$$

$$\boldsymbol{v}_m^{(i)} = f_m\left(\boldsymbol{x}_m^{(i)}\right) \qquad \boldsymbol{v}_m^{(i)} \in \mathbb{R}^n$$

Πρέπει να είναι του ίδιου μεγέθους
π.χ., $\boldsymbol{n} = 32$

Πρόβλεψη βαθμολογίας του χρήστη $\boldsymbol{j}$ στην ταινία $\boldsymbol{i}$ ως:

$$\boldsymbol{v}_u^{(j)} \cdot \boldsymbol{v}_m^{(i)}$$

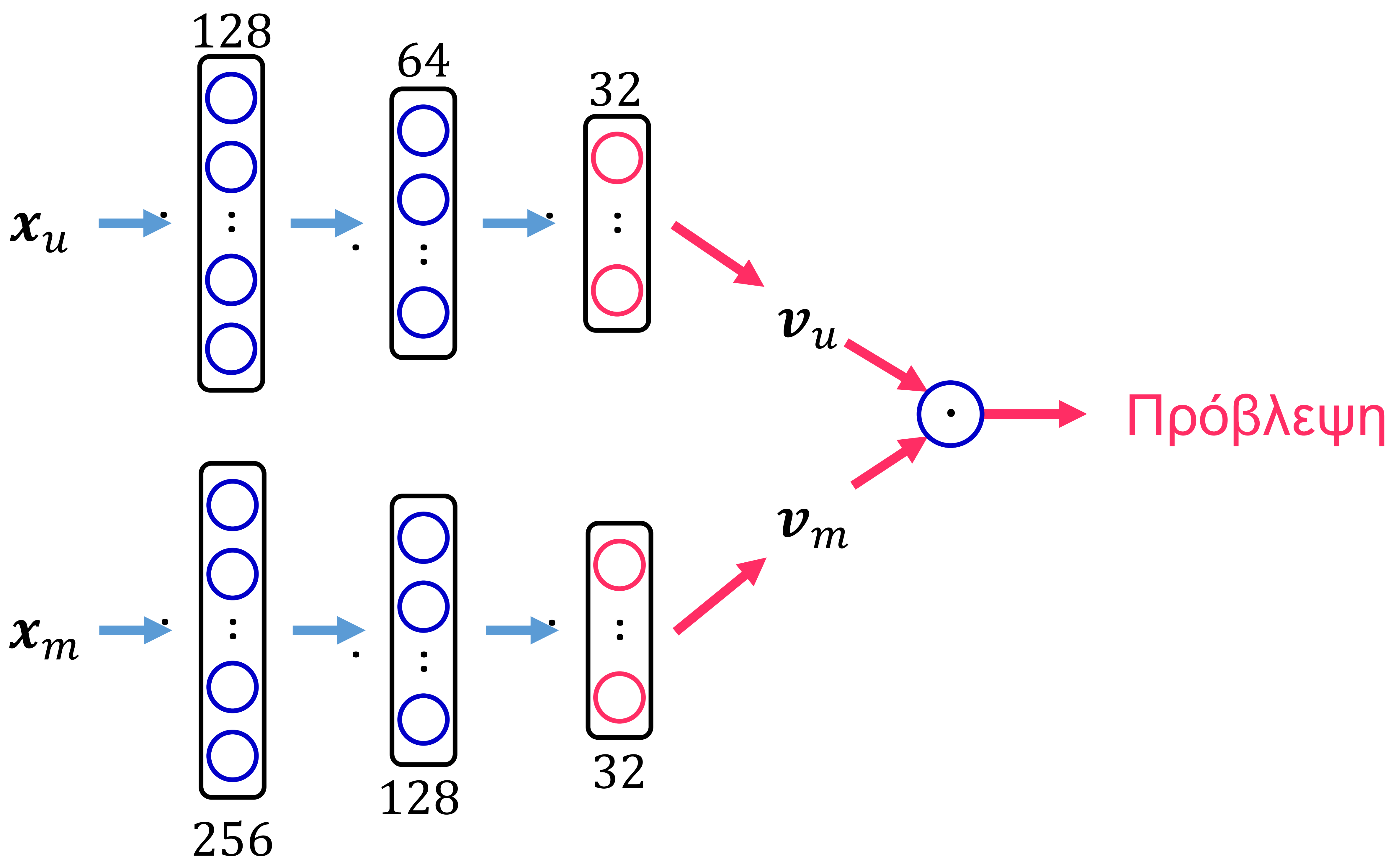
Το πρόβλημα:
Πώς να καταλήξετε σε μια κατάλληλη επιλογή των ενσωματώσεων, ώστε το **εσωτερικό γινόμενο** τους να είναι μια καλή πρόβλεψη της βαθμολογίας που j δίνει ο χρήστης στην ταινία i

Διαφάνειες βασισμένες στο Andrew Ng's Machine Learning course - Coursera

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Βαθιά μάθηση για φιλτράρισμα με βάση το περιεχόμενο



Λειτουργία κόστους

$$L(\dots) = \sum_{(i,j):r(i,j)=1} \left(\boldsymbol{v}_u^{(j)} \cdot \boldsymbol{v}_m^{(i)} - y^{(i,j)} \right)^2$$

+ όρος κανονικοποίησης ΝΔ

Εκπαίδευση

Χρήση του backprop και gradient descent

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Εύρεση σχετικών στοιχείων

Για να βρείτε ταινίες που σχετίζονται i με την ταινία, βρείτε τους πιο κοντινούς γείτονες $\boldsymbol{v}_m^{(i)}$

- Αυτό μπορεί να προ-υπολογιστεί

Πώς να βρείτε αποτελεσματικά συστάσεις από ένα μεγάλο σύνολο αντικειμένων (π.χ., εκατομμύρια);

1. Ανάκτηση:
    - Δημιουργία μεγάλου και ευρέος καταλόγου εύλογων υποψήφιων στοιχείων
        - 10 πιο παρόμοιες ταινίες με τις τελευταίες 10 ταινίες που παρακολουθεί ο χρήστης
        - Οι 10 κορυφαίες ταινίες για τις περισσότερες προβολές 3 είδη από το χρήστη
        - Οι 20 κορυφαίες ταινίες στη χώρα
    - Συνδυάστε ανακτημένα στοιχεία στη λίστα, αφαιρέστε διπλότυπα και στοιχεία που έχουν ήδη παρακολουθηθεί

2. Ταξινόμηση:
    - Λήψη λίστας που ανακτήθηκε και κατάταξη χρησιμοποιώντας το μορφωμένο μοντέλο $(\boldsymbol{v}_u^{(j)} \cdot \boldsymbol{v}_m^{(i)})$
    - Εμφάνιση κατατασσόμενων στοιχείων στον χρήστη (π.χ., top 10)

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Οι συστάσεις ως διαδοχικό πρόβλημα λήψης αποφάσεων

- Αυτό που αναφέραμε μέχρι στιγμής δεν εξετάζει ρητά τη διαδοχική φύση της αλληλεπίδρασης των χρηστών με ένα σύστημα
  - Αυτό γίνεται συνήθως με τη δημιουργία χαρακτηριστικών που προσπαθούν να καταγράψουν αυτή την αλληλεπίδραση (π.χ., τελευταία ταινία που παρακολουθήθηκε)

- Ωστόσο, αν λάβουμε ρητά υπόψη αυτή τη διαδοχική φύση, αναμένουμε ότι θα είμαστε σε θέση να μεγιστοποιήσουμε τη δέσμευση.

- Το πρόβλημα των συστημάτων Recommender μπορεί να εξεταστεί κάτω από το πρίσμα της Ενίσχυτικής Μάθησης!

- Ενεργός ερευνητικός τομέας

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Επόμενες Διαλέξεις

Μέρος 4: Ενισχυτική Μάθηση

- Εισαγωγή στη ΕΜ
- Διαδικασίες Αποφάσεων Markov & Δυναμικός Προγραμματισμός

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Σας ευχαριστούμε

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423