

Πετράκης Κωνσταντίνος ΥΔΑ ΑΜ:1041589

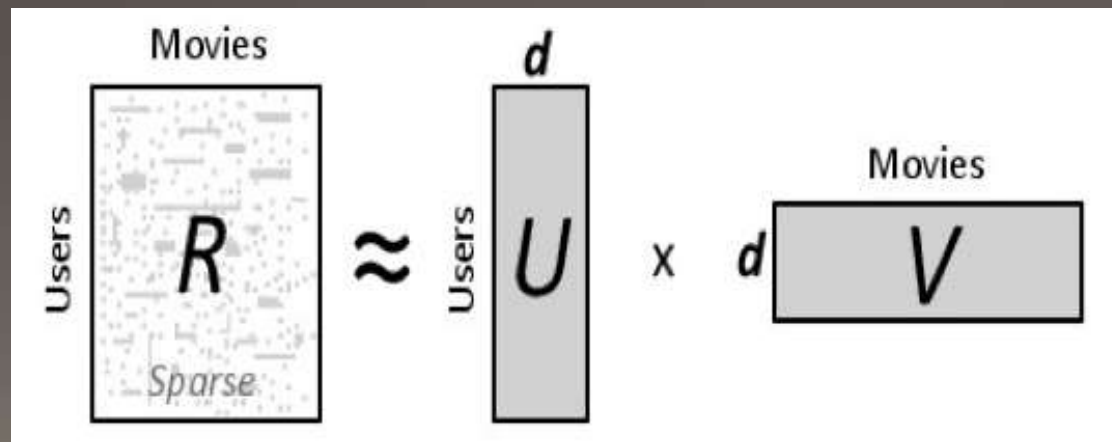
Steffen Rendle, Walid Krichene, Li Zhang, John Anderson
Neural Collaborative Filtering vs. Matrix Factorization Revisited

Τι περιλαμβάνει η εργασία

- Επανεξετάζονται τα αποτελέσματα της εργασίας ‘Neural Collaborative Filtering’ (NCF) που εισήγαγε τη χρήση MLP για την μάθηση μιας συνάρτησης ομοιότητας χρηστών-αντικείμενων.
- Χρήση του εσωτερικού γινομένου με κατάλληλες υπέρ-παραμέτρους οδηγεί σε καλύτερα αποτελέσματα από τις ομοιότητες που μαθαίνονται με MLP.
- Εμπειρική απόδειξη πως είναι δύσκολο ένα MLP να ‘μάθει’ το εσωτερικό γινόμενο
- Ζητήματα υλοποίησης και πρακτικής εφαρμογής των 2 μεθόδων σε περιβάλλοντα πραγματικού κόσμου

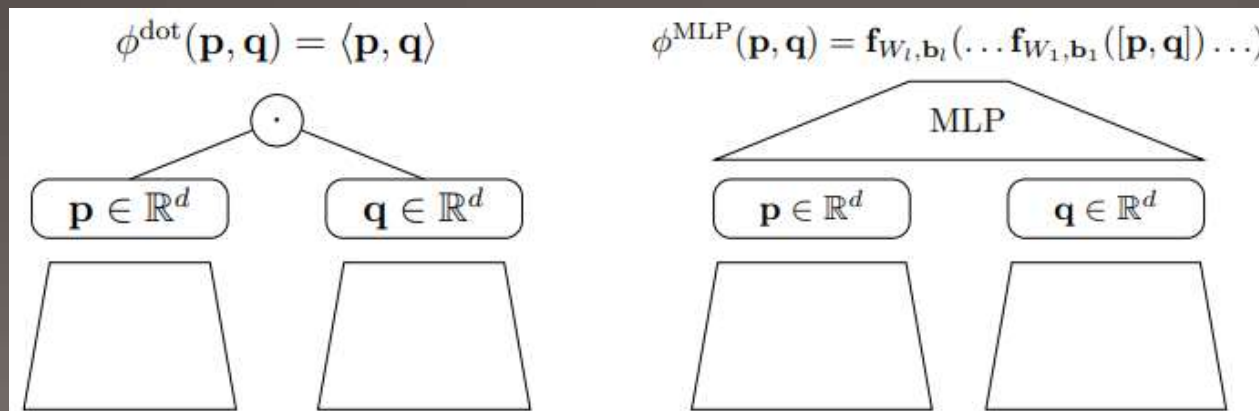
Συνεργατικό φιλτράρισμα-Τεχνικές

- Συνεργατικό φιλτράρισμα: σύσταση αντικειμένων στους χρήστες με βάση την ομοιότητα στις προτιμήσεις τους.
- Παραγοντοποίηση Μητρώου (MF): χρήστες και αντικείμενα αντιστοιχίζονται σε έναν κοινό latent χώρο χαμηλότερης διάστασης d .
 - Συνάρτηση αλληλεπίδρασης: το εσωτερικό γινόμενο των embedding διανυσμάτων χρηστών-αντικειμένων.
 - Ελεύθερες παράμετροι τα διανύσματα p, q , όπου $p \in \mathbb{R}^d$ το embedding χρήστη και $q \in \mathbb{R}^d$ το embedding αντικειμένου
- NCF: προτάθηκε η χρήση MLP για την μοντελοποίηση της συνάρτησης ομοιότητας μεταξύ χρηστών-αντικειμένων
 - Θεώρημα Καθολικής Προσέγγισης



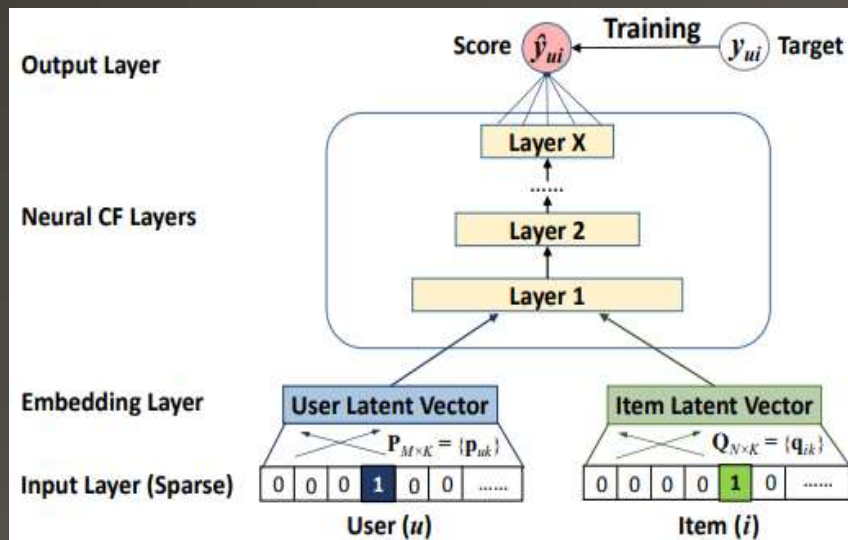
Μαθηματική Διατύπωση

- $\varphi: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ η συνάρτηση αλληλεπίδρασης
 - Τα embeddings p, q μπορεί να είναι παράμετροι του μοντέλου ή συναρτήσεις άλλων χαρακτηριστικών (π.χ. η έξοδος ενός ΤΝΔ)
- $\varphi^{dot}(p, q) = \langle p, q \rangle = p^T q = \sum_{f=1}^d p_f q_f$
 - Παραλλαγή με ρητά (explicit) biases $\varphi^{dot}(p, q) := b + p_1 + q_1 + \langle p_{[2...d]}, q_{[2...d]} \rangle$
- Ένα επίπεδο ενός MLP: $f_{W,b}(x) = \sigma(Wx + b)$, όπου $W \in \mathbb{R}^{in \times out}$, $b \in \mathbb{R}^{out}$ και $\sigma: \mathbb{R} \rightarrow \mathbb{R}$.
- Στην εργασία NCF προτάθηκε $\varphi^{MLP}(p, q) := f_{W_l, b_l}(\dots f_{W_1, b_1}([p, q]) \dots)$ και μια παραλλαγή $\varphi^{NeuMF} := \varphi^{MLP}(p_{[1...j]}, q_{[1...j]}) + \varphi^{GMF}(p_{[j+1...d]}, q_{[j+1...d]})$ όπου $\varphi^{GMF}(p, q) := \sigma(w^T(p \odot q)) = \sigma(\langle w \odot p, q \rangle) = \sigma(\sum_{f=1}^d w_f p_f q_f)$

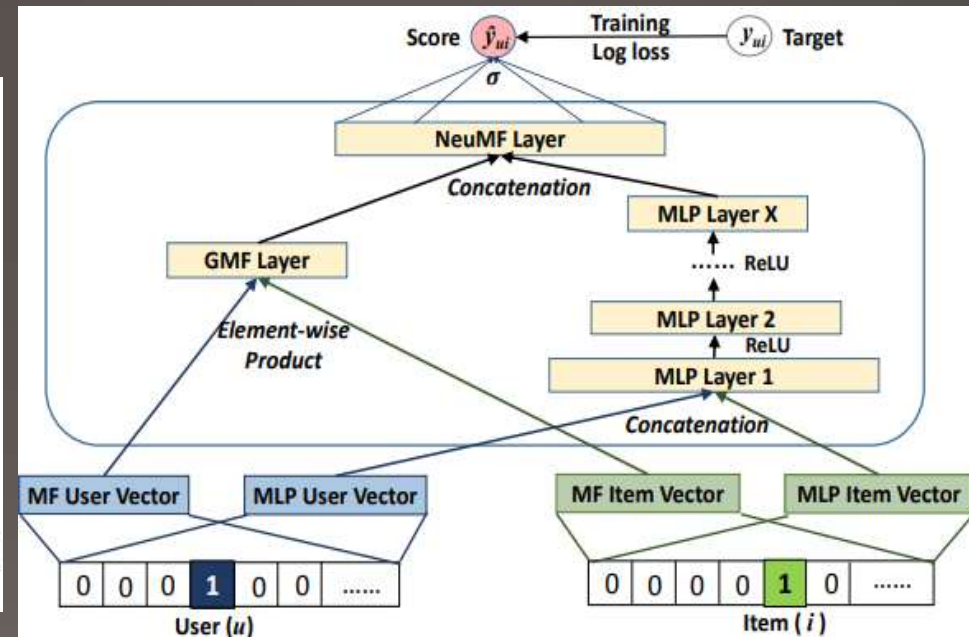


Χρήση ΤΝΔ στην εργασία NCF

MLP



NeuMF



Σύνολα Δεδομένων και Μετρικές

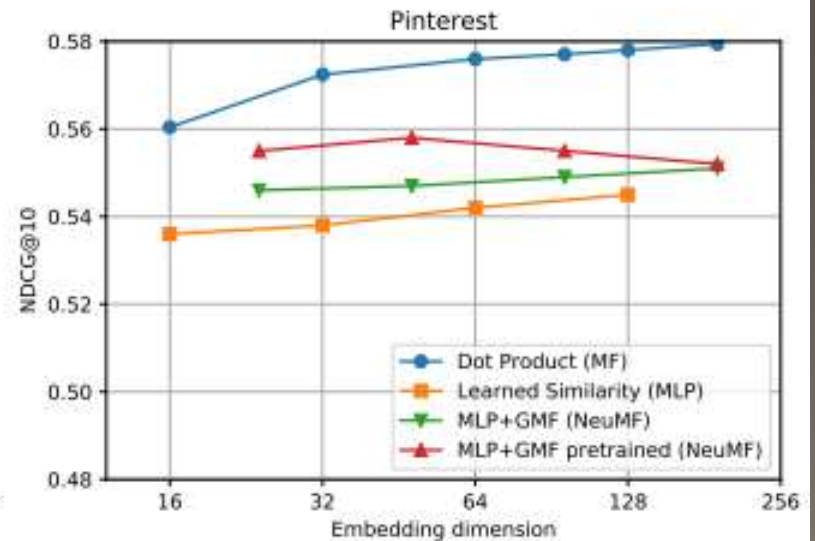
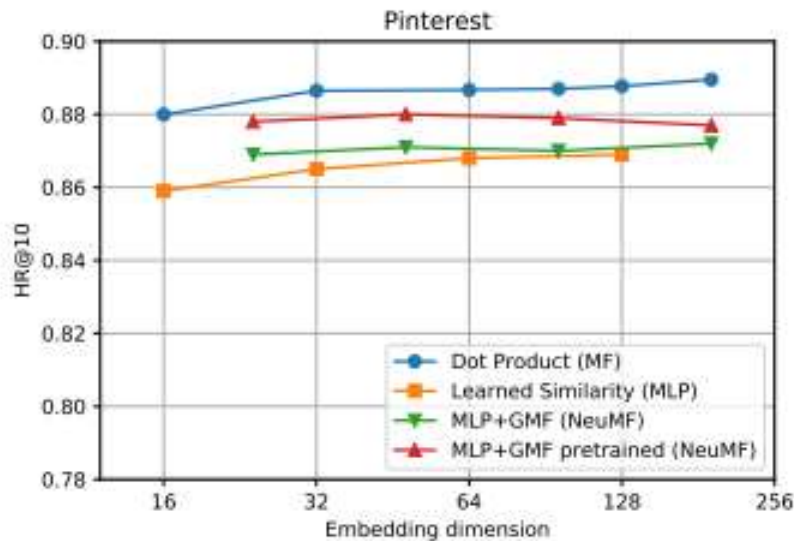
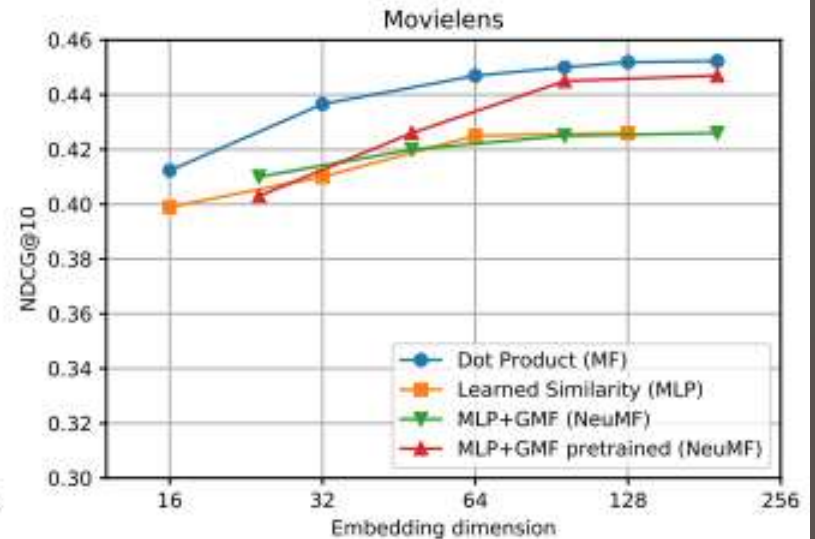
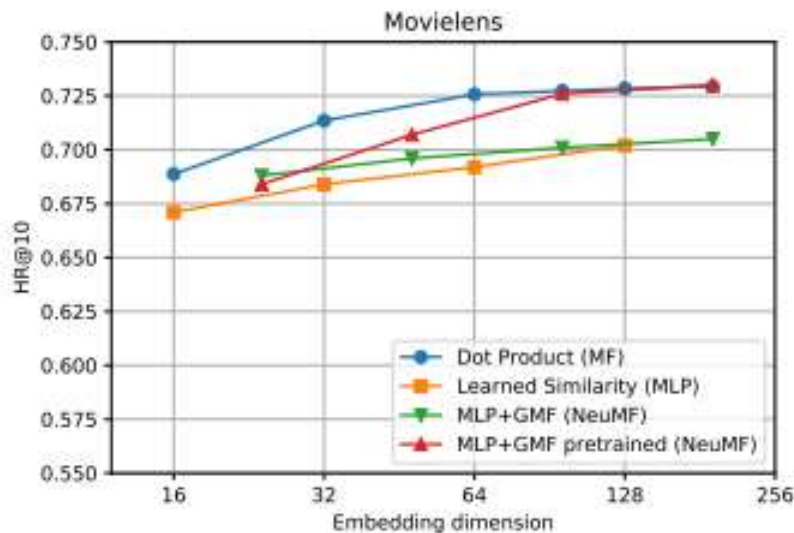
- Χρησιμοποιούνται 2 implicit feedback σύνολα δεδομένων
 - MovieLens 1M (binarized)
 - Pinterest.
- Σύνολο ελέγχου: το τελευταίο αντικείμενο αλληλεπίδρασης για κάθε χρήστη.
 - Τα υπόλοιπα αντικείμενα αποτελούν το σύνολο εκπαίδευσης.
- Αξιολόγηση: ταξινομούνται για κάθε χρήστη 101 αντικείμενα, το αντικείμενο του συνόλου ελέγχου και 100 τυχαία αντικείμενα με τα οποία δεν είχε αλληλεπίδραση, και αξιολογείται η θέση του αντικειμένου του συνόλου ελέγχου (στα πρώτα 10 αντικείμενα) με τις μετρικές:
 - Hit Ratio (Recall)
 - NDCG : $1/\log(r+1)$, r η θέση του παρακρατηθέντος αντικειμένου

Λεπτομέρειες υλοποίησης MF

Εκπαίδευση MF μοντέλου:

- Συνάρτηση σφάλματος $l(u, i, y) = -y \ln(\sigma(\varphi(p_u, q_i))) - (1 - y) \ln(1 -$

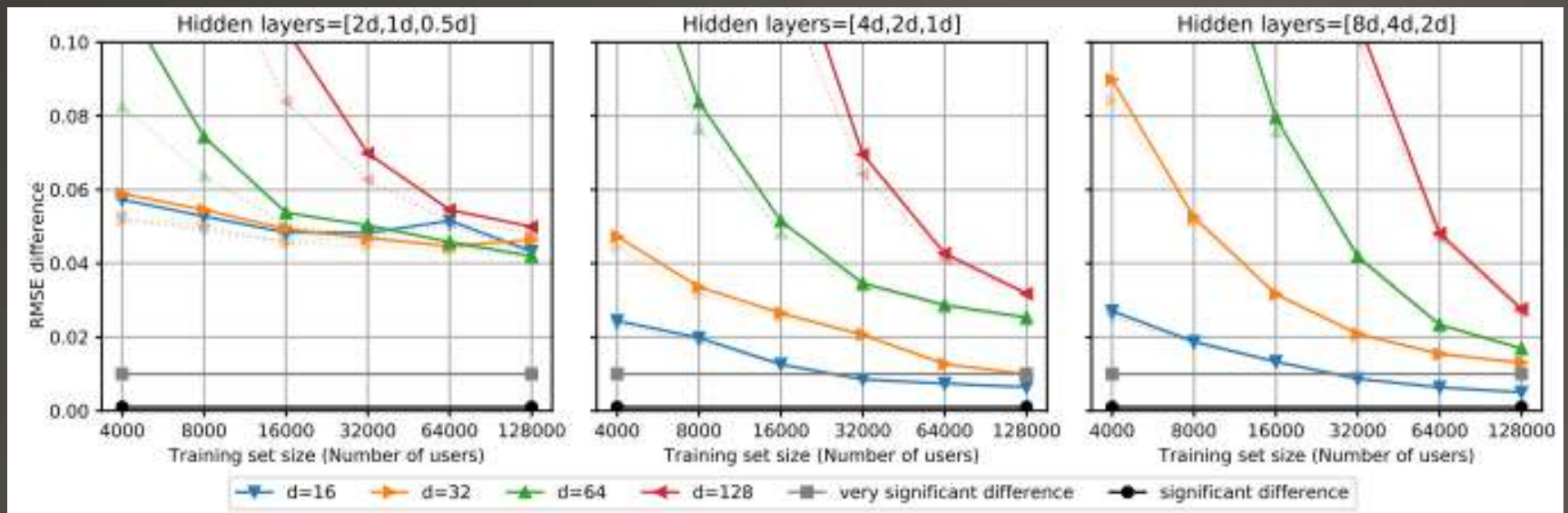
Πειραματικά Αποτελέσματα



Προσέγγιση εσωτερικού γινομένου με MLP

- Με στόχο να δουν αν η συνάρτηση ομοιότητας που μαθαίνει το MLP γενικεύει σε embeddings που δεν έχει δει στην εκπαίδευση.
 - Embeddings $p, q \in \mathbb{R}^d$, $p, q \sim \mathcal{N}(0, \sigma_{emb}^2 I)$
 - Πραγματικές τιμές ομοιότητας $y(p, q) = \langle p, q \rangle + \epsilon$, όπου $\epsilon \sim \mathcal{N}(0, \sigma_{label}^2)$
 - Χρήση MLP με συνάρτηση στόχο $\hat{y}: \mathbb{R}^{2d} \rightarrow \mathbb{R}$ και $\hat{y}(p, q)$ η προσέγγιση
- 3 συνθετικά σύνολα δεδομένων (πλειάδες $\{p, q, y\}$), 1 για εκπαίδευση και 2 για έλεγχο
 - Δειγματοληπτούν M χρήστες και N αντικείμενα \rightarrow μητρώα $P \in \mathbb{R}^{M \times d}$, $Q \in \mathbb{R}^{N \times d}$
 - Δειγματοληπτούν ομοιόμορφα 100M συνδυασμούς χρηστών-αντικειμένων
 - 90% το σύνολο εκπαίδευσης
 - 10% το 1^ο σύνολο ελέγχου
 - Δειγματοληπτούν ‘φρέσκα’ embeddings από $\mathcal{N}(0, \sigma_{emb}^2 I)$ για το 2^ο σύνολο ελέγχου
- Αρχιτεκτονική MLP: 3 κρυφά επίπεδα μεγέθους $[4h, 2h, h]$ για $h=d/2$, d , $2d$
- $\sigma_{label} = 0.85$
 - Στο Netflix prize τα καλύτερα μοντέλα έχουν $RMSE=0.85$
- $\sigma_{emb}^2 = \sqrt{\frac{1.13^2 - 0.85^2}{d}}$
 - Τετριμμένο μοντέλο (προβλέπει πάντα 0) θα έχει αναμενόμενο $RMSE$ ίσο με $\sqrt{\text{Var}(y)} = \sqrt{\sigma_{label}^2 + d\sigma_{emb}^4}$
 - Στο Netflix το τετριμμένο μοντέλο έχει $RMSE=1.13$
- Συγκρίνεται το MLP με το MF μοντέλο όπου $\hat{y}(p, q) = \langle p, q \rangle$
 - Το MF έχει αναμενόμενο $RMSE$ ίσο με σ_{label}

Αποτελέσματα τεχνητών πειραμάτων



- Για προσέγγιση χρειάζεται πλήθος δειγμάτων $O(d/\epsilon)^a$, $1 \leq a \leq 2$ και κρυφά επίπεδα με αρκετούς νευρώνες

Στην πράξη...

- MLP δεν μπορούν να χρησιμοποιηθούν σε πραγματικού χρόνου, βασισμένων στα συμφραζόμενα (context-aware) εφαρμογές
- Η χρονική πολυπλοκότητα για την σύσταση από ένα σύνολο n αντικειμένων
 - MLP: $O(d^2n)$
 - Εσωτερικό γινόμενο: $O(dn)$
 - Προσεγγίζεται αποδοτικά σε υπό-γραμμικό χρόνο

Αλλά...

- Δεν αποθαρρύνεται η χρήση των MLP
- ΤΝΔ στα οποία χρησιμοποιείται το εσωτερικό γινόμενο πριν την έξοδο
 - Αντιστοίχιση της εισόδου x σε μια αναπαράσταση (embedding) $f(x) \in \mathbb{R}^d$
 - Έξοδος $Qf(x)$ (logits της softmax), $Q \in \mathbb{R}^{n \times d}$, όπου n το πλήθος των αντικειμένων
 - Τότε $p=f(x)$ και $Qf(x)=Qp=[\langle p, q_i \rangle]_{i=1}^n$ τα score για κάθε αντικείμενο
- Χρήση MLP σαν το τελικό στάδιο άλλων ΤΝΔ
 - Εξωτερικό γινόμενο μεταξύ p και q και εφαρμογή CNN στο δισδιάστατο μητρώο που προκύπτει.

Συμπεράσματα

- Η εφαρμογή του εσωτερικού γινομένου έχει τα εξής πλεονεκτήματα:
 - Πιο εύκολη χρήση σε εφαρμογές πραγματικού κόσμου
 - Απλοποιείται η μοντελοποίηση και η διαδικασία της μάθησης
 - Δεν απαιτείται προ-εκπαίδευση,
 - Δεν υπάρχει ανάγκη για μεγάλα σύνολα δεδομένων
 - Λιγότερες παραμέτρους
 - Πιο φθηνό υπολογιστικά
 - Συμφωνία με άλλες περιοχές (Επεξεργασία φυσικής γλώσσας, Αναγνώριση εικόνας) όπου χρησιμοποιείται ευρέως
 - Έχει όμως τα καλύτερα αποτελέσματα?

Ευχαριστώ που με ακούσατε!