Deep Learning For Recommendation Systems

James Le
CSCI 736 - Neural Networks and Machine Learning
April 7th, 2020

About Me

- MS in Computer Science (Intelligent Systems)
- Member of Neural Adaptive Computing Lab
- Work Experience in FinTech and SaaS domains
- Writer on Medium: Data Science, Machine Learning, Statistics, Recommendation Systems, ML Infrastructure

Agenda

1 - Recommendation Systems Overview

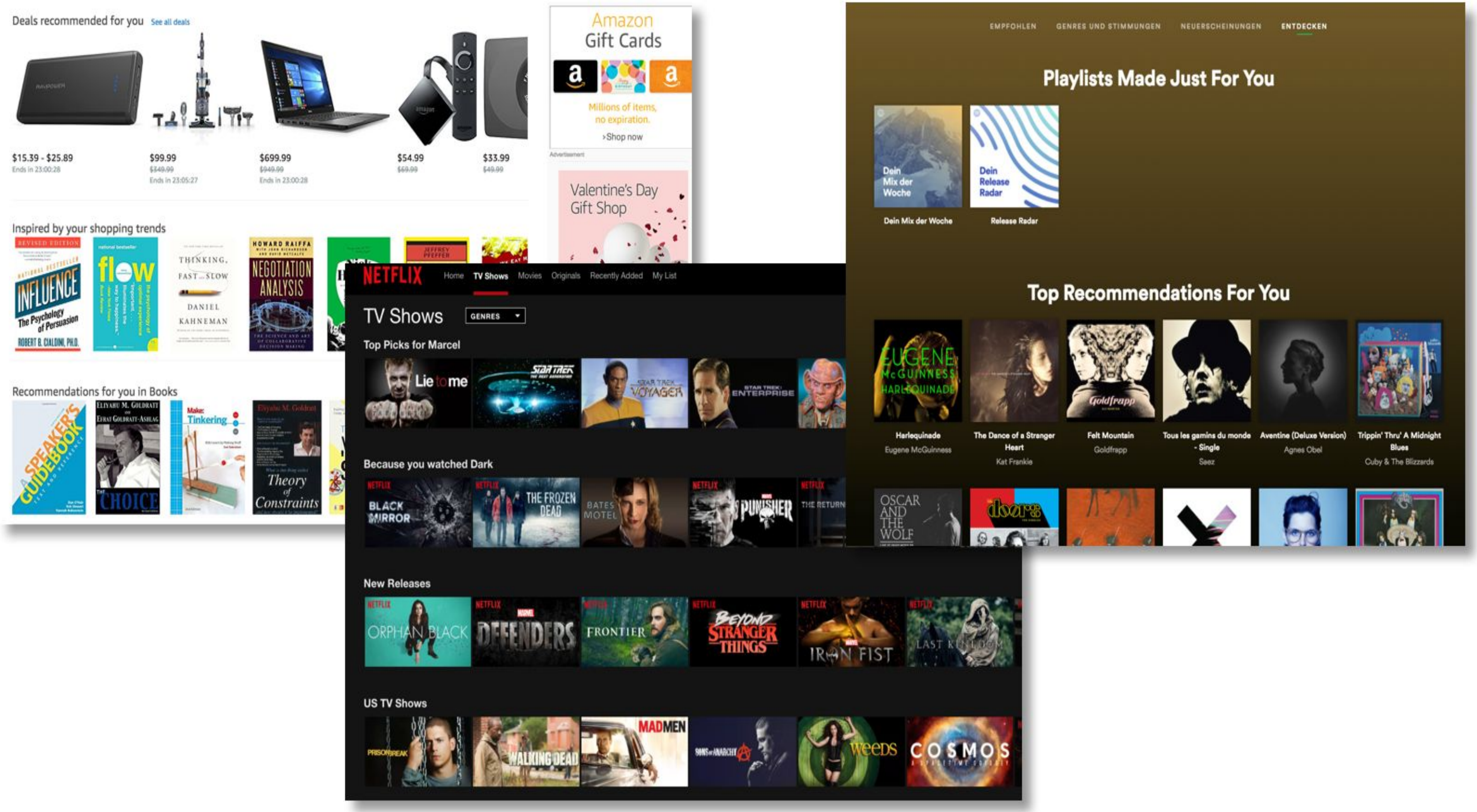
2 - Deep Learning Overview

3 - Deep Learning For Recommendations

4 - Experiments

5 - Future Directions

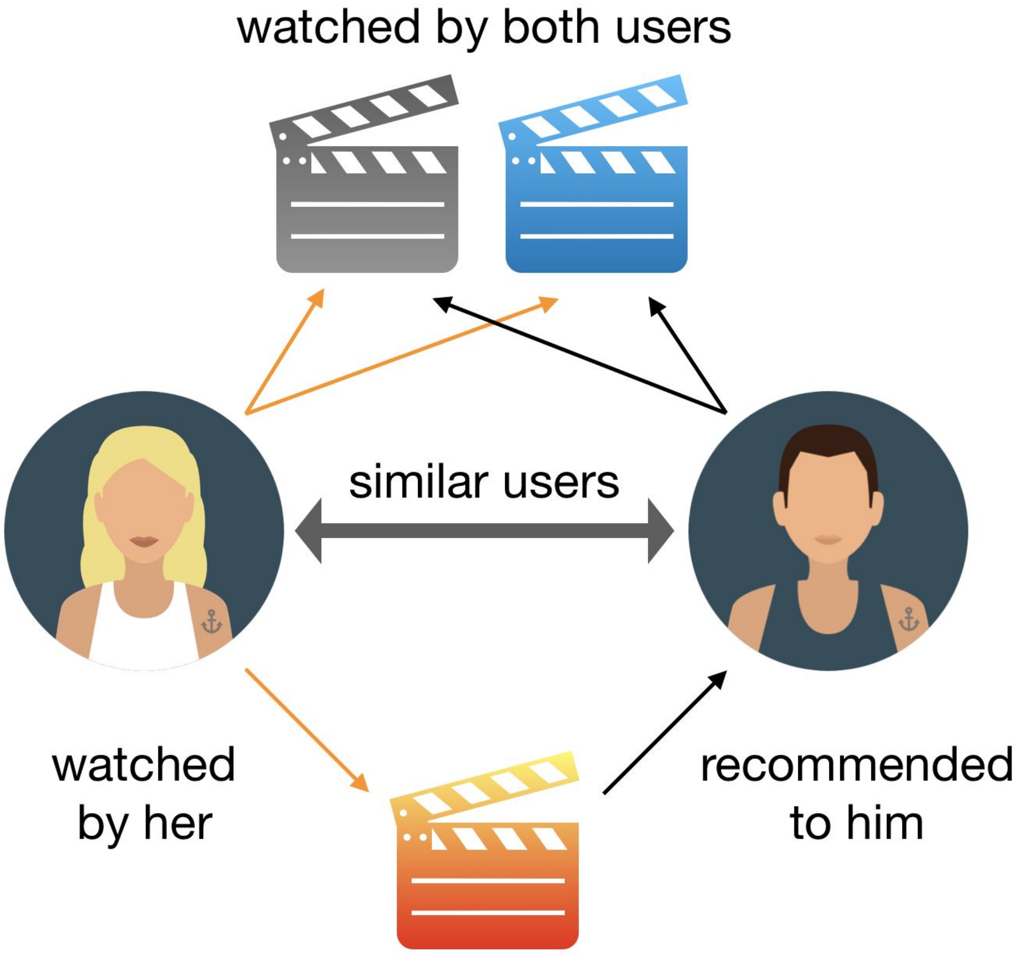
1 - Recommendation Systems Overview



Deals recommended for you
See all deals
$15.39 - $25.89
Ends in 23:00:28
$99.99
$349.99
Ends in 23:05:27
$699.99
$949.99
Ends in 23:00:28
$54.99
$69.99
$33.99
$49.99
Amazon
Gift Cards
Millions of items,
no expiration.
›Shop now
Advertisement
Valentine's Day
Gift Shop
Inspired by your shopping trends
INFLUENCE
The Psychology of Persuasion
ROBERT B. CIALDINI, PH.D.
flow
THINKING,
FAST AND SLOW
DANIEL
KAHNEMAN
HOWARD RAIFFA
NEGOTIATION
ANALYSIS
Recommendations for you in Books
SPEAKER'S GUIDEBOOK
ELIYAHU M. GOLDRATT
EFRAT GOLDRATT-ASHLAG
THE CHOICE
Make:
Tinkering
Eliyahu M. Goldratt
Theory of Constraints
NETFLIX
Home
TV Shows
Movies
Originals
Recently Added
My List
TV Shows
GENRES
Top Picks for Marcel
Lie to me
STAR TREK
STAR TREK VOYAGER
STAR TREK: ENTERPRISE
Because you watched Dark
BLACK MIRROR
THE FROZEN DEAD
BATES MOTEL
PUNISHER
New Releases
ORPHAN BLACK
DEFENDERS
FRONTIER
BEYOND STRANGER THINGS
IRON FIST
LAST KINGDOM
US TV Shows
PRISON BREAK
WALKING DEAD
MAD MEN
SONS of ANARCHY
weeds
COSMOS
EMPFOHLEN
GENRES UND STIMMUNGEN
NEUERSCHEINUNGEN
ENTDECKEN
Playlists Made Just For You
Dein Mix der Woche
Dein Release Radar
Dein Mix der Woche
Release Radar
Top Recommendations For You
EUGENE McGUINNESS
HARLEQUINADE
Goldfrapp
Harlequinade
Eugene McGuinness
The Dance of a Stranger Heart
Kat Frankie
Felt Mountain
Goldfrapp
Tous les gamins du monde - Single
Seez
Aventine (Deluxe Version)
Agnes Obel
Trippin' Thru' A Midnight Blues
Cuby & The Blizzards
OSCAR AND THE WOLF
doors

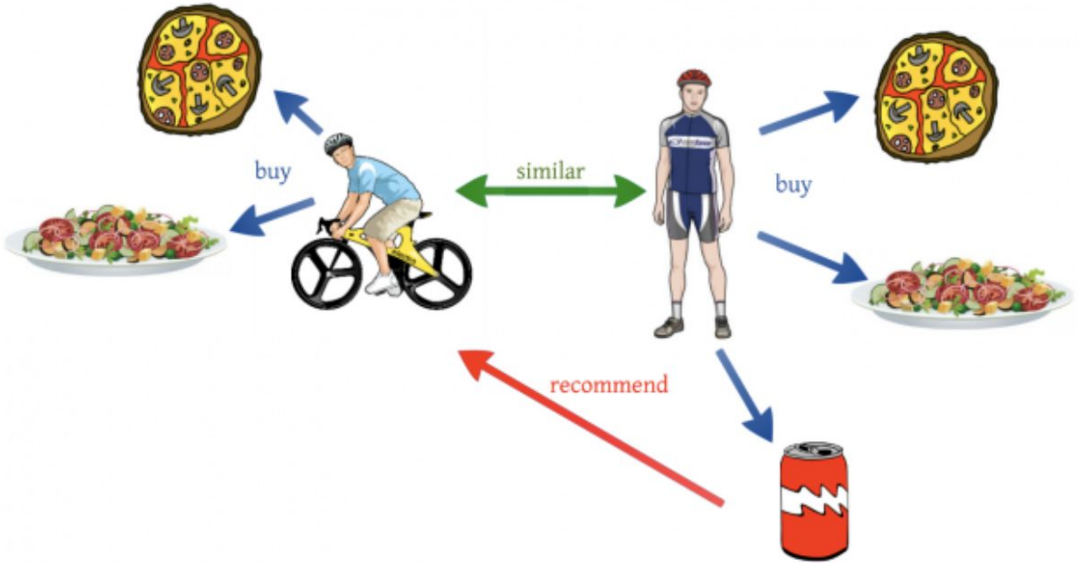
Content-Based Filtering

- Makes recommendations based on the user's purchase or consumption history
- Becomes more accurate the more actions/inputs the user takes
- By Content Similarity
- By Latent Factor Modeling
- By Popular Content Promotion



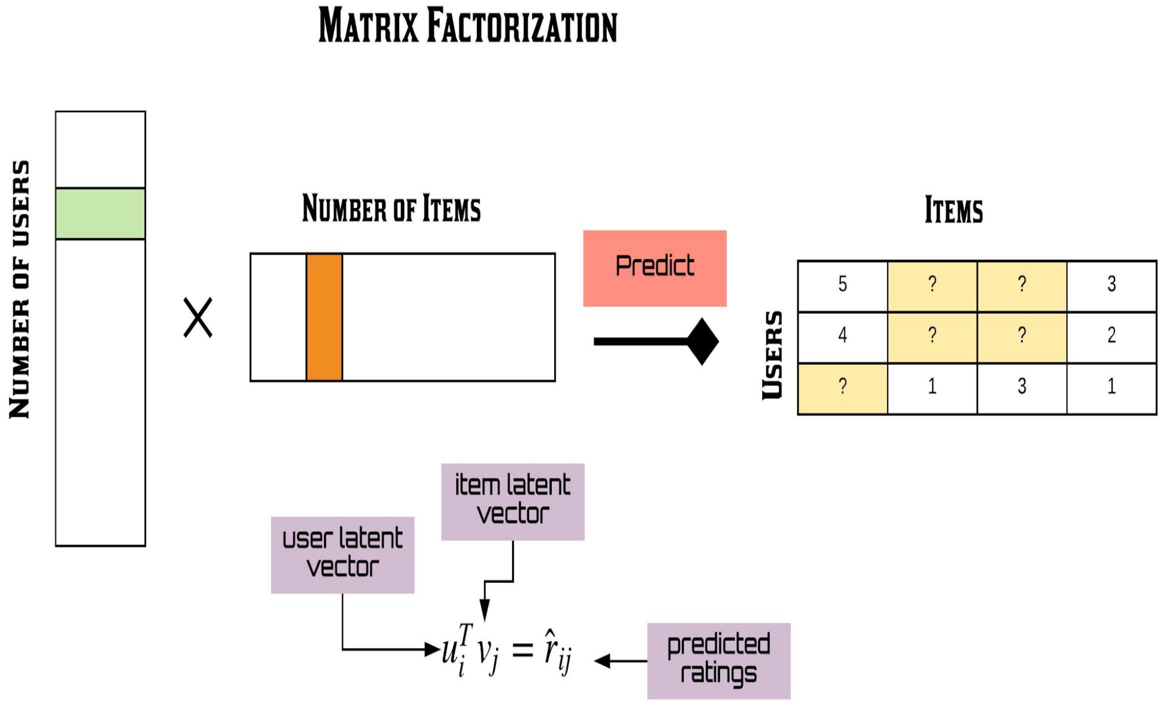
Collaborative Filtering

- Primarily makes recommendations based on inputs/actions from other people (rather than only the user for whom a recommendation is being made)
- By User Similarity
- By Association



Matrix Factorization

- The de facto standard model for collaborative filtering
- Represent user ratings as a user-item matrix
- Find two smaller matrices (called the latent/factor matrices) that approximate the full matrix
- Minimize the rating prediction errors (i.e. RMSE)
- Can be optimized via Gradient Descent
- Prediction is simple



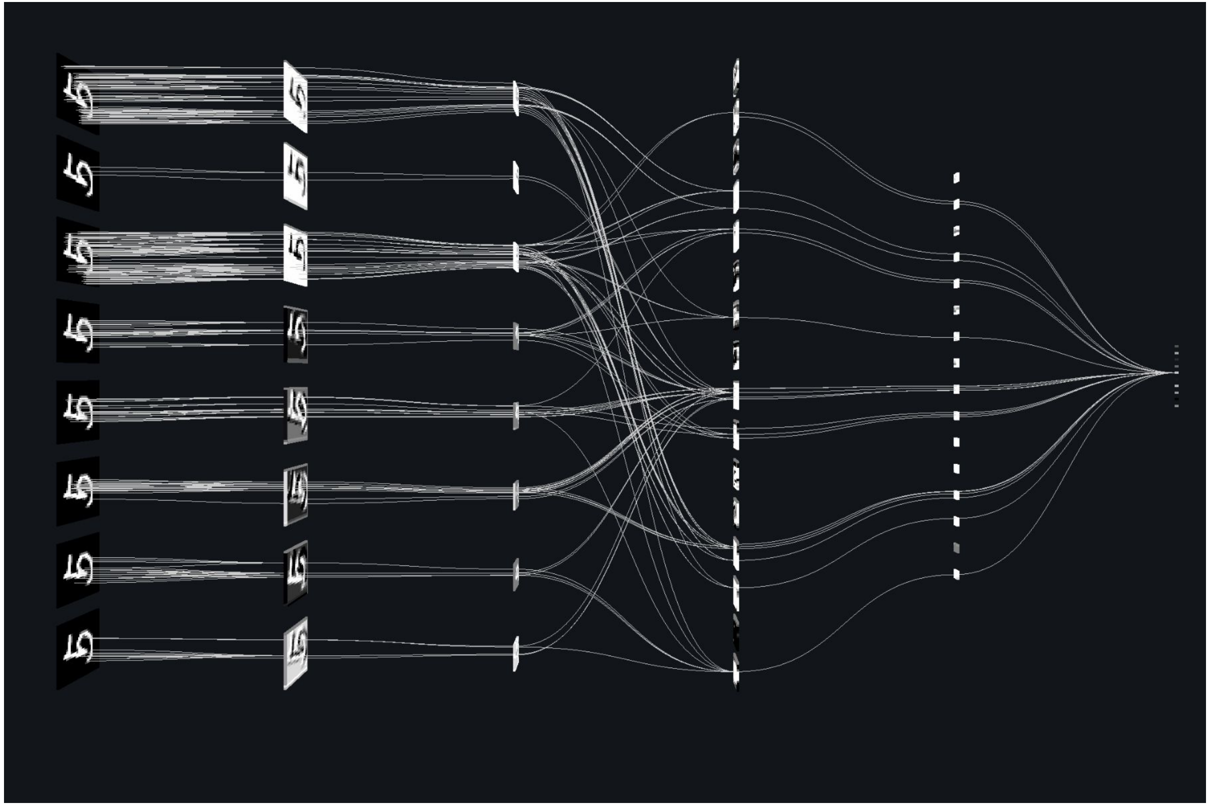
2 - Deep Learning Overview

Deep Learning Techniques

- Subfield of Machine Learning
- Learns *deep representations*
- Demonstrated success in both supervised and unsupervised learning tasks.

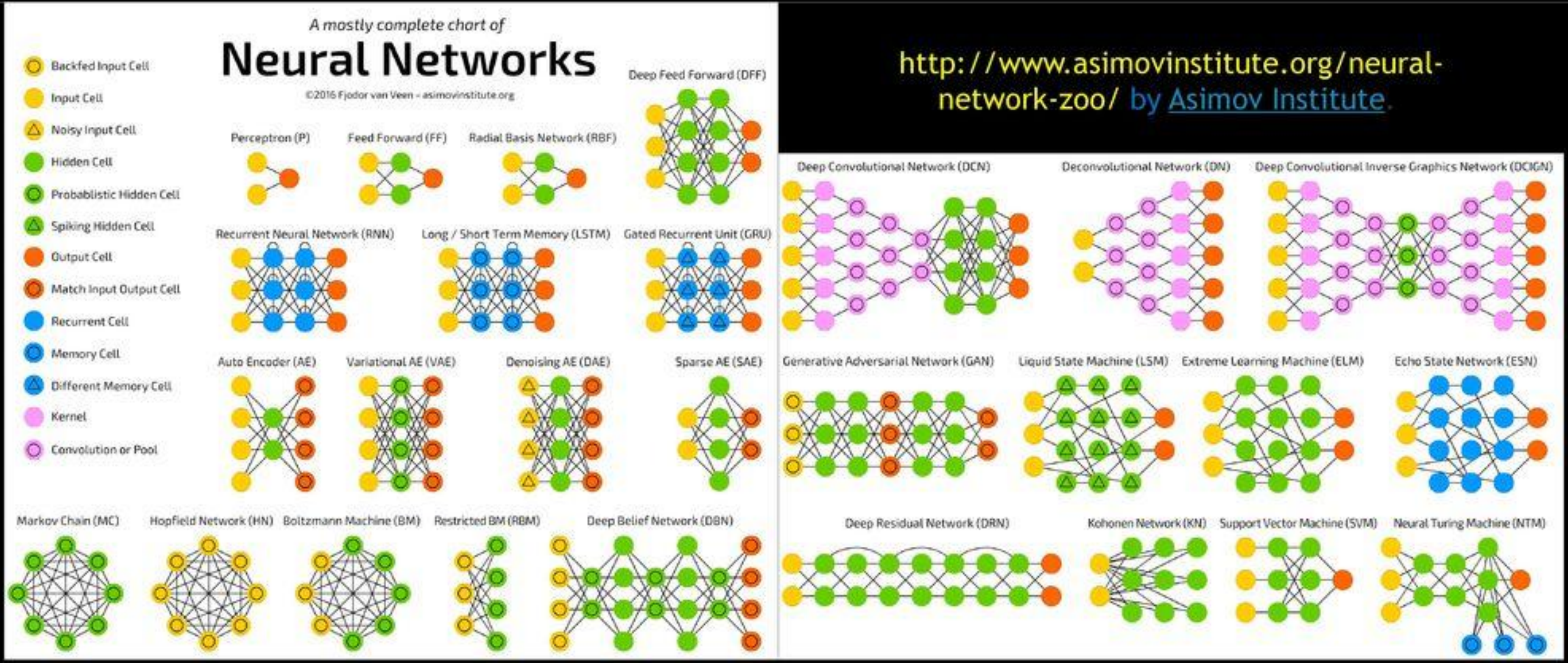
Architecture Paradigms:

1. Multilayer Perceptron
2. Autoencoder
3. Convolutional Neural Network
4. Recurrent Neural Network
5. Restricted Boltzmann Machine



Blog Post: [The 10 Neural Network Architectures Machine Learning Researchers Need To Learn]

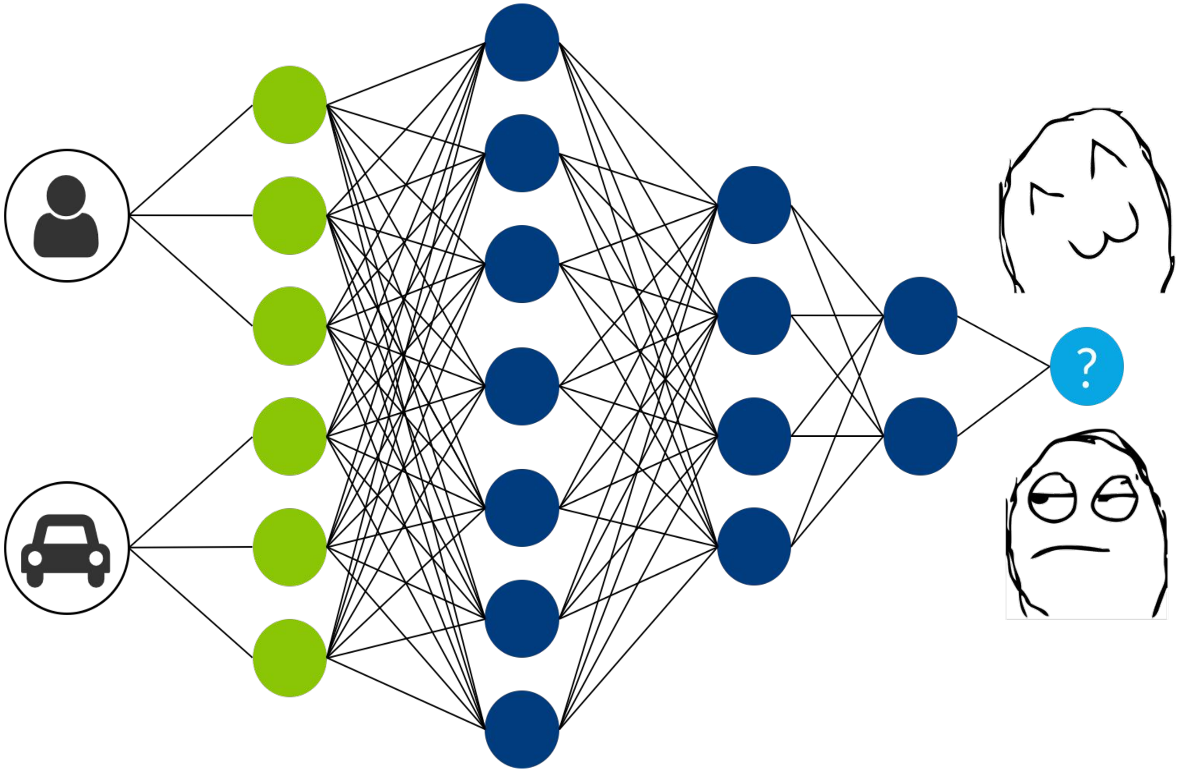
The Neural Networks Zoo



http://www.asimovinstitute.org/neural-network-zoo/ by Asimov Institute.

Why Deep Learning For Recommendations?

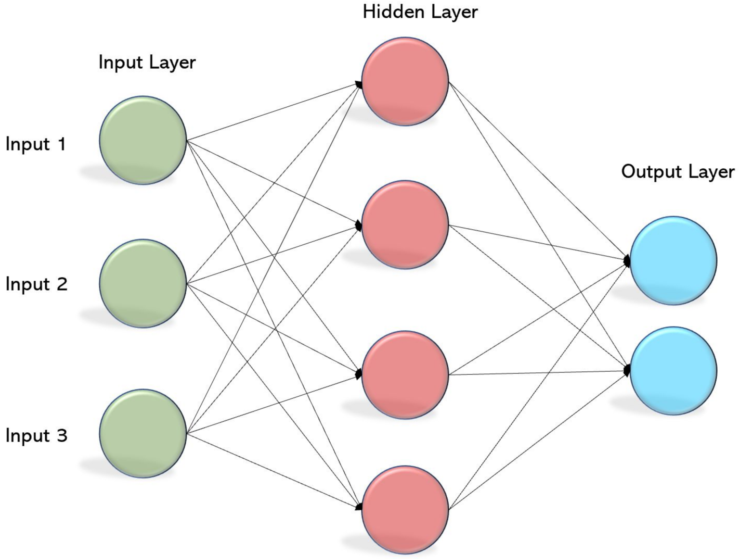
- **Nonlinear Transformation** => Dealing with complex user-item interaction patterns
- **Representation Learning** => (1) Feature engineering and (2) Multi-task learning
- **Sequence Modeling** => Mining temporal dynamics of user behavior and item evolution
- **Flexibility** => Modular frameworks with active community



3 - Deep Neural Networks For Recommendation Systems

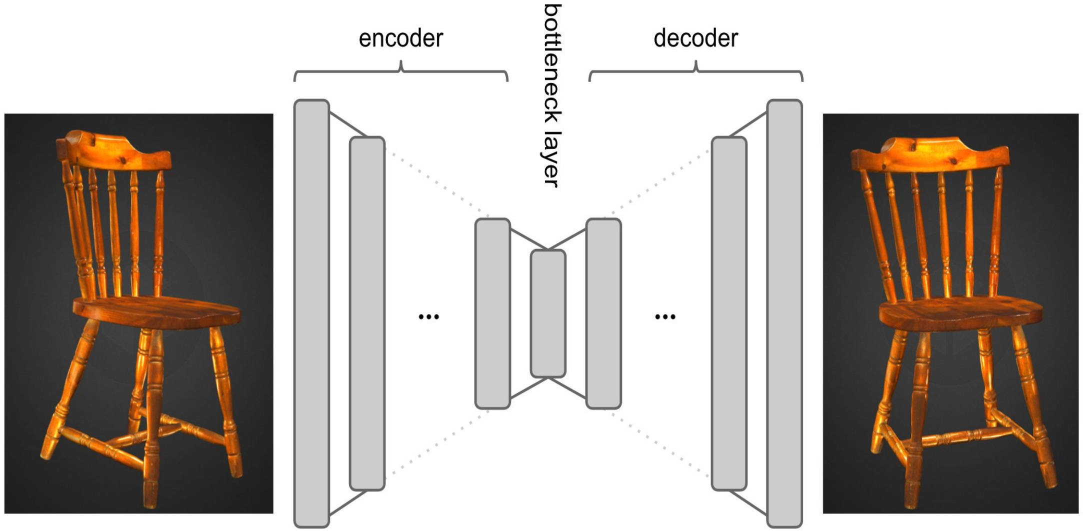
Multilayer Perceptron Based Recommendation

- MLP can add the non-linear transformation to existing recommendation approaches and interpret them into neural extensions
  - Neural Collaborative Filtering (National University of Singapore, 2017)
  - Deep Factorization Machine (Harbin Institute of Technology & Noah Research Lab, 2017)
- MLP can be used for feature representation
  - Wide and Deep Learning (Google, 2016)
  - Deep Neural Networks for YouTube Recommendations (Google, 2016)
  - Collaborative Metric Learning (Cornell Tech, 2017)



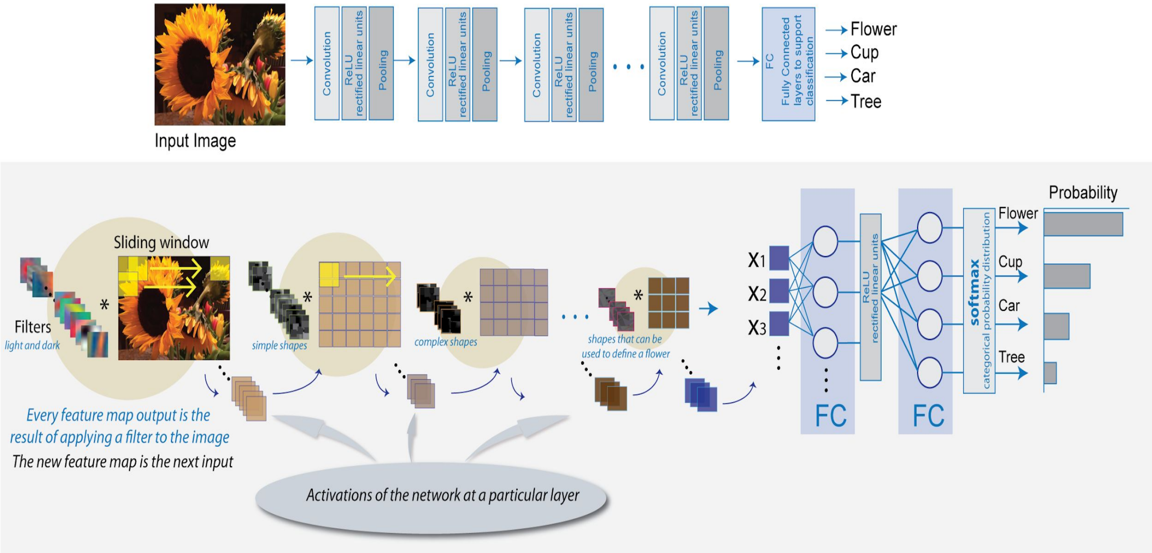
Autoencoder Based Recommendation

- AE can learn the lower-dimensional feature representation at the bottleneck layer
    - Collaborative Deep Learning (Hong Kong University of Science & Tech, 2015)
    - Collaborative Deep Ranking (Zhejiang University, 2016)
    - Deep Collaborative Filtering (Northeastern University and Adobe Research, 2015)
- AT can fill in the blanks of the user-item interaction matrix in the reconstruction layer
    - AutoRec (Australian National University, 2015)
    - Collaborative Denoising Autoencoder (Simon Fraser University and Dato Inc., 2016)
    - Multi-VAE and Multi-DAE (Netflix, Google, and MIT, 2018)



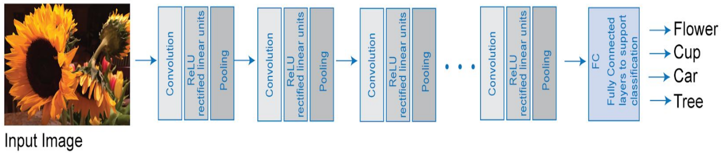
Convolutional Neural Nets Based Recommendation

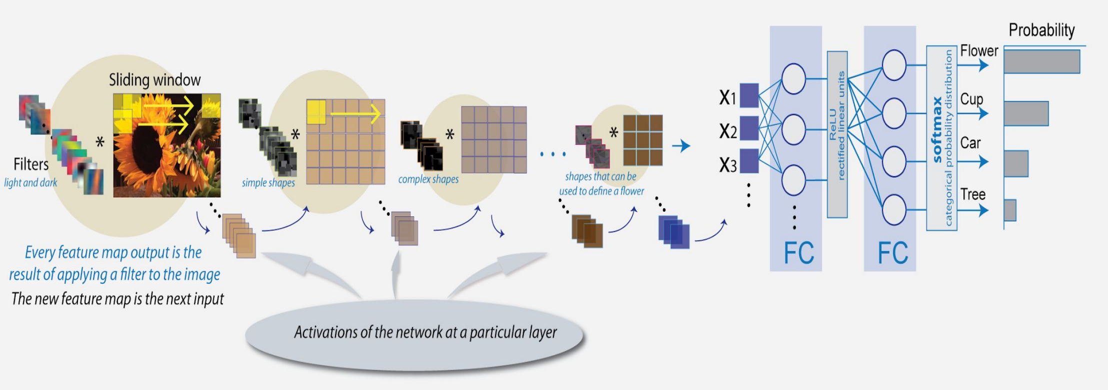
- ConvNet can extract features from images content
  - What Your Images Reveal (Arizona State and Michigan State, 2017)
  - Comparative Deep Learning of Hybrid Representations for Image Recommendations (University of Science and Tech of China, 2016)
  - ConTagNet (National University of Singapore, 2016)
- ConvNet can extract features from text content
  - DeepCoNN (U of Illinois - Chicago, 2017)
  - Automatic Recommendations Technology for Learning Resources (Central China Normal University, 2016)



Convolutional Neural Nets Based Recommendation

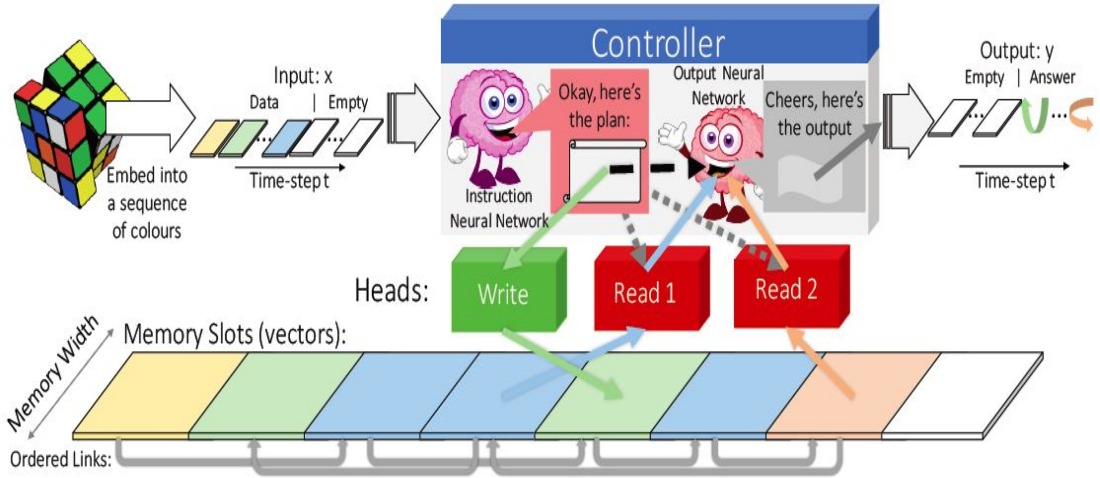
- ConvNet can extract features from audio and video content
  - Deep Content-Based Music Recommendation (Ghent University, 2013)
  - Collaborative Deep Metric Learning for Video Understanding (Google, 2018)
- ConvNet can be applied to vanilla CF
  - Outer Product-Based Neural Collaborative Filtering (National University of Singapore, 2018)
  - Personalized Top-N Sequential Recommendation via Convolutional Sequence Embedding (Simon Fraser University, 2018)
- Graph-Based ConvNet can handle the interactions in recommendation tasks
  - Graph Convolutional Matrix Completion (University of Amsterdam, 2017)
  - PinSage (Pinterest, 2018)





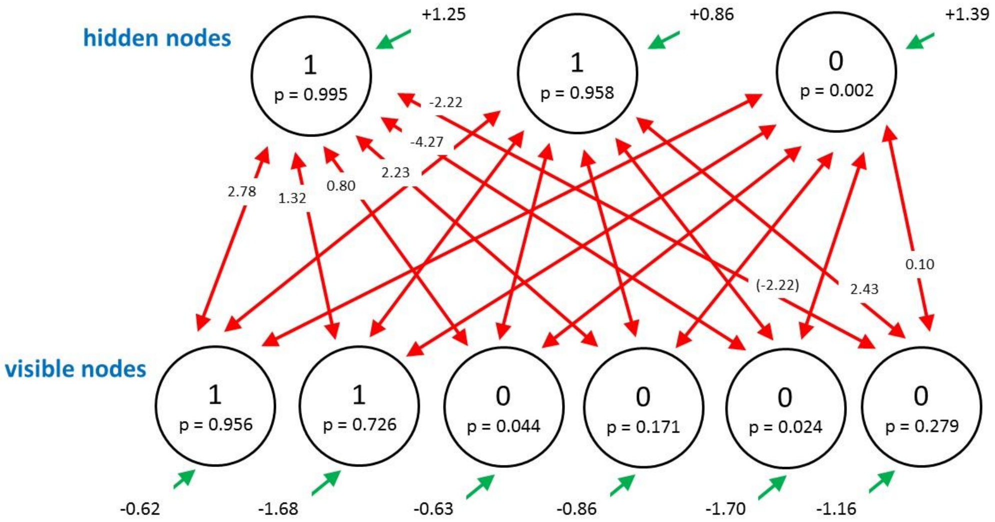
Recurrent Neural Nets Based Recommendation

- RNN can handle temporal dynamics of interactions and sequential patterns of user behaviors in session-based recommendation tasks
  - GRU4Rec (Gravity R&D, Netflix, Telefonica Research, 2016)
  - Personal Recommendation Using Deep Recurrent Neural Nets (Zhejiang University, 2016)
  - Recurrent Recommender Network (Google, LinkedIn, UTA, CMU, 2017)
- RNN can learn the side information with sequential patterns
  - Recurrent Coevolutionary Latent Feature Process for Continuous-Time Recommendation (GA Tech, 2016)
  - Ask the GRU (UMass Amherst, 2016)
  - Embedding-Based News Recommendation For Millions of Users (Yahoo Japan, 2017)



Restricted Boltzmann Machine Based Recommendation

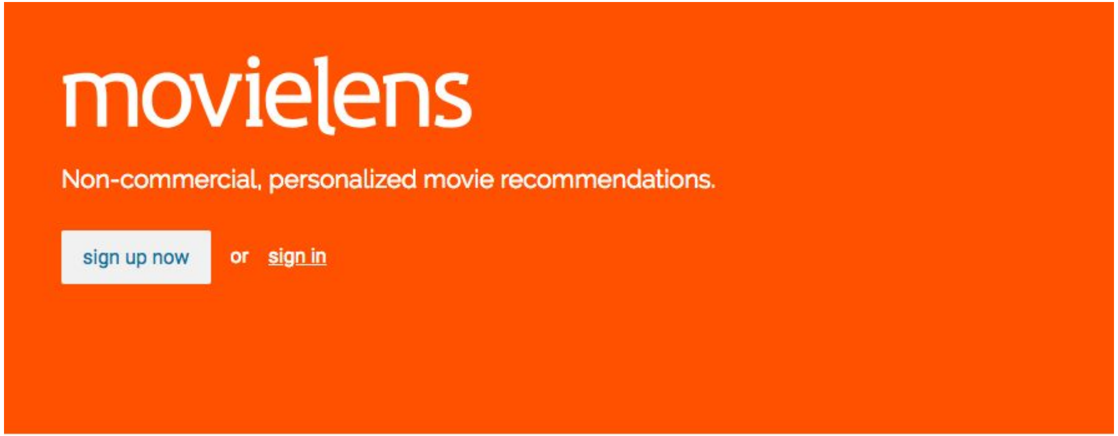
- RBMs For Collaborative Filtering (University of Toronto, 2007)
- A Non-IID Framework For Collaborative Filtering with RBMs (VMware and Qatar Computing Research Institute, 2018)
- Item Category Aware Conditional RBM Based Recommendation (Beihang University, 2015)

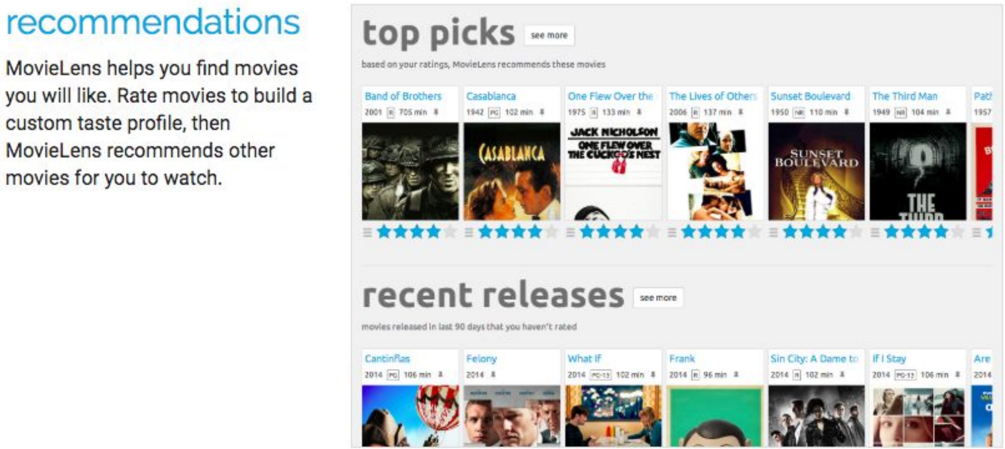


4 - Experiments

MovieLens1M Data

- 1 Million Ratings
  - UserID
  - MovieID
  - Rating
  - Timestamp
- 4,000 Movies
  - MovieID
  - Title
  - Genres
- 6,000 Users
  - UserID
  - Gender
  - Age
  - Occupation
  - Zipcodes





[Data Repository]

Matrix Factorization Experiments

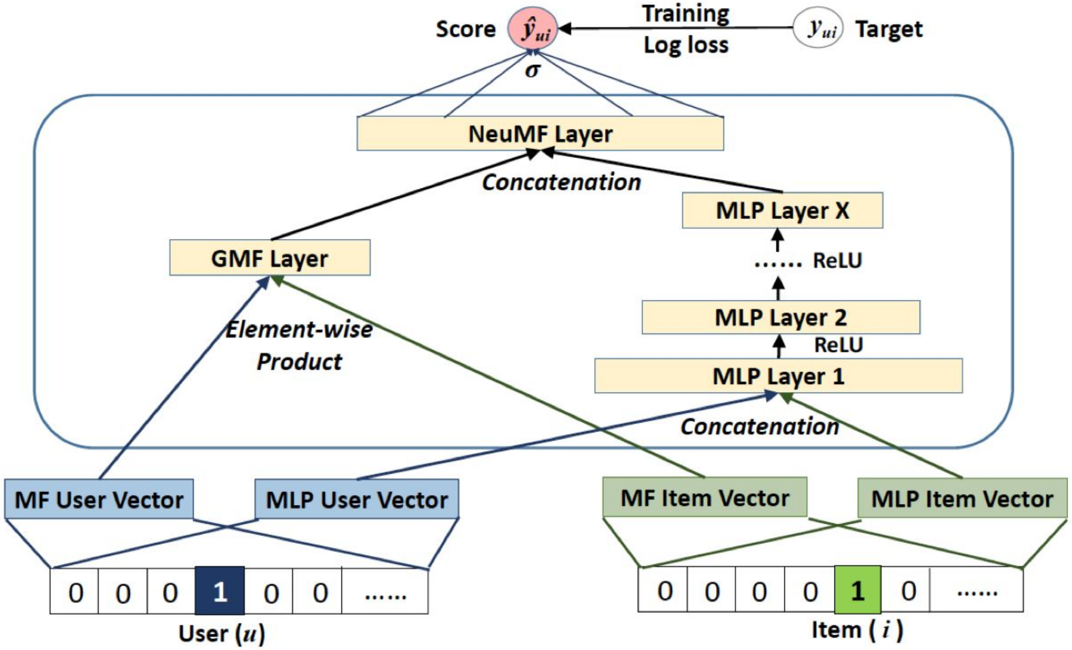
1. Vanilla Matrix Factorization
2. Matrix Factorization With Biases
3. Matrix Factorization With Side Features
4. Matrix Factorization With Temporal Features
5. Factorization Machines
6. Matrix Factorization With Mixture of Tastes
7. Variational Matrix Factorization

Model	Training Loss	Test Loss	Training Time
Vanilla MF	0.6947	0.8174	6m5s
MF Biases	0.6789	0.7895	11m38s
MF Side Features	0.6602	0.7843	13m34s
MF Temporal Features	0.7088	0.7939	18m51s
Factorization Machine	0.6542	0.8225	3m40s
MF Mixture of Tastes	0.6366	0.7878	13m44s
Variational MF	0.6206	0.8385	16m51s

Blog Post: The 7 Variants of Matrix Factorization For Collaborative Filtering

Neural Collaborative Filtering

- He, Liao, Zhang, Nie, Hu, and Chua, 2017
- 3 Layers: Generalized Matrix Factorization + Multi-Layer Perceptron = Neural Matrix Factorization
- Epochs = 50, Optimizer = Adam, Batch Size = 1024, Learning Rate = 0.001, L2 Regularizer = 0.01
- Hit Ratio @ 10 = 0.730
- Normalized Discounted Cumulative Gain @ 10 = 0.173



VAEs For Collaborative Filtering

- Liang, Krishnan, Hoffman, and Jebara, 2018
- Variational Autoencoders With Multinomial Likelihood
- Epochs = 100, Optimizer = Adam, Batch Size = 128, Learning Rate = 0.01
- Hit Ratio @ 10 = 0.837
- Normalized Discounted Cumulative Gain @ 10 = 0.403

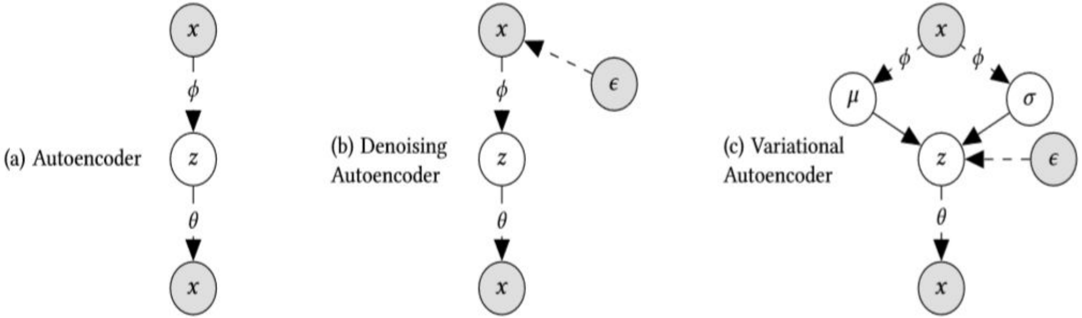


Figure 2: A taxonomy of autoencoders. The dotted arrows denote a sampling operation.

RBMs For Collaborative Filtering

- Salakhutdinov, Mnih, and Hinton, 2007
- 1 Layer of Visible Units
- 1 Layer of Hidden Units
- 1 Bias Unit
- Epochs = 200, Batch Size = 100, Hidden Units = 100
- Test Reconstruction Loss: 0.403

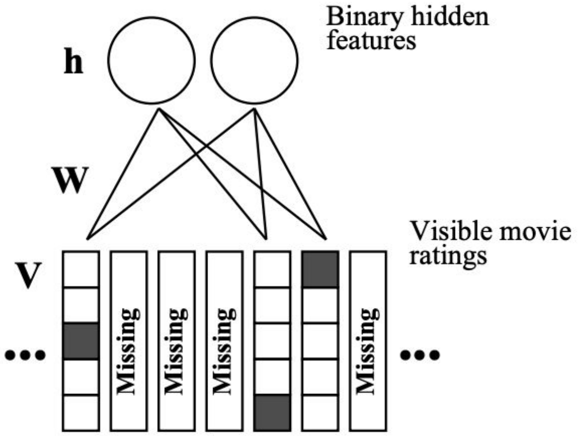


Figure 1. A restricted Boltzmann machine with binary hidden units and softmax visible units. For each user, the RBM only includes softmax units for the movies that user has rated. In addition to the symmetric weights between each hidden unit and each of the $K = 5$ values of a softmax unit, there are 5 biases for each softmax unit and one for each hidden unit. When modeling user ratings with an RBM that has Gaussian hidden units, the top layer is composed of linear units with Gaussian noise.

To Be Completed

Approaches

- Multi-Layer Perceptron Experiments
  - Neural Network Matrix Factorization
  - Wide and Deep Learning
- Autoencoders Experiments
  - AutoRec: Autoencoders Meet Collaborative Filtering
  - Denoising Autoencoders With Multinomial Likelihood
- Boltzmann Machines Experiments
  - Deep Boltzmann Machines
- Hybrid Experiments?

Datasets

- RecSys 2018 Challenge: Playlist Recommendation (Spotify)
- RecSys 2019 Challenge: Hotel Recommendation (trivago)
- RecSys 2020 Challenge: Tweet Engagement Prediction (Twitter)

5 - Future Directions

Research Directions

1 - Explainable Recommendation

- Make explainable predictions to users
- Focus on explain-ability to the practitioner
- Attentional models => Design **better** attentional mechanisms
- Neural Rating Regression with Abstractive Tips Generation For Recommendation (Chinese University of Hong Kong, JD.com, Tencent, 2017)

2 - Cross Domain Recommendation

- Domain adaptation via **transfer learning**
- A Multi-View Deep Learning Approach For Cross-Domain User Modeling in Recommendation Systems (Microsoft Research, 2015)
- A Content-Boosted Collaborative Filtering Neural Network for Cross-Domain Recommendation Systems (Microsoft Research, 2017)

Research Directions

3 - Multi-Task Learning

- Learning several tasks at a time can prevent overfitting by generalizing the shared hidden representations
- Auxiliary task provides interpretable output for explaining the recommendations
- Multi-task provides an implicit data augmentation for alleviating the sparsity problem.
- Neural Survival Recommender (LinkedIn, Amazon, 2017)
- Recommending What Video To Watch Next (Google, 2019)

4 - Scalability For Large-Scale Settings

- Incremental learning for non-stationary and streaming data
- Computation efficiency for high-dimensional tensors and multimedia data sources
- Balancing of the model complexity and scalability with the exponential growth of parameters
- **Compression** techniques (model pruning and model quantization)
- Knowledge **distillation** (Hugging Face's DistilBERT)

Research Lessons

An Opinionated Guide To ML Research (Open AI)

1 - Choosing Problems

2 - Making Continual Progress

3 - Building Knowledge of ML

Contact Info

- Email: jl1165@rit.edu
- Twitter: @le_james94
- LinkedIn: khanhnamle94
- GitHub: khanhnamle1994
- Medium: @james_aka_yale
- Website: jameskle.com

Questions?