Agenda

- 5:30 - Networking
- 6:00 - Talk
- 6:30 - Discussion and Networking

WiFi

- Name: RokkinCat Guest
- Password: makingstuff

Bathroom

- Take a bathroom key and go down a floor and the bathrooms are at the end of the hall

Deep Learning Recommender Systems

Milwaukee Machine Learning Meetup

Mitchell Henke / @MitchellHenke

About Me/This Meetup

- Software Architect
- Specialize in data, databases, APIs
- Self-taught R, Python/Keras, Machine Learning

DeadpoolKostumesmith 

@kibblesmith

Follow

Amazon is a $250 billion dollar company that reacts to you buying a vacuum by going THIS GUY LOVES BUYING VACUUMS HERE ARE SOME MORE VACUUMS

11:27 PM - 25 Apr 2016

Movie Recommender 5K

- MovieLens 10M Ratings Dataset
- Study recent research, and apply it to a product

Content-based

- Define information based on content
- Build a profile of user preferences
- Relate product information to user preferences

Collaborative Filtering

- Collect and analyze behavior
- Base recommendations on ratings from users, activity history, etc.

User-User Collaborative Filtering

- Compare users' movie ratings to each other
- Users are "similar" if they rate the same movies similarly
- Recommend movies you haven't rated, but similar users have rated highly

Item-Item Collaborative Filtering

- Compare movie ratings across users
- Movies are “similar” if users rate them similarly
- Recommend movies that are similar to ones you’ve rated highly

Deep Learning Recommenders

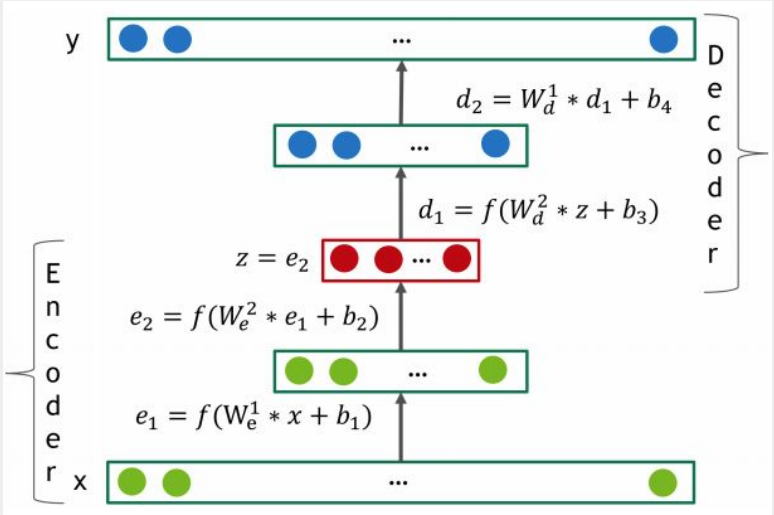
- Autoencoders
- Recurrent Neural Networks

Recommendations are Hard

- Balance between “obvious” recommendations while also providing an element of surprise
- Not all data is useful
- Machine learning is not a silver bullet

Autoencoders

- Dimensionality reduction
- Lossy compression
- Unsupervised

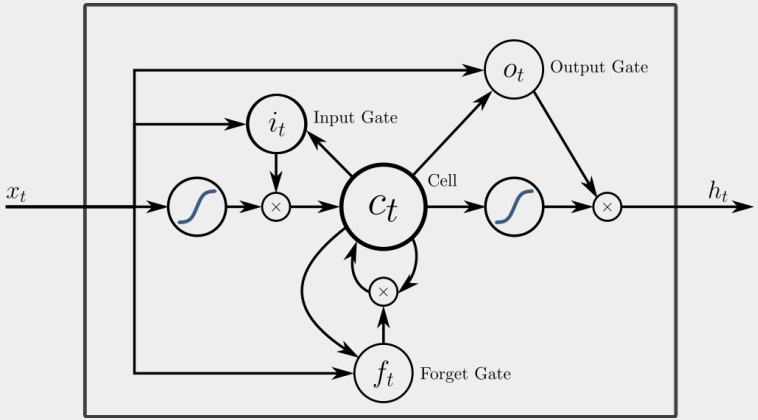


Deep Autoencoders for Recommending Movies

- Optimizes for rating prediction
- Likes to recommend movies lots of people like
- Suffers from cold-start

Recurrent Neural Networks

- Able to “remember”
- Learns what to remember, what to forget
- Recognize time-dependent relationships



Movie Recommendations as a Sequence

- Optimize for what to watch *next*, not what is *best*
- Focus on short-term instead of long-term

Lessons Learned

- Utilize multiple metrics
- Models will have weaknesses (and that’s okay)

Strengths & Weaknesses

- Autoencoder
  - Good at predicting what users will like the most
  - Tendency towards popularity
  - Cold-start
- Recurrent NN
  - Good at producing varied recommendations
  - Effective immediately
  - Doesn’t take into account ratings

Productizing Machine Learning Models

- Both models have strengths and weaknesses
- How can they be combined effectively?

Try It!

- movies.mitchellhenke.com
- Movie recommendations based on recent research
- Built in a few weeks

References:

- *Training Deep AutoEncoders for Collaborative Filtering*, Kuchaiev and Ginsburg (2017)
- *Collaborative Filtering with Recurrent Neural Networks*, Devooght and Bersini (2017)

Thanks!

movies.mitchellhenke.com