Recommendation Engine

Using Link prediction and Collaborative filtering

Prasanna Patil

Dept of Computer Science and Automation
Indian Institute of Sicence

April 28, 2019

Summary

Problem Statement

- Recommender System:- Recommending products, movies, items etc...
- Link prediction:- Predicting future connection among set of entities, given knowledge about current connections.
- Collaborative filtering:- Heavily used to make recommender system. Two types
  1. Item based
  2. User based

Relevant work

- Variational Auto Encoder based representation[1]
  - Problem: latent representation of users for collaborative filtering.
  - A VAE[2] based Encoder Decoder architecture with input to the encoder being item list of user: $x_u = [x_{u1}, x_{u2}, ..., x_{uI}]^T$.
  - Simple MLP network is used as encoder.
  - Samples a latent representation using VAE encoder and decoder tries to reconstruct the input vector.

[1] Liang, D. *et al.* Variational Autoencoders for Collaborative Filtering. *arXiv e-prints,* arXiv:1802.05814. arXiv: 1802.05814 [stat.ML] (2018).

[2] Kingma, D. P. & Welling, M. Auto-Encoding Variational Bayes. *arXiv e-prints,* arXiv:1312.6114. arXiv: 1312.6114 [stat.ML] (2013).

- Graph Convolution Networks[3]
  - Semi supervised learning over graphs.
  - Uses topological structure of graph along with additional node information to learn node representations.
  - Basic idea: Embeddings of a node are influences by embeddings of its neighbor nodes in each successive layers.

$$X_u = \mathcal{F}(X_u + \Sigma_{v \in N(u)} \mathcal{F}_e(X_v))$$

- Graph Auto Encoders for Link Prediction[4]
  - Problem: latent representation of nodes in graph to predict future interaction.
  - Also a VAE based encoder-decoder architecture that uses GCN is used as encoder.
  - Hidden representations are sampled using VAE encoder.
  - Decoder performs link prediction task over a graph.

[3] Kipf, T. N. & Welling, M. Semi-Supervised Classification with Graph Convolutional Networks. *arXiv e-prints,* arXiv:1609.02907. arXiv: 1609.02907 [cs.LG] (2016).

[4] Kipf, T. N. & Welling, M. Variational Graph Auto-Encoders. *arXiv e-prints,* arXiv:1611.07308. arXiv: 1611.07308 [stat.ML] (2016).

Approach

- Use link prediction paradigm in conjunction with collaborative filtering.
- Learn latent representation of both user and item.
  - Predict next item (or movie).
  - Predict rating alongside.
- A simple auto encoder based approach to learn representations.
- Use the underlying graph structure to learn representation jointly using Graph Auto Encoders [5].
- 2 layers of GCN are used for encoder. Embedding dimension is fixed to 128. Dropout is also used for regularization.
- Use decoder to predict rating and links as multi class classification and binary classification, respectively.

[5] More info in appendix.

Experiment setup

Dataset

- Netflix Prize (17,770 movies, 480189 user, 100480507 ratings (1-5))
- Dataset reduced to 7000 movies and 70,000 users for the purpose of experiments. Total 5456033 (5%) ratings preserved.
- For link prediction, if rating is greater than 2 then link is present otherwise absent.

Baseline

- Simple MLP based encoder decoder architecture.

Models

- Fixed initial embedding vs Learned initial embedding
- Variational Auto Encoder vs Auto Encoder.

Evaluation Metric

- Accuracy on rating classifier.
- Accuracy on link prediction classifier.

Observations

Table: Performance of Variational Auto Encoder

Model	Learned Embeddings		Fixed Embeddings	
-	Rating	Link Prediction	Rating	Link Prediction
Baseline	0.335	0.855	0.34	0.855
GAE	0.358	0.855	0.358	0.855

Table: Performance of Auto Encoder

Model	Learned Embeddings		Fixed Embeddings	
-	Rating	Link Prediction	Rating	Link Prediction
Baseline	0.478	0.860	0.435	0.866
GAE	0.484	0.860	0.455	0.861

Table: F1-Score of Graph Auto Encoders on Fixed Embeddings

Model	Fixed Embeddings	
-	Rating	Link Prediction
GVAE	0.359	0.922
GAE	0.362	0.924

Table: F1-Score of Graph Auto Encoders on Learned Embeddings

Model	Learned Embeddings	
-	Rating	Link Prediction
GVAE	0.358	0.922
GAE	0.376	0.923

Conclusion

- Graph Encoders perform $\sim 2\%$ better than MLP based encoders in both Variational Auto Encoder and Auto Encoder.
- Auto Encoder perform better than Variational Auto Encoder. This could be because there is no explicit reconstruction of input is happening in the output, in which case Auto Encoder becomes a simple neural network.
- Learned embedding out perform fixed embedding. However, fixed embedding do have an advantage in that they address cold start problem (partially).

References

- Kingma, D. P. & Welling, M. Auto-Encoding Variational Bayes. *arXiv e-prints,* arXiv:1312.6114. arXiv: 1312.6114 [stat.ML] (2013).
- Kipf, T. N. & Welling, M. Semi-Supervised Classification with Graph Convolutional Networks. *arXiv e-prints,* arXiv:1609.02907. arXiv: 1609.02907 [cs.LG] (2016).
- Kipf, T. N. & Welling, M. Variational Graph Auto-Encoders. *arXiv e-prints,* arXiv:1611.07308. arXiv: 1611.07308 [stat.ML] (2016).
- Liang, D., Krishnan, R. G., Hoffman, M. D. & Jebara, T. Variational Autoencoders for Collaborative Filtering. *arXiv e-prints,* arXiv:1802.05814. arXiv: 1802.05814 [stat.ML] (2018).

The End

Appendix

Architecture

- Let X_u be initial embedding of a node. Let $\hat{A} = D^{\frac{-1}{2}} A D^{\frac{-1}{2}}$.
- E(X_u) $= GCN_2(GCN_1(X_u))$, where $GCN_i(X_u) =$ LayerNorm($R_u * X_u + (1 - R_u)H_u$), where $H_u =$ Swish($\hat{A} X_u W_i$), Swish(x) $= x * \sigma(x)$ and $R_u = \sigma(W_o * X_u + W_n * H_u)$
- $X_\mu = \mathcal{F}_\mu(X_u)$ and $X_\sigma = \mathcal{F}_\sigma(X_u)$. Where $\mathcal{F}_\mu$ and $\mathcal{F}_\sigma$ are single layer feed forward network.
- Latent representation of node $Z_u = X_\mu + X_\sigma * \epsilon$ where $\epsilon \sim \mathcal{N}(0, I_n)$.
- Rating is predicted using $P_r = \mathcal{F}_r(Z_u, Z_m)$.
- For link prediction, the probability of ratings is summed up. More specifically, probability of link being absent is $P_r(rating = 1) + P_r(rating = 2)$ and link being present is $P_r(rating = 3) + P_r(rating = 4) + P_r(rating = 5)$.