NCF — Neural Collaborative Filtering

User item (movie). Interaction

Unique User : 200000
Unique Movie : 5000

Unique user X unique movie : 200000 X 5000
Har user ki embedding : 5000

matrix factorisation methods for Collaborative Filtering
U.VT -> SVD -> NMF…. -> disadvantages : This cannot learn non-linearity.

improved methods that can capture Non linearity
Non-linearity: Neural Collab Filtering, Two-tower Model , Bi-Encoder, Cross-Encoder, Classificatiion

additional info regarding embedding
Cross Encoder : BERT, Sentence-BERT (models) —> embeddings —>
{Neural Net} —> (Classification/Regression) head
Bi-Encoder : <sentence1, sentence2> —> <BERT, BERT> —> {cosine_losss}
—> similar sentences ka clustering and dissimilar ka things ka alag cluster

project info and methods

Interaction : If user has rated a movie (x), that is the interaction of user with the movie

Positive data for all user :
        Analysis : (Rating), Genre,
        Hard positive : Which you will definitely interact with (9/10) probs
        Soft positive : You might interact with (6/10)

Negative Data
        Hard negative : You will not at all interact with (1/10)
        Soft negative : Maybe you can click … (2/10)

Ratio of positive : negative for every user : 1:4, for every positive sample generate 4 negative samples.
Decide the ratio : Based on analysis.
If you have 2 positive interactions, then you will have 8 negative samples.

User.	Movie.	Interaction
User1.	Movie1.	1
User2.	Movie2.	1
User1.	Movie3.	0
User1.	Movie4.	0
User1.	Movie5.	0
User2.	Movie6.	0
User2	Movie3.	0
User2.	Movie4.	0
User2.	Movie5.	0