



Training Deep AutoEncoder for Collaborate Filtering

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/81432287-6078-445a-a958-440a8116b892/_0131_sent.pptx

<https://s3-us-west-2.amazonaws.com/secure.notion-static.com/6e19980b-580c-4c70-87e3-d8045d713c73/TraningDeepAutoEncodersforrCF.pdf>

Training deep auto encoders for collaborative filtering

Published on Training deep auto encoders for collaborative filtering 1. Training Deep AutoEncoders for Collaborative Filtering Marlesson Santana Prof. Dr. Anderson Soares 2. *

https://www.slideshare.net/marlessonsa/training-deep-auto-encoders-for-collaborative-filtering?from_action=save




Training Deep AutoEncoders for Collaborative Filtering
Marlesson Santana
Prof. Dr. Anderson Soares

Code

Deep Autoencoders For Collaborative Filtering

Introduction Deep Autoencoder Model Implementation An Autoencoder is a deep learning neural network architecture that achieves state of the art performance in the area of

 <https://towardsdatascience.com/deep-autoencoders-for-collaborative-filtering-6cf8d25bbf1d>



Abstract

time-split Netflix dataset에서 SOTA를 보였다. 6개의 레이어를 가지고 프리-트레이닝 없이.

- 1) 딥 오토인코더가 얇은 모델에 비해 좀 더 일반화를 잘한다는 것을 보였다.
- 2) Negative part를 가지는 비선형 활성화 함수가 딥 모델을 학습 하는게 crucial하다.
- 3) 오버피팅을 방지하기 위해서 드롭아웃 같은 레귤라이제이션 기술을 많이 쓰는게 필요하다.

새로운 알고리즘을 보였다. 그 알고리즘을 대충 보자면 반복적으로 출력을 재 사용을 한다. CF의 특징인 sparseness를 극복할 수 있다. 이 알고리즘이 학습 속도와 퍼포먼스를 향상 시켰다.

Introduction

추천 시스템은 두가지 카테고리로 분류할 수 있다 : 컨텍스트 기반과 개인화 추천

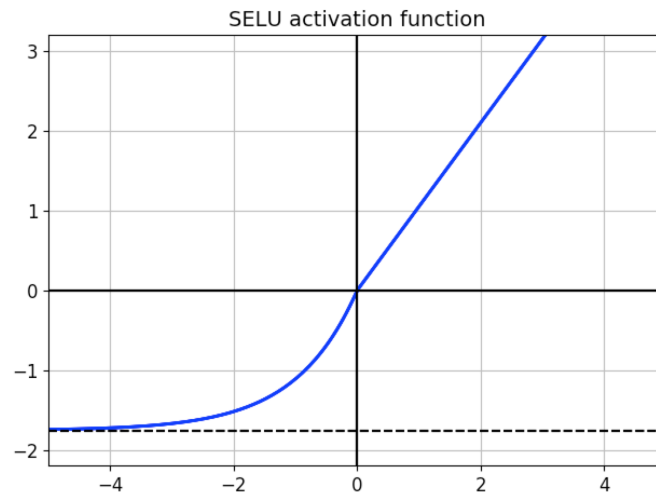
컨텍스트 기반의 추천은 위치, 날짜 그리고 시간과 같은 계정의 컨텍스트 팩터(맥락적인 요소)를 고려한다. 개인화 추천은 CF 접근법을 사용해서 아이템을 유저에게 추천한다. 이런 접근법은 유저의 관심 요소들이 예측된다. 시스템 내부에서 확보한 통계적인 특성과 다른 유저들의 선호도에 기반해서 그리고 "유사도" 라고 불리는 함축적인 정보로 깔려진 가정은 랜덤하게 선택된 두 사람보다 비슷한 특성의 두 사람이 아이템에 대해 같은 의견을 보일 가능성이 높다.

클래식한 CF 문제는 : $m * n$ 매트릭스 내 없는 칸을 추론한다. 이 칸의 인덱스를 (i, j) 라고 하면 i 번째 유저가 j 번째 아이템에 준 점수다. 성능은 RMSE로 측정된다.

Model

유저 기반 오토인코더인 U-AutoRec에서 영감을 받았다. 프리-트레이닝으로 가능하기 위해

a) SELUs (Scaled Exponential Linear Units)



b) 학습 중에 출력을 다시 먹이는 반복

오토인코더의 "목표"는 d 차원의 데이터 표현을 얻는 것이다. 기존 값과 복원 값의 차이를 최소화하는. 만약에 인코딩 과정에 데이터가 노이즈가 추가되면 오토인코더는 디노이징이라고 불린다. 오토인코더는 훌륭한 틀이다. 차원 감소와 PCA의 strict generalization 정도로 고려된다. 비선형 활성화가 없는 오토인코더와 "코드" 레이어만 있는 오토인코더는 학습으로 MSE 로스를 최적화 하면서 인코더가 PCA 변환 처럼 되게 만들 수 있다.

Forward pass and inference

포워드 패스에서 모델은 유저를 입력으로 받는다. 유저 데이터는 트레이닝 셋의 레이팅 벡터로 표현된다. 출력도 다시 레이팅 벡터로 나온다. 출력은 dense하고 모든 아이템에 대한 레이팅 예측을 가진다.

Loss Function



Recommender Model: Loss function



The optimize Masked Mean Squared Error loss (**MMSE**)

$$MMSE = \frac{m_i * (r_i - y_i)^2}{\sum_{i=0}^n m_i}$$

where, ***ri*** is actual rating, ***yi*** is reconstructed and ***mi*** is a mask function such that ***mi* = 1 if *ri* != 0 else *mi* = 0.**



11

로스 함수로 MMSE (Masked Mean Squared Error)를 사용한다. MSE 개념에 마스크 개념을 추가한다. 이 마스크는 아이템에 대한 레이팅이 있으면 1이고 없으면 0이다. 이걸 전체 mse with mask에 전체 아이템 셋에 대한sum(mask)를 나눈다.

Dense re-feeding

어차피 기존 데이터가 추론 데이터와 같아야 한다. $y = f(y)$. fixed point가 되어야 한다.

- (1) 주어진 sparse한 x에서 dense $f(x)$ 와 loss를 계산한다. (forward pass)
- (2) gradient를 계산하고 weight를 업데이트한다 (backward pass)
- (3) 구해진 $f(x)$ 로 $f(f(x))$ 를 계산한다. 이제 $f(x)$ 와 $f(f(x))$ 는 dense하고 모든 mask가 1이 된다. (2nd forward pass)
- (4) gradient를 계산하고 weight를 업데이트한다 (2nd backward pass)