

Neural Collaborative Filtering

Przemysław Rośleń 180150
Jakub Sachajko 179976

Plan prezentacji

1. **Możliwe modyfikacje**
2. **Opisy modyfikacji**
3. **Opis eksperymentu 1**
4. **Opis eksperymentu 2**

Możliwe modyfikacje

1. Przetestować modyfikacje innych parametrów algorytmu, które nie były modyfikowane we wcześniejszych eksperymentach (MLP layers, learning rate, batch size, optimizer)
2. Przetestować modele na innym zbiorze danych (pinterest-20)
3. Powtórzyć eksperymenty wykonane w ramach etapu 2, dla większej liczby epok
4. Użycie innych metryk ewaluacji algorytmu niż HR i NDCG (np. Precision, Recall, Mean Average Precision)
5. Dodanie regularyzacji do warstw MLP oraz do MF
6. Przeprowadzić eksperymenty modyfikując kilka wspomnianych wyżej hiperparametrów jednocześnie

Modyfikacje innych parametrów algorytmu

1. Istota modyfikacji:
 - a. Typ modyfikacji - modyfikacja eksperymentu
 - b. Modyfikacja polega na ewaluacji modeli modyfikując inne parametry niż autorzy artykułu (MLP layers, learning rate, batch size, optimizer)
 - c. Algorytm zostanie wykonany w ten sam sposób, jedyną zmianą będą wyniki dla innych wartości parametrów
 - d. Zmiana będzie występowała w plikach konkretnych algorytmów NCF w dodawaniu argumentów do obiektu `ArgumentParser`
2. Motywacje do wprowadzenia danej modyfikacji:
 - a. Główną motywacją do wprowadzenia tej modyfikacji jest chęć uzyskania lepszych rezultatów algorytmu
 - b. Modyfikacja pozwoli ulepszyć eksperyment. Dzięki niej uzyskamy więcej informacji co do możliwości algorytmów.
Czy wyniki mogą być jeszcze lepsze?
 - c. Dzięki zastosowaniu tej modyfikacji uzyskamy szerszy ogłęd na w jaki sposób zachowuje się algorytm dla różnych parametrów

Testowanie na innym zbiorze danych

1. Istota modyfikacji:

- a. Typ modyfikacji - modyfikacja eksperymentu
- b. Modyfikacja polega na zmianie zbioru danych na np. niewykorzystany w badaniach przez autorów pracy pinterest-20
- c. Algorytm zostanie wykonany w ten sam sposób, jedyną zmianą będą wyniki na nowym zbiorze danych
- d. Zmiana będzie występowała przy podawania parametru dataset do algorytmów, ewentualnie w przypadku potrzeby preprocessingu zmiany powinny zostać zastosowane w pliku dataset.py

2. Motywacje do wprowadzenia danej modyfikacji:

- a. Główną motywacją jest próba kontrolna na innym zbiorze danych, aby sprawdzić czy nie jest to wynik wyłącznie na wybranym datasetcie
- b. Modyfikacja pozwoli ulepszyć eksperyment. Sprawdzenie czy wyniki są zależne od konkretnego zbioru danych czy też nie.

Większa liczba epok

1. Istota modyfikacji:
 - a. Typ modyfikacji - modyfikacja eksperymentu
 - b. Modyfikacja polega na powtórzeniu eksperymentów wykonanych w ramach etapu 2 dla większej liczby epok
 - c. Algorytm niewątpliwie będzie wykonywał się dłużej
 - d. Zmiana będzie występowała w plikach konkretnych algorytmów NCF w dodawaniu argumentów do obiektu `ArgumentParser`
2. Motywacje do wprowadzenia danej modyfikacji:
 - a. Główną motywacją do wprowadzenia tej modyfikacji jest duża szansa na uzyskanie lepszych rezultatów algorytmu
 - b. Modyfikacja najprawdopodobniej pozwoli ulepszyć eksperyment, ponieważ podczas poprzednich eksperymentów z każdą kolejną epoką odnotowywaliśmy coraz lepsze rezultaty algorytmów
 - c. Dzięki zastosowaniu tej modyfikacji uzyskamy wiedzę, po ilu epokach uczenia konkretnych wariantów algorytmu, ulegają one przeuczeniu (nie doświadczyliśmy tego w poprzednich eksperymentach)

Użycie innych metryk ewaluacji algorytmu

1. Istota modyfikacji:
 - a. Typ modyfikacji - modyfikacja eksperymentu
 - b. Modyfikacja polega na dodaniu w procesie ewaluacji metryk Precision, Recall oraz Mean Average Precision
 - c. Sposób wykonania algorytmu nie ulegnie zmianie. Zmianie ulegnie jednak proces ewaluacji algorytmu
 - d. Zmiana będzie miała miejsce w pliku evaluation.py, będzie trzeba zaimplementować wyżej wymienione metryki
2. Motywacje do wprowadzenia danej modyfikacji:
 - a. Główna motywacja - Wprowadzenie MAP może pomóc w optymalizacji algorytmu pod kątem równowagi między dokładnością a kompletnością rekomendacji
 - b. Modyfikacja niewątpliwie pozwoli ulepszyć eksperyment.
 - c. Użycie dodatkowych metryk pozwoli na dokładniejsze dostrojenie algorytmów NCF
 - d. Pozyskamy szerszą wiedzę na temat jakości algorytmów NCF

Dodanie regularyzacji do warstw MLP oraz do MF

1. Istota modyfikacji:
 - a. Typ modyfikacji - modyfikacja eksperymentu
 - b. Modyfikacja polega na zmianie parametrów regularyzacji do MLP oraz MF. (reg_layers, reg_mf)
 - c. Sposób wykonania algorytmu nie ulegnie zmianie. Zmienione zostaną jedynie parametry
 - d. Zmiana będzie występowała przy podawaniu parametrów
2. Motywacje do wprowadzenia danej modyfikacji:
 - a. Główną motywacją jest dodanie warstwy regularyzacji w celu otrzymania bardziej zgeneralizowanych wyników oraz mniejszy overfitting do danych treningowych
 - b. Metoda prawdopodobnie pozwoli ulepszyć eksperyment i da bardziej zgeneralizowane wyniki
 - c. Uzyskamy wiedzę na temat implementacji oraz wpływu regularyzacji

Modyfikacja kilku hiperparametrów jednocześnie

1. Istota modyfikacji:
 - a. Typ modyfikacji - modyfikacja eksperymentu
 - b. Modyfikacja polega na przetestowaniu kilku hiperparametrów jednocześnie.
 - c. Algorytm zostanie wykonany w ten sam sposób, jedyną zmianą będą wyniki dla innych wartości parametrów
 - d. Zmiana będzie występować jedynie przy włączaniu programu i nadpisywaniu podstawowych parametrów.
2. Motywacje do wprowadzenia danej modyfikacji:
 - a. Główną motywacją do wprowadzenia tej modyfikacji jest chęć uzyskania najlepszej możliwej kombinacji parametrów dla danego problemu
 - b. Dzięki tej modyfikacji będziemy w stanie zbadać korelację wpływu zmiany wielu parametrów na wyniki algorytmu
 - c. Dzięki zastosowaniu tej modyfikacji uzyskamy szerszy ogłęd na w jaki sposób zachowuje się algorytm dla różnych parametrów

Opis eksperymentu - Modyfikacje innych parametrów algorytmu

1. Wykorzystamy zbiór danych - MovieLens
2. Kroki eksperymentu:
 - a. Wytrenujemy każdy z trzech algorytmów NCF dla 3 różnych wartości parametrów MLP layers, learning rate, optimizer
 - b. Dokonamy ewaluacji wyników uwzględniając metryki HR oraz NDCG
 - c. Porównamy na wykresach wyniki poszczególnych algorytmów. Rezultaty te porównamy z tymi, uzyskanymi w poprzednich eksperymentach
3. Zmienne niezależne: parametry **MLP layers**, **learning rate**, **optimizer**
Zmienne zależne: metryki **HR** i **NDCG**
4. Nie będzie wymagane wcześniejsze strojenie pozostałych hiperparametrów modelu. Będą one przyjmowały wartości domyślne

Opis eksperymentu - Dodanie regularyzacji do warstw MLP oraz MF

1. Wykorzystamy zbiór danych - MovieLens
2. kroki eksperymentu:
 - a. Zmiana parametrów `reg_layers` (default='[0,0,0,0]') oraz `reg_mf` (default=0). Opcje mogą zostać zmienione na liczby w postaci float z zakresu od 0 do 1 oraz w przypadku `reg_layers` można wybrać dla której warstwy zostanie przeprowadzona regularyzacja.
 - b. Wyniki uzyskane dla danych eksperymentów zostaną porównane i przedstawione na wykresach, w celu porównania.
3. Zmienne niezależne: parametry **reg_layers** oraz **reg_mf**
Zmienne zależne: metryki **HR** i **NDCG**
4. Pozostałe hiperparametry zostaną bazowe, czyli z przykładów na githubie projektu.

Źródła

1. <https://arxiv.org/pdf/1708.05031v2.pdf>
2. https://github.com/hexiangnan/neural_collaborative_filtering

Dziękujemy za uwagę