**xDeepFM: Combining Explicit and Implicit Feature Interactions for Recommender Systems (Mô hình xDeepFM: Sự kết hợp giữa các thuộc tính ẩn và tường minh cho hệ khuyến nghị)**

**Giới thiệu chung về bài báo**

Sự kết hợp giữa các đặc trưng là một phần rất quan trọng trong sự thành công của các mô hình máy học hiện nay. Kết hợp các thuộc tính một cách thủ công (bằng tay) có nhiều hạn chế như có quá nhiều thuộc tính để kết hợp, sự gia tăng nhanh chóng của dữ liệu, khó tìm ra insight dữ liệu trong hệ thống web.

Với sự phát triển mạnh mẽ của các mô hình Deep Learning, các nhà nghiên cứu đã đề xuất mô hình DNN-based factorization, mô hình này có thể học được sự ảnh hưởng của các đặc trưng cấp thấp và cấp cao (vd như cấp thấp là sự kết hợp của 2 thuộc tính, cấp cao là 3,4,5.. thuộc tính). Dù cho có khả năng học bất kì hình thái nào của dữ liệu, mô hình DNN-based factorization chỉ tạo ra các feature 1 cách ngầm định và ở mức bit-wise.

Trong bài báo này, tác giả đề xuất mạng **Compressed Interaction Network** (CIN). Tác giả kết hợp CIN với mô hình Deep Learning cổ điển thành 1 mô hình chung và đặt tên là eXtreme Deep Factorization Machine (xDeepFM).

Tác giả đã tiến hành các thử nghiệm trên ba bộ dữ liệu và thấy rằng xDeepFM hoạt động tốt hơn các mô hình hiện đại.

**Giới thiệu chung về bài toán, các hướng tiếp cận**

Các thuộc tính đóng 1 vai trò rất quan trọng trong sự thành công của các mô hình Hệ khuyến nghị. Bởi vì các thuộc tính ban đầu hiếm khi cho kết quả tốt, do đó các nhà Khoa học dữ liệu (DS) thường dành nhiều thời gian trong việc biến đổi dữ liệu để có được kết quả tốt hơn.

Một trong các phép biến đổi đặc trưng quan trọng là **cross-product transformation over categorical (cross-feature),** phép biến đổi này đo lường sự tương tác giữa các đặc trưng ban đầu.

Ví dụ như: 3-way feature *AND(user\_organization=UIT,item\_category=deeplearning, time=monday)* có giá trị 1 nếu người dùng làm tại UIT, đọc bài báo liên quan Deep learning vào thứ 2.

Có 3 nhược điểm đối với kĩ thuật **cross-feature:**

* Để có được các đặc trưng tốt thì phải đi kèm với chi phí cao. Bởi vì các đặc trưng phù hợp thường dành riêng cho 1 lĩnh vực(domain) cụ thể
* Trong các hệ thống dự đoán quy mô lớn (hệ khuyến nghị trên web), có nhiều đặc trưng thô khiến việc trích xuất tất cả các đặc trưng chéo (2-way feature) theo cách thủ công là không khả thi.
* Các cross-feature được tạo thủ công không khái quát hóa.
* Do đó mô hình có thể tự học được tương tác giữa các đặc trưng mà không sử dụng các kĩ thuật thủ công được nêu trên là điều cần thiết.

Các bài báo khác đã đề xuất đến các mô hình như:

Factorization Machine (FM): có thể học được bất kì đặc trưng cấp cao nào. Sự tổ hợp các đặc trưng có thể có lợi hoặc gây nhiễu, các đặc trưng gây nhiễu có thể làm giảm độ chính xác của mô hình

Factorisation-machine supported Neural Network (FNN) và Product-based Neural Network (PNN): có thể học các đặc trưng cấp cao nào, nhưng 2 mô hình này tập trung quá nhiều vào các đặc trưng cấp cao và bỏ qua sự chú ý đối với đặc trưng cấp thấp.

Wide&Deep và DeepFM: có thể giải quyết vấn đề của 2 mô hình trên, tuy nhiên mô hình này không đưa ra rõ ràng số bậc cụ thể trong quá trình kết hợp các đặc trưng.

**Giới thiệu chung về phương pháp, kỹ thuật áp dụng**

Với các vấn đề được nêu ở trên, tác giả đề xuất mạng **Compressed Interaction Network (CIN)** có thể học được sự tương tác giữa các đặc trưng một cách rõ ràng, và số bậc của đặc trưng được tăng cùng với độ sâu của mạng.

Diagram

Description automatically generated

Dựa trên mô hình Wide&Deep và DeepFM, tác giả kết hợp lợi ích từ tương tác rõ ràng giữa các đặc trưng bậc cao (**CIN**) với tương tác ngầm (DNN) và FM truyền thống, đồng thời đặt tên cho mô hình là **eXtreme Deep Factorization Machine (xDeepFM)**.

Diagram

Description automatically generated

Kết quả thu được:

Table

Description automatically generated

**Giới thiệu chung về bộ dữ liệu**

**Criteo Dataset:**

Đây là bộ dữ liệu nổi tiếng phát triển và đánh giá các mô hình dự đoán tỷ lệ nhấp chuột vào quảng cáo. Với một người dùng và trang đang được truy cập, mục tiêu là dự đoán khả năng anh ta sẽ nhấp vào một quảng cáo nhất định.

Here is a detailed description of the fields (they are comma-separated in the file):

f0, f1, f2, f3, f4, f5, f6, f7, f8, f9, f10, f11: feature values (dense, float)

treatment: treatment group (1 = treated, 0 = control)

conversion: whether a conversion occured for this user (binary, label)

visit: whether a visit occured for this user (binary, label)

exposure: treatment effect, whether the user has been effectively exposed (binary)

https://www.kaggle.com/datasets/mrkmakr/criteo-dataset

**Dianping Dataset**

Với hồ sơ của người dùng, thuộc tính của nhà hàng (rating, ..) và ba địa điểm được ghé gần đây nhất của người dùng, mục tiêu là dự đoán khả năng anh ta sẽ ghé thăm nhà hàng.

**B**