

HƯỚNG DẪN CÀI ĐẶT CHƯƠNG TRÌNH

Phần cài đặt chương trình trên các tập dữ liệu với nội dung mô tả ngắn MovieLens 1M, MovieLens 10M, MovieLens 20M và tập dữ liệu với mô tả trung bình CiteUlike. Đối với 3 bộ dữ liệu với mô tả ngắn thì nội dung được sử dụng để biểu diễn thông tin items bao gồm tiêu đề và loại phim. Ngoài ra sử dụng thêm bộ dữ liệu từ những từ dự án Glove trên tập từ điển tiếng Anh với dữ liệu từ nguồn Wikipedia.

Để chạy chương trình yêu cầu máy tính có cấu hình khá ổn định và cần cài đặt các thư viện:

- Tensorflow 1.4
- Python 2
- Numpy

Với ban đầu hệ số Gauss biểu diễn nội dung item là 0.01 , hệ số cho tiên nghiệm Poisson là 0.01 với các tỷ lệ dropout chạy từ 0.0, 0.1, 0.2, 0.3

Để chạy mô hình phân rã ma trận sử dụng Poisson sử dụng bộ tri thức tiên nghiệm những của từ được cài đặt trong thư mục poisson_nn_rec ta chạy thông qua các file shell như sau:

- File run_(dataset_name)_bash_5x.sh: dùng để training mô hình, kết quả sẽ được hai ma trận tương ứng là các vector ẩn của mỗi user và item được lưu trong thư mục results.
- File run_eval_(dataset_name)_bash.sh: dùng để đánh giá mô hình, sử dụng kết quả đã train được ở trên để đưa ra gợi ý và sau đó đánh giá mô hình.