

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



NGUYỄN BÁ QUYỀN

**NGHIÊN CỨU PHÂN TÍCH HÀNH VI MUA HÀNG CỦA
KHÁCH HÀNG SỬ DỤNG MẠNG NƠ-RON**

CHUYÊN NGÀNH: KHOA HỌC MÁY TÍNH

MÃ SỐ:

TÓM TẮT LUẬN VĂN THẠC SỸ KỸ THUẬT

(Theo định hướng ứng dụng)

HÀ NỘI – 2021

Luận văn được hoàn thành tại:

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Người hướng dẫn khoa học: **TS. PHẠM HOÀNG DUY**

Phản biện 1:

.....

Phản biện 2:

.....

Luận văn sẽ được bảo vệ trước Hội đồng chấm luận văn thạc sĩ tại
Học viện Công nghệ Bưu chính Viễn thông

Vào lúc:giờ ngày..... tháng Năm

MỞ ĐẦU

Dự đoán hành vi của khách hàng trong tương lai là một nhiệm vụ quan trọng để mang lại cho họ trải nghiệm tốt nhất có thể và cải thiện sự hài lòng của họ. Một ví dụ thực tế được quan sát thấy trong các hệ thống thương mại điện tử, nơi người dùng có thể tránh việc tìm kiếm thông qua một danh mục sản phẩm không thực sự cần thiết và thay vào đó họ có một bộ sản phẩm được đề xuất đáp ứng được điều họ quan tâm. Hành vi của người tiêu dùng có thể được biểu diễn dưới dạng dữ liệu tuần tự mô tả các tương tác qua thời gian, ví dụ về những tương tác này là các mặt hàng mà người dùng mua hoặc xem. Do đó, lịch sử tương tác của người dùng có thể được mô hình hóa dưới dạng dữ liệu tuần tự có đặc điểm cụ thể và có thể kết hợp với nhau thông qua khía cạnh thời gian. Để kiểm tra, nếu người dùng mua điện thoại di động mới, họ có thể mua phụ kiện cho điện thoại di động này trong tương lai gần hoặc nếu người dùng mua sách, họ có thể quan tâm đến sách của cùng tác giả hoặc thể loại. Để đưa ra dự đoán chính xác là điều rất quan trọng. Một cách phổ biến để xử lý dữ liệu này là xây dựng các tính năng thủ công để tổng hợp thông tin từ các bước trong quá khứ. Ví dụ: người ta có thể đếm số lượng sản phẩm đã mua của một danh mục cụ thể trong N ngày qua hoặc số ngày kể từ lần mua cuối cùng. Việc tạo một số tính năng được làm thủ công sẽ tạo ra một vectơ đặc trưng có thể được đưa vào một thuật toán học máy như hồi quy logistic. Mặc dù có thể đạt được kết quả tốt với phương pháp luận này, nhưng nó có một số mặt hạn chế. Đầu tiên, một phần của mối quan hệ thời gian và trình tự bị bỏ qua. Mặc dù chúng có bao gồm các tính năng chứa thông tin từ các tương tác trong quá khứ nhưng trên thực tế vẫn có thể bao gồm tất cả thông tin có trong dữ liệu thô. Chỉ các tín hiệu được mã hóa trong các tính năng này mới có thể được các mô hình dự đoán ghi lại. Thứ hai, thông thường sẽ có một tập hợp rất lớn các tính năng được tạo thủ công bằng tay. Các nhà khoa học dữ liệu có thể dành nhiều thời gian để thiết kế và thử nghiệm các tính năng mới, mà nhiều tính năng trong số đó dẫn đến không cải thiện hiệu suất dự đoán. Ngay cả khi họ có thể cải thiện, rất khó để biết liệu tập hợp các tính năng thủ công thực tế có tối ưu cho vấn đề hay không, vì vậy quá trình thử nghiệm và thêm các tính năng thủ công mới không bao giờ dừng lại hoặc dừng lại khi thuật toán đạt mức chấp nhận được. mức hiệu suất có thể khác xa so với tiềm năng thực sự. Thứ ba, trong một số trường hợp, việc tính toán các tính năng thủ công có thể dẫn đến việc xử lý trước dữ liệu tốn kém.

Với việc học sâu nhận được rất nhiều sự chú ý trong những năm qua, một cách tiếp cận mới đối với dữ liệu tuần tự của mô hình đã được khám phá. Mạng thần kinh tái tạo (RNN) rất năng động để học các mẫu tuần tự phức tạp, vì chúng có khả năng duy trì trạng thái ẩn được cập nhật bởi một hàm phi tuyến tính phức tạp được học từ chính dữ liệu. Họ có thể nắm bắt thông tin về sự phát triển của những gì đã xảy ra trong các bước thời gian trước đó. Trong những năm qua, RNN đã đạt được trình độ tiên tiến trong các vấn đề như lập mô hình ngôn ngữ, nhận dạng giọng nói, dịch máy hoặc nhận dạng chữ viết tay, các tác vụ này có một số điểm tương đồng với bài toán dự đoán các hành động trong tương lai từ dữ liệu tương tác trong quá khứ, theo nghĩa là dữ liệu được biểu diễn tuần tự.

Luận văn này là nghiên cứu các kỹ thuật khác nhau khi sử dụng RNN để dự đoán hành vi của khách hàng trong tương lai. Cụ thể hơn, tập trung vào hai khía cạnh: Nghiên cứu thuật toán nhúng có thể được sử dụng để tạo ra các đại diện mục vector hữu ích giúp cải thiện các dự đoán với RNN. Tiếp theo đánh giá và phân tích các biểu diễn vector của các lựa chọn thay thế khác nhau để tìm hiểu cách nhúng mục. Nghiên cứu cách các cơ chế chú ý có thể giúp giải thích các dự đoán của các mô hình RNN. Sau đó phân tích hiệu suất của các biến thể cơ chế chú ý khác nhau và cung cấp các ví dụ trong đó các dự đoán được giải thích bằng các nguyên tắc trong quá khứ.

CHƯƠNG 1. PHÂN TÍCH HÀNH VI NGƯỜI DÙNG

Trong chương này, chúng ta tập trung xác định các thành phần của quá trình ra quyết định mà người tiêu dùng tuân theo, các mô hình chính được giới thiệu trong nghiên cứu hành vi người tiêu dùng, nghiên cứu về lý thuyết cho việc đưa ra quyết định. Ngoài ra, trong chương này, tập trung vào việc xác định vấn đề cần nghiên cứu và phạm vi nghiên cứu. Và nêu những khó khăn gặp phải trong quá trình nghiên cứu và cách khắc phục

1.1. Khái quát về hành vi người tiêu dùng.

1.1.1. Giới thiệu chung.

Bất kỳ người nào tham gia vào quá trình tiêu dùng đều là người tiêu dùng. Người tiêu dùng là cá nhân mua để tiêu dùng cá nhân hoặc để đáp ứng nhu cầu tập thể của gia đình và nhu cầu hộ gia đình. Hành vi của người tiêu dùng có nghĩa là cách các cá nhân đưa ra quyết định sử dụng các nguồn lực sẵn có của họ như thời gian, tiền bạc, nỗ lực để tiêu dùng các sản phẩm và dịch vụ khác nhau. Nó bao gồm những gì họ mua, tại sao họ mua nó, khi họ mua nó, họ mua nó ở đâu, tần suất mua nó và tần suất sử dụng nó. Hành vi của người tiêu dùng là những hành động mà một người thực hiện trong việc mua và sử dụng các sản phẩm và dịch vụ, bao gồm các quá trình tinh thần và xã hội diễn ra trước và sau các hành động này. Leon G.Schiffman và Leslie Lazar Kanuk đã định nghĩa hành vi của người tiêu dùng “là hành vi mà người tiêu dùng thể hiện khi tìm kiếm, mua, sử dụng, đánh giá và loại bỏ các sản phẩm, dịch vụ và ý tưởng mà họ mong đợi sẽ thỏa mãn nhu cầu của họ”.

1.1.2. Các yếu tố ảnh hưởng đến hành vi tiêu dùng

Việc mua hàng của người tiêu dùng bị ảnh hưởng mạnh mẽ bởi các đặc điểm văn hóa, xã hội, cá nhân và tâm lý.

1.1.2.1. Yếu tố văn hóa.

1.1.2.2. Yếu tố xã hội

1.1.2.3. Yếu tố cá nhân.

1.1.2.4. Yếu tố tâm lý.

1.1.3. Đóng góp của luận văn và các kỹ thuật liên quan.

Đóng góp chính của luận văn này là nghiên cứu các kỹ thuật khác nhau khi sử dụng

RNN để dự đoán hành vi của khách hàng trong tương lai. Cụ thể hơn, luận văn tập trung vào hai khía cạnh:

- Nghiên cứu phép nhúng có thể được sử dụng để tạo ra các biểu diễn mục vector hữu ích giúp cải thiện các dự đoán với RNN và đánh giá, phân tích biểu diễn vector của các lựa chọn thay thế khác nhau để tìm hiểu cách nhúng mục.
- Nghiên cứu cách các cơ chế chú ý có thể giúp giải thích các dự đoán về Các mô hình RNN và phân tích hiệu suất của các cơ chế chú ý khác nhau cách biến thể và cung cấp các ví dụ trong đó các dự đoán được giải thích bằng các tương tác trong quá khứ.

Để thực hiện nghiên cứu, luận văn tập trung vào việc nghiên cứu các phép nhúng và cơ chế chú ý.

1.2. Phạm vi công việc nghiên cứu.

1.2.1. *Khái quát vấn đề.*

Phân khúc thị trường là một quá trình đòi hỏi xác định các loại nhóm người tiêu dùng đồng nhất được mô tả bởi một tập hợp các đặc điểm tương đồng, để cải thiện các hoạt động tiếp thị thông qua việc phân bổ nguồn lực và xây dựng chiến lược tùy biến tốt hơn. Khi các nhóm mục tiêu được biết đến trước, vấn đề sẽ trở thành một nhiệm vụ phân loại, theo một quá trình học tập có giám sát. Sự quan tâm ngày càng tăng trong việc xác định các nguồn khách hàng mới, buộc các tổ chức tài chính phải điều tra các phương pháp mới để phát hiện các cá nhân có xu hướng tiết kiệm tiền cao, và sử dụng trong việc chi tiêu. Các phương pháp thống kê truyền thống như phân tích phân biệt thường được sử dụng trong các nhiệm vụ phân loại, mang lại kết quả tốt. Tuy nhiên, nhu cầu thu được kết quả chính xác hơn nữa đã khiến các nhà nghiên cứu quan tâm đến các kỹ thuật phân loại không tham số như mạng lưới thần kinh nhân tạo. Mục đích chính của nghiên cứu này là phân tích kết quả thu được khi xây dựng mô hình xác định các cá nhân có cơ hội lớn để thực hiện việc mua hàng, sử dụng mạng lưới thần kinh nhân tạo.

1.2.2. *Mục đích nghiên cứu.*

Luận văn này nghiên cứu phép nhúng được sử dụng để tạo ra các đại diện mục vector hữu ích giúp cải thiện các dự đoán với RNN. Luận văn sẽ trình bày việc đánh giá và phân tích các biểu diễn vector của các lựa chọn thay thế khác nhau để tìm hiểu cách

nhúng mục. Ngoài ra, trong luận văn còn nghiên cứu cách các cơ chế chú ý có thể giúp giải thích các dự đoán của các mô hình RNN. Thêm nữa, luận văn này còn trình bày việc phân tích hiệu suất của các biến thể cơ chế chú ý khác nhau và cung cấp các ví dụ trong đó các dự đoán được giải thích bằng các nguyên tắc trong quá khứ.

1.2.3. Đối tượng và phạm vi nghiên cứu.

Luận văn tập trung vào nghiên cứu bài toán phân tích hành vi người dùng tập trung vào hành vi mua hàng nhằm xác định các cá nhân có cơ hội lớn để thực hiện việc mua hàng, sử dụng mạng lưới thần kinh nhân tạo.

1.2.4. Phương pháp nghiên cứu.

Nghiên cứu đến dữ liệu có sẵn, quy trình dọn dẹp và phương pháp phân vùng được sử dụng, mô tả các bước cấu hình được thực hiện để xây dựng các mô hình. Phân tích các kết quả thu được và kết luận về việc sử dụng mạng Nơ-ron trong ứng dụng phân tích hành vi của người tiêu dùng, đề xuất các hướng nghiên cứu tiếp theo.

CHƯƠNG 2. MẠNG RNN VÀ KỸ THUẬT PHÂN TÍCH

Trong chương này, luận văn sẽ trình bày nghiên cứu về các lý thuyết phục vụ cho luận văn, như mạng RNN, các phép nhúng tuyến tính và phi tuyến tính, các cơ chế chú ý tuyến tính và phi tuyến tính nhằm phân tích và đưa ra các dự đoán về hành vi của người dùng.

2.1. Mạng neural và các vấn đề cơ bản.

2.1.1. Tổng quan về mạng lưới thần kinh

Cách tiếp cận thông thường đối với tính toán dựa trên một tập hợp các hướng dẫn được lập trình rõ ràng và ngày tháng từ công việc của Babbage, Turing và von Neumann. Mạng nơ-ron đại diện cho một mô hình tính toán thay thế trong mà giải pháp cho một vấn đề được học từ một tập hợp các ví dụ. Nguồn cảm hứng cho mạng lưới thần kinh ban đầu xuất phát từ các nghiên cứu về cơ chế xử lý thông tin trong hệ thần kinh sinh học, đặc biệt là não người.

2.1.1.1. Mạng lưới thần kinh sinh học

2.1.1.2. Mạng lưới thần kinh nhân tạo

2.2. Phương pháp nghiên cứu.

2.2.1. RNN đường cơ sở

Đầu tiên bắt đầu với mô hình RNN không có thuật toán nhúng và cơ chế chú ý, được gọi là RNN-baseline.

2.2.2. Các phương pháp nhúng

Trong phần này, sẽ mô tả các mô hình được sử dụng để trả lời câu hỏi nghiên cứu RQ1. Sử dụng RNN-Baseline để làm cơ sở, đầu tiên tạo ra một số biến giá trị bằng cách thêm phép nhúng vào mô hình. Bảng 2.1 tóm tắt các phương pháp nhúng khác nhau đã được liệt kê và cung cấp một mô tả ngắn gọn.

2.2.2.1. Phép nhúng được đào tạo riêng với Word2vec

2.2.2.2. Phép nhúng được tìm hiểu với mô hình phân loại

2.2.2.3. Phép nhúng được tìm hiểu học riêng biệt và sau đó được điều chỉnh chung.

2.2.2.4. Dự đoán các phép nhúng. Phép nhúng được tìm hiểu riêng

2.2.2.5. Cơ chế chú ý đến các trạng thái ẩn RNN

2.2.2.6. Cơ chế chú ý tới các phép nhúng

2.3. Xử lý trước dữ liệu tuần tự

Khi xử lý dữ liệu tuần tự, chúng ta có thể nghĩ nhiều cách khác nhau để tạo các mẫu dữ liệu.

- Tạo một mẫu cho mỗi trình tự hoàn chỉnh, xem xét việc mất điểm cuối cùng cố vấn.
- Tạo một mẫu cho mỗi chuỗi hoàn chỉnh, nhưng xem xét phần trung gianments.
- Tạo một mẫu cho mọi tiền tố trước của dãy.

CHƯƠNG 3. THỰC NGHIỆM VÀ ĐÁNH GIÁ.

Trong phần này, tập trung vào việc trình bày các tập dữ liệu được sử dụng trong việc phân tích, thiết lập công cụ và các thử nghiệm được thực hiện và các chỉ số đánh giá để đo lường hiệu suất của các mô hình.

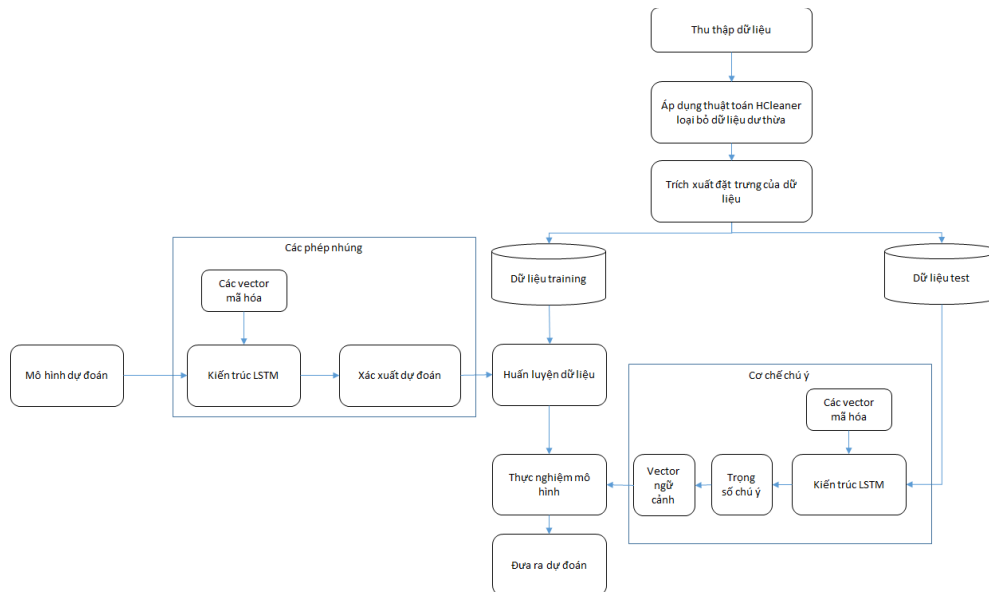
3.1. Bộ dữ liệu

Tập dữ liệu thứ hai được sử dụng là tập dữ liệu danh sách các bộ phim được bán ra và đánh giá của người dùng từ năm 2008 đến năm 2015. Tập dữ liệu bao gồm lịch sử xếp hạng các bộ phim do người dùng khác nhau đánh giá. Bộ dữ liệu xếp hạng chứa một giá trị là mốc thời gian, cho biết thứ tự mà người dùng xếp hạng các bộ phim. Trong tập dữ liệu này, các bộ phim được xếp hạng theo điểm số từ 1 đến 5. Tuy nhiên, trong luận văn này chỉ sử dụng các dữ liệu đã người dùng đã xếp hạng một bộ phim chứ không phải điểm số.

3.2. Cách thức thực nghiệm và đánh giá.

3.2.1. Cách thức thực nghiệm.

Trong luận văn này sẽ triển khai việc các phép nhúng và các cơ chế chú ý cho các mô hình bằng cách sử dụng bộ framework TensorFlow^[1], cho phép tạo và đào tạo các mạng nơ-ron sâu mà không cần cung cấp các biểu thức của gradient, vì chúng được tính toán tự động. Đối với các mô hình mà các lần nhúng được tính bằng phương pháp Skip-gram, mã nguồn sử dụng gói python Gensim^[33].



Hình 3.1. Phương pháp thực nghiệm đề xuất.

3.2.2. Cách thức đánh giá.

Trong trường hợp này, tập trung đánh giá các dự đoán ngắn hạn (phim đầu tiên được xếp hạng trong giai đoạn thử nghiệm) và dự đoán dài hạn (tất cả các phim được xếp hạng trong giai đoạn thử nghiệm). Với một chuỗi đầu vào của các bộ phim được xếp hạng trước thời gian thử nghiệm, sẽ có được xác suất dự đoán cho mỗi bộ phim được xếp hạng tiếp theo. Có thể sắp xếp từng phim theo xác suất được xếp hạng dự đoán. Sau đó, có thể đo lường hiệu suất của mô hình bằng "dự đoán trình tự thành công tại k" ($sps@k$). Biện pháp này đã được giới thiệu trong [6] và đo lường mức độ tốt của mô hình trong các dự đoán ngắn hạn. Đối với một người dùng nhất định, $sps@k$ là 1 nếu phim được xếp hạng đầu tiên trong tập hợp thử nghiệm nằm trong top k được dự đoán là phim và 0 nếu ngược lại. Cụ thể, $sps@10$ và các kết quả được báo cáo trong chương 3 tương ứng với mức trung bình của tất cả những người dùng được sử dụng.

3.3. Cài đặt phần mềm.

3.3.1. Tổng quan phần mềm.

Trong nghiên cứu này, sử dụng bộ framework Tensorflow, cho phép tạo và đào tạo các mạng neural thông qua việc tính toán tự động. Kèm theo đó sử dụng phương pháp Skip-gram thông qua việc sử dụng package Gensim. Phần tiếp theo sẽ trình bày chi tiết về hướng dẫn cài đặt phần mềm và chạy thử nghiệm.

3.3.2. Cài đặt framework Tensorflow.

3.3.3. Cài đặt package Gensim

3.4. Thực nghiệm mã nguồn và đánh giá kết quả

3.4.1. Thực nghiệm mã nguồn.

Trong phần này, luận văn sẽ trình bày cách thiết lập mã nguồn để kiểm tra và đánh giá kết quả của các phép nhúng và áp dụng các cơ chế chú ý trong việc dự đoán hành vi người dùng.

3.4.1.1. Import các thư viện của Python.

Như đã đề cập trong các phần trước, phần mã nguồn này sẽ dùng các thư viện đã được đề cập trong phần cài đặt đã nêu trước đó.

Cụ thể, mã nguồn sẽ sử dụng thư viện NumPy để xử lý dữ liệu tính toán ở mức độ cao, Pandas để phân tích dữ liệu và thực thi việc phân tích và Matplotlib để mô hình hóa dữ liệu.

```
import pandas as pd
import numpy as np
from gensim.models import Word2Vec
import random
from tqdm import tqdm
import matplotlib.pyplot as plt
%matplotlib inline
import warnings;
warnings.filterwarnings('ignore')
```

3.4.1.2. Đọc dữ liệu nguồn.

Như đã đề cập trước đó, bộ dữ liệu sẽ được sử dụng là Movielens, chứa hơn 20 triệu dòng dữ liệu đánh giá với hơn 465,000 đánh giá của hơn 27 nghìn bộ phim đánh giá bởi hơn 138 nghìn người xem.

```
df_movies = pd.read_csv('movies.csv')
df_ratings = pd.read_csv('ratings.csv')
```

3.4.1.3. Gộp dữ liệu và loại bỏ dữ liệu trống.

Hợp nhất dữ liệu có nghĩa là kết hợp hai tập dữ liệu theo cách mà mỗi hàng trong cả hai tập dữ liệu đều căn chỉnh dựa trên các thuộc tính hoặc cột chung. Ở đây, chúng ta sẽ hợp nhất bộ dữ liệu phim và xếp hạng để lấy ID phim, ID người dùng và tiêu đề phim trong một khung dữ liệu. Và loại bỏ những dòng dữ liệu có chứa các giá trị trống.

```
df = pd.merge(df_movies, df_ratings)
df.dropna(inplace=True)
```

3.4.1.4. Tiền xử lý dữ liệu.

Mục đích của việc này là làm mịn dữ liệu trước khi tiến hành phân tích. Đầu tiên, ID sẽ được thay đổi định dạng sang kiểu chuỗi và loại bỏ những UserID trùng nhau.

```
df['movieId'] = df['movieId'].astype(str)
users = df["userId"].unique().tolist()
len(users)
```

Sau khi là mịn ta thu được hơn 162 nghìn dữ liệu của người dùng, mỗi người dùng đều có lịch sử xem phim riêng.

3.4.1.5. Phân tách dữ liệu.

Để đáp ứng việc thử nghiệm performance của các mô hình, dữ liệu ban đầu cần được tách và sử dụng theo công thức 90% dữ liệu training và 10% dữ liệu thử nghiệm.

```
random.shuffle(users)
# extract 90% of user ID's
```

```

users_train = [users[i] for i in
range(round(0.9*len(users)))]
# split data into train and validation set
train_df = df[df['userId'].isin(users_train)]
validation_df = df[~df['userId'].isin(users_train)]

```

3.4.1.6. Chiến lược thực hiện.

Trong phần này, mã nguồn sẽ được thay đổi tùy theo phép nhúng hay cơ chế chú ý nào sẽ được áp dụng, trong mã nguồn phía dưới, mục đích để đưa ra các gợi ý về các bộ phim nên xem cho người xem sử dụng phép nhúng RNN-Baseline:

```

#list to capture watch history of the users
watch_train = []
# populate the list with the movie ID
for i in tqdm(users_train):
    temp = train_df[train_df["userId"] ==
i]["movieId"].tolist()
    watch_train.append(temp)

```

3.4.1.7. Đào tạo cho các mô hình.

Để đào tạo các mô hình, mã nguồn sẽ sử dụng thư viện gensim. Module này chứa bộ thư viện Word2Vec, được sử dụng để xử lý lượng dữ liệu lớn và hỗ trợ giao diện Pythonic, mã nguồn bên dưới áp dụng RNN-emb-Word2vce:

```

model = Word2Vec(window = 10, sg = 1, hs = 0, negative = 10,
alpha=0.03, min_alpha=0.0007, seed = 14)
model.build_vocab(watch_train, progress_per=200)
model.train(watch_train, total_examples =
model.corpus_count, epochs=10, report_delay=1)

```

Tiếp theo có thể in ra mô hình để kiểm tra nếu cần.

```

X = model[model.wv.vocab]
X.shape

```

3.4.2. Đánh giá kết quả.

Trong phần đánh giá này, luận văn sẽ làm rõ hai kết quả đã cho mục tiêu được đặt ra trong chương 1. Thứ nhất, chúng ta có thể thấy rằng thông qua việc sử dụng các phép nhúng dựa trên word2vec sẽ giúp cho việc tối ưu hóa vector dữ liệu tương tác của người dùng để phục vụ cho việc dự đoán hành vi trong tương lai. Trong các ví dụ trong phần 3.4.2.1, chúng ta sẽ thấy rõ hơn việc các phép nhúng sử dụng word2vec sẽ kết hợp thông tin về thời gian phát hành bộ phim, thể loại hoặc sagas cụ thể. Điều này có nghĩa là có thể có một số người dùng xem liên tiếp các bộ phim cùng thời kỳ, cùng thể loại hoặc cùng một câu chuyện.

Thứ hai, các cơ chế chú ý sử dụng các trong số chú ý đã được thử nghiệm góp phần đưa ra các dự đoán hành vi của người dùng trong tương lai dựa vào các tương tác trong quá khứ. Trong phần phân tích tại mục 3.4.2.2 sẽ cho chúng ta thấy việc dự đoán hành vi của người dùng trong tương lai cho bộ dữ liệu cụ thể đã đề cập ở phần trước và hỗ trợ việc đưa ra các chiến lược phù hợp để tiếp cận người dùng tốt hơn.

Cần lưu ý rằng việc dự đoán hành vi người dùng và việc giới thiệu phim tới người dùng có một điểm chung là đưa ra các dự đoán của người dùng trong tương lai để từ đó có những cơ chế quảng cáo, tiếp thị và kinh doanh phù hợp, tránh đi ngược lại xu hướng của người tiêu dùng. Tuy nhiên có một điểm khác biệt giữa hai hình thức này là phân tích hành vi người dùng mang ý nghĩa dự đoán rộng và phức tạp hơn so với giới thiệu phim tới người dùng.

3.4.2.1. *Đánh giá kết quả việc phân tích phép nhúng*

Trong phần này, luận văn sẽ trình bày kết quả phân tích cho các phép nhúng thu được bằng các phương pháp khác nhau. Mục đích là để quan sát các tính năng nào được thể hiện trong các mục nhúng.

3.4.2.2. *Phân tích tính tập trung*

Như đã đề cập, trọng số chú ý có thể cung cấp mức độ tập trung của mô hình vào từng phần tử của chuỗi để đưa ra dự đoán. Do đó, chúng có thể rất hữu ích để giải thích các dự đoán. Trong phần này, một số ví dụ về các dự đoán và trọng số chú ý cho các phương pháp khác nhau sẽ được đưa ra.

KẾT LUẬN VÀ ĐỊNH HƯỚNG NGHIÊN CỨU

Luận án này nghiên cứu việc sử dụng mạng thần kinh hồi quy để dự đoán hành vi của khách hàng từ dữ liệu tương tác trong quá khứ. RNN đã chứng tỏ tình hữu dụng trong việc lập mô hình dữ liệu tương tác của khách hàng để dự đoán các mặt hàng mà người dùng sẽ mua trong tương lai. Việc sử dụng kỹ thuật nhúng đã cải thiện đáng kể hiệu năng mô hình RNN đơn giản trong các dự đoán ngắn hạn và dài hạn.

Các phương pháp nhúng khác nhau có thể biểu diễn các đặc trưng khác nhau mà người tiêu dùng quan tâm như thời gian bộ phim được phát hành, thể loại hoặc câu chuyện. Kỹ thuật Word2vec có ưu thế rõ ràng hơn một chút, nhưng các phương pháp khác có thể đạt được hiệu suất tương tự với mục tiêu này. Nghiên cứu trong tương lai sẽ là khám phá các phương pháp để học cách nhúng bằng cách sử dụng dữ liệu mục (ví dụ như năm của bộ phim, đạo diễn, v.v.) hoặc sử dụng phép nhúng biểu đồ tri thức.

Cơ chế chú ý là một công cụ hữu ích để giải thích các dự đoán được thực hiện bởi các mô hình. Trong các thử nghiệm với tập dữ liệu MovieLens, sự chú ý đến các nhúng có trọng số tuyến tính cung cấp một bộ phim hoặc một vài bộ phim quan trọng nhất đối với mô hình để đưa ra dự đoán. Đặc biệt kỹ thuật này cho phép cải thiện hiệu suất dự đoán dài hạn.

Một lĩnh vực nghiên cứu quan trọng sẽ là làm thế nào để biểu diễn thời gian điện tử trong RNN. RNN có thể có ưu thế để biểu diễn dữ liệu tuần tự như trong các nghiên cứu về xử lý ngôn ngữ tự nhiên, nơi mà khái niệm thời gian không tồn tại (không có thời gian giữa các từ, chúng chỉ thường xuyên được lấy mẫu lần lượt). Tuy nhiên, trong các nhiệm vụ như biểu diễn các tương tác của khách hàng, có một khoảng thời gian giữa mỗi lần tương tác. Thời gian này có thể chứa thông tin quan trọng, vì không giống nhau nếu một lần tương tác xảy ra sau lần trước đó vài giây so với lần tương tác xảy ra sau một năm. Chúng ta đã thấy rằng RNN không chứa thông tin này và chúng chỉ có thể đại diện cho thứ tự của chuỗi các tương tác.