Proceedings of IYSC, (2022), vol. 11

Journal homepage: http://journals.sdu.edu.kz/index.php/iysw

2022 International Young Scholars' Conference

DESIGNING A RECOMMENDATION SYSTEM FOR SPECIALIZED COURSES FOR THE UNIVERSITY

D. Anefiyayeva[1], A. Talasbek[2]

[1,2]Suleyman Demirel University, Kaskelen, Kazakhstan

Abstract

All universities have compulsory courses in, some of which students have to choose for themselves. This choice will affect the student's further career as they choose not only the subject they are studying, but what they will do soon. I know how difficult it is to choose a topic for myself. The recommendation engine is one of the most popular artificial intelligence applications. Since switching to the Internet, recommender systems have become widely used in everyday life without even being noticed. Many machine learning algorithms can used to implement recommender systems, but this task attracts many researchers from around the world. First, let's consider how to classify k-means.

Keywords: Recommendation system, ml, k-means, subjects, academies

Introduction

Recommendation systems (RS) are widely used in many areas of modern life. They can also be used in education. Universities with a credit system have compulsory and elective courses. The courses, that students must choose are named compulsory, elective courses must be taken by students themselves. Compulsory subjects are primarily aimed at completing research. Electives, on the other hand, are aimed at a narrower discipline. This choice will affect your entire life as the course you complete will determine your future work. From this, we can conclude that this decision is one of the most important. Create an automated system based on this. Of course, the recommendations within the university wall are appropriate. It analyzes and emphasizes the main aspects that influence the successful delivery of the subject and recommends selected students based on this. The main goal of my research is to create a recommendation system for students to help them choose a subject based on the results obtained when they pass a psychological test. This recommender system is built using the latest machine learning algorithm technology. The rest of the paper is organized as follows: Section II describes related courses and recommender system initiatives using different techniques.

The recommendation system proposed for optional course proposals is describ in detail in Section III. The results and conclusions are in Section IV.

Literature review

This article showed how to include a recurrent neural network (RNN) used to predict behavior while preserving previously observed behavior. In the case of recommender systems, that means a series of sentences without intuitive foresight is too narrow, putting the user in a filter bubble. This article solved this problem. Predict the outcome of the courses offered to students to choose from the one that suit their interests. Create a set of models based on the course catalog (BOW) and another model based on the registration history (course2vec). Performance comparison of two models using validation datasets. Comparing the two models, user findings show a dramatic indifference to RNN policies. However, the machine learning of the course2vec model showed the best results for the task when validated offline. [1].

This article introduces some of the methods used in the course recommendation system. With the help of this system, students can know the approximate results in advance, which helps them choose to study. There is also a table and comparative analysis of the effectiveness of the educational data for each method. Then the best method was selected. And the final version of the recommended system was created. It can already be used in practice. [2].

The course recommendation system is designed to allow students to choose from a variety of subjects. However, the student's choice cannot depend solely on his or her interests. Teachers, peers, and others influence them. Existing rating matrices in the form of interpersonal relationships, rating text, and "user elements" form a multimodal data structure from multiple sources. Therefore, to make recommendations based on these different qualities, need a way to systematically combine the data. Therefore, this article proposes a hybrid recommended model that combines structured network functionality with the use of graph neural networks and interactive student actions with factorization of tensors. First, a structured graphical learning evaluation network is developed using student evaluations to characterize students, courses, commentary texts, grades, and interpersonal relationships. By examining the student's relationship structure, use a neural network with a random walk to create a vectorized representation of the student. Finally, use the Bayesian probability tensor decomposition to examine and predict the scores of students in subjects that are not attending and recognize these personalized traits as the third dimension of the evaluation tensor. Due to the small prediction error and improved recommendation accuracy, the proposed method outperforms other current matrix factorization models and neural networks (RTTF, xSVD++, DSE).[3].

Pakistani universities offer credit programs of all kinds to applicants. Pakistan Virtual University-Developed with the latest information technology. This university offers a choice of a large number of professions and specialties. Each program has subjects that the student must complete. The chosen course that suits the competencies and interests greatly affects the final grade (CGPA) of the student. In this work, a system for recommending subjects in the Virtual University was created. The system has been tested on 470 courses that are accessible, including simulated data from 2,600 people. As a result, it has been proved that grades are influenced by the average score of the student in the courses already studied and the average score in similar courses. The implemented system and its accuracy were evaluated using the mean absolute error for 100 observations. The MAE was in the acceptable range. [4].

This document describes a hybrid RS that uses Content-Driven Filtering (CBF) and Collaborative Filtering (CF) to propose the most relevant courses for students based on several factors linked to both student and course information. The genetic algorithm (GA) was created to find the best RS configuration, which takes into account the most significant criteria and the remaining factors. The pilot project employed real data from the University of Cordoba's computer science program, comprising data collected from students over three academic years and based on 2,500 records from 95 individuals and 63 courses. The examination of the most acceptable criteria for course suggestions is demonstrated by experimental findings. The experimental findings show the relevance of employing a hybrid model that includes both student and course information to increase the dependability of suggestions, as well as better performance compared to earlier models [5].

Consider all the methods that were used in the works that are described above, as well as to conduct a comparative analysis and show the results of the work. The method proposed consists of several stages: (1) after reading all the resources, identify what technique is used (2) what was dataset used (3) what classification method be used (4) what the result works. To identify methods that will help me in further work, I brought out all the algorithms used, as well as the results of the work in Table 1.

Study	Technique	Dataset	Classifier	Obtained Results	
				Metrics	
Zachary A. Pardos and	Diversity based	student course enrollments at	BOW (div)	unexpectedness	3.550

Weijie Jiang 2020[1]	algorithms (BOW, Equivalency) vs Non-diversity algorithms (Equivalency, RNN)	UC Berkeley from Fall 2008 through Fall 2017		serendipity	3.227
				novelty	3.896
			Analogy (div)	diversity	4.286
			Equivalency (non-div)	successfulness	3.619
				commonality	4.500
H. Thanh-Nhan, N. Thai-Nghe, H. Nguyen 2016[2]	KNN, MF, BMF	Grading system at Can Tho University from 1994 to 2004 in ICT	UserKNN	RMSE	0.998
			ItemKNN	RMSE	0.862
			MF	RMSE	0.862
			BMF	RMSE	0.831
Yifan Zhu, Hao Lu, Ping Qiu, Kaize Shi, James Chambua, Zhendong Ni 2020[3]	xSVD++, DSE, TENTF	Beijing Institute of Technology, BIT-UASET evaluation system	xSVD++	MAE	8.4237
				RMSE	11.9673
			DSE	MAE	7.4385
				RMSE	10.7338
			TENTF (Proposed)	MAE	6.9133
				RMSE	9.5913
Aleem Akhtar 2020[4]	Collaborative Filtering, select	Virtual University 2600 students and 470	VU-CRS	MAE	5.12

	neighbours, predict score	courses data during 4 years			
A. Esteban, A. Zafra, C. Romero 2020[5]	Recommendation System, Collaborative Filtering, Content-based Filtering	Academic survey of University Cordoba, CS, 95 students and 2500 ratings of 63 courses from 2016 to 2018	Proposed hybrid RS	RMSE	0.971
				nDCG	0.682
			CBF with clustering	RMSE	1.224
				nDCG	0.234
			User-based & item-based CF	RMSE	1.166
				nDCG	0.549
			MCSeCF	RMSE	1.595
				nDCG	0.112

Table 1. Results of related works

Based on all of the above, say with confidence that the creation of a recommendation system for the university is very relevant at the moment. The articles that were presented to us will be very helpful in our work since they showed what goals to pursue in our work, what data to collect, what algorithms to use in our work. Analysis of Table 1 shows that all work is aimed at creating a system that can prompt the student as accurately as possible on which subject will be of interest to him based on his choice of previous subjects, the average grade for them, as well as the teacher who teaches the subject.

Method and results

Data structure

Data was taken from an open source as a basis[7]. It consists of 12784 rows and 7 columns about the course: its id, name, usage period in seconds, announced time, full description, completion status, and if the course is out-of-date. In Table 3, the top 5 rows displayed are sorted in descending order by the date the course was released.

	id	Title	Duration (sec)	Release Date	Description	Assessment Status	Is retired
0	prometheusconfiguringcollect-metrics	Configuring Prometheus 2 to Collect Metrics	5071	2021-07-14	A Prometheus deployment is only as useful as t...	None	no
1	kubernetes-creating-custom-resources	Creating Custom Resources in Kubernetes 1	6450	2021-07-14	Kubernetes comes with out-of-the-box support f..	None	no
2	kubernetesmonitoring-scaling-applications	Monitoring and Scaling Applications in Kubernetes	4924	2021-07-14	Organizations these days sit on massive piles ...	None	no
3	getting-started-microsoftazure-computer-vision-api	Getting Started with Microsoft Azure Computer Vision API	2796	2021-07-14	Processing images to get returned information ...	None	no
4	test-taking-skills-microsoft-certifications	Test-Taking Skills for Microsoft Certifications	6464	2021-07-01	Microsoft certification exams test not only yo...	None	no

Table 3. Courses data first 5 rows

Data Pre-processing

After receiving the data, one of the most important steps is the preprocessing of the data. Preprocessing data improves the quality of the data and makes it easier to draw meaningful conclusions from the data. The process of cleaning up and deploying raw data suitable for building and training machine learning models is called machine learning data preprocessing. My data went through the following steps:

- Remove all null and none values
- Replace characters with empty space
- Leave only alphanumeric
- Remove stopwords
- Replace words with numbers using vectorization functions

The k-means method is a basic supervised machine learning technique that can be used to troubleshoot classification problems. The k-means method stores information only during the training phase, and when new data is received, it is grouped into clusters that closely resemble the new data. Based on the subject description, we have divided all university subjects into specific clusters. The elements in a particular cluster are contiguous. Next, the student describes his or her tendency. For example, he enjoys painting and inventing figures. Here, the recommended system provides him with elements from the first cluster, such as modeling, tools, or animation. The table below (Table 4) compares the different input parameters and outputs of the model using these parameters. After fitting the model with vector data, 30 different clusters are generated and Table 5 contains the top of the list of cluster results with the most important terms. Figure 1 shows how the sum of squares of the error decreases slightly as the size of clusters increases.

Parameters		
Cluster size	8	30
Iteration count	100	500
Centroid initialization	8	15
SSE	7583.8	7015.9
Accuracy	0.73	0.84

Table 4. Model comparison

Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4

modeling	network	azure	data	test
tools	router	cloud	hadoop	studio
rendering	protocol	storage	science	unit
components	docker	virtual	big	automated
effects	access	aws	course	framework
animation	ccnp	services	analysis	penetration
techniques	comptia	database	visualisation	verification

Table 5. Top of the list of clusters results

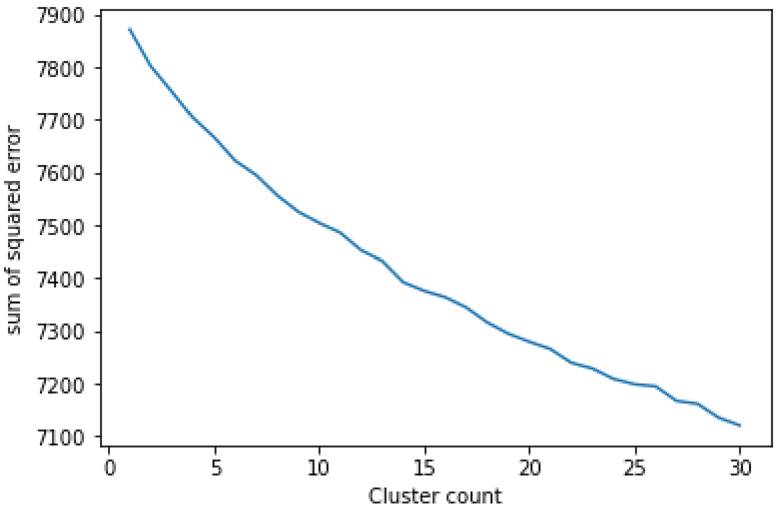

Figure 1. SSE within-cluster size

Conclusion and limitation

Based on all of the above, we are confident that creating a university recommender system is very important at this time. This task will help students find a course that suits their abilities by simply writing down the notes. Our method is to determine a more accurate cluster and output the corresponding rate. We also improved the previous method to get a more detailed clustering model. This model works well when you need to predict a course category with a fairly complete description of the subject and a description of the student's key skills, as the sum of connectivity and root-mean-squared depends on it. For future work, you can retrieve datasets from SDUs and create university recommender systems.

References

1. Z. A. Pardos, W. Jiang, "Designing for Serendipity in a University Course Recommendation System", In Proceedings of the 10th International Conference on Learning Analytics and Knowledge (LAK '20), pp 350-359, 2020
2. H. Thanh-Nhan, N. Thai-Nghe, H. Nguyen, "Methods for building course recommendationsystems", In Eighth International Conference on Knowledge and Systems Engineering (KSE), 2016
3. Y. Zhu, H.Lu, P.Qiu, K.Shi, J.Chambua, Z.Ni, "Heterogeneous teaching evaluation network based offline course recommendation with graph learning and tensor factorization", In Neurocomputing 415, pp 84–95, 2020
4. A.Akhtar, "Implementation of Course Recommender System for Virtual University of Pakistan",In Computers in Human Behavior, 2018
5. A. Esteban, A. Zafra, C. Romero, "Helping university students to choose elective courses by using a hybrid multi-criteria recommendation system with genetic optimization", In Knowledge-Based Systems, 2020
6. C. De Medio, C. Limongelli , F.Sciarrone & M.temperini , "MoodleREC: A recommendationsystem for creating courses using the moodle elearning Platform", In Computers in Human Behavior, 2019
7. Pluralsight [Online] http://api.pluralsight.com/api-v0.9/courses [Last update available 07.03.2022]. E., ”Internet memes as a phenomenon of mass culture”, 2020