Approach for Eliciting Learners' Preferences in Moocs Through Collaborative Filtering

https://doi.org/10.3991/ijet.v17i14.29887

Khadija Naji(✉), Abdelali Ibriz
Transmission and Information Processing Laboratory, Sidi Mohamed Ben Abdallah University, Fes, Morocco
khadija-naji@usmba.ac.ma

Abstract—MOOCs (massive open online courses) quickly become essential components for assuring educational continuity and supporting future life and ways of working. Therefore, there is a need for MOOCs to move away from its old model. This framework will present an algorithm-based recommendation system that will employ collaborative filtering based on MOOC learners' preferences. Collaborative filtering is a technique for anticipating a user's interests by studying the users' preferences who are similar to the individual in question. This approach ensures the analysis of many elements using the participants' rating choices. A recommendation system is becoming increasingly common in online study activities; we want to study how it might help learning and promote a more effective involvement. This study will provide an insight into the existing literature on recommendation systems for online learning and their contribution to supporting learners. We will base our proposed recommendation system on the evaluation of course content. The idea is that learners rate the courses and content they have registered for on the platform between 1 and 5. Following the rating, we extract the data into a comma-separated values (CSV) file and use Python programming to provide recommendations using data from learners with similar rating patterns. The purpose was to utilize Python programming to propose courses to different users in a text editor mode. We will use similar rating patterns via collaborative filtering to recommend courses to various learners, enhancing their learning experience and passion.

Keywords—MOOC massive open online courses, recommendation system, collaborative filtering, Python, learners' preferences

1 Introduction

Morocco's educational system has been the topic of several disputes and controversies in a wide range of subjects since the country's independence in 1956 [1]. Since independence, the educational system has faced several issues that have contributed to a decline in academic quality [2]. The differences of opinion over the measures to take contrast with their unanimity in diagnosing the problems, including uncontrolled growth of students in secondary and higher education. A lack of equipment, inferior quality, high rates of failure and drop-out levels, inadequate teacher training, inappropriate syllabuses, social inequality, problems with female school attendance, and poor coordination with the employment market [3].

The global pandemic drove social distance practice worldwide at the beginning of 2020, altering work environment dynamics and leaving possibilities like online trading, work from home, and online education [4].

Like all other countries affected by the Covid-19 pandemic, Morocco has implemented many measures to reduce the coronavirus's rapid spread as part of a prevention and preparedness strategy [5]. The country has implemented several instructions and initiatives in education to deal with the consequences of this epidemic. The Covid-19 pandemic pushed Morocco to initiate the move from face-to-face to distance schooling [6].

The current situation, combined with technological advancements, calls for the update of educational systems. The Moroccan government has recognized the value of incorporating MOOCs into its educational system. Several initiatives have succeeded in several Moroccan universities before the spread of the coronavirus [7]; other universities have been mobilized to address the delay, and MOOCs have been given top priority. Unfortunately, several factors obstruct the integration of these tools into Moroccan universities: Teachers' interest and involvement in continuous quality improvement directly linked to the use of MOOCs [8], the incentive of students to enroll in a new method of learning, and the educational use of MOOCs [9]. And so on. Following the declaration of a state of emergency, the Moroccan government imposed further measures, including prohibiting public gatherings and suspending schools and universities [10].

The pandemic, however, has become a true advocate for public education, which seems to be seizing the opportunity through the development of online educational content and the need for good collaboration at all levels. Each individual (teachers and students) within a given institution has their part to contribute to the structure to make this transition a success.

Therefore, MOOCs are critical components for ensuring educational continuity and sustaining future life and work styles [11]. This research project examines the transfer from face-to-face learning to virtual learning and its impact on students' learning. In addition, this research is part of a more significant effort to offer a new approach based on collaborative filtering to improve student safety and instill a mindset of using MOOCs as a new mode of learning. This work also aims to respond to students' reluctance to use this educational tool through preference elicitation.

Collaborative filtering algorithms are frequently used in such frameworks for the prediction of choices, preferences and / or ratings of Internet users.

2 Related work

We adopt two parts in this study: on the one hand, we discuss an overview of existing approaches for recommendation systems applied to platforms e-learning and MOOCs environments using the elicitation of learner preferences and interests. On the other hand, we integrate the collaborative filtering methodology into this study to propose a framework developed under the python programming language.

2.1 Content-based filtering

Content-based filtering provides recommendations based on the similarity of item features and user preferences. We can extract meaningful features from the course content and correlate them with the user's preferences [12].

For example, the subject "Mechanics" can be viewed as a Physics subject offered to learners who regularly study Physics. We can extract course information from the title of the topic, its description, and information provided by the editor.

User preferences can be both information that the user provided himself, for example, demographic data, and indirectly obtained information based on the user's browsing and purchase history. Many recommender systems have to deal with the cold start problem when there is not enough information gathered for a new user [13].

2.2 Knowledge-based recommendation system

If the website sells houses or vehicles, the system cannot rely on user ratings since the transactions are small and infrequent, which means there's too little data - no statistical significance. Nevertheless, how can we ensure that the user buys his ideal home from us rather than one of our competitors? [14].

In this case, you can add filters: a house in a city or a village, the number of floors and square meters, the material of the walls. After that, the recommendation system selects the most suitable houses from the catalog. In platforms of e-learning and MOOCs environment, knowledge-based systems are case-based systems that use a similarity function to access the needs of the learners and provide recommendations [15].

This type of recommendation is accurate: the user sees exactly what he wants. However, knowledge-based filtering algorithms are more difficult to think through because there can be many search parameters.

2.3 Hybrid filtering system

Recommender Systems have been developed to provide content and services compatible with users based on their preferences and interests. Hybrid recommender systems combine different approaches [16]. This way, you can get rid of most of the disadvantages of "unmixed" systems. For example, in online clothing stores, recommendations show items that are similar to those you have already viewed, as well as those purchased by users with similar tastes - that is, both content-based filtering and collaborative filtering are enabled at the same time [17].

2.4 Rated recommendation system

This system is a model of collaborative filtering mechanism where learners can assess course content. Once the scoring is complete, there will be a recommendation for learners who have achieved similar results. Assuming a new learner starts the course, the system will recommend the general functionality as they assess and progress through the course. Then the functionality of the assessment system will be in effect. As a contrast to the effectiveness of the grading system, [18], in a working paper on peer review, states that "participants are primarily interested in collective capacity building across the network, and are more likely to use feedback and grading systems honestly. While in xMOOCs, participants aim for a good personal score.

2.5 LSH and minHash

Locality Sensitive Hashing (LSH) is an indexing method. This technique seeks to generate buckets. Each bucket includes identical documents with a strong probability, while non-similar documents tend to be in distinct buckets; we may also define it as a function list, which runs individually. LSH is a technique that, for similar inputs, it would, with high probability, create outputs that fall in the same "bucket," unlike traditional hashing, where we aim to minimize hashing collisions for locality-sensitive hashing, we strive to maximize collisions for similar items. There is a possibility that a learner will be thrown into more than one bucket so that he meets several neighbors [19].

Collaborative filtering recommendation system generally uses Jaccard distance to calculate the similarity between two learners. The minHash has the property that the probability for two elements is the same is equal to the Jaccard distance between arrays. The idea of minHash is to have a hash function that is sensitive to distances (LSH). Indeed, if two points are close to each other, the probability that this function hashes them to the same bucket is high. On the other hand, if the points are far apart, the probability that they are hashed in the same bucket is low [20].

3 Proposed collaborative filtering-based algorithm

We will apply the framework proposed in this work in the MOOCs environment. The idea of this framework is to use some concepts detailed in the literature. A learner enrolled in a course will be able to give a grade from 1 to 5. These scores will thus be stored in a CSV file and subsequently exported to a directory created using a terminal. Then, the Python editor will process this file, and we perform the recommendation process using the collaborative filtering algorithm that we developed in Python and applying the prediction function (Figure 1).

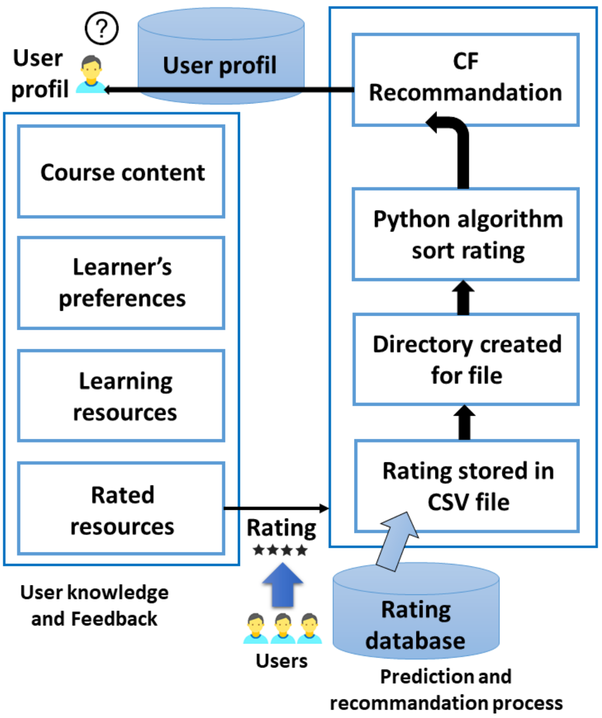

Fig. 1. A proposed CF architecture

3.1 Concepts of CF proposed system

In order to recommend personalized preference course contents to learners using a rating system, we adopted the process involved in the CF architecture. After the rating step, we will store the scores in CSV file format. This file will then be held by Python environment and extracted using the CF algorithm we have developed. We will then create a directory file using a terminal as in the CSV file. The process of course content recommendation is done based on the unique users' stored profiles in the database (Figure 1). As described in [21], the suggested method will extend the prediction function to recommend items to new target users.

3.2 Prediction function

The target user requests for a recommendation from a group of collaborators. Finding content recommendations for the target user from the collaborative users using a similarity matrix. The identification of all the collaborative users with similar rating patterns is significant. The prediction function *M*, calculates the prediction of user s for content *j* . *p* is the set number of collaborative users, *sim(s,i)* is the collaborative user set for similarity with the target user, then multiply by the rating difference of any of the contents or sub-contents *j* within the whole rating process of the collaborative user.

Then sum it up and multiply by the modulation factor q and added with the average rating n_s of our target user which will then produce the prediction rating for our target user recommendation:

$$M_{s,j} = n_s + q * (\sum_{i=1}^{p} sim(s,i) * (n_{i,j} - n_i)) \quad (1)$$

The prediction function *M* process select the target user *t* and all of the similar collaborative users, as well as similar content of interest to the learners, as shown in the far-right equations following the equality sign, n_s plus $q*$ and the brackets. Collaborative filtering and content-based recommendations are two prevalent paradigms for 'context recommendation systems and learner preference prediction[22].

According to [23], they argued that individual predictors rate items in a different manner. Some individuals rate content depending on their personal preferences for how relevant the topic is to them. With this practice, users rate some contents higher than the others, which did not reflect and predict the actual values of the contents previously scored. They resolved these anomalies and drawbacks of these rating behaviors by normalizing the contents matrix before predicting.

3.3 Recommendation's algorithm

We wrote the developed algorithm in the Python programming language. The flowchart reveals the process of course content rating and recommending content based on similarity rating. In the case of a new member, the general course content is recommended because it is new, and therefore no content has been evaluated for the features to take effect (Figure 2). The first component of the algorithm (shown in Figure 3) analyzes the CSV file, associates the user with the contents, and investigates similarities between other users who participated in the rating, as indicated in the prediction function.

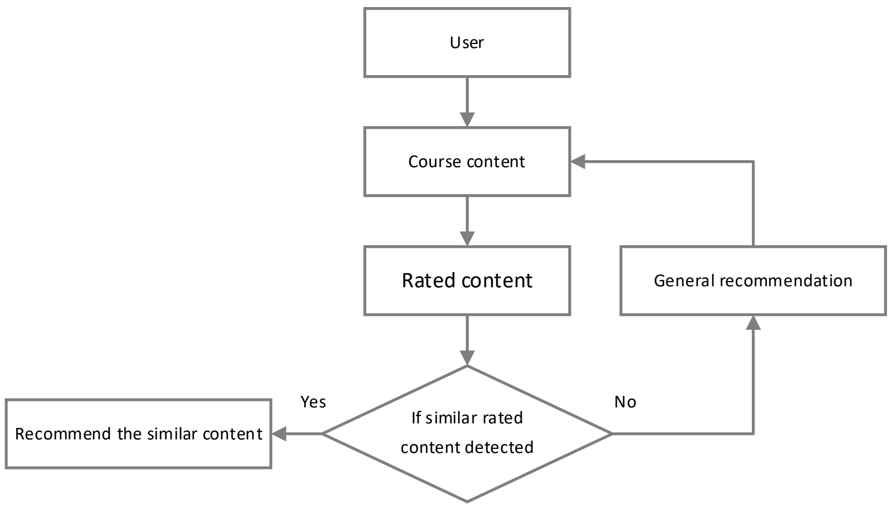

Fig. 2. Flowchart of the learner's rating process and content recommendation

```
  Recommendation's Algorithm
  Input: Data set D,
  Output: K-frequent item sets
  Parameters: data set path, support, output path
  Start:
  // The first phase
  Read data from CSV files
  Affect for each item the length
  Write lengths and the average rating per learner and
per course
  Count items frequency
  Find lowest and highest rated courses
  Show learners who rated course highest and lowest
  // The second phase: Create learner-item matrix using
scipy csr matrix
  N = length of learner Id
  M = length of course Id
  Map Ids to indices
  Map indices to Ids
  X = csr_matrix()
  return X, learner_mapper, course_mapper,
learner_inv_mapper, course_inv_mapper
  // The third phase: Find similar courses using KNN
  Define find_similar_courses function
      find_similar_courses(course_id, X, k, metric='co-
sine', show_distance=False)
      course_ind = course_mapper[course_id]
  course_ vec = X[course_ind]
      k = k + 1
      kNN = NearestNeighbors(n_neighbors=k, algo-
rithm="brute", metric=metric)
      course_vec = course_vec.reshape(1, -1)
      neighbour = kNN.kneighbors(course_vec, return_dis-
tance=show_distance)
      for i in range(0, k):
          n = neighbour.item(i)
          neighbour_ids.append(course_inv_mapper[n])
      endfor
      return neighbour_ids
  endFunction
  similar_ids = find_similar_courses(course_id, X, k=10)
  course_title = course_titles[course_id]
  Write ("Since you enroll in {course_title}")
  for i in similar_ids:
      Write (course_titles[i])
  endFor
  End
```

```
#!/usr/bin/python
import csv
import math
import random
import sys

r=2
nsteps=10
eta=0.0001
f=open('Session.csv', 'rU')
cf=csv.reader (f, delimiter=',')
n=0
A=[]
for line in cf :
m=0
for v in line[1:]:
t=float(v)
if t<6:
A.append((n,m,t))
m+=1
n+=1
if n==1000:
break
f.close()
```

Fig. 3. A screen shot of the algorithm Python part treating the CSV file

The second part of our framework tries to find the similarities between the contents and the learners and to estimate their degree within the rating collaborations as shown in the Figure 4.

```
M1=[]
    for i in xrange(n):
        M1i=[]
        for j in xrange(m):
            s1=mu + b[i] + d[j]
            for k in xrange(r):
                s1+=C[i][k]*P[j][k]
                M1.append(s1)
                M1i.append(i)
    for i,j,s in M:
        M1[i][j]=-7
    maxC=[0]*n
    maxP=[0]*m
    for i in xrange(n):
        for j in xrange(m):
            s1=M1[i][j]
            if s1>M1[i][maxC[j]]
                maxC[i]=j
            if s1>M1[maxP[j]][j]
                maxP[j]=i
    for i in xrange(n):
        sys.stdout.write('Recommend concept %3d to learner %5d (preference:%f)\n'%(maxC[i],i,M1[i][maxC[i]]))
```

Fig. 4. A screen shot of the part of our proposed python algorithm showing the recommendation similarity process

In the final stage of the process, the system recommends to the learner the content best suited to their preferences and interests. These subroutines are just small parts of the algorithm. We are still in the process of developing the whole algorithm, which calculate similarities between learners and their degree using LSH and minHash functions. We have presented only the relevant sections of our framework.

4 Conclusion

The success of online learning environments generally and in MOOCs, in particular, makes progress in the accuracy of personalized recommendation systems potentially beneficial for the satisfaction and loyalty of learners to predict course content for the choice of multiple items.

MOOC recommendations are made based on an intention to promote collaborative learning on MOOC suppliers' platforms. CF recommendation considers on interest by items of similar users, which fits well for collaborative learning in the group. Favorite scores can represent learners' interests. This study proposes a framework CF developed under the Python program to predict learners' best suitable course content. Although the proposed method gives hope for promising forecast accuracy, it would be interesting to consider research that studies the influence of factors such as the incompleteness of the database, the number of available transactions, and the ratio between available transactions and the number of items on relative performance. Because of the present work, we plan to extend our study to develop a recommendation system based on machine learning to encourage learners to enroll in the recommended courses contents using the present approach and follow the lessons with optimal involvement.

5 References

[1] M. Willis, "Politics and power in the Maghreb: Algeria, Tunisia and Morocco from independence to the Arab Spring," *Oxford Univ. Press*, 2012.

[2] F. Viguier, "A French Educational Meritocracy in Independent Morocco?," *French Polit. Cult. Soc.*, vol. 38, no. 2, pp. 148–173, 2020. https://doi.org/10.3167/fpcs.2020.380208

[3] Naji. K., I. A., and M. Y., "Adoption of MOOCs by Emerging Countries Seeking Solutions to University OvercrowdingNo Title," *Int. J. Emerg. Technol. Learn. (iJET).*, vol. 15, no. 22, pp. 78–101, 2020. https://doi.org/10.3991/ijet.v15i22.16945

[4] Y. K. Dwivedi, Hughes, C. D. L., D. C., Constantiou, I., Y., and J. S. Edwards, "Impact of COVID-19 pandemic on information management research and practice: Transforming education, work and life," *Int. J. Inf. Manage.*, vol. 55, pp. 102–211, 2020. https://doi.org/10.1016/j.ijinfomgt.2020.102211

[5] Dehbi *et al.*, "Moroccan digital health response to the Covid-19 crisis," *Front. public Heal.*, vol. 9, 2021. https://doi.org/10.3389/fpubh.2021.690462

[6] Anigri, "E-learning for smart-universities: pandemic challenges and opportunities in Morocco," *E3S Web Conf.*, vol. 297. https://doi.org/10.1051/e3sconf/202129701066

[7] J. A. I., B. K., B. R., S. Machwate, A. Miraoui, and D. Burgos, "Starting MOOCs in African University: The experience of Cadi Ayyad University, process, review, recommendations, and prospects," *IEEE Access*, vol. 8, pp. 17477–17488, 2020. https://doi.org/10.1109/ACCESS.2020.2966762

[8] N. Koukis and Jimoyiannis, "MOOCS for teacher professional development: exploring teachers' perceptions and achievements," *Interact. Technol. Smart Educ.*, 2019. https://doi.org/10.1108/ITSE-10-2018-0081

[9] A. Watted and M. Barak, "Motivating factors of MOOC completers: Comparing between university-affiliated students and general participants," *Internet High. Educ.*, vol. 37, pp. 11–20, 2018. https://doi.org/10.1016/j.iheduc.2017.12.001

[10] Y. Zarhloule, “Framing Nationalism in times of a pandemic: The Case of Morocco. The COVID-19 Pandemic in the Middle East and North Africa,” vol. 55, 2020.
[11] S. Jagannathan, “Future of Work in a Globalized World: Relevance of MOOCs for Continuous Learning,” *Handb. Distance Educ.*, pp. 366–381, 2018. https://doi.org/10.4324/9781315296135-29
[12] O. D. F. O. and S. J., “Collaborative filtering recommendation system: a framework in massive open online courses,” in *9th International Technology, Education and Development Conference*, 2015, pp. 1249–1257.
[13] U. Javed, K. Shaukat, A. Hameed, I., Iqbal, F., Mahboob Alam, T., and Luo, S. A Review of Content-Based and Context-Based Recommendation Systems. International Journal of Emerging Technologies in Learning (iJET), vol. 16(03), pp. 274–306, 2021. https://doi.org/10.3991/ijet.v16i03.18851
[14] P. Rodríguez, N. Duque, and D. Ovalle, “Multi-agent system for Knowledge-based recommendation of Learning Objects,” *Adv. Distrib. Comput. Artif. Intell. J.*, vol. 4, pp. 80–89, 2015. https://doi.org/10.14201/ADCAIJ2015418089
[15] M. Tianxing, N. Zhukova, and N. Mustafin, “A Knowledge-based Recommendation System for Time Series Classification.”
[16] G. Geetha, M. Safa, C. Fancy, and D. Saranya, “A Hybrid Approach using Collaborative filtering and Content based Filtering for Recommender System,” *J. Phys. Conf. Ser.*, vol. 1000, 2018. https://doi.org/10.1088/1742-6596/1000/1/012101
[17] X. Yang, Automatic Recommendation System of College English Teaching Videos Based on Students’ Personalized Demands. International Journal of Emerging Technologies in Learning (iJET), vol. 16(21), pp. 42–57, 2021. https://doi.org/10.3991/ijet.v16i21.26861
[18] R. O’Toole, “Pedagogical strategies and technologies for peer assessment in Massively Online Courses (MOOCs),” *Univ. Open Warwick*, 2013.
[19] X. Z. L. Qi, C. H. W. Dou, C. Yang, and J. Chen, “A two-stage locality-sensitive hashing based approach for privacy-preserving mobile service recommendation in cross-platform edge environment,” *Futur. Gener. Comput. Syst.*, vol. 88, pp. 636–643, 2018. https://doi.org/10.1016/j.future.2018.02.050
[20] Y. Shi, and X.Yang, A Personalized Matching System for Management Teaching Resources Based on Collaborative Filtering Algorithm. International Journal of Emerging Technologies in Learning (iJET), vol. 15(13), pp. 207–220, 2020. https://doi.org/10.3991/ijet.v15i13.15353
[21] D. Blazar and M. A. Kraft, “Teacher and Teaching Effects on Students’ Attitudes and Behaviors,” *Educ. Eval. Policy Anal.*, vol. 39, no. 1, pp. 146–170, 2017. https://doi.org/10.3102/0162373716670260
[22] S. S. Khanal, P. W. C. Prasad, A. Alsadoon, and A. Maag, “A systematic review: machine learning based recommendation systems for e-learning,” *Educ. Inf. Technol.*, vol. 25, pp. 2635–2664, 2020. https://doi.org/10.1007/s10639-019-10063-9
[23] J. Wang, D. A. P. Vries, and M. J. Reinders, “Unifying user-based and item-based collaborative filtering approaches by similarity fusion,” in *the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, 2006, pp. 501–508. https://doi.org/10.1145/1148170.1148257

6 Authors

Khadija Naji is State Engineer in Computer Science, From the National Higher School of Mines of Rabat, Morocco (ENSMR), and curently a Phd student At Transmission and Information Processing Laboratory LTTI, Sidi Mohamed Ben Abdellah University (SMBAU), Fez, Morocco. She is also educational engineer and referent of MOOC projects at the Elearning Center at the University Sidi Mohamed Ben Abdellah.

Abdelali Ibriz is an HDR professor (French Acronym: Habilité à Diriger des Recherches) in the Department of Computer Science at the Higher School of Technology, Sidi Mohamed Ben Abdellah University (SMBAU), Fez, Morocco.Dr.Ibriz's research interests include Computer Systems Design,Cloud Computing, Resource Allocation and Scheduling and Real-time. Currently, he is a member of the Laboratory of Innovative Technologies and Director of the Elearning Center of the Sidi Mohamed Ben Abdellah University.

Article submitted 2022-02-08. Resubmitted 2022-05-08. Final acceptance 2022-05-10. Final version published as submitted by the authors.