A content-based recommender system for choosing universities

MIFTAHUL JANNAT MOKARRAMA

SUMI KHATUN

MOHAMMAD SHAMSUL AREFIN

Recommended Citation

A content-based recommender system for choosing universities

Miftahul Jannat MOKARRAMA[1,*] , **Sumi KHATUN**[2] , **Mohammad Shamsul AREFIN**[3]
[1]Department of Computer Science and Engineering, Faculty of Engineering, Port City International University,
Chittagong, Bangladesh
[2]Department of Computer Science and Engineering, Faculty of Engineering, Bangladesh University of
Business and Technology, Dhaka, Bangladesh
[3]Department of Computer Science and Engineering, Faculty of Engineering, Chittagong University of
Engineering and Technology, Chittagong, Bangladesh

Abstract: Recommender system (RS) is a knowledge discovery and decision-making system that has been extensively used in a myriad of applications to assist people in making distinct choices from vast sources. This paper proposes a recommendation system that will help the prospective students of Bangladesh in choosing the most suitable private universities for getting admission. Since selecting the best private university does not depend merely on a few criteria or choices and making a decision considering all those criteria is not an easy task, a recommendation system can be of great assistance in this scenario for the prospective students. In this proposed recommendation system a list of top-K private universities is recommended to the students who are willing to get admitted to the private universities using content-based filtering technique. To attain this goal we considered six parameters, namely grade point average (GPA) of secondary school certificate (SSC) examination, GPA of higher secondary certificate (HSC) examination, total GPA, tuition fees, university ratings, and university rankings. Finally, we evaluated the system with a total of 947 real feedback from prospective students and obtained the accuracies of 89.05%, 95.85%, 48%, 92.32%, and 71.93% using 5 different performance metrics: precision, recall, specificity, F1 score, and balanced accuracy, respectively.

Key words: Recommender system, content-based filtering, decision-making system, information retrieval, private university admission, data normalization

1. Introduction

Recommender system (RS) is a software tool and technique that renders suggestions to the users on the basis of their requisites, penchants, or tastes. The suggestions pertain to several decision-making procedures, like the genres of the music one would love to listen, types of the items one would like to buy, categories of the online news one would like to read, etc. In recent years to withstand the information overload problems recommendation system has proven to be a valuable means as they beget recommendations based on the history of users or items using different data and knowledge [1–3].

In Bangladesh, Tertiary Education Institutes (TEIs) are the primary organizations after the secondary education system that prepare the manpower with pertinent expertness and knowledge that are prerequisites to fulfill the market and economy demands. Recently, Bangladesh has seen a rapid increase in enrollment in tertiary education [4]. In 2010, the total number of tertiary education students was 1.57 million which was almost

---

*Correspondence: mokarramacse07@gmail.com

doubled in 2017 to 2.92 million[1]. So, it is evident that the share of the youth population with postsecondary education will probably escalate in the near future and it is predicted that this will be from 11 percent in 2010 to 20 percent in 2035[2].

Currently, there are 42 public universities and 103 private universities which are mainly responsible for providing tertiary education in Bangladesh[3]. As of 2018, total 691,958 students passed in higher secondary school certificate (HSC) examination among which number of students who have achieved GPA greater than or equal to 3.5 is 285,301 (GPA 5: 25,562; $5 > \text{GPA} \geq 4$: 116,351; $4 > \text{GPA} \geq 3.5$: 143,388)[4]. But, the seating capacity of the public universities in the first year is only around 50,000[3]. From the information stated above, it is obvious that entrance into universities, particularly into public universities, is very competitive[2]. In this circumstance, according to the statistics of 2017, the rush of the prospective students to the private universities has increased to a significant value over time and constituted about 50% of the total admitted students [5].

Moreover, public universities get the priorities of the majority of students for enrollment in tertiary education after the completion of secondary education. This is due to the fact that public universities are state-run institutions that offer a great variety of subjects, hall of residence at lower expenditure, nationally and internationally recognized platforms for enthusiastic and juvenile knowledge aspirants. In addition, there is a well-known notion that students who get a chance to admit to the public universities are more meritorious and competent as the students having higher GPAs in SSC and HSC examinations can only attend in the entrance examination arranged by a particular department or a public university as whole [6–8]. On the other hand, except some top private universities, most of the private universities have confined spaces, no permanent campus, lack of proper educational facilities and run with part-time teachers. Most importantly, the monthly expense of a student in a medium-ranked private university is about $187 which is around $10 in a public university. So, there lies continuous debate and confusion among general people regarding the regulatory policies and educational excellence of the private universities with respect to their exorbitant tuition fees[5] [4].

In this scenario, if the prospective students fail to qualify in the admission test of any public university, it becomes troublesome for them to choose the best one from a list of over one hundred private universities according to their preferences. So, our task focuses mainly on recommending suitable private universities to the graduate admission seeker to combat this perplexing situation. Some researches [9, 10] showed that factors behind the choice of private universities are: university reputation, location, tuition fees, low session jam, low student's politics, credit transfer facility, demand in the job market, influences of the family members and friends, scholarship facility, and campus atmosphere. It was also found that the primary reasons behind the students' admission into private universities are: good subject choice, reduced session jam, job opportunity, and scholarship opportunity. Since in Bangladesh there is hardly any trusted or updated information center for admission assistance into the private universities, prospective students face problems in collecting or getting the proper information of universities. Even if they want to know about all the universities to compare them depending on their preferences, it becomes a long and strenuous process. For these reasons, in most cases, they

---

[1]BANBEIS (2010,2017). Bangladesh Educational Bureau [online]. Website http://www.banbeis.gov.bd/[accessed 03 September 2019].

[2]SABER (2017). Bangladesh Tertiary Education [online]. Website https://openknowledge.worldbank.org/handle/10986/28466 [accessed 03 September 2019].

[3]BES Chapter 8 (2018). Bangladesh Education Statistics [online]. Website http://data.banbeis.gov.bd/images/ban008.pdf [accessed 03 September 2019].

[4]BES Chapter 10 (2018). Bangladesh Education Statistics [online]. Website http://data.banbeis.gov.bd/images/ban010.pdf [accessed 03 September 2019].

[5]Discrimination between public and private university students [online]. Website https://bit.ly/31mF63u[accessed 02 February 2020].

get admitted into a university just because of its fame, or reduced cost, or location, etc. Sometimes the situation gets worse when students select wrong private universities for their higher education and understand the mistake after one or two years. So, to decrease the workloads of the students by helping them in choosing top-K private universities, here we proposed a recommender system that will collect real data of private universities from relevant and reliable sources and calculate similarities between the university data and prospective students' data for the recommendation.

2. Related work

Recommendation system is an information processing system that helps users search through records of knowledge by gathering various kinds of available data. It generally targets individuals who face problems in assessing the potential items from the overwhelming number of alternatives online [1]. Basically, all recommender systems follow 3 phases: (a) Information collection phase via implicit, explicit or hybrid feedback, (b) Learning phase using different machine learning algorithms, pattern recognition etc., and (c) Prediction or recommendation phase [11].

The system can be perceived in different approaches. It can be content-based, collaborative, or hybrid. A content-based recommendation system mainly focuses on 2 criteria: item and user. Records of the different attributes of an item and a user's profile may reveal valuable information before the system to analyze [12]. Content-based recommender takes into account such information and provides users more attractive items that go to their interests. So, to provide a satisfactory recommendation, firstly, there should be sufficient usable information about the user and item, and secondly, an appropriate learning algorithm needs to be applied to them [13].

On the other hand, the collaborative system tries to ascertain the feasibility of items for an individual considering not only their profile histories but also deducing the similarities with other's profiles [14]. It assumes that people show the same behavior in deciding if they expose similarity in their tastes in former times while making any purchase, giving ratings, or choosing items [15]. Hybrid based is the system that coalesce heterogeneous techniques to attain a new system. For example, a collaborative system can be conjoined with a knowledge-based system so that the cold-start problem of the collaborative system can be compensated [16].

Many researchers worked on different techniques for university recommendation. To ensure the fair and standard enrollment of students' admission in the university colleges, a novel design [17], hybrid recommender system for predicting college admission (HRSPCA) is proposed for college enrollment using hybrid recommender system integrating the techniques of data mining and rules of knowledge discovery. The system consists of 6 components: (a) Students web-portal, b) Enrollment stage, c) Auditing process, d) Track recommender (TR), e) College recommender (CR), and f) College predictor (CP). The first 3 components are used for enrolling the students in the system with proper and validated information following data mining techniques. Next, a total of 23 knowledge discovery rules (categorized into groups A to F) are used in TR and CR. The TR traces the students of the preliminary year and the CR associates them with the specialized colleges if they pass the preliminary year examination with distinction. Finally, for achieving high performance, these 2 hybrid recommenders are cascaded and CP algorithm is applied which uses historical colleges GPA students admission data.

To solve the complex systems that used more than one criterion to apply for more than one course, an admission recommender system is proposed [18] that is developed by using a hybrid method of neural network (NN), decision tree (DT), and system's own algorithm. Here, the applicants are classified into 10 groups using

DT according to some special properties. Then, NN is used to assign the applicants to the available courses that can apply. Finally, to make an ordered list of colleges and universities for each candidate, proposed Fitness equation is applied to the hidden layer in NN that can find the best courses for an applicant. This paper uses GPA, test score, candidates' interest, and their desire jobs as decision parameters.

A university admission recommender system (RSAU) is proposed using a hybrid model of neural network and C4.5 decision tree classifier [19] which has 3 sections, namely data analyzer, classifier, and visualization. In the data analyzer module, data preprocessing and OLAP analysis are done. The classifier module acts as the core design of RSAU. The primary objective of this module is to classify and predict students' admission to universities taking into consideration their secondary school records, profile data, etc. The visualization module acts as a medium between the data mining system and the users by providing a user-friendly interface for amateur users.

A college recommendation system has been developed [20] which comprises of admin module, alumni module, and user module. Here, the alumni module plays the vital role in the whole recommendation generation process. In this module, alumni of different colleges give positive and negative feedback and ratings according to some particulars like teaching, campus, placement, location, fees structure, gender wise college, establishment year, etc. To acquire the positive and negative sentiments of the reviews given by alumni, semantic analysis algorithm is used, and to rank the branches as well as colleges, combined navie Bayes and Ada-Boost algorithm is used. In the user module, a user gets recommendation based on the calculative output of alumni module and the admin module is used to supervise the whole system.

Since Bangladesh still lacks in providing such decision-making assistance to the prospective tertiary students and no noteworthy work has been done to surmount the lacking, here, a private university recommendation system is presented for Bangladeshi students which may help them in deciding which private universities are most likely to fulfill their requirements without having a lot of knowledge about them.

3. Proposed work

In Figure 1, our recommendation system architecture is brought to view. As is shown, the system works in the following sequences: at first, prospective students' data and university data are stored in real time database, then the data are preprocessed by normalization, and finally, these normalized data are sent to the recommender system for generating recommendation. Details of the system are expounded in upcoming sections.

3.1. Prospective student data

GPA is the only information record for SSC and HSC examinations. Every university requires a minimum SSC GPA and HSC GPA for the admission into the university and a prospective student must fulfill those requirements before proceeding forward. To check whether a student satisfies the requirements these basic data are taken as input from the user.

The division is the administrative region of Bangladesh. In Bangladesh, there are 8 divisions each of which is further split into several districts. In this system, students are required to choose the division(s) as location where they want to study for higher education. Keeping in mind the different choices regarding place selection, single or multiple division(s) selection option is provided.

While opting for higher education, a student may have a number of subject choices. Again, tuition fees of different subjects vary not only in different universities but also in the same university. For this reason, a prospective student has to select his or her preferable subject and maximum tuition fees that he or she can
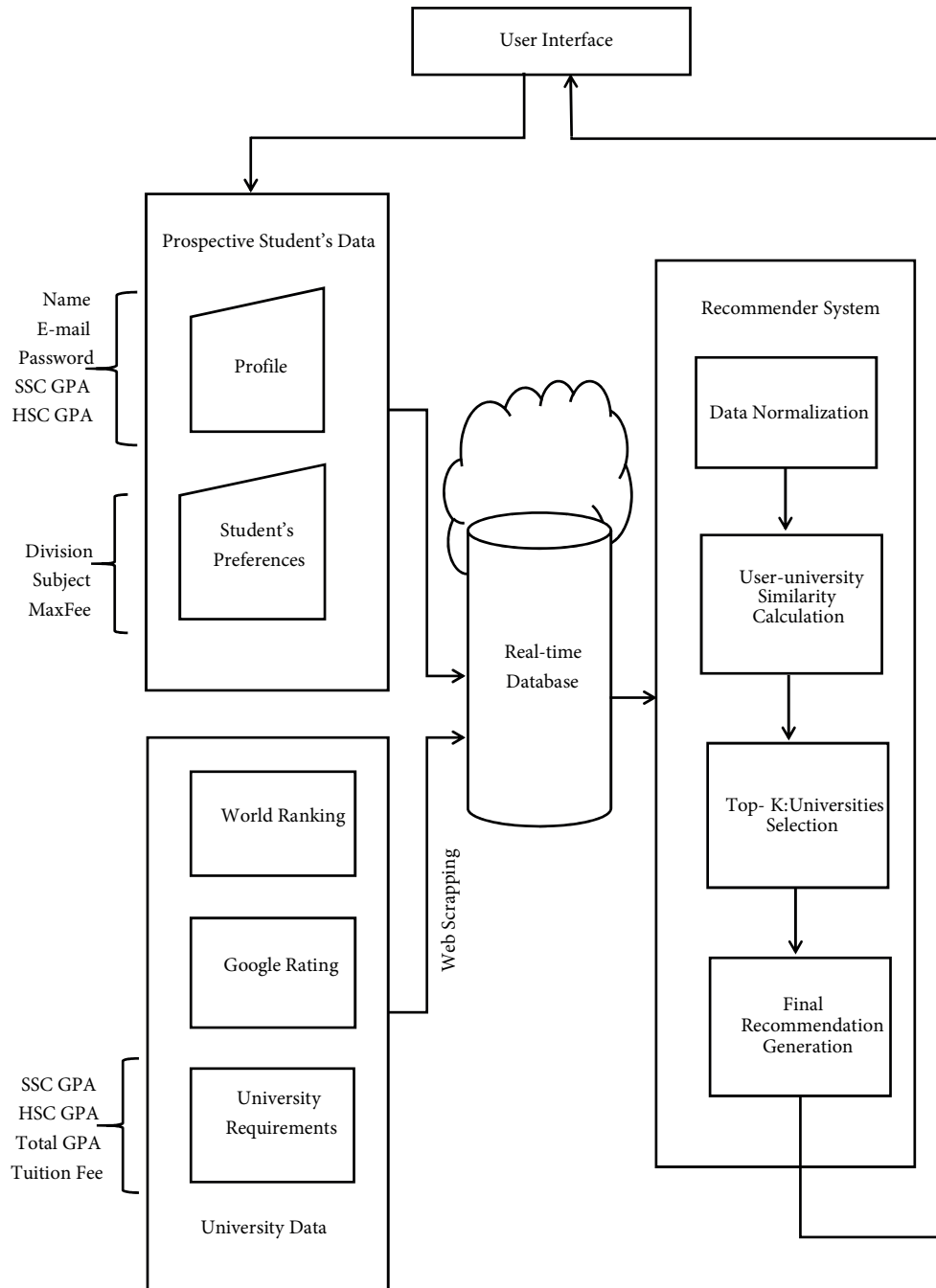
Figure 1. Architecture of the proposed system.

afford for higher study. In the system different ranges are provided for choosing tuition fees by analyzing the tuition fees requirements of private universities of Bangladesh.

3.2. University data

After collecting the prospective student's data, the system filters out private universities which fulfill the requirements. For this purpose, minimum SSC GPA and HSC GPA requirements, total GPA requirements,

subjects offered by the universities, and the tuition fees of each offered subject of each university are scrapped from the official websites of the universities. For providing more reliable and fruitful recommendation this system not only relies on the academic and financial requirements of the universities but also takes the world ranking and Google's rating of the universities into consideration.

Google's rating is an outstanding feature provided by Google that shows ratings of different locations like institutions, restaurants, hotels, etc. by considering top three review sites. Since, nowadays, people tend to weigh up ratings while making a decision; we have evaluated Google's rating of the universities in our system[6].

World rankings of college and university are rankings of institutions in higher education that have been ranked by combining and considering various factors. In most cases, rankings have been conducted by academics, organizations, newspapers, websites, or governments. At present, world rankings have become one of the driving factors in choosing universities for the aspirants who want to do higher studies whether in their homeland or in overseas. For this reason, in our system, we have incorporated the standings of private universities of Bangladesh in the world rankings by considering Webometrics ranking of world universities which is promulgated from Madrid by the Cybermetrics Lab of the Spanish National Research Council (CSIC).

Starting in 2004, the ranking sets a number of benchmarks for over 12,000 universities across the globe aiming to enhance the pedagogical and research association of the institutions on the web and elevating the open access scientific publications. These benchmarks include allegiance to teaching, research outputs, international reputation, and association with the industrial and economic communities. Thus, the ranking not only considers the vastnesses and appearances of the web pages and files of a university but also quantifies the citations received by these contents from different sites[7].

3.3. Real time data storage

The data taken from a prospective student's profile and the data scrapped from the filtered universities are then stored in a real-time database and used for further analysis, processing and up-to-date recommendation generation of the universities. In this regard, we have used cloud firestore (CF) which is a scalable, flexible, and NoSQL database for mobile, server and web development from firebase and Google cloud platform. Moreover, it is a real-time database that has the capability to the real-time data updating using synchronizing and also offers offline support that helps in developing responsive apps.

3.4. University recommender

In this phase, data stored in cloud database are used for generating recommendation in 4 steps: 1. Normalizing the data, 2. Calculating user-university similarities, 3. Selecting top-K universities and 4. Providing final recommendation. Each step is described in the next sections and the recommendation generation process is shown in Algorithm 1.

3.4.1. Normalizing the data

Since data required for calculating similarity are of different categories having varied ranges, we normalized them from 0 to 1 (0 and 1 outlines lower and upper limit of the range respectively). Furthermore, there are 2 types of data in our system. Firstly, data having a fixed maximum and minimum possible values, that is,

[6]Google Ranking of Private Universities of Bangladesh [online]. Website http://bit.ly/2t7pa8E[accessed 02 February 2020].

[7]Webometrics (2019). Ranking Web of Universities [online]. Website http://www.webometrics.info/en/Methodology[accessed 07 October 2019].

Algorithm 1 Recommending top-*K* private universities to the prospective students.

1: **Input:** selected_division, selected_subject, selected_maxfees, user_SSC_GPA, user_HSC_GPA
2: **Goal:** Top-*K* university selection for recommendation
3: Calculate user_total_GPA
4: Normalize user's data
5: **for** selected division **do**
6:    **for** each university **do**
7:       get course_fees, required_SSC_GPA, required_HSC_GPA,
        required_total_GPA, world_ranking and Google_rating
8:       **if** (selected_maxfees$\geq$course_fees && user_SSC_GPA$\geq$required_SSC_GPA
        && user_HSC_GPA$\geq$ required_HSC_GPA && user_total_GPA$\geq$
        required_total_GPA) **then**
9:          Normalize university data
10:          Calculate similarity(user, university)
11:          **do** array(similarity)
12:       **else**
13:          **do** nothing
14: Sort out the university list according to similarity in descending order
15: Select top-*K* universities from university list
16: Calculate system_ratings for top-*K* universities
17: Recommend top-*K* universities to user
18: end

SSC GPA, HSC GPA, total GPA, and rating. In Bangladesh, the maximum possible value of GPA is 5, and for total GPA it is 10 (SSC GPA + HSC GPA). For rating, values range from 0 to 5. In all these cases, we normalized the data by simply dividing it by the maximum possible value. However, normalizing the rating without considering the number of voters will tend to favor universities with smaller number of voters with extremely high ratings. For example, a university with a rating of 4.8 from 10 voters will be considered 'better' than a university with a rating of 4.6 from 10,000 voters. So, we normalized rating taking into consideration the number of votes shown as Equation 1.

$$r_i(normalized) = \frac{r_i}{5} * \frac{v_i}{tv} \tag{1}$$

where, r_i is the rating value of a university u_i in a scale of 0 to 5, v_i is the voter count for u_i, and tv is the total vote count of all the universities in the filtered list.

    Secondly, data is of variable limits (maximum fees, world ranking) which are normalized using decimal scaling data normalization technique. Decimal scaling is a technique where the decimal points of values of an attribute are moved depending on the maximum value among all values of that attribute. For example, if we have a finite set of values $Z = \{z_i\}$ of an attribute A, and if d is the number of digits of the maximum value in the set, then any value z_i of Z can be normalized by Equation 2.

$$Z_i(normalized) = \frac{Z_i}{10^d} \tag{2}$$

$$Z_i(normalized) = 1 - \frac{Z_i}{10^d} \tag{3}$$

Since increasing standing values of universities in world ranking indicate the decreasing ranks, and it is considered generally that students will prefer the universities which have less tuition fees but provide better overall

educational facilities; we have updated the normalization equation of these two attributes or indicators as shown in Equation 3.

3.4.2. Calculating user-university similarities

In a recommendation system similarity matching is one of the main tasks for getting recommendation. The basic concept of evaluating similarity is to calculate the distance between 2 objects where distance refers to the degree of dissimilarity. Therefore, the shorter the distances are, the more similar the objects are to one another and the longer the distances are, the farther the similarities are. This distance is measured by the features or the entities of the objects. Entities having similar categories are more similar, whereas differences in entities decrease the similarities [13]. There are some popular similarity algorithms like Jaccard similarity, Cosine similarity, Pearson similarity, etc. In our system, we used the cosine similarity algorithm to calculate the user-university similarities. Cosine similarity is the similarity between 2 vectors which is calculated by the cosine of the angle of 2 vectors in n-dimensional space. The similarity value ranges from –1 to 1 referring to perfectly dissimilar and perfectly similar respectively. However, in positive space, where the resultant values are neatly bounded in [0, 1], cosine similarity is particularly useful.

Let $n = \mid C_j \mid (j = 1...n)$ be the total number of criteria of our system. For similarity matching we have used 6 criteria (normalized): SSC GPA, HSC GPA, total GPA, ranking of the university, rating of the university and maximum tuition fees. In content-based student to university recommendation system these 6 features will determine the degree of similarity between a prospective student and a university.

Again, consider, $U = u_1, u_2, u_3, u_m; \ m > 0$ be the finite set of filtered universities according to the preferences of a prospective student s. Assume that, the prospective student s and a university $u_i(i = 1...m)$ have weight values w_j and w'_j for criterion C_j. Then 2 vectors S and U_i for s and u_i can be constituted respectively holding the numerical value for each criterion C_j as follows:

$$S = w_1, w_2, w_j, w_n$$
$$U_i = w'_1, w'_2, w'_j, w'_n$$

Then cosine similarity between the student and the university is computed as Equation 4. Thus, the similarity between the student and each university in the filtered list is computed.

$$Similarity_{cosine}(s, u_i) = \frac{\sum_{j=1}^{n} w_j * w'_j}{\sqrt{\sum_{j=1}^{n} (w_j)^2} * \sqrt{\sum_{j=1}^{n} (w'_j)^2}} \tag{4}$$

3.4.3. Selecting top-K universities

In the previous stage, for each prospective student we get similarity value for every university enlisted in the criteria matching list. After that, the values will be sorted in descending order for selecting top-K universities.

Let the similarity values between a student s and universities $u_i \ (u_i \in U)$ be $x_1, x_2, x_3, \ldots, x_m$ respectively, that is,

$$Similarity_{(s,u_1)} = x_1,$$
$$Similarity_{(s,u_2)} = x_2,$$
$$...$$
$$Similarity_{(s,u_m)} = x_m.$$

Suppose that after sorting we get $(x_2 > x_3 > x_m > x_1 > \cdots > x_i)$ for m similarity values. Then the university list will be found as in Table 1 following the sequence of sorted similarity values and top-K universities will be selected from the list for final recommendation.

Table 1. Sorted university list for a prospective student.

Similarity values	Sorted universities
x_2	u_2
x_3	u_3
x_m	u_m
x_1	u_1
...	...
x_i	u_i

3.4.4. Providing final recommendation

While recommending the top-K university list to a prospective student, a system-generated rating value will be associated with each university on the list which is termed as system rating. This rating value gives an abstract idea of the likeliness of each university with the user as it is calculated making use of the similarity value of each university with the user. As similarity values range within 0 to 1 in our system but it is convenient to consider the rating value within 0 to 5, the system similarity value of each university is multiplied by 5 to map this value as system rating of the university. So, the system rating value for each university is as Equation 5.

$$SystemRating_{u_i} = x_i * 5 \tag{5}$$

4. Experiment and result analysis

4.1. Experimental setup

The system is developed using Android Smartphone and personal computer having Windows 10, 8GB RAM. As IDE Android Studio and PyCharm are used where Java and Python were the adopted programming languages. Moreover, a number of dependencies were used during the whole system development process like MultiDex, Python 3.7.3, Pip, BeautifulSoup 4, Lxml, Requests, etc.

4.2. Implementation

The system is developed with real data which are collected from different sources. It is important to work with real data where possible as it can give correct and trustworthy results. Here, at first, prospective student's data are collected through registration process (Figure 2) which are then stored in NoSql database of cloud firestore. Next, different choice categories are to be selected from the category panel shown in Figure 3 according to which data are to be retrieved from the websites of the universities. In this process, the HTML or XML files that are legal or for public use could only be parsed into structured data and stored in the cloud storage. Primarily, we parsed the information of a total of 97 departments of 15 private universities of Dhaka, Chittagong, and Comilla as listed in Table 2.

After that, the system will implicitly get the world ranking and rating of each university in the filtered list from webometrics.com and Google rating respectively. When all these data are collected, normalization

Figure 2. Sign up.

Figure 3. Preference selection.

takes place and similarities between the prospective student and the universities are measured as described in

Table 2. Universities with considered departments.

| Serial number	University name	Total considered departments
1.	North South University	14
2.	East West University	12
3.	United International University	4
4.	Daffodil International University	11
5.	International Islamic university, Chittagong	3
6.	Britannia University	5
7.	Victoria University of Bangladesh	2
8.	Chittagong Independent University	6
9.	Port City International University	11
10.	University of Creative Technology Chittagong	5
11.	Ahsanullah University of Science and Technology	5
12.	International University of Business Agriculture and Technology	4
13.	Bangladesh University of Business and Technology	5
14.	Stamford University Bangladesh	6
15.	Southeast University	3

section 3.4. Finally, top-K universities are displayed along with system generated rating values (for example, system rating of East West University is 4.5) and links to the websites for detailed information (Figure 4).

4.3. Performance evaluation

A feedback option is provided (Figure 4b) for the students who have used the system and got recommendations to gauge the performance of the recommender system. To date, we have got 947 feedbacks from prospective students based on which we have evaluated the system with respect to five performance metrics: precision, recall, F1 score, specificity, and balanced accuracy. Precision, also known as positive predictive value (PPV), is a metric that measures the ratio of correctly given positive rating by the system to all the ratings given by the system. Recall (true positive rate, TPR), on the other hand, measures the ratio of correctly given positive rating by the system to all the positive rating given by the prospective students. F1 score is a measure that combines the results of precision and recall. However, precision, recall and F1 score only consider the positive rating values although the negative ratings are also good if they predict as of students. So, we also measured the true negative rate (TNR) of the rating which is known as specificity. Moreover, we calculated the balanced accuracy of the system which gives more reliable accuracy for imbalanced data sets by considering both TPR and TNR.

Now, let's look at the confusion matrix shown in Table 3 where the rating values greater than or equal to 3 are considered positive and the rating values smaller than 3 are considered negative.

$$Precision, PPV = \frac{TP}{TP + FP} = \frac{740}{740 + 91} = 0.8905 \qquad (6)$$

(a) Recommended Universities.   (b) User Feedback.
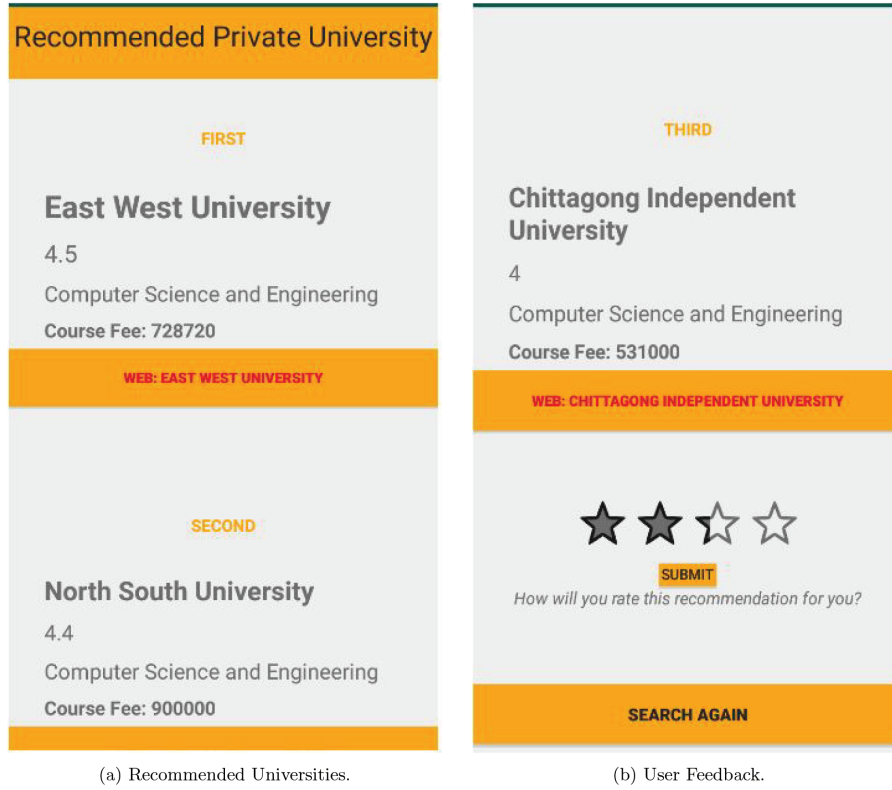
Figure 4. Recommendation and user feedback.

Table 3. Confusion matrix.

| Confusion matrix	Positive ratings given by students (rating ≥ 3)	Negative ratings given by students (rating < 3)
Positive ratings recommended by system (rating ≥ 3)	True positive, TP (740)	False positive, FP (91)
Negative ratings recommended by system (rating < 3)	False negative, FN (32)	True negative, TN (84)

$$Recall, TPR = \frac{TP}{TP + FN} = \frac{740}{740 + 32} = 0.9585 \tag{7}$$

$$F1 Score = 2 * \frac{Precision * Recall}{Precision + Recall} = 2 * \frac{0.8905 * 0.9585}{0.8905 + 0.9585} = 0.9232 \tag{8}$$

$$Specificity, TNR = \frac{TN}{TN + FP} = \frac{84}{84 + 91} = 0.4800 \tag{9}$$

$$Balanced Accuracy = \frac{TPR + TNR}{2} = \frac{0.9585 + 0.4800}{2} = 0.7193 \tag{10}$$

From the above calculations in Equations 6, 7, 8, 9 and 10 we can see that, the system exhibits 92.32% and 71.93% accuracy in F1 score and balanced accuracy respectively. In Figure 5a, variations in precision and recall values with the increase of students' feedback are shown using line graph. Also, Figure 5b shows the relative values of F1 score and balanced accuracy in the bar chart. From these graphical representations, it is evident that, with the increase of the users, the system provides more balanced accuracy.
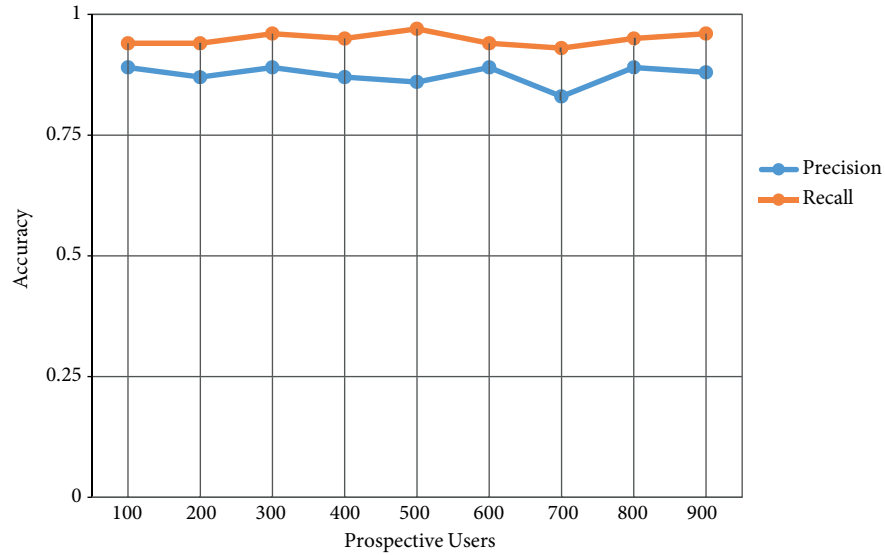
(a) Precision vs. recall.
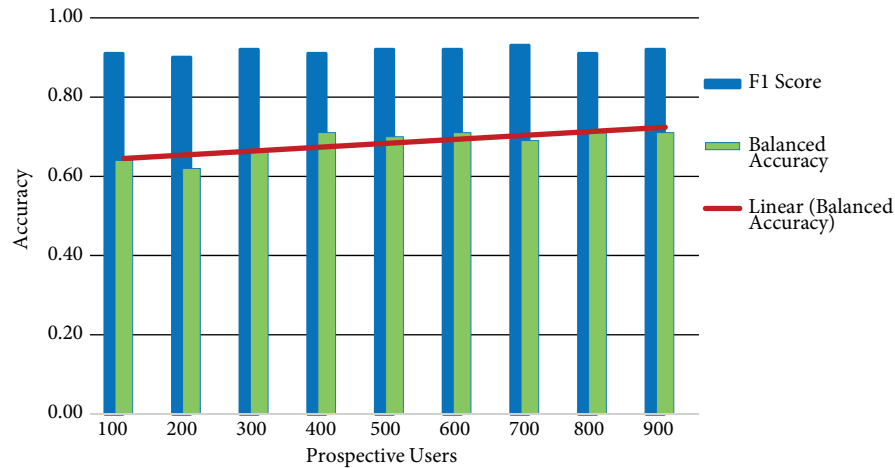
(b) F1 score vs. balanced accuracy.

Figure 5. System performance.

5. Conclusion

In the proposed system, we showed that applying content-based recommendation filtering technique by taking into consideration the university information and user preferences, a prospective student can be provided

recommendation of the most suitable universities. For this purpose, we used cosine similarity algorithm that calculates the similarities between a prospective student and a university using six parameters. Additional parameters like graduates' job statistics, scholarship opportunities for poor or meritorious students, student-teacher ratio etc. can be included for more reliable recommendation generation. Moreover, the system can be extended by integrating user feedback and making use of collaborative filtering techniques to constitute a hybrid system. The more universities are considered, system performance can be evaluated more precisely and it will be more interactive to the user.

References

[1] Ricci F, Rokach L, Bracha S, Kantor PB. Recommender Systems Handbook. Boston, USA: Springer, 2011.

[2] Lai Y, Zeng J. A cross-language personalized recommendation model in digital libraries. Electronic Library of the Institute of Scientific and Technical Information of China 2013; 31 (3): 264-277. doi: 10.1108/EL-08-2011-0126

[3] Mokarrama MJ, Arefin MS. RSF: A recommendation system for farmers. In: IEEE Region 10 Humanitarian Technology Conference (R10-HTC); Dhaka, Bangladesh; 2017. pp. 843-850.

[4] Rahman T, Nakata S, Nagashima Y, Rahman MM, Sharma U et al. Bangladesh Tertiary Education Sector Review: Skills and Innovation for Growth. USA: World Bank, 2019.

[5] Ali M, Hossain ME, Aktar J. The impact of socioeconomic status on selecting post-secondary institutions: The case of private universities in Bangladesh. European Journal of Education Studies 2018; 5 (8): 185-206. doi: 10.5281/zenodo.2483325

[6] Monem M, Baniamin HM. Higher education in Bangladesh: Status, issues and prospects. Pakistan Journal of Social Sciences (PJSS) 2010; 30 (2): 293-305.

[7] Wadood SN. The Public-Private University Debate in the Higher Education Sector in Bangladesh. Munich, Germany: Munich Personal RePEc Archive, Munich University Library, 2008.

[8] Nobi MN. A comparative study of the socioeconomic profile of public and private university students in Bangladesh. The Chittagong University Journal of Social Sciences 2012; 30 (1): 43-56.

[9] Jamil MAA, Sarker MM, Abdullah M. Students' choice criteria to select a private university for their higher education in Bangladesh. European Journal of Business and Management 2012; 4 (17): 177-185.

[10] Saif ANM, Nipa N, Siddique MA. Assessing the factors behind choosing universities for higher education: A case of Bangladesh. Global Journal of Management and Business Research 2016; 16 (11): 45-51.

[11] Isynkaye FO, Folajimi YO, Ojokoh BA. Recommendation systems: Principles, methods and evaluation. Egyptian Informatics Journal 2015; 16 (3): 261-273. doi: 10.1016/j.eij.2015.06.005

[12] Pazzani MJ, Billsus D. Content-based recommendation systems. In: Brusilovsky P, Kobsa A, Nejdl W (editor). The Adaptive Web: Methods and Strategies of Web Personalization. 1st ed. Berlin, Heidelberg: Springer, 2007, pp. 325-341.

[13] Lops P, de Gemmis M, Semeraro G. Content-based recommender systems: State of the art and trends. In: Ricci F, Rokach L, Bracha S, Kantor PB (editor). Recommender Systems Handbook. 1st ed. Boston, USA: Springer, 2011, pp. 73-105.

[14] Adomavicius G, Tuzhilin A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering 2005; 17 (6): 734-749. doi: 10.1109/TKDE.2005.99

[15] Bhatia T. User based collaborative filtering recommendation system using Apache Mahout. MS, Georgia Institute of Technology, Atlanta, GA, USA, 2016.

[16] Burke R. Hybrid web recommender systems. In: Brusilovsky P, Kobsa A, Nejdl W (editor). The Adaptive Web: Methods and Strategies of Web Personalization. 1st ed. Berlin, Heidelberg: Springer, 2007, pp. 377-408.

[17] Ragab AHM, Mashat AFS, Khedra AM. Design and implementation of a hybrid recommender system for predicting college admission. International Journal of Computer Information Systems and Industrial Management Applications 2014; 6: 35-44.

[18] Wakil K, Akram B, Kamal N, Safi A, Web recommender system for private universities' admission in Iraq: UHD case study. International Journal of e-Education, e-Business, e-Management and e-Learning 2014; 4 (5): 329-340. doi: 10.7763/ijeeee.2014.v4.348

[19] Fong S, Biuk-Aghai RP. An automated university admission recommender system for secondary school students. In: The 6th 29 International Conference on Information Technology and Applications; New York, USA; 2009. pp. 1-20.

[20] Monali MD, Dhanashri MG, Dipali MJ, Tejaswini MK, Nale RK. College recommendation system for admission college. International Research Journal of Engineering and Technology 2018; 5 (3): 1269-1272.