

A Model for Automating Course Selection for Teachers and Managers with Natural Language Processing Methods

Dr. Ertürk Erdağ

School Principal, Republic of Turkey Ministry of National Education, (ORCID: 0000-0001-8619-8879)

Abstract: In Turkey, the Ministry of National Education offers courses to teachers and managers. These courses have been practiced for many years. This method is used to train employees in many public and private organizations. In this study, a support application has been created for teachers and managers that can help them to choose from these courses. It has been observed that there is a loss of motivation if the course selection is not made correctly. For this application, which was prepared to prevent this situation, 147 teachers and managers participated in an online form in which various information was taken as multiple choice. The branch and characteristics of teachers and managers were obtained through open-ended questions. The data were compared with 1378 documents in the Ministry of National Education's Standard Education Program using Natural Language Processing (NLP) methods. In the comparison, each course and seminar was evaluated based on word frequency and word proximity and a course-based scoring was obtained. The recommended courses were created in line with the personal and professional characteristics of the teachers. The results were shared with teachers and managers again and they were asked to review the results. A survey was conducted again and 90.47% positive opinions were obtained. With this study, the time spent by the teacher in choosing a training will be minimized and an efficient training process will be ensured.

I. Introduction

The educational process is an ongoing element with many interdependent factors. While it is generally thought that the cluster that learns or is educated is usually the student, since the teacher is a natural part of this system, it is inevitable that he/she is in the education process. It is inevitable that teachers should improve themselves for both professional development and personal development [1]. Education is defined as the process of providing knowledge, skills and behaviors to people who work for a certain wage in private or legal entities' work areas [2]. The teacher's need for course is not only to overcome a deficiency, but also often arises from the desire for professional or personal motivation. The in-branch or out-of-branch contributions of education can provide this motivation in teachers [3]. In Turkey, various studies have been conducted in line with teachers' views on the effectiveness of teacher course programs. There are generally positive opinions in these studies [4], [5]. There are also studies in which it is stated that course is a good process, but structural deficiencies are identified, and that high efficiency can be achieved if the process is organized [6]. The aim of this study is to recommend the most efficient course for teachers and managers. The choice of course will be made with the information to be obtained from teachers and managers due to individual differences. The efficiency of the proposed course will be measured by asking teachers and managers again.

II. Method

For the study, 147 teachers and managers were asked various multiple-choice and open-ended questions in a digital environment. Multiple-choice questions were generally used to reveal their individual characteristics. The answers to the open-ended questions were evaluated with various methods related to natural language processing, a sub-branch of artificial intelligence, and appropriate courses were suggested on an individual basis. The aim of the study was to compare the words and phrases that teachers would use to describe their self-evaluative vocabulary and the words and phrases they would use to describe their fields with the content of the documents prepared for the Ministry of National Education Standard In-Service Training programs and to recommend the courses with the highest correlation.

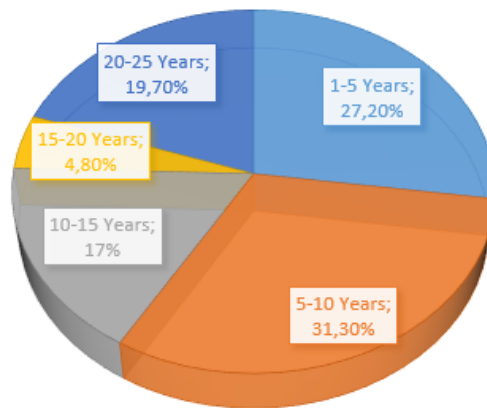


Figure 1: Years in the profession of teachers and managers

When the variable of time spent in teaching for 147 teachers is analyzed in Figure 1, it is observed that this distribution is close to equal. Although it was thought that age would have a negative effect on the results of the research in new education and methodologies such as technology and 21st century skills, the fact that the distribution was close to equal prevented any negative situation in this regard, on the contrary, the equal distribution of factors such as age and teaching experience made the research result better.

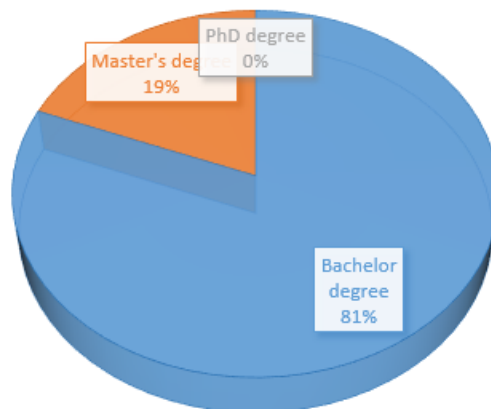


Figure 2: Graduation of teachers and managers

Figure 2 provides information on the type of the last program graduated from. In this graph, it was observed that 81% (119 teachers) were bachelor's degree graduates and 19% (28 teachers) were master's degree graduates. Although it is thought that the studies conducted during and after graduate education can change the needs of the teacher, it was observed that this situation did not have a great impact on the research.

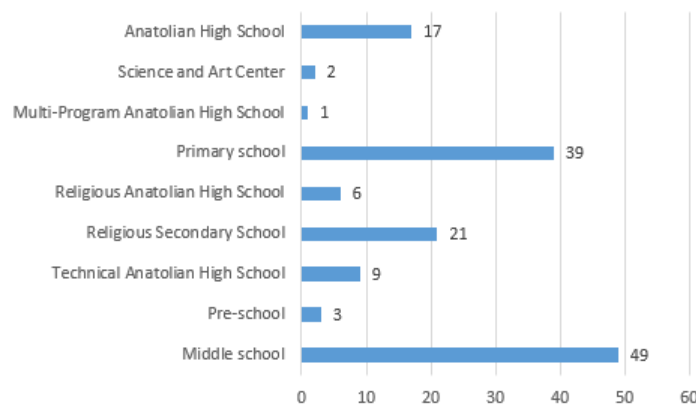


Figure 3: School type of teachers and managers

The types of institutions where teachers work are presented in Figure 3. In this graph, it is observed that there are different proportions of school types. While this situation strengthened the diversity of the scope of the research, it made it difficult to examine different situations under the same conditions. Secondary School (49), Primary School (39), Religious Secondary School (21) and Anatolian High School (17) are the groups with the highest number of teachers per institution.

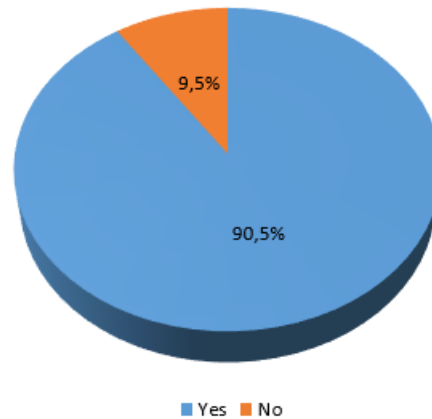


Figure 4: Participation rate in course within the branch

Figure 4 shows the results of the question whether teachers and managers have received previous course in their branches. A large portion of 90.5% of teachers and managers agreed, while 9.5% disagreed. At this point, even if it is a small percentage, it can be thought that the reason for disagreement in this section is that teachers working between 1-5 years do not have enough information about course.

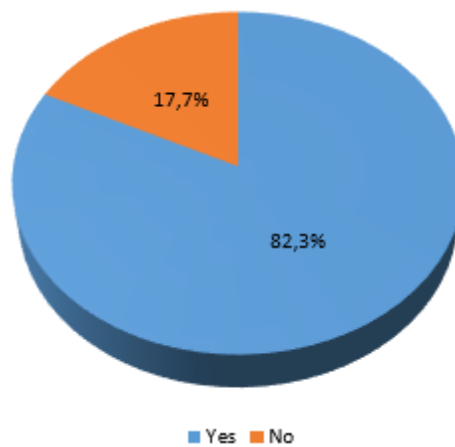


Figure 5: Participation rate in off-branch course

Figure 5 shows the rate of participation of teachers and managers in course outside their field of study. A high rate of 82.3% participated in out-of-field course. This part can be seen as a result of the increase in the number of courses and seminars based on innovative education models outside the field in recent years, and the widespread use of digital competence level both as a hobby and professionally. Today, the fact that robotic coding courses are given at the primary school level has revealed the necessity for teachers to improve themselves in this field.

T.C.
MİLLÎ EĞİTİM BAKANLIĞI
Öğretmen Yetiştirme ve Geliştirme Genel Müdürlüğü
Mesleki Gelişim Programı

ALAN	ALT ALAN	KODU
Bilişim Teknolojileri Alanı	İşletim Sistemleri ve Temel BT	1.01.01.02.003

1-ETKİNLİĞİN ADI

Bilgisayar - Temel Bilgisayar Kullanım Kursu

2-ETKİNLİĞİN AMAÇLARI

Bu kursu başarı ile tamamlayan her kursiyer;

- Bilgisayarı açıp kapar.
- Bilgisayarın temel donanımsal ve çevresel elemanları tanıır.
- Bilgisayarın bir işletim sistemi ile çalıştığını bilir.
- MS Windows'un bir işletim sistemi olduğunu bilir.
- Klasör ve dosya açmayı, kaydetmeyi, kopyalama ve silmeyi yapar.
- Depolama birimlerini bilir ve kullanır.
- Windows işletim sisteminde pencere kavramını bilir.
- Masa üstünü kişiselleştirebilir.
- Word'ün kelime işlemci olduğunu bilir.
- Word'de yeni sayfa açar ve kayıt eder.
- Word'de metinleri düzenler, metne tablo ve resim ekler.
- Power Point'in sunum programı olduğunu bilir.
- Sunuma yeni sayfa açmayı, kopyalama ve taşımayı yapar.
- Sunuma resim, müzik ve video ekler.
- Excel'de boş kitap açar.
- Excel'de çalışma sayfası ekler.
- Excel'de formüllerle dört işlemi yapar.
- İnternet gezginini bilir.
- İnternette yayın yapan sayfaları açar.
- İnternet güvenliğini bilir.
- Elektronik posta gönderir

3-ETKİNLİĞİN SÜRESİ

Faaliyetin süresi 75 ders saatidir.

4-ETKİNLİĞİN HEDEF KİTLESİ

Bakanlığımız okul/kurumlarında görev yapan yönetici, öğretmen ve diğer personel.

5-ETKİNLİĞİN UYGULAMASI İLE İLGİLİ AÇIKLAMALAR

- Bu eğitim faaliyeti personelin temel bilgisayar becerilerini geliştirmek amacıyla hazırlanmıştır.

Figure 6: Introduction page for a sample course content

All documents were converted to text files with txt extension due to the possibility of non-text data in the documents. Python programming language was used in the study. In order to obtain healthy results from the data, preprocessing was performed. Since the preprocessing stage differs according to the data set obtained, the process was used specifically for this data [7]. Within the scope of preprocessing, structures such as tables,

spaces, bullets, etc. were cleaned [8]. For the keyword method, the words in each document were separated on the basis of the relevant course and saved in a local database on SQL Server. In this process, the frequency-based values of all words were determined and recorded with the assumption that high-frequency words are important for this document [9], [10]. Long texts were shortened to a certain extent with extractive based text summarization methods. In particular, the texts were reduced by using two vowel methods [11]. The 40 words with the highest frequency related to the course presented in Figure 6 are presented in Figure 7.

WORD	FREQUENCY	WORD	FREQUENCY
ve	43	test	4
internet	9	program	4
bilir	8	öğretmen	3
temel	7	programı	3
özellikleri	7	alan	3
ayarları	7	bilgisayarı	3
eğitim	6	sistemi	3
ekleme	6	olduğunu	3
bilgisayar	5	yapar	3
dosya	5	windows	3
bu	4	açar	3
ile	4	tablo	3
bir	4	ekler	3
işletim	4	excelde	3
klasör	4	posta	3
kopyalama	4	faaliyetin	3
yeni	4	amacıyla	3
sayfa	4	için	3
resim	4	kullanılacaktır	3
elektronik	4	yöntemleri	3

Figure 7: Part of the word frequency table for sample course content

Stop words (conjunctions, pronouns, etc. that do not have an impressive meaning in the text) are shown in different colors in the table. After recording the content of each course and seminar on a word basis, the words in the open-ended questions submitted by the teachers were calculated using the Jaccard similarity criterion [12], and by adding the weights of the frequency values to this similarity value, a score was created for each teacher for the relevant course or seminar. In this section, the words in which the teachers expressed themselves and their field-based words were evaluated as a total cluster, and no distinction was made. As a result of the calculation, the scores obtained on the basis of teachers and courses were ranked from high to low. It was ensured that the recommendation list was the course or seminar with the highest five points.

A list of recommended courses was sent to each of the teachers and administrators. This was done to validate the study. 133 out of 147 teachers and administrators gave a positive opinion for the proposed courses. This proved that the study achieved a success rate of 90.47%.

III. Results

In the study, a digital survey was conducted with 147 teachers and managers working in various branches and different school types in two different categories: multiple-choice and open-ended. The data obtained were evaluated with Natural Language Processing methods. In the multiple-choice questions, in order to see and interpret the structure of the research group, various questions such as branch, years of profession, in-field and out-of-field course status were asked. Then, as the main focus point, words or word groups that would express the branch they work in on a personal basis were taken with the open-ended question method. Here, a pool of words and word groups was created and the level of closeness of this pool with the words in the document containing the content, functioning and evaluations of each course or seminar in the Ministry of National Education General Directorate of Teacher Training Standard In-Service Training programs was checked, and then the relationship between each course or seminar and the teacher was scored. The course or seminar with the highest value was determined as the most appropriate course or seminar for the teacher. It was seen that it would

be more appropriate to present five options in case the teacher had taken this course before and to create an alternative.

The results obtained eliminate the difficulty of examining the 1378 documents in detail by one person and suggest courses or seminars for the teacher according to his/her interest and ability, both in and out of the field. This study will not only save time for the teacher, but will also pave the way for a more efficient structure, which the Ministry of National Education supports both educationally and economically, and will enable the teacher and the trainer to realize a better educational process.

IV. References

- [1]. S. Altınışık, 'Hizmetiçi eğitim ve Türkiye'deki uygulama', *Kuram ve Uygulamada Eğitim Yönetimi*, vol. 7, no. 7, pp. 329–348, 1996.
- [2]. H. Taymaz, 'Hizmet içi eğitim üstüne', *Eğitim ve Bilim*, vol. 3, no. 16, 1978.
- [3]. A. Unruh and H. E. Turner, 'Supervision for change and innovation', 1970.
- [4]. B. Alan, 'Mesleğe yeni başlayan öğretmenlerin Anadolu Üniversitesindeki bir hizmet içi eğitim kursuna yönelik algılamaları', *Bilkent Üniversitesi, Ankara*, 2003.
- [5]. M. Boydak, 'Hizmet içi eğitim programlarının etkililiğinin değerlendirilmesi (Fırat, Gazi; Marmara Üniversitesi ve TÜBİTAK örneği)', 1999.
- [6]. H. İ. Yalın, 'Hizmet içi eğitim programlarının değerlendirilmesi', *Milli Eğitim Dergisi*, vol. 150, no. 3, pp. 58–68, 2001.
- [7]. F. Başkaya and İ. Aydın, 'Haber metinlerinin farklı metin madenciliği yöntemleriyle sınıflandırılması', in *2017 International Artificial Intelligence and Data Processing Symposium (IDAP)*, Sep. 2017, pp. 1–5. doi: 10.1109/IDAP.2017.8090310.
- [8]. H. Cengiz, T. Uçkan, E. Seyyarer, and A. Karci, 'Metin özetlemesi için düğüm merkezliklerine dayalı denetimsiz bir yaklaşım', *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, vol. 8, no. 3, pp. 1109–1118, 2019.
- [9]. S. Yıldırım and T. Yıldız, 'Türkçe için karşılaştırmalı metin sınıflandırma analizi', *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, vol. 24, no. 5, pp. 879–886, 2018.
- [10]. H. P. Luhn, 'The Automatic Creation of Literature Abstracts', *IBM J. Res. & Dev.*, vol. 2, no. 2, pp. 159–165, Apr. 1958, doi: 10.1147/rd.22.0159.
- [11]. E. Erdağı and V. Tunalı, 'Comparison of feature-based sentence ranking methods for extractive summarization of Turkish news texts'.
- [12]. I. Meltem and A. Y. Çamurcu, 'Web belgeleri kümelemede benzerlik ve uzaklık ölçütleri başarılarının karşılaştırılması', *Marmara Fen Bilimleri Dergisi*, vol. 20, no. 1, pp. 35–49, 2011.