

DETERMINING THE CORE PART OF SOFTWARE DEVELOPMENT CURRICULUM APPLYING ASSOCIATION RULE MINING ON SOFTWARE JOB ADS IN TURKEY

Ilkay Yelmen¹ and Metin Zontul²

¹Department of Computer Engineering,
Istanbul Technical University, Istanbul, Turkey
yelmen@itu.edu.tr

²Department of Software Engineering,
Istanbul Aydin University, Istanbul, Turkey
metinzontul@aydin.edu.tr

ABSTRACT

The software technology is advancing rapidly over the years. In order to adapt to this advancement, the employees on software development should renew themselves consistently. During this rapid change, it is vital to train the proper software developer with respect to the criteria desired by the industry. Therefore, the curriculum of the programs related to software development at the universities should be revised according to software industry requirements. In this study, the core part of Software Development Curriculum is determined by applying association rule mining on Software Job ads in Turkey. The courses in the core part are chosen with respect to IEEE/ACM computer science curriculum. As a future study, it is also important to gather the academic personnel and the software company professionals to determine the compulsory and elective courses so that newly graduated software developers can easily adapt to the software projects in the market without taking extra training.

KEYWORDS

Association Rule Mining, IEEE/ACM Computer Science Curriculum, Software Development Curriculum, Software Job Ads

1. INTRODUCTION

There are many departments that give education in software development in bachelor degree, such as Software Engineering, Computer Engineering, Computer Science or Mathematics Computer. All graduates in the market generally use software engineer, software specialist or software developer titles in Turkey. Actually, software engineering is used instead of software development for the most of time. Therefore, it is very vital to determine the common core part of the curriculums of these departments. Software development is an engineering practice that includes the topics such as design, implementation and maintenance [1]. In the last 30 years, the importance of software development has increased and it has been growing continuously [2].

This increasing importance and rapid change forced the software education to be adaptive to the market needs. Increasing costs in the software industry, applied wrong strategies, the desires to rise in quality and performance issues and so fast technology changes had revealed the need to educate experts in the field and qualified software developers. Therefore, the university-industry collaboration has gained utmost importance.

Zhengyu stated that a lot of strong software professionals were urgently required in the community but employers felt that the graduates had the software talent shortage while a considerable number of graduates could not find a suitable position [3].

In another study, Kuang and Han proposed the methods of teaching reform as guided by market demand, to update the teaching content, to optimize the teaching methods, to reform the teaching practice, to strengthen the teacher-student exchange and to promote teachers and students together because software development training could not meet the needs of the community [4].

Among the software development departments, software Engineering (SE) is the fastest-evolving engineering discipline that has ability to provide tools and methods for all areas of society [5]. This situation increases the responsibility of SE education to prepare SE professionals for the industry by providing them with skills to meet the expectations of the software industry. Innovations and improvements in the curriculum are required to bridge academia-industry gap [6] since SE education has inability to provide students with large-scale software development experiences [7]. However, only universities can produce highly skilled professionals who can satisfy the needs of software industry by taking into account different standards, frameworks and recommendations developed by interest groups [8].

A study presented software engineering education evolvement in Turkey to provide an assessment of SE curriculum in Turkish Universities with respect to IEEE/ACM guidelines given in SEEK (2004) and to provide a guideline to universities conducting an SE programme at undergraduate level to align their course curriculum with IEEE/ACM guidelines [9].

Students should have necessary background of programming experience for the study of software engineering concepts in their curriculums. In order to satisfy this condition, the current software engineering guidelines include concepts and programming paradigms that must be mastered through study and practice. The well-known guideline for software engineering curricula is recommended by IEEE/ACM. This guideline gives the standards related to course scheduling, faculty preparation, student loads, hardware and software resources, instructional materials and curriculum development. ACM published “Curriculum Guidelines for Undergraduate Degree Programs in Software Engineering (SE2004)” to provide guidance to academic institutions and accreditation agencies about what should constitute an undergraduate software engineering education [10].

This study is related to curriculum development and human resources in software development. In addition, association rule mining on software job ads is applied. There are similar studies in literature as follows. Chien and Chen developed a data mining framework for personnel selection to explore the association rules between personnel characteristics and work behaviors, including work performance and retention. Moreover, they used decision tree analysis to discover latent knowledge and extract the rules to assist in personnel selection decisions [11].

Mohsin, Ahmad, Din, Mahamud and Din proposed an intelligent model that is aimed at facilitating key workers select suitable trainees for a training program. In this study, trainees dataset was mined using association rule to discover important personality characteristics. Their model produced an efficient selection process and suitable trainees [12].

In another study, Ali and Rajamani presented the solution for selecting appropriate talented personnel resumes without risk factors using association rule mining. The practical experimental results obtained from the proposed model encouraged human resource department to take prompt decisions for recruiting talented personnel accurately without wasting interviewers' time of employer and employee. Also, they indicated that the proposed system reduced frequent resignations, improved performance of talented personnel without training cost and continuous monitoring [13].

Finally, Smith and Ali indicated that today's rapid changing and competitive environment required educators to stay alongside of the job market in order to prepare their students for the jobs being demanded. They also implied that data mining methods were suitable for this kind of analysis due to the large volume of job data generated through the web instead of the classical data analysis methods. Their study illustrated the experience with employing mining techniques to understand the trend in IT Technology jobs. At the end, collected data from an online agency was analysed to reach a conclusion about the trends in the job market [14].

In this study, the core part of Software Development Curriculum is determined by applying association rule mining on Software Job ads in Turkey. As a result, software engineering or related fields that give education in software development should include these core courses in their curriculum in order to adapt the software development industry in Turkey.

The rest of this paper is organized as follows: The second chapter makes mention of association rules and the third chapter depicts how to apply association rule mining on software ads. The fourth chapter gives results and discussion with respect to the association rule mining. Finally, the fifth chapter gives the conclusion.

2. ASSOCIATION RULES

One of the important tasks for Knowledge Discovery in data is Association Rule Mining which is a well-known procedure in data mining. In its basic structure, every association rule fulfilling the minimum support and confidence are extracted [15]. The general purpose of an association rule $A \Rightarrow B$ is to denote that records possessing attribute A also tend to possess attribute B. The aim is to find association rules which are considered sufficiently interesting as defined by one or more measures. Most common formulas for support and confidence are as follows [16]:

$$\text{Support}(A \Rightarrow B) = \frac{|A \wedge B|}{|D|} \quad (1)$$

$$\text{Confidence}(A \Rightarrow B) = \frac{|A \wedge B|}{|A|} \quad (2)$$

where $|D|$ indicates total number of records and $|A|$ refers to total number of record including A.

2.1. Apriori Algorithm

Apriori is an algorithm which is developed for common set learning mining on transactional database and association rule learning [17]. Apriori uses a level-wise search, where k -itemsets are used to explore $(k+1)$ itemsets. First, the set of frequent 1-itemsets denoted by L_1 is found by scanning the dataset to find the count for each item, and collecting those items satisfying minimum support. Then, L_1 is used to find L_2 , the set of frequent 2-itemsets, which is used to find L_3 , and so on, until no more frequent k -item sets can be found. The finding of each L_k requires one full scan of the dataset. The Apriori property is used to improve the efficiency of the level-wise generation of frequent item sets by reducing the search space [18].

The Apriori property is based on the following interpretations. By definition, if an item set I does not satisfy the minimum support threshold, \min_sup , then I is not frequent ($P(I) < \min_sup$). If an item A is added to the item set I , then the resulting item set cannot happen more frequently than I . Therefore, $I \cup A$ is not frequent either ($P(I \cup A) < \min_sup$) [18].

3. ASSOCIATION RULE MINING ON SOFTWARE JOB ADS

3.1 Collecting Data

The data used in this study is taken from one of the popular job recruitment site in Turkey and currently available ads about software are examined from 5 big cities that are Istanbul (European Side), Istanbul (Asian Side), Ankara, Izmir and Bursa. At the end feature list and sub-categories are prepared with respect to job ads for using in the data set [19].

653 software job ads in 5 big cities are collected from this job recruitment site in Excel format. Finally, 30 main features are determined for finding suitable candidates for the position as shown in Tab. A.I in Appendix.

3.2 Data Processing

After creating the initial dataset, it is observed that the year of experience in software features can be important. Consequently, 30 features and some information within these features have been separated. For example, considering the experience in programming languages on the job recruitment site, programming languages are divided into 8 pieces as C, C++, C#, Java, Php, Objective C and the other programming languages. On the other hand, experience levels are divided into 7 as 0 (Not Acknowledged), 1, 2, 3, 4, 5 and 5+ (Years). Programming language and experience level features are combined for simplicity. For example, if C# feature contains 2, at least 2 year-experience is required. In addition, as example qualifications in the special programming techniques are grouped and new features are created under new names as Software Architecture Methodologies, Software Patterns, Programming Paradigms and Other Software Development Processes as shown in Tab. A.II in Appendix. Finally, 653 ads are entered for the features in Table A.II.

Then, the data is visualized in terms of database systems, programming languages, Front-end technologies and other software technologies. As shown in Fig. 1, the leading database systems in

job ads are MS SQL, Oracle and MySQL. The mostly used programming languages are C#, Java and C++ in job ads as depicted in Fig. 2

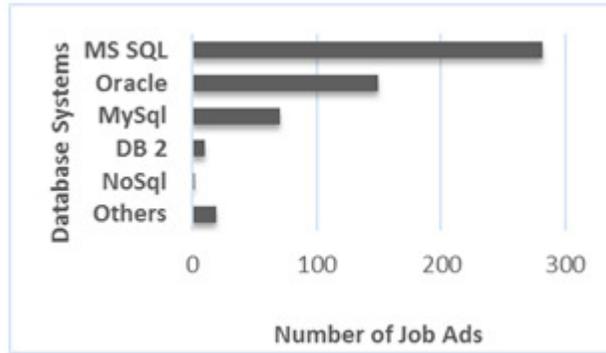


Figure 1. The leading database systems in software job ads

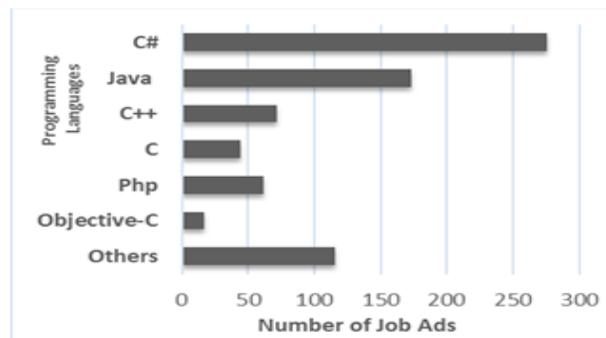


Figure 2. The leading programming languages in job ads

In similar way, Fig. 3 shows the leading front-end development technologies in job ads where Javascript is prominent technology for front-end development. The other software technologies such as software architecture, software paradigms and web services are essential as shown in Fig. 4.



Figure 3. The leading Front-end technologies in job ads

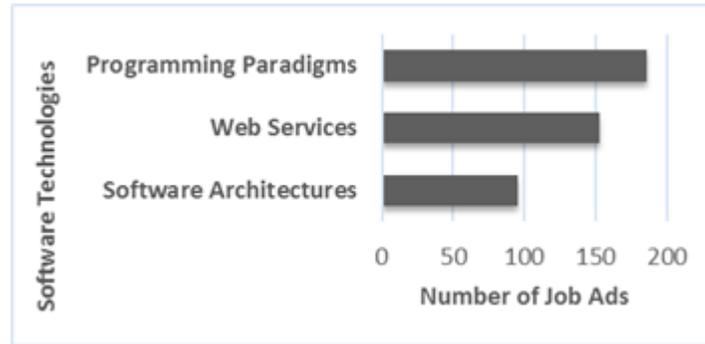


Figure 4. The leading software technologies in job ads

3.3 Applying Apriori Algorithm

Fig. 1, Fig. 2, Fig. 3 and Fig. 4 above give the frequencies of related technologies but they don't give which technologies are related to each other. Moreover, one job ad can contain more than one software technology at the same time. For this purpose, association rule mining by using apriori algorithm is applied in assessment. Different combinations trying out through the features analysis has been done. Minimum support value has been taken as 0.1 and confidence value has been taken as 0.5 in the analysis process. At the end, 54 rules are determined regarding software development.

4. RESULTS AND DISCUSSION

The rules obtained by apriori algorithm are divided into 3 parts as follows:

1. Programming languages, frameworks and databases
2. Front-end web technologies
3. Web services, software architectures and programming paradigms

There are 23 rules related to programming languages, frameworks and databases as shown Tab. 1. If the result sides of these rules are noted, Education_Level=3 (B.Sc. Degree), Position=1 (Software Specialist) are prominent results. On the left side of rule 1, {Java=0} condition means that Java is necessary but the year of experience is not important. The same thing can be said for C# if the rule 8 is considered. While 87% of ads that want Java experience require at least B.Sc. degree (Rule 1), 74% of ads that want C# experience require at least B.Sc. degree (Rule 14). It can be argued from these rules that the university degree is more important in Java than C#. Looking at rules 6 and 12, the similar result can be obtained for Oracle (79%) and MS SQL Server (75%). The rules 4 (80%), 5 (80%) show that the ads seeking for C# experience with MS.NET or ASP.NET framework experience categorize the job seekers as Software Specialists. In rule 2, the same result can be obtained for 82% of ads seeking for C# experience with MS SQL Server experience. From the rule 4, 5 and 7, it can be concluded that Job seekers knowing C# should have enough experience in MS.NET, ASP.NET framework and MS SQL Server. If the rule 8 is compared with the rule 21, it can be said that while 77% of the ads seeking for C#

categorize the job seekers as Software Specialist, the corresponding ratio for Java is 67%. From these rules, it can be decided that C# experience more valuable for the software firms in Turkey. The similar result is valid for MS SQL Server (75% in rule 13) and Oracle (71% in rule 18). Another interesting result can be obtained from the rules 22 and 23 that Oracle ads and Java ads are separated from each other. However, C# and MS SQL Server are combined in rules 2 and 15. It means that Oracle and Java are complicated technologies that cannot be known by one specialist while C# and MS SQL Server are moderate technologies that can be known by one specialist.

Table 1. Rule Extraction for Programming Languages Frameworks and Database through the Apriori Algorithm

No	Rule	Confidence Value
1	If {Java=0} ==> Education_Level =3	0.87
2	If {C#=0 \cap MSSQL=0} ==> Position =1	0.82
3	If {Other_Languages =0} ==> Position=1	0.81
4	If {C#=0 \cap AspDotNetFramework=0} ==> Position =1	0.80
5	If {C#=0 \cap DotNetFramework=0} ==> Position =1	0.80
6	If {Oracle=0} ==> Education_Level =3	0.79
7	If {AspDotNetFramework =0} ==> C#=0	0.78
8	If {C#=0} ==> Position=1	0.77
9	If {DotNetFramework=0} ==> Position =1	0.76
10	If {DotNetFramework=0} ==> Education_Level =3	0.75
11	If {AspDotNetFramework=0} ==> Position =1	0.75
12	If {MSSQL=0} ==> Education_Level =3	0.75
13	If {MSSQL=0} ==> Position =1	0.75
14	If {C#=0} ==> Education_Level=3	0.74
15	If {C#=0 \cap MSSQL=0} ==> Education_Level =3	0.73
16	If {C#=0 \cap DotNetFramework=0} ==> Education_Level =3	0.73
17	If {Other_Languages=0} ==> Education_Level =3	0.72
18	If {Oracle=0} ==> Position =1	0.71
19	If {DotNetFramework =0} ==> C#=0	0.68
20	If {AspDotNetFramework=0} ==> Education_Level =3	0.67
21	If {Java=0} ==> Position =1	0.67
22	If {Java=0} ==> Education_Level =3 \cap Position =1	0.60
23	If {Oracle=0} ==> Education_Level =3 \cap Position =1	0.56

20 rules are available about front-end web technologies as shown in Tab. 2. The rule 24 indicates that 83% of ads that want ajax knowledge require javascript as well (year of experience is not important). Moreover, 81% of ads that include ajax and software specialist together want javascript, too. It can be understood from the rules 24 and 25 that ajax and javascript should be taught together. Similar results can be obtained for the rules 26, 27, 28, 30 in a way that there is a strong relationship among ajax, jquery, javascript, HTML and CSS in terms of software development education in front-end web technology. Also, the rules 32, 34, 36, 37 and 39 imply that at least B.Sc. degree is required at most of the times for front-end development. The rules 24 and 43 show that while 83% of ads that want ajax knowledge require javascript, only 53% of ads that want javascript require ajax. From these rules, it can be extracted that javascript education is more fundamental than ajax for front-end education. Another important item for front-end development is XML as indicated in the rules 32, 33, 40 and 41. The rules 32 and 33 indicate that

the candidates knowing XML should have at least B.Sc. degree and be software specialist at an important level over 70%. The rules 40 and 41 imply the relation between XML and other web technologies HTML, CSS and web services. As 62% of ads requiring XML also want HTML and CSS (rule 40), 56% ads requiring XML want web services. These rules mean that XML is more common data format for data transfer in web platforms. Actually, a few ads include JSON data format but they are eliminated by apriori algorithm because of their low support count. It means that XML is more common in the market in Turkey.

Table 2. Rule Extraction for Front-End Web Technologies through the Apriori Algorithm

No	Rule	Confidence Value
24	If {Ajax=0} ==> Javascript=0	0.83
25	If {Ajax=0 \cap Position =1} ==> Javascript=0	0.81
26	If {Javascript=0} ==> HTML_CSS=0	0.81
27	If {Ajax=0} ==> HTML_CSS=0	0.78
28	If {jQuery=0} ==> HTML_CSS=0	0.78
29	If {Ajax=0} ==> Position =1	0.78
30	If {jQuery=0} ==> Javascript=0	0.76
31	If {Ajax=0 \cap Javascript=0} ==> Position =1	0.76
32	If {XML=0} ==> Education_Level =3	0.75
33	If {XML=0} ==> Position =1	0.71
34	If {HTML_CSS=0} ==> Education_Level =3	0.70
35	If {Javascript=0} ==> Position =1	0.68
36	If {Ajax=0} ==> Education_Level =3	0.68
37	If {Javascript=0} ==> Education_Level =3	0.67
38	If {Education_Level =3 \cap Javascript=0} ==> Position =1	0.67
39	If {HTML_CSS=0 Javascript=0} ==> Education_Level =3	0.66
40	If {XML=0} ==> HTML_CSS=0	0.62
41	If {XML=0} ==> Web_Services=0	0.56
42	If {HTML_CSS=0 \cap Javascript=0} ==> Ajax=0	0.55
43	If {Javascript=0} ==> Ajax=0	0.53

11 rules are obtained related to web services, software architectures and programming paradigms as depicted in Tab. 3. These rules mostly focus on B.Sc. degree and software specialist position. In fact, there are other ads focusing on other positions such as database administrator or software test specialist but they are eliminated because of their low support counts. This means that the most of software firms give ads focusing on software specialists as shown in rules 49, 51, 52. From the other perspective, software architectures, web services and programming paradigms (object oriented programming, functional programming etc.) topics are very special software technologies that should be carried out by software engineers or equivalents having at least B.Sc. degree as seen in rules 44, 45, 46, 47 and 48. The rules 53 and 54 imply B.Sc. degree and software specialist position at the same time. 53% ads including web services imply B.Sc. degree and software specialist position together (rule 53). The same thing is valid for programming paradigms (rule 54). The Programming paradigm stands for the styles of various programming languages such as Python, Lisp, F# and Objective-C.

Table 3. Rule Extraction for Web Services, Software Architectures and Programming Paradigms through the Apriori Algorithm

No	Rule	Confidence Value
44	If {Software_Architectures=0} ==> Education_Level =3	0.84
45	If {Web_Services=0} ==> Education_Level =3	0.83
46	If {Position=1 \cap Web_Services=0} ==> Education_Level =3	0.80
47	If {Programming_paradigms=0} ==> Education_Level =3	0.78
48	If {Position=1 \cap Programming_paradigms=0} ==> Education_Level =3	0.77
49	If {Software_Architectures=0} ==> Position =1	0.68
50	If {Programming_paradigms=0} ==> Position =1	0.68
51	If {Education_Level =3 \cap Programming_paradigms =0} ==> Position =1	0.67
52	If {Education_Level =3 \cap Web_Services=0} ==> Position =1	0.64
53	If {Web_Services=0} ==> Education_Level =3 \cap Position =1	0.53
54	If {Programming_paradigms =0} ==> Education_Level =3 \cap Position =1	0.53

With respect to the rules above, the following courses should be included in Software Development Curriculum as compulsory core courses as shown in Tab. 4. These courses are compatible with IEEE/ACM computer science curriculum where it is dictated that successfully deploying an updated computer science curriculum at any individual institution requires sensitivity to local needs [20]. The rules above correspond to the local needs in Turkey.

The rules 2, 10, 12 and 15 in Tab. 1 imply that MS SQL Server is a fundamental database system for software developers. Thus, Database Systems-I course focusing on relational database concept by MS SQL Server is included in the curriculum. In similar way, since Oracle is also very popular in the market according to the rules 6, 18, 23 in Tab. 1 Database Systems-II course focusing on Oracle should be involved. The most popular language in Turkey is C# on MS.Net framework according to the rules 4, 5, 7, 8, 9, 10, 15, 16 and 19. Because the easiest way to start to learn C# is Windows Desktop environment, Desktop Programming course should be given in the curriculum. The rules 1, 21, 22, 47, 48, 50, 51, and 54 in Tab. 1 indicate that Java language is second popular language in Turkey. Since Java includes all object-oriented principles Object-Oriented Programming course applying the basic principles on Java is recommended. Web development can be divided into two parts as front-end and back-end web development. While the rules 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42 and 43 in Tab. 2 imply Front-End Web Development course, the rules 4, 7, 11, 16, 19 in Tab. 2 involve Back-End Web Development. Front-end development should include HTML, CSS, Javascript, JQuery, Ajax, XML. Relating to the rules 4, 7, 11, 16, 19 in Tab. 1, the most popular back-end web programming is ASP.NET with C# in Turkey. As for the rules 47, 48, 50, 51, 54, the programming paradigm concept refers to various programming languages having different programming styles. As a result, Programming Language Concept lecture should be included in order to teach different programming languages such as Python, Lisp, F#, Objective-C. Finally, Software Architecture course is recommended as compulsory course since the rules 44 and 49 in Tab. 3 show the importance of software architectures focusing on multi-tier architectures and web services.

Table 4. The Core Part of Software Development Curriculum for Undergraduate Degree Programs

Year/ Semester	Related Tech.	Course Name	Related Rules
2/1	MS SQL Server	Database Systems-I	7, 12, 13, 15
2/1	C#, MS.Net Framework	Desktop Programming	5, 7, 8, 9, 10, 14, 15, 19
2/1	Java	Object Oriented Programming	1, 21, 22, 47, 48, 50, 51, 54
2/2	Oracle	Database Systems-II	6, 18, 23
2/2	HTML, CSS, JavaScript, JQuery, Ajax, XML	Front-End Web Development	24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43
3/1	ASP.NET	Back-End Web Development	4, 7, 11, 16, 19
3/1	Phyton, Lisp, F#, Objective-C	Programming Language Concepts	47, 48, 50, 51, 54
3/2	Web Services, Multi-Tier	Software Architecture	44, 49

5. CONCLUSION

Due to consistent growth in software market and rapid change in software technology, the adaption of software development curriculum is necessary with respect to criteria desired by software development industry.

This study has a contribution to the literature in a way that it applies association rule mining on software job ads to help the managers decide on the software development curriculum. In this study, the software job ads regarding the first 5 cities having intensive ads are obtained from a famous Turkish employment web site. While determining the features required for association rule mining, the years of experience on software expertise areas is considered. After applying association rule mining, the rules related to desires of software companies are achieved.

Considering the criteria owned by software job ads, it is investigated that practical part of software development education should be increased. Moreover, since the most of job ads seek the employee at minimum B.Sc. level, the importance of software engineering or related fields is increasing as well. As the companies request the new software technologies as well as fundamental programming abilities, it is vital to revise the software development curriculum at the universities. However, the most of them are very late to adapt their curriculums to the criteria of the companies. Actually, it not enough to revise the curriculums but also academic personnel in these departments should renew their knowledge on the new software technologies.

Since there are many departments related to software development such as software engineering, computer engineering etc., it is necessary to determine the core courses of software development. For this purpose, the core courses are determined by using association rule mining on software job ads and IEEE/ACM computer science curriculum. The core courses should be common for all departments related to software development because they reflect the local need of the software development companies in Turkey.

As a future study, it is also important to gather the academic personnel and the software company professionals to focus on the hot software technologies. With respect to the results obtained from these meetings, the compulsory and elective lectures should be determined so that newly graduated software developers can easily adapt to the software projects in the market. As a result, it will be easier to find a job for them and to hire a proper developer for the companies without giving extra training.

REFERENCES

- [1] Laplante, P. A.. What every engineer should know about. Taylor & Francis Group., Boca Raton, FL, 2007 [Online]. Available:<http://lib.mdp.ac.id/ebook/Karya%20Umum/Every-Engineer-Should-Know-about-Software-Engineering.pdf>
- [2] Mccracken, M., et al. A proposed curriculum for an undergraduate software engineering degree. In: Software Engineering Education Training. // In Proc. 13th Conference on. IEEE, 2000, p. 246-257.
- [3] Zhengyu, G. R. Y. Strengthening Practices and Researches in the Education of Corporation of Enterprises, Colleges and Institutions. Jiangsu Social Sciences, S2, 2007. Available: http://en.cnki.com.cn/Article_en/CJFDTOTAL-JHKX2007S2007.htm
- [4] Kuang, L. Q.; Han, X. The Research of Software Engineering Curriculum Reform. Physics Procedia. [Online] 33(2012), pp. 1762-1767. Available: <http://www.sciencedirect.com/science/article/pii/S1875389212015957>
- [5] Král, J.; Zemlicka, M. Engineering Education-A Great Challenge to Software Engineering. // In Computer and Information Science, 2008. ICIS 08. Seventh IEEE/ACIS International Conference, pp. 488-495.
- [6] Shaw, M.; Herbsleb, J. D.; Ozkaya, I. Deciding What to Design: Closing a Gap in Software Engineering Education. // Invited paper for Education and Training Track of 27th Int. Conf. on Software Engineering (ICSE 2005), 2005, pp. 607 – 608.
- [7] Su, H.; Jodis, S.; Zhang, H. Providing an integrated software development environment for undergraduate software engineering courses. // Journal of Computing Sciences in Colleges. [Online] 23, 2(2007), pp. 143-149. Available: <http://dl.acm.org/citation.cfm?id=1292453>
- [8] Jaakkola, H.; Henno, J.; Rudas, I. J. IT Curriculum as a complex emerging process. // In Computational Cybernetics, 2006. ICC 2006. IEEE International Conference on 2006, pp. 1-5
- [9] Mishra, A.; Yazici, A. An Assessment of the Software Engineering Curriculum in Turkish Universities: IEEE/ACM Guidelines Perspective. // Hrvatski časopis za odgoj i obrazovanje. [Online] 13, 1(2011), pp. 188-219. Available: <http://hrcak.srce.hr/72403>
- [10] Computing Curricula (CC) (2005). Guidelines for Associate-Degree Transfer Curriculum in Software Engineering. Available:http://www.capspace.org/committee/CommitteeFileUploads/TYC_SE_report.pdf
- [11] Chien, C. F.; Chen, L. F. Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry. // Expert Systems with applications. [Online] 34, 1(2008), pp. 280-290. Available: <http://www.sciencedirect.com/science/article/pii/S0957417406002776>

- [12] Mohsin, M. F.; Ahmad, M. F; Din, A. M.; Mahamud, K.R.K.; Din, R. An intelligent trainee selection model. // In Computer Sciences and Convergence Information Technology (ICCIT), 6th International Conference on, 2011, pp. 390-393.
- [13] Ali, M. M.; Rajamani, L. Automation of decision making process for selection of talented manpower considering risk factor: A data mining approach. // In Information Retrieval & Knowledge Management (CAMP), International Conference on 2012, pp. 39-44.
- [14] Smith, D.; Ali, A. Analyzing computer programming job trend using web data mining. // Issues in Informing Science and Information Technology. [Online] 11, 2014, pp. 203-214. Available: <http://iisit.org/Vol11/IISITv11p203-214Smith0494.pdf>
- [15] Ishibuchi, H.; Kuwajima, I.; Nojima, Y. Prescreening of candidate rules using association rule mining and Pareto-optimality in genetic rule selection. // In Knowledge-Based Intelligent Information and Engineering Systems. [Online] 4693, 2007, pp. 509-516. Available: http://link.springer.com/chapter/10.1007/978-3-540-74827-4_64
- [16] Agrawal, R.; Imieliński, T.; Swami, A. Mining association rules between sets of items in large databases. // In ACM SIGMOD Record. 22, 2(1993), pp. 207-216.
- [17] Agrawal, R.; Srikant, R. Fast algorithms for mining association rules. // In Proc. 20th int. conf. very large data bases. VLDB 1215, 1994, pp. 487-499.
- [18] Han, J.; Kamber, M. Data Mining, Southeast Asia Edition: Concepts and Techniques. Morgan Kaufmann. 2006.
- [19] Job Recruitment Website (JRW). 2014. Available: <http://www.kariyer.net/>
- [20] Curriculum Guidelines for Undergraduate Degree Programs in Computer Science (CGCS) (2013, December). Computer Science Curricula 2013 [Online]. Available: <http://www.acm.org/education/CS2013-final-report.pdf>

AUTHORS

Ilkay Yelmen received the BS degree in Software Engineering from the Istanbul Aydin University in 2013 and he is currently pursuing the MS degree in Computer Engineering at the Istanbul Technical University. His research interests are Data Mining and Natural Language Processing.



Metin Zontul received the BS Degree in Computer Engineering from Middle East Technical University, Ankara, Turkey, the MS Degree in Computer Science from Erciyes University, Kayseri, Turkey, and the PhD Degree in Numerical Methods from Cumhuriyet University, Sivas, Turkey. His research interests are software development, information systems, soft computing and data mining. He is currently a faculty member of Software Engineering Department at Istanbul Aydin University, Istanbul, Turkey.



Appendix

Table AI. Main Features and Their Sub Details in Software Job Ads

Education Level	Delphi	Sybase	Action Script	WAF	DDD	Windows Phone	Hudson	JUnit	DirectX
Not acknowledged	Objective-C	Oracle	XSLT	WWF	Comet	Blackberry	PMP	JBatis	JSON
High School	Progress-ABL	MS Sql Server	Vb Script	Data Mining	MVC	Military Status	OCP	Doctrine ORM	Dojo
Associate (Student)	TROIA	Neo4j	DHTML	OLAP	SDLC	Not acknowledged	CHH	Spring MVC	Google Closure
Associate (Graduate)	Ruby	PostgreSQL	XLS	Prmefaces	Web Services	Postponed	Framework	Django	Prototype
B.Sc. (Student)	Foreign Lang. Knowledge	Sql	Xpath	JMX	SOA	Free	PHP5	Magento	Django
B.Sc. (Graduate)	Not acknowledged	Graduated Area	XSL-FO	JPA	WS	Certificate Info	Struts	Travel Ban	OpenGL
M.Sc. (Student)	English	Mathematics	DOM	EJB	SOAP	Not acknowledged	Servlets	Not acknowledged	RUP
M.Sc. (Graduate)	German	Software Eng.	Xquery	JDO	WSDL	MCSO	Spring	Asakd	MFC
Phd (Student)	Work Experience	Computer Eng.	XSD	JPA	UDDI	MCT	Hibernate	Smoking Status	ODM
Phd (Graduate)	Not acknowledged	Electrical and Electronics Eng	XSL	JTA	RISTful	MCPD	Coherence	Not-ue	Enterprise
Operating Systems	1	Industrial Eng.	XAML	JCA	Position	ISTQB	MVC	Unspecified	Social API
Not acknowledged	2	Electronics and Comm. Eng.	Technical Skills	JAXB	Software Architect	CSTE	.NET	Age Information	Facebook
Linux	3	Management Inf. Systems	Not acknowledged	JAX-RPC	Soft. Test Specialist	CISM	Entity	Not acknowledged	Twitter
Unix	4	Statistics	Looking for	JSP	Junior Software Dev	MCDBA	Zend	20-25	OpenGraph
Windows OS	5	Physics	Personal Skills	CRM	Senior Software Dev	PMI	Codeigniter	25-30	Auxiliary Platforms
Macintosh	6	Mathematics-Computer	Not acknowledged	ERP	Medior Software Developer	MCAD	Symfony	30+	Glassfish
Office Applications	7	Computer Programming	Looking for	Special Programming Technics	Mobile Soft. Develop. Specialist	CSTE	CakePHP	IDE	SAP Basis
Not acknowledged	8	Mathematics Eng.	Position Type	N-tier	Project Manager	ISEB	Agavi	Eclipse	IIS
Asked to know	9	Comp. and Ins.Tech. Edu	Part Time	UML	Business Analyst	MCP	Cocoa Touch	Netbeans	Git
Programming Languages	10	Electrical Eng.	Full Time	MVVM	Analyst Developer	CCSE	UIKit	Visual Studio	GitHub
C#	10+	Electronics Eng.	Intern	MVC 3 Razor	Game Programmer	TMAP	Sencha	Dev C++	SVN
Java	Graphic Applications	Engineering	Volunteer	OOWDA	Team Leader	Test Automation and Management Tools	ExtJS	MATLAB	SSIS
C	Not acknowledged	Not acknowledged	Special Software Tech.	Multithreading	Automation Project Engineer	Jira	Android	Xcode	SSAS
C++	Asked to know	Web Technologies (Client-Side)	ORM	MVP	Database Specialist	Selenium IDE	OAF	Emacs	SSRS
PHP	Database	IT/ITML	WCF	Middleware	Software Developer	Apache Jmeter	EJB	Vi	Mercurial
Python	T-Sql	CSS	WPF	Nhibernate	Web Developer	SOAP UI	Silverlight	Borland	TFS
ABAP	Pl-Sql	Ajax	Asp.Net	SEO	Gender	Manis	RichFaces	Zend Studio	Mercurial
Scala	Progress	Jquery	ASP	TDD	Not acknowledged	Bugzilla	ADF	Adobe Air	TFS
Asapta(X++)	DB2	Javascript	RSS	SCRUM	Male	ClearQuest	Play!	CVS	Devexpress
COBOL	My Sql	XML	MS Azure	LINQ to SQL	Female	Mercury	Cross-platform	Libraries	LAMP
Visual Basic	Mongodb	XHTML	JMS	AIM	Mobile Platforms	Firebug	Apache cxf	ExtJS	JMS
Perl	No-SQL	HTML5	J2EE	MDD	IOS	Firebug Lite	Toplink	Timer	Nginx
Unity	Sql Lite	CSS3	Apache Tomcat	Cryptography	Android	JIRA	PhoneGap	QT	Swing

Table AII. The Final Dataset Structure for Software Job Ads

Education Level			Position Type	20-25				
Not acknowledged			Part Time	25-30				
High School		Not acknowledged	Full Time	30+				Not acknowledged
Voca. High School		1	Intern	Certificate Info				1
B.Sc.	C#		Volunteer	Not acknowledged				
M.Sc.	C		Military Status	Analysis				
PHD	PHP	2	Not acknowledged	Project Manage.		Android	UIKit	2
Experience	Java		Postponed	Development		IOS	Cocoa Touch	
Not acknowledged	C++		Free	Database		Blackberry	IOS/OS Other Frameworks	
1	Objective - C	3	Position	Security		Windows Phone	Other Frameworks	3
2	Other Languages		Not acknowledged	Test		Personal Skills	Javascript Library	
3	Oracle		Software Architect (Software Specialist)	Gender		Technical Skills	GUI Library	
4	MS SQL	4	Software Test Specialist	Not acknowledged		Visual Studio	Other Libraries	
5	DB2		Junior Software Dev.	Important		Eclipse	Facebook	
6	NO-SQL		Senior Software Dev.			Xcode	Twitter	
7	MySQL	5	Medior Software Developer			Netbeans	OpenGraph	
8	Sql Lite		Mobile Software Dev. Specialist			Zend Studio	Json	
9	Other Databases		Project Manager			Dev C++	Software Architectures	4
10	HTML/CSS	5+	Business Analyst			Other IDE	Methodologies	
10+	Ajax		Analyst Developer			.NET	Software Patterns	
Foreign Lang. Knowledge	Jquery		Game Programmer			ASP.NET MVC	Prog. Paradigms	
Not acknowledged	HTML5		Web Developer			Framework	Other Soft. Dev.	
Elementary	Other Web Tech.		Automation Project Eng.			Entity Framework	Processes	
Intermediate	The Field graduated	Statistics	Database Specialist			ASP.NET Framework	API	
Advanced	Not acknowledged	Physics	Travel Ban			.NET Other	Software tech.	5
Operating Systems	Software Engineering	Mathematics-Computer	Not acknowledged			Struts	Process Server	
Not acknowledged	Computer Eng.	Computer Programming	Asked			Servlets	Source Control	
Windows OS	Electrical and Electronics Eng.	Mathematics Eng.	Smoking Status			Spring	Other Spedal Prog. Tech.	
Windows - Linux-Unix	Industrial Eng.	Comp. and Ins. Tech. Education	Non-use			Hibernate	Services	
	Electronics Comm.	Graduates from relevant dept.	Unspecified			JUnit	Testing Tools	5+
		Electrical Eng.	Age Info			JSF	Component	
		Electronics Eng.	Not acknowledged			RichFaces	Other Auxiliary Platforms	
Windows - Linux - Unix - MAC OS	Management Information Systems	Mathematics				Other Java Fra.		
						PHPS Framework		
						Zend Framework		
						Other PHP Fra.		
						Android Framework		
						Other Android Fr.		