

# **Recommendation System of Information Technology Jobs using Collaborative Filtering Method Based on LinkedIn Skills Endorsement**

**Latifah Diah Kumalasari**  
Dian Nuswantoro University  
latifahdiah.kumalasari@gmail.com

**Ajib Susanto**  
Dian Nuswantoro University  
ajib.susanto@dsn.dinus.ac.id

**Abstract—** Students who are graduated from Informatics Engineering have wide employment opportunities in the information technology work field, such as database administrator, data scientist, UI designer, IT project manager, network engineer, system analyst, software engineer and UX designer. Each job in Information Technology field has different skill requirement for the interest of work field. Therefore, IT skill classification is needed to find out the suitable career recommendation for Informatics Engineering students. Data from IT professionals which are obtained from LinkedIn account of IT professionals will be processed as reference for students. Data are processed using K-Means Clustering algorithm to find out how is feasible IT professionals data are used as a reference. Then, Collaborative Filtering method by the K-NN algorithm is used to determine classification based on the proximity between student skills and information technology job field. The output is recommendation of information technology job field which are generated from calculate of IT student skills. Result has been tested by testing one of user that has been labeled software engineer produce a recommendation output as a software engineer.

**Keywords—** Classification, Clustering, Collaborative Filtering, IT Career Field Recommendation, K-Means, K-Nearest Neighbor

## **I. INTRODUCTION**

### **A. Background**

Informatics Engineering is a course that focuses on learning the principles of computer science and mathematical analysis for designing, developing, testing and evaluating of software [1]. Department of Informatics Engineering is a computer science major that its graduates students can work in lot branch in Informatic Technology fields such as database administrators, data scientist, UI designer, IT project manager, network engineer, UX designer, etc. It is different with other majors are like doctor education graduation student will work as a doctor, nursing graduation student will work as a nurse, pharmacy graduation student will work as a pharmacist, education major student will work as a teacher, and other majors that are their work sciences have been studied during college. However, informatics engineer students will never be enough if only have high education degree without completed by skills. Informatics engineering students have to expand their skills so they can fulfill the specifications as an expert worker in the certain IT fields. From the lot branch of IT job fields, job applicants must know what specific skills and knowledge are needed when they want job for a particular IT field. Informatics engineering students often do not know benchmark of their ability for work skill requirements in IT fields, so it result to the lack of graduates readiness for confront workplace in Informatics technology fields. Therefore, application system

recommendation career in IT fields base on IT skills and knowledge is necessary. Although every student is equally graduates from Informatic Engineering, they actually have different skills and knowledge.

There is currently a business oriented social networking website which mainly used for professional networking. That social network is LinkedIn. LinkedIn can promote professional online profile of someone to offer employment. So LinkedIn can connect users whom have not known each other in their professional scope [2]. The author will use 400 LinkedIn professional IT profiles from 8 types of IT work fields as a real reference data to make recommendations of 8 IT work fields which most suitable for students whom already have a LinkedIn account. Eight IT work fields which author mean are database administrators, data scientist, UI designer, IT project manager, network engineer, system analyst, software engineer, and UX designer. So in this research, we propose a combination of K-Means and K-NN algorithms to produce recommendations to graduates of informatics engineering based on their current expertise.

## **B. Literature Review**

Tonni Limbong in 2013 with study titled "Implementations Methods of Simple Additive Weighting (SAW) for Selection of Informatics Work Field". This research weighted on academic number values of alumni data in academics information systems database [3].

Yurmalin M. Z. in 2016 also conducted research on the SPK (decision-making system) with title "Decision Making System for Concentration and Interest of Information Technology Course in Yogyakarta Janabadra University". This research is done by found the weight value for each criteria and carried out the rank of interest majors using Fuzzy Multiple Attribute Decision Making (FMADM) method and Simple Additive Weighting (SAW) method [4].

Research titled "A Multi-Criteria Collaborative Filtering Recommender System Using Clustering and Regression Techniques" in 2016 by Mehrbakhsh Nilashi, Muhammad

Dalvi Esfahani, Morteza Zamani Roudbaraki, T. Ramayah, and Othman Ibrahim discusses about the multi-criteria recommendation system that uses many item criteria for rating to improve traditional collaborative filtering accuracy. Techniques used are regression and clustering which developed based on Expectation Maximization (EM) and Classification and Regression Tree (CART). The analysis is done on the dataset provided by Yahoo! Film and TripAdvisor show that the proposed method can improve the accuracy of Multi-Criteria Collaborative Filtering [5].

In 2017 there was research by Laode Aldhi Maulana Ramadhan, Sutardi, and Jumadil Nangi which was discussing about "Creation of E-Commerce Web at Kenime Store Shop Using Recommendation System based Collaborative Filtering with Adjusted Cosine Similarity Algorithm", where the system will provide goods recommendation to user according to his similarity with other users in pattern purchase. Results have been tested by black box method and approved that all test scenarios were completed and have not error at all [6].

Research by Sari Rahmawati, Dade Nurjanah, and Rita Rismala in 2018 regarding "Analysis and Implementation of Hybrid Approaches for Recommendation System with Knowledge-Based Recommender System and Collaborative Filtering Methods". Research combines 2 methods recommender system with weighted technique, where the prediction of knowledge-based recommender system will be combined with collaborative filtering method based on weight rule of two social types (liberal and moderate). The social aperture moderate gives weight of 75% recommendation based on user profile preference and 25% based on other user recommendations. While the liberal social aperture gives weight of 50% recommendation based on user profile and 50% based on other user recommendations. This research gets the best results for interaction prediction is on hybrid social moderate aperture with an average value of RMSE 0,347 of the value  $n=50$  [7].

Another research by Andrias Eko Wijaya and Deni Alfian in 2018 about “Laptop Recommendation System using Collaborative Filtering and Content-Based Filtering”. The study discusses the laptop recommendation system for collaborative filtering method is using adjusted-cosine similarity algorithm to calculate similarity between users, and using weighted sum algorithm to calculate rating prediction that has been given by user on an item. For the content-based filtering method is using TF-IDF algorithm to find the available contents to be recommended. The result show that manual recommendation calculation is same as the system recommendation calculation. Additionally, the execution time by the system is effected by number of items. Collaborative filtering method has an execution time longer than content-based filtering method [8].

In 2018, Jorge Valverde Rebaza and Paul Bustios also conducted a study titled “Job Recommendation Based on Job Seeker Skills”. The research focuses on job recommendations based on the the professional skills of jobseekers. The technique used is the extraction of jobseekers skills using text processing. Algorithms used for recommendation are TF-IDF and 4 types of Word2Vec (Word2Vec-CBOW, Word2Vec-SkipGram, Word2Vec-Ngrams-SkipGram and Word2Vec-Ngrams-CBOW) with data taken are from Catho and LinkedIn sites. The Result of study obtained conclusion that not all profiles are given good recommendation, because maximum value of minimum average effectiveness which occurs on TF-IDF, word2Vec-SkipGram and Word2Vec-Ngrams-CBOW is 0.96 (only 48 from 50 profiles given recommendation) [9].

## C. Operational Definition

### 1. Collaborative Filtering

Collaborative filtering is a recommendation system method based on same assuming between user’s item rating with other user’s item rating that is considered have same liking item.

Collaborative filtering is divided into 3 methods, namely :

### a) Memory-based Collaborative Filtering

Memory-based collaborative filtering is divided into two namely user-item filtering and item-item filtering. User-item filtering takes a specific user, then finds other user like that user based on rating, to recommends items liked by similar user. While, item-item filtering retrieves the item, then finds the user who likes the item to find another item that the user liked. Some techniques utilize memory-based CF are user-based Top-N recommendation algorithm and item-based Top-N recommendation algorithm [10].

### b) Model-based Collaborative Filtering

Model-based produces recommendations by creating models from learning a specific algorithm in the training process. Process of making this model is various, such as Bayesian network, clustering, and rule-based [10]

### c) Hybrid Collaborative Filtering

Hybrid collaborative filtering is a recommendation system method that combines the method of collaborative filtering with other methods, such as content-based so as to complement the advantages and disadvantages of each method so that it can be improve recommender system performance [10].

## 2. Clusterization

Cluster is a collection of data object that have similarities between one and other in same group and different objects with other group’s data. Clustering or cluster analysis grouping a set of physical or abstract objects into single class of the same object [11].

## 3. Classification

One of algorithm for classifying is K-Nearest Neighbor (K-NN). K-NN is a method that use supervised algorithm where the result of newly classified test samples based on majority of categories on K-NN. The purpose of this algorithm is classifying new objects based on trainer’s attributes and samples. The test point will be found number of K objects which closest to test point [12] .

## II. METHOD

### A. Data Preparation

IT professional data profiles which obtained from LinkedIn site consists of 400 profiles will be used as reference for making recommendations. Data profiles are still an early data which contain list of top skills, knowledge, tools and other skills from each profile.

Data of 400 IT professional profiles have collected by author through month of June 2019 comes from 50 Data Scientist profiles, 50 Database Administrator profiles, 50 IT Project Manager profiles, 50 Network Engineer profiles, 50 System Analyst profiles, 50 Software Engineer profiles, 50 User Experience Designer profiles, and 50 User Interaction Designer profiles. Each profile has list of top skills, knowledge, tools, and other skills that already have a endorsement rating provided by LinkedIn users. If point endorsement of top skills, knowledge, tools, and other skills are more higher in value, so those are the more better tops skills, knowledge, tools, and other skills are owned by user. Data 100 top ranking endorsement of top skills, knowledge, tools and other skills are used as attributes to determine recommendation of IT career field for students. Data 100 top endorsement are very factual consider that data 100 top ranking are calculated from the highest endorsements number of top skills, knowledge, tools, and other skills 50 LinkedIn IT professional profile in every IT work field.

### B. System Development Model

Author uses system design techniques with waterfall schemes. The waterfall model is often referred as a linear sequential or classic life cycle [13]. Waterfall model provides sequential flow of software approach starting from analysis, design, coding, and testing. The following is an explanation :

#### 1. Requirement Analysis

Requirement analysis phase is determine list of requirements related to recommendation system application that will be designed. Application of system

recommendation will have this following specifications :

- a) Application will use PHP tool
- b) Application will receive input of user profile which consist of education data, skills, knowledge, rating endorsement of skills obtained by user.
- c) Collaborative filtering using RapidMiner software allows clustering users of IT career fields based on endorsement skills rating.
- d) Collaborative filtering also performs KNN calculation process to obtain classification of the IT career fields for user.
- e) Application performs Euclidian Distance for recommendation process.
- f) Application will provide result a recommendation of IT career fields that best suited for user.

#### 2. Design System

In the design system stage, author will create application models. Author use PHP tools with OOP (Object Oriented Programming) approach. OOP approach requires Unified Modeling Language (UML) as an abstraction about the system will be built. UML is a popular modeling language that has a good system visualization and documentation performance [14]. UML modeling even can generate ready to implement programming codes. UML consists of Use Case, Sequence Diagram, Class Diagram and Activity Diagram.

#### 3. Coding

After design system, the next step is to do coding as PHP tool implementation. Coding must be adjusted with design system which have made.

#### 4. Testing

The final stage in waterfall scheme is testing. Code that has been created will be tested to see whether the application has sufficient the requirement analysis according the initial planning phase.

### III. RESULTS AND DISCUSSION

#### A. K-Means

K-Means algorithm is a clusterization algorithm which grouping data based on the closest centroid cluster with data. Purpose of K-Means is grouping data by maximizing data similarity in one cluster and minimizing data similarity between clusters. The size of similarity used in cluster is distance function. So, maximization of data similarity is obtained based on shortest distance between data and centroid [15].

Author uses method of data clusterization using K-Means algorithm to determine whether the data of 400 professionals IT have been collected are worthy to be used as a reference of IT career recommendation system to students.

#### Stages of clusterization :

- 1) Author only took 50 top ranking of the highest endorsements number of top skills, knowledge, tools and other skills overall from 400 professional IT to be used as a variable in clusterization.
- 2) Data 400 professionals IT (which consist of 50 user Data Scientist, 50 user Database Administrator, 50 user UI Designer, 50 user IT Project Manager, 50 user Network Engineer, 50 user System Analyst, 50 user Software Engineer and 50 user UX Designer) prepared and processed use Microsoft Excel to be processed through the RapidMiner 5.3 software.

Here is initial view of RapidMiner version 5.3

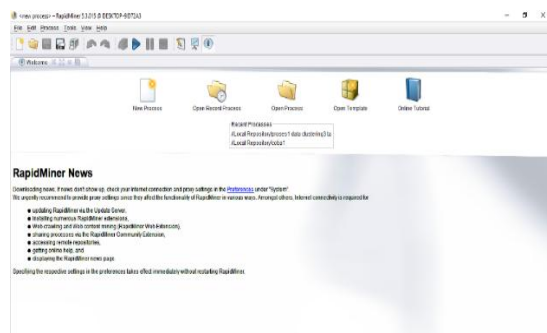


Figure 1. Initial View of RapidMiner Ver 5.3

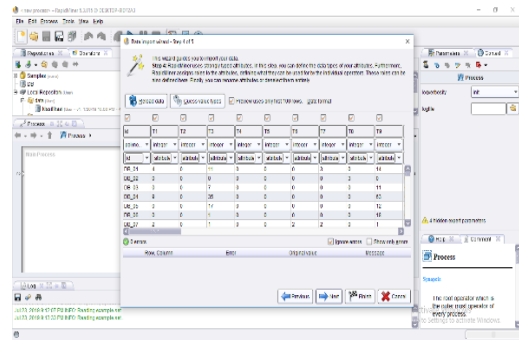


Figure 2. Excel File Configuration View

Next, configure the excel file so that data can be processed in clustering on RapidMiner. Configurations include name configuration and attribute type on table.

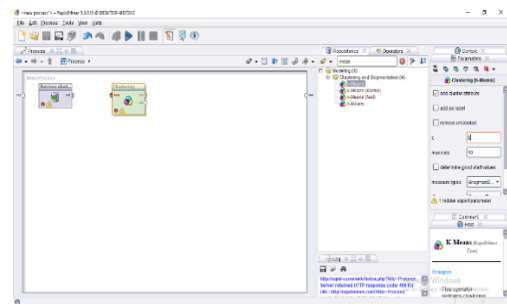


Figure 3. Input View of Excel File and K-Means Operator

Put the selected Excel file with K-Means operation. Fill the value K=8 because author will use 8 types of clusters to see 8 types of data grouping, which will be used as a reference in the next using K-NN method. Connect the Excel file with K-Means operator. Click Run for the data processing process.

Here is the result of processing using K-Means with K value = 8 and max iteration of 10 times.

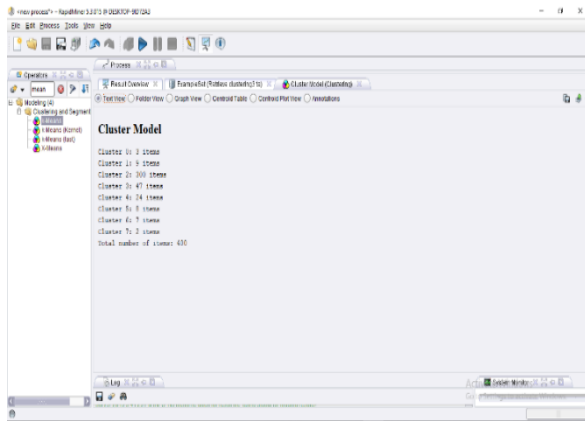


Figure 4. Clusterization Result View

Author obtained the result of clusterization as this following data grouping :

- 0 Cluster: 3 items
- 1<sup>st</sup> Cluster: 9 items
- 2<sup>nd</sup> Cluster: 300 items
- 3<sup>rd</sup> Cluster: 47 items
- 4<sup>th</sup> Cluster: 24 items
- 5<sup>th</sup> Cluster: 8 items
- 6<sup>th</sup> Cluster: 7 items
- 7<sup>th</sup> Cluster: 2 items

The following is a table 3.1 that is a cluster deployment table based on individual IT career fields.

Table 1. Deployment of Clusters Based on Career Fields

No	Career Field	C 0	C 1	C 2	C 3	C 4	C 5	C 6	C 7
1	Database Administrator	3	8	3	9				
2	Data Scientist			3	8	4	8		
3	UI Designer			2	6	2			2
4	IT Project Manager		1	4	9				
5	Network Engineer			5	0				
6	System Analyst			5	0				
7	Software Engineer			2	3	2	0	7	
8	UX Designer			2	5	2	5		

In the career area Database Administrator branched to 3 clusters of 6% C0, 16% C1, and

78% C2. The career field of Data Scientist branched into 3 cluster of 76% C2, 8% C4 and 16% C5. The career field of UI Designer branched into 3 clusters of 52% C2, 44% C3 and 4% C7. The Project Manager IT's career field branched into 2 clusters of 2% C1 and 98% C2. Network Engineer career field is only 1 cluster of 100% C2. The career filed of System Analyst is only 1 cluster of 100% C2. The career field of Software Engineer branched into 3 clusters namely 46% C2, 40% C4 and 14% C7. The career area of UX Designer branched into 2 clusters of 50% C2 and 50% C3. From the research above, it can be known that from 8 clusters of each career field just branched out to maximum of 3 clusters or not more than half the number clusters specified. So, data is worth to be used as a reference recommendation system.

### B. K-NN

K-NN is an algorithm includes a group of instance-based learning. This is one of lazy learning techniques. K-NN is done by looking for group K objects in the closest training data with object in new data or data testing [16].

In this K-NN research author use k=1,2,3,4,5,6,7,8 to determine classification of 8 IT careers areas, and to calculate the distance between user testing and data 8 IT career field by using Euclidean Distance. The smallest distance between user testing and IT career field will be used as a recommendation.

User testing SE\_19 data that will be used as an example of the K-NN calculation testing. User SE\_19 is the user who was originally labeled Software Engineer.

Calculation of distance between user SE\_19 with each area of IT career fields is calculated by comparing the distance of skills user SE\_19 and skills in each IT career fields.

**Table 2. Point Endorsements of User SE\_19's Skills, Knowledge, Tools & Other Skills**

No	Top Skills, Knowledge, Tools, and Other Skills User SE_19	Ranking Endorsement
1	Algorithms	9
2	Artificial Intelligence	15
3	Automated Planning and Scheduling	0
4	Big Data	2
5	Big Data Analytics	0
6	Bioinformatics	0
7	C++	7
8	Computational Biology and Bioinformatics	0
9	Computer Science	5
10	Constraint Programming	0
11	Data Mining	2
12	Database Management	0
13	Deep Learning	0
14	Distributed Systems	2
15	Eclipse	1
16	Hadoop	0
17	Information Retrieval	0
18	Integer Programming	0
19	Java	6
20	Knowledge Representation and Reasoning	0
21	LaTeX	1
22	Linux	0
23	Machine Learning	13
24	Natural Language Processing	0
25	NoSQL	0
26	Object Oriented Design	0
27	Object Oriented Modeling	0
28	Optimization	7
29	Program Optimization	0
30	Programming	9
31	Python	4
32	R	0
33	Reusable Software Design	0
34	Scalable System Design	0
35	Software Design	1
36	Software Engineering	4
37	SQL	0
38	Statistics	2

39	UML	0
40	Version Control	0

1. Calculation distance between user SE\_19 with each area of IT career fields is calculated by comparing the distance of skills user SE\_19 and skills of each IT career fields using Euclidian Distance formula.

$$d(x_i, c_j) = \sqrt{\sum_{i=1}^N (x_i - c_i)^2} \dots\dots\dots(1)$$

**Description :**

d(x<sub>i</sub>,c<sub>j</sub>)= Euclidian distance between i<sup>th</sup> skill with center skill i  
 x<sub>i</sub> = x value in the i<sup>th</sup> attribute  
 c<sub>j</sub> = c value on center i  
 i = individual attributes 1 to n

a. d(DS, SE\_19) =

$$\sqrt{\sum (\overline{DS} - SE_{19})^2} = 33,8900$$

b. d(DB, SE\_19) =

$$\sqrt{\sum (\overline{DB} - SE_{19})^2} = 21,7300$$

c. d(PM, SE\_19) =

$$\sqrt{\sum (\overline{PM} - SE_{19})^2} = 11,8600$$

d. d(NE, SE\_19) =

$$\sqrt{\sum (\overline{NE} - SE_{19})^2} = 4,6600$$

e. d(SA, SE\_19) =

$$\sqrt{\sum (\overline{SA} - SE_{19})^2} = 7,8200$$

f. d(SE, SE\_19) =

$$\sqrt{\sum (\overline{SE} - SE_{19})^2} = 49,3000$$

g. d(UX, SE\_19) =

$$\sqrt{\sum (\overline{UX} - SE_{19})^2} = 22,5700$$

h. d(UI, SE\_19) =

$$\sqrt{\sum (\overline{UI} - SE_{19})^2} = 27,6100$$

2. This classification stage does not use any model, because it only match based on memory. The formula for calculating this algorithm is :

$$Distance \{X, A\} = \frac{\sqrt{\sum(\bar{A} - X)^2}}{\sqrt{\sum \bar{A}}} \dots\dots(2)$$

- a. Distance (user testing, data scientist)  
 $Distance \{SE_{19}, data\ scientist\} = \frac{33,8900}{37,7444} = 0,8979$
- b. Distance (user testing, database administrator)  
 $Distance \{SE_{19}, database\ administrator\} = \frac{21,7300}{21,7411} = 0,9996$
- c. Distance (user testing, project manager)  
 $Distance \{SE_{19}, project\ manager\} = \frac{11,8600}{11,9561} = 0,9924$
- d. Distance (user testing, network engineering)  
 $Distance \{SE_{19}, network\ engineering\} = \frac{4,6600}{4,7027} = 0,9912$
- e. Distance (user testing, system analyst)  
 $Distance \{SE_{19}, system\ analyst\} = \frac{7,8200}{8,1191} = 0,9627$
- f. Distance (user testing, software engineering)  
 $Distance \{SE_{19}, software\ engineering\} = \frac{49,3000}{58,8306} = 0,8380$
- g. Distance (user testing, ux designer)  
 $Distance \{SE_{19}, ux\ designer\} = \frac{22,5700}{22,6109} = 0,9981$
- h. Distance (user testing, ui designer)  
 $Distance \{SE_{19}, ui\ designer\} = \frac{27,6100}{27,6488} = 0,9986$

3. The last stage is determination of the user testing career classification based on the skills distance which are owned by user testing with each IT career fields. Based on that calculations, user testing SE\_19 has distance with each IT career fields as follows :

**Table 4. Recommendation Results of User SE\_19**

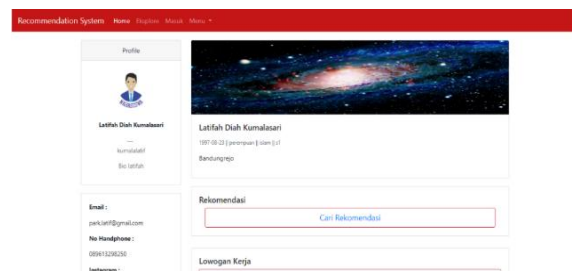
IT Career Fields	Distance	Rank
Data Scientist	0,8979	2
Database Administrator	0,9996	8
Project Manager	0,9924	5
Network Engineer	0,9912	4
System Analyst	0,9627	3
Software Engineering	0,8380	1
UX Designer	0,9981	7
UI Designer	0,9986	6

From the table above, User SE\_19 has first recommendation on software engineer, second recommendation on data scientist, third recommendation on system analyst, fourth recommendation on network engineer, fifth recommendation on project manager, sixth recommendation on UI designer, seventh recommendation on UX designer, and eighth recommendation on database administrator.

### C. Program Implementation

#### 1. Student Dashboard

Student can manage student profile data, input skills, point endorsements of skills and seek recommendations.



**Figure 5. Student Dashboard Screen**



## 2. Find Student Recommendations

On search recommendations page will show list of skills that point endorsements skills must be input by student.

No	ID Skills	Name Skills	Total Endorsement	Point Average	Add Skills
1	SK_DE_001	Agile Methodologies	19	0.30	Add
2	SK_DE_002	Java	17	0.24	Add
3	SK_DE_003	Apache	29	0.10	Add
4	SK_DE_004	Business Analysis	0	0.16	Add
5	SK_DE_005	Business Intelligence	53	1.08	Add
6	SK_DE_006	C	17	0.34	Add
7	SK_DE_007	C++	8	0.18	Add

Figure 6. Screen of Search Student Recommendation

## 3. Student Recommendation Results

On recommendation results page will produce output of recommendation according to the skills and point endorsements skills entered by student.

Rank	Job Role	LinkedIn ID
1.	Database Administrator	0.99966821751491
2.	Data Scientist	0.89792884564067
3.	Interaction Designer	0.99863836965397
4.	IT Project Manager	0.99236104167146
5.	Network Engineer	0.99124367758921
6.	Sistem Analis	0.96269328289307
7.	Software Engineer	0.83802557329012
8.	UX Designer	0.99806876150122

Figure 7. Screen of Student Recommendation Results

## IV. CONCLUSION

1. Results of clusterization by using RapidMiner tool version 5.3 show that Database Administrator users branched into 3 clusters (6% C0, 16% C1 and 78% C2). Data Scientist users branched into 3 clusters (76% C2, 8% C4 and 16% C5). UI Designer users branched into 3 clusters (52% C2, 44% C3 and 4% C7). IT Project Manager users branched into 2 clusters (2% C1 and 98% C2). Network Engineer users and System Analyst users only on C2 cluster. Software Engineer users branched into 3 clusters (46% C2, 40% C4 and 14%

C7). UX Designer users branched into 2 clusters (50% C2 and 50% C3).

2. Implementation of K-NN algorithm to determine classification use  $K=8$  has been tested on user SE\_19 which originally labeled Software Engineering also resulted in a recommendation of first rank in Software Engineering, then followed by Data Scientist, Network Engineer, Project Manager, UI Designer, UX Designer and Database Administrator.

## REFERENCES

- [1] youthmanual, "Teknik Informatika," 2015. .
- [2] P. O. Indonesia, "Mengenal LinkedIn dan Manfaatnya," 2012. .
- [3] T. Limbong, "Implementasi Metode Simple Additive Weighting (Saw) Untuk Pemilihan Pekerjaan Bidang Informatika," no. August, 2016.
- [4] C. Journal and U. Janabadra, "Sistem Pendukung Keputusan Konsentrasi dan Peminatan Prodi Teknik Informatika Universitas Janabadra Yogyakarta," vol. 3, no. 4, pp. 307–318, 2016.
- [5] O. I. Mehrbakhsh Nilashi, Mohammad Dalvi-Esfahani, Morteza Zamani Roudbaraki, T. Ramayah, "A Multi-Criteria Collaborative Filtering Recommender System Using Clustering and Regression Techniques," vol. 3, p. 7, 2016.
- [6] J. N. Laode Aldhi Maulana Ramadhan, Sutardi, "Pembuatan Web E-Commerce Pada Toko Kenime Store Menggunakan Sistem Rekomendasi Berbasis Metode Collaborative Filtering Dengan Algoritma Adjusted Cosine Similarity," vol. 3, p. 10, 2017.
- [7] dan A. P. K. Ramadhanuz A Djamal, Warih Maharani, "Analisis Dan Implementasi Metode Item-Based Clustering Hybrid Pada Recommender System," 2010.
- [8] D. A. Anderias Eko Wijaya, "Sistem Rekomendasi Laptop Menggunakan Collaborative Filtering Dan Content-Based Filtering," vol. 12, p. 17, 2018.

- [9] P. B. Jorge Valverde Rebaza, "Job Recommendation Based on Job Seeker Skills," p. 6, 2018.
- [10] T. M. K. Xiaoyuan Su, "A Survey of Collaborative Filtering Techniques," vol. 1, p. 19, 2009.
- [11] E. Irfiani and S. S. Rani, "Algoritma K-Means Clustering untuk Menentukan Nilai Gizi Balita," vol. 6, no. 4, pp. 165–172, 2018.
- [12] M. A. Nugroho and H. A. Santoso, "Klasifikasi Dokumen Komentar Pada Situs Youtube Menggunakan Algoritma K-Nearest Neighbor ( K-NN )," no. 5.
- [13] S. Susilowati and M. Informatika, "Pengembangan Sistem Informasi Manajemen Zakat , Infaq , Shadaqoh , Waqaf dan Hibah Menggunakan Metode Waterfall," vol. 19, no. 1, pp. 52–60, 2017.
- [14] A. D. Supriatna, "Systematic Design of Expert System Using Unified Modelling Language Systematic Design of Expert System Using Unified Modelling Language," 2018.
- [15] R. A. ASRONI, "Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang," vol. 18, p. 7, 2015.
- [16] A. Sulistiyo, "Penentuan Jurusan Sekolah Menengah Atas Menggunakan Metode K-Nearest Neighbor Classifier Pada SMAN 16 Semarang," p. 5.