

PERFORMANCE PREDICTING IN HIRING PROCESS AND PERFORMANCE APPRAISALS USING MACHINE LEARNING

Ali A. Mahmoud

*Department of Computer Sciences, ESE King Hussein
School of Computing Sciences,
Princess Sumaya University for Technology,
Amman, Jordan
ali.mah989@gmail.com*

Walid A. Salameh

*Department of Computer Sciences, School of Computer
Science,
Princess Sumaya University for Technology,
Amman, Jordan
walid@psut.edu.jo*

Tahani AL Shawabkeh

*Department of Computer Sciences, CS King Hussein
School of Computing Sciences,
Princess Sumaya University for Technology,
Amman, Jordan
tahanishawabkeh@gmail.com*

Ibrahim Al Amro

*Department of Computer Sciences, CS King Hussein
School of Computing Sciences,
Princess Sumaya University for Technology,
Amman, Jordan
ibrahimalamro760@gmail.com*

Abstract—In the nowadays-competitive race of finding a suitable talented, qualified, bright and potential personnel to fulfill the needed spot of a vacancy in an industry, and with the beginning of the fourth industrial revolution, employers are taking the hiring process to the digital world. As Artificial Intelligent (AI) has a high-speed computation and adaptively to big data, it used to analyses and represents the data to employers in an easy way so they can make their decisions effectively.

An upcoming challenge is raised where if the new candidates for a vacancy will give the expected performance based on the hiring criteria's or not, and how to hire a candidate that will while dealing with the hiring process? Employers are concerned with the performance evaluation of their current employees, but it is a challenge knowing the performance of new candidates before hiring.

This study is proposing a follow-up conceptual model of using Artificial Intelligent (AI) in the hiring process with the using of performance management and social screening to predict the new candidate expected performance by analyzing historical performances and conditions of employees. This method will give an additional parameter that assists the decision makers in the hiring process.

Although this method is a step forward to eliminate bad hiring, but it is requiring a huge historical data including the tracking of the performance, personal information collected from several sources like surveys and social Media and employees conditions related to the time of old and current employees, to give results that are more efficient and accurate.

Index Terms—Data Mining; Classification; Decision Tree; Job Performance; Performance Appraisals; Artificial Intelligence (AI).

I. INTRODUCTION

ONE of the most concerns to any business industry is the human's capital, which focuses on profits maximization and cost minimization (WallerStone, 1980). Human resource is a critical process in the capitalization, finding peoples with effective ideas to benefit the financial impact of the industry and avoid hiring people with non-effective ideas who can cost the industry financially. More employers than ever are struggling to fill open jobs. Forty-five percent say they cannot find the skills they need, and for large organizations (250+ employees), it is even higher with 67% reporting talent shortages in 2018. Accordingly, employers are experiencing a financial loss due to bad hiring decisions, according to a survey the average cost of one bad hire is nearly \$15,000; average cost of losing a good hire is nearly \$30,000(CareerBuilder, 2017).

Another task of the Human resources is performance management, the process where HR is keeping tracking the employee's performances and ensures the motion of productivity is meeting the industry's goals. A used evaluation technique to create a solid ground for the performance management is the Performance Appraisals, or also called Performance review or performance evaluation (Muchinsky,2012), which is a process where the duties & responsibilities of a position are compared with the standards of performance stated in the Job Description.

Heyel observes, "it is the process of evaluating the performance and qualifications of the employees in terms of the requirement of the job, for which he is employed, for purposes of administration including placement, selection for promotions, providing financial rewards and other action which require differential treatment among the members of a group as distinguished from actions affecting all members equally others regard, it as a process of estimating or judging the value excellence, qualities or status of some object, person or thing."

As for performance appraisal needed to judge the level of employee amendment to his role and to the institution, the performance management needed to control, enhance and to improve the employee performance level.

The concerning is, Will the new candidates for a vacancy give the expected performance based on the hiring criteria's or not? In addition, how to hire a candidate that will meet the performance expectation while dealing with the hiring process?

Employers are concerned with the performance evaluation of their current employees, but it is a challenge knowing the performance of new candidates before hiring. In the hiring process, finding an approximate estimation of the new candidate performance will assist the recruiters with the decision of the hire.

Normal techniques of performance predicting can be sophisticated and timely cost, knowing who is going to perform best in the required job role. Methods such as knowledge tests, cognitive tests, personality tests, reference checks, structured/unstructured interviews, work samples, and integrity tests and others are used in the normal hiring process (Warton, people Analytic).

To overcome the drain of using such routine methods in the hiring process, the HR systems integrated with artificial intelligence (AI) solutions. These artificial intelligence (AI) applications used to automate HR administration to decrease the time consumed in routine hiring tasks. John McCarthy represented artificial intelligence (AI) in 1956, he described it "Artificial Intelligence is the science and engineering of making intelligent machines, especially intelligent computer programs". In other words, artificial intelligence (AI) technologies are tools that programmed to simulate human ways of decision-making using methods such as machine learning, or advance searches algorithms or statistical analysis. More and more researches come after describing the various types of artificial intelligence (AI) techniques and their applications.

An example of artificial intelligence (AI) applications used in the HR process is Chatbots; a conversational-based tool helps manage and response in the pre-application processes (Artisan 2017). In addition, artificial intelligence (AI) machine learning software on social media or called screening a process whereby employers browse social media profiles of applicants in order to find positive or negative information that may help them decide if the candidate is suitable for an open position. Social media screening is very valuable to recruiters, hiring managers and employers, it humanizes the candidate and makes them more than just the words on their resume. There are various tools for social media screening such as LinkedIn Profile API, FAMA, and Meltwater.

Predicting performances for new candidates using artificial intelligence (AI) presented by Q. AlRadaideh and E. Alnagi in 2012 [6], using the techniques of data mining; they built a classification model using CRISP-DM (Cross Industry Standard Process for Data Mining) methodology and decision tree tool. They used a data set contains personal information collected by a survey from 3 different IT companies to build a classification that tests certain attributes that may affect job performance. This data set is processed through two main steps, learning step which analysis a set of training database instances and a testing

step which processing a different data set to estimate the classification accuracy.

This research proposing additional parameters and attributes to build a conceptual method combining the artificial intelligence (AI) techniques that proposed by AlRadaideh and Alnagi. With the using of performance management and Screening techniques, a procedure that can help decision-makers with finding and hiring the best candidate by predicting his/her performance according to generated performance patterns using Machine-learning techniques based on historical performance Appraisals of other existing employees in addition to the used data set [17].

II. LITERATURE REVIEW

Machine learning is one type of artificial intelligence methods; it is a computational program goaled to optimize a process performance based on learned information extracted from training on a data set. One use of this ML learned information is predictions in the future or description of a status.

In the proposed module by (AlRadaideh and Alnagi 2012), the data set was a table of personal information such as Age, Gender, Marital status, education degree...etc. collected among three IT companies using a survey. This information assumed to have an effect on the performance of an employee, applying ID3, C4.5 and Naïve Bayes algorithms using J4.8 in WEKA to generate a decision tree, which returned highest information gain for the attribute "JobTitle".

TABLE I
DATA SET USED IN (ALRADAIDEH AND ALNAGI 2012) MODEL

<i>Attribute</i>	<i>Description</i>	<i>Possible Values</i>
Age	Employee's Age	a (<25), b (25-29), c (30-34), d (35-40), e (>40)
Gender	Employee's Gender	Male, Female
MStatue	Employee's Marital	Single, Married
NKids	No. of Kids	0, 1, 2, 3, 4
UnivType	The type of university of graduation	Public, Private
GSpecial	General Specialization	Business, IT, English Literature, Engineering, CS, Other
Degree	Employee Education Degree	Diploma, Bachelor, High Diploma, Master, PhD
Grade	Employee Graduation	Excellent, Very good, Good,
Expyears	No. of Working Experience Years	a(none), b(<1year), c(1-5years), d(6-10years),e(>10years)
PrevCo	No. of Previous Companies the employee worked for	0,1,2,3,4,5
JobTitle	Employee's Job Title in the current company	Developer, Officer, QA, Data Entry, System Administrator, Office Manager, Technical Writer, Technical Manager Software Engineering, Accountant, Infrastructure Engineer, Department Manager, software Architect,

Analyst, Designer, Trainer,
PM, Consultant

Rank	Employee's Rank in the current company	Junior, Senior, Team Leader, Manager
ServPeriod	Service Period in the current company(years)	a(<1year), b(1-5years), c(5-10years), d(>10years)
UncomWorkcond	Working in uncomfortable conditions (employee's perspective)	Yes, No
Dissatsalrank	Existence of dissatisfaction in either salary or rank	Yes, No
Performance	Employee's performance either informed or predicted. This is a class	Accomplish, Exceed, Far Exceed

Table1: Data Set Used In (Alradaideh And Alnagi 2012) Model.

The data set has experimented in 3 phases; the first phase used the whole data set (130 instances), the second phase used the data set collected from company 1 (37 instances) and the final phase used the data set collected from company2 (38 instances). Applying the ID3, C4.5 to generate a decision tree and Naïve Bayes algorithms in each phase to build the classification model. Results are shown in Table2.

TABLE2
DATA SET USED IN (ALRADAIDEH AND ALNAGI 2012) MODEL

Method	10-Fold Cross Validation			Hold-out (60%)		
	Phase 1	Phase 2	Phase 3	Phase 1	Phase 2	Phase 3
ID3	0.369	0.378	0.5	0.365	26.7%	0.437
C4.5	0.423	0.486	0.605	0.481	0.533	0.562
Naïve Bayes	0.407	0.378	0.658	0.442	46.7%	0.687

Table2: Data Set Used In (Alradaideh And Alnagi 2012) Model.

This model predicts the evaluation of new data inputs depending on the total evaluation value of training data, it does not concentrate the prediction to a certain area of skill performance.

III. HYPOTHESIS AND DESIGN

The proposed enhancement will result a prediction not only for the performance evaluation as a total but on the level of the skill by dividing the historical evaluation to be depending on the skills areas, this addition to the attributes will raise the specificity and accuracy of the prediction process.

In order to have the maximum accuracy, the performance evaluation processed on areas separately, as an example; the employee evaluated on the level of Ethics and then evaluated on another area of skills separately.

A. The Performance evaluation process

For this study, an appraisal form taken from a local IT company, where it uses multiple inputs according to its policy, such as a code of ethics and tasks performances.

This form was taken due to its specification on the skills area in the evaluation process, each area is evaluated separately, at the end of the evaluation calculation, a scaled value of evaluation calculated transformed into a code of evaluation (Table3).

TABLE3
PERFORMANCE APPRAISAL TABLE

Performance Levels	Rating code	Rating Description	Rating Range
Exceed Expectation	EE	Exceptional performance that exceeds requirements, Performance and results achieved consistently exceed expectations for the position requirements and objectives	95% - 110%
Meets All Expectations	ME	Meets the expected performance requirements. Performance and results generally meet the expectations for the position requirements and objectives. Performance requires normal degree of supervision.	80% - 94%
Meets Most Expectations	MME	Meets some expected performance requirements. Performance and results achieved generally do not meet established objectives. Performance requires more than normal degree of supervision	55% - 79%
Below Average	BA	Significant shortfall in expected performance. Performance and results achieved consistently do not meet	0% - 54%

Table3: Performance Appraisal Table.

TABLE4
PERFORMANCE APPRAISAL TABLE AREAS

Key Performance Areas	Performance sub Area	%	Key Performance Evaluation	Rating Code
Ethics	Display passion and optimism	0.6875	0.798958333	ME
	Work as a team	0.8625		
	Invest time in continuous learning	0.825		
	Be respectable	0.825		
	Build relationship based on trust	0.825		
	Act with integrity	0.825		
	Anticipate customers needs	0.825		
	Exercise empowerment	0.6875		
	Drive innovation	0.75		
	Never let a customer down	0.825		
Competencies	Be an ambassador	0.825	0.939583333	ME
	Think and act like an owner	0.825		
	Customer Focus:	0.9625		

	Flexibility, Adaptability: Communication & Collaboration: Attention to details and adherence to procedures: Planning & Organizing: Applying Expertise and Professional Knowledge:	0.9625		
Skill a	project 1	0.9625	0.9625	EE
	Project 2	0.9625		
Skill b	Project 3	0.9625	0.89375	ME
	Project 4	0.825		
Skill c	Project 5	0.9625	0.9625	EE
	Project 6	0.9625		
Skill d	project 8	0.825	0.6875	MME
	project 9	0.55		
Skill e	Project 10	0.9625	0.89375	ME
	Project 11	0.825		
Over All		0.876934524		ME

Table4: Performance Appraisal Table Areas.

When the evaluation process finishes, the evaluation results categorized to areas in such how the employee performed in each area in that period of evaluation, then stored in the historical evaluations. These evaluation areas are bounded to the employees' statuses in the same period of evaluation, these values are processed together to initiate the bounded values and stored in the knowledge base. More the evaluations and more the statuses, more the values and new related information generated and stored in the knowledge base.

B. Employee Statuses

Employee Status is the set of information's related to an employee in the time of evaluation process such as gender, academic degree, skills, field experience, health condition...etc. all these data are collected regularly and integrated with his/her evaluation in the same period in which it collected.

This information is collected from multiple sources and maintained in the data set. The regular surveys at the time of evaluations, social media screening and areas performance evaluations are the key sources of this data set.

Certain patterns of performances learned by relating these combined evaluations and collected data. When a new candidate is applying for a vacancy, his/her information collected and bounded to the job role of application, these inputs compared with the results in the knowledge base. Based on these learned patterns, a prediction of a new candidate performance evaluation is derived, comparing a single area in the new evaluation with a similar area in previous evaluations a prediction in this certain area is also can be predicted. Such as, if an employee X with a certain status performance in project A in evaluation 2016Q4 is 55% (MME) and this project is categorized in the area of Sales skills is presented as:

Employee (x, status1, Sales, 2016Q4) is 55% and the employee performance in project B in evaluation 2017Q1 is 45% (BA) and this project categorized in the area of Sales skills as: Employee (x, status2, Sales, 2017Q1) à 45% a pattern will be learned that this employee skill of sales is

dropping and need attention. New information will be kept in the knowledge base for this employee as an example (potential weakness in Sales Skill for the employee (X, status). The variation of the Employee X status between 2016Q4 and 2017Q1 is marked as leading to this performance changes if the same variation leads to the same performance changes with other employees.

When combining a number of employees with the same criteria and same status range

i.e Employee (A,status1,Sales) → X1%

Employee (B, status1, Sales) → X2%

Employee (C, status1, Sales) → X3%

.

.

.

Employee (Z, status1, Sales) → Xn%

We can generate a pattern such as an employee (n, status1, Sales) → average from above

When a new candidate is applying for a certain vacancy, this vacancy that contains the evaluation area, the status of this candidate collected and compared with the values in the space of knowledge. The method should give an estimation of this candidate predicted performance due to the status provided.

C. Using ML Methods For Predicting

A number of machine learning methods used for prediction techniques as AIRadaideh and Alnagi Model used decision trees with ID3 and C4.5 algorithms in addition to Naïve Bayes to generate a decision tree based on one class, which is the performance.

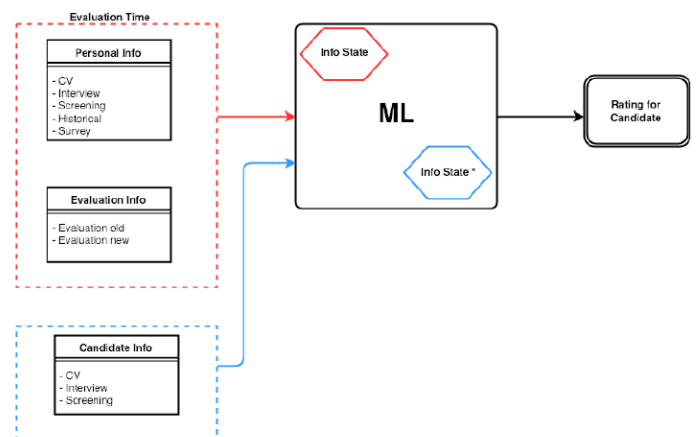


Figure 1: High-Level diagram of proposed model data flow.

This approach will use the same algorithms to find the enhancements of adding new attributes to the data set. The more steps in this experiment are to repeat the training with a different class each time, every time the generated tree will classify the data set according to a performance area, this will allow predicting the performance of the candidate in that certain area.

TABLE5
DATA SET USED IN THE PROPOSED MODEL

Attribute	Use	Attribute	Use
Age	Input	SalRange	Input

Gender	Input	UncomWorkcond	Input
MStatue	Input	Dissatsalrank	Input
NKids	Input	Screening Att1	Input
UnivType	Input	Screening Att2	Input
GSpecial	Input	Screening Att2	Input
Degree	Input	Ethics	Target
		Performance	
Grade	Input	Competencies	Target
		Performance	
Expyears	Input	Skill a	Target
		Performance	
PrevCo	Input	Skill b	Target
		Performance	
JobTitle	Input	Skill c	Target
		Performance	
Rank	Input	Skill d	Target
		Performance	
ServPeriod	Input	Skill e	Target
		Performance	

Table5: Data Set Used In Proposed Model.

The attributes Screening Att1, Screening Att2, Screening Att3 eliminated because the employees who have done the survey are anonymous.

Using WEKA (v3.8.3) explorer to generate the J48 tree with a cross-validation Fold 10, the "Skill a" selected as the class of the decision tree, "Job Title" attribute has the highest information gain as follows:

J48 pruned tree

```

-----
Job Title = Developer
| Experience years Range = e: MME (1.0)
| Experience years Range = c
|| Salary Range = c: ME (5.0/3.0)
|| Salary Range = b: ME (15.0/6.0)
|| Salary Range = a: EE (3.0/1.0)
|| Salary Range = d: BA (1.0)
|| Salary Range = e: ME (0.0)
| Experience years Range = d
|| University Type = Private: MME (3.0/1.0)
|| University Type = Public
||| Rank = TeamLeader: ME (3.0)
||| Rank = Senior: EE (2.0)
||| Rank = Junior: ME (0.0)
||| Rank = Manager: ME (0.0)
||| Rank = Architect: ME (0.0)
| Experience years Range = b: ME (3.0/1.0)
| Experience years Range = a: ME (1.0)
Job Title = Officer: EE (7.0/4.0)
Job Title = QA
| University Type = Private: MME (5.0/1.0)
| University Type = Public
|| Age = c: BA (0.0)
|| Age = a: BA (2.0)
|| Age = b: ME (3.0/1.0)
|| Age = d: BA (0.0)
|| Age = e: BA (0.0)
Job Title = DataEntry: ME (2.0)
Job Title = SysAdmin: ME (6.0/2.0)
Job Title = OfficeManager: EE (4.0/1.0)
Job Title = TechWriter
| Age = c: EE (0.0)
| Age = a: EE (5.0/2.0)
| Age = b: ME (2.0/1.0)
| Age = d: EE (0.0)

```

```

| Age = e: EE (0.0)
Job Title = TechManager: EE (4.0/1.0)
Job Title = SoftEng
| No. of Kids <= 2: ME (20.0/10.0)
| No. of Kids > 2: MME (2.0)
Job Title = Accountant: BA (1.0)
Job Title = InfraEng: EE (1.0)
Job Title = DeptManager
| Dissatisfaction with the salary or rank? = Yes: BA (2.0)

| Dissatisfaction with the salary or rank? = No
|| University Type = Private: MME (2.0/1.0)
|| University Type = Public
||| Marital Statue = MWKids: EE (3.0)
||| Marital Statue = Single: ME (3.0/1.0)
||| Marital Statue = MNKids: EE (0.0)
||| Marital Statue = Other: EE (0.0)
Job Title = SoftArch: MME (2.0/1.0)
Job Title = Analyst: ME (3.0/1.0)
Job Title = Designer
| Gender = Male: MME (2.0)
| Gender = Female: ME (2.0/1.0)
Job Title = Trainer: ME (1.0)
Job Title = PM: ME (2.0/1.0)
Job Title = Consultant: EE (3.0/1.0)
Job Title = CustSup: ME (3.0/1.0)
Job Title = GM: ME (1.0)

```

Number of Leaves : 49

Size of the tree: 62

Correctly Classified Instances 91 70 %

Incorrectly Classified Instances 39 30 %

Repeating the same for the Skills attributes and performances areas attributes to generate a separated prediction for each area.

Compared to the proposed model by AlRadaideh and Alnagi the tree output is as following:

J48 pruned tree

```

-----
Gender = Male
| Grade = Acceptable: Accomplish (2.0)
| Grade = Good
|| Salary Range = c
||| General Specialization = Business: Exceed (0.0)
||| General Specialization = IT: Exceed (2.0)
||| General Specialization = EngLiterature: Exceed (0.0)
||| General Specialization = Engineering: Exceed (2.0)
||| General Specialization = CS: Accomplish (3.0)
||| General Specialization = Other: Exceed (0.0)
|| Salary Range = b: Accomplish (10.0/2.0)
|| Salary Range = a: Accomplish (1.0)
|| Salary Range = d: Accomplish (0.0)
| Grade = Vgood: Exceed (3.0/1.0)
| Grade = Excellent: Exceed (3.0)
Gender = Female: Accomplish (12.0/1.0)

```

Number of Leaves : 13

Size of the tree : 17

Correctly Classified Instances 24 63.1579 %

Incorrectly Classified Instances 14 36.8421 %

IV. CONCLUSION

It is difficult to find a suitable talented personal to fit the job needed, normal hiring process is time-consuming and bad hiring can cause a financial loss.

Moving to the digital evolution is a trend in the hiring process; a number of solutions such as INDEED, CAREER, Google for jobs and others or a 3rd party agencies such as ARYA's solutions are now in the market using the artificial intelligence (AI) techniques to help to find the correct choice of candidates to fulfill the needed roles. Still, the recruiters are depending on attitude hiring more than skills hiring.

This paper shows that the more attributes considered in the calculation of predicting more prediction accuracy, although not much research has been worked on this idea clearly, so we worked in this model to narrow the gap between skills and attitude hiring by predicting the performance of a new candidate in multiple areas, areas that hard to be decided in the regular hiring process.

In order to have the most accurate approximation for a new candidate performance, a huge amount of data needs to be processed. More the evaluations history and employee statuses more reasonable approximation outputted. In addition to the variance of employees natures and skills.

With the deep search in the machine learning algorithms, and more about applying the Multilayers Machine learning algorithms in prediction algorithms, this approach considered as future work to be implemented instead of decision trees.

V. REFERENCES

- [1] Transforming Human Resources into Human Capital, S.Zakaria,W.Yusoff, Information Management and Business review, Vol 2,No2,PP 48-54,Feb 2011
- [2] a survey by Harris Poll on behalf of CareerBuilder between August 16 and September 15, 2017 <http://press.careerbuilder.com/2017-12-07-Nearly-Three-in-Four-Employers-Affected-by-a-Bad-Hire-According-to-a-Recent-CareerBuilder-Survey>
- [3] A Study on the Effectiveness of Performance Appraisal System and its Influence with the Socio-Demographic Factors of the Employees of a Manufacturing Industry in Tamil Nadu, International Journal of Research in Management & Business Studies (IJRMBS 2015), Vol. 2 Issue 1 Jan. - Mar. 2015
- [4] RECRUITMENT THROUGH ARTIFICIAL INTELLIGENCE: A CONCEPTUAL STUDY, International Journal of Mechanical Engineering and Technology (IJMET), Volume 9, Issue 7, July 2018, pp. 63–70, Article ID: IJMET_09_07_007
- [5] https://cdn2.hubspot.net/hubfs/1951043/Content%20Offers/Ai_Whitpaper/Artisan%20Talent%20AI%20and%20Bots%20in%20Recruiting%20Paper.pdf
- [6] Using Data Mining Techniques to Build a Classification Model for Predicting Employees Performance,Q.Alradaideh, Eman.Alnagi, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 3, No. 2, 2012
- [7] <https://ideal.com/top-recruiting-software/>
- [8] <https://goarya.com/solutions/>
- [9] The new age: artificial intelligence for human resource opportunities and functions, EY 2018

[https://www.ey.com/Publication/vwLUAssets/EY-the-new-age-artificial-intelligence-for-human-resource-opportunities-and-functions/\\$FILE/EY-the-new-age-artificial-intelligence-for-human-resource-opportunities-and-functions.pdf](https://www.ey.com/Publication/vwLUAssets/EY-the-new-age-artificial-intelligence-for-human-resource-opportunities-and-functions/$FILE/EY-the-new-age-artificial-intelligence-for-human-resource-opportunities-and-functions.pdf)

- [10] Role of Artificial Intelligence in Recruitment, Anjana Raviprolu, International Journal of Engineering Technology, Management and Applied Sciences, April 2017, Volume 5 Issue 4, ISSN 2349-4476
- [11] Effectiveness of Performance Appraisal System and its Effect on Employee Motivation, Idowu, Ayomikun O., Nile Journal of Business and Economics, (2017) 5: 15-39
- [12] The History of Artificial Intelligence, Chris Smith, Brian McGuire, Ting Huang, Gary Yang, University of Washington December 2006
- [13] Introduction To Machine Learning, 2nd edition , Ethem Alpaydm, The MIT Press Cambridge,2010
- [14] <https://ieeexplore.ieee.org/document/7457238> (Hirability in the Wild: Analysis of Online Conversational Video Resumes)
- [15] Should human resource managers use social media to screen job applicants? Managerial and legal issues in the USA - <https://www.emeraldinsight.com/doi/abs/10.1108/14636691211196941>
- [16] The Writing on the (Facebook) Wall: The Use of Social Networking Sites in Hiring Decisions - <https://link.springer.com/article/10.1007/s10869-011-9221-x>
- [17] Dataset taken from "Network International" company in Jordan