“eRReBIS” Business Intelligence based Intelligent Recommender System for e-Recruitment Process

Siwar Ayadi[a], Manel Bensassi[b] and Henda Ben Ghezala[c]

RIADI Lab, National School of Computer Science, Manouba University, Tunisia

Keywords: e-Recruitment, Business Intelligence, Content-based Recommendation, Similarity Measure, Prescriptive Analysis, Machine Learning Algorithms.

Abstract: Due to the continuous and growing spread of the corona virus worldwide, it is important, especially in the business era, to develop accurate data driven decision-aided system to support business decision-makers in processing, managing large amounts of information in the recruitment process. In this context, e-Recruitment Recommender systems emerged as a decision support systems and aims to help stakeholders in finding items that match their preferences. However, existing solutions do not afford the recruiter to manage the whole process from different points of view. Thus, the main goal of this paper is to build an accurate and generic data driven system based on Business intelligence architecture. The strengths of our proposal lie in the fact that it allows decision makers to (1) consider multiple and heterogeneous data sources, access and manage data in order to generate strategic reports and recommendations at all times (2) combine many similarity's measure in the recommendation process (3) apply prescriptive analysis and machine learning algorithms to offer adapted and efficient recommendations.

1 INTRODUCTION

E-recruitment platforms, as decision support system, has increasingly been used in the industry particularly with the epidemic widespread and accomplished clear advantages for both recruiters and job-seekers by reducing the recruitment time and advertisement cost. Data has been largely widespread in the job market.Consequently, these platforms suffer from an inappropriateness of traditional information retrieval and exploitation techniques (Al-Otaibi and Ykhlef, 2012). Recommendations for e-recruitment systems generally differ from recommendations generated in other contexts (e.g. movies, e-commerce), given that the job seekers'level of competencies and knowledge rather then their interests is key for suggesting the most appropriate position. In fact, the problem of matching jobs and candidates can be enhanced from two distinct perspectives: (a) find relevant candidates to a job opening; and (b) select the suitable jobs to a specific candidate.

Thus, the challenge of making recommendations and to develop an accurate decision support system is closely tied to which and how data are extracted, transformed and conveyed. Over the past few years, Jobs' recommendation became an important task for the modern recruitment process in order to improve manager experience.

On the other hand, existing recommender solutions do not afford the recruiter to manage the whole process. Recruiters need to capitalize on clear KPIs inside of Talent Management through guided and dynamic insights. Consequently, Explanation for recommendation systems through the deployment of a KPIs dashboard with the purpose of helping the recruitment's team make more precise decisions is underlined.

In order to support this research area, we describe, in this article, a Content based Recommender which relies on a Business Intelligence architecture, developed in a computer science company, part of a larger group, whose objective was to improve the management of e-recruitment process searching for more adequate means to attract qualified coworkers. Our proposal is based on prescriptive analytics that assists decision makers in identifying data-driven strategic decisions.

This article is organized as follows. We analyze the e-recruiting business requirements and re-

[a] https://orcid.org/0000-0002-6085-4986
[b] https://orcid.org/0000-0002-0224-6165
[c] https://orcid.org/0000-0002-6874-1388

Ayadi, S., Bensassi, M. and Ben Ghezala, H.
“eRReBIS” Business Intelligence based Intelligent Recommender System for e-Recruitment Process.
DOI: 10.5220/0011530200003318
In *Proceedings of the 18th International Conference on Web Information Systems and Technologies (WEBIST 2022)*, pages 373-380
ISBN: 978-989-758-613-2; ISSN: 2184-3252

view related recommender systems proposals for e-recruitment management in section 2. eRReBis Recommender System overview and architecture are presented in Section 3. Usage scenario and the interpretation of results are discussed in Section 4. The evaluation of efficiency of the proposal system are described in section 5. Finally, conclusions are drawn and future works are suggested in Section 5.

2 RECOMMENDATION SYSTEM IN E-RECRUITMENT: BUSINESS REQUIREMENT AND RELATED WORKS

In recent years, data are overwhelming, growing in size and in connections, creating a potential challenge of information retrieval in a complex environment. Filtering, extracting, and prioritizing relevant information is a pressed requirement for data 's users and analytics, especially, decision makers in the business management field. In this context, recommender systems have been developed to search into a large amount of data and select relevant information according to user's preferences, interest, or observed behavior about item (Karimi et al., 2018)

In order to assist the recruiter in looking for new talents and decision-maker in defining strategic guidelines for the e-recruitment process, recommendation system, as decision support system (Stohr and Viswanathan, 1998), afford a flexible support to different stakeholders. In literature, many recommender system have been proposed. Casagrande (Casagrande et al., 2017) proposes person-job recommendation system using the profile information from both candidates and jobs descriptions in order to find a good match between talents and jobs. Ramanath (Ramanath et al., 2018) explored machine learning of candidate potentials to determine a ranked list of the most relevant candidates from the LinkedIn thousands of candidates. Recently, Norbert (Jiechieu and Tsopze, 2021) followed the ideas of recommender systems and proposed a comprehensive job recommender system based on competency prediction and resume classification. He used classification model based on neural networks in order to predict high level skills imprecisely written.

Other works are dedicated to semantic modeling. For example, Paquette (Paquette, 2014) proposes a semantic modeling of skills which allows to model different level of granularity of skills. Zhao et al. (Zhao et al., 2015) proposed a combination of Named Entity Recognition and Named Entity Normalization to identify skills from texts, considering them as named entities. Cheng et al. collected the job-related information from various social media sources and constructed an inter-company job-hopping network to demonstrate the flow of talents. (Wang et al., 2013) predicted the job transition of employees by exploiting their career path data. Xu et al. (Xu et al., 2016) proposed a talent circle detection model based on a job transition network which can help the organizations to find the right talents and deliver career suggestions for job seekers to locate suitable jobs.

However, these systems are not enough sophisticated and presented some limits related to the management of e-recruitment process. In fact, enterprises require better management of the data and systems involved in decision-making. To do so, we conclude that decision-makers need to:

- Deploy an agile IT architecture that can integrate an increasing number of data sources required for decision-making, as well as external and big data sources.
- Increase benefit from collective intelligence in decision making process by exploring collaboration and participation in data analysis and planning processes to improve performance and to easily move from strategic to operational point of view (and reverse) to a more democratic style of decision-making.
- Define, pervasively use and visualize KPIs across the organization through dashboard to achieve a common foundation for decisions, align measures of success and focus data governance on important data.

From this point of view, presented works still insufficient and do not cover the whole business requirement to manage the e-recruitment process. Companies had to embrace flexible business structures and to change decision-making cultures from decisions based on intuition and experience to **data-driven decision-making** in order to identify new business opportunities, predict future trends and behavior.

On the other side, many online recruiting platforms suffer from an inappropriateness of Boolean search methods for matching applicants with job requirements. Consequently, a large number of candidates missed the opportunity of recruiting (Lang et al., 2011). Actual practices and theoretical thoughts show that this search type is insufficient for achieving a good fit between candidate aptitudes and job requirements (Färber et al., 2003).

Further, given the business requirements, recruiters need to capitalize on clear KPIs inside of Talent Management through guided and dynamic insights. Consequently, the deployment of a KPIs dashboard with the purpose of helping the recruitment's team make more precise decisions is underlined.

3 BI BASED INTELLIGENT RECOMMENDER SYSTEM FOR E-RECRUITMENT PROCESS MANAGEMENT DESIGN

3.1 Motivations

Our main goal is to provide a theoretical model along with a proof-of-concept implementation that:

- could be adopted and appropriately adapted to enhance the e-recruitment management from back to front office.
- assists decision makers in identifying data-driven strategic decision and help them to avoid the limitation of standard data analytics.

Thus, the design of our approach is dedicated to the need to exploit capabilities of BI paradigms towards the enhancement of data driven decision making in the e-recruitment management process. The blending of BI tools, machine learning algorithms and recommender system is considered promising for the improvement of data driven decision making in the e-recruitment field and for the development of solutions able to provide effective and adapted recommendations by combining different techniques of similarity measures.

3.2 eRReBIS: Conceptual Architecture

In the proposed approach (Fig 1), heterogeneous data are collected transformed according the following process

Step 1: Data Gathering: from multiple data sources that are structured by Company's developers. We consider:

- Byblos data source that contains Company's employee information.
- Application Tracking System (ATS) Database from which passive candidate profiles looking for a new opportunity are created.
- LinkedIn from which relevant Company's job data is extracted. To do so, we have opted for the web scraping (screen scraping) method, based on automatic data retrieval implemented in python.

Step 2: Data Integration: is ensured by the ETL (Extract, Transform, Load) data integration process, which takes care of collecting all the necessary information from the various sources.

Data Pre-processing and Cleaning: Heterogeneous data requires a homogenization and integration process. The first stage consists of eliminating the irrelevant data that represents a noise. In the second stage, we removed duplicated data, punctuation and numbers. The tokenization stage is an essential phase in order to remove stop words and irrelevant words. In this step,we also integrate an algorithm to measure the similarity between the descriptions of the offers and the profiles of the candidates. Transformed and adapted data to the feeding phase is stored in a specialized database "data Mart" which is integrated in the "Data Warehouse".

Step 3: Data Analysis: is the graphical representation of the performance indicators and recommendations in a way that is understandable and helps the decision maker to make strategic and relevant decisions. For this end, we analyse data with the Tabular Model which is based on a new in-memory engine for tables. The in-memory engine loads all table data into memory and answers questions directly from memory. This is very fast in terms of response time. Unlike the multidimensional model, the tabular model processes the data warehouse in a single processing block without dividing it into cubes.

Step 4: Data Visualization: whose objective is to make access to the decision much more immediate through graphic representations and indicators mapping.

Step 5: Prediction and Recommendations: Our "eRReBIS" system provides predictive analytics and guided recommendations to decision makers in order to provide additional insights. Given the fact that our training data is labeled and has output variables corresponding to the the input variables, supervised machine learning algorithm was the most suitable to our business requirement. In addition, our main goal is to support decision-maker tp predict the evolution of strategic KPI: the number of incoming employees, the number of resignations and hires,...Then, we need a regression algorithm to build a mathematical model from the training data. Then for this study, to choose the most suitable regression model to the extracted data, a comparative and evaluation study is elaborated between these two supervised and regression algorithms:

- **Random Forest Regression:** is a supervised learning algorithm that combines the prediction of several automatic learning algorithms to make a more accurate prediction than a single model.

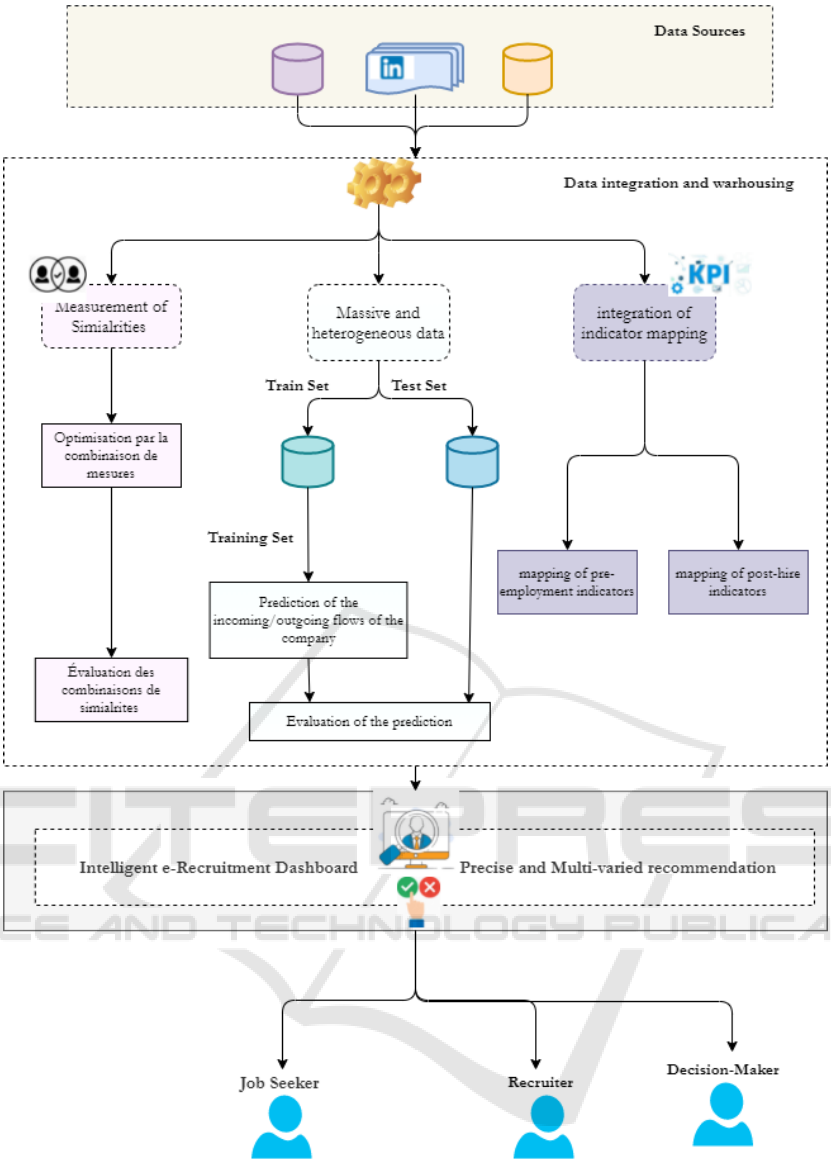

Figure 1: Conceptual Architecture.

During the learning process the algorithm builds decision trees that work in parallel.

- **Linear Regression:** is a supervised prediction algorithm, its objective is to predict from the dependent variables of the independent variables by finding a function of linear prediction y=f(x) in order to predict values which are not observed. There are two types of simple and multiple linear regression, according to the number of dependent variables. In our case we have two dependent variables (month, year) so it is the multiple linear regression

Finally, Data is represented in an e-recruitment dashboard that aims to make access to the decision much more immediate through graphic representations and indicators mapping presented in the following section.

3.3 e-Recruitment Dashboarding based on KPI Cartography

To build such a dashboard, the first step is to understand with greater detail the real pains of Talent Management and the main opportunities that could be created for the short and long term by answering these questions: *What is currently not possible to analyze? Which analyses and information would be necessary for a decision to be made comfortably? For each* *analysis, which metrics would help take the actions chosen above?* The recruitment process is divided into two phases: pre and post hire. In order to respond to manage the whole process chronology, we organise the e-recruitment KPI into:

- **Pre-hire Indicators** (calculated before the recruitment)**:** The first pre-hire indicators are specific to the application tracking system (SmartRecruters) as the recruitment process begins online. Then, pre-hire indicators are automatically calculated as soon as the recruiter receives the applications.
- **Post-hire Indicators** (calculated after the recruitment)**:** The second post-hire indicators are calculated from the company's ”Byblos” database after hiring by using the dimensions ”Personnel”, ”Contract” and their attributes.

In order to assist different stakeholders in decision making process at different levels (operational, analytical and strategic), we propose, a cartography of indicators organised as follow:

- **Operational Indicators:** which show data related to daily e-recruitment operations. The main purpose of an “operational dashboard” is to provide a comprehensive snapshot of e-recruitment performance. For example, we cite : **Number of applications** that represents the total number of applications received by the recruitment channels, **Hiring rate**, **Number of inactive** or **active employees**.
- **Analytical Indicators:** such as Expected number of hires and expected number of resignations, which use historical company data to identify trends that can influence future decision-making. The ideal audience for viewing analytical dashboards are database analysts, as they typically require a level of understanding that a typical business user may not possess.
- **Strategic Indicators:** which offers a decision-maker to track performance in relation to strategic key performance indicators to better align actions with strategy such as cooptation indicator which measures *”participative recruitment”* (number of applications from internal source). In fact, Cooptation is a recruitment method that enhances collaboration and employee participation in the recruitment process by recommending talent for a specific position.

3.4 e-Recruitment Recommender Functionalities

Figure 2 shows the use case breakdown of the eRReBIS System. The roles involved in the process include both human actors and machine system components (shown, respectively, as man icons and gray rectangles in Figure 2). The recruiter has the possibility to configure the dashboard, to select the axis of analysis and to explore operational and tactical KPI through descriptive analysis.The process of descriptive analysis is split into job seeker reporting (visualising offers list) and recruiter reporting (visualizing candidate profile). Recruiter reporting is the process in which similarity computed and all aspects related to job position and each candidate profile are visualized . The recruiter deals with lists of offers, possibly organizing them hierarchically in sub groups through recommendation engines. Then, he could visualize the most appropriate profile to a job position and conversely.

The role of the decision-maker is to deal with the strategic dashboard through predictive analysis when necessary to define the recruitment requirements and strategies.

eRReBIS System is also based on the “Recommendation engine” that is an autonomous module in charge of executing the recommendation algorithm and computing the similarity measure.It is based on predictive analysis which is a process during execution and it always follows descriptive analysis.

Recommendation engine ensures prescriptive analysis which is the process by which the job seeker, recruiter and decision-maker finally makes decisions about the most suitable position, the most suitable profile or about recruitment strategies. Based on the output of predictive analysis and the computation of the similarity's measure, recommendations are proposed to the job seeker, recruiter and decision-maker by eRReBIS solution. To do so, we propose a content based recommendation system and algorithm which:

- Creates vectors from jobs' databases and profiles.
- Combines semantic and syntactic similarity measures to reconcile documents represented in vector form.

A comparative study has been conducted, to show the efficiency of our measure compared with classical measures, is presented in section 5.0.2.

eRReBIS system was developed by combining several tools and languages. In fact, we visualize integrated data using the Power BI tool. Recommender Engine, coded with python language, was injected into the integration process deployed with Talend Data Integration (see figure 2).

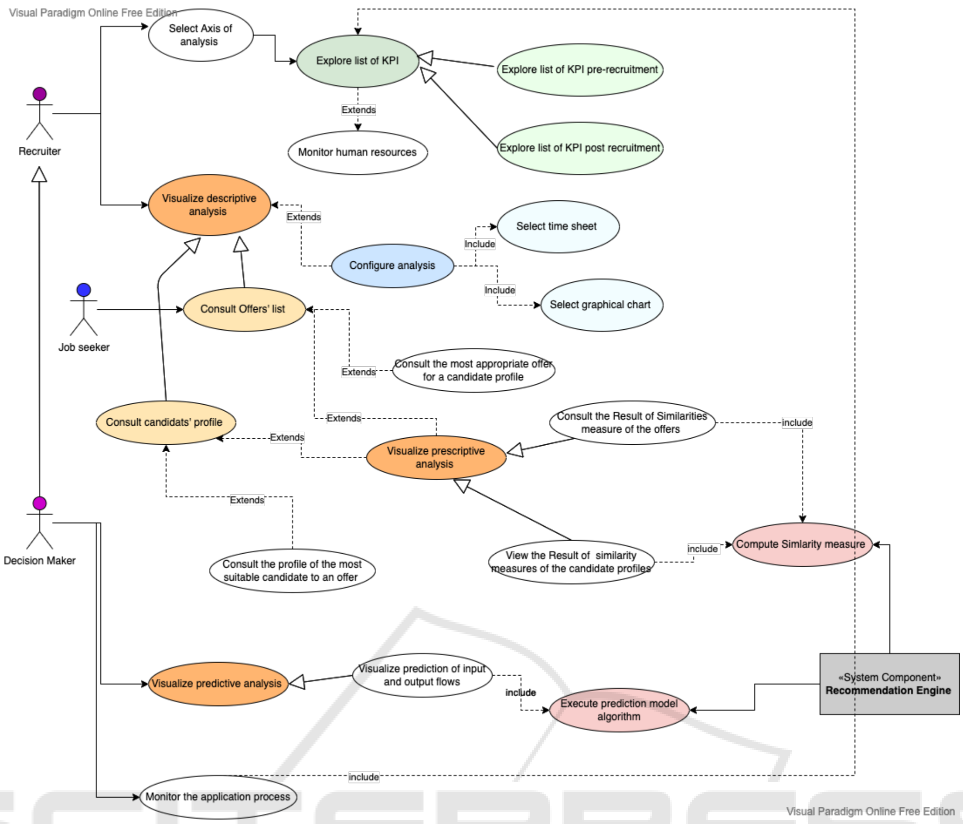

Figure 2: eRReBis Functionnalities.

4 CASE STUDY AND RESULT INTERPRETATION

Our prototype provides several views supporting filtering in the result set and integrates:

- KPI cartography for strategic, analytical and operational decisions based on straightforward charts to interpret.
- Prescriptive analysis based on predictive machine learning algorithm and adapted recommendations

Building upon these features, we divide our dashboard into 3 views to support the recommendation process:

- **Strategic View:** (figures 3 and 4) Main KPIs of Talent Management, not with the objective of a micro analysis but a general view of how the metrics are doing, being able to identify any potential issues. KPIs such as Forecast, hired, candidate to hire, and others. This can all be filtered by job or area.

- **Analytical View:** also known as tactical view, shows a more specific view for the leader of the area (recruiter). Has some analyzes which are more broken up by date for example.

Figure 3: Analysis of applications.

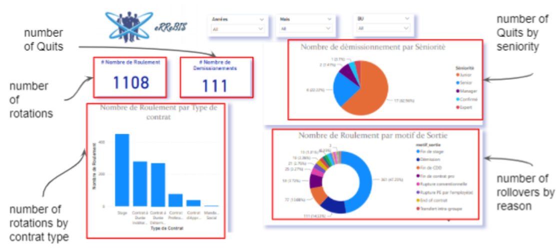

Figure 4: Turnover analysis.

A snapshot of the dashboard illustrates generic results on the offers and profiles of the candidates, such as the total number of offers, number of offers by location, as well as the total number of profiles of the candidates and the profiles by seniority

- **Operational View:** This view is designed to be used by recruiters and offer a very micro view of how each job is doing, the pipeline, forecast etc. This is all filtered by the recruiter.
- **Multi-varied Recommendations on Different Activities and Steps** (figures 5, and 6)**:** The recruiter has the possibility to choose an offer from the list extracted from the linked In. He can visualize a table that contains the profiles with the similarity measure. The most suitable profile with the highest similarity measure to his offer is shown in figure 5. The selected profile could be more adequate to another position proposed by the company. **eRReBIS** prototype offers the possibility to the recruiter to evaluate the adequation between this profile and all offered positions. He could select, as shown in the figure, the same profile resulting from the first search and visualize the similarity measure with all the offers of the company. In this way, he can visualize the most adequate offer for this profile. Comparing scores, the last one is more appropriate.

Figure 5: Analysis of the offer recommendations.

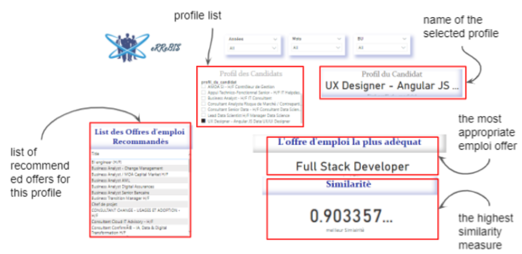

Figure 6: Analysis of Profile Recommendations.

5 EFFICIENCY EVALUATION OF eRReBIS

5.0.1 Prediction Algorithm Performance

For the evaluation process of the prediction algorithm, we used two measures : **The term R score** 1 and **The Root mean square error (RMSE)** 2

$$R^2 = \frac{\sum(Y_i - y_i)^2}{\sum(Y_i - \overline{y}_i)^2} \quad (1) \qquad RMSE = \sqrt{\frac{\sum(Y_i - y_i)^2}{n}} \quad (2)$$

The result of the comparative study is shown in the figure 7 and 8. The algorithm "Random Forest regression algorithm" that has the lowest value of RMSE and the highest value of R score was adopted.

Figure 7: R2 Results.

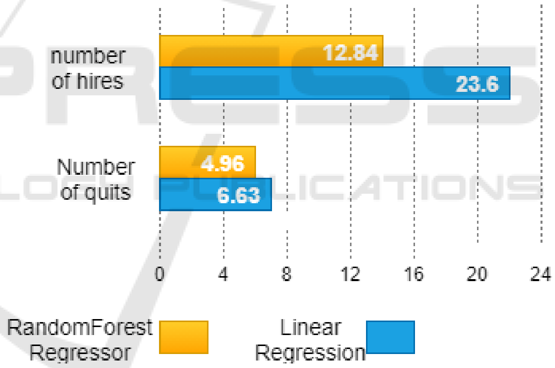

Figure 8: RMSE Results.

5.0.2 Evaluation of Similarity Measures Combination

The job seeker data is taken from Babylos and ATS database. The data set contains 24734 candidates plus 7270 active collaborators . In total we have 32004profiles, where each profile contains weighted competencies and 70 job positions. For testing purposes, we divide the user ratings dataset into training (80%) and testing sets (20%). We perform a 4-fold cross validation and results are averaged over the 4 cycles of execution. We conducted a set of experiments using four different scenarios. Scenario 1 uses the Content based Filtering method only. We conducted three sub experiments with different syntactic measures(Pearson, Cosine and Jaccard).Scenario

2 uses the semantic similarity method only, Scenario 3 uses a hybrid of syntactic and semantic similarities methods with equal contribution (both have a weight of 0.5).We present the evaluation of the proposed similarity measure compared with the other measure. Similarity measures are computed as shown in 1. When we combine cosine and semantic similarity, we obtain optimal results and it is more accurate than a single measurement. We evaluate recommendation accuracy by using Precision as shown in table 2.

Table 1: Similarity measurement results.

Profile ID	Job ID	Syntactic Similarity			Semantic Similarity	Hybrid Similarity
		Pearson	Cosine	Jaccard	Spacy NLP	Cosine and Spacy NLP
1	1	0.34	0.44	0.41	0.51	0.47
2	2	0.45	0.38	0.21	0.32	0.35
3	3	0.74	0.92	0.67	0.99	0.9
4	4	0.56	0.58	0.45	0.62	0.6
5	5	0.33	0.41	0.12	0.38	0.35

Table 2: Evaluation recommendation.

	Syntactic similarity			Semantic similarity	Hybrid similarity
	Jaccard	Pearson	Cosine	Spacy NLP	Cosine & Spacy NLP
Total of recommendations	180	166	208	195	248
Pertinent recommendations	77	68	92	84	139
Precision Value	0.42	0.4	0.44	0.43	0.56

6 CONCLUSION AND PERSPECTIVES

“eRReBIS” is data-driven strategic decision system and differs from the existing approach in that it relies mainly on data management based BI architecture. We propose Dynamic Multiviews dashboarding (strategic, tactical, operational) based on prescriptive analysis in order to assist the recruiter to identify trends, to create new business opportunities. Our prototype is a content-based recommendation engines that deploy a combination of similarity measures in order to classify the candidates. We show that by combining the similarity measures between the jobs and skills, our model provides better recommendation for both recruiter and candidate. Additionally, we also show some case studies which validate our claims. Our system is scalable and can be gradually enhanced by other massive data sources from social networks and Business platforms. Furthermore, the system architecture can be improved by other intelligent features and components of deep learning allowing the inference of strategic recommendations. Future research will focus more on deep and prediction analysis with machine learning techniques to improve the performance of recommenders and linked data profiling in the knowledge representation.

REFERENCES

Al-Otaibi, S. T. and Ykhlef, M. (2012). A survey of job recommender systems. *International Journal of Physical Sciences*, 7(29):5127–5142.

Casagrande, A., Gotti, F., and Lapalme, G. (2017). Cerebra, un système de recommandation de candidats pour l’e-recrutement. In *AISR2017*.

Färber, F., Weitzel, T., and Keim, T. (2003). An automated recommendation approach to selection in personnel recruitment.

Jiechieu, K. F. F. and Tsopze, N. (2021). Skills prediction based on multi-label resume classification using CNN with model predictions explanation. *Neural Computing and Applications*, 33(10):5069–5087. Publisher: Springer.

Karimi, M., Jannach, D., and Jugovac, M. (2018). News recommender systems–Survey and roads ahead. *Information Processing & Management*, 54(6):1203–1227. Publisher: Elsevier.

Lang, S., Laumer, S., Maier, C., and Eckhardt, A. (2011). Drivers, challenges and consequences of e-recruiting: a literature review. In *Proceedings of the 49th SIGMIS annual conference on Computer personnel research*, pages 26–35.

Paquette, G. (2014). Technology-based instructional design: Evolution and major trends. *Handbook of research on educational communications and technology*, pages 661–671. Publisher: Springer.

Ramanath, R., Inan, H., Polatkan, G., Hu, B., Guo, Q., Ozcaglar, C., Wu, X., Kenthapadi, K., and Geyik, S. C. (2018). Towards deep and representation learning for talent search at LinkedIn. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pages 2253–2261.

Stohr, E. A. and Viswanathan, S. (1998). Recommendation systems: Decision support for the information economy.

Wang, J., Zhang, Y., Posse, C., and Bhasin, A. (2013). Is it time for a career switch? In *Proceedings of the 22nd international conference on World Wide Web*, pages 1377–1388.

Xu, H., Yu, Z., Yang, J., Xiong, H., and Zhu, H. (2016). Talent circle detection in job transition networks. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 655–664.

Zhao, M., Javed, F., Jacob, F., and McNair, M. (2015). Skill: A system for skill identification and normalization. In *Twenty-Seventh IAAI Conference*.