International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

A Data-Driven Profile-based Analytics for Career Path and Upskilling Recommendations in HRS

Paul Nigel S. Abalos[1, a], Thelma D. Palaoag[1, b]

Submitted: 07/02/2024 **Revised**: 15/03/2024 **Accepted**: 21/03/2024

Abstract. This quantitative-experimental research aimed to develop a profile-based data analytics module for career path and upskilling recommendations module for the Data-Driven Human Resource Analytics System (DDHRAS) of the largest state university in Ilocos Region and address common analytics results heavily based on skills which do not fully include other useful data in some sectors, such as the academe in which the researchers belong to. Hence, the study is motivated to come up with an analytics module that is close to the university HRS by incorporating significant employee personal data, educational background, seminars and trainings, skills and competencies, board examinations, job performance score, work experience, and tenure. A scoring matrix based on human resource policies was devised and paired with Word2vec algorithm to evaluate which profile composition criteria each employee has an advantage of or needs to improve to qualify for promotions. Dataset covers 628 personnel data sheets of employees for School Years 2020 to 2023 from different campuses of the subject institution. The sequential neural network to predict career path and upskilling recommendations was utilized and resulting models were evaluated using model classification performance metrics. Experimental results show that the models developed have consistently obtained an average accuracy of 88.99% during performance validation, indicative of rigid classification performance. The implication of the developed models, when fully integrated and implemented within the existing system of the HRD, can provide an alternative faster career path and upskilling recommendations to lessen the burden of HRD evaluators, thus, save valuable resources.

Keywords: profile-based analytics, upskilling recommendations, career path prediction, sequential neural network

1. Introduction

The human resource department (HRD) has been one of the organizational departments that are constantly bombarded by voluminous data coming from numerous applicants and periodic updating of personnel data sheet (PDS) or 201 Files of existing employees. Hence, the need for a reliable human resource system, both manual and digital systems, is a necessity among organizations to deal with the storage and retrieval of personnel records and be easily available with high integrity when needed.

Usually, it is the duty of any HRD to document the profiling of each employee from the moment they are hired up to the most recent performance review as the core roles of the HRD, regardless of employee population size and industry sector, do not stop after successfully hiring individuals and place them in their respective posts to become a part of the organization, and in general, be in unison in achieving the day-to-day objectives and functions of the business (Berra & Sanna, 2022). Ideally, according to Brewster & Larsen (2021), one of the roles of the HRD is to provide an upward career path for each employee and ensure a significant list of upskilling recommendations to empower everyone and prepare them for the next, higher-level position.

However, as Lawler & Mohrman (2019) emphasize, for organizations belonging to medium to large-scale companies, determining and planning in toto with individual employees their respective next ladder of position movement is both tedious and complex task that requires sufficient time, energy, and impartiality during performance results evaluation. Hence, Schuler & Jackson (2021) underscored the growing importance of HR analytics in informing strategic decision-making within organizations and emphasized the need for HR professionals to develop data-driven upskilling and leverage analytics effectively to optimize workforce management particularly for promotions or career path planning that basically involves upskilling.

Career path planning and suggested upskilling measures have been an integral HRD activities in relation to performance assessment and outcomes-based evaluation of employees which determine employee retention, demotion, or promotion (Tzafrir, Baruch & De Beer, 2020). These manual tasks have been the focus of computerization and further integration with artificial intelligence (AI) and machine learning (ML) approaches in recent years (Schuler & Jackson, 2021).

University of the Cordilleras Baguio City, Philippines
[a] *Main author: paulnigelabalos@gmail.com*
[b] *tdpalaoag@uc-bcf.edu.ph*

In the works of Karakurum, Gupta, & Raj (2022), the authors investigated how ML algorithms can be used to personalize upskilling recommendations for employees based on their skill gaps, career goals, and organizational needs and explored the potential benefits and challenges associated with implementing such systems. This endeavor is similar to the works of Yang, Chen, & Zhou (2023) who envisioned an AI-powered HR system that utilizes ML algorithms to analyze employee data and provide personalized career development and upskilling recommendations for the potential benefits of both employees and organizations. In relation to skills gap, the works of Karakurum, et al. (2022) parallel to the research of Romero, Lopez & Arnau (2021) which explored the application of ML techniques for identifying skill gaps and recommending personalized learning paths in a learning and development context and offered valuable insights into the potential of ML for personalized skill development. The herein researchers took note of the ML algorithms used in the cited studies and explored the possibilities of incorporating personalized learning paths by way of extending the approach into generalized career paths for government workers in the academe.

Meanwhile, based on the study of Fan, Lu & Zhao (2020), the proponents suggested an ML-based framework for predicting potential career paths and recommending upskilling opportunities for employees in the IT industry and emphasized the importance of utilizing various data sources, such as employee skills, job descriptions, and internal mobility data, to improve the accuracy of recommendations. On the other hand, the research of Zhang, Chen, & Zhao (2023) emphasized the importance of explainability in ML-based recommendation systems for upskilling and career development and elaborated strategies to improve transparency and user trust in such systems, ensuring users understand the rationale behind recommended learning and career paths. The herein researcher-developers noted the addition of various data sources as well as the participative aspect of employees in putting honest and transparent profile data to improve recommendation accuracy.

As can be observed among cited related works, most recommender systems for career path planning and upskilling suggestions focus on the current employee's available skills, skills, gap, organization needs, and job descriptions, hence, such systems can be categorized as skills-based, and needs-based career path and upskilling recommendation system in contrast to profile-based system which relies not only on skills but as well as with various personal employment data sources that can optimize recommendation accuracy particularly in structured or ladderized promotions or employee movement (Xu, Zhao & Li, 2020; Singh & Singh, 2021; and Liu, Shi, & Zhao, 2022).

Meanwhile, doing an online search among existing HRS that usually provide career path suggestions and upskilling recommendations, it can be observed that most proprietary and open-source software are commonly based on employee skills and are patterned to common HRD records in the business sector. Hence, it is not easy to find a comprehensive and generalized HRS with career path and upskilling recommendations for academic institutions considering that each school has its own HR policy and employee plan in terms of promotions and in-campus career growth.

Considering the aforementioned gaps and development opportunities expounded, this study aims to develop the data analytics career path suggestions and upskilling recommendations module of the Data-Driven Human Resource Analytics System (DDHRAS) for the largest state university in the Ilocandia Region of the Philippines. It intends to incorporate significant employee personal data, educational background, relevant seminars and trainings attained, skills and competencies earned, professional board examinations passed, recent job performance score, and work experience and tenure as part of the overall employee profile that will be utilized as an input to machine learning neural network models that will provide career path suggestions and upskilling recommendations.

The output models of this research are regarded to be highly significant to the full completion of the DDHRAS simultaneously being developed by the researchers and consecutively, as an overall impact, aide the subject institution in a faster and less tedious career path planning and upskilling recommendations for both the academic and non-academic personnel.

2. Methods

2.1 Research Design and Model Development

For this study, the researchers employed the Quantitative-Experimental method of research and took into account the recommendations of authors Abu-El-Haija & Al-Khateeb (2022) that software development that integrates AI belong to quantitative method of research and the de facto research approach to be used is experimental research design.

Figure 1 exhibits the data analytics processing of this research which is split into 5 stages: data preparation, skills and career criteria enumeration, skills and career criteria scoring, modeling and prediction, and model evaluation and modelling, skills extraction, matching and scoring, job recommendation, and model evaluation.

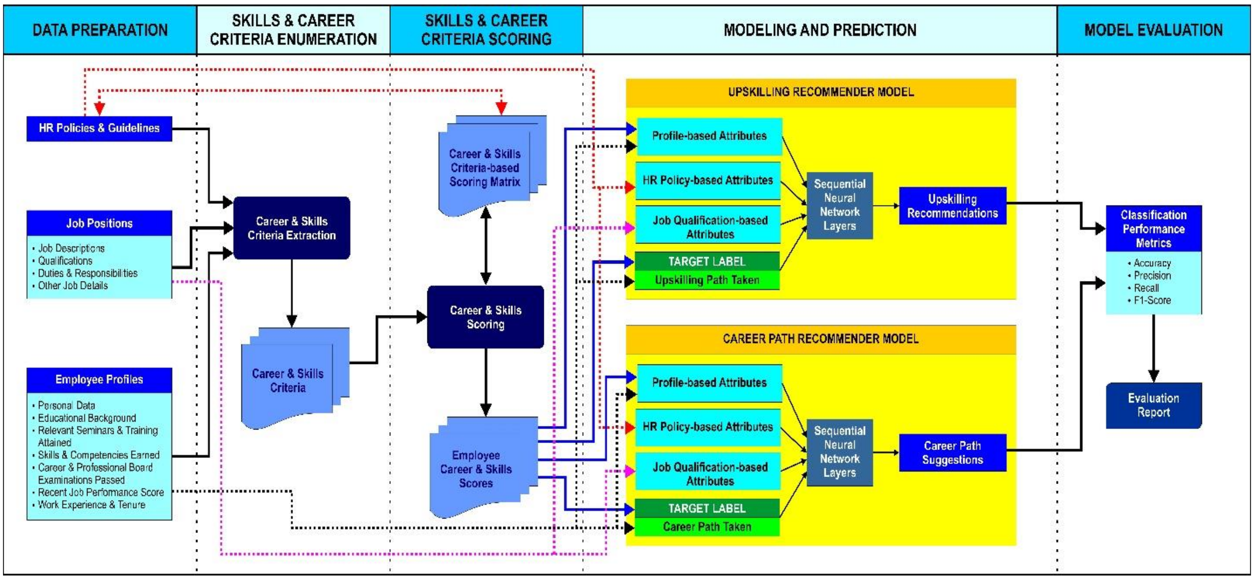


Figure 1. Data Analytics Processing

Stage 1 relates to data preparation in which collation of respective the respective human resource policies and guidelines, job positions, and personnel profile data are collated and prepared as inputs of the model datasets and criteria-based scoring mechanism devised by the researchers. Each Job Position is composed of job description, job qualifications, job duties and responsibilities, and other pertinent job details. Meanwhile, each Employee Profile is consisted of personnel information about personal data, educational background, relevant seminars and trainings attained, skills and competencies earned, career and professional board examinations passed, recent job performance score, and work experience and tenure data. Criteria are then identified and listed for each unique career and skill taken from each input data during Stage 2 of study.

Once all criteria for skills and careers are accounted for, the skills and career criteria scoring commences during Stage 3. The scoring is based on the criteria-based scoring matrix computed based on human resource policies and guidelines. After each employee's skills and career criteria-based scores are computed, Stage 4 or modeling and prediction can now start in which two ML-based recommender models are developed and utilized for career path and upskilling.

The models will provide a list of career path or upskilling recommendations and will be ranked according to classification accuracy. These recommendations are evaluated in Stage 5 using accuracy, precision, recall, and F1-score and are saved to be part of the evaluation report.

The researchers developed a front-end user interface (UI) where participating employees of the subject institution can signup, encode and update their personnel profile which includes personal data, educational background, relevant seminars and trainings attained, skills and competencies earned, professional board examinations passed, recent job performance score, and work experience and tenure. Upon encoding or updating personnel profile, the system would automatically compute or recompute each corresponding profile component criteria score of the employee.

For the computation of each employee's profile component criteria score, the researchers used the following formula:

$$Profile\ Component\ \ Criteria\ Score = \frac{Obtained\ Criteria\ Score}{Max\ Allowed\ Criteria\ Score} \quad (1)$$

With the above formula, the system would be able to identify which profile component criteria an employee has lower score, hence, can be used as initial basis to suggest relevant upskilling and career path recommendations.

The researchers made use of HTML, PHP, and Javascript programming languages and Bootstrap template for the development of the UI and user account records and their corresponding employee profile were stored in tables using the MySQL database. The personnel profile constituted the profile-based attributes of the recommender models while, career and skills per HR policies and guidelines for each job position were encoded also in the system and represented the HR policy-based attributes and job qualifications attributes of the recommender models.

2.2 Model Performance Evaluation

The researchers utilized the Model Classification Performance Metrics to evaluate the prediction performance of the models in recommending upskilling and career path suggestions. The Model Classification Performance Metrics include the four most common derivatives of confusion matrix consisted of accuracy, precision, recall and F1-score metrics.

Accuracy metric pertains to the ratio of the total number of correct predictions and is mathematically defined as:

$$Accuracy = \frac{TP+T}{n} \quad (2)$$

where TP refers to True Positives, TN represents True Negatives and n is the total number predictions, in this case, the total elements of the recommendation list. Meanwhile, precision, which is also referred to as the positive predictive value, is the number of correct predictions among the predicted successes and is characterized to be of high value if false positives are low. Precision is mathematically defined as:

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

Recall, which is also referred to as sensitivity or the true-positive rate, is the fraction of portion of correct predictions among the true successes and is also characterized to have a high value if there are few false negatives. It can be written using the following formula:

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

Finally, F1-score is the harmonic average or mean of Precision and Recall and has a direct similar effect Precision and Recall, that is, it has a higher value if Precision or recall is high and F1-score tend to be low if either Precision or Recall is low. It can be mathematically written as:

$$F1-score = 2\frac{Precision*Recall}{Precision+Recall} \quad (5)$$

3. Results and Discussion

Table 1 shows a sample criteria-based scoring for a specific job title under educational background devised by the researchers to be utilized as part of the Stage 3 of the data analytics processing.

Table 1. Sample Criteria-based Scoring

Job Title/Position/Rank:	I.T. Instructor		
Educational Background Criteria	**Sub-Rank & Max Score**		
	I	**II**	**III**
1. Graduate of B.S. Computer Science, Information Technology, or Computer Engineering	2	2	2
2. Has an unexpired TESDA related National Certification such as NCIV Programming, NCII Computer Hardware Servicing, NCII Visual Graphics Design, NCII Computer Systems Servicing, NCII 2D Animation, NCII 3D Animation, etc.	2	2	2
3. Has at least 21 Master’s units under Master in Information Technology (MIT) or Master of Science in Computer Science (MSCS)	3	3	3
4. Graduate of MIT or MSCS	5	5	5
5. Has at least 21 Doctorate Units under Doctorate in Information Technology (DIT) or Doctor of Philosophy in Computer Science (PhDCS)		5	7
6. Graduate of DIT or PhDCS		7	10
Total	**12**	**24**	**32**

Once each employee’s skills and career criteria-based scores are computed, they were fed as inputs to the two ML-based recommender models developed during Stage 4, namely, the Upskilling and Career Path Recommender Models which were consisted of: profile-based attributes which come from employee profile data; HR policy-based attributes which can be extracted from the human resource policy and guidelines data; and job qualifications-based attributes which originate from job positions data. The target class label for the Upskilling Recommender Model is the historical data of upskills that employees have taken to be promoted or get permanent job status while the target class label for the Career Path Recommender Model relates to the historical data of career path taken by employees in order to get promoted or get permanent job status.

During prediction, the models provide a list of career path or upskilling recommendations ranked according to classification accuracy. These predictions and/or recommendations are evaluated during Stage 5 using model classification performance metrics which are composed of the 4 most common derivatives of confusion matrix which include accuracy, precision, recall, and F1-score. The training and validation prediction performances of each model are stored as part of the evaluation report.

The raw data for this research came from the Human Resource Department of Pangasinan State University in which one of the researchers is currently employed. A formal letter was sent to the HR Director to get approval to collect personnel data sheets of both academic and non-academic employees from School Years 2020-2021 to 2022-2023 among the nine campuses the HRD is managing.

Table 2 presents the raw dataset sources from which the researchers obtained significant portion of the total number of hired academic and non-academic employees of the subject institution among all campuses from School Years 2020-2021 to 2022-2023 that fit the data requirements of the study.

Table 2. Employee Profile Raw Dataset Sources

#	Campus	SY 2020-2021			SY 2021-2022			SY 2022-2023			TOTAL		
		Acad	**Non**	**Total**	**Acad**	**Non**	**Total**	**Acad**	**Non**	**Total**	**Acad**	**Non**	**Total**
1	Lingayen-Main	23	12	35	25	15	40	26	14	40	**74**	**41**	**115**
2	Alaminos	8	5	13	12	7	19	15	10	25	**35**	**22**	**57**
3	Asingan	7	5	12	11	6	17	15	12	27	**33**	**23**	**56**
4	Bayambang	10	6	16	15	8	23	18	10	28	**43**	**24**	**67**
5	Binmaley	6	5	11	8	7	15	10	10	20	**24**	**22**	**46**
6	Infanta	7	5	12	10	8	18	12	10	22	**29**	**23**	**52**
7	San Carlos	11	7	18	12	9	21	18	12	30	**41**	**28**	**69**
8	Sta. Maria	10	7	17	11	9	21	17	12	29	**38**	**28**	**66**
9	Urdaneta	15	10	25	21	14	35	25	15	40	**61**	**39**	**100**
TOTAL													

Legend: **Acad**: Academic Employees; **Non**: Non-Academic Employees;

Considering that not all campuses of the subject institution offer the same courses belonging to the same industry sector, the researcher grouped the respective colleges and courses of each campus into six course sectoral categories that are common among all campuses in order to have consistency in terms of career path and upskilling industry sector. The six course sectoral categories include: 1) Agriculture, Fisheries and Livelihood Technologies (AFLT); 2) Architecture, Computing and Engineering (ACE); 3) Arts and Other Sciences (AOS); 4) Business and Public Administration (BPA); 5) Hospitality Management (HM); and 6) Teacher Education (TE).

After categorizing each college and course into the six sectoral categories, the researchers organized and grouped the total number of academic and non-academic employees per campus and per category. In order for the model to provide better accuracy and precision, personnel data sheets (PDS) of employees not belonging to the six sectoral categories were not included to be part of the raw dataset.

Table 3 lists the total personnel data sheets per sectoral categories collected by the researchers among all the campuses of the subject institution for academic employees from School Years 2020-2021 to 2022-2023 totaling to 628.

Table 3. Employee Profile Datasets Obtained per Sectoral Category

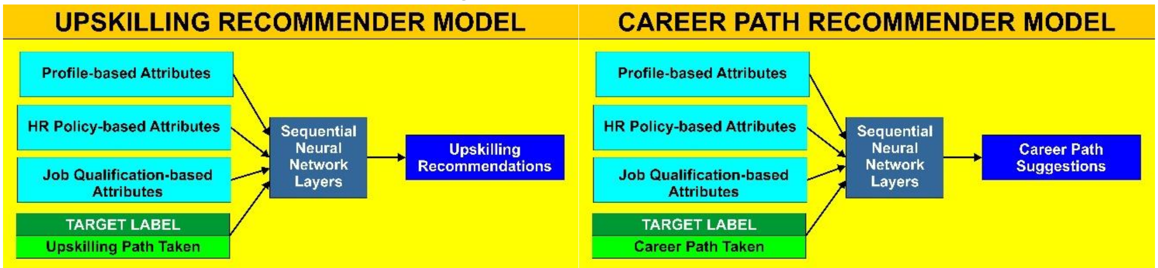
#	CAMPUS	AFLT	ACE	AOS	BPA	HM	TE	TOTAL
1	Lingayen-Main	18	19	17	19	12	30	**115**
2	Alaminos	8	9	11	12	7	10	**57**
3	Asingan	9	9	8	7	6	17	**56**
4	Bayambang	11	12	12	10	8	14	**67**
5	Binmaley	6	7	8	5	7	13	**46**
6	Infanta	9	1	11	8	7	16	**52**
7	San Carlos	12	14	11	7	10	15	**69**
8	Sta. Maria	9	8	10	12	11	16	**66**
9	Urdaneta	15	17	13	13	14	28	**100**
TOTAL		**97**	**96**	**101**	**93**	**82**	**159**	**628**

Legend: **AFLT**: Agriculture, Fisheries and Livelihood Technologies; **ACE**: Architecture, Computing and Engineering; **AOS:** Arts and Other Sciences; **BPA:** Business and Public Administration; **HM:** Hospitality Management; **TE:** Teacher Education.

In developing the upskilling and career path recommender machine learning models, the researchers utilized the Python programming language and followed the supervised learning model development concept which is composed of data cleansing, feature or attribute engineering, training-testing data splitting based on 70:30 ratio, model training, model testing, model hyperparameter adjustment, and model visualization for interpretation.

During the data cleansing, dataset rows with null feature values were discarded to avoid invalid values and prevent confusing the model with incomplete data during model training. Considering that machine learning algorithms need only numeric data types, feature engineering was employed in which both Boolean and categorical attributes were referenced using record index number from their corresponding table dataset. Meanwhile, numeric attribute scaling using MinMaxScaler function was performed to safeguard those attributes with numeric values, other than those Boolean and categorical attributes, fall between the same minimum and maximum value range to maintain equilibrium. For each PDS with multiple entries of skills, educational achievements, seminars, trainings, and job experiences, multiple rows of dataset were also allotted to accommodate several entries for each skill, educational achievement, seminar, training, and job experience. Hence, an employee's physical record corresponded to multiple rows of dataset records. At the end of the data cleansing and feature engineering steps, a final dataset composed of 41 columns for features, 2 columns for each of the models' target label, and 6,282 rows of dataset records were achieved.

Figure 2 shows the sequential neural network (SNN) diagram of the Upskilling and Career Path Recommender Models, respectively.



(a) (b)

Figure 2. SNN Diagram of Upskilling and Career Path Recommender Models

Meanwhile, Table 4 lists the final attributes of each model composed of the 30 profile-based attributes, 5 job qualifications-based attributes, 6 HR policy-based attributes, and 2 target labels, 1 for Upskilling Recommender Model and 1 for Career Path Recommender Model.

#	Attribute	Attribute Type	Data Type	Description
Profile-based Attributes				
1	Age	Numeric	Integer	Employee's age at the time of research
2	Gender	Boolean	Integer	Employee's gender (either male or female)
3	CivilStatus	Categorical	Integer	Civil Status of employee
4	Height	Numeric	Float	Employee's height in centimeter
5	Weight	Numeric	Float	Employee's weight in kilograms
6	BloodType	Categorical	Integer	Employee's blood type
7	Children	Numeric	Integer	Employee's number of children
8	SpouseOccupation	Categorical	Integer	Employee's spouse occupation
9	ElemHonor	Ordinal	Integer	Elementary honor achieved by the employee
10	HSHonor	Ordinal	Integer	High school honor achieved by the employee
11	VocationalCourse	Categorical	Integer	Vocational course of the employee
12	CollegeCourse	Categorical	Integer	College course graduated by the employee
13	MasteralCourse	Categorical	Integer	Masteral course graduated by the employee
14	MasteralHonor	Ordinal	Integer	Masteral honor achieved by the employee
15	DoctoralCourse	Categorical	Integer	Doctorate course graduated by the employee
16	DoctoralHonor	Ordinal	Integer	Doctoral honor achieved by the employee
17	ElibilityType	Categorical	Integer	Type of employee's eligibility

18	ElibilityExamRating	Numeric	Float	Eligibility examination rating obtained by employee
19	JobPosition	Categorical	Integer	Array of job positions held by the employee
20	JobTenure	Categorical	Integer	Array of corresponding job tenures in months
21	Department	Categorical	Integer	Array of corresponding departments where employee worked
22	Salary	Numeric	Float	Array of corresponding monthly salary per job held
23	SalaryGrade	Ordinal	Integer	Array of corresponding salary grade per job held
24	AppointmentStatus	Categorical	Integer	Array of corresponding appointment status per job held
25	Training	Categorical	Integer	Array of job trainings incurred
26	TrainingType	Categorical	Integer	Array of training type per training incurred
27	TrainingHours	Numeric	Float	Array of training hours per training incurred
28	Skill	Categorical	Integer	Array of special skills relevant to work
29	SkillType	Categorical	Integer	Array of skill type per relevant work skill
30	ProfileScore	Numeric	Float	Total employee's Component Profile Criteria Score
Job Qualifications-based Attributes				
31	JobTitle	Categorical	Integer	Job position title available for hiring
32	DescriptionW2VS	Numeric	Float	Word2Vector score of job description
33	RequirementsW2VS	Numeric	Float	Word2Vector score of job requirements
34	ResponsibilitiesW2VS	Numeric	Float	Word2Vector score of job duties and responsibilities
35	JobW2VS	Numeric	Float	Total Job Qualification Word2Vector Score
HR Policy-based Attributes				
36	JobPositionTitle	Categorical	Integer	Job position title based on HR policies and manuals
37	Rank	Ordinal	Integer	Job position Rank
38	CriteriaCategory	Categorical	Integer	Category of HR Policy Criteria
39	Criteria	Categorical	Integer	Actual HR Policy criteria
40	MaxScore	Numeric	Float	Maximum score that can be obtained per criteria per rank
41	CriteriaW2VS	Numeric	Float	Total Policy-based Criteria Word2Vec Score
Target Labels				
42	UpskillingRecom	Categorical	Integer	Target labels for Upskilling Recommender Model
43	CareerPathRecom	Categorical	Integer	Target labels for Career Path Recommender Model

Legend: **W2VS:** Word to Vector Score.

The 70:30 dataset training-testing split ratio was used by the researchers to allocate 70% or 4,397 rows of the dataset for training and 30% or 1,885 rows for evaluation. The researchers also made use of stratified splitting to warrant equal frequency distribution of rows for each target label both for training and testing datasets. During machine learning model training, TensorFlow and Keras libraries were imported to the Python program and were used to implement SNN or sequential neural network classification algorithm. Both input layers of Upskilling and Career Path Recommender Models were set to have 41 input shape which correspond to 41 model attributes. Likewise, both models were programmed to incorporate an initial 16 units of dense layer, each unit corresponding to 1 neuron and were later on increased to adjust for better accuracy results. The researchers also chose Rectified Linear Unit (ReLU) activation function as neuron activators which automatically compute the significant weight value of each attribute in connection with weight values of the rest of all the attributes. In contrast, the output layer of the SNN-based architecture of the models were comprised of 36 nodes for upskilling model and 17 nodes for career path model, each node matches to a particular upskilling recommendation label value or career path suggestion label value considering that the general objective function of the subject models are multi-label classification problems in nature. The researchers also setup the output layer of each model with the sigmoid activation function to allow the SNN to initiate non-linearity into the model and learn more complex decisions based on multiple label classification.

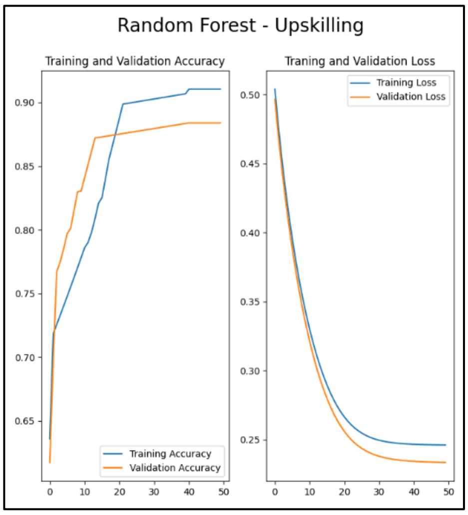
For the models' hyperparameter adjustment, the researchers utilized Adam optimizer and was set up to have an initial learning rate of 0.001 and an epoch size of 10. The Adam optimizer and the epoch size were adjusted to 0.0001 and 50, respectively, until the 85% target evaluation accuracy rate is obtained without overfitting and achieve optimum model multi-label classification performance.

Figure 3 displays the results of the Upskilling Recommender Model during machine learning model training and evaluation using sequential neural network (SNN) through the Random Forest algorithm.

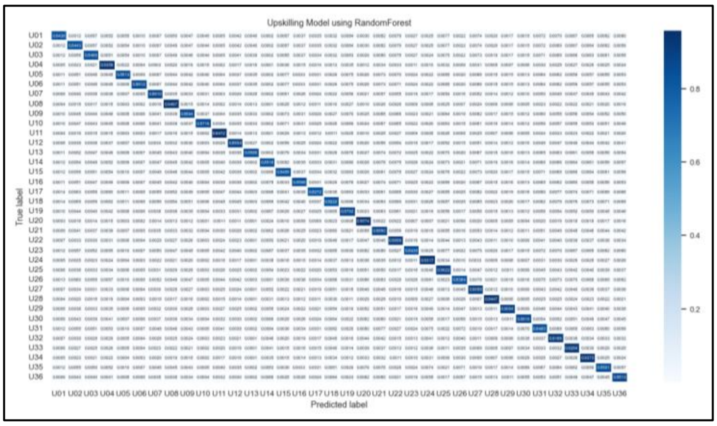
Upskilling Model Evaluation Performance Metrics:

	precision	recall	f1-score	support
0	0.8436	0.9639	0.8997	53
1	0.8443	0.8558	0.8500	53
2	0.8466	0.8623	0.8544	53
3	0.9359	0.8677	0.9005	53
4	0.8613	0.9548	0.9056	53
5	0.8612	0.9770	0.9154	53
6	0.8910	0.8923	0.8917	53
7	0.9497	0.8830	0.9152	53
8	0.8694	0.8812	0.8753	53
9	0.8719	0.9554	0.9117	53
10	0.9472	0.9203	0.9336	53
11	0.8934	0.8945	0.8939	53
12	0.8586	0.9562	0.9048	53
13	0.8516	0.8576	0.8546	52
14	0.8459	0.8504	0.8481	52
15	0.8596	0.9028	0.8806	52
16	0.8272	0.9062	0.8649	52
17	0.8222	0.8503	0.8360	52
18	0.8792	0.8474	0.8630	52
19	0.9574	0.8762	0.9150	52
20	0.8890	0.8084	0.8468	52
21	0.9099	0.8522	0.8801	52
22	0.8430	0.9252	0.8822	52
23	0.9317	0.8967	0.9139	52
24	0.9022	0.8495	0.8751	52
25	0.8364	0.9021	0.8680	52
26	0.9059	0.8477	0.8759	52
27	0.9447	0.9487	0.9467	52
28	0.9004	0.9536	0.9262	52
29	0.8819	0.9343	0.9074	52
30	0.8480	0.8212	0.8344	52
31	0.9166	0.8349	0.8739	52
32	0.9264	0.8407	0.8814	52
33	0.9373	0.8470	0.8899	52
34	0.8501	0.8418	0.8460	52
35	0.8813	0.8507	0.8657	52
accuracy			0.8839	1885
macro avg	0.8839	0.8864	0.8841	1885
weighted avg	0.8839	0.8866	0.8842	1885

(a)



(b)



(c)

Figure 3. Upskilling Recommender Model Training and Evaluation Results

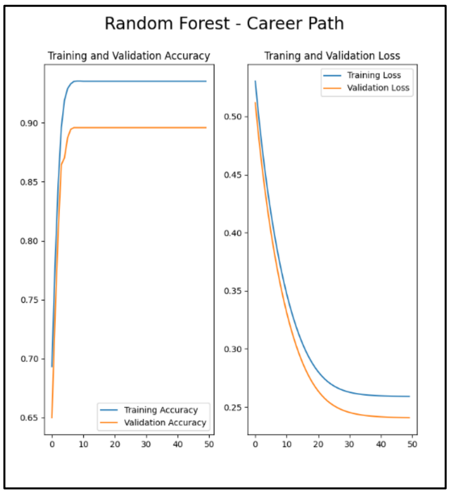
A significant equal accuracy and average precision scores of 88.39% was garnered by the upskilling recommender model during model evaluation as shown in Figure 3(a). It took about 40 epochs for the model to be trained before establishing a rigid accuracy of 88.46% while training losses started to stabilized starting epoch 30 and reached the least validation loss of 23.34% at epoch 50 as shown in Fig. 6(b). The training was halted after epoch 50 to avoid overfitting (Ying, 2019).

Meanwhile, Figure 3(c) shows the confusion matrix of the model with an average True Positive of 88.39%, minimum is at 82.22% and maximum is at 95.74. All True Positives of each upskilling label have darker blue color at the diagonal of the matrix indicating that the model can forecast upskilling label with significantly high precision (Seth, 2023) and that the selected SNN algorithm sufficiently suits the multi-label classification problem in relation to the dataset (Narkhede, 2023).

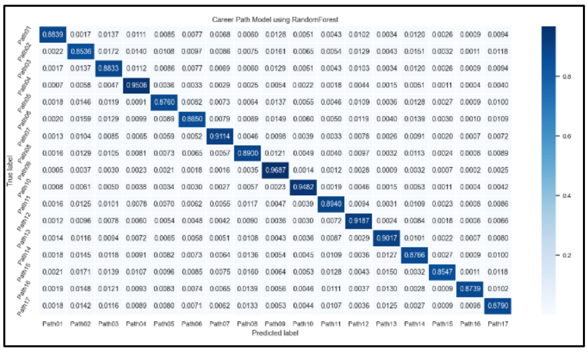
On the other hand, Figure 4 exhibits the results of the Career Path Recommender Model during machine learning model training and evaluation using sequential neural network (SNN) through the Random Forest algorithm

Career Path Model Evaluation Performance Metrics:

	precision	recall	f1-score	support
0	0.8839	0.9731	0.9264	111
1	0.8536	0.8266	0.8399	111
2	0.8833	0.8433	0.8629	111
3	0.9506	0.8756	0.9116	111
4	0.8760	0.8865	0.8812	111
5	0.8650	0.8962	0.8803	111
6	0.9114	0.9085	0.9099	111
7	0.8900	0.8663	0.8780	111
8	0.9687	0.8774	0.9208	111
9	0.9482	0.9313	0.9397	111
10	0.8940	0.9019	0.8980	111
11	0.9187	0.8905	0.9044	111
12	0.9017	0.9119	0.9068	111
13	0.8766	0.8641	0.8703	110
14	0.8547	0.9637	0.9059	111
15	0.8739	0.9763	0.9223	111
16	0.8790	0.8681	0.8735	110
accuracy			0.8958	1885
macro avg	0.8958	0.8977	0.8960	1885
weighted avg	0.8959	0.8978	0.8960	1885



(a) (b)



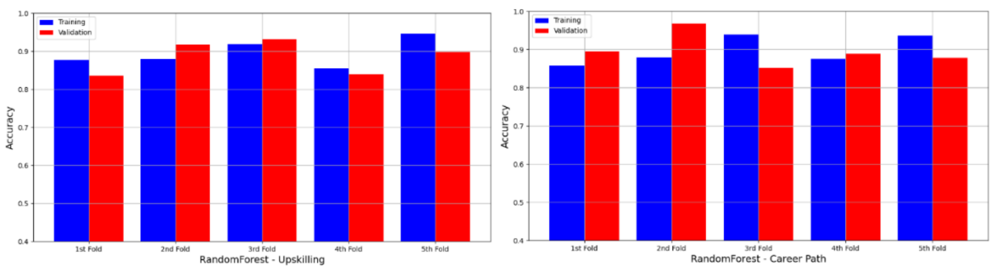
(c)

Figure 4. Upskilling Recommender Model Training and Evaluation Results

Similar to the upskilling recommender model, the career path recommender model obtained a significant evaluation performance accuracy and average precision scores of of 89.58% as shown in Figure 4(a). It took only less than 10 epochs before forming a steady accuracy of 89.57% while training losses started to stabilize during epoch 40 and attained at epoch 50 the smallest validation loss of 24.07% as displayed on Figure 4(b). Again, the training was also stopped after epoch 50 to shun overfitting as recommended in the study of Ying (2019).

The confusion matrix of the model shown on Figure 4(c) indicates an average True Positive of 89.58%, with 85.36% minimum TP value and 96.87% maximum TP value. Similar with the first model, all True Positives of each career path label have darker blue color at the diagonal of the matrix which is indicative that the developed model can forecast career paths with significantly high precision as prescribed in the study of Seth (2023) and can sufficiently fit the multi-label classification in relation to the utilized dataset parallel to the study of Narkhede (2023).

Figure 5 exhibits the results of the 5-fold cross validation conducted by the researchers to ensure the rigidity of accuracy among rows of datasets.



(a) (b)

Figure 5. 5-Fold Cross Validation Results of Models

Both the Upskilling Recommender Model and Career Path Recommender Model performed with significant accuracy during the 5-fold cross validation of each model's corresponding dataset with average accuracies of

89.65% and 88.95%, respectively. The Upskilling Recommender Model obtained a minimum accuracy of 85.14% and a maximum accuracy of 96.70% while the Career Path Recommender Model garnered a minimum accuracy of 83.50% and a maximum accuracy of 95.59%. This means that regardless of how the data will be randomized and how many attempts to create the models, both models' accuracies can be approximated by referencing the obtained minimum and maximum accuracy results of the 5-fold cross validation. This confirms that with the dataset on hand, the accuracy of the Upskilling Recommender Model cannot go down lower than at least 85% while Career Path Recommender Model's accuracy cannot go down below 83% as emphasized in the study of Nighania (2018).

4. Conclusion

The study developed a profile-based data analytics module for career path and upskilling recommendations of the Data-Driven Human Resource Analytics System (DDHRAS) for the largest state university in the Ilocandia Region of the Philippines. Experimental results have shown that the developed Upskilling Recommender Model and Career Path Recommender Model using sequential neural network and Random Forest algorithm have both obtained significant accuracy and precision during model training and evaluation and can provide upskilling and career path recommendations with significant accuracy using attributes from employee profile, human resource policies, and job qualifications.

The implication of the developed models, when fully integrated and implemented within the existing system of the HRD of the subject institution, can provide an alternative career path and upskilling recommendations with just few mouse clicks and lessen the burden of HRD evaluators, thus, save valuable resources which can then be used in strengthening quality services among its stakeholders.

References

[1] Abu-El-Haija, S., & Al-Khateeb, A. (2022). "Research methods in machine learning: A content analysis." *International Journal of Computer and Information Technology*, 11(1), 1-12.

[2] Berra, A., & Sanna, L. D. S. (2022). The evolving role of HR in the digital age: Challenges and opportunities. *International Journal of Human Resource Management*, 33(8), 2051-2074. DOI: 10.1080/09548799.2019.1604414.

[3] Brewster, C., & Larsen, H. H. (2021). The strategic role of HR in a changing world: A critical review. *Human Resource Management Journal*, 31(2), 507-534. DOI: 10.1111/hrmj.12413.

[4] Fan, Y., Lu, Y., & Zhao, X. (2020). A machine learning approach to career path prediction and upskilling recommendation in the IT industry. *Journal of Business Analytics*, 2(2), 189-202. DOI: 10.1080/24756820.2020.1730411.

[5] Karakurum, F., Gupta, A., & Raj, A. (2022). Leveraging machine learning for personalized upskilling recommendations. *Human Resource Management*, 61(3), 829-854. DOI: 10.1002/hrm.22235.

[6] Lawler, E. E., & Mohrman, S. A. (2019). The future of work and the role of HR: Challenges and opportunities. *Human Resource Management*, 58(12), 1804-1824. DOI: 10.1002/hrm.22061.

[7] Liu, Y., Shi, Y., & Zhao, P. (2022). A personalized recommendation system for upskilling and career development. *Decision Support Systems*, 151, 113588. DOI: 10.1016/j.dss.2022.113588

[8] Narkhede, S. (2023). Understanding Confusion Matrix. *Towards Data Science*. Web article: https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62.

[9] Nighania, K. (2018). Various Ways to Evaluate a Machine Learning Model's Performance. *Towards Data Science*, https://towardsdatascience.com/various-ways-to-evaluate-a-machine-learning-models-performance-230449055f15.

[10] Romero, R., López, M. H., & Arnau, J. (2021). Machine learning for skill gap analysis and recommendation in the learning and development context. *Studies in Educational Evaluation*, 73, 100670. DOI: 10.1016/j.stueduc.2021.100670

[11] Schuler, R. S., & Jackson, S. E. (2021). The changing landscape of HR analytics: A strategic perspective. *Organizational Dynamics*, 50(1), 101314. DOI: 10.1016/j.orgdyn.2020.101314.

[12] Seth, N. (2023). Estimation of Neurons and Forward Propagation in Neural Net. *Analytics Vidhya*. Web article: https://www.analyticsvidhya.com/blog/2021/04/estimation-of-neurons-and-forward-propagation-in-neural-net/.

[13] Singh, A., & Singh, P. K. (2021). Profile-based career path recommendation system using machine learning techniques. *Artificial Intelligence and Machine Learning in Healthcare*, 20(1), 1-14. DOI: 10.1007/978-981-15-7406-9_1

[14] Tzafrir, S., Baruch, Y., & De Beer, J. (2020). HR 4.0: The next-generation human resources model.

Journal of Management Studies, 57(1), 1-31. DOI: 10.1111/joms.12434.

[15] Xu, Y., Zhao, P., & Li, J. (2020). A profile-driven recommendation system for career path planning and upskilling in the big data era. *International Journal of Human Resource Management*, 31(10), 1502-1532. DOI: 10.1080/09548799.2018.1455271

[16] Yang, M., Chen, B., & Zhou, M. (2023). Towards an AI-powered HR system for career development and upskilling. *International Journal of Human Resource Management*, 34(2), 307-334. DOI: 10.1080/09548799.2021.1975650.

[17] Ying, X. (2019). An Overview of Overfitting and its Solutions. *Journal of Physics: Conference Series (IOP Science, 2019)*. DOI: 10.1088/1742-6596/1168/2/022022

[18] Zhang, Y., Chen, L., & Zhao, P. (2023). Explainable recommendation systems for upskilling and career development. *Decision Support Systems*, 152, 113610. DOI: 10.1016/j.dss.2023.113610.