JPPIPA 9(8) (2023)

Jurnal Penelitian Pendidikan IPA

Journal of Research in Science Education

http://jppipa.unram.ac.id/index.php/jppipa/index

Optimizing Learning Pathway in Human Capital Development Programs with BERT Transformer

Irvan Zidny[1*], Ira Puspitasari[2], Imam Yuadi[3]

[1] Human Resources and Development, Faculty School of Graduate, Universitas Airlangga, Surabaya, Indonesia.
[2] Information System Study Program, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia.
[3] Department of Information and Library Science, Faculty of Social Sciences and Political Science, Universitas Airlangga, Surabaya, Indonesia.

Received: June 20, 2023
Revised: July 4, 2023
Accepted: August 25, 2023
Published: August 31, 2023

Corresponding Author:
Irvan Zidny
irvan.zidny-2022@pasca.unair.ac.id

DOI: 10.29303/jppipa.v9i8.4531

© 2023 The Authors. This open access article is distributed under a (CC-BY License)

Abstract: This study presents a novel approach to assist learning analysts in identifying suitable learning pathways based on historical training data through the utilization of text mining techniques. The dataset utilized in this research comprises training data from the year 2021 and the Course Development Management Program (CDMP) catalogue. The BERT 'bert-base-nli-mean-tokens' model is employed for encoding purposes. By comparing the training data names from 2021 with the CDMP catalogue using cosine similarity and dot score, valuable insights are obtained. The findings indicate that cosine similarity is a more effective measure for interpreting the data, thereby simplifying the process for learning analysts and managers in identifying appropriate learning paths for their employees. This research provides a practical solution that leverages text mining techniques to optimize the analysis and decision-making processes in learning and development domains, enabling organizations to enhance the effectiveness and efficiency of their training programs.

Keywords: BERT transformers; Big data; Sentence similarity; Text mining

Introduction

Strategic human resources (HR) are critical to achieving an organization's overall goals and have a functional approach that is excellent for implementing development initiatives (Valmohammadi & Shahrashoob, 2022). In modern business, employees, competence, and professionalism are important factors in economic and financial activities (Bazyliuk et al., 2023). Employee retention is an important strategy for HR managers (Samuel & Chipunza, 2009). There is much evidence that employee retention is important for managers to address the increasing levels and costs of employee turnover (Aman-Ullah et al., 2022; Zayed et al., 2022). For example, factors of influence employee retention, including wage, payment length, employment agreement, labor market mobility, and participation in significant building projects (Le et al., 2023). As a result, one of a company's primary responsibilities is to assist and continuously improve its employees' capabilities, developing talents and abilities that enable individuals to swiftly adapt to new technology and the complex challenges of the modern marketplace (Bazyliuk et al., 2023), which falls under employee job satisfaction (Manuere, 2018).

Human resource management in an organization has changed dramatically in recent years. Considering the vital link between organization management strategy and human resource management has resulted in a major tendency for human resource managers to take a strategic approach (Arunprasad et al., 2023). Strategic Human Resource Management (SHRM) underlines the evolution of the HRM position from restrictive, receptive, and administrative to one that is illustrative, active, and executive (Fahim, 2018).

According to Sahoo (2011), the notion of SHRM is connected to the incorporation of HR operations with the company's concept for enterprise (Fahim, 2018). The attraction of high-quality employees from a managerial perspective, has become more important than ever

How to Cite:
Zidny, I., Puspitasari, I., & Yuadi, I. (2023). Optimizing Learning Pathway in Human Capital Development Programs with BERT Transformer. *Jurnal Penelitian Pendidikan IPA*, 9(8), 6093–6100. https://doi.org/10.29303/jppipa.v9i8.4531

before (Holtom et al., 2008). Employee attraction refers to the positive feelings and perceptions that an employee has about their job, leadership, compensation, and workplace (Togia et al., 2004). For example, essential elements such as age, education, surroundings, workplace, salary level, training and development, and employee recognition have an impact on employee job satisfaction (Bodur, 2002). Polls on demographics, practice patterns, job satisfaction, and career intentions of endocrinologists in Poland show that work on human resource planning and management needs to be improved, because based on the results, the average working week is 45 hours, and they administrate an average of 100 patients (Zgliczyński & Skowron, 2023).

Company X has conducted a lot of training and development as well as training program documentation since 2017. As the business needs of Company X have evolved, this data needs to be updated to obtain information on the training programs that have been implemented. Based on the data held by Company X, the data does not yet have a learning path to show the results or areas supported by the training program. As a result, HR managers must compare text similarities between the implementation program and the learning path to examine training programs that match the program with a proper learning path.

A series of changes that have occurred, such as globalization, an increase in knowledge work, the acceleration of technological progress, and increased competition, make it important for institutions to have superior and competitive human resources for organizational success (Kankaew et al., 2023; Li, 2023). Employers must grasp newly discovered fields in the workplace and need certain competences from employees, including a combination of technical and non-technical skills (Ternikov, 2022). The vast amount of training data collected by Company X has made connecting the right learning path to the programs that have been executed problematic. Additionally, this has made it difficult for human resource managers to analyze the suitability between training and which learning pathway is supported in achieving employee development. Therefore, to solve this problem, researchers suggest using the concept of big data to facilitate human resource management in analyzing which learning pathway is suitable for the training program.

Text mining (TM) is a way of converting unstructured data into structured information by evaluating different manuscripts such as networking platforms postings and emails, dealing with bilingual text and comparable terms in networking platforms, and acknowledging the existence of many topics in a single phrase (Chowdhary, 2020; Puri et al., 2023). TM has proven to be a powerful tool in health research, especially when combined with natural language processing (NLP) (Jensen et al., 2012). The research conducted by Islam et al. (2008) used corpus-based semantic similarity assessment of words and a normalized and updated Longest Common Subsequence (LCS) matching of strings approach to estimate text similarity. The findings demonstrated that the suggested strategy outperformed previous methods in calculating text similarity. Unlike the research by Islam et al. (2008), the researcher will use BERT base in their study to convert text into vectors, and then similarity will be calculated using cosine and distance similarity. The convergence becomes more self-adaptive due to the cosine similarity. The adjusted environment selection maximises the value of previous generation vectors while introducing unpredictability into the following iteration (Zhang et al., 2023).

This research applies Text Mining to determine the appropriate learning pathways for employee training history programs and analyze the suitability between the history program and learning pathways catalogue. Text mining utilization benefits managers as they can analyze the suitability between the employee training programs conducted and the learning paths available in the catalogue. This analysis also assists them in ensuring that the training programs conducted are genuinely in line with the needs and help in employees' career development. In addition, in the future, this can also help improve efficiency in planning and implementing employee training programs.

Method

A schematic diagram of the analytics process used in this study is presented in Figure 1.

Data Gathering

The first step is to collect training data and the Course Development Management Playbook (CDMP) catalogue. The training data comprises of programs offered by employees of Company X in 2021, whilst the CDMP catalogue data maps the company's existing learning paths. The researcher can receive this data from participants who submit training plans to the learning analyzer. The training data obtained includes all training programs that employees participated in during 2021, and there were about 154 programs implemented. In addition, there is also CDMP training catalogue data consisting of 651 CDMP data along with the Learning Pathway from the available course catalogue. In this study, this data will be used as a source to analyze the suitability of the programs that have been implemented with the appropriate learning pathway.

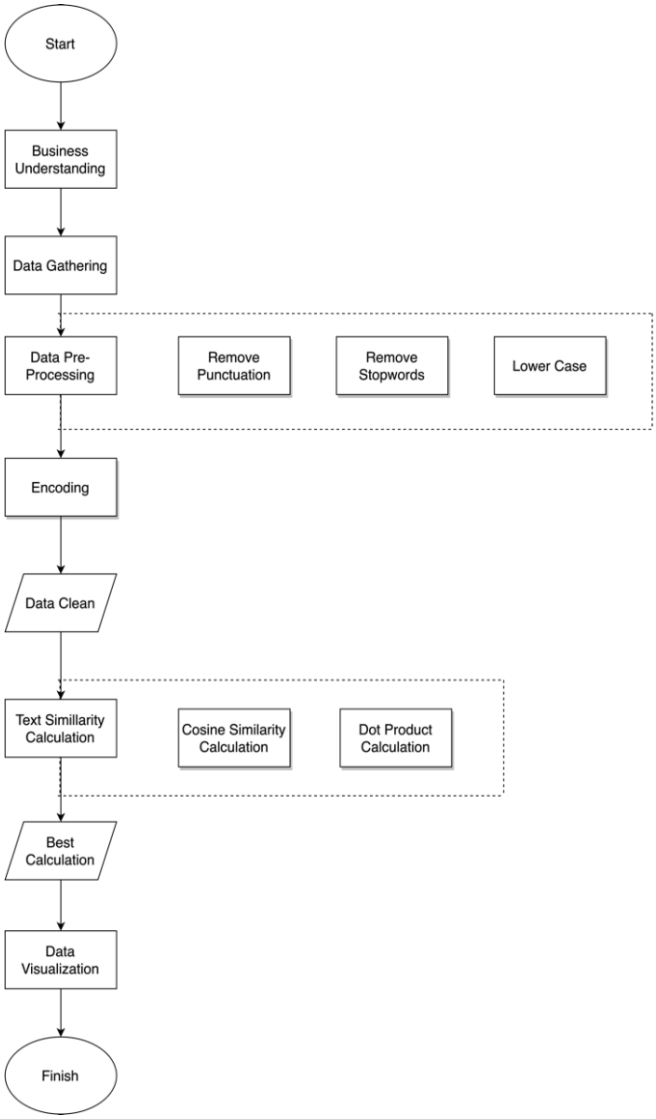

Figure 1. Methodology design of the major phases of the project

Data Preprocessing

This process includes data cleaning, format conversion, and data preparation for encoding. The steps performed in pre-processing include removing punctuation and stopwords and converting all letters to lowercase in each word in the data. This is done to facilitate the next step of encoding and data analysis.

Encoding

The model used to perform encoding on the training data is the Bidirectional Encoder Representations from Transformer (BERT) Base or 'bert-base-nli-mean-tokens'. The selected BERT is a type that can perform bidirectional language processing and obtain representations of the entire text (Ullah et al., 2022). Bert-base-nli-mean-tokens is a BERT-based model that pools mean-tokens to generate sentence embeddings and has achieved a 77.12 performance on the STS benchmark dataset, making it suitable for various NLP tasks due to its ability to capture complex relationships between words and sentences (Conneau et al., 2017). The work by Roman et al. (2021) employed word embedding algorithms such as GloVe, InferSent, and BERT to categorize citation contexts on the Citation Context Dataset, which had 10 million entries. The results indicated that the BERT approach had the highest accuracy of 89% of all examined methods. In other words, BERT has been shown to be more accurate than other word embedding approaches such as GloVe and InferSent for categorizing citation contexts with the purpose of citation.

There is evidence that pre-training language models are effective in improving various natural language processing tasks (Conneau et al., 2017; Devlin et al., 2019; Peters et al., 2018; Vaswani et al., 2017). Natural language inference (NLI), for example, is part of these tasks. NLI is an appropriate testing ground for theories of semantic representation, and it may generate rich and wide domain semantic representations with the addition of a large-scale and realistic corpus of phrase pairs tagged for entailment, contradiction, and independence (Bowman et al., 2015).

Figure 2 shows the schema applied to find similarities between the two sentences used in this study. After encoding data into numerical representations, the following step is to compute the text similarity. Algorithms such as cosine similarity and dot distance are used to determine the level of similarity between different training data. Using these algorithms, we can determine the level of similarity between documents or sentences for further analysis.

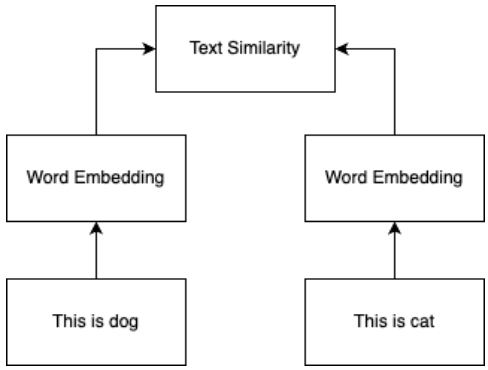

Figure 2. Text similarity design

Text Similarity Calculation

Cosine similarity computes the similarity between two vectors in an n-dimensional space by calculating cosine of the angle formed by the two vectors (Gidaris & Komodakis, 2018). In the context of text analysis, the similarity is measured using cosine similarity between two texts represented as vectors (Sun et al., 2013) using

models such as BERT. The cosine similarity function is defined as:

$$\text{Cosine similarity (A, B)} = \frac{\text{A} \cdot \text{B}}{||\text{A}||\,||\text{B}||} \quad (1)$$

In the Equation 1, A and B are vectors that represent two texts, where dot_product (A · B) is the dot product of A and B, and norm(A) and norm(B) are the norms of vectors A and B. The greater the cosine similarity score, the more similar the two sentences are. The cosine similarity value is between -1 and 1, with 1 signifying perfect resemblance between the two vectors.

Dot distance, also known as dot product, is a similarity calculation method that measures how far two vectors are in the same direction in a vector space (Qin et al., 2021). If the first vector is represented as A = [a1, a2, ..., an] and the second vector is represented as B = [b1, b2, ..., bn], then the dot distance is calculated as:

$$\text{Dot distance (A . B)} = a_1 * b_1 + a_2 * b_2 + a_3 * b_3 + \cdots + a_n * b_n \quad (2)$$

After the text similarity calculation, the next step is to translate and analyze the results to determine conclusions and recommendations for follow-up actions. In this stage, the results of text similarity calculation can provide information about the similarities or differences between different training data.

Result and Discussion

Table 1 is a sample of training history that has been conducted in 2021. There are several attributes used, namely training_name, training_objid, CFU/FU, category, type of training, training_location, provider, academy_event, and location. The "training_name" attribute is the name or title of the training held. "Training_objid" is a unique identification number for each training. "CFU/FU" is the unit of training participants. "Category" is the training category, which can include the type of exercise. The "type" attribute describes the type of training, such as online training or face-to-face training. "Training_location" indicates the physical location of the training, whether it is conducted inside or outside the office building. The "provider" attribute is the training organizer, such as the company that holds the training or the training vendor that collaborates with the company. "Academy_event" describes the sub-unit that runs the training. Finally, the "location" attribute is the physical location of the training organizer or training vendor. This table is very useful for companies to monitor and manage training programs for employees.

Table 1. Sample of Training History that has been Conducted in 2021

Training Name	Training Objid	CFU/FU	Category	Type	Training Location	Provider	Academy Event	Location
Swift Advanced	8011XXXX	DB	LAT	E-learning	BDO	LA3	DSP	Online
Business Management	8011XXXX	DB	LAT	E-learning	BDO	LA3	DSP	Online
AWS Cloud Practitioner Essentials	8011XXXX	DB	LAT	Virtual Class-room	BDO	LA3	NITS	Online
Digital Product Management	8011XXXX	DB	LAT	E-learning	BDO	LA3	DSP	Online
Head of Data	8011XXXX	DB	LAT	Virtual Class-room	BDO	LA7	DSP	Online

Table 2. Sample of Course Development Management Playbook (CDMP)

School	Lp Code	Learning Pathway	Slp Code	Sub Learning Pathway	Sslp Code	Sub Learning Pathway	Course Code	Course Name
SoDPS	DPS.1	Data Science	DPS.1.1	Math & Statistics	DPS.1.1.A	Mathematics	DPS.1.1.A.1	Introduction to Discrete Mathematics for Computer Science Specialization
SoDPS	DPS.1	Data Science	DPS.1.1	Math & Statistics	DPS.1.1.A	Mathematics	DPS.1.1.A.2	Mathematics for Machine Learning Specialization
SoDPS	DPS.1	Data Science	DPS.1.1	Math & Statistics	DPS.1.1.A	Mathematics	DPS.1.1.A.3	Mathematics for Data Science Specialization
SoDPS	DPS.1	Data Science	DPS.1.1	Math & Statistics	DPS.1.1.B	Statistics	DPS.1.1.B.1	Statistics and Probability
SoDPS	DPS.1	Data Science	DPS.1.1	Math & Statistics	DPS.1.1.B	Statistics	DPS.1.1.B.2	Business Statistics and Analysis Specialization

Table 2 is a sample catalogue of CDMP training. There are several attributes used, namely school, lp code, learning pathway, slp code, sub learning pathway, sslp code, sub-sub learning pathway, course code, and course name. The table contains data about a list of subjects (school) and learning pathways (learning pathway) in an educational institution or organization. Each learning pathway consists of three levels: learning pathway, sub-learning pathway, and sub-sub learning pathway. Each level has a unique code for identification (lp_code, slp_code, and sslp_code). In addition, this table also includes data on the subjects (course) contained in each sub-sub learning pathway. Each subject also has a unique code (course_code) and subject name (course_name). This table can help education or organization managers in preparing the right learning pathways and providing the necessary information for students or employees to choose subjects that suit their needs.

In this research, the BERT model ('bert-base-nli-mean-tokens') was used to tokenize the data during the pre-processing stage. Tokenization was performed on the training program history data and the course name catalogue data so that the data could be processed and analysed more effectively using algorithms such as cosine similarity and dot distance in determining the level of similarity between different training data. Table 3 shows the calculation of text similarity using cosine similarity, which provides results that are easier to understand compared to the calculation of text similarity using dot score in Table 4. This is because cosine similarity has a range of values from 0 to 1 that can be easily interpreted as the level of similarity which increases as it approaches 1, whereas in dot score, the greater the resulting value, the more similar the texts are, making clear interpretation more difficult.

According to Table 3, a number close to 0 indicates a low degree of similarity between the training title and the CDMP catalogue, whereas a value close to 1 suggests a high or identical level of similarity between the training title and the CDMP catalogue. For example, the training program "Swift Advance" has a cosine similarity value of 0.18097565 when compared to "Introduction to Discrete Mathematics for Computer Science Specialization". This indicates that the similarity between the two training programs is not sufficient, and therefore "Swift Advance" cannot be considered a continuation of "Introduction to Discrete Mathematics for Computer Science Specialization". Table 4 shows the calculation of text approximation using dot score, which indicates that "Swift Advance" has a value of 50.521927 when compared to "Introduction to Discrete Mathematics for Computer Science Specialization". However, since dot score is difficult to interpret clearly, cosine similarity calculations are more recommended for determining the level of similarity between different training data.

Figure 3 is a visual display created using Tableau software, which shows the cosine similarity calculation between each pair of training programs and learning paths in the CDMP catalogue. This display makes it easier for users to understand the relationship between training programs and learning paths and to determine the most suitable learning path for a particular training program. Additionally, cosine similarity calculations can be used to identify training programs that have similar topics with other training programs, enabling users to group training programs into more specific and relevant categories. This can help make better decisions in planning and developing training programs.

Table 3. Sample Text Similarity Calculation Using Cosine Similarity

	Swift advanced	Business management	Aws cloud practitioner essentials
Introduction discrete mathematics computer science specialization	0.1809757	0.3604596	0.4013383
Mathematics machine learning specialization	0.1990268	0.3868091	0.370091
Mathematics data science specialization	0.2264968	0.365726	0.417759
Statistics probability	0.442252	0.34871	0.4674507
Business statistics analysis specialization	0.2242318	0.7547472	0.4283131

Table 3. Sample Text Similarity Calculation Using Dot Score

	Swift advanced	Business management	Aws cloud practitioner essentials
Introduction discrete mathematics computer science specialization	50.521927	99.80604	105.87479
Mathematics machine learning specialization	57.232674	110.32384	100.56867
Mathematics data science specialization	64.51117	103.31636	112.43996
Statistics probability	126.36042	98.820206	126.21146
Business statistics analysis specialization	62.637527	209.11218	113.06308

The authors investigated the use of BERT Transformer 'bert-base-nli-mean-tokens' to determine suitable learning paths for implemented programs using a platform developed by the authors. According to the findings of this study, the learning platform as in Figure 3 can facilitate learning analysts in analyzing appropriate learning types efficiently. However, a limitation of the modeling results is that after creating the platform for learning analysts, further analysis is needed based on program names and corresponding learning types. Additionally, the researchers agreed that using cosine similarity calculations made it easier to identify similarities between the text of implemented programs and learning paths.

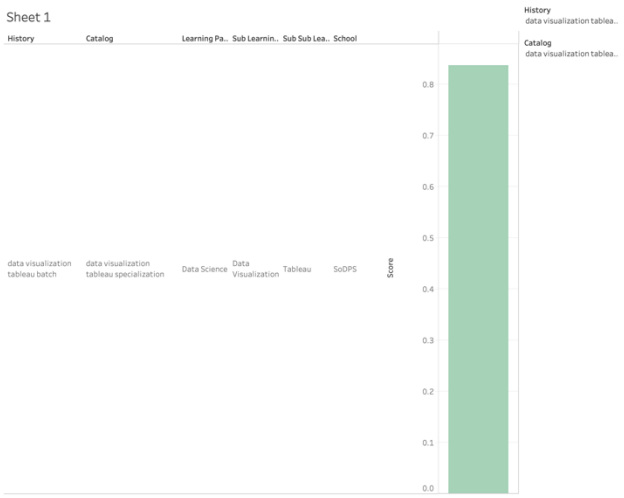

Figure 3. Dashboard visualization

Conclusion

Based on these results, the study concludes that text mining techniques, particularly BERT Transformer encoding and similarity testing, can help find the right learning paths for employee training programs. This technique also simplifies the process for learning analysts and managers to identify the appropriate learning paths for their employees. Furthermore, data visualization tools such as Tableau can assist in presenting the data in a more user-friendly manner. The study recommends that future research explore the application of other text mining techniques or different encoding models to improve the accuracy of the results. Additionally, the analysis results also provide insights and new knowledge about employees' learning preferences and needs, which can be used as a basis for evaluating and planning training programs in the future.

Acknowledgments
I'd like to thank Ms. Wewen Indriyani, the Learning Analyst of Company X, for granting permission and implementing the findings of my research in her department. I appreciate her collaboration and assistance during the study process, as well as the opportunity to help Company X establish staff training programs.

Author Contributions
The main author, Irvan Zidny, contributed to designing research, conducting research, and writing research articles. The second and third author, Ira Puspitasari and Imam Yuadi, played a role in guiding the research to writing articles.

Funding
This research received no external fundings.

Conflicts of Interest
The authors declare there are no conflicts of interest.

References

Aman-Ullah, A., Aziz, A., Ibrahim, H., Mehmood, W., & Abdullah Abbas, Y. (2022). The Impact of Job Security, Job Satisfaction and Job Embeddedness on Employee Retention: An Empirical Investigation of Pakistan's Health-Care Industry. *Journal of Asia Business Studies, 16*(6), 904–922. https://doi.org/10.1108/JABS-12-2020-0480

Arunprasad P., Dey, C., Santhanam, N., & Bin Ahmad, K. Z. (2023). Strategic HRM practices, learning orientation and learning competence: study from retail industry. *Learning Organization, 30*(4), 406–425. https://doi.org/10.1108/TLO-05-2022-0066

Bazyliuk, A., Khomenko, V., & Ilchenko, V. (2023). World Experience of Investing and Human Capital Development. In *Lecture Notes in Networks and Systems* (Vol. 487). https://doi.org/10.1007/978-3-031-08084-5_31

Bodur, S. (2002). Job Satisfaction of Health Care Staff Employed at Health Centres in Turkey. *Occup Med (Lond), 52*(6), 353-355. https://doi.org/10.1093/occmed/52.6.353. PMID: 12361997

Bowman, S. R., Angeli, G., Potts, C., & Manning, C. D. (2015). A Large Annotated Corpus for Learning Natural Language Inference. *Conference Proceedings-EMNLP 2015: Conference on Empirical Methods in Natural Language Processing*, 632–642. https://doi.org/10.18653/v1/d15-1075

Chowdhary, K. R. (2020). *Fundamentals of Artificial Intelligence*. India: Springer. https://doi.org/10.1007/978-81-322-3972-7

Conneau, A., Kiela, D., Schwenk, H., Barrault, L., & Bordes, A. (2017). *Supervised Learning of Universal Sentence Representations from Natural Language Inference Data*. New York: Cornell University. https://doi.org/10.48550/arXiv.1705.02364

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019).

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), Minneapolis, Minnesota,* 4171–4186. https://doi.org/10.18653/v1/N19-1423

Fahim, M. G. A. (2018). Strategic Human Resource Management and Public Employee Retention. *Review of Economics and Political Science, 3*(2), 20–39. https://doi.org/10.1108/reps-07-2018-002

Gidaris, S., & Komodakis, N. (2018). Dynamic Few-Shot Visual Learning without Forgetting. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 4367–4375. https://doi.org/10.1109/CVPR.2018.00459

Holtom, B. C., Mitchell, T. R., Lee, T. W., & Eberly, M. B. (2008). 5 Turnover and Retention Research: A Glance at the Past, a Closer Review of the Present, and a Venture into the Future. *The Academy of Management Annals, 2*(1), 231–274. https://doi.org/10.1080/19416520802211552

Islam, A., & Inkpen, D. (2008). Semantic Text Similarity using Corpus-Based Word Similarity and String Similarity. *ACM Transactions on Knowledge Discovery from Data, 2*(2), 1-25. https://doi.org/10.1145/1376815.1376819

Jensen, P. B., Jensen, L. J., & Brunak, S. (2012). Mining Electronic Health Records: Towards Better Research Applications and Clinical Care. *Nature Reviews Genetics, 13*(6), 395–405. https://doi.org/10.1038/nrg3208

Kankaew, K., Komolvanij, T., Paethrangsi, N., Kungwol, K., Charernnit, K., Worasuwannarak, B., & Waramontri, R. (2023). Digitalization and Crises: The Opportunity Overturned on Sustainable Human Resource Development in Airlines in Thailand. In *Strategic Human Resource Management in the Hospitality Industry: A Digitalized Economic Paradigm*. https://doi.org/10.4018/978-1-6684-7494-5.ch003

Le, D.-A., Le-Hoai, L., Le, V. H., & Dang, C. N. (2023). Factors Affecting Employee Retention in Construction: Empirical Study in the Mekong Delta Region. In *Lecture Notes in Civil Engineering* (Vol. 268). https://doi.org/10.1007/978-981-19-3303-5_21

Li, R. (2023). Three-Pillar Model of HR in the Context of Digital Transformation. In *Lecture Notes on Data Engineering and Communications Technologies* (Vol. 168). https://doi.org/10.1007/978-981-99-1157-8_43

Manuere, F. (2018). Human Capital Development Programmes and their Effect on the Job Satisfaction of Workers in Zimbabwe Urban Municipalities: The Case of Chinhoyi. *International Journal of Academic Research in Business and Social Sciences, 7*(12), 874-889. https://doi.org/10.6007/ijarbss/v7-i12/3719

Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). Deep Contextualized Word Representations. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers), New Orleans, Louisiana,* 2227–2237. https://doi.org/10.18653/v1/N18-1202

Puri, V., Mondal, S., Das, S., & Vrana, V. G. (2023). Blockchain Propels Tourism Industry—An Attempt to Explore Topics and Information in Smart Tourism Management through Text Mining and Machine Learning. *Informatics, 10*(1), 9. https://doi.org/10.3390/informatics10010009

Qin, C., Luo, X., Deng, C., Shu, K., Zhu, W., Griss, J., Hermjakob, H., Bai, M., & Perez-Riverol, Y. (2021). Deep Learning Embedder Method and Tool for Mass Spectra Similarity Search. *Journal of Proteomics, 232*. https://doi.org/10.1016/j.jprot.2020.104070

Roman, M., Shahid, A., Khan, S., Koubaa, A., & Yu, L. (2021). Citation Intent Classification using Word Embedding. *IEEE Access, 9*, 9982–9995. https://doi.org/10.1109/ACCESS.2021.3050547

Samuel, M. O., & Chipunza, C. (2009). Employee Retention and Turnover: Using Motivational Variables as a Panacea. *African Journal of Business Management, 3*(8), 410–415. https://doi.org/10.5897/AJBM09.125

Sun, W., Wang, B., Cao, N., Li, M., Lou, W., Hou, Y. T., & Li, H. (2013). Privacy-Preserving Multi-Keyword Text Search in the Cloud Supporting Similarity-Based Ranking. *ASIA CCS 2013-Proceedings of the 8th ACM SIGSAC Symposium on Information, Computer and Communications Security*, 71–81. https://doi.org/10.1145/2484313.2484322

Ternikov, A. (2022). Soft and Hard Skills Identification: Insights from IT Job Advertisements in the CIS Region. *PeerJ Computer Science, 8,* e946. https://doi.org/10.7717/PEERJ-CS.946

Togia, A., Koustelios, A., & Tsigilis, N. (2004). Job Satisfaction among Greek Academic Librarians. *Library and Information Science Research, 26*(3), 373–383. https://doi.org/10.1016/j.lisr.2004.01.004

Ullah, S., Liaqat, M., Asif, A., Khan, A., Aslam, U., & Asif, H. (2022). Deep Auto Encoder based Chatbot for Discrete Math Course. *2022 International Conference on Recent Advances in Electrical Engineering and Computer Sciences (RAEE and CS), Islamabad, Pakistan,* 1-7.

https://doi.org/10.1109/RAEECS56511.2022.9954572

Valmohammadi, C., & Shahrashoob, V. (2022). Determinants of the Human Capital Department Development Programs toward Achieving Functional Strategic Objectives: A Fuzzy Hybrid Approach. *Industrial and Commercial Training, 54*(3), 377–395. https://doi.org/10.1108/ICT-02-2021-0012

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is All You Need. *31st Converence on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA,* 5999–6009. Retrieved from https://papers.nips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

Zayed, N. M., Edeh, F. O., Islam, K. M. A., Nitsenko, V., Dubovyk, T., & Doroshuk, H. (2022). An Investigation into the Effect of Knowledge Management on Employee Retention in the Telecom Sector. *Administrative Sciences, 12*(4), 138. https://doi.org/10.3390/admsci12040138

Zgliczyński, W. S., & Skowron, A. (2023). Demographics, Practice Patterns, Job Satisfaction, and Career Plans of Endocrinologists in Poland. *Endokrynologia Polska, 74*(3), 331-339 https://doi.org/10.5603/EP.a2023.0040

Zhang, Q., Zou, D., & Duan, N. (2023). An Improved Differential Evolution using Self-Adaptable Cosine Similarity for Economic Emission Dispatch. *Energy, 283*. https://doi.org/10.1016/j.energy.2023.128421