As per UGC guidelines an electronic bar code is provided to seure your paper

International Journal for Modern Trends in Science and Technology, 9(02): 26-32, 2023
Copyright © 2023 International Journal for Modern Trends in Science and Technology
ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST0902004
Available online at: http://www.ijmtst.com/vol9issue02.html

Applications, Challenges, and Implications of data processing models in Human Capital Management

Parth Chopra

Independent Researcher, New York.

To Cite this Article
Parth Chopra. Applications, Challenges, and Implications of data processing models in Human Capital Management. International Journal for Modern Trends in Science and Technology 2023, 9(02), pp. 26-32. https://doi.org/10.46501/IJMTST0902004

Article Info
Received: 26 December 2022; Accepted: 30 January 2023; Published: 02 February 2023.

ABSTRACT

Human Capital Management (HCM) is the cohesive set of strategy and processes aimed at managing an organization's employees in the most efficient manner. HR managers make decisions that directly impact an organization's work culture, competitiveness and ability to meet goals. A qualitative study by Buyens and De Vos (2006) [1] concluded that the value of the HR manager as perceived by top managers and line managers extend beyond just the formulation and execution of HR strategy to delivering real topline business impact. Not to mention, weak HCM can cost companies real dollars. For example, the Department of Labor (US) estimated the cost of a bad hire at ~30% of their year-1 earnings.

There already exist several papers establishing a positive causal relationship between the use of data processing and machine learning models and improved HCM outcomes [2] (Mark Tomassen, 2016). Moreover, there exist multitudes of research (Chui et al., 2015) [3] answering the question "Can technology replace humans in HCM functions?" and the resounding conclusion has been that this is highly unlikely in the near future. In fact, a study by Frey and Osbourne (2017) [4] estimated the probability of computerization for over 700 occupations and ranked the HR manager position at 28/702 implying that the function would be near impossible to replace by technology. Therefore, we begin our paper with the foundation that while AI, ML and other computer science and data processing models are not poised to or even intended to replace the HR manager, there is value to be derived from leveraging these tools for better decision making.

KEYWORDS: Human Capital Management, Data Processing Applications, HCM

1. INTRODUCTION

The objective of this paper is to explore the areas of HCM where big data machine learning can create incremental value. In order to do this, we break down use cases into the following application areas (Fig. 1): Recruiting, Development, Compensation and Performance Management. Then, we will explore the challenges and pitfalls associated with relying on data to drive human personnel decisions. Finally, we will discuss mechanisms to overcome these challenges and the cost to US productivity and organizations that comes from continued reliance on human decision-making for their HCM strategy and practices.

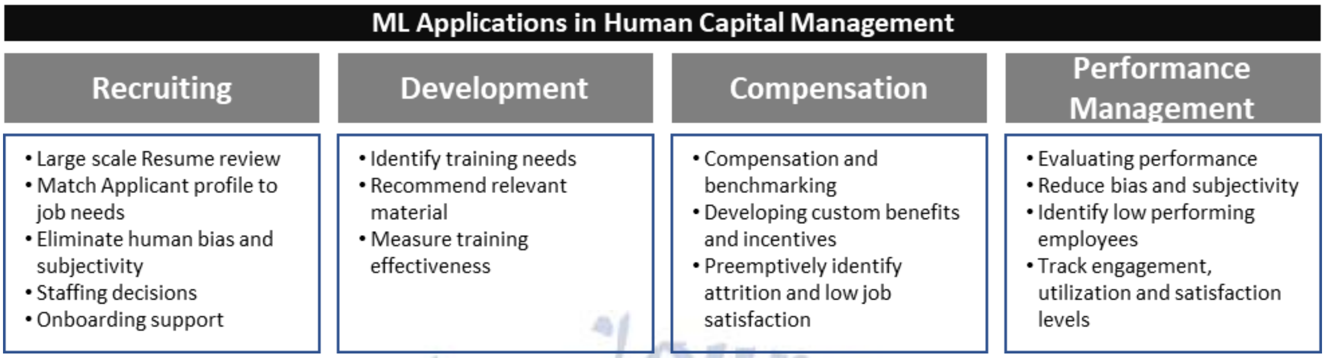

Fig. 1: ML Applications within HCM

2. Big data processing applications within the HCM life cycle

Recruiting: According to Glassdoor, an average vacancy invites over 250 resumes while typically only 4-6 candidates are interviewed. Moreover, human recruiters take only 6 seconds on average to go through an applicant resume. With companies spending an average of $4,000 to hire every new employee, this cumulatively becomes a high stakes decision where errors are expensive. Studies also show that hiring managers often rely on personal factors and cues such as handshakes, clothing style, accents etc. while judging an applicant's fit for a role hereby making sub-optimal decisions. Big data processing models, especially ones that can be trained on an organization's existing datasets around successful employees, can help HR and hiring managers analyze vast pools of applicant profiles in short spans of time and also eliminate the bias and stereotypes that comes with human decision-making, thereby creating a more diverse and competent workforce. ML models can also be trained to help supervisors make allocation and staffing decisions by identifying the projects best suited for an employee's specific interests, background and skillsets. Finally, HCM does not end once the hiring decision is made. New employees need the right onboarding resources to be set up for success and to ensure a fast ramp. Big data models can suggest the right teams and individuals for a new employee to connect with, propose relevant trainings, predict and preemptively answer common and/or critical questions required to ramp quickly. A study conducted by the Brandon Hall Group [5] discovered that robust processes around onboarding improved new-hire retention and productivity by 82% and 70% respectively.

Development: For this paper, let's assume that training and development are good for employee performance and productivity. With that assumption, let's now focus on how big data analyses can significantly improve learning outcomes for employees by assisting HR and other People Managers.

Learning Management Systems ("LMS") are used by large enterprises and small businesses alike to meet compliance requirements – such as complete anti-sexual harassment trainings. Technologically speaking, it would be easy to leverage them for additional on-the-job trainings, however people managers often either don't have the insight to assign relevant trainings, or the time to identify, assign, communicate the trainings to employees and track progress. This is precisely where intelligent next-generational LMSs can assist through their data integration capabilities.

For example, LMS combined with Big Data Analysis models have started identifying companies similar to each other – for the sake of argument, say Twitter and Facebook – and start recommending courses to their employees. Thus, when People Managers at Facebook start assigning a good course in social media advertising to their associates, the LMS would start recommending that to Twitter People Managers too. Similarly, when Twitter associates start enrolling themselves (self-enrollment) in mental health courses seemingly from the 2022-layoffs, similar courses would be recommended to Twitter associates for self-enrollment too. These help feed into overall user preferences too – to identify which courses/modules are more popular for which subsets of the industry/population. Recommendation algorithms are therefore making a people manager's job easy and increasing LMS utilization. So are actual LMS platforms

– for example, managers need only a few clicks today to assign trainings to employees, track their progress, and send reminders for overdue courses.

There are more use-cases for data analysis in LMSs [6] – algorithms are predicted to crunch data on correlations between LMS usage by learners and their future performance – implying that one day in the future, we could arrive at learning strategies that are data-backed to increase on-the-job performance, increasing productivity, and increasing GDP growth. Data rich LMSs output other actionable metrics [7] for the people manager – like rapid page exists could indicate to the manager that associates are skipping pages without imbibing learnings, learner satisfaction ratings could inform managers on the usefulness of a course, and learner proficiency testing through simulations and branching of scenarios could aid the manager in choosing best courses.

Compensation: HR professionals spend valuable time setting pay brackets, only to further spend hours negotiating pay with each employee who joins the firm. Pay transparency recently came under renewed focus after the New York city pay transparency laws were implemented on Nov 01, 2022 [8] requiring all job posts to include a compensation range.

Payroll technology companies have taken a head-start to address this problem. ADP uses anonymized data from the 30+ million payrolls that ADP pays [9], processes it using big data processing models, and outputs a compensation benchmark along 3 parameters – job title, location, and industry. For example, it outputs how much an administrative assistant would make in Chicago in the Logistics industry and slices the data by underlying components (base, bonus, commission, overtime, tips) and pay type (salaried vs. hourly).

In this example, given data analytics is used on 30+ million actual real payrolls, the compensation range output is likely to be more accurate than self-reported compensation insights, thereby freeing-up time that an HR manager may have used to research compensation and increasing their productivity.

Data analysis can also churn out employee insights on workforce change – turnover, tenure and retention. Algorithms even today are alerting HR managers to those employees who are most at-risk of quitting the firm by analyzing compensation (lower compensated employees are at higher turnover risk), tenure (lower is at higher risk), commuting time to office (higher is at higher risk) and overall retention rates at the company and industry. These can help HR managers increase compensation, offer flexible PTO or work from home benefits whenever algorithms highlight a high risk of turnover.

Performance Management: Evaluating employees' performance against goals and tasks assigned to them can be an expensive and subjective exercise and many large organizations struggle to establish structured mechanisms around providing fair, relevant feedback and associated development opportunities. A Robert Half International study [10] (2014) found that people managers spend on average 17% of working hours supervising poorly performing employees. Research indicates that big data models can reduce some of the complexity, costs and biases in employee evaluations. Aktepe and Ersoz [11] suggest use of algorithms to group performers into different groups based on factors such as performance levels, background, and job satisfaction. This could allow HR and direct managers to identify employee groups that need more training, development and encouragement resulting in increased employee satisfaction and retention rates. Newer applications of big data models have extended further into 1) determining levels of proficiency and specialization among employees through ordinal regression clustering as demonstrated by Horesh et al [12], 2) identification and elimination of biases, prejudice, stereotypes and other forms of subjectivity commonly found in human-led evaluation mechanisms through the use of natural language processing as demonstrated by Abed and El-Halees [13], 3) using classification algorithms for employee profiling in order to create custom incentives and benefits. There has also been significant research to understand how ML models can be used to preemptively identify employees that are displaying signs of low performance or satisfaction or are at greater risk of attrition by tracking their engagement and utilization levels (Sharma and Goyal

[14]). Organizations can use big data sets around working hours, promotion frequency and levels, tenure, compensation levels and even employees' personal data around age, gender, marital status, geographical location etc. to develop causal models predicting attrition and assigning value to employees. However, these strategies come with its own data privacy risks that are discussed at a later stage in this paper. Aside from privacy issues, research by Deloitte also shows that employees tend to be demotivated by organizations that monitor their online activity and usage. Moreover, low performers tend to prefer human agents or evaluators so they can explain the specifics of their circumstances and therefore it is vital to use these methods in conjunction with HR managers that are well trained to use and interpret these tools.

3. Challenges and Limitations

Big data ML models and algorithms can be overused, misused and misinterpreted in HCM decision-making processes. Moreover, employees can be demotivated by the lack of human empathy, interaction and the ability to control outcomes. Below, we outline some of the key challenges associated with 1) over reliance on ML models, 2) inadequately setup, trained or tested algorithms, 3) lack of training available to supervisors and HR managers who are the direct users of the models and 4) weak or absent mechanisms around accountability, audits and use of data.

Availability of large, clean datasets for training models: Few organizations would have access to a large enough training dataset needed to create an accurate and useful model. Not to mention, it is important to clean the dataset to avoid introducing biases or incorrect causalities in the model. For example, for a model that analyses job applicants' resumes it is important to scrub the training dataset so it does not contain names, location, gender or birthdates in order to avoid ethnic, socio-economic, gender or age bias. In 2015, retail giant Amazon discovered that their ML based model built to analyze job applicant resumes to select interview candidates was biased against women applicants, especially for technical jobs that have historically been male dominated [15]. In addition, ML-models can introduce new biases to decision-making that may not have existed within human decision-making. A Bavarian Broadcasting study [16], had an actor answer certain questions on camera keeping certain body language traits constant (such as expressions and voice modulation). An ML model trained on facial datasets then scored the actor on different traits like sincerity, extraversion, and amicability. The study found that results varied meaningfully when she wore glasses, a headscarf, or changed her background.

Unclear lines of accountability: When humans delegate tasks or decision-making to automated systems, there is a blurring of direct lines of responsibility and accountability. This can largely be broken down into 1) Distributed accountability: A human agent could be part of a complex organizational structure and therefore only partially contribute to outcomes and may or may not be trained or qualified to use the ML model that management requires them to leverage to make hiring decisions. When poor or biased hiring decisions are made, who do we hold responsible?; 2) Ignorance of automation logic: If human agents don't fully comprehend why automated systems make the decisions that they do, can they be held responsible for choices made by or suggested by the system?

Loss of jobs: Data processing AI/ML models can be most effective when used to perform low value-add tasks requiring manual, repetitive actions to create human bandwidth for more 'think-big', strategic tasks. For example, instead of an HR manager manually reading resumes or assigning training modules to new hires they could spend their time doing greater value-adding work such as develop better feedback and mentorship mechanisms or create stronger recruiting pipelines through engagements with universities and schools. This shift could result in loss of jobs that are manual or repetitive in nature and lack any specialization. However, at the same time it would create a need for more qualified and better trained HR professionals and can be viewed as an opportunity for HR managers to upskill.

Data Privacy – There is growing concern among employees that their work devices and in some cases even private devices that they may use to access work email or messaging platforms may be heavily monitored by employees. A qualitative study (Teebken and Hess [17]) found that employees had concerns around 1) collection of personal data that would be stored in databases for uncertain periods of time; 2) personal data or recorded conversations being used for unauthorized purposes or by unauthorized parties; 3) data collection errors that could result in misinterpretation (for example, offline/online status); 4) employees' inability to effectively control the collection of their personal data; and 5) lack of awareness or information around the organization's privacy policies and how it plans to use employee data. While this data can provide significant insights into employee utilization, engagement and satisfaction metrics and preemptively answer questions around training or development needs, attrition etc. there is risk around misuse and misinterpretation. There is also risk around incentivizing the wrong employee behaviors. A 2021 Gartner study [18] found that employees are twice as more likely to pretend to be engaged with work related tasks when their organizations monitor them using AI.

4. Discussion

The costs of poor recruiting, employee development and compensation decisions materialize in tangible and intangible ways including but not limited to loss of productivity (for both employees and their managers), bad customer experiences, damage to an organization's reputation, re-recruiting and re-training costs due to attrition and firing actions, legal costs etc.

In 2021, the US recorded a total GDP of $23 trillion and with a population of 332 million, this gives us an annual per capita GDP of $69,298 and a daily per capita GDP of $190. There are a total of 148 million employees in the US and SHRM's Human Capital Benchmarking Survey estimates that, on average, the US sees an annual employee turnover rate of 19% which means every year on average 28 million individuals [19] quit their job. The same study estimates that it takes on average 42 days to fill a vacant role which gives us 1.2 billion working days of lost productivity from unfilled positions. Applying the daily per Capita GDP we calculated earlier, we estimate that the US loses $224 billion or 0.97% of total GDP annually due to attrition and weak HCM mechanisms and practices. Our sensitivity analysis in Fig 2 shows that by reducing attrition rates by just 300bps and filling positions just 3 days sooner on average, we can save 21bps or $49 billion of lost GDP (productivity).

% GDP lost		No. of days positions will remain unfilled					
	0.97%	39	40	41	**42**	43	44
	16%	0.76%	0.78%	0.80%	0.82%	0.84%	0.86%
Annual Attrition Rate	17%	0.81%	0.83%	0.85%	0.87%	0.89%	0.91%
	18%	0.86%	0.88%	0.90%	0.92%	0.94%	0.97%
	19%	0.90%	0.93%	0.95%	**0.97%**	1.00%	1.02%
	20%	0.95%	0.98%	1.00%	1.02%	1.05%	1.07%
	21%	1.00%	1.02%	1.05%	1.08%	1.10%	1.13%
	22%	1.05%	1.07%	1.10%	1.13%	1.15%	1.18%

Fig. 2: Sensitivity analysis of US GDP (productivity) lost

Separately, the cost of hiring an employee in the US is estimated to average $4,129 [20]. If we assume that 94.5% of the 28 million US individuals that quit their job annually end up being hired by another organization (assuming a 5.5% unemployment rate), that puts annual hiring costs in the US at $109 billion, a cost that can be minimized through robust hiring, compensating, performance management, training and development practices. This represents resources that if invested in the better management of human capital, could reduce significant churn and loss of productivity.

While we recognize the benefits of adopting big data processing models in HCM strategy and processes, we must also address some of the key limitations laid out in this paper.

We can eliminate hiring bias introduced within algorithms by using clean datasets and conducting adequate testing of our models. One way of doing this could be to intentionally introduce bias in the system and record how the model responds to it or using larger quantities of human decision-making data so that the model has increased data points. Models should continue to be audited for errors and bias after they are launched. Organizations must also invest in training and upskilling HR managers as well as establishing the right mechanisms for accountability.

Some of the data and privacy concerns discussed earlier can be addressed by aggregating and anonymizing data, receiving explicit consent from employees around data use and establishing transparency and strict controls around storage and utilization of data for specific, predetermined purposes only. Adequate training will also play a strong role in ensuring that data is not misinterpreted or misused and that HR managers are well equipped to draw the right conclusions and where needed, treat the model's output with cautious doubt so as to prevent over-reliance. Studies show that many tracking and monitoring algorithms lack accuracy as they are unable to track offline activities such as reading, ideating, mentoring etc.

5. CONCLUSION

In this paper we have laid out the applications of big data models in HCM workstreams as well as the costs, to not only organizations but the country as a whole, of weak, error-prone HR mechanisms that rely solely on human decision-making. We have established that both human agents and ML models alone pose limitations and can be expensive, biased, manual, time consuming and sub-optimal on one end and pose data privacy risks, accountability and other concerns on the other end. Therefore, we must strive to establish mechanisms where both coexist. Technology is currently not positioned to replace human HR managers but must be used as a lever by them to remain competitive and make better decisions. It must enable the HR manager to stay focused on larger strategic tasks by dissecting actionable insights from the analysis of big data. We can break down different HR tasks and processes and leverage big data ML models to improve efficiency, reduce biases and human error-prone decision-making to improve overall outcomes and as a result augment the productivity of HR managers, supervisors, organizations, and the country as a whole.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] Dirk Buyens, Ans De Vos (2006); Perceptions of the value of the HR function.

[2] Mark Tomassen (2016); Exploring the Black Box of Machine Learning In Human Resource Management

[3] Michael Chui, James Manyika, and Mehdi Miremadi (2015); Four fundamentals of workplace automation

[4] Carl Benedikt Frey and Michael A. Osborne (2013); The Future of Employment: How Susceptible Are Jobs To Computerisation?

[5] Nancy S. Charney (2022); Onboarding: The Key to Improving Employee Retention and Productivity

[6] Christopher Pappas (2016); Types Of Big Data To Extract From Your LMS And How To Use It

[7] Christopher Pappas (2016); Types Of eLearning Analytics To Obtain From Google And Your LMS

[8] Salary Transparency in Job Advertisements, www.nyc.gov/site/cchr/media/pay-transparency.page

[9] ADP; Analytics and Benchmarking, Powered by ADP® DataCloud

[10] Robert Half International (2012); Managers Spend Nearly One Day a Week Managing Poor Performers

[11] Adnan Aktepe, Suleyman Ersoz (2012); A Quantitative Performance Evaluation Model Based On A Job Satisfaction-Performance Matrix And Application In A Manufacturing Company

[12] Raya Horesh, Kush R. Varshney, and Jinfeng Y (2016); Information Retrieval, Fusion, Completion, and Clustering for Employee Expertise Estimation

[13] Amani A. Abed, Alaa M. El-Halees (2017); Detecting Subjectivity in Staff Perfomance Appraisals by Using Text Mining: Teachers Appraisals of Palestinian Government Case Study

[14] Sharma and Goyal (2015); An Application of data mining to improve personnel performance evaluation in higher education sector in India

[15] Jeffrey Dastin (2018); Amazon scraps secret AI recruiting tool that showed bias against women

[16] Hugo Britt (2022); Good AI, Bad AI, and How To Avoid Hiring Bias In Machine Learning

[17] Mena Angela Teebken, Thomas Hess (2021); Privacy in a Digitized Workplace: Towards an Understanding of Employee Privacy Concerns

[18] Gartner (2021); Gartner HR Research Shows Organizations Are Eroding Employee Performance and Well-Being with Virtualized Office-Centric Design

[19] Apollo Technical (2022); The Cost Of A Bad Hire And Red Flags To Avoid

[20] Harvard Technology and Operations Management (2018); Machine Learning for Recruiting at Amazon – Challenges and OpportunitiesRuss J.C. (1990) Image Processing. In: Computer-Assisted Microscopy. Springer, Boston, MA. https://doi.org/10.1007/978-1-4613-0563-7_3

International Journal for Modern Trends in Science and Technology IJMTST