Artificial Intelligence in Program Evaluation: Insights and Applications

Sidney Shapiro
Dhillon School of Business, University of Lethbridge, Lethbridge, Alberta, Canada

Vinh Lam
University of Lethbridge, Lethbridge, Alberta, Canada

Abstract: *The practice note outlines six approaches to integrating artificial intelligence (AI) and machine learning (ML) into program evaluation, enhancing traditional methods with data-driven insights and improved efficiency. These approaches aim to address the growing need for evaluators to analyze complex datasets accurately while reducing manual effort. They include identifying patterns in data to uncover trends and outliers, using predictive models for forecasting outcomes, pinpointing areas for improvement by analyzing performance, simplifying data interpretation through visualizations, automating data analysis for efficiency, and leveraging dashboards for real-time monitoring and decision-making. The note highlights that while it does not offer guidance for evaluators with limited technical expertise, it provides a framework for integrating AI into evaluation practices. Successful implementation requires understanding stakeholder needs, fostering client engagement, ensuring tool usability, and maintaining effective communication. While AI can enhance evaluation quality and innovation, ethical considerations and biases in AI algorithms must be carefully addressed. These AI-powered techniques can enable more robust, evidence-based decision-making, supporting positive social impact.*

Keywords: *artificial intelligence*, *machine learning*, *program evaluation*, *data analysis*, *predictive modeling*, *data visualization*, *dashboards*, *evaluation framework*

Résumé : *La présente note de pratique décrit six façons d'incorporer l'intelligence artificielle (IA) et l'apprentissage automatique à l'évaluation de programmes, pour améliorer des méthodes traditionnelles avec de nouvelles informations fondées sur des données et une plus grande efficience. Ces approches visent à trouver une solution aux besoins croissants, pour les évaluateurs/évaluatrices, d'analyser des ensembles de données complexes avec exactitude, tout en réduisant les efforts manuels. Il est notamment question de reconnaître des motifs dans les données, pour noter des tendances et des valeurs aberrantes, en utilisant des modèles prédictifs*

Corresponding author: Sidney Shapiro, Dhillon School of Business, University of Lethbridge, 4401 University Drive, Lethbridge, Alberta, Canada, T1K 3M4; sidney.shapiro@uleth.ca

© 2024 *Canadian Journal of Program Evaluation / La Revue canadienne d'évaluation de programme*
39.2 (December / décembre), 382–391 doi:10.3138/cjpe-2024-0027

pour prévoir les résultats, en déterminant les occasions d'amélioration en analysant le rendement, en simplifiant l'interprétation des données par l'intermédiaire de visualisations, en automatisant l'analyse des données à des fins d'efficience et en tirant parti de tableaux de bord pour un suivi en temps réel et pour la prise de décision. La note souligne qu'elle ne permet pas d'orienter les évaluateurs/évaluatrices ayant une expertise technique limitée, mais offre un cadre pour l'intégration de l'IA dans les pratiques d'évaluation. Une mise en œuvre réussie exige de comprendre les besoins des intervenants, d'encourager la participation de la clientèle, de veiller à l'utilité de l'outil et de maintenir des communications efficaces. L'IA peut améliorer la qualité de l'évaluation et encourager l'innovation, mais il faut soigneusement tenir compte de considérations éthiques et des biais présents dans les algorithmes d'IA. Ces techniques fondées sur l'IA peuvent mener à une prise de décision plus robuste et fondée sur des faits probables, à l'appui d'un impact social positif.

Mots clés : *analyse des données*, *apprentissage automatique*, *cadre d'évaluation*, *évaluation de programmes*, *intelligence artificielle*, *modélisation prévisionnelle*, *tableau de bord*, *visualisation des données*

INTRODUCTION

In this practice note, we propose six ways to incorporate artificial intelligence (AI) and machine learning (ML) technologies into program evaluation practice, transforming traditional approaches into more data-driven, efficient, and insightful processes. As program evaluators face increasing demands to analyze complex datasets and derive meaningful insights, integrating AI and ML can enhance accuracy and reduce manual workload. These suggestions aim to highlight potential applications of AI for improving evaluation quality, enabling new insights, and facilitating more effective decision-making. The proposed recommendations cover a range of practices, from identifying patterns in data to automating routine tasks, ensuring a comprehensive framework for incorporating AI and ML into program evaluation. This practice note does not guide evaluators on leveraging AI; instead, it proposes ways evaluators could integrate AI into their practice.

To incorporate AI into program evaluation, evaluators can take several key steps. First, they can leverage AI to identify patterns in large datasets, revealing trends and outliers that manual analysis might miss. This can illuminate program dynamics, such as whether a health program improves participants' health over time. Predictive models can forecast future outcomes, aiding in assessing long-term impacts. This helps evaluators recommend directions for program improvement. AI can also be used to discover areas for improvement by pinpointing where a program excels or underperforms. These insights guide recommendations for program adjustments, such as revising challenging training modules.

AI-generated visualizations, including charts and dashboards, simplify data interpretation and facilitate stakeholder communication. Data analysis can be

automated with AI to reduce manual tasks and enhance efficiency, allowing evaluators to focus on high-level interpretation. Intuitive dashboards enable real-time monitoring, offering user-friendly interfaces to track program performance and adjust views as needed. Lastly, technical support and training provided by AI tools can be leveraged to ensure that even non-technical evaluators can effectively incorporate AI into their evaluations, accessing tutorials and customer support for guidance.

PRACTICE DIMENSIONS AND LITERATURE REVIEW

ML and AI are pivotal in program evaluation across various sectors. Machine learning, a subset of AI, involves using statistical techniques to enable machines to enhance their performance through experience (Izdihar et al., 2021). It is a significant aspect of AI that allows software applications to improve forecasting outcomes without explicit programming (Lohit, Mujahid, & Sai, 2022). AI and machine learning have significantly advanced in various industries, including program evaluation, enabling computers to learn and improve from data without explicit programming (Mueller, Kinoshita, Peebles, Graber, & Lee, 2022).

In the service and tourism sectors, machine learning is utilized for continuous improvement and handling of multi-dimensional data (Bangare et al., 2022), enhancing the precision of software applications in predicting outcomes (Felix & Lee, 2017). Machine learning techniques are applied in human resource management in various companies, illustrating how individuals can teach machines to perform tasks (Thu, Nguyen, & Nguyen, 2022). Machine learning models are increasingly integrated into multiple industries, including the automotive sector, to introduce intelligence in vehicular networks (Seth, Guleria, & Panda, 2022). In education, machine learning models are being developed to facilitate the smooth integration of AI in various fields of interest, showcasing the potential for enhancing learning experiences through technology (Lee & Cho, 2021). The integrations of AI in these and other sectors underscore the adaptability of machine learning algorithms in improving decision-making processes in different domains.

METHODOLOGICAL APPROACH FOR PRACTICE ANALYSIS

AI offers diverse tools that enable evaluators to conduct detailed analyses, predict outcomes, identify areas for improvement, and streamline routine tasks. By harnessing AI's ability to identify patterns in large datasets, evaluators can reveal trends and correlations that would otherwise go unnoticed, providing critical insights into program effectiveness and underlying dynamics. Predictive AI models can forecast outcomes based on current data, aiding in evaluating long-term program impacts and helping design targeted interventions to address specific challenges. AI-generated visualizations and intuitive dashboards are examined as methods to simplify data interpretation and facilitate clear communication with

stakeholders. Automation for data analysis reduces manual workload, while technical support and training ensure that even non-technical evaluators can effectively incorporate AI into their workflow. This comprehensive methodological approach aims to enhance the efficiency and accuracy of program evaluation through advanced AI techniques and tools.

The six approaches to leveraging AI are identifying patterns in data, predicting future outcomes, discovering areas for improvement, providing visualizations, automating data analysis, and offering intuitive dashboards. These steps create a framework for program evaluators to harness AI's capabilities, fostering more insightful and efficient evaluations and reporting.

Approach 1: Identify patterns in data

AI offers an unparalleled ability to analyze large datasets and uncover patterns that would be difficult to discern through manual methods (Provost & Fawcett, 2013). This capability is particularly crucial in program evaluation, where the volume and complexity of data can overwhelm traditional analytical techniques. By leveraging AI, evaluators can discover trends, correlations, and outliers, leading to a deeper understanding of program dynamics and insights into the factors contributing to success or failure.

Take, for example, a health program aimed at reducing diabetes among participants. By applying AI algorithms to patient data, evaluators can identify patterns in health outcomes over time, such as trends in blood glucose levels, medication adherence, and frequency of hospital visits. This analysis can reveal whether the program effectively manages diabetes and can pinpoint specific factors contributing to positive or negative outcomes.

The patterns uncovered by AI can guide further improvements, allowing evaluators to recommend changes that might boost program effectiveness, such as additional patient support or modifications to treatment plans. AI's pattern recognition capability can facilitate a data-driven approach to understanding and enhancing program performance.

Approach 2: Predict future outcomes

Predictive AI models can forecast future outcomes of business programs and initiatives based on existing data, providing valuable insights for evaluators. This feature helps assess long-term program impacts and anticipate upcoming trends that could influence program success (T. Davenport & Harris, 2017). For instance, in evaluating the effectiveness of a new employee training program, a business analyst could use predictive modelling to forecast the program's impact on employee performance, retention, and productivity over time.

By accurately forecasting the outcomes of the training program, the analyst can better understand the potential long-term benefits and return on investment. This could involve predicting how the program will affect key performance indicators, such as sales figures, customer satisfaction scores, or operational efficiency. With these data-driven predictions, the business can

make more informed decisions about continuing, expanding, or modifying the training program to maximize its impact.

Predictive analytics also enables evaluators to anticipate future market conditions and industry trends that could affect the success of business programs. For example, an analyst could use AI to forecast changes in customer preferences, shifts in competitor strategies, or the emergence of new technologies that could disrupt the company's operations. Armed with these insights, the business can proactively adapt its programs and initiatives to stay ahead of the curve and capitalize on upcoming opportunities.

Approach 3: Discover areas for improvement

AI can pinpoint areas where a program is underperforming or excelling. This insight allows program evaluators to make informed recommendations for improvement. For example, if a training program has high dropout rates, AI can help identify the challenging modules or other factors contributing to this trend, guiding curriculum adjustments or enhanced support.

By leveraging AI's analytical capabilities, program evaluators can better understand the factors driving program outcomes. For instance, when analyzing a training program with high dropout rates, AI can help pinpoint specific modules or aspects contributing to this trend.

AI may reveal that specific modules within the training program are particularly challenging for participants, leading to increased dropout rates. Additionally, AI can uncover other contributing factors, such as inadequate support mechanisms or outdated instructional materials. Armed with this information, program evaluators can formulate informed recommendations for improvement, such as revising the curriculum to make challenging modules more accessible or enhancing support services to address participant needs.

Approach 4: Provide visualizations

AI-generated visualizations are powerful tools to simplify data interpretation and enhance communication of findings to stakeholders. By leveraging AI's capabilities, evaluators can transform complex datasets into precise, actionable visual representations. Visual tools such as tables, charts, and graphs enable evaluators to communicate key insights and trends, facilitating informed decision-making effectively.

Graphical representations of data, such as line charts and charts with change indicators, can help stakeholders quickly grasp the program's impact. AI-powered interactive dashboards can provide real-time updates and customizable views, allowing stakeholders to explore data and uncover insights relevant to their interests. By generating visualizations, the explainability, transparency, and accessibility of evaluation results are improved so stakeholders can easily understand and engage with the findings. This capability is valuable in communicating findings and demonstrating impact to drive meaningful action and foster stakeholder buy-in.

Approach 5: Automate data analysis

Automating data analysis tasks through AI reduces manual effort, speeds up the evaluation process, and minimizes the risk of human error (T. H. Davenport & Kirby, 2016). AI systems can handle data entry, sorting, cleaning, and fundamental analysis, allowing evaluators to focus on higher-level interpretation and decision-making.

For example, in evaluating the performance of a marketing campaign, an analyst could use AI-powered automation to streamline the data analysis process. The AI system could automatically ingest and organize data from various sources, such as website analytics, social media metrics, and customer feedback. It could then perform initial analysis, such as calculating key performance indicators, identifying trends, and detecting any anomalies in the data.

With these automated tasks taken care of, the analyst can devote more time and attention to interpreting the insights, understanding the underlying drivers of campaign performance, and developing actionable recommendations to optimize future marketing efforts. This enhanced efficiency enables evaluators to manage complex, large-scale program evaluations more effectively, generating timely and reliable insights to support strategic decision-making.

Approach 6: Leveraging AI to build intuitive dashboards

Intuitive dashboards play a critical role in program evaluation by providing real-time monitoring and enabling evaluators to track program performance dynamically (Alborhamy, 2020). With the help of AI, evaluators can build intuitive dashboards that offer a user-friendly interface for accessing and analyzing data. These dashboards allow evaluators to monitor program metrics as they unfold, facilitating timely decision-making and adjustments (Solanki, Jain, & Jadiga, 2024).

AI-powered dashboards can display key performance indicators and metrics related to program outcomes, such as participant engagement, completion rates, and satisfaction levels. By visualizing this data in real time, evaluators gain valuable insights into program effectiveness and promptly identify areas for improvement. Moreover, intuitive features such as interactive filters and customizable views empower evaluators to adapt their analysis based on changing needs and priorities.

Technical support and training

AI-powered evaluation tools often have comprehensive technical support and training resources, making them accessible to non-technical program evaluators (Bulathwela, Pérez-Ortiz, Holloway, Cukurova, & Shawe-Taylor, 2024). This support ensures that evaluators can leverage the benefits of advanced analytics without requiring deep technical expertise. Access to detailed tutorials, user guides, and customer support services helps evaluators learn how to apply AI-driven tools in their program assessments effectively. For instance, many AI vendors provide interactive online tutorials that guide users through step-by-step

procedures for data preparation, model building, and results interpretation. Some vendors offer live webinars and training workshops, both in-person and virtually (Cesaro, 2021), to help evaluators develop the necessary skills and confidence to integrate AI into their evaluation workflows (Labarthe, Luengo, & Bouchet, 2018).

Examples and suggestions for the six areas discussed include AI software for specific applications without endorsing any particular option. Tools for identifying data patterns, including RapidMiner and IBM SPSS, can help evaluators identify trends and outliers in datasets. For predicting future outcomes, software such as TensorFlow and Microsoft Azure Machine Learning provides robust platforms for building predictive models. To discover areas for improvement, evaluators can use Microsoft Excel, Tableau, or Power BI to pinpoint areas where programs excel or underperform. Tools like Tableau and D3.js are excellent for generating visualizations. For automating data analysis, Alteryx and KNIME offer capabilities for automating routine data analysis tasks. Lastly, platforms like Qlik Sense and Looker enable the creation of user-friendly dashboards for real-time monitoring.

INSIGHTS GAINED AND FACTORS INFLUENCING OUTCOMES

Many tools are available for data analysis in program evaluation. Despite the variety of competing tools on the market, the primary focus should be on establishing effective connections with clients and stakeholders. Similar dashboards exploring program data can be created using standard tools (Islam, 2020), including Microsoft Excel, Microsoft PowerBI, Tableau, and many other software packages.

By prioritizing client engagement and understanding their specific needs, program evaluators can ensure that the chosen tools align with the objectives of the evaluation. This approach enhances trust and collaboration and ensures that the evaluation process is relevant and impactful. Leveraging new and advanced tools, particularly those powered by AI and machine learning, such as the Data Analyst module from ChatGPT (Hassani & Silva, 2023; Rasheed et al., 2024), can significantly enhance the quality and depth of program data analysis. These tools provide powerful capabilities for identifying trends, predicting outcomes, discovering areas for improvement, and visualizing data, enriching the insights derived from evaluations and supporting more informed decision-making.

The success or failure of program evaluation efforts often hinges on several factors, with client engagement paramount. Effective communication and collaboration with clients facilitate a clear understanding of evaluation goals and requirements, ensuring that the chosen tools meet the desired outcomes. Additionally, factors such as the usability and accessibility of tools play a crucial role in driving successful program evaluations. Intuitive and user-friendly tools enable evaluators to navigate complex datasets more efficiently, leading to more

insightful analyses and actionable findings. Ultimately, program evaluators can unlock valuable insights and drive meaningful improvements in program data analysis and evaluation outcomes by prioritizing client engagement and leveraging advanced tools.

CONCLUSIONS AND IMPLICATIONS FOR PROGRAM EVALUATION

Integrating AI and advanced data analytics tools into program evaluation practices offers significant implications for evaluation practice. AI-driven methodologies can enable program evaluators to enhance their evaluations' efficiency, accuracy, and depth; leveraging AI's capabilities to identify patterns, predict outcomes, and provide intuitive visualizations, evaluators can uncover nuanced, previously inaccessible insights through traditional methods.

AI in program evaluation underscores the importance of continuous innovation and adaptation in evaluation practices. As technology evolves and new tools emerge, evaluators must remain agile and responsive to changing needs and opportunities. Embracing AI enhances the quality of evaluations and fosters a culture of innovation within the evaluation community. Moreover, the widespread adoption of AI in program evaluation has implications for capacity-building efforts, requiring evaluators to develop new skills and competencies to leverage these advanced tools effectively.

Integrating AI into program evaluation promises to drive evidence-based decision-making and improve the effectiveness of social programs, policies, and interventions. By leveraging AI to analyze complex datasets and derive actionable insights, evaluators can better address societal challenges and contribute to positive social change. However, it is essential to approach the integration of AI in program evaluation with caution, considering ethical considerations, privacy concerns, and potential biases inherent in AI algorithms. Ethical considerations include transparency in AI decision-making, protecting sensitive information, and adhering to data protection regulations. To address these concerns, evaluators should maintain transparency by documenting and reporting their use of AI methodologies.

REFERENCES

Alborhamy, Y. (2020). *How do data dashboards affect evaluation use in a knowledge network? A study of stakeholder perspectives in the Centre for Research on Educational and Community Services (CRECS)* [Master's thesis]. University of Ottawa. https://doi.org/10.20381/ruor-25487

Bangare, M. L., Bangare, P. M., Ramirez-Asis, E., Jamanca-Anaya, R., Phoemchalard, C., & Bhat, D. A. R. (2022). Role of machine learning in improving tourism and education sector. *Materials Today: Proceedings*, *51*, 2457–2461. https://doi.org/10.1016/j.matpr.2021.11.615

Bulathwela, S., Pérez-Ortiz, M., Holloway, C., Cukurova, M., & Shawe-Taylor, J. (2024). Artificial intelligence alone will not democratise education: On educational

inequality, techno-solutionism and inclusive tools. *Sustainability*, *16*(2), 781. https://doi.org/10.3390/su16020781

Cesaro, M. (2021, November). *The future of learning* [paper presentation]. Abu Dhabi International Petroleum Exhibition & Conference, Abu Dhabi, United Arab Emirates. https://doi.org/10.2118/207989-MS.

Davenport, T., & Harris, J. (2017). Competing on analytics: Updated, with a new introduction: The new science of winning. Harvard Business Press.

Davenport, T. H., & Kirby, J. (2016). Only humans need apply: Winners and losers in the age of smart machines. Harper Business.

Felix, E. A., & Lee, S. P. (2017). Integrated approach to software defect prediction. *IEEE Access*, *5*, 21524–21547. https://doi.org/10.1109/access.2017.2759180

Hassani, H., & Silva, E. S. (2023). The role of ChatGPT in data science: How AI-assisted conversational interfaces are revolutionizing the field. *Big Data and Cognitive Computing*, *7*(2), 62. https://doi.org/10.3390/bdcc7020062

Islam, M. (2020). Data analysis: Types, process, methods, techniques and tools. *International Journal on Data Science and Technology*, *6*(1), 10–15. https://doi.org/10.11648/j.ijdst.20200601.12

Izdihar, K., Karim, M. K. A., Aresli, N. N., Radzi, S. F. M., Sabarudin, A., Yunus, M. M., Rahman, M. A. A., & Shamsul, S. (2021). *Detection of novel coronavirus from chest X-Ray radiograph images via automated machine learning and CAD4COVID* [Paper presentation]. 2021 International Congress of Advanced Technology and Engineering (ICOTEN), Taiz, Yemen. https://doi.org/10.1109/icoten52080.2021.9493542

Labarthe, H., Luengo, V., & Bouchet, F. (2018). *Analyzing the relationships between learning analytics, educational data mining and AI for education.*

Lee, Y., & Cho, J. (2021). Development of an artificial intelligence education model of classification techniques for non-computer majors. *Joiv International Journal on Informatics Visualization*, 5(2), 113–119. https://www.joiv.org/index.php/joiv/article/view/552

Lohit, V. S., Mujahid, M. M., & Sai, G. K. (2022). Use of machine learning for continuous improvement and handling multi-dimensional data in service sector. *Computational Intelligence and Machine Learning*, 3(2), 30–35. https://doi.org/10.36647/ciml/03.02.a006

Mueller, B., Kinoshita, T., Peebles, A. T., Graber, M. A., & Lee, S. (2022). Artificial intelligence and machine learning in emergency medicine: A narrative review. *Acute Medicine & Surgery,* 9(1), e740. https://doi.org/10.1002/ams2.740

Provost, F., & Fawcett, T. (2013). Data science for business: What you need to know about data mining and data-analytic thinking. O'Reilly Media.

Rasheed, Z., Waseem, M., Ahmad, A., Kemell, K.-K., Xiaofeng, W., Duc, A. N., & Abrahamsson, P. (2024). Can large language models serve as data analysts? A multi-agent assisted approach for qualitative data analysis. https://doi.org/10.48550/arXiv.2402.01386.

Seth, I., Guleria, K., & Panda, S. N. (2022). Introducing intelligence in vehicular ad hoc networks using machine learning algorithms. *ECS Transactions*, 107(1), 8395. https://doi.org/10.1149/10701.8395ecst

Solanki, A., Jain, K., & Jadiga, S. (2024). Building a data-driven culture: Empowering organizations with business intelligence. *International Journal of Computer Trends and Technology*, *72*(2), 46–55. https://doi.org/10.14445/22312803/ijctt-v72i2p109

Thu, T. N. T., Nguyen, B.-L., & Nguyen, T. (2022). Framework of using machine learning techniques in human resource management in Vietnamese companies. In Vu Dinh Khoa, Shivani Agarwal, Gloria Jeanette Rincon Aponte, Nguyen Thi Hong Nga, Vijender Kumar Solanki, Ewa Ziemba (eds), *Proceedings of the 2021 International Conference on Research in Management & Technovation*, ACSIS 28, 295–300. https://doi.org/10.15439/2021km8

CONTRIBUTOR INFORMATION

Sidney Shapiro is an assistant professor of data analytics at the University of Lethbridge, researching the areas of program evaluation data science, business analytics, and machine learning. His expertise spans research, teaching, and consulting, focusing on developing data-driven solutions and automating data tools to enhance business efficiency and effectiveness. ORCID: https://orcid.org/0000-0003-4193-1154. Email: sidney.shapiro@uleth.ca

Vinh Lam, BCom, is a data scientist with a background in business analytics, holding postgraduate and bachelor's degrees. He develops database APIs, conducts data queries, and utilizes machine learning for predictive analytics at Sofvie Inc. Vinh utilizes Python, R, SQL, and Excel and has presented on AI applications. Email: dangvinh1212@gmail.com