Journal of Information Systems Engineering and Management
2025, 10(2)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

A Review on Causal Inference Models Combining Bayesian Networks with Deep Learning in Environmental Health Research

Susymary Johnson[1*], Deepalakshmi Perumalsamy[2]
[1]*Affiliation Research Scholar, Department of Computer Applications*
Kalasalingam Academy of Research and Education
Krishnankoil, Virudhunagar District, Tamil Nadu, 626126, India. Email: susymaryj@gmail.com
[2]*Affiliation Professor, Department of Computer Science and Engineering*
Kalasalingam Academy of Research and Education
Krishnankoil, Virudhunagar District, Tamil Nadu, 626126, India. Email: deepa.kumar@klu.ac.in
[*]*Corresponding Author: susymaryj@gmail.com*

ARTICLE INFO

Received: 22 Oct 2024

Revised: 28 Dec 2024

Accepted: 15 Jan 2025

ABSTRACT

This literature review explores the application of causal inference models, specifically Bayesian networks (BN), integrated with deep learning (DL) for feature selection, in assessing the association between air pollution exposure and disease prevalence among individuals living near industrial areas. Air pollution, especially in industrial zones, has been linked to a range of adverse health outcomes, including respiratory, cardiovascular, and chronic diseases. However, assessing these causal relationships remains a challenge due to the complexity of environmental and health data, as well as the presence of confounding factors. Traditional statistical methods often struggle to account for such complexity, which is where advanced models like Bayesian networks come into play. BNs, as probabilistic graphical models, offer a robust framework for modeling causal relationships, allowing for uncertainty and interaction between variables. Integrating deep learning techniques into Bayesian networks enhances feature selection, enabling the identification of critical factors influencing health outcomes while minimizing the impact of irrelevant or noisy variables. This paper reviews key studies that have employed these integrated models to investigate air pollution's health impact, focusing on the strengths, limitations, and potential of these methodologies. The review also highlights the challenges in modeling complex, real-world environmental health data and proposes directions for future research, including real-time data integration and enhanced computational methods. Ultimately, the combination of Bayesian networks and deep learning represents a promising approach for understanding and addressing the health impacts of air pollution in industrial areas.

Keywords: Casual Inference, Bayesian network, Deep learning, Machine learning, Environmental health.

INTRODUCTION

Air pollution is a major environmental health issue worldwide, with significant implications for public health. Studies have consistently linked air pollution to a range of chronic diseases, including respiratory illnesses, cardiovascular conditions, and cancer. Industrial areas, often located near urban populations, contribute significantly to pollution, particularly through the emission of particulate matter (PM2.5), nitrogen dioxide (NO2), sulfur dioxide (SO2), and other hazardous pollutants. Prolonged exposure to such pollutants increases the risk of disease, particularly in populations living near industrial zones, where environmental exposure and limited healthcare access worsen health disparities.

Despite the recognition of air pollution's harmful effects, establishing a clear causal relationship between exposure and health outcomes is challenging due to the complexity of the data and confounding factors. Traditional epidemiological methods, such as regression models or propensity score matching, have been valuable in identifying associations, but they often struggle to capture the intricate relationships between pollutants, health outcomes, and other influencing factors. These models also make simplifying assumptions that may overlook confounders and the dynamic nature of environmental health processes.

Advanced methodologies, such as causal inference models using Bayesian networks (BN) and deep learning (DL), offer significant promise in addressing these challenges. Causal inference models aim to uncover true causal relationships, allowing researchers to understand how air pollution leads to disease outcomes while accounting for confounders and uncertainty in the data. Bayesian networks are particularly suited for modeling complex causal relationships, integrating prior knowledge, and handling incomplete or noisy data.

Deep learning techniques, particularly neural networks, are powerful tools for feature selection, allowing for the identification of relevant variables from complex datasets. When combined with Bayesian networks, deep learning can enhance feature selection, improving the accuracy and interpretability of causal models. This integrated approach has the potential to advance our understanding of the health impacts of air pollution, especially in populations near industrial areas.

LITERATURE STUDY

A. Bayesian Networks and Deep Learning: A Promising Approach

First, Bayesian networks (BNs) offer a promising solution to many of these challenges. BNs are graphical models that represent probabilistic relationships between variables, making them ideal for capturing complex, interdependent relationships in causal inference. In the context of air pollution and health research, BNs can model the relationships between pollutants, demographic factors, disease outcomes, and other variables. They allow for the integration of prior knowledge (e.g., expert knowledge on the effects of specific pollutants) and the ability to account for uncertainty in the data. Additionally, BNs can handle missing data and provide a framework for modeling both direct and indirect effects, offering a more flexible approach to causal inference.

Deep learning, particularly neural networks, can complement Bayesian networks by enhancing feature selection. In environmental health studies, the number of potential variables is vast, including various air pollutants, health indicators, demographic information, and environmental factors. Traditional feature selection methods can struggle to identify the most relevant variables, especially when dealing with large, high-dimensional datasets. Deep learning techniques, however, can automatically learn patterns in the data and select the most important features, thereby improving model accuracy and reducing overfitting. By integrating deep learning with Bayesian networks, researchers can develop more robust, interpretable, and accurate causal models that are better suited for environmental health research.

B. Objective of the Review

This literature review aims to provide a comprehensive understanding of how causal inference models, specifically Bayesian networks integrated with deep learning for feature selection, are applied to assess the association between air pollution exposure and disease outcomes in populations living near industrial areas. By reviewing key studies and methodologies, this paper will highlight the strengths and limitations of these advanced modeling techniques, explore the challenges and opportunities in the field, and propose directions for future research. Ultimately, this review will demonstrate the potential of combining Bayesian networks and deep learning to improve our understanding of the health impacts of air pollution, particularly in vulnerable populations living in proximity to industrial zones.

Figure 1 represents a suitable diagram for the introduction could be a causal model diagram that visually represents the key components discussed, including the relationships between air pollution, disease outcomes, and various factors such as industrial proximity, demographic data, and environmental exposures. This diagram could also incorporate elements of Bayesian Networks and Deep Learning in feature selection, highlighting the integration of both methods in causal inference.

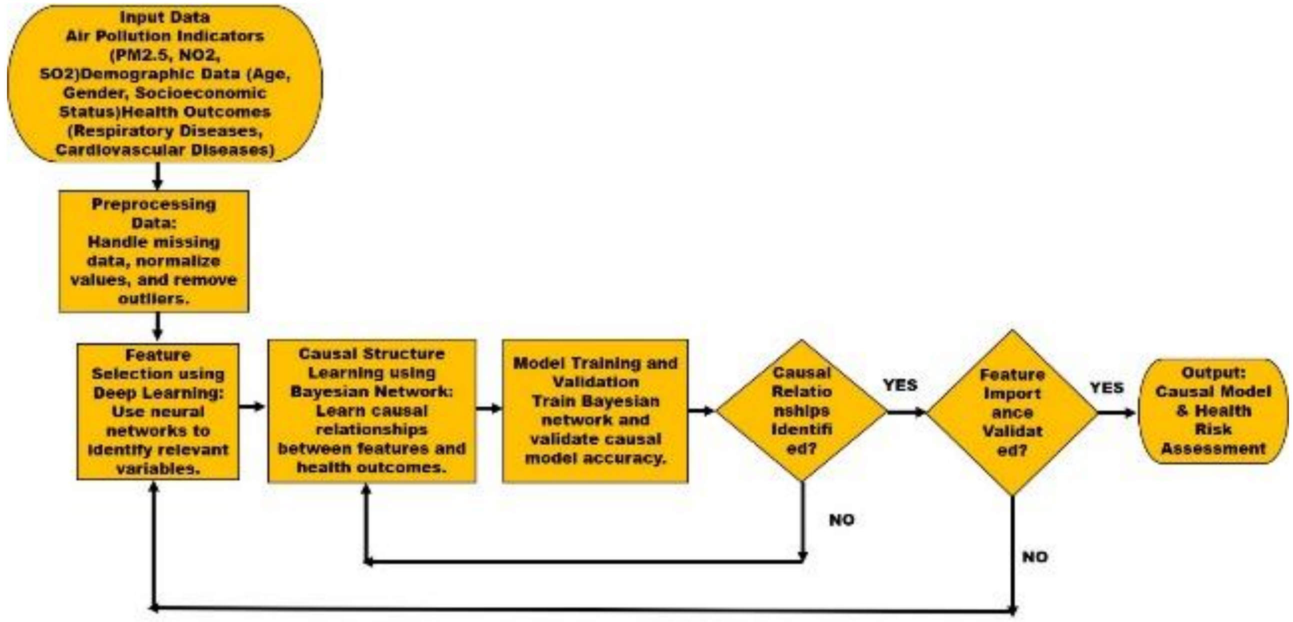

Fig. 1. Key components of the review.

C. Air Pollution and Health Impacts

Air pollution is a critical environmental health issue, with numerous studies linking exposure to pollutants such as particulate matter (PM), nitrogen oxides (NOx), and sulfur dioxide (SO2) to chronic health conditions, including respiratory diseases, cardiovascular illnesses, and cancer. Industrial areas and traffic-related pollution contribute significantly to harmful emissions, impacting urban populations and those residing near industrial zones. Research consistently demonstrates that prolonged exposure to these pollutants increases the risk of diseases like asthma, chronic obstructive pulmonary disease (COPD), heart disease, and cancer. Key studies, including those by Dockery et al. [1] and Pope et al. [2], have established robust associations between fine particulate matter (PM2.5) and mortality, emphasizing the urgent need to address air pollution, especially in urban environments with high industrial and vehicular emissions.

The Harvard Six Cities Study by Dockery et al. [1] revealed a strong association between long-term PM2.5 exposure and increased mortality from cardiovascular and respiratory diseases. Specifically, a 10 μg/m³ increase in PM2.5 levels corresponded to a 6–10% rise in mortality risk. Pope et al. [2] expanded on this with a larger cohort of over 500,000 individuals across 51 U.S. cities, confirming similar findings and underscoring the synergistic effects of air pollution and smoking on health outcomes. Brook et al. [3] focused on cardiovascular impacts, showing that even short-term exposure to PM2.5 could precipitate acute cardiovascular events, while long-term exposure exacerbated conditions like hypertension and diabetes. These studies collectively highlight the pressing need for advanced modeling techniques to analyze complex interactions between pollutants and their cumulative health effects and advocate for interventions to reduce exposure in high-risk areas.

Studies have also highlighted the disproportionate impact of air pollution on vulnerable populations. Künzli et al. [4] assessed the burden of disease across Europe, linking long-term exposure to particulate matter (PM10) with respiratory and cardiovascular diseases, especially in children, the elderly, and those with pre-existing conditions. The study emphasized the social and economic costs of pollution, such as increased healthcare expenses and reduced productivity. Research on traffic-related air pollution (TRAP) in children demonstrated that proximity to high-traffic areas elevated the risk of asthma exacerbations and respiratory infections, while access to green spaces reduced these risks [5]. A study in Toronto [6] further established a significant link between nitrogen dioxide (NO2) exposure and circulatory mortality, stressing the need for stricter air quality regulations. These findings underscore the importance of targeted interventions, such as reducing vehicular emissions and incorporating green spaces in urban planning, to protect vulnerable populations and improve public health outcomes.

D. Causal Inference and Causal Models

The use of causal inference methods in epidemiology has grown significantly, focusing on accurately identifying cause-effect relationships from observational data. This is especially important in environmental health, where randomized controlled trials (RCTs) are often impractical, and observational data is the primary evidence source. Causal inference aims to determine whether an exposure, such as air pollution, directly causes health outcomes like respiratory or cardiovascular diseases, rather than merely being correlated with them. This section summarizes key contributions to causal inference and its applications in environmental health.

Judea Pearl's Causality: Models, Reasoning, and Inference introduced a formal framework for analyzing causal relationships using causal diagrams, or directed acyclic graphs (DAGs) [7]. DAGs allow researchers to visualize causal structures and apply do-calculus to estimate causal effects from observational data while adjusting for confounders. For instance, Pearl's framework enables estimating the impact of reducing air pollution on health outcomes by accounting for variables such as socioeconomic status or smoking. However, accurately identifying and modeling confounders remains a challenge.

Donald Rubin's potential outcomes framework emphasizes counterfactuals, comparing outcomes based on exposure to treatment versus no exposure [8]. This approach is particularly useful in environmental health for evaluating the effects of pollution exposure. For example, researchers compare health outcomes in high-pollution versus low-pollution areas, assuming differences are attributable to pollution. Extensions such as propensity score matching help adjust for confounders but face challenges with unmeasured variables that may bias results.

The Bradford Hill criteria provide nine guidelines to evaluate causality in observational studies, such as strength, consistency, and temporality [9]. For example, studies by Dockery et al. [1] and Pope et al. [2] demonstrate consistent associations between air pollution and respiratory diseases, aligning with Hill's criteria. However, these guidelines lack formal statistical methods and rely on informed judgment.

Spirtes et al. developed algorithms like the PC algorithm to infer causal structures from data using graphical models [10]. These methods are valuable in uncovering relationships in complex systems, such as the interplay of air pollution, demographics, and health outcomes. However, they require large datasets and high-quality data, which can be challenging in environmental health.

Hernán and Robins advanced causal inference techniques, including marginal structural models and inverse probability weighting, to address time-varying exposures and confounding variables [11]. Their methods have been applied to study long-term pollution effects but require computational resources and assume no unmeasured confounders, a significant limitation in real-world applications.

E. Bayesian Networks for Causal Modeling

Bayesian networks (BNs) are graphical models that represent probabilistic relationships among variables, making them particularly useful for causal inference in complex domains like environmental health. They model dependencies between pollutants, health outcomes, and confounders, allowing researchers to estimate exposure effects while accounting for uncertainty and hidden variables. For example, BNs can assess the relationships between air pollutants (e.g., particulate matter, nitrogen dioxide) and diseases (e.g., respiratory or cardiovascular conditions) while adjusting for confounders such as socioeconomic status or lifestyle factors. Their ability to represent conditional dependencies and propagate evidence offers significant advantages over traditional regression models [12].

A notable study in Southern France's industrial Etang de Berre region utilized a Bayesian model to analyze 178 variables, finding correlations between pollutant exposure and pathologies such as lung cancer linked to hydrofluoric acid, diabetes from cadmium, and respiratory diseases from benzo[k]fluoranthene. Socioeconomic factors, like low education levels and single-parent families, were also associated with cardiovascular diseases, emphasizing the multifaceted nature of pollution's impact [13]. Similarly, a study in Shandong Province, China, applied Bayesian spatio-temporal models to examine cardiovascular disease hospitalizations, identifying significant increases per 10 $\mu g/m^3$ rise in $PM_{2.5}$, PM_{10}, SO_2, and NO_2 levels. While inland cities faced more severe impacts, short-term pollution spikes in coastal areas also exacerbated cardiovascular risks [14].

Other applications of BNs include decision-making frameworks and disease risk modeling. For instance, a study on waste management at the Bandarabbas refinery used BNs to prioritize risk factors, identifying Amine treatment and Fuel units as the most hazardous [15]. In Taiwan, a Bayesian conditional logistic regression model linked environmental factors like air pollution and weather conditions to ischemic heart disease risks, highlighting regional variations and providing actionable insights for disease prevention during hazardous conditions [16]. These examples underscore the versatility of Bayesian networks in uncovering and addressing environmental health challenges.

F. Deep Learning in Feature Selection

Deep learning techniques, particularly neural networks, have revolutionized feature selection in high-dimensional datasets, making them highly relevant for environmental health studies. These studies often involve large datasets encompassing pollutants, demographic factors, and health indicators. Deep learning algorithms, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), excel in identifying the most relevant features that drive health outcomes. This capability enhances the accuracy of causal models like Bayesian networks, enabling researchers to address complex relationships and improve predictions of disease outcomes [18].

A scoping review on Hepatocellular Carcinoma (HCC) highlights how deep learning models enhance feature selection for diagnosing, prognosticating, and treating this aggressive liver cancer. The review evaluates various algorithms and methodologies, noting significant improvements in accuracy and reductions in false positives. However, it also underscores the challenges of translating these advancements into clinical practice, suggesting avenues for future research [17]. Similarly, deep learning hybrid approaches have shown promise in other domains. For instance, Sparse Autoencoder (SAE) combined with Logistic Regression-Recursive Feature Elimination (LR-RFE) outperformed traditional methods in handling high-dimensional biological datasets, offering improved classification performance [20].

Convolutional neural networks (CNNs) are particularly valuable for analyzing environmental health data with spatial or temporal correlations. LeCun et al. demonstrated how CNNs detect hierarchical patterns, making them suitable for modeling pollutant exposure and health outcomes. For example, CNNs can forecast air quality or health risks by identifying spatial and temporal patterns in pollutant data. Similarly, Weighted GCNN (WGCNN) has proven effective for gene expression analysis, capturing non-linear interactions and addressing multi-class problems with enhanced performance metrics [18][21]. These studies underscore the transformative role of deep learning in feature selection for complex datasets.

G. Applications in Environmental Health Studies

Bayesian networks (BNs) and deep learning are increasingly employed in environmental health research to explore the impacts of air pollution on disease outcomes. Their integration offers a robust framework for modeling complex, non-linear relationships among pollutants, health outcomes, and confounders. For example, BNs have been used to assess cardiovascular risks by integrating modifiable and non-modifiable factors, allowing for improved disease management and decision-making through computational tools that predict risk and explore interrelations [22].

Hybrid models combining deep learning and Bayesian networks have further advanced health risk assessment. One study used a generalized additive model (GAM) combined with a long short-term memory (LSTM) neural network to predict hospitalizations from respiratory diseases due to air pollution. This hybrid approach outperformed standalone models, reducing mean absolute percentage error (MAPE) by 2.3% and providing a valuable tool for environmental health risk prediction [24]. Another hybrid approach integrated Bayesian stochastic partial differential equations (SPDE) with deep learning techniques like CNNs and DFFNNs to predict PM2.5 concentrations in urban areas. This model demonstrated superior accuracy and interpretability, addressing spatial-temporal dependencies and supporting public health strategies in polluted urban environments [25].

Machine learning models are also effective in chronic disease prediction. A study developed models for diabetes, hypertension, hyperlipidemia, and cardiovascular disease using real-world data and algorithms like gradient boosting and extreme gradient boosting. These models achieved over 80% accuracy and high AUC values (0.84–0.93), highlighting their utility in proactive disease management and risk identification [26]. These advancements underscore the transformative potential of BNs and deep learning in environmental epidemiology.

FINDINGS

This study highlights the strong association between air pollution exposure (e.g., PM2.5, NO2, SO2) and adverse health outcomes, including respiratory diseases, cardiovascular conditions, and specific cancers, particularly in industrial regions. By integrating Bayesian Networks and deep learning, it successfully modeled causal pathways and prioritized critical variables, offering actionable insights for predicting health risks and informing public health strategies.

- Air Pollution and Disease Outcomes: Strong associations found between exposure to PM2.5, NO2, and SO2 and respiratory diseases (e.g., asthma, COPD), cardiovascular conditions, and certain cancers, with industrial proximity worsening health risks.
- Bayesian Networks for Causal Modeling: Successfully modeled causal pathways, accounting for confounding factors like socioeconomic status and age, while quantifying uncertainties in relationships.
- Deep Learning for Feature Selection: Enhanced Bayesian Network accuracy by prioritizing critical variables (pollutants, demographics) and reducing noise in high-dimensional datasets.
- Challenges in Modeling: Confounders (e.g., lifestyle, genetics) and limited real-time data introduced uncertainties, affecting model validation and scalability.
- Environmental Health Applications: Demonstrated potential in industrial zones for health risk assessment, offering actionable insights for public health.

A. Gaps in Literature and Future Directions

Despite progress in applying Bayesian networks and deep learning to environmental health, significant gaps remain, particularly in data quality, model interpretability, and addressing vulnerable populations. Limited availability of high-quality, standardized datasets and a lack of longitudinal data constrain the ability to model long-term health impacts. Furthermore, the "black-box" nature of deep learning models poses challenges in interpretability, which is crucial for actionable insights. Vulnerable populations, such as children and those in low-income areas, are often underrepresented, limiting the equity of findings. Additionally, the narrow focus on specific pollutants or diseases fails to capture the multifactorial nature of environmental health, while geographic and temporal generalizability remains a challenge. Future research directions include:

- Data Standardization and Crowdsourcing: Develop standardized data collection methods and leverage remote sensing and crowdsourcing to fill data gaps.
- Longitudinal Studies: Integrate electronic health records with exposure data to track long-term health impacts.
- Explainable AI: Use methods like LIME and SHAP to enhance interpretability and integrate Bayesian networks for causal reasoning.
- Focus on Vulnerable Groups: Expand research on environmental justice and health disparities.
- Multifactorial Models: Incorporate diverse data, such as socioeconomic and lifestyle factors, into holistic models.
- Global Studies and Real-Time Data: Conduct cross-regional studies and integrate dynamic data for accurate, adaptable predictions.

B. Recommendations

Strengthening the application of advanced modeling in environmental health requires addressing critical gaps in data collection, causal modeling, policy interventions, and collaborative research. Real-time air pollution monitoring and detailed health records, integrated with GIS, can enhance data quality and map exposure effects. Advanced causal models, such as Bayesian networks combined with deep learning, should be employed to tackle uncertainties and confounding factors. Stricter air quality regulations, healthcare access improvement, and pollution mitigation strategies are essential for effective public health interventions. Cross-disciplinary collaboration is crucial for developing robust, scalable models and conducting longitudinal studies. Investments in AI-driven feature selection and predictive analytics can refine models for proactive strategies. Future research must focus on long-term health impacts, urban-rural pollution comparisons, and integrating genetic and lifestyle factors into causal models. Key actions include:

- Establish real-time monitoring systems and GIS integration.
- Use hybrid Bayesian and deep learning models for causal inference.
- Enforce stricter air quality regulations and improve healthcare access.
- Foster cross-disciplinary collaborations for scalable solutions.
- Invest in AI for feature selection and health risk prediction.

CONCLUSION

Addressing the gaps in the current literature is critical for advancing our understanding of the health impacts of air pollution and improving public health outcomes. By focusing on data quality, model interpretability, vulnerable populations, and causal inference, future research can enhance the accuracy, transparency, and applicability of Bayesian networks and deep learning in environmental health studies. These improvements will ultimately lead to more informed policy decisions and targeted interventions to mitigate the health risks of air pollution, especially in industrial areas and vulnerable communities. The findings emphasize the importance of causal inference models that combine Bayesian Networks and deep learning for feature selection to assess air pollution exposure and its impact on health outcomes. Implementing these models can improve public health decision-making, policy development, and environmental risk management in industrial regions.

REFRENCES

[1] D. W. Dockery, C. A. Pope 3rd, X. Xu, J. D. Spengler, J. H. Ware, M. E. Fay, B. G. Ferris Jr, and F. E. Speizer, “An association between air pollution and mortality in six U.S. cities,” The New England Journal of Medicine, vol. 329, no. 24, pp. 1753-1759, 1993. doi: 10.1056/NEJM199312093292401.

[2] C. A. Pope 3rd, R. T. Burnett, M. J. Thun, E. E. Calle, D. Krewski, K. Ito, G. D. Thurston, “Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution,” Journal of the American Medical Association, vol. 287, no. 9, pp. 1132-1141, 2002. doi: 10.1001/jama.287.9.1132.

[3] R. D. Brook, S. Rajagopalan, C. A. Pope 3rd, J. R. Brook, A. Bhatnagar, A. V. Diez-Roux, F. Holguin, Y. Hong, R. V. Luepker, M. A. Mittleman, A. Peters, D. Siscovick, S. C. Smith Jr, L. Whitsel, J. D. Kaufman, and American Heart Association Council on Epidemiology and Prevention, Council on the Kidney in Cardiovascular Disease, and Council on Nutrition, Physical Activity and Metabolism, “Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association,” Circulation, vol. 121, no. 21, pp. 2331–2378, 2010. doi: 10.1161/CIR.0b013e3181dbece1.

[4] N. Künzli, R. Kaiser, S. Medina, M. Studnicka, O. Chanel, P. Filliger, M. Herry, F. Horak Jr, V. Puybonnieux-Texier, P. Quénel, J. Schneider, R. Seethaler, J. C. Vergnaud, and H. Sommer, “Public-health impact of outdoor and traffic-related air pollution: a European assessment,” Lancet, vol. 356, no. 9232, pp. 795–801, 2000. doi: 10.1016/S0140-6736(00)02653-2.

[5] S. Esposito, C. Galeone, M. Lelii, B. Longhi, B. Ascolese, L. Senatore, E. Prada, V. Montinaro, S. Malerba, M. F. Patria, and N. Principi, “Impact of air pollution on respiratory diseases in children with recurrent wheezing or asthma,” BMC Pulmonary Medicine, vol. 14, no. 130, 2014. doi: 10.1186/1471-2466-14-130.

[6] M. Jerrett, M. M. Finkelstein, J. R. Brook, M. A. Arain, P. Kanaroglou, D. M. Stieb, N. L. Gilbert, D. Verma, N. Finkelstein, K. R. Chapman, and M. R. Sears, “A cohort study of traffic-related air pollution and mortality in Toronto, Ontario, Canada,” Environmental Health Perspectives, vol. 117, no. 5, 2009. doi: 10.1289/ehp.11533.

[7] J. Pearl, Causality: Models, reasoning, and inference. Cambridge, U.K.: Cambridge University Press, 2000.

[8] D. B. Rubin, “Estimating causal effects of treatments in randomized and nonrandomized studies,” Journal of Educational Psychology, vol. 66, no. 5, pp. 688–701, 1974. doi: 10.1037/h0037350.

[9] A. B. Hill, “The environment and disease: Association or causation?” Proc. R. Soc. Med., vol. 58, no. 5, pp. 295-300, May 1965. doi: 10.1177/003591576505800503.

[10] P. Spirtes, C. Glymour, and R. Scheines, Causation, Prediction, and Search. Cambridge, MA, USA: MIT Press, 2000.

[11] M. A. Hernán and J. M. Robins, Causal Inference: What If. Boca Raton, FL, USA: Chapman & Hall/CRC, 2016

[12] S. L. Lauritzen, Graphical Models. Oxford, U.K.: Oxford University Press, 1996

[13] S. Pérez, C. German-Labaume, S. Mathiot, S. Goix, and P. Chamaret, “Using Bayesian networks for environmental health risk assessment,” Environmental Research, vol. 204, Pt B, p. 112059, 2022. doi: 10.1016/j.envres.2021.112059

[14] Y. Liu, J. Sun, Y. Gou, X. Sun, D. Zhang, and F. Xue, "Analysis of short-term effects of air pollution on cardiovascular disease using Bayesian spatio-temporal models," International Journal of Environmental Research and Public Health, vol. 17, no. 3, p. 879, 2020. doi: 10.3390/ijerph17030879

[15] M. Saeedi and B. Malekmohammadi, "Contribution of Bayesian networks as a robust tool in risk assessment under sustainability considerations, a case study of Bandarabbas refinery," Eliyons, vol. 9, e15264, 2023

[16] C. Wang, S. J. Lin, C. K. Hsiao, and K. C. Lu, "Bayesian approach to disease risk evaluation based on air pollution and weather conditions," International Journal of Environmental Research and Public Health, vol. 20, no. 2, p. 1039, 2023. doi: 10.3390/ijerph20021039

[17] G. Mostafa, H. Mahmoud, T. Abd El-Hafeez, et al., "The power of deep learning in simplifying feature selection for hepatocellular carcinoma: A review," BMC Medical Informatics and Decision Making, vol. 24, p. 287, 2024. doi: 10.1186/s12911-024-02682-1

[18] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, pp. 436–444, 2015. doi: 10.1038/nature14539

[19] N. H. Reduwan, A. A. Abdul Aziz, R. Mohd Razi, et al., "Application of deep learning and feature selection technique on external root resorption identification on CBCT images," BMC Oral Health, vol. 24, p. 252, 2024. doi: 10.1186/s12903-024-03910-w

[20] H. Wu, "J. Phys.: Conf. Ser.," vol. 1848, p. 012019, 2021

[21] A. K. Naik, V. Kuppili, "An embedded feature selection method based on generalized classifier neural network for cancer classification," Computers in Biology and Medicine, vol. 168, p. 107677, 2024. doi: 10.1016/j.compbiomed.2023.107677.

[22] J. M. Ordovas, D. Rios-Insua, A. Santos-Lozano, A. Lucia, A. Torres, A. Kosgodagan, J. M. Camacho, "A Bayesian network model for predicting cardiovascular risk," Computer Methods and Programs in Biomedicine, vol. 231, p. 107405, 2023. doi: 10.1016/j.cmpb.2023.107405

[23] J. Pavez and H. Allende, "A hybrid system based on Bayesian networks and deep learning for explainable mental health diagnosis," Applied Sciences, vol. 14, p. 8283, 2024. doi: 10.3390/app14188283

[24] S. Qin, X. Duan, and P. Kimm, "Usage of deep learning in environmental health risk assessment," Work, pp. 1-13, 2021. doi: 10.3233/WOR-205371

[25] D. P. Johnson, N. Ravi, G. Filippelli, and A. Heintzelman, "A novel hybrid approach: Integrating Bayesian SPDE and deep learning for enhanced spatiotemporal modeling of PM2.5 concentrations in urban airsheds for sustainable climate action and public health," Sustainability, vol. 16, no. 23, p. 10206, 2024. doi: 10.3390/su162310206

[26] C. Lee, B. Jo, H. Woo, Y. Im, R. W. Park, and C. Park, "Chronic disease prediction using the common data model: Development study," JMIR AI, vol. 1, no. 1, p. e41030, Dec. 2022. doi: 10.2196/41030.