Review

Applied Bayesian Networks Rely on Expert Knowledge and Scarce Data Sharing

Liam Coorssen , Hamid Kalantari , Parham Afsharnia , Pouria Ramazi *

Posted Date: 18 June 2026

doi: 10.20944/preprints202606.1426.v1

Keywords: Bayesian networks; structure learning; causal discovery; reproducibility; open datasets

Review

Applied Bayesian Networks Rely on Expert Knowledge and Scarce Data Sharing

Liam Coorssen [1], Hamid Kalantari [2], Parham Afsharnia [3] and Pouria Ramazi [4,5,*]

[1] Department of Mathematics and Statistics, Brock University, St. Catharines, Canada
[2] Department of Computing Science, University of Alberta, Edmonton, Canada
[3] Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan, Iran
[4] Department of Biological Sciences, University of Calgary, Calgary, Canada
[5] Department of Mathematics and Statistics, University of Calgary, Calgary, Canada
* Correspondence: pouria.ramazi@ucalgary.ca

Highlights

What are the main findings?

- Expert knowledge was used in 56.2% of relevant Bayesian network papers.
- Only 3.5% of relevant papers shared functional dataset links.
- Score-based methods dominated reported algorithmic structure learning.

What are the implications of the main findings?

- Open dataset sharing should become routine in applied BN studies.
- Expert-informed models should be compared with automatic baselines.
- The curated dataset index supports replication and benchmarking.

Abstract

To assess how Bayesian network structures are built and learned in applied work, we screened 5,993 recent papers (2020-2025) whose abstracts mention "Bayesian (belief) network" and deemed 3,661 relevant. Among these relevant papers, expert knowledge was used in 2,059 papers (56.2%): 1,785 (48.8%) relied on expert knowledge alone, whereas 274 (7.5%) combined expert input with algorithmic structure learning through edge modification or structural constraints. Data sharing was scarce: only 129 studies (3.5%) provided functional dataset links, which we curated into an open benchmark index. These findings support a data-first norm: share datasets, prefer automatic structure-learning baselines, and document expert input. Among 1,106 papers (30.2%) using algorithms without expert knowledge, score-based flags were most common (797, 72.1%; mainly hill climbing, K2, and tabu), followed by constraint-based methods (194, 17.5%; mainly PC and Grow-Shrink), fixed or restricted-topology BN classifiers (143, 12.9%; mainly TAN and naive Bayes), and hybrid methods (131, 11.8%; mainly MMHC); bootstrapping appeared in 223 papers (6.1%). Reported practice remains concentrated around familiar algorithms.

Keywords: Bayesian networks; structure learning; causal discovery; reproducibility; open datasets

MSC: 62C10; 62-02; 62-08; 68T05; 68T37

---

1. Introduction

Bayesian networks (BNs) are increasingly being used across a range of disciplines—from ecology and systems biology to medicine and engineering—to model complex systems where uncertainty and interdependence among variables are key concerns [1]. Their ability to represent joint probability distributions through a graph-based structure makes them both intuitive and powerful. Formally, a BN is a directed acyclic graph (DAG) in which nodes correspond to random variables

doi:10.20944/preprints202606.1426.v1

and edges represent conditional dependencies. The joint probability distribution over variables X_1, ..., X_n then factorizes into the product of each variable conditional on its parent set Pa(X_i). This factorization allows BNs to capture complex interdependencies in a compact and computationally efficient way, making them well suited for both inference and prediction.

The network structure can be learned fully automatically from data. Methods fall into three families. Constraint-based algorithms (e.g., PC/PC-stable; Grow-Shrink) infer adjacencies via conditional-independence tests and then orient edges by logical rules [2–4]. Score-based procedures optimize a goodness-of-fit criterion—commonly AIC or BIC—over DAG space using heuristics such as greedy hill climbing, tabu search, or the K2 family [5–9]. Hybrid methods combine both, typically constraining candidate neighborhoods by CI tests and then refining with a score-based search, such as MMHC [10]. For broader overviews, see recent surveys [11–13].

Despite these advantages, practical implementations of BNs often fall short of their full potential. Two recurring issues are (i) relying on expert-elicited structures in lieu of automatic structure learning [1,14], and (ii) not providing the underlying datasets, which impedes replication and benchmarking. It remains an open question how widespread these practices are, which fields rely most on expert-driven models, and—among studies that perform structure learning—which algorithms and search/score configurations are most commonly used.

Prior surveys are valuable for organizing algorithms and theory, but they do not set out to quantify adoption patterns, field-level usage, or dataset sharing in the applied literature. For example, Scanagatta et al. review structure-learning families, missing-data strategies, and software [11]; Kitson et al. catalogue 74 algorithms and evaluation practices [12]; Vowels et al. emphasize modern DAG discovery, including continuous-optimization approaches [13]. Benchmark-style comparisons study accuracy and speed across algorithm families [16]. Domain-focused reviews, such as agriculture-focused work [15], or application overviews [1], synthesize uses within specific sectors. None of these map, at scale, how often expert elicitation replaces automatic learning across fields, how frequently datasets are shared with working links, or which structure-learning algorithms dominate in practice across domains.

We address this gap with a scoping review of 5,993 studies (2020-2025), coding discipline, data availability, use of expert knowledge, overlap between expert knowledge and algorithmic learning, structure-learning family, fixed or restricted-topology BN classifiers, named algorithms, and bootstrapping. We also release a curated dataset index with working links to support replication and benchmarking. Our analysis provides a field-wide quantification of these practices and a resource to standardize empirical comparisons in Bayesian network structure learning.

Before presenting the methodology, we briefly introduce frequently used structure-learning algorithms.

2. Structure Learning Algorithms

2.1. Peter-Clark (PC) Algorithm

The PC algorithm [2] is a constraint-based method that constructs the structure of a Bayesian network by identifying conditional independencies in the data. It begins with a fully connected undirected graph where all variables are connected. For each pair of variables X and Y, it searches for a conditioning set $S \subseteq V \setminus \{X, Y\}$ such that $X \perp Y \mid S$. If such a set is found, the edge X—Y is removed. The algorithm gradually increases the size of the conditioning set and then applies orientation rules, such as collider orientation and Meek's rules, to direct the remaining edges in a way that ensures a DAG structure.

2.2. K2 Algorithm

The K2 algorithm [9] is a score-based structure-learning algorithm that searches through possible parent sets in a restricted greedy manner. It assumes a known topological ordering of the nodes and seeks to identify the best parent set for each node that maximizes the posterior probability

doi:10.20944/preprints202606.1426.v1

of the network structure given the data. For each node, the method starts with an empty parent set and iteratively adds the preceding candidate parent that most increases the K2 score until no addition improves the score or a predefined limit is reached.

2.3. Hill-Climbing Search

Hill climbing [17] is a local search strategy that attempts to improve a scoring function by making small changes to a candidate network. It starts with an initial graph, generates neighbors through single-edge changes (add, delete, or reverse), evaluates each neighbor using a score such as BIC, AIC, or BDeu, and moves to the highest-scoring neighbor if it improves the score. The process repeats until no improvement is possible. The BIC score balances model fit and complexity. Hill climbing is simple and fast but may miss better structures reachable only through non-local changes.

2.4. Greedy Search

Greedy search [16] follows the same general process as hill climbing but is a broader term for algorithms that make the best immediate decision at each step without backtracking or global exploration.

2.5. Tabu Search

Tabu search [18] enhances greedy or hill-climbing methods by adding a short-term memory—a tabu list—to avoid revisiting previous solutions. It considers local moves, including moves that do not improve the score, while forbidding moves that reverse recent changes. This allows the search to escape local optima and can improve performance on large or noisy datasets.

2.6. Grow-Shrink (GS) Algorithm

The Grow-Shrink algorithm [4] is a constraint-based method that learns the Markov blanket of each variable using conditional-independence tests. In the grow phase, variables are added to the current Markov blanket when they remain dependent conditional on the current blanket. In the shrink phase, variables are removed if they become conditionally independent given the rest of the blanket. The resulting Markov blankets are combined to form the network skeleton and orient edges.

2.7. Thick-Thinning Algorithm

The Thick-Thinning algorithm [19] is a two-phase score-based method. In the thickening phase, the algorithm starts from an empty or sparse DAG and iteratively adds the edge that most improves a score while avoiding cycles. In the thinning phase, edges are removed if doing so improves or maintains the score. This approach builds a candidate structure and then prunes it for parsimony.

3. Methods

To investigate the prevalence and diversity of structure-learning approaches in Bayesian network research, a comprehensive review of the literature was conducted in Web of Science. Papers were identified between 2020 and 2025. Inclusion criteria required that papers contain at least one of the following terms in their abstract: “Bayesian belief network” or “Bayesian network”. This broad search strategy ensured coverage of a wide range of modeling approaches while maintaining focus on methods relevant to BN structure learning. In total, 5,993 papers were retrieved and accessible for review.

Following collection, an initial screening process was conducted to extract methodological features from each paper. A dataset was developed in Microsoft Excel to facilitate consistent data entry and tracking of key features across studies. Each paper was reviewed and assigned binary or categorical values across predefined variables: presence of a dataset, availability of a dataset link, use of score-based learning, constraint-based learning, or hybrid learning approaches, fixed or restricted-topology BN classifiers, reliance on expert knowledge for network construction or modification, mention of bootstrapping techniques, specific learning algorithm used, and the scoring metric applied during model training.

Each paper was analyzed, and attributes were recorded either as binary values (e.g., 1 = present, 0 = absent) or as text entries where appropriate (e.g., algorithm name or score type). For example, if a paper employed a score-based algorithm such as K2 in conjunction with the Bayesian Information Criterion (BIC), the following entries would be recorded: 1 under score-based, K2 under specific learning approach, and BIC under score used for training. Constraint-based and hybrid approaches were recorded in a similar manner. If more than one structure-learning method was used or discussed in a single paper, multiple fields could be populated. Similarly, expert knowledge was noted if the network structure was manually derived from expert surveys, if structural constraints such as required or forbidden edges were imposed, or if a learned DAG was subsequently modified based on expert feedback.

To determine whether bootstrapping was employed, each document was searched for the term "bootstrap" using the text-search functionality within the PDF viewer. If any results were returned and the context confirmed the use of a bootstrapping method, either for uncertainty quantification or model validation, a 1 was entered under the bootstrapping column.

Fixed or restricted-topology BN classifiers, such as naive Bayes, TAN, and related variants, were classified as a separate method family because their structures are fixed or partly predetermined. If a paper also reported score-based, constraint-based, or hybrid structure learning, those methods were counted under their corresponding families.

Once all available papers were annotated, the compiled data were imported into R for analysis and visualization. Descriptive statistics were used to summarize the distribution of structure-learning methods and associated features across the literature. To assess the representation of different academic domains, research fields were extracted from each paper and categorized. Fields represented by fewer than 55 papers were aggregated into a single "Other" category. A figure was generated to illustrate the distribution of BN applications across research areas.

A separate list of datasets was also compiled. This included only datasets with direct, functional links, either in the supplementary information or embedded in the body of the paper. This criterion was applied to evaluate reproducibility and accessibility across BN studies. Datasets without an accessible URL or repository reference were excluded, regardless of whether the data were described in the text.

To further explore trends in methodological preferences, totals were calculated for each major method family and for normalized named approaches, including both widely used algorithms (e.g., PC and K2) and less common techniques. These counts were summarized to identify the most frequently reported methods in the literature.

4. Results

A total of 5,993 papers were included in the review. Of these, 3,661 were retained as relevant papers for analysis of how BN structures were built or learned. The remaining 2,332 papers were excluded, including 1,239 meta-based papers. Percentages in the Results section use the 3,661 relevant papers as the denominator. Fixed or restricted-topology BN classifiers, such as TAN, naive Bayes, and related variants, were retained as a separate method family rather than merged with score-based, constraint-based, or hybrid structure-learning methods.

4.1. Disciplinary Breakdown

Bayesian networks were applied across a broad spectrum of disciplines in the total corpus of 5,993 papers (Figure 1). The largest shares of studies came from engineering (20.7%) and computer science (20.0%). Other areas, including environmental sciences and ecology (5.5%), general and internal medicine (4.3%), and various branches of the natural and social sciences, also demonstrated uptake at smaller scales.

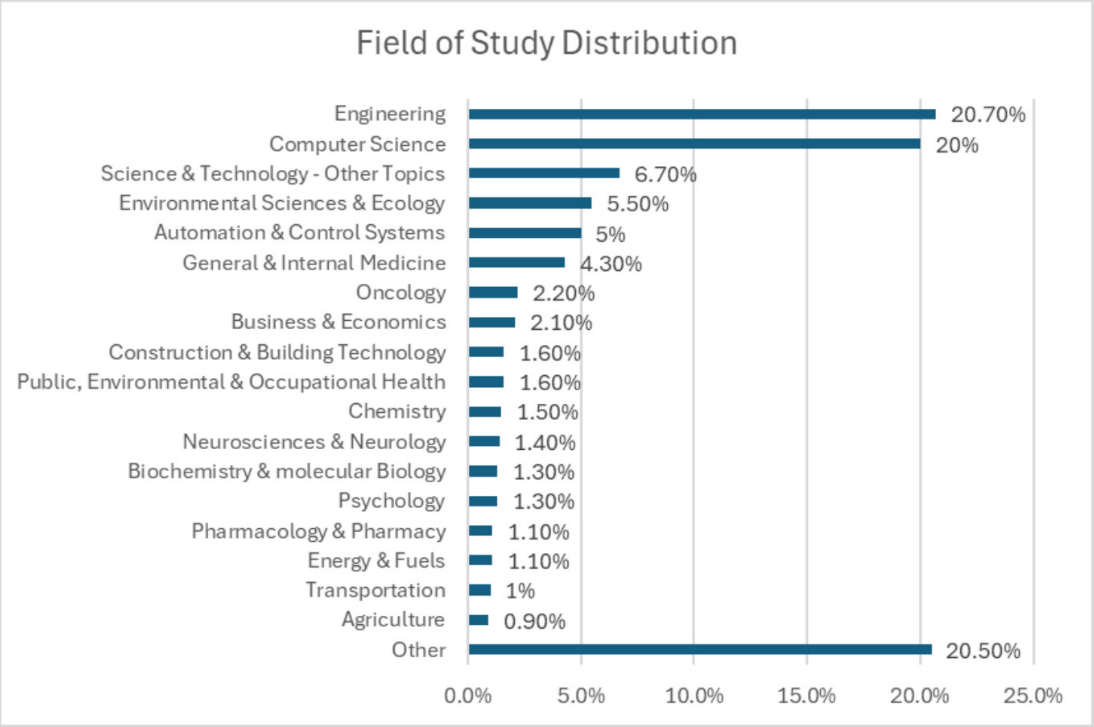

Figure 1. Distribution of papers including the terms "Bayesian belief network" and "Bayesian network" over 2020-2025.

4.2. Use of Expert Knowledge

Expert knowledge was used in 2,059 relevant papers (56.2%; Figure 2). Of these, 1,785 papers (48.8%) relied on expert knowledge without an algorithm or fixed/restricted-topology classifier, whereas 274 papers (7.5%) combined expert input with such methods. Expert input was typically used to construct the DAG directly, impose structural constraints such as required or forbidden edges, or modify an algorithmically learned DAG. Expert-driven modeling was prominent in engineering (n = 785, 21.4%) and computer science (n = 256, 7.0%).

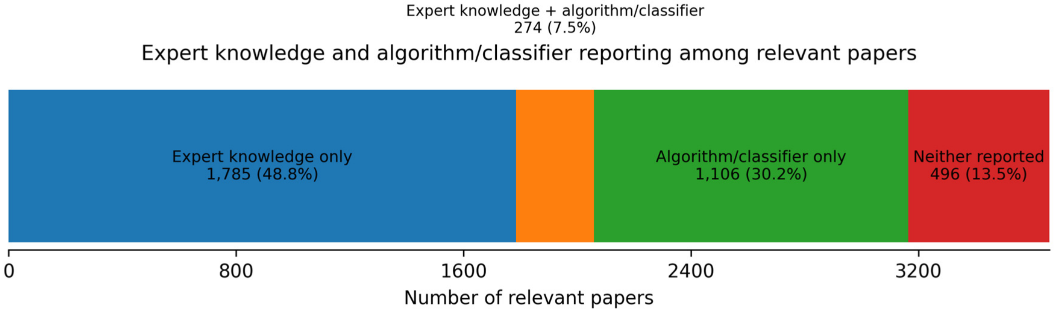

Figure 2. Mutually exclusive categories among the 3,661 relevant papers according to whether expert knowledge, an algorithm or fixed/restricted-topology classifier, both, or neither was reported.

4.3. Structure-Learning Methods

Across the 3,661 relevant papers, score-based methods were reported in 950 papers (25.9%), constraint-based methods in 236 papers (6.4%), fixed or restricted-topology BN classifiers in 164 papers (4.5%), and hybrid methods in 182 papers (5.0%). These categories are not mutually exclusive because some papers compared or combined multiple approaches. Among the 1,106 papers that

doi:10.20944/preprints202606.1426.v1

reported an algorithm or fixed/restricted-topology classifier without expert knowledge, the corresponding counts were 797 score-based, 194 constraint-based, 143 fixed/restricted-topology classifier, and 131 hybrid papers.

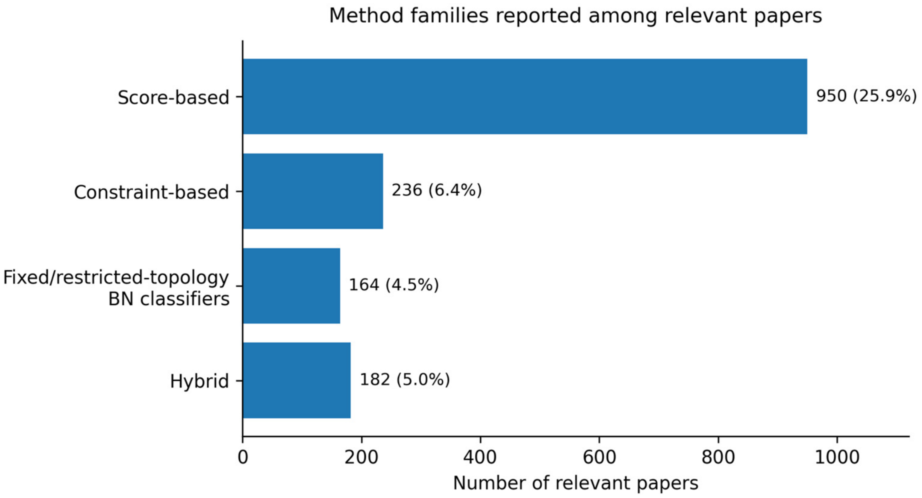

Figure 3. Method families reported among the 3,661 relevant papers. Papers may report more than one family.

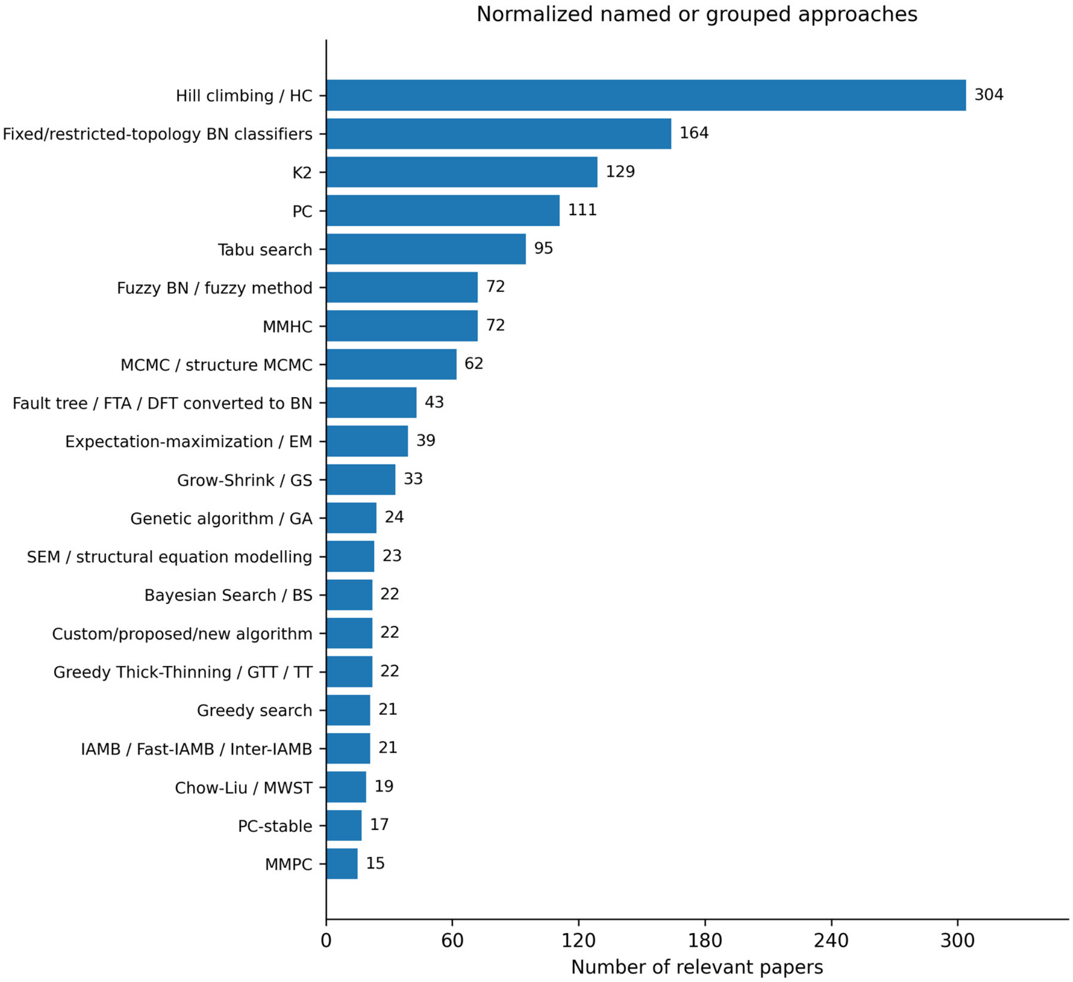

Figure 4. Counts of normalized named or grouped approaches among the 3,661 relevant papers.

doi:10.20944/preprints202606.1426.v1

The most frequently named or grouped approaches were hill climbing (n = 304, 8.3%), fixed or restricted-topology BN classifiers (n = 164, 4.5%; mainly TAN and naive Bayes), K2 (n = 129, 3.5%), PC (n = 111, 3.0%), tabu search (n = 95, 2.6%), fuzzy BN or fuzzy methods (n = 72, 2.0%), MMHC (n = 72, 2.0%), and MCMC or structure MCMC (n = 62, 1.7%). Among constraint-based or Markov-blanket methods, Grow-Shrink appeared in 33 papers (0.9%), IAMB variants in 21 papers (0.6%), PC-stable in 17 papers (0.5%), and MMPC in 15 papers (0.4%).

4.4. Bootstrapping

Bootstrapping, used to estimate model stability and confidence in structural features, appeared in 223 relevant papers (6.1%).

4.5. Reproducibility

Only 129 relevant papers (3.5%) included a readily accessible dataset, defined as having a direct link to supplementary material, a journal-hosted data file, or a data repository. This sparsity limits reproducibility and comparative benchmarking. Without access to the underlying data, it becomes difficult to evaluate the robustness of reported results, replicate findings, or compare algorithmic performance across studies. Supplementary Table S1 lists the links to the datasets in all 129 papers, together with their references and reported methods.

Table 1. Summary of main outcomes among the 3,661 relevant papers.

Outcome	Count	% of relevant papers
Relevant papers retained for analysis	3,661	100.0%
Open datasets with functional links	129	3.5%
Expert knowledge used	2,059	56.2%
Expert knowledge only	1,785	48.8%
Expert knowledge + algorithm/classifier	274	7.5%
Algorithm/classifier only	1,106	30.2%
Neither expert knowledge nor algorithm/classifier reported	496	13.5%
Score-based methods	950	25.9%
Constraint-based methods	236	6.4%
Hybrid methods	182	5.0%
Fixed or restricted-topology BN classifiers	164	4.5%
Bootstrapping	223	6.1%

Table 2. Normalized named or grouped approaches with expert-knowledge overlap.

Algorithm / approach	**Total**	**With expert knowledge**	**Without expert knowledge**
Hill climbing / HC	304	40	264
Fixed or restricted-topology BN classifiers	164	21	143
K2	129	33	96
PC	111	21	90
Tabu search	95	13	82
Fuzzy BN / fuzzy method	72	49	23
MMHC	72	11	61
MCMC / structure MCMC	62	3	59
Fault tree / FTA / DFT converted to BN	43	26	17
Expectation-maximization / EM	39	16	23
Grow-Shrink / GS	33	3	30
Genetic algorithm / GA	24	6	18
SEM / structural equation modelling	23	2	21
Bayesian Search / BS	22	7	15

Custom/proposed/new algorithm	22	1	21
Greedy Thick-Thinning / GTT / TT	22	4	18
Greedy search	21	1	20
IAMB / Fast-IAMB / Inter-IAMB	21	0	21
Chow-Liu / MWST	19	3	16
PC-stable	17	4	13
MMPC	15	4	11
ISM	14	11	3
DEMATEL	13	13	0
GES / FGES / ARGES / SCCGES	13	4	9
Dynamic Bayesian network / DBN	11	6	5
Bow-tie converted/modelled as BN	10	8	2
Delphi/expert scoring	10	9	1
RSMAX2 / restricted maximization	10	3	7

5. Discussion

Bayesian networks capture conditional dependence through directed acyclic graphs and support probabilistic reasoning and, with additional assumptions, causal interpretation. Structure learning seeks graphs that best explain observed independencies or optimize a fit-complexity criterion, with three main families—constraint-based, score-based, and hybrid—each offering different theoretical and practical trade-offs [2,20].

Across disciplines, use of Bayesian networks was broad but uneven. Engineering accounted for 20.7% and computer science for 20.0% of the screened corpus, with smaller shares in science-and-technology "other" topics, automation and control, environmental sciences and ecology, and general and internal medicine. This pattern reflects BN versatility in technical decision-making and signals growth opportunities in underrepresented areas such as psychology, economics, and public health.

Data availability remains the central barrier to comparability and cumulative progress. Only 129 relevant studies (3.5%) provided working links to the underlying datasets, limiting external validation, cross-paper benchmarking, and reuse. In this review we therefore compiled a dataset index with direct links; this resource lowers search costs, supports immediate replication, and facilitates standardized benchmarks and teaching use. We also defined strict inclusion criteria for that table—only datasets with functional links in the paper or supplement were retained—to promote reliable reuse. Community experience in machine learning shows that artifact policies and reproducibility checklists can improve transparency and research quality [22,23], and the dataset index is intended to serve that role for Bayesian network structure learning.

Expert knowledge was used in 56.2% of relevant papers, and 48.8% relied on expert knowledge without an algorithm or fixed/restricted-topology classifier. While expert constraints can improve interpretability and shrink the search space, they can also introduce bias and hinder replication unless the elicitation protocol and exact structural constraints are reported. In this review we emphasized the value of fully automatic, data-driven structure learning as the default; expert-informed models, if used, should be reported alongside an automatic baseline to quantify incremental value and expose discrepancies [14,20,21].

Methodologically, score-based methods were the most common. Constraint-based methods, hybrid methods, and fixed or restricted-topology BN classifiers were much less frequent. Among papers reporting an algorithm or fixed/restricted-topology classifier without expert knowledge, score-based methods were the largest group, followed by constraint-based methods, fixed/restricted-topology classifiers, and hybrid methods. Reporting the chosen score, hyperparameters, search settings, and any expert constraints is essential for interpretability and comparability [5–7,10,16].

The named-method distribution suggests that the field relies heavily on a small subset of familiar algorithms and software-accessible workflows. Hill climbing, fixed or restricted-topology BN classifiers, K2, PC, tabu search, MMHC, and MCMC accounted for much of the named-method reporting, whereas many exact, metaheuristic, continuous-optimization, and Markov-blanket approaches appeared only sporadically. This concentration suggests that applied BN research is shaped not only by algorithmic performance, but also by familiarity, implementation availability, and reporting conventions.

Bootstrapping appears relatively uncommon: only 6.1% of the relevant papers employed some form of resampling. Because model stability and uncertainty quantification matter in graphical models—especially for causal inference or decision support—robustness checks are important. Bootstrap resampling, or alternatives such as Bayesian model averaging, can be informative for assessing sensitivity to sampling variability, although many studies do not report such analyses.

There are limitations to this review. Our classification recorded the presence or absence of methods and practices at the paper level rather than effect sizes or head-to-head outcomes; papers using multiple approaches were counted across categories; and fixed or restricted-topology BN classifiers were grouped separately from score-based, constraint-based, and hybrid structure-learning families. The counts therefore represent reported adoption, not algorithmic performance or mutually exclusive methodological pathways. Field labels with small counts were aggregated. These choices were made a priori to balance breadth with consistency.

6. Recommendations

1. Share data and code by default, with stable links or DOIs and minimal metadata.
2. Deposit data in a trusted, domain-appropriate repository that offers persistent identifiers, rich metadata, and long-term preservation—such as Zenodo, Dryad, or Figshare; where specialist archives exist, prefer the disciplinary option [24–27].
3. Use fully automatic, data-driven structure learning as the default; if expert knowledge is used to construct, restrict, or modify the DAG, report it alongside an automatic baseline and document elicitation protocols, required or forbidden edges, and post-learning modifications in detail.
4. Document scoring and search settings explicitly (score family, hyperparameters, restarts, tabu, simulated annealing) to enable reproducibility and fair comparison.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org. The curated index of 129 open datasets is provided as Supplementary Table S1 in Excel format.

Author Contributions: Conceptualization, P.R. and H.K.; methodology, H.K., L.C., P.A., and P.R.; formal analysis, H.K., L.C., and P.A.; investigation, L.C., H.K., and P.A.; data curation, L.C.; visualization, L.C.; writing—original draft preparation, L.C.; writing—review and editing, L.C., H.K., P.A., and P.R.; supervision, P.R.; project administration, P.R.; All authors have read and agreed to the published version of the manuscript.

Funding: P.R. acknowledges support from an NSERC Discovery Grant.

Data Availability Statement: No new experimental data were generated in this study.

Acknowledgments: We thank Sam Longdon, Ethan Garrafa, Milo Boccinfuso, Rika Luo, Bennett Warkentin, Hayden James, Ada Fast, Sheraz Ahmad, Maham Taqi, and Izaak Gagnon, who supported the literature review through screening and preliminary assessments of papers.

Conflicts of Interest: The authors declare no conflicts of interest. The funder had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Iqbal K, Yin X-C, Hao H-W, Ilyas QM, Ali H. An overview of Bayesian network applications in uncertain domains. International Journal of Computer Theory and Engineering. 2015;7(6):416-427.

2. Spirtes P, Glymour C, Scheines R. Causation, prediction, and search. 2nd ed. Cambridge: MIT Press; 2000.
3. Kalisch M, Bühlmann P. Estimating high-dimensional directed acyclic graphs with the PC-algorithm. Journal of Machine Learning Research. 2007;8:613-636.
4. Margaritis D. Learning Bayesian network model structure from data. PhD thesis. Carnegie Mellon University; 2003.
5. Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 1974;19(6):716-723. https://doi.org/10.1109/TAC.1974.1100705
6. Schwarz G. Estimating the dimension of a model. Annals of Statistics. 1978;6(2):461-464. https://doi.org/10.1214/aos/1176344136
7. Chickering DM. Optimal structure identification with greedy search. Journal of Machine Learning Research. 2002;3:507-554.
8. Glover F. Tabu search—part I. ORSA Journal on Computing. 1989;1(3):190-206. https://doi.org/10.1287/ijoc.1.3.190
9. Cooper GF, Herskovits E. A Bayesian method for the induction of probabilistic networks from data. Machine Learning. 1992;9(4):309-347. https://doi.org/10.1007/BF00994110
10. Tsamardinos I, Brown LE, Aliferis CF. The max-min hill-climbing Bayesian network structure learning algorithm. Machine Learning. 2006;65:31-78. https://doi.org/10.1007/s10994-006-6889-7
11. Scanagatta M, Salmerón A, Stella F. A survey on Bayesian network structure learning from data. Progress in Artificial Intelligence. 2019;8:425-439. https://doi.org/10.1007/s13748-019-00194-y
12. Kitson NK, Constantinou AC, Guo Z, Liu Y, Chobtham K. A survey of Bayesian network structure learning. Artificial Intelligence Review. 2023;56:8721-8814. https://doi.org/10.1007/s10462-022-10351-w
13. Vowels MJ, Camgoz NC, Bowden R. D'ya like DAGs? A survey on structure learning and causal discovery. ACM Computing Surveys. 2022;55(4):1-36. https://doi.org/10.1145/3527154
14. Ramazi P, Kunegel-Lion M, Greiner R, Lewis MA. Exploiting the full potential of Bayesian networks in predictive ecology. Methods in Ecology and Evolution. 2021;12(1):135-149. https://doi.org/10.1111/2041-210X.13509
15. Drury B, Valverde-Rebaza J, Moura M-F, de Andrade Lopes A. A survey of the applications of Bayesian networks in agriculture. Engineering Applications of Artificial Intelligence. 2017;65:29-42. https://doi.org/10.1016/j.engappai.2017.07.003
16. Scutari M, Graafland CE, Gutiérrez JM. Who learns better Bayesian network structures: accuracy and speed of structure learning algorithms. International Journal of Approximate Reasoning. 2019;115:235-253. https://doi.org/10.1016/j.ijar.2019.10.003
17. Gámez JA, Mateo JL, Puerta JM. A fast hill-climbing algorithm for Bayesian networks structure learning. In: Symbolic and Quantitative Approaches to Reasoning with Uncertainty. Lecture Notes in Computer Science, vol 4664. Berlin: Springer; 2007. p. 585-597. https://doi.org/10.1007/978-3-540-75256-1_52
18. Ji J, Zhang H, Hu R, Liu C. A Tabu-search based Bayesian network structure learning algorithm. Journal of Beijing University of Technology. 2011;37(8):1274-1280. https://doi.org/10.3969/j.issn.0254-0037.2011.08.025
19. Cheng J, Bell DA, Liu W. An algorithm for Bayesian network construction from data. In: Madigan D, Smyth P, editors. Proceedings of the Sixth International Workshop on Artificial Intelligence and Statistics. 1997. p. 83-90. https://proceedings.mlr.press/r1/cheng97a.html
20. Koller D, Friedman N. Probabilistic graphical models: principles and techniques. Cambridge: MIT Press; 2009.
21. Heckerman D, Geiger D, Chickering DM. Learning Bayesian networks: the combination of knowledge and statistical data. Machine Learning. 1995;20:197-243. https://doi.org/10.1007/BF00994016
22. Pineau J, Vincent-Lamarre P, Sinha K, Larivière V, Beygelzimer A, d'Alché-Buc F, et al. Improving reproducibility in machine learning research: a report from the NeurIPS 2019 reproducibility program. Journal of Machine Learning Research. 2021;22(164):1-20.
23. Pineau J. The machine learning reproducibility checklist. 2020. https://www.cs.mcgill.ca/~jpineau/ReproducibilityChecklist.pdf
24. DataCite. DataCite metadata schema documentation for the publication and citation of research data and other research outputs. https://schema.datacite.org/

doi:10.20944/preprints202606.1426.v1

25. FAIRsharing. FAIRsharing registry. https://fairsharing.org/
26. re3data.org. Registry of research data repositories. https://www.re3data.org/
27. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data. 2016;3:160018. https://doi.org/10.1038/sdata.2016.18