

doi: 10.17586/2226-1494-2024-24-4-608-614

УДК 519.226.3

Моделирование многомерных данных с помощью композитных байесовских сетей

Ирина Юрьевна Деева¹, Каринэ Артуровна Шахкян²✉, Юрий Константинович Каминский³

^{1,2,3} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

¹ iriny.deeva@gmail.com, <https://orcid.org/0000-0001-8679-5868>

² kshahkyan@yandex.ru✉, <https://orcid.org/0009-0003-2606-431X>

³ jkaminski@niuitmo.ru, <https://orcid.org/0009-0006-6418-6117>

Аннотация

Введение. Представлен новый подход к моделированию нелинейных зависимостей, названный композитными байесовскими сетями. Основной акцент сделан на интеграции моделей машинного обучения в байесовские сети с сохранением их основополагающих принципов. Новизна предложенного подхода состоит в том, что он позволяет решить проблему несоответствия данных традиционным предположениям о зависимостях.

Метод. Представленный подход заключается в подборе разнообразных моделей машинного обучения на этапе обучения композитных байесовских сетей. Это позволяет гибко настраивать характер зависимостей в соответствии с требованиями и продиктованными характеристиками моделируемого объекта. Программная реализация подхода выполнена в виде специализированного фреймворка, описывающего все необходимые функциональные возможности. **Основные результаты.** Проведена экспериментальная оценка эффективности моделирования зависимостей между признаками. Для экспериментов выбраны для бенчмарков и из репозитория UCI для реальных данных. Эффективность предложенных композитных байесовских сетей подтверждена сравнением значений правдоподобия и показателя F1 с классическими байесовскими сетями, обученными алгоритмом Hill-Climbing. Показана высокая точность представления многомерных распределений. При этом на бенчмарках улучшение оказалось незначительным, поскольку они содержат линейные зависимости, которые хорошо моделируются классическими алгоритмами. На реальных наборах данных UCI получено улучшение правдоподобия в среднем на 30 %. **Обсуждение.** Полученные результаты могут найти применение в областях, требующих моделирования сложных зависимостей между признаками, например, в машинном обучении, статистике, задачах анализа данных, а также в конкретных предметных областях.

Ключевые слова

байесовские сети, вероятностные графовые модели, обучение параметров, модели машинного обучения, генетический алгоритм

Благодарности

Работа выполнена в рамках государственного задания Министерства науки и высшего образования Российской Федерации (проект № FSER-2024-0004).

Ссылка для цитирования: Деева И.Ю., Шахкян К.А., Каминский Ю.К. Моделирование многомерных данных с помощью композитных байесовских сетей // Научно-технический вестник информационных технологий, механики и оптики. 2024. Т. 24, № 4. С. 608–614. doi: 10.17586/2226-1494-2024-24-4-608-614

Flexible and tractable modeling of multivariate data using composite Bayesian networks

Irina Yu. Deeva¹, Karine A. Shakhkyan²✉, Yury K. Kaminsky³

^{1,2,3} ITMO University, Saint Petersburg, 197101, Russian Federation

¹ iriny.deeva@gmail.com, <https://orcid.org/0000-0001-8679-5868>

² kshahkyan@yandex.ru✉, <https://orcid.org/0009-0003-2606-431X>

³ jkaminski@niuitmo.ru, <https://orcid.org/0009-0006-6418-6117>

© Деева И.Ю., Шахкян К.А., Каминский Ю.К., 2024

Abstract

The article presents a new approach to modeling nonlinear dependencies called composite Bayesian networks. The main emphasis is on integrating machine learning models into Bayesian networks while maintaining their fundamental principles. The novelty of the approach is that it allows us to solve the problem of data inconsistency with traditional assumptions about dependencies. The presented method consists in selecting a variety of machine learning models at the stage of training composite Bayesian networks. This allows you to flexibly customize the nature of the dependencies in accordance with the requirements and dictated characteristics of the modeled object. The software implementation is made in the form of a specialized framework that describes all the necessary functionality. The results of experiments to evaluate the effectiveness of modeling dependencies between features are presented. Data for the experiments was taken from the blearn repository for benchmarks and from the UCI repository for real data. The performance of composite Bayesian networks was validated by comparing the likelihood and F1 score with classical Bayesian networks trained with the Hill-Climbing algorithm, demonstrating high accuracy in representing multivariate distributions. The improvement in benchmarks is insignificant since they contain linear dependencies that are well modeled by the classical algorithm. An average 30 % improvement in likelihood was obtained on real UCI datasets. The obtained data can be applied in areas that require modeling complex dependencies between features, for example, in machine learning, statistics, data analysis, as well as in specific subject areas.

Keywords

Bayesian networks, probabilistic graph models, parameter learning, machine learning models, genetic algorithm

Acknowledgements

The research was carried out within the state assignment of the Ministry of Science and Higher Education of the Russian Federation (project No. FSER-2024-0004).

For citation: Deeva I.Yu., Shakhkyan K.A., Kaminsky Yu.K. Flexible and tractable modeling of multivariate data using composite Bayesian networks. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2024, vol. 24, no. 4, pp. 608–614 (in Russian). doi: 10.17586/2226-1494-2024-24-4-608-614

Введение

Современные методологии искусственного интеллекта конкурируют с традиционными моделями, включая байесовские сети, которые используются для представления данных через условные распределения, определяемые направленными ациклическими графами [1]. В работе [1] рассмотрены ограничения классических байесовских сетей: гибкость (несоответствие между данными и предположениями о характере зависимости данных); контролируемость (невозможность влиять на зависимость в условных распределениях); неэффективное представление параметров распределения, часто связанное с большими таблицами условных вероятностей [2, 3]. В настоящей работе представлен новый подход — композитные байесовские сети (КБС), которые расширяют классические модели, сохраняя их достоинства и добавляя возможность моделировать все параметры с помощью методов машинного обучения. Пример такой КБС показан на рис. 1.

Большинство байесовских сетей, описанных в тематических научных работах, относятся к категории параметрических моделей, т. е. они строятся с использованием параметрических условных распределений вероятностей с фиксированным числом параметров. Среди них наиболее распространены дискретные байесовские сети, специально разработанные для моделирования дискретных признаков [4]. В то же время гауссовские байесовские сети представляют собой отдельный подтип, специально разработанный для применения в непрерывных байесовских сетях. В гауссовских байесовских сетях каждый узел использует линейные гауссовские условные распределения вероятностей [5].

Гибридные байесовские сети могут моделировать как дискретные, так и непрерывные признаки. Одним из наиболее известных типов таких сетей является

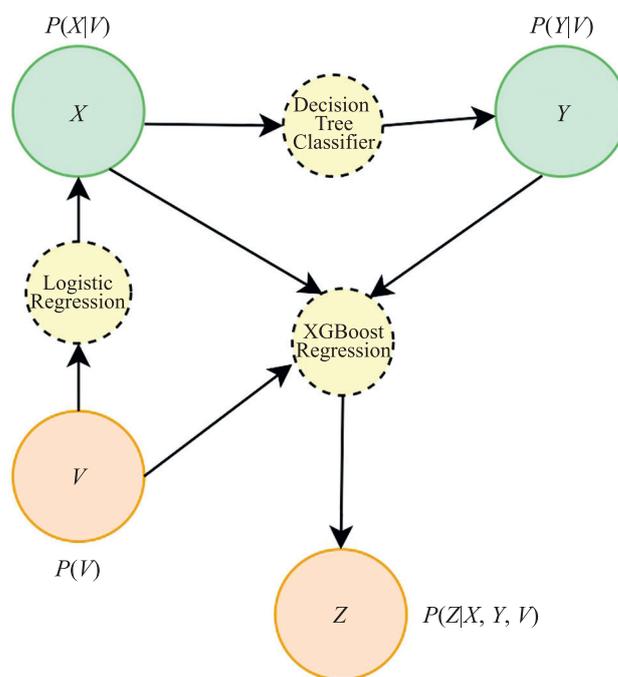


Рис. 1. Пример композитной байесовской сети.

Зеленые узлы — дискретные признаки, оранжевые — непрерывные признаки, желтые — модели, используемые для моделирования параметров условных распределений; X, Y, V, Z — узлы байесовской сети; P — вероятность; Decision Tree Classifier — классификатор дерева решений; Logistic Regression — логистическая регрессия; XGBoost Regression — XGBoost перрессия

Fig. 1. An example of a composite Bayesian network.

Green nodes are discrete features, orange nodes are continuous features, yellow nodes are models used to model the parameters of conditional distributions; X, Y, V, Z — Bayesian network nodes; P — probability

условная линейная гауссовская байесовская сеть [6]. В этой модели дискретные переменные представлены таблицами условных вероятностей, а непрерывные — линейными гауссовскими распределениями.

Исследователи также указывают на ограничения байесовских сетей, связанные с гибкостью, контролируемостью и большими таблицами условных вероятностей. Например, непараметрические байесовские сети используют оценку плотности ядра для моделирования условных распределений [7, 8], обеспечивая заметную гибкость. Однако вычислительные требования для непараметрической оценки могут быть значительными, что требует дополнительных процедур для обучения параметра сглаживания. Существуют и параметрические решения, включающие использование смесей гауссовых распределений [9] или смесей усеченных базисных функций [10]. Тем не менее, эти методы затрачивают слишком много времени. Чтобы упростить параметрическое обучение, некоторые ученые выступают за использование полупараметрических моделей [11], в которых только определенные узлы представлены в виде непараметрических моделей. Такой подход значительно сокращает время обучения и семплирования, но не решает проблему, связанную с хранением больших таблиц условных вероятностей для дискретных распределений. Для решения подобных задач также используются каузальные модели, которые описывают причинно-следственные отношения между признаками [12]. Связи в таких моделях моделируются с помощью предиктивной модели с добавлением случайного шума. Однако каузальные модели могут быть чувствительны к выбросам, в то время как КБС более устойчивы к выбросам и имеют вероятностный подход.

Новизна настоящей работы включает следующие аспекты: КБС, объединяющие полезные свойства моделей машинного обучения и классических вероятностных моделей; представление алгоритма для автоматического выбора моделей машинного обучения с помощью эволюционной оптимизации; процедуру семплирования, адаптированную для КБС; разработку программного обеспечения в виде комплексной структуры для изучения и применения КБС на различных типах данных, включая непрерывные, дискретные и смешанные данные. Эффективность предложенного метода, который устраняет вышеупомянутые ограничения, подтверждается экспериментальными результатами.

Постановка проблемы

В основе байесовских сетей лежит направленный ациклический граф, который задается как $G = \{V, E\}$. Однако в предлагаемой композитной формулировке эта структура дополнена дополнительной категорией вершин M и задается как $G = \{V, E, M\}$. В данном контексте V — набор узлов-признаков, в которых моделируются распределения вероятностей, M — набор узлов-моделей, представляющих модели машинного обучения, которые устанавливают связи с узлами-признаками, а E — набор направленных ребер. Процесс

обучения КБС формулируется как оптимизационная задача с функцией оценки F , в виде:

$$M_{opt}, E_{opt} = \operatorname{argmax}_{M, E} F(G).$$

Предлагаемый подход. В контексте КБС модели машинного обучения играют ключевую роль в вычислении параметров условных распределений, выраженных в виде $\Theta_{V_{child}|V_{parent}} = M(V_{parent})$, где V_{child} — узел дочернего признака; V_{parent} — узел родительского признака. Отметим, что Θ — параметры распределения в узле-признаке. Массив потенциальных моделей машинного обучения зависит от типов узлов-признаков, интегрированных в модель. Эти узлы могут принимать непрерывную и дискретную формы, что зависит от характеристик, присущих моделируемому признаку. Например, если два признака являются непрерывными, они могут быть связаны между собой в рамках любой модели машинного обучения, способной решить задачу регрессии. Рассчитаем параметры условного непрерывного распределения для связи между узлами X и Y , сформулированной как $X \rightarrow m \rightarrow Y$.

$$\mu_{Y|X} = m(x), \operatorname{var}_{Y|X} = (m(x) - y)^2.$$

Аналогично вычислениям параметров условных распределений получим общий вид функции оценки:

$$F(G) = \sum_{i=1}^n LL(\theta_i | v_i) + \sum_{j=1}^k LL(m_j(v_{parents_j}) | v_j), \quad (1)$$

где LL — лог-правдоподобие; n — количество узлов без родителей; θ — параметры распределения в узлах без родителей (маргинальные узлы); k — количество узлов с родителями; m — узел модели машинного обучения.

Ядро алгоритма

Для решения задачи структурного и параметрического обучения КБС в качестве стратегии оптимизации используем эволюционный алгоритм. Эволюционные алгоритмы, которые часто применяются для решения подобных задач, могут быть подходящим алгоритмом оптимизации, однако успешность результата во многом зависит от заданных генетических операторов и других параметров [13]. В данной постановке необходимо задать правила редактирования ребер графа. Редактирование ребер требует трех различных спецификаций: структуры (изменение направлений ребер графа); моделей (изменение модели); гибридное (изменение структуры и модели).

Исходя из характера моделируемых признаков, выберем набор допустимых моделей машинного обучения. Эта начальная спецификация позволяет накладывать ограничения, например, использовать монотонно возрастающие модели в соответствии с требованиями области. Такая *управляемость* полезна для работы с теми областями, где зависимости известны. Далее определим правила проверки валидности моделей, включая отсутствие циклов и уникальность признаков.

Для оптимизации опишем специальные правила редактирования моделей. Критерием завершения работы

алгоритма является стабилизация значения функции оценки. Функция оценки в формуле (1) использована в качестве метрики. После окончания подготовки начался процесс оптимизации графа, включая итеративную оценку параметров распределений в узлах-признаках для вычисления функции оценки.

Дополнительно, в рамках разработанных алгоритмов, представлен алгоритм семплирования для КБС (рис. 2). Этот алгоритм, который называется предковым семплированием от корней к листьям, разработан для работы в рамках устоявшейся структуры. При наличии наблюдаемых признаков в родительских узлах алгоритм использует модель машинного обучения, связывающую родительские и дочерние узлы, для предсказания параметров условного распределения. Для непрерывных узлов это среднее значение, а для дискретных — вычисляются условные вероятности. Затем, используя предсказанные параметры условного распределения, алгоритм выполняет семплирование. Для непрерывных узлов используются гауссовские распределения, а для дискретных — мультиномиальные.

Программная реализация

Предложенный алгоритм, предназначенный для оптимизации композитных моделей, интегрирован в более крупный фреймворк с открытым исходным ко-

дом ВАМТ¹ для моделирования байесовских сетей. Алгоритм разработан с особой точностью, обеспечивая легкое слияние процесса оптимизации с существующими компонентами фреймворка.

В рамках фреймворка ВАМТ генетический алгоритм использует ряд сложных операций — от кроссоверов, позволяющих обменивать родительские узлы-модели между композитными структурами, до мутаций, которые облегчают добавление, удаление и изменение структур, а также встраивание новых моделей. Эти операции были оптимизированы для обеспечения эффективности и точности, гарантируя, что композитные модели эволюционируют в направлении наилучшей возможной конфигурации для любого набора данных.

Важнейшим элементом эволюционной оптимизации в рамках ВАМТ является фреймворк GOLEM. Разработанный в лаборатории композитного искусственного интеллекта Университета ИТМО GOLEM² служит основой, обрабатывая генетические операции и обеспечивая правильную оптимизацию композитных моделей.

¹ [Электронный ресурс]. Режим доступа: <https://github.com/aimclub/VAMT> (дата обращения: 03.05.2024).

² [Электронный ресурс]. Режим доступа: <https://github.com/aimclub/GOLEM> (дата обращения: 03.05.2024).

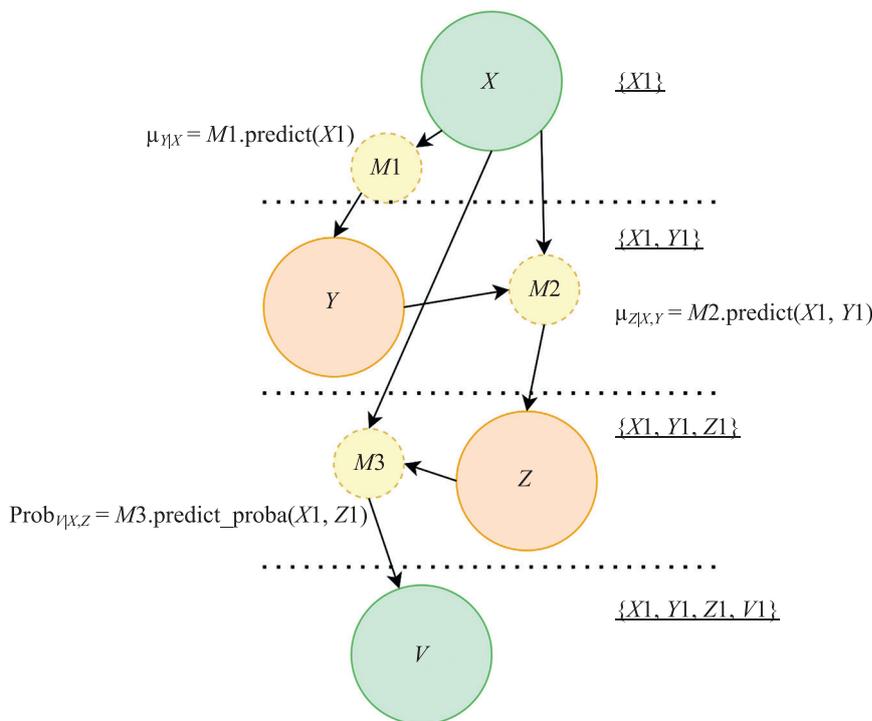


Рис. 2. Схема процесса вероятностного вывода для композитной байесовской сети.

Зеленые узлы — дискретные признаки, оранжевые — непрерывные признаки, желтые — модели, используемые для моделирования параметров условных распределений; X_1, Y_1, V_1, Z_1 — семплированные значения узлов байесовской сети X, Y, V, Z соответственно; M_1, M_2, M_3 — модели машинного обучения

Fig. 2. Schematic illustration of the probabilistic inference process for a composite Bayesian network.

Green nodes are discrete features, orange nodes are continuous features, yellow nodes are models used to model the parameters of conditional distributions; X_1, Y_1, V_1, Z_1 — sampled values of Bayesian network nodes X, Y, V, Z , respectively; M_1, M_2, M_3 — machine learning models

Одной из наиболее примечательных особенностей наших композитных моделей является эффективность использования памяти. Например, КБС из 8 узлов на основе набора данных Asia, сохраненная в формате JSON, занимает в четыре раза меньше места, чем классическая байесовская сеть. Такая разница в объеме памяти впечатляет не только для небольших сетей, но становится значительно более заметной для больших сетей, особенно для тех, которые работают со смешанными данными.

Примеры и экспериментальные исследования

Экспериментальные исследования КБС выполнены на двух наборах данных. Бенчмарки из репозитория bnlearn¹ и реальные наборы данных из репозитория UCI². Проведено сравнение результатов эксперимента с классическим подходом, когда для обучения структуры используется алгоритм Hill-Climbing [14] с функцией K2, а все распределения представлены одинаково, с помощью линейных зависимостей или таблиц условных вероятностей (HCbn) [15]. Результаты сравнивались с позиций качества структурного и параметрического обучений. Для измерения качества структурного обучения использован показатель F1 между эталонной и полученной структурами (чем F1 выше, тем более похожа полученная структура на эталонную). Для измерения качества параметрического обучения вычислено правдоподобие на тестовом наборе данных.

Алгоритм обучения КБС предлагает три режима работы. В первом режиме (одновременное обучение) структура и параметры в виде моделей машинного обучения подбираются одновременно с помощью методов эволюционной оптимизации (compositeGA). Второй режим (многокритериальное обучение) предполагает выбор структуры на основе определенного структурного критерия (например, K2 [16]), в то время как модели машинного обучения выбираются в соответствии с формулой (1) (multiGA). Во втором режиме целью является определение оптимального решения

на фронте Парето. Третий режим (последовательное обучение) использует пошаговый подход. Сначала обучается структура — обычно это делается с помощью алгоритмов типа Hill-Climbing, а затем выбираются модели машинного обучения (seqGA).

Для работы генетического алгоритма необходимо задать его гиперпараметры: количество индивидов — 20; количество поколений — 1000 (большое значение, чтобы алгоритм сошелся); вероятность мутации — 0,9; вероятность кроссовера — 0,8; критерий остановки — нет улучшений метрики в течение 10 поколений; тайм-аут — 15 мин.

Результаты экспериментов на бенчмарках. Для проведения экспериментов на бенчмарках были выбраны наборы данных, содержащие различные типы данных (дискретные и непрерывные). В табл. 1 представлены результаты обучения.

Результаты обучения показали, что КБС имеют высокое правдоподобие, но улучшение по сравнению с базовым вариантом невелико. Это объясняется спецификой используемых бенчмарков, имеющих нормальные распределения и линейные зависимости. Однако КБС выделяются в качестве структурного обучения, особенно благодаря многокритериальному подходу и использованию эволюционной оптимизации. Даже при сопоставимом правдоподобии они эффективно распознают структуры, предположительно, благодаря эффективному исследованию эволюционным алгоритмом пространства потенциальных структур.

На рис. 3 показан график сходимости для трех режимов обучения КБС. Видно, что для подхода, когда структура уже обучена и выбраны только модели машинного обучения (seqGA), алгоритм сходится быстрее.

Эксперименты на реальных данных. Чтобы подробнее изучить преимущества КБС, выполнен эмпирический анализ на реальных наборах данных UCI, включающих в себя сочетание непрерывных и дискретных признаков. Из-за отсутствия эталонных структур для этих наборов данных сравнение проводилось только на основе метрики правдоподобия. Результаты, представленные в табл. 2, продемонстрировали заметное превосходство КБС. Отметим, что постоянно выигрывает подход, при котором и структура и модели выбираются с использованием одной и той же композитной метрики. Эти результаты подчеркивают способность предло-

¹ [Электронный ресурс]. Режим доступа: <https://www.bnlearn.com/bnrepository> (дата обращения: 03.05.2024).

² [Электронный ресурс]. Режим доступа: <https://archive.ics.uci.edu/datasets> (дата обращения: 03.05.2024).

Таблица 1. Сравнение результатов обучения байесовских сетей для трех режимов композитных байесовских сетей и базового жадного алгоритма (HCbn) для бенчмарков

Table 1. Comparison of Bayesian Network learning results for three modes of composite Bayesian Network and the baseline — greedy algorithm (HCbn) for benchmarks

Наборы данных	Правдоподобие				F1			
	CompGA	multiGA	consqGA	HCbn	CompGA	multiGA	consqGA	HCbn
Asia	-451	-454	-451	-456	0,49	0,67	0,29	0,29
Cancer	-417	-419	-418	-426	0,39	0,57	0,57	0,57
Earthquake	-100	-101	-101	-93	0,57	0,93	1	1
Sachs	-1469	-1470	-1427	-1545	0,38	0,6	0,53	0,53
Sangiovese	984	1005	1013	988	0,22	0,2	0,17	0,17

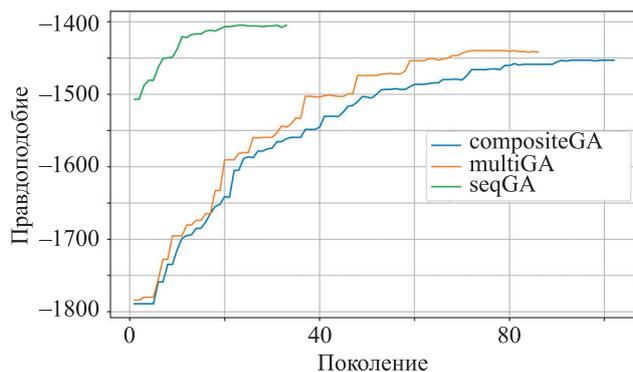


Рис. 3. Сходимость трех режимов обучения составной байесовской сети на примере набора данных Sachs

Fig. 3. Convergence of three training modes for a composite Bayesian network using the example of the Sachs dataset

Таблица 2. Сравнение результатов обучения байесовских сетей для трех режимов КБС и базового жадного алгоритма (HCbn) для наборов данных UCI

Table 2. Comparison of Bayesian Network learning results for three modes of composite Bayesian Network and the baseline — greedy algorithm (HCbn) for UCI datasets

Наборы данных	CompGA	multiGA	consqGA	HCbn
Abalone	1816	1803	1833	1732
Adult	-3527	-4666	-8361	-11041
Australian_statlog	-3143	-3196	-5022	-5146
Liver_disorders	-1382	-1494	-1759	-1809

Литература

- Handbook of Graphical Models / ed. by M. Maathuis, M. Drton, S. Lauritzen, M. Wainwright. CRC Press, 2018. 554 p. <https://doi.org/10.1201/9780429463976>
- Mascaro S., Nicholso A.E., Korb K.B. Anomaly detection in vessel tracks using Bayesian networks // *International Journal of Approximate Reasoning*. 2014. V. 55. N 1. P. 84–98. <https://doi.org/10.1016/j.ijar.2013.03.012>
- McLachlan S., Dube K., Hitman G.A., Fenton N.E., Kyrimi E. Bayesian networks in healthcare: Distribution by medical condition // *Artificial Intelligence in Medicine*. 2020. V. 107. P. 101912. <https://doi.org/10.1016/j.artmed.2020.101912>
- Friedman N., Goldszmidt M. Learning Bayesian networks with local structure // *NATO ASI Series*. 1998. V. 89. P. 421–459. https://doi.org/10.1007/978-94-011-5014-9_15
- Grzegorzczak M. An introduction to gaussian bayesian networks // *Methods in Molecular Biology*. 2010. V. 662. P. 121–147. https://doi.org/10.1007/978-1-60761-800-3_6
- Lerner U., Segal E., Koller D. Exact inference in networks with discrete children of continuous parents // *arXiv*. 2013. arXiv:1301.2289. <https://doi.org/10.48550/arXiv.1301.2289>
- Pérez A., Larrañaga P., Inza I. Bayesian classifiers based on kernel density estimation: Flexible classifiers // *International Journal of Approximate Reasoning*. 2009. V. 50. N 2. P. 341–362. <https://doi.org/10.1016/j.ijar.2008.08.008>
- Ickstadt K., Bornkamp B., Grzegorzczak M., Wiecezorek J., Sheriff M.R., Grecco H.E., Zamir E. Nonparametric Bayesian networks // *Bayesian Statistics 9*. 2011. P. 283–316. <https://doi.org/10.1093/acprof:oso/9780199694587.003.0010>
- Deeva I., Bubnova A., Kalyuzhnaya A.V. Advanced approach for distributions parameters learning in Bayesian networks with gaussian mixture models and discriminative models // *Mathematics*. 2023. V. 11. N 2. P. 343. <https://doi.org/10.3390/math11020343>
- Langseth H., Nielsen T.D., Rumí R., Salmerón A. Mixtures of truncated basis functions // *International Journal of Approximate*

женной методологии эффективно отражать нелинейные зависимости в сложных наборах данных.

Заключение

Представлена новая категория байесовских сетей, известная как композитные байесовские сети, в которых для аппроксимации параметров условных распределений используются модели машинного обучения. Представлены алгоритмы, облегчающие обучение и семплирование композитных байесовских сетей. Описан фреймворк ВАМТ, который включает в себя все основные функциональные возможности для работы с композитными байесовскими сетями. Экспериментальные результаты показали уникальную гибкость и практичность рассмотренного подхода при моделировании условных распределений, что особенно заметно на реальных данных, которые часто не соответствуют классическим предположениям о распределениях и зависимостях. Кроме того, композитный подход обеспечивает эффективное хранение обученных моделей, устраняя необходимость в больших таблицах условных вероятностей и тем самым уменьшая размер модели в четыре раза. Таким образом, композитные байесовские сети представляют собой инновационный инструмент, позволяющий активно применять вероятностные методы искусственного интеллекта к реальным наборам данных.

References

- Handbook of Graphical Models*. Ed. by M. Maathuis, M. Drton, S. Lauritzen, M. Wainwright. CRC Press, 2018, 554 p. <https://doi.org/10.1201/9780429463976>
- Mascaro S., Nicholso A.E., Korb K.B. Anomaly detection in vessel tracks using Bayesian networks. *International Journal of Approximate Reasoning*, 2014, vol. 55, no. 1, pp. 84–98. <https://doi.org/10.1016/j.ijar.2013.03.012>
- McLachlan S., Dube K., Hitman G.A., Fenton N.E., Kyrimi E. Bayesian networks in healthcare: Distribution by medical condition. *Artificial Intelligence in Medicine*, 2020, vol. 107, pp. 101912. <https://doi.org/10.1016/j.artmed.2020.101912>
- Friedman N., Goldszmidt M. Learning Bayesian networks with local structure. *NATO ASI Series*, 1998, vol. 89, pp. 421–459. https://doi.org/10.1007/978-94-011-5014-9_15
- Grzegorzczak M. An introduction to gaussian bayesian networks. *Methods in Molecular Biology*, 2010, vol. 662, pp. 121–147. https://doi.org/10.1007/978-1-60761-800-3_6
- Lerner U., Segal E., Koller D. Exact inference in networks with discrete children of continuous parents. *arXiv*, 2013, arXiv:1301.2289. <https://doi.org/10.48550/arXiv.1301.2289>
- Pérez A., Larrañaga P., Inza I. Bayesian classifiers based on kernel density estimation: Flexible classifiers. *International Journal of Approximate Reasoning*, 2009, vol. 50, no. 2, pp. 341–362. <https://doi.org/10.1016/j.ijar.2008.08.008>
- Ickstadt K., Bornkamp B., Grzegorzczak M., Wiecezorek J., Sheriff M.R., Grecco H.E., Zamir E. Nonparametric Bayesian networks. *Bayesian Statistics 9*, 2011, pp. 283–316. <https://doi.org/10.1093/acprof:oso/9780199694587.003.0010>
- Deeva I., Bubnova A., Kalyuzhnaya A.V. Advanced approach for distributions parameters learning in Bayesian networks with gaussian mixture models and discriminative models. *Mathematics*, 2023, vol. 11, no. 2, pp. 343. <https://doi.org/10.3390/math11020343>
- Langseth H., Nielsen T.D., Rumí R., Salmerón A. Mixtures of truncated basis functions. *International Journal of Approximate*

- Reasoning, 2012, V. 53, N 2, P. 212–227. <https://doi.org/10.1016/j.ijar.2011.10.004>
11. Atienza D., Larrañaga P., Bielza C. Hybrid semiparametric Bayesian networks // *TEST*, 2022, V. 31, N 2, P. 299–327. <https://doi.org/10.1007/s11749-022-00812-3>
 12. Sloman S. Causal Models: How People Think about the World and Its Alternatives. Oxford University Press, 2005. 211 p. <https://doi.org/10.1093/acprof:oso/9780195183115.001.0001>
 13. Larrañaga P., Karshenas H., Bielza C., Santana R. A review on evolutionary algorithms in Bayesian network learning and inference tasks // *Information Sciences*, 2013, V. 233, P. 109–125. <https://doi.org/10.1016/j.ins.2012.12.051>
 14. Gámez J.A., Mateo J.L., Puerta J.M. Learning Bayesian networks by hill climbing: efficient methods based on progressive restriction of the neighborhood // *Data Mining and Knowledge Discovery*, 2011, V. 22, N 1-2, P. 106–148. <https://doi.org/10.1007/s10618-010-0178-6>
 15. Behjati S., Beigy H. Improved K2 algorithm for Bayesian network structure learning // *Engineering Applications of Artificial Intelligence*, 2020, V. 91, P. 103617. <https://doi.org/10.1016/j.engappai.2020.103617>
 16. Lerner B., Malka R. Investigation of the K2 algorithm in learning Bayesian network classifiers // *Applied Artificial Intelligence*, 2011, V. 25, N 1, P. 74–96. <https://doi.org/10.1080/08839514.2011.529265>
- Reasoning*, 2012, vol. 53, no. 2, pp. 212–227. <https://doi.org/10.1016/j.ijar.2011.10.004>
11. Atienza D., Larrañaga P., Bielza C. Hybrid semiparametric Bayesian networks. *TEST*, 2022, vol. 31, no. 2, pp. 299–327. <https://doi.org/10.1007/s11749-022-00812-3>
 12. Sloman S. Causal Models: How People Think about the World and Its Alternatives. *Oxford University Press*, 2005, 211 p. <https://doi.org/10.1093/acprof:oso/9780195183115.001.0001>
 13. Larrañaga P., Karshenas H., Bielza C., Santana R. A review on evolutionary algorithms in Bayesian network learning and inference tasks. *Information Sciences*, 2013, vol. 233, pp. 109–125. <https://doi.org/10.1016/j.ins.2012.12.051>
 14. Gámez J.A., Mateo J.L., Puerta J.M. Learning Bayesian networks by hill climbing: efficient methods based on progressive restriction of the neighborhood. *Data Mining and Knowledge Discovery*, 2011, vol. 22, no. 1-2, pp. 106–148. <https://doi.org/10.1007/s10618-010-0178-6>
 15. Behjati S., Beigy H. Improved K2 algorithm for Bayesian network structure learning. *Engineering Applications of Artificial Intelligence*, 2020, vol. 91, pp. 103617. <https://doi.org/10.1016/j.engappai.2020.103617>
 16. Lerner B., Malka R. Investigation of the K2 algorithm in learning Bayesian network classifiers. *Applied Artificial Intelligence*, 2011, vol. 25, no. 1, pp. 74–96. <https://doi.org/10.1080/08839514.2011.529265>

Авторы

Деева Ирина Юрьевна — кандидат физико-математических наук, старший научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57210416999](https://orcid.org/0000-0001-8679-5868), <https://orcid.org/0000-0001-8679-5868>, iriny.deeva@gmail.com

Шахкян Каринэ Артуровна — инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0009-0003-2606-431X>, kshahkyan@yandex.ru

Каминский Юрий Константинович — инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0009-0006-6418-6117>, jkaminski@niuitmo.ru

Статья поступила в редакцию 03.05.2024
Одобрена после рецензирования 05.06.2024
Принята к печати 19.07.2024

Authors

Irina Yu. Deeva — PhD (Physics & Mathematics), Senior Researcher, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57210416999](https://orcid.org/0000-0001-8679-5868), <https://orcid.org/0000-0001-8679-5868>, iriny.deeva@gmail.com

Karine A. Shakhkyan — Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0009-0003-2606-431X>, kshahkyan@yandex.ru

Yury K. Kaminsky — Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0009-0006-6418-6117>, jkaminski@niuitmo.ru

Received 03.05.2024
Approved after reviewing 05.06.2024
Accepted 19.07.2024



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»