

Rozdział 3

Inżynieria wiedzy w konstruowaniu modeli sieci bayesowskich w medycynie

Agnieszka Oniśko

Wydział Informatyki, Politechnika Białostocka

Streszczenie: Modele sieci bayesowskich znalazły zastosowania zarówno w medycznym diagnozowaniu, jak i prognozowaniu. Inżynieria wiedzy w budowaniu modeli sieci bayesowskich sprowadza się między innymi do pozyskiwania wiedzy w danej dziedzinie, transformacji tej wiedzy do acyklicznego grafu skierowanego oraz wyznaczania warunkowych rozkładów prawdopodobieństwa. Źródłem wiedzy może być opinia eksperta, bądź też dane medyczne. Ważną rolę odgrywa w tym procesie inżynier wiedzy, który powinien opanować techniki pozyskiwania wiedzy od eksperta, rozumieć modelowaną dziedzinę, posiadać umiejętności przetwarzania danych medycznych, również w sytuacji, gdy są one niekompletne. Pomimo, że istnieją metody automatycznego wyznaczania struktury i parametrów numerycznych modeli sieci bayesowskich, w przypadku medycyny trudno jest zbudować model bez pomocy eksperta oraz inżyniera wiedzy, który ma wgląd w modelowaną dziedzinę. Większość inżynierów wiedzy uważa, że sieci bayesowskie nie są czułe na dokładność ich parametrów numerycznych i znacznie ważniejsza jest ich struktura graficzna. W niniejszym rozdziale przedstawiono wyniki eksperymentów, w których badane były: (1) wpływ struktury oraz (2) wpływ dokładności parametrów numerycznych na jakość diagnostyczną modeli sieci bayesowskich. Opisano również techniki i metody, z których warto korzystać w trakcie konstruowania modeli sieci bayesowskich, szczególnie w kontekście ich parametrów numerycznych.

Słowa kluczowe: sieci bayesowskie, inżynieria wiedzy, parametry numeryczne, czułość

Wprowadzenie

Sieci bayesowskie [1] w ciągu ostatnich trzech dekad stały się popularnym narzędziem do modelowania trudnych problemów w warunkach niepewności. Modele sieci bayesowskich znalazły zastosowania zarówno w diagnozowaniu, jak i prognozowaniu medycznym. Sieć bayesowska składa się z (1) części jakościowej, która jest acyklicznym

grafem skierowanym modelującym zmienne losowe i zależności probabilistyczne między nimi oraz z (2) części ilościowej reprezentowanej przez warunkowe rozkłady prawdopodobieństwa.

Inżynieria wiedzy w budowaniu modeli sieci bayesowskich sprowadza się między innymi do pozyskiwania wiedzy w danej dziedzinie, transformacji tej wiedzy do acyklicznego grafu skierowanego oraz wyznaczania warunkowych rozkładów prawdopodobieństwa. Ważną rolę odgrywa w tym procesie inżynier wiedzy, który powinien opanować techniki pozyskiwania istotnych informacji od eksperta, rozumieć modelowaną dziedzinę, posiadać umiejętności przetwarzania danych medycznych, również w sytuacji, gdy są one niekompletne. Pomimo, że istnieją metody automatycznego wyznaczania struktury i parametrów numerycznych modeli sieci bayesowskich, w przypadku medycyny trudno jest zbudować model bez pomocy eksperta oraz inżyniera wiedzy.

Większość inżynierów wiedzy uważa, że sieci bayesowskie nie są czułe na dokładność ich parametrów numerycznych i znacznie ważniejsza jest ich struktura graficzna. W niniejszym rozdziale przedstawiono wyniki eksperymentów, w których badane były: (1) wpływ struktury oraz (2) wpływ dokładności parametrów numerycznych na jakość diagnostyczną modeli sieci bayesowskich.

Punktem wyjścia w eksperymentach były modele diagnostyczne sieci bayesowskich wyznaczone na podstawie rzeczywistych danych medycznych. W pierwszej serii eksperymentów stopniowo (a) usuwano i (b) odwracano łuki modelu. Z kolei w eksperymentach badających wpływ dokładności parametrów numerycznych stopniowo wprowadzano różne rodzaje szumu (symetryczny, symulujący zbyt pewnego siebie i niepewnego siebie eksperta). Następnie badano wpływ modyfikacji struktury oraz dokładności parametrów numerycznych na jakość diagnostyczną modeli sieci bayesowskich.

3.1. Badanie czułości struktury modeli sieci bayesowskich

Część jakościową sieci bayesowskiej stanowi struktura modelu, czyli acykliczny graf skierowany, który może być wyznaczany w oparciu o wiedzę eksperta lub dostępne dane. W serii eksperymentów, które przeprowadzono, badano czułość struktury modeli sieci bayesowskich [2, 3]. W badaniach tych sprawdzono, jaki wpływ ma modyfikacja struktury graficznej sieci bayesowskich na jakość diagnostyczną modeli probabilistycznych.

W eksperymentach sięgnięto po 6 medycznych modeli diagnostycznych, które stworzono opierając się na danych z Irvine Machine Learning Repository [4]. Strukturę graficzną tych modeli poddano systematycznej modyfikacji. W pierwszym z eksperymentów systematycznie usuwano krawędzie modelu, z kolei w drugim eksperymentcie systematycznie odwracano kolejne krawędzie modelu. Po każdej modyfikacji struktury grafu sprawdzano jakość diagnostyczną modelu. Jakość diagnostyczną zdefiniowano

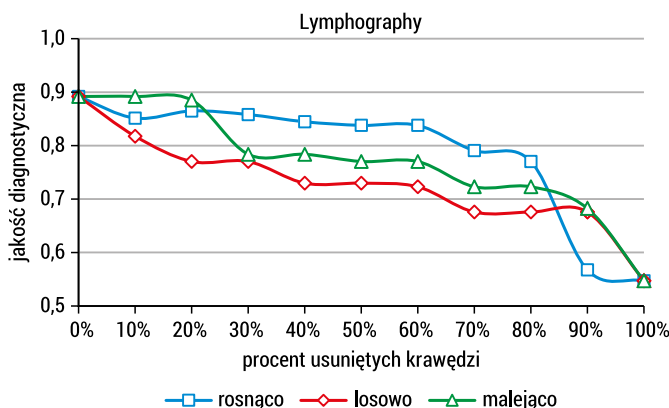
jako procent poprawnych odpowiedzi modelu sieci bayesowskiej, tzn. odpowiedź modelu była uznawana za poprawną w sytuacji, jeśli najbardziej prawdopodobna diagnoza (jednostka chorobowa z najwyższą wartością prawdopodobieństwa *a posteriori* wyznaczoną przez model) była zgodna z diagnozą zweryfikowaną w zbiorze danych. Do wyznaczenia jakości diagnostycznej zastosowano walidację krzyżową typu *leave-one-out*.

Miara mocy krawędzi w modelu sieci bayesowskiej

W eksperymentach skorzystano z miary mocy krawędzi zaproponowanej przez Koitera [5]. Miara ta wyznaczana jest na podstawie różnicy pomiędzy brzegowym prawdopodobieństwem *a posteriori* węzła, który jest dzieckiem a brzegowym prawdopodobieństwem różnych stanów węzła-rodzica. Wyznaczając odległość między tymi rozkładami skorzystano z odległości euklidesowej oraz odległości Hellingera.

Usuwanie krawędzi modelu sieci bayesowskiej

W eksperymencie tym systematycznie była modyfikowana struktura graficzna modelu przez usuwanie krawędzi grafu. Eksperyment rozpoczęto od wyznaczenia jakości diagnostycznej modelu oryginalnego, a następnie usunięto kolejno 10%, 20%, 30%,..., 90% oraz 100% krawędzi. Za każdym razem sprawdzano jakość diagnostyczną modelu. Eksperyment ten powtórzono trzy razy zmieniając za każdym razem porządek usuwania na: (1) rosnący względem miary mocy krawędzi, tj. zaczynano modyfikację od krawędzi z najniższą mocą; (2) malejący względem miary mocy krawędzi, tj. zaczynano modyfikację od krawędzi z najwyższą mocą; oraz (3) losowy.



RYSUNEK 3.1. Jakość diagnostyczna modelu Lymphography w funkcji procentu usuniętych krawędzi

Wyniki eksperymentów pokazały, że systematyczne usuwanie krawędzi grafu prowadzi do pogorszenia jakości diagnostycznej modelu sieci bayesowskiej. Usuwanie krawędzi według porządku malejącego względem miary mocy krawędzi powoduje szybszy spadek jakości diagnostycznej niż usuwanie według porządku rosnącego

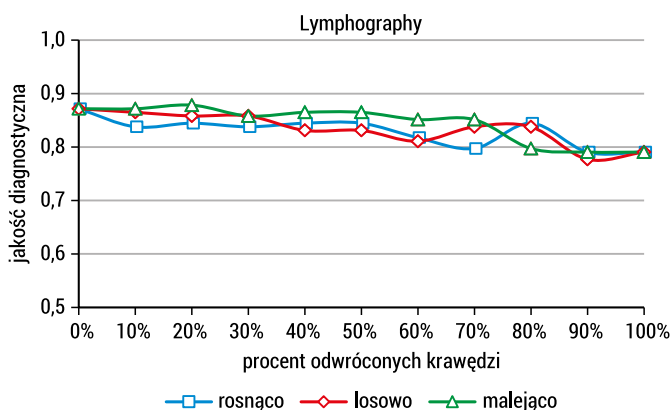
względem miary mocy krawędzi. Należy zwrócić uwagę na fakt, że nawet po usunięciu wszystkich krawędzi, jakość diagnostyczna modeli nie uległa drastycznemu pogorszeniu. Ma to związek z faktem, że jakość diagnostyczna jest determinowana poprzez powszechność choroby, która jest najbardziej prawdopodobna. Czyli w sytuacji, gdy jedna z chorób ma powszechność 80%, dokładność takiego modelu po usunięciu wszystkich krawędzi będzie równa 80%.

Wynik eksperymentu usuwania krawędzi dla modelu *Lymphography*, stworzonego w oparciu o dane dostępne na stronie repozytorium UCI przedstawia rysunek 3.1 [6].

Odwracanie krawędzi modelu sieci bayesowskiej

W eksperymencie tym systematycznie była modyfikowana struktura graficzna modelu przez odwracanie krawędzi grafu. Podobnie, jak w przypadku usuwania krawędzi, eksperyment rozpoczęto od wyznaczenia jakości diagnostycznej modelu oryginalnego, a następnie odwrócono kolejno 10%, 20%, 30%,..., 90% oraz 100% krawędzi. Ze względu na fakt, że sieć bayesowska jest acyklicznym grafem skierowanym, w niektórych sytuacjach nie było możliwe odwrócenie krawędzi. Wówczas dana krawędź oczekiwała na sytuację, w której to odwrócenie jej będzie możliwe. Eksperyment ten powtórzono trzy razy, zmieniając za każdym razem porządek usuwania krawędzi modelu, tj. porządek: (1) rosnący względem miary mocy kolejnych krawędzi; (2) malejący względem miary mocy kolejnych krawędzi oraz (3) losowy.

Wyniki tych eksperymentów pokazały, że zmiany w strukturze sieci bayesowskiej, szczególnie w przypadku odwracania kierunku krawędzi, nie mają dużego wpływu na jakość diagnostyczną modelu. Wynik eksperymentu odwracania krawędzi dla modelu *Lymphography* przedstawia rysunek 3.2.



RYSUNEK 3.2. Jakość diagnostyczna modelu *Lymphography* w funkcji procentu odwróconych krawędzi

3.2. Badanie czułości parametrów numerycznych modeli sieci bayesowskich

Część ilościową sieci bayesowskiej stanowią warunkowe rozkłady prawdopodobieństwa, które mogą być wyznaczane albo na bazie wiedzy eksperta albo na podstawie dostępnych danych. Wyznaczane wartości parametrów numerycznych sieci bayesowskiej nie są zazwyczaj wartościami dokładnymi. Badanie czułości parametrów numerycznych sieci bayesowskiej jest istotnym elementem inżynierii wiedzy w konstruowaniu tych modeli; szczególnie w sytuacji, gdy inżynier wiedzy tworzy modele sieci bayesowskich oparte na wiedzy ekspertów lub gdy parametry numeryczne modeli są estymowane z nielicznych lub niekompletnych zbiorów danych.

W ramach badań przeprowadzono szereg eksperymentów, które sprawdzały czułość parametrów numerycznych modeli sieci bayesowskich. Eksperymenty te sprowadzały się do modyfikacji oryginalnych parametrów modelu sieci bayesowskiej poprzez wprowadzanie szumu, a następnie do badania wpływu tej modyfikacji na jakość diagnostyczną modelu [7, 8, 9]. Druga grupa eksperymentów polegała na badaniu wpływu zaokrąglania parametrów numerycznych modelu sieci bayesowskiej na jej jakość diagnostyczną [9, 10]. We wszystkich eksperymentach jakość diagnostyczna była wyznaczana na podstawie walidacji krzyżowej typu *leave-one-out*.

Szum symetryczny w parametrach numerycznych sieci bayesowskich

Badanie czułości parametrów numerycznych sieci bayesowskich rozpoczęto od zreplikowania eksperymentu Pradhana i in. [11]. Wyniki eksperymentów przeprowadzonych przez Pradhana i in. pokazały, że sieci bayesowskie nie są czułe na szum wprowadzany do ich parametrów numerycznych. Eksperyment ten był przeprowadzony dla sieci CPC⁵, a jakość diagnostyczna modelu oparta była na mierze reprezentującej średnie prawdopodobieństwo *a posteriori* poprawnej diagnozy. Eksperyment Pradhana i in. powtórzyłam dla sieci bayesowskiej Hepar II [12] opierając jednakże badania o dwie miary jakości diagnostycznej: (1) miarę Pradhana, tj. średnie prawdopodobieństwo *a posteriori* poprawnej diagnozy oraz (2) miarę bazującą na prawdopodobieństwie *a posteriori* konkretnego pacjenta [7].

W eksperymencie, podobnie jak u Pradhana, szum był modelowany na podstawie rozkładu normalnego z wartościami standardowego odchylenia $\sigma \in <0,0, 3,0>$. Do każdego prawdopodobieństwa wprowadzany był szum o określonej wartości σ , a następnie badana była jakość diagnostyczna modelu. Eksperyment powtórzyłam 30 razy, za każdym razem z inną wartością σ z krokiem równym 0,1. Zaskakującym wynikiem był fakt, że jakość diagnostyczna modelu, zdefiniowana przez Pradhana, wzrastała wraz ze wzrostem poziomu szumu wprowadzanego do modelu. Fakt ten można zinterpretować tym, że szum zwiększył powszechność każdej z 11 modelowanych chorób, co spowodowało zwiększenie średniego prawdopodobieństwa *a posteriori*

⁵ Skrót CPC⁵ oznacza Computer-based Patient Case Study [13]

poprawnej diagnozy. Jednocześnie wyniki eksperymentu pokazały, że stosując inną miarę jakości diagnostycznej, opartej na prawdopodobieństwie *a posteriori* konkretnego pacjenta, obserwuje się wpływ szumu parametrów numerycznych na pogorszenie jakości modelu. W przypadku modelu Hepar II można zaobserwować systematyczne pogarszanie się jakości diagnostycznej od wartości $\sigma = 0,2$. Eksperyment ten został powtórzony dla trzech dodatkowych modeli sieci bayesowskich, które stworzono opierając się na danych z Irvine Machine Learning Repository [9]. Wyniki były spójne z tym, co zostało zaobserwowane dla modelu Hepar II, tj. jakość diagnostyczna modelu systematycznie malała wraz z poziomem szumu wprowadzanego do parametrów numerycznych.

Przeprowadzono również modyfikację tego eksperymentu, wprowadzając szum tylko do wybranych grup zmiennych [7]. Eksperyment ten przeprowadzono opierając się na modelu Hepar II. Wyróżniono cztery grupy zmiennych: testy laboratoryjne, badania przedmiotowe i podmiotowe, zmienne pochodzące z wywiadu z pacjentem, oraz zmienne reprezentujące jednostki chorobowe. W kolejnych krokach eksperymentu szum był wprowadzany tylko do wybranej grupy zmiennych, a następnie była badana jakość diagnostyczna modelu. Szum był modelowany na podstawie rozkładu normalnego z wartościami $\sigma \in <0,0, 3,0>$ z krokiem $\sigma = 1,0$. Wyniki tego eksperymentu pokazały, że szum wprowadzany do zmiennych reprezentujących testy laboratoryjne oraz jednostki chorobowe ma większy wpływ na pogorszenie jakości diagnostycznej niż szum wprowadzany do zmiennych modelujących badania przedmiotowe i podmiotowe oraz zmienne pochodzące z wywiadu z pacjentem.

Szum modelujący zachowanie eksperta w parametrach numerycznych sieci bayesowskich

Kolejna seria eksperymentów, które miały na celu badanie czułości parametrów numerycznych dotyczyła wprowadzania szumu, który modeluje zachowanie eksperta określającego parametry numeryczne sieci bayesowskiej, tj. ekspert zbyt pewny siebie lub ekspert niepewny swoich estymacji [9]. Eksperyment ten został przeprowadzony dla modelu Hepar II oraz dla trzech modeli sieci bayesowskich, które stworzono opierając się na danych z Irvine Machine Learning Repository.

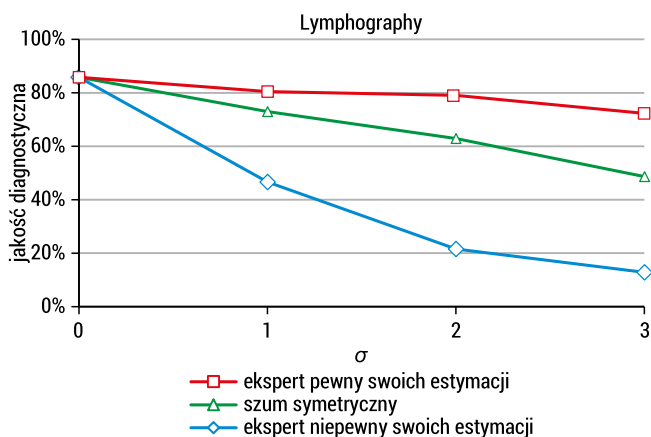
Podobnie jak w przypadku eksperymentu Pradhana i in., szum był modelowany na podstawie rozkładu normalnego ze standardowym odchyleniem $\sigma \in <0,0, 3,0>$. Z kolei sama modyfikacja parametrów numerycznych wyglądała inaczej, tzn. najmniejsze prawdopodobieństwo w danym rozkładzie było dodatkowo zmniejszane, a największe prawdopodobieństwo w rozkładzie było dodatkowo zwiększane o wartość, o którą zmniejszone było najmniejsze prawdopodobieństwo. Taka transformacja miała na celu modelowanie eksperta, który jest zbyt pewny swoich estymacji. Po wprowadzeniu szumu o określonym poziomie, badana była jakość diagnostyczna modelu. Eksperyment był powtórzony 30 razy, za każdym razem z inną wartością standardowego odchylenia σ . Podobnie wyglądał eksperyment, w którym modelowane było zachowanie eksperta, który jest niepewny swoich estymacji, z tym że największe prawdopodobieństwo w rozkładzie było zmniejszane, z kolei najmniejsze

prawdopodobieństwo w tym samym rozkładzie było zwiększane o wartość, o którą zmniejszone było największe prawdopodobieństwo. Po wprowadzeniu szumu o określonym poziomie, badana była jakość diagnostyczna modelu. Eksperyment był powtórzony 30 razy, za każdym razem z inną wartością standardowego odchylenia σ z krokiem równym 0,1.

Wyniki tych eksperymentów pokazały, że jakość diagnostyczna modelu jest najbardziej czuła na szum modelujący eksperta, który jest niepewny swoich estymacji. Z kolei szum modelujący eksperta, który jest zbyt pewny swoich estymacji ma mniejszy wpływ na pogorszenie jakości diagnostycznej modelu sieci bayesowskiej niż szum modelujący eksperta niepewnego swoich estymacji, czy też szum symetryczny, który modelowany był w eksperymencie Pradhana.

Wyniki tych eksperymentów są istotną wskazówką dla inżynierów wiedzy, którzy konstruują modele sieci bayesowskich szczególnie w sytuacji, gdy korzystają z wiedzy ekspertów lub gdy parametry numeryczne modeli są estymowane z niekompletnych zbiorów danych.

Wynik eksperymentu wprowadzania szumu do parametrów numerycznych modelu *Lymphography* przedstawia rysunek 3.3. Przedstawiono na nim trzy wykresy, które odpowiednio reprezentują szum eksperta przekonanego o swojej wiedzy, szum symetryczny oraz szum eksperta niepewnego swojej wiedzy. Jakość diagnostyczna modelu *Lymphography* jest przedstawiona w funkcji standardowego odchylenia (z krokiem $\sigma = 1,0$), które reprezentuje poziom wprowadzanego szumu.



RYSUNEK 3.3. Jakość diagnostyczna modelu *Lymphography* w funkcji poziomu szumu

Precyzja parametrów numerycznych modeli sieci bayesowskich

Kolejna seria eksperymentów, które miały na celu badanie czułości parametrów numerycznych dotyczyła badania wpływu precyzji parametrów numerycznych modeli sieci bayesowskich na ich jakość diagnostyczną [9, 10]. W badaniach tych systematycznie były zaokrąglane parametry numeryczne sieci bayesowskiej do wartości o coraz

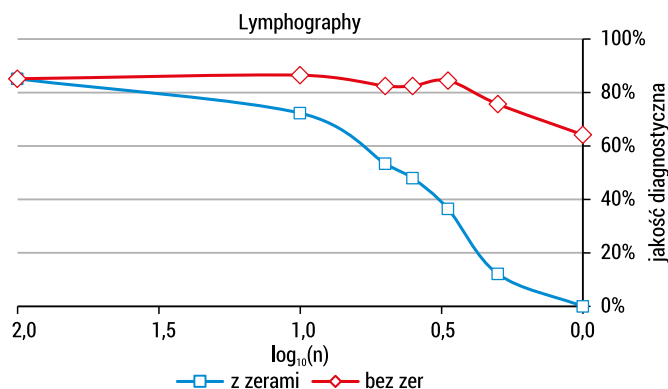
mniejszej precyzji. Eksperymenty te przeprowadzono dla 7 medycznych modeli diagnostycznych: modelu Hepar II oraz 6 modeli sieci bayesowskich, które stworzono opierając się na danych z Irvine Machine Learning Repository.

Zaokrąglanie wartości prawdopodobieństw w danym rozkładzie może prowadzić do tego, że ich suma wartości po zaokrągleniu nie jest równa 1,0, a więc zaokrąglona wartość wymagała dalszego przetworzenia. Problem ten był przedmiotem badań od ponad 200 lat – pierwsze jego analizy były przeprowadzane w okresie powstawania konstytucji amerykańskiej gdzie motywacją było opracowanie reguł sprawiedliwego podziału poszczególnych sił politycznych. Balinski & Young [14] omówili i porównali liczne algorytmy zaokrąglania. W swoich eksperymentach zastosowałam algorytm stacjonarnego zaokrąglania opracowany przez Heinricha i in. [15]. Jeden z parametrów (n) algorytmu stacjonarnego zaokrąglania określa liczbę przedziałów, w których może się znaleźć zaokrąglane prawdopodobieństwo. Na przykład, jeśli $n = 5$, wówczas zaokrąglane prawdopodobieństwo może przyjąć następujące wartości: 0,0, 0,2, 0,4, 0,6, 0,8, oraz 1,0. Jeśli z kolei $n = 1$, wówczas zaokrąglane prawdopodobieństwo może przyjąć tylko dwie wartości: 0,0 lub 1,0. W eksperymentach zaokrąglania zbadano następujące wartości parametru $n=100, 10, 5, 4, 2$, oraz 1. Każde prawdopodobieństwo w modelu było zaokrąglane do określonej dokładności, a następnie badana była jakość diagnostyczna modelu. Eksperyment był powtórzony 7 razy, tj. dla 7 różnych wartości parametru n , czyli za każdym razem model sieci bayesowskiej zawierał parametry o innej precyzji.

Wyniki przeprowadzonych eksperymentów pokazały, że jakość diagnostyczna modelu sieci bayesowskiej pogarsza się wykładniczo wraz ze zmniejszaniem precyzji parametrów numerycznych. Jednakże należy zauważyć, że pogorszenie jakości diagnostycznej wynikało głównie z zaokrąglania wartości prawdopodobieństwa do 0,0. Wartość 0,0 jest specjalną wartością, szczególnie w odniesieniu do twierdzenia Bayesa. Jeśli określone zdarzenie ma prawdopodobieństwo równe 0,0, prawdopodobieństwo *a posteriori* tego zdarzenia również będzie równe 0,0. Dlatego też zmodyfikowano ten eksperyment, eliminując z rozkładów prawdopodobieństwa wartości 0,0, tzn., za każdym razem gdy wartość miała być zaokrąglana do 0,0, była zastępowana wartością ϵ , jednocześnie od prawdopodobieństwa z najwyższą wartością w danym rozkładzie była odejmowana wartość ϵ . Parametr ϵ , w moich badaniach, przyjmował następujące wartości: 0,0001, 0,001, oraz 0,01. Po zastąpieniu prawdopodobieństw równych 0,0 wartością parametru ϵ , pogorszenie jakości diagnostycznej modelu sieci bayesowskiej wynikające z zaokrąglania parametrów było minimalne. Wartość parametru ϵ nie miała znaczącego wpływu, tzn., każda z trzech przebadanych wartości 0,0001, 0,001, oraz 0,01 dawała podobne wyniki.

Wyniki tych eksperymentów wskazują, że jakość diagnostyczna modeli sieci bayesowskich nie jest czuła na zaokrąglanie parametrów numerycznych, może jednakże nastąpić znaczące pogorszenie jakości diagnostycznej modelu w sytuacji, jeśli wartości prawdopodobieństw zostaną zaokrąglone do zera. Jest to istotna wskazówka dla inżyniera wiedzy, który powinien unikać wartości 0,0 określając rozkłady

prawdopodobieństwa w modelach sieci bayesowskich. Jest to też z pewnością cecha modeli w dziedzinie medycyny, gdzie rzadko wykluczane jest określone zdarzenie, a więc rzadko występuje prawdopodobieństwo równe 0,0.



RYSUNEK 3.4. Jakość diagnostyczna modelu Lymphography w funkcji zmiennej n

Wyniki eksperymentu zaokrąglania parametrów numerycznych modelu *Lymphography* w funkcji zmiennej n przedstawia rysunek 3.4.

Podsumowanie

Wyniki przeprowadzonych eksperymentów pokazały, że zmiany w strukturze sieci bayesowskiej, szczególnie w przypadku odwracania kierunku krawędzi, nie mają dużego wpływu na jakość diagnostyczną modelu. Okazało się, że usuwanie łuku ma silniejszy wpływ na dokładność diagnostyczną modelu niż jego odwrócenie. Dodatkowo, odwrócenie łuku ma silniejszy wpływ na jakość diagnostyczną w modelach typu naiwny Bayes i TAN⁶ niż w modelach przyczynowo-skutkowych [3].

Wyniki przeprowadzonych eksperymentów sugerują, że chociaż jakość diagnostyczna modeli sieci bayesowskich jest czuła na dokładność swoich parametrów numerycznych, małe ilości szumu nie wpływają znacznie na obniżenie jakości. Co więcej, wydaje się, że zbyttnia pewność siebie eksperta, który określa parametry numeryczne modelu jest bezpieczniejszą strategią niż brak pewności siebie.

Wyniki przeprowadzonych eksperymentów wskazują na to, że jakość diagnostyczna modeli sieci bayesowskich nie jest czuła na zaokrąglanie parametrów numerycznych. Znaczące pogorszenie jakości diagnostycznej modelu może nastąpić w sytuacji, jeśli wartości prawdopodobieństw zostaną zaokrąglone do zera. Jest to istotna wskazówka dla inżyniera wiedzy, który powinien unikać wartości 0,0 określając

⁶ TAN: Tree Augmented Naive Bayes

rozkłady prawdopodobieństwa w modelach sieci bayesowskich. Jest to też z pewnością cecha modeli w dziedzinie medycyny, gdzie rzadko występuje prawdopodobieństwo równe 0,0.

Wyniki przedstawionych eksperymentów są istotną wskazówką dla inżynierów wiedzy, którzy konstruują modele sieci bayesowskich szczególnie w sytuacji, gdy korzystają z wiedzy ekspertów lub gdy parametry numeryczne modeli są estymowane z niekompletnych zbiorów danych.

Bibliografia

- [1] Pearl J., *Probabilistic Reasoning in Intelligent Systems, Networks of Plausible Inference*, Morgan Kaufmann, 1988
- [2] Onisko A., Druzdzal M.J., *Impact of Bayesian Network model Structure on the Accuracy of Medical Diagnostic Systems*, "Proceedings of 13th International Conference on Artificial Intelligence and Soft Computing" (ICAISC'2014), red. L. Rutkowski, M. Korytkowski, R. Scherer, R. Tadeusiewicz, L. A. Zadeh, J. M. Zurada, Heidelberg, Springer, Lecture Notes in Computer Science, t. 8467, s. 167-178, Zakopane, 1-5 czerwca 2014
- [3] Mackiewicz G. and Onisko A., *The impact of structural and numerical quality of Bayesian network models on their accuracy*. "In Working Notes of the Conference on Risk and Decision-Making", Wellington, New Zealand, 13-14 listopada, 2019
- [4] Dua D. and Graff C., UCI Machine Learning Repository [<http://archive.ics.uci.edu/ml>]. Irvine, CA: University of California, School of Information and Computer Science, 2019
- [5] Koiter J.R., *Visualizing Inference in Bayesian Networks. Master's thesis*, Delft University of Technology, Delft, The Netherlands, 2006
- [6] Kononenko I., *Inductive and Bayesian learning in medical diagnosis*, "Applied Artificial Intelligence", 1993, 7, s. 317-337
- [7] Onisko A. and Druzdzal M.J., *Effect of Imprecision in Probabilities on Bayesian Network Models. An Empirical Study*. "In Working Notes of the Workshop on Model-based and Qualitative Reasoning in Biomedicine, European Conference on Artificial Intelligence in Medicine" (AIME-03), Protaras, Cypr, 19-22 października 2003
- [8] Druzdzal M.J., Onisko A., *The Impact of Overconfidence Bias on Practical Accuracy of Bayesian Network Models. An Empirical Study*, In Working Notes of the 2008 Bayesian Modelling Applications Workshop, Special Theme: How Biased Are Our Numbers? Part of the Annual Conference on Uncertainty in Artificial Intelligence (UAI-2008), Helsinki, Finland, 9 lipca 2008
- [9] Onisko A., Druzdzal M.J., *Impact of Quality of Bayesian Networks Parameters on Accuracy of Medical Diagnostic Systems, An Empirical Study*, "In Working Notes of the Workshop Probabilistic Models in Biomedicine. European Conference on Artificial Intelligence in Medicine" (AIME-11), Bled, Slovenia, 2011
- [10] Onisko A., Druzdzal M.J., *Impact of Precision of Bayesian Networks Parameters on Accuracy of Medical Diagnostic Systems*. "Artificial Intelligence in Medicine", 2013, 57(3), s. 197-206
- [11] Pradhan M, Henrion M, Provan G, del Favero B, Huang K. *The Sensitivity of Belief Networks to Imprecise Probabilities: an Experimental Investigation*, "Artificial Intelligence", 1996, 85(1-2), s. 363-97

- [12] Onisko A., Druzdzal M.J., Wasyluk H., Extension of the Hepar II Model to Multiple-Disorder Diagnosis. In: Kłopotek M, Michalewicz M, Wierzchoń S, editors. "Intelligent Information Systems", Advances in Soft Computing series. Heidelberg: Physica-Verlag (A Springer-Verlag Company); s. 303–13, 2000
- [13] Pradhan M., Provan G., Middleton B., and Henrion Max, *Knowledge engineering for large belief networks*, "In Proceedings of the Tenth Annual Conference on Uncertainty in Artificial Intelligence" (UAI-94), s. 484-490, Morgan Kaufmann Publishers, San Francisco, CA, 1994.
- [14] Balinski ML, Young HP, *Fair Representation. Meeting the Ideal of One Man, One Vote*, Yale University Press, New Haven, CT, 1982
- [15] Heinrich L, Pukelsheim F, Schwingenschlogl U., *On Stationary Multiplier Methods for the Rounding of Probabilities and the Limiting Law of the Sainte-Lague Divergence*. "Statistics and Decisions", 2005, 23, s. 117–29

Abstract

Knowledge engineering for modeling with Bayesian networks

Bayesian network models have been applied to both medical diagnosis and prognosis. Knowledge engineering for building Bayesian network models consists of domain knowledge acquisition, transforming this knowledge into an acyclic directed graph and determining conditional probability distributions. This process often requires combining data from different sources. The source of the data can be expert knowledge or objective data. Knowledge engineer plays an important role in this process and has to master the techniques of obtaining knowledge from an expert, understand the modeled field, have the ability to process medical data, also when they are incomplete. Although there are methods for automatic learning Bayesian networks from data, it is difficult to build a model without the assistance of an expert and knowledge engineer especially in the field of medicine. Most knowledge engineers believe that Bayesian networks are not sensitive to the diagnostic accuracy of their numerical parameters and that their graphical structure is much more important. In this chapter, I present the results of experiments in which the following were investigated: (1) the impact of the graphical structure and (2) the impact of the quality of numerical parameters on the diagnostic accuracy of Bayesian network models. The following chapter describes the techniques and methods to use when building Bayesian network models especially in the context of their numerical parameters.

Keywords: Bayesian networks, knowledge engineering, numerical parameters, sensitivity