

یادگیری ساختاری شبکه بیزی با به کارگیری پوشش مارکوفی در الگوریتم K_2

وحید رضایی تبار^۱، سلوا سلیمی^۲

تاریخ دریافت: ۱۳۹۴/۱۲/۱۰

تاریخ پذیرش: ۱۳۹۵/۷/۱۷

چکیده:

شبکه‌های بیزی، مدل‌های گرافیکی احتمالی هستند که رابطه علت و معلولی بین متغیرها را تعیین می‌کنند و شامل یادگیری ساختاری و یادگیری پارامتری می‌باشند. الگوریتم K_2 یکی از بهترین روش‌های یادگیری ساختار در شبکه‌های بیزی برای متغیرهای گسسته است. کارایی الگوریتم K_2 ، به شدت تحت تأثیر ترتیب متغیرهای ورودی است. بنا بر این برای رسیدن به گراف دقیقی که توصیف کننده داده‌ها باشد، یافتن الگوریتمی که ترتیب دقیق تری از عناصر به عنوان ورودی K_2 ارائه کند، مورد نیاز است. در این مقاله، نخست با استفاده از روش افزایشی-کاهشی، پوشش مارکوفی هر متغیر را یافته، سپس بر اساس فراوانی‌های شرطی و استفاده از تابع چگالی احتمال دیریکله، از بین پوشش مارکوفی هر متغیر، والدین احتمالی آن متغیر انتخاب می‌شوند. مجموعه والدین انتخابی هر رأس به عنوان ورودی الگوریتم K_2 مورد استفاده قرار می‌گیرد و شبکه بیزی به دست می‌آید. نتایج حاصل از اعمال الگوریتم پیشنهادی بر روی چند مجموعه داده معیار و مقایسه آن با روش‌های دیگر، نشان می‌دهد که الگوریتم پیشنهادی بسیار کاراتر از سایر روش‌ها است.

واژه‌های کلیدی: شبکه بیزی، الگوریتم K_2 ، پوشش مارکوفی، الگوریتم افزایشی-کاهشی

۱ مقدمه

با افزایش تعداد گره‌های آن به صورت نمایی افزایش می‌یابد و تولید همه ساختارهای ممکن و انتخاب بهترین ساختار از میان آن‌ها کاری بسیار پیچیده و زمان‌بر است (لارانیگا و همکاران [۶]). الگوریتم K_2 یکی از بهترین روش‌های یادگیری ساختاری شبکه بیزی برای متغیرهای گسسته است. این الگوریتم بر پایه ترتیب^۴ متغیرهای ورودی استوار است. دقت الگوریتم K_2 وابستگی زیادی به ترتیب ارائه شده برای متغیرها دارد و هر اندازه این ترتیب به ترتیب واقعی متغیرها نزدیک باشد، شبکه بیزی دقیق تری به دست می‌آید (کوپر و هرسکوویتس [۲]). به عنوان مثال، داشتن ترتیب A, B, C از متغیرها نشان دهنده این است که A والد احتمالی^۵ B و A, B والد احتمالی C هستند. بر این اساس، تمرکز اصلی الگوریتم‌های ارائه شده برای بهبود تابع امتیازدهی K_2 ، یافتن روشی است که منجر به رسیدن به ترتیب دقیقی از

چارچوب بیزی، روشی برای مدل‌بندی دنیای غنی و پیچیده اطراف ماست. با استفاده از این چارچوب می‌توان ساختار ذاتی یک فرایند را تفسیر نمود و اعمال مختلفی نظیر استنباط احتمالی و یادگیری را به سادگی امکان پذیر کرد. یک شبکه بیزی، گرافی متشکل از رأس‌ها (گره‌ها) و یال‌های جهت‌دار میان آن‌هاست که رأس‌ها نشان دهنده متغیرهای تصادفی‌اند (پیرل [۹]). دو نوع یادگیری در شبکه‌های بیزی مطرح است: یادگیری ساختاری و یادگیری پارامتری. در یادگیری ساختاری، گراف حاصل از روابط علت و معلولی بین متغیرها ارائه می‌شود و در یادگیری پارامتری، توزیع‌های شرطی بین متغیرها برآورد می‌شود (هکرمین [۳]). یادگیری ساختاری شبکه بیزی یک مسئله چندجمله‌ای سخت^۳ است؛ زیرا تعداد ساختارهای ممکن شبکه،

^۱ استادیار گروه آمار، دانشگاه علامه طباطبائی، تهران، ایران

^۲ دانشجوی کارشناسی ارشد رشته علوم تصمیم و مهندسی دانش، دانشگاه خوارزمی، تهران، ایران

^۳ NP-hard

^۴ order

^۵ candidate parent

^۶ feature ranking

ساختار این مقاله به شرح زیر است: در بخش ۲ به تشریح چارچوب نظری تحقیق پرداخته شده، در بخش ۳ الگوریتم کو و کیم به اختصار توضیح داده شده و پس از یافتن اشکالات این الگوریتم، در بخش ۴ به تشریح الگوریتم پیشنهادی پرداخته شده است. در نهایت، نتایج حاصل از اعمال روش های معرفی شده توسط چن و همکاران [۱]، هروشکا و ایبکن [۴]، کو و کیم [۵] و همچنین الگوریتم پیشنهادی بر روی چند مجموعه داده معیار در بخش ۵ ارائه شده و گراف حاصل از این روش ها با یکدیگر مقایسه شده است.

۲ چارچوب نظری

۱.۲ شبکه بیزی

شبکه های بیزی از ابزارهای داده کاوی و گروهی از مدل های گرافیکی احتمالی هستند که عمدتاً نشان دهنده روابط علت و معلولی بین متغیرها هستند. شبکه بیزی یک گراف جهت دار بدون دور^{۱۰} است. این گراف با مجموعه رئوس V و مجموعه یال های E که زیرمجموعه ای از $V \times V$ است، به صورت $G = (V, E)$ نشان داده می شود. هر گره گراف مربوط به یک متغیر تصادفی در دامنه است و این گراف، یک خانواده از توزیع های احتمال را بر روی متغیرهای V نشان می دهد. یال های جهت دار E نشان دهنده وابستگی بین رأس های مربوط به آن یال هستند. یک یال $X \rightarrow Y$ در گراف، رابطه بین والد X و فرزند Y را شرح می دهد. به علاوه، هر گره یک جدول احتمال شرطی دارد که احتمال هر ترکیب ممکن، یک گره و والدین خود را نشان می دهد. اگر یک گره والدی نداشته باشد، احتمال های شرطی همان احتمال های حاشیه ای آن گره خواهند بود (پیرل [۹]).

یکی از مشکلات استفاده از شبکه های بیزی، رسیدن به گرافی است که به درستی بیان کننده ارتباط میان متغیرهای مختلف در مجموعه داده باشد. بنا بر این تلاش های زیادی برای یادگیری شبکه های بیزی صورت گرفته است. در فرایند یادگیری باید بتوان ساختار شبکه بیزی و جداول احتمالات شرطی آن را تعیین کرد. برای یادگیری خود کار ساختار شبکه، یکی از روش ها بر

عناصر باشد. هروشکا و ایبکن [۴] به معرفی الگوریتمی بر اساس رتبه بندی بر پایه انتخاب ویژگی^۶ پرداخته و بر اساس آزمون آماری χ^2 و بهره اطلاع^۷ رده بندی کرده و پس از ارزیابی میزان وابستگی متغیرها به هم، ترتیبی برای متغیرها در نظر گرفته اند. از معایب این روش آن است که روابط علی میان متغیرها در نظر گرفته نمی شود؛ یعنی به وضوح مشخص نیست میان دو متغیری که وابستگی زیادی دارند، کدام یک والد احتمالی دیگری است. چن و همکاران [۱] با استفاده از ترکیب مفاهیم نظریه اطلاع و توابع جستجو، الگوریتمی برای یافتن ترتیبی از عناصر ارائه کردند. الگوریتم آن ها از سه مرحله اصلی تشکیل شده است که در دو مرحله اول با استفاده از آزمون های استقلال و اطلاع متقابل و مجموعه جداکننده^۸ ساختار اولیه گراف مشخص می شود و در مرحله آخر، یال ها با استفاده از آزمون استقلال و الگوریتم های جستجو، جهت دهی می شوند. الگوریتم کو و کیم [۵] که نسبت به سایر الگوریتم ها عملکرد بهتری دارد، بر این اساس شکل گرفته که هر فرزند تحت والد خود، فراوانی شرطی بهتری نسبت به سایر متغیرهایی دارد که والدش نیستند. بنا بر این امتیازدهی میان رئوس بر اساس فراوانی شرطی آن ها انجام شده و در نهایت بر اساس این امتیازات، والدین احتمالی هر رأس مشخص می شوند. این والدین احتمالی در نهایت به عنوان ورودی الگوریتم K^2 به کار گرفته می شوند و شبکه بیزی به دست می آید. یکی از مشکلات روش کو و کیم [۵]، تولید یال های اضافی در شبکه بیزی است.

در این مقاله برای حل مشکل روش کو و کیم [۵]، نخست پوشش مارکوفی هر رأس را می یابیم. سپس والدین احتمالی هر رأس، فقط بر اساس امتیاز میان رأس مذکور و اعضای پوشش مارکوفی آن محاسبه می شوند و به عنوان ورودی الگوریتم K^2 مورد استفاده قرار می گیرند. به عبارت دیگر، والدین احتمالی هر متغیر از بین اعضای پوشش مارکوفی آن متغیر انتخاب می شوند. نتایج حاصل از اعمال الگوریتم پیشنهادی بر روی چند مجموعه شبکه معیار^۹ نشان می دهد، الگوریتم پیشنهادی با کاهش تعداد یال های اضافی منجر به دقیق تر شدن گراف حاصل می شود و در نتیجه گراف مذکور، انطباق بیشتری با داده ها دارد.

^۶ information gain

^۸ d-separation

^۹ benchmark

^{۱۰} directed acyclic graph

مجموعه والدین متغیر را صفر در نظر بگیر؛ یعنی

$$Pa(x_i) = \cdot$$

$$P_{old} = g(x_i, Pa(x_i))$$

(تابع g متر $K2$ شبکه ییزی است.)

تا زمانی که متغیری برای ادامه دادن وجود دارد و تعداد والدین گره x_i از u کمتر است، ادامه بده.

شروع

تا زمانی که گره z از مجموعه متغیرهایی که x_i می تواند به عنوان والدین خود انتخاب کند وجود دارد، آن را طوری انتخاب کنید که عبارت $g(x_i, Pa(x_i) \cup \{z\})$ را ماکسیمم نماید.

$$P_{New} = g(x_i, Pa(x_i) \cup \{z\})$$

$$P_{New} > P_{old} \text{ اگر}$$

شروع

$$P_{old} := P_{New}$$

$$Pa(x_i) = Pa(x_i) \cup \{z\}$$

پایان

اگر متغیری که به عنوان والدین متغیر انتخاب شود در ترتیب وجود نداشت.

پایان

والدین مربوط به متغیر x_i را چاپ کنید.

پایان الگوریتم $K2$

در این الگوریتم، تابع امتیاز به صورت زیر تعریف می شود (کوپر و هرسکوویتس [۲]):

$$g(x_i, Pa(x_i)) = \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod_{k=1}^{r_i} N_{ijk}$$

که در آن

$Pa(x_i)$: مجموعه والدین رأس x_i

q_i : تعداد حالات والد رأس x_i

r_i : تعداد حالات متغیر x_i

N_{ijk} : تعداد نمونه های مجموعه داده D

که در آن x_i k امین حالت و والد آن z امین حالت خود را دارد. همچنین $N_{ij} = \sum_{k=1}^{r_i} N_{ijk}$ تعداد واحدهای نمونه از مجموعه داده است که در آن والدین متغیر x_i z امین حالت خود را دارند.

^{۱۱} greedy search algorithm

پایه تعیین وابستگی بین متغیرها بنا نهاده شده و در روش دیگر (یادگیری ساختاری) به دنبال یافتن ساختاری بهینه برای شبکه ییزی هستیم که با داده های موجود، بیشترین تطابق را داشته باشد. یادگیری ساختاری، خود دارای دو رویکرد اصلی است: رویکرد مبتنی بر قید که بر اساس تحلیل وابستگی موجود در داده ها، ساختار بهینه را می یابد و رویکرد مبتنی بر رتبه بندی- امتیازدهی که از یک تابع امتیاز برای رتبه بندی ساختارهای ممکن استفاده می کند و سپس با بهره گیری از یک الگوریتم جستجو، به دنبال کشف ساختاری با بیشترین امتیاز است. این روش نسبت به روش مبتنی بر قید، دقیق تر و خروجی آن ساختاری منحصر به فرد خواهد بود (نیلسن و یینسن [۸]).

۲.۲ الگوریتم $K2$

الگوریتم $K2$ ، الگوریتم جستجوی حریصانه ای^{۱۱} است که ساختار یک شبکه ییزی را بر اساس میزان تطابق آن با داده ها به دست می آورد. در واقع بر اساس این الگوریتم، ساختاری که تطابق بیشتری با داده ها دارد، امتیاز بیشتری کسب کرده، به عنوان بهترین ساختار توصیف کننده داده ها انتخاب می شود (کوپر و هرسکوویتس [۲]). الگوریتم $K2$ برای کاهش پیچیدگی زمانی، ترتیب اولیه ای از گره ها (متغیرها) را به عنوان ورودی در نظر می گیرد. بنا بر این یافتن ترتیبی از گره ها که والدین احتمالی هر گره را به درستی مشخص کند، در رسیدن به یک ساختار شبکه ییزی با امتیاز زیاد، بسیار با اهمیت است. بر این اساس، تمرکز اصلی الگوریتم ارائه شده، یافتن ترتیب دقیقی از گره ها است. الگوریتم $K2$ به شرح زیر است (کوپر و هرسکوویتس [۲]):

الگوریتم $K2$

- ورودی و خروجی الگوریتم
ورودی: مجموعه ای از n گره یا متغیر تصادفی گسسته، یک ترتیب روی گره ها، یک حد بالای u روی تعداد والدین هر گره و یک پایگاه داده شامل n نمونه مستقل. خروجی: مجموعه والدین مربوط به هر گره.

- شروع الگوریتم $K2$
برای i از یک تا n انجام شود.

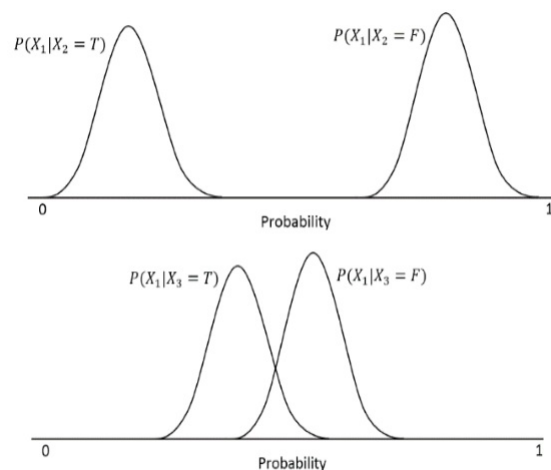
۳ روش کو و کیم در یافتن والدین احتمالی

روش کو و کیم [۵] مبتنی بر توزیع دیریکله، به صورت

$$Dir(p(c_1), \dots, p(c_{r_i}); f(c_1), \dots, f(c_{r_i})) \quad (1)$$

$$= \frac{\Gamma(\sum_{k=1}^{r_i} f(c_k))}{\prod_{k=1}^{r_i} \Gamma(f(c_k))} \prod_{k=1}^{r_i} p(c_k)^{f(c_k)-1}$$

است که در آن c_k ، k امین سطح یک متغیر، $p(c_k)$ احتمال c_k و $f(c_k)$ فراوانی c_k است. هر اندازه فاصلهٔ توابع چگالی احتمال دیریکلهٔ متغیری بر روی تغییرات والدش بیشتر باشد بیان‌کنندهٔ آن است که متغیر موجود به درستی به عنوان والد انتخاب شده است.



شکل ۱. تابع چگالی احتمال متغیر X_1 به شرط دو متغیر X_2 و X_3

بنا بر این با توجه به شکل ۱، متغیر X_2 انتخاب درست‌تری برای X_1 به عنوان والد است. برای تعیین والدین هر متغیر، تابع چگالی احتمال دیریکلهٔ آن، در صورتی که مطابق شکل ۱ و نمودار مربوط به متغیر X_2 باشد، باید امتیاز بیشتری دریافت کند. برای این منظور، در ابتدا جدول توافقی بین هر جفت متغیر تشکیل شده، فرض می‌شود که متغیر فرزند با سطوح مختلف آن در سطر و متغیر والد با سطوح مختلف آن در ستون جدول توافقی باشد. در این صورت برای تشخیص این که آیا متغیر ستون می‌تواند والد احتمالی متغیر سطر باشد به صورت زیر عمل می‌شود:

- مقدار احتمال توزیع دیریکله برای ستون اول با در نظر گرفتن فراوانی و احتمال‌های مربوط به آن محاسبه می‌شود.
- سپس مقدار احتمال توزیع دیریکله برای سایر ستون‌ها، با ثابت در نظر گرفتن احتمال‌های مربوط به فراوانی ستون

اول محاسبه می‌شود. به عبارت دیگر در رابطه (۱)، احتمال $p(c_i)$ همان احتمالات ستون اول است و فقط فراوانی‌ها از ستون‌های دیگر برای محاسبهٔ احتمال دیریکله انتخاب می‌شوند. هدف از این کار، محاسبهٔ احتمال دیریکلهٔ سایر ستون‌ها با در نظر گرفتن احتمال‌های ستون اول است.

- مقدار مینیمم و ماکسیمم قدم‌های قبل به ترتیب به عنوان c_1 و r_1 در نظر گرفته می‌شوند (شکل ۲). به عنوان مثال برای نمونه‌ای مانند شکل ۲ داریم:

$$r_1 = \frac{\Gamma(d_{11} + d_{12})}{\Gamma(d_{11})\Gamma(d_{12})} \left(\frac{d_{11}}{d_{11} + d_{12}}\right)^{d_{11}} \left(\frac{d_{12}}{d_{11} + d_{12}}\right)^{d_{12}}$$

$$c_1 = \frac{\Gamma(d_{11} + d_{12})}{\Gamma(d_{11})\Gamma(d_{12})} \left(\frac{d_{21}}{d_{21} + d_{22}}\right)^{d_{21}} \left(\frac{d_{22}}{d_{21} + d_{22}}\right)^{d_{22}}$$

- این عملیات برای همهٔ سطوح متغیر والد تکرار می‌شود. بنا بر این به تعداد سطوح متغیر والد، مقادیر r_1 و c_1 خواهیم داشت.

- در نهایت با استفاده از رابطهٔ ۲، میزان وابستگی متغیر فرزند (X_i) به متغیر والد X_j محاسبه می‌شود. در واقع رابطهٔ ۲، فاصلهٔ بین سطوح متغیر والد احتمالی و فرزند را محاسبه می‌کند، که بیشتر بودن این فاصله نشان‌دهندهٔ انتخاب صحیح متغیر والد است. هرچه میزان این فاصله بیشتر باشد نشان از تحت تأثیر بودن متغیر فرزند تحت متغیر والد است و بیان‌کنندهٔ انتخاب درست متغیر والد احتمالی است (به مثال ۱ رجوع شود).

$$score(X_j, X_i) = Dep(X_i|X_j) \quad (2)$$

$$= \sum_{k=1}^{r_j} (\log r_k - \log c_k)$$

شایان ذکر است که امتیاز میان هر دو متغیر (رأس یا گره)، با توجه به رابطه (۲)، محاسبه می‌شود و در صورتی که $score(A, B) > score(B, A)$ باشد، رأس A به عنوان والد احتمالی رأس B انتخاب شده، الگوریتم $K2$ برای هر رأس بر اساس والدین احتمالی‌اش فراخوانی و محاسبه می‌شود (منظور از $score$ ، امتیاز بین دو متغیر A و B بر اساس تابع امتیاز $K2$ است).

همان‌گونه که در جدول ۱ نشان داده شده است، برای متغیر $meldug_1$ فراوانی‌ها و احتمالات آن برای چهار متغیر دیگر محاسبه شده است. همین روند برای همه جفت متغیرها محاسبه می‌شود. شایان ذکر است که به دلیل جلوگیری از خطای صفر به توان صفر برای حالتی که فراوانی و در نتیجه احتمال نظیر آن صفر است، در جداول توافقی، هرگاه احتمالی برابر با صفر بود، ۰.۰۱ در نظر گرفته می‌شود.

برای دو متغیر $meldug_1$ (به‌عنوان فرزند) و $foto_1$ (به‌عنوان والد) که فراوانی‌ها و احتمالات آن‌ها در جدول ۱ نیز آمده است، داریم:

$$Dir(0.9339, 0.0661; 9296, 658) = -4/1295$$

$$Dir(0.9339, 0.0661; 20, 3) = -2/0381$$

$$\underbrace{Dir(0.9339, 0.0661; 21, 2)}_{\max - \min = R_1 - C_1 = 2/7936}$$

$$Dir(0.8606, 0.1304; 9296, 658) = -220/3954$$

$$Dir(0.8696, 0.1304; 20, 3) = -1/4266$$

$$\underbrace{Dir(0.8696, 0.1304; 21, 2)}_{\max - \min = R_2 - C_2 = 2/18968}$$

$$Dir(0.9130, 0.0870; 9296, 658) = -33/6381$$

$$Dir(0.9130, 0.0870; 20, 3) = -1/6672$$

$$\underbrace{Dir(0.9130, 0.0870; 21, 2)}_{\max - \min = R_3 - C_3 = 32/3764}$$

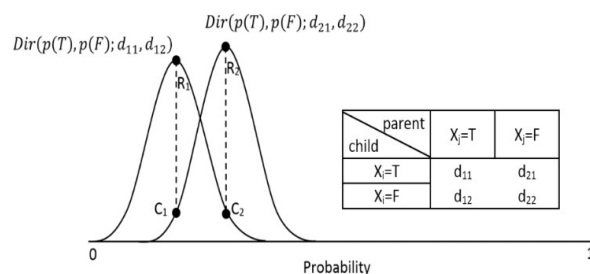
$$score(foto_1, meldug_1)$$

$$= 2/7936 + 2/18968 + 32/3764 = 254/1318$$

امتیاز میان هر دو رأس، همانند روش فوق‌الذکر محاسبه و در جدولی مانند جدول ۲ ذخیره می‌شود. در جدول مربوط، علامت ★ نشان‌دهنده والدین احتمالی است.

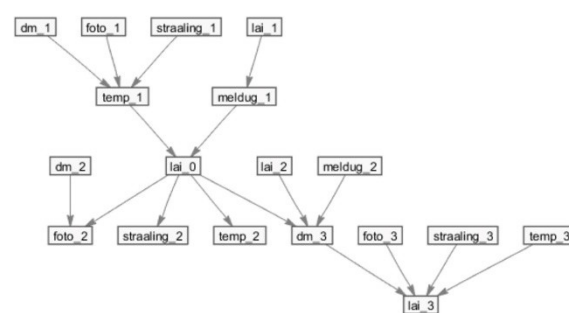
پس از تکمیل جدول، اگر $score(A, B) < score(B, A)$ برقرار باشد، B به‌عنوان والد احتمالی A انتخاب می‌شود. بنا بر این با توجه به جدول ۲، والدین احتمالی متغیر $meldug_1$ عبارت‌اند از:

$$\{dm_1, foto_1, straaling_1, temp_1, lai_2, dm_3, meldug_2, foto_3, straaling_3, temp_3\}$$



شکل ۲. تابع چگالی احتمال دیریکله، کو و کیم [۵]

مثال ۱.۳. روند کلی الگوریتم کو و کیم برای مجموعه داده car که شامل ۱۸ متغیر (رأس) و ۱۷ یال می‌باشد، در ادامه بررسی شده است. شبکه اصلی مربوط به داده‌های car در شکل ۳ نشان داده شده است. شایان ذکر است که مجموعه داده car همراه با شبکه اصلی آن در نرم‌افزار R و در بسته نرم‌افزاری $bnlearn$ موجود است. منظور از شبکه اصلی، شبکه‌ای است که الگوریتم‌های یادگیری به دنبال یافتن بیشترین تطابق با این شبکه، برای بررسی میزان کارایی خود می‌باشند.



شکل ۳. شبکه اصلی مربوط به مجموعه داده car بسته

نرم‌افزاری $bnlearn$ در نرم‌افزار R

در الگوریتم کو و کیم [۵]، در ابتدا فراوانی سطوح مختلف هر متغیر به‌زای سطوح مختلف متغیر دیگر محاسبه می‌شود و سپس احتمال رخداد آن‌ها با استفاده از توزیع دیریکله به دست می‌آید.

جدول ۱. جدول توافقی و احتمالات نظیر آن‌ها برای چهار

متغیر مجموعه داده car

meldug_1/dm_1				meldug_1/dm_1			
	۱	۲			۱	۲	
۱	۹۳۳۳	۴		۱	۰.۹۳۳۷	۰.۹۹	
۲	۶۶۳	۰		۲	۰.۰۶۶۳	۰.۰۱	

meldug_1/foto_1				meldug_1/foto_1			
	۱	۲	۳		۱	۲	۳
۱	۹۲۹۶	۲۰	۲۱	۱	۰.۹۳۴	۰.۸۷	۰.۹۱۳
۲	۶۵۸	۳	۲	۲	۰.۰۶۶	۰.۱۳	۰.۰۸۷

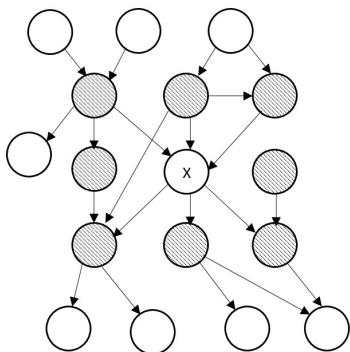
meldug_1/straaling_1				meldug_1/straaling_1			
	۱	۲			۱	۲	
۱	۱	۹۳۳۶		۱	۰.۹۹	۰.۹۳۳۷	
۲	۰	۶۶۳		۲	۰.۰۱	۰.۰۶۶۳	

meldug_1/temp_1				meldug_1/temp_1			
	۱	۲			۱	۲	
۱	۹۲۹۲	۴۵		۱	۰.۹۳۳۹	۰.۹	
۲	۶۵۸	۵		۲	۰.۰۶۶۱	۰.۱	

جدول ۲. انتخاب والد احتمالی رأس ۱-meldug از جدول امتیازات

	dm_۱	foto_۱	straaling_۱	temp_۱	lai_۱	meldug_۱	lai_۲	dm_۲	foto_۲	straaling_۲	temp_۲	lai_۳	dm_۳	foto_۳	straaling_۳	temp_۳	lai_۴
dm_۱	.	۱۳۵۰۰۰۰	۹۶۸۰۰۰۰	۱۳۵۰۰۰۰	۷۷۶۴۱۴۷	۷۱۵۰۴۰۰	۱۰۵۳۴۴۴	۱۴۵۰۰۰۰	۱۵۰۰۰۰۰	۷۶۱۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴	۱۴۵۰۰۰۰	۷۶۱۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰
foto_۱	.	.	۱۹۱۰۰۰۰	۹۱۱۵۳۴۴	۵۸۴۰۰۰۰	۷۵۴۰۰۰۰	۸۰۰۰۰۰۰	۱۰۰۰۰۰۰	۱۰۰۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴	۱۴۵۰۰۰۰	۷۶۱۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴
straaling_۱	۸۶۸۰۰۰۰	۱۳۴۸۹۷	.	۴۵۵۱۷۷۶	۳۶۱۹۸۸۳	۷۱۵۰۰۰۰	۴۰۰۰۰۰۰	۱۰۰۰۰۰۰	۱۰۰۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴	۱۴۵۰۰۰۰	۷۶۱۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴
temp_۱	۶۴۴۴۷۷۸	۷۵۱۷۷۰۰	.	.	۴۴۴۸۰۰۰	۷۵۴۴۴۴۴	۴۰۰۰۰۰۰	۱۰۰۰۰۰۰	۱۰۰۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴	۱۴۵۰۰۰۰	۷۶۱۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴
lai_۱	۷۶۵۴۰۰۰	۶۰۰۰۰۰۰	۴۵۰۰۰۰۰	۵۰۰۰۰۰۰	.	۴۴۴۴۴۴۴	۴۰۰۰۰۰۰	۱۰۰۰۰۰۰	۱۰۰۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴	۱۴۵۰۰۰۰	۷۶۱۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴
meldug_۱	۷۴۵۱۷۴	۱۱۰۰۰۰۰	۸۴۵۵۵۹۳	۷۰۰۰۰۰۰	۱۰۰۰۰۰۰	.	۴۴۴۴۴۴۴	۴۰۰۰۰۰۰	۱۰۰۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴	۱۴۵۰۰۰۰	۷۶۱۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴
lai_۲	۴۴۸۸۰۰۰	۶۹۱۰۰۰۰	۱۰۰۰۰۰۰	۷۱۴۰۰۰۰	۹۰۰۰۰۰۰	۴۸۳۹۰۰۰	.	۴۴۴۴۴۴۴	۴۰۰۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴	۱۴۵۰۰۰۰	۷۶۱۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴
dm_۲	۴۰۰۰۰۰۰	۴۰۰۰۰۰۰	۴۰۰۰۰۰۰	۴۰۰۰۰۰۰	۴۰۰۰۰۰۰	۴۰۰۰۰۰۰	۴۰۰۰۰۰۰	۴۰۰۰۰۰۰	۴۰۰۰۰۰۰	۴۰۰۰۰۰۰	۴۰۰۰۰۰۰	۴۰۰۰۰۰۰	۴۰۰۰۰۰۰	۴۰۰۰۰۰۰	۴۰۰۰۰۰۰	۴۰۰۰۰۰۰	۴۰۰۰۰۰۰
foto_۲	۹۶۸۷۷۱۲	۴۸۵۰۰۱۲	۱۰۰۰۰۰۰	۸۰۰۰۰۰۰	۸۴۸۵۶۴۵	۹۷۱۰۰۰۰	۱۰۰۰۰۰۰	۱۰۰۰۰۰۰	۱۰۰۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴	۱۴۵۰۰۰۰	۷۶۱۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴
straaling_۲	۱۷۰۰۰۰۰	۳۶۴۳۶۴۴	۷۴۵۷۷۸۹	۳۱۵۱۷۷۹	۱۳۷۲۰۰۰	۶۴۴۰۰۰۰	۶۸۸۰۰۰۰	۴۰۰۰۰۰۰	۴۰۰۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴	۱۴۵۰۰۰۰	۷۶۱۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴
temp_۲	۱۸۰۰۰۰۰	۳۷۰۰۰۰۰	۷۶۶۴۴۹۹	۴۰۰۰۰۰۰	۱۳۴۱۰۰۰	۵۶۴۴۴۹۹	۶۱۵۸۸۳۱	۴۰۰۰۰۰۰	۴۰۰۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴	۱۴۵۰۰۰۰	۷۶۱۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴
lai_۳	۸۴۶۸۴۶۹	۱۴۶۴۴۹۱	۹۴۷۷۷۱۶	۱۹۰۰۰۰۰	۱۳۶۷۱۴۵	۴۴۹۹۸۸۵	۶۴۷۷۷۵۴	۴۱۰۰۰۰۰	۴۱۰۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴	۱۴۵۰۰۰۰	۷۶۱۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴
meldug_۲	۸۶۸۰۰۰۰	۱۳۴۸۹۷	۹۶۸۵۹۹۹	۴۱۰۰۰۰۰	۳۶۱۹۸۸۳	۷۱۵۰۰۰۰	۸۰۰۰۰۰۰	۱۰۰۰۰۰۰	۱۰۰۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴	۱۴۵۰۰۰۰	۷۶۱۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴
dm_۳	۴۶۶۴۴۴۴	۵۱۸۵۶۵۵	۱۰۰۰۰۰۰	۵۸۵۱۳۶۶	۵۱۱۲۰۰۰	۱۵۴۴۴۴۴	۱۸۰۰۰۰۰	۴۰۰۰۰۰۰	۴۰۰۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴	۱۴۵۰۰۰۰	۷۶۱۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴
foto_۳	۸۶۷۱۳۱۱	۱۳۵۴۴۴۱	۹۶۷۷۷۸۸	۴۱۰۰۰۰۰	۱۱۰۰۰۰۰	۱۵۴۸۸۸۵	۱۴۶۴۴۴۴	۱۱۵۴۴۴۴	۱۱۵۴۴۴۴	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴	۱۴۵۰۰۰۰	۷۶۱۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴
straaling_۳	۸۶۶۸۴۸۸	۱۳۵۴۴۴۱	۹۶۶۸۴۸۸	۴۱۰۰۰۰۰	۱۱۰۰۰۰۰	۱۵۴۸۸۸۵	۱۴۶۴۴۴۴	۱۱۵۴۴۴۴	۱۱۵۴۴۴۴	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴	۱۴۵۰۰۰۰	۷۶۱۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴
temp_۳	۸۶۶۸۴۸۸	۱۳۵۴۴۴۱	۹۶۶۸۴۸۸	۴۱۰۰۰۰۰	۱۱۰۰۰۰۰	۱۵۴۸۸۸۵	۱۴۶۴۴۴۴	۱۱۵۴۴۴۴	۱۱۵۴۴۴۴	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴	۱۴۵۰۰۰۰	۷۶۱۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴
lai_۴	۴۵۵۴۴۴۴	۵۰۰۰۰۰۰	۱۰۰۰۰۰۰	۵۸۵۱۳۶۶	۵۱۱۲۰۰۰	۱۵۴۴۴۴۴	۱۸۰۰۰۰۰	۴۰۰۰۰۰۰	۴۰۰۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴	۱۴۵۰۰۰۰	۷۶۱۰۰۰۰	۷۸۵۴۱۴۷	۶۰۰۰۰۰۰	۱۰۵۳۴۴۴

یعنی به ازای هر رأس X و Y که عضو مجموعه رئوس باشند، X عضو مجموعه پوشش مارکوفی رأس Y است اگر و تنها اگر دو متغیر X و Y به شرط حداقل یکی از اعضای مجموعه V (به جز خود X و Y) به هم وابسته باشند.



شکل ۴. پوشش مارکوفی رأس X

در واقع پوشش مارکوفی هر رأس شامل والد، فرزند و دیگر والد فرزند (همسر) است. شکل ۴، پوشش مارکوفی رأس X را نشان می دهد. الگوریتم های مختلفی برای یافتن پوشش مارکوفی رئوس پیشنهاد شده که در این مقاله از روش افزایشی-کاهشی برای یافتن پوشش مارکوفی استفاده شده است. این الگوریتم، یکی از روش هایی است که با استفاده از آزمون های استقلال شرطی، وابستگی یا عدم وابستگی رئوس را تشخیص می دهد و بر این اساس، ساختار محلی در اطراف هر رأس به دست می آید. این الگوریتم دارای دو مرحله افزایش و کاهش است. برای یافتن همسایه های هر رأس مانند X ، نخست یک مجموعه بدون عضو مانند S در نظر گرفته و در مرحله افزایش، متغیری را که به شرط مجموعه S وابسته به X است، به S افزوده و همین روند برای تمامی متغیرهای دیگر به شرط وجود مجموعه S بررسی می شود. در این مرحله ممکن است رئوسی به مجموعه S اضافه شده باشند که در همسایگی رأس X قرار ندارند. برای مشخص شدن این رئوس، مرحله کاهش اجرا شده، اگر رأسی مانند Y در مجموعه S وجود داشته باشد که $X \perp Y | S - \{Y\}$ باشد، Y از مجموعه S حذف می شود ($\perp$ نشان دهنده استقلال است). این مرحله نیز به ازای همه اعضای مجموعه S اجرا شده و در نهایت، رئوس موجود در S نشان دهنده همسایه های (پوشش مارکوفی) رأس X می باشند (مارگاریتیس و ترون [۷]). بررسی استقلال شرطی رئوس با استفاده از آزمون آماری χ^2 صورت می گیرد که پیش تر توضیح داده شد.

در ادامه، الگوریتم K^2 امتیاز میان هر رأس و اعضای مجموعه والدین احتمالی آن محاسبه می شود و در صورتی که امتیاز زیادی به دست آید، آن متغیر به عنوان والد متغیر مذکور در نظر گرفته شده، یالی بین آن ها ایجاد می شود.

۴ روش پیشنهادی

یکی از مشکلات روش کو و کیم، ایجاد یال های اضافی است. بنا بر این برای رفع این مشکل، در ابتدا پوشش مارکوفی هر رأس را یافته، سپس والدین احتمالی از بین آن ها انتخاب می شوند.

۱.۴ استقلال شرطی

اگر X و Y و Z سه متغیر تصادفی با مقادیر گسسته باشند، می گوییم X و Y به شرط دانستن Z مستقل اند اگر و تنها اگر:

$$P(X = x_i | Y = y_j, Z = z_k) = P(X = x_i | Z = z_k) \quad (\forall x_i, y_j, z_k)$$

مهم ترین آزمون آماری که از آن برای ارزیابی مستقل بودن رأس ها استفاده می شود، آزمون آماری χ^2 است که به صورت زیر فرمول بندی می شود:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

که در آن E تعداد دفعاتی است که انتظار داشتیم یک اتفاق رخ دهد و O تعداد دفعاتی است که این اتفاق رخ داده است.

۲.۴ پوشش مارکوفی

پوشش مارکوفی هر رأس در شبکه های بیزی، دربردارنده همسایه های آن رأس می باشد. به عبارت دیگر می توان گفت، پوشش مارکوفی هر رأس، مجموعه ای از رئوس است که رأس مذکور را از سایر رئوس شبکه جدا می سازد و همه اطلاع لازم برای پیش بینی رفتار هر رأس را می توان با کمک مجموعه اعضای پوشش مارکوفی آن به دست آورد، که این اطلاع قابل استخراج از سایر متغیرها نمی باشد (پیرل [۹]). پوشش مارکوفی رأس X به صورت $MB(X)$ نشان داده می شود. اگر V مجموعه رئوس موجود در مجموعه داده ها باشد، داریم:

$$\forall X, Y \in V : (X \in MB(Y) \Leftrightarrow X \not\perp Y | V \setminus \{X, Y\})$$

الگوریتم GS^{12} برای یافتن اعضای پوشش مارکوفی رؤس

• ورودی و خروجی الگوریتم

ورودی: مجموعه داده.

خروجی: اعضای پوشش مارکوفی هر گره.

• شروع الگوریتم

مجموعه S را تهی در نظر بگیر؛ یعنی $S = \emptyset$.

مرحله افزایش:

به ازای همه متغیرهای موجود در مجموعه داده، اگر

$$X \not\perp Y | S$$

$$S \leftarrow S \cup \{Y\}$$

مرحله کاهش:

تا زمانی که رأسی مانند Y در مجموعه S وجود داشته باشد

$$X | S - \{Y\}$$

$$S \leftarrow S - \{Y\}$$

مجموعه باقی مانده، اعضای پوشش مارکوفی رأس X می باشند.

پایان الگوریتم GS

در الگوریتم پیشنهادی نیز امتیازات بین هر دو متغیر به شیوه روش کو و کیم محاسبه می شود؛ با این تفاوت که امتیاز، فقط بین هر متغیر و متغیرهای موجود در پوشش مارکوفی آن متغیر محاسبه می گردد (جدول ۳).

به عنوان مثال، پوشش مارکوفی رأس $meldug_1$ مجموعه

$$temp_1, lai_1, lai_0$$

$$score(meldug_1, temp_1) < score(temp_1, meldug_1)$$

$$score(meldug_1, lai_1) > score(lai_1, meldug_1)$$

$$score(meldug_1, lai_0) > score(lai_0, meldug_1)$$

بنا بر این فقط متغیر $temp_1$ به عنوان والد احتمالی $meldug_1$ انتخاب می شود. محاسبات در الگوریتم K^2 برای این متغیر به صورت زیر انجام می شود:

امتیاز	اضافه شدن به $meldug_1$
-۲۴۴۴.۴۴۳۹	---
-۲۴۴۶.۶۴۱۸	$temp_1$

نتیجه: هیچ رأس دیگری به عنوان والد به $meldug_1$ افزوده

نمی شود.

۵ نتیجه گیری

به دلیل این که امتیاز متغیر $meldug_1$ بدون داشتن والد، بیشتر از امتیاز حاصل از افزودن $temp_1$ به عنوان والد است، متغیر $meldug_1$ هیچ والدی ندارد. در مقایسه با روش کو و کیم، چهار یال اضافی که برای متغیر $meldug_1$ تولید شده بودند، در الگوریتم پیشنهادی به شبکه افزوده نمی شوند. بنا بر این خطای الگوریتم پیشنهادی کمتر از الگوریتم کو و کیم است.

در این بخش، نتایج حاصل از اجرای چهار الگوریتم با عناوین الگوریتم پیشنهادی، کو و کیم، چن و همکاران، و هروشکا و ایکن، بر روی مجموعه داده های $Hail, Car, Alarm, Asia, Midway$ که به ترتیب شامل ۸، ۳۷، ۱۸، ۵۶ و ۲۶ متغیر و ۸، ۴۶، ۱۷، ۶۶ و ۳۸ یال می باشند، ارائه شده است. برای ارزیابی، نخست شبکه های اصلی این مجموعه داده ها در نظر گرفته شده، سپس تعداد یال های انطباقی با جهت یکسان، تعداد یال های انطباقی با جهت معکوس، تعداد یال های اضافی و تعداد یال های گم شده برای هر چهار الگوریتم محاسبه شده است. میانگین نتایج حاصل از اعمال هر چهار الگوریتم بر نمونه های ۱۰۰۰۰ تایی، در شکل ۴ نشان داده شده است.

شایان ذکر است که مجموعه داده های ذکر شده همراه با شبکه های اصلی آن ها در بسته نرم افزاری $bnlearn$ در نرم افزار R موجود است.

با توجه به شکل ۴، الگوریتم پیشنهادی توانسته یکی از مشکلات روش کو و کیم را حل کند، که ایجاد یال های اضافی بود. با این حال در دو مجموعه داده $hail$ و $midway$ تعداد یال های صحیح نسبت به الگوریتم کو و کیم کاهش و تعداد یال های گم شده افزایش یافته است. اما همان طور که پیش تر نیز بیان شد چون هدف شبکه های بیزی رسیدن به گرافی است که در شناسایی ارتباطات میان متغیرها عملکرد دقیق تری داشته باشد، به این معنا که اگر دو متغیر با هم در ارتباط نباشند هیچ یال اضافی یا حتی معکوس میان آن دو وجود نداشته باشد، به منظور ارزیابی دقت الگوریتم ها، خطای شبکه برابر با مجموع یال های معکوس، اضافی و گم شده تعریف می شود و روشی دقیق تر است که بتواند در مجموع، تعداد این یال ها را کاهش دهد. همان طور که در شکل ۴ نیز مشخص است، خطای الگوریتم پیشنهادی در مقایسه با سایر روش ها به میزان قابل توجهی کاهش یافته است.

در این مقاله، روش پیشنهادی کو و کیم برای تعیین ترتیب بین متغیرها، به عنوان ورودی الگوریتم K^2 تعمیم داده شد. به این منظور برای هر رأس، در ابتدا اعضای پوشش مارکوفی آن رأس با استفاده از الگوریتم افزایشی-کاهشی تعیین شده، سپس والدین احتمالی رأس مذکور، از میان اعضای پوشش مارکوفی و با توجه به امتیاز میان هر جفت متغیر (رأس)، به دست می آید. در نهایت الگوریتم K^2 به کار رفته، ساختار شبکه بیزی حاصل می شود. نتایج حاصل از ارزیابی میزان دقت الگوریتم ها نشان می دهد که الگوریتم پیشنهادی با دقت بیشتری قادر به تشخیص ارتباطات صحیح بین متغیرها در مقایسه با شبکه های اصلی است.

همچنین به منظور ارزیابی میزان دقت هر یک از الگوریتم های فوق، از رابطه (۳) به صورت زیر استفاده شده است. رابطه (۳) در واقع بیان کننده نسبت تعداد یال های انطباقی با جهت یکسان به کل یال ها است.

$$accuracy = \frac{\#correct\ edge}{\#correct\ edge + \#graph\ error} \quad (3)$$

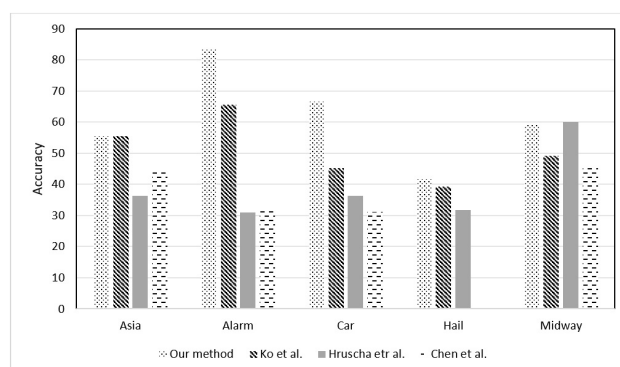
نتایج حاصل از محاسبه میزان دقت هر یک از روش ها با استفاده از رابطه (۳)، در شکل ۵ نشان داده شده است. همان طور که مشخص است، نتایج، نشان دهنده دقت زیاد الگوریتم پیشنهادی است.

جدول ۳. انتخاب والد احتمالی رأس ۱- *meldug* از جدول امتیازات پس از اعمال پوشش مارکوفی

	dm_۱	foto_۱	straaling_۱	temp_۱	lai_۱	meldug_۱	lai_۰	dm_۰	foto_۰	straaling_۰	temp_۰	lai_۰	dm_۰	foto_۰	straaling_۰	temp_۰	lai_۰	dm_۰	foto_۰	straaling_۰	temp_۰	lai_۰
dm_۱	۰	۱۳۵۲.۰۸	۹۵۸۰.۷۶۹	۱۳۵۲.۰۸	۷۷۶۴.۱۴۷	۷۱۵.۵۴۰.۷	۱۰.۵۳۹.۶۳	۱۴۵۵.۰۸	۱۵۰.۲۲۰.۴۳	۷۶۱۷.۰۰۷۸	۷۸۵۶.۹۱۶	۶۰۷۴.۵۳۰.۳	۱۰.۷۴۴.۵	۷۸۵۴۷۱۱	۷۱۴۵۷۷۲	۹۲۰.۱۹۷۵۱	۱۰.۵۸۸.۱	۱۳۵۲.۰۸	۱۵۰.۲۲۰.۴۳	۷۶۱۷.۰۰۷۸	۷۸۵۶.۹۱۶	۶۰۷۴.۵۳۰.۳
foto_۱	۰	۱۹۱.۳۳۶۲	۹۱۷۵۳.۴۳	۵۸۲۰.۹۲۸	۲۵۴.۱۳۸۸	۸۰۷۹۴.۱۸	۱۰۰.۱۱۷۵۸	۱۶۶۹۷.۸۹	۵۸۲۹.۳	۳۸۲۹.۳	۳۸۱۲۴.۲۸	۹۱۷۵۱۲۷۹	۱۹۱.۳۳۶۲	۱۵۴۰.۱۳	۱۳۹.۵۰۴	۱۸۳.۸۵۶	۲۰۸۰.۱۵	۱۹۱.۳۳۶۲	۱۵۴۰.۱۳	۱۳۹.۵۰۴	۱۸۳.۸۵۶	۲۰۸۰.۱۵
straaling_۱	۰	۴۵۵۱۷.۷۶	۴۵۵۱۷.۷۶	۴۵۵۱۷.۷۶	۴۵۵۱۷.۷۶	۴۵۵۱۷.۷۶	۴۵۵۱۷.۷۶	۴۵۵۱۷.۷۶	۴۵۵۱۷.۷۶	۴۵۵۱۷.۷۶	۴۵۵۱۷.۷۶	۴۵۵۱۷.۷۶	۴۵۵۱۷.۷۶	۴۵۵۱۷.۷۶	۴۵۵۱۷.۷۶	۴۵۵۱۷.۷۶	۴۵۵۱۷.۷۶	۴۵۵۱۷.۷۶	۴۵۵۱۷.۷۶	۴۵۵۱۷.۷۶	۴۵۵۱۷.۷۶	۴۵۵۱۷.۷۶
temp_۱	۰	۲۳۴۸۰.۹۲	۲۳۴۸۰.۹۲	۲۳۴۸۰.۹۲	۲۳۴۸۰.۹۲	۲۳۴۸۰.۹۲	۲۳۴۸۰.۹۲	۲۳۴۸۰.۹۲	۲۳۴۸۰.۹۲	۲۳۴۸۰.۹۲	۲۳۴۸۰.۹۲	۲۳۴۸۰.۹۲	۲۳۴۸۰.۹۲	۲۳۴۸۰.۹۲	۲۳۴۸۰.۹۲	۲۳۴۸۰.۹۲	۲۳۴۸۰.۹۲	۲۳۴۸۰.۹۲	۲۳۴۸۰.۹۲	۲۳۴۸۰.۹۲	۲۳۴۸۰.۹۲	۲۳۴۸۰.۹۲
lai_۱	۰	۵۲۰.۲۷۰۶	۵۲۰.۲۷۰۶	۵۲۰.۲۷۰۶	۵۲۰.۲۷۰۶	۵۲۰.۲۷۰۶	۵۲۰.۲۷۰۶	۵۲۰.۲۷۰۶	۵۲۰.۲۷۰۶	۵۲۰.۲۷۰۶	۵۲۰.۲۷۰۶	۵۲۰.۲۷۰۶	۵۲۰.۲۷۰۶	۵۲۰.۲۷۰۶	۵۲۰.۲۷۰۶	۵۲۰.۲۷۰۶	۵۲۰.۲۷۰۶	۵۲۰.۲۷۰۶	۵۲۰.۲۷۰۶	۵۲۰.۲۷۰۶	۵۲۰.۲۷۰۶	۵۲۰.۲۷۰۶
meldug_۱	۰	۱۰۹۴۱.۵۵	۱۰۹۴۱.۵۵	۱۰۹۴۱.۵۵	۱۰۹۴۱.۵۵	۱۰۹۴۱.۵۵	۱۰۹۴۱.۵۵	۱۰۹۴۱.۵۵	۱۰۹۴۱.۵۵	۱۰۹۴۱.۵۵	۱۰۹۴۱.۵۵	۱۰۹۴۱.۵۵	۱۰۹۴۱.۵۵	۱۰۹۴۱.۵۵	۱۰۹۴۱.۵۵	۱۰۹۴۱.۵۵	۱۰۹۴۱.۵۵	۱۰۹۴۱.۵۵	۱۰۹۴۱.۵۵	۱۰۹۴۱.۵۵	۱۰۹۴۱.۵۵	۱۰۹۴۱.۵۵
lai_۰	۰	۲۸۳۹۳.۴۷	۲۸۳۹۳.۴۷	۲۸۳۹۳.۴۷	۲۸۳۹۳.۴۷	۲۸۳۹۳.۴۷	۲۸۳۹۳.۴۷	۲۸۳۹۳.۴۷	۲۸۳۹۳.۴۷	۲۸۳۹۳.۴۷	۲۸۳۹۳.۴۷	۲۸۳۹۳.۴۷	۲۸۳۹۳.۴۷	۲۸۳۹۳.۴۷	۲۸۳۹۳.۴۷	۲۸۳۹۳.۴۷	۲۸۳۹۳.۴۷	۲۸۳۹۳.۴۷	۲۸۳۹۳.۴۷	۲۸۳۹۳.۴۷	۲۸۳۹۳.۴۷	۲۸۳۹۳.۴۷
dm_۰	۰	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹
foto_۰	۰	۴۸۵۰.۵۱۳	۱۰۱۲۹.۰۸	۸۰.۴۴۰.۵۵	۸۳.۸۵۶۲۵	۹۷۱.۹۳۱۳	۱۰.۵۳۵۶۵	۳۶۰.۱۸.۹۴	۰	۳۳۴.۲۷۹	۳۳۴.۲۷۹	۳۳۴.۲۷۹	۳۳۴.۲۷۹	۳۳۴.۲۷۹	۳۳۴.۲۷۹	۳۳۴.۲۷۹	۳۳۴.۲۷۹	۳۳۴.۲۷۹	۳۳۴.۲۷۹	۳۳۴.۲۷۹	۳۳۴.۲۷۹	۳۳۴.۲۷۹
straaling_۰	۰	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹
temp_۰	۰	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹
lai_۰	۰	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹
dm_۰	۰	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹
foto_۰	۰	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹
straaling_۰	۰	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹
temp_۰	۰	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹
lai_۰	۰	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹	۲۳۲۵۳۳۹

مجموعه داده	ملاک (تعداد پالها)	روش پیشنهادی	روش کو	روش هروشکا	روش چن
Asia	انطباقی با جهت درست	۵	۵	۴	۵
	گمشته	۰	۰	۱	۱/۱
	انطباقی با جهت معکوس	۳	۳	۳	۱/۹
	اضافی	۱	۱	۳	۳/۴
	خطا (جمع گمشته، معکوس و اضافی)	۴	۴	۷	۶,۳
Alarm	انطباقی با جهت درست	۴۰	۴۰	۲۲	۲۳,۹
	گمشته	۲	۲	۱	۲
	انطباقی با جهت معکوس	۴	۴	۲۳	۲۰,۱
	اضافی	۲	۱۵	۲۵	۳۰,۲
	خطا	۸	۲۱	۴۹	۵۲,۳
Car	انطباقی با جهت درست	۱۴	۱۴	۱۲	۱۱,۷
	گمشته	۱	۱	۱	۱,۷
	انطباقی با جهت معکوس	۲	۲	۴	۳,۹
	اضافی	۴	۱۴	۱۶	۲۰,۵
	خطا	۷	۱۷	۲۱	۲۵,۸
Hail	انطباقی با جهت درست	۳۰	۳۷	۳۲	-
	گمشته	۲۷	۱۹	۱۵	-
	انطباقی با جهت معکوس	۹	۱۰	۱۹	-
	اضافی	۶	۲۸	۳۵	-
	خطا	۴۲	۵۷	۶۹	-
Midway	انطباقی با جهت درست	۲۳	۲۷	۲۷	۲۳,۴
	گمشته	۱۰	۳	۳	۴,۴
	انطباقی با جهت معکوس	۵	۸	۸	۱۰,۲
	اضافی	۱	۱۷	۷	۱۳,۳
	خطا	۱۶	۲۸	۱۸	۲۷,۹

شکل ۵. نتایج حاصل از اعمال الگوریتم‌ها بر روی مجموعه داده‌های مختلف



شکل ۶. میزان دقت الگوریتم‌ها

مراجع

- [1] Chen, X. W., Anantha, G., and Lin, X. (2008). Improving Bayesian network structure learning with mutual information-based node ordering in the K2 algorithm. *IEEE Transactions on Knowledge and Data Engineering*, **20(5)**, 628-640.
- [2] Cooper, G. F. and Herskovits, E. (1992). A Bayesian method for the induction of probabilistic networks from data. *Machine learning*, **9(4)**, 309-347.
- [3] Heckerman, D. (1998). *A tutorial on learning with Bayesian networks*. Springer Netherlands.
- [4] Hruschka, E. R. and Ebecken, N. F. (2007). Towards efficient variables ordering for Bayesian networks classifier. *Data and Knowledge Engineering*, **63(2)**, 258-269.
- [5] Ko, S. and Kim, D. W. (2014). An efficient node ordering method using the conditional frequency for the K2 algorithm. *Pattern Recognition Letters*, **40**, 80-87.
- [6] Larrañaga, P., Poza, M., Yurramendi, Y., Murga, R. H., and Kuijpers, C. M. (1996). Structure learning of Bayesian networks by genetic algorithms: A performance analysis of control parameters. *IEEE transactions on pattern analysis and machine intelligence*, **18(9)**, 912-926.
- [7] Margaritis, D. and Thrun, S. (1999). *Bayesian network induction via local neighborhoods*. Carnegie-Mellon University Pittsburgh Pa Department of Computer Science.
- [8] Nielsen, T. D. and Jensen, F. V. (2009). *Bayesian networks and decision graphs*. Springer Science Business Media.
- [9] Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible reasoning*. San Mateo, CA: Kaufmann.