



مدلسازی فراگیر در آموزش الکترونیک با استفاده از شبکه های بیزین

هومن پاینده فر^۱، حسن سید رضی^۲، مسعود رهگذر^۳

^۱ دانشجوی کارشناسی ارشد، دانشگاه پیام نور
تهران، ایران
hooman.payandehfar@gmail.com

^۲ حسن سید رضی
دانشیار دانشکده کامپیوتر، دانشگاه تهران
تهران، ایران

^۳ مسعود رهگذر
استادیار دانشکده کامپیوتر، دانشگاه تهران
تهران، ایران

چکیده

در سال های اخیر، نوآوری های مهمی در سیستم های آموزشی در راستای معرفی تکنولوژی های نوینی همچون آموزش الکترونیک رخ داده است.

مساله پیش بینی رفتار فراگیر در آموزش الکترونیک، اهمیت زیادی دارد. زیرا اگر در هر مقطعی بتوان رفتار فراگیر را پیش بینی کرد، می توان محتوای آموزشی مناسب تری را به وی ارائه نمود. قبلا روش های غیر گرافیکی داده کاوی مانند طبقه بندی و قواعد انجمنی برای این منظور استفاده شده اند. ولی توجه خیلی کمی به شبکه های بیزین، به عنوان یک روش داده کاوی گرافیکی در خصوص پیش بینی رفتار فراگیر در آموزش الکترونیک شده است. لذا بر آن شدیم تا مدلی را با استفاده از شبکه بیزین به منظور پیش بینی رفتار فراگیر ارائه نماییم که هم بر اساس شهود عینی قابل درک باشد و هم اینکه قادر به پیش بینی میزان دانش فراگیر در مبحث باشد. روش های یادگیری ماشین که برای پیش بینی رفتار فراگیر بکار می روند، عموما برای طبقه بندی فراگیر استفاده می شوند، حال آنکه مدلسازی ما به وسیله شبکه بیزین، منجر به پیش بینی جوابی عددی و ملموس تر نسبت به روش های طبقه بندی می گردد. با پیش بینی میزان دانش فراگیر در مبحث، می توان قبل از به پایان رساندن مبحث، از سطح دانش و یادگیری فراگیر در آن مبحث مطلع شد و محتوای آموزشی ای که مطلوبتر اوست به وی پیشنهاد کرد. امیدواریم این مقاله گذرگاهی در جهت تحقیقات آینده باشد.

کلمات کلیدی

آموزش الکترونیک، داده کاوی، شبکه های بیزین، پیش بینی.

۱- مقدمه

مدل فراگیر توسط Holt و همکارانش [1] اینگونه تعریف شده است: "نمایشی از باورهای یک سیستم رایانه ای که در مورد فراگیر است و بنابراین، یک نمایش انتزاعی از فراگیر می باشد". بنابراین، مدل سازی فراگیران شامل مجموعه ای از طرز فکرها و برخوردهای ممکن است که یک فراگیر می تواند داشته باشد. با این حال، باید بین مدل فراگیری و مدل کاربر، که مدلی است جامع، در عمل تمایز قایل شویم. بدیهی است که فراگیران از اجزای اصلی یک سیستم آموزشی هوشمند هستند [2]. مدل کاربر در مقایسه با مدل فراگیر خیلی کلی تر است و تحقیقات آن می تواند به رفتارهایی که کلی ترند و فقط به یک حوزه خاص از رفتارهای فراگیر دلالت نمی کند، توجه داشته باشد. پس مدل کاربر، شامل حوزه هایی می شود که صرفا محدود به تدریس (آموزش) نمی شوند.

بهتر است که مدل فراگیر را در قالب طرح آموزشی و تئوری یادگیری بررسی نماییم. سه تئوری یادگیری اصلی وجود دارد که عبارتند از: رفتارگرایی، شناخت گرایی، ساخت گرایی.

قدیمی ترین تئوری یادگیری، رفتارگرایی است که به فراگیر به عنوان یک جعبه سیاه نگاه می کند. آموزنده رفتارگرا، شبیه به ماشینی است که وقتی در برابر محرکهایی قرار می گیرد، پاسخ هایی از خود نشان می دهد. کار یادگیری هنگامی صورت می گیرد که فراگیران با یک محرک خاص رو به رو شوند و از خود یک پاسخ خاص بروز دهند. پس یادگیری وقتی حاصل می شود که فراگیر مکررا با آن محرک ها رو به رو شود و پاسخ های درست وی با تشویق تقویت شود و پاسخ های غلط وی با تنبیه رد شود.

تئوری شناخت گرایی با فرض اینکه امر یادگیری امری است که شامل فراگیری ساختارهای شناختی از طریق ذخیره اطلاعات و پردازش آن اطلاعات می شود [3] با تکیه بر رفتارگرایی ساخته می شود. به عبارت دیگر در شناخت گرایی، امر یادگیری به عنوان شکل دهی و اصلاح مجدد نمایش های ذهنی حوزه مورد نظر تعریف می شود. بنابراین، فردی که از گروه یادگیران شناخت گرایی است به عنوان یک جعبه سیاه تعریف نمی شود و مدل های شناختی، این نمایش ذهنی را در او تعریف می کنند.

تئوری سوم نیز ساخت گرایی است. در حالیکه دو تئوری قبلی تئوری هایی عینی از یادگیری هستند و در آنها رفتارهایی از پیش مشخص شده و یا ساختارهای شناختی به فراگیر منتقل می شوند، تئوری سوم یک تئوری نظری است که در آن فراگیران حقیقت را بر اساس تجربه های کسب شده می سازند. دانش جدید بجای اینکه به وی انتقال داده شود، بر پایه تجارب قبلی فراگیر شکل می گیرد. ساختارهای ذهنی

موجود و نیز باورهای فراگیر برای تفسیر وقایع و اشیاء بکار می روند. بنابراین، از هر فراگیر انتظار می رود که واقعیت خاص مربوط به خودش را بسازد.

اما ITS های سنتی به همراه یک مدل فراگیر کجای این چارچوب روانشناختی جای می گیرد و ارتباط بین آنها چگونه است؟ نظریه ای که از یک تست برای ارزیابی سطح مهارت فراگیر استفاده شود مربوط به تئوری رفتارگرایی و دیدگاه رفتارگراهاست. همینطور، ITS هایی که سعی می کنند تا حالت درونی فراگیر را مدلسازی کنند در حقیقت شناخت گرا هستند. اما ساخت گراها با ITS های سنتی همخوانی ندارند. اگر هر فراگیر یک واقعیت را براساس تجارب و دانش های قبلی که خاص خودش است بسازد آنگاه فرض اینکه یک مدل فراگیر که از پیش مشخص شده، می تواند یک تعریف معقول از آن فراگیر داشته باشد، بیهوده است. لذا ارائه مدل فراگیر ما با استفاده از شبکه های بیزین هم، در قالب تئوری های رفتارگرایی و شناخت گرایی است زیرا ITS های کنونی از این دو نوع تئوری پشتیبانی می کنند.

در ادامه این مقاله، ابتدا انواع مدل های فراگیر را بیان نموده و سپس به بیان مفاهیم شبکه های بیزین می پردازیم و در نهایت مدل فراگیر پیشنهادی را با استفاده از شبکه های بیزین بیان خواهیم کرد.

۲- انواع مدل های فراگیر

انواع مدل های فراگیر عبارتند از: مدل پوششی، مدل افتراقی و مدل اختلالی [4].

۲-۱ مدل پوششی

این مدل یک مقیاس ساده از دانش را به عناصری از حوزه ای که از یک کارشناس انتظار می رود تا یادگرفته باشد تقسیم می کند و دانش حوزه ای فراگیر به عنوان یک زیرمجموعه از دانش کارشناس نمایش داده می شود.

فراگیری هر مهارت در این مدل از صفر (کاملا مبهم) تا ۱ (یادگیری در حد کارشناس) متغیر است. پس یک کارشناس با پوششی به همراه ۱ برای هر بخش از مهارت نمایش داده می شود. اگر فراگیری یک مهارت برای یک فرد ۰.۵ است این یعنی سیستم بر این باور است که فراگیر فقط دارای مهارت نیمه فراگرفته است و به تمرین زیادی نیاز دارد ولی نمی توان گفت که دارای گنگی کامل از آن مهارت است.

۲-۲ مدل افتراقی

احتمال توام $P(X)$ (همان $P(X_1, X_2, \dots, X_n)$) را داشته باشیم، آن گاه می توانیم توزیع احتمال زیر توام (Sub-joint)، $P(X_q)$ را برای $X_q \subseteq X$ با بکار بردن قاعده حاشیه ای زیر تشکیل دهیم:

$$P(X_q) = \sum_{X-X_q} P(X) \quad (2)$$

این قاعده به طور کل مجموع احتمال متغیرهایی که مورد توجه نیستند را محاسبه می کند (متغیرهایی که متعلق به $X-X_q$ می باشند) بنابراین توزیع احتمال توام، احتمال آن متغیرهایی که متعلق به X_q هستند را محاسبه نمی کند.

مهمترین پیامد پیاده سازی نظریه احتمال بیزی ذخیره سازی آن می باشد. برای روشن شدن این موضوع، یک جدول ساده نمایش $P(X_1, \dots, X_n)$ ، اگر $n=16$ و X_i ها دوتایی باشند، به $2^{16}=65536$ آرایه مختلف نیاز دارد. پس کاربرد ساده ای این نظریه بر کم بازدهی از نظر ذخیره سازی دلالت می کند.

یک روش برای برطرف کردن این کمبود جستجوی خاصیت ویژه ای از جداول احتمال شرطی $P(X_1 | X_2, \dots, X_n)$ به نام استقلال شرطی می باشد. برای تشریح این مفهوم فرض کنید توزیع احتمال توام $P(A, B, C)$ را داریم. از قاعده ضرب و خاصیت تجزیه کلی می توان بدست آورد که:

$$P(A, B, C) = P(A | B, C) P(B | C) P(C) \quad (3)$$

حال، فرض کنید که $P(A | B, C)$ دارای این خاصیت است که همیشه با $P(A | C)$ برابر می شود. این یعنی اینکه برای هر زوج (a, c) ، $P(A=a | B, C=c)$ به ازای هر مقدار B بدون تغییر می ماند. در این حالت گوئیم که A مستقل شرطی از B می باشد به ازای مقدار C داده شده.

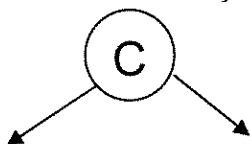
جمله اخیر را با نماد استاندارد $A \perp B | C$ بیان می کنند. (4)

بنابراین می توانیم B را از احتمال شرطی $P(A | B, C)$ حذف کنیم و رابطه (3) را بازنویسی نماییم:

$$P(A, B, C) = P(A | C) P(B | C) P(C) \quad (5)$$

از جنبه ذخیره سازی نمایش اخیر دارای اهمیت بیشتری است. توضیح بیشتر اینکه، اگر A, B, C دوتایی هایی باشند آن گاه فاکتور $P(A | B, C)$ به 2^3 درایه و $P(A | C)$ تنها به 2^2 درایه نیاز دارد.

به خاطر فواید موثر استقلال شرطی، شبکه های بیزی به سمت استقلال شرطی پیشرفت کردند. در یک شبکه بیزی متغیری که سایر متغیرها را در تجزیه مقید می کند (مثلاً C در معادله های 3 و 5، A, B را مقید کرده و B نیز در معادله 3، A را مقید کرده است) منشا (پدر یا مادر) آن شبکه می شود. این باعث ساختن یک حلقه جهت دار نمی شود بلکه یک گراف غیر چرخشی جهت دار را تشکیل می دهد که شبکه بیزی نام دارد. شکل ۱ یک شبکه بیزی برای توزیعی که در معادله 3 تعریف شد را مجسم کرده است و شکل ۲ شبکه دیگری از انواع توزیعات معادله 5 را ترسیم کرده است.



مدل افتراقی نسخه ای از مدل پوششی است. مدل افتراقی اساساً لایه ای نیز بر روی دانش مورد انتظار در نظر می گیرد. دانش مورد انتظار، دانشی است که یک فراگیر در یک نقطه خاص از زمان باید داشته باشد.

به طور کلی، مدل افتراقی از مدل پوششی کمتر سختگیری می کند زیرا فقط به خلاهای ممکن که در دانش انتظار رفته مهم اند توجه دارد و هیچ استنباطی در مورد فراگیری فراگیر که خارج از دانش انتظار رفته اوست، نمی کند.

۳-۲ مدل اختلال

یک مشکل عمده ای که به همراه مدل های پوششی و افتراقی وجود دارد، این فرضیه است که دانش فراگیر زیر مجموعه ای از دانش کارشناس است. این فرضیه، به طور کامل این واقعیت را که فراگیران می توانند حقایق و قوانین حوزه ای از دانش که کاملاً اشتباه است و یا در نتیجه تصورات غلط درباره آن حوزه پدید آمده است را استنتاج کنند، نادیده می گیرد. این دانش های نادرست کلاً به عنوان دانش ناصحیح فراگیر گفته می شوند. فلسفه مدل اختلال نمایش دادن دانش فراگیر به عنوان زیرمجموعه ای از دانش کارشناس و نیز دانش ناصحیحی است که ممکن است داشته باشد. از طریق مدل سازی درک های غلط فراگیر، سیستم بهتر قادر خواهد بود تا راه های علاج (راهکار) پیدا کند.

در ادامه ابتدا اساس شبکه های بیزی را شرح داده و سپس به ارائه مدل فراگیر پیشنهادی از نوع پوششی به وسیله شبکه بیزی خواهیم پرداخت.

۳-۱ اساس شبکه بیزی

به صورت کلی، توزیع احتمال توام با n متغیر را می توان به طور بازگشتی با استفاده از قاعده ضرب به دست آورد [5]:

(1)

$$P(X_1, X_2, \dots, X_n) = P(X_1 | X_2, \dots, X_n) P(X_2, \dots, X_n) = P(X_1 | X_2, \dots, X_n) P(X_2 | X_3, \dots, X_n) P(X_3, \dots, X_n) = P(X_1 | X_2, \dots, X_n) P(X_2 | X_3, \dots, X_n) \dots (P(X_{n-1} | X_n) P(X_n))$$

این خاصیت توزیع احتمال توام را خاصیت تجزیه کلی می نامند. هنگام استفاده از قاعده ضرب برای تشکیل توزیع احتمال توام، حاشیه سازی، توزیع احتمال توام را به توزیعی که متغیرهای آن زیرمجموعه ای از متغیرهای توزیع احتمال توام می باشد تبدیل می کند. مخصوصاً، اگر ما مجموعه پیشامدهای $X = \{X_1, X_2, \dots, X_n\}$ و توزیع

۴-۱ متغیرهای مرتبط با سنجش دانش فراگیر

برای ارزیابی دانش فراگیر از متغیرهایی در سطوح مختلف استفاده شده است که شامل موارد مفهوم، مبحث و موضوع هستند که به نوعی تداعی کننده مفاهیمی نظیر دانش، مهارت و توانایی هستند.

- یک مفهوم، عنصر ساده ای از دانش است که قابل تقسیم به اجزا کوچک تر نیست. مفاهیم به عنوان واحدهای پایه دانش مورد توجه قرار می گیرند.

برای نمایش مفهوم C از توزیع ساده باینری استفاده می کنیم که فراگیر هنگامی که مفهوم را می داند مقدارش ۱ و در غیر این صورت صفر خواهد بود. رابطه احتمال C به صورت زیر خواهد بود:

$$P(C=x) = p^x (1-p)^{1-x}$$

که p احتمال دانستن مفهوم C توسط فراگیر است و X می تواند مقادیر ۰ یا ۱ را داشته باشد.

- یک مبحث زوج مرتبی نظیر (C,w) می باشد که:
○ C مجموعه ای از عناصر $C = (C_1, \dots, C_n)$ است که مستقل از یکدیگرند.

○ $w = (w_1, \dots, w_n)$ برداری است که مقدار هر درایه آن متناسب با اهمیت هر کدام از عناصر

است و فرض بر این است که $\sum_{i=1}^n w_i = 1$.

برای محاسبه دانش فراگیر در یک مبحث، از متغیر T استفاده می شود:

$$T = \sum_{j=1}^n w_j c_j$$

۴-۲ مدل کردن ارتباطات تجمعی

برای پیچیدن در این مبحث ابتدا مفهوم فقره دانش (KI) را معرفی می کنیم که به موضوع، مبحث، مفهوم، مهارت یا موارد اینچنینی گفته می شود. ارتباط تجمعی بین یکم KI و اجزای آن KI تعریف می شود. به عنوان مثال ارتباط تجمعی می تواند بین یک موضوع و مباحث آن موضوع یا بین یک مهارت و زیرمهارت های مربوطه تعریف شود.

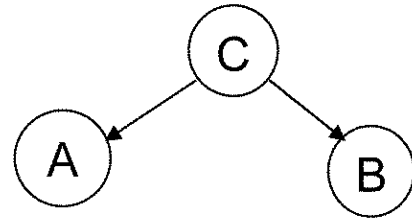
فرض کنید I یک KI است که به تعدادی KI جزئی تر مانند I_1, I_2, \dots, I_n تقسیم می شود. هر KI با یک متغیر تصادفی دودویی که می تواند دو مقدار یادگرفته شده و یادگرفته نشده به خود بگیرد، نمایش داده می شود. برای مدل کردن ارتباطات سببی بین این متغیرها دو گزینه وجود دارد:

- گزینه ۱: فرض کنیم که فقره های خاص بر روی دانستن فقره های عام تر تاثیر دارند.



شکل ۱ یک شبکه ی بیزی برای

$$P(A,B,C) = P(A | B,C)P(B | C)P(C)$$



شکل ۲ یک شبکه بیزی برای

$$P(A,B,C) = P(A | C)P(B | C)P(C)$$

برای داشتن یک شبکه بیزی پیشرفته ابتدا باید در راستای تعریف استقلال شرطی قدم برداریم قبل از اینکه به توزیعات احتمال شرطی حقیقی بپردازیم.

یک خاصیت استقلال شرطی اصلی شبکه های بیزی این است که هر متغیر X به طور شرطی از غیر اولادهای خود (NonDescendent ND(X)) مستقل است، البته به شرط داشتن والدینش $PA(X)$:

$$X \perp\!\!\!\perp ND(X) | PA(X) \quad (6)$$

یعنی اینکه، اگر والدین متغیرها معلوم باشند، آن گاه هرگونه اطلاعات بیشتری درباره ی گره هایی که در مسیر جهت دار از X قرار دارند بی-رابط (بی فایده) خواهد بود. این خاصیت را خاصیت جهت دار مارکوف از شبکه های بیزی نامند.

سوالی که اینجا بوجود می آید این است که چگونه باید از این خاصیت جهت دار مارکوف برای کاهش دادن اندازه نمایش کلی (ذخیره) استفاده کرد. قانون ضرب، ترتیب متغیرهایی که تجزیه می شوند را محدود نمی کند. بنابراین اگر بتوان ترتیبی از گره ها یافت که هر متغیر در تجزیه تنها مشروط به والدین و نااولادهای خود باشد، آن گاه نااولادها به صورت تصادفی از معادله حذف شده و هر متغیر تنها مشروط به والدین خود می باشد. یعنی، هدف ما استفاده از استقلال شرطی برای ساده کردن معادله ۱ به صورت زیر است:

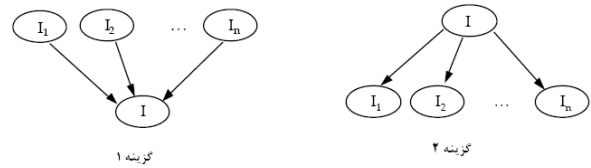
$$P(X_1, \dots, X_n) = P(X_1 | PA(X_1))P(X_2 | PA(X_2)) \dots P(X_n | PA(X_n)) \\ = \prod_{i=1}^n P(X_i | PA(X_i))$$

۴-۳ مدل سازی فراگیر با استفاده از شبکه بیزین

در مدلی که می خواهیم ارائه دهیم، باید متغیرهای مرتبط با سنجش دانش فراگیر، ارتباطات سببی و پارامترهای آن را بیان نماییم.

- از جنبه تعریف پارامتر، گزینه ۱ پیچیده تر است چون در گزینه ۱ تعداد پارامترها نمایی و در گزینه ۲ خطی است (عیب).

- گزینه ۲: فرض کنیم که دانستن یک فقره بر روی دانستن فقره های تشکیل دهنده آن تاثیر دارد. این گزینه ها به صورت گرافیکی در شکل ۳ نشان داده شده است.



شکل ۳ گزینه های مختلف برای مدل کردن ارتباطات سببی

در مرحله بعد باید ساختار پارامترهای مورد نیاز برای هر کدام از گزینه ها مورد بررسی قرار گیرد:

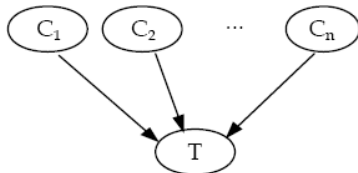
- در گزینه ۱، پارامترهای مورد نیاز شامل این موارد هستند: احتمال های پیشین هر I_i یعنی $\{P(I_i), i=1, \dots, n\}$ ، و توزیع احتمال های شرطی I یعنی $P(I | \{I_1, \dots, I_n\})$ بر اساس والدین آن است که باعث تولید $2^n + n$ مقدار می شود.
 - در گزینه ۲ این پارامترها مورد نیازند: احتمال پیشین I ، $P(I)$ و احتمال های شرطی $\{P(I_i | I), i=1, \dots, n\}$ که در این حالت $2n+1$ مقدار تولید خواهد شد.
- بررسی و تحلیل نحوه عملکرد شبکه در گزینه های مختلف نیز می تواند جالب باشد:
- در گزینه ۱، یادگرفتن فقره I_j احتمال یادگرفتن فرزندش یعنی I را تغییر می دهد.
 - در گزینه ۲، شواهد یادگیری I_j احتمال والدین آن یعنی I را تغییر می دهد که در نتیجه باعث تغییر احتمال مربوط به سایر فرزندان ($i \neq j$) می شود.

بنابراین تفاوت های اساسی بین این دو گزینه است که در گزینه ۲، یادگیری فقره I_j بر یادگیری بقیه فقره های I_i در همان سطح و ($i \neq j$) تاثیر می گذارد. دلایل محکمی برای تشخیص اینکه کدام یک از این دو گزینه بهتر است، وجود ندارد. در ادامه در طراحی این مدل بیزین با توجه به موارد زیر از سیستم آموزش الکترونیک گزینه ۱ استفاده شده است.

- از زاویه دید ارایه مسایل آموزشی، در گزینه ۱ یادگیری فراگیر به عنوان یک مساله تدریجی نگریسته شده است که با مدل های واقعی تطابق بیشتری دارد.
- به نظر نمی رسد که یادگیری I_i در یادگیری I_j تاثیرگذار باشد، موردی که در گزینه ۲ مطرح می شود.

۳-۴ ارتباط بین مفاهیم و مباحث

همانگونه که ذکر شد دانستن مفاهیم یک مبحث، بر دانستن آن مبحث تاثیر دارد. BN مرتبط با ارتباطات موجود در ساختار این مساله در شکل ۴ نشان داده شده است.

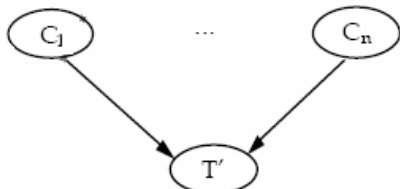


شکل ۴ BN مربوط به مدل سازی مربوط به ارتباطات بین مفاهیم و مباحث

- برای این شبکه باید پارامترهای زیر را تعریف نمود:
- احتمال های پیشین هر کدام از مفاهیم $\{p_i, i=1, \dots, n\}$.
 - احتمال های شرطی $P(T | \{C_i\}, i=1, \dots, n)$ با استفاده از مفاهیم:
- $$P(T=x | (\{C_i=1\}_{i \in S}, \{C_j=0\}_{j \notin S})) = \begin{cases} 1 & \text{if } x = \sum_{i \in S} w_i \\ 0 & \text{otherwise} \end{cases}$$

که در آن $S = \{j \in \{1, \dots, n\} \dots \text{where } C_j = 1\}$. وضعیت متغیر T به این صورت تفسیر می شود: اگر T مقدار $x \in [0,1]$ داشته باشد، این بدین معنا خواهد بود که اگر از فراگیر خواسته شود دانش خود را در n موقعیت ارایه کند وی $n \times x$ موقعیت را با موفقیت پاسخ خواهد داد.

برای درآوردن مدل فراگیر ارائه شده در قالب مدل پوششی، رفتار BN نشان داده شده در شکل ۴ را می توان با یک BN معادل شبیه سازی نمود که آن را تعریف می کنیم و در شکل ۵ نیز نشان داده ایم.



شکل ۵ BN معادل

در این شبکه تمام متغیرها به صورت باینری هستند و مقادیر آنها به صورت ۰ و ۱ می باشد. پارامترهای شبکه نیز به صورت زیر تعریف می شوند:

سیستم آموزش الکترونیک (به طور خاص بخش حل سوالات و استفاده فراگیر از دانش فرد در این راستا و پیش بینی نتیجه آموزش وی) در این مقاله نشان داده شده است. مدل فراگیر، قلب سیستم آموزشی هوشمند می باشد که می توان با روش های داده کاوی اقدام به مدلسازی فراگیر نمود. روش های غیر گرافیکی داده کاوی که به روش های یادگیری ماشین موسومند، می توانند به منظور طبقه بندی جهت پیش بینی رفتارهای آتی فراگیران بکار روند ولی در این میان، به روش های گرافیکی داده کاوی مانند شبکه های بیزین، توجه خیلی کمی جهت پیش بینی رفتارهای فراگیران شده است. لذا در این مقاله مدل سازی فراگیر را با استفاده از شبکه بیزین انجام دادیم که از آن برای پیش بینی میزان دانش فراگیر در مبحث می توان استفاده نمود و به عنوان کارهای آتی این مقاله، می توان به بکارگیری این مدل فراگیر ارائه شده، در یک سیستم آموزش الکترونیک واقعی جهت شخصی سازی پروفایل آموزشی فراگیران و ارائه محتوای آموزشی مناسبتر جهت یادگیری بهتر فراگیر، اشاره نمود.

مراجع

- [1] Tsinakos A., *Wiki and Student Models- a handshake in Distance Education field?*, Proceeding of the 3rd WSEAS/IASME International Conference on ENGINEERING EDUCATION, Vouliagmeni, Greece, July 11-13, 2006.
- [2] Romero C., *Educational data mining: A survey from 1995 to 2005*, Expert Systems with Applications: An International Journal, 2007.
- [3] Good T. and Brophy J., *Educational psychology: A realistic approach* (4th Ed.) White Plains, NY: Longman, 1990.
- [4] Urban. M., *Intelligent Tutoring Systems: An Historical Review in the Context of the Development of Artificial Intelligent and Educational Psychology*, 2003.
- [5] Jiang S. and Zhang H., *Full Bayesian Network Classifiers*, Proceeding of the 23rd International Conference on Machine Learning, Pittsburgh, 2006.

• احتمال های پیشین برای C_i ، $P(C_i) = p_i$ ، $i = 1, \Lambda, n$

• احتمال شرطی T' بر اساس مقادیر C_i :

$$P(T' | \{C_1, K, C_n\}) = \sum_{i \in S} w_i$$

متغیر تصادفی باینری T' در عالم واقعی معنای روشنی ندارد. هدف از معرفی این متغیر همان گونه که در قضیه ۱ شرح داده می شود، تعیین مقادیری است که متغیر تصادفی پیوسته T به خود می گیرد. به این ترتیب ما می توانیم BN توضیح داده شده در شکل ۵ را به جای BN شکل ۴ استفاده نماییم به گونه ای که تمام گره های آن مقدار باینری دارند و به این ترتیب توصیف و استفاده از آن را آسان تر نماییم همچنین سبب می گردد که مدل سازی فراگیر در این سیستم آموزش الکترونیک، در قالب مدل استاندارد پوششی که قبلاً آنرا توضیح دادیم، درآید.

قضیه ۱:

فرض نمایید متغیرهای تصادفی C_1, K, C_n مجموعه ای از مقادیر را به خود می گیرند و برای S زیرمجموعه $\{1, K, n\}$ ، برای هر $i \in S$ داریم $C_i = 1$ و برای هر $i \notin S$ نیز داریم $C_i = 0$. سپس متغیر تصادفی T مقدار x را خواهد داشت، اگر و تنها اگر احتمال پسینی که متغیر تصادفی T' مقدار ۱ را بگیرد x باشد یعنی:

$$P^*(T = x) = 1 \Leftrightarrow P^*(T' = 1) = x$$

اثبات: فرض کنید S زیرمجموعه ای از $\{1, K, n\}$ است که $C_i = 1$ برای هر $i \in S$ و $C_i = 0$ برای هر $i \notin S$. حال اگر x مقداری باشد که T می گیرد یعنی $x = \sum_{i \in S} w_i$ ، حال احتمال پسینی که متغیر تصادفی T' مقدار ۱ می گیرد اینگونه است:

$$P^*(T' = 1) = \sum_{i \in S} w_i = x$$

برای نشان دادن طرف کافی رابطه، فرض کنید x احتمال پسینی باشد که T' با آن ۱ می شود. حال لزوماً باید $x = \sum_{i \in S} w_i$ و لذا خواهیم

$$P^*(T = x) = 1$$

با تشکیل شبکه بیزین فوق، برای آنکه سیستم آموزش هوشمند، میزان دانش فراگیر در مبحث T' را پیش بینی نماید، ابتدا توزیع احتمال توام در شبکه بیزین فوق را برای وی تشکیل داده و در نهایت به وسیله حاشیه سازی توزیع احتمال توام توسط پارامترهای معرفی شده در فوق، میزان دانش فراگیر در مبحث که عددی بین ۰ تا ۱ خواهد بود را پیش بینی می کند.

۵- نتیجه گیری

شبکه های بیزین از روش های داده کاوی گرافیکی است که نحوه استفاده از آن برای مدل کردن رفتارهای کاربر در یک