ارزشیابی تطبیق‌پذیر و تخمین سطح دانش با استفاده از مدل دانش و به کمک شبکه بیزی

احمد کاردان[1] و صمد کاردان[2]

چکیده: در سیستم‌های مدیریت یادگیری[1] همواره ارزشیابی میزان دانش یادگیران بعنوان یک مسأله مهم مطرح است. روش‌های ابتدایی ارزشیابی دانش یادگیران تنها به محاسبه امتیاز کسب شده در آزمون‌هایی بسنده می‌کنند که سؤالات آنها به صورت ماشینی قابل تصحیح باشند. مشکلی که در اینجا ممکن است رُخ دهد پاسخ‌دهی صحیح به صورت تصادفی توسط کاربر بدون داشتن دانش کافی[2]، و یا پاسخ‌دهی اشتباه بر اثر عدم دقت[3] با وجود برخورداری از دانش کافی می‌باشد. از طرف دیگر به منظور بهبود فرایند یادگیری لازم است که نتایج آزمون‌ها به صورت تفصیلی و با تعیین نقاط ضعف و قوت کاربر مشخص شود، و بنا بر این فقط یک امتیازدهی کلی برای ارزشیابی دانش کاربر کافی نمی‌باشد. برای بدست آوردن سطح دانش کاربر به صورت تفکیک‌شده لازم است در مواردی که سؤالات چند مفهوم را مورد پرسش قرار می‌دهند تشخیص دهیم ضعف در دانستن کدام مفهوم آموزشی عامل پاسخ نادرست کاربر شده است[4]. در این پژوهش روشی بهبودیافته برای ارزشیابی دانش یادگیران ارائه شده است. در این روش مفاهیم آموزشی با استفاده از یک ساختار سلسله مراتبی اهداف آموزشی مدل‌سازی می‌شوند و از مدل دانش یادگیرنده برای تخمین میزان دانش او استفاده می‌شود. این مدل مبتنی بر شبکه‌های بیزی[5] است و با در نظر گرفتن ساختار اهداف آموزشی ارائه شده است. در روش پیشنهادی برای ارزشیابی یادگیرنده از سؤالاتی استفاده شده است که با توجه به اهداف آموزشی تهیه شده و پس از سطح‌بندی بر اساس میزان پیچیدگی به کار گرفته شده‌اند. در این پژوهش با ارائه الگوریتمی‌جدید برای روزآمد کردن گره‌ها در شبکه بیزی، تا حد امکان تأثیر پاسخ‌های صحیح تصادفی و نیز پاسخ‌های غلط ناشی از عدم دقت، در ارزیابی نهایی کاهش یافته‌اند. در نهایت با بکارگیری مدل دانش ارائه شده و استفاده از ساختار اهداف آموزشی سلسله مراتبی راه حل مؤثری برای سه مسئلهٔ مطرح در ارزشیابی یادگیران یعنی پاسخ‌های صحیح تصادفی، پاسخ‌های غلط ناشی از عدم دقت، و تشخیص ضعف‌های یادگیرنده، ارائه شده است.

کلمات کلیدی: تخمین سطح دانش یادگیران، ارزشیابی تطبیق‌پذیر، مدل دانش کاربر، شبکه‌های بیزین، آموزش الکترونیکی، یادگیری الکترونیکی

۱- مقدمه

یکی از مسائل مهم آموزشی ارزشیابی میزان دانش یادگیران است. این مسأله با شروع آموزش کلاسیک به صورت ساخت‌یافته‌تری پیگیری شد و انواع امتحانات برای سنجش دانش یادگیران طراحی شد. در سیستم آموزش سنتی نتایج این امتحانات به عنوان ملاکی برای تصمیم‌گیری در مورد سطح دانش و مهارت کسب شده و توانایی فرد برای دریافت آموزش در سطوح بالاتر مورد استفاده قرار می‌گیرد. تحقیقات جدیدتر در روان‌شناسی تعلیم و تربیت نشان دادند که برگزاری امتحان به شیوه سنتی نمی‌تواند همواره ارزیابی دقیقی از سطح دانش و یا مهارت کسب شده توسط یادگیرنده به دست دهد.

با مطرح شدن سیستم‌های آموزش الکترونیکی امکانات جدیدی به منظور انجام ارزشیابی یادگیران ارائه شد. این روش‌ها در ابتدا مشابه روش‌های سنتی بودند و تنها قالبی تحت وب از شیوه‌های مرسوم را ارائه می‌کردند. به این معنا که ارزشیابی با تعدادی سؤال از پیش طراحی‌شده انجام می‌شد. تنها تفاوت در نمره‌دهی خودکار توسط سیستم در مورد سؤالاتی بود که پاسخ مشخصی داشتند. به مرور مکانیزم‌های جدیدی مطرح شدند که به صورت موردی

تاریخ دریافت مقاله ۸۷/۱۱/۶ و تاریخ تصویب نهایی ۸۸/۳/۲۵

[1] **استادیار،** آزمایشگاه فنّاوری‌های پیشرفته در آموزش الکترونیکی دانشگاه صنعتی امیرکبیر (نویسنده مسئول)، پست الکترونیکی: aakardan@aut.ac.ir

[2] **کارشناس ارشد،** دانشکده مهندسی کامپیوتر و فنّاوری اطلاعات، دانشگاه صنعتی امیرکبیر

باعث ایجاد امکان تولید سؤالات به صورت نیمه پویا شدند. به عنوان نمونه به سیستم ارائه شده در مرجع یک می‌توان اشاره کرد که برای درس برنامه سازی کامپیوتر سؤالات را به شکل قطعه کدهایی با متغیرهای پارامتری دریافت می‌کند. این سیستم برای هر فرد مقادیر پارامترها را به صورت تصادفی تولید می‌کند و در نتیجه برای هر یادگیرنده سؤالی با پاسخ متفاوت ارائه می‌دهد. این سیستم پاسخ را بر اساس مقادیر پارامترهای به‌کار رفته محاسبه می‌کند و با پاسخ یادگیرنده مقایسه می‌کند.

در ادامه روند تکامل فرایند ارزشیابی در سیستم‌های آموزش الکترونیکی موضوع مدل کردن دانش کاربران مطرح گردید [۲]. در فرایند مدل کردن دانش کاربران با عدم قطعیت روبرو هستیم و لذا ساختارهایی که بتوان با استفاده از آن‌ها به خوبی عدم قطعیت را مدل کرد، مورد نیاز می‌باشند. یکی از ساختارهای مورد توجه برای مواجهه با عدم قطعیت شبکه های بیزی می‌باشند. مدل‌های مبتنی بر شبکه بیزی به راحتی قابل فهم توسط انسان می‌باشند و لذا در سیستم‌های آموزش الکترونیکی به فراوانی استفاده می‌شوند [۳-۱۱].

از مسائل مطرح در ارزشیابی کاربران در سیستم‌های مدیریت آموزشی، پاسخ‌دهی صحیح به صورت تصادفی توسط کاربر بدون داشتن دانش کافی، و یا پاسخ‌دهی اشتباه بر اثر عدم دقت با وجود داشتن دانش کافی می‌باشد. از مدل‌های ارائه شده برای دانش کاربران به منظور برطرف کردن این دو مشکل استفاده شده است [۷]. مدلی که در این پژوهش به کار رفته نیز از نوع شبکه بیزی می‌باشد و در ادامه مقاله مزایای مدل ارائه شده در مقایسه با مدل‌های قبلی بررسی می‌گردد.

ارزشیابی به غیر از کاربرد در فرایند سنجش و امتیازدهی دانشجویان، به منظور بهبود فرایند یادگیری و برطرف کردن نقاط ضعف کاربران نیز استفاده می‌شود. بدین منظور لازم است که نتایج آزمون‌ها به صورت تفصیلی و با تعیین نقاط ضعف و قوت کاربر مشخص شود، و بنابراین فقط یک امتیازدهی کلی برای ارزشیابی دانش کاربر کافی نمی‌باشد. استفاده از ساختار اهداف آموزشی راهکاری مناسب برای تقسیم‌بندی کل یک مبحث آموزشی به اجزای آموزشی تشکیل دهنده آن ارائه می‌دهد. در مرجع یازده از ساختار مهارت تفکیک شده برای ارزیابی نقاط ضعف و قوت دانش‌آموزان استفاده شده است. به منظور بدست آوردن سطح دانش کاربر به صورت تفکیک شده لازم است در مواردی که سؤالات در سطح پیشرفته بوده و چند مفهوم را مورد پرسش قرار می‌دهند، تشخیص دهیم ضعف در کدام مفهوم آموزشی عامل پاسخ نادرست کاربر شده است. با استفاده از ساختار اهداف آموزشی چند لایه پیشنهادی و به‌کارگیری آن در مدل دانش پیشنهادی و الگوریتم پیشنهاد شده برای روزآمد کردن مقادیر گره‌ها در مدل دانش، می‌توان با دقت مناسبی نقاط ضعف و قوت هر یادگیرنده را مشخص نمود و با استفاده از آموزش کمکی نقاط ضعف را برطرف کرد.

در ادامه، ابتدا در بخش ۲ مسائل مطرح در ارزشیابی یادگیران در آموزش الکترونیکی مورد بررسی دقیق قرار می‌گیرد. در بخش ۳ به موضوع ساختار اهداف آموزشی سلسله مراتبی پرداخته می‌شود، و بخش ۴ مبحث مدل‌کردن دانش یادگیران با استفاده از شبکه‌های بیزی را ارائه می‌دهد. در بخش ۵ روش پیشنهادی که ترکیبی است از مباحث مطرح در دو بخش ۳ و۴ ارائه می‌گردد. در این بخش مدل دانش مبتنی بر ساختار سلسله مراتبی اهداف آموزشی و الگوریتم مربوط به روزآمد کردن گره‌ها به منظور برطرف کردن سه مشکل مطرح شده در ابتدای مقاله که عبارتند از: پاسخ‌های صحیح تصادفی، پاسخ‌های غلط ناشی از عدم دقت و تشخیص ضعف‌های یادگیرنده، ارائه خواهند شد. در بخش ۶ چگونگی کاربرد این روش به منظور انجام سبک جدیدی از ارزشیابی تطبیق‌پذیر و مزایای این سبک ارائه شده است، و در بخش ۷ نتایج حاصل از بکارگیری روش پیشنهادی به منظور ارزشیابی دانشجویان در یک سیستم تدریس تحت وب، بررسی می‌شود. در بخش ۸ نتایج بدست آمده مورد بررسی قرار گرفته و پیشنهادهایی برای ادامه کار ذکر می‌گردد.

۲- مسائل مطرح در ارزشیابی آموزشی

به منظور ارزشیابی یادگیران در سیستم‌های مدیریت آموزشی و به جهت خودکار شدن فرایند ارزشیابی، معمولاً از سؤالاتی با ساختار مشابه سؤالات چند گزینه‌ای استفاده می‌شود. در این سؤالات پاسخ‌های کاربر در قالب تعداد

محدودی پاسخ از پیش تعیین شده دریافت می‌شود. در هنگام پاسخ‌دهی به این گونه سؤالات کاربر می‌تواند پاسخ صحیح را بدون داشتن دانش کافی برای پاسخ‌دهی به سؤال و به صورت تصادفی انتخاب کند که این امر دقت ارزشیابی را کاهش می‌دهد. در حالتی دیگر ممکن است کاربری بر اثر عدم دقت و با وجود داشتن دانش کافی، به سؤال پاسخ اشتباه دهد. این امر سبب ضعیف نشان داده شدن این کاربر در نتایج ارزشیابی می‌شود.

در کلاس درس عادی مدرسین با بهره‌گیری از شناخت قبلی خود از توانایی درسی یادگیرنده اقدام به اصلاح امتیاز دهی حاصل از ارزشیابی، که می‌تواند تحت تأثیر دو عامل بالا باشد، می‌کنند. در سیستم‌های آموزش الکترونیکی نیز می‌توان از شیوه‌ای مشابه استفاده نمود. به این صورت که سیستم با توجه به پاسخ‌های کاربر به سؤالات پیشین مربوط به همان هدف آموزشی می‌تواند تاثیر یک سؤال نادرست که با احتمال بالا در نتیجه عدم دقت بوده را کاهش دهد و یا هنگامی‌که کاربر اکثر سؤالات مربوط به یک هدف آموزشی را نادرست پاسخ داده است امتیاز پاسخ درستی که احتمال می‌رود به صورت تصادفی داده شده باشد را کاهش دهد.

در سیستم‌های آموزشی جدید ارزشیابی به غیر از کاربرد در فرایند سنجش و امتیاز دهی دانشجویان، به منظور مدل کردن دانش اولیه کاربران و نیز کمک در بهبود فرایند یادگیری و برطرف کردن نقاط ضعف کاربران نیز استفاده می‌شود [۱۲-۱۴]. بدین منظور لازم است که پس از انجام ارزشیابی سیستم به صورت تفصیلی سطح دانش کاربر را در هریک از بخش‌های موضوع مورد ارزشیابی و با تعیین نقاط ضعف و قوت کاربر مشخص سازد. در این گونه کاربردها فقط یک امتیازدهی کلی برای ارزشیابی دانش کاربر کافی نمی‌باشد.

به منظور تقسیم بندی مباحث آموزشی روش‌های گوناگونی استفاده شده است که عموماً بر اساس سیستم مورد استفاده و نوع مطلب انجام شده است به عنوان مثال در مرجع یازده از روش تقسیم‌بندی بر اساس مهارت‌های ریاضی مورد نیاز و خرد کردن هر مهارت به اجزای آن، استفاده شده است. با استفاده از روش‌های مطرح در طراحی آموزشی[۶] کل محتوای مورد آزمون به اجزای خردتری به نام اهداف آموزشی تقسیم می‌شوند و با استفاده از ساختار اهداف آموزشی می‌توان دانش کاربر را در هر یک از اهداف آموزشی به صورت تفکیک شده بدست آورد.

به منظور بدست آوردن سطح دانش کاربر به صورت تفکیک شده لازم است در مواردی که سؤالات در سطح پیشرفته بوده و چند مفهوم را مورد پرسش قرار می‌دهند، تشخیص دهیم ضعف در کدام مفهوم آموزشی عامل پاسخ نادرست کاربر شده است. به مثال زیر در این درباره توجه کنید. در حالت عادی هنگامی‌که سؤالی مانند Q1 بیش از یک هدف آموزشی را مورد پرسش قرار می‌دهد، به عنوان مثال OB1 و OB2 در این صورت پاسخ اشتباه به این سؤال به طور مساوی باعث کاهش تخمین دانش کاربر در مورد هردو هدف آموزشی می‌شود. به این ترتیب اگر کاربر سه سؤال قبلی مرتبط با OB1 را درست پاسخ داده باشد و تنها یک سؤال از ۴ سؤال مرتبط با OB2 را درست پاسخ داده باشد تأثیر سؤال Q1 به صورت زیر خواهد بود.

$$3 \text{ of } 3.5 = 86\% \rightarrow \text{OB1} \;;\; 1 \text{ of } 4.5 = 22\% \rightarrow \text{OB2}$$

در حالی که با توجه به این امر که کاربر به تمامی‌سؤالات قبلی مرتبط با OB1 پاسخ صحیح داده است و منطقی است تا احتمال بیشتری داده شود که کاربر به دلیل ضعف در OB2 موفق به پاسخ دادن به این سؤال نشده است. با در نظر گرفتن این موضوع تأثیر منفی[۷] این سؤال به نسبت تخمین قبلی برای این دو هدف آموزشی بین آن‌ها تقسیم می‌شود و به صورت عبارت (۱) خواهد بود.

$$\left.\begin{array}{l} 100\% \rightarrow \text{OB1} \\ 25\% \rightarrow \text{OB2} \end{array}\right\} \Rightarrow \begin{cases} Blame(OB1) = (25/125) = 0.2 \\ Blame(OB2) = (100/125) = 0.8 \end{cases} \quad (۱)$$

حال با تأثیر محاسبه شده جدید تخمین نسبت داده شده به هر یک از دو هدف آموزشی روزآمد می‌شود.

3 of 3.2 = 94% → OB1 ; 1 of 4.8 = 21% → OB2

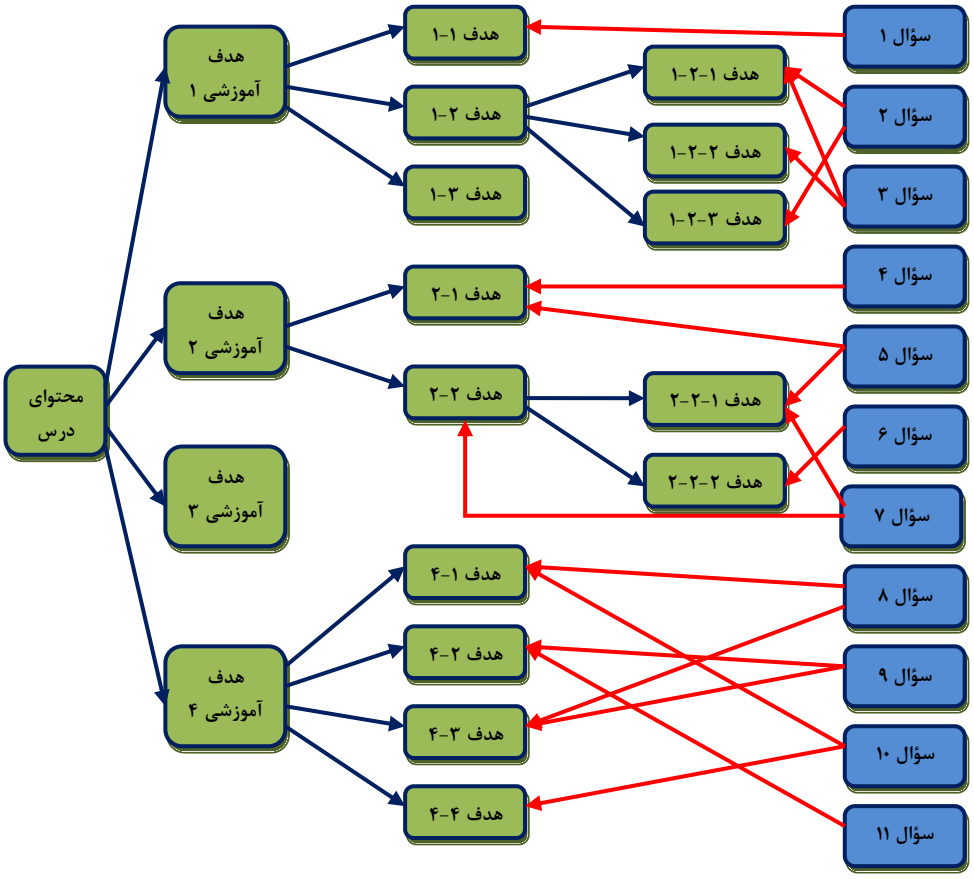
در مواردی از این دست که سؤالات در سطح پیشرفته بوده و چند مفهوم را مورد پرسش قرار می‌دهند، تشخیص این که ضعف در کدام مفهوم آموزشی عامل پاسخ نادرست کاربر شده است با استفاده از مدل دانش مبتنی بر شبکه بیزی امکان‌پذیر است.

۳- ساختار اهداف آموزشی

ساختارهای چند لایه با هدف خرد کردن موضوع مورد آموزش می‌تواند کمک مؤثری در شناخت نقاط ضعف و قوت یادگیرنده ارائه دهند.

در مراجع یازده و پانزده از ساختارهای چند لایه برای تقسیم‌بندی مهارت‌های ریاضی مورد نیاز استفاده شده‌است. در این کار از مفهوم اهداف آموزشی که یکی از مفاهیم پرکاربرد در روانشناسی آموزش و پرورش می‌باشد استفاده گردیده است. در این رابطه مروری بر طراحی آمورشی با تأکید بر اهداف آموزشی در مرجع شانزده آمده است و در مرجع هفده به اهداف تعلیم و تربیتی در طراحی قطعات آموزشی توجه شده است. همچنین در مراجع هجده و نوزده اصول حاکم بر طراحی آموزشی دروس الکترونیکی از جوانب مختلف و از جمله با تأکید بر اهداف آموزشی مورد بررسی قرار گرفته است.

در شکل (۱) یک ساختار سلسله مراتبی نمونه را مشاهده می‌کنید.



شکل ۱ ساختار مدل مفهومی‌محتوای درسی

به منظور ایجاد این ساختار، با استفاده از روش‌های مطرح در طراحی آموزشی کل محتوای مورد آزمون به اجزای خردتری به نام اهداف آموزشی تقسیم می‌شوند. در مرحله بعد، در صورت لزوم این اهداف نیز به زیر اهداف ریزتری تقسیم می‌شوند. در نهایت پس از شکستن مفاهیم آموزشی در حد مورد نیاز، ساختار سلسله مراتبی اهداف آموزشی تشکیل می‌شود. در لایه انتهایی این ساختار سؤالات دسته بندی شده و تعیین می‌گردد که هر سؤال با کدام هدف یا هدف‌های آموزشی مرتبط است و در نهایت هر سؤال بر اساس اهداف آموزشی مرتبط با آن برچسب‌گذاری می‌شود. همان‌گونه که در شکل مشاهده می‌شود سؤالات می‌توانند به یک یا چند هدف آموزشی مرتبط باشند. این حالت در مورد سؤالات پیشرفته که همزمان نیاز به داشتن چندین مهارت برای حل و یا دانستن چندین هدف آموزشی برای تحلیل و پاسخ گویی دارند، اتفاق می‌افتد.

۴– مدل دانش مبتنی بر شبکه بیزی

شبکه‌های بیزی یکی از ساختار شناخته شده در هوش مصنوعی برای مواجهه با عدم قطعیت می‌باشد [۲۰]. از شبکه‌های بیزی به دو منظور تشخیص و پیش‌بینی می‌توان استفاده نمود. به منظور پیش‌بینی، با مشاهده علت‌ها احتمال وقوع یک واقعه محاسبه می‌شود. در تشخیص، که مورد استفاده ما نیز می‌باشد، با مشاهده بروز یک واقعه می‌توان احتمال بروز آن را بر اساس هر یک از عوامل مؤثر محاسبه کرد و علت را تشخیص داد. در شکل (۲) یک نمونه شبکه بیزی را مشاهده می‌کنید.

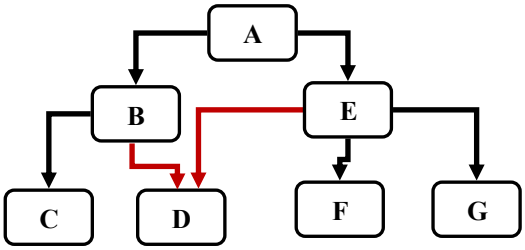
در اینجا با یک مثال فرایند تشخیص را بررسی می‌کنیم. فرض کنید مقادیر احتمال اولیه زیر برای A و B موجود باشند،

$P(A) = 0.7; P(B)=0.6$

در این صورت مقدار P(C) با توجه به جدول احتمال شرطی آن، به صورت زیر محاسبه خواهد شد.

جدول احتمال شرطی گره D

B	E	D=T	D=F
F	F	0.2	0.8
F	T	0.5	0.5
T	F	0.6	0.4
T	T	0.95	0.05



شکل ۲ یک شبکه بیزی جدول احتمال شرطی گره D

$$P(D) = P(D|\sim B,\sim E)\times P(\sim B)\times P(\sim E) + P(D|B,\sim E)\times P(B)\times P(\sim E) +$$
$$P(D|\sim B,E)\times P(\sim B)\times P(E) + P(D|B,E)\times P(B)\times P(E) = 0.05\times 0.3\times 0.4 + 0.6\times 0.7\times 0.4 +$$
$$0.5\times 0.3\times 0.6 + 0.95\times 0.7\times 0.6 = 0.663$$

حال فرض کنید در شبکه بالا واقعه D اتفاق افتاده است. می‌خواهیم احتمال B را با دانستن این واقعه روزآمد کنیم. محاسبات به صورت زیر خواهد بود.

شبکه‌های بیزی برای تقسیم تأثیر منفی پاسخ نادرست بین مهارت‌های مرتبط با آن استفاده شده است. در مرجع یازده یک مقدار ثابت برای احتمال پاسخ تصادفی و نیز احتمال پاسخ نادرست به دلیل عدم دقت در نظر گرفته شده است.

$$P(D \mid B,E) = P(D) \times P(E); P(B) = 0.7; P(E) = 0.6$$

$$P(B \mid D = true) = \frac{P(B)P(D \mid B)}{P(D)} = \frac{P(B) \times P(D \mid B,E) \times P(E) + P(B) \times P(D \mid B, \sim E) \times P(\sim E)}{P(D)} =$$

$$\frac{0.7 \times 0.95 \times 0.6 + 0.7 \times 0.6 \times 0.4}{0.663} = 0.869$$

احتمال E نیز به طریق مشابه محاسبه شده و برابر است با:

$$P(E \mid D = true) = \frac{0.489}{0.663} = 0.738$$

همان‌طور که مشاهده کردید بروز واقعه D که وابسته به علت‌های B و E بود باعث افزایش احتمال این دو علت شد. و نیز توجه داشته باشید که احتمال B، به دلیل اینکه احتمال اولیه آن و نیز میزان وابستگی D به آن بیشتر از E افزایش بیشتری داشته است. فرایند روزآمد کردن گره‌ها به صورت بازگشتی ادامه یافته و در مرحله بعد گره A روزآمد خواهد شد.

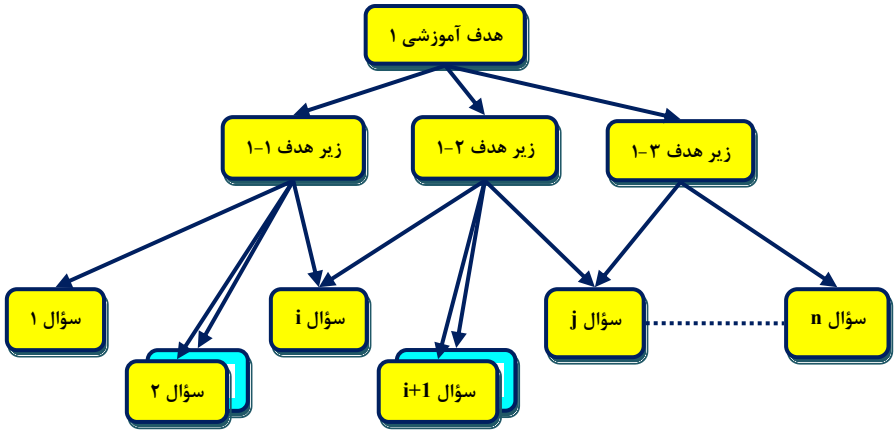
در مورد مدل‌کردن دانش یادگیران با استفاده از شبکه‌های بیزی سؤالات به عنوان وقایع محسوب می‌شوند. با هر پاسخ صحیح به یک سؤال، به این دلیل که کاربر برای پاسخ دادن صحیح به آن نیاز به دانستن اهداف آموزشی مربوطه دارد؛ احتمال آگاه بودن یادگیرنده از اهداف آموزشی آن سؤال افزایش می‌یابد. در انتها هنگامی‌که کلیه سؤالات مربوط به یک هدف آموزشی توسط یادگیرنده پاسخ داده شدند یک احتمال نهایی برای گره ریشه محاسبه می‌شود که نشان دهنده سطح دانش یادگیرنده از آن هدف آموزشی می‌باشد.

مزایای اصلی استفاده از شبکه‌های بیزی در این است که به سادگی می‌توان مسایلی نظیر احتمال پاسخ تصادفی، و نیز احتمال پاسخ‌های غلط ناشی از عدم دقت را در جدول احتمال شرطی مربوط به هر گره اعمال کرد. همچنین به سادگی می‌توان با اعمال الگوریتم روزآمدکردن روی یک گره و فراخوانی آن به صورت بازگشتی، احتمال نسبت داده شده به تمامی‌گره‌های والد آن را نیز روزآمد کرد. در [۷] از یکی از مسائل استفاده از شبکه‌های بیزی نسبت دهی یک مقدار احتمال اولیه به گره‌ها است. در این مورد در [۷و۱۱] احتمال اولیه گره‌ها برابر ۰/۵ در نظر گرفته شده است. در [۸] میانگین امتیاز دانشجویان قبلی به عنوان احتمال اولیه در نظر گرفته شده است.

۵- مدل پیشنهادی

در این بخش مدل دانش پیشنهادی ما که براساس ساختار سلسله مراتبی اهداف آموزشی و با استفاده از شبکه بیزی ساخته شده است و نیز الگوریتم پیشنهادی برای محاسبه مقادیر جدول احتمال شرطی گره‌ها را بررسی می‌کنیم.

مدل دانش: مدل پیشنهادی مجموعه‌ای از شبکه‌های بیزی است که هریک مربوط به یک هدف آموزشی می‌باشد. برگ‌های شبکه سؤالات مربوط به آن هدف آموزشی می‌باشند، لایه بالاتر را، زیر هدف‌های آن هدف آموزشی تشکیل می‌دهند و در انتها گره ریشه قرار دارد که احتمال یادگیری این هدف آموزشی را نشان می‌دهد. نمونه‌ای از این ساختار را در شکل (۳) مشاهده می‌نمایید.



شکل ۳ شبکه بیزی مربوط به یک هدف آموزشی

احتمال اوّلیهٔ گره‌ها: همان گونه که در بخش ۴ توضیح داده شد، برای محاسبه اثر هر سؤال در مدل هدف‌های آموزشی نیاز به مقدار احتمال اولیه گره‌های غیر برگ می‌باشد. بدین منظور ما در این کار از ۲ تا ۳ سؤال ابتدایی برای هر هدف آموزشی برای نسبت دهی احتمال اولیه استفاده کردیم. این سؤالات تنها یک هدف آموزشی را مورد پرسش قرار داده و ساده‌ترین سؤالات هر بخش می‌باشند. احتمال اولیه از محاسبه نمره خام پاسخ به این سؤالات بدست می‌آید. به عنوان مثال اگر کاربری برای هدف آموزشی Ob1 به ۲ سؤال از ۳ سؤال ابتدایی پاسخ صحیح دهد مقدار ۰/۶۶ به گره مربوط به Ob1 نسبت داده می‌شود.

نحوهٔ روزآمد شدن گره‌ها: همان گونه که در بخش ۴ توضیح داده شد، با پاسخ یادگیرنده به هر سؤال مدل(های) مربوط به هدف یا اهداف آموزشی آن سؤال مطابق روش ذکر شده در بخش ۴ روزآمد می‌شود. این روش روزآمد کردن دو قابلیّت زیر را تأمین می‌کند:

۱- برای هر گره سؤال با پاسخ صحیح، تاثیر مثبت آن پاسخ، بین گره‌های والد آن گره سؤال، به نسبت میزان فعلی احتمال گره‌های والد توزیع می‌شود. به مثال بخش ۴ مراجعه شود.

۲- برای هر گره سؤال با پاسخ نادرست، تأثیر منفی آن پاسخ، بین گره‌های والد آن گره سؤال، به نسبت عکس میزان فعلی احتمال گره‌های والد توزیع می‌شود. محاسبات مشابه مثال بخش ۴ و برای ~D خواهد بود. این شیوه باعث می‌شود تا نقاط ضعف در اهداف آموزشی برای هر سؤال به خوبی مشخص شود.

برای مسأله پاسخ‌های صحیح تصادفی و پاسخ‌های نادرست ناشی از عدم دقت، از جدول احتمال شرطی استفاده می‌شود. بر خلاف [۱۱] در این پژوهش مقادیر اعمال شده ثابت نیستند بلکه متناسب با مقدار فعلی گره‌های والد تغییر می‌کنند. برای توضیح بیشتر، یک مثال ارائه می‌شود. سؤال پنج گزینه‌ای Q1 را در نظر بگیرید که دارای دو گره والد Sob1 و Sob2 می‌باشد. مقدار پایه احتمال پاسخ تصادفی (انتخاب جواب درست ازبین ۵ گزینه) برای این سؤال ۰/۲۰ می‌باشد. این به این معناست که در جدول احتمال شرطی برای حالتی که پاسخ درست باشد باید احتمال ۰/۸ قرار داده شود. ما در این کار از یک مکانیزم سه سطحی برای محاسبه میزان احتمال پاسخ تصادفی (c) استفاده کردیم. در این مکانیزم سه سطح دانش بر اساس مقدار احتمال گره والد و یا حاصلضرب گره‌های والد (کمتر از ۰/۵، بین ۰/۵ و ۰/۸ و بیشتر از ۰/۸) تعریف می‌شود. در سطح یک دانش کاربر ناچیز تلقی شده و مقدار c دو برابر مقدار پایه در نظر گرفته می‌شود. در سطح دو مقدار دانش کاربر متوسط است و مقدار c برابر مقدار پایه در نظر گرفته می‌شود. برای سطح سه، مقدار c برابر ۰/۷۵ مقدار پایه در نظر گرفته می‌شود.

برای مواجهه با پاسخ‌های نادرست ناشی از عدم دقت، نیز از همین مکانیزم سطح بندی با تغییر مرز سطوح استفاده می‌شود (کمتر از ۰/۷، بین ۰/۷ و ۰/۹ و بیش از ۰/۹). برای

این سه سطح مقدار احتمال خطای ناشی از عدم دقت (s) به ترتیب ۰/۰۵ و ۰/۱ و ۰/۱۵ خواهد بود. این احتمال به این معنا است که اگر پاسخ کاربر نادرست باشد به جای امتیاز صفر با در نظر گرفتن احتمال عدم دقت، درصدی از امتیاز به او داده می‌شود. در اینجا با افزایش سطح دانش فرد احتمال این‌که پاسخ نادرست چنین فردی ناشی از عدم دقت او بوده باشد افزایش می‌یابد. به این ترتیب برای سؤال Q1 مقادیر اختصاص یافته بر اساس پاسخ کاربر مطابق جدول (۱) خواهد بود.

جدول ۱ مقادیر گره Q1 با احتساب پاسخ تصادفی و عدم دقت

پاسخ کاربر	مقدار گره
صحیح	if (p(SOb1)*p(SOb2))>0.8 = 0.8 else if (p(SOb1)*p(SOb2))>0.5 = 0.7 otherwise = 0.6
نادرست	if (p(SOb1)*p(SOb2))>0.9 = 0.2 else if (p(SOb1)*p(SOb2))>0.5 = 0.1 otherwise = 0.05

این اصلاحات باعث بهبود قابل توجه نتایج نسبت به مرجع یازده گردید که در بخش ۷ توضیح داده می‌شود.

۶- تخمین سطح دانش تطبیق یافته

ارزشیابی تطبیق‌پذیر[۸] تاکنون مورد توجه زیادی از جانب محققین قرار گرفته است [۲۱-۲۵]. در روش‌های کنونی سؤالات بر اساس سطح دشواری تقسیم بندی می‌شوند. به منظور انتخاب سؤال با سطح دشواری مناسب، سطح دانش کاربر به دو روش تخمین زده می‌شود. در روش اول برای تخمین سطح دانش یادگیرنده از تعدادی سؤال اولیه که از سطح ساده شروع شده و به تدریج دشوارتر می‌شوند، استفاده می‌شود. سپس سیستم بر اساس پاسخ‌های اولیه یادگیرنده اقدام به انتخاب باقی مانده سؤالات در سطح دشواری متناسب با سطح دانش یادگیرنده می‌نماید. هر پاسخ صحیح، سطح تخمینی را افزایش و هر پاسخ نادرست این مقدار را کاهش می‌دهد. این سؤالات به منظور بهبود دقت ارزشیابی استفاده می‌شوند تا سرانجام الگوریتم مقدار نسبتا ثابتی را برای سطح دانش یادگیرنده بیابد و به آزمون پایان دهد [۱۵].

در روش دوم مقداری پیش‌فرض برای سطح دانش یاد گیرنده، با استفاده از فعالیت‌های قبلی وی در سیستم، و از جمله میزان مطالعه محتوای درسی مربوطه برای هر یادگیرنده اختصاص یافته است. سیستم با استفاده از این تخمین اولیه اقدام به انتخاب سؤالات خواهد کرد، و سپس با توجه به پاسخ‌های یادگیرنده، همانند روش بالا مقداری نهایی برای سطح دانش کاربر محاسبه خواهد شد [۲۴].

در این دو روش همواره این سؤال مطرح است که آیا آزمون بطور منصفانه عمل کرده است؟ برای توضیح این مسأله به مثال زیر توجه کنید. الگوریتم تطبیق پذیری یک سیستم ارزشیابی تطبیق‌پذیر به منظور سادگی با قوانین زیر فرض می‌شود.

۱- ارزشیابی از ۲ سطح پایین‌تر از سطح اولیه شروع شده و با ۲ پاسخ صحیح به سؤالات این سطح، تا رسیدن به سطح اولیه افزایش می‌یابد.
۲- اگر در هر سطحی کاربر پاسخی نادرست داد، در صورتی که سطح اولیه بالاتر از آن بود سطح اولیه به آن سطح کاهش می‌یابد.
۳- سطح اولیه و سطوح بالاتر با ۳ سؤال صحیح تثبیت می‌شوند.
۴- آزمون با ۳ پاسخ نادرست در سطح n پایان یافته و سطح n-1 به عنوان سطح دانش یادگیرنده خواهد بود.

دو یادگیرنده با سطح دانش یکسان پس از این‌که به یکی از دو روش بالا سطح اولیهٔ دانش آن‌ها تخمین زده شود وارد مرحلهٔ پاسخگویی به سؤالات تطبیق یافته می‌شوند. یاد گیرنده A با تخمین اولیه ۷۰ از صد نمره وارد سیستم می‌شود و به سؤالاتی از سطح ۵۰ و ۶۰ و ۷۰ امتیاز، پاسخ صحیح می‌دهد، و سپس به سه سؤال از سطح ۸۰ پاسخ نادرست می‌دهد. در اینجا سیستم با ۱۰ سؤال سطح این کاربر را ۷۰ ارزیابی می‌کند. حال یادگیرنده B را در نظر بگیرید که با تخمین اولیه ۶۰ وارد سیستم می‌شود. این کاربر در ابتدا در حین پاسخ دادن به یک سؤال از سطح ۴۰ دچار لغزش شده و به علت عدم دقت به این سؤال پاسخ نادرست می‌دهد. حال سیستم سؤالات دیگری از سطح ۴۰ را در مقابل او قرار می‌دهد. در صورتی که الگوریتم کاملی با تعداد سؤال نامحدود در سیستم وجود داشته باشد

یادگیرنده B ناچار است با تعداد ۹ سؤال با پاسخ صحیح، به سطح ۶۰ باز گردد. حال این یادگیرنده باید به ۳ سؤال از سطح ۷۰ نیز پاسخ صحیح دهد و نیز سؤالات سطح ۸۰ را نیز پاسخ دهد تا سطح ۷۰ برای او تثبیت شود و حال فرض کنید که هیچ یک از یادگیرندگان A و B نمی‌توانند به یکی از سؤالات سطح ۷۰ درصد توسط موجود پاسخ دهند و این سؤال تنها از یادگیرنده B پرسیده می‌شود و در نتیجه امتیاز کسب شده یادگیرنده B، برابر ۶۸ می‌گردد. به این ترتیب یادگیرنده B با تعداد ۱۱پاسخ صحیح از ۱۶ سؤال به نتیجه یادگیرنده A با ۹ پاسخ صحیح از ۱۲سؤال نیز دست نیافت. با توجه به اینکه تعداد سؤالاتی که B پاسخ داده ۲۵ درصد بیش از یادگیرنده A است احتمال بروز اشتباه ناشی از عدم دقت توسط وی بیشتر است و نیز در نتیجه سؤالات بیشتر، دانش وی بیشتر مورد آزمون قرار گرفته است و با وجود تلاش بیشتر وی نتیجه‌ای تقریبا مشابه یادگیرنده A کسب کرده است، از طرف دیگر سؤالی از یادگیرنده B پرسیده شد که هیچ یک از دو نفر از آن اطلاع نداشتند و درنتیجه با وجود سطح دانش کاملاً برابر، امتیاز یادگیرنده B کمتر از یادگیرنده A گردید که منصفانه نیست. برای درک بهتر غیر منصفانه بودن شرایط فرض کنید که یادگیرنده B همان یادگیرنده A بوده است که بار دیگر در آزمون شرکت کرده است.

باید توجه داشته باشیم که علاوه بر مشکل بالا در موارد عادی نیز تنها در صورتی امتیاز کسب شده افراد مختلف در یک سیستم قابل مقایسه‌اند که، از سؤالاتی که سطح دشواری آن‌ها در آزمون دیگری قبلاً استاندارد سازی شده باشد استفاده کنیم و این امر همواره امکان پذیر نیست. در این پژوهش ما روشی متفاوت را در پیش گرفتیم، به این معنا که سؤالات آزمون‌ها برای تمامی کاربران یکسان انتخاب شدند و تطبیق‌پذیری در امتیازدهی به پاسخ‌های افراد انجام گردید. در این روش سؤالات از ساده به دشوار مرتب شده‌اند و کاربران به تمامی سؤالات پاسخ می‌گویند و به سؤالات نادرست نمره منفی به صورت معمول تعلق نمی‌گیرد. امروزه حذف نمره منفی امری عادی در آزمون‌های استاندارد شده است زیرا این امر باعث می‌شود تا استرس آزمون برای کاربران کاهش یافته و نیز تمامی‌توان آن‌ها در آزمون به نمایش گذاشته شود.

نوآورهای اصلی این پژوهش ارائه مدل دانش مبتنی بر ساختار اهداف آموزشی و الگوریتم مربوط به روزآمد کردن گره‌ها در آن می‌باشد. علاوه بر آن‌ها در اینجا ما ایده امتیازدهی تطبیق‌پذیر برای یک آزمون با سؤالات ثابت را مطرح می‌کنیم که می‌تواند جایگزینی برای شیوه فعلی ارزشیابی تطبیق‌پذیر گردد. این روش امتیازدهی با در نظر گرفتن احتمال پاسخ نادرست با وجود داشتن دانش کافی و احتمال پاسخ صحیح تصادفی به صورت متغیر در طول امتیازدهی به سؤالات از ساده به دشوار صورت می‌گیرد. این امتیازدهی تطبیق‌پذیر با بکارگیری مدل بیزی ارائه شده برای دانش کاربر و الگوریتم پیشنهادی برای روزآمد کردن گره‌ها در آن امکان‌پذیر می‌شود.

از مزایای روش پیشنهادی در این کار میزان منصفانه بودن آزمون برای افراد مختلف را افزایش می‌دهد. زیرا بر خلاف دو روش دیگری که در این مقاله از آنها یادآوری شد سؤالات برای همهٔ کاربران یکسان خواهد بود و از طرف دیگر تعداد سؤالات برابر، تلاش یکسانی را می‌طلبد. مزیّت مهم دیگر این است که تنها یک برآورد اوّلیه از سطح دشواری سؤالات کافی است و نیازی به به استانداردسازی سؤالات در آزمون‌های قبل نمی‌باشد.

۷– روش تحقیق

به منظور ارزیابی روش پیشنهاد شده از سیستم تدریس تحت وب - سیستمی‌که تنها قسمت ارائه محتوای آموزشی و ارزشیابی را به صورت تحت وب پشتیبانی می‌کند- پیاده‌سازی شده در آزمایشگاه فناوری‌های پیشرفته در آموزش الکترونیکی استفاده گردید. این سیستم برای آموزش بخشی از یکی از دروس دوره مجازی دانشگاه صنعتی امیرکبیر به ۶۱ نفر از دانشجویان مورد استفاده قرار گرفت. برای مقایسه دقت و ارزیابی عملکرد روش پیشنهادی، دانش یادگیران با استفاده از این روش به صورت بخش به بخش در طول دوره آموزش تخمین زده شد. در ادامه دانش یادگیران با استفاده از یک امتحان جامع که به امتیازدهی آن به روش سنتی انجام شد نیز موردانداره گیری قرار گرفت.

در این پژوهش ابتدا یک درس الکترونیکی تحت عنوان تجارت الکترونیکی طراحی و برای بخش‌های مختلف آن

آزمون‌هایی توسط استاد درس طراحی شد. برای اجرای دقیق این مرحله، اطلاعات لازم درباره هدف پژوهش در اختیار استاد درس قرار گرفت و آزمون‌های طراحی شده به صورت ماژول‌های مستقل و قابل ارائه از طریق شبکه آموزش الکترونیکی طراحی و پیاده‌سازی شدند. بنابراین، یادگیرنده پس از هر مرحله یادگیری محتوای آموزشی، قادر به برگزاری آزمون مربوطه می‌باشد. آزمون‌ها کوتاه و به صورت تست‌های چندگزینه‌ای طراحی شدند، و در هنگام برگزاری آزمون کلیه مشخصات یادگیرنده، پاسخ‌های داده شده، و اطلاعات زمانی آزمون ثبت گردید. سطح دانش یادگیرندگان بر اساس آزمون‌های مرحله‌ای صورت پذیرفت و برای ارزیابی رضایت بخش بودن نتایج بدست آمده، در پایان دورهٔ آموزشی یک آزمون جامع نیز از یادگیرندگان گرفته شد. نهایتاً تحلیل نتایج حاصل از این آزمون و نتایج حاصل از آزمون‌های مرحله‌ای برای یافتن سطح رضایت بخش بودن روش پیشنهادی مورد استفاده قرار گرفتند. بطور روشن می‌توان روش اجرایی این پژوهش را در مراحل زیر خلاصه نمود:

- انجام آزمون های کوتاه در طول استفاده از سیستم
- تخمین میزان دانش یادگیرندگان برای هریک از مفاهیم آموزشی موجود در مدل و با استفاده از آزمون‌های کوتاه
- انجام آزمون جامع در پایان دوره و ارزشیابی آن با شیوه سنتی و احتساب نمره منفی برای پاسخ‌های غلط
- برگزاری آزمون جامع و تفکیک سؤالات این آزمون به مفاهیم مورد نیاز برای پاسخ دادن به آن‌ها
- تخمین احتمال پاسخ دادن به هر سؤال با استفاده از مدل و محاسبه نمره آزمون با استفاده از تخمین‌ها
- مقایسه نمرات کسب شده از آزمون‌ها با نمرات حاصل از روش تخمین

۸- نتایج و بحث

دورهٔ آموزشی مورد نظر در این پژوهش در طول مدت یک ماه به صورت مستمر و با برنامه‌ریزی زمانی از قبل تعیین شده به اجرا درآمد. کلیهٔ دانشجویان نسبت به مراحل مختلف دوره و وظایفی که به عهده داشتند توجیه شدند و در طی مدت برگزاری دورهٔ از طریق پست الکترونیکی مورد حمایت و پشتیبانی فنّی قرار گرفتند. اطلاعات آماری مربوط به بررسی انجام شده در جدول (۲) آمده است.

جدول ۲ نتائج آماری آزمون

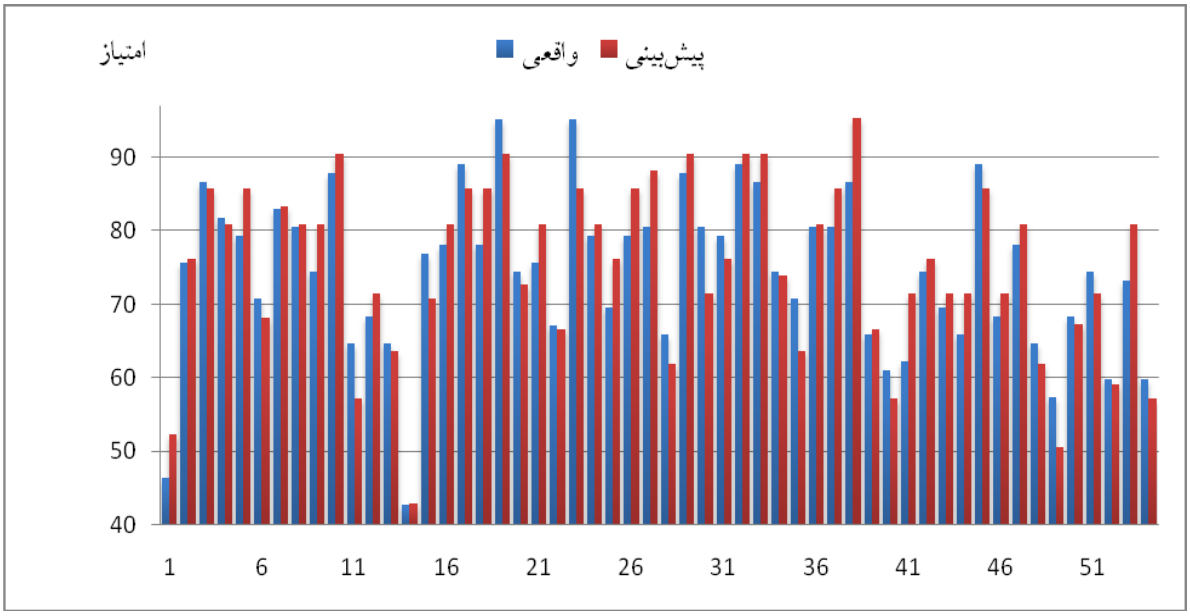
تعداد دانشجویان	۵۴ نفر
تعداد سؤالات طول دوره	۱۷۵ سؤال
تعداد سؤالات امتحان جامع	۸۲ سؤال

سپس نتایج امتیاز کسب شده در امتحان جامع با امتیاز پیش بینی شده توسط سیستم مورد مقایسه قرار گرفتند. بدین منظور فرایند زیر برای هر یادگیرنده انجام شد. ابتدا امتیاز آزمون جامع از ۱۰۰ محاسبه شد. سپس سطح دانش یادگیران بطور کلی و بر اساس میانگین وزن دهی شده برای سطح دانش یادگیرنده (با توجه به نسبت سؤالات هر هدف آموزشی به کل سؤالات آزمون) برای تمام اهداف آموزشی محاسبه گردید. از آنجا که این احتمال عددی بین صفر و یک است، به صورت درصد تبدیل شد و با امتیاز کسب شده در آزمون مقایسه گشت. برای بررسی نتایج بدست آمده، قدر مطلق تفاوت بین امتیاز واقعی و امتیاز پیش‌بینی شده محاسبه و مورد بررسی آماری قرار گرفت. نتایج این بررسی در جدول (۳) نشان داده شده است. محاسبات آماری بطور بسیار شفاف نشان می‌دهند که نمرات تخمین زده شده با نمرات حاصل از امتحان مشابهت قابل توجهی دارند. در آزمون‌ها از آزمون t-student برای بررسی اینکه میانگین دو گروه تفاوت معنی‌دار آماری دارند استفاده شد، و در این حالت فرض صفر این است که دو گروه میانگین مشابه دارند، و این همان هدف پژوهش انجام گرفته است که بتواند این امر را نشان دهد. مقدار p-value برای این آزمون 0.159782 شده است و به این ترتیب فرض صفر قابل رد کردن نمی‌باشد، و میانگین نمرات پیش بینی شده با نمرات کسب شده در آزمون جامع مشابه هستند.

نمودار نشان داده شده در شکل (۴) امتیازات واقعی کسب شده و امتیازات پیش‌بینی شده را نشان می‌دهد.

جدول ۳ مقایسه نتایج پیش بینی شده با نتایج واقعی

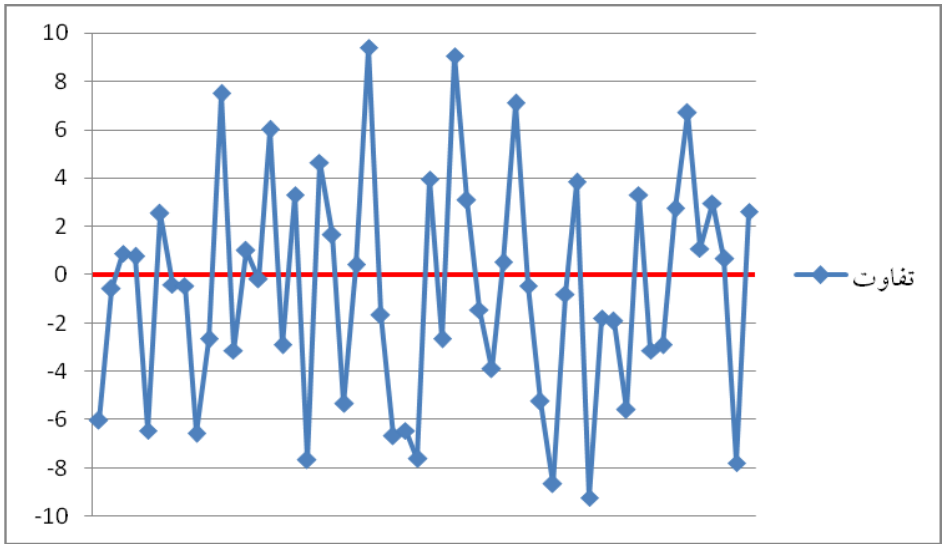
شاخص	مقدار (از ۱۰۰)
بیشترین تفاوت	۹/۴۰۷
میانگین قدرمطلق تفاوت‌ها	۳/۸۱۵
انحراف معیار تفاوت‌ها	۲/۷۵۶۷
میانگین امتیاز کسب شده	۷۴/۳۹
میانگین امتیاز پیش‌بینی شده	۷۵/۰۳۲



شکل ۴ نمودار امتیازات دانشجویان

همان‌گونه که مشاهده می‌کنید پیش‌بینی انجام شده تا حد بسیار مناسبی با نتایج آزمون جامع انجام شده مطابقت دارد به طوری که بالاترین میزان تفاوت کمتر از ۱۰ درصد است و میانگین قدر مطلق تفاوت‌ها نیز در حد ۳/۸ درصد با انحراف معیار ۲/۷۶ می‌باشد. نمودار تفاوت‌ها بر حسب درصد را در شکل (۵) مشاهده می‌نمایید. چگونگی توزیع تفاوت‌ها در جدول (۴) آمده است.

این نتایج در مقایسه با تنها مرجع موجود که از این روش برای ارزیابی کار خود استفاده کرده است مرجع یازده بهبود قابل توجهی را نشان می‌دهد. در مرجع یازده میانگین میزان خطا ۱۴/۵۲ درصد می‌باشد. عوامل بهبود حاصل در این پژوهش را می‌توان اینگونه تحلیل نمود. در مدل ارائه شده در این کار اهداف آموزشی تا حد امکان و بصورت پویا خرد شده است در حالی که در مرجع ذکر شده لایه‌های مهارت بصورت ایستا به کار گرفته شده‌اند. نمودار نشان داده شده در شکل (۴) امتیازات واقعی کسب شده و امتیازات پیش‌بینی شده را نشان می‌دهد. در این پژوهش الگوریتم به کار رفته برای روزآمد کردن گره‌ها به طور



شکل ۵ مقایسه امتیازات پیش‌بینی شده با نتایج واقعی برحسب درصد

جدول ۴ توزیع تفاوت‌ها در نتایج بدست آمد

	تعداد	درصد
تفاوت بین صفر تا ۵ از ۱۰۰	۳۵	۶۴/۸۱۵
تفاوت بین ۵ تا ۱۰ از ۱۰۰	۱۹	۳۵/۱۸۵
تفاوت بیش از ۱۰ از ۱۰۰	۰	۰
جمع	۵۴	۱۰۰

خاص با در نظر گرفتن توانایی تشخیص نقاط ضعف و قوت به کار رفته است و در نتیجه دقت تخمین دانش در سؤالاتی که با چندین هدف آموزشی در ارتباط هستند افزایش یافته است. علاوه بر آن محاسبه احتمالات مربوط به پاسخ صحیح تصادفی و پاسخ نادرست ناشی از عدم دقت، به گونهٔ مؤثری بهبود یافته است. در پایان باید توجه داشت که با توجه به ماهیّت درس مورد استفاده در پژوهش فعلی میزان سؤالات با اهداف آموزشی متعدد نسبتاً کم می‌باشد که می‌تواند عامل مؤثری در بهبود نتایج باشد.

۹- نتیجه‌گیری

در این مقاله ما به سه موضوع مطرح در تخمین دانش و ارزشیابی یادگیران که عبارتند از پاسخ‌های صحیح تصادفی، پاسخ‌های غلط ناشی از عدم دقت و تشخیص ضعف‌های یادگیرنده پرداختیم. مدل دانش پیشنهادی ما که بر اساس ساختار سلسله مراتبی اهداف آموزشی و توسط شبکه بیزی پیاده سازی می‌شود راه حلی مناسب برای این سه مسأله می‌باشد. در ادامه ما روش جدیدی را پیشنهاد نمودیم که بر اساس مدل و روش‌های ارائه شده در این مقاله عمل می‌کند. نتایج بسیار امیدوارکننده‌ای از پیاده‌سازی روش پیشنهادی بدست آمد که راه را برای ادامه کار هموار می‌سازد.

در این کار با استفاده از شبکه‌های بیزی از سؤالات به عنوان وقایعی استفاده شد که مدل‌کردن دانش یادگیران را امکان‌پذیر ساخت. همچنین در روش پیشنهادی نشان داده شد که با هر پاسخ صحیح به یک سؤال، احتمال آگاه بودن یادگیرنده از اهداف آموزشی مربوط به آن سؤال افزایش می‌یابد. در نتیجه پس از پاسخگویی به تمام سؤالات مربوط به یک هدف آموزشی، می‌توان احتمال نهایی را محاسبه کرد و نشان داد که سطح دانش یادگیرنده در آن هدف آموزشی چه می‌باشد. علاوه بر این، در این کار نشان داده

شد که با استفاده از ساختار سلسله مراتبی اهداف آموزشی می‌توان مفاهیم آموزشی را مدل‌سازی نمود و با استفاده از مدل دانش یادگیرنده، میزان دانش او را با دقّت قابل توجهی تخمین زد. از نوآورهای این پژوهش می‌توان به ارائه مدل دانش مبتنی بر ساختار اهداف آموزشی، ارائه الگوریتمی برای روزآمد کردن گره‌ها در شبکه بیزین، و ایده امتیازدهی تطبیق‌پذیر برای یک آزمون با سؤالات ثابت را مطرح نمود. بنابراین، تا حد امکان می‌توان تأثیر پاسخ‌های صحیح تصادفی و یا پاسخ‌های غلط ناشی از عدم دقت را در ارزیابی نهایی کاهش داد.

از مواردی که می‌تواند به عنوان ادامه فعالیت‌ها و در ارتباط با این پژوهش تلقی شود ارائه تطبیق‌پذیر مواد آموزشی توسط سیستم آموزش الکترونیکی و بر اساس سطح دانش تخمینی می‌باشد. هم‌اکنون ادامه این کار در آزمایشگاه فناوری‌های پیشرفته در آموزش الکترونیکی در دانشگاه صنعتی امیرکبیر در حال اجر است. همچنین می‌توان روش پیشنهادی را با دخیل کردن سایر خصوصیات کاربر، و به عنوان مثال وضعیت احساسی کاربر، در جهت تطبیق‌پذیری دقیق و جامع‌تر ارتقاء بخشید.

پی نوشت

[1] Learning Management Systems
[2] Guessing
[3] Slipping
[4] Credit –Blame Problem
[5] Bayesian
[6] Instructional Design
[7] Blame
[8] Addaptive Assessment

مراجع

[1] Brusilovsky P. and Sosnovsky S., *Individualized exercises for self-assessment of programming knowledge: An evaluation of quizpack*, ACM Journal of Educational Resources in Computing, Vol. 5, No. 6, **2005**.

[2] VanLehn K. and Martin J., *Evaluation of an assessment system based on Bayesian student modeling*, International Journal of Artificial Intelligence and Education, Vol. 8,No 2, **1997**, pp. 179-221.

[3] Martin J. and Vanlehn K., *Student assessment using Bayesian nets*, International Journal of Human-Computer Studies, Vol. 42,No. 6, **1995**, pp. 575–591.

[4] Conati C. and Vanlehn K., *POLA: A student modeling framework for probabilistic on-line assessment of problem solving performance*, In Proceedings of UM-96, 5th International Conference on User Modeling, User Modeling, Inc., **1996**, pp. 75-82.

[5] VanLehn K. and Martin J., *Evaluation of an assessment system based on Bayesian student modeling*, International Journal of Artificial Intelligence and Education,Vol. 8, No. 2, **1997**, 179-221.

[6] Mayo M. and Mitrovic A., *Optimizing ITS behavior with Bayesian networks and decision theory*, International Journal of Artificial Intelligence in Education, Vol. 12, **2001**, pp. 124–153.

[7] Conati C., Gertner A. and Vanlehn K., *Using bayesian networks to manage uncertainty in student modeling*, User Modeling and User-Adapted Interaction, Vol. 12,No. 4, **2002**, pp.371-417.

[8] Desmarais M.C. and Pu X., *A bayesian inference adaptive testing framework and its comparison with item response theory*, International Journal of Artificial Intelligence in Education, Vol. 15, **2005**, pp. 291-323.

[9] Vanlehn K., Lynch C., Schulze K., Shapiro J. A., Shelby R., Taylor L., Treacy D., Weinstein A., and Wintersgill M., *The andes physics tutoring system: Lessons learned*, International Journal of Artificial Intelligence in Education, Vol. 15,No. 3, **2005**.

[10] Bekele R. and Menzel W., *A Bayesian Approach to Predict Performance of a Student (BAPPS): A case with ethiopian students*, In Proceedings of International Conference of Artificial Intelligence and Applications, AIA-2005, Vienna, Austria, **2005**.

[11] Pardos Z.A., Heffernan N.T., Anderson B. and Heffernan C., *Using fine-grained skill models to fit student performance with Bayesian networks*, On-line Proceedings of the Workshop on Educational Data Mining at the Eighth International Conference on Intelligent Tutoring Systems, Taiwan, **2006**, pp.5-12.

[12] Arroyo I, Beal C., Murray T., Walles R. and Woolf B. *Wayang outpost: Intelligent tutoring for high stakes achievement tests*, In Proceedings of the 7th International Conference on Intelligent Tutoring Systems (ITS2004), Springer Berlin / Heidelberg, **2004**, pp. 468-477.

[13] Razzaq L. and Heffernan N.T., *Towards designing a user-adaptive web-based e-learning system*, In CHI '08 Extended Abstracts on Human Factors in Computing Systems, Florence, Italy, April, **2008**,pp. 3525-3530.

[14] Kardan A. and Kardan S., *Learning object tendency: A new Concept for adaptive learning improvement*, In Proceedings of 3rd International Conference of Virtual Learning (ICVL 2008), Constanta, Romania, **2008**, pp. 237-248.

[15] Collins J., Greer J., Huang S., *Adaptive assessment using granularity hierarchies and Bayesian nets, In: Frasson C., Gauthier G., Lesgold A.(eds) Intelligent tutoring systems*, Third International Conference, ITS'96 ,June, Montreal, Canada, **1996**, pp. 569-577.

[16] Molenda M., *Historical and philosophical foundations of instructional design: A north american view, In: R. D. Tennyson, F. Schott, N. Seel and S. Dijkstra, (Eds.)*, Instructional design: International Perspective: Theory, research, and models. Vol. 1, pp. 41-54. Mahwah, NJ: Erlbaum.

[17] Salaberry M.R., *Pedagogical design of computer mediated communication tasks: Learning objectives and technological capabilities*, Modern Language Journal, Vol. 84, part 1, **2000**, pp. 28-37.

[18] Boyle T., *Design principles for authoring dynamic, reusable learning objects*, Australian Journal of Educational Technology, Vol. 19, part 1, **2003**, pp. 46-58.

[19] Gagne R.M., Wager W.W., Golas K.C., Keller J.M., *Principles of instructional design*, 5th edition, Wadsworth Publishing, **2004**.

[20] Mitchell T.M., *Machine learning*, McGraw-Hill, **1997,** pp. 154-184.

[21] Birenbaum M., *Toward adaptive assessment - the student's angle*, Studies in Educational Evaluation, Vol. 20, No.2, **1994**, pp. 239-255.

[22] Dowling C.E., Hockemeyer C., Ludwig A.H., *Adaptive assessment and training using the neighborhood of knowledge states*, Lecture Notes in Computer Science, Issue 1086, **1996**, pp. 578-586.

[23] Challis D., *Committing to quality learning through adaptive online assessment*, Assessment and Evaluation in Higher Education, Vol. 30,No. 5, **2005**, pp. 519-528

[24] Hatzilygeroudis I., Koutsojannis C., Papavlasopoulos C., Prentzas J., *Knowledge-based adaptive assessment in a web-based intelligent educational system*, In Advanced Learning Technologies 2006, Sixth International Conference on, IEEE, **2006**, pp. 651 – 655.

[25] Sitthisak O., Gilbert L. and Davis H.C., *Towards a competency model for adaptive assessment to support lifelong learning*, In: TENCompetence Workshop on Service Oriented Approaches and Lifelong Competence Development Infrastructures, January, Manchester, UK, **2007**.