Archive of SID

«نشریه علمی آموزش و ارزشیابی»
سال دوازدهم ـ شماره ۴۵ ـ بهار ۱۳۹۸
ص. ص. ۱۶۶-۱۵۱

تشخیص خطاهای رویه‌ای دانش‌آموزان در جمع و تفریق اعداد اعشاری با استفاده از شبکه‌های بیزی[۱]

هاشم هوشیار[۲]، علی دلاور[۳]،
نورعلی فرخی[۴]، جلیل یونسی[۵]

تاریخ دریافت مقاله: ۱۳۹۷/۱۰/۱۲
تاریخ پذیرش نهایی مقاله: ۱۳۹۸/۰۴/۱۲

چکیده

اعداد اعشاری یکی از موضوعات مهم و پرکاربرد در آموزش ریاضی است. با این وجود، دانش‌اموزان در کار با اعداد اعشاری با سختی‌هایی مواجه هستند. یک روش جهت بررسی علت سختی کار با اعداد اعشاری برای دانش‌آموزان، تحلیل خطاهای آنها است؛ اما تشخیص خطاها در مهارت‌های رویه‌ای به خاطر ذات بی‌ثباتشان کار دشواری است. این مطالعه، مسئله تشخیص خطاهای رویه‌ای دانش‌آموزان را در زمینه جمع و تفریق اعداد اعشاری، توسط پیشنهاد و ارزیابی یک رویکرد مبتنی بر احتمال با استفاده از شبکه‌های بیزی مورد بررسی قرار می‌دهد. این رویکرد، یک شبکه علّی بین خطاهای رویه‌ای در جمع و تفریق اعداد اعشاری با سؤالات آزمون در نظر می‌گیرد. این پژوهش از نظر هدف، کاربردی و از نظر نحوه گردآوری اطلاعات، کمی می‌باشد. جامعه آماری، کلیه دانش‌آموزان پایه ششم شهر بیرجند در سال تحصیلی ۹۷-۹۶ بودند که از بین آنها تعداد ۴۰۷نفر با استفاده از روش نمونه‌گیری خوشه‌ای چندمرحله‌ای به عنوان نمونه انتخاب شدند. عملکرد شبکه با استفاده از دو نوع از موقعیت‌های آزمون، یکی استفاده از داده‌های دوتایی (نمره‌گذاری صحیح – غلط) و دیگری استفاده تشخیصی از پاسخ‌های غلط توسط یک آزمون چندگزینه‌ای مورد ارزیابی قرار گرفت. نتایج نشان داد شبکه بیزی با استفاده از داده‌های دوتایی عملکرد ضعیفی در تشخیص خطاها داشت اما استفاده تشخیصی از پاسخ‌های غلط دانش‌آموزان، عملکرد شبکه را به طور قابل ملاحظه‌ای بهبود بخشید که ضریب توافق کاپا برای آن در جمع و تفریق اعداد اعشاری به بالای ۹۰ درصد رسید. این نتایج پیشنهاد می‌کند تشخیص قابل اعتماد خطاهای رویه‌ای در جمع و تفریق اعداد اعشاری می‌تواند با استفاده از چارچوب شبکه‌های بیزی با استفاده از پاسخ‌های غلط دانش‌آموزان به دست آید.

واژگان کلیدی: شبکه‌های بیزی، خطاهای رویه‌ای، جمع اعداد اعشاری، تفریق اعداد اعشاری

۱- این مقاله برگرفته از رساله دکتری می‌باشد.

۲- دانشجوی دکتری سنجش و اندازه گیری، دانشگاه علامه طباطبائی.
۳- استاد گروه سنجش و اندازه گیری، دانشگاه علامه طباطبائی (نویسنده مسئول: delavar@atu.ac.ir).
۴- دانشیار گروه سنجش و اندازه گیری، دانشگاه علامه طباطبائی.
۵- استادیار گروه سنجش و اندازه گیری، دانشگاه علامه طباطبائی.

www.SID.ir

Archive of SID

Diagnostic of Students' procedural errors in Addition and subtraction of decimals Using Bayesian Networks

Hashem Hooshyar, Ali Delavar
Noorali Farrokhi, Jalil Younesi

Data of recript: 2019.01.02
Data of acceptance: 2019.07.03

Abstract

Decimal numbers are one of the most important and useful topics in mathematical education. However, students have trouble with decimal numbers. One way to investigate the reason of difficulty in working decimals for the students is through error analysis. Diagnosis of errors in procedural skills is difficult because of their unstable nature. This study investigates the diagnosis of procedural errors in students' decimals addition and subtraction performance, by proposing and evaluating a probability-based approach using Bayesian networks. This approach assumes a causal network relating hypothesized decimals addition and subtraction errors to the observed test items. This study is practical and quantitative. The population of this study included all the sixth-grade students of Birjand city in academic year 2017 – 2018. A sample consisted of 407 students was selected through multi-stage sampling method. Network performance was evaluated by two types of testing situations: using binary data (scored as correct or incorrect) and diagnostic use of wrong answers by a multi-choice test. Results showed Bayesian network with binary data had poor performance in errors diagnosis but diagnostic use of students' wrong answers improved network performance. The Kappa agreement rate for the Bayesian network with wrong answers reached above 90%. Our results suggest that reliable diagnosis of errors can be achieved by using a Bayesian network framework with students' wrong answers.

Keywords: Bayesian networks, procedural errors, addition of decimals, subtraction of decimals

www.SID.ir

Archive of SID

مقدمه

اعداد اعشاری همواره یکی از موضوعات مهم و پرکاربرد در آموزش ریاضی است. با این وجود، پژوهش‌ها نشان می‌دهند که یادگیری اعداد اعشاری و مفاهیم مرتبط با آن، برای دانش‌آموزان بسیار دشوار است و اغلب آنها با انواع بدفهمی‌هایی روبرو هستند که تا سنین بزرگسالی باقی می‌مانند (ایزوتانی، مک لارن و آلتمن[۱]، ۲۰۱۰). نتایج مطالعات بین‌المللی ریاضیات و علوم (تیمز) در سال‌های اخیر، بیانگر خطاها و بدفهمی‌های دانش‌آموزان در زمینه اعداد و عملیات ریاضی و کار با اعداد اعشاری بوده‌اند (استینل[۲]، ۲۰۰۴ به نقل از بخشعلی‌زاده، ۱۳۹۲). نتایج مطالعات بر روی دانش‌آموزان ایرانی بر اساس آزمون‌های تیمز سال‌های ۲۰۰۳، ۲۰۰۷ و ۲۰۱۱ نشان می‌دهد که بیشترین مشکل دانش‌آموزان در بین سؤالات آزمون، مربوط به بدفهمی درک مفهوم عدد اعشاری و مقایسه اعشاری‌ها می‌باشد. همچنین در خصوص عملیات روی اعداد اعشاری نتایج نشان می‌دهد شیوع بدفهمی در جمع و تفریق اعداد اعشاری در بین دانش‌آموزان ایرانی، دو برابر بیشتر از دانش‌آموزان خارجی می‌باشد (بخشعلی‌زاده، ۱۳۹۲).

یک راه برای بررسی این که چرا کار با اعداد اعشاری برای دانش‌آموزان سخت است، تحلیل خطاهای دانش‌آموزان می‌باشد. یک فرض معمول در بررسی خطاهای دانش‌آموزان این است که هر فرد الگوهای خطای باثبات و منظمی دارد؛ زیرا افراد رویه‌های اشتباهی را که آموخته‌اند یا ابداع کرده‌اند را دنبال می‌کنند (بروان و بورتن[۳]، ۱۹۷۸؛ کاکس[۴]، ۱۹۷۵). در عمل تشخیص خطاها در مهارت‌های رویه‌ای[۵] یک فرد دشوار است، زیرا همواره اشتباهات محاسباتی در اجرای یک الگوریتم اشتباه نیز همانند اجرای یک الگوریتم درست رخ می‌دهند. به همین دلیل، این خطاها بی‌ثباتند، یعنی اینکه نشانه یک خطای مشاهده شده در یک فرد ممکن است در سؤالات بعدی یا در جلسه بعدی آزمون دیده نشود. (لی و کورتر[۶]، ۲۰۱۱) در نتیجه بعضی از محققین هشدار داده‌اند که تشخیص قابل اعتماد این خطاها بر اساس عملکرد دانش‌آموزان نباید مورد انتظار باشد (بروان و بورتن، ۱۹۷۸؛ بروان و ون لن[۷]، ۱۹۸۰).

یکی از مورد استفاده‌ترین ابزارها برای کنترل و سنجش عدم قطعیت و بی‌ثباتی یک پیشامد، با کاربردهای موفق در هوش مصنوعی، علوم کامپیوتر، علوم تصمیم و مهندسی، شبکه‌های بیزی[۸] می‌باشد. نظریه زیربنایی شبکه‌های استنباط بیزی به وسیله پیرل[۹] (۱۹۸۸) و لوریزن، اسپیگل‌هالتر[۱۰] (۱۹۸۸) توسعه یافت. شبکه‌های بیزی به طور موفقی برای ارزیابی خرده‌مهارت‌ها[۱۱] و شایستگی‌ها در زمینه‌های

1. Isotani, McLaren, and Altman
2. Steinle
3. Brown and Burton
4. Cox
5. Procedural skills
6. Lee and Corter
7. Brown and VanLehn
8. Bayesian Networks
9. Pearl
10. Lauritzen and Spiegelhalter
11. Subskills

www.SID.ir

Archive of SID

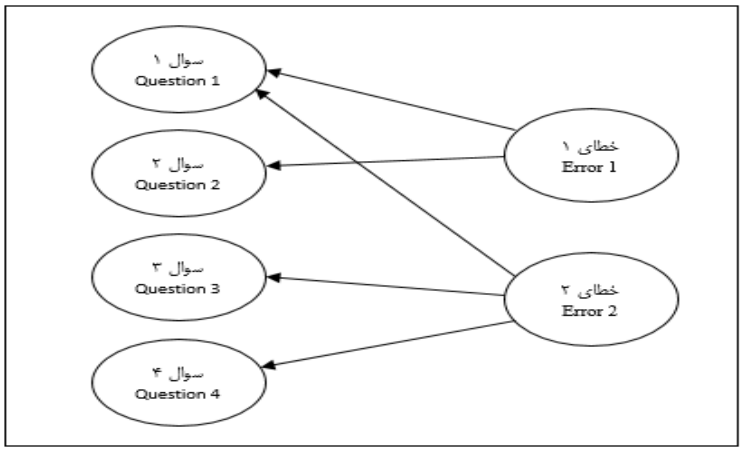
مختلف شامل حل مسئله فیزیک (ون لن و مارتین[۱]، ۱۹۹۸)، تفریق اعداد مخلوط (میسلوی[۲]، ۱۹۹۵)، دانش انگلیسی (آلموند و میسلوی[۳]، ۱۹۹۹) و معلومات دانش‌آموزان در سیستم آموزش مبتنی بر کامپیوتر (کوناتی، گرتنر و ون لن[۴]، ۲۰۰۲؛ نودلمن، مودلی و برمن[۵]، ۲۰۱۹) به کار گرفته شده‌اند. در سال‌های اخیر، علاقه به کاربرد شبکه‌های بیزی در سنجش‌های آموزشی به نظر می‌رسد که در حال رشد می‌باشد (به عنوان مثال، آلموند و همکاران، ۲۰۱۵؛ کاسر[۶] و همکاران، ۲۰۱۷).

شبکه بیزی یک گراف بدون دور جهت‌دار[۷] (DAG) با مجموعه گره‌های[۸] E و مجموعه یال‌های[۹] V که به صورت G=(E,V) نمایش داده می‌شود. گره‌ها در شبکه بیزی نشان‌دهنده متغیرهای تصادفی گسسته یا پیوسته (یا ترکیب هر دو) در حوزه موردنظر و یال‌های جهت‌دار بین زوج متغیرها، بیانگر وابستگی‌های مستقیم آنها هستند. میزان رابطه بین متغیرها با تخصیص توزیع‌های احتمال شرطی به هر گره مشخص می‌شود. اولین مرحله در ساخت یک شبکه بیزی، طراحی بخش گرافیکی آن است. پس از ایجاد ساختار شبکه، نوبت به برآورد پارامترهای شبکه بیزی می‌رسد. این پارامترهای وزنی باید به منظور ایجاد یک شبکه بیزی، از قبل تعیین شوند. وزن‌ها ممکن است به چند روش تعیین شوند: با استفاده از فراوانی‌های مشاهده شده احتمالات حاشیه‌ای و شرطی در یک مجموعه داده، یا با استفاده از یک تئوری، یا استفاده از برآوردهای ذهنی کارشناسان. وقتی که ساختار یک شبکه بیزی تعیین می‌شود و مقادیر پارامترهای وزنی روابط علّی (احتمالات شرطی) فراهم می‌شوند، شبکه می‌تواند برای استنباط یا تشخیص مورد استفاده قرار گیرد. وقتی که یک شبکه بیزی برای تشخیص مورد استفاده قرار می‌گیرد، شواهد به وسیله تعیین مقادیر گره‌های مشخص (در اینجا متناظر با سؤالات آزمون) وارد و در نتیجه، احتمالات پسین بر روی یک مجموعه از گره‌های مورد بررسی (در اینجا متناظر با خطاها) برآورد می‌شوند. در محاسبه احتمالات پسین روی یک متغیر، شبکه یک فرآیند را که انتشار شبکه[۱۰] نامیده می‌شود را به انجام می‌رساند. یک مثال از شبکه بیزی بر اساس مسئله تشخیص خطاهای رویه‌ای در تصویر ۱ نشان داده شده است.

1. VanLehn and Martin
2. Mislevy
3. Almond and Mislevy
4. Conati, Gertner and VanLehn
5. Nudelman, Moodley and Berman
6. Käser
7. Directed Acyclic Graph
8. Nodes
9. Arcs
10. Network propagation

www.SID.ir

Archive of SID



تصویر ۱. یک مثال از شبکه بیزی برای تشخیص خطاها

Figure 1

An example of a Bayesian network for errors diagnosis

چارچوب شبکه‌های بیزی برای مدل‌سازی روابط بین خطاها و سؤالات آزمون به چند دلیل مناسب به نظر می‌رسد. اول این که، روابط علّی بین خطاها و عملکرد در سؤالات آزمون می‌توانند مدل و برآورد شوند. دوم، بی‌ثباتی خطاها می‌توانند به شیوه‌ای طبیعی به وسیله چارچوب احتمالاتی یک شبکه بیزی بیان شود. سوم، درجه‌های مختلفی از قدرت در روابط بین خطاهای مشخص و سؤالات می‌توانند به وسیله احتمالات شرطی مربوط بیان شوند. بالاخره، اطلاعات پیشین درباره فراوانی خطاهای مختلف می‌توانند از طریق احتمالات پیشین وقوع خطا در یک گروه مرجع مورد استفاده قرار بگیرد. (لی و کورتر، ۲۰۱۱)

برای تشخیص خطاهای رویه‌ای دانش‌آموزان از اطلاعات مربوط به عملکرد آزمودنی در یک آزمون مشخص استفاده می‌شود. واضح است که فرمت سؤال در کارایی و موفقیت سنجش‌های تشخیصی که برای هر دانش‌آموز اجرا می‌شود، نقش مؤثری دارد. پژوهش‌هایی همچون ون لن (۱۹۸۱) از سؤالات بازـپاسخ برای به دست آوردن اطلاعات تشخیصی استفاده کردند. از طرف دیگر، آزمون‌های سرنوشت‌ساز در مقیاس بزرگ تمایل دارند برای سهولت در مدیریت آزمون از جمله نمره‌گذاری آزمون، از سؤالات چندگزینه‌ای استفاده کنند. یک راه برای تحقق هر دو هدف، طراحی سؤالات چندگزینه‌ای است که در آن پاسخ‌های غلط برای تشخیص خطاهای دانش‌آموز تدوین شده باشند. استفاده از پاسخ‌های غلط به عنوان شواهد در آزمون‌های تشخیصی در مطالعات مختلف امتحان شده‌اند (دی لا تور[1]، ۲۰۰۹؛ لی و کورتر، ۲۰۱۱).

1. De La Torre

www.SID.ir

Archive of SID

با توجه به توضیحات فوق، هدف این مطالعه بررسی ظرفیت شبکه‌های بیزی برای تشخیص خطاهای رویه‌ای دانش‌آموزان در جمع و تفریق اعداد اعشاری است. در این راستا دو سناریوی معمول در آزمون‌های تشخیصی مورد بررسی قرار گرفت: تشخیص بر اساس اطلاعات دوتایی آزمون (نمره‌گذاری صحیح و غلط سؤالات) و تشخیص بر اساس پاسخ‌های غلط دانش‌آموزان به سؤالات آزمون. لذا سؤال تحقیق در این مطالعه عبارت است از این که آیا خطاهای رویه‌ای دانش‌آموزان در زمینه جمع و تفریق اعداد اعشاری می‌تواند با استفاده از شبکه‌های بیزی تشخیص داده شود و کدام یک از دو شبکه حاصل، عملکرد بهتری دارد؟

روش کار

این پژوهش از نظر هدف، کاربردی و از نظر نحوه گردآوری اطلاعات کمی می‌باشد که در آن از شبکه‌های بیزی به منظور تشخیص خطاهای رویه‌ای دانش‌آموزان پایه ششم در زمینه جمع و تفریق اعداد اعشاری استفاده شد. جامعه آماری، کلیه دانش‌آموزان پایه ششم شهر بیرجند در سال تحصیلی ۹۷-۹۶ بودند که از بین آنها تعداد ۴۰۷ نفر با استفاده از روش نمونه‌گیری خوشه‌ای چندمرحله‌ای از بین مدارس دخترانه و پسرانه به عنوان نمونه انتخاب شدند. با بررسی مقالات، تعداد پنج خطای رویه‌ای معمول برای جمع اعداد اعشاری و تعداد هفت خطای رویه‌ای معمول برای تفریق اعداد اعشاری به شرح زیر شناسایی شد:

خطاهای رویه‌ای در جمع اعداد اعشاری

۱- زیر هم قرار ندادن ممیزها (نقاط اعشار) و مرتب نکردن ارقام در سمت چپ یا راست
(به عنوان مثال؛ ۴/۹۷=۳/۴۵+۱۵/۲)

۲- زیر هم قرار دادن ممیزها اما مرتب نکردن ارقام اعشار در سمت راست
(به عنوان مثال؛ ۱۱/۲۹=۳/۴+۸/۲۵)

۳- جمع کردن ارقام بدون توجه به گروه‌بندی مجدد با اضافه کردن یک ستون در جواب
(به عنوان مثال؛ ۷/۱۲۸=۵/۷۴+۲/۵۴)

۴- عدم انتقال رقم دهگان هنگام گروه‌بندی مجدد
(به عنوان مثال؛ ۷/۲۸=۵/۷۴+۲/۵۴)

۵- وارد کردن ممیز در پاسخ به صورت اشتباه
(به عنوان مثال؛ ۱/۱۶۵=۳/۴+۸/۲۵)

www.SID.ir

Archive of SID

خطاهای رویه‌ای در تفریق اعداد اعشاری

۱- زیر هم قرار ندادن ممیزها (نقاط اعشار) و مرتب نکردن ارقام در سمت چپ یا راست
(به عنوان مثال؛ ۱۳/۲=۱/۲۴—۱۳/۶)

۲- زیر هم قرار دادن ممیزها اما مرتب نکردن ارقام اعشار در سمت راست
(به عنوان مثال؛ ۱/۷۱=۳/۴—۳/۷۵)

۳- تفریق عدد کوچکتر از عدد بزرگتر وقتی قرض گرفتن لازم است
(به عنوان مثال؛ ۱۲/۱۹=۱۴/۲۲—۲۶/۴۱)

۴- توقف قرض گرفتن در صفر (قرض گرفتن از صفر در ستون یکان انجام می‌شود اما صفر در ستون دهگان بدون تغییر باقی می‌ماند. تفریق در ستون صدگان مجدداً با کاهش برای ستون دهگان به صورت صحیح انجام می‌گیرد.
(به عنوان مثال؛ ۳/۲۸۷=۲/۳۲۴—۵/۶۰۱)

۵- قرض گرفتن بدون کاهش (به عنوان مثال؛ ۳/۳۸۷=۲/۳۲۴—۵/۶۰۱)

۶- در نظر نگرفتن صفر در سمت راست ارقام اعشار وقتی تعداد ارقام اعشار عدد اول کمتر است.
(به عنوان مثال؛ ۱/۴۲=۶/۰۲—۷/۴)

۷- وارد کردن ممیز در پاسخ به صورت اشتباه
(به عنوان مثال؛ ۳/۵=۳/۴—۳/۷۵)

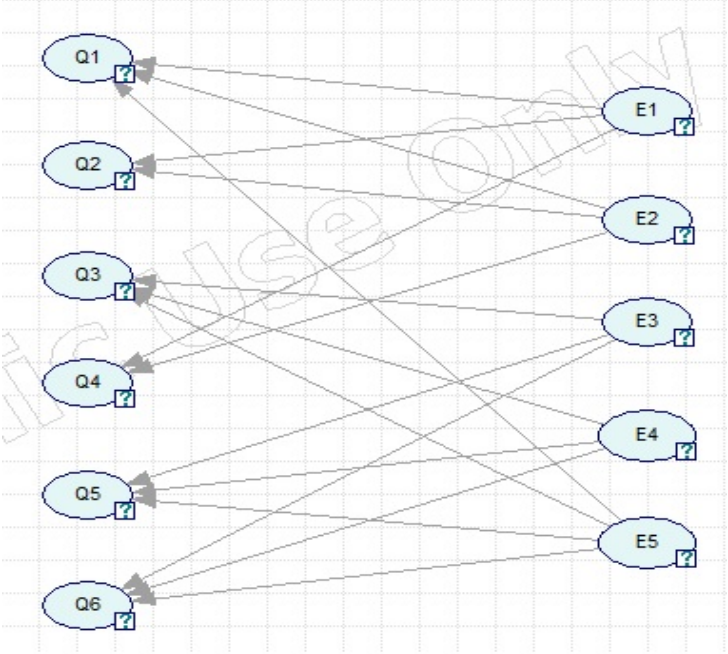
پس از شناسایی و تعیین خطاهای رویه‌ای دانش‌آموزان در زمینه جمع و تفریق اعداد اعشاری، به منظور تعیین حضور یا عدم حضور خطاها در هر یک از دانش‌آموزان از یک آزمون باز-پاسخ شامل ۸ سؤال در زمینه جمع و ۱۰ سؤال در زمینه تفریق اعداد اعشاری استفاده شد. سؤالات این آزمون به گونه‌ای طراحی شدند که هر یک از خطاهای رویه‌ای شناسایی شده، توسط حداقل پنج سؤال مورد بررسی قرار گرفت. بر اساس مطالعه کاکس (۱۹۷۵) یک دانش‌آموز به عنوان دارنده یک خطای رویه‌ای تعریف شد که حداقل سه تا از پاسخ‌های اشتباهش به سؤالات، پاسخ تولید شده توسط آن خطا را نشان داد. از این تشخیص‌ها به عنوان متغیرهای ملاک بیرونی استفاده شد. به منظور ساختن و ارزیابی مدل نیز از یک آزمون چهارگزینه‌ای که گزینه‌های غلط آن مربوط به خطاهای دانش‌آموزان بودند، شامل ۶ سؤال مربوط به جمع اعداد اعشاری و ۱۳ سؤال مربوط به تفریق اعداد اعشاری استفاده گردید. سؤالات آزمون‌ها با توجه به اهداف موردنظر با استفاده از کتاب درسی ریاضی، مقالات مرتبط و آزمون‌های معتبر طراحی گردید. جهت روایی محتوایی آزمون‌ها از نظرات اساتید صاحب‌نظر در آموزش ریاضی و معلمان ریاضی استفاده شد.

در راستای اهداف تحقیق، در این مطالعه دو مدل شبکه بیزی بر حسب نوع اطلاعات حاصل از سؤالات آزمون (اطلاعات دوتایی در مقابل اطلاعات پاسخ مشخص) ساخته شدند. پیش‌بینی خطاها در شبکه

www.SID.ir

Archive of SID

پاسخ دوتایی (شبکه ۱) بر اساس نتایج حاصل از روش نمره‌گذاری صحیح–غلط سؤالات آزمون و در شبکه پاسخ–مشخص (شبکه ۲) بر اساس برداری از پاسخ‌های دانش‌آموزان به سؤالات آزمون، ساخته شدند. تصویر ۲ و ۳ به ترتیب شبکه ۱ مربوط به جمع و تفریق اعداد اعشاری را نشان می‌دهد. تنها تفاوت شبکه ۲ برای جمع و تفرق اعداد اعشاری نسبت به شبکه ۱ در این است که هر گره سؤال در شبکه ۱ با چهار گره مربوط به چهار پاسخ محتمل به آن سؤال جایگزین شده است.



تصویر ۲. شبکه ۱ در جمع اعداد اعشاری. در نمودار، E نماد خطا و Q نماد سؤال است.

Figure 2

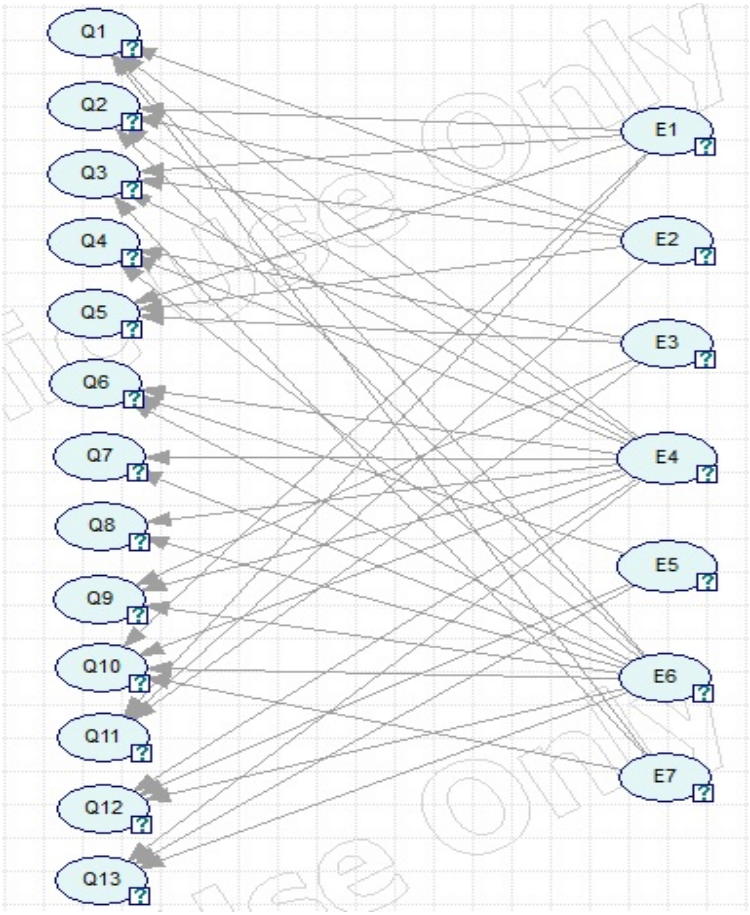
Network 1 for addition of decimals. In the diagram, E is symbol of Error and Q is symbol of Question.

برای ارزیابی مدل، گروه نمونه به صورت تصادفی به دو گروه شامل گروه مدل‌سازی (۲۰۳ نفر) و گروه آزمایش (۲۰۴ نفر) تقسیم شدند. از گروه مدل‌سازی برای برآورد پارامترهای شبکه استفاده شد. در این مرحله مقادیر پارامترها (توزیع‌های احتمال ساده و شرطی) به طور مستقیم با استفاده از داده‌های تجربی شامل خطاهای شناسایی شده برای هر یک از آزمودنی‌ها در گروه مدل‌سازی برآورد شد. بنابراین داده‌های مورد استفاده در مرحله مدل‌سازی شامل بردار داده‌های عملکرد دانش‌آموزان در سؤالات آزمون چهارگزینه‌ای و مقادیر ملاک مشاهده شده دانش‌آموزان (بر اساس آزمون باز–پاسخ) در متغیرهای مربوط به خطاها بودند. پس از برآورد پارامترهای شبکه، احتمالات پسین برای داده‌های گروه آزمایش محاسبه شد. در این مرحله، بردار داده‌های عملکرد هر یک از دانش‌آموزان گروه آزمایش در سؤالات آزمون، وارد مدل شده و احتمالات پسین برای خطاها تولید شدند. برای ارزیابی شبکه، درصد دانش‌آموزانی که به طور

www.SID.ir

Archive of SID

صحیح روی هر خطا طبقه‌بندی شدند به عنوان معیار عملکرد شبکه در نظر گرفته شد. برای پیش‌بینی حضور یا عدم حضور خطاهای معین، از نقاط برش برای احتمالات پسین هر خطا برای هر یک از دانش‌آموزان استفاده شد. نقاط برش در این مطالعه، نقاط برش ثابت شامل ۰/۲۵، ۰/۵، ۰/۷۵ و نقطه برش متغیر شامل نرخ پایه خطا در گروه مدل‌سازی انتخاب شدند. نرخ پایه خطا در گروه مدل‌سازی در جمع اعداد اعشاری برای خطاهای ۱ تا ۵ به ترتیب ۰/۰۳۴، ۰/۰۵۴، ۰/۰۳، ۰/۰۴۹، ۰/۰۲ و در تفریق اعداد اعشاری برای خطاهای ۱ تا ۷ به ترتیب ۰/۰۳۴، ۰/۰۵۴، ۰/۰۲، ۰/۰۴۴، ۰/۰۴۹، ۰/۱۲۸، ۰/۱۰۸ بود. سپس خطاهای پیش‌بینی شده شبکه با خطاهای مشاهده شده مورد مقایسه قرار گرفتند. جهت تحلیل داده‌ها از نرم افزارهای GeNIe و SPSS22 استفاده گردید و سطح معنی‌داری P<۰/۰۵ در نظر گرفته شد.



تصویر ۳. شبکه۱ در تفریق اعداد اعشاری. در تصویر، E نماد خطا و Q نماد سؤال است.

Figure 3

Network 1 for subtraction of decimals. In the diagram, E is symbol of Error and Q is symbol of Question.

www.SID.ir

Archive of SID

نتایج

درصد خطای واقعی و درصد خطای پیش‌بینی شده به وسیله دو شبکه در گروه آزمایش برای خطاهای ۱ تا ۵ در جمع اعداد اعشاری و خطاهای ۱ تا ۷ در تفریق اعداد اعشاری در جدول ۱ آمده است. نتایج نشان می‌دهد درصد خطای مشاهده دانش‌آموزان در جمع و تفریق اعداد اعشاری به طور معقولی توسط شبکه‌ای که از پاسخ‌های اشتباه استفاده کرده است (شبکه ۲) پیش‌بینی شده است. یعنی این که درصد خطاهای پیش‌بینی شده توسط شبکه ۲ هم در جمع و هم در تفریق اعداد اعشاری نزدیک‌ترین برازش را به درصد خطاهای واقعی مشاهده شده نشان داد. در حقیقت شبکه‌ای که از پاسخ‌های اشتباه دانش‌آموزان به عنوان داده‌های ورودی استفاده کرده صرف نظر از نقاط برش به کار رفته، بهتر از شبکه با اطلاعات دوتایی سؤال عمل کرده است. این نتایج به این نتیجه منجر می‌شود که تشخیص خطا وقتی پاسخ‌های اشتباه دانش‌آموز به عنوان شواهد مورد استفاده قرار می‌گیرند نتیجه‌بخش‌تر است. نتایج نشان می‌دهد شبکه ۱ در پیش‌بینی درصد خطای دانش‌آموزان در جمع و تفریق اعداد اعشاری عملکرد ضعیفی داشته است. عملکرد نسبی شبکه ۲ در بین نقاط برش مختلف یکسان می‌باشد.

درصدهای طبقه‌بندی هر شبکه با استفاده از نقاط برش مختلف در جدول ۲ گزارش شده است. درصدهای طبقه‌بندی ارائه شده، درصد توافق بین خطاهای مشاهده شده و خطاهای پیش‌بینی شده می‌باشند. درصدهای طبقه‌بندی ارائه شده در جدول ۲ قابل توجه به نظر می‌رسند. اما همانطور که مشخص است برای یک پدیده نادر، درصد توافق به دلیل تعداد زیاد توافق‌های منفی می‌تواند بزرگ باشد. بنابراین علاوه بر درصد طبقه‌بندی خام، ضریب توافق کاپای کوهن نیز مورد محاسبه قرار گرفت. مقادیر ضریب کاپا برای پیش‌بینی‌های مدل در جدول ۲ در داخل پرانتز نشان داده شده‌اند.

همانطور که در جدول ۲ مشاهده می‌شود میانگین درصدهای طبقه‌بندی تولید شده توسط شبکه ۱ و ۲ در بین خطاها در جمع و تفریق اعداد اعشاری بالاتر از ۹۳ درصد هستند. اما میانگین مقادیر کاپا برای شبکه۱ در جمع و تفریق اعداد اعشاری پایین بوده است. بهبود نرخ پیش‌بینی شبکه۲ در مقایسه با شبکه۱ از نظر مقادیر کاپا قابل ملاحظه است. میانگین مقادیر کاپا در بین خطاها در جمع و تفریق اعداد اعشاری بیشتر از ۸۵ درصد بود. نتایج گزارش شده در جدول ۲ نشان می‌دهد که عملکرد نسبی دو شبکه در پیش‌بینی خطاها در جمع و تفریق اعداد اعشاری بر حسب ضرایب توافق کاپا به طور قابل ملاحظه‌ای متفاوت است ($P<۰/۰۵$ در نقاط برش مختلف). شبکه ۲ عملکرد به مراتب بهتری از شبکه ۱ در پیش‌بینی خطاهای رویه‌ای دانش‌آموزان در جمع و تفریق اعداد اعشاری داشته است. استفاده از پاسخ‌های اشتباه به عنوان اطلاعات تشخیصی برای پیش‌بینی خطاهای رویه‌ای باعث بهبود قابل توجهی در عملکرد طبقه‌بندی شبکه برای کلیه خطاها شده است.

www.SID.ir

Archive of SID

جدول ۱. درصد پایه خطای مشاهده شده و درصد خطای پیش‌بینی شده در نمونه آزمایشی به وسیله دو مدل شبکه با نقاط برش مختلف

Table 1

The Observed Error Base percentages and Predicted Error percentages in the testing Sample by the Two Network Models with Various Cutoff Points

خطا Error	درصد خطای مشاهده شده Observed base percentage of error	شبکه ۱ Network 1				شبکه ۲ Network 2			
		نقطه برش Cutoff point				نقطه برش Cutoff point			
		نرخ پایه base rate	0.25	0.5	0.75	نرخ پایه base rate	0.25	0.5	0.75
جمع اعداد اعشاری	Addition of Decimal Numbers								
خطای ۱ Error 1	2	12.3	11.8	0	0	2	2	2	2
خطای ۲ Error 2	11.3	12.3	12.3	11.8	0	9.8	9.8	9.8	9.8
خطای ۳ Error 3	2.9	6.9	6.9	0	0	2.9	2.9	2.9	2.9
خطای ۴ Error 4	4.4	6.9	6.9	5.4	0	3.9	3.4	3.4	3.4
خطای ۵ Error 5	2.5	10.8	6.4	0	0	5.4	2.9	2.9	2.5
تفریق اعداد اعشاری	Subtraction of Decimal Numbers								
خطای ۱ Error 1	2	0	0	0	0	2.9	2.5	2.5	2
خطای ۲ Error 2	11.3	1	1	1	1	11.3	11.3	10.8	10.8
خطای ۳ Error 3	12.7	2.9	1.5	1.5	1.5	14.7	14.7	13.2	13.2
خطای ۴ Error 4	5.9	2	1.5	1.5	0	5.9	5.9	5.9	5.9
خطای ۵ Error 5	7.8	0.5	0.5	0.5	0.5	12.3	9.3	9.3	9.3
خطای ۶ Error 6	9.8	1.5	1.5	0	0	9.3	9.3	9.3	9.3
خطای ۷ Error 7	2.5	0.5	0.5	0.5	0.5	9.8	6.4	4.9	2.9

www.SID.ir

Archive of SID

جدول ۲. درصد طبقه‌بندی صحیح در نمونه آزمایشی برای دو مدل شبکه با نقاط برش مختلف

Table 2

Correct Classification percentages in the testing Sample for the Two Network Models with Various Cutoff Points

خطا Error	نرخ پایه base rate		0.25		0.50		0.75	
	شبکه ۱ Network 1	شبکه ۲ Network 2	شبکه ۱ Network 1	شبکه ۲ Network 2	شبکه ۱ Network 1	شبکه ۲ Network 2	شبکه ۱ Network 1	شبکه ۲ Network 2
جمع اعداد اعشاری Addition of Decimal Numbers								
خطای ۱ Error 1	89.7 (0.25)	100 (1)	90.2 (0.26)	100 (1)	98 (0)	100 (1)	98 (0)	100 (1)
خطای ۲ Error 2	96.1 (0.81)	98.5 (0.92)	96.1 (0.81)	98.5 (0.92)	96.6 (0.83)	98.5 (0.92)	88.7 (0)	98.5 (0.92)
خطای ۳ Error 3	96.1 (0.58)	100 (1)	96.1 (0.58)	100 (1)	97.1 (0)	100 (1)	97.1 (0)	100 (1)
خطای ۴ Error 4	95.6 (0.59)	98.5 (0.82)	95.6 (0.59)	99 (0.87)	95.1 (0.47)	99 (0.87)	95.6 (0)	99 (0.87)
خطای ۵ Error 5	91.7 (0.34)	97.1 (0.61)	96.1 (0.54)	98.5 (0.72)	97.5 (0.28)	98.5 (0.72)	97.5 (0)	99 (0.80)
میانگین Mean	93.8 (0.51)	98.8 (0.87)	94.8 (0.56)	99.2 (0.90)	96.9 (0.32)	99.2 (0.90)	95.4 (0)	99.3 (0.92)
سطح معنی‌داری Sig.	0.016		0.016		0.016		0.008	
تفریق اعداد اعشاری Subtraction of Decimal Numbers								
خطای ۱ Error 1	98 (0)	99 (0.80)	98 (0)	99.5 (0.89)	98 (0)	99.5 (0.89)	98 (0)	100 (1)
خطای ۲ Error 2	89.7 (0.15)	100 (1)	89.7 (0.15)	100 (1)	89.7 (0.15)	99.5 (0.98)	89.7 (0.15)	99.5 (0.98)
خطای ۳ Error 3	88.7 (0.15)	99.5 (0.98)	88.7 (0.19)	98 (0.92)	88.7 (0.19)	99.5 (0.98)	88.7 (0.19)	98 (0.92)
خطای ۴ Error 4	94.1 (0.23)	100 (1)	94.6 (0.25)	100 (1)	94.6 (0.25)	100 (1)	94.1 (0)	100 (1)
خطای ۵ Error 5	92.6 (0.11)	95.6 (0.76)	92.6 (0.11)	98.5 (0.91)	92.6 (0.11)	98.5 (0.91)	92.6 (0.11)	98.5 (0.91)
خطای ۶ Error 6	89.7 (0.06)	99.5 (0.97)	89.7 (0.06)	99.5 (0.97)	90.2 (0)	99.5 (0.97)	90.2 (0)	99.5 (0.97)
خطای ۷ Error 7	98 (0.33)	92.6 (0.38)	98 (0.33)	96.1 (0.54)	98 (0.33)	96.6 (0.52)	98 (0.33)	98.5 (0.72)
میانگین Mean	93 (0.15)	98 (0.84)	93 (0.16)	98.8 (0.89)	93.1 (0.15)	99 (0.89)	93 (0.11)	99.1 (0.93)
سطح معنی‌داری Sig.	0.001		0.001		0.001		0.001	

www.SID.ir

Archive of SID

توجه: برای هر خانه جدول، عدد اول درصد طبقه‌بندی صحیح دانش‌آموزان در هر خطا و عدد داخل پرانتز مقدار کاپای کوهن را نشان می‌دهد. مقادیر سطح معنی‌داری مربوط به آزمون من‌ویتنی برای مقایسه دو شبکه بر حسب ضریب توافق کاپا می‌باشد.

به منظور تعیین قدرت شبکه‌ها در پیش‌بینی همزمان خطاها، الگوهای خطای مشاهده شده با الگوهای خطای پیش‌بینی شده توسط شبکه برای هر دانش‌آموز مقایسه شدند. مقادیر درصد طبقه‌بندی صحیح و ضریب توافق کاپا برای مقادیر برش مختلف در جدول ۳ آمده است. نتایج نشان می‌دهد مشابه با حالت پیش‌بینی تکی خطاها، عملکرد شبکه ۲ بهتر از شبکه ۱ بود.

جدول ۳. درصد طبقه‌بندی صحیح دانش آموزان بر اساس الگوی خطا در نمونه آزمایشی برای دو مدل شبکه با نقاط برش مختلف

Table 3

Correct Classification percentages in the testing Sample for the Two Network Models with Various Cutoff Points

شبکه Network	نرخ پایه base rate	نقطه برش Cutoff point 0.25	0.5	0.75
جمع اعداد اعشاری	Addition of Decimal Numbers			
شبکه ۱ Network 1	78.9 (0.42)	78.9 (0.42)	90.2 (0.69)	80.9 (0)
شبکه ۲ Network 2	95.1 (0.86)	96.6 (0.90)	96.6 (0.90)	96.6 (0.90)
تفریق اعداد اعشاری	Subtraction of Decimal Numbers			
شبکه ۱ Network 1	62.3 (0.10)	63.2 (0.12)	62.3 (0.08)	61.8 (0.08)
شبکه ۲ Network 2	94.1 (0.91)	94.6 (0.92)	96.1 (0.94)	88.2 (0.82)

توجه: برای هر خانه جدول، عدد اول درصد طبقه‌بندی صحیح دانش‌آموزان و عدد داخل پرانتز مقدار کاپای کوهن را نشان می‌دهد.

بحث و نتیجه‌گیری

بر اساس نتایج حاصل از این مطالعه، عملکرد شبکه بیزی با استفاده از پاسخ‌های دوتایی صحیح–غلط در پیش‌بینی خطاهای رویه‌ای دانش‌آموزان در زمینه جمع اعداد اعشاری نسبتاً ضعیف و در زمینه تفریق اعداد اعشاری بسیار ضعیف بود. این در حالی است که لی و کورتر (۲۰۱۱) در مجموع عملکرد شبکه بیزی با پاسخ‌های دوتایی را در تشخیص خطاهای رویه‌ای دانش‌آموزان در تفریق اعداد طبیعی را نسبتاً مطلوب ارزیابی کردند. این نتیجه ممکن است به دلیل شیوع کم خطاهای رویه‌ای در نمونه مورد مطالعه باشد همچنان که در مطالعه لی و کورتر عملکرد شبکه بیزی با پاسخ‌های دوتایی در پیش‌بینی یکی از خطاها که از شیوع پایینی برخوردار بود مشابه عملکرد شبکه بیزی با پاسخ‌های دوتایی در پیش‌بینی خطاهای مطالعه حاضر بود.

www.SID.ir

Archive of SID

نتایج نشان داد که شبکه‌های بیزی با استفاده از پاسخ‌های غلط می‌تواند برای تشخیص خطاهای رویه‌ای هم در زمینه جمع و هم در زمینه تفریق اعداد اعشاری ولو اینکه این خطاها بی‌ثباتند به کار برده شوند. بر اساس نتایج، دقت تشخیص خطاها در مهارت‌های رویه‌ای زمانی که پاسخ‌های غلط دانش‌آموزان به عنوان شواهد مورد استفاده قرار گرفتند بهبود چشمگیری را نشان داد. این نتایج با مطالعه لی و کورتر (۲۰۱۱) هم‌خوانی دارد. اطلاعات مربوط به پاسخ‌های غلط معمولاً با سؤالات آزمون باز–پاسخ به دست می‌آیند. اما این مهم با آزمون‌های چندگزینه‌ای ساخته شده برای این هدف نیز می‌تواند فراهم شود. طراحی سؤالات آزمون چندگزینه‌ای با پاسخ‌های غلط انتخابی که مشخصه بدفهمی یا خطاهای رویه‌ای هستند یک ابزار مفید در توسعه آزمون‌های بزرگ مقیاس خواهد بود که می‌تواند برای تشخیص بدفهمی‌ها و خطاها در مهارت‌های رویه‌ای استفاده شوند. از طرفی سنجش تشخیصی الگوهای خطای دانش‌آموزان این امکان را برای معلمان فراهم می‌آورد که به هر دانش‌آموز آموزش هدفمند ارائه داده و مساوات را میان یک جمعیت ناهمگن از دانش آموزان ترویج دهند (واتسون[1] و همکاران، ۲۰۱۸). بنابراین روش‌شناسی پیشنهاد شده در این پژوهش می‌تواند به طراحی آموزش‌های فردی اثربخش کمک کند.

این پژوهش همانند دیگر پژوهش‌ها، با محدودیت‌هایی روبه‌رو بوده است. یافته‌های این پژوهش تجربی، محدود به تشخیص خطاهای رویه‌ای در زمینه جمع و تفریق اعداد اعشاری در دانش‌آموزان پایه ششم شهرستان بیرجند است؛ بنابراین، پیشنهاد می‌شود در تعمیم نتایج این پژوهش جوانب احتیاط رعایت و تحقیقات مشابهی در زمینه سایر مهارت‌های رویه‌ای چه در زمینه اعداد اعشاری و چه در سایر زمینه‌ها در جوامع مختلف انجام و با نتایج این مطالعه مقایسه گردد. همچنین در پژوهش حاضر، تنها برخی از محتمل‌ترین خطاهای رویه‌ای در جمع و تفریق اعداد اعشاری مورد بررسی قرار گرفت که با توجه به تعداد زیاد این خطاها، بررسی عملکرد شبکه‌های بیزی در تشخیص سایر خطاهای رویه‌ای اشاره شده در متون تحقیق، می‌تواند موضوع تحقیق پژوهش‌های آتی باشد. برخی از محققان (تمپلین[2] و همکاران، ۲۰۰۸؛ لی و کورتر، ۲۰۱۱) اشاره می‌کنند که تشخیص خطاهای رویه‌ای می‌تواند با یک مدل علی که دانش خرده‌مهارت‌ها را به عنوان فاکتورهای علی اثرگذار بر خطاها در نظر می‌گیرد، بهبود یابد. لذا پیشنهاد می‌شود این موضوع در پژوهش‌های مکمل دیگر، با وارد کردن گره‌های مربوط به خرده‌مهارت‌های زیربنایی هر یک از خطاها در شبکه بیزی مورد بررسی قرار گیرد.

1. Watson
2. Templin

www.SID.ir

Archive of SID

منابع References

بخشعلی‌زاده، ش. (۱۳۹۲). *شناسایی بدفهمی‌های رایج دانش‌آموزان پایه چهارم در حوزه محتوایی ریاضی*، سازمان پژوهش و برنامه ریزی، پژوهشگاه مطالعات آموزش و پرورش.

Almond, R. G., & Mislevy, R. J. (1999). Graphical models and computerized adaptive testing. *Applied Psychological Measurement*, 23, 223-237.

Almond, R. G., Mislevy, R. J., Steinberg, S. L., Yan, D., & Williamson, D. M. (2015). Bayesian Networks in Educational Assessment. New York, NY: Springer.

Bakhshalizadeh, Sh., & Broojerdian, N. (2014). *Identifying common misconceptions in fourth-grade students in the field of mathematical content*. Organization for Research and Planning. Institute for Education Studies. [In Persian]

Brown, J. S., & Burton, R. B. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science*, 2, 155-192.

Brown, J. S., & VanLehn, K. (1980). Repair theory: A generative theory of bugs in procedural skills. *Cognitive Science*, 4, 379-426.

Conati, C., Gertner, A., & VanLehn, K. (2002). Using Bayesian networks to manage uncertainty in student modeling. *User Modeling and User-Adapted Interactions*, 12, 371-417.

Cox, L. S. (1975). Systematic errors in the four vertical algorithms in normal and handicapped populations. *Journal for Research in Mathematics Education*, 6, 202-220.

De La Torre, J. (2009). A cognitive diagnosis model for cognitively based multiple-choice options. *Applied Psychological Measurement*, 33, 163-183.

Isotani, S., McLaren, B., & Altman, M. (2010). Towards Intelligent Tutoring with Erroneous Examples: A Taxonomy of Decimal Misconceptions. Proc. of the Int. *Conference on Intelligent Tutoring Systems*. LNCS 6095, 346-348.

Käser, T., Klingler, S., Schwing, A. G., & Gross, M. (2017). Dynamic Bayesian Networks for Student Modeling. *IEEE Transactions on Learning Technologies*, 10(4), 450-462.

Lauritzen, S. L., & Spiegelhalter, D. J. (1988). Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society*, Series B, 50,157-224.

Lee, J., & Corter, J. E. (2011). Diagnosis of subtraction bugs using Bayesian Networks. *Applied Psychological Measurement*, 35(1), 27-47.

Li, X. (2006). *Cognitive analysis of students' errors and misconceptions in variables, equations and functions*. Ph.D. Desertation, Texas A&M University.

Mislevy, R. J. (1995). *Probability-based inference in cognitive diagnosis*. In P. Nichols, S. Chipman, &R. Brennan (Eds.), Cognitively diagnostic assessment (pp. 43-71). Mahwah, NJ: Erlbaum.

www.SID.ir

Archive of SID

Nudelman, Z., Moodley, D., & Berman, S. (2019). *Using Bayesian Networks and Machine Learning to Predict Computer Science Success*. Paper presented at the ICT Education, Cham.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems*. San Mateo, CA: Morgan Kaufmann.

Steinle, V., Stacey, K., & Chambers, D. (2002). *Teaching and Learning about Decimals [CD-ROM]: Department of Science and Mathematics Education*, The University of Melbourne. Online sample http://extranet.edfac.unimelb.edu.au/DSME/decimals.

Templin, J. L., Henson, R. A., Templin, A. E., & Roussos, L. (2008). Robustness of hierarchical modeling of skill association in cognitive diagnosis models. *Applied Psychological Measurement*, 32, 559-574.

VanLehn, K. (1981). *Bugs are not enough: Empirical studies of bugs, impasses and repairs in proceduralskills (Technical Report)*. Palo Alto, CA: Xerox Palo Alto Research Center.

VanLehn, K. (1990). *Minds bugs: The origins of procedural misconceptions*. Cambridge, MA: MIT Press.

VanLehn, K., & Martin, J. (1998). Evaluation of an assessment system based on Bayesian student modeling. *International Journal of Artificial Intelligence in Education*, 8, 179-221.

Watson, S. M. R., Lopes, J., Oliveira, C., & Judge, S. (2018). Error patterns in Portuguese students' addition and subtraction calculation tasks: implications for teaching. *Journal for Multicultural Education*, 12(1), 67-82.

www.SID.ir