Archive of SID

Published by *Vesnu Publications*

DOI: 10.22122/him.v17i3.4139

مجله مدیریت اطلاعات سلامت

سال هفدهم/ شماره سوم/ مرداد و شهریور ۱۳۹۹

به کارگیری الگوریتم خوشه‌بندی C– میانگین فازی در تشخیص میزان اضطراب

فرشته پارساپور[۱]، جاوید پیمانی[۲]، محمد خان‌بابایی[۳]

مقاله پژوهشی

چکیده

مقدمه: یکی از مهم‌ترین گام‌ها در پیشگیری و کنترل اختلالات اضطرابی، تشخیص آن‌ها در مراحل اولیه توسط روان‌پزشک است. هدف از انجام پژوهش حاضر، ارایه روشی جهت تشخیص شدت اضطراب با استفاده از الگوریتم خوشه‌بندی C- میانگین فازی بود. همچنین، تأثیر هر مشخصه در سنجش اضطراب و خوشه‌بندی مراجعه‌کنندگان تعیین گردید.

روش بررسی: این مطالعه از نوع توصیفی بود. مجموعه داده‌های مرتبط به پرونده ۳۰۰ نفر از مراجعه‌کنندگان به سه کلینیک روان‌پزشکی در شهر تهران، بر اساس پرسش‌نامه آزمون BAI (Beck Anxiety Inventory) تهیه شد. سپس الگوریتم خوشه‌بندی C- میانگین فازی، به بخش‌بندی مراجعه‌کنندگان و تعیین میزان اضطراب آن‌ها در هر خوشه پرداخت. این الگوریتم به طور مجزا بر روی هر مشخصه نیز اعمال گردید.

یافته‌ها: مراجعه‌کنندگان به کلینیک روان‌پزشکی به چهار خوشه با برچسب‌های «فاقد اضطراب، اضطراب خفیف، اضطراب متوسط و اضطراب شدید» تقسیم شدند. الگوریتم خوشه‌بندی C- میانگین فازی با دقت ۹۰/۶۶ درصد، به تشخیص اضطراب مراجعه‌کنندگان پرداخت. با اجرای این الگوریتم بر روی هر مشخصه، تأثیر مشخصه‌ها در سنجش اضطراب و خوشه‌بندی مراجعه‌کنندگان نیز تعیین شد.

نتیجه‌گیری: الگوریتم خوشه‌بندی C- میانگین فازی، شدت اضطراب مراجعه‌کنندگان را با دقت بالایی تشخیص می‌دهد. بخش‌بندی بیماران با رویکرد خوشه‌بندی و بر اساس مهم‌ترین مشخصه‌ها، می‌تواند ابزار مفیدی جهت تصمیم‌گیری روان‌پزشک در تشخیص شدت اضطراب در مراحل اولیه آن باشد.

واژه‌های کلیدی: تحلیل خوشه‌ای؛ منطق فازی؛ اضطراب

پیام کلیدی: به کارگیری روش خوشه‌بندی فازی در تعیین سطح اضطراب بیماران و شناسایی مهم‌ترین مشخصه‌های سنجش اضطراب، باعث افزایش دقت در تشخیص صحیح اضطراب و صرفه‌جویی در زمان و هزینه می‌شود.

دریافت مقاله: ۱۳۹۹/۳/۲۹ پذیرش مقاله: ۱۳۹۹/۵/۳ تاریخ انتشار: ۱۳۹۹/۵/۱۵

ارجاع: پارساپور فرشته، پیمانی جاوید، خان‌بابایی محمد. **به کارگیری الگوریتم خوشه‌بندی C- میانگین فازی در تشخیص میزان اضطراب.** مدیریت اطلاعات سلامت ۱۳۹۹؛ ۱۷ (۳): ۱۰۳-۹۷

مقدمه

تشخیص اختلالات روانی به دلیل نشانه‌های ذهنی و درونی آن پیچیده است. با گذر زمان ممکن است این پیچیدگی پیش‌رونده شود و تشخیص اولیه توسط روان‌پزشک را دشوار سازد (۱). تشخیص و درمان اضطراب توسط پزشک اغلب کمتر از آمار واقعی مبتلایان می‌باشد (۲). میزان تشخیص اضطراب در مراقبت‌های اولیه بسیار کم، در اضطراب فراگیر خالص، حدود ۳۴/۴ درصد و برای اضطراب فراگیر همراه با افسردگی، حدود ۴۳ درصد برآورد شده است (۳). بنابراین، افزایش دقت تشخیص اضطراب توسط پزشک در مراحل اولیه ضروری به نظر می‌رسد.

در دو دهه گذشته، تکنیک‌های داده‌کاوی همچون شبکه عصبی مصنوعی ANN (Artificial Neural Network)، درخت تصمیم (Decision Tree)، شبکه Bayesian و ماشین بردار پشتیبان SVM (Support Vector Machine)، کاربرد زیادی در حوزه روان‌شناسی داشته است (۴). داده‌کاوی می‌تواند از طریق کشف الگوهای ارزشمند پنهان در میان انبوه داده‌های مرتبط با پرونده‌های پزشکی بیماران، به شناسایی علل وقوع، تشخیص، پیش‌بینی و درمان بیماری‌ها از جمله اضطراب کمک نماید (۵).

خوشه‌بندی، یکی از تکنیک‌های مشهور داده‌کاوی به شمار می‌رود. فرایند گروه‌بندی مجموعه‌ای از مشاهدات به خوشه‌هایی از مشاهدات مشابه، خوشه‌بندی نام دارد. مطالعات مختلف، از خوشه‌بندی برای بخش‌بندی بیماران استفاده کرده‌اند، اما در روش‌های کلاسیک خوشه‌بندی، هر مشاهده با احتمال برابر با ۱ یا صفر به خوشه‌ای تعلق دارد؛ در حالی که در مواردی ممکن است مقادیر مشاهدات و تعلق آن‌ها به خوشه‌ها از یک عدم قطعیت برخوردار باشد (۶).

مقاله حاصل تحقیق مستقل بدون حمایت مالی و سازمانی است.

۱- کارشناس ارشد، روان‌شناسی بالینی، گروه روان‌شناسی بالینی، دانشکده روان‌شناسی، واحد کرج، دانشگاه آزاد اسلامی، البرز، ایران

۲- استادیار، علوم اعصاب، گروه روان‌شناسی بالینی، دانشکده روان‌شناسی، واحد کرج، دانشگاه آزاد اسلامی، البرز، ایران

۳- دکتری تخصصی، مدیریت فن‌آوری اطلاعات، گروه مدیریت فن‌آوری اطلاعات، دانشکده مدیریت و اقتصاد، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

نویسنده طرف مکاتبه: فرشته پارساپور؛ کارشناس ارشد، روان‌شناسی بالینی، گروه روان‌شناسی بالینی، دانشکده روان‌شناسی، واحد کرج، دانشگاه آزاد اسلامی، البرز، ایران

Email: fparsapour1@gmail.com

www.SID.ir

Archive of SID

در روش خوشه‌بندی فازی، درجه عضویت و تعلق هر مشاهده به هر یک از خوشه‌ها، عددی بین صفر و ۱ است. این روش در مقایسه با روش خوشه‌بندی کلاسیک، با واقعیت محیط پیرامون انطباق بیشتری دارد (۷). با توجه به این که تحقیقات و کاربردهای روان‌شناسی از عدم قطعیت برخوردار است، داشتن انعطاف‌پذیری (فازی بودن مشاهدات)، باعث کاهش خطای تحلیل در پژوهش‌های روان‌شناختی می‌شود (۹، ۸).

مطالعات گوناگونی از تکنیک‌های داده‌کاوی در تشخیص اضطراب استفاده کرده‌اند. در تحقیق شجاعی استبرق و همکاران با استفاده از شبکه Bayesian و بر اساس عوامل شناختی- رفتاری و نشانه‌های قابل مشاهده، اضطراب اجتماعی با دقت ۸۹/۸ درصد در گروه آزمایش تشخیص داده شد (۱۰). در پژوهش Malladi با استفاده از ANN و بر اساس برخی متغیرها از جمله اشتغال، ساعات کاری، احساس عصبانیت، ناتوانی در کنترل نگرانی‌ها و...، اضطراب فراگیر با دقت ۹۰/۳۲ درصد در گروه آزمایش تشخیص داده شد (۱۱). در مطالعه Sau و Bhakta نیز با استفاده از الگوریتم CatBoost و بر اساس متغیرهای مربوط به سلامت، جمعیت‌شناختی و شغلی، اضطراب با دقت ۸۲/۶ درصد تشخیص داده شد (۱۲).

بررسی تحقیقات پیشین نشان می‌دهد که تمرکز بر روش‌های خوشه‌بندی، در تشخیص اضطراب نسبت به سایر روش‌های داده‌کاوی در حال گسترش است؛ چرا که در تعداد مشاهدات کم، روش‌های بدون ناظر همچون خوشه‌بندی، عملکرد بهتری نسبت به روش‌های همراه با ناظر همچون درخت تصمیم‌گیری دارد (۱۳). از طرف دیگر، پدیده‌های روانی، ماهیتی مبهم و غیر دقیق دارند و مرز آن‌ها مشخص نیست و مقدار آن‌ها به جای قطعی بودن، در یک طیف معینی به طور مثال بین صفر تا ۱ به طور غیر قطعی تعیین می‌شود (۸). بنابراین، روش‌های خوشه‌بندی فازی می‌توانند به حل مسایل مبهم و غیر دقیق در روان‌شناسی کمک کنند. برخی پژوهش‌ها در حوزه به کارگیری تکنیک خوشه‌بندی فازی، در تشخیص اختلالات روانی انجام شده است.

Yu و Lin در مطالعه خود، با استفاده از الگوریتم‌های خوشه‌بندی C- میانگین فازی و K- میانگین، شدت افسردگی را در نمونه‌ای از افراد بر اساس مشخصه‌های مقیاس BDI (Beck Depression Inventory) پیش‌بینی کردند. در مقایسه با تشخیص روان‌پزشک، الگوریتم C- میانگین فازی به عنوان مدل بهینه انتخاب گردید (۱۴). در تحقیق حسن‌پور و همکاران با استفاده از الگوریتم خوشه‌بندی C- میانگین فازی و بر اساس اطلاعات حاصل از پرسش‌نامه Y-BOCS (Yale–Brown Obsessive Compulsive Scale)، بیماران مبتلا به اختلال وسواس فکری- عملی به دو تا شش خوشه تقسیم شدند که با توجه به نتایج شاخص‌های خوشه‌بندی، دسته‌بندی با دو خوشه بهترین نتایج را ارایه داد (۱۵). در پژوهش پورمحمدی و ملکی، با استفاده از روش خوشه‌بندی C- میانگین فازی و بر اساس مجموعه داده‌های واکنش‌های فیزیولوژیکی افراد در شرایط رانندگی، معیاری پیوسته از میزان استرس افراد در طول رانندگی ارایه گردید که می‌توان به کمک آن، میزان استرس فرد در طول رانندگی را در هر لحظه تشخیص داد (۱۶).

در مطالعه حاضر از الگوریتم C- میانگین فازی برای خوشه‌بندی مراجعه‌کنندگان به سه کلینیک روان‌شناسی برای تشخیص میزان اضطراب بر اساس آزمون BAI (Beck Anxiety Inventory) استفاده گردید. الگوریتم C- میانگین فازی شباهت بسیاری به الگوریتم مشهور K- میانگین دارد. از طرف دیگر، این الگوریتم کاربرد بسیار وسیعی در تحقیقات مختلف دارد. در پژوهش حاضر، مراجعه‌کنندگان به چهار خوشه تقسیم شدند و برای هر خوشه یک برچسب در خصوص میزان اضطراب تعیین شد. سپس نتایج حاصل از اجرای الگوریتم با تشخیص یک پزشک متخصص مقایسه گردید و نشان داد که الگوریتم C- میانگین فازی، مراجعه‌کنندگان را با دقت بالایی خوشه‌بندی می‌کند. هدف از انجام مطالعه حاضر، ارایه روشی برای تشخیص شدت اضطراب با استفاده از تکنیک خوشه‌بندی C- میانگین فازی و تعیین تأثیر هر مشخصه در سنجش اضطراب و خوشه‌بندی مراجعه‌کنندگان بود.

روش بررسی

این تحقیق از نوع توصیفی بود. ابتدا مجموعه اطلاعات مرتبط به پرونده مراجعه‌کنندگان به سه کلینیک تخصصی روان‌پزشکی جامی، معتمد و تریتا در شهر تهران جمع‌آوری گردید. مجموعه داده‌ها شامل یک بانک اطلاعاتی مقادیر تکمیل شده ۲۳ مشخصه برای ۳۰۰ مراجعه‌کننده بود. این اطلاعات بر اساس مقیاس BAI در زمستان سال ۱۳۹۸ توسط مراجعه‌کنندگان تکمیل شد. معیار ورود نمونه‌ها به پژوهش، تکمیل مقیاس BAI در پرونده بیمار بود. روایی و پایایی این آزمون به ترتیب ۰/۷۲ و ۰/۸۳ گزارش شده است (۱۸، ۱۷).

مشخصه‌ها، شدت اضطراب را برای هر یک از مراجعه‌کنندگان تعیین نمود. این مشخصه‌ها در واقع همان سؤالات آزمون تعیین شدت مقیاس BAI بودند. پس از آماده‌سازی و پیش‌پردازش داده‌ها، مجموعه داده‌ها برای اجرا توسط الگوریتم خوشه‌بندی C- میانگین فازی به کار گرفته شد.

مشخصه‌های مطرح در مقیاس BAI شامل کرختی، داغی، لرزش پا، ناتوانی در آرامش، ترس از وقوع حادثه بد، سرگیجه و منگی، تپش قلب و نفس‌نفس زدن، حالت متغیر (بی‌ثبات)، وحشت‌زده، عصبی، احساس خفگی، لرزش دست، لرزش بدن، ترس از دست دادن کنترل، به سختی نفس کشیدن، ترس از مردن، ترسیده (حالت ترس)، سوء هاضمه و ناراحتی در شکم، غش کردن (از حال رفتن)، سرخ شدن صورت و عرق کردن (نه در اثر گرما) بود (۱۹).

در مرحله پیش‌پردازش داده‌ها، متغیرهای اسمی کدگذاری شدند. در این مرحله، داده‌های مرتبط به برخی از مراجعه‌کنندگان ناقص بود و در نتیجه، از مجموعه داده حذف شدند. در تحلیل داده‌ها، از مشخصه‌های جمعیت‌شناختی استفاده نشد و مشخصه‌های سن و جنسیت نیز فقط در ابتدای پرسش‌نامه مورد سؤال قرار گرفت.

مراحل اجرای الگوریتم خوشه‌بندی C- میانگین فازی (۲۰) در ادامه آمده است. در ابتدا همه اطلاعات شامل مقادیر مشخصه‌های ۳۰۰ مراجعه‌کننده به صورت جدولی به عنوان ورودی الگوریتم در نظر گرفته شد. در ادامه، تعداد خوشه‌ها تعیین گردید. تعداد خوشه‌ها برابر با تعداد دسته‌بندی داده‌ها (۴ نوع بیمار بر اساس داده‌های ورودی) می‌باشد. این خوشه‌ها شامل «خوشه اول: فاقد اضطراب، خوشه دوم: اضطراب خفیف، خوشه سوم: اضطراب متوسط و خوشه چهارم: اضطراب شدید» بود (۲۱). سپس مرکز هر خوشه به طور تصادفی مشخص و فاصله هر داده (مراجعه‌کننده) تا مرکز خوشه تعیین شد. تعیین این فاصله با استفاده از روش فاصله اقلیدسی صورت گرفت.

فاصله اقلیدسی دو داده عبارت از جذر جمع توان دوم تفاضل مقادیر مشخصه‌های دو داده است (۲۲). در ادامه، به تعیین میزان درجه عضویت هر مراجعه‌کننده به هر خوشه بر اساس فاصله اقلیدسی پرداخته شد. هر مراجعه‌کننده که بیشترین تعلق و عضویت را به خوشه مورد نظر داشته باشد، در همان خوشه

http://him.mui.ac.ir
www.SID.ir

Archive of SID

قرار می‌گیرد. میزان درجه عضویت، عددی بین صفر و ۱ است که در الگوریتم خوشه‌بندی C- میانگین فازی محاسبه می‌گردد. این مراحل تا زمانی تکرار می‌شود که داده‌های مراجعه‌کنندگان در یک خوشه بیشترین شباهت را نسبت به هم داشته باشند. این الگوریتم در نرم‌افزار MATLAB R2015a اجرا و نتایج ارایه گردید.

پس از اجرای الگوریتم، نتایج به دست آمده با نتایج تشخیص یک پزشک متخصص به عنوان یک فرد خبره مقایسه شد. مقایسه به صورت یک به یک برای هر مراجعه‌کننده انجام گرفت؛ بدین ترتیب که پزشک برای هر مراجعه‌کننده، تشخیصی را بر اساس برچسب اختصاصی بین ۱ تا ۴ ارایه داد. این تشخیص به کمک الگوریتم خوشه‌بندی C- میانگین فازی نیز انجام شد. در نهایت، برچسبی که پزشک متخصص مشخص کرده بود، با برچسب تعیین شده توسط الگوریتم برای هر مراجعه‌کننده مقایسه گردید.

در پژوهش حاضر، علاوه بر خوشه‌بندی مراجعه‌کنندگان با استفاده از روش C- میانگین فازی، این الگوریتم به طور جداگانه بر روی هر مشخصه مقیاس BAI نیز اجرا و تأثیر هر مشخصه در سنجش اضطراب و خوشه‌بندی مراجعه‌کنندگان تعیین شد.

ملاحظات اخلاقی در مطالعه در نظر گرفته شد. در کلیه مراحل تحقیق حفظ محرمانگی اطلاعات مراجعه‌کنندگان، اولویت اصلی کار بود. اطلاعات فردی مراجعه‌کنندگان که از طریق آن هویت مراجعه‌کننده شناسایی می‌شود، در مجموعه داده‌های تهیه شده مورد تحلیل قرار نگرفت. هر یک از شرکت‌کنندگان بر اساس یک کد منحصر به فرد به طور محرمانه شناسایی شدند.

یافته‌ها

الگوریتم خوشه‌بندی C- میانگین فازی برای بخش‌بندی مراجعه‌کنندگان در سنجش میزان اضطراب آن‌ها به کار گرفته شد. مراجعه‌کنندگان شامل ۱۳۸ مرد و ۱۶۲ زن بود. ۱۴۴ نفر از آن‌ها در بازه سنی ۱۶ تا ۳۰ سال و ۱۵۶ بیمار در بازه سنی ۳۱ تا ۵۵ سال قرار داشتند. پس از اجرای الگوریتم خوشه‌بندی، مراجعه‌کنندگان بر اساس درجه عضویت در چهار خوشه بخش‌بندی شدند. شکل ۱ میزان تعلق همه مراجعه‌کنندگان را در خوشه‌ها نشان می‌دهد. به طور نمونه، مراجعه‌کننده اول با چه درجه عضویتی متعلق به کدام خوشه است.

در ادامه، به مقایسه نتایج حاصل از اجرای الگوریتم C- میانگین فازی با نظرات پزشک متخصص پرداخته شد. برای مقایسه نتایج از شاخص‌های دقت (Accuracy)، صحت (Precision)، حساسیت (Sensitivity) و خاص‌بودن (Specificity) که طبق روابط ۱ تا ۴ محاسبه شدند، استفاده گردید (۲۴، ۲۳) که در روابط مذکور، TP (True Positive) بیانگر تعداد مراجعه‌کنندگانی است که در یک خوشه از اضطراب قرار گرفته‌اند و الگوریتم خوشه‌بندی نیز آن‌ها را در آن خوشه قرار داده است. TN (True Negative) بیان‌کننده تعداد مراجعه‌کنندگانی است که در خوشه دیگری از اضطراب قرار گرفته‌اند و الگوریتم خوشه‌بندی نیز آن‌ها را در آن خوشه قرار داده است. FP (False Positive) نشان دهنده تعداد مراجعه‌کنندگانی است که در یک خوشه از اضطراب قرار گرفته‌اند، اما الگوریتم آن‌ها را به اشتباه در خوشه‌های دیگر قرار داده است. FN (False Negative) نیز بیانگر تعداد مراجعه‌کنندگانی است که در خوشه‌های دیگری از اضطراب قرار گرفته‌اند، اما الگوریتم آن‌ها را به اشتباه در خوشه مورد نظر خود قرار داده است.

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN}$$ رابطه ۱

$$Percision = \frac{TP}{TP+FP}$$ رابطه ۲

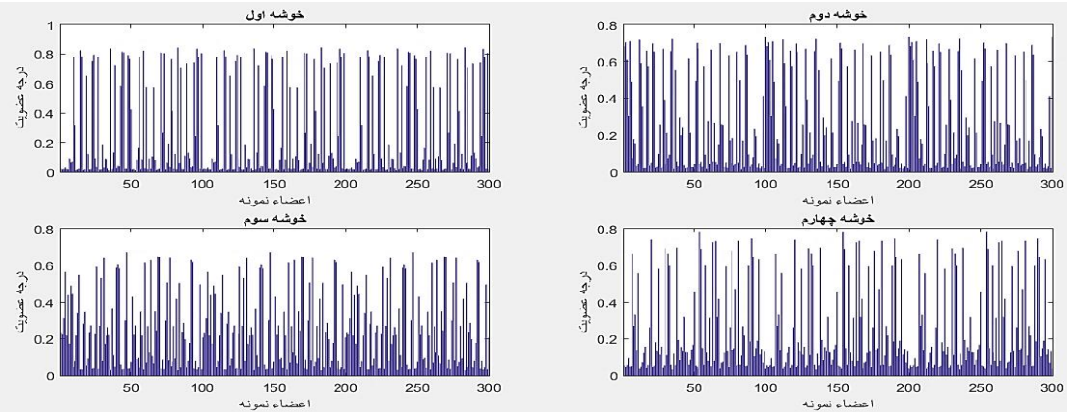
$$Sensitivity = \frac{TP}{TP+FN}$$ رابطه ۳

$$Specivity = \frac{TP}{FP+TN}$$ رابطه ۴

دقت نهایی الگوریتم C- میانگین فازی در مقایسه با نظر پزشک متخصص در خوشه‌بندی مراجعه‌کنندگان برای تشخیص اضطراب آن‌ها، ۹۰/۶۶ درصد به دست آمد. بر این اساس، صحت، حساسیت و خاص بودن به ترتیب ۸۷/۶۲، ۹۳/۹۶ و ۸۲/۷۸ درصد گزارش گردید.

برای مقایسه الگوریتم C- میانگین فازی با نظر پزشک متخصص از ماتریس درهم‌ریختگی استفاده شد.

در شکل ۲، قطر اصلی نشان دهنده خوشه‌بندی صحیح الگوریتم در مقایسه با نظر پزشک است. به طور نمونه، ۱۳۸ مراجعه‌کننده فاقد اضطراب بودند که الگوریتم خوشه‌بندی نیز آن‌ها را به درستی تشخیص داد. ۶ مراجعه‌کننده در خوشه اول قرار گرفتند؛ یعنی توسط پزشک فاقد اضطراب تشخیص داده شدند، اما الگوریتم آن‌ها را در خوشه دوم قرار داد؛ بدین معنی که آن‌ها دارای اضطرب خفیف بودند.

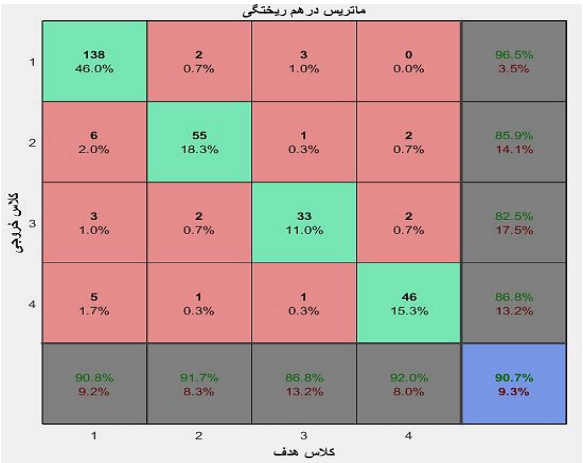


شکل ۱: درجه تعلق مراجعه‌کنندگان به هر یک از خوشه‌ها

محور افقی: بیماران، محور عمودی: درجه تعلق هر بیمار به خوشه

http://him.mui.ac.ir
www.SID.ir

Archive of SID



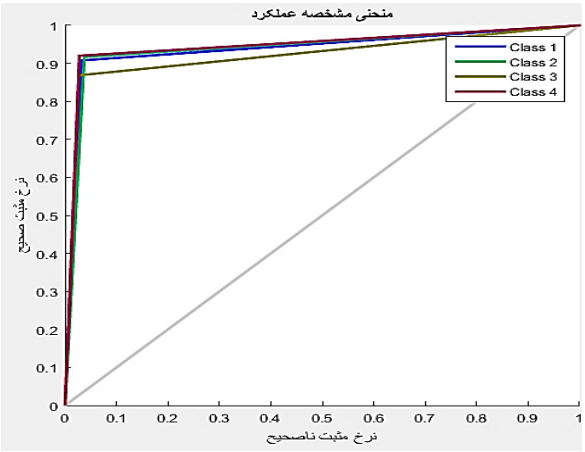
شکل ۲: ماتریس درهم‌ریختگی الگوریتم C- میانگین فازی

تعداد مراجعه‌کنندگانی که به کمک الگوریتم به طور صحیح و ناصحیح خوشه‌بندی شده بودند، در جدول ۱ آمده است.

جدول ۱: عملکرد الگوریتم C- میانگین فازی در خوشه‌بندی

خوشه	تعداد مراجعه‌کنندگان در هر خوشه	تعداد مراجعه‌کنندگان به طور صحیح خوشه‌بندی شده	تعداد مراجعه‌کنندگان به طور ناصحیح خوشه‌بندی شده
۱	۱۵۲	۱۳۸	۱۴
۲	۶۰	۵۵	۵
۳	۳۸	۳۳	۵
۴	۵۰	۴۶	۴

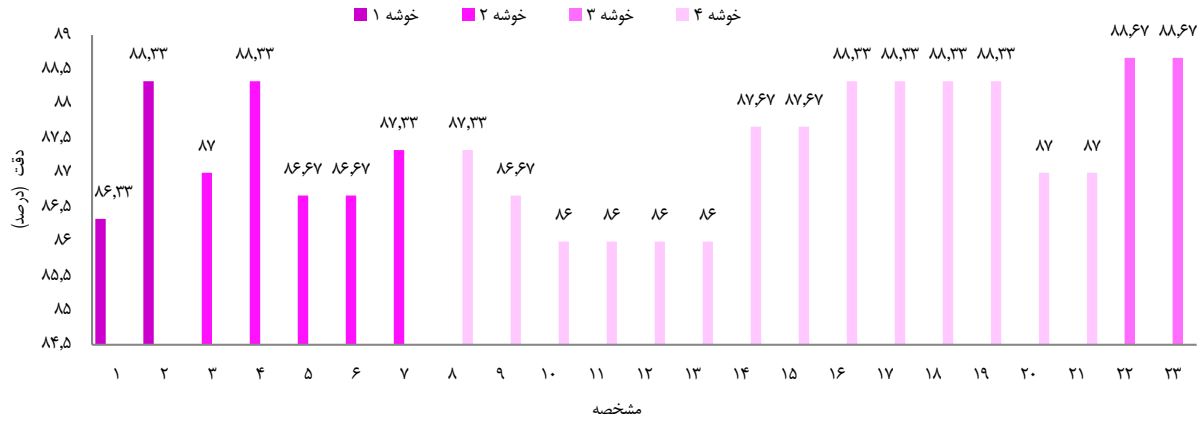
شکل ۳ به ارایه منحنی مشخصه عملکرد الگوریتم C- میانگین فازی برای هر خوشه از مراجعه‌کنندگان پرداخت. هر چقدر سطح زیر منحنی بیشتر باشد، دقت عملکرد الگوریتم بیشتر است. همان‌طور که مشاهده می‌شود، الگوریتم توانست مراجعه‌کنندگان متعلق به خوشه ۴ (یعنی بیمارانی که دارای اضطراب شدید هستند) را نسبت به سایر مراجعه‌کنندگان در خوشه‌های دیگر با دقت بیشتر بخش‌بندی کند.



شکل ۳: منحنی مشخصه عملکرد الگوریتم C- میانگین فازی

پژوهش حاضر به بررسی تأثیر مشخصه‌ها در سنجش اضطراب و خوشه‌بندی مراجعه‌کنندگان پرداخت که شکل ۴ گویای این موضوع می‌باشد.

دقت مشخصه، درصد مراجعه‌کنندگانی را نشان می‌دهد که خوشه آن به‌درستی توسط آن مشخصه پیش‌بینی شده است. علاوه بر این، در ارزیابی هر مشخصه، خوشه‌ای که تعداد نمونه‌های خوشه‌بندی صحیح آن بیشتر از سایر خوشه‌ها بود، مشخص گردید. به طور نمونه، شکل ۴ نشان داد که در ارزیابی مشخصه‌های ۸ تا ۲۱، تعداد نمونه‌های خوشه‌بندی صحیح خوشه چهارم (اضطراب شدید) بیشتر از سایر خوشه‌ها می‌باشد و در این بین، چهار مشخصه ۱۶ تا ۱۹ [ترس از دست دادن کنترل، به سختی نفس کشیدن، ترس از مردن، ترسیده (حالت ترس)] دقت بیشتری در سنجش اضطراب داشتند.



شکل ۴: تأثیر هر مشخصه سنجش اضطراب در خوشه‌بندی

http://him.mui.ac.ir
www.SID.ir

Archive of SID

بحث

نتایج پژوهش حاضر نشان داد که الگوریتم C- میانگین فازی توانست مراجعه‌کنندگان را بر اساس مشخصه‌های اضطراب، با دقت بالایی خوشه‌بندی کند. با استفاده از این الگوریتم، درجه عضویت و تعلق هر یک از مراجعه‌کنندگان به خوشه مورد نظر و در نتیجه، میزان اضطراب آن‌ها به صورت بیان فازی (فاقد اضطراب، خفیف، متوسط و شدید) تعیین گردید. الگوریتم C- میانگین فازی قادر بود بیمارانی را که میزان اضطراب آن‌ها شدید است را نسبت به سایر مراجعه‌کنندگان، با دقت بالاتری شناسایی نماید. همچنین، تأثیر هر مشخصه در سنجش اضطراب و خوشه‌بندی مراجعه‌کنندگان تعیین شد که این مهم می‌تواند برای سنجش سطح اضطراب بیماران توسط پزشک مورد توجه قرار گیرد.

برخی مطالعات در به کارگیری داده‌کاوی به منظور سنجش اضطراب بیماران انجام شده است که از آن جمله می‌توان به تحقیق Cardenas و همکاران (۲۵) اشاره کرد. آن‌ها از الگوریتم K- میانگین برای تفکیک دانش‌آموزان مضطرب به شش خوشه استفاده کردند، اما از تعداد اندک (۱۱۶ نمونه) برای خوشه‌بندی استفاده شد (۲۵). با توجه به کیفی بودن مفهوم اضطراب، پژوهش مورد نظر از مفاهیم فازی بهره نبرده و نتایج خوشه‌بندی به صورت غیر فازی ارایه شده است. یک عدم توازن بین تعداد بیماران در خوشه‌ها برقرار می‌باشد که نمی‌تواند ارایه دهنده بخش‌بندی مناسبی از آن‌ها باشد. در نهایت، اهمیت و تأثیر هر مشخصه آزمون Beck برای بیماران در خوشه‌ها تعیین نشده است.

Iliou و همکاران در مطالعه خود، به مقایسه روش پیشنهادی خود با روش پیش‌پردازش تحلیل مؤلفه‌های اساسی در پیش‌بینی سطح اضطراب دانشجویان پرداختند. سپس نتایج را با هفت تکنیک طبقه‌بندی مقایسه کردند و به این نتیجه رسیدند که روش پیشنهادی در تکنیک SVM، دارای بالاترین دقت به میزان ۹۸ درصد است. در تحقیق آن‌ها، تعداد کم ۱۱۳ نمونه برای طبقه‌بندی به کار گرفته شد و از مفاهیم سیستم‌های فازی برای کاربردی‌تر کردن نتایج استفاده نگردید. Iliou و همکاران از تکنیک‌های خوشه‌بندی نیز در تحلیل داده‌های خود استفاده نکردند و به ارزیابی تأثیر هر یک از مشخصه‌ها بر سطح اضطراب پرداخته نشد (۲۶).

تحقیق Husain و همکاران از تکنیک جنگل تصادفی به منظور پیش‌بینی اختلال اضطراب فراگیر در زنان استفاده کرد و دقت مدل طبقه‌بندی، ۹۳/۷۹ درصد گزارش گردید. ۱۱۲ نمونه برای مدل به کار گرفت شد. آن‌ها تعداد نمونه کمی را برای محاسبات خود لحاظ کردند و از روش‌های خوشه‌بندی و مفاهیم فازی استفاده نکردند (۲۷).

به عنوان وجوه تمایز، اول این که در پژوهش حاضر از مفاهیم فازی برای بیان مقادیر مشخصه‌ها و حالات بیماران مضطرب استفاده شد. دوم این که مطالعه حاضر به تعیین تأثیر هر یک از مشخصه‌های آزمون Beck در سنجش اضطراب و خوشه‌بندی مراجعه‌کنندگان پرداخت.

می‌توان در آینده از مشخصه‌های جمعیت‌شناختی که به نظر می‌رسد نقش قابل توجهی در سنجش اضطراب داشته باشد و در پژوهش حاضر وارد نشد، استفاده نمود. همچنین، می‌توان حیطه نمونه‌ها را به کلینیک‌های دیگری افزایش داد تا یافته‌های جدیدی حاصل شود.

نتیجه‌گیری

نتایج تحقیق حاضر نشان داد که این الگوریتم از دقت بالایی در تشخیص سطح اضطراب بیماران برخوردار است. می‌توان متوجه شد که هر مراجعه‌کننده با چه درجه عضویتی متعلق به کدام خوشه از شدت اضطراب است. همچنین، میزان تأثیر هر مشخصه در سنجش اضطراب و خوشه‌بندی مراجعه‌کنندگان تعیین می‌گردد. روش ارایه شده در پژوهش حاضر می‌تواند به پزشک کمک کند تا از طریق خوشه‌بندی مراجعه‌کنندگان، شدت اضطراب آن‌ها را مشخص و تأثیر هر مشخصه در خوشه‌بندی را تعیین نماید.

پیشنهادها

پیشنهاد می‌شود پزشک در ابتدا مراجعه‌کنندگان را با یک درجه عضویت به صورت فازی متعلق به یک خوشه از اضطراب در نظر بگیرد. سپس بر اساس هر خوشه، به تعیین مشخصه‌های اضطراب برای هر یک از مراجعه‌کنندگان بپردازد. بنابراین، نیازی نیست تا همه مشخصه‌های اضطراب برای بیمار مورد ارزیابی قرار گیرد. همچنین، مراجعه‌کنندگان جدید بر اساس نوع خوشه و مشخصه‌های مرتبط به آن خوشه، شناسایی و شدت اضطراب آن‌ها تعیین شود.

تشکر و قدردانی

بدین وسیله از کلیه کسانی که در انجام این تحقیق همکاری نمودند، تشکر و قدردانی به عمل می‌آید.

تضاد منافع

در انجام پژوهش حاضر، نویسندگان هیچ‌گونه تضاد منافعی نداشته‌اند.

References

1. Mukherjee S, Ashish K, Hui N, Chattopadhyay S. Modeling depression data: Feed forward neural network vs. radial basis function neural network. Am J Biomed Sci 2014; 6(3): 166-74.
2. Wittchen HU, Jacobi F. Size and burden of mental disorders in Europe--a critical review and appraisal of 27 studies. Eur Neuropsychopharmacol 2005; 15(4): 357-76.
3. Wittchen HU, Kessler RC, Beesdo K, Krause P, Hofler M, Hoyer J. Generalized anxiety and depression in primary care: prevalence, recognition, and management. J Clin Psychiatry 2002; 63(Suppl 8): 24-34.
4. Jegan C. Heart attack prediction system using fuzzy C means classifier. IOSR J Comput Eng 2013; 14(2): 23-31.
5. Jooriyan N, Ashoori M. Predicting the effectiveness of preeclampsia medications based on dose and method of drug consumption using data mining. Iran J Obstet Gynecol Infertil 2014; 17(123): 13-22.
6. de Oliveira JV, Pedrycz W. Advances in fuzzy clustering and its applications. Chichester, UK: John Wiley and Sons; 2007. 3-424.
7. Pedrycz W. Knowledge-Based Clustering: From Data to Information Granules. Hoboken, NJ: John Wiley and Sons; 2005. 1-297.

www.SID.ir

Archive of SID

8. Ghanbari H, Safaeian M, Toozandeh Jani H, Talebian Sharif J. Application of fuzzy sets in psychological studies. Proceedings of the International Conference on Psychology, Consultation and Education; 2017 Mar 21; Mashhad, Iran. 86-374. [In Persian].
9. Dubois D, Prade H. Fundamentals of Fuzzy Sets. New York, NY: Springer US; 2012. 4-637
10. Shojaei Estabragh Z, Riahi Kashani MM, Jeddi Moghaddam F, Sari S, Taherifar Z, Moradi Moosavy S, et al. Bayesian network modeling for diagnosis of social anxiety using some cognitive-behavioral factors. Netw Model Anal Health Inform Bioinform 2013; 2(4): 257-65.
11. Malladi S. An approach of artificial neural networks for prediction of generalized anxiety disorder. Int J Res Comput Appl Robot 2015; 33(3): 118-24.
12. Sau A, Bhakta I. Screening of anxiety and depression among the seafarers using machine learning technology. Inform Med Unlocked 2019; 16: 100228.
13. Yang MS, Hwang PY, Chen DH. Fuzzy clustering algorithms for mixed feature variables. Fuzzy Set Syst 2004; 141(2): 301-17.
14. Yu V, Lin YH. Applications of fuzzy theory on health care: an example of depression disorder classification based on FCM. WSEAS Trans Inf Sci Appl 2008; 5(1): 31-6.
15. Hasanpour H, Asadi S, Ghavamizadeh MR, Daraeian A, Ahmadiani A, Shams J, et al. A critical appraisal of heterogeneity in Obsessive-Compulsive Disorder using symptom-based clustering analysis. Asian J Psychiatr 2017; 28: 89-96.
16. Pourmohammadi S, Maleki A. A fuzzy C-means clustering approach for continuous stress detection during driving. Signal and Data Processing 2018; 14(4): 129-42. [In Persian].
17. Hossein KH, Mousavi AS. Psychometric properties of the Persian version of Beck Anxiety Inventory (BAI). Tehran Univ Med J 2008; 66(2): 136-40. [In Persian].
18. Golden J, Conroy RM, O'Dwyer AM. Reliability and validity of the Hospital Anxiety and Depression Scale and the Beck Depression Inventory (Full and Fast Screen scales) in detecting depression in persons with hepatitis C. J Affect Disord 2007; 100(1-3): 265-9.
19. Blacker D. Psychiatric Rating Scales. In: Sadock BJ, Sadock V, editors. Comprehensive textbook of psychiatry. 8th ed. Philadelphia, PA: Lippincott Williams and Wilkins; 2005. 929-59.
20. Bezdek JC. Cluster validity with fuzzy sets. Journal of Cybernetics 1973; 3(3): 58-73.
21. Beck AT, Steer RA. Manual for the Beck Anxiety Inventory. San Antonio, TX: Psychological Corporation; 1990.
22. Fazel Zarandi MH, Faraji MR, Karbasian M. An exponential cluster validity index for fuzzy clustering with crisp and fuzzy data. Scientia Iranica 2010; 17(2): 95-110.
23. Ghazanfari M, Alizadeh S, Teimorpour B. Data mining and knowledge discovery. 2nd ed. Tehran, Iran: Publication of Iran University of Science and Technology; 2011. 14-398. [In Persian].
24. Han J, Kamber M, Pei J. Data mining: Concepts and techniques. Waltham, MA: Morgan Kaufmann Publishers; 2012. 1-625.
25. Cardenas F, Hernandez E, Zenil A, Redondo A. Application method of data mining using the K means algorithm for the determination of stress level in high school students using the Beck Depression Inventory. Ciencia Huasteca Boletin Científico de la Escuela Superior de Huejutla 2020; 8(15): 1-8.
26. Iliou T, Konstantopoulou G, Stephanakis I, Anastasopoulos K, Lymberopoulos D, Anastassopoulos G. Iliou machine learning data preprocessing method for stress level prediction. In: Iliadis L, Maglogiannis I, Plagianakos V, editors. Artificial intelligence applications and innovations. Cham, Switzerland: Springer International Publishing; 2018.351-61.
27. Husain W, Xin LK, Rashid NA, Jothi N. Predicting generalized anxiety disorder among women using random forest approach. Proceedings of the 3rd International Conference on Computer and Information Sciences (ICCOINS); 2016 Aug 15-17s; Kuala Lumpur, Malaysia. 37-42.

www.SID.ir

Archive of SID

Health Information Management

Published by *Vesnu Publications*

Vol 17/ No 3/ August & September 2020

DOI: 10.22122/him.v17i3.4139

Using Fuzzy C-means Clustering Algorithm to Diagnose the Severity of Anxiety

Fereshteh Parsapour[1], Javid Peymani[2], Mohammad Khanbabaei[3]

Original Article

Abstract

Introduction: Diagnosing anxiety in the early stages by psychiatrists is one of the important steps in preventing and controlling these types of disorders. This study endeavors to present a method to diagnose the severity of anxiety using fuzzy C-means clustering (FCM) algorithm. Moreover, the influence of each feature on measuring anxiety and clustering of clients is determined.

Methods: This was a quantitative and descriptive study with a dataset including 300 clients related to three psychiatric clinics in Tehran, Iran provided based on the Beck Anxiety Inventory (BAI). Then, the FCM algorithm was utilized to segment the clients and determine the severity of their anxiety in each cluster. Additionally, this algorithm was employed for each feature separately.

Results: The psychiatric clinics' clients were divided into four clusters with the labels including no, minimal, moderate, and severe anxiety. Using the FCM algorithm, the anxiety of the clients was diagnosed with 90.66% accuracy. Moreover, as a result of implementing the algorithm on each feature, the influence of the features on measuring anxiety and clustering of clients was determined.

Conclusion: The FCM algorithm diagnosed the anxiety of clients with a high accuracy. Segmenting patients by the clustering approach and based on the important features can be a dependable instrument for psychiatrists to make a decision in diagnosing the severity of anxiety in the early stages.

Keywords: Cluster Analysis; Fuzzy Logic; Anxiety

Received: 18 June, 2020 Accepted: 24 July, 2020 Published: 05 Aug., 2020

Citation: Parsapour F, Peymani J, Khanbabaei M. **Using Fuzzy C-means Clustering Algorithm to Diagnose the Severity of Anxiety.** Health Inf Manage 2020; 17(3): 97-103.

Article resulted from an independent research without financial support.

1- MSc, Clinical Psychology, Department of Clinical Psychology, School of Psychology, Karaj Branch, Islamic Azad University, Alborz, Iran
2- Assistant Professor, Neuroscience, Department of Clinical Psychology, School of Psychology, Karaj Branch, Islamic Azad University, Alborz, Iran
3- PhD, Information Technology Management, Department of Information Technology Management, School of Management and Economics, Science and Research Branch, Islamic Azad University, Tehran, Iran
Address for correspondence: Fereshteh Parsapour; MSc, Clinical Psychology, Department of Clinical Psychology, School of Psychology, Karaj Branch, Islamic Azad University, Alborz, Iran; Email: fparsapour1@gmail.com

www.SID.ir