

نمونه‌ها در امتداد هر مسیر به انجام می‌رسد. در واقع از دیدگاه فازی برای ساده‌سازی ساختار درخت تصمیم‌گیری استفاده می‌شود.

در طی این مقاله پس از ایجاد درخت تصمیم‌گیری بر اساس اطلاعات موجود در داده‌های آموزشی تلاش خواهد شد که با بکارگیری اطلاعات موجود در ساختار درخت تصمیم‌گیری، مجموعه قواعد فازی اما- اگر توصیف‌کننده درخت استخراج شود. در ادامه الگوریتمی پیشنهاد خواهد شد که ضمن کاهش تعداد قواعد فازی، تا حد ممکن بخش مقدم هر قاعده فازی را نیز ساده‌سازی کند. در واقع هدف ارائه الگوریتمی است که قواعد فازی را از درخت تصمیم‌گیری استخراج کند و سپس با توجه به ساختار درخت تصمیم‌گیری به ساده‌سازی آن بپردازد بگونه‌ای که در نهایت منجر به افزایش صحت طبقه‌بندی داده شود.

۲- ویژگی داده پزشکی، داده فازی و معرفی بانک اطلاعات مورد استفاده:

۱-۲- داده‌های پزشکی و فازی:

داده‌های پزشکی ذاتاً پیچیده‌اند؛ بدین صورت که در اغلب رکوردهای پزشکی با شماری از انواع داده مواجه می‌شویم و شامل مواردی هستند که تاریخچه بیمار، آزمایشات فیزیکی، تست‌های آزمایشگاهی، گزارشات آسیب‌شناسی، گزارشات تصویری و ثبت ضربان قلب را شرح می‌دهند.

از داده فازی برای در نظر گرفتن عدم قطعیت موجود در داده‌ها و نزدیک نمودن اینگونه مفاهیم به بینش یک متخصص خبره استفاده می‌شود. بنابراین برای بیان متوسط بودن دمای بدن بجای استفاده از یک عدد خاص، می‌توان بیان متوسط بودن را بجای نمایش بازه ۳۶ تا ۳۷ درجه استفاده نمود؛ و یا اینکه برای نمایش بالابودن دمای بدن بجای استفاده از یک عدد خاص، می‌توان از نمایش بازه بزرگتر از ۳۷ درجه با بیان دمای بالا استفاده نمود؛ هر دو مورد ذکر شده در بالا دارای دیدگاه فازی می‌باشند. در بیشتر کاربردها برای فازی نمودن اعداد، از توابع تعلق استخراج‌شده بر مبنای دانش فرد خبره و یا دانش استخراج‌شده از داده، استفاده می‌شود.

۲-۲- ساختار بانک اطلاعاتی مورد استفاده:

بانک اطلاعات بکارگرفته‌شده در این مقاله [۴]، توسط شارون سامر از دانشگاه کانزاس و لیندا وولری از دانشگاه میسوری جمع‌آوری شده است و توسط جری گریزمالا از دانشگاه کانزاس در دسترس قرار گرفته است. این بانک اطلاعاتی حاوی داده‌های فازی است که عمل فازی‌سازی توسط متخصص بر روی آن انجام شده است (بر اساس تعریف هر متغیر در ذیل). در واقع پایگاه داده بصورت فازی شده (داده‌های با بیان گفتاری) در دسترس قرار دارد. این بانک شامل ۹۰ بیمار است که به بخش مراقبت بعد از عمل جراحی آورده شده‌اند و از هر یک از بیماران ۸ ویژگی خاص اندازه‌گیری شده است تا براساس این ویژگیها تصمیم‌گیری شود که این بیماران قابلیت انتقال به منزل، بخش عمومی و یا قسمت مراقبت‌های ویژه بیمارستان را دارا می‌باشند. این ویژگی‌ها عبارتند از:

۱-۲-۲- دمای داخلی بدن بیمار (L-CORE):

این ویژگی به سه دسته تقسیم و برای هر دسته یک برچسب در نظر گرفته شده است:

۳- High (> 37)

۲- Mid (≥ 36 and ≤ 37)

۱- Low (< 36)

۲-۲-۲- دمای سطحی بدن بیمار (L-SURF):

دمای سطحی بدن بیمار به سه دسته تقسیم و برای هر دسته یک برچسب در نظر گرفته شده است:

۳- High (> 36.5)

۲- Mid (≥ 35 and ≤ 36.5)

۱- Low (< 35)

۲-۲-۳- درصد اکسیژن خون بیمار (L-O2):

به چهار دسته تقسیم و برای هر دسته یک برچسب در نظر گرفته شده است:

۴- Excellent (≥ 98)

۳- Good (≥ 90 and < 98)

۲- Fair (≥ 80 and < 90)

۱- Poor (< 80)

۲-۲-۴- آخرین اندازه‌گیری فشار خون بیمار (L-BP):

این ویژگی، به سه دسته تقسیم و برای هر دسته یک برچسب در نظر گرفته شده است:

۳- High ($> 130/90$)

Mid ($\leq 130/90$ and $\geq 90/70$) - 2

Low ($< 90/70$) - 1

2-2-5- پایداری دمای سطحی بدن بیمار (SURF-STBL):
به سه دسته تقسیم و برای هر دسته یک برچسب در نظر گرفته شده است:

Stable - 3

Mod-Stable - 2

Unstable - 1

2-2-6- پایداری دمای مرکزی بدن بیمار (CORE-STBL):
این مشخصه به سه دسته تقسیم و برای هر دسته یک برچسب در نظر گرفته شده است:

Stable - 3

Mod-Stable - 2

Unstable - 1

2-2-7- پایداری آخرین فشار خون بیمار (BP-STBL):
به سه دسته تقسیم و برای هر دسته یک برچسب در نظر گرفته شده است:

Stable - 3

Mod-Stable - 2

Unstable - 1

2-2-8- آسودگی خیال در ترخیص بیمار:

که توسط عدد صحیحی بین 0 تا 20 نشان داده می شود.
این داده‌ها برای اولین بار در مراجع [6,5] استفاده شد.

3- مرور تحقیقات گذشته به همراه نتایج:

در این زیربخش چند روش متفاوت که بر روی بانک اطلاعات مورد استفاده در این مقاله کار نموده به همراه نتایج آنها جهت طبقه‌بندی بیماران مرور خواهند شد. هدف از طبقه‌بندی، این است که بتوان در فضای ویژگی‌های شرح داده شده در زیربخش 2-2، قسمتی که بیمار باید به آن منتقل شود را مشخص نماییم.

3-1- روش LERS در طبقه‌بندی بیماران:

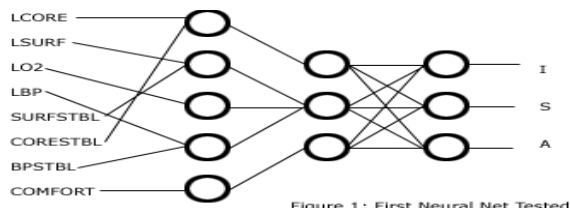
سیستم یادگیری LERS (یادگیری توسط مثالها بر پایه مجموعه‌های سخت) بر اساس قوانین مجموعه‌ها، توسط یک الگوریتم 4 مرحله‌ای و انتخاب یک ویژگی تصمیم که قادر به تصمیم‌گیری‌های تجاری و یا تصمیماتی در حیطه پزشکی است، به ما اجازه تولید قوانینی را می‌دهد که دارای

ویژگی‌های خاصی در سمت راست قانون می‌باشند. بعنوان مثال، اگر ویژگی تصمیم ما با توجه به زیربخش 2-2 کلاس دمای بدن باشد؛ چه بالا، چه میانه و چه پایین آنگاه می‌توانیم بیاموزیم که چه ویژگی‌هایی بالا، میانه و پایین را با استفاده از ویژگی‌های دیگر تعریف می‌کنند. روش LERS برای طبقه‌بندی بیماران موجود در این پایگاه داده منجر به صحت 48٪ بر روی کل داده شد [6,5].

3-2- روش شبکه عصبی در طبقه‌بندی بیماران [7]:

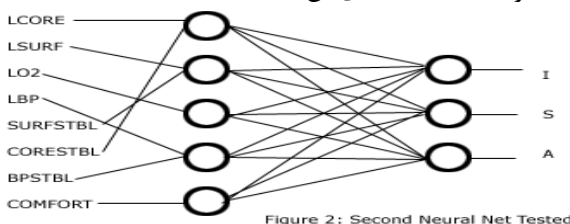
این مقاله با اختیار کردن شبکه عصبی با سه گره خروجی و حداقل کردن تابع خطای RMS، محل انتقال بیمار را برای کاربر نمایش می‌دهد. بمنظور نمایش معماری شبکه عصبی، سه طراحی برگرفته شده از مرجع [7] در شکل‌های 1، 2 و 3 نشان داده شده‌اند و نمایش مناسب از لحاظ زمان کمتر آموزش ساختار شکل 3 می‌باشد. در هر طراحی، بدلیل نیاز به وجود تمامی اطلاعات جهت تصمیم‌گیری محل قرارگیری بیمار، گره‌های خروجی به گره‌های لایه قبل از خروجی کاملاً متصل هستند. این مقاله از معیار صحت برای ارزیابی عملکرد طبقه‌بندی با استفاده از شبکه عصبی استفاده می‌کند.

شبکه اول سه لایه بوده، متغیرهای سطح اکسیژن و مطلوبیت لایه میانی را تشکیل می‌دهند و دارای بیشترین زمان آموزش و خطای RMS است. صحت این ساختار 40.2٪ است:



شکل 1: اولین شبکه عصبی مورد آزمایش [7]

شبکه دوم دو لایه بوده و دارای کوچکترین خطای RMS است و صحت طبقه‌بندی آن 46٪ است:



شکل 2: دومین شبکه عصبی مورد آزمایش [7]

شبکه سوم سه لایه بوده، متغیرهای سطح اکسیژن و مطلوبیت لایه میانی را تشکیل می‌دهند و دارای کمترین زمان آموزش است و صحت آن ۴۰.۲٪ است.

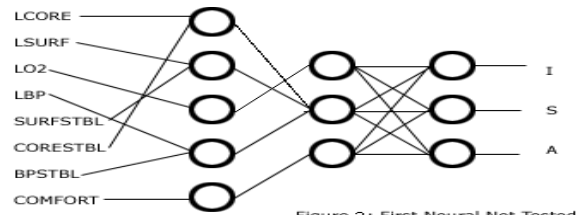


Figure 3: First Neural Net Tested

شکل ۳: سومین شبکه عصبی مورد آزمایش [۷]

ویژه بیمارستان انتقال یافتند ۲ نفر، تعداد افرادی که به منزل ترخیص شدند ۲۴ نفر و تعداد افراد منتقل شده به بخش عمومی ۶۴ نفر است. از آنجاییکه پایگاه داده دردسترس بصورت فازی شده در اختیار بود، مقادیر ویژگیها همه بصورت گفتاری بودند و در واقع الگوریتم ارائه شده، یک سیستم استنتاج فازی-گفتاری می‌باشد. در این سیستم فازی بدلیل عدم دسترسی به مقادیر واقعی ویژگیها و نیز گفتاری بودن محل مراقبت از بیمار، از بخشهای واسطه فازی ساز و واسطه غیرفازی ساز خروجی استفاده نشده است و تنها قسمت میانی یا تصمیم گیرنده اصلی (بخش استنتاج فازی) در طراحی سیستم بکار رفته است که براساس یکسری قوانین اگر و آنگاه تولید می‌شود. در واقع الگوریتم طراحی شده با دیدگاهی که توسط متخصص بکارگرفته می‌شود قرابت دارد.

جدول ۱: ماتریس Confusion بدست آمده با روش LHCM

		تشخیص بدست آمده با الگوریتم LHCM		
		I.C.U	بخش عمومی	منزل
تصمیم واقعی	I.C.U	۱	۱	۰
	بخش عمومی	۲	۴۱	۲۱
	منزل	۰	۱۸	۶

۴-۲- بررسی ساختار مجموعه قوانین تولیدشده در برنامه:

بانک اطلاعات موجود دارای یک ماتریس ۸*۹۰ (که حاوی اطلاعات مربوط به ۹۰ بیمار و مقادیر مختلف اندازه گیری شده برای ۸ ویژگی می‌باشد) و یک بردار ۱*۹۰ (که شامل سه برچسب مختلف برای این ۹۰ بیمار است، بدین صورت که هر یک از این برچسبها نشان می‌دهد که بیمار مربوطه برحسب تشخیص پزشک به کدام قسمت باید انتقال یابد) است. تعیین قواعد فازی بگونه‌ای که ارتباط مناسب و مطلوب را بین ویژگیهای ثبت شده از بیمار و برچسب خروج همان بیمار از قسمت احیاء اتاق عمل برقرار نماید دارای اهمیت است. همچنین این قواعد، با ورود بیمار جدید به عنوان داده ورودی باید قادر به تشخیص درست محل مناسب ترخیص برای او باشند. بنابراین ۹۰ داده این بانک اطلاعات بصورت تصادفی به دو دسته ۷۰ و ۳۰ درصدی

۳-۳- روش C-Means سخت و C-Means سخت گفتاری در طبقه بندی بیماران [۸]:

در روش C-Means بصورت تصادفی برای هر خوشه یک نماینده برگزیده می‌شود؛ یک نمونه از فضای داده‌ها بصورت تصادفی انتخاب شده و براساس فاصله بین مراکز خوشه‌ها رقابت ایجاد و این داده اندکی به سمت برنده جابجا می‌شود. این مقاله روش C-Means سخت (HCM) را به روش C-Means سخت گفتاری بهبود بخشیده و قادر به گروه بندی بردارهای گفتاری، بر پایه ویژگیهای همسان آنها و ارزیابی نتیجه می‌باشد. ملاک توقف الگوریتم C-Means سخت گفتاری، قرارگرفتن بر روی حداکثر مقدار شمارنده (۸۰) و یا عدم تجانس نماینده خوشه‌ها به کمتر یا مساوی ۰.۰۰۰۱ است. این مقاله از ماتریس Confusion بر روی کل ۹۰ داده و صحت استخراج شده از این ماتریس جهت ارزیابی عملکردش استفاده می‌کند. این ماتریس برای روش LHCM در جدول ۱ آورده شده است. صحت الگوریتم HCM برابر با ۵۱.۱٪ و برای الگوریتم LHCM برابر با ۵۳.۳٪ گزارش شد [۸].

۴- سیستم استنتاج فازی-گفتاری (روش پیشنهادی):

۴-۱- توضیح کلی سیستم پیشنهادی:

در الگوریتم ارائه شده در این مقاله، برحسب محل انتقال بیمار به او برچسبی نسبت داده می‌شود (برچسب ۰ برای انتقال به بخش مراقبتهای ویژه، برچسب ۱ برای ترخیص و برچسب ۲ برای انتقال به بخش عمومی). از ۹۰ نفر موجود در این بانک اطلاعات تعداد افرادی که به قسمت مراقبتهای

تقسیم‌بندی شدند؛ بدین صورت که ۷۰٪ داده برای استخراج قواعد فازی سیستم و ۳۰٪ داده برای تست کارآئی سیستم بصورت تصادفی از هر سه برچسب خروجی مطلوب سیستم انتخاب شدند. تعداد بیماران با برچسبهای خروجی متفاوت، در دو دسته ۳۰٪ و ۷۰٪، در جدول ۲ قابل مشاهده است. در ادامه این مقاله f1 تا f7 بترتیب متنظر با ویژگیهای L-CORE، L-SURFE، L-O2، L-BP، SURF-STBL، CORE-STBL و BP-STBL توضیح داده شده در زیربخش ۲-۲ هستند.

جدول ۲: دسته‌بندی بیماران با برچسبهای خروجی متفاوت در دو دسته ۷۰٪ (آموزش) و ۳۰٪ (تست):

برچسب خروج بیمار	۷۰٪ بیماران	۳۰٪ بیماران
I.C.U (برچسب ۰)	۱ نفر	۱ نفر
منزل (برچسب ۱)	۱۷ نفر	۷ نفر
بخش (برچسب ۲)	۴۵ نفر	۱۹ نفر

برطبق توضیحات داده شده و جدول ۱، الگوریتم سه مرحله‌ای زیر را جهت بررسی ساختار مجموعه قوانین تولیدشده مدنظر قرار می‌دهیم و در نرم‌افزار Matlab نسخه ۷.۴ پیاده‌سازی می‌کنیم:

۴-۲-۱- مرحله اول: برنامه جستجو

بدلیل تنوع بیش از اندازه داده‌های ورودی، دسته‌بندی مناسب و دستی بر روی آنها دشوار است. این مرحله می‌تواند نمایش مناسبی برای دسته‌بندی این داده‌ها جهت کاهش تعداد قوانین موردنیاز برای آموزش این سیستم را فراهم نماید. نمایش درختی برای ۷۰٪ بیماران که در تعیین قواعد فازی نقش دارند، دارای اهمیت ویژه‌ای است که در این مرحله به انجام می‌رسد. شکل ۴، درخت ایجاد شده برای افرادی که ترخیص خواهند شد را نشان می‌دهد. مشابه این درخت، برای دسته‌بندی ۷۰٪ بیمارانی که دارای سایر برچسبهای خروجی دیگر هستند، نیز مورد نیاز است. با اجرای این برنامه در واقع می‌توانیم امکان هرس کردن و کاهش تعداد قوانین لازم برای بیان این سیستم طبقه‌بندی را طبق توضیحات بعدی فراهم نماییم. شایان ذکر است که در شکل ۴، اعداد نشان‌داده‌شده در دایره‌های نمایشگر تعداد افراد و اعداد نشان‌داده شده بر روی خطوط، نمایشگر مقادیر مختلف ۸ ویژگی f1 تا f7 می‌باشد.

۴-۲-۲- مرحله دوم: الگوریتم نوشتن قوانین آنگاه و اگر:

از روی درخت‌های بدست آمده از زیربخش ۴-۲-۱ (براساس ۷۰٪ داده‌ها) قوانین آنگاه و اگر را بشرح زیر بدست می‌آوریم:

۴-۲-۱- گام اول: تعیین اولیه قوانین

در این گام جهت تعیین اولیه مجموعه قوانین، با توجه به درختهای منتج از برنامه جستجو در زیربخش ۴-۲-۱ از الگوریتم شکل ۵ استفاده می‌شود.

احتمال هر ویژگی را از تقسیم‌نمودن تعداد مقادیر مختلف یک ویژگی، بر کل تعداد داده‌های آن ویژگی خاص بدست می‌آوریم که در جداول ۳ و ۴ به تفکیک برای دو برچسب (۱) و (۲) در خروجی آورده شده‌اند. شایان ذکر است برای برچسب خروجی (۰) به دلیل کمی تعداد آن، احتمال هر ویژگی مدنظر قرار نگرفته است.

۴-۲-۲- گام دوم: اصلاح نمودن قوانین

هدف از این گام، اصلاح قوانین منتج از گام اول جهت فعال کردن تنها یک قانون به ازاء ورود هر بیمار از دسته آموزشی سیستم می‌باشد. این مرحله با توجه به داده آموزشی به انجام می‌رسد. بدلیل مشابهت مقادیر ویژگیهای برخی بیماران چند قانون بصورت همزمان فعال می‌شدند، که برای رفع این مشکل از مقدار ویژگی هشتم که یک ویژگی کیفی بود در قواعد فازی استفاده نمودیم.

در اصلاح برخی قوانین نیز، حتی اضافه کردن مقدار این ویژگی بدلیل نوع داده‌های موجود در بانک اطلاعاتی و حضور داده‌هایی با مقادیر ویژگی کاملاً یکسان که در دو دسته مختلف در خروجی جای گرفته بودند، کمکی به ما ننمود و این امر موجب ایجاد خطای آموزش شد. شایان ذکر است که بدلیل کمی تعداد بیمارانی که برچسب خروجی (۰) داشتند، بصراحت نمی‌توان اذعان داشت که این برچسب به درستی آموزش دیده است یا خیر.

۴-۲-۳- گام سوم: بررسی نتایج خروجی بر روی قوانین اصلاح شده

هدف از این گام بررسی نتایج طبقه‌بندی ۳۰٪ بیماران، بر روی مجموعه قوانین استخراج شده از ۷۰٪ بیماران است (ارزیابی الگوریتم پیشنهادی). بدین منظور در این برنامه، قوانین اصلاح شده در مراحل قبل توسط ۳۰٪ داده ورودی

که به بخش مراقبت‌های ویژه (I.C.U) منتقل می‌شوند با توجه به تعداد کم داده نمی‌توان اظهار نظر خاصی کرد.

جدول ۵: ماتریس Confusion بدست آمده با روش پیشنهادی مقاله

		تشخیص بدست آمده با الگوریتم پیشنهادی		
		I.C.U	بخش عمومی	منزل
واقعیت	I.C.U	۱	۰	۱
	بخش عمومی	۰	۶۳	۱
	منزل	۰	۹	۱۵

همچنین با تقسیم نمودن تعداد دفعات تکرار هر ویژگی در قوانین بدست آمده از گام یک و دو الگوریتم بر تعداد کل قوانین آموزشی موجود، طبق نمودار ۱، توانستیم میزان مشارکت ویژگی‌های بانک اطلاعات را در روش استنتاج فازی-گفتاری (روش پیشنهادی) بدست آوریم که نماینده‌ای از میزان اهمیت ویژگیها است.

۶- مقایسه نتایج با سایر روشها:

در روش شبکه‌عصبی از ۳ داده بدلیل ناقص بودن و عدم دارا بودن مقدار، استفاده نشده‌است. در روش ما همه داده‌ها یا در آموزش و یا در ارزیابی سیستم شرکت نموده‌اند و ناقص بودن برخی از ویژگیها در الگوریتم ارائه شده توسط ما نیازی به حذف داده مربوط به آنها را ندارد.

در روش C-means سخت نیز بدلیل استفاده از روش نگاشت قبل از اجرا، نتایج قابل قبول نمی‌باشد و اگر روش غیرفازی نمودن تغییر کند نتیجه هم متفاوت خواهد بود. حال آنکه الگوریتم پیشنهادی ما بدون نیاز به هر گونه نگاشتی و فقط بر اساس اطلاعات بدست آمده از داده به طبقه‌بندی می‌پردازد.

مقایسه روش استنتاج فازی-گفتاری (روش پیشنهادی) با روشهای LERS، شبکه‌عصبی و روش C-Means در نمودار ۲ آورده شده است. همانگونه که قابل مشاهده است، روش استنتاج فازی-گفتاری بیشترین میزان صحت در تعیین بخش مراقبت از بیمار را در بین روشهای ارائه شده را دارا می‌باشد.

روش پیشنهادی مقاله با قابلیت اصلاح کردن بر مبنای درخت تصمیم‌گیری، نسبت به روش‌های یادشده دارای

که از قبل به صورت اتفاقی از بانک اطلاعات انتخاب می‌شوند، مورد آزمایش قرار می‌گیرد. در این گام ماتریس Confusion محاسبه می‌شود و با کمک درایه‌های قطر اصلی آن تطبیق برچسب تعیین شده توسط متخصص و نتیجه حاصل از قواعد استخراج شده و همچنین صحت طبقه‌بندی استخراج و بررسی می‌شود.

جدول ۳: احتمال مقادیر مختلف هر ویژگی برای برچسب (۱) در خروجی و روی ۷۰٪ داده مورد آزمایش:

مقدار ویژگی	f _۱	f _۲	f _۳	f _۴	f _۵	f _۶	f _۷
۱	٪۲۳	٪۲۹	٪۰	٪۰	٪۴۱	٪۱۷	٪۱۷
۲	٪۷۰	٪۵۸	٪۰	٪۷۰	٪۰	٪۰	٪۱۷
۳	٪۷	٪۱۳	٪۴۷	٪۳۰	٪۵۹	٪۸۳	٪۶۶
۴			٪۵۳				

جدول ۴: احتمال مقادیر مختلف هر ویژگی برای برچسب (۲) در خروجی و روی ۷۰٪ داده مورد آزمایش:

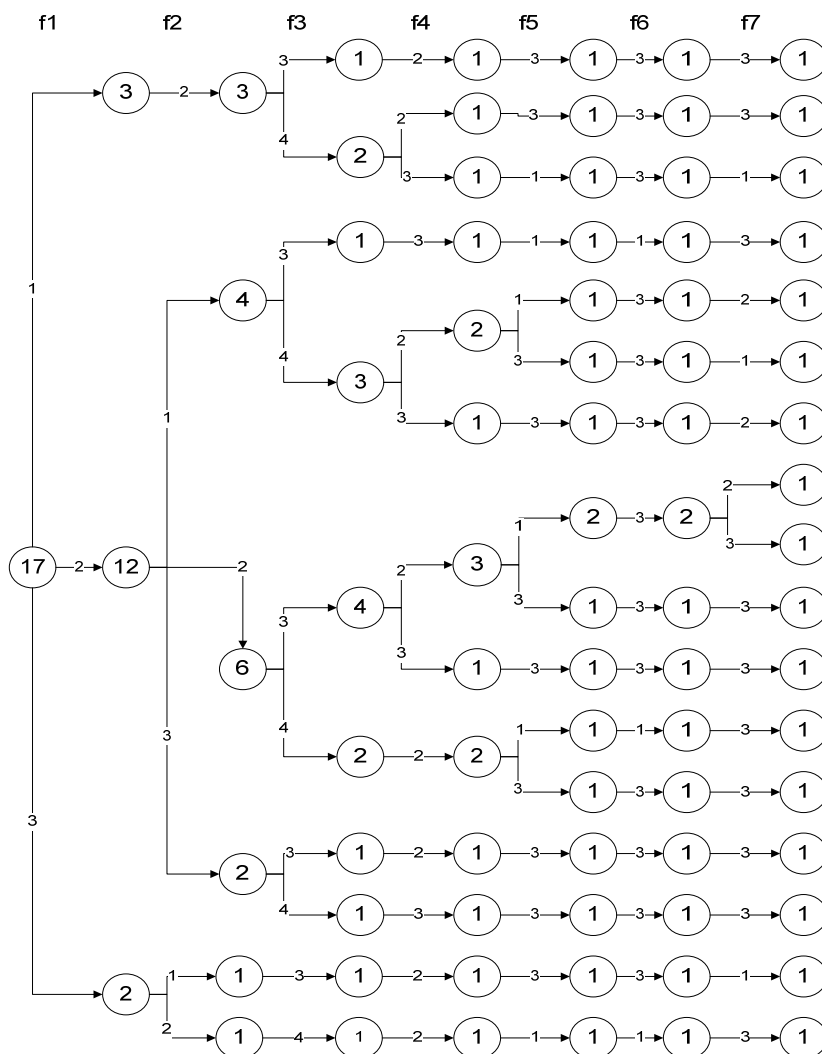
مقدار ویژگی	f _۱	f _۲	f _۳	f _۴	f _۵	f _۶	f _۷
۱	٪۲۰	٪۳۳	٪۰	٪۶	٪۵۱	٪۶	٪۲۴
۲	٪۶۴	٪۴۲	٪۰	٪۵۷	٪۰	٪۲	٪۳۵
۳	٪۱۶	٪۲۶	٪۵۵	٪۳۷	٪۴۹	٪۹۲	٪۴۱
۴			٪۴۵				

۵- نتایج:

بمنظور ارزیابی الگوریتم ارائه شده، ماتریس Confusion و صحت بدست آمده از آن یک بار برای نتیجه ۳۰٪ داده تست و بار دیگر برای کل داده استخراج شد.

با اجرا نمودن این الگوریتم برای ۳۰٪ داده عده‌ای از افراد تست، برچسب خروجی واقعی خود را فعال نمی‌نمایند و این خطایی است که از آن با نام خطای آموزش این الگوریتم یاد می‌کنیم. در این حالت به صحت ۶۶.۶٪ در طبقه‌بندی بیماران دست یافتیم.

با توجه به محدودیت فضا و بمنظور مقایسه با نتایج روش زیربخش ۳-۳ فقط ماتریس Confusion مربوط به کل داده در جدول ۵ آورده شده است. صحت طبقه‌بندی برای کل داده از این ماتریس برابر با ۸۷.۷٪ است. در رابطه با افرادی



شکل ۴: درخت رسم شده برای دسته‌بندی ۷۰٪ بیمارانی که دارای برچسب خروجی (۱) بوده‌اند.

مزیت‌های زیر می‌باشد:

۱- بدلیل عدم نیاز به وجود داده‌های عددی و کافی دانستن تنها ویژگیهای گفتاری (فازی شده)، سیستم فازی طراحی شده بسیار راحت‌تر از دیگر روش‌ها قابلیت طراحی و اجرا را دارا است.

۲- دسته‌بندی داده‌ها به دو دسته ۷۰٪ و ۳۰٪ جهت آموزش و ارزیابی سیستم، نتایج مناسبی را فراهم کرد.

۳- برنامه جستجو که بشرح آن پرداختیم، امکان دسته‌بندی به هر شکل و در هر بانک اطلاعات مشابهی را در اختیار قرار می‌دهد و آسانی کار در نوشتن قوانین را مهیا می‌سازد.

۴- استفاده از روش هرس نمودن (توسط برنامه جستجو) و روش رشد (شروع از یک ساختار حداقل قوانین و به تدریج گسترش دادن در حد نیاز) در نوشتن قوانین مربوط، ترکیب

مناسبی در بهینه‌سازی قوانین توصیف‌کننده مجموعه می‌باشد. در واقع تولید قوانین و هرس کردن، بصورت خودکار و همزمان به انجام می‌رسد.

۵- وجود داده‌های نامشخص (NaN) که در مقادیر بعضی ویژگیها وجود داشتند، در ایجاد قواعد سیستم مشکلی ایجاد نمی‌نماید.

۶- درجه اهمیت هر ویژگی در طبقه‌بندی بر اساس توضیحات زیربخش نتایج فراهم شد.

۷- جمع‌بندی:

روش پیشنهادی در این مقاله برای طبقه‌بندی داده‌های فازی که با بیان گفتاری در پایگاه داده موجود بودند، ارائه شد. از جمله ویژگیهای روش ذکر شده، ایجاد درخت داده

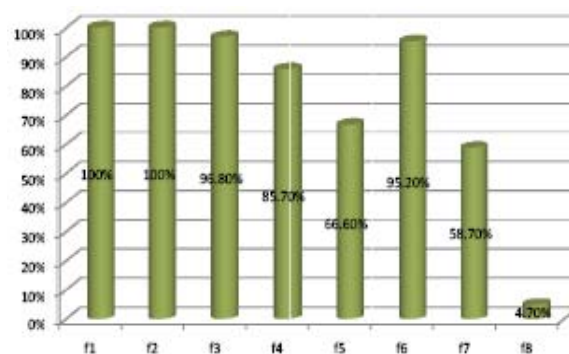
- [2] J. Mingers. "An Empirical Comparison of Pruning Methods for Decision Tree Induction". Machine Learning 4, 1989, pp. 227-243.
- [3] M. Dong, R. Kothari. "Look-Ahead Based Fuzzy Decision Tree Induction". IEEE Transaction on Fuzzy Systems, 2001
- [4] A. Asuncion, D.J. Newman. UCI Machine Learning Repository [http://archive.ics.uci.edu/ml/datasets/Post-operative+Patient]. Irvine, CA: University of California, School of Information and Computer Science.
- [5] A. Budihardjo, J. Grzymala-Busse, L. Woolery, "Program LERS_LB ۲/۵ as a tool for knowledge acquisition in nursing", Proceedings of the 4th Int. Conference on Industrial & Engineering Applications of AI & Expert Systems, pp. ۷۳۵-۷۴۰, ۱۹۹۱.
- [6] L. Woolery, J. Grzymala-Busse, S. Summers, A. Budihardjo, The use of machine learning program LERS_LB ۲/۵ in knowledge acquisition for expert system development in nursing. Computers in Nursing ۹, pp. ۲۲۷-۲۳۴, ۱۹۹۱.
- [7] mitesh trivedi, "Neural network applications-postoperative patient discharge", cas'۰۳.
- [8] Sansanee Auephanwiriayukul, and Nipon Theera-Umpon, "Comparison of Linguistic and Regular Hard C-Means in Postoperative Patient Data", Journal of Advanced Computation Intelligence and Intelligent Informatics, Vol. ۸, No. ۶ pp. ۵۹۹-۶۰۵, ۲۰۰۴.



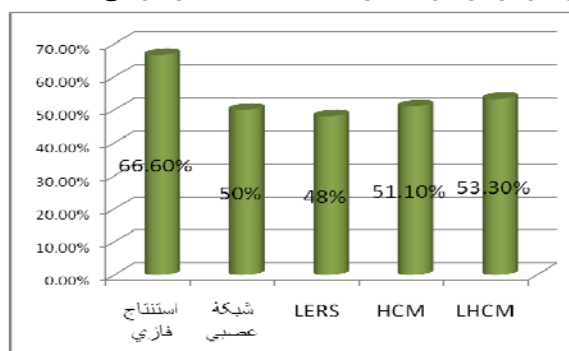
شکل ۵: الگوریتم تعیین اولیه مجموعه قوانین

بر اساس ۷۰٪ داده است که از این درخت برای ایجاد قواعد فازی، هرس نمودن قواعد و ساده‌تر کردن آنها استفاده می‌شود. همچنین روش تعیین قواعد ذکرشده را می‌توان براحتی بر روی داده‌های فازی یا پایگاه داده‌هایی که توسط یکی از روشهای فازی‌سازی، فازی شده‌است اعمال کرد. روش پیشنهادی با صحت بالاتری نسبت به سایر روشهایی که تاکنون بر روی پایگاه داده مرجع [۴] اعمال شده بود، عمل نمود. همچنین این روش به ناقص بودن و عدم مشخص بودن بعضی از ویژگیها حساس نبود. همچنین در روش پیشنهادی، هر چه تعداد داده و تنوع حالت‌های هر گروه بیشتر باشد منجر به نتایج بهتری در طبقه‌بندی می‌شود.

نمودار ۱: میزان مشارکت (بر حسب درصد) ویژگیها در روش استنتاج فازی-گفتاری:



نمودار ۲: میزان صحت طبقه‌بندی (بر حسب درصد) در پنج روش مختلف اعمال شده به پایگاه داده مورد استفاده در این مقاله. صحت عملکرد استنتاج فازی (روش پیشنهادی مقاله) در صورت در نظر گرفتن کل داده از ۶۶.۶٪ به ۸۷.۷٪ افزایش می‌یابد.



۸- مراجع:

- [1] L. Breiman, J.H. Friedman, R.A. Olsen & C.J. Stone. "Classification and Regression Trees". Wadsworth, 1984.