

تشخیص خودکار نارسایی‌های صوتی از طریق پردازش صدا

بهرام وزیرنژاد

محمدحسن مرادی

دانشگاه صنعتی امیرکبیر، دانشکده مهندسی پزشکی
{bvazirnezhad, mhmoradi}@cic.aut.ac.ir

چکیده

۱- مقدمه

در این تحقیق به منظور تشخیص خودکار نارسایی‌های صوتی، ویژگی‌هایی مرتبط با هدف، از سیگنال صوتی واکه کشیده شده^۱ استخراج نمودیم. این ویژگی‌های تمایزگر در صورتی که با یک طبقه‌بندی‌کننده مناسب مورد ارزیابی قرار گیرند، نتایج خوبی قابل انتظار خواهد بود. دو روش طبقه‌بندی ویژگی‌ها به دو کلاس بیمار و طبیعی مورد توجه قرار گرفته است. طبقه‌بندی با استفاده از طبقه‌بندی‌کننده بیزین با مدل گوسی و درخت تصمیم که یک روش آماری است صورت گرفت. و برای هر کدام از گروه زن‌ها و مرد‌ها یک طبقه‌بندی‌کننده بصورت جداگانه تعلیم داده شد. دادگان مورد استفاده شامل ۳۲۹ صدای بیمار و ۵۸ صدای طبیعی بود. از هریک از فایل‌های صوتی در فریم‌های ۴۰ میلی‌ثانیه‌ای، طول گام واکه با استفاده از روش‌های آنالیز دنباله کپستروم و یا روش خودهمبستگی، بعلاوه ضریب مربوط به اغتشاش دامنه استخراج گردید. با استفاده از مقادیر فرکانس پایه استخراج شده از فریم‌های موجود در هریک از فایل‌های صوتی مقادیر میانگین فرکانس پایه، بالاترین و پایین‌ترین فرکانس پایه و انحراف معیار فرکانس پایه و پارامترهای اغتشاش فرکانسی و اغتشاش دامنه به عنوان مشخصه‌های تمایزگر برای طبقه‌بندی استخراج گردیدند. برای تعیین میزان کارایی طبقه‌بندی‌کننده‌ها از روش ارزیابی متقابل^۲ استفاده گردید. که درخت تصمیم کارایی خوبی از خود نشان داد.

واژه‌های کلیدی: نارسایی‌های صوتی، پردازش صداهای

پاتولوژیک، ناتوانایی‌های گفتاری

نارسایی‌های صوتی شامل انواع گوناگونی با منشاءهای فیزیکی یا عصب-عضله‌ای و یا صرفاً عصبی می‌باشند، این عوامل بصورت مستقیم کیفیت صدای تولیدی توسط بیمار را تحت تاثیر قرار می‌دهند. در حال حاضر ابزارهای تشخیصی نظیر استروبوسکوپی ویدئویی^۳، کیموگرافی ویدئویی^۴ و سیستم‌های نمایشگر حرکات تارهای صوتی در اختیار گفتاردرمانگرها قرار گرفته است. لاکن روش‌های ذکرشده تهاجمی، وقت‌گیر و گران می‌باشند[۱].

واکه‌های کشیده‌شده که با طول گام و سطح بلندی ذاتی توسط شخص تولید شده‌اند، از نظر کلینیکی برای تشخیص بسیاری از بیماری‌های گفتاری مورد توجه است. بسیاری از محققان عنوان نموده‌اند که واکه‌های کشیده‌شده می‌توانند منعکس‌کننده بسیاری از نارسایی‌ها و نکات غیرطبیعی مانند اغتشاشات فرکانسی سیکل به سیکل در واکه، و اغتشاشات دامنه یا قطع غیرارادی صدا یا لرزش صدا باشند. این موارد می‌توانند بیانگر ضعف در کنترل فشار هوای تولیدی توسط ریه‌ها یا نارسایی‌های حنجره یا دیگر جهاز صوتی و یا حتی مشکلات عصبی باشند، باید عنوان نمود چنین تغییراتی در واکه کوتاه مدت رخ نمی‌دهد[۱، ۲، ۳].

در این مقاله به گزارش روشی مطمئن و ساده برای تشخیص زود هنگام و حتی ارزیابی روند درمان نارسایی‌های صوتی می‌پردازیم. مزایای این روش در مقایسه با روش‌های دیگر طی موارد زیر عنوان می‌گردد.

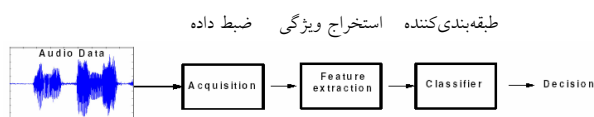
³ Videostroboscopy

⁴ Videokymography

¹ Sustained Vowel

² Cross-Validation

ویژگی و طبقه‌بندی‌کننده می‌باشد. نمودار بلوکی فرآیند طبقه‌بندی در شکل زیر آمده است:



شکل ۱- نمودار بلوکی فرآیند طبقه‌بندی صدای پاتولوژیک

۲-۱- دادگان

برای ضبط داده از بیمار خواسته می‌شود تا به مدت حدود ۳ الی ۵ ثانیه به تولید بدون وقفه واکه /a/ اقدام نماید. صدای تولید شده با فرکانس نمونه‌برداری ۴۴۱۰۰ هرتز و ۱۶ بیتی و بصورت تک‌کاناله ضبط می‌گردد و جهت پردازش در رایانه ذخیره می‌گردد. در این تحقیق از دادگان برجسب‌خورده 'Disordered Voice Database Model 4337' تهیه شده در آزمایشگاه گفتار و صدا، مرکز تحقیقات بیمارستان تخصصی گوش و چشم ماساچوست، جهت تعلیم طبقه‌بندی‌کننده‌ها استفاده گردید. این دادگان شامل تعداد ۳۸۷ فایل صوتی واکه کشیده شده /a/ می‌باشد. بیمارانی که صدای آنها در این دادگان در اختیار است مبتلا به نارسایی‌های صوتی از نظر بیماری‌های جهاز صوتی و یا سیستم عصبی بوده‌اند. تعداد فایل‌های صوتی موجود در دادگان برای مردان و زنان در دو دسته طبیعی و پاتولوژیک در جدول ۱ ذکر شده است.

جدول ۱- تعداد فایل‌های صوتی موجود در دادگان برای مردان و زنان در

دو دسته طبیعی و پاتولوژیک

جنسیت	تعداد نمونه‌های پاتولوژیک	تعداد نمونه‌های طبیعی	مجموع
مردان	۱۴۱	۲۲	۱۶۳
زنان	۱۸۸	۳۶	۲۲۴

۲-۲- بخش استخراج ویژگی

سیگنال صوتی واکه کشیده شده منبع خوبی برای تشخیص انواع نارسایی‌های صوتی است. عوارض ناشی از نارسایی‌های صوتی بصورت اغتشاشات و نوسانات فرکانس پایه و یا دامنه در طول سیگنال بروز می‌نماید. قطع غیرارادی صدا و یا لرزش صدا از دیگر نشانه‌های نارسایی‌های صوتی است، که ناشی از عملکرد نامناسب عضلات و یا جهاز صوتی دخیل در کار تولید صوت می‌باشد [۵]. برای نمونه در شکل ۲ دو نمونه از

- این روش غیرتهاجمی بوده و اجرای آن برای درمانگر و بیمار بسیار ساده می‌باشد، تنها لازم است بیمار صدای /a/ را برای مدت چند ثانیه تولید نماید.
- با استفاده از این روش امکان ساخت یک سیستم تشخیص خودکار رایانه‌ای وجود دارد.
- تشخیص شیء‌گرا^۵ است چراکه بر مبنای مقادیر ویژگی‌های صوتی انجام می‌گیرد.

- این روش برای ارزیابی روند پیشرفت درمان مثل جراحی، درمان‌های دارویی یا فرآیندهای بازتوانی، قابل استفاده می‌باشد. و می‌توان با استفاده از مقادیر ویژگی‌های صوتی از میزان پیشرفت روند درمان ارزیابی کمی بدست آورد.

در این روش مشخصه‌هایی از سیگنال صوتی تولید شده توسط افراد مورد آزمایش استخراج می‌گردند، این مشخصه‌ها اطلاعاتی در مورد میزان اغتشاش طول گام و نیز اغتشاش دامنه در طول سیگنال در اختیار می‌گذارند. تحقیقات زیادی نشان داده‌است که اغتشاشات طول گام و دامنه سیگنال در افراد طبیعی با بیماران متفاوت است [۴]. این مشخصه‌ها در صورتیکه در اختیار یک طبقه‌بندی‌کننده مناسب قرار گیرند، می‌توان نتایج خوبی را انتظار داشت. در این مقاله سعی شده از ویژگی‌هایی استفاده شود که بیشترین اختلاف را بین بیماران افراد طبیعی دارا باشند. این ویژگی‌ها عمدتاً بیانگر میزان اغتشاشات فرکانسی و دامنه‌ای و نیز نوسانات موضعی فرکانس و دامنه می‌باشند. در قسمت طبقه‌بندی‌کننده از دو طبقه‌بندی‌کننده بیزین و درخت تصمیم استفاده گردید. بدلیل ساختار داده‌ها استفاده از طبقه‌بندی‌کننده بیزین با دو مدل گوسی به نتایج خوبی منجر نشد، اما استفاده از درخت تصمیم که قابلیت مدلسازی مرزهای تصمیم غیرخطی را برای داده‌های با ساختار پیچیده داراست نتایج خوبی حاصل نمود.

۲- تشریح فرآیند تشخیص نارسایی‌های صوتی

در این بخش قسمت‌های مختلف طبقه‌بندی‌کننده را بررسی می‌نماییم. این قسمت‌ها شامل بخش‌های ضبط داده، استخراج

⁵ Objective

روش‌های ذکر شده به علاوه ضریب مربوط به دامنه استخراج گردید. مقادیر استخراج شده از تمام فریم‌ها در بردارهایی قرار می‌گیرند. سپس مقدار میانگین فرکانس پایه و انحراف معیار فرکانس‌های پایه و انحراف معیار ضریب دامنه، با توجه به فرمول‌های ۱، ۲ و ۳ محاسبه می‌شوند:

$$avgF_0 = \frac{1}{N} \sum_{i=1}^N F_0(i) \quad (1)$$

$$STD = \sqrt{\frac{1}{N} \sum_{i=1}^N (F_0(i) - avgF_0)^2} \quad (2)$$

$$STDa = \sqrt{\frac{1}{N} \sum_{i=1}^N (A(i) - avgA)^2} \quad (3)$$

مقادیر اغتشاشات محلی فرکانس jitter و اغتشاشات محلی دامنه shimmer نیز بصورت زیر محاسبه می‌شوند [۱].

$$jitter = \frac{1}{N-1} \sum_{i=1}^{N-1} \frac{F_0(i+1) - F_0(i)}{\frac{1}{2}[F_0(i+1) + F_0(i)]} \times 100\% \quad (4)$$

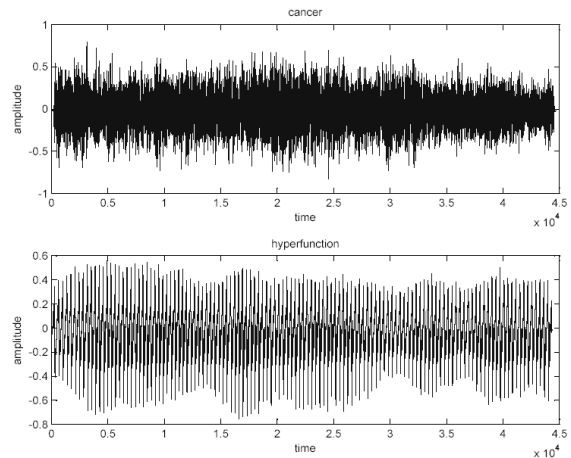
$$shimmer = \frac{1}{N-1} \sum_{i=1}^{N-1} \frac{A(i+1) - A(i)}{\frac{1}{2}[A(i+1) + A(i)]} \times 100\% \quad (5)$$

در جدول ۲ ویژگی‌های مورد استفاده به همراه توصیف کوتاه آنها ملاحظه می‌شود.

جدول ۲- توصیف ۶ مشخصه استخراج شده از سیگنال

گروه مشخصه	علامت اختصاری	تعریف مشخصه
فرکانس پایه	avgF0	میانگین فرکانس پایه در طول سیگنال
	STD	انحراف معیار فرکانس پایه در طول سیگنال
اغتشاشات محلی فرکانسی	Jitter	ضریب اغتشاش فرکانسی بین فریم‌های ۴۰ میلی‌ثانیه‌ای
	F0l	پایین‌ترین فرکانس پایه دیده شده در تمام فریم‌های ۴۰ میلی‌ثانیه‌ای
	F0h	بالاترین فرکانس پایه دیده شده در تمام فریم‌های ۴۰ میلی‌ثانیه‌ای
اغتشاش دامنه	STDa	انحراف معیار دامنه قله‌های موجود در فریم‌های ۴۰ میلی‌ثانیه‌ای
اغتشاشات محلی دامنه	Shimmer	ضریب اغتشاشات دامنه بین فریم‌های ۴۰ میلی‌ثانیه‌ای

سیگنال واکه /a/ در دو بیمار را نشان می‌دهد. سیگنال بالایی مربوط به بیمار مبتلا به سرطان حنجره و سیگنال پایینی مربوط به بیمار مبتلا به Hyperfunction است.



شکل ۲- دو نمونه سیگنال پاتولوژیک واکه کشیده شده /a/، سیگنال بالایی مربوط به بیمار مبتلا به سرطان حنجره و سیگنال پایینی مربوط به بیمار مبتلا به Hyperfunction است. اغتشاشات فرکانسی و نوسانات دامنه در هر دو سیگنال کاملاً مشهود است. چنین حالتی در سیگنال طبیعی وجود ندارد.

با توجه به آنچه ذکر شد، باید ویژگی‌ها را طوری تعریف نماییم که بتواند بین افراد طبیعی و بیمار تمایز ایجاد نماید. این ویژگی‌ها را طوری طراحی نمودیم که بتواند معیاری کمی از اغتشاشات فرکانسی و نوسانات دامنه‌ای در اختیار قرار دهد. ویژگی‌هایی که به عنوان مشخصه‌های تمایزگر از سیگنال استخراج گردید، شامل ۷ ویژگی میانگین فرکانس پایه، بالاترین فرکانس پایه، پایین‌ترین فرکانس پایه، انحراف معیار فرکانس پایه در طول سیگنال، پارامتر jitter یا ضریب اغتشاشات محلی فرکانسی و انحراف معیار ضریب دامنه در فریم‌ها و پارامتر shimmer ضریب اغتشاشات محلی دامنه، بوده است.

برای استخراج فرکانس پایه می‌توان از روش‌های آنالیز کپستروم و یا روش آنالیز تابع خودهمبستگی استفاده نمود که با توجه به پیاده‌سازی صورت گرفته، استفاده از هریک از این دو روش تفاوتی در نتایج ایجاد نمی‌نماید. فریم‌های مورد استفاده برای استخراج ویژگی‌ها با طول ۴۰ میلی‌ثانیه‌ای در نظر گرفته شد، و از این فریم‌ها فرکانس پایه با یکی از

$$\frac{p(x|w_n) > p(w_p)}{p(x|w_p) < p(w_n)} \quad (8)$$

ماتریس کوواریانس و بردار میانگین مقادیر ویژگی‌ها، بصورت جداگانه برای مردان و زنان در دو دسته طبیعی و پاتولوژیک استخراج شد و تابع تمایز خطی فوق برای دو طبقه مردان و زنان شکل گرفت. طبقه‌بندی کننده بیزین با دو مدل تابع چگالی احتمال گوسی برای دو کلاس طبیعی و پاتولوژیک، منجر به یک تابع تمایز خطی می‌گردد. بنابراین در صورتی که کلاس‌ها در فضای ویژگی بصورت خطی قابل تفکیک نباشند، با دو مدل گوسی برای دو کلاس نمی‌توان تفکیک مناسبی صورت داد، مگر آنکه تعداد مدل‌های گوسی را بیشتر نمود، به این ترتیب با چند تابع تمایزگر خطی می‌توان فضای ویژگی را به تعداد مناسبی زیر فضا افراز نمود. و به نتایج خوبی دست یافت اینکار مستلزم خوشه‌بندی بردارها در فضای ویژگی است [۶].

۲-۳-۲- طبقه‌بندی با استفاده از درخت تصمیم

درخت‌های تصمیم، ابزاری جهت طبقه‌بندی دادگان می‌باشند. در مرحله تعلیم، درخت تصمیم ساختار خود را با استفاده از دادگان تعلیم شکل می‌دهد. در ساخت یک درخت تصمیم چند نکته حائز اهمیت است: اول انتخاب سؤالاتی که به تفکیک اجزای دادگان تعلیم می‌پردازند. برخی از الگوریتم‌ها می‌توانند بصورت خودکار سؤالات را طراحی نمایند. انتخاب سؤالات به گونه‌ای صورت می‌گیرد که حداکثر کاهش آنتروپی محقق شود. نکته بعدی نحوه توسعه و رشد درخت تصمیم به منظور نمایش بهتر دادگان تعلیم است. توسعه درخت تصمیم تا جایی انجام می‌گیرد که شرط توقف تقسیم گره‌ها تحقق یابد. این شرط تعداد کمتر از حد آستانه عناصر موجود در آن گره و یا کاهش آنتروپی زیر حد آستانه موردنظر می‌باشد. گره‌هایی که دیگر قابل تفکیک نیستند به گره‌های نهایی موسومند. بعد از تعلیم و ساخت درخت تصمیم، هرس آن به منظور تعمیم دادن آن به دادگان غیر از تعلیم صورت می‌گیرد، چراکه ممکن است درخت تصمیم بیش از حد به دادگان تعلیم تطابق یابد و خاصیت تعمیم نداشته باشد، که در جهت رفع این اشکال هرس درخت تصمیم در فاز نهایی صورت می‌گیرد. درخت‌های تصمیم توانایی مدلسازی مرزهای تصمیم‌گیری غیرخطی و پیچیده را دارند. یک درخت تصمیم

بعد از آنکه ویژگی‌های مرتبط با نوع کار انتخاب و استخراج شدند، مساله انتخاب طبقه‌بندی کننده مناسب از اهمیت ویژه‌ای برخوردار است. چنانچه ساختار داده‌ها برایمان مشخص باشد، انتخاب طبقه‌بندی کننده مناسب راحتتر خواهد بود. برای طبقه‌بندی ویژگی‌های استخراج شده به دو دسته طبیعی یا پاتولوژیک از دو نوع طبقه‌بندی کننده بیزین با مدل تابع چگالی احتمال شرطی گوسی و درخت تصمیم استفاده نمودیم. باید اشاره کنیم با توجه به تفاوت ذاتی بردارهای ویژگی برای گروه زنان و مردان، طبقه‌بندی کننده‌ها برای هرگروه بصورت جداگانه تعلیم داده‌شد. استفاده از درخت تصمیم با توجه به اینکه قابلیت مدلسازی غیرخطی مرزهای تصمیم‌گیری را در فضای ویژگی‌ها داراست، نتایج خوبی حاصل نمود. لاکن طبقه‌بندی کننده بیزین با دو مدل گوسی بدلیل ایجاد تنها یک مرز تصمیم خطی، نتایج خوبی دربرداشت. در این بخش به توضیح مختصری از طبقه‌بندی کننده‌های مورد استفاده می‌پردازیم.

۲-۳-۱- طبقه‌بندی کننده بیزین با مدل گوسی

در چنین طبقه‌بندی کننده‌ای سعی داریم با توجه به بردار ویژگی محتمل‌ترین کلاس را برای چنین برداری انتخاب نماییم به عبارتی با توجه به فرمول زیر کلاس k مربوط به بردار ویژگی را انتخاب می‌نماییم.

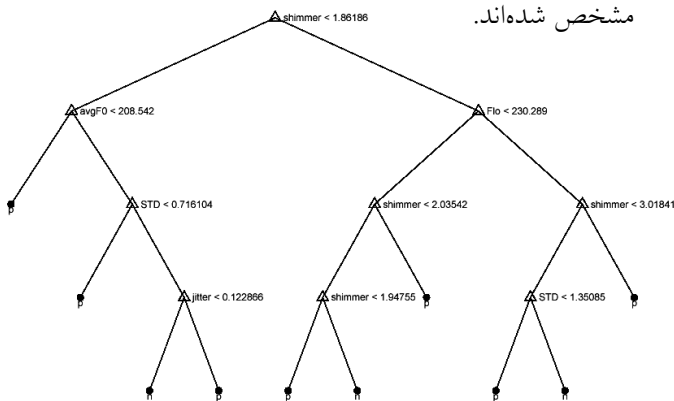
$$k = \arg \max_i p(w_i | x) \quad (6)$$

از آنجائیکه احتمال پیشین وقوع هر کلاس $p(w_i)$ و تابع چگالی احتمال شرطی بردار ویژگی به شرط کلاس $p(x | w_i)$ قابل تخمین است، می‌توانیم با توجه به قاعده بیز احتمال شرطی کلاس به شرط بردار ویژگی را پیدا نماییم.

$$p(w_i | x) = \frac{p(x | w_i)p(w_i)}{p(x)} \quad (7)$$

برای یک مساله شناسایی الگوی دو کلاسه فرمول ۶ می‌تواند به شکل زیر نوشته شود. p اندیس مربوط به کلاس بیمار و n اندیس مربوط به گروه طبیعی است.

بنابراین یک درخت برای طبقه‌بندی صدای زنان و یک درخت برای طبقه‌بندی صدای مردان ایجاد شد. در شکل ۳ بعنوان نمونه درخت تعلیم داده شده برای گروه زنان ملاحظه می‌شود. سؤال مطرح شده در بالای هر گره با توجه به یکی از ویژگی‌های استخراج شده ملاحظه می‌شود. و گره‌های نهایی با دو برچسب n به معنای طبیعی و p به معنای پاتولوژیک مشخص شده‌اند.



شکل ۳- درخت تصمیم طبقه‌بندی کننده گروه زنان

استفاده از درخت برای طبقه‌بندی سیگنال‌ها نتایج خوبی را در برداشت که در بخش بعد این نتایج را ذکر می‌نمایم.

۳- نتایج

برای تعیین میزان کارایی طبقه‌بندی‌کننده‌ها از روش ارزیابی متقابل استفاده شد. این روش ارزیابی در مواقعی که حجم دادگان برای تعلیم و تست کم باشد مناسب است [۶]. در این روش دادگان به ۱۰ قسمت تقسیم شد، که بصورت تصادفی انتخاب گردیدند، حجم داده در هر یک از ۱۰ قسمت تقریباً مساوی در نظر گرفته می‌شود. برای درخت تصمیم طبقه‌بندی، در هر یک از زیرمجموعه‌ها نسبت تعداد عناصر موجود در دو کلاس مساوی در نظر گرفته می‌شود. سپس هر یک از زیرمجموعه‌ها برای تست کنار گذاشته می‌شود و درخت برای باقیمانده داده (۹۰٪ بقیه) تعلیم می‌یابد. سپس در هر بار تعلیم و تست تابع هزینه طبقه‌بندی نادرست محاسبه می‌شود و در نهایت تابع هزینه میانگین و انحراف معیار تابع هزینه، با توجه به ۱۰ بار تعلیم و تست محاسبه می‌شود. مقدار تابع هزینه درخت مساوی مجموع هزینه گره‌ها ضربدر احتمال رسیدن به آن گره می‌باشد. و هزینه هر گره مساوی تعداد

مثالی از یک فرآیند تصمیم‌گیری و طبقه‌بندی چندمرحله‌ای است. این ابزار بجای آنکه تمام مشخصه‌ها را برای طبقه‌بندی بطور همزمان مورد نظر قرار دهد، زیرمجموعه‌های مختلفی از ویژگی‌ها را در سطوح مختلف درخت بررسی می‌نماید. برای دادگان‌های با ساختار پیچیده که مطمئن هستیم مرزهای تصمیم در فضای ویژگی‌ها برای طبقه‌بندی غیرخطی هستند، و درعین حال می‌توان نواحی تصمیم را با مجموعه‌ای از نواحی ساده‌تر تخمین زد از درخت تصمیم استفاده می‌کنیم، بعلاوه استفاده از درخت تصمیم برای طبقه‌بندی ساختار دادگان را برایمان روشن‌تر می‌نماید [۷].

در گره آغازین یکسری داده برچسب‌خورده جهت تعلیم در اختیار الگوریتم رشد درخت قرار می‌گیرد و درخت با توجه به معیار کاهش آنتروپی بهترین سؤال را با توجه به ویژگی‌ها برای تفکیک دادگان انتخاب می‌نماید. به عبارتی الگوریتم سؤالی را انتخاب می‌کند که خالص‌ترین دو گره در مرحله بعدی ایجاد گردند. آنتروپی هر گره پایانی t بصورت زیر محاسبه می‌گردد. Y متغیر تصادفی مورد نظر می‌باشد.

$$H_t(Y) = -\sum_i p(w_i | t) \log p(w_i | t) \quad (9)$$

آنتروپی وزن‌دار شده بصورت زیر محاسبه می‌گردد. که $P(t)$ احتمال رسیدن به گره پایانی t است.

$$\bar{H}_t(Y) = H_t(Y)P(t) \quad (10)$$

اختلاف آنتروپی موضعی قبل و بعد از تفکیک داده‌ها به دو گره جدید بصورت زیر محاسبه می‌شود.

$$\Delta \bar{H}_t(q) = \bar{H}_t(Y) - (\bar{H}_t(Y|q) + \bar{H}_t(Y|q)) = \bar{H}_t(Y) - \bar{H}_t(Y|q) \quad (11)$$

میزان کاهش آنتروپی برای هر سؤال محاسبه می‌گردد و در نهایت سؤالی برای تفکیک دادگان برگزیده می‌شود که بیشترین کاهش آنتروپی موضعی را ایجاد نماید.

$$q^* = \arg \max_q (\Delta \bar{H}_t(q)) \quad (12)$$

آنتروپی نهایی درخت بر اساس فرمول زیر محاسبه می‌گردد.

$$\bar{H}(Tree) = \sum_t \bar{H}_t(Y) \quad (13)$$

به منظور طبقه‌بندی سیگنال‌ها به دو دسته طبیعی و پاتولوژیک از درخت تصمیم استفاده شد. برای هر یک از گروه زنان و مردان فرآیند تعلیم درخت بصورت جداگانه انجام شد.

دست یافته‌اند. درخت تصمیم و شبکه عصبی قابلیت مرزبندی غیرخطی در فضای ویژگی‌ها را دارا هستند. اما استفاده از درخت تصمیم این مزیت را دارد که ساختار داده‌ها در فضای ویژگی‌ها را برایمان واضح‌تر می‌نماید، و می‌توان از نحوه شکل‌گیری درخت تصمیم به میزان اهمیت ویژگی‌ها برای طبقه‌بندی پی برد.

سپاسگذاری

از مسؤولین آزمایشگاه زبان‌شناسی دانشکده ادبیات و علوم انسانی دانشگاه تهران به ویژه دکتر محمود بیجن‌خان بدلیل دراختیار گذاشتن دادگان لازم برای انجام این تحقیق، بسیار سپاسگذاریم.

مراجع

- [1] M.N. Vieira, F.R. McInnes, M.A. Jack, "On the influence of laryngeal pathologies on acoustic and electroglottographic jitter measures,". *J. Acoustical Society of America*, Vol. 111, No. 2, pp.1045 -55, February, 2002.
- [2] C. H. Espinosa, M. F. Redondo, "Diagnosis of Vocal and Voice Disorders by the Speech Signal,". *Proc. Int. conf. on neural networks*. Como, Italy, Vol.4, p . 4253, July 24-27,2000 .
- [3] A. Fourcin, "Patterns, Structure and Measurement in Voice Pathology,". *IEEE Trans. On Biomedical Engineering*, Vol. 45, No. 3, pp. 300-312, March, 2000
- [4] K. Umapathy, S. Krishnan, V. Parsa and D. Jamieson, "Time-Frequency Modeling and Classification of Pathological Voices,". *Proc. 2nd Joint EMBS/BMES Conf.* Houston, Texas, U.S.A, pp.116 -117, October 23-26,2002 .
- [5] B. R. Gerratt, J. Kreimann and N. A. Barroso, "Comparing Internal and External Standards in Voice Quality Judgments,". *J. of Speech and Hearing Research*, Vol.36 , pp.14 -20,1993 .
- [6] A.Webb, *Statistical Pattern Recognition*. Arnold publishers, pp.193 -213,2003 .
- [7] J. Ngan, *Decision Trees for Data Classification Based on Information Theory*, Master of Science Project, December 10,2000 .

طبقه‌بندی‌های اشتباه دیده‌شده در آن گره می‌باشد. در جداول زیر نتایج ارزیابی برای هر یک از درخت‌های تصمیم گروه زنان و مردان با تعداد گره‌های نهایی متفاوت (که با توجه به حد آستانه هرس ایجاد می‌گردند) ملاحظه می‌شوند.

جدول ۳- نتایج ارزیابی با روش Cross-Validation

نتایج تست درخت تصمیم گروه زنان		
تعداد گره‌های نهایی	انحراف معیار تابع هزینه	میانگین تابع هزینه بر اساس طبقه‌بندی اشتباه
۱۰	۰/۰۱۴۴	۰/۰۶۶۷
۸	۰/۰۱۴۴	۰/۰۶۶۷
نتایج تست درخت تصمیم گروه مردان		
تعداد گره‌های نهایی	انحراف معیار تابع هزینه	میانگین تابع هزینه بر اساس طبقه‌بندی اشتباه
۱۱	۰/۰۲۰۱	۰/۱۱۰۵
۷	۰/۰۱۸۹	۰/۱۰۰۰

۴- نتیجه‌گیری

سیگنال واکه کشیده شده منبع ارزشمندی برای تشخیص نارسایی‌های صوتی است. در این تحقیق ۷ ویژگی به منظور طبقه‌بندی سیگنال واکه /a/ به کلاس‌های طبیعی و پاتولوژیک، استخراج گردید. برخی از این ویژگی‌ها از مقادیر فرکانس پایه در فریم‌های متوالی سیگنال گفتار محاسبه می‌شوند. تغییرات کوتاه مدت و بلند مدت فرکانس پایه در طول سیگنال ویژگی خوبی برای تشخیص بیماری‌های گفتاری است. لرزش‌های نامنظم اندام تولیدی در بیماران گفتاری بصورت نوسانات کوتاه مدت دامنه که حالتی شبه‌نویز دارد، بروز می‌کند. با استفاده از هر ۷ ویژگی و بهره‌گیری از یک درخت تصمیم به درصد صحت بالای ۸۹٪ برای مردان و درصد صحت ۹۳٪ برای زنان دست یافتیم. در محاسبه این نتایج، تعداد گره‌های نهایی را طوری انتخاب نمودیم که بالاترین درصد صحت طبقه‌بندی بدست آید. محققان دیگر با استفاده از شبکه عصبی و حدود ۳۰ ویژگی به درصد صحت حدود ۹۱٪