



Investigating the Phonatory and Articulatory Features in Diagnosis of Parkinson's Disease using Optimized Extracted Features by Genetic Algorithm

Najafi, Mohammad Bahador¹ / Vali, Mansour^{2*}

¹ - M.Sc. Student, Department of Electrical Engineering, K. N. Toosi University of Technology, Tehran, Iran

² - Assistant Professor, Department of Electrical Engineering, K. N. Toosi University of Technology, Tehran, Iran

ARTICLE INFO

DOI: 10.22041/IJBME.2020.116028.1530

Received: 20 October 2019

Revised: 28/1/2020-23/3/2020

Accepted: 22 May 2020

KEYWORDS

Parkinson's Disease
Phonatory Features
Articulatory Features
Genetic Algorithm

ABSTRACT

After Alzheimer, Parkinson's disease is known as the most common malignant disease of the nervous system. One of the common obstacles of this disease is the expansion of speech disorders. Since the speech production in humans is made by combination of vibration of the vocal cords (phonatory section) and then passage through the resonator in vocal tract (articulatory section), it is expected that both of these sections to be impaired. In this study, by using a noninvasive method, it is intended to diagnose Parkinson's disease from speech signal of each subject; for this purpose, using 3 sustain vowels in Persian language recorded from 48 people (27 people with Parkinson's disease and 21 healthy people), it has been evaluated to assess the extent of damage to both phonatory and articulatory sections. The phonatory model can include features such as jitter, shimmer, fundamental frequencies, opening and closing cycling time of the glottal pulses. On the other hand, for the articulatory section, features such as first, second, and third formants, zero crossing rates, MFCCs, and LPC are investigated. In this study, 38 feature categories were extracted and four statistical parameters of mean, standard deviation, skewness and kurtosis were calculated. Genetic Algorithm was used to identify the optimum features. Then, using the SVM, KNN and the Decision Tree classifiers, the optimum extracted features are classified to determine whether a person is patient or healthy. Finally for the main aim of this study, the results of both phonatory and articulatory sections were compared and challenged. The results of this study showed that phonatory features with accuracy of $96.1 \pm 1.2\%$ were more useful than articulatory section in diagnosing of Parkinson. Also it was proved that vowel /u/ has more significant role in the diagnosis of Parkinson's disease compared to other vowels by accuracy of 97.6%.

***Corresponding Author**

Address	Department of Electrical Engineering, K. N. Toosi University of Technology, Tehran, Iran			
Postal Code	163151355	Tel	+98-21-84062405	
E-Mail	mansour.vali@eetd.kntu.ac.ir		Fax	+98-21-88402066





بررسی نقش ویژگی‌های آوایی و تلفظی در تشخیص بیماری پارکینسون با استفاده از ویژگی‌های استخراج شده بهینه توسط الگوریتم ژنتیک

نجفی، محمدبهادر^۱ / ولی، منصور^{۲*}

^۱ - دانشجوی کارشناسی ارشد مهندسی پزشکی بیوالکتریک، دانشکده‌ی مهندسی برق و کامپیوتر، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران

^۲ - استادیار، گروه بیوالکتریک، دانشکده‌ی مهندسی برق و کامپیوتر، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران

مشخصات مقاله

شناسه‌ی دیجیتال: 10.22041/IJBME.2020.116028.1530

پذیرش: ۲ خرداد ۱۳۹۹

باننگری: ۱۳۹۹/۱/۴-۱۳۹۸/۱/۱/۸

ثبت در سامانه: ۲۸ مهر ۱۳۹۸

چکیده

واژه‌های کلیدی

بیماری پارکینسون بعد از آلزایمر به عنوان رایج‌ترین بیماری مخرب سیستم عصبی شناخته می‌شود. یکی از عوارض شایع این بیماری، به وجود آمدن اختلالات گفتاری است. با توجه به این که تولید گفتار در انسان شامل تولید صوت در اثر ارتعاش تارهای صوتی (بخش آوایی) و سپس عبور آن از فیلتر لوله‌ی صوتی (بخش تلفظی) است، انتظار می‌رود هر کدام از این دو بخش دچار اختلال شوند. در این تحقیق با استفاده از یک روش غیرتهاجمی و به کمک سیگنال گفتار فرد، به تشخیص بیماری پارکینسون پرداخته شده است. بدین منظور از گویش ۳ واکه‌ی کشیده‌ی زبان فارسی توسط ۴۸ نفر (۲۷ نفر مبتلا به بیماری پارکینسون و ۲۱ نفر سالم) استفاده شده است تا میزان تخریب دو بخش تلفظی و آوایی ارزیابی شود. از ویژگی‌های مرتبط با بخش آوایی تولید گفتار می‌توان به جیتر، شیمیر، فرکانس گام و طول زمانی باز و بسته شدن پالس‌های چاکنایی و از ویژگی‌های بخش تلفظی گفتار می‌توان به فرمتهای اول، دوم و سوم، نرخ عبور از صفر، MFCC و LPC اشاره کرد. در این تحقیق، در مجموع ۳۸ دسته‌ی ویژگی استخراج شده و چهار پارامتر آماری میانگین، انحراف معیار، ضریب چولگی و ضریب کشیدگی از روی آن‌ها محاسبه شده است. در ادامه از الگوریتم ژنتیک برای شناسایی ویژگی‌های بهینه استفاده شده و شناسایی بیماری پارکینسون با به کارگیری طبقه‌بندهای SVM، KNN و درخت تصمیم‌گیر انجام شده است. به عنوان شاخصه‌ی اصلی این پژوهش، نتایج مربوط به دو بخش آوایی و تلفظی مورد مقایسه و چالش قرار گرفته است. نتایج حاصل از این مطالعه نشان داده که ویژگی‌های آوایی با صحت $1/2 \pm 96/1\%$ نسبت به ویژگی‌های تلفظی در تشخیص بیماری پارکینسون نقش مفیدتری داشته و هم‌چنین واکه‌ی /او/ با میزان صحت $97/6\%$ بهترین عمل کرد را در تشخیص بیماری پارکینسون نسبت به سایر واکه‌ها داشته است.

*نویسنده‌ی مسئول

نشانی: دانشکده‌ی مهندسی برق و کامپیوتر، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران

تلفن: +۹۸-۲۱-۸۴۰۶۲۴۰۵

۱۶۳۱۵۱۳۵۵

دورنگار: +۹۸-۲۱-۸۸۴۰۲۰۶۶

mansour.vali@eed.kntu.ac.ir

پست الکترونیک



۱- مقدمه

بیماری پارکینسون^۱ (PD) برای اولین بار توسط دکتر جیمز پارکینسون در سال ۱۸۱۷ میلادی کشف شده است [۱]. این بیماری که به عنوان فلج آشفته^۲ نیز شناخته می‌شود، ناشی از تخریب گسترده‌ی بخشی از جسم سیاه یعنی بخش متراکم^۳ در مغز است که فیبرهای عصبی ترشح کننده‌ی دوپامین را به هسته‌ی دم‌دار و پوتامن می‌فرستد [۲، ۳]. این بیماری از طریق نشانه‌هایی مانند سختی بخش اعظم ساختمان عضلانی بدن، لرزش غیر ارادی نواحی درگیر حتی در زمان استراحت، مشکل جدی در آغاز حرکت، ناتوانی در حفظ موقعیت به دلیل اختلال در رفلکس‌های وضعی و نشانه‌های حرکتی شامل دیس‌فاژی (اختلال در بلع)، اختلالات تکلم، اختلال در راه رفتن و خستگی تشخیص داده می‌شود [۳].

تخریب صوتی جز اولین نشانه‌های شایع بیماری پارکینسون بوده و با توجه به اهمیت تشخیص زودهنگام این بیماری نورولوژیکی، آنالیز تغییرات ایجاد شده در سیگنال‌های صوتی، یکی از روش‌های مطرح در زمینه‌ی تفکیک بیماران پارکینسون به شمار می‌رود [۴]. اکثر مبتلایان به پارکینسون دچار اختلالات گفتاری شده و در طی آن علائمی مانند صدای تنفس در هنگام گفتار، بلندی یک‌نواخت صدا، صدای با روح خشن و کلفت، مکث نامناسب، تلفظ نادرست و لرزش صدا آشکار می‌شود [۳، ۵]. در مطالعات متعددی بی‌ثباتی در آواسازی و هم‌چنین گفتار همراه با مکث در این بیماران گزارش شده است که این امر سبب نامفهوم شدن گفتار می‌شود [۳]. تاکنون مطالعات تجربی زیادی با تمرکز بر تشخیص پارکینسون و با استفاده از الگوی گفتاری انجام شده است. اسوالد و هم‌کارانش [۶] با مطالعه روی افراد مبتلا به پارکینسون نشان دادند که استفاده از یک روش غیرتهاجمی و ارزان که در تست اولیه در بیش از ۹۰٪ مواقع موثر بوده است، به تشخیص بیماری پارکینسون کمک می‌کند. این روش شامل نظارت بر الگوهای گفتاری بیمار است که به طور خاص حرکت الگوهای زبان و فک را بررسی کرده و علاوه بر دارا بودن پتانسیل ردیابی پیش‌رفت این بیماری، اثربخشی درمان این اختلال را نیز ارزیابی می‌کند. از آن زمان تا کنون مطالعات گسترده‌ای برای تشخیص پارکینسون از روی گفتار انجام شده است. در این ارزیابی‌ها از نمونه‌های صوتی مانند ادای حروف الفبا، واژه‌های پایدار، شمارش اعداد، خواندن جملات کوتاه و خواندن جملات بلند استفاده شده است. در دو مطالعه [۷، ۸] با استفاده از پایگاه داده‌ی موجود در مرکز تحقیقاتی

UCI شامل نمونه‌هایی صوتی از گفتار افراد در قالب جملات کوتاه، شمارش اعداد و واژه‌ی کشیده [۷]، به بررسی این موضوع پرداخته شده که نقش کدام یک از نمونه‌های صوتی در تمایز افراد مبتلا به پارکینسون و افراد سالم بیش‌تر است. آن‌ها به این نتیجه رسیدند که واژه‌های کشیده شده عمل‌کرد بهتری در تمایز افراد سالم و بیمار دارند. با توجه به نقش مهم واژه‌ها در تشخیص‌های بالینی در بیماری‌های سیستم عصبی از جمله پارکینسون و آلزایمر، در مطالعات متعددی روی استفاده از واژه‌های کشیده شده و پایدار برای تشخیص مراحل اولیه‌ی بیماری تمرکز شده [۹] به طوری که در یک دهه‌ی گذشته تعداد تحقیق‌های انجام شده روی تاثیرات واژه‌های کشیده شده و واژه‌های پایدار بر تشخیص بیماری پارکینسون، افزایش چشم‌گیری داشته است. در این مطالعه نیز سعی شده است تا با استفاده از واژه‌های کشیده شده در زبان فارسی به تشخیص بیماری پارکینسون در مراحل اولیه پرداخته شود. سیگنال‌های صوتی یک فرد سالم (بالا) و یک فرد مبتلا به پارکینسون (پایین) در هنگام ادای واژه‌ی /اوو/ کشیده شده، در شکل (۱) ارائه شده است. همان‌طور که اشاره شد، فرد مبتلا به پارکینسون هنگام ادای واژه‌ی کشیده شده، دارای اختلالاتی از جمله عدم یک‌نواختی، مکث نامناسب و لرزش در صدا است. طیف زمانی-فرکانسی (اسپکتروگرام) دو فرد سالم و مبتلا به پارکینسون در شکل (۲) نشان داده شده است. در این شکل پیوستگی هارمونیکی ادای واژه در سیگنال فرد سالم نسبت به بیمار پارکینسون قابل مشاهده است.

در مطالعه‌ی داریا هم‌لینگ و هم‌کارانش [۹] با استفاده از واژه‌های کشیده شده و واژه‌های پایدار و ادای آن توسط فرد مشکوک به پارکینسون به تشخیص خودکار این بیماری پرداخته شده است. در این مطالعه از ۱۰۰ نمونه‌ی انسانی (۵۰ فرد مبتلا به پارکینسون و ۵۰ فرد سالم) استفاده شده و از این افراد تنها واژه‌ی پایدار /آ/ (به انگلیسی a) توسط میکروفن ضبط شده است. هم‌لینگ در مطالعات خود اثبات کرده که واژه‌های کشیده شده نسبت به واژه‌های پایدار نقش موثرتری در تشخیص بیماری پارکینسون دارند. در مطالعه‌ی آزادی و هم‌کارانش [۱۱] با استفاده از واژه‌ی پایدار /آ/ و استخراج ۱۲ دسته‌ی ویژگی از سیگنال گفتاری به جداسازی گروه سالم از گروه بیمار پرداخته شده است. در این آزمایش با استفاده از روش تسکین^۴ [۱۰] به بهینه‌سازی ۱۳۲ ویژگی استخراج شده از سیگنال‌های صوتی پرداخته شده و در نهایت با استفاده از

^۱ Pars Compacta^۲ Relief Method^۱ Parkinson's Disease^۲ Paralysis Agitans

همان‌طور که اشاره شد، تولید گفتار در انسان تحت تاثیر ویژگی‌های دو بخش آوایی^۲ و تلفظی^۳ گفتار است. هدف دیگر این پژوهش پاسخ به این پرسش مهم است که کدام یک از این دو بخش، در اثر بیماری پارکینسون بیش‌تر دچار اختلال شده و برای شناسایی بیماری پارکینسون از روی گفتار از ویژگی‌های کدام بخش بهتر می‌توان بهره برد.

۲- پایگاه داده

مجموعه‌ی دادگان جهت انجام این آزمایش شامل ۴۸ نفر از فارسی‌زبانان ساکن شهر تهران (۲۱ فرد سالم (۱۰ مرد و ۱۱ زن) و ۲۷ فرد مبتلا به پارکینسون (۱۵ مرد و ۱۲ زن)) است. این مجموعه در بیمارستان حضرت رسول اکرم (ص) دانشگاه تهران ضبط شده و صحت آن توسط کمیته‌ی متخصصین علوم اعصاب این بیمارستان تایید شده است [۱۲]. خصوصیات اولیه‌ی بیماران پارکینسون و افراد سالم در جدول (۲) ارائه شده است. از این ۴۸ نفر خواسته شده تا به ادای واژه‌های زبان فارسی پرداخته و اصوات آن‌ها توسط میکروفن ضبط شده است. داده‌های صوتی با نرخ نمونه‌برداری ۴۴/۱ کیلوهرتز ثبت شده و به طور کلی ۱۴۴ قطعه‌ی صوتی با میانگین طول زمانی برابر با ۱۰/۳±۲/۱ ثانیه از این افراد ضبط شده است.

جدول (۲) - مشخصات آماری افراد شرکت‌کننده در این تحقیق

جنسیت	مرد	زن	مجموع
پارکینسون	۱۵	۱۲	۲۷
سالم	۱۰	۱۱	۲۱
سن (سال)	۶۱/۶±۸/۹	۵۹/۳±۷/۳	۶۰/۶±۸/۱
مدت بیماری (سال)	۸/۶±۴/۵	۱۱/۴±۷/۶	۹/۸±۶/۱

۳- روش پیاده‌سازی

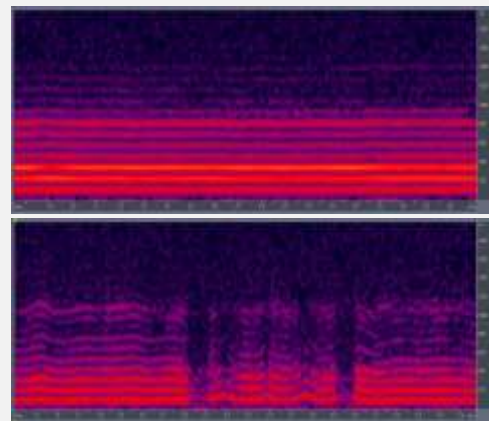
تمام شبیه‌سازی‌های این پژوهش در محیط نرم‌افزار MATLAB نسخه‌ی ۲۰۱۸b انجام شده و استخراج ویژگی‌های مربوط به سیگنال‌های صوتی افراد سالم و مبتلا به پارکینسون صرفاً توسط نرم‌افزار Open SMILE نسخه‌ی ۲۰۱۶ که مستقیماً از سایت انتشار دهنده‌ی این نرم‌افزار تهیه شده صورت گرفته است [۱۳].

پس از ثبت اصوات افراد و آماده‌سازی دادگان، به پیش‌پردازش سیگنال پرداخته شده است. به دلیل وجود نوفه‌های ناخواسته روی سیگنال اصلی، از ابزارهای حذف نوفه استفاده شده است.

طبقه‌بند^۱ ماشین بردار پشتیبان، صحت تفکیک ۹۵/۹۳٪ برای جداسازی افراد بیمار و سالم گزارش شده است. در اکثر مطالعات انجام شده تنها از یک واژه برای تشخیص این بیماری استفاده شده و این در حالی است که هر کدام از واژه‌های موجود در زبان‌های مختلف می‌تواند نقش متفاوتی در تشخیص این بیماری داشته باشد. در تحقیق حاضر این فرضیه مورد بررسی و آزمایش قرار گرفته است که کدام یک از سه واژه‌ی کشیده شده در زبان فارسی نقش موثرتری در تشخیص بیماری پارکینسون دارد. سه واژه‌ی کشیده در زبان فارسی به همراه نگارش آن‌ها در زبان انگلیسی در جدول (۱) ارائه شده است.



شکل (۱) - سیگنال‌های صوتی یک فرد سالم (بالا) و یک فرد مبتلا به پارکینسون (پایین) در هنگام ادای واژه‌ی /اوو/ کشیده در حوزه‌ی زمان



شکل (۲) - اسپکتروگرام سیگنال صوتی یک فرد سالم (بالا) و یک فرد مبتلا به پارکینسون (پایین) در هنگام ادای واژه‌ی /اوو/ کشیده

جدول (۱) - سه واژه‌ی کشیده در زبان فارسی به همراه

فونتیک آن‌ها در زبان انگلیسی

نگارش فارسی	نگارش انگلیسی	واک‌دار (V) / بی‌واک (U)
/آ/	[ā]	v
/ای/	[i]	v
/او/	[u]	v

^۳ Phonatory

^۱ Classifier

^۲ Articulatory

جدول (۳) - ویژگی‌های استخراج شده از واژه‌های کشیده

تعداد	ویژگی‌های آوایی	تعداد	ویژگی‌های تلفظی
۱	Pitch	۱۳	Formant Frequency 1:3
۲	Shimmer	۱۶	Rasta PLP CC1 : 5
۳	Jitter	۲۱	ZCR
۴	Opening Cycling Time	۲۲	MCR
۵	Closing Cycling Time	۲۳	Abs Max ZCR
۶	HNR	۲۴	Max ZCR
۷	HNR(dB)	۲۵	Min ZCR
۸	Intensity	۳۸	MFCC 1 : 13
۹	Loudness		
۱۰	RMS Energy		
۱۱	Squared Energy		
۱۲	Log Energy		

برای بخش آوایی، ویژگی‌هایی مانند پیچ گفتار^۵ (پریود گام)، جیت^۶ (تغییرات فرکانس سیگنال)، شیم^۷ (تغییرات دامنه‌ی سیگنال)، طول زمانی باز و بسته شدن تارهای صوتی^۸ در هر سیکل، شدت^۹ و بلندی^{۱۰} صوت، نسبت هارمونیک به نویز^{۱۱}، نسبت نویز به هارمونیک^{۱۲} و انرژی RMS^{۱۳} استخراج شده است. ویژگی‌های جیت^۶ و شیم^۷ از پرکاربردترین ویژگی‌های پردازش گفتار هستند. از جیت^۶ به عنوان تغییرات سیکل به سیکل^{۱۴} یا پریود گام نام برده شده [۱۴، ۱۵]، اغلب برای گزارش آن از واحد میلی‌ثانیه یا درصد استفاده شده که مقدار نرمال آن در بازه‌ی ۰/۲ تا ۱ درصد قرار داشته و برای به دست آوردن میزان تغییرات آن از رابطه‌ی (۲) استفاده می‌شود که در آن، T_i پریود گام در سیکل i -ام و N تعداد کل سیکل‌های صوت است.

$$\text{Jitter (absolute)} = \frac{1}{N-1} \sum_{i=1}^{N-1} |T_i - T_{i+1}| \quad (2)$$

از شیم^۷ به عنوان تغییرات دامنه در سیکل‌های متوالی صوت با واحد اندازه‌گیری دسی‌بل نام برده شده و با استفاده از رابطه‌ی (۳) محاسبه می‌شود که در آن A_i دامنه‌ی سیکل i -ام سیگنال و N تعداد کل سیکل‌های پریود گام در سیگنال است [۱۵].

$$\text{Shimmer (dB)} = \frac{1}{N-1} \sum_{i=1}^{N-1} |20 \log(A_{i+1}/A_i)| \quad (3)$$

در ابتدا به دلیل وجود صدای "هوف" ناشی از میکروفون، نویز فرکانس بالا مشاهده شده و برای رفع آن از یک فیلتر پایین‌گذر باترورث با مرتبه‌ی ۴ استفاده شده است. در ادامه حذف نویز DC از روی سیگنال صوتی صورت گرفته است. همچنین در برخی اصوات، به مدت چند ثانیه سکوت در ابتدا و انتها وجود داشته که آن‌ها نیز حذف شده است. در ادامه‌ی مرحله‌ی پیش‌پردازش به منظور بهادان به فرکانس‌های بالا، از فیلتر پیش‌تاکید استفاده شده است. بدین منظور یک فیلتر FIR به عنوان فیلتر بالاگذر (فیلتر پیش‌تاکید) در نظر گرفته شده است که می‌تواند با استفاده از فیلتر مرتبه‌ی اول در رابطه‌ی زیر روی سیگنال x اعمال شود.

$$y[n] = x[n] - \alpha x[n-1], \quad 0.95 < \alpha < 1 \quad (1)$$

در این رابطه، n بیان‌گر نمونه‌های سیگنال است و بر اساس این روش امکانی فراهم شده تا فرمنت‌های دوم و سوم بهتر از قبل وجود خود را در سیگنال ابراز کرده و در مرحله‌ی استخراج ویژگی، ویژگی فرمنت اول تا فرمنت سوم از این سیگنال‌ها استخراج شود.

۳-۱- استخراج ویژگی‌های صوتی

در این بخش ویژگی‌هایی به عنوان مشخصه‌های تمایزگر از سیگنال استخراج شده (شامل ۳۸ دسته‌ی ویژگی)، از هر کدام از آن‌ها پارامترهای آماری میانگین^۱، انحراف معیار^۲، ضریب چولگی^۳ و ضریب برجستگی^۴ (کشیدگی) محاسبه شده (به جز ویژگی‌های دینامیک جیت^۶ و شیم^۷ که فقط یک مقدار از کل هر گویش به دست آمده و پارامترهای آماری برای آن‌ها موضوعیت ندارد) و در کل ۱۴۶ ویژگی به دست آمده است. دلیل انتخاب این دسته‌ی ویژگی‌ها، مطالعه‌ی بسیاری از پژوهش‌های پیشین در زمینه‌ی استخراج ویژگی مربوط به اصوات افراد مبتلا به پارکینسون می‌باشد. همچنین در این مقاله سعی شده است تا تمام ویژگی‌های مشاهده شده در مقالات پیشین جمع‌آوری شده و مورد بررسی و تحلیل قرار گیرد. در ادامه، این ۳۸ دسته‌ی ویژگی به سه دسته‌ی ویژگی‌های تلفظی، آوایی و ترکیبی (ویژگی‌های تلفظی و آوایی) تقسیم‌بندی شده است که در جدول (۳) قابل مشاهده می‌باشد.

^۵ Opening and Closing of Glottal Cycles

^۶ Voice Intensity

^۷ Voice Loudness

^۸ Harmonic to Noise Ratio (HNR)

^۹ Noise to Harmonic Ratio (NHR)

^{۱۰} Root Mean Square Energy (RMS)

^{۱۱} Cycle-to-Cycle Variation

^۱ Mean

^۲ Standard Deviation

^۳ Skewness

^۴ Kurtosis

^۵ Pitch

^۶ Jitter

^۷ Shimmer

از ضریب صفرم (MFCC₀) استفاده نشده و به جای آن لگاریتم انرژی کل سیگنال استخراج می‌گردد. در این پژوهش ۱۳ ویژگی MFCC (L=۱۲) از سیگنال واکه‌ی هر گویش استخراج شده است (برای مطالعه‌ی دقیق‌تر در مورد نحوه‌ی استخراج این ویژگی به مراجع [۱۷، ۲۰] مراجعه شود).

تکنیک پیش‌بینی کننده‌ی خطی ادراکی (PLP) سیگنال گفتار در سال ۱۹۹۰ توسط هرمانسکی مطرح شده است [۲۱]. امروزه روش ضرایب PLP بعد از MFCC پرکاربردترین روش در استخراج پارامتر از سیگنال گفتار به شمار می‌رود [۲۲]. تحلیل PLP مشابه تحلیل MFCC است اما سه تفاوت اساسی با آن دارد به طوری که به جای مقیاس مل از مقیاس بارک و به جای تبدیل کسینوسی گسسته از مدل تمام‌قطب رگرسیون خودکار^۷ استفاده کرده و نحوه‌ی اعمال تابع غیرخطی بلندی صدا روی انرژی‌های استخراج شده از بانک فیلتر نیز در آن متفاوت است.

۳-۲- انتخاب ویژگی بهینه توسط الگوریتم ژنتیک

پس از استخراج ویژگی‌ها، به بحث بهینه‌سازی ویژگی‌ها توسط الگوریتم ژنتیک^۸ پرداخته شده است. الگوریتم ژنتیک یک روش تصادفی بهینه‌یابی پرکاربرد است که توسط جان هالند در سال ۱۹۹۲ طراحی شده است [۲۳].

در الگوریتم ژنتیک ارائه شده در این مقاله، ابتدا کروموزوم‌ها به عنوان جمعیت اولیه ایجاد شده، پس از ایجاد چندین جواب تصادفی برای مساله، با استفاده از تابع هزینه‌ی تعیین شده، تکامل الگوریتم با جمعیت اولیه شروع شده، سپس در هر نسل مناسب‌ترین کروموزوم‌ها انتخاب شده و این رویکرد تا زمانی ادامه پیدا کرده است که نتیجه‌ی مطلوب به دست آمده و یا تعداد تکرارها به حداکثر مقدار خود برسد [۲۴].

۳-۳- طبقه‌بند

در این پژوهش به منظور عمل‌کرد هر چه بهتر طبقه‌بند در تمایز افراد مبتلا به پارکینسون از افراد سالم و هم‌چنین تفکیک مناسب ویژگی‌های آوایی و تلفظی از سه طبقه‌بند ماشین بردار پشتیبان، درخت تصمیم‌گیر و k-نزدیک‌ترین همسایگی استفاده شده است. در ادامه عمل‌کرد این سه طبقه‌بند به اختصار توضیح داده شده است. خوانندگانی که با طبقه‌بندهای معرفی شده آشنایی لازم را دارند، می‌توانند از مطالعه‌ی این بخش صرف نظر کرده و بخش بعدی را مطالعه نمایند.

ویژگی شیمر به دلیل نمایش دامنه‌ی یک سیگنال در هر لحظه می‌تواند در تشخیص‌های بالینی از اهمیت بالایی برخوردار باشد. برای افراد سالم (افرادی که در تکلم آن‌ها هیچ‌گونه خللی وجود ندارد)، میزان تغییرات شیمر کم‌تر از ۰/۵ دسی‌بل است [۱۵] در حالی که انتظار می‌رود میزان جیتر و شیمر برای افراد مبتلا به پارکینسون بیش‌تر از افراد سالم باشد.

با تمرکز روی مفاهیم تغییرات نامنظم کانال صوتی، می‌توان به استخراج دو ویژگی طول زمانی باز و بسته شدن تارهای صوتی در سیکل‌های متوالی سیگنال پرداخت. انتظار می‌رود که این ویژگی در بیماران مبتلا به پارکینسون با افراد سالم متفاوت باشد. نسبت هارمونیک به نویز و نسبت نویز به هارمونیک نیز از مهم‌ترین ویژگی‌های موجود در سیگنال گفتار بوده که برای شناسایی میزان نویزی بودن سیگنال هارمونیک واکه‌های گفتار از آن‌ها استفاده می‌شود. این ویژگی‌ها را می‌توان بخشی از سومین گروه بزرگ اندازه‌گیری دیسفونی در نظر گرفت که هدف آن تشخیص رفتار سیگنال با استفاده از روش‌های نسبت سیگنال به نویز^۱ است [۱۶].

در بخش تلفظی، ویژگی‌هایی نظیر فرمنت اول تا سوم، ضرایب PLP^۲، ZCR^۳، MCR^۴ و MFCC^۵ استخراج شده است. در این پژوهش به دلیل انتخاب واکه‌های کشیده شده برای تشخیص پارکینسون، انتظار می‌رود که ویژگی فرمنت‌های اول، دوم و سوم نقش به‌سزایی در دینامیک تلفظ^۶ داشته باشند به طوری که با استفاده از آن‌ها بتوان تمایز مطلوبی بین فرد مبتلا به پارکینسون و فرد سالم ایجاد کرد.

ضرایب کپسترال فرکانس مل (MFCC) از رایج‌ترین ویژگی‌های استخراج شده در پردازش گفتار به شمار رفته [۱۷] که برای اولین بار در سال ۲۰۰۹ از آن برای تشخیص بیماری پارکینسون استفاده شده است [۱۸، ۱۹]. این ویژگی به محاسبه‌ی میزان انرژی سیگنال گفتار در هر باند فرکانس پرداخته و با استفاده از رابطه‌ی (۴) محاسبه می‌شود.

$$MFCC_n = \sum_{k=1}^N E_k \cos \left[\frac{n(k-0.5)\pi}{N} \right], n = 0 \dots L \quad (4)$$

رابطه‌ی (۴) نحوه‌ی استخراج ضرایب صفر تا L-ام MFCC را مشخص کرده (معمولاً L بین ۱۲ تا ۱۶) که در آن E_k میانگین انرژی k-امین فیلتر و N تعداد کل فیلترها (معمولاً بین ۲۰ تا ۳۰ فیلتر میان‌گذر به صورت یک بانک فیلتر) می‌باشد. عموماً

^۵ Mel Frequency Cepstral Coefficients

^۶ Articular Dynamics

^۷ Autoregressive

^۸ Genetic Algorithm (GA)

^۱ Signal to Noise Ratio (SNR)

^۲ Perceptual Linear Prediction (PLP)

^۳ Zero Crossing Rate (ZCR)

^۴ Mean Crossing Rate (MCR)

مورد استفاده قرار می‌گیرد. این الگوریتم یک طبقه‌بند مبتنی بر نمونه با ناظر بوده که اساس کار آن بر مقایسه‌ی میزان شباهت نمونه‌ی جدید با نمونه‌های موجود در دیتاست اولیه (آموزشی) استوار است. جهت به دست آوردن میزان شباهت دو نمونه از معیار فاصله‌ی دو نمونه استفاده می‌شود و روش‌های متفاوتی برای به دست آوردن فاصله‌ی بین دو نمونه در فضای جست‌وجو وجود دارد که از جمله‌ی آن‌ها می‌توان به فاصله‌ی اقلیدسی^۶، فاصله‌ی مینکوفسکی^۷ و فاصله‌ی منهتن^۸ اشاره کرد. برای تشخیص کلاس نمونه‌ی جدید، k عدد از نزدیک‌ترین اعضای مجموعه‌ی آموزش نسبت به نمونه‌ی جدید انتخاب شده و کلاسی که دارای بیش‌ترین عضو در این k عضو باشد به نمونه‌ی جدید تعلق می‌گیرد [۳۱].

۴- یافته‌ها و بحث

در این بخش به بررسی یافته‌های به دست آمده از این مطالعه پرداخته شده است. شایان ذکر است که این نتایج می‌تواند علاوه بر تبیین یک روش مهندسی برای تحلیل بیماری، برای تشخیص ارگان‌های تحت تاثیر بیماری پارکینسون و انتخاب پروتکل درمانی متناسب با آن نیز مورد استفاده قرار گیرد.

۴-۱- نتایج حاصل از انتخاب بهینه‌ی ویژگی توسط

الگوریتم ژنتیک

در این تحقیق با ایجاد یک ماتریس بازنمایی با ابعاد 144×146 (۱۴۴ گویش واکه‌ها با ۱۴۶ ویژگی) مرحله‌ی استخراج ویژگی انجام شده و سپس فرایند اجرای الگوریتم ژنتیک برای انتخاب ویژگی بهینه، با جمعیت اولیه‌ی ۲۰۰ کروموزوم آغاز شده است. هدف الگوریتم ژنتیک این است که با استفاده از تابع هزینه و به کمک طبقه‌بند SVM و اعتبارسنجی LOSO، میزان دقت الگوریتم به حداکثر و میزان خطا به کم‌ترین مقدار رسانده شود. خطای تابع هزینه به صورت رابطه‌ی (۵) محاسبه می‌شود.

$$Err = 1 - \max(Accuracy) \quad (5)$$

در هر مرحله، کروموزوم‌ها بر اساس تابع هزینه‌ی تعیین شده، مرتب‌سازی شده، سپس توسط تابع‌های آمیزش^۹ و جهش^{۱۰} (به ترتیب ۸۰٪ و ۱۲٪) جمعیت جدید تولید شده و مقدار ۸٪

۳-۳-۱- طبقه‌بند ماشین بردار پشتیبان

ماشین بردار پشتیبان^۱ یکی از روش‌های یادگیری با نظارت است که بر مبنای نظریه‌ی یادگیری آماری بوده و از آن برای طبقه‌بندی و رگرسیون استفاده می‌شود. این الگوریتم توسط وپنیک ارائه شده [۲۵] که بر اساس ایده‌ی حداقل‌سازی ریسک شکل گرفته است. ایده‌ی اساسی این طبقه‌بند، یافتن یک ابر صفحه‌ای بهینه به عنوان سطح تصمیم‌گیری است به گونه‌ای که حاشیه‌ی بین دو دسته را بیشینه کند. اگر داده‌ها به طور خطی جداپذیر نباشد، با یک کرنل غیرخطی به فضای با ابعاد بالاتر منتقل شده و ابر صفحه‌ی بهینه در آن فضا تعیین می‌شود [۲۶، ۲۷]. ایده‌ی اصلی SVM این است که با فرض جداپذیری خطی کلاس‌ها، ابر صفحاتی که قادر به جدا نمودن کلاس‌ها از هم باشند را به دست می‌آورد. در مساله‌هایی که داده‌ها به صورت خطی جداپذیر نباشند با استفاده از هسته‌های غیرخطی، داده‌ها به فضای با ابعاد بیش‌تر نگاشت شده تا امکان جدا کردن آن‌ها به صورت خطی در این فضای جدید فراهم گردد. در این پژوهش برای SVM از هسته‌های RBF^۲ و خطی و هم‌چنین از الگوریتم‌های اعتبارسنجی مانند LOSO^۴ استفاده شده است.

۳-۳-۲- طبقه‌بند درخت تصمیم‌گیر

درخت تصمیم‌گیری، یک نمودار ترسیمی است که روند تصمیم‌گیری یا سلسله‌ای از تصمیم‌ها را نمایش می‌دهد. این طبقه‌بند یک ابزار کمکی برای تصمیم‌گیری بوده و از یک گراف یا مدل درختی که تصمیمات و عواقب محتمل آن‌ها را نشان می‌دهد تشکیل شده است. الگوریتم‌های درخت تصمیم، از ره‌یافت‌های تصمیم‌گیری چندمرحله‌ای هستند [۲۸]. ایده‌ی اصلی در تصمیم‌گیری چندمرحله‌ای این است که تصمیم‌های سنگین و پیچیده به تصمیم‌های کوچک‌تر تبدیل شده و با ترکیب این تصمیم‌های ساده، تصمیم نهایی حاصل شود. در سال‌های اخیر به منظور ساخت درخت تصمیم، الگوریتم‌های بسیاری ارائه شده است که از مهم‌ترین آن‌ها می‌توان به CART، SPRINT و ID3 C4.5 اشاره کرد [۲۹، ۳۰].

۳-۳-۳- طبقه‌بند k-نزدیک‌ترین همسایگی

الگوریتم k-نزدیک‌ترین همسایگی^۵ (KNN) از ساده‌ترین الگوریتم‌های داده‌کاوی و طبقه‌بندی است که به طور گسترده

^۶ Euclidean Distance

^۷ Minkowski Space

^۸ Taxicab Geometry

^۹ Cross Over

^{۱۰} Mutation

^۱ Support Vector Machine (SVM)

^۲ Hyperplane

^۳ Radial Base Function (RBF)

^۴ Leave One Subject Out (LOSO)

^۵ K-Nearest Neighbors Algorithm (KNN)

صحت^۲، حساسیت^۳ و تشخیص^۴ استفاده شده که از طریق روابط زیر محاسبه می‌شوند [۳۲].

$$ACC = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \quad (۶)$$

$$SEN = \frac{T_P}{T_P + F_N} \quad (۷)$$

$$SPC = \frac{T_N}{T_N + F_P} \quad (۸)$$

در تمام روابط فوق از شاخص‌هایی استفاده شده است که برای ارزیابی مورد استفاده قرار می‌گیرند. شاخص TP بیان‌گر شخص بیماری که به درستی بیمار تشخیص داده شده، FP بیان‌گر شخص سالمی که به اشتباه بیمار تشخیص داده شده، TN بیان‌گر شخص سالمی که به درستی سالم تشخیص داده شده و FN بیان‌گر شخص بیماری که به اشتباه سالم تشخیص داده شده می‌باشد. هر یک از ویژگی‌های بهینه در سه طبقه‌بند SVM، Decision Tree و KNN مورد ارزیابی قرار گرفته است. بدین منظور هر طبقه‌بند ۱۰ بار با مقادیر اولیه‌ی متفاوت به روش LOSO تعلیم داده شده و برای هر ویژگی ۱۰ نتیجه‌ی مختلف به دست آمده که میانگین و انحراف معیار این نتایج در جدول (۴) ارائه شده است.

در رویکرد اول به بررسی ویژگی‌های آوایی پرداخته شده است. تعداد ۸ ویژگی بهینه‌ی آوایی انتخاب شده توسط GA، در ابتدا به ورودی سه طبقه‌بند مورد نظر در این پژوهش اعمال شده است. طبق جدول (۴)، برای تمام ویژگی‌های آوایی، بهترین نتیجه توسط طبقه‌بند SVM، با صحت $96/1 \pm 1/2$ درصد و حساسیت $97/4 \pm 1/5$ درصد گزارش شده است. هم‌چنین در بین ویژگی‌های بخش آوایی، ویژگی شیمیر با صحت $81/7 \pm 1/1$ ، عمل‌کرد بهتری در طبقه‌بندی از خود نشان داده است.

در رویکرد دوم، این فرایند برای ویژگی‌های تلفظی انجام شده و بهترین نتایج با استفاده از طبقه‌بند SVM با صحت $90 \pm 1/2$ ٪ و حساسیت $93/2 \pm 1/8$ ٪ به دست آمده است. در بین ویژگی‌های تلفظی نیز ویژگی MFCC دارای بالاترین صحت و دقت است. با مقایسه‌ی صحت، حساسیت و دقت ویژگی‌های آوایی و تلفظی می‌توان به این نتیجه‌ی بالینی مهم دست یافت که ویژگی‌های آوایی نسبت به ویژگی‌های تلفظی نقش مهم‌تری و تعیین‌کننده‌تری در تشخیص پارکینسون ایفا می‌کنند. هم‌چنین با شواهد به دست آمده انتظار می‌رود که میزان اختلالات به وجود آمده در بخش آوایی نسبت به بخش تلفظی

باقی‌مانده‌ی جمعیت تولیدی نیز توسط الگوریتم انتخاب بر مبنای رقابت^۱، انتخاب شده است. کروموزوم‌هایی که می‌توانند عمل‌کرد تابع هزینه را به حداکثر برسانند، باقی‌مانده و بقیه حذف شده و این روند تا زمانی که تغییرات تناسب کروموزوم‌ها کم‌تر از ۰/۰۲ شده یا الگوریتم به تعداد از پیش تعیین شده‌ای از تکرارها (در این تحقیق، ۵۰) برسد ادامه یافته است. در نهایت، هر کروموزوم با حداکثر میزان بهینگی در تابع هزینه انتخاب شده و تعداد ویژگی‌های آن کروموزوم به عنوان ویژگی‌های منتخب در نظر گرفته شده است. در این پژوهش دو آزمایش طراحی شده است. در آزمایش اول، الگوریتم ژنتیک تنها روی ویژگی‌های آوایی گفتار اعمال شده و بنابراین طول کروموزوم‌های باینری به طول ۴۲ ویژگی آوایی انتخاب شده است. پس از ۱۰ مرتبه تکرار الگوریتم ژنتیک در نهایت ۸ ویژگی از بین ۴۲ ویژگی آوایی بیش‌ترین تکرار را در بین ویژگی‌های برنده داشته است. ویژگی‌هایی چون جیتز، شیمیر، میانگین طول زمانی باز و بسته شدن تارهای صوتی، میانگین و انحراف معیار HNR، ضریب چولگی شدت انرژی صوت و انرژی RMS به عنوان ویژگی‌های بهینه شناخته شده که فهرست آن‌ها در جدول (۴) ارائه شده است. در آزمایش دوم، الگوریتم ژنتیک تنها روی ویژگی‌های تلفظی گفتار اعمال شده و بنابراین طول کروموزوم‌های باینری نیز به طول ۱۰۴ ویژگی تلفظی انتخاب شده است. پس از ۱۰ مرتبه تکرار الگوریتم در نهایت ۶ ویژگی از بین ۱۰۴ ویژگی تلفظی بیش‌ترین تکرار را در بین ویژگی‌های برنده داشته است. ویژگی‌هایی چون میانگین فرمونت اول و سوم، میانگین MFCC1,4، انحراف معیار MFCC8 و ضریب کشیدگی PLP CC1 در دسته‌ی ویژگی‌های بهینه قرار گرفته که فهرست آن‌ها نیز در جدول (۴) ارائه شده است. در پایان ترکیب ویژگی‌های بهینه‌ی آوایی و تلفظی در سه طبقه‌بند مختلف ارزیابی شده است. نتایج به دست آمده نشان می‌دهد که در بین پارامترهای آماری استخراج شده از ویژگی‌های صوتی، پارامتر میانگین با بیش‌ترین حضور در ویژگی‌های بهینه، نقش موثرتری نسبت به سایر ویژگی‌های آماری داشته و هم‌چنین در بین ویژگی‌های بهینه‌ی استخراج شده، تعداد ویژگی‌های مربوط به بخش آوایی از بخش تلفظی بیش‌تر است.

۴-۲- طبقه‌بندی و ارزیابی

یکی از نکات مهم در فرایند ارزیابی مدل‌های مختلف، آزمودن مدل است. در این پژوهش برای ارزیابی طبقه‌بندها از سه معیار

^۱ Sensitivity (True Positive Rate) (SEN)

^۲ Specificity (SPC)

^۱ Selection Tournament

^۲ Accuracy (ACC)

بیش تر باشد. در واقع عوارض شایع این بیماری، بخش آوایی گفتار انسان را بیش تر از بخش تلفظی تحت تاثیر قرار می‌دهد. در رویکرد سوم ویژگی‌های آوایی و تلفظی با هم ترکیب شده و تمام آن‌ها به ورودی طبقه‌بندها اعمال شده است. طبق جدول (۴)، طبقه‌بند SVM با صحت $1/8 \pm 97/4\%$ و حساسیت $1/2 \pm 98/7\%$ بهترین عمل کرد را برای تفکیک افراد سالم و بیمار داشته است. با توجه به این که ویژگی‌های آوایی نسبت به ویژگی‌های تلفظی عمل کرد بهتری از نظر صحت برای تشخیص پارکینسون دارند و با توجه به عمل کرد مناسب (صحت و دقت) بخش تلفظی، می‌شود ادعا کرد که نمی‌توان به طور کلی از نقش ویژگی‌های تلفظی در برابر ویژگی‌های آوایی صرف نظر نمود به طوری که طبق جدول (۴) ترکیب ویژگی‌های این دو بخش باعث بهبود هرچه بیشتر عمل کرد طبقه‌بندی شده است. در ادامه به تحلیل و بررسی این موضوع پرداخته شده است که کدام یک از سه واکه‌ی کشیده شده‌ی زبان فارسی (آ/ا، ای/ا و

ا/ا) نقش موثرتری در تشخیص بیماری پارکینسون دارند. بدین منظور با تفکیک این واکه‌ها در افراد، از هر فرد ۴۸ واکه‌ی آ/ا، ۴۸ واکه‌ی او/ا و ۴۸ واکه‌ی ای/ا مورد آزمایش قرار گرفته است. با اعمال ویژگی‌های بهینه‌ی مربوط به این واکه‌های صوتی به طبقه‌بندها و به کمک الگوریتم اعتبارسنجی LOSO این موضوع مورد بررسی قرار گرفته و نتایج آن در جدول (۵) ارائه شده است. در بررسی‌های انجام شده این نتیجه به دست آمده که واکه‌ی او/ا کشیده با صحت $97/6\%$ و حساسیت $98/4\%$ در طبقه‌بند SVM موثرترین نقش را در تشخیص بیماری پارکینسون داشته و بعد از آن واکه‌ی آ/ا بهترین عمل کرد (دقت و صحت) را داشته است. از طرف دیگر با توجه به نزدیک بودن میزان صحت این سه واکه می‌توان به این نتیجه‌ی مهم اشاره کرد که نمی‌توان به صورت قطعی واکه‌ی او/ا را به عنوان واکه‌ی برتر در تشخیص بیماری پارکینسون در نظر گرفت و از نقش کلیدی سایر واکه‌ها صرف نظر کرد.

جدول (۴) - فهرست نتایج طبقه‌بندی ویژگی‌های بهینه شده در سه دسته‌ی ویژگی‌های آوایی، تلفظی و ترکیبی

بخش	ویژگی	ماشین بردار پشتیبان			درخت تصمیم‌گیری			k-نزدیک‌ترین همسایگی		
		SPC	SEN	ACC	SPC	SEN	ACC	SPC	SEN	ACC
آوایی	Jitter	88/7±2/7	68/3±1/9	79/5±1/4	72/1±2/5	66/4±3/2	68/1±2/1	64/5±1/2	60/8±2/4	70/2±2/6
	Shimmer	95/8±2/9	72/8±2/6	81/7±1/1	80/5±2/3	70/2±4/3	75/3±1/8	70/2±2/7	62/2±3/3	82/4±2/5
	open cycling time (std)	84/6±2/6	60/9±1/3	73/2±2/5	75±2/1	68/4±2/0	70/2±3/6	69/4±2/4	61/9±2/7	75/2±1/4
	close cycling time (mean)	92/1±1/1	71/2±1/9	78/8±1/5	80/4±3/6	76/9±2/1	77/5±2/1	79±1/8	69/3±1/2	79/2±1/6
	HNR (mean)	88/3±3/19	69/5±2/7	78/7±1/6	74/7±1/5	69/1±2/9	73/6±3/1	75/6±2/7	70/2±1/6	76/1±2/2
	HNR (std)	90/2±2/5	68/2±1/6	79/5±1/1	70/3±2/8	62/9±3/5	65/3±1/2	61±4/3	57/1±2/4	65/5±1/9
	Intensity (skewness)	86±2/3	59/4±1/9	72/2±1/9	70±2/2	59/8±3/6	62/2±2/6	65/3±1/3	58/2±2/7	71/6±2/8
	RMS energy (mean)	89/2±1/2	66/8±1/7	77/8±2/6	71/3±1/6	60/4±2/4	65/2±1/7	62/7±2/7	57/5±3/6	70/7±3/2
	All Phonatory	95/8±1/9	97/4±1/5	96/1±1/2	94/2±1/8	94/2±1/8	94/7±2/6	93/9±2/8	96/9±1/2	91/2±1/7
تلفظی	Formant 1 (mean)	84/2±1/6	65/2±2/8	75/1±2/5	72/5±1/5	65/6±3/7	71/4±2/7	70/8±1/7	67/4±2/6	76/1±2/1
	Formant 3 (mean)	81/6±2/3	64/4±1/6	73/6±2/6	68/1±2/4	58/7±4/8	60/6±2/4	62/9±3/9	57/5±3/1	61/8±3/7
	MFCC 1 (mean)	90±2/2	70/9±1/1	77/2±3/7	66/4±3/9	62/5±2/9	64/3±3/4	61/1±3/1	60/9±2/4	71/3±2/4
	MFCC8 (STD)	81/4±1/9	65/5±2/6	73/7±2/3	63/8±1/8	63/4±1/7	63/4±1/7	68/2±2/4	57/2±2/5	65±1/6
	MFCC 4 (mean)	73/6±2/7	58/3±1/2	69/5±1/5	57/3±2/7	58/5±4/1	57/3±2/7	62/4±3/7	59±2/1	65/7±2/4
	PLP CC1 (Kurtosis)	76/7±1/6	57±1/9	68/3±1/6	62±2/8	62±2/8	62/1±2/3	64/5±2/7	61±3/9	66/2±3/0
	All Articulatory	88/4±1/3	93/2±1/8	90±1/2	84/3±1/6	86/2±2/6	86/2±2/6	83/8±1/8	87/9±2/4	81/1±3/4
	All Features	95/5±2/2	98/7±1/2	97/4±1/8	94/2±1/3	94/2±1/3	95/7±2/7	94/3±2/1	94/3±2/1	91/8±1/9

جدول (۵) - فهرست نتایج طبقه‌بندی افراد بیمار و سالم بر اساس واکه‌های کشیده

واکه‌های کشیده	ماشین بردار پشتیبان			درخت تصمیم‌گیری			k-نزدیک‌ترین همسایگی		
	SPC	SEN	ACC	SPC	SEN	ACC	SPC	SEN	ACC
آ/ا	95/2	93/1	96/1	91/5	90/4	94/7	92/5	93/2	93/9
ای/ا	93/7	90/0	94/6	91/2	86/4	93/7	92/8	92/6	92/3
او/ا	100	98/4	97/6	95/3	93/1	95/9	93/7	92/1	92/1

از طریق یک تماس تلفنی و با ادای واژه‌های مورد نظر، در دو دسته‌ی افراد سالم یا افراد مشکوک مبتلا به بیماری پارکینسون دسته‌بندی شود.

۶- مراجع

- [1] J. Parkinson, "An essay on the shaking palsy", London: Sherwood, Neely and Jones, 1817.
- [2] J. E. Hall, Guyton and Hall, "Textbook of Medical Physiology (Guyton Physiology)", 2018.
- [3] A.Elbaza, L.Carcaillond, S.Kab, F.Moisan, "Epidemiology of Parkinson's disease", Neuroepidemiology, 2015.
- [4] B. T. Harel, M. S. Cannizzaro, H. Cohen, N. Reilly, P. J. Snyder, "Acoustic characteristics of Parkinsonian speech: a potential biomarker of early disease progression and treatment", Journal of Neurolinguistics, 2004.
- [5] L. O. Ramig, C. Fox, and S. Sapir, "Speech treatment for Parkinson's Disease," Expert Rev. Neurotherapeutics, vol. 8, no 2, pp. 299-311, 2008.
- [6] T. Oswald, "New noninvasive tool targets Parkinson's disease", Michigan State University, Science & Technology, 2012.
- [7] E. Sakar, B. Isenkul, M. Sakar, C.O. Sertbas, A. Gurgun, F. Delil, S. Apaydin, H. Kursun, O. "Collection and Analysis of a Parkinson Speech Dataset with Multiple Types of Sound Recordings", IEEE Journal of Biomedical and Health Informatics, vol. 17(4), pp. 828-834, 2013.
- [8] N. Afza, M. Challa, J. Mungara, "Speech Processing Algorithm for Detection of Parkinson's Disease", International Journal of Engineering Research & Technology, 2013.
- [9] D. Hemmerling, J. R. Arroyave, A. Skalski, J. Gajda, E. Noth, "Automatic detection of Parkinson's disease based on modulated vowels", INTERSPEECH, San Francisco, IS161062, 2016.
- [10] K. Kira, L. A. Rendell, "A Practical Approach to Feature Selection", Computer Science, 1992.
- [11] H.R. Azadi, M.A. Khalilzade, M.R. Akbarzade-T, H.R. Kobravi, F. Rezaeitalab, A. Ziafati, A. Noei-S, N. Shahsavanpour. "Optimal Feature Selection and Comparison for Automatic Detection of Parkinson's Disease Using Speech Signal", 2017.
- [12] F. Majdinasab, S. Karkheiran, M. Soltani, N. Moradi, G. Shahidi, "Relationship Between Voice and Motor Disabilities of Parkinson's Disease", Journal of Voice, 2015.
- [13] F. Eyben, F. Wenginger, F. Gross, B. Schuller, "Recent Developments in openSMILE, the Munich Open source Multimedia Feature Extractor", Proceedings of the 21st ACM International Conference on Multimedia Barcelona, Spain, 2013.

۵- نتیجه‌گیری

در این پژوهش به بررسی نقش ترکیبی و جداگانه‌ی هر کدام از بخش‌های آوایی و تلفظی گفتار انسان برای تفکیک افراد مبتلا به بیماری پارکینسون از افراد سالم پرداخته شده است. بدین منظور از ۴۸ نفر (۲۷ نفر مبتلا به پارکینسون و ۲۱ نفر سالم) خواسته شده است تا به ادای سه واژه‌ی کشیده شده در زبان فارسی بپردازند. در ادامه ۳۸ دسته‌ی ویژگی از واژه‌های صوتی استخراج شده که به منظور اهداف مورد مطالعه در این پژوهش به سه دسته‌ی ویژگی‌های آوایی، تلفظی و ترکیبی تقسیم‌بندی شده است. سپس با به کارگیری الگوریتم ژنتیک، ویژگی‌های بهینه‌ی انتخاب شده در سه دسته‌ی آوایی، تلفظی و ترکیبی جهت طبقه‌بندی افراد به ورودی سه طبقه‌بند مورد مطالعه اعمال شده است. با استفاده از طبقه‌بند SVM، بهترین نتایج بخش آوایی با صحت $96/1 \pm 1/2$ درصد و حساسیت $97/4 \pm 1/5$ درصد و بهترین نتایج بخش تلفظی با صحت $90 \pm 1/2$ درصد و حساسیت $93/2 \pm 1/8$ درصد به دست آمده است. با مقایسه‌ی میزان صحت، حساسیت و دقت ویژگی‌های آوایی و تلفظی مشاهده می‌شود که ویژگی‌های آوایی می‌توانند نقش مهم‌تر و تعیین‌کننده‌تری در تشخیص بیماری پارکینسون داشته باشند. در حالت ویژگی‌های ترکیبی نیز طبقه‌بند SVM با صحت $97/4 \pm 1/8$ درصد و حساسیت $98/7 \pm 1/2$ درصد بهترین نتیجه را برای تفکیک افراد سالم و افراد مبتلا به پارکینسون ارائه داده است. با تحلیل و بررسی سه واژه‌ی کشیده شده‌ی /آه/، /ای/ و /او/ مشاهده شده است که واژه‌ی /او/ با صحت $97/6$ درصد، واژه‌ی /آه/ با صحت $96/1$ درصد و واژه‌ی /ای/ با صحت $94/6$ درصد به ترتیب بهترین عمل‌کرد را در طبقه‌بندی افراد بیمار و سالم داشته‌اند.

۵-۱- محدودیت‌ها و پیشنهادهای پژوهش

از محدودیت‌های این پژوهش و همچنین سایر پژوهش‌های انجام شده در زمینه‌ی تشخیص پارکینسون می‌توان به کوچک بودن حجم پایگاه داده اشاره کرد. تعداد کم دادگان یکی از چالش‌برانگیزترین مباحث در راستای سنجش عمل‌کرد الگوریتم‌های تشخیص پارکینسون می‌باشد. دادگان با حجم پایین (معمولاً کم‌تر از ۱۰۰ نمونه از افراد مبتلا به پارکینسون) در اغلب مطالعات انجام شده با نتایج طبقه‌بندی متفاوت، مورد مطالعه و بررسی قرار گرفته است. انتظار می‌رود با افزایش حجم پایگاه داده برای تحلیل و تشخیص پارکینسون بتوان از نظر صحت و دقت به نتایج بهتری در طبقه‌بندی دست یافت و در آینده یک سیستم نرم‌افزاری جامع ایجاد نمود تا فرد بیمار تنها



- [23] J. H. Holland, "Adaptation in Natural and Artificial Systems", Cambridge, MIT Press, 1992.
- [24] M. Nosrati, R. Karimi, "A Survey on Usage of Genetic Algorithms in Recent Steganography Researches", World Applied Programming, 2011.
- [25] V. Vapnik, "The nature of statistical learning theory", Springer, 1999.
- [26] K. Diamantaras, S. Kung, "Principal Component Neural Networks", Wiley, New York, 1996.
- [27] B. Schölkopf, C. J. Burges and A. J. Smola, "Advances in Kernel Methods: Support Vector Learning", Cambridge, MA: MIT Press, 1999. K.I. Diamantaras, S.Y. Kung, Principal Component Neural Networks, Wiley, New York, 1996.
- [28] C. W. Olanow, W. C. Koller, "An algorithm (decision tree) for the management of Parkinson's disease", Neurology, 1998.
- [29] L. Yongming, Y. Liuyang, W. Pin, Z. Cheng, "Classification of Parkinson's disease by Decision Tree Based Instance Selection and Ensemble Learning Algorithms", Journal of Medical Imaging and Health Informatics, 2017.
- [30] S. Bind, A. K. Tiwari, A. K. Sahani, "A Survey of Machine Learning Based Approaches for Parkinson Disease Prediction", International Journal of Computer Science and Information Technologies, 2015.
- [31] Z. Qin, A.T. Wang, C. Zhang, S. Zhang, "Cost-sensitive classification with k-nearest neighbors", Wang, M. (ed.) KSEM, vol. 8041, pp. 112-131. Springer, Heidelberg, 2013.
- [32] A. Jafari, Classification of Parkinson's disease patients using nonlinear phonetic features and Mel-frequency cepstral analysis, Biomedical Engineering: Applications, Basis and Communications, 2013.
- [14] I. R. Titze, "Principles of voice production", National Center for Voice and Speech, 2000.
- [15] M. Farrús, J. Hernando, P. Ejarque, "Jitter and Shimmer Measurement for Speaker Recognition", Signal Theory and Communications, 2000.
- [16] C. A. Ferrer, E. González, M. E. Hernández-Díaz, "Evaluation of Time and Frequency Domain-Based Methods for the Estimation of Harmonics-to-Noise-Ratios in Voice Signals", Progress in Pattern Recognition, Image Analysis and Applications, 2006.
- [17] A. Benba, A. Jilbab, A. Hammouch, "Voice analysis for detecting persons with Parkinson's disease using MFCC and VQ", Recent Advances in Electrical Engineering and Computer Science, 2014.
- [18] R. Frail, J. Godino-Llorente, N. Saenz-Lechon, V. Osma-Ruiz, C. Fredouille, "MFCC-based remote pathology detection on speech transmitted through the telephone channel", Proc Biosignals, 2009.
- [19] A. Jafari, "Classification of Parkinson's disease patients using nonlinear phonetic features and Mel-frequency cepstral analysis, Biomed", Eng. Appl. Basis Commun, Vol. 52, no. 4, 2013.
- [20] A. Tsanas, "Accurate telemonitoring of Parkinson's disease symptom severity using nonlinear speech signal processing and statistical machine learning", Thesis in partial fulfillment of the requirements for the degree of Doctor of Philosophy, 2012.
- [21] H. Hermans, "Perceptual linear predictive (PLP) analysis of speech", Acoustical Society of America, 1990.
- [22] N. Dave, "Feature Extraction Methods LPC, PLP and MFCC in Speech Recognition", International Journal for Advance Research in Engineering and Technology, 2013.