|  |  |  |
| --- | --- | --- |
|  | دانشگاه تهران  پردیس دانشکده­های فنی  دانشکده مهندسی برق و کامپیوتر |  |
|  | | |
|  | | |
| قالب گزارش پایانی پروژه­های کارشناسی  به همراه توضیحات در هر بخش  پايان‌نامه براي دريافت درجه کارشناسی  در رشته مهندسی .... گرايش ..... | | |
| **نام**  **شماره دانشجویی**  8101 | | |
| **استاد راهنما:**  **......** | | |
|  | | |
| **بهمن‌ماه 139** | | |
|  | | |



|  |
| --- |
| **تعهدنامه اصالت اثر** |
| **باسمه تعالي** |
| **اينجانب ......... تائيد مي كنم كه مطالب مندرج در اين پایان نامه حاصل تلاش اينجانب است و به دستاوردهاي پژوهشي ديگران كه در اين نوشته از آنها استفاده شده است مطابق مقررات ارجاع گرديده است. اين پایان نامه قبلاٌ براي احراز هيچ مدرك هم سطح يا بالاتر ارائه نشده است.**  **كليه حقوق مادي و معنوي اين اثر متعلق به دانشكده فني دانشگاه تهران مي باشد.**  **نام و نام خانوادگي دانشجو :**  **امضاي دانشجو :** |

**تقديم**[[1]](#footnote-1) **به: (اختياري)**

**..................................................**

**اين صفحه اختياري است و اختصاص به ذكر نام فرد يا افرادي دارد كه پايان نامه به آنها تقديم مي‌گردد. محتوای متن این صفحه به اختیار نگارنده است که باید از نظر تحصیلات تکمیلی دانشگاه نیز بلامانع باشد.**

**تشكر و قدرداني**[[2]](#footnote-2)**: (اختياري)**

..........................................................................................................

اين صفحه نیز اختياري است. عنوان اين صفحه به فاصله چهار سطر از بالا و به صورت وسط چين نوشته شده و متن اصلي نیز به فاصله دو سطر در زير آن نوشته مي‌شود. رعایت اخلاق در نگارش در این قسمت از مقاله نیز دارای اهمیت است. شما باید از کسانی که واقعاً به شما در راستای تحقیق‌تان کمک کرده‌اند، تشکر کنید. تشکر و نام بردن از آنها را لزوماً نباید به معنای تصدیق آنها از کار شما برداشت كرد. سعی کنید بصورت مشخص از کمک دیگران تشکر کنید.

**چکيده**[[3]](#footnote-3)

حداکثر یک صفحه شامل تعریف مساله و روند کنونی در زمینه موضوع پروژه و کار انجام شده توسط شما .....

**کلمات کلیدی:**

**فهرست مطالب**

[فصل 1: مقدمه و بيان مساله 1](#_Toc208953147)

[1-1- مقدمه 2](#_Toc208953148)

[1-2- تاريخچه‌ای از موضوع تحقيق 4](#_Toc208953149)

[1-2-1- معماری‌های رایج مدل‌ها در تحلیل صدای بیماران پارکینسون 4](#_Toc208953150)

[1-2-2- ویژگی های رایج در تحلیل صدای بیماران پارکینسون 6](#_Toc208953151)

[1-2-3- داده ها 10](#_Toc208953152)

[1-3- شرح مسئله تحقيق 13](#_Toc208953153)

[1-4- تعريف موضوع تحقيق 13](#_Toc208953154)

[1-5- اهداف و آرمان‌های کلی تحقيق 13](#_Toc208953155)

[1-6- روش انجام تحقیق 13](#_Toc208953156)

[1-7- ساختار پايان‌نامه 13](#_Toc208953157)

[فصل 2: مفاهيم اوليه و پیش زمینه .... 15](#_Toc208953158)

[2-1- مقدمه 16](#_Toc208953159)

[2-2- بخش اول : .... 16](#_Toc208953160)

[2-2-1- مقدمه‌ای بر ... 16](#_Toc208953161)

[2-2-2- سيستم‌های غيرقطعی 16](#_Toc208953162)

[2-2-2-1- واحد توليد ..... 16](#_Toc208953163)

[2-2-2-2- ....... 17](#_Toc208953164)

[2-2-3- ..... 17](#_Toc208953165)

[2-3- خلاصه و جمع بندی 17](#_Toc208953166)

[فصل 3: مدلسازی/شبیه سازی/طراحی /.... و ..... 18](#_Toc208953167)

[3-1- مقدمه 19](#_Toc208953168)

[3-2- روش پیشنهادی برای .... 19](#_Toc208953169)

[3-3- ابزارها/.... و ..... مورد نیاز 19](#_Toc208953170)

[3-4- معيار ارزيابی 20](#_Toc208953171)

[3-5- نتايج بدست آمده از ... 20](#_Toc208953172)

[3-6- تحليل نتايج 20](#_Toc208953173)

[3-7- خلاصه و جمع‌بندی 20](#_Toc208953174)

[فصل 4: پیاده سازی .... 21](#_Toc208953175)

[4-1- مقدمه 21](#_Toc208953176)

[4-2- خلاصه و جمع‌بندی 22](#_Toc208953177)

[فصل 5: جمع‌بندی، نتيجه‌گيری و پيشنهادها 23](#_Toc208953178)

[5-1- جمع‌بندی 24](#_Toc208953179)

[5-2- نتيجه‌گيری 24](#_Toc208953180)

[5-2-1- نوآوری / دستاوردها 24](#_Toc208953181)

[5-2-2- محدودیتها 24](#_Toc208953182)

[5-2-3- پيشنهادها 24](#_Toc208953183)

[فصل 6: مراجع 25](#_Toc208953184)

[پيوست‌ها 27](#_Toc208953185)

فهرست شکل­ها

[شکل (2-1) نمونه‌ای از انجام عمليات .... به کمک ... [16]. 14](#_Toc208952555)

فهرست جدول‌ها

[جدول 1 - مجموعه داده های بیماری پارکینسون شامل وظایف گفتاری 9](#_Toc208952556)

فهرست علائم اختصاري

|  |  |
| --- | --- |
| Dopamine Transporter scans | DAT |
| Artificial Intelligence | AI |
| Parkinson’s Disease | PD |
| Harmonic to noise ratio | HNR |
| Mel-frequency cepstral coefficients | MFCC |
| Deep Learning | DL |
| Machine Learning | ML |
| Convolutional Neural Network | CNN |
| Short-time Fourier transform | STFT |
| Linear Frequency Cepstral Coefficient | LFCC |
| Unified Parkinson's Disease Rating Scale | UPDRS |
| Support Vector Machine | SVM |
| Sparse Autoencoder | SAE |
| Random Forest | RF |
| Multilayer Perceptron | MLP |
| K-Nearest Neighbor | KNN |
| Long Short-Term Memory | LSTM |
| Healthy Control | HC |
| Group Wise Scaling | GWS |
| SHapley Additive exPlanations | SHAP |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

فصل 1

1. مقدمه و بيان مساله

در اين فصل نخست به بيان مقدمات کار، تاريخچه‌ای کوتاه از مساله تحقيق و روش کلی تحقيق پرداخته، سپس مساله و موضوع مورد بررسی در اين پایان‌نامه و اهداف و آرمان‌های کلی تحقيق را بيان می­کنید و در نهايت به ساختار پايان‌نامه‌ی پيش رو اشاره خواهيد کرد.

* 1. مقدمه

بیماری‌های عصبی از مهم‌ترین عوامل ناتوانی در سطح جهانی محسوب می‌شوند و بیماری پارکینسون یکی از شایع‌ترین این اختلالات است. بر اساس مطالعات جهانی، این بیماری بین سال‌های ۱۹۹۰ تا ۲۰۱۶ سریع‌ترین رشد را در میان اختلالات عصبی از نظر مرگ‌ومیر و ناتوانی[[4]](#footnote-4) داشته است. پارکینسون نوعی اختلال تخریب عصبی است که به‌دلیل مرگ سلول‌های مولد دوپامین در ناحیه‌ی توده‌ی سیاه[[5]](#footnote-5) مغز رخ می‌دهد. دوپامین یک انتقال‌دهنده‌ی عصبی حیاتی است که با انتقال سیگنال‌ها در مغز، حرکات بدنی روان و کنترل‌شده را هماهنگ می‌کند. این بیماری علاوه بر مشکلات حرکتی مانند سفتی عضلات[[6]](#footnote-6)، کندی حرکت[[7]](#footnote-7) و لرزش[[8]](#footnote-8)، با علائم غیر حرکتی نظیر اختلالات خواب، مشکلات گوارشی و کاهش توانایی شناختی همراه است و به‌تدریج توانایی عملکرد فرد را کاهش می‌دهد .(Armstrong & Okun, 2020)

تشخیص بیماری پارکینسون در مراحل اولیه به‌دلیل شباهت علائم آن با سایر اختلالات عصبی چالش‌برانگیز است .(Tolosa et al., 2021)بروز بالینی علائم حرکتی بیماری پارکینسون معمولاً تنها پس از از دست رفتن قابل‌توجه این سلول‌ها (۶۰ تا ۸۰ درصد) رخ می‌دهد، وضعیتی که ممکن است آسیب غیرقابل‌برگشت ایجاد شده باشد .(Valarmathi et al., 2025)میانگین زمان دستیابی به تشخیصی با دقت ۹۰ درصد می‌تواند تا ۲.۹ سال پس از بروز علائم به طول انجامد، در حالی که دقت اولیه‌ی تشخیص در پنج سال نخست ممکن است تنها حدود ۵۳ درصد باشد. افزون بر این، روش‌های تشخیصی تأییدکننده مانند اسکن انتقال‌دهنده‌ی دوپامین[[9]](#footnote-9) (DAT) یا تحلیل مایع مغزی–نخاعی می‌توانند تهاجمی، پرهزینه و زمان‌بر باشند و همین امر استفاده‌ی گسترده از آن‌ها را برای غربالگری زودهنگام محدود می‌سازد. این تأخیر در تشخیص بسیار حیاتی است، چرا که بیماری پارکینسون دارای یک مرحله‌ی پیش‌بالینی طولانی است که ممکن است تا چند دهه ادامه یابد، و درمان‌ها زمانی بیشترین اثربخشی را دارند که در مراحل اولیه آغاز شوند .(Ozbolt et al., 2022)

تحقیقات نشان داده‌اند که قریب به 90% بیماران پارکینسون دچار اختلالات صوتی یا دیسفونی[[10]](#footnote-10) هستند که این تغییرات می‌تواند از نخستین نشانه‌های بیماری باشد. این تغییرات صوتی عمدتاً به اختلالات در فوناسیون و تولید آواها مربوط می‌شود .(Fagherazzi et al., 2021) پاتوفیزیولوژی بیماری پارکینسون بر روی عضلات کنترل‌کننده‌ی حنجره، دیافراگم و زبان اثر می‌گذارد و منجر به شرایطی به نام دیس‌آرتریا[[11]](#footnote-11) می‌شود. این وضعیت با مجموعه‌ای از تغییرات صوتی بروز می‌کند که شامل موارد زیر است:

* هیپوفونیا[[12]](#footnote-12) (گفتار آهسته یا با شدت صدای پایین)
* یکنواختی در زیر و بمی و شدت صدا[[13]](#footnote-13) (صدایی یکنواخت و شبیه ربات)
* کیفیت صدای نفس‌دار یا خشن و لرزش صوتی
* تلفظ نادقیق و مکث‌ها یا تردیدهای مکرر (Momeni et al., 2025)

همچنین، این تغییرات به شکل دقیق تر شامل کاهش انرژی در فرکانس‌های بالای طیف هارمونیک و اشکالات در تلفظ دقیق حروف صدادار و بی‌صدا است که در نهایت باعث کاهش وضوح گفتار بیماران می شود. اگرچه این تغییرات صوتی در مراحل اولیه، اغلب از سوی بیماران و پزشکان نادیده گرفته‌می‌شود، اما مطالعات نشان داده‌اند که در ۷۸% از بیماران مبتلا به پارکینسون در مراحل ابتدایی، تغییرات قابل‌اندازه‌گیری در ویژگی‌های صوتی قابل مشاهده‌است .(Fagherazzi et al., 2021)

در سال‌های اخیر، هوش‌مصنوعی نقش مهمی در حوزه‌ی سلامت ایفا کرده و پیشرفت آن باعث افزایش دقت و سرعت در فرآیندهای تشخیصی شده‌است. با توسعه حسگرهای پوشیدنی و دستگاه‌های هوشمند، امکان جمع‌آوری داده‌های سلامت فراهم شده‌است. امروزه، با گسترش استفاده از گوشی‌های هوشمند و اتصال میلیاردها دستگاه به اینترنت، نظارت بر سلامت از راه دور تسهیل شده و فرصت‌های جدیدی برای بهبود خدمات پزشکی و کمک به متخصصان ایجاد شده‌است. این پیشرفت‌ها مسیر را برای روش‌های تشخیصی غیرتهاجمی، کم‌هزینه و در دسترس هموار ساخته که می‌توانند نقش مؤثری در بهبود تشخیص و مدیریت بیماری پارکینسون داشته‌باشند (Costantini et al., 2023). روش‌های هوش مصنوعی (AI) به‌طور ویژه برای تحلیل این سیگنال‌های صوتی پیچیده و تمایز بین صداهای سالم و پاتولوژیک[[14]](#footnote-14) مناسب هستند، حتی زمانی که این تغییرات هنوز برای گوش انسان قابل‌تشخیص نباشند (Ozbolt et al., 2022).

* 1. تاريخچه‌ای از موضوع تحقيق
     1. معماری‌های رایج مدل‌ها در تحلیل صدای بیماران پارکینسون

مرور ادبیات نشان می‌دهد که روندی تکاملی از الگوریتم‌های سنتی یادگیری ماشین به سمت معماری‌های پیچیده‌تر یادگیری عمیق و مدل‌های تجمیعی (Ensemble) وجود داشته است؛ به‌طوری که هر دسته از این مدل‌ها برای انواع خاصی از مجموعه‌ویژگی‌ها مناسب‌تر هستند.

#### طبقه‌بندهای سنتی یادگیری ماشین:

در مطالعاتی که از بردارهای ویژگی طراحی‌شده یا مهندسی‌شده استفاده می‌کنند، ماشین بردار پشتیبان[[15]](#footnote-15) (SVM) انتخاب غالب است؛ چراکه برای فضاهای با ابعاد بالا بسیار مؤثر بوده و به‌ویژه در شرایطی که تعداد ویژگی‌ها بیشتر از تعداد نمونه‌ها باشد، عملکرد مناسبی دارد. به‌طور مشخص، Hoq و همکاران (2021) از SVM استفاده کردند زیرا این مدل قادر است راه‌حل یکتایی ارائه دهد و در مسائل بهینه‌سازی محدب[[16]](#footnote-16) از شبکه‌های عصبی عملکرد بهتری نشان دهد. آن‌ها با ترکیب SVM و یک Sparse Autoencoder (SAE) برای استخراج ویژگی به دقت 93.5% دست یافتند(Hoq et al., 2021) .

مدل‌های متداول دیگر شامل جنگل تصادفی[[17]](#footnote-17) (RF)، نزدیک‌ترین همسایه[[18]](#footnote-18) (KNN)، XGBoost و پرسیپترون چندلایه[[19]](#footnote-19) (MLP) هستند که به‌عنوان معیارهای مرجع مهم شناخته می‌شوند. هرچند این مدل‌ها مبنای قوی و تفسیرپذیری بالایی دارند، اما در شناسایی الگوهای پیچیده و غیرخطی به اندازه یادگیری عمیق مؤثر نیستند(Hoq et al., 2021) .

#### معماری‌های یادگیری عمیق (CNN):

با حرکت به سمت استفاده از اسپکتروگرام‌ها، شبکه‌های عصبی کانولوشنی (CNNs) به استاندارد غالب تبدیل شدند. معماری‌هایی مانند AlexNet و Inception V3 اغلب در چارچوب یادگیری انتقالی به‌کار گرفته می‌شوند؛ جایی که مدل ابتدا روی یک پایگاه داده‌ی تصویری بزرگ مانند ImageNet پیش‌آموزش داده می‌شود. این رویکرد توانایی پیشینی مدل در شناسایی الگوهای بصری بنیادی (مانند لبه‌ها، بافت‌ها و اشکال) را به‌کار می‌گیرد و آن را برای طبقه‌بندی تصاویر اسپکتروگرام بسیار کارآمد می‌سازد.

این تکنیک با موفقیت توسط Iyer و همکاران (2023) (Iyer et al., 2023)و Rahmatallah و همکاران (2025) (Rahmatallah et al., 2025)به‌کار گرفته شد که پایداری آن را در داده‌های باکیفیت پایین (نظیر ضبط‌های تلفنی) نشان دادند، و همچنین توسط Zahid و همکاران (2020) (Zahid et al., 2020)که از AlexNet به‌عنوان یک استخراج‌کننده عمیق ویژگی برای دستیابی به دقت 99.7% در ترکیب با MLP استفاده کردند.

#### مدل‌های ترکیبی و تجمیعی (Hybrid & Ensemble):

از آنجا که هیچ مدلی به‌تنهایی کامل نیست، پیشرفته‌ترین پژوهش‌ها اغلب از ترکیب رویکردهای مختلف برای بهره‌گیری از نقاط قوت مکمل آن‌ها استفاده می‌کنند.

##### مدل‌های ترکیبی (Hybrid Models):

این دسته معماری‌های مختلف را در یک خط پردازشی یکپارچه ادغام می‌کنند. نمونه‌ی برجسته، معماری CNN-LSTM است که توسط Islam و همکاران (2025) ارائه شد. در این مدل، CNN ابتدا ویژگی‌های مکانی[[20]](#footnote-20) را از اسپکتروگرام استخراج کرده و سپس [[21]](#footnote-21)LSTM توالی آن‌ها را در طول زمان تحلیل می‌کند تا وابستگی‌های زمانی[[22]](#footnote-22) شناسایی شوند (Islam et al., 2025).

رویکرد نوآورانه‌ی دیگر، شبکه عصبی عمیق مبتنی بر ویژگی است که از یک SAE برای استخراج بدون‌ناظر ویژگی‌ها استفاده کرده و سپس با یک DNN به طبقه‌بندی می‌پردازد. این روش در مدل‌سازی الگوهای غیرخطی در داده‌های صوتی بسیار مؤثر گزارش شده است .(Valarmathi et al., 2025)

##### مدل‌های تجمیعی (Ensemble Models):

این مدل‌ها خروجی چندین مدل مستقل را ترکیب می‌کنند تا پایداری و دقت کلی افزایش یابد. Majda-Zdancewicz و همکاران (2021) یک مدل تجمیعی پشته‌ای[[23]](#footnote-23) طراحی کردند که خروجی بهترین SVM مهندسی ویژگی و CNN یادگیری عمیق را با هم ادغام کرد. این رویکرد ترکیبی بالاترین دقت آن‌ها (96.6%) را به دست آورد و نشان داد که ترکیب نقاط قوت پارادایم‌های مختلف مدلسازی، راهبردی بسیار کارآمد است .(Majda-Zdancewicz et al., 2021)

* + 1. ویژگی های رایج در تحلیل صدای بیماران پارکینسون

رویکردهای محاسباتی اولیه بر کمّی‌سازی علائم صوتی شناخته‌شده‌ی [[24]](#footnote-24)PD از طریق مجموعه‌ای از ویژگی‌های آکوستیکی دستی[[25]](#footnote-25) تمرکز داشتند. این ویژگی‌ها شامل معیارهای جیتر[[26]](#footnote-26) (نوسان فرکانس)، شیمر[[27]](#footnote-27) (نوسان دامنه) و نسبت هارمونیک به نویز[[28]](#footnote-28) (HNR) بودند. (Carrón et al., 2021) می توان این ویژگی ها را شاخص های اغتشاش ساده[[29]](#footnote-29) هم نامید که به مجموعه ویژگی‌هایی گفته می‌شود که میزان ناپایداری یا تغییرات ریز سیگنال صوتی را از یک چرخه ارتعاش تارهای صوتی تا چرخه بعدی اندازه‌گیری می‌کنند. مطالعات اولیه با استفاده از این ویژگی‌ها و مدل‌های یادگیری ماشین پیشگام بودند و امکان‌پذیری شناسایی PD بر اساس صدا را نشان دادند.

با این حال، این ویژگی‌های سنتی محدودیت‌های جدی دارند. آن‌ها سیگنال‌های صوتی پیچیده و پویا را به مقادیر آماری منفرد کاهش می‌دهند؛ فرآیندی که اطلاعات زمانی را از بین می‌برد و الگوهای ظریف حیاتی برای تشخیص را حذف می‌کند. علاوه بر این، ارتباط آن‌ها در مدل‌های مدرن و با کارایی بالا مورد تردید قرار گرفته است؛ Carrón و همکاران (2021) اشاره کردند که در فرآیند انتخاب ویژگی بر روی یک پایگاه‌داده کنترل‌شده، "ویژگی‌های متداول مانند جیتر، شیمر یا HNR چندان مرتبط نیستند" (Carrón et al., 2021). این یافته نشان می‌دهد که ویژگی‌های جامع‌تر اغلب قدرت تفکیک بیشتری دارند. انتقاد بنیادی‌تر این است که ویژگی‌هایی مانند جیتر و شیمر برای بخش پایدار واکه‌ها[[30]](#footnote-30) طراحی شده‌اند و در صورت کاربرد نادرست روی گفتار متصل و پویا که شامل آهنگ طبیعی و همخوان‌های بی‌صدا است، می‌توانند به نتایج نادرست منجر شوند(Ozbolt et al., 2022). این موضوع، نیاز به مجموعه ویژگی‌های انعطاف‌پذیرتر را برجسته می‌سازد.

به دلیل محدودیت‌های ویژگی‌های آکوستیک ساده، پژوهش در این حوزه به سمت ویژگی‌های طیفی حرکت کرده که هم جامع‌تر هستند و هم با فیزیولوژی سیستم شنوایی انسان همخوانی دارند. در این میان، دو روش ضرایب کپسترال فرکانس مِل (MFCC)[[31]](#footnote-31) و مل‌اسپکتروگرام[[32]](#footnote-32) به عنوان نمایش‌های کلیدی بیشتر استفاده شده‌اند.

ضرایب MFCC به دلیل توانایی در مدل‌سازی ویژگی‌های مجرای صوتی انسان، به یکی از پایه های اصلی تحلیل صوت تبدیل شده است. برخلاف شاخص های اغتشاش ساده، MFCC با شبیه‌سازی ادراک غیرخطی گوش انسان (بر اساس مقیاس مل) محاسبه می‌شود. این ویژگی‌ها به تغییرات رزونانسی مجرای صوتی که مستقیماً تحت تأثیر اختلالات گفتاری در بیماری پارکینسون قرار می‌گیرند، حساسیت بالایی دارند. تغییرات رزونانسی مجرای صوتی یعنی تغییراتی که در ویژگی‌های آکوستیکی حاصل از شکل و طول و وضعیت مجرای صوتی (شامل حلق، دهان، بینی) به وجود می‌آید و باعث تغییر محل و شدت فرکانس‌های تشدید یا همان فرمانت‌ها[[33]](#footnote-33) می‌شود. فرمانت‌ها در واقع فرکانس‌های تشدید مجرای صوتی هستند که هنگام تولید صدا توسط شکل خاص دهان، حلق و بینی ایجاد می‌شوند. شماره‌گذاری فرمانت‌ها به ترتیب از کم‌ترین فرکانس شروع می‌شود. F1 بیشتر به باز بودن دهان و ارتفاع زبان بستگی دارد. F2 بیشتر به جلو یا عقب بودن زبان مربوط است. F3 به بعد اغلب برای ویژگی‌های ظریف‌تری مثل شکل لب‌ها و لهجه مهم‌ هستند. همان‌طور که Carrón و همکاران (۲۰۲۱) اشاره کرده‌اند، بیماران پارکینسون اغلب «موقعیت نامتقارن زبان در تولید واکه‌ها و کاهش مساحت فضای واکه» را نشان می‌دهند که MFCC قادر به ثبت این پدیده است (Carrón et al., 2021). موقعیت نامتقارن زبان در تولید واکه‌ها یعنی زبان در هنگام گفتن مصوت‌ها به‌صورت متوازن در دهان قرار نمی‌گیرد و کنترل حرکتی دقیق کمتری دارد. کاهش مساحت فضای واکه به این معناست که فاصله آکوستیکی بین واکه‌های اصلی کوچک‌تر می‌شود؛ چون بیمار نمی‌تواند واکه‌ها را با وضوح و تمایز زیاد ادا کند.

یک فراتحلیل[[34]](#footnote-34) نشان داده که MFCCها «بیشترین همبستگی را با شدت آسیب‌های صوتی دارند» که اهمیت بالینی آن را تأیید می‌کند. برتری این ویژگی‌ها در مطالعات متعدد به اثبات رسیده است. به عنوان مثال، ماجدا-زدانچویچ و همکاران (۲۰۲۱) در یک مقایسه مستقیم نشان دادند که یک روش مبتنی بر مهندسی ویژگی (با استفاده از LFCC[[35]](#footnote-35)، نوعی از ضرایب کپسترال) و انتخاب ویژگی، با دقت ۹۵.۵٪ از یک شبکه عصبی کانولوشنی (CNN) با دقت ۸۶.۴٪ عملکرد بهتری دارد. (Majda-Zdancewicz et al., 2021) این نتیجه حاکی از آن است که یک مجموعه ویژگی مبتنی بر اصول بیولوژیکی تولید گفتار می‌تواند حتی از یک مدل پیچیده انتها-به-انتها[[36]](#footnote-36) نیز قدرتمندتر عمل کند. به همین دلیل، MFCC هسته مرکزی بسیاری از سیستم‌های موفق تشخیصی جدید، از جمله مدل‌های ارائه‌شده توسط اسلام و همکاران (۲۰۲۵) (Islam et al., 2025)، شن و همکاران (۲۰۲۵) (Shen et al., 2025)، و زاهد و همکاران (۲۰۲۰) (Zahid et al., 2020) را تشکیل می‌دهد.

با ظهور یادگیری عمیق[[37]](#footnote-37)، به‌ویژه شبکه ی عصبی کانولوشنی [[38]](#footnote-38)، نیازمند شیوه‌ای جدید برای نمایش سیگنال‌های صوتی به صورت تصویر بود. این امر منجر به استفاده گسترده از اسپکتروگرام‌ها[[39]](#footnote-39) شد که از طریق تبدیل فوریه کوتاه‌مدت (STFT)[[40]](#footnote-40) تولید می‌شوند و نمایش دوبعدی غنی از تغییرات محتوای فرکانسی سیگنال صوتی در طول زمان فراهم می‌کنند. قدرت اسپکتروگرام‌ها به‌ویژه در نقش آن‌ها در یادگیری انتقالی [[41]](#footnote-41)با مدل‌های یادگیری عمیق نمایان می‌شود. از آنجا که اسپکتروگرام‌ها را می‌توان به‌عنوان تصاویر در نظر گرفت، پژوهشگران قادرند از معماری‌های قدرتمند CNN از پیش آموزش‌دیده مانند AlexNet و Inception V3 بهره ببرند؛ معماری‌هایی که بر روی میلیون‌ها تصویر عمومی از پایگاه داده‌ای مانند ImageNet آموزش دیده‌اند. این پیش‌آموزش به مدل‌ها امکان می‌دهد الگوهای بنیادی بصری (مانند لبه‌ها، بافت‌ها و شکل‌ها) را بیاموزند که به‌خوبی قابل انتقال به وظیفه‌ی شناسایی الگوهای بالینی در ساختار یک اسپکتروگرام هستند. (Rahmatallah et al., 2025; Zahid et al., 2020)

اگرچه اسپکتروگرام‌های مقیاس خطی[[42]](#footnote-42) مفید هستند، مل-اسپکتروگرام‌ها به عنوان گزینه برتر برای تحلیل گفتار شناخته شده‌اند. با تغییر محور فرکانس به مقیاس مل، این نمایش‌ها باندهای فرکانسی پایین‌تر را که بیشترین اطلاعات گفتاری در آن متمرکز است، اولویت می‌دهند و در واقع سیستم شنوایی انسان را بازتاب می‌دهند. رحمت‌الله و همکاران (2025) شواهد تجربی ارائه کردند که مل-اسپکتروگرام‌ها در مقایسه با نسخه‌های خطی، "افزایش معناداری[[43]](#footnote-43)" در عملکرد طبقه‌بندی[[44]](#footnote-44) ایجاد می‌کنند (Rahmatallah et al., 2025). با توجه به اطلاعات مبتنی بر ادراک، مل-اسپکتروگرام‌ها به ورودی استاندارد برای مدل‌های یادگیری عمیق پیشرفته در شناسایی PD تبدیل شده‌اند، همان‌طور که در کارهای اسلام و همکاران (2025) (Islam et al., 2025)، رحمت‌الله و همکاران (2025) (Rahmatallah et al., 2025) و توگی و همکاران (2024) (Tougui et al., 2024) مشاهده می‌شود.

* + 1. مجموعه ی داده ها

#### واکه پایدار /a/ در مقابل گفتار پیوسته

بیشتر مطالعات این حوزه، از تولید پایدار واکه /a/ به‌عنوان وظیفه اصلی گفتاری استفاده می‌کنند. این انتخاب باعث به وجود آمدن چندین مزیت در مقابل گفتار پیوسته می شود:

* سادگی و دسترس‌پذیری: این وظیفه برای بیماران ساده است، زیرا نیازی به خواندن یا به خاطر سپردن چیزی ندارد و جمع‌آوری داده را در میان جمعیت‌های متنوع آسان و سازگار می‌کند (Carrón et al., 2021).
* تمرکز بر آواسازی: واکه‌های پایدار، اطلاعات مستقیمی درباره لرزش و سفتی ساختارهای تنفسی و گلوتال[[45]](#footnote-45) ارائه می‌دهند و مکانیسم‌های بنیادی تولید صدا را از حرکات پیچیده‌تر مفصلی جدا می‌سازند (Ozbolt et al., 2022). ساختار گلوتال به ناحیه‌ای از حنجره گفته می‌شود که شامل گلوتیس[[46]](#footnote-46)، یعنی فضای بین دو تار صوتی است.
* استقلال زبانی: از آنجا که ویژگی‌های آواسازی به تلفظ خاص زبان وابسته نیستند، مدل‌های آموزش‌دیده بر اساس آن‌ها، اغلب چندزبانه یا مستقل از زبان در نظر گرفته می‌شوند. در یک مقایسه وظایف گفتاری، سوپا و همکاران (2022) نتیجه گرفتند که واکه پایدار وظیفه گفتاری ترجیحی برای ارزیابی جهانی PD است، زیرا یک وظیفه گفتاری مستقل از زبان و فرهنگ به شمار می‌رود (Suppa et al., 2022).

#### بررسی پایگاه‌های داده عمومی

پیشرفت یادگیری ماشین (ML) برای شناسایی PD به‌شدت تحت تأثیر دسترسی به پایگاه‌های داده عمومی بوده است. با این حال، ماهیت و قالب این داده‌ها بر نوع پرسش‌های پژوهشی قابل بررسی و اعتبار نتایج به دست‌آمده تأثیر دارد.

پایگاه‌های داده اولیه و تأثیرگذار، مانند Oxford Parkinson’s Disease Detection Dataset (Little, 2007) و Parkinson’s Telemonitoring Dataset (Tsanas & Little, 2009) که در مخزن داده‌های دانشگاه کالیفرنیا، ارواین (UCI) قرار دارند، نقش اساسی در اثبات این مفهوم داشتند که ویژگی‌های صوتی می‌توانند برای طبقه‌بندی PD مورد استفاده قرار گیرند. این مجموعه‌ها معمولاً شامل داده‌های پردازش‌شده از اندازه‌گیری‌های صوتی تعداد محدودی از افراد بودند که وظایف ساده‌ای مانند کشش صدای مصوت /a/ را انجام می‌دادند. به‌طور مشابه، Parkinson’s Disease Classification Dataset (Sakar et al., 2019) از دانشگاه استانبول با ارائه گروه بزرگ‌تری از افراد، بر همان وظیفه متمرکز بود و مجموعه‌ای از ویژگی‌های چندبُعدی مشتق‌شده از تکنیک‌های پیشرفته پردازش سیگنال را در اختیار قرار داد.

یک تغییر پارادایم اساسی با معرفی مطالعه mPower رخ داد(Bot et al., 2016)؛ این پروژه از یک اپلیکیشن تلفن هوشمند برای جمع‌آوری داده‌های عظیم (از جمله نمونه‌های صوتی) از هزاران شرکت‌کننده در محیط‌های واقعی روزمره استفاده کرد. با درک محدودیت‌های پایگاه‌های تک‌زبانه و تک‌وظیفه‌ای، جامعه پژوهشی به سمت ایجاد پایگاه‌های داده غنی‌تر و متنوع‌تر حرکت کرده است تا با ارائه ی این تنوع و فایل های صوتی به جای ماتریس های ویژگی، قابلیت تعمیم مدل‌های ML ارتقا یابد. مجموعه‌هایی مانند PD-Italian(Dimauro & Girardi, 2019)، Neurovoz (اسپانیایی کاستیلین) (Mendes-Laureano et al., 2024)، PC-GITA (اسپانیایی کلمبیایی) (Orozco-Arroyave et al., 2014)، (Jaeger et al., 2019) MDVR-KCL، و UAMS (Iyer et al., 2023) شامل زبان‌های گوناگون و انواع وظایف گفتاری هستند.

جدول 1 - مجموعه داده های بیماری پارکینسون شامل وظایف گفتاری

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **نام مجموعه داده** | **منبع اصلی** | **تعداد افراد**  **(بیمار/سالم)** | **تعداد نمونه** | **نوع داده** | **وظایف اصلی گفتاری** | **هدف اصلی** |
| Oxford PD Detection | (Little, 2007) | 23 / 8 | 195 | ویژگی‌های پردازش‌شده | واکه کشیده /a/ | طبقه‌بندی |
| PD Telemonitoring | (Tsanas & Little, 2009) | 42 / 0 | 5,875 | ویژگی‌های پردازش‌شده | واکه کشیده /a/ | رگرسیون پیش‌بینی نمره UPDRS[[47]](#footnote-47) |
| Istanbul PD Classification | (Sakar et al., 2019) | 188 / 64 | 756 | ویژگی‌های پردازش‌شده | واکه کشیده /a/ | طبقه‌بندی |
| mPower Study | (Bot et al., 2016) | 1,000 / 5,000 | 65,000 | صدای خام | واکه کشیده /a/ | طبقه‌بندی و پایش |
| PD-Italian | (Dimauro & Girardi, 2019) | 28 / 37 | 1,662 | صدای خام | خواندن متن، تکرار هجا | طبقه‌بندی و تحلیل گفتار |
| Neurovoz | (Mendes-Laureano et al., 2024) | 54 / 58 | 2,903 | صدای خام | واکه‌ها، DDK، جملات، مونولوگ | طبقه‌بندی و تحلیل گفتار |
| PC-GITA | (Orozco-Arroyave et al., 2014) | 50 / 50 | 100 | صدای خام | گفتار خودانگیخته (روال‌های روزانه) | طبقه‌بندی |
| MDVR-KCL | (Jaeger et al., 2019) | 21 / 16 | 37 | صدای خام | خواندن متن | طبقه‌بندی |

* + 1. ملاحظات روش‌شناختی و سوگیری‌ها

با وجود پیشرفت در نمایش ویژگی‌ها، اعتبار یافته‌های پژوهشی وابسته به روش‌شناسی دقیق است. پژوهش‌هایی نشان داده‌اند که بسیاری از نتایج با کارایی بالا در ادبیات ممکن است به دلیل نقص‌های روش‌شناختی بزرگ‌نمایی شده باشند. Ozbolt و همکاران (2022) یک تحلیل جامع از این مسائل ارائه کردند و نشان دادند چگونه می‌توانند به یافته‌های خوش‌بینانه و غیرقابل بازتولید منجر شوند. (Ozbolt et al., 2022)

* **اعتبارسنجی و نشت داده:** یک خطای حیاتی در بسیاری از مطالعات، استفاده از اعتبارسنجی متقابل رکوردی است؛ جایی که چندین ضبط از یک فرد می‌تواند هم در مجموعه آموزش و هم در آزمون ظاهر شود. از آنجا که مدل‌ها می‌توانند ویژگی‌های خاص صدای افراد را بیاموزند، این امر به نشت داده منجر می‌شود. هم Ozbolt و همکاران (2022) و هم Momeni و همکاران (2025) بر ضرورت تقسیم سختگیرانه بر اساس فرد برای اطمینان از تعمیم‌پذیری مدل به افراد ناشناخته، و نه صرفاً شناسایی صداهای آشنا، تأکید کرده‌اند .(Momeni et al., 2025; Ozbolt et al., 2022)
* **سوگیری جمعیتی**: مجموعه داده عمومی اغلب دارای عدم‌توازن جمعیتی قابل توجه هستند؛ برای مثال، گروه PD معمولاً بسیار مسن‌تر از گروه [[48]](#footnote-48)HC است. از آنجا که سن بر صدا تأثیر می‌گذارد، یک طبقه‌بند ممکن است صرفاً بین "صداهای پیر" و "صداهای جوان" تمایز قائل شود، نه بین "PD" و "HC"، که به نتایج سوگیرانه منجر می‌شود. Ozbolt و همکاران (2022) به‌طور تجربی نشان دادند که دقت طبقه‌بندی همواره با افزایش تفاوت میانگین سنی بین گروه‌های PD و HC افزایش می‌یابد. برای کاهش این مسئله، Momeni و همکاران (2023) یک تکنیک مقیاس‌گذاری درون‌گروهی[[49]](#footnote-49) (GWS) توسعه دادند که ویژگی‌ها را در زیرگروه‌های جمعیتی استاندارد می‌کند و این کار دقت طبقه‌بندی را تا 9.5% افزایش داد .(Momeni et al., 2025; Ozbolt et al., 2022)
* **وابستگی به پایگاه داده و فقدان تعمیم‌پذیری:** مدل‌ها به ندرت بر روی پایگاه‌های داده‌ای غیر از آنچه بر روی آن آموزش دیده‌اند آزمایش می‌شوند. آزمایش‌های میان-پایگاهی انجام‌شده توسط Ozbolt و همکاران (2022) و Carrón و همکاران (2021) نشان داد که مدلی که بر روی یک پایگاه داده آموزش دیده، هنگام آزمون بر روی پایگاه دیگر عملکرد به‌مراتب ضعیف‌تری دارد. این ناسازگاری دامنه، ناشی از تفاوت‌های زبانی، تجهیزات ضبط و جمعیت‌شناسی، چالشی عمده برای تعمیم‌پذیری است (Carrón et al., 2021; Ozbolt et al., 2022)
  + 1. شفافیت و کیفیت داده در تشخیص پارکینسون

پرداختن به چالش‌های فوق برای توسعه مدل‌های قابل استفاده بالینی حیاتی است. پژوهش‌های اخیر تنها بر دقت تمرکز ندارند، بلکه بر ایجاد ابزارهای منصفانه، شفاف و مقاوم نیز تأکید دارند.

* **تفسیرپذیری:** برای مقابله با ماهیت "جعبه سیاه"[[50]](#footnote-50) یادگیری عمیق،Shen و همکاران (2025) و Momeni و همکاران (2023) از روش (SHAP) SHapley Additive exPlanations برای تفسیر تصمیمات مدل استفاده کردند. این کار الگوهای بالینی همخوان، مانند اهمیت "الگوی صدای یکنواخت با مکث‌های مکرر" در پیش‌بینی PD را آشکار ساخت و اعتماد به خروجی‌های مدل را افزایش داد .(Momeni et al., 2025; Shen et al., 2025)
* **کیفیت داده:** داده‌های واقعی اغلب پر از نویز هستند. Momeni و همکاران (2023) یک گام تشخیص ناهنجاری با استفاده از الگوریتم جنگل جداسازی برای شناسایی و حذف خودکار ضبط‌های خراب از پایگاه داده جمع‌سپاری‌شده mPower افزودند. به‌طور مشابه، Carrón و همکاران (2021) کاهش قابل توجه عملکرد را هنگام انتقال از شرایط ضبط کنترل‌شده به کنترل‌نشده نشان دادند (از 92% به 71% دقت)، که نیاز به مدل‌های مقاوم یا پروتکل‌های جمع‌آوری داده نیمه‌ کنترل‌شده را برجسته می‌کند .(Carrón et al., 2021; Momeni et al., 2025)
  1. شرح مسئله تحقيق

با وجود پیشرفت‌های چشمگیر در استفاده از یادگیری عمیق برای تشخیص بیماری پارکینسون (PD) از طریق سیگنال‌های صوتی، سه چالش اساسی همچنان باقی است که انگیزه اصلی این پژوهش را تشکیل می‌دهند:

* **پیچیدگی محاسباتی و هزینه بالا:** آموزش مدل‌های پیشرفته، به‌ویژه شبکه‌های عصبی عمیق، فرآیندی زمان‌بر و پرهزینه است که نیازمند منابع سخت‌افزاری قدرتمند می‌باشد. این موضوع، استفاده از این مدل‌ها را در محیط‌های بالینی با منابع محدود یا برای کاربردهای نیازمند پاسخ سریع، دشوار می‌سازد.
* **چالش تفسیرپذیری**[[51]](#footnote-51)**:** بسیاری از مدل‌های با دقت بالا، ماهیت جعبه سیاه دارند؛ یعنی نمی‌توان به سادگی درک کرد که مدل بر اساس کدام ویژگی‌ها تصمیم‌گیری می‌کند. این عدم شفافیت، مانعی جدی برای جلب اعتماد پزشکان و پذیرش این ابزارها در فرآیندهای تشخیص بالینی است. و همچنین ارائه ی این تفسیرپذیری باعث می شود که متوجه شویم، مدل با توجه به کدام ویژگی تصمیم می گیرد که چه طور طبقه بندی را انجام دهد.
* **عدم تعمیم‌پذیری به شرایط واقعی:** اغلب مدل‌ها با داده‌های تمیز و کنترل‌شده در محیط آزمایشگاهی آموزش داده می‌شوند. اما عملکرد آن‌ها هنگام مواجهه با داده‌های جمع‌آوری‌شده در شرایط واقعی، مانند پایگاه‌داده mPower که حاوی نویز محیطی و تنوع زیاد است، به شدت افت می‌کند. این شکاف میان عملکرد آزمایشگاهی و کارایی در دنیای واقعی، یک مسئله کلیدی در مسیر کاربردی‌سازی این فناوری است.

بنابراین، مسئله اصلی این تحقیق، یافتن یک نقطه بهینه میان دقت تشخیصی، سرعت آموزش، تفسیرپذیری و استحکام مدل در مواجهه با داده‌های واقعی است.

* 1. تعريف موضوع تحقيق

موضوع این تحقیق، طراحی، پیاده‌سازی و ارزیابی یک مدل یادگیری ماشین کارآمد برای تشخیص بیماری پارکینسون از طریق سیگنال‌های صوتی است.

این مدل با تمرکز بر استخراج و ترکیب ویژگی‌های مبتنی بر ادراک شنوایی، یعنی ضرایب کپسترال فرکانس مِل (MFCC) و مل-اسپکتروگرام، از واکه پایدار /a/ توسعه یافته است. هدف اصلی، دستیابی به مدلی است که ضمن حفظ دقت بالا، از نظر زمان آموزش و پیچیدگی محاسباتی بهینه باشد، تصمیمات آن قابل تفسیر بوده و بتواند عملکرد خود را بر روی داده‌های جمع‌آوری‌شده در شرایط واقعی (مانند mPower) اثبات کند.

* 1. اهداف و آرمان‌های کلی تحقيق

آرمان کلی: توسعه یک ابزار تشخیصی سریع، دقیق و تفسیرپذیر برای بیماری پارکینسون که بتواند به عنوان یک سیستم پشتیبان تصمیم‌گیری در سناریوهای بالینی واقعی به کار گرفته شود.

برای دستیابی به این آرمان، اهداف مشخص زیر تعریف شده‌اند:

۱. پیاده‌سازی یک خط لوله کارآمد برای پردازش صوت: استخراج همزمان ویژگی‌های MFCC و مل-اسپکتروگرام از پایگاه‌داده‌های صوتی UAMS، PD-Italian و mPower.

۲. طراحی و آموزش یک مدل بهینه: توسعه یک مدل یادگیری ماشین که بتواند با زمان آموزش کمتر، به دقتی قابل رقابت با مدل‌های عمیق و پیچیده‌تر دست یابد.

۳. ارائه تفسیرپذیری برای مدل: استفاده از تکنیک‌هایی مانند SHAP برای تحلیل و درک تصمیمات مدل و شناسایی مهم‌ترین ویژگی‌های صوتی مؤثر در تشخیص بیماری.

۴. ارزیابی استحکام مدل (Robustness): آزمون عملکرد نهایی مدل بر روی داده‌های چالش‌برانگیز و نمونه‌برداری‌شده در شرایط واقعی از پایگاه‌داده mPower.

۵. تحلیل مقایسه‌ای عملکرد: مقایسه نتایج مدل پیشنهادی (از نظر دقت، سرعت و تفسیرپذیری) با معیارهای عملکردی مدل‌های پیشرفته‌تر در ادبیات پژوهش.

* 1. روش انجام تحقیق

این تحقیق بر اساس یک متدولوژی کمی و ساختاریافته در پنج مرحله اصلی انجام خواهد شد:

۱. جمع‌آوری و پیش‌پردازش داده‌ها: داده‌های صوتی خام از سه پایگاه‌داده عمومی UAMS، PD-Italian و mPower جمع‌آوری می‌شوند. پیش‌پردازش شامل حذف نویز، نرمال‌سازی دامنه و قطعه‌بندی سیگنال‌ها برای اطمینان از کیفیت و یکنواختی داده‌ها خواهد بود.

۲. استخراج ویژگی: در این مرحله، از هر نمونه صوتی واکه پایدار /a/، دو دسته ویژگی استخراج می‌شود: (الف) مل-اسپکتروگرام‌ها به عنوان نمایش تصویری-طیفی سیگنال و (ب) بردارهای ویژگی MFCC به عنوان نمایش فشرده و مبتنی بر ادراک از طیف صدا.

۳. طراحی و آموزش مدل: یک مدل یادگیری ماشین بهینه طراحی خواهد شد. فرآیند آموزش با استفاده از اعتبارسنجی متقابل sujeto-محور (subject-wise cross-validation) انجام می‌شود تا از نشت داده و تخمین بیش از حد خوش‌بینانه عملکرد جلوگیری شود. زمان آموزش به دقت ثبت و با مدل‌های دیگر مقایسه خواهد شد.

۴. تحلیل تفسیرپذیری: پس از آموزش و اعتبارسنجی مدل، از الگوریتم SHAP برای تحلیل خروجی‌ها استفاده می‌شود. این تحلیل مشخص می‌کند که کدام ویژگی‌ها (از MFCC یا نواحی خاصی در مل-اسپکتروگرام) بیشترین تأثیر را در تشخیص بیماری داشته‌اند.

۵. ارزیابی و آزمون نهایی: در نهایت، عملکرد مدل با معیارهایی مانند دقت (Accuracy)، امتیاز F1 و AUC بر روی داده‌های تست، به‌ویژه زیرمجموعه‌ای چالش‌برانگیز از پایگاه‌داده mPower، سنجیده می‌شود تا استحکام و قابلیت تعمیم آن در شرایط واقعی ارزیابی گردد.

* 1. ساختار پايان‌نامه

در فصل دوم، شامل بررسی تعاريف اساسی مربوط به حوزه‌ی … ، مفاهيم اوليه و اجزای اساسی …. ، مروری بر پيشينه‌ی تحقيق و پيش‌زمينه‌های مورد نياز برای درک هرچه بهتر …. خواهيم داشت.

فصل سوم در برگيرنده‌ی توضیح مربوط به ساختار/سیستم/معماری/مدل/روش/... پیشنهادی و پیاده سازی/شبیه­سازی/... شده است.

در فصل چهارم در مورد روش استفاده شده/ابزارها/محيط پیاده سازی/شبیه سازی/..... و معيار ارزيابی نتايج بدست آمده صحبت خواهيم کرد. همچنين اين فصل ارائه دهنده‌ی نتايج .... خواهد بود. در اين فصل علاوه بر ارائه و تحليل نتايج، در مورد ويژگی‌های/مزایا/دستاوردها/.... صحبت خواهيم کرد.

در نهايت، در فصل پنجم، نتيجه‌گيری‌های کلی حاصل شده در اين تحقيق، پیاده سازی­ها/ نوآوری‌های/ .... انجام شده و محدوديت‌ها مورد بحث قرار می­گیرد و پيشنهادهایی برای ادامه‌ی مسير به علاقمندان اين حوزه‌ی ارائه خواهد شد.

فصل 2

1. مفاهيم اوليه و پیش زمینه ....

در فصل پيش رو مقدمات، مفاهيم اوليه و پيش‌زمينه‌هايي را که جهت درک هر چه بهتر موضوع­های مطرح شده در اين پایان‌نامه مورد نياز است، از مفاهيم مربوط به .... تا ... ارائه خواهد شد.

مقدمه

* 1. بخش اول : ....
     1. مقدمه‌ای بر ...

جهت روشن‌تر شدن نحوه آمدن شکل در متن و نحوه ارجاع به شکل و درصورتی که شکل از کار دیگری کپی شده باشد نحوه ارجاع به مرجع اصلی آن، به ‏شکل (2-1) توجه کنید.



نمونه‌ای از انجام عمليات .... به کمک ... [16].

* + 1. سيستم‌های غيرقطعی[[52]](#footnote-52)

جهت آشنایی و نحوه نوشتن پاورقی برای اولین بار که یک کلمه جدید را ترجمه می­کنید به مثال زیربخش 2-2-2 (غیرقطعی) توجه کنید.

* + - 1. واحد توليد .....

برای آشنایی با نحوه نوشتن فرمول و ارجاع به آن به مثال زیر توجه کنید: احتمال اينکه يک عدد تصادفی R توليد شده توسط ... برابر مقدار مشخص Q در اين مجموعه شود با رابطه (1) تعیین می شود [17]:

رابطه (1)

* + - 1. .......

* + 1. .....

......

* 1. خلاصه و جمع بندی

در اين فصل با مفاهيم اوليه‌ و پيش‌زمينه‌هایی که جهت درک هرچه بهتر ...........

فصل 3

1. مدلسازی/شبیه سازی/طراحی /.... و .....

فصل سوم در برگيرنده‌ی توضیح مربوط به ساختار/سیستم/معماری/مدل/روش/... پیشنهادی و پیاده سازی/شبیه­سازی/... شده است.

* 1. مقدمه

در اين فصل نخست به معرفی ......

* 1. روش پیشنهادی برای ....

جهت آشنایی با نحوه تولید جدول، نوشتن بالانویس و نحوه ارجاع به آن آشنا شوید به مثال زیر توجه کنید. .... ‏جدول (3-1) محدوده‌ی تغييرات پارامترهای در نظر گرفته شده در ..... را نشان می‌دهد.

بازه‌ی تغييرات پارامترهای مختلف در .....

|  |  |
| --- | --- |
| *Parameters* | *Ranges of parameters* |
| LFSR Bits | 3, 4, 5, 6, 7, 8, 9, 10 |
| Number of XORs | 1 to LFSR\_Bits-1 |
| LFSR initial values | 111111, 101010, Random value |
| Length of Stochastic Bit streams | 8, 16, 32, 64, 128, 256 |
| Input test values | 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 |

همانطور که در ‏جدول (3-1) نشان داده شده است، .......

* 1. ابزارها/.... و ..... مورد نیاز
  2. معيار ارزيابی[[53]](#footnote-53)

* 1. نتايج بدست آمده از ...
  2. تحليل نتايج
  3. خلاصه و جمع‌بندی

فصل سوم به طور عمده در برگيرنده‌ی ....

فصل 4

1. پیاده سازی ....

پس از ...انجام شده در فصل قبل، در اين فصل ....

* 1. مقدمه
  2. خلاصه و جمع‌بندی

......

**فصل 5**

1. جمع‌بندی، نتيجه‌گيری و پيشنهادها
   1. جمع‌بندی

در اين تحقيق در گام نخست ما ....

* 1. نتيجه‌گيری
     1. نوآوری / دستاوردها

....

* + 1. محدودیتها

...

* + 1. پيشنهادها

.....

**فصل6**

1. مراجع

مراجع

[1] A. Alaghi and J. P. Hayes, “Survey of Stochastic Computing,” *ACM Trans. Embed. Comput. Syst.*, vol. 12, no. 2s, pp. 1–19, 2013.

[2] W. Qian, X. Li, M. D. Riedel, K. Bazargan, and D. J. Lilja, “An Architecture for Fault-Tolerant Computation with Stochastic Logic,” *IEEE Transactions on Computers*, vol. 60, no. 1. pp. 93–105, 2011.

….

پيوست‌ها

پيوست الف:

**Abstract:**

*....*

**Keywords:**

|  |  |  |
| --- | --- | --- |
|  | University of Tehran |  |
| College of Engineering  School of Electrical and Computer Engineering  **Thesis Title** | | |
| A thesis submitted to the Undergraduate Studies Office  In partial fulfillment of the requirements for  The degree of …. in  ….. | | |
| **By:**  **....**  **Supervisor:**  **…..** | | |

Armstrong, M. J., & Okun, M. S. (2020). Diagnosis and treatment of Parkinson disease: a review. *Jama*, *323*(6), 548–560.

Bot, B. M., Suver, C., Neto, E. C., Kellen, M., Klein, A., Bare, C., Doerr, M., Pratap, A., Wilbanks, J., & Dorsey, E. (2016). The mPower study, Parkinson disease mobile data collected using ResearchKit. *Scientific data*, *3*(1), 1–9.

Carrón, J., Campos-Roca, Y., Madruga, M., & Pérez, C. J. (2021). A mobile-assisted voice condition analysis system for Parkinson’s disease: Assessment of usability conditions. *Biomedical engineering online*, *20*(1), 114.

Costantini, G., Cesarini, V., Di Leo, P., Amato, F., Suppa, A., Asci, F., Pisani, A., Calculli, A., & Saggio, G. (2023). Artificial intelligence-based voice assessment of patients with Parkinson’s disease off and on treatment: machine vs. deep-learning comparison. *Sensors*, *23*(4), 2293.

Dimauro, G., & Girardi, F. (2019). Italian parkinson's voice and speech.

Fagherazzi, G., Fischer, A., Ismael, M., & Despotovic, V. (2021). Voice for health: the use of vocal biomarkers from research to clinical practice. *Digital biomarkers*, *5*(1), 78–88.

Hoq, M., Uddin, M. N., & Park, S.-B. (2021). Vocal feature extraction-based artificial intelligent model for Parkinson’s disease detection. *Diagnostics*, *11*(6), 1076.

Islam, M., Akter, K., Hossain, M. A., & Dewan, M. A. A. (2025). Pd-net: Parkinson’s disease detection through fusion of two spectral features using attention-based hybrid deep neural network. *Information*, *16*(2), 135.

Iyer, A., Kemp, A., Rahmatallah, Y., Pillai, L., Glover, A., Prior, F., Larson-Prior, L., & Virmani, T. (2023). A machine learning method to process voice samples for identification of Parkinson’s disease. *Scientific Reports*, *13*(1), 20615.

Jaeger, H., Trivedi, D., & Stadtschnitzer, M. (2019). Mobile Device Voice Recordings at King’s College London (MDVR-KCL) from both early and advanced Parkinson’s disease patients and healthy controls. *Zenodo*.

Little, M. (2007). Parkinsons [Data set]. UCI Machine Learning Repository. In.

Majda-Zdancewicz, E., Potulska-Chromik, A., Jakubowski, J., Nojszewska, M., & Kostera-Pruszczyk, A. (2021). Deep learning vs feature engineering in the assessment of voice signals for diagnosis in Parkinson’s disease. *Bulletin of the Polish Academy of Sciences Technical Sciences*, e137347–e137347.

Mendes-Laureano, J., Gómez-García, J. A., Guerrero-López, A., Luque-Buzo, E., Arias-Londoño, J. D., Grandas-Pérez, F. J., & Godino-Llorente, J. I. (2024). NeuroVoz: a Castillian Spanish corpus of parkinsonian speech. *Scientific data*, *11*(1), 1367.

Momeni, N., Whitling, S., & Jakobsson, A. (2025). Interpretable parkinson’s disease detection using group-wise scaling. *IEEE Access*.

Orozco-Arroyave, J. R., Arias-Londoño, J. D., Vargas-Bonilla, J. F., Gonzalez-Rátiva, M. C., & Nöth, E. (2014). New Spanish speech corpus database for the analysis of people suffering from Parkinson's disease. Lrec,

Ozbolt, A. S., Moro-Velazquez, L., Lina, I., Butala, A. A., & Dehak, N. (2022). Things to consider when automatically detecting Parkinson’s disease using the phonation of sustained vowels: analysis of methodological issues. *Applied Sciences*, *12*(3), 991.

Rahmatallah, Y., Kemp, A. S., Iyer, A., Pillai, L., Larson-Prior, L. J., Virmani, T., & Prior, F. (2025). Pre-trained convolutional neural networks identify Parkinson’s disease from spectrogram images of voice samples. *Scientific Reports*, *15*(1), 7337.

Sakar, C. O., Serbes, G., Gunduz, A., Tunc, H. C., Nizam, H., Sakar, B. E., Tutuncu, M., Aydin, T., Isenkul, M. E., & Apaydin, H. (2019). A comparative analysis of speech signal processing algorithms for Parkinson’s disease classification and the use of the tunable Q-factor wavelet transform. *Applied Soft Computing*, *74*, 255–263.

Shen, M., Mortezaagha, P., & Rahgozar, A. (2025). Explainable artificial intelligence to diagnose early Parkinson’s disease via voice analysis. *Scientific Reports*, *15*(1), 11687.

Suppa, A., Costantini, G., Asci, F., Di Leo, P., Al-Wardat, M. S., Di Lazzaro, G., Scalise, S., Pisani, A., & Saggio, G. (2022). Voice in Parkinson's disease: a machine learning study. *Frontiers in neurology*, *13*, 831428.

Tolosa, E., Garrido, A., Scholz, S. W., & Poewe, W. (2021). Challenges in the diagnosis of Parkinson's disease. *The Lancet Neurology*, *20*(5), 385–397.

Tougui, I., Zakroum, M., Karrakchou, O., & Ghogho, M. (2024). Transformer-based transfer learning on self-reported voice recordings for Parkinson’s disease diagnosis. *Scientific Reports*, *14*(1), 30131.

Tsanas, A., & Little, M. (2009). Parkinsons telemonitoring. *UCI Machine Learning Repository*, *10*, C5ZS3N.

Valarmathi, P., Suganya, Y., Saranya, K., & Shanmuga Priya, S. (2025). Enhancing parkinson disease detection through feature based deep learning with autoencoders and neural networks. *Scientific Reports*, *15*(1), 8624.

Zahid, L., Maqsood, M., Durrani, M. Y., Bakhtyar, M., Baber, J., Jamal, H., Mehmood, I., & Song, O.-Y. (2020). A spectrogram-based deep feature assisted computer-aided diagnostic system for Parkinson’s disease. *IEEE Access*, *8*, 35482–35495.

1. Dedication [↑](#footnote-ref-1)
2. Acknowledgements- [↑](#footnote-ref-2)
3. Abstract [↑](#footnote-ref-3)
4. Mortality and disability [↑](#footnote-ref-4)
5. Substantia nigra [↑](#footnote-ref-5)
6. muscle rigidity [↑](#footnote-ref-6)
7. bradykinesia [↑](#footnote-ref-7)
8. tremors [↑](#footnote-ref-8)
9. Dopamine Transporter scans [↑](#footnote-ref-9)
10. dysphonia [↑](#footnote-ref-10)
11. dysarthria [↑](#footnote-ref-11)
12. Hypophonia [↑](#footnote-ref-12)
13. Mono-pitch and mono-loudness [↑](#footnote-ref-13)
14. Pathologic [↑](#footnote-ref-14)
15. Support vector machine [↑](#footnote-ref-15)
16. Convex optimality problems [↑](#footnote-ref-16)
17. Random forest [↑](#footnote-ref-17)
18. K-nearest neighbors [↑](#footnote-ref-18)
19. Multilayer perceptron [↑](#footnote-ref-19)
20. Spatial [↑](#footnote-ref-20)
21. Long short-term memory [↑](#footnote-ref-21)
22. Temporal dependencies [↑](#footnote-ref-22)
23. Stacked ensemble [↑](#footnote-ref-23)
24. Parkinson’s disease [↑](#footnote-ref-24)
25. Manual acoustic features [↑](#footnote-ref-25)
26. Jitter [↑](#footnote-ref-26)
27. Shimmer [↑](#footnote-ref-27)
28. Harmonic to noise ratio [↑](#footnote-ref-28)
29. Perturbation measures [↑](#footnote-ref-29)
30. Vowels [↑](#footnote-ref-30)
31. Mel-frequency cepstral coefficients [↑](#footnote-ref-31)
32. Mel-spectrogram [↑](#footnote-ref-32)
33. Formants [↑](#footnote-ref-33)
34. Meta-analysis [↑](#footnote-ref-34)
35. Linear frequency cepstral coefficient [↑](#footnote-ref-35)
36. End-to-end model [↑](#footnote-ref-36)
37. Deep learning [↑](#footnote-ref-37)
38. Convolutional Neural Network [↑](#footnote-ref-38)
39. Spectrograms [↑](#footnote-ref-39)
40. Short-time fourier transform [↑](#footnote-ref-40)
41. Transfer learning [↑](#footnote-ref-41)
42. Linear spectrograms [↑](#footnote-ref-42)
43. Statistically significant gain [↑](#footnote-ref-43)
44. Classification [↑](#footnote-ref-44)
45. Glottal [↑](#footnote-ref-45)
46. Glottis [↑](#footnote-ref-46)
47. Unified Parkinson's disease rating scale [↑](#footnote-ref-47)
48. Healthy control [↑](#footnote-ref-48)
49. Group wise scaling [↑](#footnote-ref-49)
50. Black box [↑](#footnote-ref-50)
51. Interpretability [↑](#footnote-ref-51)
52. Stochastic [↑](#footnote-ref-52)
53. Evaluation metric [↑](#footnote-ref-53)