|  |  |
| --- | --- |
| نام و نام خانوادگی | **امیرحسین پورداود – امیرغرقابی – محمد ابوذری – مهدی سلیمانی زادگان** |
| شماره دانشجویی | **810101120 – 810102217 – 810101087 - 810102176** |
| تاریخ ارسال گزارش | **۱۴۰2.11.۰4** |

|  |  |  |
| --- | --- | --- |
|  | **به نام خدا**  **دانشگاه تهران**  **دانشکده‌ مهندسی برق و کامپیوتر** |  |
| **درس یادگیری ماشین**  **گزارش پروژه پایانی** | | |

**فهرست**

[**قسمت 1**. **پیش پردازش داده ها** 5](#_Toc157098266)

[1.1. تمیزکردن داده ها 5](#_Toc157098267)

[1.1.1 Resample: 5](#_Toc157098268)

[Trim 1.1.2 5](#_Toc157098269)

[Normalize1.1.3 5](#_Toc157098270)

[1.1.4 PreemPhasis 6](#_Toc157098271)

[2.1. استخراج ویژگی داده ها 7](#_Toc157098272)

[2.1.1 MCFF 7](#_Toc157098273)

[2.1.2 Chorma\_stft 7](#_Toc157098274)

[2.1.3 Spectral\_Contrast 8](#_Toc157098275)

[2.1.4 Spectral\_Centrolid 8](#_Toc157098276)

[2.1.5 Zero\_Crossing\_rate 8](#_Toc157098277)

[2.1.6 Piptrack 8](#_Toc157098278)

[2.1.7 Harmonic 8](#_Toc157098279)

[2.1.8 Energy 8](#_Toc157098280)

[2.1.9 Beat\_track 8](#_Toc157098281)

[2.1.10 Mean Frequency 8](#_Toc157098282)

[**قسمت ۲**. **طبقه‌بندی داده ها** 9](#_Toc157098283)

[1.2. مدل های پیشنهادی 9](#_Toc157098284)

[2.2. روش های بهبود نتیجه 9](#_Toc157098285)

[3.2. نتیجه‌گیری و مقایسه 10](#_Toc157098286)

[**قسمت ۳** **– خوشه بندی داده ها** 10](#_Toc157098287)

[1.3. مدل های پیشنهادی 10](#_Toc157098288)

[2.3. روش های بهبود نتیجه 10](#_Toc157098289)

[3.3. نتیجه گیری و مقایسه 10](#_Toc157098290)

[مدل **(ASR) Automatic Speech Recognition** 11](#_Toc157098291)

**شکل‌ها**

شکل 1 Fine tuning در آموزش شبکه های عصبی **Error! Bookmark not defined.**

**جدول‌ها**

[جدول 1. عنوان جدول نمونه **Error! Bookmark not defined.**](#_Toc115453324)

# **قسمت 1**. **پیش پردازش داده ها**

* همانطورکه مید

## 1.1. تمیزکردن داده ها

پیش پردازش داده‌ها در داده‌های صوتی تأثیر قابل توجهی دارد. این فرآیند شامل مجموعه‌ای از تکنیک‌ها و الگوریتم‌ها است که به منظور بهبود کیفیت و قابلیت استفاده از داده‌های صوتی انجام می‌شود. تاثیرات اصلی پیش پردازش داده‌های صوتی عبارتند از:

### 1.1.1 Resample:

ibrosa.resample یک تابع در کتابخانه‌ی librosa است که برای بازنمونده کردن یک سیگنال صوتی به یک نرخ نمونه‌برداری مختلف استفاده می‌شود. این تابع با گرفتن یک سیگنال صوتی و نرخ نمونه‌برداری مطلوب به عنوان ورودی، سیگنال صوتی بازنمونه‌برداری شده را برمی‌گرداند. استفاده از این تابع در وظایف پردازش صوتی متداول است که نرخ نمونه‌برداری برای تحلیل یا مقایسه‌ی بعدی نیاز به تغییر دارد.

### Trim 1.1.2

تابع trim در پردازش داده‌های صوتی به معنی برش (کوتاه کردن) قسمت های ابتدایی و/یا انتهایی یک سیگنال صوتی است. این عملیات معمولاً برای حذف بخش‌های بی‌فایده یا ناکارآمد از سیگنال صوتی استفاده می‌شود و می‌تواند بهبودی در کارایی الگوریتم‌های پردازش صوتی و موارد دیگر ایجاد کند. عموماً می‌توان از تفاوت زمان شروع و پایانی یک سیگنال صوتی با توجه به معیارهای مشخصی (مانند آستانه بالا/پایین، توان، انرژی، آمارهای زمانی و فرکانسی و غیره) استفاده کرده و قسمت‌های غیرمربوط را حذف نمود. این عملیات برای حذف سکوت‌ها، نویزها و قسمت‌هایی از سیگنال صوتی که در تحلیل یا استفاده‌ی بعدی مورد نیاز نیستند، مفید می‌باشد.

### Normalize1.1.3

نرمال‌سازی یک تکنیک متداول در پیش پردازش داده‌های صوتی است که به هدف تغییر مقیاس و مقادیر سیگنال صوتی ورودی به یک محدوده استاندارد استفاده می‌شود، معمولاً بین 0 تا 1 یا -1 تا 1. این فرآیند به تطبیق مقیاس و دامنه مقادیر ویژگی‌ها یا مقادیر ورودی کمک می‌کند، که برای الگوریتم‌های یادگیری ماشین و وظایف پردازش سیگنال مفید است.

در این مسله سعی شده است داده های صوتی با استفاده از معادل سازی در با (max-min) /X-min مقایس بندی شوند

### 1.1.4 PreemPhasis

پری‌امفاسیس (Preemphasis) یک تکنیک است که در پیش پردازش داده‌های صوتی به کار می‌رود. این تکنیک برای بهبود بخش‌های خاص فرکانسی صدا و افزایش قابلیت درک و کیفیت کلی سیگنال استفاده می‌شود. هدف اصلی از پری‌امفاسیس، تاکید بر بخش‌های فرکانسی بالا و کاهش بخش‌های فرکانسی پایین می‌باشد.

در داده‌های صوتی، انرژی سیگنال معمولاً در محدوده‌ی فرکانسی پایین قرار دارد. با اعمال پری‌امفاسیس، ما از طریق تقویت بزرگی مولفه‌های فرکانسی بالا نسبت به مؤلفه‌های فرکانسی پایین، کیفیت سیگنال را بهبود می‌بخشیم. برای این کار، با اعمال فیلتر پاس‌بالا به سیگنال، فرکانس‌های بالاتر تقویت و فرکانس‌های کمتر ضعیف می‌شوند.

با اعمال پری‌امفاسیس ، می‌توانیم نسبت سیگنال به نویز را افزایش دهیم، تعادل طیفی را افزایش دهیم و عملکرد الگوریتم‌های پردازش گفتار بعدی مانند تشخیص گفتار یا شناسایی گوینده را بهبود بخشیم. اجزای فرکانس بالا تقویت شده می تواند به حفظ جزئیات مهم و کاهش تأثیر نویز یا اعوجاج در طول تجزیه و تحلیل گفتار یا انتقال کمک کند.

## 2.1. استخراج ویژگی داده ها

استخراج ویژگی‌ها در داده‌های صوتی یکی از مهمترین مراحل در پردازش سیگنال صوتی و برای دستیابی به نتایج دقیق در وظایف پردازش صوتی است. در واقع، استخراج ویژگی‌ها به معنای تبدیل سیگنال صوتی از فرمت زمانی به فرمت فضایی است، که داده‌های صوتی را قابل تحلیل و استفاده‌ی بیشتر می‌کند.

استخراج ویژگی‌ها به منظور بهبود دقت تشخیص گفتار و تشخیص صدا و همچنین بهبود عملکرد مدل‌های یادگیری ماشین مورد استفاده قرار می‌گیرد. بسته به وظایف مورد نظر، ویژگی‌های مختلفی مانند: طول پالس، نرخ تکرار و کوتاهی‌های متوسط سیگنال، انرژی و توان سیگنال، باندهای طیفی، شتاب، سرعت تغییر و ... استخراج می‌شوند. هر یک از این ویژگی‌ها، نشان دهنده‌ی خصوصیات خاصی از سیگنال صوتی هستند و می‌توانند به نتایج بهتر در وظایف تشخیص و تفکیک صدا، ترجمه صوتی و دیگر وظایف پردازش را عامل شود.

در زیر فیچر های مطرح شده در این پروژه را مطرح مینماییم

### 2.1.1 MCFF

MFCC یا مجموعه ویژگی‌های همینوسیکل فرکانسی، یک روش استخراج ویژگی از سیگنال صوتی است که معمولاً برای تشخیص و تمایز دادن بین صداها استفاده می‌شود. این روش بر اساس تجزیه سیگنال صوتی به بخش‌های مختلف با استفاده از نمونه‌های زمانی کوتاه تقسیم می‌شود. پس از تقسیم سیگنال، معیارهای آماری مانند انرژی، میانگین و واریانس برای هر بخش محاسبه شده و سپس با استفاده از تبدیل فوریه معیارهای زمانی به معیارهای فرکانسی تبدیل می‌شوند. در نهایت، ضرایب MFCC حاصل به عنوان بردار ویژگی برای تشخیص و تمایز دادن بین صداها استفاده می‌شوند.

### 2.1.2 Chorma\_stft

Chroma\_STFT یک روش استخراج ویژگی از سیگنال صوتی است که برای تشخیص و تمایز دادن بین آکوردهای موسیقی استفاده می‌شود. این روش بر اساس تبدیل فوریه کوتاه مدت زمانی (STFT) سیگنال صوتی استخراج می‌شود. در این روش، ابتدا سیگنال صوتی تقسیم می‌شود و برای هر بخش کوتاه زمانی STFT محاسبه می‌شود. سپس با استفاده از STFT، انرژی هر باند فرکانسی در هر بخش محاسبه می‌شود. در نهایت، از انرژی هر باند فرکانسی برای محاسبه ماتریس کروما استفاده می‌شود.

ماتریس کروما نشان می‌دهد که در هر بخش زمانی، آکوردهای مختلف موسیقی در چه میزان وجود دارند. این ماتریس به طور معمول از 12 ستون تشکیل شده است که هر ستون نمایانگر یک آکورد موسیقی است. ارزش هر سلول در ماتریس نشان دهنده حضور یا عدم حضور آکورد مربوطه در بخش زمانی موردنظر است. در کل، Chroma\_STFT یک روش مفید است که برای تحلیل و تمایز آکوردهای موسیقی در سیگنال صوتی استفاده می‌شود.

### 2.1.3 Spectral\_Contrast

Spectral\_Contrast یک روش استخراج ویژگی از سیگنال صوتی است که برای تشخیص تفاوت‌های طیفی در آن استفاده می‌شود. این روش بر اساس تبدیل فوریه کوتاه مدت زمانی (STFT) سیگنال صوتی استخراج می‌شود.

در این روش، ابتدا سیگنال صوتی تقسیم می‌شود و برای هر بخش کوتاه زمانی STFT محاسبه می‌شود. سپس با استفاده از STFT، طیف فرکانسی برای هر بخش محاسبه می‌شود. در Spectral\_Contrast، برای هر باند فرکانسی معیارهای مختلفی مانند میانگین و واریانس طیف در ناحیه آن باند محاسبه می‌شوند. سپس با استفاده از این معیارها، کنتراست طیفی بین باندهای مختلف محاسبه می‌شود.

مقادیر کنتراست طیفی نشان می‌دهد که در هر بخش زمانی، تفاوت‌های طیفی در باندهای مختلف برجسته است یا خیر. این ویژگی می‌تواند برای تشخیص و تمایز دادن بین سیگنال‌های صوتی با مشخصات طیفی متفاوت مانند صداهای موسیقی ، گفتار و غیره استفاده شود..

در کل، Spectral\_Contrast یک روش مفید است که برای تشخیص ویژگی‌های طیفی در سیگنال صوتی و تمایز دادن بین آنها استفاده می‌شود.

### 2.1.4 Spectral\_Centrolid

در مورد استخراج ویژگی های داده های صوتی، مرکز طیفی (Spectral Centroid) یک ویژگی است که به طور کلی استفاده می شود. این ویژگی نقطه مرکز ثقل طیف سیگنال صدا را نشان می دهد. برای محاسبه آن، از میانگین وزن داده های فرکانسی موجود در سیگنال استفاده می شود، که وزن ها توسط بردارهای میدان طیفی مربوطه تعیین می شود. مرکز طیفی اطلاعاتی درباره توزیع محتوای فرکانسی یک سیگنال صدا ارائه می دهد. مقدار بالاتر مرکز طیفی نشان دهنده این است که بیشتر از انرژی سیگنال در فرکانس های بالاتر تمرکز شده است، در حالی که مقدار پایین تر نشان دهنده تمرکز بر فرکانس های پایین تر است. مرکز طیفی می تواند به عنوان یک ویژگی در برنامه های مختلف مانند تشخیص گفتار، طبقه بندی سبک موسیقی و طبقه بندی رویدادهای صوتی مورد استفاده قرار گیرد. این ویژگی می تواند برخی از ویژگی های طیفی یک صدا را نشان دهد و به تمایز بین انواع مختلف صداها بر اساس محتوای فرکانسی آنها کمک کند.

سه فیچر مختلف را میتوان از این داده استخراج نمود که میانگین وزن های فرکانسی ، ماکزیمم آن ها و همچنین مقدار میانه این داده بعنوان سه فیچر مجزا برای این تحلیل به کار گرفته شده است.

### 2.1.5 Zero\_Crossing\_rate

Zero Crossing Rate (نرخ گذار صفر) یکی از ویژگی‌های استخراج شده از داده‌های صوتی است. این ویژگی نشان می‌دهد که چقدر سیگنال صوتی از صفر عبور می‌کند. وقتی که سیگنال صوتی از مثبت به منفی یا از منفی به مثبت تغییر می‌کند، یک گذار صفر رخ می‌دهد. نرخ گذار صفر می‌تواند اطلاعاتی درباره ریتم و تناوب سیگنال صوتی ارائه دهد. در سیگنال‌های با ریتم بالا، نرخ گذار صفر بیشتر می‌شود به این معنی که سیگنال صوتی بین مثبت و منفی بیشتر عوض می‌شود. در سیگنال‌هایی با ریتم کمتر، نرخ گذار صفر کمتر می‌شود زیرا سیگنال صوتی بین مثبت و منفی کمتر تغییر می‌کند. استفاده از نرخ گذار صفر در برنامه‌های پردازش صوتی می‌تواند در شناسایی الگوها، تشخیص گفتار، تشخیص خواننده یا سخنران و حتی تحلیل سیگنال‌های موسیقی مفید باشد.

### 2.1.6 Piptrack

Piptrack یک ویژگی در پردازش و تحلیل گفتار است که برای تخمین فرکانس بنیادی (یا پیچیدگی یا تن طبیعی) یک سیگنال صوتی استفاده می‌شود. فرکانس بنیادی نشان دهنده نت مورد شنیدار و تن طبیعی یک صدا است. Piptrack بر پایه مفهوم تحلیل هارمونیک استوار است، جایی که اجزای هارمونیکی در طیف فرکانسی یک سیگنال صوتی شناسایی می‌شوند. این روش، قله‌های طیفی یا پیک‌های موجود در دامنه فرکانس را تحلیل کرده و حرکت آنها در طول زمان را پیگیری می‌کند. با پیگیری این قله‌ها، Piptrack فرکانس بنیادی را تخمین می‌زند و اطلاعاتی درباره تغییرات تن طبیعی در سیگنال صوتی ارائه می‌دهد. این ویژگی رایج در وظایفی مانند تشخیص گفتار، سنتز گفتار و تحلیل موسیقی استفاده می‌شود. این ویژگی به استخراج اطلاعات مربوط به فرکانس بنیادی، شناسایی الگوهای ملودیک و تمایز صداها یا سازهای مختلف بر اساس فرکانس بنیادی آنها کمک می‌کند.

### 2.1.7 Harmonic

هارمونیک‌ها در تشخیص ویژگی‌های صدا در تحلیل گفتار نقش بسیار مهمی دارند. در گفتار، هارمونیک‌ها ضرب‌های صحیحی از فرکانس بنیادی (کمترین جزء فرکانسی) سیگنال صوتی هستند.

هارمونیک‌ها برای نشان دادن پوشش طیفی سیگنال صوتی استفاده می‌شوند و برای شناسایی تن (پیچیدگی) صدا بسیار حائز اهمیت هستند. حضور هارمونیک‌ها در سیگنال نشان می‌دهد که صدا دوره‌ای است، و عدم حضور هارمونیک‌ها نشان می‌دهد که صدا بی‌دوره یا نویزی است.

تشخیص هارمونیک‌ها در برنامه‌های مختلفی مانند تشخیص گفتار، تشخیص گوینده و تحلیل صدای خواننده استفاده می‌شود. این ویژگی در شناسایی فرکانس بنیادی، تعیین کیفیت صدا، و تشخیص حضور برباده‌های صدایی یا ناهنجاری‌های صدایی کمک می‌کند.

به طور خلاصه، تشخیص هارمونیک‌ها در تشخیص ویژگی‌های صدا بسیار حیاتی است زیرا اطلاعاتی درباره تن و پوشش طیفی سیگنال صوتی ارائه می‌دهند.

### 2.1.8 Energy

اندازه گیری میزان انرژی یک سیگنال دارای اطلاعات زیادی میباشد از جمله تشخیص تن صدا برای بیان یک جمله ، همچنین قدرت همچنین صدا های با انرژی بیشتر قابلیت این را دارند که بتوانند راحتتر تشخیص داده شوند و تاثیر نویز در آن ها کمتر بوده و بعنوان یک فیچر میتوان از آن ها استفاده نمود.

### 2.1.9 Beat\_track

Beat tracking یا تشخیص ضرب در موسیقی یک ویژگی از تحلیل صدا است که به کمک آن می‌توان فاصله زمانی بین ضربات را در موسیقی تشخیص داد. این ویژگی معمولاً در تحلیل پردازشی موسیقی بکار می‌رود. در ویژگی های صدا، تشخیص ضرب در موسیقی با استفاده از الگوریتم های مختلفی انجام می‌شود که توسط تحلیل موجک، تبدیل فوریه، تحلیل انرژی و غیره انجام می‌شوند.

تشخیص ضرب در موسیقی در ویژگی های صدا، همچنین می‌تواند در تحلیل ویژگی های صدای انسانی نیز مفید باشد. به عنوان مثال، با تشخیص ضرب در موسیقی با استفاده از روش های مختلف، می توان به بررسی تغییرات فاصله‌های ضربات در زمان در صداهای انسانی پرداخت.

از دیگر کاربردهای تشخیص ضرب در موسیقی به عنوان یک قابلیت صدای هوشمند در تلفن همراه و دستگاه های پخش موسیقی می توان اشاره کرد. با افزودن کاربرد تشخیص ضرب در موسیقی به این دستگاه ها، می توان به طور خودکار تنظیم های مختلف را برای پخش موسیقی انجام داد. به طور مثال، با تشخیص ضرب در موسیقی، دستگاه پخش موسیقی می تواند به طور خودکار حالت های مختلفی را اعم از پخش در حالت شلوغی یا آرامش و... به طور خودکار انتخاب کند.

# **قسمت ۲**. **طبقه‌بندی داده ها**

پیشبینی جنسیت گوینده از روی داده های صوتی

## 1.2. مدل های پیشنهادی

* + توضیح مختصر درباره علت انتخاب مدل‌های طبقه‌بندی

(یکسان بودن نسبت جنسیت در داده های آموزشی و تست)

## 2.2. روش های بهبود نتیجه

* + توضیح مختصر درباره علت و نتیجه استفاده از روش‌های تدریس شده در درس مانند روش‌های کاهش بعد و normalizationدر جهت بهبود کارایی مدل های طبقه بند
  + برای بهبود نتایج خود میتوانید از انواع تکنیک های Dimensionality ،Normalization
  + Reductionو ٍ Ensemble Learningنیز استفاده کنید

## 3.2. نتیجه‌گیری و مقایسه

* + ذکر و تحلیل معیار هایی نظیر F1 score ،recall ،precisionو غیره، برای هر کدام از مدل‌های طبقه‌بند
  + ماتریس آشفتگی، ، ROC Curveتحلیل میزان خطای هر کلاس
  + تاثیر استفاده از بهبود های گفته شده

# **قسمت ۳** **– خوشه بندی داده ها**

## 1.3. مدل های پیشنهادی

**برای خوشه‌بندی داده‌ها، مدل‌های مختلفی وجود دارند. مانند:**

* + **1- خوشه‌بندی سلسله مراتبی (Hierarchical Clustering ):**

**این مدل به ما این امکان را می‌دهد که به صورت سلسله مراتب ، خوشه‌ها را بسازیم. این امر می‌تواند مفید باشد اگر بخواهیم ساختار دقیق‌تری از داده‌هایمان را بدست آوریم.**

* + **2- DBSCAN (Density-Based Spatial Clustering of Applications with Noise):**

**این الگوریتم بر اساس چگالی نقاط به جای تعداد خوشه‌ها عمل می‌کند. این الگوریتم می‌تواند با وجود داده‌های نویزی و اندازه‌های متفاوت خوشه‌ها مقاوم‌تر باشد**

* + **3- الگوریتم K-Means :**

**در انجام این پروژه از روش الگوریتم K-Means استفاده کرده ایم.**

**الگوریتم K-Means با انجام مراحل مقداردهی اولیه و به‌روزرسانی تا همگرایی، داده‌ها را به خوشه‌های مختلف تقسیم کرده و مراکز خوشه‌ها را به موقعیت بهینه رسانده است.**

**با استفاده از روش‌های بهبودی مانند کاهش بعد و نرمال‌سازی، کیفیت و عملکرد مدل بهبود یافته است.**

## **2.3. روش های بهبود نتیجه**

* + **1-کاهش بعد : (Dimensionality Reduction)**

**استفاده از روش‌های کاهش بعد مانند PCA (Principal Component Analysis) یا t-SNE می‌تواند کمک کند تا داده‌های پیچیده‌تر را در یک فضای کمتر ابعاد قرار دهیم و الگوریتم خوشه‌بندی بهبود یابد.**

* + **2- نرمال‌سازی (Normalization):**

**نرمال‌سازی ویژگی‌ها می‌تواند به ازدیاد کارایی الگوریتم‌های خوشه‌بندی کمک کند. این عمل باعث می‌شود تا واحد مقیاس ویژگی‌ها یکسان شود و الگوریتم بهتر بتواند تطابق مناسبی انجام دهد.**

* + **3- انتخاب تعداد خوشه مناسب:**

**استفاده از روش‌هایی مانند نمودار پراکندگی خوشه‌ها و ارزیابی معیارهایی نظیر silhouette score برای انتخاب بهترین تعداد خوشه ممکن است به بهبود نتایج بیانجامد.**

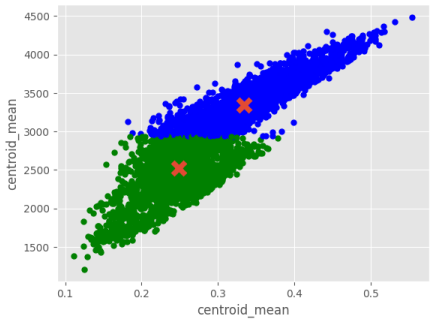
* + **4- استفاده از معیارهای ارزیابی:**

**معیارهای ارزیابی مانند دقت (accuracy) یا F1-score می‌توانند کمک کنند تا عملکرد مدل‌های خوشه‌بندی را بهبود بخشید.**

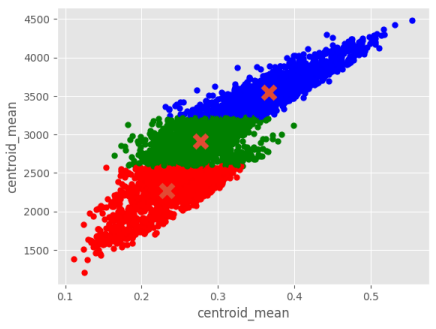
* + **5- استفاده از روش‌هایی مانند نمودار پراکندگی خوشه‌ها و ارزیابی معیارهایی نظیر silhouette score برای انتخاب بهترین تعداد خوشه ممکن است به بهبود نتایج بیانجامد.**

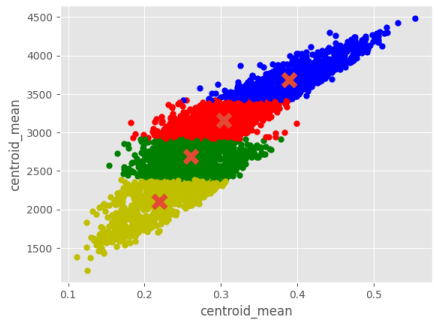
**با اعمال این روش‌ها، می‌توانید بهبودهای چشمگیری در کارایی مدل‌های خوشه‌بندی خود داشته باشید.**

* + **در اینجا ما از خوشه‌بندی k-means با تعداد خوشه‌های مختلف (2و 3 و 4) با استفاده از ویژگی‌های 'zero\_crossings\_mean' و 'centroid\_mean' از مجموعه داده هایمان استفاده کرده‌ایم. هر خوشه با رنگهای مختلف هستند و برای هر خوشه نشانگر 'x' برای مراکز خوشه‌ها نمایش داده شده است.**
  + **نتایج هر حالت:**
  + **حالت 2خوشه:**

****

* + **داده به دو خوشه تقسیم شده است، نقاط سبز یک خوشه و نقاط آبی خوشه دیگر را نمایان می‌کنند.**
  + **به نظر می‌آید که خوشه‌ها تفکیک واضحی دارند و مراکز با 'x' مشخص شده‌اند.**
  + **حالت 3خوشه:**

****

* + **داده حالا به سه خوشه تقسیم شده است، با نقاط سبز، آبی و قرمز به ترتیب هر خوشه را نمایان می‌کنند.**
  + **افزودن یک خوشه اضافه به دقت بیشتر در خوشه‌بندی منجر شده و الگوریتم به نحوی الگوهای دقیق‌تری از داده را ضبط کرده است.**
  + **حالت 4خوشه:**
  + **داده به چهار خوشه تقسیم شده است، با نقاط سبز، آبی، قرمز و زرد به ترتیب هر خوشه را نمایان می‌کنند.**
  + **با اضافه کردن یک خوشه دیگر، دقت بیشتری در تجزیه و تحلیل داده‌ها حاصل شده و هر خوشه یک زیرمجموعه خاص از نقاط را نمایان می‌کند.**

## **3.3. نتیجه گیری و مقایسه**

* + **نتیجه‌گیری و مقایسه مدل‌های خوشه‌بندی با استفاده از الگوریتم K-Means:**
  + **1- میزان پراکندگی:**

**میزان پراکندگی درون خوشه: (Intra-Cluster Dispersion) این معیار نشان دهنده این است که داده‌های درون هر خوشه به چه اندازه نزدیک به یکدیگر هستند.**

* + **2- میانگین پراکندگی میان خوشه: (Inter-Cluster Dispersion) این معیار نشانگر این است که چقدر خوشه‌ها از یکدیگر دور هستند.**
  + **3- تحلیل ویژگی‌های هر خوشه:**

**برای هر تعداد خوشه، نماینده بودن یک خوشه از ویژگی‌های مشترک داده‌های درون آن خوشه نتیجه می‌شود.**

**به عنوان مثال، اگر دو خوشه تشخیص داده شوند (K=2)، ممکن است یک خوشه نماینده داده‌های مرتبط با ویژگی‌های صوتی مردان باشد، و دیگری نماینده داده‌های مرتبط با ویژگی‌های صوتی زنان.**

* + **4- تعداد خوشه‌ها:**

**برای هر تعداد خوشه انتخابی، میزان شباهت داده‌های درون یک خوشه و تفاوت بین خوشه‌ها بررسی می شود.**

**ممکن است برخی داده‌ها در یک خوشه به دلیل شباهت در ویژگی‌ها گروه‌بندی شده باشند. تفاوت‌های بین خوشه‌ها ممکن است به تفکیک صحیح میان دسته‌ها کمک کند.**

* + **مزیت های استفاده از الگوریتم : K-Means**
  + **1- سادگی و سرعت بالا:**

**K-Means یک الگوریتم سریع و ساده است که به راحتی قابل استفاده و تفسیر است.**

* + **2- کارایی در مقیاس‌های بزرگ:**

**برای مجموعه داده‌های بزرگ، K-Means به خوبی مقیاس می‌شود و به سرعت می‌تواند خوشه‌بندی را انجام دهد.**

* + **3- قابلیت مشخص کردن تعداد خوشه‌ها:**

**قابلیت انتخاب تعداد خوشه‌ها توسط کاربر یک امکان مهم است.**

* + **4- استفاده از K-Means با توجه به سادگی، سرعت، و قابلیت مشخص کردن تعداد خوشه‌ها، مزایای زیادی دارد. با توجه به مطالب بالا، این الگوریتم می‌تواند در مسائل مختلفی که نیاز به خوشه‌بندی دارند، مفید باشد.**
  + **در این پروژه با به کارگیری الگوریتم K-Means و با افزایش تعداد خوشه ها از دو به سه و سپس چهارتا خوشه، دقت طبقه بندی افزایش و همچنین میزان پراکندگی درون خوشه کاهش یافت.**

# مدل **(ASR) Automatic Speech Recognition**

* + میزان خطای مدل ASR و بررسی همبستگی این خطا به ویژگی ها
  + تحلیل همبستگی احتمالی میزان خطا با ویژگی های هر صوت (جنسیت، لهجه و لحن)
  + میزان خطای مدل را بر اساس آماره WER7را ذکر کرده
  + همبستگی میزان خطا را با ویژگی هایی نظیر سرعت صحبت کردن و نویز محیطی را نیز به صورت کیفی بررسی کرده