

به نام خدا

دانشگاه تهران  
پردیس دانشکده های فنی  
دانشکده مهندسی برق و کامپیوتر



## گزارش نهایی

درس یادگیری ماشین

فاطمه نائینیان ۸۱۰۱۹۸۴۷۹

پریسا اصغری توانا ۸۱۰۱۹۸۳۵۲

هومن ابوالفتحی ۸۳۰۴۹۹۰۰۸

تیرماه ۱۴۰۱

## فهرست مطالب

۱	مقدمه	۵
۲	مسیر ذهنی و چالش ها	۶
۳	تمیز سازی داده ها	۹
۱.۳	مشکلات اولیه	۹
۲.۳	تمیز سازی CSV اولیه	۹
۳.۳	کاهش نویز	۱۰
۴.۳	استخراج ویژگی	۱۰
۱.۴.۳	معرفی ویژگی ها	۱۰
۲.۴.۳	دسته اول ویژگی ها	۱۲
۳.۴.۳	دسته دوم ویژگی ها	۱۲
۵.۳	نرمالایز کردن داده ها	۱۲
۶.۳	تولید CSV نهایی و تفکیک داده های تست و آموزش	۱۳
۴	طبقه بندی	۱۴
۱.۴	KNN	۱۶
۱.۱.۴	طبقه بندی جنسیت با KNN با کمک Spectrogram Features	۱۷
۲.۱.۴	طبقه بندی جنسیت با KNN با کمک Rythmic Features	۱۹
۳.۱.۴	طبقه بندی احساس با KNN با کمک Spectrogram Features	۲۲
۴.۱.۴	طبقه بندی احساس با KNN با کمک Rythmic Features	۲۴
۲.۴	SVM	۲۶
۱.۲.۴	طبقه بندی جنسیت با SVM با کمک Spectrogram Features	۲۷
۲.۲.۴	طبقه بندی جنسیت با SVM با کمک Rythmic Features	۲۹
۳.۲.۴	طبقه بندی احساس با SVM با کمک Spectrogram Features	۳۱
۴.۲.۴	طبقه بندی احساس با SVM با کمک Rythmic Features	۳۳

۳۴	.....	MLP	۲.۴
۳۵	طبقه بندی جنسیت با MLP با کمک Spectrogram Features	۱.۳.۴	
۳۷	طبقه بندی جنسیت با MLP با کمک Rythmic Features	۲.۳.۴	
۴۰	طبقه بندی احساس با MLP با کمک Spectrogram Features	۲.۳.۴	
۴۲	طبقه بندی احساس با MLP با کمک Rythmic Features	۴.۳.۴	
۴۴	..... LOGISTIC REGRESSION	۴.۴	
	طبقه بندی جنسیت با LOGISTIC REGRESSION با کمک	۱.۴.۴	
۴۵	..... Spectrogram Features		
	طبقه بندی جنسیت با LOGISTIC REGRESSION با کمک	۲.۴.۴	
۴۷	..... Rythmic Features		
	طبقه بندی احساس با LOGISTIC REGRESSION با کمک	۳.۴.۴	
۵۰	..... Spectrogram Features		
	طبقه بندی احساس با LOGISTIC REGRESSION با کمک	۴.۴.۴	
۵۲	..... Rythmic Features		
۵۵	..... جمع بندی طبقه بندی ها	۵.۴	
۵۹		خوشبندی	۵
۵۹	..... روش های خوشبندی	۱.۵	
۵۹	..... Homogeneity	۱.۱.۵	
۶۰	..... Mutual Information	۲.۱.۵	
۶۰	..... V_Score	۳.۱.۵	
۶۰	..... Silhouette	۴.۱.۵	
۶۱	..... Davis Bouldin	۵.۱.۵	
۶۱	..... Calinski-Harabasz	۶.۱.۵	
۶۱	..... K_means	۲.۵	
۶۲	خوشبند K=2 برای ویژگی های Spectrogram	۱.۲.۵	
۶۳	خوشبند K=2 برای ویژگی های Rythmic	۲.۲.۵	
۶۵	خوشبند K=4 برای ویژگی های Spectrogram	۳.۲.۵	

۶۶	خوشه بند K=4 برای ویژگی های Rythmic	۴.۲.۵
۶۸	خوشه بند K=10 برای ویژگی های Spectrogram	۵.۲.۵
۶۹	خوشه بند K=10 برای ویژگی های Rythmic	۶.۲.۵
۷۱	خوشه بند GMM	۳.۵
۷۱	خوشه بند K=2 GMM برای ویژگی های Spectrogram	۱.۳.۵
۷۲	خوشه بند K=2 GMM برای ویژگی های Rythmic	۲.۳.۵
۷۴	خوشه بند K=4 GMM برای ویژگی های Spectrogram	۳.۳.۵
۷۵	خوشه بند K=4 GMM برای ویژگی های Rythmic	۴.۳.۵
۷۷	خوشه بند K=10 GMM برای ویژگی های Spectrogram	۵.۳.۵
۷۸	خوشه بند K=10 GMM برای ویژگی های Rythmic	۶.۳.۵
۸۰	خوشه بند Agglomerative Hierarchical	۴.۵
۸۰	خوشه بند K=2 Hierarchical برای ویژگی های Spectrogram	۱.۴.۵
۸۱	خوشه بند K=2 Hierarchical برای ویژگی های Rythmic	۲.۴.۵
۸۳	خوشه بند K=4 Hierarchical برای ویژگی های Spectrogram	۳.۴.۵
۸۴	خوشه بند K=4 Hierarchical برای ویژگی های Rythmic	۴.۴.۵
۸۶	خوشه بند K=10 Hierarchical برای ویژگی های Spectrogram	۵.۴.۵
۸۷	خوشه بند K=10 Hierarchical برای ویژگی های Rythmic	۶.۴.۵
۸۹	نتیجه گیری خوشه بندی	۵.۵

## ۱ مقدمه

ما در این پژوهه قصد داریم که یک الگوریتم یادگیری ماشین را به هدف تحلیل و پردازش، طبقه‌بندی و خوشه‌بندی داده‌ها، روی مجموعه بزرگی از صوت‌ها اجرا کنیم. در ابتدا برای انجام این کار به جمع آوری داده‌ها پرداختیم و حدود ۱۶ هزار صوت جمیع آوری شد. پس از مرحله جمع آوری داده، با مدلسازی و خوشه‌بندی سعی می‌شود تا توسط هر صوت، جنسیت و احساس آن شخص تشخیص داده شود. در این گزارش، نتایج مدلسازی را بررسی خواهیم کرد.

## ۲ مسیر ذهنی و چالش ها

از مهم‌ترین ابزار ارتباطی انسان‌ها با یکدیگر می‌توان به ابراز احساسات اشاره کرد. در سیستم‌های طبیعی تشخیص این احساسات حیاتی است و به صورت غریزی انجام می‌گیرد؛ اما تشخیص احساسات توسط سیستم‌های کامپیوتری به سادگی امکان‌پذیر نیست. با توجه به پیشرفت و استفاده‌ی گسترده از روش‌های یادگیری ماشین در زندگی روزمره، یکی از مهم‌ترین کاربردهای این دانش را می‌توان تشخیص احساسات انسانی دانست. از آسان‌ترین روش‌هایی که می‌توان با استفاده از آن ابراز احساسات کرد، ارتباط کلامی است.

در این پژوهه قصد داریم با کمک داده‌های صوتی و استفاده از الگوریتم‌های یادگیری ماشین، مدلی را برای تشخیص احساسات ارائه کنیم. اولین قدم برای انجام این پژوهه، جمع‌آوری داده است. برای این قسمت از نوزده شخص خواسته‌ایم تا ده جمله‌ی متفاوت و کوتاه را با چهار احساس مختلف بیان کنند. این احساسات عبارتند از: خوشحالی، عصبانیت، ناراحتی و خنثی. این مجموعه‌ی داده، در واقع هم داده‌های آموزش و آزمون ما را تشکیل خواهد داد.

به صورت کلی می‌توانیم ویژگی این داده‌ها را از منظر کیفی و کمی بررسی کنیم. چهار ویژگی برای یک صوت در نظر گرفته می‌شود: زیرایی (pitch)، طنین (timbre)، مدت (duration) و شدت (intensity). در واقع خوشحالی، ناراحتی یا هر احساسی، بیان کیفی برهم‌کنش چهار ویژگی صدا هستند. به عنوان مثال افراد خوشحال صدایی با شدت بالا و زیرتر نسبت به افراد ناراحت خواهند داشت؛ یا افراد عصبانی بیانی سریع و با شدت بالا دارند. همچنین برای افراد ناراحت می‌توان بیان کرد که کلمات را آرامتر و طولانی‌تر بیان خواهند کرد. اگر بخواهیم این ویژگی‌ها را در گذر سن یک شخص در صدای او بررسی کنیم، می‌توان گفت که به صورت کلی با افزایش سن شخص، صدای او بم‌تر خواهد شد. البته احتمالاً این بم شدن ابتدا افزاینده خواهد بود و سپس در دوره‌ی میانسالی کاهشی خواهد شد. به طوری که صدای یک فرد مسن کمی زیرتر از صدای همان شخص در دوران میانسالی خواهد بود. از منظر جنسیت نیز می‌توان این مشخصه‌ها را بررسی کرد. به طور کلی می‌توان بیان کرد که زنان صدایی نازک‌تر از مردان دارند. البته که این موضوع به صورت مطلق صادق نیست و گاهی ممکن است معکوس باشد.

	Sadness	Happiness	Anger	Normal
Average pitch	Slightly low	Much higher	Very much higher	-
Pitch range	Slightly narrow	Much wider	Much wider	-
Intensity	Lower	Higher	Higher	-
Speech rate	Slightly slower	Faster or Slower	Slightly faster	-
Articulation	Slurring	Normal	Tense	-

شکل ۱: بیان کیفی چندی از مشخصه های بیان به ازای چهار احساس مختلف

دشواری های تشخیص احساسات وقتی تشدید خواهد شد که ویژگی های بیان شده حتی درون اشخاص با سن و جنس یکسان نیز متفاوت است. به طور کلی صدای تمامی افراد در هر چهار ویژگی بیان شده، با دیگران متفاوت است. مقایسه بیان اشخاص با یکدیگر، با توجه تفاوت های بیانی آنها، گاهی احتمال خطرا بالا خواهد برد. پیشتر بیان شد که عصبانی بودن افراد، بیان آنها را شدیدتر خواهد کرد اما اگر شخصی دارای بیان عادی شدید باشد، موضوع را پیچیده خواهد کرد چرا که دیگر نمیتوان تنها با بررسی شدت بیان این دو شخص به احساس آنها پی ببریم. این موضوع درمورد زیرایی صدا نیز صادق است به صورتی که اگر تنها زیرایی صدا را در تشخیص خوشحالی افراد در نظر بگیریم، احتمالاً بیان عادی یک شخص با صدای زیر را خوشحال قلمداد خواهیم کرد. همچنین اینکه افراد غالباً سرعت های بیان متفاوتی دارند، بررسی مدت خوانش یک جمله نیز نمی تواند به تنها یی کافی باشد. به طور مثال ممکن است یک فرد عصبانی که به صورت عادی کند صحبت می کند، با بیان خنثی اشتباه گرفته شود. بنابراین نباید صرفاً یک ویژگی را در بررسی احساس بیان افراد نظر بگیریم.

یکی دیگر از چالش های این پروژه، نحوه جمع آوری مجموعه داده بود. با توجه به اینکه ما از افراد عادی داده برداری کردیم، احتمال آنکه این افراد در بیان احساسات خود افراط یا تفریط کرده باشند، وجود دارد. بنابراین اگر داده ها کیفیت لازم را نداشته باشند، احتمال آن که مدل ما خطای داشته باشد، بیشتر خواهد شد. البته که خطای در تمامی مدل های ارائه شده در مسائل یادگیری ماشین وجود دارد و تنها باید آن را کم کنیم. دیگر عامل مخرب را می توان نویز دانست. در سیستم ها معمولاً نسبت سیگنال به نویز (SNR)، از مهم ترین مشخصات یک سیگنال به شمار می رود که کاهش آن به معنای ناکارآمد بودن سیگنال اصلی است. ما در جمع آوری این مجموعه داده تلاش کردیم تا مقدار نویز را به حداقل برسانیم. البته در صورت وجود نویز می توان آن را به نحوی از بین برد که در بخش مربوطه به آن پرداخته خواهد شد.

تا به اینجای گزارش به تفاوت‌های موجود در صدای افراد مختلف و چالش‌های جمع‌آوری داده پرداختیم؛ اما باید توجه داشته باشیم که ویژگی‌هایی می‌توانند مورد استفاده قرار بگیرند که بتوان آن‌ها را به صورت اعداد بیان کرد و کامپیوترها توانایی تحلیل آن‌ها را داشته باشند. به عبارت دیگر باید بتوانیم ویژگی‌های صدا را کمی کنیم. در ادامه به بیان چندی از کمیت‌هایی می‌پردازیم که در مدل ما مورد استفاده قرار خواهند گرفت. به طور کلی می‌توان ویژگی‌های مختلفی را در دو حوزه‌ی زمان و فرکانس برای سیگنال‌های صوت تعیین کرد. معمولاً ویژگی‌های فرکانسی گویاتر و کاربردی‌تر هستند. انرژی و بازه‌ی زمانی، دو مورد از مهمترین مشخصه‌های زمانی سیگنال‌ها به شمار می‌روند. همچنین در حوزه‌ی فرکانس می‌توان از ویژگی‌هایی چون بازه‌ی فرکانسی، فرکانس غالب، فرکانس‌های غالب دوم یا چندم، میانگین فرکانسی و ... استفاده کرد. باید به این موضوع توجه کنیم که سیگنال صوت یک سیگنال زمانی است و برای انتقال آن به حوزه‌ی فرکانس کافیست تبدیل فوریه‌ی آن را محاسبه کنیم. لازم به ذکر است که انتخاب ویژگی‌های مناسب را شاید بتوان از مهم‌ترین اقدامات ارائه‌ی یک مدل پیش‌بینی کننده‌ی کارا دانست. از صفات یک ویژگی مناسب می‌توان به عدم هم‌پوشانی دسته‌های مختلف به ازای آن ویژگی اشاره کرد. این موضوع را می‌توان به این صورت بیان کرد که برای مثال اگر از ویژگی الف برای جداسازی احساسات مختلف از یکدیگر استفاده می‌کنیم، این ویژگی باید کم‌ترین شباهت را در بین احساسات مختلف داشته باشد. با انتخاب این ویژگی‌ها در واقع ما ماتریس ویژگی‌های سیستم را تعیین کرده‌ایم. به بیان دیگر ما یک مدل طبقه‌بند را پیشنهاد خواهیم داد که با بررسی این ویژگی‌ها به عنوان ورودی مدل، برچسب آن داده که در واقع احساس آن جمله است را تعیین خواهد کرد.

به طور کلی تشخیص احساسات موضوع پیچیده‌ای است. هر احساس ویژگی‌های خاص خود را در بیان فرد ایجاد خواهد کرد و طبقه‌بندی این احساسات امر پیچیده‌ای خواهد بود کما اینکه گاهی انسان‌ها در تشخیص احساسات دچار اشتباه می‌شوند. این پیچیدگی وقتی بیش‌تر خواهد شد که بخواهیم مدلی کامپیوتری برای این تشخیص ارائه دهیم. ما در این پژوهه برآنیم تا با ارائه‌ی یک مدل بر مبنای روش‌های یادگیری ماشین و با استفاده از ویژگی‌های مختلف صدا، چه در حوزه‌ی زمان و چه در حوزه‌ی فرکانس، طبقه‌بندی این احساسات را میسر کنیم. این مدل با استفاده از ماتریس ویژگی‌هایی که تعیین می‌کنیم و داده‌های جمع‌آوری شده، تشخیص خواهد داد که یک جمله‌ی جدید با چه احساسی بیان شده است.

## ۳ تمیز سازی داده ها

### ۱.۳ مشکلات اولیه

ابتدا فایل ها را دانلود کرده و سعی شد با کتابخانه librosa صوت ها را در کد پایتون وارد کنیم. در این مرحله متوجه شدیم که تعدادی از صوت ها دچار مشکل فرمت هستند و ایراد داشتند. پس ابتدا در فایل Finding corrupted.py صوت هایی که دچار مشکل فرمت بودند و صوت هایی که طولانی بودند و طول بیشتر از ۷.۵ ثانیه داشتند را به عنوان صوت های ایراد دار در یک فolder دیگر ذخیره کردیم. صوت هایی که طول بیشتر ۷.۵ ثانیه داشتند را کلا کنار گذاشتیم و آنهایی که فرمت اشتباه داشتند که حدود سه هزار صوت بود را با نرم افزار WINFF به فرمت wav تبدیل کردیم. در نهایت یک فolder اصلی داریم که فرمت همه صوت ها درست است و طول اضافه ندارند. پس از این به بعد از این فolder به عنوان صوت های مرجع استفاده کردیم.

### ۲.۳ تمیز سازی CSV اولیه

۱. اگر دو تا Voice ID یکسان باشند یکی از آنها حذف شود.
۲. اگر چندین Voice ID شبیه هم بودند، همه آنها حذف می شود.
۳. برای همه Voice ID ها صوت متناظر آنها وجود داشته باشد. اگر برای یک Voice ID صوت متناظر آن وجود نداشت، اطلاعات آن صوت از CSV حذف می شود.
۴. در قسمتی از CSV دو ستون مربوط به Text ID و Emotion جا به جا شده بودند که به صورت دستی این دو ستون را جا به جا کردیم.
۵. همچنین در ستون SEX گاهی با حروف بزرگ و گاهی با حروف کوچک نام گذاری شده بود. برای اینکه این موضوع را بطرف کنیم جنسیت را با دو عدد ۰ و ۱ جایگذاری کردیم.
۶. در ستون Text ID دسته ای از اعداد به جای اینکه از ۱ تا ۱۰ باشند، از ۰ تا ۹ بودند. در این مورد نیز همه Text ID ها به علاوه ۱ کردیم تا مشکل بر طرف شود.

### ۳.۳ کاهش نویز

در مرحله گرداوری صوت ها، نویز و صدای اضافه زیادی وجود دارد پس لازم تا نویز ها را کاهش دهیم. با کاهش نویز دقت طبقه بند بیشتر می شود و خطای کمتر می شود. نویز موجود در صوت ها سبب می شود تا اطلاعات اصلی موجود در صوت کم اهمیت تر شوند و تاثیر کمتری در مدل داشته باشد. با کمک کتابخانه noisereduce و در فایل Denoising.py نویز را کاهش میدهیم. سپس همه صوت های کاهش نویز یافته را در یک فolder جدید ذخیره میکنیم تا در مرحله استخراج ویژگی از آن استفاده کنیم.

### ۴.۳ استخراج ویژگی

#### ۱.۴.۳ معرفی ویژگی ها

\* Zero Crossing Rate: نرخ عبور صفر، نرخ تغییرات علامت در طول یک سیگنال است؛ یعنی نرخی که در آن، سیگنال از مثبت به منفی یا به صورت بر عکس تغییر می کند. این ویژگی، هم در تشخیص گفتار و هم در بازیابی اطلاعات موسیقی به شدت مورد استفاده قرار گرفته است.

\* Spectral Centroid (مرکز طیفی): این ویژگی نشان می دهد که میانگین وزن دار برای یک داده صوتی در کجای آن قرار دارد. و از طریق میانگین وزنی فرکانس های موجود در داده صوتی محاسبه می شود. اگر فرکانس های موسیقی در طول آهنگ یکسان باشند، مرکز طیفی، حول یک مرکز خواهد بود؛ و اگر فرکانس های بالایی در انتهای داده صوتی وجود داشته باشد، مرکز به سمت انتهای آن متمایل می شود.

$$\text{Centroid} = \frac{\sum_{n=0}^{N-1} f(n)x(n)}{\sum_{n=0}^{N-1} x(n)} \quad (1)$$

\* Spectral Rolloff: بین ویژگی، بیانگر فرکانس قسمت هایی از سیگنال است که کمتر از درصد مشخصی از کل انرژی طیفی هستند.

\* Chroma Frequencies: ویژگی های کرومایک نمایش جالب و قدرتمند برای داده صوتی است که در آن کل طیف در ۱۲ bin نمایش داده می شود. که نشان دهنده

۱۲ نیم تون (یا کرومما) متمایز اکتاو موسیقی است. از آنجایی که در موسیقی، نت‌هایی که دقیقاً یک اکتاو از هم فاصله دارند بسیار شبیه یکدیگر هستند، دانستن توزیع کرومما حتی بدون فرکانس مطلق (یعنی اکتاو اصلی) هم می‌تواند اطلاعات مفیدی در مورد داده صوتی بدهد. و حتی ممکن است شbahت‌هایی را آشکار کند که در طیف اصلی مشخص نیست. در هر فریم، pitch‌های متفاوت وجود دارد که کیفیت یک لحن خاص را از بقیه صدای‌های داخل یک اکتاو جدا می‌کند و تفاوت‌های ادراکی گام‌ها در یک اکتاو و یکسانی ادراکی گام‌هایی که با یک یا چند اکتاو کامل از هم جدا شده‌اند را توصیف می‌کند.

\* Mel Frequency Cepstral Coefficients (MFCCs): این ویژگی یکی از مهم‌ترین روش‌ها برای استخراج ویژگی سیگنال صوتی است و به طور عمده در هنگام کاربر روی سیگنال‌های صوتی استفاده می‌شود. MFCCs یک سیگنال مجموعه کوچکی از ویژگی‌ها (معمولأ حدود ۲۰-۱۰) هستند که به طور خلاصه شکل کلی یک پوشش طیفی را توصیف می‌کنند. این ویژگی مهمترین ویژگی است و مشابه شناوی انسان عمل می‌کند.

\* Spectral Bandwidth: این ویژگی واریانس از مرکز طیف را نشان می‌دهد و با انرژی پخش شده در باند‌های فرکانسی متناسب است.

\* RMS: این ویژگی با توان موثر سیگنال مرتبط است و مقدار rms را برای هر فریم بر می‌گرداند.

\* Tempo: تندا به میزان تندا یا کندی در هر قطعه موسیقی گفته می‌شود. چون شکل‌های مختلف نت‌های موسیقی در بین خود دارای ارزش‌های نسبی هستند و طول زمانی آن‌ها در آثار مختلف موسیقی با هم تفاوت دارند؛ از این رو اصطلاحاتی به زبان ایتالیایی در شروع هر اثر موسیقایی به کار می‌رود و تندا‌های گوناگون در دوره‌های مختلف تاریخ موسیقی را تا حدود زیادی مشخص می‌کنند. از متُرم برابر تعیین سرعت نواختن استفاده می‌شود و واحد اندازه گیری تندا در موسیقی، ضرب در دقیقه یا Beats Per Minute است.

\* Rhythm: یا ریتم نخستین عامل اصلی در موسیقی است که به آن زندگی و حرکت میبخشد. ریتم عبارت است از دیرند های گوناگون صداها و سکوت ها در یک مجموعه صوتی. نت های مختلف موسیقایی از منظر دیرند، وقتی در کنار هم قرار میگیرند، ریتم یک موسیقی را تشکیل میدهند. سکوت ها و تاکید ها نیز در خلق ریتم شریک هستند.

### ۲.۴.۳ دسته اول ویژگی ها

در حالت اول، ویژگی های زیر را انتخاب میکنیم:

Zero Crossing, Spectral Centroid, Spectral Rolloff, Chroma Frequencies, MFCCs

باید توجه کرد که از ویژگی ها باید میانگین بگیریم تا ویژگی را برای کل صوت پیدا کنیم (این ویژگی ها برای هر فریم محاسبه می شود).

پس به این ترتیب به ازای هر صوت، یک ویژگی Zero Crossing و یک ویژگی Spectral Centroid و یک ویژگی Spectral Rolloff و دوازده ویژگی Chroma Frequencies و پنجاه ویژگی MFCCs داریم.

در نهایت ۶۵ ویژگی خواهیم داشت که نتایج آنها در یک CSV ذخیره شده است.

### ۳.۴.۳ دسته دوم ویژگی ها

در حالت دوم ویژگی Fourier tempogram را استفاده میکنیم. این ویژگی ۱۹۳ داده به ازای هر صوت به ما میدهد که اطلاعات آن را در CSV ذخیره میکنیم.

## ۵.۳ نرمالایز کردن داده ها

پس از اسخراج ویژگی ها و پیدا کردن CSV اولیه، ویژگی ها را نرمال میکنیم. برای این کار از MinMaxScaler کتابخانه sklearn استفاده کردیم. این الگوریتم دیتا را استانداردسازی می کند. یعنی به صورتی تغییر می دهد که داده ها دارای توزیع گاووسی شوند و میانگین صفر و واریانس یک بگیرند. ما رنج کمینه و بیشینه استانداردسازی را ۰ و ۱ قرار دادیم تا داده ها نرمال هم بشوند و در محدوده بین صفر و یک قرار بگیرند. بدین صورت عمل میکند که همه داده های ویژگی ها را از کمترین داده آن ویژگی کم می کند و حاصل را بر اختلاف بزرگترین و کوچک ترین داده آن ویژگی، تقسیم می کند.

این کار باعث می شود که وزن همه داده ها با هم برابر شود و اثر یکی بر دیگری برتری نداشته باشد. اهمیت این موضوع زمانی مشخص می شود که یکی ویژگی ها در یک محدوده نباشند و یک یا چند تا از آنها تاثیر بیشتری روی برچسب نهایی بگذارد، اگر این ویژگی ها با برچسب ناسازگاری داشته باشند، سبب ایجاد خطای زیادی می شوند و دقت کم می شود. به طور خلاصه تاثیر این کار این است که مدل را از سوگیری روی یک ویژگی خاص از داده ها نجات می دهد، حجم محاسبات عددی را بسیار کاهش می دهد و سرعت مدل را بسیار بهبود می بخشد و در نهایت دقت بهتری را برای مدل نتیجه می دهد.

### ۶.۳ تولید CSV نهایی و تفکیک داده های تست و آموزش

حال بعد از اینکه ویژگی ها را استخراج کردیم، لازم است تا یک CSV نهایی تولید کنیم. در این فایل باید در کنار ویژگی های استخراج شده، label هایی که در CSV داشتیم، یعنی جنسیت (SEX) و احساس (Emotion) و متن (Text ID) به هر سطر متناظر آن صوت اضافه کنیم. در نهایت داده ها به داده های آموزش و تست تقسیم میکنیم و برای آنها CSV میسازیم. با کمک این تکنیک ارزیابی عملکرد بهتری از الگوریتم یادگیری ماشین خواهیم داشت و میتوانیم از صحت عملکرد مدل مطمئن شویم. با این روش، مدل خود را روی داده های آموزش اجرا می کنیم و fit می کنیم و مدل خود را دیگر روی داده آزمون، fit نمی کنیم. سپس با اجرای مدل خود روی داده آزمون، عملکرد مدل خود را می سنجیم.

باید مقدار داده های آموزش و آزمون به حدی باشد که از طرفی داده برای آموزش مدل کم نیاید و دچار نحسی ابعاد نشویم و از طرفی داده آموزش آنقدر زیاد نباشد که مدل ما روی داده های نویزی هم آموزش بینند؛ به صورت خلاصه، میزان داده ها باید به گونه ای باشد که مدل ما overfit و underfit نکند.

برای این کار، چند نفر را جدا میکنیم تا هیچ اطلاعاتی از آنها در داده های آموزش نباشد و صرفا در داده های تست وجود داشته باشند. سپس داده ها را شافل میکنیم. سپس داده های آموزش و تست را با الگوریتم train\_test\_split کتابخانه sklearn جدا میکنیم و آن چند نفری که جدا کرده بودیم را به داده های تست اضافه میکنیم و دوباره شافل میکنیم. در نهایت برای داده های تست و آموزش چهار CSV خواهیم داشت که شامل X\_train، Y\_train، X\_test، Y\_test می شود.

## ۴ طبقه بندی

یک بار بر اساس جنسیت و یک بار هم بر اساس احساسات، طبقه بندی را انجام میدهیم. همچنین یک بار برای دیتاست spectrum features و یک بار هم برای دیتاست tempogram features طبقه بندی را انجام میدهیم. برای طبقه بندی، از ۴ روش Logistic ، KNN و SVM ، MLP استفاده میکنیم. ابتدا با استفاده از روش cross validation، هایپر پارامتر های مدل Regression را انجام میدهیم. که بهترین accuracy را دارند، بدست می آوریم. سپس مدل را با استفاده از داده های train، اموزش داده و عملکرد آن را برای داده های test و train می سنجیم. معیار های گزارش شده: اولین و مهمترین روش ارزیابی طبقه بند ها را میتوان ماتریس درهم ریختگی یا همان confusion matrix نامید. این ماتریس نشان میدهد که چه تعداد از داده ها به درستی و چه تعداد از آنها نادرست طبقه بندی شده اند. ماتریس درهم ریختگی یک ماتریس مرتبی است که تعداد سطر های آن به تعداد کلاس های داده ها است. سطر ها در این ماتریس نماینده ی برچسب های واقعی است و ستون ها نماینده ی برچسب های پیش بینی شده توسط طبقه بند مورد استفاده است. پس با توجه به سطر و ستون هر درایه میتوان عددی که در آن درایه نوشته شده است را تفسیر کرد. با استفاده از همین ماتریس چند نسبت برای ارزیابی نتایج تعریف میشود که در ادامه توضیحی کوتاه راجع به چندی از آنها بیان خواهیم کرد: (از این پس تشخیص های صادق و مثبت را با TP، صادق و منفی را با TN، کاذب و مثبت را با FP و کاذب و منفی را با FN نشان میدهیم. حساسیت یا recall به نسبت تشخیص های مثبت و صادق به جمع تشخیص های مثبت و صادق با تشخیص های منفی و کاذب گفته میشود. این پارامتر از خروجی های برنامه ای است.

$$Recall = \frac{TP}{TP + FN}$$

بازیابی اطلاعات یا precision به نسبت تشخیص های مثبت و صادق به جمع تشخیص های مثبت و صادق با تشخیص های مثبت و کاذب گفته میشود. این پارامتر هم در خروجی برنامه قرار دارد.

$$Precision = \frac{TP}{TP + FP}$$

امتیاز f1 یا همان f1\_score به نسبت تشخیص های مثبت و صادق به مجموع تشخیص های مثبت و صادق و نصف تشخیص های مثبت و کاذب با منفی و کاذب گفته میشود. این کمیت نیز در

خروجی نمایش داده شده است.

$$f1 - Score = \frac{TP}{TP + \frac{FP+FN}{2}}$$

و در نهایت میتوان مهمترین کمیت ارزیابی عملکرد طبقه بند را دقت یا accuracy محسوب کرد. تعریف دقت به نسبت تمامی تشخیص های صادق، اعم از مثبت یا منفی، به کل تعداد نمونه ها است. این کمیت به صورت تکی به ازای هر روش طبقه بندی اعلام شده است. دقت های بدست آمده در الگوریتم های ما، برای جنسیت به مراتب بالاتر و بهتر از احساسات بوده است.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

لازم به ذکر است که در ماتریس های درهم ریختگی چند کلاسه، تعریف تشخیص صادق واضح است. اما برای تشخیص های کاذب میتوان اینگونه بیان کرد که در یک کلاس خاص، تنها پیش‌بینی تعلق یا عدم تعلق به همان کلاس، صادق محسوب میشود و اگر اتلاف به هر کلاس دیگری باشد، کاذب محسوب میشود. (One against all) در نهایت معیار دیگری از ارزیابی عملکرد یک طبقه بند، نمودار ROC یا Receiver operating characteristic یا همان منحنی مشخصه ای عملکرد سیستم است. در این نمودار نسبت تشخیص های صادق و مثبت (همان حساسیت یا recall) به عنوان تابعی از نسبت تشخیص های کاذب و مثبت (نسبت تشخیص های مثبت و کاذب به جمع تشخیص های مثبت و کاذب و منفی و صادق) رسم میشود. باید در نظر داشته باشیم که در یک تشخیص تصادفی یا به عبارتی بدترین طبقه بند، این دو نسبت با هم مساوی هستند. به عبارت دیگر نمودار مشخصه عملکرد سیستم برای یک طبقه بند تصادفی یک خط مبدا گذرا شیب ۱ است. از طرفی به صورت شهودی میتوان به این نتیجه رسید که طبقه بندی بهتر است که نسبت حساسیت به نسبت تشخیص های کاذب و مثبت بیشتر باشد. پس در واقع هر چه نمودار از خط مبدا گذرا شیب ۱ فاصله ای بیشتری داشته باشد، عملکرد بهتری داشته است.

$$False - Positive - Rate(FPR) = \frac{FP}{FP + TN}$$

لازم به بیان مجدد است که در طبقه بندی نیاز به تقسیم مجموعه داده ها به دو مجموعه ای مجزای آموزش و آزمون داریم. از مجموعه ای آموزش تنها برای آموزش طبقه بند استفاده

میکنیم؛ سپس از مجموعه‌ی آزمون برای ارزیابی طبقه بند استفاده میکنیم. به این صورت که برچسب های این مجموعه توسط طبقه بند اعلام میشوند و با دانستن برچسب های واقعی این مجموعه، اقدام به ارزیابی نتایج با توجه به سنجه های بیان شده میکنیم. پس تمامی نتایج فوق از اعمال طبقه بند بر روی مجموعه‌ی آزمون استخراج شده اند.

برای تعریف novelty curves ابتدا نیاز داریم تا fourier tempogram را بررسی کنیم. برای صورت زیر تعریف میشود:

$$\Delta(t) := \sum_{k=1}^K |Y(t+1, k) - Y(t, k)| \geq 0$$

که در آن:

$$Y(t, k) := \log(1 + C \cdot |X(t, k)|)$$

و  $X(t, k)$  در واقع k امین ضریب تبدیل فوریه برای سیگنال زمانی  $x(t)$  است و  $C$  یک ثابت است. حال با دانستن fourier tempogram، میتوانیم novelty curve را تعریف کنیم:

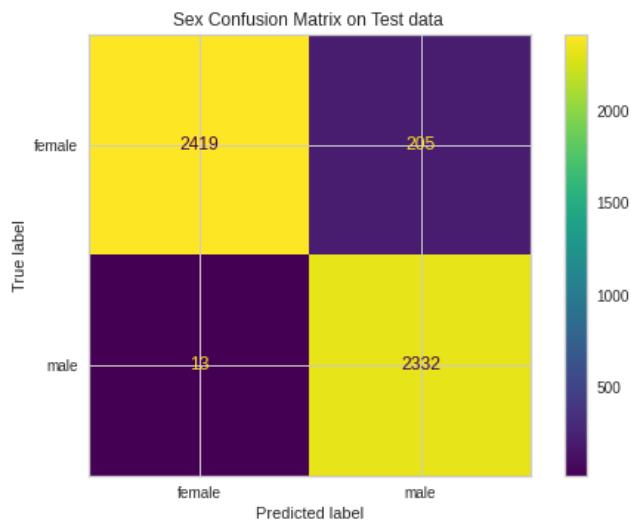
$$\text{Fourier Tempogram}(t, w) = \sum_{n \in Z} \Delta(n) \cdot W(n - t) \cdot e^{-2\pi i w n}$$

## KNN ۱.۴

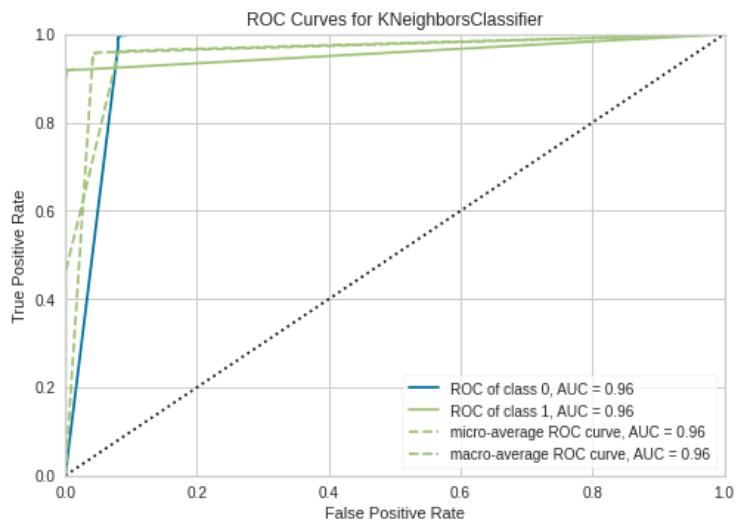
در طبقه بند KNN یا K nearest neighbors گیری در مورد برچسب هر داده، تعداد همسایه های نزدیک به آن داده است. به این ترتیب که برای هر داده تعداد معینی از همسایه های آن داده بررسی میشوند. برچسب نهایی داده با توجه به برچسب همسایه های آن تعیین میشود. با اجرای این الگوریتم، مرز تصمیم تعیین خواهد شد. باید در نظر داشت تعداد همسایه های مورد بررسی یک هایپر پارامتر برای این روش محسوب میشود. از دیگر پارامتر هایی که میتوان در این روش تغییر داد، تعریف فاصله است. برای این روش میتوان هر دو تعریف اقلیدسی یا منهتن را از فاصله بیان کرد که این موضوع نیز در برنامه بررسی شده است.

## ۱.۱.۴ طبقه بندی جنسیت با KNN با کمک Spectrogram Features

```
best parameters{'algorithm': 'auto', 'metric': 'manhattan', 'n_neighbors': 4}
max validation score0.9955124640940339
accuracy score on Test data for Sex : 0.9561
Sex classification report on test data
      precision    recall   f1-score   support
female        0.99     0.92     0.96    2624
male         0.92     0.99     0.96    2345
accuracy          0.96     0.96     0.96    4969
macro avg       0.96     0.96     0.96    4969
weighted avg    0.96     0.96     0.96    4969
```



شکل ۲: confusion matrix برای جنسیت با داده های آزمون با KNN با ویژگی های Spectrogram

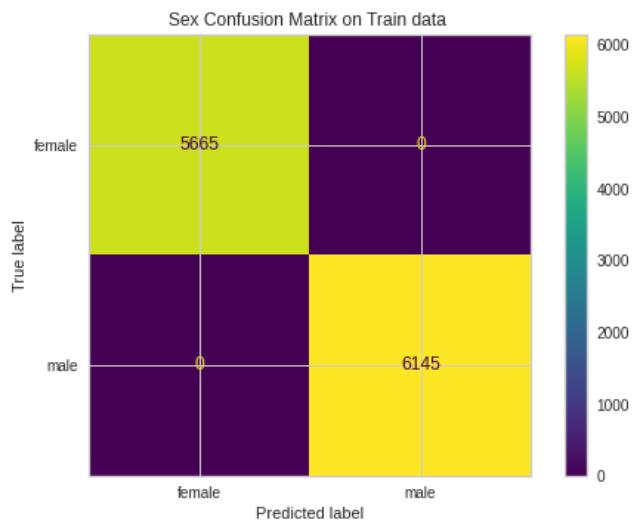


شکل ۳: نمودار ROC برای جنسیت با طبقه بند KNN با ویژگی های Spectrogram

```

accuracy score on Train data for Sex : 1.0000
Sex classification report on train data
precision      recall   f1-score   support
female         1.00      1.00      1.00      5665
male           1.00      1.00      1.00      6145
accuracy       1.00      1.00      1.00     11810
macro avg      1.00      1.00      1.00     11810
weighted avg   1.00      1.00      1.00     11810

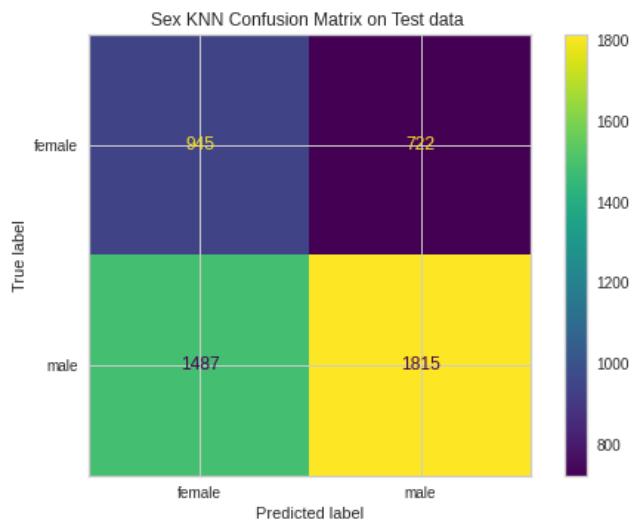
```



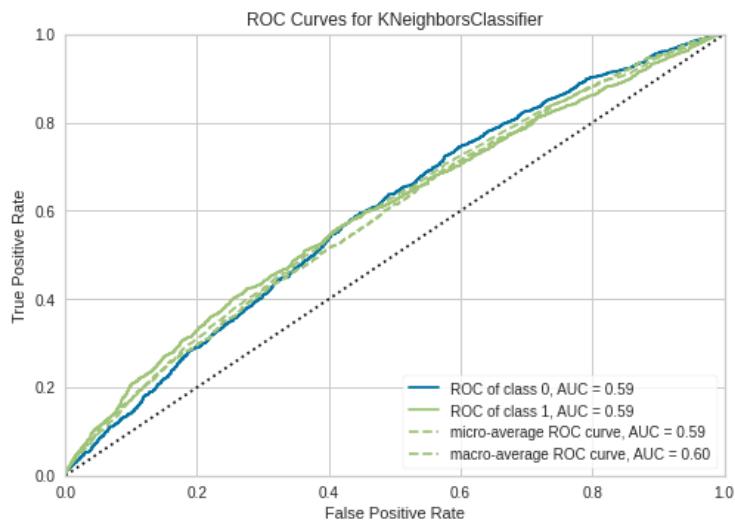
شکل ۴ : confusion matrix برای جنسیت با داده های آموزش با KNN با ویژگی های Spectrogram

#### ۲.۱.۴ طبقه بندی جنسیت با KNN با کمک Rythmic Features

```
best parameters{'algorithm': 'auto', 'metric': 'euclidean', 'n_neighbors': 14}
max validation score 0.5778999423852507
Sex KNN accuracy score on Test data for Sex : 0.5554
Sex KNN classification report on test data
      precision    recall   f1-score   support
  female       0.39     0.57     0.46    1667
    male       0.72     0.55     0.62    3302
  accuracy        0.55     0.56     0.54    4969
  macro avg       0.55     0.56     0.54    4969
  weighted avg    0.61     0.56     0.57    4969
```



شکل ۵: نمودار confusion matrix برای جنسیت با داده های آزمون با KNN با ویژگی های rhythmic

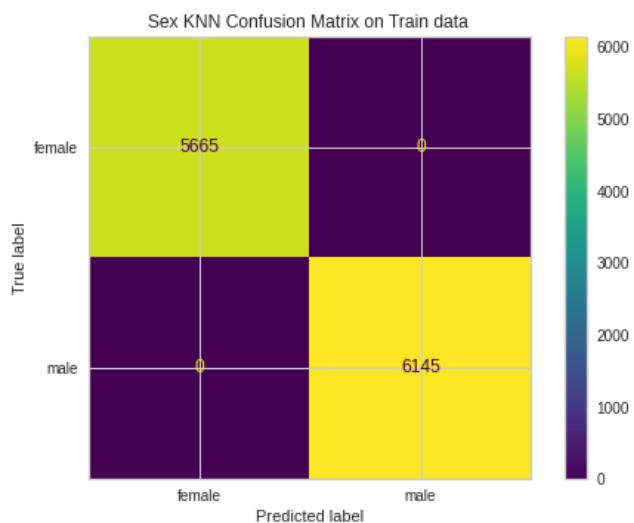


شکل ۶: نمودار ROC برای جنسیت با طبقه بند KNN با ویژگی های rhythmic

Sex KNNaccuracy score on Train data for Sex : 1.0000

Sex KNN classification report on train data

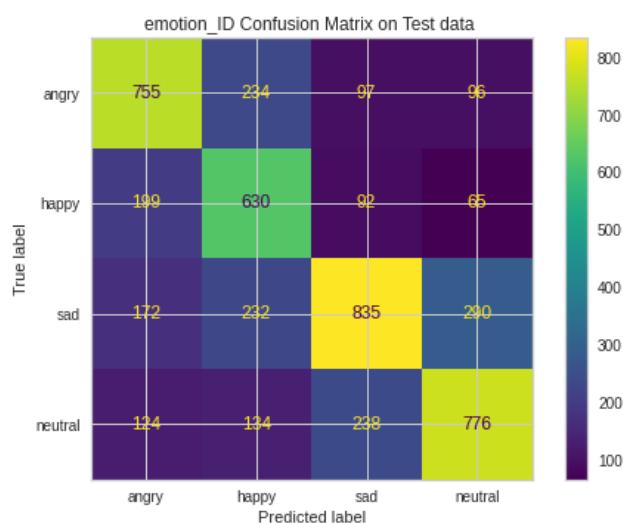
	precision	recall	f1-score	support
female	1.00	1.00	1.00	5665
male	1.00	1.00	1.00	6145
accuracy			1.00	11810
macro avg	1.00	1.00	1.00	11810
weighted avg	1.00	1.00	1.00	11810



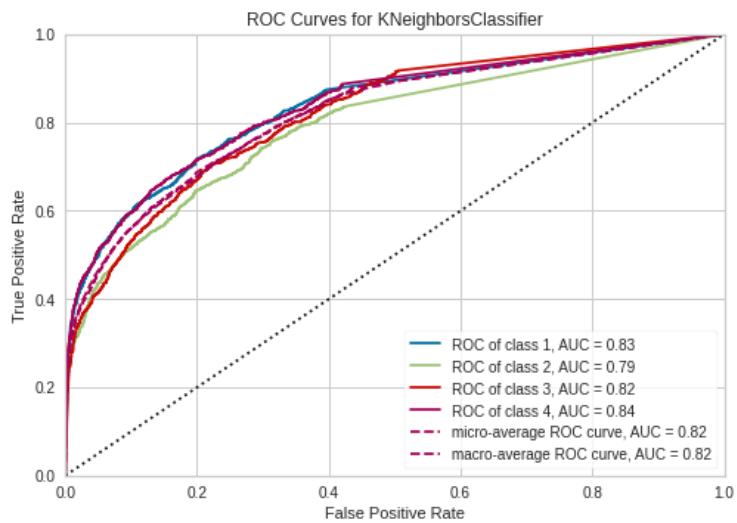
شکل ۷: confusion matrix برای جنسیت با داده های آموزش با KNN با ویژگی های rhythmic

### ۳.۱.۴ طبقه بندی احساسات با KNN با کمک Spectrogram Features

```
accuracy score on Test data for emotion_ID : 0.6029
emotion_ID classification report on test data
precision    recall    f1-score   support
angry        0.60      0.64      0.62      1182
happy        0.51      0.64      0.57      986
sad          0.66      0.55      0.60      1529
neutral      0.63      0.61      0.62      1272
accuracy      0.60      0.61      0.60      4969
macro avg    0.60      0.61      0.60      4969
weighted avg 0.61      0.60      0.60      4969
```



شکل ۸: confusion matrix برای احساسات با داده های آزمون با KNN با ویژگی های Spectrogram

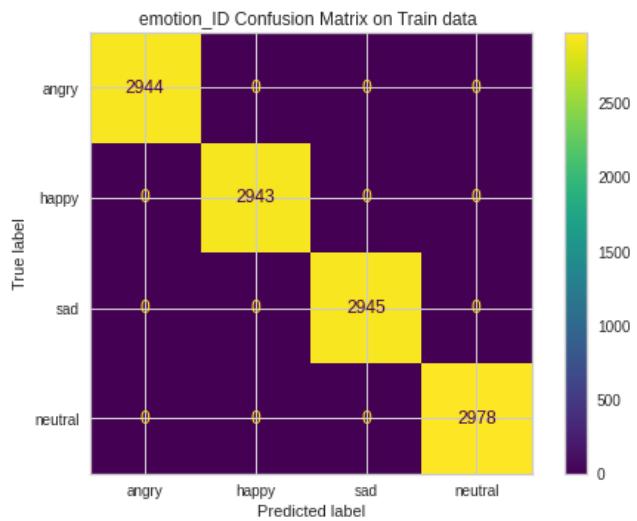


شکل ۹: نمودار ROC برای احساسات با طبقه بند KNN با ویژگی های Spectrogram

```

accuracy score on Train data for emotion_ID : 1.0000
emotion_ID classification report on train data
      precision    recall  f1-score   support
angry       1.00     1.00     1.00     2944
happy       1.00     1.00     1.00     2943
sad         1.00     1.00     1.00     2945
neutral     1.00     1.00     1.00     2978
accuracy
macro avg       1.00     1.00     1.00    11810
weighted avg     1.00     1.00     1.00    11810

```



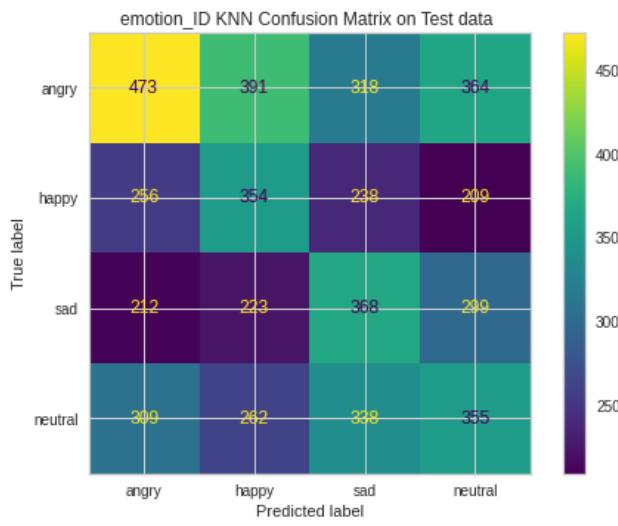
شکل ۱۰: confusion matrix برای احساسات با داده های آموزش با KNN با ویژگی های Spectrogram

#### طبقه بندی احساس با کمک KNN با Rythmic Features ۴.۱.۴

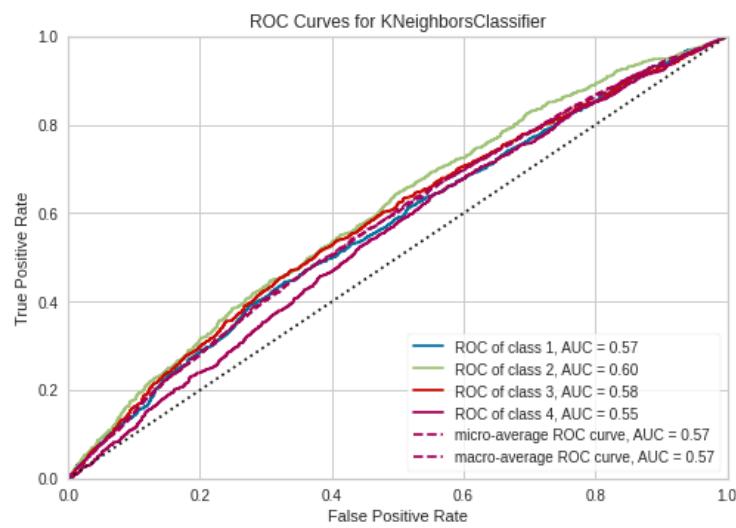
```

best parameters{'algorithm': 'auto', 'metric': 'euclidean', 'n_neighbors': 13}
max validation score 0.3090603452548369
emotion_ID KNN accuracy score on Test data for emotion_ID : 0.3119
emotion_ID KNN classification report on test data
      precision    recall   f1-score   support
angry        0.38     0.31     0.34    1546
happy        0.29     0.33     0.31    1057
sad          0.29     0.33     0.31    1102
neutral      0.29     0.28     0.29    1264
accuracy           0.31     0.31     0.31    4969
macro avg       0.31     0.31     0.31    4969
weighted avg    0.32     0.31     0.31    4969

```



شکل ۱۱: confusion matrix برای احساسات با داده های آزمون با KNN با ویژگی های rhythmic

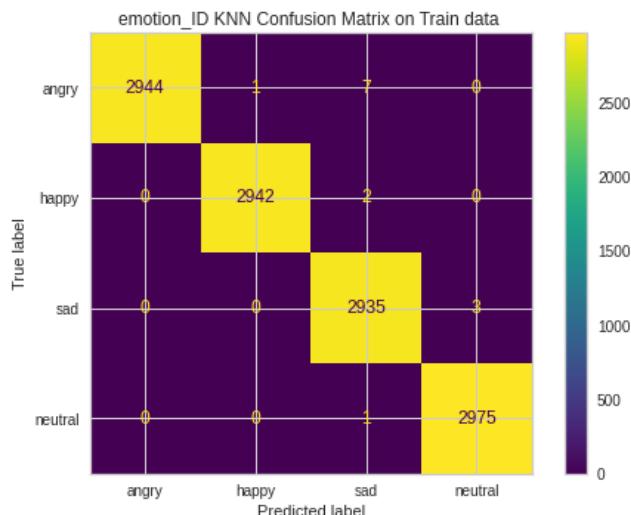


شکل ۱۲: نمودار ROC برای احساسات با طبقه بند KNN با ویژگی های rhythmic

```
emotion_ID KNNaccuracy score on Train data for emotion_ID : 0.9988
```

```
emotion_ID KNN classification report on train data
      precision    recall   f1-score   support
```

angry	1.00	1.00	1.00	2952
happy	1.00	1.00	1.00	2944
sad	1.00	1.00	1.00	2938
neutral	1.00	1.00	1.00	2976
accuracy			1.00	11810
macro avg	1.00	1.00	1.00	11810
weighted avg	1.00	1.00	1.00	11810



شکل ۱۳: confusion matrix برای احساسات با داده های آموزش با KNN با ویژگی های rhythmic

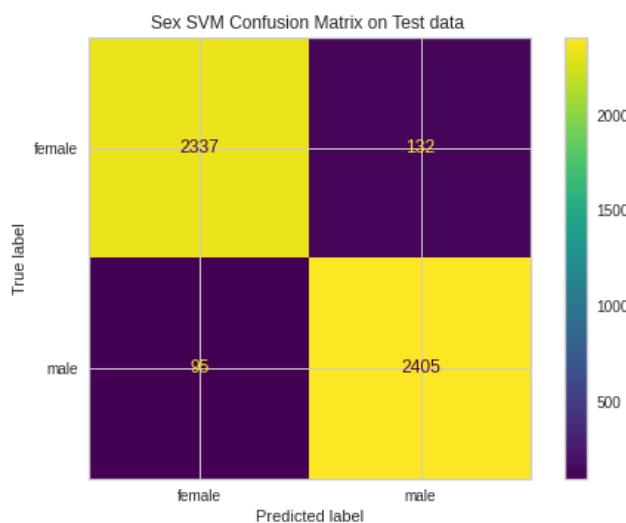
## SVM ۲.۴

ایده‌ی کلی SVM یا Support vector (ماشین بردار پشتیبان) به این صورت است که مرز تصمیم در ناحیه‌ای قرار بگیرد که بیشترین فاصله را از نزدیکترین داده‌های دو کلاس به خط داشته باشد. در واقع این داده‌های نزدیک همان بردارهای پشتیبان هستند. با توجه به همین تعریف میتوان متوجه شد که این روش به تعداد داده‌های هر کلاس توجهی ندارد و صرفاً با توجه به بردارهای پشتیبان، مرز تصمیم را بیان میکند (موضوعی که در مقایسه با logistic regression

متفاوت است). البته این تعریف در مثالی صادق است که دو کلاس کاملاً از یکدیگر مجزا باشند. ماشین بردار پشتیبان با اندکی تغییر قابل اعمال بر روی هر نوع ترکیبی از داده‌ها است که در اینجا از بیان تعاریف ریاضی آنها خودداری می‌کنیم. در این روش هایپر پارامترهایی وجود دارد که ما در برنامه به ازای ترکیب‌های مختلفی از آنها، طبقه‌بندی را اعمال کردیم و در نهایت بهترین این ترکیب‌ها را گزارش داده‌ایم.

## ۱.۲.۴ طبقه‌بندی جنسیت با SVM با کمک Spectrogram Features

```
best parameters{'C': 4, 'gamma': 'scale', 'kernel': 'rbf'}
max validation score 0.9555461473327689
Sex SVM accuracy score on Test data for Sex : 0.9543
Sex SVM classification report on test data
precision    recall    f1-score   support
female       0.96      0.95      0.95      2469
  male       0.95      0.96      0.95      2500
accuracy           0.95      0.95      0.95      4969
macro avg       0.95      0.95      0.95      4969
weighted avg     0.95      0.95      0.95      4969
```

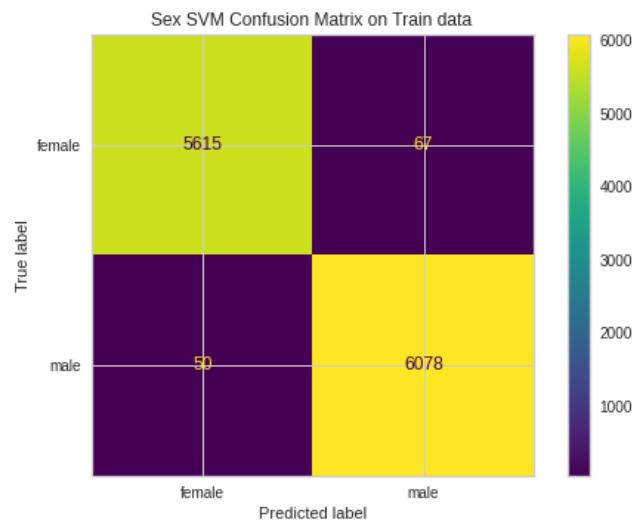


شکل ۱۴: confusion matrix برای جنسیت با داده‌های آزمون با SVM با ویژگی‌های Spectrogram

Sex SVM accuracy score on Train data for Sex : 0.9901

Sex SVM classification report on train data

	precision	recall	f1-score	support
female	0.99	0.99	0.99	5682
male	0.99	0.99	0.99	6128
accuracy			0.99	11810
macro avg	0.99	0.99	0.99	11810
weighted avg	0.99	0.99	0.99	11810



شکل ۱۵: confusion matrix برای جنسیت با داده های آموزش با SVM با ویژگی های Spectrogram

## ۲.۲.۴ طبقه بندی جنسیت با SVM با کمک Rythmic Features

```
best parameters{'C': 1, 'gamma': 'auto', 'kernel': 'rbf'}
max validation score 0.6043183742591024
Sex SVM accuracy score on Test data for Sex : 0.6146
Sex SVM classification report on test data
precision    recall    f1-score   support
female       0.57      0.61      0.59      2273
male        0.65      0.62      0.63      2696
accuracy          0.61      0.61      0.61      4969
macro avg       0.61      0.61      0.61      4969
weighted avg    0.62      0.61      0.62      4969
```

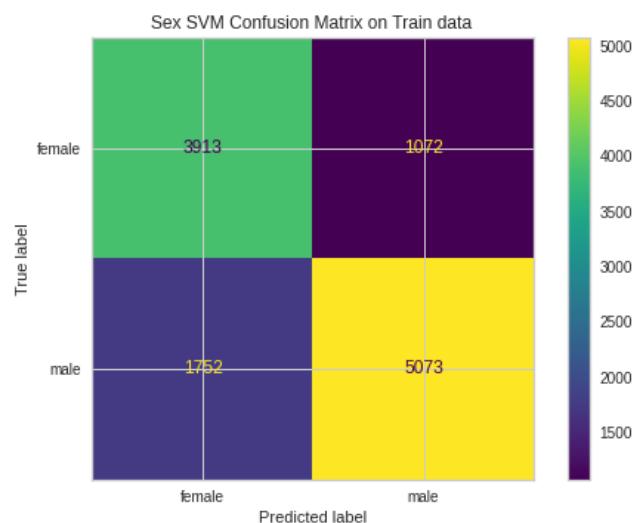


شکل ۱۶: confusion matrix برای جنسیت با داده های آزمون با SVM با ویژگی های rhythmic

Sex SVM accuracy score on Train data for Sex : 0.7609

Sex SVM classification report on train data

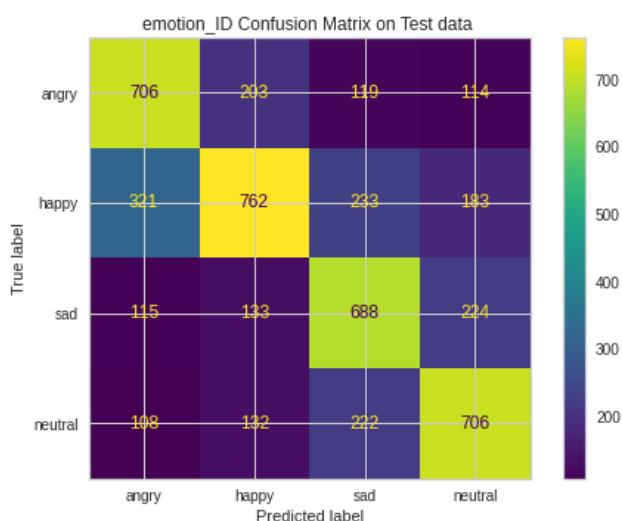
	precision	recall	f1-score	support
female	0.69	0.78	0.73	4985
male	0.83	0.74	0.78	6825
accuracy			0.76	11810
macro avg	0.76	0.76	0.76	11810
weighted avg	0.77	0.76	0.76	11810



شکل ۱۷: confusion matrix برای جنسیت با داده های آموزش با SVM با ویژگی های rhythmic

### ۳.۲.۴ طبقه بندی احساسات با SVM با کمک Spectrogram Features

```
best parameters{'C': 4, 'gamma': 'auto', 'kernel': 'rbf'}
max validation score 0.5669780280956226
accuracy score on Test data for emotion_ID : 0.5760
emotion_ID classification report on test data
      precision    recall   f1-score   support
angry        0.56     0.62     0.59     1142
happy        0.62     0.51     0.56     1499
sad          0.55     0.59     0.57     1160
neutral      0.58     0.60     0.59     1168
accuracy           0.58         -         -     4969
macro avg       0.58     0.58     0.58     4969
weighted avg    0.58     0.58     0.58     4969
```



شکل ۱۸: confusion matrix برای احساسات با داده های آزمون با SVM با ویژگی های Spectrogram

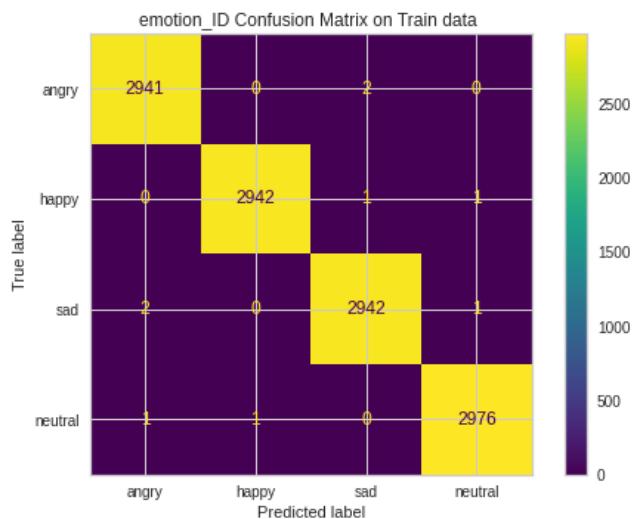
```

accuracy score on Train data for emotion_ID : 0.9992
emotion_ID classification report on train data
      precision    recall   f1-score   support

      angry      1.00      1.00      1.00      2943
      happy      1.00      1.00      1.00      2944
      sad       1.00      1.00      1.00      2945
      neutral     1.00      1.00      1.00      2978

      accuracy          1.00      1.00      1.00      11810
      macro avg       1.00      1.00      1.00      11810
      weighted avg    1.00      1.00      1.00      11810

```



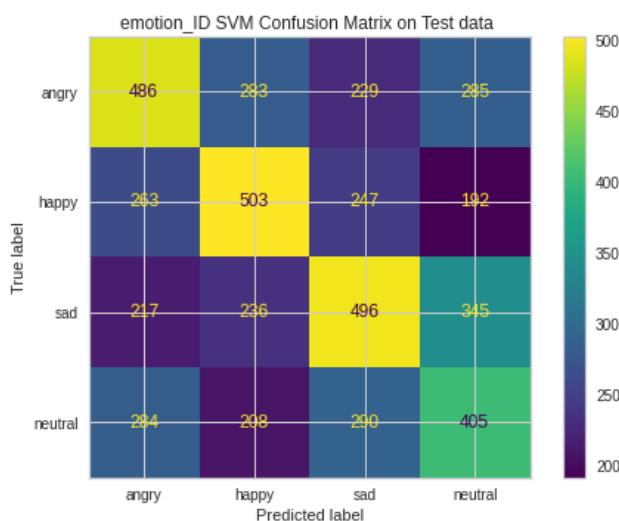
شکل ۱۹: confusion matrix برای احساسات با داده های آموزش با SVM با ویژگی های Spectrogram

## ۴.۲.۴ طبقه بندی احساسات با SVM با کمک Rythmic Features

```

best parameters{'C': 1, 'gamma': 'scale', 'kernel': 'rbf'}
max validation score0.3746824724809484
emotion_ID SVMaccuracy score on Test data for emotion_ID : 0.3804
emotion_ID SVM classification report on test data
      precision    recall   f1-score   support
angry        0.39     0.38     0.38    1283
happy        0.41     0.42     0.41    1205
sad          0.39     0.38     0.39    1294
neutral      0.33     0.34     0.34    1187
accuracy           0.38     0.38     0.38    4969
macro avg       0.38     0.38     0.38    4969
weighted avg    0.38     0.38     0.38    4969

```

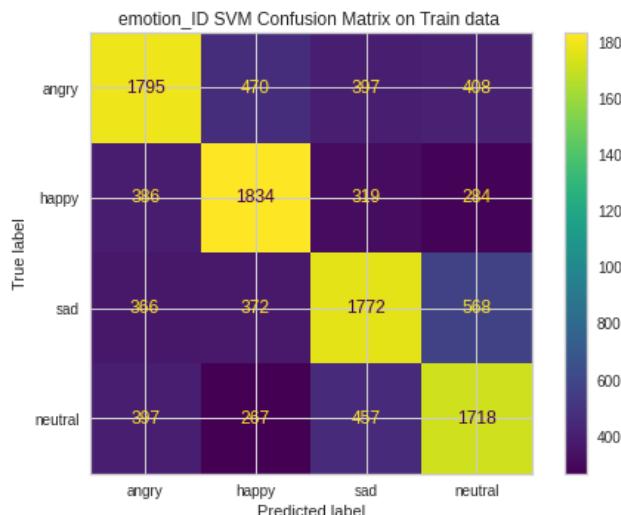


شکل ۲۰: confusion matrix برای احساسات با داده های آزمون با SVM با ویژگی های rhythmic

```
emotion_ID SVMaccuracy score on Train data for emotion_ID : 0.6028
```

```
emotion_ID SVM classification report on train data
```

	precision	recall	f1-score	support
angry	0.61	0.58	0.60	3070
happy	0.62	0.65	0.64	2823
sad	0.60	0.58	0.59	3078
neutral	0.58	0.61	0.59	2839
accuracy			0.60	11810
macro avg	0.60	0.60	0.60	11810
weighted avg	0.60	0.60	0.60	11810



شکل ۲۱: confusion matrix برای احساسات با داده های آموزش با SVM با ویژگی های rhythmic

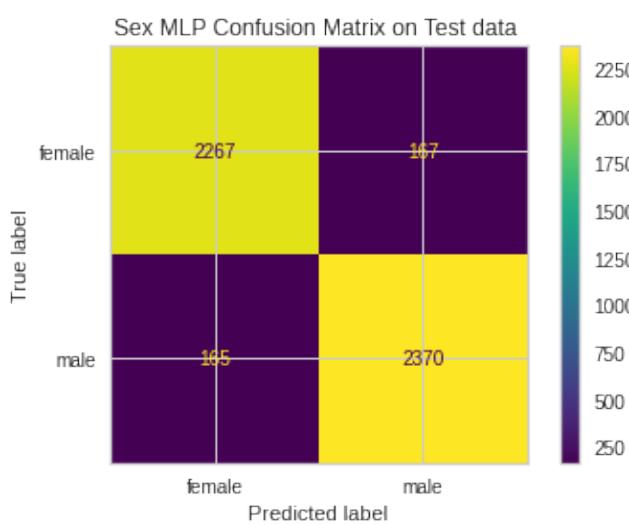
## MLP ۳.۴

این روش که مخفف multilayer perceptron است، یکی از انواع شبکه های عصبی محسوب میشود. شبکه های عصبی را میتوان از بهترین روش های یادگیری ماشین دانست که طبیعتاً میتوانند عمل طبقه بندی را نیز به خوبی انجام دهند. به طور خلاصه در این شبکه ها با استفاده از تعریفی از خطأ، خطای هر داده‌ی محاسبه میشود و در ادامه وزن های شبکه‌ی عصبی مجدد تعیین میشود. شبکه های عصبی جزئیات زیادی دارند که بیان تمامی آنها از حوصله‌ی این متن

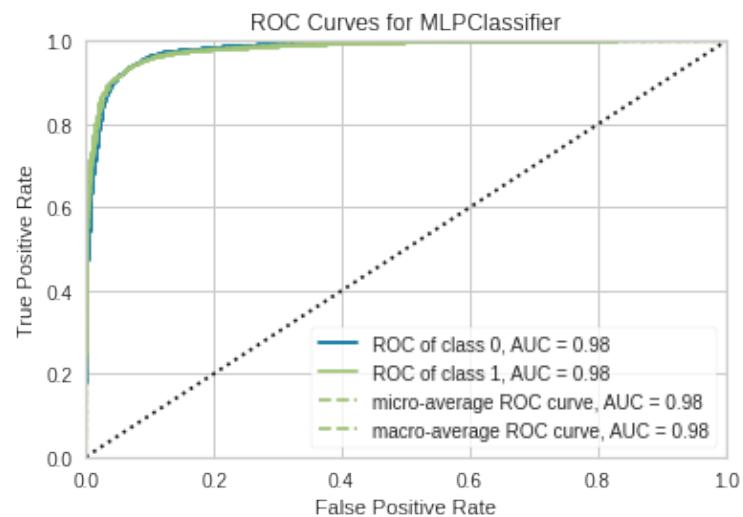
خارج است و تنها دو تا از مهمترین آنها را توضیح میدهیم. شاید بتوان مهمترین پارامتر شبکه های عصبی و اختصاصا MLP را تعداد لایه های مخفی آن دانست. لایه های مخفی در واقع لایه هایی هستند که در بین ورودی و خروجی قرار میگیرند. این تعداد که در واقع هایپر پارامتر این روش محاسب میشود، در برنامه تعیین شده. در دیگر موارد مهم در شبکه های عصبی، تعداد نورون های هر کدام از لایه های مخفی است. برای این پارامتر عدد ۱۰۰ را در نظر گرفتیم.

#### ۱.۳.۴ طبقه بندی جنسیت با MLP با کمک Spectrogram Features

```
best parameters{'activation': 'relu', 'hidden_layer_sizes': 19, 'solver': 'adam'}
max validation score0.9347163420829805
/usr/local/lib/python3.7/dist-packages/sklearn/neural_network/_multilayer_perceptron
ConvergenceWarning,
Sex MLP accuracy score on Test data for Sex : 0.9332
SexMLP classification report on test data
precision    recall   f1-score   support
female      0.93     0.93     0.93    2434
male        0.93     0.93     0.93    2535
accuracy          0.93     0.93     0.93    4969
macro avg      0.93     0.93     0.93    4969
weighted avg    0.93     0.93     0.93    4969
```



شکل ۲۲: confusion matrix برای جنسیت با داده های آزمون با MLP با ویژگی های Spectrogram



شکل ۲۳: نمودار ROC برای جنسیت با طبقه بند MLP با ویژگی های Spectrogram

```

Sex MLPaccuracy score on Train data for Sex : 0.9670
Sex MLP classification report on train data
precision      recall      f1-score   support
female         0.96        0.97        0.97      5629
male           0.97        0.97        0.97      6181
accuracy       0.97        0.97        0.97     11810
macro avg      0.97        0.97        0.97     11810
weighted avg   0.97        0.97        0.97     11810

```



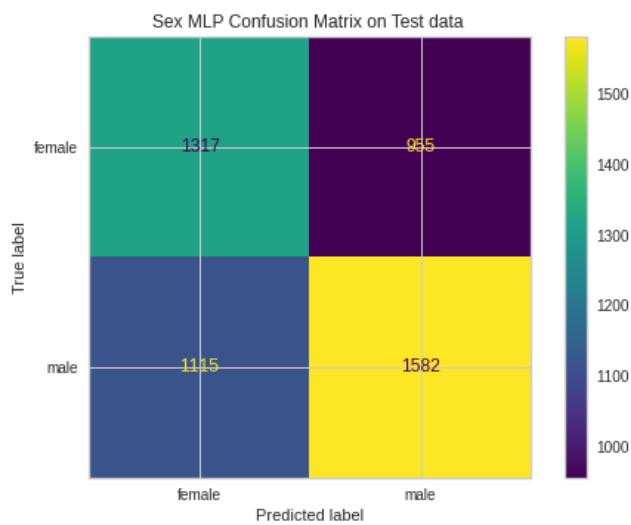
شکل ۲۴: confusion matrix برای جنسیت با داده های آموزش با MLP با ویژگی های Spectrogram

#### ۲.۳.۴ طبقه بندی جنسیت با MLP با کمک Rythmic Features

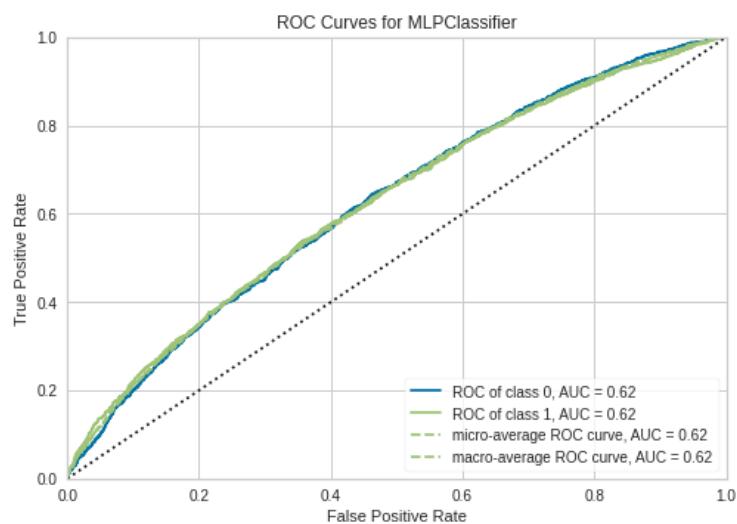
```

best parameters{'activation': 'logistic', 'hidden_layer_sizes': 15, 'solver': 'adam'}
max validation score0.5935646192091842
/usr/local/lib/python3.7/dist-packages/sklearn/neural_network/_multilayer_perceptron.py:59:
ConvergenceWarning,
Sex MLP accuracy score on Test data for Sex : 0.5834
Sex MLP classification report on test data
precision    recall    f1-score   support
female      0.54      0.58      0.56     2272
          male      0.62      0.59      0.60     2697
accuracy           0.58      0.58      0.58     4969
macro avg       0.58      0.58      0.58     4969
weighted avg    0.59      0.58      0.58     4969

```



شکل ۲۵: confusion matrix برای جنسیت با داده های آزمون با MLP با ویژگی های rhythmic

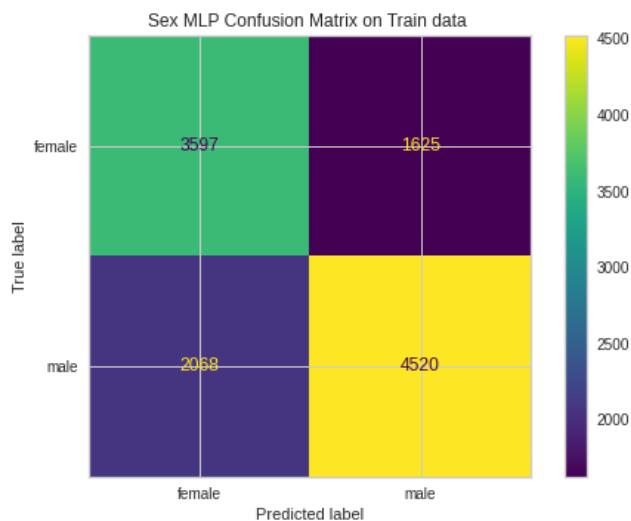


شکل ۲۶: نمودار ROC برای جنسیت با طبقه بند MLP با ویژگی های rhythmic

Sex MLP accuracy score on Train data for Sex : 0.6873

Sex MLP classification report on train data

	precision	recall	f1-score	support
female	0.63	0.69	0.66	5222
male	0.74	0.69	0.71	6588
accuracy			0.69	11810
macro avg	0.69	0.69	0.69	11810
weighted avg	0.69	0.69	0.69	11810



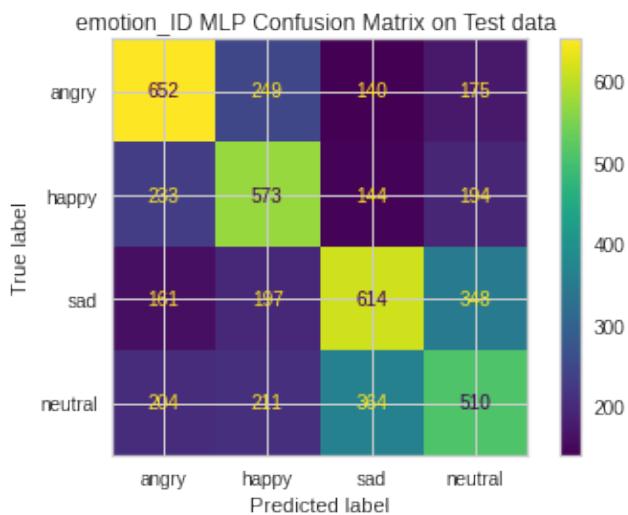
شکل ۲۷: confusion matrix برای جنسیت با داده های آموزش با MLP با ویژگی های rhythmic

### ۳.۳.۴ طبقه بندی احساسات با MLP با کمک Spectrogram Features

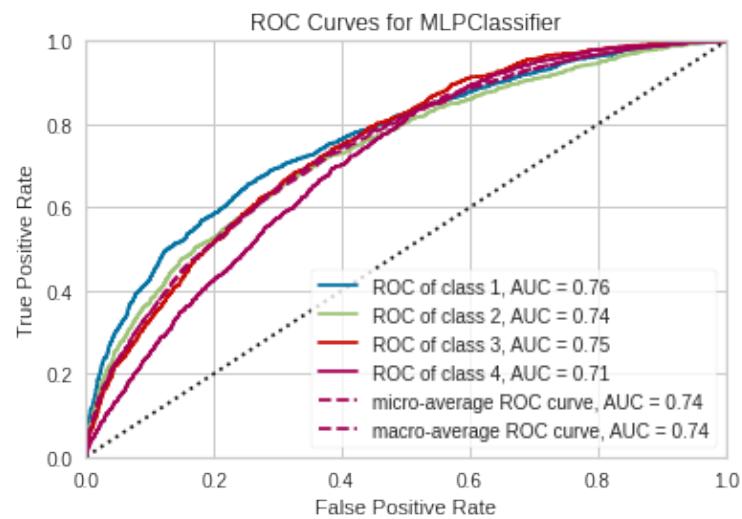
```

best parameters{'activation': 'relu', 'hidden_layer_sizes': 19, 'solver': 'adam'}
max validation score0.4635901778154107
/usr/local/lib/python3.7/dist-packages/sklearn/neural_network/_multilayer_perceptron.py:696:
    ConvergenceWarning,
emotion_ID MLPaccuracy score on Test data for emotion_ID : 0.4727
emotion_IDMLP classification report on test data
precision    recall   f1-score   support
angry       0.52      0.54      0.53     1216
happy       0.47      0.50      0.48     1144
sad         0.49      0.47      0.48     1320
neutral     0.42      0.40      0.41     1289
accuracy                           0.47      4969
macro avg       0.47      0.47      0.47     4969
weighted avg    0.47      0.47      0.47     4969

```



شکل ۲۸: confusion matrix برای احساسات با داده های آزمون با MLP با ویژگی های Spectrogram

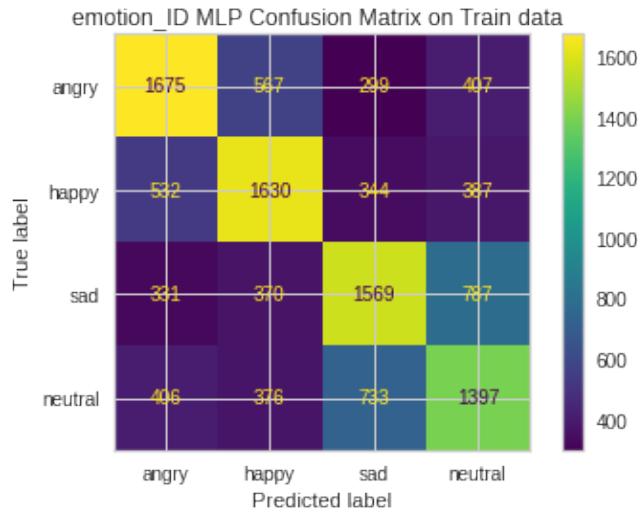


شکل ۲۹: نمودار ROC برای احساسات با طبقه بند MLP با ویژگی های Spectrogram

```

emotion_ID MLPaccuracy score on Train data for emotion_ID : 0.5310
emotion_ID MLP classification report on train data
      precision    recall   f1-score   support
angry        0.57     0.57     0.57     2948
happy        0.55     0.56     0.56     2893
sad          0.53     0.51     0.52     3057
neutral      0.47     0.48     0.47     2912
accuracy
macro avg       0.53     0.53     0.53    11810
weighted avg    0.53     0.53     0.53    11810

```

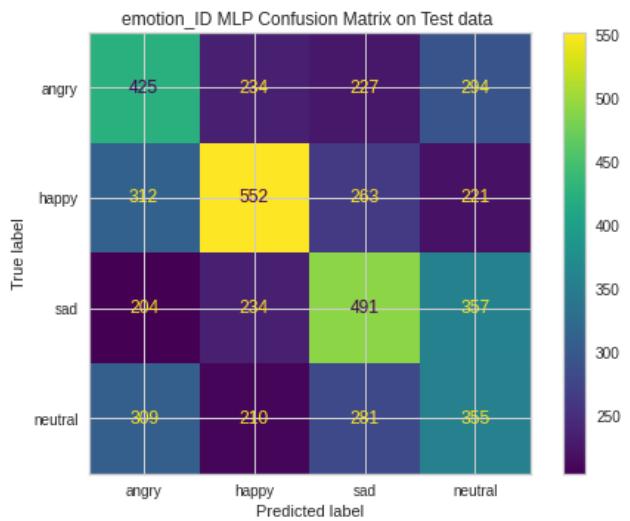


شکل ۳۰: confusion matrix برای احساسات با داده های آموزش با MLP با ویژگی های Spectrogram

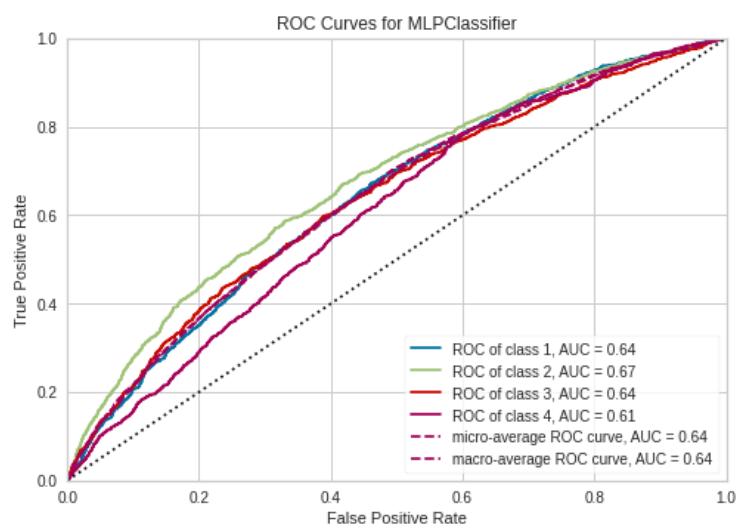
#### ۴.۳.۴ طبقه بندی احساس با کمک MLP با Rythmic Features

```

best parameters{'activation': 'logistic', 'hidden_layer_sizes': 16, 'solver': 'sgd'}
max validation score0.37298866789467994
/usr/local/lib/python3.7/dist-packages/sklearn/neural_network/_multilayer_perceptron.
    ConvergenceWarning,
emotion_ID MLPaccuracy score on Test data for emotion_ID : 0.3669
emotion_ID MLP classification report on test data
      precision    recall   f1-score   support
      angry       0.34      0.36      0.35     1180
      happy       0.45      0.41      0.43     1348
      sad        0.39      0.38      0.39     1286
      neutral     0.29      0.31      0.30     1155
      accuracy         -       0.37      0.37     4969
      macro avg     0.37      0.36      0.37     4969
      weighted avg  0.37      0.37      0.37     4969
  
```



شکل ۳۱: confusion matrix برای احساسات با داده های آزمون با MLP با ویژگی های rhythmic

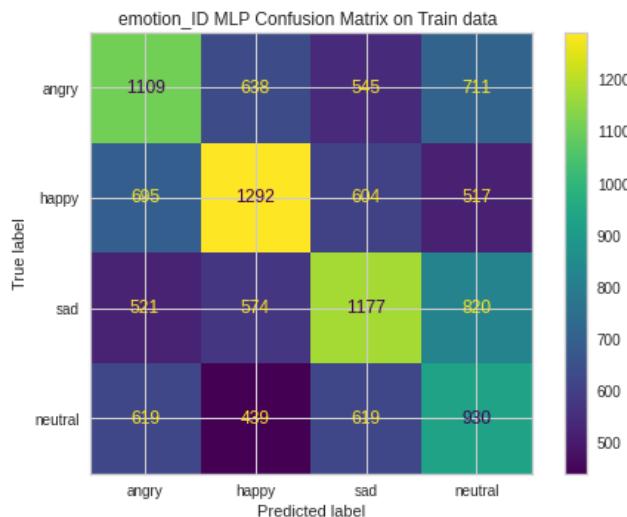


شکل ۳۲: نمودار ROC برای احساسات با طبقه بند MLP با ویژگی های rhythmic

```
emotion_ID MLPaccuracy score on Train data for emotion_ID : 0.3817
```

```
emotion_ID MLP classification report on train data
      precision    recall   f1-score   support
```

angry	0.38	0.37	0.37	3003
happy	0.44	0.42	0.43	3108
sad	0.40	0.38	0.39	3092
neutral	0.31	0.36	0.33	2607
accuracy			0.38	11810
macro avg	0.38	0.38	0.38	11810
weighted avg	0.38	0.38	0.38	11810



شکل ۳۳: confusion matrix برای احساسات با داده های آموزش با MLP با ویژگی های rhythmic

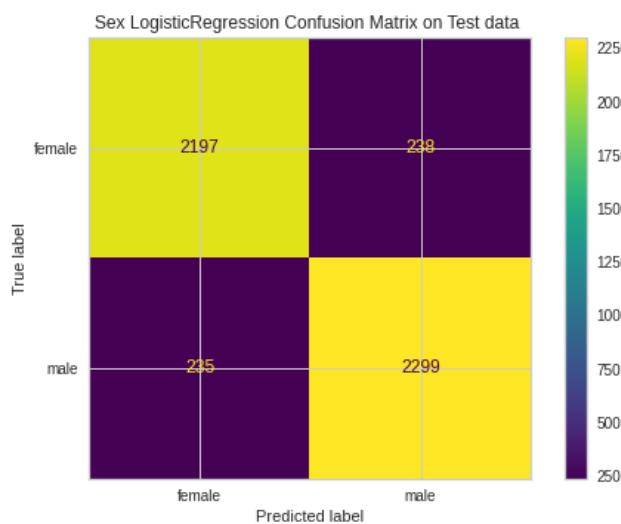
## LOGISTIC REGRESSION ۴.۴

در این طبقه بند، با توجه به قانون بیز، رابطه ای استخراج خواهد شد که احتمال هر کلاس را در هر کجای صفحه‌ی مختصات اندازه‌گیری می‌کند. با استفاده از قانون بیز تابع ای با عنوان تابع سیگموید به دست خواهد آمد که خروجی آن از جنس احتمال خواهد بود. برای هر نقطه، کلاس نهایی آنی خواهد بود که احتمال بزرگتری داشته باشد. در این روش با توجه به میزان داده های موجود در قسمت از صفحه‌ی مختصات، احتمال آن کلاس خاص از داده ها، در آن

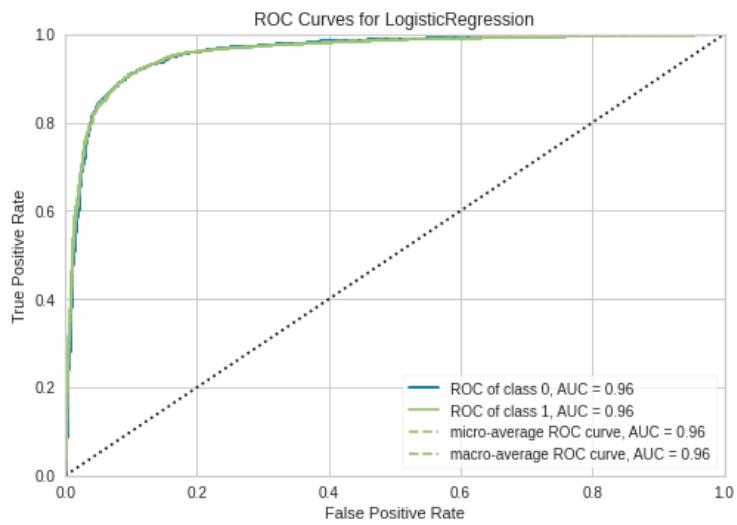
قسمت خاص از صفحه‌ی مختصات، بالاتر خواهد رفت. این روش به وزن داده‌ها توجه می‌کند. تعداد داده‌های بیشتر، احتمال را بیشتر می‌کند. این موضوع در مقایسه با SVM متفاوت است.

#### ۱.۴.۴ طبقه‌بندی جنسیت با LOGISTIC REGRESSION با کمک Spectrogram Features

```
best parameters{'solver': 'newton-cg'}
max validation score0.9089754445385265
Sex LogisticRegressionaccuracy score on Test data for Sex : 0.9048
Sex LogisticRegression classification report on test data
precision    recall   f1-score   support
female      0.90      0.90      0.90      2435
male        0.91      0.91      0.91      2534
accuracy          0.90      0.90      0.90      4969
macro avg       0.90      0.90      0.90      4969
weighted avg    0.90      0.90      0.90      4969
```



شکل ۳۴: confusion matrix برای جنسیت با داده‌های آزمون با LOGISTIC REGRESSION با ویژگی‌های Spectrogram

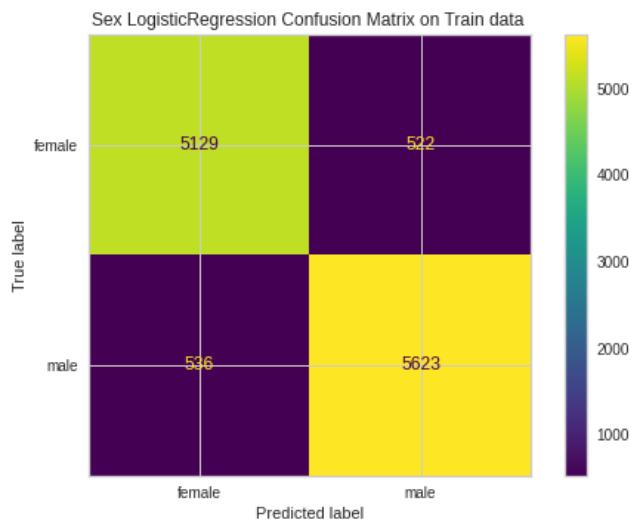


شکل ۳۵: نمودار ROC برای جنسیت با طبقه بند LOGISTIC REGRESSION با ویژگی های Spectrogram

```

Sex LogisticRegression accuracy score on Train data for Sex : 0.9104
Sex LogisticRegression classification report on train data
precision    recall   f1-score   support
female       0.91      0.91      0.91      5651
male        0.92      0.91      0.91      6159
accuracy           0.91      0.91      0.91     11810
macro avg       0.91      0.91      0.91     11810
weighted avg    0.91      0.91      0.91     11810

```



شکل ۳۶: confusion matrix برای جنسیت با داده های آموزش با LOGISTIC REGRESSION با ویژگی های Spectrogram

#### ۲.۴.۴ طبقه بندی جنسیت با کمک LOGISTIC REGRESSION Rythmic Features

```

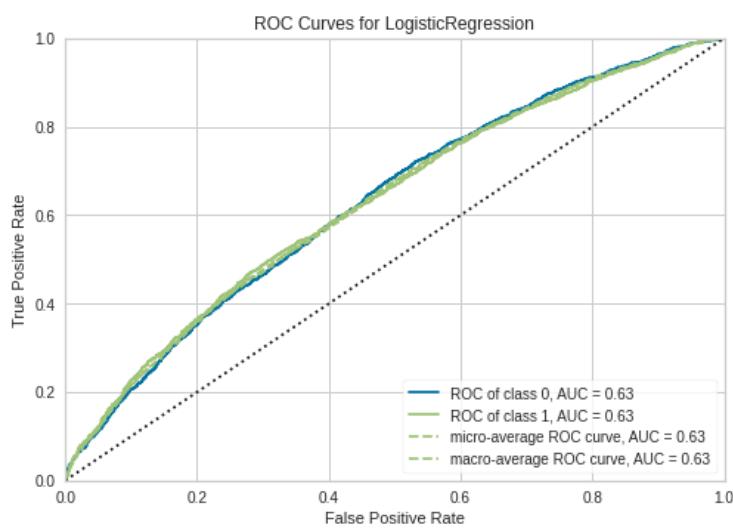
best parameters{'solver': 'saga'}
max validation score0.5909398814563929
Sex LogisticRegressionaccuracy score on Test data for Sex : 0.5854
Sex LogisticRegression classification report on test data
precision    recall    f1-score   support
female      0.49      0.59      0.54     2008
          male      0.68      0.58      0.63     2961

accuracy                           0.59      4969
macro avg       0.58      0.59      0.58     4969
weighted avg    0.60      0.59      0.59     4969

```



شکل ۳۷: confusion matrix برای جنسیت با داده های آزمون با LOGISTIC REGRESSION با ویژگی های rhythmic

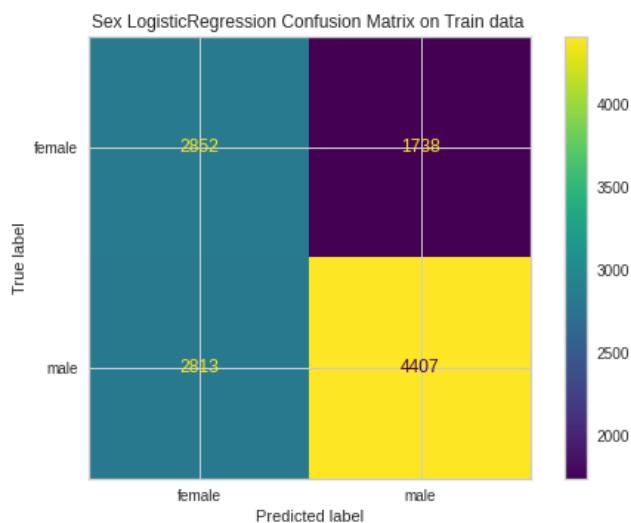


شکل ۳۸: نمودار ROC برای جنسیت با طبقه بند LOGISTIC REGRESSION با ویژگی های rhythmic

Sex LogisticRegression accuracy score on Train data for Sex : 0.6146

Sex LogisticRegression classification report on train data  
precision recall f1-score support

female	0.50	0.62	0.56	4590
male	0.72	0.61	0.66	7220
accuracy			0.61	11810
macro avg	0.61	0.62	0.61	11810
weighted avg	0.63	0.61	0.62	11810



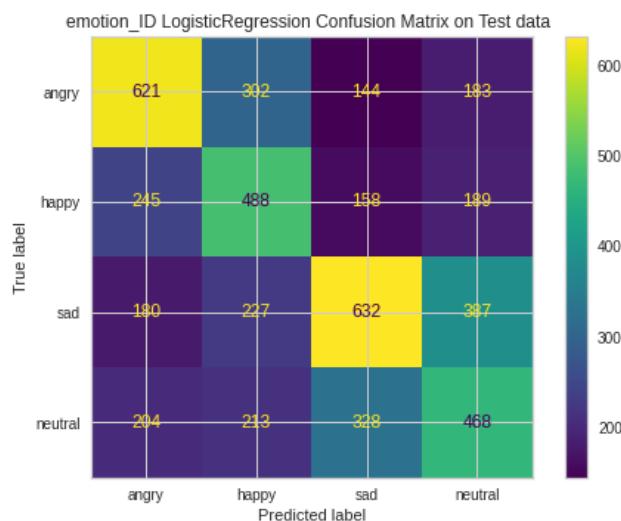
شکل ۳۹: confusion matrix برای جنسیت با داده های آموزش با LOGISTIC REGRESSION با ویژگی های rhythmic

## ۳.۴.۴ طبقه بندی احساس با LOGISTIC REGRESSION با کمک Spectrogram Features

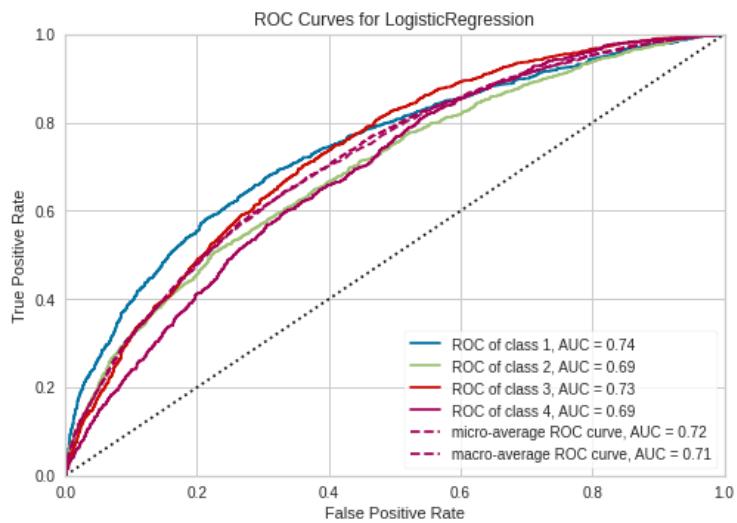
```

best parameters{'solver': 'newton-cg'}
max validation score 0.4292972057578323
emotion_ID LogisticRegression accuracy score on Test data for emotion_ID : 0.4446
emotion_ID LogisticRegression classification report on test data
precision    recall   f1-score   support
angry        0.50     0.50     0.50      1250
happy        0.40     0.45     0.42      1080
sad          0.50     0.44     0.47      1426
neutral      0.38     0.39     0.38      1213
accuracy           0.44     0.44     0.44      4969
macro avg       0.44     0.44     0.44      4969
weighted avg    0.45     0.44     0.45      4969

```



شکل ۴۰: confusion matrix برای احساسات با داده های آزمون با LOGISTIC REGRESSION با ویژگی های Spectrogram

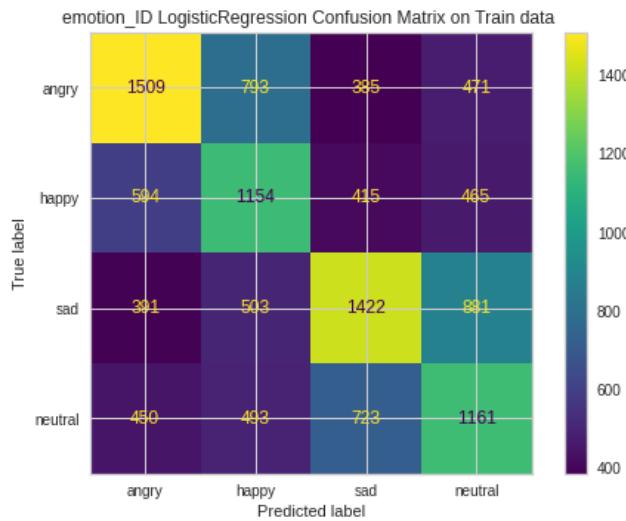


شکل ۴۱: نمودار ROC برای احساسات با طبقه بند LOGISTIC REGRESSION با ویژگی Spectrogram های

```

emotion_ID LogisticRegressionaccuracy score on Train data for emotion_ID : 0.4442
emotion_ID LogisticRegression classification report on train data
      precision    recall   f1-score   support
angry        0.51     0.48     0.49     3158
happy        0.39     0.44     0.41     2628
sad          0.48     0.44     0.46     3197
neutral      0.39     0.41     0.40     2827
accuracy           0.44     0.44     0.44    11810
macro avg       0.44     0.44     0.44    11810
weighted avg    0.45     0.44     0.45    11810

```



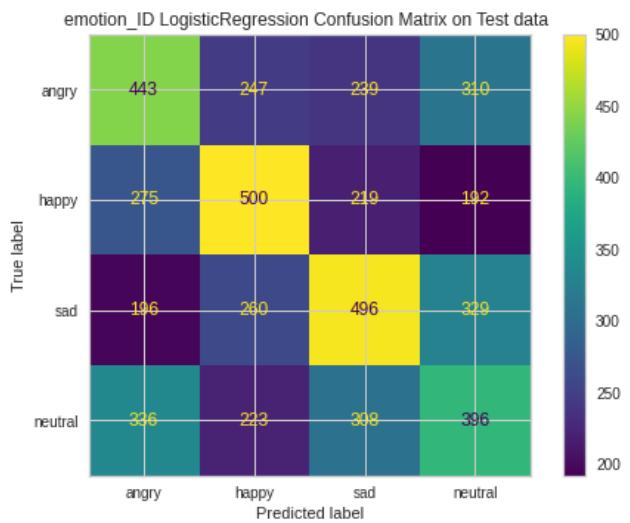
شکل ۴۲: confusion matrix برای احساسات با داده های آموزش با LOGISTIC REGRESSION با ویژگی های SPECTROGRAM

#### ۴.۴.۴ طبقه بندی احساس با LOGISTIC REGRESSION با کمک Rythmic Features

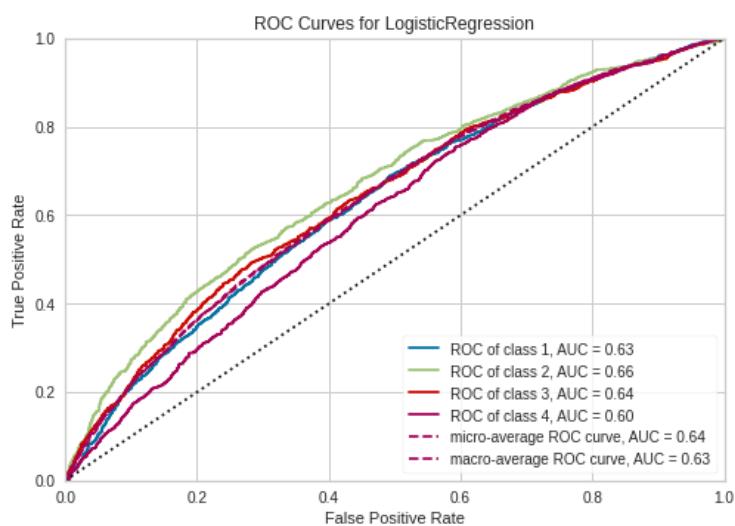
```

best parameters{'solver': 'newton-cg'}
max validation score 0.3690093141405589
emotion_ID LogisticRegression accuracy score on Test data for emotion_ID : 0.3693
emotion_ID LogisticRegression classification report on test data
      precision    recall   f1-score   support
angry       0.35      0.36      0.36     1239
happy       0.41      0.42      0.41     1186
sad         0.39      0.39      0.39     1281
neutral     0.32      0.31      0.32     1263
accuracy          0.37      0.37      0.37     4969
macro avg       0.37      0.37      0.37     4969
weighted avg     0.37      0.37      0.37     4969

```



شکل ۴۲: confusion matrix برای احساسات با داده های آزمون با LOGISTIC REGRESSION با ویژگی های rhythmic



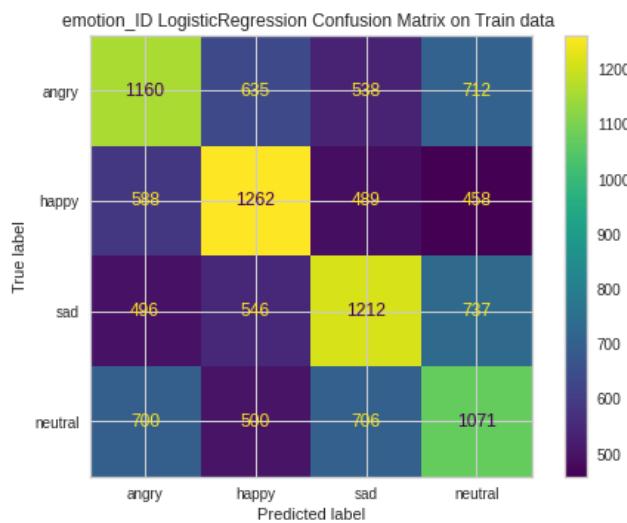
شکل ۴۴: نمودار ROC برای احساسات با طبقه بند LOGISTIC REGRESSION با ویژگی های rhythmic

```

emotion_ID LogisticRegressionaccuracy score on Train data for emotion_ID : 0.3984
emotion_ID LogisticRegression classification report on train data
      precision    recall   f1-score   support
angry        0.39     0.38     0.39    3045
happy        0.43     0.45     0.44    2797
sad          0.41     0.41     0.41    2991
neutral      0.36     0.36     0.36    2977

accuracy         0.40     0.40     0.40    11810
macro avg       0.40     0.40     0.40    11810
weighted avg    0.40     0.40     0.40    11810

```



شکل ۴۵: confusion matrix برای احساسات با داده های آموزش با LOGISTIC REGRESSION با ویژگی های rhythmic

## ۵.۴ جمع بندی طبقه بندی ها

	KNN	SVM	MLP	Logistic Regression
Test/Sex accuracy/Spectrogram Features	0.96	0.95	0.93	0.9
Train/Sex accuracy/Spectrogram Features	1	0.99	0.97	0.91
Test/emotion_ID accuracy/Spectrogram Features	0.6	0.58	0.47	0.44
Train/emotion_ID accuracy/Spectrogram Features	1	1	0.53	0.44
Test/Sex accuracy/Rhythmic Features	0.56	0.61	0.58	0.59
Train/Sex accuracy/Rhythmic Features	1	0.76	0.69	0.61
Test/emotion_ID accuracy/Rhythmic Features	0.31	0.38	0.37	0.37
Train/emotion_ID accuracy/Rhythmic Features	1	0.6	0.38	0.4

شکل ۴۶: مقایسه دقت طبقه بندها برای دو نوع ویژگی

Spectrogram Features:

Sex

همان طور که میبینیم هر چهار مدل، به خوبی Train شده اند و داده های Test را با accuracy بالایی طبقه بندی کرده اند.

Emotion\_ID

همان طور که میبینیم، با اینکه در مدل های KNN و SVM عملکرد به خوبی روی Train آموخته دیده اند و با اینکه از روش k fold cross validation استفاده کردیم، مدل ها Logistic Regression overfit شده و در داده های Test دقت کمی دارند. مدل های MLP و در Training data هم دقت خوبی ندارند و بر روی Test data، به accuracy زیر 0.5 رسیده اند. بنابر این با استفاده از Spectrogram Features توانستیم جنسیت افراد را به خوبی کلاس بندی کنیم. همچنین به کمک KNN و SVM نیز توانستیم با دقتی بالاتر از 0.5 احساسات را طبقه بندی کنیم.

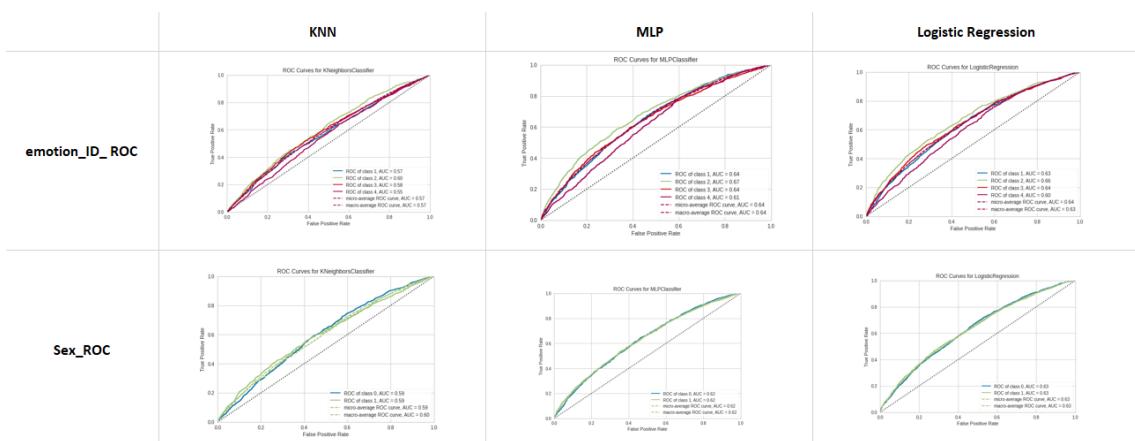
Tempogram Features:

Sex

همان طور که میبینیم در هر چهار مدل دقت Test data پایین و به 0.5 نزدیک است. همچنین در MLP و Logistic Regression دقت Train data نیز پایین است و با این فیجرها مدل به خوبی train نشده است. طبقه بند KNN نیز دچار overfit شده است.

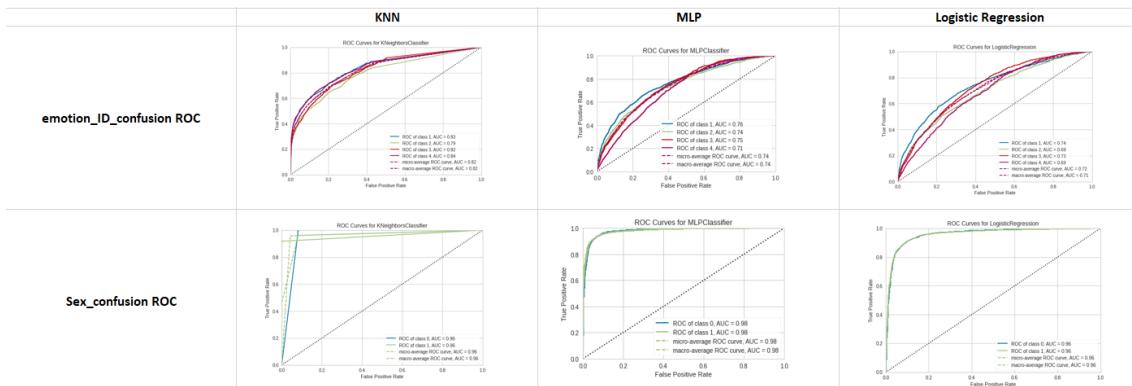
Emotion\_ID

همان طور که میبینیم دقت تمام مدل ها زیر ۰.۵ است و نمی توانند داده ها را classify کنند. همچنین KNN دچار overfit شده و سایر مدل ها برای Train data هم دقت پایینی دارند. بنابر این با استفاده از Tempogram Features نتوانستیم طبقه بند خوبی بسازیم.



شکل ۴۷: مقایسه نمودار ROC برای Rythmic Features

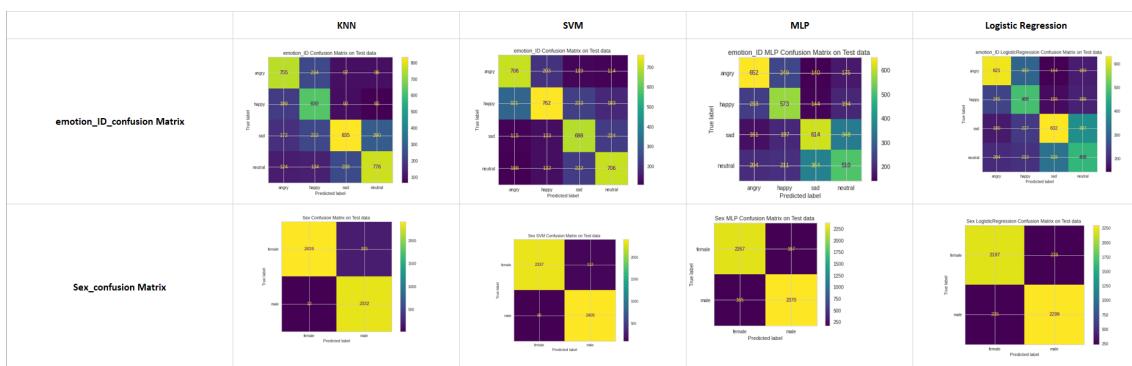
همانطور که میبینیم نمودار مدل ها بسیار به نیمساز نزدیک است و این نشان دهنده ی عملکرد ضعیف مدل ها هم در طبقه بندی جنسیت و هم در طبقه بندی احساسات است. بنابر این با استفاده از Tempogram Features نتوانستیم طبقه بند خوبی بسازیم.



شکل ۴۸: مقایسه نمودار ROC برای Spectrogram Features

همانطور که میبینیم برای جنسیت، نمودار ROC برای هر سه طبقه بند، فاصله‌ی زیادی از نیم ساز دارد و TPR به سرعت افزایش می‌یابد. که نشان دهنده‌ی عملکرد خوب طبقه بند‌ها در کلاس‌بندی جنسیت است.

برای احساسات نیز میبینیم که بیشترین فاصله از نیمساز را KNN دارد و پس از آن، MLP و Logistic Regression بیشترین فاصله‌های را دارند.



شکل ۴۹: مقایسه confusion matrix برای Spectrogram Features

همانطور که میبینیم در طبقه‌بندی جنسیت جمع قطر اصلی بیشتر از طبقه‌بندی احساسات

است. و نتایج قبلی را برای عملکرد طبقه بند ها تایید میکند. همچنین با دقت به emotion\_ID confusion matrix میبینیم که misclassification بین دو احساس عصبانی و خوش حال و همچنین بین دو احساس neutral و sad زیاد اتفاق افتاده است. با توجه به اینکه در هنگام خوش حالی و ناراحتی هیجانات و صوت انسان بالاست و در هنگام ناراحتی و خشی هیجانات و صوت انسان پایین است، ممکن است طبقه بند ها به خوبی عصبانیت و خوش حالی را از یکدیگر و ناراحتی و خشی بودن را از یکدیگر نتوانسته اند تشخیص دهند.

## ۵ خوشه بندی

### ۱.۵ روش های خوشه بندی

خوشه بندی، فرآیندی است که به کمک آن می‌توان مجموعه‌ای از اشیاء را به گروه‌های مجزا افراز کرد. هر افزار یک خوشه نامیده می‌شود. اعضاء هر خوشه با توجه به ویژگی‌هایی که دارند به یکدیگر بسیار شبیه هستند و در عوض میزان شباهت بین خوشه‌ها کمترین مقدار است. در چنین حالتی هدف از خوشه بندی، نسبت دادن برچسب‌هایی به اشیاء است که نشان دهنده عضویت هر شیء به خوشه است.

برای خوشه بندی، با استفاده از روش PCA، ابعاد داده‌ها را به بعد کاهش میدهیم سپس با استفاده از ۳ روش Agglomerative Clustering، GMM Clustering، K\\_means، به ۲، ۴، و ۱۰ خوشه، طبقه بندی می‌کنیم. پس از آن معیارهای ارزشیابی هر خوشه بند را گزارش کرده و خوشه‌ها را در ۲ و ۳ بعد رسم می‌کنیم. برای رسم خوشه‌ها، ابتدا بعد نقاط را به ۲ و ۳ جهتی که بیشترین واریانس را دارند، نگاشت کرده و سپس نمودار را رسم می‌کنیم. برای بررسی معیارهای supervised، برای کلاستر با تعداد خوشه‌های ۲، ۴، و ۱۰ به ترتیب لیل‌های جنسیت، احساسات و شماره متن را میدهیم. زیرا حدس میزان شباهت و تفاوت داده‌ها با یکدیگر بر این اساس را باشد. معیارهای گزارش شده عبارتند از:

#### Homogeneity ۱.۱.۵

این معیار supervised است و نشان می‌دهد که تا چه میزان سمپل‌های داخل هر خوشه به یکدیگر شباهت دارند. معادله آن به صورت زیر در می‌آید:

$$H(C|K) = - \sum_{c,k} \frac{n_{ck}}{N} \log \left( \frac{n_{ck}}{n_k} \right)$$

$$h = 1 - \frac{H(C|K)}{H(C)}$$

که مقدار  $\frac{n_{c_k}}{n_k}$  نسبت لیل های  $c$  در کلاستر  $k$  ام به تمام سمپل های کلاستر  $k$  ام می باشد. بین ۰ و ۱ می باشد. هر چه  $h$  به ۱ نزدیک تر باشد، کلاستر ها خلوص بیشتری دارند.

#### Mutual Information ۲.۱.۵

این معیار supervised است و میزان شباهت بین کلاستر و سمپل ها را نشان میدهد و به روش زیر محاسبه می شود:

$$MI(U, V) = \sum_{i=1}^{|U|} \sum_{j=1}^{|V|} \frac{|U_i \cap V_j|}{N} \log \frac{N|U_i \cap V_j|}{|U_i||V_j|}$$

که  $|U_i|$  و  $|V_j|$  به ترتیب نشان دهنده تعداد سمپل ها در خوشه  $i$  ام و تعداد سمپل ها در کلاس  $j$  ام هستند.

#### V\_Score ۳.۱.۵

این معیار همان mutual information می باشد که به صورت arithmetic نرمال شده است.

#### Silhouette ۴.۱.۵

این یک معیار unsupervised است و میانگین Silhouette Coefficient تمام سمپل ها را محاسبه میکند. که به صورت زیر محاسبه می شود:

$$s = \frac{b-a}{\max(a,b)}$$

که مقدار  $a$  میانگین فاصله‌ی درون کلاسی (mean intra-cluster distance) و  $b$  میانگین نزدیکترین خوش است.

### Davis Bouldin ۵.۱.۵

این یک معیار unsupervised است و برابر نسبت بین یک معیار درون کلاسی و یک معیار برون کلاسی می‌باشد. و به صورت زیر محاسبه می‌شود:

$$DB = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} R_{ij}$$

$$R_{ij} = \frac{s_i + s_j}{d_{ij}}$$

که  $s_i$  برابر است با میانگین فاصله‌ی بین هر نقطه از کلاستر  $i$  ام و مرکز آن کلاستر و  $d_{ij}$  نیز برابر است با فاصله‌ی بین مرکز کلاستر  $j$ ، هرچه این مقدار کم تر باشد، فاصله‌ی بین خوش‌ها بیشتر و فاصله‌ی درون خوش‌ای کمتر است در نتیجه کلاستر عملکرد بهتری دارد.

### Calinski-Harabasz ۶.۱.۵

این یک معیار unsupervised است و جمع نسبت پراکندگی بین کلاسترها به پراکندگی درون کلاسترها را حساب می‌کند. این معیار هرچه بیشتر باشد یعنی کلاستر عملکرد بهتر داشته است.

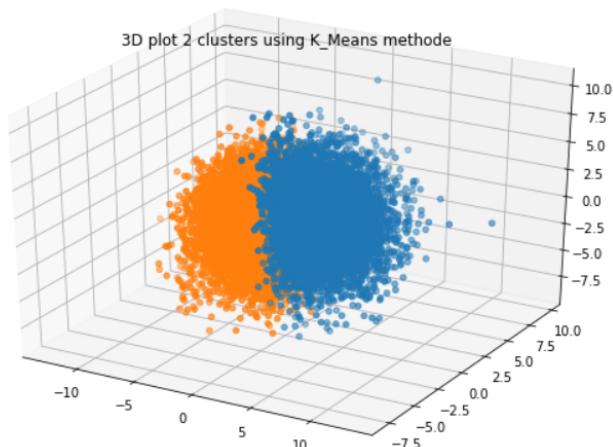
## ۲.۵ خوش بند K\_means

در این روش، ابتدا تعداد خوش‌های هدف را مشخص کرده و به تعداد آن، نقاطی تصادفی به عنوان مرکز خوش‌ها در نظر می‌گیریم. سپس در هر مرحله، فاصله‌ی نقاط تا مرکز خوش‌ها را بدست آورده و هر نقطه را به نزدیکترین مرکز خوش‌اش نسبت میدهیم. سپس مجدداً مرکز خوش‌ها را محاسبه می‌کنیم. تا زمانی که نقاط داخل خوش‌ها بین دو مرحله‌ی متوالی تغییر نکند این الگوریتم را ادامه میدهیم.

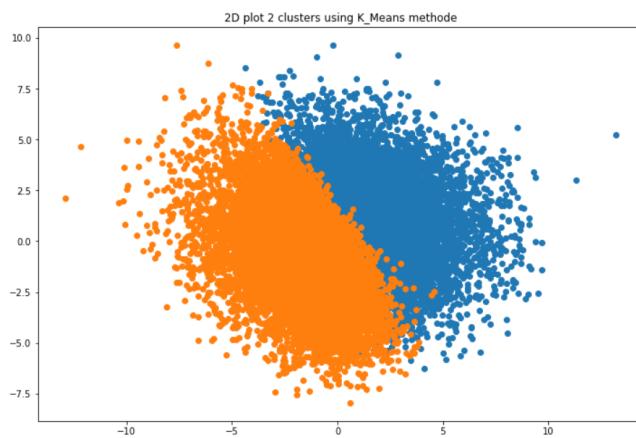
فاصله‌ی بین هر دو نقطه را، فاصله‌ی اقلیدسی تعریف می‌کنیم.

### 1.2.5 خوش‌بند K-means برای ویژگی‌های Spectrogram

```
2 Clusters K_Means Clustering Report:  
homogeneity: 0.3595045201318311  
mutual information: 0.24666282197612507  
v_measure: 0.35783030011062417  
silhouette: 0.09568201236485985  
davies_bouldin: 2.563309580358709  
calinski_harabasz: 2310.970441127032
```



شکل ۵۰: نمودار ۳ بعدی خوش‌بندی k=2 برای ویژگی‌های Spectrogram



شکل ۵۱: نمودار ۲ بعدی خوش بندی k=2 برای ویژگی های Spectogram با K\_means

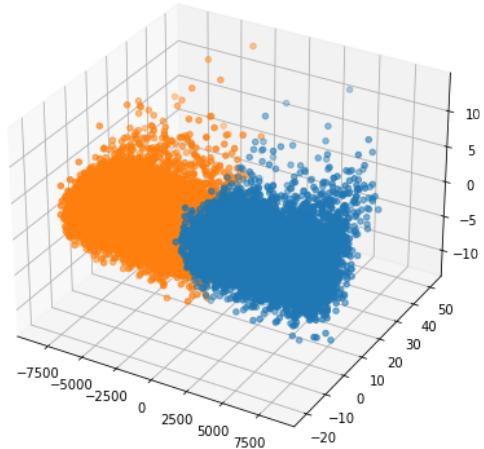
## Rythmic خوش بند K=2 برای ویژگی های ۲.۲.۵

```

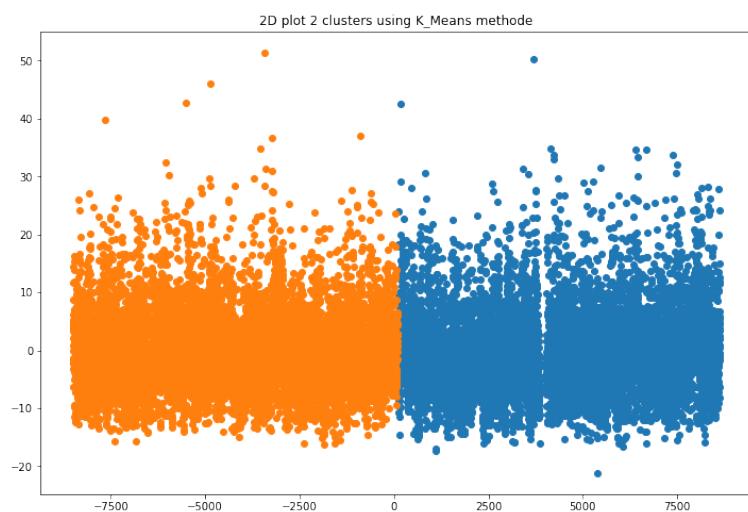
2 Clusters K_Means Clustering Report:
homogeneity: 9.347688910378505e-05
mutual information: 6.477416245126433e-05
v_measure: 9.350413278363775e-05
silhouette: 0.6225009244963325
davies bouldin: 0.5020791865815077
calinski_harabasz: 50108.47868110464

```

3D plot 2 clusters using K\_Means methode



شکل ۵۲: نمودار ۳ بعدی خوشه بندی K\_means برای ویژگی های Rythmic

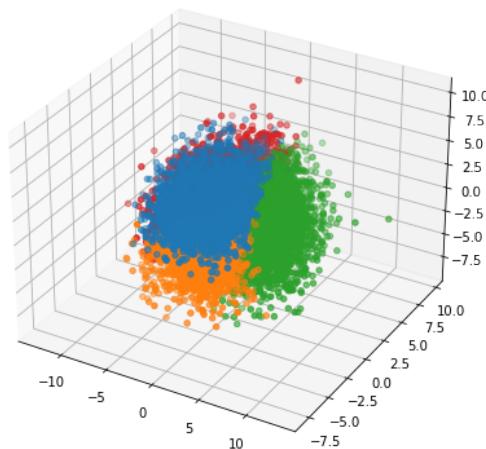


شکل ۵۳: نمودار ۲ بعدی خوشه بندی K\_means برای ویژگی های Rythmic

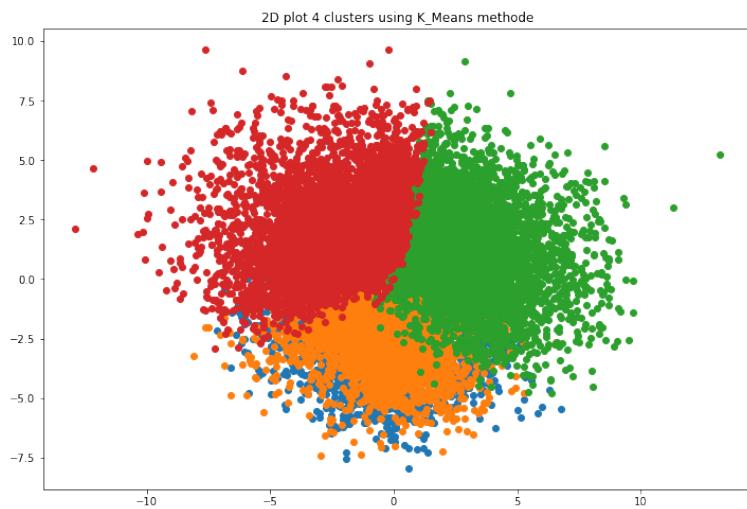
### ۳.۲.۵ خوش بند Spectrogram برای ویژگی های K=4 با K\_means

```
4 Clusters K_Means Clustering Report:  
homogeneity: 0.007448215921773789  
mutual information: 0.010252217695813373  
v_measure: 0.007421733864669898  
silhouette: 0.09737394201334638  
davies bouldin: 1.9706100354598362  
calinski_harabasz: 2069.435439275602
```

3D plot 4 clusters using K\_Means methode



شکل ۵۴: نمودار ۳ بعدی خوش بندی Spectrogram برای ویژگی های K=4 با K\_means

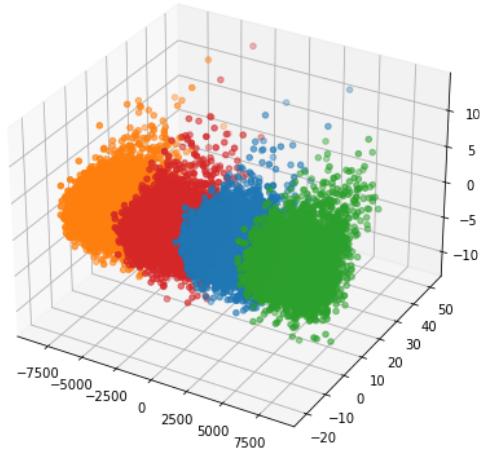


شکل ۵۵: نمودار ۲ بعدی خوشه بندی K\_means با k=4 برای ویژگی های Spectrogram

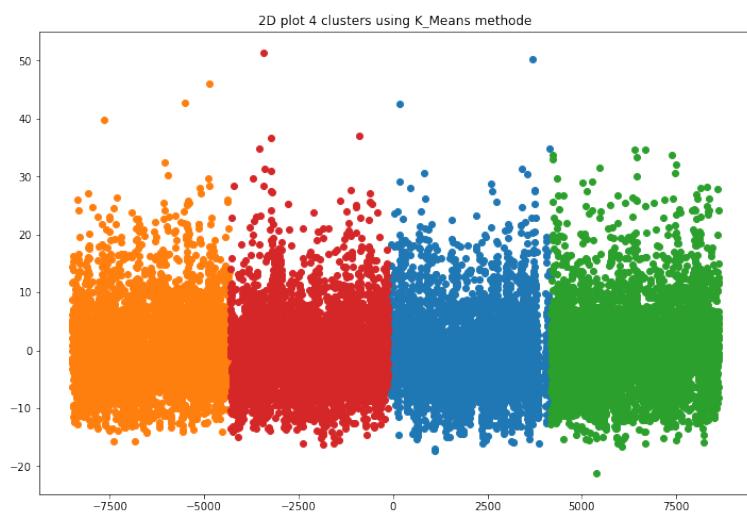
#### ۴.۲.۵ خوشه بند K=4 برای ویژگی های Rythmic

```
4 Clusters K_Means Clustering Report:
homogeneity:  1.9635344204990015e-05
mutual information:  2.7211116838066207e-05
v_measure:  1.963204405815643e-05
silhouette:  0.56988162354792
davies bouldin:  0.5002577206117109
calinski_harabasz:  85015.70023450584
```

3D plot 4 clusters using K\_Means methode



شکل ۵۶: نمودار ۳ بعدی خوشه بندی k=4 برای ویژگی های Rythmic

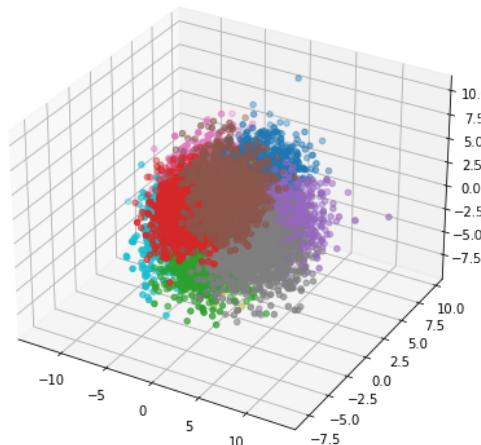


شکل ۵۷: نمودار ۲ بعدی خوشه بندی k=4 برای ویژگی های Rythmic

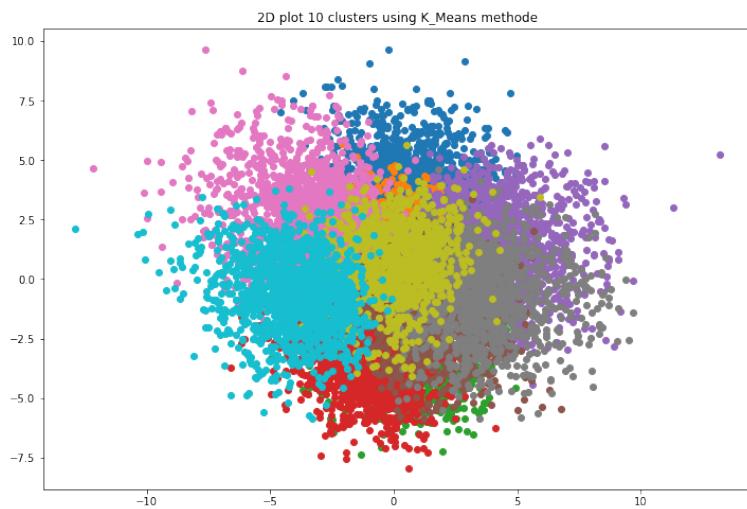
## 5.2.5 خوش بندی سپکتограм برای $K=10$ که با $K\_means$ بدستور Spectrogram ایجاد شده است

```
10 Clusters K_Means Clustering Report:  
homogeneity: 0.009585722481751156  
mutual information: 0.02174621614616053  
v_measure: 0.009514515295645534  
silhouette: 0.08611754116724342  
davies bouldin: 2.026602100694459  
calinski_harabasz: 1425.6071426553622
```

3D plot 10 clusters using K\_Means methode



شکل ۵۸: نمودار ۳ بعدی خوش بندی سپکتogram برای  $K=10$  که با  $K\_means$  بدستور Spectrogram ایجاد شده است



شکل ۵۹: نمودار ۲ بعدی خوشه بندی K\_means با k=10 برای ویژگی های Spectrogram

#### ۶.۲.۵ خوشه بند K\_means با K=10 برای ویژگی های Rythmic

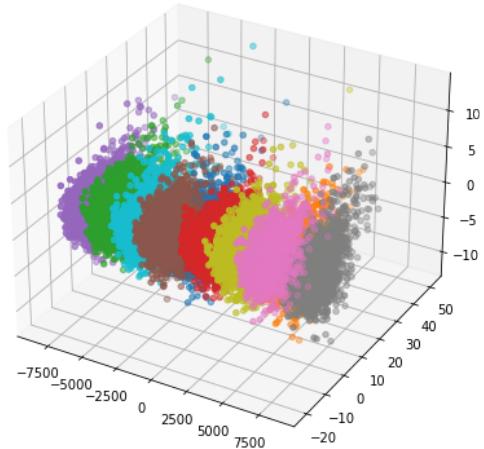
---

```

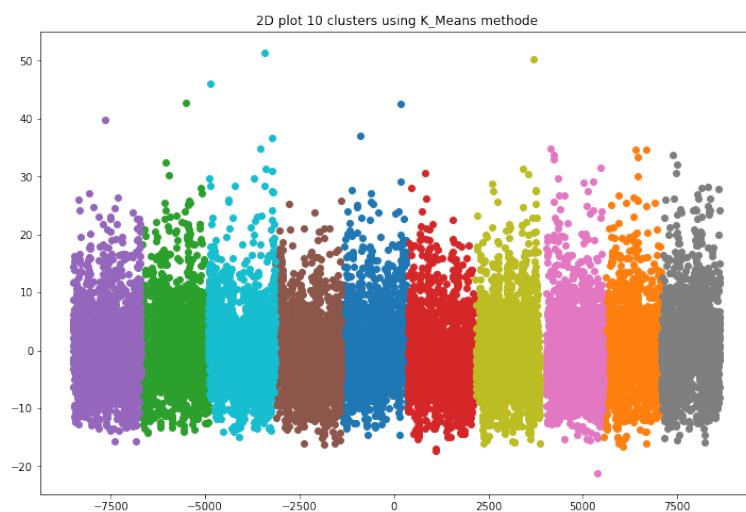
10 Clusters K_Means Clustering Report:
homogeneity:  0.00010731530567649716
mutual information:  0.0002468621944140617
v_measure:  0.00010726362040916856
silhouette:  0.5331613085489404
davies bouldin:  0.4982987787353855
calinski_harabasz:  188756.64844631977

```

3D plot 10 clusters using K\_Means methode



شکل ۶۰: نمودار ۳ بعدی خوشه بندی K\_means با k=10 برای ویژگی های Rythmic



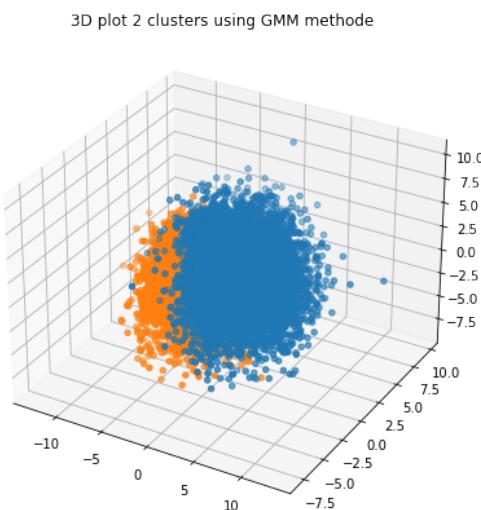
شکل ۶۱: نمودار ۲ بعدی خوشه بندی K\_means با k=10 برای ویژگی های Rythmic

### ۳.۵ خوشه بند GMM

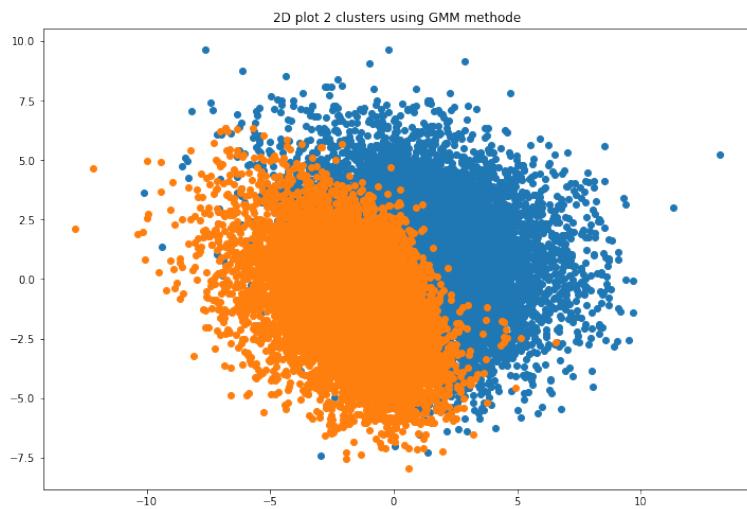
در این روش توزیع داده را جمع وزن دار تعدادی توزیع گوسی در نظر میگیرم و به کمک الگوریتم بهینه سازی EM، تلاش میکنیم تا پارامتر های میانگین و واریانس توزیع های گوسی و همچنین وزن هر توزیع را برای یک نقطه بدست اوریم. هرچه وزن یک توزیع برای یک نقطه بیشتر باشد، احتمال تعلق نقطه به آن توزیع بیشتر است. بر همین اساس میتوانیم خوشه بندی را انجام دهیم.

#### ۱.۳.۵ خوشه بند GMM با $K=2$ برای ویژگی های Spectrogram

```
2 Clusters GMM Clustering Report:  
homogeneity: 0.448571063299629  
mutual information: 0.3104220747356492  
v_measure: 0.4484042213461966  
silhouette: 0.08044868934030508  
davies bouldin: 2.7758116701924007  
calinski_harabasz: 1967.6210870161979
```



شکل ۶۲: نمودار ۳ بعدی خوشه بندی GMM با  $k=2$  برای ویژگی های Spectrogram

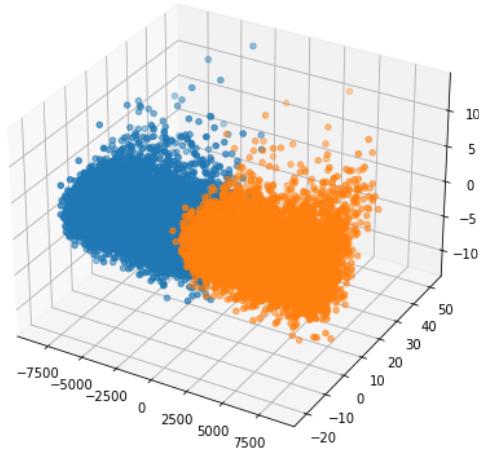


شکل ۶۳: نمودار ۲ بعدی خوشبندی GMM با  $k=2$  برای ویژگی های Spectrogram

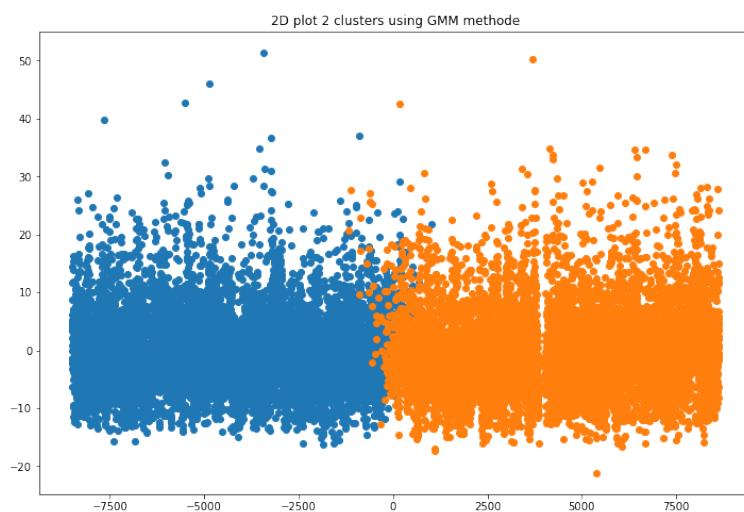
### ۲.۳.۵ خوشبند GMM با $K=2$ برای ویژگی های Rythmic

```
2 Clusters GMM Clustering Report:
homogeneity: 1.0332491146322358e-05
mutual information: 7.1578657250825906e-06
v_measure: 1.0334086346891977e-05
silhouette: 0.6183566622441287
davies bouldin: 0.504419068697261
calinski_harabasz: 49427.19856319352
```

3D plot 2 clusters using GMM methode



شکل ۶۴: نمودار ۳ بعدی خوشه بندی GMM با  $k=2$  برای ویژگی های Rythmic

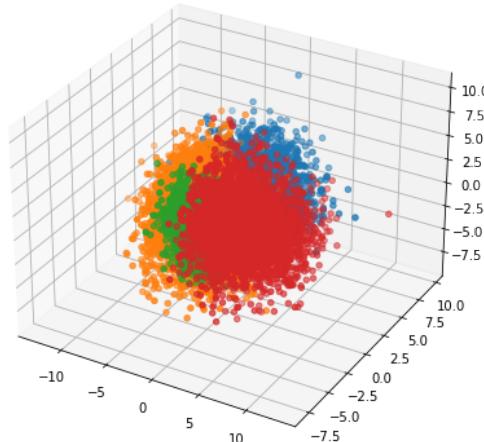


شکل ۶۵: نمودار ۲ بعدی خوشه بندی GMM با  $k=2$  برای ویژگی های Rythmic

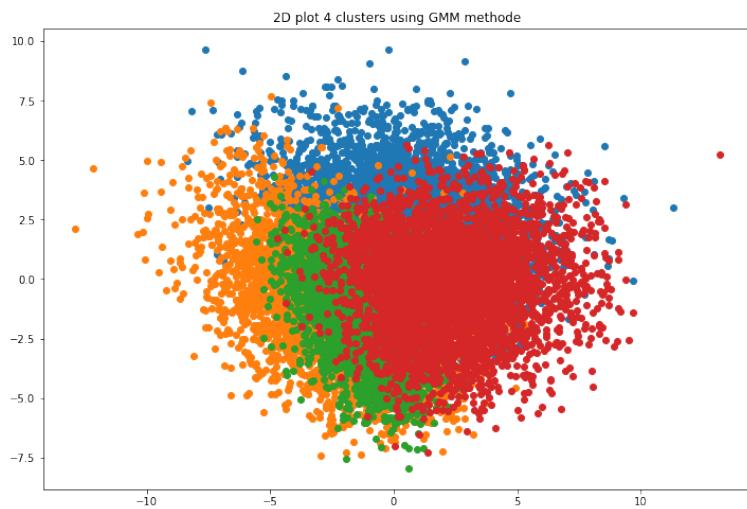
### ۳.۳.۵ خوش بند GMM با $K=4$ برای ویژگی های Spectrogram

```
4 Clusters GMM Clustering Report:  
homogeneity: 0.011426451689904566  
mutual information: 0.015746975689321943  
v_measure: 0.011392668417095905  
silhouette: 0.040576598985306085  
davies bouldin: 3.333660568213911  
calinski_harabasz: 1331.7280988088512
```

3D plot 4 clusters using GMM methode



شکل ۶۶: نمودار ۳ بعدی خوش بندی k=4 GMM برای ویژگی های Spectrogram



شکل ۶۷: نمودار ۲ بعدی خوشبندی GMM با  $k=4$  برای ویژگی های Spectrogram

#### ۴.۳.۵ خوشبند GMM با $K=4$ برای ویژگی های Rythmic

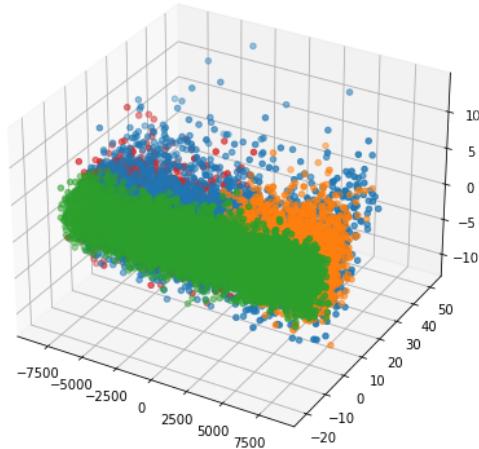
---

```

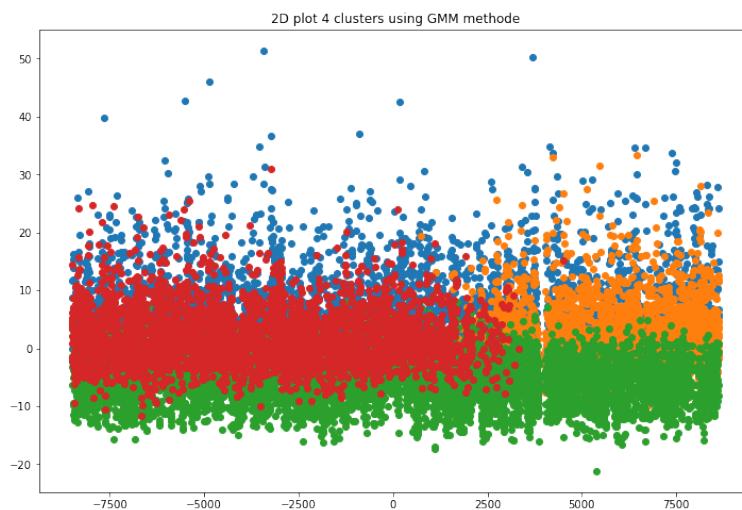
4 Clusters GMM Clustering Report:
homogeneity:  0.0037227444006303844
mutual information:  0.004884498117875974
v_measure:  0.003620348006346737
silhouette:  -0.10420579305183235
davies bouldin:  9.545581525023596
calinski_harabasz:  2601.2576443846738

```

3D plot 4 clusters using GMM methode



شکل ۶۸: نمودار ۳ بعدی خوشه بندی GMM با  $k=4$  برای ویژگی های Rythmic

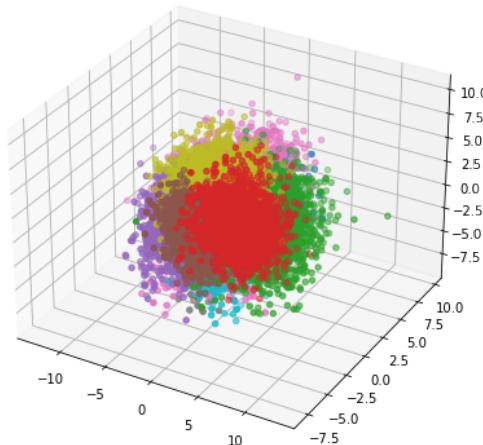


شکل ۶۹: نمودار ۲ بعدی خوشه بندی GMM با  $k=4$  برای ویژگی های Rythmic

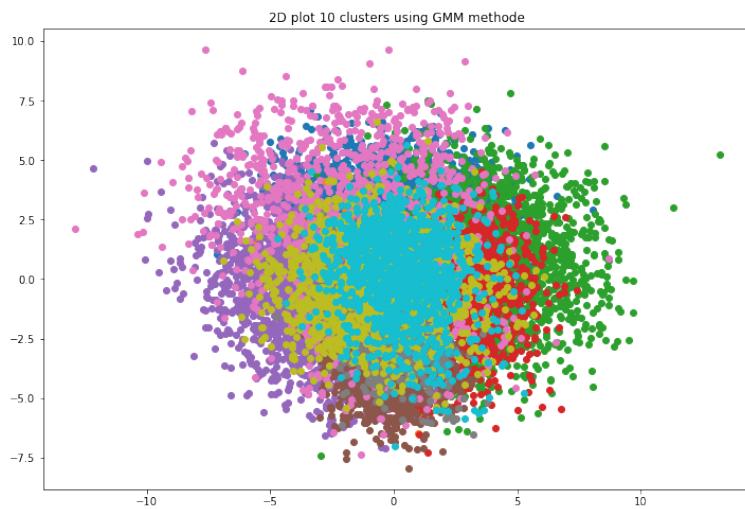
### ۵.۳.۵ خوش بند GMM با $K=10$ برای ویژگی های Spectrogram

```
10 Clusters GMM Clustering Report:  
homogeneity: 0.006042848018637367  
mutual information: 0.01373224660178999  
v_measure: 0.006003114402541006  
silhouette: 0.009368863140224124  
davies bouldin: 2.849722302732162  
calinski_harabasz: 857.4311814661856
```

3D plot 10 clusters using GMM methode



شکل ۷۰: نمودار ۳ بعدی خوش بندی GMM با  $k=10$  برای ویژگی های Spectrogram

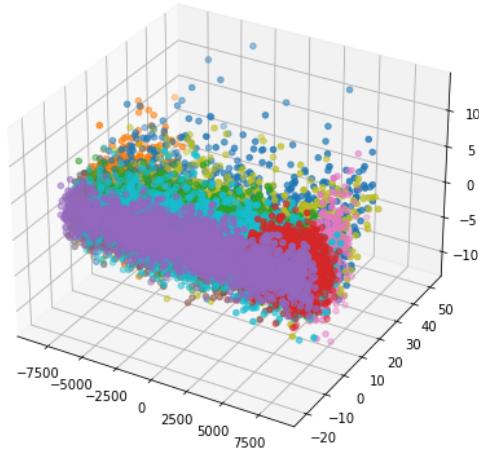


شکل ۷۱: نمودار ۲ بعدی خوش بندی GMM با  $k=10$  برای ویژگی های Spectrogram

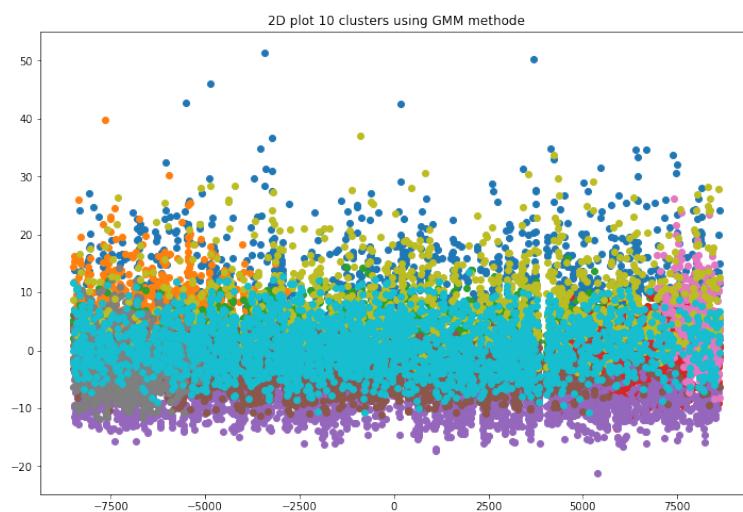
### ۶.۳.۵ خوش بند GMM با $K=10$ برای ویژگی های Rythmic

```
10 Clusters GMM Clustering Report:
homogeneity:  0.078502823267307
mutual information:  0.1711081910047719
v_measure:  0.07635000096483446
silhouette:  -0.26831787780046557
davies bouldin:  19.674136053509418
calinski_harabasz:  1572.8504632108438
```

3D plot 10 clusters using GMM methode



شکل ۷۲: نمودار ۳ بعدی خوشه بندی GMM با  $k=10$  برای ویژگی های Rythmic



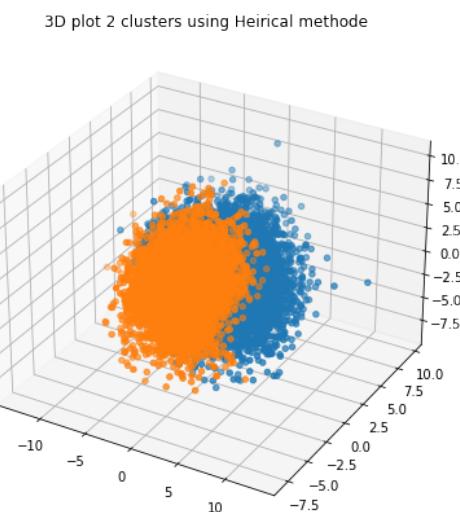
شکل ۷۳: نمودار ۲ بعدی خوشه بندی GMM با  $k=10$  برای ویژگی های Rythmic

## ۴.۵ خوشه بند Agglomerative Hierarchical

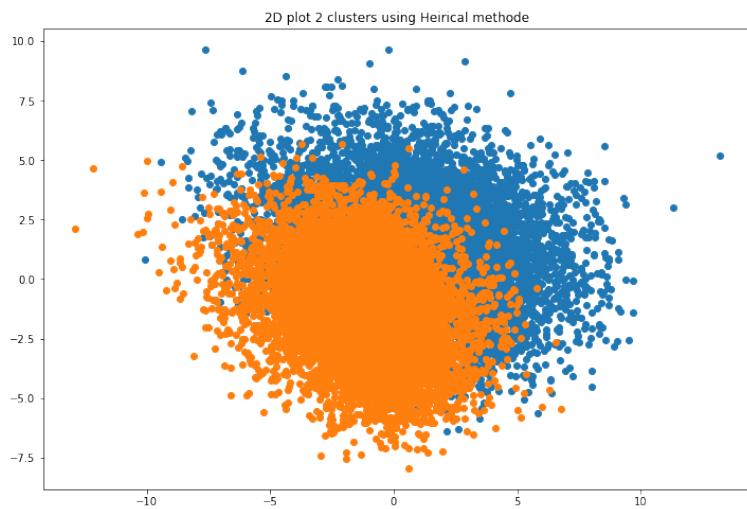
در این روش که یک روش hierarchical می باشد، ابتدا هر سمپل را یک خوشه در نظر می گیریم. سپس در هر مرحله دو خوشه که کمترین شباهت را دارند با یکدیگر ترکیب می کنیم تا به تعداد خوشه های مورد نظر برسیم. اگر تعداد خوشه ها مشخص نباشد، این الگوریتم تازمانی که تمام داده ها در یک خوشه قرار بگیرند میتواند ادامه پیدا کند.

### ۱.۴.۵ خوشه بند Spectrogram برای ویژگی های K=2 Hierarchical

```
2 Clusters Heirical Clustering Report:  
homogeneity: 0.26590389599226716  
mutual information: 0.18400134399048537  
v_measure: 0.2657973273516683  
silhouette: 0.07369087487201653  
davies bouldin: 3.144820496407376  
calinski_harabasz: 1550.581999661632
```



شکل ۷۴: نمودار ۳ بعدی خوشه بندی Hierarchical برای ویژگی های k=2



شکل ۷۵: نمودار ۲ بعدی خوشه بندی Hierarchical با  $k=2$  برای ویژگی های Spectrogram

#### ۲.۴.۵ خوشه بند Hierarchical با $K=2$ برای ویژگی های Rythmic

```
2 Clusters Heirical Clustering Report:  

homogeneity: 8.473796991863614e-05  

mutual information: 5.8086257304301725e-05  

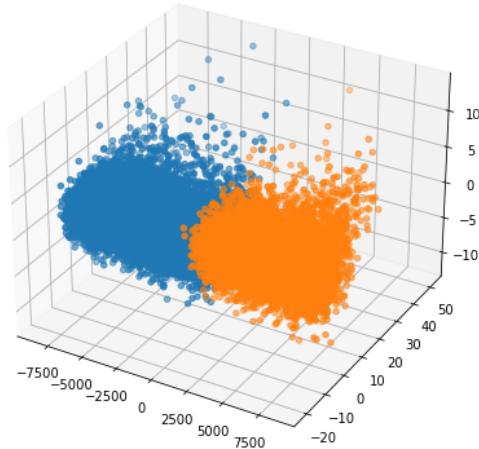
v_measure: 8.430393071075607e-05  

silhouette: 0.6169157007725594  

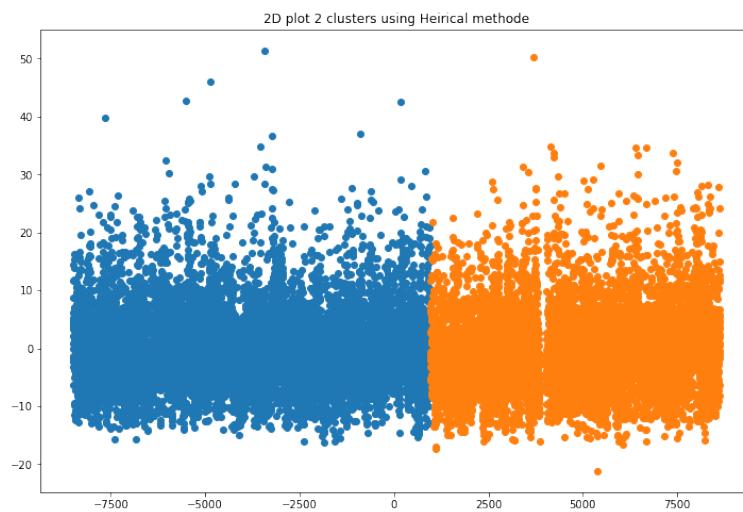
davies bouldin: 0.4988927517610284  

calinski_harabasz: 48077.62765708181
```

3D plot 2 clusters using Hierical methode



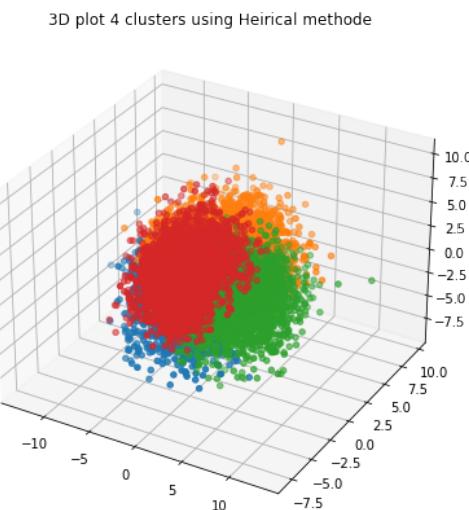
شکل ۷۶: نمودار ۳ بعدی خوشه بندی Hierarchical با  $k=2$  برای ویژگی های Rythmic



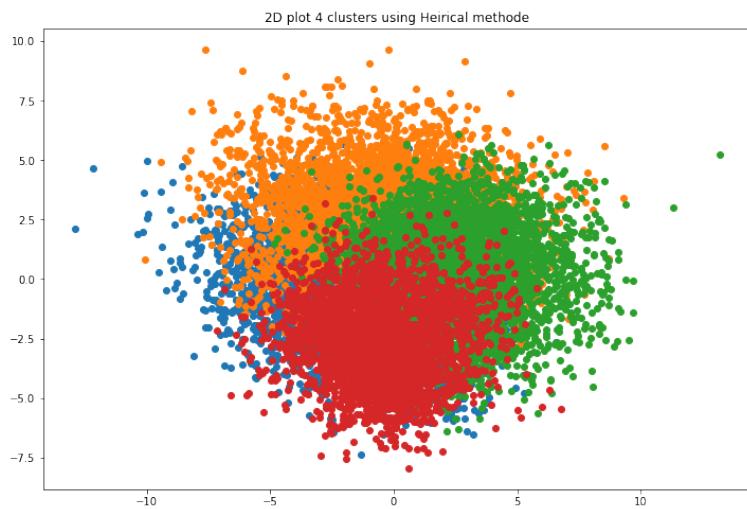
شکل ۷۷: نمودار ۲ بعدی خوشه بندی Hierarchical با  $k=2$  برای ویژگی های Rythmic

### ۳.۴.۵ خوش بند Spectrogram برای ویژگی های K=4

```
4 Clusters Heirical Clustering Report:  
homogeneity: 0.008770061304515013  
mutual information: 0.01199050511943782  
v_measure: 0.00870928916124969  
silhouette: 0.07047417865643608  
davies bouldin: 2.2713952092608336  
calinski_harabasz: 1601.0185486627877
```



شکل ۷۸: نمودار ۳ بعدی خوش بندی Spectrogram برای ویژگی های Hierarchical با k=4

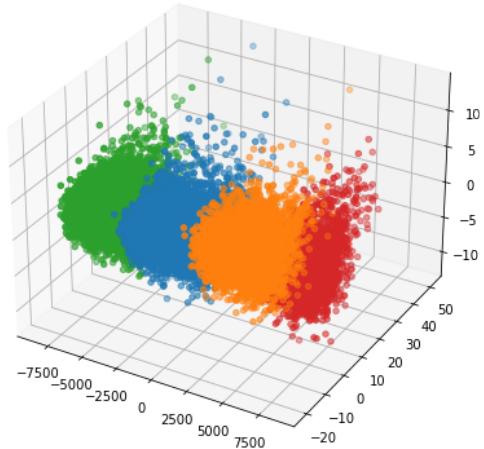


شکل ۷۹: نمودار ۲ بعدی خوشبندی های k=4 برای ویژگی های Spectrogram با Hierarchical

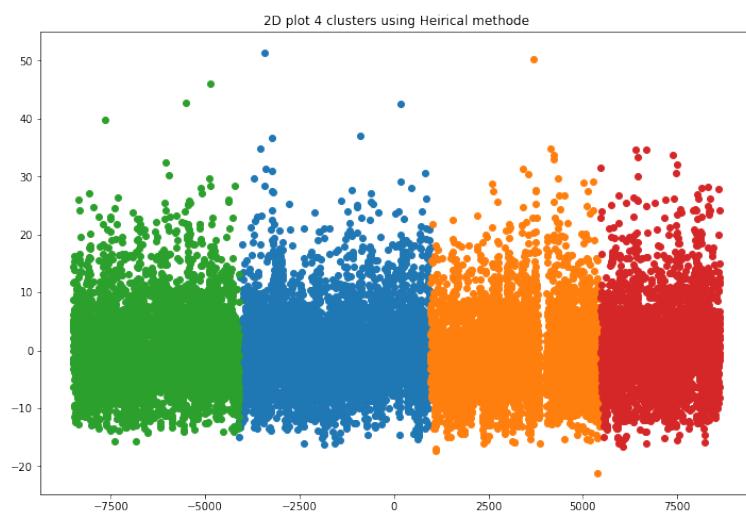
#### ۴.۴.۵ خوشبندی های K=4 برای ویژگی های Rythmic

```
4 Clusters Heirical Clustering Report:  
homogeneity: 2.1069337668601936e-05  
mutual information: 2.8908244996156407e-05  
v_measure: 2.0960594866415532e-05  
silhouette: 0.5531192105728246  
davies bouldin: 0.5029180476259563  
calinski_harabasz: 77321.42964466655
```

3D plot 4 clusters using Hierical methode



شکل ۸۰: نمودار ۳ بعدی خوشه بندی Hierarchical با  $k=4$  برای ویژگی های Rythmic

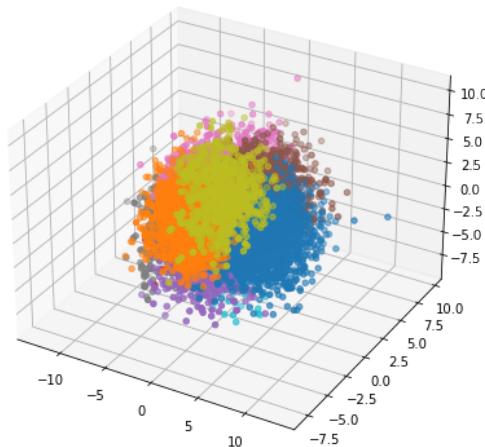


شکل ۸۱: نمودار ۲ بعدی خوشه بندی Hierarchical با  $k=4$  برای ویژگی های Rythmic

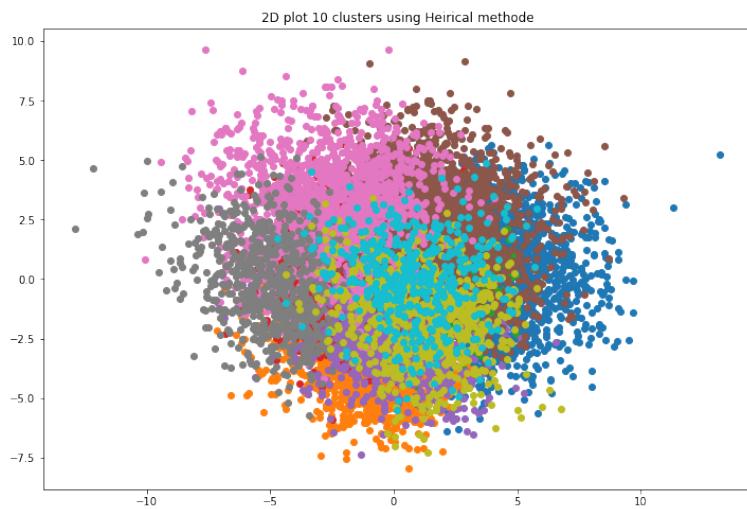
## ۵.۴.۵ خوش بند Spectrogram برای ویژگی های $K=10$ با Hierarchical

```
10 Clusters Heirical Clustering Report:  
homogeneity: 0.012340504951492526  
mutual information: 0.0268790828521112  
v_measure: 0.015081934967497478  
silhouette: 0.04684539752204872  
davies bouldin: 2.274038999272027  
calinski_harabasz: 1039.6122055080625
```

3D plot 10 clusters using Heirical methode



شکل ۸۲: نمودار ۳ بعدی خوش بندی Spectrogram برای ویژگی های  $k=10$  با Hierarchical

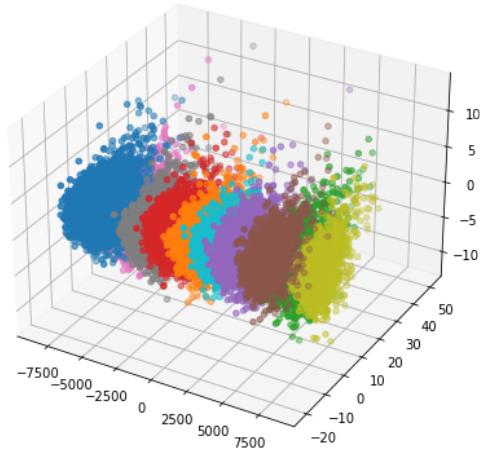


شکل ۸۳: نمودار ۲ بعدی خوشبندی Hierarchical با  $k=10$  برای ویژگی های Spectrogram

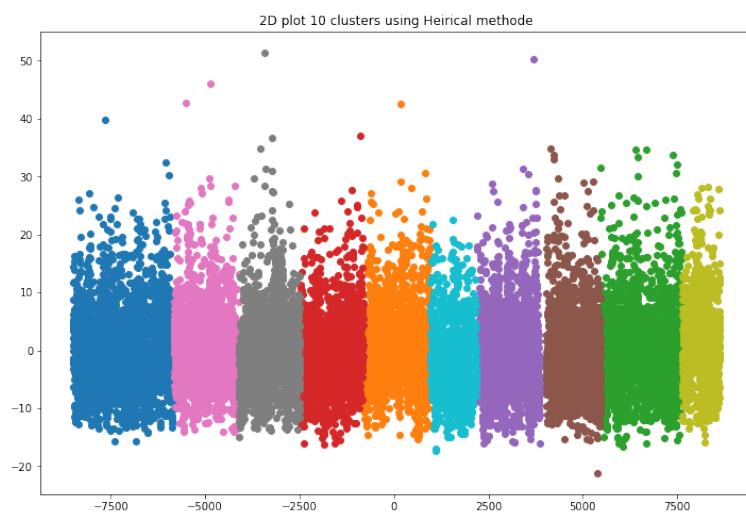
#### ۶.۴.۵ خوشبند **Hierarchical** با $K=10$ برای ویژگی های Rythmic

```
10 Clusters Heirical Clustering Report:
homogeneity: 3.4459583706814454e-05
mutual information: 7.816786496609579e-05
v_measure: 4.27768377629473e-05
silhouette: 0.5065855786630756
davies bouldin: 0.49809261757562134
calinski_harabasz: 156118.23882583197
```

3D plot 10 clusters using Heirical methode



شکل ۸۴: نمودار ۳ بعدی خوشه بندی Hierarchical با  $k=10$  برای ویژگی های Rythmic

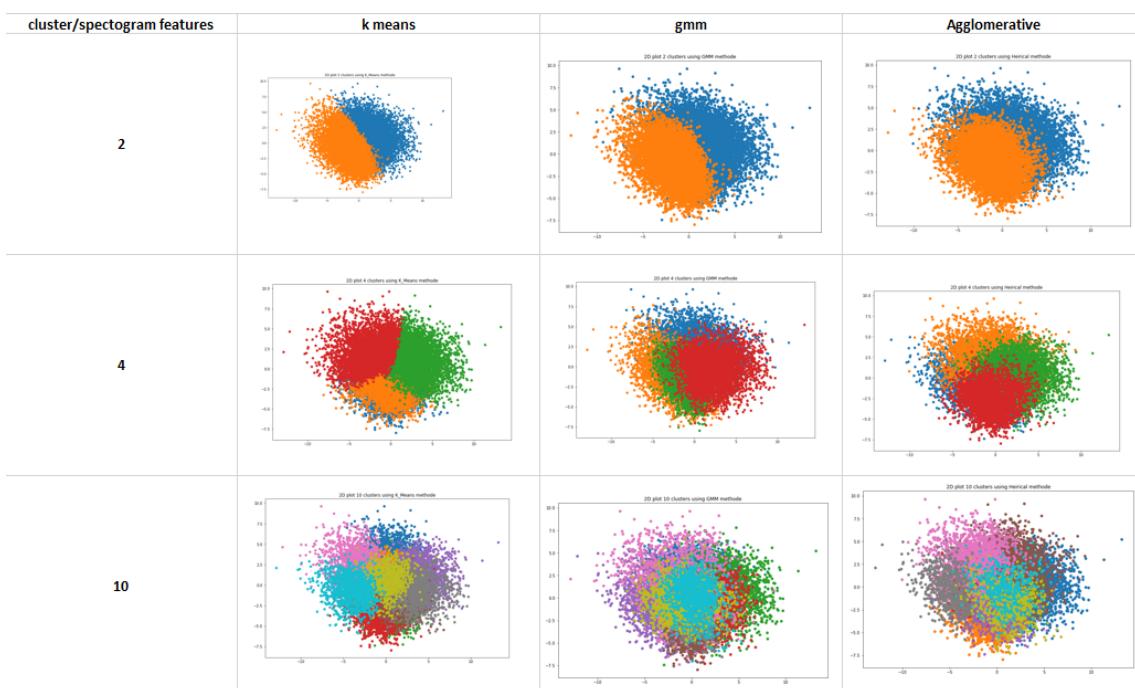


شکل ۸۵: نمودار ۲ بعدی خوشه بندی Hierarchical با  $k=10$  برای ویژگی های Rythmic

## ۵.۵ نتیجه گیری خوش بندی

cluster/spectrogram features	k means	gmm	Agglomerative
silhouette	0.095	0.08	0.0736
calinski_harabasz	2310.637	1967.621	1550.581
davies bouldin:	2.563	2.775	3.144
4cluster/spectrogram features	k means	gmm	Agglomerative
silhouette	0.097	0.04	0.07
calinski_harabasz	2069.435	1331.728	1601.018
davies bouldin:	1.9706	3.333	2.271
10cluster/spectrogram features	k means	gmm	Agglomerative
silhouette	0.086	0.009	0.046
calinski_harabasz	1425.607	857.431	1039.612
davies bouldin:	2.026	2.849	2.274

شکل ۸۶: جمع بندی خوش بندی Spectrogram features



شکل ۸۷: نمودار همه حالت های خوش بندی برای Spectrogram features

:K=2

همان طور که میبینیم مقدار calinski\_harabasz و silhouette در kmeans >gmm > ag- این ترتیب بر عکس davies bouldin باشد و برای معیار glomerative

ها نیز دو خوش دو در kmeans و gmm بهتر جدا شده اند. بنابر این می توان نتیجه گرفت عملکرد kmeans > gmm > agglomerative بوده است.

:K=4

همانطور که میبینیم معیار calinski\_harabasz و silhouette در kmeans بیشترین مقدار را داشته و davies bouldin نیز در kmeans کمترین مقدار را دارد. همچنین در شکل خوش ها نیز میبینیم که در kmeans خوش ها بهتر تقسیم شده اند. البته در هر سه روش، خوش ها به خوبی تقسیم نشده اند و خوش های قرمز به خوبی مشخص شده است.

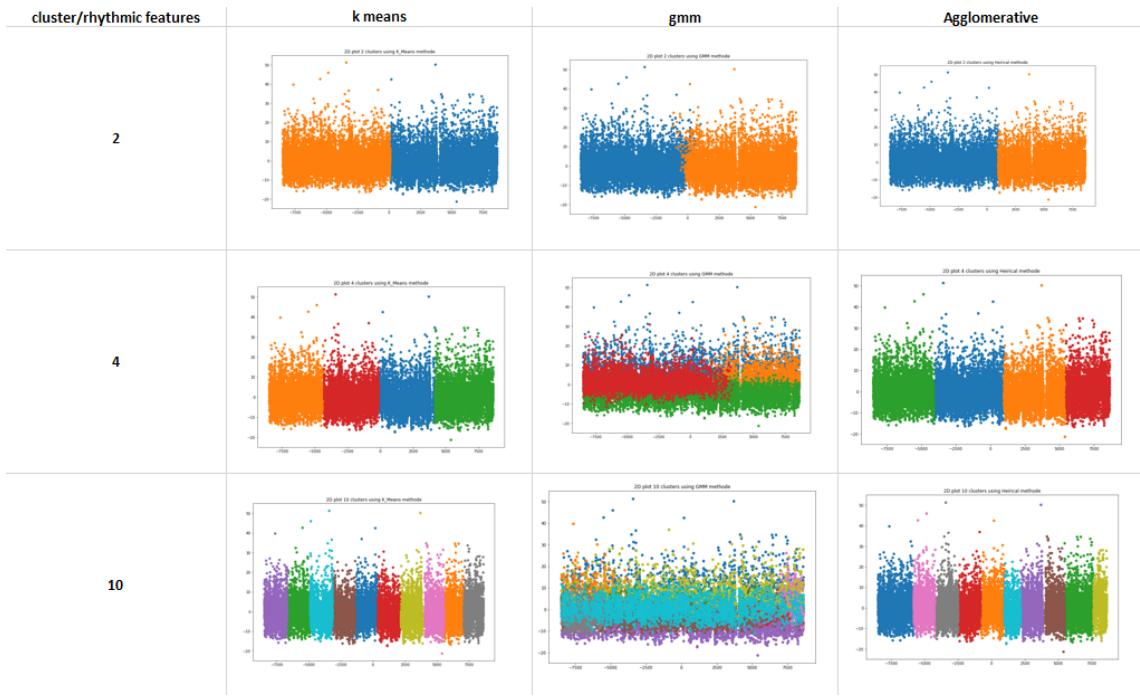
:K=10

همان طور که میبینیم مقدار calinski\_harabasz و silhouette در kmeans > agglomerative > gmm می باشد و برای معیار davies bouldin این ترتیب برعکس است. همچنین در شکل ها نیز ترتیب مشخص بودن مرز خوش ها به همین صورت می باشد. بنابر این می توان نتیجه گرفت عملکرد kmeans > gmm > agglomerative است. البته به طور کلی در هر سه روش مرز بین خوش ها به خوبی مشخص نیست.

به طور کلی میبینیم که در هر روش با افزایش تعداد خوش ها، عملکرد خوش بند کاهش پیدا کرده و خوش ها به خوبی قابل جداسازی نیستند.

	k means	gmm	Agglomerative
2cluster/rhythmic features silhouette	0.622	0.618	0.616
calinski_harabasz	50188.478	49427.198	0.498
davies bouldin:	0.50207918	0.504	48077.627
4cluster/rhythmic features silhouette	k means	gmm	Agglomerative
calinski_harabasz	0.569	-0.104	0.553
davies bouldin:	85815.7	2601.257	77321.429
10cluster/rhythmic features silhouette	0.5	9.545	0.502
calinski_harabasz	k means	gmm	Agglomerative
davies bouldin:	0.533	-0.268	0.506
188756.648	188756.648	1572.85	156118.238
0.498	0.498	19.674	0.498

شکل ۸۸: جمع بندی خوش بندی Rhythmic features



شکل ۸۹: نمودار همه حالت های خوش بندی برای Rhythmic features

:K=2

همان طور که میبینیم مقدار calinski\_harabasz و silhouette در kmeans > gmm > agglomerative می باشد و برای معیار davies bouldin این ترتیب برعکس است. در شکل خوش ها نیز دو خوش دو در kmeans و gmm بهتر جدا شده اند. بنابر این می توان نتیجه گرفت عملکرد kmeans > gmm > agglomerative بوده است. البته در کل هر سه خوش بند به خوبی عملکرده اند و داده ها را داسته بندی کرده اند.

:K=4

همان طور که میبینیم مقدار calinski\_harabasz و silhouette در kmeans> agglomer- ative >gmm می باشد و برای معیار davies bouldin این ترتیب برعکس است. همچنین در شکل ها نیز دو خوش بند kmeans و agglomerative به خوبی داده ها را خوش بندی کرده اند ولی در gmm خوشی آبی، به خوبی تشکیل نشده است. بنابر این می توان نتیجه گرفت عملکرد kmeans > gmm > agglomerative بسیار پایینتر از دو خوش

بند دیگر است.

:K=10

همان طور که میبینیم مقدار silhouette در calinski\_harabasz و agglomer-  
ative>gmm می باشد و برای معیار bouldin davies این ترتیب برعکس است. همچنین در  
شکل ها نیز دو خوشه بند kmeans و agglomerative به خوبی داده ها را خوشه بندی کرده  
اند ولی در gmm خوشه ای آبی ، به خوبی تشکیل نشده است. بنابر این می توان نتیجه گرفت  
عملکرد kmeans>gmm بسیار پایینتر از دو خوشه  
بند دیگر است.

در این فیچر ها ، هر دو خوشه بند k=2,4,10 و agglomerative kmeans به خوبی عمل  
میکنند و با افزایش k ، عملکرد آنها کمی کاهش می یابد. خوشه بند gmm نیز در k=2 به خوبی  
عمل میکند و به ازای k=10 ، k=4 برخی از خوشه ها را به خوبی جدا نمیکند. با مقایسه ای-  
Rhythmic features و mic features Spectrogram features نیز میبینیم که در خوشه بندی  
داده ها بهتر جدا شده اند و با افزایش k ، مرز بین دسته خیلی بهم نریخته است. در بین سه خوشه  
بند نیز به طور کلی kmeans عملکرد را داشته است. و پس از آن agglomerative به  
طور کلی بهتر از gmm عمل کرده است.