

آرمان سمیعی

۹۵۳۱۰۳۹

دکتر راستی

گزارش پروژه

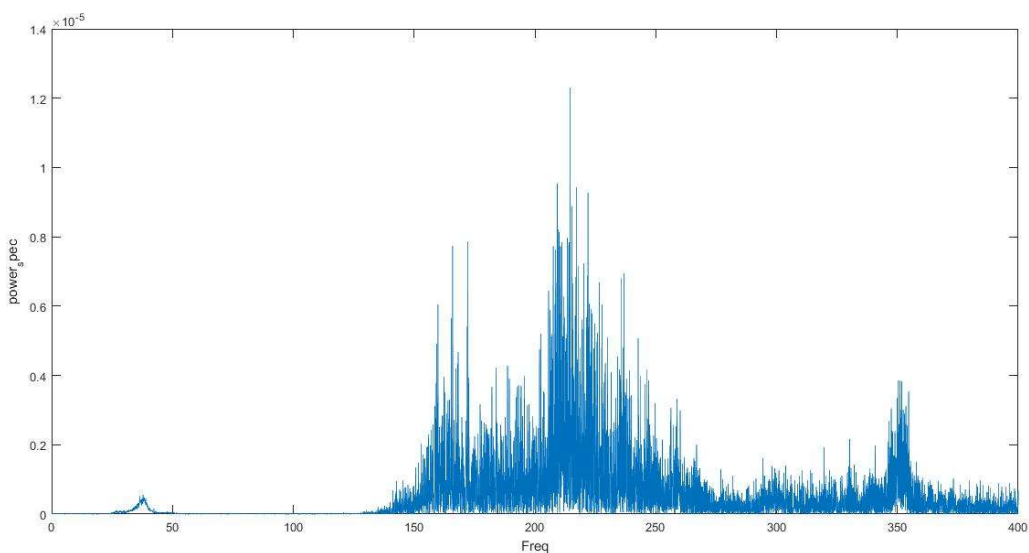
ابتدا صدا را خوانده و ویژگی هایش را در fs و data ذخیره میکنیم. Fs (۴۴۱۰۰ هرتز) فرکانس نمونه برداری است و data نمونه ها هستند. ابتدا میانگین را از داده ها کم میکنیم تا مقدار دامنه را کم کنیم چون شیفست در راستای محور y تاثیری در مولفه های فرکانسی ندارد.

سپس به دلیل مختلط بودن و یکی نبودن مقیاس ها در حوزه زمان و فرکانس وجود مولفه های فرکانسی در نیمه چپ محور فرکانسی باید یک نرمال سازی روی تبدیل فوریه انجام دهیم. بنابراین قدرمطلق تبدیل فوریه را میگیریم تا اندازه مولفه ها به دست بیاید و برای یکی کردن مقیاس در حوزه فرکانس این مقدار را بر تعداد نمونه ها تقسیم میکنیم. اگر این مقدار را به توان دو برسانیم طیف توانی سیگنال در حوزه فرکانس به دست می آید. اما چون نیمه سمت چپ را حذف کردیم [باید در یک دو هم ضرب کنیم اما چون برای هدف ما فرقی ندارد این کار را نکردیم.

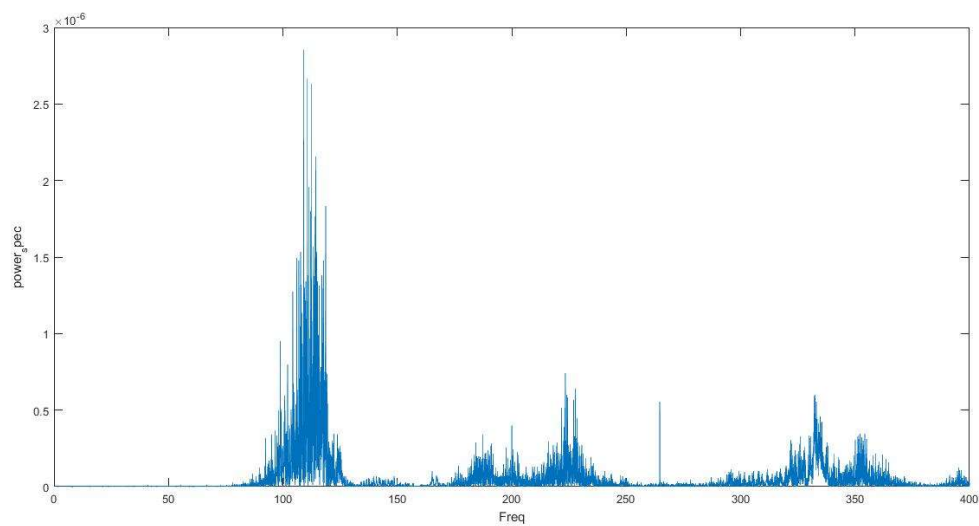
سپس محور فرکانسی را به صورت قدم های fs/N از ۰ تا fs تعریف میکنیم .

به این ترتیب نمودار ها به دست می آیند. فایل های [plot voices](#) و [plot fft](#) برای این کار هستند.

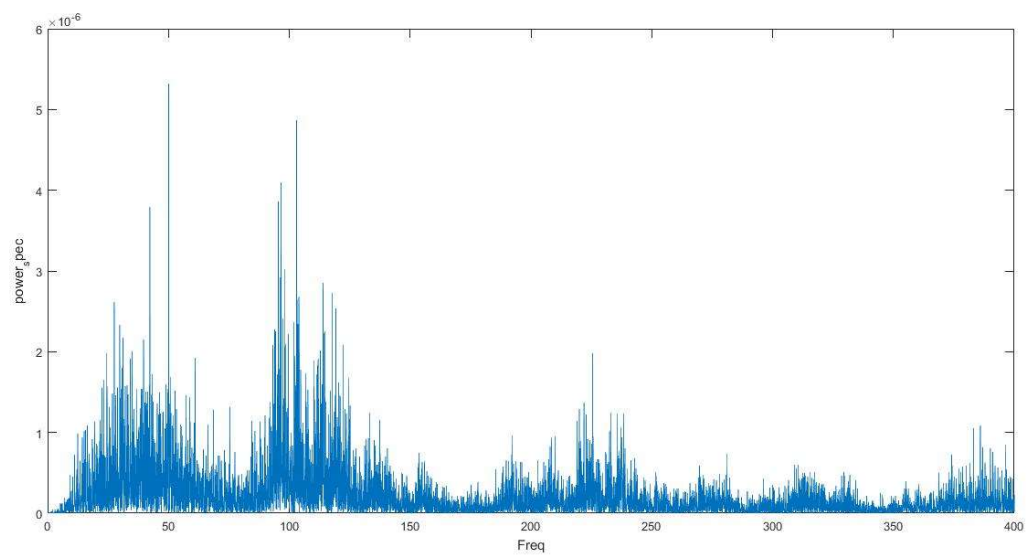
v0.mp3 – ۱



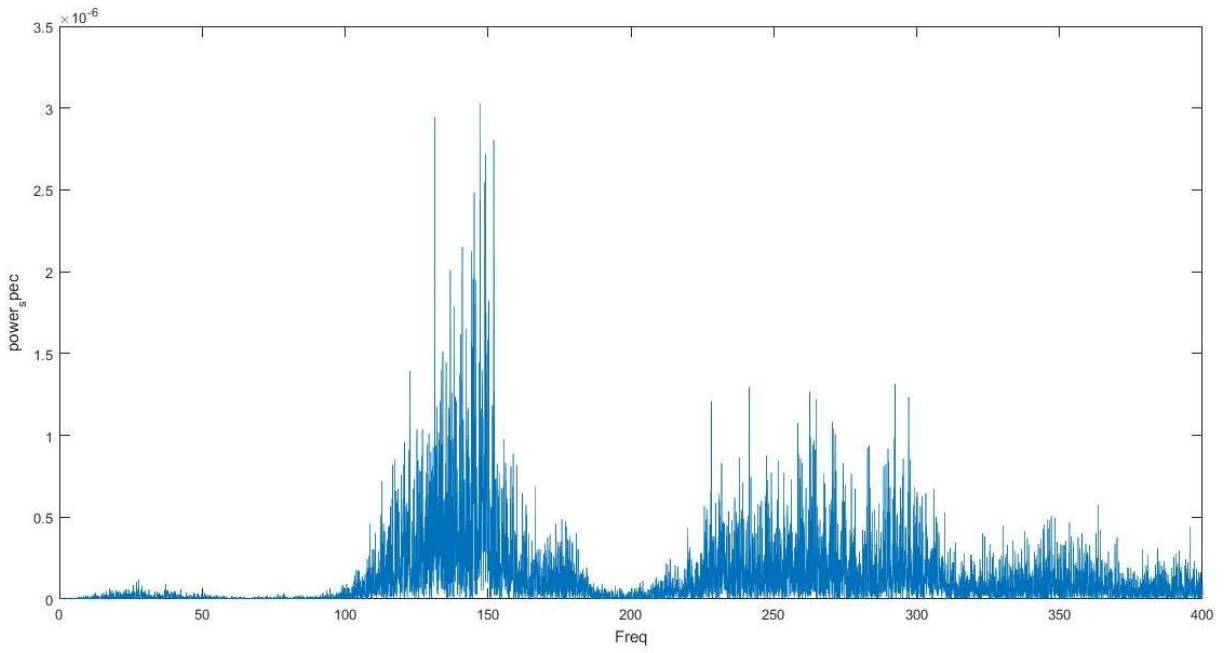
v1.mp3 – Γ



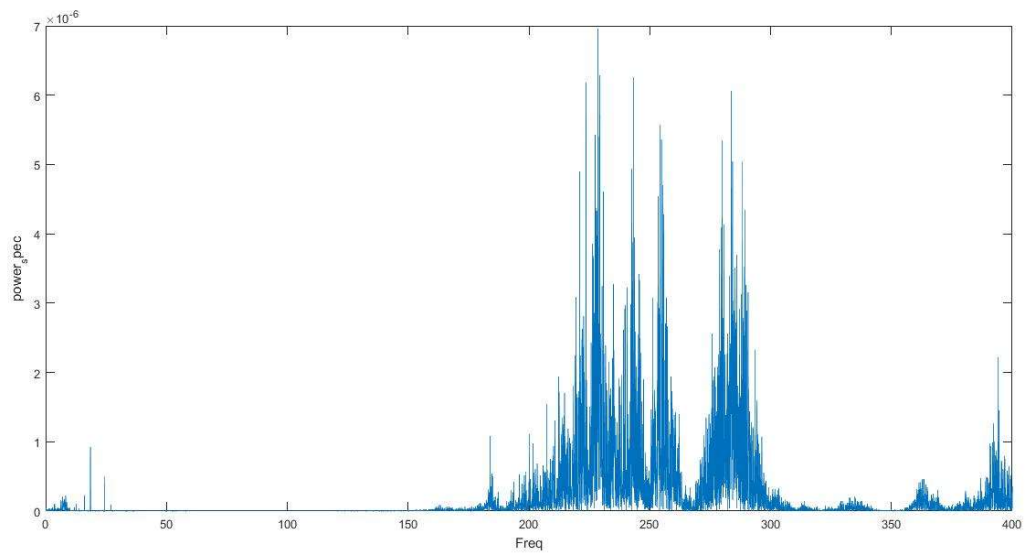
v2.mp3 – Ψ



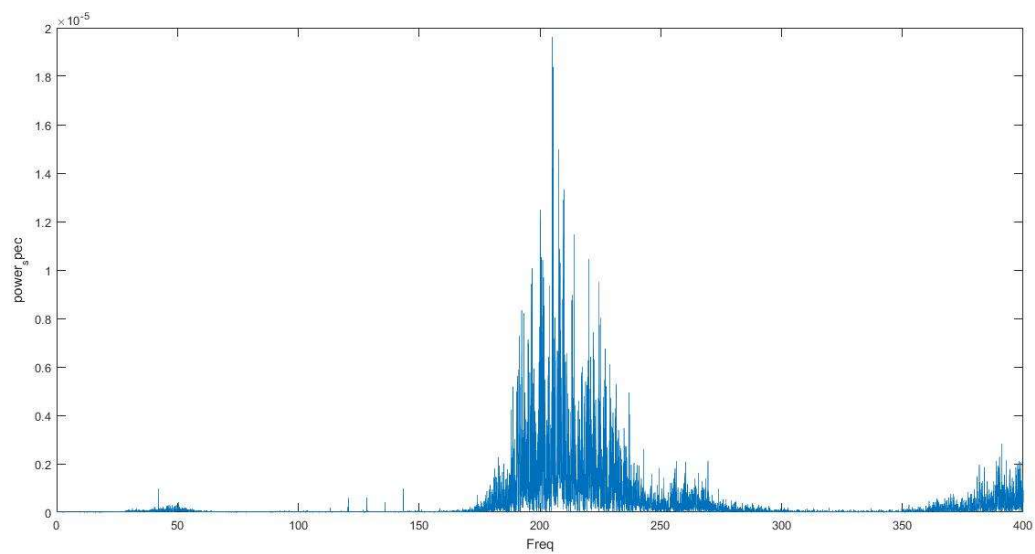
v3.mp3 – ϕ



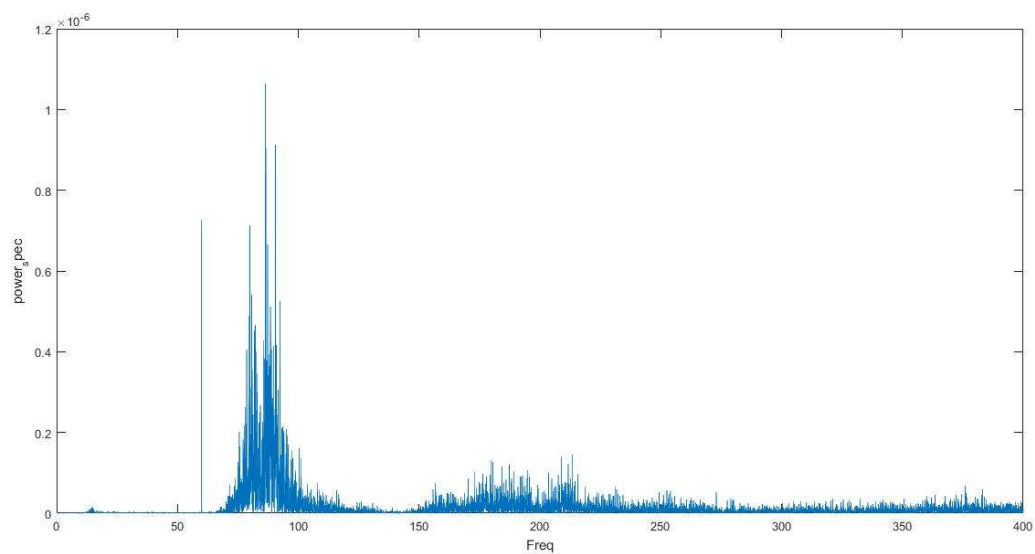
v4.mp3 - 0



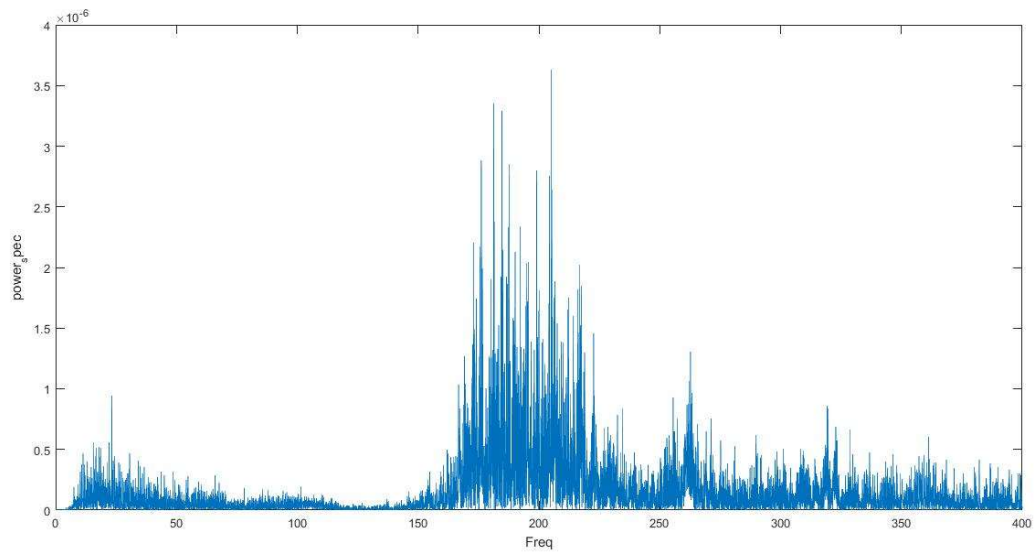
v5.mp3 – ξ



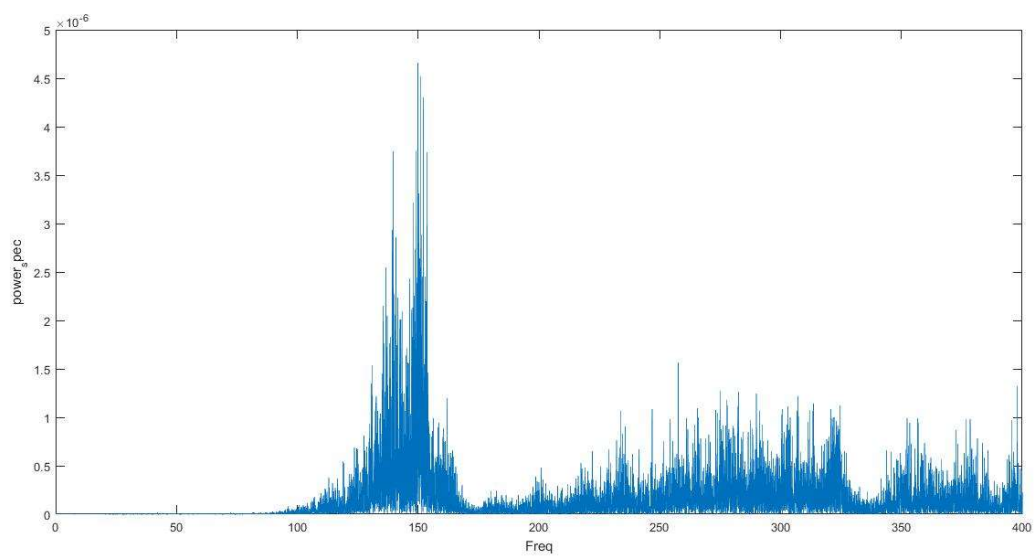
v6.mp3 – V



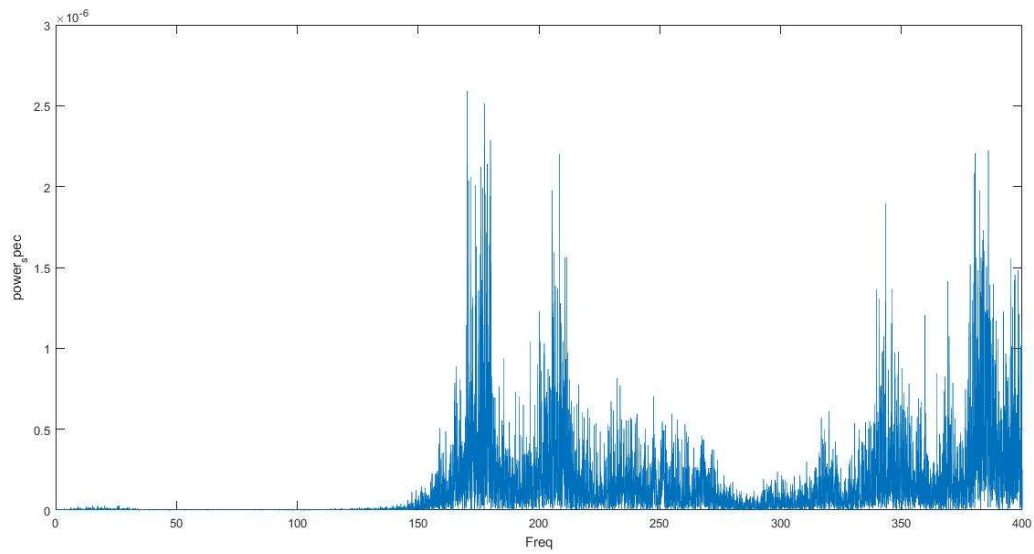
v7.mp3 – Λ



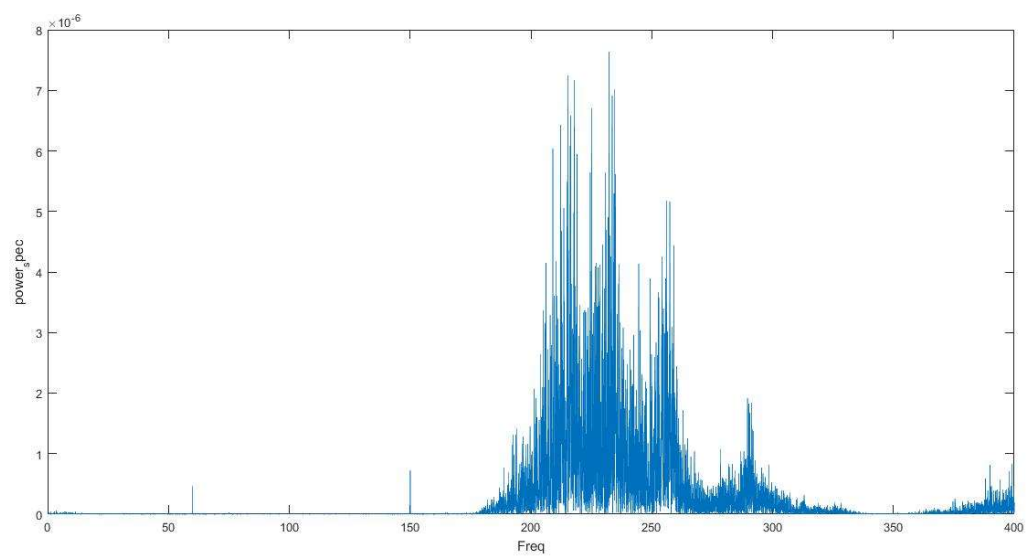
v8.mp3 - 9



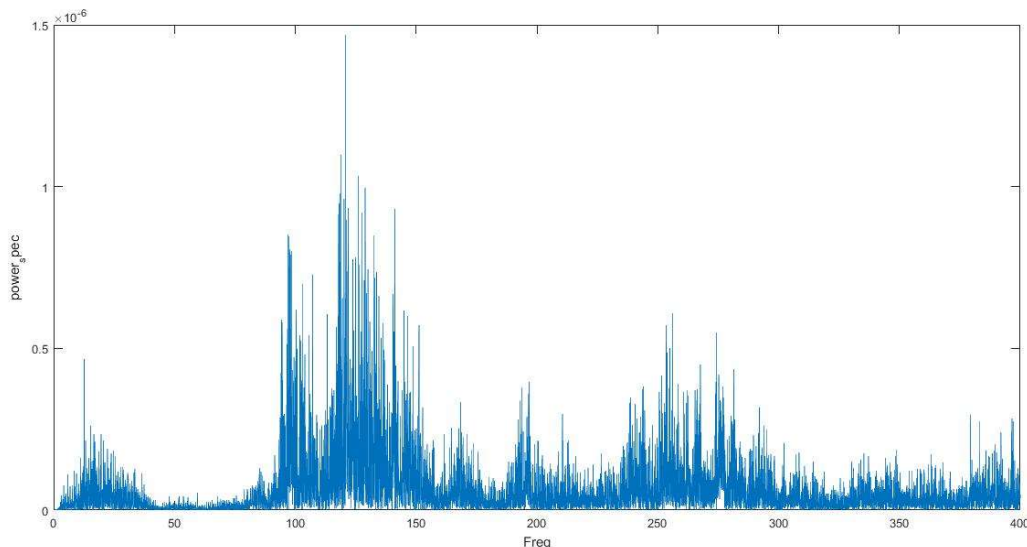
v9.mp3 - 1.



v10.mp3 - l1



v11.mp3 - l2



برای تشخیص جنسیت فرد از روی فرکانس ماکسیمم توان ، اندازه تفاضل این مقدار و ۱۲۲ هرتز(فرکانس پیک مردان) و اندازه تفاضل این مقدار و ۲۱۲ هرتز (فرکانس پیک زنان) و هر کدام که کمتر بود یعنی بیشتر به آن فرکانس پیک نزدیک است و بنابراین جنسیتش تعیین میشود. برای این کار ابتدا با تابع [peak_freq](#) مقدار فرکانس پیک را بر میگردانیم. سپس در تابع [gender_detection](#) همان کارهای گفته شده برای تشخیص جنسیت انجام میشود و مقدار 'male' یا 'female' برمیگردد. سپس فایل [labeling](#) براساس مقداری که تابع [gender_detection](#) برمیگرداند پیشوند female یا male را برای ویس ها اضافه میکند. و [file_labeling](#) این پیشوند را به اسم ویس ها اضافه میکند.

راهکار افزایش دقت تشخیص جنسیت :

برای این کار راهکاری که من در نظر گرفتم این است که به جای دخیل کردن تنها یک فرکانس پیک از یک بازه فرکانسی استفاده کرد. همان طور که در تعریف پروژه آمده فرکانس صدای مردان بین ۸۵ تا ۱۸۰ هرتز و برای زنان ۱۶۵ تا ۲۵۵ هرتز است. پس میتوان دو سیگما روی طیف توان سیگنال حساب کرد. یکی در بازه ی ۸۵ تا ۱۸۰ هرتز و یکی در بازه ی ۱۶۵ تا ۲۵۵ هرتز تابع [integ_freq](#) این کار را انجام میدهد . اگر جمع اول بیشتر بود مرد است و اگر دومی بیشتر بود زن است. تابع [gender_detection&labeling_powerful](#) هم تشخیص جنسیت و هم تعیین پیوند های فایل های یک دایرکتوری را برعهده دارد. سپس در [Files_labeling_powerful](#) این پیشوند ها به اول اسم ویس ها اضافه میشود.