آموزش پروژه‌محور پردازش صوت و گفتار با پایتون

فصل اول: سیگنال، ریاضیات و ابزارها

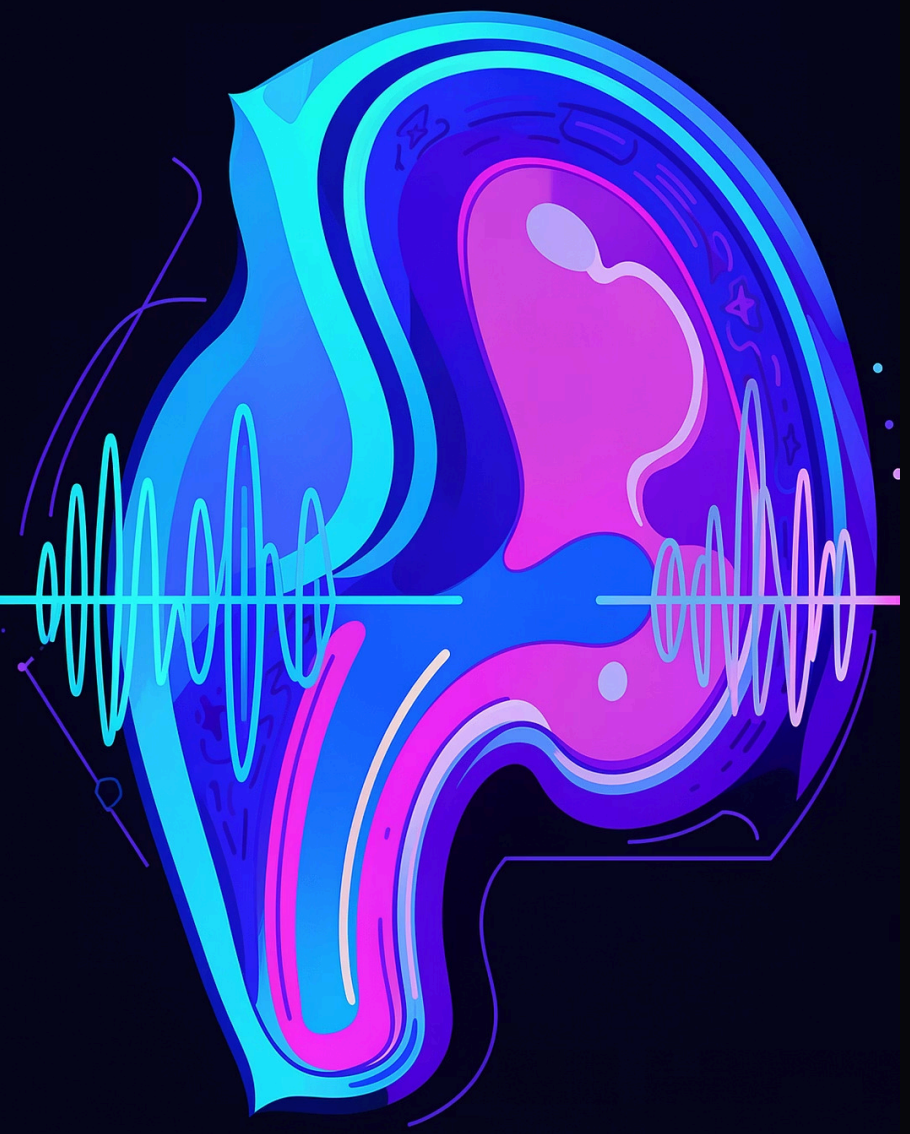
درس 4

اثر انگشت صوتی

طراح و مدرس: مصطفی کرمانی‌نیا

تحت نظارت علمی: دکتر مهدی سیفی‌پور

Made with GAMMA



سیستم شنوایی انسان: ما چطور می‌شنویم؟

گوش انسان یک ابزار خطی نیست!

تفاوت درک فرکانس

گوش ما حساسیت بسیار بالایی به فرکانس‌های پایین (مثل صدای گفتار) دارد، اما در فرکانس‌های بالا (مثل سوت کشیدن) تفاوت‌ها را به خوبی تشخیص نمی‌دهد.

نیاز به یک مدل جدید

ابزارهای ریاضی مثل تبدیل فوریه همه‌چیز را خطی می‌بینند. برای اینکه کامپیوتر مثل یک «انسان» بشنود، باید صدای ریاضی را به صدای انسانی ترجمه کنیم!

Made with GAMMA

نقشه راه: خلق اثر انگشت صوتی

شنوایی انسان
چرا درک ما از صدا خطی نیست؟

مقیاس مل (Mel Scale)
پل ریاضی بین انسان و ماشین

تولد MFCC
از اسپکتروگرام تا استخراج ویژگی‌های طلایی

کدتایم!
استخراج ۱۳ ویژگی جادویی در پایتون

Made with GAMMA

مقیاس مل (Mel Scale)؛ پل ارتباطی انسان و ماشین

درک شهودی فرمول

این فرمول قرار است فرکانس‌های پایین را با دقت بالا (زیر ذره‌بین) و فرکانس‌های بالا را با دقت کمتر بررسی کند.

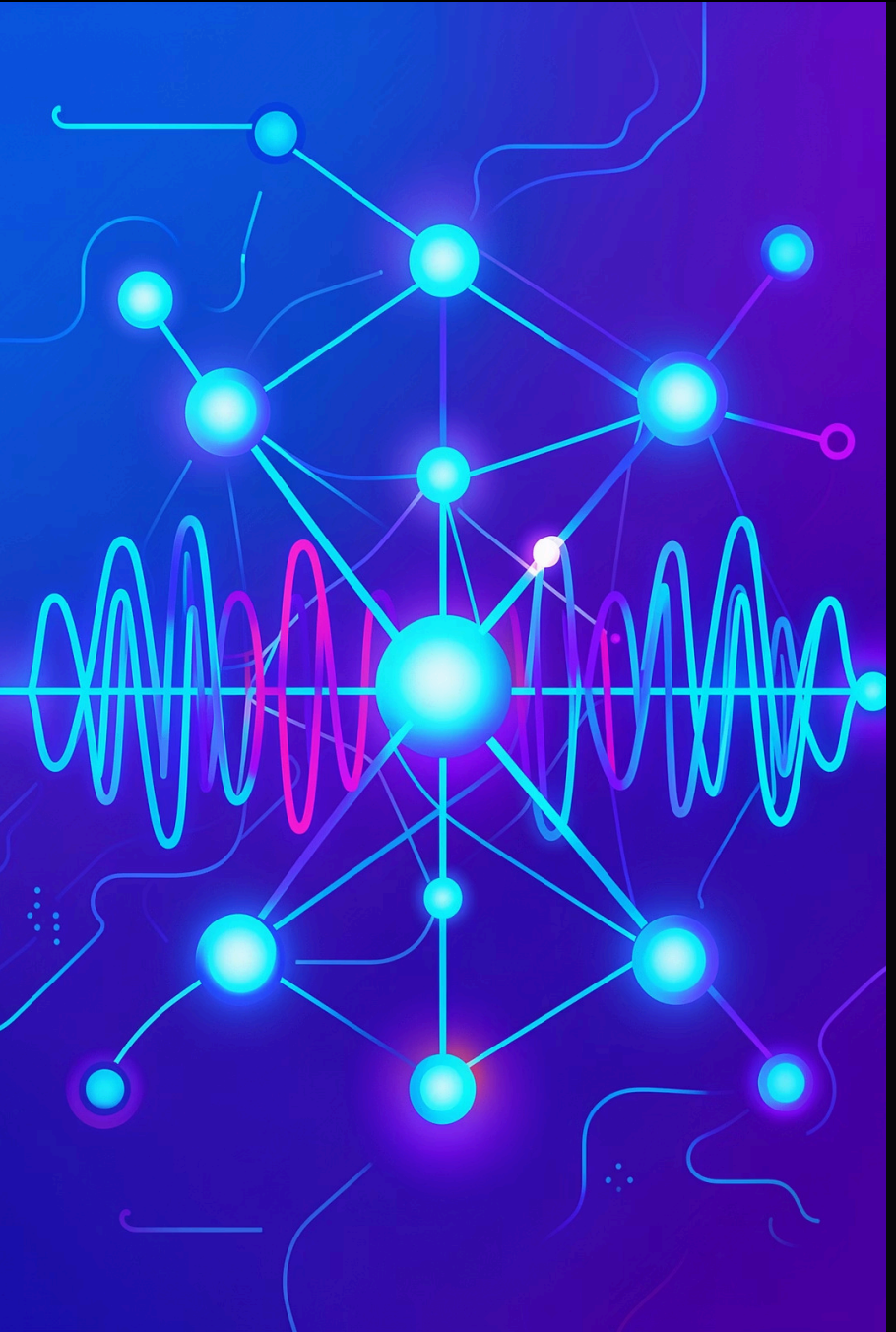
اهلی کردن ریاضیات

این فرمول به کامپیوتر یاد می‌دهد که دقیقاً مانند سیستم شنوایی انسان به فرکانس‌ها وزن بدهد و نویزهای فرکانس بالای غیرضروری را نادیده بگیرد.

فرمول مقیاس مل

$$m = 2595 \times \log_{10}\left(1 + \frac{f}{700}\right)$$

Made with GAMMA



جادوی MFCC چیست؟

استخراج ویژگی (Feature Extraction)

ویژگی‌های MFCC دقیقاً همان اعداد طلایی هستند که هوش مصنوعی برای درک کلام، تشخیص گوینده و درک احساسات به آن‌ها نیاز دارد.

چه بلایی سر سیگنال می‌آید؟

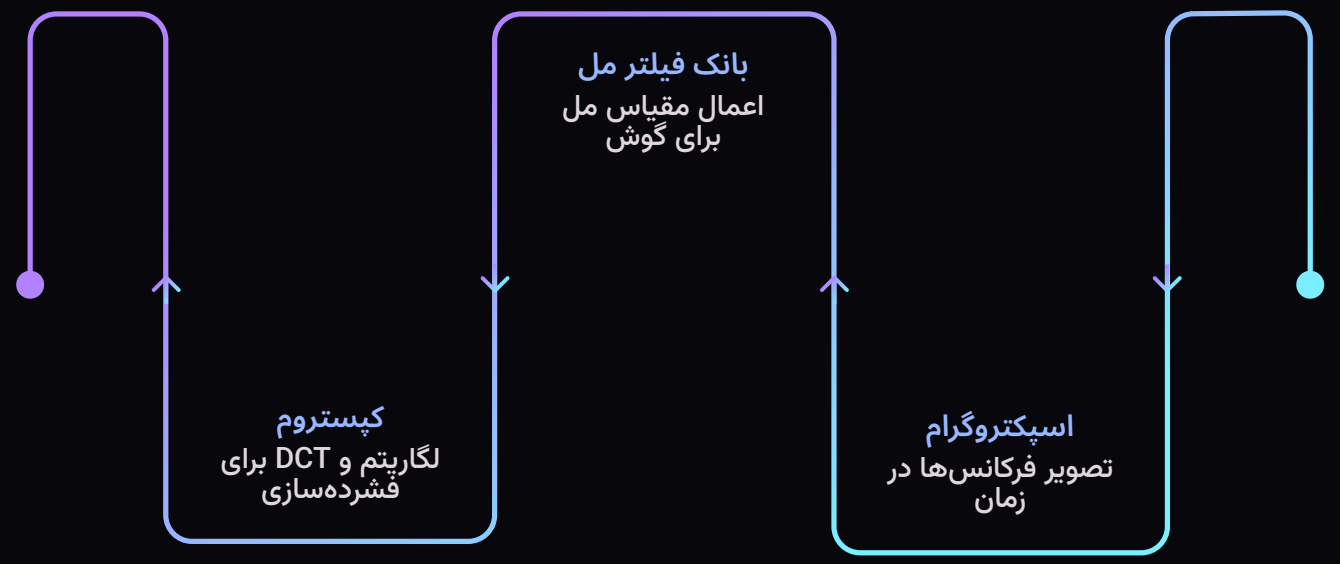
تمام اطلاعات اضافی MFCC را دور می‌ریزد و ساختار حنجره و مجرای صوتی گوینده را در قالب تنها ۱۳ تا ۴۰ عدد فشرده‌سازی می‌کند.

کجا استفاده می‌شود؟

دستیارهای صوتی (مثل Siri)، تشخیص احساسات از روی صدا، و سیستم‌های امنیتی احراز هویت گوینده.

Made with GAMMA

از اسپکتروگرام تا کپستروم (Cepstrum)

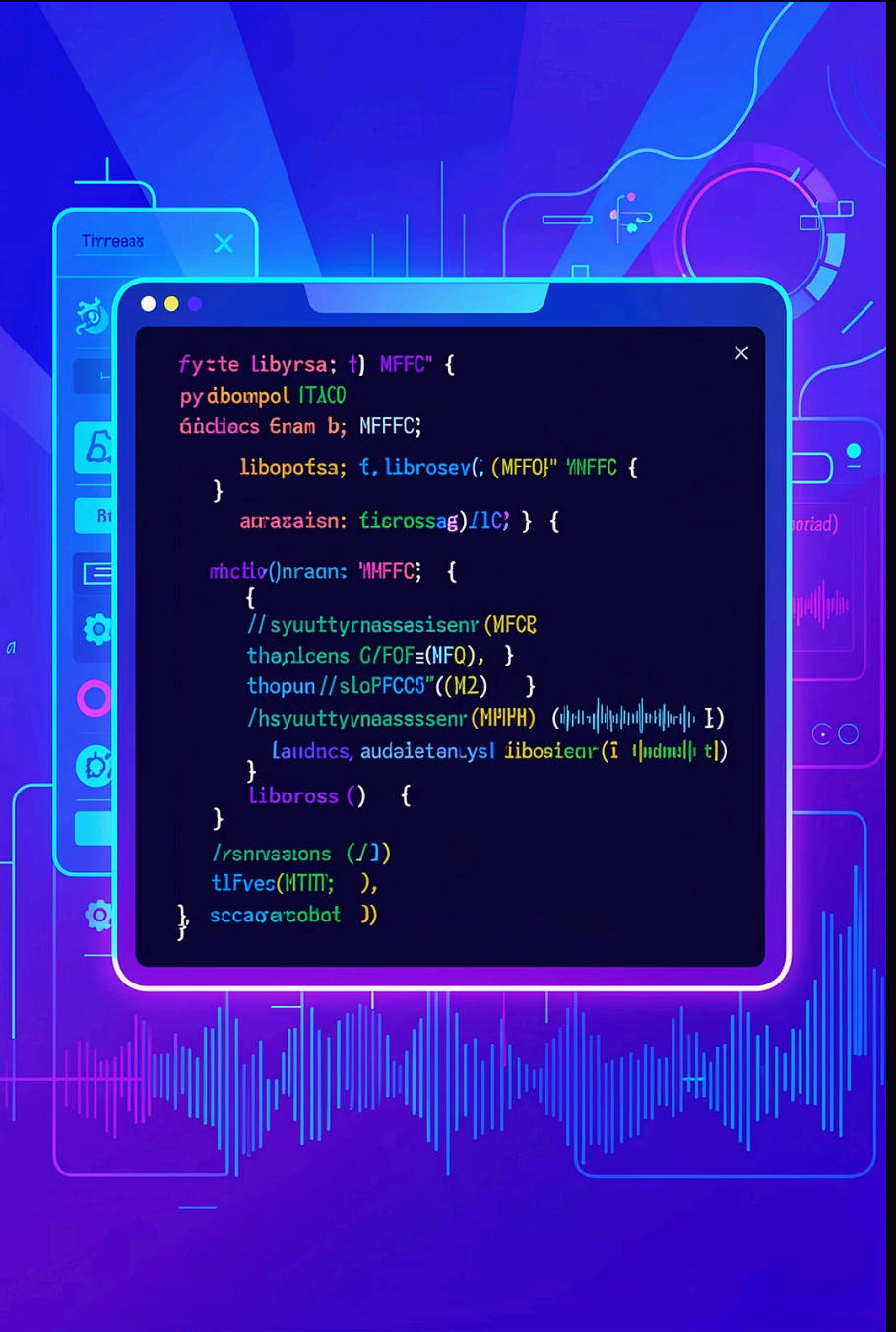


اسپکتروگرام: تصویر فرکانس‌ها در زمان (که در درس قبل با آن آشنا شدیم).

بانک فیلتر مل (Mel Filterbank): اعمال مقیاس مل روی اسپکتروگرام برای شبیه‌سازی دقیق گوش انسان.

کپستروم (Cepstrum): گرفتن لگاریتم و انجام یک تبدیل نهایی (DCT) روی داده‌ها برای فشرده‌سازی حداکثری و استخراج اثر انگشت خالص صدا.

Made with GAMMA



کدتایم!

استخراج اثر انگشت در عمل

محاسبه MFCC

استخراج ۱۳ ویژگی اصلی صدا با `librosa.feature.mfcc()`

رسم Heatmap

مصورسازی ماتریس ویژگی‌ها با `librosa.display.specshow()`

تحلیل الگوها

مشاهده اثر انگشت صوتی و تفاوت کلمات به صورت بصری

Made with GAMMA