



به نام خدا



گروه هوش مصنوعی، دانشکده مهندسی  
کامپیوتر

# گفتار پردازشی رقمی

## Digital Speech Processing

---

حمیدرضا برادران کاشانی

پاییز ۱۴۰۱



# اهداف درس:

## ❖ آشنایی با مباهات تئوریک در زمینه های زبانشناسی و پردازش گفتار

- مفاهیم آواشناسی و واج شناسی
- تولید و درک گفتار
- انواع ویژگی های گفتاری
- روش های مدلسازی سیگنال گفتار

- پردازش سیگنال رقمی
- روش های آماری شناسایی الگو
- روش های یادگیری ماشین
- ....

## ❖ آشنایی با کاربردهای پردازش گفتار

- ❖ بازشناسی گفتار (تبدیل گفتار به متن)
- ❖ سنتز گفتار (تبدیل متن به گفتار)
- ❖ شناسایی هویت گوینده
- ❖ بهسازی کیفیت گفتار
- ❖ فشرده سازی گفتار

Hamidreza Baradaran Kashani



# سرفصل مطالب

❖ مقدمه ای بر صوت و گفتار

❖ مفاهیم زبانشناسی

❖ آواشناسی، واج شناسی، آناتومی و اجزا سیستم تولید گفتار، آواهای زبان، مختصه های تولید آواها

❖ مدلسازی سیستم تولید و درک گفتار

❖ پیش پردازش گفتار

❖ مروری بر روش های استخراج ویژگی های سیگنال گفتار

❖ انرژی، نرخ عبور از صفر، اتوکورولیشن، تابع میانگین تفاضل دامنه، آنالیز کپسترال، آنالیز پیشگویی خطی

❖ تخمین طیف، فرمنت و گام

*Hamidreza Baradaran Kashani*



# سرفصل مطالب

- ❖ آشکارسازی فعالیت صوتی (VAD)
- ❖ مروری بر روش های طبقه بندی و مدلسازی
- ❖ پیچش زمانی پویا (DTW)، چندی سازی برداری (VQ) و مدل مخلوط گاوسی (GMM) و شبکه های عصبی
- ❖ شناسایی هویت گوینده با استفاده از صوت (Speaker Recognition)
- ❖ بازشناسی گفتار (Speech Recognition)
- ❖ سنتز گفتار یا تولید گفتار مصنوعی (Speech Synthesis)
- ❖ بهبود کیفیت گفتار (Speech Enhancement)

Hamidreza Baradaran Kashani



# مراجع درس

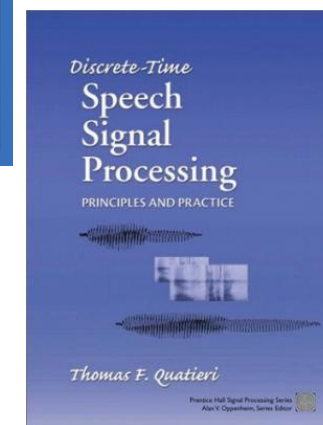
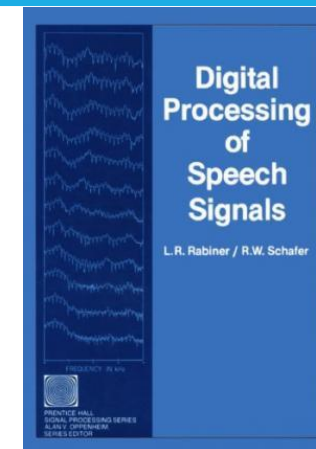
❖ کلیه کتاب ها قابل دانلود از سایت <https://z-lib.org/> هستند

❖ *Digital Processing of Speech Signal*, Rabiner 1978.

❖ *Discrete Time Processing of Speech Signal*, Deller, Hansen and Proakis, 1993.

❖ *Discrete-Time Speech Signal Processing*, Quatieri, 2001.

❖ *Spoken Language Processing*, Huang, Acero, Hon, 2001.



Hamidreza Baradaran Kashani



# مراجع درس

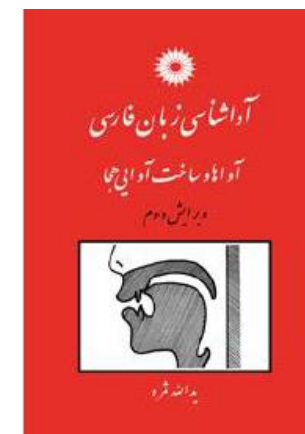
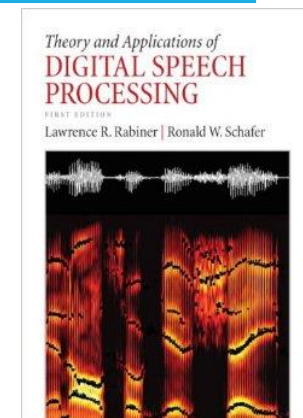
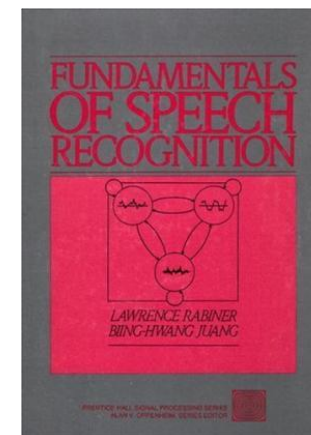
❖ *Fundamental of Speech Recognition*, Rabiner and Juang, 1993.

❖ *Theory and Applications of Digital Speech Processing*, Rabiner and Schafer, 2011.

❖ *Handbook of Speech Processing*, Benesty, Sondhi, Huang, 2008.

❖ آواشناسی زبان فارسی، ویرایش دوم، یدالله ثمره، ۱۳۷۸

❖ مبانی پردازش گفتار، سید سعید آیت، انتشارات پیام نور



Hamidreza Baradaran Kashani



# ابزارهای مورد نیاز

❖ **Wavesurfer**

❖ **Praat**

❖ **CoolEdit**

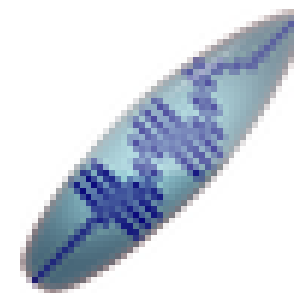
❖ **Matlab**

❖ **Voicebox**

❖ **Librosa**

❖ <https://librosa.org/doc/latest/index.html>

❖ A python package for music and audio analysis



*Hamidreza Baradaran Kashani*



# نحوه ارزیابی درس

## ❖ تکالیف درسی (نوشتاری و پیاده سازی)

- ❖ تمامی تمرینات بایستی طبق دستورالعمل بیان شده توسط تدریس یار درس ارسال شوند.
- ❖ تمرینات دارای پیاده سازی بایستی هم شامل کد و هم شامل گزارش مربوط به تحلیل نتایج باشد.

## ❖ سمینار / پروژه پیاده سازی

- ❖ دو مقاله از مقالات جدید در سه سال اخیر که توسط استاد درس تایید شود.
- ❖ یک مقاله فقط ارائه بصورت پاورپوینت / مقاله دوم شامل ارائه پاورپوینت / گزارش کتبی / پیاده سازی

## ❖ آزمون پایان ترم

Hamidreza Baradaran Kashani