

دانشکده علوم و فنون نوین

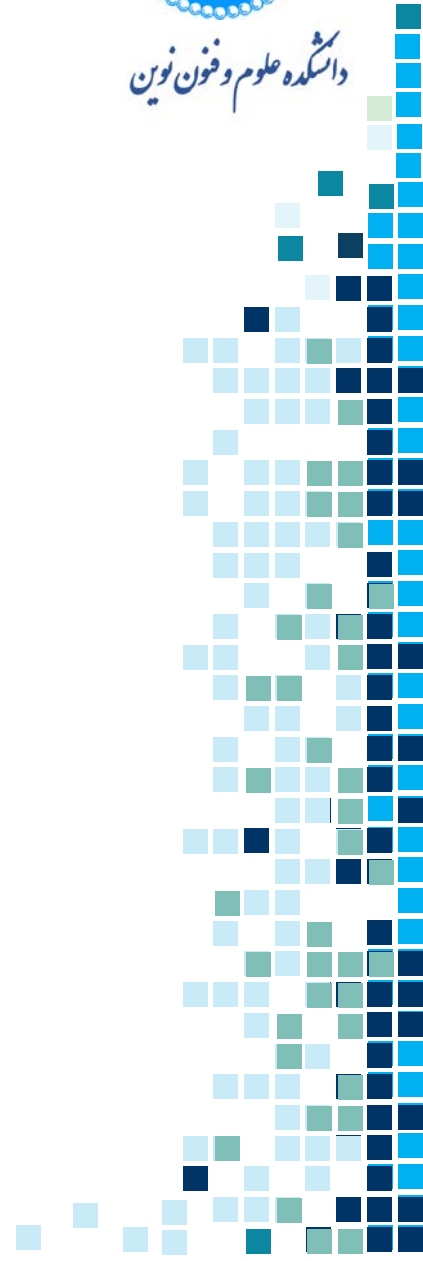
پردازش گفتار

Speech Processing

هادی ویسی

h.veisi@ut.ac.ir

نیمسال دوم ۱۴۰۰-۱۴۰۱



معرفی درس . . .

□ زمان و مکان

◀ شنبه و دوشنبه، ساعت ۸:۰۰ الی ۹:۳۰

◀ در سامانه eLearn

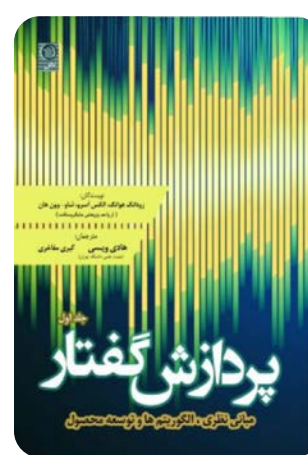
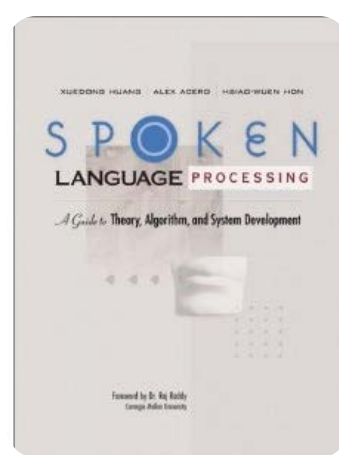
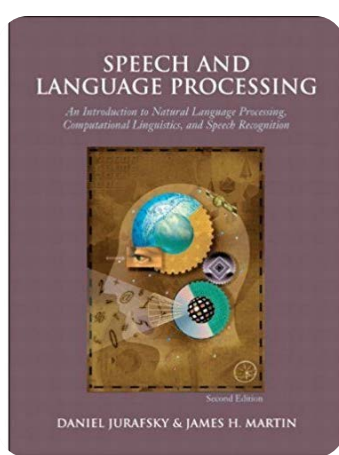
□ وب سایت

◀ dsp.ut.ac.ir

□ هدف

- ◀ مروری بر مفاهیم پایه پردازش گفتار و کاربردهای آنها
- ◀ بیان اصول و روش‌های پردازش سیگنال گفتار دیجیتال
- ◀ مبانی یادگیری ماشین و بازشناسی الگو: شبکه عصبی
- ◀ مروری بر روش‌های بازشناسی، سنتز، بهسازی گفتار و کدینگ
- ◀ فعالیت‌های تمرینی با رویکرد کاربردی

معرفی درس . . .



منابع

◀ هادی ویسی، کبری مفاخری، پردازش گفتار: مبانی نظری، الگوریتم ها و توسعه محصول (جلد اول)، نص، ۱۳۹۹

X Huang, A Acero, HW Hon, Spoken Language Processing, Prentice Hall, Upper Saddle River, NJ, USA, ۲۰۰۱.

◀ Daniel Jurafsky and James H. Martin, *Speech and Language Processing*, Pearson Education (۲nd edition, ۳rd Ed), ۲۰۰۸, ۲۰۲۱.

◀ Lawrence R. Rabiner, Ronald W. Schafer, Theory and Applications of Digital Speech Processing, Pearson, ۲۰۱۱.

◀ J. R. Deller, J. G. Proakis, J. H. Hansen, Discrete Time Processing of Speech Signals, Prentice Hall PTR Upper Saddle River, NJ, USA, ۱۹۹۳.

معرفی درس: ارزیابی . . .

□ تمرین (برای هر موضوع)

- ◀ همفکری و همکاری در یافتن پاسخ سوالها توصیه می‌شود
- ◀ در صورت کپی بودن یکی یا چند مورد از پاسخها، کل نمره آن تمرین در نظر گرفته نمی‌شود.
- ◀ تمرین‌های دارای پیاده‌سازی، باید هم شامل کدها و هم شامل گزارش مربوطه باشد
- ◀ تاخیر در تحویل
 - ◁ ارسال پاسخ حداکثر تا ساعت ۲۳:۵۹ مهلت تعیین شده
 - ◁ هر یک ساعت تاخیر در ارسال پاسخها (از یک ثانیه تا ۶۰ دقیقه!)، کسر یک درصد نمره به عنوان جریمه تاخیر
- ◀ ارسال پاسخ تمرینها
 - ◁ تنها به صورت الکترونیکی و به ایمیل استاد درس است.
 - ◁ همه فایل‌های مرتبط با یک تمرین را در یک فایل فشرده شده
 - ◁ فرمت نام‌گذاری فایل ارسالی: `Speech_Family_StNo_HW#`
 - مثال: `Speech_Ahmadi_830496001_HW1.rar`
- ◀ وزن تمرین‌های مختلف با هم برابر نیست



معرفی درس: ارزیابی . . .

□ آزمونک (کويز)

◀ در کلاس‌های غیرحضورى نداريم، در صورت حضورى بودن خواهيم داشت!

□ امتحان ميانه ترم

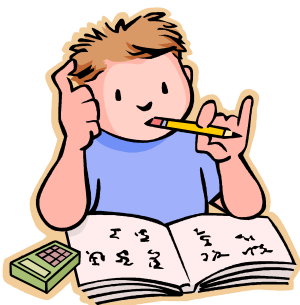
◀ مطالب تدریس شده تا تاریخ اعلام شده برای این امتحان

◀ زمان: دوشنبه ۱۴۰۱/۰۲/۰۵ ساعت ۸:۰۰

□ امتحان پايان ترم

◀ شامل کلیه مطالب تدریس شده، از جمله مطالب میان ترم

◀ زمان: مطابق برنامه دانشگاه



معرفی درس: ارزیابی . . .

□ پروژه: یک مورد

- ◀ پروژه کاربردی دارای پیاده‌سازی در Python یا سایر زبان‌های برنامه‌نویسی
- ◀ عدم نیاز به نوآوری: پیاده‌سازی یک پژوهش (مقاله، پایان‌نامه و ...) موجود
- ◀ علاوه بر کد برنامه، گزارش مکتوب (به صورت تایپ شده)، داده‌ها و منابع هم تحویل گرفته می‌شود
- ◀ در صورت مساعد شدن شرایط کرونایی، تحویل به صورت حضوری است
- ◀ آخرین زمان پیشنهاد موضوع توسط دانشجو: روز دوشنبه ۱۴۰۱/۰۲/۰۵
- ◀ تحویل پروژه: اولین هفته بعد از آخرین امتحان پایان‌ترم
- ◀ موضوع الزاما مرتبط با مباحث درس باشد
- ◀ برخی موضوع‌های پیشنهادی

◀ تشخیص احساس در گفتار با استفاده از یادگیری عمیق

◀ تشخیص گفتار برای تعداد کلمات محدود

◀ تبدیل متن به گفتار با استفاده از شبکه‌های عمیق مانند مبدل‌ها یا GAN

◀ بهسازی گفتار و حذف نویز با شبکه‌های عصبی عمیق

◀ بازشناسی دیداری-شنیداری (Audio-Visual) گفتار



معرفی درس: ارزیابی . . .

□ حضور و مشارکت در کلاس (نمره اضافی)

◀ مشارکت فعال در بحث‌های کلاس

□ مقاله (نمره اضافی)

◀ مقاله ارسال شده مورد قبول است

◀ به هر قیمتی مقاله ننویسید!

◀ همکاری با این درس در نوشتن مقاله را به اطلاع استاد خود برسانید

□ بازنگری نمره‌ها و برگه‌ها

◀ در زمان تحویل پروژه درس (به صورت حضوری)

◀ اولین دوشنبه بعد از آخرین امتحان پایان ترم



معرفی درس: ارزیابی . . .

عنوان	وزن	توضیح
تمرین	۵۰%	بعد از هر موضوع (وزن تمرین‌ها برابر نیست)
آزمونک (کويز)	-	در شرایط غیرحضورى کلاس‌ها ندارم
امتحان میان‌ترم	۱۵%	دوشنبه ۱۴۰۱/۰۲/۰۵ ساعت ۸:۰۰
امتحان پایان‌ترم	۲۰%	از کل مطالب درس، مطابق برنامه دانشگاه در صورت داشتن آزمونک وزن این آزمون ۱۵% خواهد بود
پروژه	۱۵%	موضوع اختیاری، تعیین موضوع تا شنبه ۱۴۰۱/۰۲/۰۵ تحويل پروژه: اولین دوشنبه بعد از آخرین امتحان پایان‌ترم
حضور و مشارکت کلاس (نمره اضافی)	۵%	پاسخ دادن به سوالات حین تدریس و مشارکت در بحث‌های کلاس
مقاله (نمره اضافی)	۱۵%	مقاله ارسال شده به مجله/کنفرانس مورد قبول است



معرفی درس . . .



دستیاران آموزشی

پدرام عبدزاده

pabdzadeh@gmail.com

آتنا عرب احمدی

atena.arabahmadi@ut.ac.ir

برگزاری کلاس‌های حل تمرین و رفع اشکال

به ویژه راهنمایی پیاده‌سازی و کدنویسی

معرفی درس: سرفصل‌ها . . .

- مروری بر مفاهیم و کاربردهای پردازش گفتار
- مروری بر پردازش سیگنال دیجیتال
 - ◀ مبانی سیگنال‌ها و سیستم‌ها
 - ◀ تبدیل فوریه و Z
- مروری بر آمار و احتمال
 - ◀ نظریه احتمال و توزیع‌های مختلف
 - ◀ نظریه تخمین
- مروری بر یادگیری ماشین
 - ◀ مبانی یادگیری ماشین
 - ◀ شبکه عصبی و یادگیری عمیق

معرفی درس: سرفصل‌ها . . .

□ ساختار زبان گفتاری

- ◀ تولید گفتار و درک گفتار (اندام‌ها، واک‌داری، طیف‌نگار و فرمنت)
- ◀ آواشناسی و واج‌شناسی: فارسی و انگلیسی
- ◀ هجاها و واژه‌ها

□ نمایش سیگنال گفتار

- ◀ مدل منبع-فیلتر
- ◀ تحلیل فوریه کوتاه‌مدت
- ◀ تحلیل LPC
- ◀ تحلیل کپستروم
- ◀ روش MFCC

معرفی درس: سرفصل‌ها . . .

□ بازشناسی گفتار

- ◀ پیچیدگی‌ها
- ◀ روش‌های ارزیابی
- ◀ روش‌های بازشناسی گفتار
- ◀ انطباق زمانی پویا (DTW)
- ◀ شبکه عصبی مصنوعی (ANN) و یادگیری عمیق
- ◀ مدل مخفی مارکوف (HMM)

معرفی درس: سرفصل‌ها . . .

□ سنتز گفتار

- ◀ روش‌ها و مسائل پردازش متن
- ◀ مشکلات پردازش متن
- ◀ تحلیل نوایی
- ◀ روش‌های تولید سیگنال گفتار (سنتز)
 - ◀ سنتز فرمندی
 - ◀ سنتز پیوندی و انتخاب واحد (Unit Selection/Concatenative)
 - ◀ سنتز آماری پارامتری
 - ◀ سنتز مبتنی بر یادگیری عمیق

معرفی درس: سرفصل‌ها . . .

□ بهسازی گفتار

- ◀ روش‌های بهسازی گفتار (تک کاناله)
- ◀ شناخت نویز
- ◀ معیارهای ارزیابی
- ◀ روش تفریق طیفی
- ◀ تخمین‌گر کمترین خطای مربعات میانگین (MMSE) (طیف: فیلتر وینر، اندازه طیف و لگاریتم اندازه طیف)
- ◀ بهسازی گفتار مبتنی بر یادگیری (HMM و یادگیری عمیق)

□ کد کردن (کدینگ) گفتار

- ◀ روش‌های خانواده PCM
- ◀ روش CELP



معرفی درس

زمان بندی

متناسب شرایط و سطح کلاس، و همچنین تغییرات پیش بینی نشده در زمان بندی، ممکن است سرفصل مطالب و یا زمان بندی های کلاس مقداری تغییر داشته باشد

هفته	تاریخ	موضوع	توضیح
۱	۲۳ و ۲۴ / ۱۴۰۰/۱۱/۲۵	معرفی درس	
۲	۱۴۰۰/۱۱/۳۰ و ۱۴۰۰/۱۲/۰۲	مروری بر مفهوم و کاربردهای پردازش گفتار ساختار زبان گفتاری	
۳	۱۴۰۰/۱۲/۰۹ و ۰۷	ساختار زبان گفتاری	
۴	۱۴۰۰/۱۲/۱۶ و ۱۴	مروری بر پردازش سیگنال دیجیتال	
۵	۱۴۰۰/۱۲/۲۳ و ۲۱	تمرین مروری بر پردازش سیگنال دیجیتال	
۶	۱۴۰۰/۱۲/۲۸	مروری بر آمار، احتمال و روش های تخمین	
۷	۱۴۰۱/۰۱/۱۵	مروری بر آمار، احتمال و روش های تخمین	
۸	۱۴۰۱/۰۱/۲۲ و ۲۰	نمایش سیگنال گفتار و استخراج ویژگی	
۹	۱۴۰۱/۰۱/۲۹ و ۲۷	تمرین نمایش سیگنال گفتار و استخراج ویژگی	
۱۰	۱۴۰۱/۰۲/۰۵ و ۰۳	آزمون میان ترم	
۱۱	۱۴۰۱/۰۲/۱۲ و ۱۰	مروری بر بازشناسی الگو و یادگیری ماشین	
۱۲	۱۴۰۱/۰۲/۱۹ و ۱۷	تمرین مروری بر شبکه عصبی	
۱۳	۱۴۰۱/۰۲/۲۶ و ۲۴	یادگیری عمیق	
۱۴	۱۴۰۱/۰۲/۳۱ و ۱۴۰۱/۰۳/۰۲	تمرین یادگیری عمیق در پردازش گفتار	
۱۵	۱۴۰۱/۰۳/۰۹ و ۰۷	بازشناسی گفتار: معرفی و DTW	
۱۶	۱۴۰۱/۰۳/۱۶ و ۱۴	تمرین بازشناسی گفتار: مدل مخفی مارکوف (HMM)	
۱۷	۱۴۰۱/۰۳/۲۳ و ۲۱	بازشناسی گفتار: یادگیری عمیق	

۱۴۰۱/۰۲/۰۳: شهادت حضرت علی
 ۱۴۰۱/۰۲/۱۲: عید فطر
 ۱۴۰۱/۰۳/۱۴: ...