

پردازش گفتار Speech Processing

هادی ویسی

h.veisi@ut.ac.ir

دانشگاه تهران - دانشکده علوم و فنون نوین

نیمسال دوم ۱۳۹۹-۱۳۹۸



معرفی درس ...

○ زمان و مکان

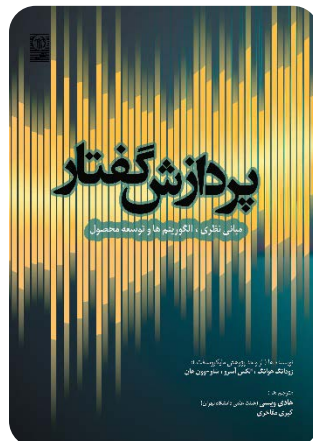
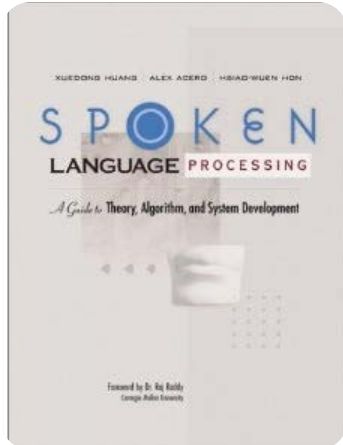
- شنبه و دوشنبه، ساعت ۸:۰۰ الی ۹:۳۰
- دانشکده علوم و فنون نوین

○ 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



- X Huang, A Acero, HW Hon, Spoken Language Processing, Prentice Hall, Upper Saddle River, NJ, USA, 2001.
- J. R. Deller, J. G. Proakis, J. H. Hansen, Discrete Time Processing of Speech Signals, Prentice Hall PTR Upper Saddle River, NJ, USA, 1993.
- Lawrence R. Rabiner, Ronald W. Schafer, Theory and Applications of Digital Speech Processing, Pearson, 2011.

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
 - در صورت داشتن یک روز تاخیر در ارسال پاسخها (از یک ثانیه تا ۲۴ ساعت!)، ۲۵٪ نمره آن تمرین به عنوان جریمه تاخیر
 - در صورت تاخیر دو روزه ۵۰٪ نمره مربوطه به عنوان جریمه تاخیر
 - پس از دو روز نمره‌ای در نظر گرفته نمی‌شود
- 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)

- از مطالب هر بخش
- ممکن است بدون اطلاع قبلی باشد

• امتحان میان ترم

- مطالب تدریس شده تا تاریخ اعلام شده برای این امتحان
- زمان: دوشنبه ۱۳۹۹/۰۲/۰۱ - ساعت ۸:۰۰

• امتحان پایان ترم

- شامل کلیه مطالب تدریس شده (از جمله مطالب امتحان میان ترم)

- زمان: مطابق برنامه دانشگاه، فعلا ۱۳۹۹/۰۳/۲۶



معرفی درس ...

○ ارزیابی ...

• پروژه: یک مورد

- پروژه کاربردی دارای پیاده‌سازی در MATLAB، Python یا سایر زبان‌های برنامه‌نویسی
- علاوه بر کد برنامه، گزارش مکتوب (به صورت تایپ شده)، داده‌ها و منابع هم تحویل گرفته می‌شود
- تحویل به صورت حضوری است
- آخرین زمان تعیین موضوع پروژه: روز دوشنبه ۲۶/۱۲/۱۳۹۸
- تحویل پروژه: اولین هفته بعد از آخرین امتحان پایان‌ترم (شنبه ۰۷/۰۴/۱۳۹۹)

• 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)
- دوشنبه شنبه ۰۷/۰۴/۱۳۹۹



○ ارزیابی

عنوان	وزن	توضیح
تمرین	۳۰٪	بعد از هر موضوع (وزن تمرین‌ها برابر نیست)
آزمونک (کويز)	۱۰٪	ممکن است بدون اطلاع قبلی باشد
امتحان میان‌ترم	۲۰٪	دوشنبه ۱۳۹۹/۰۲/۰۱ ساعت ۸:۰۰
امتحان پایان‌ترم	۲۵٪	از کل مطالب درس (مطابق تقویم دانشگاه، فعلا ۱۳۹۹/۰۳/۲۶)
پروژه	۱۵٪	موضوع اختیاری تعیین موضوع تا دوشنبه ۱۳۹۸/۱۲/۲۶ تحويل پروژه: اولین شنبه بعد از آخرین امتحان پایان‌ترم (۱۳۹۹/۰۴/۰۷)
مقاله (نمره اضافی)	۱۵٪	مقاله ارسال شده به مجله/کنفرانس مورد قبول است



معرفی درس ...

○ دستیار آموزشی

• وریا فتحی

○ verya.fathy@gmail.com





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 بازشناسی الگو
 - مبانی
 - درخت تصمیم
 - شبکه عصبی
 - یادگیری عمیق



معرفی درس ...

○ سرفصل‌ها ...

- ساختار زبان گفتاری

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

- نمایش سیگنال گفتار

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

- بازشناسی گفتار

- پیچیدگی‌ها
- روش‌های ارزیابی



معرفی درس ...

○ سرفصل‌ها ...

• بازشناسی گفتار

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

• 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



معرفی درس ...

○ سرفصل‌ها

• بهسازی گفتار

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

• 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	توضیحات
۱	۱۳۹۸/۱۱/۱۴ و ۱۲	معرفی درس	
۲	۱۳۹۸/۱۱/۲۱ و ۱۹	مروری بر مفهوم و کاربردهای پردازش گفتار	
۳	۱۳۹۸/۱۱/۲۸ و ۲۶	مروری بر پردازش سیگنال دیجیتال	
۴	۱۳۹۸/۱۲/۰۳ و ۰۵	مروری بر پردازش سیگنال دیجیتال	کويز
۵	۱۳۹۸/۱۲/۱۰ و ۱۲	مروری بر آمار، احتمال و روش های تخمین	تمرین
۶	۱۳۹۸/۱۲/۱۷ و ۱۹	مروری بر بازشناسی الگو و یادگیری ماشین	
۷	۱۳۹۸/۱۲/۲۴ و ۲۶	مروری بر درخت تصمیم	
۸	۱۳۹۹/۰۱/۱۸ و ۱۶	مروری بر شبکه عصبی	کويز
۹	۱۳۹۹/۰۱/۲۵ و ۲۳	مروری بر شبکه عصبی	تمرین
۱۰	۱۳۹۹/۰۱/۳۰	یادگیری عمیق	
	۱۳۹۹/۰۲/۰۱	ساختر زبان گفتاری	
		امتحان میان ترم	
۱۱	۱۳۹۹/۰۲/۰۸ و ۰۶	ساختر زبان گفتاری	
۱۲	۱۳۹۹/۰۲/۱۵ و ۱۳	نمایش سیگنال گفتار و استخراج ویژگی	تمرین
۱۳	۱۳۹۹/۰۲/۲۲ و ۲۰	نمایش سیگنال گفتار و استخراج ویژگی	کويز
۱۴	۱۳۹۹/۰۲/۲۹ و ۲۷	بازشناسی گفتار: معرفی و DTW	
۱۵	۱۳۹۹/۰۳/۰۳ و ۰۵	بازشناسی گفتار: مدل مخفی مارکوف (HMM)	
۱۶	۱۳۹۹/۰۳/۱۰ و ۱۲	سنتز گفتار: معرفی و پردازش متن	تمرین