به نام خدا

دانشکده سامانه‌های هوشمند

پردازش گفتار (۱۴۸-۰۵-۸۳)

نیم‌سال دوم ۱۴۰۴-۱۴۰۳

تمرین شماره ۲

تاریخ تحویل: ۱۴۰۴/۰۱/۳۱

---

۱. (۴۰٪) [نظری- پردازش سیگنال]

الف. جدول زیر را تکمیل کنید (برای هر ستون، پاسخ "بله/خیر" داده و به‌صورت مختصر دلیل را توضیح دهید).

سیستم	خطی	تغییرناپذیر با زمان	علّی
$y(t) = 3x(t)\cos(t)$			
$y(t) = \sqrt{x^2(t)}$			
$y(t) = \int_{t}^{t+1} x(\lambda)d$			
$y[n] = 2(x[n+1]u[n] - x[n]) + 1$			
$y[n] = \begin{cases} +1, x[n] \geq 0 \\ -1, x[n] < 0 \end{cases}$			

ب. تبدیل فوریه سیگنال‌های زیر را محاسبه کنید.

$$[e^{-a\,t}cos(\omega_0 t)]u(t), \qquad a > 0$$

$$e^{3t}\,sin(2t)$$

$$\frac{\sin(\pi t)\sin(2\pi t)}{\pi t}$$

پ. خروجی یک سیستم خطی تغییرناپذیر با زمان (LTI) علّی به ورودی x(t) از طریق معادله دیفرانسیل زیر مرتبط است:

$$\frac{d}{dt}y(t) + 2y(t) = x(t)$$

پاسخ فرکانسی سیستم $H(\omega) = \frac{Y(\omega)}{X(\omega)}$ را تعیین کنید و اندازه و فاز $H(\omega)$ را رسم کنید.

h.veisi@ut.ac.ir

به نام خدا

پردازش گفتار (۱۴۸-۰۵-۸۳)

نیم‌سال دوم ۱۴۰۴-۱۴۰۳

تمرین شماره ۲

تاریخ تحویل: ۱۴۰۴/۰۱/۳۱

دانشکده سامانه‌های هوشمند

اگر $x(t) = e^{-t}u(t)x(t)$ باشد، تبدیل فوریه خروجی $Y(\omega)$ را بیابید.

y(t) را برای ورودی داده‌شده در بخش قبل محاسبه کنید.

---

۲. (۳۰٪) **[عملی- تحلیل اثرات تغییر طول داده و فرکانس‌ها]**

با استفاده از روش **افزودن صفر در حوزه فرکانس**، سیگنال زیر را به نرخ نمونه‌برداری بالاتر تبدیل کنید.

$$x(t) = 6\cos(2\pi 8.5t) \text{ Volts}, \quad \text{for } 0 \le t < 2 \text{ seconds}$$

این سیگنال در ابتدا با نرخ $f_s = 32$ نمونه در ثانیه نمونه‌برداری شده است. هدف این است که آن را به نرخ نمونه‌برداری جدید $f_{s_{\text{new}}} = 256$ نمونه در ثانیه افزایش دهید و ۲ ثانیه داده (۶۴ نمونه) تولید کنید. مراحل زیر را دنبال کنید:

۱. سیگنال را با نرخ ۳۲ نمونه در ثانیه تولید کنید.

۲. تبدیل فوریه گسسته (DFT) این سیگنال را محاسبه کنید.

۳. مقدار مربوط به فرکانس $0.5f_s$ را صفر قرار دهید.

۴. تعداد مناسبی صفر در اطراف $0.5f_s$ درج کنید و اطمینان حاصل کنید که تقارن مزدوج مختلط حول $0.5f_{s_{\text{new}}}$ حفظ شده است.

۵. نتیجه را به درستی مقیاس‌بندی کرده و تبدیل معکوس را محاسبه کنید. بررسی کنید که قسمت موهومی بسیار کوچک باشد (در حالت ایده‌آل صفر، اما معمولاً خطای گرد کردن وجود دارد). قسمت حقیقی را استخراج کنید.

۶. سیگنال نمونه‌برداری‌شده جدید را رسم کنید و در همان نمودار، نمونه‌های اصلی نمایش دهید.

ب) فرآیند بالا را تکرار کنید، اما این بار به جای تولید ۶۴ نمونه (۲ ثانیه داده)، **80 نمونه (۲.۵ ثانیه داده)** تولید کنید و سپس با استفاده از روش افزودن صفر، نرخ نمونه‌برداری را به ۲۵۶ نمونه در ثانیه افزایش دهید. بررسی کنید که در نزدیکی $0.5f_s$ (نصف نرخ نمونه‌برداری اولیه) چه

h.veisi@ut.ac.ir

اتفاقی می‌افتد و توضیح دهید که چگونه در روش افزایش نرخ نمونه‌برداری این مسئله را مدیریت می‌کنید. نتایج خود را تحلیل و تفسیر کنید.

---

۳. (۳۰٪) [عملی- تحلیل فوریه و نوشتن کد تبدیل فوریه گسسته خودتان]

در این سوال، شما باید تبدیل فوریه گسسته (DFT) خود را پیاده‌سازی کنید و از آن برای تجزیه و تحلیل یک سیگنال واقعی به منظور اندازه‌گیری نوسانات عصبی از طریق طیف توان استفاده کنید. لطفاً در هنگام انجام تمرینات در نوت‌بوک، مراحل زیر را تکمیل کنید:

۱. تابع تبدیل فوریه گسسته (DFT) را با استفاده از سینوسی‌های مختلط و ضرب داخلی همانطور که در کلاس توضیح داده شد، پیاده‌سازی کنید.
۲. پیاده‌سازی DFT خود را بر روی یک سیگنال واقعی برای دو واج /f/ و /e/ که توسط خودتان با نرخ نمونه برداری ۱۶ کیلوهرتز و ۱۶ بیت ضبط شده است، اعمال کرده و نمودارهای توان طیف و فاز آن را رسم کنید.
۳. مشاهده‌های کلیدی خود را در مورد نوسانات عصبی در طیف توان توضیح دهید.

اطمینان حاصل کنید که کد خود را به طور کامل مستند کرده و مشاهده‌های خود را در سلول‌های مربوطه توضیح دهید. پس از تکمیل تمام مراحل، نوت‌بوک را ارسال کنید.

---

۴. (۲۵٪ امتیازی) [پیاده‌سازی کاربردی: چت‌بات صوتی فارسی]

الف) در این سوال می‌خواهیم یک چت‌بات شبیه زیگپ (نسخه اندروید و PWA و تحت وب آن در سایت Zigap.ir است) بسازیم. بدین صورت که ابتدا در خود پایتون برنامه‌ی شما صدا (با فرض اینکه شما فارسی حرف می‌زنید) را ضبط می‌کند (برای این کار پیشنهاد می‌شود از کتابخانه‌ی pyaudio کمک بگیرید) سپس صدای ضبط شده باید به یک مدل تشخیص گفتار که فارسی را پشتیبانی می‌کند مثل مدل هزار، داده شود تا تبدیل به متن شود.

https://github.com/hezarai/hezar

h.veisi@ut.ac.ir

حال باید از api یکی از هوش مصنوعی ها مانند chatgpt یا gemini یا copilot یا DeepSeek یا ... استفاده کنید. پیشنهاد می‌شود از api زیر که نیاز به api key ندارد استفاده کنید

https://github.com/xtekky/gpt4free

حال متن تبدیل شده خود را چاپ کرده و به مدل دهید و جواب بات را نیز دریافت کنید و چاپ کنید. برنامه شما باید هر بار از کاربرد بپرسد که می‌خواهد ادامه دهد و صدای جدید ضبط کند یا از برنامه خارج شود.

در انتها باید تمام گفتگوها ما بین کاربر و چت بات در یک فایل متنی (.txt) ذخیره شود.

در گزارش خود بیاورید که با تغییر پارامترهای مختلف در هنگام ضبط صدا مانند sample_rate(frame_rate). آیا در هنگام تبدیل صدا به متن تفاوتی پیدا می‌شود یا خیر. می‌توانید برای سنجش این کار چند بار یک جمله ی ثابت را با پارامتر های مختلف تست کنید) همه پارامترها را نام ببرید و برای هر کدام با ثابت نگه داشتن بقیه حداقل دو مقدار را امتحان کنید.

ب) می‌توان برای هر بار استفاده از برنامه ابتدا از کاربر اسم کاربر و پسورد خاص خود را بگیریم و در دیتابیس چک کنیم اگر قبلا این کاربر و پسورد دارای گفتگو است، ابتدا گفتگو قبلی را بخوانیم و به بات نشان دهیم، سپس در اخر برنامه نیز گفتگوهای جدید به این گفتگو اضافه شوند و در فایل txt مربوط به آن کاربر نوشته شوند.

نکته: استفاده از هر کتابخانه‌ای در این سوال کاملا مجاز است.

h.veisi@ut.ac.ir