



---

۱. (۱۰٪) [آوانویسی] دنباله واجی جمله "به جای آنکه به تاریکی لعنت بفرستید، یک شمع روشن کنید" را با هریک از واحدهای آوایی خواسته شده در زیر بنویسید (برای یکسان بودن علائم مورد استفاده برای واج‌نویسی از جدول موجود در فایل ارائه با عنوان "علائم IPA" برای فارسی استفاده کنید).

الف) واج

ب) سیلاب (هجا)

---

۲. (۱۵٪) [تحلیل شکل موج] شکل موج‌های زیر (با فرکانس نمونه‌برداری 10KHz) مربوط به سه کلمه انگلیسی "shoot"، "cease" و "room" می‌باشند. ابتدا مشخص کنید هر شکل مربوط به کدام کلمه است؛ سپس قسمت‌های مختلف واجی هر کلمه را در نمودار مربوط به آن تعیین نمایید.

بسم خدا

پروازش گفتار (۸۳-۰۵-۱۴۸)

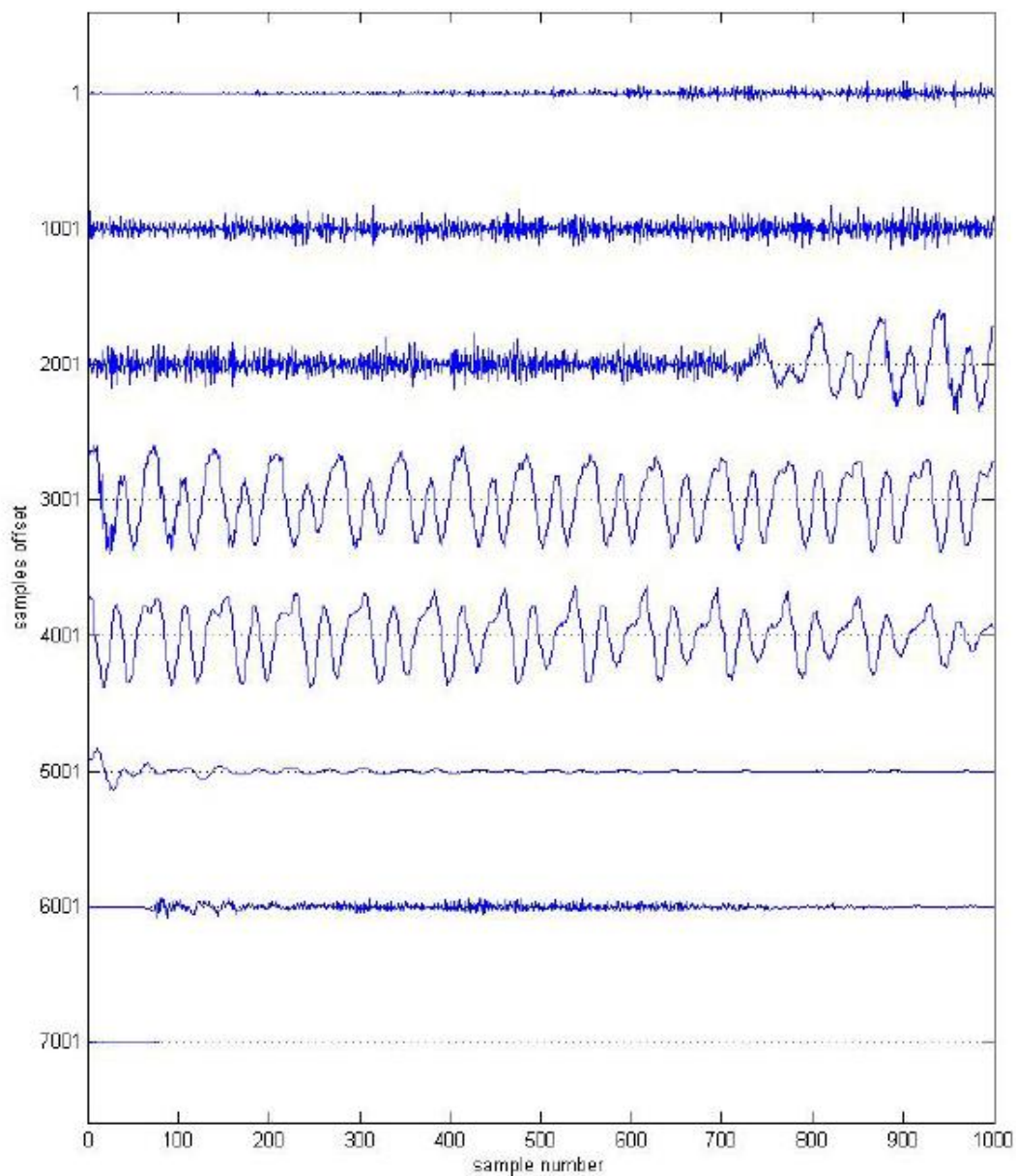
نیمسال اول ۱۳۹۶-۱۳۹۷

تاریخ تحویل: ۱۳۹۶/۱۰/۲۰

تمرین شماره ۳



دانشکده علوم و فنون نوین



شکل ۳-الف: شکل موج

بسم خدا

پردازش گفتار (۸۳-۰۵-۱۴۸)

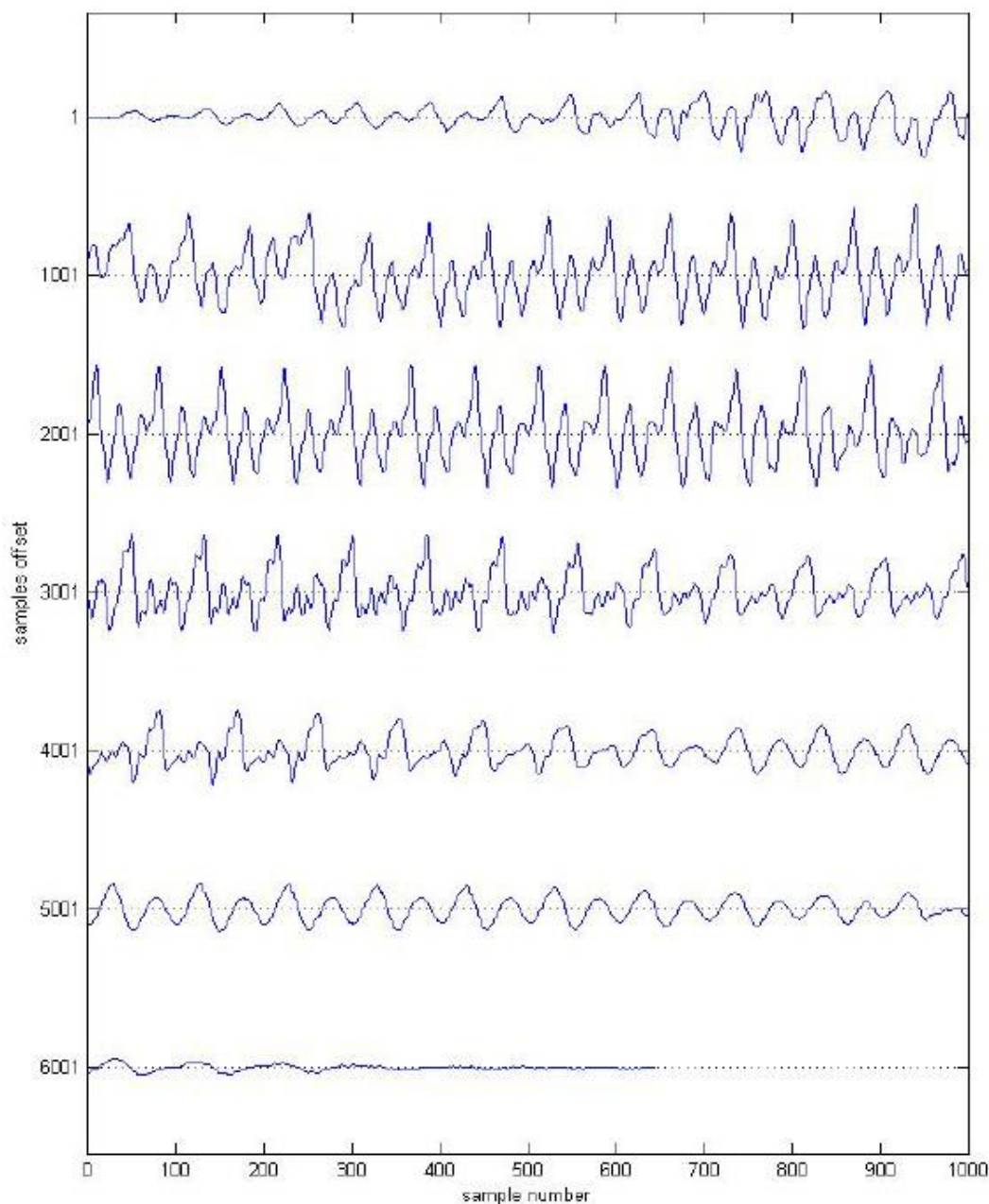
نیم سال اول ۱۳۹۶-۱۳۹۷

تاریخ تحویل: ۱۳۹۶/۱۰/۲۰

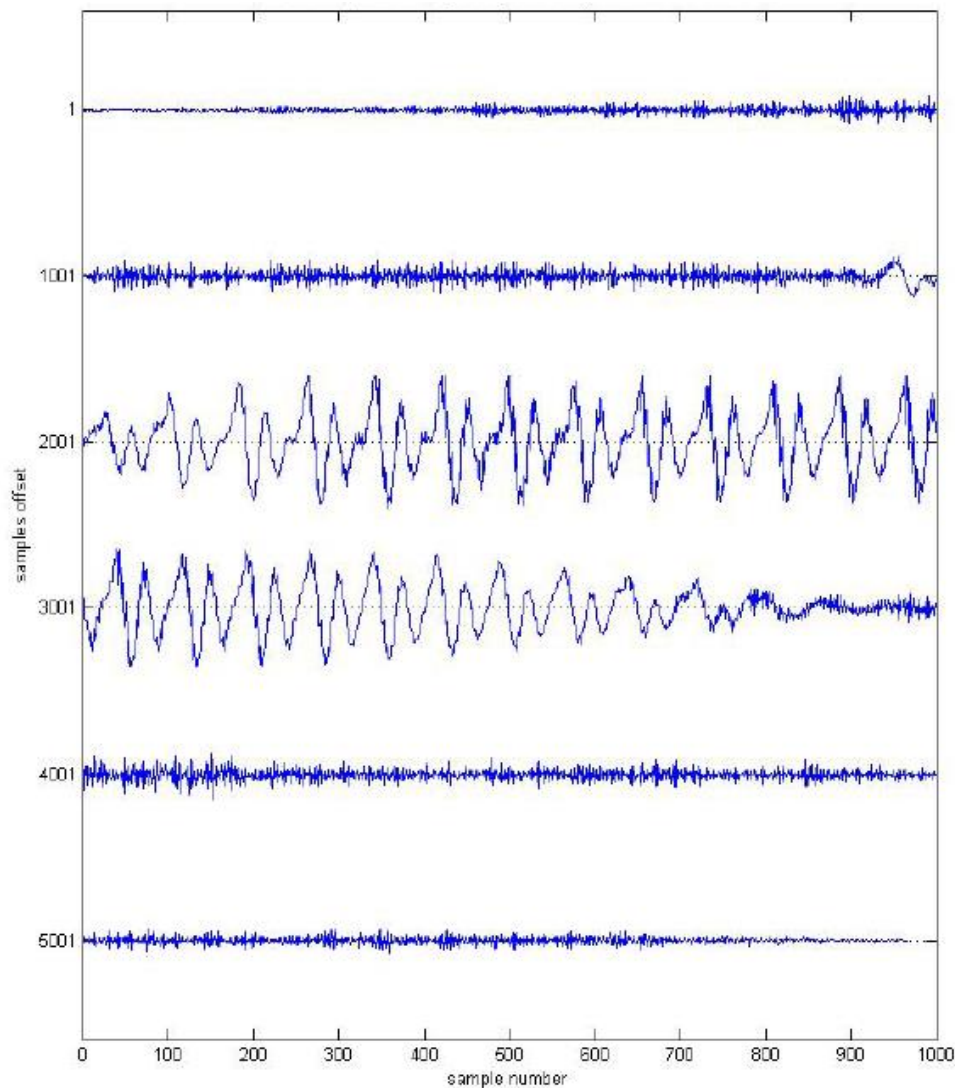
تمرین شماره ۳



دانشکده علوم و فنون نوین

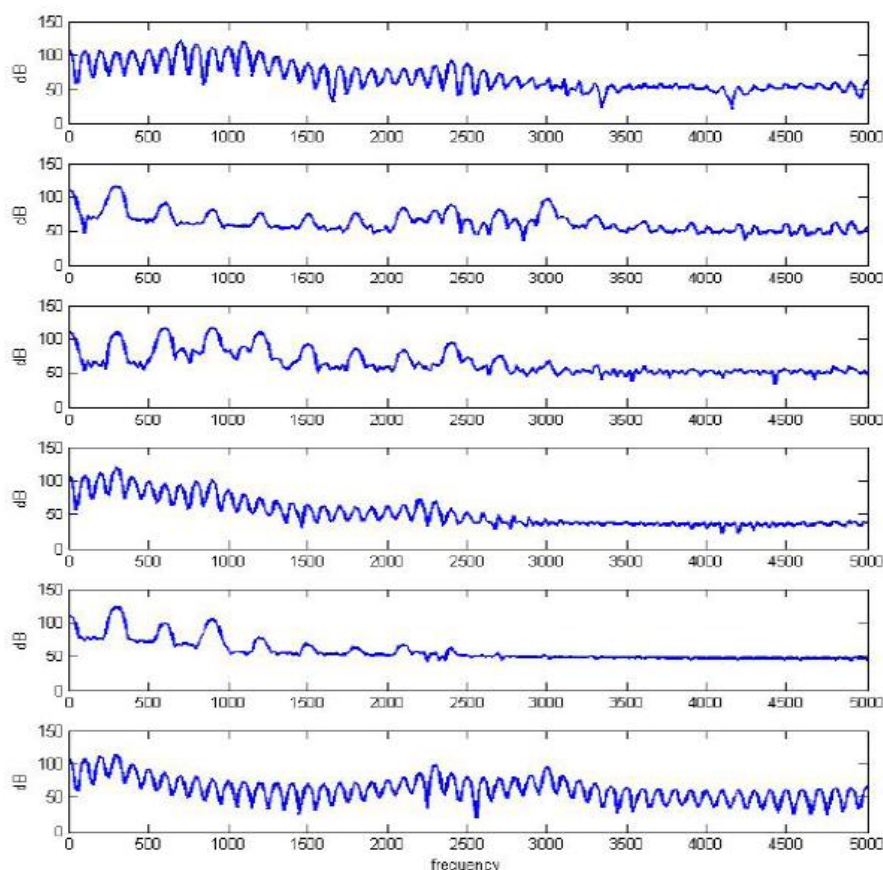


شکل ۳-ب: شکل موج



شکل ۳- ج: شکل موج

۳. (۹٪) [تشخیص واج از روی طیف] شکل زیر نمودار طیف مربوط به واکلهای /a/ /i/ و /u/ را نشان می‌دهد که پس از اعمال پنجره با طول و نوع مناسب ایجاد شده‌اند. برای هریک از واکلهای ممکن است بیش از یک نمودار نشان داده شده باشد. مشخص کنید هریک از این نمودارها مربوط به چه واکلهای می‌باشد. دلیل خود را بنویسید.



۴. (۱۶٪) [تحلیل اسپکتروگرام] با استفاده از یکی از نرم افزارهای نمایش و تحلیل سیگنال گفتار (مانند CoolEdit، Adobe Audition و WaveSurfer) اسپکتروگرام دسته واج های زیر را مشاهده کنید و برداشت خود را در هر مورد بیان کنید. برای این کار، می توانید صدای خود را از طریق نرم افزارهای مذکور ضبط کنید و یا نمونه گفتاری را که دربرداشته واج مورد نظر است، استفاده کنید.

الف) مقایسه واکه های /a/ و /o/

ب) مقایسه بین دو هم خوان سایشی /s/ و /z/

ج) مقایسه بین دو هم خوان انسدادی /p/ و /b/

د) تفاوت دو هم خوان خیشومی /m/ و /n/ در دو کلمه "نار" و "مار"





۵. (۵۰٪) [پیاده سازی: تشخیص اعداد انگلیسی با شبکه عصبی و ویژگی های MFCC] مشابه تمرین قبل از شبکه عصبی پس انتشار خطا (MLP) برای تشخیص اعداد انگلیسی صفر تا نه استفاده کنید. برای این کار از دادگان صوتی تمرین شماره ۲ استفاده کنید و برای پیاده سازی شبکه عصبی می توانید از ابزارهای آماده بهره بگیرید. ابتدا تمام فایل های صوتی را هم طول کنید؛ برای این کار همه فایل ها را با اضافه کردن صفر به اول و آخر آن به اندازه طولانی ترین فایل در آورید. ساختار شبکه عصبی را ثابت در نظر بگیرید؛ تعداد نرون های ورودی را برابر با تعداد کل ویژگی های فایل، تعداد نرون های میانی را تقریباً نصف تعداد نرون های ورودی و تعداد نرون های خروجی را به تعداد دسته ها (۱۰ دسته) در نظر بگیرید. مقدار نرخ یادگیری را عدد کوچکی قرار دهید.

الف) از ویژگی های MFCC با تعداد ۱۲ ویژگی استاتیک و ویژگی های دینامیک (به همراه انرژی کلا ۳۹ ویژگی برای هر فریم)، پنجره همینگ و فریم های با طول ۲۰ میلی ثانیه با ۵۰٪ هم پوشانی استفاده کنید. دقت (درصد تعداد اعدادی که درست تشخیص داده شده) و ماتریس درهم ریختگی (confusion matrix) مجموعه آزمون را به عنوان نتیجه گزارش کنید. برای محاسبه MFCC از کد گذاشته شده روی سایت در متلب استفاده کنید.

ب) برای هر کدام از فایل ها، ویژگی ها را با روش CMS نرمال کرده، شبکه را دوباره آموزش دهید و نتایج را بدست آورید. نتایج حاصل از این روش را با نتایج بخش الف مقایسه کنید.

ج) نتایج بدست آمده در این تمرین را با بهترین نتایج تمرین ۲ (بر حسب دقت) در یک جدول با همدیگر مقایسه کنید.

راهنمایی - ماتریس درهم ریختگی (Confusion Matrix): برای ارزیابی دقیق تر، به جای محاسبه خطای کلی، خطاهای مربوط به هر عدد را به تفکیک و به صورت زیر در یک



ماتریس (که آن را ماتریس درهم ریختگی یا ماتریس خطا می گویند) نمایش دهید. هر سطر این ماتریس بیانگر دسته های واقعی و هر ستون آن بیانگر دسته های تشخیص داده شده توسط روش شماست. در هر سلول مقدار دقت دسته بندی را بیان کنید. به عنوان مثال در جدول زیر، در ۹۳٪ نمونه های عدد صفر، به درستی به عنوان صفر تشخیص داده شده است، در ۰٫۵٪ حالت ها به عنوان یک و .... در ۱٫۲٪ حالت ها معادل نه تشخیص داده شده است. بدیهی است که جمع اعداد هر سطر برابر با ۱۰۰ خواهد بود و میانگین اعداد سطر اصلی میانگین کل دقت روش شما می باشد.

دسته تشخیص داده شده					دسته واقعی
صفر	یک	...	نه		
صفر	۹۳	۰٫۵	...	۱٫۲	
یک	۰٫۰	۹۶٫۵٪	...	۱٫۶	
.	.	.	.	.	
نه	۰٫۶	۱٫۴	۹۴٫۲	.	