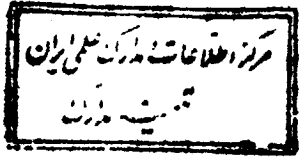


بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

٢٧٥٠١



۱۳۷۸ / ۱۷ ۵

دانشگاه علم و صنعت ایران  
دانشکده مهندسی کامپیوتر

# بازشناسی گفتار پیوسته فارسی بکمک شبکه‌های عصبی

۱۴۹۹۰

نگارش: منیژه منتظریان

اساتید راهنما: دکتر احمد اکبری، دکتر عادل رحمانی

تابستان ۱۳۷۸

۲۷۸۰۱

### تقدیم به مادر،

به او که به هر آنچه می‌نگرد نیکو می‌شود و  
شنیدن آوازش دل را از مهر و صفا می‌آکند

### و تقدیم به پدر،

که خشم و خودستایی از برابرش می‌گریزند  
و آنگاه که به لبخند لب می‌گشاید، نه به وصف می‌آید  
و نه دریاد می‌گنجد.

## چکیده

گفتار محصول سیستمهای تولید و درک گفتار و مغز انسان است. انسان همیشه از طریق گفت و شنود توانسته است ارتباط بهتری با محیط خود برقرار کند. بنابراین اگر بتوان از کامپیوتر بصورت سمعی و بصری بهره گرفت، تحول بزرگی در استفاده از آنها بوجود می آید.

در این پروژه، بمنظور طراحی روشهایی در بازشناخت گفتار پیوسته فارسی، شبکههای عصبی بعنوان ابزار مدلسازی انتخاب شده اند. در بخش نخست اجرای پروژه، باهدف دستیابی به یک روش مناسب مرزبندی، سه روش تعیین مرکزکلمات برگزیده و مقایسه شده اند. سپس، از یک الگوریتم ساده و کارا در استخراج هجاهای عبارات استفاده شده است. بمنظور محدود ساختن دامنه گسترده لغات فارسی، کلمات را به اعداد دورقمی فارسی محدود ساخته ایم. بنابراین، از این پس هر جاکه از مضامین "عبارت" و "کلمه" استفاده شده است، منظور اعداد دورقمی فارسی است. در بخش بعدی پروژه، ضرایب کپسترال و دلتا کپسترال از هجاهای بدست آمده از عبارات تلفظ شده توسط یک گوینده زن، استخراج می شوند. بمنظور دستیابی به ویژگی از گفتار که در تعلیم مدل شبکه عصبی بازشناخت بهترین کارایی را داشته باشد، از ضرایب بدست آمده در تشکیل سه گونه متفاوت بردار ویژگی استفاده می شود.

بمنظور مقایسه عملکرد مدل‌های شبکه با روشهای مختلف تعلیم، در انجام مدلسازی از دو شبکه MLP و کوهونن بهره برده ایم. ساختارهای مختلفی از این دو مدل را تعلیم داده و نتایج حاصل از عملکرد آنها را با یکدیگر مقایسه می کنیم.

## کلید واژگان

شناسایی مرزها، آستانه انرژی، شبکه عصبی، قطعه بندی خودکار، هجا، بازشناسی گفتار، پرسپترون چندلایه، شبکه خودسازمانده کوهونن، آموزش بدون نظارت، آموزش بانظارت

## تقدیر و تشکر :

در انجام مراحل مختلف این پروژه، راهنماییهای ارزشمند اساتید محترم همواره تعیین کننده مسیر پیشرفت تحقیق بوده است. بدینوسیله زحمات ایشان را ارج نهاده، بر خود لازم میدانم از اساتید محترم جناب آقای دکتر احمد اکبری و جناب آقای دکتر عادل رحمانی که بارهنگامیهای خویش در مدت تحصیل و در مراحل مختلف انجام پروژه، مرا یاری نموده‌اند، صمیمانه تشکر و قدردانی نمایم.

تابستان ۷۸

فهرست اشکال

فهرست جداول

## فصل اول - مسئله باز شناخت گفتار

۲	۱-۱) مقدمه .....
۴	۲-۱) مکانیزم تولیدگفتار .....
۶	۳-۱) مدل ریاضی تولیدگفتار .....
۷	۴-۱) اجزای تشکیل دهنده یک سیستم باز شناسی .....
۷	۱-۴-۱) نمونه برداری از سیگنال صوتی .....
۹	۲-۴-۱) پیش پردازش .....
۹	۳-۴-۱) تعیین ابتدا و انتهای عبارات و قطعه بندی آن .....
۱۰	۴-۴-۱) استخراج ویژگی از سیگنال گفتار .....
۱۱	۱-۴-۴-۱) آنالیز بانک فیلتر .....
۱۲	۲-۴-۴-۱) آنالیز پیشگوی خطی .....
۱۶	۳-۴-۴-۱) آنالیز کیسترال .....
۱۷	۴-۴-۴-۱) مشتقات ضرایب کیسترال .....

۱۸	..... ۱-۴-۴-۵ استفاده از مقیاس MEL در آنالیز کپسترال
۲۲	..... ۱-۴-۵ انطباق الگو
۲۳	..... ۱-۴-۶ پردازش زبان
۲۴	..... ۱-۵ عوامل موثر در باز شناخت گفتار
۲۶	..... ۱-۶ ارزیابی کارایی سیستمهای باز شناسی گفتار
۲۷	..... ۱-۷ جمع بندی

## فصل دوم - روشهای شبکه عصبی در باز شناخت گفتار

۲۹	..... ۲-۱ مقدمه
۳۱	..... ۲-۲ شبکههای عصبی بیولوژیکی
۳۳	..... ۲-۳ ساختار نرون عصبی
۳۶	..... ۲-۴ مروری بر مدل های ANN
۳۹	..... ۲-۵ پرسپترون چند لایه
۴۵	..... ۲-۶ چند گیر بردار یادگیری
۴۹	..... ۲-۷ جمع بندی

## فصل سوم - تعیین مرزکلمات و قطعه‌بندی سیگنال گفتار

۵۱	..... (۱-۳) مقدمه
۵۲	..... (۲-۳) آوای زبان فارسی
۵۴	..... (۳-۳) تعیین مرزکلمات
۵۴	..... (۱-۳-۳) بکارگیری شبکه‌های عصبی
۵۷	..... (۲-۳-۳) استفاده از مقادیر آستانه انرژی
۶۲	..... (۳-۳-۳) ارزیابی روشها
۶۴	..... (۴-۳) قطعه‌بندی سیگنال گفتار
۶۴	..... (۱-۴-۳) انتخاب واحد تشخیص مناسب
۶۶	..... (۲-۴-۳) روش پوش محدب
۶۹	..... (۵-۳) جمع‌بندی

## فصل چهارم - بکارگیری شبکه MLP گفتار

۷۲	..... (۱-۴) مقدمه
۷۴	..... (۲-۴) الگوریتم تعلیم انتشاریه عقب (BP)
۸۰	..... (۳-۴) دادگان



۸۱	..... استخراج بردارهای ویژگی. (۱-۳-۴)
۸۳	..... پیش تاکید کردن (۱-۱-۳-۴)
۸۳	..... قاب بندی. (۲-۱-۳-۴)
۸۴	..... پنجره گذاری. (۳-۱-۳-۴)
۸۵	..... استخراج ضرایب کپسترال و دلتا کپسترال (۴-۱-۳-۴)
۸۵	..... فشرده سازی ضرایب (۵-۱-۳-۴)
۸۶	..... ساختار شبکه (۴-۴)
۸۷	..... گرامر (۵-۴)
۸۹	..... جمع بندی (۶-۴)

### فصل پنجم - شبکه خودسازمانده کوهونن و رویه تعلیم ناظر

۹۱	..... مقدمه (۱-۵)
۹۲	..... مدل خودسازمانده کوهونن. (۲-۵)
۹۲	..... اصول شبکه خودسازمانده کوهونن (۱-۲-۵)
۹۶	..... تعمیم شبکه خودسازمانده کوهونن (۲-۲-۵)
۹۸	..... رویه تعلیم ناظر (۳-۵)

۹۹	.....	۱-۳-۵) گونه‌های دیگری از شبکه LVQ
۱۰۳	.....	۴-۵) شبکه عصبی کوهونن برای تهیه جدول هجاها
۱۰۳	.....	۱-۴-۵) مشخصات ساختار شبکه و دادگان مورد استفاده
۱۰۳	.....	۲-۴-۵) آموزش بدون نظارت
۱۰۵	.....	۳-۴-۵) آموزش بانظارت
۱۰۵	.....	۵-۵) جمع بندی

### فصل ششم - بررسی نتایج و پیشنهاد برای ادامه کار

۱۱۰	.....	۱-۶) جمع بندی کارهای انجام شده و نتایج آنها
۱۱۵	.....	۲-۶) پیشنهاد برای ادامه کار
۱۱۷	.....	مراجع

شکل (۱-۱) نمودار ساده‌ای از سیستم صوتی انسان	۵
شکل (۲-۱) طرحی از جهاز صوتی انسان	۵
شکل (۳-۱) مدل گسسته زمانی از فرآیند تولیدگفتار	۷
شکل (۴-۱) بلوک دیاگرام سیستم بازشناسی گفتار	۸
شکل (۵-۱) بلوک دیاگرام بانک فیلتر	۱۱
شکل (۶-۱) کدکننده LP	۱۴
شکل (۷-۱) اثر P در تخمین پوش طیف سیگنال	۱۵
شکل (۸-۱) مراحل محاسبه ضرایب کپسترال	۱۷
شکل (۹-۱) مقیاس MEL	۱۹
شکل (۱۰-۱) نمودار پهنای باند بحرانی	۲۰
شکل (۱۱-۱) استفاده از فیلترهای باند بحرانی در محاسبه ضرایب مل کپسترال	۲۱
شکل (۱-۲) نرونهاى بیولوژیکی	۳۱
شکل (۲-۲) یک نرون مصنوعی	۳۳
شکل (۳-۲) نمونه‌هایی از توابع آستانه در نرونهاى مصنوعی	۳۴
شکل (۴-۲) برچسب‌گذاری نرون مصنوعی	۳۵

شکل (۵-۲) خروجی سلول $\pi_k$ در پاسخ به ورودی $y'$ .....	۳۵
شکل (۶-۲) جابجایی تابع آستانه با استفاده از ورودی معین .....	۳۶
شکل (۷-۲) دودسته از شبکه‌های عصبی .....	۳۸
شکل (۸-۲) الف- پرسپترون ب- ابرصفحه تصمیم .....	۴۰
شکل (۹-۲) الف- دوکلاس جداپذیر خطی ب- دوکلاس جداپذیر غیرخطی .....	۴۱
شکل (۱۰-۲) الگوریتم یادگیری پرسپترون .....	۴۲
شکل (۱۱-۲) الف- لایه حاوی ۳ پرسپترون ب- مرز تعیین شده بوسیله پرسپترون .....	۴۴
شکل (۱۲-۲) مسئله XOR .....	۴۴
شکل (۱۳-۲) نواحی تصمیم MLP با استفاده از آستانه پله غیرخطی .....	۴۶
شکل (۱۴-۲) معماری LVQ .....	۴۶
شکل (۱۵-۲) الگوریتم یادگیری غیرناظر تک برنده برای آموزش LVQ .....	۴۸
شکل (۱-۳) بلوک دیاگرام مراحل پیاده‌سازی دسته‌بند بکمک شبکه MLP .....	۵۶
شکل (۲-۳) تعیین مرز قطعات سکوت ، واکنش و بیواک با استفاده از شبکه MLP .....	۵۸
شکل (۳-۳) تعیین مرز کلمات مجزا با استفاده از شبکه MLP .....	۵۹

- شکل (۳-۴) الگوریتم تشخیص ابتدا و انتهای کلمات با استفاده از آستانه‌های انرژی و نرخ عبور از صفر .. ۶۱
- شکل (۳-۵) تعیین مرز کلمات مجزا با استفاده از آستانه‌های انرژی و نرخ عبور از صفر..... ۶۱
- شکل (۳-۶) استفاده از مقادیر آستانه انرژی در مرز بندی کلمات ..... ۶۲
- شکل (۳-۷) تعیین مرز کلمات مجزا با استفاده از مقادیر آستانه انرژی ..... ۶۳
- شکل (۳-۸) روش پوش محدب..... ۶۷
- شکل (۳-۹) نمایش مرز بندی هجاهادر اعداد دو رقمی فارسی ..... ۶۸
- شکل (۴-۱) فلوچارت عملکرد الگوریتم BP ..... ۷۸
- شکل (۴-۲) مراحل استخراج بردارهای ویژگی ..... ۸۲
- شکل (۴-۳) نحوه قاب بندی سیگنال گفتار ..... ۸۴
- شکل (۴-۴) ساختار شبکه MLP ..... ۸۷
- شکل (۵-۱) الگوریتم یادگیری کوهونن ..... ۹۴
- شکل (۵-۲) نگاهت ویژگی در لایه خروجی یک بعدی ..... ۹۵
- شکل (۵-۳) شبکه دوبعدی کوهونن ..... ۹۷
- شکل (۵-۴) نحوه نامگذاری سلولهادر شبکه کوهونن ..... ۱۰۷

جدول (۱-۳) نتایج ارزیابی سه روش مرزبندی	۶۴
جدول (۲-۳) نتایج قطعه‌بندی اعداد دورقمی فارسی به روش پوش محدب	۶۹
جدول (۳-۳) دسته‌های استخراج شده از اعداد دورقمی به روش پوش محدب	۷۰
جدول (۱-۴) نتایج تعلیم شبکه MLP	۸۸
جدول (۲-۴) جدول هم‌جای‌های مشابه	۸۸
جدول (۱-۵) نتایج آزمایشات انجام شده بکمک شبکه کوهونن و رویه آموزش بانظارت	۱۰۶
جدول (۱-۶) نتایج اجرای رویه تفریق طیفی قبل از انجام مرزبندی	۱۱۱
جدول (۲-۶) نتایج آموزش شبکه MLP با استفاده از ضرایب مختلف	۱۱۴

فصل اول

مسئله بازشناسی گفتار