بسم الله الرحمن الرحيم

بسمه تعالی

مدیریت تحصیلات تکمیلی

تعهدنامه اثر اصالت

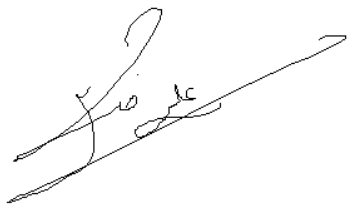
اینجانب **علیرضا اسمخانی** متعهد می‌شوم که مطالب مندرج در این پایان‌نامه که حاصل کار پژوهشی اینجانب است و دستاوردهای پژوهشی دیگران که در این پژوهش از آن‌ها استفاده شده است، مطابق مقررات ارجاع و در فهرست منابع و مآخذ ذکر گردیده است. این پایان‌نامه قبلاً برای احراز هیچ مدرک هم سطح یا بالاتر ارایه نشده است. در صورت اثبات تخلف (در هر زمان) مدرک تحصیلی صادر شده توسط دانشگاه از اعتبار ساقط خواهد شد.

کلیه حقوق مادی و معنوی این اثر متعلق به دانشگاه تربیت دبیر شهید رجایی می‌باشد.

نام ونام خانوادگی دانشجو

علیرضا اسمخانی

امضاء



دانشکده مهندسی برق و کامپیوتر

بازشناسی ارقام دست‌نویس فارسی با استفاده از اختلاط خبره‌ها

نگارش

علیرضا اسمخانی

استاد راهنما: دکتر رضا ابراهیم پور

استاد مشاور: دکتر محمدشمس اسفندآبادی

پایان نامه برای دریافت درجه کارشناسی ارشد

در رشته مهندسی برق– الکترونیک

تیر ۱۳۸۹

شماره: ۱۶۰۴۱۶
تاریخ: ۸۹/۴/۲۸
پیوست:

دانشگاه تربیت دبیر شهید رجایی

صورتجلسهٔ دفاع پایان‌نامه تحصیلی دوره کارشناسی ارشد

با تأییدات خداوند متعال و با استعانت از حضرت ولی عصر (عج) جلسه دفاع از پایان‌نامه کارشناسی‌ارشد آقای **علیرضا اسمخانی** رشتهٔ: **مهندسی‌برق- الکترونیک** با عنوان **بازشناسی ارقام دست‌نویس فارسی با استفاده از اختلاط خبره‌ها**، که در تاریخ ۸۹/۴/۱۵ با حضور هیأت محترم داوران در دانشگاه تربیت دبیر شهید رجائی برگزار گردید و نتیجه به شرح زیر اعلام گردید.

☐ قبول (بادرجه عالی امتیاز 19,25) ☐ دفاع مجدد ☐ مردود.

۱ ـ عالی (۲۰ – ۱۸)

۲ ـ بسیار خوب (۱۷/۹۹ – ۱۶)

۳ ـ خوب (۱۵/۹۹ – ۱۴)

۴ ـ قابل قبول (۱۳/۹۹ – ۱۲)

اعضاء	نام و نام خانوادگی	مرتبه علمی	امضاء
استاد راهنما	دکتر رضا ابراهیم پور	استادیار	
استاد مشاور	دکتر محمدشمس‌اسفندآبادی	استادیار	
استاد داور داخلی	دکتر شهریار شیروانی‌مقدم	استادیار	
استاد داور خارجی	دکتر علی‌احمدی	استادیار	
نماینده تحصیلات تکمیلی	دکتر جلال ولی‌الهی	استادیار	

دکتر علی اکبر مطیع بیرخدی
رئیس دانشکده مهندسی برق

تهران ،لویزان ،کدپستی: ۱۵۸۱۱ – ۱۶۷۸۸
صندوق پستی: ۱۶۳ – ۱۶۷۸۵
تلفن: ۹- ۲۲۹۷۰۰۶۰ فکس: ۲۲۹۷۰۰۳۳
Email:sru@sru.ac.ir
www.srttu.edu

تقدیم به :

پدر و مادر عزیزم

که همواره از راهنمایی‌های ارزشمندشان بهره‌مند بوده‌ام.

همسر عزیزم

که همواره یار و یاور صمیمی و مشوق من در امر ادامه تحصیل بوده است.

تشکر و قدردانی

بر خود لازم می‌دانم تا از زحمات تمامی عزیزانی که مرا در انجام این پایان‌نامه یاری نموده‌اند، خصوصاً اساتید گرانقدر جناب آقای دکتر رضا ابراهیم‌پور و جناب آقای دکتر محمدشمس اسفندآبادی که در طول مدت انجام پایان‌نامه علاوه بر راهنمایی‌های ارزشمندشان، نکات اخلاقی، علمی و عملی فراوان به من آموختند تشکر وقدردانی نمایم. همچنین از مرکز تحقیقات مخابرات ایران که در اجرای این پایان-نامه ما را مورد حمایت و پشتیبانی خود قرار داده‌اند، صمیمانه متشکرم.

چکیده

بازشناسی ارقام دست‌نویس یکی از مسائل مهم در بازشناسی الگو است. در زمینه‌ی تشخیص ارقام دست‌نویس فارسی در دو حوزه‌ی روش‌های استخراج ویژگی و استفاده از طبقه‌بندها تحقیقات زیادی صورت گرفته است. انتخاب روش استخراج ویژگی به عنوان مهم‌ترین عامل در بازشناسی الگو و به منظورکاهش ابعاد داده‌های ورودی مطرح است. برای بازشناسی الگو، در تحقیقات متعددی به منظور بهبود کارایی طبقه‌بندی، افزایش نرخ بازشناسی و افزایش قابلیت اعتماد در سیستم از ترکیب طبقه-بندها استفاده شده است. ترکیب طبقه‌بندها راهکاری برای بهبود کارایی در مسائل پیچیده طبقه‌بندی است. این پیچیدگی می‌تواند ناشی از محدود بودن تعداد الگوها، همپوشانی کلاس‌ها، بالا بودن بعد ویژگی‌ها و وجود نویز قابل ملاحظه در نمونه‌ها باشد.

تاکنون روش‌های ترکیب متعددی در بازشناسی ارقام دست‌نویس پیشنهاد شده است که از بین آنها می‌توان به روش‌های ترکیب میانگین، بیشینه، کمینه و حاصلضرب اشاره نمود. طبقه‌بندهای موجود با یک الگوریتم تکراری، سعی در بازشناسی الگو دارند.

در این رساله هدف این است که با استفاده از یک روش ترکیب با نام اختلاط خبره‌ها به نرخ بازشناسی بالاتری نسبت به روش‌های ترکیب ارائه شده در زمینه بازشناسی ارقام دست‌نویس فارسی دست یافت. در این مدل، مسأله پیچیده‌ی محاسباتی به تعدادی مسأله ساده‌تر تقسیم شده و در نهایت حل مسأله اصلی با ترکیب حل مسائل کوچکتر امکان‌پذیر می‌شود. به عبارت دیگر این مدل، با توزیع وظیفه یادگیری بین تعدادی طبقه‌بند و تقسیم فضای ورودی به مجموعه‌ای از زیر فضاها، قدرت تعمیم طبقه‌بندها را افزایش می‌دهد. سپس با ترکیب وزن‌دار خروجی طبقه‌بندها به خروجی نهایی بهینه دست می‌یابد. برای کاهش احتمال به دام افتادن طبقه‌بندها در کمینه‌های محلی نیز تغییراتی اعمال شده است. از جمله این تغییرات افزودن ضریب ممنتم در روابط اصلاح وزن‌های طبقه‌بندها و تغییر ساختار طبقه‌بندها از پرسپترون چند لایه به تابع پایه شعاعی است. با اعمال این تغییرات سعی در افزایش سرعت همگرایی و نرخ بازشناسی بوده است، و نتایج آزمایشات نیز حاکی از حصول این امر است.

کلمات کلیدی: اختلاط خبره‌ها، بازشناسی ارقام دست‌نویس فارسی، اجتماع شبکه‌های عصبی، ادغام طبقه‌بند، ترکیب چند طبقه‌بند، یادگیری معلم محور

فهرست مطالب

عنوان.....صفحه

پیشگفتار.....۱

فصل اول: بازشناسی الگو

۱-۱-بازشناسی الگو.....۴

۱-۲-روش‌های بازشناسی الگو.....۶

۱-۲-۱-روش‌های حسی-تجربی.....۶

۱-۲-۲-روش‌های ساختاری یا نحوی.....۷

۱-۲-۳-روش‌های ریاضی.....۷

۱-۲-۳-۱-روش‌های آماری.....۸

۱-۲-۳-۲-روش‌های غیرآماری.....۸

۱-۲-۴-مقایسه تجربی طبقه‌بندها.....۹

۱-۲-۴-۱-انتخاب مجموعه آموزش.....۹

۱-۲-۴-۲-انتخاب مجموعه آزمایشی.....۱۰

۱-۲-۴-۳-تصادفی بودن ذاتی الگوریتم یادگیری.....۱۰

۱-۲-۵-ادغام اطلاعات در بازشناسی الگو.....۱۰

۱-۲-۶-جایگاه سامانه‌های طبقه‌بندی مرکب در بازشناسی الگو.....۱۲

فصل دوم: سیر تحول بازشناسی نوری نویسه‌ها

۲-۱-مقدمه.....۱۵

۲-۲-سیر تحول عمومی *OCR*.....۱۵

۲-۲-۱-پیدایش *OCR*.....۱۵

۲-۲-۲-روند تحقیقات.....۱۶

۲-۳- سیر تحول *OCR* عربی.....۱۹

۲-۴-- سیر تحول *OCR* فارسی.....۲۰

۲-۵-انواع سیستم‌های بازشناسی نوری کاراکتر از لحاظ نوع الگوی ورودی.....۲۱

۲-۶-معرفی بخش‌های مختلف یک سامانه بازشناسی نوری کاراکتر.....۲۲

۲-۶-۱-پیش پردازش.....۲۲

۲-۶-۱-۱-کاهش نویز.....۲۲

۲-۶-۱-۲-نرمالیزه کردن داده‌ها.....۲۲

۲-۶-۱-۳-فشرده‌سازی.....۲۳

۲-۶-۲-استخراج ویژگی.....۲۳

۲-۶-۲-۱-گشتاور.....۲۴

۲-۶-۲-۲-هیستوگرام.....۲۷

۲-۶-۲-۳-ویژگی‌های جهتی.....۲۷

۲-۶-۲-۴-ویژگی‌های ناحیه‌ای..... ۳۰
۲-۶-۲-۵-تحلیل اجزاء اصلی (PCA)..... ۳۰
۲-۶-۳-پایگاه داده جهت ارزیابی روش‌های بازشناسی ارقام..... ۳۴
۲-۶-۴-طبقه‌بندی..... ۳۵
۲-۶-۵-نتیجه‌گیری..... ۳۶
فصل سوم: الگوریتم‌های یادگیری تجمعی و الگوریتم‌های ایجاد شبکه‌های عصبی گوناگون
۳-۱- مقدمه..... ۳۸
۳-۲- اهمیت ترکیب طبقه‌بندها ۳۹
۳-۳- الگوریتم‌های یادگیری تجمعی..... ۴۰
۳-۴- الگویتم‌های ایجاد شبکه‌های عصبی گوناگون..... ۴۱
۳-۵- نیازهای اساسی ترکیب..... ۴۲
۳-۶- استراتژی کاهش همبستگی..... ۴۲
۳-۶-۱- مکانیزم‌های مختلف یادگیری..... ۴۲
۳-۶-۲- طرق مختلف نمایش الگوها..... ۴۳
۳-۶-۳- تقسیم‌بندی مجموعه آموزشی..... ۴۳
۳-۶-۴- برچسب گذاری‌های متفاوت در یادگیری..... ۴۶
۳-۷- مدل‌های الگوی یادگیری تجمعی..... ۴۶
۳-۷-۱- مدل ایستا..... ۴۶
۳-۷-۱-۱- روش میانگین‌گیری دسته جمعی..... ۴۶
۳-۷-۲- مدل پویا..... ۴۷
۳-۷-۲-۱- مدل اختلاط خبره‌ها..... ۴۷
۳-۷-۲-۲- اختلاط خبره‌های سلسله مراتبی..... ۴۷
۳-۸ نتیجه‌گیری..... ۴۸
فصل چهارم : آزمایش‌های انجام شده و نتایج آن
۴-۱- مقدمه..... ۵۰
۴-۲-شبکه‌های عصبی..... ۵۰
۴-۳- قواعد متداول برای ترکیب خروجی طبقه‌بندها..... ۵۲
۴-۳-۱- ترکیب خروجی طبقه‌بندها با روش‌های هوشیار به کلاس..... ۵۴
۴-۳-۱-۱- روش‌های میانگین وزن‌دار مرتب شده..... ۵۴
۴-۳-۱-۲- روش‌های رأی‌گیری..... ۵۵
۴-۴- بازشناسی ارقام دست‌نویس فارسی با استفاده از روش اختلاط خبره‌ها..... ۵۶
۴-۴-۱- اختلاط خبره‌ها..... ۵۶
۴-۴-۲- پایگاه داده برای ارزیابی و پیاده سازی ساختار ارائه شده..... ۵۷
۴-۴-۳- نحوه بازنمایی تصاویر ارقام به خبره‌ها..... ۵۹
۴-۴-۴- آزمایش‌ها و نتایج..... ۶۰
۴-۴-۵- نتیجه‌گیری..... ۶۵

۴-۵- ساختار بهبود یافته اختلاط خبره‌ها با افزودن واحد گشتاور..... ۶۵
۴-۵-۱- واحد گشتاور..... ۶۶
۴-۵-۲- تغییرات اعمال شده بر روی ساختار اختلاط خبره‌های معمول..... ۶۷
۴-۵-۳- پایگاه داده برای ارزیابی و پیاده‌سازی ساختار ارائه شده..... ۶۸
۴-۵-۴- نحوه بازنمایی تصاویر ارقام به خبره‌ها..... ۶۸
۴-۵-۵- آزمایش‌ها و نتایج..... ۶۹
۴-۶- اختلاط خبره‌های تابع پایه شعاعی..... ۷۲
۴-۶-۱- پایگاه داده برای ارزیابی و پیاده سازی ساختارهای ارائه شده..... ۷۴
۴-۶-۲- نحوه بازنمایی تصاویر ارقام به خبره‌ها..... ۷۵
۴-۶-۳- نتایج تجربی..... ۷۵
۴-۶-۴- نتیجه‌گیری..... ۷۷
فصل پنجم: نتیجه‌گیری و پیشنهادها
۵-۱- مقدمه..... ۷۹
۵-۲-جمع‌بندی نهایی..... ۷۹
۵-۳-کارهای آینده..... ۸۱
فهرست مقالات مستخرج از پایان‌نامه..... ۸۲
فهرست مراجع..... ۸۳
واژه‌نامه‌نامه فارسی به انگلیسی..... ۸۷
واژه‌نامه‌نامه انگلیسی به فارسی..... ۸۹

فهرست جدول‌ها

شماره جدول **صفحه**

جدول۲-۱: گشتاور زرنیکی **۲۶**

جدول۴-۱: تغییرات نرخ بازشناسی ساختار اختلاط خبره ها نسبت به تعداد نورنهای لایه میانی طبقه بند های پایه و تعداد نورون لایه میانی شبکه میانجی **۶۰**

جدول۴-۲: تغییرات نرخ بازشناسی نسبت به تعداد نورنهای لایه میانی طبقه بند های پایه برای روشهای رای گیری، قواعد بیشینه، کمینه، میانگین و حاصلضرب **۶۱**

جدول۴-۳: معیار ارزیابی طبقه بند **۶۴**

جدول۴-۴: ماتریس سر در گمی برای بهترین ساختار اختلاط خبره ها **۶۴**

جدول۴-۵: توپولوژی بهترین ساختار برای اختلاط خبره ها با واحد گشتاور و بدون واحد گشتاور **۶۹**

جدول۴-۶: ماتریس سر در گمی برای بهترین توپولوژی اختلاط خبره ها با واحد گشتاور **۷۱**

جدول۴-۷: ماتریس سر در گمی برای بهترین توپولوژی اختلاط خبره ها بدون واحد گشتاور **۷۱**

جدول۴-۸: توپولوژی بهترین ساختار برای اختلاط خبره ها با ساختار تابع پایه شعاعی و اختلاط خبره ها با ساختار پرسپترون چند لایه **۷۵**

جدول۴-۹: ماتریس سر در گمی برای بهترین توپولوژی اختلاط خبره ها با ساختار تابع پایه شعاعی **۷۶**

جدول۴-۱۰: ماتریس سر در گمی برای بهترین توپولوژی اختلاط خبره ها با ساختار پرسپترون چند لایه ۷۶

فهرست شکل‌ها

شماره شکل	صفحه
شکل۱-۱: مراحل بازشناسی الگو	۴
شکل۱-۲: روش‌های غیرآماری، آماری و ساختاری برای بازشناسی حرف A	۹
شکل ۱-۳: یک تقسیم‌بندی ممکن برای ادغام اطلاعات	۱۲
شکل۲-۱: عملگرهای سوبل	۲۸
شکل۲-۲: عملگرهای کریش	۲۸
شکل۲-۳: جهت پایه برای گرادیان	۲۹
شکل۲-۴: تصویر کردن یک جهت گرادیان به دو جهت پایه	۲۹
شکل۲-۵: نحوه محاسبه ویژگی ناحیه‌ای	۳۰
شکل۲-۶: محور های جدید با توجه به بردارهای ویژه	۳۱
شکل۲-۷: داده‌ها	۳۳
شکل۲-۸: داده‌های اصلی به همراه داده‌های بازسازی شده با یک بردار ویژگی	۳۴
شکل۲-۹: نمونه‌هایی از مجموعه ارقام هدی	۳۵
شکل۲-۱۰: نمونه‌هایی از مجموعه ارقام ۸۶۰۰ نمونه‌ای	۳۵
شکل۳-۱: بلوک دیاگرام یک سیستم ترکیبی	۴۱
شکل۳-۲: بلوک دیاگرام سیستم اختلاط خبره‌ها	۴۵
شکل۳-۳: ساختار شماتیک میانگین‌گیری دسته جمعی	۴۶
شکل۳-۴: ساختار یک سیستم اختلاط خبره‌های سلسله مراتبی	۴۸
شکل۴-۱: مدل ترکیبی اختلاط خبره‌ها	۵۷
شکل۴-۲: نمونه‌هایی از تصاویر ارقام مجموعه داده هدی	۵۹
شکل ۴-۳: ویژگی مکان‌های مشخصه برای یک نقطه از تصویر عدد ۷	۶۰
شکل ۴-۴: ساختار اختلاط خبره‌ها با دو شبکه عصبی پرسپترون یک لایه و شبکه‌میانجی	۶۶
شکل ۴-۵: تسریع در همگرایی با افزودن واحد گشتاور	۶۶
شکل۴-۶: نمودار تغییرات نرخ بازشناسی بر حسب تعداد نورونهای لایه‌میانی	۷۰
شکل۴-۷: نمودار نرخ بازشناسی ارقام بر حسب تعداد تکرار های آموزش ساختار	۷۰
شکل۴-۸: مدل ترکیبی اختلاط خبره‌ها با ساختار تابع پایه شعاعی	۷۳

پیشگفتار

پیداش علوم و فنون جدید، جوامع بشری را با شکل‌های مختلفی از اطلاعات روبرو نموده است. سطح توسعه یک جامعه را می‌توان با مقدار اطلاعات و دانش تولید شده در آن ارزیابی کرد. تولید فزاینده اطلاعات به شکل‌های مختلف صورت می‌گیرد و با درجات متفاوتی از پیچیدگی‌ها همراه است و در نتیجه نیاز به سیستم‌های پردازش اطلاعات به‌صورت روز افزون افزایش یافته است. یکی از مسائل مهم در طراحی سیستم‌های مدرن اطلاعاتی، بازشناسی الگوهاست. یکی از زیر شاخه‌های بازشناسی الگو، سامانه بازشناسی نوری کاراکتر است که خود به دو زیر بخش، بازشناسی ارقام دست‌نوشته و بازشناسی حروف دست‌نوشته تقسیم می‌شود. یکی از دلایل تفکیک کاراکترها به ارقام و حروف از منظر بازشناسی الگو، سهولت در طبقه‌بندی[1] و تمرکز برای استخراج ویژگی‌های مناسب برای هر کدام است. اغلب مردم در سنین کودکی خواندن و نوشتن را می‌آموزند، انسان همزمان با فرآیند رشد خود مهارت‌های خواندن و نوشتن را به‌خوبی می‌آموزد و توانایی خواندن اکثر نوشته‌های پرینت شده یا دست‌نویس را دارد. همچنین اغلب مردم در خواندن متون پررنگ، کمرنگ، واژگون، متون با قلم‌های فانتزی و تبلیغات، متون با تزئینات و جلوه‌های خاص حتی گاهی با قلم‌خوردگی، مشکلی ندارند. در مقابل علی‌رغم پنج دهه تحقیقات متمرکز در حوزه بازشناسی نوری کاراکتر، مهارت خواندن توسط رایانه عقب‌تر از نوع بشر است. اغلب سامانه‌های بازشناسی نوری کاراکتر[2] توانایی خواندن کامل و دقیق متون ساده و یا متون دست‌نوشته را ندارند.

با توجه به کاربرد گسترده بازشناسی ارقام دستنویس، نظیر خواندن مبلغ چک بانکی، خواندن پلاک‌خودرو، خواندن کدپستی و سایر اطلاعات فرمها، در این زمینه تحقیقات زیادی صورت گرفته که مبتنی بر روش‌های استخراج ویژگی و استفاده از طبقه‌بند آموزش‌پذیر است. استخراج ویژگی به منظور کاهش ابعاد داده های ورودی مطرح است. برای بازشناسی حروف و ارقام، ویژگی‌های ناحیه‌ای، گشتاورهای هندسی، گشتاورهای زرنیکی، توصیفگرهای فوریه، گشتاورهای پایا، هیستوگرام‌نما و ویژگی مکانهای مشخصه پیشنهاد شده‌اند. انتخاب نوع ویژگی به کاربرد وابسته است. معمولا با یک ارزیابی تجربی از داده‌های مورد نظر، ویژگی مناسب‌تر مشخص می‌شود.

در شناسایی الگو به منظور بهبود کارایی طبقه‌بندی، افزایش نرخ بازشناسی و افزایش قابلیت اعتماد در سیستم، از ترکیب طبقه‌بندها استفاده می شود. ترکیب طبقه‌بندها راهکاری برای بهبود کارایی در مسائل پیچیده‌ی طبقه‌بندی است. پیچیدگی در طبقه بندی می‌تواند ناشی از محدود بودن تعداد الگو ها،

[1] - Classification
[2] - Optical Character Recognition(OCR)

همپوشانی کلاس‌ها، بالا بودن بعد ویژگی‌ها و وجود نویز قابل ملاحظه در نمونه‌ها باشد. نتایج تئوری و تجربی ارائه شده در مقالات متعدد نشان می‌دهند که ترکیب چند طبقه‌بند، زمانی مفید واقع می شود که طبقه‌بندهای مختلف نرخ خطای نسبتا کمی داشته باشند و خطاهای آنها با یکدیگر متفاوت باشد و یا به عبارت دیگر تصمیم‌های دو طبقه بند، غیرهمبسته[1] باشند. دو روش کلی در ترکیب طبقه‌بندها وجود دارد: ترکیب[2]، انتخاب[3]. در ترکیب، فرض شده است که هر طبقه‌بند روی تمام فضای ویژگی آموزش داده شده و نظر همه‌ی اعضا در تشریح و تشخیص نتیجه‌ی نهایی استفاده می‌شود، روش‌هایی مانند رای‌گیری، قواعد بیشینه، کمینه، میانگین و حاصلضرب از این قبیل هستند. در روش انتخاب فرض بر این است که هر طبقه‌بند، قسمتی از فضای ویژگی را آموزش می بیند و نتیجه نهایی از تجمع تصمیم یک یا چند طبقه بند خاص حاصل می شود، اختلاط خبره‌ها که اولین بار توسط ژاکوب، مطرح شد از این قسم است. لذا در این پایان‌نامه جهت حل این مشکل و بالا بردن نرخ بازشناسی ارقام دست‌نوشته فارسی تلاشهایی صورت گرفته که روند کلی آن به صورت زیر است:

در فصل اول مباحثی در زمینه بازشناسی الگو و مراحل آن مطرح شده است. در فصل دوم مقدمه-ای بر سامانه بازشناسی نوری کاراکتر و روند بازشناسی ارقام دست‌نویس به عنوان یک زیر مجموعه‌ای از سامانه بازشناسی الگو مطرح گردیده است. در فصل سوم الگوریتم‌های یادگیری تجمعی و الگوریتم-های ایجاد شبکه‌های عصبی گوناگون توضیح داده شده است. در فصل چهارم سه روش ترکیب برای بالا بردن نرخ بازشناسی ارقام دستنویس فارسی مطرح شده و فصل پنجم به نتیجه‌گیری و پیشنهادها اختصاص دارد.

[1] - Uncorrelated
[2] - Fusion
[3] - Selection

فصل اول

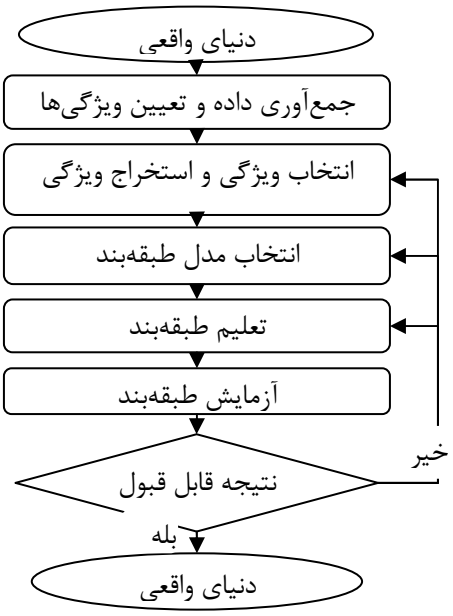
بازشناسی الگو

۱-۱- بازشناسی الگو

یادگیری ماشینی یک شاخه مهم از گرایش هوش‌مصنوعی است که هدف آن تعلیم یک ماشین است به‌طوری‌که بتواند تجربیات و نمونه‌های موجود را یاد بگیرد. حاصل این یادگیری ایجاد یک مدل طبقه‌بندی است که بر اساس آن ماشین می‌تواند نمونه‌هایی را که در آینده می‌بیند و مشابه نمونه‌های موجود هستند در کلاس مناسب خود قرار دهد. امروزه روش‌های بازشناسی الگو، به عنوان یک شاخه از یادگیری ماشینی، کاربردهای فراوانی در زمینه‌های مختلف علمی و صنعتی پیدا کرده‌اند.

در حالت کلی هر توصیف کیفی یا کمی از یک موضوع را می توان یک الگو نامید. الگو می تواند خود یک شیء باشد و یا ساختار اجزاء یک شیء و روابط بین آنها را توصیف کند. هدف از بازشناسی الگو می‌تواند شناسایی منطقه‌ای خاص در تصویر، تشخیص آوا یا کلمه در صدا یا تشخیص هویت باشد. در حال حاضر از تکنیک‌های بازشناسی الگو در بسیاری از کاربردهای صنعتی، پردازش مستندات، تشخیص هویت و بسیاری از زمینه‌های دیگر استفاده می‌شود.

در فرآیند بازشناسی الگو، الگوهای ورودی در کلاس‌ها و دسته‌های از پیش تعیین شده‌ای طبقه‌بندی می‌شوند. روند بازشناسی الگو، در شکل(۱-۱) نشان داده شده است.



شکل۱-۱: مراحل بازشناسی الگو[۱]

چنانچه شکل(۱-۱) نشان می‌دهد، اولین گام در بازشناسی الگو، جمع‌آوری تعداد مناسبی نمونه از الگوهای مورد نظر است. این بخش زمان زیادی از فرآیند طراحی سامانه بازشناسی الگو را به خود اختصاص می‌دهد و گاهی اوقات با مشکلاتی همراه است. در این پایان‌نامه از مجموعه داده‌های استاندارد موجود بهره گرفته شده است. پس از جمع آوری نمونه‌های لازم، باید اقدام به انتخاب نوع ویژگی کرد. انتخاب نوع ویژگی‌ها نیازمند دانش اولیه در مورد الگوهاست. انتخاب باید به گونه‌ای باشد که ویژگی‌ها، وجه اشترک الگوهای متعلق به یک کلاس و وجه تمایز الگوهای کلاس‌های مختلف باشند. توانمندی ویژگی برای جداسازی نمونه‌های کلاس‌های مختلف، معیار انتخاب آن است.

گام بعدی، استخراج ویژگی است. استخراج ویژگی به معنی به‌دست آوردن خصوصیات الگوها به‌نحوی است که این خصوصیات وجه اشترک الگوهای متعلق به یک کلاس و وجه تمایز الگوهای کلاس‌های مختلف باشند. استخراج ویژگی یک نگاشت از فضای الگوها به فضای ویژگی‌ها است که نتیجه آن پیدایش یک بردار ویژگی X است که خصوصیات الگو را به صورت معنی‌دارتر و مناسب‌تری برای مرحله طبقه‌بندی آماده می‌کند. ویژگی‌های استخراج شده از الگوها ممکن است کیفی (توصیفی) یا کمی[1](مقداری) باشند. ویژگی‌های کیفی تعداد محدودی حالت دارند و برای معرفی کیفی الگو استفاده می‌شوند. به عنوان مثال آفتابی، برفی، بارانی ویژگی‌های کیفی برای توصیف هوا هستند. ویژگی‌های کمی، معمولا یک مقدار پیوسته(تعداد حالات نامحدود) دارند و الگوها را به صورت مقادیر عددی (کمی) توصیف می‌کنند. به عنوان مثال، مقدار گاز دی‌اکسیدکربن موجود در یک لیتر هوا، یک ویژگی کمی برای معرفی وضعیت هواست. تبدیل ویژگی‌های کیفی به کمی نیازمند یک روش یا متدولوژی است. به عنوان مثال برگزاری یک امتحان روشی است برای کمی کردن میزان یادگیری. در مورد ویژگی‌های کمی، مولفه‌های بردار ویژگی با اندازه‌گیری به‌دست می‌آیند. هر کمیت اندازه‌گیری شده، یک ویژگی از شیء مورد نظر را بیان می‌کند. این کار عملا یک فرآیند کدگذاری است چرا که از نظر هندسی، هر شیء را می‌توان به عنوان یک نقطه در فضای اقلیدسی در نظر گرفت. انتخاب روش استخراج ویژگی یک عامل مهم در کارایی سامانه بازشناسی‌الگو است. در کاربردهای عملی بازشناسی الگو، معمولا از داده‌های خام ویژگی‌های متفاوتی استخراج می‌شوند و مجموعه آنها برای بازشناسی الگو استفاده می‌شوند.

ویژگی‌هایی که از الگوها استخراج می‌شوند دارای اهمیت یکسان نیستند و برخی از آنها نسبت به بقیه مهمترند. در مرحله انتخاب ویژگی، با انتخاب ویژگی‌هایی که بیشترین اهمیت را در جداسازی کلاس‌های الگو دارند، هم بعد بردار ویژگی کاهش داده می‌شود و هم کیفیت توصیف الگو توسط ویژگی‌ها بهتر می‌شود.

انتخاب نوع مدل طبقه‌بندی، مرحله بعدی در بازشناسی الگو است. طبقه‌بند به عنوان هسته اصلی یک سامانه بازشناسی الگو مطرح است. مدل طبقه‌بندی، هر الگوی ناشناخته را بر اساس ویژگی‌های آن، به یکی از کلاس‌های شناخته شده نسبت می دهد. به بیان ریاضی، به هر نگاشتی از فضای nبعدی ویژگی‌ها، به فضای K بعدی برچسب‌های کلاسی، یک طبقه‌بندی گفته می شود.

[1] - Quantative Feature

$$D: R^n \rightarrow \Omega \qquad (۱-۱)$$

که در آن $\Omega = \{\omega_1, \omega_2, ..., \omega_k\}$ مجموعه برچسب‌های کلاسی الگوهاست. یک طبقه‌بند، میزان تعلق بردار ویژگی $X \in R^n$ به هر یک از کلاس‌های الگو را به صورت یک عدد حقیقی بیان می‌کند. مدل‌های طبقه‌بندی متعددی برای بازشناسی الگو پیشنهاد شده‌اند که می توان به طبقه‌بند بیز، شبکه‌های عصبی، ماشین بردار پشتیبان[1] و طبقه بند کمترین فاصله اشاره کرد. پس از انتخاب مدل طبقه‌بندی، باید پارامترهای آن را مشخص کرد. روش‌های مختلفی برای تعیین این پارامترها ارائه شده‌اند. گروهی از این روش‌ها بر اساس خواص آماری، مانند بردار میانگین و ماتریس کوواریانس نمونه‌های آموزشی هر کلاس در طبقه‌بند بیز، پارامترهای طبقه‌بندی را تعیین می‌کنند. گروه دیگری از روش‌ها وجود دارند که با استفاده از الگوریتم‌های یادگیری تکراری، مانند آموزش یک طبقه‌بند شبکه عصبی بر روی نمونه‌های یادگیری، پارامترهای طبقه‌بندی را تعیین می‌کنند. در روش‌های گروه دوم، پارامترهای طبقه‌بندی در طول فرآیند یادگیری، تعیین می‌شوند. پس از آنکه مدل طبقه‌بندی کامل شد، با استفاده از نمونه‌های آزمایشی، سامانه بازشناسی الگو اعتبارسنجی می‌شود. اگر نرخ بازشناسی مورد نیاز بدست نیامد یک یا چند تا از کارهای زیر را انجام داد.

۱. ویژگی‌های استخراج شده از الگو یا ویژگی‌های انتخاب شده از بین ویژگی‌های قبلی را بهبود داد.
۲. مدل طبقه‌بندی را عوض کرد یا پارامترهای آن را تغییر داد.
۳. یادگیری(تعلیم) طبقه‌بند را کامل‌تر کرد.

کارهای فوق آنقدر تکرار می‌شوند تا تضمین کافی جهت یک تعمیم معتبر حاصل شود.

۱-۲- روش‌های بازشناسی الگو

تکنیک‌های بازشناسی الگو، یک شیء یا یک رخداد فیزیکی را به یک یا چند کلاس الگوی از قبل معلوم نسبت می‌دهند. بنابراین یک سامانه بازشناسی الگو می‌تواند به عنوان یک قاعده تصمیم‌گیری خودکار در نظر گرفته شود که با اندازه‌گیری‌های انجام شده روی الگو آن‌را به یک کلاس نسبت می‌دهد. روش‌های انتساب الگو به یکی از کلاس‌ها، به سه گروه عمده تقسیم می‌شوند. روش‌های حسی-تجربی[2]، روش‌های نحوی یا ساختاری[3] و روش‌های ریاضی[4].

۱-۲-۱- روش‌های حسی-تجربی

این روش‌ها بر مبنای تجربیات طراح استوار هستند و یک سامانه مبتنی بر این روش‌ها شامل مجموعه‌ای از قواعد تجربی است که با توجه به خصوصیات الگوها حاصل شده‌اند. به عنوان مثال سامانه بازشناسی دست‌نوشته‌های فارسی(مبتنی بر روش‌های حسی-تجربی) شامل قواعدی حسّی

[1] - Support Vector Machine(SVM)
[2] - Heuristic Methods
[3] - Syntactic(Structural) Methods
[4] - Mathematical Methods

است که مخصوص الگوهای دست‌نویس فارسی است و ممکن است برای بازشناسی دست‌نوشته‌های لاتین مناسب نباشد. این قواعد به عنوان مثال بر مبنای تعداد و توالی حرکات قلم، تعداد و موقعیت نقاط و علائم و تعداد زیرکلمات هستند که به صورت تجربی به‌دست می آیند. به عنوان مثال، در زبان فارسی، اگر توالی نقاط یک کلمه به صورت سه نقطه در بالا، یک نقطه در بالا و یک نقطه در پایین باشد این کلمه "شنبه" است. کارایی یک سامانه بازشناسی الگو مبتنی بر روش‌های حسی-تجربی بستگی به گستره تجربیات طراح دارد.

۱-۲-۲- روش های ساختاری یا نحوی

در خیلی از مسائل پیچیده، تعداد ویژگی‌های مورد نیاز ممکن است بسیار زیاد باشند. در چنین شرایطی می توان یک الگو را مجموعه‌ای از چند الگوی ساده‌تر در نظر گرفت. هر الگوی ساده می‌تواند از چندین زیرالگوی ساده‌تر تشکیل شده باشد. ساده‌ترین زیرالگوها را عناصرپایه[1] می‌گویند. بیان الگو بر حسب عناصر پایه، اساس شناسایی ساختاری الگو را تشکیل می‌دهد. ویژگی‌های ساختاری می‌توانند به صورت صریح با روابط ریاضی یا به صورت ضمنی با یک گرامر زبان بیان شوند. الگوها جملاتی از یک زبان هستند که عناصر پایه، الفبای آن هستند. این جملات با توجه به گرامر زبان شاخته می‌شوند. با استفاده از الفبای زبان و گرامر آن می توان مجموعه‌ی پیچیده‌ای از الگوها را با مجموعه‌ی کوچکی از عناصر پایه و قواعد گرامری زبان بیان کرد.گرامرهای هر کلاس با توجه به نمونه‌های آموزشی همان کلاس ساخته می‌شوند. روش‌های ساختاری، زمانی مفید هستند که الگوها، ساختار های قابل تعریفی به‌وسیله مجموعه‌ای از قواعد داشته باشند. به‌دلیل وجود الگوهای نویزی در شناسایی عناصر پایه و قواعد گرامری، پیاده سازی ساختاری دارای مشکلات زیادی است.

۱-۲-۳- روش‌های ریاضی

در روش‌های ریاضی بازشناسی الگو، هر الگوی $X = \{x_1, x_2, ..., x_k\}$ به صورت نقطه‌ای در فضای n بعدی مدل شده و با توجه به مقدار ویژگی‌هایش به یکی از کلاس‌ها نسبت داده می‌شود. ویژگی‌ها به گونه‌ای انتخاب می‌شوند که بردارهای نمونه‌های مربوط به کلاس‌های مختلف، ناحیه‌های متفاوتی از فضا را اشغال کنند. روش‌های ریاضی، بر مبنای توابع تصمیم[2] (جداساز) برای کلاس‌های الگو بنا شده‌اند. برای یک مساله طبقه‌بندی K کلاسه نیاز به پیدا کردن K تابع تصمیم $d_1(x), d_2(x), ..., d_k(x)$ با خاصیت زیر است.

$$if \quad x \in \omega_i \; then \;\; d_i(x) > d_j(x) \quad for \quad j = 1,2,...,k \;\; , j \neq i \qquad (۲-۱)$$

به عبارت دیگر الگوی ناشناخته x به کلاس ω_i متعلق است اگر با جایگزینی x در توابع تصمیم فوق، مقدار عددی $d_i(x)$ بزرگتر از بقیه باشد. در این‌صورت، مرز تصمیم جدا کننده کلاس ω_i از ω_j به صورت زیر خواهد بود:

---

[1] - Primitive Elements
[2] - Decision(Discriminant) functions

$$d_{ij}(x) = d_i(x) - d_j(x) = 0 \qquad (۳-۱)$$

برای الگوهای متعلق به کلاس ω_i مقدار $d_{ij}(x)$ مثبت و برای الگوهای متعلق به کلاس ω_j منفی است. برای تعیین تابع تصمیم هر کلاس معمولا از نمونه‌های آموزشی آن کلاس استفاده می‌شود. روش های ریاضی شامل دو گروه آماری[1] و غیر آماری[2] هستند. از این دو گروه با عنوان پارامتریک(غیر آموزش‌پذیر[3]) و غیر پارامتریک (آموزش‌پذیر) نیز یاد می‌شود. در روش‌های پارامتریک، پارامترهای تابع تصمیم مستقیما از خصوصیات آماری نمونه‌های آموزشی تعیین می‌شوند و در روش‌های غیر پارامتریک، پارامترهای تابع تصمیم معمولا در یک فرآیند یادگیری تعیین می‌شوند. بسیاری از تکنیک‌های ارائه شده در روش‌های پارامتریک مبتنی بر فرض نرمال بودن توابع توزیع الگوها هستند. در صورتی که تعداد نمونه‌های موجود در هر کلاس کم باشد، معمولا روش‌های غیر پارامتریک بهتر عمل می‌کنند.

۱-۲-۳-۱-روش‌های آماری

روش‌های آماری بر مبنای قواعد طبقه‌بندی ریاضی با بهره‌گیری از خصوصیات آماری الگوها استوار هستند. در این روش‌ها، مرزهای تصمیم‌گیری به‌وسیله توزیع‌های آماری نمونه‌ها تعیین می‌شوند. طبقه‌بندهای الگو برای روش‌های آماری عموما بر مبنای قاعده تصمیم بیز[4] کار می‌کنند. با به‌کارگیری این قاعده، اگر تابع چگالی احتمال هر یک از کلاس‌های الگو و همچنین احتمال پیشین هر کلاس مشخص باشد، سطح تصمیم بهینه پیدا می شود.

۱-۲-۳-۲-روش‌های غیرآماری

روش‌های غیرآماری بر مبنای یک چارچوب ریاضی بدون استفاده صریح از خصوصیات آماری الگوها استوار هستند. در این روش‌ها، ابتدا فرم کلی مرز جدا کننده (مانند توابع خطی یا درجه دوم) مشخص شده و سپس پارامترهای آن با استفاده از نمونه‌های آموزشی تعیین می‌شود. الگوریتم‌های یادگیری تکراری، نمونه‌ای از روش‌های غیر آماری هستند که در آنها توابع تصمیم از الگوهای آموزشی و در یک فرآیند تکرار تولید می‌شوند. یعنی پس از انتخاب نوع تابع تصمیم، ضرایب مربوط به آن در یک فرآیند تکراری، با استفاده از نمونه‌های آموزشی، تعیین می‌شوند. شبکه‌های عصبی پرسپترون چند لایه از متداولترین روش های غیرآماری برای بازشناسی الگو هستند. برخی مراجع روش‌های غیرآماری را با عنوان روش های عصبی نیز معرفی کرده‌اند. در شکل(۱-۲) روش‌های غیرآماری، آماری و ساختاری برای بازشناسی حرف A نشان داده شده‌اند. روش آماری از تابع چگالی‌احتمال ویژگی‌ها استفاده می‌کند و روش ساختاری با تبدیل الگو به سه عنصر پایه به بازشناسی آن می‌پردازد.

[1] - Statistical
[2] - Deterministic
[3] - Nontrainable
[4] - Bayes Decision