بسم الله الرحمن الرحيم

دانشگاه بیرجند
دانشکده مهندسی برق وکامپیوتر

عنوان:

بهبود نرخ بازشناسی ارقام دستنویس فارسی با استفاده از روش‌های ادغام در سطح ویژگی

نگارش:
حامد فضل اللهی آقاملک

استادراهنما: دکتر سید محمد رضوی

استاد مشاور: دکتر ناصر مهرشاد

پایان نامه برای دریافت درجه کارشناسی ارشد
در رشته برق_ الکترونیک

حامی:
موسسه تحقیقات ارتباطات و فناوری اطلاعات

شهریور ماه ۱۳۹۱

تعهدنامه اصالت اثر

اینجانب **حامد فضل اللهی آقاملک** تایید می‌نمایم که مطالب مندرج در این پایان‌نامه حاصل کار پژوهشی این جانب بوده و به دستاوردهای پژوهشی دیگران که در این نوشته از آن‌ها استفاده شده است، مطابق مقررات، ارجاع گردیده است. این پایان‌نامه قبلا برای احراز هیچ مدرک هم‌سطح یا بالاتر ارائه نشده است.

همه حقوق مادی و معنوی این اثر متعلق به دانشکده مهندسی برق وکامپیوتر دانشگاه بیرجند می‌باشد.

نام و نام خانوادگی: **حامد فضل اللهی آقاملک**

امضاء

تقدیم به

پدر و مادر عزیزم

که در پیمودن این مسیر

مرا یاری کردند.

تشکر و قدردانی

سپاس خداوند عزوجل که توفیق برداشتن قدمی کوچک در راه فراگیری ذره‌ای از علوم این عالم پهناور را به این حقیر عطا فرمود. امید است شایسته نور معرفت حضرتش نیز بگرداند.

از تمامی دوستانی که بنده را در به انجام رسیدن این پژوهش یاری و همراهی نموده‌اند کمال تشکر را دارم. بیش از همه از اساتید محترم و گرامی جناب آقای دکتر رضوی و جناب آقای دکتر مهرشاد به دلیل راهنمایی‌های بی دریغشان سپاسگزارم. همچنین جا دارد از پدر و مادر عزیزم به پاس تمامی فداکاری‌ها و مهربانی‌هایشان تشکر کنم. این پایان نامه به حمایت موسسه تحقیقات ارتباطات و فناوری اطلاعات می باشد و در اینجا جا دارد که کمال تشکر و سپاسگزاری را از این مرکز داشته باشم.

حامد فضل اللهی آقاملک

شهریور ماه ۱۳۹۱

چکیده

هدف یک سیستم بازشناسی الگو، قرار دادن الگوها با کمترین خطا، در کلاس مربوط به خودشان است. بازشناسی ارقام دستنویس فارسی یکی از مسائل مهم در حوزه بازشناسی الگو می باشد. تحقیقات در این زمینه چندین دهه است که آغاز شده است و هنوز هم در حال تحول می‌باشد. در سیستم های اولیه بازشناسی الگو از یک ویژگی و یک طبقه بند استفاده می‌شد. این سیستمها برای بازشناسی الگوهای پیچیده و کاربردهای زمان حقیقی مشکل داشتند. استفاده از ادغام اطلاعات به منظور افزایش کارایی سیستمهای بازشناسی الگو، یک مسئله متداول در کاربردهای مختلف می‌باشد. در این ایده با استفاده از ترکیب و ادغام چند بردار ویژگی و تشکیل یک بردار ویژگی جدید، توصیف بیان الگو، برای طبقه بند آسان تر خواهد شد. در این پژوهش، چند هدف دنبال شده است. اول اینکه هزینه و زمان بازشناسی، نسبت به ترکیب طبقه بندها بهبود داده شود. هدف دیگر اینکه نرخ بازشناسی بهتری را نسبت به ترکیب طبقه بندها بدست آورده شود. در این تحقیق، از سه طبقه بند، طبقه بند بیز، طبقه بند ۳ همسایه نزدیکتر و طبقه بند کمترین فاصله استفاده شده است. از یازده روش استخراج ویژگی متفاوت برای بازشناسی ارقام دستنویس فارسی استفاده شده است که عبارتند از، ویژگی هیستوگرام افقی، هیستوگرام عمودی، فاصله(Distance)، پایه گرادیان(Gradient base)، شمارش تلاقی(Cross count)، تبدیل کسینوسی گسسته(DCT)، تبدیل موجک(DWT)، تبدیل گابور، مکان مشخصه، زونینگ یا ناحیه بندی، و یک روش پیشنهادی هیستوگرام پنجره ای. برای پیدا کردن وزن‌های بهینه برای ادغام ویژگی از دو الگوریتم بهینه سازی، الگوریتم ژنتیک(GA) و الگوریتم جمیعت ذرات(PSO) استفاده شده است. تابع برازندگی در این الگوریتم ها تعداد خطاهای طبقه بندی کننده می باشد و هدف کمینه کردن خطای طبقه بندی کننده می باشد. از دو روش کاهش ابعاد ویژگی، روش آنالیز مولفه اصلی(PCA) و تحلیل تفکیک کننده خطی (LDA) برای کاهش بعد بردار ویژگی استفاده شده است. دیتابیسی که در این پایان نامه از آن استفاده شده، دیتابیس هدی می باشد. این دیتابیس شامل ۱۰۲۳۵۲نمونه می باشد که از ۶۰،۰۰۰ نمونه برای آموزش طبقه بندی کننده و۲۰،۰۰۰نمونه جهت آزمایش استفاده شده است. از خوشه بندی (k-means) برای کم کردن تعداد نمونه های آموزشی استفاده شده است این کار کمک به سریعتر شدن زمان پردازش برای بازشناسی می شود.

کلمات کلیدی: ادغام ویژگی، الگوریتم بهینه سازی ژنتیک، الگوریتم بهینه سازی گروه ذرات، بازشناسی ارقام دستنویس فارسی، ترکیب طبقه بندها

فهرست مطالب

فهرست شکل ها .....ذ

فهرست جداول .....ز

فصل ۱- مقدمه ۱

۱-۱- بازشناسی الگو..... ۱

۱-۲- روشهای بازشناسی الگو..... ۴

۱-۲-۱- روش های حسی – تجربی.....۵

۱-۲-۲- روش های ساختاری یا نحوی ۵

۱-۲-۳- روش های ریاضی ۵

۱-۳- ارزیابی یک طبقه بند ۸

۱-۳-۱- دقت طبقه بند برای کلاس **ωi** ۹

۱-۳-۲- فراخوانی طبقه بند روی کلاس **ωi**.....۹

۱-۳-۳- معیار F ۹

۱-۳-۴- نرخ بازشناسی کلی یا دقت طبقه بند یا قابلیت اعتماد بازشناسی.....۹

۱-۴- هدف پایان نامه و ساختار آن ۱۰

۱-۵- جمع بندی ۱۰

فصل ۲- مروری بر کارهای گذشته و روش های استخراج ویژگی ۱۱

۲-۱- باز شناسی شناسه های(حروف و ارقام)دستنویس ۱۱

۲-۱-۱- پیش پردازش ۱۲

۲-۱-۲- روش های استخراج ویژگی ۱۴

۲-۱-۳- فرایند آموزش ۲۸

۲-۱-۴- فرایند بازشناسی ۲۸

۲-۲- مروری بر تحقیقات انجام شده در زمینه بازشناسی ارقام دستنویس ۲۸

۲-۳- ادغام اطلاعات در بازشناسی الگو.....۳۳

۲-۴- جمع بندی ۳۵

فصل ۳- ادغام ویژگی و الگوریتم پیشنهادی ۳۶

۳-۱- مقدمه ۳۶

۳-۲- طبقه بندی ۳۶

۳-۲-۱- طبقه بند بیز ۳۷

۳-۲-۲- طبقه بند K همسایه نزدیکتر.....۳۸

۳-۳- روش های کاهش افزونگی ۴۰

۳-۳-۱- آنالیز مولفه اصلی ۴۱

۳-۳-۲- روش تحلیل تفکیک کننده خطی..... ۴۹
۳-۴- الگوریتم های بهینه سازی ۴۹
۳-۴-۱- الگوریتم ژنتیک..... ۴۹
۳-۴-۲- الگوریتم بهینه سازی جمعیت ذرات..... ۵۱
۳-۵- روش خوشه بندی K-Means ۵۲
۳-۶- نظریه ترکیب اطلاعات ۵۳
۳-۶-۱- سطوح ترکیب اطلاعات..... ۵۳
۳-۶-۲- مدل ترکیب ویژگی و تصمیم (FDF) ۵۴
۳-۷- روشهای ادغام ویژگی ۵۵
۳-۷-۱- انتخاب ویژگی ۵۷
۳-۷-۲- استخراج ویژگی ۵۷
۳-۷-۳- ترکیب ویژگی ۵۷
۳-۸- عملگرهای ترکیب اطلاعات ۵۸
۳-۹- جمع بندی ۵۹
فصل ۴- نتایج و آزمایشات ۶۰
۴-۱- جمع بندی ۷۴
فصل ۵- نتیجه گیری و پیشنهادات ۷۵
مراجع و منابع ۷۷
واژه نامه انگلیسی به فارسی ۸۱

فهرست شکل ها

شکل ۱- ۱: مراحل سیستم بازشناسی الگو ۲
شکل ۱- ۲: روش های غیرآماری، آماری و ساختاری برای بازشناسی حرف A ۷
شکل ۲- ۱: ماسک مورفولوژی برای هموارسازی کانتور ۱۳
شکل ۲- ۲: عملیات نازک سازی، (الف) تصویر رقم قبل از نازک سازی، (ب) تصویررقم نازک شده ۱۴
شکل ۲- ۳: طریقه استخراج ویژگی از تصویر به روش DWT ۱۸
شکل ۲- ۴: اضافه کردن فضای سفید جهت قرار گرفتن رقم در داخل دایره واحد ۲۰
شکل ۲- ۵: ویژگی های جهتی کانتور و ویژگی های نقاط خمش ۲۲
شکل ۲- ۶: استخراج ویژگی زونینگ ۲۲
شکل ۲- ۷: نحوه محاسبه ویژگی مکان مشخصه برای یک نقطه از تصویر ۲۳
شکل ۲- ۸: نحوه محاسبه ویژگی هیستوگرام پنجره ای از تصویر ۲۴
شکل ۲- ۹: نحوه محاسبه ویژگی شمارش تلاقی از تصویر ۲۵
شکل ۲- ۱۰: نحوه محاسبه ویژگی فاصله از تصویر ۲۵
شکل ۲- ۱۱: برخی از ویژگی های توپولوژی در حروف لاتین ۲۶
شکل ۲- ۱۲: ماسک های استفاده شده برای ویژگی پایه گرادیان در تصاویر ارقام ۲۷
شکل ۲- ۱۳: مراحل استخراج ویژگی پایه گرادیان ۲۸
شکل ۲- ۱۴: مقایسه ارقام عربی، فارسی و هندی ۳۱
شکل ۲- ۱۵: یک تقسیم بندی ممکن برای ادغام اطلاعات ۳۴
شکل۳- ۱: تعلق الگوی ورودی x به کلاس سیاه با طبقه بند ۵ همسایه نزدیکتر ۳۸
شکل۳- ۲: انتخاب محورهای جدید برای داده های دو بعدی ۴۲
شکل۳- ۳: داده های دوبعدی اولیه که قرار است PCA برروی آنها اعمال شود ۴۵
شکل۳- ۴: داده های نرمال سازی شده (با کم شدن میانگین) به همراه بردارهای ویژگی ماتریس کواریانس ..۴۶
شکل۳- ۵: داده های بدست آمده از تبدیل PCA با انتخاب مهمترین بردار ویژگی ۴۷
شکل۳- ۶: داده های بازیابی شده از تبدیل PCA با انتخاب مهمترین بردار ویژگی ۴۸
شکل۳- ۷: فلوچارت الگوریتم ژنتیک ۵۰
شکل۳- ۸: ترکیب اطلاعات در سطح داده ۵۳
شکل۳- ۹: ترکیب اطلاعات در سطح ویژگی ۵۴
شکل۳- ۱۰: ترکیب اطلاعات در سطح تصمیم ۵۴
شکل۳- ۱۱: مدل کلی Feature Decision Fusion (FDF) ۵۵
شکل۳- ۱۲: مدل سلسله مراتبی روش های متفاوت از ادغام ویژگی ۵۶
شکل۴- ۱: ساختار ادغام تک ویژگی زونینگ با وزن دهی فردی و برداری ۶۳

شکل۴- ۲: (الف)مقدار خطا بازشناسی و وزن دهی برداری بهینه برای ویژگی زونینگ با الگوریتم ژنتیک.
(ب) مقدار خطا بازشناسی و وزن دهی فردی بهینه برای ویژگی زونینگ با الگوریتم ژنتیک ۶۳
شکل۴- ۳: ساختار ترکیب ویژگی با وزن دهی برداری و طبقه بند بیز ۶۴
شکل۴- ۴: مقدار خطا بازشناسی و وزن دهی برداری بهینه برای ویژگی گابور با الگوریتم ژنتیک(الف) و PSO(ب). ۶۴
شکل۴- ۵: ساختار ترکیب ویژگی با وزن دهی فردی و برداری و طبقه بند کمترین فاصله ۶۵
شکل۴- ۶: (الف) مقدار خطا بازشناسی و وزن دهی فردی بهینه برای ویژگی مکان مشخصه با الگوریتم ژنتیک و (V=1). (ب) مقدار خطا بازشناسی و وزن دهی فردی بهینه برای ویژگی گرادیان پایه با الگوریتم ژنتیک و (W=1) . (ج) مقدار خطا بازشناسی و وزن دهی فردی بهینه برای ویژگی های مکان مشخصه(W) و گرادیان پایه(V) با الگوریتم ژنتیک. ۶۶
شکل۴- ۷: (الف) مقدار خطا بازشناسی و وزن دهی برداری بهینه برای ویژگی مکان مشخصه با الگوریتم ژنتیک و (V=1). (ب) مقدار خطا بازشناسی و وزن دهی برداری بهینه برای ویژگی گرادیان پایه با الگوریتم ژنتیک و (W=1) . (ج) مقدار خطا بازشناسی و وزن دهی برداری بهینه برای ویژگی های مکان مشخصه(W) و گرادیان پایه(V) با الگوریتم ژنتیک. (د)مقدار خطا بازشناسی و وزن دهی برداری بهینه برای ویژگی های مکان مشخصه(W) و گرادیان پایه(V) با الگوریتم PSO. ۶۷
شکل۴- ۸: ساختار ترکیب طبقه بندها ۶۸
شکل۴- ۹: ساختار ترکیب ویژگی با وزن دهی برداری به ویژگی ها و طبقه بند بیز ۶۸
شکل۴- ۱۰: مقدار خطا بازشناسی و وزن دهی برداری بهینه برای سه ویژگی متفاوت با الگوریتم ژنتیک ۶۸
شکل۴- ۱۱: ساختار ترکیب ویژگی با وزن دهی برداری به ویژگی ها و طبقه بند بیز ۶۹
شکل۴- ۱۲: مقدار خطا بازشناسی و وزن دهی برداری بهینه برای سه ویژگی متفاوت با الگوریتم ژنتیک ۶۹
شکل۴- ۱۳: ساختار ترکیب ویژگی با وزن دهی برداری به ویژگی ها و طبقه بند بیز ۷۰
شکل۴- ۱۴: مقدار خطا بازشناسی و وزن دهی برداری بهینه برای سه ویژگی متفاوت با الگوریتم ژنتیک ۷۰
شکل۴- ۱۵: ساختار ترکیب ویژگی سری و موازی با طبقه بند بیز ۷۱
شکل۴- ۱۶: ساختار ترکیب طبقه بندها با طبقه بند بیز ۷۱
شکل۴- ۱۷: بلوک دیاگرام Feature Decision Fusion (FDF) با طبقه بند بیز ۷۲
شکل۴- ۱۸: نمونه های از ارقام دستنویس فارسی از دیتا بیس هدی ۷۴

فهرست جداول

جدول ۴- ۱: نتایج به دست آمده بر روی ارقام با استفاده از ۳طبقه بند متفاوت و ۱۱ ویژگی مختلف (٪) ... ۶۰
جدول ۴- ۲: نتایج به دست آمده روی ارقام با طبقه بندهای متفاوت و ویژگی های مختلف با کاهش افزونگی PCA و LDA (٪) ۶۱
جدول ۴- ۳: نتایج به دست آمده روی ارقام با استفاده از خوشه بندی K_Means بر روی نمو نه های آموزشی (٪) ۶۲
جدول ۴- ۴: نرخ بازشناسی ادغام ویژگی سری زونینگ ۶۳
جدول ۴- ۵: نرخ بازشناسی ادغام ویژگی سری گابور وDCT ۶۴
جدول ۴- ۶: نرخ بازشناسی ادغام ویژگی سری مکان مشخصه و پایه گرادیان ۶۷
جدول ۴- ۷: نرخ بازشناسی ادغام ویژگی سری با سه ویژگی متفاوت ۶۹
جدول ۴- ۸: نرخ بازشناسی ادغام ویژگی سری با سه ویژگی متفاوت ۷۰
جدول ۴- ۹: نرخ بازشناسی ادغام ویژگی سری با سه ویژگی ۷۱
جدول ۴- ۱۰: نرخ بازشناسی ادغام ویژگی سری و موازی با سه ویژگی متفاوت ۷۲
جدول ۴- ۱۱: نرخ بازشناسی مدل Feature Decision Fusion (FDF) ۷۲
جدول ۴- ۱۲: نتایج روشهای مختلفی که در زمینه بازشناسی ارقام دستنویس انجام شده است ۷۳

فصل ۱– مقدمه

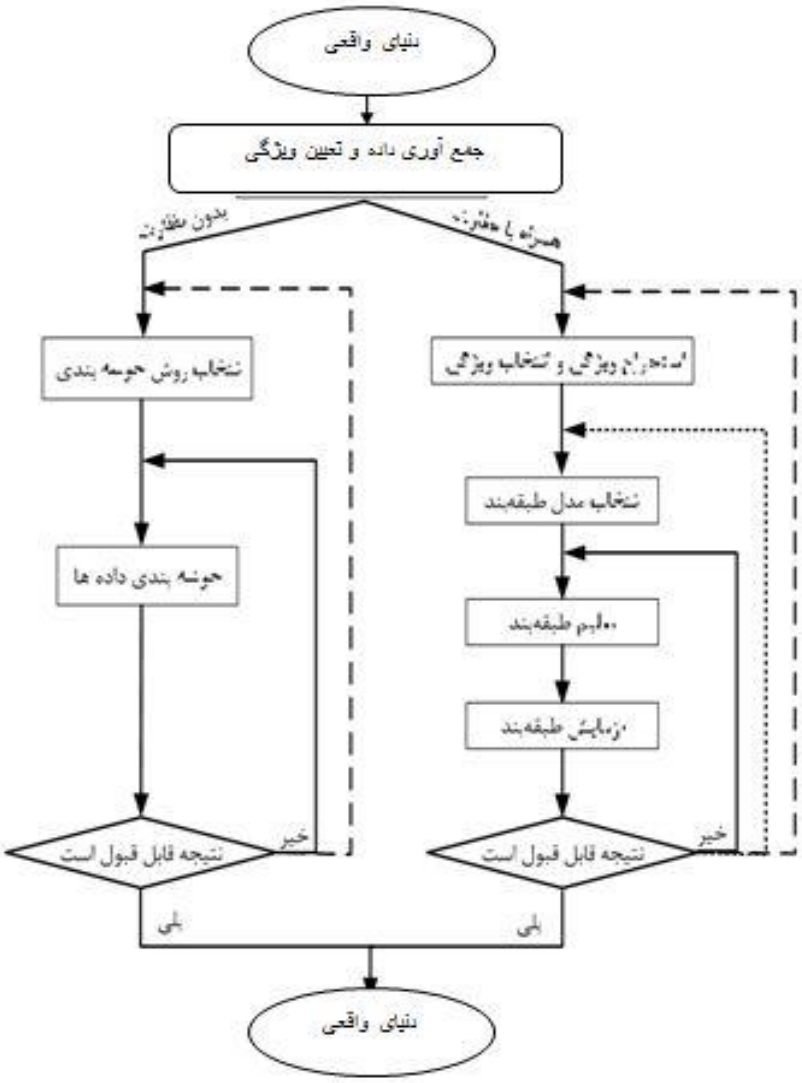
۱–۱– بازشناسی الگو

یادگیری ماشینی یک شاخه مهم از گرایش هوش مصنوعی است که هدف آن تعلیم یک ماشین است به طوری که بتواند تجربیات و نمونه های موجود را یاد بگیرد. حاصل این یادگیری ایجاد یک مدل طبقه بندی است که بر اساس آن ماشین می تواند نمونه هایی را که در آینده می بیند و مشابه نمونه های موجود هستند در کلاس مناسب خود قرار دهد. امروزه روش های بازشناسی الگو، به عنوان یک شاخه از یادگیری ماشینی، کاربردهای فراوانی در زمینه های مختلف علمی و صنعتی پیدا کرده اند[۱].

بازشناسی الگو امری است که در زندگی روزمره ما به وفور رخ می دهد. مانند بازشناسی و تشخیص صدای افراد مختلف، بازشناسی چهره افرادی که با آنها روبرو می شویم و یا تشخیص نوع دستخط افراد مختلف. در واقع بازشناسی الگو امری پیچیده است که بدلیل قابلیتهای بالای مغز انسان به آسانی و در حداقل زمان انجام می شود و ما به دلیل آنکه همواره و به طور ناخودآگاه انواع مختلفی از این بازشناسی را انجام می دهیم، از پیچیده بودن این فرایند بی خبریم. اما اهمیت توانایی انسان در بازشناسی الگو هنگامی مشخص می شود که انسان سعی می کند که ساده ترین نوعهای بازشناسی را که براحتی توسط مغز انسان انجام می شود، توسط رایانه انجام دهد[۲].

در حالت کلی هر توصیف کیفی یا کمی از یک موضوع را می توان یک الگو نامید[۳]. الگو می تواند خود یک شیء باشد و یا ساختار اجزاء یک شیء و روابط بین آنها را توصیف کند. هدف از بازشناسی الگو می تواند شناسایی منطقه ای خاص در تصویر، تشخیص آوا یا کلمه در صدا یا تشخیص هویت باشد. در حال حاضر از تکنیک های بازشناسی الگو در بسیاری از کاربردهای صنعتی، پردازش مستندات، تشخیص هویت و بسیاری از زمینه های دیگر استفاده می شود[۱].

در فرآیند بازشناسی الگو، الگوهای ورودی در کلاس ها و دسته های از پیش تعیین شده ای طبقه بندی می شوند[۴]. مراحل بازشناسی الگو، (در شکل ۱-۱) نشان داده شده است.



شکل ۱- ۱: مراحل سیستم بازشناسی الگو [۱]

چنانچه شکل (۱-۱) نشان می دهد، اولین گام در بازشناسی الگو، جمع آوری تعداد مناسبی نمونه از الگوهای مورد نظر است. این بخش زمان زیادی از فرآیند طراحی سیستم بازشناسی الگو را به خود اختصاص می دهد و گاهی اوقات با مشکلاتی همراه است. پس از جمع آوری نمونه های لازم، باید اقدام به انتخاب نوع ویژگی کرد. انتخاب نوع ویژگی ها نیازمند دانش اولیه در مورد الگوهاست. انتخاب باید به گونه ای باشد که ویژگی ها، وجه اشتراک الگوهای متعلق به یک کلاس و وجه تمایز الگوهای کلاس های مختلف باشند. توانمندی ویژگی برای جداسازی نمونه های کلاس های مختلف، معیار انتخاب آن است[۱].

پس از تعیین نوع ویژگی، باید روش یادگیری را انتخاب کرد. روش یادگیری می تواند از نوع بدون نظارت، همراه با نظارت و یا ترکیبی باشد. در روش یادگیری همراه با نظارت، هر الگو از مجموعه داده با یک برچسب کلاسی همراه است. هدف این است که بر اساس نمونه های موجود، مدل طبقه بندی را طوری بسازیم که بتواند نمونه های را که تاکنون ندیده است با کمترین خطا در کلاس مربوط به خودشان

دسته بندی کند. در یادگیری بدون نظارت، الگوها بر چسب کلاسی ندارند و بر اساس شباهتشان در دسته های یکسان قرار می گیرند[۱].

مرحله بعدی، استخراج ویژگی است. استخراج ویژگی به معنی به دست آوردن خصوصیات الگو ها به نحوی است که این خصوصیات وجه اشتراک الگوهای متعلق به یک کلاس و وجه تمایز الگوهای کلاس های مختلف باشند. استخراج ویژگی یک نگاشت از فضای الگوها به فضای ویژگی ها است که نتیجه آن پیدایش یک بردار ویژگی X است که خصوصیات الگو به صورت معنی دارتر و مناسب تری برای مرحله طبقه بندی آماده می کند. ویژگی های استخراج شده از الگوها ممکن است کیفی[1] (توصیفی) یا کمی[2] (مقداری) باشند. ویژگی های کیفی تعداد محدودی حالت دارند و برای معرفی کیفی الگو استفاده می شوند. به عنوان مثال آفتابی، برفی، بارانی ویژگی های کیفی برای توصیف هوا هستند[۱]. ویژگی های کمی، معمولا یک مقدار پیوسته (تعداد حالات نامحدود) دارند و الگوها را به صورت مقادیر عددی (کمی) توصیف می کنند. به عنوان مثال، مقدار گاز دی اکسیدکربن موجود در یک لیتر هوا، یک ویژگی کمی برای معرفی وضعیت هواست. تبدیل ویژگی های کیفی به کمی نیازمند یک روش یا متدولوژی است[۱]. به عنوان مثال برگزاری یک امتحان روشی است برای کمی کردن میزان یادگیری. در مورد ویژگی های کمی، مؤلفه های بردار ویژگی با اندازه گیری به دست می آیند. هر کمیت اندازه گیری شده، یک ویژگی از شیء مورد نظر را بیان می کند. این کار عملا یک فرآیند کدگذاری است چرا که از نظر هندسی، هر شیء را می توان به عنوان یک نقطه در فضای اقلیدسی در نظر گرفت. انتخاب روش استخراج ویژگی یک عامل مهم در کارایی سیستم های بازشناسی الگو است[۵]. در کاربردهای عملی بازشناسی الگو، معمولا از داده های خام ویژگی های متفاوتی استخراج می شوند و مجموعه آنها برای بازشناسی الگو استفاده می شوند[۶].

ویژگی هایی که از الگوها استخراج می شوند دارای اهمیت یکسان نیستند و برخی از آنها نسبت به بقیه مهمترند. در مرحله انتخاب ویژگی، با انتخاب ویژگی هایی که بیشترین اهمیت را در جداسازی کلاس های الگو دارند، هم بعد بردار ویژگی کاهش داده می شود و هم کیفیت توصیف الگو توسط ویژگی ها بهتر می شود.

انتخاب روش استخراج ویژگی به عنوان مهمترین فاکتور در کارایی سیستم بازشناسی الگو مطرح است. برای شناسایی حروف و ارقام ویژگیهای ناحیه ای، گشتاورهای هندسی، گشتاورهای زرنیکی، توصیفگرهای فوریه، ثابتهای گشتاوری، هیستوگرام نما و توصیف گراف، پیشنهاد شده اند. معمولا با یک ارزیابی تجربی از داده های مورد نظر، ویژگیهای مناسب تر مشخص می شوند.

مرحله بعدی در بازشناسی الگو، انتخاب نوع مدل طبقه بندی است. طبقه بندی به عنوان هسته اصلی یک سامانه بازشناسی الگو مطرح است[۴]. مدل طبقه بندی، هر الگوی ناشناخته را بر اساس ویژگی های آن، به یکی از کلاس های شناخته شده نسبت می دهد. به بیان ریاضی، به هر نگاشتی از فضای n بعدی ویژگی ها، به فضای k بعدی برچسب های کلاسی، یک طبقه بندی گفته می شود.

[1] Qualitative Features
[2] Quantative Features

$$D: R^n \to \Omega \qquad (۱-۱)$$

که در آن $\Omega = \{\omega_1, \omega_2, ..., \omega_k\}$ مجموعه برچسب های کلاسی الگوهاست. یک طبقه بند، میزان تعلق بردار ویژگی $x \in R^n$ به هر یک از کلاس های الگو را به صورت یک عدد حقیقی بیان می کند. مدل های طبقه بندی متعددی برای بازشناسی الگو پیشنهاد شده اند که می توان به طبقه بند بیز، شبکه های عصبی، kهمسایه نزدیکتر، ماشین بردار پشتیبان[۱] و طبقه بند کمترین فاصله[۲] اشاره کرد. پس از انتخاب مدل طبقه بندی، باید پارامترهای آن را مشخص کرد. روش های مختلفی برای تعیین این پارامترها ارائه شده اند. گروهی از این روش ها بر اساس خواص آماری، مانند بردار میانگین و ماتریس کوواریانس نمونه های آموزشی هر کلاس در طبقه بند بیز، پارامترهای طبقه بندی را تعیین می کنند. گروه دیگری از روش ها وجود دارند که با استفاده از الگوریتم های یادگیری تکراری، مانند آموزش یک طبقه بند شبکه عصبی بر روی نمونه های یادگیری، پارامترهای طبقه بندی را تعیین می کنند. در روش های گروه دوم، پارامترهای طبقه بندی در طول فرآیند یادگیری، تعیین می شوند. پس از آنکه مدل طبقه بندی کامل شد، با استفاده از نمونه های آزمایشی، سامانه بازشناسی الگو اعتبارسنجی می شود. اگر نرخ بازشناسی مورد نیاز بدست نیامد، می توان یک یا چند تا از کارهای زیر را انجام داد[۱].

۱. ویژگی های استخراج شده از الگو یا ویژگی های انتخاب شده از بین ویژگی های قبلی را بهبود داد.
۲. مدل طبقه بندی را عوض کرد یا پارامترهای آن را تغییر داد.
۳. یادگیری (تعلیم) طبقه بند را کامل تر کرد.

کارهای فوق آنقدر تکرار می شوند تا تضمین کافی جهت یک تعمیم معتبر حاصل شود.
در یادگیری بدون نظارت، پس از جمع آوری داده های کافی از الگوهای مورد نظر، یک روش خوشه بندی انتخاب می شود(به عنوان مثال روش K میانگین[۳]) و بر اساس آن، داده ها خوشه بندی می شوند. اگر نتیجه خوشه بندی قابل قبول نباشد، با تغییر روش خوشه بندی یا تغییر پارامترهای خوشه بندی، روند کار ادامه پیدا می کند تا نتیجه مورد نظر حاصل شود.

۱-۲- روشهای بازشناسی الگو

تکنیک های بازشناسی الگو، یک شیء یا یک رخداد فیزیکی را به یک یا چند کلاس الگوی از قبل معلوم نسبت می دهند. بنابراین یک سامانه بازشناسی الگو می تواند به عنوان یک قاعده تصمیم گیری خودکار در نظر گرفته شود که با اندازه گیری های انجام شده روی الگو آن را به یک کلاس نسبت می دهد. روش

[1] Support Vector Machine (SVM)
[2] Minimum Distance Classifier
[3] K-Means

های انتساب الگو به یکی از کلاس ها، به سه گروه عمده تقسیم می شوند. روش های حسی تجربی[1]، روش های نحوی یا ساختاری[2] و روش های ریاضی[3][۱].

۱-۲-۱- روش های حسی - تجربی

این روش ها بر مبنای تجربیات طراح استوار هستند و یک سامانه مبتنی بر این روش ها شامل مجموعه ای از قواعد تجربی است که با توجه به خصوصیات الگوها حاصل شده اند. به عنوان مثال سامانه بازشناسی دست نوشته های فارسی (مبتنی بر روش های حسی- تجربی) شامل قواعدی حسی است که مخصوص الگوهای دست نویس است و ممکن است برای بازشناسی دست نوشته های لاتین مناسب نباشد. این قواعد به عنوان مثال بر مبنای تعداد و توالی حرکات قلم، تعداد و موقعیت نقاط و علائم و تعداد زیر کلمات هستند که به صورت تجربی به دست 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 روش های حسی - تجربی بستگی به گستره تجربیات طراح دارد[۳].

۱-۲-۲- روش های ساختاری یا نحوی

در خیلی از مسائل پیچیده، تعداد ویژگی های مورد نیاز ممکن است بسیار زیاد باشند. در چنین شرایطی می توان یک الگو را مجموعه ای از چند الگوی ساده تر در نظر گرفت. هر الگوی ساده می تواند از چندین زیر الگوی ساده تر تشکیل شده باشد. ساده ترین زیر الگوها را عناصر پایه[4] می گویند. بیان الگو بر حسب عناصر پایه، اساس شناسایی ساختاری الگو را تشکیل می دهد. ویژگی های ساختاری می توانند 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 شناخته می شوند. با استفاده از الفبای زبان و گرامر آن می توان مجموعه پیچیده ای از الگوها را با مجموعه ی کوچکی از عناصر پایه و قواعد گرامری زبان بیان کرد. گرامرهای هر کلاس با توجه به نمونههای آموزشی همان کلاس ساخته می شوند. روش های ساختاری، زمانی مفید هستند که الگوها، ساختارهای قابل تعریفی به وسیلهٔ مجموعه ای از قواعد داشته باشند. به دلیل وجود الگوهای نویزی در شناسایی عناصر پایه و قواعد گرامری، پیاده سازی ساختاری دارای مشکلات زیادی است[۳].

۱-۲-۳- روش های ریاضی

در روش های ریاضی بازشناسی الگو، هر الگوی $x=\{x_1, x_2, ..., x_n\}$ به صورت نقطه ای در فضا n بعدی مدل شده و با توجه به مقدار ویژگی هایش به یکی از کلاس ها نسبت داده می شود. ویژگی ها به

[1] Heuristic Methods
[2] Syntactic (Structural) Methods
[3] Mathematical Methods
[4] Primitive Elements

گونه ای انتخاب می شوند که بردارهای نمونه های مربوط به کلاس های مختلف، ناحیه های متفاوتی از فضا را اشغال می کنند. روش های ریاضی، بر مبنای توابع تصمیم[1] (جداساز) برای کلاس های الگو بنا شده اند. برای یک مسأله طبقه بندی K کلاسه نیاز به پیدا کردن K تابع تصمیم $d_1(x), d_2(x), ..., d_k(x)$ با خاصیت زیر است.

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

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 این صورت، مرز تصمیم جدا کننده کلاس ω_i از ω_j به صورت زیر خواهد بود.

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

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 نمونه های آموزشی آن کلاس استفاده می شود.

روش های ریاضی شامل دو گروه آماری[2] و غیر آماری[3] هستند. از این دو گروه با عنوان پارامتریک (غیر آموزش پذیر[4]) و غیر پارامتریک (آموزش پذیر[5]) نیز یاد می شود[۴]. در روش های پارامتریک، پارامترهای تابع تصمیم مستقیما از خصوصیات آماری نمونه های آموزشی تعیین می شوند و در روش های غیر پارامتریک، پارامترهای تابع تصمیم معمولا در یک فرآیند یادگیری تعیین می شوند. بسیاری از تکنیک های ارائه شده در روش های پارامتریک مبتنی بر فرض نرمال بودن توابع توزیع الگوها هستند. در صورتی که تعداد نمونه های موجود در هر کلاس کم باشد، معمولا روش های غیر پارامتریک بهتر عمل می کنند[۷].

۱-۲-۳-۱- روش های آماری

روش های آماری بر مبنای قواعد طبقه بندی ریاضی با بهره گیری از خصوصیات آماری الگوها استوار هستند. در این روش ها، مرزهای تصمیم گیری به وسیله توزیع های آماری نمونه ها تعیین می شوند. طبقه بندی های الگو برای روش های آماری عموما بر مبنای قاعده تصمیم بیز[6] کار می کنند. با به کار گیری این قاعده، اگر تابع چگالی احتمال هر یک از کلاس های الگو و همچنین احتمال پیشین هر کلاس مشخص باشد، سطح تصمیم بهینه پیدا می شود[۱].

---

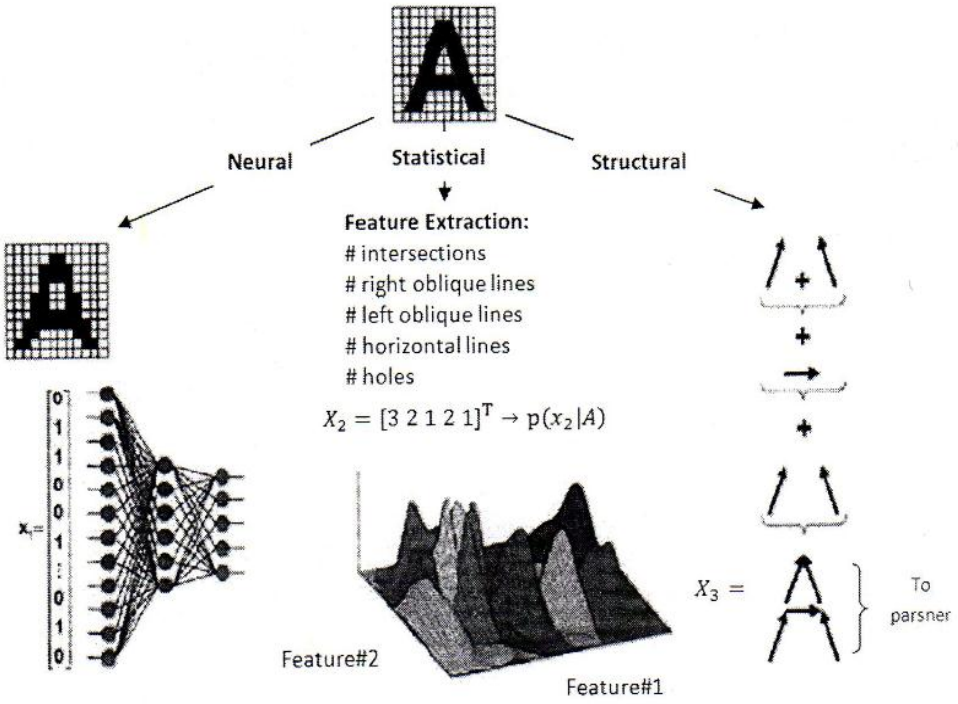
[1] Decision(Discriminant)Functions
[2] Statistical
[3] Deterministic
[4] Non Trainable
[5] Trainable
[6] Bayes Decision

۱-۲-۳-۲- روش های غیر آماری

روش های غیر آماری بر مبنای یک چارچوب ریاضی بدون استفاده صریح از خصوصیات آماری الگوها استوار هستند. در این روش ها، ابتدا فرم کلی مرز جدا کننده (مانند توابع خطی یا درجه دوم) مشخص شده و سپس پارامترهای آن با استفاده از نمونه های آموزشی تعیین می شود. الگوریتم های یادگیری تکراری، نمونه ای از روش های غیر آماری هستند که در آنها توابع تصمیم از الگوهای آموزشی و در یک فرآیند تکرار تولید می شوند. یعنی پس از انتخاب نوع تابع تصمیم، ضرایب مربوط به آن در یک فرآیند تکراری، با استفاده از نمونه های آموزشی، تعیین می شوند. شبکه های عصبی پرسپترون چند لایه از متداولترین روش های غیر آماری برای بازشناسی الگو هستند. برخی مراجع روش های غیر آماری را با عنوان روش های عصبی نیز معرفی کرده اند[۷]. در شکل (۱-۲) روش های غیر آماری، آماری و ساختاری برای بازشناسی حرف A نشان داده شده اند. روش آماری از تابع چگالی احتمال ویژگی ها استفاده می کند و روش ساختاری با تبدیل الگو به سه عنصر پایه به بازشناسی آن می پردازد.

انتخاب روش بازشناسی الگو به پارامترهایی نظیر تعداد کلاس های الگو، تنوع آنها، اعوجاج و نویز در الگوها، تعداد ویژگی ها، تعداد نمونه های یادگیری، سرعت و دقت مورد نظر بستگی دارد[۱].

در این پایان نامه از طبقه بند بیز، طبقه بند ۳ همسایه نزدیکتر و طبقه بند کمترین فاصله برای بازشناسی الگو استفاده خواهیم کرد. بنابراین در فصل سوم به بررسی و معرفی بیشتر این روش های طبقه بندی می پردازیم.



شکل ۱- ۲: روش های غیرآماری، آماری و ساختاری برای بازشناسی حرف A [۷]

۱–۳– ارزیابی یک طبقه بند

پس از آنکه مدل طبقه بندی کامل شد، با استفاده از نمونه های آزمایشی، سیستم بازشناسی الگو اعتبار سنجی می شود. اگر نرخ بازشناسی مورد نیاز بدست نیامد باید یک یا چند تا از کارهای زیر را انجام داد

الف – ویژگیهای استخراج شده از الگو یا ویژگیهای انتخاب شده از بین ویژگی های قبلی رابهبود داد .

ب – مدل طبقه بندی را عوض کرد یا پارامترهای آن را تغییر داد .

ج - یادگیری طبقه بند را کامل تر کرد .

کارهای فوق آنقدر تکرار می شوند تا تضمین کافی جهت یک تعمیم معتبر حاصل شود[۱] . یکی از معیارهای رایج برای ارزیابی یک طبقه بند، استفاده از ماتریس سردرگمی[1] آن است . این ماتریس با استفاده از عملکرد طبقه بند روی نمونه های آزمایشی ساخته می شود . برای k کلاس الگو ، بدون احتساب کلاس واپس زده[2]، این ماتریس یک ماتریس $K \times K$ است که در آن درایه C_{ij} بیانگر تعداد الگوهایی از کلاس ω_i است که طبقه بند آنها را در کلاس ω_j قرار داده است .

$$CM = [C_{ij}]_{K\times K} = \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1j} & \dots & C_{1K} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ C_{i1} & C_{i2} & \dots & C_{ij} & \dots & C_{iK} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ C_{K1} & C_{K2} & \dots & C_{Kj} & \dots & C_{KK} \end{bmatrix} \qquad (۴-۱)$$

با احتساب کلاس واپس زده، این ماتریس دارای اندازه $k \times (k+1)$ خواهد بود که در آن کلاس $(k+1)$ ام کلاس واپس زده است. این کلاس شامل الگوهایی است که طبقه بند در مورد کلاس آن ها اظهارنظر نمی کند . این ماتریس در حقیقت سابقه قبلی طبقه بند روی نمونه های آموزشی را نشان می دهد و حاوی اطلاعات مفیدی در مورد رفتار طبقه بند است. به عنوان مثال مجموع عناصر سطر i ام این ماتریس تعداد الگوهای آموزشی کلاس ω_i را نشان می دهد. سطرهای این ماتریس (پس از تقسیم عناصر بر مجموع عناصر آن سطر) مفهوم دقت[3] و ستون های آن (پس از تقسیم عناصر بر مجموع عناصر آن ستون) مفهوم فراخوانی[4] رادارند[۱].

بعضی از معیارهایی که از این ماتریس قابل استخراج هستند به شرح زیر می باشند .

[1] Confusion Matrix
[2] Rejected Class
[3] Precision
[4] Recall

۱-۳-۱- دقت طبقه بند برای کلاس ω_i

بیانگر آن است که چه میزان از داده های مربوط به کلاس ω_i درست تشخیص داده شده اند . برای محاسبه این معیار کافی است درایه متناظر با این کلاس در قطر ماتریس سردرگمی را بر مجموع عناصر سطر مربوط به آن تقسیم کنیم یعنی :

$$P_i = \frac{C_{ii}}{\sum_{k=1}^{k} C_{ik}} \qquad (۵-۱)$$

بر اساس این معیار ، خطای حذف شدگی[1] به صورت $eo_i = 1 - P_i$ تعریف می شود و بیانگر آنستکه چه میزان از داده های مربوط به کلاس ω_i به صورت نادرست تشخیص داده شده اند .

متوسط دقت طبقه بند روی کل کلاسها نیز به صورت رابطه زیر تعریف می شود و بیانگر آنستکه به طور متوسط چه میزان از داده ها توسط این طبقه بند به درستی تشخیص داده شده اند.

$$P_{av} = \frac{\sum_{i=1}^{K} P_i}{K} \qquad (۶-۱)$$

۱-۳-۲- فراخوانی طبقه بند روی کلاس ω_i

بیانگر این است که چه تعداد از الگوهایی که به کلاس ω_i تخصیص یافته اند واقعا متعلق به خود آن کلاس بوده اند . برای محاسبه آن کافی است درایه متناظر با این کلاس در قطر ماتریس سردرگمی را بر مجموع عناصر ستون مربوط به آن تقسیم کنیم یعنی :

$$R_i = \frac{C_{ii}}{\sum_{i=1}^{k} C_{ik}} \qquad (۷-۱)$$

در بعضی مقالات از پارامترهای دقت و فراخوانی با عنوان صحت و اعتبار طبقه بند یاد شده است .

۱-۳-۳- معیار F

این معیار بیشتر در بحث بازیابی[2] مطرح است و ترکیبی از دقت و فراخوانی طبقه بند است و به صورت زیر تعریف می شود :

$$F_i = \frac{2 \times P_i \times R_i}{P_i + R_i} \qquad (۸-۱)$$

۱-۳-۴- نرخ بازشناسی کلی یا دقت طبقه بند یا قابلیت اعتماد بازشناسی

اعتماد بازشناسی[3] برابر مجموع عناصر روی قطر اصلی ماتریس سردرگمی تقسیم بر مجموع کل عناصر آن است یعنی:

[1] Omission
[2] Retrieval
[3] Recognition Reliability