دانشگاه آزاد اسلامی

واحد تهران جنوب

دانشکده فنی و مهندسی

پژوهش برای درس روش پژوهش و ارائه

مهندسی (کامپیوتر– نرم افزار)

عنوان:

دسته بندی ، شناسایی اشیاء ، تحلیل احساسات در یادگیری ماشین

استاد:

جناب آقای حسین کی نژاد

نام و نام خانوادگی دانشجویان/شماره دانشجویی:

لیلی حسین زاده/۹۶۲۵۵۱۱۴۸۱

شیرین شغلی/۹۶۲۵۵۱۲۱۲۳

بهار کفتانی /۹۶۲۵۵۱۲۵۲۷

تیر۱۳۹۹

بسم الله الرحمن الرحیم

فهرست مطالب

عنوان مطالب	شماره صفحه
چکیده	1
مقدمه	2
Classification	6
کاربرد های classification	6
الگوریتم KNN	9
مزایا و معایب KNN	12
Normalization داده ها	12
کد MATLAB الگوریتم KNN	13
روش الگوریتم KNN مبتنی بر وزن فاصله	16
کاربرد های الگوریتم KNN	17
الگوریتم دسته بندی کننده بیز ساده	18
نظریه بیزین	19
فرضیه با حداکثر احتمال پسین (Maximum A Posteriori Probability)	20
فرضیه با بیشینه احتمال (Maximum Likelihood)	21
Naïve Bayes Classifier	21
یک مثال ساده از این الگوریتم	23

فهرست مطالب

عنوان مطالب	شماره صفحه
مزایا و معایب این دسته بندی	26
کاربرد های الگوریتم بیز ساده	27
الگوریتم K-means	28
انواع روش های خوشه یابی	29
شیوه کارکرد الگوریتم K-means	29
مزایا و معایب الگوریتم K- means	30
تفکیک تصویر با استفاده از K-means	31
تشخیص الگو	38
انواع تشخیص الگو	38
استخراج ویژگی	39
الگوها و کلاس الگوها	40
ویژگی های تصویر	41
استخراج ویژگی های شکلی	42
آشکارساز لبه	44
عملگرهای آشکارساز لبه	46
الگوریتم sobel	47
آستانه گیری	48

فهرست مطالب

عنوان مطالب	شماره صفحه
مقدمه ای بر شبکه عصبی	49
عصب چیست؟	50
سیناپس چیست؟	50
مقایسه اجمالی رایانه و شبکه عصبی	53
پیش پردازش تصویر	57
استخراج ویژگی ها	59
پرسپترون	62
قانون دلتا	66
الگوریتم gradient descent	66
محاسبه گرادیان	67
تست شبکه عصبی	68
تحلیل احساسات	79
فصل اول(مقدمه)	79
مقدمه	79
بیان مساله	80
ویژگی های مساله و مسایل پیش رو	81
خلاصه ی مدل پیشنهادی	82

فهرست مطالب

عنوان مطالب	شماره صفحه
ساختار پژوهش جاری	82
فصل دوم(مفاهیم پایه)	83
مقدمه	83
اصطلاحات، مفاهیم و تعاریف پایه تحلیل احساسات	83
نمودار معماری کلی یک سامانه ی تحلیل احساسات	87
الگوریتم های مطرح در یادگیری ماشین	88
فصل سوم(مروری بر کارهای گذشته)	94
تحلیل احساسات در سطح متن	94
تحلیل احساسات در سطح جمله	95
تحلیل احساسات بر اساس ویژگی	95
تحلیل احساسی رقابتی	96
دستیابی به لغتنامه	97
پژوهش های گذشته در زمینه ی تشخیص محتوای فیلم ها	97
کاربردهای تحلیل احساسات	98
فصل چهارم(مدل پیشنهادی)	99
راهکار پیشنهادی	99
طرح کلی راهکار پیشنهادی	100

فهرست مطالب

عنوان مطالب	شماره صفحه
استخراج ویژگی ها	104
طبقه بندی تک برچسب بر مبنای ویژگی های انتخاب شده در روش های اول و دوم	109
ارزیابی و تخمین خطا	110
فصل پنجم(ارزیابی و تحلیل نتایج)	111
انتخاب ژانرهای مناسب	111
نتایج حاصل از اجرا	114
مقایسه ی نتایج	119
اطلاعات کلی و یافته های جانبی	119
فصل ششم(نتیجه گیری)	121
نتیجه گیری	121
مدل سازی زبان فارسی	123
سیستم خبره ی پیشنهاد دهنده ی فیلم بر اساس علاقه	123
داده کاوی مکانی	124
پیشنهاد برای پژوهش در حوزه ی تجزیه و تحلیل احساسات	124
نتیجه گیری نهایی	126
منابع و مراجع	127

چکیده مطالب:

انسان همواره در تکاپوی دسته بندی داده های ذهنی خود بوده است تا بتواند به عملکرد بهتری در کار خود دست یابد . بدیهی است در موضوع یادگیری ماشین که تلاش ما بر این است تا ماشین ها را قادر به الگو برداری از ذهن انسان کنیم، دسته بندی داده های ورودی برای بهبود عملکرد امر مهمی تلقی میشود. در ابتدای این پژوهش سعی شده است ضمن معرفی مفاهیم اولیه ی کلاس بندی، به الگوریتم هایی همچون KNN و k means بپردازیم.در قسمت شناسایی تصویر هدف، شناسایی اشیای هندسی با استفاده از ویژگی های شکلی میباشد. در ابتدا اشیای هندسی متفاوت، به شبکه ی عصبی شناسانده و آموزش داده میشوند و سپس با استخراج ویژگی های شکلی و نیز تست شبکه توسط داده های متفاوت از آنچه که در مرحله ی آموزش به شبکه ارائه شد، راستی آزمایی صورت میگیرد. در قسمت نهایی این پژوهش، تلاش کردیم تا به تحلیل احساسات و دیگر الگوریتم های به کار رفته در بحث یادگیری ماشین بپردازیم. به طور کلی آنالیز احساسات یا کاویدن نظرات، به فرآیند کشف خودکار عقاید و نظرات نویسندگان در مورد یک موجودیت خاص گفته میشود . ما با در نظر گرفتن زیرنویس فیلم ها و همچنین تکیه بر الگوریتم های مطرح در یادگیری ماشین، اقدام به طبقه بندی فیلم ها به ژانر آنها نموده ایم.

مقدمه:

هدف یادگیری ماشین این است که رایانه‌ها و سامانه‌ها بتوانند به تدریج و با افزایش داده‌ها کارایی بهتری در انجام وظیفهی مورد نظر پیدا کند .گسترهی این وظیفه می‌تواند از تشخیص خودکار چهره با دیدن چند نمونه از چهرهی مورد نظر تا فراگیری شیوهی گام‌برداری روبات‌های دوپا با دریافت سیگنال پاداش و تنبیه باشد .

طیف پژوهش‌هایی که در یادگیری ماشین می‌شود گسترده‌است .به لحاظ نظری پژوهش‌گران بر آن‌اند که روش‌های یادگیری تازه‌ای به وجود بیاورند و امکان‌پذیری و کیفیت یادگیری را برای روش‌های‌شان مطالعه کنند و در سوی دیگر عده‌ای از پژوهش‌گران سعی می‌کنند روش‌های یادگیری ماشینی را بر مسایل تازه‌ای اعمال کنند. البته این طیف گسسته نیست و پژوهش‌های انجام‌شده دارای مؤلفه‌هایی از هر دو روی‌کرد هستند .

یادگیری ماشین کمک فراوانی به صرفه جویی در هزینه‌های عملیاتی و بهبود سرعت عمل تجزیه و تحلیل داده‌ها می‌کند. به عنوان مثال در صنعت نفت و پتروشیمی با استفاده از یادگیری ماشین، داده‌های عملیاتی تمام حفاری‌ها اندازه‌گیری شده و با تجزیه و تحلیل داده‌ها، الگوریتم‌هایی تنظیم می‌شود که در حفاری‌های بعدی استخراج پربازده و بهینه‌تری داشته باشیم[1]

یکی از تقسیم‌بندی‌های متداول در یادگیری ماشینی، تقسیم‌بندی بر اساس نوع داده‌های در اختیار کارگزار هوشمند است. به سناریوی زیر توجه کنید :

فرض کنید به تازگی رباتی سگ‌نما خریده‌اید که می‌تواند توسط دوربینی دنیای خارج را مشاهده کند، به کمک میکروفن‌هایش صداها را بشنود، با بلندگوهایی با شما سخن بگوید (گیریم محدود) و چهارپایه‌اش را حرکت دهد. هم‌چنین در جعبهی این ربات دستگاه کنترل از راه دوری وجود دارد که می‌توانید انواع مختلف دستورها را به ربات بدهید. در پاراگراف‌های آینده با بعضی از نمونه‌های این دستورها آشنا خواهید شد .

اولین کاری که می‌خواهید بکنید این است که اگر ربات شما را دید خرناسه بکشد اما اگر غریبه‌ای را مشاهده کرد با صدای بلند عوعو کند. فعلاً فرض می‌کنیم که ربات توانایی تولید آن صداها را دارد اما هنوز چهرهی شما را یادنگرفته‌است. پس کاری که می‌کنید این است که جلوی چشم‌های‌اش قرار می‌گیرید و به کمک کنترل از راه دورتان به او دستور می‌دهید که چهره‌ای که جلوی‌اش می‌بیند را با خرناسه‌کشیدن مربوط کند. این‌کار را برای چند زاویهی مختلف از صورت‌تان انجام می‌دهید تا مطمئن باشید که ربات در صورتی که شما را از مثلاً نیم‌رخ ببیند بهتان عوعو نکند. هم‌چنین شما چند چهرهی غریبه نیز به او نشان می‌دهید و چهرهی غریبه را با دستور عوعوکردن مشخص می‌کنید .در این حالت شما به کامپیوتر ربات گفته‌اید که چه ورودی را به چه خروجی مربوط کند. دقت کنید که هم ورودی و هم خروجی مشخص است و در اصطلاح خروجی برچسب‌داراست. به این شیوهی یادگیری، یادگیری با نظارت می‌گویند .

اینک حالت دیگری را فرض کنید. برخلاف دفعهی پیشین که به ربات‌تان می‌گفتید چه محرکه‌ای را به چه خروجی ربط دهد، این‌بار می‌خواهید ربات خودش چنین چیزی را یاد بگیرد. به این صورت که اگر شما را دید و خرناسه کشید به نحوی به او پاداش دهید (مثلاً به کمک همان کنترل از راه دورتان) و اگر به اشتباه به شما عوعو کرد، او را تنبیه کنید (باز هم با همان کنترل از راه دورتان). در این حالت به ربات نمی‌گویید به ازای هر شرایطی چه کاری مناسب است، بلکه اجازه می‌دهید ربات خود کاوش کند و تنها شما نتیجهی نهایی را تشویق یا تنبیه می‌کنید. به این شیوهی یادگیری، یادگیری تقویتی می‌گویند .

در دو حالت پیش قرار بود ربات ورودی را به خروجی مرتبط کند. اما گاهی وقت‌ها تنها می‌خواهیم ربات بتواند تشخیص دهد که آنچه می‌بیند (یا می‌شنود و...) را به نوعی به آنچه پیش‌تر دیده‌است ربط دهد بدون این‌که به‌طور مشخص بداند آن‌چیزی که دیده شده‌است چه چیزی است یا این‌که چه کاری در موقع دیدنش باید انجام دهد. ربات هوش‌مند شما باید بتواند بین صندلی و انسان تفاوت قایل شود بی‌آنکه به او بگوییم این نمونه‌ها صندلی‌اند و آن نمونه‌های دیگر انسان. در این‌جا برخلاف یادگیری با نظارت هدف ارتباط ورودی و خروجی نیست، بلکه تنها دسته‌بندی‌ی آن‌ها است. این نوع یادگیری که به آن یادگیری بی نظارت می‌گویند بسیار مهم است چون دنیای ربات پر از ورودی‌هایی است که کسی برچسبی به آن‌ها اختصاص نداده اما به وضوح جزئی از یک دسته هستند .

یادگیری بی‌نظارت را می‌توان به صورت عمل کاهش بعد در نظر گرفت .

از آن‌جا که شما سرتان شلوغ است، در نتیجه در روز فقط می‌توانید مدت محدودی با رباتتان بازی کنید و به او چیزها را نشان دهید و نام‌شان را بگویید (برچسب‌گذاری کنید). اما ربات در طول روز روشن است و داده‌های بسیاری را دریافت می‌کند. در این‌جا ربات می‌تواند هم به خودی خود و بدون نظارت یاد بگیرد و هم این‌که هنگامی که شما او را راه‌نمایی می‌کنید، سعی کند از آن تجارب شخصی‌اش استفاده کند و از آموزش شما بهره ی بیش‌تری ببرد .ترکیبی که عامل هوش‌مند هم از داده‌های بدون برچسب و هم از داده‌های با برچسب استفاده می‌کند به یادگیری نیمه نظارتی می‌گویند[1]

دسته بندی

- **Classification**

در حوزه ی علوم کامپیوتر و علم پردازش داده ، دسته بندی کردن اطلاعات برای پردازش سریعتر آن ها امری مرسوم و قابل استناد است . حال علم یادگیری ماشیین که نیز با داده های وسیعی سر و کار دارد نیز از این قاعده مستثنی نیست .

رده بندی و یا دسته بندی یکی از شاخه های بسیار پر کاربرد حوزه یادگیری ماشین است . در این شاخه روش ها و الگوریتم هایی معرفی میشود . این الگوریتم ها هر کدام کار کرد خود را در مسائل متفاوت دارند و ما باید با توجه به شرایط مسئله تصمیم بگیریم که از کدام روش استفاده کنیم تا کارکرد و بازدهی بالاتری داشته باشیم . پس به شکل کلی هدف از این کار ایجاد برچسب بر روی بخشی از نمونه های داده ای است و تعیین تکلیف کلی برای باقی نمونه های ورودی است . [2]

داده های ورودی میتوانند از انواع داده های پیوسته و یا گسسته تشکیل شده باشند. با این وجود میتوان نتتیجه گرفت که عمل رده بندی بر روی داده های گسسته و داده های پیوسته دارای روش های متفاوتی خواهد بود .

رده بندی بر روی داده های پیوسته را رگرسیون و رده بندی بر روی داده های گسسته را دسته بندی میگوند که هر کدام دارای روش ها و الگوریتم های متفاوتی هستند . [2,5]

- **کاربرد های classification**

بحث دسته بندی یکی از مباحث اصلی و پایه ای در علم یادگیری ماشین است ؛ به نوعی در زمره ی یکی از پایه ای ترین پردازش ها در این مقوله است و در اکثر زیر شاخه های یادگیری

ماشین به نوعی از الگوریتم های این روش استفاده میشود به نوعی هر کجا که کار با داده ها مطرح باشد استفاده از روشی برای دسته بندی آن ها کارا خواهد بود .[2]

از کار برد های مهم آن میتوان به مواردی همچون :

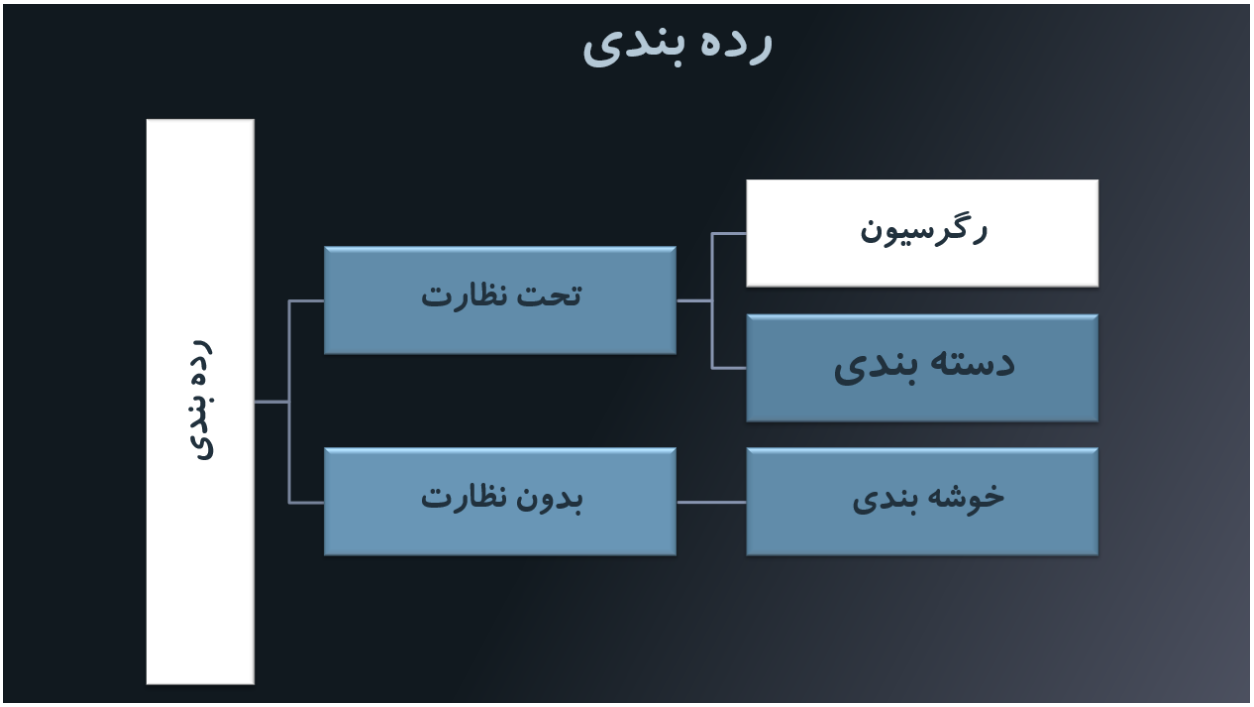
- **تشخیص صفحات هرز در محیط وب ها** [2,21]
- **تشخیص تقلب در سیستم های مالی و بیمه ای و ...** [2]
- **تشخیص انواع بیماری ها در حوزه پزشکی و سلامت** [2]
- **پیشبینی های مختلف** [4]
- **تشخیص صوت** [3]
- **تشخیص و پردازش چهره**[3]
- **تشخیص موارد نامتعارف**(outlier detection)[3]

شایان ذکر است برخی از موارد نام برده شده تنها با روش های دسته بندی قابل حل نیستند و نیاز به دانستن روش ها و الگوریتم ها دیگری نیز خواهد بود .

همانطور که پیش از این مطرح شد ، پردازش ها در این حوزه بر دو دسته صورت میگرد پردازش های تحت نظارت (supervised) **و یا بدون نظارت** (unsupervised) **.**

پردازش های تحت نظارت همانگونه که پیشتر گفته شد بر اساس پیوسته و یا گسسته بودن داده ها به دو دسته ی رگرسیون و یا دسته بندی گفته میشود .

و پردازش های بدون نظارت نیز در دسته ی رده بندی های خوشه یابی قرار میگیرند.(شکل ۱)[5]

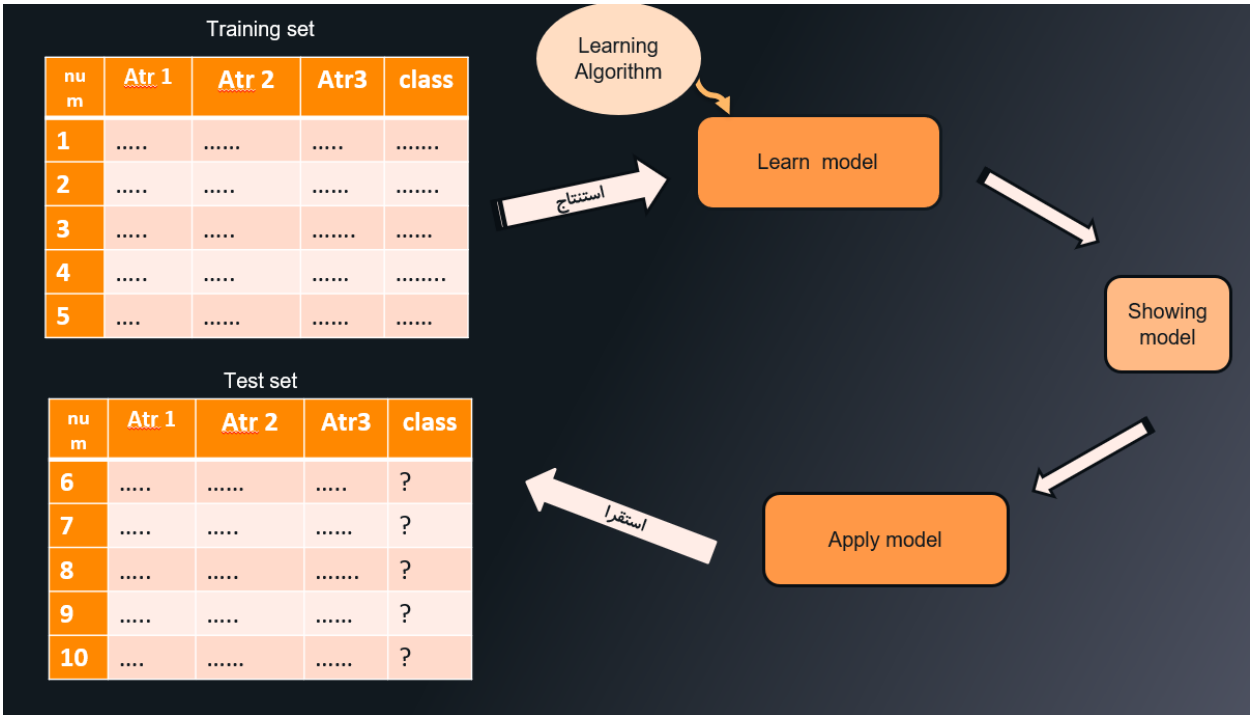


(شکل ۱ نمای کلی ای از رده بندی)

در این تحقیق ما به معرفی الگوریتم هایی از شاخه دسته بندی و خوشه یابی میپردازیم .

با تفاسیر گفته شده تا به اینجا از classification **(دسته بندی) یک فرآیند دو مرحله ایست و در دو مرحله قابل انجام است . در گام اول میبایست تا مدلی برای رده بندی انتخاب کنیم برای انجام این کار باید ابتدا رده ها و یا دسته ها را مشخص کنیم ، یعنی مشخص شود که قرار است داده هارا به چند دسته تقسیم کنیم ؛ سپس لازم است تا رده ی هر نمونه ی داده ای آموزشی مشخص شود و برچسب گذاری شود و در انتها میبایست مدل انتخاب شده مان برای انجام دسته بندی را نمایش دهیم که نمایش آن به صورت مجموعه ای از قواعد ، درخت تصمیم و فرمول های ریاضی و ... خواهد بود .**[2]

در گام دوم باید از مدل رده بندی انتخاب شده استفاده کنیم به منظور تشخیص رده برای هر نمونه ی داده ای تست و در نهایت هم باید دقت مدل استفاده شده را ارزیابی کنیم . به گونه ای که برچسب اتخاذ شده برای هر نمونه را با برچسب واقعی آن ها مقایسه کنیم . با محاسبه درصد داده های تست که به درستی برچسب گذاری شده اند میتوان پی به دقت مدل انتخاب شده برد . (شکل ۲)[2]



(شکل ۲ نمای کلی از مراحل دسته بندی داده های گسسته)

- **الگوریتم KNN**

اولین الگوریتمی که به تحلیل آن خواهیم پرداخت الگوریتم KNN (k-nearest neighbor) است . این الگوریتم در شمار الگوریتم های تحت نظارت برای دسته بندی داده های گسسته است . روشی بسیار ساده و برای مسائل بسیاری کارا خواهد بود .اما پیش از آن که وارد جزئیات مطالب این الگوریتم شویم باید به مفهوم دیگری بپردازیم .

الگوریتم knn در شمار الگوریتم های یادگیری مبتنی بر نمونه هاست . یعنی پردازش های ما براساس تک نمونه هاست .[2]

- **مدل یادگیری حریص و تنبل (Eager & Lazy)**

  در مدل یادگیر lazy باید منتظر باشیم تا زمان دریافت یک در خواست از کاربر برای شروع یادگیری که الگوریتم های KNN و استدلال های مبتنی بر مورد نیز در ایم دسته قرار میگیرند .

در مدل یادگیری Eager باید پیش از آن که از کاربر درخواستی بیاید مدلی بر اساس نمونه های داده آموزشی تولید بشود .[2]

در الگوریتم KNN ایده ی اصلی آن بر اساس انتساب تجمیعی از برچسب های همسایگان به داده ی تست وارد شده میباشد و مبنای کار آن براساس فواصل است .[2]]

شیوه ی کار در این الگوریتم به شرح زیر خواهد بود :

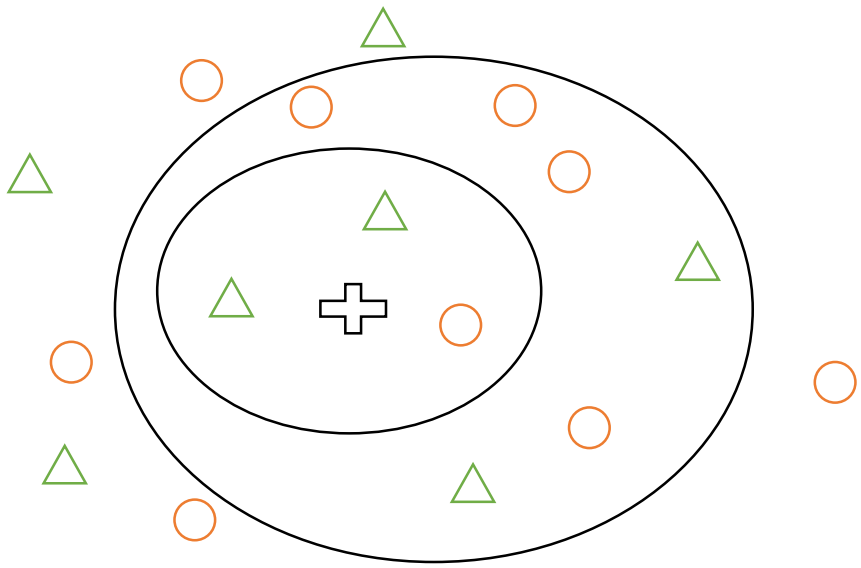
در ابتدای امر میبایست داده آموزشی و تست جدا شده باشند .

باید برچسب ها و گروه های مشخص شده برای داده های آموزشی جداشده مشخص بشود .

فاصله داده ی تست وارد شده را تا هر یک ار داده های آموزشی باید بدست آورد .

بر اساس K تعیین شده باید K نمونه ی آموزشی که نزدیک ترین فاصله را تا داده ی تست دارند ، مشخص شود .

در بین k نزدیک ترین همسایه به داده ی تست آن رده ای که آراء بیشتری را داشته باشد به عنوان برچسب نهایی داده ی تست مشخص خواهد شد . (شکل ۳)



(شکل ۳ مثالی از الگوریتم knn با وجود دو رده ی مثلث و دایره و داده ی تست+)

همانطور که در شکل ۳ مشخص شده است اگر فرض بشود که داده های آموزشی به دو رده ی مثلث و دایره تقسیم بشود . با فرض آن که K عدد ۳ باشد باید سه داده ی تستی را که از باقی فاصله ی کمتری را دارد در یک گروه قرار داد . که در این حالت با توجه به این که رده ی مثلث دارای اعضای بیشتری در این گروه دارد پس برچسب نهایی داده ی تست وارد شده ، مثلث خواهد بود .

از مطالب بالا میتوان نتیجه گرفت دو امر مهم در این الگوریتم بدست آوردن فاصله و تعیین مقدار K خواهد بود . [2]

برای بدست آوردن فاصله میتوان از روش های اقلیدسی و منهتن استفاده کرد .[2]

در روش اقلیدسی داریم ؛ اگر فرض شود به ازای داده ی آموزشی q با ویژگی های

$q_1,\ldots,q_n$ داده ی تست p وجود داشته باشد که دارای ویژگی های $p_1,\ldots,p_n$ است ، آنگاه خواهیم داشت [2]:

$$d(x,y) = \sqrt{\sum_{i=1}^{n} (p_{i-q_i})^2}$$

و اما در روش منهتن داریم [2]:

$$d(x,y) = |q_1 - p_1| + |q_2 - p_2| + \cdots + |p_n - q_n|$$

در تعیین مقدار K نیز باید به مواردی توجه داشت[2] :

اگر K خیلی کوچک باشد :

حساسیت به وجود داده های نویزی بیشتر خواهد شد .

نیاز به داده های آموزشی کمتر داریم .

امکان تشخیص دقیق تر جزئیات فضای راه حل .

اگر K خیلی بزرگ باشد [2] :

حساسیت به وجود داده های نویزی کمتر خواهد بود .

امکان تشخیص بهتر کلاس های گسسته

ممکن است یک همسایگی نقاطی از سایر کلاس هارا نیز شامل شود .

و درحالت کلی باید تعیین مقدار K تعادل برقرا شود و اگر مقدار K بسیار بزرگ باشد و اندازه داده آموزشی بسیار زیاد باشد ؛ الگوریتم KNN به روش بیزین بسیار نزدیک میشود.[2]

- **مزایا و معایب KNN :**

از مزایای این روش میتوان به آموزش بسیار سریع و از دست نرفتن اطلاعات و همچنین یادگیری توابع پیچیده بسادگی نام برد .[2]

از معایب این روش نیز باید به نیاز به حافظه زیاد ، تاثیر پذیری از ویژگی های غیر مهم در داده ها و زمان اجرای زیاد در هنگام پردازش نمونه های جدید اشاره کرد .[2]

- **Normalization داده ها**

نرمالیزه کردن و یا هنجار سازی داده ها به معنی این است که داده ها در یک بازه ی مشخصی تعریف شوند . به طور مثال یعنی داده ها بین دو مقدار 1, 0 **تعریف کنیم . از این کار صرفا برای راحت تر شدن کار استفاده میشود .** [2]

این کار به روش های مختلفی انجام میگیرد . یک روش مرسوم در این کار استفاده از فرمول Min_max **است :**

$$\acute{x}_t = \frac{x_{t-\min(x)}}{\max(x) - \min(x)}$$

- **کد MATLAB الگوریتم KNN**

در این برنامه سعی شده مجموعه داده ی Iris **را مورد پردازش و دسته بندی قرار دهیم .**

مجموعه داده ی iris **حاوی ۱۵۰ رکورد از داده است . هر رکورد حاوی ۴ ویژگی است پس به عبارتی** iris dataset **یک مجموعه داده حاوی ۱۵۰ تا سطر و ۵ ستون است (۴ ستون ویژگی + ۱ ستون کلاس که رده یا دسته ی هر داده را مشخص میکند)**[20]

```
clc ;
clear all;
close all;
%%%%%%%%%%%%%%%%%%%%%%%%%%%% normalization
load iris_data.txt;
dataset = iris_data;
for i=1:4
    Min = min(dataset(:,i));
    Max = max(dataset(:,i));
    dataset(:,i)=(dataset(:,i)-min)/(Max-Min);
end

```

(شکل ۴)

در این تیکه کد در ابتدا داده های iris را در ماتریس dataset قرار میدهیم و بعد توسط حلقه ی for داده های را با فرمول Min_Max نرمالایز میکنیم .(شکل ۴)

```
[row,col]=size(dataset);
number_of_class=3 ;
k=3;

train= zeros(floor(0.7*row),5); %%%%%105 dade
test = zeros(floor(0.3*row),5); %%%%%45 dade

```

(شکل ۵)

حال باید مقدار K را تعیین کنیم و در این مثال بنا به خواست خودمان ۷۰ درصد از کل داده ها را در مجموعه train به عنوان داده های آموزشی و ۳۰ درصد را در مجموعه test به عنوان داده های تست و یا ورودی قرار میدهیم .(شکل ۵)

```
m2 =row;
for i=1:floor(0.7*row)
    r=(randi(m2-1)+1)
    train(i,:)=dataset(r,:);
    dataset(r,:)=[];
    m2=m2-1;
end

test =dataset;
```

(شکل ۶)

در این مرحله باید از بین داده های اصلی ۷۰ درصد آن را یعنی ۱۰۵ داده را به شکل رندوم داخل دیتاست train قرار دهیم و هر رکورد را که وارد مجموعه train میشود از مجموعه دیتا ست حذف میشود .

در نهایت هر چه که در مجموعه dataset باقی مانده بود را به عنوان داده ی تست به مجموعه تست انتقال دهیم . (شکل ۶)

```
[TrainCnt,ee]=size(train);
[TestCnt,n]= size(test);
predictClass = zeros(TestCnt,1);

for i=1:TestCnt
    dist=zeros(TrainCnt,1);
    dis2=zeros(TrainCnt,2);
    for j=1:TrainCnt
        dist(j,1)=sqrt(((test(i,1)-train(j,1))^2)+((test(i,2)-train(j,2))^2) + ((test(i,3)-train(j,3))^2) +((test(i,4)-train(j,4))^2))
        dis2(j,1)= dist(j,1);
        dis2(j,2)=j;
    end
    dist=sort(dist);
    CntofClass = zeros(number_of_class,1);
```

(شکل ۷)

در این قسمت از برنامه به ترتیب فاصله ی هر داده ی تست را تا تک تک داده های train باید بدست آوریم و در ستون اول جدول dist و dist2 قرار میدهیم . در نهایت هم باید ماتریس distرا مرتب کنیم .

ماتریس dist : ماتریسی با یک ستون و به تعداد داده های train نیز سطر داریم.

ماتریس dist2 : ماتریسی با دو ستون و به تعداد داده های آموزشی سطر دارد . در ستون اول فاصله و در ستون دوم رده ی مشخص شده مقابلش قرار میگیرد . (شکل ۷)

```
    dist=sort(dist);
    CntofClass = zeros(number_of_class,1);

    for z=1:k
        Nearest = find(dis2(:,1)==dist(Z));
        ClsN = train(Nearest,5);
        CntofClass(ClsN)=CntofClass(ClsN)+1;
    end

    class =find(CntofClass==max(CntofClass));
    predictClass(i)=class;
end
```

(شکل ۸)

و در نهایت میبایست K نمونه ای را که کمترین فاصله را از داده ی تست وارد شده دارند را بررسی کرد و کلاس با تعداد نمونه ی بیشتر را به آن نسبت دهند . (شکل۸)

- **روش KNN مبتنی بر وزن فاصله**

برای بهبود عملکرد روش KNN **وزنی را برای هر** K **همسایگی در نظر میگیریم. این وزن بر اساس فاصله دیتا های آموزشی تا داده ی تست ، مورد بررسی ایجاد میشود . یعنی فاصله معکوس وزن خواهد بود .** [2]

$$\hat{f}(x_q) \leftarrow \underset{v \in V}{\operatorname{argmax}} \sum_{i=1}^{k} w_i \delta(v, f(x_i)), \text{ where}: w_i = \frac{1}{d(x_q, x_i)^2}$$

پس میتوان نتیجه گرفت هر چه قدر فاصله کمتر باشد وزن آن داده بیشتر خواهد بود .

این الگوریتم در حل مسائل مختلفی برای استنتاج استقرایی بکار میرود و همچنین این روش در مسئله ها که نویز بیشتری دارند مقاوم تر است و در مواردی که حجم داده آموزشی بالا باشد بسیار مفید تر خواهد بود .[2]

و به عنوان پیشنهاد میتوان در این روش برای ویژگی های موثر تر وزن بالاتری را در نظر بگیریم که اهمیت آن هارا در الگوریتم مشاهده بکنیم .[2]

- **کاربرد های الگوریتم KNN**

یکی از کاربرد های این الگوریتم، در پژوهشی تحت عنوان پیشبینی عوامل موثر بر بقای ۵ ساله ی کلیه ی پیوندی در داده های نامتوازن . در این تحقیق سعی شده با استفاده از مدل های الگوریتم ژنتیک و الگوریتم KNN بررسی شود تا چه عواملی در بقای کلیه های پیوندی به مدت ۵ شال بعد از پیوند تاثیر گذار است و پس از پژوهش ها و مدل سازی ها به این نتیجه رسیدند که بقای کلیه ی

پیوندی در روش پیش نمونه برداری با دقت ۹۶,۸ در صد و در روش زیر نمونه برداری با دقت ۸۹,۶ درصد پیش بینی شده است همچنین عواملی چون وزن و سن دهنده و گیرنده ، اوره ی قبل از پیوند ، کراتین قبل از پیوند و ... تاثیر مهمی دارد . [4]

- **الگوریتم دسته بندی کننده ی بیز ساده (Naïve bayes)**

استدلال بیزین روشی بر پایه ی احتمالات است و جز یادگیری های مبتنی بر نظریه احتمال است . اساس این روش به این صورت است که یرای هر مقداری یک توزیع احتمالی وجود دارد که با مشاهده یک داده جدید و استنتاج در مورد احتمالات آن ، میتوان تصمیمات بهینه ای در خصوص آن اتخاذ کرد . [2]

این روش ، بسیار کارا و مفید خواهد بود و در کار های مختلفی همچون دسته بندی متن میتواند مفید واقع شود .

براساس تحقیقات صورت گرفته ،روش یادگیری بیزین ، با روش های کارایی مثل درخت تصمیم و شبکه عصبی میتواند قابل مقایسه باشد . [7,1]

نگرش یادگیری بیزین به شکل زیر است :

۱) دانش موجود در خصوص مسئله ای را توسط احتمالات خیره میکنیم .

- **برای اینکار ، مقادیر کیفی به شکل توزیع احتمال مدل میشود . این مدل دارای پارامتر های ناشناخته ای خواهد بود .**

- **برای هر یک از مقادیر ناشناخته ،توزیع احتمال پیشینی در نظر گرفته میشود . که بازگو کننده باور ما به محتمل بودن هر یک از این مقادیر ، بدون مشاهده نمونه های داده است .** [2]

۲) داده را جمع آوری کنیم .

۳) با مشاهده داده ها ، مقدار توزیع احتمال پسین را محاسبه میکنیم . [2]

۴) به کمک این احتمال پسین ، در خصوص نمونه های جدید ، پیش بینی انجام میدهیم. [2]

- **ویژگی های یادگیری بیزین**

یکی از ویژگی های یادگیری بیزین یادگیری تدریجی است به صورتیکه هر نمونه به صورت جزئی میتواند احتمال درست بودن یک فرضیه را افزایش و یا کاهش داد .

فرضیه های حاصل از روش های بیزین ، قادر به پیشبینی احتمالی هستند .

داده های جدید را میتوان با ترکیب وزنی چندین فرضیه دسته بندی نمود .[2]

- **نظریه ی بیزین**

در یادگیری ماشین در این نظریه ، در بین مجموعه از نظریه های H ما به دنبال بهترین فرضیه هستیم که در مورد داده ی آموزشی Dصدق کند . [2]

که برای تعیین بهترین فرضیه یک روش این است که به دنبال محتمل ترین فرضیه باشیم. محتمل ترین فرضیه فرضیه ای ست که با داشتن داده ی آموزشی D و احتمال قبلی در مورد فرضیه های مختلف میتوان انتظار داشت . [2]

اگر فرض بشود که فضای فرضیه ی Hو مجموعه داده آموزشی D وجود داشته باشد داریم ؛یادگیری بیزین بر مبنای نظریه احتمال بیزین است. این نظریه ، امکان محاسبه احتمال پسین را بر مبنای احتمالات اولیه می آورد[2] **:**

$$P(h|D) = \frac{P(h)P(D|h)}{P(D)}$$

P(h) = احتمال پیشینی که فرضیه h قبل از مشاهده مثال آموزشی D ، داشته است (احتمال پیشین) در صورت نبود این احتمال میتوان به تمام فرضیه ها مقدار یکسانی را نسبت داد.

P(D) = احتمال پیشین مشاهده داده آموزشی D است.

P(D|h) = احتمال مشاهده داده آموزشی Dبه فرض آنکه فرضیه h صادق باشد . [2]

در این یادگیر ماشین هدف تعیین مقدار P(h|D) است . یعنی احتمال اینکه با مشاهده داده آموزشی D فرضیه h صادق باشد . به این رابطه احتمال پسین (Posterior Probability)**.**

احتمال پیشین مستقل از داده آموزشی و احتمال پسین ، متأثر از داده آموزشی است .

در رابطه ی بالا دیده میشود که هر چه قدر مخرج بزرگتر شود مقدار P(h|D) **نیز کاهش می یابد و این به این معناست که هر چه احتمال مشاهده** D **مستقل از فرضیه** h **بیشتر باشد ، به این معنا خواهد بود که شواهد کمتری در حمایت از فرضیه** h **در اختیار داریم .**

- **فرضیه با حداکثر احتمال پسین (Maximum A Posteriori Probability)**

تعین محتمل ترین فرضیه بر اساس حداکثر احتمال پسین (MAP)را از بین مجموعه فرضیات ممکن H به شکل زیر است :

$$h_{max} = arg \max_{h \in H} \frac{P(D|h)P(h)}{p(D)}$$

چون P(D) مستقل از h است میتوان مخرج را حذف کنیم .

$$h_{max} = arg \max_{h \in H} P(D|h)P(h)$$

- **فرضیه با بیشینه احتمال** (Maximum Likelihood)

اگر هیچ اطلاعاتی در مورد P(h) **وجود نداشته باشد؛ یا به عبارتی** P(h) **حذف شود میتوان فرض کرد که تمام فرضیه ها دارای احتمال پیشین یکسانی هست .** [2]

در اینصورت میتوان فقط مقدار P(D|h) **را برای فرضیه با حداکثر احتمال در نظر گرفت .**

این مقدار میزان احتماال داده D **نسبت به فرض** h**نامیده میشود و هر فرضیه ای که این مقدار را بیشینه کند** Maximum Likelihood (ML) **نامیده میشود :**

$$h_{ML} = arg \max_{h \in H} P(D|h)$$

- **Naïve Bayes Classifier**

این روش در مسائلی کاربردی است که نمونه Xتوسط ترکیب عطفی ویژگی های قابل توصیف بوده و این ویژگی ها به صورت شرطی و مستقل از یکدیگر باشند . تابع هدف f(x) بتواند هر مقداری را از مجموعه محدود V داشته باشد و همچنین مجموعه مثال های آموزشی نسبتا زیادی در دست باشد .[2]

تابع هدف زیر را در نظر بگیرید f :X → V که در آن هر نمونه x توسط بردار ویژگی به صورت مقابل توصیف میشود .(a_n ,..., a_2, a_1) . هدف از این تابع این است که به ازای یک نمونه ی جدید رده ی آن نمونه مشخص شود . در روش بیزین، برای حل مسیله محتمل ترین مقدار هدف، V_{MAP} ، محاسبه میشود .[2]

$$V_{MAP} = arg \max_{v_j \in V} P(V_j | a_1, \dots, a_n)$$

حال طبق نظریه ی بیزین داریم :

$$V_{MAP} = arg \max_{v_j \in V} \frac{P(a_1, \dots, a_n | V_j) P(V_j)}{P(a_1, \dots, a_n)}$$

$$= argmax P(a_1, \dots, a_n | V_j) P(V_j)$$

حال با استفاده از داده‌های آموزشی سعی می‌کنیم دو جمله ی معادله ی بالا را تخمین بزنیم. محاسبه از روی داده‌های آموزشی به این صورت که میزان تکرار V_j در داده‌ها چقدر است، آسان می‌باشد. اما

محاسبه ی جملات مختلف $P(a_1, \dots, a_n | V_j)$ به این صورت قابل قبول نخواهد بود مگر اینکه حجم بسیار بسیار زیادی از داده‌های آموزشی در اختیار داشته باشیم.[2]

مشکل اینجاست که تعداد این جملات برابر تعداد نمونه‌های ممکن ضرب در تعداد مقادیر تابع هدف می‌باشد .

$$V_{NB} = argmax\, P(V_j) \prod_{i=1}^{n} P(a_i | V_j)$$

- یک مثال ساده از این الگوریتم

فرض کنیم مجموعه ای از داده هارا داریم برای این که برر سی کنیم آیا شرایط برای انجام بازی در روز های مشخص شده مناسب هست یا خیر ؟[2]

در این مثال ۱۴ نمونه ویا روز مورد بررسی قرار گرفته است . که هر مورد دارای ۴ ویژگی آب و هوایی است و برای هر روز با برچسب yes/No مشخص شده که آیا اون روز برای ورزش تنیس مناسب هست یا نه . و قرار است در انتها به ازای نمونه ی داده ای جدید بررسی شود که آیا بازی تنیس در آن روز میتواند انجام شود و یا خیر ؟(شکل ۹)

day	outlook	Tempreture	Humidity	wind	Play tennis
day1	sunny	Hot	High	Weak	NO
day2	Sunny	Hot	High	Strong	NO
day3	Overcast	Hot	High	Weak	YES
day4	Rain	Mild	High	Weak	YES
day5	Rain	Cool	Normal	Weak	YES
Day6	Rain	Cool	Normal	Strong	NO
Day1	Overcast	Cool	Normal	Strong	YES
Day8	Sunny	Mild	High	Weak	NO
day9	Sunny	Cool	Normal	Weak	YES
day10	Rain	Mild	Normal	Weak	YES
day11	Sunny	Mild	Normal	Strong	YES
day12	Overcast	Mild	High	Strong	YES
Day13	Overcast	Cool	Normal	Weak	YES
Day14	Rain	Mild	High	Strong	NO

(شکل ۹ بررسی وضعیت بازی تنیس در روز های مختلف)

همان طور که در رابطه بیزین مطرح شد ما نیاز داریم:

به احتمال هر کدام از رده ها (که در اینجا باید احتمال دو رده ی Yes/No **بررسی شود)**

به علاوه احتمال هر کدام از ویژگی ها نسبت به رده ها .

Outlook	Temperature	Humidity	Windy	Class
sunny	hot	high	false	N
sunny	hot	high	true	N
overcast	hot	high	false	P
rain	mild	high	false	P
rain	cool	normal	false	P
rain	cool	normal	true	N
overcast	cool	normal	true	P
sunny	mild	high	false	N
sunny	cool	normal	false	P
rain	mild	normal	false	P
sunny	mild	normal	true	P
overcast	mild	high	true	P
overcast	hot	normal	false	P
rain	mild	high	true	N

شکل بالا وضعیت بهبود یافته ی داده ها هست . در این حالت احتمال Positive بودن ((P(P)) برابر 9/14 و احتمال Negative بودن (p(N)) برابر 5/14 است .

برای هر یک از ویژگی های نیز این محاسبات را داریم :

Outlook	
P(sunny|p)=2/9	P(sunny|n)= 3/5
P(overcast|p)=4/9	P(overcast|n)=0
P(rain|p)=3/9	P(rain|n)=2/5
Temperature	
P(hot|p)=2/9	P(hot|n)=2/5
P(mild|p)=4/9	P(mild|n)=2/5
P(cool|p)=3/9	P(cool|n)=1/5
Humidity	
P(high|p)=3/9	P(high|n)=2/5
P(normal|p)=6/9	P(normal|n)=2/5
windy	
P(true|p)=3/9	P(true|n)=3/5
P(false|p)=6/9	P(false|n)=2/5

(جدول ۱)

حال میخواهیم با این روش مثال زیر را رده بندی کنیم :

X:(outl=sunny,temp=cool,Hum=high,wind=strong)

با اعمال رابطه بیزین داررریم :

$$V_{NB} = arg \max_{v_k \in [yes,no]} P(V_j) \prod_{i=1}^{n} P(a_i|V_j)$$

$$= arg \max_{v_k \in [yes,no]} \begin{matrix} P(v_k)P(outl = sunny|v_k)P(tmp = cool|v_k) \\ P(hum = high|v_k)P(wind = strong|v_k) \end{matrix}$$

برای محاسبه این احتمال باید ۱۰ مقدار احتمال دیگر را تخمین بزنیم . زیرا ۴ ویژگی و احتمال یک رده برابر است با ۵ محاسبه و چون دو رده داریم در نهایت ۱۰ محاسبه داریم.

و محاسبه نهایی به شکل زیر است .

$$P(yes)P(sunny|yes)P(cool|yes)P(high|yes)P(strong|yes) = 0.0054$$

$$P(no)P(sunny|no)P(cool|no)P(high|no)P(strong|no) = 0.0206$$

و همان طور که میبینیم با داشتن چنین شرایطی احتمال انجام شدن بازی تنیس در آن آب و هوا منفی هست .

- **مزایا و معایب این دسته بندی**

مزایا :

این الگوریتم سرعت بسیار خوبی دارد . از این الگوریتم میتوان برای پیشبینی در بیش از دو دسته استفاده کرد .این الگوریتم میتواند نسبت به الگوریتم رگرسیون لجستیک بهتر عمل کند البته با وجود شرط استقلال پیشامد ها . این الگوریتم برای حجم داده ی کوچک بسیار مناسب عمل میکند . نیاز به حجم داده ی بالایی نیست . [6]

معایب :

اگر دسته و رده ای را داشته باشیم که در نمونه تصادفی حضور نداشته باشد ، دسته بند بیز ساده احتمال آن را صفر محاسبه میکند و هیچ عضوی را در این گروه قرار نخواهد داد .

اگر متغیرهای پیشگو دارای وابستگی باشند، دسته‌بند بیز به خوبی عمل نمی‌کند. در دنیای واقعی بیشتر مواقع با چنین متغیرهای مواجه هستیم.[6]

- **کاربرد های الگوریتم بیز ساده**

پیش بینی بلادرنگ (Real Time Prediction) : **به علت سرعت بالای این الگوریتم و علاقه آن به یادگیری در سییستم های بلا درنگ میتواند به خوبی عمل کند .**

دسته بندی متن (تشخیص هرزنامه) : بسیاری از شرکت ها برای طبقه بندی اسناد در پست الکترونیکی خود از این الگوریتم استفاده میکنند .

سیستم های توصیه گر : یکی دیگر از کاربرد های الگوریتم بیز ساده در این سیستم ها است به گونه ای که با زیر نظر گرفتن یک مشتری و بررسی فعالیت های قبلی آن میتوان انتخاب های آتی را به او پیشنهاد کرد . [6]

- **الگوریتم K-Means**

این الگوریتم از جمله الگوریتم های بدون نظارت و در دسته ی یادگیری خوشه یابی قرار میگیرد . در فرآیند خوشه یابی میتوانیم نمونه ها را بر اساس میزان شباهتشان به یک دیگر گرهبندی کنیم . در واقع نمونه های در یک گروه شباهت زیادی باهم داشته باشند ولی اعضای گروه های مختلف دارای حداقل میزان شباهت به یک دیگر را داشته باشند . فی الواقع در این شیوه شباهت با فاصله رابطه عکس دارد.

اما در این مباحث یک ابهام ذاتی وجود دارد و ان به این معناست که نمیتوان به شکل دقیق با مشاهده یک سری از داده ها مشخص کرد که به چند دسته باید تقسیم شوند که این سوال با توجه به کاربرد در مسئله قابل پاسخگویی خواهد بود .

از جمله کاربرد های خوشه یابی میتوان در فرآیند های بازیابی اطلاعات وب ، داده کاوی ، بیوانفورماتیک ، بازاریابی ، مدیریت و سنجش از راه دور و .... اشاره کرد .

قبل از بررسی این الگوریتم باید به تعدادی تعریف توجه بکنیم .

فاصله(Distance) **: میزان عدم شباهت بین نمونه های داده را گویند .**

مرکز خوشه (Cluster Centroid)**: یک نمونه داده ای که مقادیر هر یک از ویژگیهای آن برابر است با متوسط مقادیر نمونه های مختلف داده در ان ویژگی ها .**

- **انواع روش های خوشه یابی**

روش های متعددی برای خوشه بندی وجود دارد ولی دو نمونه ی مهم آن از این قرار است [2]:

روش های مبتنی بر افراز (Partitioning):

بر اساس ایجاد افراز های مختلف و ارزیابی آنها بر اساس شاخصی همچون مجموع مربع خطاها ؛ از جمله روش های آن K-means، K-medoids

روش های سلسله مراتبی (Hierarchical):

ایجاد تفکیک های سلسله مراتبی از مجموعه داده بر اساس شاخصی معین ؛ و از جمله روش های آن AGNES، DIANA

- **شیوه کارکرد الگوریتم K-means**

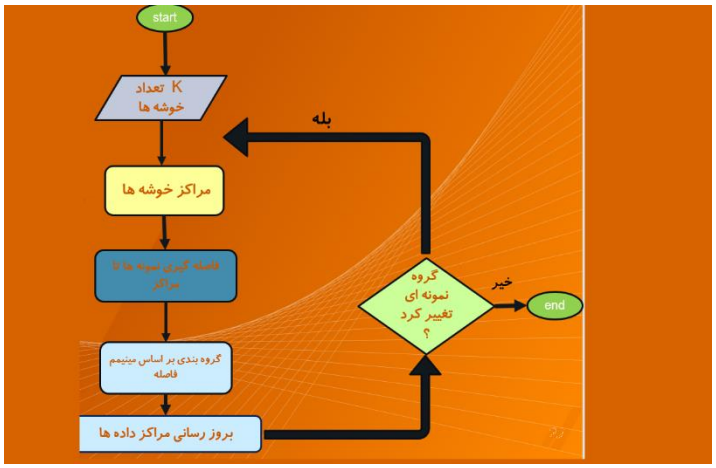
این الگوریتم اولین بار توسط MacQueen **در سال ۱۹۶۷ معرفی شد. و مبنای کلی آن فاصله تا مراکز خوشه ها است .** [2]

در این الگوریتم درابتدا باید بر اساس K **و تعداد رده ها مرکز خوشه مشخص کنیم این مراکز خوشه میتوانند به شکل رندوم از بین داده های ورودی بدست آیند و یا به شکل رندوم ذو داده به عنوان مرکز خوشه معرفی بکنیم .**

در مرحله بعد داده ها ی تست باید به ترتیب و جدا جدا وارد شوند ، سپس فاصله هر یک از آن ها را تا مراکز خوشه باید جدا جدا بدست آید و فاصله با هر کدام که کمتر بود داده ی مربوطه به آن دسته تعلق دارد .

پس از آن که تمام داده ها گروه بندی شدند لازم است که مراکز خوشه به روز شوند برای این کار باید از تمام ویژگی های اعضای هر یک از خوشه ها میانگین گیر شود .

پس از میانگین گیری باید دوباره فاصله ها اندازه گیری شوند ؛ اگر طی این مرحله از اندازه گیری فواصل داده ای وجد نداشت که از خوشه ای به خوشه دیگری انتقال پیدا کرده باشد ؛ میتوان عملیات را خاتمه داد در غیر این صورت باید مراحل گفته شده در بالا مجددا تکرار شود .(شکل ۱۰)



(شکل ۱۰ فلوچارت الگوریتم K-means)

- **مزایا و معایب این روش**

از مزایای این روش میتوان گفت ؛کار آمدی نسبی اجرا با هزینه $O(tkn)$ **که در آن** n**تعداد نمونه ها ، K برابر با تعداد خوشه ها و t برابر با تعداد دفعات تکرار است و معمولا** $k,t<<n$ **.**

و از معایب آن هم میتوان اشاره کرد ، نیاز به تعیین تعداد خوشه ها (K) داریم . در برخورد با داده های نویزی و پرت مشکل داریم . در تشخیص خوشه های غیر محدب ناتوانیم و امکان گیر کردن در کمینه های محلی وجود دارد .

روالی مشخص برای محاسبه ی اولیه ی مراکز خوشه ها وجود ندارد همچنین علی رقم اینکه خاتمه پذیری الگوریتم بالا تضمین شده است ولی جواب نهایی آن واحد نبوده و جواب نهایی به انتخاب خوشه های اولیه بستگی دارد . [2]

- **تفکیک تصویر با استفاده از K-means**

یکی از کاربرد های اساسی این الگوریتم در تفکیک تصویر و در قسمت بنید تصاویر توسط این الگوریتم هست . که کد متلب آن را نمایش خواهیم داد .

```
clc;
clear all;
close all;

%%%%%%%%%%%%%%%%%%%% Definition Matrix %%%%%%%%
%%
x1=imread('2.jpg');

x=imresize(x1,[300,300]);

%%
%%%%%%%%%%%%%%%%%%%%%%random weight%%%%%%%%%%%%%

imshow(x

x=double(x);
 [m,n]=size(x);
len=m;
```

(شکل ۱۱)

در این تیکه کد سعی داریم تا عکسی را بخوانیم و ابعاد عکس را نرمال کردیم و در متغیر x قرار میدهیم به گونه ای که این متغیر یک ماتریس حاوی ۳۰۰ سطر و ستون است ، که به این شکل خواهد

بود . (شکل ۱۱و۱۲)

	1	2	3	4	5	6	7	8	9
1	5	97	232	232	235	238	238	239	24
2	7	138	247	243	244	245	247	250	25
3	10	142	248	247	251	252	251	253	25
4	10	147	253	249	253	253	253	254	25
5	12	145	253	250	252	254	254	254	25
6	13	144	255	250	253	254	252	253	25
7	14	146	255	249	252	252	251	253	25
8	12	144	255	250	249	248	250	253	25
9	14	143	253	249	250	248	251	253	25
10	13	145	254	250	252	251	252	254	25
11	12	145	253	249	249	250	251	252	25
12	13	143	250	247	248	249	248	250	24
13	12	147	250	244	245	247	249	251	24
14	15	146	251	245	248	248	251	251	25
15	13	147	250	246	250	252	251	251	25
16	9	149	251	246	249	250	251	252	25
17	14	150	249	245	248	249	250	251	25
18	12	152	249	245	247	249	249	251	25

(شکل ۱۲ متغیر x)

حال باید مقدار Kرا تعیین کنیم و مراکز خوشه ها را از بین داده ها به شکل رندوم بدست آوریم . در این قسمت متغیر هایی مثل m (برای قرار گیری مراکز خوشه ها) و متغیر mm (برای قرار دادن تعداد اعضای هر خوشه و جمع مقادیر هر خوشه) نیز تعریف شده است . (شکل۱۳)

```
k=2;
m=zeros(k,1);
mm=zeros(k,2);


for i=1:k
    m(i,1)=x(randi(len-1),randi(len-1));
end
```

(شکل ۱۳)

در این تیکه از برنامه اول با دادن بعد سومی به متغیر x این امکان را پیدا میکنیم که برای هر خانه بتوانیم شماره خوشه و یا رده ی تعیین شده اش را هم قرار دهیم و سپس متغیر ای با نام temp تعریف میشود که یک سطر و دو ستون دارد که بعدا در اندازه گیری فواصل به کمکمان می آید .(شکل ۱۴)

```
x(:,:,3)=0;
x2=x;

temp=zeros(1,k);
s=0;
gamha=0;
```

(شکل ۱۴)

در این قسمت سعی میکنیم فاصله ی هر یک از داده هارا تا دو مرکز خوشه به دست آوریم. در اینجا از شیوه فصله منهتن استفاده شده است . سپس توسط تابع find() کمترین فاصله را بدست می آوریم . و در نهایت بعد سوم متغیر x را با شماره ی کلاس پر میکنیم .(شکل ۱۵)

```
while s==0
    gamha=gamha+1;
    for i=1 :len
        for i2=1 :len
        for j=1:k
            temp(j)=sqrt(abs(m(j,1)-x(i,i2))^2);
        end
        z1=find(temp==min(temp));
        x(i,i2,3)=z1(1);
        end
    end
```

(شکل ۱۵)

در این تیکه کد سعی میکنیم با استفاده ا میانگین گیر یمراکز خوشه را به روز رسانی کنیم و مراکز به روز شده را در ماتریس m قرار میدهیم . (شکل ۱۶)

```
%%%%%%%%%%%%%%%%mohasebe jame har cluster va tedad an %%%%%%%%%%
for i=1:len
    for i2=1:len
        mm(x(i,i2,3),1)=mm(x(i,i2,3),1)+x(i,i2);
        %calculet the numeber of pixle in each cluster
        mm(x(i,i2,3),2)=mm(x(i,i2,3),2)+1;
    end
end
for l=1:k
    m(l,1)=mm(l,1)/mm(l,2);
end
```

(شکل ۱۶)

حال باید چک کنیم ببینیم آیا حلقه به پایان رسیده است یا نه ؟که این کار را با چک کردن x و x2 انجام میدهیم. X2 را پیش از این تعریف کرده بودیم و به عبارتی مقدار x را در آن ذخیره کردیم . (شکل ۱۷)

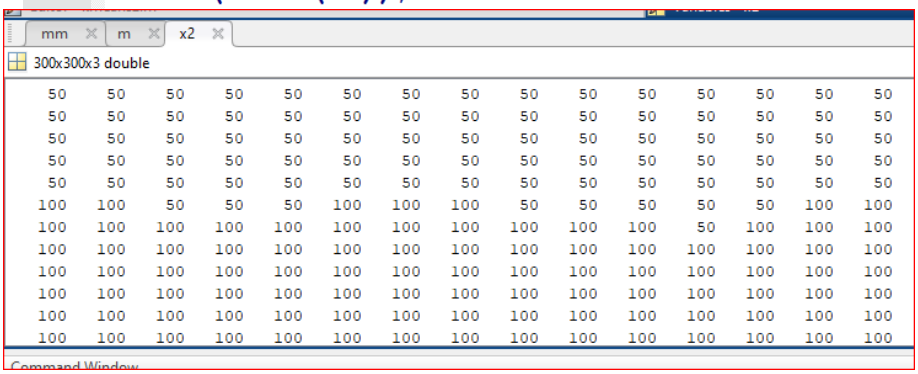
```
        if(x==x2)
            s=1;
        else
            x2=x;
        end

end
```

(شکل ۱۷)

در این قسمت با توجه به رده ی مشخص شده برا یهر خانه در X باید مقادیر آنهارا تغییر دهیم و در نهایت در x2 قرار میدهیم ؛ تا بشود بر روی کیفیت تصویر نهایی تغیرر اعمال شده را مشاهده کرد . در شکل ۱۹ وضعیت ماتریس x2 را مشاهده میکنید . (شکل ۱۸)

(شکل ۱۸)

```
for i=1:len
    for i2=1:len
        x2(i,i2)=x(i,i2,3)*50;
    end
end
x2(:,:,2:3)=[];
figure
imshow(uint8(x2));
```



(شکل ۱۹)

و حال تصویر ۲۰ نشان دهنده عکس خام ، قبل از اعمال تغییرات است و همچنین عکس ۲۱ نیز تغییرات را بعد از دسته بندی نمایش میدهد و همان طور که میبینید تضاویر به دو قسمت افراز شده اند ، قسمت های روشن تر در عکس اصلی بعد از اعمال الگوریتم تیره تر و قسمت های تیره بعد از تغییرا ت روشن تر شده اند .

(شکل ۲۰)

(شکل ۲۱)

شناسایی اشیا

تشخیص الگو

تشخیص الگو شاخه‌ای از مبحث یادگیری ماشینی است. می‌توان گفت تشخیص الگو، دریافت داده‌های خام و تصمیم گیری بر اساس دسته‌بندی داده‌ها است. بیشتر تحقیقات در زمینه‌ی تشخیص الگو در رابطه با «یادگیری نظارت شده» یا «یادگیری بدون نظارت» است. روش‌های تشخیص الگو، الگوهای مورد نظر را از یک مجموعه داده‌ها با استفاده از دانش قبلی در مورد الگوها یا اطلاعات آماری داده‌ها، جداسازی می‌کند. الگوهایی که با این روش دسته‌بندی می‌شوند، گروه‌هایی از اندازه‌گیری‌ها یا مشاهدات هستند که نقاط معینی را در یک فضای چند بعدی تشکیل می‌دهند. این ویژگی اختلاف عمده‌ی تشخیص الگو با تطبیق الگو است، که در آنجا الگوها با استفاده از موارد کاملا دقیق و معین و بر اساس یک الگوی مشخص، تشخیص داده می‌شوند. تشخیص الگو و تطبیق الگو از بخش‌های اصلی مبحث پردازش تصویر به خصوص در زمینه‌ی بینایی ماشین هستند.[7,8]

انواع تشخیص الگو

نیاز به سیستم‌های اطلاعاتی بهبود یافته بیشتر از قبل مورد توجه قرار گرفته‌است چرا که اطلاعات، عنصری اساسی در تصمیم سازی است و جهان در حال افزایش دادن مقدار اطلاعات در فرم‌های مختلف با درجه‌هایی از پیچیدگی است. یکی از مسائل اصلی در طراحی سیستم‌های اطلاعاتی مدرن، تشخیص الگو به‌طور اتوماتیک است. تشخیص به عنوان یک صفت اصلی انسان بودن است. یک الگو، توصیفی از یک شی است. یک انسان دارای یک سیستم اطلاعاتی سطح بالاست که یک دلیل آن داشتن قابلیت تشخیص الگوی پیشرفته است. بر طبق طبیعت الگوهای مورد تشخیص، عملیات تشخیص به دو گونه‌ی اصلی تقسیم می‌شوند[7,9] :

تشخیص آیتم‌های واقعی

این ممکن است به عنوان تشخیص سنسوری معرفی شود که تشخیص الگوهای سمعی و بصری را دربر می‌گیرد.

تشخیص الگوهای زمانی و فضایی

این فرآیند تشخیص، شناسایی و دسته‌بندی الگوهای فضایی و الگوهای زمانی را دربر می‌گیرد. مثال‌هایی از الگوهای فضایی کارکترها، اثر انگشت‌ها، اشیاء فیزیکی و تصاویر هستند. الگوهای زمانی شامل فرم‌های موجی گفتار، سری‌های زمانی و ... هستند.

به طور کلی طراحی یک سیستم تشخیص الگو چندین مسئله‌ی اصلی را دربر می‌گیرد:

۱) طریقه نمایش داده‌ها ۲) استخراج ویژگی ۳) تعیین رویه‌ی تصمیم بهینه

اول از همه، ما بایستی در مورد نمایش داده‌های ورودی تصمیم بگیریم. دومین مسئله در تشخیص الگو، استخراج ویژگی‌ها یا صفات خاصی از داده‌های ورودی دریافت شده و کاهش ابعاد بردارهای الگوست. این مورد اغلب به عنوان مسئله‌ی پیش پردازش و استخراج ویژگی معرفی می‌شود. عناصر ویژگی‌های اینتراست که برای همه‌ی کلاس‌های الگو مشترک هستند می‌توانند حذف شوند. اگر یک مجموعه‌ی کامل از ویژگی‌های تشخیص برای هر کلاس از داده‌های اندازه‌گیری شده تعیین شود، تشخیص و دسته‌بندی الگوها، دشواری کمتری را دربر خواهد داشت. تشخیص اتوماتیک ممکن است به یک فرایند تطبیق ساده یا یک جدول جستجو کاهش یابد. به هر حال در بسیاری از مسائل تشخیص الگو، در عمل، تعیین یک مجموعه‌ی کامل از ویژگی‌های تشخیص اگر غیرممکن نباشد دشوار است. مسئله‌ی سوم در طراحی سیستم تشخیص الگو تعیین رویه‌های تصمیم بهینه است که در فرآیند شناسایی و دسته‌بندی مورد نیاز واقع می‌شود. پس از آنکه داده‌های مشاهده شده از الگوها جمع‌آوری شد و در فرم نقاط الگو یا بردارهای اندازه‌گیری در فضای الگو بیان شد، ما ماشینی را می‌خواهیم تا تصمیم بگیرد که این داده به کدام کلاس الگو تعلق دارد[7,9].

استخراج ویژگی

استخراج ویژگی (Feature extraction)**: فرآیندی است که در آن با انجام عملیاتی بر روی داده‌ها، ویژگی‌های بارز و تعیین‌کننده‌ی آن مشخص می‌شود.**

هدف استخراج ویژگی این است که داده‌های خام به شکل قابل استفاده‌تری برای پردازش‌های آماری بعدی درآیند و نیز کاهش دادن بیشتر داده‌ها به مجموعه‌ای از ویژگی‌هاست، که باید به اغتشاشاتی چون شرایط نورپردازی، موقعیت دوربین، نویز و اعوجاج ایمن باشند.

نمونه‌هایی از استخراج ویژگی عبارت‌اند از:

- **انجام آشکارسازی لبه.**
- **استخراج ویژگی‌های گوشه‌ای.**
- **استخراج تصاویر چرخش از نقشه‌های ژرفا.**
- **بدست آوردن خطوط تراز و احتمالاً گذر از صفرهای خمش.**

روش‌های مختلف استخراج ویژگی بنا به فلسفه‌ی پشت سرشان ممکن است یک یا چند کار زیر را انجام دهند:

1. **حذف نوفه‌ی داده‌ها**
2. **جداسازی اجزای مستقل داده‌ها**
3. **فروکاهی ابعاد برای تولید بازنمایی مختصرتر**
4. **افزایش بعد برای تولید بازنمایی جدایی‌پذیرتر**

برای اینکه از روی الگوهای یک تصویر هویت و یا خالق آن تصویر مشخص شود باید یک سری مشخصات عام و یا خاص از دل تصویر بیرون کشیده شود که به این کار استخراج ویژگی گفته می‌شود. به‌عنوان مثال در تشخیص امضاء به وسیله‌ی پردازش تصویر یک سری ویژگی ها (مانند شیب خط ها) از تصویر اسکن شده‌ی امضاء بیرون کشیده می‌شود، که به وسیله‌ی آن می‌توان صاحب امضاء را تشخیص داد[7,9].

الگوها و کلاس‌های الگوها

تشخیص الگو می‌تواند به‌عنوان دسته‌بندی داده‌های ورودی در کلاس‌های شناخته شده به‌وسیله‌ی استخراج ویژگی‌های مهم یا صفات داده تعریف شود. یک کلاس الگو، یک دسته‌ی متمایز شده به‌وسیله‌ی برخی صفات و ویژگی‌های مشترک است. ویژگی‌های یک کلاس الگو، صفات نوعی هستند که بین همه الگوهای متعلق به آن کلاس مشترک هستند. ویژگی‌هایی که تفاوت‌های بین کلاس‌های الگو را بیان می‌کنند اغلب به‌عنوان ویژگی‌های اینتراست شناخته می‌شوند. یک الگو، توصیفی از یکی از اعضای دسته است که ارائه دهنده‌ی کلاس الگو می‌باشد. برای راحتی، الگوها معمولا به‌وسیله‌ی یک بردار نمایش داده می‌شوند.

برای دسترسی آسان به ویژگی‌های تصویر، باید به‌جای استفاده از مقادیر پیکسل‌ها، با استفاده از روش-هایی خاص ویژگی‌های تصویر را استخراج و در قالبی نمایش دهیم که قابل مقایسه باشند. به این روش‌ها استخراج ویژگی می‌گویند. خروجی این روش‌های خاص عدد یا برداری است که یکی از ویژگی‌های تصویر را معرفی می‌کند، که به آنها کلاس ویژگی یا توصیف‌گر می‌گویند (مانند هیستوگرام رنگ). به برداری که از کنار هم قرار دادن این کلاس‌های ویژگی به‌دست می‌آید بردار ویژگی تصویر گفته می‌شود. با توجه به عملکرد و ساختار سامانه‌های بازیابی، مشخص است که کارایی آن به‌شدت تحت تاثیر ویژگی-های استخراجی از تصاویر می‌باشد پس می‌توان با استفاده از روش‌های متداول در بهبود کارایی ویژگی‌ها بهبود قابل توجهی در نتایج بازیابی به‌دست آورد. در سال‌های اخیر برای ایجاد روش‌های کارآمد در بازیابی تصویر بر پایه‌ی محتویات آن تلاش‌های زیادی شده‌است خصوصا موضوع ارتقای کارایی ویژگی‌ها شامل تطبیق ویژگی‌ها، بهبود ویژگی‌ها و گسسته‌سازی آنهاست.[7]

ویژگی‌های تصویر

ویژگی‌های تصویر را می‌توان از لحاظ نحوه‌ی توصیف تصویر به سه دسته تقسیم کرد:

ویژگی‌های عمومی یا سطح پایین:

این ویژگی‌ها معرف کلیات تصویر هستند و توانایی توصیف اشیا یا مفاهیم موجود در تصویر را ندارند مانند رنگ و بافت.

ویژگی‌های معنایی یا سطح بالا:

این ویژگی‌ها به بیان مفاهیم و اشیای موجود در تصویر می‌پردازند و غالبا از روش‌های قطعه‌بندی برای این منظور استفاده می‌کنند.

ویژگی‌های مختص کاربردهای خاص:

نظیر ویژگی‌های توصیفی چهره در سیستم‌های بازیابی تصاویر مربوط به چهره.[7,9].

استخراج ویژگی‌های شکلی

شکل یک شی، مشخصه‌های پیکربندی سطح آن است که بوسیله‌ی مرز آن نمایش داده می‌شود. تشخیص شکل یکی از روش‌هایی است که به‌واسطه‌ی آن درک انسان از محیط انجام می‌گیرد. این امر در بازیابی تصاویر مبتنی بر محتوا بسیار مهم می‌باشد چرا که معادل ناحیه‌ی مورد نظر در تصاویر است. در پردازش تصویر شکل تصویر باینری شامل مرز اشیا بوده که پس از قطعه‌بندی به‌دست می‌آید.

در سیستم‌های بازیابی تصویر مبتنی بر محتوا که برای حوزه‌ای خاص همچون ابزاری جهت علامت‌گذاری تجاری طراحی شده‌اند، قطعه‌بندی شکل می‌تواند به‌صورت خودکار و موثر صورت گیرد اما این روش‌ها مورد مناسبی برای بازیابی تصویر مبتنی بر محتوای دارای پایگاه داده‌ی نامتجانس نیستند که در این مورد قطعه‌بندی شکل ممکن است مشکل و در برخی نقاط غیرممکن باشد. ویژگی‌های شکل بر طبق تکنیک-های بکار رفته دسته‌بندی می‌شوند که مبتنی بر مرز و ناحیه هستند. تکنیک‌های مبتنی بر مرز ناحیه‌ی شکل را با استفاده از بازنمایی مشخص خارجی آن مرز توصیف می‌کنند، برای مثال پیکسل‌های در راستای مرز اشیا جزو این دسته از ویژگی‌ها هستند.

بنابراین روش‌های مبتنی بر شکل به دو دسته تقسیم می‌شوند:

1- **روش اول بر مبنای لبه**
2- **روش دوم بر مبنای ناحیه می باشد.**

در روش‌هایی که بر مبنای ناحیه هستند تنها از جزئیات لبه و پیرامون اشیا استفاده می‌کنند و داخل شکل به‌طور کلی نادیده گرفته می‌شود، در حالی‌که در روش‌هایی که بر مبنای ناحیه هستند به محتوای داخلی شکل اهمیت داده می‌شود. توصیف‌گر فوریه و گشتاورهای ثابت دو نمونه‌ی موفق از این دو دسته به حساب می‌آیند. شناسایی اشیای عمومی همه منظوره به دلیل دانش محدود ما بسیار مشکل است زیرا متاسفانه شکل یک شی بستگی زیادی به زاویه‌ی دید و فاصله دارد. اما شکل حاوی اطلاعات مفیدی می‌باشد که باید به نحوی مورد استخراج قرار گیرد. در سیستم‌های بازیابی تصویر مبتنی بر محتوا، استفاده از شکل توسعه‌ی کمتری نسبت به کاربرد رنگ و بافت داشته‌است زیرا استخراج و نمایش شکل نسبت به رنگ و بافت پیچیده‌تر می‌باشد. دلیل دیگر اینست که مدل‌های ریاضی که درک انسان از اشکال مختلف را به نحو مناسب مدل می‌کنند، وجود ندارد و بعلاوه اکثر بازنمایی‌های شکل نسبت به تبدیلات هندسی مانند دوران، بزرگ‌نمایی و انتقال بی‌تفاوت نمی‌باشند. بی‌تفاوتی ویژگی‌های استخراجی نسبت به تبدیلات هندسی شاخص مهمی در انتخاب آنهاست زیرا اشیای مشابه در تصویرهای متفاوت اغلب تغییر می‌کنند. [9]

توصیف‌گرهای ساده‌ی مبتنی بر مرز شکل شامل مساحت، محیط، تراکم، خروج از مرکز، فاصله‌ی زاویه‌ای و جهت‌یابی هستند و توصیف‌گرهای پیچیده مبتنی بر مرز شکل، شامل توصیف‌گر فوریه، توصیف‌گر شبکه‌ای، کدهای زنجیره‌ای و گشتاورهای آماری است. می‌توان دلیل ایجاد این سیستم‌ها را این‌طور بیان کرد که به‌علت تغییر نورپردزای و رنگ در تصویر، روش‌های مبتنی بر رنگ و رنگی-مکانی قادر به تشخیص شباهت بین دو تصویر نیستند. بدین جهت بود که رده‌ی دیگری از سیستم‌های CBIR **و مستقل از رنگ و نورپردازی که به آن‌ها سیستم‌های مبتنی بر شکل اطلاق می‌شوند، ایجاد شدند.**

طرح اساسی در سیستم‌های مبتنی بر شکل، به‌دست آوردن تخمینی از شکل اشیای موجود در تصویر است که بر اساس این اطلاعات، بردار ویژگی تصویر ساخته می‌شود. تکنیک‌های نمایه‌سازی مبتنی بر شکل را می‌توان به دو بخش عمده تقسیم‌بندی کرد:

دسته‌ی اول، تکنیک‌هایی که برای ساخت بردار ویژگی نیاز به قطعه‌بندی تصویر دارند و دسته‌ی دوم، تکنیک‌هایی که بدون قطعه‌بندی با استخراج پاره‌ای از فاکتورهای شکلی از تصویر، بردار ویژگی را می‌سازند. روش‌های مبتنی بر قطعه‌بندی تصویر ابتدا ناحیه‌ی همگن تصویر را شناسایی می‌کنند و سپس بردار ویژگی از روی اطلاعات مربوط به نواحی نظیر مساحت نواحی و یا مرزهای نواحی تصویر استخراج می‌گردند. اما به دو علت حیطه‌ی کاربرد آن‌ها محدود است:

اول اینکه قطعه‌بندی دقیق تصویر و تشخیص شی اصلی در تصویر، سربار محاسباتی زیادی دارد که کارایی سیستم را به شدت کاهش می‌دهد، دوم اینکه قطعه‌بندی برای تصاویر فرکانس بالا امکان‌پذیر نیست.

دسته‌ی دوم تکنیک‌های مبتنی بر شکل هستند که سعی دارند با توجه به ماهیت زمان‌گیر بودن کار قطعه-بندی، بدون آن پاره‌ای از فاکتورهای شکلی را از تصویر استخراج و بردار ویژگی تصویر را از روی آن بسازند. یکی از روش‌های مبتنی بر شکل که بدون استفاده از قطعه‌بندی تصویر عمل می‌کند، روش مبتنی بر لبه می‌باشد. مطلوبیت استفاده از لبه‌ها به دو علت است:

اولا حجم لبه‌های تصویر نسبت به حجم پیکسل‌های تصویر بسیار کم بوده بنابراین به‌علت کاهش اطلاعات تحت پردازش می‌توان سیستمی کارا با استفاده از لبه‌ها ساخت، ثانیا لبه‌ها به‌علت استفاده از همبستگی میان پیکس‌ها حجم زیادی از اطلاعات تصویر را دربر دارند و همچنین از آنجایی که از دیدگاه درک انسانی نیز اشکال بواسطه‌ی لبه‌هایشان از یکدیگر تمییز داده می‌شوند پس می‌توان سیستمی دقیق با استفاده از لبه‌ها ساخت.[7]

در طی عمل بازیابی تصویر مبتنی بر ویژگی شکل لازم است سه مسئله حل گردد:

از آنجایی‌که معمولا شکل به طور خاص به شی داخل تصویر وابسته است بنابراین ویژگی معنایی شکل نسبت به بافت بسیار قویتر است. برای به دست آوردن پارامترهای واقعی شکل، باید در ابتدا تصویر را قطعه‌بندی کرد، پس به این ترتیب قطعه‌بندی تصویر تاثیر بسیاری بر ویژگی شکل دارد.

توصیف واقعی در واقع یک مسئله‌ی پیچیده است چرا که حواس انسان ترکیبی از دریافت شبکیه‌ای و داشتن زندگی واقعی اوست. بنابراین نمی‌توان یک مدل که با حس غیرواقعی انسان منطبق باشد را یافت.

در زوایای مختلف دید انسان‌ها تفاوت بسیار زیادی وجود دارد. به‌منظور تطابق دقیق شکل نیاز به نگهداری ثابت تصویر در طی حرکت بدون دور درشت‌نمایی و دوران است که این نیز خود یک مسئله است.

آشکارساز لبه

هدف آشکارسازی لبه نشان‌گذاری نقاطی از یک تصویر است که در آنها شدت روشنایی به تندی تغییر می‌کند. تغییرات تند در خصوصیات تصویر معمولاً نماینده‌ی رویدادهای مهم و تغییرات در خصوصیات محیط هستند. شناسایی لبه یک محدوده‌ی تحقیقاتی در پردازش تصویر و استخراج ویژگی است.

لبه‌ها ممکن است وابسته به دیدگاه باشند – یعنی می‌توانند با تغییر نقطه‌ی دید تغییر کنند، و نوعاً هندسه-ی صحنه‌ی اجسامی که جلوی همدیگر را گرفته‌اند و مانند آن را نشان می‌دهند یا ممکن است نابسته به دیدگاه باشند – که معمولاً نمایانگر ویژگی‌های اجسام دیده شده همچون نشان‌گذاری‌ها و شکل سطح باشند. در دو بعد و بالاتر مفهوم تصویر باید در نظر گرفته شود.

یک لبه‌ی نوعی ممکن است (برای نمونه) مرز میان یک بخش قرمزرنگ و یک بخش سیاه رنگ باشد، حال آن‌که یک خط می‌تواند تعداد کمی پیکسل‌های ناهم‌رنگ در یک زمینه‌ی یکنواخت باشد. در هر سوی خط یک لبه وجود خواهد داشت. لبه‌ها نقش مهمی در کاربردهای پردازش تصویر دارند.

لبه، مرز بین نواحی با خواص نسبتاً متفاوت سطح خاکستری است. نظریه‌ی پایه در بیشتر روش‌های آشکارسازی لبه، محاسبه یک عملگر مشتق محلی است. در این مقطع توجه شود که لبه (گذر از تاریک به روشن) به صورت یک تغییر آرام، نه سریع، سطح خاکستری مدل می‌شود. این مدل نشان می‌دهد که معمولاً لبه‌های تصاویر رقمی بر اثر نمونه‌برداری، کمی مات می‌شوند.

مشتق اول مقطع سطح خاکستری در لبه‌ی جلویی گذر، مثبت است، در لبه‌ی عقبی آن منفی است و همان طور که مورد انتظار است، در نواحی با سطح خاکستری ثابت صفر است. مشتق دوم برای قسمتی از گذر که در طرف تیره‌ی لبه است، مثبت است، برای قسمت دیگر گذر که در طرف روشن لبه است، منفی است، و در نواحی با سطح خاکستری ثابت، صفر است. بنابراین، از بزرگی مشتق اول می‌توان برای تعیین این که آیا پیکسل در روی لبه قرار دارد، استفاده کرد. مشتق دوم در نقطه‌ی وسطی هر گذر سطح خاکستری یک عبور از صفر دارد. عبور از صفرها راهی قوی برای تعیین محل لبه‌های تصویر فراهم می‌آورند. اندازه‌ی مشتق اول تصویر در هر نقطه برابر بزرگی گرادیان است. مشتق دوم نیز با استفاده از لاپلاسین به دست می‌آید.

اگر یک لبه را به عنوان تغییر در شدت روشنایی که در طول چند پیکسل دیده می‌شود در نظر بگیریم، الگوریتم‌های آشکارسازی لبه به‌طور کلی مشتقی از این تغییر شدت روشنایی را محاسبه می‌کنند.

محاسبه‌ی مشتق اول

تعداد زیادی از عملگرهای آشکارسازی لبه بر پایه‌ی مشتق اول شدت روشنایی کار می‌کنند، یعنی با گرادیان شدت روشنایی داده‌های اصلی سروکار داریم.

محاسبه‌ی مشتق دوم

برخی دیگر از الگوریتم‌های آشکارسازی لبه بر اساس مشتق دوم شدت روشنایی کار می‌کنند که در واقع نرخ تغییرات گرادیان شدت روشنایی است و برای آشکارسازی خط‌ها بهترین است، زیرا هر خط یک لبه‌ی دوگانه است، بنابراین در یک سوی خط یک گرادیان روشنایی و در سوی دیگر گرادیان مخالف آن دیده می‌شود. پس می‌توانیم منتظر تغییر بسیار زیاد در گرادیان شدت روشنایی در محل یک خط باشیم. برای یافتن خط‌ها می‌توانیم گذر از صفرهای تغییر گرادیان را در نتایج جستجو کنیم[7,10].

عملگرهای آشکارسازی لبه

- مرتبه‌ی نخست: رابرتز، پرویت، سوبل، کنی، اسپیسک
- مرتبه‌ی دوم: لاپلاسین، مار–هیلدرث

اکنون، عملگر کنی و پس از آن مار–هیلدرث بیشترین کاربرد را دارند. عملگرهای زیادی تاکنون منتشر شده‌اند اما هیچ‌یک برتری قابل ملاحظه‌ای بر عملگر کنی در شرایط کلی نداشته‌اند.

لبه در واقع مرز میان یک شکل (Object) و پس‌زمینه (Background) می‌باشد که برای پردازش تصاویر باید لبه‌ها مشخص شوند تا محدوده‌ی تصویر تعیین گردد.

در عملیات لبه‌برداری ورودی یک تصویر به فرمت intensity **می‌باشد و در خروجی تصویر** binary **داده می‌شود، که در تصویر حاصل مرزهای بیرونی تصویر به صورت** 1 **و مرزهای داخل به صورت** 0 **نشان داده می‌شود.**

```
I=rgb2gray(i1);
Bw=edge(I,'sobel')
```

Edge **لبه‌ها را در تصاویر** intensity **پیدا می‌کند، این تابع یک تصویر باینری یا** intensity **را به عنوان ورودی می‌گیرد و یک تصویر باینری** bw **به همان اندازه‌ی تصویر اولی برمی‌گرداند، که جاهایی که تابع لبه‌ها را در تصویر پیدا می‌کند، در تصویر خروجی** 1 **می‌کند و جاهایی دیگر را** 0 **قرار می‌دهد**[7,10].

برخی از الگوریتم‌های لبه‌یابی

- **الگوریتم** soble
- **الگوریتم** canny
- **الگوریتم** Roberts
- **الگوریتم** Prewitt
- **الگوریتم** zero-cross

الگوریتم Sobel

این متد لبه‌ها را با استفاده از تخمین زدن مشتق پیدا می‌کند، که لبه‌ها را در آن نقاطی بر می‌گرداند که گرادیان تصویر I ، max **است.**

```
Bw= edge(I,'sobel',thresh)
```

مقدار thresh **یک میزان آستانه را برای این متد مشخص می‌کند.**

این تابع (edge) **را از همه‌ی لبه‌هایی که قوی‌تر (بیشتر) از** thresh **نیستند چشم پوشی می‌کند و اگر ما مقدار این** thresh **را مشخص نکنیم یا اگر** thresh **خالی باشد** {[]} **، تابع** edge **خود به‌طور اتوماتیک مقداری را انتخاب می‌کند.**

Bw=edge(I,'sobel',thresh,direction)

در این directin , syntax **جهت را مشخص می‌کند، یعنی رشته‌ای است که مشخص می‌کند که این تابع لبه‌های افقی یا عمودی و یا هر دو را جستجو کند که به‌طور پیش فرض هر دو را جستجو می‌کند.**

(horizontal) **افقی**:

(Vertical) **عمودی**:

Bw=edge(I,'sobel',…,options)

در این دستور تابع یک رشته‌ی اختیاری به‌عنوان ورودی می‌گیرد که رشته‌ی nothinning **سرعت عملیات الگوریتم را بالا می برد، به این علت که در مرحله‌ی نازک شدن لبه‌های اضافی می‌گذرد (می‌پرد) و اگر رشته‌ی** nothinning **را انتخاب کنیم، الگوریتم لبه‌های نازک شده را نیز درخواست می‌کند.**

[Bw,thresh]=edge(I,'sobel',…)

این دستور، مقدار آستانه را برمی‌گرداند.

[Bw,thresh,gv,gh]=edge(….)

در این دستور، لبه‌های افقی و عمودی (gv,gh) **را با توجه به عملگرهای گرادیان برمی‌گرداند.**

آستانه‌گیری

تشخیص آستانه (thresholding) **یکی از اولین مراحلِ تعداد زیادی از برنامه‌های پردازش تصویر است که در این مرحله با پیدا کردن یک آستانه، زمینه‌ی عکس از اجسام جدا می‌گردد.**

هنگامی که مشتق را حساب کردیم، گام بعدی اعمال یک آستانه برای کشف نقاطی که بخشی از یک لبه هستند می‌باشد. هر چه آستانه کمتر باشد، خط‌های بیشتری آشکارسازی می‌گردند و نتایج نسبت به نویز و ویژگی‌های نامرتبط تصویر بیشتر حساس می‌شوند، از سوی دیگر یک آستانه‌ی زیاد ممکن است خط‌های ضعیف یا بخش‌هایی از خط‌ها را از دست بدهد.

یک مصالحه‌ی معمول، آستانه‌گیری با پسماند است. این روش از چندین آستانه برای جستجوی لبه‌ها بهره می‌جوید. با آستانه‌ی بالایی جستجو را برای پیدا کردن ابتدای خط‌ها آغاز می‌کنیم. هنگامی که یک نقطه‌ی آغاز داریم، مسیر لبه را درون تصویر پیکسل به پیکسل با نشانه‌گذاری پیکسل‌هایی که از آستانه‌ی پایینی بالاترند پی می‌گیریم و تنها هنگامی که مقدار از آستانه‌ی پایینی پایین‌تر رود آن را پایان می‌دهیم. این رهیافت بر اساس این گمان است که لبه‌ها به احتمال زیاد در مسیرهای پیوسته قرار دارند و دنبال کردن بخش ضعیفی از لبه‌ای که از پیش دیده‌ایم ممکن می‌کند، بدون آنکه پیکسل‌های نویزی را به‌عنوان لبه نشانه‌گذاری کنیم[11].

مقدمه‌ای بر شبکه‌ی عصبی

شبکه های عصبی مصنوعی با الهام از شبکه های عصبی مغزی ایجاد شده اند. کاملترین شبکه های مغزی در پستانداران و به ویژه انسان یافت می شوند. به همین دلیل شناخت ساختار پایه ای شبکه های عصبی مغزی و آ شنایی با نحوه ی عملکرد آن، کمک شایانی به درک بهتر مفاهیم مرتبط با شبکه های ع صبی مصنوعی خواهد نمود.

شبکه های عصبی مصنوعی یا Artificial Neural Networks **الگوهایی برای پردازش اطلاعات هستند که با تقلید از شبکه ی عصبی مغزی انسان ساخته شده اند، سیستم پردازش اطلاعات در شبکه های عصبی**

مصنوعی دارای ساختاری جدید است. این شبکه ها از تعداد زیادی عناصر پردازنده ی متصل به هم تشکیل یافته اند، که به صورت هماهنگ با یکدیگر کار می کنند. درک بهتر چگونگی عملکرد شبکه های عصبی مصنوعی نیازمند آشنایی با شبکه ی عصبی مغزی و عملکرد سلولهای آن است.

ایده‌ی شبکه‌ی عصبی مصنوعی با تحقیق درباره‌ی سلول‌های مغزی انسان شروع شد. مغز بشر از میلیونها نورون عصبی منحصر به‌فرد تشکیل شده است و این رشته‌های عصبی به اشکال و اندازه‌های مختلف تغییر می‌کنند[7].

عصب چیست؟

بافت‌هایی که عصب نامیده می شوند، اجتماعی از نورون‌ها هستند که اطلاعات و پیام‌های الکتروشیمیایی را از یک قسمت بدن به قسمت دیگر منتقل می‌کنند. در اینجا برای برانگیزش سلول لازم است تا آشفتگی از حدی مشخص که آستانه یا سرحد نامیده می‌شود تجاوز کند.

پس از برانگیزش، نورون‌ها پیام‌هایی برای نورون‌های دیگر می‌فرستند، و خود را برای دریافت اطلاعات و پاسخ به اطلاعات دریافتی آماده می‌کنند. مغز قادر است برای شناسایی الگوها و تفکیک الگوهای ناقص، آموزش ببیند. به‌علاوه، شبکه‌ی مغز با از دست دادن برخی از نورون‌ها، عملکرد خود را از دست نخواهد داد.

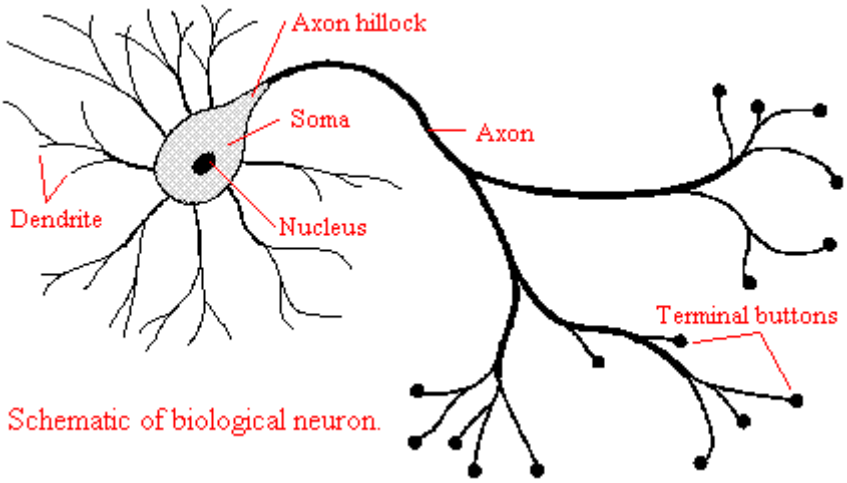
از دیدگاه کلاسیک، نورون یک واحد تخمین‌گر ساده است که سیگنال‌ها را از طریق راه‌های ورودی رشته‌ای شکل به‌نام دندریت (کلمه‌ای یونانی به معنی درختچه)، دریافت و ترکیب می‌کند. مغز به-عنوان یک سیستم پردازش اطلاعات از صد میلیارد نورون به‌هم مرتبط با ده میلیون میلیارد ارتباط تشکیل شده‌است[7].

سیناپس چیست؟

محل تلاقی آکسون یک سلول به دندریت‌های سلول دیگر را سیناپس (کلمه‌ای یونانی به معنی تماس) می‌نامند. یک سیناپس مرکب از پایه‌های پیش سیناپ سی، شکاف سیناپ سی، و پایه‌های پ سا سیناپ سی می‌باشد.

سیناپس‌ها واحدهای ساختاری و عملیاتی هستند که ارتباط بین نرون‌ها را برقرار می سازند. سیناپس‌ها انواع مختلفی دارند که یکی از مهمترین آنها سیناپس شیمیایی است. در سیناپس‌ها آکسون یک سلول به دندریت سلول بعدی متصل نیست، ولی فاصله‌ی آنها آنقدر کم است که می‌توانند اطلاعاتی را مبادله کنند.

زمانی که یک پتانسیل تحریک به انتهای یک آکسون می‌رسد، موجب آزاد شدن یک ماده‌ی شیمیایی به‌نام پیام‌رسان نورونی از انتهای آکسون می‌شود، و پس از نفوذ در سیناپس‌ها، دندریت نرون دیگری را تحریک می‌کند. این ارتباط عامل اصلی پردازش اطلاعات در مغز انسان است.



نمایی از یک نرون

همان‌طور که در ابتدا گفته شد، شبکه‌های عصبی مصنوعی با تقلید از شبکه‌های عصبی مغزی ساخته شده-اند، و اجزای تشکیل دهنده‌ی شبکه‌های عصبی مغزی در شبکه‌های عصبی مصنوعی نیز یافت می‌شوند. این اجزا به‌نحوی شایسته باهم در ارتباطند، هر نرون در شبکه از فعالیت سایر نرون‌ها متاثر می‌باشد، و

به همین دلیل، ویژگی‌های منحصر به‌فرد پردازش مغزی را دارا می‌باشد. همین قابلیت باعث می‌شود هنگامی که شبکه‌ی عصبی مصنوعی در شرایط جدید قرار بگیرند، پاسخ مناسبی ارایه دهند.

سه مشخصه‌ی مشترک میان شبکه‌های عصبی مصنوعی و شبکه‌های مغزی عبارتند از:

ساختار پایه‌ای شبکه‌ی عصبی از تعداد زیادی واحد پردازنده به اسم نرون تشکیل شده‌است.
این ساختارهای پردازنده به طول وسیعی به‌هم متصل می‌باشند.
نرون‌ها به‌صورت دسته‌جمعی فعالیت می‌کنند.

چرا از شبکه های عصبی استفاده می کنیم؟

شبکه‌های عصبی با توانایی قابل توجه خود در استنتاج نتایج از داده‌های پیچیده می‌توانند در استخراج الگوها و شناسایی گرایش‌های مختلفی که شناسایی آن‌ها برای انسان‌ها و کامپیوتر بسیار دشوار است استفاده شوند. از مزایای شبکه‌های عصبی می‌توان موارد زیر را نام برد:

1. یادگیری تطبیقی: توانایی یادگیری یعنی اینکه چگونه وظایف خود را بر اساس اطلاعات داده شده به آن و یا تجارب اولیه انجام دهد که در واقع آن‌را اصلاح شبکه گویند.

2. خودسازماندهی: یک شبکه‌ی عصبی مصنوعی بتواند به‌صورت خودکار سازماندهی و ارائه داده‌هایی که در طول آموزش دریافت کرده را انجام دهد. نورون‌ها با قاعده‌ی یادگیری سازگار شده و لذا پاسخ به ورودی تغییر می‌یابد.

3. عملگرهای بی‌درنگ: محاسبات در شبکه‌ی عصبی مصنوعی می‌تواند به‌صورت موازی و به وسیله‌ی سخت‌افزارهای مخصوصی که طراحی و ساخت آن برای دریافت نتایج بهینه‌ی قابلیت‌های شبکه عصبی مصنوعی است، انجام شود.

4. تحمل خطا: با ایجاد خرابی در شبکه مقداری از کارایی کاهش می‌یابد ولی برخی امکانات آن با وجود مشکلات بزرگ همچنان حفظ می‌شود.

5. دسته‌بندی: شبکه‌های عصبی قادر به دسته‌بندی ورودی‌ها برای دریافت خروجی مناسب می‌باشند.

6. تعمیم‌دهی: این خاصیت شبکه را قادر می‌سازد تا تنها با برخورد با تعداد محدودی نمونه، یک قانون کلی از آن را به دست آورده، نتایج این آموخته‌ها را به موارد مشاهده از قبل نیز تعمیم دهد. توانایی‌ای که در صورت نبود آن سامانه باید بی‌نهایت واقعیت‌ها و روابط را به خاطر بسپارد.

7. پایداری- انعطاف پذیری: یک شبکه‌ی عصبی هم به حد کافی پایدار است تا اطلاعات فراگرفته‌ی خود را حفظ کند و هم قابلیت انعطاف و تطبیق را دارد و بدون از دست دادن اطلاعات قبلی می‌تواند موارد جدید را بپذیرد[7].

مقایسه‌ی اجمالی رایانه و شبکه‌ی عصبی

یک شبکه‌ی عصبی به‌طور کلی با یک کامپیوتر سنتی در موارد زیر تفاوت دارد:

1. شبکه‌های عصبی دستورات را به‌صورت سری اجرا نکرده، شامل حافظه‌ای برای نگهداری داده و دستورالعمل نیستند.

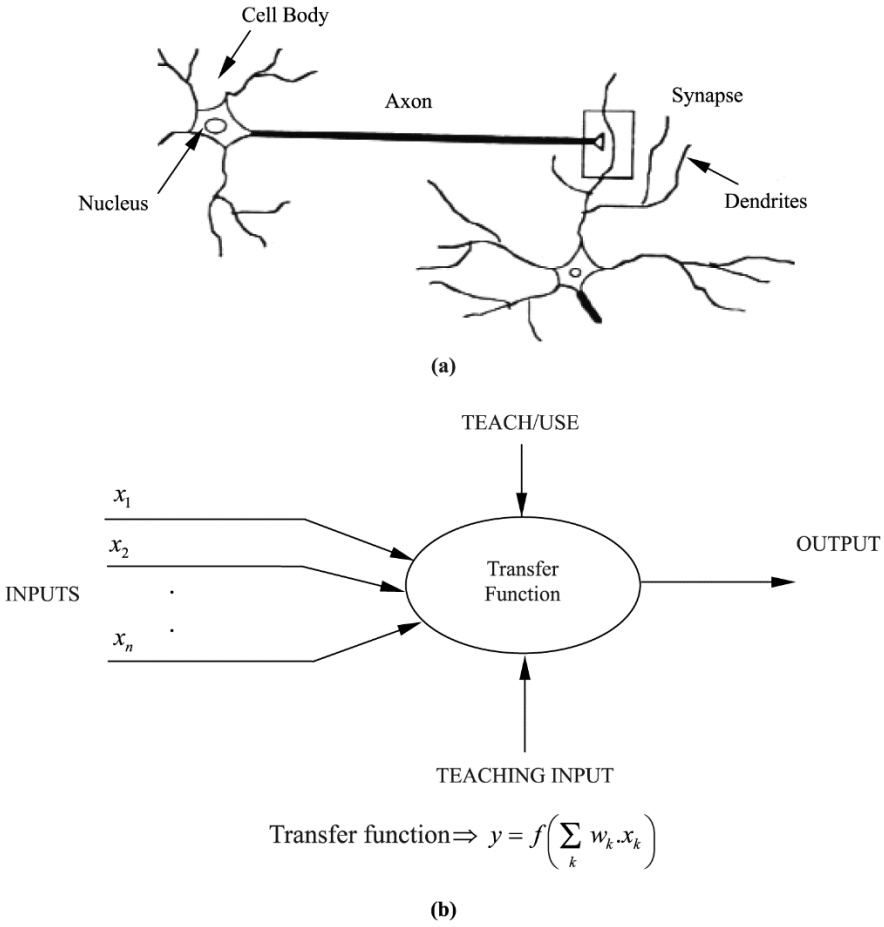
2. به مجموعه‌ای از ورودی‌ها به‌صورت موازی پاسخ می‌دهند.

3. بیشتر با تبدیلات و نگاشت‌ها سروکار دارند تا الگوریتم‌ها و روش‌ها.

4. شامل ابزار محاسباتی پیچیده نبوده، از تعداد زیادی ابزارساده که اغلب کمی بیشتر از یک جمع وزن‌دار را انجام می‌دهند تشکیل شده‌اند.

شبکه‌های عصبی شیوه‌ای متفاوت برای حل مسئله دارند. کامپیوترهای سنتی از شیوه‌ی الگوریتمی برای حل مسئله استفاده می‌کنند که برای حل مسئله مجموعه‌ای از دستورالعمل‌های بدون ابهام دنبال می‌شود. این دستورات به زبان سطح بالا و سپس به زبان ماشین که سامانه قادر به تشخیص آن می‌باشد تبدیل می‌شوند. اگر مراحلی که کامپیوتر برای حل مسئله باید طی کند از قبل شناخته شده نباشند و الگوریتم مشخصی وجود نداشته باشد، سامانه توانایی حل مسئله را ندارد. کامپیوترها می‌توانند خیلی سودمندتر

باشـــند اگر بتوانند کارهایی را که ما هیچ پیش‌زمینه‌ای از آنها نداریم انجام دهند. شـــبکه‌های عصـــبی و کامپیوترها نه تنها رقیب هم نیستند بلکه می‌توانند مکمل هم با شند. کارهایی وجود دارند که بهتر ا ست از روش الگوریتمی حل شوند و همین‌طور کارهایی وجود دارند که جز از طریق شبکه‌ی عصبی مصنوعی قابل حل نمی‌باشـــند و البته تعداد زیادی نیز برای بدســـت آوردن بازده حداکثر، از ترکیبی از روش‌های فوق استفاده می‌کنند. به‌طور معمول یک کامپیوتر سنتی برای نظارت بر شبکه‌ی عصبی استفاده می‌شود. شبکه‌های عصبی معجزه نمی‌کنند، اگر به‌طور محسوس استفاده شوند کارهای عجیبی انجام می‌دهند[7].



مقایسه‌ی نورون مصنوعی و نورون طبیعی

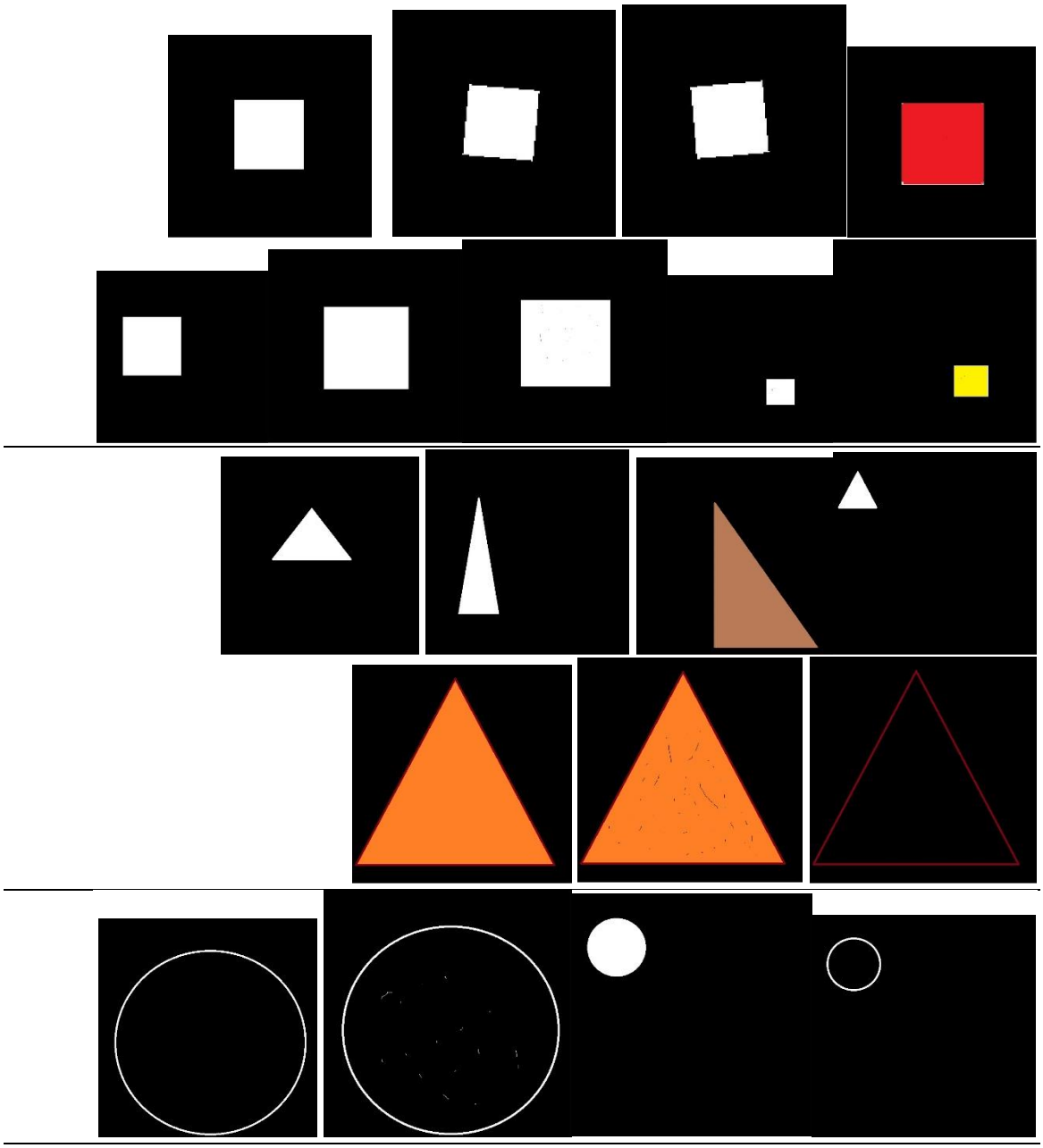
شبکه‌های عصبی اطلاعات را به روشی مشابه با کاری که مغز انسان انجام می‌دهد پردازش می‌کنند. آنها از تعداد زیادی از عناصر پردازشی (سلول عصبی) که فوق العاده بهم پیوسته‌اند تشکیل شده‌است که این عناصر به‌صورت موازی باهم برای حل یک مسئله‌ی مشخص کار می‌کنند. شبکه‌های عصبی با مثال کار می‌کنند و نمی‌توان آنها را برای انجام یک وظیفه‌ی خاص برنامه ریزی کرد. مثال‌ها می‌بایست با دقت انتخاب شوند در غیر این‌صورت زمان سودمند، تلف می‌شود و یا حتی بدتر از این شبکه ممکن است نادرست کار کند. امتیاز شبکه‌ی عصبی این است که خودش کشف می‌کند که چگونه مسئله را حل کند، عملکرد آن غیرقابل پیش‌گویی است.

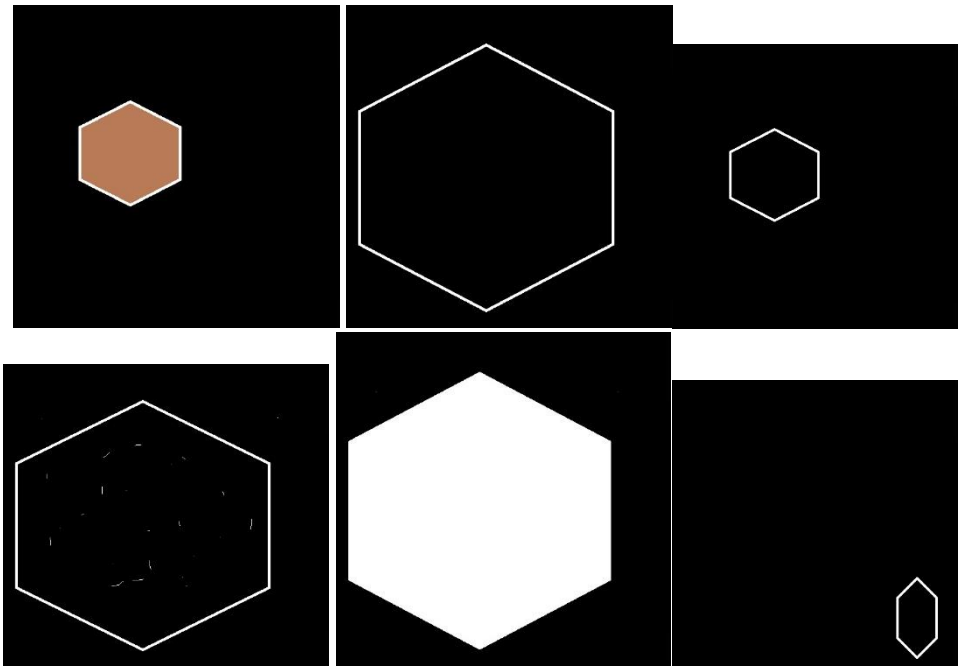
همان‌طوری که قبلا اشاره شد، شبکه‌ی عصبی مصنوعی الهام گرفته از شبکه‌ی عصبی انسان می‌باشد. جهت بیان ساده‌ی مراحل یادگیری، آموزش مثال‌هایی را ارائه می‌نماییم:

فرض می‌کنیم برای شناساندن یک کتاب به یک کودک، ده مدل کتاب را در اندازه‌ها و رنگ جلدهای متفاوت را نشان داده و آموزش می‌دهیم و انتظار داریم که با نشان دادن یک کتابی دیگر که متفاوت اما برگرفته از کتاب‌هایی که کودک قبلا دیده بدور از در نظر گرفتن عنوان و محتوای کتاب، آن را نیز به عنوان کتاب تشخیص دهد.

فرض می‌کنیم استادی یک ترم به دانشجویان آموزش (Train) **می‌دهد سپس جهت ارزیابی دانشجویان یک کوئیز از آن‌ها می‌گیرد تا هم دیدگاه استاد در مورد دانشجویان مشخص گردد و هم بفهمد که تدریسش کیفیت خوبی داشته است یا نه. درصورتی که نتیجه‌ی کلی ارزیابی ضعیف باشد دانشجویان را برای بازیادگیری آماده می‌نماید** (validation)**. در نهایت برای ارزیابی پایانی آزمون پایان ترم به‌عمل می‌آید** (test)**.**

هدف اصلی ما در این ارائه تشخیص اشکال هندسی می‌باشد و بدین منظور می‌بایستی دیتاستی متشکل از انواع اشکال هندسی در اندازه، رنگ و زوایای چرخشی متفاوت را آماده نماییم. لذا از محیط paint **جهت آماده‌سازی دیتاست استفاده نمودیم که در زیر نمونه‌هایی از آن‌ها را مشاهده می‌نمایید.**





بخشی ار دیتاست اشکال هندسی

در کل فرایند تشخیص اشکال در 5 گام قابل انجام می باشد:

پیش پردازش تصویر

استخراج ویژگی‌های اشکال

ساخت شبکه‌ی عصبی مناسب با ویژگی‌های استخراج شده

آموزش شبکه‌ی عصبی مربوطه

تست شبکه‌ی عصبی با نمونه های جدید

پیش‌پردازش تصویر

در این فاز تصاویر اولیه که کاملا به‌هم‌ریخته و دارای اطلاعات نامرتبط و غیرسودمندی هستند را برای یادگیری شبکه‌ی عصبی و استخراج ویژگی آماده سازی می‌کنیم. اشکال را نمی‌توان قبل از پیش‌پردازش

به شبکه‌ی عصبی آموزش داد، لذا برای اینکه به یک ورودی مناسب تبدیل شوند بایستی دنباله‌ای از یک سری از فرآیندها را تحت عنوان پیش‌پردازش طی کنند. این پیش پردازش شامل موارد مختلفی به شرح ذیل می‌شود:

- **حذف نویز**
- **جدا سازی تصویر از پس زمینه**
- **تبدیل تصویر به تصویر باینری**
- **استانداردسازی ابعاد همه‌ی تصاویر**
- **و...**

مراحل فوق توسط تابعی تحت عنوان getdata() **انجام می‌گیرد. با استفاده از حلقه‌ی** for **عمل خواندن تصاویر به شرح ذیل صورت می‌گیرد:**

```
for i2=1:20
if i2==6 || i2==11 || i2==16
j=1;
end
if i2<6
i=imread(['trainPic\a',num2str(j),'.jpg']);

elseif i2<11
i=imread(['trainPic\b',num2str(j),'.jpg']);
elseif i2<16
i=imread(['trainPic\c',num2str(j),'.jpg']);
```

سپس تصاویر خوانده شده با استفاده از دستور زیر از حالت رنگی به حالت سیاه و سفید تبدیل می‌گردد پس عمل سیاه و سفید سازی تصاویر ورودی یک عمل استاندارد سازی در مرحله‌ی عمل پیش‌پردازش می‌باشد.

```
a=rgb2gray(i);
```

سپس با استفاده از تابع removeback **تصویر یا ناحیه‌ی مفید را از تصویر پس‌زمینه جدا می‌کنیم یعنی نواحی غیرمفید حذف می‌گردد که این نیز یک عمل پیش‌پردازش به شرح ذیل می‌باشد.**

```
function [ x ] = removeBack( x )
```

```
x(:,1:5)=[];
x(:,end-5:end)=[];
x(1:5,:)=[];
x(end-5:end,:)=[];
while x(1,:)==0
x(1,:)=[];
end
while x(:,1)==0
x(:,1)=[];
end
while x(end,:)==0
x(end,:)=[];
end
while x(:,end)==0
x(:,end)=[];
end
x(:,1)=[];
x(:,end)=[];
x(1,:)=[];
x(end,:)=[]
```

استخراج ویژگی‌ها

از پارامترهای مختلفی می‌توان برای استخراج ویژگی استفاده کرد. مثلا برای تشخیص مربع از مثلث با یک ویژگی ساده‌ی تعداد اضلاع و یا با استفاده از تعداد زاویه‌های شکل می‌توان این کار را انجام داد.

مربع 4 ضلع دارد ولی مثلث 3 ضلع – برای زاویه نیز همین‌طور. اما برای تشخیص مربع از مستطیل دیگر این ویژگی‌ها مناسب نمی‌باشند و برای تفکیک این دو شکل (مربع و مستطیل) باید دنبال ویژگی‌های دیگری باشیم. پس باید دنباله‌ای از ویژگی‌ها را انتخاب کنیم تا با داشتن آن‌ها بتوانیم همه‌ی اشکال را از هم متمایز کنیم.

ویژگی هایی که ما در این ارائه انتخاب کردیم به‌شرح ذیل می باشند:

- **تعداد زاویه ها**
- **خروجی حاصل از الگوریتم لبه‌یابی**
- **تعداد اضلاع**

- حجم تصویر نسبت به کل شکل
- خروجی حاصل از الگوریتم hough transform
- و...

در مرحله‌ی بعدی با استفاده از تابع "canny" عمل لبه‌یابی صورت می‌گیرد که یکی از ویژگی‌های اشکال در شبکه‌ی عصبی می‌باشد. بعد از اینکه شکل لبه‌یابی شد، احتمال دارد شکل لبه‌یابی شده دارای نویز بوده و در نتیجه استاندارد نمی‌باشد، لذا از الگوریتم‌های مورفولوژیک جهت استانداردتر کردن آن استفاده می‌کنیم.

شناسایی لبه در روش **Canny** شامل شش مرحله ست:

مرحله‌ی اول فیلتر کردن تصویر اولیه و حذف نویز از تصویر است. برای این منظور فیلتر گوسین با یک ماسک ساده قابل استفاده‌است که به‌طور انحصاری در الگوریتم Canny استفاده می‌شود.

مرحله‌ی دوم یافتن لبه‌های قوی با استفاده از gradient magnitude در هر نقطه می‌باشد که برای این منظور معمولاً از ماسک سوبل استفاده می‌گردد.

$$|G|=|Gx|+|Gy|$$

مرحله‌ی سوم بدست آوردن جهت لبه‌های تصویر، با استفاده از مقدار gradient در راستای x و y است که در مرحله‌ی قبل محاسبه شد. فرمول زیر برای محاسبه‌ی جهت لبه‌ها استفاده می‌شود.

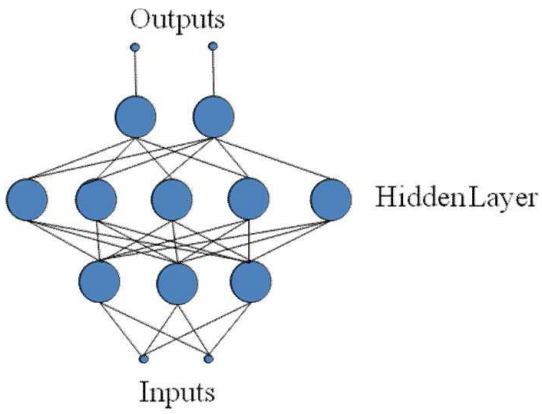
$$theta = invtan (Gy / Gx)$$

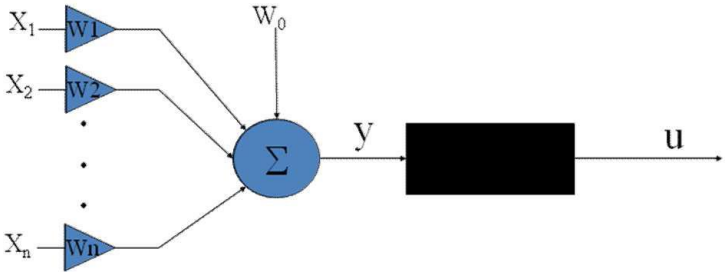
مرحله‌ی چهارم اختصاص دادن جهت‌های قابل قبول در تصویر به جهت‌های به‌دست آمده‌است. برای هر پیکسل در تصویر فقط 4 جهت 0 و 45 و 90 و 135 درجه امکان‌پذیر می‌کنیم. بنابراین جهت‌های به-دست آمده را به یکی از این 4 جهت Map می‌کنیم.

مرحله‌ی پنجم nonmaximum suppression است. این مرحله جهت لبه‌ها را بررسی کرده و از بین آن‌ها لبه‌هایی را که نمی‌شود به‌عنوان لبه در نظر گرفت را حذف می‌کند. این مرحله یک خط باریک را در تصویر نهایی می‌دهد.

مرحله‌ی ششم آستانه‌گیری در تشخیص لبه است. در Canny از روشی به نام hysteresis استفاده می‌شود. برای این منظور دو آستانه‌ی بالا و پایین را تعریف می‌کنیم. هر پیکسل که دارای gradient بیشتر از حد بالا باشد به‌عنوان لبه پذیرفته می‌شود و در صورتی که دارای مقدار کمتر از حد پایین باشد رد می‌گردد و در صورتی که دارای مقداری بین این دو حد داشته باشد در صورتی پذیرفته می‌شود که یکی از همسایه‌های آن پذیرفته شده باشد.

ساخت و آموزش شبکه‌ی عصبی



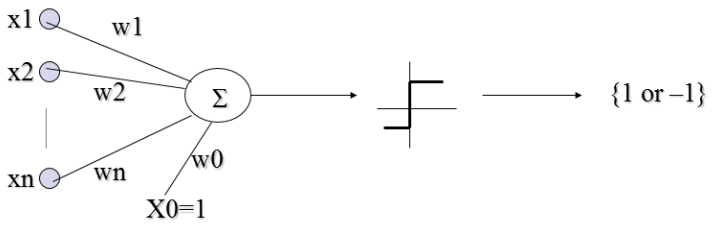


پرسپترون چند لایه

برای آموزش سیستم ما نیاز به یک شبکه‌ی عصبی پرسپترون داریم چون عمل Classification را به سادگی انجام می‌دهد. یک سیستم با تعدادی ورودی و تعدادی خروجی و تعدادی لایه‌ی مخفی است.

پرسپترون

نوعی از شبکه‌ی عصبی بر مبنای یک واحد محاسباتی به نام پرسپترون ساخته می‌شود. یک پرسپترون برداری از ورودی‌های با مقادیر حقیقی را گرفته و یک ترکیب خطی از این ورودی‌ها را محاسبه می‌کند. اگرمقدار حاصل از یک مقدار آستانه بیشتر بود خروجی پرسپترون برابر با 1 و در غیر اینصورت معادل 1– خواهد بود.



شماتیک پرسپترون

یادگیری یک پرسپترون

خروجی پرسپترون توسط رابطه‌ی زیر مشخص می‌شود:

$$O(x_1, x_2, \ldots, x_n) = \begin{cases} 1 & \text{if } w_0 + w_1x_1 + w_2x_2 + \ldots + w_nx_n > 0 \\ -1 & \text{otherwise} \end{cases}$$

که برای سادگی می‌توان آن‌را به‌صورت زیر نشان داد:

$$O(X) = Sgn(WX) \quad \text{where} \quad Sgn(y) = \begin{cases} 1 & \text{if } y > 0 \\ -1 & \text{otherwise} \end{cases}$$

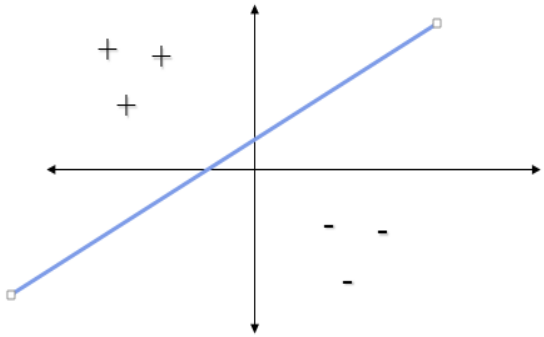
ورودی‌های شبکه‌ی عصبی بسته به تعداد ویژگی‌های انتخاب شده دارد. از آنجایی‌که ورودی‌های هر نورون یا یک است و یا صفر، پس ما باید ویژگی‌ها را به نوع باینری تبدیل کنیم. در این سیستم ما 2500 نورون ورودی داریم و تعداد خروجی‌های ما برابر با تعداد اشکالی است که قرار است سیستم یاد بگیرد (فعلا 4 مورد) و 10 لایه‌ی مخفی داریم. پس نیاز به یک ماتریس وزن داریم ماتریس وزن برابر با

4*2500 می باشد. در حین آموزش ماتریس وزن در حال تغییر می باشد و در نهایت به یک مقدار ثابت می‌رسیم.

توانایی پرسپترون

پریسپترون را می‌توان به‌صورت یک سطح تصمیم hyperplane در فضای n بعدی نمونه‌ها در نظر گرفت. پرسپترون برای نمونه‌های یک طرف صفحه مقدار 1 و برای مقادیر طرف دیگر مقدار 1- بوجود می‌آورد.

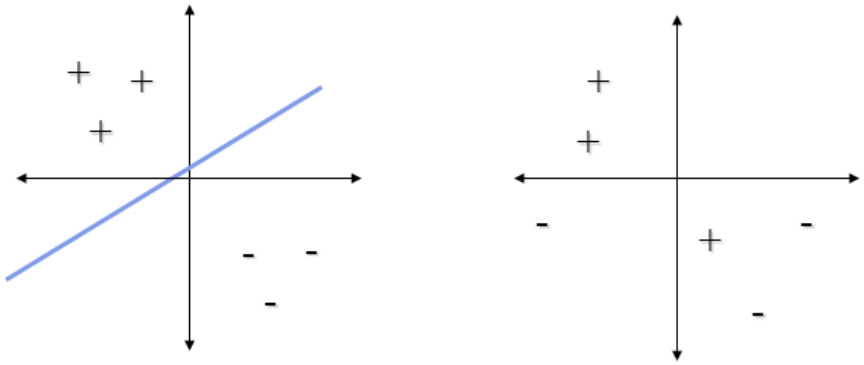
Decision boundary (WX = 0)



جداسازی خطی

توابعی که پرسپترون قادر به یادگیری آنها می‌باشد

یک پرسپترون فقط قادر است مثال‌هایی را یاد بگیرد که به‌صورت خطی جداپذیر باشند. این‌گونه مثال‌ها مواردی هستند که به‌طور کامل توسط یک hyperplane قابل جداسازی می‌باشند[7].



جداسازی خطی (تصویر سمت چپ)، جداسازی غیرخطی (تصویر سمت راست)

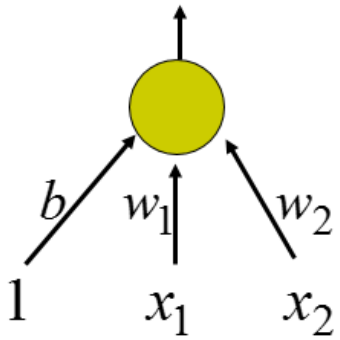
اضافه کردن بایاس

افزودن بایاس موجب می‌شود تا استفاده از شبکه‌ی پرسپترون با سهولت بیشتری انجام شود.

برای اینکه جهت یادگیری بایاس نیازی به استفاده از قانون دیگری نداشته باشیم بایاس را به‌صورت یک ورودی با مقدار ثابت 1 در نظر گرفته و وزن W_0 را به آن اختصاص می‌دهیم.

$$\hat{y} = b + \sum_i x_i w_i$$

$$\hat{y} = w_0 + \sum_{i=1} x_i w_i$$



بایاسینگ پرسپتون

آموزش پرسپترون

چگونه وزن‌های یک پرسپترون واحد را یاد بگیریم به نحوی که پرسپترون برای مثال‌های آموزشی مقادیر صحیح را ایجاد نماید؟

دو راه مختلف برای این امر وجود دارد که به شرح ذیل می‌باشند:

- **قانون پرسپترون**
- **قانون دلتا**

الگوریتم یادگیری پرسپترون

مقادیری تصادفی به وزن‌ها نسبت می‌دهیم. پرسپترون را به تک تک مثال‌های آموزشی اعمال می‌کنیم. اگر مثال غلط ارزیابی شود مقادیر وزن‌های پرسپترون را تصحیح نموده و روند ذیل را پی می‌گیریم.

آیا تمامی مثال‌های آموزشی درست ارزیابی می‌شوند:

- **بله ← پایان الگوریتم.**
- **خیر ← به مرحله‌ی 2 برمی‌گردیم.**

قانون پرسپترون

برای یک مثال آموزشی $X = (x1, x2, \ldots, xn)$**، در هر مرحله وزن‌ها بر اساس قانون پرسپتون به‌صورت زیر تغییر می‌کنند:**

$$wi = wi + \Delta wi \qquad (3\text{–}5)$$

که در آن

$$\Delta wi = \eta\,(t - o)\,xi \qquad (3\text{–}6)$$

t = target output

o = output generated by the perceptron

η = constant called the learning rate (e.g., 0.1)

اثبات شده‌است که برای یک مجموعه مثال جداپذیر خطی این روش همگرا شده و پرسپترون قادر به جداسازی صحیح مثال‌ها خواهد شد[7].

قانون دلتا (Delta Rule)

الگوریتم یادگیری با استفاده از قانون دلتا به‌صورت زیر می‌باشد:

به وزن‌ها مقدار تصادفی نسبت می‌دهیم.

تا رسیدن به شرایط توقف مراحل زیر را ادامه می‌دهیم.

- **هر وزن** w_i **را با مقدار صفر عدددهی اولیه می‌کنیم.**
- **برای هر مثال وزن** w_i **را به‌صورت زیر تغییر می‌دهیم:**

$$\nabla w_i = \nabla w_i + \eta\,(t - o)\,x_i \qquad (3\text{-}8)$$

- **مقدار** w_i **را به‌صورت زیر تغییر می‌دهیم:**

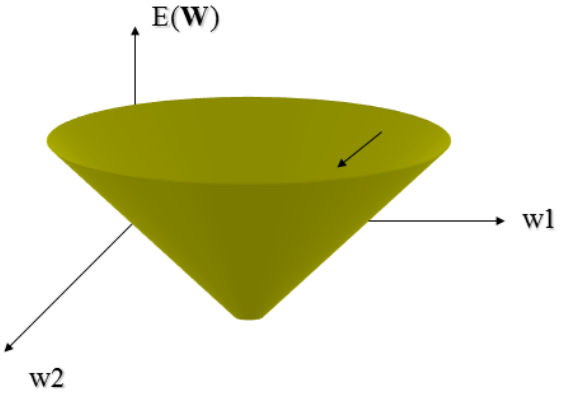
$$w_i = w_i + \nabla w_i \qquad (3\text{-}9)$$

تا خطا بسیار کوچک شود[7].

الگوریتم (gradient descent)

با توجه به نحوه‌ی تعریف E **سطح خطا به‌صورت یک سهمی خواهد بود. ما به‌دنبال وزن‌هایی هستیم که حداقل خطا را داشته باشند. الگوریتم** gradient descent **در فضای وزن‌ها به‌دنبال برداری می‌گردد که**

خطا را حداقل کند. این الگوریتم از یک مقدار دلخواه برای بردار وزن شروع کرده و در هر مرحله وزن‌ها را طوری تغییر می‌دهد که در جهت شیب کاهشی منحنی فوق، خطا کاهش داده شود.



شماتیک گرادیان نزولی

ایده‌ی اصلی: گرادیان همواره در جهت افزایش شیب E عمل می‌کند.

گرادیان E نسبت به بردار وزن w به‌صورت زیر تعریف می‌شود:

$$E(W) = [\, E'/w0,\ E'/w1,\ \ldots,\ E'/wn \,] \qquad (3\text{–}10)$$

که در آن E (W) یک بردار و E' مشتق جزئی نسبت به هر وزن می‌باشد.

برای یک مثال آموزشی X = (x1, x2, …, xn) در هر مرحله وزن‌ها بر اساس قانون دلتا به‌صورت زیر تغییر می‌کند:

$$wi = wi + \Delta wi \qquad (3\text{–}11)$$

Where $\Delta wi = -\eta\ E(\mathbf{W})/wi$

η: learning rate (e.g., 0.1)

علامت منفی نشان دهنده‌ی حرکت در جهت کاهش شیب می‌باشد.

محاسبه‌ی گرادیان

با مشتق گیری جزئی از رابطه‌ی خطا می‌توان به‌سادگی گرادیان را محاسبه نمود:

$$E'(W)/ w_i = \Sigma_i (t_i - O_i) (-x_i) \qquad (3-12)$$

لذا وزن‌ها طبق رابطه‌ی زیر تغییر خواهند کرد:

$$\Delta wi = \eta \Sigma i (ti - oi) xi \qquad (3-13)$$

تست شبکه‌ی عصبی

همان‌طوری که در کد متلب ذیل مشاهده می‌کنید ما برای مثال چهار نوع شکل را به سیستم آموزش داده و در نهایت برای تست الگوریتم، بایستی داشته باشیم:

اگر خروجی [1 -1 -1 -1] **بود یعنی شکل ورودی مربع می باشد،**

اگر خروجی [-1 1 -1 -1] **بود یعنی شکل ورودی مثلث می باشد،**

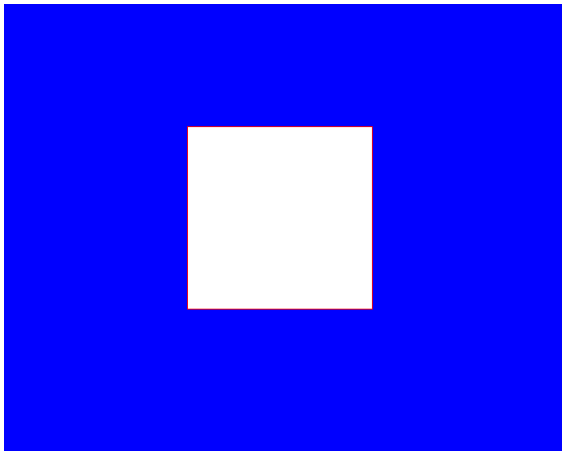
و اگر [-1 -1 1 -1] **بود دایره و اگر** [-1 -1 -1 1] **بود یعنی شکل ورودی شش ضلعی می‌باشد در غیر این‌صورت شکل ارائه شده آسیب دیده یا نویزی تشخیص داده می‌شود.**

```
for k=1:4
for i=1:2500
y=y+test(i)*w(i,k);
end
y=y+b(k);
if(y>=0)
result(k)=1;
else
result(k)=-1;
end
y=0;
end
disp(['Your Output array is : ']);
disp(result);
```

```
if result== [ 1    -1    -1    -1    ]
msgbox('Your test element is <morabaa>');
elseif result== [ -1    1    -1    -1    ]
msgbox('Your test element is <mosalas>');
elseif result==[-1    -1   1    -1     ]
msgbox('Your test element is <Circle>');
elseif result==[ -1    -1    -1    1    ]
msgbox('Your test element is <6 zelii>');
else
msgbox('Your test element is very Damaged');
end
```

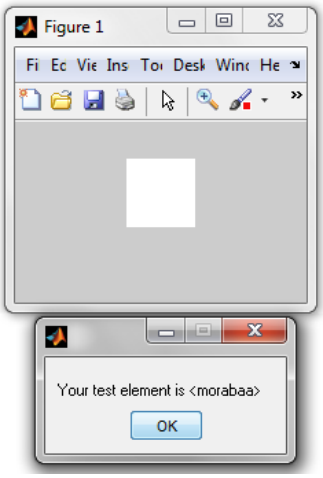
در این ارائه ما با استفاده از دو نوع کد نویسی مجزا که اولی توسط روش یادگیری با ناظر و دومی با استفاده از استخراج ویژگی‌های شکلی را ارائه نمودیم که در زیر نتایج حاصل را به همراه تصاویر تست (ورودی) مشاهده می فرمایید:

نتایج حاصل از کد نویسی شماره‌ی 1



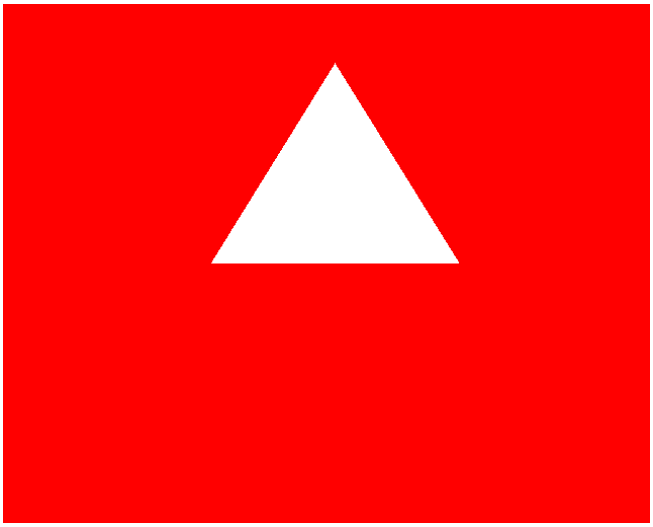
الف

```
Your Output array is :
     1    -1    -1    -1
>>
```

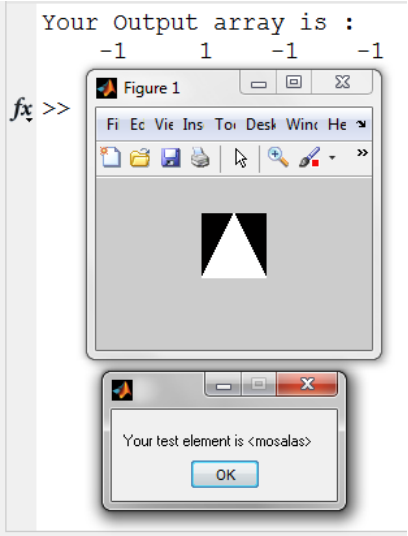


ب

(الف) ورودی، (ب) خروجی

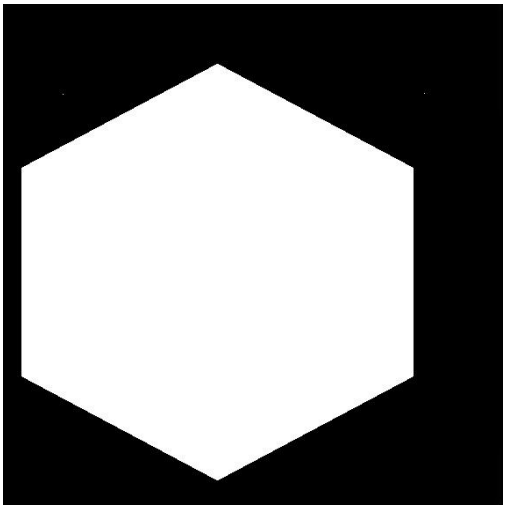


الف

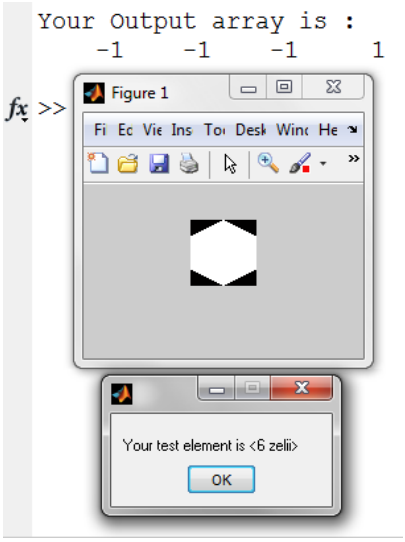


ب

(الف) ورودی، (ب) خروجی

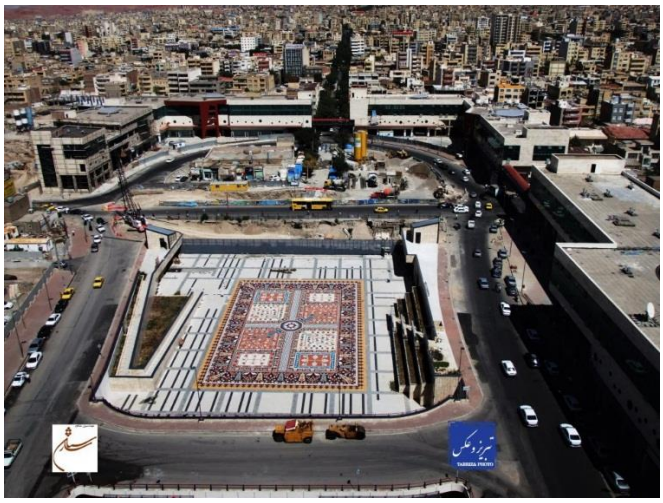


الف

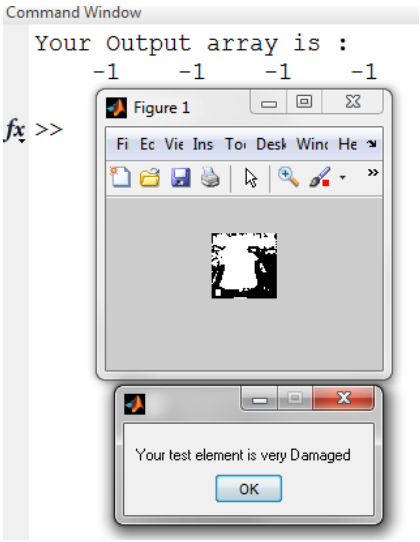


ب

(الف) ورودی، (ب) خروجی



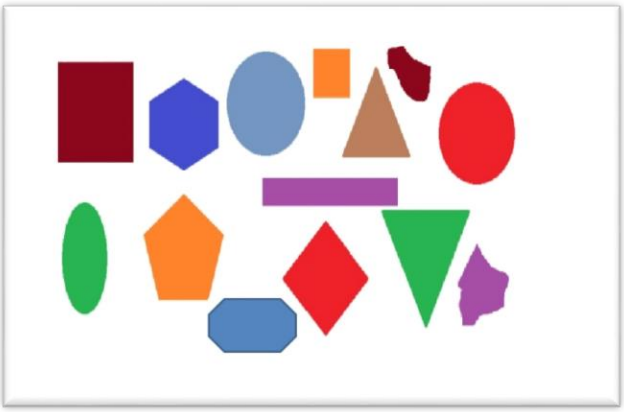
الف



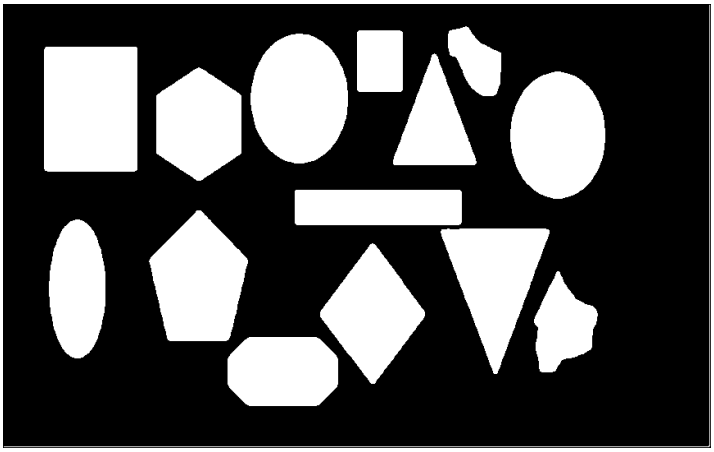
ب

(الف) ورودی، (ب) خروجی

نتایج حاصل از کد نویسی شماره‌ی 2



الف



ب

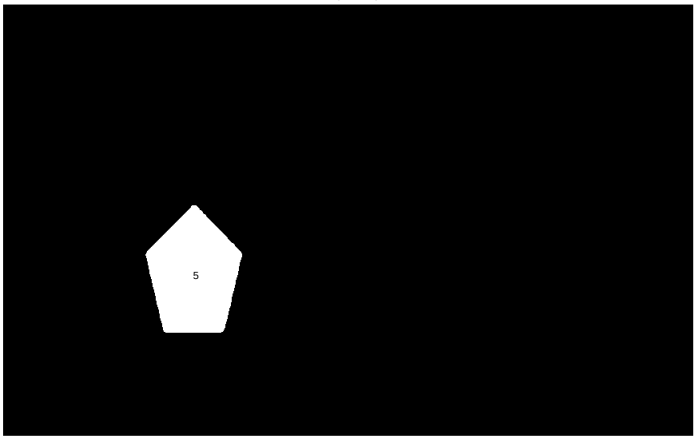
Square Image

پ

Ellipse Image

ت

Pentagon Image

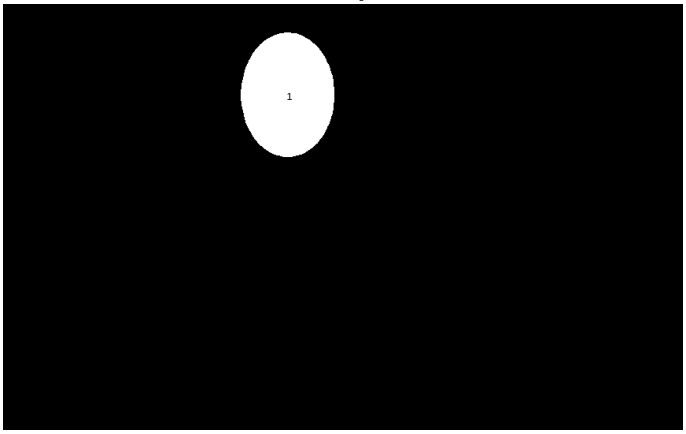


ث

Hexagon Image

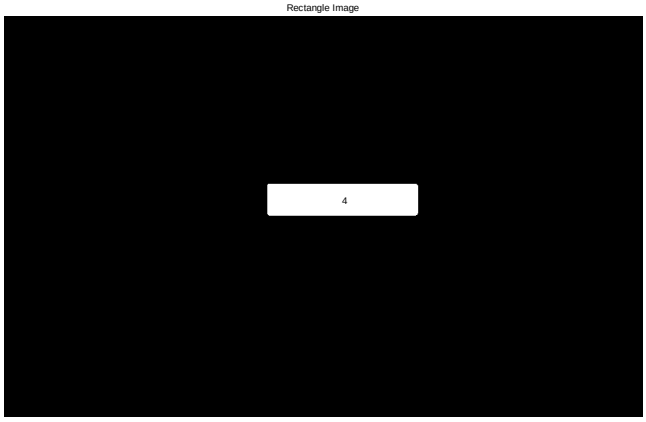
ح

Circle Image



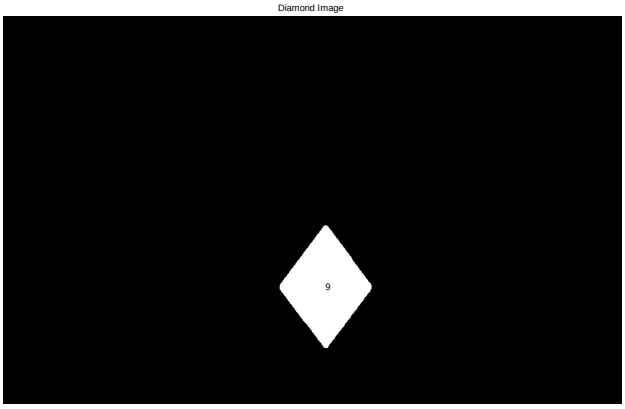
خ

Rectangle Image



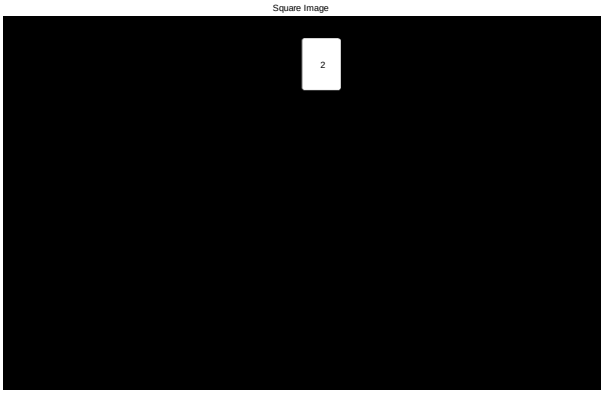
ج

Diamond Image

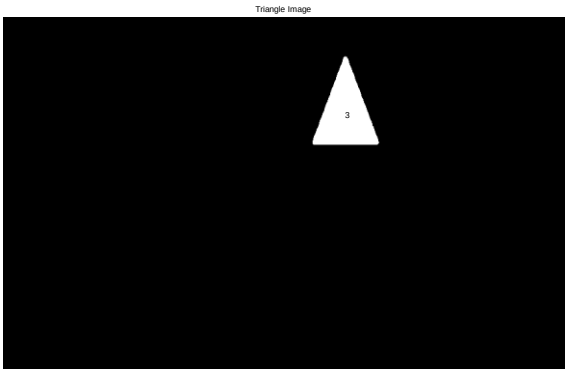


ج

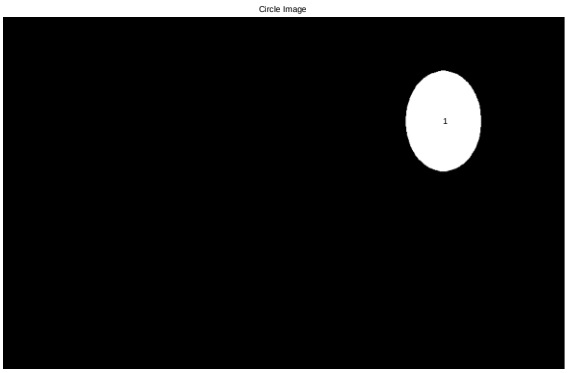
Square Image



د



ذ



ر

(الف) تصویر ورودی، (ب) تصویر باینری شده، (پ) تا (ر) جداسازی اشکال موجود در تصویر اولیه.

شایان ذکر است که الگوریتم شبکه‌ی عصبی استفاده شده در این ارائه قابلیت تعمیم‌پذیری برای هر الگوی جدیدی را دارد. در این الگوریتم ما دارای 2500 نورون ورودی و تعداد 4 خروجی و 10 لایه‌ی مخفی (کمک به یادگیری بهتر برای به روز رسانی وزن‌ها) استفاده گردیده که شامل یک ماتریس به ابعاد 2500×4 **می‌باشد. و از ویژگی‌هایی شامل: تعداد خطوط، تعداد زوایا، مرکز ثقل، لبه‌ها، محیط، تبدیل هاف استفاده شده‌است.**

تحليل احساسات

فصل اول (مقدمه)

مقدمه:

آنالیز احساسات یا کاویدن نظرات ، به فرآیند محاسباتی و الگوریتمیک مطالعه ی نظرات، احساسات و برخوردها در رابطه با موجودیت ها، افراد اتفاقات، مشکلات، موضوعات و تمامی زیرمجموعه های آنها میباشد .این پژوهش، ازنظر فنی پر چالش درعین حال در عمل بسیار پرکاربرد میباشد .به عنوان مثال بنگاه های اقتصادی بسیار علاقه مند هستند که از نظرات، عقاید و میزان رضایت مندی مشتریانشان نسبت به محصولاتشان آگاه شوند .به همین ترتیب مشتریان نیز علاقه مند هستند که نظرات سایر مشتریان را نسبت به یک کالا یا سرویس، قبل از خرید یا استفاده از آن بدانند.

یک شرکت تجاری، دارای محصولات متنوعی میباشد، این شرکت میخواهد نظرات مشتریانش را در رابطه با یکی از محصولات تولیدی خود بداند .مشتریان احتمال دارد که نظرات خود را از چندین روش ابراز نمایند .اولین راه بیان نظرات به واحد روابط عمومی آن شرکت به صورت شفاهی میباشد .دومین روش ارسال ایمیل به شرکت موردنظر و بیان نظرات میباشد و آخرین روش بیان نظرات در شبکه های اجتماعی ، تالارهای گفتگو و سامانه های پرسش و پاسخ متعدد و پراکنده در سطح شبکه ی جهانی وب میباشد. به دلایل اجتماعی، ازجمله علاقه به دیده شدن، کسب احساس صاحب نظر بودن، ابراز تجربیات شخصی و مشابه آن، کاربران علاقه دارند اطلاعات و نظراتشان را به وسیله ی روش سوم بیان نمایند . یعنی در مکانی که سایر کاربران بتوانند به توانایی های فردی شخص صاحب نظر پی ببرند .این روش

به دلیل گستردگی روزافزون شبکه های اجتماعی، تالارهای گفتگو و سامانه های پرسش و پاسخ، بسیار مورد اقبال کاربران قرارگرفته است. این رفتار کاربران، فرآیند جذب نظرات، انتقادات و پیشنهادها را برای شرکت تجاری مذکور، بسیار سخت و پیچیده مینماید .مشتریان احتمالاً در شبکه های مختلف، به زبان های مختلف و در مورد بخش های مختلفی از محصول موردنظر، نظر خواهند داد .در ضمن نظرات آنها در بخش های مختلف امکان دارد مثبت، منفی و یا ممتنع - خنثی - باشد .اولین راهی به ذهن مدیران شرکت میرسد، استفاده از نیروی انسانی بسیار زیاد برای رصد نمودن شبکه های اجتماعی مختلف و استخراج اطلاعات آنها میباشد. اما راهکار پیشنهادی، استفاده از یک فرآیند مکانیزه میباشد، تا نظرات کاربران را از مکان های مختلف جمع آوری نموده، سپس این نظرات را با استفاده از ویژگی های محصول دسته بندی نماید و تشخیص دهد نظر هر مشتری نسبت به هر یک از آن ویژگی ها و درنهایت برآیند نظرات مشتریان در رابطه با محصول موردنظر و ویژگی های آن چه بوده است.

مثال بیان شده در بالا، حداقل انتظار از یک سامانه ی آنالیز احساسات و کاوش نظرات میباشد. اگر فرآیند ذکر شده، به صورت مکانیزه انجام پذیرد، علاوه بر هدایت صاحبان محصول برای بهبود محصولات خویش، به مشتریان جدید نیز برای کسب اطلاعات و خرید هوشمندانه تر کمک مینماید.[15]

بیان مساله:

در این پژوهش، هدف تجزیه و تحلیل متون زیرنویس فیلم ها در راستای تشخیص و طبقه بندی آنها بر اساس ژانر فیلم ها میباشد. تمرکز اصلی بر ارائه ی فرآیندی مبتنی بر تجزیه و تحلیل احساسات و فهم ارتباط میان لغات برای ارائه ی یک دسته بندی دقیق تر میباشد.

دستاورد این پژوهش، ارائه ی یک مدل طبقه بندی فیلم ها به ژانر آنها بر مبنای متن زیرنویس آنها میباشد که تاکنون فعالیتی در این زمینه انجام نپذیرفته است.

ویژگی های مساله و مسایل پیش رو:

در بررسی یک محتوای چندرسانه ای یعنی محتوای شامل صدا، تصویر و متن، هریک دارای نقاط مثبت و منفی میباشند .متون زیرنویس فیلم ها، به دلیل اینکه قابل تبدیل به متن خالص هستند، از لحاظ بار پردازشی و سرعت بررسی بسیار بهینه تر از سایر بخش های یک محتوای چندرسانه ای میباشند .از سوی دیگر در پردازش هایی مانند تجزیه و تحلیل مفاهیم و احساسات، این کار به دلیل وجود لغات و عبارات در متن بسیار سریع و بهینه قابل انجام میباشد. از طرف دیگر متون زیرنویس فیلم ها دارای یک گسستگی ذاتی نیز میباشند، به گونه ای که ممکن است در بخشی از فیلم صحبتی انجام نشده باشد و بنابراین متنی نیز برای آن وجود نداشته باشد و در این حین، موضوع و سناریوی فیلم تغییر نموده باشد .این در حالی است که متن به صورت یکپارچه ارائه گردیده است و به صورت یکپارچه تجزیه و تحلیل میگردد. مسئله ی دیگر تداخل بسیار زیاد ژانرهای فیلم ها است .طبق محاسبه ی انجام شده، به طور میانگین، هر فیلم دارای دو ژانر میباشد .یک فیلم حداقل دارای یک ژانر میباشد و در فیلم های چند ژانری، پایگاه IMDb که به عنوان مرجع اطلاعات فیلم ها شناخته میشود، هیچگونه اطلاعاتی در رابطه با ژانر اصلی فیلم نداده است .این اتفاق تداخل و خطا را در طبقه بندی فیلم ها افزایش میدهد.

آخرین مسئله، سلیقه ای بودن تعیین ژانرها میباشد که در اکثر متون پژوهشی که در مورد طبقه بندی فیلم ها بر اساس ژانر آنها منتشر گردیده است به این مطلب اشاره شده است .به این معنی که استاندارد خاصی برای طبقه بندی فیلم ها به ژانر آنها وجود ندارد و این امر توسط مجموعه ی وسیعی از بینندگان تعیین میگردد .به همین دلیل جدا از مبحث یادگیری ماشین، انسانها نیز در بعضی مواقع در حدس ژانر فیلم ها با خطا عمل مینماید. [12]

خلاصه ی مدل پیشنهادی:

در این پژوهش دو راهکار انتخاب ویژگی ها بیان گردیده است، راهکار اول با استفاده جمله بندی زیرنویس ها و سپس برچسب گذاری ادات سخن در هریک از جملات و در نهایت انتخاب صفات، قیود و اسامی از میان ادات سخن مختلف موجود میباشد. [16]

در راهکار دوم انتخاب ویژگی ها علاوه بر مراحل انجام شده در راهکار اول، با تعریف پارامتر حد آستانه ی شباهت ، کلمات یکتای هر ژانر مشخص گردیدند، سپس سه مرتبه با انتخاب 10، 15 و 20 عدد از پرتکرارترین کلمات انتخاب شده، به عنوان ویژگی های انتخابی هر ژانر، با استفاده از محاسبه ی اطلاعات متقابل نقطه به نقطه برای هر کلمه ی یک متن زیرنویس با ژانر نامشخص و تمامی ویژگی های هریک از ژانرهای موجود، اقدام به پیشنهاد دو ژانر برای یک متن زیرنویس با ژانر نامشخص مینماییم.

ساختار پژوهش جاری:

در این پژوهش ابتدا در فصل دوم، مفاهیم مقدماتی تجزیه و تحلیل احساسات و الگوریتم های مطرح در یادگیری ماشین بررسی میگردند .برخی از کارهای انجام پذیرفته در زمینه ی تجزیه و تحلیل احساسات و همچنین تشخیص ژانر تصاویر متحرک، در فصل سوم بررسی میگردند .در فصل چهارم مدل پیشنهادی برای طبقه بندی متون زیرنویس فیلم ها به ژانر آنها و همچنین مجموعه ی داده یMetaSub ارائه میگردند .ارزیابی مدل پیشنهادی در فصل پنجم صورت پذیرفته است و در نهایت، فصل ششم، به نتیجه گیری و پیشنهاد پژوهش های آتی اختصاص داده شده است. [12]

فصل دوم(مفاهیم پایه)

مقدمه:

هدف یادگیری ماشین این است که کامپیوتر(در کلی ترین مفهوم آن)، بتواند به تدریج و با افزایش داده ها بازدهی بالاتری در وظیفه مورد نظر پیدا کند. طیف پژوهش هایی که در یادگیری ماشینی مورد بررسی قرار میگیرند بسیار گسترده است. یکی از جدیدترین مسائل، تحلیل احساسات میباشد .تحلیل احساسات به دلیل کاربرد فراوان و همچنین مسائل پر چالش و حل نشده، از یادگیری ماشین نیز به عنوان یک زیرمجموعه استفاده خواهد نمود.

اصطلاحات، مفاهیم و تعاریف پایه تحلیل احساسات:

در این بخش به بیان تعدادی از تعاریف و اصطلاحات مهم و پایه ای در موضوع آنالیز احساسات پرداخته میشود .در موارد مورد نیاز، مثالی برای شفاف تر شدن مفاهیم بیان گردیده است .بدیهی است که توضیح جامع همه ی مفاهیم، از حوزه ی این نوشته خارج می باشد .اما سعی شده است در مطالب به صورت مطلوبی، مفاهیم را پوشش دهند.

تعریف 1 (موجودیت):

یک موجودیت، در اصل یک محصول، خدمت، فرد، اتفاق، سازمان ویا یک موضوع میباشد .این موجودیت را به وسیله ی یک زوج مرتب و به شکل $e : (T , W)$ نمایش میدهند .که در این زوج مرتب، T ، نشان دهنده ی یک سلسله مراتب از زیر بخش های موجودیت e میباشد W .در این زوج

مرتب، نمایش دهنده ی مجموعه ی ویژگی های موجودیت e می باشد .هر یک از زیر بخش های e نیز، خود دارای یک مجموعه از ویژگی ها میباشند. [15]

مثال: یک خودرو از یک برند خودروسازی، یک موجودیت به حساب می آید .این خودرو یک سلسله مراتبی از زیر بخش ها را در خود دارد .مانند درب، موتور، جعبه دنده، فرمان، تهویه مطبوع، سیستم ترمز، تایر، باتری و غیره .از طرف دیگر، ویژگی هایی نیز دارد، مانند حداکثر سرعت، وزن، تعداد دنده، سایز لاستیک، طول عمر باتری و غیره.

بر اساس تعریف 1 یک موجودیت را میتوان به صورت درختی نمایش داد .در این درخت،گره ریشه، همان نام موجودیت است .هر گره غیر ریشه نیز یک زیر بخش از موجودیت اصلی است و هر یال بین گره ها، نشانگر ارتباط زیرمجموعه ای است .در ضمن، هر گره دارای یک مجموعه ای از ویژگی ها میباشد .یک احساس/نظر میتواند در مورد هریک از گره های این درخت و در مورد هریک از ویژگی های آن گره باشد. [12]

تعریف بیان شده، دارای پیچیدگی میباشد، به همین دلیل در عمل، تعریف فوق را به دو دلیل، به شکل زیر ساده سازی مینمایند .اولاً، فرآیند پردازش زبان طبیعی، کاری سخت و پیچیده میباشد .درصورتی که نیاز به پردازش زبان طبیعی با جزئیات زیاد، - همانند آنچه درتعریف 1 گفته شد - این کاری بسیار سخت خواهد بود .ثانیاً، برای یک کاربر معمولی، استفاده از یک ساختار سلسله مراتبی، سخت و پیچیده خواهد بود .بنابراین برای سادگی، ارتفاع درخت را به دو سطح کاهش میدهیم و از اصطلاح جدید جنبه ،برای بیان زوج مرتب(زیر بخش و ویژگی) استفاده میکنیم .در درخت ساده سازی شده، ریشه هنوز، موجودیت اصلی است .درحالی که گره های سطح دوم، جنبه های مختلف یک موجودیت را بیان میکنند.

احساسات / نظرات، از نظر علمی به دو دسته تقسیم بندی میشوند؛ احساسات/نظرات معمولی و احساسات/نظرات مقایسه ای، احساسات/نظرات معمولی، که عموماً در پژوهش ها با واژه ی احساسات/نظرات، مورداستفاده قرار میگیرند .احساسات و نظرات مقایسه ای، در مورد بیش از یک

موجودیت و در مورد ابعاد مختلف آنها به مقایسه میپردازد .به عبارت دیگر حداقل یک جنبه از حداقل دو موجودیت را بررسی مینماید.

تعریف 2 (احساس/نظر):

یک عقیده یا نظر، یک پنج بخشی به شکل (E(i), A(i,j), OO(i,j), H(k), T(m)) میباشد . در این پنج بخشی، E(i) ، نام موجودیتی است که در مورد آن نظر/عقیده بیان شده است A(i,j) .، موضوعی است که قرار است در آن موضوع، در مورد موجودیت i، نظر/عقیده بیان شود OO(i,j,k) .نشان دهنده ی نوع جهت گیری است که در مورد موضوع j ام از موجودیت i ام بیان گردیده است H(k) .بیان کننده(صاحب) آن نظر/عقیده میباشد .پنجمین بخش، T(m) زمانی است که این نظر/عقیده توسط فرد k بیان شده است.[15]

پارامترهای مطرح شده در تعریف2، پارامترهای ضروری در استخراج اطلاعات یک نظر/عقیده/احساس است. در صورت نبود هر یک از آنها، نظر/عقیده/احساس ارزش کافی برای استخراج اطلاعات را دارا نیست .از طرف دیگر امکان اضافه نمودن ویژگی های جدیدی مانند سن و جنسیت نیز به پنج بخش بیان شده درتعریف 2 وجود دارد و به غنی تر شدن اطلاعات و تصمیم گیری کمک میکند و البته بر پیچیدگی فرآیند نیز میافزاید. [15]

تعریف 3 (طبقه بندی احساسات):

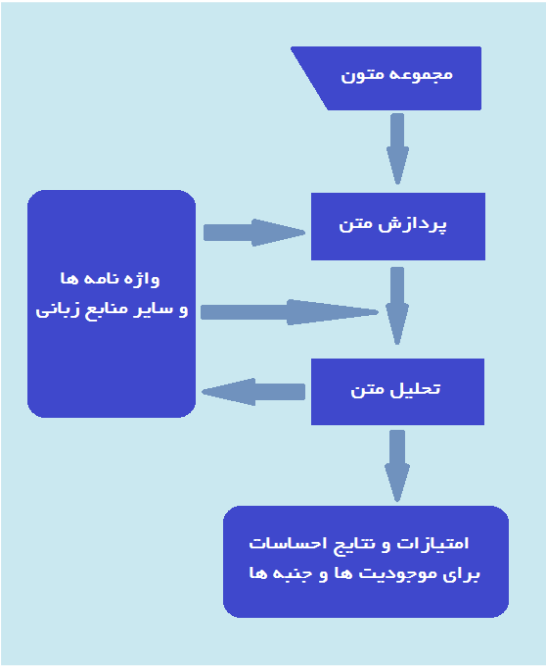
به فرآیند، تشخیص ویژگی های کلیدی و لغات اصلی متن و دسته بندی آنها بر اساس احساسات مختلف، طبقه بندی احساسات گویند.

در طبقه بندی جملات، دو کلاس پایه ای با توجه به موضوع جملات بسیار متداول است: جملات مفعولی که شامل اطلاعات مبتنی بر واقعیت و جملات فاعلی که شامل دیدگاه های صریح کاربران، باورها و

دیدگاه ها در مورد یک موضوع مشخص هستند .در این بخش بیشتر بر روی تحلیل جملات مفعولی تمرکز شده است .به عنوان مثال یک دیدگاه در مورد یکی از هتل های شهر منهتن آورده شده است.

"اتاق اصلی بسیار جادار، تمیز و بسیار مرتب چیده شده بود .کارمند بخش پذیرش، پیشخدمت و نظافتچی بسیار کمک میکردند .سیستم تهویه ی هوای مطبوع عملکرد بسیار مناسبی داشت و تختخواب با سایز بسیار بزرگ بهترین چیزی بود که تاکنون تجربه کرده بودم. ساختمان و اتاق ها در مقابل صدا کاملاً مناسب عایق بندی شده بودند و در همسایگی هتل بهترین مراکز خرید و رستوران ها موجود است و همچنین دسترسی به بزرگراه نیز میسر است .تنها "مشکل "موجود دسترسی به اینترنت پرسرعت بود که فقط در طبقات 8 تا 12 در دسترس بود" . [15]

به طورکلی این دیدگاه درباره ی این هتل مثبت بوده است و به وجه های مختلف هتل شامل: سیستم تهویه، برخورد کارکنان، تختخواب ، اطراف هتل و دسترسی به اینترنت اشاره میکند. سامانه های تحلیل احساسات باید قادر به امتیازدهی به کل دیدگاه به همان صورت که به هر یک از ویژگی های هتل امتیاز تعلق میگیرند، باشد.



نمودار معماری کلی یک سامانه ی تحلیل احساسات:

ورودی سیستم مجموعه ای از متون با فرمت های مختلف میباشد.متون این مجموعه نیازمند پیش پردازش های زبان شناختی از قبیل نشانه گذاری، ریشه یابی، برچسب گذاری، استخراج موجودیت و استخراج روابط بین کلمات است. سیستم ممکن است از یک با چند واژه نامه و منابع زبان شناختی استفاده کند .بخش اصلی سیستم ماژول تحلیل متن است که از منابع زبان شناختی جهت تفسیر احساسات متون استفاده میکند .تفاسیر ممکن است به همه ی متون الصاق شود (تحلیل اطلاعات در سطح متن)، به جملات نسبت داده شود(در سطح جملات)یا به یکی از ویژگی ها نسبت داده شود(در سطح ویژگی ها). این تفاسیر خروجی سیستم میباشد و به کاربر به صورت های گوناگون میتواند نمایش داده شود. [12]

پیشفرض بسیار مهم در تعاریف، مثال ها و مفاهیم این فصل، در بحث طبقه بندی احساسات، به قرار زیر است:

پیشفرض: در طبقه بندی احساسات، فرض بر این است هر متن، نظرات و احساسات راجع به یک محصول و نوشته شده توسط یک نویسنده را بررسی میکند.

ساده ترین روش تحلیل احساسات تحلیل بر مبنای متون میباشد. متن شامل یک دیدگاه در مورد یک شی اصلی که توسط نویسنده توصیف شده است میباشد .دو رهیافت اساسی در تحلیل احساسات در سطح متون وجود دارد :یادگیری با ناظر و یادگیری بدون ناظر.

اکثر روش های موجود برای طبقه بندی احساسات در سطح متن، بر اساس یادگیری با ناظر ، میباشند . گرچه روش هایی نیز بر اساس یادگیری بدون ناظر ارائه شده اند که در بخش اول پژوهش به مفهوم آن پرداخته شد.

الگوریتم های مطرح در یادگیری ماشین:

1)دسته بندی کنندهNaive Bayes :

Naïve Bayes، از روش های مطرح یادگیری با ناظر است .هدف ایجاد قانونی است که قادر باشد اعضای جدید را در یکی از دسته های موجود قرار دهد و این کار را تنها با داشتن بردارهایی از متغیرهای توصیف کننده اجسام جدید انجام دهد .مسائلی از این نوع، دسته بندی با ناظر نام دارند و روش های بسیاری برای ایجاد چنین قوانینی به وجود آمده اند .یک روش بسیار مهم روش بیز ساده است که بیز سطحی و بیز مستقل نیز نامیده میشود .این روش به دلایل متعددی اهمیت دارد .از جمله؛ ساخت آن بسیار ساده است و نیازی به برنامه های تخمین پارامتر تکرارشونده پیچیده ندارد .یعنی از آن میتوان برای مجموعه داده های وسیع استفاده کرد و در نهایت این روش معمولاً بسیار موفق عمل میکند .بیز ساده، ممکن است بهترین دسته بندی کننده ممکن در یک کاربرد خاص نباشد اما اغلب میتوان به قدرت و عملکرد عالی آن اطمینان نمود.

این مدل از قدیمی ترین الگوریتم های دسته بندی رسمی است و هنوز حتی در ساده ترین شکل بسیار مؤثر است .از این مدل در دسته بندی متون و جداسازی هرزنامه ها به طور گسترده استفاده میشود. [12]

2)درخت تصمیم:

ساختار درخت تصمیم در یادگیری ماشین، یک مدل پیشبینی کننده میباشد که حقایق مشاهده شده در مورد یک پدیده را به استنتاج هایی در مورد مقدار هدف آن پدیده نقش میکند. فن یادگیری ماشین برای استنتاج یک درخت تصمیم از داده ها، یادگیری درخت تصمیم نامیده میشود که یکی از رایج ترین روش های داده کاوی است. درختان تصمیم قادر به تولید توصیفات قابل درک برای انسان، از روابط موجود در یک مجموعه ی داده ای هستند و میتوانند برای وظایف دسته بندی و پیشبینی بکار روند .این

فن به شکل گستردهای در زمینه های مختلف همچون تشخیص بیماری دسته بندی گیاهان و استراتژی های بازاریابی مشتری بکار رفته است. [17]

اهداف اصلی درخت های تصمیم گیری دسته بندی کننده در ادامه بیان شده اند: [14]

- **داده های ورودی را تا حد ممکن درست دسته بندی کنند.**
- **دانش یادگیری شده از داده های آموزشی را به گونه ای عام سازی کنند که داده های دیده نشده را با بالاترین دقت ممکن دسته بندی کنند.**
- **در صورت اضافه شدن داده های آموزشی جدید، بتوان به راحتی درخت تصمیم گیری را گسترش داد(دارای خاصیت افزایشی باشند).**
- **ساختار درخت حاصل به ساده ترین شکل ممکن باشد.**

انواع درختان تصمیم:

- **درخت دسته بندی: هنگامی که خروجی یک درخت، یک مجموعه ی گسسته از یک مجموعه مقادیر ممکن است؛ به آن درخت دسته بندی ، میگوییم(مثلاً مؤنث یا مذکر، برنده یا بازنده.)این درختها تابع** $X \rightarrow C$ **را بازنمایی میکنند که در آن** C **مقادیر گسسته میپذیرد.** [17]

- **درخت برگشت :هنگامی که بتوان خروجی درخت را یک عدد حقیقی در نظر گرفت آن را درخت برگشت مینامیم(مثلاً قیمت خانه یا طول مدت اقامت یک بیمار در یک بیمارستان).این درختان اعداد را در گره های برگ پیشبینی میکنند و میتوانند از مدل رگرسیون خطی یا ثابت (یعنی میانگین)یا مدل های دیگر استفاده کنند.**

- **درخت** CART : **CART نامی است که به هر دو روال بالا اطلاق میشود. نام**CART **سرنام کلمات درختان برگشت و دسته بندی است.**

- **درختان خوشه ای: فقط نمونه ها را در گره های برگ گروه بندی میکنند.**

طرز کار درخت تصمیم:

درختان تصمیم، نمونه ها را با مرتب کردن آنها در درخت از گره ی ریشه به سمت گره های برگ دسته بندی میکنند .هر گره ی داخلی در درخت، صفتی از نمونه را آزمایش میکند و هر شاخه ای که از آن گره خارج میشود متناظر یک مقدار ممکن برای آن صفت میباشد .همچنین به هر گره ی برگ، یک دسته بندی منتسب میشود .هر نمونه، با شروع از گره ریشه - ی درخت و آزمایش صفت مشخص شده توسط این گره و حرکت در شاخه ی متناظر با مقدار صفت داده شده در نمونه، دسته بندی میشود .این فرآیند برای هر زیردرختی که گره ی جدید ریشه ی آن میباشد تکرار میشود. [14]

3) ماشین بردار پشتیبان(SVM):

ماشین بردار پشتیبان یا SVM ، یکی از روش های یادگیری باناظر است که از آن برای طبقه بندی و رگرسیون استفاده میکنند .این روش از جمله روش های نسبتاً جدیدی است که در سال های اخیر کارایی خوبی نسبت به روش های قدیمی تر از جمله شبکه های عصبی پرسپترون، برای طبقه بندی نشان داده است. مبنای کاری دسته بندی کننده SVM دسته بندی خطی داده ها است.

یکی از روش هایی که در حال حاضر به صورت گسترده برای مسئله دسته بندی مورداستفاده قرار میگیرد، روش ماشین بردار پشتیبان است . شاید به گونه ای بتوان محبوبیت کنونی روش ماشین بردار پشتیبان را با محبوبیت شبکه های عصبی در دهه گذشته مقایسه کرد .علت این قضیه نیز قابلیت استفاده این روش در حل مسائل گوناگون میباشد، درحالیکه روش هایی مانند درخت تصمیم گیری را نمیتوان به راحتی در مسائل مختلف به کار برد. [18]

این الگوریتم دارای مبانی نظری قوی و بی نقصی میباشد، فقط به تعداد کافی، نمونه احتیاج دارد و به تعداد ابعاد مسئله حساس نیست .همچنین متدهایی کارآمد برای یادگیریSVM به سرعت در حال

رشد هستند .در یک فرایند یادگیری که شامل دو کلاس میباشد، هدفSVM پیدا کردن بهترین تابع برای طبقه بندی میباشد به نحوی که بتوان اعضای دو کلاس را در مجموعه داده ها از هم تشخیص داد . معیار بهترین طبقه بندی به صورت هندسی مشخص میشود، برای مجموعه داده هایی که به صورت خطی قابل تجزیه هستند .به طور حسی آن مرزی که به صورت بخشی از فضا تعریف میشود یا همان تفکیک بین دو کلاس به وسیله ابرصفحه تعریف میشود. همین تعریف هندسی به ما اجازه میدهد تا کشف کنیم که چگونه مرزها را بیشینه کنیم ولو اینکه تعداد بیشماری ابرصفحه داشته باشیم و فقط تعداد کمی، شایستگی راه حل برایSVM دارند دلیل اینکه SVM روی بزرگترین مرز برای ابرصفحه پافشاری میکند این است که قضیه قابلیت عمومیت بخشیدن به الگوریتم را بهتر تأمین میکند .این نه تنها به کارایی طبقه بندی و دقت آن روی داده های آزمایشی کمک میکند، فضا را نیز برای طبقه بندی بهتر داده های آتی مهیا میکند.

یکی از مشکلات SVM پیچیدگی محاسباتی آن است .بااینحال این مشکل به طور قابل قبولی حل شده است .یک راه حل این است که یک مسئله بهینه سازی بزرگ را به یک سری از مسائل کوچکتر تقسیم کرد که هر مسئله شامل یک جفت با دقت انتخاب شده از متغیرها که مسئله به طور مؤثر بتواند از آنها بهره ببرد .این پروسه تا زمانی که همه ی این قسمت های تجزیه شده حل شوند ادامه خواهد داشت. [18]

پارامترهای الگوریتم ماشین بردار پشتیبان چگونه تنظیم می‌شوند؟

تنظیم پارامترها برای الگوریتم ماشین بردار پشتیبان به طور مؤثر باعث بهبود کارایی مدل می‌شود. در ادامه برخی از پارامترهای مهمی که تأثیر بیشتری بر کارایی مدل دارند، یعنی «کرنل»، «گاما» و «C» مورد بررسی قرار می‌گیرند. [13]

کرنل: این پارامتر پیش از این مورد بررسی قرار گرفت. گزینه‌های گوناگونی شامل«linear» ، «rbf»، «poly»برای کرنل وجود دارند و در حالت پیش‌فرض کرنل روی پارامتر rbf قرار دارد rbf و poly برای خط جداساز غیر راست مفید هستند.

پیشنهاد می‌شود در صورت زیاد بودن تعداد ویژگی‌ها (۱۰۰۰ <)از کرنل خطی استفاده شود

گاما (gamma) : ضریب کرنل برایrbf ، polyو sigmoid است .هرچه مقدار گاما بیشتر باشد، الگوریتم تلاش می‌کند برازش را دقیقاً بر اساس مجموعه داده‌های تمرینی انجام دهد و این امر موجب تعمیم یافتن خطا و وقوع مشکل بیش برازش (Over-Fitting) می‌شود.

C :پارامتر جریمه C ، برای جمله خطا است. این پارامتر همچنین برقراری تعادل بین مرزهای تصمیم‌گیری هموار و طبقه‌بندی نقاط داده تمرینی را کنترل می‌کند.

مزایا و معایب الگوریتم ماشین بردار پشتیبان:

- **مزایا**
  - **حاشیه جداسازی برای دسته‌های مختلف کاملاً واضح است.**
  - **در فضاهای با ابعاد بالاتر کارایی بیشتری دارد.**
  - **در شرایطی که تعداد ابعاد بیش از تعداد نمونه‌ها باشد نیز کار می‌کند.**

- **یک زیر مجموعه از نقاط تمرینی را در تابع تصمیم‌گیری استفاده می‌کند)که به آن‌ها بردارهای پشتیبان گفته می‌شود)، بنابراین در مصرف حافظه نیز به صورت بهینه عمل می‌کند.**

- **معایب**

- **هنگامی که مجموعه داده‌ها بسیار بزرگ باشد، عملکرد خوبی ندارد، زیرا نیازمند زمان آموزش بسیار زیاد است.**
- **هنگامی که مجموعه داده نوفه (نویز) زیادی داشته باشد، عملکرد خوبی ندارد و کلاس‌های هدف دچار همپوشانی می‌شوند.**
- **ماشین بردار پشتیبان به طور مستقیم تخمین‌های احتمالاتی را فراهم نمی‌کند و این موارد با استفاده از یک اعتبارسنجی متقابل (Cross Validation) پرهزینه پنج‌گانه انجام می‌شوند. این امر با روش SVC موجود در کتابخانه scikit-kearn پایتون، مرتبط است.** [13]

فصل سوم(مروری بر کارهای گذشته)

پنج مشکل ویژه در تحلیل احساسات:

- تحلیل احساسات در سطح متن
- تحلیل احساسات در سطح جمله
- تحلیل احساسات با توجه به ویژگی ها
- تحلیل احساسات مقایسه ای
- فراگیری واژه نامه ی احساسات

تحلیل احساسات در سطح متن:

ساده ترین روش تحلیل احساسات تحلیل بر مبنای متون میباشد. متن شامل یک دیدگاه در یک شی اصلی که توسط نویسنده توصیف شده است میباشد .دو رهیافت اساسی در تحلیل احساسات در سطح متون وجود دارد :یادگیری با ناظر و یادگیری بدون ناظر.

رهیافت بدون نظارت در تحلیل احساسات در سطح متون، بر مبنای جهت گیری معنایی یک سری عبارات ویژه در کل متن شکل گرفته است .اگر میانگین جهت گیری معنایی این عبارات ویژه از یک حد آستانه ی از پیش تعیین شده بیشتر باشد متنِ شامل عبارات ویژه، در دسته ی مثبت و در غیر این صورت در دسته ی منفی دسته بندی خواهد شد.

رهیافت با نظارت فرض بر این دارد که تعداد محدودی دسته موجود است و برای هر دسته به تعداد کافی نمونه آموزشی وجود دارد .ساده ترین حالت هنگامی است که تنها دو دسته ی مثبت و منفی موجود باشد. با توجه به داده های آموزشی، سیستم با یکی از روش های متداول دسته بندی مانند SVM ،

Naïve Bayes، Logistic Regression یا KNN شروع به یادگیری میکند(به طور مثال در سیستم امتیازدهی پنج ستاره وبسایت آمازون). [19]

تحلیل احساسات در سطح جمله:

یک متن ممکن است شامل دیدگاه های مختلفی حتی در مورد یک موجودیت خاص باشد. فرض بر این است که هویت موجودیت مورد بحث در جمله را میدانیم .فرض دیگر این است که در هر جمله تنها یک دیدگاه موجود است. قبل از تحلیل جمله ، باید فاعلی یا مفعولی بودن جمله تعیین گردد و تنها جملات فاعلی تجزیه و تحلیل میشوند.

(در طبقه بندی جملات، دو کلاس پایه ای با توجه به موضوع جملات بسیار متداول است: جملات مفعولی که شامل اطلاعات مبتنی بر واقعیت و جملات فاعلی که شامل دیدگاه های صریح کاربران، باورها و دیدگاه ها در مورد یک موضوع مشخص هستند.)

پس از اتمام طبقه بندی جملات به جملات فاعلی میتوانیم آنها را به دو کلاس مثبت و منفی تقسیم بندی کنیم برای پیدا کردن جهت گیری معنایی متن است . تحقیقات اخیر از جمله نشان داده است که بهتر است انواع مختلف جملات را با روش های مختلفی اداره کنیم .جملاتی که نیازمند استراتژی های خاصی هستند عبارتند از :جملات شرطی، جملات سؤالی و کنایه ها. تشخیص کنایه ها بسیار مشکل است و بیشتر در متون سیاسی به کاربرده میشود. [19]

تحلیل احساسات بر اساس ویژگی:

تحلیل بر اساس ویژگی یک مسئله ی تحقیقاتی است که تمرکز آن بر همه ی عبارات احساسی در یک متن و همه ی ویژگی هایی از موجودیت است که در متن به آنها اشاره شده است.

دو رهیافت قبلی هنگامی که در هر جمله تنها یک دیدگاه بیان شده باشد عملکرد قابل قبولی دارند . اگرچه در اکثر دیدگاه ها مردم در مورد موجودیت هایی صحبت میکنند که دارای صفات و وجوه مختلف بسیار زیادی هستند و افراد مختلف دیدگاه های مختلفی را در مورد هرکدام از صفات و ویژگی دارند . این اتفاق بیشتر در مورد محصولات خاص مانند خودرو، دوربین، گوشی های هوشمند و یا حتی داروهای پزشکی اتفاق میافتد.

به عنوان مثال: نمایشگر آن نور، وضوح بالا و زاویه ی دید زیادی دارد .ولی مشکلاتی نیز دیده میشود، سایز کوچک آن خوش دست بودن و کاربردپذیری.

دسته بندی این نظر هر چند به یکی از دو کلاس مثبت یا منفی باعث از دست رفتن اطلاعات ارزشمندی درون آن میشود، نگارنده ویژگی های مختلفی از دستگاه را مورد بررسی قرار داده است که شامل سرعت، راحتی کار، کیفیت صفحه، میزان حافظه ی در دسترس و ... میشود. بعضی از این ویژگی ها دارای بار احساسی مثبت و بعضی نیز دارای بار احساسی منفی هستند. [19]

تحلیل احساسی رقابتی:

در بسیاری از موارد کاربران به صورت مستقیم نظرات خود را بیان نمیکنند بلکه نظرات قابل مقایسه ای مانند مثال زیر که از یک پایگاه اینترنتی آورده شده است را بیان میکنند:

"C300 **خیلی بهتر از** magna **به نظر می آید .من با هندا سیویک رانندگی کرده ام و به نظر کنترل بهتری نسبت به** TSX **ندارد**". [19]

هدف سیستم تحلیل احساسات در این موارد شناسایی جملاتی است که شامل دیدگاه های مقایسه ای و استخراج بهترین مورد در هرکدام از دیدگاه ها است. در یکی از اولین مقالات در تحلیل احساسات رقابتی اشاره شده که با استفاده از تعداد کمی از کلمات ربطی میتوان تمام جملات رقابتی را پوشش داد:

- **صفات و قیود رقابتی مانند** more , less :
- **صفات و قیود عالی مانند** most, least :

دستیابی به لغتنامه:

واژه نامه ی احساسات مهم ترین و حیاتی ترین منبع برای الگوریتم های تحلیل احساسات است. در اینجا به طور خلاصه رهیافتی برای به دست آمدن واژه نامه ارائه میکنیم. سه گزینه برای دستیابی به لغتنامه وجود دارد:

الف)رهیافت پیاده سازی به صورت انسانی که در آن افراد واژه نامه را تهیه میکنند.

ب) رهیافت مبتنی بر دیکشنری که مجموعه ای از لغات پایه و لغات وابسته به آن در فرهنگ لغات گسترش مییابد.

ج) رهیافت واژه نامه مبتنی بر پیکره ی زبانی که با استفاده از تعداد زیادی متن در یک حوزه ی خاص گسترش مییابد.

مطمئناً رهیافت پیاده سازی انسانی به راحتی ممکن نخواهد بود و در هر حوزه ی خاص نیازمند صرف هزینه های زمانی بسیاری خواهد بود .در ادامه روی دو رهیافت دیگر تمرکز خواهیم کرد. [12]

پژوهش های گذشته در زمینه ی تشخیص محتوای فیلم ها:

تاکنون پژوهش هایی برای تشخیص محتوای فیلم ها و همچنین طبقه بندی فیلم ها بر اساس ژانر آنها صورت پذیرفته است. روش ارائه شده در این پژوهش به این صورت است که هر فیلم، ابتدا به بخش هایی به نام قاب های کلیدی تقسیم میگردد . سپس ویژگی های هر قاب کلیدی استخراج میگردد . سپس با تبدیل این ویژگی ها به بردار عددی، با استفاده از روش خوشه بندی K-Means **، اقدام به دسته بندی و کشف مشابهت فیلم ها مینمایند.**

در پژوهش دیگری، اقدام به طبقه بندی فیلم ها بر اساس ویژگی های صدا و تصویر فیلم صورت میپذیرد .در این پژوهش، فیلم ها به دو دسته پرتحرک و غیر پرتحرک طبقه بندی میگردند .سپس

کلاس غیر پرتحرک را به سه زیرکلاس وحشت ، کمدی و درام تقسیم بندی مینماید .کلاس پرتحرک نیز به دو زیرکلاس آتش سوزی /انفجار و سایر، تقسیم بندی میگردد.

دسته بندی فیلم های غیرمتحرک به سه زیر دسته ی وحشت، کمدی و درام، بر اساس پارامتر میزان نورپردازی صحنه صورت میپذیرد .بر طبق این پژوهش، میزان و توزیع نور در مقایسه با سایه و تاریکی و میزان رنگ صحنه نشان دهنده ی میانگین بصری یک صحنه میباشد. بر این اساس، دو تقسیم بندی صورت میگیرد:

1-نورپردازی روشن: نورپردازی روشن یک صحنه به این معنی است که یک فیلم دارای مقدار زیادی نور روشن میباشد .این صحنه ها عموماً دارای تضاد کمی هستند و اختلاف میان روشن ترین نور و تاریک ترین نور کم است .فیلم های با نورپردازی روشن، عموماً شاد هستند و کمتر درام هستند .همچنین فیلم های کمدی نیز از نورپردازی روشن استفاده مینمایند.

2- نورپردازی تاریک: در این نوع نورپردازی، پس زمینه و بخشی از صحنه عموماً تیره هستند .در فیلم های با نورپردازی تاریک، نرخ تضاد بسیار زیاد است .نورپردازی تاریک، بیشتر در فیلم های درام و ترسناک استفاده میگردد.

در این پژوهش، تشخیص زیرکلاس های فیلم های پرتحرک که شامل آتش سوزی/انفجار و سایر بود، به وسیلهی جلوه های تصویری و صوتی فیلم صورت میپذیرد .تشخیص انفجار با استفاده از جهش های ناگهانی در نمودار هیستوگرام صوت فیلم صورت میپذیرد .با استفاده از تشخیص رنگ ها در صحنه های فیلم، سعی در تشخیص آتش سوزی در فیلم نموده اند. [12]

کاربردهای تحلیل احساسات:

بیشترین کاربرد تحلیل احساسات در حوزه ی تحلیل دیدگاه های مشتریان یک محصول یا یک سرویس است .وبسایت های زیادی که خلاصه ای از دیدگاه های کاربران و ویژگی های متناسب محصولات را فراهم میکنند، موجود هستند .یک مثال قابل ذکر در بالا سرویس جستجوی محصولات گوگل است .

توییتر و فیسبوک نقطه ی کاربردهای تحلیل احساسات هستند .بیشترین کاربرد تحلیل احساسات در جهان، نظارت بر توجه و دیده شدن یک برند خاص در شبکه های اجتماعی مانند فیسبوک و توییتر است.

تحلیل احساسات میتواند ارزش قابل توجهی برای کاندیداهای انتخاباتی به همراه داشته باشد .تحلیل احساسات مدیران کمپین های انتخاباتی را قادر میسازد تا نظرات رأی دهندگان در مورد مسائل مختلف و سخنرانی های کاندیداهای رقیب دنبال کنند.

یکی دیگر از کاربردهای مهم تحلیل احساسات تحلیل بازارهای مالی است.

یکی از کاربردهایی که در آن از تحلیل احساسات رقابتی استفاده شده است ارزیابی ساختار بازار خودروهای سدان و همچنین داروهای ضد دیابت است. [12]

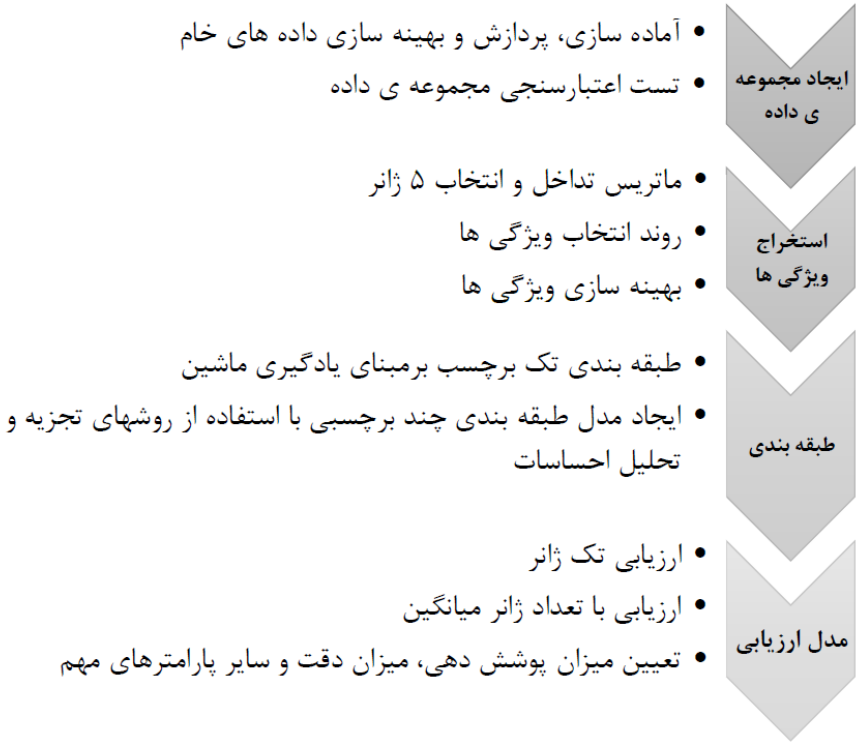
فصل چهارم(مدل پیشنهادی)

راهکار پیشنهادی(مدل جامع طبقه بندی فیلم ها بر اساس ژانر، با استفاده از روش های تجزیه و تحلیل احساسات):

در این بخش به بیان مدل پیشنهادی پرداخته میشود .ازاین رو این راهکار یک مدل نامیده میشود که هدف مورد دستیابی ما، فراتر از یک الگوریتم برای حل یک مسئله است .همانطور که در بخش دوم از همین فصل به آن اشاره شد، مسئله ی موردپژوهش، نیازمند چندین راهکار برای هر یک از بخش های آماده سازی، پردازش، آزمون و ارزیابی میباشد .از سوی دیگر به دلیل اینکه این پژوهش در نوع خود،

برای اولین بار در حال انجام است، هیچگونه مجموعه ی داده ای برای آن موجود نیست .ازاین رو بر آن شدیم که برای اولین بار، اقدام به ایجاد یک مجموعه داده کنیم که به زودی برای دسترسی سایر پژوهشگران قرار داده خواهد شد .بنابراین این مدل از چندین جهت شامل نوآوری میباشد .اول اینکه خود مسئله، مسئله ی جدیدی است .دوم مجموعه ی داده ی ایجادشده، مجموعه ی جدیدی است که تاکنون وجود نداشته است .سوم روش به کار برده شده برای تجزیه و تحلیل احساسات و توجه به مفاهیم لغات به جای استفاده ی محض از الگوریتم های یادگیری ماشین چه برای تشخیص و دسته بندی ویژگی ها و چه برای آنالیز و ارزیابی، راهکاری نو و از جدیدترین حوزه های پژوهش در جهان محسوب میگردد .در این مدل، یک طبقه بندی چند برچسبی نیز ارائه میگردد .در فصل های قبل به طور مفصل درباره این موضوع و ابعاد آن توضیح داده شد .در این بخش ابتدا دید کلی مدل پیشنهادی مورد بررسی قرار میگیرد، سپس به بیان جزییات هر یک از اجزای این دید پرداخته میشود .در شکل زیر بخش های کلی مدل پیشنهادی بیان گردیده است .هریک از این زیر بخش ها در ادامه به تفصیل بررسی گردیده اند. [12]

طرح کلی راهکار پیشنهادی:



در ادامه هریک از بخش ها به تفصیل توضیح داده میشود:

ایجاد مجموعه ی داده MetaSub :

برای ایجاد یک مجموعه ی داده، مراحل مشخصی باید طی شود، تا از طرفی، نیازمندی ها داده محور پژوهش برآورده شود و از طرف دیگر این مجموعه ی داده، صحت سنجی و راستی سنجی گردد تا نتایج به دست آمده صحیح و قابل استفاده برای دیگر پژوهشگران نیز باشد. مجموعه ی داده ی ایجادشده، MetaSub نامگذاری گردید . [12]

نیازسنجی مجموعه ی داده ی MetaSub :

این مجموعه ی داده، از طرفی باید دارای زیرنویس فیلم ها باشد، از طرف دیگر بایستی حاوی اطلاعات فرا داده ی این فیلم ها باشد به گونه ای که منبعی غنی برای داده کاوی بر بستر متن فیلم ها باشد .بر اساس یک اصل اولیه از داده کاوی، هرچه اطلاعات در زمینه های مختلف از یک موضوع واحد بیشتر باشد، احتمال دستیابی به داده های غنی پنهان شده در این اطلاعات بیشتر میباشد. بر اساس مطالعات انجام شده، تاکنون مجموعه ی داده ای که این اطلاعات را به صورت یکپارچه ارائه نماید موجود نیست . بنابراین اقدام به طراحی این مجموعه داده نمودیم. پس از نیازسنجی، پارامترهای موردنیاز در مجموعه ی داده ی MetaSub به صورت زیر تعیین گردیدند:

- **زیرنویس فیلم با شماره شناسایی یکتا بر اساس**OpenSubtitles
- **نام فیلم**
- **شماره شناسایی یکتای فیلم در**IMDb
- **نام کارگردان یا کارگردانان**
- **شماره شناسایی یکتای کارگردان یا کارگردانان در**IMDb
- **نام نویسنده یا نویسندگان**

- **شماره شناسایی یکتای نویسنده یا نویسندگان درIMDb**
- **نام تهیه کننده یا تهیه کنندگان**
- **شماره شناسایی تهیه کننده یا تهیه کنندگان فیلم درIMDb**
- **سال ساخت فیلم**
- **مدت زمان فیلم**
- **نام بازیگران اصلی فیلم**
- **شماره شناسایی یکتای بازیگران درIMDb**
- **ژانر یا ژانرهای فیلم**
- **آدرس یکتای فیلم در پایگاهIMDb**

جمع آوری اطلاعات و داده های مورد نیاز:

بزرگترین پایگاه اطلاعات فیلم ها، IMDb نام دارد .در این پایگاه، اطلاعات بسیار زیادی از فیلم ها، از جمله لیست تمامی فیلم ها، بازیگران، تهیه کنندگان، ژانر فیلم ها وجود دارد .این پایگاه تا پایان ماه سپتامبر 2014 میلادی، حاوی 1815541 عنوان فیلم میباشد .از سوی دیگر بزرگترین پایگاه زیرنویس فیلم ها، پایگاه OpenSubtitles می باشد .در این پایگاه تا پایان سال 2012 میلادی، برای 30370 فیلم، زیرنویس قرار داده شده است .گروه ها و افراد مختلف متناسب با نوع کدگذاری فیلم ها، کیفیت تولید و زبان مورداستفاده، زیرنویس های متعددی را برای هر فیلم در این پایگاه قرار داده اند .تهیه ی زیرنویس این فیلم ها به صورت دسته ای، از سمت پایگاه OpenSubtitles با دریافت 200 زیرنویس در روز محدود شده است .پایگاه پژوهشی Opus حاوی حجم بسیار زیادی از مجموعه های داده می باشد .این پایگاه زیرنویس های پایگاه OpenSubtitles را در زبان های مختلف و همچنین به صورت مجموعه ی زبان های مشترک، تا پایان سال 2012 در اختیار قرار داده است .مجموعه ی زیرنویس های به زبان انگلیسی این پایگاه تا پایان سال 2012 دارای حجمی در حدود 20 گیگابایت میباشد. [12]

مجموعه ی تهیه شده از پایگاه Opus ، حاوی 300000 فایل زیرنویس با فرمت xml ، مربوط به 30370 فیلم یکتا میباشد .طبق مستندات ارائه شده توسط پایگاه Opus ، مشخص گردید، این زیرنویس ها بر اساس شماره ی شناسایی یکتای پایگاه OpenSubtitle شماره گذاری گردیده اند .لازم به ذکر است که پایگاه OpenSubtitles جزو پایگاه های مسدود شده از جانب سازمان تنظیم مقررات رادیویی جمهوری اسلامی ایران می باشد و دسترسی به این پایگاه از کشور ایران ممکن نیست.

بنابراین اقدام به بدست آوردن نگاشت شماره شناسایی یکتای *OpenSubtitles* و شماره شناسایی یکتای *IMDb* ، نمودیم .این نگاشت به ما این امکان را میدهد که اطلاعات جامع پایگاه IMDb را به فایل های زیرنویس پایگاه Opus که برگرفته از پایگاه OpenSubtitles میباشد مرتبط نماییم .با پیاده سازی اسکریپت هایی، طی فرآیندی پیچیده و طولانی، صفحات وب هر فیلم در OpenSubtitlesو IMDb به یکدیگر مرتبط گردیدند، و اطلاعات خام موردنیاز در بخش اول شکل تأمین گردید .به دلیل حجم بالای دریافت و پردازش، در فاز اول، اطلاعات 11087 عدد از زیرنویس فیلم ها به همراه اطلاعات فراداده ی آنها به مجموعه ی داده ی MetaSub اضافه گردید .بیان سایر جزئیات فرآیند پیچیده و زمان بر ایجاد مجموعه ی داده و مسائل مشکلات در روند ایجاد مجموعه ی داده ی MetaSub خارج از مبحث این پژوهش میباشد .در کل برای فرآیند یافتن داده های مورد نظر، تهیه داده های خام، ارزیابی اولیه، استخراج ارتباط میان داده ها، پیاده سازی بخش های مختلف، تلفیق و کنار هم قرار دادن داده ها از منابع مختلف، اشکالزدایی و ارزیابی نهایی بیش از 750 ساعت زمان صرف گردیده است. [12]

اعتبارسنجی مجموعه ی داده یMetaSub :

اعتبارسنجی مجموعه ی داده ی MetaSub ، به طور عمده وابسته به داده ی اصلی آن یعنی زیرنویس ها است که از پایگاه Opus تهیه شده است .این پایگاه یک مرجع بزرگ از داده های خام میباشد . چندین مقاله ی پژوهشی نیز بر پایه ی مجموعه ی داده ی زیرنویس این پایگاه منتشر گردیده است . بنابراین صحت زیرنویس ها و شماره گذاری آنها مورد تأیید میباشد .از طرف دیگر اسکریپت های پیاده سازی شده برای مرتبط نمودن این زیرنویس ها با اطلاعات IMDb آنها بر مبنای یک الگوریتم قابل

پیشبینی پیاده سازی شده اند و رفتار غیرقابل پیشبینی ندارند .بنابراین صحت اطلاعات ارائه شده در MetaSub قابل اثبات میباشد. علاوه بر موارد ذکر شده 2000 رکورد انتخاب شده به شکل تصادفی، به صورت دستی بررسی و صحت سنجی گردید 4000 رکورد نیز به شکل تصادفی انتخاب گردید و به صورت خودکار و با استفاده از کد نویسی بررسی گردید .در پایان، تمامی رکوردهای بررسی شده، به طور کامل و صحیح، با نمونه ی قبلی مطابقت داشتند و به درستی ایجاد گردیده اند .لازم به ذکر است که، برخی اطلاعات فیلم ها توسط پایگاه IMDb ارائه نگردیده است. [12]

استخراج ویژگی ها:

پس از آماده سازی مجموعه ی داده ی MetaSub ، مهمترین بخش استخراج ویژگی های مناسب برای طبقه بندی زیرنویس فیلم ها میباشد .در این مدل، دو روش برای انتخاب ویژگی های مناسب از متن زیرنویس ارائه میگردد، که روش دوم تکامل یافته ی روش اول میباشد.

روش اول :استخراج ادات سخن مؤثر بر مبنای الگوهای پردازش زبان طبیعی و سپس استخراج ویژگی ها برمبنای تبدیل لغات به بردار، TF-IDF و فاکتور فیلترینگ InfoGain

روش دوم :استخراج ادات سخن مؤثر بر مبنای الگوهای پردازش زبان طبیعی و سپس استخراج لغات یکتای نزدیک به هر ژانر بر مبنای حد آستانه ی شباهت

از روش های ذکرشده در راستای استخراج ویژگی های مناسب ارائه گردیده است که در ادامه با جزئیات و الگوریتم اجرایی، تشریح خواهد گردید .در روش های ارائه شده در بالا، روش دوم تکامل یافته ی روش اول میباشد و بر مبنای موفقیت های روش اول تعریف گردیده است. به طورکلی در روش اول از فن های متن کاوی استفاده گردیده است و در روش دوم سعی شده است ویژگی ها متناسب با فن های تجزیه و تحلیل احساسات انتخاب گردند .در ادامه خواهیم دید که برای هریک از این روش های استخراج ویژگی ها، چندین روش طبقه بندی زیرنویس ها بر اساس ژانر آنها ارائه میگردد که در فصل بعدی نتایج هر یک از آنها به تفکیک ارائه گردیده است .هدف نهایی این پژوهش، برداشتن قدم اول

در راستای میل نمودن به سمت روشی مبتنی بر مفاهیم لغات، احساسات کاوی و میزان نزدیکی لغات برای طبقه بندی فیلم ها به ژانرهای مختلف و بر اساس متن هر فیلم میباشد .به این معنی که جدا از روش های سنتی متنکاوی، که مبتنی بر روش هایی بدون درک مفهوم لغات بودند، مدل جامعی ارائه گردد که هم در انتخاب ویژگی ها و هم در طبقه بندی متون زیرنویس، هر چه بیشتر به مفاهیم لغات اهمیت بدهد.

روش اول استخراج ویژگی ها:

در راستای کاهش ویژگی های مورد استفاده و تشخیص ویژگی های مناسب و با توجه به اینکه مهم ترین فاکتور در تشخیص ژانر یک فیلم، نوع احساساتی است که به بیننده منتقل مینماید .ادات سخن مناسب برای این موضوع، صفت، قید و اسم میباشد .بنابراین ابتدا اقدام به استخراج صفت ها، قیدها و اسم ها مینماییم .این کار توسط ابزار ارائه شده توسط دانشگاه استنفورد، انجام پذیرفت. [19]

برای طبقه بندی مناسب متون با استفاده از الگوریتم های موجود یادگیری ماشین باید پس از تخصیص وزن مناسب به کلمات موجود در یک متن با استفاده از فیلتر StringToWordVector **آنها را به یک بردار عددی مناسب تبدیل نماییم .نکته ی قابل توجه اینجاست که متون به صورت مجموعه ای از جملات طولانی است که باید در ابتدا آنها را به کلمات مجزا تقسیم نماییم .این عمل با استفاده از فیلتر** StringToWordVector **در ابزار** Weka **قابل انجام است .**

در ادامه، در الگوریتم ۴-۱، روش ارائه‌شده برای انتخاب ویژگی‌ها با روش اول انتخاب ویژگی بیان گردیده است.

```
Model: Feature Selection Model
Input: Genre Classified Subtitles from METASUB Dataset
Method:
        For each genre G
                FOR each subtitle
                        Extract Subtitle Sentences
                        FOR each sentence
                        Use POS Tagging
                        END FOR
                        Extract good Parts of Speeches as candidate features
                END FOR
                Convert G.AllDocuments to Numeral Vector
                Use TF-IDF and IFOGAIN for selecting features
        END FOR
Output: Feature Seed Set for All Genres
```

الگوریتم ۴-۱- انتخاب ویژگی ها براساس فاکتور INFOGAIN و فیلتر StringToWordVector

روش دوم استخراج ویژگی ها:

در راهکار بیان شده در روش اول استخراج ویژگی ها، ویژگی های انتخاب شده برای هر ژانر، دارای نرخ تکرار بالا هستند، اما ممکن است این کلمات در چندین ژانر دارای نرخ تکرار بالا باشند، بنابراین، این اشتراک باعث میگردد کلاس ها به یکدیگر نزدیکتر گردند، دقت طبقه بندی کاهش پیدا کند و همپوشانی ژانرها افزایش یابد .ازاین رو، ویژگی های انتخابی بایستی هرچه بیشتر، به کلمات اختصاصی هر ژانر نزدیک گردند .برای این منظور در این روش سعی شده است ارتباط میان کلمات هر ژانر در نظر گرفته شود . از طرفی نمیتوان صرف اینکه یک کلمه در ژانر دیگری نیز تکرار شده است، آن کلمه را از مجموعه ی ویژگی های ژانر جاری خارج نمود .درنتیجه بایستی نقطه ای به عنوان حد آستانه ی شباهت تعریف نمود .این آستانه، میزان شبیه بودن ژانرها با داشتن این کلمه را مشخص مینماید .به

عبارت دیگر این معیار شرط قرار گرفتن یک کلمه در لیست ویژگی های کاندید یک ژانر میباشد . ویژگی هایی که وظیفه دارند تا حد امکان ژانرها را از لحاظ مفهومی از یکدیگر جدا سازند. [19]

در ادامه حد آستانه ی شباهت را به شکل زیر تعریف می نماییم:

$$STL(w) = (GenreCount - 1) * \left(\frac{Freq(w, MainGenre)}{\sum_{i=1}^{RemGenreCount} STC * Freq(w, Genre_i)} \right)$$

در معادله بالا ضریب GenreCount مشخص کننده ی تعداد کل کلاس ها میباشد که در این پژوهش کلاس ها، ژانرها میباشند MainGenre ،. ژانر اصلی است، به عبارت دیگر نمایانگر ژانری است که قصد دارد کلمه ی w را به عنوان یکی از ویژگی های کاندید خود معرفی نماید .تابع Freq(w,g)مشخص کننده ی نرخ تکرار کلمه ی w در ژانر g میباشد. RemGenreCount تعداد ژانرهای باقیمانده میباشد . به عبارت دیگر، تمامی ژانرها به جز ژانر اصلی. STC را به عنوان ضریب آستانه ی شباهت تعریف مینماییم .این ضریب تعیین کننده میزان اختلاف در نرخ تکرار برای انتخاب شدن یک کلمه به عنوان ویژگی کاندید یک ژانر میباشد .هرچه این ضریب افزایش یابد میزان اختلاف در نرخ تکرار یک کلمه در ژانر اصلی با نرخ تکرار همان کلمه در هر یک از سایر ژانرها افزایش مییابد .بنابراین با کم نمودن این ضریب تعداد زیادی کلمه عموماً با نرخ تکرار بالا و همچنین با ضریب اشتراک بیشتر با سایر ژانرها به دست خواهد آمد و در مقابل با زیاد نمودن این ضریب تعداد کمتری کلمه با نرخ تکرار بالا و همچنین با ضریب اشتراک کمتر با سایر ژانرها به دست می آید .هر دو مورد ذکرشده میتواند نامناسب عمل نماید به طور مثال در حالت کم بودن STC کلمات آزادانه تر به لیست ویژگی های کاندید اضافه میگردند، بنابراین شاهد لغات نامربوط به ژانر خواهیم بود .از سوی دیگر با افزایش STC کلمات به صورت سختگیرانه تری به لیست ویژگی های کاندید اضافه میگردند، در این حالت نیز ممکن است برخی کلمات مهم ژانر حذف گردند .بنابراین بایستی تا حد امکان، حد وسطی از هر دو بعد را به عنوان مقدار STC در نظر بگیریم .ضریب STC متناسب با نوع مسئله، ژانرهای انتخابی، تعداد لغات، اندازه ی مسئله،به صورت تجربی ارزیابی و انتخاب میگردد .در ادامه به تحلیل مقدار تابع STL(w) میپردازیم. [12]

در صورت کسر، مقدار تکرار کلمه ی w در ژانر اصلی قرارگرفته شده است .این مقدار، در (GenreCount-1) ضرب شده است .هدف این است که نرخ تکرار کلمه ی w از نرخ تکرار این کلمه ضربدر STC بیشتر شود .که در اینجا درصورتی که نرخ تکرار w در ژانر اصلی را در تعداد سایر ژانرها که w با آنها مقایسه میگردد ضرب نماییم و از طرف دیگر حاصل جمع نرخ تکرارw ضربدر STC را استخراج نماییم .درصورتی که این نسبت بزرگتر از 1 گردد، کلمه ی w به لیست ویژگی های کاندید اضافه میگردد.

پس از مراحل بالا، ویژگی های کاندید به ترتیب نزولی نرخ تکرار مرتب میگردند، ویژگی هایی که دارای بیشتر از سه حرف هستند جداسازی میشوند و سپس n عدد از پرتکرارترین آنها جداشده و به عنوان سبد ویژگی های ژانر معرفی میگردند .الگوریتم زیر روند انتخاب ویژگی ها را در این روش نمایش میدهد.

```
Model: Sentiment Analysis Based Feature Selection Algorithm
Input: Genre Classified Subtitles from METASUB Dataset
Method:
        FOR each Genre G
                FOR each subtitle S in G
                        Convert XML subtitles into plain text
                        Extract Subtitle Sentences
                        FOR each sentence E in S
                                Apply POS Tagging
                                Remove StopWords
                                Extract Adjectives, Adverbs and Nouns
                                FOR each word w in adj,adv,noun
                                        Calculate STL(w)
                                        IF(STL(w) > 1 && w.size > 3)
                                                Add w to G.Candidate_Features_List
                                        END IF
                                END FOR
                        END FOR
                END FOR
                Sort G.Candidate_Features_List Descending
                Add top N candidate features from G.Candidate_Features_List to
        G.Final_Features_List
        END FOR
Output: Feature Seed Set for All Genres
```

الگوریتم ۴-۲- مدل انتخاب ویژگی ها بر مبنای لغات یکتا با اعمال حد آستانه ی مشابهت

الگوریتم بالا علاوه بر حذف لغات بی اثر ، به دلیل ساختار یافتن کلمات یکتای پیاده سازی شده در این الگوریتم، به صورت خودکار، کلمات بی اثر اختصاصی زیرنویس ها را نیز حذف مینماید. به طور مثال کلمه ی life در هیچ یک از لیست های کلمات بی اثر، به عنوان کلمه ی بی اثر شناخته نمیشود .اما در این الگوریتم به صورت خودکار به عنوان کلمات غیر یکتا یا بی اثر شناخته میشوند و از کلیه ی لیست ها خارج میگردد .به عبارتی در مبحث تحلیل زیرنویس ها، کلماتی بیاثر هستند که با نسبت یکسانی، در همه ی ژانرها تکرار شده اند .این کلمات نمیتوانند به ما در تشخیص و طبقه بندی ژانرها کمکی نمایند .این نسبت همان STC میباشد. [12,19]

طبقه بندی تک برچسب بر مبنای ویژگی های انتخاب شده در روش های اول و دوم:

طبقه بندی در روش های انتخاب ویژگی اول و دوم برمبنای روش های متداول یادگیری ماشین در مدل های متنکاوی صورت میپذیرد .در طبقه بندی میزان دقت، وابسته به میزان جدا بودن کلاس ها و همچنین میزان شمول ویژگی های انتخاب شده به نمایندگی از کل کلاس میباشد .در بخش قبل روش های انتخاب ویژگی بیان گردید .در فصل دوم الگوریتم های مطرح در یادگیری ماشین توضیح داده شد .در ادامه الگوریتم های استفاده شده در طبقه بندی بر مبنای ویژگی های انتخاب شده در روش های اول و دوم ارائه میگردد .در فصل بعد نتایج حاصل از هریک از الگوریتم ها بیان میگردد.

- الگوریتم یادگیری بیزین ساده
- الگوریتم یادگیری بیزین چندجمله ای
- الگوریتم یادگیری شبکه های بیزین
- الگوریتم یادگیری K نزدیکترین همسایه
- درخت تصمیم گیری J48

ارزیابی و تخمین خطا:

برای ارزیابی دقت عملکرد روش های پیشنهادی، از روش های معمول استفاده میگردد که ادامه بیان خواهند شد .همان طور که در بخشهای قبل نشان داده شد، فیلم های مجموعه ی داده ی MetaSub **به طور میانگین دارای دو ژانر میباشند که هر فیلم حداقل یک ژانر و حداکثر چهار ژانر را دارا میباشد .از طرف دیگر مرجع ارائه ی این اطلاعات یعنی** IMDb **هیچ یک از ژانرهای یک فیلم را به عنوان ژانر اصلی بیان ننموده است .در این پژوهش از پارامترهای زیر برای ارزیابی نتایج استفاده گردیده است.** [12]

- Precision
- Recall
- F Measure

فصل پنجم (ارزیابی و تحلیل نتایج)

انتخاب ژانرهای مناسب:

پس از بررسی مجموعه ی داده ی MetaSub ، اطلاعات ژانرهای این مجموعه به قرار زیر هستند.

جدول ۵-۱- بررسی ژانر زیرنویس‌های موجود در مجموعه‌ی داده‌ی MetaSub

Unique Movies Count		11’087	
Action	1’776	Romance	1’700
Sci-Fi	567	Horror	957
Thriller	1’126	Adventure	1’157
Comedy	3’377	Music	384
Crime	1’608	Documentary	727
Drama	5’868	Western	174
Animation	430	War	378
Family	490	Fantasy	614
History	397	Biography	349
Film-Noir	75	Reality-TV	10
Sport	184	Game-Show	9
Talk-Show	3	News	4
Mystery	774	Musical	626
Average Genres per Movie		2	

همان طور که در جدول مشخص است، برخی از ژانرها از تعداد بسیار کمی برخوردار هستند و همانطور که قبلاً بیان گردید، فیلم ها از تداخل بسیار بالایی در ژانرها برخوردار هستند .بنابراین، پس ازاین

مرحله، مقرر گردید ماتریس تداخل برای ژانرهای بالاتر از 900 نمونه تشکیل گردد و 5 ژانری که کمترین تداخل را با سایر ژانرها دارند انتخاب گردند .در ادامه در معادله، که نشان دهنده ی تداخل هریک از ژانرهای روش ایجاد جدول پرتکرار با یکدیگر هستند بیان گردیده است. [12]

$$M[i,j] = \frac{Cooccurence(\text{MovieGenre[i]}, \text{MovieGenre[i]})}{\text{GenreCount(MovieGenre[i])} + \text{GenreCount(MovieGenre[j])} - Cooccurence(\text{MovieGenre[i]}, \text{MovieGenre[i]})}$$

ماتریس M که نمایش دهنده ی تداخل وزن دار ژانرهای پرتکرار می باشد آمده است:

جدول ۵-۲- ماتریس M نشان‌دهنده‌ی میزان تداخل وزن‌دار ژانرهای با تکرار بالا

	Action	Thriller	Comedy	Crime	Drama	Romance	Adventure	Horror	Sci-Fi
Action	1	0.086485	0.069753	0.148286	0.096543	0.024764	0.201557	0.031321	0.087239
Thriller	0.086485	1	0.014875	0.125566	0.067786	0.014722	0.019652	0.103755	0.060777
Comedy	0.069753	0.014875	1	0.088428	0.153031	0.171435	0.061827	0.032889	0.024416
Crime	0.148286	0.125566	0.088428	1	0.158351	0.027968	0.026736	0.024770	0.010218
Drama	0.096543	0.067786	0.153031	0.158351	1	0.171336	0.055122	0.028791	0.018680
Romance	0.024764	0.014722	0.171435	0.027968	0.171336	1	0.024749	0.008732	0.008452
Adventure	0.201557	0.019652	0.061827	0.026736	0.055122	0.024749	1	0.013423	0.068816
Horror	0.031321	0.103755	0.032889	0.024770	0.028791	0.008732	0.013423	1	0.101952
Sci-Fi	0.087239	0.060777	0.024416	0.010218	0.018680	0.008452	0.068816	0.101952	1

اکنون برای تشخیص میزان تداخل هر ژانر با ژانرهای دیگر بر اساس معادله ی زیر عمل مینماییم، این معادله، محاسبه کننده میزان تداخل تجمعی وزن دار هر ژانر با سایر ژانرها است.

$$Genre\ Conflict\ Score(i) = (\sum_{j=2}^{10} M\,[i,j]) - 1$$

حاصل معادله ی بالا در جدول زیر آمده است:

جدول۵-۳- میزان تداخل تجمعی وزن‌دار هر ژانر با سایر ژانرها (مقدار کمتر، بهتر می‌باشد)

Genre	Conflict Score
Action	0.746
Thriller	0.493
Comedy	0.616
Crime	0.610
Drama	0.749
Romance	0.452
Adventure	0.471
Horror	0.345
Sci-Fi	0.380

همانطور که در جدول بالا قابل مشاهده است، میزان تداخل ژانرها بسیار نزدیک به هم میباشد، بااینحال، پنج ژانر با کمترین میزان تداخل و با حداقل 900 نمونه، انتخاب میگردند، این ژانرها عبارتند از:

- **ژانر وحشت**
- **ژانر جنایی**

- **ژانر ماجراجویی**
- **ژانر عاشقانه**
- **ژانر هیجان انگیز**

پس از مشخص شدن ژانرها، از هر ژانر 1000 نمونه از مجموعه ی داده ی MetaSub برداشت گردید، تنها از ژانر وحشت 926 نمونه برداشت گردید، بنابراین درمجموع داده های مورد پژوهش 4926 نمونه میباشند. [12]

نتایج حاصل از اجرا:

در این بخش نتایج حاصل از اجرا، را بر هریک از روش های انتخاب ویژگی بررسی مینماییم. برای یادآوری، دو روش انتخاب ویژگی ها را مجدداً بیان مینماییم: [12,1]

- **روش اول :استخراج ادات سخن مؤثر بر مبنای الگوهای پردازش زبان طبیعی و سپس استخراج ویژگی ها برمبنای تبدیل لغات به بردار، TF-IDF و فاکتور فیلترینگ InfoGain**
- **روش دوم :استخراج ادات سخن مؤثر بر مبنای الگوهای پردازش زبان طبیعی و سپس استخراج لغات یکتای نزدیک به هر ژانر بر مبنای حد آستانه ی شباهت**

نتایج حاصل از اجرای روش اول انتخاب ویژگی ها(روش های متداول متن کاوی):

در این مرحله، با در نظر گرفتن روش اول انتخاب ویژگی ها، الگوریتم های زیربر روی داده ها اعمال گردید که در ادامه نتایج حاصل از اجرای هر یک بیان گردیده است .روش طبقه بندی، به این شکل بوده است که هر مرتبه، با 90٪ از داده ها یادگیری انجام شده است و با 10% از داده ها که به صورت تصادفی از مجموعه ی زیرنویس ها انتخاب میگردند، آزمایش انجام شده است .سپس داده های آزمایش دوباره به مجموعه بازگردانده میشوند .این روند 10 مرتبه انجام می پذیرد آن را 10FOLD مینامند. [12]

الگوریتم شبکه های بیز:

در ادامه و در جدول زیر، جزئیات پارامترهای ارزیابی الگوریتم شبکه های بیزین برحسب 5 کلاس، بیان گردیده است.

جدول ۵-۴- جزئیات پارامترهای ارزیابی الگوریتم شبکه‌های بیزین برحسب ۵ کلاس

Precision	Recall	F-Measure	Class Name
0.632	0.564	0.596	Adventure
0.509	0521	0.515	Crime
0.401	0.607	0.483	Horror
0.636	0.648	0.642	Romance
0.337	0.198	0.25	Thriller
0.506			Average Precision

در جدول بعدی ماتریس اغتشاش آمده است .در هر سطر از این ماتریس کلاس واقعی بیان گردیده است و در ستون ها کلاس هایی که هر سطر به آنها طبقه بندی گردیده است.

جدول ۵-۵- ماتریس اغتشاش الگوریتم شبکه‌های بیز

Adventure	Crime	Horror	Romance	Thriller	
564	71	217	92	56	Adventure
50	521	139	122	168	Crime
96	69	562	65	134	Horror
93	64	164	648	31	Romance
90	299	321	92	198	Thriller

به همین ترتیب جدول های بالا را برای الگوریتم بیز ساده ، الگوریتم بیز ساده چندجمله ای، الگوریتم K نزدیکترین همسایه و الگوریتم درخت تصمیم گیری J48 نیز رسم میکنیم.

نتایج حاصل از اجرای روش دوم انتخاب ویژگی ها(مبتنی بر تجزیه و تحلیل احساسات):

همانطور که در فصل قبل بررسی گردید، روش دوم انتخاب ویژگی بر مبنای یافتن لغاتی اختصاصی برای هر ژانر میباشد .طبقه بندی ارائه شده نیز برمبنای میزان نزدیکی لغات مهم یک متن زیرنویس جدید با ژانر نامشخص، به ویژگی های انتخاب شده برای هر ژانر میباشد .این روش که الگوریتم آن در فصل قبل ارائه گردید، در اجرا بسیار زمان بر میباشد. در این پژوهش، مجموعه لغات منتخب زیرنویس مجهول، سه مرحله با مجموعه ی ویژگی ها مقایسه گردید .با این تفاوت که در مرحله ی اول مجموعه ی ویژگی های هر ژانر دارای 10 عنصر، در مرحلهی دوم مجموعه ی ویژگی های هر ژانر دارای 15 عنصر و در مرحله ی سوم، مجموعه ویژگی های هر ژانر دارای 20 عنصر تعیین گردید .که در ادامه، جداول مربوط به هر مرحله ارائه گردیده اند. [19]

نتایج حاصل از طبقه بندی بر اساس روش دوم انتخاب ویژگی ها با 10 عنصر به ازای هر ژانر:

در جدول زیر جزئیات پارامترهای ارزیابی الگوریتم طبقه بندی فیلم ها به ژانر آنها بر مبنای اطلاعات متقابل نقطه به نقطه بر اساس روش دوم انتخاب ویژگی به صورت چند برچسبی با استفاده از 10 ویژگی، برحسب 5 کلاس، با آزمودن 600 نمونه به صورت 6 مرتبه ی 100 تایی بیان گردیده است.

جدول ۵-۱۴- جزئیات پارامترهای ارزیابی الگوریتم طبقه‌بندی فیلم‌ها به ژانر آن‌ها بر مبنای اطلاعات متقابل نقطه‌به‌نقطه بر اساس روش دوم انتخاب ویژگی به‌صورت چند برچسبی بر اساس ۱۰ ویژگی به ازای هر ژانر

Precision	Recall	F-Measure	Class Name
0.61	0.547	0.576	Adventure
0.536	0.594	0.563	Crime
0.582	0.572	0.576	Horror
0.576	0.572	0.573	Romance
0.544	0.564	0.553	Thriller
0.553			Average Precision

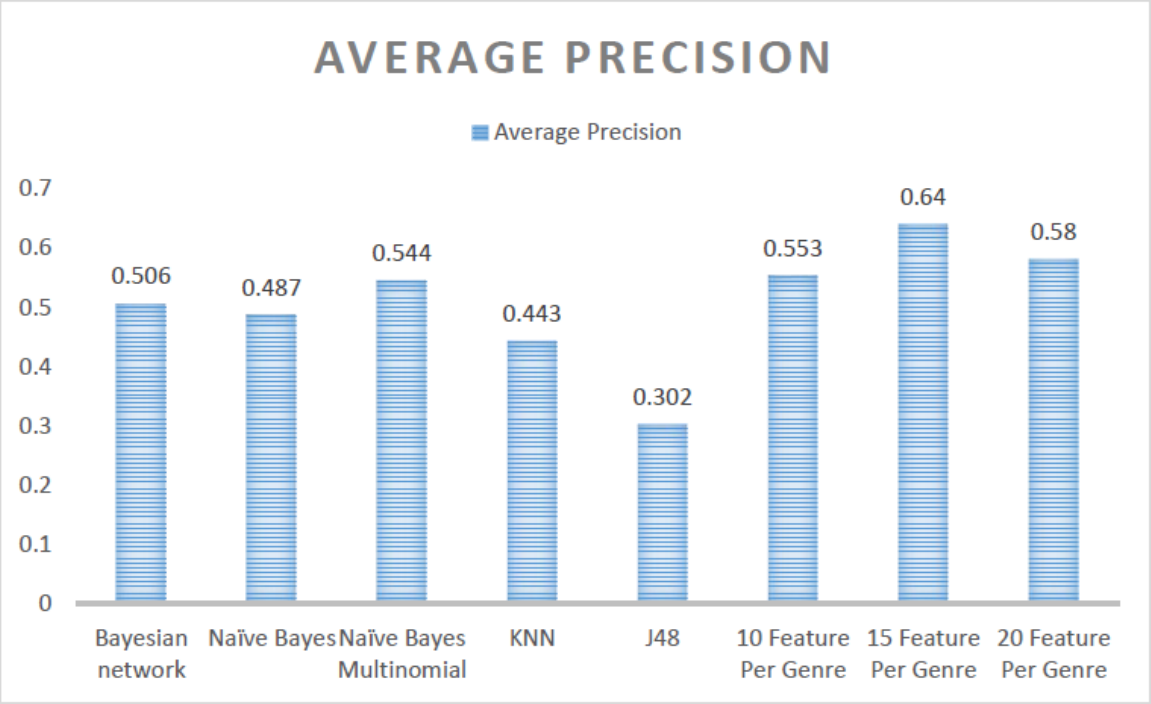
در جدول بعدی ماتریس اغتشاش آمده است .در هر سطر از این ماتریس کلاس واقعی بیان گردیده است و در ستون ها کلاس هایی که هر سطر به آنها طبقه بندی گردیده است.

جدول ۵-۱۵- ماتریس اغتشاش الگوریتم طبقه‌بندی فیلم‌ها به ژانر آن‌ها بر مبنای اطلاعات متقابل نقطه‌به‌نقطه بر اساس روش دوم انتخاب ویژگی به‌صورت چند برچسبی بر اساس ۱۰ ویژگی به ازای هر ژانر

Adventure	Crime	Horror	Romance	Thriller	
69	13	15	15	14	Adventure
12	66	7	14	12	Crime
9	17	67	15	9	Horror
11	17	15	79	16	Romance
12	10	11	14	61	Thriller

به همین ترتیب جداول بالا را با 15 عنصر و 20 عنصر نیز رسم میکنیم. [12]

مقایسه ی نتایج:



شکل ۵-۳- مقایسه‌ی مقدار دقت میانگین تمامی الگوریتم‌های بکار رفته در این پژوهش

اطلاعات کلی و یافته های جانبی:در ادامه برخی اطلاعات کلی از بخشهای مختلف این پژوهش ارائه میگردد. [12]

جدول ۵-۲۰- میزان لغات استفاده‌شده در هر یک از روش‌های طبقه‌بندی ارائه‌شده در این پژوهش

میانگین تعداد لغات در زیرنویس‌های خام	۱۸،۳۵۸ لغت (۱۰۰ ٪)
میانگین تعداد لغات در زیرنویس‌ها پس از استخراج صفات، قیود و اسامی استفاده‌شده در الگوریتم‌های متن‌کاوی متداول	۲،۸۹۶ لغت (۱۵/۷۷ ٪)
میانگین تعداد لغات استخراجی از یک زیرنویس با ژانر نامعلوم برای استفاده در الگوریتم ۳ از فصل ۴	۷۵ لغت (۰/۴ ٪)
میانگین تعداد لغات هر ژانر (بدون اعمال STC)	۵۷،۵۵۳ لغت
میانگین تعداد لغات هر ژانر (با اعمال STC)	۲۱،۲۸۸ لغت

یافته های جانبی:

در این پژوهش علاوه بر یافته های اصلی، در حین مراحل دادهکاوی، برخی اطلاعات جانبی به دست آمده است .این اطلاعات شامل نام متداول افراد در هر یک از ژانرها میباشد .به این معنی که به طور مثال افراد در ژانر وحشت عموماً چه نام هایی دارند .یا به طور مثال در ژانر جرم و جنایت بیشتر در چه شهرهایی اتفاق میافتد. [12]

جدول ۵-۲۱ – اسامی یکتای افراد، به‌دست‌آمده از بررسی کلمات یکتای ژانرها

ژانر هیجان‌انگیز	ژانر عاشقانه	ژانر ماجراجویی	ژانر جنایی	ژانر وحشت
Michael	Kate	William	Paul	Carla
Charlotte	Maria	Robin	Charlie	Richie
Philip	Eve	Wendy	Tony	Samantha
Henri	Linda	Dorothy	Joey	Megan
Mandy	Brian	Harry	Albert	Sharon

جدول زیر برخی کلمات استفاده شده به عنوان ویژگی برای هر ژانر را نمایش میدهد . کلمات حاصل به دست آوردن کلمات یکتای هر ژانر با اعمال حد آستانه ی شباهت و به طور خودکار به دست آمده اند.

جدول ۵-۲۲- برخی کلمات یکتای به‌دست‌آمده به‌عنوان ویژگی برای هر ژانر به‌طور خودکار و با اعمال حد آستانه‌ی شباهت

ژانر هیجان‌انگیز	ژانر عاشقانه	ژانر ماجراجویی	ژانر جنایی	ژانر وحشت
Senator	Love	Captain	Lieutenant	Vampire
Committee	Girl	King	Court	Horror
American	Mother	Master	Crime	Dracula
Journalist	Happy	Ship	Jail	Godzilla
Active	Beautiful	Lord	Client	Zombie
Maintenance	Heart	Fire	Judge	Wicked
Rob	Darling	Power	Bullet	Spritual
Lone	Wedding	War	Gang	Supernatural
Rocker	Marriage	Earth	Jury	Frightening
Moody	Sweet	Gold	Drug	Ghost
Express	Kiss	Boat	Attorney	Formula

در این جدول نام برخی شهرها و مکانها به همراه ژانر مربوطه بیان شده است:

جدول ۵-۲۳- برخی اسامی شهر، کشور و مکان به‌دست‌آمده به‌عنوان ویژگی برای هر ژانر به‌طور خودکار و با اعمال حد آستانه‌ی شباهت

ژانر وحشت	ژانر جنایی	ژانر ماجراجویی	ژانر عاشقانه	ژانر هیجان‌انگیز
Frankenstein	Boston	England	Paris	Moscow
Chelsea	Colombia	Spain	France	Soviet

فصل ششم(نتیجه گیری)

نتیجه گیری:

در این پژوهش، آنالیز احساسات معرفی گردید، برخی از مفاهیم پایه بررسی شد و درنهایت مروری بر فعالیت های گذشته در این حوزه انجام شد .با توجه به تعدد مسائل تحقیقاتی چالش برانگیز و تنوع کاربرد آن، این حوزه در سالهای اخیر، یکی از فعال ترین حوزه ها بوده است .در حقیقت پهنه ی این حوزه، از علوم رایانه تا علوم مدیریت گسترده شده است .سپس به بیان مدل پیشنهادی برای طبقه بندی

فیلم ها به ژانر آنها بر اساس زیرنویس فیلم پرداخته شد و در پایان نتایج ارزیابی روش پیشنهادی ارائه گردید.

به دلیل جدید بودن این زمینه ی تحقیقاتی، همچنین نقص در اکثر راهکارها، ناکافی بودن مجموعه داده های موجود و در بعضی مواقع، عدم وجود مجموعه داده برای زمینه هایی از تحلیل احساسات، فرآیند تحلیل احساسات دارای پیچیدگی هایی گردیده است .از طرف دیگر نوع پژوهش هایی این چنین، دارای ضریب خطا هست به این دلیل که حتی درصدی از انسان ها نیز قادر به تشخیص برخی احساسات در معنی اصلی آنها نیستند .چیزی که در زبان طبیعی، گاهی اوقات از آن به سوءبرداشت نام میبرند.

درنتیجه یکی از بخش های مهم و پیچیده ی این فرآیند، بخش ارزیابی نتایج میباشد .که در پژوهش های مختلف به شکل های مختلفی انجام پذیرفته است.

هدف از ارائه ی این پژوهش، ارائه ی راهکاری برای طبقه بندی محتواهای چندرسانه ای میباشد .در دنیای امروز که با حجم قابل توجهی از تولید محتوا روبرو هستیم، وجود راهکاری سریع و کارآمد برای تعیین مفهوم فیلم ها بدون مشاهده ی کل فیلم، میتواند کاربردهای فراوانی را در پی داشته باشد .در این پژوهش سعی شده است، قدمی در این راستا برداشته شود .با ارائه ی الگوریتم ها و روش های انتخاب ویژگی و مختلف سعی در هرچه بیشتر مشخص شدن موضوع و ابعاد آن گردید.

نوآوری اصلی این پژوهش، بررسی و تشخیص ژانر فیلم ها بر اساس زیرنویس آنها بود که تاکنون هیچگونه مقاله ای از این مبحث به چاپ نرسیده است .دیگر نوآوری این پژوهش ارائه مجموعه ی داده ی MetaSub **میباشد .این مجموعه با بیش از** 11000 **زیرنویس فیلم به همراه نام فیلم، کارگردان، نویسنده، تهیه کننده، بازیگران، ژانر، مدت زمان و سال ساخت فیلم منبع مناسبی برای داده کاوی در حوزه ی چندرسانه ای میباشد .از دیگر نقاط قوت این پژوهش میتوان به استفاده از تجزیه و تحلیل احساسات در سنجش و طبقه بندی فیلم ها به ژانر آنها اشاره نمود که تجزیه و تحلیل احساسات یکی از مباحث جدید در پژوهش میباشد.** [12]

تمامی فعالیت های تحلیل احساسات، بسیار چالش برانگیز هستند .درک و دانش ما از این مسئله و راه حل های آن هنوز بسیار محدود میباشد .مهم ترین دلیل آن این است که ریشه ی تمامی این مسائل، به مسئله ی پردازش زبان طبیعی بازمیگردد و پردازش زبان طبیعی، هیچ مسئله ی آسانی ندارد .به هرحال تلاش های مهمی در این حوزه صورت پذیرفته است .شاهد این مسئله، حجم گسترده ی گروه های نوبنیادی است که دائم در حال ارائه ی راهکارهای جدید درزمینه ی تحلیل احساسات هستند .همچنین شبکه های اجتماعی و تجارت ها در حال ایجاد راهکارهایی برای تحلیل احساسات کاربران و مشتریان خویش هستند .این یک نیاز واقعی و عظیم در صنعت میباشد .به این دلیل که هر کمپانی تجاری، علاقه مند است که بداند مشتریان محصولاتش، چه نظرات و احساساتی نسبت به محصولات آنها و رقبایشان دارند .این نیازها و همچنین چالش های فنی، قطعاً این زمینه را برای سال های آینده زنده و شکوفا نگاه خواهند داشت.

مدل سازی زبان فارسی:

در این پژوهش، زبان مورد بررسی زبان انگلیسی بود .مدل استفاده شده برای پردازش زبان طبیعی و برچسب گذاری ادات سخن در جمله، به دلیل ماژولار بودن، قابلیت تهیه ی مدل تجزیه و تحلیل ادات سخن برای زبان فارسی را دارا میباشد .میتوان با استفاده از چنین مدلی تعداد بسیار زیادی از محتواهای چندرسانه ای به زبان فارسی را با روش ارائه شده در این پژوهش بررسی و کاوش نمود.

سیستم خبره ی پیشنهاد دهنده ی فیلم بر اساس علاقه:

یکی از کاربردهای کاملاً موردنیاز برای افراد، در دوران جاری که با حجم عظیم تهیه ی محتواهای چندرسانه ای مواجه هستیم، کشف سلیقه ی آنها می باشد .به این شکل که کاربر با بیان برخی از فیلم هایی که مشاهده نموده است، بتواند سایر فیلم هایی که با سلیقه ی او سازگار هستند را به عنوان پیشنهاد دریافت نماید .با توجه به حجم بالای زیرنویس های موجود در مجموعه ی داده ی MetaSub که در

این پژوهش تهیه و ارائه گردید، این امکان وجود دارد که کاربران نیاز و سلیقه ی خود از فیلم را مطرح نمایند، سپس سیستم پیشنهاد گر موردبحث، با مطالعه و تجزیه و تحلیل زیرنویس فیلم ها، متناسب با سلیقه ی کاربر، به وی فیلم پیشنهاد دهد یا او را از مشاهده ی فیلمی بازدارد. [12]

داده کاوی مکانی:

مدل پیشنهادی دیگر، استفاده از روش های داده کاوی مکانی بر روی زیرنویس فیلم ها میباشد .داده کاوی را علم استخراج اطلاعات مخفی شده در درون اطلاعات روزمره میدانند .یکی از اطلاعاتی که در زیرنویس ها کمتر موردتوجه قرار میگیرد، زمانبندی دقیق جملات و دیالوگ های فیلم میباشد .به گونه ای که در هرلحظه از فیلم، دیالوگ در حال صحبت، قابل ردگیری میباشد .میتوان ایده های فراوانی برای استفاده از این قابلیت زیرنویس های موجود در مجموعه ی داده MetaSub **ارائه داد.** [12]

پیشنهاد برای پژوهش در حوزه ی تجزیه و تحلیل احساسات:

علاوه بر کاربردها و پژوهش های جدید پیشنهاد شده در بخش قبل که مربوط به پژوهش جاری بود، مسائلی برای بهینه سازی حوزه ی جدید تجزیه و تحلیل احساسات وجود دارد که به صورت اختصار مطرح میگردند:

- **نیاز به مدل سازی بهتری در جملات ترکیبی وجود دارد. مدل سازی بهتر در سطح جمله به معنی افزایش دقت محاسبات در همه ی عبارات احساسی ، شیفت دهنده ها و ساختار جملات است.**
- **هر محصول دارای نام های مختلفی است که در متون مختلف یا حتی یک متن از آنها استفاده میشود.مشکل نامگذاری بدون ابهام اشیا هنوز برطرف نشده است.**

- یک متن در مورد اشیا گوناگونی بحث میکند ،تشخیص این امر که متن به کدامیک از اشیا مربوط است بسیار حیاتی است .نتایج مربوط به این امر تاکنون رضایت بخش نبوده اند.

– اگرچه رهیافت هایی که از روش های طبقه بندی برای تشخیص کنایه ها استفاده میکنند، وجود دارد ولی این رهیافت ها هنوز با سامانه های خودکار تشخیص احساسات هماهنگ و یکپارچه نشده اند.

- متون دارای داده های ناهماهنگ مانند غلط های املایی ، اشکالات گرامری و ... یک چالش بزرگ برای سامانه های تحلیل احساسات هستند.

- سامانه های تحلیل احساسات کنونی بر روی جملات فاعلی تأکید دارند و از جملات مفعولی چشم پوشی نموده اند . نیاز به الگوریتم هایی که برای امتیازدهی به جملات مفعولی احساس میشود .جملات مفعولی با فرکانس بسیار بالایی در مقالات خبری تکرار میشوند.

- متون فارسی، یکی از مهم ترین بخش هایی است که میتوان کاوش های متعددی درزمینه ی تحلیل احساسات در مورد آن انجام داد .این مورد کمتر توسط پژوهشگران موردمطالعه قرارگرفته شده است . در این متون میتوتن به وسیله ی ترجمه ی ماشینی به تحلیل احساسات پرداخت .همچنین میتوان با ادات سخن و زبان شناسی فارسی برای تحلیل احساسات اقدام نمود. [12]

نتیجه گیری نهایی:

به طور خلاصه تا به اینجا متوجه شدیم که دسته بندی چه اثری میتواند در روند یادگیری ماشین بکند . چندین الگوریتم کلاس بند را با هم بررسی کردیم . اما باید توجه داشت که هیچ یک از الگوریتم ها به طور صد در صد برای شرایط های گوناگون مسئله ما کارا و مفید نخواهند بود و شاید ترکیب آنها بتواند کارایی و عملکرد بهتری را در مسائل به وجود آورد . در حال حاضر این موضوع همچنان توسط بسیاری از محققان حوزه ی علوم کامپیوتر در حال تحقیق و بررسی است و همچنان به دنبال الگوریتم هایی کارآمد تر و مفید تر جهت بهبود یادگیری ماشین ها هستند . در ادامه نگاهی به شبکه های عصبی انداختیم ؛ شبکه‌های عصبی مصنوعی ابزاری توانمند و مبتنی بر شبکه های عصبی بیولوژیکی می‌باشند. قابلیت‌های مغز انسان مانند قابلیت یادگیری، تعمیم، خلاقیت، پردازش موازی و ... باعث شد که ساخت ماشینی که این قابلیت‌ها را داشته باشد هدف اصلی بسیاری از دانشمندان و محققان قرار گیرد. با وجود رشد گسترده‌ی توان محاسباتی دیجیتال پردازشگرهای همه منظوره، به‌دلیل قابلیت‌های فراوانی که شبکه‌های عصبی مصنوعی با الهام از شبکه‌ی عصبی انسان دارند، برای کاربردهای خاص مثل پردازش تصویر،آنالیز صدا، شناسایی الگو و ... به کار می‌رود. همچنین به تحلیل احساسات پرداختیم که تمامی فعالیت های تحلیل احساسات، بسیار چالش برانگیز هستند. درک و دانش ما از این مسئله و راه حل های آن هنوز بسیار محدود میباشد. مهمترین دلیل آن این است که ریشه ی تمامی این مسائل، به مسئله ی پردازش زبان طبیعی بازمیگردد و پردازش زبان طبیعی، هیچ مسئله ی آسانی ندارد. به هرحال تلاش های مهمی در این حوزه صورت پذیرفته است. همچنین شبکه های اجتماعی و تجارت ها در حال ایجاد راهکارهایی برای تحلیل احساسات کاربران و مشتریان خویش هستند. این یک نیاز واقعی و عظیم در صنعت میباشد. به این دلیل که هرکمپانی تجاری، علاقه مند است که بداند مشتریان محصولاتش، چه نظرات و احساساتی نسبت به محصولات آنها و رقبایشان دارند.

منابع و مراجع:

[1] : Tom M. Mitchell ,” Machine Learning” , McGraw-Hill Science/Engineering/Math , (march 1,1997)

[2]: Dr.amirhossein Keyhaniipoor(www.fardars.org) ,(n.d) , https://faradars.org/courses/fvml9505-machine-learning

[3] :Mahmudi Aznave Ahmad ,(1393),Shahid Beheshti university , http://facultymembers.sbu.ac.ir/a_mahmoudi/

[4] :Emami Nasibe , Hosseini Zeinab , پیشبینی و تعیین عوامل موثر بر بقای پنج ساله ی کلیه ی پیوندی در داده های نامتوازن با رویکرد فرا ابتکاری و یادگیری ماشین پرورش علائم و داده ها , , NO. 4 ,pp :85-93 , 1397

[5] : Erfaan Ebrahim esfahani , (1396/8/20), https://bit.ly/2CHe8Lu

[6] : Dr.arman ribaad (www.faradars.org) ,(1398/2/8), https://bit.ly/30fnW7H

[7] Fundamentals of neural networks architectuees, algorithms and applications, by Laurene Fausett,1994

[8] Michael Oren, Constantine Papageorgiou, Pawan Sinha, Edgar Osuna, Tomaso Poggio, "Pedestrian Detection Using Wavelet Templates", CBCL and AI Lab, Cambridge, MA 02139, june 1997. Computer vision and pattern recognition, Puerto, MIT.

[9] Imani, M.; Ghassemian, H. "Band Clustering-Based Feature Extraction for Classification of Hyperspectral Images Using Limited Training Samples" Geoscience and Remote Sensing Letters, IEEE Year: 2014, Volume: 11, Pages: 1325 – 1329.

[10] Wenlong Fu; Johnston, M.; Mengjie Zhang "Low-Level Feature Extraction for Edge Detection Using Genetic Programming" Cybernetics, IEEE Transactions on, Year: 2014, Volume: 44, Pages: 1459 – 1472.

[11] Najjar, M.A.; Karlapudi, S.; Bayoumi, M.A."Memory-Efficient Architecture for Hysteresis Thresholding and Object Feature Extraction" Image Processing, IEEE Transactions on, Year: 2011, Volume: 20, Pages: 3566 – 3579.

[12]Ali Samimi, “Tabaghe bandi va tahlil ehsasat ba estefade az alogoritm haye machine learning”, Zemestan 1396.

[13]Elham Hesaraki, ordibehesht 1397 (www.faradars.org)

https://blog.faradars.org/understaing-support-vector-machine-example-code/

https://blog.faradars.org/support-vector-machine-simplified/

[14] Dr. Farshid Shirafkan (www.faradars.org)

https://faradars.org/courses/fvdm9406s08-decision-trees-in-machine-learning

[15] L. Bing and L. Zhang, Minig text data, A survey of opinion mining and sentiment analysis. Chicago, IL: Springer Science, 2012.

[16] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and trends in information retrieval*, vol. 2, pp. 1-135, 2008.

[17] R. Kohavi and J. R. Quinlan, "Data mining tasks and methods: Classification: decision-tree discovery," in *Handbook of data mining and knowledge discovery*, 2002, pp. 267-276.

[18] T. Mullen and N. Collier, "Sentiment Analysis using Support Vector Machines with Diverse Information Sources," in *EMNLP*, 2004, pp. 412-418.

[19] H. Yu and V. Hatzivassiloglou, "Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences," in *Proceedings of the 2003 conference on Empirical methods in natural language processing*, 2003, pp. 129-136.

[20] : (no.N), (n.d) , https://archive.ics.uci.edu/ml/datasets/iris

[21] : Hesaraki Elham (www.faradars.org),(1397/7/22), https://blog.faradars.org/introduction-to-machine-learning/