دانشکده مهندسی کامپیوتر

تشخیص ناهنجاری با استفاده از شبکه‌های عمیق

گزارش سمینار کارشناسی ارشد
در رشته مهندسی کامپیوتر-گرایش هوش مصنوعی و رباتیک

نام دانشجو:
علی نادری پاریزی

استاد راهنما:
دکتر محسن سریانی

اردیبهشت ماه ۱۴۰۱


بسم الله الرحمن الرحیم

چکیده

تشخیص ناهنجاری مسئله مهمی است که در زمینه‌های تحقیقاتی گوناگون مورد مطالعه قرار می‌گیرد و کاربردهای بسیار زیادی دارد. یک نیاز مرسوم در حوزه تجزیه و تحلیل داده‌های دنیای واقعی، پی بردن به این است که بدانیم کدام نمونه‌ها از نقطه نظر تشابه رفتار و ظاهر با اکثریت نمونه‌های موجود بسیار متفاوت هستند. این تفاوت می‌تواند به دلیل خطای اندازه‌گیری در هنگام جمع آوری داده‌ها باشد. گاهی اوقات این تفاوت می‌توانند نشان دهنده وجود پدیده‌ای ناشناخته باشد که در پشت‌پرده جامعه آماری مورد مطاالعه در حال رخ دادن است و ما از آن بی‌خبر هستیم.

در علم داده اصطلاح ناهنجاری به داده‌ای تعلق می‌گیرد از نقطه‌نظر یک معیار تشابه تعریف شده، میزان تشابه آن با سایر دادگان موجود بسیار کم باشد. برای مثال اگر عکس رادیولوژی فردی که بیماری ریوی دارد را با عکس‌های رادیولوژی گرفته شده از ریه افراد سالم مقایسه کنیم متوجه تفاوت این عکس با سایر عکس‌ها خواهیم شد. این عدم تشابه در دادگان، مشخص می‌کند که فرد دچار بیماری ریوی است. درواقع پزشکان با مشاهده این عدم شباهت‌ها به وجود بیماری پی می‌برند. عمل مقایسه دادگان می‌تواند به وسیله کامپیوتر نیز انجام شود که موضوع این سمینار است.

در این سمینار تلاش شده روش‌های مبتنی بر یادگیری عمیق برای تشخیص ناهنجاری را برسی کنیم. از آنجا که کاربرد این موضوع در حوزه‌های گوناگون بسیار وسیع است و مقالات بسیار متعددی در رابطه با کاربردی‌های مختلف به چاپ رسیده، سعی کردیم حوزه سمینار را محدود کرده و ضمن معرفی انواع کاربردهای مسئله تشخیص ناهنجاری، به بررسی روش‌هایی بپردازیم که در رابطه با کاربرد پردازش تصویر و بینایی کامپیوتر هستند. با توجه به تعدد مقالات در سال‌های اخیر و وجود مقالات جامع در این حوزه، بیشتر مقالات جدید که در سال‌های ۲۰۱۹ میلادی و بعد از آن منتشر شده‌اند را بررسی کنیم و برای باقی روش‌ها به ارجاع دهی به مقالات دیگر اکتفا کنیم.

واژه‌های کلیدی: تشخیص ناهنجاری، پردازش تصویر، شبکه‌های عمیق

فهرست مطالب

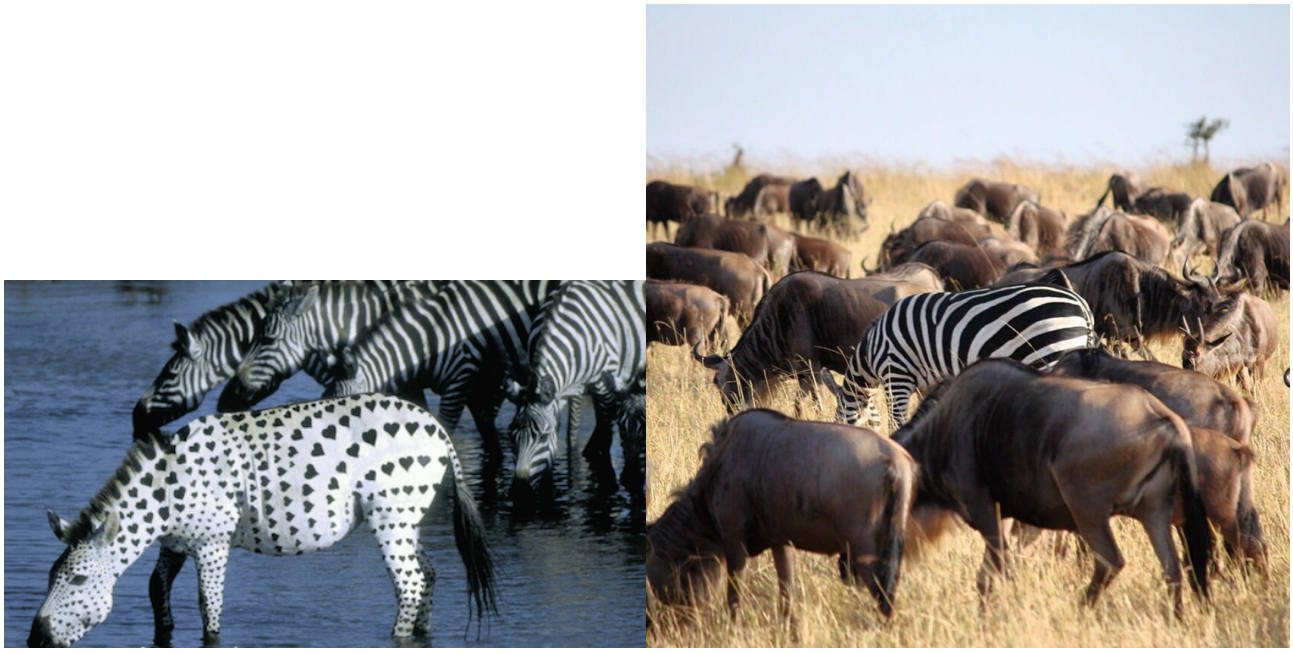
فهرست تصاویر

فهرست جداول

فصل ۱

مقدمه

تشخیص ناهنجاری[۱] مسئله مهمی است که در زمینه‌های تحقیقاتی گوناگون مورد مطالعه قرار می‌گیرد و کاربردهای بسیار زیادی دارد. یک نیاز مرسوم در حوزه تجزیه و تحلیل داده‌های دنیای واقعی، پی بردن این است که بدانیم کدام نمونه‌ها از نقطه نظر تشابه رفتار و ظاهر با اکثریت نمونه‌های موجود بسیار متفاوت هستند. این تفاوت می‌تواند به دلیل خطای اندازه‌گیری در هنگام جمع آوری داده‌ها باشد. گاهی اوقات این تفاوت می‌توانند نشان دهنده وجود پدیده‌ای ناشناخته باشد که در پشت‌پرده جامعه آماری مورد مطاالعه در حال رخ دادن است و ما از آن بی‌خبر هستیم.



شکل ۱.۱: مثالی از تفاوت دادگان ناهنجار و نوین

در کنار ناهنجاری‌ها، دادگان دیگری نیز وجود دارند که با دادگان عادی متفاوت‌اند امّا این تفاوت به اندازی کافی زیاد نیست. به این دادگان اصطلاحا دادگان نوین[۲] گفته می‌شود. دادگان نوین درواقع دادگانی هستند که در دسته دادگان عادی قرار می‌گیرند اما چون هنوز کشف نشده‌اند به نظر می‌رسد که با دادگان عادی تفاوت داشته باشند. برای مثال، اکثر ببرهای دیده شده و شناخته شده به رنگ نارنجی و با خطوط راه راه سیاه هستند و دیدن ببر سفید برای ما تعجب آور خواهد بود. امّا همه به خوبی می‌دانیم که ببر سفید درواقع یک ببر است که فقط رنگ آن غیرعادی است و نباید آن را در دسته جدایی

[1] Anomaly detection
[2] Novelties

از حیوانات قرار داد.

در ادامه این فصل پس از تعریف ناهنجاری در دادگان، به بیان کاربردهای این بحث در حوزه‌های مختلف می‌پردازیم. سپس یک تعریف معیار که مرتبط با حوزه مورد نظر ما که همان پردازش تصویر است ارائه می‌دهیم. پس از تعریف حوزه مورد مطالعه و بررسی اهمیت موضوع، به توضیح ساختار کلی گزارش این سمینار خواهیم پرداخت.

۱.۱ مسئله تشخیص ناهنجاری

تشخیص ناهنجاری که با عنوان تشخیص دادگان خارج از محدوده[۳] نیز شناخته می‌شود، به عملیاتی گفته می‌شود که طی آن به آشکارسازی نمونه‌هایی از مجموعه دادگان می‌پردازد که تفاوت زیادی با اکثریت دادگان موجود دارد. در واقع، اینجا تفاوت به معنی متفاوت بودن مشخصات و ویژگی‌های این نمونه‌ها با الگوی معمول موجود در مجموعه دادگان است. این مسئله یک موضوع فعال تحقیق در دهه‌های اخیر بوده که تقریبان از سال ۱۹۶۰ میلادی تا کنون مورد مطالعه قرار گرفته است [۶]. کاربردهای تشخیص ناهنجاری بسیار وسیع است و در حوزه‌های گوناگونی مورد استفاده قرار می‌گیرد.

ناهنجاری‌ها انواع مختلفی دارند که بسته به کاربرد و مفاهیم مختلف تعریف می‌شوند. به طور کلی می‌توان برای ناهنجاری‌ها سه نوع مختلف درنظر گرفت که عبارت اند از ناهنجاری نقطه‌ای[۴]، ناهنجاری مفهومی[۵]، ناهنجاری مجموعه‌ای[۶]. اکثرکارهای انجام شده در متون علمی در مورد ناهنجاری نقطه‌ای بحث شده است. در این گونه ناهنجاری دادگان به صورت نقاطی در فضا درنظر گرفته می‌شوند و دادگان ناهنجار، نقاطی در فضای مورد نظر هستند که با دیگر دادگان فاصله دارند و رفتاری تصادفی از خود نشان می‌دهند که اغلب تفسیر خاصی ندارند. برا مثال مبلغ بسیار بالای تراکنش در یک رستوران یک تراکنش غیر عادی به حساب می‌آید که با در نظر گرفتن آن در فضای بازنمایی دادگان این نقطه شباهتی به دیگر دادگان نخواهد داشت. دسته دوم ناهنجاری‌های مفهومی هستند که در این دسته مفهموم داده در یک مکان و یا زمان مختلف می‌تواند به صورت ناهنجاری درنظر گرفته شود. برای مثال عبور وسیله نقلیه در خیابان یک امر طبیعی است اما عبود وسایل نقلیه در مسیر عابرین پیاده یک پدیده غیرعادی است. نوع سوم ناهنجاری‌ها که اصطلاحا ناهنجاری مجموعه‌ای گفته می‌شود،مفهوم ناهنجاری از در یک سلسله از رویدادها دنبال می‌کند در حالی که هر رویداد یک داده کاملا عادی است. برای مثال در دنباله تراکنش‌های یک کارت اعتبار وجود چندین تراکنش یکسان با فاصل زمانی بسیار کم مشکوک است.

۲.۱ ساختار کلی روش‌های تشخیص ناهنجاری

اگر بخواهیم روش‌های تشخیص ناهنجاری را به صورت عمومی توصیف کنیم، می‌توانیم بگوییم که این روش‌ها از سه بخش اصلی تشکیل شده‌اند. بخش اوّل یادگیری بازنمایی داده‌ها[۷] است. در این مرحله نگاشتی از دادگان ورودی به فضایی معیین آموخته می‌شود. این نگاشت را می‌توان به صورت تابعی مانند زیر تعریف کرد.

$$f(.;\theta) : x \to y \tag{۱.۱}$$

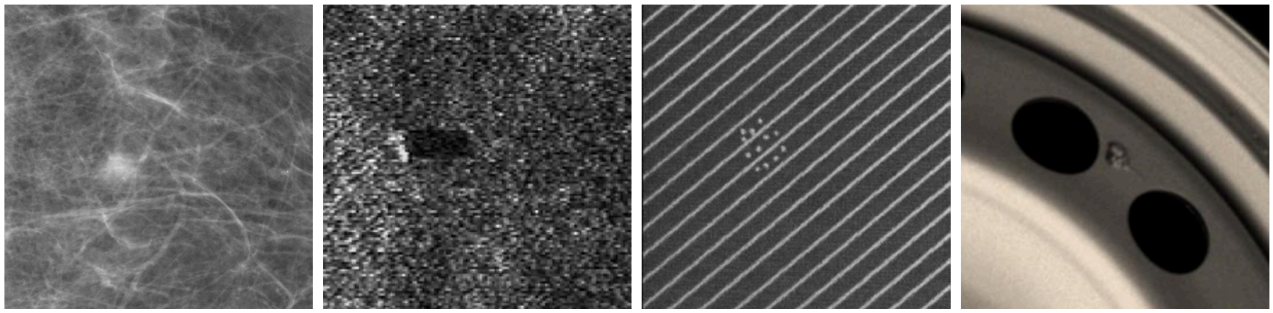
[3] Outlier detection
[4] Point anomaly
[5] contextual anomalies
[6] collective anomalies
[7] Data representation

May-22	1:14 pm	FOOD	Monaco Café	$1,127.80
May-22	2:14 pm	WINE	Wine Bistro	$28.00
...				
Jun-14	2:14 pm	MISC	Mobil Mart	$75.00
Jun-14	2:05 pm	MISC	Mobil Mart	$75.00
Jun-15	2:06 pm	MISC	Mobil Mart	$75.00
Jun-15	11:49 pm	MISC	Mobil Mart	$75.00
May-28	6:14 pm	WINE	Acton shop	$31.00
May-29	8:39 pm	FOOD	Crossroads	$128.00
Jun-16	11:14 am	MISC	Mobil Mart	$75.00
Jun-16	11:49 am	MISC	Mobil Mart	$75.00

Point Anomaly

Collective Anomaly

شکل ۲.۱: ناهنجاری نقطه‌ای و دنباله‌ای [۳]



به ترتیب از سمت چپ، توده سرطان سینه، مین زیردریایی، نقص رنگ‌آمیزی کاشی تولید شده در کارخانه،نمونه نقص موجود در چرخ خودرو.

شکل ۳.۱: مثال‌هایی از ناهنجاری در تصاویر [۵]

در بخش دوّم به تعریف یک معیار سنجش پرداخته می‌شود که برای ارزیابی خروجی مرحله قبل استفاده می‌شود. این معیار با دریافت خروجی مرحله اوّل یک امتیاز برای سنجش میزان تعلق داده ورودی به دسته ناهنجار اختصاص می‌دهد که به آن امتیاز ناهنجاری[8] گوییم.

$$d(f(x);\eta) : f(x) \to d \ , \ d \in \mathbb{R} \tag{۲.۱}$$

در آخر نیز با درنظر گرفتن یک مقدار آستانه δ، به تصمیم‌گیری در مورد داده ورودی با توجه به امتیاز اختصاص داده شده در مرحله دوّم پرداخته می‌شود.

$$\begin{cases} anomaly & d \geq \delta \\ not\ anomaly & d < \delta \end{cases}$$

با توجه به این تعریف، رویکردهای موجود می‌توانند انواع زیر را داشته باشند:

۱. غیر پارامتری: نیازی به یادگیری θ و η و δ نیست.

۲. یک مرحله‌ای: تنها یکی از مجموعه پارامترهای موجود θ یا η یادگرفته می‌شوند.

۳. دو مرحله‌ای: هر دو مجموعه پارامتر θ و η به صورت مستقل و جداگانه یادگرفته می‌شوند.

۴. ادغامی[9]: هر دو مجموعه پارامتر θ و η باهم یادگرفته می‌شوند.

درصورت عدم وجود برچسب‌های دادگان موجود، ناچار به استفاده از روش بدون ناظر هستیم که در آن از هیچ گونه اطلاعاتی در مورد ماهیت دادگان استفاده نمی‌شود. در این گونه مواقع معمولا δ از پیش تعریف شده است و یا همراه با η یادگرفته می‌شود. در حالتی که تنها بخشی از دادگان برچسب خورده باشند و باقی برچسب نخورده، می‌توانیم از رویکرد یادگیری با نظارت ضعیف استفاده کرد. در این مورد نیز مقدار آستانه می‌تواند با استفاده از تنظیم دقیق مدل بدست آید.

۳.۱ ساختار گزارش

در فصل اوّل این سمینار به معرفی حوزه سمینار و تعریف مسئله پرداخته شد و در فصل دوّم به تعریف مفاهیم و اصطلاحات استفاده شده در این حوزه خواهیم پرداخت. فصل سوّم نیز در رابطه با بررسی کارهای مرتبط با این سمینار و معرفی و بررسی جزئی از روش‌ها و مقالات موجود چاپ شده در سال‌های اخیر خواهد پرداخت. در ابتدای فصل سوّم پس از معرفی کارهای مرتبط یک دسته‌بندی از روش‌های موجود ارائه می‌گردد و در ادامه، ترتیب معرفی و بررسی روش‌های موجود بر طبق این دسته‌بندی خواهد بود. در نهایت یک جمع بندی و نتیجه گیری کلی از روش‌های موجود در هر دسته انجام می‌دهیم و پیشنهاداتمان را در رابطه با استفاده از این روش‌ها بسته به کاربرد مورد نظر ارائه می‌کنیم. در فصل آخر گزارش پیشنهادات خود را درباره کارهای آینده این حوزه ارائه کرده و در نهایت پیشنهاد انجام پروژه کارشناسی ارشد را که در راستای همین سمینار است معرفی می‌کنیم.

[8] Anomaly score
[9] Integrated

فصل ۲

مروری بر روش‌های سنتی

اگر به‌یاد داشته باشید، در ابتدای فصل یک به این نکته اشاره شد که مسئله تشخیص ناهنجاری، یک موضوع فعال تحقیق در چند دهه اخیر است که یکی از مقالات معتبر چاپ شده آن مربوط به دهه ۱۹۶۰ میلادی می‌شود. از این رو، در طی این مدت بسیاری از روش‌ها برای یافتن دادگان خارج از محدوده معرفی و توسعه داده شده‌اند که از یادگیری عمیق استفاده نمی‌کنند. این روش‌ها به صورت عمده دادگان را مجموعه‌ای از نقاط در یک فضای چند بعدی فرض می‌کنند و تلاش آنها برای این است که نقاط خارج از محدوده را در این فضا با توجه به ویژگی‌ها و مشخصات دیگر نقاط آشکار کنند. عمدتاً این اینگونه روش‌ها را می‌توان از نقطه‌نظر ایده اصلی به سه دسته کلی استفاده از رده‌بندی، معیار فاصله و مدل‌های آماری تقسیم کرد[۱]. در ادامه به مرور کلی این روش‌ها خواهیم پرداخت. با توجه به اینکه تمرکز ما بر بررسی کامل این روش‌ها نیست پیشنهاد می‌شود برای آشنایی بیشتر با این‌گونه روش‌ها به مقاله چاندولا و همکاران مراجعه کنید [۴].

دسته‌بندی روش‌های سنتی در تشخیص ناهنجاری			
مبتنی بر	خلاصه ایده	انواع	روش‌های شناخته شده
رده‌بندی	یادگیری یک مرز تفکیک میان دادگان عادی و ناهنجار	یک کلاسه	One-class SVM SVDD
		چند کلاسه	-
معیار فاصله	اقدام به تعریف یک معیار فاصله می‌کند تا دادگان عادی را از دادگان ناهنجار جدا کند	فاصله تا نزدیک ترین همسایه	LOC[2] COF
		خوشه بندی و سنجش فاصله تا نزدیک ترین خوشه	K-means CBLOF
		استفاده از تصویر سازی نقاط در فضایی با ابعاد کمتر	PCA Isolation Forest
مدل آماری	دادگان عادی در نواحی پر احتمال مدل آماری قرار می‌گیرند	روش‌های پارامتری	Gausian Mixture Model
		روش‌های غیر پارامتری	Kernel destiny estimator

جدول ۱.۲: دسته‌بندی روش‌های سنتی

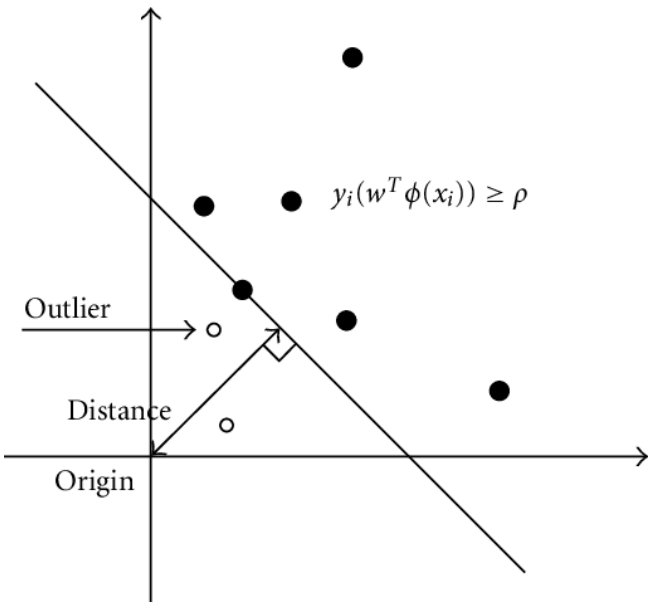
[۱] ر.ک جدول ۱.۲

۱.۲ روش‌های مبتنی بر رده‌بندی

همانطور که در ابتدای این بخش گفته شد، یکی از ایده‌های کلی در روش‌های مورد استفاده برای تشخیص ناهنجاری استفاده از ایده رده‌بندی است. در اینگونه روش‌ها تلاش می‌شود یک مرز تفکیک میان دادگان عادی و دادگان ناهنجار رسم شود. اگر چنین مرزی وجود داشته باشد، می‌توانیم با استفاده از الگوریتم‌های رده‌بند موجود اقدام به یافتن این مرز کرد و سپس با استفاده از مدل آموزش دیده اقدام به آشکارسازی داده‌های ناهنجار کرد. همانطور که مشخص است در این گونه روش‌ها تنها یک دسته برای دادگان تعریف می‌شود که آن دسته دادگان عادی است. دیگر دادگانی که در این دسته قرار نمی‌گیرند به عنوان دادگان عادی در نظر گرفته می‌شوند. البته استفاده از رویکرد رده‌بندی چند کلاسه نیز در صورت وجود برچسب برای تمامی دادگان امکان پذیر است امّا استفاده از این روش کمتر مرسوم است. یکی از معروف ترین روش‌های مورد استفاده دسته بند بردار پشتیبان یک کلاسه[۳] است.

در ماشین بردار پشتیبان ما به دنبال یافتن یک ابر صفحه جدا کننده میان دو دسته داده موجود هستیم. در الگوریتم بردار پشتیبان یک کلاسه ما درواقع به دنبال یافتن صفحه ای هستیم که دادگان معمول در یک طرف این صفحه قرار بگیرند. در این روش تلاش می‌شود صفحه مورد نظر تا حد امکان به نقاد داده نزدیک باشند. پس از رسم این صفحه، دادگانی که به مبدا مختصات نزدیک تر هستند در دسته ناهنجاری‌ها قرار می‌گیرند [۸].

در اینجا تابع نگاشتی که باید یاد کرفته شود همان تابع کرنل در ماشین بردار پشتیبان است و تابع امتیاز ناهنجاری نیز به صورت اندازه فاصله از مبدا مختصات تعریف می‌شود. شکل ۱.۲ این روش را به تصویر کشیده است. توجه داشته باشید که در اینجا تنها یک دسته برای رده‌بندی تعریف می‌شود که آن دسته دادگان عادی است، پس نیازی به وجود برچسب برای تمامی دادگان نیست و رویکرد ما در اینجا به صورت کاملا بدون ناظر خواهد بود.

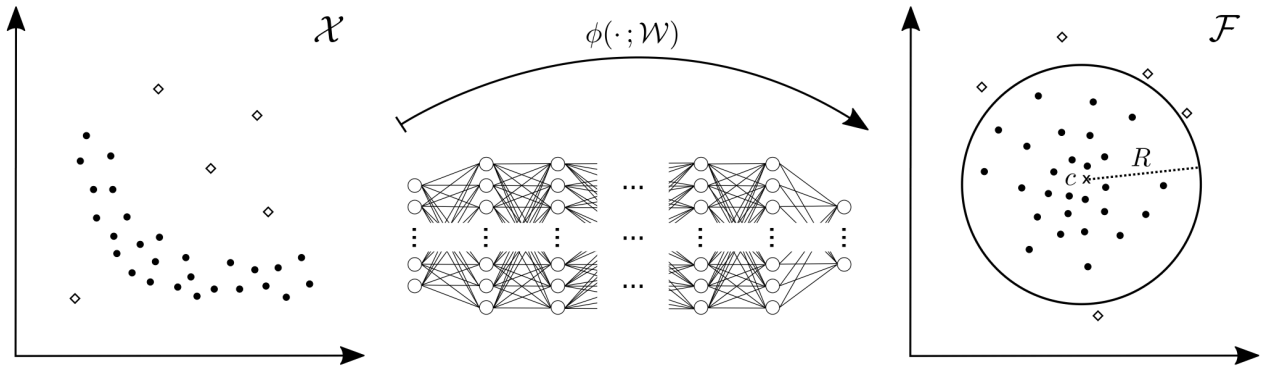


شکل ۱.۲: ماشین بردار پشتیبان یک کلاسه

نمونه دیگری از روش‌های مورد استفاده برای آشکارسازی ناهنجاری که از رویکرد رده‌بندی استفاده می‌کند، بردار پتیبان توصیفگر داده[۴] است. در این روش سعی می‌شود کره‌ای با کوچک ترین اندازه ممکن حول دادگان موجور رسم شود. پس از

[3] One-class SVM
[4] Support Vector Data Description (SVDD)

رسم این کره، دادگانی که در خارج از آن قرار می‌گیرند به عنوان داده ناهنجار شناخته خواهند شد [۷] .



شکل ۲.۲: بردار پشتیبان توصیفگر داده عمیق [۷]

ازجمله مزیت‌های این رویکرد،آموزش سریع، و دقت بهتر آن در مواقعی است که دادگان برچسب خورده در اختیار هستند. و از معایب این روش در هنگام استفاده از رده‌بندی چند کلاسه می‌توان به نیاز برای چندین دسته داده عادی یاد کرد. همچنین این این رویکردها نیاز به تعیین ابر پارامتر برای مدل یادگیری دارند.

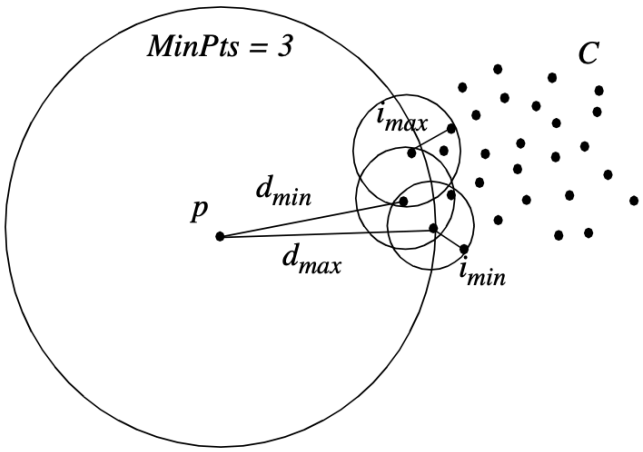
۲.۲ روش‌های مبتنی بر معیار فاصله

اگر به دادگان موجود را به صورت نقاطی بازنمایی شده بر روی صفحه مختصات نگاه کنیم، می‌توانیم از معیار فاصله نقاط از یکدیگر به تصمیم‌گیری در مورد دادگان بپردازیم. در اینگونه رویکردها معمولا اقدام به تعریف یک معیار فاصله می‌کنند تا دادگان عادی را از دادگان ناهنجار جدا کنند. یک نمونه روش معروف که در این دسته می‌گنجد روش معروف عامل پرت محلی[۵] است. در این روش میانگین فاصله هر نقطه از همسایگان محلی محاسبه شده و اگر این میانگین از یک مقدار آستانه بیشتر باشد،داده به عنوان داده ناهنجار شناخته می‌شود. برای سادگی کار، میانگین فاصله نقطه تا تمام همسایگان را بر میانگین فاصله میان همسایگان نقطه محاسبه شده و مقدار آستان برابر با عدد یک درنظر گرفته می‌شود [۲]. در استفاده از این روش نیز نیازی به وجود برچسب دادگان نیست همچنین این روش پارامتری برای یادگیری ندارد و در دسته روش‌های بدون پارامتر نیز قرار می‌گیرد. در واقع این گونه روش‌ها معمولا به صورت بدون ناظر هستند.

۳.۲ روش‌های مبتنی بر مدل آماری

ایده اصلی در این دسته از رویکردها بدین صورت است که، دادگان عادی همواره احتمال رخدادن بالایی دارند، در نتیجه در نواحی از مدل مدل آماری قرار می‌گیرند که احتمال وقوع آنها بیشتر است. برای مثال در روش مدل خطی پویا[۶] ابتدا دادگان را از فضای ورودی به یک فضای از پیش تعیین شده نگاشت می‌کنیم. سپس با استفاده از مدل بدست آمده سعی در پیشبینی مقدار ورودی با توجه به دیگر دادگان موجود می‌کنیم. در اینجا امتیاز ناهنجاری میزان تفاوت مقدار پیشبینی شده و مقدار حقیقی داده است. اگر مقدار اختلاف از یک مقدار آستانه از پیش تعیین شده، که با استفاده از آزمایش با دادگان برچسب خورده بدست آمده، بیشتر باشد، به دسته دادگان ناهنجار تعلق می‌گیرد.

[5] Local Outlier Factor
[6] Dynamic liner model



شکل ۳.۲: نمایش کلی روش عامل پرت محلی [۲]

فصل ۳

روش‌های مبتنی بر یادگیری عمیق

در این فصل ابتدا به معرفی مدل‌های پایه‌ای یادگیری عمیق خواهیم پرداخت که در تشخیص ناهنجاری مورد استفاده قرار می‌گیرند. ینگونه مدل‌ها، پایه و اساس خیلی از روش‌های ارائه شده هستند و آشنایی با آنها به درک بهتر مطلب کمک بسیار زیادی خواهد کرد. پس از معرفی ساختار مورد بحث نمونه‌هایی از کارهای انجام شده که از آن استفاده می‌کنند را به اختصار معرفی خواهیم کرد. جدول ۱.۳ لیستی از تمام روش‌های مورد بحث در این بخش را جمع آوری کرده است.

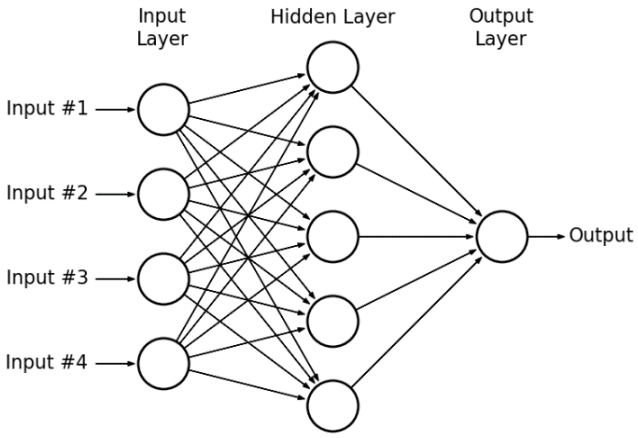
مدل‌های پایه مورد استفاده در روش‌های عمیق برای تشخیص ناهنجاری		
نام مدل	مقاله مرجع	مزیت استفاده
AE		
VAE		
SAE		
DCAE		
DTS		
GAN		

جدول ۱.۳: الگوریتم‌های عمیق مورد استفاده در تشخیص ناهنجاری

۱.۳ مدل پرسپترون چند لایه‌ای

مدل پرسپترون چند لایه‌ای[1] یکی از ابتدایی ترین و مهم‌ترین مدل‌هایی است که می‌توان آنرا نقطه شروعی بر تمام روش‌های عمیق موجود در حال حاضر دانست. این مدل درواقع شبکه‌ای از نورون‌های عصبی مصنوعی[2] است که لایه‌های آن صورت کاملا متصل با یکدیگر ارتباط دارند. این شبک دارای حداقل سه لایه ورودی، مخفی و خروجی است که در آن به غیر از نورون‌های لایه ورودی، باقی نورون‌ها دارای تابع فعال سازی غیر خطی هستند. این مدل برای یادگیری بازنمایی غیر خطی دادگان ورودی معرفی شده و در بسیاری از روش‌های تشخیص ناهنجاری کاربرد دارد.

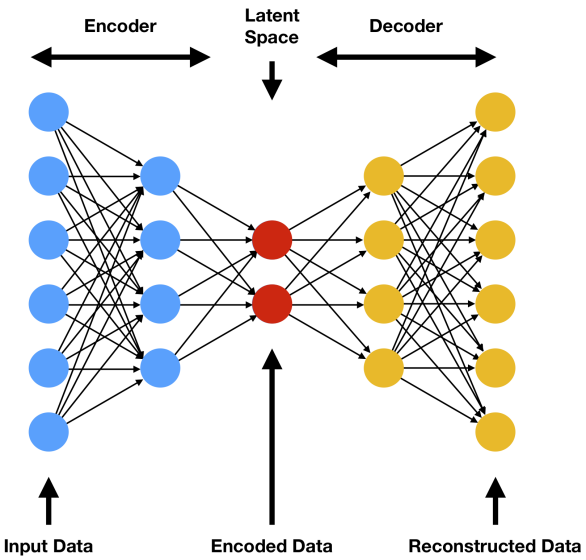
[1] Multilayer Perceptron
[2] Perceptron



شکل ۱.۳: مدل پرسپترون چند لایه

۲.۳ خود رمزکننده

خودرمز کننده[۳]ها نوعی از شبکه‌های عصبی هستند که از روش پس انتشار[۴] برای یادگیری ویژگی‌های مفهومی استفاده می‌کنند. این شبکه‌ها به صورت دو مرحله‌ای اقدام به یادگیری می‌کنند که به ترتیب رمزنگاری و رمزگشایی نام دارند. در مرحله اول داده ورودی به شبکه رمز کنند داده می‌شود و رمز کننده داده ورودی را به یک فضا با ابعاد پایین نگاشت می‌کند. به این فضا به اصطلاح فضای باقی‌مانده یا فضای z می‌گویند. در مرحله دوم، بازنمایی بدست آمده وارد شبکه رمز کنند شده تا داده از فضای باقی مانده دوباره به فضای ورودی باز گردانده شود. آنچه که انتظار می‌رود آن است که خروجی مدل با آنچه در ورودی به مدل داده شده بسیار شبیه باشند. در این صورت قسمت رمز کنند تواسنسته بازنمایی خوبی از داده را در فضای باقی مانده ایجاد کند [۱].



شکل ۲.۳: مدل خود رمز کننده

اگر بخواهیم کارکرد مدل شکل ۳.۳ را با فرمول ریاضی توصیف کنیم، با در نظر داده X به عنوان ورودی مدل، رمز کنند

[3] AutoEncoder
[4] Backpropagation

با گرفتن این ورودی، آن را به فضای باقی مانده و به نقطه z نگاشت می‌کند. اگر تابع رمز کننده را f بنامیم معادله مرحله اول به صورت زیر خواهد بود.

$$f(X, \theta_1) : X \to z \tag{۱.۳}$$

که در اینجا ابعاد فضای z از ابعاد فضای ورودی X کمتر است. این بدان معنی است که در اینجا عمل کاهش ابعاد ورودی صورت گرفته است. اگر رمزگشا را مانند تابعی درنظر بگیریم و آن‌را g بنامیم، این تابع با دریافت ورودی z، اقدام به بازسازی داده ورودی می‌کند.

$$g(z, \theta_2) : z \to X \tag{۲.۳}$$

در کاربردهای تشخیص ناهنجاری معمولا در هنگام استفاه از این معماری، سعی می‌شود از تابع خطای مقایسه ورودی و خروجی مدل برای آموزش مدل استفاده کنند و در فرایند آموزش تنها از دادگان عادی استفاده شود. ایده اصلی در این گونه روش‌ها این است که با توجه به اینکه مدل تنها با دادگان عادی آموزش دیده است، دادگانی که توسط این مدل نتوانند به خوبی بازسازی شوند دارای ناهنجاری بوده‌اند. در واقع در اینجا تابع خطا که همان تابع امتیاز ناهنجاری است به صورت زیر تعریف می‌شود.

$$L(X, g(f(x))) = d \tag{۳.۳}$$

پس از آموزش مدل مقدار آستانه δ برای بدست آوردن بهترین نتیجه با آزمون و خطا و یا روش‌های دیگر مانند استفاده از نمودار حساسیت و دقت تعیین می‌شود.

خودرمز کننده‌ها باید به تغییرات دادگان ورودی حساس باشند تا بتوانند با دقت مطلوب داده رمز شده را بازسازی کنند. همچنین این حساسیت نباید به اندازه‌ای باشد که باعث بشود مدل بجای یادگیری عملکرد مناسب، به حفظ کردن دادگان ورودی بپردازد و دچار بیش‌برازش بشود. برای دستیابی به چنین توازنی، انواع مختلفی از خودرمز کننده‌ها معرفی شده‌اند که با افزودن یک مقدار تنظیم کننده[۵] به تابع خطای اصلی معرفی شده، بدست می‌آیند.

$$L(X, g(f(X))) + regulizer \tag{۴.۳}$$

۱.۲.۳ خود رمز کننده خلوت

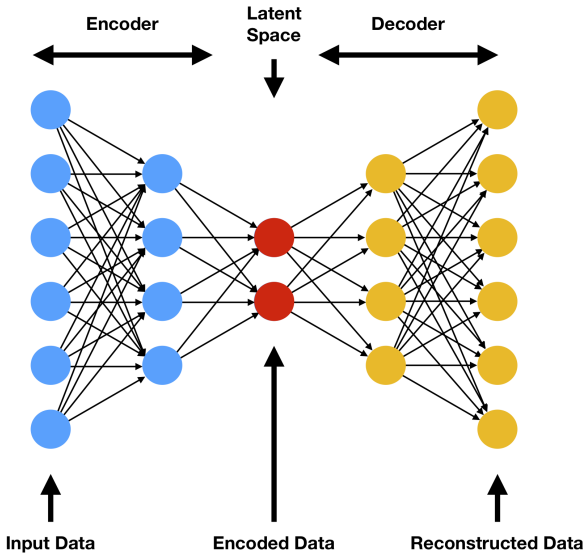
خود رمز کننده خلوت[۶] یکی از انواع خودرمز کننده ها است. ایده اصلی این گونه رمز کننده‌ها این است که نورون‌ها لایه مخفی اگر تعدادشان کمتر از تعداد نورون‌های لایه ورودی باشد شاید نتوانند به خوبی مفاهیم پیچیده را یاد بگیرند. در نتیجه پیشنهاد می‌شود در لایه مخفی تعداد نورون‌های بیشتری قرار گیرند اما از تابع فعال سازی خلوت[۷] استفاده کنند. برای دستیابی به چنین هدفی میتوان از دو نوع تنظیم کنند در تابع خطای مدل استفاده کرد. نوع اول استفاده از تنظیم کننده نرم یک است[۸] که معادله تابع خطا به صورت زیر خواهد بود.

$$L(X, g(f(X))) + \lambda \sum_{i}^{n} |a^{(h)}| \tag{۵.۳}$$

[5] Rgulizer
[6] Sparse AutoEncoder (SAE)
[7] Sparse
[8] L1-Rgulizer

و نوع دوم استفاده از KL-Divergance به عنوان تنظیم کننده است.

$$L(X, g(f(X))) + \sum_{j}^{n} KL(\rho||\hat{\rho}) : \hat{\rho} = \frac{۱}{m} \sum_{i} [a_i^{(h)}(x)] \tag{۶.۳}$$



شکل ۳.۳: مدل خود رمز کننده خلوت

فصل ۴

کارهای آینده

۱.۴ نتیجه گیری

۲.۴ مسائل باز و کارهای قابل انجام

۳.۴ موضوع پیشنهادی برای پایان نامه

کتاب‌نامه

[1] Bhuvaneshwari, M., Kanaga, E. Grace Mary, Anitha, J., Raimond, Kumudha, and George, S. Thomas. Chapter 7 - a comprehensive review on deep learning techniques for a bci-based communication system. In N, Pradeep, Kautish, Sandeep, and Peng, Sheng-Lung, editors, *Demystifying Big Data, Machine Learning, and Deep Learning for Healthcare Analytics*, pages 131–157. Academic Press, 2021.

[2] Breunig, Markus M., Kriegel, Hans-Peter, Ng, Raymond T., and Sander, Jörg. Lof: Identifying density-based local outliers. In *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, SIGMOD '00, pages 93–104, New York, NY, USA, 2000. Association for Computing Machinery.

[3] Chalapathy, Raghavendra and Chawla, Sanjay. Deep learning for anomaly detection: A survey. 01 2019.

[4] Chandola, Varun, Banerjee, Arindam, and Kumar, Vipin. Anomaly detection: A survey. *ACM Comput. Surv.*, 41(3), jul 2009.

[5] Ehret, Thibaud, Davy, Axel, Morel, Jean-Michel, and Delbracio, Mauricio. Image anomalies: A review and synthesis of detection methods. *Journal of Mathematical Imaging and Vision*, 61(5):710–743, 2019.

[6] Grubbs, Frank E. Procedures for detecting outlying observations in samples. *Technometrics*, 11:1–21, 1969.

[7] Ruff, Lukas, Vandermeulen, Robert, Goernitz, Nico, Deecke, Lucas, Siddiqui, Shoaib Ahmed, Binder, Alexander, Müller, Emmanuel, and Kloft, Marius. Deep one-class classification. In Dy, Jennifer and Krause, Andreas, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 4393–4402. PMLR, 10–15 Jul 2018.

[8] Schölkopf, Bernhard, Williamson, Robert, Smola, Alex, Shawe-Taylor, John, and Platt, John. Support vector method for novelty detection. In *Proceedings of the 12th International Conference on Neural Information Processing Systems*, NIPS'99, pages 582–588, Cambridge, MA, USA, 1999. MIT Press.

Abstract

Anomaly detection is a well studied problem in varios fileds of sciance.

Departmant of computer engineering

Deep learning for anomaly detection

Master seminar report
Computer engineering - Artificial intelligence and robotics

Student name:
Ali Naderi Parizi

Professor:
Dr. Mohsen Soryani

April 2022