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 بررسی 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 اخیر 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

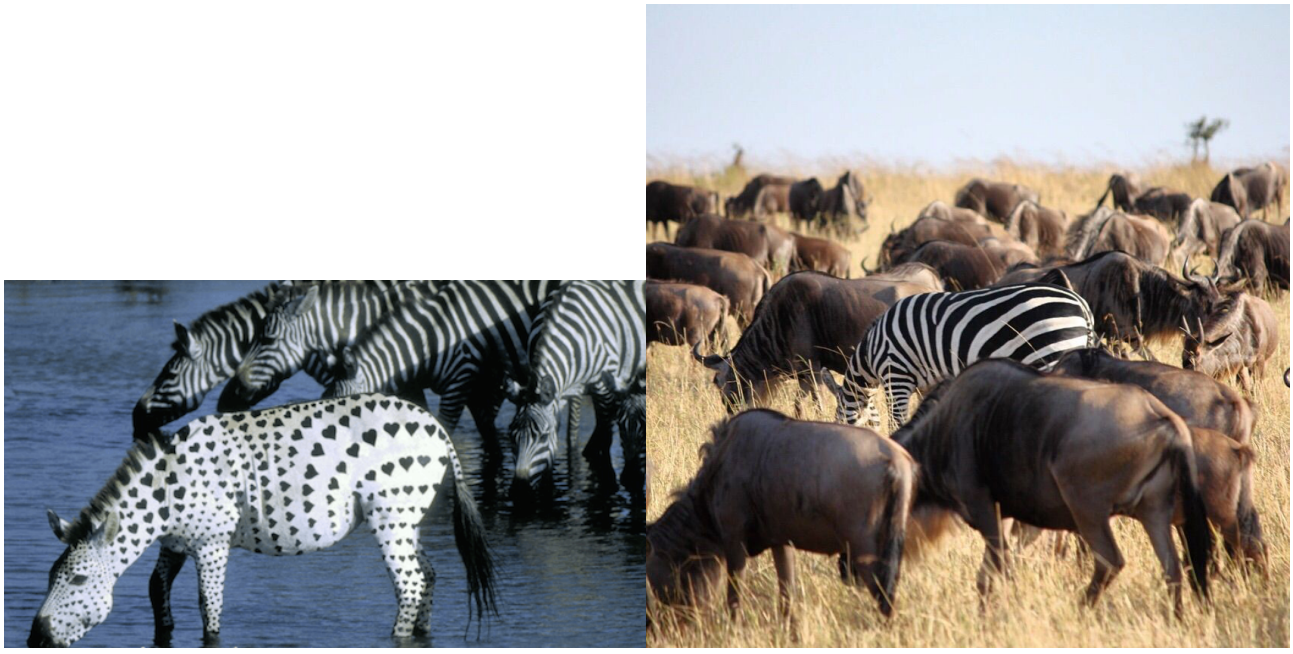
فهرست تصاویر

فهرست جداول

فصل ۱

مقدمه

تشخیص ناهنجاری[1] 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[2] 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 تردد 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 فواصل زمانی بسیار کم مشکوک است.

۲.۱ جنبه‌های مختلف تشخیص ناهنجاری

مسئله تشخیص ناهنجاری را جنبه‌های مختلفی می‌توان مورد بررسی قرار داد. برای مثال می‌توان روش‌های موجود را بر اساس ماهیت دادگان موجود مورد بررسی قرار داد و با توجه به نوع داده انواع روش‌ها را دسته‌بندی کرد. برای نمونه می‌توان ماهیت دادگان را به دودسته کلی، دنباله‌ای (مانند صدا، موسیقی، فیلم، متن و ...) غیر دنباله‌ای (مانند عکس، مشخصات بیمار و ...) تقسیم کرد. و یا بر اساس تعداد ویژگی‌های داده ورودی به دو دسته ابعاد پایین و ابعاد بالا تقسیم کرد. همچنین می‌توان روش‌های تشخیص ناهنجاری‌ها را از دید در دسترس بود برچسب دادگان مورد استفاده بررسی کرد. اما باید توجه داشت که

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

پدیده‌های ناهنجار اصولا کم اتفاق می‌افتند و تعداد آنها در دادگان موجود کم است. با این حال می‌توان روش‌های تشخیص ناهنجاری را از دید رویکرد بر اساس در دسترس بودن برچسب دادگان به سه دسته باناظر، با نظارت ضعیف و همچنین بدون ناظر تقسیم کرد.

۳.۱ کاربردهای مسئله تشخیص ناهنجاری

برای درک اهمیت و کاربرد مسئله تشخیص ناهنجاری می‌توان به حجم مقالات چاپ شده در این حوزه و دامنه وسیع موضوعات تحقیقاتی اشاره کرد که حول این موضوع انجام شده و یا در حال انجام است. در این قسمت برخی از کاربردهای مسئله تشخیص ناهنجاری را به تفکیک حوزه‌های کاربردی مختلف می‌آوریم.

۱.۳.۱ امنیت سیستم و تشخیص نفوذ

تشخیص نفوذ در کاربرد امنیت سایبری که عمل تشخیص و اطلاع پیدا کردن از دسترسی‌های غیر مجاز به شبکه و یا سامانه‌های رایانه‌ای است می‌تواند یکی از کاربردهای مسئله تشخیص ناهنجاری باشد. در اینگونه مسائل با بررسی گزارش‌های سیستم در طول زمان به عنوان داده ورودی به بررسی این قضیه می‌پردازند. همانطور که مشخص است، نوع ناهنجاری در این جا میتواند از دو نوع دنباله‌ای و یا مفهومی باشد.

۲.۳.۱ تشخیص جعل اسناد و کلاهبرداری

تشخیص مدارک جعلی در حوزه‌های مختلف مانند هویتی، بانکی، بیمه، کارت اعتباری و غیره بسیار کارآمد است. در اینگونه کاربردها نیز مدارک از جنبه‌های مختلفی با یکدیگر مقایسه می‌شوند تا مدارک جعلی از مدارک حقیقی تشخیص داده شوند. برای مثال، در جعل تراکنش‌های بانکی، میتوان با بررسی تاریخچه تراکنش‌ها، به عنوان داده ورودی، به یافتن تراکنش‌های غیر مجاز و جعلی پرداخت.

۳.۳.۱ سلامت و پزشکی

بررسی گزارش‌های پزشکی یک حوزه بسیار فعال در علم کامپیوتر و مهندسی پزشکی بوده است. مقایسه و بررسی این گزارش‌ها از دید مسئله تشخیص ناهنجاری نیز بسیار مورد مطالعه قرار گرفته و کاربردهای فراوانی دارد. برای مثال در بررسی تصاویر پزشکی می‌توان از دید مسئله تشخیص ناهنجاری به یافتن بیماری‌ها و نواقص بیمار و علت بیماری پرداخت. همچنین بررسی گزارش علائم بیمار مانند ظربان قلب، سیگنال‌های مغز، فشار خون و غیره توسط دستگاه‌های پزشکی با هدف آگاهی از شرایط بحرانی و کنترل شرایط بیمار بسیار مناسب است. در این نوع کاربردها دادگان به صورت دنباله‌ای از رویدادها به عنوان داده ورودی مورد بررسی قرار می‌گیرند تا در صورت بروز علائم و شرایط حیاطی غیر طبیعی از پیش‌آمدن اتفاقات ناگوار جلوگیری کنند.

۴.۳.۱ سامانه‌های هوشمد و اینترنت اشیا

در سیستم‌های خانه هوشمند، سامانه‌های خودکار و اینترنت اشیا معمولا بسیاری از حسگرها و دستگاه‌ها با استفاده از شبکه‌هایی به هم متصل شده‌اند که برای بررسی وضعیت کلی سیستم و اطمینان از کارکرد صحیح سیستم می‌توان رویدادهای سامانه را در طول زمان مورد بررسی و ارزیابی قرار داد. کاربرد مسئله تشخیص ناهنجاری در اینجا بررسی گزارش‌های سامانه در طی زمان برای پی‌بردن به اتفاق افتادن شرایط نامتعادل و خطاهای سامانه است.

۵.۳.۱ نظارت ویدیویی و سیستم‌های امنیتی

دوربین‌های امنیتی در بسیاری از مکان‌ها برای بالابردن امنیت و همچنین نظارت بر افراد و وضعیت کلی مکان مورد استفاده قرار می‌گیرند اما بررسی و نظارت بر فیلم‌های ظبط شده توسط این دوربین‌ها کار بسیار دشوان و وقتی‌گیری است که در مقیاس وسیع این امر نزدیک به غیر ممکن می‌شود. برای مثال نظارت کارآمد دوربین‌های موجود در سطح شهر تهران برای کنترل ترافیک کار بسیار دشواری است و در صورتی که بخواهیم این کار را با استفاده از منابع انسانی انجام دهیم کار وقت و منابع بسیاری را طلب می‌کند. یکی از کاربردهای مسئله تشخیص ناهنجاری در این حوزه بررسی ویدیوها و تلاش برای یافتن پدیده‌های غیر عادی است. برای مثال تشخیص ناهنجاری در تشخیص عبور غیرمجاز وسایل نقلیه، تشخیص تخلف‌های رانندگی، بررسی امنیت مکان‌های عمومی، وضعیت خط تولید کارخانه برای یافتن کالاهای معیوب و کاربردهای دیگری از این قبیل بسیار مورد استفاده قرار می‌گیرد.

۶.۳.۱ خودروهای خودران

یکی دیگر از حوزه‌های بسیار پرطرفدار در سال‌های اخیر ساخت خودروهای خودران و رانندگی خودکار وسایل نقلیه مختلف است. در این گونه سسیتم‌ها نیز می‌توان با بررسی وضعیت حسگرها و دوربین‌های نصب شده بر روی وسیله نقلیه به بررسی خطرات احتمالی و شرایط غیرعادی مسیر درحال عبور پرداخت. با توجه به اینکه شرایط غیر عادی در رانندگی که منجر به تصادف و خطر شود به ندرت اتفاق می‌افتند و همچنین این شرایط می‌توانند به صورت‌ها و شکل‌ها مختلف روی دهند، استفاده از روش‌های تشخیص ناهنجاری در این کاربردها بسیار مورد پسند پژوهشگران این حوزه قرار گرفته است.

۴.۱ 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.

$$f(.;\theta) : x \to y \tag{۱.۱}$$

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[۸] 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:

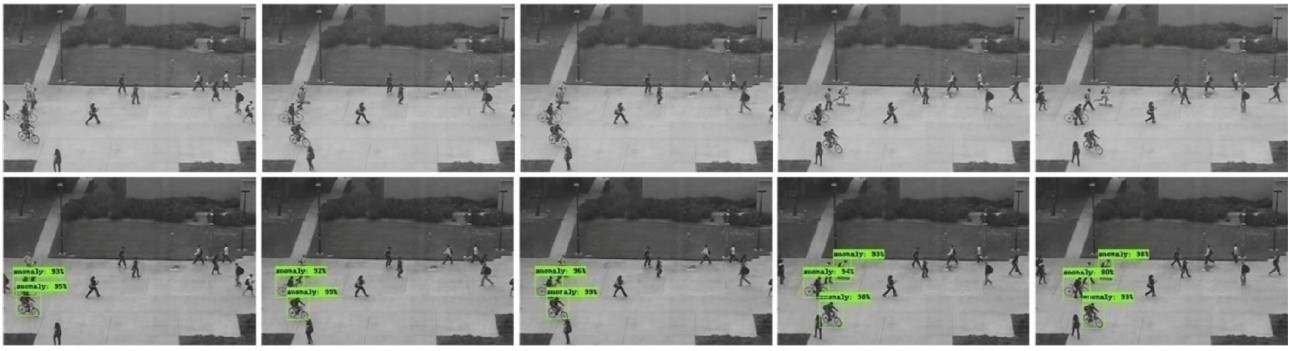
[7] Data representation
[8] Anomaly score

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

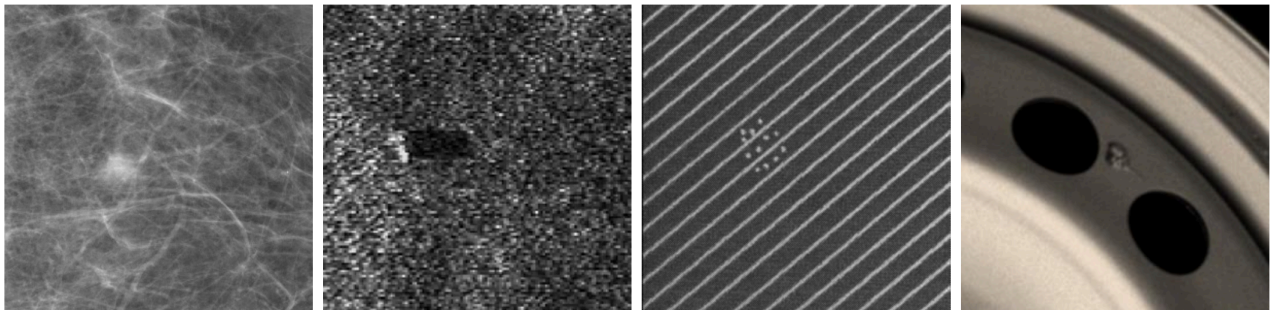
Point Anomaly

Collective Anomaly

شکل ۲.۱: ناهنجاری نقطه‌ای و دنباله‌ای [۳]



شکل ۳.۱: ناهنجاری در کاربرد نظارت ویدیو [۷]



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

شکل ۴.۱: مثال‌هایی از ناهنجاری در تصاویر [۵]

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[۹]: 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.

۵.۱ ساختار گزارش

در فصل اوّل این سمینار به معرفی حوزه سمینار و تعریف مسئله پرداخته شد و در فصل دوّم به بررسی کلی روش‌ها سنتی در مسئله تشخیص ناهنجاری خواهیم پرداخت. فصل سوّم نیز در رابطه با بررسی روش‌های عمیق مورد استفاده در مقالات روز و معرفی کارهای مرتبط با این سمینار و بررسی جزئی از روش‌ها و مقالات موجود چاپ شده در سال‌های اخیر خواهد خواهد بود. در فصل سوّم پس از معرفی هر یک از روش‌ها، چند نمونه مقاله به روز و مرتبط با آن دسته معرفی می‌شوند. در نهایت، در فصل چهارم، مسائل باز و کارهای آینده این حوزه معرفی شده و چند نمونه پیشنهاد برای پروژه نهایی مطرح می‌شود.

[9] Integrated

فصل ۲

مروری بر روش‌های سنتی

اگر به یاد 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			
رویکرد	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

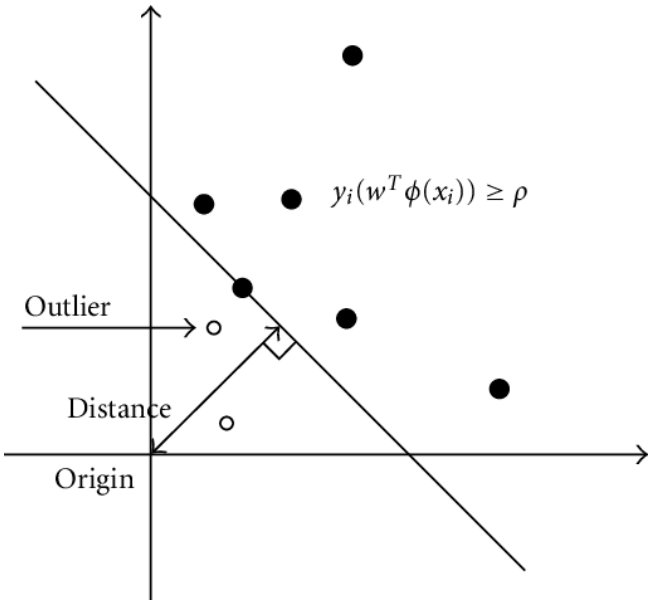
[۱] ر.ک جدول ۱.۲

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 [۱۰].

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.

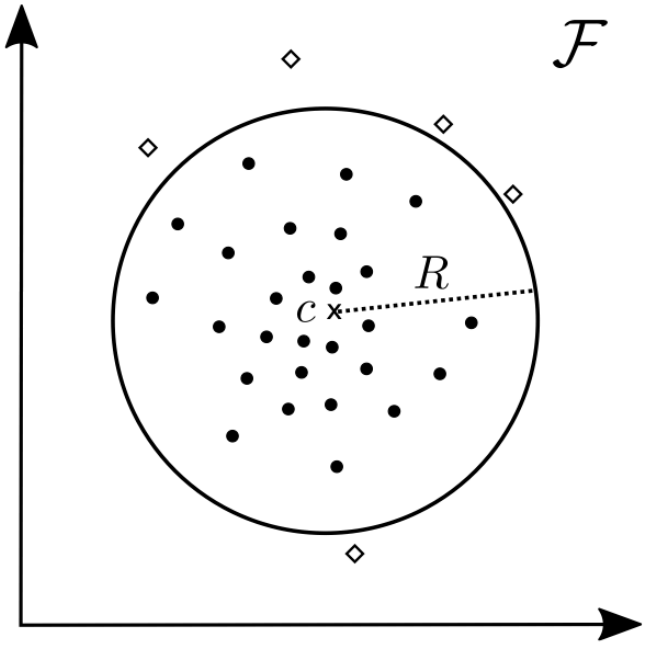


شکل ۱.۲: ماشین بردار پشتیبان یک کلاسه

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 [۸] .



شکل ۲.۲: بردار پشتیبان توصیفگر داده عمیق [۸]

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 رویکردها نیاز به تعیین ابرپارامتر برای مدل یادگیری دارند.

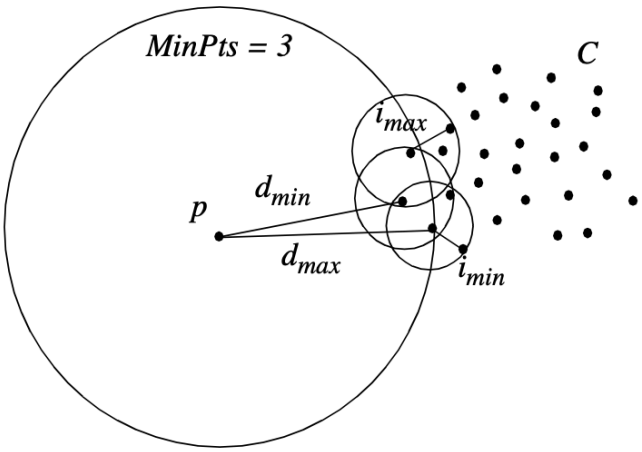
۲.۲ روش‌های مبتنی بر معیار فاصله

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

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



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

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.

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 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		
SAE		
DAE		
RDA		
VAE		
DCAE		
DTS		
GAN		

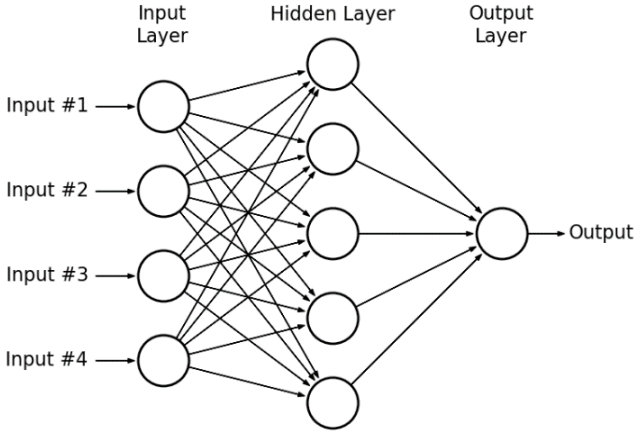
جدول ۱.۳: الگوریتم‌های عمیق مورد استفاده در تشخیص ناهنجاری

۱.۳ استفاده از ساختارهای عمیق

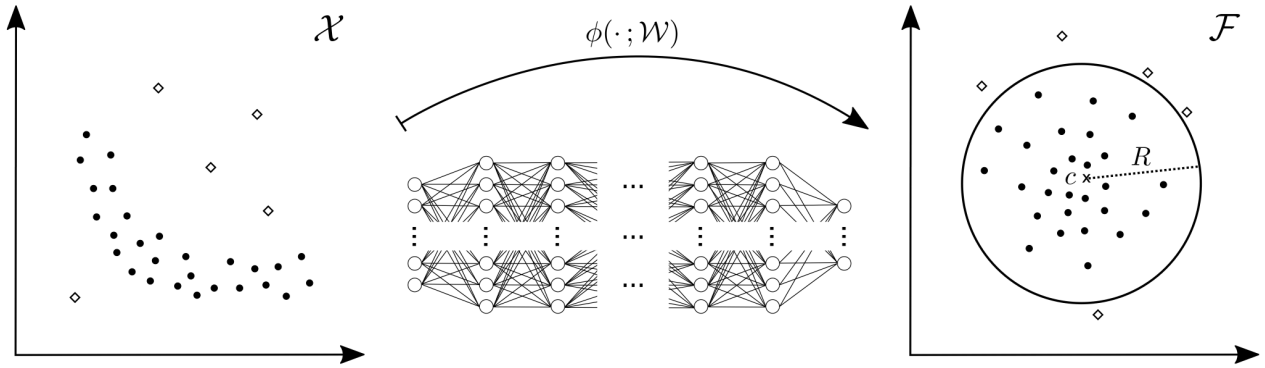
یکی از ابتدایی ترین ایده‌هایی که در مورد استفاده از روش‌های سنتی موجود با توجه به معرفی و پیشرفت ساختارهای عمیق به ذهن می‌رسد، استفاده از این ساختارها در روش‌های سنتی است. ساختارهای عمیق با توجه به قابلیت بالای یادگیری ترکیب‌های غیر خطی گوناگون، می‌توانند به عنوان تابع نگاشت دادگان در روش‌های سنتی استفاده شوند تا بتوانند بازنمایی بسیار بهتری از دادگان را برای انجام عملیات امتیازدهی و تشخیص ناهنجاری بدست آورند. 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 شبکه 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

[1] Multilayer Perceptron
[2] Perceptron

در بسیاری از روش‌های تشخیص ناهنجاری کاربرد دارد. برای مثال در روش بردار پشتیبان توصیفگر داده که در فصل دوم معرفی شد می‌توان بجای تابع $f(.;\theta)$ که مسئول نگاش دادگان به فضایی معین برای بدست آورد بازنمایی خوبی از دادگان است از یک شبکه عمیق مانند مدل پرسپترون چندلایه استفاده کرد. این مدل به دلیل توانایی یادگیری نگاشت غیر خطی دادگان می‌تواند بازنمایی بهتری از دادگان را برای مرحله دوم محاسبات که همان عمل امتیاز دهی به نقاط است بدست آورد. راف و همکاران با استفاده از این ایده، روش بردار پشتیبان توصیف‌گر داده عمیق را معرفی کردن که در مقایسه با روش‌های سنتی عملکرد بسیار بهتری را از خود نشان داده است [۸]. این روش در مقایسه با برخی روش‌های عمیق نیز نتایج بسیار مطلوبی را دربر داشته است که در بخش مقایسه روش‌ها این نتایج آورده خواهند شد.



شکل ۱.۳: مدل پرسپترون چند لایه



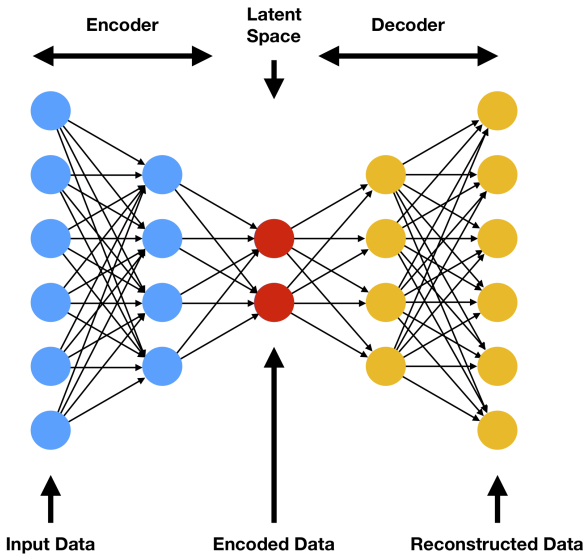
شکل ۲.۳: بردار پشتیبان توصیفگر داده عمیق [۸]

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.

[3] AutoEncoder
[4] Backpropagation

به این فضا به اصطلاح فضای باقی‌مانده[5] یا فضای z می‌گویند. در مرحله دوم، بازنمایی بدست آمده وارد شبکه رمزگشا 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 توانسته بازنمایی خوبی از داده را در فضای باقی مانده ایجاد کند [۱].



شکل ۳.۳: 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 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.

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

در کاربردهای تشخیص ناهنجاری معمولا در هنگام استفاده 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.

$$L(X, g(f(x))) = d \tag{۳.۳}$$

[5] Latent space

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 بخاطر سپاری دادگان ورودی بپردازد و دچار بیش‌برازش[۶] بشود. برای دستیابی به چنین توازنی، انواع مختلفی از خودرمز کننده‌ها معرفی شده‌اند که با افزودن یک مقدار تنظیم کننده[۷] به تابع خطای اصلی معرفی شده، بدست می‌آیند.

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

۳.۳ خود رمز کننده SAE

خود رمز کننده SAE[۸] یکی از انواع خودرمز کننده ها است. ایده اصلی این گونه رمز کننده‌ها این است که، با توجه به اینکه تعداد نورون‌ها لایه مخفی به اندازه کافی زیاد نباشند 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 ترتیبی داده شود تا این نورون‌ها تاحد ممکن کم استفاده شوند و یا به اصطلاح، به صورت خلوت[۹] فعال سازی آنها صورت بگیرید. برای دستیابی به چنین هدفی می‌توان از تنظیم کنند[۱۰] در تابع خطای مدل استفاده کرد. نوع اول استفاده از تنظیم کننده نرم یک است[۱۱] که معادله تابع خطا به صورت زیر خواهد بود.

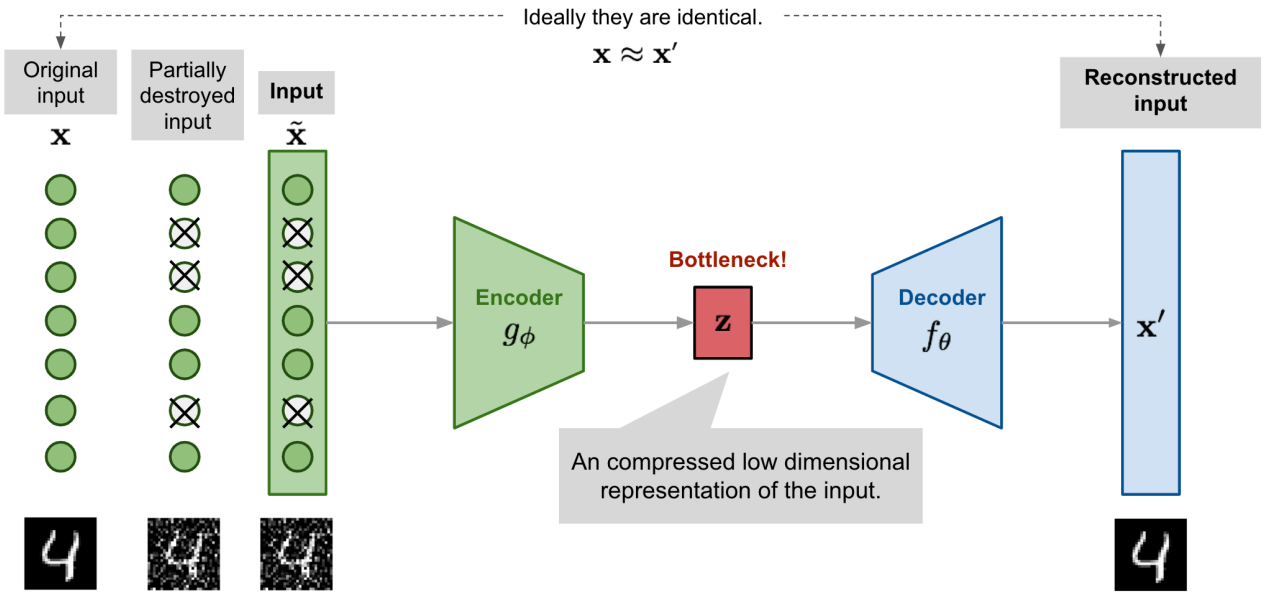
$$L(X, g(f(X))) + \lambda \sum_{i}^{n} |a^{(h)}| \tag{۵.۳}$$

با استفاده از این نوع تنظیم کننده، چون تابع نرم یک استفاده شده، در طی فرایند یادگیری سعی می‌شود وزن یال‌های متصل به نورنوها تاجای امکان صفر باشد و با صفر شدن این وزن‌ها، درواقع نورون‌های کمتری در فرایند محاسبه استفاده می‌شوند.

۴.۳ خود رمز کننده حذف نویز

نوع دیگری از خود رمز کننده که می‌توان از آن برای حذف نویز در داده استفاده کرد را به اصطلاح خودرمز کننده حذف نویز[۱۲] نام دارد. تفاوت این مدل با حالت کلی خود رمز کننده‌ها در فرایند آموزش مدل است. در این مدل داده ورودی ابتدا با استفاده از یک تولید کننده نویز، نویزی می‌شود. سپس به مدل خودرمز کننده داده می‌شود. شبکه نهایی باید بتواند نویز اضافه شده با تصاویر را حذف کند. برای انجام این کار یک راه ساده، تعریف تابع خطا به صورت مقایسه خروجی مدل و ورودی اصلی بدون نویز است. شبکه باید تلاش کند تا اختلاف تصویر باز سازی شده و تصویر اصلی را به حداقل برساند. پس از آموزش این مدل، شبکه قادر خواهد بود تا هرگونه ناهنجاری در داده که در اینجا همان نویز موجود در دادگان است را حذف کند. سپس با مقایسه مقدار خروجی و ورودی مکان‌هایی که تفاوت زیادی بایکدیگر دارند به احتمال تعلق به دسته ناهنجاری در آنها زیاد است.

[6] Overfit
[7] Rgulizer
[8] Sparse AutoEncoder (SAE)
[9] Sparse
[10] Regulizer
[11] L1-Rgulizer
[12] Denoising Auto Encoder



شکل ۴.۳: مدل خود رمز کننده حذف نویز

۵.۳ خودرمزکننده RDA

خودرمزکننده‌هایی که تا کنون معرفی شدند، در مرحله آموزش مدل تنها از دادگان عادی و بدون ناهنجاری استفاده می‌کردند و دادگان ناهنجار تنها زمان آزمون مدل استفاده می‌شدند. حال اگر بخواهیم دادگان ناهنجار را نیز در فرآیند آموزش مدل دخیل کنیم باید روش جدیدی را معرفی کنیم. خودرمز کننده مقاوم[۱۳] درواقع از ایده تجزیه و تحلیل مؤلفه بنیادی مقاوم[۱۴] برگرفته شده است. در روش تجزه و تحلیل مؤلفه بنیادی مقاوم، دادگان ورودی با استفاده از دو ماتریس مرتبه پایین[۱۵] و خلوت[۱۶] نمایش داده می‌شوند.

$$X = L + S \tag{۶.۳}$$

که در اینجا L نمایش داده ورودی در ابعاد پایین‌تر است و S قسمتی از دادگان است که نمی‌تواند توسط L به خوبی نمایش داده شود. این دوماتریس تحت شرط بهینه‌سازی و تابع هدف زیر آموزش داده می‌شوند [۱۷] .

$$||X - L - S||_F^2 = 0 \tag{۷.۳}$$

$$min_{L,S}||L||_* + \lambda||S||_1 \tag{۸.۳}$$

این روش نیز سعی دارد دادگان ورودی را به استفاده از دو ماتریس نمایش دهد که ماتریس اول بازنمایی بدست آمده توسط خودرمزکننده است و قسمت دوم نمایانگر ناهنجاری‌هایی است که نمی‌توانند توسط خودرمزکنند به خوبی بازنمایی

[13] Robust Deep AutoEncoder
[14] Robust PCA
[15] Low rank
[16] Sparse

[۱۷] در اینجا $||.||_F$ نرم frobenius و $||.||_*$ جمع مقادیر یکتا(singular value) است.

شوند.

$$X = L_D + S \tag{۹.۳}$$

اگر رمزکنند و رمزگشا را به عنوان دو تابع f و g در نظر بگیریم، معادل بهینه سازی مدل به صورت زیر خواهد بود.

$$min_\theta ||L_D - G_\theta(F_\theta(L_D))||_۲ + \lambda ||S||_۱ \tag{۱۰.۳}$$

که شرایط زیر باید در فرایند بهینه سازی صدق کند:

$$X - L_D - S = ۰ \tag{۱۱.۳}$$

فرایند امتیاز دهی به ناهنجاری در این نوع خودرمز کننده مشابه روش اصلی خواهد بود. در این جا S در واقع همان ناهنجاری‌های موجود در دادگان هستند که پس از تکمیل فرایند آموزش می‌توانیم از آن استفاده کنیم. این روش در مقایسه با روش سنتی در کاربرد تشخیص ناهنجاری حدود ۷۰ درصد بهتر عمل کرده است [۱۱].

۶.۳ خود رمز کننده VAE

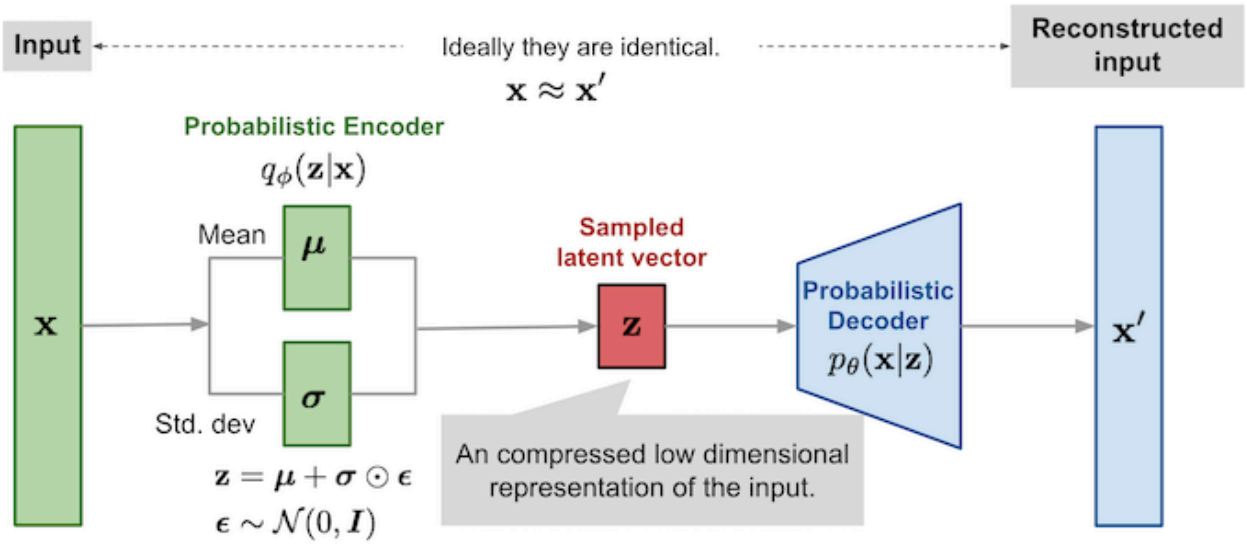
مشکل که خودرمز کننده‌هایی که تا کنون معرفی کردیم در این است که نگاشت دادگان به فضای باقی‌مانده به صورت قطعی صورت می‌گیرد. یعنی هر نقطه از فضای ورودی به یک مقطه معین از فضای باقی‌مانده نگاشته می‌شود. از طرف دیگر اگر یک نقطه را به صورت تصادفی در فضای باقی‌مانده، مانند z' را در نظر بگیریم، نمی‌توان به طور قطع گفت که این نقطه به کدام دسته از نقاط تعلق خواهد گرفت. در واقع خودرمز کننده‌هایی که تا کنون مطالعه کردیم به خوبی دادگان ورودی را به فضایی با ابعاد دیگر نگاشت میکردند اما در هیچ یک از این روش‌ها ما اختیاری برا کنترل روند و نحوه این نگاشت نداشتیم. انواع مختلف این رمز کننده‌ها نیز بسته به نیاز، نگاشت‌های گوناگونی را در اختیار ما قرار می‌دادند تا مناسب کاربرد انتخاب شده باشند. برای اینکه در طی فرایند یادگیری ما بر روی نحوه نگاش دادگان به فضای باقی‌مانده کنترل داشته باشیم، نوع دیگری از خودرمزکننده‌ها تحت عنوان خودرمز کننده VAE[۱۸] معرفی شده است [؟]. در این روش بجای یادگیری نگاشت گسسته و قطعی دادگان به فضای باقیمانده، سعی می‌شود توزیع دادگان در فضای باقیمانده یاد گرفته شود که در این صورت دادگان در این فضا توضیع پیوسته‌ای خواهند داشت و عمل درون‌یابی نقاط در این فضا کار راحت‌تری خواهد بود.

در این روش بجای تلاش برای کمینه کردن اختلاف ورودی مدل با خروجی بازسازی شده توسط مدل، سعی می‌شود تا احتمال درستنمایی نهایی[۱۹] حداکثر شود. معادل تابع بهینه سازی این مدل به صورت زیر تعریف می‌شود.

$$log(p(x)) \geq log(p(x)) - KL(q_\phi(z|x)||p(z)) \tag{۱۲.۳}$$

$$log(p(x)) - KL(q_\phi(z|x)||p(z)) = E_{z\sim q_\phi(x)} log P_\phi(x|z) - KL(q_\phi(z|x)||p(z)) \tag{۱۳.۳}$$

[18] Variational AutoEncoder
[19] Marginal likelihood



شکل ۵.۳: مدل خود رمز کننده variational

$$maximize\ E_{z\sim q_\phi(x)} logP_\phi(x|z) - KL(q_\phi(z|x)||p(z)) \tag{۱۴.۳}$$

در معادله (۱۴.۳) قسمت اول برای حداکثر کردن احتمال داده باز سازی شده است. قسمت دوم که درواقع می‌توان آن را به عنوان تنظیم کننده معادله درنظر گرفت، تلاش می‌کند تا توزیع دادگان در فضای بازنمایی z بسیار مشابه توزیع دادگان ورودی باشند. بنابراین بازنمایی دادگان در فضای باقیمانده بر خلاف مدل پایه به صورت غیر قطعی[۲۰] خواهد بود. همچنین دادگان بازسازی شده و همچنین امتیاز ناهنجاری برای دادگان نیز غیر قطعی و به صورت احتمال خواهند بود.

۷.۳ شبکه‌های مولد رقابتی(GAN)

شبکه‌های مولد رقابتی[۲۱]از دو قسمت اصلی تشکیل شده‌اند که به صورت رقابتی با یکدیگر آموزش می‌بینند. هر یک از این دو قسمت، سعی دارند عملکرد طرف مقابل را با بالا بردن کیفیت کار خود به چالش بکشند. بخش اول این مدل که مول[۲۲] نام دارد، مسئولیت تولید داده مصنوعی را بر عهده دارد. این قسمت با گرفتن یک بردار ورودی از فضای باقیمانده، داده‌ای مصنوعی را تولید می‌کند. خروجی این قسمت به همرای یک نمونه از دادگان آموزش برای مقایسه و داوری جهت تشخیص مصنوعی و یا حقیقی بودن به قسمت دوم مدل که تصمیم گیرنده[۲۳] نام دارد وارد می‌شوند. بخش دوم باید بتواند به داده حقیق که از دادگان آموزش دریافت کرده است برچسب حقیق و به داده تولید شده توسط بخش مولد برچسب مصنوعی بودن را اختصاص دهد. آموزش این مدل ها به صورت نوبتی صورت می‌گیرد و با شروع از بخش دوم، وزن ها در بخش دیگر ثابت می‌مانند و پس از چند مرحله که عملکرد این قسمت بهبود یافت، تغییر وزن‌ها در آن بخش متوقف شده و وزن‌های بخش دیگر آموزش می‌بینند و پس از بهبود عملکرد قسمت بعدی این چرخه ادامه پیدا می‌کند. تابع خطای مورد استفاده در مدل‌های مولد پایه به صورت زیر است.

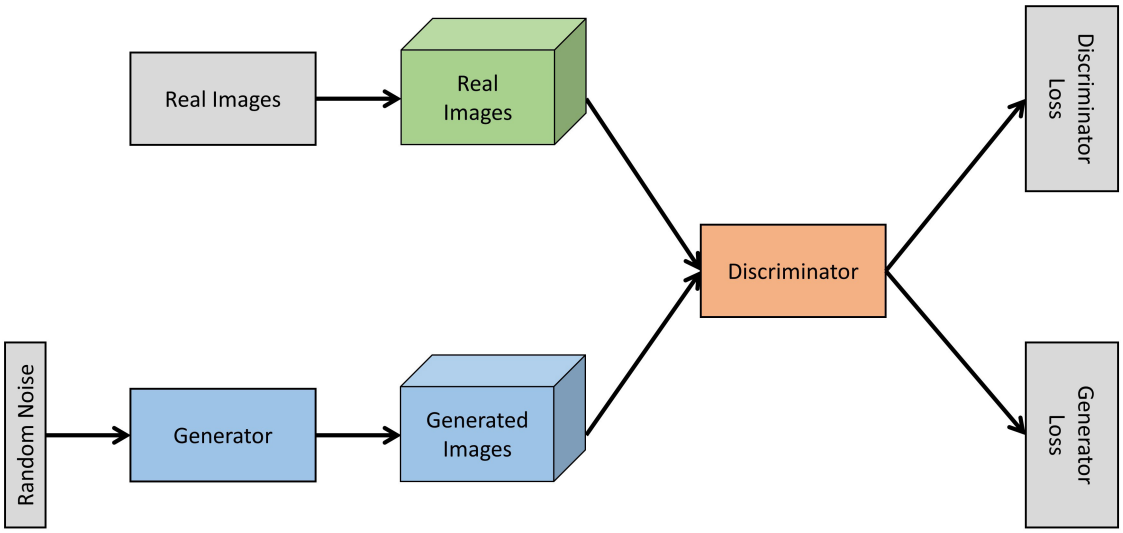
[20] Stochastic
[21] Generative Adverserial Networks
[22] Generator
[23] Discriminator

$$min_G max_D V(D,G) = E_{X\sim p_{data}(x)}[logD(x)] + E_{z\sim p_z(z)}[log(۱ - D(G(z)))] \quad (۱۵.۳)$$

برای استفاده از این مدل در تشخیص ناهنجاری، استفاده از تابع خطای تصمیم گیرنده به عنوان تابع امتیاز ناهنجاری می‌تواند مفید باشد. در اینصورت، تابع تصمیم گیرنده $D(X)$ وظیفه نگاشت دادگان به فضای تشخیص ناهنجاری را بر عهده دارد و تابع خطای این قسمت از مدل که به صورت زیر تعریف می‌شود به عنوان تابع امتیاز ناهنجاری بکار خواهد رفت.

$$d(x) = log(۱ - D(X)) \quad (۱۶.۳)$$

میزان آستانه تصمیم گیری δ نیز می‌تواند با استفاده از آزمون و خطا و یا با استفاده از منحنی حساسیت و دقت تعیین گردد.



شکل ۶.۳: شبکه مولد رقابتی

برای اینکه بتوانیم از شبکه مولد نیز در این مسئله کمک بگیریم، می‌توانیم در فرایند آموزش دادگان، بجای انتخاب تصادفی یک نقطه از فضای z به عنوان ورودی شبکه مولد، با استفاده از یک رمز کننده دادگان ورودی را ابتدا به رمز کننده بدهیم تا بازنمایی دادگان در فضای باقیمانده بدست آید و سپس این داده را به عنوان ورودی به شبکه مولد بدهیم تا با این بازنمایی اقدام به تولید داده مصنوعی کند. چیزی که در اینجا توقع داریم این است که داده تولید شده توسط تابع مولد، بسیار شبیه به داده ورودی رمز کننده باشد. در این صورت تابع بهینه سازی مدل به صورت زیر خواهد بود.

$$min_\theta ||G(E(X,\theta)) - X|| + \lambda log(۱ - D(G(E(X,\theta)))) \quad (۱۷.۳)$$

در این معادله پارامتر λ یک ابر پارامتر مدل است که به صورت دستی تعیین می‌شود. این روش در سال ۲۰۱۸ توسط چلگ و همکاران تحت عنوان AnoGan معرفی شد [۹].

۸.۳ مدل‌های جریانی

مدلهای جریانی[۲۴] مشابه خود رمز کننده VAE سعی می‌کنند تا توزیع دادگان ورودی را بیاموزند. در این رویکرد، سعی می‌شود تا با استفاده از یک دنباله از توابع مبدل معکوس پذیر که دادگان را از فضای باقیمانده به فضای ورودی نگاشت می‌کنند، توزیع دادگان آموخته شود. یکی از ایراداتی که بر خودرمز کننده VAE در یادگیری توزیع دادگان وارد است، در نظر گرفتن توزیع مانند گوسی برای دادگان است.

۹.۳ بررسی کارهای انجام شده

حال که به انواع مختلف روش‌های پایه‌ای آشناشدیم، می‌توانیم به استفاده از این روش‌ها در کاربردهای مختلف بپردازیم. اگر دادگان ما به صورت تصویر، صوت و یا ویدیو باشند، با تغییر لایه‌های کاملا متصل به شبکه‌های کانولوشنی می‌توانیم از این معماری برای این کاربردها استفاده کنیم. همچنین برای دادگان دنباله‌ای مانند متن، سری‌های زمان و غیره نیز میتوانیم با جایگذاری شبکه‌های بازگشتی به جای شبکه کاملا متصل از این روش‌ها استفاده کنیم.

[24] Flow based models

فصل ۴

کارهای آینده

همانطور که در فصل یک با اهمیت کاربرد روش‌های تشخیص ناهنجاری در حوزه‌های مختلف آشنا شدیم و در فصل دوم و سوم علاوه بر مطالعه روش‌های موجود و مناسب برای مسئله تشخیص ناهنجاری، به بررسی کارهای انجام شده در این حوزه پرداختیم حال وقت آن رسیده که مسائل و چالش‌های باز این حوزه را مرور کنیم.

۱.۴ مسائل باز و کارهای قابل انجام

در مسائل تشخیص ناهنجاری یکی از چالش‌های مهم، در دسترس نبودن دادگان برچسب خورده کافی و مناسب یادگیری است. با توجه به اینکه ناهنجاری‌ها ذاتا به ندرت اتفاق می‌افتند و همچنین مفهوم ناهنجاری در کاربردهای گوناگون متفاوت است، استفاده از روش‌های با ناظر در این مورد مسئله‌ای چالش برانگیز است. یکی از روش‌های پرکاربرد در تشخیص ناهنجاری، یادگیری بازنمایی دادگان عادی است که در فصل سوم به برخی از آنها پرداختیم. در استفاده از این روش‌ها، این احتمال وجود دارد که دادگان آموزش ممکن است دارای خطا باشند که این مورد ممکن است در فرایند یادگیری تاثیر بگذارد. همچنین در بسیاری از کاربردها مرز مشخص و دقیقی میان دادگان عادی و ناهنجار وجود ندارد. از این رو یافتن روش‌هایی که در مقابل خطا مقاوم باشند یکی از مسائل و چالش‌های بازی در این حوزه است. برای مقابله با چالش کمبود دادگان ناهنجار، ابداع روش‌های داده‌افزایی و بازنمایی دادگان که نیاز به تعداد کمی داده داشته باشند نیز یکی از مسائل باز است.

۲.۴ موضوع پیشنهادی برای پایان نامه

با توجه به کاربرد وسیع مسئله تشخیص ناهنجاری در حوزه پزشکی و مشکل کمبود دادگان برچسب خورده به دلیل چالش‌هایی که در این حوزه وجود دارد، استفاده از روش‌های تشخیص ناهنجاری در این حوزه بسیار مناسب است. در پردازش تصاویر پزشکی به دلیل منحصر به فرد بودن بافت بدن اشخاص مختلف و همچنین نادر بودن بیماری‌ها در میان افراد می‌توان از دید مسئله تشخیص ناهنجاری به وجود توده‌های سرطانی در تصاویر پزشکی نگاه کرد و به بررسی مسئله از این دید پرداخت که می‌تواند به عنوان پیشنهادی برای پروژه پایانی بررسی شود.

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] Murugan, B.S., Elhoseny, Mohamed, Shankar, K., and Uthayakumar, J. Region-based scalable smart system for anomaly detection in pedestrian walkways. *Comput. Electr. Eng.*, 75(C):146–160, may 2019.

[8] 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.

[9] Schlegl, Thomas, Seeböck, Philipp, Waldstein, Sebastian M., Schmidt-Erfurth, Ursula, and Langs, Georg. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In Niethammer, Marc, Styner, Martin, Aylward, Stephen, Zhu, Hongtu, Oguz, Ipek, Yap, Pew-Thian, and Shen, Dinggang, editors, *Information Processing in Medical Imaging*, pages 146–157, Cham, 2017. Springer International Publishing.

[10] 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.

[11] Zhou, Chong and Paffenroth, Randy C. Anomaly detection with robust deep autoencoders. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '17, page 665–674, New York, NY, USA, 2017. Association for Computing Machinery.

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