



دانشکده فیزیک

پایان نامه دوره کارشناسی ارشد

در رشته فیزیک

یافتن اندازه گیری های نیمه بهینه برای تشخیص درهم تنیدگی با استفاده از

شبکه های عصبی خود رمز گذار

نام دانشجو:

محمد یوسف پور

استاد راهنما:

دکتر صادق رئیسی

بهار ۱۴۰۰

به پاس تعبیر عظیم و انسانی شان از کلمه اثار و از خودگذشتگی
به پاس محبت های بی دریغشان که هرگز فروکش نمی کند

این مجموعه را به پدر و مادر عزیزم تقدیم می کنم

شکرتقدردانی

بسیار شایسته است از استاد محترم جناب آقای دکتر صادق ریسی، به خاطر زحمات بی دریغ، دلسوزانه و مؤثر در روند انجام این پایان نامه کمال تشکر و قدردانی را داشته باشم.

محمد یوسف پور

بهار ۱۴۰۰

چکیده

درهم تنیدگی یکی از منابع کلیدی علوم اطلاعات کوانتومی است که شناسایی وضعیت های درهم تنیده را برای طیف گسترده ای از فناوری ها و پدیده های کوانتومی ضروری می کند. این مشکل به عنوان مساله جدپذیری نیز شناخته می شود و NP-Hard است. در حالت ایده آل ، می خواهیم شاهدان بهینه بیابیم که با کمترین تعداد اندازه گیری ممکن، بتوانند حالت های درهم تنیده را تشخیص دهند.

در این پایان نامه، ما پیشنهاد می کنیم از شبکه های عصبی خود رمزگذار برای یافتن اندازه گیری های نیمه بهینه برای تشخیص درهم تنیدگی و ساخت شاهدان جدید درهم تنیدگی استفاده کنیم که می توانند حالت های درهم تنیده را با داده های ناکامل تشخیص دهند، به عنوان مثال فقط با سه اندازه گیری. شاهد شبکه عصبی، حالت ها را به عنوان ورودی می گیرد و در خروجی مشخص می کند که آیا حالت در خروجی درهم تنیده است یا خیر. در اینجا ما روش خود را روی سیستم دو کیوبیتی اعمال می کنیم، اما این روش می تواند برای سیستم هایی با ابعاد فضای هیلبرت بالاتر تعمیم یابد.

واژه های کلیدی: درهم تنیدگی، اطلاعات کوانتومی، مساله جدپذیری، یادگیری ماشین، هوش مصنوعی

فهرست مطالب

۱	فهرست مطالب
۵	فهرست تصاویر
۷	فهرست جداول
۹	۱ پیش گفتار
۹	۱.۱ مقدمه
۱۰	۲.۱ حالت و سیستم کوانتمی
۱۰	۱.۲.۱ حالت کوانتمی
۱۰	۲.۲.۱ حالت کوانتمی سیستم مرکب
۱۰	۳.۱ درهم تنیدگی
۱۱	۴.۱ مساله جداپذیری
۱۱	۵.۱ شاهد درهم تنیدگی
۱۲	۶.۱ معیار Peres – Horodecki (PPT)
۱۳	۷.۱ نمایش حالت های دوکیوبیت
۱۴	۸.۱ بررسی سایر رویکرد های موجود
۱۵	۹.۱ ساختار پایان نامه و علامت گذاری
۱۷	۲ یادگیری ماشینی
۱۷	۱.۲ مقدمه
۱۷	۲.۲ یادگیری ماشین چیست؟
۱۹	۳.۲ انواع یادگیری ماشین

۱۹	یادگیری نظارت شده	۱.۳.۲
۱۹	یادگیری نظارت نشده	۲.۳.۲
۱۹	یادگیری شبه نظارت شده	۳.۳.۲
۲۰	شبکه های عصبی	۴.۲
۲۰	شبکه های خود رمزگذار	۵.۲
۲۱	تشخیص ناهنجاری	۶.۲
۲۲	متریک های ارزیابی	۷.۲
۲۴	ماتریس درهم ریختگی	۱.۷.۲
۲۴	نتیجه گیری	۸.۲
۲۷	تشخیص حالت های درهم تنیده از جداپذیر	۳
۲۷	مقدمه	۱.۳
۲۷	تولید داده ها	۲.۳
۲۸	بررسی آماری داده ها	۳.۳
۳۰	روش های کلاسیک	۴.۳
۳۲	اهمیت ویژگی ها	۱.۴.۳
۳۳	شبکه های عصبی	۵.۳
۳۴	شبکه عصبی خود رمزگذار	۶.۳
۳۴	شبکه عصبی با تعداد ویژگی کمتر	۷.۳
۳۷	تقارن استوانه ای	۸.۳
۴۰	پایداری و اثر نوفه	۹.۳
۴۱	یادگیری نظارت نشده	۱۰.۳
۴۲	یادگیری شبه نظارت شده	۱۱.۳
۴۳	نتیجه گیری	۱۲.۳
۴۵	جمع بندی و پیشنهادات	۴
۴۵	جمع بندی	۱.۴
۴۶	پیشنهادات	۲.۴
۴۷	بررسی تقارن های دیگر	۱.۲.۴

۴۷	ابعاد بالاتر	۲.۲.۴
۴۷	ماشین بولتزمن و مدل های تولید کننده	۳.۲.۴
۴۸	روش های دیگر یادگیری ماشینی	۴.۲.۴

فهرست تصاویر

- ۱.۱ شاهد درهم تنیدگی فضای حالت ها را به دو زیر مجموعه افراز میکند و می تواند بخشی یا همه ی حالت های در هم تنیده را تشخیص دهد. شاهد درهم تنیدگی می تواند خطی یا غیر خطی باشد (شکل آ). معیار پرز-هورودکی یک شاهد لازم و کافی برای سیستم دو کیبیت است. این شاهد فضای حالت ها را به طور کامل به دو بخش فقط درهم تنیده و فقط جداپذیر افراز می کند (شکل ب). ۱۲
- ۱.۲ تفاوت برنامه نویسی کلاسیک با یادگیری ماشین ۱۸
- ۲.۲ ارتباط بین هوش مصنوعی، یادگیری ماشین، شبکه های عصبی، یادگیری عمیق و علم داده. [۱] ۱۸
- ۳.۲ معماری انواع شبکه های عصبی ۲۰
- ۴.۲ نحوه تشخیص ناهنجاری با استفاده از برازش توزیع گاوسی به داده ها ۲۲
- ۵.۲ مصالحه دقت و بازیابی. با تغییر آستانه که معادل جابجایی مرز تشخیص است، می توان به دقت بالاتر یا بازیابی بالاتر دست یافت. ۲۴
- ۱.۳ هیستوگرام ویژگی های مختلف. همه ی ۱۵ ویژگی دارای توزیع یکسان شبه گاوسی هستند. ۲۹
- ۲.۳ ماتریس پخشی ویژگی های مختلف نسبت به هم. ۲۹
- ۳.۳ ماتریس همبستگی. این ماتریس همبستگی ویژگی های مختلف نسبت به همدیگر را نشان می دهد. ۳۰
- ۴.۳ اهمیت ویژگی ها برای مدل درخت تصادفی. هر کدام از ویژگی ها برای تعیین نتیجه اهمیت متفاوتی در مدل دارند. ۳۲
- ۵.۳ معماری شبکه عصبی پیشخور ارایه شده برای این مساله. تعداد گره ها و لایه ها در شکل نمایش داده شده است. ۳۳

- ۶.۳ معماری شبکه ی خود رمزگذار ارایه شده برای پیدا کردن دو اندازه گیری غیر خطی. . . . ۳۵
- ۷.۳ معماری شبکه ی عصبی برای تشخیص تعداد کمتر اندازه گیری خطی ۳۶
- ۸.۳ نتایج ارزیابی مدل. صحت مدل شبکه ی عصبی با تعداد اندازه گیری خطی کمتر، برای حالت بدون تقارن و حالت با تقارن استوانه ای در شکل بر حسب تعداد اندازه گیری نشان داده شده است. ۳۸
- ۹.۳ ضرایب لایه ی اول برای هر سه گره. این ضرایب در واقع همان ضرایب خطی ماتریس های پاولی اندازه گیری بهینه پیدا شده هستند. همچنین، بزرگی اندازه این ضرایب، نشانگر اهمیت هر یک از ویژگی ها در تعیین نتیجه است. ۳۹
- ۱۰.۳ نمودار صحت شبکه عصبی پیش خور نسبت به انحراف معیار نوفه گاوسی جمع شونده . . ۴۱

فهرست جداول

۱.۲	ماتریس درهم ریختگی	۲۵
۱.۳	نتایج مربوط به روش های کلاسیکی یادگیری ماشینی از جمله تعداد نمونه ی یادگیری، زمان آموزش، زمان آزمایش، و صحت ارزیابی	۳۱
۲.۳	ماتریس درهم ریختگی شبکه ی پیشخور	۳۵

فصل ۱

پیش گفتار

۱.۱ مقدمه

در سال های اخیر، شاهد کاربرد های هوش مصنوعی در زمینه های مختلف علوم بوده ایم که منجر به حل مسائل بسیار پر کاربردی در علم و فناوری شده است. از این دست می توان به کاربرد هوش مصنوعی در علوم پایه مانند فیزیک، بیولوژی و شیمی اشاره کرد [۲]. با پیشرفت شگرف هوش مصنوعی و به وجود آمدن روش هایی مانند یادگیری ماشینی، قدرت حل مسائل پیچیده تر با آن نیز افزایش یافته است [۳].

درهم تنیدگی یکی از منابع کلیدی علوم اطلاعات کوانتومی است که کاربرد های بسیاری در رمزنگاری کوانتومی، تلپورت کوانتومی، و میکروسکوپی کوانتومی دارد.

در نتیجه شناسایی حالت های درهم تنیده برای طیف گسترده ای از فناوری ها و پدیده های کوانتومی ضروری است. این مساله که به عنوان مشکل جداپذیری نیز شناخته می شود [۴]، از مسائل معروف در زمینه ی اطلاعات کوانتومی است که حل آن به روش های تحلیلی با استفاده از کامپیوتر می تواند بسیار زمان بر و عملا غیر ممکن باشد [۵]. در این پروژه با استفاده از مفاهیم و ابزار های هوش مصنوعی به بررسی این مسئله پرداخته شده است تا بتوان با استفاده از ابزار های یادگیری ماشینی، حالت های درهم تنیده را تشخیص دهیم. از این رو، ضروری است تا در ابتدا مفاهیم مورد نیاز در هر دو زمینه، به اختصار توضیح داده شوند.

۲.۱ حالت و سیستم کوانتومی

۱.۲.۱ حالت کوانتومی

در تئوری اطلاعات کوانتومی، حالت یک سیستم فیزیکی، به صورت یک ماتریس چگالی ρ عضو فضای هیلبرت سیستم نشان داده می شود. ماتریس چگالی، خواص زیر را دارد [۶]:

$$\begin{aligned}\rho^\dagger &= \rho \\ \text{Tr}(\rho) &= 1 \\ \rho &\geq 0\end{aligned}\tag{۱.۱}$$

۲.۲.۱ حالت کوانتومی سیستم مرکب

حالت کوانتومی یک سیستم مرکب به صورت یک ماتریس چگالی عضو ضرب تنسوری فضاهای هیلبرت زیرسیستم های آن است.

$$\mathcal{H} = \bigotimes_{i=1}^n \mathcal{H}_i\tag{۲.۱}$$

۳.۱ درهم تنیدگی

در سیستم های مرکب، پدیده ای فیزیکی به نام درهم تنیدگی^۱ ممکن است رخ بدهد به طوری که حالت فیزیک هر یک از این زیر سیستم ها را نمیتوان به صورت مستقل از زیر سیستم های دیگر توصیف کرد، و اندازه گیری های ویژگی های فیزیکی زیر سیستم ها با یکدیگر همبستگی تقریبی دارد [۷]. درهم تنیدگی منبع مهمی در علوم اطلاعات کوانتومی محسوب می شود و همه ی حالت های مرکب این خاصیت را ندارند [۸]. حالت کوانتومی که درهم تنیده نیست را جداپذیر می گویند. درهم تنیدگی را می توان به صورت کمی درآورد و به روش های مختلف میزان درهم تنیدگی یک حالت را به دست آورد [۹].

برای یک سیستم دو بخشی^۲، فضای هیلبرت متناظر به صورت $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ است. در این حالت، یک ماتریس چگالی جداپذیر خوانده می شود اگر و تنها اگر بتوان آن را به صورت ترکیب محدبی^۳ از حالت

^۱Entanglement

^۲bi-partite

^۳convex combination

های ضربی نوشت. به زبان ریاضی، حالت های جداپذیر به شکل زیر تعریف می شوند [۱۰]:

$$\rho_{AB} = \sum_i p_i \rho_A^{(i)} \otimes \rho_B^{(i)}, \sum_i p_i = 1, p_i \geq 0 \quad (3.1)$$

حالت هایی که جداپذیر نباشند، حالت های درهم تنیده گفته می شوند.

۴.۱ مساله جداپذیری

در بخش قبل، تعریف حالت جداپذیر و درهم تنیده را گفتیم. حال سوال اینجاست که چگونه تشخیص دهیم یک حالت مشخص، درهم تنیده یا جداپذیر است. به عنوان مثال، برای یک سیستم دو بخشی با حالت خالص، می توان با تجزیه اشمیت^۱ به راحتی فهمید که آیا حالت درهم تنیده یا جداپذیر است. اگر در این تجزیه، فقط یکی از ضرایب غیر صفر بود، حالت جداپذیر است، و در غیر این صورت حالت قطعا درهم تنیده است [۱۱]، [۱۲]. با این حال، برای حالت های مخلوط، نمایش تجزیه ی معادله ی ۳.۱ ممکن است یکتا نباشد [۱۱]، [۱۲]. بنابراین، اگر تجزیه ای به این صورت پیدا شود که ضرب حالت های زیر سیستم ها باشد ولی شرط $\sum_i p_i = 1$ یا $p_i \geq 0$ برقرار نشود، نمی توان نتیجه گرفت که حالت جداپذیر نیست. چون ممکن است نمایش دیگری از این تجزیه وجود داشته باشد که شرط های معادله ی ۳.۱ را برقرار سازد. به همین خاطر، در حالت کلی، مشخص کردن اینکه یک ماتریس چگالی درهم تنیده است یا جداپذیر از نظر محاسباتی مساله سختی است [۹]، اگرچه در بعضی حالت های خاص، روش های ساده تحلیلی وجود دارد [۴]. به این مساله در علوم اطلاعات کوانتومی، مساله ی جداپذیری^۲ گفته میشود، و در حالت کلی، نشان داده شده است که از نظر محاسباتی یک مساله NP-Hard است [۵].

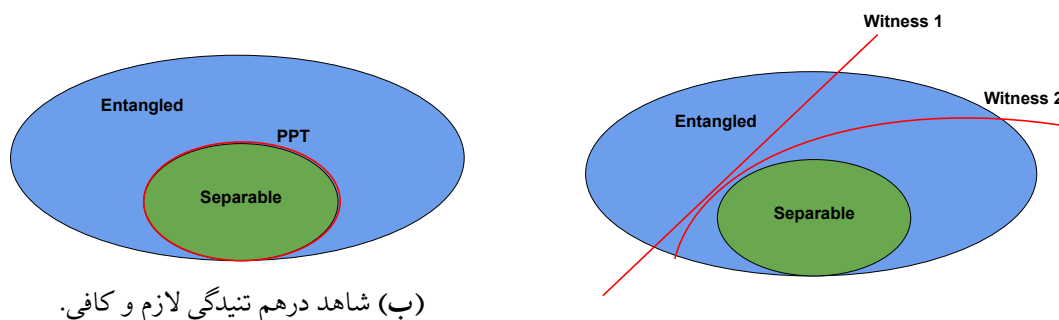
۵.۱ شاهد درهم تنیدگی

شاهد درهم تنیدگی^۳، معیاری برای تشخیص حالت های درهم تنیده است. این شاهد ها در واقع یک شرط کافی برای درهم تنیدگی هستند، اما باید توجه کرد یک شاهد درهم تنیدگی، لزوما شرط لازمی برای درهم تنیده بودن نیست. بنابراین اگر شروط شاهد درهم تنیدگی برآورده نشود، نمی توان نتیجه گرفت که حالت جداپذیر است. [۷]، ولی اگر شروط شاهد برآورده شود، می توان نتیجه گرفت که حالت حتما درهم تنیده است.

¹Schmidt decomposition

²Separability Problem

³Entanglement witness



(آ) شاهد درهم تنیدگی خطی و غیر خطی.

شکل ۱.۱: شاهد درهم تنیدگی فضای حالت ها را به دو زیر مجموعه افراز میکند و می تواند بخشی یا همه ی حالت های درهم تنیده را تشخیص دهد. شاهد درهم تنیدگی می تواند خطی یا غیر خطی باشد (شکل آ). معیار پرز-هورودکی یک شاهد لازم و کافی برای سیستم دو کیوبیت است. این شاهد فضای حالت ها را به طور کامل به دو بخش فقط درهم تنیده و فقط جداپذیر افراز می کند (شکل ب).

به زبان ریاضی، برای هر حالت درهم تنیده ρ ، یک عملگر هرمیتی w وجود دارد به طوری که $\text{Tr}(w\rho) < 0$ و $\text{Tr}(w\sigma) \geq 0$ برای هر حالت جداپذیر σ [۴]. به این عملگر w یک شاهد درهم تنیدگی گفته میشود، چرا که هر گاه $\text{Tr}(w\rho) < 0$ شود، قطعاً این حالت درهم تنیده است.

فرص کنید که فضای حالت های کوانتومی مطابق شکل ۱.۱ آ باشد. آنگاه یک شاهد درهم تنیدگی فضای حالت ها را به دو زیر فضا افراز میکند، که یک بخش فقط شامل حالت های درهم تنیده، و بخش دیگر شامل حالت های جداپذیر و درهم تنیده است. پس با هر شاهد درهم تنیدگی، ممکن است فقط بخشی از حالت های درهم تنیده تشخیص داده شوند. در حالتی که شاهد درهم تنیدگی، یک شرط لازم و کافی باشد، شاهد فضا را به دقیقاً دو زیر فضای درهم تنیده و جداپذیر افراز میکند، یا به عبارتی می تواند همه ی حالت های درهم تنیده را تشخیص دهد (شکل ۱.۱ ب).

۶.۱ معیار Peres – Horodecki (PPT)

در حالتی که ابعاد زیر فضاهای هیلبرت سیستم مرکب، رابطه ی $d_A \times d_B \leq 6$ را برآورده کنند، (به طوری که d_A و d_B ابعاد زیر فضاهای متناظر A و B باشد) یک شرط لازم و کافی برای جداپذیری به نام معیار Peres-Horodecki یا PPT ارایه شده است [۱۳].

در ابعاد بالاتر، این شرط فقط یک شرط لازم برای جداپذیری است و دیگر شرط کافی نیست.

این شرط بیان میکند که برای سیستم های دو بخشی، اگر ترانهاده جزئی^۱ آن سیستم نسبت به یکی از

¹Partial Transpose

زیر سیستم هایش یک مقدار ویژه منفی داشته باشد، آن حالت حتما در هم تنیده است. این نتیجه مستقل از انتخاب زیر سیستم برای گرفتن ترانهاده جزئی است [۱۳].

به عنوان مثال ماتریس چگالی ρ که زیر مجموعه فضای $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ است را در نظر بگیرید. این حالت را می توان به صورت زیر نمایش داد:

$$\rho = \sum_{ij\mu\nu} p_{\mu\nu}^{ij} (|i\rangle\langle j|) \otimes (|\mu\rangle\langle\nu|) \quad (4.1)$$

ترانهادی جزئی این سیستم نسبت به بخش B به صورت زیر می شود:

$$\begin{aligned} \rho^{TB} &= \sum_{ij\mu\nu} p_{\mu\nu}^{ij} |i\rangle\langle j| \otimes (|\mu\rangle\langle\nu|)^T \\ &= \sum_{ij\mu\nu} p_{\mu\nu}^{ij} |i\rangle\langle j| \otimes |\nu\rangle\langle\mu| \end{aligned} \quad (5.1)$$

در نتیجه برای یک حالت جداپذیر داریم:

$$\rho_{AB}^{TB} = \sum_i p_i \rho_A^{(i)} \otimes \rho_B^{T(i)} = \sum_i p_i \rho_A^{(i)} \otimes \rho_B^{(i)} \quad (6.1)$$

که این ماتریس مثبت معین است [۱۳]. بنابراین اگر $\rho_{AB}^{TB} < 0$ باشد، حالت حتما در هم تنیده است. برای سیستم های دو کیوبیتی معیار ساده تری بر اساس معیار PPT توسط Augusiak ارایه شده است [۱۴]. این معیار بر اساس درمیان ماتریس ترانهاده جزئی است و یک شرط لازم و کافی است.

$$\det(\rho_{AB}^T) > 0 \iff separable \quad (7.1)$$

۷.۱ نمایش حالت های دوکیوبیت

در حالتی که سیستم دو کیوبیتی است، یعنی از دو زیر سیستم با بعد دو، تشکیل شده است، می توان حالت سیستم را به صورت زیر نمایش داد [۶]:

$$\rho = \frac{1}{4} \sum_{ij} \Gamma_{ij} \sigma_i \otimes \sigma_j \quad (۸.۱)$$

که در این رابطه $\sigma_1, \sigma_2, \sigma_3$ ماتریس های پاولی هستند و $\sigma_0 = \mathbb{I}$ عملگر همانی است. مقدار چشم داشتی $\langle \sigma_k \otimes \sigma_l \rangle$ به این صورت محاسبه می شود:

$$\begin{aligned} \text{Tr}(\rho(\sigma_k \otimes \sigma_l)) &= \frac{1}{4} \sum_{ij} \text{tr}(\Gamma_{ij}(\sigma_i \otimes \sigma_j)(\sigma_k \otimes \sigma_l)) \\ &= \frac{1}{4} \Gamma_{kl} \text{Tr}(\mathbb{I} \otimes \mathbb{I}) = \Gamma_{kl} \end{aligned} \quad (۹.۱)$$

در نتیجه مقادیر Γ_{ij} ها را می توان به صورت زیر به دست آورد:

$$\Gamma_{ij} = \langle \sigma_i \otimes \sigma_j \rangle \quad (۱۰.۱)$$

همانطور که مشاهده می شود، این مقادیر Γ_{ij} چیزی جز مقدار چشم داشتی اندازه گیری های $(\sigma_i \otimes \sigma_j)$ نیستند.

با توجه به اینکه مقدار چشم داشتی $\mathbb{I} \otimes \mathbb{I}$ روی سیستم مقدار بدیهی $\Gamma_{..} = ۱$ را نتیجه می دهد، می توان با انجام ۱۵ اندازه گیری، مضارب Γ_{ij} را به دست آورد و در نتیجه حالت یک سیستم دو کیوبیت را به طور کامل توصیف کرد. با این وجود انجام ۱۵ اندازه گیری در آزمایشگاه هزینه بر و زمان بر است، و لازم است بتوانیم با تعداد کمتری آزمایش، حالت های در هم تنیده را تشخیص دهیم. در این پایان نامه، ما با استفاده از روش های یادگیری ماشینی، روش هایی را ارائه می دهیم که بتوان بخش قابل توجه ای از حالت های در هم تنیده را با تعداد اندازه گیری های خیلی کمتر تشخیص دهیم.

۸.۱ بررسی سایر رویکرد های موجود

روش های متعددی برای تفکیک حالت های جدپذیر و درهم تنیده وجود دارد که از آن ها می توان به روش های کلاسیکی از جمله بازسازی توموگرافیکی^۱ حالت های کوانتومی^{۱۵}، معیار ماتریس کورایانس^{۱۶}، معیار همزمانی^۲ [۱۷]–[۱۹]، عملگر های شاهد^۳ [۴]، [۲۰]، و تفکیک با استفاده از تنظیمات جمعی

¹Tomographic

²Concurrence

³Witness Operators

خاص^۱ [۲۱] اشاره کرد. همچنین روش هایی مبتنی بر یادگیری ماشین نیز ارائه شده است [۲۲] - [۳۰]. در برخی کار های دیگر روش هایی با استفاده از الگوریتم های قابل اجرا روی کامپیوتر های کوانتومی ارائه کرده اند [۳۱]، [۳۲].

باید خاطر نشان کرد در تمامی روش های یادگیری ماشینی پیشین که در بالا ذکر شده اند، این تفکیک با داشتن توصیف کامل سیستم انجام شده است. به عنوان مثال برای حالت دوکیوبیت نیاز دارند که تمامی ۱۵ اندازه گیری انجام شود. اما در روش پیشنهادی این پایان نامه فقط با استفاده از نتایج تعداد اندکی اندازه گیری این کار انجام می شود.

لازم به ذکر است که اگر توصیف کامل سیستم موجود نباشد، نمی توان حالت های تفکیک پذیر و درهم تنیده را به صورت ۱۰۰٪ کامل از یکدیگر تفکیک کرد، و فقط می توان قسمتی از آن ها را با یقین به صورت درهم تنیده یا جدا پذیر تشخیص داد. پس در واقع روش پیشنهادی این پایان نامه، شاهد های درهم تنیدگی برای تشخیص قابل توجه حالت های درهم تنیده را با تعداد کمتری اندازه گیری ارائه می دهد. اگرچه نشان می دهیم با همین روش با داشتن توصیف کامل سیستم، (یعنی هر ۱۵ اندازه گیری) می توان حالت های درهم تنیده و تفکیک پذیر را به صورت تقریباً کامل نیز از یکدیگر تفکیک کرد. همچنین نشان می دهیم با تغییر آستانه^۲، این روش می تواند معیار جداپذیری^۳ با تعداد کمتری اندازه گیری نیز ارائه دهد.

۹.۱ ساختار پایان نامه و علامت گذاری

در این پایان نامه، ابتدا به مفاهیم یادگیری ماشینی مورد نیاز برای بررسی این مسئله در فصل ۲ پرداخته می شود. این فصل به یادگیری مفاهیم ضروری یادگیری ماشین و شبکه های عصبی و روش های بررسی نتایج آن ها می پردازد. سپس در فصل سوم ابتدا به توضیح نحوه تولید داده ها و بررسی های آماری آن ها می پردازیم. در ادامه سعی می کنیم با مدل های یادگیری ماشین کلاسیک و شبکه های عصبی، ابتدا با داشتن توصیف کامل سیستم به تفکیک تقریباً کامل حالت های جداپذیر و درهم تنیده برسیم. سپس مدل های جدیدی برای تشخیص حالت های درهم تنیده با تعداد کمتری آزمایش غیر خطی، و روش دیگر با تعداد کمتری آزمایش خطی، ارائه کرده و نتایج هر یک را بررسی می کنیم. پس از آن با فرض تقارن های آزمایشگاهی، مدل را مجدد آموزش می دهیم و نتایج بهبود یافته ای به دست می آوریم. پس از آن، تعدادی از مدل های یادگیری نظارت نشده و شبه نظارت شده را بررسی می کنیم. در آخر نیز به پایداری سیستم و اثر نوفه های مخرب بر آن می پردازیم. در فصل آخر نیز، به جمع بندی مطالب می پردازیم و پیشنهادات موجود برای ادامه کار تبیین می شود.

¹Collective settings

²Threshold

³Separability Criterion

همچنین در این پایان نامه، عملگرها و ماتریسها به جز ماتریسهای چگالی با حروف بزرگ نشان داده شده اند (به طور مثال U نمایانگر یک عملگر و ρ نمایانگر یک ماتریس چگالی می باشد) و برای بردارها نمایش برا-کتی دیراک (به عنوان مثال $|\psi\rangle$) در نظر گرفته شده است.

فصل ۲

یادگیری ماشینی

۱.۲ مقدمه

در این فصل به بیان مفاهیم ضروری هوش مصنوعی شامل یادگیری ماشین نظارت شده و نظارت نشده، شبکه های عصبی و شبکه های خود رمزگذار می پردازیم. از هر کدام از این روش ها، مفاهیم تعدادی را توضیح می دهیم و سپس نحوه ی ارزیابی مدل از جمله صحت^۱، دقت^۲، بازیابی^۳، ماتریس درهم ریختگی^۴، و تفاوت های هر یک گفته می شود.

۲.۲ یادگیری ماشین چیست؟

یادگیری ماشین زیر مجموعه ای از هوش مصنوعی است که با استفاده از داده های ارایه شده توسط کاربر، به یادگیری مساله ای می پردازد. از این روش معمولاً برای تعیین خروجی مطلوب، با دادن ورودی مورد نظر استفاده می شود. تفاوت عمده ی آن با الگوریتم های کلاسیکی این است که در الگوریتم های کلاسیکی، برنامه نویس روش عملکرد ماشین بر داده های ورودی را طراحی و پیاده سازی می کند. اما در یادگیری ماشین، اینکه چگونه داده های ورودی را به خروجی مطلوب تبدیل کند، به ماشین داده نمی شود و خود ماشین طبق الگوریتم و روش هایی، نحوه تبدیل ورودی به خروجی را یاد میگیرد (شکل ۱.۲^۵). مثلاً با دیدن مثال های آماده شده و یا دسته بندی داده ها با توجه به یک سری ویژگی مشترک.

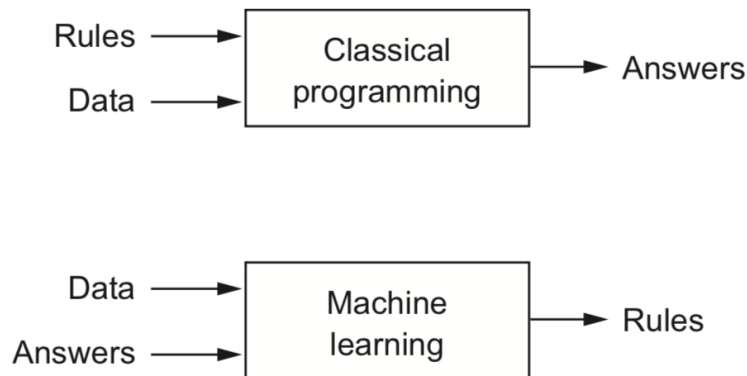
¹Accuracy

²Precision

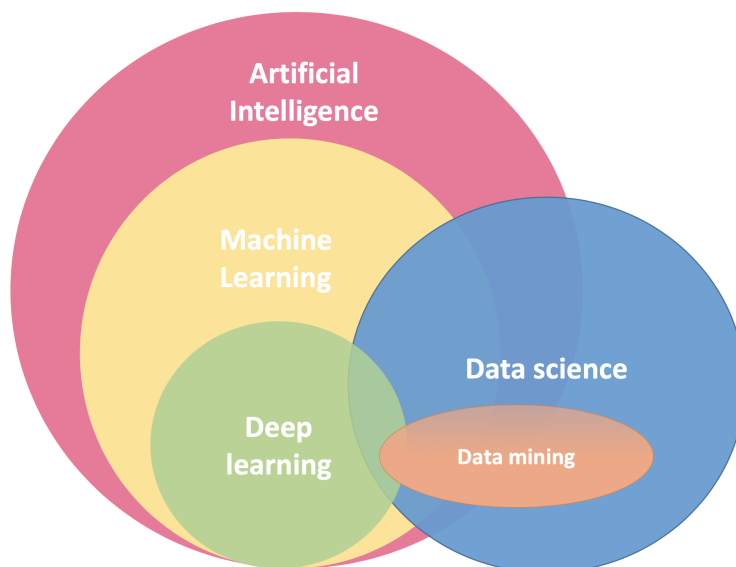
³Recall

⁴Confusion Matrix

⁵<https://medium.com/@artin.sinani/dont-be-intimidated-it-s-only-machine-learning-158f703ab387>



شکل ۱.۲: تفاوت برنامه نویسی کلاسیک با یادگیری ماشینی



شکل ۲.۲: ارتباط بین هوش مصنوعی، یادگیری ماشینی، شبکه های عصبی، یادگیری عمیق و علم داده. [۱]

الگوریتم های بسیار زیادی در یادگیری ماشینی وجود دارد، که دسته ای از آن ها از روش های ریاضی به نسبت قدیمی تر استفاده می کنند [۳۳]، [۳۴] و دسته ای از آن ها اخیراً محبوبیت زیادی پیدا کرده اند که به شبکه های عصبی معروف هستند [۳]. این شبکه ها که معمولاً به صورت لایه لایه هستند، می توانند تعداد اندکی لایه داشته باشند (که به آن ها یادگیری کم عمق می گویند) و یا تعداد زیادی لایه دارند (که به آن ها یادگیری عمیق [۳۵] گفته می شود).

یادگیری ماشینی کاربرد های بسیار زیادی در همه ی علوم از جمله علوم کامپیوتر، علوم زیستی و بیولوژی، فیزیک، شیمی، زبان شناسی اقتصاد و غیره دارد. در علم داده نیز با استفاده از یادگیری ماشینی، با بررسی داده های ساختار یافته یا ساختار نیافته، به کسب اطلاعاتی از مساله پرداخته می شود. در شکل ۲.۲ ارتباط هوش مصنوعی، یادگیری ماشینی، شبکه های عصبی، یادگیری عمیق و علم داده نشان داده شده است.

۳.۲ انواع یادگیری ماشین

در این بخش به سه دسته کلی روش های یادگیری ماشینی با توجه به برجسب گذاری داده ها می پردازیم.

۱.۳.۲ یادگیری نظارت شده

اگر دسته بندی داده های آزمایش، از قبل به نوعی مشخص شده باشد، یا به عبارتی همه ی آن ها برجسب گذاری شده باشند، و الگوریتم ماشین از این برجسب ها برای یادگیری استفاده کنند، به این روش یادگیری نظارت شده می گویند. تعدادی از این روش ها که در این پایان نامه استفاده شده اند عبارتند از: Ridge, Random Forest, Support Vector, Bagging, Gradient Boosting و غیره. همچنین دسته ای از شبکه های عصبی پیشخور^۱ (FFNN)، مانند شبکه های عصبی پیچشی^۲ (CNN) و شبکه های عصبی بازگشتی^۳ (RNN) نیز از دسته ی یادگیری نظارت شده هستند.

۲.۳.۲ یادگیری نظارت نشده

اگر الگوریتم برای دسته بندی، فقط از ویژگی های هر نمونه استفاده کند و استفاده ای از برجسب داده ها نکند، به آن ها یادگیری نظارت نشده می گویند. تعدادی از این روش ها که در این پایان نامه استفاده شده اند عبارتند از: Birch، DBSCAN، K-means، OPTICS. همچنین از روش های شبکه عصبی نظارت نشده می توان به Boltzman Machine ها اشاره کرد، که دسته ای مهم از روش های یادگیری نظارت نشده هستند.

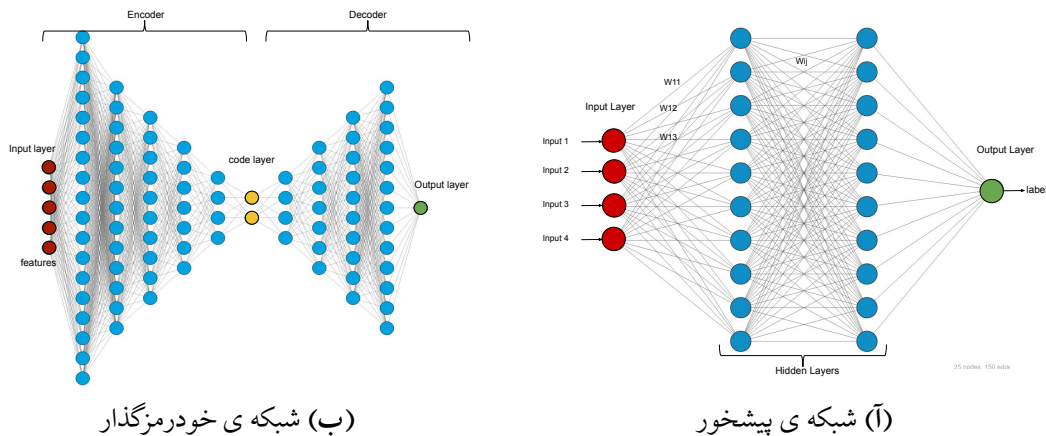
۳.۳.۲ یادگیری شبه نظارت شده

دسته ای از روش ها هم هستند که ترکیب این دو اند. یعنی قسمتی از داده ها برجسب گذاری، و قسمت دیگری برجسب گذاری نشده اند. این روش ها با استفاده از برجسب های موجود، سعی در بهتر کردن نتیجه دارند. روش های شبه نظارت شده ی مورد استفاده در این پایان نامه عبارتند از LabelPropagation، LabelSpreading، و Anomaly Detection. در ادامه مبانی تعدادی از این روش ها به اختصار توضیح داده می شود.

^۱Feedforward Neural Network

^۲Convolutional Neural Network

^۳Recursive Neural Network



(ب) شبکه ی خودرمزگذار

(آ) شبکه ی پیشخور

شکل ۳.۲: معماری انواع شبکه های عصبی

۴.۲ شبکه های عصبی

این روش متشکل از شبکه ای از گره ها و یال ها است که به صورت لایه لایه مطابق شکل ۳.۲ است [۳] و به این صورت عمل می کند که ویژگی های موجود به صورت پارامتر های اولیه به لایه ی اول شبکه داده می شود. گره ها هر کدام مانند حافظه ای هستند که می توانند یک عدد اعشاری را در خود ذخیره کنند. هر یک از یال ها، در واقع ضریب ضرب شونده در گره لایه ی قبل است. در نتیجه مقدار اولیه هر گره در لایه ی بعد، ترکیب خطی از مقادیر گره های لایه های قبل است. پس از به دست آمدن این ترکیب خطی، یک تابع غیر خطی بر این مقدار اعمال می شود که به این تابع غیر خطی، تابع فعال سازی^۱ گفته می شود. توابع متفاوتی برای تابع فعال سازی استفاده می شود که از جمله آن ها می توان به توابع زیر اشاره کرد:

$$\text{ReLU} = \max \{0, x\} \quad (۱.۲)$$

$$\text{Sigmoid} = \frac{1}{1 + e^{-x}} \quad (۲.۲)$$

$$\text{Gaussian} = e^{-x^2} \quad (۳.۲)$$

۵.۲ شبکه های خود رمزگذار

شبکه های خود رمزگذار، نوعی از شبکه های عصبی هستند که به طرز بهینه ای داده ها را کدگذاری می کنند و بعد فضا را کاهش می دهند [۳۶]. معماری این شبکه ها به این صورت است که ابتدا تعداد گره های در

^۱Activation Function

لایه های اول زیاد و در لایه های بعدی کم می شود تا اینکه به تعداد کمی گره ی مطلوب برسد. سپس مجدداً تعداد گره ها شروع به افزایش می کند تا به تعداد گره ی مطلوب برسد. در شکل ۳.۲ ب^۱ معماری این شبکه به تصویر کشیده شده است. در این صورت لایه ی میانی که دارای کمترین گره است، به صورت یک گلوگاه برای جریان اطلاعات عمل می کند. در نتیجه، تمامی اطلاعاتی که ورودی را به خروجی تبدیل می کند، مجبور است از گلوگاه عبور کند و به عبارتی در این تعداد کم گره ها کد شود. به قسمت اول این شبکه که اطلاعات ورودی را در تعداد کم گره ی گلوگاه کد می کند، کدگذار، و به قسمت دوم این شبکه که با استفاده از مقادیر گره های گلوگاه، خروجی را بازسازی می کند، کدگشا گفته می شود.

۶.۲ تشخیص ناهنجاری

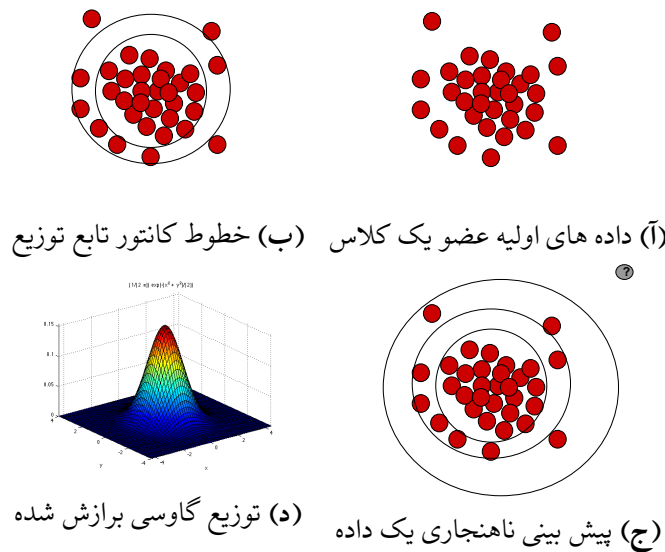
در علم بررسی داده، به تشخیص حالت ها یا اتفاقات نادر که با غالب مشاهدات پیشین منطبق نیست، و مقدار زیادی تفاوت دارد، تشخیص ناهنجاری می گویند [۳۷]. این مساله کاربرد های بسیاری از جمله تشخیص تقلب (مثلاً تقلب بانکی)، تشخیص نقص مقاومت سازه، مسائل پزشکی، خطاهای متنی و غیره دارد [۳۸]. دسته ای از تشخیص ناهنجاری با یادگیری ماشینی انجام می شود، که از آن ها می توان به تشخیص ناهنجاری نظارت نشده و تشخیص ناهنجاری شبه نظارت شده اشاره کرد.

در تشخیص ناهنجاری نظارت نشده، ماشین با فرض اینکه اکثر داده ها توزیع نرمال دارند، و فقط ناهنجاری ها خارج از این توزیع هستند، یک توزیع نرمال بهینه برای اکثریت داده ها پیدا میکند، و دسته ای که خارج از این توزیع هستند به عنوان ناهنجاری شناخته می شوند.

در روش توزیع ناهنجاری شبه بهینه، داده های یک کلاس به ماشین داده می شود، و ماشین با برازش کردن یک منحنی نرمال به صورت بهینه، آموزش داده می شود. در این حالت ماشین فقط با داده های یک کلاس آموزش داده شده است، و داده هایی که برچسب کلاس دیگر را دارند مشخص نیستند. به همین خاطر این یک روش شبه نظارت شده است. سپس داده های آزمایشی به ماشین داده می شوند و ماشین با توجه به اینکه این داده در کجای این توزیع قرار می گیرد، آن را به عنوان یک داده ی معمولی (هم برچسب با کلاس آموزش داده شده) یا به عنوان یک داده ی پرت یا ناهنجاری (غیر هم برچسب با کلاس آموزش داده شده) پیش بینی می کند.

در شکل ۴.۲ نحوه ی عملکرد این روش با مثالی آورده شده است. همانگونه که در شکل ۴.۲ آ مشاهده می شود، دسته ای از داده های هم برچسب از یک کلاس، داده می شود. سپس ماشین با توجه به توزیع داده ها در فضا، یک توزیع نرمال برای این داده ها پیدا می کند. خطوط کانتور این توزیع در شکل ۴.۲ ب و نمودار

¹<https://laptrinhx.com/autoencoders-3518090441/>



شکل ۴.۲: نحوه تشخیص ناهنجاری با استفاده از برازش توزیع گاوسی به داده ها

سه بعدی آن در شکل ۴.۲ آورده شده است. در نهایت اگر داده ای برای پیش بینی به ماشین داده شود، مقدار تابع گاوسی برای نقطه ی جدید محاسبه می شود و اگر از آستانه ای کمتر شد، این داده به عنوان یک ناهنجاری در نظر گرفته می شود (شکل ۴.۲ ج)

۷.۲ متریک های ارزیابی

در یادگیری ماشینی، مفاهیم متفاوتی برای بررسی نتایج عملکرد مدل ها بر اساس مفاهیم آماری استفاده می شود که هر یک معنی به خصوصی دارند. در این قسمت تعدادی از مفاهیم استفاده شده در این پایان نامه را شرح می دهیم.

فرض کنید داده های مساله دو دسته ی A و B هستند، و ماشین بر روی این داده ها آموزش داده می شود. سپس دسته ای از داده ها به ماشین آموزش یافته، داده می شود تا آن ها را برچسب گذاری کند. در این صورت برای پیش بینی برچسب A چهار حالت مختلف رخ می دهد (جدول ۱.۲). در صورتی که داده از کلاس A باشد و ماشین هم به درستی برچسب کلاس A را پیش بینی کند، این یک پیش بینی مثبت درست (TP) است.^۱ اگر داده از کلاس B باشد و ماشین هم مجدداً به درستی کلاس B را پیش بینی کند، حالت منفی درست (TN)^۲ رخ می دهد. اگر برچسب داده A باشد و ماشین اشتباهاً آن را به عنوان کلاس B برچسب

^۱True-Positive

^۲True-Negative

گذاری کند، حالت منفی نادرست (FN) ^۱ رخ داده است. و در آخر آگه برچسب داده از نوع B باشد اما ماشین اشتباها به عنوان کلاس A تشخیص بدهد، در این صورت مثبت نادرست (FN) ^۲ که به آژیر کاذب نیز معروف است رخ داده است.

حال با استفاده از این چهار مفهوم مختلف، انواع متریک های تعریف شده برای ارزیابی نتایج مدل را تعریف می کنیم. معروف ترین متریک برای مدل های یادگیری ماشینی صحت ^۳ است. صحت ماشین در واقع نسبت تعداد حالت هایی هستند که ماشین برچسب آن ها را به درستی تشخیص داده، نسبت به تعداد کل داده ها است. به عبارتی صحت می شود:

$$\text{Accuracy} = \frac{TP + TN}{P + N} = \frac{TP + TN}{TP + TN + FP + FN} \quad (۴.۲)$$

این متریک معمولاً برای زمانی که داده های هر دو کلاس فراوانی تقریباً مشابهی دارند مناسب است. اما فرض کنید می‌خواهیم دستگاهی داشته باشیم که قاتلی را از میان افراد یک شهر تشخیص دهد. تعداد قاتلین موجود در شهر، خیلی خیلی کمتر از تعداد کل افراد شهر است. بنابراین اگر ماشین همه را بی گناه اعلام کند، با توجه به این متریک، درصد خیلی بالایی از پیش بینی هایش درست است (فرض کنید ۹۹/۹٪). اما آیا این پیش بینی ارزشی دارد؟ قطعاً خیر. پس برای ارزیابی چنین مدلی به متریک دیگری احتیاج داریم.

متریک معروف دیگر، دقت ^۴ است. دقت در مثال قبل به این معنا است که اگر ماشین، شخصی را به عنوان قاتل اعلام کند، به چه احتمالی این شخص واقعا قاتل است. به عبارتی اگر ماشین تعداد n نفر را قاتل اعلام کند، چند نفر از این n نفر واقعا قاتل هستند. این متریک همانگونه که دیده می شود، در این مساله خیلی مهم است. به این خاطر که اگر دقت پیش بینی پایین باشد، فرد بی گناهی به عنوان قاتل شناسایی می شود. تعریف ریاضی دقت بنا به مفاهیم گفته شده به صورت زیر است:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (۵.۲)$$

در آخر، متریک بازیابی ^۵ به معنی تعداد داده تشخیص داده شده به عنوان کلاس A نسبت به کل تعداد کلاس A است. در مثال قبل، بازیابی در واقع تعداد افراد قاتل تشخیص داده شده، نسبت به کل تعداد قاتلین شهر است. به صورت ریاضی، بازیابی عبارت است از:

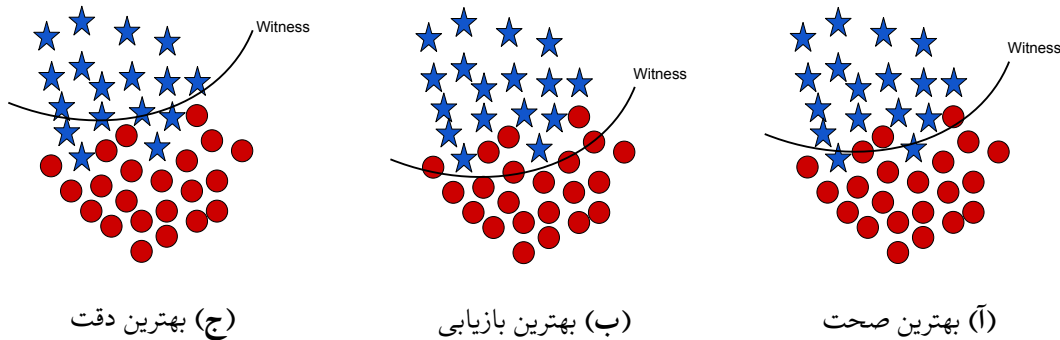
^۱False-Negative

^۲False-Negative

^۳Accuracy

^۴Precision

^۵Recall



شکل ۵.۲: مصالحه دقت و بازیابی. با تغییر آستانه که معادل جابجایی مرز تشخیص است، می توان به دقت بالاتر یا بازیابی بالاتر دست یافت.

$$\text{Recall} = \frac{TP}{P} \quad (۶.۲)$$

متریک های دقت و بازیابی با دیگر مصالحه دارند. به این معنی که می توان با افزایش آستانه ی مدل برای شناسایی A، میزان دقت را بالا برد، اما معمولاً در این حالت میزان بازیابی کم می شود. همچنین اگر آستانه مدل برای تشخیص حالت A را کم کنیم، مدل بازیابی بیشتری پیدا میکند، اما معمولاً دقت کم می شود. پس با تنظیم آستانه مدل، می توان به مقدار دقت دلخواه یا بازیابی دلخواه دست یافت. این مساله در شکل ۵.۲ قابل مشاهده است.

۱.۷.۲ ماتریس درهم ریختگی

ماتریس درهم ریختگی^۱ یا ماتریس خطا، جدولی برای پیش بینی مدل در حوزه ی یادگیری ماشینی است. در این ماتریس هر سطر بیانگر یک کلاس تشخیص داده شده توسط ماشین، و هر ستون بیانگر یک کلاس واقعی داده ها است. در حالت ایده آل که همه ی حالت ها توسط ماشین درست تشخیص داده شده اند، این ماتریس قطری می شود. در جدول ۱.۲ این ماتریس بر حسب چهار حالت گفته شده در بخش قبل توصیف شده است.

۸.۲ نتیجه گیری

در این فصل مفاهیم و ابزارهای مورد نیاز در حوزه ی یادگیری ماشینی برای بررسی مساله جدایی پذیری بیان شد. در فصل بعد، به ارزیابی روش تشخیص حالت های درهم تنیده با استفاده از یادگیری ماشینی به کمک مفهوم های مطرح شده پرداخته می شود.

¹Confusion Matrix

	P	N
Predicted P	TP	FP
Predicted N	FN	TN

جدول ۱.۲: ماتریس درهم ریختگی

فصل ۳

تشخیص حالت های درهم تنیده از جدپذیر

۱.۳ مقدمه

با توجه به پیش نیاز های گفته شده در فصل قبل، حال آماده ایم تا به مسئله ی اصلی بپردازیم. در ادامه، ابتدا به نحوه تولید داده ها و برخی بررسی های آماری انجام گرفته روی آن ها می پردازیم. سپس مدل های یادگیری ماشینی کلاسیکی را آموزش می دهیم و نتایج آن ها و زمان یادگیری هر یک را با یکدیگر مقایسه می کنیم. در ادامه با استفاده از روش شبکه های عصبی و شبکه های خودرمزگذار مدل را با تعداد کمی اندازه گیری آموزش می دهیم. سپس با استفاده از تقارن استوانه ای به بررسی مجدد موضوع می پردازیم. همچنین اثر نوفه بر این داده ها را بررسی کرده ایم. در آخر روش تشخیص ناهنجاری^۱ که از روش های یادگیری نظارت نشده است را بر روی داده ها بررسی می کنیم.

۲.۳ تولید داده ها

برای این که بتوانیم از یادگیری ماشین استفاده کنیم، نیاز است چند میلیون حالت کوانتومی برای این کار استفاده شود. ما این داده ها را به صورت مصنوعی، با انتخاب تصادفی از فضای هیلبرت سیستم تولید کردیم. نحوه ی تولید داده ها با استفاده از کتابخانه ی qutip در پایتون به این صورت است که یک ماتریس هرمیتی تصادفی H که 4×4 است انتخاب می کنیم. سپس $P = H^\dagger H$ را محاسبه کرده و ماتریس نرمال شده ی N را به صورت $N = \frac{P}{\text{Tr } P}$ محاسبه می کنیم. با استفاده از این ماتریس و تابع Qobj کتابخانه ی qutip یک ماتریس چگالی برای سیستم دوکیوبیتی تولید می کنیم.

¹Anomaly Detection

پس از آنکه با این روش ۵ میلیون ماتریس چگالی تصادفی تولید کردیم، ویژگی های مورد نظر که همان Γ_{ij} ها بودند را توسط معادله ی ۱۰.۱ محاسبه کردیم:

$$\Gamma = \begin{bmatrix} \langle \sigma_x \otimes \sigma_x \rangle & \langle \sigma_x \otimes \sigma_y \rangle & \langle \sigma_x \otimes \sigma_z \rangle & \langle \sigma_x \otimes \mathbb{I} \rangle \\ \langle \sigma_y \otimes \sigma_x \rangle & \langle \sigma_y \otimes \sigma_y \rangle & \langle \sigma_y \otimes \sigma_z \rangle & \langle \sigma_y \otimes \mathbb{I} \rangle \\ \langle \sigma_z \otimes \sigma_x \rangle & \langle \sigma_z \otimes \sigma_y \rangle & \langle \sigma_z \otimes \sigma_z \rangle & \langle \sigma_z \otimes \mathbb{I} \rangle \\ \langle \mathbb{I} \otimes \sigma_x \rangle & \langle \mathbb{I} \otimes \sigma_y \rangle & \langle \mathbb{I} \otimes \sigma_z \rangle & \langle \mathbb{I} \otimes \mathbb{I} \rangle \end{bmatrix}$$

سپس این ویژگی ها را به صورت یک سطر 1×15 با حذف ۱۶ امین ویژگی بدیهی ۱ حاصل از $\mathbb{I} \otimes \mathbb{I}$ در آوردیم. این کار برای تمامی ۵ میلیون ماتریس چگالی انجام شد.

در قسمت اول این پروژه، از روش یادگیری نظارت شده برای تفکیک حالت ها استفاده کردیم. بنابراین در این روش ها نیاز است که برچسب^۱ هر یک از داده ها مشخص شود (که به معنی جداپذیر بودن یا درهم تنیده بودن است). با توجه به اینکه در حالت دوکیوبیت این کار با معادله ی ۷.۱ و داشتن تمامی ۱۵ ویژگی امکان پذیر است، ما توانستیم تمامی داده ها را برچسب گذاری کنیم. برچسب گذاری به صورت عدد ۰ برای جداپذیر ها و برچسب عدد ۱ برای درهم تنیده ها صورت گرفت.

تمامی این برچسب ها در ستون آخر (۱۶ ام) برای هر یک از ماتریس های چگالی ذخیره شد. پس داده نهایی به صورت یک ماتریس با ۵ میلیون سطر و ۱۶ ستون به دست آمد. این داده ها را به صورت عمومی در Gitlab بر روی اینترنت^۲ قرار داده ایم و برای عموم قابل دسترس کرده ایم.

۳.۳ بررسی آماری داده ها

بعد از تولید داده ها، نمودار های هیستوگرام را برای متغیر های متفاوت بررسی کردیم و اطمینان حاصل کردیم که حالت ها فضا را کاملاً پوشش داده اند و توزیع های فراوانی مقارنی نسبت به ابعاد فضا دارند. همانطور که در شکل ۱.۳ قابل مشاهده است، هر یک از ویژگی ها دارای توزیع شبه نرمال می باشند.

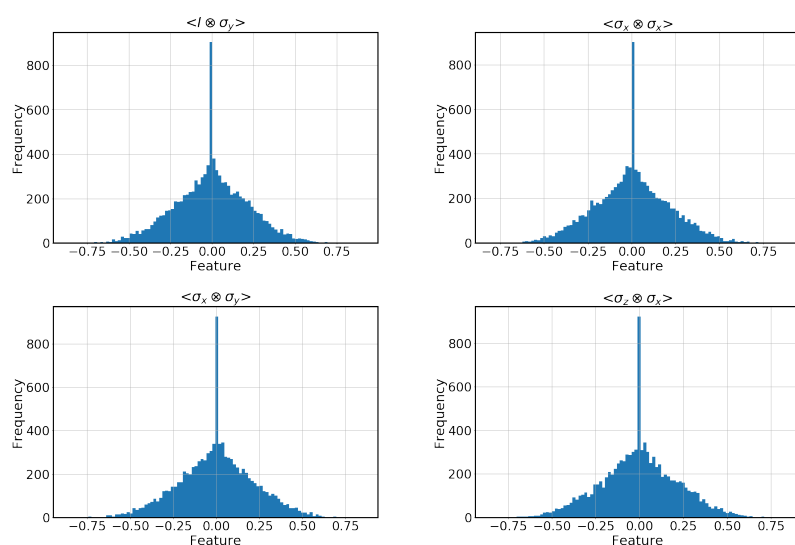
همچنین ماتریس های همبستگی^۳ این داده ها در شکل ۳.۳ آورده شده است. همانطور که مشاهده می شود هیچ گونه همبستگی خطی بین داده ها دیده نمی شود. در ماتریس پخشی^۴ هم هیچ الگوی مشخصی دیده نمی شود و برای هر دو ویژگی متفاوت دارای روند یکسانی است (شکل ۲.۳).

¹Label

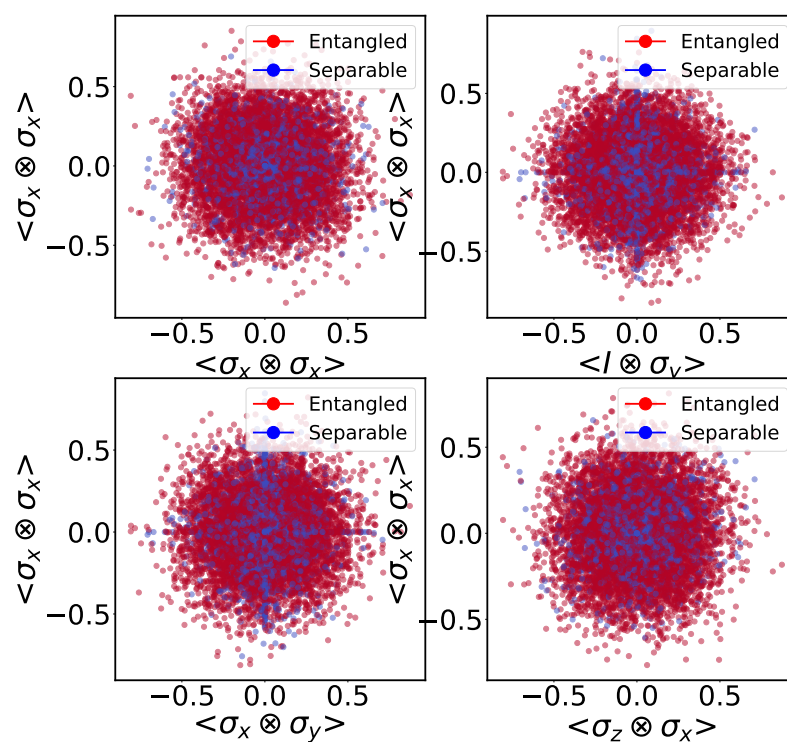
²<https://gitlab.com/quantum-machine-learning/autoencoder-entanglement-detection-optimizer>

³Correlation Matrix

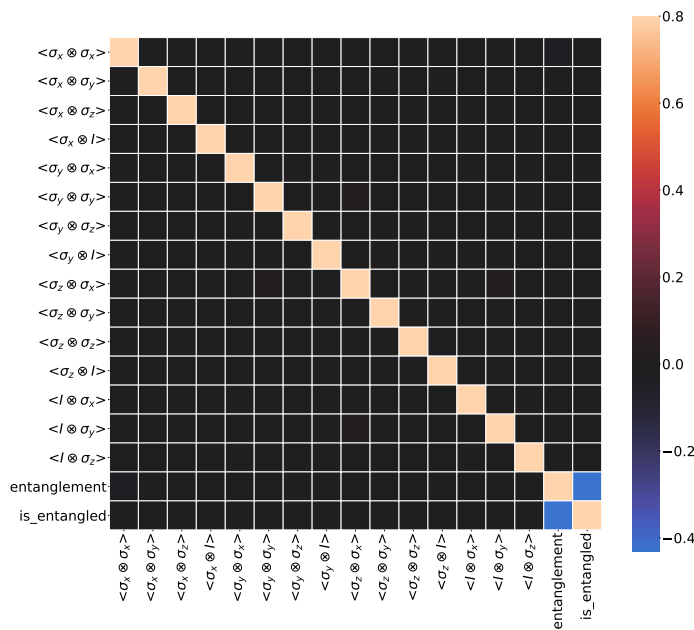
⁴Scatter Matrix



شکل ۱.۳: هیستوگرام ویژگی های مختلف. همه ی ۱۵ ویژگی دارای توزیع یکسان شبه گاوسی هستند.



شکل ۲.۳: ماتریس پخش ویژگی های مختلف نسبت به هم.



شکل ۳.۳: ماتریس همبستگی. این ماتریس همبستگی ویژگی های مختلف نسبت به همدیگر را نشان می دهد.

۴.۳ روش های کلاسیک

در این بخش با استفاده از روش های کلاسیکی یادگیری ماشینی، به تفکیک حالت های جداپذیر و درهم تنیده پرداخته ایم. روش انجام شده در اینجا به دو دسته کلی تقسیم می شود. دسته اول مساله دسته بندی^۱ با استفاده از برچسب ۰ و ۱ (درهم تنیده یا جداپذیر بودن است) در این مساله میزان درهم تنیدگی به ماشین داده نمی شود. در دسته دوم، یک مساله ی رگرسیون برای تخمین دترمینان ماتریس ترانواده جزئی انجام می دهیم، به طوری که ۱۵ ویژگی به همراه دترمینان ماتریس ترانواده جزئی به ماشین داده می شود، و ماشین با استفاده از این اعداد سعی در یادگیری میزان درهم تنیده بودن میکند. در آخر، زمانی که ماشین آموزش داده شده حالت جدیدی را پیش بینی کرد، با توجه به مثبت یا منفی بودن این عدد، درهم تنیده بودن یا جدا پذیر بودن حالت را استنباط می کنیم. پس در واقع از یک رگرسور، یک تفکیک گر^۲ می سازیم.

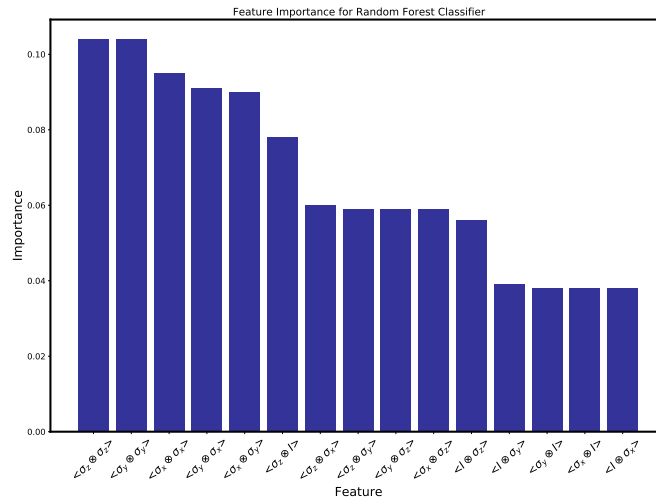
هر یک از این دسته ها با روش های متعددی انجام شده است که نتایج هر یک و زمان یادگیری و پیش بینی آن ها در جدول ۱.۳ آورده شده است.

¹Classification

²Classifier

Accuracy	Prediction time	Fit time	Samples	Estimation Method
۰/۹۹۶۹	۰/۰۶	۸۵/۱۳	۲۰۰۰۰	Linear Regression + polynomial transform
۰/۹۹۱۰	۰/۰۴	۱۱/۹۶	۲۰۰۰۰	Ridge Regression + polynomial transform
۰/۸۸۲۸	۷/۱۷	۴۸۹/۲۷	۱۰۰۰۰۰	Random Forest Classification
۰/۸۷۹۶	۴/۶۳	۳۱/۲۴	۵۰۰۰۰	Support Vector Classification
۰/۸۷۶۲	۰/۴۳	۵/۲۵	۱۰۰۰۰	Support Vector Regression
۰/۸۴۷۴	۰/۲۲	۱۱/۷۶	۵۰۰۰۰	Random Forest Classification
۰/۸۴۷۰	۰/۳۵	۷۲/۶۵	۵۰۰۰۰	Random Forest Regression
۰/۸۴۶۸	۰/۰۳	۱/۴۶	۲۰۰۰۰	Linear classification with SGD training
۰/۸۳۲۵	۰/۲۰	۷/۶۵	۱۰۰۰۰	Support Vector Classification (degree = 4)
۰/۸۳۱۹	۰/۰۲	۵/۱۸	۵۰۰۰۰	Bagging Classification
۰/۸۲۷۱	۰/۰۱	۷/۰۱	۵۰۰۰۰	Gradient Boosting Classification
۰/۸۲۶۷	۰/۰۳	۱۲/۳۸	۲۰۰۰۰	Ridge Classification
۰/۸۲۱۳	۰/۰۰۳	۰/۰۱۷	۵۰۰۰۰	Gaussian Naive Bayes
۰/۸۱۴۵	۰/۱۰	۰/۴۸	۵۰۰۰۰	Extra Trees Classification
۰/۸۱۱۰	۰/۰۵	۴/۲۳	۵۰۰۰۰	Ada Boost Classification
۰/۷۹۹۴	۸/۴۳	۰/۰۴۷	۵۰۰۰۰	K-Neighbors Classification
۰/۷۸۱۶	۰/۳۸	۱/۵۴	۱۰۰۰۰	K-Neighbors Regression
۰/۷۶۱۸	۰/۰۰۶	۱/۳۴	۵۰۰۰۰	Decision Tree Regression

جدول ۱.۳: نتایج مربوط به روش های کلاسیکی یادگیری ماشینی از جمله تعداد نمونه ی یادگیری، زمان آموزش، زمان آزمایش، و صحت ارزیابی



شکل ۴.۳: اهمیت ویژگی ها برای مدل درخت تصادفی. هر کدام از ویژگی ها برای تعیین نتیجه اهمیت متفاوتی در مدل دارند.

۱.۴.۳ اهمیت ویژگی ها

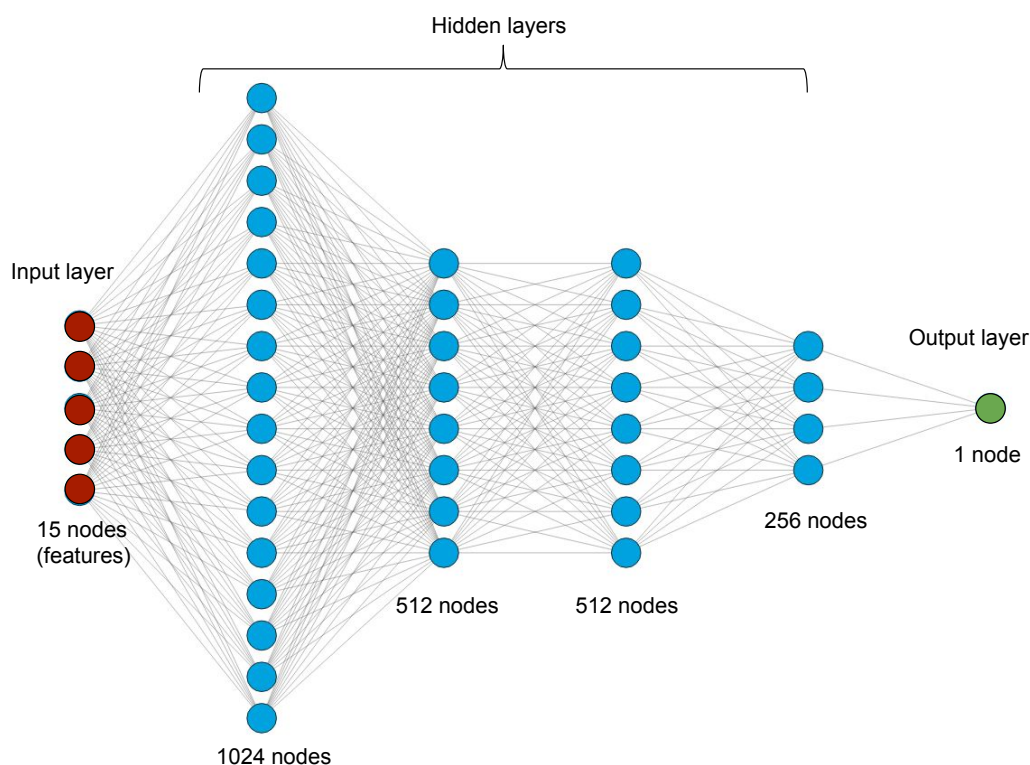
در این روش های کلاسیکی، از اندازه ضرایب به دست آمده می توان اهمیت هر یک از ویژگی ها را به دست آورد. همچنین روش های درختی اهمیت ویژگی ها را ارایه می دهند. هرچه اهمیت یک ویژگی بیشتر باشد، حاوی اطلاعات بیشتری برای تفکیک مساله به دست می دهد. بنابراین اگر ویژگی های با اهمیت بیشتر حذف شوند، صحت تفکیک افت بیشتری می کند.

در بررسی صورت گرفته مشاهده کردیم که همه ی ویژگی ها به یک اندازه مهم نیستند و برخی از آن ها حاوی اطلاعات بیشتری برای تفکیک حالت های جداپذیر و درهم تنیده هستند. ویژگی هایی که شامل ضرب تانسوری دو ماتریس پاولی یکسان بودند، همه در رده ی اول اهمیت هستند. این ها عبارتند از:

$$\sigma_x \otimes \sigma_x, \sigma_y \otimes \sigma_y, \sigma_z \otimes \sigma_z$$

پس از این ها، ویژگی هایی که شامل ضرب تانسوری دو ماتریس پاولی در ابعاد مختلف هستند، دارای اهمیت رده ی دوم می باشند، به عنوان مثال $\sigma_x \otimes \sigma_z$ ، $\sigma_x \otimes \sigma_y$. در آخر نیز ویژگی هایی که شامل ضرب یک ماتریس پاولی و یک عملگر همانی اند در رده سوم و آخر اهمیت هستند.

این اهمیت ویژگی ها برای یک نمونه از روش تفکیک تصادفی درختی در شکل ۴.۳ آورده شده است.



شکل ۵.۳: معماری شبکه عصبی پیشخور ارایه شده برای این مساله. تعداد گره ها و لایه ها در شکل نمایش داده شده است.

۵.۳ شبکه های عصبی

یکی دیگر از روش های یادگیری ماشینی، روش شبکه های عصبی^۱ و یادگیری عمیق^۲ است. در ابتدا برای اینکه بفهمیم آیا این روش قادر به یادگیری بهینه با تعداد اندازه گیری کمتر است، این موضوع را بررسی می کنیم که آیا اصلاً شبکه قادر است با داشتن هر ۱۵ اندازه گیری (توصیف کامل سیستم)، حالت های درهم تنیده و جدپذیر را از یکدیگر تفکیک کند یا خیر. در ادامه خواهیم دید که شبکه می تواند به طور تقریباً کامل، تمامی حالت های درهم تنیده و جدپذیر را از یکدیگر تفکیک کند.

معماری های متفاوتی از شبکه های عصبی برای یادگیری انجام شد و یکی از معماری هایی که بهترین نتیجه را منجر میشود در شکل ۵.۳ آورده شده است. همانطور که مشاهده می شود، در این معماری در ابتدا تعداد گره های بیشتری استفاده شده، و به مرور در لایه ها تعداد گره ها کمتر می شوند. تابع فعالسازی استفاده شده در این معماری تابع ReLU است.

¹Neural Networks

²Deep learning

بعد از آموزش شبکه با ۸۰٪ از ۵ میلیون داده، شبکه با ۲۰٪ داده های دیگر که مدل برچسب آن داده ها را ندیده است، آزمایش می شود و صحت برچسب پیش بینی شده با برچسب واقعی مقایسه می شود. با این کار، مشاهده کردیم که مدل آموزش داده شده، بیشتر از ۹۹٪ داده ها درست پیش بینی می کند. اگرچه در ظاهر این پیش بینی حالت ها توسط ماشین، با توجه به داشتن معیار PPT بدون ارزش افزوده ای بنظر می رسد، اما نتیجه ی جالبی که از این مشاهده منجر می شود این است که ماشین بدون اینکه اطلاعی از معیار PPT داشته باشد، فقط با دیدن نمونه هایی از حالت های درهم تنیده و جداپذیر، توانسته است این معیار یا معیاری هم ارز PPT را بدست آورد که نتیجه ی حیرت انگیزی است.

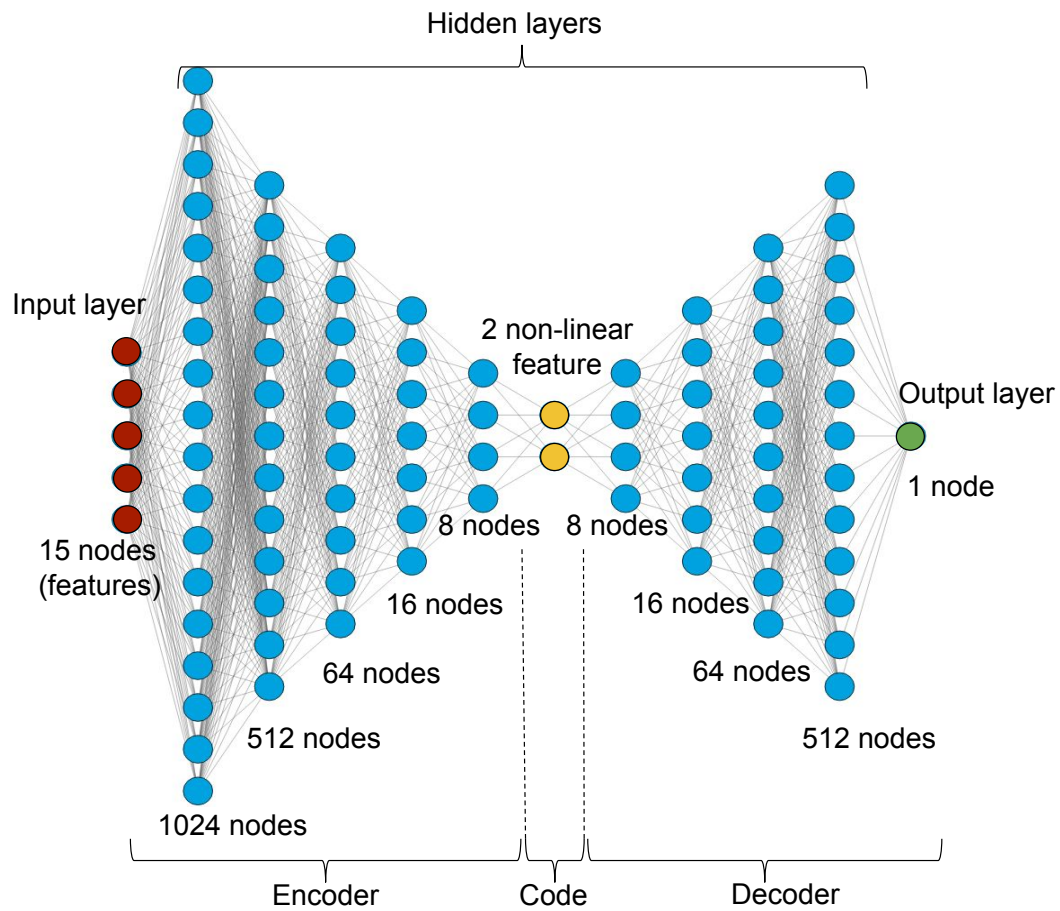
۶.۳ شبکه عصبی خود رمزگذار

در قسمت قبل، توانستیم با داشتن ۱۵ اندازه گیری، به طور تقریباً کامل حالت ها را از یکدیگر تفکیک کنیم. در این قسمت، معماری دیگری بر اساس شبکه های عصبی خود رمزگذار ارائه می دهیم که بتواند با تعداد کمتری اندازه گیری این کار را انجام دهد. معماری ارائه شده در شکل ۶.۳ آورده شده است. در بخش میانی این شبکه، یک گلوگاه دو گره ای ایجاد شده است، تا همه ی اطلاعات مورد نیاز برای تشخیص در هم تنیدگی، از این گلوگاه عبور کند. این شبکه را مجبور کرده که بخش اول که به عنوان کدگذار شناخته می شود، یعنی از لایه ی اول تا گلوگاه، و بخش دوم که به عنوان کدگشا شناخته میشود، یعنی از گلوگاه تا خروجی را به صورت بهینه برای تفکیک حالت ها تنظیم کند. با توجه به اینکه دو گره میانی، هر یک تابع خطی ویژگی های ورودی هستند، که در واقع هر یک بیانگر آزمایش غیر خطی از اندازه گیری های پاولی هستند. پس از آنکه مدل به طور کامل آموزش یافت، می توان تابع های مورد نظر را با توجه به بخش اول شبکه به دست آورد. همچنین با داشتن مقادیر این اندازه گیری ها، می توان از بخش کدگشای شبکه استفاده کرد و نتیجه ی درهم تنیده بودن یا جداپذیر بودن حالت را به دست آورد.

در این حالت شبکه موفق شد به صحت ۹۸٪ در تفکیک حالت ها دست پیدا کند. برای اطلاعات دقیق تر، ماتریس درهم ریختگی در جدول ۲.۳ آورده شده است.

۷.۳ شبکه عصبی با تعداد ویژگی کمتر

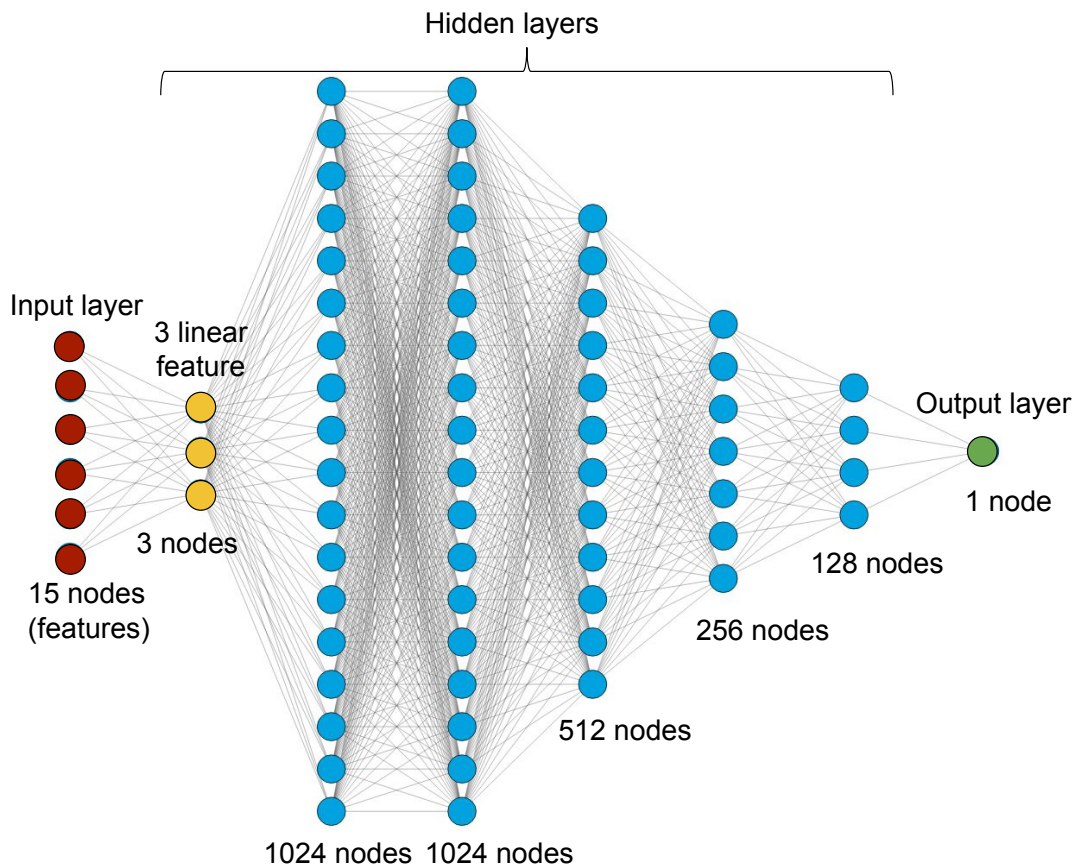
در بخش قبل توانستیم با دو اندازه گیری غیر خطی، حالت های درهم تنیده از حالت های جداپذیر را تشخیص دهیم. اما با توجه به اینکه اندازه گیری های غیر خطی در آزمایشگاه چالش برانگیز هستند، در این بخش معماری دیگری را ارائه می دهیم که بتواند با تعداد کمتری اندازه گیری خطی بهینه این کار را تا حد قابل قبولی انجام



شکل ۶.۳: معماری شبکه ی خود رمزگذار ارایه شده برای پیدا کردن دو اندازه گیری غیر خطی.

	separable pred.	entangled pred.
separable true	۹۹/۷۲٪ (۳۱۴۸۶۵)	۰/۲۸٪ (۸۷۶)
entangled true	۰/۱۶٪ (۱۱۰۵)	۹۹/۸۴٪ (۶۸۳۱۵۴)

جدول ۲.۳: ماتریس درهم ریختگی شبکه ی پیشخور



شکل ۷.۳: معماری شبکه ی عصبی برای تشخیص تعداد کمتر اندازه گیری خطی

دهد.

معماری شبکه عصبی ارایه شده در شکل ۷.۳ آورده شده است. همانطور که مشاهده می شود، در لایه اول، به هر تعداد آزمایش خطی که مورد نیاز است، گره می گذاریم. در این مثال سه گره آورده شده است. نکته قابل توجه دیگر این است که تابع فعال سازی لایه ی اول تابع همانی در نظر گرفته شده است، و لایه های دوم به بعد فقط از تابع غیر خطی ReLU استفاده می کنند. لذا مقادیر گره های لایه ی اول هر کدام یک ترکیب خطی از مقادیر ورودی هستند. یعنی معماری شبکه، ماشین را مجبور می کند ترکیب خطی بهینه ای برای تشخیص حالت های درهم تنیده پیدا کند. با توجه به اینکه مقادیر ورودی، خود مقدار چشم داشتی اندازه گیری های ضرب ماتریس های پاولی می باشند، پس مقدار گره های لایه ی اول نیز هر کدام معادل یک اندازه گیری ترکیب خطی از ضرب ماتریس های پاولی هستند. با به دست آمدن ضرایب یال های لایه ی اول می توان این ترکیب خطی، و در نتیجه عملگر مورد نیاز برای تشخیص در هم تنیدگی را به دست آید.

این کار برای تعداد اندازه گیری های خطی متفاوت تکرار شد و نتایج در شکل ۸.۳ نمایش داده شده است. همانگونه که انتظار می رود، با افزایش تعداد اندازه گیری ها، صحت مدل افزایش می یابد تا در نهایت

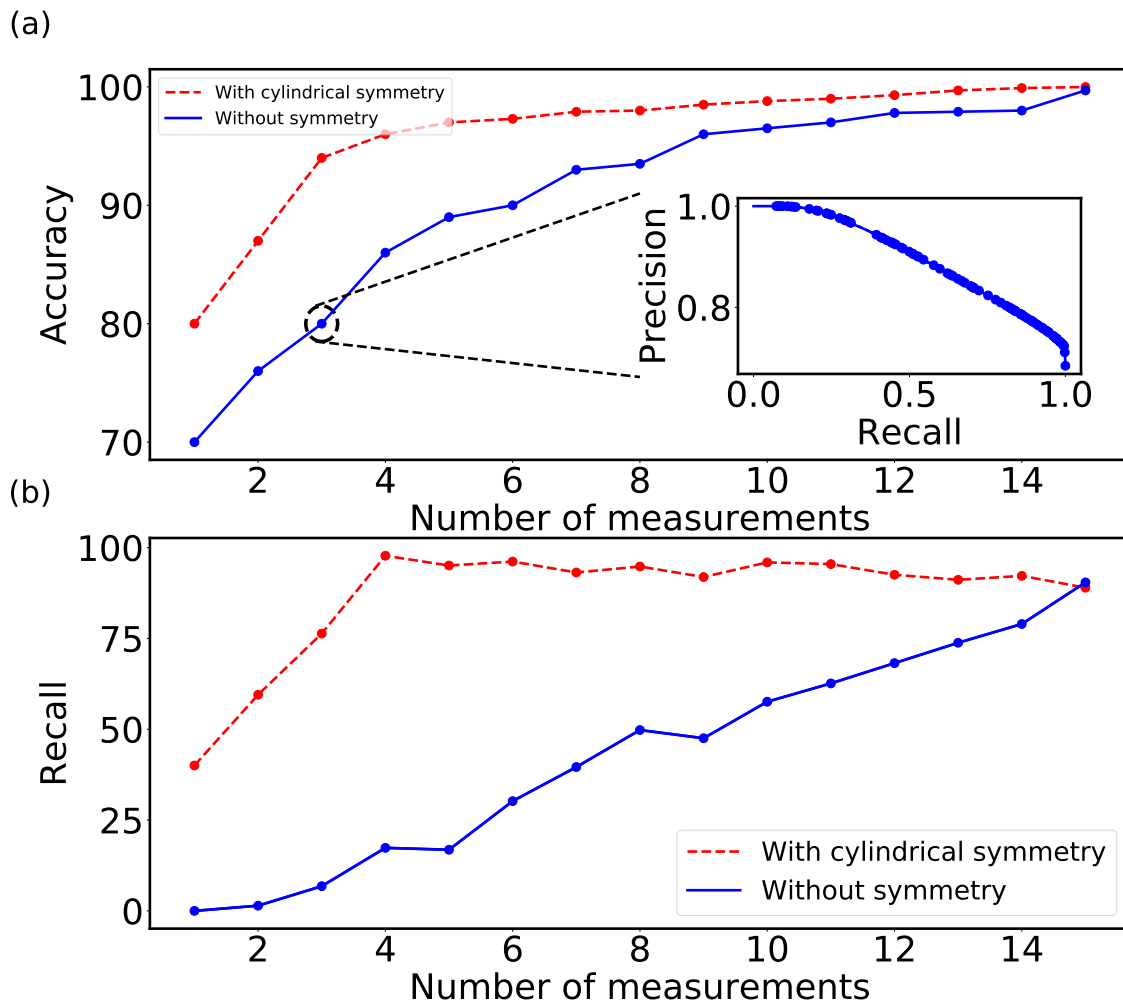
با داشتن ۱۵ اندازه گیری، می تواند به طور تقریباً کامل حالت ها را تفکیک کند. به طور مثال با داشتن فقط ۳ اندازه گیری خطی، مدل می تواند صحت تقریباً ۸۰٪ را به دست آورد. توجه کنید که با وجود مصالحه بین دقت و بازیابی، می توان آستانه، را به گونه ای تغییر داد که دقت تا مقدار دلخواه بالا برود. مثلاً در صورت نیاز دقت ۹۵٪، بازیابی مدل ۶۰٪ می شود. این بدین معنی است که اگر بخواهیم مدل هر حالتی را که درهم تنیده تشخیص دهد، به احتمال ۹۵ درصد واقعاً درهم تنیده باشد، مدل می تواند ۶۰٪ حالت های درهم تنیده را تشخیص دهد. یا به عنوان مثال، اگر بخواهیم با قطعیت تقریباً کامل حالت های درهم تنیده را تشخیص دهیم، می توانیم آستانه را طوری تنظیم کنیم که دقت بیشتر از ۹۹/۹۹٪ باشد. در این صورت، هر حالت تشخیص داده شده توسط ماشین به عنوان حالت درهم تنیده، با قطعیت تقریباً کامل، واقعاً درهم تنیده است. پس به نوعی ماشین به عنوان یک شاهد درهم تنیدگی عمل میکند. با تنظیم چنین آستانه ای، ماشین به بازیابی ۲۱٪ می رسد، که یعنی می تواند با قطعیت تقریباً کامل ۲۱٪ حالت های درهم تنیده را با فقط سه اندازه گیری خطی تشخیص دهد.

نتیجه ی جالب دیگری که در این بخش به دست آمد، بررسی ضرایب لایه ی اول شبکه عصبی است. این ضرایب در واقع همان ضرایب ترکیب خطی عملگر مورد نظر از ضرب ماتریس های پاولی است. از طرف دیگر، اندازه ی ضریب هر یک از این عملگر ها، به معنی اهمیت بیشتر آن عملگر در تشخیص حالت های درهم تنیده است. نتیجه ی به دست آمده، با نتایج به دست آمده از بخش اهمیت ویژگی های روش های کلاسیکی کاملاً تطابق دارد. همانطور که مشاهده می شود، مهم ترین ویژگی ها در این بخش نیز ضرب تانسوری دو ماتریس پاولی یکسان بودند، یعنی $\sigma_x \otimes \sigma_x, \sigma_y \otimes \sigma_y, \sigma_z \otimes \sigma_z$.

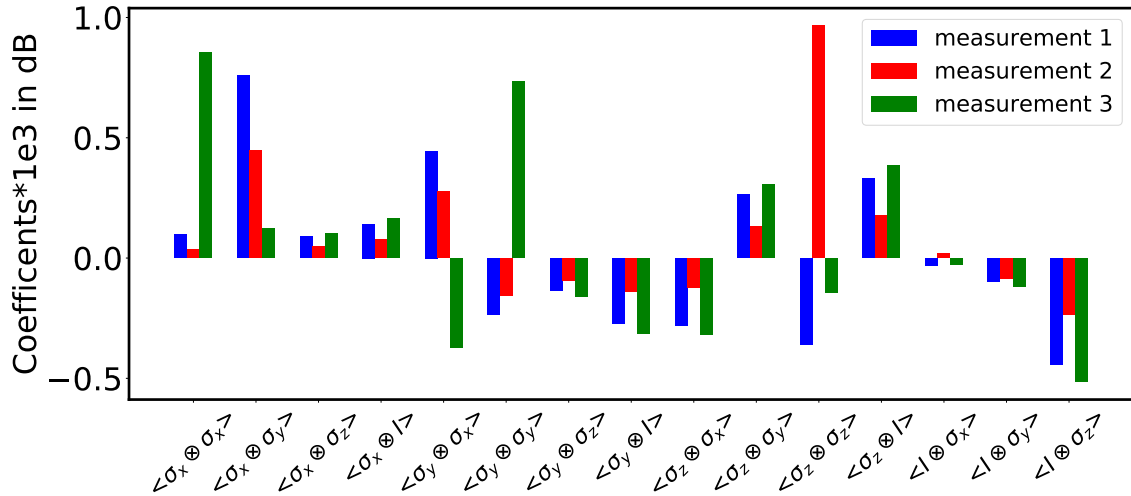
۸.۳ تقارن استوانه ای

حالت های کوانتمی که در آزمایشگاه استفاده می شوند، معمولاً تقارن های خاصی دارند که معمولاً از تقارن های دستگاه تولید کننده ی آن ها ناشی می شود. این تقارن ها، شرایطی را روی ماتریس های چکالی الزام می کنند که باعث می شود گستره ی آن ها خاص تر شود. در نتیجه ممکن است اطلاعات تقارن حالت ها، به تشخیص درهم تنیدگی کمک کند، چرا که تقارن ها، اطلاعات اضافه ای در مورد ویژگی ها ارایه می دهند. در این قسمت ما تقارن استوانه ای را مورد بررسی قرار داده ایم. در این جا محور تقارن استوانه ای z در نظر گرفته شده است.

$$(R_z(\theta) \otimes R_z(\theta)) \rho (R_z(\theta)^\dagger R_z(\theta)^\dagger) = \rho \quad (1.3)$$



شکل ۸.۳: نتایج ارزیابی مدل. صحت مدل شبکه ی عصبی با تعداد اندازه گیری خطی کمتر، برای حالت بدون تقارن و حالت با تقارن استوانه ای در شکل بر حسب تعداد اندازه گیری نشان داده شده است.



شکل ۹.۳: ضرایب لایه ی اول برای هر سه گره. این ضرایب در واقع همان ضرایب خطی ماتریس های پاولی اندازه گیری بهینه پیدا شده هستند. همچنین، بزرگی اندازه این ضرایب، نشانگر اهمیت هر یک از ویژگی ها در تعیین نتیجه است.

برای یک ماتریس چگالی با تقارن استوانه ای، معادله ی ۱.۳ برقرار است. در نتیجه باید شرط زیر برای همه ی زوایای θ برقرار باشد:

$$\forall \theta \quad [R_z(\theta) \otimes R_z(\theta), \rho] = 0 \quad (2.3)$$

که در آن عملگر دوران به این صورت است:

$$R_z(\theta) = e^{i\theta\sigma_z} \implies (\forall \theta \quad [e^{i\theta\sigma_z} \otimes e^{i\theta\sigma_z}, \rho] = 0) \quad (3.3)$$

با ساده سازی معادله ی ۳.۳ به رابطه ی زیر می رسیم:

$$[\sigma_z \otimes \mathbb{I} + \mathbb{I} \otimes \sigma_z, \rho] = 0 \quad (4.3)$$

با جایگذاری $\rho = \frac{1}{4} \sum a_{ij} \sigma_i \otimes \sigma_j$ نتیجه می گیریم:

$$\sum a_{ij} [\sigma_z, \sigma_i] \otimes \sigma_j + \sum a_{ij} \sigma_i \otimes [\sigma_z, \sigma_j] = 0 \quad (5.3)$$

با توجه به رابطه ی $[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k$ که بین ماتریس های پاولی برقرار است داریم:

$$2i \sum_j a_{xj} \sigma_y \otimes \sigma_j - a_{yj} \sigma_x \otimes \sigma_j + a_{jx} \sigma_j \otimes \sigma_y - a_{jy} \sigma_j \otimes \sigma_x = 0 \quad (6.3)$$

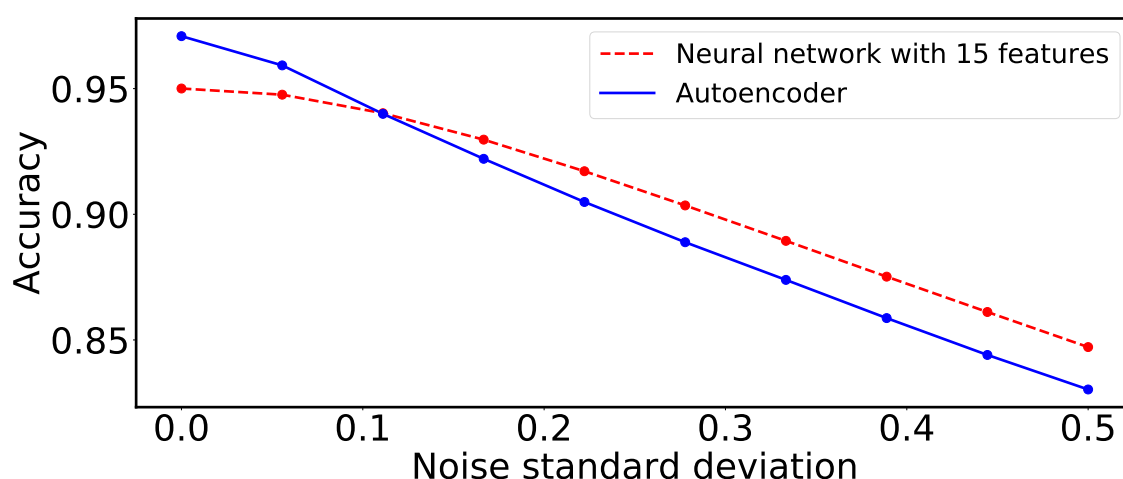
ماتریس های چگالی را با الگوریتمی که در کد آمده است، به صورتی پیدا می کنیم که تمام شرایط گفته شده در بالا را برآورده کند. سپس تمامی مدل های ارایه شده در قسمت های قبل را مجدداً با این داده ها آموزش می دهیم. در این حالت نتایج بهتری از قبل به دست می آوریم. نتایج حاصل شده با وجود تقارن استوانه ای، در شکل ۸.۳ آورده شده است. همانگونه که مشاهده می شود، در این حالت، با ۴ اندازه گیری خطی، می توان تقریباً به تفکیک کامل دست یافت. همچنین فقط با یک اندازه گیری خطی به صحت ۹۰٪ دست یافتیم.

۹.۳ پایداری و اثر نوفه

نتایج اندازه گیری های موجود در آزمایشگاه، همواره همراه با نوفه مزاحم هستند که از محیط اطراف یا حتی خود دستگاه اندازه گیری ناشی می شوند. بنابراین باید مطمئن شویم نتایج به دست آمده در بخش قبل، تا حد قابل قبولی در مقابل نوفه مقاوم هستند. چراکه اگر با وجود اندکی نوفه، و تغییر اندک پارامترهای ورودی، صحت شاهد ارایه شده توسط شبکه عصبی به طرز قابل توجهی افت کند، این مدل ها در عمل خیلی کاربردی نیستند.

در این قسمت، ما نوفه جمع شونده با توزیع گاوسی، که از نوفه های معمول آزمایشگاه هست را بررسی کرده ایم. نوفه به صورت یک عدد تصادفی با توزیع گاوسی تولید شد، و با ویژگی های موجود جمع می شود. میانگین این توزیع را صفر در نظر گرفته، و همچنین این کار را با انحراف از معیارهای متفاوت برای این توزیع تکرار کرده ایم. سپس در هر مرحله صحت شبکه عصبی را مجدداً آزمایش کرده، و نتایج را در شکل ۱۰.۳ نشان داده ایم.

همانطوری که مشاهده می شود، صحت شبکه عصبی به طور آرامی با افزایش نوفه افت می کند، و افت ناگهانی دیده نمی شود. به طور مثال با اضافه کردن نوفه با انحراف معیار ۰/۱، صحت فقط حدود ۴ درصد افت میکند و همچنان بالای ۹۰٪ باقی می ماند (داده های مساله همگی بین ۱- و ۱ هستند).



شکل ۱۰.۳: نمودار صحت شبکه عصبی پیش خور نسبت به انحراف معیار نوفه گاوسی جمع شونده

۱۰.۳ یادگیری نظارت نشده

در تمامی روش های قبل، ما از یادگیری نظارت شده برای آموزش مدل استفاده کردیم. یعنی برچسب داده ها را نیز داشتیم، و هدف بیشتر انجام پیش بینی با وجود تعداد کمتر آزمایش بود. در این قسمت ما از روش های یادگیری نظارت نشده استفاده می کنیم. که نیازی به دانستن برچسب داده ها ندارند. هرچند انتظار داریم که دقت این روش ها، از روش های نظارت شده به مقدار قابل توجهی کمتر باشد، اما این روش ها برای تعمیم به ابعاد بالاتر که نمی توان به راحتی با معیاری مانند PPT برچسب گذاری کرد، ارزشمند هستند. البته ما در این پایان نامه، ابعاد بالاتر را بررسی نکرده ایم، و این قسمت فقط به عنوان اثبات ایده^۱ ارایه شده است.

ابتدا ما روش های کلاسیکی یادگیری نظارت نشده را بررسی کرده ایم. در این روش ها، فقط ۱۵ ستون اول داده که شامل ویژگی ها هستند استفاده شده، و هیچ استفاده ای از برچسب ها نشده است. همچنین در این روش ها، از توصیف کامل (هر ۱۵ ویژگی) ماتریس چگالی استفاده شده است، چرا که هدف ما در این بخش، پیش بینی با توصیف ناکامل نیست. نتایج این روش ها در جدول ۱۰.۳ آورده شده است. همانطور که مشاهده می شود، روش OPTICS بیشترین صحت را بین این مدل ها به دست آورد.

Accuracy	Unsupervised
۷۳%	K-means
۷۴%	DBSCAN
۷۷%	OPTICS
۶۸%	Birch

¹Proof of Concept

۱۱.۳ یادگیری شبه نظارت شده

در روش یادگیری شبه نظارت شده، برچسب همه ی داده ها مشخص نیست، بلکه فقط برخی از این داده ها برچسب دارند. تفاوت این روش با روش یادگیری نظارت نشده، این است که این روش ها از برچسب های موجود داده ها استفاده می کنند، و احتمالاً منجر به دقت بالاتری نسبت به روش های نظارت نشده دارند. هرچند نسبت به روش های نظارت شده، همچنان انتظار دقت کمتری را داریم.

این روش ها نیز برای تعمیم به ابعاد بالاتر بسیار مطلوب هستند. به این خاطر که می توان برخی از حالت های ابعاد بالاتر را با استفاده از شاهد ها و یا روش های دیگر برچسب گذاری کرد. مثلاً معیار PPT با توجه به اینکه یک شرط لازم، و یک شاهد در ابعاد بالاتر است، می تواند مقداری از حالت های در هم تنیده را تشخیص دهد. همچنین حالت های خالص را می توانیم با روش تجزیه اشمیت برچسب گذاری کنیم.

علاوه بر این، در این قسمت، مقدار نسبتاً زیادی حالت جداپذیر تولید کرده ایم. روش کار به این صورت است که ابتدا دو ماتریس چگالی تصادفی تولید می کنیم و آن ها را در هم ضرب تانسوری می کنیم. با این کار یک سری داده تولید می کنیم. سپس یک آرایه ی n عضوی در نظر می گیریم (ما در اینجا $n = 10$ انتخاب کردیم)، و ده عدد تصادفی مثبت تولید کرده و آن ها را به جمع آن ده عدد تقسیم می کنیم. به این صورت ده عدد متفاوت که جمع آن ها واحد می شود به دست آورده ایم. سپس با استفاده از این ده عدد، و ده ماتریس تولید شده از قسمت قبل، یک ترکیب خطی از این ماتریس ها میسازیم، که چون ترکیب خطی از ماتریس های جداپذیر هستند، این ها نیز جداپذیرند. سپس با استفاده از این ماتریس های چگالی، می توانیم تمام پانزده ویژگی را محاسبه کنیم و به همه ی این داده ها برچسب جداپذیر بودن می زنیم و آن ها را ذخیره می کنیم.

در این بخش ما از تعدادی از روش های شبه نظارت شده استفاده کردیم، و نتایج این مدل ها در جدول ۱۱.۳ آورده ایم.

یک روش یادگیری شبه نظارت شده دیگر روش تشخیص ناهنجاری است که با روش های شبه نظارت شده قبلی نسبتاً متفاوت است. از این نظر که در این روش، با داشتن برچسب فقط یک نوع داده این کار صورت می گیرد. یعنی داده های مورد نیاز آموزش این مدل، همه فقط حالت های جداپذیر، و یا همه فقط حالت های درهم تنیده هستند. سپس اگر یک حالت جدید را بخواهیم پیش بینی کنیم، این روش با مقایسه با حالت های آموزش دیده شده، تصمیم میگیرد که آیا این مشابه داده های قبلی (و در نتیجه هم برچسب با آن) است یا اینکه یک ناهنجاری (و در نتیجه دارای برچسب متفاوت است).

در این جا ما از داده های جداپذیر که در قسمت قبل روش تولید آن ها رو توضیح دادیم، استفاده کرده و مدل را آموزش داده ایم. سپس مدل را با زیر مجموعه ای تصادفی از داده های اصلی که برچسب آن ها را نیز

می دانیم، آزمایش کردیم. نتیجه به دست آمده در جدول ۱۱.۳ قابل مشاهده است.

Accuracy	Semisupervised
۸۱%	Anomaly Detection
۶۴%	LabelPropagation
۶۷%	LabelSpreading

۱۲.۳ نتیجه گیری

در این فصل ابتدا به نحوه ی تولید داده ها و بررسی آماری آن ها پرداختیم. سپس انواع مدل های یادگیری ماشینی کلاسیکی و شبکه عصبی روی داده ها با داشتن توصیف کامل سیستم بررسی شد و دیدیم که چندی از این مدل ها می توانند به تفکیک کامل دست یابند. سپس دو معماری جدید مبتنی بر شبکه های خود رمزگذار برای تشخیص درهم تنیدگی با تعداد آزمایش کمتر خطی و غیر خطی ارائه کردیم. همچنین اثر نوفه بر این سیستم ها و اهمیت ویژگی ها برای تشخیص درهم تنیدگی را بررسی کردیم. در آخر به بررسی چند مدل یادگیری نظارت نشده و شبه نظارت شده پرداختیم. در فصل بعد، به جمع بندی نتایج و پیشنهادات برای ادامه ی کار می پردازیم.

فصل ۴

جمع بندی و پیشنهادات

۱.۴ جمع بندی

در این پایان نامه، قصد داشتیم تا حالت های درهم تنیده و جداپذیر در یک سیستم دو کیوبیتی را با استفاده از ابزار های یادگیری ماشینی از یکدیگر تفکیک کنیم. در فصل اول، به بیان مفاهیم مورد نیاز حوزه اطلاعات کوانتومی، و تبیین مساله به زبان ریاضی پرداختیم. در فصل دوم مفاهیم هوش مصنوعی و یادگیری ماشینی مورد استفاده در این پایان نامه به اختصار توضیح داده شد. سپس در فصل سوم، روش تولید داده ها، بررسی آماری آن ها، و مدل های یادگیری ماشینی ارایه شد و نتایج آن ها ارزیابی و مقایسه شد.

در ابتدا، سعی کردیم بتوانیم با داشتن توصیف کامل سیستم (نتایج ۱۵ اندازه گیری روی سیستم دو کیوبیت)، ماشین را آموزش دهیم تا بتواند حالت های جداپذیر و درهم تنیده را تفکیک کند و به معیار PPT یا معیاری هم ارز آن دست پیدا کند. ابتدا این کار را ما با استفاده از روش های کلاسیکی یادگیری ماشینی انجام دادیم و نتایج آموزش و همچنین زمان آموزش و ارزیابی را در جدولی مقایسه کردیم. مشاهده شد که بعضی روش ها مانند Ridge توانستند به صورت تقریباً کامل (صحت ۹۹/۷٪) حالت ها را از دیگر تفکیک کنند همچنین اهمیت هر یک از این اندازه گیری ها را در این روش های کلاسیکی مورد بررسی قرار دادیم و دیدیم بعضی از این اندازه گیری ها اهمیت بیشتری از بقیه در تعیین نتیجه ی درهم تنیده یا جداپذیر بودن دارند. سپس با روش های شبکه ی عصبی این کار را مجدداً تکرار کردیم و مشاهده شد که با داشتن توصیف کامل سیستم، آن ها نیز قادر هستند به تفکیک نسبتاً کامل (صحت ۹۹/۹٪) برسند.

سپس مدل دیگری بر اساس شبکه های خود رمزگذار ارایه کردیم که موفق شد فقط با دو اندازه گیری غیر خطی به صحت بالای ۹۹٪ برسند. با توجه به اینکه اندازه گیری های غیر خطی در آزمایشگاه چالش برانگیز

و هزینه بر هستند، مدل دیگری ارایه دادیم که بتواند تعداد کمی اندازه گیری خطی بهینه برای تشخیص بخش قابل توجهی از حالت های در هم تنیده پیدا کند. به عنوان مثال توانستیم فقط با ۳ اندازه گیری خطی، به صحت ۸۰٪ برسیم. می توان با تغییر آستانه ماشین، دقت تشخیص حالت های درهم تنیده را تا حد مطلوب بالا برد. مثلاً اگر دقت ۹۹/۹٪ مورد نیاز باشد، ماشین بازیابی ۲۱٪ پیدا می کند. در این صورت ماشین به عنوان یک شاهد درهم تنیدگی عمل می کند.

با توجه به اینکه حالت های تولید شده در آزمایشگاه، معمولاً دارای تقارن های مشخصی هستند، ما مجدداً مراحل قبل را برای حالت های با تقارن استوانه ای تکرار کردیم و نتایج بهتری کسب کردیم. در این حالت ماشین فقط با دو اندازه گیری خطی صحت ۶۰٪ و با ۴ اندازه گیری خطی به صحت ۹۸٪ رسید. همچنین با توجه به اینکه نتایج اندازه گیری های انجام شده در آزمایشگاه همیشه همراه با نوفه هستند، ما مدل های قبلی را با وجود اثر نوفه جمع شونده گاوسی بررسی کردیم و مشاهده شد که مدل ها پایداری مطلوبی در صورت وجود نوفه دارند.

در ادامه فصل ۳، از روش های یادگیری نظارت نشده و شبه نظارت شده هم استفاده کرده و آن مدل ها را ارزیابی کردیم. در این مدل ها، روش OPTICS از روش های نظارت نشده و روش Anomaly Detection از روش شبه نظارت شده، بهترین نتایج را حاصل کردند و به ترتیب صحت ۷۷٪ و ۸۱٪ دارند. هرچند صحت این دسته از مدل ها، از حالت نظارت شده کمتر است، اما قابلیت تعمیم به ابعاد بالاتر، برای این مدل ها راحت تر است.

تمامی مدل های ارایه شده در این پایان نامه، به عنوان اثبات ایده ی استفاده از شبکه های عصبی و شبکه های خود رمزگذار در پیدا کردن اندازه گیری های نیمه بهینه برای تشخیص درهم تنیدگی هستند، و باید در نظر داشت که پژوهش ما محدودیت منابع محاسباتی مشخصی داشت. صحت هر یکی از این مدل ها، می تواند با تنظیم فرآپارامترهای^۱ شبکه و افزایش تعداد داده های آموزش، و لایه های شبکه بهبود یابد. مشخص نیست تا چه میزان می توان این مدل ها را بهبود داد و تا چه حد صحت هر یک را افزایش داد.

۲.۴ پیشنهادات

در ادامه، پیشنهاداتی برای ادامه ی کار مطرح می شود.

¹Hyper-parameters

۱.۲.۴ بررسی تقارن های دیگر

یکی از پیشنهاد های موجود برای ادامه ی کار، این است که تقارن های دیگر مانند تقارن کروی یا تقارن های آینه ای نیز بررسی شود. برای این کار لازم است تا شروط لازم این تقارن ها به صورت معادله هایی برای ویژگی های ماتریس چگالی به دست آید، و سپس مشابه الگوریتم ارایه شده در کد، ابتدا عملگر مورد نیاز برای متقارن سازی حالت ها پیدا شده، و روی حالت های تولید شده اثر داده شود که حالت هایی با تقارن مورد نظر ساخته شود. سپس می توان مدل های ارایه شده در این پایان نامه را مجدداً روی آن تقارن ها نیز بررسی کرد.

۲.۲.۴ ابعاد بالاتر

در این پایان نامه، به سیستم دو کیوبیت پرداخته شد، اما اکثر روش های ارایه شده در این پایان نامه، قابل تعمیم به ابعاد بالاتر هستند. برای این کار لازم است که شاهد های متفاوتی بر داده های تولید شده اعمال شوند تا هر یک بخشی از این حالت ها را به صورت درهم تنیده تشخیص دهند. حالت های خالص را نیز می توان با تجزیه ی اشمیت برچسب گذاری کرد. همچنین با روش ارایه شده در فصل سوم، می توان ماتریس های چگالی جداپذیر تولید کرد. یک چالش دیگر ابعاد بالاتر، مشقت بعدچندی^۱ است. با افزایش ابعاد زیر فضا ها، بعد سیستم مرکب دویخشی به صورت $O(N^2)$ سریعاً رشد می کند و داده های موجود برای یادگیری در فضا بسیار تنگ و پراکنده می شوند. در نتیجه به تعداد بسیار بیشتری داده برای یادگیری نیاز است. علاوه بر این، تعداد ویژگی های هر ماتریس چگالی نیز به به شدت افزایش می یابد و این حجم زیاد از داده های با تعداد ویژگی بالا، برای یادگی به منابع محاسباتی زیادی احتیاج دارد.

۳.۲.۴ ماشین بولتزمن و مدل های تولید کننده

یکی دیگر از روش های بسیار معروف یادگیری نظارت نشده، ماشین بولتزمن است. روش های نظارت نشده بررسی شده در این پایان نامه، نتوانستند صحت بسیار زیادی در تشخیص درهم تنیگی به دست آورند، و ممکن است بتوان با ماشین بولتزمن نتایج مطلوب تری به دست آورد. علاوه بر این، می توان از مدل های تولید کننده^۲ استفاده کرد و حالت های درهم تنیده و جداپذیر برای یادگیری با روش های نظارت شده و شبه نظارت شده با این روش ها تولید کرد.

^۱Curse of Dimensionality

^۲Generative Models

۴.۲.۴ روش های دیگر یادگیری ماشینی

معماری های بسیار زیادی برای شبکه های عصبی در یادگیری مسائل مختلف ارائه شده است، مانند معماری Inception یا معماری Dueling، روش های Actor-Critic مانند A2C و A3C و یا روش های جدید دیگر که ممکن است بتوانند نتایج قابل قبولی در حل مساله ی جداپذیری ارائه دهند.

- [1] M. Kulin, T. Kazaz, E. De Poorter, and I. Moerman, “A survey on machine learning-based performance improvement of wireless networks: Phy, mac and network layer,” *Electronics*, vol.10, no.3, 2021.
- [2] M. Taddeo and L. Floridi, “How ai can be a force for good,” *Science*, vol.361, no.6404, pp.751–752, 2018.
- [3] Y.-Y. Chen, Y.-H. Lin, C.-C. Kung, M.-H. Chung, and I.-H. Yen, “Design and implementation of cloud analytics-assisted smart power meters considering advanced artificial intelligence as edge analytics in demand-side management for smart homes,” *Sensors*, vol.19, p.2047, May 2019.
- [4] M. Horodecki, P. Horodecki, and R. Horodecki, “Separability of mixed states: necessary and sufficient conditions,” *Physics Letters A*, vol.223, no.1, pp.1–8, 1996.
- [5] S. Gharibian, “Strong np-hardness of the quantum separability problem,” *Quantum Info. Comput.*, vol.10, p.343–360, Mar. 2010.
- [6] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. USA: Cambridge University Press, 10th ed. , 2011.
- [7] B. M. Terhal, “Detecting quantum entanglement,” *Theoretical Computer Science*, vol.287, no.1, pp.313–335, 2002. Natural Computing.
- [8] S. M. Cohen, “Understanding entanglement as resource: Locally distinguishing unextendible product bases,” *Phys. Rev. A*, vol.77, p.012304, Jan 2008.
- [9] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, “Quantifying entanglement,” *Phys. Rev. Lett.*, vol.78, pp.2275–2279, Mar 1997.
- [10] J.-L. Li and C.-F. Qiao, “A necessary and sufficient criterion for the separability of quantum state,” *Scientific Reports*, vol.8, Jan. 2018.
- [11] A. Ekert and P. L. Knight, “Entangled quantum systems and the schmidt decomposition,” *American Journal of Physics*, vol.63, pp.415–423, May 1995.
- [12] S. Sciara, R. L. Franco, and G. Compagno, “Universality of schmidt decomposition and particle identity,” *Scientific Reports*, vol.7, Mar. 2017.

- [13] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, “Quantum entanglement,” *Rev. Mod. Phys.*, vol.81, pp.865–942, Jun 2009.
- [14] R. Augusiak, M. Demianowicz, and P. Horodecki, “Universal observable detecting all two-qubit entanglement and determinant-based separability tests,” *Phys. Rev. A*, vol.77, p.030301, Mar 2008.
- [15] D. Lu, T. Xin, N. Yu, Z. Ji, J. Chen, G. Long, J. Baugh, X. Peng, B. Zeng, and R. Laffamme, “Tomography is necessary for universal entanglement detection with single-copy observables,” *Physical Review Letters*, vol.116, June 2016.
- [16] O. Gühne, P. Hyllus, O. Gittsovich, and J. Eisert, “Covariance matrices and the separability problem,” *Physical Review Letters*, vol.99, Sep 2007.
- [17] W. K. Wootters, “Entanglement of formation of an arbitrary state of two qubits,” *Phys. Rev. Lett.*, vol.80, pp.2245–2248, Mar 1998.
- [18] P. Rungta, V. Bužek, C. M. Caves, M. Hillery, and G. J. Milburn, “Universal state inversion and concurrence in arbitrary dimensions,” *Phys. Rev. A*, vol.64, p.042315, Sep 2001.
- [19] J. I. de Vicente, “Lower bounds on concurrence and separability conditions,” *Phys. Rev. A*, vol.75, p.052320, May 2007.
- [20] B. M. Terhal, “Bell inequalities and the separability criterion,” *Physics Letters A*, vol.271, no.5, pp.319–326, 2000.
- [21] L. Pezzè, Y. Li, W. Li, and A. Smerzi, “Witnessing entanglement without entanglement witness operators,” *Proceedings of the National Academy of Sciences*, vol.113, pp.11459–11464, Sept. 2016.
- [22] Y. Liu, X. Zhang, M. Lewenstein, and S.-J. Ran, “Entanglement-guided architectures of machine learning by quantum tensor network,” *arXiv e-prints*, p.arXiv:1803.09111, Mar. 2018.
- [23] J. Gao, L.-F. Qiao, Z.-Q. Jiao, Y.-C. Ma, C.-Q. Hu, R.-J. Ren, A.-L. Yang, H. Tang, M.-H. Yung, and X.-M. Jin, “Experimental machine learning of quantum states,” *Phys. Rev. Lett.*, vol.120, p.240501, Jun 2018.
- [24] Y.-C. Ma and M.-H. Yung, “Transforming bell’s inequalities into state classifiers with machine learning,” *npj Quantum Information*, vol.4, July 2018.
- [25] S. Lu, S. Huang, K. Li, J. Li, J. Chen, D. Lu, Z. Ji, Y. Shen, D. Zhou, and B. Zeng, “Separability-entanglement classifier via machine learning,” *Phys. Rev. A*, vol.98, p.012315, Jul 2018.
- [26] J. Gray, L. Banchi, A. Bayat, and S. Bose, “Machine-learning-assisted many-body entanglement measurement,” *Phys. Rev. Lett.*, vol.121, p.150503, Oct 2018.

- [27] D.-L. Deng, “Machine learning detection of bell nonlocality in quantum many-body systems,” *Phys. Rev. Lett.*, vol.120, p.240402, Jun 2018.
- [28] Y. Levine, D. Yakira, N. Cohen, and A. Shashua, “Deep learning and quantum entanglement: Fundamental connections with implications to network design,” 2017.
- [29] Y. Liu, X. Zhang, M. Lewenstein, and S.-J. Ran, “Entanglement-guided architectures of machine learning by quantum tensor network,” 2018.
- [30] P.-H. Qiu, X.-G. Chen, and Y.-W. Shi, “Detecting entanglement with deep quantum neural networks,” *IEEE Access*, vol.7, pp.94310–94320, 2019.
- [31] X.-D. Cai, D. Wu, Z.-E. Su, M.-C. Chen, X.-L. Wang, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, “Entanglement-based machine learning on a quantum computer,” *Phys. Rev. Lett.*, vol.114, p.110504, Mar 2015.
- [32] E. Behrman, V. Chandrashekar, Z. Wang, C. Belur, J. Steck, and S. Skinner, “A quantum neural network computes entanglement,” 03 2002.
- [33] L. Wang, ed. . *Support Vector Machines: Theory and Applications*. Springer Berlin Heidelberg, 2005.
- [34] R. F. de Mello and M. A. Ponti. *Machine Learning*. Springer International Publishing, 2018.
- [35] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol.521, pp.436–444, May 2015.
- [36] P. Baldi, “Autoencoders, unsupervised learning, and deep architectures,” in *Proceedings of ICML Workshop on Unsupervised and Transfer Learning* (I. Guyon, G. Dror, V. Lemaire, G. Taylor, and D. Silver, eds.), vol.27 of *Proceedings of Machine Learning Research*, (Bellevue, Washington, USA), pp.37–49, JMLR Workshop and Conference Proceedings, 02 Jul 2012.
- [37] V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection: A survey,” *ACM Comput. Surv.*, vol.41, July 2009.
- [38] V. Hodge and J. Austin, “A survey of outlier detection methodologies,” *Artificial Intelligence Review*, vol.22, pp.85–126, Oct. 2004.