

ذهن مشابه سازه و ماده

۱- مقدمه

هوش مصنوعی (AI) از ابتدای تأسیس خود، هدفمندی پیوسته به سمت تقلید از قابلیت‌های انسانی را داشته است. تلاش‌ها برای شبیه‌سازی و تداعی از هوش انسانی در سیستم‌های مصنوعی، همواره یکی از اهداف اصلی پژوهش‌های مرتبط با هوش مصنوعی بوده است. همچنین، تلاش‌های موثرتر و کارآمدتر جهت تطابق با میزان بالاتر هوش انسانی، به‌ویژه در زمینه‌هایی که با تعامل انسانی ارتباط دارند، اهمیت چشمگیری دارد.

در این مقاله، به بررسی موضوع مشابهت هوش مصنوعی با ابعاد فیزیکی سازه و ماده پرداخته و رویکردهای مختلفی در این زمینه مورد بررسی قرار گرفته‌اند مورد ارزیابی و تجزیه و تحلیل قرار می‌گیرند. همچنین، اهمیت این مفهوم در مسیر حرکت به سوی هویت انسان‌نمای هوش مصنوعی و تأثیرات آن در تحقق هدف نهایی شبیه‌سازی هوش انسانی را مورد بررسی و بحث قرار خواهیم داد. در پایان، با نگاهی به آینده، پتانسیل‌ها و چالش‌های ادامه تحقیقات در این زمینه را بررسی می‌کنیم.

۲- نگاهی به گذشته و ابعاد هوش مصنوعی

هوش مصنوعی یکی از حوزه‌های پیشرفته‌ای از علوم کامپیوتر است که به ایجاد سیستم‌ها و برنامه‌های کامپیوتری توانایی تفکر و یادگیری مشابه انسان را می‌بخشد.

۱-۱-۲ شبکه‌های عصبی مصنوعی (ANNs)

این روش‌ها تا حد زیادی بر مبنای ساختار مغز انسان ساخته شده‌اند و به طور گسترده در تشخیص الگو، تصویربرداری، پردازش زبان طبیعی و دسته‌بندی داده‌ها مورد استفاده قرار می‌گیرند.

۲-۱-۲ یادگیری ماشینی (Machine Learning)

این روش‌ها بر اساس الگوریتم‌هایی که به کامپیوتر اجازه می‌دهند از داده‌ها یاد بگیرد، کار می‌کنند. الگوریتم‌هایی مانند درخت تصمیم، ماشین بردار پشتیبان (SVM) و کاهش بعد از این دسته هستند.

۳-۱-۲ یادگیری عمیق (Deep Learning)

این روش‌ها از شبکه‌های عصبی ژرف با تعداد لایه‌های زیاد برای حل مسائل پیچیده مانند تشخیص تصویر و ترجمه ماشینی استفاده می‌کنند.

۲-۲ شبکه‌های عصبی مصنوعی (ANNs)

شبکه‌های عصبی مصنوعی (Artificial Neural Networks) یکی از اصولی‌ترین و پرکاربردترین روش‌های هوش مصنوعی هستند که تا حد زیادی تشابه با ساختار نورون‌های مغز انسان دارند. این شبکه‌ها از چندین لایه از واحدهای پردازشی که به نورون‌ها یا نودها معروف‌اند، تشکیل شده‌اند و توسط وزن‌هایی که مشخص می‌کنند چگونه ورودی‌ها تبدیل به خروجی‌ها می‌شوند، کنترل می‌شوند. به طور کلی، هدف از آموزش شبکه‌های عصبی، یادگیری الگوها و ارتباطات پیچیده در داده‌ها است.

شبکه‌های عصبی مصنوعی می‌توانند به صورت گسترده در مسائل مختلفی مانند تشخیص الگو، تصویربرداری، ترجمه ماشینی، پردازش زبان طبیعی، پیش‌بینی و غیره مورد استفاده قرار گیرند. آن‌ها از توانایی یادگیری از داده‌های آموزشی برای بهبود عملکرد در وظایف پیچیده بهره می‌برند.

شبکه‌های عصبی مصنوعی (ANNs) یک مدل ریاضی است که بر اساس ساختار نورون‌ها و اتصالات میان آن‌ها، برای حل مسائل پیچیده استفاده می‌شود. این شبکه‌ها از لایه‌های مختلفی از نورون‌ها تشکیل شده‌اند که اطلاعات از ورودی به خروجی منتقل می‌کنند. هر نورون در لایه متصل به نورون‌های لایه قبلی و لایه بعدی است و وزن‌هایی که اتصالات میان آن‌ها را تنظیم می‌کنند، نقش مهمی در عملکرد شبکه دارند.

یک معماری ساده از یک شبکه‌ی عصبی به نام شبکه‌های عصبی پرسپترون (Perceptrons) می‌تواند به صورت زیر توصیف شود:

۱- ورودی (Input Layer)

این لایه وظیفه دریافت داده‌های ورودی (مانند ویژگی‌ها یا داده‌های تصویری) را دارد. هر نود در این لایه به یک ویژگی ورودی اختصاص دارد.

۲- لایه‌های مخفی (Hidden Layers)

این لایه‌ها برای پردازش و استخراج ویژگی‌های پیچیده‌تر از داده‌های ورودی استفاده می‌شوند. هر نورون در لایه‌های مخفی با ورودی‌های خود وزن‌هایی دارد و خروجی‌هایی تولید می‌کند که به لایه بعدی انتقال داده می‌شود.

۳- خروجی (Output Layer)

این لایه نتیجه نهایی شبکه را تولید می‌کند. بسته به نوع مسئله، هر نود ممکن است با یک کلاس یا یک ویژگی مرتبط باشد.

هر نورون در شبکه با یک تابع فعال‌سازی (Activation Function) مشخص می‌کند که آیا خروجی آن باید فعال شود یا نه. یکی از توابع فعال‌سازی معمولی، تابع سیگموئید (Sigmoid) است که میانگین خروجی را محدود به بازه ۰ تا ۱ می‌کند. فرمول تابع سیگموئید به صورت زیر است:

$$f(z) = \frac{1}{1 + e^{-z}} \quad (۱)$$

وزن‌ها در شبکه‌های عصبی مصنوعی از اهمیت بالایی برخوردار هستند، زیرا طریقه تنظیم آن‌ها تأثیر زیادی بر عملکرد شبکه دارد. هدف اصلی در آموزش شبکه عصبی، بهینه‌سازی وزن‌ها به نحوی است که خروجی مدل نزدیک‌ترین مقدار به مقدار مورد انتظار باشد.

یکی از الگوریتم‌های معروف برای آموزش شبکه‌های عصبی، پس‌انتشار خطا (Backpropagation) است. این الگوریتم به شبکه اجازه می‌دهد تا خطا بین خروجی مورد انتظار و خروجی پیش‌بینی شده را به عقب منتقل کرده و وزن‌ها را به گونه‌ای تنظیم کند که این خطا کمینه شود.

۱-۲-۲ تغییر وزن‌ها در الگوریتم پس‌انتشار خطا

$$w_{ij} = \eta \cdot \delta_j \cdot x_i \quad (۲)$$

در این فرمول:

w_{ij} نشان‌دهنده تغییر وزن میان نورون i در لایه قبلی و نورون j در لایه فعلی است.

η نرخ یادگیری (learning rate) است که مشخص‌کننده مقداری است که وزن‌ها در هر مرحله تغییر می‌کنند.

δ_j نشان‌دهنده خطای نورون j است که بر اساس اختلاف بین خروجی مورد انتظار و خروجی پیش‌بینی شده محاسبه می‌شود.

الگوریتم پس‌انتشار خطا در چند مرحله اجرا می‌شود:

انتشار پیشرو: در این مرحله، ورودی‌ها از لایه ورودی به طرف جلو منتقل می‌شوند و خروجی‌های مراحل مختلف محاسبه می‌شوند.

محاسبه خطا: خطا بین خروجی مورد انتظار و خروجی پیش‌بینی شده محاسبه می‌شود.

انتشار خطا: خطا به عقب از لایه خروجی به لایه‌های قبلی منتقل می‌شود و خطاهای مربوط به هر نورون در لایه‌های مختلف محاسبه می‌شوند.

تغییر وزن‌ها: با استفاده از فرمول فوق و خطاهای محاسبه شده، وزن‌ها به عقب تغییر داده می‌شوند.

۲-۲-۲ تابع هدف (Objective Function)

تابع هدف (Objective Function) در شبکه‌های عصبی مصنوعی (ANNs) نشان‌دهنده معیاری است که برای ارزیابی کیفیت پیش‌بینی‌ها و تطابق میان خروجی‌های مدل و خروجی‌های مورد انتظار مورد استفاده قرار می‌گیرد. هدف اصلی از تغییر و بهینه‌سازی وزن‌ها در شبکه‌های عصبی، کمینه کردن مقدار تابع هدف است تا مدل بهترین پیش‌بینی‌ها را ارائه دهد.

در مسائل مختلف، توابع هدف متفاوتی مورد استفاده قرار می‌گیرند. به عنوان مثال، در مسائل رگرسیون، تابع هدف معمولاً به معیار میانگین مربعات خطا بین خروجی مدل و خروجی مورد انتظار مرتبط است. در مسائل طبقه‌بندی، تابع هدف می‌تواند به معیارهایی مانند تابع خطا لگاریتمی (Cross-Entropy) مرتبط باشد. به طور کلی، هدف از انتخاب تابع هدف این است که مدل به گونه‌ای بهینه‌سازی شود که پیش‌بینی‌های آن به حداکثر تطابق با داده‌های واقعی برسد.

۲-۲-۳ تابع هدف میانگین مربعات خطا در رگرسیون

اگر فرض کنید y_i ورودی واقعی (خروجی مورد انتظار) برای نمونه i ام و \hat{y}_i خروجی پیش‌بینی شده توسط مدل باشد، تابع هدف میانگین مربعات خطا به صورت زیر تعریف می‌شود:

$$J = \frac{1}{2m} \sum_{i=1}^m (y_i - \hat{y}_i)^2 \quad (3)$$

n تعداد نمونه‌ها است.

y_i خروجی واقعی نمونه i ام.

\hat{y}_i خروجی پیش‌بینی شده توسط مدل برای نمونه i ام.

۲-۲-۴ تابع هدف خطا لگاریتمی (Cross-Entropy) در طبقه‌بندی

اگر y_{ij} نشان‌دهنده برچسب واقعی برای نمونه i ام و کلاس j باشد و \hat{y}_{ij} برچسب پیش‌بینی شده توسط مدل باشد، تابع هدف خطا لگاریتمی به صورت زیر تعریف می‌شود:

$$J = -\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^C y_{ij} \log(\hat{y}_{ij}) \quad (4)$$

در اینجا:

n تعداد نمونه‌ها است.

C تعداد کلاس‌ها.

y_{ij} برچسب واقعی کلاس j برای نمونه i ام.

\hat{y}_{ij} احتمال پیش‌بینی شده توسط مدل برای کلاس j و نمونه i ام.

۵-۲-۲ نورون‌ها

هر نورون در شبکه‌های عصبی مصنوعی با ترکیب خطی از ورودی‌ها و وزن‌ها کار می‌کند و سپس این ترکیب خطی از طریق تابع فعال‌سازی (Activation Function) تبدیل به خروجی نورون می‌شود.

فرمول عمومی برای محاسبه خروجی یک نورون به صورت زیر است:

$$z = \sum_{i=1}^n w_i x_i + b \quad (5)$$

در اینجا:

z خروجی ترکیب خطی و وزن‌دار نورون است.

n تعداد ورودی‌ها یا ویژگی‌ها.

w_i وزن مربوط به ورودی x_i .

x_i ورودی i ام.

b تغییرپذیری (bias) که به ترکیب خطی اضافه می‌شود.

پس از محاسبه ترکیب خطی، این مقدار به تابع فعال‌سازی اعمال می‌شود تا خروجی نورون تولید شود.

۶-۲-۲ تابع فعال‌سازی

تابع فعال‌سازی به نوعی تصمیم می‌گیرد که آیا نورون فعال شود و خروجی غیرصفر تولید کند یا خیر. این تابع معمولاً یک تابع غیرخطی است و وظیفه آن تزریق غیرخطیت و تولنایی یادگیری الگوهای پیچیده در شبکه‌ها است.

یکی از توابع فعال‌سازی معروف، تابع سیگموئید (Sigmoid) است، که به صورت زیر تعریف می‌شود:

$$f(z) = \frac{1}{1 + e^{-z}} \quad (6)$$

این تابع مقدار ورودی خطی را به بازه $[0,1]$ تبدیل می‌کند و به عنوان یک مدل احتمالی‌سازی نیز استفاده می‌شود.

همچنین، توابع فعال‌سازی دیگری نیز وجود دارند که به صورت غیرخطی عمل می‌کنند، مانند تابع ReLU (Rectified Linear Unit) که در شبکه‌های عصبی عمیق به طور گسترده استفاده می‌شود.

$$f(z) = \max(0, z) \quad (7)$$

۳-۲ یادگیری ماشین

یادگیری ماشینی یک رشته از هوش مصنوعی است که به ماهیتی از انجام کارها توسط ماشین‌ها بدون برنامه‌ریزی دقیق انسانی می‌پردازد. در این رویکرد، ماشین‌ها با تحلیل داده‌ها و الگوریتم‌های مختلف، قادر به یادگیری الگوها، ارتباطات و قوانین موجود در داده‌ها ورودی هستند و بهبود یافته‌هایی را در تصمیم‌گیری و پیش‌بینی ارائه می‌دهند.

یادگیری ماشینی به دو دسته اصلی تقسیم می‌شود:

۱- یادگیری نظارت‌شده (Supervised Learning)

در این نوع یادگیری، ماشین با تعدادی داده ورودی و خروجی متناظرشان آموزش داده می‌شود تا بتواند به طور خودکار و بدون دستور کار، برای ورودی‌های جدید خروجی مناسب تولید کند. این نوع یادگیری برای مسائل پیش‌بینی و طبقه‌بندی مورد استفاده قرار می‌گیرد.

۲- یادگیری بدون نظارت (Unsupervised Learning)

در این نوع یادگیری، ماشین بدون تمام داده‌های خروجی آموزش می‌بیند و تلاش می‌کند تا ساختارها، الگوها یا ارتباطات مخفی در داده‌ها را شناسایی کند. این نوع یادگیری به تجزیه و تحلیل داده‌ها و کشف اطلاعات مخفی مورد استفاده قرار می‌گیرد.

به عنوان مثال، در مقاله خود می‌توانید نحوه استفاده از یادگیری ماشینی برای تشخیص اشیاء در تصاویر (که به عنوان مثال یکی از مسائل پرکاربرد در دنیای واقعی است) را توضیح دهید. به طور کلی، ابتدا با داده‌های ورودی و خروجی آموزشی شبکه‌های عصبی مصنوعی را آموزش داده و سپس با داده‌های جدید، مدل را تست و پیش‌بینی‌های آن را ارزیابی می‌کنید.

۴-۲ یادگیری عمیق (Deep Learning)

یادگیری عمیق به استفاده از شبکه‌های عصبی عمیق با تعداد زیادی لایه‌های مخفی (یعنی شبکه‌هایی با ساختار عمیق‌تر از شبکه‌های کلاسیک) برای تبدیل ویژگی‌های سطح پایین به ویژگی‌های سطح بالا و در نهایت به پیش‌بینی‌های دقیق‌تر می‌پردازد. این روش اغلب به تشکیل یک سلسله انتقال ویژگی‌ها از لایه به لایه به نام "یادگیری نشانه (Feature Learning)" معروف است.

یکی از مزایای بزرگ یادگیری عمیق، قدرت آن در استخراج ویژگی‌های پیچیده از داده‌های ورودی است. به عنوان مثال، در تشخیص تصاویر، شبکه‌های عمیق می‌توانند به صورت خودکار ویژگی‌هایی مانند لبه‌ها، تشکیل‌دهنده‌ها و مفاهیم پیچیده‌تر مانند چهره‌ها را از تصاویر استخراج کنند.

۳- کاربردهای هوش مصنوعی

۳-۱ تشخیص الگو و تصویربرداری

هوش مصنوعی در تشخیص الگوها و تصویربرداری در زمینه‌هایی مانند تشخیص چهره‌ها، تشخیص بیماری‌های پزشکی از تصاویر پرتودهی و حتی تشخیص تقلب در مالیات به کار می‌رود.

۳-۲ پردازش زبان طبیعی (NLP)

هوش مصنوعی به کمک روش‌های پردازش زبان طبیعی می‌تواند متن‌ها را تحلیل و درک کند. این کاربرد در ترجمه ماشینی، تحلیل احساسات متن‌ها و پرسش و پاسخ خودکار مورد استفاده قرار می‌گیرد.

۳-۳ خودرانی و خودکارسازی

هوش مصنوعی در توسعه خودروهای خودران، کارخانه‌های خودکار و سیستم‌های خودمختار کاربرد دارد. بهینه‌سازی و تحلیل داده‌ها: در تجزیه و تحلیل داده‌های بزرگ، هوش مصنوعی می‌تواند الگوریتم‌های بهینه‌سازی پیچیده را اجرا کرده و الگوهای مخفی را کشف کند.

۴- استفاده از توابع مدل‌سازی در هوش مصنوعی

در زمینه هوش مصنوعی، استفاده از توابع مدل‌سازی یکی از مهم‌ترین اصول در تحقیقات و پیشرفت‌های اخیر به‌شمار می‌آید. توابع مدل‌سازی به عنوان مدل‌های محاسباتی تعریف می‌شوند که به واسطه آنها، می‌توان از داده‌ها و اطلاعات موجود در محیط‌ها الگوها و روابط مخفی را کشف کرد و از این طریق، پیش‌بینی‌ها و تصمیم‌گیری‌های هوشمندانه‌تری انجام داد.

در فرآیند یادگیری، توابع مدل‌سازی به کمک اطلاعات آموزشی، تلاش می‌کنند به‌طور خودکار الگوها و قوانین موجود در داده‌ها را فراگیرند. این فرآیند به گونه‌ای است که با تغییرپذیری و تطابق با داده‌ها، مدل‌ها قادر به بهترین تطبیق با ورودی‌ها و تولید خروجی‌های مطلوب می‌شوند.

مدل‌های هوش مصنوعی می‌توانند در انواع مسائل کاربرد داشته باشند؛ از مسائل ساده‌تر مانند پیش‌بینی نتایج یک آزمایش تا مسائل پیچیده‌تر مثل تشخیص الگوها در تصاویر پزشکی یا ترجمه متون به زبان‌های مختلف. هرچه توابع مدل‌سازی هوش مصنوعی پیچیده‌تر باشند، قدرت و دقت آنها در تحلیل داده‌ها و ارائه پاسخ‌ها به مسائل مختلف بیشتر می‌شود.

از مهم‌ترین توابع مدل‌سازی در هوش مصنوعی می‌توان به شبکه‌های عصبی، ماشین‌های پشتیبانی بردار (SVM) و رگرسیون خطی اشاره کرد. این توابع با توجه به خصوصیات مسئله و نوع داده‌ها، به عنوان مدل‌هایی قوی و کارآمد مورد استفاده قرار می‌گیرند.

در نتیجه، استفاده از توابع مدل‌سازی در هوش مصنوعی، باعث امکان یادگیری، تفسیر و تحلیل دقیق‌تر داده‌ها می‌شود و به سیستم‌های هوش مصنوعی امکان پیش‌بینی‌ها و تصمیم‌گیری‌های هوشمندانه‌تر را می‌دهد. همچنین، استفاده از توابع مدل‌سازی، توانایی شبیه‌سازی توانایی‌های انسانی در سیستم‌های مصنوعی را بهبود می‌بخشد و این موضوع، مسیری مهم و مبتکر در تحقق هدف تداعی از هوش انسانی در هوش مصنوعی را به نمایش می‌گذارد.

۵- هوش مصنوعی و ارتباط با مکانیک مواد

در مکانیک مواد برای بررسی و درک رفتار مواد آن را تحت آزمایش تنش-کرنش قرار می‌دهند. مواد در دسته‌های ترد و شکننده و نرم و شکل پذیر یا کشسان قرار می‌گیرند. مواد ترد مثل بتن، ماده پس از عبور از حد الاستیک، خاصیت شکل پذیری ندارد و به گسیختگی می‌رسند. در مواد نرم مثل فولاد که خاصیت شکل پذیری دارند، تحت تنش، ابتدا ماده به صورت الاستیک عمل می‌کند و با برداشتن بار از روی ماده، ماده به حالت اول باز می‌گردد، اگر تنش‌ها از حد الاستیک فراتر رود ماده حالت الاستیک نخواهد داشت و تغییر شکل دائمی پیدا خواهد کرد. در این مرحله ماده در برابر تغییر شکل مقاومت می‌کند و پس از آن اگر باز هم میزان تنش بیشتر شود ماده تا حالتی پیش خواهد رفت که گسیختگی در آن روی خواهد داد.

بعد فرا مادی انسان (روح یا ذهن) را مشابه یک ماده نرم در نظر داشته باشید که ماده در حالت الاستیک خاصیت ارتجاعی دارد، یعنی تحت تنش‌ها ماده تغییر شکل می‌دهد و پس از برداشتن بار از روی ماده و حذف تنش ماده به حالت اول خود باز می‌گردد. اگر برای ذهن ابعادی مشابه ابعاد مادی که با هم متعامد در نظر بگیریم، می‌توانیم بین این ابعاد تاثیراتی را در نظر داشته باشیم. همچنین تحت تنش‌های متفاوت ماده می‌تواند حالت ارتجاعی داشته باشد، به شیب نمودار تنش-کرنش در این حالت مدول یلنگ یا مدول الاستیسیته می‌گویند که نشان دهنده خاصیت ارتجاعی ماده است، در ذهن هم این حالت اتفاق می‌افتد. هنگامی که انسان تحت تنش‌های احساسی قرار می‌گیرد از خود کرنش‌هایی نشان می‌دهد و در صورت حذف این تنش‌ها به حالت ابتدایی خود باز می‌گردد که در بردارنده حالت ارتجاعی برای ذهن انسان است. همچنین میزان این تاثیر ابعاد بر روی هم می‌تواند میزان خلاقیت را در ذهن‌های متفاوت تداعی کند.

اگر به یک سازه و ساز و کار آن بنگریم می‌توانیم مشابهت آن را با نورون‌ها و انتقال اطلاعات در بین نورون‌ها و رسیدن به هدف دریابیم. نیروها به عنوان ورودی بر گره‌ها و مفاصل سازه و عضوها وارد می‌شوند و پس از انتقال نیروها بین اعضا جابه‌جایی‌ها و تنش‌ها و کرنش‌ها شکل می‌گیرند.



شکل (۱) - سازه و شباهت آن با ساختار مغز (انتقال اطلاعات از طریف اعضا و مفاصل)

۶- روابط تنش- کرنش ذهن

تنش و کرنش در مواد در حالت الاستیک با هم رابطه ای خطی دارند، از این رو می توان آن را به صورت زیر نمایش داد.

$$(\varepsilon - \varepsilon_0) = C\sigma \quad (۸)$$

که در آن σ نشان دهنده تنش، ε نشان دهنده کرنش، ε_0 کرنش اولیه و C ماتریس نرمی ماده است. ماتریس نرمی برای مواد الاستیک و همسان یک ماتریس متقارن و وابسته به مدول الاستیسیته و ضریب پواسن است. این ماتریس میزان تاثیر پذیری ابعاد از یکدیگر و رابطه تنش و کرنش را نمایش می دهد. کرنش های اولیه در مدل ذهن نشان دهنده تمایلات اولیه است. یعنی در حالتی که ذهن تحت تاثیر هیچ تنشی قرار ندارد میزان کرنش ها برابر کرنش اولیه خواهد بود. به عنوان مثال در حالتی عادی انسان میل به زیبایی، اخلاق خواد داشت.

۷- استفاده مهندسی مواد در بهبود هوش مصنوعی

استفاده از مهندسی مواد در طراحی هوش مصنوعی به ارتقاء و بهینه‌سازی عملکرد این مدل‌ها کمک می‌کند. با تحلیل دقیق ویژگی‌ها و رفتارهای مدل هوش مصنوعی با استفاده از روابط مکانیک مواد، می‌توان به ایجاد مدل‌های هوش مصنوعی با عملکرد بهتر و منحصر به فرد دست یافت. این امر به واسطه استفاده از ماتریس نرمی و استحکام مخصوص به امکان تنظیم ویژگی‌های اخلاقی و عاطفی و غیره هوش مصنوعی در هر مدل خاص کمک می‌کند. در نهایت، تغییرات کرنش‌های اولیه همچنان به ما این امکان را می‌دهند که تمایلات را مدل‌سازی کنیم. این ارتباطات مهم بین مهندسی مواد و هوش مصنوعی، زمینه‌ای را برای بهبود و پیشرفت هرچه بیشتر مدل‌های هوش مصنوعی فراهم می‌کنند و ما را به سوی راه‌حل‌های نوآورانه‌تر و کارآمدتر در این حوزه هدایت می‌کنند.

۸- رابطه سازی در روش اجزای محدود

یک جسم مجازی را شامل چندین ویژگی در نظر بگیرید، آنگاه فرضیه‌های زیر را برای این جسم و ویژگی‌های آن در نظر می‌گیریم.

الف- تغییر وضعیت جسم بر اثر تغییر ویژگی‌های موجود قابل اغماض است (این فرضیه موجب ایجاد روابط خطی کرنش-تغییر ویژگی می‌شود و همچنین اجازه می‌دهد تا معادلات تعادل بر اساس هندسه اولیه جسم مطرح شود).

ب- رفتار ماده تشکیل دهنده جسم خطی است (این فرض به روابط خطی تنش-کرنش منجر می‌شود)

فرض کنید تابع مدل‌سازی شامل ویژگی‌های x و y و z است و خروجی آن متاثر از تمام ویژگی‌ها نسبت به هر ویژگی است.

$$\begin{aligned}u_x &= \alpha_0 + \alpha_1 x + \alpha_2 y + \alpha_3 z \\u_y &= \beta_0 + \beta_1 x + \beta_2 y + \beta_3 z \\u_z &= \gamma_0 + \gamma_1 x + \gamma_2 y + \gamma_3 z\end{aligned}\quad (9)$$

می‌توانیم تابع را به صورت ماتریسی بازنویسی کنیم

$$\{U\} = [P]\{W\} \quad (10)$$

که در آن W بردار وزن و P ماتریس ویژگی است.

اگر تغییرات کوچک مدل نسبت به هر ویژگی را کرنش بنامیم، می‌توانیم این تغییرات را به صورت زیر نمایش دهیم.

$$\{\varepsilon\} = [L]\{U\} \quad (11)$$

L ماتریس عملگر مشتق است، این تغییرات حاصل از تنش‌های وارد بر جسم مجازی مذکور است.

با توجه به فرض‌های اولیه روابط تنش-کرنش برای یک ماده مجازی الاستیک خطی به صورت زیر در نظر گرفته می‌شود.

$$(\varepsilon - \varepsilon_0) = C\sigma \quad (12)$$

در رابطه بالا ε کرنش، ε_0 کرنش اولیه، σ تنش و C را ماتریس نرمی می‌نامیم، حالت معکوس این رابطه به صورت زیر است.

$$\sigma = D(\varepsilon - \varepsilon_0) \quad (13)$$

در این رابطه D ماتریس استحکام ماده مجازی است.

۸-۱ قضیه انرژی پتانسیل کل مینیمم

جسمی مجازی را که دچار تنش-کرنش است را در نظر بگیرید. برای چنین جسمی تابعی به صورت رابطه زیر برای کرنش‌ها بیان شده است.

$$V(\varepsilon_x, \varepsilon_y, \dots) = \frac{1}{2} \varepsilon^T D_m \varepsilon = \frac{1}{2} \int \varepsilon^T D_m \varepsilon d\Omega \quad (14)$$

تابع V را کرنش ذخیره شده در واحد حجم جسم مجازی است. ماتریس D_m همان ماتریس استحکام ماده مجازی است. تغییرات V بر حسب تغییرات کرنش برابر است با δV که با بسط دادن آن داریم

$$\delta V = \frac{1}{2} (\delta \varepsilon^T D_m \varepsilon + \varepsilon^T D_m \delta \varepsilon) \quad (15)$$

چنانچه ماتریس استحکام متقارن باشد، با فرض رفتار الاستیک خطی می‌توان نوشت

$$\delta \varepsilon^T D_m \varepsilon = \varepsilon^T D_m \delta \varepsilon = \sigma^T \delta \varepsilon \quad (16)$$

بنابراین

$$\delta V = \sigma^T \delta \varepsilon = \sigma_x \delta \varepsilon_x + \sigma_y \delta \varepsilon_y + \dots \quad (17)$$

δV به صورت زیر قابل محاسبه است

$$\delta V = \frac{\partial V}{\partial \varepsilon_x} \delta \varepsilon_x + \frac{\partial V}{\partial \varepsilon_y} \delta \varepsilon_y + \dots \quad (18)$$

انرژی کرنشی ذخیره شده در کل جسم مجازی به صورت زیر است

$$V_t = \iiint_{volume} V d(volume) \quad (19)$$

با استفاده از قضیه کار مجازی داریم

$$\delta W_I = \delta V_t = \iiint_{volume} \delta V \, d(volume) \quad (20)$$

اکنون یک چگالی پتانسیل نیروی Ω را که از دو قسمت زیر تشکیل شده است، در نظر می‌گیریم

$$B = -b_x u - b_y v + \dots \quad (21)$$

$$P = -\bar{p}_x u - \bar{p}_y v + \dots$$

در معادلات (21)، B پتانسیل نیروهای جسمی در داخل جسم و P پتانسیل نیروهای سطحی در روی مرز مکانیکی جسم است. تغییرات این دو کمیت بر حسب تغییرات مکان‌ها به صورت زیر است.

$$-\delta B = b_x \Delta u + b_y \Delta v + \dots \quad (22)$$

$$-\delta P = \bar{p}_x \Delta u + \bar{p}_y \Delta v + \dots$$

با توجه به قضیه کار مجازی می‌توان نوشت

$$\iiint_{volume} \delta V \, d(volume) + \iiint_{volume} \delta B \, d(volume) + \iint_{S_\sigma} \delta P \, ds = 0 \quad (23)$$

یا

$$\delta \pi_P = 0 \quad (24)$$

که در آن داریم

$$\begin{aligned} \pi_P &= \iiint_{volume} V \, d(volume) + \iiint_{volume} B \, d(volume) + \iint_{S_\sigma} P \, dS \\ &= V_t + \Omega_t \end{aligned} \quad (25)$$

در رابطه (25)، Ω انرژی پتانسیل کل نیروها و π انرژی پتانسیل کل سیستم است. تعادل این سیستم را می‌توان به صورت زیر بیان کرد

$$\delta \pi_P = \delta V_t + \delta \Omega_t = 0 \quad (26)$$

با کمینه شدن پتانسیل کل، جسم در حالت تعادل قرار می‌گیرد. لذا با صفر شدن مشتق انرژی پتانسیل کل نسبت به متغیرهای مکانی (ویژگی‌ها)، دستگاه معادله حاکم به دست می‌آید. توابع مکانی بر حسب توابع درونیاب به صورت رابطه زیر قابل توصیف است.

$$\{U\} = [\phi]\{D\} \quad (27)$$

اگر ماتریس L را اپراتور مشتق باشد روابط تغییر مکان-کرنش به صورت رابطه زیر حاصل می‌شود.

$$\{\varepsilon\} = [L]\{U\} = [L][\phi]\{D\} = [B]\{D\} \quad (28)$$

از این رو چگالی انرژی پتانسیل V مطابق رابطه زیر به دست می‌آید.

$$V = \frac{1}{2} \{D\}^T \int [B]^T [D_m] [B] d\Omega \{D\} \quad (29)$$

با جای گذاری روابط بدست آمده در رابطه (۱۸) رابطه (۲۲) تشکیل می شود.

$$\begin{aligned} \pi = \frac{1}{2} \{D\}^T \int_{\Omega} [B]^T [D_m] [B] d\Omega \{D\} - \{D\}^T (\{P_J\} + \sum [\phi] \{P_{ie}\} \\ + \int_{\Omega} [\phi] \{P_b\} d\Omega + \int_{\Gamma} [\phi] \{P_T\} d\Gamma) \end{aligned} \quad (30)$$

با کمینه کردن این رابطه نسبت به تغییر مکان معادله زیر حاصل می شود.

$$\begin{aligned} \int_{\Omega} [B]^T [D_m] [B] d\Omega \{D\} - (\{P_J\} + \sum [\phi] \{P_{ie}\} + \int_{\Omega} [\phi] \{P_b\} d\Omega \\ + \int_{\Gamma} [\phi] \{P_T\} d\Gamma) = [S] \{D\} - \{P\} = 0 \end{aligned} \quad (31)$$

در رابطه بالا Ω نشان دهنده حجم و Γ نشان دهنده سطح است. ماتریس S را ماتریس سختی جزء و P و D را بردار نیرو و جابه جایی گویند. مقادیر P_J , P_{ie} , P_b و P_T به ترتیب نیروهای گرهی، نیروهای نقطه‌ای، نیروهای حجمی و نیروهای سطحی هستند.

۲-۸ تابع شکل

با استفاده از روابط درونیاب لاگرانژی توابع جزء شش وجهی-هشت گرهی را می توان به صورت زیر در مختصات محلی نوشت.

$$\varphi_i = \frac{1}{8} (1 - \xi_i \xi)(1 - \eta_i \eta)(1 - \zeta_i \zeta), \quad i = 1, 2, \dots, 8 \quad (32)$$

۳-۸ اپراتور مشتق

اپراتور مشتق و مشتق توابع شکل به صورت زیر هستند.

$$L = \begin{bmatrix} \frac{\partial}{\partial x} & 0 & 0 \\ 0 & \frac{\partial}{\partial y} & 0 \\ 0 & 0 & \frac{\partial}{\partial z} \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} & 0 \\ 0 & \frac{\partial}{\partial z} & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} & 0 & \frac{\partial}{\partial x} \end{bmatrix} \quad (33)$$

$$B_i = \begin{bmatrix} \frac{\partial \varphi_i}{\partial x} & 0 & 0 \\ 0 & \frac{\partial \varphi_i}{\partial y} & 0 \\ 0 & 0 & \frac{\partial \varphi_i}{\partial z} \\ \frac{\partial \varphi_i}{\partial y} & \frac{\partial \varphi_i}{\partial x} & 0 \\ 0 & \frac{\partial \varphi_i}{\partial z} & \frac{\partial \varphi_i}{\partial y} \\ \frac{\partial \varphi_i}{\partial z} & 0 & \frac{\partial \varphi_i}{\partial x} \end{bmatrix} \quad (34)$$

۴-۸ ژاکوبین و تبدیل مختصات

از آنجایی که توابع شکل بر حسب مختصات طبیعی بیان شده است، برای تبدیل مختصات از ژاکوبی استفاده می‌شود.

$$J = \sum_{i=1}^n \begin{bmatrix} \frac{\partial \varphi_i}{\partial \xi} x_i & \frac{\partial \varphi_i}{\partial \xi} y_i & \frac{\partial \varphi_i}{\partial \xi} z_i \\ \frac{\partial \varphi_i}{\partial \eta} x_i & \frac{\partial \varphi_i}{\partial \eta} y_i & \frac{\partial \varphi_i}{\partial \eta} z_i \\ \frac{\partial \varphi_i}{\partial \zeta} x_i & \frac{\partial \varphi_i}{\partial \zeta} y_i & \frac{\partial \varphi_i}{\partial \zeta} z_i \end{bmatrix} \quad (35)$$

$$\begin{pmatrix} \frac{\partial \varphi_i}{\partial x} \\ \frac{\partial \varphi_i}{\partial y} \\ \frac{\partial \varphi_i}{\partial z} \end{pmatrix} = J^{-1} \begin{pmatrix} \frac{\partial \varphi_i}{\partial \xi} \\ \frac{\partial \varphi_i}{\partial \eta} \\ \frac{\partial \varphi_i}{\partial \zeta} \end{pmatrix} \quad (36)$$

۹- رویکرد

هر جزء در روش المان محدود معادل یک نورون است. از طریق درجات آزادی و انتقال نیروها و جابه‌جایی با سایر نورون‌ها ارتباط برقرار می‌کند. مختصات محلی نیز معادل تابع فعال‌ساز در روش‌های پیشین است. ماتریس استحکام نیز مقادیر وزن و بایاس تابع هستند. در این روش برای آموزش مدل هوش مصنوعی، نیازمند تغییراتی در رابطه‌سازی هستیم. از آنجایی که ماتریس استحکام به متغیرهای انتگرال وابسته نیست، رابطه (۳۱) را می‌توانیم به صورت زیر نمایش دهیم.

$$[M_D] \int_{\Omega} [B]^T [B] d\Omega \{D\} = \{P\} \rightarrow [M_D] [S_c][D] = [P] \quad (37)$$

$$= [M_D] [D_s] = \{P\}$$

که در آن ماتریس M_D ماتریس ساختاری مواد است.

با استفاده از روش های بهینه سازی ماتریس استحکام M_D به عنوان وزن های سیستم معادلاتی بدست خواهند آمد. ماتریس S_c که ماتریس ساختار است وظیفه لایه بندی بین نوروں ها و تابع فعال سازی را بر عهده دارد. با جایگذاری ورودی مقادیر پیش بینی در درجات آزادی مشلبه درجه آزادی ورودی بدست می آید. با محاسبه ماتریس استحکام M_D ماتریس سختی K از روابط زیر قابل محاسبه است.

$$[K] = [M_D] \int_{\Omega} [B]^T [B] d\Omega \quad (38)$$

پس از آموزش مدل به این صورت با محاسبه ماتریس سختی می توان به وسیله مدل دسته بندی و پیش بینی کرد.

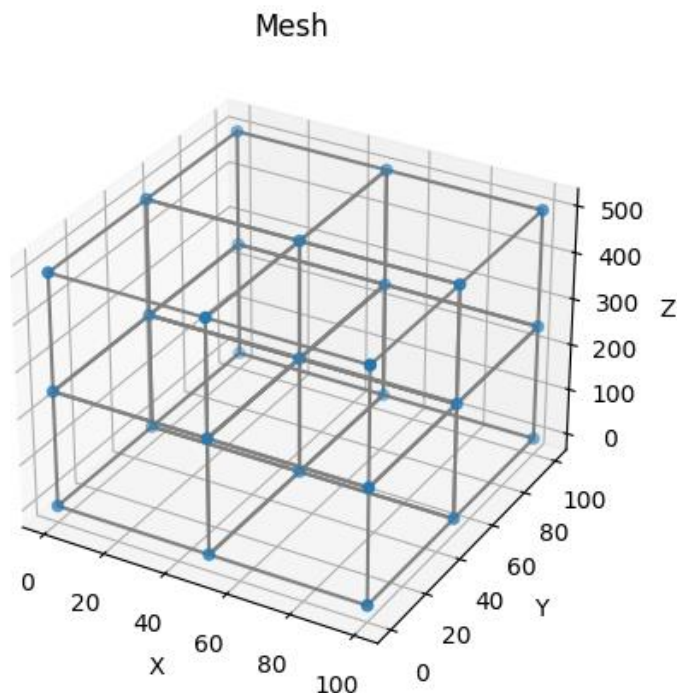
$$[K] \{D\} = \{P\} \quad (39)$$

ماتریس مواد از رابطه ی زیر قابل استخراج است.

$$[M] = ([B]^T)^{-1} [M_D] [B]^T \quad (40)$$

۱۰- بررسی رویکرد

برای بررسی روش و رویکرد، پیش بینی آلودگی هوا در بلژیک و به طور مشخص تر، آلودگی هوای ایجاد شده بر اثر گوگرد دی اکسید (SO_2) را مورد بررسی قرار داده ایم. برای مدل سازی داده ها و روش از زبان پایتون بهره گرفته شده است و در حل مسائل از روش عددی نقاط گوس استفاده شده است. مقادیر زمانی را به عنوان ورودی به یکی از درجات آزادی که در شکل (۲) آمده اعمال کرده ایم. مقادیر هدف مقادیر آلودگی هوا بلژیک است که در یکی دیگر از گره های سازه فرضی اعمال شده است. شبکه بندی مدل در مطابق شکل (۱) انجام شده است. برای فرایند بهینه سازی از روش رهایی پویا که از معادل سازی مسائله با نوسانات فنر و میرایی آن بهره میبرد استفاده شده است که سرعت بسیار بالایی در حل مسائل دستگاه معادلات سازه ای دارد.



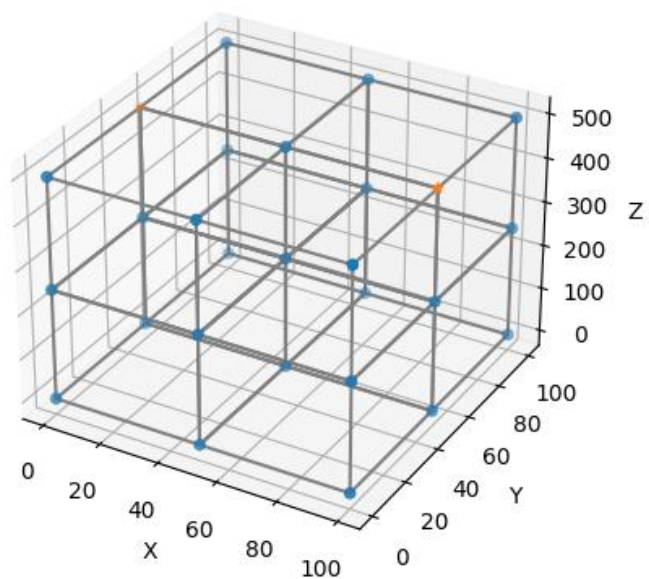
شکل (۲) - شبکه بندی فرضی سازه به عنوان فضای آموزش شبکه عصبی

این شبکه شامل دو ردیف در هر راستا عمود بر هم است و تعداد هشت شش وجهی هشت گرهی را داده است. هر کدام از گره ها سه درجه آزادی دارد و هر شش وجهی ایجاد ۲۴ درجه آزادی می کند که در برخی گره ها با شش وجهی مجاور اشتراک دارد که وظیفه انتقال پیام در شبکه های عصبی را انجام می دهند. هر شش وجهی معادل یک لایه در شبکه عصبی است و هر گره یک نورون را نشان خواهد داد. بنابراین هر شش وجهی یک لایه است و هشت نورون دارد و هر نورون شامل سه ویژگی است.

تکه گاه ها شرایط مرزی را نشان می دهند که در شکل (۴) محل آنها نمایش داده شده است. لازم به ذکر است که تعداد داده های آموزشی ۵۴۶ و تعداد داده های اعتبار سنجی ۱۸۴ و تعداد داده های آزمون ۳۶۵ عدد است. برای مشاهده بهتر تغییرات نرخ یادگیری پایین انتخاب شده است.

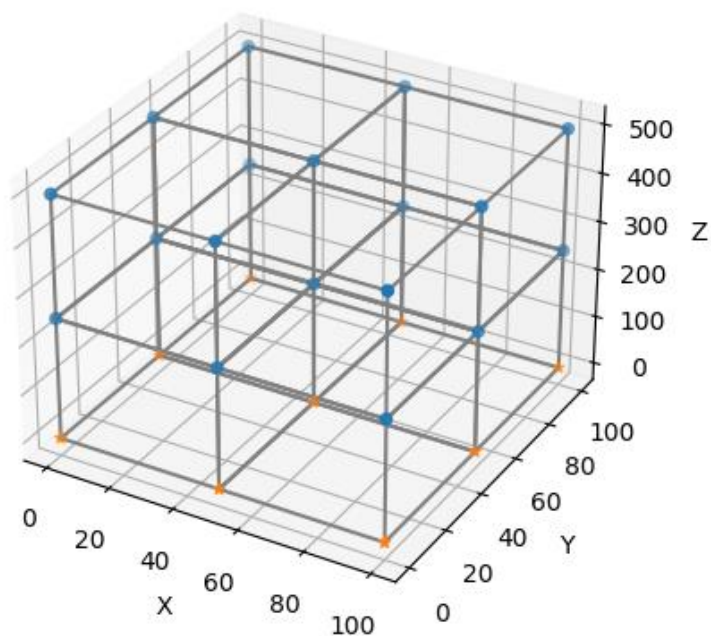
مقادیر نهایی هزینه برای داده های آموزشی ۰.۰۰۴۸۰۵۵۹۹۱۲۱۹۸۴۸۴۵ و برای داده های اعتبار سنجی ۰.۰۰۰۱۷۵۳۸۲۱۲۲۹۴۳۴۵۱۸۸ بدست آمده است. این مقادیر به دلیل نرخ پایین آموزش کمی بزرگتر از حالت معمول هستند. تعداد دوره ها برای انجام آموزش ۱۳۶۳ عدد بوده است. با افزایش نرخ یادگیری مدت زمان اجرای کد ۱۳.۷۹۷۵ ثانیه بوده است. مقادیر هزینه برای داده های آموزشی ۰.۰۰۰۰۱۲۲۶۳۳۵۲۱۳۱۹۴۳۶۲۵ و برای داده های اعتبار سنجی ۰.۰۰۰۰۱۴۵۹۰۱۲۲۵۳۵۴۵۲۰۴۷ بوده است. تعداد دوره های آموزش در این حالت ۱۲ دوره بوده است.

Effect Position

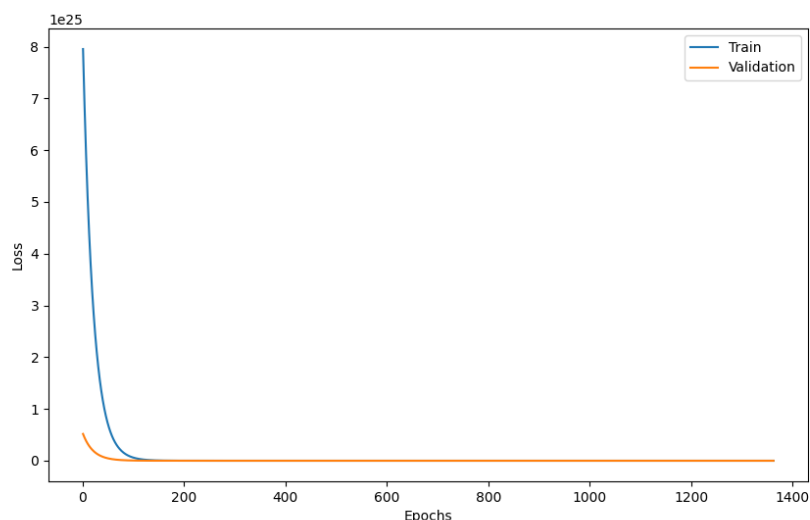


شکل (۳) - اعمال ورودی و هدف در نقاط مشخص شده به عنوان نیرو و جابه‌جایی

Support Position

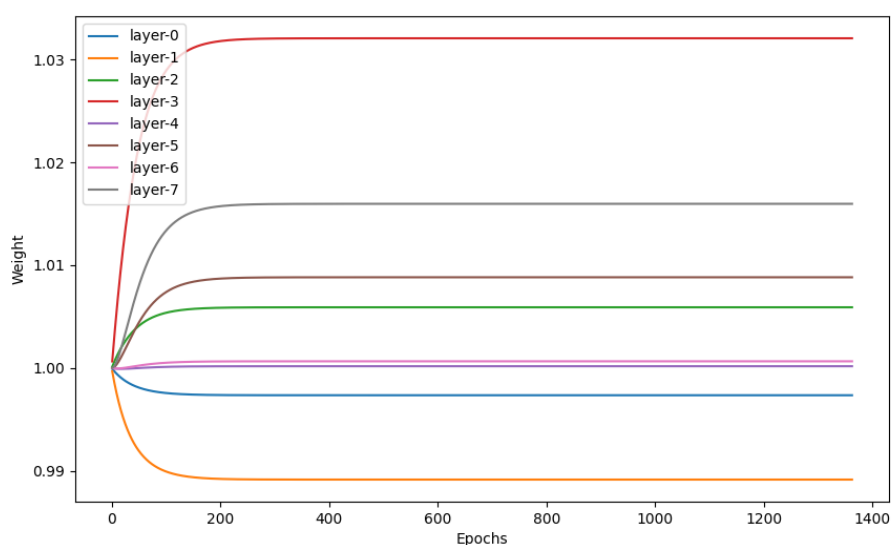


شکل (۴) - محل اعمال تکیه‌گاه‌ها بر سازه فرضی



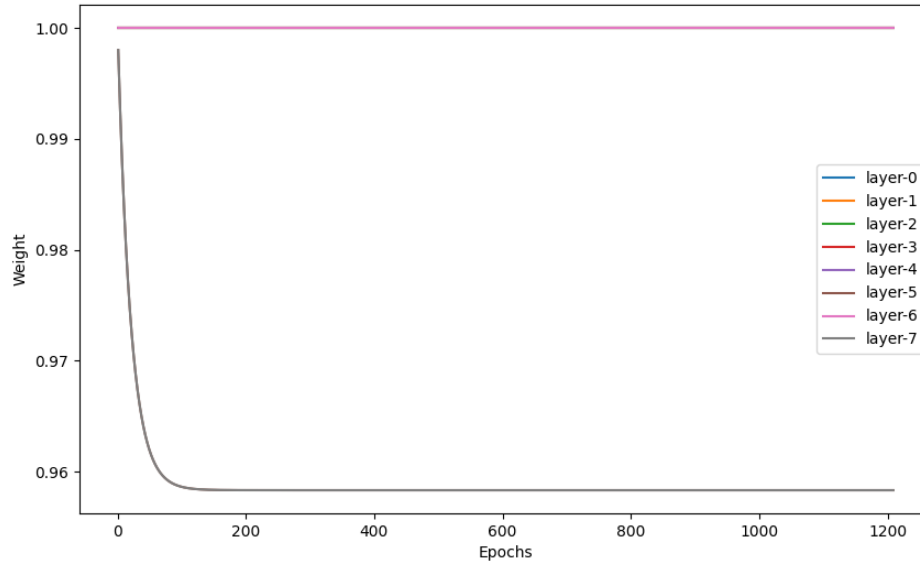
شکل (۵) - نمودار تابع هزینه در برابر تعداد دوره ها با نرخ آموزش پایین

شکل بالا مقایسه نرخ اعتبار سنجی را در برابر داده های آموزشی و در طول روند آموزش نمایش می دهد.



شکل (۶) - مقادیر میانگین وزن هر لایه (شش وجهی)

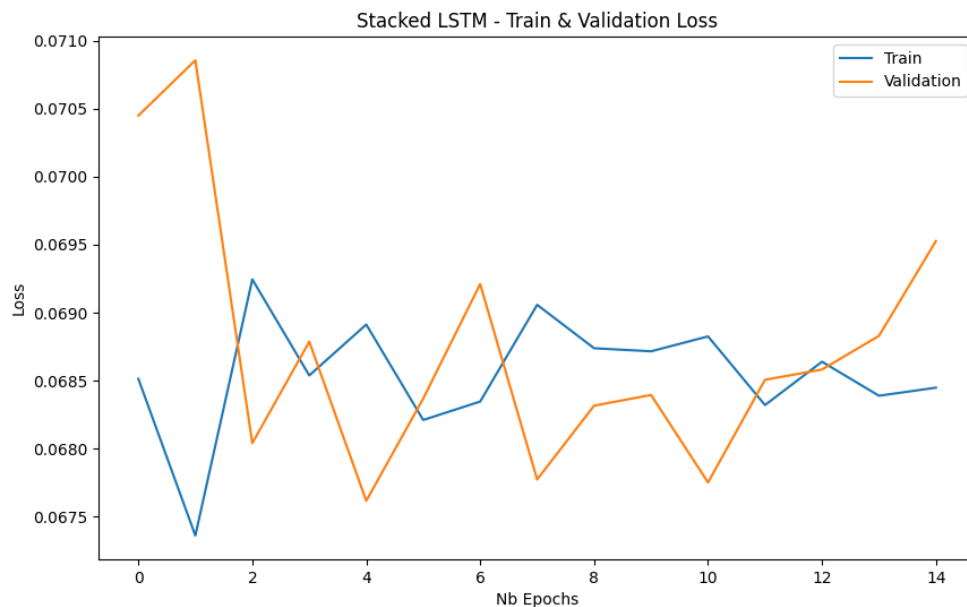
مقادیر وزن برای هر شش وجهی (لایه عصبی) که توزیع وزن ها را در طی فرایند آموزش نشان می دهد. در عدم استفاده از ماتریس ساختار S_e و تاثیر آن بر روی ورودی ها D مقادیر وزن در طی آموزش به صورت شکل (۷) است. مقایسه نتایج مشخص کننده تاثیر ماتریس ساختار بر توزیع وزن ها و دخالت تمام لایه ها و انتقال اطلاعات بین لایه ها و نورون هاست.



شکل (۷) – مقادیر وزن هر لایه در عدم استفاده از ماتریس ساختار

توزیع وزن ها به جلوگیری از فراموشی داده ها در اثر تغییرات ناگهانی مقادیر وزن موثر است و تمام نرون ها را در طی فرایند آموزش موثر می کند.

برای مشخص شدن دقت روش داده های آموزشی با روش شبکه های عصبی LSTM مقایسه می کنیم. این شبکه عصبی شامل هشت لایه و یک لایه ورودی با یک ویژگی و یک لایه Danse است.



شکل (۸) – مقادیر هزینه داده های آموزشی و اعتبار سنجی در طی فرایند آموزش

مدت زمان اجرای کد ۲۴۴.۷۷۸۴ ثانیه و بهترین مقادیر هزینه برای داده های اعتبار سنجی ۰.۰۶۷۶۱۷۸۳۳۶۱۴۳۴۹۳۷ است.

۱۱- نتیجه گیری

مقایسه مقادیر اعتبار سنجی، بهبود عملکرد ۹۹.۹۸ درصدی و ۵۵۱۳ برابری را نسبت به روش LSTM نشان می دهد. توزیع وزن ها به فرایند آموزش کمک می کند و مشکل از دست رفتن اطلاعات را از میان برمی دارد. ساختار بر روی مسیر انتقال اطلاعات موثر است و شکل سازه می تواند مسیر انتقال را تعیین کند. تعداد ورودی ها محدودیت ندارد و فرایند حل مسئله بسیار سریع است. هم چنین از لحاظ زمانی ۱۷.۷۴ برابر سریع تر عمل کرده است. بررسی تنش ها و کرنش های حاصل از این روش نیز می توانند اطلاعات کمکی را برای حل مسائل در اختیار ما قرار دهند و همچنین مقادیر کرنش اولیه می تواند در جهت دهی و تمایلات در مسائل هوش انسانی بر حل مسائل تاثیر به سزایی داشته باشد.

۱۲- دید آینده

با توجه به این موضوع که روش بر اساس استفاده از ماتریس مواد به عنوان وزن های شبکه عصبی است می توان از این روش برای ساختارهای فیزیکی نیز استفاده کرد، به عنوان مثال معکبی از قطعات کوچک تر ایجاد کرد که با استفاده از آن بتوان مسائل ریاضی مانند ضرب و جمع را بر روی اعداد دودویی انجام داد. شکل سازه بر روی مسیر انتقال اطلاعات تاثیر دارد و با استفاده از آن میتوان مسائل گوناگون را با استفاده از یک مدل حل کرد. تعداد بالای گره ها امکان دریافت تعداد زیادی از ویژگی ها را به عنوان ورودی فراهم می کند. با افزایش تعداد درجات آزادی چرخشی به روش می توان تعداد ویژگی ها را برای هر گره از ۳ به ۶ افزایش داد. افزایش تعداد ابعاد امکانات بیشتری را برای حل مسائل هوش مصنوعی به ما می دهد. به عنوان مثال افزایش بعد زمان امکان دریافت چند پیش بینی را تحت بررسی شرایط زمانی برای ما فراهم خواهد کرد، یا در مسائل هوش انسانی شرایط برای ایجاد اختیار بین چندین نتیجه که از زمان های متفاوت حاصل شده را برای ما فراهم می کند.

منابع

- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
- Russell, S. J., & Norvig, P. (2016). Artificial Intelligence: A Modern Approach. Pearson.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.
- Jurafsky, D., & Martin, J. H. (2020). Speech and Language Processing. Pearson.
- Domingos, P. (2012). A few useful things to know about machine learning. Communications of the ACM, 55(10), 78-87.
- McKinsey Global Institute. (2017). Artificial Intelligence: The Next Digital Frontier? McKinsey & Company.

Haykin, S. (2009). Neural Networks and Learning Machines. Pearson.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

Nielsen, M. A. (2015). Neural Networks and Deep Learning: A Textbook. Determination Press.

Bishop, C. M. (1995). Neural Networks for Pattern Recognition. Oxford University Press.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. Neural Networks, 61, 85-117.

Nielsen, M. A. (2015). Neural Networks and Deep Learning: A Textbook. Determination Press.

Haykin, S. (2009). Neural Networks and Learning Machines. Pearson.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

Bishop, C. M. (1995). Neural Networks for Pattern Recognition. Oxford University Press.

كد

<https://github.com/systbs/ai-prism>