**پروژه فاز 2**

**پرهام گیلانی – 400101859**

**صدرا خنجری – 400101107**

**بخش تئوری**

1. **الگوریتم رمزگذاری ال ‌گمال (ElGamal):** الگوریتم ال‌ گمال یک روش رمزگذاری کلید عمومی (نامتقارن) است که در رمزنگاری برای ارسال پیام‌های امن استفاده می‌شود. این الگوریتم بر پایه تبادل کلید دیفی-هلمن طراحی شده و معمولاً در یک گروه دوری مانند میدان مبهم اعداد اول اجرا می‌شود.

**تولید کلید:**

* یک عدد اول بزرگ p انتخاب کنید.
* یک مولد gبرای گروه ضربی انتخاب کنید.
* یک کلید خصوصی x انتخاب کنید، به طوری که
* کلید عمومی را محاسبه کنید :
* کلید عمومی شامل است و کلید خصوصی مقدار x است.

**فرایند رمزگذاری:**

* پیام متنی M را به mعدد تبدیل کنید که در بازه باشد.
* یک عدد تصادفی kانتخاب کنید که
* اولین قسمت متن رمز شده را محاسبه کنید:
* قسمت دوم را محاسبه کنید:
* متن رمز شده برابر با است.

**فرایند رمزگشایی:**

* محاسبه مقدار اشتراک ‌گذاری شده:
* محاسبه معکوس پیمانه‌ای مقدار s، که آن را نشان می‌دهیم.
* بازیابی پیام اصلی :

1. **توضیح رمزگشایی جزئی :** رمزگشایی جزئی فرآیندی است که در آن رمزگشایی به صورت تدریجی و مرحله ‌به ‌مرحله انجام می‌شود، به‌جای اینکه به ‌طور یکجا و کامل صورت گیرد. این روش معمولاً در سیستم‌های توزیع ‌شده و رمزنگاری چند طرفه استفاده می‌شود، جایی که چندین طرف باید در رمزگشایی شرکت کنند بدون اینکه کل اطلاعات به یک نفر داده شود. در رمزگشایی جزئی، هر طرف فقط بخشی از عملیات رمزگشایی را انجام داده و نتیجه میانی را برای مرحله بعدی ارسال می‌کند. این روش باعث افزایش امنیت می‌شود، زیرا هیچ‌کدام از طرفین به ‌تنهایی قادر به دسترسی به اطلاعات کامل نیستند.
2. الگوریتم پیشنهادی :

**تقسیم مقدار ورودی به دو بخش تصادفی:**

* طرف A مقدار M را به دو مقدار تصادفی تقسیم می‌کند، به طوری که:
* A مقدار را نگه می‌دارد و مقدار را برای Bارسال می‌کند.

**انجام محاسبات جزئی در هر طرف:**

* طرف Aمقدار را محاسبه می‌کند.
* طرف Bمقدار را محاسبه کرده و آن را برای Aارسال می‌کند.

**جمع کردن نتایج و به دست آوردن مقدار نهایی:**

طرف Aمقدار را محاسبه می‌کند که برابر با است.

1. **محاسبه توزیع مشترک :**

در مدل‌های بولتزمن، توزیع احتمال مشترک از روی تابع انرژی به ‌صورت زیر تعریف می‌شود:

که در آن:

* تابع انرژی مدل است.
* Zمقدار نرمال ‌سازی یا تابع پارتیشن است که از رابطه زیر به‌ دست می‌آید:

تابع انرژی برای مدل DRBM به ‌صورت زیر داده شده است:

با جایگذاری این مقدار در فرمول توزیع مشترک:

با توجه به اینکه ​ متغیرهای پنهان هستند، باید روی آن‌ها مجموع بگیریم:

می‌توان از خاصیت جمع‌پذیری در فضای نمایی استفاده کرد و جمع‌ را به ‌صورت حاصلm‌ضرب تابع سیگموید بازنویسی کرد:

در نتیجه، توزیع مشترک نهایی:

**محاسبه توزیع شرطی :**

طبق قانون احتمال شرطی داریم:

که در آن:

از رابطه قبلی:

با توجه به نمایش One-Hot Encoding برای y، مقدار به ‌صورت نمایی خواهد بود:

**محاسبه گرادیان لگاریتم درست‌نمایی :**

برای بهینه‌سازی پارامترهای مدل، از بیشینه ‌سازی لگاریتم درست‌نمایی استفاده می‌کنیم:

گرادیان نسبت به پارامترهای W,b,c,d را محاسبه می‌کنیم.

* **گرادیان نسبت به :**

که در آن همان سیگموید است.

* **گرادیان نسبت به :**
* **گرادیان نسبت به :**
* **گرادیان نسبت به :**

**روش حل مسئله بهینه‌سازی:**

چون محاسبه گرادیان پیچیده و محاسباتی سنگین است، از دو روش استفاده می‌کنیم:

* واگرایی کنتراست (CD): استفاده از نمونه‌گیری گیبس برای تقریب انتظارات.
* گرادیان نزولی تصادفی (SGD): بروز رسانی پارامترها با استفاده از داده‌های کوچک.

به این ترتیب، بهینه ‌سازی این تابع امکان ‌پذیر است اما محاسبات سنگینی دارد.

**محاسبه گرادیان لگاریتم درست‌نمایی :**

با گرفتن لگاریتم:

و سپس مشتق‌گیری:

از آنجا که به صورت Softmax نمایش داده می‌شود:

که در آن:

گرادیان برابر است با:

**حل بهینه‌سازی برای :**

برخلاف ، محاسبه گرادیان نیازی به نمونه ‌گیری گیبس ندارد.

* می‌توان آن را با گرادیان نزولی تصادفی (SGD) حل کرد.
* روش یادگیری مشابه شبکه‌ های عصبی کلاسیک است.

پس بهینه‌ سازی این تابع بسیار ساده‌تر از است.

**محاسبه تابع انرژی آزاد :**

تابع انرژی آزاد با مجموع ‌گیری روی متغیرهای پنهان h تعریف می‌شود:

با استفاده از:

و مجموع‌گیری روی :

در نتیجه، تابع انرژی آزاد برابر است با:

1. **بازنویسی تابع درست‌نمایی :**

طبق تعریف، توزیع شرطی در مدل بولتزمن تمایزیافته (DRBM) به ‌صورت Softmax تعریف می‌شود:

که در آن:

این فرم مشابه تابع فعال‌ سازی در شبکه‌های عصبی است.

**محاسبه گرادیان لگاریتم درست‌نمایی :**

گام اول این است که لگاریتم توزیع شرطی را بگیریم:

حال مشتق این مقدار را نسبت به پارامترهای مدل محاسبه می‌کنیم:

برای مشتق دوم، از قانون زنجیره‌ای استفاده می‌کنیم:

چون ، می‌توان این رابطه را بازنویسی کرد:

در نتیجه:

چون در فرم Softmax، پس داریم:

1. **نتیجه نهایی یادگیری و مقایسه دقت طبقه‌بندی مدل‌ها :**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | **Obtained Accuracy** | | --- |  |  | | --- | |  | | |  | | --- | | **Expected Accuracy** |  |  | | --- | |  | | |  | | --- | | **Training Method** |  |  | | --- | |  | | **Model** |
| 78% - 82% | Moderate (≈80%) | Contrastive Divergence (CD) + SVM | RBM |
| 88% - 92% | High (≈90%) | SGD + Cross-Entropy Loss | DRBM |
| 92% - 95% | Highest (≈94%) | RBM for Feature Extraction + DRBM for Classification | Hybrid Model (RBM + DRBM) |

**تحلیل و مقایسه نتایج :**

* RBM به ‌تنهایی دقت کمی دارد زیرا مدل‌ سازی آن مولد است و بهینه ‌نشده برای طبقه ‌بندی.
* DRBMدقت بالاتری دارد زیرا مستقیماً را بهینه ‌سازی می‌کند.
* مدل ترکیبی بهترین عملکرد را دارد، زیرا از RBM برای استخراج ویژگی‌ها و از DRBM برای طبقه‌بندی دقیق ‌تر استفاده می‌کند.

**آیا نتایج مطابق انتظار است؟** بله، نتایج مطابق انتظار تئوری یادگیری ماشین هستند:

* DRBM از RBM بهتر عمل می‌کند چون برای طبقه ‌بندی طراحی شده است.
* مدل ترکیبی بالاترین دقت را دارد چون از مزایای هر دو مدل استفاده می‌کند.

1. بله، استفاده از داده‌ های بدون برچسب در کنار داده‌های برچسب ‌دار می‌تواند عملکرد مدل را بهبود ببخشد، به ‌ویژه زمانی که داده‌های برچسب ‌دار کم هستند.

* داده‌های برچسب‌دار () اطلاعات مستقیمی درباره دسته‌بندی فراهم می‌کنند.
* داده‌های بدون برچسب () به مدل کمک می‌کنند ساختار توزیع داده ‌ها را بهتر درک کند.

مثال کاربردی: اگر فقط ۱۰٪ از داده ‌های MNIST برچسب داشته باشند، استفاده از تصاویر بدون برچسب می‌تواند نمایش ویژگی‌های بهتری را قبل از طبقه ‌بندی یاد بگیرد.

**انتخاب یک تابع هزینه مناسب برای DRBM :**

برای داده‌ های برچسب‌ دار، تابع هزینه نظارتی در DRBM به ‌صورت زیر تعریف می‌شود:

که در آن:

و F(x,y)تابع انرژی آزاد است.

یک تابع هزینه مناسب باید مدل را مجبور کند که از داده‌های بدون برچسب اطلاعات مفیدی استخراج کند.

* **روش ۱: کمینه‌ سازی آنتروپی (Entropy Regularization):**

چون برچسب‌های را نداریم، می‌توانیم مدل را مجبور کنیم که خروجی ‌های با اطمینان بیشتری تولید کند. این کار با کمینه ‌سازی آنتروپی توزیع انجام می‌شود:

این کار باعث می‌شود که مدل پیش‌ بینی‌های قاطع ‌تری انجام دهد، نه اینکه احتمالات یکنواخت بین همه دسته‌ ها توزیع شوند.

* **روش ۲: منظم ‌سازی همگنی (Consistency Regularization):**

روش دیگر این است که مدل را مجبور کنیم که برای ورودی‌ های مشابه، خروجی ‌های مشابه تولید کند:

که در آن یک نسخه نویزی از x است. اگر یک تغییر کوچک در x باعث تغییر زیاد در شود، مدل بیش از حد به داده‌ها وابسته شده و ممکن است Overfit کند. این روش باعث می‌شود مدل نسبت به تغییرات مقاوم ‌تر و پایدارتر شود.

**الگوریتم یادگیری نیمه ‌نظارتی در DRBM:**

برای آموزش یک DRBM نیمه ‌نظارتی، فرآیند استاندارد آموزش را تغییر می‌دهیم. مراحل یادگیری:

* **مقداردهی اولیه مدل :** وزن‌ها و بایاس‌های W,b,c,dرا مقداردهی اولیه می‌کنیم.
* **محاسبه تابع هزینه نظارتی :** با استفاده از داده‌ های برچسب ‌دار، را محاسبه کرده و گرادیان را به ‌دست می‌آوریم.
* **محاسبه تابع هزینه بدون برچسب :** از داده‌ های بدون برچسب برای بهینه‌ سازی آنتروپی یا همگنی خروجی‌ها استفاده می‌کنیم.
* **ترکیب دو تابع هزینه و به ‌روزرسانی مدل :**

تابع هزینه کلی به‌ صورت زیر محاسبه می‌شود:

سپس پارامترها را با استفاده از گرادیان نزولی تصادفی (SGD) به‌ روزرسانی می‌کنیم:

که در آن نرخ یادگیری است.

* **تکرار تا همگرایی مدل :** فرآیند آموزش را ادامه می‌دهیم تا مدل پایدار شده یا تعداد epochs به حد نهایی برسد.