دانشکده مهندسی برق و کامپیوتر

گروه مهندسی کامپیوتر

عنوان:

پیاده‌سازی و ارزیابی مقایسه‌ای طبقه‌بندهای خطی SVM و Softmax بر روی مجموعه داده CIFAR-۱۰

استاد محترم:

دکتر مریم شعاران

دانشجو:

امیررضا میرزاجانی اسکوئی

مرداد ۱۴۰۴

نام خانوادگی دانشجو: میرزاجانی اسکوئی	**نام:** امیررضا
عنوان: پیاده‌سازی و ارزیابی مقایسه‌ای طبقه‌بندهای خطی SVM و Softmax بر روی مجموعه داده CIFAR-۱۰	
استاد محترم: دکتر مریم شعاران	
مقطع تحصیلی: کارشناسی ارشد **رشته:** مهندسی مکانیک **گرایش:** طراحی کاربردی **دانشگاه:** تبریز **دانشکده:** مهندسی مکانیک **تعداد صفحات:** ۱۶	
کلیدواژه‌ها: طبقه‌بندی تصاویر، ماشین بردار پشتیبان(SVM) ، طبقه‌بندSoftmax ، گرادیان کاهشی تصادفی(SGD) ، CIFAR-۱۰	
چکیده: این گزارش به پیاده‌سازی و ارزیابی مقایسه‌ای دو طبقه‌بند خطی بنیادین، ماشین بردار پشتیبان (SVM) و Softmax، برای مسئله طبقه‌بندی تصاویر بر روی مجموعه داده CIFAR-۱۰ می‌پردازد. در این پروژه، توابع هزینه و گرادیان تحلیلی هر دو مدل به صورت برداری‌شده[۱] پیاده‌سازی و صحت عملکرد آن‌ها از طریق بررسی گرادیان عددی تأیید گردید. مدل‌ها با استفاده از الگوریتم گرادیان کاهشی تصادفی (SGD) آموزش داده شده و هایپرپارامترهای کلیدی شامل نرخ یادگیری و قدرت تنظیم‌گری، با استفاده از یک مجموعه داده اعتبارسنجی بهینه‌سازی شدند. نتایج نهایی نشان داد که طبقه‌بند SVM با دستیابی به دقت ۳۴.۱٪ بر روی مجموعه داده آزمون، عملکرد بهتری نسبت به طبقه‌بند Softmax با دقت ۲۹.۶٪ داشت. این پژوهش ضمن نمایش موفقیت‌آمیز فرآیند پیاده‌سازی و آموزش طبقه‌بندهای خطی، محدودیت‌های ذاتی این مدل‌ها را در مواجهه با داده‌های تصویری پیچیده و غیرخطی برجسته می‌سازد.	

[۱] vectorized

۱- مقدمه

طبقه‌بندی تصاویر به عنوان یکی از مسائل بنیادین در حوزه بینایی کامپیوتر و یادگیری ماشین، به فرآیند تخصیص یک برچسب از میان مجموعه‌ای از کلاس‌های از پیش تعریف‌شده به یک تصویر ورودی اطلاق می‌شود. یکی از رویکردهای اصلی برای حل این مسئله، بهره‌گیری از طبقه‌بندهای خطی است که با یادگیری یک تابع امتیازدهی خطی، مرزهای تصمیم‌گیری بین کلاس‌های مختلف را مدل‌سازی می‌کنند. این مدل‌ها، علی‌رغم ساختار ساده، سنگ‌بنای الگوریتم‌های پیچیده‌تری همچون شبکه‌های عصبی را تشکیل می‌دهند. در این پروژه، از مجموعه داده معتبر CIFAR-۱۰ که شامل تصاویر رنگی در ابعاد کوچک و متعلق به ۱۰ کلاس متنوع است، به عنوان بستر آزمایش استفاده می‌شود.

هدف اصلی این پژوهش، پیاده‌سازی، آموزش، و ارزیابی مقایسه‌ای دو مورد از برجسته‌ترین طبقه‌بندهای خطی، یعنی ماشین بردار پشتیبان (SVM) و Softmax، می‌باشد. این فرآیند شامل پیاده‌سازی توابع هزینه و گرادیان تحلیلی آن‌ها به دو روش مبتنی بر حلقه[۱] و برداری‌شده[۲] است. صحت پیاده‌سازی گرادیان‌ها از طریق روش بررسی گرادیان عددی تأیید شده و مدل‌ها با استفاده از الگوریتم بهینه‌سازی گرادیان کاهشی تصادفی[۳] آموزش داده می‌شوند. همچنین، هایپرپارامترهای کلیدی هر مدل، از جمله نرخ یادگیری و قدرت تنظیم‌گری، با ارزیابی عملکرد بر روی یک مجموعه داده اعتبارسنجی تنظیم گردیده و در نهایت، عملکرد بهینه دو طبقه‌بند بر روی داده‌های آزمون با یکدیگر مقایسه خواهد شد.

۲- روش شناسی

در این بخش، مجموعه داده مورد استفاده و الگوریتم‌های به کار رفته برای به تفصیل شرح داده می‌شوند.

۲-۱ مجموعه داده CIFAR-۱۰

مجموعه داده CIFAR-۱۰ یکی از دیتاست‌های استاندارد و پرکاربرد در حوزه بینایی ماشین برای ارزیابی الگوریتم‌های طبقه‌بندی تصاویر است. این مجموعه داده شامل ۶۰,۰۰۰ تصویر رنگی با ابعاد کوچک 32×32 پیکسل می‌باشد. این تصاویر به ۱۰ کلاس مجزا تقسیم‌بندی شده‌اند که هر کلاس شامل ۶,۰۰۰ تصویر است. کلاس‌ها عبارتند از: هواپیما، اتومبیل، پرنده، گربه، گوزن، سگ، قورباغه، اسب، کشتی و کامیون. کل مجموعه داده به دو بخش اصلی تقسیم شده است: ۵۰,۰۰۰ تصویر برای مجموعه آموزشیو ۱۰,۰۰۰ تصویر برای مجموعه تست[۴] که امکان آموزش و ارزیابی مدل را فراهم می‌آورد.

پس از معرفی مشخصات کلی مجموعه داده CIFAR-۱۰، در شکل ۱ نمونه‌هایی از تصاویر موجود در هر یک از ده کلاس نمایش داده شده است. همانطور که مشاهده می‌شود، این تصاویر با ابعاد کوچک 32×32 پیکسل، چالش‌های قابل توجهی را برای طبقه‌بندی به همراه دارند. تنوع بالا در زاویه دید، شرایط نوری، پس‌زمینه و موقعیت اشیاء در تصاویر، پیچیدگی مسئله را افزایش می‌دهد.

---

[۱] naive
[۲] vectorized
[۳] Stochastic Gradient Descent

شکل ۱- نمونه تصاویر از مجموعه داده CIFAR-۱۰

۲-۲ پیش‌پردازش داده‌ها[۱]

پیش از شروع فرآیند آموزش، داده‌های تصویری طی چندین مرحله به فرمت مناسب برای طبقه‌بندهای خطی تبدیل شدند که در ادامه به آن‌ها اشاره شده است.

۲-۲-۱ تغییر ساختار[۲]

هر تصویر با ابعاد $32\times32\times3$ به یک بردار ویژگی تک‌بعدی با طول ۳۰۷۲ تبدیل شد. این کار با مسطح‌سازی آرایه سه‌بعدی پیکسل‌ها انجام گرفت.

۲-۲-۲ نرمال‌سازی [۳]

برای متمرکز کردن داده‌ها حول مبدأ، تصویر میانگین محاسبه و از تمام تصاویر مجموعه آموزشی، اعتبارسنجی و آزمون کسر گردید. لازم به ذکر است که تصویر میانگین فقط بر اساس داده‌های آموزشی محاسبه شد تا از نشت اطلاعات از مجموعه آزمون به فرآیند آموزش جلوگیری شود.

۲-۲-۳ افزودن بُعد بایاس[۴]

یک ستون با مقدار ثابت ۱ به انتهای هر بردار ویژگی اضافه شد. این تکنیک که به ترفند بایاس معروف است، به مدل اجازه می‌دهد تا پارامتر بایاس را به عنوان بخشی از ماتریس وزن‌ها یاد بگیرد و ابعاد نهایی هر بردار ویژگی به ۳۰۷۳ افزایش یافت.

۲-۳ پیاده‌سازی طبقه‌بندها[۵]

---

[۱] Data Preprocessing
[۲] Reshaping
[۳] Normalization
[۴] Bias Dimension
[۵] Classifier Implementation

۲-۳-۱ طبقه‌بند ماشین بردار پشتیبان[۱] (SVM)

ماشین بردار پشتیبان یک طبقه‌بند خطی قدرتمند است که هدف آن یافتن یک مرز تصمیم‌گیری بهینه بین کلاس‌های مختلف است. در حالت چندکلاسه، این مدل تلاش می‌کند تا برای هر نمونه داده، امتیاز کلاس صحیح نه تنها از امتیاز سایر کلاس‌ها بیشتر باشد، بلکه این برتری با یک حاشیه اطمینان (margin) مشخص حفظ شود. تابع هزینه SVM به گونه‌ای طراحی شده است که مدل را به دلیل نقض این حاشیه اطمینان جریمه کند.

تابع هزینه برای یک طبقه‌بند SVM چندکلاسه، که به آن Hinge Loss چندکلاسه نیز گفته می‌شود، به همراه عبارت تنظیم‌گری L2 به صورت زیر تعریف می‌شود:

$$L = \frac{1}{N}\sum_{i=1}^{N} L_i + \frac{1}{2}\lambda \sum_{k,l} W_{k,l}^2 \qquad (۲-۱)$$

که به‌صورت رابطه (۲-۲) نیز می‌توان نشان داد.

$$L = \underbrace{\frac{1}{N}\sum_{i=1}^{N}\sum_{j\neq y_i} \max\left(0, f\left(x_i;W\right)_j - f\left(x_i;W\right)_{y_i} + 1\right)}_{\text{Data Loss}} + \underbrace{\lambda R(W)}_{\text{Regularization Loss}} \qquad (۲-۲)$$

این تابع از دو بخش اصلی تشکیل شده است: خطای داده و خطای تنظیم‌گری.

- **بخش اول: خطای داده[۲]**

این بخش، که هسته اصلی تابع هزینه SVM را تشکیل می‌دهد، میزان عدم انطباق پیش‌بینی‌های مدل با برچسب‌های واقعی را اندازه‌گیری می‌کند و رابطه زیر نمایانگر آن است.

$$L_{data} = \frac{1}{N}\sum_{i=1}^{N} L_i \qquad (۲-۳)$$

که در آن به صورت زیر داریم:

$$L_i = \sum_{j\neq y_i} \max(0, \underbrace{f\left(x_i;W\right)_j}_{\text{score of wrong class}} - \underbrace{f\left(x_i;W\right)_{y_i}}_{\text{score of correct class}} + \underbrace{1}_{\text{safety margin}}) \qquad (۲-۴)$$

عبارت $f\left(x_i;W\right)_k$ نشان‌دهنده امتیازی است که مدل با پارامترهای W به نمونه ورودی x_i برای کلاس k تخصیص می‌دهد. همچنین $f\left(x_i;W\right)_{y_i}$ امتیاز تخصیص داده شده به کلاس صحیح (y_i) برای نمونه i.

در ادامه برای راحتی کار $f\left(x_i;W\right)_k$ و $f\left(x_i;W\right)_{y_i}$ را به ترتیب با s_j و s_{y_i} نشان خواهیم داد. در نتیجه فرمول را به‌صورت (۲-۵) داریم:

$$L_i = \sum_{j\neq y_i} \max(0, s_j - s_{y_i} + 1) \qquad (۲-۵)$$

- **بخش دوم: خطای تنظیم‌گری[۳]**

---

[۱] Support Vector Machine
[۲] Data Loss
[۳] Regularization Loss

این بخش به منظور کنترل پیچیدگی مدل و جلوگیری از پدیده بیش‌برازش به تابع هزینه اضافه می‌شود.

$$L_{reg} = \lambda R(W) \quad (۶-۲)$$

که در آن به صورت زیر داریم:

$$R(W) = \sum_{k,l} W_{k,l}^2 \quad (۷-۲)$$

که اجزاء این فرمول در حالت جمع‌بندی به‌صورت زیر بیان شده است:

- L: مقدار کل تابع هزینه که هدف بهینه‌سازی، کمینه کردن آن است.
- N: تعداد کل نمونه‌های آموزشی در یک بسته[۱].
- L_i: میزان خطای محاسبه شده برای نمونه آموزشی i-ام.
- s_j: امتیازی که مدل برای نمونه i-ام به کلاس ناصحیح j اختصاص می‌دهد. این امتیاز از حاصل ضرب $X_i \cdot W_j$ به‌دست می‌آید.
- s_{y_i}: امتیای که مدل برای نمونه i-ام به کلاس صحیح آن یعنی y_i اختصاص می‌دهد.
- 1: این مقدار ثابت، حاشیه اطمینان[۲] مطلوب است.
- $\max(0,\ldots)$: این عبارت هسته اصلی Hing Loss است. اگر امتیاز کلاس صحیح حداقل به اندازه ۱ واحد از امتیاز یک کلاس ناصحیح بیشتر شود ($s_{y_i} \geq s_j + 1$)، عبارت داخلی max منفی یا صفر شده و در نتیجه هزینه برای آن مقایسه صفر در نظر گرفته می‌شود. در غیر این صورت، مدل به اندازه فاصله‌ای که تا رسیدن به این حاشیه دارد، جریمه می‌شود.
- $\lambda \sum_{k,l} W_{k,l}^2$: این عبارت جریمه تنظیم‌گری $L2$ است. λ یک هایپرپارامتر است که قدرت این جریمه را کنترل می‌کند و $\sum_{k,l} W_{k,l}^2$ مجموعه مربعات تمام وزن‌های مدل است. این بخش از پیچیدگی بیش از حد مدل جلوگیری کرده و به بهبود قابلیت تعمیم آن کمک می‌کند. ضریب $\frac{1}{2}$ نیز یک قرارداد ریاضی برای ساده‌سازی فرمول گرادیان است.

✓ **گرادیان تابع هزینه SVM**

برای آموزش مدل و به‌روزرسانی وزن‌ها از طریق الگوریتم گرادیان کاهشی، از مشتق تابع هزینه نسبت به وزن‌ها (W) استفاده می‌شود. گرادیان L_i نسبت به ستون‌های ماتریس وزن به صورت زیر است:

---

[۱] batch

[۲] safety margin

گرادیان نسبت به وزن‌های کلاس ناصحیح (W_j):

$$\nabla w_j L_i = I\left(s_j - s_{y_i} + 1 > 0\right) \cdot X_i \tag{۸-۲}$$

گرادیان نسبت به وزن‌های کلاس صحیح (W_{y_i}):

$$\nabla w_{y_i} L_i = -\left(\sum_{j \neq y_i} I\left(s_j - s_{y_i} + 1 > 0\right)\right) \cdot X_i \tag{۹-۲}$$

در این روابط، $I(\ldots)$ تابع نشانگر[۱] است که اگر شرط داخل آن برقرار باشد مقدار ۱ و در غیر این صورت مقدار صفر دارد. به عبارت دیگر، برای هر کلاس ناصحیحی که حاشیه اطمینان را نقض کند، گرادیان باعث می‌شود وزن‌های آن کلاس از نمونه X_i فاصله بگیرند، در حالی که وزن‌های کلاس صحیح به اندازه مجموع این تخطی‌ها به سمت X_i نزدیک می‌شوند. گرادیان نهایی با افزودن گرادیان بخش تنظیم‌گری (λW) تکمیل می‌گردد.

۲-۳-۲ طبقه‌بند Softmax

طبقه‌بند Softmax، همانند ماشین بردار پشتیبان (SVM) ، یک مدل طبقه‌بندی خطی است که با استفاده از یک ماتریس وزن Wو یک بردار بایاس b، یک تابع امتیازدهی خطی $s = Wx + b$ را یاد می‌گیرد. تفاوت اصلی این دو مدل در نحوه تفسیر امتیازهای خروجی (s) است.

در طبقه‌بند Softmax، این امتیازات به عنوان لگاریتم غیرنرمال احتمال برای هر کلاس در نظر گرفته می‌شوند. سپس، با استفاده از تابع Softmax، این امتیازات خام به یک توزیع احتمال معتبر تبدیل می‌شوند. خروجی این تابع برای هر کلاس، عددی بین ۰ و ۱ است و مجموع این احتمالات برای تمام کلاس‌ها دقیقاً برابر با ۱ خواهد بود.

فرآیند آموزش مدل با هدف به حداکثر رساندن احتمال پیش‌بینی‌شده برای کلاس صحیح انجام می‌شود که این کار از طریق کمینه‌سازی تابع هزینه آنتروپی متقاطع صورت می‌گیرد.

۱. تبدیل امتیاز به احتمال با تابع Softmax

ابتدا، تابع Softmax امتیازات (s) را به احتمالات نرمال‌شده (P) تبدیل می‌کند:

$$P\left(y = k \mid x_i; W\right) = \frac{e^{s_k}}{\sum_j e^{s_j}} \tag{۱۰-۲}$$

اجزای این فرمول عبارتند از:

- s_k: امتیاز خام محاسبه شده توسط مدل برای کلاس k.

[۱] indicator function

- e^{s_k}: تابع نمایی که برای مثبت کردن تمام امتیازات به کار می‌رود.
- $\sum_j e^{s_j}$: مجموع امتیازات نمایی‌شده برای تمام کلاس‌ها، که به عنوان یک عامل نرمال‌سازی عمل می‌کند.

۲. محاسبه هزینه برای یک نمونه

تابع هزینه آنتروپی متقاطع برای یک نمونه‌ی منفرد (L_i)،به صورت لگاریتم منفیِ احتمال پیش‌بینی‌شده برای کلاس صحیح (y_i) تعریف می‌شود:

$$L_i = -\log\left(P\left(y = k \mid x_i; W\right)\right) = -\log\left(\sum_j e^{s_j}\right) \quad (۱۱-۲)$$

این تابع هزینه زمانی کمینه می‌شود که احتمال پیش‌بینی‌شده برای کلاس صحیح به ۱ نزدیک شود (در نتیجه، لگاریتم آن به صفر میل می‌کند). برعکس، اگر این احتمال به صفر نزدیک شود، مقدار خطا به سمت بی‌نهایت افزایش می‌یابد.

۳. تابع هزینه کل

در نهایت، تابع هزینه کل از میانگین خطای داده‌ها به اضافه‌ی یک عبارت تنظیم‌گری $L2$ به دست می‌آید:

$$L = \frac{1}{N}\sum_{i=1}^{N} L_i + \frac{1}{2}\lambda \sum_{k,l} W_{k,l}^2 \quad (۱۲-۲)$$

در این فرمول، N تعداد نمونه‌های آموزشی و λ ضریب قدرت تنظیم‌گری است.

✓ **گرادیان تابع هزینه Softmax**

یکی از ویژگی‌های برجسته تابع هزینه آنتروپی متقاطع، سادگی و زیبایی گرادیان آن است. گرادیان L_i نسبت به وزن‌های مربوط به کلاس k یعنی ستون k-ام ماتریس W به صورت زیر محاسبه می‌شود:

$$\nabla w_k L_i = \left(p_k - I\left(k = y_i\right)\right) \cdot x_i \quad (۱۳-۲)$$

اجزای این فرمول به شرح زیر هستند:

- p_k: احتمال پیش‌بینی شده توسط مدل برای کلاس k (خروجی تابع softmax)
- $I(\ldots)$: تابع نشانگرکه اگر شرط داخل آن برقرار باشد (یعنی k کلاس صحیح باشد) مقدار ۱ و در غیر این صورت مقدار صفر دارد.
- x_i: بردار ورودی نمونه i-ام.

این فرمول به زیبایی نشان می‌دهد که گرادیان، همان بردار ورودی x_i است که با خطای احتمال $\left(p_k - p_k^*\right)$ مقیاس شده است، جایی که p_k^* توزیع احتمال واقعی است (یک بردار one-hot که در اندیس کلاس صحیح مقدار ۱ و در بقیه جاها صفر است). این گرادیان ساده و شهودی، فرآیند بهینه‌سازی مدل را با روش‌هایی مانند گرادیان کاهشی[۱] بسیار کارآمد می‌کند.

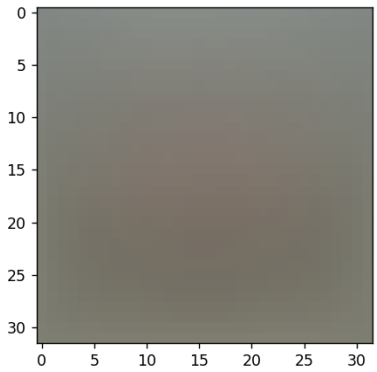
[۱] Gradient Descent

۳- نتایج

در این بخش، نتایج حاصل از پیاده‌سازی طبقه‌بند SVM و Softmax بررسی و در نهایت مقایسه می‌شوند.

۳-۱- نتایج طبقه‌بند SVM

پس از اجرای مراحل پیش‌پردازش، که شامل تغییر ساختار ابعادی، نرمال‌سازی با تصویر میانگین و افزودن بُعد بایاس بود، مجموعه داده‌های نهایی با ابعاد (۴۹۰۰۰,۳۰۷۳) برای آموزش، (۱۰۰۰,۳۰۷۲) برای اعتبارسنجی و آزمون، و (۵۰۰,۳۰۷۲) برای توسعه آماده شدند. تصویر میانگین به درستی بر اساس داده‌های آموزشی محاسبه شد که در شکل ۲ نشان داده شده است.



شکل ۲- تصویر میانگین

به منظور صحت‌سنجی اولیه پیاده‌سازی تابع هزینه، مقدار خطای اولیه برای مدل آموزش ندیده محاسبه گردید. مقدار به دست آمده برابر با ۸.۷۱۱۶۶۱ می‌باشد.

- ✓ **بررسی گرادیان[۱]**

به منظور اطمینان از صحت محاسبات گرادیان تحلیلی (فرمول مشتق)، نتایج آن با یک گرادیان عددی (که از طریق تقریب محاسبه می‌شود) در ابعاد مختلف مقایسه گردید. طبق نتایج به‌دست آمده، خطای نسبی (relative error) بین دو روش برای اکثر ابعاد بسیار ناچیز و در مرتبه ۱e-۱۱ بود. این تطابق بالا، صحت پیاده‌سازی گرادیان تحلیلی ما را تأیید می‌کند.

- ✓ **مقایسه عملکرد نسخه‌های Naive و Vectorized**

در ادامه، پیاده‌سازی برداری‌شده (vectorized) با پیاده‌سازی مبتنی بر حلقه (naive) مقایسه شد تا از صحت و کارایی آن اطمینان حاصل شود. هر دو نسخه، مقدار خطای اولیه کاملاً یکسانی برابر با ۸.۷۱۱۶۶۱ را محاسبه کردند. این مقدار با مقدار تئوری مورد انتظار (حدود ۹.۰) برای یک مدل آموزش ندیده با ۱۰ کلاس مطابقت دارد و صحت هر دو پیاده‌سازی را نشان می‌دهد.

[۱] Gradient Check

در نتیجه این تطابق، Loss difference و Gradient difference بین دو نسخه دقیقاً صفر بود. این در حالی است که نسخه vectorized بهبود سرعت قابل توجهی (تقریباً ۱۰ برابر) را به نمایش گذاشت و محاسبات را در ۰.۰۱ ثانیه در مقابل ۰.۱ ثانیه نسخه naive انجام داد.

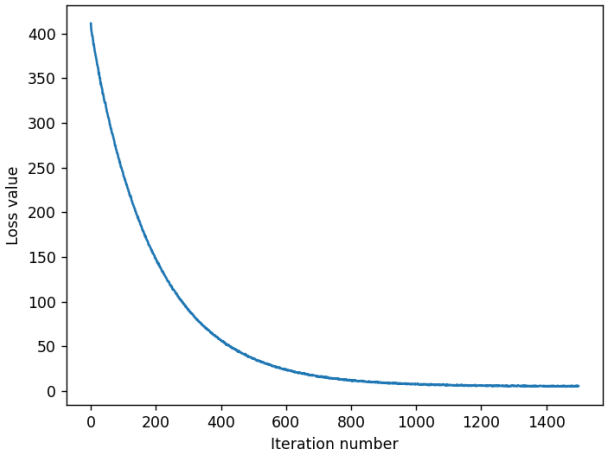
علاوه بر صحت، نسخه vectorized بهبود چشمگیری در سرعت اجرا نشان داد و محاسبات را تقریباً ۱۰ برابر سریع‌تر از نسخه naive انجام داد (به عنوان مثال، ۰.۱ ثانیه در مقابل ۰.۰۱ ثانیه). این نتایج اهمیت بالای پیاده‌سازی برداری‌شده در محاسبات یادگیری ماشین را نشان می‌دهد.

```
Naive loss: ۸,۷۱۱۶۶۱e+۰۰ computed in ۰,۱۰۴۸۳۶s
Vectorized loss: ۸,۷۱۱۶۶۱e+۰۰ computed in ۰,۰۱۰۰۲۷s
```

✓ **نتایج آموزش اولیه مدل**

پیش از انجام جستجوی کامل هایپرپارامترها، یک مدل SVM با مجموعه‌ای از پارامترهای ثابت (نرخ یادگیری ۱e-۷ و قدرت تنظیم‌گری ۲.۵e۴) به تعداد ۱۵۰۰ تکرار آموزش داده شد تا عملکرد پایه و روند همگرایی مدل بررسی شود.

خروجی فرآیند آموزش، کاهش موفقیت‌آمیز و پیوسته هزینه (Loss) را نشان داد. مقدار هزینه از ۴۰۸.۴۹ در تکرار اول به ۵.۶۱ در تکرار پایانی کاهش یافت. این روند همگرایی به وضوح در شکل ۳ قابل مشاهده است.



شکل ۳- منحنی کاهش هزینه برای مدل SVM

پس از اتمام آموزش، مدل بر روی مجموعه داده‌های آموزشی و اعتبارسنجی ارزیابی شد و به نتایج زیر دست یافت:

دقت آموزش[۱]: ٪۳۷.۷

دقت اعتبارسنجی[۲]: ٪۳۸.۷

✓ **آموزش و تنظیم هایپرپارامترها**

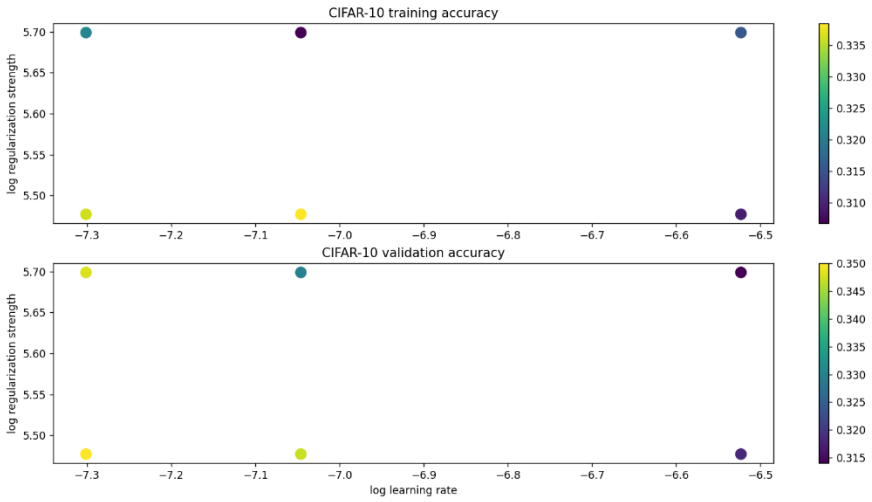
[۱] Training Accuracy
[۲] Validation Accuracy

به منظور یافتن بهترین هایپرپارامترها، یک جستجوی شبکه‌ای[1] بر روی مقادیر مختلف نرخ یادگیری و قدرت تنظیم‌گری انجام شد. نتایج کامل این جستجو در جدول ۱ خلاصه شده است.

جدول ۱- نتایج حاصل از تنظیم هایپرپارامتر برای طبقه‌بند SVM

نرخ یادگیری (Learning Rate)	قدرت تنظیم‌گری (Regularization)	دقت آموزش (Train Accuracy)	دقت اعتبارسنجی (Validation Accuracy)
۵.۰۰۰۰۰۰e-۰۸	۳.۰۰۰۰۰۰e+۰۵	۰.۳۳۶۶۵۳	۰.۳۵۴۰۰۰
۵.۰۰۰۰۰۰e-۰۸	۵.۰۰۰۰۰۰e+۰۵	۰.۳۲۶۰۸۲	۰.۳۴۷۰۰۰
۹.۰۰۰۰۰۰e-۰۸	۳.۰۰۰۰۰۰e+۰۵	۰.۳۳۰۳۲۷	۰.۳۴۵۰۰۰
۹.۰۰۰۰۰۰e-۰۸	۵.۰۰۰۰۰۰e+۰۵	۰.۳۱۳۲۰۴	۰.۳۲۰۰۰۰
۳.۰۰۰۰۰۰e-۰۷	۳.۰۰۰۰۰۰e+۰۵	۰.۳۱۰۵۷۱	۰.۳۲۰۰۰۰
۳.۰۰۰۰۰۰e-۰۷	۳.۰۰۰۰۰۰e-۰۷	۰.۲۷۷۰۲۰	۰.۲۷۸۰۰۰

همانطور که در جدول مشاهده می‌شود، بهترین دقت بر روی مجموعه داده اعتبارسنجی (Validation Accuracy) برابر با ۰.۳۵۴۰۰۰ بود. این نتیجه با نرخ یادگیری ۵e+۸ و قدرت تنظیم‌گری ۵e+۳ به دست آمد و این مدل به عنوان مدل بهینه انتخاب گردید. نتایج این جستجو به صورت بصری نیز در شکل ۴ نمایش داده شده است.



شکل ۴- نتایج اعتبارسنجی متقابل برای هایپرپارامترهای مختلف

[1] Grid Search

در نهایت، مدل بهینه انتخاب‌شده بر روی مجموعه داده آزمون (test set) که تا این لحظه در هیچ‌یک از مراحل آموزش یا تنظیم دخالتی نداشته، ارزیابی شد. دقت نهایی طبقه‌بند SVM در این آزمون برابر با ۰.۳۴۱۰۰۰ محاسبه گردید.

وزن‌های یادگرفته‌شده توسط این مدل در شکل ۵ بصری‌سازی شده‌اند. این تصاویر، الگوهای انتزاعی را نشان می‌دهند که مدل برای هر کلاس آموخته است. در این الگوها می‌توان برخی ویژگی‌های کلی مانند رنگ غالب (سبز برای قورباغه، آبی برای هواپیما) را تشخیص داد که نشان‌دهنده یادگیری ویژگی‌های بصری ساده توسط مدل است.

شکل ۵- بصری‌سازی ماتریس وزن‌های بهینه مدل SVM برای هر یک از ۱۰ کلاس.

۳-۲- نتایج طبقه‌بند Softmax

در این بخش، نتایج حاصل از صحت‌سنجی، فرآیند آموزش و تنظیم هایپرپارامترها، و ارزیابی نهایی طبقه‌بند Softmax ارائه و تحلیل می‌گردد.

عملکرد پیاده‌سازی تابع هزینه و گرادیان softmax پیش از آموزش اصلی مورد ارزیابی قرار گرفت.

- ✓ بررسی اولیه مقدار خطا (Sanity Check):

مقدار خطای اولیه برای مدل آموزش ندیده ۲.۳۳۶ محاسبه شد که با مقدار تئوری مورد انتظار $-\log(0.1) \approx 2.302$ تطابق بسیار خوبی دارد.

- ✓ بررسی گرادیان (Gradient Check):

صحت گرادیان تحلیلی با مقایسه با گرادیان عددی تأیید شد و خطای نسبی (relative error) بسیار کوچکی (در مرتبه ۸-e۱) مشاهده گردید.

مقایسه نسخه‌ها: نتایج نشان داد که پیاده‌سازی‌های naive و vectorized خروجی‌های کاملاً یکسانی تولید می‌کنند (Loss s difference: ۰.۰۰۰۰۰۰ و Gradient difference: ۰.۰۰۰۰۰۰)، که صحت نسخه بهینه vectorized را که در آموزش اصلی استفاده شد، تضمین می‌کند.

✓ **نتایج آموزش اولیه مدل**

برای یافتن بهترین هایپرپارامترها، یک جستجوی شبکه‌ای بر روی مقادیر مختلف نرخ یادگیری و قدرت تنظیم‌گری انجام شد. خروجی‌های دقیق هر مرحله از آموزش (iteration) نشان‌دهنده همگرایی موفق مدل‌ها بود. نتایج نهایی این جستجو در جدول ۲ خلاصه شده است.

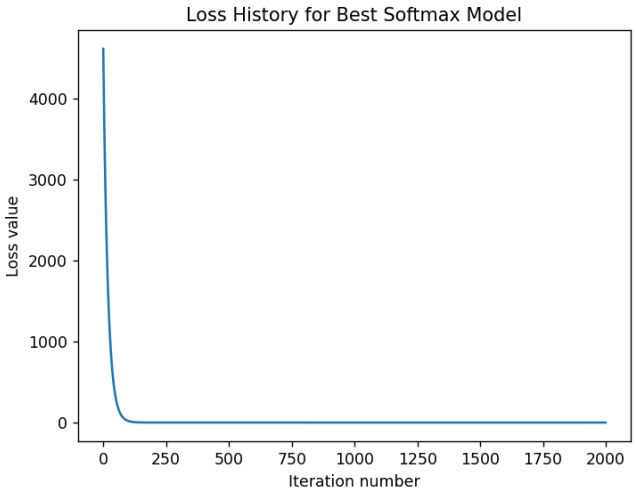
جدول ۲- نتایج حاصل از تنظیم هایپرپارامتر برای طبقه‌بند Softmax

نرخ یادگیری (Learning Rate)	قدرت تنظیم‌گری (Regularization)	دقت آموزش (Train Accuracy)	دقت اعتبارسنجی (Validation Accuracy)
۵.۰۰۰۰۰۰e-۰۸	۳.۰۰۰۰۰۰e+۰۵	۰.۲۷۳۹۱۸	۰.۲۸۷۰۰۰
۵.۰۰۰۰۰۰e-۰۸	۵.۰۰۰۰۰۰e+۰۵	۰.۲۶۱۸۷۸	۰.۲۷۱۰۰۰
۹.۰۰۰۰۰۰e-۰۸	۳.۰۰۰۰۰۰e+۰۵	۰.۲۷۵۸۹۸	۰.۲۸۴۰۰۰
۹.۰۰۰۰۰۰e-۰۸	۵.۰۰۰۰۰۰e+۰۵	۰.۲۷۱۸۵۷	۰.۲۷۵۰۰۰
۳.۰۰۰۰۰۰e-۰۷	۳.۰۰۰۰۰۰e+۰۵	۰.۲۵۵۶۳۳	۰.۲۷۰۰۰۰
۳.۰۰۰۰۰۰e-۰۷	۳.۰۰۰۰۰۰e-۰۷	۰.۲۷۰۴۰۸	۰.۲۸۶۰۰۰

بر اساس نتایج جدول بالا، بهترین دقت اعتبارسنجی به میزان ۰.۲۸۷۰۰۰ با نرخ یادگیری ۵e-۸ و قدرت تنظیم‌گری ۳e+۵ به دست آمد.

مدل بهینه انتخاب‌شده از مرحله قبل، بر روی مجموعه داده آزمون (test set) ارزیابی گردید و به دقت نهایی ۰.۲۹۶۰۰۰ دست یافت.

برای یافتن بهترین هایپرپارامترها، یک جستجوی شبکه‌ای بر روی مقادیر مختلف نرخ یادگیری و قدرت تنظیم‌گری انجام شد. همانطور که در شکل ۶ مشاهده می‌شود، منحنی هزینه برای بهترین مدل، یک روند کاهشی سریع و همگرایی موفق را نشان می‌دهد.



شکل ۳- منحنی کاهش هزینه برای مدل Softmax

وزن‌های یادگرفته‌شده توسط این مدل در شکل ۴ بصری‌سازی شده‌اند. این تصاویر نمایانگر الگوهای انتزاعی هستند که مدل برای هر کلاس آموخته است. همانند مدل Softmax، در این الگوها نیز می‌توان برخی ویژگی‌های کلی مانند شکل دو رنگ "اتومبیل" یا رنگ آبی "هواپیما" را تشخیص داد که نشان‌دهنده یادگیری ویژگی‌های بصری ساده توسط مدل است.

شکل ۵- بصری‌سازی ماتریس وزن‌های بهینه مدل Softmax برای هر یک از ۱۰ کلاس.

نتیجه گیری

در این پروژه، دو طبقه‌بند خطی بنیادی، ماشین بردار پشتیبان (SVM) وSoftmax ، به منظور طبقه‌بندی تصاویر مجموعه داده CIFAR-۱۰ پیاده‌سازی و ارزیابی شدند. پس از پیاده‌سازی توابع هزینه و گرادیان به صورت برداری‌شده و تأیید صحت عملکرد آن‌ها، مدل‌ها با استفاده از الگوریتم گرادیان کاهشی تصادفی و هایپرپارامترهای بهینه‌سازی‌شده، آموزش دیدند. نتایج نهایی نشان داد که طبقه‌بند SVM با دستیابی به

دقت ۳۴.۱٪ بر روی مجموعه داده آزمون، عملکرد بهتری نسبت به طبقه‌بند Softmax با دقت ۲۹.۶٪ از خود به نمایش گذاشت.

این یافته‌ها ضمن تأیید کارایی طبقه‌بندهای خطی در یادگیری الگوهای پایه از داده‌های تصویری (با عملکردی به مراتب بهتر از حدس تصادفی)، بر محدودیت ذاتی این مدل‌ها در مواجهه با ماهیت پیچیده و غیرخطی چنین مسائلی تأکید دارد. برای دستیابی به دقت بالاتر، استفاده از مدل‌های با ظرفیت بیشتر و ساختار غیرخطی، مانند شبکه‌های عصبی کانولوشنی (CNNs)، به عنوان گام بعدی پژوهش پیشنهاد می‌گردد.

Surname: Mirzajani	**Name:** Amirreza	
Title: Implementation and Comparison of SVM and Softmax Linear Classifiers on the CIFAR-۱۰ Dataset		
Supervisor: Dr. M. Shoaran		
Degree: Master of Science	**Major:** Mechanical Engineering	**Field:** Applied Design
University: University of Tabriz	**Faculty:** Mechanical Engineering	
Pages: ۱٦		
Keywords: Image Classification, Support Vector Machine (SVM), Softmax Classifier, Stochastic Gradient Descent (SGD), CIFAR-۱۰		

Abstract: This report addresses the implementation and comparative evaluation of two fundamental linear classifiers, the Support Vector Machine (SVM) and Softmax, for the problem of image classification on the CIFAR-۱۰ dataset. In this project, the loss functions and analytical gradients for both models were implemented in a vectorized manner, and their correctness was verified through numerical gradient checking. The models were trained using the Stochastic Gradient Descent (SGD) algorithm, and key hyperparameters, including learning rate and regularization strength, were optimized using a validation dataset. The final results showed that the SVM classifier, by achieving an accuracy of ۳٤,۱٪ on the test set, performed better than the Softmax classifier with an accuracy of ۲۹,٦٪. This research, while successfully demonstrating the implementation and training process of linear classifiers, highlights the inherent limitations of these models when faced with complex and non-linear image data.

University of Tabriz

Faculty of Electrical and Computer Engineering
Department of Computer Engineering

Title:
Implementation and Comparison of SVM and Softmax Linear Classifiers on the CIFAR-۱۰ Dataset

Professor:
Dr. M. Shoaran

Researcher:
Amirreza Mirzajani

August ۲۰۲۵