

به نام خدا
درس مبانی یادگیری عمیق
تمرین سری سوم
استاد درس : دکتر مرضیه داوودآبادی
دستیاران : سحر سرکار، فائزه صادقی، پریسا ظفری،
شایان موسوی نیا
دانشگاه علم و صنعت ایران، دانشکده مهندسی کامپیوتر
نیمسال اول تحصیلی ۱۴۰۲ - ۱۴۰۳



مهلت تحویل : ۱۴۰۲/۰۹/۰۳

لطفاً به نکات موجود در سند قوانین انجام و تحویل تمرین ها دقت فرمایید.

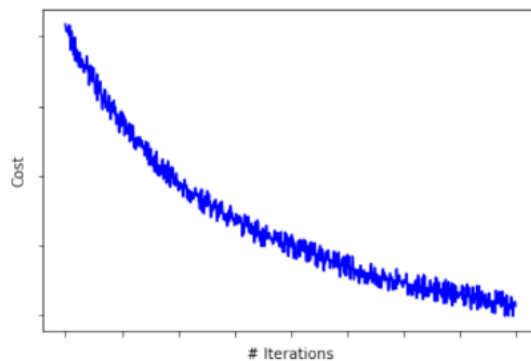
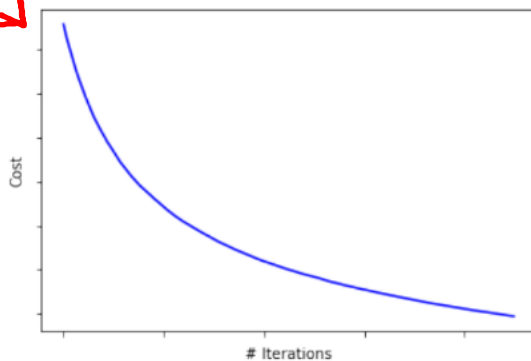
۱. سوالات زیر را با توجه به مبحث بهینه‌سازها پاسخ دهید (۱۵ نمره).

الف) استفاده از نرخ یادگیری بسیار بالا چه مشکلاتی را ایجاد می‌کند؟ چگونه می‌توان این مشکلات را تشخیص داد؟ به جواب خوب همگرا نشود حتی و اگر شود

ب) استفاده از نرخ یادگیری بسیار پایین چه مشکلاتی را ایجاد می‌کند؟ چگونه می‌توان این مشکلات را تشخیص داد؟ دیر همگرا شود

پ) نقطه زینی چیست؟ دو الگوریتم *Adam* و *SGD* را در برخورد با این نقاط مقایسه کنید. مزایا و معایب هر کدام را بنویسید.

ت) شکل زیر کاهش هزینه را (با افزایش تکرارها) زمانی که از دو الگوریتم بهینه‌سازی مختلف برای آموزش استفاده می‌شود، نشان می‌دهد. کدام یک از این نمودارها با استفاده از شیب نزولی دسته‌ای به عنوان الگوریتم بهینه‌سازی و کدام یک مربوط به استفاده از شیب نزولی دسته‌ای کوچک^۱ است؟ توضیح دهید.



¹Mini-batch

۲. ورودی یک لایه هم‌گشتی^۲ (X) با ابعاد سه در سه را در نظر بگیرید. فیلتر F با ابعاد دو در دو روی ورودی X اعمال شده است. روی خروجی این لایه هم‌گشتی، یک لایه ادغام میانگین سراسری^۳ اعمال می‌شود که خروجی نهایی یک عدد خواهد شد. باتوجه به اینکه گرادیان تابع ضرر نسبت به این خروجی نهایی که یک عدد است، ۱ می‌شود، با استفاده از الگوریتم پس‌انتشار^۴ خطا، گرادیان‌های این لایه‌ی هم‌گشتی را به دست آورید (برای محاسبه می‌توانید از فرمول‌های این مرجع استفاده کنید) (۱۵ نمره).

۳	۴	۵
۲	۱	-۳
۴	-۲	۰

X

۲	۰
-۳	۱

F

۳. لطفاً به سوالات زیر پاسخ دهید.

الف) تعداد پارامترهای شبکه زیر را محاسبه کنید (محاسبات خود را برای هر لایه به‌طور کامل یادداشت کنید) (۱۰ نمره).

```
model = Sequential()
model.add(Input(shape=(500, 7)))
model.add(Conv1D(filters=16, kernel_size=3, activation="relu"))
model.add(MaxPool1D())
model.add(Conv1D(filters=32, kernel_size=5, activation="relu"))
model.add(MaxPool1D())
model.add(Conv1D(filters=64, kernel_size=5, activation="relu"))
model.add(MaxPool1D())
model.add(Flatten())
model.add(Dense(units=128, activation="relu"))
model.add(Dense(units=5, activation="softmax"))
```

^۲Convolutional layer

^۳Global average pooling (GAP)

^۴Backpropagation

✓ (ب) لایه هم‌گشتی دو بعدی ($Conv2D$) و سه‌بعدی ($Conv3D$) را مقایسه کرده و کاربرد لایه $Conv3D$ را ذکر کنید (۱۰ نمره امتیازی).

✓ ۴. در این سوال قصد داریم دادگان تومور مغزی را دسته‌بندی کنیم. لطفا موارد خواسته شده را در نوت‌بوک $DL - Q4.ipynb$ را پیاده‌سازی کنید. اهداف این سوال به ترتیب موارد زیر می‌باشد (۲۰ نمره):

- خواندن تصاویر دادگان که در یک مسیر مشخص قرار داده شده‌اند و استفاده از آن‌ها برای آموزش و ارزیابی شبکه.

- ساختن یک شبکه عصبی هم‌گشتی با معماری یکسان با $Sequential API$ و $Functional API$ $convolutional$

✓ ۵. با در نظر گرفتن یک کاربرد واقعی مانند دسته‌بندی تصاویر، یک مثال عملی ارائه دهید که ویژگی‌های منحصر به فرد لایه‌های هم‌گشتی (مزایا) منجر به عملکرد مفیدی می‌شود؟ به طور عکس، آیا می‌توانید یک حالت تصور کنید که ویژگی‌های لایه هم‌گشتی چالش‌هایی (معایب) را ایجاد می‌کنند؟ توضیح دهید که چرا و تأثیر آن بر کارایی مدل چیست (۱۰ نمره).

✓ ۶. فیلترهای 1×1 در شبکه‌های عصبی هم‌گشتی ویژگی‌های منحصر به فردی دارند که عملکرد شبکه را تحت تأثیر قرار می‌دهند. در این سوال قصد داریم تأثیر فیلترهای 1×1 را بررسی کنیم (۱۵ نمره). الف) هدف استفاده از فیلترهای 1×1 در شبکه‌های عصبی هم‌گشتی چیست و چگونه به کاهش تعداد نقشه‌های ویژگی در حین حفظ ویژگی‌های مهم کمک می‌کنند؟

ب) پس از اعمال فیلتر 1×1 ، نقشه ویژگی چه اطلاعاتی ارائه می‌دهد؟
پ) این نقشه ویژگی چگونه از تصویر اصلی یا فیلترهای دیگر با اندازه‌های مختلف متفاوت است؟
ت) در چه مدل‌هایی از این نوع فیلتر استفاده شده است؟
ث) آیا حالت‌هایی وجود دارد که استفاده از فیلترهای 1×1 ممکن است مفید نباشد؟ اگر بله، دلیل چیست؟
ج) یک مدل هم‌گشتی ساده با فیلتر 1×1 پیاده‌سازی کنید و تأثیر آن را روی یک ورودی دلخواه با مقایسه اندازه ورودی و خروجی مدل نشان دهید.

۷. یک شبکه عصبی هم‌گشتی که شامل یک ماژول $Inception$ برای طبقه‌بندی تصویر با استفاده از مجموعه داده $CIFAR - 10$ است، پیاده‌سازی کنید. در اجرای خود، حتماً به موارد زیر توجه کنید (۱۵ نمره):

- ساختار و هدف ماژول $Inception$ را در این شبکه توضیح دهید.

- چگونه اندازه پارامتر گام^۵ در لایه‌های هم‌گشتی بر ابعاد فضایی نگاشت^۶ ویژگی‌ها تأثیر می‌گذارد؟
اینجا 0 گرفته پس تغییر نمی‌کنه ابعاد
- ویژگی‌های کلیدی و عملیات لایه‌های هم‌گشتی مورد استفاده در شبکه خود و اهمیت آنها در استخراج ویژگی را شرح دهید.
- انجام پیاده‌سازی صرفاً با استفاده از مدل‌های آماده در کتابخانه‌های *keras* و نظیر آنها بدون مشکل است.
- هدف این تمرین رسیدن به دقت بالای ۸۰ درصد در آموزش است. برای رسیدن به این هدف از همه روش‌هایی که در درس آموخته‌اید استفاده کنید (۱۰ نمره امتیازی).

⁵Stride

⁶Mapping