

به نام خدا



دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده برق و کامپیوتر



شبکه های عصبی مصنوعی و یادگیری عمیق

مینی پروژه 1

اردیبهشت ماه 1400

فهرست سوالات

- سوال 1 – مفاهیم تئوری 3
- سوال ۲ CNN 4
- سوال 3 Data Augmentation 6
- سوال 4 Transfer Learning 7

سوال 1 – مفاهیم تئوری

1. همان طور که می‌دانیم به منظور بهینه‌سازی شبکه‌های عصبی از روش گرادیان کاهشی¹ استفاده می‌شود.

- ا. در مورد مشکلاتی که استفاده از این روش می‌تواند به وجود آورد توضیح دهید.
ب. در مورد عملکرد هر یک از روش‌های زیر صحبت کنید و بیان نمایید که آیا می‌توانند مشکلات روش گرادیان کاهشی را حل نمایند یا خیر؟

- روش Momentum برای گرادیان کاهشی

- روش Adam

- روش AdaDelta

2. بیش برآزش² در طبقه‌بندی را توضیح دهید. برای حل این مشکل از سه روش norm, drop out, penalty و early stopping استفاده می‌شود، در مورد هر یک از روش‌های بیان شده توضیح دهید و بیان کنید که روش‌های مذکور چگونه می‌توانند این مشکل را حل نمایند.

3. می‌دانیم که شبکه‌های عصبی تنها با دو لایه مخفی خاصیت general function approximator بودن خود را حفظ می‌نمایند. با این وجود، چرا گاهی بیش از دو لایه مخفی برای شبکه‌ها استفاده می‌شود؟

4. در طراحی شبکه‌های عصبی عمیق چرا در انتخاب تعداد لایه‌ها یا تعداد پارامتر هر لایه محدود هستیم؟

5. در یک شبکه عصبی اگر تمامی وزن‌های اولیه یکسان تعیین شوند چه مشکلی ایجاد می‌شود؟

6. تفاوت vanishing gradient و exploding gradient را بیان نمایید.

¹ Gradient Descent
² Overfitting

سوال ۲ CNN

یکی از معروفترین داده‌هایی که برای سنجش صحت الگوریتم‌های طبقه‌بندی مورد استفاده قرار می‌گیرد، داده‌ی CIFAR10 می‌باشد که دارای 60000 تصویر در 10 کلاس مختلف است. برای کسب اطلاعات و دانلود داده به <https://www.cs.toronto.edu/~kriz/cifar.html> مراجعه کنید

در این تمرین شما باید شبکه CNN سه لایه‌ای (هر لایه شامل چند فیلتر و یک Pool است) برای طبقه‌بندی داده‌های مذکور طراحی کنید و موارد زیر را بررسی نمایید.

1. مشخصات شبکه عصبی خود را بنویسید.
 - اندازه پنجره‌های convolution، اندازه stride و تعداد filterها در هر لایه
 - توابع فعال‌سازی مورد استفاده
 - اندازه لایه‌های fully connected
 - تابع loss و روش بهینه‌سازی مورد استفاده
 - اندازه mini-batch مورد استفاده
2. شبکه عصبی‌ای که طراحی کرده‌اید را اجرا و نمودار دقت را بر حسب اپیاک برای داده‌های یادگیری و تست رسم کنید.
 - بهترین دقت به دست آمده را برای داده‌های آموزشی و تست چه مقدار است و بعد از چند اپیاک به دست می‌آید؟
3. کارایی شبکه طراحی شده را با 0، 1 و 2 لایه مخفی بررسی نمایید و نمودار دقت را برای داده‌های یادگیری و تست رسم کنید.
4. از توابع فعال‌ساز مختلف (sigmoid، tanh و RELU) استفاده کنید. گرافی با سه نمودار برای سه تابع فعال‌ساز تولید کنید. بعد Y گراف نشان‌دهنده‌ی دقت طبقه‌بندی بر روی داده‌های تست و بعد X نشان‌دهنده تعداد اپیاک‌ها بر روی داده‌های آموزشی می‌باشد. نمودارها را برای 10 اپیاک رسم کنید (اگرچه 10 اپیاک برای همگرا شدن کافی نیست). نتیجه را تحلیل کرده و بیان نمایید که کدام تابع فعال‌ساز بهتر است؟ چرا؟
5. از روش‌های بهینه‌سازی gradient descent و adam استفاده کنید. گرافی با دو نمودار ایجاد کنید. بعد Y گراف نشان‌دهنده‌ی دقت طبقه‌بندی بر روی داده‌های تست و بعد X نشان‌دهنده تعداد اپیاک‌ها بر روی داده‌های آموزشی می‌باشد. نمودارها را برای 10 اپیاک رسم کنید نتیجه را تحلیل نمایید.

6. حجم داده‌ها را کاهش دهید، به‌طوری که در هر کلاس تنها 600 نمونه وجود داشته باشد. حال شبکه را بر روی این داده‌ها اجرا کنید و نمودار دقت بر حسب ایپاک رسم کنید و نتیجه را تحلیل کنید.
7. به جای دو عمل کانولوشن در هر لایه یک عمل کانولوشن با کرنل بزرگتر قرار دهید و عملکرد را بررسی کنید. جایگزین کردن کانولوشن با کرنل بزرگ با چند کانولوشن با کرنل کوچک چه امتیازاتی می‌تواند داشته باشد؟
8. تاثیر افزودن dropout بعد از هر تابع فعالساز در شبکه‌ی خود را بررسی کنید. سعی کنید درصد بهینه dropout را برای شبکه‌ی خود پیدا کنید. دلیل بهبود در عملکرد شبکه چیست؟

سوال 3 Data Augmentation

از جمله مشکلاتی که در بسیاری از موارد وجود دارد، محدود بودن دیتاست و یا نامتوازن بودن تعداد داده ها در هر کلاس است. این مشکلات سبب ایجاد اختلال در عملکرد شبکه می شود. روش های گوناگونی برای حل این مشکلات وجود دارد. یکی از این روش ها Data Augmentation است که در این سؤال با آن آشنا می شود.

1. در مورد Data Augmentation تحقیق کنید و توضیح دهید که نحوه عملکرد این روش چگونه است و تبدیل هایی که در آن استفاده می شود را شرح دهید. آیا از این روش برای داده های تست استفاده می شود؟ علت را شرح دهید.
2. با استفاده از چند تبدیل، از تصویر test.jpg، 10 نمونه مصنوعی ایجاد کنید و آن ها را به همراه تصویر اصلی نمایش دهید.
3. همانند سؤال 2 از این تمرین، دیتاست Cifar10 را بارگذاری کنید سپس از کلاس های گربه و سگ در داده های آموزش، 4500 تصویر را حذف کنید. با این کار 90 درصد از تصاویر این دو کلاس حذف می شوند و تعداد تصاویر آموزش به 41000 می رسد. پس از حذف تصاویر، داده های جدید را با بهترین شبکه بدست آمده در سؤال 2 آموزش دهید. پس از آموزش مدل، ماتریس آشفتگی¹ آن را برای داده های تست نمایش دهید. چه اختلالی در عملکرد شبکه رخ داده است؟
4. برای حل مشکل پیش آمده، با استفاده از Data Augmentation و ترکیبی از تبدیل های مختلف آن، تعداد تصاویر در کلاس های سگ و گربه را به 5000 برسانید تا در مجموع 50000 تصویر آموزش وجود داشته باشد. سپس داده های جدید را با بهترین شبکه به دست آمده در سؤال 2 آموزش و پس از آن ماتریس آشفتگی داده های تست را برای مدل آموزش دیده نمایش دهید. چه بهبودی نسبت به ماتریس آشفتگی بخش 3 ایجاد شده است؟ تفاوت ها با بخش 3 را شرح دهید.

نکته: برای Data Augmentation در این سؤال تنها مجاز به استفاده از ImageDataGenerator از کتابخانه Keras هستید.

¹ Confusion Matrix

سوال 4 Transfer Learning

تا کنون مدل‌های معروفی از شبکه‌های عصبی کانولوشنی به وجود آمده‌اند که در این سوال تصمیم بر آن است که یک مدل مطالعه و به کمک Trasfer Learning پیاده‌سازی شود.

برای یافتن مدل بایستی باقی مانده‌ی مجموع ارقام آخر شماره دانشجویی اعضای گروه را بر 10 حساب کنید و بر اساس آن مدل را انتخاب نمایید (بدیهتا در صورت انجام پروژه به صورت فردی رقم آخر بیانگر شماره مدل می‌باشد).

مدل	رقم آخر شماره دانشجویی
EfficientNet	0
MobileNet	1
Inception	2
ResNet	3
DenseNet	4
ShuffleNet	5
NASNet	6
VGG	7
Xception	8
AlexNet	9

1. مختصرا موارد زیر را در مورد مدل خود توضیح دهید:

- معماری شبکه
- کاربرد و مزایا
- ساینز تصویر ورودی
- آیا تصویر ورودی نیاز به پیش‌پردازش اولیه دارد؟ چرا؟ اگر بله چه پیش‌پردازشی؟
- ساینز خروجی و معنای آن

2. Transfer Learning به چه معناست و چه زمانی استفاده می‌شود؟

3. شبکه انتخابی را به کمک Transfer Learning پیاده نمایید.

4. لیست اشیایی که شبکه شما می‌تواند شناسایی کند چیست؟

5. ابتدا یک عکس رنگی با کیفیت مناسب از محیط اطراف خود بگیرید که منطبقا این عکس باید شامل اشیای قابل تشخیص توسط شبکه شما باشد، سپس در صورت لزوم عکس را پیش پردازش کنید و به شبکه‌ای ک در بخش 3 پیاده‌سازی کردید داده و 3 شی را به ترتیب با بیشترین احتمال و همراه با مقدار احتمالی‌شان در خروجی نشان دهید.

در گزارش بخش 5 تصویر را نیز به همراه نتیجه در فایل گزارش آورد و همچنین تصویر به همراه کد به گونه‌ای که قابل اجرا باشد، ارسال شود.

نکات:

- مهلت تحویل این مینی پروژه 24 اردیبهشت است.
- گزارش را در قالب تهیه شده که روی صفحه درس در Elearn بارگذاری شده، بنویسید.
- گزارش شما در فرآیند تصحیح از اهمیت ویژه‌ای برخوردار است. لطفاً تمامی نکات و فرض‌هایی که برای پیاده‌سازی‌ها و محاسبات خود در نظر می‌گیرید را در گزارش ذکر کنید.
- در گزارش خود برای تصاویر زیرنویس و برای جداول هم بالانویس اضافه کنید.
- الزامی به ارائه توضیح جزئیات کد در گزارش نیست. اما باید نتایج به دست آمده را گزارش و تحلیل کنید.
- دستیاران آموزشی ملزم به اجرا کردن کدهای شما نیستند. بنابراین هرگونه نتیجه و یا تحلیلی که در شرح سوال از شما خواسته شده است را به طور واضح و کامل در گزارش بیاورید. در صورت عدم رعایت این مورد، بدیهی است که از نمره تمرین کسر می‌شود.
- در صورت مشاهده تقلب امتیاز تمامی افراد شرکت‌کننده در آن، 100- لحاظ می‌شود.
- برای انجام تمرین‌ها و مینی پروژه‌ها، تنها زبان برنامه نویسی مجاز Python است.
- استفاده از کدهای آماده برای تمرین‌ها به هیچ‌وجه مجاز نیست. اما برای مینی پروژه‌ها فقط برای قسمت‌هایی از کد و به عنوان راهنمایی برای پیاده‌سازی، می‌توانید از کدهای آماده استفاده کنید.
- نحوه محاسبه تاخیر به این شکل است: مهلت ارسال بدون جریمه تا تاریخ اعلام شده و پس از آن به مدت هفت روز تا 31 اردیبهشت بارگذاری ممکن است و در نهایت، پس از بازه تاخیر نمره تکلیف صفر خواهد شد. به ازای هر روز تاخیر 5 درصد جریمه می‌شوید.
- لطفاً گزارش، فایل کدها و سایر ضمایم مورد نیاز را با فرمت زیر در سامانه مدیریت دروس بارگذاری نمایید.

PROJECT#1_[Lastname]_[StudentNumber].zip

- در صورت وجود هرگونه ابهام یا مشکل می‌توانید از طریق رایانامه‌های زیر با دستیاران آموزشی مربوطه آقای شعاع حقیقی (سوال 3) و خانم حسن ابراهیمی (سوال 1 و 2) و خانم کاویانی (سوال 4) در تماس باشید:

afsaneh.h.ebrahimi@gmail.com

bahareh.kaviani75@gmail.com

a.shoaahaghighi@gmail.com