



به نام خدا
درس مبانی یادگیری عمیق
تمرین سری دوم
استاد درس : دکتر مرضیه داوودآبادی
دستیاران : سحر سرکار، فائزه صادقی، حسن حماد
دانشگاه علم و صنعت ایران، دانشکده مهندسی کامپیوتر
نیمسال اول تحصیلی ۱۴۰۲ - ۱۴۰۳

مهلت تحویل : ۱۴۰۲/۰۸/۱۶

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

۱. به سوالات زیر پاسخ دهید. (۱۵ نمره)

الف) مشکل بیش‌برازش^۱ و کم‌برازش^۲ در شبکه‌های عصبی را توضیح دهید.

ب) فرض کنید مدلی داریم که قبلاً آموزش دیده است. چگونه می‌توانیم بیش‌برازش مدل را تشخیص دهیم؟

پ) یکی از راه‌های جلوگیری از بیش‌برازش استفاده از *Dropout* است. فرض کنید مقادیر یکی از لایه‌های یک شبکه عصبی و ماسک *Dropout* به صورت زیر باشند. مقادیر نهایی این لایه بعد از اعمال *Dropout* را در مرحله‌ی آموزش و آزمون محاسبه کنید.

جدول ۲: *Dropoutmask*

۱	۰	۰	۱
۰	۱	۱	۰
۰	۱	۱	۰
۱	۰	۰	۱

جدول ۱: *Output*

۱.۶	-۰.۷	-۰.۲	۱.۹
-۲.۳	۲.۵	۲.۵	-۰.۹
-۰.۵	۳.۲	۳.۷	-۰.۴
۱.۳	-۰.۴	-۲.۶	۱.۲

¹Overfitting

²Underfitting

۲. لطفا سوالات زیر را به صورت کامل پاسخ دهید. (۲۰ نمره)

الف) یکی از الگوریتم‌هایی که در حوزه یادگیری ماشین مورد استفاده قرار می‌گیرند، الگوریتم نزدیک‌ترین همسایگی^۳ است. برای مطالعه بیشتر درباره‌ی این الگوریتم می‌توانید به این [لینک](#) مراجعه کنید. توضیح دهید که با تغییر مقدار K ، بایاس و واریانس چه تغییری می‌کنند.

ب) درست یا غلط بودن گزاره‌های زیر را مشخص کنید و دلیل پاسخ خود را نیز بیان کنید.

- استفاده از منظم‌سازی، ممکن است باعث تضعیف عملکرد مدل شود.

- اضافه کردن تعداد زیاد ویژگی‌های^۴ جدید، باعث جلوگیری از بیش‌برازش می‌شود.

- با زیاد کردن ضریب منظم‌سازی، احتمال بیش‌برازش بیشتر می‌شود.

پ) فرض کنید یک مدلی داریم و برای جلوگیری از بیش‌برازش می‌خواهیم از منظم‌سازی $L1$ و $L2$ استفاده کنیم. برای این کار چهار آزمایش اجرا کرده و نتایج به دست آمده به صورت زیر است. با توجه به این نتایج، مشخص کنید در هر آزمایش از کدام منظم‌سازی استفاده شده است (دلیل انتخاب خود را توضیح دهید).

- $W_{exp1} = [0.26, 0.25, 0.25, 0.25]$

- $W_{exp2} = [1, 0, 0, 0]$

- $W_{exp3} = [13.3, 23.5, 53.2, 5.1]$

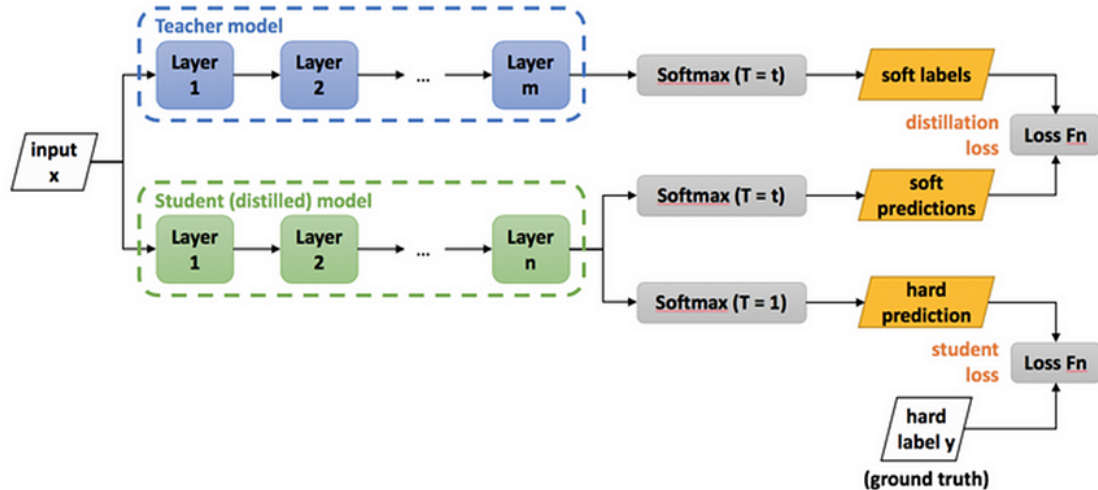
- $W_{exp4} = [0.5, 1.2, 8.5, 0]$

³K-Nearest Neighbors (KNN)

⁴Features

۳. به سوالات زیر پاسخ دهید. (۱۵ نمره)

الف) فرایند تقطیر دانش^۵ چیست و به چه منظور استفاده می‌شود؟
 ب) معماری شکل زیر برای استخراج دانش از شبکه‌ی *teacher* و انتقال آن به شبکه‌ی *student* پیشنهاد شده است. روند یادگیری آن را توضیح دهید.



پ) وزن‌های شبکه‌ی *student* با توجه به کدام تابع ضرر^۶ به‌روزرسانی خواهند شد؟

۴. در نوت‌بوک پیوست شده، کد آماده‌ای قرار داده شده که تنها نیاز است سلول‌ها را اجرا کنید و نتایج به دست آمده از بهینه‌سازهای متفاوت را با هم مقایسه کرده و تحلیل نمایید (لطفاً تحلیل خود را به علاوه توابع استفاده شده به صورت کامل توضیح دهید). (۱۰ نمره)

۵. در نوت‌بوک داده شده، موارد خواسته شده را پیاده‌سازی کنید. در این نوت‌بوک هدف آموزش یک شبکه ساده *MLP* برای یادگیری مجموعه داده *FashionMNIST* می‌باشد. (۴۰ نمره)
 الف) در این بخش مدل خود را تعریف کنید. تعداد لایه‌ها و تعداد نوروں‌های هر لایه بر عهده شماست. برای تابع ضرر از *CrossEntropy* و برای بهینه‌ساز از *SGD* استفاده کنید. سپس قسمت آموزش مدل را تکمیل کنید. در قسمت تست نیز خروجی مدل برای چند عکس موجود در داده‌های تست را به دست آورده و با برچسب واقعی مقایسه کنید.

۶. ب) مدل تعریف شده در قسمت الف را تغییر دهید تا شبکه شما دچار بیش‌برازش شود. دلیل بیش‌برازش شبکه در این مرحله را توضیح دهید. همچنین نموداری رسم کنید که میزان خطای حین آموزش و آزمون را با هم مقایسه کند.

پ) حال تلاش کنید فقط با داده‌افزایی^۷، شبکه بیش‌برازش شده را بهبود دهید. برای مطالعه بیشتر

^۵Knowledge Distillation

^۶Loos Function

^۷Data Augmentation

درباره‌ی داده‌افزایی در *PyTorch* می‌توانید از این [لینک](#) راهنمایی بگیرید. حداقل دو مورد از تبدیلات توضیح داده شده را پیاده‌سازی کرده و نتایج حاصل را تحلیل کنید.

ت) با استفاده از منظم‌سازی $L1$ یا $L2$ (به انتخاب خود) شبکه را بهبود دهید و نتایج را تحلیل کنید.

ث) (امتیازی) با استفاده از ترکیبی از داده‌افزایی، منظم‌سازی و *Dropout* شبکه را بهبود و بیان کنید که چه ترکیبی از این‌ها باعث بهبود حداکثری می‌شود. (۱۵ نمره)