

به نام خدا



دانشکده مهندسی کامپیوتر

درس شبکه‌های عصبی

دکتر ناصر مزینی

تمرین سری دوم

دستیاران آموزشی:

فائزه صادقی

غزاله محمودی

تاریخ تحویل:

۱۴۰۱/۸/۷

نکات تکمیلی:

۱. پاسخ سوالات را به صورت کامل در یک فایل PDF و به همراه کدها در فرمت ipynb. در یک فایل فشرده به شکل HW2_StudentID.zip قرار داده و تا زمان تعیین شده بارگذاری نمایید.
 ۲. برای پیاده سازی ها زبان پایتون پیشنهاد می شود، لازم به ذکر است توضیح کدها و نتایج حاصله باید در فایل PDF آورده شوند و به کد بدون گزارش نمره ای تعلق نخواهد گرفت.
 ۳. در مجموع تمام تمرین ها، تنها ۷۲ ساعت تاخیر در ارسال پاسخ ها مجاز است و پس از آن به ازای هر روز تاخیر ۱۰ درصد از نمره کسر می گردد.
 ۴. چنانچه دانشجویی تمرین را زودتر از موعد ارسال کند و ۷۰ درصد از نمره را کسب کند، تا سقف ۴۸ ساعت به ساعات مجاز تاخیر دانشجو اضافه می گردد.
 ۵. لطفا منابع استفاده شده در حل هر سوال را ذکر کنید.
 ۶. تمرین ها باید به صورت انفرادی انجام شوند و حل گروهی تمرین مجاز نیست.
 ۷. ارزیابی تمرین ها بر اساس صحیح بودن راه حل، گزارش های کامل و دقیق، بهینه بودن کدها و کپی نبودن می باشد.
 ۸. لطفا برای انجام تمرین زمان مناسب اختصاص دهید و انجام آن را به روزهای پایانی موکول نکنید.
 ۹. سوالات خود را می توانید در گروه مربوطه مطرح نمایید (لطفا از پرسیدن سوالات درسی به صورت شخصی خودداری فرمایید، زیرا سوالات بقیه ی دانشجویان هم می تواند مشابه سوالات شما باشد و پرسیدن در فضای عمومی مفیدتر واقع می شود).
- موفق باشید.

۱- به سوالات زیر پاسخ دهید.

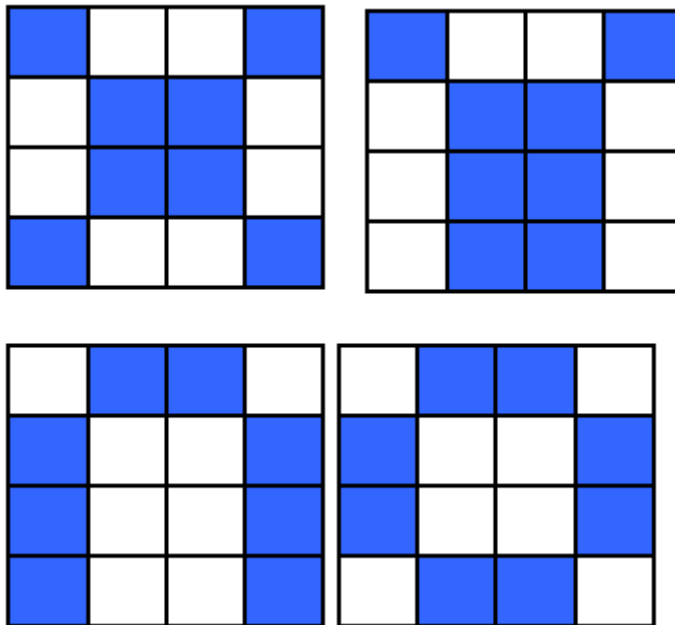
أ. مفهوم overfit در شبکه‌های عصبی را توضیح دهید. چه زمانی می‌گوییم شبکه overfit شده است؟

ب. چه روش‌های برای رفع مشکل overfit پیشنهاد می‌کنید؟ به طور کامل توضیح دهید.

ت. مفهوم underfit در شبکه‌های عصبی را توضیح دهید.

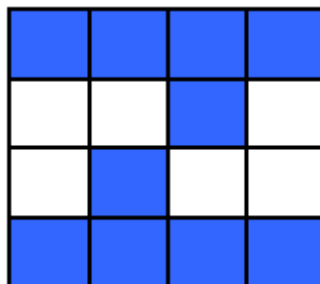
ث. برای دسته‌بندی دو کلاس چه تابع فعالساز و تابع ضرری را برای لایه خروجی پیشنهاد می‌دهید؟ علت پیشنهادتان را شرح دهید.

۲- فرض کنید ۴ الگوی زیر موجود است. با استفاده از یک شبکه MLP تفکیک دو الگو بالا از دو الگو پایین را انجام دهید.

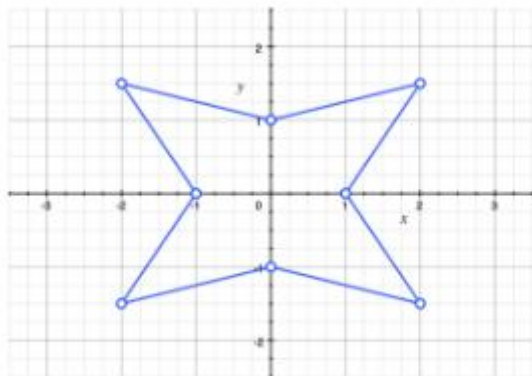


أ. ساختار و پارامترهای شبکه را پیشنهاد دهید.

ب. به نظر شما پاسخ شبکه برای الگوی جدید زیر چیست؟

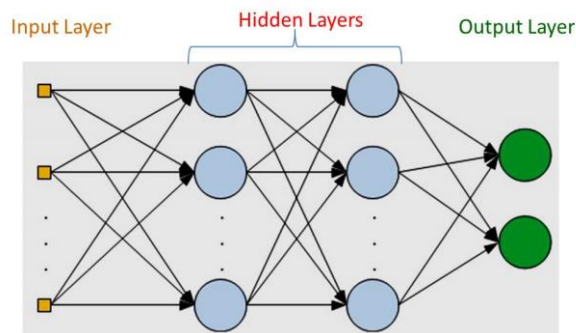


۳- شبکه‌ای را در نظر بگیرید که ورودی آن x, y (مختصات صفحه) می‌باشد و تنها یک لایه پنهان دارد. آیا امکان پذیر است که این شبکه به ازای ورودی‌های موجود در ناحیه زیر ۱ و در بقیه حالات صفر باشد؟ آیا تعداد نورون‌های لایه پنهان می‌تواند تعیین کننده باشد؟ پاسخ خود را با ذکر دلیل توضیح دهید.



۴- در این بخش، ویژگی‌های یک مدل یادگیری عمیق اولیه به نام پرسپترون چندلایه (MLP) را بررسی و پیاده‌سازی خواهید کرد. اساساً هدف MLP یادگیری یک نقشه برداری غیرخطی از ورودی‌ها به خروجی‌ها است. ما می‌توانیم این نگاشت را به صورت $y = f(x; \theta)$ نشان دهیم، که در آن x ورودی و θ بردار تمام پارامترهای شبکه است.

همانطور که در شکل مشاهده می‌کنید، هر شبکه MLP از یک لایه ورودی، یک لایه خروجی و یک یا چند لایه پنهان در بین آنها تشکیل شده است. هر لایه از یک یا چند سلول به نام نورون تشکیل شده است. در هر نورون، حاصل ضرب نقطه‌ای بین ورودی‌های سلول و بردار وزن محاسبه می‌شود. نتیجه حاصل ضرب نقطه‌ای سپس از طریق یک تابع غیرخطی (تابع فعال‌سازی به عنوان مثال \tanh یا sigmoid) عبور می‌کند و خروجی نورون را به ما می‌دهد.



در این تمرین ورودی به صورت ماتریس با اندازه $M * b$ که b اندازه batch و M تعداد feature هاست. در ادامه خروجی لایه i ام به صورت زیر محاسبه می‌شود:

$$A^i = f(A^{i-1}w^i + b^i)$$

که با فرض اینکه لایه $(i-1)$ و i ام به ترتیب n و p نورون داشته باشند، ابعاد وزن و بایاس به صورت زیر است:

$$w^{n \times p}, b^{1 \times p}$$

با توجه به این توضیحات اولیه با مراجعه به نوت بوک پیوست شده، با مطالعه توضیحات هر بخش توابع داده شده را کامل کنید و سپس به سوالات زیر پاسخ دهید.

ا. چرا توابع فعالسازی باید غیرخطی باشند؟ آیا هر تابع غیرخطی را به عنوان تابع فعالسازی می توان استفاده کرد؟

ب. اگر در ابتدای کار شبکه عصبی به جای اینکه وزن ها را به صورت تصادفی مقدار دهی کنیم، مقدار اولیه شان را صفر قرار دهیم، چه اتفاقی می افتد؟

ت. آیا در گام های الگوریتم گرادیان کاهشی همواره گرادیان کاهش می یابد؟

ث. با توجه به نمودار دقت و تابع ضرر مدل، مدل خود را از نظر overfit , fit , underfit بررسی کنید.

۵- (امتیازی) با استفاده از کتابخانه Keras یک پرسپترون چندلایه طراحی کنید تا عملیات دسته بندی را بر روی بر روی دیتاست mnist انجام دهد.