



به نام خدا
دانشگاه تهران
دانشکده مهندسی برق و کامپیوتر



درس شبکه‌های عصبی و یادگیری عمیق

تمرین شماره یک

طراحان تمرین:

سوال اول و دوم: رضا براتی

سوال سوم و چهارم: علی فرتوت

مهلت ارسال: ۱۱ آبان ۱۴۰۴

4	سوال ۱. نرون McCulloch-Pitts
5	۱-۱. منطق آستانه‌ای (۴۰ امتیاز)
5	۲-۱. Majority و Parity (۶۰ امتیاز)
6	سوال ۲. Madaline و Adaline
6	۱-۲. Breast Cancer روی Adaline
6	۱-۱-۲. از SSE تا LMS (تئوری) (۱۰ امتیاز)
6	۲-۱-۲. آماده‌سازی داده و تحلیل اولیه (۸ امتیاز)
6	۳-۱-۲. پیاده‌سازی Adaline از صفر (۱۷ امتیاز)
6	۴-۱-۲. ارزیابی و مصورسازی (۱۰ امتیاز)
7	۵-۱-۲. مدل‌های پایه و بحث (۵ امتیاز)
7	۲-۲. Madaline (Rule II) بر روی داده‌های Moons و Circles
7	۱-۲-۲. پیاده‌سازی Madaline با قانون دوم (۱۵ امتیاز)
7	۲-۲-۲. داده‌ها و تنظیمات آموزشی (۱۰ امتیاز)
7	۳-۲-۲. ارزیابی و نمودارها (۱۵ امتیاز)
7	۴-۲-۲. مطالعه ابرپارامترها (۱۰ امتیاز)
9	سوال ۳. رگرشن با استفاده از mlp
9	۱-۳. دریافت دیتاست و انجام پیش‌پردازش‌های لازم (۲۵ نمره)
9	۲-۳. پیاده‌سازی یک شبکه عصبی ساده (۱۵ نمره)
10	۳-۳. آموزش شبکه عصبی ساده (۲۵ نمره)
10	۴-۳. طراحی و آموزش شبکه عصبی پیچیده‌تر (۲۰ نمره)
10	۵-۳. مقایسه (۱۵ نمره)
10	نکات
11	سوال ۴. طراحی شبکه AutoEncoder
11	۱-۴. دریافت دیتاست و انجام پیش‌پردازش‌های لازم (۱۰ نمره)
11	۲-۴. پیاده‌سازی (۲۵ نمره)
11	۳-۴. آموزش شبکه (۳۵ نمره)

11	۴-۴. ارزیابی شبکه (نمره ۳۰)
11	۴-۴-۱. انکود کردن داده‌ها
11	۴-۴-۲. کلاسترینگ با K-Means
12	۴-۴-۳. تحلیل خوشه‌ها
12	۴-۴-۴. کاهش ابعاد و t-SNE
12	۴-۴-۵. مقایسه با داده‌های خام
12	۴-۴-۶. ارزیابی کیفیت خوشه‌بندی
12	نکات

جدول‌ها

9	جدول 1. ساختار مدل شبکه عصبی
10	جدول 2. هایپرپارامترهای شبکه عصبی

۱-۱. منطق آستانه‌ای (۴۰ امتیاز)

یک نورون آستانه‌ای را در نظر بگیرید که خروجی آن برابر یک است هرگاه جمع وزن‌دار ورودی‌ها به علاوه bias نامنفی باشد و در غیر این صورت صفر. بدون کدنویسی، وزن‌ها و bias را برای گیت‌های AND و OR و NOT تعیین کنید، سپس برای هر کدام جدول درستی را بنویسید و استدلال بیاورید که چرا این تنظیمات درست کار می‌کنند. در ادامه، توضیح دهید که چرا یک نورون آستانه‌ای فقط می‌تواند توابع بولی جداسازی‌پذیر خطی را نمایش دهد؛ توضیح شما باید شهودی باشد و به هندسهٔ مرز تصمیم اشاره کند.

۱-۲. Majority و Parity (۶۰ امتیاز)

به صورت هندسی توضیح دهید که چرا تابع Parity با سه ورودی یا همان XOR سه‌تایی با یک نورون آستانه‌ای قابل نمایش نیست؛ استدلال شما باید به این نکته تکیه کند که نقاط با جمع فرد یک‌ها در دو سوی مرزهای خطی قابل جداسازی هم‌زمان نیستند. سپس کوچک‌ترین شبکهٔ دولایهٔ آستانه‌ای را که این تابع را نمایش می‌دهد طراحی کنید؛ نمودار سادهٔ معماری را رسم کنید و وزن‌ها و bias های عددی پیشنهادی خود را بنویسید تا نشان دهید چگونه خروجی درست تولید می‌شود. در ادامه برای تابع Majority با پنج ورودی یک شبکهٔ دولایه طراحی کنید و به صورت مستدل بحث کنید حداقل چند نورون در لایهٔ پنهان لازم است تا این تابع به درستی پیاده شود. در این بخش نیز نیازی به کدنویسی نیست و تأکید بر تحلیل و طراحی روشن است.

سوال ۲. Madaline و Adaline

در این سؤال دو تجربه طبقه‌بندی انجام می‌دهید: نخست با Adaline روی داده Breast Cancer Wisconsin که یک مسئله دودویی است، و سپس با Madaline Rule II روی داده‌های مصنوعی غیرخطی moons و circles. منظور از moons دو هلال در هم‌تنیده با برجسب متفاوت است و منظور از circles دو دایره هم‌مرکز با برجسب متفاوت. در هر بخش، داده را با نسبت هفتاد به سی به آموزش و آزمون تقسیم کنید و seed را صفر بگذارید؛ نرمال‌سازی ویژگی‌ها فقط بر پایه آمار مجموعه آموزش انجام شود. همه نمودارهای روند آموزش روی داده آموزش رسم شوند و همه معیارهای نهایی فقط روی داده آزمون گزارش شوند.

۲-۱. Breast Cancer روی Adaline

۲-۱-۱. از SSE تا LMS (تئوری) (۱۰ امتیاز)

با شروع از مجموع مربعات خطا برای اهداف منفی یک و مثبت یک، قانون به‌روزرسانی دسته‌ای LMS را استخراج کنید و ارتباط آن را با معادلات نرمال توضیح دهید. قانون به‌روزرسانی آنلاین را نیز بنویسید و توضیح دهید افزودن جریمه L2 چگونه به ridge می‌انجامد.

۲-۱-۲. آماده‌سازی داده و تحلیل اولیه (۸ امتیاز)

داده را با استفاده از دستور `sklearn.datasets.load_breast_cancer` لود کنید و با نسبت هفتاد به سی و seed برابر صفر به آموزش و آزمون تقسیم کنید. فقط با دیتای آموزش نرمال‌سازی کنید. توازن کلاس‌ها و نیز میانگین و انحراف معیار ویژگی‌های نرمال‌شده را در مجموعه آموزش گزارش کنید.

۲-۱-۳. پیاده‌سازی Adaline از صفر (۱۷ امتیاز)

Adaline را با پارامترهای نرخ یادگیری، تعداد دوره‌ها و جریمه اختیاری L2 پیاده‌سازی کنید. یک جست‌وجوی کوچک روی نرخ یادگیری انجام دهید، برای نمونه میان یک هزارم، پنج هزارم و دو صدم، با پنجاه دوره یا توقف زود هنگام در صورت واگرایی. در هر دوره خطای SSE یا MSE و نیز دقت آموزش را ثبت کنید. پیاده‌سازی روشن و خوانا کافی است؛

۲-۱-۴. ارزیابی و مصورسازی (۱۰ امتیاز)

بر مبنای آزمون، ماتریس درهم‌ریختگی، دقت، recall، precision، F1 و ROC-AUC را برای بهترین مدل گزارش کنید. سپس ویژگی‌های آزمون را با PCA به دو بُعد ببرید و مرز تصمیم خطی Adaline را در این فضا رسم کنید و دست‌کم سه نمونه false positive و سه نمونه false negative را مشخص نمایید. در پایان این بخش دو

نمودار آموزش را بیاورید: خطا بر حسب دوره و دقت بر حسب دوره، به گونه‌ای که منحنی‌های مربوط به همه نرخ‌های یادگیری روی یک نمودار دیده شوند.

۲-۱-۵. مدل‌های پایه و بحث (۵ امتیاز)

Logistic Regression و LDA را با همان پیش‌پردازش آموزش دهید. یک جدول مقایسه کوتاه از معیارها ارائه کنید و در حداکثر ده خط بحث کنید که Adaline در چه شرایطی بهتر یا بدتر از این پایه‌های احتمالاتی عمل می‌کند.

۲-۲. (Rule II) Madaline بر روی داده‌های Moons و Circles

۲-۲-۱. پیاده‌سازی Madaline با قانون دوم (۱۵ امتیاز)

یک شبکه دولا به با m واحد Adaline در لایه پنهان و یک خروجی نهایی با تابع sign پیاده‌سازی کنید؛ منظور از sign این است که خروجی نهایی مثبت یکی تولید می‌کند اگر ورودی آن نامنفی باشد و در غیر این صورت منفی. الگوریتم Rule II را اجرا کنید: برای هر نمونه اشتباه طبقه‌بندی شده آن واحدهای پنهان را که تغییر علامت یا وزن‌شان بیشترین کاهش خطا را ایجاد می‌کند شناسایی کنید و یک به‌روزرسانی حریصانه انجام دهید و این فرایند را تکرار نمایید. یک نمونه شبکه‌کد را در گزارش بیاورید و پیاده‌سازی را در نوت‌بوک نگه دارید.

۲-۲-۲. داده‌ها و تنظیمات آموزشی (۱۰ امتیاز)

برای ایجاد داده در این قسمت از توابع make_moons و make_circles استفاده کنید. داده moons را با ششصد نمونه و نویز بیست و پنج صدم و seed صفر بسازید. داده circles را با ششصد نمونه و نویز یک‌دهم و seed صفر بسازید. تقسیم هفتاد به سی و نرمال‌سازی بر پایه آموزش انجام شود. تنظیمات آموزشی ساده و سبک باشد: تعداد واحدهای پنهان را میان سه و پنج بررسی کنید، نرخ یادگیری را ثابت و برابر یک‌هزارم بگیرید، جریمه L_2 را میان صفر و یک‌هزارم بررسی کنید. در هر دوره خطا و دقت آموزش را ثبت کنید.

۲-۲-۳. ارزیابی و نمودارها (۱۵ امتیاز)

برای هر داده بهترین مدل را بر اساس دقت آزمون انتخاب کنید. بر مبنای آزمون، ماتریس درهم‌ریختگی، دقت، precision، recall، F1 و ROC-AUC را گزارش دهید. مرز تصمیم را در فضای ورودی ترسیم کنید و نقاط اشتباه را مشخص نمایید. توضیح دهید ابرصفحه‌های پنهان چگونه ترکیب می‌شوند تا مرز نهایی را بسازند. نمودار خطا بر حسب دوره و دقت بر حسب دوره را برای مدل برگزیده در پایان این بخش بیاورید.

۲-۲-۴. مطالعه ابرپارامترها (۱۰ امتیاز)

نتایج را برای هر داده در یک جدول خلاصه کنید؛ میانگین و انحراف معیار را روی سه seed یعنی صفر و یک و دو گزارش دهید. دربارهٔ بیش‌برازش و کم‌برازش و ارتباط آن با پیچیدگی مرز تصمیم توضیح دهید.

سوال ۳. رگرشن با استفاده از mlp

در این سوال شما وظیفه طراحی یک شبکه عصبی برای مسئله رگرسیون از دیتاست Housing Prices دارید. دیتاست «California Housing Prices» شامل اطلاعاتی درباره قیمت خانه‌ها در ایالت کالیفرنیا است. این مجموعه داده برای تحلیل‌های آماری و یادگیری ماشین طراحی شده و می‌تواند برای پیش‌بینی قیمت خانه‌ها، تحلیل روند بازار مسکن و مدل‌سازی رگرسیون استفاده شود. ویژگی‌های موجود در دیتاست شامل متغیرهایی مانند مساحت خانه، تعداد اتاق‌ها، سن ساختمان و موقعیت جغرافیایی است.

۳-۱. دریافت دیتاست و انجام پیش‌پردازش‌های لازم (۲۵ نمره)

در ابتدا دیتاست ([لینک](#)) را دریافت کنید. سپس برحسب لیست زیر پیش‌پردازش‌های گفته شده را اعمال کنید. علت هر یک از پیش‌پردازش‌های گفته شده را بیان کنید. (حتماً دیتاست را به صورت دو دسته آموزش و ارزیابی تقسیم کنید.)

- انتخاب یک استراتژیک برای مدیریت missing values
- حذف کردن Outliers از قسمت target column یا همان قیمت
- تبدیل فیچرهای categorical به یک مقدار قابل قبول.

۳-۲. پیاده‌سازی یک شبکه عصبی ساده (۱۵ نمره)

به جدول زیر توجه کنید. در این جدول معماری شبکه عصبی داده شده است. آن را با کتابخانه مورد نظر پیاده سازی کنید. تعداد پارامترهای آن را گزارش دهید.

جدول ۱. ساختار مدل شبکه عصبی

NN Module	Parameters
Linear	In = input_size, Out = 8
(2) ReLU (Activate Function)	
(3) Linear	In = 8, Out = 1

جدول 2. هایپرپارامترهای شبکه عصبی

Hyper Parameters	Value
Batch Size	64
Train, Validation, Test ratio	10% - 10% - 80%
Optimizer	Adam
Learning Rate	0.1
Loss Function	MSE

۳-۳. آموزش شبکه عصبی ساده (۲۵ نمره)

حال شبکه عصبی پیاده‌سازی شده را برای ۶۰ اپیاک آموزش دهید. معیارهای R^2 ، MAE، MSER، و RMSE را توضیح دهید و کاربرد هر کدام را بیان کنید. سپس معیارهای گفته شده را گزارش دهید. نمودار loss را نیز گزارش کنید.

۳-۴. طراحی و آموزش شبکه عصبی پیچیده تر (۲۰ نمره)

حال با توجه به شبکه عصبی گفته شده یک شبکه پیچیده‌تری طراحی کنید. که با تعداد اپیاک برابر، عملکرد بهتری نسبت به شبکه ساده داشته باشد. (این عملکرد بهتر باید در معیارهای گزارش شده قابل شهود باشد) تعداد پارامترهای آن را گزارش دهید.

۳-۵. مقایسه (۱۵ نمره)

هایپرپارامتر، نمودارها و معیارهای دو شبکه را در کنار هم مقایسه کنید و نتیجه‌گیری خودتان را از آن‌ها بیان کنید. همچنین نمودار Actual vs predicted [1] را برای شبکه عصبی پیچیده تر گزارش دهید.

نکات

- تمامی معیارهای R^2 ، MAE، MSER، و RMSE باید توسط خودتان پیاده‌سازی شود.
- توجه شود که نمودار خطا باید برای مجموعه آزمون و ارزیابی، و معیارها برای هر سه مجموعه گزارش شوند.
- برای کسب نمره کامل بخش ۳ معیار R^2 باید حداقل به ۰.۶ برسد.
- برای کسب نمره کامل بخش ۴ معیار R^2 باید حداقل به ۰.۷ برسد.

سوال ۴. طراحی شبکه AutoEncoder

شبکه‌های عصبی AutoEncoder نوعی مدل یادگیری عمیق هستند که هدف آن‌ها یادگیری نمایش فشرده و معنادار از داده‌ها به صورت خودکار است. در پروژه بازطراحی تصاویر دیتاست [Fashion MNIST](#)، از AutoEncoder برای کاهش ابعاد و استخراج ویژگی‌های مهم تصاویر استفاده می‌شود تا بتوان تصاویر اصلی را با کمترین اتلاف کیفیت بازسازی کرد. این رویکرد به‌خصوص در فشرده‌سازی داده‌ها، حذف نویز و یادگیری ویژگی‌های غیرخطی بسیار کاربردی است و می‌تواند به بهبود عملکرد مدل‌های تشخیص و دسته‌بندی کمک کند.

۴-۱. دریافت دیتاست و انجام پیش‌پردازش‌های لازم (۱۰ نمره)

دیتاست گفته شده را دریافت کنید. مطمئن شوید داده‌های آموزش و آزمون قبل از ورود به مدل نرمالایز می‌شوند. تحلیل کنید چرا اصولاً داده‌ها باید نرمالایز شوند.

۴-۲. پیاده‌سازی (۲۵ نمره)

معماری شبکه AutoEncoder را تحلیل و بررسی کنید، سپس آن را با استفاده از لایه‌های Fully Connected پیاده‌سازی نمایید. برای هر کدام از بخش‌های شبکه (انکودر و دیکودر) سه لایه Fully Connected کافی است. پیاده‌سازی و تنظیم هایپرپارامترهای شبکه به عهده خودتان است. در این پیاده‌سازی می‌توانید از تابع فعال‌سازی ReLU استفاده کنید، تابع خطا را با MSE تعریف کنید و الگوریتم بهینه‌سازی AdamW را به کار ببرید.

۴-۳. آموزش شبکه (۳۵ نمره)

شبکه را آموزش دهید و نمودار خطا آن را برای مجموعه آموزش و تست گزارش دهید. ۵ نمونه از ۵ کلاس مختلف از مجموعه تست رندوم انتخاب کنید و تصویر بازسازی شده به همراه عکس حقیق دیتاست را گزارش دهید و مقایسه کنید.

۴-۴. ارزیابی شبکه (۳۰ نمره)

پس از آموزش AutoEncoder، مراحل زیر را انجام دهید.

۴-۴-۱. انکود کردن داده‌ها

با استفاده از انکودر شبکه آموزش دیده، تمامی داده‌های تست را به فضای فشرده شده (فضای ویژگی) تبدیل کنید.

۴-۴-۲. کلاسترینگ با K-Means

با مقدار $K=10$ فضای ویژگی را با الگوریتم K-Means خوشه‌بندی کنید.

۴-۴-۳. تحلیل خوشه‌ها

ماتریس درهم‌ریختگی (**Confusion Matrix**) را بین برچسب‌های واقعی و خوشه‌های پیش‌بینی‌شده محاسبه و رسم کنید. تحلیل کنید که هر خوشه بیشتر به کدام کلاس اصلی تعلق دارد و کدام کلاس‌ها اشتراک بیشتری در خوشه‌ها دارند.

۴-۴-۴. کاهش ابعاد و t-SNE

t-SNE را برای نمایش دوبعدی فضای انکودشده اجرا کرده و نتایج را یک‌بار با رنگ‌بندی بر اساس برچسب‌های واقعی (True Labels) و یک‌بار با رنگ‌بندی بر اساس خوشه‌بندی (K-Means Predicted Clusters) نمایش دهید.

۴-۴-۵. مقایسه با داده‌های خام

در این مرحله هدف آن است که عملکرد خوشه‌بندی در فضای ویژگی فشرده‌شده‌ی AutoEncoder با خوشه‌بندی در فضای اولیه‌ی داده‌ها (فضای خام ورودی) مقایسه شود. برای این منظور، تمامی مراحل خوشه‌بندی، Confusion Matrix، و t-SNE که در مراحل قبل برای داده‌های انکودشده انجام شد، باید برای داده‌های اصلی بدون استفاده از شبکه عصبی نیز تکرار شوند.

۴-۴-۶. ارزیابی کیفیت خوشه‌بندی

برای هر دو حالت (با و بدون AutoEncoder) معیارهای Adjusted Rand Index و Adjusted Mutual Information را محاسبه و مقایسه کنید.

نکات

- مقدار ARI باید حداقل ۰.۳ و مقدار AMI باید حداقل ۰.۵۰ باشد.
- تمامی معیارها باید توسط خودتان پیاده‌سازی شود.