



مبانی یادگیری ماشین

نیم سال دوم ۰۱-۰۲

مدرس: دکتر حامد ملک

دانشکده‌ی مهندسی کامپیوتر

موعد تحویل: ۱۴۰۲/۰۲/۳۰

تمرین سری پنجم و ششم

مسائل تئوری

مسئله‌ی ۱.

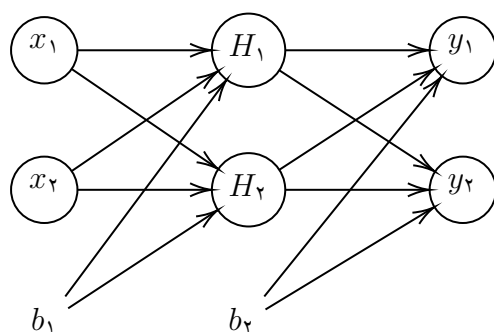
توضیح دهید که چطور می‌توان مقدار مناسب برای تعداد دسته‌ها را در یک الگوریتم یادگیری بدون نظارت^۱ پیدا کرد.

مسئله‌ی ۲.

روشی ارائه کنید تا بتوان با آن تعداد رنگ‌های مختلف موجود در یک عکس را کاهش داد در حالی که همچنان محتوای عکس قابل تشخیص باشد. جزئیات اینکه چگونه به هر پیکسل رنگ جدید را اختصاص می‌دهید را در راه حل به صورت کامل توضیح دهید.

مسئله‌ی ۳.

در شبکه عصبی زیر با یک لایه پنهان، عملیات پس انتشار^۲ را یک بار انجام دهید و تمام بردار وزن‌های موجود را بروزرسانی کنید و مقدارشان را بنویسید. مقدار واقعی y_1 برابر ۰.۱۰ و مقدار واقعی y_2 برابر ۰.۹۹ است. همچنین نرخ یادگیری برابر ۰.۵ است. همه توابع فعال‌ساز نیز سیگموئید^۳ هستند. از تابع میانگین مربع خطا به عنوان تابع هزینه استفاده کنید.



$$\begin{aligned}x_1 &= 0.05 & x_2 &= 0.10 \\b_1 &= 0.35 & b_2 &= 0.60 \\w_{11} &= 0.15 & w_{12} &= 0.25 \\w_{21} &= 0.20 & w_{22} &= 0.30 \\u_{11} &= 0.40 & u_{12} &= 0.5 \\u_{21} &= 0.45 & u_{22} &= 0.55\end{aligned}$$

مسئله‌ی ۴.

نحوه کارکرد شبکه‌های عصبی پیچشی^۴ را توضیح دهید.

¹Unsupervised learning

²back propagation

³Sigmoid

⁴Convolutional neural network

مسئله‌ی ۵.

در رابطه با فقط یک مورد از بهینه‌سازهای شبکه عصبی زیر توضیح دهید.

- RMSprop
- Adam
- Adagrad

پروژه‌های کدی

پروژه‌ی ۱. تحلیل خوشه‌ای

الف) الگوریتم K-means را خودتان و بدون استفاده از الگوریتم آماده آن پیاده‌سازی کنید.

ب) برای دیتاست‌های موجود در این [لینک](#) با استفاده از الگوریتمی که پیاده‌سازی کردید، با سعی و خطا بهترین مقدار k را پیدا کنید و نمودار نقطه‌ای^۵ آن را به صورت مناسب نمایش دهید. همچنین مرکز هر کدام از دسته‌ها در این نمایش مشخص باشد.

پ) با توجه به اینکه نقطه اولیه در این الگوریتم به صورت تصادفی انتخاب می‌شود، خروجی را به ازای حداقل یک نقطه تصادفی دیگر اجرا کنید و با پاسخ قبلی مقایسه کنید.

ت) در مورد الگوریتم DBSCAN مطالعه کنید و با استفاده از کتابخانه دلخواهتان الگوریتم از پیش پیاده‌سازی شده آن را بر روی دیتاست‌های داده‌شده اجرا کنید.

پروژه‌ی ۲. گل‌های آکسفورد

مجموعه داده‌ای از ۱۰۲ دسته‌بندی گل‌های مختلف وجود دارد که در بریتانیای کبیر سابق وجود دارند. این دیتاست از این [لینک](#) قابل دسترسی است. با استفاده از شبکه‌های عصبی مصنوعی^۶ و شبکه‌های عصبی پیچشی مدلی طراحی کنید که بتواند این ۱۰۲ دسته را دسته‌بندی کند. (لازم به ذکر است برای حل این مسئله فقط باید از یکی از کتابخانه‌های pytorch یا tensorflow استفاده کنید.) موارد زیر را پیاده‌سازی کنید.

- ابتدا یک مدل با استفاده از شبکه‌های عصبی مصنوعی طراحی کنید که بتواند این ۱۰۲ کلاس را دسته‌بندی کند. معماری خود را گزارش کنید.
- در این مرحله از بهینه‌سازی که پیش‌تر آن را توضیح دادید استفاده کنید و نتیجه استفاده از آن را بنویسید.
- تمامی نمودارهای مربوط به کاهش خطای آموزش، خطای تست و افزایش دقت آموزش و تست را گزارش دهید. (برای داده تست می‌توانید از داده اعتبارسنجی نیز استفاده کنید.)
- حال مسئله را با استفاده از شبکه‌های عصبی پیچشی حل کنید و نتایج و نمودارها را مجدداً گزارش کنید.

^۵Scatter plot

^۶Artificial neural networks