به نام خدا



دانشگاه صنعتی امیرکبیر

(پلی‌تکنیک تهران)

درس شبکه‌های عصبی

استاد صفابخش

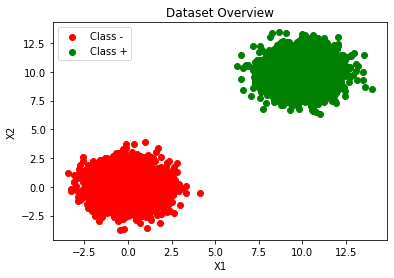
تمرین اول

علیرضا مازوچی

۴۰۰۱۳۱۰۷۵

سوال ۱

برای ایجاد مجموعه‌داده خطی جداپذیر دو توزیع دو بعدی گاوسی با کواریانس یکسان و میانگین‌های و درنظر گرفتم. این مجموعه‌داده با احتمال بسیار بسیار بالا خطی جداپذیر است؛ چراکه فاصله دو مرکز برابر با ده واریانس در هر دو بعد است و در توزیع نرمال 99.9٪ داده‌ها در فاصله‌ای به اندازه چهار انحراف معیار از میانگین قرار گرفته‌اند. مجموعه‌داده می‌تواند چیزی مانند تصویر زیر باشد.



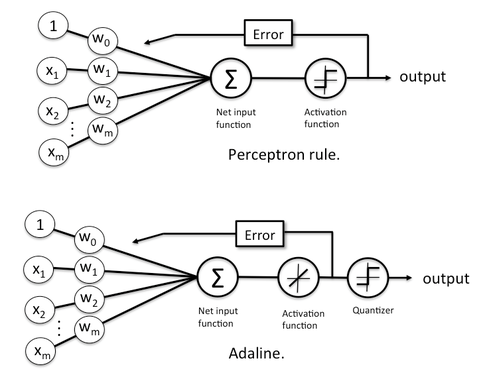
لازم به ذکر است که یک کلاس مثبت با برچسب عددی +۱ و یک کلاس منفی با برچسب عددی ۱- درنظر گرفته شده است.

سوال ۲

باتوجه به آنکه ابعاد مسئله ورودی دوبعدی است، ورودی هر دو واحد پردازشگر هم دوبعدی خواهد بود (). برای مقادیر اولیه وزن‌ها اعداد تصادفی بین 1- تا ۱+ درنظر گرفته‌ام. اما تابع فعالیت دو نورون متفاوت است که جلوتر توضیح داده می‌شود.

تفاوت دو نورون به این شرح است[[1]](#footnote-1)[[2]](#footnote-2):

* پرسپتورون از یک تابع فعالیت استفاده می‌کند که خروجی را باینری ( یا به طور مشابه قطبی) می‌کند. این خروجی باینری همان خروجی نهایی نورون است. خطا و درنتیجه بروزرسانی وزن‌ها هم که برای این نورون درنظر گرفته می‌شود برپایه خروجی گسسته خواهد بود. در آدالاین به طور کلاسیک ابتدا یک تابع فعالیت همانی روی ورودی خالص (net input) لحاظ می‌شود که عملا بی‌تاثیر است. خروجی تابع فعال‌سازی این قسمت که خروجی خطی (linear output) نامیده می‌شود برای محاسبه خطا و بروزرسانی وزن‌ها استفاده می‌شود. پس در اینجا تابع خطا برخلاف پرسپتورون پیوسته خواهد بود. نهایتا از یک تابع sign در خروجی خطی نورون استفاده می‌شود تا یک خروجی قطبی (bipolar output) ایجاد شود که این خروجی پیشبینی مدل به ازای داده ورودی است و برای خطا و بروزرسانی وزن استفاده نمی‌شود. خلاصه این تفاوت در تصویر زیر مشخص است:

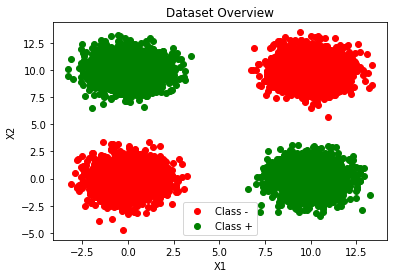


* تفاوت دیگری هم که بین دو نورون وجود دارد برای حالت درجه دو است (بر اساس تدریس کلاس). در نورون پرسپتورون درجه دوم از و استفاده می‌شود. درحالی که در آدالاین مرتبه دوم علاوه بر دو جمله مذکور از هم کمک گرفته می‌شود.

سوال ۳

سوال ۴

باتوجه به آنکه در صورت سوال به صورت غیرمستقیم درخواست شده است که مجموعه‌داده از نوع XOR باشد، من چهار توزیع گاوسی در چهار نقطه از فضا با کواریانس در نظر گرفتم. میانگین این چهار توزیع عبارت است از ، ،، و . توزیع اول و چهارم برای کلاس و دو توزیع دیگر به کلاس دیگر متعلق است. مشابه استدلال‌های ارائه‌شده در سوال ۱ می‌توان با اطمینان بسیار بالا مطمئن شد که داده‌ها این چهار توزیع باهم ترکیب نخواهند شد. مجموعه‌داده به شکل زیر به وجود می‌آید. بدیهی است که مجموعه خطی جداپذیر نیست ولی استفاده از نورون با درجات بالا ممکن است مشکل را حل کند.



سوال ۵

سوال ۶

1. <https://sebastianraschka.com/faq/docs/diff-perceptron-adaline-neuralnet.html> [↑](#footnote-ref-1)
2. <https://datascience.stackexchange.com/q/36368/121138> [↑](#footnote-ref-2)