



هوش مصنوعی

بهار ۱۴۰۲

استاد: محمد مهدی سمیعی

دانشگاه صنعتی شریف

دانشکده مهندسی کامپیوتر

گردآوردندگان: علی مهربانی، سوگند صالحی، امیررضا میرزایی

Supervised Learning مهلت ارسال: { ۱۱ خرداد (تئوری) } { ۱۲ خرداد (عملی) }

تمرین هفتم

- مهلت ارسال پاسخ هر کدام از بخش‌های تئوری و عملی، تا ساعت ۲۳:۵۹ روزهای مشخص شده است.
- در طول ترم امکان ارسال با تاخیر پاسخ وجود ندارد و پاسخ‌هایی که بعد از زمان تعیین شده ارسال شوند، پذیرفته نخواهند شد.
- همکاری و هم‌فکری شما در انجام تمرین مانعی ندارد اما پاسخ‌های هر کس حتما باید توسط خود او نوشته شده باشد.
- در صورت هم‌فکری و یا استفاده از هر منابع خارج درسی، نام هم‌فکران و آدرس منابع مورد استفاده برای حل سوال مورد نظر را ذکر کنید.
- لطفا تصویری واضح از پاسخ سوالات نظری بارگذاری کنید. در غیر این صورت پاسخ شما تصحیح نخواهد شد.

سوالات نظری (۱۰۰ نمره)

۱. (۱۰ نمره) به سوالات زیر پاسخ کوتاه دهید.
 - الف) در صورتی که یکی از ورودی‌های softmax به بی‌نهایت میل کند خروجی را محاسبه کنید.
 - ب) فرض کنید یک مدل Perceptron با ورودی‌های w_1 و w_2 داریم. آیا می‌توانید وزن‌های مدل را طوری تعیین کنید که خروجی آن برابر xor ورودی‌ها باشد؟ این موضوع را درباره‌ی and و or نیز بررسی کنید.
۲. (۳۰ نمره) در این سوال قصد داریم با استفاده از روش Naive Bayes احتمال spam بودن یا نبودن ایمیل‌ها بر اساس وجود کلماتی مشخص در آن‌ها را بررسی کنیم. به همین منظور مجموعه داده زیر را در نظر بگیرید که مربوط به ۱۰ ایمیل مختلف است. با استفاده از روش Naive Bayes و داده‌های موجود به سوالات زیر را پاسخ دهید:

Email	w_1	w_2	w_3	w_4	w_5	Is Spam
E_1	1	0	0	1	1	Yes
E_2	1	1	0	0	1	Yes
E_3	0	1	1	1	0	No
E_4	1	1	1	0	1	Yes
E_5	0	0	0	1	1	No
E_6	1	0	1	1	1	No
E_7	1	1	0	1	1	Yes
E_8	0	1	0	0	1	Yes
E_9	0	0	1	0	0	No
E_{10}	1	0	0	0	1	Yes

الف) آیا مدل ایمیل یازدهم را spam تشخیص می‌دهد؟

Email	w_1	w_2	w_3	w_4	w_5	Is Spam
E_{11}	1	1	1	1	1	?

ب) حال فرض کنید می‌دانیم در ایمیل یازدهم با وضعیت وجود کلمات زیر، ایمیل spam خواهد بود. با استفاده از Naive Bayes , spam بودن آن را بررسی کنید. مشاهده خواهید کرد نتیجه بدست آمده با واقعیت تطبیق ندارد. دلیل این عدم تطابق را توضیح دهید.

Email	w_1	w_2	w_3	w_4	w_5	Is Spam
E_{11}	1	1	0	0	0	?

ج) آیا با استفاده از روش laplace smoothing مشکل قسمت قبل حل می‌شود؟ با اعمال این روش احتمال قسمت ب را دوباره محاسبه کنید.

۳. (۳۰ نمره) روش logistic regression برای یک مسئله binary classification را در نظر بگیرید که شامل دو متغیر x_1 و x_2 است. می‌دانیم احتمال قرارگیری در دسته اول ($y=1$) به صورت زیر تعریف می‌شود.

$h_{\theta}(x) = P(y = 1|x_1, x_2) = \sigma(z) = \frac{1}{1+e^{-z}}$ where $z = w_1x_1 + w_2x_2 + w_0$
همچنین می‌دانیم تابع هزینه برای یک داده خاص به صورت زیر تعریف می‌شود.

$$Cost(y, h_{\theta}(x)) = -y \ln(\sigma(z)) - (1 - y) \ln(1 - \sigma(z))$$

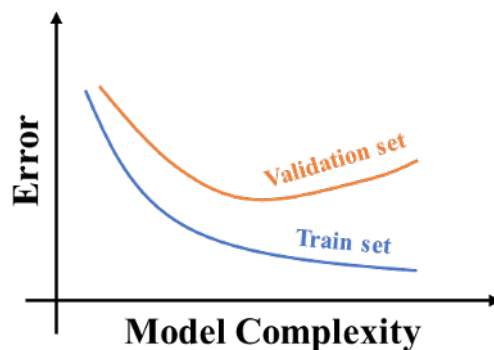
حال فرض کنید به جای استفاده از این تابع هزینه، از تابع زیر استفاده کنیم.

$$Cost(y, h_{\theta}(x)) = (y - h_{\theta}(x))^2$$

الف) مشتق تابع cost را نسبت به پارامترهای مدل بدست بیاورید سپس نحوه‌ی آپدیت پارامترها برای یک مجموعه داده در هر مرحله از الگوریتم gradient descent را بنویسید.

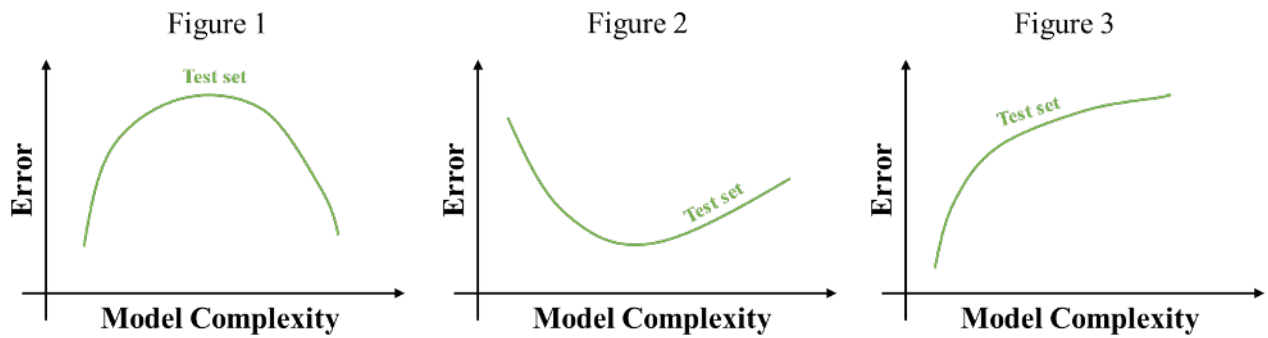
ب) چرا در واقعیت از تابع cost اول به جای دوم استفاده می‌کنیم.

۴. (۱۵ نمره) نمودار زیر خطای مدل آموزش دیده بر روی داده‌های train و validation را نشان می‌دهد. محور افقی نمودار پیچیدگی مدل انتخاب شده را نشان می‌دهد. با توجه به این نمودار به سوالات زیر پاسخ دهید.



الف) مفاهیم overfitting و underfitting را به اختصار توضیح دهید و بگویید هر مفهوم در کدام ناحیه از نمودار بالا قبل مشاهده است.

ب) با توجه به نمودار بالا انتظار می‌رود خطای مدل بر روی داده‌های تست شبیه به کدام یک از نمودارهای زیر باشد؟ دلیل خود را شرح دهید.



۵. (۱۵ نمره) فرض کنید پس از آموزش دسته‌بند خود متوجه شده‌اید که مدل روی داده تست دقت خوبی ندارد در هر مورد مشخص کنید که آیا انجام این کار می‌تواند مشکل را حل کند و یا مشکل را بدتر می‌کند و یا تاثیر خاصی ندارد.
- الف) افزایش داده‌ی آموزش
 - ب) کاهش پارامتر regularization
 - ج) افزایش پارامتر regularization
 - د) استخراج ویژگی‌های بهتر از داده‌ها

سوالات عملی (۱۳۵+۱۵ نمره)

۱. (۱۳۵+۱۵ نمره) برای حل سوالات عملی به دفترچه سوالات عملی مراجعه کنید.