



یادگیری ماشین برای بیوانفورماتیک

بهار ۱۴۰۰

اساتید: دکتر سلیمانی، دکتر شریفی زارچی

تمرین چهارم

شبکه‌های عصبی عمیق

مهلت ارسال: ۱۵ خرداد

• مهلت ارسال پاسخ تا ساعت 23:59 روز مشخص شده است.

• هم‌کاری و هم‌فکری شما در انجام تمرین مانعی ندارد اما پاسخ ارسالی 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 + ۲۵ نمره)

۱. (۵ + ۱۰ نمره) می‌خواهیم حاصل کانولوشن دو تصویر x و y ، هر یک به ابعاد $N \times N$ را محاسبه کنیم؛ به این صورت که

$$(x \otimes y)[m, n] = \sum_{m'=0}^{N-1} \sum_{n'=0}^{N-1} x[m', n'] y[m - m', n - n']$$

(آ) (۵ نمره) مطلوب است محاسبه‌ی مرتبه زمانی محاسبه‌ی این مقدار.

(ب) (۱۰ نمره امتیازی) با معرفی محاسبات میانی، راه‌کاری پیشنهاد کنید که بتوان با آن مقدار فوق را در مرتبه زمانی $O(N^2 \log(N))$ محاسبه کرد و جایگیری راه‌کار خود در این مرتبه زمانی را اثبات کنید. بعضاً از فرمولیشن‌های این چینی برای کاهش هزینه‌های محاسباتی شبکه‌های کانولوشنال استفاده می‌شود. (راهنمایی: از تبدیل فوریه استفاده کنید)

۲. (۲۵ نمره) می‌خواهیم بر دادگان $S = \{(x_i, y_i) | x_i \in \mathbb{R}^d, y_i \in \{0, 1\}\}$ و $|S| = n$ دسته‌بند خطی $(\hat{y} = w^T x)$ به منظور کمینه‌سازی تابع هزینه $J(w) = \frac{1}{n} \|Xw - Y\|_2^2$ را بیابیم که X در آن ماتریسی بوده که سطرهايش را بردارهای x_i تشکیل داده‌اند و Y برداری ستونی باشد که در آن $Y(i) = y_i$ باشد.

(آ) (۷ نمره) بر حسب اندازه‌ی n ، در چه شرایطی حتما پاسخی برای صفرکردن تابع هزینه وجود دارد؟

(ب) (۸ نمره) تصور کنید W مجموعه‌ی بردارهای w ای باشند که J را کمینه می‌کنند، نشان دهید $\dim(W)$ کوچک‌تر یا مساوی n است.

(ج) (۱۰ نمره) فرض کنید برای حل این مسئله از الگوریتم بهینه‌سازی Stochastic Gradient Descent بهره‌گیریم. نشان دهید پاسخ‌نهایی به دست آمده از این الگوریتم در صورتی که $\dim(W) > 0$ و $\text{rank}(X) = n$ و بردار $w_0 = 0$ باشد، w_* ای بوده که از میان فضای n بعدی جواب کمترین نرم اقلیدسی را دارد. (این مسئله به SGD's Implicit Regularization مشهور است.)

۳. (۱۵ نمره) در حالت کلی نشان دهید خروجی هر شبکه‌ی عصبی با توابع فعال‌سازی خطی (و m لایه‌ی مخفی) را می‌توان با شبکه‌ای عصبی دیگر بدون لایه‌ی مخفی تشکیل داد.

۴. (۱۵ نمره) می‌دانید که هر رابطه‌ی منطقی را می‌توان به کمک شیوه‌های گوناگون چون جدول کارنو به قالب جمعی از ضرب‌ها یا Sum of Products نوشت، با توجه به این موضوع به سوالات زیر پاسخ دهید.

(آ) (۵ نمره) نشان دهید که چگونه می‌توان هر رابطه‌ی منطقی را با یک شبکه‌ی عصبی ۲ لایه (یک لایه مخفی) و تابع فعال‌ساز $ReLU$ نمایش داد.

(ب) (۵ نمره) یک شبکه‌ی عصبی دو لایه (یک لایه مخفی) با تعداد پارامترهایی از مرتبه‌ی 2^n برای نمایش تابع XOR با n ورودی پیشنهاد دهید.

(ج) (۵ نمره) می‌توان با افزایش عمق شبکه، تعداد پارامترهای لازم برای نمایش تابع هدف را کاهش داد. بدین منظور یک شبکه‌ی عصبی برای نمایش تابع XOR با n ورودی و تعداد پارامترهایی از مرتبه‌ی n و لایه‌هایی از مرتبه‌ی $\log n$ پیشنهاد دهید.

۵. (۱۵ نمره) مزایا و معایب استفاده از هر یک از توابع فعال‌سازی sigmoid و tanh و ReLU را بیان کنید و این توابع را از نظر هزینه محاسباتی و موضوع vanishing gradient بررسی کنید.

۶. (۱۵ نمره) قصد داریم شبکه‌ای طراحی کنیم که با ورودی گرفتن سه عدد صحیح، تعیین کند که کدام یک بزرگ‌تر است. اگر x_1, x_2 و x_3 سه ورودی و y_1, y_2 و y_3 سه خروجی شبکه باشند، ضمن طراحی شبکه‌ای با عمق و عرض دلخواه و تابع فعال‌سازی ReLU برای این کار، وزن‌ها و بایاس‌های شبکه را به گونه‌ای تعیین کنید که اگر x_i بزرگ‌ترین عدد ورودی باشد، مقدار y_i عددی مثبت و در غیر این صورت برابر صفر شود. همچنین در حالتی که بیش از یک عدد بیشینه باشد انتظار داریم مقدار خروجی متناظر با هر بیشینه مقداری مثبت شود.

۷. (۱۰ نمره) توضیح دهید که چگونه می‌توان L_2 Decay را تحت بهینه‌سازی SGD برابر با Weight Decay تصور کرد. (راهنمایی)

۸. (۱۵ نمره امتیازی) امروزه شبکه‌های کانولوشنال برای انواع و اقسام مسائل مربوط به بینایی ماشین به صورت همه‌جانبه استفاده شده و همه‌جا وجود دارند. این در حالیست که درک دقیقی از آنکه چگونه در مورد تصاویر ورودی خود تصمیم‌گیری کرده و برای تخمین‌های خود به کدام بخش از تصویر توجه می‌کنند وجود ندارد. هرچند مصورسازی فیچرهای تشخیص‌داده‌شده در لایه‌ی نخستین آنها ممکن است، اعمال این کار برای لایه‌های میانی شبکه سود چندانی نداشته و به ادراک عمل‌کرد آنها کمک نمی‌کند. بدین منظور ابزارهای گوناگونی برای بررسی عمل‌کرد درونی این شبکه‌ها پیشنهاد شده و به کار گرفته می‌شود. به انتخاب خود از میان مقالات زیر دو ابزار را انتخاب کرده و راه‌کار پیشنهادی آنها را به صورت خلاصه توصیف کنید.

Guided Grad-Cam •

Integrated Gradients •

Guided Backpropagation •

DeepLIFT •

Deconvolution •

SHAP •

سوالات عملی (۱۰۰ نمره)

۱. (۱۰۰ نمره) در این سوال قصد داریم علاوه بر پیاده‌سازی یک CNN در فریم‌ورک pytorch، به بررسی یادگیری انتقالی (Transfer Learning) که یکی از مهم‌ترین روش‌های حال حاضر است و باعث تسریع فرآیند یادگیری به خصوص در مسائل مربوط به تصاویر می‌باشد. در پایان نیز به بررسی تفسیر پذیری شبکه‌های عصبی با استفاده از روش GradCAM خواهیم پرداخت که یکی از روش‌های معروف برای بدست آوردن Attention بدون تغییر ساختار شبکه است. از تصویر بدست آمده در این مرحله می‌توان استفاده‌های زیادی کرد، مانند: بررسی اینکه آیا شبکه برای تصمیم‌گیری به همان نقاطی نگاه میکند که انسان نگاه میکند و یا حتی کشف ویژگی‌های جدیدی برای پیش‌بینی. دیتاست مورد استفاده در این تمرین مربوط به تصاویر MRI از مغز می‌باشد و هدف تشخیص وجود یا عدم وجود تومور مغزی است. برای پاسخ‌دهی به این سوال به نوت‌بوک مربوطه مراجعه کنید.