



هوش مصنوعی

بهار ۱۴۰۰

استاد: محمدحسین رهبان

دانشگاه صنعتی شریف

دانشکده مهندسی کامپیوتر

مباحث فصل پنج تا هفت آزمون پایانترم مهلت ارسال: ۸ تیر، ساعت ۱۳

- مهلت ارسال پاسخ تا ساعت ۱۳:۰۰ سه‌شنبه ۸ تیرماه است. هیچ ارسالی پس از این زمان پذیرفته نخواهد شد.
- هر گونه هم‌فکری ممنوع بوده و پاسخ شما باید کاملاً حاصل تفکر و به نگارش خودتان باشد.
- امتحان به صورت کتاب و اینترنت باز است، با این حال جواب همه‌ی سوالات باید به بیان خودتان بوده و مشاهده‌ی مشابهت‌های غیر عادی به منظره‌ی تقلب در نظر گرفته خواهد شد. همچنین منابع استفاده شده برای پاسخ‌دهی به هر یک از سوال (در صورت وجود) باید مشخصاً ذکر شود.
- لطفاً تصویری واضح از پاسخ سوالات خود بارگذاری کنید. در غیر این صورت پاسخ شما تصحیح نخواهد شد.
- امتحان در مجموع شامل ۵ نمره‌ی امتیازی بوده و دریافت ۷۰ نمره از ۷۵ نمره‌ی آن برای دریافت نمره‌ی کامل کفایت می‌کند.

مسائل (۷۰+۵ نمره)

۱. (۱۲ نمره) سه جعبه داریم که در هر کدام مطابق جدول پایین تعدادی گوی سفید و سیاه قرار دارند. در هر مرحله یک جعبه به ما داده می‌شود، ما هم بدون نگاه کردن به داخل جعبه یک گوی برمی‌داریم و رنگ آن را یادداشت می‌کنیم و گوی را در جعبه قرار می‌دهیم. سپس بر اساس جدول احتمالات داده شده جعبه‌ی بعدی داده می‌شود و این روند ادامه می‌یابد. توجه کنید که در هر مرحله نمی‌دانیم چه جعبه‌ای به ما داده شده است و تنها رنگ گویی که انتخاب می‌کنیم را می‌دانیم.

	جعبه‌ی سه	جعبه‌ی دو	جعبه‌ی یک
سیاه	۴	۱	۲
سفید	۳	۳	۲

	۱	۲	۳
۱	۰.۳	۰.۶	۰.۵
۲	۰.۲	۰.۱	۰.۵
۳	۰.۵	۰.۳	۰

آ) (۴ نمره) مدل HMM را با مشخص کردن متغیرهای مشاهده شده و پنهان رسم کنید.

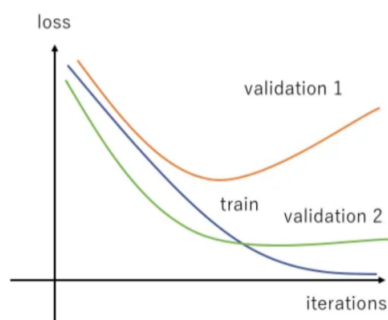
ب) (۸ نمره) با فرض اینکه اولین جعبه با احتمالات زیر انتخاب شود و زنجیره

$$E_0 = 0, E_1 = 1, E_2 = 1$$

از گوی‌ها را مشاهده کرده باشیم، از الگوریتم Viterbi استفاده کنید و محتمل‌ترین زنجیره از جعبه‌ها را بدست آورید. (در زنجیره عدد ۰ به معنای رنگ سفید است و عدد ۱ به معنای رنگ سیاه)

جعبه‌ی سه	جعبه‌ی دو	جعبه‌ی یک
۰.۶	۰.۲	۰.۲

۲. (۱۲ نمره) فرض کنید دو مجموعه‌ی ارزیابی (validation) برای بررسی آموزش مجموعه‌ای در اختیار داریم. منحنی زیر مقدار loss را برای این دسته‌ها نشان می‌دهد. برای هر یک از حالات زیر تفسیری از منحنی ارائه دهید و برای اصلاح وضعیت راه حلی ارائه دهید:



- (۶ نمره) دو مجموعه‌ی validation از توزیع یکسان آمده‌اند.
- (۶ نمره) دو مجموعه‌ی validation از توزیع متفاوت آمده‌اند.

۳. (۱۰ نمره) مجموعه نمونه‌های آموزشی زیر را در نظر بگیرید.

A	B	Class
True	True	+
True	True	+
False	True	-
False	False	+
True	False	-
True	False	-

- (۵ نمره) آنتروپی این مجموعه از نمونه‌های آموزشی با توجه به دسته‌بندی تابع هدف چقدر است؟
- (۵ نمره) بهره‌ی اطلاعات ویژگی B برای این نمونه‌های آموزشی چقدر است؟

۴. (۱۰ نمره) یک پزشک قصد دارد با استفاده از یک شبکه عصبی بیمار بودن یا نبودن افراد را پیش‌بینی کند. او برداری از ویژگی‌ها (x) را به ورودی شبکه داده و در صورت تشخیص اشتباه شبکه، جریمه‌ای به آن اختصاص می‌دهد. اگر شبکه یک فرد بیمار را اشتباهاً سالم پیش‌بینی کند به اندازه c_{FN} و اگر فرد سالم را به عنوان فرد بیمار پیش‌بینی کند به اندازه c_{FP} جریمه می‌شود. پیش‌بینی درست نیز هزینه‌ای ندارد. فرض کنید شبکه استفاده شده یک شبکه بدون لایه‌های پنهان و تابع activation سیگموئید در لایه خروجی است. در واقع:

$$\hat{y} = \sigma(\mathbf{w} \cdot \mathbf{x})$$

(آ) (۵ نمره) رابطه‌ی تابع هزینه را به طور ریاضی بنویسید (داده‌ها شامل n نمونه از $\mathbf{x}^{(1)}$ تا $\mathbf{x}^{(n)}$ هستند). آیا این تابع برای آموزش شبکه عصبی مناسب است؟ چرا؟

(ب) (۵ نمره) فرض کنید یک تابع جایگزین به صورت زیر تعریف شود:

$$E(\mathbf{w}) = c_{FN} \sum_{i|y^{(i)}=1} (\hat{y}^{(i)} - y^{(i)})^2 + c_{FP} \sum_{i|y^{(i)}=0} (\hat{y}^{(i)} - y^{(i)})^2$$

این تابع چه چیزی را اندازه‌گیری می‌کند؟ گرادین آن را نسبت به \mathbf{w} محاسبه کنید.

۵. (۱۲ نمره) یک لایه را در یک شبکه CNN در نظر بگیرید که یک feature map با ابعاد 100×100 ورودی می‌گیرد (به طور مثال یک عکس سیاه و سفید). در هرکدام از حالات زیر، تعداد ویژگی‌هایی که باید بهینه شوند را محاسبه کنید. در محاسبات خود بایاس را نیز در نظر بگیرید.

- (آ) (۳ نمره) لایه مورد بررسی یک لایه convolution است که اندازه فیلتر آن برابر 1×1 است و طول گامی (stride) ۱ دارد که ۱۰۰ عدد feature map خروجی می‌دهد.
- (ب) (۳ نمره) لایه مورد بررسی یک لایه convolution است که اندازه فیلتر آن 10×10 است و طول گامی (stride) مساوی با ۵ دارد که ۵۰ عدد feature map خروجی می‌دهد.
- (ج) (۳ نمره) لایه‌ی مورد نظر یک لایه‌ی max-pooling است که اندازه‌ی فیلتر آن 2×2 می‌باشد و دارای گام (stride) ۲ می‌باشد.
- (د) (۳ نمره) لایه مورد بررسی یک لایه fully-connected است که دارای ۱۰ نرون می‌باشد.

۶. (۱۹ نمره) یک عدد بین ۱ تا n به طور یکنواخت انتخاب شده‌است. در هر گام می‌توانیم یک حدس بزنیم و به ما گفته می‌شود که عدد ما بزرگتر، کوچکتر یا برابر با عدد انتخاب شده است. سعی داریم با کمترین تعداد حدس، عدد را پیدا کنیم.

- (آ) (۶ نمره) مدل MDP این مسئله را به ازای $n = 3$ رسم کنید. (۷ حالت متفاوت)
- (ب) (۶ نمره) سیاستی را در نظر بگیرید که اعداد را به ترتیب از بزرگ به کوچک حدس می‌زند. این سیاست را ارزیابی کنید. به طور دقیق شما باید توابع V_1^π ، V_2^π و V_3^π حساب کنید. (با شروع از هر حالتی با این استراتژی پس از سه حرکت به یک وضعیت نهایی می‌رسیم، لذا می‌توانید بررسی کنید که سه مرحله برای محاسبه‌ی V^π کافیست.)
- (ج) (۴ نمره) با توجه به مقادیر به دست آمده، سیاست را بهبود دهید.
- (د) (۳ نمره) سیاست جدید را ارزیابی کرده و نشان دهید بهینه است.
- در تمام قسمت‌های سوال ارزش‌های اولیه صفر بوده و مقدار γ برابر یک می‌باشد.