



آزمون میان‌ترم

بخش اول آزمون (۳۰ دقیقه و ۳۶ امتیاز)

استفاده از اینترنت و ابزارهای هوشمند مانند ChatGPT به هیچ‌وجه مجاز نیست.

۱ پرسش یک

درستی یا نادرستی هریک از گزاره‌های زیر را با ذکر دلیل مشخص کنید.

۱. طبقه‌بند بیز، بهترین طبقه‌بندی است که می‌توان برای جداسازی یک مسئله دوکلاسه طراحی کرد.
۲. استفاده از رویکرد بیز برای تخمین پارامترهای توزیع، می‌تواند مانع از بیش‌برازش (Overfitting) شود.
۳. استفاده از معیار Information Gain برای ساخت درخت در شرایطی که بعضی از ویژگی‌ها حالات زیادی دارند، مناسب نیست.
۴. هر شبکه عصبی چندلایه با توابع فعال‌ساز خطی در لایه‌های پنهان می‌تواند به عنوان یک شبکه عصبی بدون هیچ لایه پنهانی نمایش داده شود.

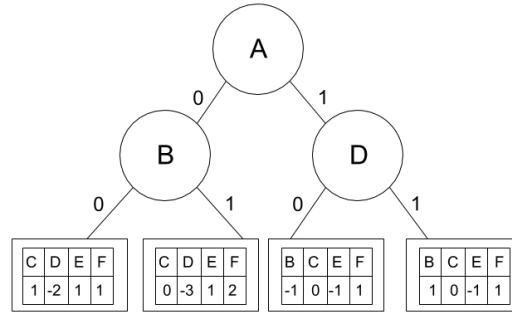
۲ پرسش دو

برای بهره‌برداری از ویژگی‌های مطلوب طبقه‌بندهای درخت تصمیم و پرسپترون، سارا الگوریتم جدیدی به نام «درخت پرسپترون» ایجاد کرده که ویژگی‌های هر دو را ترکیب می‌کند. درخت‌های پرسپترونی شبیه به درخت‌های تصمیم هستند؛ اما هر leaf node به جای مکانیزم رأی اکثریت، شامل یک پرسپترون است.

برای ایجاد یک درخت پرسپترون؛ اولین مرحله، اجرای یک الگوریتم یادگیری درخت تصمیم معمولی (مانند ID3) و انجام تقسیم‌بندی بر اساس ویژگی‌ها تا رسیدن به عمق حداکثر مشخص شده است. هنگامی که به عمق حداکثر می‌رسیم، در هر leaf node، یک پرسپترون روی ویژگی‌های باقی‌مانده که هنوز در آن شاخه استفاده نشده‌اند، آموزش داده می‌شود. دسته‌بندی یک نمونه جدید از طریق مراحل مشابهی انجام می‌شود. ابتدا نمونه از طریق درخت تصمیم بر اساس مقادیر ویژگی‌هایش گذر می‌کند. وقتی به یک leaf node می‌رسد، پیش‌بینی نهایی با اجرای پرسپترون متناظر در آن گره انجام می‌شود.

فرض کنید که دارای مجموعه داده‌ای با ۶ ویژگی دودویی $\{A, B, C, D, E, F\}$ و دو برچسب خروجی $\{-1, 1\}$ هستید. یک درخت پرسپترون با عمق ۲ روی این مجموعه داده‌ها در شکل ۱ آمده است. وزن‌های پرسپترون نیز در leaf node‌ها آمده است (فرض کنید که بایاس برای هر پرسپترون $b = 1$ است).

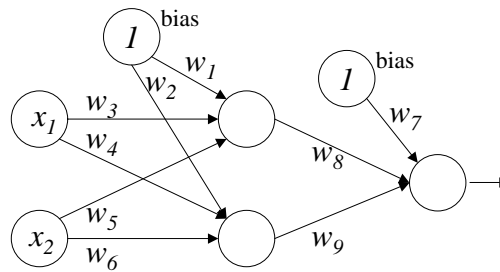
۱. برای نمونه $x = [1, 1, 0, 1, 0, 1]$ ، درخت پرسپترون داده‌شده چه برچسب خروجی‌ای را پیش‌بینی می‌کند؟
۲. آیا مرزتصمیم درخت پرسپترون همواره خطی است؟ برای مقادیر کوچک حداکثر عمق، کیفیت آموزش درخت تصمیم و درخت پرسپترون را با ذکر دلیل مقایسه کنید. آیا تفاوتی دارند؟



شکل ۱: درخت پرسپترون با عمق دو.

۳ پرسش سه

شبکه عصبی آورده شده در **شکل ۲** را برای یک مسئله طبقه بندی دو کلاسه در نظر بگیرید. فرض کنید که لایه های میانی از تابع فعال ساز خطی ($h(z) = cz$) و لایه خروجی از تابع سیگموئید ($g(z) = \frac{1}{1+e^{-z}}$) استفاده می کند. این شبکه می خواهد یک تابع برای $P(Y = 1 | X, w)$ که در آن $X = (x_1, x_2)$ و $W = (w_1, w_2, \dots, w_9)$ است را یاد بگیرد.



شکل ۲: شبکه عصبی سوال سوم.

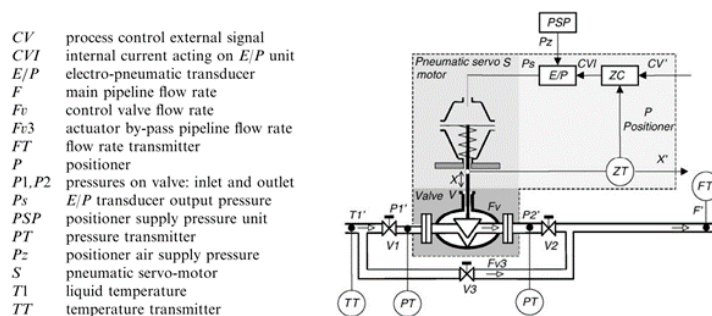
۱. خروجی شبکه عصبی $P(Y = 1 | X, w)$ را بر حسب پارامترهای شبکه (W, x) و ثابت c نوشته و مرز تصمیم نهایی را به دست آورید.
۲. آیا می توان یک شبکه عصبی بدون لایه مخفی به دست آورد که معادل شبکه عصبی فوق باشد؟ در صورت وجود، شبکه پیشنهادی تان را رسم کنید.

بخش دوم آزمون (۱۱۰ دقیقه و ۶۴ امتیاز)

- استفاده «متعارف» از هر وسیله و کمکی (غیر از موجود زنده!) مجاز است.
- لازم است که پیوند گوگل کولب و یا پیوند گوگل درایو فایل IPYNB، کدهای خود را در حالتی که خروجی سلول کدها کاملاً مشخص باشد، در ابتدای گزارش خود قرار دهید.
- تنها یک فایل PDF برای گزارش و یک فایل IPYNB، برای کدها قابل پذیرش است.
- کد شما تنها در صورتی بررسی می‌شود که به نتیجه آن در گزارش خود به‌صورت صریح و مشخص و به‌همراه توضیحات کافی اشاره کرده باشید.
- توصیه می‌شود سوالات مربوط به این بخش را یک‌بار از ابتدا تا انتها به‌صورت دقیق مطالعه بفرمایید و سپس به حل سوالات بپردازید.

۴ پرسش چهار

مجموعه داده DAMADICS که مربوط به سیستم آورده شده در شکل ۳ است را در نظر بگیرید.



شکل ۳: سیستم مجموعه داده DAMADICS.

فایل‌های این مجموعه‌داده را از طریق [گوگل درایو](#) دانلود کنید. این داده‌ها **۱۷** **بُعد** دارند (ستون اول شمارهٔ داده است) که در طول دو روز از سیستم اخذ شده‌اند. در **جدول ۱** مشخصات داده‌های اخذشده آورده شده است.

جدول ۱: مشخصات داده‌های اخذشده (اگر کنجکا و شدید، می‌توانید زوم کنید:)).

Column	Actuator	Variable Symbol	Variable Description	Range	Units
1	1	-	Time stamp	0 – 86399	s
2	1	P51_05	P1 - juice pressure (valve inlet)	0 – 1000	kPa
3	1	P51_06	P2 - juice pressure (valve outlet)	0 – 1000	kPa
4	1	T51_01	T - juice temperature (valve outlet)	50 – 150	°C
5	1	F51_01	F - juice flow (1 st evaporator inlet)	0 – 500	m ³ /h
6	1	LC51_03CV	CV - control value (controller output)	0 – 100 ²	%
7	1	X51_03X	X - servomotor rod displacement	0 – 100 ²	%
8	1	LC51_03PV	PV - process value (juice level in 1 st evaporator)	0 – 100 ³	%
9	1	TC51_05	Juice temperature (1 st evaporator inlet)	50 – 150	°C
10	1	T51_08	Juice temperature (1 st evaporator outlet)	50 – 150	°C
11	1	D51_01	Juice density (1 st evaporator inlet)	0 – 25	Bx
12	1	D51_02	Juice density (1 st evaporator outlet)	13 – 41	Bx
13	1	F51_02	Steam flow	1 – 100	t/h
14	1	PC51_01	Steam pressure	100 – 300	kPa
15	1	T51_06	Steam temperature	50 – 150	°C
16	1	P51_03	Vapour pressure	0 – 250	kPa
17	1	T51_07	Vapour temperature	50 – 150	°C

در تاریخ ۹ نوامبر ۲۰۰۱ عیوبی در زمانی مشخص و مطابق با مشخصات **جدول ۲** به سیستم وارد شده است. در تاریخ ۱۷ نوامبر ۲۰۰۱ مجدداً همان عیوب در سیستم رخ داده‌اند.

جدول ۲: مشخصات داده‌های روز ۹ نوامبر ۲۰۰۱.

Item	Fault tag	Sample	Date	Fault description
1	f18	58800 – 59800	October 30, 2001	Partly opened bypass valve
2	f16	57275 – 57550	November 9, 2001	Positioner supply pressure drop
3	f18	58830 – 58930	November 9, 2001	Partly opened bypass valve
4	f18	58520 – 58625	November 9, 2001	Partly opened bypass valve
5	f18	54600 – 54700	November 17, 2001	Partly opened bypass valve
6	f16	56670 – 56770	November 17, 2001	Positioner supply pressure drop

فرض کنید داده‌های ۹ نوامبر در اختیار شماست و می‌خواهید برای داده‌های ۱۷ نوامبر سیستم تشخیص عیب مبتنی بر شبکه‌های عصبی طراحی کنید تا بتواند عیوب را تشخیص داده و تفکیک کند. این شبکه را طراحی کنید و فرآیند اجرایی و نتایجی که گزارش می‌کنید حتماً شامل تمام موارد زیر باشد و به‌صورت صریح و مشخص، پاسخ نهایی‌ای که به هر مورد مطرح‌شده اشاره دارد را مشخص کنید:

۱. تعیین و انجام تقسیم‌بندی مناسب برای داده‌ها (آموزش، اعتبارسنجی و آزمون)، و تعیین و انجام روش‌های پیش‌پردازشی مناسب (از بین مواردی هم‌چون: نرمال‌سازی، استانداردسازی، حذف یا جایگزینی داده‌های خالی، حل مشکل عدم توازن در کلاس‌ها و... هرکدام که لازم می‌دانید را انجام دهید).
۲. لازم است که در گزارش خود گزارشی از تعداد و توزیع درصدی داده‌ها در قسمت‌های مختلف (آموزش، اعتبارسنجی و آزمون) و در کلاس‌های مختلف را نمایش دهید.
۳. استفاده از روش‌های آموخته‌شده برای استخراج ویژگی، انجام حداقل ۵ روش آن، رسم ماتریس هم‌بستگی، تحلیل آن و حذف برخی ویژگی‌ها در صورت لزوم (با توضیح). استفاده از روش‌های فرکانسی مانند FFT نمره امتیازی دارد.
۴. رسم شبکه عصبی و کل ساختار پیشنهادی از ابتدا تا انتها در قالب یک بلوک‌دیگرام و مشخص‌کردن مواردی مانند تعداد لایه‌های پنهان، توابع فعال‌ساز لایه‌های میانی و لایه خروجی، تابع اتلاف و بهینه‌ساز انتخابی و... به‌صورت آیت‌بندی‌شده و مشخص.
۵. استفاده از حداقل یکی از تکنیک‌های لایه دورریز^۱، برنامه‌ریز نرخ یادگیری^۲ و توقف زودهنگام^۳ الزامی است.
۶. رسم نمودار تابع اتلاف و دقت در فرآیند آموزش برای داده‌های آموزش و اعتبارسنجی برای حداقل دو حالت مختلف از شبکه (مثلاً تغییر در تعداد لایه‌های پنهان و یا تغییر در بهینه‌ساز).
۷. دقت داشته باشید که فقط یک عدد نمودار خروجی برای تابع اتلاف مجاز است. یعنی تمام مقایسه‌های مربوط به توابع اتلاف روی یک نمودار واحد و با رنگ‌های مختلف انجام شود.
۸. رسم ماتریس درهم‌ریختگی به‌صورت درصدی برای داده‌های آزمون و نمایش [classification_report](#) به تفکیک هر کلاس.
۹. تکرار مرحله پیاده‌سازی با استفاده از یکی از روش‌های K-Fold Cross-validation یا Stratified K-Fold Cross-validation و نمایش ماتریس درهم‌ریختگی به ازای هر فُلد و به‌صورت کلی.
۱۰. اوزان مربوط به یکی از مدل‌های خود (ترجیحاً بهترین مدل) را ذخیره کنید و روی گوگل‌درایو بارگذاری کنید. سپس، در یک سلول کد دستوراتی بنویسید که فایل وزن را با استفاده از [gdown](#) فراخوانی کند و فرآیند تست را روی داده نمونه تست که آن هم با [gdown](#) فراخوانی شده اجرا کند.
۱۱. پیاده‌سازی تمام موارد فوق ممکن بود؟ کدام یک عملی نبود و انجام ندادید؟ چرا؟ اگر به هر دلیلی نتوانستید پیاده‌سازی را با استفاده از دیتاست خواسته‌شده در این سوال انجام دهید، می‌توانید پیاده‌سازی را با استفاده از یکی از دیتاست‌هایی که در طول درس آموزش دیدید انجام دهید و حداکثر ۴۰ درصد از نمره این سوال را دریافت کنید. تشخیص میزان نمره دریافتی بر اساس دیتاست انتخابی شما و برعهده مصحح خواهد بود.

¹Dropout Layer

²Learning Rate Scheduler

³Early Stopping