

به نام خدا



درس هوش مصنوعی و سیستم‌های خبره

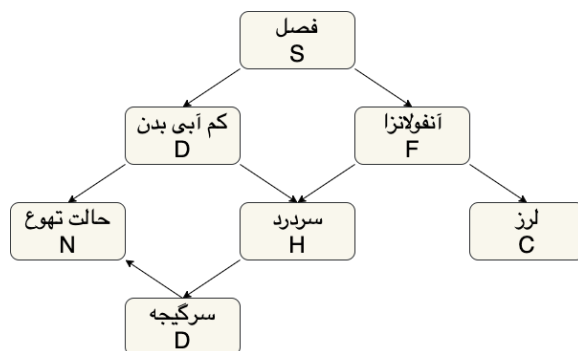
تمرین هشتم

طراحان: محمدامین چینی‌فروشان، مرتضی شهبابی
مدرس درس: جناب آقای دکتر محمدی

مهلت ارسال: ۱۴۰۱/۰۸/۰۹

سوال ۱:

هفت متغیر تصادفی فصل (S)، کم آبی بدن (D)، آنفولانزا (F)، سردرد (H)، لرز (C)، حالت تهوع (N)، سرگیجه (Z) را همراه با گراف جهت دار آن در شکل زیر در نظر بگیرید:



گراف جهت دار متغیرها

همچنین مقادیر احتمالات شرطی میان این متغیرها به صورت زیر است:

جدول مقادیر احتمالات شرطی متغیرها

$P(S = \text{پاییز})$		$P(S = \text{بهار})$	
0.5		0.5	
	$P(F = \text{true} S)$	$P(F = \text{false} S)$	
$S = \text{پاییز}$	0.4	0.6	
$S = \text{بهار}$	0.1	0.9	
	$P(D = \text{true} S)$	$P(D = \text{false} S)$	
$S = \text{پاییز}$	0.1	0.9	
$S = \text{بهار}$	0.3	0.7	
	$P(C = \text{true} F)$	$P(C = \text{false} F)$	
$F = \text{true}$	0.8	0.2	
$F = \text{false}$	0.1	0.9	
	$P(H = \text{true} F, D)$	$P(H = \text{false} F, D)$	
$F = \text{true}, D = \text{true}$	0.9	0.1	
$F = \text{true}, D = \text{false}$	0.8	0.2	
$F = \text{false}, D = \text{true}$	0.8	0.2	
$F = \text{false}, D = \text{false}$	0.3	0.7	
	$P(Z = \text{true} H)$	$P(Z = \text{false} H)$	
$H = \text{true}$	0.8	0.2	
$H = \text{false}$	0.2	0.8	
	$P(N = \text{true} D, Z)$	$P(N = \text{false} D, Z)$	
$D = \text{true}, Z = \text{true}$	0.9	0.1	
$D = \text{true}, Z = \text{false}$	0.8	0.2	
$D = \text{false}, Z = \text{true}$	0.6	0.4	
$D = \text{false}, Z = \text{false}$	0.2	0.8	

حال احتمال هر یک از رخداد‌های زیر را محاسبه کنید.

الف) $P(F = \text{true})$

ب) $P(F = \text{true} \mid S = \text{پاییز})$

ج) $P(F = \text{true} \mid S = \text{پاییز}, H = \text{true}, D = \text{true})$

سوال ۲:

مجموعه داده‌ی Iris یکی از معروف‌ترین دیتاست‌های داده کاوی است. در این دیتاست اطلاعات مختلفی در مورد گیاهان از جمله طول گلبرگ و کاسبرگ و ... آورده شده‌است و با استفاده از این اطلاعات، سعی می‌شود تا نوع گیاه را از بین سه دسته گیاه موجود در این دیتاست تشخیص بدهند. در این سوال هدف انجام برخی اصلاحات روی دیتاست و به دست آوردن برخی احتمالات مربوط به آن است.

الف) برای آشنایی بیشتر با دیتاست، می‌توانید لینک زیر را مطالعه کنید.

<https://www.kaggle.com/datasets/uciml/iris>

ب) به نوت بوک iris.ipynb مراجعه کنید و عملیات‌های مختلفی که روی دیتاست انجام شده را مطالعه کنید.

ج) دو احتمال زیر را محاسبه کنید:

- احتمال اینکه PetalWidth در دسته‌ی خیلی کوچک باشد، به شرط آنکه نوع گل از نوع Iris-setosa باشد ($P(\text{PetalWidth} = \text{XS} \mid \text{Species} = \text{Iris-setosa})$).

- احتمال اینکه نوع گل از دسته‌ی Iris-virginica باشد به شرط آنکه SepalLength آن در دسته بزرگ باشد ($P(\text{Species} = \text{Iris-virginica} \mid \text{SepalLength} = \text{L})$).

د) باتوجه به تعداد ویژگی‌ها و تعداد داده‌های آزمایش، آیا امکان تشکیل جدول توزیع احتمال توأم وجود دارد یا خیر؟ پاسخ خود را شرح دهید.

قوانین:

۱. تمرین ها به صورت فردی انجام شوند و حل گروهی تمرین ها مجاز نیست.
۲. نمره شما بر اساس گزارش راه طی شده برای حل مسئله و پاسخ صحیح خواهد بود لذا از هرگونه اطناب در گزارش پرهیز و به موارد خواسته شده به صورت کامل پاسخ دهید.
۳. برای تحویل تمرین یک فایل zip شامل گزارش حل سوالات، با نام [HW8_ID_NAME] در سامانه LMS بارگذاری کنید.