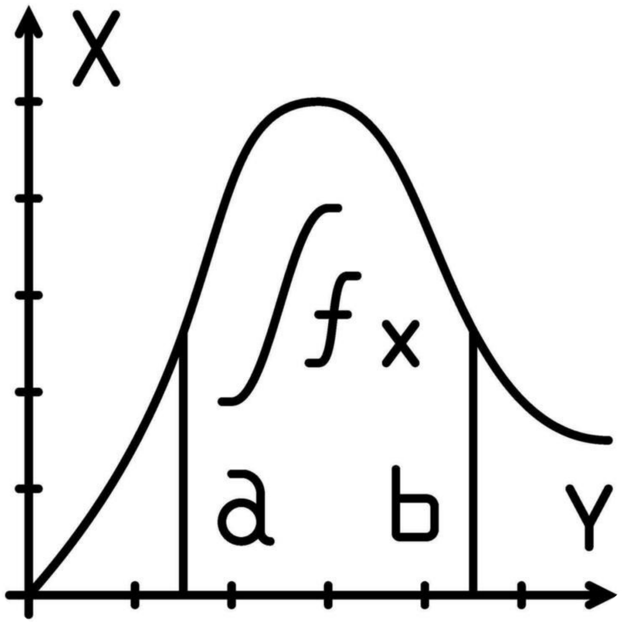
مبانی آمار و احتمال

هادی عاشری



مقدمه

- آمار و احتمال یکی از پایه‌های اساسی در یادگیری ماشین برای تحلیل داده‌ها، مدل‌سازی عدم قطعیت و پیش‌بینی در داده‌ها است.
- استفاده از احتمالات برای مدل‌سازی پیش‌بینی کلاس‌ها یا دسته‌بندی داده‌ها، مانند رگرسیون لجستیک.
- یادگیری آماری شامل یادگیری بیزی یا تحلیل توزیع داده‌ها.
- مدل‌سازی احتمالی برای شناسایی و مدیریت داده‌های نویزی.
- تخمین پارامترهای مدل‌های یادگیری ماشین.
- بررسی اعتبار فرضیات و نتایج مدل با آزمون‌های آماری.

نظریه‌ی احتمال

- نظریه‌ی احتمال شاخه‌ای از ریاضیات است که به مطالعه رویدادهای تصادفی و عدم قطعیت می‌پردازد.
- الگوریتم‌های مبتنی بر احتمال به صراحت عدم قطعیت موجود در داده‌ها یا نتایج پیش‌بینی را مدل می‌کنند.
- الگوریتم‌های مبتنی بر احتمال معمولاً نیاز به فرضیات مشخص درباره توزیع داده‌ها دارند.
- خروجی الگوریتم‌های مبتنی بر احتمال به صورت مقادیر احتمال که نشان‌دهنده سطح اطمینان در پیش‌بینی هاست، ارائه می‌شود.
- الگوریتم‌های مبتنی بر احتمال می‌توانند در برابر داده‌های نویزی یا اطلاعات ناقص مقاوم‌تر و انعطاف‌پذیرتر عمل کنند.

Iran Annual AI Award

فیلتر کالمن در حسگرهای رباتیک

- در رباتیـک، حسـگرها ماننـد GPS یـا شتاب‌سـنج‌ها داده‌هـای نویزی ای ارائه می‌کنند.
- برای تخمین موقعیت، فیلتر کالمن با ترکیـب مـدل دینـامیکی روبات و داده‌های حسگر، عدم قطعیـت هـر انـدازه‌گیری را بـه صورت یک ماتریس کوواریانس مدل می‌کنـد و بـرآورد بهبـود یافته‌ای از موقعیت روبات به دست می‌آید.
- نادیده گرفتن عدم قطعیت باعث تجمع خطاها، عدم انطباق با تغییـرات محیطـی، فقـدان قابلیـت اعتمـاد بـه پیش‌بینی‌هـا می‌شود.

شبیه‌سازی مونت کارلو در مسائل مالی

- در نظر بگیرید می‌خواهید ارزش یک سرمایه‌گذاری مثلاً یک سهم خاص در یک سال آینده را بسنجید.
- بجای یک پیش‌بینی نقطه‌ای، هزاران سناریوی ممکن با روش مونت کارلو ارزیابی می‌شود.
- با استفاده از شبیه‌سازی مونت کارلو می‌توانید بفهمید که در یک سال آینده، قیمت سهم احتمالاً در چه محدوده‌ای خواهد بود و چقدر از این پیش‌بینی مطمئن هستید.
- قدرت تحلیل ریسک و امکان مدیریت پورتفو به وجود می‌آید.

فضای احتمال

- در نظریه‌ی احتمال، فضای احتمال یک سه‌گانه $(\Omega, \mathcal{F}, P)$ است:
- فضای نمونه (Ω) مجموعه‌ای از تمام نتایج ممکن یک آزمایش تصادفی.
- فضای پیشامد $(\mathcal{F})$ مجموعه‌ای از پیشامدهاست که یک پیشامد/رویداد زیرمجموعه‌ای فضای نمونه که به عنوان رویداد در نظر گرفته می‌شوند.
- تابع احتمال (P) تابعی است که به هر پیشامد، یک مقدار احتمال اختصاص می‌دهد و طبق آکسیوم‌های کولموگروف عمل می‌کند.

Iran Annual AI Award

فضای احتمال پرتاب تاس

- برای مثال پرتاب تاس سالم، اجزای اصلی نظریه احتمال به صورت زیر بیان می‌شود:
- فضای نمونه شامل اعداد ۱ تا ۶ است.

$$\Omega = \{۱، ۲، ۳، ۴، ۵، ۶\}$$

- هر زیرمجموعه از فضای نمونه یک پیشامد است مانند:
  - مشاهده عدد زوج
  - مشاهده عدد بزرگتر از ۴.
- به هر وجه از ۶ وجه تاس، احتمال $\frac{۱}{۶}$ نسبت داده می‌شود. بنابراین برای پیشامد زوج بودن داریم:

$$A = \left\{۲، ۴، ۶\right\}$$

$$P(A) = P(۲) + P(۴) + P(۶) = \frac{۱}{۲} + \frac{۱}{۲} + \frac{۱}{۲} = \frac{۳}{۶}$$

Iran Annual AI Award

فضای احتمال پرتاب سکه

- برای مثال پرتاب تاس سالم، اجزای اصلی نظریه احتمال به صورت زیر بیان می‌شود:
  - فضای نمونه شامل اعداد رو یا پشت است.

$$\Omega = \{\text{رو، پشت}\}$$

  - هر زیرمجموعه از فضای نمونه یک پیشامد است .

$$\mathcal{F} = \{\{\Phi\}, \{\text{پشت}\}, \{\text{رو}\}, \{\text{رو، پشت}\}\}$$

  - تابع احتمال:

$$P(\{\Phi\}) = ۰, \mathrm{P}(\Omega) = \mathrm{P}\left(\{\text{رو، پشت}\}\right) = ۱,$$

$$P\left(\{\text{پشت}\}\right) = ۰.۵, P\left(\{\text{رو}\}\right) = ۰.۵$$

Iran Annual AI Award

اصول احتمال کولموگروف

- اصل اول:
  - احتمال هر پیشامد/رویداد یک مقدار نامنفی است:

$$P(E) \in \mathbb{R}, P(E) \geq 0 \; \forall E \in \mathcal{F}$$

  - تضمین می‌کند که مقادیر احتمال معنادار و از صفر کمتر نیستند.

Iran Annual AI Award

اصول احتمال کولموگروف

اصل دوم:

- احتمال رخداد کل فضای نمونه برابر با ۱ می باشد.

$$P(\Omega) = 1$$

- اطمینـان حاصـل می‌کنـد کـه تمـام احتمـالات رخـدادهای ممکـن جمعـاً نشان‌دهنده‌ی قطعیت (۱) هستند.

اصول احتمال کولموگروف

- اصل سوم:
  - اگر E_1 و E_2 و ... پیشامدهایی ناسازگار شمارش‌پذیر از فضای نمونه باشند:

$$\forall i,j\ E_i \cap Ej = \emptyset \ \Rightarrow\ P\left(\bigcup_{n=1}^{\infty}(E_1 \cup E_2 \cup \cdots)\right) = \sum_{i=0}^{\infty} P(E_i)$$

  - امکان محاسبه‌ی احتمال وقوع ترکیب‌های غیر متداخل از رویدادها را فراهم می‌کند.

Iran Annual AI Award

قوانین ابتدایی احتمال

- جمع احتمال‌ها: اگر دو رویداد A و B ناسازگار باشند (یعنی همزمان اتفاق نیفتند)، احتمال وقوع حداقل یکی از آن‌ها برابر است با:

$$P(A \cup B) = P(A) + P(B)$$

- ضرب احتمال‌ها: اگر دو رویداد A و B مستقل باشند، احتمال وقوع همزمان آن‌ها برابر است با:

$$P(A \cap B) = P(A)P(B)$$

- مکمل یک رویداد: احتمال عدم وقوع یک رویداد A برابر است با:

$$P(A') = 1 - P(A)$$

Iran Annual AI Award

مثال پرتاب تاس

- جمع احتمال‌ها: احتمال پیشامد (C) اینکه عدد تاس بزرگتر مساوی ۵ یا کوچکتر مساوی ۲ باشد:

$$A = \{۵، ۶\}, B = \{۱، ۲\} \Rightarrow$$

$$P(C) = P(A \cup B) = P(A) + P(B) = \frac{۲}{۶} + \frac{۲}{۶} = \frac{۲}{۳}$$

- مکمل رویداد: احتمال پیشامد (D) اینکه عدد تاس کوچکتر از ۵ یا بزرگتر از ۲ باشد:

$$D = \{۳،۴\} \Rightarrow$$

$$P(D) = P(\bar{C}) = 1 - P(C) = ۱ - \frac{۲}{۳} = \frac{۱}{۳}$$

مثال پرتاب تاس و سکه

- یک تاس پرتاب می‌کنی و عدد زوج بزرگتر از ۲ بیاید:

$$A = \{۴، ۶\} \Rightarrow P(A) = \frac{۲}{۶}$$

- همزمان یک سکه پرتاب می‌کنی و شیر بیاید:

$$B = \{\text{شیر}\} \Rightarrow P(B) = \frac{۱}{۲}$$

- چون تاس و سکه هیچ تأثیری بر هم ندارند، رویدادها مستقل‌اند:

$$P(A \cap B) = P(A)P(B) = \frac{۲}{۶} \times \frac{۱}{۲} = \frac{۱}{۶}$$

تعریف متغیر تصادفی

- یک متغیر تصادفی تابعی است که به هر نتیجه از یـک آزمـایش تصـادفی، یک عدد حقیقی اختصاص می‌دهد.
- اگر فضای نمونه را با Ω نشان دهیم، آنگاه متغیر تصادفی X به صـورت زیـر تعریف می‌شود:

$$X: \Omega \to \mathbb{R}$$

- در آزمایش پرتاب سکه، متغیر تصادفی X می‌تواند چنین تعیین شود:

$$X(\text{رو}) = 1, X(\text{پشت}) = 0$$

Iran Annual AI Award

متغیر تصادفی گسسته

- متغیر تصادفی گسسته تنها می‌تواند تعدادی محدود یا شمارا مقدار به خود بگیرد (مانند اعداد صحیح یا مجموعه‌ای محدود از اعداد).
- تابع جرم احتمال: برای هر مقدار ممکن x، احتمال وقوع آن به صورت $P(X = x)$ تعریف می‌شود.
- جمع‌بندی احتمال: مجموع احتمالات تمامی نتایج ممکن برابر با ۱ است:

$$\sum_i P(X = x_i) = 1$$

متغیر تصادفی گسسته در یادگیری ماشین

- در مدل‌های طبقه‌بندی متغیر تصادفی گسسته کلاس خروجی است.
  - اسپم/نرمال، مرد/زن، خردسال/نوجوان/جوان/میانسال/پیر
- در مدل‌های یادگیری ماشین برای تبلیغات دیجیتال، تعداد کلیک‌هایی که کاربران روی یک تبلیغ انجام می‌دهند یک متغیر تصادفی گسسته است.
- در تشخیص بیماری یا پیش‌بینی موفقیت یک استراتژی سرمایه‌گذاری، خروجی مدل‌ها می‌تواند یک متغیر تصادفی گسسته باشد (مانند "موفق" یا "ناموفق").

Iran Annual AI Award

مثال پرتاب دوسکه سالم

- در پرتاب دو سکه، متغیر تصادفی X تعداد پشت‌ها را نشان می‌دهد:

$$\Omega = \{\text{رو} - \text{رو} \text{ ، } \text{رو} - \text{پشت} \text{ ، } \text{پشت} - \text{رو} \text{ ، } \text{پشت} - \text{پشت}\}$$

$$X \in \{0, 1, 2\}$$

- تابع احتمال:

$$P(X = 0) = \frac{1}{4},$$

$$P(X = 2) = \frac{1}{4},$$

$$P(X = 1) = \frac{1}{2}$$

متغیر تصادفی پیوسته

- متغیر تصادفی پیوسته می‌تواند هر مقداری را در یک بازه‌ی عددی بگیرد؛ به بیان دیگر، مجموعه مقادیر آن پیوسته و معمولاً بـین دو مقـدار حـدی قرار دارد.
- برای هر نقطه مشخص به دلیل بی‌نهایت بودن تعداد نمونه‌ها، مقـدار جـرم احتمال قابل تعریف نیست (صفر است).
- تابع چگالی احتمال: احتمال اینکه متغیر در یک محدوده خاص قـرار بگیـرد، از طریق انتگرال‌گیری از تابع چگالی تعریف می‌شود:

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

Iran Annual AI Award

متغیر تصادفی پیوسته در یادگیری ماشین

- درجه حرارت محیط در پیش‌بینی‌های آب‌وهوا
- سطح قند خون در مدل‌های تشخیصی پزشکی
- میزان فروش روزانه یک فروشگاه
- سرعت خودرو در تحلیل‌های حمل‌ونقل

مقایسه متغیرهای تصادفی گسسته و پیوسته

ویژگی	متغیر تصادفی گسسته	متغیر تصادفی پیوسته
تعریف	مقدارهایی که قابل شمارش هستند و معمولاً تعداد محدودی از حالات ممکن دارند.	مقدارهایی که در یک بازه پیوسته قرار دارند و می‌توانند هر عدد حقیقی را بپذیرند.
مقادیر ممکن	عددهای گسسته مانند 0، 1، 2، …	عددهای پیوسته مانند 3.14، 7.89، …
مثال‌ها	تعداد کلیک‌های کاربران، تعداد دسته‌های یک تصویر، تعداد خطاهای مدل	دمای هوا، سرعت خودرو، میزان فروش، فشار خون
احتمال	تابع جرم احتمال	تابع چگالی احتمال
نمایش نمودار	نمایش در نمودار نقاط منفرد روی محور مختصات	یک منحنی پیوسته روی محور مختصات

توزیع‌های رایج: نرمال

- توزیع نرمال یا توزیع گاوسی یک توزیع پیوسته است که معمولاً به شکل یک منحنی زنگوله (bell curve) نمایش داده می‌شود:
- این توزیع با دو پارامتر میانگین (μ) و انحراف معیار (σ) تعریف می‌شود.
- فرمول تابع چگالی احتمال:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- منحنی یکنواخت، همگام از سمت چپ به میانگین افزایش و از میانگین به سمت راست کاهش می‌یابد.

Iran Annual AI Award

ویژگی‌های توزیع نرمال

- منحنی نرمال نسبت به میانگین متقارن است؛ به عبارت دیگر، میانگین، میانه و مد برابرند.
- شکل کلاسیک زنگوله‌ای، که نشان‌دهنده توزیع احتمال در اطراف میانگین است.
- کنترل می‌کند که داده‌ها تا چه حد از میانگین فاصله دارند؛ انحراف معیار کوچک، تراکم بیشتر داده‌ها در اطراف میانگین را نشان می‌دهد.
- احتمال مشاهده مقادیر بسیار دور از میانگین به سرعت کاهش می‌یابد.

Iran Annual AI Award

اهمیت توزیع نرمال در تحلیل داده

- بسیاری از آزمون‌های آماری و روش‌های استنباطی (مانند آزمون‌های z gt) بر اساس فرض نرمال بودن داده‌ها طراحی شده‌اند.
- قضیه حد مرکزی: این قضیه تضمین می‌کند که توزیع میانگین نمونه‌ها حتی در برخی شرایط متفاوت به توزیع نرمال نزدیک می‌شود.
- خطاهای اندازه‌گیری و نویز در بسیاری از پدیده‌های طبیعی و تجربی با توزیع نرمال مدل می‌شوند.
- پیش‌پردازش یا نرمال‌سازی داده‌ها با هدف بهبود عملکرد الگوریتم‌ها از تکنیک‌های مبتنی بر ویژگی‌های توزیع نرمال بهره می‌برند.
- توزیع نرمال به عنوان یکی از مباحث پایه در آمار، نقش اساسی در مدل‌های تحلیلی و یادگیری ماشین دارد.

Iran Annual AI Award

توزیع‌های رایج: توزیع برنولی

- توزیع برنولی یک توزیـع احتمـال گسسـته اسـت کـه تنهـا دو حالـت ممکـن دارد: موفقیت (1) و عدم موفقیت (0).
- احتمال رخداد موفقیت p (رویداد 1)
- احتمال رخداد عدم موفقیت 1-p (رویداد ٠)
- فرمول تابع چگالی احتمال:

$$P(X=x)=\begin{cases} p & if\ x=1 \\ 1-p & if\ x=0 \end{cases}$$

$$=p^x(1-p)^{1-x}$$

- مدل‌سازی آزمایش‌هایی که نتیجه آنها تنها دو حالت دارند (مثلاً روشـن/خاموش، قبول/رد، موفق/ناموفق)

Iran Annual AI Award

اهمیت توزیع برنولی در تحلیل داده

- بسیاری از مسائل در یادگیری ماشین و تحلیل داده (مانند طبقه‌بندی باینری، تصمیم‌گیری‌های دوجمله‌ای) مبتنی بر رویدادهای باینری هستند.
- توزیع برنولی، بلاواسطه به توزیع دو جمله‌ای (Binomial) تعمیم داده می‌شود که در تحلیل تعداد موفقیت‌ها در تعدادی آزمایش مستقل کاربرد دارد.
- مدل‌هایی مانند رگرسیون لجستیک بر روی داده‌های باینری اعمال می‌شوند و در بسیاری از مسائل طبقه‌بندی مشارکتی بر اساس احتمال موفقیت و شکست ساخته می‌شوند.
- با داشتن تنها یک پارامتر (p)، تحلیل‌گران می‌توانند به راحتی احتمال موفقیت را تخمین زده و از آن در مدل‌های پیش‌بینی بهره ببرند.

توزیع‌های رایج: توزیع دوجمله‌ای

- بررسی موفقیت/شکست در n آزمایش مستقل با احتمال موفقیت p.
- احتمال رخداد موفقیت p (رویداد 1)
- احتمال رخداد عدم موفقیت 1-p (رویداد ٥)
- فرمول تابع چگالی احتمال:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

- در مسائل طبقه‌بندی دودویی مانند تشخیص ایمیل هرزنامه، تعداد نمونه‌هایی که به درستی طبقه‌بندی شده‌اند (موفقیت) در یک مجموعه داده به‌عنوان یک رویداد دوجمله‌ای در نظر گرفته می‌شود.

توزیع‌های رایج: توزیع پواسون

- تعریف: توزیع پواسون یک توزیع احتمال گسسته است که احتمال وقوع تعداد معینی از رویدادها در یک بازه زمانی یا مکانی که رویدادها مستقل از هم اتفاق می‌افتند و با نرخ میانگین ثابت رخ می‌دهند را مدل‌سازی می‌کند.
- فرمول تابع جرم احتمال:

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \qquad k = 0, 1, 2, \cdots$$

- پارامتر λ: نرخ متوسط وقوع رویدادها در بازه مورد نظر
- تنها مقادیر صحیح غیرمنفی (0، 1، 2، ...) را می‌پذیرد.
- فرض می‌شود که رویدادها به صورت مستقل از یکدیگر رخ می‌دهند.

اهمیت توزیع پواسون در تحلیل داده

- برای تحلیل تعداد دفعات وقوع رویدادهایی مانند تماس‌های ورودی در یک مرکز تماس، خطاهای نرم‌افزاری، یا تعداد تصادفات یک شهر در یک بازه زمانی، بسیار مناسب است.
- در مواردی که رویدادها نادر رخ می‌دهند ولی تحلیل آن‌ها اهمیت دارد (مثلاً خرابی‌های بحرانی در سیستم‌های صنعتی یا سازمانی)، توزیع پواسون کاربرد ویژه‌ای دارد.
- کمک می‌کند تا نرخ وقوع رویدادها λ به عنوان معیاری برای پیش‌بینی‌های آینده استخراج شود.

Iran Annual AI Award

توزیع‌های رایج: توزیع نمایی

- تعریف: توزیع نمایی یک توزیع احتمال پیوسته است که زمان بین رویدادهایی که به‌طور مستقل و با نرخ ثابت رخ می‌دهند را مدل می‌کند.
- تابع چگالی احتمال: برای متغیر تصادفی X با نرخ وقوع رویداد λ داریم:

$$f(x) = \lambda e^{-\lambda x}, \qquad \forall x \geq 0$$

- احتمال رخداد یک رویداد در آینده، صرف نظر از طول وقایع گذشته ثابت است (بی حافظگی)، به این معنا که:

$$P(X > s + t \mid X > s) = P(X > t)$$

- احتمال وقوع زمان‌های بسیار بلند به صورت نمایی کاهش می‌یابد.

Iran Annual AI Award

اهمیت توزیع نمایی در تحلیل داده

- کاربرد گسترده در تحلیل زمان‌های انتظار، مانند فاصله بین وقوع خطاهای سیستم، تماس‌های ورودی و زمان‌های سرویس دهی.
- در یک فرایند پواسون، زمان بین رویدادها (فاصله‌های زمانی) دارای توزیع نمایی است.
- ارزیابی عمر مفید تجهیزات و دستگاه‌ها (تحلیل قابلیت اعتماد)
- تحلیل زمان‌های انتظار مشتریان در سیستم‌های خدماتی.

احتمال شرطی

- احتمال شرطی ابزاری اساسی برای به‌روزرسانی باورها با توجه به شواهد جدید می‌باشد.
- در بسیاری از الگوریتم‌های یادگیری ماشین (مثلاً دسته‌بندی با استفاده از (Naive Bayes) و تحلیل‌های آماری مورد استفاده قرار می‌گیرد.

Iran Annual AI Award

احتمال شرطی برای متغیرهای گسسته

- برای دو رویداد A و B در فضای نمونه‌های گسسته، احتمال شرطی بیان می‌شود به صورت:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, P(A) \geq 0$$

- فرض کنید یک تاس منصفانه داریم. احتمال آمدن عدد ۶ به شرط آنکه عدد زوج آمده باشد:
  - رویداد آمدن عدد فرد (A)
  - رویداد آمدن عدد ۶ (B)

$$P(A) = \frac{3}{6}, P(A \cap B) = \frac{1}{6}, P(B|A) = \frac{1}{3}$$

- کاربردهای متعدد در مسائل طبقه‌بندی و مدلسازی رویدادهای شمارشی در تحلیل داده وجود دارد.

Iran Annual AI Award

احتمال شرطی برای متغیرهای پیوسته

- برای دو رویداد A و B در فضای نمونه‌های گسسته، احتمال شرطی بیان می‌شود به صورت:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

- به جای احتمال نقطه‌ای، از انتگرال‌گیری روی بازه‌ها استفاده می‌شود:

$$P(a \leq X \leq b|Y = y) = \int_a^b f_{X|Y}(x|y)dx$$

- کاربرد ویژه در استنباط بیزی، رگرسیون‌های احتمالاتی و تحلیل وابستگی‌های پیچیده میان متغیرها در داده‌ها دارد.

Iran Annual AI Award

اهمیت احتمال شرطی در تحلیل داده

- احتمال شرطی اجازه می‌دهد تا با دریافت اطلاعات جدید، احتمال وقوع رویدادها به‌روز شود (اصول استنباط بیزی).
- در داده‌های پیچیده، بسیاری از متغیرها به طور مستقیم یا غیرمستقیم بر یکدیگر تأثیر می‌گذارند؛ استفاده از احتمال شرطی می‌تواند این وابستگی‌ها را تشخیص داده و مدل کند.
- الگوریتم‌هایی مانند Naive Bayes، شبکه‌های بیزی و مدل‌های پنهان مارکوف برای طبقه‌بندی و تشخیص الگو، از احتمال شرطی به عنوان عنصر اصلی استفاده می‌کنند.
- شناسایی ارتباط بین متغیرها و تعیین تأثیر یک متغیر بر متغیر دیگر از طریق محاسبه احتمال‌های شرطی امکان‌پذیر می‌شود.

Iran Annual AI Award

رویدادهای مستقل

- تعریف: دو رویداد A و B مستقل هستند اگر وقوع یکی تـأثیری بـر احتمـال وقـوع دیگری نداشته باشد. به عبارت دیگر:

$$P(A \cap B) = P(A)P(B)$$

- در آزمایش تصادفی پرتاب دو سکه، نتیجه سکه اول هیچ تـاثیری بـر نتیجـه سـکه دوم ندارد.

- عدم تغییر احتمال رویداد A با اطلاع از وقوع B (و بالعکس).

- هنگام شروع آموزش یک شـبکه عصبی، وزن‌هـا اغلـب بـه صـورت تصـادفی از یـک توزیـع مشـخص (ماننـد گاوسـی یـا یکنواخـت) مقـداردهی اولیـه می‌شـوند. ایـن مقـداردهی بـه‌طور مسـتقل بـرای هـر وزن انجـام می‌شـود تـا شـبکه در مسـیرهای بهینه‌سازی بدون جهت‌گیری اولیه آغاز به کار کند.

رویدادهای وابسته

- تعریف: دو رویداد A و B وابسته هستند اگر وقوع یکی احتمال وقوع دیگری را تغییر دهد. رابطه شرطی آن به صورت زیر بیان می‌شود:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, P(A) \geq 0$$

- در انتخاب کارت از یک دسته بدون جایگذاری، نتیجه انتخاب کارت اول، احتمال انتخاب کارت دوم را تغییر می‌دهد.
- در مدل GPT احتمال تولید یک کلمه به کلماتی که قبل از آن تولید شده وابسته است.
- مثلاً در جمله "من به ... می‌روم"، احتمال کلمه بعدی به زمینه جمله و کلمات قبلی بستگی دارد.
- مکانیزم توجه در مدل‌های زبانی بزرگ برای وابستگی‌های طولانی‌مدت

Iran Annual AI Award

رویدادهای متقارن

- تعریف: رویدادهای متقارن به آن دسته از رویدادهایی گفته می‌شود که در یک ساختار یا توزیع احتمال، هر یک از نتایج یک نمونه‌فضا با احتمال یکسان رخ می‌دهند.
  - پرتاب سکه سالم
  - پرتاب تاس سالم
- بهره‌گیری از تقارن باعث کاهش پیچیدگی محاسبات و تحلیل‌های آماری می‌شود.
- در مدل‌های سریع و کاربردی یادگیری ماشین، شرایط متقارن می‌تواند به ساده‌سازی مسئله کمک کند.
- افزایش داده با تغییرات متقارن در مدل‌های بینایی ماشین

Iran Annual AI Award

رویدادهای ناسازگار

- دو رویداد ناسازگار، یا متناقص، به رویدادهایی گفته می‌شوند که به هیچ‌وجه نمی‌توانند همزمان رخ دهند:

$$P(A \cap B) = 0$$

- در پرتاب یک تاس، رویداد "مشاهده 5" و رویداد "مشاهده 6" ناسازگار هستند؛ زیرا نمی‌توان در یک پرتاب هر دو عدد را مشاهده کرد.
- انتساب برچسب در روش one-hot-encoding

Iran Annual AI Award

www.iaaa.ai

support@iaaa.ai

021-91096992

iaaa.event

iaaa_ai

با تشـــکر
از توجه شما