مفاهیم و کاربردهای هوش مصنوعی در حسابداری

عم داده و هوش مصنوعی



بهرام صادقی بی غم

عضو هیات علمی دانشگاه و مشاورصنعت

و تجارت در حوزه هوش مصنوعی و هوش تجاری

وبسایت رسمی www.sadeghibigham.ir/

تاریخ واژه ها

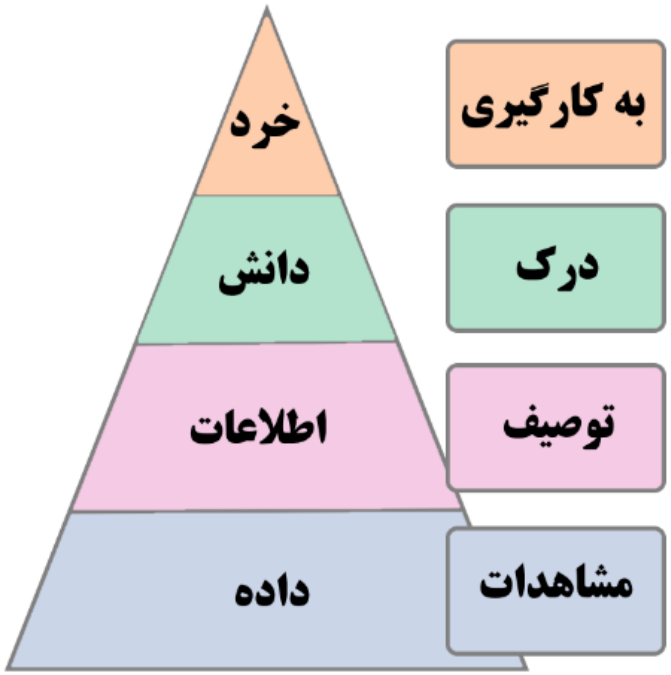
تی اس الیوت در سال ۱۹۳۴ میلادی شعری سروده است که مبنای تقسیم بندی خاصی شده است.
Thomas Stearns Eliot

کجاست آن زندگی که لابلای روزمرگی گم کرده ایم؟
کجاست خردی که لابلای دانش گم کرده ایم؟
کجاست دانش که لابلای اطلاعات گم کرده ایم؟



کلیات

- داده: مقدار اندازه گیری شده بدون امکان قضاوت
- اطلاعات: داده های ساختارمند و قابل مقایسه. چه کسی؟ چه چیزی؟ نفر چندم؟
- دانش: اطلاعات قابل درک و تفسیر به قصد نتیجه گیری یا پیش بینی. چه چیزهایی باهم؟ از چه نوعی است؟ بعدا چه می شود؟
- خرد: دانش جامع و چندبعدی و پاسخ به چرایی پدیده ها



داده

- داده ها: از طریق اندازه گیری های انجام شده از جهان (اشیا، رویدادها، ...) و مقادیر متغیرهای مختلف ایجاد می شوند و به خودی خود و به تنهایی معنای خاصی ندارند. لیستی از اعداد (مثلا ۱۷و ۴و ۹۸و ۴۵و ۱۱.۵) را میتوانیم به عنوان یک مجموعه داده ببینیم و طبیعتا نمیتوان هیچ برداشتی از آن کرد. حتی لیست نمرات یک درس (مثلا ۱۱و ۱۸و ۱۵و ۹و ۱۲) هم میتواند بعنوان داده در نظر گرفته شود به شرطی که هیچ قضاوتی راجع به نمرات و مقایسه آنها باهم و یا حتی مقایسه با نمره ی ۱۰ یا ۲۰ نکنیم.



داده، نفت جدید است

اطلاعات

اطلاعات: داده‌هایی حاوی معنی هستند که پردازش شده‌اند تا برای انسان معنادار باشند.
اطلاعات داده هایی هستند که ارتباط بین آنها و یا ارتباطشان با هدفی مشخص است (بالی و همکاران ۲۰۰۹). ممکن است نشان دهنده روندی و یا الگویی در یک دوره زمانی معین باشند. با داشتن اطلاعات می توانیم به سوالاتی از قبیل "چه کسی"، "چه چیزی"، "چه زمانی"، و .... پاسخ دهیم. لیست نمراتی که در قسمت داده ها مثال زده شد، میتواند به معنی اطلاعات باشد. چوان میتوانیم بفهمیم چند نفر درس را گذرانده اند و یا مردود شده اند، نمره ی نفر سوم چه مقدار تا نمره کامل اختلاف دارد، کمترین نمره چند است، و چندین اطلاعات دیگر.
داده‌های سازمان‌یافته، ساختارمند و دارای زمینه (کانتکست) است. در واقع، داده‌ای است که ساختاردهی و دارای زمینه (Context) شده است.



نمونه‌هایی از اطلاعات:

- گزارشی از فروش که پرفروش‌ترین محصولات را برجسته می‌کند
- مقاله‌ای که یافته‌های پژوهشی را خلاصه می‌کند
- نموداری که همبستگی بین دو متغیر را نشان می‌دهد

دانش

دانش: نتیجهٔ تحلیل و تفسیر اطلاعات به‌منظور کشف الگوها، روندها و روابط است. نتیجه اش قضاوت و بینش است اطلاعاتی است که می توان بر اساس آن عملی انجام داد، پیش بینی کرد و یا نتیجه ای گرفت. دانش نوعی آگاهی در مورد یک پدیده، فرد، مکان، رویداد، شی، مساله یا هر چیز دیگری است که حاصل داشتن اطلاعات، یادگیری، تجربه و شهود است و میتواند در نتیجه ی درک مفاهیم، روابط، تجزیه و تحلیل ایجاد شود.. در مثال مربوط به نمره ی درس، نتیجه گیری هایی مانند: "آن چهار نفر که همیشه باهمند، نمراتشان دقیقا با هم یکی است!"، "نمرات کم اغلب مربوط به کسانی است که بیشتر جلسات را غایب بوده اند"، ... از نوع دانش هستند.



داده/اطلاعات/دانش

سطح	تعریف	ارزش افزوده
داده(Data)	واقعیت‌ها و اعداد خام، بدون معنا و زمینه	پایه اولیه برای پردازش
اطلاعات(Information)	داده‌های پردازش‌شده و سازمان‌یافته که معنا دارند	برای تصمیم‌گیری و تحلیل مفیدند
دانش(Knowledge)	ترکیب اطلاعات با تجربه، درک و قضاوت	برای اقدام مؤثر و تصمیم هوشمندانه

- به بیان ساده:

داده جمع‌آوری می‌شود،
اطلاعات تحلیل می‌شود،
و **دانش** به‌کار گرفته می‌شود.

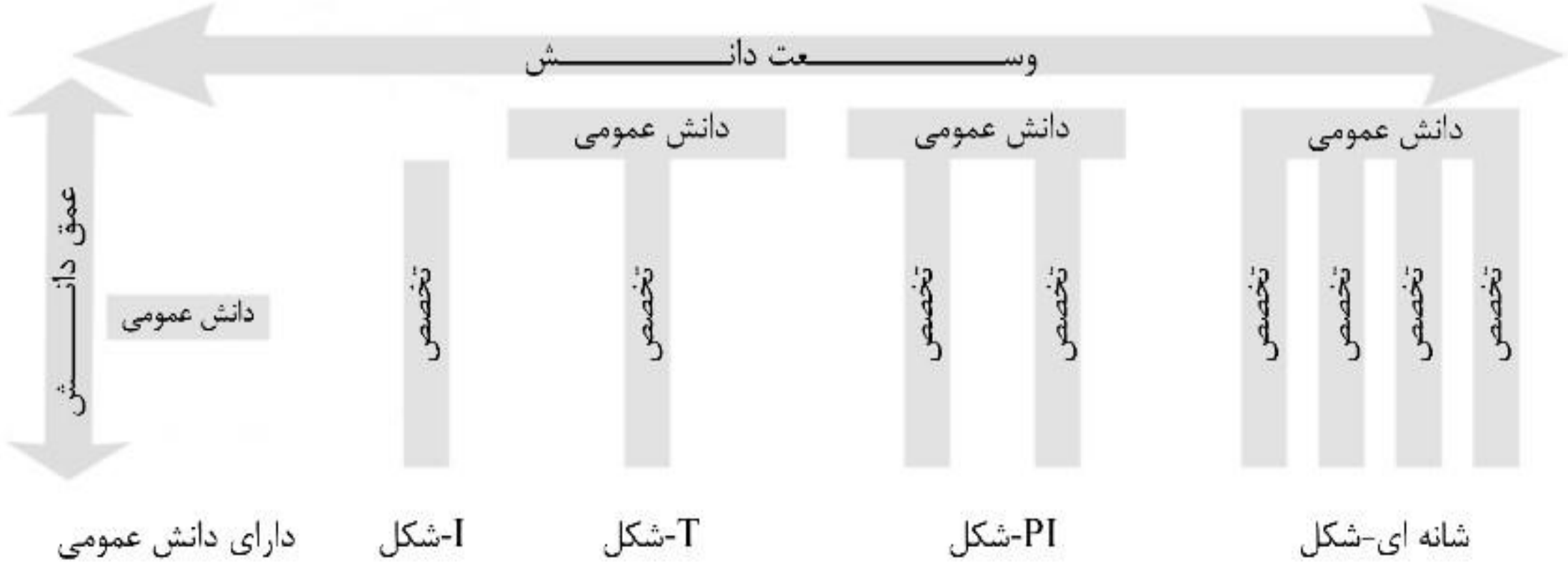
خرد

- خرد: خرد پاسخ دادن به "چرا"ها با استفاده از یک یا چند دانش عمیق است و داشتن خرد منجر به رفتار صحیح و مناسب می شود. دانش فرد خردمند جامع است. در برخی نوشته ها از عبارت بینش و یا حکمت نیز استفاده شده است و گاها با مفهوم "دانش در مورد دانشها" نیز بیان می شود. با این اوصاف میتوان به تفاوت فرد دانشمند و خردمند پی برد و متوجه شد که هر خردمندی دانشمند است ولی هر دانشمندی لزوما یک خردمند نیست. فرد خردمند از دانشی که از نمرات یک دانش آموز به دست می آورد، با کسب دانش در مورد استعداد و علاقه دانش آموز و همچنین با استفاده از دانش دیگری که از آینده ی کاری و شغلی رشته های دانشگاهی دارد نهایتا بهترین پیشنهاد را برای انتخاب رشته دانشگاهی او ارایه میدهد.



خِرَد توانایی تصمیم‌گیری آگاهانه و اقدام مؤثر بر اساس درک و فَهم دانش زیربنایی است.

مرز بین دانش و خرد



نمونه‌ای از زندگی روزمره: پایش تناسب اندام

داده:(Data) : ساعت هوشمند تعداد گام‌ها، ضربان قلب و مدت زمان خواب را جمع‌آوری می‌کند.
اطلاعات:(Information) : اپلیکیشن ساعت هوشمند داده‌ها را سازمان‌دهی کرده و آن‌ها را در قالبی قابل درک نمایش می‌دهد، مانند تعداد گام‌های روزانه، میانگین ضربان قلب و ساعات خواب در هر شب.
دانش:(Knowledge) : تحلیل و قضاوت، الگوهایی را آشکار کند؛ مثلاً افزایش تعداد گام‌ها باعث بهبود کیفیت خواب می‌شود یا بین ضربان قلب و شدت تمرین ارتباط وجود دارد.
خِرَد:(Wisdom) : درک این الگوها به شما کمک می‌کند تصمیم‌های آگاهانه‌تری دربارهٔ تنظیم برنامه ورزشی، عادات خواب و سایر جنبه‌های سبک زندگی خود بگیرید تا سلامت و تناسب‌اندام‌تان را بهبود دهید.



نمونه بازاریابی: کمپین بازاریابی دیجیتال

داده:(Data) : داده‌های خام شامل تعاملات کاربران با محتوای بازاریابی است، مانند کلیک‌ها، بازدیدها، لایک‌ها، اشتراک‌گذاری‌ها و خریدها.

اطلاعات:(Information) : داده‌ها سازمان‌دهی و ساختارمند می‌شوند و شاخص‌هایی مانند نرخ کلیک، نرخ تبدیل به سرنخ، هزینه جذب هر مشتری (CPA) را در هر کانال بازاریابی ارائه می‌دهند.

دانش:(Knowledge) : تحلیل اطلاعات کمپین و مقایسهٔ آن با گذشته، الگوهایی را آشکار می‌کند؛ مانند مؤثرترین قالب‌های تبلیغاتی، بهترین زمان‌های انتشار، و ویژگی‌های کلیدی مشتریانی که بیشترین نرخ تبدیل را دارند.

خِرَد:(Wisdom) : دانش استخراج شده، کمک می‌کند تا تصمیم‌های آگاهانه‌ای دربارهٔ هدف‌گیری بخش‌های خاصی از مخاطبان، تنظیم پیام‌های تبلیغاتی و تخصیص منابع اتخاذ کند



نمونه فروش: مدیریت ارتباط با مشتری(CRM)

داده:(Data) : داده‌های خام CRM شامل سرنخ ها و تعاملات انفرادی مشتریان با تیم فروش است؛ مانند تماس‌های تلفنی، ایمیل‌ها، جلسات، پیشنهادها و فروش های نهایی‌شده

اطلاعات:(Information) : داده‌ها به شکل ساختارمند ارائه می‌شوند و شاخص‌هایی مانند نرخ تبدیل سرنخ‌ها، میانگین اندازهٔ معامله، طول چرخه فروش، و ارزش طول عمر مشتری (CLV) را به تفکیک صنایع و بخش‌های مختلف مشتریان نشان می‌دهند.

دانش:(Knowledge) : تحلیل و تفسیر اطلاعات الگوهایی را آشکار می‌کند؛ مانند مؤثرترین تکنیک‌های فروش، بهترین زمان برای پیگیری مشتریان، و ویژگی‌های کلیدی مشتریان باارزش بالا.

خرَد:(Wisdom) : تصمیم‌های آگاهانه برای اولویت‌بندی مشتریان باارزش‌تر، شخصی‌سازی روش فروش و تخصیص بهینهٔ منابع تا درآمد و رضایت مشتری حداکثر شود.



نمونه محصول: توسعهٔ اپلیکیشن موبایل

داده:(Data) : داده‌های خام شامل تعاملات کاربران با اپلیکیشن است، مانند کلیک روی دکمه‌ها، مشاهده صفحات، مدت زمان استفاده از اپ، و همچنین بازخوردها

اطلاعات:(Information) : شاخص‌هایی مانند میانگین مدت کاربری، فراوانی استفاده از قابلیت‌ها، نرخ نگهداشت کاربران **(Retention Rate)** و موضوعات متداول بازخورد در بخش‌های مختلف کاربران

دانشKnowledge : مانند ویژگی‌هایی که بیشتر درخواست می‌شوند، علل نارضایتی کاربران و عوامل کلیدی که موجب تعامل و وفاداری آن‌ها می‌شود.

خِرَد:(Wisdom) : درک نیازها، ترجیحات و نقاط درد کاربران به مدیر محصول کمک می‌کند تا تصمیم‌های آگاهانه‌ای برای اولویت‌بندی توسعهٔ قابلیت‌ها، بهبود تجربه کاربری بگیرد



نمونه مالی: برنامه‌ریزی مالی(Financial Planning)

داده:(Data) : داده‌های خام شامل تراکنش‌ها و سوابق مالی انفرادی است، مانند فاکتورهای فروش، رسیدهای هزینه، سوابق حقوق و دستمزد، و ارزیابی دارایی‌ها.

اطلاعات:(Information) : داده‌های مالی در قالب صورت‌های مالی و گزارش‌ها سازمان‌دهی می‌شوند و شاخص‌هایی مانند درآمد، سود خالص، جریان نقدی، حاشیهٔ سود ناخالص و بازده سرمایه‌گذاری (ROI) را ارائه می‌دهند.

دانش:(Knowledge) : تحلیل و تفسیر اطلاعات مالی الگوهایی را آشکار می‌کند؛ مانند شناسایی سودآورترین خطوط محصول، عوامل اصلی هزینه، و مؤلفه‌های کلیدی موثر بر درآمد و هزینه‌ها

خِرَد:(Wisdom) : درک سلامت مالی شرکت و شرایط بازار، تصمیم‌های آگاهانه‌ای برای اولویت‌بندی سرمایه‌گذاری‌ها، مدیریت هزینه‌ها و تخصیص بهینهٔ منابع وسودآوری و رشد حداکثری



انواع داده

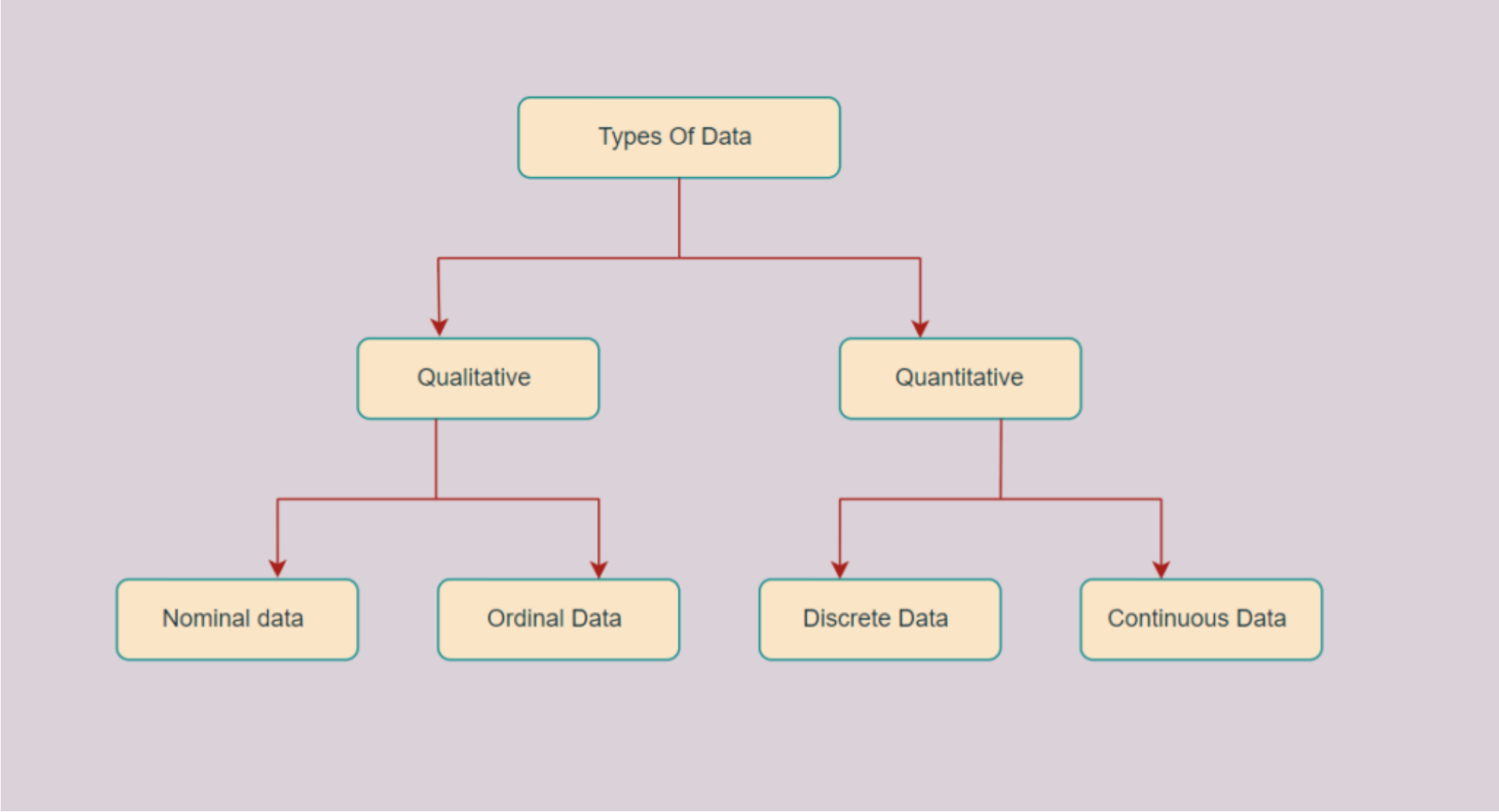
۱- داده‌های کمی یا عددی(Quantitative / Numerical Data)

- تعریف :داده‌هایی که به صورت عددی هستند و می‌توان عملیات ریاضی روی آن‌ها انجام داد.
- شامل:
  - گسسته :(Discrete) اعداد شمارشی که محدود هستند، مثل تعداد دانش‌آموزان در کلاس.
  - پیوسته :(Continuous) اعداد قابل اندازه‌گیری با دقت دلخواه، مثل قد، وزن، دما.
- مثال :۵ نفر، ۲۳.۷ درجه، ۱۲۰ کیلوگرم

۲- داده‌های کیفی(Qualitative / Categorical Data)

- تعریف :داده‌هایی که ویژگی‌ها یا دسته‌ها را نشان می‌دهند و عددی نیستند.
- شامل:
  - اسمی :(Nominal) فقط نام یا دسته هستند، بدون ترتیب. مثال: رنگ‌ها (قرمز، آبی، سبز)
  - ترتیبی :(Ordinal) دسته‌ها دارای ترتیب هستند، ولی فاصله بین آن‌ها مشخص نیست.
    مثال: سطح تحصیلات (دیپلم لیسانس فوق‌لیسانس)

ساختار کلی


Types Of Data
Qualitative
Quantitative
Nominal data
Ordinal Data
Discrete Data
Continuous Data

داده های کمی (عددی)

داده‌های عددی داده‌هایی هستند که مقدار آن‌ها به‌صورت عدد بیان می‌شود و می‌توان آن‌ها را اندازه‌گیری، مقایسه و روی آن‌ها عملیات ریاضی (مانند جمع، تفریق، میانگین، و انحراف معیار) انجام داد.



داده های گسسته/پیوسته

داده‌های گسسته

داده‌های گسسته شامل مقادیری هستند که اعداد صحیح یا کامل هستند. برای مثال، تعداد کل دانش‌آموزان یک کلاس، یک داده گسسته است. این داده‌ها نمی‌توانند به مقادیر کسری یا اعشاری تقسیم شوند. داده‌های گسسته قابل شمارش بوده و دارای مقادیر محدود هستند و تقسیم‌بندی دلخواه آن‌ها به اجزا کوچکتر از واحد امکان‌پذیر نیست. این داده‌ها معمولاً با نمودار میله‌ای، خط عددی یا جدول فراوانی نمایش داده می‌شوند.

نمونه‌های داده‌های گسسته: تعداد افراد، روزهای هفته،

داده‌های پیوسته

داده‌های پیوسته به شکل اعداد اعشاری یا کسری هستند. داده‌های پیوسته اطلاعاتی را نشان می‌دهند که می‌توان آن‌ها را به سطوح کوچکتر (تا حد دلخواه) تقسیم کرد. متغیر پیوسته می‌تواند هر مقداری در یک بازه مشخص داشته باشد.

نمونه‌های داده‌های پیوسته: قد یک فرد، سرعت، مدت زمان صرف‌شده، قیمت

داده های کیفی

داده‌های کیفی یا دسته‌ای نوعی از داده‌ها هستند که نمی‌توان آن‌ها را به شکل عددی اندازه‌گیری یا شمارش کرد. این داده‌ها بر اساس دسته‌بندی مرتب می‌شوند، نه بر اساس عدد. به همین دلیل به آن‌ها داده‌های دسته‌ای (Categorical Data) نیز گفته می‌شود. این داده‌ها می‌توانند شامل صوت، تصویر، نماد یا متن باشند. برای مثال، جنسیت یک فرد (مرد، زن) زبان، مقصد، نظر (موافق، مخالف، یا بی‌طرف)، رنگ‌ها



داده های اسمی/ترتیبی

داده‌های اسمی(Nominal Data)

داده‌های اسمی برای برچسب‌گذاری متغیرها بدون هیچ ترتیب یا ارزش کمی استفاده می‌شوند. برای مثال، رنگ یک داده اسمی است، زیرا نمی‌توان یک رنگ را با رنگ دیگر مقایسه کرد. با داده‌های اسمی نمی‌توان عملیات عددی انجام داد یا داده‌ها را بر اساس ترتیب مرتب کرد. این داده‌ها هیچ ترتیب معناداری ندارند و مقادیر آن‌ها در دسته‌های متمایز قرار می‌گیرد.

نمونه‌های داده‌های اسمی: رنگ مو (بلوند، قرمز، قهوه‌ای، مشکی و غیره)، وضعیت تأهل، ملیت، جنسیت، رنگ چشم

داده‌های ترتیبی(Ordinal Data)

داده‌های ترتیبی دارای ترتیب طبیعی هستند که تنها توالی را نشان می‌دهند و مقادیر آن‌ها موقعیت نسبی در یک مقیاس را نشان می‌دهد. در این داده‌ها ترتیب معنا دارد، اما نمی‌توان عملیات ریاضی روی آن‌ها انجام داد. نمونه‌های داده‌های ترتیبی: رضایت مشتری نمرات امتحان، رتبه‌ب در یک مسابقه، دهک اقتصادی، تحصیلات

داده ها از نظر ساختار

داده‌های ساخت‌یافته(Structured Data)

داده‌هایی که در قالب جدول با ردیف‌ها و ستون‌های مشخص سازمان‌دهی شده‌اند.

- **مثال: پایگاه‌های داده رابطه‌ای، فایل‌های CSV ، جداول اکسل .جدول، پایگاه داده، مقادیر عددی و متنی با قالب مشخص**

داده‌های نیمه‌ساخت‌یافته(Semi-structured Data)

داده‌هایی که برخی ساختار دارند، اما به اندازه داده‌های ساخت‌یافته منظم نیستند.

- **مثال: فایل‌های XML ، ایمیل‌ها.**

داده‌های غیرساخت‌یافته(Unstructured Data)

- **داده‌های فاقد ساختار مشخص که پردازش آن‌ها پیچیده‌تر است**
- **مثال: متن آزاد، تصاویر، ویدیوها، صداها.**

داده های ساخت یافته

داده‌های ساخت‌یافته شبیه یک صفحه گسترده مرتب که هر قطعه اطلاعات در ردیف‌ها و ستون‌های مشخص قرار گرفته است (شامل رکوردها و ویژگی ها)، مانند لیست مخاطبین تلفن یا بودجه‌ای که در جدول تنظیم می‌کنید. جستجو و مدیریت آن‌ها آسان‌تر است.

	A	B	C	D
1	Product	Qtr 1	Qtr 2	Grand Tota
2	Chocolade	$744.60	$162.56	$907.16
3	Gummibarchen	$5,079.60	$1,249.20	$6,328.80
4	Scottish Longbreads	$1,267.50	$1,062.50	$2,330.00
5	Sir Rodney's Scones	$1,418.00	$756.00	$2,174.00
6	Tarte au sucre	$4,728.00	$4,547.92	$9,275.92
7	Chocolate Biscuits	$943.89	$349.60	$1,293.49
8	Total	$14,181.59	$8,127.78	$22,309.37

داده های غیر ساخت یافته

داده‌های غیرساخت‌یافته برعکس داده‌های ساخت‌یافته هستند. بیشتر شبیه یک کشوی نامرتب پر از اقلام مختلف که برخی از آن‌ها دشوار به دسته‌بندی هستند.
این داده‌ها شامل ایمیل‌ها، ویدیوها، فایل‌های صوتی و نظرات مشتریان می‌شوند. تحلیل این داده‌ها معمولاً دشوارتر است، زیرا اطلاعات به شکل قابل پیش‌بینی سازمان‌دهی نشده‌اند،



داده های نیمه ساخت یافته

داده‌های نیمه‌ساخت‌یافته تا حدودی سازماندهی شده‌اند، اما همه چیز از یک قالب سخت و ثابت پیروی نمی‌کند مانند یک ایمیل با موضوع و متن اصلی، ساختار پایه (فرستنده، موضوع، متن) در آنها وجود دارد. نمونه‌ه دیگر صفحات وب است.



دسته بندی از منظر اندازه گیری

نوع داده	تعریف	ویژگی‌های کلیدی	مثال
اسمی (Nominal)	داده‌های طبقه‌ای بدون ترتیب ذاتی	فقط برابری یا نابرابری معنا دارد	جنسیت (زن/مرد)، رنگ، نوع خودرو
ترتیبی (Ordinal)	داده‌های طبقه‌ای با ترتیب، ولی بدون فواصل دقیق	ترتیب معنا دارد، اما فاصله‌ها یکسان نیست	سطح رضایت (کم، متوسط، زیاد)
فاصله‌ای / بازه‌ای (Interval)	داده‌هایی با فواصل معنی‌دار ولی بدون صفر مطلق	تفاضل‌ها معنی دارد، نسبت‌ها ندارد	دما بر حسب سانتی‌گراد
نسبتی (Ratio)	داده‌هایی با صفر مطلق و فواصل یکسان	هم تفاضل و هم نسبت معنی دارد	وزن، قد، درآمد
دودویی (Binary)	داده‌هایی با دو حالت ممکن (معمولاً خاص نوع اسمی)	فقط دو مقدار (۰ و ۱)	درست/نادرست، بله/خیر، روشن/خاموش

دسته بندی از منظر محتوا

نوع داده	تعریف	مثال‌ها	چالش‌ها
داده مکانی / جغرافیایی (Spatial / Geographic Data)	داده‌هایی که دارای موقعیت در فضا هستند (مختصات، ناحیه، سطح)	نقشه‌ها، GPS، مرزها، مسیرها	وابستگی فضایی، دقت مکانی، مقیاس
داده زمانی(Temporal Data)	داده‌هایی که در طول زمان جمع‌آوری شده‌اند یا وابسته به زمان هستند	قیمت روزانه، دمای ساعتی	فصلی بودن، روند، تأخیر
داده مکانی-زمانی(Spatio-Temporal Data)	داده‌هایی با بعد مکان و زمان	مسیر خودرو، تغییر دما در نقشه زمانی	مدل‌سازی دو بعدی (مکان + زمان)
داده متنی(Text Data)	داده‌هایی به‌صورت متن طبیعی	مقالات، توییت‌ها، گزارش‌ها	ابهام زبانی، نیاز به پردازش زبان طبیعی
داده شبکه‌ای / گرافی (Network / Graph Data)	داده‌هایی که روابط بین گره‌ها را نشان می‌دهند	شبکه اجتماعی، ارتباط میان سرورها	مقیاس‌پذیری، تراکم، جهت‌دار بودن
داده چندرسانه‌ای (Multimedia Data)	داده‌های تصویری، صوتی یا ویدیویی	تصاویر پزشکی، ویدیوهای نظارتی	حجم بالا، ویژگی‌های پیچیده
داده ترکیبی / ناهمگن (Heterogeneous Data)	ترکیب چند نوع داده باهم	سامانه‌های هوشمند شهری (متن + مکان + زمان)	همسان‌سازی ساختار و مقیاس

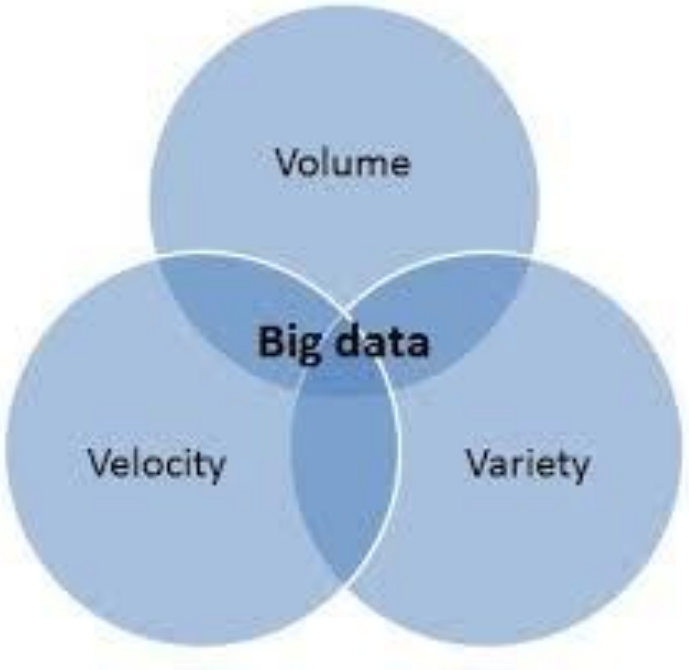
کلان داده یا Big Data

برخی از مهم‌ترین انواع داده‌های بزرگ

- داده‌های تراکنشی :(Transactional data) مانند سوابق فروش و تراکنش‌های مالی
- داده‌های ماشین :(Machine data) مانند داده‌های حسگر، لاگ‌های سرور و داده‌های اینترنت اشیا

داده‌های اجتماعی :(Social data) مانند پست‌های شبکه‌های اجتماعی و محتواها

- داده‌های متنی :(Text data) مانند ایمیل‌ها، اسناد



Three Vs of Big Data

۱. حجم :(Volume) مقدار بسیار زیاد داده‌ها

۲. تنوع :(Variety) انواع مختلف داده‌ها (ساخت‌یافته، نیمه‌ساخت‌یافته و غیرساخت‌یافته)

۳. سرعت :(Velocity) سرعت تولید و جریان سریع داده‌ها

داده ها از منظر کاربرد در هوش مصنوعی

نوع هوش مصنوعی	داده‌های ورودی معمول
یادگیری ماشین (Machine Learning)	داده‌های ساخت‌یافته و کمی، متنی، برچسب‌گذاری شده
یادگیری عمیق (Deep Learning)	داده‌های حجیم، تصاویر، ویدیو، صوت، متن
NLP (پردازش زبان طبیعی)	متن، گفتار، ترجمه، چت‌ها
بینایی ماشین (Computer Vision)	عکس، ویدیو، نقشه‌ها، تصاویر پزشکی
توصیه‌گرها (Recommendation Systems)	داده‌های رفتاری کاربران، امتیازدهی، تاریخچه خرید

وظایف استاندارد در علم داده

- مدلسازی مساله
- خوشه بندی
- پیش‌بینی (شامل طبقه بندی و رگرسیون)
- تشخیص ناهنجاری (تشخیص داده پرت).
- کاوش قواعد انجمنی
- ..

سهم علوم و فنون مختلف



Any question?