

NAUKA O DANYCH

NIEZBĘDNIK ANALITYKA



STATYSTYKA (ANG. STATISTICS)

Nauka, która zajmuje się zbieraniem, analizowaniem i interpretacją zjawisk zachodzących np. w gospodarce. Zamiast koncentrować się na pojedynczych przypadkach, statystyka wykorzystuje metody, które umożliwiają analizę zbiorów danych, co pozwala na opisywanie i wnioskowanie o cechach całej grupy (populacji) na podstawie badanych próbek.

ALGORYTM (ANG. ALGORITHM)

Zbiór jasno określonych instrukcji służących do wykonania zadania lub rozwiązania problemu. W informatyce i matematyce, algorytm dokładnie opisuje kroki niezbędne do osiągnięcia określonego wyniku. Algorytmy są kluczowym elementem w tworzeniu modeli, ponieważ definiują sposób, w jaki model uczy się z danych i dokonuje przewidywania (predykcji).

ŚREDNIA ARYTMETYCZNA (ANG. MEAN)

Miara statystyczna, która jest obliczana poprzez zsumowanie wszystkich wartości w zbiorze danych i podzielona przez liczbę elementów w tym zbiorze.

MODEL PREDYKCYJNY (ANG. PREDICTIVE MODEL)

Model matematyczny lub statystyczny wykorzystywany do przewidywania prawdopodobnych wyników na podstawie wcześniejszych danych. Wykorzystuje się go do przewidywania wartości nieznanymi zmiennymi (cech) lub przyszłych zdarzeń. Budowanie modelu predykcyjnego wymaga zastosowania algorytmów statystycznych lub uczenia maszynowego do danych treningowych, które pomagają odkryć zależności między zmiennymi.

SZTUCZNA INTELIGENCJA (ANG. ARTIFICIAL INTELLIGENCE)

Dziedzina informatyki, która umożliwia maszynom wykonywanie zadań, które tradycyjnie wymagały ludzkiego intelektu. Sztuczna inteligencja działa przez naśladowanie sposobu myślenia ludzi, ucząc się i rozumiejąc informacje. Wykorzystuje algorytmy, które są zestawami instrukcji do przetwarzania danych i generowania odpowiedzi.

MEDIANA (ANG. MEDIAN)

Wartość środkowa w uporządkowanym zestawie danych. Aby ją znaleźć, należy posortować wszystkie obserwacje od najmniejszej do największej. Jeśli liczba obserwacji jest nieparzysta, mediana jest wartością, która znajduje się na środku listy. Jeśli liczba obserwacji jest parzysta, mediana jest średnią arytmetyczną dwóch środkowych wartości.

BAZA DANYCH (ANG. DATABASE)

Narzędzie służące do zbierania, przechowywania i zarządzania informacjami. Baza danych składa się z rekordów, czyli zbiorów danych, które są ze sobą powiązane. Te rekordy są uporządkowane w tabelach, gdzie każda tabela zawiera zbiór rekordów tego samego rodzaju. Każdy rekord w tabeli ma takie same pola, nazywane również kolumnami, które określają rodzaj i cechy danych zapisanych w rekordzie.

UCZENIE MASZYNOWE (ANG. MACHINE LEARNING)

Dziedzina sztucznej inteligencji, która zajmuje się tworzeniem algorytmów zdolnych do uczenia się i dostosowywania się do danych bez konieczności bezpośredniego programowania. Uczenie maszynowe działa poprzez analizę i interpretację dużych zbiorów danych, aby wydobyc z nich użyteczne wzorce lub przewidywania. Algorytmy uczenia maszynowego uczą się z dostępnych danych i dostosowują swoje parametry.

GENERATYWNA SZTUCZNA INTELIGENCJA (ANG. GENERATIVE ARTIFICIAL INTELLIGENCE)

Rodzaj technologii wykorzystujący sztuczną inteligencję, który skupia się na generowaniu nowych treści, danych lub obiektów, które mogą być nieodróżnialne od stworzonych przez człowieka. Wykorzystuje zaawansowane algorytmy uczenia maszynowego, aby analizować istniejące dane i na ich podstawie tworzyć nowe, unikatowe wyniki.

ODCHYLENIE STANDARDOWE (ANG. STANDARD DEVIATION)

Miara rozproszenia danych wokół średniej arytmetycznej. Określa, jak bardzo poszczególne wartości w zbiorze danych różnią się od średniej. Im większe odchylenie standardowe, tym większe zróżnicowanie wartości w zbiorze.

CHMURA OBLICZENIOWA (ANG. CLOUD COMPUTING)

Usługa dostarczająca zasoby komputerowe, takie jak moc obliczeniowa, pamięć, przechowywanie danych czy aplikacje, poprzez internet na żądanie i według wybranego modelu płatności. Umożliwia ona użytkownikom dostęp do danych i oprogramowania z każdego miejsca na świecie, gdzie jest dostęp do sieci.

BIG DATA

Termin używany do opisu ogromnych, złożonych zbiorów danych, które ze względu na swoją wielkość, szybkość powiększania lub różnorodność wymagają specjalnych metod przetwarzania w celu wydobycia wartościowych informacji.

DATA MINING

Proces wydobywania danych, czyli analizy dużych zbiorów danych w celu odkrywania wzorców, trendów i innych użytecznych informacji, które nie są dostrzegalne na pierwszy rzut oka.



SGH



Projekt realizowany z Narodowym Bankiem Polskim w ramach programu edukacji ekonomicznej