

Mineração de Dados

Trabalho 2

Arthur do Prado Labaki - 11821BCC017

13-04, 2023

GBC212

Estudo do Conceito

Variáveis binárias, discretas e contínuas se referem à natureza dos atributos ou variáveis que podem ser medidas em uma pesquisa ou estudo. Em mineração de dados, as variáveis binárias, discretas e contínuas são utilizadas para descrever as características dos dados coletados. Com isso, temos:

As variáveis binárias são usadas para representar dados categóricos com apenas duas opções possíveis, como "sim" ou "não", "verdadeiro" ou "falso", etc. Esses dados são comumente usados em análises de classificação ou decisão.

Já as variáveis discretas são usadas para representar dados categóricos com um número finito ou contável de opções possíveis, como o número de filhos de uma pessoa ou o número de pessoas em uma família. Esses dados são comumente usados em análises de frequência ou de associação.

As variáveis contínuas, por outro lado, são usadas para representar dados numéricos que podem assumir qualquer valor dentro de um intervalo contínuo, como a altura ou o peso de uma pessoa. Esses dados são comumente usados em análises estatísticas e de regressão.

Também existem as variáveis qualitativas, que são usadas para descrever características dos dados que não podem ser medidas numericamente, como a cor dos olhos ou o gênero. As variáveis nominais, que não possuem uma ordem natural, são comumente usadas em análises de agrupamento ou classificação, enquanto as variáveis ordinais, que possuem uma ordem natural, são usadas em análises de tendência ou de ordenação.

Por fim, as variáveis quantitativas são usadas para descrever características dos dados que podem ser medidas numericamente, como a altura de uma pessoa ou o salário. As variáveis intervalares, que não possuem um zero absoluto significativo, são usadas em análises de correlação ou regressão, enquanto as variáveis proporcionais, que possuem um zero absoluto significativo, são usadas em análises de razão ou proporção.

Resolução do Exercício 1)

Classifique os atributos a seguir como binário, discreto ou contínuo. Classifique-os também como qualitativo (nominal ou ordinal) ou quantitativo (intervalo ou proporção). Se algum caso parece ambíguo, explique as interpretações possíveis. Exemplo: Idade em anos. Resposta: Discreto, Quantitativo, Proporção.

Resposta

- a) Tempo em termos de AM ou PM - Binário, Qualitativo, Nominal;
- b) Claridade como uma medida da intensidade da luz - Contínuo, Quantitativo, Proporção;
- c) Claridade como uma medida para julgamento das pessoas - Discreto, Qualitativo, Ordinal;
- d) Ângulos como uma medida em graus entre 0° e 360° - Contínuo, Quantitativo, Proporção;
- e) Medalhas de bronze, prata e ouro como premiação nas Olimpíadas - Discreto, Qualitativo, Ordinal;
- f) Altura acima do nível do mar - Contínuo, Quantitativo, Proporção;
- g) Número de pacientes em um hospital - Discreto, Quantitativo, Proporção;
- h) Número de ISBN de livros - Discreto, Qualitativo, Nominal;
- i) Relação com a passagem da luz em termo dos seguintes valores: opaco, translúcido, transparente - Discreto, Qualitativo, Nominal;
- j) Rank militar - Discreto, Qualitativo, Ordinal;
- k) Distância do centro do Campus - Contínuo, Quantitativo, Proporção;
- l) Densidade de uma substância em gramas por centímetro cúbico - Contínuo, Quantitativo, Proporção;
- m) Senha para retirar produto - Discreto, Qualitativo, Ordinal.

Explicação

- a) Tempo em termos de AM ou PM: é binário, pois há apenas duas opções, manhã (AM) ou tarde (PM). É qualitativo, nominal, pois não há uma ordem lógica para os valores.
- b) Claridade como medida da intensidade da luz: é contínuo, pois pode assumir qualquer valor num intervalo contínuo. É quantitativo, proporção, porque tem uma unidade de medida padrão e a relação entre os valores é significativa, ou seja, o dobro de claridade significa o dobro da intensidade de luz.
- c) Claridade como medida para julgamento das pessoas: é discreto, pois assume apenas valores inteiros e limitados (por exemplo, numa escala de 1 a 10). É qualitativo, ordinal, porque há uma ordem lógica para os valores.
- d) Ângulos como medida em graus entre 0º e 360º: é contínuo, pois pode assumir qualquer valor num intervalo contínuo. É quantitativo, proporção, porque tem uma unidade de medida padrão e a relação entre os valores é significativa, ou seja, um ângulo de 180 graus é a metade de um ângulo de 360 graus.
- e) Medalhas de bronze, prata e ouro como premiação nas Olimpíadas: é discreto, pois assume valores inteiros e limitados (apenas 3). É qualitativo, ordinal, porque há uma ordem lógica para os valores.
- f) Altura acima do nível do mar: é contínuo, pois pode assumir qualquer valor num intervalo contínuo. É quantitativo, proporção, porque tem uma unidade de medida padrão e a relação entre os valores é significativa, ou seja, uma diferença de 100 metros significa uma diferença de 100 metros em relação ao nível do mar.
- g) Número de pacientes em um hospital: é discreto, pois assume valores inteiros e limitados (por exemplo, de 0 a alguns milhares). É quantitativo, proporção, porque tem uma unidade de medida padrão e a relação entre os valores é significativa, ou seja, o dobro de pacientes significa o dobro de pessoas no hospital.
- h) Número de ISBN de livros: é discreto, pois assume valores inteiros e limitados (por exemplo, um livro pode ter apenas um único número de ISBN). É qualitativo, nominal, pois não há uma ordem lógica para os valores.
- i) Relação com a passagem da luz em termos dos seguintes valores: opaco, translúcido,

transparente: é discreto, pois assume apenas valores inteiros e limitados (três). É qualitativo, nominal, pois não há uma ordem lógica para os valores.

j) Rank militar: é discreto, pois assume valores inteiros e limitados (por exemplo, de 1 a 5 estrelas). É qualitativo, ordinal, porque há uma ordem lógica para os valores.

k) Distância do centro do Campus: é contínuo, pois pode assumir qualquer valor num intervalo contínuo. É quantitativo, proporção, porque tem uma unidade de medida padrão e a relação entre os valores é significativa, ou seja, uma distância de 100 metros significa uma diferença de 100 metros em relação ao centro do campus.

l) Densidade de uma substância em gramas por centímetro cúbico: é contínuo, pois pode assumir qualquer valor num intervalo contínuo. É quantitativo, proporção, porque tem uma unidade de medida padrão e a relação entre os valores é significativa, ou seja, uma densidade duas vezes maior significa que há o dobro da massa por volume.

m) Senha para retirar produto: é discreto, pois normalmente é formado por um número inteiro de um até mil. É qualitativo, ordinal, pois existe uma ordem lógica para os valores.