

Scale	True Zero	Equal Intervals	Order	Category	Example
Nominal	No	No	No	Yes	Marital Status, Sex, Gender, Ethnicity
Ordinal	No	No	Yes	Yes	Student Letter Grade, NFL Team Rankings
Interval	No	Yes	Yes	Yes	Temperature in Fahrenheit, SAT Scores, IQ, Year
Ratio	Yes	Yes	Yes	Yes	Age, Height, Weight

Aula 9

Tipos de Dados

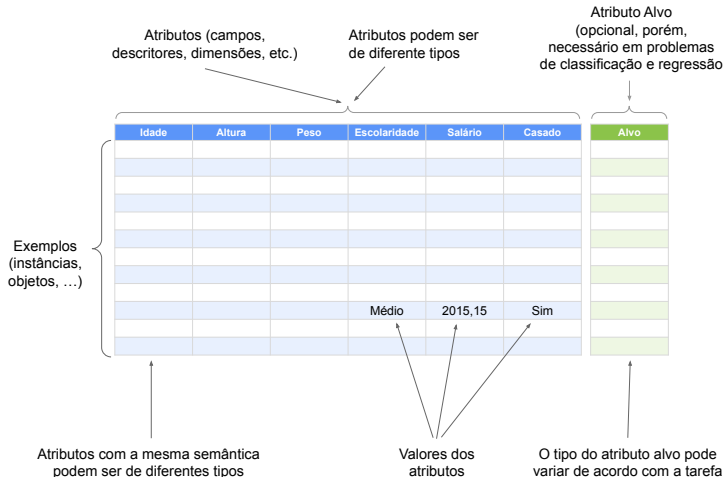
Introdução

- Conhecer as características dos dados torna fácil algumas tarefas
 - Preencher valores ausentes
 - Tratar valores ruidosos
 - Utilizar medidas de distância
 - Conceitos sobre como os algoritmos devem considerar os dados no aprendizado
 - ...

O que é um Atributo?

- Conjuntos de dados são constituídos de objetos de dados
- Os objetos de dados também são conhecidos por
 - Exemplos
 - Instâncias
 - Tuplas
 - ...
- Os objetos podem ser consumidores, itens de venda, valores de venda, pacientes, universidades, estudantes, professores, cursos, ...
- Um atributo é um campo do objeto que representa uma característica do objeto

O que é um Atributo?



O que é um Atributo?

- Os atributos também são conhecidos por
 - Característica
 - Dimensão
 - Coluna (se pensarmos em termos de tabela)
 - ...
- Exemplos de atributos de consumidores: ID, nome, endereço, ...
- Um conjunto de atributos utilizado para descrever um dado objeto é chamado de vetor de atributos (ou vetor de características)
- O tipo de um atributo é determinado pelo conjunto dos possíveis valores: **nominal**, **binário**, **ordinal** ou **numérico**

- **Nominal significa “relacionado a nomes”**
- Os valores de um atributo nominal são símbolos ou nomes de “coisas”
- Cada valor representa uma categoria, código, estado, ...
- Exemplo de valores de atributos nominais
 - Cor de cabelo: preto, castanho, ruivo, branco, ...
 - Relacionamento: solteiro, casado, divorciado, ...
 - Ocupação: professor, dentista, programador, ...
- **Operações matemáticas nos valores de atributos nominais não são significativas**
- Também não faz sentido calcular média ou mediana
- **Pode-se calcular a moda**

- Um atributo binário é um atributo nominal com apenas duas categorias ou estados: **0** ou **1**, **presença** ou **ausência**, **verdadeiro** ou **falso**, **sim** ou **não**, **positivo** ou **negativo**
- Também conhecidos por atributos booleanos
- Exemplo de valores de atributos binários
 - Fumante: não
 - Estudante: sim
 - Resultado Teste Médico: positivo

- **Um atributo ordinal é um atributo com valores que possuem uma ordem ou ranking**
- A magnitude entre os valores não é conhecida
- Exemplo de valores de atributos ordinais
 - Idade: adolescente, adulto, idoso
 - Altura: baixa, média, alta
 - Temperatura: baixa, alta
- **Moda e Mediana pode ser utilizada**

- Atributos ordinais podem também ser obtidos pela discretização de quantidades numéricas por meio da criação de intervalos em um número finitos de categorias ordenadas
 - Altura
 - $[0 - 1, 50] \rightarrow \textit{baixo}$
 - $]1, 50 - 1, 70] \rightarrow \textit{média}$
 - $]1, 70 - \infty[\rightarrow \textit{alta}$
 - Idade
 - $[0 - 10] \rightarrow \textit{criança}$
 - $[11 - 21] \rightarrow \textit{adolescente}$
 - $[22 - 59] \rightarrow \textit{adulto}$
 - $[60 - \infty[\rightarrow \textit{idoso}$

- **Os valores dos atributos numéricos são quantidades mensuráveis, representados por valores inteiros ou reais**
- Pode-se computar média, mediana, moda, ...
- Os atributos podem ser do tipo *escala intervalar* e *escala de razão*
- **Atributos de escala intervalar**
 - Os valores possuem ordem e podem ser positivos, 0 ou negativos
 - Permitem a comparação e quantificação entre os valores

● Atributos de escala intervalar

- Taxas entre números não são significativos, portanto, operações como divisões e multiplicações não podem ser executadas diretamente.
 - Temperatura Celsius: 2° não significa o dobro da temperatura de 1°
 - Calendário: 10 d.c. não significa o dobro de anos de 5 d.c.

● Atributos de escala de razão

- É um atributo numérico com um ponto 0 inerente
- Taxas entre número são significativas
 - Temperatura Kelvin
 - Anos de experiência
 - Idade
 - ...

● Atributos discretos

- Tem um número finito de valores
- Atributos nominais, binários e ordinais se enquadram nesta categoria
- Exemplos
 - Cor de cabelo
 - Fumante
 - Teste Médico
 - Idade

● Atributos contínuos

- Os valores dos atributos são números reais
- Atributos numéricos se enquadram nesta categoria
- Exemplos
 - Temperatura
 - Peso
 - Altura

Tabela: Tabela comparativa entre os atributos

Tipo Atr.	Descrição	Exemplos	Operações	
Discreto (Categórico)	Nominal	Os valor de um atributo nominal apenas provê informação para distinguir informação entre os objetos (=, ≠)	Cor de cabelo Relacionamento	Moda, entropia, correlação, χ^2
	Ordinal	Os valores de um atributo ordinal provêm informação para ordenar os objetos (<, >)	Notas (A,B,C) Altura (pequeno, médio, alto)	Mediana, percentis, correlação de ranking
Contínuo (Numérico)	Intervalar	A diferença entre os valores são significativas (+, -)	Datas de calendário Temperatura em Celsius	Média, desvio padrão, Correlação de Pearson
	Razão	Diferenças e taxas são significativas (*, /)	Temperatura Kelvin Quantidades monetárias	Média geométrica, média Harmônica, variação percentual

Material Complementar

- Tipos de Variáveis

http://www.ernestoamaral.com/docs/dcp030d-101/var_reg.pdf

- ANALISANDO DADOS (Cap. 1)

https://epge.fgv.br/we/Graduacao/Estatistica1/2009/2?action=AttachFile&do=get&target=Estatistica_descritiva.pdf

- TIPOS BÁSICOS DE DADOS

<https://pt.slideshare.net/alexandrend/introduo-anlise-de-dados-aula-02-tipos-bsicos-de-dados>

Imagem do Dia



Inteligência Artificial
<http://lives.ufms.br/moodle/>

Rafael Geraldeli Rossi
rafael.g.rossi@ufms.br

Slides baseados em [Han et al., 2011] e [Tan et al., 2005]

Referências Bibliográficas I



Han, J., Kamber, M., and Pei, J. (2011).

Data Mining: Concepts and Techniques.

The Morgan Kaufmann Series in Data Management Systems.
Elsevier.



Tan, P.-N., Steinbach, M., and Kumar, V. (2005).

Introduction to Data Mining.

Addison-Wesley.