[Flash! 08]



Oitavo seminário de pesquisas em curso do Nomads.usp | 2020 Programa de Pós-Graduação em Arquitetura e Urbanismo - IAU USP Nomads.usp - Núcleo de Estudos sobre Habitares Interativos

Apresentação [Flash! 08 - MÉTODO]

Modelagem da informação e estratégias metodológicas para o aprendizado de máquina

Pesquisa de Mestrado em curso [2019-2022]

Inteligência Artificial (IA) e Arquitetura: caminhos possíveis de aplicação e melhoria no processo de projeto

Pesquisador

Christian Jhulian Braga Quesada

Orientadora

Profa. Dra. Anja Pratschke

Delimitação da pesquisa em três grandes áreas de interesse:

Representação do Conhecimento

Aprendizado de Máquina

Tomada de Decisões

Delimitação da pesquisa em três grandes áreas de interesse:

Representação do Conhecimento Aprendizado de Máquina

Tomada de Decisões

Delimitação da pesquisa em três grandes áreas de interesse:



Delimitação da pesquisa em três grandes áreas de interesse:



Linguagem de Programação

Delimitação da pesquisa em três grandes áreas de interesse:



Linguagem de Programação

Linguagem de Baixo Nível

mais próxima da linguagem de máquina

Delimitação da pesquisa em três grandes áreas de interesse:

Representação do Conhecimento

Aprendizado de Máquina

Decisões

Linguagem de Programação

Linguagem de Baixo Nível

mais próxima da linguagem de máquina

Ex. comandos de máquina (código binário) 01001101 01110101 01101110

Delimitação da pesquisa em três grandes áreas de interesse:

Representação do Conhecimento

Aprendizado de Máquina

Tomada de Decisões

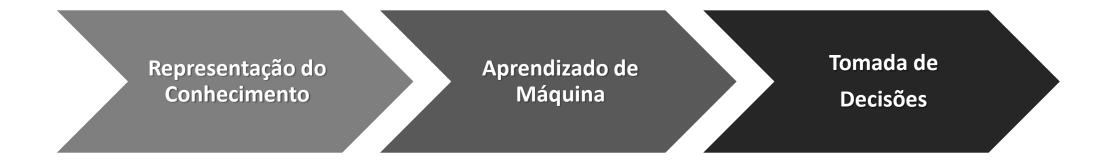
Linguagem de Programação

Linguagem de Baixo Nível mais próxima da linguagem de máquina

Ex. comandos de máquina (código binário) 01001101 01110101 01101110

Linguagem de Alto Nível que se aproxima mais da linguagem humana

Delimitação da pesquisa em três grandes áreas de interesse:



Linguagem de Programação

Linguagem de Baixo Nível

mais próxima da linguagem de máquina

Ex. comandos de máquina (código binário) 01001101 01110101 01101110

Linguagem de Alto Nível

que se aproxima mais da linguagem humana

Ex. Java; C#; Ruby; Python; Swift; JavaScript; PHP 1 | print ("Esse é um exemplo em Python!")

Delimitação da pesquisa em três grandes áreas de interesse:



Linguagem de Programação

Linguagem de Baixo Nível

mais próxima da linguagem de máquina

Ex. comandos de máquina (código binário) 01001101 01110101 01101110

Linguagem de Alto Nível

que se aproxima mais da linguagem humana

Ex. Java; C#; Ruby; Python; Swift; JavaScript; PHP 1 | print ("Esse é um exemplo em Python!")

BIM - Building Information Modeling

"[...] definimos BIM como uma <u>tecnologia de modelagem</u> e um <u>conjunto associado de processos</u> para produzir, comunicar e analisar modelos de construção." (EASTMAN et al., 2014)

Modelos de Construção são caracterizados por:

- Componentes de construção que são representados com representações digitais inteligentes (objetos) que "sabem" o que eles são, e que podem ser associados com atributos (gráficos e de dados) computáveis e regras paramétricas.
- Componentes que <u>incluem dados</u> que descrevem como eles se comportam conforme são necessários para análises e processos de trabalho, por exemplo, <u>quantificação</u>, especificação e análise energética.
- <u>Dados consistentes e não redundantes</u> de forma que as modificações nos dados dos componentes sejam representadas em todas as visualizações dos componentes.
- <u>Dados coordenados</u> de forma que todos as visualizações de um modelo sejam representadas de maneira coordenada.

print ("AI4BIM")

Artificial Intelligence for Building Information Modeling

- 8. O futuro: construindo com BIM
- 8.5. Impulsionadores de <u>mudanças e impactos do BIM até 2020</u>
- [...] Outra área técnica que pode introduzir desenvolvimentos adicionais que influenciam sistemas BIM é aquela a que se faz referência como <u>inteligência artificial</u>. Ferramentas BIM são plataformas convenientes para o <u>renascimento</u> do desenvolvimento de <u>sistemas especialistas</u> para uma gama de propósitos, como <u>verificação de conformidade com códigos, revisões</u> <u>qualitativas, ferramentas inteligentes para comparação de versões, guias e tutoriais de projetos.</u>

(EASTMAN et al., 2014)

Aprendizado de Máquina

A capacidade de aprendizado é considerada essencial para um comportamento inteligente. Atividades como memorizar, observar e explorar situações para aprender fatos, melhorar habilidades motoras/cognitivas por meio de práticas e organizar conhecimento novo em representações apropriadas podem ser consideradas atividades relacionadas ao aprendizado.

Os algoritmos de aprendizado de máquina aprendem a induzir uma função ou hipótese capaz de resolver um problema a partir dos dados que representam instâncias do problema a ser resolvido. A essa propriedade de <u>uma hipótese continuar a ser válida para novos objetos</u> dá-se o nome de capacidade de generalização de hipótese. Ou seja, para ser útil quando aplicada a novos dados, uma hipótese precisa apresentar <u>boa capacidade de generalização</u>. A esse processo de indução de uma hipótese (ou aproximação de função) a partir da experiência passada dá-se o nome Aprendizado de Máquina (AM).

Aprendizado de Máquina

Em AM [Aprendizado de Máquina], computadores são programados para aprender com a experiência passada. Para tal, empregam um princípio de inferência denominado indução, no qual se obtêm conclusões genéricas a partir de um conjunto particular de exemplos. Esses dados formam um conjunto, simplesmente denominado como conjunto de dados.

Quando um algoritmo de aprendizado de máquina está aprendendo a partir de um conjunto de dados de treinamento, ele está procurando uma hipótese, no espaço de possíveis hipóteses, capaz de descrever as relações entre os objetos que melhor se ajuste aos dados de treinamento. Dessa forma, cada algoritmo de AM utiliza uma forma ou uma representação para descrever a hipótese induzida.

Viés de Representação

Por exemplo, <u>redes neurais</u> artificiais representam uma hipótese por um conjunto de valores reais, <u>associado aos pesos das conexões da rede</u>, enquanto as <u>árvores de decisão</u> utilizam uma <u>estrutura de árvore em que cada nó interno é representado por uma pergunta</u> referente ao valor de um atributo e cada nó externo está associado a uma classe. Dessa forma, a representação utilizada pode <u>definir a preferência ou viés de representação do algoritmo</u>, podendo restringir o conjunto de hipóteses que podem ser induzidas.

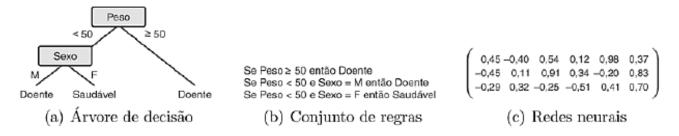


Figura 1.1 Diferentes vieses de representação.

Viés de Busca

Além do viés de representação, os algoritmos de AM possuem também um <u>viés de busca</u>. O viés de busca de um algoritmo é a <u>forma como o algoritmo busca a hipótese</u> que melhor se ajusta aos dados de treinamento, ou seja, ele define como as hipóteses são pesquisadas no espaço de hipóteses. <u>Por exemplo</u>, alguns algoritmos que se utilizam da representação por árvores de decisão, <u>podem ter como viés de busca a preferência por árvores de decisão com poucos nós.</u>

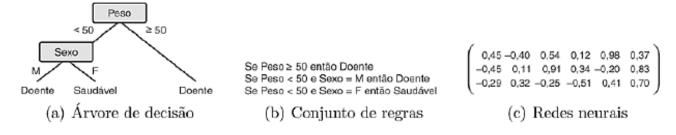
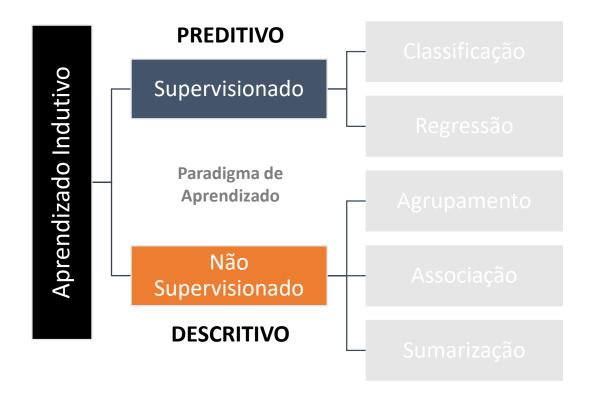
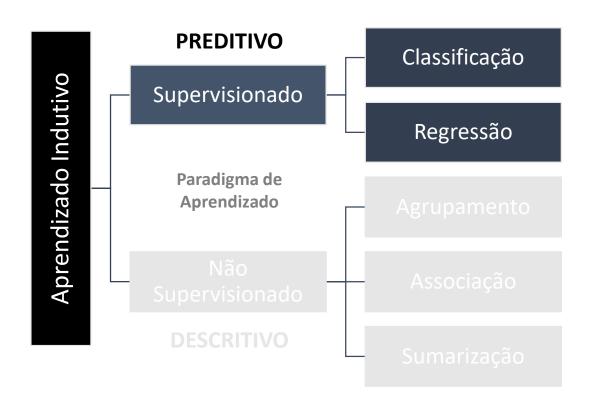


Figura 1.1 Diferentes vieses de representação.



Algoritmos de AM tem sido amplamente utilizados em diversas tarefas, que podem ser organizadas de acordo com diferentes critérios. Um deles diz respeito ao <u>paradigma de aprendizado a ser adotado para lidar com a tarefa.</u>

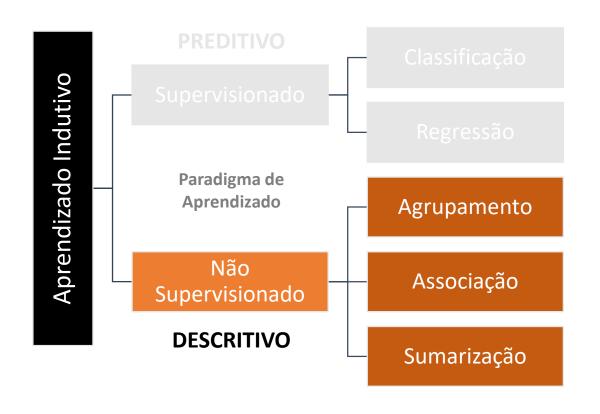
No topo aparece o aprendizado indutivo, processo pelo qual são realizadas as generalizações a partir dos dados. Tem-se em seguida os tipos de aprendizado supervisionado (preditivo) e não supervisionado (descritivo).



Em tarefas de previsão, a meta é encontrar uma função (modelo ou hipótese) a partir dos dados de treinamento que possa ser utilizada para prever um rótulo ou valor que caracterize um novo exemplo, com base nos valores dos seus atributos de entrada. Para isso, cada objeto do conjunto de treinamento deve possuir atributos de entrada e saída.

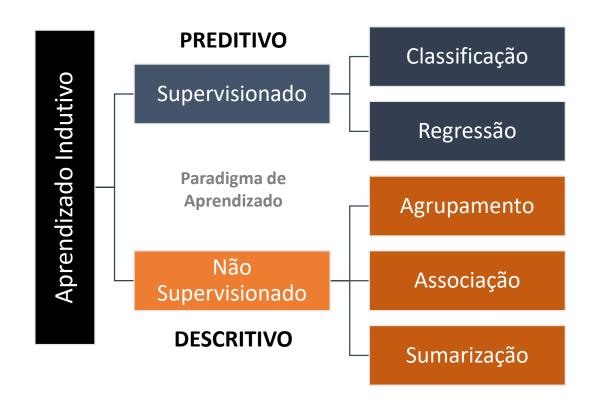
O termo supervisionado vem da simulação da presença de um "supervisor externo", que conhece a saída (rótulo) desejada para cada exemplo (conjunto de valores para os atributos de entrada). Com isso, o supervisor externo pode avaliar a capacidade da hipótese induzida de predizer o valor de saída para novos exemplos.

As tarefas supervisionadas se distinguem pelo tipo de rótulos dos dados: discreto, no caso de classificação; e contínuo, no caso de regressão.



Em tarefas de descrição, a meta é explorar ou descrever um conjunto de dados. Os algoritmos de AM utilizados nessas tarefas não fazem uso do atributo de saída. Por isso, seguem o paradigma de aprendizado não supervisionado.

As tarefas descritivas são genericamente divididas em: agrupamento, em que os dados são agrupados de acordo com sua similaridade; associação, que consiste em encontrar padrões frequentes de associações entre os atributos de um conjunto de dados; e sumarização, cujo objetivo é encontrar uma descrição simples e compacta para um conjunto de dados.



Deve ser observado que, apesar dessa divisão básica de modelos em preditivos e descritivos, <u>um modelo preditivo</u> também provê uma descrição compacta de um conjunto de dados e um modelo descritivo pode prover previsões após ser validado.

Algumas tarefas de aprendizado não se enquadram na hierarquia apresentada [...] três dessas tarefas são: aprendizado <u>semissupervisionado</u> (Zhu et al., 2009), aprendizado <u>ativo</u> (Settles, 2012) e aprendizado <u>por reforço</u> (Sutton e Barto, 1998).

Pré-processamento de dados

Técnicas de pré-processamento de dados são frequentemente utilizadas para <u>melhorar a qualidade dos dados</u> por meio da eliminação ou <u>minimização dos problemas</u> citados. Essa melhora pode facilitar o uso de técnicas de AM [Aprendizado de Máquina], levar à construção de modelos mais fiéis à distribuição real dos dados, reduzindo sua complexidade computacional, tornar mais fáceis e rápidos o ajuste de parâmetros do modelo e seu posterior uso. Isso pode, adicionalmente, <u>facilitar a interpretação dos padrões extraídos pelo modelo.</u>

Técnicas de pré-processamento de dados são úteis não apenas porque podem minimizar ou eliminar problemas existentes em um conjunto de dados, mas também porque **podem tornar os dados mais adequados para sua utilização** por um determinado algoritmo de AM. **Por exemplo, alguns algoritmos de AM trabalham apenas com valores numéricos.**

De acordo com o pesquisador Júlio Cesar Pereira (2004), <u>a assunção de premissas (postulados, proposições, axiomas) é inerente à estrutura do conhecimento científico e condiciona o entendimento de objetividade e verdade, ou seja, de representação do real.</u> Repensar essas premissas deve auxiliar o pesquisador a conceber suas estratégias de investigação.

A distinção entre ser e atributo, de que se ocupa a filosofia, constitui-se em elemento importante para que o pesquisador defina a delimitação de seu objeto de estudo.

A classificação das variáveis em tipos e escalas envolvem o estabelecimento de premissas de relação entre atributos de um objeto e a representação simbólica desses atributos. A escala atribui rótulos numéricos aos atributos e é definida pelo pesquisador. Disto resulta que a fidedignidade da representação que os números fazem dos atributos é um juízo primário do investigador, o qual se apoia no seu conhecimento do objeto.

QUANTITATIVA QUALITATIVA

DISCRETA NOMINAL

CONTÍNUA ORDINAL

QUANTITATIVA

QUALITATIVA

DISCRETA

<u>Números inteiros</u>, sem frações, como em contagens. Constituem um conjunto finito.

CONTÍNUA

Números que podem assumir valores

<u>fracionários.</u> Normalmente têm intervalos de valores conhecido, mas um conjunto infinito de valores possíveis.

NOMINAL

ORDINAL

Ilustração: ArchiCAD 22

QUANTITATIVA

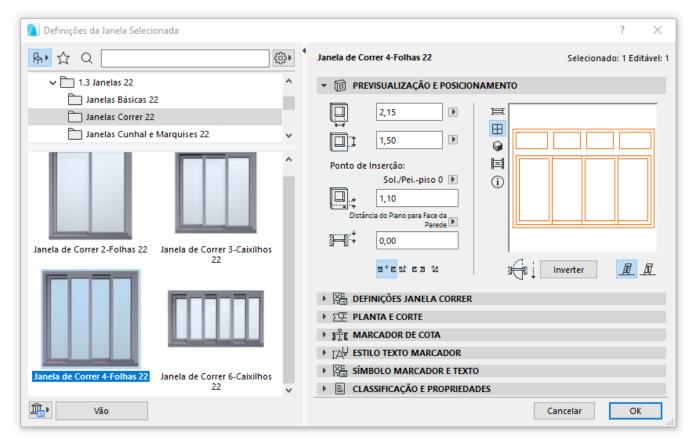
DISCRETA

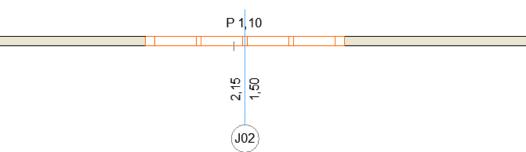
<u>Números inteiros</u>, sem frações, como em contagens. Constituem um conjunto finito.

CONTÍNUA

Números que podem assumir valores

<u>fracionários.</u> Normalmente têm intervalos de valores conhecido, mas um conjunto infinito de valores possíveis.





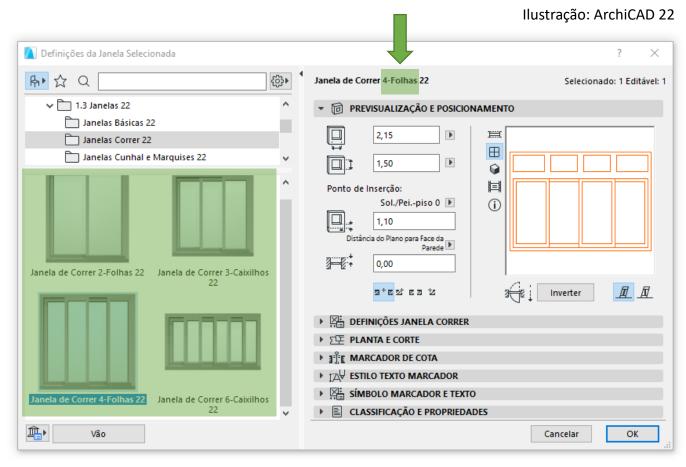
QUANTITATIVA

DISCRETA

Números inteiros, <u>sem frações</u>, como em contagens. Constituem um <u>conjunto finito</u>.

CONTÍNUA

Números que podem assumir valores fracionários. Normalmente têm intervalos de valores conhecido, mas um conjunto infinito de valores possíveis.



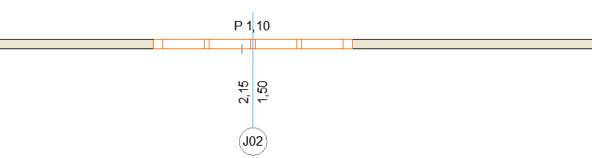


Ilustração: ArchiCAD 22

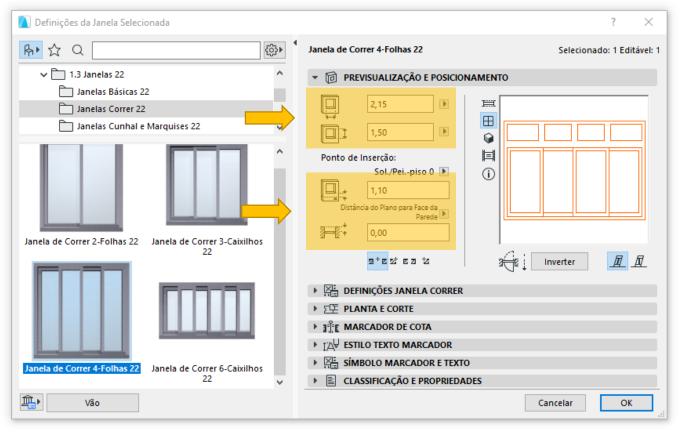
QUANTITATIVA

DISCRETA

Números inteiros, sem frações, como em contagens. Constituem um conjunto finito.

CONTÍNUA

Números que podem assumir <u>valores</u> <u>fracionários.</u> Normalmente têm intervalos de valores conhecido, mas um conjunto infinito de valores possíveis.



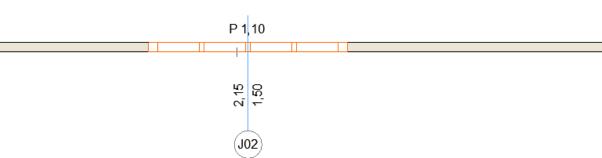


Ilustração: ArchiCAD 22

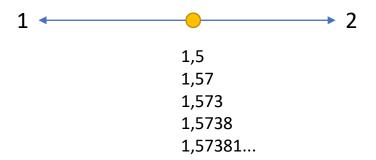
QUANTITATIVA

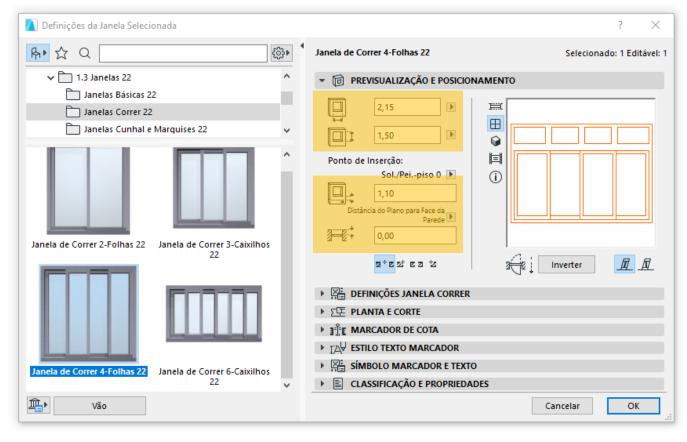
DISCRETA

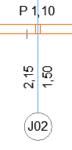
Números inteiros, sem frações, como em contagens. Constituem um conjunto finito.

CONTÍNUA

Números que podem assumir valores fracionários. Normalmente têm intervalos de valores conhecido, mas um conjunto infinito de valores possíveis.







QUANTITATIVA

DISCRETA

Números inteiros, sem frações, como em contagens. Constituem um conjunto finito.

CONTÍNUA

Números que podem assumir valores fracionários. Normalmente têm intervalos de valores conhecido, mas um conjunto infinito de valores possíveis.

QUALITATIVA

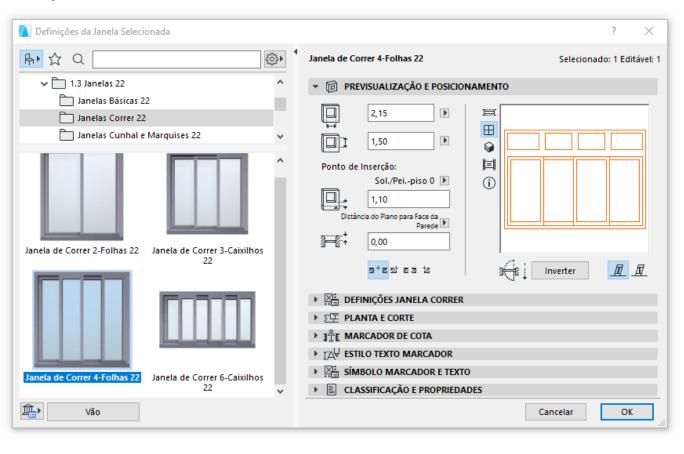
NOMINAL

<u>Categorias</u>, sendo que cada categoria é <u>independente, sem relação com as outras.</u>

ORDINAL

<u>Categorias</u>, sendo que cada categoria <u>mantém</u> <u>uma relação de ordem com as outras.</u>

Ilustração: ArchiCAD 22



P 1,10

QUALITATIVA

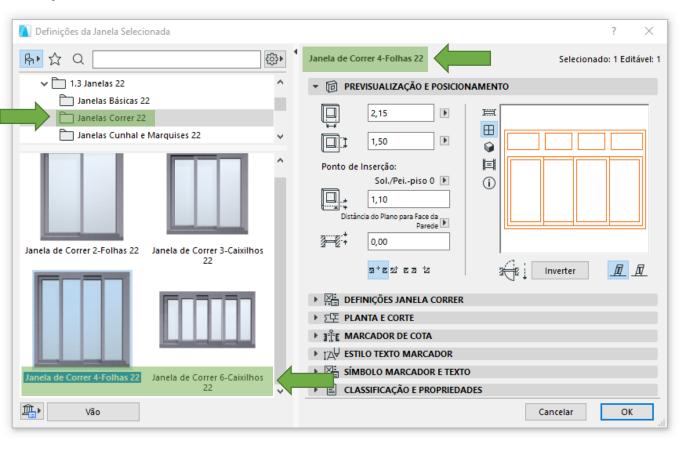
NOMINAL

<u>Categorias</u>, sendo que cada categoria é <u>independente, sem relação com as outras.</u>

ORDINAL

<u>Categorias</u>, sendo que cada categoria <u>mantém</u> <u>uma relação de ordem com as outras.</u>

Ilustração: ArchiCAD 22



P 1 10

QUALITATIVA

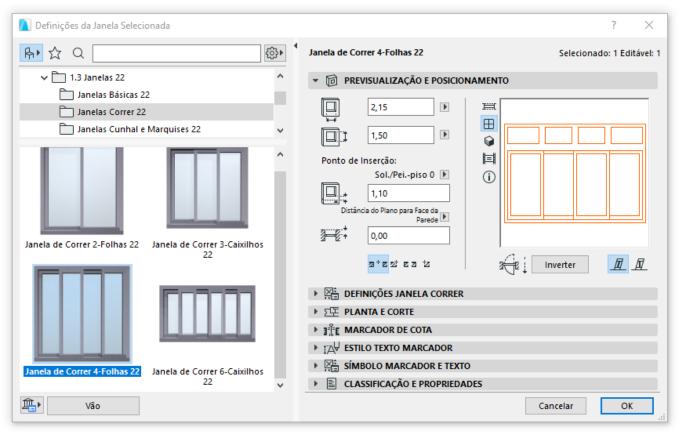
NOMINAL

Categorias, sendo que cada categoria é independente, sem relação com as outras.

ORDINAL

Categorias, sendo que cada categoria mantém uma relação de ordem com as outras.

Ilustração: ArchiCAD 22



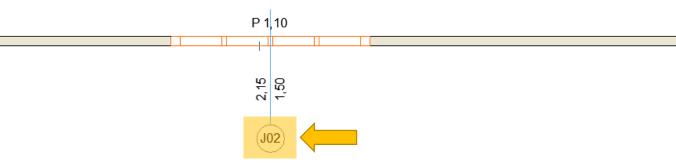
QUALITATIVA

NOMINAL

Categorias, sendo que cada categoria é independente, sem relação com as outras.

ORDINAL

Categorias, sendo que cada categoria mantém uma relação de ordem com as outras.



QUANTITATIVA

DISCRETA

Números inteiros, sem frações, como em contagens. Constituem um conjunto finito.

CONTÍNUA

Números que podem assumir valores fracionários. Normalmente têm intervalos de valores conhecido, mas um conjunto infinito de valores possíveis.

QUALITATIVA

NOMINAL

Categorias, sendo que cada categoria é independente, sem relação com as outras.

ORDINAL

Categorias, sendo que cada categoria mantém uma relação de ordem com as outras.

QUANTITATIVA

QUALITATIVA

DISCRETA

Números inteiros, sem frações, como em contagens. Constituem um conjunto finito.

CONTÍNUA

Números que **podem assumir valores fracionários.** Normalmente têm intervalos de valores conhecido, mas um conjunto infinito de valores possíveis.

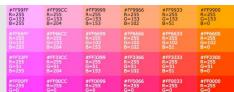
NOMINAL

<u>Categorias</u>, sendo que cada categoria é <u>independente</u>, sem relação com as outras.

ORDINAL

Categorias, sendo que cada categoria mantém uma relação de ordem com as outras.

Por exemplo, um evento qualitativo que comporta medida quantitativa contínua é a cor. Embora seja frequentemente medida por variável categórica nominal (rosa, roxo, laranja ou vermelho), ela comporta mensurações contínuas para matiz e saturação.



Flash! 08 | Caracterização de Dados | Tipos de Escala |

Nominal Ordinal

Intervalar Proporcional

A escala nominal mede atributos que só conhecem relações de equivalência (=, ≠) e, dessa forma, quaisquer que sejam seus códigos numéricos, as análises a ser consideradas não devem exceder esse nível. [...] Quaisquer números podem ser usados para a codificação dos atributos, respeitada a relação biunívoca entre atributos e códigos numéricos (um mesmo atributo não pode ter dois códigos, nem um código pode corresponder a dois atributos distintos).

reflexiva [$\forall x (x = x)$] simétrica (se x = y, então y = x) transitiva (se x = y e y = z, então x = z) **Ordinal**

Proporcional

A **escala nominal** mede atributos que só conhecem relações de equivalência (=, ≠) e, dessa forma, quaisquer que sejam seus códigos numéricos, as análises a ser consideradas não devem exceder esse nível. [...] Quaisquer números podem ser usados para a codificação dos atributos, respeitada a relação biunívoca entre atributos e códigos numéricos (um mesmo atributo não pode ter dois códigos, nem um código pode corresponder a dois atributos distintos).

reflexiva [$\forall x (x = x)$] simétrica (se x = y, então y = x) transitiva (se x = y e y = z, então x = z)

Ordinal

A escala ordinal, por sua vez, mede atributos que se distinguem em grau ou intensidade, de forma que, além das relações de igualdade/desigualdade, podem-se reconhecer relações de ordem (>, <). Tem sentido de direção definido, mas as unidades de mensuração são desconsideradas. Tampouco tem um valor nulo definido; um eventual código zero não terá valor de nulidade, mas de posição entre eventuais códigos negativos e positivos. [...]

irreflexiva [$\forall x (x > x)$ não é verdade] assimétrica [(se x > y, então y > x) não é verdade transitiva (se x > y e y > z, então x > z).

A **escala nominal** mede atributos que só conhecem relações de equivalência (=, ≠) e, dessa forma, quaisquer que sejam seus códigos numéricos, as análises a ser consideradas não devem exceder esse nível. [...] Quaisquer números podem ser usados para a codificação dos atributos, respeitada a relação biunívoca entre atributos e códigos numéricos (um mesmo atributo não pode ter dois códigos, nem um código pode corresponder a dois atributos distintos).

reflexiva [
$$\forall x (x = x)$$
]
simétrica (se x = y, então y = x)
transitiva (se x = y e y = z, então x = z)

Ordinal

A **escala ordinal**, por sua vez, mede atributos que se distinguem em grau ou intensidade, de forma que, além das relações de igualdade/desigualdade, podem-se reconhecer relações de ordem (>, <). Tem sentido de direção definido, mas as unidades de mensuração são desconsideradas. Tampouco tem um valor nulo definido; um eventual código zero não terá valor de nulidade, mas de posição entre eventuais códigos negativos e positivos. [...]

irreflexiva [$\forall x (x > x)$ não é verdade] assimétrica [(se x > y, então y > x) não é verdade] transitiva (se x > y e y > z, então x > z).

Ordinal

Intervalar

Proporcional

A escala intervalar, mede atributos de forma a que os intervalos desta representem quantidades regulares de atributo; a escala é uma função linear de atributos. Além da relação biunívoca entre atributos e códigos numéricos da escala e do sentido de orientação da medida, tem-se a definição da unidade de mensuração. A escala intervalar tem um zero, mas ele é um ponto arbitrado para a origem das unidades de medida e não tem correspondência com a situação zero de atributo. [...] As propriedades da escala intervalar incluem todas as propriedades de aritmética, que terão sentido se aplicadas a diferenças entre os valores da escala (intervalos), mas não aos valores propriamente ditos.

Intervalar

A escala intervalar, mede atributos de forma a que os intervalos desta representem quantidades regulares de atributo; a escala é uma função linear de atributos. Além da relação biunívoca entre atributos e códigos numéricos da escala e do sentido de orientação da medida, tem-se a definição da unidade de mensuração. A escala intervalar tem um zero, mas ele é um ponto arbitrado para a origem das unidades de medida e não tem correspondência com a situação zero de atributo. [...] As propriedades da escala intervalar incluem todas as propriedades de aritmética, que terão sentido se aplicadas a diferenças entre os valores da escala (intervalos), mas não aos valores propriamente ditos.

Ordinal

Proporcional (Racional)

A escala proporcional (racional) mede atributos de forma que os acréscimos em atributos sejam representados por acréscimos proporcionais em valores da escala. A razão entre dois valores da escala corresponde à razão entre dois valores dos atributos, isto é, se um atributo é o dobro do outro, seu valor na escala é igualmente o dobro do outro. A escala é uma função linear do atributos e a origem é comum, ou seja, **há um zero real.** [...] As propriedades da escala proporcional (racional) incluem todas as **propriedades** da aritmética, sem restrição. O pesquisador está, portanto, assumindo que seu evento tenha comportamento aritmético, que é a premissa mais forte que ele pode adotar para a escolha de uma escala.

Flash! 08 | Caracterização de Dados | Tipos de Escala |

Ex. Escala Intervalar

Atributo de temperatura	Escala em Célsius	Intervalo entre as medidas	Escala em Fahrenheit	Intervalo entre as medidas	Razão entre intervalos	Razão entre as medidas
А	0 °C	B - A = 5	32 °F	B - A = 9	°F/°C = 9/5	32°F / 0°C = !
В	5 °C	C - B = 5	41 °F	C - B = 9	°F/°C = 9/5	41°F / °5C = 8,2
С	10 °C	D - C = 5	50 °F	D - C = 9	°F/°C = 9/5	50°F / 10°C = 5,0
D	15 °C	E - D = 5	59 °F	E - D = 9	°F/°C = 9/5	59°F / 15°C = 3,9
E	20 °C		68 °F			68°F / 20°C = 3,4

Ex. Escala Proporcional

Atributo de peso	Escala em Quilogramas	Intervalo entre as medidas	Escala em Libras	Intervalo entre as medidas	Razão entre intervalos	Razão entre as medidas
А	0 Kg	B - A = 2	0,0 Lb	B - A = 4,4	Lb/Kg = 2,2	0,0Lb/0Kg = !
В	2 Kg	C - B = 2	4,4 Lb	C - B = 4,4	Lb/Kg = 2,2	4,4Lb/2Kg = 2,2
С	4 Kg	D - C = 2	8,8 Lb	D - C = 4,4	Lb/Kg = 2,2	8,8Lb/4Kg = 2,2
D	6 Kg	E - D = 2	13,2 Lb	E - D = 4,4	Lb/Kg = 2,2	13,2Lb/6Kg = 2,2
E	8 Kg		17,6 Lb		Lb/Kg = 2,2	17,6Lb/8Kg = 2,2

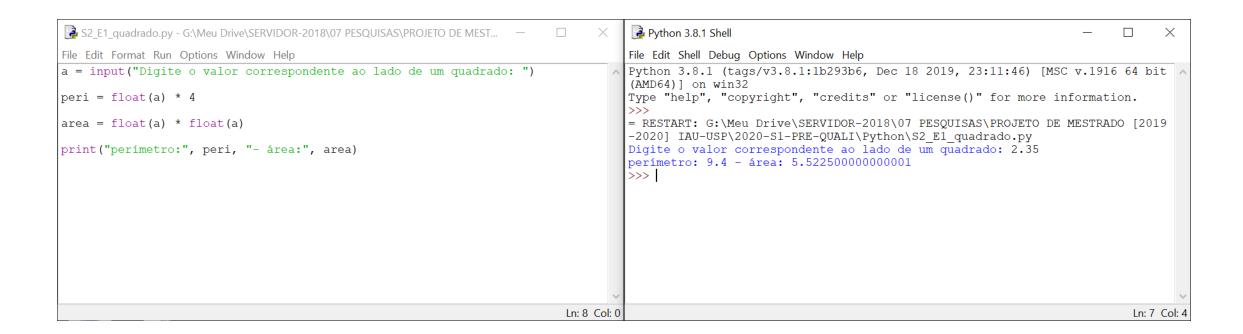
Ex. Escala Intervalar

Atributo de temperatura	Escala em Célsius	Intervalo entre as medidas	Escala em Fahrenheit	Intervalo entre as medidas	Razão entre intervalos	Razão entre as medidas
А	0 °C	B - A = 5	32 °F	B - A = 9	°F/°C = 9/5	32°F / 0°C = !
В	5 °C	C - B = 5	41 °F	C - B = 9	°F/°C = 9/5	41°F / °5C = 8,2
С	10 °C	D - C = 5	50 °F	D - C = 9	°F/°C = 9/5	50°F / 10°C = 5,0
D	15 °C	E - D = 5	59 °F	E - D = 9	°F/°C = 9/5	59°F / 15°C = 3,9
E	20 °C		68 °F			68°F / 20°C = 3,4

Ex. Escala Proporcional

Atributo de peso	Escala em Quilogramas	Intervalo entre as medidas	Escala em Libras	Intervalo entre as medidas	Razão entre intervalos	Razão entre as medidas
А	0 Kg	B - A = 2	0,0 Lb	B - A = 4,4	Lb/Kg = 2,2	0,0Lb/0Kg = !
В	2 Kg	C - B = 2	4,4 Lb	C - B = 4,4	Lb/Kg = 2,2	4,4Lb/2Kg = 2,2
С	4 Kg	D - C = 2	8,8 Lb	D - C = 4,4	Lb/Kg = 2,2	8,8Lb/4Kg = 2,2
D	6 Kg	E - D = 2	13,2 Lb	E - D = 4,4	Lb/Kg = 2,2	13,2Lb/6Kg = 2,2
E	8 Kg		17,6 Lb		Lb/Kg = 2,2	17,6Lb/8Kg = 2,2

= RESTART: C:\Exemplo\ Programação em Python.py



A assunção de premissas (postulados, proposições, axiomas) é inerente à estrutura do conhecimento científico e condiciona o entendimento de objetividade e verdade, ou seja, de representação do real. Repensar essas premissas deve <u>auxiliar o pesquisador a conceber suas</u> estratégias de investigação.

A distinção entre ser e atributo, de que se ocupa a filosofia, constitui-se em elemento importante para que o pesquisador defina a delimitação de seu objeto de estudo.

(PEREIRA, 2004, p. 40)

Dessa forma, <u>antes de aplicarmos algoritmos de AM a um conjunto de dados (seja ele qual for)</u> <u>é importante que esses dados sejam analisados e pré-processados.</u> Essa análise e pré-processamento, pode ser realizada por diversas técnicas, e <u>permite uma melhor compreensão</u> da distribuição dos dados e pode dar suporte à escolha de formas de abordagem do problema.

(FACELI et al., 2019, p. 28)

REFERÊNCIAS BIBLIOGRÁFICAS

EASTMAN, Chuck; TELCHOLZ, Paul; SACK, Rafael; LISTON, Hathleen. Manual de BIM: um guia de modelagem da informação da construção para arquitetos, engenheiros, gerentes, construtores e incorporadores. [tradução: Cervantes Gonçalves Ayres Filho... et al.]; Revisão técnica: Eduardo Toledo Santos. - Dados eletrônicos. Porto Alegre: Bookman, 2014.

FACELI, Katti; LORENA, Ana Carolina; GAMA, João; CARVALHO, André Carlos Ponce de Leon Ferreira de. **Inteligência artificial: uma abordagem de aprendizado de máquina.** Reimpr. – Rio de Janeiro: LTC, 2019.

PEREIRA, Júlio Cesar Rodrigues. **Análise de dados qualitativos: estratégias metodológicas para as Ciências da Saúde, Humanas e Sociais.** 3 ed. 1. Reimpr. – São Paulo: Editora da Universidade de São Paulo (EDUSP), 2004.