

Instituto de Computação - UNICAMP - Universidade Estadual de Campinas

MC536 - Banco de Dados, Teoria e Prática - 2º Semestre de 2018

Professor: André Santachè

Nome: Thiago Dong Chen                      RA: 187560                      Turma A

Nome: Thiago Lima Costa                      RA: 187592                      Turma C

Grupo: Thiago<sup>2</sup>

### Resumo:

A partir de uma base de dados coletada - um conjunto de atributos biomecânicos de pacientes - com dados sobre a forma e orientação da pelve e da coluna lombar, tem-se como objetivo a classificação de um problema ortopédico de um paciente entre três classes: Normal, Espondilolistese e Hérnia de Disco. Para isso, o trabalho utilizará *Machine Learning* para prever se os novos pacientes possuem problema ortopédico mediante suas características.

### Requisitos:

- Consultar, incluir, excluir e alterar pacientes.
- Consultar, incluir, excluir e alterar características de pacientes (relativos ao seu problema, como Incidência Pélvica, Versão Pélvica, Ângulo de Lordose Lombar, etc).
- Consultar, incluir, excluir e alterar classificações para os pacientes (como Normal, Hérnia de Disco e Espondilolistese).
- Consultar, incluir, excluir e alterar sintomas comuns de classificações (como Fraqueza Muscular, Dormência na Perna, etc).
- Permitir emissão de relatórios estatísticos (por meio de consultas) com os dados dos pacientes, suas características e classificações.

Link para o projeto no GitHub: <[https://github.com/thiagolim4/bd\\_thiago2](https://github.com/thiagolim4/bd_thiago2)>

### Fonte:

<<https://www.kaggle.com/faizunnabi/orthopedic-symptoms-classification/data>>

Site: Kaggle

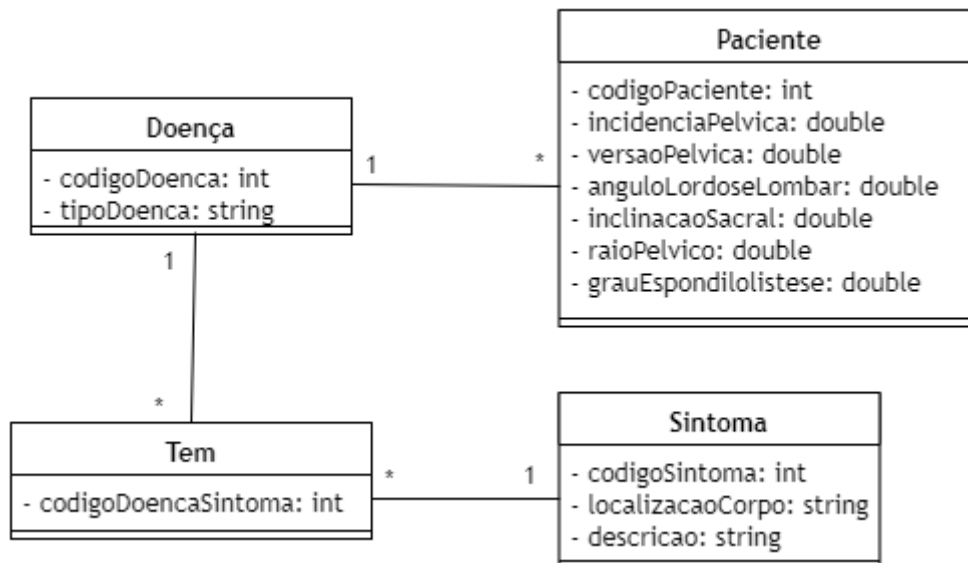
Acesso em 20/08/2018

Kaggle é uma comunidade de cientistas de dados e conteúdo para *Machine Learning*. A Kaggle começou oferecendo competições de *Machine Learning* e

também oferece uma plataforma de dados públicos, um banco em nuvem para ciência de dados e educação de inteligência artificial de forma compacta.

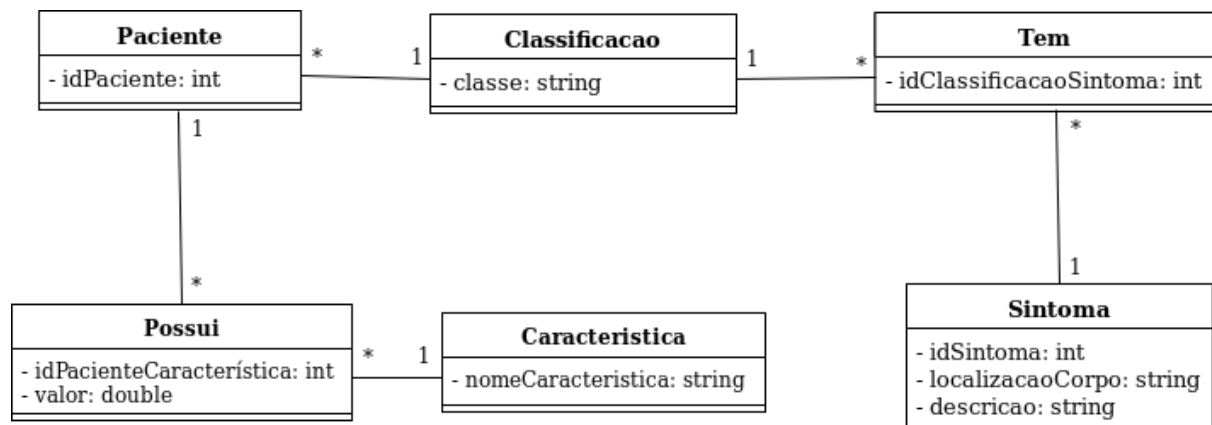
### Modelagem Conceitual (etapa 01):

Diagrama UML



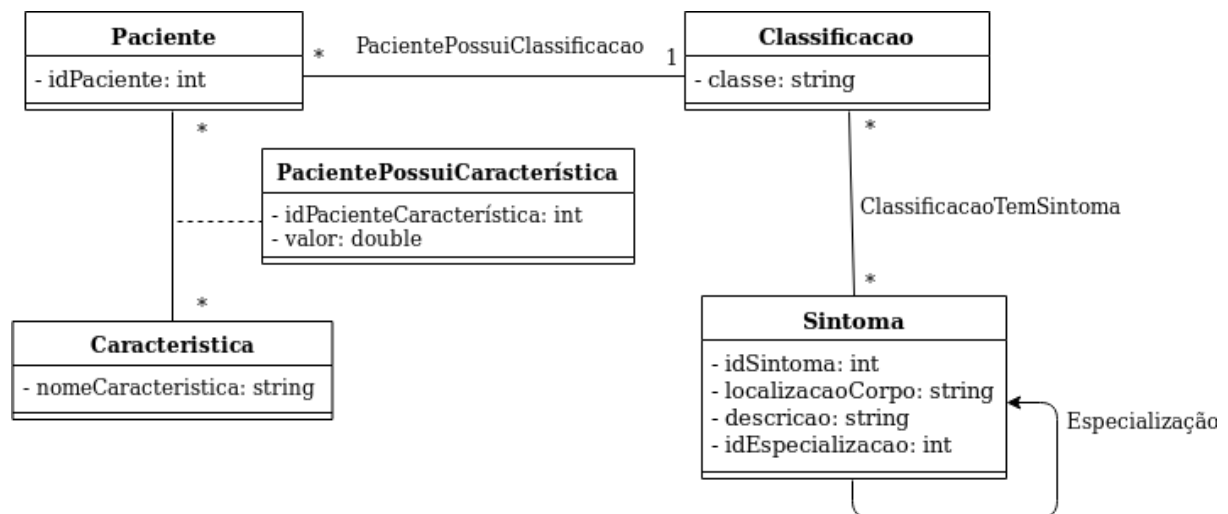
### Modelagem Conceitual (etapa 02):

Diagrama UML



### Modelagem Conceitual (etapa 03):

Diagrama UML



## Modelo Lógico

Paciente(idPaciente, classe)

CHE: classe -> classe (Classificacao)

Sintoma(idSintoma, localizacaoCorpo, descricao, idEspecializacao)

CHE: idEspecializacao -> Sintoma (idSintoma)

Classificacao(classe)

ClassificacaoTemSintoma(idClassificacaoTemSintoma, classe, idSintoma)

CHE: idSintoma -> idSintoma (Sintoma)

CHE: classe -> classe (Classificacao)

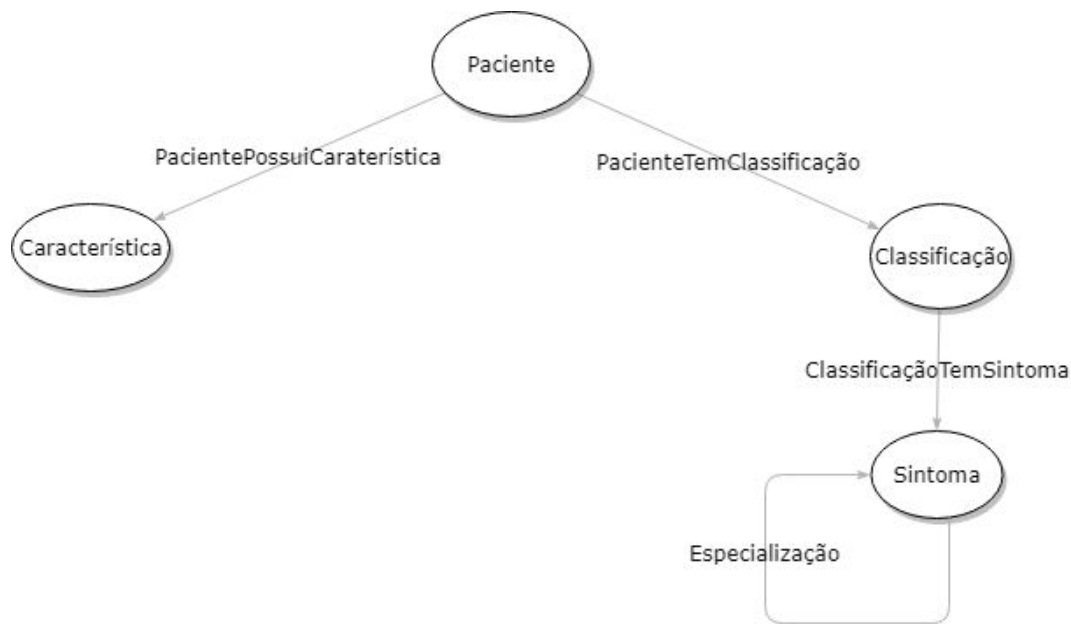
Caracteristica(nomeCaracteristica)

PacientePossuiCaracteristica(idPacienteCaracteristica, idPaciente, nomeCaracteristica, valor)

CHE: idPaciente -> idPaciente (Paciente)

CHE: nomeCaracteristica -> nomeCaracteristica (Caracteristica)

## Modelo em Grafo



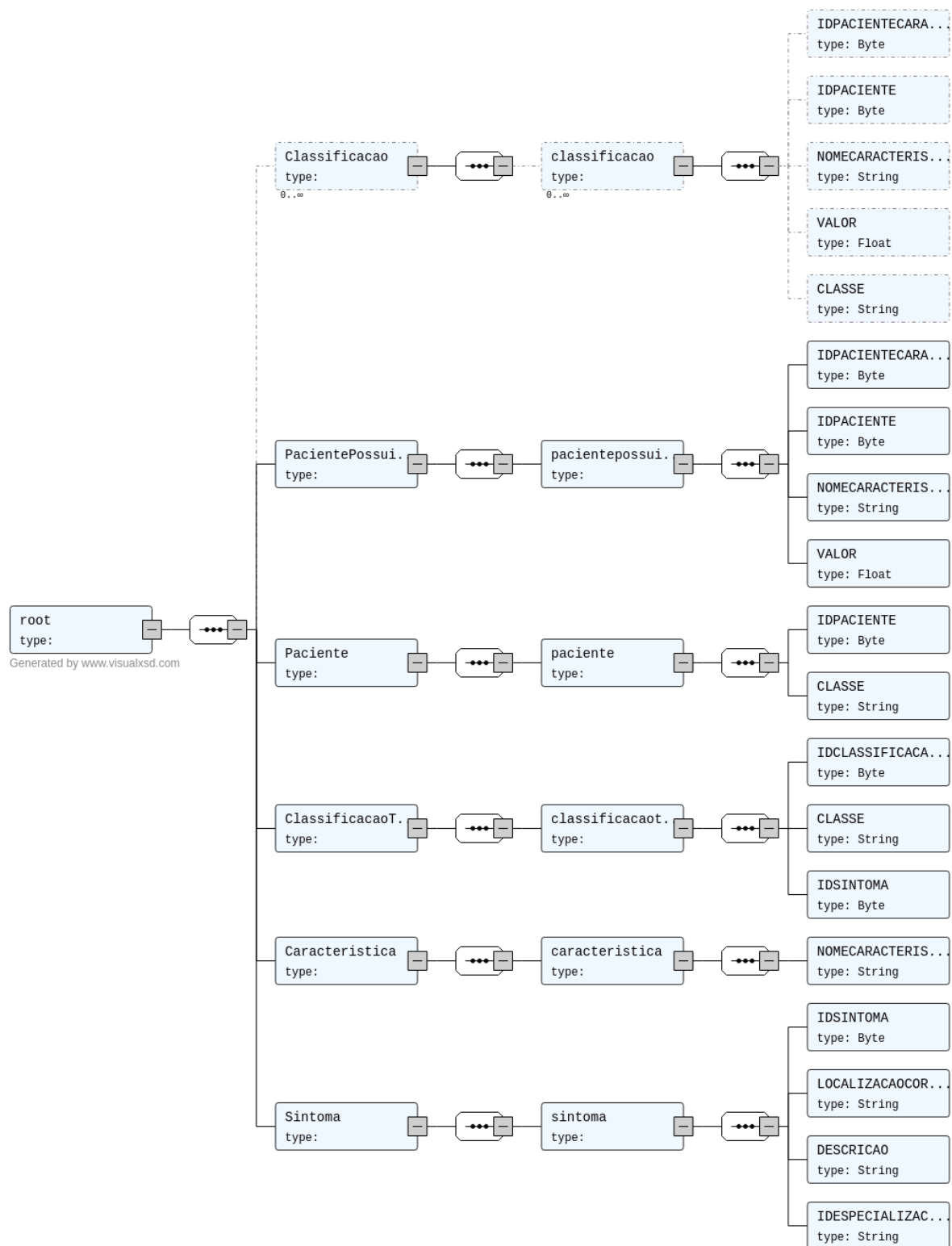
As mudanças ocorridas da Modelagem Conceitual da Etapa 01 para a Etapa 02 foram:

- Modificação nos atributos da classe Paciente: agora ela possui somente o identificador do paciente chamado idPaciente, os atributos referentes à situação do mesmo (como Incidência Pélvica, Versão Pélvica, etc) foram transferidos para outras classes a fim de que se pudesse, caso surgisse uma nova característica (novo tipo de medição da Pelve, por exemplo) seria possível adicioná-la ao banco;
- Seguindo a modificação do Paciente, foram adicionadas duas novas classes: Possui e Característica. Cada Paciente possui várias características (N:N), sendo que o valor do relacionamento entre o Paciente e sua Característica fica na classe intermediária Possui - por exemplo, João (Paciente) tem uma Incidência Pélvica (Característica) de valor 60° (Possui) e Maria (Paciente) tem uma Incidência Pélvica (Característica) de 70° (Possui);
- Mudança no nome da classe Doença para Classificação e tipoDoença para classe, pois havia certos pacientes classificados como “Normal”, o que não caracterizava um problema;
- Adição do atributo idEspecialização na classe Sintoma, pois agora os sintomas participarão de uma hierarquia (os sintomas poderão se

especializar em partes do corpo, por exemplo, região da perna e subsintoma na região da coxa);

- A modelagem para grafos do modelo ficou: nós da classe Paciente apontando para nós Característica e Classificação, nós Classificação apontando para nós Sintoma e, por fim, nós Sintomas apontando para si mesmos (em uma hierarquia). Seguindo as modificações em Sintoma, o modelo de grafos é adequado no sentido de que haverão Sintomas e diversos Subsintomas, e assim, diferentemente de sempre haver um conjunto fixo das outras arestas, não há a mesma certeza para a hierarquia de Sintomas, podendo seguir um caminho indefinido.

## Modelo Lógico etapa 4 para XML



Na etapa 04 do trabalho, não houve mudança do modelo conceitual, já que houve apenas uma mudança na modelagem lógica do problema, pois as queries utilizadas foram de web semântica como o uso de SPARQL e XQUERY.

### **Vantagens da web-semântica**

Uma das vantagens da web-semântica é o fato de não necessitar de uma estrutura fixa como o modelo relacional. Isso permite a flexibilização na hora da modelagem. Dependendo do problema, o modelo hierárquico permite uma modelagem mais simples do que o modelo convencional. E a maior vantagem em relação ao modelo relacional, o XML por exemplo permite que a máquina interprete os dados de maneira mais simplificada, fornecendo suporte ao programador.