

Nomes:

Giovanna Vendramini
Vitoria Dias Moreira Pinho

RA:

173304
188511

Trabalho 1 - Etapa 2

Chronos Parte 1

I. Introdução

A tireóide é a glândula localizada na parte inferior do pescoço, responsável pela produção de hormônios que controlam o metabolismo das células, garantindo assim, o funcionamento correto do corpo. Seus principais hormônios são o T3 e o T4, sendo que é controlada por hormônios liberados por outras glândulas. Quando os níveis de T3 e T4 estão baixos, o hipotálamo libera o TRH, provocando a também liberação do TSH pela glândula pituitária, que estimula a tireóide. Quando há um desequilíbrio nessa função, ocorrem os distúrbios da tireóide.

II. Resumo

Devido a alta incidência de pessoas com distúrbios da tireóide, o estudo sobre o diagnóstico e a prevenção do problema tornam-se cada vez mais necessários.

Dessa forma, o trabalho terá como intuito realizar o diagnóstico de pacientes com base em alguns valores de seus exames, como TSH, T3, T4 total e livre, além de prever a possibilidade do indivíduo apresentar algum distúrbio na tireóide de acordo com suas informações gerais como idade, sexo, gravidez, se apresenta bócio, entre outros.

III. Requisitos

A. Funcionais

- Diagnosticar se o paciente apresenta hipotireoidismo e hipertireoidismo com base em exames clínicos;
- Prever, de acordo com informações gerais do paciente como idade e sexo, se ele possui tendência a apresentar distúrbios da tireóide;
- Caso seja detectado desequilíbrio nos hormônios da tireóide, diagnosticar em qual categoria este pertence, dentre: hipotireoidismo, hipotireoidismo primário, hipotireoidismo compensado, hipotireoidismo secundário, hipertireoidismo, T3 tóxico, bócio, toxicidade secundária.

B. Não funcionais

- Deve apresentar um alto grau de confiabilidade, sendo capaz de acertar a maioria dos diagnósticos.

IV. Base de Dados

UCI Machine Learning Repositor

(<http://archive.ics.uci.edu/ml/datasets/thyroid+disease>) é um site que disponibiliza diversas Bases de Dados para uso em machine learning.

As bases fornecidas para distúrbios de tireóide são do Instituto Garvan, na Austrália. Foram selecionadas três bases:

1. new-thyroid:
 - a. Apresenta três casos: indivíduos normais, que possuem hipotireoidismo ou hipertireoidismo.
 - b. 6 Atributos;
 - c. 215 Instâncias;
 - d. Atributos específicos: T3, T4, TSH; ...
2. allhyper:
 - a. Apresenta se o indivíduo possui algum tipo de hipertireoidismo;
 - b. 30 atributos;
 - c. 2800 instâncias;
 - d. Atributos gerais: idade, sexo, gravidez, se realizou cirurgia da tireóide; ...
 - e. Atributos específicos: TSH, T3, T4; ...
3. allhypo:
 - a. Apresenta se o indivíduo possui algum tipo de hipotireoidismo;
 - b. 30 atributos;
 - c. 2800 instâncias;
 - d. Atributos gerais: idade, sexo, gravidez, se realizou cirurgia da tireóide, ...
 - e. Atributos específicos: TSH, T3, T4, ...

V. Ferramenta de Análise

A ferramenta de análise escolhida foi o Weka, de forma que o Weka permitirá a predição da doença a partir das bases de dados.

VI. Desenvolvimento do Modelo Conceitual

A) 1ª Versão

A primeira versão do modelo conceitual foi criada levando em consideração apenas a organização dos dados em tabelas. Dessa forma, havia três tipos de pacientes, referentes às três tabelas e aos exames que eles realizavam em cada uma delas. Por isso, não representava de forma correta um modelo conceitual.

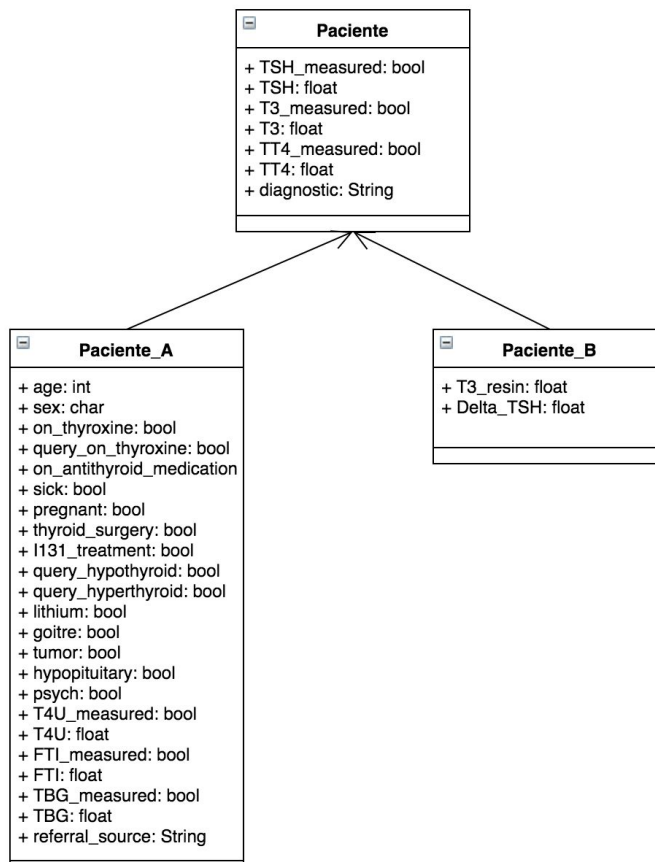


Fig.1: Primeira versão do modelo conceitual em UML.

B) 2ª Versão

Na segunda versão, o modelo já tomou uma nova forma e se aproximou melhor de um modelo conceitual. No entanto, ainda havia pontos a serem melhorados, como a classe doença que tinha dois herdeiros: hipertireoidismo e hipotireoidismo, e a classe exame que apresentava muitos atributos. As classes herdeiras hipertireoidismo e hipotireoidismo faziam com que essas instâncias de doença se tornassem classes, o que não estava correto. Já o exame se tornava algo muito complexo e de difícil atualização, com itens muito específicos.

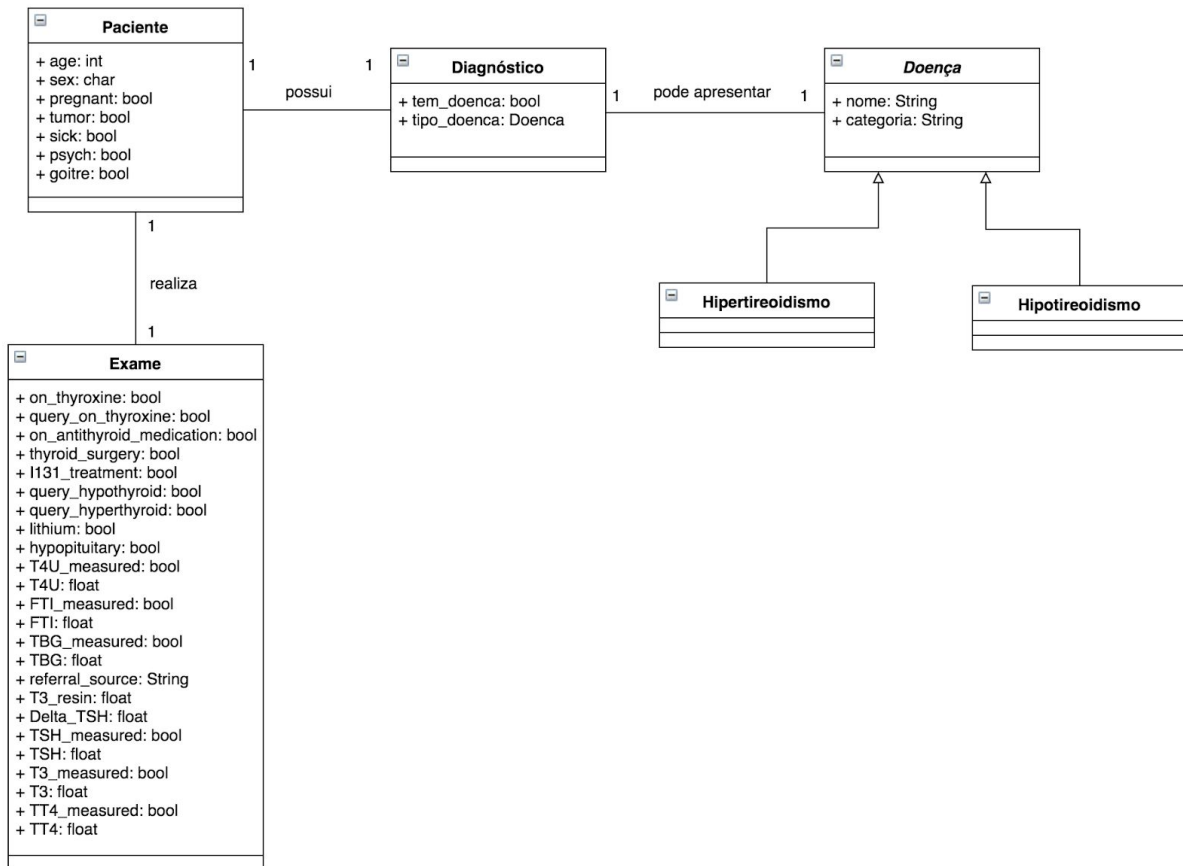


Fig.2: Segunda versão do modelo conceitual em UML.

C) 3ª Versão

A versão final corrige os problemas da versão dois, eliminando as classes herdeiras de doença e criando uma nova classe item, que instancia cada um dos itens do exame separadamente. Nesse novo modelo, exame se torna uma agregação de itens de exame.

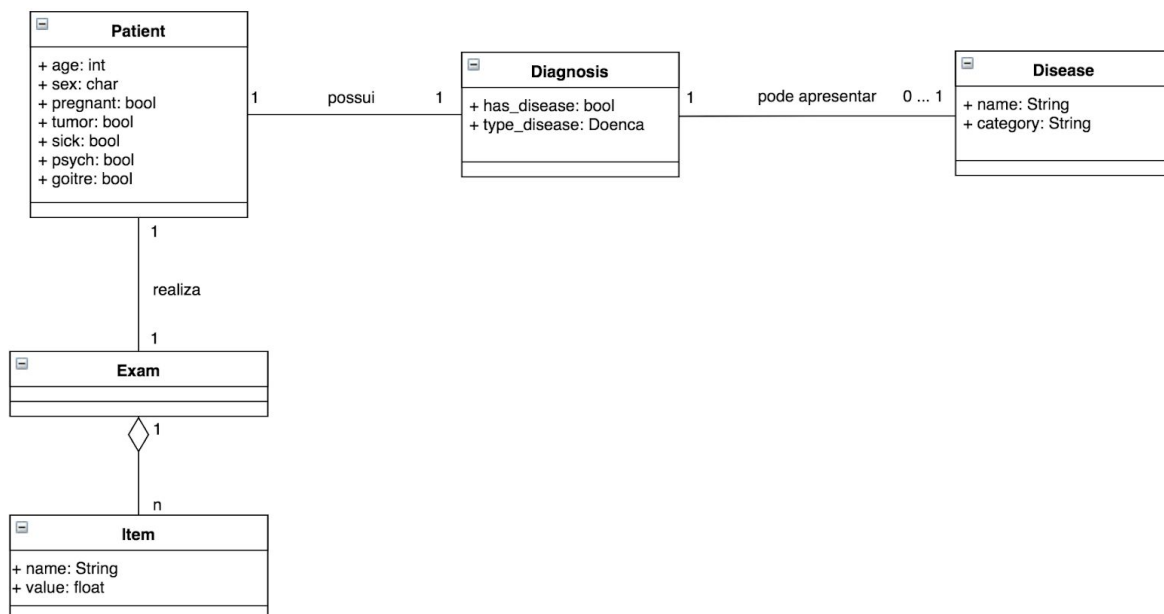


Fig.3: Diagrama final do modelo conceitual em UML.

VII. Modelo Lógico

O modelo relacional pode ser organizado nas relações abaixo.

DISEASE (id_disease, name, category)

ITEM (id_item, id_patient, name, result, value)

CHE: id_patient para patient(id_patient)

PATIENT (id_patient, id_disease, age, sex, pregnant, tumor, sick, psych, goitre)

CHE: id_disease para disease(id_disease)

É possível dizer que o modelo relacional se mostrou bem condizente com o modelo conceitual. As modificações adotadas foram a adição de chaves primárias e chaves estrangeiras para o relacionamento entre as tabelas e exclusão das tabelas Diagnosis e Exam. A justificativa para a exclusão da primeira se dá porque esta dificultaria o processo de consulta direta da doença de determinado paciente, apenas para apresentar a informação de um diagnóstico negativo. Ou seja, a motivação para sua construção era indicar os casos em que o paciente não foi diagnosticado com nenhum distúrbio da tireóide, problema que foi resolvido ao atribuir ao paciente diretamente o id_disease (da tabela Disease), e permitindo que tal campo seja nulo (caso em que has_disease seria negativo). Já a exclusão da tabela Exam apresenta uma justificativa análoga, do ponto de vista de modelo conceitual a atribuição de itens de exame ao paciente dificulta o entendimento e organização do problema. Entretanto, do ponto de vista lógico atribuir o item de exame diretamente ao identificador do paciente torna direto o acesso à informação,

facilitando a geração de consultas e sem perda de informação (já que não havia nenhuma informação referente apenas ao exame).