**LEONARDO BALBI DA COSTA**

**TRIAGEM MÉDICA OCUPACIONAL: Sistema de análise de risco à saúde de colaboradores.**

Aracruz

11/2021

**Resumo**

Uma das grandes preocupações dos profissionais da medicina do trabalho é manter um controle de acompanhamento da força de trabalho, visando minimizar os impactos financeiros e laborais causados pelo absenteísmo. Desta forma, ainda recorrem à algumas ferramentas mais tradicionais, para que o histórico de saúde de cada trabalhador seja mantido, mas ainda baseado em um princípio de causa e efeito, o que quase nunca aponta um possível problema futuro. Existem ainda casos de erros de diagnóstico em decorrência da relação médico-paciente, da deficiência ou falha na realização dos exames ou erro na avaliação dos resultados. O uso da inteligência artificial presente em diversos ramos de atividades como comércio eletrônico, sistemas bancários, mídias sociais, dentre outras, tem reforçado o aumento da aplicação desta no ramo da saúde. Sendo assim a análise dos dados busca apontar para a evolução de doenças causadas pelo sedentarismo e/ou maus hábitos praticados, possibilitando um diagnóstico antecipado e dando a oportunidade de um tratamento precoce, bem como a adoção de medidas preventivas nos casos menos críticos. O presente trabalho visa apresentar o desenvolvimento de um software para avaliação de dados e classificação do estado de saúde dos indivíduos, além de recomendar práticas de exercícios e alimentação saudável a depender da classificação obtida.

**Palavras-chave**: absenteísmo; inteligência artificial; sistema especialista.

**1. INTRODUÇÃO**

Um dos fatores primordiais às grandes corporações, no que tange a sobrevivência ao mercado extremamente competitivo, é a manutenção e a retenção de uma mão de obra qualificada, para que se possa manter o cumprimento de prazos e metas a serem atingidas, prezando a qualidade de serviço ofertado. Contudo, é preciso que as organizações passem a avaliar mais de perto, não somente a qualidade de vida no ambiente de trabalho, mas também o perfil clínico de sua força de trabalho.

Sendo assim, a ausência do trabalhador por razões médicas têm sido um agravante presente no ambiente corporativo, mesmo antes da primeira revolução industrial, sendo necessário uma melhor gestão da identificação das causas e avaliação dos custos a este atribuído, visto que a ausência do trabalhador em suas atividades laborais implica em perda de produtividade, sobrecarregamento da força de trabalho presente, aumento de custos com horas extras dentre outros fatores, que consequentemente, impactarão no resultado final da corporação frente à seus clientes.

Portanto, cada vez mais engajadas em estudos de dados, as grandes corporações tendem a investir em alternativas computacionais para mitigar ou até eliminar, problemas relativos à ausência total ou parcial no período de trabalho. Com o uso da Inteligência Artificial torna-se possível uma avaliação antecipada, através de dados clínicos da força de trabalho, dos riscos à saúde apresentados por cada colaborador, e que seja tomada a ação para que tal indivíduo adote medidas preventivas, visando a garantia de sua saúde, de acordo com padrões estabelecidos pela organização mundial da saúde (OMS) para a execução de suas funções laborais.

O objetivo geral desta pesquisa é desenvolver uma ferramenta de software desktop, para auxiliar equipes de saúde ocupacional na identificação antecipada de indivíduos que estejam a desenvolver e/ou agravar problemas de saúde.

Portanto, os seguintes objetivos específicos se fazem necessários.

Na etapa inicial, será feita uma análise temática e do problema em relação às ferramentas similares já existentes no mercado, com o objetivo de identificar as variáveis que são utilizadas por estas, bem como identificar variáveis ainda não utilizadas para melhor performance da ferramenta a ser desenvolvida, estabelecendo uma estrutura básica do sistema.

Na etapa intermediária, será feita a identificação dos stakeholders, com os quais serão feitas as entrevistas, com o objetivo de identificar as necessidades destes e obter uma validação prévia das variáveis identificados para o desenvolvimento. Em continuidade será feito o levantamento dos requisitos que embasarão a modelagem do sistema, por meio dos diagramas de caso de uso, de sequência e de classes.

Na etapa final, será feita a escrita do código juntamente com o desenvolvimento das telas do sistema. Em paralelo à esta etapa, será executada a modelagem do banco de dados, para proporcionar uma etapa inicial de testes simulados do funcionamento das funcionalidades (software) e transações do sistema (software, hardware).

* 1. **JUSTIFICATIVA**

Este estudo visa orientar os indivíduos a terem hábitos saudáveis mesmo que dentro do ambiente corporativo, uma vez que a utilização de uma mão de obra mais qualificada, atrelada ao uso da tecnologia, proporciona um ambiente menos exaustivo, porém relativamente propício ao sedentarismo. Somado a isso, as corporações podem adotar políticas de hábitos saudáveis por meio de programas de saúde, visando garantir a redução de ausência de trabalhadoresno ambiente corporativo.

Como profissionais da área de saúde e em atendimento a requisitos legais o médico e o enfermeiro do trabalho desempenham funções importantes no controle de saúde de funcionários de empresas, de qualquer que seja o ramo, cumprindo o que está descrito no PCMSO. Estes profissionais acompanham os pacientes, de forma a orientá-los às melhores práticas de saúde. Contudo, na maioria das vezes o fazem baseados em resultados, não agrupados ou pontualmente isolados, o que pode incorrer em erro de diagnóstico.

Através de um sistema de software é possível correlacionar os dados obtidos e cruzá-los para que sejam definidos os melhores tipos de tratamento a cada funcionário, bem como a vantagem de um diagnóstico precoce de problemas de saúde.

**2. REFERENCIAL TEÓRICO**

Este capítulo discute aspectos importantes relativos em se tratando de triagem médica ocupacional e o uso da tecnologia no auxílio à área de saúde. Estando este dividido em duas partes: Na seção 2.1 é apresentada a definição de absenteísmo e explanação deste como custo oculto das corporações; Na seção 2.2 é feita a apresentação e explanação sobre inteligência artificial e as vantagens do uso desta na área da saúde.

2.1 - ABSENTEÍSMO

Segundo Souza (2007), absenteísmo é uma palavra de origem francesa, *absentéisme* que tem como significado a ausência do trabalhador ao seu ambiente de trabalho por inúmeras razões. Nota-se que o autor explana que o absenteísmo está ligado a todas as formas e/ou fatores que levam os indivíduos a se ausentarem na empresa não distinguindo estes entre voluntários, ou seja, os programados como férias e folgas e involuntários que se caracterizam pela falta não programada, sendo estes na maioria das vezes por problemas de saúde e/ou pessoais.

Segundo Chiavenato (2014, p. 81),

“Absenteísmo ou ausentismo é a frequência e/ou duração do tempo de trabalho perdido, quando os colaboradores não comparecem ao trabalho. O absenteísmo constitui a soma dos períodos em que os colaboradores se encontram ausentes do trabalho, seja por falta, atraso ou algum motivo interveniente.”

Diversos autores evidenciam que o absenteísmo onera às corporações em uma grande quantidade de perda financeira, uma vez que lidam com o problema de forma diária. Associados a isto, a quantificação dos custos com os benefícios por doença afeta diretamente o produto nacional bruto (PNB), podendo este ser elevado caso a taxa de absenteísmo por empregado/por ano diminuísse por apenas um dia, sendo assim a sua consequência, a preocupação de empresários uma vez que afetam diretamente no balanço e faturamento das corporações.

Segundo Dall´inha (2006), de forma geral o absenteísmo exprime a inadequação do empregado ao ambiente laboral, na maioria das vezes com apresentação do atestado médico, como forma legar de garantir a redução dos rendimentos mensais e/ou gratificações devido à sua ausência, que nem sempre está diretamente ligada à saúde.

De acordo com Andrade e Cardoso (2012, p. 130),

“Os trabalhadores estão sujeitos a condições de trabalho que podem gerar sofrimento, tensão emocional, insatisfação, irritação, insônia, envelhecimento prematuro, aumento do adoecimento e morte por doenças cardiovasculares e outras doenças crônico-degenerativas como as osteomusculares”.

Sendo assim a implementação de uma estratégia bem elaborada, de controle de absenteísmo torna-se extremamente relevante no ambiente corporativo, visando a melhora motivacional do trabalhador em relação à corporação e/ou à atividade exercida por este, promovendo desta forma a qualidade de vida dentro do ambiente de trabalho.

2.2 - INTELIGÊNCIA ARTIFICIAL

A inteligência artificial no início já foi vista como área de valor pouco prático visto que a grande maioria dos problemas podiam ser resolvidos por meios de programação dos passos necessários a resolução do problema. Contudo, a partir da década de 70, houve grande disseminação da resolução de problemas práticos com o uso de técnicas de IA através da aquisição de conhecimento de especialistas de um dado domínio, sendo conhecidos como Sistemas Especialistas ou Sistemas Baseados em Conhecimento.

Segundo Silva (2019, p. 13),

“O termo inteligência artificial representa um software diferente dos demais, pois é inteligente e visa fazer os computadores realizarem funções que eram exclusivamente dos seres humanos, por exemplo, praticar a linguagem escrita ou falada, aprender, reconhecer expressões faciais, etc. Seu campo tem um longo histórico e muitos avanços, como o reconhecimento de caracteres ópticos, que atualmente são considerados de rotina”.

Frente a crescente complexidade dos problemas e do volume dos dados gerados, torna-se evidente a necessidade de ferramentas computacionalmente mais sofisticadas e autônomas e com a menor intervenção humana. Para tal, as técnicas aplicadas devem ser capazes de criar por si próprias, uma função e/ou hipótese a partir da experiência passada de forma a resolver o problema. Ao processo de indução de uma hipótese dá-se o nome de Aprendizagem de Máquina (AM).

Segundo Russel e Norvig (2013, p. 29)

Um agente estará aprendendo se melhorar o seu desempenho nas tarefas futuras de aprendizagem após fazer observações sobre o mundo. A aprendizagem pode variar do corriqueiro, como anotar um número de telefone, até o profundo, como mostrado por Albert Einstein, que inferiu uma nova teoria para o universo.

Algoritmos de AM são vastamente utilizadas na resolução de diversos tipos de tarefas, que são classificadas como Preditivas e Descritivas. Tarefas preditivas são aquelas, onde os algoritmos de AM são aplicados a um conjunto de dados de treinamento para induzir um modelo capaz de predizer se um novo objeto, através da representação de dados de seus atributos, o valor do seu atributo alvo. Em geral para as tarefas preditivas, os algoritmos se utilizam do paradigma de aprendizado Supervisionado (Classificação e Regressão), ou seja, de um agente externo que conhece o valor verdadeiro dos atributos do novo objeto em questão. Já as tarefas descritivas são aquelas onde o algoritmo de AM faz a extração de padrões dos valores preditivos do conjunto de dados, não fazendo uso de um agente supervisor, estes algoritmos se utilizam do paradigma de aprendizagem não supervisionado (Agrupamento, Associação e Sumarização).

A área da saúde é uma das maiores utilizadoras das técnicas de AM com ferramentas computacionais desenvolvidas para as mais diversas aplicações como auxilio para diagnóstico médico, auxilio a exames clínicos (eletrocardiogramas, eletroencefalogramas, ressonância magnética, raios X, mamografias), dosagem de medicamentos, monitoramento de pacientes, dentre outros. Diferentemente de outras áreas de aplicação, a utilização da inteligência artificial no campo da saúde tem um custo relevante para o paciente. Este custo nada tem com a escala financeira, uma vez que diagnosticar um paciente como saudável, quando na verdade este se encontra doente é pior do que diagnosticar o mesmo paciente como doente, mas em estágio inicial, visto que o diagnóstico errado pode ter graves consequências.

**3. METODOLOGIA**

Esta seção descreverá as etapas necessárias pra o desenvolvimento do presente trabalho. O trabalho consiste no desenvolvimento de um sistema de software baseado em Sistema Especialista (SE) e Raciocínio Baseado em Casos (RBC) para auxiliar na identificação de colaboradores, que possivelmente, possam desenvolver ou tenham tendência a desenvolver doenças ocupacionais, para que uma ação prevencionista seja tomada adiantadamente. Para tal, será adotado o modelo de desenvolvimento incremental, intercalando as atividades de especificação, desenvolvimento e validação. Como metodologia, será adotado a metodologia ágil *Lean Software Development* (LD) tendo como foco principal o princípio **P1** (Evitar desperdícios) que visa evitar atividades sem valor ao cliente.

Para a construção do sistema proposto, será feita uma entrevista com um médico e/ou enfermeiro ocupacionais, para se estabelecer as variáveis necessárias a serem utilizadas pelo SE e pelo RBC. Através desta entrevista, objetiva-se a identificação dos requisitos funcionais (RF) e não-funcionais (RNF) do sistema de forma a garantir que estes atendam aos reais propósitos da aplicação de software. Em seguida, faz-se necessário a coleta de dados de triagem médica ocupacional e questionários de saúdes para carga no banco de dados e para uso do RBC na aprendizagem e a identificação de causas raízes dos problemas de saúde tais como alcoolismo e tabagismo.

A pesquisa bibliográfica relacionada a saúde ocupacional será feita, para um entendimento mais profundo do assunto, que se faz necessário para a posterior modelagem do SE e para a criação das árvores de decisão e das regras que serão validadas no sistema. Nesta etapa, uma vez que já definidos os RF e RNF, será iniciada a modelagem do projeto através da UML (*Unified Modeling Language*) com diagrama de caso de uso, diagramas de classes, diagrama de sequência e modelagem do banco de dados.

Após a modelagem do sistema, será iniciado o desenvolvimento das telas do sistema baseados nos requisitos previamente identificados na entrevista com o especialista. Por fim, será feita a codificação e teste da aplicação de software e apresentação dos resultados obtidos através das pesquisas bibliográficas e do funcionamento do sistema.

**Quadro 1** - Atividades para Construção de Projeto de Software.

|  |  |
| --- | --- |
| **Atividade** | **Artefatos resultantes** |
| Engenharia de Requisitos | Escolha de Stakeholders e entrevista levantamento de requisitos funcionais e não-funcionais. |
| Elaboração do Projeto de Software | diagrama de caso de uso, diagrama de classes, diagrama de sequência e modelo para banco de dados. |
| Tecnologias a serem utilizadas e os requisitos de implantação | Linguagem de programação *JAVA* e banco de dados relacional *MySQL*. Operação somente em SO *Windows* e *Java Virtual Machine* deve estar instalada. |

**4. DESENVOLVIMENTO**

4.1 – Introdução

O desenvolvimento do projeto se deu, primeiramente com a escolha dos stakeholders após uma discussão sobre o processo de automação de trabalhos relacionados a coleta de dados de saúde de uma população de trabalhadores *Offshore*, por meio de entrevista com o médico e enfermeiro do trabalho, para identificar quais dados seriam mais relevantes no estudo. Em etapa seguinte foi executado o levantamento dos requisitos funcionais e não funcionais do sistema a ser desenvolvido, embasando a modelagem do software com casos de usos, diagrama de classes e diagrama de sequências. O desenvolvimento do código bem como das telas foi executado em conjunto com a modelagem do banco de dados, o que proporcionou a execução de testes iniciais.

4.2 - Escopo do Produto

O sistema de software deve permitir que usuário seja capaz de obter informações pessoais e de saúde (fornecidos através de relatórios) ou salvos em um banco de dados, de onde posteriormente, os dados serão coletados para análise pelo sistema

4.3 - Público-Alvo

Tal sistema de software terá como público alvo, clínicas de medicina ocupacional e/ou empresas, que queiram analisar os dados clínicos de seus funcionários, de forma a garantir um melhor monitoramento da condição de saúde de seus colaboradores.

4.4 - Convenções

O usuário será classificado como toda pessoa que tenha interação com o sistema de software, seja para utilização na alimentação e análise dos dados ou para manutenção (corretiva e/ou evolutiva) do sistema de software. Paciente estará classificado como toda pessoa que será incluída, alterada, ou excluída no sistema de software em questão.

4.5 - Prioridade dos Requisitos

A prioridade dos requisitos do sistema de software estarão classificados em:

* **Essencial**: requisito sem o qual o sistema não entra em funcionamento. São requisitos imprescindíveis, devendo ser disponibilizados na implantação do sistema.
* **Importante**: requisito sem o qual o sistema entra em funcionamento, mas de forma não satisfatória. Não impedem a implantação do sistema, mas devem ser implementados o mais breve possível.

4.6 - Ambiente Operacional

O sistema proposto para desenvolvimento é um sistema *Desktop*, desenvolvido exclusivamente para sistema operacional *Windows*.

4.7 - Restrições de Projeto e Implementação

O sistema proposto não funcionará em ambientes *Linux* e *MacOS*. Os requisitos mínimos são: Memória RAM de 2GB e espaço livre de 50MB no disco rígido.

4.8 - Tecnologias Utilizadas

O sistema foi desenvolvido na linguagem JAVA e a escolha se deu devido a versatilidade de ser utilizadas em sistemas desktop, web e mobile. O banco de dados utilizados foi o MySQL por ser de código aberto e gratuito. Para a modelagem dos diagramas UML (*Unified Modeling Language*), a ferramenta utilizada foi a Astah UML por ser de código aberto e por permitir a modelagem dos mais diversos diagramas, pela facilidade na utilização da ferramenta por sua interface gráfica amigável e intuitiva. A modelagem do banco de dados, se deu através da ferramenta BRModelo devido ao foco no modelo conceitual no desenvolvimento do banco de dados, ao contrário da maioria das ferramentas disponíveis no mercado, além de permitir alterações na estrutura do modelo, caso necessário. O desenvolvimento de telas e da programação do sistema foi feito através na IDE (*Integrated Development Environment*) NetBeans por possuir um conjunto de ferramentas que dão suporte ao desenvolvedor, o ajudando a focar mais no código do que em implementações secundárias e por escolha pessoal devido a familiaridade com tal IDE.

**4.9 DOCUMENTO DE DEFINIÇÃO DE REQUISITOS (DDR)**

4.9.1 – Objetivo

Este documento tem por objetivo concentrar e definir a identificação dos requisitos para o sistema a ser desenvolvido, fornecendo a equipe de projetos, informações pertinentes para implementação do sistema, bem como para realização de testes. O processo de elicitação se deu através de entrevista com profissionais da área de saúde, buscando um enfoque nos problemas enfrentados por estes durante análise e/ou interpretação dos dados de saúde coletados.

4.9.2 - Definições, Acrônimos e Abreviações.

Um requisito é uma condição ou uma capacidade com o qual o sistema deve estar de acordo, expressando as necessidades do cliente (CASTRO, 2014). Podem ser dos seguintes tipos:

**RF (Requisito Funcional):** Definem as funcionalidades do Módulo Publicação e Vendas a serem implementadas pelos desenvolvedores na construção do sistema, a fim de possibilitar que os usuários realizem suas tarefas e satisfaçam os requisitos de negócio (CASTRO, 2014).

**RNF (Requisito Não-Funcional):** Relacionam os aspectos de qualidade desejada (requisitos não-funcionais de qualidade), como confiabilidade, eficiência, portabilidade, usabilidade ou qualquer outra característica que o sistema deva atender, como padrões, regulamentos e contratos com os quais o sistema deve ter conformidade (CASTRO, 2014).

4.9.3 - Características do Usuário (Stakeholders)

* **Administrador:** pessoa responsável pela manutenção do sistema de forma geral, bem como o único responsável pela inclusão de novos usuários e seus respectivos níveis de acesso.
* **Médico:** responsável por operar o sistema de acordo com o nível de acesso disponibilizado, podendo este alterar dados e/ou excluí-los.
* **Atendente:** pessoa que interage com o sistema tão quanto o Médico, porém não deve ter as funcionalidades de alterar e excluir dados disponível para o nível de acesso.

4.9.4 - Requisitos Funcionais

**RF01 –** Fazer Login**.**

A aplicação de software deve permitir aos usuários acesso somente por meio de login e senha.

**ATOR:** Usuário **Prioridade:** Essencial

**RF02 –** Cadastrar dados funcionário

A aplicação de software deve permitir o cadastro de dados pessoais dos funcionários com entrada destes na tela de cadastro.

**ATOR:** Usuário **Prioridade:** Essencial.

**RF03 –** Buscar dados funcionário

A aplicação de software deve permitir a busca por dados de um funcionário através de um ID.

**ATOR:** Usuário **Prioridade:** Essencial.

**RF04 –** Limpar dados da tela

A aplicação de software deverá permitir a limpeza dos dados da tela de funcionários após uma busca ou cadastro dos dados.

**ATOR:** Usuário **Prioridade:** Essencial

**RF05 –** Alterar dados funcionário

A aplicação de software deverá permitir a alteração dos dados do funcionário somente se o nível de acesso for o de médico

**ATOR:** Médico **Prioridade:** Essencial

**RF06 –** Excluir dados funcionário

A aplicação de software deverá permitir a alteração dos dados do funcionário somente se o nível de acesso for o de médico

**ATOR:** Médico **Prioridade:** Essencial

**RF07 –** Cadastrar usuário

A aplicação de software deverá o cadastro e/ou exclusão de usuários do sistema somente pelo administrador. Todo usuário deve estar cadastrado como funcionário no sistema.

**ATOR:** Administrador **Prioridade:** Essencial

**RF08 –** Gerar ASO

A aplicação de software deverá permitir a geração do atestado de saúde ocupacional mediante o ID do funcionário.

**ATOR:** Atendente/Médico **Prioridade:** Essencial

**RF09 –** Imprimir ASO

A aplicação de software deverá permitir a impressão do atestado de saúde ocupacional.

**ATOR:** Atendente/Médico **Prioridade:** Essencial

**RF10 –** Enviar ASO

A aplicação de software deverá permitir o envio do atestado de saúde ocupacional por correio eletrônico.

**ATOR:** Atendente/Médico **Prioridade:** Essencial

**RF11 –** Carregar CSV

O sistema deverá permitir que um arquivo CSV seja carregado quando de uma quantidade muito grande de dados, evitando assim um trabalho manual de inserção dos dados.

**ATOR:** Atendente **Prioridade:** Importante

**RF12 –** Preencher anamnese

O sistema deverá permitir o cadastro de dados clínicos e de riscos dos funcionários com entrada destes na tela de anamnese.

**ATOR:** Atendente **Prioridade:** Essencial

**RF13 –** Buscar anamnese funcionário

A aplicação de software deve permitir a busca por dados da anamnese de um funcionário através de um ID.

**ATOR:** Usuário **Prioridade:** Essencial.

**RF14 –** Limpar anamnese da tela

A aplicação de software deverá permitir a limpeza dos dados da tela de anamnese após uma busca ou cadastro dos dados.

**ATOR:** Usuário **Prioridade:** Essencial

**RF15 –** Alterar anamnese funcionário

A aplicação de software deverá permitir a alteração dos dados da anamnese do funcionário somente se o nível de acesso for o de médico.

**ATOR:** Médico **Prioridade:** Essencial

**RF16 –** Excluir anamnese funcionário

A aplicação de software deverá permitir a alteração dos dados da anamnese do funcionário somente se o nível de acesso for o de médico.

**ATOR:** Médico **Prioridade:** Essencial

4.9.5 - Requisitos Não Funcionais de Qualidade

**RNF01** – Interface Amigável

O design de layout da aplicação de software deverá ser em português, não contendo palavras que possam causar ambiguidade ou de difícil compreensão. Caso haja algum erro na autenticação dos dados, o sistema não permitirá que o LOGIN seja efetuado e exibirá uma mensagem de erro que houve um erro na informação inserida.

**Prioridade:** Importante

**RNF02** – Backup

O sistema realizará salvamento dos dados a cada 4 horas. Este processo contempla a cópia de todas as informações contidas no SGBD.

**Prioridade:** Importante

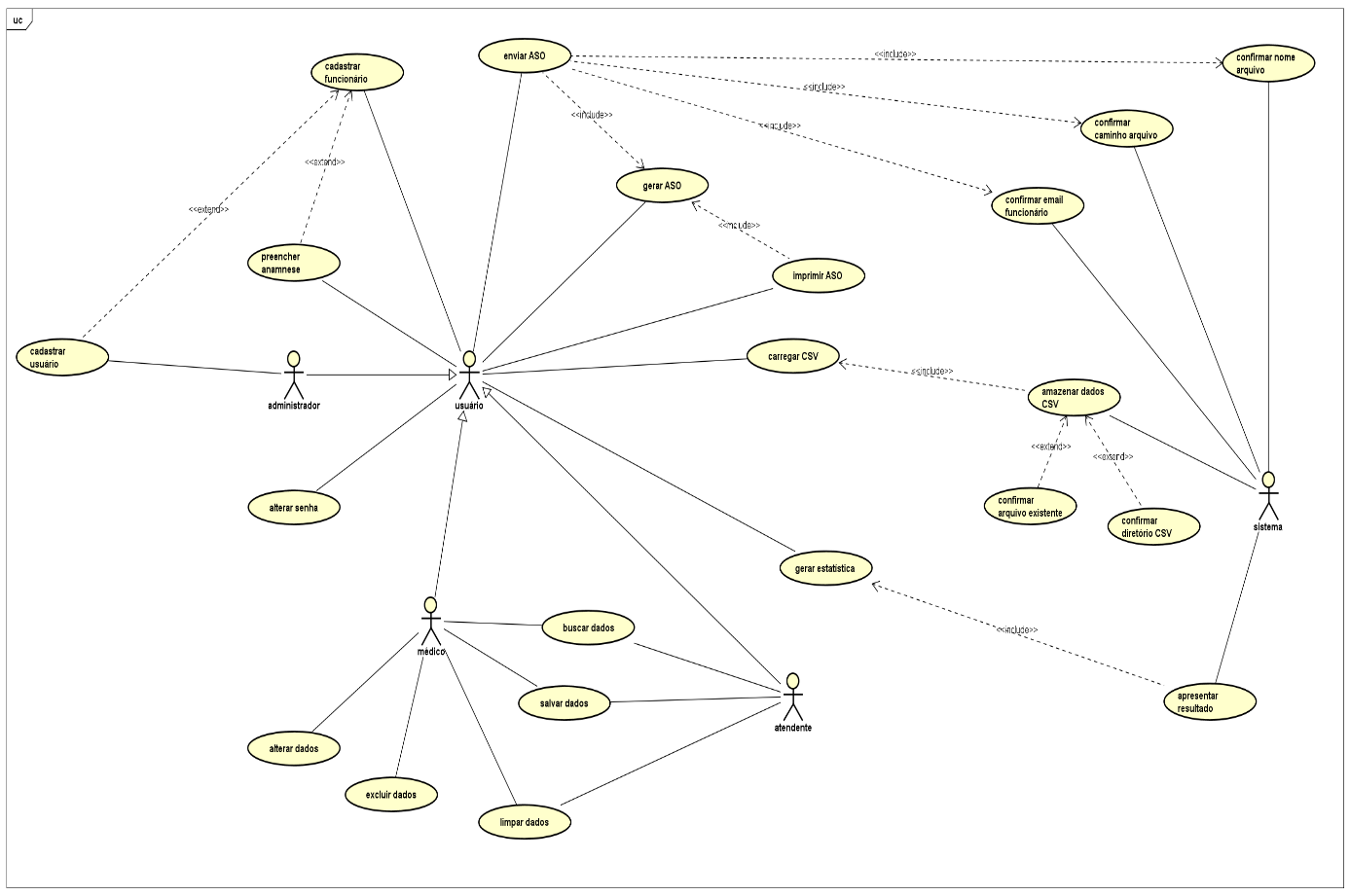
**RNF03** – Performance

As ações como inclusão, exclusão, alteração ou busca de dados não deverão levar mais do 5s em seu processamento.

**Prioridade:** Importante.

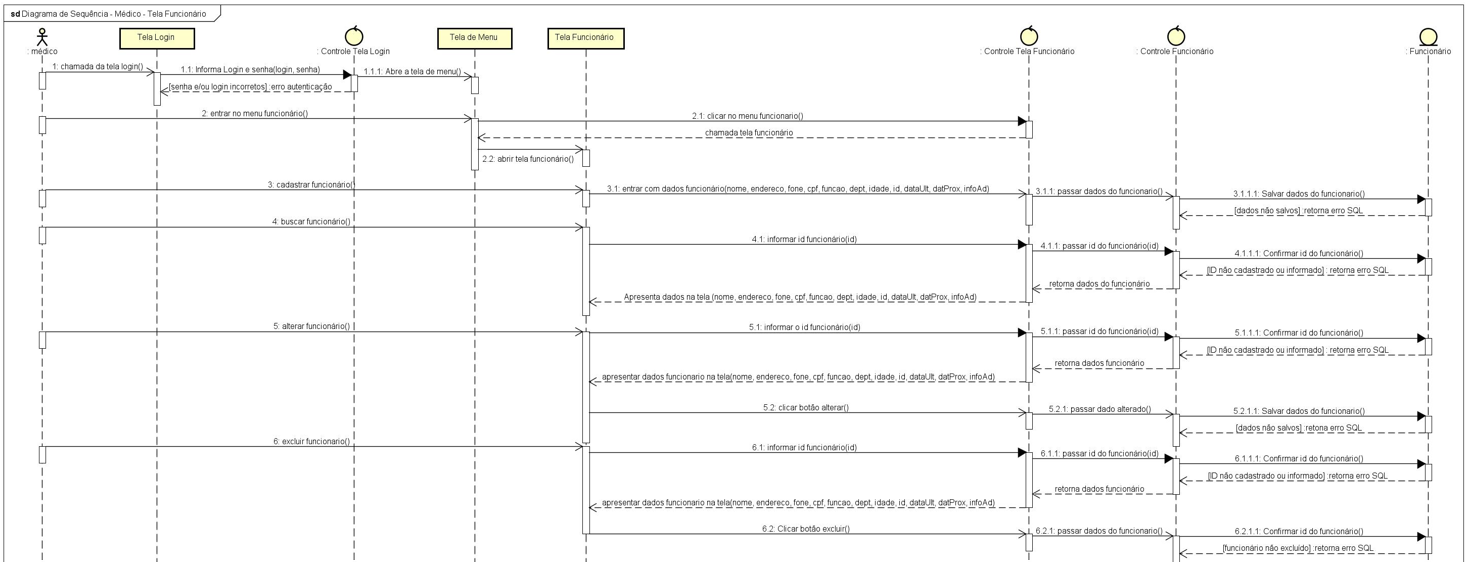
4.10 - Modelagem do Sistema – UML

Esta seção apresentará os diagramas de caso de uso e diagrama de sequência modelados durante o desenvolvimento do sistema SIGSMO. O diagrama de classes encontra-se no **Anexo** 1.



**Figura 1** – Diagrama de Caso de Uso

Este diagrama representa os passos percorridos pelo sistema percorre internamente para a funcionalidade do botão FUNCIONÁRIO.

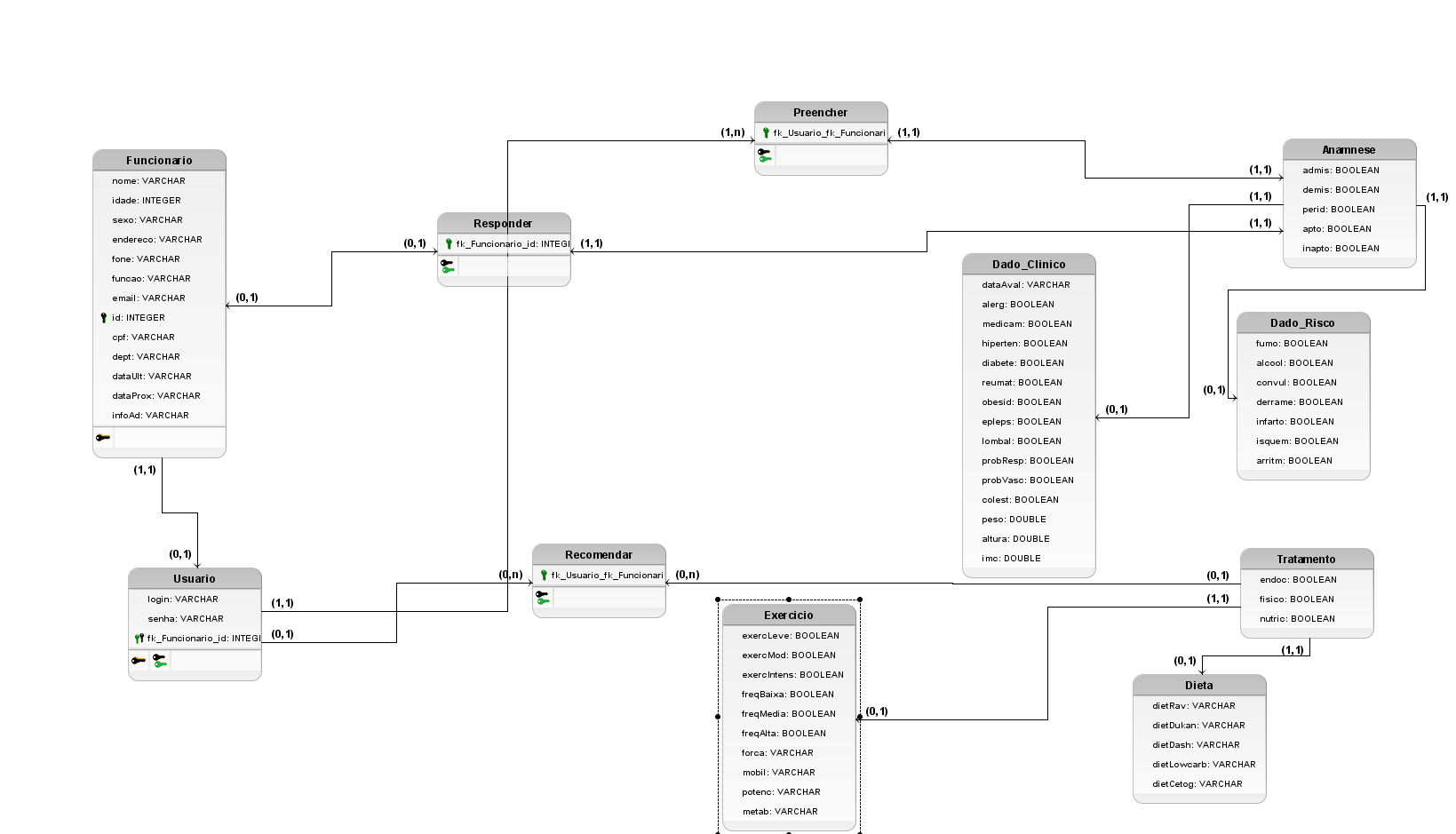


**Figura 2** – Diagrama de Sequência.

4.11 - Modelagem do Banco de Dados

Para a modelagem do banco de dados utilizou-se a ferramenta brModelo para que fosse possível a construção de um modelo lógico do pretendido banco de dados.

Em seguida, utilizou-se da ferramenta mySQL para a criação do banco de dados com o uso de query e possível alterações nas tabelas, usando o processo de engenharia reversa para a geração do MER (Modelo Entidade Relacionamento).



**Figura 3** – Modelo Lógico do Banco de Dados.

4.12 - JAVA Machine Learning

Para a classificação das instâncias dos funcionários, foi utilizada a API JavaML e o algoritmo de árvore de decisão J48. Os dados foram carregados a partir de um arquivo (Data Source) para a aplicação e parte dos dados foi selecionado para treino. A segunda parte do arquivo de dados foi destinado ao teste para determinar a precisão da classificação dos funcionários de acordo com o risco.

Uma vez que estes dados tenham sido carregados para a classe ClassificarInst, o algoritmo então é executado classificando as instâncias e determinando o percentual de instâncias classificadas correta e incorretamente. Na tela do sistema SIGSMO, temos ainda os valores representados e o percentual de classificação, conforme apresentado no **Anexo 2**.

Em geral, a performance do método de aprendizagem de máquina é medida pelo indicador de acurácia, precisão e chamada baseados na matriz de confusão. A matriz de confusão resume os resultados de uma comparação da classe prevista e da classe correta na matriz. Em sumo, mostra a frequência de classificação para cada classe do modelo.

Como podemos observar na matriz apresentada no **Anexo 3**, para a classificação das instâncias de “a=nível baixo” tivemos o total de 74 TPs (*true positive*) positivos verdadeiros, ou seja, no conjunto real, as classes que estamos buscando foram corretamente previstas. Para a classificação das instâncias “b=baixo\_limitrofe” tivemos o total de 2 TPs, 4 FNs (False Negative) falsos negativos, ou seja, no conjunto real a classe que não estamos buscando prever foi prevista incorretamente e 1 FP (False Positive) falso positivo, ou seja, no conjunto real a classe que estamos buscando prever foi prevista incorretamente. Para a classificação das instâncias “c=moderado” tivemos 15 verdadeiros positivos (TPs) e somente 2 falsos negativos (FPs). Para a classificação “d=moderado\_limitrofe” tivemos 16 verdadeiros positivos (TPs) e 2 falsos negativos (FNs). E finalizando a matriz, para a classificação de “e=alto” tivemos 15 verdadeiros positivos (TPs) e 1 falso negativo (FN).

De acordo com o obtido na matriz de confusão, podemos observar que esta está quase que na sua totalidade diagonal, com mais imprecisão na classificação de nível b=baixo\_limitrofe. Contudo, as demais classificações apresentam resultados mais precisos, o que corrobora com o conceito de entropia que associa a ideia de que se o resultado de um experimento é incerto, então maior informação se obtém na observância da ocorrência deste.

**5. CONCLUSÕES**

Para o desenvolvimento do sistema proposto foi necessário a integração e colaboração de profissionais de diferentes expertises, para que fosse necessário o alcance do conhecimento específico dentro do campo de avaliação na medicina ocupacional, sendo este fator importante para o desenvolvimento do sistema.

O desenvolvimento das telas se deu somente depois de uma longa discussão sobre os itens que seriam necessários para a avaliação, fase inicial de coleta dos requisitos funcionais, passando ainda por algumas alterações para completude de sua utilização pelo especialista. Buscou-se desenvolver uma interface de usuário entendível, simples e organizada, facilitando o uso por parte do usuário. Para uma maior confiabilidade dos resultados práticos do sistema desenvolvido, utilizou-se de uma coleta de dados reais de indivíduos como amostra, montando assim, a base de dados inicial que compõe o sistema SIGSMO.

Apesar de já existirem outras ferramentas no mercado com o mesmo propósito, o diferencial do sistema proposto é a funcionalidade voltada para análise dos dados dentro do campo da estatística, o que intrinsicamente, colabora com o campo econômico da corporação que venha a utilizar tal ferramenta. Isto se dá, por um contexto de avaliação do absenteísmo como custo oculto às organizações e que medidas podem ser tomadas para obter melhores resultados. Contudo, um problema encontrado foi o fato de a grande maioria estar somente disponível para sistema operacional *Windows*.

Com relação aos objetivos deste trabalho, foi possível identificar correlação entre os indivíduos com doenças associadas à maus hábitos, através do questionário preenchido através do sistema e posterior classificação do nível de risco. Tal resultado é evidenciado através da funcionalidade de geração do ASO (Atestado de Saúde Ocupacional), no campo de recomendações. Entretanto, uma dificuldade encontrada durante a fase inicial do projeto foi a disponibilidade de colaboradores que aceitassem participar da pesquisa inicial, com o fornecimento de dados pessoais, causando assim uma limitação na quantidade de dados fornecidos. Como continuidade deste trabalho, pretende-se aplicar o desenvolvimento com uso de linguagem natural utilizados nos prontuários, para uma análise mais precisa do estado de saúde de cada indivíduo. Ainda como melhoria, pretende-se que o sistema seja desenvolvido em um *builder* mais atual e com aplicação de persistência de dados, tornando-o assim, um sistema mais robusto em quantidade de informações disponíveis para uso. O código fonte encontra-se disponível em <https://github.com/BalbiLeo/projeto-final-de-curso>.

**REFERÊNCIAS**

ANDRADE, P. S. de.; CARDOSO, T. A. de O. Prazer e dor na docência: Revisão bibliográfica sobre a Síndrome de Burnout, Ver. Saúde soc. [online], vol. 21, n. 1, p. 130, 2012. Disponível em: [https://www.scielo.br/pdf/sausoc/v21n1/13.pdf.](https://www.scielo.br/pdf/sausoc/v21n1/13.pdf.%20)  Acesso em 28/03/2021

CASTRO, Eduardo J. R. de; CALAZANS, Angélica T. S.; PALDÊS, Roberto A.; GUIMARÃES, Fernando A. Engenharia de Requisitos: Um enfoque prático na construção de software orientado ao negócio. Editora Bookess, 2014.

CHIAVENATO, Idalberto. **Gestão de pessoas: o novo papel dos recursos humanos nas organizações** - 4. ed. -- Barueri, SP: Manole, 2014.

COULTHARD, Phil; FARR, George – **Java for RPG programmers** – 3rd edition – Toronto – Canada, 2006.

COPPIN, Bem. **Inteligência artificial** / Ben Coppin; tradução e revisão técnica Jorge Duarte Pires Valério. - [Reimpr.]. - Rio de Janeiro: LTC, 2013.

DALL´INHA, G. R. **A influência das práticas de recursos humanos sobre o absenteísmo e a rotatividade:** um estudo de caso. 2006. Dissertação (Mestrado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina: Florianópolis, 2006, p. 26-30.

Documento de Requisitos – Essencial ao Desenvolvimento de Software. Online – Disponível em: <https://www.devmedia.com.br/artigo-engenharia-de-software-10-documento-de-requisitos/11909> Acesso em 03/04/2021.

GRIGOREV, Alexey. **Mastering Java for Data Science** – Packt Publishing – Birmingham, 2017.

IAN, Witten H.; FRANK, Eibe; HALL Mark A. – DATA MINING – Practical Machine Learning Tools and Techniques – 3rd edition – Morgan Kaufmann, 2011 – p. 203 – 2016.

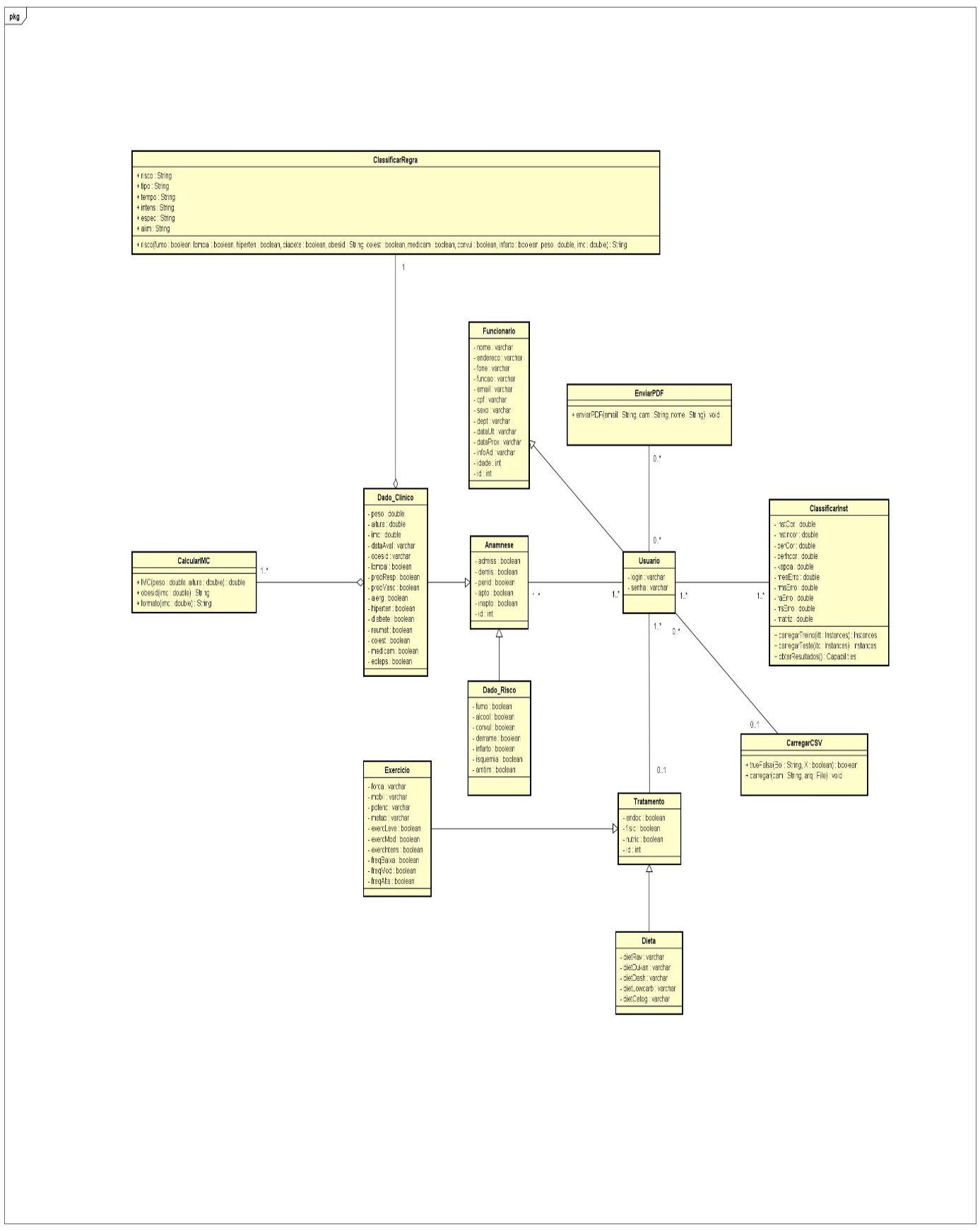
RUSSEL, Stuart J.; NORVIG, Peter. **Inteligência artificial** – Rio de Janeiro: Elsevier, 2013.

ROKACH, Lior.; MAIMON, Oded. DATA MINING WITH DECISION TREE – 2nd edition – World Scientific – 2015. – p. 77 – 81.

SOUZA, Luís Fernando Quinteiro de. [Absenteísmo no serviço público](https://jus.com.br/artigos/9204/absenteismo-no-servico-publico). **Revista Jus Navigandi**, ISSN 1518-4862, Teresina, [ano 11](https://jus.com.br/revista/edicoes/2006), [n. 1243](https://jus.com.br/revista/edicoes/2006/11/26), [26](https://jus.com.br/revista/edicoes/2006/11/26) [nov.](https://jus.com.br/revista/edicoes/2006/11) [2006](https://jus.com.br/revista/edicoes/2006). Disponível em: <https://jus.com.br/artigos/9204>. Acesso em: 27/03/2021

ANEXOS

**Anexo 1** – Diagrama de classes do Sistema SIGSMO.



**Anexo 2** – Quadro Estatístico do Sistema SIGSMO.



**Anexo 3** – Matriz de Confusão.

