

# JNT-FACIT BUSINESS AND TECHNOLOGY JOURNAL - ISSN: 2526-4281 QUALIS B1



**MINERAÇÃO DE TEXTO EM RELATÓRIO  
OPERACIONAL DO PRONTUÁRIO ELETRÔNICO PARA  
AUXÍLIO NA TRIAGEM DA SÍNDROME METABÓLICA**

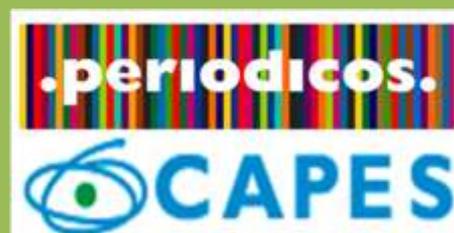
**TEXT MINING IN OPERATIONAL REPORT OF  
ELECTRONIC RECORD TO ASSIST THE SCREENING OF  
METABOLIC SYNDROME**

**Murilo de Paula FLAVIO**  
Centro Universitário Tocantinense Presidente  
Antônio Carlos (UNITPAC)  
E-mail: murilopflavio@gmail.com

**Dario Farias NUNES**  
Centro Universitário Tocantinense Presidente  
Antônio Carlos (UNITPAC)  
Email: dariofariasnunes@hotmail.com

**Rafaella Dias GALVÃO**  
Centro Universitário Tocantinense Presidente  
Antônio Carlos (UNITPAC)  
Email: rafaelladgalvao@gmail.com

**Rodolfo Lima ARAÚJO**  
Centro Universitário Tocantinense Presidente  
Antônio Carlos (UNITPAC)  
Email: rodolfo.araujo@unitpac.edu.br



## RESUMO

**Tema:** A Síndrome Metabólica é um termo médico caracterizado pela combinação de vários fatores de risco importantes para doenças cardiovasculares, sendo estes a hipertensão, diabetes, dislipidemia e obesidade. Considerada uma doença crônica não transmissível necessita de abordagens multifatoriais dentro do espectro de tratamento, para a redução da morbimortalidade. **Objetivo:** Obter por meio de um processo automatizado de mineração de texto, um algoritmo de classificação que auxilia a triagem da Síndrome Metabólica, na atenção primária, vide atenção básica em saúde. **Metodologia:** Trata-se de uma pesquisa de cunho exploratório, tendo-se realizado a mineração de texto a respeito das condições de saúde de 39 pacientes a partir do Relatório Operacional do Prontuário Eletrônico de uma Unidade Básica de Saúde, no período entre agosto de 2021 e novembro de 2021. Em seguida, fez-se uma correlação com os dados recuperados de forma manual, validando a efetividade e a precisão do algoritmo. **Resultados:** observou-se íntima relação entre as doenças crônicas e a Síndrome Metabólica, a precisão e a efetividade dos algoritmos de mineração de dados e texto tiveram como obstáculo fatores relacionados à linguagem de programação do Prontuário Eletrônico, impossibilitou a obtenção dos dados sem interferência humana. **Conclusão:** A ciência de dados e as ferramentas aliadas a ela dão início a uma nova era no cuidado a saúde, causando impactos na forma como lidamos com os dados clínicos dos pacientes e como manejamos as condições de saúde. O manejo correto das doenças crônicas associadas à Síndrome Metabólica é um importante fator redutor de Doenças Cardiovasculares.

**Palavras-chave:** Síndrome Metabólica; Mineração de Texto; Ciência de Dados; Saúde.

## ABSTRACT

**Theme:** Metabolic Syndrome is a medical term characterized by the combination of several important risk factors for cardiovascular diseases, which are hypertension, diabetes, dyslipidemia and obesity. Considered a non-communicable chronic disease, it requires multifactorial approaches within the treatment spectrum to reduce morbidity and mortality. **Objective:** To obtain, through an automated text mining process, a classification algorithm that helps the screening of Metabolic Syndrome, in primary care. **Methodology:** This is an

**Murilo de Paula FLAVIO; Dario Farias NUNES; Rafaella Dias GALVÃO; Rodolfo Lima ARAÚJO. Mineração de Texto em Relatório Operacional do Prontuário Eletrônico Para Auxílio na Triagem da Síndrome Metabólica. Facit Business And Technology Journal. QUALIS B1. ISSN: 2526-4281 <http://revistas.faculdefacit.edu.br/index.php/JNT>. Out/Nov - 2021. Ed. 31; V. 2. Págs. 259-268.**

exploratory research, having carried out a text mining about the health conditions of 39 patients from the Operational Report of the Electronic Medical Record of a Basic Health Unit, in the period between August 2021 and November 2021. Then, a correlation was made with the data retrieved manually, validating the effectiveness and accuracy of the algorithm. **Results:** there was a close relationship between chronic diseases and Metabolic Syndrome, the accuracy and effectiveness of data and text mining algorithms had as an obstacle factors related to the programming language of the Electronic Citizen's Record, which made it impossible to obtain the data without human interference. **Conclusion:** Data science and the tools associated with it usher in a new era in healthcare, impacting the way we handle clinical patient data and how we manage healthcare conditions. The correct management of chronic diseases associated with Metabolic Syndrome is an important reducing factor for Cardiovascular Diseases.

**Keywords:** Metabolic Syndrome; Text Mining; Data Science; Health.

## INTRODUÇÃO

A Síndrome Metabólica (SM) é caracterizada por um grupo de fatores de risco inter-relacionados, incluindo obesidade abdominal, hiperglicemia, hipertensão e dislipidemia. A condição está associada a um risco aumentado de diabetes mellitus tipo 2 (DM2), eventos cardiovasculares e mortes.

Essa nova doença crônica não transmissível (DCNT) tornou-se o maior risco à saúde do mundo moderno. Embora tenha começado no mundo ocidental, com a disseminação do estilo de vida ocidental, tornou-se agora um problema verdadeiramente global. A prevalência da síndrome metabólica é frequentemente maior na população urbana de alguns países em desenvolvimento do que em seus homólogos ocidentais.

O diagnóstico clínico e a avaliação laboratorial são baseados nos critérios do National Cholesterol Education Program's Adult Treatment Panel III (NCEP-ATP III) e da I Diretriz Brasileira de Diagnóstico e Tratamento da Síndrome Metabólica, associado a isto é feita a identificação de fatores de risco cardiovascular. De acordo com a Organização Mundial de Saúde, os fatores de risco mais importantes para a morbimortalidade relacionada às doenças crônicas não transmissíveis (DCNT) são: hipertensão arterial

**Murilo de Paula FLAVIO; Dario Farias NUNES; Rafaella Dias GALVÃO; Rodolfo Lima ARAÚJO. Mineração de Texto em Relatório Operacional do Prontuário Eletrônico Para Auxílio na Triagem da Síndrome Metabólica. Facit Business And Technology Journal. QUALIS B1. ISSN: 2526-4281 <http://revistas.faculadefacit.edu.br/index.php/JNT>. Out/Nov - 2021. Ed. 31; V. 2. Págs. 259-268.**

sistêmica, hipercolesterolemia, ingestão insuficiente de frutas, hortaliças e leguminosas, sobrepeso ou obesidade, inatividade física e tabagismo. Cinco desses fatores de risco estão relacionados à alimentação e à atividade física e três deles têm grande impacto no aparecimento da Síndrome Metabólica.

A predisposição genética, a alimentação inadequada e a inatividade física estão entre os principais fatores que contribuem para o surgimento da SM, cuja prevenção primária é um desafio mundial contemporâneo, com importante repercussão para a saúde. Destaca-se o aumento da prevalência da obesidade em todo o Brasil e uma tendência especialmente preocupante do problema em crianças em idade escolar, em adolescentes e nos estratos de mais baixa renda. A adoção precoce por toda a população de estilos de vida relacionados à manutenção da saúde, como dieta adequada e prática regular de atividade física, preferencialmente desde a infância, é componente básico da prevenção da SM.

Diante de tantos fatores de risco, os métodos de mineração de dados e de textos têm recebido atenção considerável como novas estratégias para capturar a interdependência das variáveis de saúde ao fazer previsões de eventos de saúde complexos, como anteriormente citado a quantidade de fatores que podem acarretar a SM. Neste estudo, buscamos desenvolver uma abordagem desse tipo para fornecer apoio aos profissionais de saúde na triagem da SM por meio da ciência de dados.

261

## **OBJETIVOS**

Obter por meio de um processo automatizado de mineração de texto, um algoritmo de classificação que auxilia a triagem da Síndrome Metabólica, na atenção primária, vide atenção básica em saúde. Entende-se que atividades de triagem têm o objetivo de identificar o risco de vida em pacientes mediante a análise de seus sinais vitais.

## **METODOLOGIA**

Tratando-se de uma pesquisa de cunho exploratório, o método para execução das atividades propostas foi pautado em um levantamento bibliográfico nas plataformas digitais: Scielo, PubMed e Google Acadêmico, com os descritores: Síndrome Metabólica, Mineração de Dados, Mineração de Texto e Python.

Posteriormente foi gerado um relatório operacional contendo dados do período entre agosto de 2021 a novembro de 2021, para acompanhamento das condições de saúde

**Murilo de Paula FLAVIO; Dario Farias NUNES; Rafaella Dias GALVÃO; Rodolfo Lima ARAÚJO. Mineração de Texto em Relatório Operacional do Prontuário Eletrônico Para Auxílio na Triagem da Síndrome Metabólica. Facit Business And Technology Journal. QUALIS B1. ISSN: 2526-4281 <http://revistas.faculdefacit.edu.br/index.php/JNT>. Out/Nov - 2021. Ed. 31; V. 2. Págs. 259-268.**

de 39 pacientes pertencentes a uma equipe da Estratégia de Saúde da Família, extraído do Prontuário Eletrônico do Cidadão de uma Unidade Básica de Saúde, localizada em Araguaína - Tocantins. Durante a coleta dos dados, estes foram randomizados, compilados e subdivididos de forma anônima.

Feito isto, utilizou o Scikit-learn, um módulo na linguagem de programação Python que integra uma ampla gama de algoritmos de aprendizado de máquina. A importância das variáveis clínicas nas previsões do aprendizado de máquina criado pelo módulo foi calculada com base na metodologia usada pelo modelo específico. Neste estudo, a classificação foi determinada segundo características do classificador Gradient-Boosted Decision Trees (GB), que é uma técnica de aprendizado de máquina para aperfeiçoar o valor preditivo de um modelo por meio de etapas sucessivas no processo de aprendizado. Em adição, uma correlação com os dados recuperados de forma manual foi realizada, analisando a efetividade e a precisão do algoritmo.

Este estudo foi submetido ao Comitê de Ética e Pesquisa (CEP), mas com dispensa do Termo de Consentimento Livre e Esclarecido, uma vez que os dados foram obtidos através do Prontuário Eletrônico, não tendo a necessidade de entrar em contato direto com os participantes, e preservando também o anonimato dos pacientes cujos dados clínicos foram avaliados. Deste modo foi deferida a dispensa do TCLE. Entretanto, a pesquisa seguiu todos os aspectos contidos na Resolução do Conselho Nacional de Saúde (CNS) nº 466/12, a qual regulamenta pesquisas envolvendo seres humanos.

## **RESULTADOS E DISCUSSÃO**

O perfil da população estudada constituiu-se de 39 pacientes com problemas e condições de saúde, considerados como ativos no relatório operacional gerado no Prontuário Eletrônico do Cidadão (PEC) entre o período de agosto de 2021 a novembro de 2021. Os dados obtidos por meio do relatório foram salvos em tabelas que continham informações sobre Hipertensão, Diabetes, IMC e Tabagismo.

Foi observado que dos 39 pacientes, trinta (76.92%) se enquadravam nos critérios diagnósticos da Síndrome Metabólica. Do total trinta eram portadores de Hipertensão Arterial e Diabetes e nove (23.07%) possuíam somente Diabetes Melitus Tipo 2, os pacientes possuíam pelo menos uma doença crônica. (Tabela 1)

A pré-diabetes e o diabetes tipo 2 são frequentes manifestações de um distúrbio subjacente muito mais amplo, incluindo a SM. Embora a Síndrome Metabólica e o diabetes tipo 2 frequentemente coexistam, aqueles com SM sem diabetes correm um risco significativo de desenvolver diabetes. Existem evidências que apoiam a investigação clínica para identificar e tratar as pessoas, não apenas aquelas com hiperglicemia, mas também aqueles com os fatores de risco CV associados que compõem a síndrome metabólica, na esperança de reduzir significativamente a morbimortalidade.

O acompanhamento do IMC destes pacientes mostrou que em relação à amostragem total, doze (30.76 %) possuíam obesidade grau I (30 – 34,9), doze (30.76 %) obesidade grau II (35 – 39,9), nove (23.07%) estavam dentro da classificação como peso normal (18,5 – 24,9), três (07.79%) sobrepesos (25 – 29,9) e três (07.79%) com dados corrompidos ou não mensurados. (Tabela 2 e Figura 1)

Os trinta pacientes que corresponderam com os critérios da SM, dezoito (60%) eram mulheres e doze (40%) homens. Entre as dezoito mulheres, nove (50%) eram tabagistas, enquanto os homens não existiam pacientes tabagistas.

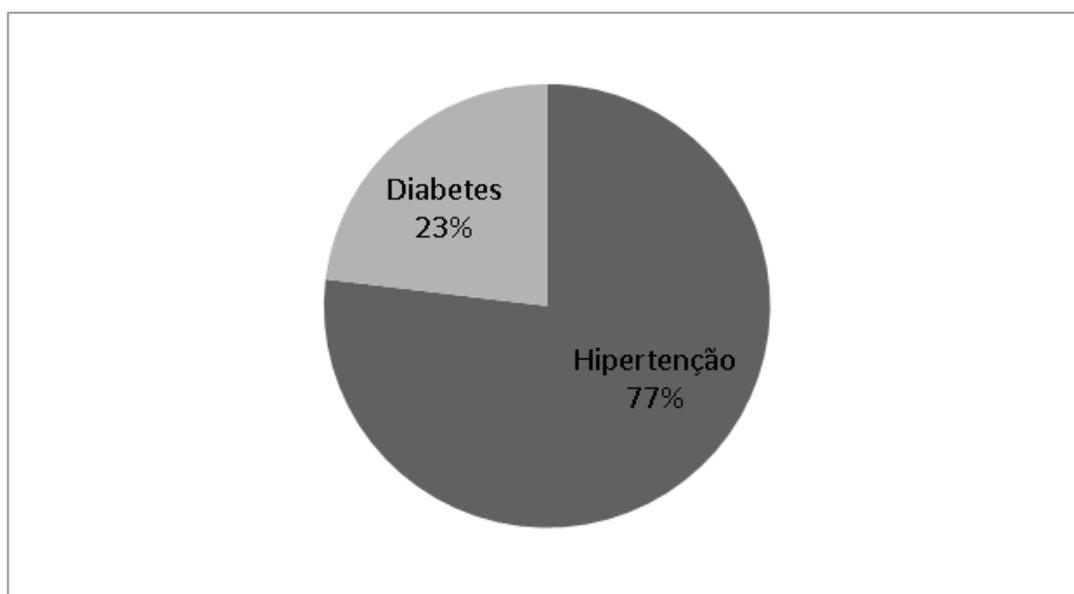
Uma consequência importante do tabagismo é o desequilíbrio hormonal que leva primeiro ao acúmulo de gordura central e, depois, à resistência à insulina. A última condição pode representar uma importante ligação entre o tabagismo e o risco de doenças cardiovasculares. Os fumantes atuais têm as taxas de mortalidade mais altas em todos os grupos de IMC e circunferência abdominal, a cessação do tabagismo está associada a um risco de mortalidade significativamente menor. Quanto mais tempo alguém for ex-fumante, menor será o risco de mortalidade. Embora a perda de peso possa reduzir o risco de mortalidade em fumantes atuais, a cessação do tabagismo está relacionada a uma redução mais forte do risco.

As diretrizes ATP-III reconhecem que a síndrome metabólica aumenta o risco de doença cardíaca coronária e o valor prognóstico da síndrome metabólica para doença cardíaca coronária incidente, doença cardiovascular, mortalidade cardiovascular, e a mortalidade por todas as causas foi estabelecida em diferentes contextos clínicos. O agrupamento de fatores de risco da síndrome de resistência à insulina predisse tanto doença coronariana quanto acidente vascular cerebral. A síndrome metabólica, definida pelo ATP-III, é um preditor independente de eventos cerebrovasculares e cardíacos. Em pacientes hipertensos com a síndrome metabólica há um risco aumentado de duas manifestações

diferentes de doença aterosclerótica, e a resistência à insulina, por si só, pode acelerar o desenvolvimento da aterosclerose.

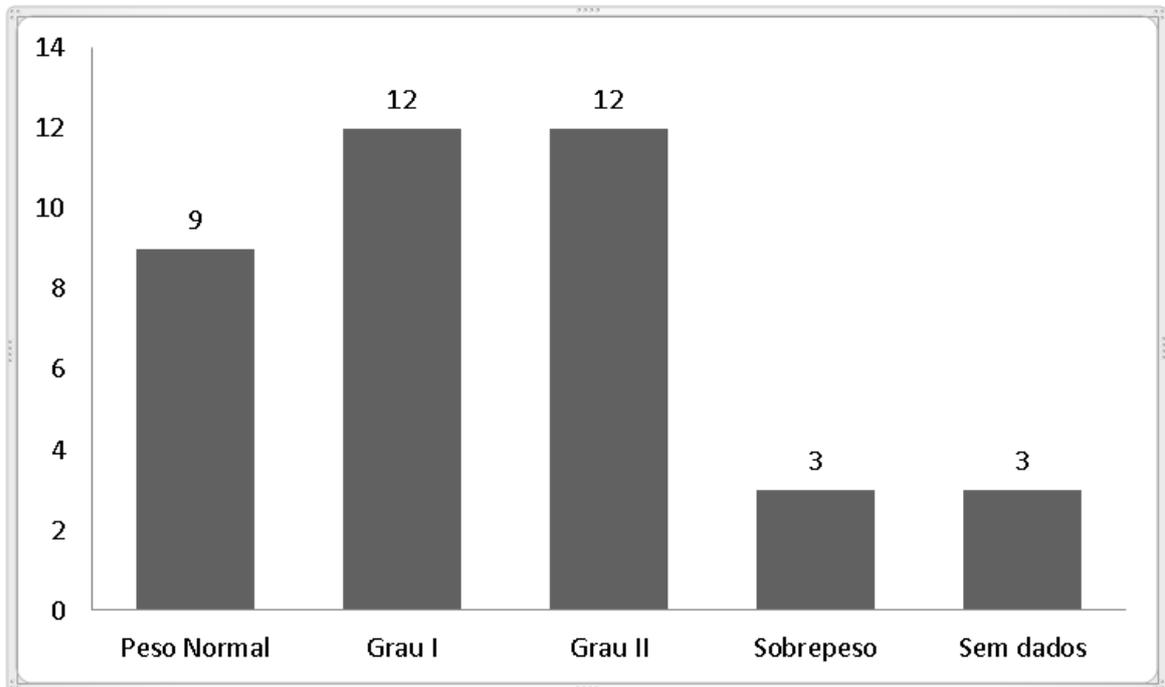
De fato, a síndrome metabólica tem sido associada à aterosclerose carotídea progressiva, bem como à aterosclerose coronariana. Por outro lado, indivíduos com síndrome metabólica tendem a apresentar disfunção endotelial sistêmica e inflamação subclínica crônica, que são cada vez mais reconhecidos como poderosos fatores de risco para eventos cardíacos e cerebrovasculares.

**Gráfico 1. Distribuição percentual de condições de saúde de acordo com Relatório Operacional do Prontuário Eletrônico do Cidadão.**



**Fonte:** PEC (2021).

**Gráfico 2: Classificação da obesidade segundo o IMC dos pacientes diagnosticados com Síndrome Metabólica.**



Fonte: PEC, 2021.

**Tabela 1: Classificação da obesidade segundo a OMS a partir do IMC**

IMC	Classificação
< 18,5	Magreza
18,5 – 24,9	Saudável
25,0 – 29,9	Sobrepeso
30,0 – 34,9	Obesidade Grau I
35,0 – 39,9	Obesidade Grau II (severa)
? 40,0	Obesidade Grau III (morbida)

Fonte: WHO (2000).

## CONCLUSÃO

Os resultados obtidos nesta pesquisa retrataram o perfil de condições de saúde e problemas ativos dos pacientes atendidos na Unidade Básica de Saúde em Araguaína - TO. A observação das condições de saúde dos pacientes demonstrou relação entre as doenças crônicas e outras comorbidades associadas. A correlação entre o diagnóstico de SM segundo os critérios do National Cholesterol Education Program's Adult Treatment Panel III (NCEP-ATP III) e da I Diretriz Brasileira de Diagnóstico e Tratamento da Síndrome Metabólica, revela que o cálculo de risco cardiovascular e o manejo adequado dos fatores de riscos associados a SM são importantes ferramentas na prevenção de Doenças Cardiovasculares e na melhoria das estatísticas a respeito da morbimortalidade destes pacientes.

Em relação à mineração de texto e dados, encontraram-se barreiras de caráter organizacional na linguagem de construção da página web, onde foram identificados muitos elementos da classe “hidden”, impedindo assim a captura dos seus conteúdos textuais. Espera-se que em futuras pesquisas esses empecilhos possam ser superados, para que a ciência de dados possa ser empregada visando a um avanço tecnológico para a saúde pública, possibilitando previsões, gerenciamento de riscos e muitas outras aplicabilidades possibilitando ações de impacto no manejo de diversas doenças e problemas relacionados a saúde.

Os recursos e tecnologias empregados e criados durante a escrita deste artigo, serão continuamente desenvolvidos, para que posteriormente, sejam julgados na intenção de promover inovações e melhorias para a Atenção Primária a Saúde.

## REFERÊNCIAS

ALBERTI, K. George MM; ZIMMET, Paul; SHAW, Jonathan. The metabolic syndrome - a new worldwide definition. **The Lancet**, v. 366, n. 9491, p. 1059-1062, 2005.

ALBERTI, K. G. M. M. et al. Harmonizing the metabolic syndrome: a joint interim statement of the international diabetes federation task force on epidemiology and prevention; national heart, lung, and blood institute; American heart association; world heart federation; international atherosclerosis society; and international association for the study of obesity. **Circulation**, v. 120, n. 16, p. 1640-1645, 2009.

**Murilo de Paula FLAVIO; Dario Farias NUNES; Rafaella Dias GALVÃO; Rodolfo Lima ARAÚJO. Mineração de Texto em Relatório Operacional do Prontuário Eletrônico Para Auxílio na Triagem da Síndrome Metabólica. Facit Business And Technology Journal. QUALIS B1. ISSN: 2526-4281 <http://revistas.faculdefacit.edu.br/index.php/JNT>. Out/Nov - 2021. Ed. 31; V. 2. Págs. 259-268.**

BLOOMGARDEN, Zachary T. Definitions of the insulin resistance syndrome: the 1st World Congress on the Insulin Resistance Syndrome. **Diabetes care**, v. 27, n. 3, p. 824-830, 2004.

GALLAGHER, Emily Jane; LEROITH, Derek; KARNIELI, Eddy. The metabolic syndrome—from insulin resistance to obesity and diabetes. **Endocrinology and metabolism clinics of North America**, v. 37, n. 3, p. 559-579, 2008.

GUARIZI, Débora Deflim; OLIVEIRA, Eliane Vendramini. Estudo da Inteligência Artificial aplicada na área da saúde. In: **Colloquium Exactarum**. 2014. p. 26-37.

HENKIN, Leora et al. Cigarette smoking, environmental tobacco smoke exposure and insulin sensitivity: the Insulin Resistance Atherosclerosis Study. **Annals of epidemiology**, v. 9, n. 5, p. 290-296, 1999.

JOTHI, Neesha et al. Data mining in healthcare—a review. **Procedia computer science**, v. 72, p. 306-313, 2015.

KAVAKIOTIS, Ioannis et al. Machine learning and data mining methods in diabetes research. **Computational and structural biotechnology journal**, v. 15, p. 104-116, 2017.

KOSTER, Annemarie et al. The combined relations of adiposity and smoking on mortality. **The American journal of clinical nutrition**, v. 88, n. 5, p. 1206-1212, 2008.

LAKKA, Timo A. et al. Sedentary lifestyle, poor cardiorespiratory fitness, and the metabolic syndrome. **Medicine & Science in Sports & Exercise**, 2003.

LIESE, Angela D.; MAYER-DAVIS, Elizabeth J.; HAFFNER, Stefen M. Development of the multiple metabolic syndrome: an epidemiologic perspective. **Epidemiologic reviews**, v. 20, n. 2, p. 157-172, 1998.

NATIONAL CHOLESTEROL EDUCATION PROGRAM (US). EXPERT PANEL ON DETECTION; TREATMENT OF HIGH BLOOD CHOLESTEROL IN ADULTS. Third report of the National Cholesterol Education Program (NCEP) **Expert Panel on detection, evaluation, and treatment of high blood cholesterol in adults** (Adult Treatment Panel III). The Program, 2002.

RAJA, Uzma et al. Text mining in healthcare. Applications and opportunities. **J Healthc Inf Manag**, v. 22, n. 3, p. 52-6, 2008.

REAVEN, Gerald; TSAO, Philip S. Insulin resistance and compensatory hyperinsulinemia: the key player between cigarette smoking and cardiovascular disease?. **Journal of the American College of Cardiology**, v. 41, n. 6, p. 1044-1047, 2003.

SAKLAYEN, Mohammad G. The global epidemic of the metabolic syndrome. **Current hypertension reports**, v. 20, n. 2, p. 1-8, 2018.

**Murilo de Paula FLAVIO; Dario Farias NUNES; Rafaella Dias GALVÃO; Rodolfo Lima ARAÚJO. Mineração de Texto em Relatório Operacional do Prontuário Eletrônico Para Auxílio na Triagem da Síndrome Metabólica. Facit Business And Technology Journal. QUALIS B1. ISSN: 2526-4281 <http://revistas.faculdefacit.edu.br/index.php/JNT>. Out/Nov - 2021. Ed. 31; V. 2. Págs. 259-268.**

SCHILLACI, Giuseppe et al. Prognostic value of the metabolic syndrome in essential hypertension. **Journal of the American College of Cardiology**, v. 43, n. 10, p. 1817-1822, 2004.

TERAN-GARCIA, Margarita; BOUCHARD, Claude. Genetics of the metabolic syndrome. **Applied physiology, nutrition, and metabolism**, v. 32, n. 1, p. 89-114, 2007.

WENG, Stephen F. et al. Can machine-learning improve cardiovascular risk prediction using routine clinical data?. **PloS one**, v. 12, n. 4, p. e0174944, 2017.

WORLD HEALTH ORGANIZATION. The world health report 2002: reducing risks, promoting healthy life. **World Health Organization**, 2002.