( ) **RECURSOS PRÓPRIOS**

Categorias:

( ) Administração

( ) Atenção Domiciliar

( ) Qualidade Assistencial

( X ) **ATENÇÃO À SAÚDE**

Categorias:

( ) Atenção Integral à Saúde

( ) Atenção Primária à Saúde

( X ) Promoção à Saúde e Prevenção de Riscos e Doenças

( ) **REGULAÇÃO EM SAÚDE E AUDITORIA**

( X ) **TECNOLOGIA DA INFORMAÇÃO EM SAÚDE**

O trabalho completo deverá seguir as orientações:

**Título:**

**Reconhecimento de padrões no comportamento dos diabéticos da Unimed Vitória**

**Nome completo dos autores**:

Delermando Branquinho Filho[[1]](#footnote-1)

**Resumo**:

(Na língua do texto, deverá apresentar, de forma concisa, os objetivos, a metodologia e os resultados alcançados, não ultrapassando 250 palavras)

**Introdução:**

**Objetivo:**

O objetivo desse trabalho é encontrar a distância entre o comportamento de usuários diagnosticados como diabéticos e compará-los com os demais usuários, os quais, não sabemos se são diabéticos.

**Metodologia:**

Para um melhor entendimento das metodologias usadas neste projeto, segmentamos em dois procedimentos, a primeiro se refere à Ciência de Dados, onde algoritmos foram usados para determinar as componentes, do comportamento, com maior contribuição na variabilidade dos dados.

O algoritmo usado para determinar essa variabilidade foi o PCA (*Principal Component Analysis,* do inglês Análise dos Componentes Principais). Na verdade, é um procedimento matemático para converter um conjunto de observações. Muitas vezes o conjunto de variáveis, possivelmente correlacionadas, estão em um conjunto de variáveis com valores linearmente não correlacionadas, o que chamamos de componentes principais. A quantidade de componentes principais é sempre menor ou igual ao número de variáveis da amostra.

O procedimento matemático define que o primeiro componente principal tem a maior variância, ou seja, é o responsável pela maior variabilidade dos dados. Essa variabilidade decai para cada um dos componentes seguintes sob a restrição de ser ortogonal a (i.e., não correlacionado com) aos demais componentes. É importante notar que os componentes principais são garantidamente independentes apenas se os dados forem normalmente distribuídos (conjuntamente). O PCA é sensível à escala relativa das variáveis originais. Dependendo da área de aplicação, o PCA é também conhecido como transformada de Karhunen-Loève (KLT) discreta, transformada de Hotelling ou decomposição ortogonal própria (POD).

O PCA foi inventado em 1901 por Karl Pearson.[1] Agora, é mais comumente usado como uma ferramenta de Análise Exploratória de Dados e para fazer modelos preditivos. PCA pode ser feito por decomposição em autovalores (Valores Próprios) de uma matriz covariância, geralmente depois de centralizar (e normalizar ou usar pontuações-Z) a matriz de dados para cada atributo.[2] Os resultados de PCA são geralmente discutidos em termos pontuações (scores) de componentes, também chamados de pontuações de fatores (os valores de variável transformados correspondem a um ponto de dado particular), e carregamentos (loadings), i.e., o peso pelo qual cada variável normalizada original deve ser multiplicada para se obter a pontuação de componente.[3]

O PCA é a mais simples das verdadeiras análises multivariadas por autovetores (Vetores Próprios). Com frequência, sua operação pode ser tomada como sendo reveladora da estrutura interna dos dados, de uma forma que melhor explica a variância nos dados. Se visualizarmos um conjunto de dados multivariados em um espaço de alta dimensão, com 1 eixo por variável, o PCA pode ser usado para fornecer uma visualização em dimensões mais baixas dos mesmos dados, uma verdadeira "sombra" do objeto original quando visto de seu ponto mais informativo. Isto é feito usando-se apenas os primeiros componentes principais, de forma que a dimensionalidade dos dados transformados é reduzida.

O PCA é fortemente ligado à análise de fatores (Factorial Analysis); de fato, alguns pacotes estatísticos propositadamente confluem as técnicas. A verdadeira análise de fatores faz suposições diferentes sobre a estrutura subjacente dos dados e encontra os autovetores de uma matriz levemente diferente.

**Resultados/Discussão:**

**Conclusão:**

**Referências Bibliográficas:**

Elemento obrigatório, constitui uma lista ordenada dos documentos efetivamente citados no texto. As referências têm espaçamento simples e duplo entre si, são apresentadas em ordem alfabética, por autor, e alinhadas somente à margem esquerda

1. Físico pela PUC-Goiás com MBAE em Gestão e Economia Empresarial na Universidade Federal do Rio de Janeiro, Mestre pela Universidade Federal de Goiás e Assessor de Informações Corporativas na Unimed Vitória, Vitória, ES, Brasil. E-mail: <dfilho@unimedvx.com.br> [↑](#footnote-ref-1)