



UNIVERSIDADE
FEDERAL
DE PERNAMBUCO



**Centro de
Informática**
UFPE

Universidade Federal de Pernambuco

Centro de Informática - CIn

Predição de Readmissão Hospitalar para Pacientes com Diabetes
Business Understanding

Grupo:

**Matheus Henrique
Gilberto Medeiros
Vinícius Barbosa
Walmir Bispo**

Professor:

Leandro Maciel Almeida

Abril de 2023

Conteúdo

1	Introdução	2
2	Descrição dos dados	2
3	Business Understanding	2
3.1	Restrições de Projeto	3
3.2	Determinar Objetivos de Negócio	3
3.3	Avaliar a Situação	3
3.4	Determinar Objetivos	4
3.5	Plano do Projeto	4

1 Introdução

Diabetes é uma doença crônica em que a pessoa sofre de um nível prolongado de glicose no sangue. A diabetes é afetada pela altura, raça, gênero, idade, mas uma das principais razões é considerada a concentração de açúcar. Uma análise atual de um grande banco de dados clínicos foi realizada para examinar padrões históricos de cuidados com diabetes em pacientes admitidos em um hospital dos Estados Unidos e para informar direções futuras que possam levar a melhorias na segurança do paciente. A redução de readmissões hospitalares precoces é uma prioridade política com o objetivo de melhorar a qualidade dos cuidados de saúde. Neste estudo de caso, veremos como machine learning pode ajudar a reduzir os problemas causados por readmissões.

2 Descrição dos dados

O conjunto de dados representa 10 anos (1999-2008) de cuidados clínicos em 130 hospitais e redes de prestação de serviços integrados nos EUA [1]. Inclui mais de 50 recursos que representam resultados do paciente e do hospital. As informações foram extraídas do banco de dados para encontros que atendiam aos seguintes critérios:

1. É um encontro de internação (uma admissão hospitalar)
2. É um encontro diabético, ou seja, durante o qual qualquer tipo de diabetes foi registrado no sistema como diagnóstico
3. O tempo de internação foi de pelo menos 1 dia e no máximo 14 dias
4. Testes laboratoriais foram realizados durante o encontro
5. Medicamentos foram administrados durante o encontro.

Os dados contêm atributos como raça, gênero, idade, tempo no hospital, especialidade médica do médico admitente, número de testes laboratoriais realizados, resultado do teste HbA1c, diagnóstico, número de medicamentos, medicamentos para diabetes, número de pacientes ambulatoriais, pacientes internados, etc.

Todo o problema é sobre classificar os pacientes em três classes:

1. Sem readmissão
2. Readmissão <30 dias: Altamente correlacionada com a qualidade do atendimento do centro de saúde.
3. Readmissão > 30 dias: Possui menos correlação com a qualidade do atendimento do centro de saúde e pode ser devido a outros fatores externos.

3 Business Understanding

As readmissões podem ser classificadas em três categorias: inevitáveis, planejadas e não planejadas. As readmissões inevitáveis são altamente previsíveis, devido à condição do paciente (câncer, metástase). Em segundo lugar, as readmissões planejadas são diretamente prescritas pelo profissional de saúde ao paciente (ou seja, check-up). Por último, as não planejadas são definidas como readmissões que não deveriam ter ocorrido, dada a avaliação do profissional de saúde, e poderiam ter sido evitadas se o cuidado adequado fosse dado ao paciente após a alta. As readmissões inevitáveis e planejadas já são altamente antecipadas. No entanto, prever as readmissões não planejadas é de grande interesse devido à sua incerteza inerente.

Readmissões não planejada é o tipo mais útil ao avaliar a qualidade do atendimento de um hospital, uma vez que destaca erros de diagnóstico ou tratamento do profissional. Além de ser um indicador central da qualidade do atendimento, readmissões não planejadas também constituem um problema financeiro para as nações.

Estima-se que 9.3% da população dos Estados Unidos tenham diabetes, sendo que 28% desses casos não foram diagnosticados. A taxa de readmissão de pacientes diabéticos em 30 dias varia de 14,4% a 22,7%. Estima-se que as taxas de readmissão além de 30 dias após a alta hospitalar sejam ainda mais altas, com mais de 26% dos pacientes diabéticos sendo readmitidos dentro de 3 meses e 30% dentro de 1 ano [2]. Os custos associados à hospitalização de pacientes diabéticos nos EUA foram de US\$ 124 bilhões, dos quais cerca de US\$ 25 bilhões foram atribuídos a readmissões em 30 dias, assumindo uma taxa de 20% de readmissão.

3.1 Restrições de Projeto

Interpretabilidade - A interpretabilidade é sempre importante no domínio da saúde. Se o modelo prevê que um paciente será readmitido, mas não consegue explicar por que chegou a essa conclusão, o médico ficará sem saber o motivo dessa decisão e também não poderá explicar ao paciente por que ele precisa ser readmitido. Isso pode criar muitos inconvenientes tanto para o médico quanto para o paciente.

A latência não é estritamente importante - A maioria das aplicações relacionadas à saúde não é dependente de latência.

O custo de classificação incorreta é alto - Se o paciente que não precisa ser readmitido for classificado pelo modelo como "sim para readmissão", isso colocará uma carga financeira sobre o paciente. Se o paciente precisar ser readmitido, mas o modelo classificar como "não para readmissão", isso causará custos de readmissão para o hospital. Portanto, a taxa de classificação incorreta deve ser o mais baixa possível.

3.2 Determinar Objetivos de Negócio

O objetivo final deste projeto é propor um modelo de previsão para readmissões não planejadas de pacientes diabéticos em hospitais dos Estados Unidos dentro de 30 dias após a alta hospitalar. A análise se baseará no dataset público [1]. Em termos mais práticos, o estudo pode permitir que os centros de saúde antecipem e abordem melhor as readmissões não planejadas, ao mesmo tempo em que melhora a qualidade do atendimento e a eficiência de custos.

3.3 Avaliar a Situação

O conjunto de dados **Diabetes 130-US hospitals for years 1999-2008** contém informações de pacientes diabéticos que foram tratados em 130 hospitais nos EUA entre 1999 e 2008. O conjunto de dados contém 50 recursos (colunas) que incluem informações como idade, sexo, etnia, diagnóstico, tratamento, resultados e outros dados médicos relacionados. As informações foram coletadas de forma retrospectiva a partir de registros hospitalares.

Algumas das suposições e restrições deste projeto incluem:

1. Os dados coletados podem não ser representativos de todos os pacientes diabéticos
2. Os dados são limitados a hospitais nos EUA entre 1999 e 2008
3. A precisão das informações coletadas depende da qualidade dos registros hospitalares

Os riscos e contingências deste projeto incluem:

1. Os dados podem conter informações sensíveis do paciente e precisam ser protegidos adequadamente.
2. A análise dos dados pode revelar informações que possam levar a decisões inadequadas na área médica.

3.4 Determinar Objetivos

Os objetivos do projeto incluem :

1. Identificar correlações entre diferentes fatores e a admissão de pacientes diabéticos em hospitais.
2. Identificar os tratamentos mais eficazes para pacientes com diabetes com base nas informações disponíveis.
3. Identificar padrões e tendências nos dados que possam ser usados para melhorar o tratamento e prevenção de diabetes.
4. Prever o tempo de internação dos pacientes diabéticos com base em informações disponíveis.

Os critérios de sucesso do projeto incluem:

1. Identificar as variáveis mais relevantes para prever a admissão de pacientes diabéticos em hospitais.
2. Identificar os tratamentos mais eficazes para pacientes com diabetes com base em evidências científicas.
3. Desenvolver modelos preditivos precisos para prever o tempo de internação dos pacientes diabéticos.

3.5 Plano do Projeto

O plano para alcançar os objetivos do projeto incluem as seguintes etapas:

1. Análise exploratória dos dados para entender as características dos pacientes diabéticos e identificar possíveis correlações entre diferentes variáveis
2. Análise estatística para identificar os tratamentos mais eficazes para pacientes com diabetes e prever o tempo de internação dos pacientes.
3. Desenvolvimento de modelos preditivos utilizando técnicas de aprendizado de máquina e validação cruzada para garantir

Referências

- [1] “UCI Machine Learning Repository: Diabetes 130-US hospitals for years 1999-2008 Data Set — archive.ics.uci.edu.” <https://archive.ics.uci.edu/ml/datasets/Diabetes+130-US+hospitals+for+years+1999-2008>. [Accessed 05-Apr-2023].
- [2] “Statistics About Diabetes | ADA — diabetes.org.” <https://diabetes.org/about-us/statistics/about-diabetes>. [Accessed 05-Apr-2023].