

Competição 1

Predição de cobertura de um plano de saúde

Alexandre da Silva¹

Enzo Cunha²

1. alexandre_silva_costa@discente.ufg.br
2. enzcdn@discente.ufg.br

Material elaborado em parceria com a
professora Nádia Felix

2022


INF

INSTITUTO DE
INFORMÁTICA



Sumário

- ▶ **Contextualização e descrição do problema**
- ▶ **Descrição do conjunto de dados**
- ▶ **Algoritmos utilizados**
- ▶ **Avaliação do modelo**



Contextualização e Descrição do problema

Contextualização

- ▶ Operadora
- ▶ Prestador
- ▶ Beneficiário
- ▶ Auditor

Descrição do Problema

- ▶ Auditor gera custo administrativo
- ▶ Sistemas atuais não fazem previsão
- ▶ Criar um modelo de predição de cobertura de um plano de saúde



Descrição do conjunto de dados

Conjunto de Dados

- ▶ Mais de 200 mil instâncias e 31 tipos de atributos
- ▶ Atributos relevantes sem dados foram preenchidos com valor zero

Atributos relevantes

Figura 1: Atributos Relevantes para a predição.

COLUNA	DESCRIÇÃO	OBSERVAÇÃO
QT_TEMPO_DOENCA	TEMPO DE DOENÇA	CAMPO NUMÉRICO
CD_GUIA_REFERENCIA		CAMPO NUMERICO
DS_UNIDADE_TEMPO_DOENCA	UNIDADE DE TEMPO DA DOENÇA	DIAS, MESES OU ANOS (DE ACORDO COM O CAMPO QT_TEMPO_DOENCA)
DS_TIPO_INTERNACAO	TIPO DE INTERNAÇÃO	CLÍNICA, CIRÚRGICA, OBSTÉTRICA, PSIQUIÁTRICA ETC.
DS_CARATER_ATENDIMENTO	CARÁTER DE ATENDIMENTO	ELETIVO OU URGENTE
DS_TIPO_ATENDIMENTO	TIPO DE ATENDIMENTO	TIPO DE ATENDIMENTO DAS GUIAS SADT: EXAMES, PRONTO SOCORRO, QUIMIOTERAPIA ETC.
DS_INDICACAO_ACIDENTE	INDICAÇÃO DE ACIDENTE	ACIDENTE DE TRABALHO, TRÂNSITO, OUTROS OU NÃO ACIDENTE
DS_TIPO_PREST_SOLICITANTE	HOSPITAL, CLINICA DE IMAGEM, CLINICA, MEDICO	
DS_TIPO_GUIA	DESCRIÇÃO DO TIPO DE GUIA	SP/SADT: EXAMES E TERAPIAS INTERNAÇÃO: CIRURGIAS OU INTERNAÇÕES CLÍNICAS PRORROGAÇÃO: COMPLEMENTO DA INTERNAÇÃO



Algoritmos utilizados

Algoritmos utilizados

▶ Naive Bayes

- Simples de implementar
- Rápido
- Aceita dados categóricos

▶ Random Forest

- Pode lidar com dados faltando
- Robusto a outliers



Evolução do modelo

Evolução do modelo

- ▶ **Encoder**
- ▶ **Select K Best**
- ▶ **Random Forest**
- ▶ **Grid Search**
- ▶ **Comparação com notebook de exemplo**



Avaliação do modelo

Avaliação no conjunto de validação

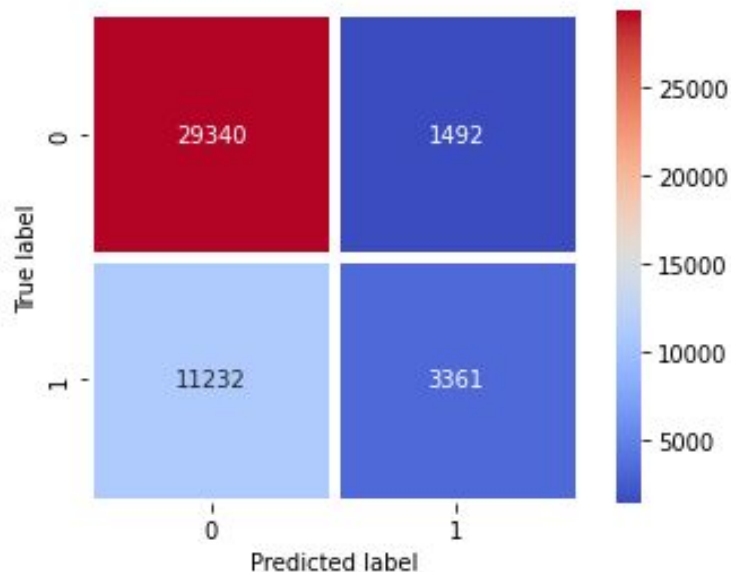


Figura 2: Matriz de confusão

	precision	recall	f1-score	support
Autorizado	0.72	0.95	0.82	30832
Negado	0.69	0.23	0.35	14593
accuracy			0.72	45425
macro avg	0.71	0.59	0.58	45425
weighted avg	0.71	0.72	0.67	45425

Figura 3: Scores

Avaliação no conjunto de teste

Figura 3: Primeira tentativa válida



Figura 4: Última entrega



Obrigado!

INF
INSTITUTO DE
INFORMÁTICA



UFG
UNIVERSIDADE
FEDERAL DE GOIÁS