

# Resultados observados

Tabela 1: Desempenho sem PCA

Resultados da Atividade 2, utilizando os dados sem aplicação de PCA. Para os classificadores DMC, 1-NN e MaxCorr, foi selecionada a normalização que apresentou o melhor desempenho médio, no caso, sem normalização.

	Média	Mínimo	Máximo	Mediana	Desvio Padrão	Matrizes Singulares	Tempo (s)
Quadrático	6.24	0.00	21.21	6.06	3.85	750	25.21
Variante 1(Tikhonov)	78.67	63.64	90.91	78.79	6.72	0	9.28
Variante 2(Pooled)	6.42	0.00	27.27	6.06	5.48	0	13.33
Variante 3(Friedman)	6.73	0.00	21.21	6.06	4.08	2	10.50
Variante 4(Naive Bayes)	7.39	0.00	21.21	6.06	3.84	716	22.31
MaxCorr	76.85	63.64	90.91	75.76	7.37	0	0.02
DMC	75.88	57.58	90.91	75.76	7.52	0	0.03
1-NN	78.18	60.61	90.91	78.79	7.30	0	0.08

Questão 1

O que se pode concluir sobre os desempenhos dos classificadores avaliados?

Os classificadores quadráticos (**Quadrático**, **Variante 2**, **Variante 3**, **Variante 4**) apresentaram, em sua maioria, um desempenho severamente comprometido. Isso ocorre devido ao número de atributos ser muito superior ao número de amostras por classe, tornando a estimativa da matriz de covariância instável.

Notavelmente, a **Variante 1 (Tikhonov)** foi a única exceção entre os modelos quadráticos, demonstrando que sua forma de regularização foi eficaz.

Em contrapartida, os classificadores que não dependem diretamente da inversão da matriz de covariância por classe, como **DMC**, **1-NN** e **MaxCorr**, mostraram-se muito mais robustos a este cenário, alcançando acurácias significativamente superiores e, em torno de 75% a 78%.

Questão 2

Qual deles teve o melhor desempenho em relação à taxa de acerto? E em relação ao tempo?

Com base nos resultados da Tabela 1:

- **Taxa de Acerto:** O classificador com o melhor desempenho médio foi a **Variante 1 (Tikhonov)**.
- **Tempo de Execução:** O classificador mais eficiente computacionalmente foi o **MaxCorr**.

Questão 3

Houve problemas de inversão das matrizes de covariância? Se sim, para quais classificadores? Este problema foi contornado por alguma das variantes avaliadas? Se sim, descreva sucintamente o mecanismo usado para resolvê-lo.

**Sim**, houve problemas de inversão de matriz. Conforme a coluna "Matrizes Singulares", o problema ocorreu para o **Classificador Quadrático** e para a **Variante 4 (Naive Bayes)**.

Isso ocorreu porque, sem regularização adequada, a matriz de covariância ( $C_i$ ) demonstrou-se singular. Na implementação, caso a tentativa de inversão direta (`np.linalg.inv`) falhe, para que o algoritmo pudesse prosseguir e gerar um resultado, foi utilizado um fallback para a **pseudo-inversa** (`np.linalg.pinv`), o que permitiu a conclusão do teste, ainda que com baixa acurácia.

As demais variantes contornaram o problema com sucesso através dos seguintes mecanismos de regularização:

- **Variante 1 (Tikhonov):** Adiciona um múltiplo da matriz identidade ( $\lambda I_p$ ) à matriz de covariância ( $C_i + \lambda I_p$ ), onde  $0 \leq \lambda \ll 1$ . Isso aumenta os valores da diagonal principal, tornando a matriz mais estável e garantindo sua invertibilidade.
- **Variante 2 (Pooled):** Utiliza uma única matriz de covariância agregada ( $C_{pool}$ ), calculada como uma soma ponderada das covariâncias de todas as classes. Essa matriz é estimada com um número muito maior de amostras, tornando-a mais robusta e menos propensa à singularidade.
- **Variante 3 (Friedman):** Realiza uma combinação entre a matriz de covariância individual de cada classe ( $C_i$ ) e a matriz agregada ( $C_{pool}$ ).

Tabela 2: Desempenho com PCA (Sem Redução)

Resultados da Atividade 4, com aplicação de PCA sem redução de dimensionalidade.

	Média	Mínimo	Máximo	Mediana	Desvio Padrão	Matrizes Singulares	Tempo (s)
Quadrático	6.24	0.00	21.21	6.06	3.85	750	2.65
Variante 1(Tikhonov)	78.67	63.64	90.91	78.79	6.72	0	1.13
Variante 2(Pooled)	6.24	0.00	21.21	6.06	3.85	750	6.53
Variante 3(Friedman)	6.24	0.00	21.21	6.06	3.85	750	2.92
Variante 4(Naive Bayes)	6.24	0.00	21.21	6.06	3.85	750	1.94
MaxCorr	76.85	63.64	90.91	75.76	7.37	0	0.01
DMC	75.88	57.58	90.91	75.76	7.52	0	0.02
1-NN	78.18	60.61	90.91	78.79	7.30	0	0.05

Questão 4

**(i) O que se pode concluir sobre os desempenhos? Houve alguma mudança? (ii) O classificador quadrático e a Variante 4 tiveram desempenho equivalente?**

(i) A aplicação de PCA apenas para descorrelacionar os atributos teve um impacto geral negativo no desempenho da maioria das variantes de regularização (Variantes 2 e 3). A estabilidade dessas variantes, que dependem da combinação de covariâncias, foi comprometida, como evidenciado pelo surgimento de matrizes singulares.

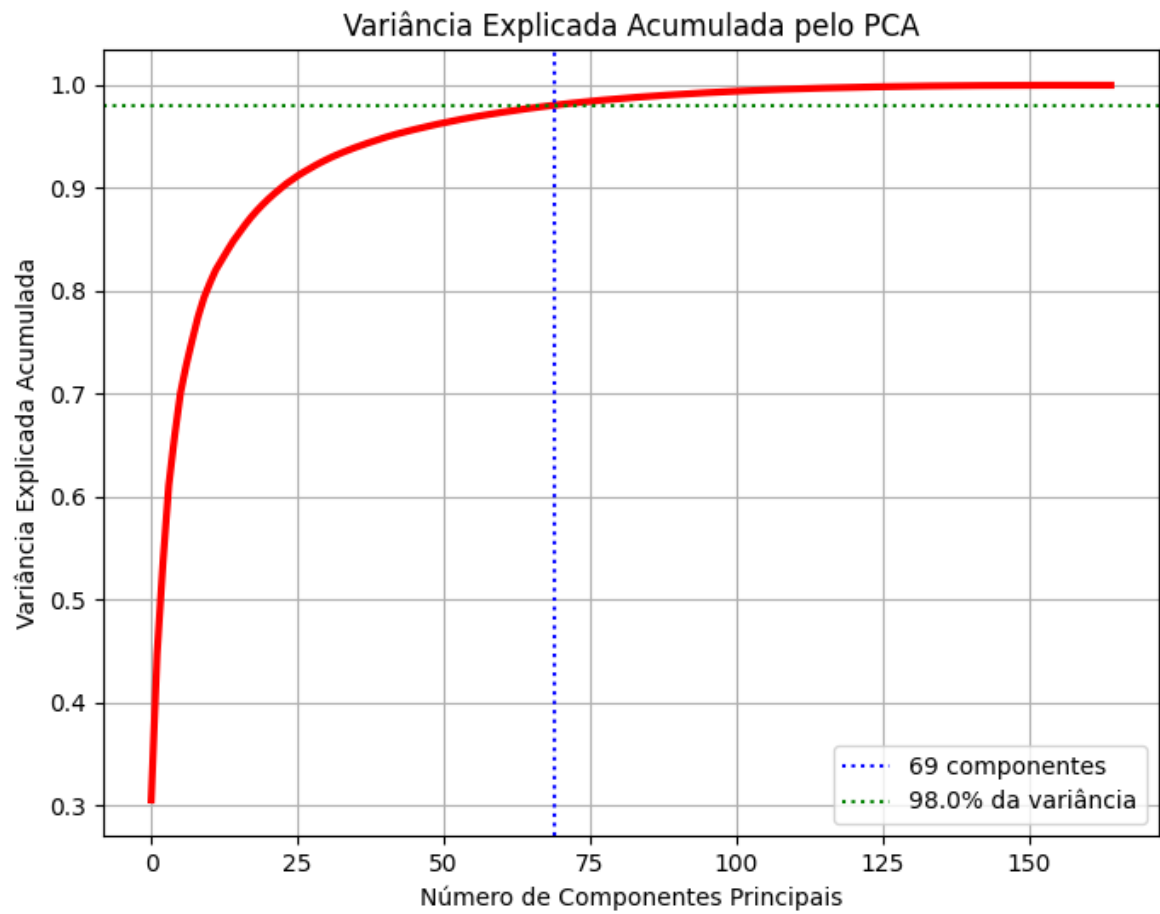
(ii) Sim, os classificadores Quadrático e Variante 4 (Naive Bayes) apresentaram desempenhos idênticos.

Tabela 3: Desempenho com PCA (Com Redução)

Questão 5

Qual foi a dimensão de redução **q** escolhida, de modo a preservar 98% da informação?

Com base no gráfico da variância explicada acumulada, o número de componentes principais (**q**) necessário para preservar pelo menos 98% da variância original dos dados é **q = 69**.



Resultados da Atividade 6, com aplicação de PCA com redução de dimensionalidade para **q=69**.

	Média	Mínimo	Máximo	Mediana	Desvio Padrão	Matrizes Singulares	Tempo (s)
Quadrático	6.79	0.00	21.21	6.06	4.00	0	0.26
Variante 1(Tikhonov)	79.03	66.67	90.91	78.79	6.35	0	0.22
Variante 2(Pooled)	92.36	84.85	100.00	90.91	4.46	0	0.26
Variante 3(Friedman)	91.94	81.82	100.00	90.91	4.64	0	0.28
Variante 4(Naive Bayes)	70.06	57.58	84.85	69.70	6.52	0	0.16
MaxCorr	76.91	63.64	90.91	75.76	7.27	0	0.01
DMC	76.00	57.58	90.91	75.76	7.47	0	0.01
1-NN	77.09	60.61	87.88	78.79	7.18	0	0.03

Questão 6

O que se pode concluir sobre os desempenhos com a realização da redução de dimensionalidade via PCA?

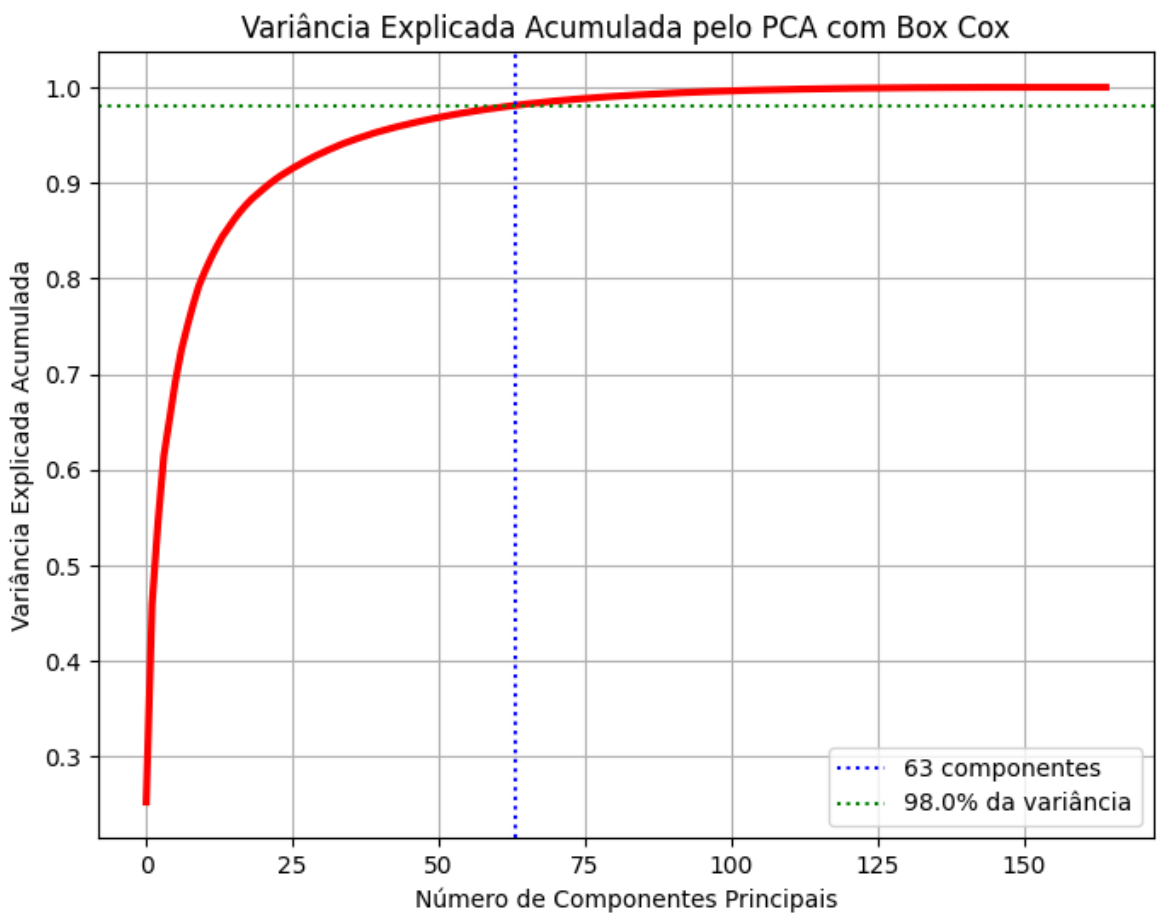
A redução de dimensionalidade via PCA para  $q=69$  resultou em uma **melhora de desempenho substancial e generalizada**.

- **Melhora e Estabilidade:** Todas as variantes quadráticas tiveram um ganho de performance, com acurácias médias relativamente altas e a eliminação de matrizes singulares.
- **Eficiência:** Houve também uma redução significativa no tempo de execução dos modelos mais complexos, pois os cálculos matriciais foram realizados em um espaço de dimensão muito menor.
- **Piora Mínima:** Apenas o classificador **1-NN** apresentou uma leve queda no desempenho (de 78.18% para 77.09%). Isso pode ocorrer pois a redução, embora preserve a maior parte da variância, pode descartar informações sutis que eram úteis para a sua fronteira de decisão.

Tabela 4: Desempenho com Box-Cox e PCA

Questão 7

Houve alguma mudança nos desempenhos com a aplicação da transformação BOX-COX juntamente com PCA?



Sim, houve uma melhora geral no desempenho. Primeiramente, a aplicação da transformação Box-Cox antes do PCA resultou em uma representação de dados ainda mais compacta, reduzindo o número de componentes necessários para 98% de variância para  $q = 63$ .

Tabela com os resultados da Atividade 7, após aplicação da transformação Box-Cox e PCA com redução de dimensionalidade.

	Média	Mínimo	Máximo	Mediana	Desvio Padrão	Matrizes Singulares	Tempo (s)
Quadrático	6.67	0.00	21.21	6.06	4.37	0	0.24
Variante 1(Tikhonov)	87.58	75.76	93.94	87.88	4.50	0	0.21
Variante 2(Pooled)	96.00	84.85	100.00	96.97	3.50	0	0.24
Variante 3(Friedman)	96.06	84.85	100.00	96.97	3.39	0	0.25
Variante 4(Naive Bayes)	70.42	48.48	87.88	69.70	8.10	0	0.15
MaxCorr	84.12	72.73	93.94	84.85	4.27	0	0.01
DMC	84.06	72.73	93.94	84.85	5.65	0	0.01
1-NN	88.91	75.76	96.97	89.39	5.13	0	0.03

Observando a Tabela 4, a transformação Box-Cox trouxe um ganho adicional de acurácia para todos os classificadores, com exceção do Quadrático puro. A variante 3 foi a que apresentou o melhor resultado.

Questão 8

- Índices de desempenho para o primeiro modelo:

	Média	Mínimo	Maximo	Mediana	Desvio Padrão
acuracia (%)	5.72	0.00	13.89	5.56	3.29
taxa_falsos_positivos (%)	100.00	100.00	100.00	100.00	0.00
taxa_falsos_negativos (%)	0.00	0.00	0.00	0.00	0.00
sensibilidade (%)	98.00	0.00	100.00	100.00	14.14
precisao (%)	5.72	0.00	13.89	5.56	3.29

- Índices de desempenho para o segundo modelo:

	Média	Mínimo	Maximo	Mediana	Desvio Padrão
acuracia (%)	70.22	50.00	88.89	69.44	8.75
taxa_falsos_positivos (%)	26.06	8.82	48.57	25.71	8.88
taxa_falsos_negativos (%)	87.33	0.00	100.00	100.00	29.22
sensibilidade (%)	10.67	0.00	100.00	0.00	26.41
precisao (%)	1.91	0.00	20.00	0.00	4.44

Podemos concluir que, o primeiro modelo demonstrou uma alta capacidade de detecção da classe intrusa, alcançando uma sensibilidade de 98% e uma taxa de falsos negativos de 0%, cumprindo com eficácia sua principal função de segurança. Contudo, essa segurança foi obtida através de uma estratégia restritiva, que resultou em uma taxa de falsos positivos de 100%, rejeitando todos os usuários autorizados e comprometendo a aplicação prática do sistema. Por outro lado, o segundo modelo apresentou uma acurácia geral superior e uma taxa de falsos positivos mais baixa, mostrando maior flexibilidade ao conceder acesso à maioria dos usuários autorizados. Entretanto, essa usabilidade veio com um custo de segurança, evidenciado pela alta taxa de falsos negativos, indicando que o modelo falhou em detectar o intruso na maioria das ocasiões.