

Relatório do trabalho computacional 2

Inteligência artificial e Sistemas inteligentes

Alunos:

Pedro Lucas Sousa Barreto, 2220318.

Professor: Paulo Cirilo Souza Barbosa.

Data: Sábado dia 4 de novembro de 2023 (17/09/2023).

Turma (IA): T296-16/17.

Turma (SI): T951-09/19.

1. A tarefa em mãos.

a. Resumo.

O artigo em mãos visa fazer uma análise pragmática e estatística de um conjunto de modelos de inteligência artificial com intuito de medir suas capacidades tanto preditoras quanto de classificação para um dado conjunto de dados, a fim de determinar quais modelos melhor se adequa ao mesmo, para que possamos fazer hipóteses sobre qual modelo teria a melhor performance pertinente a tarefas adjacentes de predição ou classificação.

b. Introdução.

Este artigo tem como objetivo realizar uma análise pragmática e estatística de três modelos de inteligência artificial: Perceptron Simples e Adaline. A análise visa medir suas capacidades tanto em termos de predição quanto de classificação, aplicando-os a conjuntos de dados específicos. O objetivo é determinar qual desses modelos se adequa melhor aos conjuntos de dados em questão, possibilitando a geração de hipóteses sobre o desempenho mais adequado para tarefas de predição e classificação.

Modelos de regressão, como Perceptron Simples e Adaline, são utilizados para prever valores futuros com base em cálculos regressores que estabelecem relações entre variáveis de entrada (regressoras) e variáveis dependentes (valores previstos). A escolha do modelo adequado para um determinado conjunto de dados é uma tarefa desafiadora, exigindo uma ponderação cuidadosa das características específicas do problema em questão.

Ao longo deste artigo, exploraremos as capacidades desses modelos em três conjuntos de dados distintos, empregando-os tanto em tarefas de predição quanto de classificação. Nosso objetivo é realizar uma análise abrangente do potencial desses modelos em termos de precisão, visando minimizar erros e, assim, fornecer insights valiosos para a seleção de modelos adequados às necessidades de uma determinada atividade.

2. Desenvolvimento.

a. Pré-visualização dos dados e hipóteses

i. Dados sintetizados:

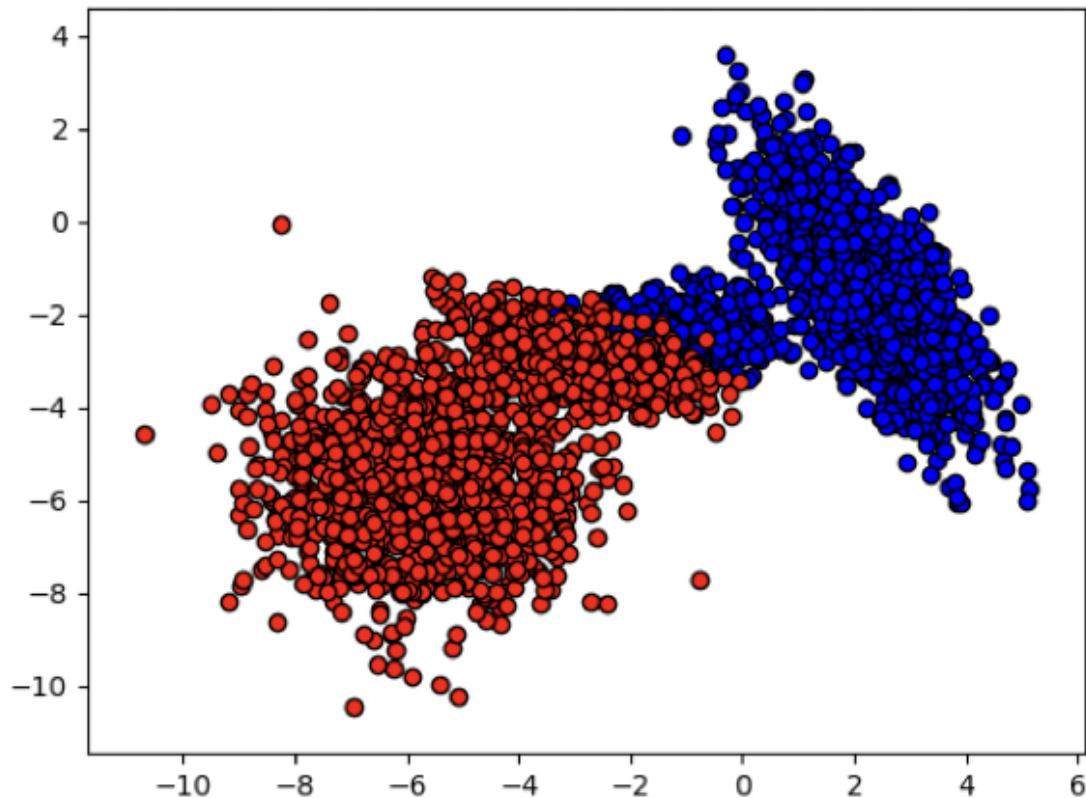


Figura 1 - Dados sintetizados. Fonte: elaborado pelos autores.

A partir da visualização desses dados, podemos observar, em primeiro lugar, que eles são adequados para modelos de Perceptron simples devido ao fato de serem linearmente separáveis. Ao observar mais detalhadamente, podemos identificar a presença de dois clusters para cada classe, sendo que os dois clusters menores são onde as classes se encontram e apresentam maior mistura. Apesar dessa mistura, ambos os modelos devem ter uma alta taxa de acurácia devido à natureza dos dados.

b. Organização dos dados e quantidade de rodadas.

i. Divisão entre conjunto de teste e conjunto de treino.

Foi escolhida a separação 80/20 (oitenta para vinte) para os conjuntos de dados fornecidos, onde 80% (oitenta por cento) dos dados foram separados para o conjunto de treino, e os 20% (vinte por cento) restantes para o teste e validação dos modelos.

ii. Embaralhamento dos dados.

As amostras presentes no conjunto de dados fornecido foram embaralhadas e aleatorizadas visto que são modelos generalistas, o que ajudará a reduzir o viés tornando-o um modelo mais robusto e capaz.

iii. Quantidade de rodadas.

Foi-se determinado que a quantidade de rodadas para treinamento dos modelos seria de 100 rodadas para cada modelo.

c. Definição dos hiperparâmetros.

i. Passo de aprendizagem (η): No caso foi escolhido um valor de 0.001 para η , isso implica que o processo de ajuste dos pesos será feito com incrementos pequenos, o que pode tornar o treinamento mais lento, porém mais estável. No entanto, o valor exato de η depende da natureza dos dados e do problema em questão, e a escolha ideal pode variar. A seleção do valor de η é uma tarefa empírica (obtida a partir de experiência) e muitas vezes requer experimentação para determinar o valor mais adequado.

ii. Precisão (ϵ): A precisão, ϵ , é uma medida da diferença entre o erro quadrático médio (EQM) em duas épocas sucessivas durante o treinamento do ADALINE. O treinamento do ADALINE é interrompido quando a diferença entre o EQM atual e o EQM anterior é menor ou igual a ϵ . A precisão escolhida foi $1e-15$, o que significa que o modelo deve atingir uma taxa de erro muito baixa para que a fase de treino seja encerrada.

iii. Máximo de epochs: O número máximo de épocas foi de 1000, mas é uma decisão empírica que define o limite de iterações durante o treinamento. Essa escolha depende do contexto do problema e da disponibilidade de tempo e recursos computacionais.

d. Modelos implementados.

i. Perceptron Simples (PS)

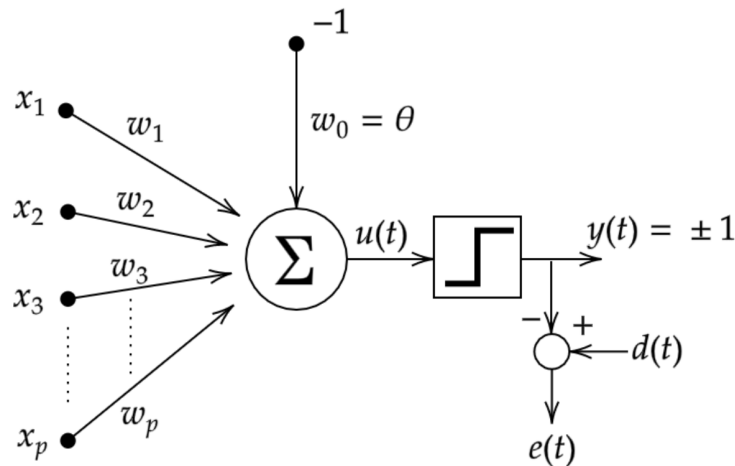


Figura 2 - arquitetura do neurônio da rede PS. Fonte: Barbosa, 2023.

O Perceptron Simples ou PS é um dos primeiros modelos de redes neurais artificiais, proposto por Frank Rosenblatt em 1958. Ele representa uma versão inicial de como os neurônios do cérebro podem ser simulados em computadores. O PS é um neurônio de McCulloch-Pitts com uma regra de aprendizagem, tornando-o capaz de aprender e realizar tarefas de classificação simples. A regra de aprendizagem é o mecanismo central que o torna inteligente, e o PS serviu como base para o desenvolvimento de redes neurais mais avançadas no futuro.

ii. ADALINE

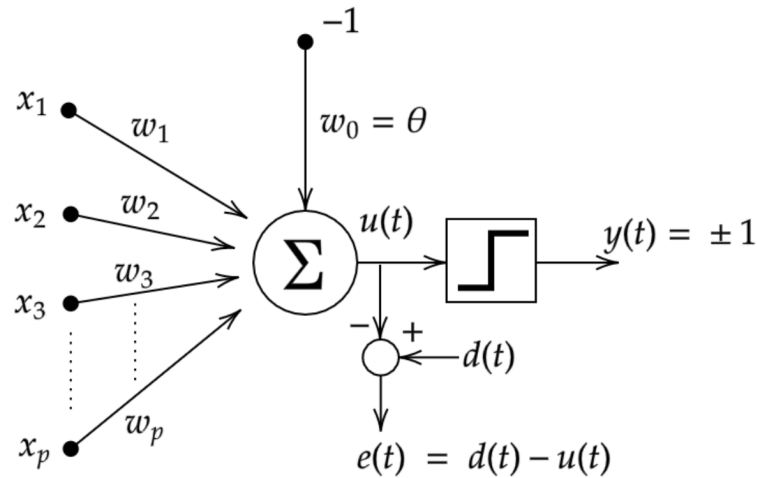


Figura 3 - arquitetura do neurônio da rede ADALINE. Fonte: Barbosa, 2023.

O Adaline (Adaptive Linear Element) é um algoritmo adaptativo proposto por Bernard Widrow e Ted Hoff em 1960. Ele é uma extensão do Perceptron original e apresenta a capacidade de ajustar seus parâmetros por meio de regras de atualização recursiva, incluindo a "Regra de Widrow-Hoff", LMS (Least Mean Squares) e a "Regra Delta". A maior diferença entre ele e o PS é a ordem que " $e(t)$ " é calculado.

iii. MLP

3. Resultados.

A avaliação de desempenho dos modelos apresentados envolverá uma análise detalhada com base em várias métricas, considerando a natureza dos dados em questão. Essas métricas serão calculadas a partir de uma matriz de confusão, uma ferramenta essencial na avaliação de modelos de classificação. A matriz de confusão é construída a partir dos seguintes valores:

- A. VP (Verdadeiro Positivo): Representa a quantidade de predições corretas para a condição positiva.
- B. VN (Verdadeiro Negativo): Indica a quantidade de predições corretas para a condição negativa real.
- C. FP (Falso Positivo): Refere-se à quantidade de predições incorretas para a condição real negativa.

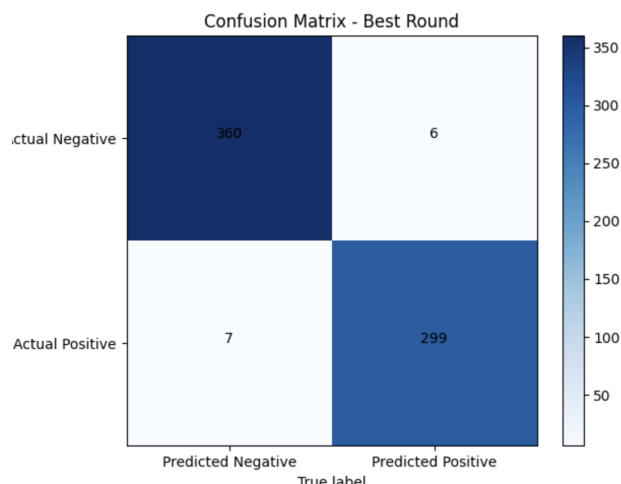
D. FN (Falso Negativo): Representa a quantidade de predições feitas de maneira errônea para a condição real positiva.

A partir dessa matriz de confusão, é possível extrair diversas medidas, das quais destacamos as seguintes:

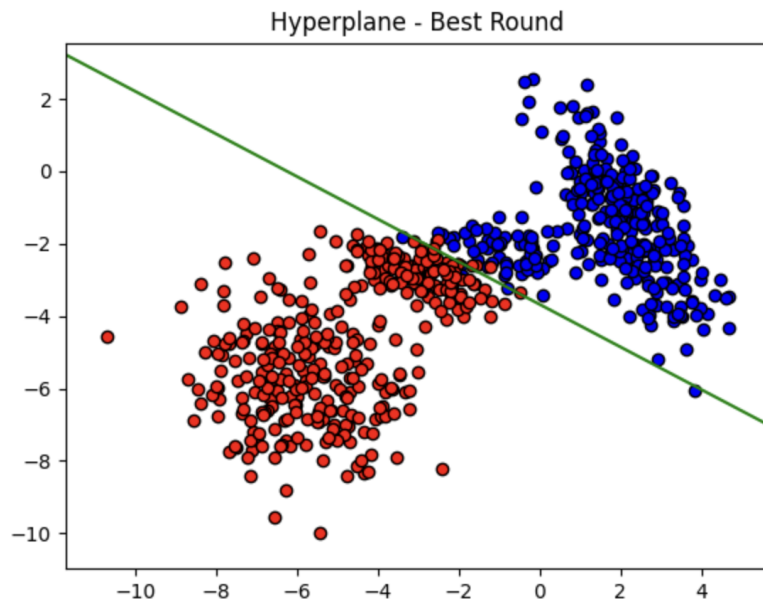
- E. Acurácia: Medida da proporção de predições corretas em relação ao total de predições. A fórmula é dada por $\text{Acurácia} = (VP + VN) / (VP + VN + FP + FN)$.
- F. Sensibilidade: Também conhecida como taxa de verdadeiros positivos, a sensibilidade é a proporção de verdadeiros positivos em relação a todos os casos reais positivos. A fórmula é $\text{Sensibilidade} = VP / (VP + FN)$.
- G. Especificidade: Refere-se à taxa de verdadeiros negativos, representando a proporção de verdadeiros negativos em relação a todos os casos reais negativos. A fórmula é $\text{Especificidade} = VN / (VN + FP)$.

1. Resultados Para o Perceptron Simples:

a. melhor round:

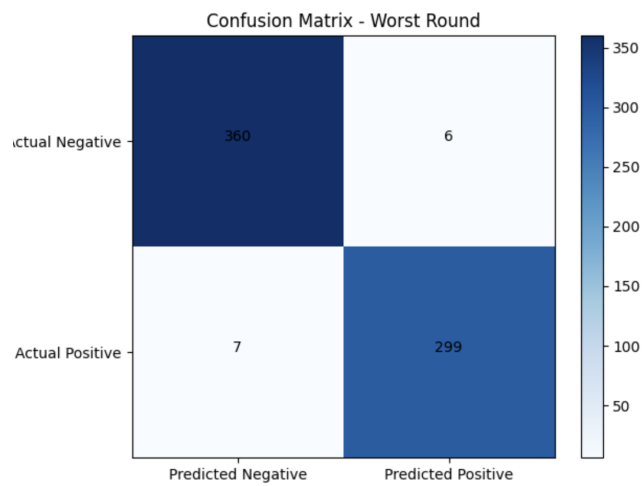


i.

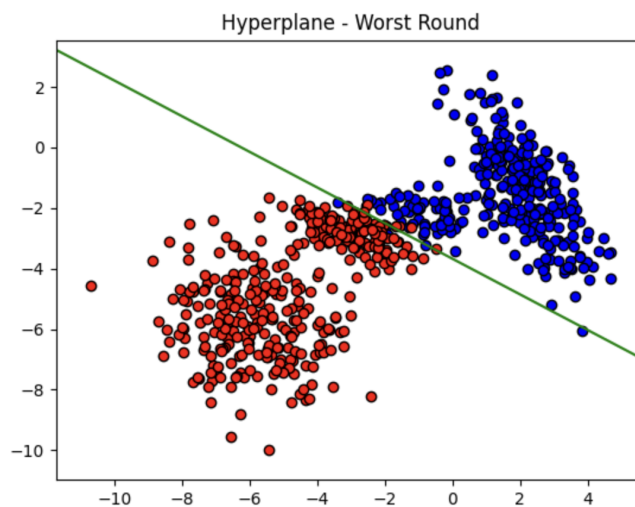


ii.

b. pior round:



i.



ii.

c. Estatísticas:

Resultados ADALINE				
	Mean	Desv. Padrão	max	min
Acuracia	98%	0	98%	98%
Sensibilidade	0.97	1.11E-16	0.97	0.97
Especificidade	0.99	1.11E-16	0.99	0.99
EPOCH	316			

i.

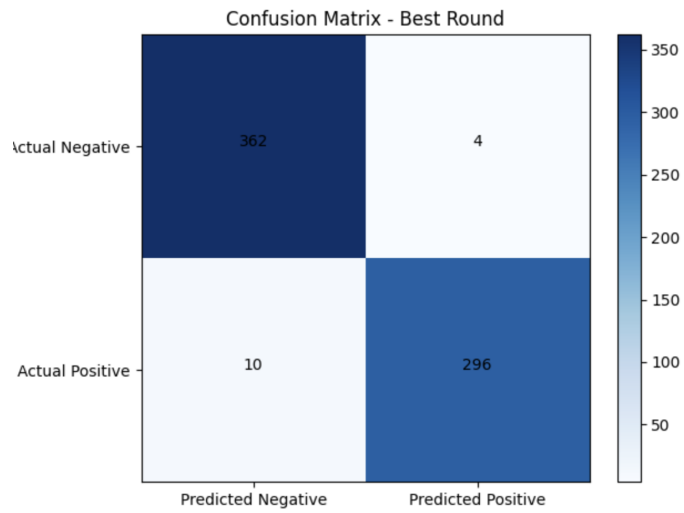
2. Resultados Para o ADALINE:

a. Estatísticas:

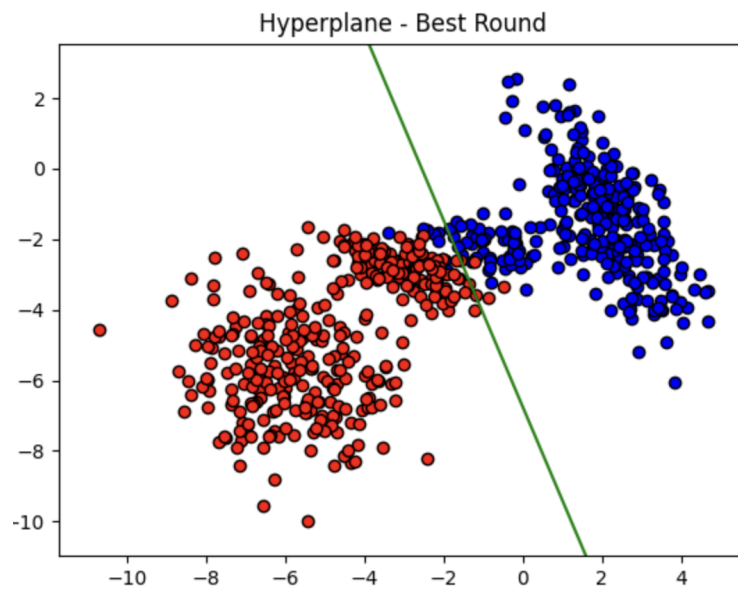
Resultados ADALINE				
	Mean	Desv. Padrão	max	min
Acuracia	98%	0	98%	98%
Sensibilidade	0.97	1.11E-16	0.97	0.97
Especificidade	0.99	1.11E-16	0.99	0.99

i.

b. melhor round:

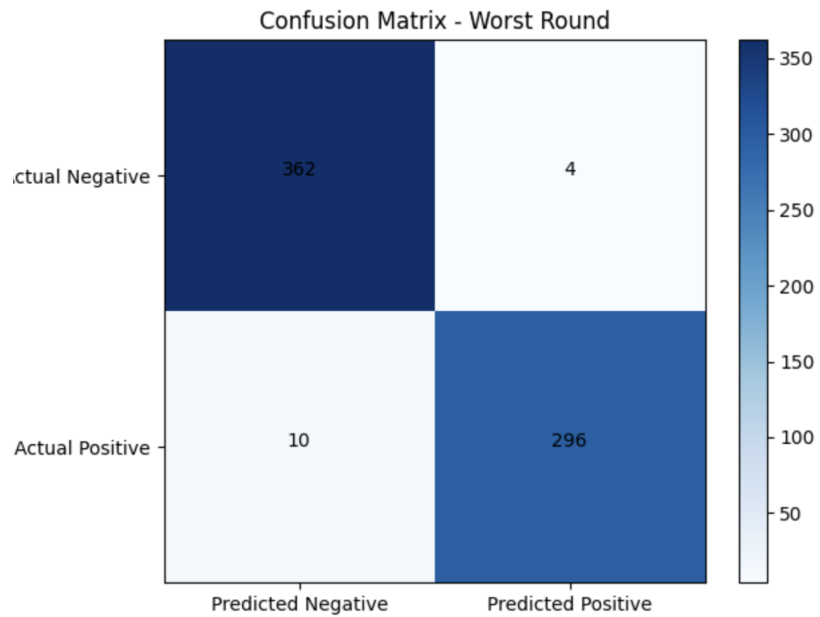


i.

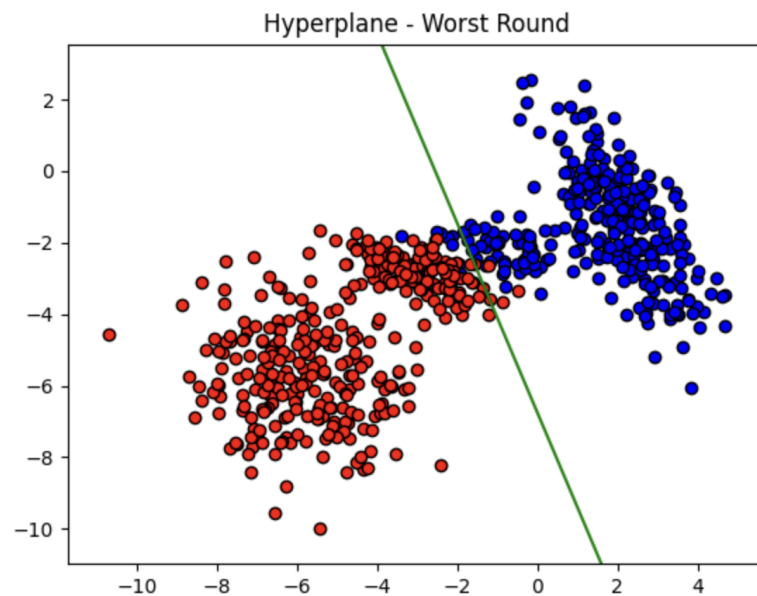


ii.

c. pior round:



i.



ii.

4. Conclusão

Após várias rodadas de experimentos, ficou claro que tanto o Perceptron quanto o Adaline apresentaram um desempenho excepcionalmente bom em tarefas de classificação. Eles foram capazes de realizar classificações com alta precisão, identificando corretamente a maioria dos casos positivos e negativos.

No entanto, é importante destacar que esses modelos são mais adequados para casos em que os dados são linearmente separáveis, ou seja, quando as classes podem ser separadas por uma linha reta. Para problemas mais complexos que envolvem fronteiras de decisão não-lineares, esses modelos podem não ser a melhor escolha.

Nossas análises demonstraram que, em cenários linearmente separáveis, o Perceptron e o Adaline são confiáveis, consistentes e capazes de alcançar altos níveis de precisão na classificação. Suas habilidades de delimitação das fronteiras de decisão entre as classes são notáveis, tornando-os ferramentas valiosas em aplicações que atendem a essa condição específica.

5. Referências.

- Barbosa, Paulo Cirillo Souza. "Inteligência Artificial Computacional - T296." Centro de Ciências Tecnológicas - CCT, Universidade de Fortaleza, Fortaleza, Ceará, Brasil, 1 de outubro de 2023.