



Aluno

- Vinícius Almeida Barros - 18/0028758

1 Apresentação

Nesse exercício, o objetivo é demonstrar como o treinamento com dados ruidosos pode aumentar a capacidade de generalização de uma rede neural artificial (RNA). Nosso conjunto de dados base são 16 caracteres, os quais podem ser vistos na figura 1. Cada caractere é representado por uma imagem binária 9x7 (63 pixels).

0 1 2 3 4 5 6 7 8 9 A B C D E F

Figura 1: Caracteres.

A esses caracteres foram adicionados diferentes níveis de ruído. Abaixo, pode-se ver o conjunto de caracteres base com 10% de ruído.

0 1 2 3 4 5 6 7 8 9 A B C D E F

Figura 2: Caracteres com 10% de ruído.

2 Método

Nesse trabalho serão criadas 18 RNAs feedforward com as mesmas condições iniciais, que receberão conjuntos de treinamento distintos. Na tabela 1, podem ser vistos tais conjuntos. O número n vai receber os seguintes valores: 1, 10 e 30 e representa o conjunto de amostras por nível de ruído. Por exemplo, serão criadas 3 RNAs tipo 5. As RNAs tipo 5 tem um conjunto de treinamento com $96n$ amostras no total. Ou seja, haverá 3 RNAs tipo 5 com 96, 960 e 2.880 amostras no total fazendo parte do conjunto de dados de treinamento ou 16, 160 e 480 amostras por nível de ruído. Vale ressaltar que todas as amostras para o nível de ruído 0 são idênticas, então realisticamente a quantidade de amostras distintas para a RNA5 é dada por $16 + 80n$. A mesma lógica se aplica às outras RNAs.

Como dito anteriormente, todas as redes neurais são criadas com as mesmas condições iniciais e todas elas com apenas três camadas na seguinte configuração: camada de entrada com 63 neurônios, camada escondida com 50 neurônios e camada de saída com 16 neurônios. O algoritmo de aprendizado utilizado foi o gradiente descendente estocástico com taxa de aprendizagem igual a 0,7.

RNA	Ruido					
	0%	10%	20%	30%	40%	50%
0	16n	0	0	0	0	0
1	16n	16n	0	0	0	0
2	16n	16n	16n	0	0	0
3	16n	16n	16n	16n	0	0
4	16n	16n	16n	16n	16n	0
5	16n	16n	16n	16n	16n	16n

Tabela 1: Conjunto de treinamento das RNAs.

3 Resultados

Como exemplo, podemos analisar evolução da acurácia da RNA5 para o conjunto de treino e o conjunto de testes para $n = 1$ e $n = 30$, como pode ser visto na figura 3. Como pode ser notado, para $n = 1$ a acurácia da RNA5 atinge 100% rapidamente para o conjunto de treinamento, mas se mantém em torno de 65% para o conjunto de teste. Ou seja, no caso $n = 1$, a RNA5 mapeou perfeitamente o conjunto de treinamento, o que implica em uma perda generalização. Diz-se portanto que esse modelo está sobre-ajustado. Já para o caso $n = 30$, a acurácia para o conjunto de treinamento e teste se mantiveram próximas e em torno de 84%, o que comprova que esse modelo está conseguindo melhor generalizar para casos ainda não vistos. A evolução da acurácia para demais casos pode ser vista na figura 5.

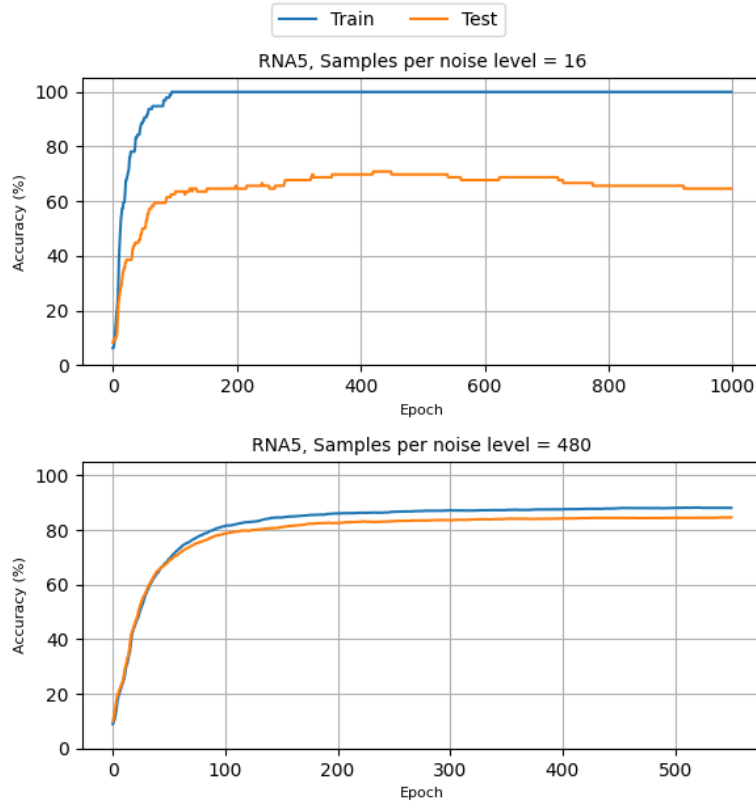


Figura 3: Evolução da acurácia da RNA5 para $n = 1$ e $n = 30$

A figura 4 mostra a acurácia das 18 RNAs criadas para diferentes níveis de ruído. Vê-se claramente a piora do desempenho das RNAs com o aumento do ruído de forma mais acentuada quanto menor foi a quantidade de amostras por nível de ruído em seus treinamentos.

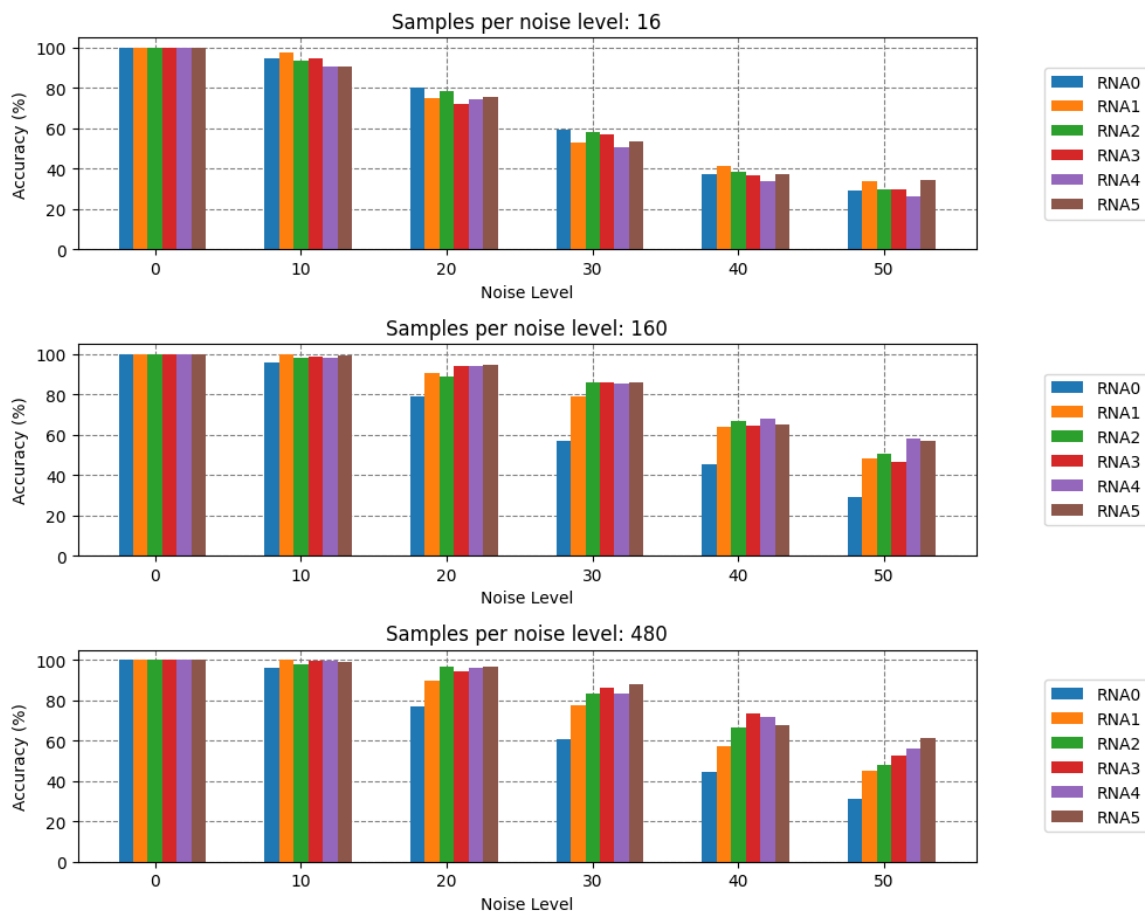


Figura 4: Acurácia da RNAs para diferentes níveis de ruído.

4 Conclusão

Como foi visto nos resultados, as RNAs treinadas com caracteres com ruído desenvolvem maior capacidade de extrapolação do que as RNAs treinadas com caracteres normais. Foi percebido também que a capacidade de extrapolação melhora com o aumento da quantidade de amostras distintas de treinamento, devido tornar mais difícil para que a RNA fique sobre-ajustada.

5 Anexos

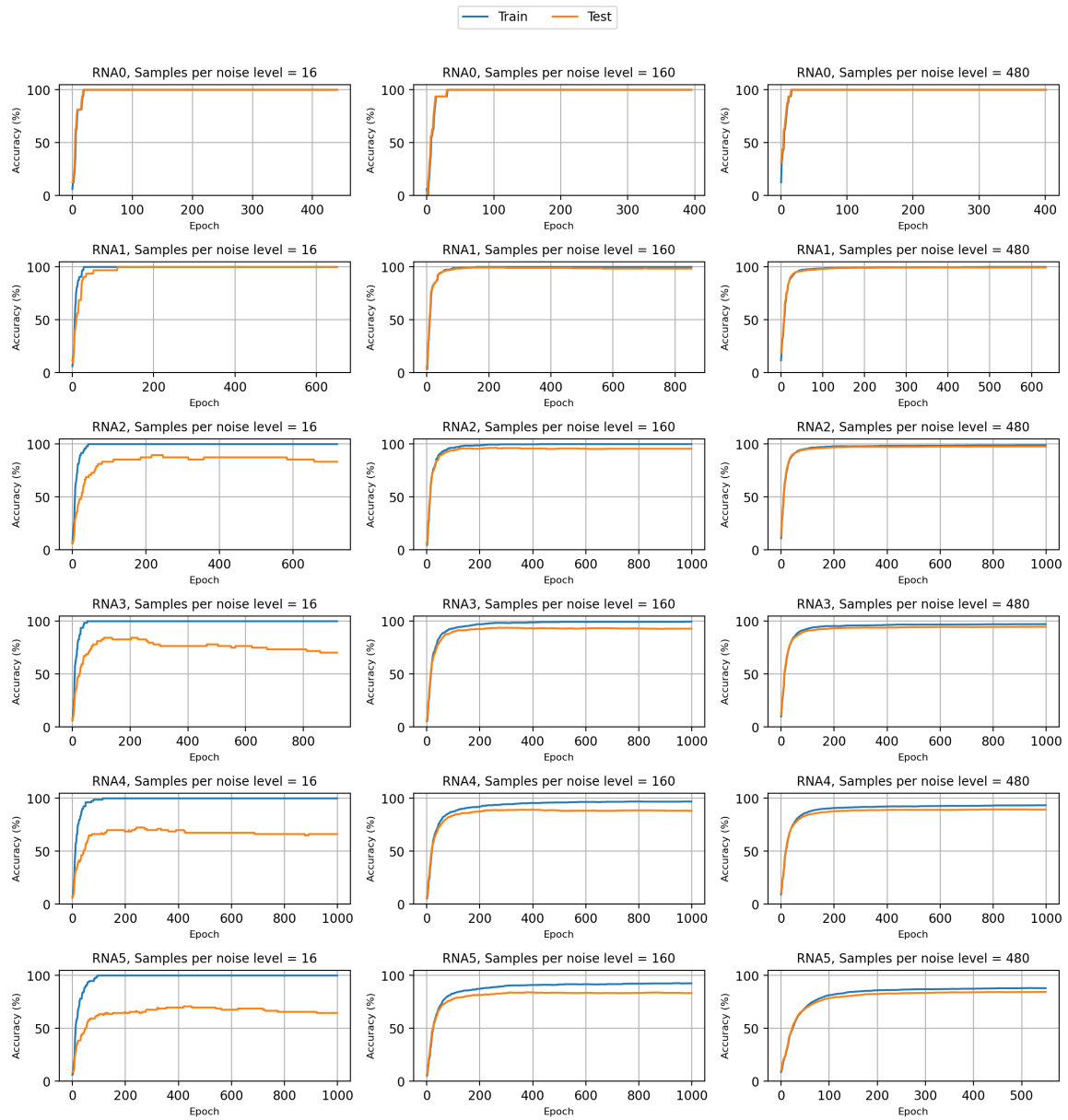


Figura 5: Evolução da acurácia da RNAs para conjuntos de treino e de teste.