



UNIVERSIDADE FEDERAL DO PARANÁ

GIOVANA SPADER SPEZZATTO

MICROESTADOS DE RECORRÊNCIA COMO FERRAMENTA PARA  
CLASSIFICAÇÃO DE SISTEMAS DINÂMICOS VIA APRENDIZADO DE  
MÁQUINA

CURITIBA

2026

GIOVANA SPADER SPEZZATTO

MICROESTADOS DE RECORRÊNCIA COMO FERRAMENTA PARA  
CLASSIFICAÇÃO DE SISTEMAS DINÂMICOS VIA APRENDIZADO DE  
MÁQUINA

Dissertação apresentada ao Programa de Pós-Graduação em Física do Setor de Ciências Exatas da Universidade Federal do Paraná como requisito parcial para a obtenção de título de Mestre em Física.

Orientador: Prof. Dr. Thiago de Lima Prado

CURITIBA

2026

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)  
UNIVERSIDADE FEDERAL DO PARANÁ  
SISTEMA DE BIBLIOTECAS – BIBLIOTECA DE CIÊNCIA E TECNOLOGIA

Spezzatto, Giovana Spader

Microestados de recorrência como ferramenta para classificação de sistemas dinâmicos via aprendizado de máquina / Giovana Spader Spezzatto. – Curitiba, 2026.

1 recurso on-line : PDF.

Dissertação (Mestrado) - Universidade Federal do Paraná, Setor de Ciências Exatas, Programa de Pós-Graduação em Física.

Orientador: Thiago de Lima Prado

1. Entropia. 2. Desordem. 3. Sistemas dinâmicos diferenciais. 4. Aprendizado do computador. I. Universidade Federal do Paraná. II. Programa de Pós-Graduação em Física. III. Prado, Thiago de Lima. IV. Título.



MINISTÉRIO DA EDUCAÇÃO  
SETOR DE CIÊNCIAS EXATAS  
UNIVERSIDADE FEDERAL DO PARANÁ  
PRÓ-REITORIA DE PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO FÍSICA - 40001016020P4

## TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação FÍSICA da Universidade Federal do Paraná foram convocados para realizar a arguição da dissertação de Mestrado de **GIOVANA SPADER SPEZZATTO**, intitulada: **"Microestados de Recorrência como Ferramenta para Classificação de Sistemas Dinâmicos via Aprendizado de Máquina"**, sob orientação do Prof. Dr. THIAGO DE LIMA PRADO, que após terem inquirido a aluna e realizada a avaliação do trabalho, são de parecer pela sua Aprovação no rito de defesa.

A outorga do título de mestra está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 20 de Fevereiro de 2026.

THIAGO DE LIMA PRADO

Presidente da Banca Examinadora

SIDNEY TIAGO SILVA

Avaliador Externo (UNIVERSIDADE FEDERAL DO PARANÁ - DEPARTAMENTO DE FÍSICA)

JOSE DANILO SZEZECH JÚNIOR

Avaliador Interno (UNIVERSIDADE ESTADUAL DE PONTA GROSSA)

## AGRADECIMENTOS

Agradeço primeiramente a Deus, por me conceder força, perseverança e sabedoria ao longo de toda esta trajetória, tornando possível a realização deste trabalho.

Aos meus pais, Nelci e Ivonete, expresso minha profunda gratidão por tudo o que sempre fizeram por mim, incluindo o apoio constante e a priorização dos estudos, fundamentais para minha formação pessoal e acadêmica.

Ao meu namorado, João Gregório, agradeço pelo carinho, compreensão e apoio ao longo de todo o desenvolvimento deste trabalho, especialmente nos momentos mais desafiadores.

Agradeço à minha psicóloga, Theresinha, pelo acolhimento, escuta atenta e apoio emocional ao longo deste período, fundamentais para a manutenção do equilíbrio, da saúde mental e da perseverança durante a realização deste trabalho.

Agradeço ao meu orientador, Prof. Dr. Thiago de Lima Prado, pela orientação, disponibilidade, incentivo e pelas discussões que contribuíram de forma decisiva para o desenvolvimento desta pesquisa.

Agradeço ao Programa de Pós-Graduação em Física da Universidade Federal do Paraná pelo suporte acadêmico e institucional oferecido durante a realização deste trabalho.

Agradeço também à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro concedido, essencial para a vida acadêmica.

Por fim, agradeço a todos os colegas, professores, amigos e familiares que me apoiaram ao longo dessa jornada.

*“A maior felicidade que você pode nos dar  
é a de podermos ver você se  
dedicando aos estudos.”*

Nelci e Ivonete

## RESUMO

Neste trabalho, investiga-se o uso de microestados de recorrência para a classificação de sistemas dinâmicos por meio de técnicas de aprendizado de máquina. A metodologia baseia-se na decomposição dos gráficos de recorrência em pequenas estruturas binárias, denominadas microestados, a partir das quais são construídas distribuições de probabilidade que caracterizam estatisticamente as séries temporais analisadas. A partir dessas distribuições, são definidos quantificadores como a entropia de microestados e medidas de desordem associadas à organização dos padrões de pesos da rede. São considerados sistemas dinâmicos caóticos discretos e contínuos, incluindo o mapa de Bernoulli generalizado, o mapa Logístico, o mapa de Hénon e o sistema de Lorenz, além de processos estocásticos representados por ruídos coloridos. As séries temporais brutas, bem como as probabilidades dos microestados de recorrência associadas a elas são utilizadas como dados de entrada em redes neurais do tipo *Multi-Layer Perceptron*. A análise concentra-se na relação entre o limiar de recorrência  $\varepsilon$ , a entropia de microestados, o quantificador de desordem e o desempenho de classificação das redes neurais. Observa-se uma correspondência sistemática entre os valores de  $\varepsilon$  associados a máximos da entropia de microestados e aqueles que produzem maiores acurácias de classificação. Além disso, a evolução da matriz de pesos da rede neural é analisada conforme a desordem de cada elemento, permitindo caracterizar diferentes regimes de organização do aprendizado em função da representação dos dados de entrada.

Palavras-chave: Microestados de Recorrência; Entropia; Desordem; Sistemas Dinâmicos; Aprendizado de Máquina.

## ABSTRACT

In this work, the use of recurrence microstates for the classification of dynamical systems through machine learning techniques is investigated. The methodology is based on the decomposition of recurrence plots into small binary structures, referred to as microstates, from which probability distributions are constructed to statistically characterize the analyzed time series. From these distributions, quantifiers such as microstate entropy and disorder measures associated with the organization of the network weight patterns are defined. Discrete and continuous chaotic dynamical systems are considered, including the generalized Bernoulli map, the Logistic map, the Hénon map, and the Lorenz system, in addition to stochastic processes represented by colored noises. The raw time series, as well as the probability distributions of the associated recurrence microstates, are used as input data for Multi-Layer Perceptron neural networks. The analysis focuses on the relationship between the recurrence threshold  $\varepsilon$ , microstate entropy, the disorder quantifier, and the classification performance of the neural networks. A systematic correspondence is observed between the values of  $\varepsilon$  associated with maxima of microstate entropy and those yielding higher classification accuracies. In addition, the evolution of the neural network weight matrix is analyzed in terms of the disorder of each element, allowing different regimes of learning organization to be characterized as a function of the input data representation.

Keywords: Recurrence Microstates; Entropy; Disorder; Dynamical Systems; Machine Learning.

## LISTA DE ILUSTRAÇÕES

FIGURA 1 – Gráficos de recorrência . . . . .	20
FIGURA 2 – Microestados de recorrência . . . . .	22
FIGURA 3 – Gráfico de $S(\varepsilon)$ . . . . .	24
FIGURA 4 – Divisão de classes de microestados . . . . .	25
FIGURA 5 – Representação visual dos quantificadores da abordagem via microestados de recorrência . . . . .	34
FIGURA 6 – Primeira arquitetura . . . . .	35
FIGURA 7 – Segunda arquitetura . . . . .	36
FIGURA 8 – Matriz de conexões entre os neurônios da primeira camada oculta (amarelo) e da segunda camada oculta (verde). . . . .	37
FIGURA 9 – Acurácia e entropia em função do parâmetro $\varepsilon$ . . . . .	40
FIGURA 10 – Evolução da matriz de pesos ao se utilizar os dados brutos como dados de entrada da rede neural . . . . .	41
FIGURA 11 – Evolução da matriz de pesos ao se utilizar os quantificadores da abordagem via microestados de recorrência como dados de entrada da rede neural . . . . .	42
FIGURA 12 – Acurácia de treinamento . . . . .	43
FIGURA 13 – Acurácia de treinamento do sistema $\beta x$ , utilizando os dados brutos (D. Brutos), bem como as probabilidades associadas a diferentes tamanhos de microestados ( $N = 2$ , $N = 3$ e $N = 4$ ) como dados de entrada. . . . .	43
FIGURA 14 – Desordem calculada sobre os elementos da matriz de pesos . . . . .	45
FIGURA 15 – Evolução da matriz de pesos da rede indicada na Figura 7, ao longo das 1000 épocas de treinamento, utilizando-se os dados brutos dos sistemas Logístico, Lorenz e Ruídos Coloridos como dados de entrada. . . . .	53

FIGURA 16 – Evolução da matriz de pesos da rede indicada na Figura 7, ao longo das 1000 épocas de treinamento, utilizando-se as probabilidades associadas a cada microestado do mapa Logístico como dados de entrada. . . . .	54
FIGURA 17 – Evolução da matriz de pesos da rede indicada na Figura 7, ao longo das 1000 épocas de treinamento, utilizando-se as probabilidades associadas a cada microestado do sistema Lorenz como dados de entrada. . . . .	55
FIGURA 18 – Evolução da matriz de pesos da rede indicada na Figura 7, ao longo das 1000 épocas de treinamento, utilizando-se as probabilidades associadas a cada microestado dos ruídos coloridos como dados de entrada. . . . .	56
FIGURA 19 – Acurácia de treinamento do mapa Logístico, do Sistema de Lorenz e dos Ruídos Coloridos, utilizando os dados brutos (D.Brutos), bem como as probabilidades associadas a diferentes tamanhos de microestados ( $N = 2$ , $N = 3$ e $N = 4$ ) como dados de entrada. . . . .	57

## LISTA DE TABELAS

TABELA 1 – Microestados de recorrência . . . . .	21
TABELA 2 – Acurácias de teste . . . . .	43

## SUMÁRIO

	<b>Lista de ilustrações . . . . .</b>	<b>8</b>
	<b>Lista de tabelas . . . . .</b>	<b>10</b>
<b>1</b>	<b>INTRODUÇÃO . . . . .</b>	<b>13</b>
<b>2</b>	<b>OBJETIVOS . . . . .</b>	<b>15</b>
2.1	Objetivo geral . . . . .	15
2.2	Objetivos específicos . . . . .	15
2.3	Organização da dissertação . . . . .	15
<b>3</b>	<b>REVISÃO BIBLIOGRÁFICA . . . . .</b>	<b>16</b>
3.1	BREVE CONTEXTUALIZAÇÃO HISTÓRICA . . . . .	16
3.2	SISTEMAS DINÂMICOS . . . . .	17
3.2.1	Recorrência . . . . .	18
3.2.2	Gráficos de recorrência . . . . .	19
3.2.3	Microestados de recorrência . . . . .	21
3.2.4	Entropia de microestados . . . . .	21
3.2.5	Desordem (DISREM) . . . . .	24
3.3	A importância dos computadores . . . . .	26
3.3.1	Aprendizado de máquina . . . . .	27
3.3.2	Multi-Layer Perceptrons . . . . .	29
<b>4</b>	<b>METODOLOGIA . . . . .</b>	<b>31</b>
4.1	DADOS GERADOS . . . . .	31
4.2	TRANSFORMAÇÃO DE ESPAÇO: DADOS BRUTOS PARA MICROESTADOS . . . . .	33
4.3	ARQUITETURA DA REDE DE APRENDIZADO DE MÁQUINA	35
<b>5</b>	<b>RESULTADOS E DISCUSSÃO . . . . .</b>	<b>38</b>
5.1	ANÁLISE DO LIMAR DE RECORRÊNCIA . . . . .	38

5.2	ANÁLISE DA MATRIZ DE PESOS . . . . .	39
5.2.1	Microestados de recorrência como uma solução para o overfitting	41
5.2.2	Desordem na dinâmica dos pesos . . . . .	44
<b>6</b>	<b>CONSIDERAÇÕES FINAIS . . . . .</b>	<b>47</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>49</b>
	<b>APÊNDICE A – RESULTADOS ADICIONAIS . . . . .</b>	<b>53</b>

## 1 INTRODUÇÃO

A caracterização de sistemas dinâmicos complexos constitui um dos desafios centrais da física contemporânea. Sistemas determinísticos podem exibir comportamentos altamente irregulares e imprevisíveis, como ocorre no caos determinístico, enquanto sistemas estocásticos apresentam flutuações associadas a processos aleatórios (1). Distinguir e classificar diferentes sistemas dinâmicos a partir de séries temporais finitas é uma tarefa não trivial, especialmente na presença de ruído e de alta dimensionalidade.

Ao longo das últimas décadas, diversas técnicas foram desenvolvidas para extrair informações relevantes de séries temporais, como a reconstrução do espaço de fase (2), o cálculo de expoentes de Lyapunov (3) e a análise de recorrência (4). Em particular, os gráficos de recorrência destacam-se por fornecer uma representação visual das propriedades da dinâmica, permitindo identificar padrões característicos associados a comportamentos periódicos, caóticos ou estocásticos.

Mais recentemente, a introdução dos microestados de recorrência ampliou essa abordagem ao permitir uma análise local e estatística das estruturas presentes nos gráficos de recorrência (5). Esses microestados reduzem a dimensionalidade dos dados ao mesmo tempo em que preservam as informações essenciais sobre o sistema, na forma de distribuição de probabilidades. Associada a essa representação, a entropia de microestados, bem como o quantificador de desordem (6) emergem para complementar a análise acerca da complexidade dinâmica.

Paralelamente, o avanço do aprendizado de máquina tem proporcionado novas ferramentas para a análise e classificação automática de sistemas. Redes neurais artificiais, em especial os perceptrons multicamadas (MLPs) (7), têm se mostrado eficazes na identificação de padrões em grandes conjuntos de dados tendo seus desempenhos fortemente atrelados à qualidade dos atributos de entrada (8).

Neste contexto, o presente trabalho investiga a utilização dos microestados de recorrência como uma ponte entre a teoria de sistemas dinâmicos e o aprendizado de máquina. Entre os principais objetivos, destaca-se a análise de como as distribuições de probabilidade associadas a cada microestado, bem como a entropia de micros-

tados, podem ser empregadas para se extrair as informações essenciais de sistemas determinísticos caóticos e processos estocásticos (ruídos coloridos). A utilização desses novos dados como entrada de um algoritmo de classificação supervisionada baseado em uma rede neural do tipo MLP interfere diretamente na otimização de seu desempenho, quando comparada à utilização dos dados brutos. Adicionalmente, busca-se compreender a organização interna das redes neurais durante o processo de aprendizado, por meio da análise do quantificador de desordem "*disorder index via symmetry in recurrence microstates*" (DISREM).

Dentre os resultados obtidos, destacam-se aqueles que evidenciam o potencial de análises baseadas nos microestados de recorrência como altamente informativas, permitindo elevadas taxas de acerto na distinção entre diferentes sistemas dinâmicos. Assim, foi possível estabelecer um critério objetivo para a escolha do parâmetro  $\varepsilon$ , baseado na maximização da entropia de microestados, o que garante que a diversidade estrutural da dinâmica é máxima, favorecendo a extração de informação significativa das séries temporais.

Outro resultado relevante refere-se à análise da desordem associada aos elementos da matriz de pesos do MLP ao longo das épocas de treinamento. Observou-se que treinamentos realizados com dados brutos como entrada da rede apresentam valores baixos de desordem por grande parte da arquitetura, perdendo capacidade de generalização, caracterizando o *overfitting*.

Por outro lado, um regime intermediário, no qual coexistem pesos com baixas e altas desordens, indica a presença de especializações localizadas, preservando a variabilidade nas demais conexões. Esse padrão foi observado quando se utilizaram dados de entrada baseados em microestados, sugerindo que essa representação favorece um equilíbrio entre especialização e diversidade interna da rede, condição associada a um desempenho mais robusto e generalizável.

## 2 OBJETIVOS

### 2.1 Objetivo geral

Investigar a eficácia dos microestados de recorrência como ferramenta para a classificação de sistemas dinâmicos, por meio de técnicas de aprendizado de máquina.

### 2.2 Objetivos específicos

- Implementar a transformação de séries temporais em representações baseadas em microestados de recorrência;
- Avaliar a relação entre a entropia de microestados em função do limiar de recorrência  $\varepsilon$  e a acurácia de classificação de redes neurais do tipo *Multi-Layer Perceptron* para diferentes sistemas dinâmicos;
- Investigar a influência do tipo e do tamanho do conjunto de dados, bem como o número de classes no desempenho do algoritmo de classificação;
- Analisar a evolução da matriz de pesos da rede neural como forma de compreender a dinâmica interna do processo de aprendizado.

### 2.3 Organização da dissertação

A seção 3 desta dissertação aborda os fundamentos teóricos relacionados a sistemas dinâmicos, gráficos de recorrência, microestados e quantificadores de entropia e desordem, além dos fundamentos relacionados à estrutura de aprendizado de máquina utilizada. Na seção 4 é descrita a metodologia empregada, incluindo os sistemas gerados e a arquitetura das redes neurais utilizadas. A seção 5 é dedicada à apresentação e discussão dos resultados obtidos, com ênfase no desempenho da classificação dos sistemas, na análise do limiar de recorrência  $\varepsilon$  e na análise da desordem da matriz de pesos. Por fim, a seção 6 reúne as conclusões e considerações finais, bem como perspectivas para trabalhos futuros.

### 3 REVISÃO BIBLIOGRÁFICA

#### 3.1 BREVE CONTEXTUALIZAÇÃO HISTÓRICA

Desde os primórdios da civilização humana, a curiosidade sobre o funcionamento do universo moldou o desenvolvimento do conhecimento científico. A busca pela compreensão do funcionamento dos astros, da dinâmica dos fenômenos naturais e do comportamento dos seres vivos impulsionou uma série contínua de revoluções intelectuais, estabelecendo a base para o pensamento científico.

Os primeiros registros escritos de algo semelhante à filosofia natural remetem às antigas civilizações do Egito e da Mesopotâmia, cujo surgimento é estimado entre 3500 e 3000 a.C. A partir do estudo dos hieróglifos egípcios e da escrita cuneiforme mesopotâmica, hoje é possível compreender que, apesar de a visão de mundo dessas sociedades estar fortemente condicionada por elementos míticos e religiosos, havia um notável interesse pelo mundo físico, especialmente nas áreas de astronomia, matemática e medicina (9).

Na Grécia antiga, os filósofos da natureza foram pioneiros em propor explicações racionais, afastando-se das narrativas míticas tradicionais e buscando princípios universais (*archái*) que estruturassem o cosmos (10). A consolidação de uma linguagem matemática rigorosa foi essencial, sendo Eudoxo (408–355 a.C.) um dos nomes responsáveis por estabelecer as bases, e Euclides (c. 300 a.C.) por formalizá-las em Os Elementos (11). A partir de então, a matemática tornou-se a principal ferramenta para descrever proporções, trajetórias e relações geométricas.

Mais adiante, a queda do Império Romano do Ocidente deu início a um período de grande fragmentação política e cultural, atualmente conhecido como Idade Média (476-1453 d.C.). Enquanto isso, estudiosos árabes e persas como Al-Kindi, Al-Farabi, Avicena (Ibn Sina) e Averróis (Ibn Rushd) foram os grandes responsáveis por conservar as ideias gregas e introduzir novas concepções matemáticas e experimentais, que foram transmitidas ao longo dos séculos seguintes (9).

O Renascimento (c. 1350-1600 d.C.) marcou o ressurgimento científico na Europa, sendo a astronomia o palco de notáveis transformações. Nicolau Copérnico (1473–1543) desafiou a visão geocêntrica ao propor o modelo heliocêntrico, realo-

cando o Sol no centro do universo (12). Mais tarde, Johannes Kepler (1571–1630) transformou essa proposta enunciando as leis do movimento planetário, com base nas observações precisas de Tycho Brahe (13).

Galileu Galilei (1564–1642 d.C.), por sua vez, consolidou o método experimental e aplicou a matemática à descrição do movimento terrestre. Ao estudar a queda dos corpos e o movimento dos projéteis, Galileu demonstrou que o comportamento físico podia ser descrito por leis quantitativas.

Impulsionado por esse avanços, Isaac Newton (1643–1727) desenvolveu a obra *Philosophiae Naturalis Principia Mathematica* (1687), unificando a física terrestre e celeste sob as mesmas leis universais do movimento e da gravitação. Ao definir tempo, espaço, massa e força em termos puramente matemáticos, Newton estabeleceu o cosmos como um sistema ordenado e previsível (14).

### 3.2 SISTEMAS DINÂMICOS

De forma geral, um sistema dinâmico pode ser definido como um conjunto de variáveis cujo estado evolui ao longo do tempo, segundo regras matemáticas determinísticas ou estocásticas. No caso de sistemas discretos, estas regras costumam ser funções, enquanto para sistemas contínuos, equações diferenciais ordinárias (EDOs) (15, 1). Em ambos os casos, a dinâmica do sistema pode ser representada por trajetórias em um espaço abstrato, chamado de espaço de fase, que contém todas as suas possíveis configurações.

O desenvolvimento da mecânica newtoniana representa um dos marcos mais importantes da história, sendo a base para a compreensão da física clássica e para o estudo dos sistemas dinâmicos. Neste contexto, a perspectiva determinística da mecânica clássica moldou por séculos a compreensão da evolução de certas quantidades com o passar do tempo, bem como o entendimento de o estado presente de um dado sistema poder influenciar seu estado futuro.

Uma nova concepção surgiu com Henri Poincaré (1854-1912 d.C.), ao investigar o problema dos três corpos e demonstrar que sistemas determinísticos podiam exibir um comportamento imprevisível devido à extrema sensibilidade às condições iniciais. Esse resultado inaugurou a noção de caos determinístico, um regime em que pequenas variações no estado inicial geram trajetórias radicalmente distintas,

impossibilitando previsões a longo prazo, ainda que as leis sejam conhecidas (16).

A partir da década de 1960, este tipo de comportamento passou a ser melhor compreendido com os trabalhos de Edward Lorenz (17), Benoît Mandelbrot (18), Stephen Smale (19) e outros pioneiros. Deste modo, houve a consolidação da Dinâmica Não Linear como um campo distinto da Física. A partir de então, o foco dos estudos sobre estes sistemas mudou da busca pela solução analítica das trajetórias para a caracterização das propriedades globais do espaço de fase e dos seus atratores, levando ao desenvolvimento de diversas técnicas.

Dentre elas, podem-se citar a reconstrução do espaço de fase (Teorema de Takens) (2), a estimação de dimensões fractais (e.g., Dimensão de Correlação) (20) e o cálculo dos Expoentes de Lyapunov (3), que assumiram um papel central na caracterização do grau de complexidade e do determinismo presentes em sistemas reais.

### 3.2.1 Recorrência

Ao discutir o comportamento de sistemas dinâmicos e a recorrência das trajetórias no espaço de fases, Poincaré (1890, p. 6, tradução nossa) (16) afirmou:

Reconheci que, nesse caso, os três corpos voltarão uma infinidade de vezes tão próximos quanto se deseje de sua posição inicial, salvo se as condições iniciais do movimento forem excepcionais. <sup>1</sup>

Essa passagem evidencia um dos princípios fundamentais da Dinâmica, em sistemas conservativos limitados (e.g., o problema dos três corpos com energia constante): a recorrência.

O Teorema da Recorrência de Poincaré postula que, em sistemas dinâmicos que preservam o volume no espaço de fase (sistemas hamiltonianos) ou que são dissipativos, mas com atratores de dimensão finita, quase todo ponto em um subconjunto de medida positiva do espaço de fase retornará infinitas vezes à

---

<sup>1</sup> "J'ai reconnu que dans ce cas les trois corps repasseront une infinité de fois aussi près que l'on veut de leur position initiale, à moins que les conditions initiales du mouvement ne soient exceptionnelles." POINCARÉ, H. Sur le problème des trois corps et les équations de la dynamique. Paris: Gauthier-Villars, 1890. Tome I, p. 6.)

vizinhança de sua posição inicial. Esta propriedade estabelece a base teórica para a análise de estabilidade e de comportamento a longo prazo das trajetórias.

Contudo, é importante ressaltar que a recorrência em sistemas caóticos determinísticos não implica previsibilidade, mas sim uma revisitação do atrator. Assim, devido à sensibilidade às condições iniciais próximas, as trajetórias permanecem confinadas, mesmo que divirjam exponencialmente.

### 3.2.2 Gráficos de recorrência

Na década de 80, a compreensão do caos dependia principalmente da reconstrução de atratores no espaço de fase, o que se tornava algo de difícil visualização e interpretação em sistemas de alta dimensionalidade. Desta forma, em 1987, Eckmann *et al.* exploraram de forma gráfica e intuitiva o conceito de recorrência, propondo representar as vezes em que o sistema retorna, ou se aproxima, de estados anteriores através dos *Recurrence Plots* (RPs) (4).

Partindo do Teorema da Recorrência de Poincaré (16), constrói-se uma matriz binária, chamada de Matriz de Recorrência  $\mathbf{R}$ . Esta matriz tem tamanho  $\mathcal{N} \times \mathcal{N}$ , onde  $\mathcal{N}$  é o número de vetores de estado e cada elemento  $R_{i,j}$  é definido pela Equação 3.1, onde o parâmetro  $\varepsilon$  define a distância máxima no espaço de fase entre dois pontos  $\vec{x}_i$  e  $\vec{x}_j$ , calculada via norma euclidiana, para serem considerados recorrentes. Além disso, a função de passo de *Heaviside* atua como um operador de limiarização, transformando o espaço contínuo de distâncias em uma resposta binária.

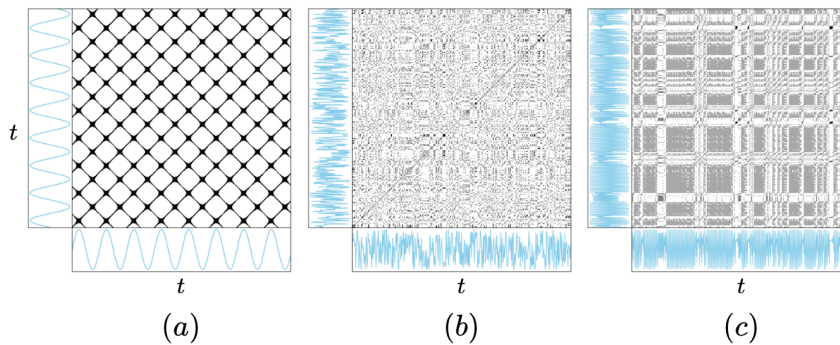
$$R_{i,j}(\varepsilon) = \Theta(\varepsilon - \|\vec{x}_i - \vec{x}_j\|) = \begin{cases} 1, & \text{se } \|\vec{x}_i - \vec{x}_j\| \leq \varepsilon, \\ 0, & \text{caso contrário.} \end{cases} \quad (3.1)$$

Desta forma, o RP se forma a partir de diversos pontos pretos e brancos, indicando se há presença ( $R_{i,j} = 1$ ) ou ausência ( $R_{i,j} = 0$ ) de recorrência, respectivamente. O resultado é um gráfico quadrado, simétrico em relação à diagonal principal  $i = j$ , onde  $i$  e  $j$  são índices de tempo.

A Figura 1 apresenta alguns exemplos de gráficos de recorrência de um Seno, um Ruído Branco do Mapa Logístico, sendo estes exemplos de sistemas periódicos, aleatórios e caóticos, respectivamente. No caso do seno, o gráfico de recorrência apresenta estruturas altamente regulares, com linhas diagonais longas e paralelas,

refletindo a repetição cíclica dos estados no espaço de fase e a previsibilidade da dinâmica. Já no caso do ruído branco, observa-se um padrão mais disperso homogêneo, sem estruturas organizadas, indicando ausência de correlação temporal e comportamento puramente aleatório. Por fim, o mapa logístico exibe um regime caótico, caracterizado pela sensibilidade às condições iniciais e apresentando linhas diagonais curtas, fragmentadas e padrões complexos.

FIGURA 1 – Gráficos de recorrência



Representação de gráficos de recorrência de (a) seno, (b) ruído branco e (c) mapa logístico.

Fonte: Elaboração própria (2026), inspirada em (21).

Para quantificar objetivamente as informações contidas nos RPs, uma série de técnicas foram desenvolvidas, em especial o método *Recurrence Quantification Analysis* (RQA) (22, 21). O RQA converte as informações contidas no RP em métricas estatísticas, como a Taxa de Recorrência (RR), o Determinismo (DET) e a Entropia de Shannon (RQA-ENT), amplamente utilizadas para distinguir entre regimes estocásticos, periódicos e caóticos.

Apesar de fornecerem medidas muito úteis, tais quantificadores são computacionalmente custosos, sensíveis à escolha de parâmetros e pouco eficientes para distinguir dinâmicas semelhantes (23). Nesse contexto, abordagens baseadas em microestados de recorrência emergem como uma extensão natural da RQA tradicional, permitindo uma caracterização mais detalhada da geometria da matriz de recorrência.

### 3.2.3 Microestados de recorrência

Partindo-se do pressuposto de que pequenas estruturas locais presentes na matriz de recorrência carregam informações fundamentais sobre a dinâmica do sistema correspondente, surge uma nova classe de métricas: os Microestados de Recorrência, conhecidos em inglês como *Recurrence Microstates* (RM).

Os microestados consistem em pequenas matrizes binárias de tamanho  $N \times N$ , obtidas a partir do RP, resultando em  $N^* = 2^{N^2}$  configurações distintas possíveis. Para o caso específico onde  $N = 2$ , tais microestados têm a forma matricial da Equação 3.2.

$$\begin{bmatrix} R_{i,j} & R_{i,j+1} \\ R_{i+1,j} & R_{i+1,j+1} \end{bmatrix} \quad (3.2)$$

Adicionalmente, para este caso específico há  $N^* = 16$  diferentes microestados possíveis, representados matricialmente conforme a Tabela 1 e visualmente conforme a Figura 2.

TABELA 1 – Microestados de recorrência

$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ (a)	$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ (b)	$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ (c)	$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ (d)	$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ (e)	$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ (f)	$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ (g)	$\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$ (h)
$\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$ (i)	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ (j)	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ (k)	$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ (l)	$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ (m)	$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ (n)	$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ (o)	$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ (p)

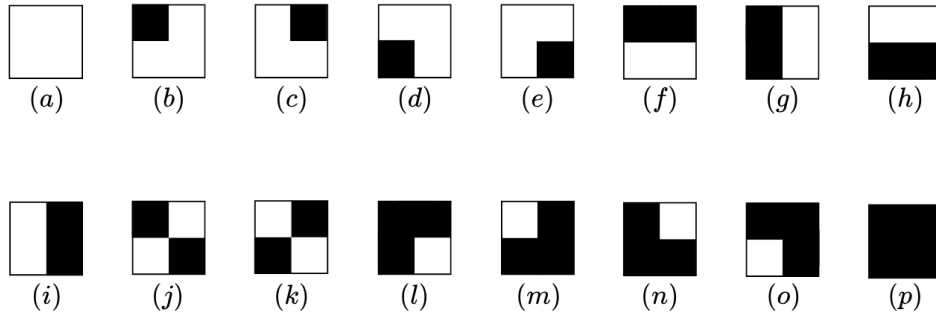
Representação matricial dos microestados de recorrência para  $N = 2$ .

Diferentemente dos métodos tradicionais, que examinam o gráfico de recorrência como um todo, a abordagem utilizando os RMs permite análises locais e estatísticas dessas estruturas elementares que compõem o gráfico (24, 5).

### 3.2.4 Entropia de microestados

Devido à natureza local dos RMs, ao invés de varrer cada posição possível do RP, pode-se selecionar aleatoriamente uma amostra de  $\bar{N}$  microestados. Assim,

FIGURA 2 – Microestados de recorrência



Representação visual dos microestados de recorrência para  $N = 2$ , conforme a tabela 1.

Fonte: Elaborado pela própria autora (2026).

pode-se analisar o mesmo a partir de um aspecto estatístico, permitindo associar a cada um deles uma probabilidade, conforme a equação (3.3)

$$P_i = \frac{n_i}{\bar{N}}. \quad (3.3)$$

Assim,  $P_i$  é a probabilidade de ocorrência do microestado  $i$ , dada pelo número de vezes  $n_i$  em que este apareceu em  $\bar{N}$  amostras.

Adicionalmente, outro quantificador de recorrência pôde ser desenvolvido por Corso *et al.* (24), baseando-se na Entropia de Shannon (25), que mede a incerteza de uma distribuição de probabilidades, segundo a equação (3.4)

$$S(N^*) = - \sum_{i=1}^{i=N^*} P_i \ln P_i. \quad (3.4)$$

A entropia  $S(N^*)$  é então calculada sobre todos os microestados possíveis, mais especificamente em função das probabilidades  $P_i$  de ocorrência de cada um.

Sabendo que maximizar a entropia é sinônimo de considerar todos os microestados como igualmente prováveis e que o número de microestados admissíveis é  $N^*$ , então a probabilidade associada a cada um deles, neste caso, é dada pela Equação 3.5

$$P_i = \frac{1}{N^*}. \quad (3.5)$$

É possível estimar analiticamente o máximo da entropia ao substituirmos este valor na Equação 3.4, resultando na Equação 3.6.

$$\begin{aligned}
 S_{max} &= - \sum_{i=1}^{N^*} \left( \frac{1}{N^*} \right) \ln \left( \frac{1}{N^*} \right), \\
 &= -N^* \left( \frac{1}{N^*} \right) \ln \left( \frac{1}{N^*} \right), \\
 &= - \ln \left( \frac{1}{N^*} \right), \\
 &= - (\ln(1) - \ln(N^*)), \\
 &= \ln(N^*).
 \end{aligned} \tag{3.6}$$

Assim, de forma geral, a entropia de microestados está limitada por  $0 \leq S(N^*) \leq \ln(N^*)$ . Ressalta-se que, como a distribuição dessas probabilidades altera-se conforme o raio de vizinhança é expandido ou contraído, a entropia  $S$  torna-se uma função dependente do parâmetro  $\varepsilon$ . Logo, podemos escrever  $S(\varepsilon)$ , normalizando-a ao dividir por  $S_{max}$ , conforme a Equação 3.7

$$S(\varepsilon) = \sum_{i=1}^{i=N^*} \frac{P_i \ln P_i}{\ln N^*}. \tag{3.7}$$

Sendo a Figura 3 a representação desta curva  $S(\varepsilon)$ .

Percebe-se que, para valores de  $\varepsilon$  muito pequenos, a entropia  $S(\varepsilon) \approx 0$ , justificado pelo fato de que nenhum estado do sistema recorreria, então no caso específico onde  $N = 2$ , teríamos praticamente todos os microestados do tipo (a) na Figura 2. Semelhantemente, para valores muito grandes de  $\varepsilon$ , obtemos  $S(\varepsilon) \approx 0$ , uma vez que todos os estados do sistema recorreriam, ou seja, teríamos praticamente todos os microestados do tipo (p) em 2.

Entre esses dois extremos, a entropia atinge valores elevados para limiares de recorrência intermediários, indicando uma maior diversidade na distribuição de microestados. Nesse regime, diferentes microestados ocorrem com probabilidades praticamente iguais, refletindo uma representação mais rica da dinâmica e capturando de forma eficiente a desordem do sistema.

Sob a ótica da entropia de informação, o comportamento de  $S(\varepsilon)$  fornece um critério objetivo para a escolha do limiar de recorrência  $\varepsilon$ , sendo relevante no contexto

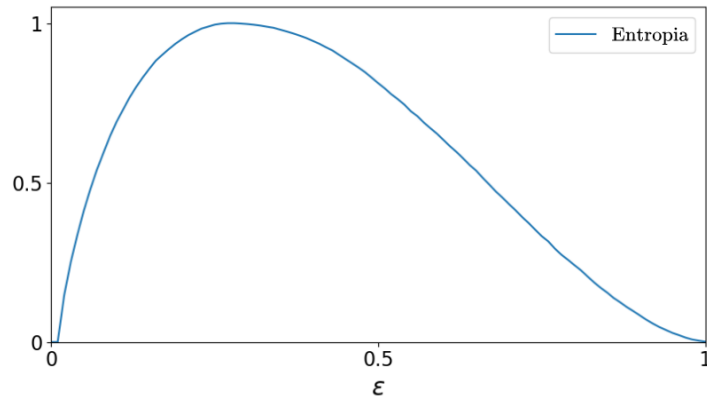
FIGURA 3 – Gráfico de  $S(\varepsilon)$ 

Gráfico normalizado da entropia  $S$  em função do parâmetro  $\varepsilon$ .

Fonte: Elaborado pela própria autora (2026).

de aprendizado de máquina, uma vez que representações com maior conteúdo informacional tendem a fornecer entradas menos redundantes, impactando diretamente a otimização do desempenho dos algoritmos de classificação supervisionada.

### 3.2.5 Desordem (DISREM)

Com base na metodologia dos microestados de recorrência, Flauzino *et al.* desenvolveram um quantificador de desordem que avalia diretamente o grau de aleatoriedade presente nos dados (6).

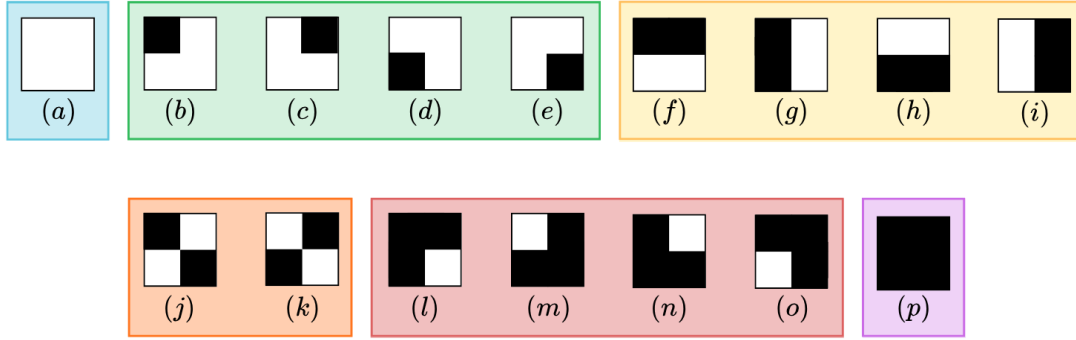
A ideia central do método é comparar pequenos trechos  $\mathbf{X}_{(n)}$  e  $\mathbf{X}_{(m)}$  de um sinal entre si e registrar se seus valores são próximos ou não, construindo uma matriz de recorrência  $\mathbf{M}$  associada a eles, de tamanho  $N \times N$  conforme a Equação 3.1. Assim, para valores de  $N$  consideravelmente menores que o tamanho da série, denominamos  $\mathbf{M}$  como um microestado de recorrência

$$\mathbf{M} = \mathbf{R}(\mathbf{X}_{(n)}, \mathbf{X}_{(m)}). \quad (3.8)$$

Em um regime de máxima desordem, a ordem dos pontos dentro dos blocos de dados não é relevante, de modo que diferentes rearranjos de um mesmo padrão de recorrência ocorrem com a mesma frequência. A partir dessa propriedade, definem-se as classes de microestados, nas quais todos os microestados de uma

mesma classe apresentam probabilidades de ocorrência equiprováveis. A figura Figura 4 ilustra as 6 classes possíveis para os microestados de tamanho  $N = 2$ .

FIGURA 4 – Divisão de classes de microestados



Microestados de tamanho  $N = 2$  agrupados em diferentes classes.

Fonte: Elaborado pela própria autora (2026).

Para quantificar o grau com que essa equiprobabilidade é satisfeita em dados reais e finitos, utiliza-se a Entropia de Shannon associada à distribuição de probabilidades dos microestados dentro de cada classe. Para uma dada classe  $\mathcal{M}_a$ :

$$\xi_a(\varepsilon) = - \sum_{\mathbf{M} \in \mathcal{M}_a} p(\mathbf{M}) \ln p(\mathbf{M}) \quad (3.9)$$

Essa entropia atinge seu valor máximo  $\ln m_a$  quando todos os  $m_a$  microestados da classe são igualmente prováveis. A entropia global é então obtida pela média normalizada das entropias de todas as classes relevantes

$$\xi(\varepsilon) = \frac{1}{A} \sum_{a=1}^A \frac{\xi_a(\varepsilon)}{\ln m_a}, \quad (3.10)$$

onde  $A$  é o número total de classes que contribuem efetivamente.

Define-se, então, ao maximizar  $\xi(\varepsilon)$  com relação à  $\varepsilon$ , o índice de desordem via microestados de recorrência (DISREM) como

$$\Xi = \max_{\varepsilon} \xi(\varepsilon), \quad (3.11)$$

com  $0 \leq \Xi \leq 1$ , onde  $\Xi = 1$  indica que os dados são altamente desordenados, como um ruído branco não correlacionado, enquanto valores menores refletem a presença de correlações temporais, estruturas determinísticas ou dependência dinâmica.

Dessa forma, o parâmetro  $\Xi$  fornece uma medida normalizada do grau de desordem dinâmica presente na série temporal, permitindo comparar sistemas distintos sob uma mesma métrica.

### 3.3 A importância dos computadores

A história da ciência está profundamente ligada à necessidade de realizar cálculos cada vez mais complexos. Conforme descrito na seção 3.1, desde a antiguidade, problemas de natureza astronômica, geométrica e mecânica exigiam métodos sistemáticos de cálculo. No entanto, à medida que os modelos científicos se aperfeiçoaram para descrever fenômenos mais realistas, envolvendo múltiplas variáveis, relações não lineares e equações cuja solução analítica nem sempre era possível, o cálculo manual passou a ser insuficiente (26).

Nesse contexto, ao longo do século XIX, surge a Máquina de Diferenças de Charles Babbage, considerada a primeira calculadora capaz de automatizar um método matemático completo para fins científicos, especificamente o cálculo sistemático de tabelas numéricas por meio do método das diferenças finitas.

Já em meados do século XX, durante a Segunda Guerra Mundial, a crescente demanda por cálculos balísticos por parte do Exército dos Estados Unidos evidenciou novamente as limitações dos métodos tradicionais de cálculo. Nesse cenário, John W. Mauchly e J. Presper Eckert desenvolveram o ENIAC (*Electronic Numerical Integrator and Computer*), considerado um dos primeiros computadores eletrônicos de grande escala, capaz de realizar milhares de operações por segundo, superando significativamente os dispositivos até então utilizados (26, 27).

Paralelamente a esses avanços tecnológicos, Alan Turing estabeleceu os fundamentos teóricos da computação ao demonstrar que qualquer processo algorítmico pode ser formalizado e executado por uma máquina universal (hoje chamada de Máquina de Turing), fornecendo a base conceitual necessária para compreender os limites e as possibilidades da computação (28).

Após a Segunda Guerra, os computadores passaram a ser adotados por universidades e centros de pesquisa científica, inicialmente para resolver equações diferenciais e simular fenômenos físicos. Durante as décadas seguintes, o aumento da capacidade de processamento permitiu que cientistas e engenheiros explorassem

problemas cada vez mais complexos, especialmente para simular sistemas dinâmicos caóticos e não lineares, comuns em áreas como a física, biologia e engenharia. Assim, o computador, que antes era visto como uma ferramenta auxiliar, tornou-se um dos principais pilares da pesquisa científica (29).

### 3.3.1 Aprendizado de máquina

O aumento da capacidade computacional das últimas décadas, como processadores mais rápidos, maior capacidade de armazenamento e o desenvolvimento de arquiteturas paralelas, mudou completamente a forma como problemas científicos são abordados.

Em 1943 Warren McCulloch e Walter Pitts (lógico-matemático) criaram um modelo matemático simples de neurônio, a fim de provar que era possível descrever o funcionamento do cérebro humano usando lógica e matemática (30). Este modelo, que estabeleceu a base conceitual das redes neurais artificiais, pode ser descrito por

$$y = \begin{cases} 1, & \text{se } \sum_i w_i x_i \geq \theta, \\ 0, & \text{caso contrário.} \end{cases} \quad (3.12)$$

Assim, cada  $x_i$  representa o estado do neurônio de entrada  $i$  (se  $x_i = 1$  o neurônio está ativo, se  $x_i = 0$ , está inativo), o peso  $w_i$  quantifica a influência de cada neurônio de entrada  $i$  sobre o neurônio analisado e  $\theta$  é o limiar necessário para ativação. Desta forma, se a soma ponderada das entradas for o suficiente para ativar o neurônio analisado,  $y = 1$ , do contrário,  $y = 0$ .

Apesar de sua grande importância, o modelo de McCulloch e Pitts possuía limitações significativas, uma vez que os pesos sinápticos eram fixos e definidos manualmente. Em 1958, Rosenblatt contornou este problema introduzindo o *Perceptron*, cujo ajuste automático dos pesos ocorriam por meio de um algoritmo de aprendizado supervisionado (31). O neurônio, em si, continua muito parecido com o de McCulloch-Pitts

$$y = \begin{cases} 1, & \text{se } \varphi \left( b + \sum_i w_i x_i \right), \\ 0, & \text{caso contrário.} \end{cases} \quad (3.13)$$

Há, agora, uma generalização das entradas, que podem ser números reais. Além disso,  $\varphi(\cdot)$  é a função de ativação responsável por ativar (ou não) o neurônio e  $b$  é o *bias*, que ajusta o modelo para se adaptar melhor aos dados fornecidos.

A principal inovação do perceptron está na regra de aprendizado supervisionado. Dado um conjunto de dados rotulados  $(\mathbf{x}, d)$ , onde  $d$  é a saída desejada ( $d \in \{0,1\}$ ), os pesos são atualizados iterativamente conforme

$$w_i^{t+1} = w_i^t + \eta(d - y)x_i, \quad (3.14)$$

sendo  $\eta$  a taxa de aprendizado,  $(d - y)$  o erro de classificação e  $t$  a iteração de treinamento.

A principal limitação deste modelo está ligada ao fato de ele somente atuar em sistemas linearmente separáveis. Apesar disso, o *Perceptron* foi fundamental para o desenvolvimento de redes neurais multicamadas, dos algoritmos de retropropagação e do próprio aprendizado de máquina moderno.

Na década de 1980 o termo *machine learning* passou a ser amplamente utilizado para designar um conjunto de métodos cujo objetivo é viabilizar que modelos computacionais aprendam a partir de dados, em vez de seguir regras previamente programadas.

A introdução do algoritmo de retropropagação do erro (*backpropagation*) possibilitou o treinamento eficiente de redes neurais com múltiplas camadas, superando as limitações dos modelos anteriores. Humelhart *et al.* (32) consideraram em seu trabalho redes do tipo feedforward (dados se movem em uma única direção, da entrada para a saída), compostas por múltiplas camadas de neurônios artificiais descritos por uma equação análoga à Equação 3.13, com a diferença de que  $\varphi$  deve ser diferenciável na Equação 3.15

$$y_i = \varphi \left( b_i + \sum_j w_{ij}x_j \right) = \varphi(z_i). \quad (3.15)$$

Desta forma, o erro global da rede é definido por uma função custo, como o erro quadrático médio da Equação 3.16

$$E = \frac{1}{n} \sum_i^n (y_i - d_i)^2. \quad (3.16)$$

E a atualização dos pesos, neste caso, ocorre conforme a Equação 3.17

$$w_{ij}^{t+1} = w_{ij}^t - \eta \frac{\partial E}{\partial w_{ij}}. \quad (3.17)$$

Assim, o atualmente conhecido como *Multi-Layer Perceptron* se tornou uma das arquiteturas mais tradicionais e amplamente utilizadas no aprendizado de máquina supervisionado.

### 3.3.2 Multi-Layer Perceptrons

O Perceptron Multicamadas (MLP, do inglês *Multilayer Perceptron*) é uma rede neural do tipo *feedforward*, na qual os neurônios são organizados em camadas (em geral uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída), sem realimentação, garantindo que a informação flua de maneira unidimensional.

Cada neurônio do MLP realiza uma combinação linear das entradas, ponderada por pesos ajustáveis, seguida da aplicação de uma função de ativação não linear. A introdução da não linearidade é essencial, pois permite que a rede represente relações complexas entre os dados. Funções de ativação como a sigmoide, a tangente hiperbólica e a *ReLU* são algumas das escolhas possíveis, de acordo com as características do problema e dos dados (33). Para problemas de classificação com  $K$  classes, como é o caso deste trabalho, a função de ativação sigmoide descrita na Equação 3.18 é frequentemente utilizada na camada de saída, pois mapeia os valores em um intervalo  $(0,1)$ , permitindo interpretá-los como probabilidades

$$\varphi(z_i) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}. \quad (3.18)$$

O processo de treinamento do MLP é realizado, em geral, por meio do algoritmo de retropropagação do erro (*backpropagation*), no qual o erro cometido na camada de saída é propagado de volta para atualizar os pesos da rede. Esse ajuste é feito utilizando métodos de otimização baseados em gradiente, como o gradiente descendente. Assim, a eficiência do processo depende fortemente da escolha de hiperparâmetros, como taxa de aprendizado, número de camadas e quantidade de neurônios por camada.

Os MLPs possuem uma alta capacidade de modelar relações não lineares complexas, servindo como ponto de partida para soluções mais sofisticadas do

aprendizado de máquina e da inteligência artificial, bem como sendo amplamente utilizados em problemas de classificação, regressão e reconhecimento de padrões em diversas áreas do conhecimento (34, 35, 36).

Apesar de suas vantagens, o MLP apresenta limitações importantes. Dentre elas, está a possibilidade de sobreajuste (*overfitting*), que ocorre quando o modelo tende a memorizar o conjunto de treinamento, perdendo capacidade de generalização para novos dados.

Neste contexto, a abordagem utilizando os Microestados de Recorrência ganha notoriedade ao fornecer uma representação compacta e informativa de séries temporais complexas, melhorando a capacidade de distinção entre sistemas dinâmicos, ainda que estes sejam descritos por parâmetros muito semelhantes (8).

## 4 METODOLOGIA

### 4.1 DADOS GERADOS

Para este trabalho foram gerados dados conforme os seguintes sistemas:

#### Mapa de Bernoulli Modificado

Referenciado como mapa  $\beta x$ , é um sistema dinâmico discreto unidimensional definido por

$$x_{n+1} = \beta x_n \pmod{1}, \quad (4.1)$$

onde o parâmetro  $\beta$  varia de acordo com  $1.99 \leq \beta \leq 4.99$ , exibindo comportamento caótico. É a generalização do *Bernoulli shift*, que corresponde ao caso particular onde  $\beta = 2$  (37).

#### Mapa Logístico

Originalmente introduzido no contexto da dinâmica populacional, descrevendo a evolução de uma população sujeita à limitação de recursos (38). É um sistema dinâmico discreto unidimensional definido por

$$x_{n+1} = rx_n(1 - x_n), \quad (4.2)$$

de modo que o parâmetro  $r$  varia de acordo com  $3.95 \leq r \leq 3.99$ , onde o sistema opera predominantemente em regime caótico, com algumas pequenas janelas periódicas.

#### Mapa de Hénon

É um sistema dinâmico discreto, bidimensional e não linear, introduzido por Michel Hénon em 1976 (39), conforme as seguintes equações

$$\begin{cases} x_{n+1} = 1 - ax_n^2 + y_n, \\ y_{n+1} = bx_n. \end{cases} \quad (4.3)$$

Neste caso, tem-se os parâmetros  $b = 0.3$  e  $1.1 \leq a \leq 1.2$ , implicando em um regime majoritariamente caótico.

### Sistema de Lorenz

É um sistema dinâmico contínuo de três equações diferenciais ordinárias, proposto por Edward N. Lorenz em 1963 para modelar de forma simplificada a convecção atmosférica (17). As seguintes EDOs descrevem o sistema

$$\begin{cases} \dot{x} &= \sigma(y - x), \\ \dot{y} &= x(\rho - z) - y, \\ \dot{z} &= xy - \beta z. \end{cases} \quad (4.4)$$

Seus parâmetros são  $\sigma = 10$ ,  $\beta = \frac{8}{3}$  e  $\rho$  variando conforme  $28.49 \leq \rho \leq 37.99$ , implicando predominantemente em um regime caótico.

### Ruídos Coloridos

A base teórica por trás dos ruídos coloridos foi estabelecida por Wiener em seu trabalho sobre processos estocásticos, onde modela matematicamente o movimento browniano (movimento aleatório de partículas suspensas em um fluido) via processo de Wiener (40). Nesse contexto, o ruído branco é definido formalmente como a derivada temporal do processo de Wiener.

Posteriormente, Mandelbrot e Van Ness generalizaram o movimento browniano clássico ao introduzir o movimento browniano fracionário (18), estabelecendo uma relação direta entre leis de potência no espectro de frequência e memória de longo alcance no tempo, dada por

$$S(f) \propto \frac{1}{f^\alpha}, \quad (4.5)$$

onde  $S(f)$  é a densidade espectral de potência,  $f$  é a frequência e  $\alpha$  é o expoente espectral, que quantifica o grau de correlação temporal do processo. Assim, diferentes valores de  $\alpha$  implicam em diferentes tipos de ruído. Por exemplo, se  $\alpha = 0$  obtém-se um ruído branco (totalmente aleatório, sem memória), se  $\alpha > 0$ , obtém-se ruídos correlacionados (memória de longo alcance), já se  $\alpha < 0$ , os ruídos são denominados anticorrelacionados (memória de curto alcance). Para este trabalho, utilizou-se  $\alpha \in [-2, 2]$ , gerando os dados conforme o trabalho de Jens Timmer e Michel König (41).

## 4.2 TRANSFORMAÇÃO DE ESPAÇO: DADOS BRUTOS PARA MICROESTADOS

Dados brutos provenientes de sistemas dinâmicos, sejam eles simulados numericamente ou obtidos experimentalmente, apresentam uma série de limitações que dificultam sua análise direta.

Dentre estes problemas estão a alta dimensionalidade, que aumenta consideravelmente o custo computacional, bem como a sensibilidade a ruídos, que pode mascarar a dinâmica do sistema. Especificamente no caso de sistemas dinâmicos caóticos, trajetórias que começam próximas divergem exponencialmente ao longo do tempo, dificultando a comparação direta entre os mesmos.

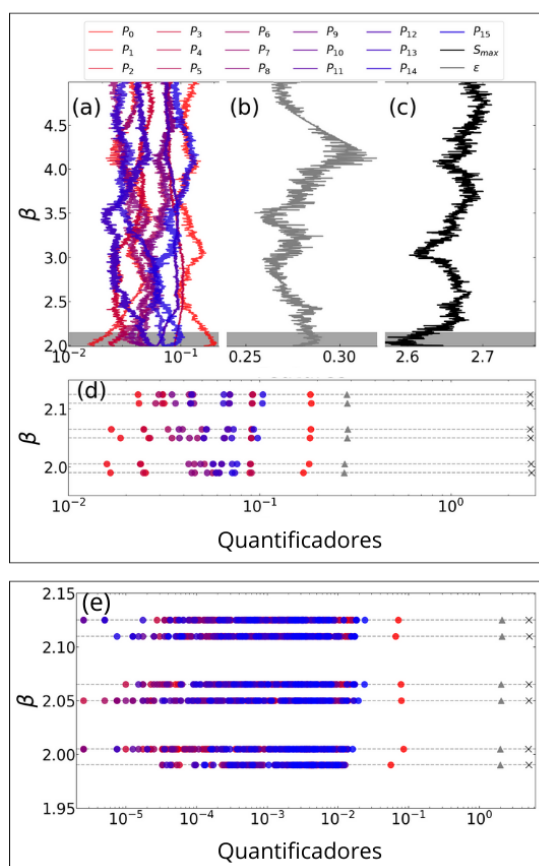
Do ponto de vista do aprendizado de máquina, o uso direto de dados brutos tende a um *overfitting* e a uma baixa capacidade de generalização. Assim, torna-se necessário um pré-processamento dos dados, reduzindo a dimensionalidade, aumentando a robustez ao ruído e capturando propriedades essenciais da dinâmica. Neste trabalho, este pré-processamento é feito através dos microestados de recorrência, descrito na subseção 3.2.3.

A Figura 5 apresenta os quantificadores provenientes desta nova representação para  $N = 2$ , especificamente para o mapa  $\beta x$ , sendo eles as probabilidades associadas a cada microestado (a), a entropia de recorrência (c) e o parâmetro  $\varepsilon$  que maximiza esta entropia (b). De modo geral, utilizam-se as probabilidades como dados de entrada da máquina.

Na parte (d) da Figura 5 está uma melhor visualização das diferenças dos quantificadores para três pares de  $\beta$ . Tais quantificadores representam as probabilidades associadas a cada microestados (círculos), a entropia máxima (letra "x") e o  $\varepsilon$  que maximiza esta entropia (triângulo). Nota-se que valores de  $\beta$  relativamente distantes apresentam grande diferença visual, enquanto valores dentro de um mesmo par se tornam mais difíceis de se diferenciar.

Ao recordarmos da subseção 3.2.3 que, quanto maior o tamanho dos microestados, maior o número de configurações possíveis, chegaremos à conclusão de que maior é o número de quantificadores. Em (e) da Figura 5, temos evidenciado esse grande aumento de quantificadores para  $N = 3$ , implicando em há  $N^* = 2^{3^2} = 512$  probabilidades. Assim, aumentar o tamanho dos microestados permite uma me-

FIGURA 5 – Representação visual dos quantificadores da abordagem via microestados de recorrência



Quantificadores oriundos da abordagem via microestados de recorrência. Primeiramente, para  $N = 2$ : No gráfico (a) estão as probabilidades associadas a cada microestado, no (b) está uma pequena faixa de variação do  $\epsilon$  e em (c) a entropia de recorrência, conforme varia-se o parâmetro  $\beta$  para o mapa  $\beta x$ . O gráfico (d) mostra cada um destes quantificadores para três pares de  $\beta$ , onde cada par corresponde a valores muito semelhantes entre si. Agora, para  $N=3$ : O gráfico (e) é análogo ao gráfico (d), porém, para um tamanho maior de microestados.

Fonte: Imagem retirada de (8).

lhor diferenciação dos parâmetros  $\beta$ , ainda que seus valores sejam relativamente próximos (8).

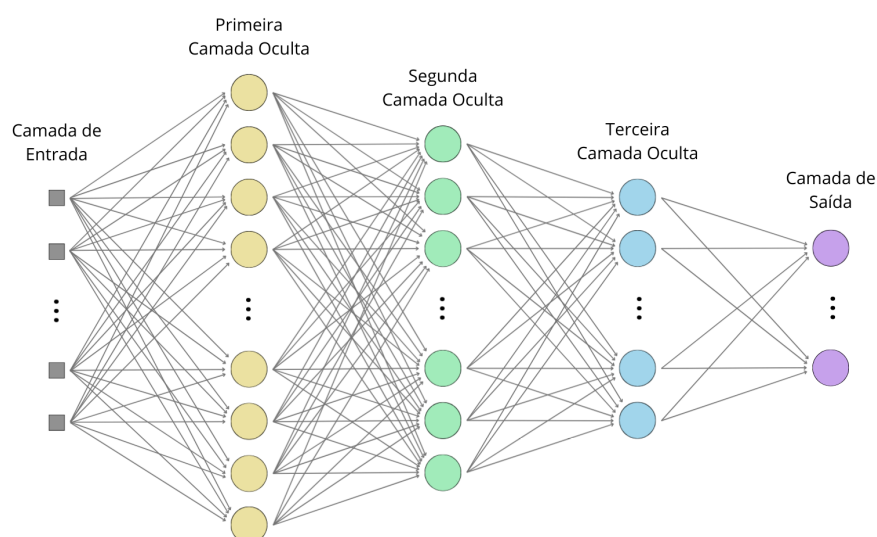
### 4.3 ARQUITETURA DA REDE DE APRENDIZADO DE MÁQUINA

Em todas as análises foram utilizadas redes do tipo *Multi-Layer Perceptron*, variando-se apenas o número de camadas ocultas e de neurônios por camada. O treinamento da rede neural foi realizado por meio do algoritmo de retropropagação do erro, implementado pela interface *Keras* e a biblioteca *TensorFlow*, juntamente com o otimizador Adam, adequado para problemas que são grandes em termos de dados e parâmetros.

#### Primeira arquitetura

Para se investigar o limiar de recorrência  $\varepsilon$ , utilizou-se uma rede com uma camada de entrada, três camadas ocultas e uma camada de saída. A camada de entrada possuía 64 neurônios e este número diminuía gradualmente em cada camada oculta, até chegar na camada de saída, que continha o número de neurônios equivalente ao número de possíveis classificações. Tal arquitetura pode ser entendida como um modelo de "funil", comprimindo progressivamente a informação, a fim de favorecer a generalização do modelo (7). A Figura 6 ilustra este modelo.

FIGURA 6 – Primeira arquitetura



Primeira arquitetura de rede neural abordada, constituída por um MLP com uma camada de entrada, três camadas ocultas e uma camada de saída.

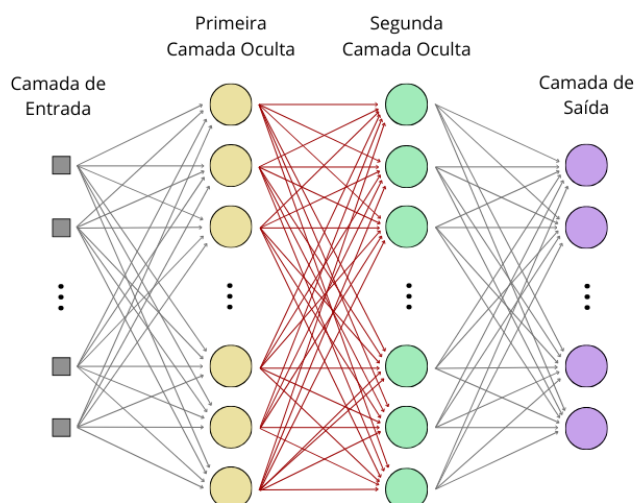
Fonte: Elaborado pela própria autora (2026).

Essa primeira análise se fez por meio de 100 épocas de treinamento, a uma taxa de aprendizagem de 0.001.

### Segunda arquitetura

Para se investigar a evolução da matriz de conexões entre os neurônios ao longo das épocas de treinamento utilizou-se uma rede mais simples, com uma camada de entrada, duas camadas ocultas e uma camada de saída, conforme a Figura 7. Ambas as camadas ocultas possuíam 64 neurônios, a fim de se gerar uma matriz de pesos quadrada.

FIGURA 7 – Segunda arquitetura



Segunda arquitetura de rede neural abordada, constituída por um MLP com uma camada de entrada, duas camadas ocultas e uma camada de saída.

Fonte: Elaborado pela própria autora (2026).

As conexões destacadas em vermelho na Figura 7 estão representadas na Figura 8 onde cada elemento  $w_{k,n}$  é o peso da conexão entre o neurônio  $k$  da primeira camada oculta com o neurônio  $n$  da segunda camada oculta.

Esta estrutura foi utilizada para um treinamento em 1000 épocas, com uma taxa de aprendizado de 0.0001. Cada uma das épocas resultava em uma matriz de pesos (Figura 8) diferente, permitindo uma análise da "dinâmica" da mesma.

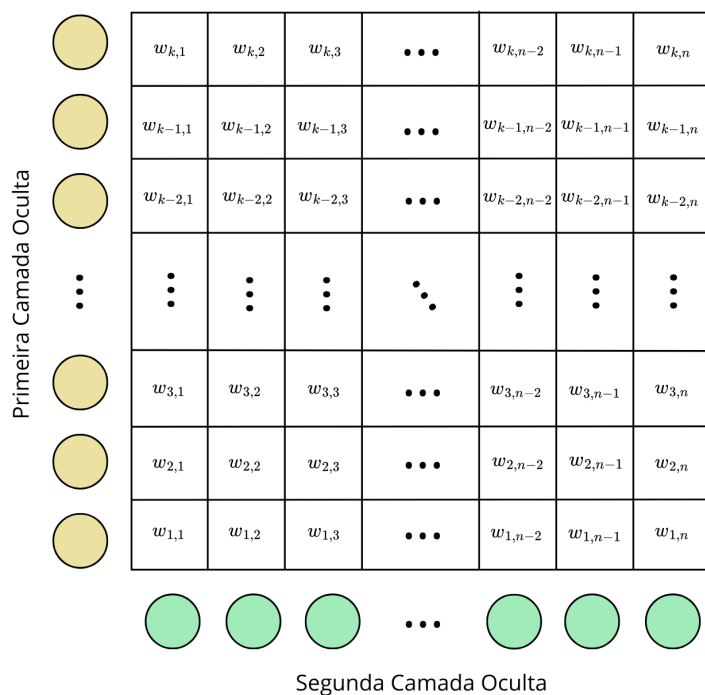


FIGURA 8 – Matriz de conexões entre os neurônios da primeira camada oculta (amarelo) e da segunda camada oculta (verde).

## 5 RESULTADOS E DISCUSSÃO

Nesta seção são apresentados e discutidos os principais resultados obtidos ao longo deste trabalho. A análise é conduzida com foco em dois aspectos centrais:

- A relação entre o limiar de recorrência  $\varepsilon$ , a entropia de microestados e a acurácia de classificação do algoritmo de aprendizagem de máquina;
- A investigação da estrutura interna do aprendizado por meio da análise da matriz de pesos da rede neural.

Conceitualmente, o aprendizado de máquina é empregado neste trabalho como um instrumento para avaliar a qualidade, do ponto de vista informacional, das representações baseadas em microestados de recorrência. Nesse contexto, um melhor desempenho de classificação indica a preservação características dinâmicas relevantes do sistema. Além da análise do desempenho, busca-se compreender de que maneira diferentes representações influenciam a dinâmica interna da rede neural, evidenciando os fatores que tornam a abordagem baseada em microestados particularmente eficaz.

### 5.1 ANÁLISE DO LIMIAR DE RECORRÊNCIA

Primeiramente foi analisada a acurácia da máquina ao classificar os parâmetros dos três mapas caóticos via aprendizado supervisionado. Para este fim, foi utilizada a arquitetura exemplificada na Figura 6, com séries temporais de  $K = 1000$  pontos, considerando-se um amplo intervalo do limiar de recorrência  $0 < \varepsilon < 0.5$ , com o objetivo de se analisar sua relação com a acurácia da máquina.

Foram considerados três sistemas dinâmicos caóticos discretos, sendo eles o mapa  $\beta x$  (Equação 4.1), o mapa logístico (Equação 4.2) e o mapa de Hénon (Equação 4.3), com diferentes números de classes, variando entre 10, 20 e 40 valores igualmente espaçados de seus respectivos parâmetros de controle:  $\beta$  para o mapa  $\beta x$ ,  $r$  para o mapa Logístico e  $a$  para o mapa de Hénon.

A partir disso, foram obtidos conjuntos de dados com  $W = 400$ ,  $W = 800$  e  $W = 1600$  sistemas, permitindo avaliar a robustez dos resultados frente ao aumento da complexidade do problema de classificação.

A Figura 9 sumariza estas análises em seus gráficos de acurácia em conjunto com os mapas de calor. As linhas distinguem os diferentes mapas investigados, sendo (a1–f1) associados ao mapa  $\beta x$ , (a2–f2) ao mapa logístico e (a3–f3) ao mapa de Hénon. Já as colunas diferenciam as configurações relativas ao número de sistemas e à divisão em classes, de modo que os painéis (a) e (d) correspondem a 400 sistemas distribuídos em 10 classes, (b) e (e) a 800 sistemas distribuídos em 20 classes e (c) e (f) à 1600 sistemas distribuídos em 40 classes.

Observa-se que, embora a acurácia da máquina diminua à medida que o número de parâmetros aumenta, os melhores resultados de classificação ocorrem sistematicamente para valores de  $\varepsilon$  associados à máxima entropia de recorrência, mostrando uma forte correlação entre essas duas quantidades.

Pode-se interpretar, portanto, a entropia como um indicador da quantidade de informação relevante (42), estabelecendo um critério objetivo para a escolha do parâmetro  $\varepsilon$  e garantindo maior eficiência ao se analisar os dados. O desenrolar desta análise, feita durante o mestrado da presente autora, resultou no artigo (8).

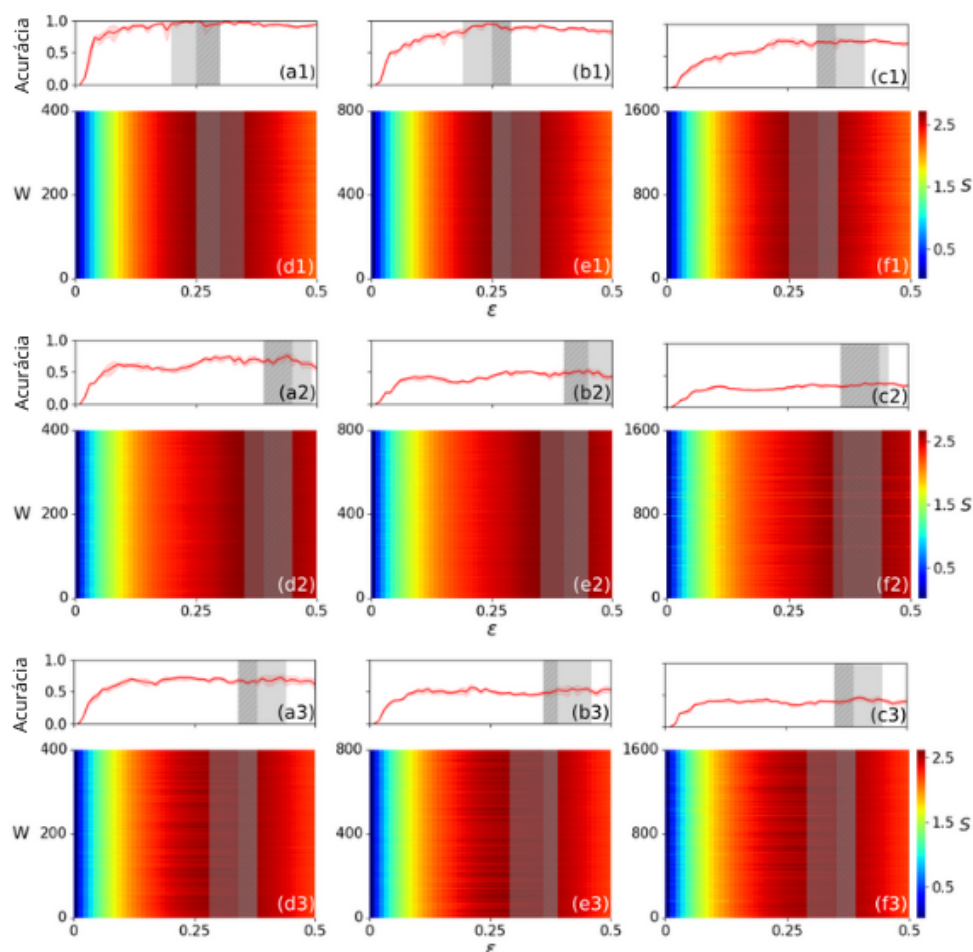
## 5.2 ANÁLISE DA MATRIZ DE PESOS

Com o objetivo de investigar o processo de aprendizado da rede neural, foi realizada uma análise da evolução da matriz de pesos que conecta as duas camadas ocultas de uma arquitetura simplificada de MLP, conforme o que está representado na Figura 7.

Adotou-se, portanto, o mesmo número de neurônios nas duas camadas ocultas ( $k = n = 64$ ), garantindo que a matriz de pesos fosse quadrada ( $64 \times 64$ ). Embora essa configuração não seja muito eficiente, ela permite uma visualização mais direta da estrutura da matriz de pesos, apresentada na Figura 8. O treinamento foi realizado ao longo de 1000 épocas, utilizando uma taxa de aprendizado reduzida ( $10^{-4}$ ), de modo a acompanhar de forma mais gradual a evolução das conexões sinápticas.

A Figura 10 mostra a evolução da matriz de pesos ao longo das épocas de

FIGURA 9 – Acurácia e entropia em função do parâmetro  $\varepsilon$



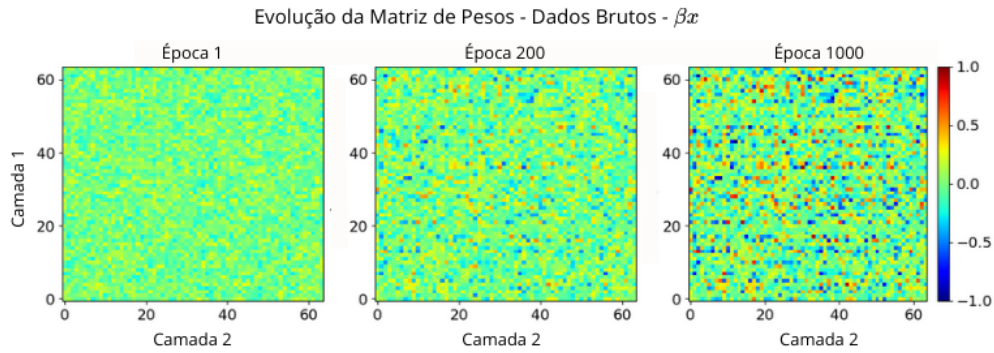
Acurácia da máquina e entropia de recorrência em função do limiar de recorrência  $\varepsilon$ . As áreas sombreadas e tracejadas indicam a sobreposição entre os intervalos de máxima entropia de recorrência e máxima acurácia. As linhas (1), (2), e (3) correspondem aos sistemas  $\beta x$ , Logístico e Hénon, respectivamente. Já as colunas (a,d), (b,e) e (c,f) correspondem a diferentes números de sistemas e classes, sendo eles  $W = 400$  e  $Labels = 10$ ,  $W = 800$  e  $Labels = 20$ ,  $W = 1600$  e  $Labels = 40$ , respectivamente. Em todos os casos, foram utilizados como dados de entrada da rede as probabilidades associadas a cada microestado, o parâmetro  $\varepsilon$  e a entropia.

Fonte: Imagem retirada de (8).

treinamento ao utilizar os dados brutos do sistema  $\beta x$  como dados de entrada da rede. Já a Figura 11, mostra a evolução da matriz de pesos ao longo das épocas

de treinamento ao utilizar as probabilidades associadas a cada microestado, para microestados de tamanho  $N = 2$ ,  $N = 3$  e  $N = 4$ , do sistema  $\beta x$  como dados de entrada da rede.

FIGURA 10 – Evolução da matriz de pesos ao se utilizar os dados brutos como dados de entrada da rede neural



Evolução da matriz de pesos da rede indicada na Figura 7, ao longo das 1000 épocas de treinamento, utilizando-se os dados brutos do sistema  $\beta x$  como dados de entrada.

Fonte: Elaborado pela própria autora (2026).

Observa-se que, nas primeiras épocas, os pesos apresentam uma distribuição aparentemente homogênea de pesos, refletindo o caráter aleatório da inicialização da rede. À medida que o treinamento progride, a rede passa a privilegiar determinadas conexões em detrimento de outras.

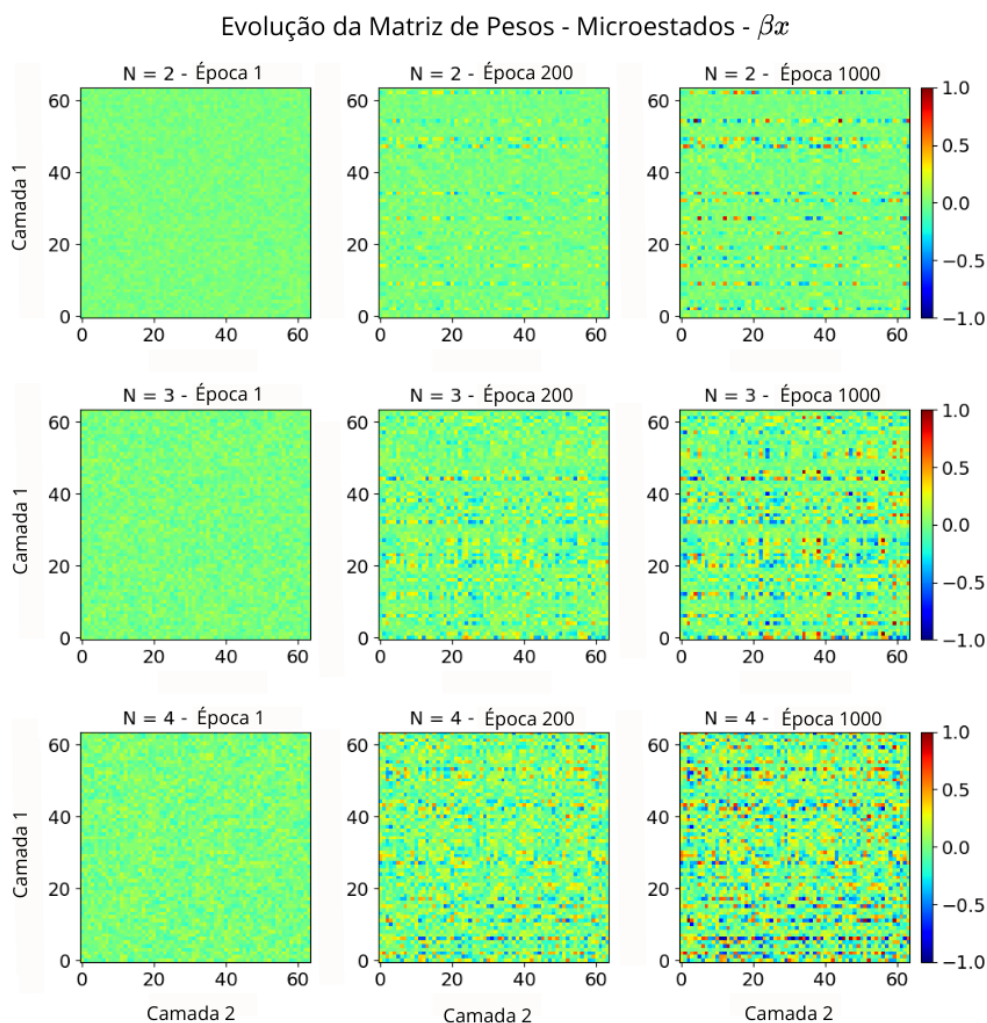
O interessante é que, ao se utilizar as probabilidades associadas a cada microestado de recorrência como dados de entrada, há o surgimento de padrões mais estruturados, concentrados em certas regiões da matriz de pesos, quando comparado ao uso direto de séries temporais brutas. Adicionalmente, para valores maiores de  $N$ , os padrões ficam ainda mais evidentes.

Esse processo pode ser interpretado como uma reorganização interna da rede, reforçando a ideia de que essa representação fornece informações diferenciadas sobre a dinâmica dos sistemas analisados.

### 5.2.1 Microestados de recorrência como uma solução para o overfitting

Seguindo com a abordagem da rede da Figura 7, gerou-se o gráfico da acurácia de treinamento ao longo das épocas. A Figura 13 ilustra para o sistema  $\beta x$ .

FIGURA 11 – Evolução da matriz de pesos ao se utilizar os quantificadores da abordagem via microestados de recorrência como dados de entrada da rede neural



Evolução da matriz de pesos da rede indicada na Figura 7, ao longo das 1000 épocas de treinamento, utilizando-se as probabilidades associadas a cada microestado do sistema  $\beta x$  como dados de entrada.

Fonte: Elaborado pela própria autora (2026).

A análise dessa figura mostra que, para os dados brutos, ainda que a acurácia de treinamento se mostre elevada, há quedas bruscas, indicando que a rede aprende padrões muito específicos de subconjuntos dos dados, não construindo uma representação global estável. Em contraste, quando se utilizam os micros-

FIGURA 12 – Acurácia de treinamento

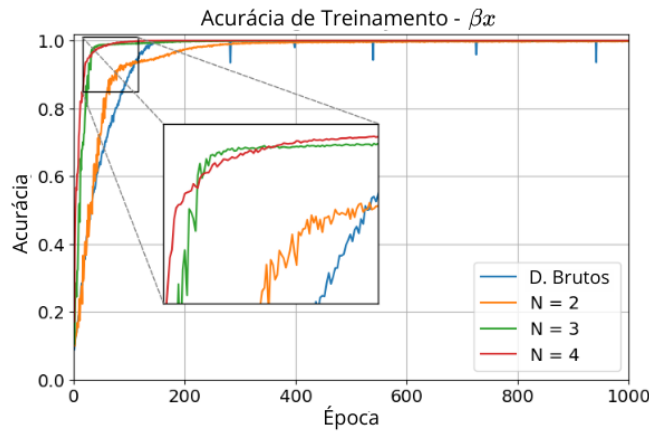


FIGURA 13 – Acurácia de treinamento do sistema  $\beta x$ , utilizando os dados brutos (D. Brutos), bem como as probabilidades associadas a diferentes tamanhos de microestados ( $N = 2$ ,  $N = 3$  e  $N = 4$ ) como dados de entrada.

Fonte: Elaborado pela própria autora (2026).

tados de recorrência, observa-se uma acurácia de treinamento consistentemente elevada, acompanhada de uma convergência mais rápida do processo de aprendizado, evidenciando que essa representação fornece informações mais robustas sobre a dinâmica do sistema. Essa premissa se confirma com a Tabela 2, que mostra as diferenças entre as acurácias de teste dos dados brutos e dos diferentes tamanhos de microestados. As figuras para os demais sistemas se encontram na Figura 19 presente no Apêndice A.

TABELA 2 – Acurácias de teste

Tipo de Dado	$\beta x$	Logístico	Lorenz	Ruídos Coloridos
Dados Brutos	9.80%	10.49%	10.89%	27.00%
$N = 2$	99.59%	91.60%	51.10%	96.59%
$N = 3$	99.89%	99.29%	100.00%	98.39%
$N = 4$	99.69%	99.69%	100.00%	96.29%

Acurácias de teste para cada sistema e tipo de dado.

Fonte: Elaborado pela própria autora (2026).

Observa-se que a utilização dos microestados de recorrência promove um aumento expressivo na acurácia do algoritmo de classificação, ou seja, embora tanto

os dados brutos quanto os microestados de recorrência tenham apresentado elevado desempenho durante a fase de treinamento, a redução significativa da acurácia no conjunto de teste para os dados brutos indica a ocorrência de sobreajuste (*overfitting*).

O fato de elevados desempenhos serem obtidos mesmo com arquiteturas relativamente simples de redes neurais sugere que o ganho observado está exclusivamente associado à qualidade da representação dos dados, e não à complexidade da rede, reforçando o papel central dos microestados na captura das propriedades dinâmicas do sistema.

### 5.2.2 Desordem na dinâmica dos pesos

Com o objetivo de investigar de forma mais profunda a organização interna do processo de aprendizado, analisou-se a dinâmica temporal dos pesos da rede neural ao longo das épocas de treinamento por meio do quantificador de desordem baseado em microestados de recorrência (DISREM), conforme descrito na subseção 3.2.5.

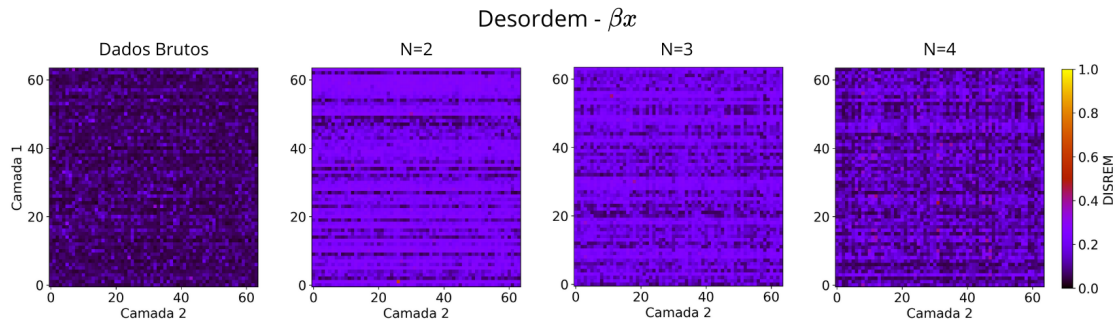
No contexto do presente trabalho, cada elemento da matriz de pesos é tratado como uma série temporal, cuja evolução ocorre ao decorrer das 1000 épocas de treinamento. Dessa forma, a aplicação do DISREM às trajetórias temporais dos pesos permite quantificar o grau de organização ou correlação introduzido pelo aprendizado em cada conexão entre as camadas ocultas da rede.

De acordo com a definição do DISREM, valores elevados do quantificador indicam que os microestados de recorrência associados à série são aproximadamente equiprováveis, caracterizando um comportamento compatível com independência estatística entre eventos. Em contrapartida, valores reduzidos de desordem refletem a quebra dessa simetria estatística, indicando a presença de correlações temporais e estrutura na dinâmica analisada.

Assim, ao aplicar o DISREM à evolução dos pesos, valores baixos de desordem podem ser interpretados como evidência de que o respectivo peso participa ativamente de um processo de aprendizado estruturado, enquanto valores elevados sugerem uma adaptação difusa ou marginal. A Figura 14 mostra a desordem calculada para  $\mathcal{M}$  de tamanho  $4 \times 4$ , conforme a Equação 3.8, para o sistema  $\beta x$ .

Percebe-se que conexões mais estáveis apresentam valores reduzidos de desordem.

FIGURA 14 – Desordem calculada sobre os elementos da matriz de pesos



Desordem de cada elemento da matriz de pesos, ao longo das 1000 épocas de treinamento. Os títulos de cada gráfico representam os respectivos dados de entrada da rede, para o sistema  $\beta x$ .

Fonte: Elaborado pela própria autora (2026).

Quando os dados brutos são utilizados como dados de entrada, esse comportamento ocorre de forma generalizada em toda a rede, de modo que o processo de aprendizado tende a se tornar excessivamente especializado em padrões específicos dos dados de treinamento, reduzindo a capacidade de generalização do modelo.

Ao utilizar a representação baseada em microestados de recorrência, observa-se que apenas um subconjunto das conexões apresenta baixa desordem. Essas conexões passam a concentrar a informação relevante da dinâmica, enquanto as demais mantêm maior variabilidade, permitindo que a rede preserve flexibilidade e capacidade de generalização.

Além disso, à medida que o número de microestados aumenta, observa-se uma concentração progressiva de conexões de baixa desordem. Esse efeito promove uma especialização ainda maior da rede, conduzindo a um aumento da acurácia de classificação, enquanto preserva simultaneamente a capacidade de generalização e a redução do risco de *overfitting*.

Esses resultados reforçam a ideia de que o aprendizado em redes neurais pode ser interpretado como um processo dinâmico no qual, ao se utilizar o método dos microestados de recorrência, a organização das conexões emerge da interação entre a estrutura estatística dos dados de entrada e os mecanismos de adaptação da rede, implicando em uma distribuição não homogênea da desordem e refletindo a presença de padrões dinâmicos relevantes. Assim, a análise baseada no DISREM

---

fornece uma ferramenta conceitualmente consistente e quantitativamente robusta para investigar a organização interna do aprendizado, complementando as métricas tradicionais de desempenho e acurácia.

## 6 CONSIDERAÇÕES FINAIS

Neste trabalho, investigou-se o uso dos microestados de recorrência como uma ferramenta para a classificação de sistemas dinâmicos por meio de técnicas de aprendizado de máquina. A abordagem proposta baseia-se na extração de padrões locais dos gráficos de recorrência e na análise estatística desses padrões, permitindo uma representação compacta e informativa das dinâmicas relativas às séries temporais. Ao longo do estudo, demonstrou-se que essa representação preserva informações essenciais da dinâmica, mesmo em cenários caracterizados por alta complexidade.

Os resultados obtidos para diferentes sistemas dinâmicos discretos evidenciaram que há uma forte correlação entre o valor do limiar de recorrência que maximiza a entropia de microestados e aquele que maximiza a acurácia de classificação das redes neurais. Esse comportamento consistente reforça o papel da entropia como um critério objetivo e fisicamente interpretável como informação para a escolha de parâmetros no processo de análise e classificação.

A investigação da dinâmica da matriz de pesos revelou padrões distintos de aprendizado dependendo do tipo de dado utilizado como entrada da rede, enquanto o quantificador de desordem forneceu uma interpretação adicional sobre a organização interna desse processo de aprendizado.

Quando os dados brutos são utilizados como entrada, observa-se uma redução generalizada da desordem, caracterizando um regime de superespecialização e comprometendo a capacidade de generalização do modelo. Em contraste, a representação baseada em microestados de recorrência induz uma especialização seletiva das conexões, preservando maior variabilidade nas demais. Essa organização resulta em um equilíbrio entre especialização e flexibilidade, permitindo que a rede neural alcance maiores valores de acurácia.

Esses resultados reforçam a interpretação do aprendizado em redes neurais como um processo dinâmico de auto-organização, fortemente condicionado pela estrutura estatística dos dados de entrada. A introdução dos microestados de recorrência modifica qualitativamente esse processo, evidenciando padrões dinâmicos relevantes

que orientam o aprendizado.

Por fim, este trabalho contribui para a interface entre sistemas dinâmicos não lineares e aprendizado de máquina ao propor uma metodologia que integra fundamentos físicos, estatísticos e computacionais. Como perspectivas futuras, destacam-se a aplicação da abordagem a outros sistemas dinâmicos e dados experimentais, bem como a investigação de diferentes arquiteturas de redes neurais.

## REFERÊNCIAS

- 1 STROGATZ, S. H. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. [S.l.]: Perseus Books Publishing, 1994. (Studies in nonlinearity).
- 2 TAKENS, F. Detecting strange attractors in turbulence. In: RAND, D.; YOUNG, L.-S. (Ed.). *Dynamical Systems and Turbulence*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1981. p. 366–381. ISBN 978-3-540-38945-3.
- 3 WOLF, A.; SWIFT, J. B.; SWINNEY, H. L.; VASTANO, J. A. Determining Lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena*, v. 16, n. 3, p. 285–317, 1985. ISSN 0167-2789.
- 4 ECKMANN, J.-P.; KAMPHORST, S. O.; RUELLE, D. Recurrence plots of dynamical systems. *Europhysics Letters*, v. 4, n. 9, p. 973, nov 1987. Disponível em: <https://doi.org/10.1209/0295-5075/4/9/004>.
- 5 PRADO, T. L.; CORSO, G.; LIMA, G. Z. dos S.; BUDZINSKI, R. C.; BOARETTO, B. R. R.; FERRARI, F. A. S.; MACAU, E. E. N.; LOPES, S. R. Maximum entropy principle in recurrence plot analysis on stochastic and chaotic systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, v. 30, n. 4, p. 043123, 2020.
- 6 FLAUZINO, J. V. V.; PRADO, T. L.; MARWAN, N.; KURTHS, J.; LOPES, S. R. Quantifying disorder in data. *Phys. Rev. Lett.*, American Physical Society (APS), v. 135, n. 9, p. 097401, aug 2025.
- 7 HAYKIN, S. O. *Neural Networks and Learning Machines*. 3. ed. Upper Saddle River, NJ: Pearson, 2008.
- 8 SPEZZATTO, G. S.; FLAUZINO, J. V. V.; CORSO, G.; BOARETTO, B. R. R.; MACAU, E. E. N.; PRADO, T. L.; LOPES, S. R. Recurrence microstates for machine learning classification. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, v. 34, n. 7, p. 073140, 07 2024. ISSN 1054-1500.
- 9 GRANT, E. *A History of Natural Philosophy, From the Ancient World to the Nineteenth Century*. [S.l.]: Cambridge University Press, 2007.
- 10 BARNES, J. *The Presocratic Philosophers*. London, England: Routledge, 1982. (Arguments of the Philosophers).
- 11 EUCLIDES. *Os elementos*. [S.l.]: Unesp, 2009.

- 12 COPÉRNICO, N. *De Revolutionibus Orbium Coelestium*. Nuremberg: Petreius, J., 1543.
- 13 DAVIS, A. E. L. Kepler translated. *The British Journal for the History of Science*, v. 28, n. 3, p. 345–349, 1995.
- 14 NEWTON, I. *Philosophiæ naturalis principia mathematica*. London: Apud G. & J. Innys. (Landmarks of science, v. 1). Acesso em: 22 dez. 2025. Disponível em: <https://cudl.lib.cam.ac.uk/view/PR-ADV-B-00039-00001/1>.
- 15 BROWN, R. J. *A modern introduction to dynamical systems*. London, England: Oxford University Press, 2018.
- 16 POINCARÉ, H. *Les Méthodes Nouvelles de la Mécanique Céleste. Tome I*. Paris: Gauthier-Villars, 1890. Digitalização disponível no acervo do Institut Mittag-Leffler. Disponível em: <https://www.mittag-leffler.se/app/uploads/2023/03/finalmemoir.pdf>.
- 17 LORENZ, E. N. Deterministic nonperiodic flow. *J. Atmos. Sci.*, American Meteorological Society, v. 20, n. 2, p. 130–141, mar 1963.
- 18 MANDELROT, B. B.; NESS, J. W. V. Fractional brownian motions, fractional noises and applications. *SIAM Review*, SIAM, v. 10, n. 4, p. 422–437, 1968.
- 19 SMALE, S. Differentiable dynamical systems. In: *The Mathematics of Time: Essays on Dynamical Systems, Economic Processes, and Related Topics*. New York: Springer, 1980. p. 1–82.
- 20 GRASSBERGER, P.; PROCACCIA, I. Measuring the strangeness of strange attractors. *Physica D: Nonlinear Phenomena*, v. 9, n. 1, p. 189–208, 1983. ISSN 0167-2789.
- 21 MARWAN, N.; ROMANO, M. C.; THIEL, M.; KURTHS, J. Recurrence plots for the analysis of complex systems. *Physics Reports*, v. 438, n. 5-6, p. 237–329, jan. 2007. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0370157306004066>.
- 22 ZBILUT, J. P.; GIULIANI, A.; WEBBER, C. L. Recurrence quantification analysis and principal components in the detection of short complex signals. *Physics Letters A*, v. 237, n. 3, p. 131–135, 1998. ISSN 0375-9601. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0375960197008438>.

- 23 ROLINK, J.; KUTZ, M.; FONSECA, P.; LONG, X.; MISGELD, B.; LEONHARDT, S. Recurrence quantification analysis across sleep stages. *Biomedical Signal Processing and Control*, v. 20, p. 107–116, 2015. ISSN 1746-8094. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1746809415000634>.
- 24 CORSO, G.; PRADO, T. L.; LIMA, G. Z. S.; KURTHS, J.; LOPES, S. R. Quantifying entropy using recurrence matrix microstates. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, AIP Publishing, v. 28, n. 8, 2018.
- 25 SHANNON, C. E. A mathematical theory of communication. *The Bell System Technical Journal*, v. 27, n. 3, p. 379–423, July 1948. Reprinted with corrections in October 1948, pp. 623–656.
- 26 GOLDSTINE, H. H. *The Computer from Pascal to von Neumann*. Paperback. Princeton: Princeton University Press, 1980. 5th paperback printing with a new preface, 1993. ISBN 0-691-02367-0.
- 27 MUSEUM, C. H. *Timeline of Computer History*. 2025. <https://www.computerhistory.org/timeline/>. Acesso em: 12 dez. 2025.
- 28 TURING, A. M. On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society*, p. 230–265, 1936.
- 29 CERUZZI, P. E. *A history of modern computing*. 2. ed. London, England: MIT Press, 2003. (History of Computing).
- 30 MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, Springer Nature, v. 5, p. 115–133, 1943.
- 31 ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol. Rev.*, American Psychological Association (APA), v. 65, n. 6, p. 386–408, nov 1958.
- 32 RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. *Nature*, Springer Science and Business Media LLC, v. 323, n. 6088, p. 533–536, oct 1986.
- 33 Keras. *Activation Functions – Keras API*. 2026. <https://keras.io/api/layers/activations/>. Acesso em: 10 mar. 2026.

- 34 PATEL, J. L.; GOYAL, R. K. Applications of artificial neural networks in medical science. *Curr. Clin. Pharmacol.*, Bentham Science Publishers Ltd., v. 2, n. 3, p. 217–226, sep 2007.
- 35 CARRASQUILLA, J.; MELKO, R. G. Machine learning phases of matter. *Nat. Phys.*, Springer Science and Business Media LLC, v. 13, n. 5, p. 431–434, may 2017.
- 36 GUEST, D.; CRANMER, K.; WHITESON, D. Deep learning and its application to LHC physics. *Annu. Rev. Nucl. Part. Sci.*, Annual Reviews, v. 68, n. 1, p. 161–181, oct 2018.
- 37 OTT, E. *Chaos in Dynamical Systems*. 2. ed. Cambridge, England: Cambridge University Press, 2002.
- 38 MAY, R. M. Simple mathematical models with very complicated dynamics. *Nature*, Springer Science and Business Media LLC, v. 261, n. 5560, p. 459–467, jun 1976.
- 39 HÉNON, M. A two-dimensional mapping with a strange attractor. *Commun. Math. Phys.*, Springer Science and Business Media LLC, v. 50, n. 1, p. 69–77, feb 1976.
- 40 WIENER, N. Differential-space. *Journal of Mathematics and Physics*, v. 2, n. 1-4, p. 131–174, 1923. Acesso em: 15 dez. 2025. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/sapm192321131>.
- 41 TIMMER, J.; KOENIG, M. On generating power law noise. *Astronomy and Astrophysics*, v. 300, p. 707–710, 1995.
- 42 BRILLOUIN, L. *Science and Information Theory*. [S.l.]: Dover Publications, 2004. (Dover Phoenix Editions). ISBN 9780486439181.

## APÊNDICE A – RESULTADOS ADICIONAIS

Conforme a análise descrita na seção 5.2, a Figura 15 representa a evolução das matrizes de pesos da rede indicada na Figura 7 ao longo das 1000 épocas de treinamento, utilizando-se os dados brutos como dados de entrada, para os sistemas Logístico, Lorenz e para os Ruídos Coloridos, respectivamente.

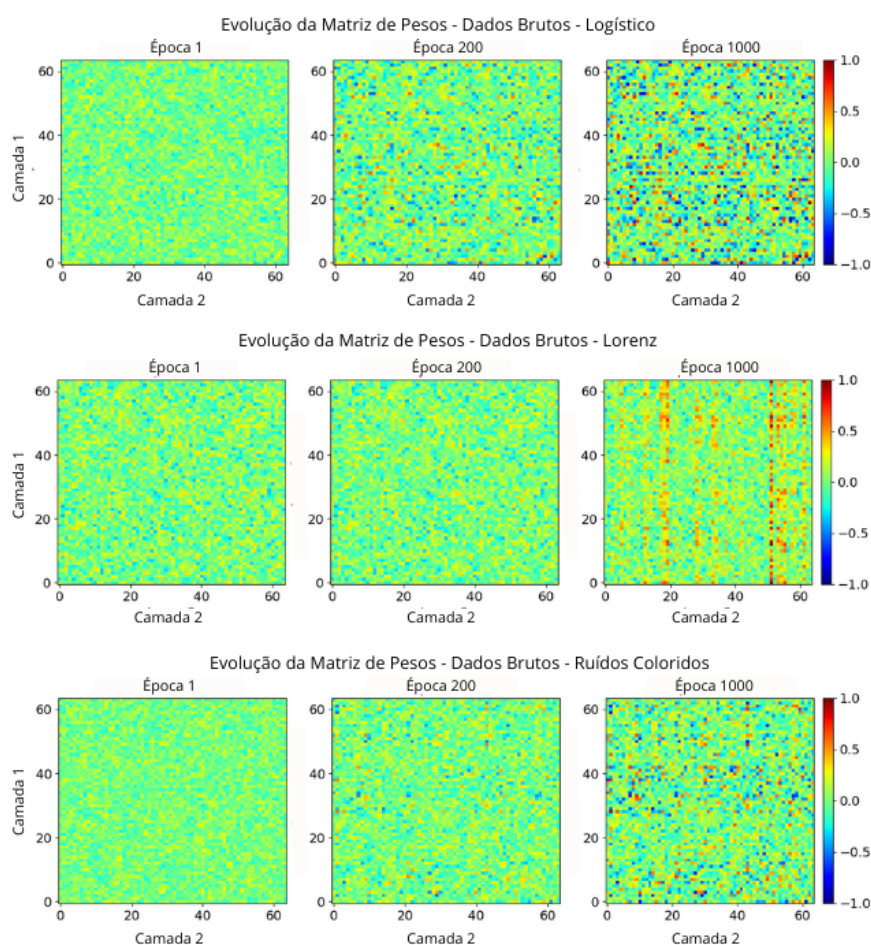


FIGURA 15 – Evolução da matriz de pesos da rede indicada na Figura 7, ao longo das 1000 épocas de treinamento, utilizando-se os dados brutos dos sistemas Logístico, Lorenz e Ruídos Coloridos como dados de entrada.

Observando estas matrizes, é possível perceber que a variação dos pesos utilizando-se os dados brutos como dados de entrada, ocorre de forma bem dis-

tribuída por toda a rede. A única exceção a este caso ocorre para o sistema de Lorenz, provavelmente pela escolha do integrador ou do passo para a análise de microestados, uma vez que trata-se de um sistema contínuo.

As imagens da Figura 16, mostra a evolução da matriz de pesos ao longo das 1000 épocas de treinamento da rede indicada na Figura 7, ao se utilizar as probabilidades associadas a cada microestado para o mapa Logístico. Cada coluna representa uma época escolhida e cada linha o tamanho do microestado:  $N = 2$ ,  $N = 3$  e  $N = 4$ .

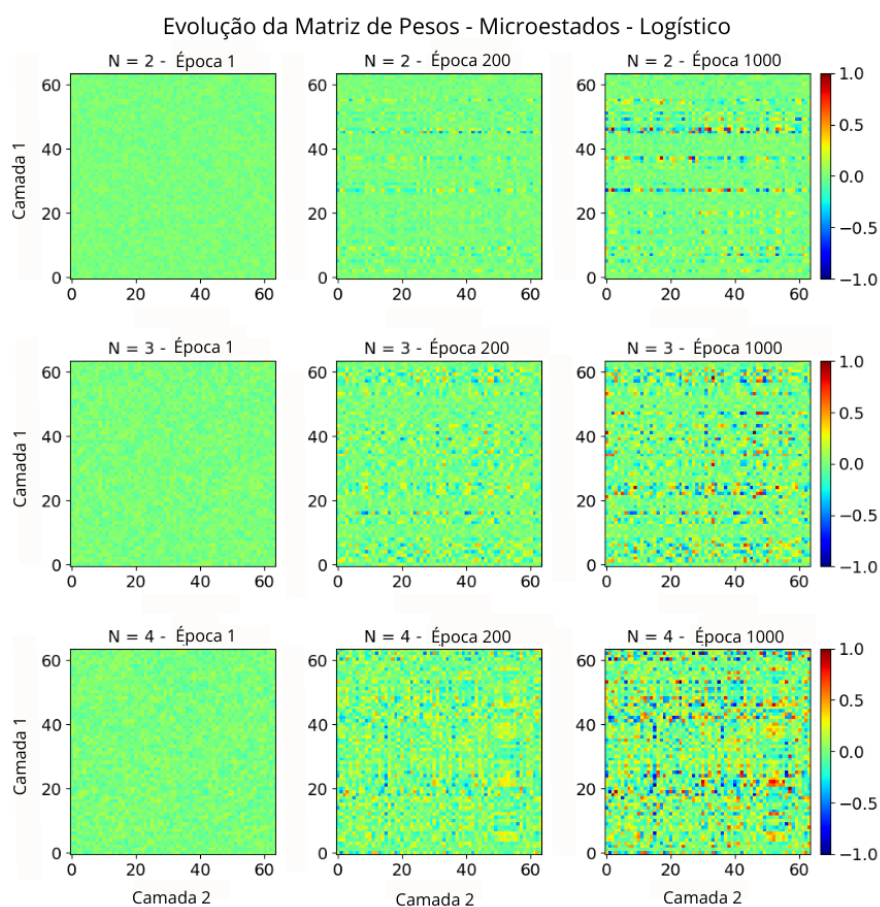


FIGURA 16 – Evolução da matriz de pesos da rede indicada na Figura 7, ao longo das 1000 épocas de treinamento, utilizando-se as probabilidades associadas a cada microestado do mapa Logístico como dados de entrada.

De forma análoga à descrição da Figura 16, a Figura 17 e a Figura 18 apresentam a evolução das matrizes de pesos para o sistema contínuo Lorenz e para os ruídos

coloridos, respectivamente.

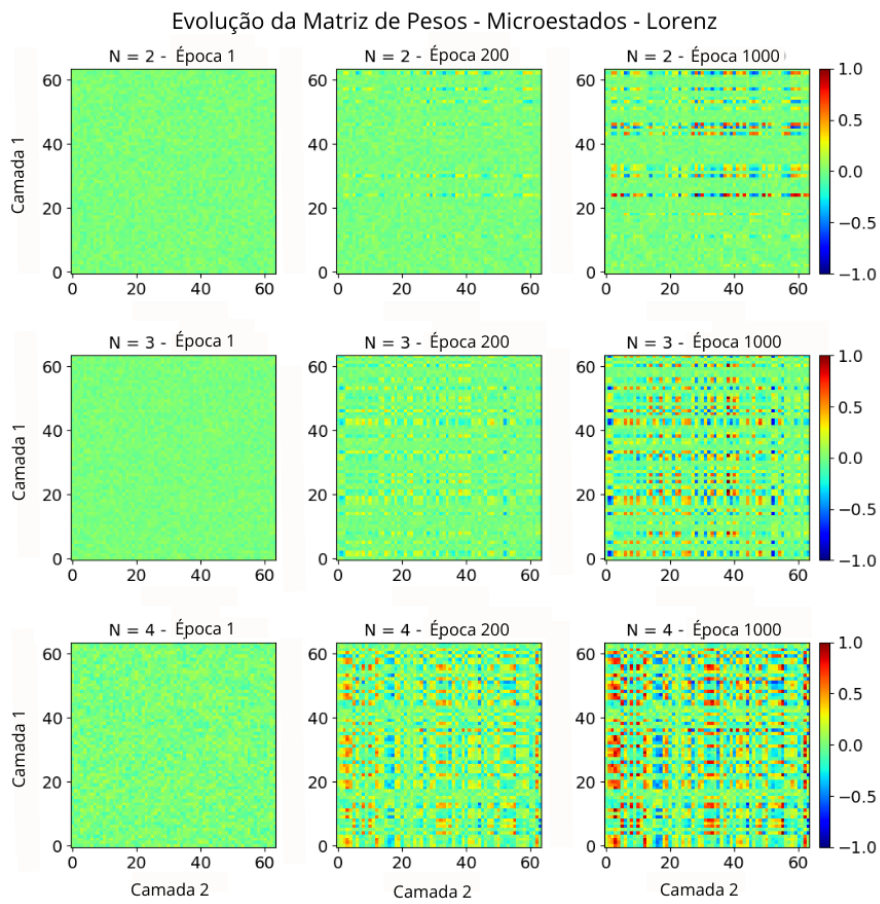


FIGURA 17 – Evolução da matriz de pesos da rede indicada na Figura 7, ao longo das 1000 épocas de treinamento, utilizando-se as probabilidades associadas a cada microestado do sistema Lorenz como dados de entrada.

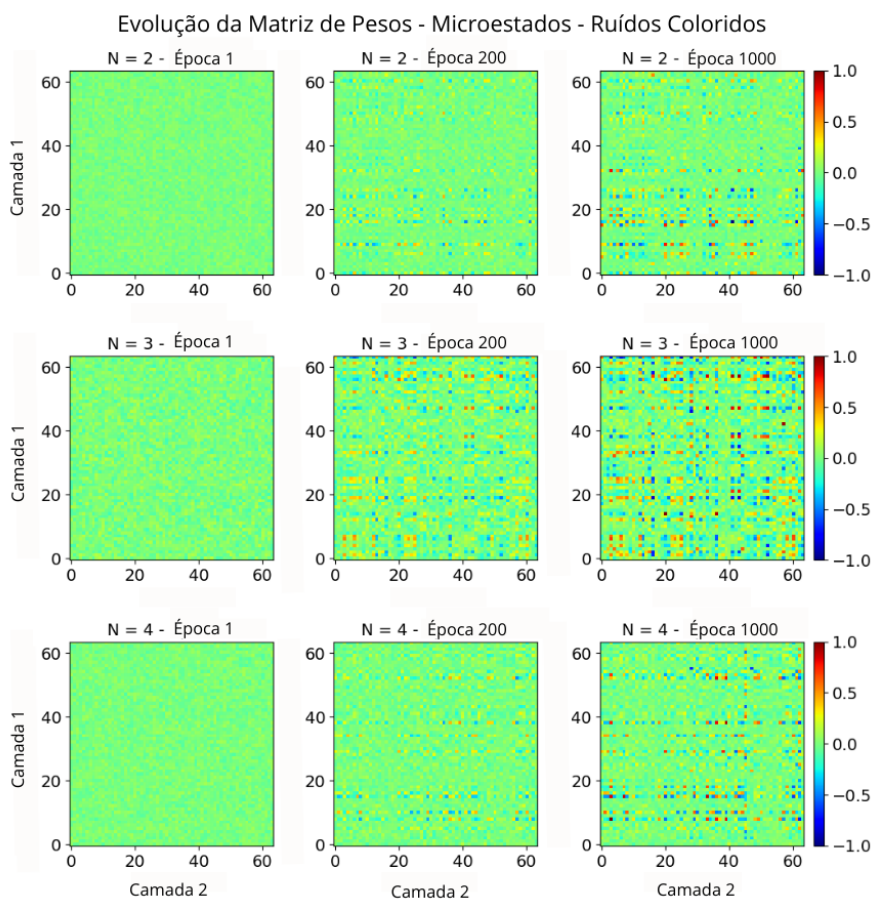


FIGURA 18 – Evolução da matriz de pesos da rede indicada na Figura 7, ao longo das 1000 épocas de treinamento, utilizando-se as probabilidades associadas a cada microestado dos ruídos coloridos como dados de entrada.

Ao se analisar a evolução da matriz de pesos utilizando-se as probabilidades associadas a cada microestado como dados de entrada, percebe-se que a distribuição de pesos ocorre de forma concentrada em determinadas regiões, exibindo certos padrões visuais, com linhas e colunas mais destacadas.

Acerca da acurácia de treinamento, os gráficos presentes na Figura 19 mostram a convergência dos dados de treinamento ao longo das 1000 épocas, para o mapa Logístico, sistema de Lorenz e para os Ruídos Coloridos.

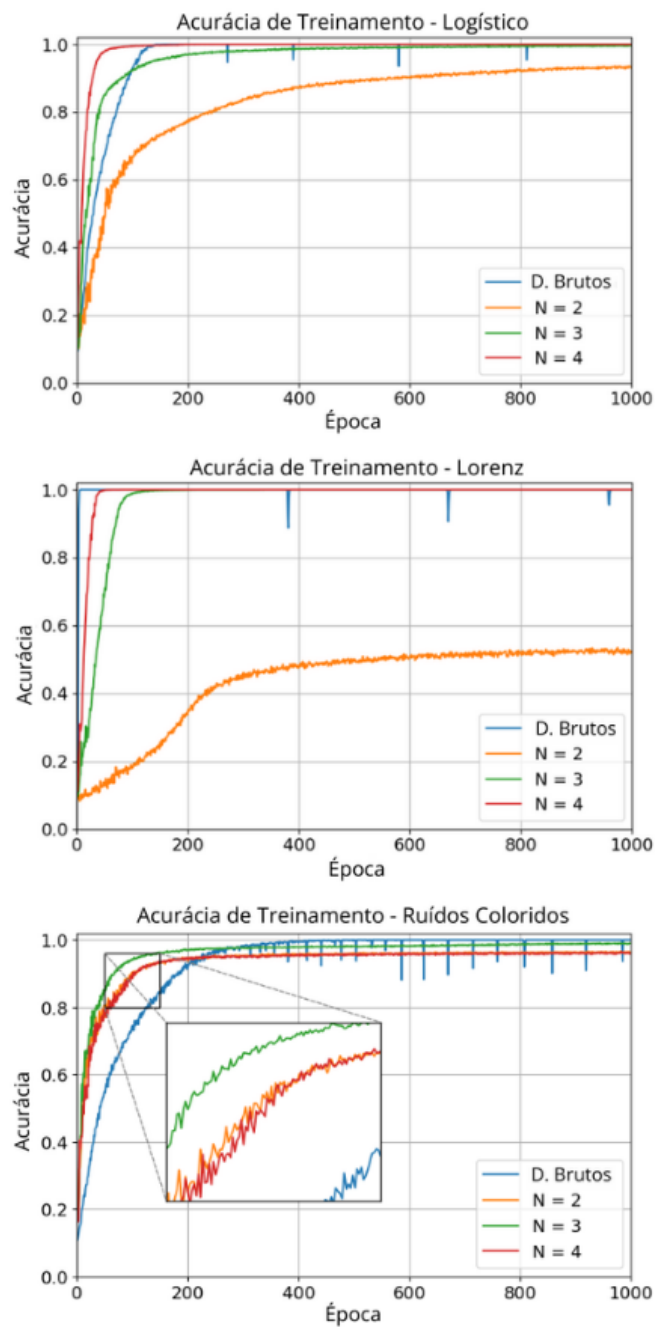


FIGURA 19 – Acurácia de treinamento do mapa Logístico, do Sistema de Lorenz e dos Ruídos Coloridos, utilizando os dados brutos (D.Brutos), bem como as probabilidades associadas a diferentes tamanhos de microestados ( $N = 2$ ,  $N = 3$  e  $N = 4$ ) como dados de entrada.

Percebe-se que, de modo geral, tanto para os dados brutos quanto para as

probabilidades associadas a cada microestado como dados de entrada, a convergência é próxima dos 100% e ocorre de forma relativamente rápida (poucas épocas). Entretanto, nota-se que as curvas referentes aos dados brutos (azul) sofrem quedas bruscas, o que indica uma instabilidade no processo de aprendizado.