

Redes Neurais para Interpretar o Sinal de Neutrinos de Supernovas do Colapso de Estrelas Massivas

Isadora da Silva Espíndola e André da Silva Schneider

Universidade Federal de Santa Catarina

Resumo

Estrelas massivas terminam sua “vida” com um colapso do seu núcleo, dando origem a uma protoestrela de nêutrons e em muitos casos uma explosão conhecida como supernova. Se uma supernova do colapso do núcleo de uma estrela massiva ocorrer na Via-Láctea ou em uma galáxia próxima, milhares a milhões de neutrinos serão detectados na Terra em um intervalo de alguns segundos. Neste trabalho treinamos uma rede neural tipo LSTM (*Long short-term memory* ou memória de curto longo prazo) para relacionar a evolução temporal da massa gravitacional de uma protoestrela de nêutrons aos neutrinos que esta emite nos primeiros dois segundos após o seu surgimento. Os dados de emissão de neutrinos foram obtidos de simulações numéricas de colapso do núcleo de estrelas massivas com simetria esférica (1D) para 18 estrelas progenitoras, duas equações de estado e dois fatores de aquecimento que emulam efeitos multidimensionais em simulações 1D. Testamos a nossa rede LSTM em dados de 49 simulações que não foram utilizadas no treinamento e demonstramos que, em 18 estrelas, a rede neural é capaz de reproduzir a evolução da massa gravitacional da protoestrela de nêutrons com desvios pequenos ($R^2 \gtrsim 0,80$). Discutimos aprimoramentos que ainda podem ser feitos para a rede neural a fim de que esta possa ser utilizada em futuras detecções. Até o momento, este é o primeiro trabalho desse tipo na literatura e serve como parametrização para futuros trabalhos.

Abstract

Massive stars end their life cycles when their inert core collapses and a protoneutron star (PNS) is born, often followed up by a supernova explosion. If such supernova were to occur in the Milky Way or a nearby galaxy, thousands to millions of neutrinos would be detected on laboratories on Earth. In this work we use a *Long short-term memory* (LSTM) neural network to infer the gravitational mass of a PNS from the neutrinos it emits in the first two seconds after its formation. Neutrino data was obtained from spherically symmetric (1D) simulations of core-collapse supernovae of 18 progenitor stars employing two equations of state as well as two heating factors to emulate multidimensional effects in 1D simulations. We tested our LSTM network in data from core-collapse simulations of 49 stars not used in the training and showed that, in 18 cases, the neural network is able to reproduce the gravitational mass evolution of the PNS with only small deviations ($R^2 \gtrsim 0.80$). We discuss how the neural network can be improved and the steps still needed so that it can be used to analyze future neutrino detections. To the best of our knowledge, this is the first work of its kind and will be useful as a parameterization of future works.

Palavras-chave: supernovas, neutrinos, redes neurais.

Keywords: supernovae, neutrinos, neural networks.

DOI: [10.47456/Cad.Astro.v7nEspecial.53342](https://doi.org/10.47456/Cad.Astro.v7nEspecial.53342)

1 Introdução

As supernovas de colapso de núcleo (CCSNe, *Core Collapse Supernovae*) marcam o fim da evolução de estrelas com massas $\gtrsim 10 M_{\odot}$ ¹. Ao longo da sua existência, estas estrelas evoluem através de ciclos sucessivos de fusão nuclear que sinteti-

zam elementos sucessivamente mais pesados, resultando na formação de um núcleo inerte de ferro e níquel.

O aumento da densidade e temperatura do núcleo da estrela causa a dissociação dos núcleos de ferro e níquel e a formação de nêutrons através do decaimento beta. Desta forma, a pressão interna não é suficiente para contrabalançar

¹O símbolo $M_{\odot} = 1,989 \times 10^{30}$ kg denota a massa do Sol.

a pressão gravitacional e o núcleo da estrela entra em colapso. Esse colapso é interrompido quando o núcleo da estrela atinge densidades extremas, $\rho \gtrsim 3 \times 10^{14} \text{ g cm}^{-3}$, e a força nuclear forte impede que este continue contraindo, nascendo assim uma protoestrela de nêutrons (PNS, do inglês *protoneutron star*) [1].

Devido a perda de sustentação, as camadas mais externas da estrela que ainda estão caindo em direção ao centro atingem a PNS e são ricocheteadas, gerando uma onda de choque que pode ejetar o envelope estelar em uma explosão supernova. A fase após a formação da onda de choque é denominada pós-rebote² e o momento em que a matéria em queda livre é impulsãoada pela PNS é conhecido por momento do rebote³. Uma explosão ocorre se os neutrinos produzidos na superfície da protoestrela de nêutrons depositarem energia suficiente na onda de choque para reavivá-la [2–4]. Caso contrário, as camadas mais externas da estrela voltam a cair sobre a PNS e forçam esta a se tornar um buraco negro estelar [1, 5].

Explosões de supernovas do colapso de estrelas massivas têm sido estudadas teoricamente desde os anos 1930 [6, 7] e através de simulações computacionais desde os anos 1960 [8]. Décadas de estudos computacionais sugerem que, dentre os mecanismos propostos para as explosões de CCSNe, o mais plausível seja via “reaquecimento por neutrinos”. Esse mecanismo foi proposto por Bethe e Wilson em 1985 [2] e continua sendo bastante estudado [3, 9, 10]. Como os neutrinos desempenham um papel fundamental em todas as fases do colapso e da explosão estelar, eles são uma peça chave para o estudo da matéria em regimes de densidade e temperatura extremas. Ademais, além das ondas gravitacionais, neutrinos são a única forma de obter informações diretas da PNS, mesmo quando esta colapsa em um buraco negro [11].

A associação entre a emissão de neutrinos em uma CCSNe [7, 8] durante o nascimento de uma estrela de nêutrons [6] foi corroborada pela detecção de 24 neutrinos por três experimentos diferentes no evento SN1987A [12–14]. Devido aos avanços nas últimas décadas, caso ocorra uma CCSN na Via Láctea nos próximos anos, centenas de mi-

lhares de neutrinos serão detectados por diversos detectores de neutrinos em funcionamento [15].

Considerando o exposto acima, o objetivo deste trabalho é utilizar redes neurais em dados gerados por simulações de CCSNe para obter a evolução temporal da massa gravitacional M_{PNS} de uma PNS a partir do sinal de neutrinos desta. No caso de um sinal de neutrinos que termina abruptamente, o que ocorre no caso de um “decaimento” da PNS em um buraco negro, a massa gravitacional neste momento limita a pressão máxima para matéria nuclear estável, uma grandeza bastante sensível à equação de estado da matéria densa e quente na PNS [16].

Estabelecer estas propriedades da PNS nos primeiros momentos após seu nascimento seria uma importante ferramenta para restringir a equação de estado da matéria nuclear em temperatura finita. Isso é possível, visto que a compacticidade⁴ da PNS, $\mathcal{C} = GM_{\text{PNS}}/R_{\text{PNS}}c^2$, é governada pela relação entre a pressão e a densidade da matéria na PNS e que a taxa e a energia total emitida em neutrinos dependem de M_{PNS} , R_{PNS} , e $\dot{M}$ [9].

Neste trabalho, estamos especialmente interessados em entender como a possível detecção de neutrinos de uma supernova pode ser utilizada para interpretar os mecanismos que ocorrem em uma protoestrela de nêutrons. Nesse sentido, construímos uma rede neural artificial baseada em arquitetura LSTM (*Long Short-Term Memory*) [17], projetada para capturar e manter informações ao longo do tempo, permitindo que ela aprenda dependências de longo prazo de séries temporais de propriedades físicas e, com isso, seja capaz de prever propriedades da PNS. Neste primeiro momento, focamos na massa gravitacional da PNS, mas em trabalhos futuros pretendemos inferir, simultaneamente à massa gravitacional M_{PNS} , o raio R_{PNS} da PNS e a taxa de acreção de massa $\dot{M}$ desta.

O artigo está organizado conforme descrito a seguir. As simulações numéricas de CCSNe são discutidas na Seção 2, enquanto a rede neural e seu treinamento são discutidos na Seção 3. Resultados para propriedades da PNS obtidas a partir da detecção de neutrinos de eventos simulados são discutidos na Seção 4. Conclusões e perspectivas estão na Seção 5.

²Traduzido livremente do inglês *post bounce*.

³Traduzido livremente do inglês *bounce time*.

⁴Razão massa-raio tornada adimensional pelas constantes gravitacional G e velocidade da luz c .

2 Simulações

Os dados utilizados para treinar a rede neural foram obtidos a partir de simulações feitas com o código FLASH [18–21], utilizando simetria esférica. A simulação acompanha uma estrela do estágio pré-colapso, momento em que o núcleo fica instável ao colapso gravitacional, até a formação de uma PNS e seus primeiros segundos de evolução.

A partir das simulações, obtemos o espectro de neutrinos emitidos durante o colapso da estrela massiva e evolução da PNS. Tanto o sinal de neutrinos quanto a estrutura e evolução da PNS dependem da equação de estado da matéria densa que, nas faixas de densidade e temperatura encontradas em uma PNS, ainda não é bem determinada. Logo, utilizamos duas equações de estado disponíveis na literatura que estão próximas aos limites superior e inferior de “rigidez” da equação de estado⁵ esperados de uma combinação de modelos teóricos, observações astrofísicas, e experimentos em laboratório [22].

As equações de estado utilizadas foram a DD2 [23], mais “dura”, e SFHo [24], mais “macia”⁶. Equações de estado mais duras, em geral, resultam em estrelas com raios maiores para a mesma massa além de massas gravitacionais máximas mais altas para as estrelas de nêutrons. Além disso, o sinal de neutrinos é sensível a interação destes com a matéria densa. As seções de choque das diferentes interações foram computadas para cada equação de estado utilizando o código NULIB [25] separando os neutrinos em 12 grupos com energias espaçadas logaritmicamente entre 1 e 300 MeV, seguindo as Referências [16, 25]. Embora o espectro de neutrinos seja contínuo, computacionalmente a discretização em poucos grupos se faz necessária. A necessidade de separar os neutrinos desta maneira se dá pelo fato de que calcular a seção de choque tem um custo computacional relativamente alto. Ao discretizar o espectro em 12 grupos, as seções de choque são pré-calculadas para as diferentes faixas de energia, limitando o número de equações a serem resolvidas. O número de grupos foi escolhido para

garantir a precisão numérica ao simular o colapso do núcleo ao mesmo tempo que minimiza os recursos computacionais utilizados por cada simulação.

A estrutura de uma estrela massiva pré-colapso do núcleo depende de sua massa inicial, composição (metalicidade), velocidade de rotação, se está em um sistema binário, dentre outros fatores. Neste trabalho, focamos em 18 estrelas progenitoras de CCSNe isoladas e sem rotação das Referências [26–28], frequentemente utilizadas em simulações de CCSNe. As estrelas escolhidas tinham massa inicial entre 11 e 40 $M_{\odot}$ e compacticidade no momento em que o colapso inicia⁷ $10^{-3} \leq \xi_{2,5 M_{\odot}} \leq 1$, testando os casos mais comuns de estrelas massivas.

Enfatizamos aqui que a evolução das estrelas progenitoras de supernova desde o momento em que as reações nucleares iniciam em seu centro até o momento em que o colapso gravitacional inicia foram realizadas pelo código de evolução estelar KEPLER [29] e descritas em detalhes nas Referências [26–28]. Essas simulações de evolução estelar supõem uma massa e composição inicial para a estrela, além de um conjunto de reações nucleares possíveis. A evolução da estrela é, então, simulada das condições iniciais até o início do colapso gravitacional, quando o núcleo da estrela se contrai tão rápido que um tratamento relativístico do sistema (ainda que aproximado) e inclusão de interações da matéria com neutrinos são indispensáveis. Assim, a configuração final de uma estrela massiva obtida por um código de evolução estelar, disponibilizada pelos autores Referências [26–28], é utilizada como condição inicial para códigos de simulação de colapso de núcleos de estrelas massivas, como fizemos neste trabalho usando o código FLASH. Esse é um procedimento padrão da literatura. Simulações de colapso de núcleo de estrelas massivas são, em geral, identificadas pela massa inicial da estrela no momento que as primeiras reações nucleares começam, além de outras propriedades, como metalicidade, quando pertinente. Entretanto, como algumas estrelas massivas costumam perder boa parte de sua massa por ventos estelares durante suas vidas, no momento do colapso do núcleo suas

⁵A rigidez de uma equação de estado denota o quão rápido a pressão varia com o aumento da densidade e/ou temperatura.

⁶As tabelas usadas nas simulações estão disponíveis em stellarcollapse.org [1].

⁷A compacticidade $\xi_M = (M/M_{\odot})/R(1000 \text{ km})$ determina quão compacta a região de raio R contendo uma massa M da estrela é no estágio pré-colapso [1].

massas são, em geral, significativamente menores que sua massa inicial.

O último ajuste feito nas simulações realizadas neste trabalho foi o de multiplicar os coeficientes de interação de neutrinos com a matéria por um *fator de aquecimento* $f_{\text{heat}} = 2,0$ e $3,0$. Isso se deve ao fato de que simulações de simetria esférica raramente explodem em supernovas devido a ausência de turbulência [1]. Esse fator de aquecimento f_{heat} aumenta a energia depositada pelos neutrinos atrás do choque para que, em alguns casos, ocorram explosões mesmo em simulações com simetria esférica. Assim, o colapso de cada uma das 18 estrelas progenitoras foi simulada quatro vezes considerando a combinação de duas equações de estado (DD2 e SFHo) e de dois fatores de aquecimento ($f_{\text{heat}} = 2,0$ e $3,0$), totalizando 72 simulações. Pretendemos ampliar o número de simulações em trabalhos futuros para incluir outras equações de estado, estrelas progenitoras, e fatores de aquecimento ou outros métodos de se gerar explosões em simulações de simetria esférica [30]. Utilizar simulações com simetria axial ou em 3D sem simetrias impostas tem um custo computacional ainda muito alto para serem viáveis. O volume das simulações incluiu os 10^8 m mais internos de cada estrela divididos em 160 blocos de comprimentos radiais iguais. Cada bloco é subdividido em até 13 níveis de resolução adaptativa, similar ao da Referência [16], resultando em uma resolução máxima de aproximadamente 76,3 m.

Nossas simulações determinaram a evolução das PNSs por 2,50 s ou até as PNSs formarem um buraco negro, no caso disso ocorrer em um tempo mais curto do que 2,50 s. Dentro deste limite de tempo estão inclusos, nos casos em que não houve formação de um buraco negro, pelo menos 2,0 s de tempo pós-rebote, visto que do início do colapso até a formação da PNS se passam no máximo 0,5 s. Embora as PNSs continuem evoluindo após alguns segundos, limitamos nosso estudo a este intervalo pois, geralmente, alguns segundos após o colapso (1) a variação nas grandezas observáveis ocorre mais lentamente, (2) a luminosidade dos neutrinos será mais baixa, o que pode dificultar a detecção destes neutrinos na Terra, e (3) as simulações são menos confiáveis no caso das PNSs estarem com seu interior mais frio⁸.

⁸Conforme a temperatura e a entropia da PNS diminuem, degenerescências nas equações de estado tornam

Para o uso dos dados na rede neural, Seção 3, dividimos o tempo pós-rebote de cada simulação em 2048 intervalos iguais. Embora cada intervalo varie entre uma simulação e outra, o intervalo será no máximo de duração $\Delta t \sim 1$ ms. Os dados de saída monitorados de cada simulação para uso na rede foram o tempo pós-rebote t e as energias médias $\langle \epsilon_\nu \rangle$ e luminosidades L_ν de cada tipo de neutrino, $\nu = \nu_e, \bar{\nu}_e, \nu_x$. Desta forma, os dados de entrada utilizados na rede tem dimensão (72, 2048, 7), combinando o número de simulações (72), o comprimento da janela temporal (2048) e o número de características (7, discutidas abaixo) utilizadas para treinar a rede.

3 Redes Neurais

Redes Neurais Artificiais são técnicas computacionais que apresentam um modelo matemático inspirado na estrutura neural biológica inteligente e que adquirem conhecimento através das experiências as quais foram expostas. Já desde as décadas de 1940 e 1950 há registros do estudo de redes neurais artificiais e um marco em 1958: a criação do *Perceptron*, o neurônio artificial, por Rosenblatt [31].

Nesse sentido, na década de 1980 temos o surgimento das redes neurais recorrentes (*RNNs* – *Recurrent Neural Network*) modernas, capazes de executar a mesma tarefa para todos os elementos de uma sequência, com a saída sendo dependente dos cálculos realizados anteriormente. Sendo essa muito utilizada em dados de processamento de linguagem, reconhecimento de voz e análise de sequências. No entanto, na década seguinte, descobriu-se que a arquitetura da rede tinha limitações à sequências mais longas causando os chamados esquecimento e explosão de gradientes (*vanishing gradient and exploding gradients*). Mais tarde, Hochreiter e Schmidhuber (1997), propuseram uma RNN capaz de resolver esta deficiência das RNNs tradicionais, a LSTM.

Uma rede LSTM (*Long Short-Term Memory*) é

as simulações bastante sensíveis ao passo de tempo utilizado e ao número de grupos de neutrinos utilizados. Logo, os passos de tempo devem ser decrescidos para evitar problemas numéricos, fazendo com que simulações longas tenham um custo computacional significativamente mais alto que um simples fator proporcional ao tempo de evolução da PNS.

um tipo especial de rede neural recorrente projetada para capturar e manter informações ao longo do tempo, permitindo que ela “aprenda” dependências de longo prazo em séries temporais [17]. Ela tem uma estrutura interna composta de células de memória e portões que controlam o fluxo de informações, decidindo o que reter, atualizar ou descartar.

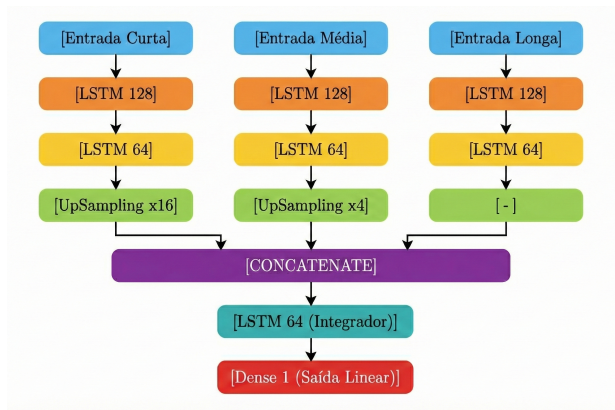


Figura 1: Esquema da arquitetura do modelo de LSTM.

Construímos uma LSTM como a da Figura 1, utilizando a interface de programação Keras [32] implementada na interface TensorFlow [33] em Python utilizando três janelas temporais (ou *sliding windows*) com comprimentos distintos: uma janela “longa” contendo uma sequência de 2048 intervalos de tempo, outra janela média com 1536 sequências de 512 intervalos, e uma janela curta com 1920 sequências de 128 intervalos. Cada uma dessas sequências é utilizada para treinar a rede LSTM em caminhos inicialmente independentes. Os dados de cada janela passam por 128 neurônios na entrada e, em seguida, por 64 neurônios. Para possibilitar o alinhamento dos dados das três janelas da rede após a análise pelas camadas de neurônios de entrada fazemos uma superamostragem (*upsampling*) das janelas curta e média, repetindo os dados destas janelas de maneira ordenada 16 e 4 vezes, respectivamente. Desta forma, após a sobreamostragem, os dados de saída das análises de janelas de comprimentos distintos tem o mesmo comprimento e as sequências temporais ficam alinhadas novamente. Feito isso, o resultado da análise das três LSTMs é concatenado, analisado por mais uma camada de 64 neurônios e, por fim, utiliza-se uma camada densa para saída com um neurônio que prevê uma sequência temporal da propriedade de interesse; no

caso deste trabalho, a massa gravitacional das protoestelas de nêutrons. Nosso objetivo foi criar um modelo que capturasse correlações temporais em três escalas distintas (curto, médio e longo prazo) com a mesma resolução e combinasse essas informações para realizar uma previsão.

Para o treinamento da rede, como já mencionado, fornecemos 7 *características* de entrada de cada uma das 72 simulações: tempo pós-rebote t da simulação, energias média $\langle \epsilon_\nu \rangle$ e luminosidades L_ν de cada tipo de neutrino, $\nu = \nu_e, \bar{\nu}_e, \nu_x$, utilizados para prever a massa gravitacional M_{grav} da PNS. Antes do treinamento, os dados de cada característica e da propriedade a ser prevista foram normalizados linearmente para terem valores entre 0 e 1 ao entrar na rede. Do total de amostras, dividimos aproximadamente 70% para o treinamento (51 simulações) e 30% para a validação (21 simulações), padrão frequentemente utilizado na literatura.

Enfatizamos a importância de distinguir os conceitos de conjuntos de validação e teste (Seção 4). O conjunto de validação é utilizado durante o treinamento para monitorar o desempenho da rede e orientar a escolha de hiperparâmetros, tais como taxa de aprendizado, número de neurônios, tamanho das janelas temporais e taxa de desativação, contribuindo para a diminuição de sobreajustes. Esses hiperparâmetros foram construídos e ajustados empiricamente a partir da literatura e das métricas de validação. Já o conjunto de teste é reservado exclusivamente para a avaliação final, utilizando dados independentes do processo de treinamento + validação para estimar a capacidade de generalização do modelo.

A arquitetura de camadas LSTM utiliza 128 unidades ocultas iniciais com uma taxa de dropout de 0,2 para reduzir o risco de *overfitting*. Uma taxa de dropout moderada (0,2) melhora a generalização, evitando que haja coadaptação dos neurônios durante o treinamento. Além disso, o modelo foi compilado utilizando o otimizador Adam [34] e a função de perda Erro Quadrático Médio (MSE, *Mean Squared Error*)

$$\mathcal{L}_{\text{MSE}} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (1)$$

onde y_i é o rótulo verdadeiro da amostra i e $\hat{y}_i$ é a previsão feita pelo modelo. O $\mathcal{L}$ representa a

função perda entre y_i e $\hat{y}_i$, sobre n elementos do conjunto total de amostras de tamanho n , *i.e.*, um conjunto de n linhas do conjunto de dados. No contexto de nosso trabalho, y_i representa os valores de massa verdadeiros (simulação) e $\hat{y}_i$ os valores previstos pela LSTM para o mesmo conjunto de dados.

Além disso, um valor de parada antecipada foi implementado para interromper o treinamento quando a perda de validação não tivesse uma melhora por 50 épocas consecutivas, ou seja, o modelo demonstrar que parou de aprender, reduzindo assim o risco de sobreajuste. O modelo foi treinado com um tamanho de lote de 32 e 1000 épocas.

Concluída a etapa de treinamento, a rede foi testada com um outro conjunto de dados simulados para testar a previsão para massa gravitacional das PNSs a partir do sinal de neutrinos desse novo conjunto. As previsões da rede foram então comparadas com o valor obtido pelas simulações.

O desempenho da rede com o conjunto de teste foi avaliado utilizando as seguintes métricas:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (2)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (3)$$

$$\text{MAPE} = \frac{100}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i}. \quad (4)$$

4 Resultados

A avaliação foi conduzida em um conjunto de teste composto por 49 simulações de colapso de núcleo de supernovas, abrangendo um espectro de massas de progenitoras entre 12 e 60 $M_\odot$ em intervalos de 1 $M_\odot$ utilizando estrelas progenitoras da Referência [35] simuladas com o programa KEPLER [29], nenhuma das quais a rede teve acesso durante o treinamento e/ou previamente. Enquanto as simulações de treinamento e validação são inteiramente novas, as simulações do teste são da Referência [16] e (1) utilizaram uma equação de estado diferente das utilizadas no treinamento e validação, *SRO “baseline”* [4], e (2) não incluíram nenhum fator de aquecimento. Logo, com exceção das estrelas menos massivas $M \leq 17 M_\odot$,

todas as simulações do teste resultaram na formação de um buraco negro. Além do mais, embora algumas estrelas deste conjunto de simulações teste tenham a mesma massa inicial que estrelas do conjunto utilizado no treinamento, as características da estrela no momento pré-colapso, que servem como condição inicial para as simulações do colapso, eram diferentes devido a escolhas feitas no código de evolução estelar.

4.1 Desempenho

A Tabela 1 apresenta o resumo estatístico global das métricas de desempenho obtidas no conjunto de teste. O modelo alcançou um erro percentual absoluto médio (MAPE) de 5,71 % para a predição da massa gravitacional, valor considerado satisfatório diante da complexidade associada ao processo de acreção de massa no regime pós-rebote e ao número limitado de simulações utilizados para o treino e validação do modelo. Esse comportamento é consistente com o RMSE médio de 0,126 $M_\odot$, indicando que, em termos absolutos, os desvios das predições em relação às simulações permanecem dentro de uma faixa que consideramos aceitável para a evolução da proto-estrela de nêutrons.

Diferentemente de abordagens voltadas a sistemas binários, como as das Referências [36,37] que utilizam simulações de fusão de sistemas binários para estabelecer a massa limite de uma estrela de nêutrons fria a partir do colapso imediato em um buraco negro, não há um método bem estabelecido para inferir a massa de uma PNS isolada como proposto aqui. Apesar de o trabalho de Nagakura & Vartanyan (2022) também utilizar sinais de neutrinos do colapso de uma estrela massiva para prever a evolução da massa gravitacional e raio de uma PNS [38], existem diferenças fundamentais entre este trabalho e o nosso. Nagakura & Vartanyan (2022) utilizam 18 simulações de progenitoras “leves” ($9 \leq M/M_\odot \leq 25$) em 2D para obter ajustar os dados da evolução da massa e do raio da PNS com funções polinomiais cujos coeficientes dependem da energia total emitida em neutrinos, consideram efeitos de detectores e modelos de oscilação de neutrinos que ignoramos neste trabalho. Por outro lado, Nagakura & Vartanyan (2022) utilizam uma única equação de estado em seus estudos (SFHo, também utilizada

aqui) e suas previsões mostram bastante variação na maioria dos casos analisados em relação aos das simulações utilizadas na análise. Logo, como nosso trabalho é único na metodologia implementada e deverá servir como parâmetro para trabalhos futuros.

Embora o coeficiente de determinação⁹ médio ($R^2 = 0,596$) apresente limitações quando interpretado isoladamente em sistemas não lineares, sua análise em conjunto com as demais métricas revela desempenho consistente: 94 % (46 estrelas) das simulações exibiram correlação positiva entre valores reais e preditos. Em particular, aproximadamente 49 % da amostra atingiu níveis elevados de fidelidade ($R^2 > 0,75$), nos quais o RMSE assume valores mínimos da ordem de $0,0181 M_\odot$.

Por outro lado, os 16 % dos casos (8 estrelas) com $R^2 \leq 0,2$ correspondem a situações em que o erro da rede se mostrou significativamente superior ao observado no restante do conjunto, com o RMSE alcançando valores de até $0,28 M_\odot$. Esses desvios concentram-se predominantemente nos extremos de massa das progenitoras, como nas simulações de massas iniciais mais baixa, como $12 M_\odot$, e nas mais massivas, próximas a massas iniciais de $60 M_\odot$. Isso se deve ao fato de que nenhuma das estrelas utilizadas no treinamento da rede tem compacticidade tão alta (baixa) quanto as estrelas mais (menos) massivas dos casos de teste.

Tabela 1: Resumo estatístico global das métricas de predição para as 49 simulações de teste. As unidades de medidas das métricas abaixo são dadas por RMSE e MAE, em $M_\odot$, MAPE adimensional (%) e, por fim, R^2 adimensional e com limite superior igual à 1.

Métrica	Média	Desvio	Mín.	Máx.
RMSE	0,1264	0,0632	0,0181	0,2757
MAE	0,1215	0,0618	0,0142	0,2642
MAPE	5,7145	2,5740	0,7480	11,2973
R^2	0,5959	0,3590	-0,5119	0,9950

Na Tabela 2 listamos, dentre as 49 simulações de colapso testadas, os quatro melhores e os qua-

tro piores desempenhos para a rede neural ao prever a evolução da massa gravitacional da estrela. Notamos que embora a rede neural tenha conseguido prever muito bem a evolução da massa gravitacional em alguns casos, $R^2 \sim 1$, em outros casos o desempenho não foi satisfatório, $R^2 \lesssim 0$.

Tabela 2: Desempenho individual das simulações: quatro melhores e quatro piores casos segundo o coeficiente de determinação (R^2) da massa gravitacional. A nome mX de cada simulação se refere a massa inicial X da estrela.

Simulação	RMSE	MAPE	R^2
gray!15 <i>Melhores desempenhos</i>			
m41	0,0181	0,75	0,9950
m43	0,0196	0,78	0,9941
m39	0,0369	1,49	0,9826
m40	0,0383	1,54	0,9809
gray!15 <i>Piores desempenhos</i>			
m15	0,1039	6,51	-0,5119
m12	0,0977	6,13	-0,1737
m18	0,1229	6,86	-0,1569
m60	0,2757	11,30	0,0672

Nas Figuras 2 e 3, são apresentadas as matrizes de confusão binadas (*binned confusion matrix*) por intervalos de massas. Para obter estas matrizes, primeiro dividimos o tempo pós-rebote de cada simulação em 2048 intervalos iguais. Em seguida, para cada simulação determinamos a massa gravitacional mínima e máxima da PNS considerando tanto a simulação quanto a previsão da rede. O alcance entre estes valores mínimo e máximo é dividido em 10 faixas (*bins*) de massas. Finalmente, para cada um dos 2048 intervalos de tempo, a matriz de confusão compara a massa obtida na simulação com aquela prevista pela rede neural e adiciona um ponto na matriz de confusão. Assim, quanto maiores os valores próximos à diagonal principal, melhor é o desempenho da rede, já que indicam que a massa prevista pela rede está no mesmo intervalo que a massa da simulação. Esse método de aferição de desempenho é frequentemente utilizado em modelos de classificação em *aprendizado de máquina* e aqui foi adaptado para analisar dados de uma série temporal.

Podemos observar que para a simulação m43, Figura 2 para a estrela com massa inicial $43 M_\odot$,

⁹O coeficiente de determinação R^2 é definido como $R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}} = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$ com SS_{res} e SS_{tot} somas dos quadrados dos resíduos e a soma dos quadrados total, no qual $\bar{y}$ é a massa gravitacional média da PNS ao longo de toda a sequência temporal analisada.

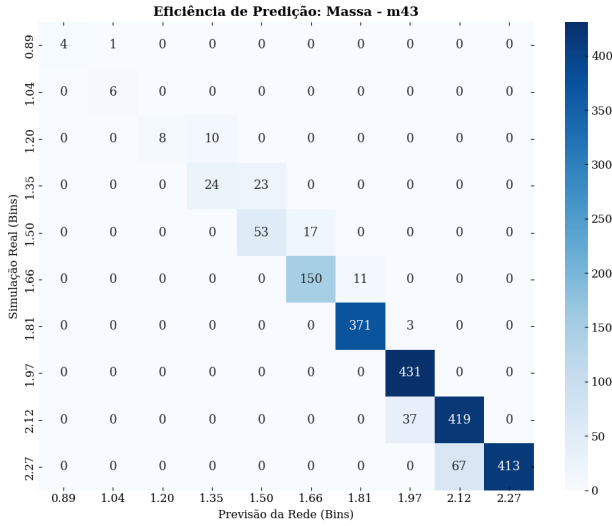


Figura 2: Matriz de confusão dos valores de massa gravitacional predito pelo nosso modelo de LSTM para uma estrela progenitora de $43 M_{\odot}$.

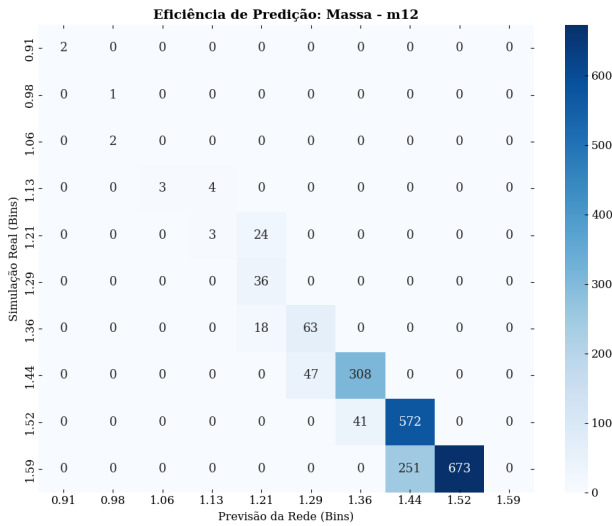


Figura 3: Matriz de confusão dos valores de massa gravitacional predito pelo nosso modelo de LSTM para uma estrela progenitora de $12 M_{\odot}$.

os valores na matriz de confusão (que somados resultam em 2048) estão distribuídos em sua maioria ao longo da diagonal principal evidenciando o bom ajuste do modelo. Isso é confirmado na Figura 4 que compara as séries temporais para a massa gravitacional pós-rebote da PNS obtida na simulação e prevista pela rede neural. Já para a simulação m12, estrela com massa inicial $12 M_{\odot}$, podemos identificar sua tendência de sobreajuste. O sobreajuste ocorre quando o modelo memoriza o ruído e os detalhes dos dados de treino, limitando sua capacidade para genera-

lizar o aprendizado para novas amostras. Neste caso, embora o erro no treinamento e validação decaiam, mostrando que há aprendizado, a rede demonstra um resultado para a estrela de massa $12 M_{\odot}$ não satisfatório em comparação a estrela de massa $43 M_{\odot}$.

4.2 Predições

As predições apresentadas nas Figuras 4 e 5 contém uma área hachurada que é um intervalo de predição de 68 %, representando a incerteza de um desvio padrão (1σ) em torno dos valores estimados pela rede. São calculados limites superior e inferior da área a partir da soma e subtração do produto entre o escore-Z bicaudal e o erro padrão dos resíduos (σ) sobre as estimativas pontuais ($\hat{y}_i$). Aqui o erro padrão dos resíduos quantifica a dispersão das observações reais em relação às previsões do modelo. Esse erro padrão é ajustado pelos graus de liberdade do sistema através da expressão $\sigma = \sqrt{\sum (y_i - \hat{y}_i)^2 / (n - 2)}$, onde n é o número total de pontos utilizados na comparação entre os dados da simulação e valores previstos pela rede. Deste modo temos $L_s = \hat{y}_i + Z \cdot \sigma$ e $L_i = \hat{y}_i - Z \cdot \sigma$, os limites superior e inferior, respectivamente.

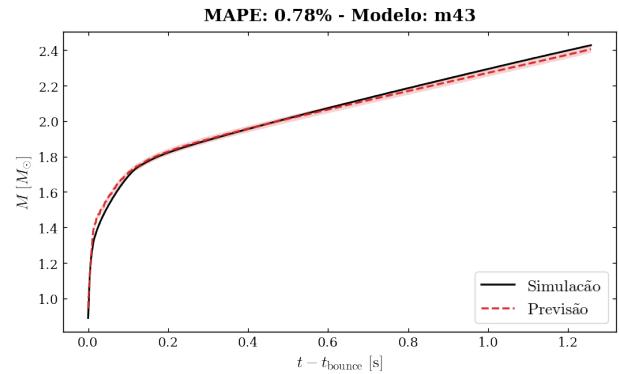


Figura 4: Comparativo entre os valores de predição da rede (curva tracejada na cor vermelha) e o verdadeiro (curva sólida na cor preta) (simulação) em função do tempo de simulação pós-rebote para uma estrela progenitora de $43 M_{\odot}$.

A Figura 4 apresenta a evolução temporal para a estrela progenitora de $43 M_{\odot}$ pós-rebote. Observa-se que a rede LSTM captura com precisão a inclinação do acréscimo de massa gravitacional da PNS, mesmo na região inicial de pós-rebote, no qual há mais flutuações. Já a Figura 5 apresenta a evolução para a progenitora

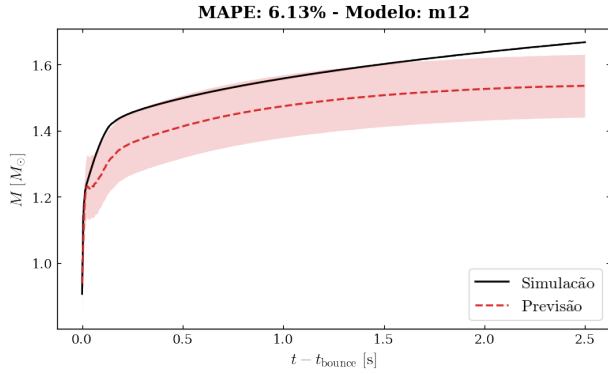


Figura 5: Comparativo entre os valores de predição da rede (curva tracejada na cor vermelha) e o verdadeiro (curva sólida na cor preta) (simulação) em função do tempo de simulação pós-rebote para uma estrela progenitora de $12 M_{\odot}$.

de $12 M_{\odot}$ pós-rebote, que não é adequadamente descrita pela rede. Vemos que, imediatamente o pós-rebote é bem descrito, no entanto, com a alta taxa de acreção de matéria a rede não foi capaz de descrever a curva de crescimento, levando à uma saturação de massa inferior ao esperado.

Na Figura 6 fazemos uma comparação entre os valores previstos pela rede neural e valores de saída das simulações para as 49 estrelas do conjunto de teste. Assim, observamos que existe uma tendência da rede em subestimar os valores previstos para a massa gravitacional das PNS quanto estas atingem $M \gtrsim 1,75 M_{\odot}$.

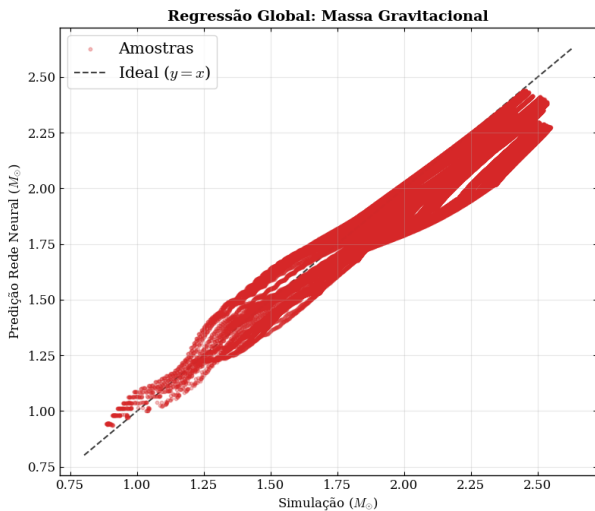


Figura 6: Comparativo entre os valores de predição da rede e valores de saída das simulações para as 49 estrelas do conjunto de teste. Observamos uma tendência de subestimar os valores previstos para as massas $M_{\text{grav}} \gtrsim 1,75 M_{\odot}$.

4.3 Análise de resíduos

O resíduo corresponde à diferença entre os valores observados e os valores estimados pelo modelo. Neste trabalho, adota-se o erro relativo percentual¹⁰, apresentado nas Figuras 7 e 8.

De maneira geral, a distribuição dos resíduos, Figura 7, apresenta pico próximo de 5%, indicando um viés global significativo, como observado na Figura 6 para $M_{\text{grav}} \gtrsim 1,75 M_{\odot}$. Este pico em um valor > 0 (à direita) com leve assimetria positiva sugere que para alguns regimes o modelo subestima o valor de massa já que temos um valor predito menor que o valor da simulação. As caudas longas são consistentes com os valores RMSE serem maior que o MAE e também o R^2 negativo em algumas simulações. Assim, podemos inferir que a rede faz boas predições, mas ainda precisa de alguns ajustes além de dados de treinamento que explore melhor o espaço de fases de compactidades possíveis para estrelas massivas.

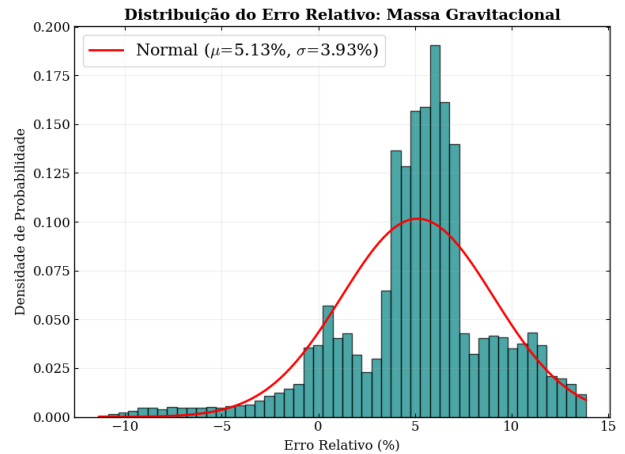


Figura 7: Histograma de distribuição de erros na predição da massa gravitacional das PNSs em comparação com uma curva gaussiana. Distribuição de frequências dos erros relativos percentuais para a massa gravitacional. A curva sólida em vermelho representa o ajuste Gaussiano teórico com média $\mu = 5,13\%$ e desvio padrão $\sigma = 3,93\%$. A concentração da densidade de probabilidade em torno de zero confirma a precisão da rede LSTM.

Os resíduos exibem comportamento aproximadamente normal na região central, com desvios significativos nas caudas evidenciados pelo gráfico quantis-quantis (Q-Q plot), Figura 8. O Q-Q plot é construído a partir da comparação entre os

$$^{10}\text{Erro \%} = \frac{y_i - \hat{y}_i}{y_i} \times 100$$

quantis empíricos dos resíduos (erros) ordenados e os quantis teóricos de uma distribuição de referência (distribuição gaussiana). A proximidade dos pontos à reta $y = \mu + \sigma x$ indica concordância com a distribuição teórica, i.e., o comportamento ideal seria que os valores de erros, pontos em azul, estivessem distribuídos ao longo da reta, em vermelho. No caso deste trabalho, a proximidade à reta no regime central indica um comportamento aproximadamente gaussiano dos erros, enquanto os desvios nas extremidades revelam erros percentuais extremos associados a regimes físicos específicos: tentativa de reproduzir com a rede LSTM condições de estrelas de compacticidade mais alta e mais baixa das que utilizadas no treinamento. Esses resultados corroboram com os indicadores já citados, mostrando que a rede possui boa capacidade preditiva, mas com limitações de generalização e uma tendência de prever massa levemente menores que as das simulações.

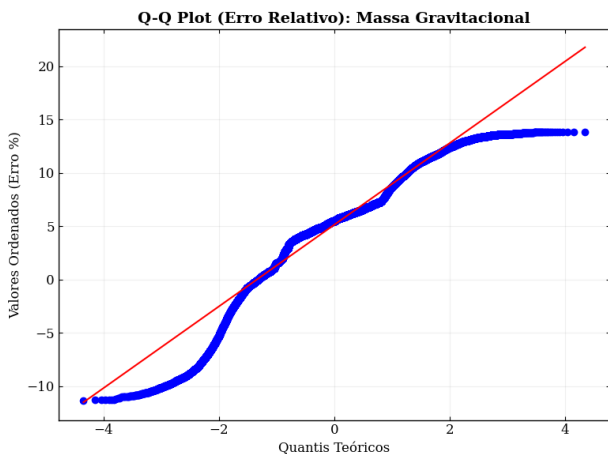


Figura 8: Gráfico Quantil-Quantil (Q-Q Plot) dos erros relativos da massa gravitacional. O eixo vertical representa os resíduos ordenados: valores positivos indicam subestimação pelo modelo ($\text{Real} > \text{Predito}$), enquanto valores negativos indicam superestimação ($\text{Real} < \text{Predito}$). A aderência dos pontos à linha de referência teórica, especialmente no intervalo central, valida a natureza Gaussiana dos erros e a estabilidade estatística da rede LSTM.

5 Conclusões

Neste estudo, desenvolvemos uma rede neural de memória de curto e longo prazo, LSTM, para prever a massa gravitacional de protoestrelas de nêutrons de supernovas do colapso do núcleo de estrelas massivas utilizando o espectro de neu-

trinos oriundo destas. Dada a inexistência de dados reais para o treinamento da rede neural, utilizamos dados de 72 simulações numéricas obtidos com o código de hidrodinâmica relativística FLASH para treinar a rede. Em seguida, testamos as previsões da rede em prever a massa gravitacional de uma protoestrela de nêutrons em 49 simulações de outro trabalho que não foram utilizadas no treinamento [16]. Concluímos que o modelo é satisfatório na maior parte do domínio de nosso conjunto de teste, embora tenha uma tendência de subestimar a massa gravitacional quando esta está no regime $M \gtrsim 1,75 M_{\odot}$. Em aproximadamente metade da amostra, obtivemos níveis elevados de fidelidade ($R^2 \gtrsim 0,75$), nos quais o RMSE assume valores mínimos da ordem de $0,018 M_{\odot}$. O RMSE é essencial porque ele penaliza mais severamente os grandes desvios e mantém a unidade física ($M_{\odot}$).

Com a análise das distribuições de erro, nota-se uma concentração próxima de zero e boa concordância com uma distribuição normal no regime central. No entanto a rede encontra dificuldade em prever os limites superior e inferior de massas das progenitoras. Esses resultados indicam elevada precisão média do modelo, com limitações localizadas. Além disso, observamos a tendência da rede em prever valores de massa gravitacional da PNS menores do que os estimados pela simulação. Isso pode ser devido à falta de simulações que cubram o espaço de parâmetros de estrelas com estruturas pré-colapso semelhantes às de nossos conjuntos de dados de treinamento e validação ou, ainda, a necessidade de uma parametrização específica para estes casos.

Novas investigações seguem sendo realizadas examinando o impacto de diferentes equações de estado e variação do fator de aquecimento nas simulações de supernovas. Nosso objetivo final é usar redes neurais para prever uma variedade de propriedades de PNSs, como raio e taxa de acreção de massa, que poderiam ser determinadas a partir de detecções de neutrinos no caso de uma supernova galáctica. Determinar as propriedades do PNS no primeiro momento após seu nascimento seria uma ferramenta inestimável para restringir a equação de estado da matéria nuclear em temperatura finita, já que o raio da PNS e sua massa gravitacional são sensíveis às propriedades da matéria densa e quente no interior da PNS.

Além de utilizar mais simulações para treinar o modelo aumentando o número de equações de estado e estrelas progenitoras utilizadas, pretendemos também construir uma outra rede neural LSTM que, ao invés de utilizar o espectro de neutrinos gerados pelo código FLASH como características de entrada para a rede neural, utiliza os dados da simulação processados pela ferramenta SNEWPY [15, 39]. Essa ferramenta utiliza os dados de emissão de neutrinos gerados por uma simulação de colapso de estrela massiva para estimar quantos neutrinos em diferentes faixas de energia seriam observados por detectores distintos na Terra considerando a distância para a supernova além de modelos para as oscilações de neutrinos. Esses dados de simulação pós-processados por SNEWPY serão mais realistas da situação que cientistas encontrarão no caso de uma supernova galáctica do que os dados gerados pelo código FLASH, ou outro simulador de colapsos de estrelas massivas, como feito neste trabalho. Assim, a rede neural pode ser diretamente utilizada no caso de uma detecção de neutrinos de uma CCSNe para determinar propriedades de uma PNS nos primeiros instantes após o seu nascimento.

A evolução inicial da PNS e o sinal de neutrinos emitidos dependem da estrutura estelar pré-colapso e da equação de estado da matéria densa e quente. Este trabalho é um primeiro passo para que, havendo uma supernova de colapso de núcleo na Via Láctea, possamos determinar a evolução da massa bariônica e do raio da PNS simultaneamente afim de impor limites a características da equação de estado da matéria nas condições extremas encontradas em uma PNS. No futuro, também pretendemos utilizar outras redes neurais como GRU [40], *TRANSFORMERS* [41] e U-NET [42] para determinar as vantagens e limitações de cada uma delas para o problema proposto.

Agradecimentos

Agradecemos ao Professor Aldo von Wangelin e Juliana Arrais por discussões sobre estratégias de escolha de arquiteturas de rede neurais e tratamento dos dados. O presente trabalho foi realizado com apoio da Coordenação de Aper-

feiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Sobre os autores

I. S. Espíndola (isadora.espindola@posgrad.ufsc.br) é bacharel em física com ênfase em astrofísica pela Universidade Federal do Rio Grande do Sul e Mestranda do Programa de Pós-Graduação em Física da Universidade Federal de Santa Catarina.

A. S. Schneider (schneider.andre@ufsc.br) é Professor do Departamento de Física na Universidade Federal de Santa Catarina com Ph.D. pela Indiana University - Bloomington e especialista em Astrofísica Nuclear e Física Computacional.

Referências

- [1] E. O'Connor e C. D. Ott, A new open-source code for spherically symmetric stellar collapse to neutron stars and black holes, *Classical and Quantum Gravity* **27**(11), 114103 (2010).
- [2] H. A. Bethe e J. R. Wilson, Revival of a stalled supernova shock by neutrino heating, *The Astrophysical Journal* **295**, 14 (1985).
- [3] H.-T. Janka, Explosion mechanisms of core-collapse supernovae, *Annual Review of Nuclear and Particle Science* **62**, 407 (2012).
- [4] A. S. Schneider et al., Equation of state effects in the core collapse of a 20-MJ star, *Physical Review C* **100**, 055802 (2019).
- [5] R. Fernandez et al., Mass ejection in failed supernovae: variation with stellar progenitor, *Monthly Notices of the Royal Astronomical Society* **476**(2), 2366 (2018).
- [6] W. Baade e F. Zwicky, Cosmic rays from super-novae, *Proceedings of the National Academy of Sciences* **20**(5), 259 (1934).
- [7] G. Gamow e M. Schoenberg, Neutrino Theory of Stellar Collapse, *Phys. Rev.* **59**, 539 (1941).

- [8] S. A. Colgate e M. H. Johnson, Hydrodynamic Origin of Cosmic Rays, *Phys. Rev. Lett.* **5**, 235 (1960).
- [9] C. D. Ott et al., The Progenitor Dependence of Core-collapse Supernovae from Three-dimensional Simulations with Progenitor Models of $12 - 40 M_{\odot}$, *The Astrophysical Journal Letters* **855**(1), L3 (2018).
- [10] A. Burrows e D. Vartanyan, Core-collapse supernova explosion theory, *Nature* **589**(7840), 29 (2021).
- [11] H.-T. Janka, Neutrino Emission from Supernovae, in *Handbook of Supernovae* (Springer International Publishing, Cham, 2016), 1–30.
- [12] K. Hirata et al., Observation of a neutrino burst from the supernova SN1987A, *Phys. Rev. Lett.* **58**, 1490 (1987).
- [13] R. M. Bionta et al., Observation of a neutrino burst in coincidence with supernova 1987A in the Large Magellanic Cloud, *Phys. Rev. Lett.* **58**, 1494 (1987).
- [14] E. Alexeyev et al., Detection of the neutrino signal from SN 1987A in the LMC using the INR Baksan underground scintillation telescope, *Physics Letters B* **205**(2), 209 (1988).
- [15] K. Scholberg, Supernova Neutrino Detection, *Annual Review of Nuclear and Particle Science* **62**(Volume 62, 2012), 81 (2012).
- [16] A. Schneider et al., Equation of state and progenitor dependence of stellar-mass black hole formation, *The Astrophysical Journal* **894**(1), 4 (2020).
- [17] S. Hochreiter e J. Schmidhuber, Long short-term memory, *Neural Computation* **9**(8), 1735 (1997).
- [18] B. Fryxell et al., FLASH: An Adaptive Mesh Hydrodynamics Code for Modeling Astrophysical Thermonuclear Flashes, *The Astrophysical Journal Supplement Series* **131**(1), 273 (2000).
- [19] A. Dubey et al., Extensible Component-Based Architecture for FLASH, a Massively Parallel, Multiphysics Simulation Code, *Parallel Computing* **35**(10-11), 512 (2009).
- [20] S. M. Couch, THE DEPENDENCE OF THE NEUTRINO MECHANISM OF CORE-COLLAPSE SUPERNOVAE ON THE EQUATION OF STATE, *The Astrophysical Journal* **765**(1), 29 (2013).
- [21] E. P. O'Connor e S. M. Couch, Exploring Fundamentally Three-dimensional Phenomena in High-fidelity Simulations of Core-collapse Supernovae, *The Astrophysical Journal* **865**(2), 81 (2018).
- [22] J. Margueron, R. Hoffmann Casali e F. Gulminelli, Equation of state for dense nucleonic matter from metamodeling. I. Foundational aspects, *Phys. Rev. C* **97**, 025805 (2018).
- [23] M. Hempel et al., NEW EQUATIONS OF STATE IN SIMULATIONS OF CORE-COLLAPSE SUPERNOVAE, *The Astrophysical Journal* **748**(1), 70 (2012).
- [24] A. W. Steiner, M. Hempel e T. Fischer, CORE-COLLAPSE SUPERNOVA EQUATIONS OF STATE BASED ON NEUTRON STAR OBSERVATIONS, *The Astrophysical Journal* **774**(1), 17 (2013).
- [25] E. O'Connor, AN OPEN-SOURCE NEUTRINO RADIATION HYDRODYNAMICS CODE FOR CORE-COLLAPSE SUPERNOVAE, *The Astrophysical Journal Supplement Series* **219**(2), 24 (2015).
- [26] S. E. Woosley, A. Heger e T. A. Weaver, The evolution and explosion of massive stars, *Rev. Mod. Phys.* **74**, 1015 (2002).
- [27] S. Woosley e A. Heger, Nucleosynthesis and remnants in massive stars of solar metallicity, *Physics Reports* **442**(1), 269 (2007). The Hans Bethe Centennial Volume 1906-2006.
- [28] T. Sukhbold et al., CORE-COLLAPSE SUPERNOVAE FROM 9 TO 120

- SOLAR MASSES BASED ON NEUTRINO-POWERED EXPLOSIONS, *The Astrophysical Journal* **821**(1), 38 (2016).
- [29] T. A. Weaver, G. B. Zimmerman e S. E. Woosley, Presupernova evolution of massive stars., *The Astrophysical Journal* **225**, 1021 (1978).
- [30] S. M. Couch, M. L. Warren e E. P. O'Connor, Simulating Turbulence-aided Neutrino-driven Core-collapse Supernova Explosions in One Dimension, *The Astrophysical Journal* **890**(2), 127 (2020).
- [31] F. Rosenblatt, The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain, *Psychological Review* **65**(6), 386 (1958).
- [32] F. Chollet et al., Keras, <https://keras.io> (2015).
- [33] M. Abadi et al., TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems (2015). Software available from tensorflow.org. Disponível em <https://www.tensorflow.org/>.
- [34] D. P. Kingma e J. Ba, Adam: A Method for Stochastic Optimization (2017). Disponível em <https://arxiv.org/abs/1412.6980>, ArXiv:1412.6980.
- [35] T. Sukhbold, S. E. Woosley e A. Heger, A High-resolution Study of Presupernova Core Structure, *The Astrophysical Journal* **860**(2), 93 (2018).
- [36] K. Hotokezaka et al., Binary neutron star mergers: Dependence on the nuclear equation of state, *Phys. Rev. D* **83**, 124008 (2011).
- [37] R. Kashyap et al., Numerical relativity simulations of prompt collapse mergers: Threshold mass and phenomenological constraints on neutron star properties after GW170817, *Phys. Rev. D* **105**, 103022 (2022).
- [38] H. Nagakura e D. Vartanyan, Efficient method for estimating the time evolution of the proto-neutron star mass and radius from a supernova neutrino signal, *Monthly Notices of the Royal Astronomical Society* **512**(2), 2806 (2022). ArXiv:<https://academic.oup.com/mnras/article-pdf/512/2/2806/43158747/stac383.pdf>.
- [39] A. L. Baxter et al., Snepy: A data pipeline from supernova simulations to neutrino signals, *The Astrophysical Journal* **925**(2), 107 (2022).
- [40] K. Cho et al., Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation, CoRR **abs/1406.1078** (2014). Disponível em <http://arxiv.org/abs/1406.1078>, ArXiv:1406.1078.
- [41] A. Vaswani et al., Attention Is All You Need, CoRR **abs/1706.03762** (2017). Disponível em <http://arxiv.org/abs/1706.03762>, ArXiv:1706.03762.
- [42] O. Ronneberger, P. Fischer e T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)* (Springer, 2015), 234–241.