



Previsão do VaR e ES utilizando metodologias híbridas: LSTM e modelo de volatilidade

Palavras-chave: Volatilidade, GARCH, LSTM

Aluno: Vinicius Rezende Bardelin - IMECC, UNICAMP

Orientador: Carlos Cesar Trucios Maza - IMECC, UNICAMP

Introdução

A mensuração adequada do risco é um elemento central na regulação e estabilidade do sistema financeiro. Diante da natureza volátil dos mercados, agentes financeiros e reguladores recorrem a medidas quantitativas para estimar potenciais perdas e garantir níveis adequados de segurança patrimonial. Entre as medidas de risco mais utilizadas nas últimas décadas, destacam-se o Valor em Risco (VaR) e a Perda Esperada (ES) [Hotta et al., 2024]. Embora amplamente adotado, o VaR apresenta limitações conhecidas, como a não coerência [Artzner et al., 1999], o que motivou a adoção da ES como medida complementar, recomendada pelo Comitê de Basileia por sua capacidade de capturar perdas além do VaR. A eficácia dessas medidas depende diretamente da precisão dos modelos utilizados em sua estimação, reforçando a importância do desenvolvimento e avaliação de métodos e modelos que capturem adequadamente a dinâmica dos dados.

Modelos de séries temporais vêm sendo largamente aplicados na previsão de volatilidade, VaR e ES, com destaque para a família GARCH [Bollerslev, 1986] e suas extensões, bem como recentemente, os modelos da família GAS [Harvey, 2013a]. Por outro lado, com os avanços recentes em aprendizado de máquina, abordagens baseadas em redes neurais, especialmente as redes LSTM, têm ganhado destaque por sua capacidade de capturar padrões não lineares e dependências de longo prazo em séries financeiras. Mais recentemente, modelos híbridos que combinam estruturas GARCH com redes LSTM têm mostrado desempenho superior às suas versões isoladas na previsão de volatilidade e do VaR Kim and Won [2018], Roszyk and Ślepaczuk [2024], Kakade et al. [2022].

Neste contexto, este trabalho compara os modelos híbridos GARCH-LSTM, MSGARCH-LSTM e GAS-LSTM com suas versões individuais (GARCH, MSGARCH e GAS), avaliando seu desempenho na previsão de volatilidade, VaR e, de forma menos explorada na literatura, da ES. A principal contribuição do estudo está na incorporação dos modelos MSGARCH e GAS em combinações com redes LSTM, abordagem ainda não explorada até o momento. A aplicação considera a série de retornos da ação PETR4, utilizando como proxy de volatilidade o estimador de Parkinson, baseado nos preços máximo e mínimo diários.

Metodologia

Os modelos de volatilidade utilizados neste estudo foram o GARCH Bollerslev [1986], MSGARCH Haas et al. [2004] e GAS Creal et al. [2013], Harvey [2013b]. Em todos eles, adotou-se a distribuição t-Student padronizada, a fim de acomodar a presença de caudas pesadas, uma característica comum em séries de retornos financeiros.

O modelo GARCH é o benchmark na modelagem de volatilidade, sendo uma das abordagens mais utilizadas para modelar a heterocedasticidade condicional em retornos financeiros.

O modelo MSGARCH é indicado para contextos em que há quebras estruturais na dinâmica da volatilidade, pois permite a alternância entre diferentes regimes, refletindo distintas fases de comportamento do mercado ao longo do tempo.

A classe de modelos GAS permite modelar parâmetros que variam ao longo do tempo com base no score da função de verossimilhança. Essa flexibilidade é útil quando a estrutura do modelo não é evidente, conforme observado em Troster et al. [2019].

A abordagem baseada em aprendizado profundo adotada neste estudo foi a rede LSTM Hochreiter and Schmidhuber [1997]. Trata-se de uma arquitetura de rede neural recorrente projetada para preservar informações relevantes por longos intervalos de tempo, sendo especialmente eficaz na identificação de padrões complexos em séries temporais financeiras, como não linearidades e dependências de longo prazo Géron [2019], Olah [2015].

Como entrada para a LSTM, foram utilizadas as previsões dos modelos GARCH, MSGARCH e GAS, os retornos da PETR4, estimativas da volatilidade de Parkinson (Parkinson [1980]), todos com valores defasados ($t - 1$ e $t - 5$). A motivação para essa integração é combinar os pontos fortes dos modelos estatísticos na

captura dos fatos estilizados da volatilidade com a flexibilidade das LSTM em modelar estruturas dinâmicas mais complexas. Na Figura 1, apresentamos a estrutura dos modelos híbridos construídos.

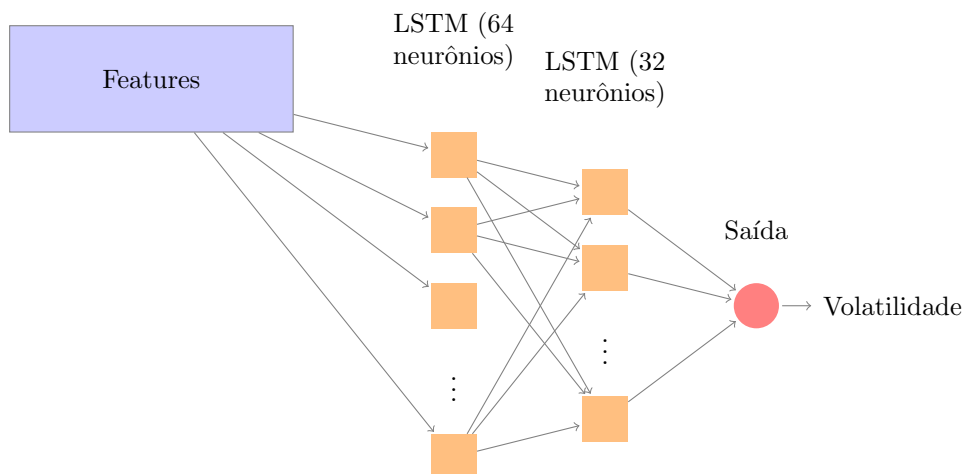


Figura 1: Estrutura dos modelos híbridos.

Para comparar os modelos individuais e híbridos, avaliamos o erro nas previsões de volatilidade por meio das funções de perda robustas MSE e QLIKE Patton [2011]. Já para o VaR e ES, são utilizados testes de calibração (UC, CC, DQ, VQR) e funções score (QL, FZG, NZ, AL). Veja por exemplo Hotta et al. [2024] e Trucíos and Taylor [2023].

Por fim, as previsões um passo à frente do VaR e da ES foram geradas utilizando uma janela deslizante de 1500 dias, combinada com o procedimento de Simulação Histórica Filtrada (FHS), conforme proposto por Barone-Adesi et al. [1999].

Resultados e Discussão

O desempenho preditivo dos modelos foi avaliado fora da amostra (2000 dias) com base nas métricas MSE e QLIKE, utilizando como referência a volatilidade de Parkinson. A Tabela 1 apresenta os erros de previsão obtidos para cada modelo.

Tabela 1: Erros de previsão dos modelos avaliados. O MSE foi multiplicado por 10^6 para facilitar a interpretação.

Modelo	MSE	QLIKE
GARCH	1,84	0,51
GARCH-LSTM	0,95	0,42
MSGARCH	1,71	0,47
MSGARCH-LSTM	0,96	0,43
GAS	1,71	0,47
GAS-LSTM	1,02	0,44
LSTM	0,91	0,48

Os modelos híbridos superaram suas versões individuais, apresentando reduções consistentes nos valores de MSE e QLIKE. O GARCH-LSTM, por exemplo, reduziu o MSE de 1,84 para 0,95 (redução de 48%) e o QLIKE de 0,51 para 0,42 (redução de 17,6%) em relação ao GARCH. Resultados semelhantes foram observados para os demais pares: o MSGARCH-LSTM apresentou MSE de 0,96 (redução de 43,8%) e QLIKE de 0,43 (redução de 8,5%) em relação ao MSGARCH, enquanto o GAS-LSTM obteve MSE de 1,02 (redução de 40,4%) e QLIKE de 0,44 (redução de 6,4%) frente ao GAS. O modelo LSTM puro apresentou o menor MSE entre todos (0,91), mas com QLIKE superior ao dos híbridos.

A Figura 2 ilustra as previsões de volatilidade geradas pelos modelos GARCH e GARCH-LSTM ao longo do período fora da amostra. Observa-se que o modelo híbrido GARCH-LSTM acompanha mais de perto a dinâmica da volatilidade de Parkinson, com menor amplitude de erro em diversos momentos. Esse comportamento está em concordância com os resultados apresentados, que indicaram reduções significativas nas métricas de MSE e QLIKE em relação ao GARCH puro.

Os resultados do exercício de backtesting, apresentados na Tabela 2, indicam que os modelos híbridos apresentaram desempenho limitado nos níveis de risco 1% e 2,5%, com taxas de violação (% Hits) superiores ao esperado e rejeições nos testes de calibração. No nível de 5%, contudo, o modelo GAS-LSTM se destacou,

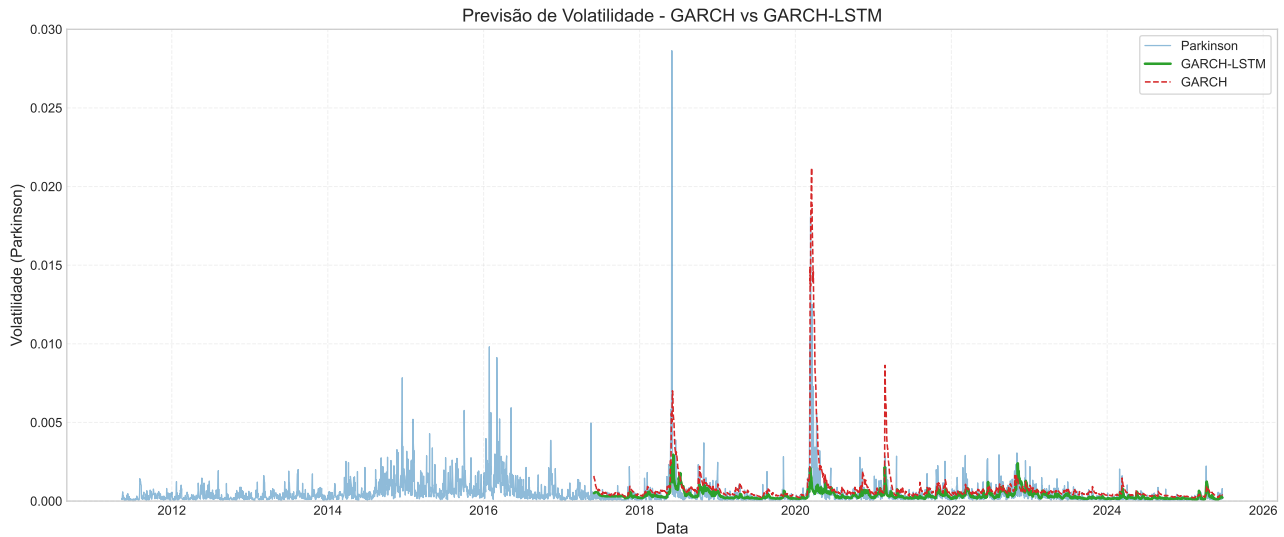


Figura 2: Previsão de volatilidade com os modelos GARCH e GARCH-LSTM no período fora da amostra.

alcançando cobertura próxima ao ideal (4,96%) e não rejeitando a hipótese nula nos testes de calibração¹. Por outro lado, os modelos individuais, como GARCH e GAS, tenderam a subestimar o risco nesse nível de significância.

Tabela 2: Resultados do backtesting para os modelos avaliados.

Modelo	% Hits	UC	CC	DQ	VQR	CoC	ER	ESR
1%								
GARCH	1,26%	0,268	0,064	0,022	0,74	0,398	0,567	0,807
MSGARCH	1,45%	0,055	0,032	0,000	0,293	0,166	0,574	0,05
GAS	1,21%	0,359	0,380	0,543	0,000	0,175	0,598	0,384
LSTM	2,3%	0,000	0,000	0,000	0,000	0,001	0,511	0,282
GARCH-LSTM	2,1%	0,000	0,000	0,000	0,033	0,003	0,432	0,213
MSGARCH-LSTM	1,6%	0,013	0,012	0,000	0,055	0,1	0,807	0,325
GAS-LSTM	1,95%	0,000	0,000	0,000	0,003	0,009	0,981	0,177
2,5%								
GARCH	2,46%	0,917	0,114	0,146	0,983	0,649	0,525	0,437
MSGARCH	2,91%	0,251	0,051	0,010	0,883	0,341	0,407	0,355
GAS	2,58%	0,830	0,418	0,204	0,043	0,139	0,492	0,150
LSTM	3,85%	0,000	0,000	0,001	0,055	0,007	0,794	0,264
GARCH-LSTM	3,6%	0,003	0,002	0,000	0,058	0,03	0,772	0,164
MSGARCH-LSTM	3,4%	0,014	0,014	0,001	0,125	0,083	0,86	0,140
GAS-LSTM	3,81%	0,000	0,001	0,003	0,021	0,1	0,949	0,68
5%								
GARCH	4,12%	0,064	0,035	0,036	0,000	0,111	0,604	0,047
MSGARCH	4,17%	0,224	0,011	0,008	0,528	0,389	0,666	0,047
GAS	4,04%	0,139	0,262	0,159	0,000	0,037	0,412	0,003
LSTM	5,65%	0,191	0,248	0,056	0,531	0,389	0,737	0,192
GARCH-LSTM	4,9%	0,837	0,188	0,442	0,757	0,978	0,94	0,045
MSGARCH-LSTM	4,85%	0,757	0,166	0,127	0,933	0,930	0,954	0,033
GAS-LSTM	4,96%	0,930	0,397	0,156	0,852	0,992	0,864	0,597

A Figura 3 apresenta as curvas de VaR estimadas pelos modelos GAS, LSTM puro e GAS-LSTM, nos níveis de significância de 1%, 2,5% e 5%.

¹Para mais detalhes sobre os resultados das funções escore, consulte o relatório final.

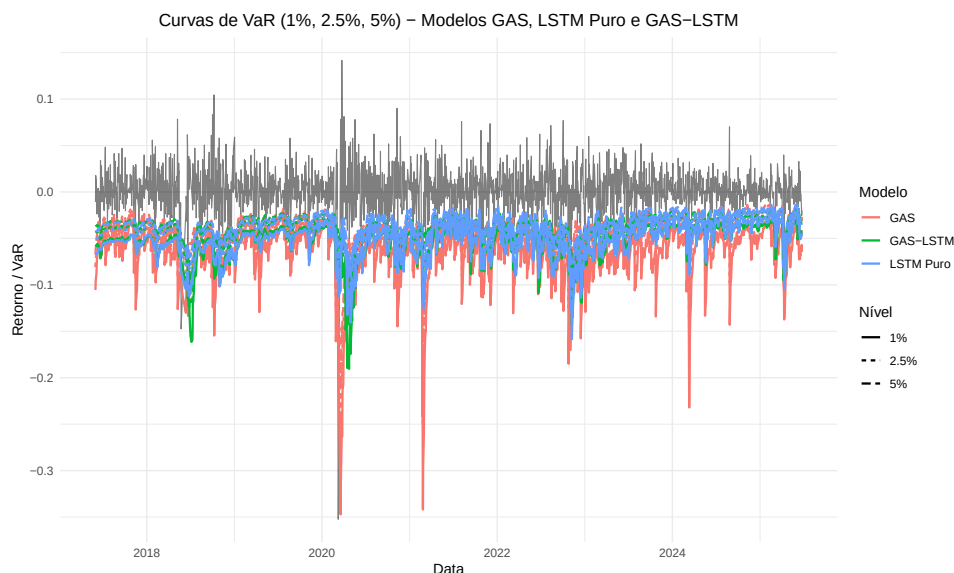


Figura 3: Curvas de VaR para os modelos GAS, LSTM Puro e GAS-LSTM nos níveis de 1%, 2,5% e 5%.

Conclusões

Os resultados encontrados indicam que a combinação de modelos estatísticos com redes neurais LSTM pode aprimorar a previsão de volatilidade e contribuir para um gerenciamento de risco mais eficaz. Os modelos híbridos apresentaram reduções expressivas nas funções de perda MSE e QLIKE em comparação às suas versões individuais, demonstrando ganhos consistentes ao incorporar estruturas LSTM. No entanto, o desempenho dos híbridos mostrou-se limitado nos níveis de risco 1% e 2,5%, com taxas de violação acima do ideal e falhas nos testes de calibração. Já no nível de 5%, o modelo GAS-LSTM destacou-se com cobertura adequada. A introdução dos modelos MSGARCH e GAS em arquiteturas híbridas representa uma nova contribuição na literatura e se mostra uma abordagem promissora para a estimativa de medidas de risco como VaR e ES em contextos financeiros práticos.

Bibliografia

- P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3): 203–228, 1999. doi: <https://doi.org/10.1111/1467-9965.00068>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-9965.00068>.
- G. Barone-Adesi, K. Giannopoulos, and L. Vosper. Var without correlations for portfolios of derivative securities. *Journal of Futures Markets*, 19:583–602, 1999. URL <https://api.semanticscholar.org/CorpusID:51802636>.
- T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3): 307–327, 1986. ISSN 0304-4076. doi: [https://doi.org/10.1016/0304-4076\(86\)90063-1](https://doi.org/10.1016/0304-4076(86)90063-1). URL <https://www.sciencedirect.com/science/article/pii/0304407686900631>.
- D. Creal, S. J. Koopman, and A. Lucas. Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28(5):777–795, 2013. doi: <https://doi.org/10.1002/jae.1279>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.1279>.
- A. Géron. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, 2019. ISBN 978-1492032649.
- M. Haas, S. Mittnik, and M. S. Paoletta. A new approach to markov-switching garch models. *Journal of Financial Econometrics*, 2(4):493–530, 09 2004. ISSN 1479-8409. doi: 10.1093/jjfinec/nbh020. URL <https://doi.org/10.1093/jjfinec/nbh020>.
- A. C. Harvey. *Dynamic Models for Volatility and Heavy Tails: With Applications to Financial and Economic Time Series*. Econometric Society Monographs. Cambridge University Press, 2013a.
- A. C. Harvey. *Dynamic Models for Volatility and Heavy Tails: With Applications to Financial and Economic Time Series*. Econometric Society Monographs. Cambridge University Press, 2013b.

- S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Comput.*, 9(8):1735–1780, 1997. doi: 10.1162/neco.1997.9.8.1735.
- L. Hotta, C. Trucíos, and M. Zevallos. A note about calibration tests for var and es. In *Time Series and Wavelet Analysis: Festschrift in Honor of Pedro A. Morettin*, pages 59–80. Springer, 2024.
- K. Kakade, I. Jain, and A. K. Mishra. Value-at-risk forecasting: A hybrid ensemble learning garch-lstm based approach. *Resources Policy*, 78:102903, 2022. ISSN 0301-4207. doi: <https://doi.org/10.1016/j.resourpol.2022.102903>. URL <https://www.sciencedirect.com/science/article/pii/S0301420722003476>.
- H. Y. Kim and C. H. Won. Forecasting the volatility of stock price index: A hybrid model integrating lstm with multiple garch-type models. *Expert Systems with Applications*, 103:25–37, 2018. ISSN 0957-4174. doi: <https://doi.org/10.1016/j.eswa.2018.03.002>. URL <https://www.sciencedirect.com/science/article/pii/S0957417418301416>.
- C. Olah. Understanding lstm networks. 2015. URL <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>.
- M. Parkinson. The extreme value method for estimating the variance of the rate of return. *The Journal of Business*, 53(1):61–65, 1980. ISSN 00219398, 15375374. URL <http://www.jstor.org/stable/2352357>.
- A. J. Patton. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1):246–256, 2011.
- N. Roszyk and R. Ślepaczuk. The hybrid forecast of sp 500 volatility ensembled from vix, garch and lstm models. Working Papers 2024-13, Faculty of Economic Sciences, University of Warsaw, 2024. URL <https://EconPapers.repec.org/RePEc:war:wpaper:2024-13>.
- V. Troster, A. K. Tiwari, M. Shahbaz, and D. N. Macedo. Bitcoin returns and risk: A general garch and gas analysis. *Finance Research Letters*, 30:187–193, 2019. ISSN 1544-6123. doi: <https://doi.org/10.1016/j.frl.2018.09.014>. URL <https://www.sciencedirect.com/science/article/pii/S1544612318304859>.
- C. Trucíos and J. W. Taylor. A comparison of methods for forecasting value at risk and expected shortfall of cryptocurrencies. *Journal of Forecasting*, 42(4):989–1007, 2023. doi: <https://doi.org/10.1002/for.2929>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/for.2929>.