
Previsão do VaR e ES utilizando o modelo combinado de regressão conjunta com rede neural recorrente LSTM

Bolsista: Vinicius Rezende Bardelin

Orientador: Carlos Cesar Trucios Maza

Instituto de Matemática, Estatística e Computação Científica, UNICAMP, Campinas, SP.

Resumo

Este trabalho visa utilizar uma metodologia híbrida na estimação do valor-em-risco (VaR) e expected shortfall (ES). O modelo híbrido será composto pelo modelo de regressão conjunta, bem como por redes neurais recorrentes. A proposta será avaliada utilizando dados reais de criptomoedas através de um cenário de janela rolante. Os códigos implementados serão gratuitamente disponibilizados em um repositório público no GitHub.

1 Introdução

O risco associado a uma posição financeira é de sumo interesse para investidores, instituições financeiras e agências regulatórias, tornando necessárias metodologias para sua estimação e previsão. Nesse sentido, duas medidas de risco amplamente empregadas no mercado financeiro são o Valor em Risco (VaR) e a Perda Esperada (ES). Com efeito, como apontado por [13], o VaR se tornou uma medida padrão para quantificar o risco de um ativo financeiro, devido a sua simplicidade conceitual. No entanto, apesar da sua fácil interpretação, o VaR possui sérios problemas conceituais, entre eles, o fato dele não ser uma medida coerente, veja [2, 3]. Para lidar com as desvantagens conceituais do VaR, [2] propôs o uso do ES, que satisfaz as condições de uma medida de risco coerente. Adicionalmente, o que esperamos de uma medida quantificadora de risco é que ela seja elicitável [9]. Embora o ES seja uma medida consistente, infelizmente não é elicitável, tornando sua estimação e avaliação fora da amostra uma tarefa desafiadora. Porém, [7] encontrou

que o VaR e o ES são conjuntamente elicitáveis, superando as limitações antes descritas. Por fim, com base no cenário de funções conjuntamente consistentes, [5] propuseram um modelo de regressão conjunta para estimar o VaR e o ES, o qual será empregado neste trabalho.

No contexto do mercado de criptomoedas, mais especificamente no mercado do Bitcoin, alguns estudos mostram que o comportamento dos retornos diários do Bitcoin apresentam uma estrutura de correlação serial de longo prazo, bem como dinâmicas não lineares em seus dados, de modo que os modelos estatísticos tradicionais utilizados para previsão de volatilidade podem falhar em capturar esse comportamento complexo [1, 4, 11]. Nesse sentido, algumas metodologias de aprendizado profundo tem sido amplamente empregadas no mercado das criptomoedas. Em particular, o modelo de rede neural LSTM é o modelo mais utilizado para a análise de dados sequenciais, uma vez que é capaz de capturar estruturas complexas e dependências de longo prazo, que é o caso dos retornos das criptomoedas [6].

Assim, este trabalho tem por objetivo estudar a combinação de um modelo de regressão conjunta com um modelo de rede neural LSTM, o denominado modelo LSTM-J-C, proposto por [4], para a previsão do VaR e ES em dados de criptomoedas.

2 Objetivos

- Familiarizar o bolsista com conceitos básicos de finanças, principalmente, medidas de risco.
- Aprender métodos clássicos de previsão de medidas de risco.
- Estudar conceitos de redes neurais.
- Aplicar metodologias híbridas na previsão de medidas de risco.

3 Métodos

3.1 Estimação conjunta do VaR e ES

3.1.1 Definição do Valor em Risco e ES

O valor em risco (VaR) é a perda máxima de uma posição financeira dentro de um horizonte de tempo e de uma probabilidade dada. Suponha que no tempo t estamos interessados em quantificar

o risco associado a uma posição financeira pelos próximos k períodos de tempo. Denotando por $X_t(k)$ a variável aleatória representando a posição financeira pelos próximos k períodos, se $F_t(x)$ é a distribuição acumulada de $X_t(k)$ e p a probabilidade de perda da posição financeira, então, o VaR é definido como $\text{VaR}_{1-p} = \inf\{x | F_t(k) \geq 1 - p\}$. Assim, pela definição de distribuição acumulada, como $F_t(k) \geq 1 - p$, temos $\mathbb{P}[X_t(k) \leq \text{VaR}_{1-p}] \geq 1 - p$. [12].

O ES de uma posição financeira foi proposta por [3] e representa a perda esperada dessa posição dado que a perda é maior do que o VaR [12]. Em termos matemáticos, se X é a variável aleatória representando a posição financeira em um tempo k e $F(x)$ é a distribuição acumulada de X , então, o ES de X é definido como $\text{ES}_{1-p} = \mathbb{E}[X | X > \text{Var}]$ [13].

3.1.2 O modelo de regressão conjunta para estimação do VaR e ES

Como proposto por [5], denotando por $Q_\alpha(Y)$ o quantil condicional de Y ao nível α e a queda esperada de Y por $\text{ES}_\alpha(Y)$, para algum $\alpha \in (0, 1)$ e considerando as covariáveis X_Q e X_E , então, o modelo de regressão conjunta é expresso por

$$Y = X_Q^T \theta_0^Q + \mu^Q; \quad Y = X_E^T \theta_0^E + \mu^E.$$

onde $Q_\alpha(\mu^Q | X) = 0$ e $\text{ES}_\alpha(\mu^E | X) = 0$. Os parametros do modelo são $\theta_0 = (\theta_0^Q, \theta_0^E)'$. Assim, o VaR e o ES condicionais de Y dado as covariáveis X_Q e X_E são

$$Q_\alpha(Y | X) = X_Q^T \theta_0^Q; \quad \text{ES}_\alpha(Y | X) = X_E^T \theta_0^E.$$

Portanto, as estimativas dos parâmetros θ_0 podem ser obtidas minimizando o valor esperado da função escore do VaR e ES.

Dessa forma, se \hat{Q}_i denota a estimativa do VaR advinda modelo de predição individual i e $\hat{\text{ES}}_i$ denota a estimativa do ES advinda modelo de predição individual i , então, o modelo preditivo de regressão combinada é definido como

$$\text{VaR}_{1-\alpha}(Y_t | X) = X_q^T \theta_q = \theta_0^Q + \theta_1^Q \hat{Q}_1 + \theta_2^Q \hat{Q}_2; \quad \text{ES}_{1-\alpha}(Y_t | X) = X_E^T \theta^E = \theta_0^E + \theta_1^E \hat{\text{ES}}_1 + \theta_2^E \hat{\text{ES}}_2.$$

3.2 Modelo LSTM

Os modelos utilizados para a análise de séries temporais normalmente utilizam uma estrutura de regressão linear para modelar a variável resposta. Embora essa estrutura de regressão linear

funcione para muitos casos, ela pode não capturar dinâmicas mais complexas, que exibam por exemplo, correlação serial de longo prazo ou ainda uma estrutura não linear [4]. Felizmente, com o advento de técnicas de aprendizado de máquina, especialmente no campo do aprendizado profundo, o modelo de rede neural recorrente LSTM, proposto por [10], tem sido muito utilizado, devido a sua capacidade de identificar dependências de longo prazo em dados sequenciais. Assim, introduzimos a célula LSTM, denotada por $\text{LSTM}(x_t, h_{t-1})$, como:

$$\mathbf{i}_{(t)} = \sigma(\mathbf{W}_{xi}^T \mathbf{x}_{(t)} + \mathbf{W}_{hi}^T \mathbf{h}_{(t-1)} + \mathbf{b}_i) \quad (1a)$$

$$\mathbf{f}_{(t)} = \sigma(\mathbf{W}_{xf}^T \mathbf{x}_{(t)} + \mathbf{W}_{hf}^T \mathbf{h}_{(t-1)} + \mathbf{b}_f) \quad (1b)$$

$$\mathbf{o}_{(t)} = \sigma(\mathbf{W}_{xo}^T \mathbf{x}_{(t)} + \mathbf{W}_{ho}^T \mathbf{h}_{(t-1)} + \mathbf{b}_o) \quad (1c)$$

$$\mathbf{g}_{(t)} = \tanh(\mathbf{W}_{xg}^T \mathbf{x}_{(t)} + \mathbf{W}_{hg}^T \mathbf{h}_{(t-1)} + \mathbf{b}_g) \quad (1d)$$

$$\mathbf{c}_{(t)} = \mathbf{f}_{(t)} \otimes \mathbf{c}_{(t-1)} + \mathbf{i}_{(t)} \otimes \mathbf{g}_{(t)} \quad (1e)$$

$$\mathbf{y}_{(t)} = \mathbf{h}_{(t)} = \mathbf{o}_{(t)} \otimes \tanh(\mathbf{c}_{(t)}) \quad (1f)$$

onde $\mathbf{i}_{(t)}$ representa o portão de entrada (*input gate*), $\mathbf{f}_{(t)}$ o portão de esquecimento (*forget gate*), $\mathbf{o}_{(t)}$ o portão de saída (*output gate*), $\mathbf{g}_{(t)}$ é a estrutura da célula que analisa as entradas $\mathbf{x}_{(t)}$ e o estado anterior de curto prazo $\mathbf{h}_{(t-1)}$, $\mathbf{c}_{(t)}$ é o estado de longo prazo e $\mathbf{y}_{(t)}$ é a saída da célula LSTM. Adicionalmente, $\sigma(\cdot)$ é a função de ativação logística e $\tanh(\cdot)$ é a função de ativação tangente hiperbólica. Por fim, as matrizes \mathbf{W} são as matrizes de pesos e os vetores \mathbf{b} são os vieses de cada neurônio, onde as matrizes e vieses constituem justamente o conjunto de parâmetros a serem estimados. Um esquema da célula LSTM é apresentado na Figura 1.

3.3 Modelo LSTM-J-C

O modelo combinado de regressão conjunta com rede LSTM, denominado LSTM-J-C, proposto por [4], é construído combinando os modelos de regressão conjunta, proposto por [5], com a rede neural recorrente LSTM, proposta por [10]. Um esquema do modelo LSTM-J-C é apresentado na Figura 2, onde os dados de entrada, R_t , são os log-retornos diários, que por sua vez alimentam

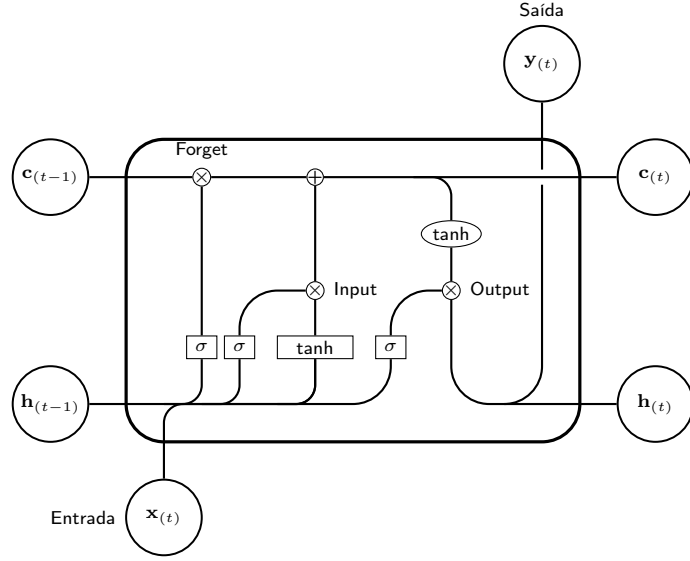


Figura 1: *Célula LSTM. As estruturas em forma de quadrado representam camadas totalmente conectadas com suas respectivas funções de ativação. Os nós descritos por Forget, Input e Output, referem-se aos portões de esquecimento, entrada e saída, respectivamente. Além disso, os sinais de multiplicação e adição dos nós indicam a operação feita pelo portão. Fonte: [8].*

o modelo de regressão individual, produzindo estimativas do VaR e do ES, denotadas por \hat{Q}_i e \widehat{ES}_i , onde essas estimativas são posteriormente filtradas pelo modelo LSTM, produzindo as novas estimativas, \hat{Q}_{il} e \widehat{ES}_{il} , de modo que \hat{Q}_{il} e \widehat{ES}_{il} são finalmente inputadas no modelo preditivo de regressão conjunta proposto por [5]. Em termos matemáticos, o modelo LSTM-J-C é definido como

$$\text{VaR}_{1-\alpha}(Y_t|X) = X_Q^T \theta^Q = \theta_0^Q + \theta_1^Q \hat{Q}_{1l} + \theta_2^Q \hat{Q}_{2l}; \quad \text{ES}_{1-\alpha}(Y_t|X) = X_E^T \theta^E = \theta_0^E + \theta_1^E \widehat{ES}_{1l} + \theta_2^E \widehat{ES}_{2l}.$$

O objetivo desse modelo composto é compreender a dinâmica conjunta do VaR e do ES, tomando vantagens das características inerentes a estrutura do modelo LSTM, permitindo uma estimação mais robusta, no sentido de identificar estruturas complexas em longos de períodos de tempo.

4 Cronograma de atividades

O cronograma de atividades é apresentado na Figura 3.

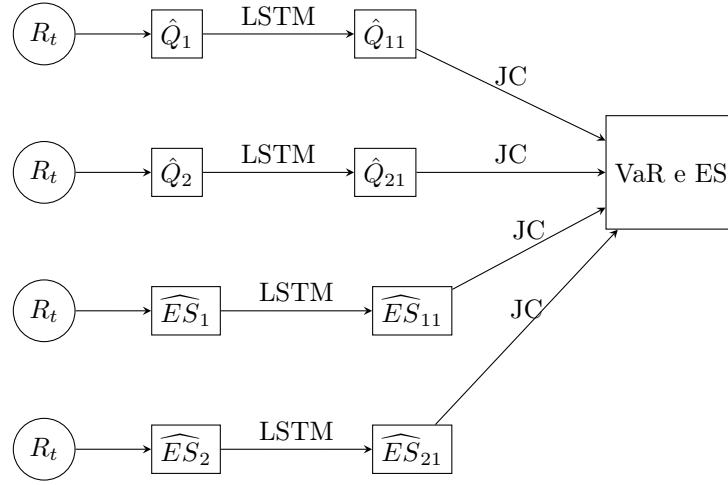


Figura 2: *Modelo LSTM-J-C. As legendas JC representam a aplicação do modelo de regressão conjunta para produzir as estimativas do Modelo LSTM-J-C. Fonte: [4].*

Referências

- [1] A. Bariviera et al. “Some stylized facts of the Bitcoin market”. Em: **Physica A: Statistical Mechanics and its Applications** 484 (2017), pp. 82–90. DOI: <https://doi.org/10.1016/j.physa.2017.04.159>.
- [2] Artzner et al. “Coherent Measures of Risk”. Em: **Mathematical Finance** 9.3 (1999), pp. 203–228. DOI: <https://doi.org/10.1111/1467-9965.00068>.
- [3] Artzner et al. “Thinking Coherently”. Em: **Risk** 10 (1997), pp. 68–71.
- [4] Lu et al. “Risk measurement in Bitcoin market by fusing LSTM with the joint-regression-combined forecasting model”. Em: **Kybernetes** 52.4 (2023). ISSN: 0368-492X.
- [5] T. Dimitriadis e S. Bayer. “A joint quantile and expected shortfall regression framework”. Em: **Electronic Journal of Statistics** 13.1 (2019), pp. 1823–1871. DOI: 10.1214/19-EJS1560. URL: <https://doi.org/10.1214/19-EJS1560>.
- [6] T. Fischer e C. Krauss. “Deep learning with long short-term memory networks for financial market predictions”. Em: **European Journal of Operational Research** 270.2 (2018), pp. 654–669. DOI: <https://doi.org/10.1016/j.ejor.2017.11.054>.
- [7] T. Fissler, J. Ziegel e T. Gneiting. **Expected Shortfall is jointly elicitable with Value at Risk - Implications for backtesting**. Papers. arXiv.org, jul. de 2015.

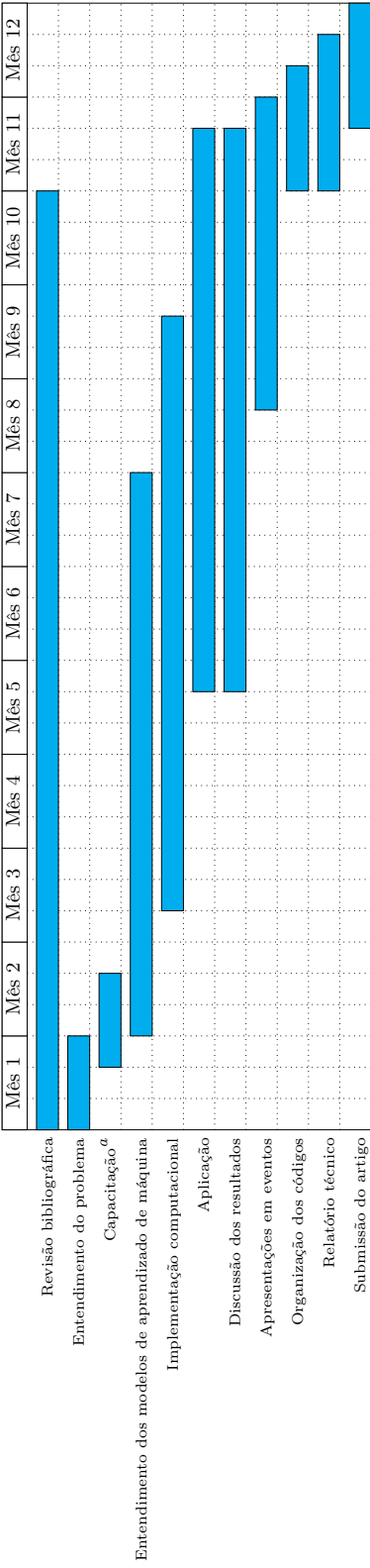


Figura 3: Cronograma mensal.

^aApresentar o uso de tecnologias úteis no desenvolvimento da pesquisa tais como GitHub, Kanban, Google Scholar, Portal de Periódicos CAPES, etc.

- [8] A. Géron. **Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems**. Sebastopol, CA: O'Reilly Media, 2019. ISBN: 978-1492032649.
- [9] T. Gneiting. “Making and Evaluating Point Forecasts”. Em: **Journal of the American Statistical Association** 106.494 (2011), pp. 746–762. DOI: 10.1198/jasa.2011.r10138.
- [10] S. Hochreiter e J. Schmidhuber. “Long Short-Term Memory”. Em: **Neural Comput.** 9.8 (1997), pp. 1735–1780. DOI: 10.1162/neco.1997.9.8.1735.
- [11] Yonghong J., He N. e Weihua R. “Time-varying long-term memory in Bitcoin market”. Em: **Finance Research Letters** 25 (2018), pp. 280–284. DOI: <https://doi.org/10.1016/j.frl.2017.12.009>.
- [12] Ruey S. Tsay. **An Introduction to Analysis of Financial Data with R**. 1st. Wiley Publishing, 2012. ISBN: 0470890819.
- [13] Yasuhiro Yamai e Toshinao Yoshiba. “Value-at-risk versus expected shortfall: A practical perspective”. Em: **Journal of Banking Finance** 29.4 (2005). Risk Measurement, pp. 997–1015. ISSN: 0378-4266. DOI: <https://doi.org/10.1016/j.jbankfin.2004.08.010>.