

# SIMULAÇÃO DE MERCADO FOCADA EM DADOS: SÉRIES FINANCEIRAS

## INFORME SETORIAL

### Encontrando agulhas em pequenos palheiros: simulação de mercado em pequenos ambientes de dados

#### Bloomberg

Na análise de dados contemporânea, *big data* e os muitos “Vs” são temas comuns, incluindo variedade, velocidade e volume, e a ampla gama de ferramentas e técnicas que evoluem para lidar com ambientes ricos em dados. No entanto, se os dados forem escassos, difíceis de obter ou incompletos, que tipo de abordagem será mais eficaz? Blanka Horvath, do King’s College London e The Alan Turing Institute, abordou essa questão no seminário BBQ (Bloomberg Quant) para um público online. Em seu trabalho, a simulação de mercado focada em dados e baseada em redes neurais fornece um método flexível para modelar séries temporais financeiras sem fazer suposições sobre a dinâmica estocástica que ocorre em segundo plano. A chamada simulação de mercado generativa é independente do modelo, embora as escolhas específicas influenciem as características dos caminhos na simulação. Se o modelo Black Scholes for usado para treinamento, o processo resultará em estratégias de *hedging* sob os parâmetros do Black Scholes. Se o modelo de Heston for escolhido, as estratégias se adaptam. Levando esta observação um passo adiante, Horvath pergunta: “E se os dados de treinamento forem semelhantes aos dados do próprio mercado?” Na verdade, explorar esta questão produz resultados realistas que

não sofrem pelas deficiências dos modelos de Black Scholes ou Heston.

A modelagem generativa é construída em torno de um tema — o teorema da aproximação universal, que sustenta que se houver nós suficientes na rede, é possível aproximar qualquer função ou distribuição. Então, quer a distribuição seja uniforme ou normal padrão, ela será mapeada para a distribuição de retornos alvo vista nos dados. As formas mais comuns de modelagem generativa são as máquinas de Boltzmann restritas (RBM, na sigla em inglês) e, mais recentemente, as redes generativas adversas (GANs, na sigla em inglês), que apresentam bom desempenho em imagens. No entanto, este recurso é caro e desnecessário ao trabalhar com dados de séries temporais financeiras. Em vez disso, os Autoencoders Variacionais (VAEs, na sigla em inglês) são os melhores candidatos neste caso, tendo uma vantagem distinta por serem parcimoniosos e capazes de operar em ambientes onde poucos dados estão disponíveis.

Naturalmente, o mercado conta com diversas situações nas quais existe uma grande quantidade de dados, mas há desafios quando este não é o caso ou os dados são privados e não podem ser acessados facilmente. Nestes casos, restrições e gargalos podem ocorrer ao usar dados reais para o mecanismo de hedging. Além disso, ao modelar um processo estocástico por completo, a observação é realizada em um padrão temporal individual, e haverá lacunas onde faltar clareza sobre o que aconteceu entre os marcos temporais. Um modelo generativo fornece flexibilidade em cada ponto de observação, sem fazer suposições sobre o que acontece entre eles. A análise agrupa-se rapidamente, mesmo com um pequeno conjunto de dados.

Para avaliar o desempenho, Horvath executou um modelo generativo baseado em retornos e um modelo generativo baseado em caminhos, e calculou os retornos diários, semanais e mensais. Ao definir uma configuração baseada em retornos,

elementos de duas distribuições são explorados e, se um teste bilateral for aplicado, suas origens se tornarão evidentes — sejam eles originados das mesmas distribuições ou de distribuições diferentes. Além disso, o teste pode dizer com um determinado nível de confiança se os elementos se originam do mesmo processo estocástico ou de processos diferentes, embora ainda seja robusto com relação a dados ausentes ou dados irregulares de amostra. Horvath descobriu que o desempenho era satisfatório em termos de deep hedging; um possível caso de uso é a negociação de alta frequência, na qual as observações podem variar amplamente de um dia para o outro. Pesquisas futuras abordarão caminhos de mercado, envolvendo trabalho com VAE condicional e pós-processamento.

Na sessão de perguntas e respostas, o anfitrião do evento, Bruno Dupire, questionou sobre os conceitos de assinaturas, de decomposição e de comparação de resultados em diferentes mercados. Horvath sugeriu o trabalho de Terry Lyons sobre computação determinística com rough path — teoria da década de 90 que propõe uma abordagem para equações diferenciais e integrais ao longo de caminhos irregulares. Claramente, o mercado está em constante movimento, de forma que tanto o ruído quanto os eventos extremos dificultam a interpretação. No entanto, a modelagem generativa pode capturar caudas longas (fat tails) — quando projetarmos a partir do caminho do mercado, as caudas estarão lá, afirma.

## **Núcleo de Inteligência – Sedet**

### **Edição 287 - Em 22 de novembro de 2021**

Os textos do conteúdo exposto neste informativo não são de autoria do Governo do Estado do Ceará.