

# Previsão da superfície de volatilidade do Ibovespa utilizando técnicas de Machine Learning

---

Daniel Caueh Dunaiski Figueira Leal

A decorative vertical bar on the left side of the slide, featuring a golden-yellow background with various 3D financial symbols like the dollar sign (\$), yen sign (¥), and Euro sign (€) in a repeating pattern.

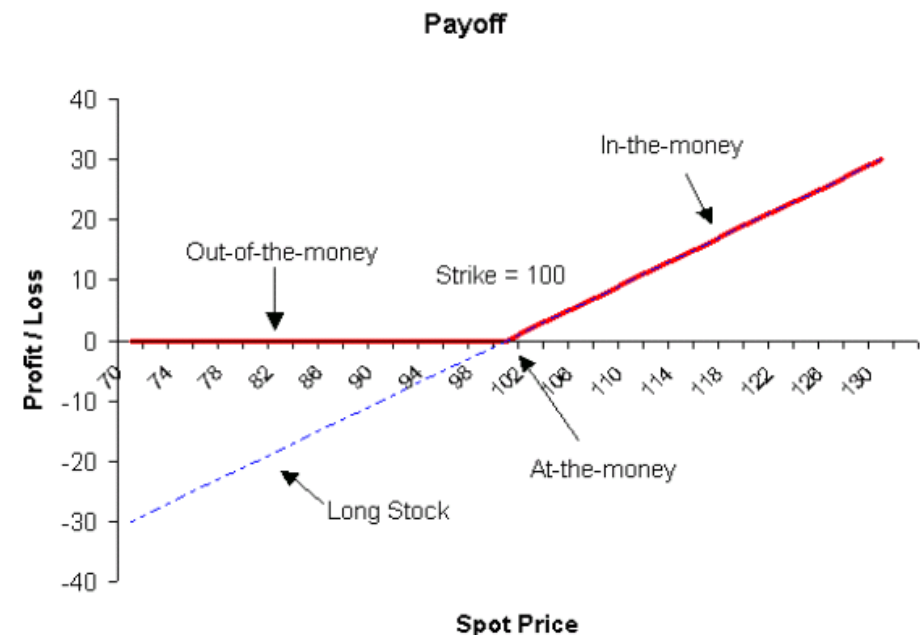
# Trabalhos Relacionados

---

- Huynh, S. (2018). Modelling and Forecasting Implied Volatility Using Neural Network.
- Liu, S.; Oosterlee, C. W.; Bohte, S. M.(2018). Pricing options and computing implied volatilities using neural networks.
- Johnsson, O. (2018). Predicting Stock Index Volatility Using Artificial Neural Networks. An empirical study of the OMXS30, FTSE100 & S&P/ASX200.

# Opções

- Opção é um ativo financeiro que dá ao titular o direito, mas não a obrigação, de comprar ou vender um determinado ativo a um determinado preço a um determinado preço.
- Compra: Call / Venda: Put
- Ativo objeto/underlying: determinado ativo
- Strike: determinado preço
- Vencimento: determinado preço



# Precificação de opções

---

- Black-Scholes (1973)
- Nobel de Economia 1997

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0$$

# Precificação de opções

---

- $S_0$ : Spot do ativo objeto (preço hoje)
- $K$ : Strike
- $\sigma$ : Volatilidade
- $r$ : Taxa livre de risco
- $q$ : Taxa de dividend/cost of carry
- $t$ : Tempo até o vencimento

# Precificação de opções

---

$$C = S_0 e^{-qt} * N(d_1) - K e^{-rt} * N(d_2)$$

$$P = K e^{-rt} * N(-d_2) - S_0 e^{-qt} * N(-d_1)$$

$$d_1 = \frac{\ln\left(\frac{S_0}{K}\right) + t\left(r - q + \frac{\sigma^2}{2}\right)}{\sigma\sqrt{t}}$$

$$d_2 = d_1 - \sigma\sqrt{t}$$



# Gregas

---

- Sensibilidade em relação algum parâmetro

$$\Delta = \frac{\partial C}{\partial S_0} = e^{-qt} * N(d_1)$$

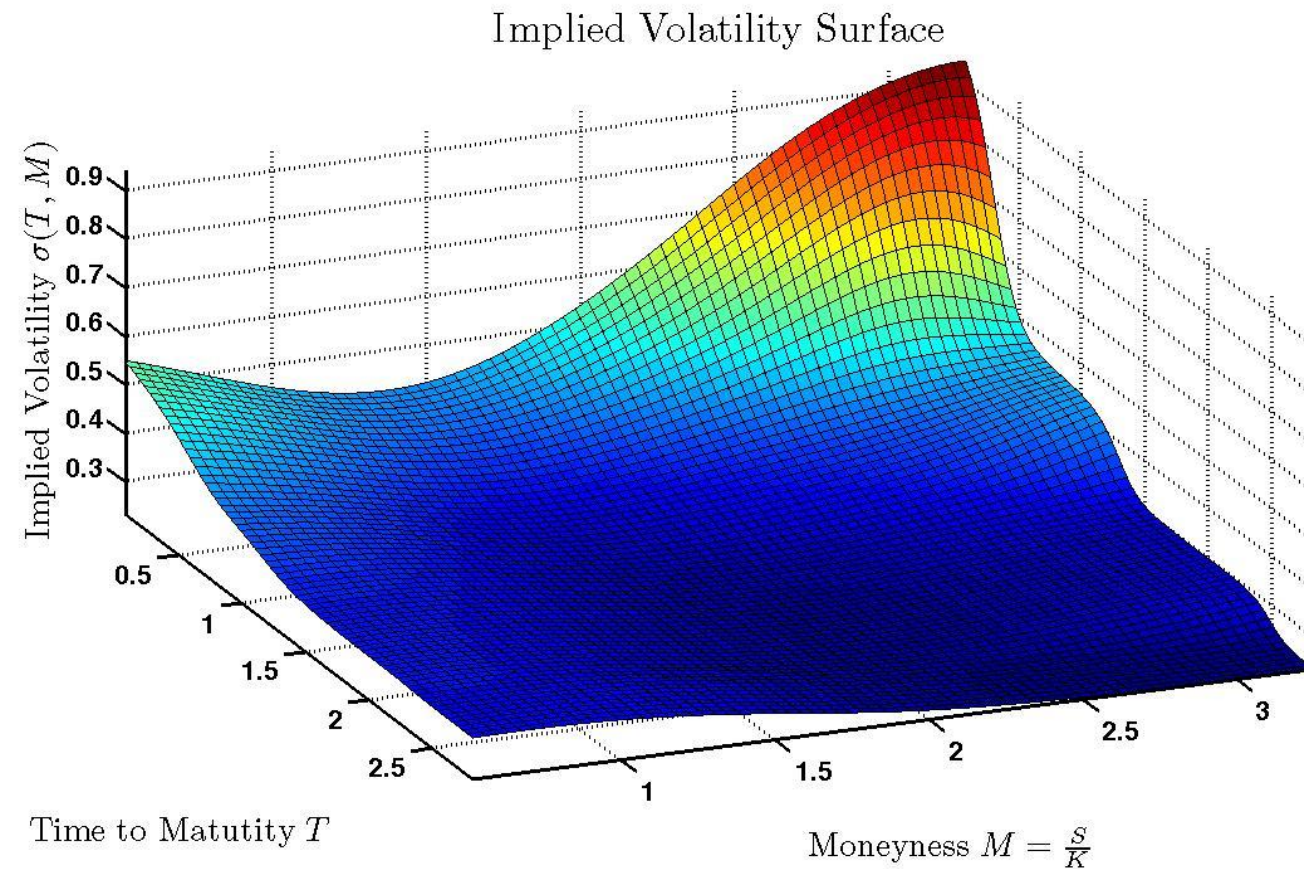
$$\nu = \frac{\partial C}{\partial \sigma} = \frac{1}{100} S_0 e^{-qt} \sqrt{t} \frac{1}{\sqrt{2\pi}} e^{-\frac{d_1^2}{2}}$$

$$m = \frac{\ln\left(\frac{F}{K}\right)}{\sigma\sqrt{t}}$$

$$\Gamma = \frac{\partial^2 C}{\partial S_0^2} = \frac{e^{-qt}}{S_0 \sigma \sqrt{t}} \frac{1}{\sqrt{2\pi}} e^{-\frac{d_1^2}{2}}$$

$$\rho = \frac{\partial C}{\partial r} = \frac{1}{100} K t e^{-rt} N(d_2)$$

# Superfície de Volatilidade





# Ibovespa

---

- Principal índice de ações brasileiras
- Simula uma carteira teórica contendo as ações listadas mais negociadas
- Vários derivativos negociados sobre o Ibovespa

ITUB4	9.231
VALE3	8.202
PETR4	7.382
BBDC4	7.131
B3SA3	5.405
PETR3	5.218
ABEV3	4.105
BBAS3	3.550
ITSA4	3.357
JBSS3	2.426








# Dataset

---
































- Arquivos disponíveis no FTP da Bolsa:
  - Indicadores Econômicos e Agropecuários
  - Prêmio de Referência para Opções sobre Ações
  - Superfície de Volatilidade em Delta para opções sobre Ações
  - Taxas de Mercado para Swaps
- Problema: Falta de clareza, arquivos diários, formato .ex\_, atualização
- Data inicial: 04/11/2016

## Índice de /

Nome	Tamanho	Data da modificação
 agropecuarios/		06/04/2004 11:54:00
 Ativos/		22/06/2010 21:49:00
 BD/		25/04/2010 19:11:00
 BdDocs/		15/12/2016 13:32:00
 BdDocs2/		18/08/2015 15:50:00
 BdDocs3/		15/04/2014 08:55:00
 BDPREGAO_BACEN/		26/11/2015 12:06:00
 CadastroIndicadores/		23/04/2007 15:09:00
 CCbancosHabilitados/		22/06/2018 03:03:00
 CCTaxa/		10/09/2019 16:01:00
 CCVolume/		10/09/2019 16:01:00
 Cenario	4.4 kB	15/10/2010 19:58:00
 Cenarios/		15/08/2014 17:43:00
 CenariosAgro/		15/08/2014 18:58:00
 CenariosSwap/		15/08/2014 17:46:00
 CodigosISINCPRs/		10/09/2019 16:04:00
 CodigosISINDerivativos/		10/09/2019 16:04:00
 CodigosISINSwap/		10/09/2019 16:06:00
 Coletiva/		29/08/2013 19:02:00
 ContratosCadastrados/		20/09/2017 21:32:00
 ContratosCadastradosDescompactados/		20/09/2017 21:31:00
 ContratosCadastradosTeste/		17/07/2006 09:36:00
 ContratosConvertidos/		17/07/2006 09:36:00
 ContratosEletronico/		10/09/2019 17:55:00
 ContratosEnergia/		17/07/2006 09:35:00
 ContratosPregaoAfterHour/		25/08/2017 16:57:00
 ContratosPregaoAjuste/		25/08/2017 16:23:00
 ContratosPregaoArbitragem/		25/08/2017 12:03:00
 ContratosPregaoAtualizado/		25/08/2017 18:32:00
 ContratosPregaoFinal/		25/08/2017 16:52:00
 ContratosPregaoFinalTeste/		17/07/2006 09:27:00
 ContratosPregaoLeilaoBacen/		02/08/2018 16:38:00
 ContratosPregaoPrevia/		25/08/2017 16:54:00

## Índice de /TaxasSwap/

 [diretório pai]

Nome	Tamanho	Data da modificação
 backup_TS040601.backup	141 kB	01/06/2004 17:29:00
 TS000114.ex_	19.9 kB	14/01/2000 19:46:00
 TS000117.ex_	19.8 kB	17/01/2000 19:20:00
 TS000118.ex_	19.8 kB	18/01/2000 17:47:00
 TS000119.ex_	19.8 kB	19/01/2000 20:39:00
 TS000120.ex_	19.9 kB	20/01/2000 19:18:00
 TS000121.ex_	19.9 kB	21/01/2000 19:12:00
 TS000124.ex_	19.9 kB	24/01/2000 19:15:00
 TS000126.ex_	19.8 kB	26/01/2000 18:51:00
 TS000127.ex_	19.8 kB	27/01/2000 19:03:00
 TS000128.ex_	19.8 kB	28/01/2000 19:52:00
 TS000131.ex_	19.8 kB	31/01/2000 19:27:00
 TS000201.ex_	19.9 kB	01/02/2000 20:05:00
 TS000202.ex_	19.8 kB	02/02/2000 19:23:00
 TS000203.ex_	19.9 kB	03/02/2000 18:52:00
 TS000204.ex_	19.8 kB	04/02/2000 19:04:00
 TS000207.ex_	19.9 kB	07/02/2000 18:59:00
 TS000208.ex_	19.8 kB	08/02/2000 18:57:00
 TS000209.ex_	19.9 kB	09/02/2000 19:06:00
 TS000210.ex_	19.8 kB	10/02/2000 18:58:00
 TS000211.ex_	19.8 kB	12/02/2000 07:41:00
 TS000214.ex_	19.9 kB	14/02/2000 19:01:00
 TS000215.ex_	19.8 kB	15/02/2000 19:12:00
 TS000216.ex_	19.9 kB	16/02/2000 19:14:00
 TS000217.ex_	19.8 kB	17/02/2000 19:40:00
 TS000218.ex_	19.8 kB	18/02/2000 19:38:00
 TS000221.ex_	19.9 kB	21/02/2000 19:23:00
 TS000222.ex_	19.9 kB	22/02/2000 16:19:00
 TS000223.ex_	19.9 kB	23/02/2000 19:01:00
 TS000224.ex_	19.9 kB	24/02/2000 16:11:00
 TS000225.ex_	19.9 kB	25/02/2000 18:57:00

# Problemas do Dataset

---

- Data de vencimento constante → Tempo até o vencimento muda todos os dias
- Solução → Criar pontos teóricos com vencimento constante: 1M, 3M, 6M, 9M, 12M, 18M
- Interpolação na Variância (também utilizado para precificação de derivativos balcão)

$$\sigma(T) = \sqrt{\frac{T - T_i}{T_{i+1} - T_i} \frac{T_{i+1}}{T} \sigma^2(T_{i+1}) + \frac{T_{i+1} - T}{T_{i+1} - T_i} \frac{T_i}{T} \sigma^2(T_i)}$$

# Problemas do Dataset

---

- Também é necessário interpolação da taxa de juros pré e do futuro de índice.

$$r = \left\{ \left( 1 + \frac{r_{ant}}{100} \right)^{\frac{ndu_{ant}}{252}} * \left[ \frac{\left( 1 + \frac{r_{post}}{100} \right)^{\frac{ndu_{post}}{252}}}{\left( 1 + \frac{r_{ant}}{100} \right)^{\frac{ndu_{ant}}{252}}} \right]^{\frac{(ndu - ndu_{ant})}{(ndu_{post} - ndu_{ant})}} \right\}^{\frac{252}{ndu}}$$



A decorative vertical bar on the left side of the slide, featuring a golden-yellow background with various financial symbols (dollar sign, yen sign, euro sign, pound sign, etc.) and currency notes in a 3D, embossed style.

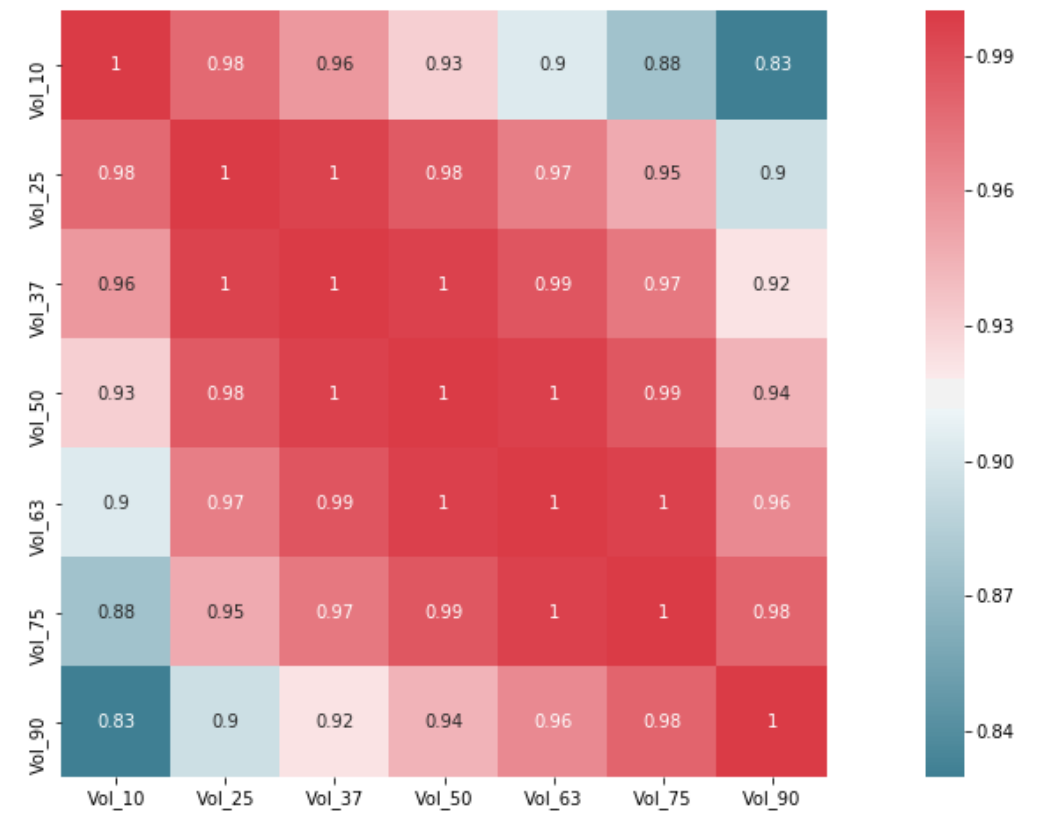
# Features iniciais

---

- Tempo até o vencimento
- Spot Ibov
- Spot CDI
- Futuro Ibov
- Futuro DI
- Delta
- Strike
- $r$
- $q$
- $d1$
- $d2$
- Gamma
- Vega
- Rho
- Moneyness

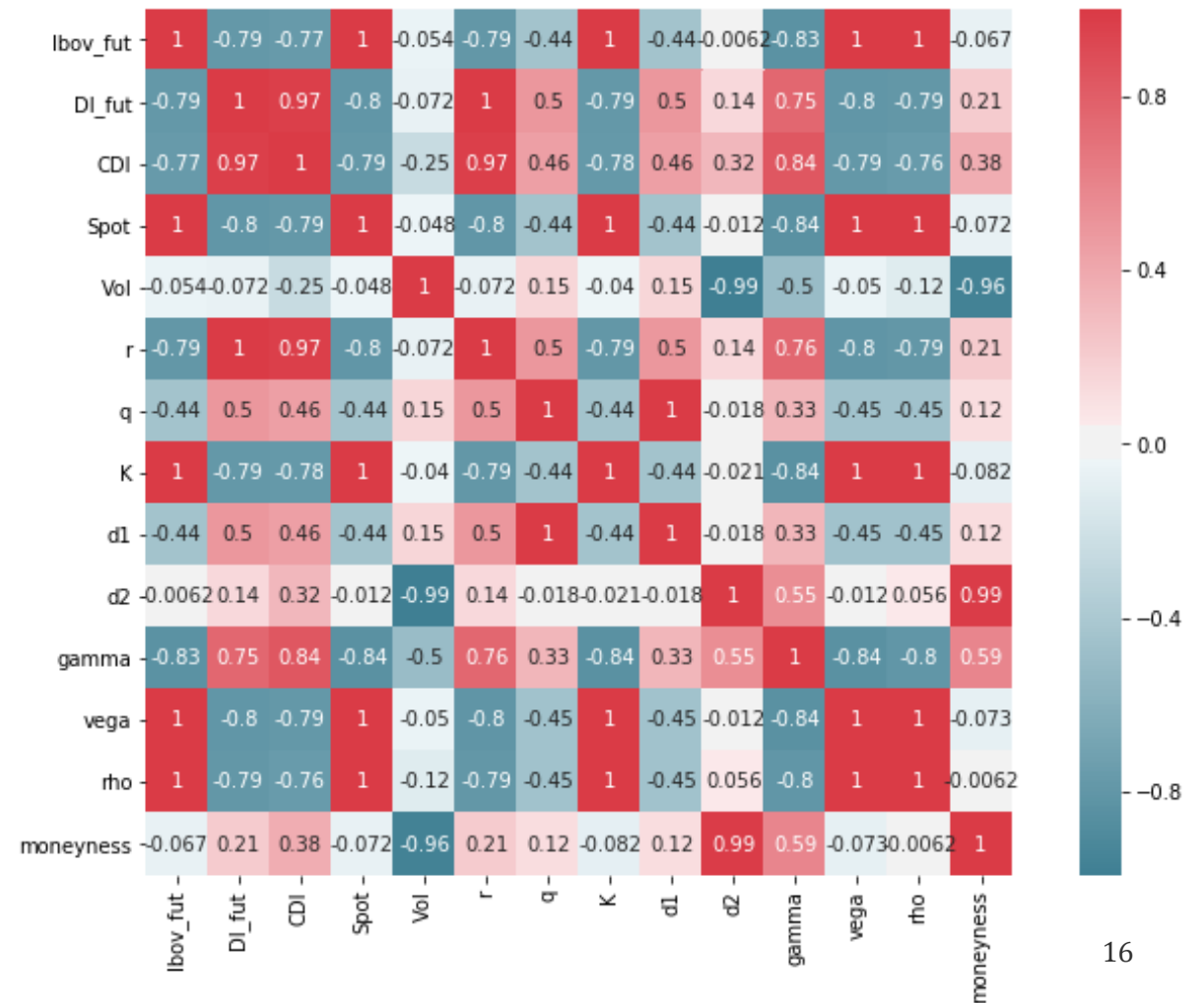
# Problemas features iniciais

- Algumas features possuem altíssima correlação, principalmente comparando vencimentos/Deltas diferentes
- Solução: Considerar apenas um par tempo até vencimento/Delta
- 12M / 50 Delta



# Problemas features iniciais

- Correlações muito fortes nas features auxiliares
- Solução: remover/transformar features
- Features finais:
  - Vol
  - Spot
  - Moneyness
  - Ibov Futuro/Ibov Spot



# Split do Dataset

---



**Split temporal**



**70% treino**

04/Nov/2016 até  
29/Out/2018



**30% teste**

30/Out/2018 até  
30/Ago/2019





# Tentativas de solução

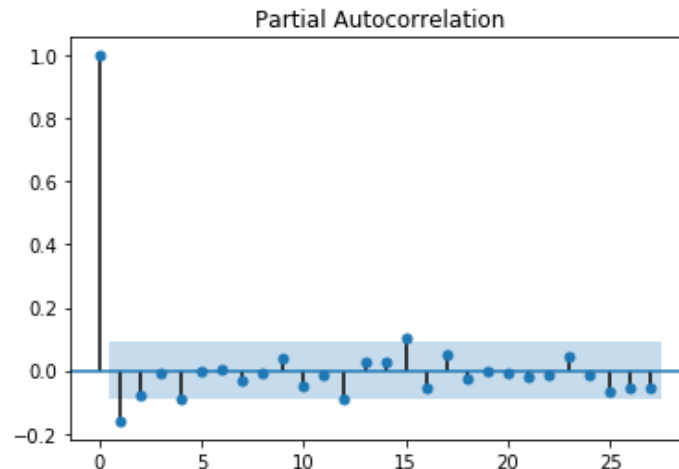
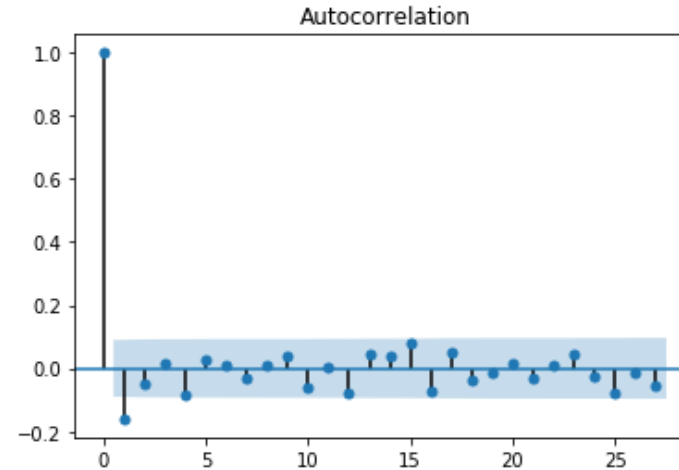
---

- Univariado
  - ARIMA
  - ANN
  - RNN GRU
- Multivariado
  - VAR
  - ANN
  - RNN GRU x2



# Univariado

- ARIMA
  - Dados não estacionários
  - Parâmetros iniciais dados pelos gráficos
  - Busca refinada em torno desses valores buscando menor AIC
  - Ordem final (2,12)





# Univariado

---

- ANN

- Input: 10 dias anteriores
- Camadas:
  - 20 Neurônios, ReLU, Lasso
  - 10 Neurônios, ReLU, Lasso
  - Saída, 1 Neurônio, Linear, Lasso

- RNN

- Input: 10 dias anteriores
- Camadas:
  - 50 Neurônios GRU, Tanh
  - Saída, 1 Neurônio



# Multivariado

---

- VAR
  - Vector AutoRegressive
  - Extensão do modelo AR para casos multivariados
  - Busca pelo número ideal de lags
  - Melhor AIC  $\rightarrow$  7 dias
- ANN
  - Input: 10 dias anteriores
  - Camadas:
    - 60 Neurônios, ReLU, Lasso
    - 40 Neurônios, ReLU, Lasso
    - 20 Neurônios, ReLU, Lasso
    - Saída, 1 Neurônio, Linear, Lasso



# Multivariado

---

- RNN com todos os dados
  - Input: 10 dias anteriores, todos os dados
  - Camadas:
    - 40 Neurônios GRU, Tanh
    - Saída, 1 Neurônio Linear
- RNN apenas com os auxiliares
  - Input: 10 dias anteriores, apenas os dados auxiliares.
  - Camadas:
    - 10 Neurônios, Tanh
    - 5 Neurônios, Tanh
    - Saída, 1 Neurônio, Linear

A decorative vertical bar on the left side of the slide, featuring a golden-yellow background with various currency symbols (Dollar, Euro, Pound, Yen, etc.) in a 3D, embossed style.

# Resultados

---

- Métricas utilizadas:
  - RMSE: Erro Quadrático Médio
  - $R^2$
  - MAE: Erro Absoluto Médio
  - MAPE: Erro Percentual Absoluto Médio



# Resultados

---

	Univariado			Multivariado			
	ARIMA	ANN	RNN	VAR	ANN	RNN	RNN só aux
RMSE	0.475	<b>0.444</b>	0.567	0.509	0.569	<b>0.464</b>	0.863
R <sup>2</sup>	0.854	0.869	<b>0.875</b>	0.866	0.842	<b>0.885</b>	0.745
MAE	0.350	<b>0.317</b>	0.462	0.392	0.455	<b>0.342</b>	0.660
MAPE	1.629	<b>1.471</b>	2.156	1.841	2.104	<b>1.592</b>	3.083

# Resultados

---

- Univariado > Multivariado
  - Variáveis auxiliares não servem para explicar o problema, e geram ruído
  - B3 utiliza dados de corretoras para montar a superfície oficial
  - Altamente dependente das negociações do dia e da qualidade das informações das corretoras
  - RNN só com variáveis auxiliares teve os piores resultados → O problema tem fortes características auto regressivas.
- ANN > ARIMA
  - Técnicas modernas conseguem resultado melhor que técnicas clássicas
  - Input da ANN é mais cru que do ARIMA: não-estacionariedade
  - Maior complexidade e não linearidade levou a um resultado melhor

A decorative vertical bar on the left side of the slide, featuring a golden-yellow background with various 3D currency symbols like the dollar sign (\$), yen sign (¥), and euro sign (€) in a repeating pattern.

# Trabalhos Futuros

---

- Conseguir dados sobre a volumetria do mercado
- Melhorar a arquitetura das redes, principalmente as RNN
- Tentar retreinar o modelo durante o período de teste utilizando uma quantidade limitada de dados → Dados muito antigos não tem grande influência quanto dados recentes
- Previsão para períodos maiores que um dia

# Bibliografia

---

Huynh, S. (2018). Modelling and Forecasting Implied Volatility Using Neural Network.

Liu, S.; Oosterlee, C. W.; Bohte, S. M. (2018). Pricing options and computing implied volatilities using neural networks.

Johnsson, O. (2018). Predicting Stock Index Volatility Using Artificial Neural Networks. An empirical study of the OMXS30, FTSE100 & S&P/ASX200.

Hutchinson, J. M.; Lo, A. W.; Poggio, T. (1994). A Nonparametric Approach to Pricing and Hedging Derivative Securities Via Learning Networks

Hahn, T. (2014). Option Pricing Using Artificial Neural Networks : an Australian Perspective.

Stafford, D. (2018). Machine Learning in Option Pricing.

Hamid, S. A. (2004). Primer on Using Neural Networks for Forecasting Market Variables.

Hull, J. (2012), "Options, Futures, and Other Derivatives". Boston, Prentice Hall

Black, F., Scholes, M. (May - Jun., 1973), "The Pricing of Options and Corporate Liabilities", Journal of Political Economy 81, no. 3 , pp 637-654

Gujarati, D. N., Porter, D. C. (2009), "Basic econometrics", Boston, Mass: McGraw-Hill

Goodfellow, I. J. , Bengio, Y., Courville, A. (2016), "Deep Learning", MIT Press

Iason Ltd. (2007), "The Implied Volatility Surfaces". Disponível em: <[http://www.iasonltd.com/wpupload/Download/All/2013 The implied volatility surfaces.pdf](http://www.iasonltd.com/wpupload/Download/All/2013%20The%20implied%20volatility%20surfaces.pdf)>. Acesso em: 11/11/2019

Bolsa de Mercadorias e Futuros (2004), "Metodologia para a apuração de curvas de preços e de spreads teóricos de títulos públicos". Disponível em: <<http://www.bmf.com.br/bmfbovespa/pages/clearing1/ativos/documentos/pdf/Metodologia>>. Acesso em: 11/11/2019

A decorative vertical bar on the left side of the slide, featuring a golden-yellow background with various 3D financial symbols like the dollar sign (\$), yen sign (¥), and Euro sign (€) in a repeating pattern.

# Alvo de publicação

---

- Alvos:
  - BBR - Brazilian Business Review
  - EBF - Encontro Brasileiro de Finanças
  - EnANPAD - Encontro da Associação Nacional de Pós-Graduação e Pesquisa em Administração
  - SEMEAD – Seminários em Administração – FEA USP



A decorative vertical bar on the left side of the slide, featuring a gold color and a pattern of various currency symbols (dollar, euro, yen, etc.) in a 3D, embossed style.

# Perguntas/Sugestões