

Universidade Estadual de Maringá

Ciência da Computação

9788/01 - Introdução à inteligência artificial

Prof. Wagner Igarashi

Correlação entre notícias financeiras e
índice bovespa com *Python*

Thiago Yasunaka 103069

Maringá - PR

Problema a ser resolvido

- Analisar preço de ativos em um período de 24 dias
- Analisar notícias relacionadas ao Bovespa em um período de 20 dias
- Dataset para treinamento da linguagem natural do mercado financeiro
 - disponível em: <https://www.kaggle.com/>
- Treinar o dataset usando a biblioteca TextBlob
- Obter uma nota para cada notícia usando com o classificador de *Naive Bayes*
- Gerar uma análise de correlação entre notícias e preço do índice Bovespa

Requisitos Principais

- Python3
- TextBlob
- NewsAPI
- Yahoo Finance

Como executar?


- Abrir o diretório *src*
- Executar no terminal *python3 main.py*
- Será exibido os gráficos, são três no total. Ao fechar um gráfico é exibido o outro *(Um exemplo de uso está no final da apresentação)*

Código Fonte




```
1 def getArticles():
2     startDate = (datetime.datetime.now() -
3                 datetime.timedelta(days=constants.WEEK_DAYS)).strftime("%Y-%m-%d")
4     endDate = datetime.datetime.now().strftime("%Y-%m-%d")
5
6     url = ('https://newsapi.org/v2/everything?'
7           'q=ibovespa&'
8           'language=pt&'
9           'sortBy=popularity&'
10          'from='+startDate+'&'
11          'to='+endDate+'&'
12          'apiKey=04f7c388507549e0984bf138d88af82e')
13
14     response = requests.get(url)
15     page1 = response.json()['articles']
16
17     return np.concatenate((page1, []))
```

Código Fonte



```
1 def converterCSVToTextBlob():
2     path = '../data/dataset.csv'
3     with open(path, 'r') as file:
4         reader = csv.reader(file)
5         train = []
6         for row in reader:
7             phrase = row[1]
8             status = row[0] # negative, neutral, positive
9             if(status == 'negative'):
10                 train.append((phrase, 'neg'))
11             if(status == 'positive'):
12                 train.append((phrase, 'pos'))
13     return train
14
```

Código Fonte



```
1 def getStockDetails():
2     startDate = (datetime.datetime.now() -
3                 datetime.timedelta(days=constants.STOCK_QUANTITY_DAYS)).strftime("%m-%d-%Y")
4     endDate = datetime.datetime.now().strftime("%m-%d-%Y")
5
6     df = web.DataReader(f'^BVSP', data_source='yahoo',
7                        start=startDate, end=endDate)
8     opens = np.array(df['Open'])
9     closes = np.array(df['Close'])
10
11     variations = []
12     for i in range(len(opens)):
13         variations.append((closes[i] - opens[i]) / opens[i])
14
15     return variations[len(closes)-constants.NUMBER_OF_VARIATIONS:len(closes)]
```

Código Fonte



```
1 classifier = NaiveBayesClassifier(train)
```

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$


Código Fonte



```
1 translation = translator.translate(text=text)
2 prob_dist = classifier.prob_classify(translation)
3 pos_prob = round(prob_dist.prob("pos"), 2)
4 neg_prob = round(prob_dist.prob("neg"), 2)
```

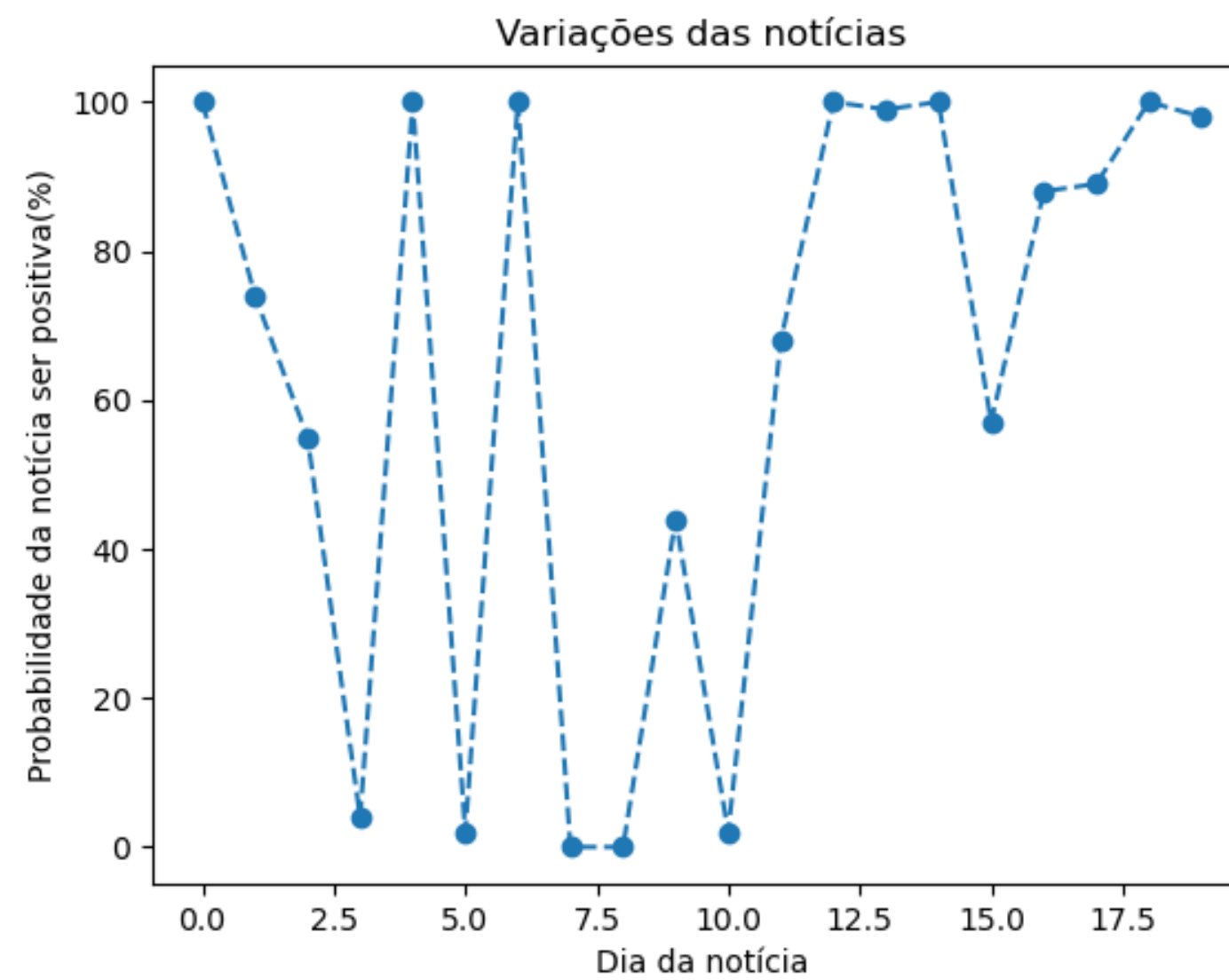
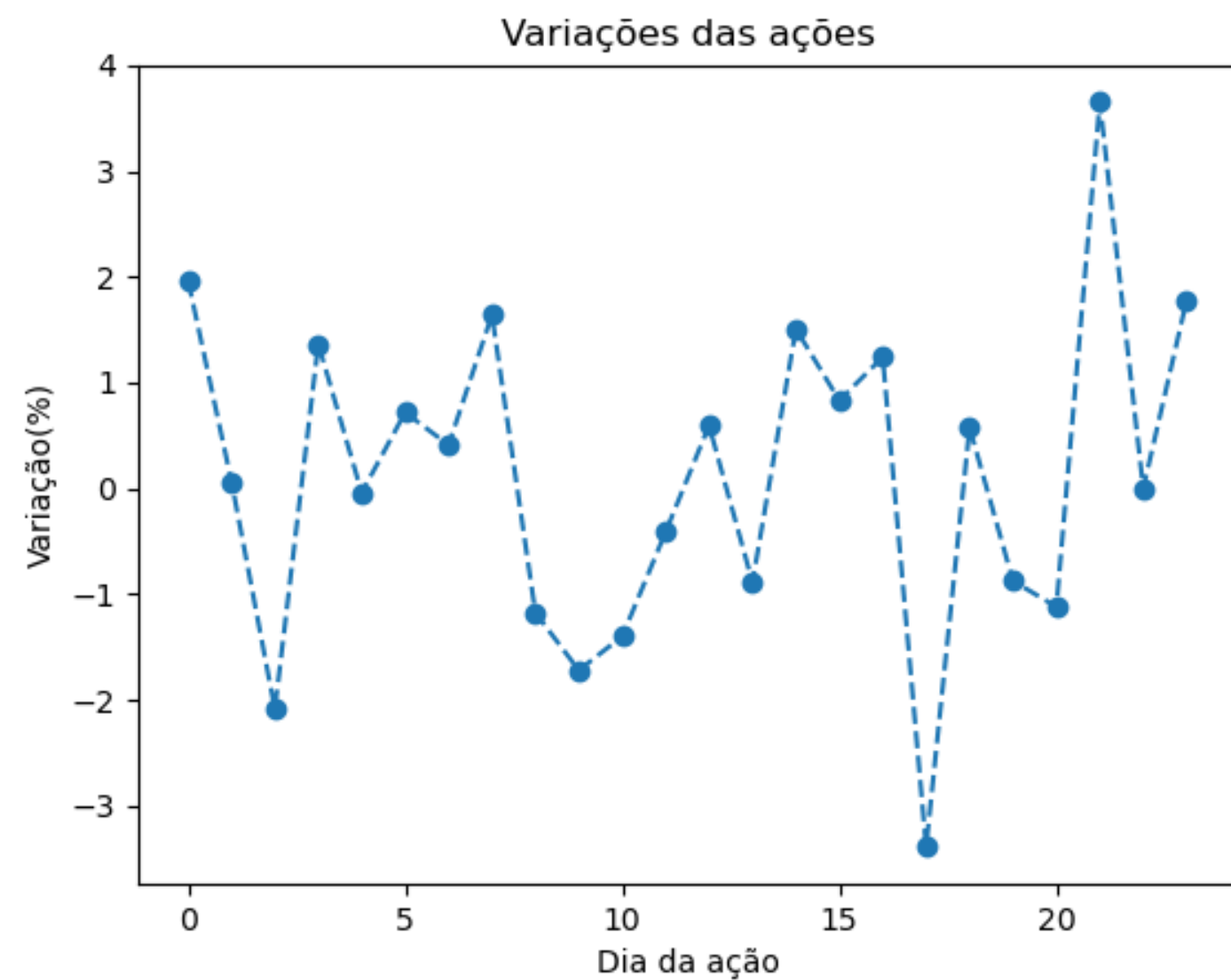

Código Fonte

Coeficiente de correlação de Pearson

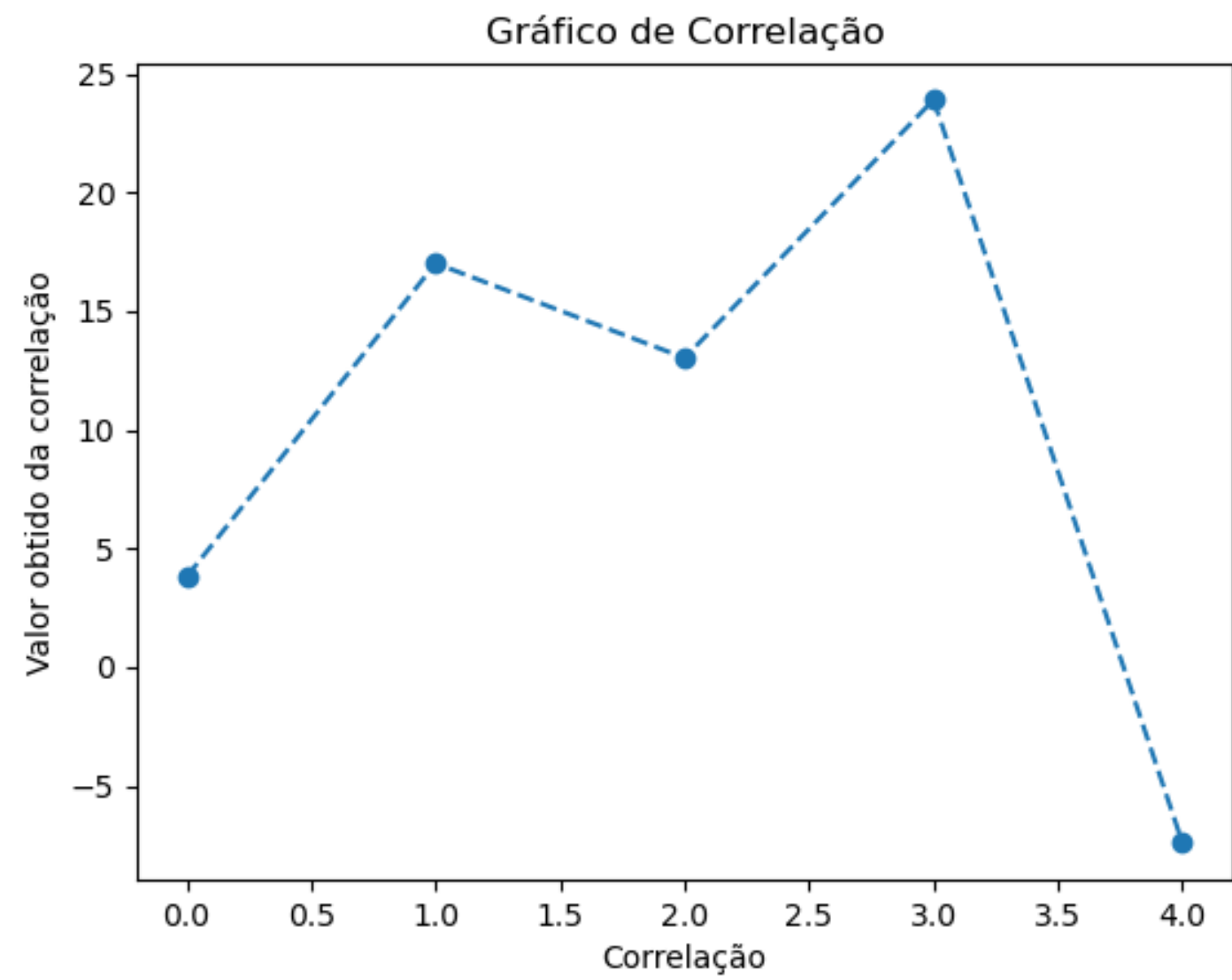


```
1 def generateCorrelationForStocks(variations, news):
2     corrs = []
3     for i in np.arange(constants.NUMBER_OF_CORRELATIONS):
4         start = i
5         end = i + constants.NUMBER_OF_NEWS
6         range = np.arange(start, end)
7         corr = np.corrcoef(news, np.take(variations, range))
8         corrs.append(corr[0][1])
9     return corrs
```

Resultados



Resultados



Bibliografia

- R. Gandhi. Naive bayes classifier. <https://towardsdatascience.com/naive-bayes-classifier-81d512f50a7c>. Accessed: 2021-12-07.
- A. Mill. Tudo o que você precisa saber sobre economia. 3. Gente, 2017.
ISBN 8545201699.
- S. J. Russel and P. Norvig. Inteligência Artificial. 3. Campus, 2010. ISBN 978-0136042594