



Universidade Estadual de Maringá



Departamento de Informática

# ANÁLISE DE SENTIMENTO: UM ESTUDO DA CORRELAÇÃO DE NOTÍCIAS E OS VALORES DAS AÇÕES DE EMPRESA

*Discente: Eduardo H. K. Shibukawa*

*Orientador: Prof. Wagner Igarashi, Dr.*

*Membros da banca:*

*Prof. Yandre Maldonado e Gomes da Costa, Dr.*

*Profa. Valéria Delisandra Feltrim, Dra.*

Maringá, 2017

# Sumário

## **Agradecimentos**

Introdução

Objetivos

Fundamentação Teórica

Implementação

Resultados

Considerações finais

Referências



# Agradecimientos

# Sumário

Agradecimentos

**Introdução**

Objetivos

Fundamentação Teórica

Implementação

Resultados

Considerações finais

Referências

# Introdução

- Em meio à crise que nossa economia vem passando, muitas pessoas têm procurado outros meios de conseguir renda extra, uma dessas formas são os investimentos
- As ações são um tipo de investimento no qual se necessita de conhecimento financeiro e além disso, do conhecimento sobre a influência do mercado nas empresas em que se quer investir.

# Introdução

Diante dos fatos apresentados:

- Na questão da influência do mercado nas empresas, este trabalho irá focar o impacto das notícias na variação do preço de ações.

# Sumário

Agradecimentos

Introdução

**Objetivos**

Fundamentação Teórica

Implementação

Resultados

Considerações finais

Referências

# Geral

Análise de sentimentos de notícias, com a finalidade de verificar a possível correlação entre as notícias sobre uma empresa e as variações do preço de suas ações.



# Específicos

- Extração de notícias de sites.
- Análise de sentimento das notícias extraídas.
- Análise da correlação dos dados da análise de sentimento com os dados de valorização ou desvalorização das mesmas.

# Sumário

Agradecimentos

Introdução

Objetivos

**Fundamentação Teórica**

Implementação

Resultados

Considerações finais

Referências

# Mercado financeiro

- O sistema financeiro é composto por instituições econômicas que tem como objetivo a intermediação entre poupadores e investidores (SELAN, 2015)
- Os poupadores depositam e emprestam dinheiro para instituições econômicas, e estas utilizam este dinheiro para financiar alguns setores da economia que estão precisando de recursos. (BERTOLO, 2002)

# Mercado de ações

- O mercado onde se negocia ações.
- As ações são a menor parcela de capital de uma empresa. (BERTOLO, 2002)
- O preço das ações está relacionado diretamente à sua oferta e procura. Quanto maior sua a procura mais seu preço irá aumentar. (BERTOLO, 2002)
- Sua procura pode estar relacionada a vários fatores tais como: político, econômico, estratégias das empresas, etc. (BERTOLO, 2002)

# **Análise técnica e fundamentalista**

- A análise técnica se concentra no estudo do mercado de ações, utiliza de varias ferramentas, tais como indicadores e desenhos gráficos na identificação de tendências. (Murphy, 1999)
- A análise fundamentalista tem como foco fatores externos, indicadores financeiros da empresa, o mercado econômico, etc. (Murphy, 1999)

# Inteligência artificial

Pensando como um humano	Pensando racionalmente
<p>“O novo e interessante esforço para fazer os computadores pensarem (...) <i>máquinas com mentes</i>, no sentido total e literal.” (Haugeland, 1985)</p> <p>“[Automatização de] atividades que associamos ao pensamento humano, atividades como a tomada de decisões, a resolução de problemas, o aprendizado...” (Bellman, 1978)</p>	<p>“O estudo das faculdades mentais pelo uso de modelos computacionais.” (Charniak e McDermott, 1985)</p> <p>“O estudo das computações que tornam possível perceber, raciocinar e agir.” (Winston, 1992)</p>
Agindo como seres humanos	Agindo racionalmente
<p>“A arte de criar máquinas que executam funções que exigem inteligência quando executadas por pessoas.” (Kurzweil, 1990)</p> <p>“O estudo de como os computadores podem fazer tarefas que hoje são melhor desempenhadas pelas pessoas.” (Rich and Knight, 1991)</p>	<p>“Inteligência Computacional é o estudo do projeto de agentes inteligentes.” (Poole <i>et al.</i>, 1998)</p> <p>“AI... está relacionada a um desempenho inteligente de artefatos.” (Nilsson, 1998)</p>

Algumas definições de inteligência artificial, organizadas em quatro categorias. (RUSSELL; NORVIG, 2013)

# Análise de sentimento

Segundo Liu (2012), a análise de sentimento tem como função identificar a opinião expressada sobre um assunto em particular e pode ser dividida em 3 níveis:

- **Nível de documento:** tem como objetivo classificar a opinião de um documento inteiro, este nível de análise leva em consideração que cada documento expressa uma opinião sobre uma única entidade.
- **Nível de sentença:** tem como objetivo analisar se a sentença foi expressada como positiva, negativa ou neutra, sendo fortemente relacionada com a análise de subjetividade citada anteriormente.
- **Nível de entidade ou aspecto:** é baseada na ideia de que uma opinião consiste em um sentimento positivo ou negativo, e um alvo. Seu objetivo é descobrir o sentimento de entidades e suas diferentes características, é o nível mais complexo de análise.

# Aprendizado de máquina

Segundo Barber (2016), os dois principais campos da aprendizagem de máquina podem ser considerado o:

- **Aprendizado supervisionado:** tem o foco de melhorar a acurácia dos resultados da predição.
- **Aprendizado não supervisionado:** encontrar de descrições plausíveis sobre os dados



# Sumário

Agradecimentos

Introdução

Objetivos

Fundamentação Teórica

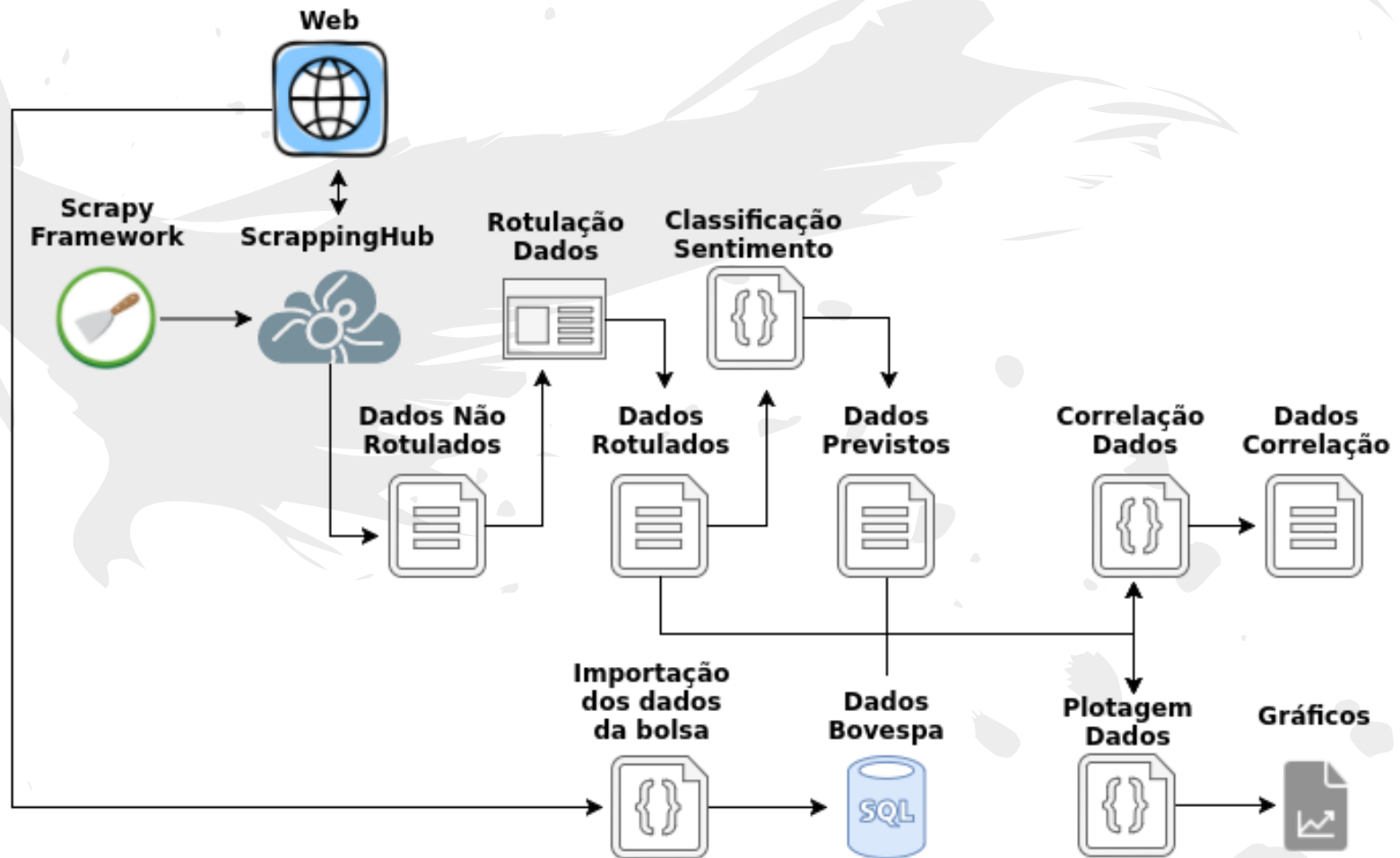
**Implementação**

Resultados

Considerações finais

Referências

# Arquitetura



# Importação dos dados das ações



python



SQLite



# Importação dos dados das ações

```
class LeitorPeewee(object):
    def ler(self, arquivo):
        registros = []
        arquivo = open(arquivo, "r")
        for linha in arquivo.readlines():
            tipo_registro = int(linha[:2])
            if tipo_registro == 1:
                registros.append(new_registro01(linha))
        return registros

    def escrever(registros):
        criar_tabela()
        atual = 0
        for registro in registros:
            atual += 1
            print("Salvando {}".format(atual))
            registro.save()

leitor = LeitorPeewee()
registros = leitor.ler("/home/eduardo/UEM/GIT/TCC/data/COTAHIST_A2017.TXT")
registros = filter(lambda r: r.codigo_negociacao.startswith('PETR4'), registros)

escrever(registros)
```

# Importação dos dados das ações

```
def new_registro01(linha):  
    reg01 = Registro01()  
    reg01.data_pregao = int(linha[2:10])  
    reg01.codigo_bdi = linha[10:12].strip()  
    reg01.codigo_negociacao = linha[12:24].strip()  
    reg01.tipo_mercado = int(linha[24:27])  
    reg01.nome_resumido = linha[27:39].strip()  
    reg01.especificacao = linha[39:49].strip()  
    reg01.prazo_termo = linha[49:52].strip()  
    reg01.moeda = linha[52:56].strip()  
    reg01.preco_abertura = to_float(linha[57:69])  
    reg01.preco_maximo = to_float(linha[69:82])  
    reg01.preco_minimo = to_float(linha[82:95])  
    reg01.preco_medio = to_float(linha[95:108])  
    reg01.preco_ultimo = to_float(linha[108:121])  
    reg01.preco_ofc = to_float(linha[121:134])  
    reg01.preco_ofv = to_float(linha[134:147])  
    reg01.total_negocios = int(linha[147:152])  
    reg01.quantidade_titulos = int(linha[152:170])  
    reg01.volume_titulos = to_float(linha[170:188])  
    reg01.preexe = to_float(linha[188:201])  
    reg01.indicador_cp = int(linha[201:202])  
    reg01.data_vencimento = int(linha[202:210])  
    reg01.fator_cotacao = int(linha[210:217])  
    reg01.ptoexe = to_float(linha[217:230])  
    reg01.codigo_isin = linha[230:242].strip()  
    reg01.dismes = int(linha[242:245])  
    return reg01
```

# Importação dos dados das ações

```
class Registro01(peewee.Model):
    data_pregao = peewee.DateField()
    codigo_bdi = peewee.CharField(max_length=2)
    codigo_negociacao = peewee.CharField(max_length=12)
    tipo_mercado = peewee.IntegerField()
    nome_resumido = peewee.CharField(max_length=12)
    especificacao = peewee.CharField(max_length=10)
    prazo_termo = peewee.CharField(max_length=3)
    moeda = peewee.CharField(max_length=4)
    preco_abertura = peewee.DecimalField(max_digits=11, decimal_places=2)
    preco_maximo = peewee.DecimalField(max_digits=11, decimal_places=2)
    preco_minimo = peewee.DecimalField(max_digits=11, decimal_places=2)
    preco_medio = peewee.DecimalField(max_digits=11, decimal_places=2)
    preco_ultimo = peewee.DecimalField(max_digits=11, decimal_places=2)
    preco_ofc = peewee.DecimalField(max_digits=11, decimal_places=2)
    preco_ofv = peewee.DecimalField(max_digits=11, decimal_places=2)
    total_negocios = peewee.IntegerField()
    quantidade_titulos = peewee.IntegerField()
    volume_titulos = peewee.DecimalField(max_digits=16, decimal_places=2)
    preexe = peewee.DecimalField(max_digits=11, decimal_places=2)
    indicador_cp = peewee.IntegerField()
    data_vencimento = peewee.DateField()
    fator_cotacao = peewee.IntegerField()
    ptoexe = peewee.DecimalField(max_digits=11, decimal_places=2)
    codigo_isin = peewee.CharField(max_length=12)
    dismes = peewee.IntegerField()
```

# Extração de notícias



Scrapy



python

# Extração de notícias

```
class NoticiasItem(scrapy.Item):
    titulo = scrapy.Field()
    conteudo = scrapy.Field()
    data_atualizacao = scrapy.Field()

class G1NoticiasSpider(Spider):
    name = 'g1_noticias'

    def start_requests(self):
        query = 'allintitle:petrobras site:g1.globo.com/economia'
        date_min = '9/1/2017'
        date_max = '9/30/2017'
        pages = 4
        for i in range(0, pages):
            start = i * 10
            url =
                'https://www.google.com.br/search?q={query}&tbs=cdr:1,cd_min:{date_min},cd_max:{
                date_max}&start={start}'.format(
                    query=query,
                    date_min=date_min,
                    date_max=date_max,
                    start=start
                )
            yield Request(url)
```



# Extração de notícias

```
def parse(self, response):
    news = response.xpath('//*[ @class="r"]')
    for new in news:
        g1_url = new.xpath('..//a/@href').extract_first()
        if '/url?' in g1_url:
            g1_url = self.format_google_url(g1_url)

        if 'blog' not in g1_url:
            yield Request(str(g1_url), callback=self.parse_g1)

def parse_g1(self, response):
    titulo = "".join([
        response.xpath('//*[ @class="content-head__title"]/text()').extract_first(),
        '\n',
        response.xpath('//*[ @class="content-head__subtitle"]/text()').extract_first()
    ])
    data_atualizacao = response.xpath('//*[ @class="content-publication-
data__updated"]/time/text()').extract_first()
    conteudo = response.xpath('//*[contains(concat(' ', normalize-space(@class), ' '),
'content-text')]/text()').extract()

    if not data_atualizacao:
        data_atualizacao = ''

    l = ItemLoader(item=NoticiasItem(), response=response)
    l.add_value('titulo', titulo)
    l.add_value('data_atualizacao', data_atualizacao)
    l.add_value('conteudo', conteudo)

    yield l.load_item()
```

# Extração de notícias

The screenshot displays the Scrapinghub web interface. At the top, the 'scrapinghub' logo is on the left, and a search bar and navigation menu (Scrapy Cloud, Portia, Crawlera, Datasets, Help) are in the center. On the right, a user profile for 'Eduardo Henrique Kenji Shibukawa' is shown. The left sidebar contains a navigation menu with sections: 'Noticias' (2 spiders, 0 members), 'JOBS' (Dashboard, Periodic Jobs), 'SPIDERS' (Dashboard, Code & Deploys, Settings), 'PROJECT' (Usage Stats, Activity, Members, Settings), and 'ADDONS' (Addons Setup). The main area is titled 'Job Dashboard' and shows sections for 'Container Group: Def...', 'Next Jobs' (0), and 'Running Jobs' (0). A modal dialog titled 'Run' is open in the center, showing the 'Current version: 04733a3-master'. It has input fields for 'Spiders' (containing 'g1\_noticias'), 'Job Units' (set to 1), and 'Priority' (set to Normal). There are also expandable sections for 'Tags' and 'Arguments'. At the bottom of the dialog are 'CANCEL' and 'RUN' buttons. The background interface is dimmed.

# Extração de notícias

The screenshot displays the Scrapyhub web interface. At the top, the 'scrapinghub' logo is on the left, followed by a search bar and navigation links: 'Scrapy Cloud', 'Portia', 'Crawlera', 'Datasets', and 'Help'. On the right, a user profile for 'Eduardo Henrique Kenji Shibukawa' is shown with a notification bell icon.

The main content area is titled 'Noticias' and 'Job 12'. It includes buttons for 'WATCH', 'CLONE THIS JOB', and 'STOP THIS JOB'. Below these are tabs for 'Job', 'Items 12', 'Requests 16', 'Log 13', 'Stats 17', and 'Console'. The 'Job Items' section is active, showing a list of items. The first item is 'Item 0' from '2017-12-05 16:58:28 UTC'. A dropdown menu is open over the 'Item 0' row, showing options: 'Get as', 'CSV', 'JSON', 'JSON Lines', and 'XML'. Other buttons like 'EXPORT', 'PUBLISH', 'SAMPLES', 'SHOW SCRAPED FIELDS', 'DOWNLOAD', and 'COMMENT' are also visible.

On the left sidebar, there are sections for 'JOBS', 'SPIDERS', 'PROJECT', and 'ADDONS', each with sub-links like 'Dashboard', 'Code & Deploys', 'Settings', 'Usage Stats', 'Activity', 'Members', and 'Addons Setup'.

The 'Job Items' table shows the following content for 'Item 0':

Item 0	2017-12-05 16:58:28 UTC
conteudo	<p>A Petrobras surpreendeu em mais ao apresentar ofertas muito superiores ao mínimo exigido. Nas 2ª e 3ª Rodadas de Partilha do petróleo, a estatal chegou a oferecer 80% de óleo excedente à União, quando o mínimo exigido era de 10,34%.</p> <p>O presidente da Petrobras, Pedro Parente, justificou que foram consideradas duas condições para a oferta: o valor máximo que a empresa poderia pagar juntamente com suas parceiras e, principalmente, o alto grau de competitividade pelas respectivas áreas.</p> <p>"Nessa situação, nós não podíamos nos dar ao luxo de perder essas oportunidades razão pelo qual fomos nos valores que estávamos autorizados a pagar porque não se conhece o valor que o competidor vai oferecer", argumentou o executivo.</p> <p>A Petrobras arrematou os blocos da Bacia de Santos Entorno de Sapinhoá e Peroba e o bloco Alto de Cabo Frio Central, na Bacia de Campos. Parente destacou que a companhia se uniu a "empresas de primeira grandeza" do setor petrolífero mundial para garantir a aquisição destes blocos.</p> <p>Para disputar Sapinhoá, que tinha 10,34% de excedente mínimo de óleo exigido, a brasileira se uniu em consórcio com a Shell, do Reino Unido, e a es</p>

# Rotulação dos dados



# Rotulação dos dados

Leitor de notícias

Carregar Arquivo

Gravar CSV

Gravar Dados

Notícias

03/09/2017 - Petrobras prorroga acordo coletivo de trabalho até novembroAcordo valerá até início da reforma trabalhista

03/09/2017 - Petrobras assina acordo com banco de desenvolvimento chinêsAcordo deve permitir à estatal buscar empréstimo de 10 bilhões de dólares

04/09/2017 - Petrobras faz novo reajuste e gasolina já sobe mais de 10% em setembroFuracão Harvey, que afetou pr...

05/09/2017 - Petrobras inicia processo de venda de fatia de 90% na Transportadora Associada de GásA Petrobras tem...

05/09/2017 - Aprovados em concurso da Petrobras serão chamados imediatamente, diz gerente de RHPrimeiro concurs...

06/09/2017 - Petrobras anuncia reajuste médio de 12,2% no gás de cozinha a partir desta quartaSe reajuste for integ...

12/09/2017 - Petrobras inicia processo para vender ativos de fertilizantesVendas são da subsidiária integral Araucária I...

12/09/2017 - Petrobras pagará R\$ 12,8 bilhões para cobrir rombo do Petros em 18 anosDéficit do fundo de pensão é d...

13/09/2017 - Funcionários da Petrobras terão que contribuir com até R\$ 3,6 mil mensais para cobrir rombo do PetrosVa...

14/09/2017 - Petrobras propõe reajuste salarial de 1,73% e corte na remuneração de hora extraData-base da categ...

18/09/2017 - Petrobras busca levantar US\$ 2 bilhões para substituir dívida cara Iniciativa é parte de plano da estatal d...

19/09/2017 - Eletrobras vence Petrobras na Justiça e libera operação de termelétrica no AmazonasPetrobras se recusa...

20/09/2017 - Governo vai atuar junto à Petrobras para acionar térmicas de baixo custo, diz ONSObjetivo é evitar o aci...

20/09/2017 - Petrobras pode ser credora do governo em US\$ 12 bilhões por cessão onerosa, diz ministroPresidente da...

21/09/2017 - Petrobras inicia mediação extrajudicial com Sete BrasilSete Brasil foi criada para ser fornecedora de sonda...

21/09/2017 - Bovespa fecha em queda com pressão de Vale e Petrobras Ibovespa caiu 0,53%, a 75.604 pontos, apes...

22/09/2017 - Refinaria de Pasadena retomará operações até o fim da semana, diz PetrobrasA refinaria havia sido fech...

25/09/2017 - Investidores aceitam trocar US\$ 6,2 bi em dívida da Petrobras por novos títulosDetentores de US\$ 6,2 bi...

25/09/2017 - Petrobras eleva em 6,9% preço do gás de cozinha, 2ª alta do mêsReajuste passa a valer já na terça-feir...

26/09/2017 - Petrobras vai participar de leilões com estratégia de recompor portfólio de exploração, diz presidente No...

27/09/2017 - Petrobras e Exxon pagam R\$ 3,4 bilhões por blocos na 'franja do pré-sal'Pedro Parente, presidente da Pe...

28/09/2017 - Conselho da Petrobras aprova registro de IPO da BR DistribuidoraOferta pública de ações ainda depende...

Data Atualização

03/09/2017

Sentimento

☐ Negativo

☒ Neutro

☐ Positivo

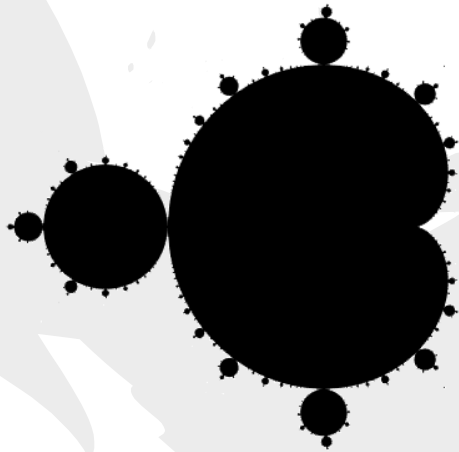
Título

Petrobras prorroga acordo coletivo de trabalho até novembro  
Acordo valerá até início da reforma trabalhista, que flexibiliza os horários de trabalho e algumas restrições, reduzindo os custos para as empresas.

Conteúdo

A Petrobras vai prorrogar o atual acordo coletivo de trabalho com os sindicatos até novembro, quando a...  
, informou a estatal neste domingo (3).  
"Para que a negociação ocorra com tranquilidade e em respeito aos empregados e às entidades sindicais, a Petrobras prorrogará o acordo coletivo de trabalho vigente até o dia 10 de novembro, véspera da data em que entra em vigor a nova legislação trabalhista", informou a empresa por email.  
A reforma trabalhista aprovada em junho flexibiliza os horários de trabalho e algumas restrições, reduzindo os custos trabalhistas para as empresas.  
A reforma enfrentou dura oposição de sindicatos pois reduz o poder dessas entidades e elimina a obrigatoriedade do imposto sindical, ao mesmo tempo que prevê a prevalência do acordo entre empregados e empregadores sobre a legislação.

# Análise de sentimentos das notícias



TextBlob



python

# Análise de sentimentos das notícias

```
def get_translated_blob(text):
    blob = TextBlob(text)
    return blob.translate(from_lang='pt-br', to="en")

def sentiment(dataset):
    sentimento_titulo = []
    sentimento_conteudo = []
    for index, row in dataset.iterrows():
        blob_titulo = get_translated_blob(row['titulo'])
        blob_conteudo = get_translated_blob(row['conteudo'])
        sentimento_titulo.append(blob_titulo.sentiment)
        sentimento_conteudo.append(blob_conteudo.sentiment)
    return sentimento_titulo, sentimento_conteudo

dataset = pd.concat([ get_dataset('noticias_outubro.csv')], ignore_index=True)

sentimento_titulo, sentimento_conteudo = sentiment(dataset)

output = pd.DataFrame(data={
    "data":dataset["data_atualizacao"],
    "dtitulo":dataset["titulo"],
    "sentimento": sentimento_conteudo
})

output.to_csv(
os.path.join(os.path.dirname(__file__), '../data/output', 'sentimentos_text_blob.csv')
)
```

# Análise de sentimentos das notícias

```
dataset_train = pd.concat([
    get_dataset_g1_rotulado('noticias_maio.csv'),
    get_dataset_g1_rotulado('noticias_junho.csv'),
    get_dataset_g1_rotulado('noticias_julho.csv'),
    get_dataset_g1_rotulado('noticias_agosto.csv'),
    get_dataset_g1_rotulado('noticias_setembro.csv'),
    ], ignore_index=True
)

dataset_test = pd.concat([
    get_dataset_g1_rotulado('noticias_outubro.csv'),
    ], ignore_index=True
)

print("Limando textos")
train = clean_text_field(dataset_train, 'conteudo', False)
test = clean_text_field(dataset_test, 'conteudo', False)

print("Criando bag of words")
from sklearn.model_selection import train_test_split
feature_train, feature_test, _ = create_bag_of_words(train, test)
X_train, X_test, y_train, y_test = train_test_split(feature_train, dataset_train['sentimento'],
test_size = 0.2, random_state=1000)
```



# Análise de sentimentos das notícias

```
print("Treinando classificador")
nayve_classifier = train_nayve(X_train, y_train)

print("Cross validation")
cross_validation(nayve_classifier, X_train, y_train)

print("Testando dataset de treino")
testar("Treino", nayve_classifier, X_test, y_test)

print("Testando dataset de teste real")
testar("Outubro", nayve_classifier, feature_test, dataset_test['sentimento'])

print("Gerando CSV")
generate_csv(nayve_classifier, dataset_test, feature_test)

print("Finalizado")
```

# Correlação de sentimento X ações



***SciPy***



python

# Correlação de sentimento X ações

```
def gerar_correlacao(dataset, arquivo_salvar):  
    valores_acao, sentimentos = gerar_valores_e_sentimentos(dataset)  
    import operator  
  
    max_range = 1  
    min_range = -1  
  
    sentimentos_normalizado = normalizar(sentimentos, (min_range, max_range))  
    acoes_normalizado = normalizar(valores_acao, (min_range, max_range))  
  
    x = []  
    y = []  
    x_normalizado = []  
    y_normalizada = []  
  
    for data in sorted(sentimentos.keys()):  
        x.append(sentimentos[data])  
        x_normalizado.append(sentimentos_normalizado[data])  
        y.append(valores_acao[data])  
        y_normalizada.append(acoes_normalizado[data])
```

# Correlação de sentimento X ações

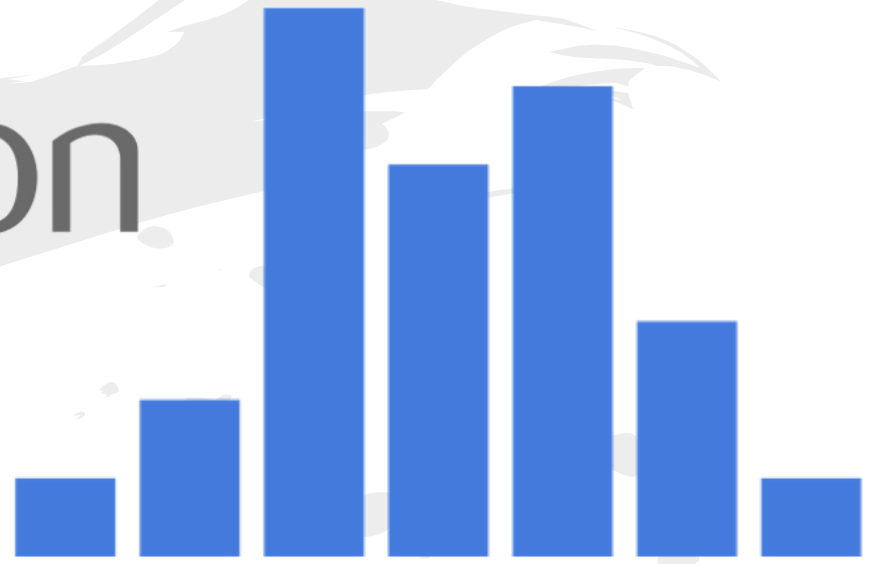
```
correlacao = get_correlacao(  
    x,  
    y  
)  
correlacao_normalizado = get_correlacao(  
    x_normalizado,  
    y_normalizada  
)  
  
return correlacao, correlacao_normalizado
```

```
def get_correlacao(x, y):  
    return pearsonr(y, x)
```

# Plotagem dos dados



python



plotly

# Sumário

Agradecimentos

Introdução

Objetivos

Fundamentação Teórica

Implementação

**Resultados**

Considerações finais

Referências

# Dados das ações

BOCOTAHIST.2017BOVESPA 20171004

012017010202AALR3	010ALLIAR	ON	NM	R\$	0000000001462000000000148800000000014400000000001458000000000146000000000
012017010210ABCB2	010ABC BRASIL	DIR	PRE N2	R\$	0000000000285000000000002850000000000285000000000285000000000285000000000
012017010202ABCB4	010ABC BRASIL	PN	EJS N2	R\$	00000000001340000000000013520000000000130800000000013300000000001331000000000
012017010296ABCB4F	020ABC BRASIL	PN	EJS N2	R\$	00000000001345000000000013480000000000131500000000013290000000001326000000000
012017010212ABCP11	010FII ABC IMOBCI	ER		R\$	00000000001183000000000011830000000000110100000000011620000000001136000000000
012017010202ABEV3	010AMBEV S/A	ON	EJ	R\$	00000000001634000000000016660000000000162600000000016420000000001631000000000
012017010296ABEV3F	020AMBEV S/A	ON	EJ	R\$	0000000000164100000000001666000000000016270000000001645000000000166000000000
012017010262ABEV3T	030AMBEV S/A	ON	EJ	016R\$	0000000000165500000000001656000000000016490000000001653000000000165000000000
012017010262ABEV3T	030AMBEV S/A	ON	EJ	045R\$	00000000001656000000000016570000000000165600000000016560000000001657000000000
012017010262ABEV3T	030AMBEV S/A	ON	EJ	060R\$	00000000001691000000000016940000000000169100000000016920000000001694000000000

Data pregão	Código BDI	Código Negociação	Nome Resumido	Moeda	Preço Ultimo
20170102	02	PETR4	PETROBRAS	R\$	14.66
20170102	96	PETR4F	PETROBRAS	R\$	14.65
20170102	62	PETR4T	PETROBRAS	R\$	14.75
20170102	62	PETR4T	PETROBRAS	R\$	14.82
20170102	62	PETR4T	PETROBRAS	R\$	14.95
20170102	62	PETR4T	PETROBRAS	R\$	15.11
20170103	02	PETR4	PETROBRAS	R\$	15.5
20170103	96	PETR4F	PETROBRAS	R\$	15.44
20170103	62	PETR4T	PETROBRAS	R\$	15.69
20170103	62	PETR4T	PETROBRAS	R\$	15.58

# Extração dos dados

"_type"	"conteudo"	"data_atualizacao"	"titulo"
"NoticiasItem"	", O ministro de Minas e Energia, ..."	"20/09/2017 18h15 "	"Petrobras pode ser credora ..."
"NoticiasItem"	", Detentores de US\$ 6,22 bilhões em ..."	"25/09/2017 11h03 "	"Investidores aceitam trocar ..."
"NoticiasItem"	", Os lances bilionários oferecidos ..."	"27/09/2017 15h26 "	"Petrobras e Exxon pagam R\$ ..."
"NoticiasItem"	", A Petrobras deverá gastar R\$ 12,8 ..."	"12/09/2017 20h31 "	"Petrobras pagará R\$ 12,8 ..."
"NoticiasItem"	", A subsidiária de geração da estatal ..."	"19/09/2017 13h48 "	"Eletrobras vence Petrobras ..."
"NoticiasItem"	", A Petrobras informou nesta terça ..."	"05/09/2017 20h43 "	"Petrobras inicia processo de ..."
"NoticiasItem"	", A Petrobras iniciou processo para a ..."	"12/09/2017 06h59 "	"Petrobras inicia processo para ..."
"NoticiasItem"	", A Petrobras informou nesta quinta ..."	"21/09/2017 19h12 "	"Petrobras inicia mediação ..."
"NoticiasItem"	", Os concursos ,, com possibilidade de , ..."	"05/09/2017 06h00 "	"Aprovados em concurso ..."
"NoticiasItem"	", A Petrobras planeja levantar US\$ 2 ..."	"18/09/2017 13h06 "	"Petrobras busca levantar ..."



# Extração dos dados

```
1 {
2   {
3     "conteudo": [
4       " ",
5       "O ministro de Minas e Energia, Fernando Coelho Filho, declarou nesta quarta-feira (20) que a Petrobras poderia ser credora de mais de US$ 12 bilhões com o governo federal ao final do processo do contrato de cessão onerosa. O presidente da empresa, Pedro Parente, disse que a petroleira quer concluir a renegociação o mais rápido possível. ",
6       " ",
7       " ",
8       "A cessão onerosa refere-se ao contrato da União e Petrobras para a exploração de 5 bilhões de barris de óleo equivalente sem licitação na época da capitalização da companhia, em 2010. Contudo, naquele momento, a petroleira pagou à União o equivalente a US$ 42,5 bilhões. ",
9       " ",
10      "O contrato previa também uma renegociação anos depois para atualização de variáveis como câmbio e preços do petróleo, o que poderá resultar em pagamentos para a estatal ao fim do processo. ",
11      " ",
12      "O governo federal avalia como razoável uma estimativa de US$ 12 bilhões, feita pelo banco UBS, sobre o total que a Petrobras teria que receber ao final do processo de renegociação com a União, disse o ministro, em entrevista à Bloomberg, nesta quarta-feira. ",
13      " ",
14      "\"Eu acho que vai ser mais do que isso, mas quanto mais eu não sei. Não gosto de colocar números\", disse o ministro, referindo-se a uma estimativa do UBS. ",
15      " ",
16      "Esse valor estimado pelo UBS, no entanto, está distante do que a Petrobras quer receber, segundo o ministro. ",
17      " ",
18      "Em Nova York, o presidente da Petrobras afirmou que a companhia está aguardando \"o governo indicar os seus negociadores\" para poder \"começar o processo formal de negociação\". ",
19      " ",
20      "\"De nossa parte, estamos absolutamente preparados para começar e terminar essa negociação... queremos terminar o mais cedo possível. Quando será isso, não depende só de nós\", disse o presidente da Petrobras. ",
21      " ",
22      "Paralelamente, Parente disse que o oferta pública inicial de ações (IPO, na sigla em inglês) da BR Distribuidora está avançando. ",
23      " ",
24      " ",
25      " ",
26      " ",
27      " ",
28      " ",
29      " ",
30      " ",
31    ],
32    "type": "NoticiasItem",
33    "data_atualizacao": [
34      "20/09/2017 18h15 "
35    ],
36    "titulo": [
37      "Petrobras pode ser credora do governo em US$ 12 bilhões por cessão onerosa, diz ministro\nPresidente da estatal diz que quer concluir a renegociação com o governo o mais rápido possível."
38    ]
39  },
40 }
```

# Rotulação dos dados

data_atualizacao	titulo	conteudo	sentimento
02/05/2017	Justiça libera venda de área ...	A Petrobras informou nesta terça-feira (2) que foi suspens ...	1
02/05/2017	Presidente da Petrobras diz q...	O fraco desempenho dos preços do petróleo neste ano não de ...	1
03/05/2017	Decreto regulamenta direito d...	Um decreto publicado no Diário Oficial da União nesta quar ...	-1
08/05/2017	Representante da Petrobras na...	O representante local da Petrobras da Bolívia está em pris ...	-1
08/05/2017	Justiça extingue ação contra ...	A Justiça Federal de Sergipe atendeu a pedido da Petrobras ...	1
09/05/2017	Petrobras diz que ainda não e...	A Petrobras informou nesta terça-feira (9) em comunicado q ...	0
10/05/2017	Bovespa tem forte alta puxada...	O principal índice da bolsa paulista opera em alta nesta q ...	1
10/05/2017	Petrobras coloca refinaria de...	A Petrobras incluiu na sua carteira de ativos à venda a re ...	-1
11/05/2017	Petrobras tem lucro de R\$ 445...	A Petrobras registrou lucro de R\$ 445 bilhões no 1º trim i...	1
11/05/2017	Bovespa fecha em alta à esper...	O principal índice da bolsa paulista fechou em alta nesta ...	1

# Análise de sentimento de notícias

data	titulo	sentimento	polaridade
02/10/2017	Petrobras anuncia US\$ 63 bi em operações de pagamento renegociação e contratação ...	18.80%	46.61%
02/10/2017	Ministro de Minas e Energia prevê que Petrobras será privatizada: É um caminho F ...	-4.17%	46.46%
03/10/2017	Não estamos tratando disso diz ministro sobre privatização da Petrobras Fernando ...	-1.81%	31.11%
04/10/2017	Comitê do setor elétrico pedirá que Petrobras forneça combustível para térmicas ...	6.50%	40.25%
04/10/2017	Petrobras está perto de atingir meta de desalavancagem A companhia que tem reduz ...	0.31%	29.06%
05/10/2017	Indústria naval deve recorrer de decisão da ANP que libera Petrobras para constr ...	5.37%	36.70%
05/10/2017	Petrobras vai analisar pedido do governo por ajuda a térmicas mas sem subsídio O ...	6.31%	32.82%
05/10/2017	Petrobras vê como positiva flexibilização de conteúdo local em reserva do pré-sa ...	3.21%	19.23%
10/10/2017	Petrobras eleva preço do botijão de gás em 129% a partir desta quarta Estatal es ...	-1.20%	45.94%
11/10/2017	Cade vê como complexa compra de ativos da Petrobras por mexicana e pede estudos ...	0.00%	22.50%
11/10/2017	TCU bloqueia bens de Dilma por prejuízo à Petrobras com compra de Pasadena Além ...	3.36%	31.16%
16/10/2017	Petrobras pede registro de companhia aberta para BR Distribuidora Pedido de IPO ...	-8.57%	34.29%
17/10/2017	Moodys eleva nota da Petrobras e muda perspectiva para estável Agência cita redu ...	11.77%	41.30%
18/10/2017	Produção da Petrobras no Brasil sobe 28% em setembro ante agosto Na comparação c ...	-2.45%	37.00%
19/10/2017	Governo exclui fatia da Petrobras na Braskem de programa de desestatização Medid ...	0.00%	24.00%
19/10/2017	Carf decide a favor da Petrobras em processo de R\$ 78 bilhões diz estatal Estata ...	1.11%	2.22%
24/10/2017	Petrobras quer vender a BR Distribuidora 'o mais rápido possível' Empresa tem me ...	0.81%	46.20%
24/10/2017	Petrobras fará parceria para disputar blocos do pré-sal em leilão na sexta-feira ...	12.25%	55.80%
26/10/2017	Petrobras aprova adesão a Nível 2 de governança da bolsa e corte de gerências Pa ...	12.53%	45.32%
26/10/2017	Petrobras pode incluir ativos de logística em vendas de refinarias; detalhes dev ...	17.10%	54.89%
26/10/2017	Petrobras será 'seletiva' mas 'muito firme' no leilão do pré-sal diz Pedro Paren ...	18.77%	52.49%
26/10/2017	Petrobras adere ao programa de regularização de dívida não tributária Segundo a ...	2.90%	39.61%
27/10/2017	Petrobras 'não podia se dar ao luxo de perder essa oportunidade' diz Parente sob ...	12.41%	38.96%
27/10/2017	Com lances altos Petrobras leva 3 blocos do pré-sal e oferece até 80% da produçã ...	11.45%	48.39%
31/10/2017	Petrobras busca parcerias para concluir obras do Comperj diz Parente Presidente ...	7.20%	37.36%
31/10/2017	Parceria da Petrobras com empresa britânica BP deve envolver troca de ativos diz ...	10.94%	43.05%

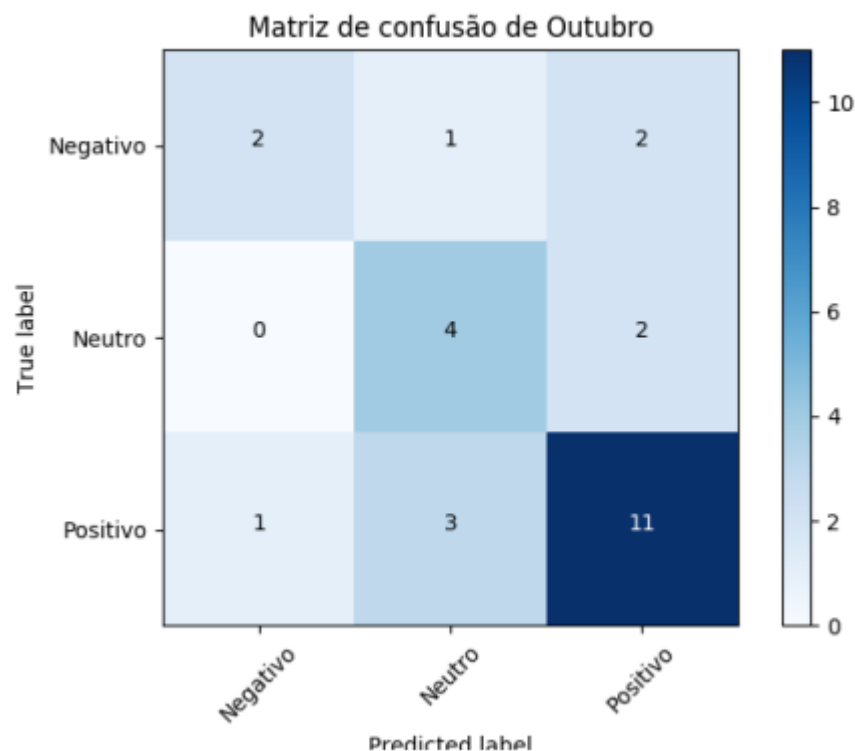
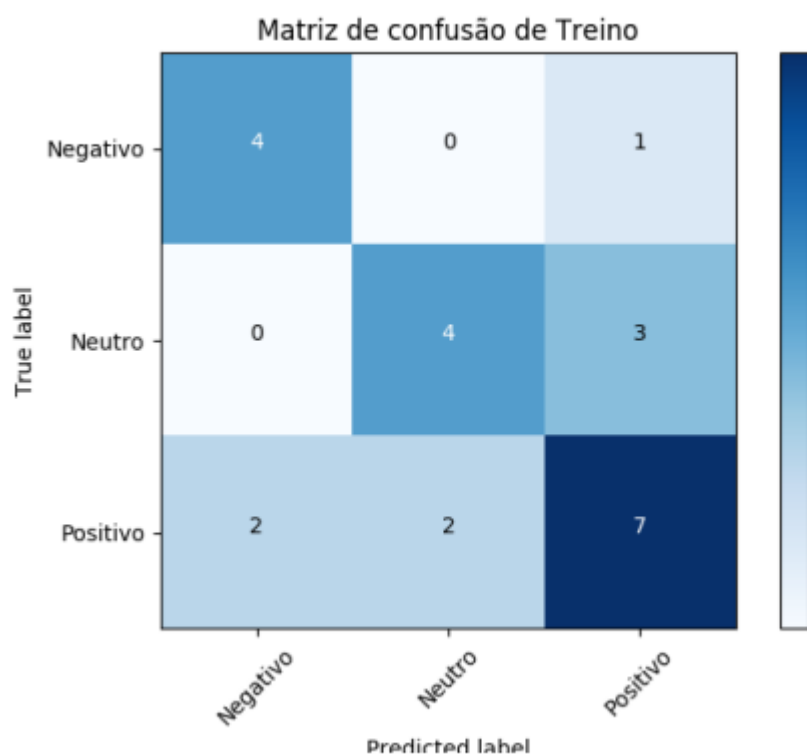
Tabela 4: Sentimento lexicon

# Análise de sentimento de notícias

data	titulo	sentimento
02/10/2017	Petrobras anuncia US\$ 63 bi em operações de pagamento renegociação e contratação de dívida Empresa pagou antec ...	1
02/10/2017	Ministro de Minas e Energia prevê que Petrobras será privatizada: É um caminho Fernando Coelho Filho afirmou n ...	1
03/10/2017	Não estamos tratando disso diz ministro sobre privatização da Petrobras Fernando Coelho Filho (Minas e Energia ...	1
04/10/2017	Comitê do setor elétrico pedirá que Petrobras forneça combustível para térmicas paradas Comitê destacou que a ...	1
04/10/2017	Petrobras está perto de atingir meta de desalavancagem A companhia que tem reduzido trimestre a trimestre sua ...	1
05/10/2017	Indústria naval deve recorrer de decisão da ANP que libera Petrobras para construir casco de Libra no exterior ...	0
05/10/2017	Petrobras vai analisar pedido do governo por ajuda a térmicas mas sem subsídio O Comitê de Monitoramento do Se ...	1
05/10/2017	Petrobras vê como positiva flexibilização de conteúdo local em reserva do pré-sal Na véspera ANP flexibilizou ...	0
10/10/2017	Petrobras eleva preço do botijão de gás em 129% a partir desta quarta Estatal estima que preço ao consumidor f ...	0
11/10/2017	Cade vê como complexa compra de ativos da Petrobras por mexicana e pede estudos Venda da Petroquímica Suape e ...	1
11/10/2017	TCU bloqueia bens de Dilma por prejuízo à Petrobras com compra de Pasadena Além da ex-presidente decisão ating ...	-1
16/10/2017	Petrobras pede registro de companhia aberta para BR Distribuidora Pedido de IPO inclui também aval para realiz ...	1
17/10/2017	Moodys eleva nota da Petrobras e muda perspectiva para estável Agência cita redução do endividamento da petrol ...	1
18/10/2017	Produção da Petrobras no Brasil sobe 28% em setembro ante agosto Na comparação com setembro de 2016 houve qued ...	-1
19/10/2017	Governo exclui fatia da Petrobras na Braskem de programa de desestatização Medida também exclui a fatia da pet ...	-1
19/10/2017	Carf decide a favor da Petrobras em processo de R\$ 78 bilhões diz estatal Estatal divulgou comunicado nesta qu ...	1
24/10/2017	Petrobras quer vender a BR Distribuidora 'o mais rápido possível' Empresa tem meta de vender ativos que somam ...	1
24/10/2017	Petrobras fará parceria para disputar blocos do pré-sal em leilão na sexta-feira Presidente da estatal adianto ...	0
26/10/2017	Petrobras aprova adesão a Nível 2 de governança da bolsa e corte de gerências Patamar intermediário de governa ...	1
26/10/2017	Petrobras pode incluir ativos de logística em vendas de refinarias; detalhes devem sair no 1º tri diz Parente ...	0
26/10/2017	Petrobras será 'seletiva' mas 'muito firme' no leilão do pré-sal diz Pedro Parente ANP pôs à venda 8 blocos de ...	0
26/10/2017	Petrobras adere ao programa de regularização de dívida não tributária Segundo a estatal adesão resultará em im ...	1
27/10/2017	Petrobras 'não podia se dar ao luxo de perder essa oportunidade' diz Parente sobre alta oferta nos leilões Est ...	0
27/10/2017	Com lances altos Petrobras leva 3 blocos do pré-sal e oferece até 80% da produção à União No regime de partilh ...	0
31/10/2017	Petrobras busca parcerias para concluir obras do Comperj diz Parente Presidente da estatal revelou a expectati ...	1
31/10/2017	Parceria da Petrobras com empresa britânica BP deve envolver troca de ativos diz Parente Estatal divulgou cart ...	1

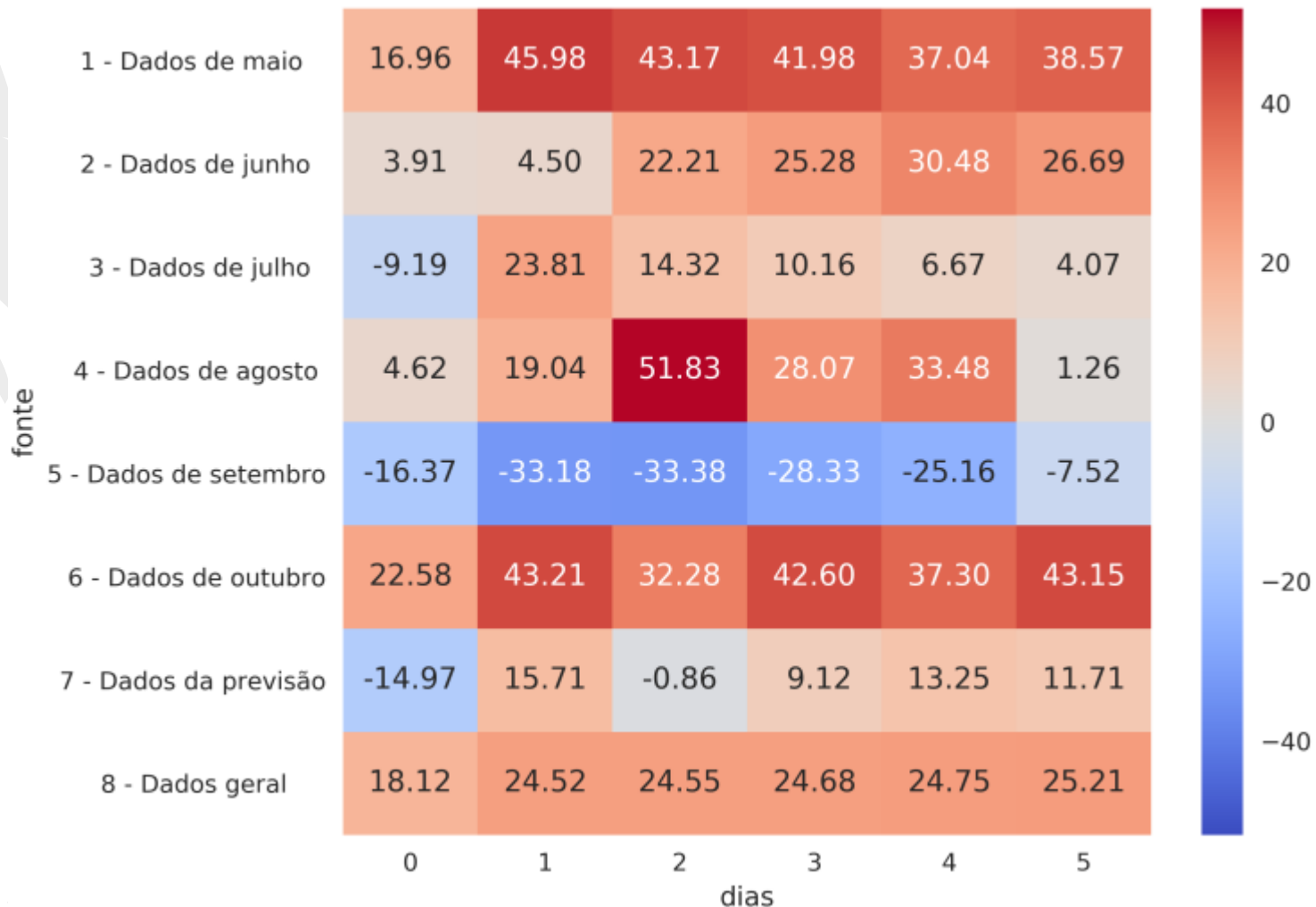
Tabela 5: Sentimento scikit

# Análise de sentimento de notícias



Fonte dos dados	Kappa	Precisão	Recall	FScore
Treino	45.24%	65.66%	66.93%	65.97%
Outubro	39.69%	63.33%	60.00%	60.16%

# Correlação de sentimento X ações





# Sumário

Agradecimentos

Introdução

Objetivos

Fundamentação Teórica

Implementação

Resultados

**Considerações finais**

Referências

# Considerações finais

- Conseguimos resultados positivos com a análise de sentimento das notícias extraídas, porém ainda não possuem um nível de confiança o suficiente para utilização em conjunto com a correlação dos dados.
- Existe uma pequena correlação entre os sentimentos das notícias e os valores das ações.



# Dificuldades

- Dificuldade em encontrar dados referentes a sentimentos de notícias rotulados.
- Problemas com a extração dados causadas pelas regras configuradas no arquivo *robot.txt*.
- Rotulação dos dados.

# Trabalhos Futuros

- Utilização de filtros mais abrangentes, como a utilização de mais empresas como referência, assim como um período maior.
- Extração de dados de mais fontes como a Exame, InfoMoney, entre outras fontes.
- Verificação da correlação de dados dos dados com as ações utilizando outras fontes de dados além do sentimento, como dados financeiros da empresa entre outros.

# Trabalhos Futuros

- Utilização de uma base de dados diferentes para a correlação, um exemplo seria utilizar os comentários das notícias.
- Utilização de pesos para os sentimentos das notícias mais importantes.
- Rotulação dos dados por mais de 1 pessoa.



**MUITO OBRIGADO!**

# Sumário

Agradecimentos

Introdução

Objetivos

Fundamentação Teórica

Proposta de protótipo

Implementação

Resultados

Considerações finais

**Referências**

# Referências

ALVES, D. S. Uso de técnicas de computação social para tomada de decisão de compra e venda de ações no mercado brasileiro de bolsa de valores. Universidade de Brasília,(2015 <http://repositorio.unb.br/handle/10482/19345>), 2015.

AULERIO. Dicionario Aurelio. 2017. Disponível em: <<https://dicionariodoaurelio.com/>>.

BARBER, D. Bayesian Reasoning and Machine Learning. [S.l.]: Cambridge University Press, 2016.

BERTOLO, L. Mercado Financeiro. [s.n.], 2002. Disponível em: <<http://lbertolo.tripod.com/MATEMATICAFINANCEIRA.pdf>>.

FISCHER, A. et al. Mercado de Valores Mobiliários Brasileiro. Macgraw-Hill, 2014. Disponível em: <[www.cvm.gov.br/menu/investidor/publicacoes/livros.html](http://www.cvm.gov.br/menu/investidor/publicacoes/livros.html)>.

FONSECA, J. W. F. da. Análise e Decisão de Investimentos. [S.l.]: IESDE Brasil, 2009. HU, M.; LIU, B. Mining and summarizing customer reviews. In: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2014. Disponível em: <<http://doi.acm.org/10.1145/1014052.1014073>>.

# Referências

JURAFSKY, D.; MARTIN, J. H. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. MIT Press, 2009. Disponível em: <<http://www.mitpressjournals.org/doi/pdf/10.1162/089120100750105975>>.

LIU, B. Sentiment analysis and opinion mining. Morgan & Claypool Publishers, 2012. Disponível em: <<https://www.cs.uic.edu/~liub/FBS/SentimentAnalysis-and-OpinionMining.pdf>>.

MITTAL, A.; GOEL, A. Stock prediction using twitter sentiment analysis. Stanford University, CS229 (2011 <http://cs229.stanford.edu/proj2011/GoelMittalStockMarketPredictionUsingTwitterSentimentAnalysis.pdf>), v. 15, 2012.

MONARD, M. C.; BARANAUSKAS, J. A. Sistemas inteligentes: fundamentos e aplicações. In: . Manole Ltda, 2003. cap. 4. Disponível em: <<http://dcm.ffclrp.usp.br/~augusto/publications/2003-sistemas-inteligentes-cap4.pdf>>.

MURPHY, J. J. Technical Analysis Of The Financial Markets. [S.l.: s.n.], 1999. 1 – 239 p. ISSN 00314439. ISBN 0735200661 9780735200661.

# Referências

MURPHY, J. J. Technical Analysis Of The Financial Markets. [S.l.: s.n.], 1999. 1 – 239 p. ISSN 00314439. ISBN 0735200661 9780735200661.

NUIJ, W. et al. An automated framework for incorporating news into stock trading strategies.

IEEE Transactions on Knowledge and Data Engineering, v. 26, n. 4, p. 823–835, 2014. ISSN 10414347.

POOLE, D. L.; MACKWORTH, A. K. Artificial Intelligence: foundations of computational agents. [S.l.]: Cambridge University Press, 2010.

RUSSELL, S.; NORVIG, P. Inteligência Artificial. 3. ed. [S.l.]: Elsevier, 2013.

SCIPY. Scipy | scipy.stats.pearsonr. 2017. Disponível em: <<https://docs.scipy.org/doc/scipy-0.19.1/reference/generated/scipy.stats.pearsonr.html>>.



# Referências

SCRAPY. Scrapy | A Fast and Powerful Scraping and Web Crawling Framework. 2017. Disponível em: <<https://scrapy.org/>>.

SELAN, B. Mercado Financeiro. Universidade Estácio de Sá, 2015. Disponível em: <<https://profhubert.yolasite.com/resources/LIVRO\%20PROPRIETARIO\%20-\%20Mercado\%20financeiro.pdf>>.

TSYTSARAU, M.; PALPANAS, T. Survey on mining subjective data on the web. Data Mining and Knowledge Discovery. MIT Press, 2012. Disponível em: <<http://www.mitpressjournals.org/doi/pdf/10.1162/089120100750105975>>.