
Análise de Desempenho e Valuation das Empresas do S&P 500 com IA Generativa

1. Título

Análise de Desempenho e Valuation das Empresas do S&P 500 com IA Generativa

2. Integrantes

- Daniel Aloia Plastina – RA: 10375596
 - Gustavo Henrique Fernandes Vieira – RA: 10400715
 - João Victor Silva de Luca Gonçalves – RA: 10401977
-

3. Resumo

Este projeto tem como objetivo aplicar técnicas de Inteligência Artificial, utilizando a API ChatGPT, para a análise de desempenho financeiro e valuation das empresas listadas no índice S&P 500. Com um dataset que reúne informações fundamentais – como TICKER, Name, Sector, Price, Price/Earnings (P/E), Dividend Yield, Earnings/Share, 52 Week Low/High, Market Cap, EBITDA, Price/Sales, Price/Book e SEC Filings – a proposta é automatizar a geração de resumos, comparações e insights preditivos. O resultado esperado é um sistema de apoio à decisão que auxilie investidores e analistas na interpretação dos dados financeiros e na identificação de oportunidades de mercado.

4. Introdução

a. Contextualização

No cenário atual dos mercados financeiros, a análise de grandes volumes de dados é essencial para a identificação de oportunidades de investimento e o monitoramento do desempenho das maiores empresas dos EUA. O índice S&P 500 reúne as principais empresas, distribuídas em diversos setores, e oferece uma visão abrangente do mercado.

Com a crescente complexidade dos dados financeiros, torna-se indispensável utilizar ferramentas automatizadas para sintetizar informações e gerar insights estratégicos.

b. Justificativa

A análise manual de métricas financeiras é demorada e sujeita a erros, o que pode comprometer a tomada de decisão. A aplicação da API ChatGPT permite a automação da extração de insights e a elaboração de relatórios dinâmicos, otimizando a interpretação dos indicadores e facilitando comparações entre empresas e setores.

c. Objetivo

Desenvolver uma solução de análise de dados de negócios que:

- Integre e normalize dados financeiros das empresas do S&P 500;
- Gere resumos automatizados e insights relevantes (como comparações de P/E, EBITDA, dividend yield, etc.);
- Ofereça uma interface interativa para visualização e filtragem dos dados;
- Auxilie investidores e analistas a identificar oportunidades e tendências no mercado financeiro.

d. Opção do Projeto

Opção de Negócios – Utilizar a API ChatGPT para implementar uma solução que integre análise financeira e geração de insights automatizados, transformando dados brutos em relatórios de fácil interpretação.

5. Descrição do Problema

Investidores e analistas enfrentam o desafio de interpretar múltiplas métricas financeiras provenientes de diversas fontes, o que exige um processo intensivo de coleta, normalização e interpretação dos dados. O grande volume e a diversidade dos indicadores – como P/E, dividend yield, EBITDA, entre outros – dificultam a comparação direta entre empresas de setores distintos. Assim, a principal dificuldade reside em:

- Integrar dados heterogêneos e normalizá-los;
- Comparar indicadores de empresas de diferentes setores;
- Gerar relatórios que sintetizem tendências e identifiquem oportunidades de mercado de forma ágil e precisa.

6. Aspectos Éticos e Responsabilidade

- **Transparência:** Os métodos e processos aplicados na análise serão documentados para que os resultados possam ser auditados.
 - **Confiabilidade:** A solução se baseará em dados públicos e reconhecidos, garantindo a veracidade das informações.
 - **Responsabilidade:** Os insights gerados pela API ChatGPT serão apresentados como apoio à decisão, sem substituir a análise crítica dos profissionais.
 - **Privacidade:** Embora os dados sejam públicos, os processos de tratamento dos dados serão realizados respeitando as normas de proteção de dados e ética na pesquisa.
-

7. Dataset e Análise Exploratória e Preparação dos Dados (Exemplo em Python)

Descrição do Dataset

Nome: S&P 500 Companies with Financial Information

Conteúdo:

- **TICKER:** Código identificador da ação.
- **Name / NAME:** Nome completo da empresa.
- **Sector / SECTOR:** Setor de atuação da empresa.
- **Price / PRICE:** Preço atual da ação.
- **Price/Earnings (P/E):** Relação entre o preço da ação e os lucros por ação.
- **Dividend Yield (DEVIDEND Y):** Percentual de dividendos pagos em relação ao preço da ação.
- **Earnings/Share:** Lucro por ação (EPS).
- **52 Week Low / 52 Week High:** Menor e maior preço negociado nos últimos 52 semanas.
- **Market Cap:** Capitalização de mercado.
- **EBITDA:** Lucro operacional antes de juros, impostos, depreciação e amortização.
- **Price/Sales:** Relação entre o preço da ação e a receita por ação.
- **Price/Book:** Relação entre o preço da ação e o valor contábil por ação.
- **SEC Filings:** Registros e informações de conformidade submetidos à SEC.

Origem

Os dados são obtidos de fontes públicas, como o EDGAR, SEC e bases financeiras dos EUA, e representam informações atualizadas das empresas listadas no S&P 500.

Análise Exploratória e Preparação dos Dados – Exemplo em Python

```
# Cabeçalho do arquivo: analise_exploratoria.ipynb
# Histórico de Alterações:
```

```
# 17/03/2025 - Daniel Aloia Plastina - Criação do esqueleto inicial do notebook: importação de bibliotecas e carregamento do dataset.
# 21/03/2025 - Gustavo Henrique Fernandes Vieira - Inclusão de análises descritivas e verificação de dados faltantes.
# 25/03/2025 - João Victor Silva de Luca Gonçalves - Implementação de visualizações: boxplots, scatter plots e análise de outliers.
# 28/03/2025 - Daniel Aloia Plastina - Ajustes finais: normalização dos dados e comentários explicativos.
```

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Carregar o dataset
df = pd.read_csv('constituents-financials_csv.csv')

# Exibir as primeiras linhas
print("Primeiras linhas do dataset:")
print(df.head())

# Estatísticas descritivas
print("\nEstatísticas Descritivas:")
print(df.describe())

# Verificação de dados faltantes
print("\nDados faltantes por coluna:")
print(df.isnull().sum())

# Visualização 1: Distribuição do P/E por Setor
plt.figure(figsize=(12,6))
sns.boxplot(x='SECTOR', y='P/E', data=df)
plt.title('Distribuição do P/E por Setor')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()

# Visualização 2: Relação entre Market Cap e EBITDA (colorido por Setor)
plt.figure(figsize=(12,6))
sns.scatterplot(x='Market Cap', y='EBITDA', hue='SECTOR', data=df)
plt.title('Relação entre Market Cap e EBITDA')
plt.tight_layout()
plt.show()

# Normalização da coluna P/E para comparações
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
df['P/E_norm'] = scaler.fit_transform(df[['P/E']])
print("\nExemplo de normalização do P/E:")
print(df[['P/E', 'P/E_norm']].head())
```

8. Metodologia e Resultados Esperados

Metodologia

1. Análise Exploratória e Preparação dos Dados:

- Utilizar Python (Pandas, NumPy) para explorar estatísticas descritivas e identificar outliers.
 - Gerar visualizações comparativas por setor e entre diferentes métricas utilizando Matplotlib e Seaborn.
 - Normalizar os dados para permitir comparações justas entre empresas.
2. **Integração com a API ChatGPT:**
- Desenvolver scripts que enviem partes do dataset para a API ChatGPT, solicitando resumos e insights, por exemplo:
Prompt: "Resuma os pontos fortes e os riscos da empresa [Name] considerando seu P/E, EBITDA e dividend yield."
 - Automatizar a geração de relatórios que combinem os insights da análise exploratória com os resumos gerados pela API.
3. **Desenvolvimento do Dashboard Interativo:**
- Criar uma interface usando Dash ou Streamlit que permita ao usuário filtrar os dados por setor, comparar métricas e visualizar os relatórios automatizados.
 - Integrar gráficos dinâmicos e resumos gerados para facilitar a interpretação dos dados.

Resultados Esperados

- **Relatórios Automatizados:** Resumos gerados pela API ChatGPT que destaquem os pontos fortes, fraquezas e tendências de cada empresa ou setor.
- **Insights Comparativos:** Identificação de oportunidades de investimento, como empresas com P/E abaixo da média do setor ou alta volatilidade identificada pelos dados de 52 Week Low/High.
- **Dashboard Interativo:** Uma interface visual que permita ao usuário explorar os dados, aplicar filtros e visualizar gráficos e relatórios automatizados.
- **Validação:** Testes de usabilidade e precisão dos relatórios, com feedback de investidores e analistas financeiros.

9. Referências

1. OpenAI. How can I access the ChatGPT API? Agosto. 2024. Disponível em: <https://help.openai.com/en/articles/7039783-how-can-i-access-the-chatgpt-api>. Acesso em: 18 de março de 2025.
 2. CAMELO, L. API do ChatGPT: como usar todo o poder da inteligência artificial? Agosto. 2023. Disponível em: <https://pluga.co/blog/api-chatgpt/>. Acesso em: 18 de março de 2025.
 3. Dados financeiros públicos: EDGAR, SEC, e bases financeiras dos EUA.
 4. Devlin, J., et al. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.
 5. Vaswani, A., et al. (2017). Attention is All You Need.
-

10. Bibliografia

- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.
 - Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). Attention is All You Need.
 - OpenAI. (2024). How can I access the ChatGPT API?
 - Camelo, L. (2023). API do ChatGPT: como usar todo o poder da inteligência artificial?
 - Publicações e relatórios sobre análise financeira e valuation disponíveis em bases públicas (ex.: EDGAR, SEC).
-