

Trabalho de Visualização

Descrição do trabalho até o momento:

Nesse pequeno resumo irei descrever as etapas do trabalho de uma forma mais objetiva, deixando o conteúdo em si para o artigo final. Farei uma breve descrição de cada etapa e indicarei onde me encontro e o que pretendo fazer nas fases que virão.

Descrição das etapas:

1. Escolha dos dados e objetivos iniciais:

Os dados escolhidos para esse projeto foram uma tabela com os índices da bolsa Dow Jones e uma tabela com as manchetes das principais notícias do mundo ranqueada pelos usuários do Reddit no período de 08/2008 à 08/2016.

Esses dados foram utilizados no trabalho realizado na disciplina de KDD onde buscou-se descobrir se haviam alguma relação entre os tópicos das notícias e as variações na bolsa. Esse trabalho não permitiu encontrar tais relações devido às diversas dificuldades em classificar as notícias por tópico e descobrir se o impacto era positivo ou negativo.

O objetivo desse trabalho de visualização é buscar outras formas de representar os dados da bolsa de valores. Gráficos fornecidos por outras ferramentas podem auxiliar:

- Identificar tendências e comportamentos mais restritos da bolsa, como períodos de baixa ou alta contínuos;
- Identificar comportamentos semanais, mensais ou anuais;
- Construir uma interface que associe as principais notícias do dia a um intervalo ou ponto dia específico;
- Utilizar filtros para selecionar dias, semanas ou meses com diferentes níveis de variação.

2. Classificação dos dados:

A primeira etapa consistiu na classificação dos dados, seguida de uma apresentação em sala de aula para discussões.

3. Escolha da ferramenta:

Buscou-se encontrar uma ferramenta que não necessita-se aprender outra linguagem de programação para manipulá-la. Como a preferência foi realizar o trabalho em Python, resolveu utilizar a ferramenta Plotly. Essa ferramenta é uma API gráfica para algumas linguagens de programação como R e Python. Uma sessão sobre a ferramenta estará presente no artigo final.

4. Utilização da ferramenta:

Nessa etapa foi instalada a ferramenta no local de trabalho e começou-se a realizar alguns tutoriais para adquirir os conceitos básicos de como funciona o Plotly em Python, como gerar gráficos com ele, como utilizar os modo online e offline.

A ferramenta disponibiliza alguns gráficos voltados para a área financeira. Oferece algumas visualizações para ilustrar o comportamento da bolsa como:

- Time series

- OHLC (Open High Low Close) Charts
- Candlestick Charts
- Funnel Charts

Todas essas visualizações foram utilizadas com os dados da bolsa Dow Jones do trabalho e o resultado estará simplificado no relatório final.

5. Novas descobertas e novas análises

As visualizações obtidas na etapa anterior poderão servir de base para a continuação do projeto. Atualmente me encontro nessa etapa. Com base nas ferramentas disponíveis pela API estou verificando a possibilidade de realização dos tópicos descritos nos objetivos. Ainda não comecei a preparar os filtros, visualizações em períodos e ligação com as notícias.

6. Trabalhos futuros e outras ideias

Nessa etapa será discutido o que não conseguiu ser implementado no trabalho. Os motivos serão explicados e será discutido o que necessita ser feito para conseguir tais objetivos. Poderão aparecer alguns novos tópicos e ideias para uma possível sequência desse trabalho como projeto de mestrado. A utilização de outras bolsas de valores como a B3 não está descartada.