**Análise por repetição de comentários**

A análise de padrões numéricos, como todas as análises desenvolvidas nesse projeto, se baseia no **princípio de ondas de Elliot,** que afirma que os fenômenos sociais e a especulação financeira são cíclicos.

Nessa análise identificamos a semelhança entre o que é publicado em uma rede social sobre a Bitcoin atualmente, e o que já foi publicado no passado. A partir dessa semelhança, e do valor histórico da Bitcoin, obtemos uma estimativa de sua tendência futura.

Nesse projeto analisamos publicações na rede social **Reddit**, na sessão específica de debates sobre a Bitcoin.

A análise detalhada é composta pelos seguintes passos:

**Passo 1: Listamos as mensagens publicadas nas últimas 24 horas.**

Na tabela a seguir estão as mensagens trocadas nas últimas 24 horas.

[Inserir mensagens trocadas nas últimas 24 horas]

**Passo 2: Separação em frases menores.**

Utilizamos o motor de processamento de linguagem natural Spacy para separar mensagens longas em sentenças menores.

Tome por exemplo, a mensagem a seguir:

<Inserir mensagem de maior tamanho + quantidade de frases>

Essa mensagem será dividida nas seguintes sentenças:

<Inserir lista de frases para a mensagem acima>

**Passo 3: Realizamos um tratamento nas frases, removendo caracteres especiais, links e etc.**

<Inserir lista de frases pós tratamentos>

**Passo 4: Analisamos ocorrências anteriores de cada frase e obtemos nosso resultado.**

Para cada frase, analisamos suas ocorrências anteriores.

Analisamos também o valor da Bitcoin no dia seguinte a data de postagem de cada ocorrência anterior da frase.

Para cada ocorrência em que o valor da Bitcoin no dia seguinte foi de pelo menos 5% maior do que no dia de postagem da frase, contamos como uma ocorrência Positiva.

Para cada ocorrência em que o valor da Bitcoin no dia seguinte foi de pelo menos 5% menor do que no dia de postagem da frase, contamos como uma ocorrência Negativa.

Ocorrências em que o valor da Bitcoin no dia seguinte teve uma variação menor do que 5% são ignoradas.

<Inserir tabela com: Frase (Com link pro post); Data da última Ocorrencia Anterior (Com link pro post); Quantidade de quedas anteriores; Quantidade de aumentos anteriores; Tendencial Geral>

Por fim, o resultado da análise será baseado se a maioria das frases que demonstraram tendência tiveram tendência positiva ou negativa.

Nessa análise específica, a tendência foi: <Inserir resultado da análise por repetição de comentários>

A análise consolidada, disponível no dahsboard, terá aumento de 5% a análise por repetição de comentário e a análise semântica tiverem resultado positivo.

Por analogia, A análise consolidada terá queda de 5% se ambas as análise por repetição de comentário e a análise semântica tiverem resultado negativo.