BHA (Bitcoin Hashtag Analyser)

Overview:

BHA will be divided into two easily differentiable parts.

The first, the Intelligence Engine, will run for a few months looking through data from news and social media sites then looking for articles related to bitcoin. Then it will look at a list of sentimental words and look at the occurrences of those words on both, days when there is a drop in price and days when there is a rise in price. Through this process, it will attempt to accurately categorise words in the sentimental words list by what price movement they are usually associated with and by how much.

The second part of BHA is called the prediction engine. It will first go through several social media and news sites and look for occurrences of the word Bitcoin or related phrases. It will then analyze the surrounding text and through multiple ways, determine the sentiment behind the text.

This will allow BHA to then with that information and other information, to determine the price movement of Bitcoin by comparing to data gathered by the intelligence engine.

Process:

Prediction Engine:

Go through a list of social media sites and news sites.

Get any text that contains bitcoin related material in it by looking at a list of bitcoin related words from a file.

Convert the text to a unified format.

Compare occurrences of words in the influential word files made by the Intelligence Engine to the occurrences of words found in the articles and social media posts.

In detail:

Get the top 1000 words in the social media posts and articles (on an occurrence basis) get the effect that each of those words has on the price of bitcoin then average out all of the effects and if the result of doing so is a positive number then the price will be predicted to go up and vice-versa.

Get the sentiment of the top 1000 words as well using a pure sentiment database and also using a python sentiment module

Get the search trends for bitcoin and depending on that BHA will also be able to get a pretty good determination as to whether or not the price will rise or fall.

Intelligence Engine:

Go through a list of social media sites and news sites.

Get any text that contains bitcoin related material in it by looking at a list of bitcoin related words from a file.

Convert the text to a unified format.

Get current bitcoin daily change in dollars using an average from multiple exchanges.

Gets the top 1000 words (on an occurrence basis) then it goes through each word in the top 1000 and it:

Adds the word to a influential words file and then adds (occurrence \* price movement of the day) to the list of values for the word in the file and if it already exists then it just adds the (occurrence \* price movement of the day) to the word.

Notes:

This same method would be applied to any instrument, in theory.

This has not been tested and so it is totally theoretical, I am going to build it soon or at least make a POC.

Please leave comments.

Thanks,

Alex :)

Social Media and News sites that BHA will be using:

|  |  |
| --- | --- |
| Name | Home Page URL |
| Twitter | <https://twitter.com/> |
| Facebook | <http://www.facebook.com/> |
| LinkedIn | <http://www.linkedin.com/> |
| Google+ | <https://plus.google.com/> |
| Tumblr | <https://www.tumblr.com/> |
| Pinterest | <https://www.pinterest.com/> |
| Instagram | <http://www.instagram.com/> |
| Reddit | <http://www.reddit.com/> |
| Google News | <http://news.google.com/> |
| Yahoo News | <http://coindesk.com> |
| CNN | <http://www.cnn.com/> |
| New York Times | <http://www.nytimes.com/> |
| Fox News | <http://www.foxnews.com/> |
| NBC news | <http://www.nbcnews.com/> |
| Daily Mail | <http://www.dailymail.co.uk/> |
| The Guardian | <http://www.washingtonpost.com/> |
| BCC News | <http://news.bbc.co.uk/> |

Related Words to Bitcoin:

Cryptocurrencies

Cryptocurrency

Digital currency

Digital currencies

Virtual currency

Virtual currencies

Virtual currency

Bitcoin

Lightcoin

Altcoin

Blockchain

Crypto

Btc

Ltc