Research Project Documentation

vanGoethem, Joren   Maerten, Andreas

October 19, 2021

Contents

1 Data collection

Data   For our training data we used both the crypto price data [Open, Low, High, Close, Volume] along with tweets and rss feeds related to cryptocurrencies.

2 Data preprocessing

Normalization   The raw price data is not suitable for training a neural network with. We need to normalize all the data so the neural network can properly adjust to the data. for this we will be using the percentage change from one data point to the next.

the following formula can be used to calculate the percentage change, where P is the percentage change, V1 is the previous value and V2 is the value of which we want to calculate the percentage change over V1. the result can be a positive or a negative value.

$P = \frac{V_2 - V_1}{V_1} \times 100$

Labeling   We will probably need to manually label tweets and rss feed data unless we can find a substantial kaggle dataset that is already labeled. Both tweets and rss articles can probably be inserted into the same model to predict whether the tweet or article is positive or negative towards a specific cryptocurrency.

3 AI models

Deep Reinforcement Learning

Long Short Term Memory

Sentiment Analysis