Using the news articles the whole process of predicting the forecast of exchange rate will be automated and for this the research paper has used specific news articles as a dataset which are written by professionals following strict rules, which will make the news appropriate to foreign exchange and remove all the unwanted news headlines like of sports, politics etc.

As usually the traders see the news headlines and predict whether the rate of the particular currency will go up or down. The same thing will be done by the model, which will see the news headlines and predict, whether the currency rate will increase or decrease and at which rate?

They have news headlines with its timestamp containing the date and the time with seconds as in an hour they will receive 40 headlines. They used four hundred records consisting of a sequence of two to five words as the keywords to for the model.

For the model the occurrence of keywords in the dataset will be counted in weights to find the most effective news headline which can predict whether the currency rate will be steady, will go up, or will be down. For calculating currency rate for current time period t, they will use the data from time periods t-1 and t-2, as they will give the best accuracy.

They have made rules, by which they are able to handle continuous attributes and do not rely on Boolean tests. They have therefore more expressive power by retaining the strength of rule classifiers using comprehensible models and relatively fast learning algorithms.

The dataset they have used is not mentioned in the research paper, so I will try to find a dataset or any website where by using web scraping we can get the news headlines.

Research paper link: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=8be9790d0c4b7bceeed174ad818c6f0b36f738ad>