**A Methodology for Securities and Cryptocurrency Trading Using Exploratory Data Analysis and Artificial Intelligence**

This paper discusses securities and cryptocurrency trading using artificial intelligence (AI) in the sense that it focuses on performing Exploratory Data Analysis (EDA) on selected technical indicators before proceeding to modelling, and then to develop more practical models by introducing new reward loss function that maximizes the returns during training phase. The results of EDA reveal that the complex patterns within the data can be better captured by discriminative classification models and this was endorsed by performing back-testing on two securities using Artificial Neural Network (ANN) and Random Forests (RF) as discriminative models against their counterpart Naïve Bayes as a generative model. To enhance the learning process, the new reward loss function is utilized to retrain the ANN with testing on AAPL, IBM, BRENT CRUDE and BTC using auto-trading strategy that serves as the intelligent unit, and the results indicate this loss superiorly outper forms the conventional crossentropy used in predictive models. The overall results of this work suggest that there should be larger focus on EDA and more practical losses in the research of machine learning modelling for stock market prediction applications.

**EXISTING SYSTEM:**

If profitability is the main prediction objective, formalizing the problem as classification suits it more than regression. There are two main hypotheses in stock market prediction as discussed by Ican in . The first hypothesis suggests that stock prices move in a stochastic manner and cannot be predicted (Efficient Market Hypothesis). On the other hand, the other hypothesis claims that stock markets are predictable, at least to a certain degree, and prediction turns into long-term profits. As a matter of fact, the literature is rich with articles that support the latter hypothesis. Specifically, the research is ongoing to investigate the feasibility of probabilistic classification models in terms of profits generation capabilities. One of the main challenges related to modelling stock markets for prediction purposes is the selection of features set for training, and which models optimally fit with this set. Features and model selection philosophy is concerned about the method with which the features were selected, that is whether they are selected based on performing any type of analysis or arbitrary selection. Proper features selection does not only contribute to building more accurate and reliable AI models.

**DISADVANTAGES OF EXISTING SYSTEM:**

1. that stock prices move in a stochastic manner and cannot be predicted.
2. the work was not proceeded in testing their efficiency in the application

**Algorithm:**

**PROPOSED SYSTEM:**

The novel contribution of this work is proposing a methodology for AI trading with systematic approach of developing and training the models. This mainly constitutes of setting and testing two hypotheses claiming that performing EDA and introducing new reward loss result in higher returns. . To enhance the learning process, the new reward loss function is utilized to retrain the ANN with testing on AAPL, IBM, BRENT CRUDE and BTC using auto-trading strategy that serves as the intelligent unit, and the results indicate this loss superiorly outper forms the conventional crossentropy used in predictive models. The overall results of this work suggest that there should be larger focus on EDA and more practical losses in the research of machine learning modelling for stock market prediction applications.

**ADVANTAGES OF PROPOSED SYSTEM:**

* 1. Retrieving the original data; that is the assets opening, closing, high and low prices over the considered time period for model development.
* 2. Man-crafted features extraction (technical indicators).
* 3. Numerical data analysis by utilizing features selection methods.
* 4. Exploratory data analysis (visualization) to assess the relationship between selected features and classes, data statistical distributions and time dependency of the data.
* 5. Proposing and training appropriate probabilistic models.
* 6. Assessing the selected trained models in terms of accuracy and profits (test of first hypothesis).
* 7. Retrain models with a reward loss function that has more practical sense for stock prediction applications.

**Algorithm:** **artificial intelligence, machine learning, probabilistic modelling, classification models, artificial neural network, random forests, naïve bayes**

**SYSTEM REQUIREMENTS:**

**HARDWARE REQUIREMENTS:**

* System : Intel Core i3.
* Hard Disk : 1TB.
* Monitor : 15’’ LED
* Input Devices : Keyboard, Mouse
* Ram : 8GB.

**SOFTWARE REQUIREMENTS:**

* Operating system : Windows 10.
* Coding Language : Python
* Tool : PyCharm, Visual Studio Code
* Database : SQLite

**REFERENCE:**

Date of Conference: 21-23 May 2021 Date Added to IEEE Xplore: 22 June 2021 ISBN Information: Electronic ISBN:978-1-7281-7029-9 CD:978-1-7281-7027-5 DVD ISBN:978-1-7281-7028-2 Print on Demand(PoD) ISBN:978-1-7281-7030-5 INSPEC Accession Number: 20712735 DOI: 10.1109/INCET51464.2021.9456351 Publisher: IEEE Conference Location: Belagavi, India