The International Young Researchers' Conference
DOI: 10.34614/iyrc2021f26

Finding ESG's Aptitude for Projecting Financial Value by Novel Machine Learning

Daniel Pyeong Kang Kim
The Stony Brook School, USA
thezenithcoltd@gmail.com

Abstract
In the last few years, Environmental, social and Corporate Governance (ESG) has established its place as a measure that discloses intangible assets or liabilities of a company. Prominent as it is, there is some skepticism regarding whether ESG serves as a suitable tool for analyzing the financial prospect of an investment. While many papers concerning this metric advocate the use of ESG with their study, some claim that there are better alternatives to ESG. Thus, this paper seeks to investigate the extent by which a company's investment can be predicted with its ESG ranking, and the accuracy of ESG in doing so compared with that of other financial features. The main form of analysis used was Exploratory Data Analysis, which was employed to show any existing correlations between monetary traits. The data used for this analysis, namely the companies' financial features, was extracted from company performances in the NYSE market for the last 10 years. With 21 financial features included in each company's data, the study extracted or removed certain types of features according to their accuracy. To determine the accuracy and capability of several pipelines in classifying investments, the paper adopted eight Machine Learning Classifiers. Though these classifiers yielded similar accuracies amongst themselves, the pipelines showed a sharp distinction: algorithm classifiers containing ESG in the train process displayed a substantially higher accuracy than those without ESG. This paper demonstrates that ESG is a comprehensive, valid instrument for investors to evaluate the accurate investment worthiness of an entity.

Keywords: Environmental Social and Corporate Governance, machine learning,

Introduction

Background

Receiving striking attention from contemporary analysts, ESG is an acronym that considers the environmental, social, and corporate governance aspects of a company. Since financial or income statements do not disclose the company's performance on these aspects, ESG helps to view a company's progress from a different perspective[1]. The term was first coined in the 2006 United Nations *Principles for Responsible Investment (PRI)* report. From then on, this relatively novel index facilitated the exhibition of intangible assets to the public. With the rising interest in sustainable investment, such information is becoming more invaluable to investors who desire a profound understanding of a corporation. The escalating popularity and prevalence of ESG investing have even resulted in serious discussions on placing a mandatory ESG disclosure for many companies[2].

Objective

An advantage of this metric is that it is a priceless tool for investments that adhere to social responsibility. Even so, people are skeptical of

ESG's capacity to accurately reflect the monetary worth of a company. Some say it is merely a superficial report and does not represent the true financial value of an entity. It is true that the factors of ESG portray a company's ability to cope with times of turmoil[3]. Nevertheless, it may be exaggerated to claim ESG and monetary success display a causal relationship. Thus, this paper intends to display how accurately the ESG metric can determine monetary growth.

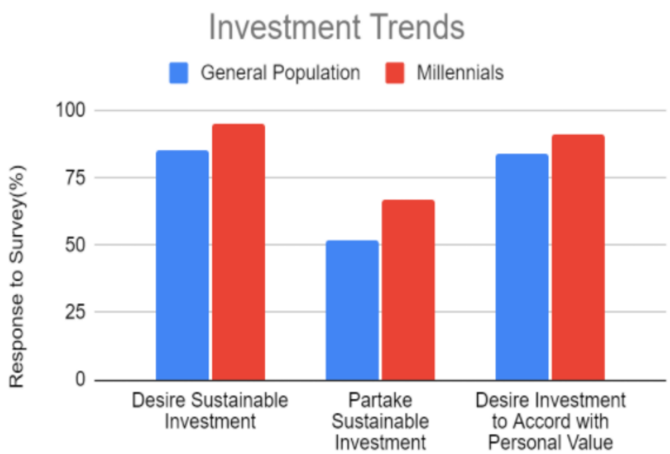


FIGURE 1. Investment trends from general population and millennials

Related Works

To understand other evaluations of ESG, other papers regarding the measurement were considered for general comprehension.

Having the most resemblance to this paper, attempts to display the link between ESG profiles and financial performances. The authors of the paper developed and used a sophisticated machine learning algorithm to identify the possible link. With their machine learning algorithm, the team used data from *Capitalization-weighted MSCI World Index USD* and ESG ratings from *Sustainalystics*. The final result of the study showed that though a link between the two exists, it can only be accessed with non-linear techniques[4].

Similar to the previous study, *Does Good ESG Lead to Better Financial Performances by Firms? Machine Learning and Logistic Regression Models of Public Enterprises in Europe* tries to assess the accuracy of financial indicators such as ROE and ROA while identifying ESG's effect on financial performance. To calculate the accuracy of ROE and ROA, the team accessed company data from *Thomson Reuters Eikon*. The team incorporates Random Forest, Support Vector Regression, Ridge Regression, and Inferential Model to reach its objective. In the end, they concluded that ROE and ROA were accurate, which in turn supported the link of ESG with Financial performance

Unlike the two research papers mentioned above that tested the validity of ESG, *ESG2Risk: A Deep Learning Framework from ESG News to Stock Volatility Prediction* is a more specific analysis that focuses on ESG with the value volatility of a company. The main source the paper derives data from was general ESG information from news-flows. It primarily utilizes Bayesian learning and a Transformer-based language model to analyze this given data and the validity of their language model. The team discovered that the Transformer-based language model successfully predicts future volatility of stock return, thus identifying the return and the risk of a company.

Surprisingly, there was also research like *Mind the gap! Machine learning, ESG metrics and sustainable investment* that tried to replace ESG with another measurement that was more transparent and exhaustive. The team extracted data from *EURO STOXX 300* and *MSCLESG Research* to demonstrate the newly-created data's efficiency and accuracy. The paper constructs their index's validity with MATLAB built-in regression, Linear regression, CAPM, and Birr model. The final conclusion of the company was that more information could be perceived with ML techniques compared to available ESG indicators.

Materials and Methods

Light Gradient Boosting Machine

Gradient Boosting Decision Tree(GBDT), which includes XGboosting implicates the trade off problem between computation time and efficiency. Therefore, Light Gradient Boosting Machine (LGBM) solves those problems by implementing new algorithms which are Gradient Based One SIde Sampling(GOSS) and Exclusive Feature Bundling(EFB). Furthermore, unlike XGboosting which undergoes level wise method, LGBM undergoes leaf wise one. GOSS first eliminates the data with low gradients and then calculates the whole information gain from the rest of the data. EFB groups the mutual exclusive variables into one bundle. An elaborately designed variable-search algorithm can produce the same thing as a histogram of individual variables by grouping them.

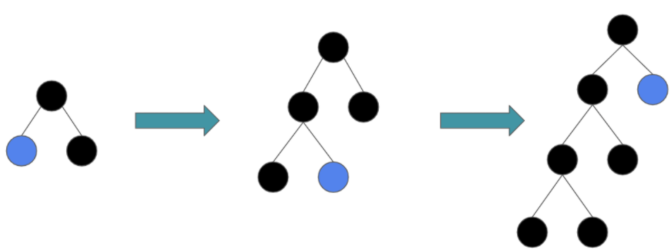
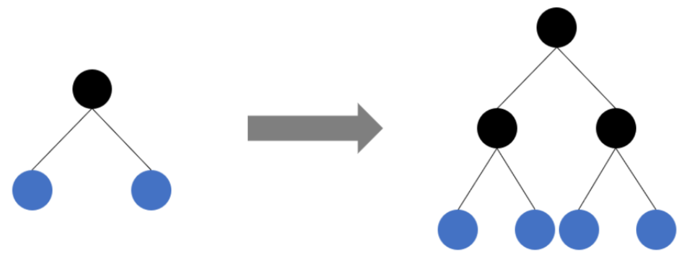


FIGURE 2. Overall architecture of light gradient boosting machine

Catboost

Catboost belongs to the boosting algorithm, which focuses on handling categorical features in the given dataset. Yandex researchers developed this algorithm and it outperformed various boosting algorithms including xgboosting and gradient boosting. As the catboost performs random permutations on ordered boosting during the training process, it can prevent overfitting efficiently compared to existing boosting algorithms[]. Likewise, catboost builds a level-wise tree-like BFS algorithm, which is also suitable for the XGboosting. In order to accelerate the permutation speed, the catboost performs categorical feature combinations, which is about bundling multiple identical features into one feature, based on the information gain. At last, lgbm or XGboosting are sensitive to the hyper parameter tuning to prevent the over fitting. However, the catboost overcomes the overfitting through the algorithms, therefore, it no longer



concentrates on hyper parameter tuning.

FIGURE 3. Overall architecture of catboost algorithm

Data Description

The data used for this research, created by Immanol Recio Erquicia, is an exhaustive data that includes random investments with financial ratios and ESG ranking. Extracting information from the 10 last years of the NYSE market, the data consists of 405,258 companies with 21 financial aspects of a company including inflation, investment suitability, expected return, and ESG ranking. With the comprehensive scale of data, various pipelines were able to be devised

Experimental pipeline

The accuracy of predicting whether an investment was bad or good was based on three samples: data with all features of a company excluding the ESG ranking, data where features were extracted by the correlation function, and data with exclusively the ESG ranking. The first of the three was relatively standard with Train test split and Standard scaler steps resulting in the Machine learning classifiers. The second process involved calculating the total correlation and selecting features with comparatively high correlations, which are 'inflation' and 'nominal-return'. This was followed by the usual phases of Train test split and Standard scaler. In the third pipeline, the ESG feature extracting process was added from the first pipeline to learn the effectiveness of the ESG ranking in predicting investment suitability.

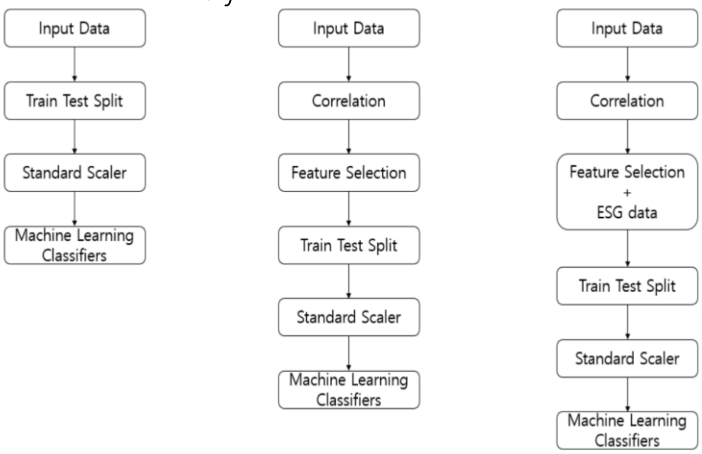


FIGURE 5. Three different experiment pipelines from our research

Results

Concurring with many preceding studies on ESG's relation with financial performance, this study displays a strong relationship between ESG ranking and stock investment. To find which pipeline showed the best overall performance, this paper adopted 8 machine learning classifiers: DecisionTreeClassifier, LogisticRegression, GradientBoosting, AdaBoostClassifier, RandomForestClassifier, XGBClassifier, LGBMClassifier, and ExtraTreeClassifier. The accuracy mentioned in the experiment is how precisely each machine learning classifier was able to predict the investment suitability of a company, given its financial features.

The first pipeline, where all the features were used to identify investment adequacy, showed imposing results. The accuracy of DecisionTreeClassifier was 100, LogisticRegression was 88.06, GradientBoosting was 100, AdaBoostClassifier was 100, RandomForestClassifier was 100, XGBClassifier was 100, LGBMClassifier was 100, ExtraTreeClassifier was 99. Though there were some anomalies to how accurate the machine learning classifiers were, the overall precision was almost flawless.

Counterintuitively, the next pipeline, where three relatively higher correlations were extracted looking at the heatmap below, displayed deficient outcomes. The accuracy of DecisionTreeClassifier was 77.65, Logistic Regression was 67.37, Gradient Boosting was 76.13, AdaBoostClassifier 75.54, RandomForestClassifier was 83.2, XGBClassifier was 75.95, LGBMCLassifier was 76.72, and ExtraTreeClassifier was 83.13. None of these classifiers in the second pipeline were able to yield better results than those in the first pipeline.

Outperforming the aforementioned two pipelines, the last pipeline, which used ESG ranking and the two features with the highest correlation to evaluate an investment, rendered almost flawless results. All of the machine learning classifiers except one, the logistic regression, displayed an accuracy of 100, meaning that almost all classifiers could identify which investments were good or bad once they were given the ESG ranking of the company.

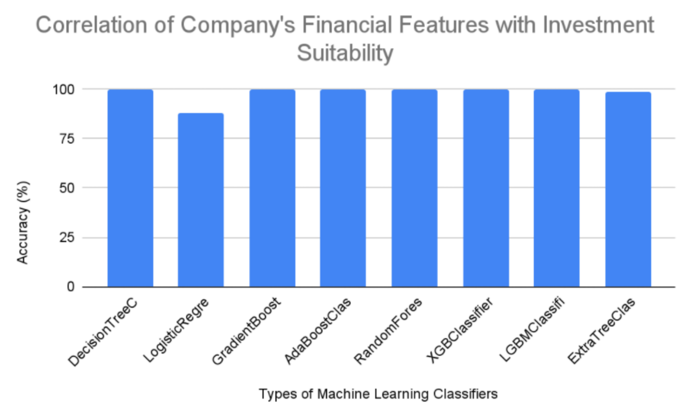


FIGURE 6. Accuracy comparison among various machine learning models based on the first pipeline

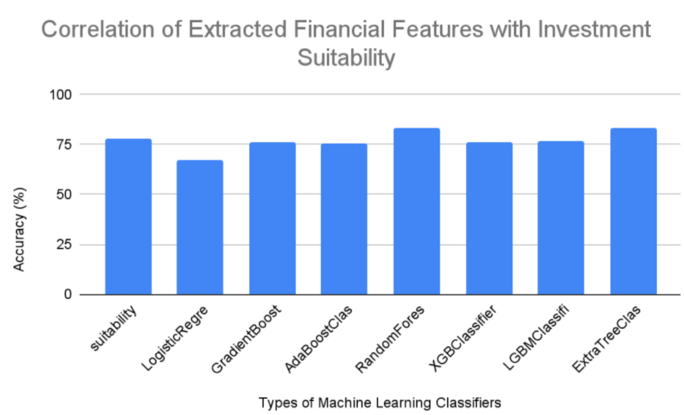


FIGURE 7. Accuracy comparison among various machine learning models based on the second pipeline

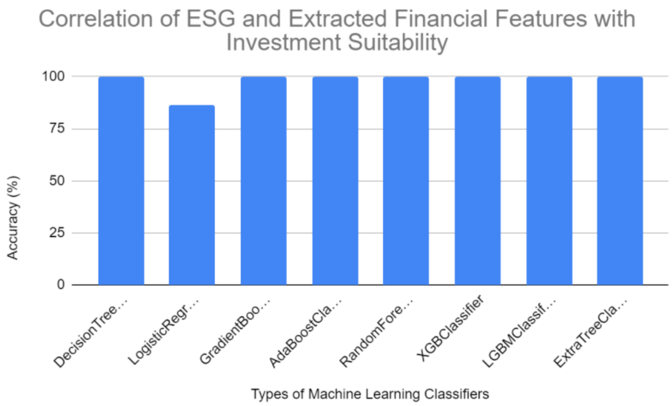


FIGURE 8. Accuracy comparison among various machine learning models based on the third pipeline

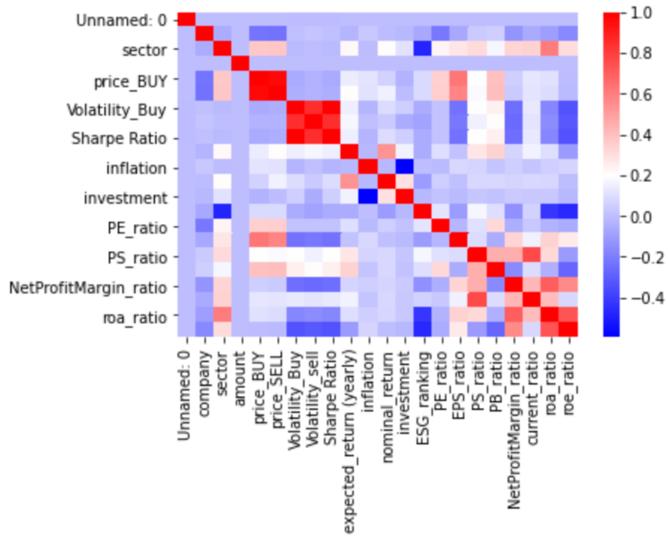


FIGURE 9. Heatmap of the correlation matrix for the feature selection

Discussions

Principal Finding

The most remarkable finding of our experiment was the drastic variation of accuracy in the second and third pipeline. As mentioned before, the second pipeline consists of financial features that have a high correlation relative to other financial features. As ESG ranking was not part of these extracted features, the second pipeline essentially shows the performance of machine learning algorithms without the ESG ranking. Given that the major discrepancy between the two pipelines is the ESG ranking in the train and test process, the difference in accuracy explicitly substantiates ESG's effectiveness when assessing an investment. Surprisingly, ESG seems to have a superior correlation with investment suitability than that of other conventionally employed company features such as PE or PS ratio. Unlike the previous researches mentioned in related work section, our research proved the importance of ESG variable by comparing the result through adding the ESG variable to the group of variables with high correlation result. Furthermore, our experiment discovered a limitation of feature importance function from various tree based machine learning models.

Limitations

A major limitation of the study, though, was that the feature importance, which identifies the features that contribute the most to a certain result, of ESG was very low. Some algorithms even experienced an importance score of almost 0 for ESG, as shown in Figure 10. Nonetheless, it is implausible to assert that the high accuracy of ESG ranking with regard to investment suitability is a mere coincidence. Thus, though this result could be worrying for the advocates of ESG when viewed from a superficial level, it is logical to conclude that the low importance score is a drawback of the feature importance itself rather than the accuracy and validity of ESG.

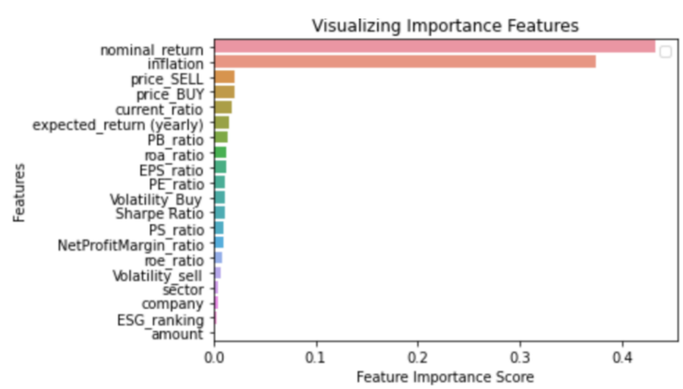


FIGURE 10. Feature importance from the light gradient boosting machine

Another limitation is the relatively low, 65 percent, accuracy rate when only the ESG ranking was used in the train and test process. This limitation shows that when investors are analyzing a company, using solely the ESG ranking will not produce the best results. Thus, we are not sure as to why classifiers with and without the ESG ranking yield stark results.

To properly find the feature and target relationship, there has to be further research by means of XAI deep learning.

Conclusions

In short, this research paper can conclude and substantiate the utility of ESG even when analyzing a firm's financial performance.

Exploiting 8 different algorithm classifiers, the paper found a perfect correlation between ESG ranking, with two other financial features, and investment, yielding an accuracy of 100% for seven out of the eight classifiers. On the other hand, when the algorithm used train-test split for solely other financial variables, such as nominal return or PE ratio, there was a noticeable difference with accuracies ranging from high sixties to low eighties.

With the surge of interests on investment, this finding is critical for investors who are concerned with investing responsibly while making a profit. Not only does ESG represent environmental and social impact along with management structure, it also determines a company's current and future aptitude to make profit.

References

400k NYSE random investments + financial ratios. (2021, January 18). Kaggle. https://www.kaggle.com/imanolrecioerquicia/400k-nyse-random-investments-financial-ratios

Atkins, B. (2021, June 29). Demystifying esg: Its history & current status. Forbes. https://www.forbes.com/sites/betsyatkins/2020/06/08/demystifying-esgits-history--current-status/?sh=107c37802cdd

Casalicchio, G., Molnar, C., & Bischl, B. (2019). Visualizing the Feature Importance for Black Box Models. Machine Learning and Knowledge Discovery in Databases, 655–670. https://doi.org/10.1007/978-3-030-10925-7_40

Cheon, M.J., & Lee, O.(2021). Detecting olfactory impairment through objective diagnosis : catboost classifier on EEG data. Journal of Theoretical and Applied Information Technology, 99(20), 3596-3604

De Franco, C., Geissler, C., Margot, V., & Monnier, B. (2020). ESG investments: Filtering versus machine learning approaches. arXiv preprint arXiv:2002.07477.

De Lucia, C., Pazienza, P., & Bartlett, M. (2020). Does good ESG lead to better FINANCIAL performances by Firms? Machine learning and logistic regression models of public enterprises in Europe. Sustainability, 12(13), 5317. https://doi.org/10.3390/su12135317

Dorogush, A. V., Ershov, V. & Gulin, A. (2018). CatBoost: gradient boosting with categorical features support.. CoRR, abs/1810.11363.

ESG Investing and Analysis. (n.d.). CFA Institute. Retrieved August 26, 2021, from https://www.cfainstitute.org/en/research/esg-investing

Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G. Z. (2019). XAI—Explainable artificial intelligence. Science Robotics, 4(37), eaay7120. https://doi.org/10.1126/scirobotics.aay7120

Guo, T. (2020). ESG2Risk: A Deep Learning Framework from ESG News to Stock Volatility Prediction. SSRN Electronic Journal. Published. https://doi.org/10.2139/ssrn.3593885

Jia, Y., Jin, S., Savi, P., Gao, Y., Tang, J., Chen, Y., & Li, W. (2019). GNSS-R Soil Moisture Retrieval Based on a XGboost Machine Learning Aided Method: Performance and Validation. Remote Sensing, 11(14), 1655. https://doi.org/10.3390/rs11141655

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. NIPS.

Lanza, A., Bernardini, E., & Faiella, I. (2020). Mind the gap! Machine Learning, ESG Metrics and Sustainable Investment. SSRN Electronic Journal. Published. https://doi.org/10.2139/ssrn.3659584

Morgan Stanley. (2019). Sustainable Signals: Individual Investor Interest Driven by Impact, Conviction and Choice.

Ostroumova, L., Gusev, G., Vorobev, A., Dorogush, A.V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. NeurIPS.

Daniel Kim
The Stony Brook School, New York, USA
danielpkkim1001@gmail.com

[1] ESG Investing and Analysis. (n.d.). CFA Institute. Retrieved August 26, 2021, from https://www.cfainstitute.org/en/research/esg-investing

[2] Atkins, B. (2021, June 29). Demystifying esg: Its history & current status. Forbes. https://www.forbes.com/sites/betsyatkins/2020/06/08/demystifying-esgits-history--current-status/?sh=107c37802cdd.

[3] Morgan Stanley. (2019). Sustainable Signals: Individual Investor Interest Driven by Impact, Conviction and Choice.
[4] De Franco, C., Geissler, C., Margot, V., & Monnier, B. (2020). ESG investments: Filtering versus machine learning approaches. arXiv preprint arXiv:2002.07477.

[5] De Lucia, C., Pazienza, P., & Bartlett, M. (2020). Does good ESG lead to better FINANCIAL performances by Firms? Machine learning and logistic regression models of public enterprises in Europe. Sustainability, 12(13), 5317. https://doi.org/10.3390/su12135317

[6] Guo, T. (2020). ESG2Risk: A Deep Learning Framework from ESG News to Stock Volatility Prediction. SSRN Electronic Journal. Published. https://doi.org/10.2139/ssrn.3593885

[7] Lanza, A., Bernardini, E., & Faiella, I. (2020). Mind the gap! Machine Learning, ESG Metrics and Sustainable Investment. SSRN Electronic Journal. Published. https://doi.org/10.2139/ssrn.3659584

[8] Jia, Y., Jin, S., Savi, P., Gao, Y., Tang, J., Chen, Y., & Li, W. (2019). GNSS-R Soil Moisture Retrieval Based on a XGboost Machine Learning Aided Method: Performance and Validation. Remote Sensing, 11(14), 1655. https://doi.org/10.3390/rs11141655

[9] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. NIPS.

[10] Ostroumova, L., Gusev, G., Vorobev, A., Dorogush, A.V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. NeurIPS.

[11] Dorogush, A. V., Ershov, V. & Gulin, A. (2018). CatBoost: gradient boosting with categorical features support.. CoRR, abs/1810.11363.

[12]
[13] 400k NYSE random investments + financial ratios. (2021, January 18). Kaggle. https://www.kaggle.com/imanolrecioerquicia/400k-nyse-random-investments-financial-ratios

[14] Casalicchio, G., Molnar, C., & Bischl, B. (2019). Visualizing the Feature Importance for Black Box Models. Machine Learning and Knowledge Discovery in Databases, 655–670. https://doi.org/10.1007/978-3-030-10925-7_40

[15] Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G. Z. (2019). XAI—Explainable artificial intelligence. Science Robotics, 4(37), eaay7120. https://doi.org/10.1126/scirobotics.aay7120