Leeds University
Business School

UNIVERSITY OF LEEDS

Assessed Coursework Coversheet

For use with *individual* assessed work

Student ID Number:	2	0	1	7	1	0	3	4	5
Module Code:	LUBS5990M								
Module Title:	Machine Learning in Practice								
Module Leader:	Aritad Choicharoon								
Declared Word Count:	3485								

Please Note:

Your declared word count must be accurate, and should not mislead. Making a fraudulent statement concerning the work submitted for assessment could be considered academic malpractice and investigated as such. If the amount of work submitted is higher than that specified by the word limit or that declared on your word count, this may be reflected in the mark awarded and noted through individual feedback given to you.

It is not acceptable to present matters of substance, which should be included in the main body of the text, in the appendices ("appendix abuse"). It is not acceptable to attempt to hide words in graphs and diagrams; only text which is strictly necessary should be included in graphs and diagrams.

Table of Contents

MACHINE LEARNING MODEL FOR PREDICTING ICO SUCCESS

Introduction: Business Understanding of Initial Coin Offerings (ICOs)

Initial Coin Offerings (ICOs) are a modern way for companies, especially startups, to get funding. Instead of going through traditional methods like banks or venture capitalists, companies use blockchain technology to issue digital tokens to investors worldwide. These tokens can give investors certain benefits. ICOs are attractive because they open investment opportunities to more people. But they also come with risks, such as unclear regulations and the chance of fraud, due to the anonymous nature of blockchain transactions (Fisch, 2019). However, the success of an ICO largely depends on whether investors believe in the project, which can be affected by how clear the company's idea is and whether the team behind it seems trustworthy and capable (Huang et al., 2022).

ICObench serves as a pivotal platform in this ecosystem, providing a detailed database and rating system that helps investors and businesses alike evaluate and compare various ICOs. Utilizing machine learning to predict the success of an ICO can significantly help investors by enabling more informed decision-making (Xu et al., 2021).

The purpose of this technical report is to analyze and predict the performance of ICO companies, evaluating their project success and potential returns. It considers various factors such as technical, economic, fundraising teams, and market influences. The report aims to provide valuable insights for both fundraising teams and investors.

1. **Fundraising Teams:** Improved fundraising efficiency through targeted marketing, realistic goals, and optimized resource allocation (Alchykava & Yakushkina, 2021).
2. **Investors:** Prediction allows investors to assess project viability, make informed choices, and diversify their portfolios to reduce risk and potentially increase returns (Rasskazova et al., 2019; Alchykava & Yakushkina, 2021).

The report will test and compare machine learning models to identify key factors that predict ICO success. The result proposes a data-driven approach to enhance investment strategies and campaign success rates.

Data Understanding

The dataset contains 2767 rows and 16 features which contain details related to ICO projects.

Variable	Data Type	Description
ID	Categorical (Nominal)	Unique identifier for each fundraising project.
success	Categorical (Binary)	Indicates if the ICO was successful ('Y' or 'N').
startDate	Date	Starting date of the fundraising campaign.
endDate		Ending date of the fundraising campaign.
coinNum	Numerical (Continuous)	Number of blockchain coins to be issued.
priceUSD		Price of each blockchain coin in US dollars.
distributedPercentage		Percentage of coin distributed to investors.
teamSize	Numerical (Discrete)	Number of team members in the fundraising project.
countryRegion	Categorical (Nominal)	Country or region of the fundraising team/company.
platform		Blockchain platform used for the ICO.
brandSlogan	Text	Textual slogan of the fundraising team/company.
rating	Numerical (Ordinal)	Rating score for the project quality.
minInvestment	Binary (Categorical)	Indicator if a minimal investment amount is set.
hasVideo		Indicates presence of a video on the campaign page.
hasGithub		Indicates presence of a Github page.
hasReddit		Indicates presence of a Reddit page.

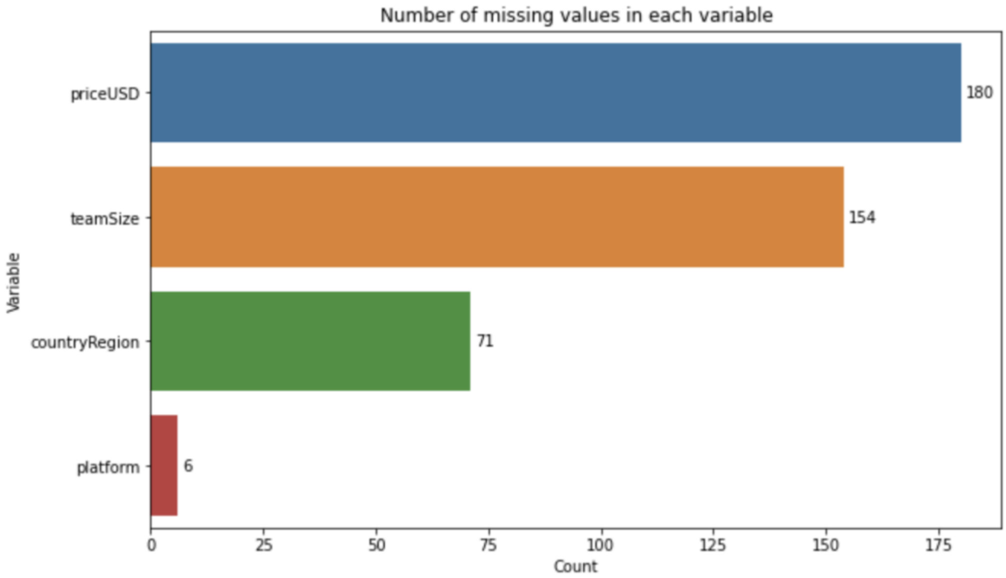
Table 1 Data Dictionary

ID Column

Dropping the **ID column** in data analysis can prevent overfitting and simplify datasets, as ID typically does not contribute meaningful information to the analysis.

Missing Values Analysis

- priceUSD: 180 NA values (6.51% of the total data)
- countryRegion: 71 blank values (2.57% of the total data)
- teamSize: 154 NA values 5.57% of the total data)
- platform: 6 blank values (0.21% of the total data)



Data Integrity and Validity Results

1. **Duplicates**: No duplicate entries were found in the dataset, which is good for data integrity.
2. **Date Validation**: There are 771 instances where the startDate is later than the endDate. This is a significant issue as it violates the logical constraints of time sequencing in the data.

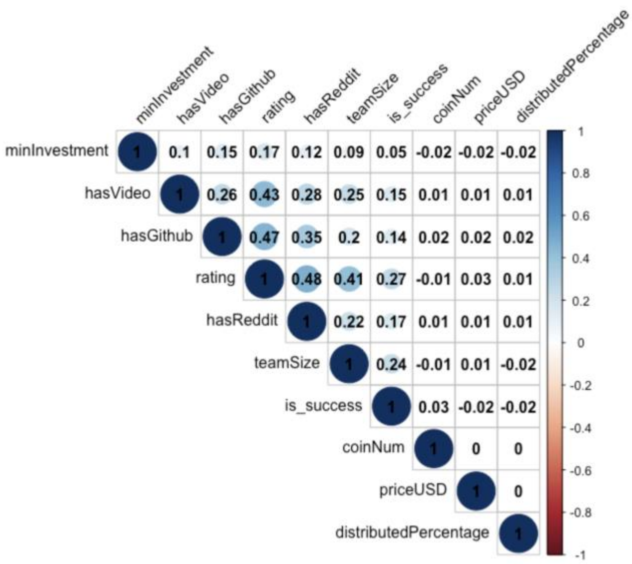


Figure 1 Correlation Plot

Strong Negative Correlation between **`rating`** and **`is_success`**: This suggests that more critically assessed projects tend to have lower success rates.

Positive Correlations involving **`hasGithub`**:Variables like **`hasGithub`** show positive correlations with several other variables, such as **`hasReddit`** and **`teamSize`**. This could indicate that projects with a GitHub presence tend to also have broader social media engagement and larger teams.

Data Cleaning and Preparation

Preparation of Categorical Features

success

- The "success" column is a binary target variable indicating whether an ICO project achieved its funding goal (success) or not (failure).
- Out of 2767 projects, 37.2% of ICO projects in the dataset were successful in reaching their funding targets, while 62.8% did not.

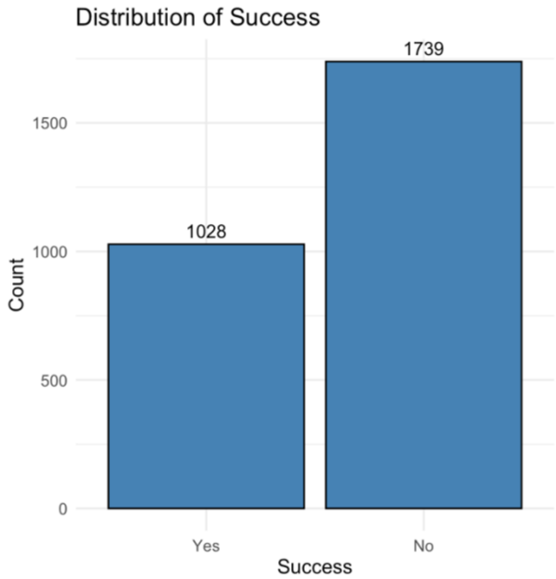


Figure 2 Distribution of Success

countryRegion

Preparation

The following Preparation steps on the countryRegion feature:

- **Lowercasing Country Names:** All country names were converted to lowercase to ensure uniformity and eliminate case sensitivity issues.
- **Trimming White Spaces:** Extraneous white spaces at the beginning or end of the country names were removed.
- Country names were replaced with their more commonly recognized forms (e.g., ‘curaçao’ to ‘curacao’, ‘méxico’ to ‘mexico’).
- **Replacement of Null Values:** Null values were replaced with “usa” (use mode for categorical) to handle missing data.
- **Correction of Errors in Spacing and Grammar:** Any errors in spacing and grammar were corrected to ensure the accuracy of the data.

Feature Engineering

To assess the impact of regulatory environments on ICO success (Offshore Protection, 2023), two binary features were added to the dataset: 'ico_friendly' and 'is_USA'. The 'ico_friendly' feature indicates whether a project's country is known to support ICO activities.

Addition of Feature ico_friendly: A binary feature 'ico_friendly' was added to the dataset where 1 indicates the country is 'Singapore", "Switzerland", "Cayman Islands", "United Arab Emirates", and 0 otherwise.

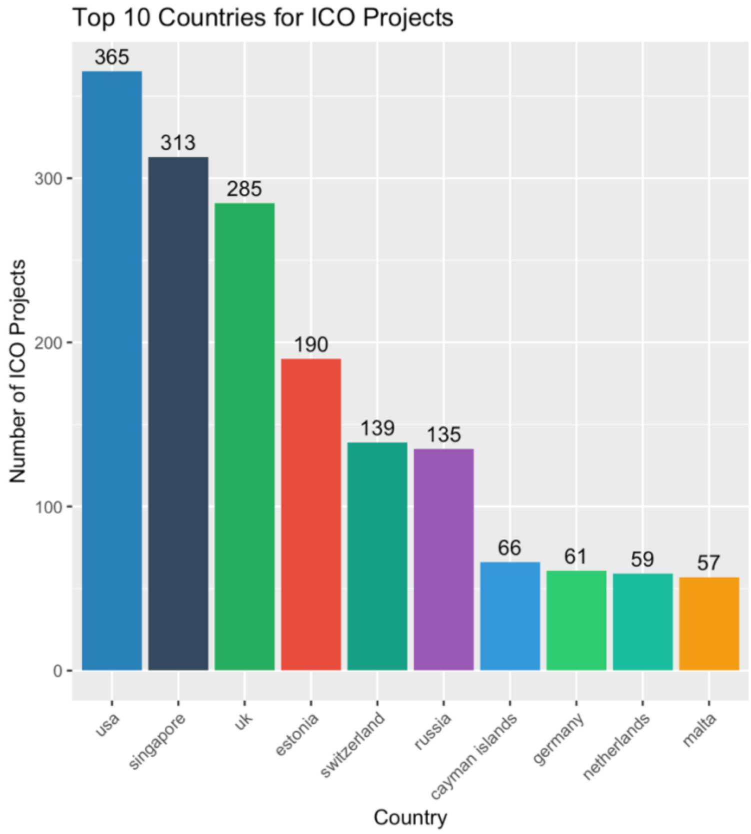


Figure 3 Top 10 Countries For ICO

platform

Preparation

The following Preparation steps on the 'platform' feature:

Lowercasing: All platform names were converted to lowercase to eliminate case sensitivity.
Trimming: Removed unnecessary trailing spaces.
Standardizing: To increase consistency, platform name variations were consolidated:

- 'btc' changed to 'bitcoin'
- 'eth', 'etherum', etc. unified to 'ethereum'
- 'x11 blockchain' simplified to 'x11'
- 'stellar protocol' shortened to 'stellar'
- 'pos+pow', 'pos,pow', etc. unified to 'pos_pow'

Null Values: Null values were replaced with "ethereum" (use mode for categorical) to handle missing data.

This transformation reduced the number of unique platforms from 130 to 96, demonstrating the effectiveness of the standardization.

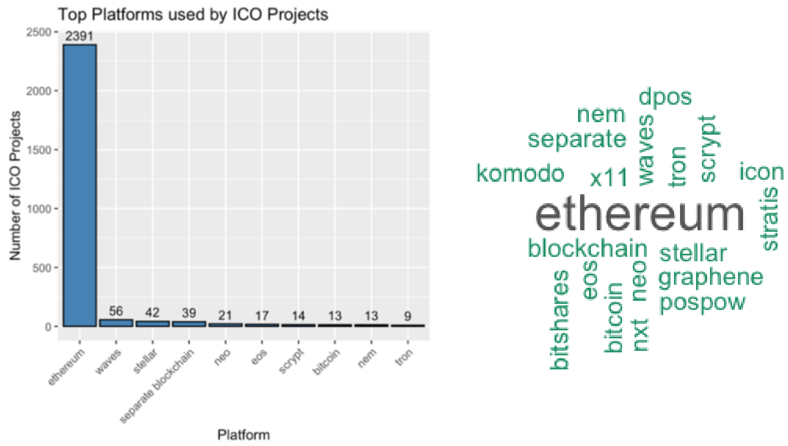


Figure 4 Plots of the platform.

Feature Engineering

As evident from the analysis 86 % of ethereum as the majority of ICOs platforms were launched.

- **Adding Feature: 'is_ethereum'**: A binary feature was created. '1' indicates projects using the Ethereum platform, '0' otherwise.

Preparation of Numerical Features

coinNum

Preparation

- The wide distribution of coin issuance: 12 to 22.6 quadrillion.
- Median issuance: 180 million coins.
- Highly skewed distribution due to extreme outliers.
- Remove ICOs issuing more than ten billion coins to reduce outlier impact.

Feature Engineering

- Logarithmic transformation was applied to 'coinNum' to normalize the distribution and reduce skewness.

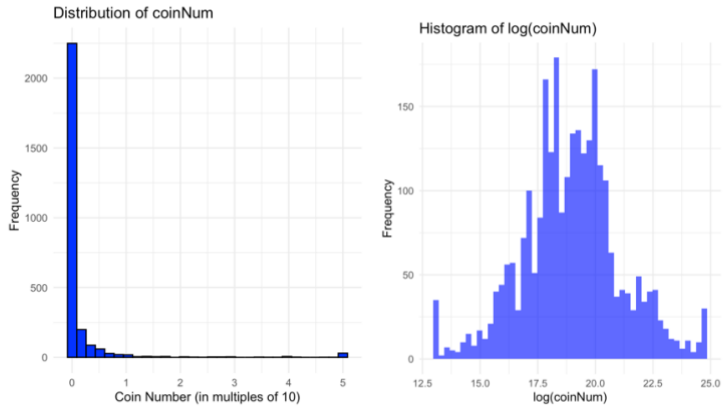


Figure 5 Before and after log transformation of coinNum

teamSize

- The average ICO team has 13 members, with the most common team size being 8. The median team size is 12.
- Team sizes vary between 1 and 75 individuals.
- 154 records with missing values for the "teamSize" feature and handle use multiple imputations.
- Mild outliers were found in the "teamSize" data and adjusted using Winsorization.

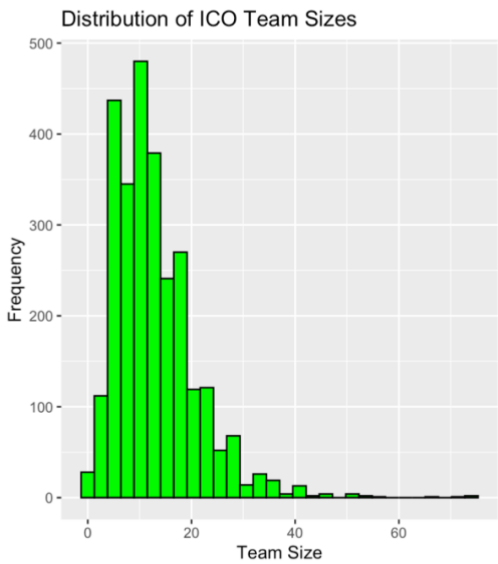


Figure 6 Bar plot of teamSize

distributedPercentage, priceUSD

Preparation

- **distributedPercentage:** Values exceeding 1 were divided by 100 to make within range and remove 266.25, 869.75.
- **priceUSD:**
  - Extreme outliers were addressed by filtering out coins with prices above $4.
  - Right skewed data.

Feature Engineering

- **value_Distributed** (priceUSD * distributedPercentage): A new feature was created to estimate the total value of coins intended for distribution to investors. This feature helps assess the overall scale of ICO fundraising goals.

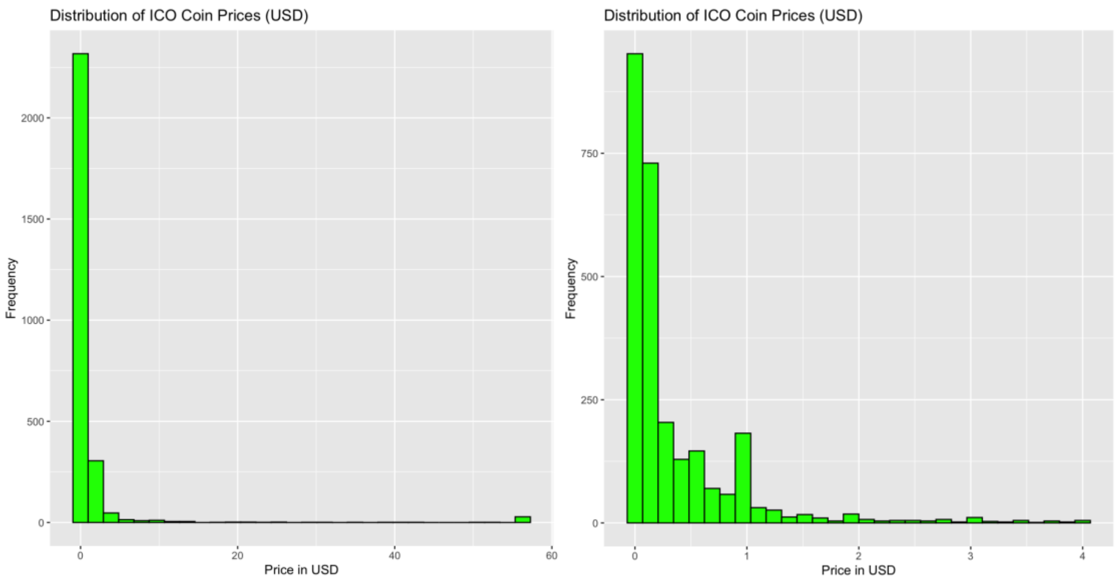


Figure 7 Distribution of PriceUsd before and after

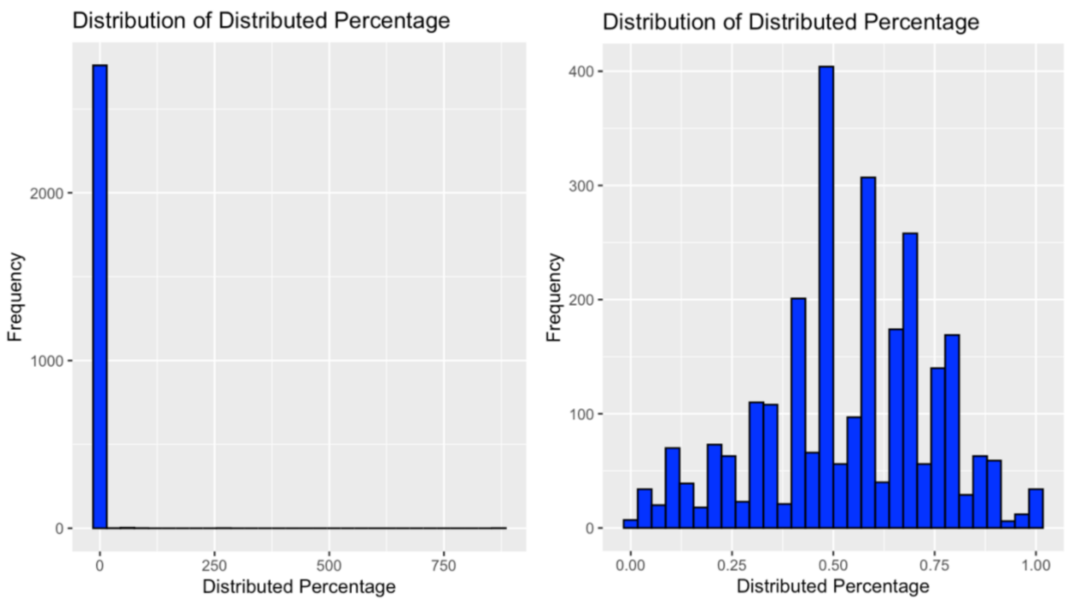


Figure 8 Before and After of pre-processing in distributedpercentage

rating

- The 'rating' feature has no missing values.
- The 'rating' feature shows a slightly right-skewed distribution.
- The mean of the 'rating' feature is 3.13, and the median is 3.1, indicating most ICOs are perceived as having average quality.

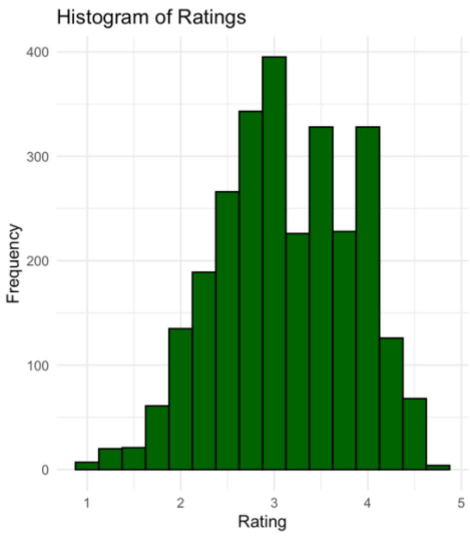


Figure 9 Distribution of Rating

minInvestment

No Preparation and feature engineering is required.

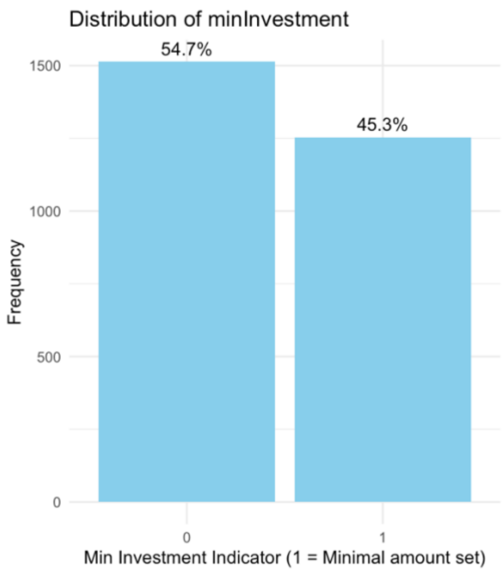


Figure 10 Percentage Distribution of minInvestment

hasVideo, hasGithub, hasReddit

No pre-processing is required.

Feature Engineering

A new feature, **'hassocial'**, was created to indicate whether an ICO has a social media presence, with a value of 1 if the ICO has any social media platforms (video, GitHub, or Reddit) and 0 otherwise.

The creation of the **'is_social'** feature was inspired by the strong correlations observed in Figure 1 between **hasVideo**, **hasGithub**, and **hasReddit** in the correlation plot, suggesting that these social media-related features collectively contribute significant information about an ICO's online engagement presence.

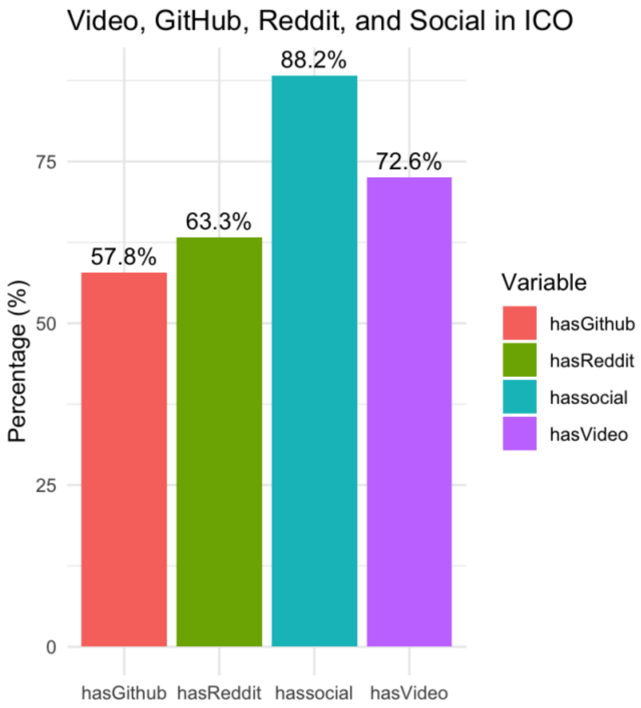


Figure 11 Percentage Distribution

Preparation of Date Features

startDate & endDate

Preparation

The following Preparation steps were taken on the ‘startDate’ and ‘endDate’ features:

- **Date Conversion:** The ‘startDate’ and ‘endDate’ features were converted to the “%d/%m/%Y” format to ensure proper handling of date values.
- **Swapping Date Values:** For ICOs with a ‘startDate’ greater than their ‘endDate’, the date values were swapped. This ensured that the ‘startDate’ represented the earlier date, while the ‘endDate’ represented the later date, maintaining the chronological order of events.

Feature Engineering

- **ICO Duration:** A new variable called ‘duration days’ was created to represent how long the campaign lasted. This was computed by subtracting the start date from the end date.
- **ICO Start and End Dates:** New variables for the start and end dates of the ICO were created to measure timing aspects more precisely.
    1. \`startDay\` and \`startMonth\` were derived from the \`startDate\` to capture the day and month when the ICO campaign begins.
    2. Similarly, \`endDay\` and \`endMonth\` were extracted from the \`endDate\` to note the day and month when the ICO campaign concludes.

These features help in analysing seasonal and temporal effects on ICO success.

Preparation of Text Features

brandsolgan

Preparation

The following Preparation steps were taken on the ‘brandSlogan’ feature:
1. **Corpus Creation**: Constructed from the 'brandSlogan' data.
2. **Lowercasing**: Converted all text to lowercase for uniformity.
3. **Cleaning**: Removed non-letter characters, stopwords, and extra spaces.
4. **Lemmatization**: Simplified words to their root forms.

Feature Engineering

To assess the impact of brand slogans on ICO success, the length of each campaign’s slogan was computed and added as another feature:

- **Sentimental Analysis**: Using the `sentimentr` package in R, sentiment analysis was conducted on brand slogans to classify them: 0 (non-positive) or 1 (positive).
- **sloganLength**: feature, capturing the character count of each brand's slogan, introduces bias by implying that longer slogans correlate with ICO success, potentially skewing predictions.
- **Document-Term Matrix (DTM):** Initially considered using a DTM of brand slogans to capture textual details, this idea was dropped to avoid the curse of dimensionality which could complicate the model unnecessarily.

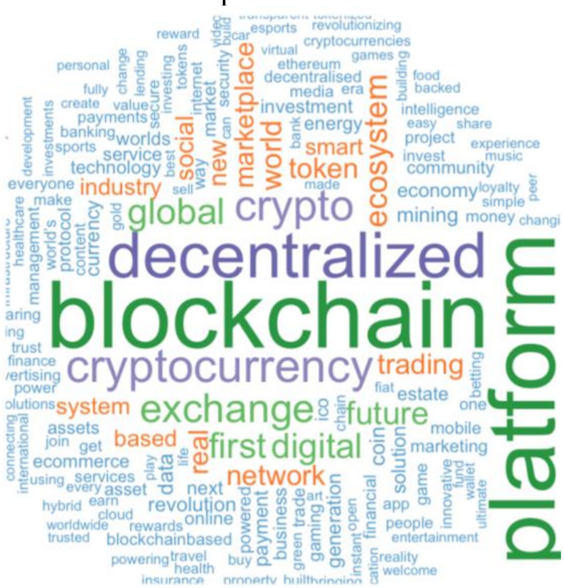


Figure 12 Word Cloud of brandSlogan

External Data Integration

After completing the necessary data Preparation steps for the given dataset, data on GDP, unemployment, and CPI inflation is retrieved from the World Bank using the WDI API (WDI, 2020) for the period from 2010 to 2020. The specific indicators fetched are GDP in current US dollars unemployment rate as a percentage of total labour force, and annual

CPI inflation. Additionally, historical daily mean prices of Bitcoin were obtained from coinmarketcap (Bitcoin, 2023), reflecting the cryptocurrency market trends relevant to ICO performance.

Within the ICO dataset, ISO 3-letter country codes and the year information were derived from the ‘startDate’ to facilitate data merging. The ICO dataset was improved by aligning ICO start dates with the corresponding Bitcoin values on those dates. Subsequently, it was merged with the GDP, unemployment, and CPI datasets using country codes and year as keys. This integration provides a comprehensive view by correlating ICO outcomes with economic indicators and Bitcoin market conditions, offering a deeper understanding of factors that could impact ICO success.

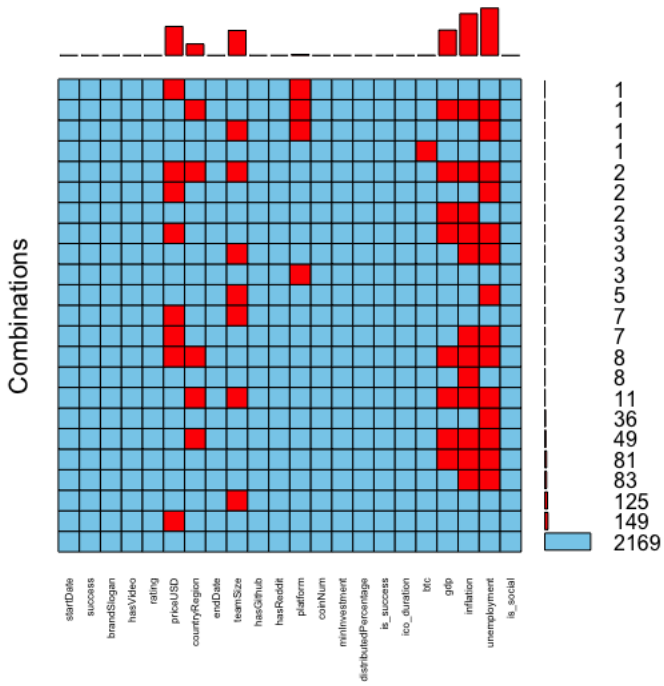


Figure 13 Missing value after External data addition

After integrating external data, we addressed missing values by creating five imputed datasets using the multiple imputation approach. Specifically, we applied the Classification and Regression Trees (CART) method to numerically impute missing values across various features including `hasVideo`, `rating`, and `teamSize`, among others, using the `mice` package in R.

Data Preparation

Now, we'll prepare the data through encoding. Z-score encoding normalizes numerical features, while frequency encoding transforms the 'countryRegion' variable.

Data Encoding

By subtracting the mean and dividing by the standard deviation, Z-score encoding normalizes all **numerical features**, guaranteeing a mean of zero and a variance of one.

This is significant because it helps algorithms converge more quickly and accurately by standardizing the range of numerical input.

Frequency encoding of the '**countryRegion**' categorical variable involves replacing each category with the dataset's frequency or proportion of occurrences. This is significant because it transforms categorical data into a numerical format that models can interpret while preserving the information about the prevalence of each category.

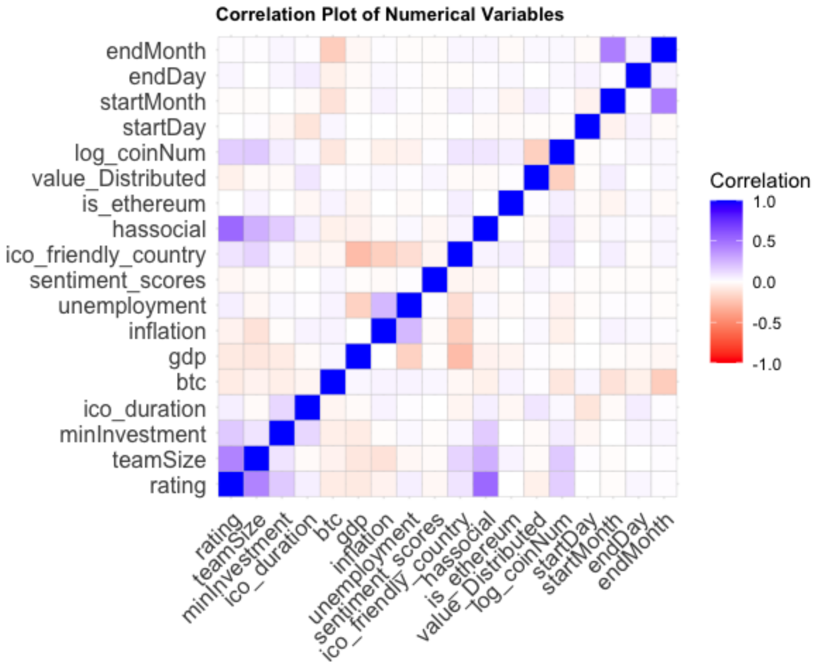


Figure 14 Correlation Plot of a numerical variable.

The correlation plot shows that `**endMonth**` and `**startMonth**` are highly correlated, suggesting redundancy. To reduce multicollinearity, `**endMonth**` will be dropped.

Train Test Split

To improve model accuracy and prevent overfitting, drop the following features from the dataset: startDate, endDate, is_success, hasVideo, hasGithub, hasReddit, coinNum, priceUSD, distributedPercentage, and platform. These features are considered irrelevant to the prediction task.

Then, Split the dataset into training and testing sets in an **80-20 ratio** to ensure sufficient data for training while also providing a robust set for model validation.

Feature Selection

Applied a Random Forest model using the Gini impurity criterion (**min impurity decrease**) to assess and rank the importance of features. This method improves feature selection by ensuring that only large divides based on Gini impurity are considered.

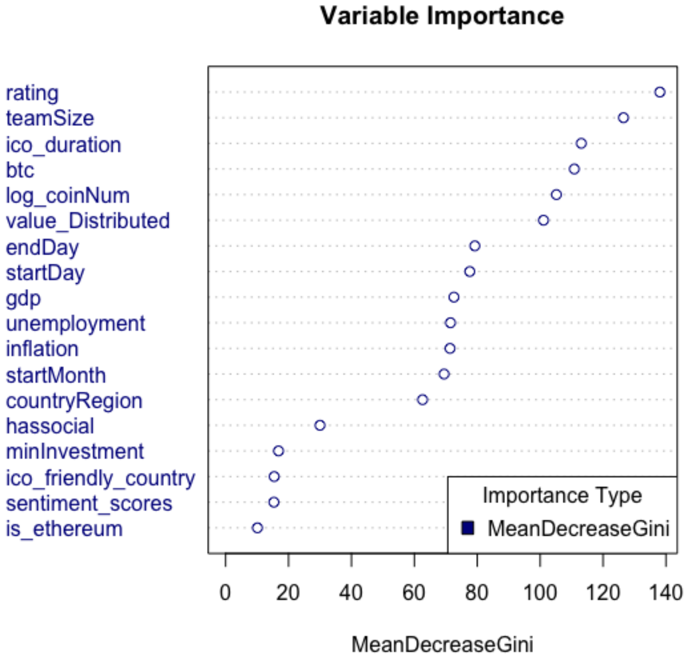


Figure 15 Variable importance.

We will use the 95% cumulative frequency cut-off from the Random Forest importance scores to select the top 14 features for final modelling and evaluation purposes.

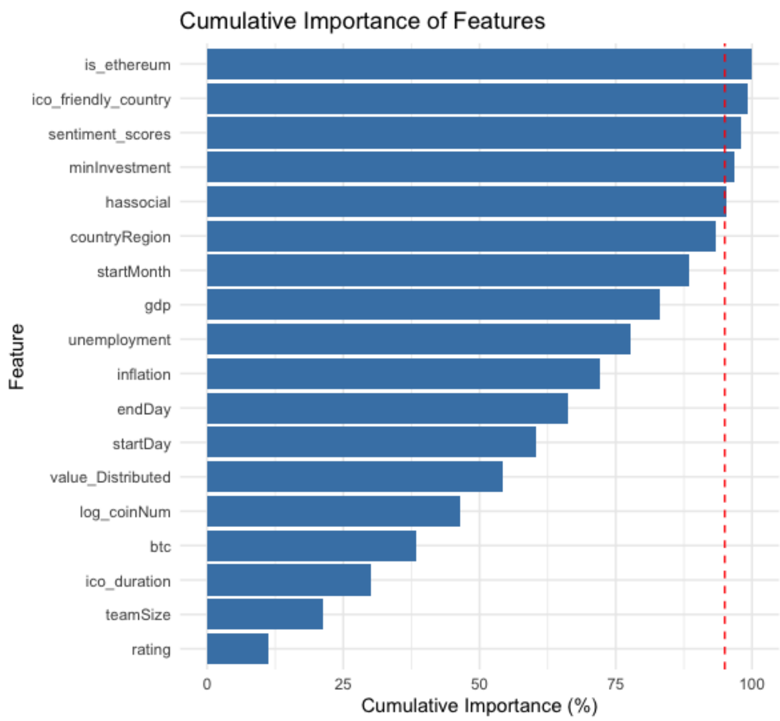


Figure 16 Cumulative Importance

Top features: rating, teamSize, ico duration, btc (bitcoin price), log_coinNum, value_Distribution, endday, startday, gdp, inflation, unemployment, countryRegion, startMonth, countryRegion, hassocial.

Modelling

Now proceed with using the top 14 features for our modelling. These features have been selected based on their importance and relevance. Now, we'll incorporate them into our model.

Logistic Regression

Logistic regression is crucial for predicting ICO success as it effectively handles binary outcomes, providing probabilities that help assess the likelihood of achieving fundraising goals based on various predictor variables (Hosmer et al., 2013).

After **Hyperparameter** tuning using 4-fold cross-validation identified optimal values: `best_alpha = 0.6` and `best_lambda = 0.02511886`.

```
Coefficients:
                    Estimate Std. Error z value Pr(>|z|)
(Intercept)       -2.538e+00  5.550e-01  -4.573 4.81e-06 ***
rating             6.801e-01  8.567e-02   7.938 2.06e-15 ***
teamSize           3.517e-02  6.580e-03   5.344 9.08e-08 ***
ico_duration      -2.731e-03  7.373e-04  -3.704 0.000212 ***
btc                4.970e-05  1.811e-05   2.745 0.006059 **
log_coinNum       -2.396e-02  2.358e-02  -1.016 0.309562
value_Distributed -3.021e-05  1.265e-04  -0.239 0.811255
startDay           1.135e-02  4.949e-03   2.294 0.021784 *
endDay            -9.909e-03  4.744e-03  -2.089 0.036724 *
inflation         -2.024e-02  2.658e-02  -0.761 0.446463
unemployment      -8.431e-03  1.659e-02  -0.508 0.611337
gdp               -3.485e-14  9.475e-15  -3.678 0.000235 ***
startMonth        -6.132e-02  1.431e-02  -4.285 1.83e-05 ***
countryRegion      3.125e+00  1.215e+00   2.571 0.010127 *
hassocial          2.858e-01  1.105e-01   2.586 0.009705 **
```

Figure 17 Logistic regression summary.

From Figure 16 logistic regression output indicates that variables such as **`rating`**, **`teamSize`**, **`ico_duration`**, **`btc`**, and **`gdp`** significantly impact the model ($p < 0.05$), with each coefficient estimate showing the effect size and direction on the likelihood of the outcome variable, adjusting for other factors in the model. The model's AIC score is 2661.6, indicating effectiveness in fitting the data.

	Accuracy	Precision	Recall	F1 Score	AUC
Logistic Regression	69.92	73.60	84.21	78.55	69.80
Logistic Regression Optimised	71.10	72.78	88.91	80.04	70.47

Figure 18 Performance metric of Logistic Regression

Decision Trees

The C5.0 decision tree algorithm is critical for predicting ICO success because it can efficiently handle large datasets and provide interpretable results that highlight key decision points influencing outcomes (Salzberg, 1994).

Hyperparameter tuning of the C5.0 decision tree algorithm using 9-fold cross-validation determined the optimal number of trials to be 11.

	Accuracy	Precision	Recall	F1 Score	AUC
C5.0	67.93	75.41	75.62	75.51	67.06
C5.0 Optimised	69.02	72.61	84.48	78.10	69.44

Figure 19 Performance metric of Decision tree.

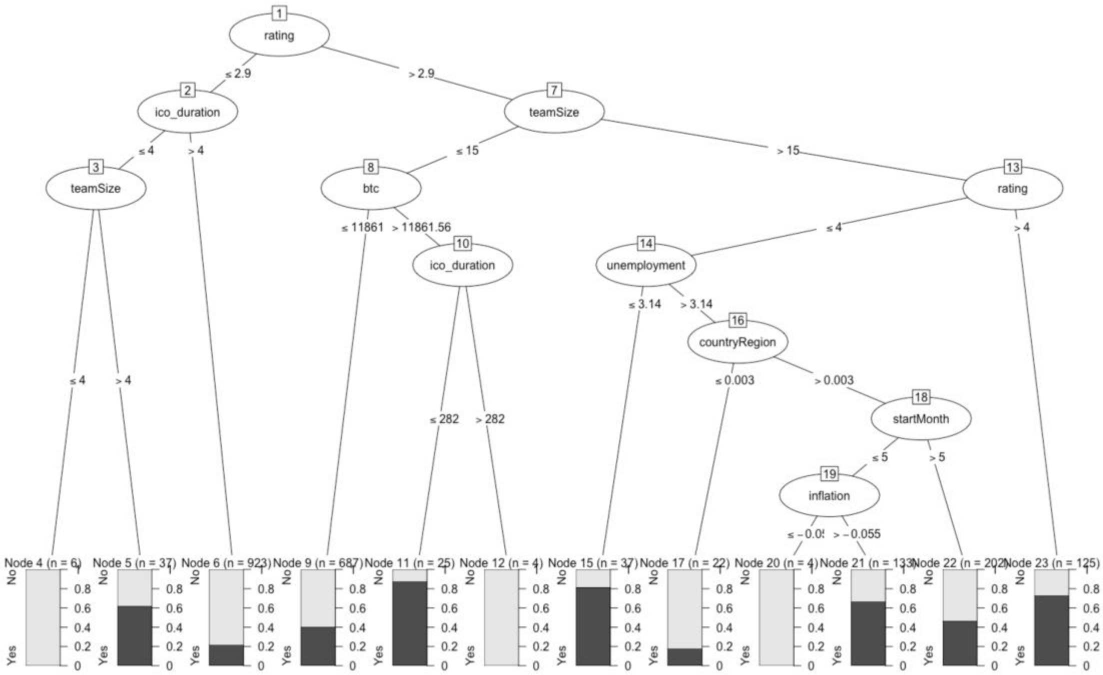


Figure 20 C5.0 output.

Primary splits on `rating` and further on `teamSize`, `btc`, and `ico_duration`, among others. Key findings include:

- Lower ratings combined with longer ICO durations typically predict failure.
- Higher ratings with smaller teams predict failure unless coupled with high Bitcoin prices and shorter ICO durations.
- Larger teams generally predict success, particularly with higher ratings.
- The tree has a 30.9% error rate and utilizes `rating` and `teamSize` as its most influential attributes for predicting ICO success.

Random Forest

Random Forest is essential for predicting ICO success as it combines multiple decision trees to reduce overfitting and enhance predictive accuracy, making it highly effective for complex datasets with numerous variables (Breiman, 2001).

Hyperparameter tuning using 7-fold cross-validation identified the optimal number of mtry = 256 and inbag = 47 (square root of #train set).

	Accuracy	Precision	Recall	F1 Score	AUC
Random Forest	67.93	42.40	54.72	47.78	67.01
Random Forest Optimised	69.56	43.97	57.93	50.00	70.12

Figure 21 Performance metric of Random Forest.

Support Vector Machine (SVM)

Support Vector Machines (SVM) are pivotal for predicting ICO success due to their ability to effectively handle high-dimensional data and maximize the margin between decision boundaries, ensuring robust classification even in complex datasets (Cortes & Vapnik, 1995).

For SVM **hyperparameter tuning** using 5-fold cross-validation, the optimal kernel identified was `polydot`.

	Accuracy	Precision	Recall	F1 Score	AUC
SVM	70.65	42.40	60.90	50.10	70.27
SVM Optimised	71.19	42.40	62.30	50.74	72.98

Figure 22 Performance metric of SVM.

Artificial Neural Network (ANN)

Artificial Neural Networks (ANN) are crucial for predicting ICO success due to their ability to model complex nonlinear relationships and patterns in large datasets, which enhances the predictive accuracy and generalization over traditional methods (Haykin, 1999).

Hyperparameter tuning of the ANN algorithm, conducted using 5-fold cross-validation, determined that the optimal configuration consists of two hidden layers with 1 and 4 nodes respectively, with decay =0.0001 This configuration effectively balances model complexity and performance.

	Accuracy	Precision	Recall	F1 Score	AUC
ANN	65.53	44.40	59.02	50.74	66.93
ANN Optimised	72.01	73.38	47.92	57.97	74.17

Figure 23 Performance metric of ANN.

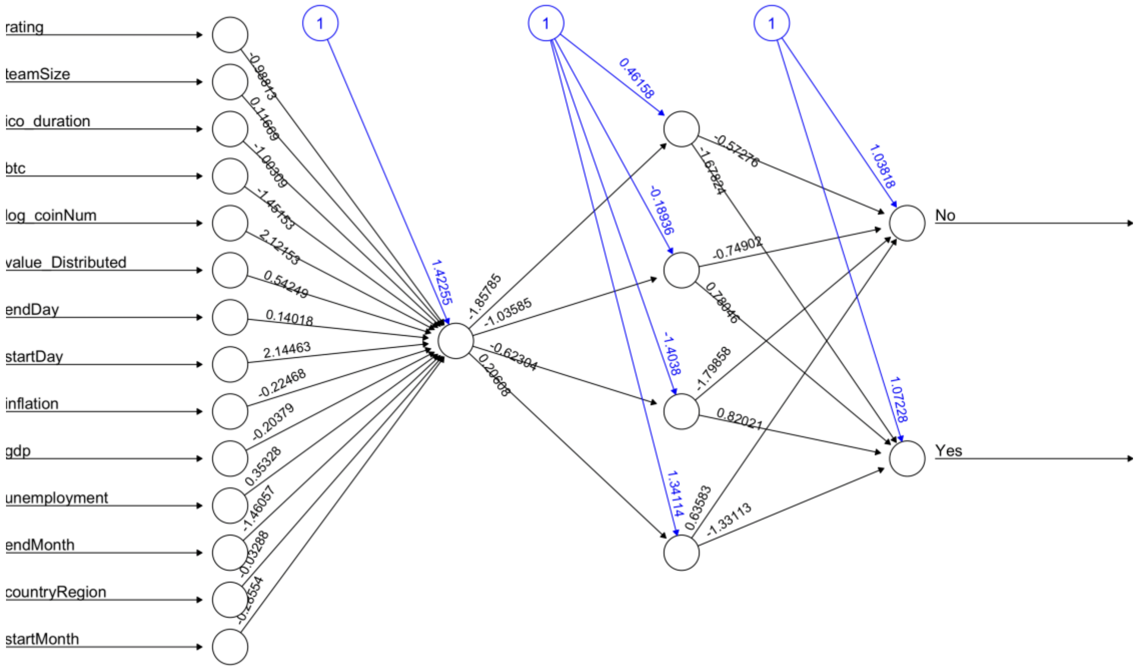


Figure 24 ANN output.

Figure 23 illustrates how the ANN utilizes the universal approximation theorem to converge predictions. The network can approximate any continuous function through this theorem, capturing complex relationships within the data. The activation function, a sigmoid given the binary classification task (indicated by the "Yes" and "No" outputs), enables the network to handle non-linearities effectively.

Conclusion

Performance Metrics Selection

1. **For Fundraising Teams**: **Precision** should be prioritized to minimize resource misallocation by reducing the number of unsuccessful ICOs inaccurately predicted to succeed, thereby maximizing resource efficiency (Alchykava & Yakushkina, 2021).
2. **For Investors**: **Recall** should be prioritized to minimize the risk of missing out on potentially successful ICOs by reducing the number of successful ICOs incorrectly predicted to fail, thereby maximizing investment opportunities. Additionally, the **AUC** should be emphasized to assess the overall effectiveness of the model at various threshold levels, ensuring that investors can accurately gauge the probability of ICO success (Rasskazova et al., 2019; Alchykava & Yakushkina, 2021).

Both groups could also consider:

- **Accuracy** for a general sense of model performance across all predictions.
- **F1 Score**, which balances as a harmonic mean of precision and recall, is useful when looking for a single metric that reflects both the cost of false positives and the cost of false negatives.

Model Selection

1. **Fundraising Teams**: Would benefit from **Logistic Regression and Decision Tree C5.0** due to their ability to balance accuracy and interpretability, helping teams understand what factors contribute to ICO success (Alchykava & Yakushkina, 2021).
2. **Investors**: Prefer **SVM**, **Random Forest** (ensemble model) and **ANN** are favoured for their high predictive accuracy, crucial for accurately assessing project success and potential returns. However, these models come with trade-offs such as higher training costs and lower interpretability, which can complicate understanding and explaining the decision-making process (Rasskazova et al., 2019; Alchykava & Yakushkina, 2021).

Comparative Analysis

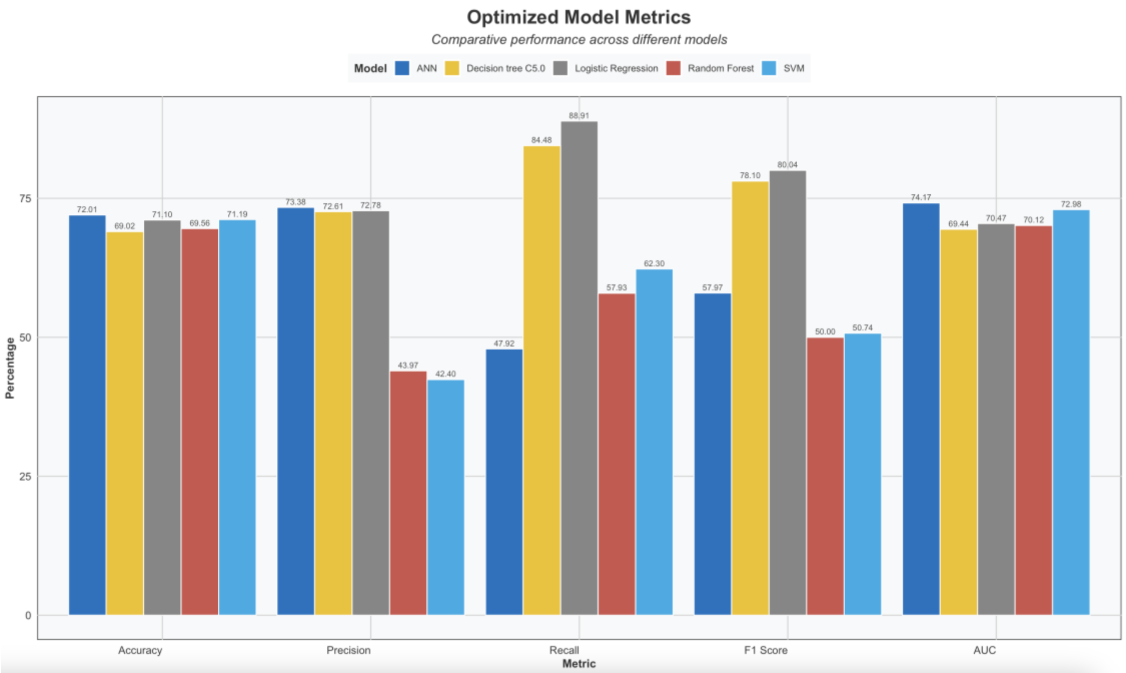


Figure 25 Comparative Analysis plot.

For fundraising teams and investors based on the optimized models, Logistic Regression and Decision Tree C5.0 emerge as the top choices for fundraising teams due to their higher Precision (72.78% and 72.61% respectively) and Recall (88.91% and 84.48% respectively), ensuring efficient resource allocation and reliable identification of successful ICOs. For investors, the ANN model, with the highest AUC at 74.17%, offers superior discrimination between successful and unsuccessful ICOs, despite its lower Recall (47.92%). SVM also presents a feasible option for investors with the second-highest AUC (72.98%), indicating robustness in handling complex datasets.

Recommendation

Fundraising teams should incorporate Decision Tree and Logistic Regression models into their ICObench platform, utilizing Precision as the key metric for effective deployment. This approach will refine decision-making and adjust dynamically to changing market conditions. Meanwhile, investors are advised to deploy ANN, focusing on the AUC score to gauge model effectiveness. By aligning these specific models and metrics with their respective goals, both fundraising teams and investors can significantly boost their ability to predict and capitalize on ICO success.

References

1. Bitcoin 2023. Bitcoin. *CoinMarketCap*. [Online]. [Accessed 2 May 2024]. Available from: https://coinmarketcap.com/currencies/bitcoin/historical-data/.

2. Breiman, L. 2001. *Machine learning*. **45**(1), pp.5–32.

3. Cortes, C. and Vladimir Vapnik 1995. Support-vector networks. *Machine learning*. **20**(3), pp.273–297.

4. Fisch, C. 2019. Initial coin offerings (ICOs) to finance new ventures. *Journal of business venturing*. **34**(1), pp.1–22.

5. Haykin, S.S. 1999. *Neural Networks*. Upper Saddle River, N.J. : Prentice Hall.

6. Hosmer, D.W., Lemeshow, S. and Sturdivant, R.X. 2013. *Applied Logistic Regression* [Online]. Wiley. [Accessed 8 May 2024]. Available from: https://onlinelibrary.wiley.com/doi/book/10.1002/9781118548387.

7. Huang, W., Silvio Vismara and Wei, X. 2022. Confidence and capital raising. *Journal of corporate finance*. **77**, pp.101900–101900.

8. Maryia Alchykava and T. Yakushkina 2021. ICO Performance: Analysis of Success Factors. *Intelligent Memory Systems*. [Online]. [Accessed 24 April 2024]. Available from: https://www.semanticscholar.org/paper/ICO-Performance%3A-Analysis-of-Success-Factors-Alchykava-Yakushkina/c57cd36d46ca2f5fe06008c691fa30aac7257c6e.

9. Offshore Protection 2023. Top 9 countries for ICO | Eternity Law International. *Eternity Law International* [Accessed 26 April 2024]. Available from: https://www.eternitylaw.com/news/top-9-countries-for-ico/.

10. Salzberg, S.L. 1994. C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993. *Machine learning*. **16**(3), pp.235–240.

11. WDI 2020. World Development Indicators | DataBank. *Worldbank.org*. [Online]. [Accessed 1 May 2024]. Available from: https://databank.worldbank.org/source/world-development-indicators.