M&A Takeover Success Classification For Merger Arbitrage

Jerry Loh, Wilson Widyahana, Isabel Teo, Tan Jing Jie, Isabel Teo, Lee Chin Juan

National University of Singapore
NUS FinTech Society

Abstract

In this project, we attempt to train and test a binary classifier to predict the deal status of announced mergers and acquisitions (M&A) on publicly-listed US equities. Given price, fundamental and financial data on merger targets and acquirors at the time of announcement, we attempt to predict the final status of the announced deal, and take the long position for instances where a success is predicted, while avoiding mergers that are unlikely to complete.

Introduction

We first clarify the idea of merger arbitrage. When a company decides to acquire a publicly-listed company, it typically announces its intention to purchase shares of the target company at a market premium. This represents an arbitrage opportunity: If the acquiror is willing to purchase shares of the target at a premium, then purchasing shares at any price lower than the announced price results in profit, because the share price will eventually converge to the acquisition price by the takeover time.

However, while the price of the target stock typically runs up close to the announced price post-announcement, the market price does not reach exactly the announced price. This is because mergers can fail to complete even after announcement: The market expresses this deal-break probability via the difference between the price the target trades at and the deal price.

As such, more uncertain deals typically manifest themselves via a larger difference between the post-announcement and acquisition price. Given this idea, if we are able to on average predict takeover success (and hence price the merger) more accurately than the market, we will be able to generate profit (modulo of any transaction costs).

This is done by predicting the status of M&A announcements, and trading the mean-reversion factor if our model predicts a success.

Related Works

Karatas and Hirsa's *Predicting Status of Pre and Post M&A Deals Using Machine Learning and Deep Learning Techniques* explore important features for takeover success prediction, namely firm and deal information, which we would utimately use in the feature selection process. They also explore and compare the performance of classification methods with nonlinear decision boundaries beyond the typical logit and probit models found in older literature, which guided parts of our methodology.

Data

In this project, we used data from Refinitiv's SDC Platinum M&A database, which provided us coverage of 3532 unique M&A announcements over the last 10 years, with transaction values over $1 million and publicly-listed target companies. Post-cleaning, we were left with 107 features, consting of financial ratios, fundamental and price data of both target and acquiror firms wherever available. Our in-sample data spans 2012/01/01 to 2019/01/01, and our out-of-sample data spans 2019/01/01 to 2022/01/01.

In the data processing stage, we drop features with greater than 50% missing values, and impute missing data with a k-Nearest Neighbours algorithm. The choice of neighbours was selected after examining the distributions of our features.

Exploratory Analysis

We analyse and compare the duration from merger announcement to merger completion or failure: In the training set, we found that the median days to successful mergers was 130, while the median days to merger withdrawals was 114. However, the variance of the distribution of merger success time was much lower than that of failures, and the 75 percentile time taken for merger success was at 199 days.

This provides us the motivation for our exit strategy: As more time elapses from the announcement day, the probability of merger success must be revised lower. As a typical merger arbitrage strategy is leveraged, holding such a position for a long time is sub-optimal. As such, we consider imposing a time limit for any position to be held.

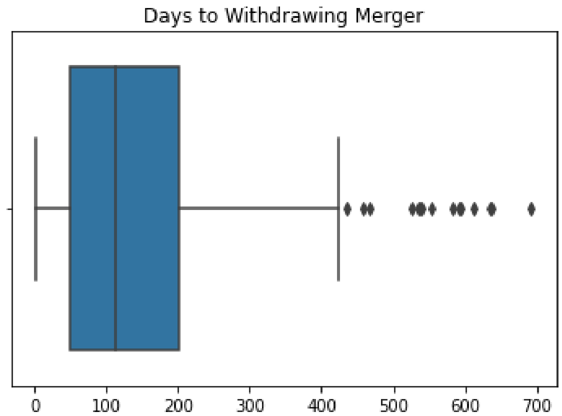

Figure 1: Boxplot of days from announcement to merger failure

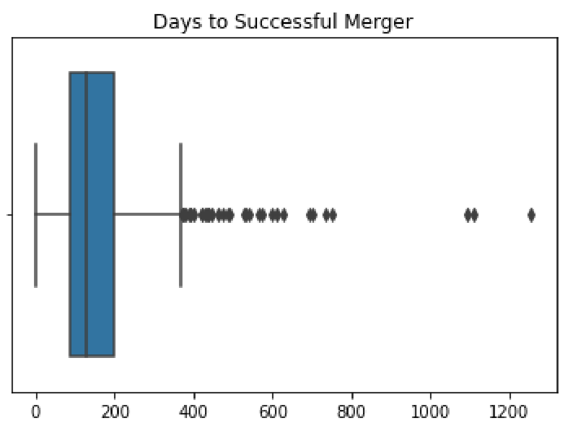

Figure 2: Boxplot of days from announcement to merger success

We also observe that our dataset is positively skewed towards successful mergers. This is because the majority of announced takeovers result in success, giving us an imbalanced dataset with 88.1% of entries being successful mergers. This affects our choice of performance metrics to optimise for: Due to this imbalance, the merger arbitrage strategy is one with strongly negatively skewed returns. It is thus wise to minimise the false positive rate of our predictive model. We also consider the usage of synthetic data creation methods, such as SMOTE, GAN or VAEs in the training step.

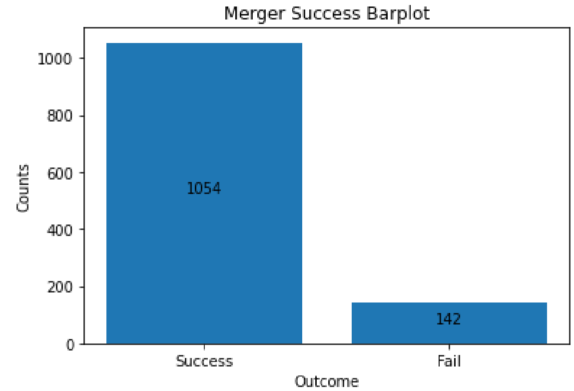

Figure 3: Frequency of successful and withdrawn mergers in training set

Feature Selection

We apply feature selection techniques, namely the chi-squared test, extra-trees classifier and mutual information regression. Similar to the prevailing literature, we found that factors relating to the financial health of target companies were most important to predicting merger success.

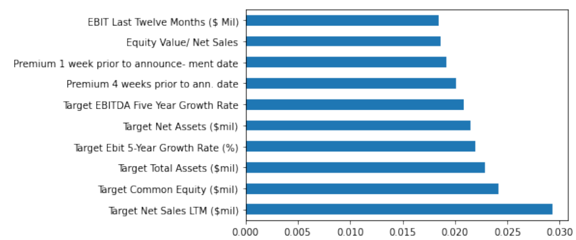

Figure 4: Extra-trees classifier results

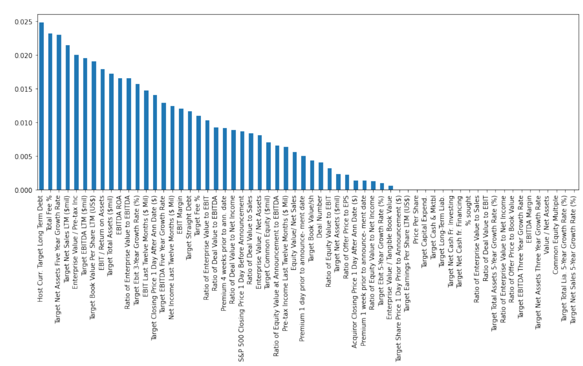

Figure 5: Mutual information feature selection results

Implementation

Given the highly imbalanced dataset, we measure the performance of our trained models against a dummy classifier which predicts merger success all the time. On the test set, the dummy classifier produced an accuracy of 88.1% (since

88.1% of mergers are successful), with a false positive rate of 100%. As mentioned in previous sections, we aim to minimise the false positive rate of our model.

We first explore logistic regression models, which produced slightly higher accuracy, as well as 77.4% false positive rate. Our 1-nearest-neighbours classifier had a 77.4% false positive rate. Optimising hyper-parameters using grid search for precision, out support vector and decision tree models produced lower 0.65% and 58.1% false positive rates.

Based on the above results, we selected a majority-voting classifier with SVM, 2-nearest-neighbours, logistic and decision tree classifiers. This produced the best result in the test set, as well as the unseen out-of-sample data, with higher precision (92.3% vs 88.1%) and lower false positive rates (58.3% vs 100%) compared to the dummy classifier.

We map these predictions on the out-of-sample data, and proceed with our backtest.

```
----True OOS Performance Classifier Performance----
Accuracy:  0.877
Precision:  0.923
Recall:  0.94
ROC/AUC:  0.678
False Positive Rate:  0.583
Confusion Matrix:
 [[ 15  21]
 [ 16 250]]

----Dummy (Always predict true) OOS Performance Classifier Performance----
Accuracy:  0.881
Precision:  0.881
Recall:  1.0
ROC/AUC:  0.5
False Positive Rate:  1.0
Confusion Matrix:
 [[  0  36]
 [  0 266]]
```

Figure 6: Our model performance vs dummy classifier on out-of-sample data

Backtesting

Since takeover announcements are discrete events that are unequally-spaced in time, it makes little sense to compute backtest performance or total P/L over time. Instead, we compare the return distributions of each event in the out-of-sample data for our strategy against the dummy strategy.

We choose a cut-off date of 60 days for any open position to be automatically closed regardless of future merger success due to the holding period leverage concerns highlighted previously.

We first obtain the performance of the benchmark dummy strategy of going long all announced mergers: The annualised mean total return of successful and withdrawn mergers was 2.0% and 0.26% respectively, with a combined average of 0.42%.

Using our model's prediction, our annualised mean total return of successful and withdrawn mergers was 2.22% and 0.26% respectively, with a combined average of 0.43%. This is a only a marginal improvement over the dummy model, partly because both models will go long on most announced takeovers. However, our model's ability to better detect false positives likely resulted in higher mean returns for successful mergers.

```
Model Performance:
Success mean Return: 0.02239
Failure mean Return: 0.00358
Combined mean Return: 0.00532
```

Figure 7: Model annualised mean return performance

```
Dummy Performance:
Success mean Return: 0.01997
Failure mean Return: 0.00331
Combined mean Return: 0.00495
```

Figure 8: Dummy annualised mean return performance

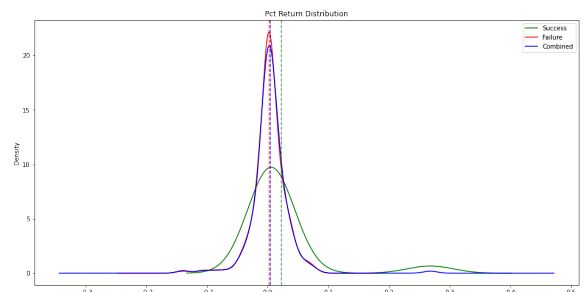

Figure 9: Return distribution for strategy using model predictions

Conclusions

Even with some ability to detect false positives, the market is already extremely efficient in pricing announced takeovers. As such, it is difficult to significantly improve upon a naive ”long all mergers” strategy.