Piotroski Score Analysis

Min Yi Li | Zheng Gong
Qiuyi Guan | Huansha You

Outline

1. What is Piotroski F-Score
2. Model Implementation
    a. Compute Piotroski
    b. Data Exploration
    c. Random Forest Classification Model
    d. Model Performance Analysis
3. Improvements

Piotroski F-Score

Piotroski F-Score

- The Piotroski score is a discrete score between 0-9 that reflects nine criteria used to determine the strength of a firm's financial position
- It is a very popular metric used to judge value stocks
- If a company has a score of 8 or 9, it is considered a good value. If the score adds up to between 0-2 points, the stock is considered weak

Piotroski's Indicators

Profitability - Prefer stocks with strong value and good cash flows

1. Positive Net Income
2. Positive return on assets in the current year
3. Positive operating cash flow in the current year
4. Cash flow from operations being greater than net Income

Leverage, Liquidity and Source of Funds - Prefer companies with low long-term debt and better ability to pay off debt

5. Lower ratio of long term debt in the current period, compared to the previous year
6. Higher current ratio this year compared to the previous year
7. No new shares were issued in the last year

Operating Efficiency - The ability of a company to increase its prices

8. A higher gross margin compared to the previous year
9. A higher asset turnover ratio compared to the previous year

Project Implementation

Two Methods to Compute Piotroski

	Attributes used	Data sources
Original	Post-processed financial ratios	• Home/WRDS/Subscriptions/FinancialRatios SuitebyWRDS/FinancialRatiosFirmLevel
New	True values using raw fundamental data	• Home/Get Data/CRSP/Annual Update/CRSP/Compustat Merged • Home/Get Data/WRDS Analytics/WRDS Analytics/Financial Ratios Suite by WRDS/Financial Ratios/ Firm Level by WRDS (Beta)

Raw Data Overview

Calculate Piotroski score using **financial ratios** (original):

gpm	roa	roe	debt_invcap	curr_debt	ocf_lct	lt_ppent	dltt_be	debt_assets	de_ratio	curr_ratio	at_turn
0.353	0.214	0.123	0.001	0.675	1.110	0.903	0.001	0.202	0.253	2.872	1.362
0.351	0.209	0.120	0.001	0.672	1.058	0.894	0.001	0.204	0.256	2.810	1.340

Calculate Piotroski score using **raw fundamental data** (new):

actq	atq	cogsq	cshoq	dlttq	lctq	saleq	niy	oancfy	roa	roe	at_turn
281.308	709.313	148.700	18.689	1.334	90.171	225.008	25.893	31.228	0.214	0.123	1.362
333.708	742.926	178.970	18.692	1.265	105.569	278.724	50.355	68.606	0.209	0.120	1.340

Quarterly

Monthly

Modified Code for New Method

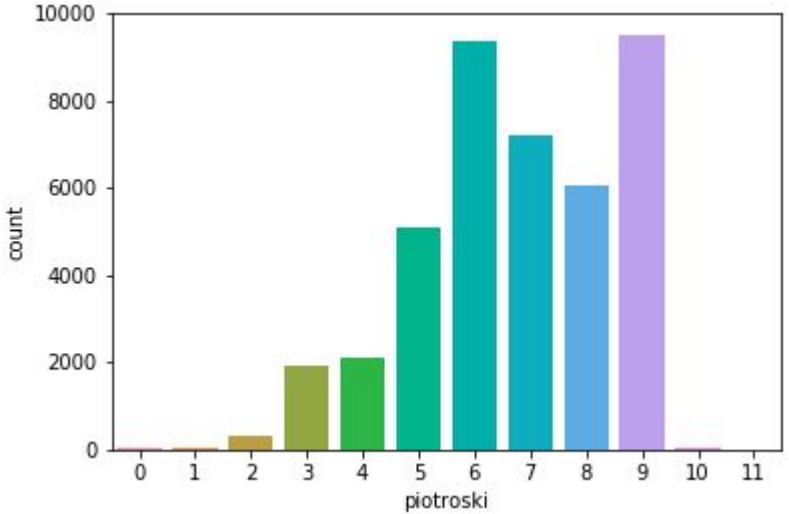
1. $NIY > 0$
2. $ROA > 0$
3. $OANCFY > 0$
4. $OANCFY > NIY$
5. $\left[\frac{DLTTQ}{ATQ}\right](t) < \left[\frac{DLTTQ}{ATQ}\right](t-4)$
6. $\left[\frac{ACTQ}{LCTQ}\right](t) < \left[\frac{ACTQ}{LCTQ}\right](t-4)$
7. $CSHOQ(t) \leq CSHOQ(t-4)$
8. $[SALEQ - COGSQ](t) > [SALEQ - COGSQ](t-4)$
9. $at\_turn(t) > at\_turn(t-12)$

```
n = 4
df1['piotroski_curr'] = np.where((df1.permno == df1.permno.shift(n)) &
                                 ((df1.actq/df1.lctq)>(df1.actq.shift(n)/df1.lctq.shift(n))), 1, 0)
```

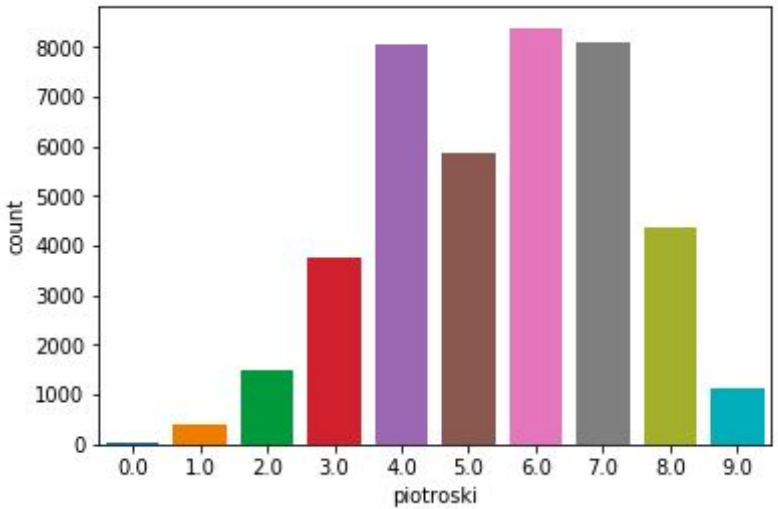
```
n=12
df2['piotroski_at_turn'] = np.where((df2.permno == df2.permno.shift(n))&
                                    (df2.at_turn>df2.at_turn.shift(n)), 1, 0)
```

Data Exploration - Original vs. New

Group by each entries Piotroski Value



Original



New

Random Forest Classification

Create Random Forest Classification model to predict daily return of SPX

- 0 for negative return
- 1 for positive return

Features:

- Log of Piotroski Score (PIO)
- Log of Return on Equity (ROE)

Data Range: 2015-01 to 2018-12

Stocks:

- IJR ETF - Small Cap
- IJH ETF - Mid Cap

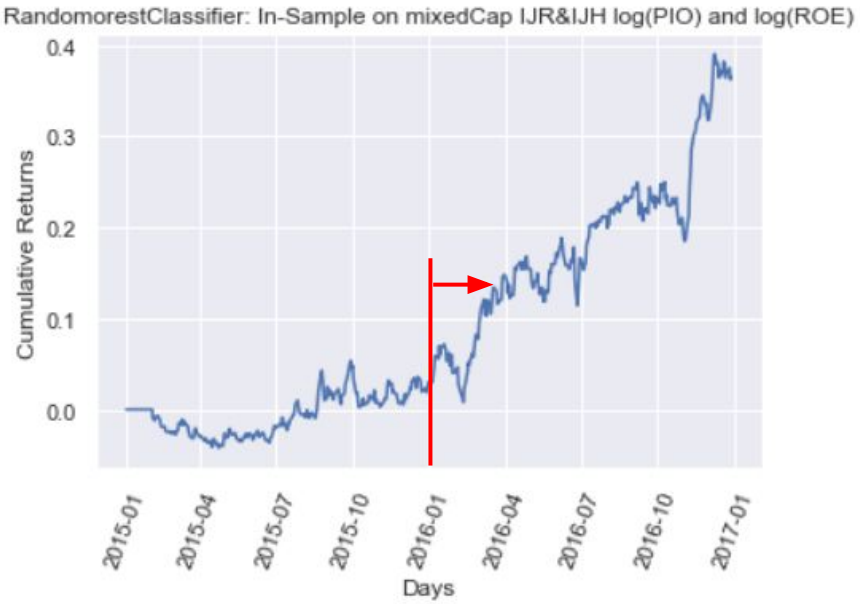
Feature Importance

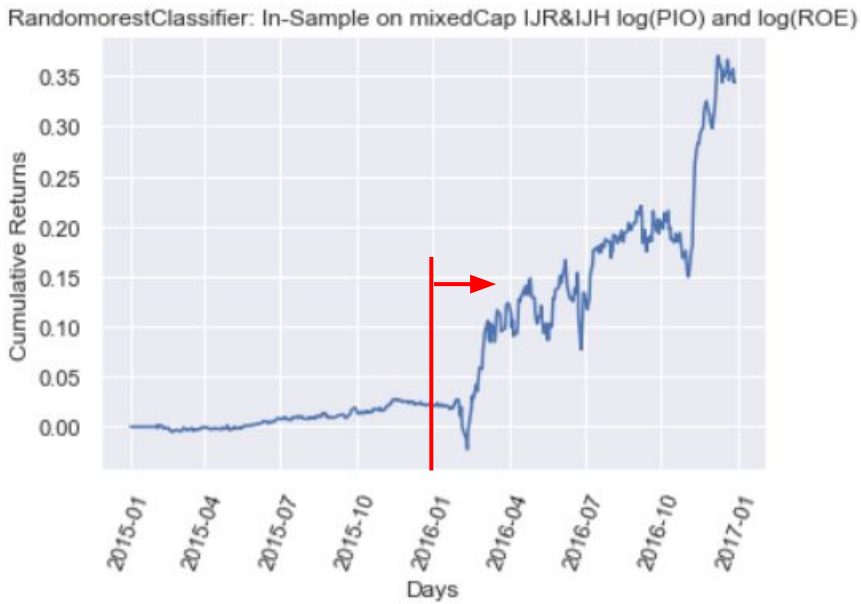
	Log PIO	Log ROE
Original	41%	59%
New	74%	26%

- Similar importance in original method with slightly more depending on ROE
- PIO is much more important in new method

Model Performance (In-Sample)

Model performance: 56% training accuracy

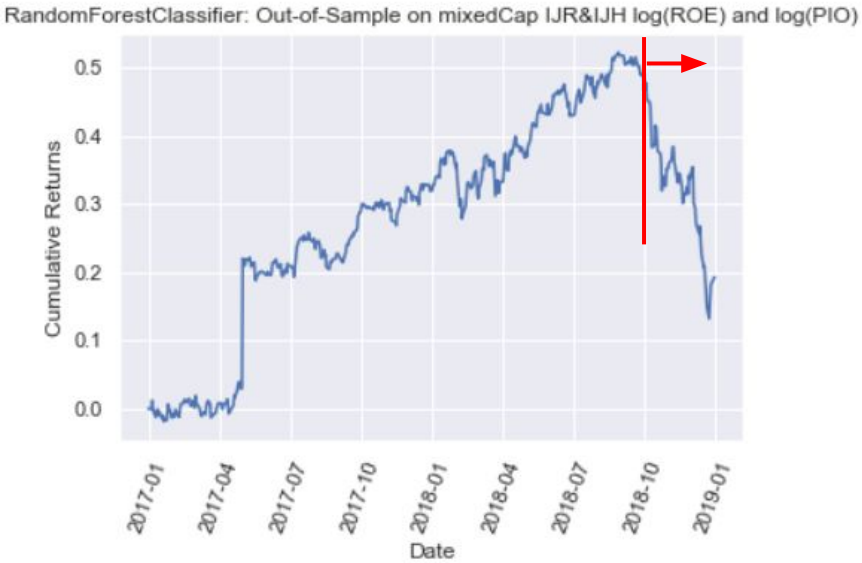


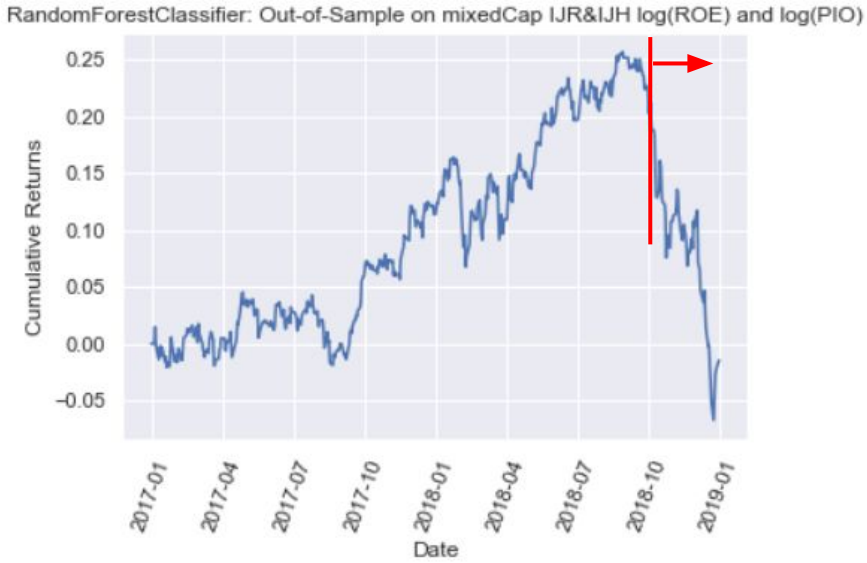


Similar trend since 2016; flat in 2015 because of shift effect when calculating Piotroski

Model Performance (Out-of-Sample)

Model performance: 48% testing accuracy





Similar decreasing trend since 2018 but new method drops below zero

Performance Analysis and Improvements

Performance Analysis

In General:

1. Data set limitations
2. Combined small and medium cap stocks
3. Combined sectors
4. Only two features were leveraged for Random Forest Classification, not realizing it's full potential

Reasons for low performance (new):

5. Monthly and quarterly data were combined during implementation
6. Used backfill and forward fill for the NaN values

Reasons for low performance (original):

7. Ratio cross-calculations and lack of some indicators in original method

Improvements

- Add more historical data
- Add more predictive factors as features
- Tuning the parameters of Random Forest Model
- Separate small and medium cap stocks
- Compare average values of Piotroski score in different sectors (e.g. manufacturing, finance, etc.)
- Implement with other cross-sectional models such as logistic regression

References

- https://www.investopedia.com/terms/p/piotroski-score.asp
- https://www.moneycrashers.com/piotroski-f-score-value-investing-model-stocks/
- https://economictimes.indiatimes.com/wealth/invest/the-piotroski-score-what-this-method-for-screening-stocks-tells-us-and-how-it-is-calculated/articleshow/53020837.cms

Thank you!

Q&A

Next Step

- Because of the shift effect, the Piotroski score in 2015 were underestimated
- Add one more year data (2014) to avoid shift effect
- See whether the overall cumulative returns could perform better