Doubly predictions: “JUSTness” and returns

Jun Zhai

February 3, 2018

Abstract

In the spirit of turning data science for social good, I propose a project that builds (doubly) predictions of company's JUST scores and market returns in both directions. Specifically, given a time series of market returns, we predict the JUST scores using neural network by extracting the stochastic features of the market returns; given a JUST score, we use a transition function to predict the time series of market returns as a stochastic process.

1 Introduction

Being a business of social good is often considered as altruism associated with financial loss. JUST Capital, a nonprofit organization, collects information and rankings to quantify the extent to which corporations perform on issues of most importance to the public as well as their associated profits. Their data resource offers us a great opportunity to crack the shell of the connections between being a just company and market returns.

What have we known? Based on the previous research by JUST Capital, the following conclusions have been drawn:

- JUST companies exhibit higher return and lower risk
- Worker treatment, leadership and ethics are the key stock drivers
- The largest companies are NOT the only ones that can afford justness, despite a very moderate size bias toward larger companies

The conclusions above are sourced from "What Drivers of Corporate Responsibility Generate Alpha?" by H. Cortina, October 2017; "Does company size drive JUST Capital?s ranking?" by H. Cortina, March 2017.

Motivation Often intuitively, we link social good to business sacrifice: the more "JUST" a company becomes, the more loss of business it has to face. Is this really true? In this project, we aim to use data to unveil the truth.

Data We obtained data for the Corporate Performance Scores which has 47 components, and the time series of the associated market returns over 767 periods (after cleanup) corresponding to 896 companies from JUST Capital.

2 Methodology

The goal of this project is to make predictions of JUST scores and market returns in both directions; that is, given either of them, we predict the other.

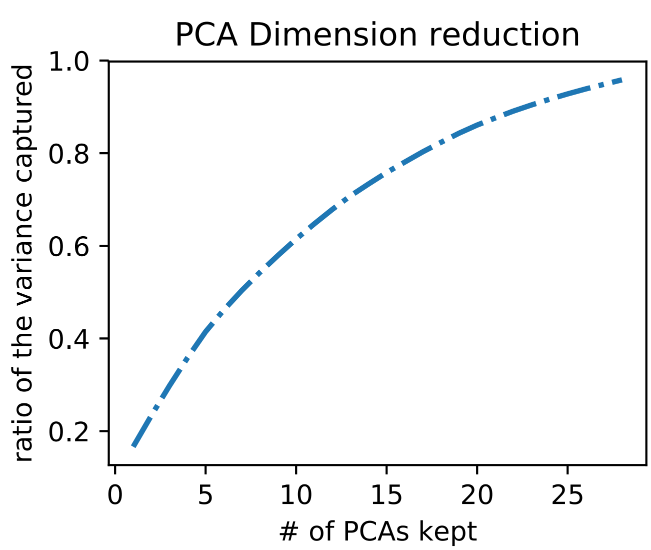

Figure 1: The percentages of variance captured by the number of PCAs.

Given JUST scores, predict market returns

1. We first apply PCA (Principal Component Analysis) to reduce the original JUST score designed by JUST Capital of 47 components to 28 PCAs with 98% of variance captured (Fig.1), of which the 1st PCA captures 21% of the variance. The purpose of this step is to turn the feature vector of each company into a 1 by 28 vector instead of 1 by 47 as the original data.

2. Second, we perform a hierarchical clustering based on the euclidean distance metric for the 896 companies represented by the reduced 1 by 28 feature vectors.

3. Third, we discretize the time series of market returns into N (an arbitrary number) states and label them from 1 to N. In this way, the original time series of market returns will be represented as a temporal sequence of discrete states.

4. Fourth, for each cluster of companies, we collected 2 stochastic properties from the discretized time series of market returns, which are the PDF (Probability Density Function) and the Markov transition matrix.

5. Fifth, after going through the previous 4 steps (Step 1 to 4), we have a PDF and Markov transition matrix for each category of JUST scores. Therefore for prediction, given a JUST score (1 by 47 vector), we first find its reduced feature vector (1x28) and find whichever category it belongs to. With the category known, we predict the time sequence of market returns over consecutive periods using the Markov transition matrix and simulating it as a stochastic process with a given initial condition.

Given market returns, predict JUST scores

1. Similarly as the prediction for market returns, for each company in the training sample, we first discretize the time series of market returns and compute its PDF and Markov transition matrix.

2. Second, we use the PDF and the Markov transition matrix (flattened) calculated from Step 1 as the feature vector of market returns time series, and the corresponding JUST score represented by the 28 PCAs as the labeled outputs to train a Multi-Layer Perceptron (MLP) Regressor.

3. Third, once the MLP Regressor is trained, given a time series of market returns, we compute its PDF and Markov transition matrix, and input them into the MLP Regressor to predict the JUST score.

3 Results and Insights

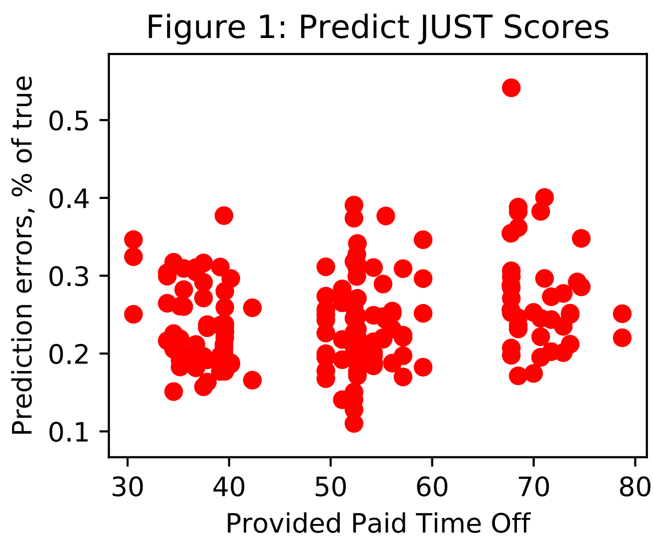

Figure 2: The prediction errors for different companies, quantified as the Euclidean distance between each prediction and true value divided by the magnitude of the true value, plotted against Provided Paid Time Off, PTO.

Predict JUST scores Fig.2 shows the prediction error for different companies, quantified as the Euclidean distance between the prediction and true value divided by the magnitude of the true value, plotted against Provided Paid Time Off, PTO, one of the components of JUST score. We see that the companies can be categorized into 3 baskets based on PTO, each of which has similar averaged prediction errors, about 25%.

Predict market returns Fig.3 shows the predicted time sequence of market returns over different periods for a company (ID 40, JUST_100) as an example. We see that certain pattern sequences (as guided by the green dashed line) are captured by the Markov model.

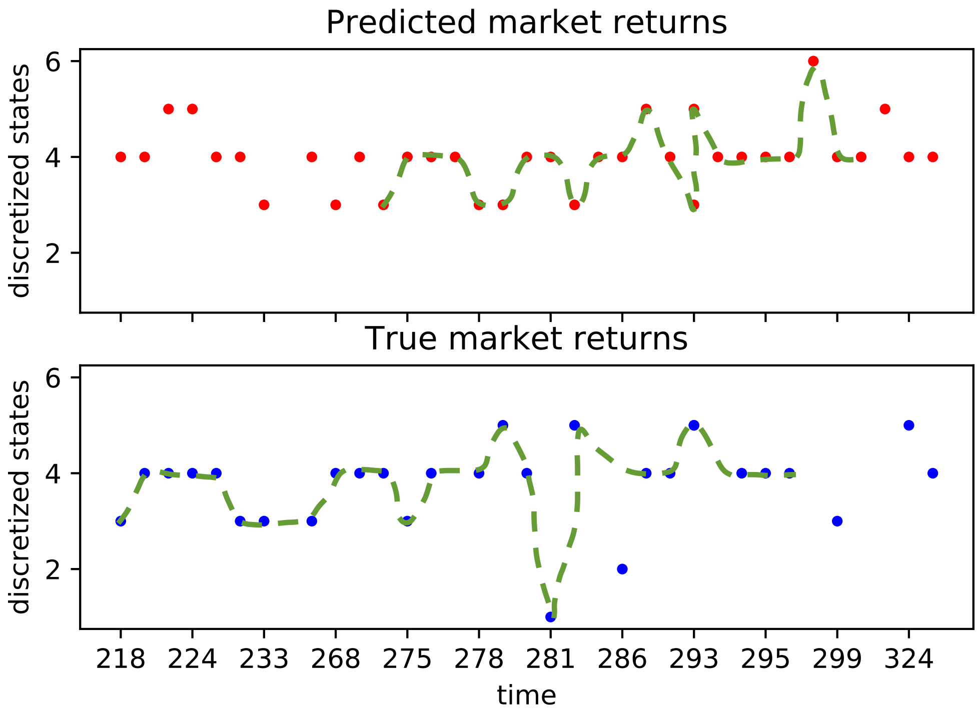

Figure 3: The predicted time sequence of market returns over different periods for the JUST_100 company with ID 40.

4 Further work plan

The following steps will be completed with the indicated time frames.

Quantify the predicability of the Markov model and estimate its temporal upper bound for trustable predictions, *4 hours* A straightforward way, which has been tested to work well in my previous project, is calculating the Euclidean distance between the predictions and the true observations of market returns and then compare it to the standard deviation of the true observations. When the Euclidean distance approaches the standard deviation of the true observations, the Markov model is judged to lose predicability.

Deeper exploration of the weights associated with the feature vectors of the Market returns, *3 hours* We have access to the weights in the trained Multi-Layer Perceptron Regressor (MLPR) and thus should be able to squeeze out some insights regarding the influences of certain statistical features of the market returns on JUST scores. For example, does more volatility lead to a higher or lower JUST score? Are there any state transitions that favor a higher JUST score? etc.

Exploring more clustering techniques to categorize the companies, *4 hours* The current hierarchical clustering based on the euclidean distance metric puts 98% of the companies into one basket, and some other outliers into separate baskets, which does not tell us too much insight about the categories. Also, this clustering result does not help simplify the computation of transition matrices for the prediction of market returns.

Experimenting a "collaborative" transition matrix shared by companies in the same basket to alleviate the scenarios of lacking data, *8 hours* For current preliminary result shown in Fig.3, we calculated a transition matrix for each of the companies. This can face severe challenges when the time series is not long enough so that not all states have been realized. Instead, after we successfully cluster the companies into different baskets, we can experiment to see if all of the companies in the same basket can share with one transition matrix for market return predictions.

Acquiring more data from JUST Capital to make the model and transition matrix more robust and accurate, *8 hours* Currently I only have 767 periods (after cleanup) corresponding to 896 companies, 60% of which are for training and the rest for testing. Analyses show that prediction results have not yet equilibrate with the sample size or the length of time series. Therefore obtaining more data for training is necessary for increasing robustness and accuracy of the models.