

# Predictive Model on Forensic Testing – Window Dressing

Madhavi Pilli, Tanmoy Choudhury and Vipin Singh  
Fidelity Asset Management Technology

## Abstract

Forensic testing techniques have long been considered a best practice by the SEC <sup>1</sup>, which expects registered investment companies and investment advisers to incorporate forensic testing into their compliance programs and Fidelity is no exception. This paper examines predictive model for certain type of trading activity called Window dressing on internal trading data set. Goal of this project is to build a predictive model to analyze Mutual fund holdings over a period of time and predict possible Window dressing events based on

- Buying and selling activity before and after the quarter end, in that order would indicate window dressing
- Analyze the portfolio performance against the standard benchmarks

## 1 Introduction

Forensic testing becomes increasingly popular among investment companies as it strengthens in house compliance. Fig.1 shows the various boundaries of investment wrongdoing can occur as part of operations.

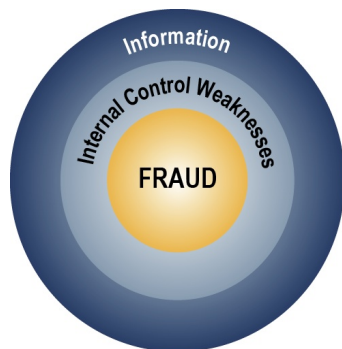


Fig. 1 Forensic Control

**Fraud:** These findings are related to detecting hidden schemes or arrangements within large sets of data. They usually have a deep financial impact.

**Internal control weaknesses:** These findings are related to direct violations of current internal or external compliance regulations. These types of findings normally have a deep regulatory impact, which can have a related financial impact.

**Information:** These findings are related to unintentional errors and inefficiencies by stakeholders and processes that can be discovered through forensic testing. These types of findings can help an organization operate more efficiently.

There are several benefits of Forensic testing. Some of them are

- Identifying potential weaknesses in a fund's or adviser's compliance program that are not readily evident with daily transactional monitoring;
- Allowing for analysis of large data sets (or full populations) as opposed to limited sampling;
- Identifying patterns of data that represent anomalies with respect to expected outcomes;
- Assistance in determining whether a firm's activities are consistent with their compliance policies and procedures and whether controls are effective
- Providing statistical analysis capabilities to review complex transactions or events

<sup>1</sup> Forensic Measures for Funds and Advisers, November 14, 2007 ([https://www.sec.gov/info/cco/forensic\\_testing.pdf](https://www.sec.gov/info/cco/forensic_testing.pdf))

Forensic testing solutions go beyond day-to-day controls and often need to intersect multiple data points to be successful (e.g. transactional data, external data, personal data). Analytic techniques employed in forensic testing have been successfully applied to other areas within asset management firms, including internal audit, risk and business functions.

Forensic testing solutions can be of various types. The major types are

- Portfolio Pumping
- Window Dressing
- Front Running

With a good understanding of the process underlying each category and building a predictive model based on historical data set, financial firms will be in a better position to prevent its occurrence thus avoid any reputation risk scenario.

This proposal will primarily look into Window dressing category to identify and build a predictive model.

## Window Dressing

A strategy used by mutual fund and portfolio managers near the year or quarter end to improve the appearance of the portfolio/fund performance before presenting it to clients or shareholders. To window dress, the fund manager will sell stocks with large losses and purchase high flying stocks near the end of the quarter. These securities are then reported as part of the fund's holdings <sup>2</sup>. So, if we typecast it, the act of increasing holding at the beginning of the quarter and decreasing holding at the end of the quarter for a security (Fig. 2) can be of one type (Type I) and decreasing holding at the beginning of quarter and increasing holding at the end quarter (Fig. 3) can be of other type (Type II). Loosely, this also means that mutual fund disclosure frequency is directly proportional to this activity. Starting May 2004, SEC changed the mandatory disclosure frequency from every six months to every

quarter with the goal to prevent various portfolio manipulation activities and study suggests that window dressing activity drops to 8.72% from 11.73% <sup>3</sup>.

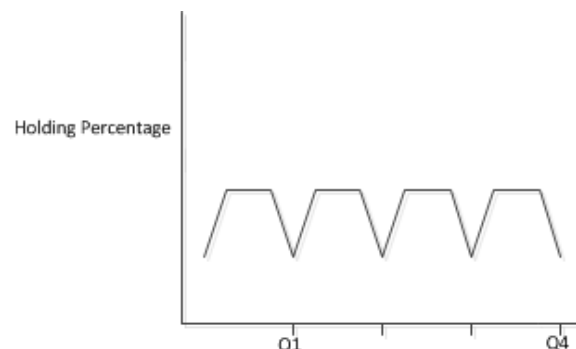


Fig.2 Type I (Decreasing holding end of Quarter)

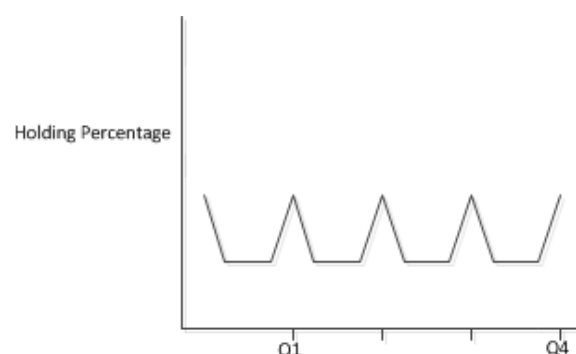


Fig.3 Type II (Increasing holding end of Quarter)

Now is it frequent for individually managed fund or team managed fund? Recent study on Window dressing (e.g., see Patel, Sarkissian, 2013) suggest that team managed funds sell less extreme losing stocks and buy less extreme winning stocks than their single-managed counterparts<sup>4</sup>.

<sup>2</sup> <http://www.investopedia.com/terms/w/windowdressing.asp>

<sup>3</sup> Xiaolu, Wang, 2014, Mutual Fund Window Dressing: Prevalence, Flow Reaction, and Limited Attention

<sup>4</sup> Saurin, Patel; Sergei, Sarkissian, 2013, Deception and managerial Structure: A Joint Study of Portfolio Pumping and Window Dressing Practices

## 2 Methodology & Approach

Set of data attributes for a period of 20 months (refer Appendix) is chosen from Fidelity's internal trading and holdings system to build a training and test dataset. Data cleansing was performed to the training data to take care of missing values and outliers. SQL is used for pre-processing of dataset and then exported into csv format. The following three different types of classification models have been used on the training data set

- Decision Tree\* (J48)
- Naïve Bayes\*\*
- Support Vector Machine (SVM)\*\*\*
  - One class SVM (OC-SVM)

Upon achieving the most possible accuracy on each classification model based on tuning various parameter associated, the model was then applied to test data to record predictions and test the model.

Weka software tool is used to predict and test various models.

The dataset identified did not require extensive cleansing. As part of preprocessing activity in Weka, it was required to remove couple of attribute to reduce noise in the dataset along with discretization to deepen the decision tree.

## 3 Observation

After finalizing the training data set, it was loaded in to Weka for preprocessing. Training dataset used for this project contains around 18 attributes including the target indicator to determine the outlier. The attributes considered are

- Fund Manager Identifier
- Mutual Fund Id
- Window Dressing Type code
- Holding Effective Dt
- End of Month Holding Rank
- Max Holding Rank

- Min Holding Rank
- Max Total Holding %
- Min Total Holding %
- Max NAV (Buy Trade)
- Max NAV (Sell Trade)
- CMP Benchmark Rank
- Share Benchmark Rank
- Fund Performance Group Code
- Loser
- Winner
- Total Holding %
- Window Dressing Indicator

Discretization was done on End of Month holding rank attribute along with modifying window dressing indicator from numerical to nominal. Observations on the data distribution among attributes are shown in Fig.4.

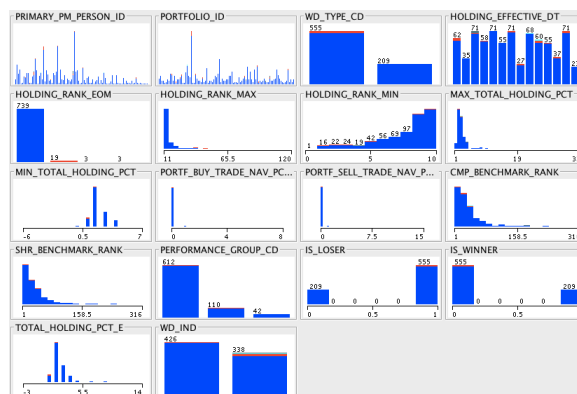


Fig.4 Data distribution visualization in Weka

\*Decision tree learning uses a decision tree as a predictive model which maps observations about an item to conclusions about the item's target value. It is one of the predictive modelling approaches used in statistics, data mining and machine learning. ([http://en.wikipedia.org/wiki/Decision\\_tree\\_learning](http://en.wikipedia.org/wiki/Decision_tree_learning))

\*\*The Naive Bayes Classifier technique is based on the so-called Bayesian theorem and is particularly suited when the dimensionality of the inputs is high. ([http://en.wikipedia.org/wiki/Naive\\_Bayes\\_classifier](http://en.wikipedia.org/wiki/Naive_Bayes_classifier))

\*\*\*Support Vector Machine (SVM) is primarily a classifier method that performs classification tasks by constructing hyper planes in a multidimensional space that separates cases of different class labels. SVM supports both regression and classification tasks and can handle multiple continuous and categorical variables. (<http://www.statsoft.com/textbook/support-vector-machines>)

## Decision Tree (J48)

J48 algorithm was used on this training data set with a cross validation of 5 folds and observed 77% accuracy (Fig.5)

```

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      593      77.6178 %
Incorrectly Classified Instances    171      22.3822 %
Kappa statistic                    0.5235
Mean absolute error                 0.317
Root mean squared error            0.4834
Relative absolute error             64.2411 %
Root relative squared error        81.2174 %
Total Number of Instances         764

=== Detailed Accuracy By Class ===
          TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
Weighted Avg.   0.776   0.278   0.826   0.776   0.76      0.76      1

=== Confusion Matrix ===
      a  b  <-- classified as
419  7  |  a = 0
164 174 |  b = 1

```

Fig.5 Model prediction on training data (J48)

Setting binary split parameter to true, improved predication a bit (around 79%). We then apply the model to test data set and observe 85% accuracy (Fig.6)

	Instance	% Correct
Total Number	308	
Correctly Classified	262	85.06%
Incorrectly Classified	46	14.94%

Fig.6 Model output on test data (J48)

Weka's built in feature selection method which automatically evaluates the training data set for best fit based on various algorithms was also used. Implementing the suggestions to training data did not provide much benefit for this use cases.

## Naïve Bayes

Running Naïve Bayes model on default parameter did not procedure better prediction (around 60%). Tuning supervised discretization parameter yields 69% model prediction (Fig.7).

	Instance	% Correct
Total Number	308	
Correctly Classified	257	83.44%
Incorrectly Classified	51	16.56%

Fig.7 Model output on test data (Naïve Bayes)

## Support Vector Machine

The following two SVM classifications were applied on training dataset.

- Classification SVM Type 1 (also known as C-SVM classification)
- Classification SVM Type 2 (also known as nu-SVM classification)

For SVM type 1, an average of 61% prediction accuracy was observed. The model predication did not improve much when different parameter settings were used. SVM type 2 was used with kernel parameter radial basis function, which shows 68% accuracy (Fig.8). Changing different kernel parameter did not produce improved result.

Fig.8 Model output on training data (SVM type 2)

When applying the model to test dataset 72% accuracy was observed (Fig.9), much lesser than the other two algorithms.

	Instance	% Correct
Total Number	308	
Correctly Classified	222	72.08%
Incorrectly Classified	86	27.92%

Fig.9 Model output on test data (SVM type 2)

## One Class Support Vector Machine

Using One Class SVM with default parameters did not produce good results. When tuned the model further to Kernel Type=RBF, normalize=True and nu=0.1, the model gave

much better accuracy of 84.57% (Fig.10). This is almost closer to the accuracy of J48 model.

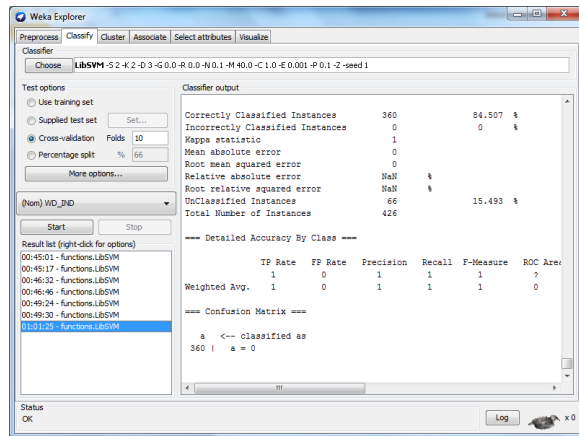


Fig.10 Model output on training data (One SVM)

Applying this model to test dataset 76.30% accuracy was observed (Fig.11), less than J48 but better than Naïve Bayes algorithm.

	Instance	% Correct
Total Number	308	
Correctly Classified	235	76.30%
Incorrectly Classified	73	23.70%

Fig.11 Model output on test data (OC-SVM)

## Oracle Data Miner

Along with Weka, oracle data miner was also explored for model prediction. For this exercise Oracle SQL developer data miner extension was used. The oracle data mining package is composed of the following

- A Project
- A workflow
- A workflow component (like source, transformation, target component etc.)

Basic steps in creating workflow are

- Define source data (accept customized query)
- Discretization of values (using the Transform component)
- Explore data which will give you different data profiling statistics like average value, % of null values, etc. ( similar to Weka Explorer)

- Specify the class build (classification, association, etc.).
- Run test data against all the models in the class build
- Build model based on source data (training data set)
- Compare the result of all the data mining models (lift factor, predictive confidence %, cost to run the model etc.)
- Apply the data model to test data
- Specify output component in order to store the data mining model prediction in database

Using oracle SQL developer data miner extension, predication models were built in the oracle database and applied to test data. To help oracle data miner, two tables were created to represent training and test data respectively. Data miner workflow was built based on these tables to construct model on training data and then apply it to test data set. The SQL developer data miner has feature that does classification for 4 models simultaneously, showing predicative confidence, average accuracy, overall accuracy and cost information.

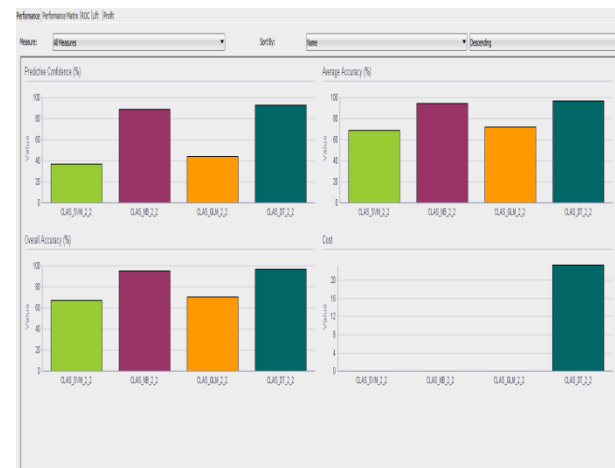


Fig.12 Model Performance on Oracle data miner

Findings show prediction% difference between Weka and Oracle data miner (Fig.13) to the same training data set.

Models	Correct Predictions %	Correct Predictions Count	Total Case Count	Total Cost
CLAS_NB_2_2	94.8276	275	290	0
CLAS_DT_2_2	95.8621	278	290	23.2474
CLAS_GLM_2_2	70	203	290	0
CLAS_SVM_2_2	66.8966	194	290	0

Fig.13 Model prediction on Oracle data miner

Applying the model to test data set, % accuracy (Fig.14) observed as below

	DT	%	NB	%	SVM	%
Total Number	327*		327		327	
Correctly Classified	313	95.72	285	87.16	268	81.96
Incorrectly Classified	14	4.28	42	12.84	59	18.04

Fig.14 Model accuracy on test data

Both the tools provided greater predication % using decision tree model. However in the case of oracle data miner there was no clear information on the specific algorithm used in decision tree. Once the data miner repository was installed, building and applying the model via oracle SQL developer was pretty quick. One class SVM was not explored in Oracle Data miner as part of this project scope.

## 4 Conclusion

A good predictive model can be built and applied to identify various types of forensic testing categories, in this case, window dressing activities. Comparison of tools used for this data mining problem is out of scope for this project. Further analysis is required to establish a conclusion on the best tool for this data mining project. However, a fair assumption can be made that oracle data miner can be a natural choice to evaluate different models if dataset is primarily Oracle based. Decision tree model provides a greater efficiency for dataset used for this project.

## 5 Appendix

### Data Attributes

Attribute	Description
Fund Manager Identifier	An identifier that uniquely identifies a fund manager primarily responsible for this portfolio
Window Dressing Type code	The act of increasing holdings at the beginning of the month and decreasing holdings at the end of the month for a security the fund manager likes (refer to as Type I WD). The act of decreasing holdings at the beginning of the month and increasing holdings at the end of the month for a security the fund manager does not like (Type II WD)
Holding Effective Dt	The date the holding was effective, as it was reported by accounting
End of Month Holding Rank	The ordinal rank of this security in terms of total market value, relative to all other securities held, as of the last day of the month.
Max Holding Rank	The highest rank enjoyed by this security, in terms of market value relative the remaining securities held, across the reporting period
Min Holding Rank	The lowest rank enjoyed by this security, in terms of market value relative the remaining securities held, across the reporting period
Max Total Holding %	The largest percentage of holdings that this security has enjoyed during the reporting period (in terms of total market value).
Min Total Holding %	The smallest percentage of holdings that this security has enjoyed during the reporting period (in terms of total market value).
Max NAV (Buy Trade)	Net Asset Value. The largest % of buy trade against fund NAV for the last five trading days of the month.
Max NAV (Sell Trade)	The largest % of sell trade against fund NAV for the last five trading days of the month.
CMP Benchmark Rank	The ordinal rank of this security in terms of weight of this security in the compensation benchmark, relative to all other securities market value, as of the last day of the period. Null if this security is not a constituent of the compensation benchmark.
Share Benchmark Rank	The ordinal rank of this security in terms of weight of this security in the shareholder benchmark, relative to all other securities market value, as of the last day of the period. Null if this security is not a constituent of the shareholder benchmark.
Fund Performance Group Code	High/Middle/Low - A categorization of portfolio's performance, relative to other portfolios sharing the same objective, for this reporting period. Either 'High', 'Low', or 'Middle' for portfolios that exceed the mean return by greater than one standard deviation, less than one standard deviation, or between 1 and -1 standard deviations from the mean, respectively.

Loser	Set to 1 if this securities total return over the period trails the benchmark's return over the same period. Zero otherwise.
Winner	Set to 1 if this securities total return over the period exceeds the benchmark's return over the same period. Zero otherwise.
Total Holding %	The percentage of total holding market value that is represented by this security on the holding effective date
Window Dressing Indicator	Model will predict/tag this attribute