Informeta, L.L.C.

Detecting Anomalous Trading Behavior

Version 1.0, 11 August 2006

Matthew Johnson
matthew.johnson@informeta.net

1 Civic Center Plaza, Suite 540
Poughkeepsie, NY 12601
www.informeta.net

1 Real-World Example

Informeta's artificially intelligent software engine, Mentys™, identifies and fixes data anomalies in real time processes. Mentys is intended to be completely data-driven; that is, it derives patterns and relationships directly from the data without additional expert input. Mentys accomplishes this using robust, non-parametric statistical methods and cutting-edge, unsupervised learning algorithms.

In our experience, however, there often are important relationships that cannot be gleaned from the data alone for one reason or another. The data samples may not be sufficiently large, or some relationships may depend on factors external to the known data. In order to account for these situations, Mentys also is designed to support the use of *prior knowledge*, which is any additional information about the problem that can be provided by application experts. There are a number of different ways that prior knowledge could be incorporated into a Mentys solution.

It is most effective to input prior knowledge directly using Mentys' built-in framework for specifying constraints on its knowledge base. This can mean explicitly specifying a causal relationship between two variables, or perhaps expressing some expected mathematical or logical relationship among variables. An example of a knowledge base constraint is the requirement that the “Bid” price always be less than the corresponding “Ask” price. Other uses of prior knowledge include removing unchanging or irrelevant data fields, and selecting mathematical data transformations.

For many solutions, it is not necessary to provide certain kinds of prior knowledge directly to the Mentys engine. Frequently, the best choice is to apply the prior knowledge either before or after invoking Mentys. Creating a specialized pre- or post-filter can help to reduce noise in the analysis and free up Mentys to focus on relationships that can be derived from the data itself.

2 Real-World Example

In early 2006, we investigated foreign currency trades for a leading financial services company (“the Client”). In general terms, our goal was to determine if any traders were involved in some sort of illegal or unethical trading. Specifically, we tried to determine whether particular traders intentionally sold above or below the market, especially when such behavior appeared to be unfavorable to the Client. We also looked for patterns in the customer field that might indicate an accomplice.

We recognized the importance of this effort, since one rogue trader operating inappropriately over a short period of time can potentially wipe out years' worth of company trading profits. The output from Mentys included an audit trail listing details of each transaction that appeared to be anomalous.

2.1 Performance considerations

The elapsed time of this engagement was approximately 6 weeks. Most of this time was spent setting up the initial test scenarios, applying prior knowledge, and tuning the system parameters. The actual run time of the Mentys process was on the order of one hour.

The preliminary work to develop this kind of solution needs to be done just once per client. Regular operation requires only executing the program and reviewing the results with application experts (i.e., auditors). We also recommend periodically refreshing the knowledge base with the latest available data and recalibrating the system. It may be practical to perform these actions overnight or during the weekend if the data contains thousands of variables and/or millions of records.

2.2 Results

In total, we analyzed about 60,000 trades made by 55 different traders over a six-month period. Our final analysis determined that one trader exhibited a pattern of behavior different than that of his peers. We noted a higher incidence of deals made off the market rate, and identified two specific customers that appeared to be involved in a disproportionate number of the anomalous trades. We found these differences to be highly statistically significant (greater than 99.9% confidence).

The charts below highlight some of the differences between the suspect trader and his peers, as well as a concentrated view of the suspect's most anomalous trades.

Figure 1. Comparison of deviation by percentile.

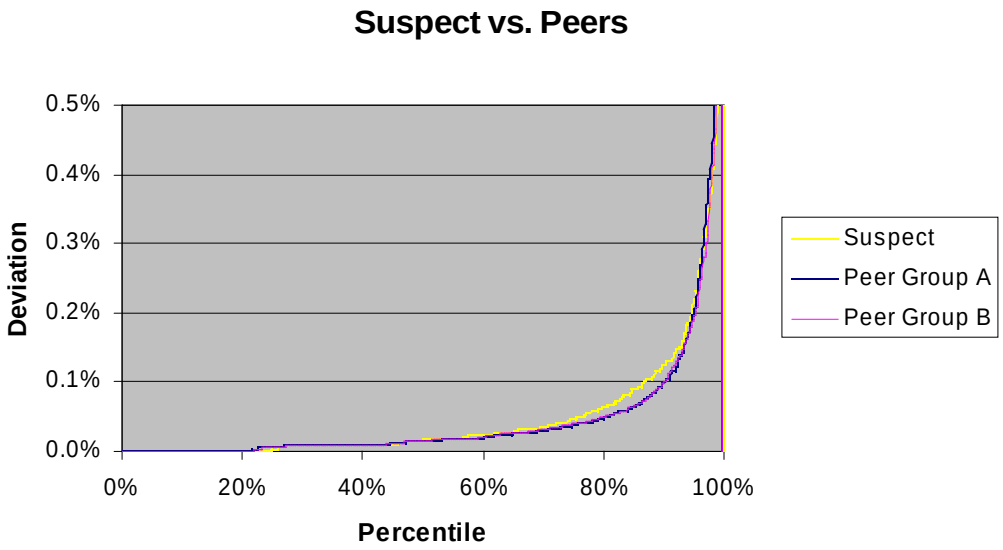
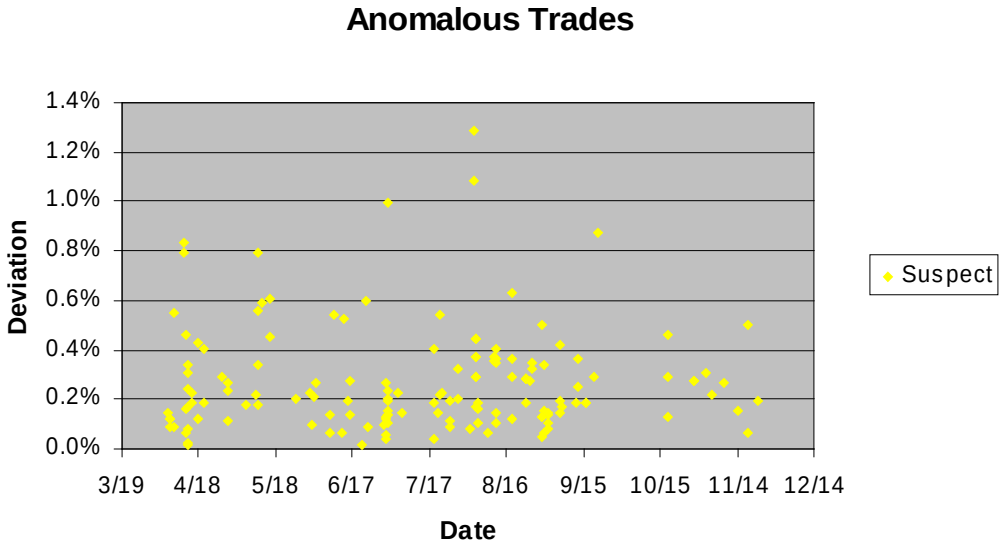


Figure 2. Off-the-market trades made by the suspect.



2.3 Prior knowledge

Our analysis of the trading data described above initially used no additional prior knowledge from the Client. This was due to both legal issues of disclosure and our unfamiliarity with the problem domain. The initial results indicated that prior knowledge outside the historical data would positively influence our conclusions. Further consultation with the Client provided additional prior knowledge about the data, which allowed us to greatly improve the accuracy of the Mentys predicted exchange rates.

Below are several examples of prior knowledge that was ultimately provided by the Client.

1. Several traders, not including the suspect, exhibited unusual patterns of behavior that differed significantly from the others we analyzed.
   - ➔ The Client confirmed that, indeed, these traders were either "options traders" or operated within specialized portfolios such as emerging markets or treasuries.
   - ➔ Knowledge of legitimate business reasons for otherwise unusual trade patterns allowed us to classify and evaluate traders in groups.
2. Certain customers appeared in many unusual trades, regardless of which trader was involved in the deal.
   - ➔ The Client confirmed that these customers were internal branches within the Client company itself.
   - ➔ According to the Client's recommendation, we subsequently ignored all trades involving these customers.
3. Other characteristics of the various traders indicated that, in some cases, several distinct trader IDs referred to a single trader.
   - ➔ The Client confirmed the instances we identified as well as others.
   - ➔ We subsequently condensed the data set using a complete mapping of ID-to-Trader.
4. We initially relied on daily averages of the inter-bank rates to support our spot rate predictions, though the Client indicated that detailed intra-day rates also may be available from external sources.
   - ➔ "Market counter-parties" are large companies that cannot expect to transact with one another off the market rate; e.g., trades between Citi and Chase will be "fair trades."
   - ➔ The Client later identified specific market counter-parties among the various customers; deals involving market counter-parties are always made at market rate and cannot be manipulated.
   - ➔ We extrapolated the intra-day rate curves from all trades involving market counter-parties.

Future work in this problem domain, for other clients, would require the kinds of prior knowledge described above. We believe that we can improve the Mentys analysis by including additional prior knowledge, such as:

1. Any information that can help match corresponding buy and sell transactions.
   - ➔ This would be extremely useful when attempting to characterize off-the-market trades in terms of whether or not they benefit the client.

2. More information about the customer field, including whether or not multiple labels can refer to the same customer (as noted above for traders).
   - ➔ This sort of prior knowledge could help us classify customers and determine the likelihood that a given customer is complicit in a fraudulent transaction.
3. For infrequently traded currencies we did not have sufficient data to reliably predict the expected exchange rates.
   - ➔ Additional information regarding market rates, either from the client or from an external source, must be obtained.
4. Information that can help to track a given trader's running balance, including inventory and profit/loss.
   - ➔ This kind of information would allow us to reliably determine whether or not a particular trader appeared to be acting in the best interest of the client.

2.4 Conclusions

Our primary goal was to find evidence of suspicious behavior, which we accomplished successfully. We also anticipate that the use of Mentys with embedded prior knowledge, as described above, could have a number of related tangible benefits.

Some of the trades that we identified to be anomalous may be simply the result of unsatisfactory (or, conversely, exceptional) trading performance rather than the result of malfeasance. By recognizing particular trade patterns that yield unusual profit or loss, the Client could promote good trading practices and reduce poor ones.

We note that expert prior knowledge of the nuances of good or bad trading practices might also be used to inform future Mentys analyses.