dotData

Case Study:

Eliminating $5M in Retail Fraud with Predictive Analytics

How a regional multi-store chain leveraged machine learning automation to identify and minimize fraud

5M+ ESTIMATED IMPACT OF FRAUD PREVENTION

200+ NUMBER OF RETAIL LOCATIONS

1 DAY TIME NECESSARY TO CHANGE FRAUD POLICIES POST-MODEL

Recent years have brought on numerous challenges for the retail industry. From pandemic-enforced closures and redundancies to recent upswings in looting and POS fraud, brick-and-mortar businesses compete with a surging e-commerce sector while defending their bottom line and fighting for customer loyalty. Providing a welcoming environment for shoppers while maintaining security and preventing theft can prove a delicate balancing act.

THE CHALLENGE OF RETAIL FRAUD

Fraud impacting convenience store business.
Retail crime and fraudulent transactions continue to impact convenience retailers financially. The 2019 National Retail Federation's (NRF) annual Retail Security Survey identified $61.7 billion in revenue loss through crime alone.

For each fraudulent dollar, the NACS survey finds, U.S. retailers lose $3.60 (up from $3.16 pre-COVID).

Association For Convenience & Fuel Retailing (NACS)

ABOUT THE CLIENT

Industry: Retail
Locations: 200+
Employees: 1000+
Annual Revenue: $500M+

CHALLENGES

- 200+ locations across multiple states and en employee to store ratio of 4:1 meant that monitoring fraud manually was difficult.
- Manual forensic analysis of shrinkage was taking too long and consuming valuable store manager time.
- Lack of actionable store-level insights was not possible on a large scale across all 200+ locations.

SOLUTION

- dotData's proprietary Feature Engineering (AutoFE) technology evaluated billions of data points across multiple categories - simultaneously.
- Combination of AutoFE and Automated Machine Learning (AutoML) technologies allowed the business to update fraud prevention measures in less than one day.

According to the NACS survey, 2021 saw a 15% increase in retail fraud cost and volume in Canada and the US. The COVID-19 pandemic has exacerbated the pre-existing trend of increasing retail theft.

The Prevalent Types of Fraud:

Consumer fraud is focused mainly on two practices: "sweethearting" and return fraud. Sweethearting involves tampering with pricing or misuse of discounts or coupons.

Return fraud consists of partly consumed goods returned for refunds (clothing retailers are especially susceptible to this type of fraud). Return fraud can also occur when consumers return goods bought on sale and return them when the price increases to the pre-sale level. Retail returns increased during the pandemic, with the NRF estimating $10.30 of value lost to fraud out of every $100 in merchandise accepted during returns by a retailer.

Employee Fraud includes cases where an employee steals goods (also known as shrinkage) and wrongly prices them, typically to assist consumer fraud. Employee assisted fraud has become commonplace, duping employers out of billions of dollars annually. Employee-assisted fraud can include marking up goods, pocketing the difference, or selling items at a discount to friends.

According to CNBC employee theft alone costs US businesses over $50 billion a year.

CLIENT CHALLENGES

Why is it difficult to detect fraud?
The client faced a dilemma. Identifying and preventing fraud would involve constant vigilance and up-to-date knowledge of the fraudulent activities that customers and employees could perpetrate. With an average employee-to-store ratio of a little over 4:1, there wasn't the time or opportunity to be hyper-vigilant or keep up to speed with trending scams and fraudulent activities.

With so much revenue lost to employee fraud, managers were spending too much time monitoring staff behavior or performing forensic analyses to discover the source of lost revenue. And even when they had the opportunity to look for evidence within the data, the client quickly found it an

enormous task. Performing statistical analysis did not seem feasible with millions of data points, dozens of variables, and no clear signal amongst the noise.

Because of the lack of clear and actionable insights into the type of fraud to prioritize in each store, installing fraud-prevention measures across 200+ locations across multiple states was not feasible from a corporate point of view. Our client needed a more focused, efficient, and automated fraud detection and prevention strategy.

DOTDATA: AN AI-POWERED SOLUTION

Advancements in AI make it ideal for aiding early fraud detection and prevention. According to a 2019 Forbes article, 80% of fraud specialists believe AI helps reduce payments fraud, while 63.6% of financial institutions utilizing AI claim it can prevent fraudulent activity and cite it as their preferred tool. Although AI is relatively new in retail, almost a quarter of retailers already use AI-empowered fraud detection measures.

AI's significant advantage over manual labor is that it works quickly at scale. AI systems can analyze billions of data points across months of transaction histories, inventory data, store locations, product categories, pricing information, etc. AI systems can interrogate a wide variety of data to identify trends, generating fraud propensity scores that help predict and prevent inappropriate activity.

Corporate Level:
With AI generating fraud propensity models, cross-referenced by store or product item, the retailer can create and deploy workable fraud identification and prevention plans across all its stores. Automation of the analytical process means little or no employee labor during the analysis phase, so there's time to educate store owners, managers, and security personnel on effective fraud prevention methods.

Store Level:
Fraud propensity scores per item are provided to owners to identify where the most significant revenue risks lie and alter inventory or take preventative measures. Preventive measures might include item positioning, security tagging, CCTV, in-store warning signage, or the hiring of trained security personnel.

RESULTS

- $5M+ in estimated cost recovery from fraud prevention measures instituted as a result of dotData.
- The entire process, from data analysis to model development involved 0 lines of code and was performed by one data scientist.
- Fraud prevention measures were instituted in less than one day as a result of the model output.

LEARN MORE

https://dotdata.com

https://dotdata.com/products

TECHNICAL CHALLENGES IN DEVELOPING AI SOLUTIONS

Multiple variables contribute to fraud propensity, such as store location, specific item, and time of day, to name just three. This multi-variable aspect makes it hard to correlate fraudulent activity with particular factors.

Since the pandemic, there have been surges in retail-related criminal activities to make matters more complex. These may not correlate with pre-pandemic trends. For this reason, it is essential to monitor fraud patterns carefully and avoid making false assumptions.

The client needed AI solutions to identify multiple trends, temporary patterns, and consistent problems. In short, the client needed to see the forest for the trees. Pattern recognition is something that AI excels in when well-targeted and trained through machine learning (AutoML) and feature engineering. dotData knew they had just the tools their retail company needed.

Why dotData? dotData uses automated machine learning and other strategies for pattern recognition.

Challenge1: Multiple fraud patterns:
With 200+ stores, more than 1,000 employees, and thousands of items per outlet, crunching the data was always difficult for the retailer. The company wasn't even sure where to identify significant risk areas.

The client had noted a pattern of negative transactions on certain items, but the source of this unusual feature was unknown. There were too many factors to consider manually, and there wasn't the time to engage in deep, manual analysis.

How dotData solved this challenge:
dotData used a revolutionary, proprietary AutoFE system to analyze billions of data points across multiple categories: transaction data, inventories, store information, product and category specifics, calendar data, etc. The AutoFE system analyzed the available data and identified "feature patterns" – recurring fraud signals hidden in the dataset.

By using AI automation, dotData was able to save the retailer untold hours of intricate and error-

prone manual labor. Crucially, the AutoFE system was able to pick out patterns and signals far more subtle than those a human analyst could spot.

Example: A high ticket item was returned several times a day.

The example above was an indicator of potential fraud. The client wanted to know what early indicators to look for to prevent such fraudulent transactions from recurring. dotData examined the relevant variables (time of day, employee ID, transaction data) and provided the necessary correlation points.

Challenge 2: Fraud patterns change.
Since the start of the COVID-19 pandemic, fraudulent activity has increased, and new patterns of criminal activity have emerged. For instance, social distancing measures often meant stores operating with a skeleton staff, making opportunistic theft easier.

In addition, a shift towards click-and-collect or BOPIS (buy online pick-up in-store) has resulted in new types of online fraud, causing inventory loss and POS (point of sale) fraud.

These changing patterns of crime made it vital to find ways to quickly update data sources so that they could be used to identify new sources of risk before they negatively impacted the client's revenue.

How dotData Solved this Challenge:
dotData used AutoFE and AutoML to continually update their fraud detection models by analyzing high volumes of real-time data. These technologies proved so efficient that it took less than a day for the organization to update its fraud detection and prevention measures.

Each data model could be collated for future use, with no new intelligence lost. Different models could be applied to various items, transactions, and store locations. New fraud patterns were continuously identified and added to the business intelligence repository. With each iteration, it became harder for fraudsters to spot an opportunity.

LEARN MORE ABOUT DOTDATA:

dotData was designed exclusively for small and mid-sized organizations looking to leverage AI and Machine Learning to help build predictive analytics models to grow their business. Learn more by visiting our website:

https://dotdata.com
https://dotdata.com/blog

About dotData

dotData solves the biggest challenge of organizations of any size: Turning raw business data into valuable and meaningful data marts ready for Machine Learning (ML), Artificial Intelligence (AI), and traditional data analytics deployments and applications. dotData provides solutions tailored to the needs of companies that are just getting started with predictive analytics and companies with more mature data engineering processes. Our core technology allows companies to automatically convert data from data warehouses and data lakes into data marts and feature tables by exploring the relationships between varied data tables with hundreds of columns and millions of rows. Our global customers have used our platforms to accelerate their ML, AI, and Advanced Analytics adoption, achieving rapid ROI by lowering their dependence on scarce, costly expert resources.

Forrester recognized dotData as a leader in ML and AI in 2019, and CRN named dotData to its emerging vendors' list in for four years running and was named a CB Insights Top 100 AI Startups for 2020. The AI breakthrough awards recognized dotData as the "best machine learning platform" for 2019, and Fortune 50 clients around the Globe rely on dotData to help them accelerate their ML, AI, and Advanced Analytics projects. For more information, visit www.dotdata.com, and join the conversation on Twitter and LinkedIn.