TIBCO

Model Ops for Dynamic Pricing in Insurance

Background

Companies stuck with traditional dynamic pricing software are struggling to compete in today's digital economy. These black-box point solutions leave little room to react to market trends and regulations, or keep up with the increasing volume and complexity of data. The insurance industry especially has started to move away from these outdated solutions towards artificial intelligence (AI) fueled dynamic pricing systems that enable them to transform, differentiate, and compete. These modern AI solutions enable insurers to quickly react to market changes and stay competitive.

However, using AI-based dynamic pricing solutions is not happening as much as insurance companies would like. The main reason companies are not deploying these solutions is due to the complexities of getting AI into production. In fact, IDC found that less than 50 percent of the best models actually get deployed and become part of the everyday business. Forbes reported that successful deployment rates are less than 10 percent.

In this paper, we describe how the auto insurance unit of a large TIBCO customer used the TIBCO Model Ops solution to develop and deploy (or operationalize) a working and successful AI-powered dynamic pricing system. The essence of operationalizing data science lies in deploying and integrating machine learning (ML) models with business applications to derive as much value as possible from predictions.

Goals of an AI-based Dynamic Pricing Tool in Insurance

In insurance, the goal is to acquire and retain customers with the highest potential lifetime value while covering underwriting costs, handling demand for coverage, and dealing with unexpected external factors such as natural disasters. A dynamic pricing strategy, in which a business sets prices for services based on market demands and other factors, helps insurers adjust price based on potential risks, customers' willingness to pay, competitor pricing, and other variables. Only systems using AI models are adept at handling these changing environmental factors.

The company's goal was to optimize pricing and understand the types of customers it attracts and wants to attract. IT also wanted to decrease fraud across the organization. Traditionally, fraud and customer value are extremely difficult to calculate in live environments. The company found that TIBCO was the only provider that could optimize pricing, provide a deep understanding of customer types and value, and prevent fraud, all at the same time.

The innovative pricing model solution built with TIBCO included configurations for price optimization and customer acquisition and retention. It also showcased price elasticity data and modeling of acceptance as a function of price. Following, we'll take you step by step through the TIBCO Model Ops process that the company used to build and get an innovative dynamic pricing model into production.

Operationalizing Data Science in Four Steps

The obstacles preventing most companies from benefiting from AI are many. Most organizations today develop models in many different systems and there is a need to properly govern, deploy, and track these assets. In addition, data scientists tend to focus on how to build a better model or apply the latest technique and are not aware of or trained on, the process-centric work required to get an analytic model efficiently into production. Also, models are made of both code and data that evolve differently overtime, rendering the models useless if they are not retrained and maintained.

Without Model Ops, your organization's ability to efficiently move analytic models from development into deployment at scale is limited by silos, manual processes, disagreement on how value will be measured, and lack of alignment by key players.

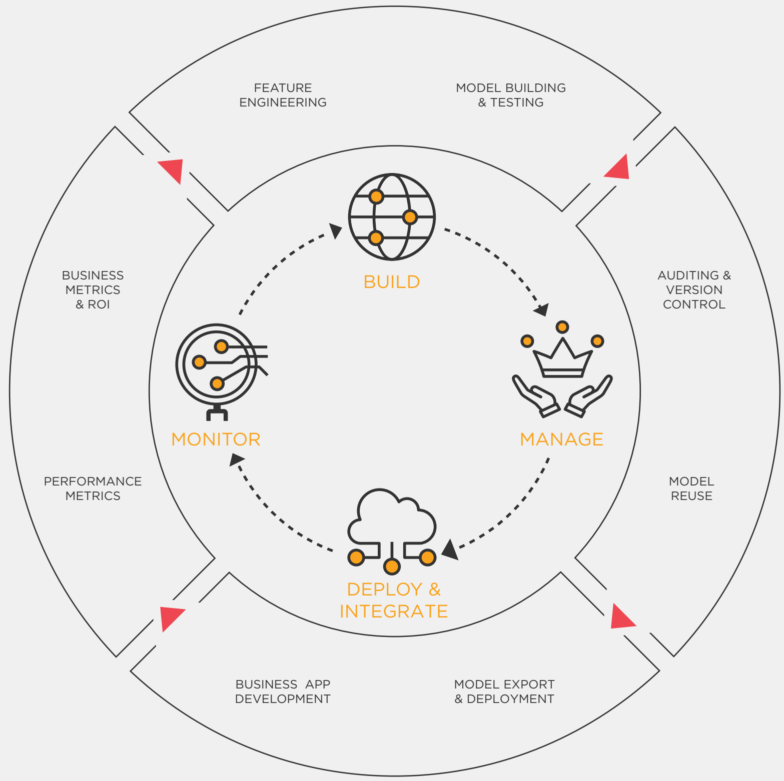

Figure 1. Model Ops, Operationalizing Data Science

The TIBCO Model Ops solution was designed to containerize the production lifecycle of ML models. To produce actionable insights and optimize pricing for insurance quotes, the company focused on four main steps to get its models into production: Build, Manage, Deploy, and Monitor.

Step 1: Build the Models

For dynamic pricing, customer models predict the probability of churn among existing customers and of acquisition of new customers. The automobile insurance company built its models using historical acceptance/rejection rates of insurance premium quotes, customer demographics, and vehicle attribute data (i.e. vehicle age, vehicle type, make, model, etc.).

Model Training and Selection

Data scientists then trained models using techniques that included generalized linear regression (GLM) with regularization via elastic net balancing between ridge regression and lasso methods. In many industries, there are regulations and laws that prevent the use of certain variables in models (gender, race, etc.). For this example, coefficients for explanatory variables were constrained during the modeling to minimize the risk of violating these laws. Customer lifetime value, commission (profitability), and price elasticity (sensitivity to a price discount for individual customers) were used to segment the customer base and obtain optimal tables for commission and discount.

Feature Engineering and Banding

To ensure the models would be as accurate as possible, the company analyzed quotes from its data warehouse to generate features and banding logic. TIBCO software would also use the banded quotes later in the model ops process for retraining the GLM models.

Step 2: Manage the Models

The insurance company wisely regarded the model not merely as a piece of static code running inside an application but in terms of its lifecycle from inception to ongoing monitoring. Without monitoring, the company knew that models could decay based on constantly changing data, leading to invalid results.

Auditing, Approvals, and Version Control

In most cases, models had to be rebased on fresh data to account for variance in market conditions since the previous model training. That meant saving and governing several versions of the models according to regulatory requirements. It was also necessary to keep the models available for auditing and compliance assessments.

Reusing and Repurposing Models from a Centrally Managed Repository

To obtain churn or conversion probabilities, and for purposes of compliance and auditing, the GLM modeling objects were stored in a database table as PMML objects. (As an alternative, models and rules can be stored as artifacts within other TIBCO or customer applications.)

The objects were used for real-time model inference against customer attribution data such as age, vehicle type, no-claims discount, and customer membership duration in years. TIBCO software detected when a new model or commission table was marked active in the data store, then downloaded and replaced its currently cached objects with the new, active versions.

Step 3: Deploy/Integrate the Models

The essence of operationalizing data science lies in deploying and integrating ML models with business applications. If ML models are not integrated with business applications, they cannot be applied to real world situations or handle the business problems at hand.

Pushing Models into the Real World

To calculate the premium price and discount, users enter their information via a web portal. The dynamic pricing application aggregates the streaming data, performs the ETL process in real time, calculates the probability of a user's quote converting to a real sale, and calculates the corresponding commission and commission discount. This is achieved by scoring the current model in production, by passing the grouped factors and levels to the PMML adapter that scores the model on the fly. This end-to-end process takes less than 20 minutes from user data entry to a quote returned to the user. This response time can be tuned to handle any reasonable input velocity via available elastic scaling.

The closed-loop, continuous learning architecture is outlined in Figure 2. Bottom left shows the quote manager request-response, which is served by the online dynamic pricing engine.

This engine includes the current statistical pricing model (GLM or machine learning model such as GBM, random forest, etc.), which is rapidly scored on the incoming data stream (in less than 20ms). Quotes and acceptance/rejection are written to the data warehouse (top left of diagram). As data accumulates, a model rebasing process is triggered. This includes model updates and automated dashboard creation with model summaries (for both champion, challenger model), explanations, and diagnostics. The web-based interactive model summary dashboard is automatically sent to stakeholders, including heads of analytics, pricing, and IT (top right of diagram).

If the rebased challenger model is accepted, it is pushed to the production quote manager system (middle of diagram). Quotes and acceptance/rejection information are also maintained in the TIBCO Spotfire Data Streams in-memory data mart, where real-time streaming data may be visually analyzed by interested parties, such as sales and marketing departments interested in the progress of various marketing campaigns in play.

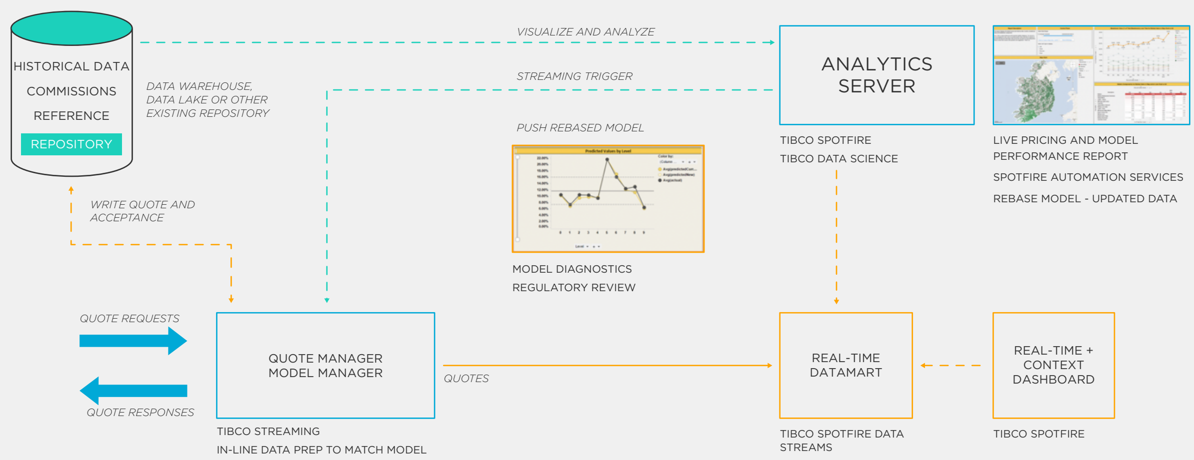

Figure 2. TIBCO algorithmic insurer real-time pricing solution architecture

Step 4: Monitor the Models

Within the TIBCO software, quotes, acceptances, and rejections trigger model retraining. Updated models are reviewed, approved, and ultimately promoted into production. The process can also be automated within the TIBCO software.

Performance Metrics

The ongoing assessment of model performance was based on a set of Gini coefficients. If outside allowable limits, models are re-evaluated automatically using a genetic algorithm for searching the parameter space. Candidate models are regularized with elastic net.

In diagnosing the model, the company assessed the local predictive power and accuracy across the predictor space. Particular regions of the predictor space were assessed in accordance with recent trends to identify areas of concern. Model diagnostic metrics (AIC, BIC, ROC AUC), visualizations (ROC curves, lift charts) and statistical hypothesis test results were weighed.

Retraining and Remodeling

To retrain models, recent data from quotes and acceptance/rejection is assembled, then blended with historical quotes. The accumulated data gets passed to a TIBCO Data Science workflow for retraining.

Updating the models includes assessment of business goals, variable importance, over-fit, bias/outliers, and regulatory perspectives. Segments and factors summarizing the differential model attributes (current versus prior model) are evaluated in a champion-challenger setting. Questions like "Which customers are affected and how?" are constantly assessed.

A triggering mechanism, based on a specific threshold or an external scheduler, kicks off the model management retraining process. The mechanism can be scheduled with user-defined intervals (daily, weekly, etc.) or it can be activated based on events, like reaching a threshold in the amount of data accumulated, in the conversion rate or in the log-loss ratio. The triggered job assembles pertinent model performance, explainability, and diagnostic data. It includes contextual factors like customer risk, quote history, and demographics. Once retraining is complete, TIBCO software sends an email to appropriate business analysts to alert them.

Using the analytic dashboard generated by automation software, the heads of Pricing and IT are also able to review status, add comments, and approve pushing the model update into production. The approval process provides for detailed review, escalation, and resolution, though in most situations this process takes only a matter of minutes or hours.

The Results

Before the insurance company operationalized data science, it had used the same, market-dominant analytics software for many years. Change was far riskier than simply hiring people who knew how to use the software. But since embarking on the TIBCO Model Ops process, the company can hire data scientists with no coding or math background. Within three months, they are able to release models.

Because the TIBCO Model Operations software is easy to learn, the data scientists can give the task of building insurance risk models to people who understand the business, modeling, and math better than IT does. They have at their disposal the computational power to run models for fraud identification and embedded customer value.

Also, whereas all of the company's analytics expertise used to reside in IT, users are empowered and connected across departments now. In case of a sudden rise or drop in sales, for instance, business users would normally approach IT, ask them to look into it and request their findings, with no idea how long it would take to get a useful answer. Now, those same users can simply click on relevant metrics and run a model that sheds light on the change in sales volume. If business users want to dive into the data or perform predictive analysis, they have the opportunity.

Finally, being able to update its models in live environments offers the insurance company real-time predictability and informed decision-making. The insurance company has always had predictive models, but they were subject to three-month update cycles. With model ops and TIBCO software, data goes out to models almost as soon as it comes in, updating them in real time. Rather than generating simple predictions, the company is able to answer questions like "If we increase or decrease discounts, what is the effect on volume and profitability?" It can now understand the total opportunities and risks in the market and make informed decisions.

Even more deeply, the company can ask questions about what to do differently, how to change pricing, and where in the organization to make changes. With a clear view of issues like segmentation, fraud modeling, and underwriter profit, the company plans to take advantage of model ops for long-term predictability of call center capacity, customer relationship management (CRM), campaigns, and ROI.

Global Headquarters
3307 Hillview Avenue
Palo Alto, CA 94304
+1 650-846-1000 TEL
+1 800-420-8450
+1 650-846-1005 FAX
www.tibco.com

05May2020