THE ESSENTIAL GUIDE TO **TIME SERIES FORECASTING**

Part I: The Value of Autonomous Forecasting

INTRODUCTION

Companies across every industry are generating massive amounts of data covering all sorts of metrics. By analyzing the data and finding patterns, these organizations can discover interesting insights about their business. Much of the focus has been on using data insights to improve day to day operations and cybersecurity. Now companies are turning their attention to utilizing business metrics to support better business decision making.

Accurate business forecasts are one of the most important aspects of corporate planning. Machine learning-based forecasting is the most efficient and effective way for a business to project growth of the business and demands on resources.

An emerging field of data science uses time series metrics to enable automated forecasting of what will happen in two business areas: growth and demand. Growth forecasting predicts future sales and revenue, answering questions like, “How much money can we expect to take in next month, next quarter, next year? How many products will we sell?” Demand forecasting predicts how many resources will be needed in the future—resources being things like workers, inventory, supplies, raw materials, monetary funds, and so on, all to meet the forecasted demand.

These are enormously challenging questions to answer using only brain power and rudimentary tools like spreadsheets due to the numerous factors that go into forecasting. There might be hundreds, thousands or even millions of metrics and events that help a business determine what it can expect to see in the future vis-à-vis what has happened in the past. Machine learning (ML) applied to time series data is a much more efficient and effective way to analyze the data, apply a forecasting algorithm, and derive an accurate forecast.

Anodot was founded in 2014 with the purpose of creating a commercial system for real-time analytics and automated outlier detection. To that we have added the means to do autonomous forecasting to predict business growth and demand. It is AI-powered forecasting in a turn-key experience, meaning the solution can be used without needing to have a data scientist to forecast the time series metrics. Anodot’s Autonomous Forecast™ automatically manages the machine learning required to create, train, tune and deploy a customized forecasting model.

Our technology has been built by a core team of highly skilled and experienced data scientists and technologists who have developed and patented numerous machine learning algorithms to isolate and correlate issues across multiple parameters in real time. Now this team has mastered the enormously complex challenge of "productizing" the process of ML-based forecasting such that the end user need not do anything more than point to the time series data source(s), set forth the forecasting parameters, and consume the highly accurate results.

This document, Part I in a three-part series, covers the basics about automated forecasting – what it is, why it's needed, and how some companies are already using it. Part II of this white paper series covers the design principles of creating a forecasting system, as well as the human element of forecasting. Part III delves into the system architecture that Anodot has chosen to deploy for this solution, and a discussion of what it takes to build this type of solution (i.e., build versus buy).

WHAT IS BUSINESS FORECASTING?

Business forecasting is the process of using time series data to develop an educated estimate or prediction of future developments in business such as revenue, sales, and demand for resources and product deliverables. A forecast is based on historical data of a given metric, plus other relevant factors that may be considered. Accurate forecasts are one of the most important aspects of corporate planning. They allow the organization to plan how to budget its funds and allocate its resources efficiently, and they enable decisions such as what strategies to adopt for growth.

Companies like to keep a close eye on their expectations for revenue, preferring to forecast quarterly, or even monthly, if possible, and then comparing the forecasts to even more timely actual results. This gives an opportunity to make necessary course corrections sooner rather than later. One CEO even acknowledges analyzing actual results against forecasts on a weekly basis because it gives his organization many more frequent opportunities to make corrections, leading to better business results.

Many CFOs would like to forecast their revenue projections more frequently but don't do so because their current manual process requires too much time and effort. Business planners and data analysts typically use a spreadsheet for this purpose, but a spreadsheet is a tool that is limited in the amount of data it can handle, the dimensions of data that can be analyzed (typically only two or three at a time), and the data modelling that can be applied. Consequently, the forecasts are high level and often look one year ahead because they won't be updated very frequently. Trying to break the forecasts down into greater granularity – as in, "What are the monthly revenue projections for each of our product lines?" – is too cumbersome and the results are usually not very accurate.

Some companies are fortunate to have a data science team at their disposal to conduct one-time forecasts that are much more accurate than spreadsheet-based predictions. Even at this, the forecast is not "productized" in that there is no underlying system that allows the forecast to be made over and over again at the user's will. Instead, the data scientists perform a one-time analysis that results in a single forecast of a single metric. The forecasting task may require some sophisticated programming and training of algorithms, but the data scientists aren't building a permanent system that can be used long-term. It's a "one and done" effort that might actually take weeks or months to arrive at a forecast. The next time there is a need to forecast the same metric, the process is often repeated from scratch.

The fact is, manual forecasts are complicated, are not timely nor repeatable, and definitely do not scale to offer frequent forecasts or forecasts of numerous metrics.

MACHINE LEARNING-BASED FORECASTING

Anodot's solution Autonomous Forecast now makes it possible to do machine learning-based forecasting in a turn-key manner. Minimal input is needed from the user: a link to the time series data for the metric to be forecast; additional factors to be considered in the prediction (optional); and what is to be forecast, over what horizon. All the work to determine an outcome is self-contained within Anodot's SaaS-based system.[1]

This solution uses an ensemble of machine learning algorithms, including deep learning, to automatically optimize forecasts. It selects a model that's uniquely suited to the chosen business metrics from a library of predictive analytics algorithms. Data feedback is used to train the model for the highest possible accuracy.

In addition to a company's internal data, relevant external data and events can be input into the model to further improve the accuracy of the forecast. For example, consider the case of a home improvement warehouse store that wants to forecast its sales over the next quarter. Sales of homes in the surrounding areas of the store have an impact on the store's sales, as new homeowners are likely to shop for items such as paint, landscaping materials, general household goods, and the like. The store can correlate its own historical sales data with external real estate turnover data to get a more accurate prediction of the store's sales in the coming quarter. This type of modeling would be extremely difficult to do in a manual forecasting process.

Each metric that a company wants to forecast must go through a process to ultimately get to a customized forecast model that can be run again and again or even continuously. The process is depicted in the illustration in **Figure 1**. This is, essentially, the blueprint for creating a time series forecast.

As the image shows, there are several phases to the forecasting process, the most important of which is the training phase.

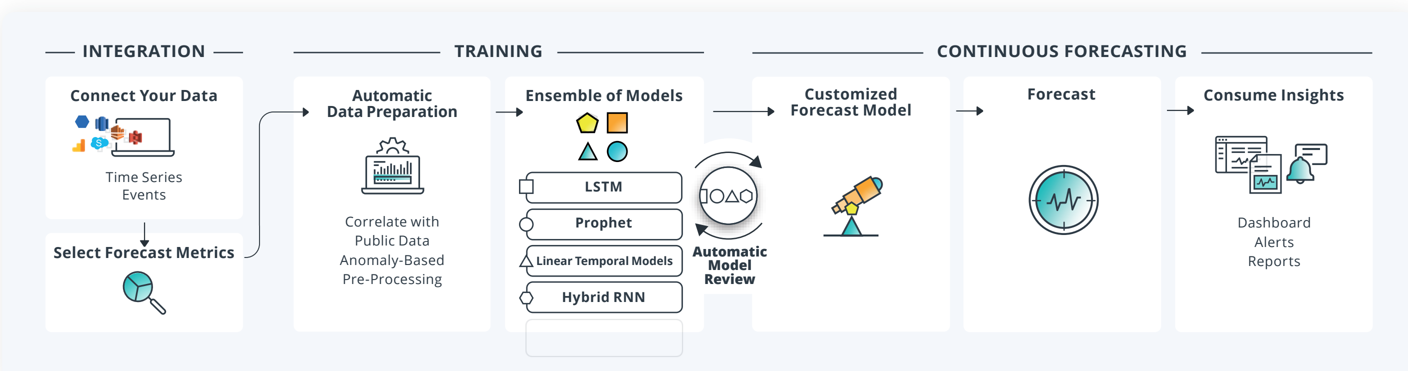


Figure 1. *The Process of Building and Using a ML-Based Customized Forecast Model.*

1. Autonomous Forecast is offered as a service hosted in the cloud, but a customer can choose to install all the product components on-premise is necessary for regulatory compliance purposes.

The Integration Phase

At this first stage, the company brings all the data it desires to work with into the Anodot system. The data can come from a wide variety of sources, as long as it is time series data with sufficient data points to cover at least two “cycles” of the metrics; e.g., multiple years of sales data in order to forecast the next year ahead, or multiple months of service usage to project the need for capacity next month. The number of metrics and events can range into the millions, if necessary (and available). The company selects the specific metric to forecast and specifies the future time horizon; e.g., a week, a month, a quarter, a year, etc. This is the extent of the company’s involvement until the forecast results are available for consumption.

The Training Phase

The first step of this phase involves preparing the data to be run through a variety of algorithms. Internal data is correlated with additional relevant internal and public data, if desired, and scrutinized for anomalies that need to be accounted for in some way in order to not affect the accuracy of the forecast.

The second step involves running the data through numerous learning algorithms and testing to see which ones are most accurate in predicting the future values of that time series. Under the philosophy that many expert opinions are better than just one, Anodot uses an ensemble of the most accurate data models and averages or otherwise combines their results. This yields a customized model that produces the most accurate possible forecast for that metric. After all, accuracy is the most important thing in forecasting, and this is the goal of training.

The Continuous Forecasting Phase

Once the forecast model is derived, it is saved and stored for future use. The organization can use it periodically or in a continuous mode to act on real-time actual data to always be updating the forecast. At the same time, the forecast model is always being fined tuned through an automatic review process.

The Consumption Phase

The company can consume the forecast insights in reports, in a dashboard, as alerts if certain thresholds are met, as input to other systems, etc.

Part II of this white paper series goes into much more detail about the automated forecasting process.

THE ADVANTAGES OF (AND CHALLENGES TO) AN AUTONOMOUS FORECASTING SYSTEM

When compared to the traditional manual methods of business forecasting, a ML-based system has many benefits—the most important of which is accuracy of the forecast. Also, this type of system is able to scale to accommodate enormous amounts of metrics and events, and the forecast models are persistent so that they can be used continuously or repeatedly without the need for data scientists.

We humans attempt to take multiple factors into consideration when we make a forecast but we are limited in our ability to process data from many factors. Even a talented team of data analysts can only handle about four things at a time together to try to figure out whether there's a pattern there that will help make the forecast more precise. But say there are 100 potential factors to consider and the information is input into one spreadsheet or database. Even then, the ability to slice and dice the data to attempt to find relevant patterns is very, very limited.

Machine learning algorithms can take in as inputs the time series the user is trying to forecast, as well as a lot of other potentially relevant data that might help in forecasting that time series. Say an eCommerce company wants to forecast sales of shoes. There's a high likelihood that past patterns of sales of shoes help make this forecast, but there's other related information that might help predict what the sales of shoes will be. For example, sales of socks over the same time period and the number of new users who have registered on the eCommerce website also might have an impact on sales of shoes. In addition, knowing that the store is running a back-to-school promotion in August is important because promotions often help increase sales temporarily.

The advantage of machine learning models is that they can take those inputs – as many as the user wants – and they can figure out patterns at very high dimensional spaces without effort. That's one of the main motivations that makes us believe that these types of models can do a better job than humans at being accurate for forecasting.

The other big advantage of a forecasting system is that, once created, the customized forecasting model is persistently stored so that it can be used repeatedly or continuously. This is definitely not a trivial matter and is discussed in detail in Part III of this white paper series. An extensive system architecture must be built to ensure that models persist for re-use on demand and are not available solely for a one-time use.

CHALLENGES

There can be challenges with the data available as input to the machine learning system. For example, if a company has only one year of sales data available and it wants to forecast the next year of sales, there isn't enough data to make an accurate forecast. At least two full years' worth of sales data – and preferably more than just two years – must be input for the model learning process. What's more, if there is seasonality in the data, there must be enough cycles of the data to account for this seasonality. If a company sells shoes and it knows that sales usually spike during back to school time, the system requires a couple of years of sales data that shows that typical spike. Otherwise, the expected spike in the data pattern isn't obvious.

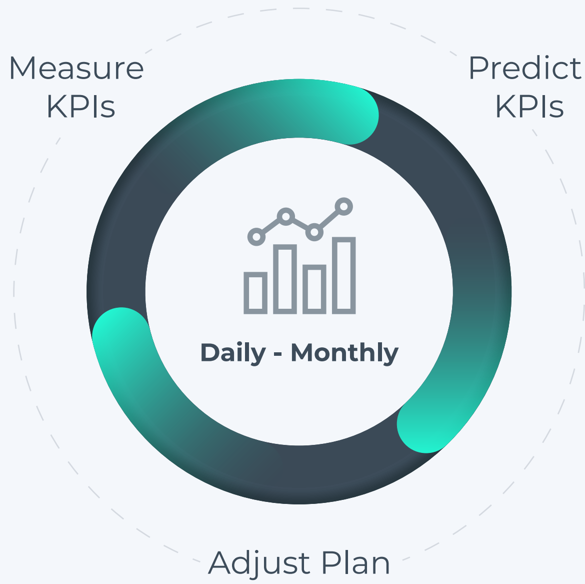
The ability to incorporate hundreds to millions of factors into training the machine learning models is a big advantage over manual forecasting processes, but this capability also has its challenges. Going back to our example of forecasting sales of shoes, suppose the company has a thousand different factors that might influence the forecast. One of the challenges of building a machine learning-based forecasting model is discovering which ones really are influencing factors and which ones are not. The system design must include the ability to discard the irrelevant factors and keep only those that have a potential for helping create a more accurate forecast.

One element of using time series data in machine learning is that there is always a finite amount of data; historical sales data only goes back so far and no further. In training these models with a lot of dimensions or a lot of potential factors, a problem called the Curse of Dimensionality can be encountered. Though it has a funny name, the curse of dimensionality is a legitimate obstacle in machine learning practices. As more data dimensions are added to the machine learning efforts, and given there is a finite amount of data, there is a point where we start having diminishing return in terms of accuracy from adding another data dimension. How Anodot chooses to deal with this situation is described in detail in Part II of this white paper series.

USE CASES FOR GROWTH FORECASTING

Every successful company plans for sustainability and growth. Forecasting the growth path helps companies set their short- and long-term business objectives and make important decisions to help them reach their goals. Short-term forecasts are important in quarterly and annual budget planning and for ensuring that daily business operations help achieve long-term goals. Long-term forecasts help with planning big objectives such as expansion of the company, whether it's into new markets, new facilities or new geographic regions.

Early adopters of our Autonomous Forecast solution have found numerous use cases for growth forecasting. Here are just a few of them.



REVENUE PREDICTION

The most common use of growth forecasting is for accurately predicting revenue over a year, quarter, or perhaps even a month. Every company needs the discipline of revenue forecasting to help develop a budget and ensure that it is not spending more money than it takes in. Moreover, public companies must provide their revenue forecasts to market analysts who use the information to develop investment guidance. Even private companies use revenue projections to encourage investment from venture capitalists, or perhaps to entice acquisition by another company.

Historically, revenue forecasting has been done manually, usually by a data analysis team studying past performance to extrapolate future revenue projections. It's a tedious process, usually resulting in high-level projections that are infrequently updated—maybe quarterly, at best. Accuracy is often an issue; the forecast must be accurate enough so that it isn't far off from what is actually going to happen. Arriving at an accurate forecast is a challenge when limited data metrics go into the formulaic process.

One CFO admitted that when everything is going well and the company is growing and growing, the lack of accuracy in a forecast can be forgiven if revenues overshoot the forecast. The problem is when actual revenues fall short of the forecast, and the company's budget and spending is based on the higher, incorrect number. Many CFOs have had their career derailed when this happened.

Anodot's Autonomous Forecast solution provides highly accurate growth projections because the calculations are based on much more data than simply overall historical performance. External factors can be considered as well; for example, when a new competitor enters the market, or when macro-economic conditions change. What's more, Anodot's forecasts can be very granular, getting down to projections for a specific region or even further for the top 10 or 25 customers in every region.

For example, one of Anodot's Autonomous Forecast customers was able to break their forecast down to multiple dimensions, looking at the revenue projections for the top 100 customers. By looking at what each of these customers would likely produce, the forecasts became much more operational. If a customer showed an expected drop-off, the company could closely plan the management of that account and hopefully reverse the trend toward lower revenue. Having finer granularity forecasts done more frequently creates more opportunity to better manage (and hopefully increase) the revenue streams.

SPENDING ON CUSTOMER ACQUISITION

For most businesses, there is a cost to acquiring new customers. Known as the customer acquisition cost (CAC), this figure is tied to the cost of marketing, advertising and other promotions the company must run in order to attract new customers. It's similar to the revenue prediction but here it is even more operational. If a company can forecast its revenue and also how much it is spending on customer acquisition, then its actions on marketing programs can be much more detailed, targeted and effective.

Accurately forecasting CAC depends on taking in historical data of how much a company spent per channel of marketing, such as digital channels (e.g., Facebook ads, Google ad words, etc.) as well as traditional channels like radio, TV, print and events. The company would calculate how much it spends on all these channels and how many customers were acquired based on each channel (conversion rates).

Performance metrics from each of these channels help create a more accurate forecast. For example, with Facebook ads, metrics could include how many views the ad had, how many likes, how many comments, how many clicks, etc. Such engagement metrics help forecast how many customers will be acquired through that channel. Anodot's machine learning algorithms can take in a lot of metrics and determine correlations among them to provide a highly accurate forecast on the cost of customer acquisition.

CUSTOMER CHURN

A cellular service provider wants to forecast the number of customers it can expect to cancel their subscription over the next quarter. This is called customer churn and it happens to all types of businesses that offer subscription-based services. Understanding the potential for loss of customers is critical to knowing how this can impact expected revenue. This knowledge also is important for making decisions that might possibly help to retain customers, such as lowering prices or offering more services at the current price level. In most cases, retaining existing customers is more cost-effective than attempting to acquire new customers.

Churn can be forecast not just for overall numbers but also per geography or by different types of customers like family plans. Each churn prediction could lead to different actions that can be taken to try to hold onto customers.

To forecast how much churn the service provider can expect, the company can look at past churn patterns and base its forecast just on that, but this method often falls short and is not very accurate. The results can be much more accurate if the company takes into consideration a number of influencing factors that indicate customers are likely to leave the service.

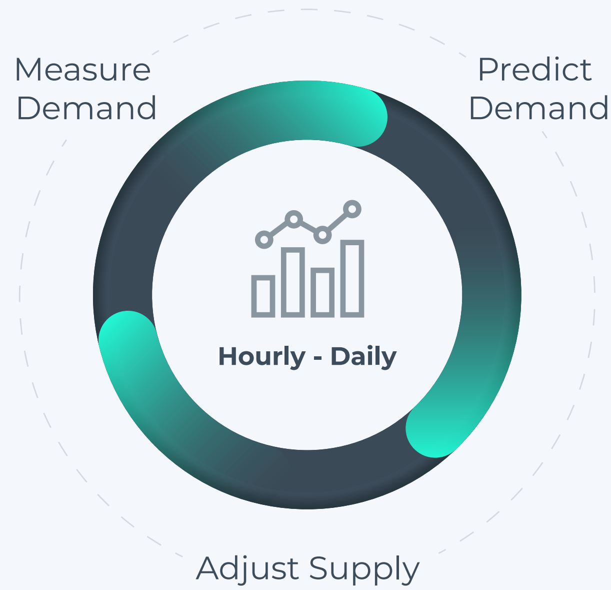
For example, how many support calls did the company get in the last month? On a daily basis, how many support calls came into the company's customer service system? If the company observes an uptick in support calls, then it's likely to see more churn in the coming month. It doesn't have to be true but the algorithm will figure out, based on historical patterns, if the correlation exists or not.

Anodot's system can use even more metrics to make the forecast. For example, how many outages in the company's network occurred in a certain region? Were there performance issues with the network for all or a subset of the subscribers? How many customer complaints showed up in social media? What pricing schemes do competitors offer? Many different metrics can go into the Anodot algorithms as input and the algorithms will figure out if they are relevant or not.

This cellular service provider learned that when training a forecasting model to learn how to forecast, entering information about competitors' events and marketing campaigns allows the model to learn the effect of competitors on its own amount of churn. For example, last month a competitor launched a big promotion for a family plan. The year before they had some other marketing campaign. The algorithm can learn whether those external activities had an effect on the company's churn rates. As long as the pattern is there, the algorithm should be able to recognize it.

USE CASES FOR DEMAND FORECASTING

Anodot has been working with a wide range of early adopters across numerous industries who are improving their ability to predict demand by using our Autonomous Forecast system. Here are just a few of the many use cases that companies are discovering.



FINTECH AND CURRENCY MANAGEMENT

A growing number of financial technology (fintech) companies are getting into the eCommerce business of facilitating payments between parties through online funds transfer. Many of these companies operate in multiple countries or even globally, which introduces the complexity of dealing with multiple currencies and exchange rates. There's a need to accurately forecast the demand for funds in order to ensure there is money in the bank when customers want to withdraw their funds, without allocating too much money for this purpose, which takes away the opportunity to earn interest on the float. Also, the company can lose money on currency exchange rates. Determining the optimal bank balances based on true demand is a challenge.

Anodot has been working with a payments company that has multiple bank accounts in multiple countries and multiple currencies. The company's treasury manager wants to know, how much money do I need to allocate to each currency and each country, and should the allocations be every day or every week? These are very complex questions because there are many factors involved, including the currency exchange rate and the amount of time to execute a transfer of funds to various countries. One of the questions is to establish how much money people are going to withdraw from their accounts in the coming days.

Anodot Demand Forecast is able to help this payments company solve their case of supply and demand: the demand of the customers who are withdrawing funds from their accounts and how the company supplies just the right amount of cash to cover those accounts and support the customers without tying up too much cash at any given time.

In a similar use case, banks with automated teller machines (ATMs) want to know how much money to put in each machine in order to cover customer demand, but without putting too much money in the devices and losing the opportunity to use the funds for other purposes. A large bank typically has more than 5,000 ATMs. An ATM may have up to $200,000 in it, but let's assume $50,000 is the average amount. That's $250 million locked up in these machines. With an ML-based forecast for each ATM that reduces the amount of cash in it by just 10%, the bank could have $25 million extra funds available to lend, invest and profit from each day. With a 25% improvement (which is not hard to achieve), it is already over $60 million extra funding on the banks books.

NETWORK CAPACITY PLANNING

In the telecom world, a cell is the geographic area that is covered by a single base station in a cellular network. A network for wireless communications is comprised of a large number of base stations to efficiently use the radio spectrum to cover the service area. Each cell can handle a certain number of simultaneous connections and support a finite bandwidth. As the population density within a geographic area grows and more people are using applications like YouTube, Facebook and online games, that cell can quickly reach its maximum capacity (connections or internet browsing), which is the point where calls cannot go through or applications on users' phones become slow—resulting in poor service.

The installation of new cells and base stations has to be planned months in advance, as it takes time to acquire and install the equipment. So, telcos need to accurately forecast the demand for the capacity of their network in order to plan for future needs. This is especially important in a geographic region where the population is rising rapidly, such as through new home building or even a new building opening up that will be host to a lot of people during certain hours—say, an office, a hospital or a school.

Most telcos today use a manual process to forecast demand for cell capacity. Once a quarter, a team comes together to analyze the bandwidth and usage of individual cells in a given area. They do the capacity planning using spreadsheets; it's complex and often not very accurate.

Anodot simplifies the process by not only looking at historical internal usage data, but also by including external data such as housing starts in a given region, and openings of or future plans for population-dense facilities such as hospitals, schools and office buildings. By observing or anticipating population changes, the telco can more accurately plan for capacity demand with sufficient lead time for installation of new cells or base stations.

Telco companies will need to put their capacity planning efforts on steroids as the world begins migrating to 5G as the cellular transmission standard. The Internet of Things is expected to have a huge impact on 5G demand, and thus telcos will need to prepare well in advance to avoid capacity constraints and will need to provide frequent or even continuous updates to their demand forecasts.

INVENTORY PLANNING FOR RETAIL

There's more pressure than ever on retail stores to get it right when it comes to inventory planning. No retail establishment wants to lose sales because items are out of stock, or conversely, be stuck with a bunch of items that just don't sell at full price. In this hyper-sensitive climate for retail, the store or company that does a poor job of inventory planning will not stay competitive for long.

Consider the case of a nationwide home improvement store. Inventory planning is a complex task that needs to take into consideration many factors, such as seasonal sales (garden plants in spring, snow shovels in winter); geographic location (those snow shovels won't sell well in Florida); local real estate conditions (older homes need repairs and updates, newly constructed homes need accessories and basic home supplies); the pace of area home sales (new homeowners want a fresh coat of paint); and so on. There are some things that every store across the country will need to stock (like generators and grills), and other things that only certain stores would carry (like brackets for hurricane window coverings). Stores in upscale zip codes will want to stock higher-end products like designer paints and fixtures. Even regional events like floods and hail storms have an impact on what a store can sell and should stock.

Anodot can take all these considerations into account and produce forecast demands on a store basis or regional basis. Categories of products – paints, plumbing, electrical, etc. – can be projected. The more accurate the inventory planning for each store, the higher the customer satisfaction will be, and most importantly, the higher the product margins and revenue will be for the retailer.

Another industry that has a strong need for capacity planning is shipping/product delivery. As eCommerce continues to grow, more items are being delivered directly to customers. Whether it's a package from Amazon or a meal from DoorDash, the operating companies need to accurately project the need for their delivery services in order to schedule enough drivers and vehicles and to know there is sufficient capacity for packages in the vehicles. Amazon might use efficient small vans to deliver packages during slow-season months but move up to large-capacity trucks in the run-up major holidays.

SUMMARY

Accurate business forecasts are one of the most important aspects of corporate planning. Companies have traditionally made their forecasts via a cumbersome manual process, often using poorly suited tools like spreadsheets or databases. Now the emerging discipline of machine learning-based forecasting offers the most efficient and effective way for an organization to project growth of the business and demands on its resources.

Anodot has "productized" the process of ML-based forecasting by developing a sophisticated system to automatically create, train, tune and deploy a customized forecasting model in a turnkey experience. Creating this system was no trivial matter; the design principles and system architecture are described in detail in parts II and III of this white paper series. The resulting product, Anodot Autonomous Forecast, enables business users to create a customized forecast model for a given business metric and use that model on demand or continuously to create an extremely accurate forecast of the metric.

In part II of this series, learn more about time series forecasting

Download Part II

For more information, please contact Anodot:

North America
669-600-3120
info.us@anodot.com

International
+972-9-7718707
info@anodot.com