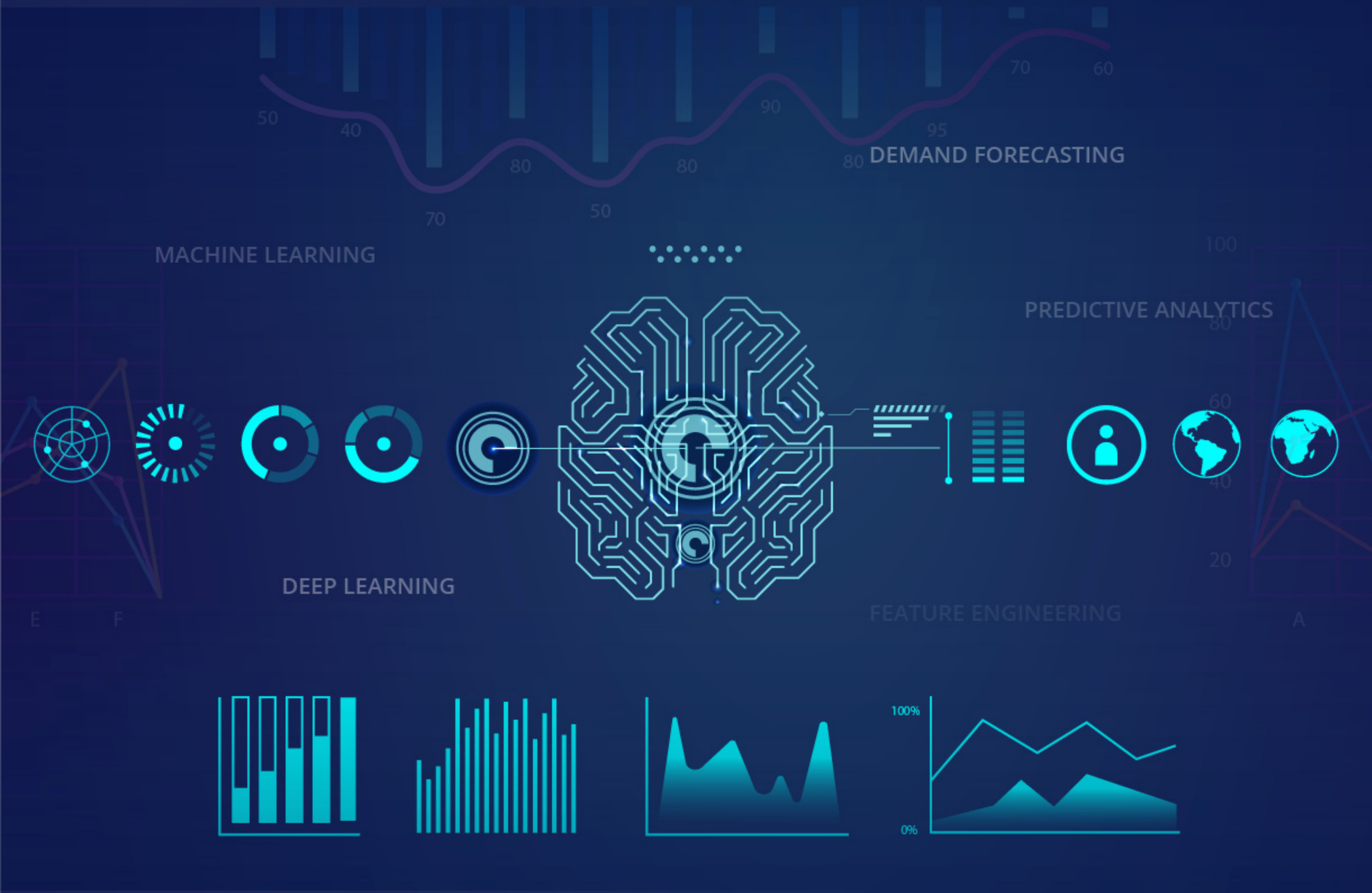
MACHINE LEARNING IN DEMAND FORECASTING FOR RETAIL



written by

Liudmyla Taranenko

Data Science Engineer at MobiDev

mobidev

Upwork 2016 · Upwork 2017 · Upwork 2018

Best Upwork Software Development Agency In Ukraine

COMPLEX SOFTWARE DEVELOPMENT

WITH A FOCUS ON INNOVATION

WEB & CLOUD INFRASTRUCTURE

AR & MOBILE APPS

IOT & HARDWARE INTEGRATION

DATA SCIENCE & MACHINE LEARNING

For Startups

For Emerging companies

For Enterprises

Guaranteed delivery
on time and on budget
No surprises

You can **adapt to evolving business needs and increase ROI** with our flexible, proven processes

Top US-level quality for **1/3 the price** to bring **3x features** to your product

350+ Products launched

100% Approval rating by Upwork

300+ English speaking professionals

Find more at **www.mobidev.biz**

info@mobidev.biz

+1 888 380 0276

UNITED STATES OFFICE
Atlanta, Georgia

UNITED KINGDOM OFFICE
Sheffield

ENGINEERING OFFICES IN UKRAINE
Kharkiv, Chernivtsi, Mykolaiv

Table of Contents:

I know for sure that human behavior could be predicted with data science and machine learning. People lie—data does not. Taking a look at human behavior from a sales data analysis perspective, we can get more valuable insights than from social surveys.

In this article, I want to show how machine learning approaches can help with customer demand forecasting. Since I have experience in building forecasting models for retail field products, I'll use a retail business as an example.

Moreover, considering uncertainties related to the COVID-19 pandemic, I'll also describe how to enhance forecasting accuracy.

And don't worry if your business's focus isn't on retail. The main goal of this article is to describe the logic of how machine learning can be applied in demand forecasting both in a stable environment and in crisis.

What Is Demand Forecasting in Machine Learning?

Machine learning techniques allow for predicting the amount of products/services to be purchased during a defined future period. In this case, a software system can learn from data for improved analysis.

Compared to traditional demand forecasting methods, machine learning:

- Accelerates data processing speed
- Provides a more accurate forecast
- Automates forecast updates based on the recent data
- Analyzes more data
- Identifies hidden patterns in data
- Creates a robust system
- Increases adaptability to changes

According to technology trends in the retail sphere, demand forecasting is often aimed to improve the following processes:

- **Supplier relationship management**. By having the prediction of customer demand in numbers, it's possible to calculate how many products to order, making it easy for you to decide whether you need new supply chains or to reduce the number of suppliers.
- **Customer relationship management**. Customers planning to buy something expect the products they want to be available immediately. Demand forecasting allows you to predict which categories of products need to be purchased in the next period from a specific store location. This improves customer satisfaction and commitment to your brand.
- **Order fulfillment and logistics.** Demand forecasting features optimize supply chains. This means that at the time of order, the product will be more likely to be in stock, and unsold goods won't occupy prime retail space.
- **Marketing campaigns.** Forecasting is often used to adjust ads and marketing campaigns and can influence the number of sales. Sophisticated machine learning forecasting models can take marketing data into account as well.
- **Manufacturing flow management.** Being part of the ERP, time series-based demand forecasting predicts production needs based on how many goods will eventually be sold.

Design Algorithm for ML-Based Demand Forecasting Solution

When initiating the demand forecasting feature development, it's recommended to understand the workflow of ML modeling. This offers a data-driven roadmap on how to optimize the development process.

Let's review the process of how we approach ML demand forecasting tasks.

Step 1. Brief Data Review

The first task when initiating the demand forecasting project is to provide the client with meaningful insights.

Related article: Business Analysis Deliverables List For Software Development Projects

The process includes the following steps:

1. Gather available data
2. Briefly review the data structure, accuracy, and consistency
3. Run a few data tests and pilots

In my experience, a few days is enough to understand the current situation and outline possible solutions.

Step 2. Setting Business Goals and Success Metrics

This stage establishes the client's highlights of business aims and additional conditions to be taken into account. Our team provides data science consulting to combine it with the client's business vision. The goal is to achieve something similar to:

"I want to integrate the demand forecasting feature so to forecast sales and plan marketing campaigns."

Success metrics offer a clear definition of what is “valuable” within demand forecasting. A typical message might state:

“I need a machine learning solution that predicts demand for […] products, for the next [week/month/a half-a-year/year], with […]% accuracy.”

These points will help you to identify what your success metrics look like. You will want to consider the following:

Product Type/Categories

What types of products/product categories will you forecast? Different products/services have different demand forecasting outputs. For example, the demand forecast for perishable products and subscription services coming at the same time each month will likely be different.

Time Frame

What is the length of time for the demand forecast?

Short-term forecasts are commonly done for less than 12 months - 1 week/1 month/6 month. These forecasts may have the following purposes:

- Uninterrupted supply of products/services
- Sales target setting and evaluating sales performance
- Optimization of prices according to the market fluctuations and inflation
- Finance maintenance
- Hiring of required specialists

Long-term forecasts are completed for periods longer than a year. The purpose of long-term forecasts may include the following:

- Long-term financial planning and funds acquisition
- Decision making regarding the expansion of business
- Annual strategic planning

Accuracy

What is the minimum required percentage of demand forecast accuracy for making informed decisions?

Implementing retail software development projects, we were able to reach an average accuracy level of 95.96% for positions with enough data. The minimum required forecast accuracy level is set depending on your business goals.

The example of metrics to measure the forecast accuracy are MAPE (Mean Absolute Percentage Error), MAE (Mean Absolute Error) or custom metrics.

General Metrics

- MAPE (Mean Absolute Percentage Error) - a statistical measure of how accurate a forecast system is. It measures accuracy as a percentage:

$$\mathrm{M} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$$

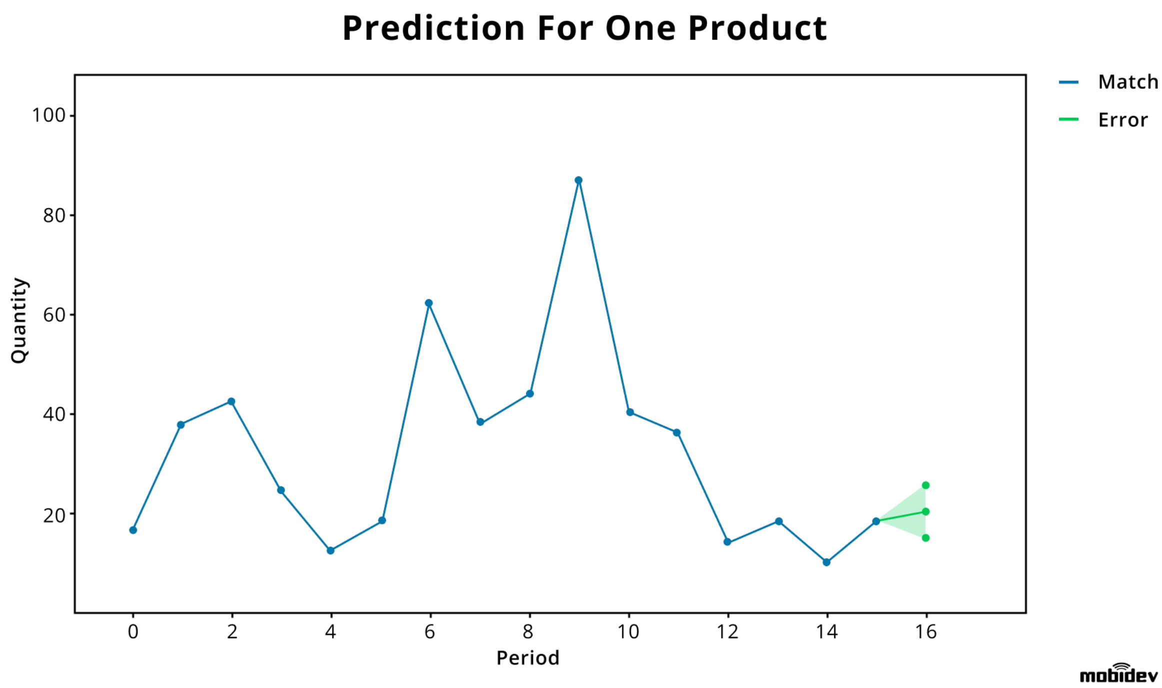
  Where A_t is the actual value and F_t is the forecast value. It is the most common measure used to forecast error and works best if there are no extremes to the data (and no zeros).
- MAE (Mean Absolute Error) - a measure of difference between two continuous variables. Assume *X* and *Y* are variables of paired observations that express the same phenomenon. They're comparisons of predicted versus observed.

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|$$

Custom Metrics:

The percentage of products with an error of less than 1, 3, or 5.

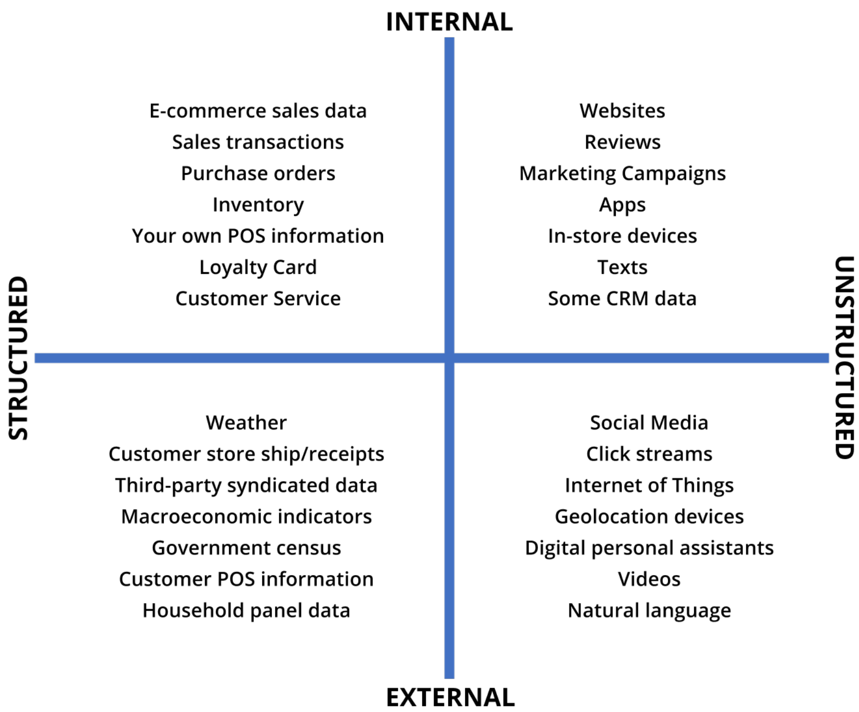
The example of the Prediction Error (highlighted in green):



Step 3. Data Preparation & Understanding

Regardless of what we'd like to predict, data quality is a critical component of an accurate demand forecast.

This following data could be used for building forecasting models:



Source: IBF (Institute of Business Forecasting and Planning)

Data Quality Parameters

When building a forecasting model, the data is evaluated according to the following parameters:

- **Consistency.** The collected data should be contradictory. The mechanism of collecting and storing the data should exclude data contradictions and be stable.
- **Accuracy.** The data should be exact. It should not have any ambiguity and uncertainty. The accuracy of the demand forecasting model directly depends on how accurate the data is.

- **Validity.** Every organization defines its own meaning of the "validity" term. This data type should be customized according to the business's purpose. For example, analyzing people's gender and nationality for input options is limited, and other information is not permitted. Any information other than this will be considered as invalid according to a system requirement.
- **Relevance.** The data has to be collected at the right time. Untimely data distorts the forecast's accuracy.
- **Accessibility.** When gathering data sources, it's possible to face legal and regulatory constraints. The level of access to data should be proper and legitimate. This allows for more qualitative data processing when building forecasting models.
- **Completeness.** The overall forecast picture can be displayed only if the data is comprehensive. The requirements for "completeness" are explicitly defined for each case.
- **Detalization.** It's not recommended to summarize and aggregate data when collecting and sorting it. Before receiving any data, it's important to define any requirements you have for the level of detail for your data. If your data isn't detailed enough, the forecasting model could provide invalid results.

Related article: Natural Language Processing (NLP) Use Cases for Business Optimization

In reality, the data collected by companies often isn't ideal. This data usually needs to be cleaned, analyzed for gaps and anomalies, checked for relevance, and restored. When developing POS application for our retail clients, we use data preparation techniques that allow us to achieve higher data quality.

Data Preparation Tactics:

- **Getting More Data**. For some hypotheses, there may be a need to add external data like currency exchange course or weather data over a period of time. Another example is to use a video of users' in-store behavior.
- **Inventing More Data**. Some existing data could be augmented or permuted. There are certain probabilistic techniques to generate "synthetic data." In some cases, this may be a good option. This technique is commonly used when there is a lack of real data.
- **Data Cleaning**. There are often some incomplete data elements that are not able to add value. This data can be removed—but there may be sensitive information required for making forecasts. In cases like that, data can be fixed. This process is called data cleaning.
- **Data Resampling.** Data can be resampled to change its size or distribution. For example, using over-sample/under-sample observations of a specific type, it's possible to represent data in a dataset better or use a smaller data sample to speed the process up.
- **Problem Reframing.** Reframing the problem involves exploring alternative perspectives on what is to be predicted. This technique helps to capture different information and turn the prediction results into a more valuable forecast for a business. Problem reframing is used when the existing data does not allow you to achieve the expected output.
- **Data Rescaling.** Normalization or standardization approaches can improve the algorithms' performance when using distance measures, weighted inputs, or other numeric input variables.
- **Data Transformation.** By passing data through the exponential function or making it more Gaussian, it's possible to change data distribution. This approach may help a learning algorithm recognize features in the data more easily.

- **Data Projection.** Projection and unsupervised clustering allow you to create a new, compressed representation of a dataset**.** This method is ideal when you have a large amount of data. Use it if finding hidden correlations is important, or you want to group data and highlight its main characteristics.
- **Feature Selection.** To determine whether all input data is equally important, use feature importance and feature selection methods. This may be helpful when creating new views of the data to explore with modeling algorithms.
- **Feature Engineering.** This tactic is all about creating and adding new data features. If there are any attributes like dates or categories, they can be decomposed into multiple new values. For example, you could make sales performance analysis by the day, month, or year.

Once the data is cleaned, generated, and checked for relevance, we structure it into a comprehensive form. Below, you can see an example of the minimum required processed data set for demand forecasting:

Transactions Forecasting

Date	Product name	Quantity (Target)
2019-06-06 04:00:15	lime	500
2019-06-06 10:00:00	lime	100
2019-06-06 10:30:11	apple	250
2019-06-06 10:35:08	orange	300
2019-06-06 10:40:01	banana	750
...	...	...

Data understanding is the next task once preparation and structuring are completed. It’s not modeling yet but an excellent way to understand data by visualization.

Below you can see how we visualized *the data understanding process:*

Product Sales For The Last Year

Step 4. Machine Learning Models Development

There are no "one-size-fits-all" forecasting algorithms. Often, demand forecasting features consist of several machine learning approaches.

Related article: [Optical Character Recognition Based on Machine Learning Technology]

The choice of machine learning models depends on several factors, such as business goal, data type, data amount and quality, forecasting period, etc.

Here I describe those machine learning approaches when applied to our retail clients. But if you have already read some articles about demand forecasting, you might discover that these approaches work for most demand forecasting cases.

Time Series Approach

This involves processed data points that occur over a specific time that are used to predict the future. Time series is a sequence of data points taken at successive, equally-spaced points in time.

The major components to analyze are:

- Trend
- Seasonality
- Irregularity
- Cyclicity

The analysis algorithm involves the use of historical data to forecast future demand. That historical data includes trends, cyclical fluctuations, seasonality, and behavior patterns.

In the retail field, the most applicable time series models are the following:

1. ARIMA (auto-regressive integrated moving average) models aim to describe the autocorrelations in the time series data. When planning short-term forecasts, ARIMA can make accurate predictions. By providing forecasted values for user-specified periods, it clearly shows results for demand, sales, planning, and production.

2. SARIMA (Seasonal Autoregressive Integrated Moving Average) models are the extension of the ARIMA model that supports univariate time series data involving backshifts of the seasonal period.

3. Exponential Smoothing models generate forecasts by using weighted averages of past observations to predict new values. The essence of these models is in combining Error, Trend, and Seasonal components into a smooth calculation.

DEMAND FORECASTING CASE STUDY: VENUE MANAGEMENT & POS SYSTEM

Client:

SmartTab, San Francisco-based POS provider

Product:

Cloud-based venue management & POS system to high volume bars, restaurants, and nightclubs

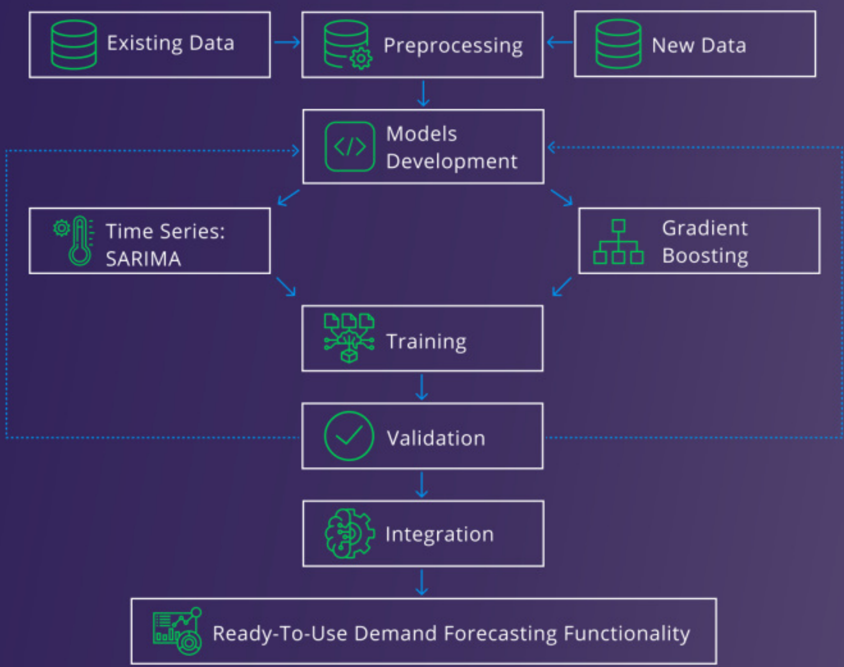
Business Goals:

Forecast revenue, track sales performance

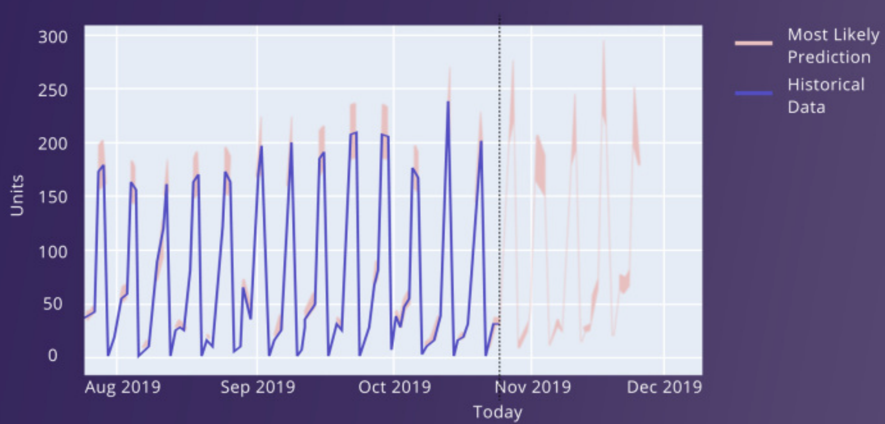
We Predict:

How many drinks, meals, and snacks will be sold within the next week/month/half-a-year in each restaurant, bar, and nightclub (180 in general)

DEMAND FORECASTING MODELS INTEGRATION



FORECAST ACCURACY



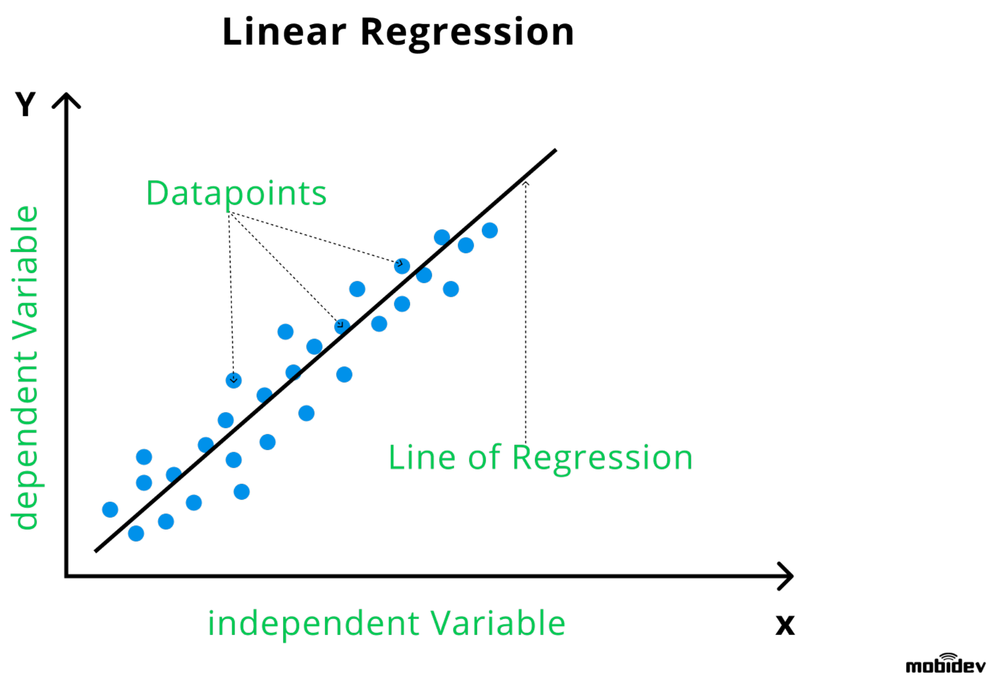
mobidev

Let's say you want to forecast the demand for vegetables in the next month. For a time series approach, you require historical sale transaction data for at least the previous three months. If you have historical data about seasonal products - vegetables in our case - the best choice will be the SARIMA model. The forecast error, in that case, maybe around 10-15%.

Linear Regression Approach

Linear regression is a statistical method for predicting future values from past values. It can help determine underlying trends and deal with cases involving overstated prices.

An example of Linear Regression:



This regression type allows you to:

1. Predict trends and future values through data point estimates.
2. Forecast impacts of changes and identify the strength of the effects by analyzing dependent and independent variables.

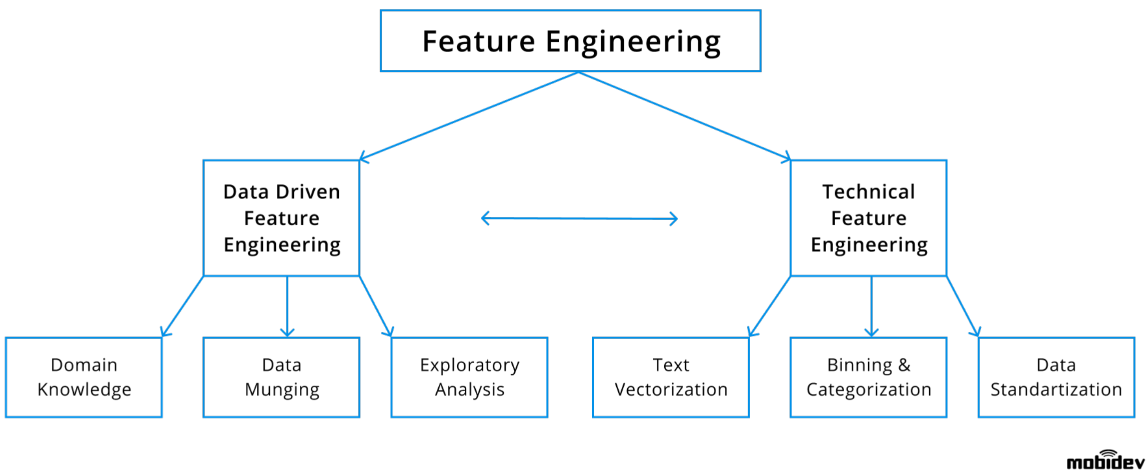
Let's say you want to calculate the demand for tomatoes based on their cost. Assuming that tomatoes grow in the summer and the price is lower because of high tomato quantity, the demand indicator will increase by July and decrease by December.

The information required for such type forecasting is historical transaction data, additional information about specific products (tomatoes in our case), discounts, average market cost, the amount in stock, etc. The forecast error may be 5-15%.

Feature Engineering

Feature engineering is the use of domain knowledge data and the creation of features that make machine learning models predict more accurately. It enables a deeper understanding of data and more valuable insights.

An example of Feature Engineering Approach:



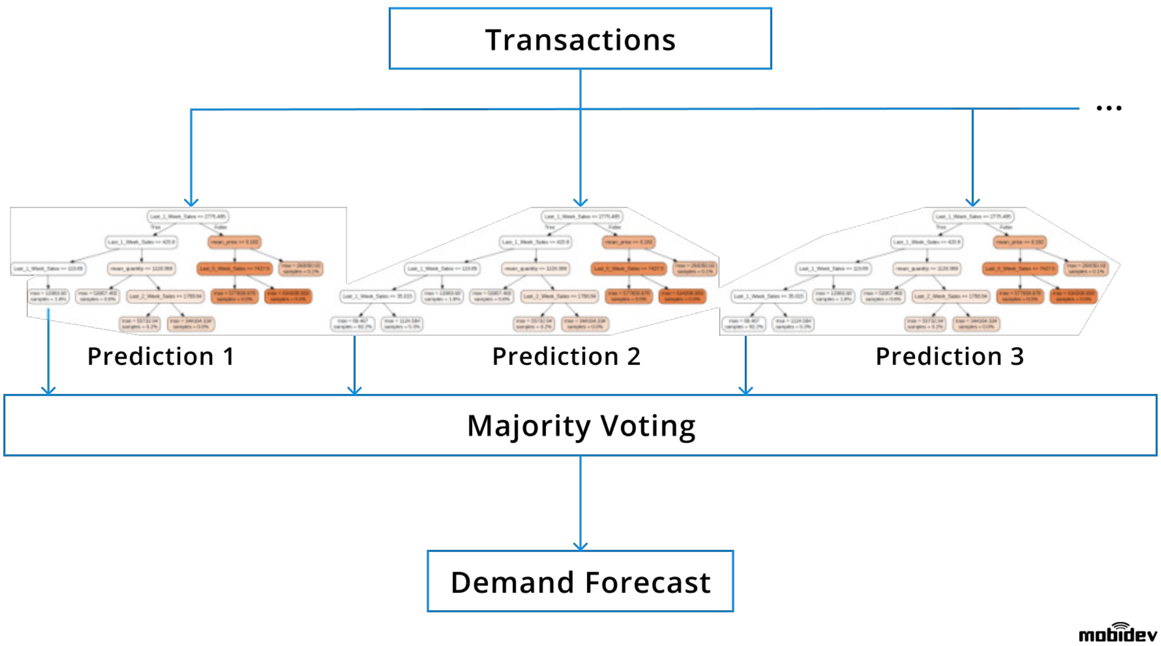
Since feature engineering is creating new features according to business goals, this approach is applicable in any situation where standard methods fail to add value. In **custom ML modeling**, a data scientist builds new features from existing ones to achieve higher forecast accuracy or to get new data.

Random Forest

The basic idea behind the random forest model is a decision tree. The decision tree approach is a data mining technique used for data forecasting and classification. The decision tree method itself does not have any conceptual understanding of the problem. It learns from the data we provide it.

Random forest is the more advanced approach that makes multiple decision trees and merges them together. By taking an average of all individual decision tree estimates, the random forest model results in more reliable forecasts.

An example of the Random Forest approach:



Random forest can be used for both classification and regression tasks, but it also has limitations. The model may be too slow for real-time predictions when analyzing a large number of trees.

If you have no information other than the quantity data about product sales, this method may not be as valuable. In such cases, the time series approach is superior.

DEMAND FORECASTING CASE STUDY: RETAIL ERP & POS

Client:

Comcash, US-based B2B product company with 20 years in the retail sphere

Product:

Cloud-based ERP & POS desktop and mobile software

Business Goal:

Provide the Comcash's clients with the forecast of sales for planning marketing campaigns

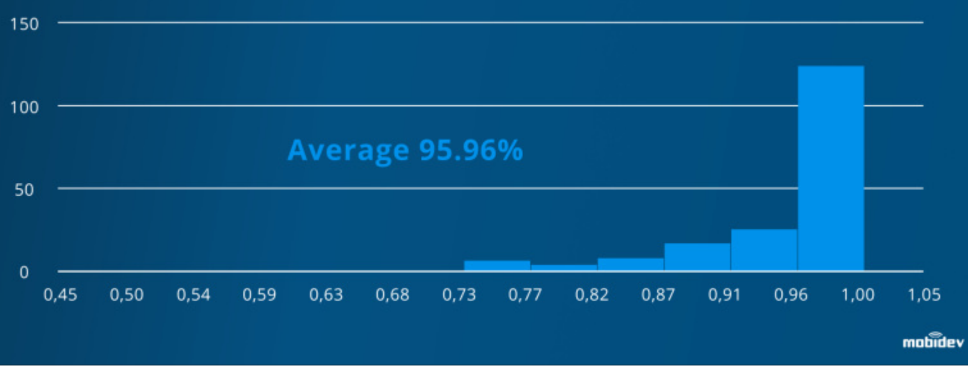
We predict:

How much food, clothes, and computer equipment will be sold in each supermarket within the next week

DEMAND FORECASTING MODELS INTEGRATION

FORECAST ACCURACY

Accuracy Distribution For All Chains



Step 5. Training & Deployment

Training

Once the forecasting models are developed, it's time to start the training process. When training forecasting models, data scientists usually use historical data. By processing this data, algorithms provide ready-to-use trained model(s).

Validation

This step requires the optimization of the forecasting model parameters to achieve high performance. By using a cross-validation tuning method where the training dataset is split into ten equal parts, data scientists train forecasting models with different sets of hyperparameters. The goal of this method is to figure out which model has the most accurate forecast.

Improvement

When researching the best business solutions, data scientists usually develop several machine learning models. Since models show different levels of accuracy, the scientists choose the ones that cover their business needs the best.

The improvement step involves the optimization of analytic results. For example, using model ensemble techniques, it's possible to reach a more accurate forecast. In that case, the accuracy is calculated by combining the results of multiple forecasting models.

Deployment

This stage assumes the forecasting model(s) integration into production use. We also recommend setting a **pipeline to aggregate new data** to use for your next AI features. This can save you a lot of data preparation work in future projects. Doing this also increases the accuracy and variety of what you could be able to forecast.

COVID-19 as an Anomaly: How to Forecast Demand in Crisis

When integrating demand forecasting systems, it's essential to understand that they are vulnerable to anomalies like the COVID-19 pandemic. It means that machine learning models should be upgraded according to current reality.

As the demand forecasting model processes historical data, it can't know that the demand has radically changed. For example, if last year, we had one demand indicator for medical face masks and antiviral drugs, this year, it would be completely different.

In that case, there might be several ways to get an accurate forecast:

- **Collect data about new market behavior.** Once the situation becomes more or less stable, develop a demand forecasting model from scratch.
- **Apply feature engineering approach.** By processing external data: news, a current market state, price index, exchange rates, and other economic factors, machine learning models are capable of making more up-to-date forecasts.
- **Upload the most recent POS data.** The period of a loadable dataset might vary from one to two months, depending on the products' category. In this way, we can timely detect shifts in demand patterns and enhance forecast accuracy.
- **Apply the transfer learning approach.** If there are any gathered historical data about past pandemics or similar behavior shifts, we can take them and predict demand in the context of the current crisis.
- **Apply information cascade modeling approach.** Combining the most recent POS data with the cascade modeling, the demand forecasting system can identify herd patterns of human behavior. In other words, we can forecast how people will make buying decisions according to the behavior patterns of most people.

- **Apply natural language processing (NLP) approach.** NLP technology enables the processing of real comments from social networks, media platforms, and other available social sources. By utilizing text mining and sentiment analysis approaches, NLP models gather samples of customer's conversations. This method allows the detection of people's preferences, choices, sentiment, and behavior shifts.

Machine learning is not limited to demand forecasting. The future potential of this technology depends on how well we take advantage of it. Today, we work on demand forecasting technology and understand what added value it can deliver to modern businesses. Still, we never know what opportunities this technology will open for us tomorrow.

CONTACT US

info@mobidev.biz

mobidev.biz