Machine Learning Methods for Sales Predcition

# regression analysis :

* One of the most common methods used to predict sales is **regression analysis.** This method involves using historical sales data to train a model that can predict future sales. The model can take into account factors such as **past sales, marketing campaigns, and economic indicators** to make its predictions.

# time series analysis :

* Another popular method for predicting sales is **time series analysis**. This method involves using historical sales data to identify patterns and trends in sales over time. The model can then use these patterns to make predictions about future sales. This method is particularly useful for predicting sales in seasonal industries, such as retail and tourism.

# decision tree-based algorithms :

* Another approach is using **decision tree-based algorithms** like **Random Forest, Gradient Boosting** etc. These algorithms are particularly useful when there are many factors that can influence sales, such as product features, customer demographics, and market conditions. The algorithm can help identify the most important factors and use them to make predictions.

# neural networks :

* In addition to these methods, machine learning can also be used to predict sales through the use of **neural networks.** Neural networks are a type of machine learning algorithm that can learn to recognize patterns in data. They can be trained on large amounts of sales data and can make predictions about future sales.

# 1.DATA PREPROCESSING IN SALES DATASET

# Data preprocessing is an important step in the data mining process that involves cleaning and transforming raw data to make it suitable for analysis. Some common steps in data preprocessing include:

# Data Cleaning:

# This involves identifying and correcting errors or inconsistencies in the data, such as missing values, outliers, and duplicates. Various techniques can be used for data cleaning, such as imputation, removal, and transformation.

# Data Integration:

# This involves combining data from multiple sources to create a unified dataset. Data integration can be challenging as it requires handling data with different formats, structures, and semantics. Techniques such as record linkage and data fusion can be used for data integration.

# Data Discretization:

# This Involves Dividing Continuous Data Into Discrete Categories Or Intervals. Discretization Is Often Used In Data Mining And Machine Learning Algorithms That Require Categorical Data. Discretization Can Be Achieved Through Techniques Such As Equal Width Binning, Equal Frequency Binning, And Clustering.

And there are many more data preprocessing techniques that helps in obtaining of proper data set

2.CHOOSING THE RIGHT MODEL

The two types of sales forecasting process are generally split into two groups: quantitative sales forecasting and qualitative sales forecasting

QUANTITATIVE SALES FORECASTING :

The so called quantitative methods of sales forecasting are those used with the availability of historical sales data that can be extrapolated to predict future revenue. These methods rely more on sound, mathematical equation than opinionated judgement from expert peers.

some of the most popular techniques include:

•Trend analysis : The idea here is that through the study of past sales data you can pick up on certain trends that with reason, could be used to predict similar fluctuations in the future. This could either be from seasonality, random factor analysis and economic demand.

•Exponential Smoothing : Probably considered the most accurate and widely used for an accurate sales forecasting process it makes an exponentially considered average of past sales to try and predict future revenue.

•Simple Moving Average : This Technique Requires The Sales Manager To Extrapolate Sales Data From A “Dynamic” Set Period Of Time; A Rolling Window Of Maybe 2,3 Or Maybe Even 6 Months.

So if you **have sales data available** research the pros and cons for each of these techniques to find one that best suits your business model.

3. **SALES FORECASTING TOOLS**

### **\*using spreadsheets for sales forecasting :**

Organizations typically use spreadsheets for [sales forecasting](https://www.clari.com/blog/sales-forecasting-guide/) when they are smaller or don’t have enough resources to purchase a more sophisticated platform. Using spreadsheets is easy to share with the team; however, the spreadsheets become obsolete the moment data gets manually added in. Without an [automation tool to automatically add this data](https://www.clari.com/products/product-overview/) in real time, the forecasting approach is never truly accurate.

### \* **using crm for sales forecasting :**

### Using your crm for sales forecasting seems like a pretty logical plan. I mean it’s serving as your system of record to track and manage every opportunity, so it should be easy to develop a forecast, right? Organizations typically use crm reports for sales forecasting when manual spreadsheets become too laborious to upkeep and teams need a tool to aggregate those numbers.

### \* **Using a Revenue Operations Platform for Sales Forecasting :**

With the emergence of AI and automation, new tools are available that make [sales forecasting](https://www.clari.com/blog/sales-forecasting-guide/) more connected, efficient and predictable. [Revenue operations platforms](https://www.clari.com/products/product-overview/) bring a new way to generate revenue with the same level of transparency and rigor that companies expect from any other mission-critical business process (ERP, supply chain, etc…).

## 4. **Techniques for Sales Forecasting**

### **\* Top-Down Sales Forecasting Method :**

The top down sales forecasting technique starts by identifying your total addressable market or TAM for each business segment. It takes a higher-level approach to viewing your business. According to the [Corporate Finance Institute](https://corporatefinanceinstitute.com/resources/knowledge/modeling/top-down-forecasting/), your TAM is developed by researching market valuations from reputable sources, such as Gartner. You then estimate how much market share you’ll be able to capture and the revenue you’ll be able to acquire. Calculating revenue is done by multiplying the TAM by the percentage of market share you think you’ll be able to achieve

### **\*Bottom-up Sales Forecasting Method :**

On the opposite spectrum is bottom-up sales forecasting, which starts with the reps instead of the managers and actual opportunities in play instead of models. In this sales forecast method, every rep calls their number based on the opportunities they have in pipeline. Managers work with reps to [inspect pipeline](https://www.clari.com/blog/5-data-driven-methods-for-inspecting-sales-pipeline/), understand projections and look through sales activity data in order to justify the forecast.

The benefits of a bottoms up sales forecast is that it’s based on real-world opportunities happening in real-time. Reps and managers who are in tune with each of their deals can provide a more accurate sales forecast based on the opportunities in play.

### **\*Qualitative Sales Forecasting Method :**

Qualitative [sales forecasting](https://www.clari.com/blog/sales-forecasting-guide/) involves the estimation of sales performance based on long-time expertise or well-versed industry knowledge. According to [AccountingTools](https://www.accountingtools.com/articles/what-is-qualitative-forecasting.html" \t "_blank), qualitative sales forecasting “...relies upon the knowledge of highly experienced employees and consultants to provide insights into future outcomes.”

### **\*Quantitative Sales Forecasting Method :**

Quantitative forecasting model uses historic sales data to [calculate accurate forecasts](https://www.clari.com/blog/sales-forecasting-accuracy/). It’s based on past performance and can be done in two ways ([Chegg](https://www.chegg.com/homework-help/definitions/quantitative-forecasting-methods-31)). The first method is a [time-series model](https://www.clari.com/blog/time-series-sales-forecasting-predictable-revenue/) which looks for patterns in the data to build the forecast and predict where you’ll land based on current [sales pipeline coverage](https://www.clari.com/blog/pipeline-coverage-best-practices/). The second method is called the associative model and it uses assumptions to build a linear regression-based forecast. The linear regression shows where you’ll end up based on those assumptions.

Organizations choose to use quantitative forecasting because it’s the first step in being data-driven. However, a major drawback from using this method is that it is highly dependent on an accurate data set, analysis of historic performance and complete visibility into pipeline. With these insights in hand, it’s possible for organizations to land within 2% of their forecast, just weeks into the quarter.

5.VALIDATING CORRECT MODEL

You might find that you try out several different sales forecast methods before settling on one. Or you may find that you start out with one, but soon outgrow it. The important thing to remember here is that you have to look at where your business is today and where you want it to go.

And, unless you’re using a revenue operations platform like [Clari](https://www.clari.com/), the reality is that all of these methods typically require manual data collection and consistent, cooperative effort from the sales reps, as well as very accurate data that's inputted manually — which rarely happens in the real world.

Let Clari take the guessing game out of sales forecasting and set you up for success using real data-driven techniques.

Start developing forecasting as a real-time function that guides your business, instead of scrambling to consolidate spreadsheets, trying to make sense of disjointed reports and wasting time with manual data input.

Clari adds a whole new layer to [sales management](https://www.clari.com/solutions/teams/sales/) beyond what your CRM can provide. By leveraging the automated activity capture, as well as the AI-powered insights that let customers improve team productivity, you’ll be able to drive and convert more pipeline, forecast the business and reduce churn.

Validation is very vital in choosing whether the method or program we are using is right interms of proper execution.

**6.ACCURACY AND PRECISION**

Since accuracy, precision, and recall are numerical measurements, you can conveniently use them to **track the model quality over time**. The only major limitation is the need for **true labels** in production.

To quickly calculate and visualize accuracy, precision, and recall for your machine learning models, you can use Evidently, an [open-source Python library](https://github.com/evidentlyai/evidently) that helps evaluate, test, and monitor ML models in production.

You will need to prepare your dataset that includes predicted values for each class and true labels and pass it to the tool. You will instantly get an interactive report that includes a [confusion matrix](http://www.evidentlyai.com/classification-metrics/confusion-matrix), accuracy, precision, recall metrics, [ROC curve](https://www.evidentlyai.com/classification-metrics/explain-roc-curve) and other visualizations. You can also integrate these model quality checks into your production pipelines.

To improve the accuracy of a prediction model, we can:

1. Add more data
2. Treat missing and outlier values
3. Perform feature engineering
4. Perform feature selection
5. Try multiple algorithms
6. Tune hyperparameters
7. Use ensemble methods
8. Use cross validation

Likewise the from obtaining data to measuring accuracy of the FUTURE SALES PREDICTION model is implemented in pyhton programming language with help of machine learning and

APPLIED DATA SCIENCE operations and techniques.