Machine Learning
with bisees

GAIN AN UNDERSTANDING OF THE BUSINESS PROBLEMS FACED BY YOUR CUSTOMERS, LEARN THE BENEFITS OF EXEPNOML, AND ENSURE THAT ML IS THE RIGHT SOLUTION FOR THE PROBLEM.

Google Cloud Partner

What is ML AND HOW DOES IT BENEFIT OUR CUSTOMERS?

Machine learning (ML) is a subfield of artificial intelligence (AI).

The goal of ML is to make computers learn from the data that you give them. Instead of describing the action the computer should take, the code provides an algorithm that adapts based on examples of intended behavior. The resulting program, consisting of the algorithm and associated learned parameters, is called a trained model.

The use of ML is becoming a necessity for businesses. A well-executed ML strategy can put your customer's business ahead of the competition, expediting the competitive edge needed to succeed in their market.

Customer expectations of products and services have evolved. They don't like waiting and want things fast. They expect you to know them and be available at all times, and they don't want to repeat themselves when they contact you. They require **personalized services** and **higher-quality products**.

ML allows you to deliver on these expectations and help businesses make data work for them by:

- ✔ **Solving problems** without explicitly codifying the solution

- ✔ **Building solutions** that improve themselves over time

- ✔ **Simplifying complexity**

- ✔ Using **predictive analytics** at scale

Use Case 1

Your company is investigating whether ML could be used to **reliably forecast demand**. One of your main challenges in gaining a competitive edge is the ability to predict demand and optimum inventory levels for products that are expected to become popular with customers with high cross-sell ratio, which accounts for a majority of their sales and revenue.

Since your company does not have the resources to build an internal analytics team to create effective ML/AI models, you are unable to make accurate forecasts for trending products. This resulted in a **missed opportunity** to stock right products, the season's best-selling and most highly demanded toy, leading to poor customer experience.
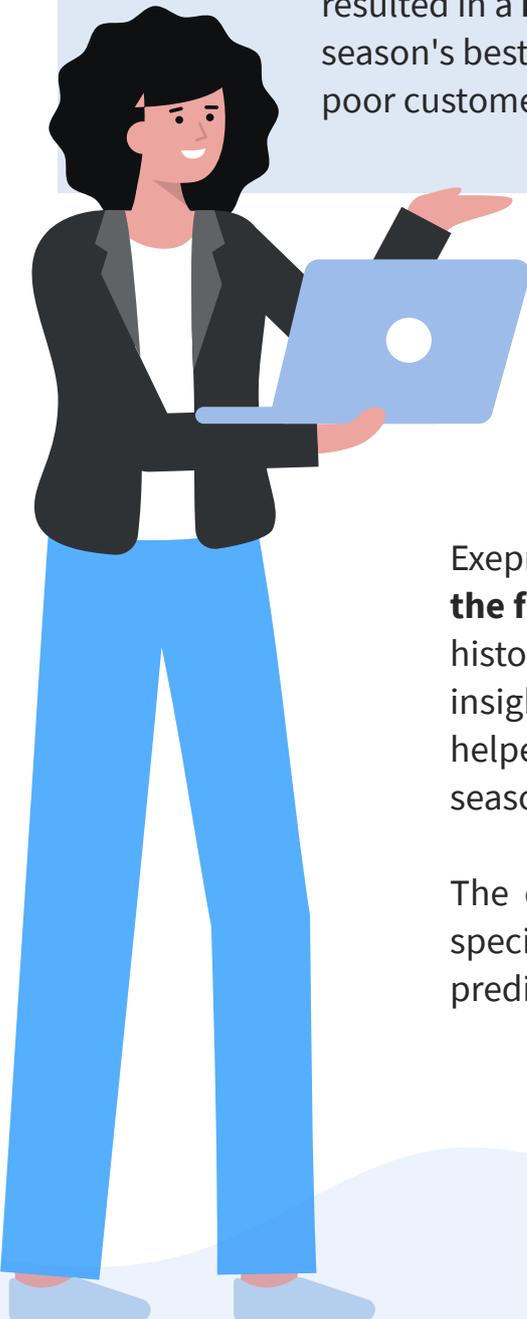
The solution

ExepnoML has ready-made **ML models that would predict the future sales** of specific products using a combination of historical sales data and seasonal trend information. The insights gained by these machine learning algorithms also helped to forecast predicted demand for inventory, making seasonal and trend-based supply decisions simpler.

The choice of the best-performing ML model is based on specific measurements, used to help provide more accurate predictions for your business.

bisees

Gaining intelligence and insight from data

Machine Learning is the new ground for **gaining a competitive edge** and **creating business value**.

According to a recent MIT study, companies that use ML and analytics are more likely to **make and execute data-driven decisions faster than competitors**.

ML helps **generate insight** and **predict business outcomes using data** (even without knowing what questions to ask), allowing for predictive and prescriptive customer solutions.

ML leads to **faster, more confident decisions** by removing doubt and gut instinct from business decisions.

Business benefits of ExepnoML

BISEES Exepno ML Engine is a managed service that enables developers and data scientists to build and bring superior machine learning models to production. ExepnoML offers training and prediction services, which can be used together or individually. ExepnoML is a proven service used by enterprises to solve problems surrounding the following:
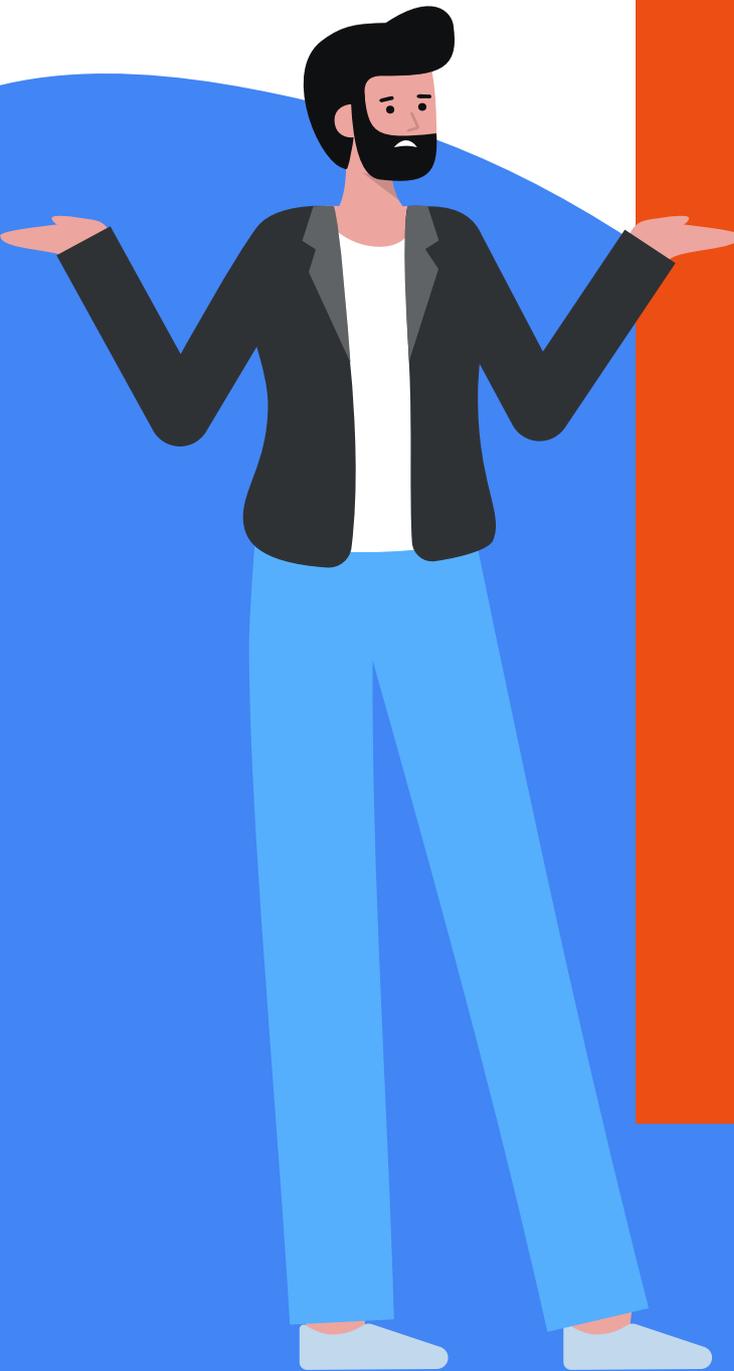
✓ **Scale**
Gain instant access to thousands of machines with ExepnoML.

✓ **Speed**
ExepnoML is run on GCP which is known for its best performance for AI workloads with custom ML hardware Cloud TPUs.

✓ **Quality**
ExepnoML can provide pre-trained AI building blocks to solve business needs with the highest quality.

✓ **Customization**
Easy-to-customize ML models with Bisees domain expertise can increase your team's productivity.

The importance of data to ML

Focus on the data critical to successful ML models, and learn about the building and training of an ML model based on a given algorithm.

Why is data so important to ML?

No data means no machine learning (ML). To create an ML solution, the pivotal necessity is to have the right set of data. It is only through the data that patterns are recognized and the ML models are trained.

Training a machine learning model requires a large number (thousands or more) of data examples. These examples must be relevant to the problem you are trying to solve. If you don't have the data you need, shift your focus to collecting the appropriate data and finding adequate data sources.

Managing data at scale is important for every company with a digital connection to customers, employees, or stakeholders. Big data is what allows businesses to **understand their markets, identify their best performers, and better serve and sell to their customers**. This is true for companies of all sizes: whether they have 100 or 100 million customers, businesses need to know who their best customers are and what they need. Data stored or received in real time needs to be analyzed effectively to be translated into insights. These valuable insights help make data-driven decisions faster and maximize your business advantage.
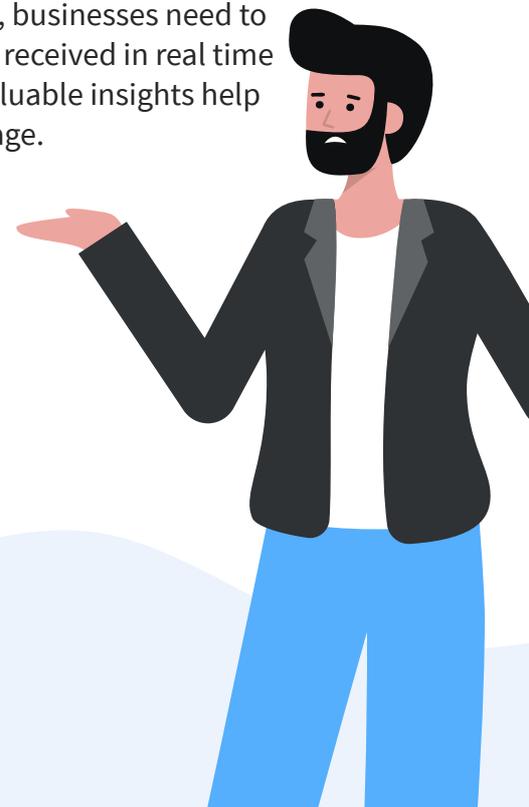
CHALLENGES OF WORKING WITH DATA

There are several challenges in this journey of data management and analytics:

Data often goes **unsynced, unused, and untapped** which prevents them being used for business advantage. Your business may have lots of historical sales data, but only when you combine it with seasonal trend information to build an ML model could you use it to predict the future sales and forecast demand.

Data silos and **managing data across legacy systems** are the top two challenges that enterprises face. Moreover, there is tremendous growth in big data, where 90% of enterprise data is unstructured. With big data growing in enterprises, other challenges include adding new skills to existing staff or recruiting new staff.
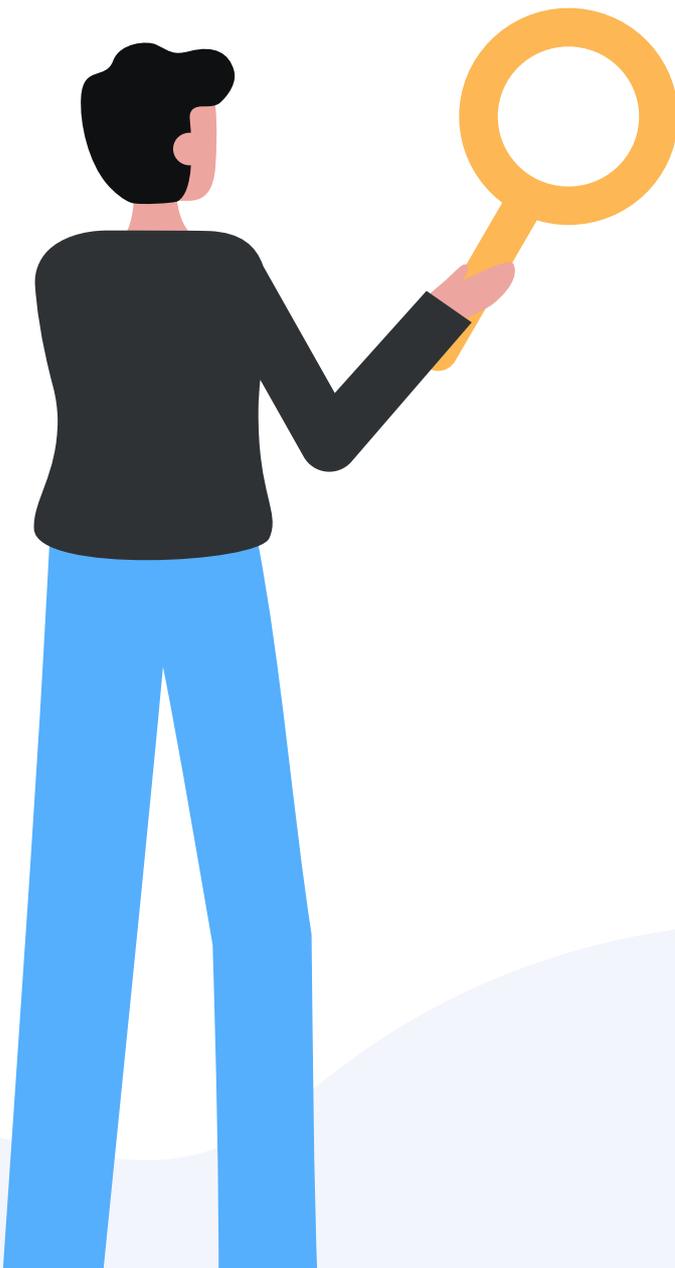
To stay relevant and ahead in this dynamic market, **businesses need to build a data strategy across all data silos**.

Checklist for 'Good' Data

The following checklist presents essential guidelines for assessing and preparing data for an ML project. This will help determine whether your data meets the minimum requirements for ML. The questions below are the basic requirements and 'rules of thumb'; that apply to most ML solutions involving structured (for example, tabular) and/or unstructured (for example, text) data.

Does the data include information that can predict the target?

For example, to detect a fraudulent credit card transaction, you'll need to have information on the transaction amount, the transaction location, the card holder's primary location, and so on. Data such as the name of the cardholder is not useful.

Does the granularity of training and prediction match?

For example, if your customer wants to predict the weather temperature hourly, but they have been recording the temperature once a day or once a week, their data doesn't have enough signal. Similarly, your customer cannot predict a very granular characteristic based on aggregates. You can help them build a model to convert blurry images to sharp images, but they first need to give it some sharp images as training data to learn from.

Is it already labeled data (i.e., training examples)?

Your customer will need labeled data.

If the target is categorical, some labeled data will be needed for each of the categories.

For example, in a spam email classification problem, some examples are required, both of correctly identified spam emails and of correctly identified non-spam emails.

If the target is numerical, some labeled data will be needed for different ranges of values.

For example, if it is necessary to predict the height of a person from her photo, some photos of tall persons, of average-height persons, and of short persons—all with the correct heights assigned to them—are required.

Is the data correct/accurate?

The data provided to the ML model must be correct to ensure that the model is trained accurately, therefore establishing good performance results.

Often, organizations have "dirty data" that isn't accurate. For example, timestamps might be incorrect. If data relies on user entry, this may not be reliable or accurate enough to use.

Is there enough data?

Enough data is needed to train a model.

Customers will need hundreds of thousands of examples to use machine learning models. However, the number of examples required varies by both the number of features and the algorithm selected.

Is the data easily accessible by the team and machines performing the ML?

If there are strict restrictions that prevent easy access, these should be addressed before the ML tasks start.

Can the data be read fast enough?

It can be frustrating to wait days to generate features and/or labels each time a change is made to the data preparation pipeline. For example, suppose your customer has a massive database of sensor readings that can be accessed using an API. If they want to generate a new feature that is the sum of all sensor readings, they'll want to be able to do so within a reasonable amount of time.

Is there documentation for each field of data?

Can a brief description for each field be provided, including its name and type, what it represents, how its value is measured, when the data is collected, whether the original values can be updated, and its applicability to the use case?

Are the missing values a small percentage of the fields of interest?

Note that missing values can be represented in different forms even for the same field: " ", None, NULL, NaN, 0, -1, 9999, and so on. The more missing values, the less useful the data. Is there enough useful data left if only data with no missing values is considered?

DATA ANALYTICS STAGES AND USE CASES

After you've made sense of your data, the next step in this journey toward ML is to know what questions to ask, and what insights to generate, by adopting advanced analytics.

Descriptive and diagnostic analytics are mainly used in business intelligence (BI) and reporting but have a low level of sophistication.

DESCRIPTIVE ANALYTICS

Describes the state of the business. Helps to answer questions such as:

- How is the business performing?
- Who are the best customers?
- What are the best-selling products?

DIAGNOSTIC ANALYTICS

Analysis that helps to diagnose issues and their root cause. Used to answer questions such as:

- What caused a decline in sales?
- Why did a region miss its target?

DATA ANALYTICS STAGES AND USE CASES

PREDICTIVE ANALYTICS

Forward-looking analysis to anticipate the future. Helps to answer questions such as:

- What is our sales forecast for the next quarter?
- Which customers are likely to default?
- Which prospects are most likely to buy our product?

PRESCRIPTIVE ANALYTICS

Gives a clear recommendation on the best course of action. Helps to answer questions such as:

- How should we invest our money?
- What is the best next move? (AlphaGo)
- What is the best route to my destination?

ML and artificial intelligence (AI) thrive on **predictive and prescriptive analytics** because they have a higher level of sophistication. Although ML is not the solution to every business problem, when a use case for ML is identified, predictive and prescriptive analytics provide vital information for building an ML solution.

The keys to successful ML
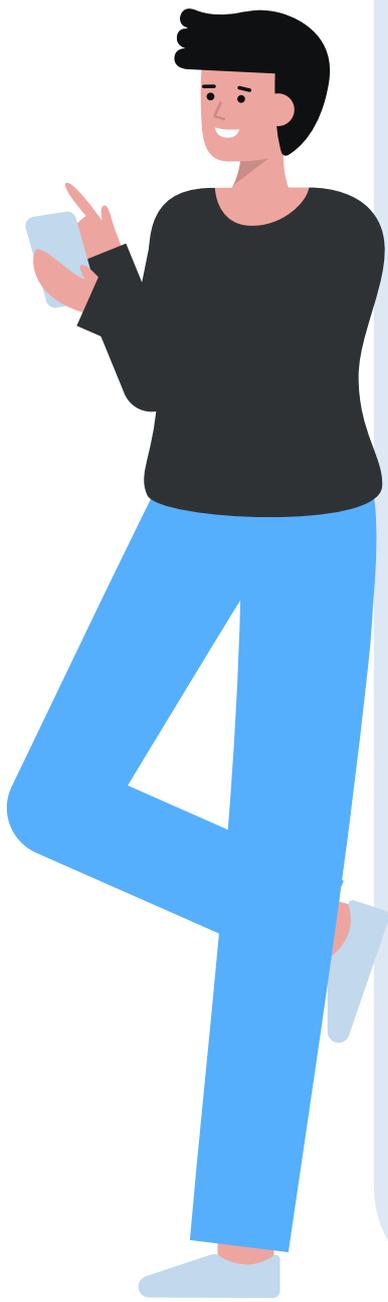
So what makes the difference between a well-executed, successful ML project & an average ML project?

The quality of an ML project typically depends on three main things.

Click on the tabs below to learn about each of the three main reasons.

- Data
- ML model
- Computation

Training data

Training data is the historical data labeled with the value you are trying to predict. The more training the model is able to do, the better the model performs. For example, if you want to detect fraud, you need many historical examples of transactions labeled as fraudulent or not fraudulent.

Machine learning is a way of **replacing complex heuristics with programs that learn from data**. ExepnoML can help your business to improve performance by combing human intelligence with machine intelligence.

ExepnoML will support and bring value to your company's journey from business need to ML solution deployment

Traditional business intelligence (BI) and reporting involves descriptive and diagnostic analytics, whereas machine learning (ML) and artificial intelligence (AI) includes predictive and prescriptive analytics—offering a more sophisticated solution. Using ExepnoML and its AI algorithms, along with your data and domain knowledge, will result in a successful and well-executed ML project.

| Vertical	Stream	Use-case
Retail	Sales	Demand Forecasting
		New Store Location
		Customer Retention
		Product Recommendation
		Pricing Optimization
		Planogram Design
	Inventory	Inventory Optimization
		Product Cannibalisation
		Procurement Bid Value Decision
	Backend & Service	Sales Report and Analysis
		Chatbots
		Voice & Visual Search

Problem Definition

Demand Forecasting is the practice of looking at the historical sales data and predicting the demand that will be generated for a future period based on current market situations. Critical business operations like cash flow and procurement are dependent on the demand forecast being correct. Demand forecasting can be of a multitude of types, including, but not limited to, short term, long term, macro demand, micro demand, internal and external forecasting. In a retail scenario, demand forecasting for sales over the next quarter is a long-established business process.

Solution Methodology

With ExepnoML, this whole module becomes a plug and play tool on a dashboard. Of the multiple quantitative and qualitative methodologies designed to address this specific problem, machine learning algorithms have stood out as the most effective.

This process involves an initial load of creating data sets, defining features, creating a use-case specific model, training and tuning that model and then deploying it for store managers and procurement managers to consume. This is where ExepnoML comes in and automates the whole process except a few key business decisions which await manual input. The first point-of-contact is ExepnoETL which reads your data from various sources like POS, company database, third party data and collates them together on a GCP data warehousing engine. Having collected the data and and creating a master database, ExepnoML starts to apply multiple algorithms from its library on the data provided to see which sort of algorithm suits it more. Once that decision is made, the algorithm is trained with the custom data provided by you to make it specific to your use case. The model is tuned and tested by the ExepnoML backend and then deployed for the user to test out it's validity.
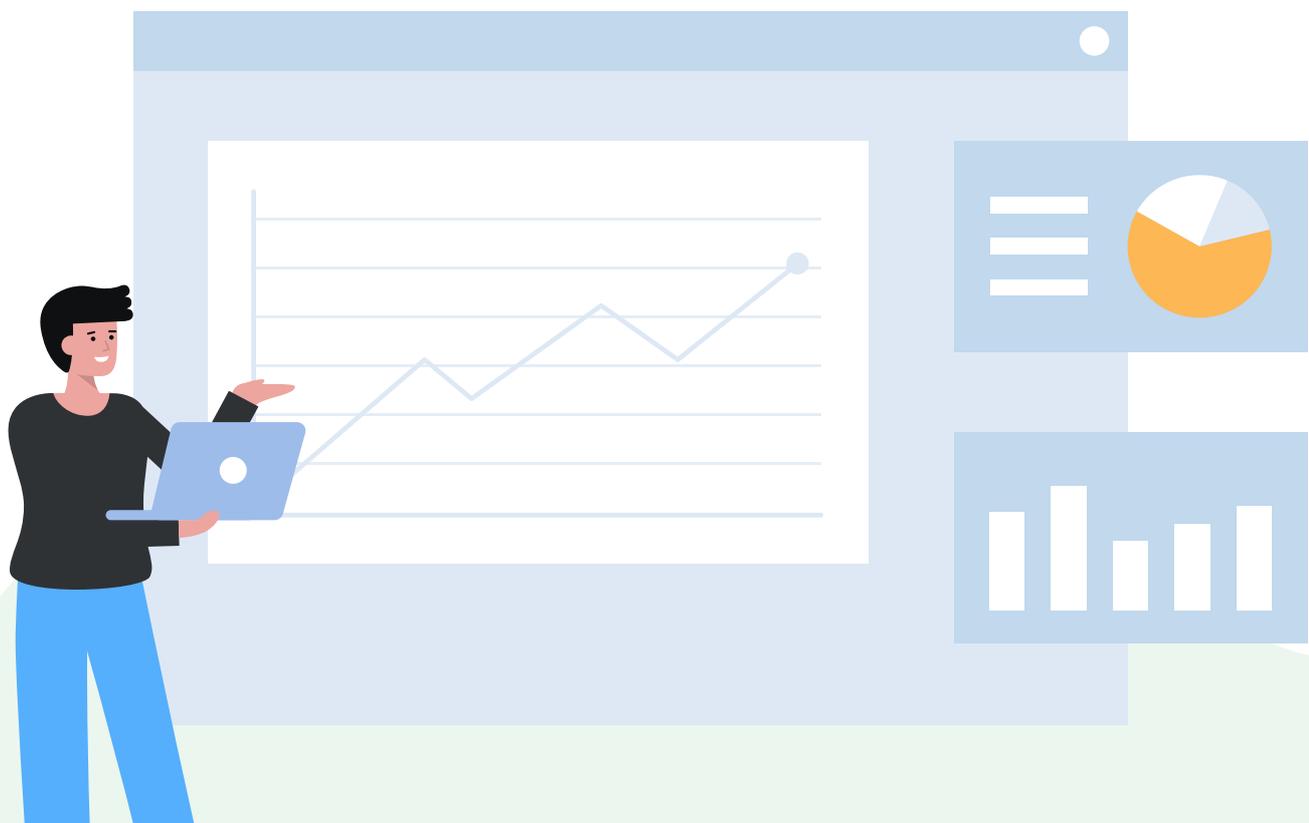
Metric comparison

Once a model is built, the data will need to be determined as 'good enough' before deployment, alongside the consideration of various variables. For example, one might consider what is the cost of training the model further with newly procured data and how long that might take. They might also consider that the model is as general as it could be given the scenario, and that any further training would increase the chances of over-fitting. With these decision being made, one needs a specific metric for comparison. Fortunately, machine learning metrics are abundant and very specific. For some models, F-Score is the best metric given a case, while others might have accuracy as the best indicator of the value of the machine learning algorithm. ExepnoML makes the process of viewing these values and making a smart decision very simple. It uses a dashboard for ML algorithm using MLFlow in the backend so that any user can view the details of the algorithm's metrics. One could even go on to provide business rules here which would be taken into account before deployment. For example, one might decide that a model is not useful unless it has an accuracy of above 0.97, while someone else may consider this accuracy as an indicator of the overfitting problem. ExepnoML simplifies this process by enabling users to make their own decisions to best fit their business.

Deployment and Usage

Once a model is built and decided on, for a use-case, the deployment is a highly-available, low latency webapp with standard APIs constructed on the GCP cloud. A user could go every three months, in this case, and upload a new dataset with the sales of the last 3 months and other relevant details such a store location and regional holidays to get their own prediction. ExepnoML makes this process very simple enabling users to gather actionable insight from their own data. Users could also choose to retrain their model at a later time to see if the new model out-performs the existing one and in that case they could choose to deploy the new model with minimal downtime.
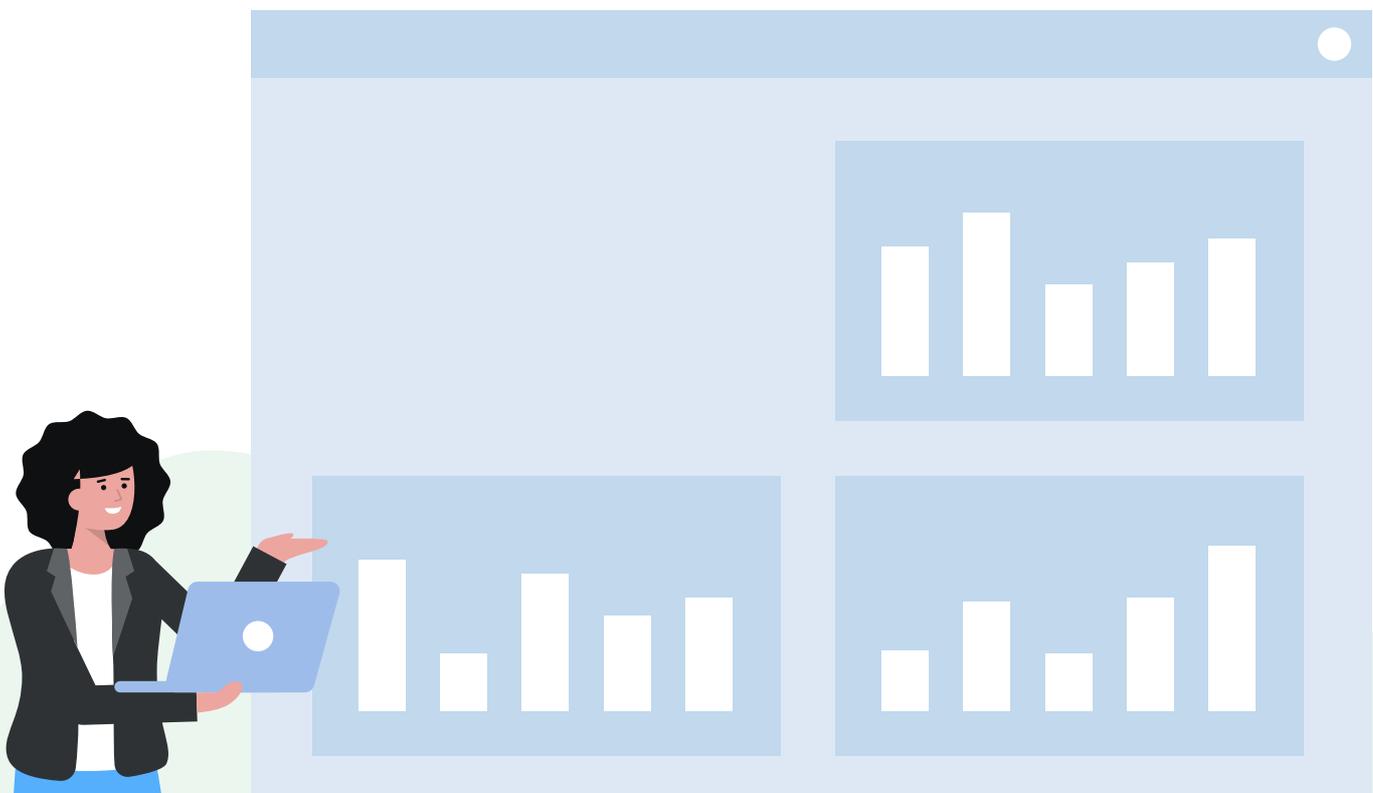
bisees

Contact Us

Email: info@bisees.com
Phone: 00353 1 410 0600

HQ:
Guinness Enterprise Centre
Taylor's Ln,
Dublin 8,
Ireland

R&D Office:
Bradfield Centre
184 Cambridge Science Park
Milton Road, Milton
Cambridge, CB4 0GA

Google Cloud
Partner