# Machine Learning

SYSPRO 8

Reference Guide

SYSPRO™
Simplifying your Success

# CONTENTS

## AI Machine Learning

# AI Machine Learning

Machine learning forms part of the *SYSPRO Artificial Intelligence* module. It uses specific algorithms and statistics to examine historical data. The program then uses the data patterns to reveal trends and predict future outcomes, benefiting management by presenting the big business picture.

Although these predictions require minimal human intervention, they rely heavily on the data quality and the attributes of the SQL statement. Reliable predictions will support business processes and improve decision making by learning from past experiences.

# Exploring

You access the SYSPRO Artificial Intelligence module from the SYSPRO Avanti website (client). Accessing and processing of information is then done via the **IIS** and **Avanti Web Server**. The **Web Server** communicates with the **SYSPRO 8 Machine Learning** service (AI Layer) and the **SYSPRO 8 e.net Communications Load Balancer** service (SYSPRO App Server).

The **SYSPRO 8 Machine Learning** service can have 2 instances installed, with different endpoints for Training and Prediction. It sends all communications through to the **SYSPRO 8 e.net Communications Load Balancer** service, which then calls business objects via e.Net to read and write data to the system-wide database. Data is then added to (or retrieved) from the database and the communication is sent back to *SYSPRO Avanti* and displayed on the user interface (graphically depicted by the bi-directional arrows).

# AI Architecture Diagram
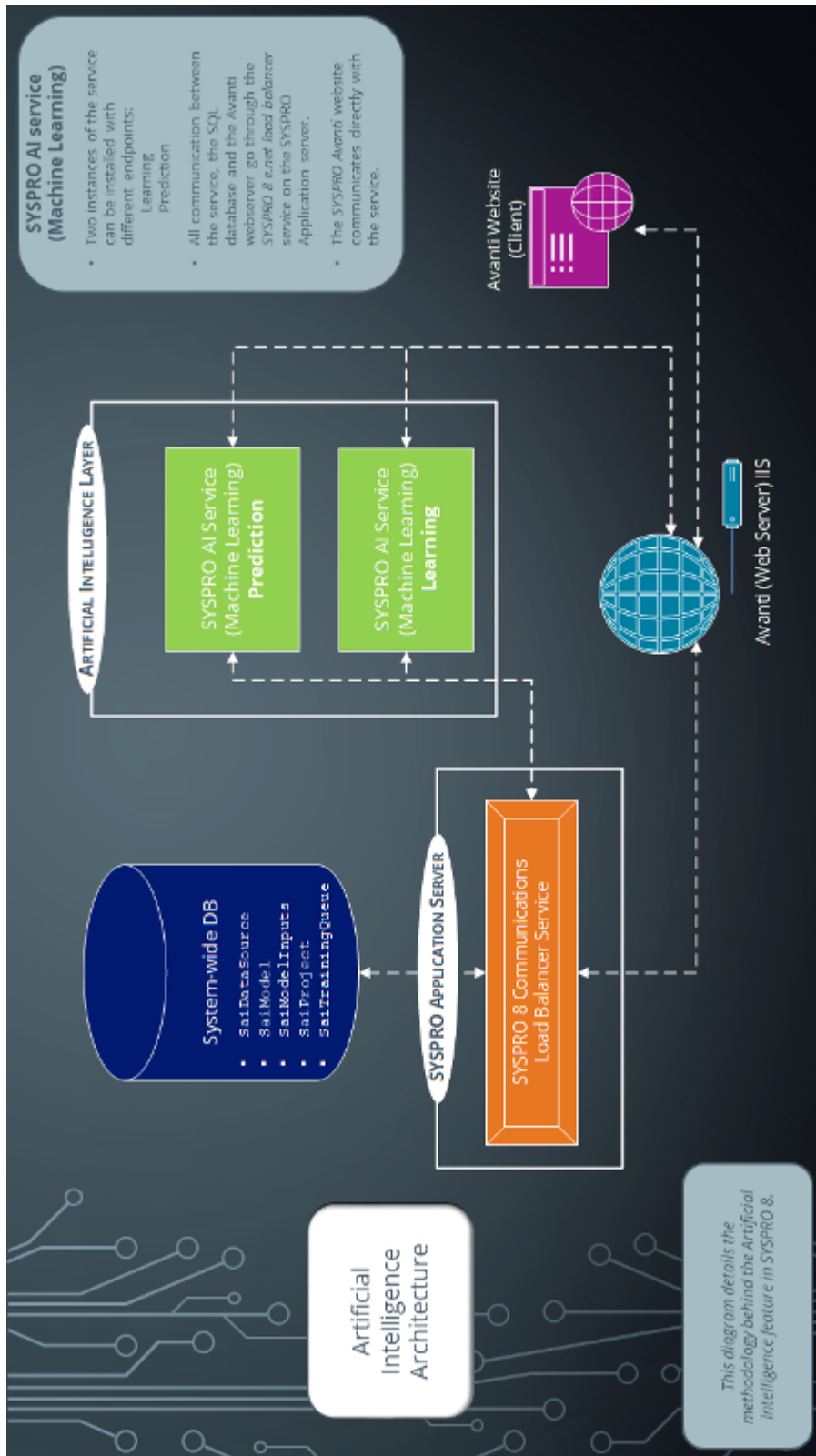
# Navigation

- A machine learning project can only be created in **SYSPRO Avanti**.

  *SYSPRO Avanti > SYSPRO Artificial Intelligence > AI Administrator*

# Terminology used

## Project

A machine learning project consists of a data-source, a SQL statement, and machine learning training options.

The project tells the SYSPRO Artificial Intelligence module what data must be used for training, and provides various options to control the training process. The SYSPRO Artificial Intelligence module uses the SQL statement to query the data-source. It then learns the patterns in the data to produce a model.

## Data source

A data-source is a reusable link to an SQL Server instance hosting the database and consists of the address of the server, login credentials, and other options specific to the server in question.

The actual dataset that used in the machine learning project, is specified in the SQL query within a project.

The data-source can be setup once, and then reused by any user to quickly train a model based on data on that specific SQL Server.

## Training

Training a model involves examining patterns in the data using various machine learning algorithms.

The data can be thought of as containing 1 or more input columns ($x$), and 1 output column ($y$). Training is the process of learning how $x$ maps to $y$.

Symbolically, if we say `f(x)` = $y$, then f is some function that can map $x$ to $y$. In mathematics, we are usually given `f` and $x$, and asked to calculate $y$. In supervised machine learning, we give the computer $x$ and $y$, and ask it to learn `f`.

## Model

A machine learning model is a reusable block of code that can make predictions on new, or unseen data. The model is a binary object that cannot be directly inspected.

The **SYSPRO Artificial Intelligence** module takes care of loading models. To make a prediction, you have to provide the same number & type of columns on which the model was trained. The model will then produce a prediction.

*For example:* You may have trained the project to predict a value $y_1$, based on the input values of $a_1$, $b_1$, $c_1$. If you provide new values of $a_2$, $b_2$, $c_2$, the model will output a new value of $y_2$, based on those new input values.

## Prediction

A machine learning prediction requires a trained model.

When performing a prediction, the **SYSPRO Artificial Intelligence** module uses a new set of input data to forecast or foresee a possible scenario or outcome. This can be considered a more advanced form of a what-if analysis.

Rather than relying on simple linear relationships, like a pro-rata calculation, a machine learning prediction can predict non-linear, or multi-dimensional relationships.

## Prediction end point

This is the rest endpoint address of the **SYSPRO Artificial Intelligence** service (e.g. `http://localhost:30238/SYSPROMLE/rest`) that is used by the **SYSPRO Machine Learning** module to perform predictions.

## Training end point

This is the training endpoint address to the **SYSPRO Artificial Intelligence** service (e.g. `http://localhost:30238/SYSPROMLE/rest`) that is used to train projects to generate a model that can be used to perform predictions.

# Starting

## Restrictions and Limits

- A machine learning project can only be created in *SYSPRO Avanti*.

## Setup Options

- Once you have installed machine learning, you need to configure the **Prediction end point** and **Training end point** within the **System Setup** program.

# Solving

## FAQs

### Can I create more data sources?

Data sources are automatically created for every SYSPRO company, but you can create your own data sources for data that resides outside of the SYSPRO database.

### Can I train more than one project at a time?

No, due to the intensity of the training process, only one project can be trained at a time.

You can, however, queue projects for training, which will automatically start the next project as soon as the previous project is completed.

### Can I edit and delete the SYSPRO models?

No, the shipped models cannot be edited or deleted.

You can, however, train a shipped model, which then creates a copy of the model that can be used for predictions.

### What do the records in the Model information section represent?

The **Raw count testing** figure is the total number of records used for the selected prediction results.

The **Raw count total** figure is the total number of records returned from the SQL statement.

# Using

## Tasks

### How to configure machine learning

The configuration of **SYSPRO Machine Learning** has to be done within the core SYSPRO product.

1. Open the **System Setup** program and navigate to the **Artificial Intelligence** tab.

2. At the **Machine learning** section, enter the **Prediction end point** and **Training end point** REST addresses (e.g. `http://localhost:30238/SYSPROMLE/REST` where `localhost` indicates the server name and `30238` indicates the port number).

   Separate endpoints for training and predicting ensures better responsiveness, as especially the training endpoint may seem unresponsive when training projects.

3. Save your changes.

### How to create a machine learning project

A machine learning project can only be created in *SYSPRO Avanti*.

1. Open *SYSPRO Avanti*.

2. From the menu, select **SYSPRO Artificial Intelligence** and **AI Administrator**.

   This loads the **SYSPRO Machine Learning** program.

3. Create a new learning project.

4. Create a new data source (if you don't have one).

5. Create a new project.

6. Add the machine learning project to the training queue.

   Once the training is done, the training queue displays a **Training completed** status and a model is created that can be used for predictions.

7. Specify input values for the model to run a prediction.

### How to create or edit a data source

1. Open *SYSPRO Avanti*.

2. From the menu, select **SYSPRO Artificial Intelligence** and **Setup**, then select **Data Source List**.

---

3. Select **New Data Source** to create your own data source or select an existing one to edit it.

4. Enter the data source details and test the connection.

   SYSPRO only supports the ODBC data connection.

---

# Referencing

## New data source

| Field | Description |
| --- | --- |
| Name | Indicate a name for the data source. |
| Description | Indicate a description for the data source. |
| Type | This defaults to ODBC. |
| Connection string | Enter the connection string. This must be in the following format:<br><br>`DRIVER={SQL Server};SERVER={SERVERNAME};DATABASE={Database Name};UID={user};PWD={Password}`<br><br>Select **Test connection** to confirm that the connection is working. |

# New Project

| Field | Description | | |
|---|---|---|---|
| Project | Select the project model. | | |
| | | Classification | Select this if you want the prediction result to be a condition (e.g. good or bad). The output column is usually a text field. Accuracy (the closeness of a measured value to a standard or known value) is used to determine the condition. |
| | | Regression | Select this if you want the prediction result to be a quantity or an amount (e.g. days late). The output column is usually a numeric field. Mean squared error (average squared difference between the estimated values and what is estimated) is used to determine the regression. |
| Name | Indicate the project name. | | |
| Description | Enter a description for the project. | | |
| Data source | Select the data source from the list. Data sources are automatically created for every SYSPRO company in your database. You can create your own data sources for data that lies elsewhere. | | |
| Select statement | Enter the select statement (this can be a SQL query) that will extract the required data from the database. Select **Execute query** to run the query. | | |

| Field | Description |
|---|---|
| Output column | Once the query is executed, the list of column names is populated. |
| | Select the column that contains the primary data you want to query and use for prediction. |
| | For classification models, this is usually a text field. |
| | For regression models, this is always a numeric field. |

# Project details

| Field | Description | | |
|---|---|---|---|
| Model | This displays the project model. | | |
| | | Classification | Select this if you want the prediction result to be a condition (e.g. good or bad). Accuracy (the closeness of a measured value to a standard or known value) is used to determine the condition. |
| | | Regression | Select this if you want the prediction result to be a quantity or an amount (e.g. days late). Mean squared error (average squared difference between the estimated values and what is estimated) is used to determine the regression. |
| Name | This displays the project name. | | |
| Description | This displays the project description. | | |
| Copy to new version | This lets you copy an existing project to create a new version. Indicate a name for the new project and select **Copy**. | | |
| Data source | This displays the data source. | | |
| Select statement | This displays the select statement or SQL query. Select **Execute query** to run the query. | | |
| Output column | This displays the output column name. For classification models, this is usually a text field. For regression models, this is always a numeric field. | | |

# Advanced Options

| Field | Description |
|---|---|
| **Training options** | This applies to the entire machine learning project. |
| Random state | |
| Scoring method | |
| Test ratio | |
| **Toolkit setup** | |
| Generations | |
| Population size | |
| Offspring | |
| Mutation rate | |
| Cross over rate | |
| Number of folds to evaluate | |
| Sub sample | |
| Max time | The maximum time is five minutes by default. |
| Max evaluation time | |
| Configuration dictionary | |
| Early stop | |

# Predict

The fields on this pane change depending on the fields used in the SQL statement.

| Field | Description |
|-------|-------------|
| Run | Executes the script to calculate predictions. |

**SYSPRO™**
Simplifying your Success