

# Machine Learning

SYSPRO 8

Reference Guide

Published: May 2021



# CONTENTS

## AI Machine Learning

Exploring .....	1
Starting .....	6
Solving .....	8
Using .....	16
Referencing .....	20

# 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

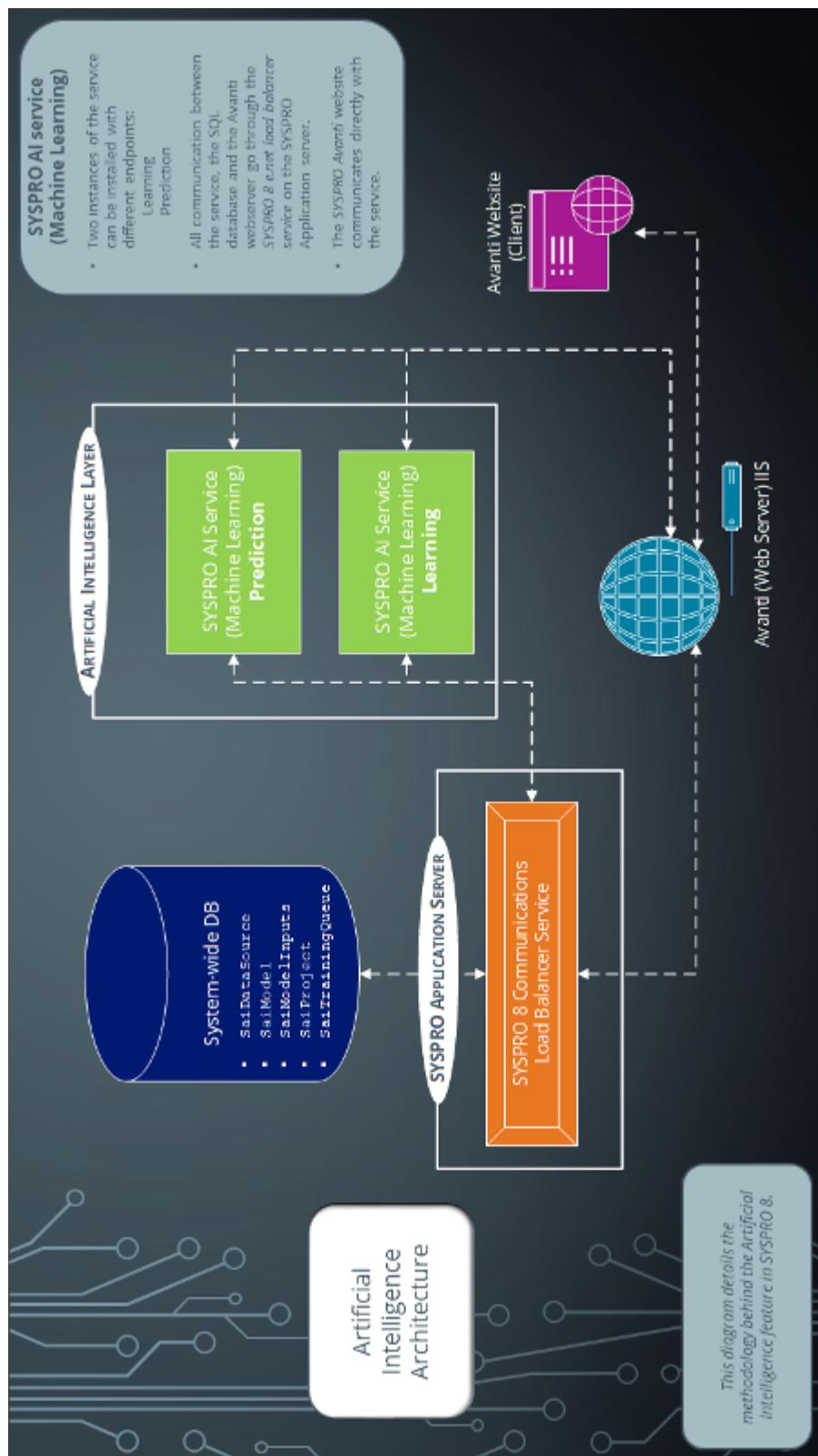
Machine Learning, the engine behind artificial intelligence, uses specific algorithms and statistics to examine historical data. This gives computers the ability to learn from large amounts of business and industry data sets in order to provide analysis and insights, which is increasingly required as part of the decision making process.

Business insights can be used in conjunction with machine learning projects by creating tiles with defined KPI thresholds. The KPIs are defined using the **Insight Tile KPI Definition** program in SYSPRO, whereas the actual tiles are created in **SYSPRO Avanti**. The tiles display a visual representation of the prediction results and indicate the proximity of the threshold, which can facilitate informed business decisions.

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



The following sample projects are shipped with the product to help kick-start the machine learning process. They are located in the `Harmony_SaiProject` file in the `\Base\Harmony\Standard` folder:

## Regression

- LCT days late

This predicts how late a shipment may be, measured in days.

- PO days late

This predicts the number of days a purchase order line may be late.

- Chance order will be late

This displays a percentage that predicts the chance a purchase order line may be late.

- Customer invoice pay days

This predicts the number of days a customer may take to pay an invoice.

## Classification

- Job status

This predicts whether a job may be completed early or late.

- Lost sales reason

This prediction determines whether sales may be lost and what the reason for losses would be.

- Problems with sales order

This predicts the most likely reason why a sales order is returned.

- Quote success

This predicts how many quotes should be accepted by the customer and result in sales orders.

- Stock code profitability

This predicts how profitable a stock code should be by location.

- Supplier performance

This predicts how well a supplier should perform in terms of deliveries, and indicates the likelihood of the delivery being on time and in full.

- AP invoice payment prediction

This predicts by when a supplier should pay an invoice.

- Customer classification

This predicts how profitable a customer would be.

## Anomaly Detection

- Purchase order line anomaly

This detects anomalies on purchase order lines based on the selected columns.

- Sales order line anomaly

This detects anomalies on sales order lines based on the selected columns.

## Navigation

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

*SYSPRO Avanti > SYSPRO Artificial Intelligence*

## Terminology

### Project (ML & AI)

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 (ML & AI)

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 (ML & AI)

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 and an active project.

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

## Prerequisites Technology

The following technology prerequisites are applicable to using this feature:

- Microsoft .NET Framework 4.6
- SYSPRO 8 e.net Communications Load Balancer
  - A valid endpoint must be configured in the **Setup Options** program of **SYSPRO 8**.
- SYSPRO 8 Machine Learning



This service can be installed on any server as long as the **SYSPRO 8 e.net Communications Load Balancer** endpoint is configured correctly in the service's configuration file.

- SYSPRO Avanti
- SYSPRO 8 Avanti Initialization Service



You can use the **SYSPRO Installer Application** to install these requirements.

## Configuring

Once you have installed the **SYSPRO 8 Machine Learning** service, the following setup options must be configured to use this feature:

### System Setup

*SYSPRO Ribbon bar > Setup > General Setup*

### Artificial Intelligence

- Prediction end point
- Training end point

## Installation considerations

- If you have two installations of the **SYSPRO 8 Machine Learning** service installed on different servers, then you can configure SYSPRO to use a **TRAINING END POINT** and a **PREDICTION END POINT**.

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

## Security

You can secure this feature by implementing a range of controls against the affected programs. Although not all these controls are applicable to each feature, they include the following:

- You restrict operator access to *activities* within a program using the **Operator Maintenance** program.
- You can restrict operator access to the *fields* within a program (configured using the **Operator Maintenance** program).
- You can restrict operator access to *functions* within a program using passwords (configured using the **Password Definition** program). When defined, the password must be entered before you can access the function.
- You can restrict access to the eSignature *transactions* within a program at operator, group, role or company level (configured using the **eSignature Setup** program). Electronic Signatures provide security access, transaction logging and event triggering that gives you greater control over your system changes.
- You can restrict operator access to *programs* by assigning them to groups and applying access control against the group (configured using the **Operator Groups** program).
- You can restrict operator access to *programs* by assigning them to roles and applying access control against the role (configured using the **Role Management** program).

## Restrictions and Limits

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

# Solving

## System messages

### Error messages

#### Error generating output columns

**Error Generating Output Columns** X

**!** The data source did not return a valid data set.  
The error message is "NoneType' object has no  
attribute 'cursor\_description'"

**OK**

#### Cause

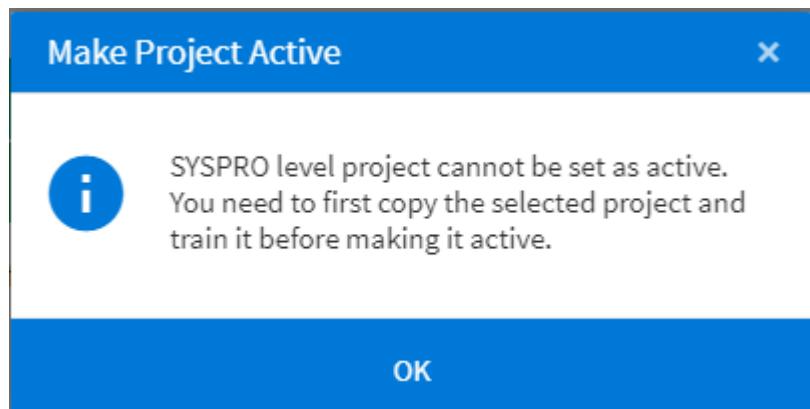
This message is displayed when you save an AI project in the **SYSPRO Machine Learning** program if the SQL select statement returns only one or zero rows of data.

#### Solution

Ensure that the SQL statement used to create the project returns more than one row of data, as projects require at least 2 rows of data to be trained.

## Informational messages

### Make Project Active



#### Synopsis

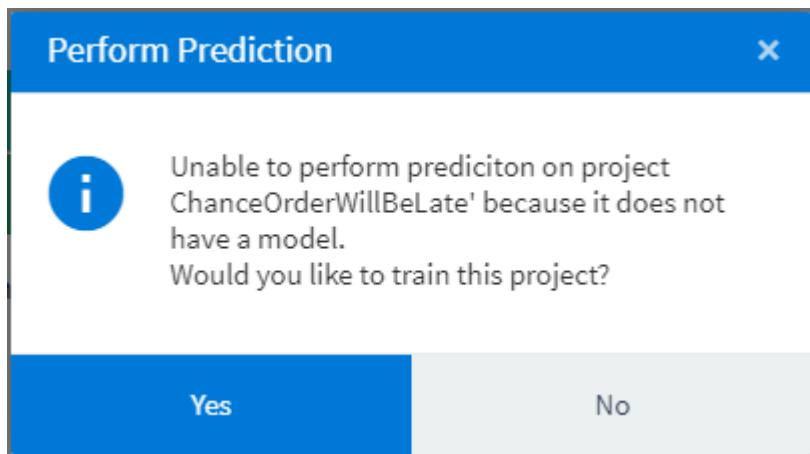
This message is displayed when you attempt to activate a project that hasn't been trained yet.

#### Solution

Before activating the machine learning project, train it so that it gets to know the data.

1. Open **SYSPRO Avanti**.
2. From the menu, select **SYSPRO Artificial Intelligence** and select **AI Administrator**.  
The SYSPRO Artificial Intelligence application is displayed.
3. Select the project you want to train from the **AI Projects** list.
4. Select **Train Project** to add the project to the training queue.  
The **Training Queue window** is displayed.
5. Once the training is done, the status **Training completed** is displayed and a model is created that can be used for predictions.

## Perform Prediction



### Synopsis

This message is displayed when you attempt to perform a prediction on a project that hasn't been trained yet.

### Solution

Before using the machine learning project for predictions, it must be trained so that it gets to know the data.

1. Open **SYSPRO Avanti**.
2. From the menu, select **SYSPRO Artificial Intelligence** and select **AI Administrator**.  
The SYSPRO Artificial Intelligence application is displayed.
3. Select the project you want to train from the **AI Projects** list.
4. Select **Train Project** to add the project to the training queue.  
The **Training Queue window** is displayed.
5. Once the training is done, the status **Training completed** is displayed and a model is created that can be used for predictions.

## FAQs

### Working with projects and data sources

#### Is it necessary to create more data sources?

You don't need to create your own data sources, as they are automatically created for every SYSPRO company.

You can create additional data sources for data that resides outside of the SYSPRO database using the **Data source List** window (*SYSPRO Avanti > SYSPRO Artificial Intelligence > Setup > Data Source List*).

#### 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 projects?

No, the shipped projects can't be edited or deleted.

You can, however, create a copy of a shipped project. The copied project can then be trained and the model used for predictions.

#### What is the difference between saving the project on system, company or role level?

The level on which the project is saved, determines which models are available for users and how the predictions are done.

Models that are saved on role level, are only available for operators who are grouped into that role.

Models that are saved on company level, are available for all operators in that company. Models that are saved on system level are available for all users across companies.

When predicting an outcome on operator level, the system uses the model that is saved for the current operator role. If there is no model for that operator role, the system first determines if a model exists for the current company and then whether a model exists on system-wide level.

## Why is my machine learning project not displayed in the Insight Tile KPI Definition program?

This may be due to one of the following reasons:

- The ML project must be active.
- The ML project must be trained.
- The ML project must be a **regression** type project.

## Why can't I delete the projects in the SYSPRO folder?

The projects listed in the SYSPRO folder are sample projects that are shipped with the product.

You can only copy sample projects to create a new version, but you can't delete or edit them. Only the copied versions can be edited and deleted.



Sample projects are located in the `Harmony_SaiProject` file in the `\Base\Harmony\Standard` folder.

## Why is there no data in the Feature importance graph?

The model may need to be retrained in order to display the feature importance percentage.

## Working with models

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

The **Model information** pane is used to display basic statistics about the trained project. To compare models and identify whether one model is better than another, some objective metric is required.

The ML Engine supports various metrics that are defined in the project using the **Scoring method**. The data set is split into a training set and a testing set. The ML Engine learns from the training set and is not allowed to see the testing set. Once a model is created, its performance on unseen data can be measured by passing the testing set into the model and comparing predicted values in the training set. This process of comparing predicted values to actual values can be used to objectively score the model.

The scores presented to the user are the result of this comparison.

- The **Scoring method** function is used to evaluate the quality of the given pipeline for the classification problem.
- The **Total set score** indicates how the model scored over the entire data set i.e. both the training and testing data sets combined.
- The **Possible error** percentage indicates how the model performed on the testing data set and is, therefore, a better indication of the prediction performance on unseen data.
- The **Test Observations** figure is the total number of records used for the selected prediction results.
- The **Training Observations** figure is the total number of records returned from the SQL statement.

## Project scoring metrics

### Which scoring metrics can be used for regression type projects?

The following scoring metrics can be used in regression projects:

- neg\_mean\_squared\_error
- neg\_mean\_absolute\_error
- neg\_mean\_absolute\_error
- r2

### Which scoring metrics can be used for classification type projects?

The following scoring metrics can be used in classification projects:

- accuracy
- adjusted\_rand\_score
- average\_precision
- balanced\_accuracy
- f1
- f1\_macro
- f1\_micro
- f1\_samples
- f1\_weighted
- neg\_log\_loss
- precision
- precision\_macro
- precision\_micro
- precision\_samples
- precision\_weighted
- recall
- recall\_macro
- recall\_micro
- recall\_samples
- recall\_weighted
- roc\_auc

## Which scoring metrics can be used for anomaly type projects?

There are no scoring methods for anomaly type projects. Anomaly detectors can't be scored as they are unsupervised machine learning algorithms and don't use analogous scoring metrics.

# Using

## Process

### 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.

## How to link a prediction to tile with KPIs

1. Access the **AI Administrator** program from within **SYSPRO Avanti**.  
This loads the **SYSPRO Machine Learning** program, where you can train and activate projects.
2. Access the **Insight Tile KPI Definition** program from within SYSPRO 8 to set operator permissions and define KPIs for the tile.
3. Access the program or location in **SYSPRO Avanti** where you want to place the tile.
4. Select **Customize** from the toolbar.
5. Select **Design Web View**.  
The Visual Designer is displayed.
6. Select a layout section and then select to add a tile.  
The **Add tile** window is displayed.
7. Select **Artificial Intelligence** and the tile you want to add.
8. Configure the tile by adding titles, additional parameters and styling.
9. Select **Add tile**.

You will return to the program where the tile is displayed.



You can only link an active, trained machine learning project to a tile with KPIs.

If you haven't activated the project in the **SYSPRO Machine Learning** program, you won't be able to add KPIs against the project.

# Available scoring methods per project type

This lists the scoring methods that can be used per project type.



There are no scoring methods for anomaly type projects. Anomaly detectors can't be scored as they are unsupervised machine learning algorithms and don't use analogous scoring metrics.

## Classification projects

The following scoring metrics can be used in classification projects:

- accuracy
- adjusted\_rand\_score
- average\_precision
- balanced\_accuracy
- f1
- f1\_macro
- f1\_micro
- f1\_samples
- f1\_weighted
- neg\_log\_loss
- precision
- precision\_macro
- precision\_micro
- precision\_samples
- precision\_weighted
- recall
- recall\_macro
- recall\_micro
- recall\_samples
- recall\_weighted
- roc\_auc

## Regression projects

The following scoring metrics can be used in regression projects:

- neg\_mean\_squared\_error
- neg\_mean\_absolute\_error
- neg\_mean\_absolute\_error
- r2

## Status Codes

### AI and Machine Learning Statuses

The status of a machine learning project indicates the current stage of the project and how it can be used.

Status	Description
0 – IN DEVELOPMENT	The machine learning project is new or currently in development and has not been trained yet.
5 – READY FOR USE	<p>The machine learning project was trained successfully and a model was produced.</p> <p>The project hasn't been activated, so the project model can't be used for predictions yet.</p>
1 – TRAINING	The machine learning project is currently being trained by the AI engine.
10 – ACTIVE	The machine learning project is active and the model for this project can be used for predictions.
99 – ERROR	<p>The AI engine ran into an error while training the project.</p> <p>The detailed error can be viewed from the training queue.</p>

## Hints and Tips

- The KPI definitions for machine learning projects must be created in SYSPRO 8 using the [Insight Tile KPI Definition](#) program.
- Only **Regression** type machine learning projects can be linked to a tile with KPIs, as values are compared to thresholds to determine the outcome.

# Referencing

## Menu and Toolbar

Field	Description
<b>AI projects in training queue</b>	This opens the <b>Training Queue</b> pane and lists all projects that are currently queued for training.
<b>Models available for predictions</b>	This opens the <b>List of Models</b> pane where all trained models are listed that can be used for predictions.
<b>Import project</b>	This lets you import a project that was previously downloaded. The <b>Import Project</b> window is displayed.
<b>New machine learning project</b>	This lets you add a new machine learning project.
<b>Save project</b>	This lets you save the new, or changes made to the existing machine learning project.
<b>Delete project</b>	This lets you delete the existing machine learning project.

## Project information

Field	Description
Model	<p>This indicates the project model.</p> <p><i>Classification</i></p> <p>This model predicts a condition (e.g. good or bad) and the <b>Output column</b> is usually a text field.</p> <p><i>Regression</i></p> <p>This model predicts the quantity or an amount (e.g. days late).</p> <p><i>Anomaly</i></p> <p>This model predicts a data irregularity or inconsistency.</p>
Name	This displays the project name.
Description	This displays the project description.
Data source	This displays the data source.

Field	Description
Select statement	This displays the select statement (usually a SQL query) that extracts the required data from the database.
Output column	<p>This is populated using the <b>Generate output columns</b> button on the <b>Select statement</b> field.</p> <p>Once this field is populated, you can select the column that contains the primary data you want to query and use for prediction.</p> <div style="border: 1px solid #0070C0; padding: 10px; margin-top: 10px;">  For classification models, this is usually a text field.            For regression models, this is always a numeric field.         </div>
<b>Train Project</b>	<p>This lets you train the project, i.e. the project examines and learns the data.</p> <div style="border: 1px solid #0070C0; padding: 10px; margin-top: 10px;">  Predictions can only be made on trained projects.         </div>
<b>Copy to new version</b>	<p>This lets you create a new version of an existing project on <b>system</b>, <b>company</b> or <b>role</b> level.</p> <div style="border: 1px solid #0070C0; padding: 10px; margin-top: 10px;">  You typically copy a shipped project listed in the <b>SYSPRO</b> category to use that project as a basis to work with.            You can also copy a project located in the <b>System</b> or <b>Company</b> category to create a new version of an existing project.         </div>
<b>Advanced options</b>	<p>This displays the <b>Advanced option</b> pane, where you can select additional training and toolkit options.</p>
<b>Activate project</b>	<p>This lets you activate a project.</p> <div style="border: 1px solid #0070C0; padding: 10px; margin-top: 10px;">  SYSPRO projects that are shipped with Avanti must be copied to a new version, before they can be activated.            Only active projects can be used for predictions.         </div>
<b>Deactivate project</b>	<p>This lets you deactivate a project that is no longer used for predictions.</p>

Field	Description
<b>Download project</b>	This lets you download the selected project, so that you can later import and train it. Once downloaded, you can find the .proj file in your Windows <b>Downloads</b> folder.

## Model information

This displays the basic statistics about the selected trained project.

Some objective metric is required to compare and identify the better model. The ML Engine supports various metrics that are defined in the project using the **Scoring method**.

The data set is split into a training set and a testing set. The ML Engine learns from the training set and isn't allowed to see the testing set. Once a model is created, its performance on unseen data can be measured by passing the testing set into the model and comparing predicted values in the training set. This process of comparing predicted values to actual values can be used to objectively score the model. The displayed scores are the result of this comparison.

Field	Description
Scoring method	This indicates the method used to evaluate the quality of the given pipeline for the classification problem.
Total set score	This indicates how the model scored over the entire data set i.e. both the training and testing data sets combined.
Possible error	This indicates how the model performed on the testing data set only, and is therefore a good indication of how well it will predict the right outcome on unseen data.
Test observations	This indicates the total number of records used for the selected prediction results.
Training observations	This indicates the total number of records returned from the SQL statement.
<b>Perform prediction</b>	This lets you perform a prediction on the trained model. <div style="border: 1px solid #ccc; padding: 10px; margin-top: 10px;">  You can only perform predictions on trained models in active projects.           </div>

## Feature importance

This displays the **Feature importance** graph containing a score for each SQL column in the AI project - indicating which data most affects the predicted value.

The score percentage against each column indicates the frequency according to which the data was used for predictions and therefore the significance of the data in the project. This information allows you to evaluate the data used in the project and delete rarely-used columns.

The **Feature importance** graph and a link to the corresponding data is also displayed when selecting an AI tile in **SYSPRO Avanti**.



You must retrain existing projects to recreate the compact model and include the feature importance.

# New Project

This window is displayed when you select the **New machine learning project** button on the toolbar.

Field	Description
<b>Save project</b>	This lets you save the project.
Project	<p>Select the type of project model you want to create.</p> <p><i>Classification</i></p> <p>Select this if you want the prediction result to be a condition (e.g. good or bad).</p> <p>The <b>Output column</b> is usually a text field.</p> <p>Accuracy (the closeness of a measured value to a standard or known value) is used to determine the condition.</p> <p><i>Regression</i></p> <p>Select this if you want the prediction result to be a quantity or an amount (e.g. days late).</p> <p>The <b>Output column</b> is usually a numeric field.</p> <p>Mean squared error (average squared difference between the estimated values and what is estimated) is used to determine the regression.</p> <p><i>Anomaly</i></p> <p>Select this model if you want to predict an irregularity or inconsistency within the data source.</p> <p>There are no scoring methods for anomaly type projects. Anomaly detectors can't be scored as they are unsupervised machine learning algorithms and don't use analogous scoring metrics.</p>
Name	Indicate the project name.
Description	Enter a description for the project.
Data source	<p>Select the data source from the list.</p> <p>Data sources are automatically created for every SYSPRO company in your database.</p> <div style="border: 1px solid #ccc; padding: 10px; margin-top: 10px;">  You can create your own data sources for data that lies outside of SYSPRO.         </div>

Field	Description
Status	This indicates the status of the selected project.
Select statement	<p>Enter the select statement (usually a SQL query) that will extract the required data from the database.</p> <p>Select <b>Generate output columns</b> to populate the column names in the <b>Output column</b> field.</p>
Output column	<p>Once this field is populated, you can select the column that contains the primary data you want to query and use for prediction.</p> <div style="border: 1px solid #ccc; padding: 10px; margin-top: 10px;">  For classification models, this is usually a text field.            For regression models, this is always a numeric field.         </div>

## Advanced Options

The advanced machine learning options are intended for data scientists and allow an increased training time for projects with a lot of data.

Field	Description
<b>Training options</b>	This applies to the entire machine learning project.
Random state	<p>This ensures that TPOT (Tree-based Pipeline Optimization Tool) gives you the same results each time you run it against the same data set using that seed.</p> <p>The data type is an integer and can be set to null.</p>
Scoring method	<p>This function is used to evaluate the quality of the given pipeline for the classification problem.</p> <div style="border: 1px solid #ccc; padding: 10px; margin-top: 10px;">  The available regression and classification functions are listed in the Using section.         </div>
Test ratio	<p>This indicates the percentage of the data set that is used to test the trained model that is generated.</p> <p><i>For example:</i></p> <p>0.2 means that 20% of the data that is returned from the select statement will be used for testing.</p>

Field	Description
<b>Toolkit setup</b>	
Generations	<p>This indicates the number of iterations required to run the pipe line optimization process.</p> <p>This defaults to 100 and must be a positive number.</p> <div style="border: 1px solid #0070C0; padding: 10px; margin-top: 10px;">  The more generations, the better TPOT works, as it evaluates the population size + generation * offspring_size pipelines in total.         </div>
Population size	<p>This indicates the number of individuals to retain in the generic programming population of every generation.</p> <p>This defaults to 100, but the greater the number, the better the prediction outcome.</p>
Offspring	<p>This indicates the number of offspring to produce in each genetic programming generation.</p> <p>This has a default value of 100, and must be a positive number.</p>
Mutation rate	<p>The mutation rate for the genetic programming algorithm indicates how many pipelines apply random changes to every generation.</p>
Cross over rate	<p>This indicates how many pipelines to breed for every generation.</p> <p>The range is [0.0, 1.0], with 0.1 being the default value.</p> <div style="border: 1px solid #0070C0; padding: 10px; margin-top: 10px;">  The mutation_rate + crossover_rate must not exceed 1.0.         </div>
Number of folds to evaluate	<p>This indicates the folds that evaluate each pipeline in k-fold cross validation during the optimization process.</p> <p>This is the cv number and must be an integer. The default value is 5.</p>
Sub sample	<p>This indicates the fraction of training samples that are used during the optimization process.</p> <p>It must be in the range [0.0, 1.0], with 1.0 being the default value.</p>
Max time	<p>This indicates the time the machine learning engine spends on optimizing the pipeline.</p> <p>This must be a positive integer.</p>

Field	Description
Max evaluation time	<p>This indicates the time the machine learning engine spends on evaluating a single pipeline.</p> <p>This must be a positive integer.</p> <div style="border: 1px solid #ccc; padding: 10px; margin-top: 10px;">  Setting this to a higher number allows the machine learning engine to consider pipelines that are more complex, and increases the time it takes to train the project.         </div>
Configuration dictionary	This lets you customize the operators and parameters that TPOT searches in the optimization process.
Early stop	<p>This indicates the number of generations TPOT uses to check if there are improvements in the optimization process.</p> <p>If no improvements are found, the optimization process ends.</p>

## Save For

This pane is displayed when you select a shipped project (in the SYSPRO folder) and then select the **Copy to a new version** button.

This lets you duplicate the project to create various versions, while keeping the original project intact. You can then edit the project versions by changing the SQL statement and/or data source.

Field	Description
Save for	<p>This lets you specify the level of the project, which can be:</p> <ul style="list-style-type: none"> <li>■ System</li> <li>■ Company</li> <li>■ Role</li> </ul>
Company	This lets you select the company for which the project will be saved.
Role	This lets you select the role for which the project will be saved.
Data source	This lets you select the data source that will be used for the project.

# Import project

The window is displayed when you select the **Import project** button.

Field	Description
Import path	<p>This lets you navigate to the path where the project was downloaded.</p> <p>Once you have selected the .proj file, the remaining fields are auto-populated.</p>
Save for	This lets you select the level of the project.
Company	This indicates the company that was specified when the project was created, if the project was created at company level.
Role	This indicates the role that was specified when the project was created, if the project was created at role level.
Data source	This indicates the data source that is used for the project.
ProjectVersion	This indicates the version of the project.
Train project	Select this to add the project to the <b>AI Training Queue</b> as soon as it is imported.
Model	This indicates the type of project model for this project.
Name	This indicates the name of the project and can be edited.
Description	This indicates the description of the project and can be edited.
Select statement	<p>Enter the select statement (usually a SQL query) that will extract the required data from the database.</p> <p>Select <b>Generate output columns</b> to populate the column names in the <b>Output column</b> field.</p>
Output column	<p>Once this field is populated, you can select the column that contains the primary data you want to query and use for prediction.</p> <div style="border: 1px solid #ccc; padding: 10px; margin-top: 10px;">  For classification models, this is usually a text field.            For regression models, this is always a numeric field.         </div>

## New data source

This pane is displayed when you select **New Data source** from the **List of Data Sources** window.

Data sources you have created, are saved in the [SaiDataSource](#) table in the system-wide database.

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	<p>Enter the connection string. This must be in the following format:</p> <p>DRIVER={ SQL Server} ; SERVER={ SERVERNAME } ; DATABASE={ Database Name } ; UID={ user } ; PWD={ Password }</p> <p>Select <b>Test connection</b> to confirm that the connection is working.</p>

# Sample project details

Chance order will be late

Field	Description
Description	<p>This displays a percentage that predicts the chance a purchase order line may be late.</p> <p>This is based on all orders placed in the past as well as current information about the supplier.</p>
Model type	Regression
Output column	IsLate
SQL query	<pre> select PTL.IsLate , PMH.OrderStatus, PMH.NextDetailLine, PMH.ExchangeRate, PMH.TermsCode, APS.CurrentBalance, APS.SupplierClass, APS.PyrEndBalance, APS.YtdAmount1099, APS.NumMonthsZero, APS.LastChequeVal from (Select top (1000) PMH.PurchaseOrder, case when sum(datediff(day, MLatestDueDate, MLastReceiptDat)) &gt; 0 then 1 else 0 End as IsLate from PorMasterDetail as PMD inner join PorMasterHdr as PMH on(PMD.PurchaseOrder=PMH.PurchaseOrder) where (PMD.MLatestDueDate is not null)and(PMD.MLastReceiptDat is not null) group by PMH.PurchaseOrder) PTL join PorMasterHdr PMH on PMH.PurchaseOrder = PTL.PurchaseOrder join ApSupplier APS on APS.Supplier = PMH.Supplier </pre>

## LCT days late

Field	Description
Description	This predicts how late a shipment may be, measured in days.
Model type	Regression
Output column	DaysLate

Field	Description
SQL query	<pre> select datediff(day,LH.EstArrivalDate, LH.ActArrivalDate) as DaysLate, case when LTRIM(RTRIM(LH.PlaceOfShipment)) = '' then null else lower(LH.PlaceOfShipment) end as PlaceOfShipment, case when LTRIM(RTRIM(LH.PlaceOfDestination)) = '' then null else lower(LH.PlaceOfDestination) end as PlaceOfDestination, LH.Route, LR.DaysBeforeEta, LR.DemurrageDays, LR.DaysAfterEta from LctShipmentHdr LH left join LctRoute LR on LH.Route = LR.Route where LH.ActArrivalDate is not null </pre>

## PO days late

Field	Description
Description	This predicts the number of days a purchase order line may be late.
Model type	Regression
Output column	DaysLate
SQL query	<pre> select PMH.Supplier, PMD.MStockCode, PMD.MWarehouse, PMD.MOrderQty, PMD.MReceivedQty, PMD.MOrigDueDate, PMD.MLatestDueDate, PMD.MPrice, PMD.MProductClass, datediff(day, MLatestDueDate, MLastReceiptDat) as DaysLate from PorMasterDetail as PMD inner join PorMasterHdr as PMH on(PMD.PurchaseOrder=PMH.PurchaseOrder) where (PMD.MLatestDueDate is not null)and(PMD.MLastReceiptDat is not null) </pre>

## Customer invoice pay date

Field	Description
Description	This predicts the number of days a customer may take to pay an invoice.
Model type	Regression
Output column	DaysToPayEstimate

Field	Description
SQL query	<pre> ;With TheInvoicePayments as (     select         IP.Customer,         IP.Invoice,         max(IP.JournalDate) as MaxJournalDate     from ArInvoicePay IP with (nolock)     group by IP.Invoice, IP.Customer ) select     I.Customer,     I.Invoice,     I.InvoiceDate,     I.OrigDiscValue,     I.InvoiceBal1,     I.InvoiceBal2,     I.InvoiceBal3,     I.CurrencyValue,     C.Area,     C.ApplyLineDisc,     C.ApplyOrdDisc,     C.BalanceType,     C.Contact,     C.Creditlimit,     C.CreditStatus,     C.Currency,     C.DateLastPay,     C.DateLastSale,     C.HighestBalanc,     C.HighInvDays,     C.Name,     C.MinimumOrderValue,     C.PaymentsAllowed,     C.PriceCode,     C.ShippingInstrs,     C.Salesperson,     C.Salesperson1,     C.Salesperson2,     C.Salesperson3,     C.CustomerOnHold,     case         when             IP.MaxJournalDate is null             then Datediff(day, I.InvoiceDate, Dateadd(day, 90, I.InvoiceDate))         when I.InvoiceBal1 &gt; 0 OR I.InvoiceBal2 &gt; 0 OR I.InvoiceBal3 &gt; 0             then Datediff(day, I.InvoiceDate, Dateadd(day, 60, I.InvoiceDate))         else Datediff(day, I.InvoiceDate, IP.MaxJournalDate)     end as DaysToPayEstimate From ArInvoice with (nolock) left join TheInvoicePayments IP with (nolock) on (I.Customer = IP.Customer and I.Invoice = IP.Invoice) left join ArCustomer C with (nolock) on (I.Customer = C.Customer) where I.DocumentType = 'I'</pre>

## Job status

Field	Description
Description	<p>This predicts whether a job may be completed early or late.</p> <p>The outcome of the prediction displays <b>On time</b> or <b>Late</b>.</p>
Model type	Classification

Field	Description
Output column	JobStatus
SQL query	<pre>select top 1000 JobDescription, JobClassification, StockCode, Warehouse, JobTenderDate, JobDeliveryDate, JobStartDate, OrigDueDate, case when (ActCompleteDate&gt;OrigDueDate) then 'Late' else 'On time' end as JobStatus from [WipMaster] where (OrigDueDate is not null)</pre>

## Lost sales reason

Field	Description
Description	This prediction determines whether sales may be lost and what the reason for losses would be.
Model type	Classification
Output column	ReasonLost
SQL query	<pre>select CostValue, QuantityLost, QuantityRequested, QuantityUom, Price, PriceUom, RC.Description as ReasonLost, IW.QtyAllocated, IW.QtyOnHand, IW.QtyOnOrder, IW.QtyAllocatedWip, IW.QtyDispatched, IW.MinimumQty, IW.MaximumQty, IW.SafetyStockQty, IW.ReOrderQty, CM.CreditLimit, CM.HighestBalance, CM.OutstOrdVal, CM.NumOutstOrd, CM.ApplyOrdDisc, CM.BackOrdReqd from [dbo].[SorLostSales] LS join [dbo].TblSoReason RC on RC.ReasonCode = LS.ReasonCode join [dbo].InvWarehouse IW on IW.StockCode = LS.StockCode and IW.Warehouse = LS.Warehouse join [dbo].ArCustomer CM on CM.Customer = LS.Customer</pre>

## Problem with sales order

Field	Description
Description	<p>This predicts the most likely reason why a sales order is returned.</p> <p>This is based on the reasons past sales order items were returned.</p>
Model type	Classification
Output column	ProblemDesc
SQL query	<pre> select     RP.ProblemDesc,     SD.MOrderQty,     SD.MWarehouse,     SD.MBin,     SD.MPrice,     IM.Supplier,     M.Mass,     IM.Volume from [dbo].[RmaDetail] RD join SorDetail SD on SD.SalesOrder = RD.SalesOrder and SD.SalesOrderLine = SD.SalesOrderLine join RmaProblemCodes RP on RP.ProblemCode = RD.ProblemCode join InvMaster IM on SD.MStockCode = IM.StockCode </pre>

## Quote success

Field	Description
Description	<p>This predicts how many quotes should be accepted by the customer and result in sales orders.</p> <p>The model returns either <code>SALE</code> or <code>NOSALE</code> based on historical quotations and the subsequent conversion into sales orders.</p>
Model type	Classification
Output column	QuoteStatus

Field	Description
SQL query	<pre> declare @startDate datetime = dateadd(year, -2, current_timestamp); declare @endDate datetime = current_timestamp; with _quoteMaster as ( select QM.Quote, QM.DateEnquiry, QM.DateDeliveryReq from QotMaster as QM where (QM.DateEnquiry between @startDate and @endDate) ) ,_quoteOffer as ( select Quote, QuoteVersion, Line, MQuantity, MPrice, MDiscountValue, MNetValue, MCost, MLeadTime from QotOffer ) ,_quoteDetail as( select Quote, QuoteVersion, Line, ProductClass, MStockCode from QotDetail), _quoteSalesLines as( select QO.Quote, QO.QuoteVersion, QO.Line, QO.MQuantity, QO.MPrice, QO.MDiscountValue, QO.MNetValue, QO.MCost, QO.MLeadTime, QD.ProductClass, QD.MStockCode from _quoteOffer as QO join _quoteDetail as QD on((QO.Quote=QD.Quote)and(QO.QuoteVersion=QD.QuoteVersion)and(QO.Line=QD.Line))), _quoteSalesLinesSummary as( select QSL.Quote, QSL.QuoteVersion, count(QSL.Line) as TotalLines, sum(QSL.MQuantity) as MQuantity, sum(QSL.MPrice) as MPrice, sum(QSL.MDiscountValue) as MDiscountValue, sum(QSL.MNetValue) as MNetValue, sum(QSL.MCost) as MCost, max(QSL.MLeadTime) as MLeadTime, count(distinct QSL.ProductClass) as UniqueProductClasses, count(distinct QSL.MStockCode) as UniqueStockCodes from _quoteSalesLines as QSL group by QSL.Quote, QSL.QuoteVersion), _quoteSales as( select SalesOrder, Quote, QuoteVersion, Customer, Salesperson, ReqShipDate, Area, </pre>

Field	Description
	<pre> CashCredit, case when ltrim(CompanyTaxNo)='' then 'NON-VENDOR' else 'VENDOR' end as TaxStatus from QotSoMasterHdrwhere (ReqShipDate between @startDate and @endDate) select top (50000) QS.SalesOrder, QS.Quote, QS.QuoteVersion, QS.Customer, QS.Salesperson, QS.ReqShipDate, QS.Area, QS.CashCredit, QS.TaxStatus, QSLSTotalLines, QSLSMQuantity, QSLSMPrice, QSLSMDiscountValue, QSLSMNetValue, QSLSMCost, QSLSMLeadTime, case when (QSLS.Quote is null) then 'NOSALE' else 'SALE' end as QuoteStatus from _quoteSales as QS left join _quoteSalesLinesSummary as QSLS on((QS.Quote=QSLS.Quote)and(QS.QuoteVersion=QSLS.QuoteVersion))); </pre>

## Stock code profitability

Field	Description
Description	<p>This predicts how profitable a stock code should be by location.</p> <p>The stock code will be grouped in a bad, medium, good or great category, which is based on the profitability and returns of similar stock codes. The categories are hard coded as follows:</p> <ul style="list-style-type: none"> <li>■ Below 0.</li> <li>■ Between 0 and 10000.</li> <li>■ Between 10000 and 100000.</li> <li>■ Above 100000.</li> </ul>
Model type	Classification
Output column	ProfitCategory

Field	Description
SQL query	<pre> with _profitByStockCode as (   select     IM.StockCode,     sum(SD.MOrderQty*(SD.MPrice-SD.MUnitCost))     as TotalProfit   from SorDetail SD   join InvMaster IM on IM.StockCode = SD.MStockCode   where SD.MStockCode is not null   group by IM.StockCode ) , profitStats as (   select     avg(TotalProfit) as Avg_TotalProfit,     min(TotalProfit) as Min_TotalProfit,     max(TotalProfit) as Max_TotalProfit     from _profitByStockCode as SCP ) , detailedData as (   select     SCP.StockCode,     IM.AlternateUom,     IM.OtherUom,     IM.ConvMulDiv,     IM.ConvFactAltUom,     IM.Mass,     IM.Volume,     IM.Supplier,     IM.ProductClass,     IM.Buyer,     IM.LeadTime,     cast(SCP.TotalProfit as decimal(18,2)) as     TotalProfit   from _profitByStockCode as SCP   join InvMaster as IM   on IM.StockCode = SCP.StockCode ) select   StockCode,   AlternateUom,   OtherUom,   ConvMulDiv,   ConvFactAltUom,   Mass,   Volume,   Supplier,   ProductClass,   Buyer,   LeadTime,   --TotalProfit,   case   when TotalProfit &lt;0 then 'BAD'   when TotalProfit between 0 and 10000 then 'MEDIUM'   when TotalProfit between 10000 and 100000 then   'GOOD'   else 'GREAT'   end as ProfitCategory   from detailedData as DD </pre>

## Supplier performance

Field	Description
Description	<p>This predicts how well a supplier should perform in terms of deliveries, and indicates the likelihood of the delivery being on time and in full.</p> <p>The possible prediction outcomes are as follows:</p> <ul style="list-style-type: none"> <li>■ Delivers late, in full.</li> <li>■ Delivers late, low stock.</li> <li>■ Delivers late, with extra.</li> <li>■ Delivers on time, in full.</li> <li>■ Delivers on time, low stock.</li> <li>■ Delivers on time, with extra stock.</li> <li>■ Delivers early, in full.</li> <li>■ Delivers early, low stock.</li> <li>■ Delivers early, with extra.</li> </ul>
Model type	Classification
Output column	<a href="#">SupplierRanking</a>

Field	Description
SQL query	<pre> select PMH.Supplier, PMD.MStockCode as StockCode, PMD.MStockDes as [Description], PMD.MWarehouse as Warehouse, PMD.MOrderQty as OrderQty, PMD.MReceivedQty as ReceivedQty, PMD.MLatestDueDate as LatestDueDate, PMD.MLastReceiptDat as LastReceiptDat, PMD.MProductClass as ProductClass, PMD.MCompleteFlag as CompleteFlag, PMD.MOrigDueDate as OrigDueDate, PMD.MReschedDueDate as ReschedDueDate, PMH.ShippingLocation, PMH.[State] as [State], PMH.ShippingInstrs, case when MLastReceiptDat &gt; MLatestDueDate and MReceivedQty = MOrderQty then 'Delivers late, in full' when MLastReceiptDat &gt; MLatestDueDate and MReceivedQty &lt; MOrderQty then 'Delivers late, low stock' when MLastReceiptDat &gt; MLatestDueDate and MReceivedQty &gt; MOrderQty then 'Delivers late, with extra' when MLastReceiptDat = MLatestDueDate and MReceivedQty = MOrderQty then 'Delivers on time, In full' when MLastReceiptDat = MLatestDueDate and MReceivedQty &lt; MOrderQty then 'Delivers on time, low stock' when MLastReceiptDat = MLatestDueDate and MReceivedQty &gt; MOrderQty then 'Delivers on time, With extra stock' when MLastReceiptDat &lt; MLatestDueDate and MReceivedQty = MOrderQty then 'Delivers early, in full' when MLastReceiptDat &lt; MLatestDueDate and MReceivedQty &lt; MOrderQty then 'Delivers early, low stock' when MLastReceiptDat &lt; MLatestDueDate and MReceivedQty &gt; MOrderQty then 'Delivers early, with extra' end as SupplierRanking From PorMasterDetail PMD join PorMasterHdr PMH on PMD.PurchaseOrder = PMH.PurchaseOrder where MStockCode &lt;&gt; '' </pre>

## AP invoice payment

Field	Description
Description	This predicts by when a supplier should pay an invoice. The outcome of the prediction displays <b>On time</b> or <b>Late</b> .
Model type	Classification
Output column	Status

Field	Description
SQL query	<pre> with _invoiceMaster as (   select     Supplier,     Invoice,     InvoiceDate,     DueDate,     OrigInvValue,     Currency,     CurrencyValue,     from ApInvoice ) , _invoicePaid as (   select     Supplier,     Invoice,     EntryNumber,     TrnType,     JournalDate,     TrnValue,     TrnYear,     TrnMonth,     PostValue,     PostCurrency   from ApInvoicePay   where TrnType='P' ) select top(500)   IM.Supplier,   IM.InvoiceDate,   IM.DueDate,   IM.OrigInvValue,   IM.Currency  ,case when (IP.JournalDate&gt;IM.DueDate) then 'Late' else 'On time' end as Status from _invoiceMaster as IM join _invoicePaid as IP on(   (IM.Supplier = IP.Supplier)   and(IM.Invoice = IP.Invoice) ) </pre>

## Customer classification

Field	Description
Description	<p>This predicts how profitable a customer would be.</p> <p>The customer will be grouped in a bad, medium, good or great category, which is based on the profitability and returns of similar customers.</p> <p>The categories are calculated using the mean profitability of all customers and the maximum profitability.</p>
Model type	Classification
Output column	ProfitCategory

Field	Description
SQL query	<pre> with _profitByCustomer as (select SM.Customer, sum(SD.MOrderQty*(SD.MPrice-SD.MUnitCost))as TotalProfit from SorDetail SD join SorMaster SM on SM.SalesOrder = SD.SalesOrder where SD.MStockCode is not null group by SM.Customer), profitStats as( select avg(TotalProfit)*0.2 as Avg_TotalProfit, min(TotalProfit) as Min_TotalProfit, max(TotalProfit) as Max_TotalProfitfrom _profitByCustomer as SCP), detailedData as( select AC.Customer, AC.Name, AC.CreditLimit, AC.Salesperson, AC.Area, AC.Branch, AC.TermsCode, AC.OutstOrdVal, AC.NumOutstOrd, AC.HighestBalance, AC.SoldPostalCode, AC.ShipPostalCode, AC.RouteCode, AC.StoreNumber, AC.MasterAccount, cast(SCP.TotalProfit as decimal(18,2)) asTotalProfitfrom _profitByCustomer as SCP join ArCustomer as ACon AC.Customer = SCP.Customer ) select CreditLimit, Salesperson, Area, Branch, TermsCode, SoldPostalCode, ShipPostalCode, RouteCode, StoreNumber, MasterAccount, TotalProfit, case when TotalProfit &lt;0 then 'bad' when TotalProfit between 0 and Avg_TotalProfit then 'medium' when TotalProfit between Avg_TotalProfit andMax_TotalProfit-Avg_TotalProfit then 'good' else 'great' end as ProfitCategory from detailedData as DD join profitStats on 1=1 order by TotalProfit </pre>

## Sales order line anomaly

Field	Description
Description	This detects anomalies on sales order lines based on the selected columns.
Model type	Anomaly
Output column	The <b>Output column</b> doesn't apply to anomaly projects.

Field	Description
SQL query	<pre> select MStockCode, MWarehouse, MOrderQty, MOrderUom, MPrice, MPriceUom, MCustRequestDat from SorDetail where (MStockCode is not null) and (MWarehouse is not null) and (MOrderQty is not null) and (MOrderUom is not null) and (MPrice is not null) and (MPriceUom is not null) and (MCustRequestDat is not null) </pre>

## Purchase order line anomaly

Field	Description
Description	This detects anomalies on purchase order lines based on the selected columns.
Model type	Anomaly
Output column	The <b>Output column</b> doesn't apply to anomaly projects.
SQL query	<pre> select MStockCode, MWarehouse, MOrderQty, MOrderUom, MPrice, MPriceUom, MLatestDueDate from PorMasterDetail where (MStockCode is not null) and (MWarehouse is not null) and (MOrderQty is not null) and (MOrderUom is not null) and (MPrice is not null) and (MPriceUom is not null) and (MLatestDueDate is not null) </pre>



[www.syspro.com](http://www.syspro.com)

Copyright © SYSPRO. All rights reserved.  
All brand and product names are trademarks or  
registered trademarks of their respective holders.