



Polynomial Neural Network Discovery Client User Guide

Version 1.3

Table of contents

Table of contents	2
1. Introduction	3
1.1 Overview.....	3
1.2 PNN algorithm principles.....	3
1.3 Additional criteria.....	3
1.4 PNN algorithm features.....	4
1.5 PNN data format.....	5
2. General information	6
2. General information	6
2.1 System requirements.....	6
2.2 Package contents.....	6
2.3 Installation and software activation.....	6
2.4 How to get PNN Discovery Client.....	6
2.5 Technical support.....	6
3. Using PNN Discovery Client	7
3.1 Getting started.....	7
3.2 Project-oriented approach.....	8
3.3.1 Opening project.....	8
3.3.2 Creating new project.....	8
3.3.3 Saving project.....	8
3.3 PNN framework toolbar and menu.....	8
3.4.1 Toolbar buttons.....	8
3.4.2 Additional menu items.....	9
3.4 Panels.....	9
3.5.1 Table View.....	9
3.5.2 HTML View.....	12
3.5.3 Chart View.....	14
3.5.4 Plot View.....	17
3.5 Project options.....	19
3.6 Calculations.....	22
3.7 Import and export.....	25
3.7.1 Text Import Wizard.....	25
3.7.2 Microsoft Excel Import Wizard.....	27
3.7.3 Export Wizard.....	30

1. Introduction

1.1 Overview

PNN is self-organizing multi-layered iterative Polynomial Neural Network algorithm. It is one of the most promising solutions from GMDH-type neural network algorithms family. PNN supports both linear and non-linear polynomial regression models.

Polynomial Neural Network (PNN) algorithm can be used to analyze complex data sets, find out internal data relationships and present knowledge about these relationships in the form of mathematical description (polynomial regression equation).

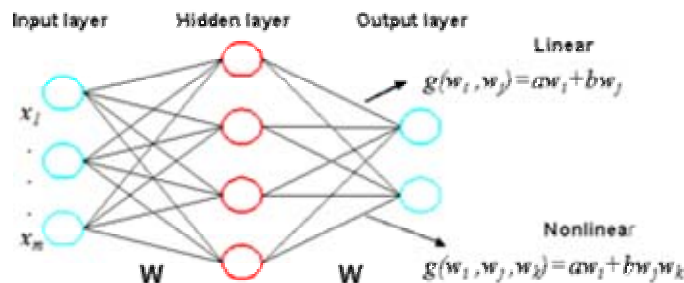
Application area of PNN algorithm may include any domain where sets of observation data should be analyzed and data relationship models should be built. Chemistry (QSAR), economical systems analysis, stocks and financial market instruments, insurance risks study and medical diagnostics can be considered as good examples.

1.2 PNN algorithm principles

To be processed with PNN algorithm, the data have to be organized in a specific way. Let us give definitions for terms used throughout this guide. A set of observed quantitative parameters $\{x_1, x_2, \dots, x_m\}$ forms the vector of input variables x . Each observation x (input) has an "already known" result y (output). The purpose of the algorithm is to find a subset of variables $\{x_{i1}, x_{i2}, \dots, x_{ik}\}$ and a model $y = f_i(x_{i1}, x_{i2}, \dots, x_{ik})$ in the set of polynomials of s -power that minimizes some criterion value (CR).

After the model is built, you can use it for prediction of yet "unknown" result y that corresponds to newly observed vector x .

The algorithm is implemented as a GMDH-type Neural Network. First layer generates the models $y = g(x_i, x_j, x_k)$, where x_i, x_j, x_k are input variables. Next layers generate models as $y = g(w_i, w_j, w_k)$, where w_i, w_j, w_k are the models of previous layers that minimize criterion value (CR). See figure below.



1.3 Additional criteria

Least-squares estimation criterion $\min_{\beta_1, \dots, \beta_m} \sum_{i=1}^n \left(y_i - \sum_{j=1}^m \beta_j f_{ji} \right)^2$ is the best for models with normal distribution of errors. Still it is well known that this criterion is sensitive to outliers in the data set. In case outliers are present, Robust PNN uses modified least-squares estimation criterion:

$$\min_{\beta_1, \dots, \beta_m} \sum_{i=1}^n \rho(y_i - \sum_{j=1}^m \beta_j f_{ji})$$

Instead of raising to the second power, function $\rho()$ is used: $\rho(z) = \begin{cases} (1/2)z^2 & \text{if } |z| \leq C \\ C|z| - (1/2)C^2 & \text{if } |z| > C \end{cases}$

The standard criteria of model estimation are:

Rest Minimal Square Errors $RMSE = \sqrt{\frac{1}{n} \sum (y_{pred} - y_{actual})^2}$

$$\sum (y_{pred} - \bar{y}_{pred})(y_{actual} - \bar{y}_{actual})$$

Coefficient of correlation $r^2 = \frac{\sum (y_{pred} - \bar{y}_{pred})(y_{actual} - \bar{y}_{actual})}{\sqrt{\sum (y_{pred} - \bar{y}_{pred})^2 \sum (y_{actual} - \bar{y}_{actual})^2}}$

Cross validated criterion $q^2 = 1 - \frac{\frac{1}{n} \sum (y_{pred} - y_{actual})^2}{\frac{1}{n} \sum (y_{actual} - y_{mean})^2}$

These criteria are based on minimal squares and sensitive to outliers as well as MSE method itself. Therefore, in case outliers are present, Robust PNN uses modifications of standard $RMSE$ and q^2 criteria to characterize models more adequately. The modified criteria use the same $\rho()$ function as least-squares criterion.

$$RMSE = \sqrt{\frac{1}{n} \sum \rho(y_{pred} - y_{actual})}$$

$$q^2 = 1 - \frac{\frac{1}{n} \sum \rho(y_{pred} - y_{actual})}{\frac{1}{n} \sum \rho(y_{actual} - y_{mean})}$$

1.4 PNN algorithm features

The main features of the algorithm can be summarized as follows:

- **Fast learning.** The transforms with two coefficients only are used, for example $\mathbf{g}(\mathbf{w}_i, \mathbf{w}_j) = \mathbf{a} \cdot \mathbf{w}_i + \mathbf{b} \cdot \mathbf{w}_j$ in the linear case. Irrespectively of the power of resulting model and the number of terms the second order matrices are only inverted. The latter enables fast learning of the algorithm.
- **Results in the parametric form.** The polynomial structures are coded using vector of simple numbers that provides the presentation of the results in the parametric form of nonlinear equation.
- **Complexity control against overfitting of model.** Contrary to traditional GMDH algorithm, using of Gedel numbers for model structure coding allows us to control the complexity of each intermediate model during selection. We consider as complexity the vector $(\mathbf{power}, \mathbf{c})^T$, where **power** is the power of the polynomial and **c** is the number of terms. It gives us the possibility to restrict the class of the models under consideration by $\mathbf{power}(\mathbf{w}_i) < \mathbf{p}$ and to search models among the polynomials with power less than **p**.

The maximum complexity is defined by the user or can be automatically selected using a full cross-validation method.

- **Twice-hierarchical neural net structure.** Twice-hierarchical neural net structure is important feature of PNN. One of the problems of traditional GMDH algorithms is that power of polynomials increases too fast. The control of complexity gives us an opportunity to implement the iteration procedure without an increase of the power of polynomials or/and the number of terms.

Twice-hierarchical neural net includes Internal and External nets. Internal iterative procedure provides a wide search under the fixed complexity. External procedure controls the complexity, i.e. the number of the terms and the power of the polynomials.

Moreover, the twice-hierarchical neural net structure provides the convergence of the coefficients. The models, which are calculated as a result of several transformations, have the coefficients that are close to the appropriate regression coefficients.

- **Robust estimation.** To use algorithm in the presence of large errors (outliers) we have developed the PNN algorithm for robust nonlinear (M-regression) model identification (**RPNN**). It allowed us to improve stability of PNN algorithm.

More details about PNN algorithm you can find [here...](#)

1.5 PNN data format

PNN algorithm deals with matrix dataset. Dataset is formally divided into three parts.

First set of rows is involved in “learning” of PNN algorithm and searching best models. These rows are named *training* rows.

Second set of rows could be involved in models estimation and ranking only. We call these rows *test subset*.

And third set of rows could be involved in building approximation values of output values using result models. We call these rows *validation subset*.

2. General information

2.1 System requirements

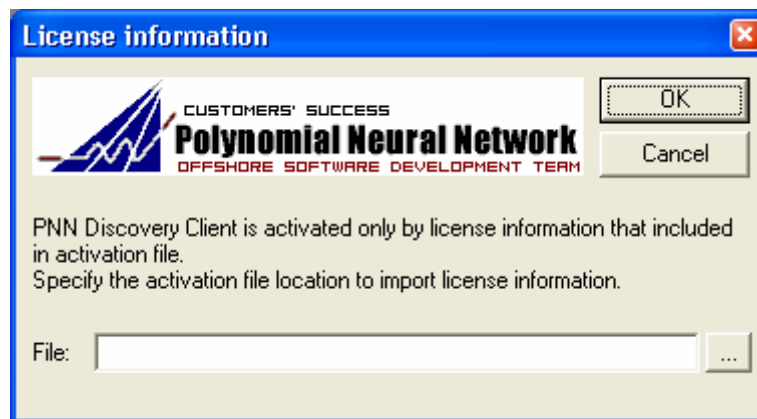
- IBM-Compatible PC with Pentium processor (recommended: Pentium II or better)
- Windows 98, Windows Me, Windows NT 4.0 SP6 or later, Windows 2000 or Windows XP
- Minimum 32 MB RAM (recommended: 64 MB RAM or higher)

2.2 Package contents

- Pnn.exe – PNN Discovery Client
- Pnnb.exe – console application to run batch scenarios using XML files.

2.3 Installation and software activation

Start setup.exe and follow instruction of the installation program. To activate software, enter the full path to the activation file with license information (See figure below).



If you need to renew your license information, click on “Change Activation file...” in Help menu.

2.4 How to get PNN Discovery Client

The majority of PNN products is sold and distributed online. When you purchase from us, you are purchasing registration details (so that you may register software that must be downloaded and installed) with technical support from our friendly staff. The activation file with license information will be sent to you by e-mail.

To download trial version of this product and receive activation file, you have to register on our [site](#) . Trial version is free for using during 30 days from the registration date.

For more details, visit [PNN site](#).

2.5 Technical support

You can send all your questions and problem-solving request to our [staff](#).

Visit our [FAQ](#) and [Tutorial](#) pages to get more information about PNN algorithm and PNN discovery Client.

3. Using PNN Discovery Client

3.1 Getting started

PNN Discovery Client is a full-fledged standalone application running under Windows operation system. PNN Discovery Client allows users to prepare spreadsheet data, find out the best regression models and view obtained results as tabular data or special chart presentations.

PNN Discovery Client offers a range of tools to export data and results in Microsoft Excel workbook, text, HTML and XML format.

The PNN Discovery Client layout consists of two panels (see figure 3.1 below):

- Data view panel
- Results view panel

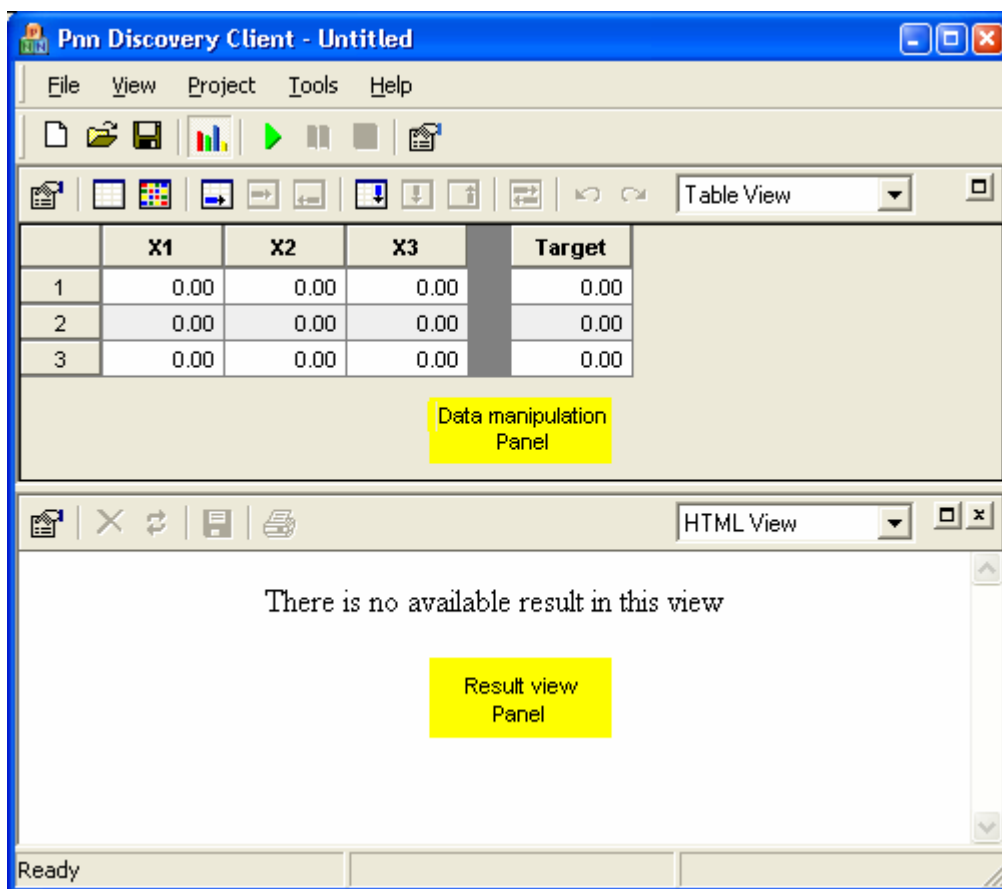


Figure 3.1

Each panel allows selecting a view best suited for one's needs. Figure 3.1 shows default appearance of panels: "Table View" in Data panel and "HTML View" in Results panel.

Toolbar buttons provide quick access to general commands and functions (see figure 3.2).

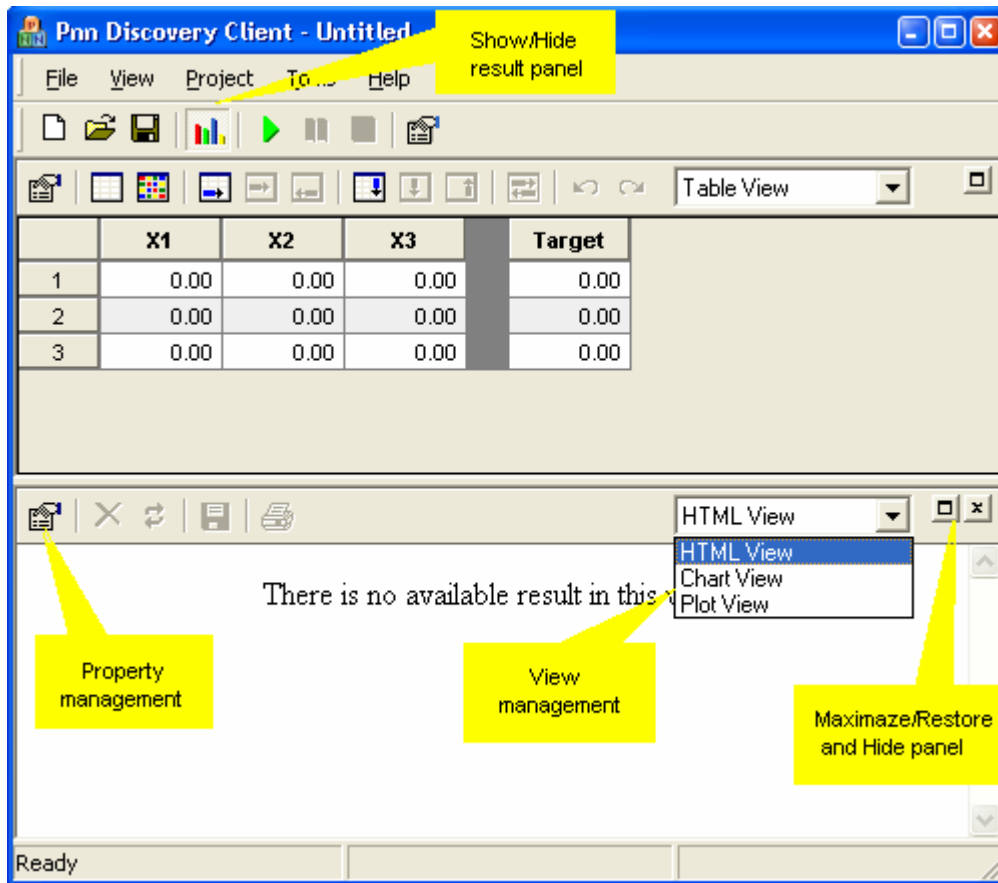



Figure 3.2


3.2 Project-oriented approach

PNN Discovery Client has project-oriented design for managing separate analyses (projects) in multi-threading environment under a single PNN framework. PNN project contains original data, results of analysis and configuration settings. When saved on disk, the PNN project file by default has the extension “.pnn”.


3.3.1 Opening project

Click on  button in the toolbar or select “File”->“Open...” menu command to open previously saved project file. Alternatively, double click on the file in Windows Explorer. Most recent projects can be accessed from the “File”->“Recent Files” submenu.

3.3.2 Creating new project

Click on  button in the toolbar or select “File”->“New” menu command to open new project. New project will be opened in the separate window.










3.3.3 Saving project

Click on  button in the toolbar or select “File”->“Save” menu command to save project file. Standard Windows Save File dialog appears.

3.3 PNN framework toolbar and menu

3.4.1 Toolbar buttons



-  **New Project** – create new project in separate window
-  **Open Project file** – open previously saved project file
-  **Save Project** – save current project data, settings and results in the “.pnn” file
-  **Show/hide resulting panel** – show or hide resulting panel in the framework.
-  **Search models** – start PNN algorithm calculations
-  **Stop** – break algorithm execution
-  **Pause** – suspend/proceed algorithm execution
-  **Settings** – set PNN algorithm parameters
-  **Apply models** – input dataset substitution to resulting models and calculating approximation values

3.4.2 Additional menu items

Some functionality is available only through PNN framework menu:

- **Options** – general settings for PNN Discovery Client framework
- **Split Vertically** – you can choose to split main framework vertically or horizontally
- **Import External Data** – invokes Wizard to import data from plain text files, XML files and Microsoft Excel workbooks
- **Export To** – invokes Wizard to export data and/or results to one of the following formats: tab delimited plain text (*.txt), comma delimited plain text (*.csv), XML (.xml) and Microsoft Excel workbook (.xls)

3.4 Panels

Every panel has its own toolbar that allows you to select and customize the view. On the Data panel you can add, remove and edit rows and columns of the spreadsheet. Current version supports the only view for Data panel:

- **Table View**

Results panel provides an access to results presentation in different ways. The following views are available for Results panel:

- **HTML View**
- **Chart View**
- **Plot View**

Data panel is always open in the PNN framework and cannot be closed. Data panel can be hidden. Results panel can be closed or hidden anytime.


3.5.1 Table View

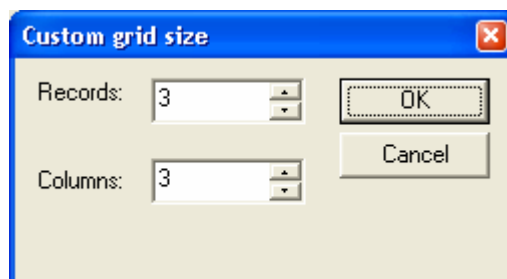
Table View displays data sheet that allows user add/delete/modify rows and columns. Redo/Undo functionality is also available. (See figure 3.5.1).

	Col1	Col2	Col3	Col4	Col5	Target
1	1.	2.	5.	0.	3.	38.
2	1.	3.	1.	4.	8.	520.
3	3.	1.	0.	8.	4.	200.
4	2.	2.	2.	0.	4.	139.
5	5.	1.	2.	1.	8.	2571.
6	2.	5.	1.	2.	9.	1457.
7	8.	0.	2.	5.	18.	46666.
8	5.	1.	1.	1.	7.	1723.
9	2.	3.	4.	5.	25.	31360.
10	4.	3.	2.	1.	9.	2929.
11	9.	5.	1.	9.	23.	109517.
12	8.	1.	5.	0.	9.	5832.

Figure 3.5.1

Columns of the data sheet represent independent variables and rows are observations of these variables. Last column of the data sheet is named “target” column and contains observable function value. The data sheet should have at least 3 rows and 3 columns (not including target column). Default size of Table View data sheet is 3x3.

The  button in the control toolbar allows you to set custom initial size of the data sheet (see figure below).




The  button in the control toolbar clears all data and restores initial size of the data sheet.

Table View Toolbar

Table below describes Table View toolbar functionality:












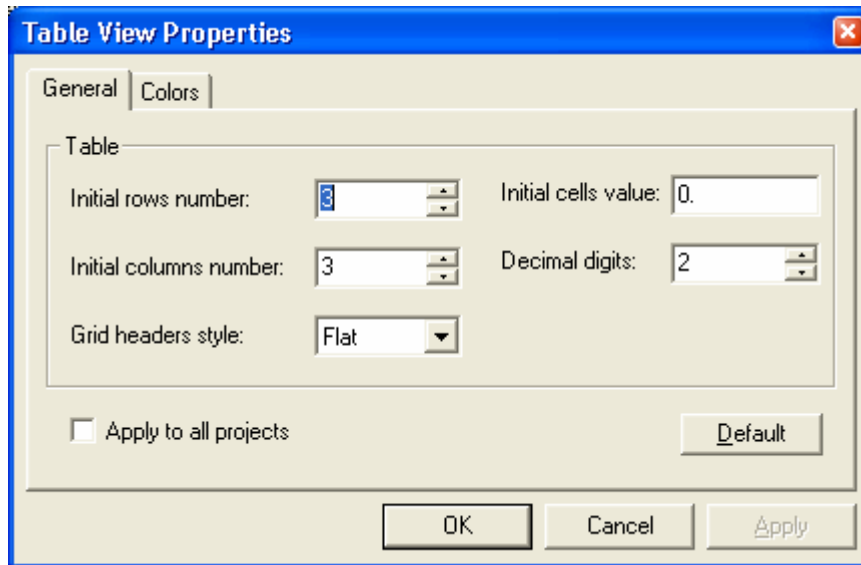
-  **Custom grid.** Set custom initial size of the data sheet
-  **Clear / Reset table.** Clear all data and restore initial size of the data sheet
-  **Add row.** Add new row at the end of the data sheet
-  **Insert row.** Insert row before current selection in the data sheet. Some cell or rows must be selected in the data sheet
-  **Delete row.** Delete selected row(s). Some cell or rows must be selected in the data sheet
-  **Add column.** Add new column as last variable column
-  **Insert column.** Insert column before current selection in the data sheet. Some cell or columns must be selected in the data sheet
-  **Delete column.** Delete selected column(s). Some cell or columns must be selected in the data sheet
-  **Swap column with Target.** Swap selected column with target column
-  **Undo last changes**
-  **Redo last changes**

Table View Properties

Click on  button in the control toolbar to invoke Table View Properties dialog. Or you can invoke it from the Table View context menu. Table View Properties dialog is shown below.



General tab:

Initial rows size – initial number of rows in the data sheet

Initial columns size – initial number of columns in the data sheet

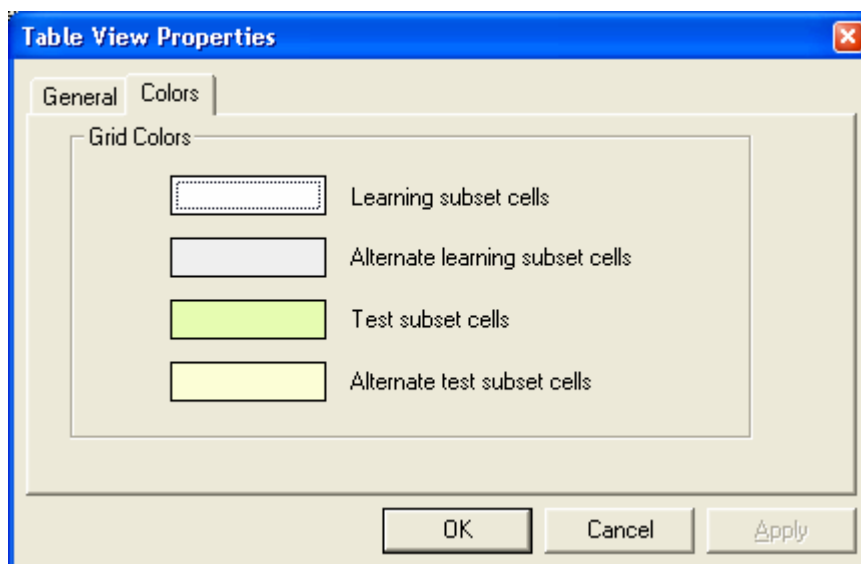
Initial cells value – initial cell value in the data sheet

View grid headers – grid headers style (flat or raised headers)

Decimal digits – number of decimal digits after decimal point

Apply to all projects – apply current settings for new projects

Default – reset all values to default state



Colors tab:

Learning subset cells – background color for cells that used for algorithm learning

Alternate learning subset cells – alternate background color for cells that used for algorithm learning

Test Learning subset cells – background color for cells that used only for testing

Alternate Test subset cells – alternate background color for cells that used only for testing

3.5.2 HTML View

HTML View displays formatting report with results of PNN algorithm calculations (See figure 3.5.2).

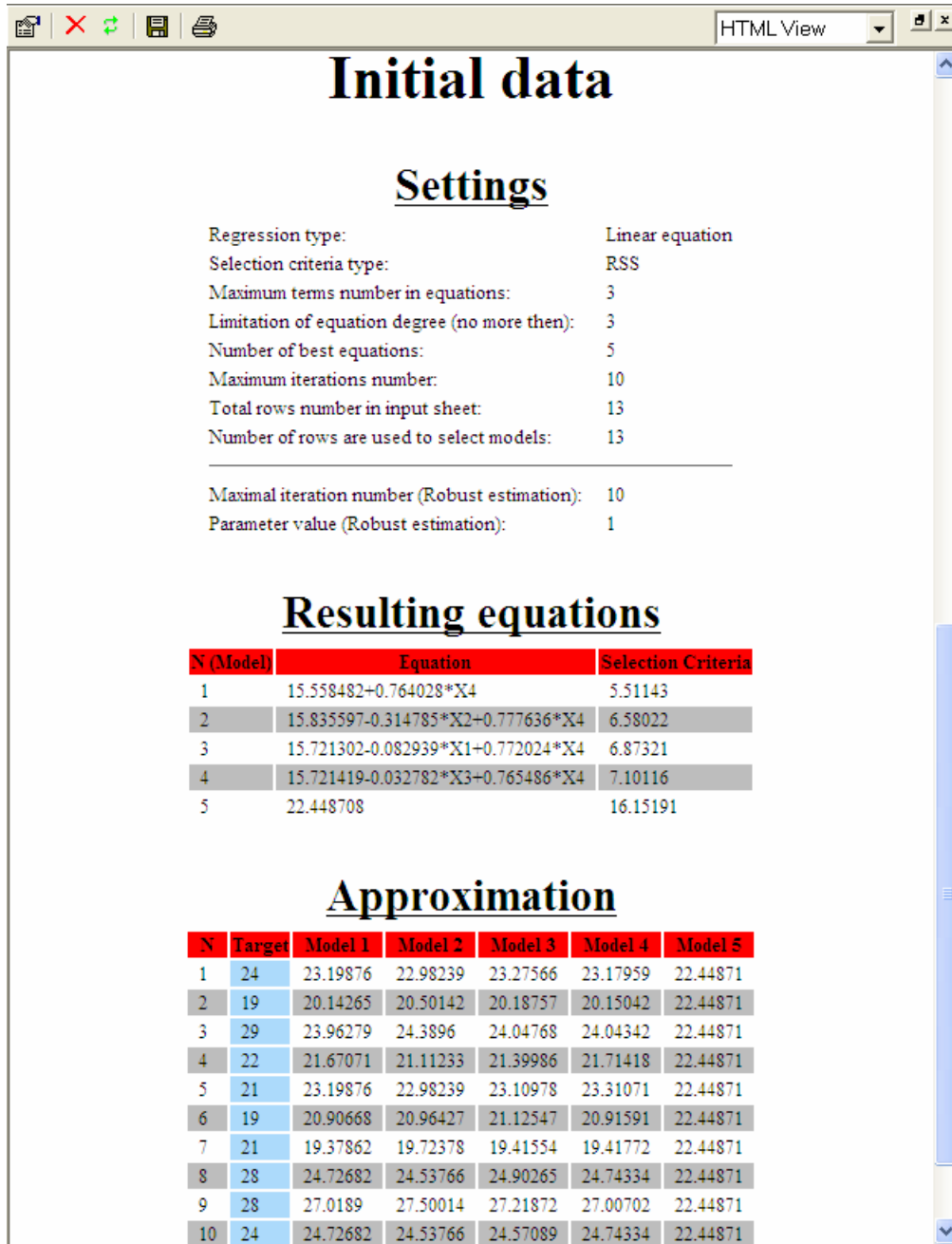


Figure 3.5.2

The report includes the following sections showing the result of best models searching:

- **Settings** – list of algorithm settings
- **Input data sheet** – data sheet records
- **Resulting equations** – list of resulting regressions and criteria values for regressions and test subset (*Main criteria* – criteria values for models; *Testing criteria* – criteria calculated based on test subset)
- **Approximation** – comparison with target values

- **Additional external criteria** – additional criteria to estimate quality of obtained regressions

Also, the report includes two additional sections represented the model’s applicability to the input dataset:





- **Resulting approximation** – input dataset substitution to the resulting equations
- **Input dataset** – input dataset records

Report in HTML View can be generated in 2 formats: as HTML and as plain text.

The  button in the HTML View control toolbar allows you to save report in .html file.

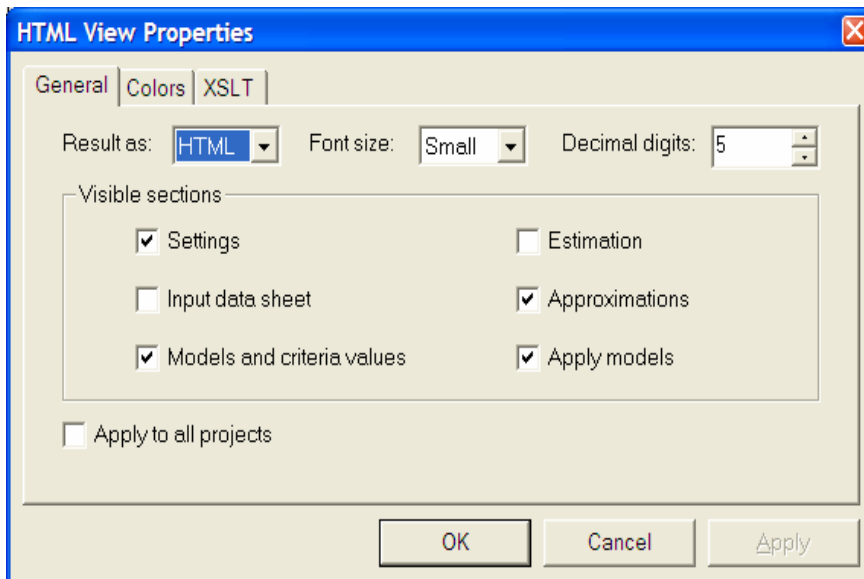
Click on  button in the HTML View control toolbar to print current report.

HTML View Toolbar

-  **Clear content.** Clear all resulting report in HTML View. To view the same resulting report again reload HTML View control in the control combo box in Results panel
-  **Refresh.** Refresh reporting layout
-  **Save As HTML file.** Save resulting report in the HTML file in the file system
-  **Print.** Print resulting report

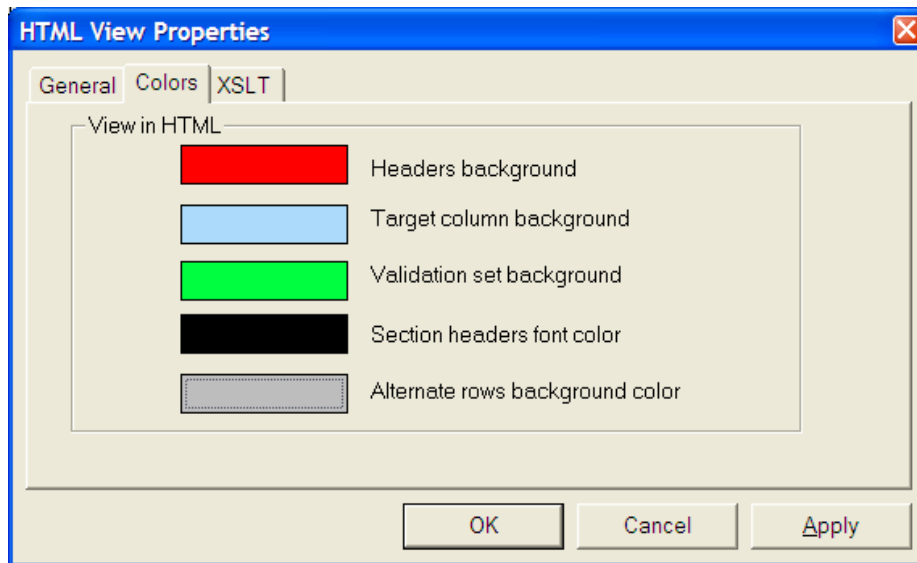
HTML View Properties

Click on  button in the control toolbar to invoke HTML View Properties dialog. Or you can invoke it from the HTML View context menu.



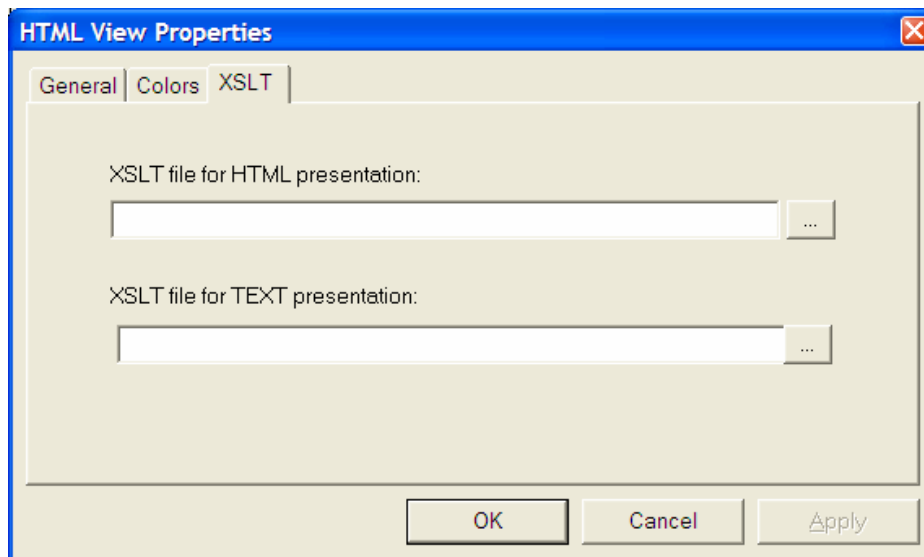
General tab:

- Result as** – switch report format between HTML and plain text
- Font size** – set font size: Large, Normal or Small
- Settings** – enable/disable Settings section in the report
- Input data sheet** – enable/disable Data sheet section in the report (disabled by default)
- Models and criteria values** - enable/disable Models section in the report
- Approximations** - enable/disable Approximation section in the report
- Estimation** - enable/disable Estimation section in the report
- Decimal digits** – number of decimal digits after decimal point
- Apply models** – display models applicability to the input dataset in the report
- Apply to all projects** – apply current settings for new projects



Colors tab:

- Table headers** – background color for table headers in the report
- Target column** – background color for target column values
- Validation set** – background color for validation rows
- Section headers** – color for section headers
- Background** – background color for regular cells



XSLT tab:

- XSLT file for HTML presentation** – XSLT file location to make custom HTML report presentation based on PNN Discovery Client XML standard.
- XSLT file for TEXT presentation** – XSLT file location to make custom TEXT report presentation based on PNN Discovery Client XML standard.

The XSLT file templates (pnn.xslt and pnn.txt.xslt) are included into the PNN Discovery Client installation.

3.5.3 Chart View

Chart View shows deviation of approximation values from target values in resulting models (See figure 3.5.3).

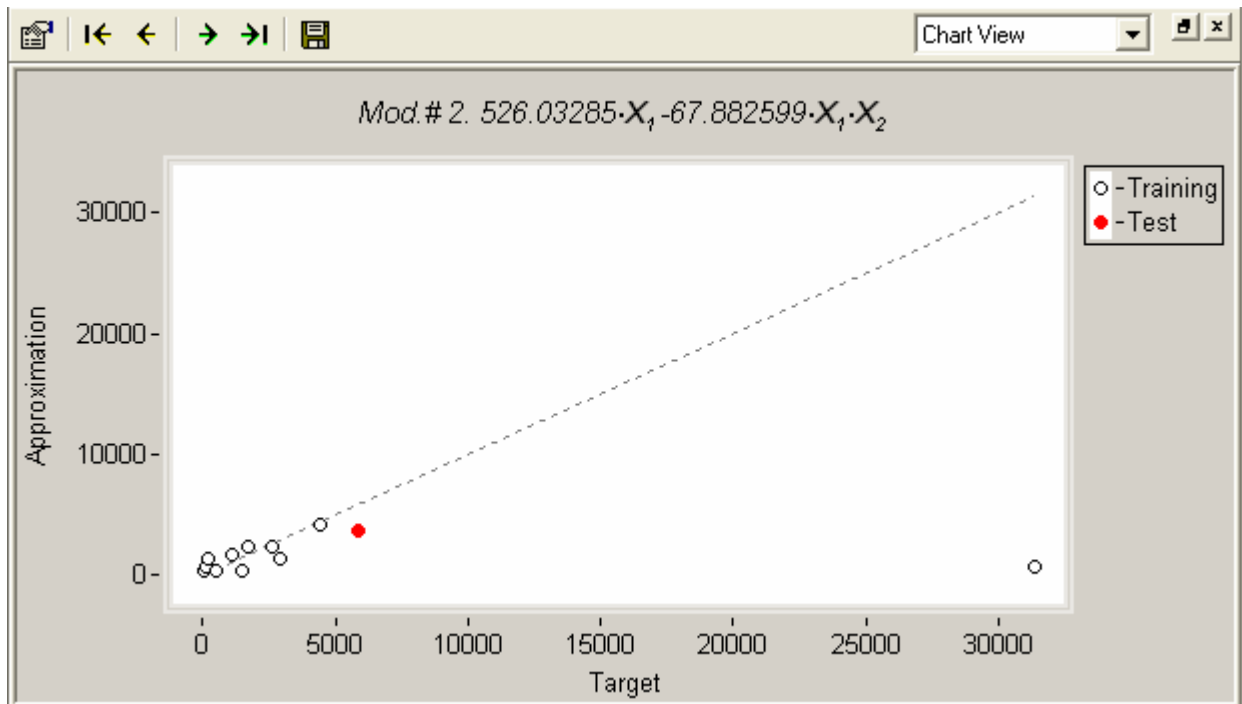


Figure 3.5.3

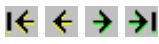
You can navigate throughout the models using  buttons on the control toolbar.

Chart View Toolbar






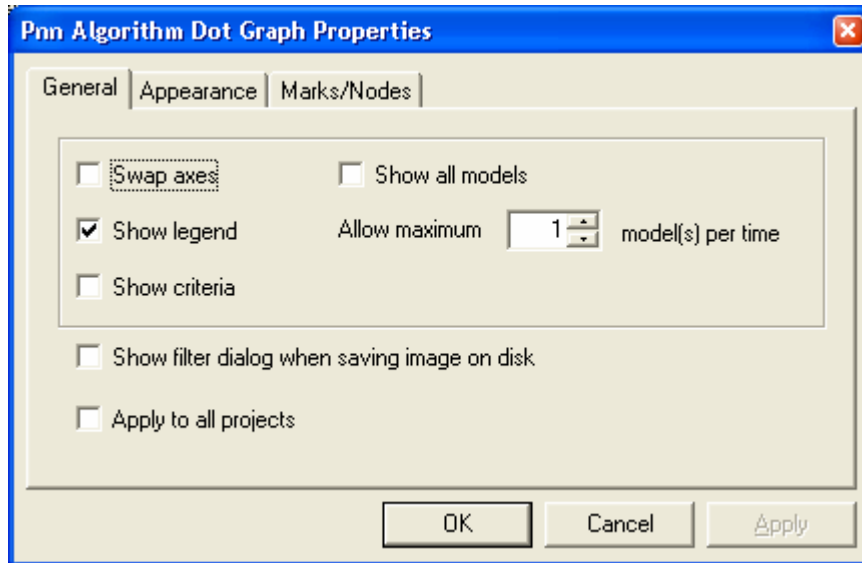
-  **First chart.** Show chart for first model
-  **Previous chart.** Show chart for previous model
-  **Next chart.** Show chart for next model
-  **Last chart.** Show chart for last model
-  **Save image.** Save chart as bitmap file

Chart View properties

Click on  button in the control toolbar to invoke Chart View Properties dialog. Or you can invoke it from Chart View context menu.



General tab:

Swap axes – swap “Approximation” and “Target” axes

Show legend – show legend for chart symbols

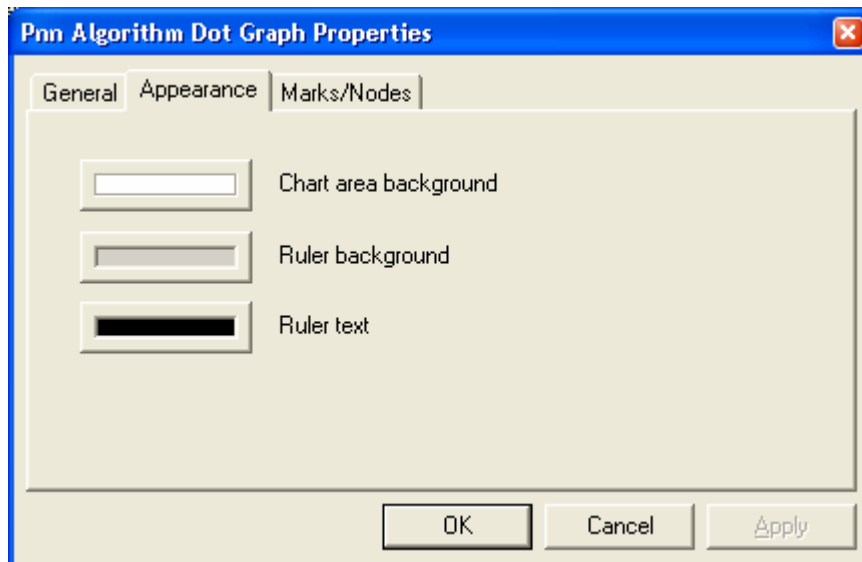
Show all models – show charts for all models

Allow maximum – maximum number of charts shown in the view simultaneously

Show criteria – show criteria value below the chart

Show filter dialog when saving image on disk – show dialog for custom save options

Apply to all projects – apply current settings for new projects

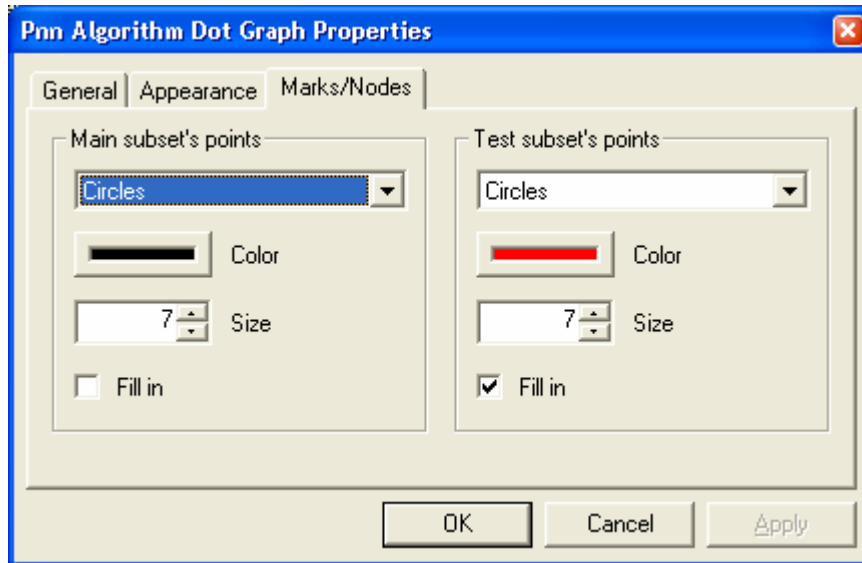


Appearance tab:

Chart area background – color of background inside chart area

Ruler background – color of background around chart

Ruler text – color for all labels on the chart




Marks/Nodes tab:

Main subset's points – type, color and size of marks for main subset points

Test subset's points - type, color and size of marks for test subset points

3.5.4 Plot View

Plot View allows comparison between approximation models and target shown as functions of their arguments (See figure 3.5.4).

You can navigate throughout the models using  buttons on the control toolbar.

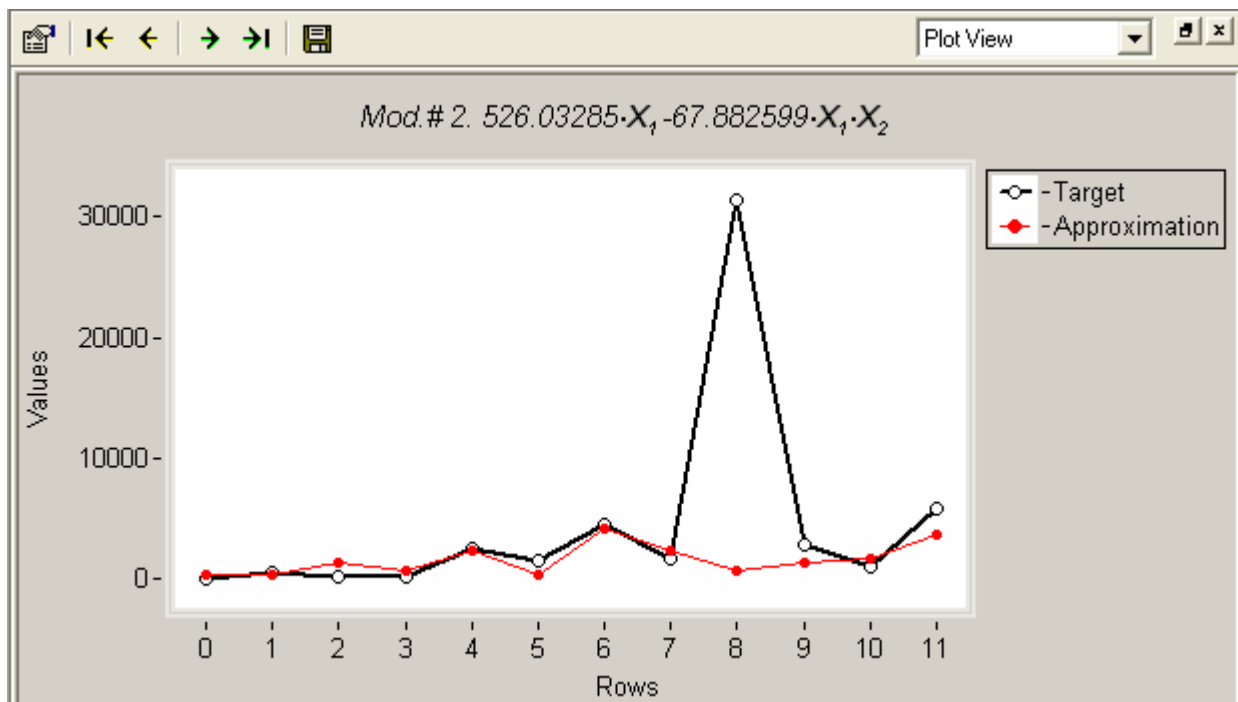







Figure 3.5.4

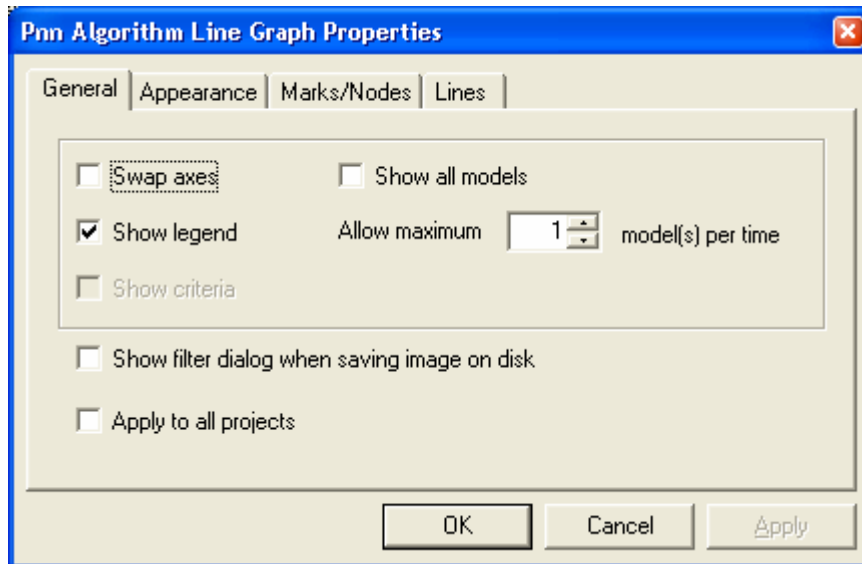
Plot View Toolbar

 **First chart.** Show chart for first model

-  **Previous chart.** Show chart for previous model
-  **Next chart.** Show chart for next model
-  **Last chart.** Show chart with last model
-  **Save image.** Save plot as bitmap file

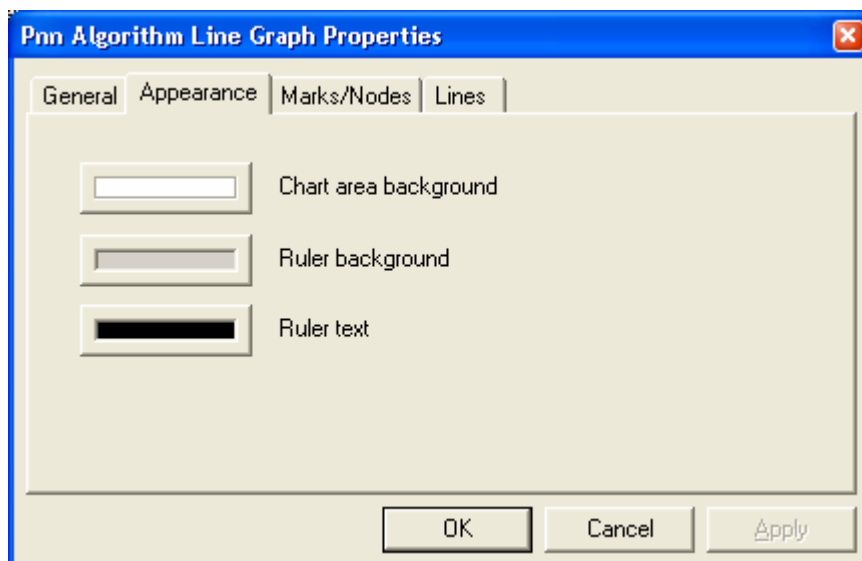
Plot View properties

Click on  button in the control toolbar to invoke Chart View Properties dialog. Or you can invoke it from Chart View context menu.



General tab:

- Swap axes** – swap “Approximation” and “Target” axes
- Show legend** – show legend for plot symbols
- Show all models** – show plots for all models on a single chart
- Allow maximum** – maximum number of plots shown on a single chart
- Show criteria** – show criteria value below the plot
- Show filter dialog when saving image on disk** – show dialog for custom save options
- Apply to all projects** – apply current settings for new projects

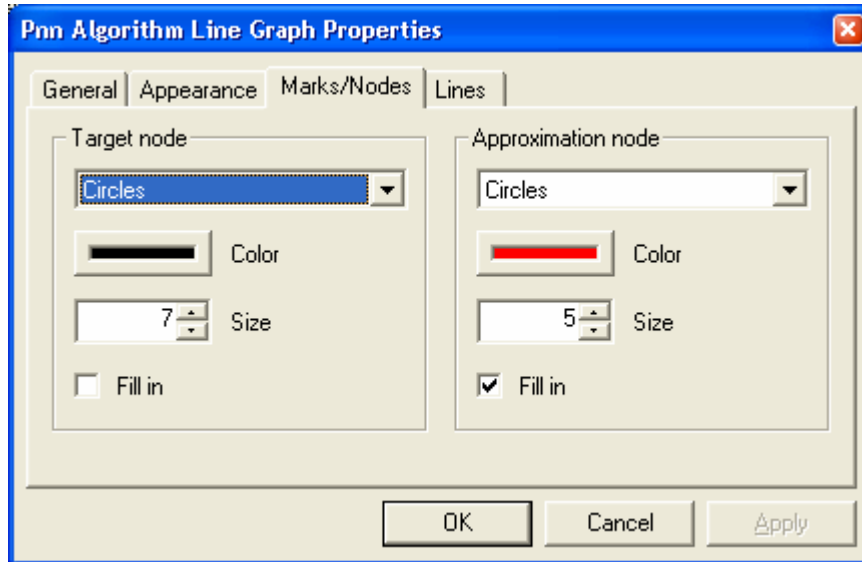


Appearance tab:

Chart area background – color of background inside plot area

Ruler background – color of background around plot

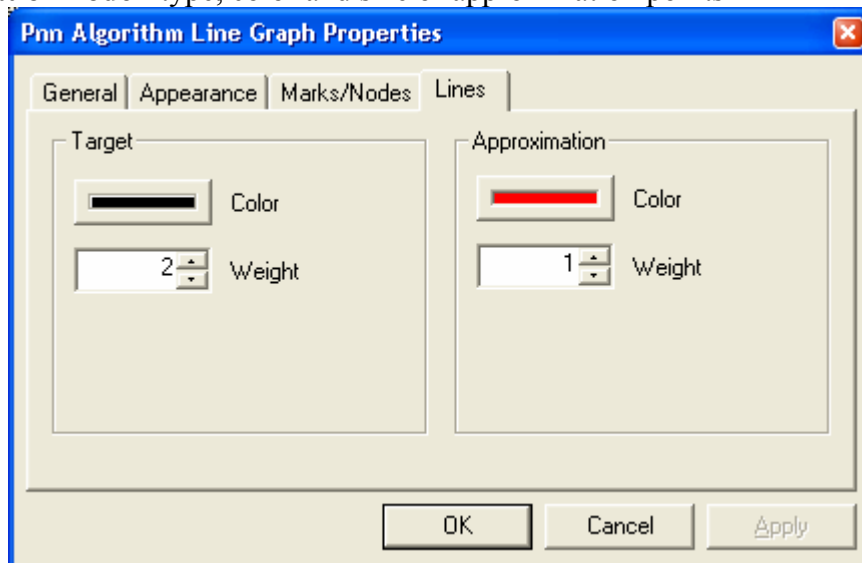
Ruler text – color for all labels on the plot



Marks/Nodes tab:

Target node – type, color and size of target points

Approximation node - type, color and size of approximation points



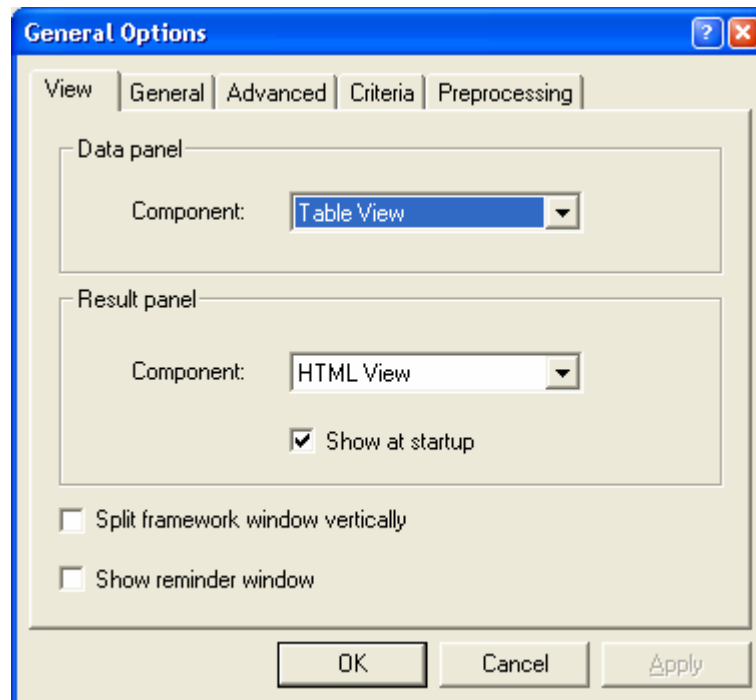
Lines tab:

Target – color and thickness of line that links target values

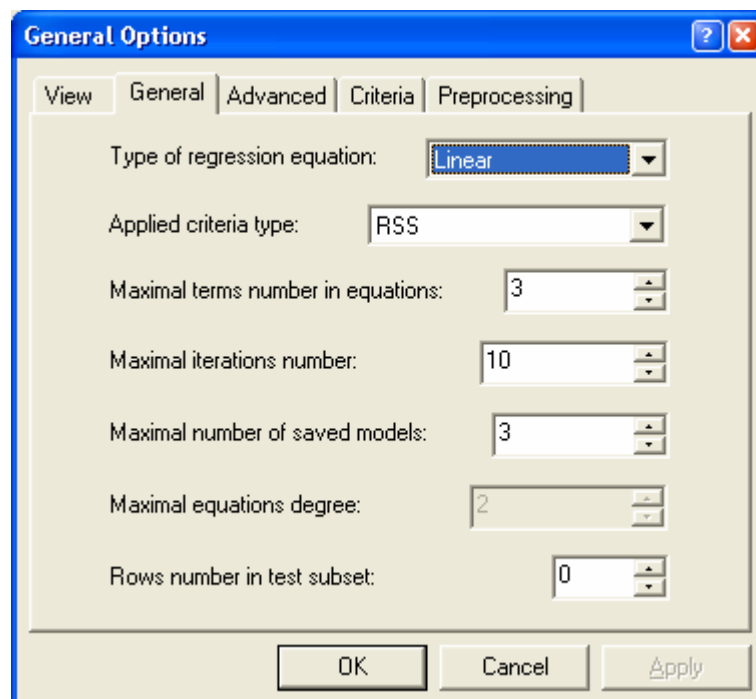
Approximation – color and thickness of line that links approximation values

3.5 Project options

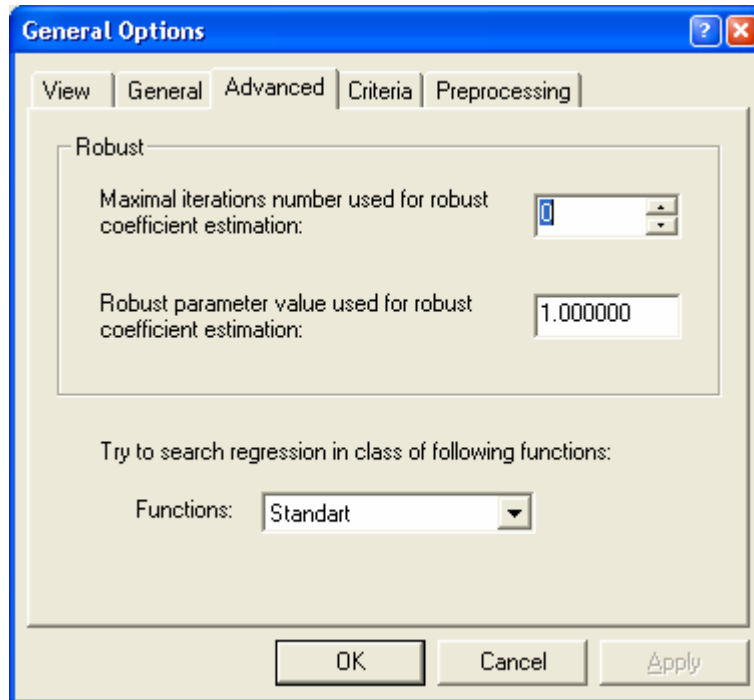
PNN Discovery Client framework menu command “**View**”->”**Options...**” allows you to customize general setting for newly created projects.



On **View** tab you can select Views that will be displayed in Data panel and Result panel by default. Also you can select the layout for main framework window. Check “Split framework window vertically” checkbox to split work area vertically. Check “Show reminder window” checkbox to show “Save Project” dialog every time you close the project.



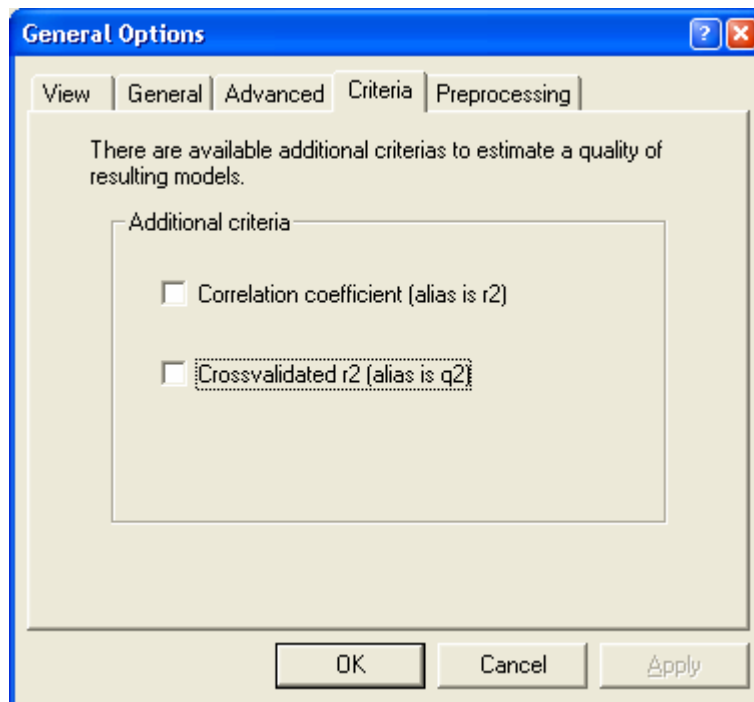
On **General** tab you can set default parameters for PNN algorithm. See details below in **Calculations** section.



On **Advanced** tab you can set default advanced parameters for PNN. See details below in **Calculations** section.

On **Criteria** tab you can set additional external criteria to estimate quality of resulting regressions. Two criteria are available:

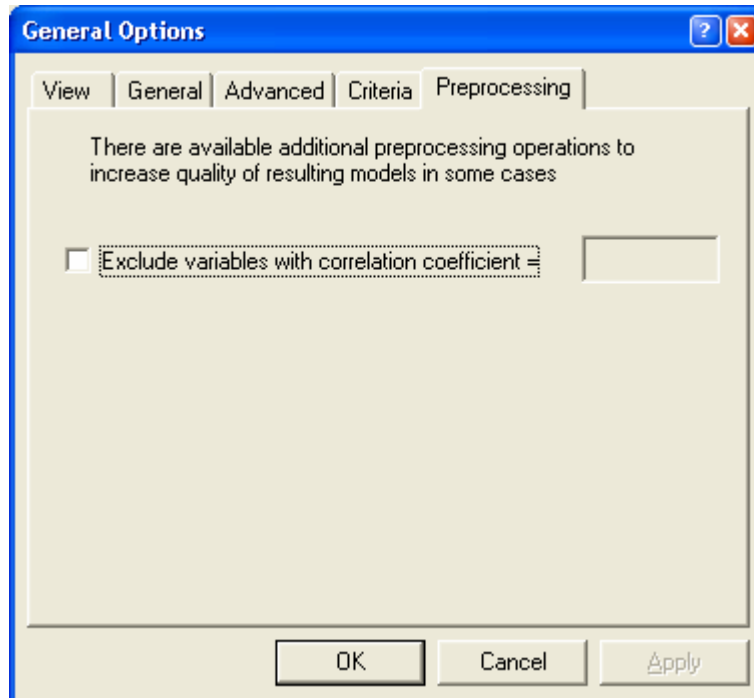
- Correlation coefficient (well-known as r2 criteria)




- Cross validated correlation coefficient (like SIBYL q2 estimation)

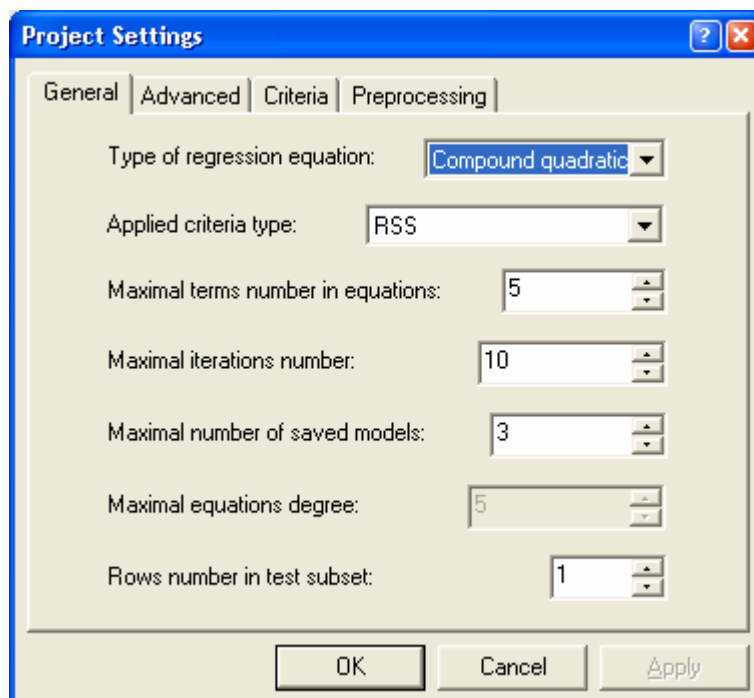
On **Preprocessing** tab you can specify additional preprocessing operations for initial data to increase quality of resulting regressions. The following preprocessing operations are available now:

- Exclude from searching process variables with specified value of correlation coefficient



3.6 Calculations

To achieve better results and higher accuracy of regression models you may consider adjusting parameters of PNN algorithm. You may access current parameters of PNN algorithm by clicking  button in the PNN Discovery Client framework toolbar (see figure below).



Type of regression equation – type of polynomials equations that will be targeted by PNN algorithm. There are 4 types of regression supported by PNN algorithm, which cover all possible types of polynomial equations.

Criteria type – type of criteria used for estimation and ranking of best models. Now, 2 types of criteria are supported: RSS (excluding values in validation rows) and RR (including only values in validation rows).

Maximal terms number in equations – number of terms in equation defines complexity of the regression models

Maximal iterations number – Number of iterations used in PNN algorithm on each complexity layer

Maximal number of saved models – number of models that will be rotated in the course PNN algorithm calculation

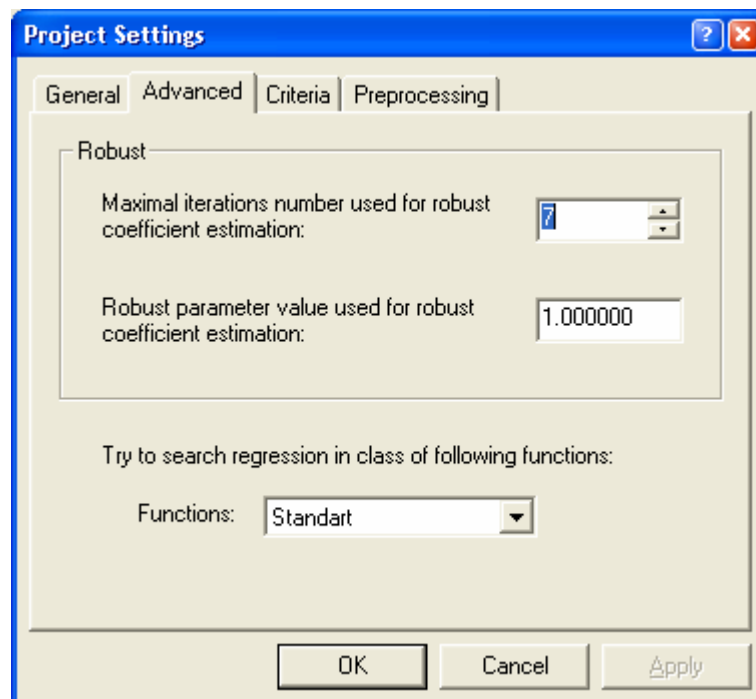
Maximal equations degree – degree of polynomial equation

Rows number in test subset – Number of rows excluded from the training set (see 1.4 for more details)

For robust models estimation you have to set parameters for robust calculations. Zero value for number of robust iterations disables robust models estimation.

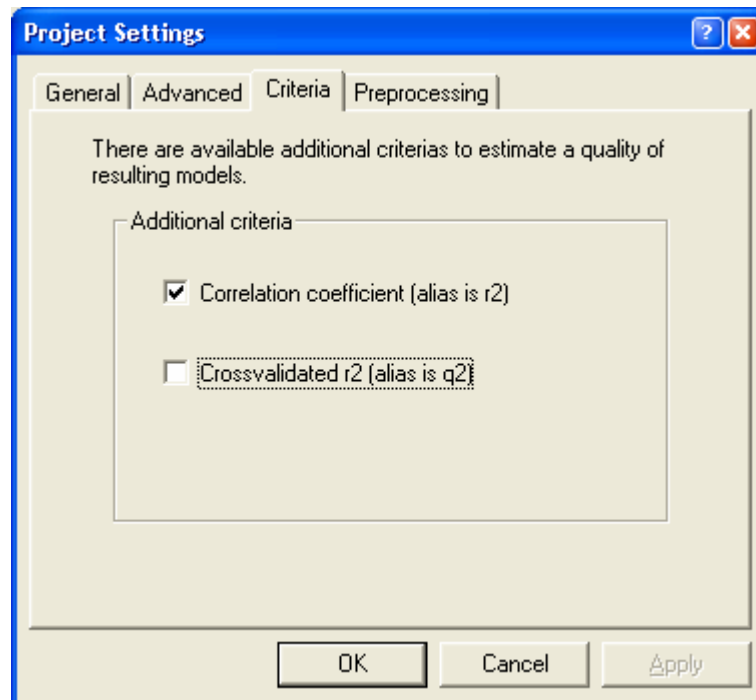
In addition, it is possible to look for regression models with specific term types:

- Quadratic fractional terms - terms like $x^{n/2}$, $n \geq 0, n < 0$ and values in spreadsheet must be positive
- Logarithm terms - terms like $\log^n x$, $n \geq 0$ and values in spreadsheet must be positive
- Fractional terms - terms like $1/x$ and values in spreadsheet must be nonzero.



Maximal iterations number used for robust coefficient estimation – internal parameter of M-estimation. Increase in number of iterations for robust estimation leads to increase in time of models searching.

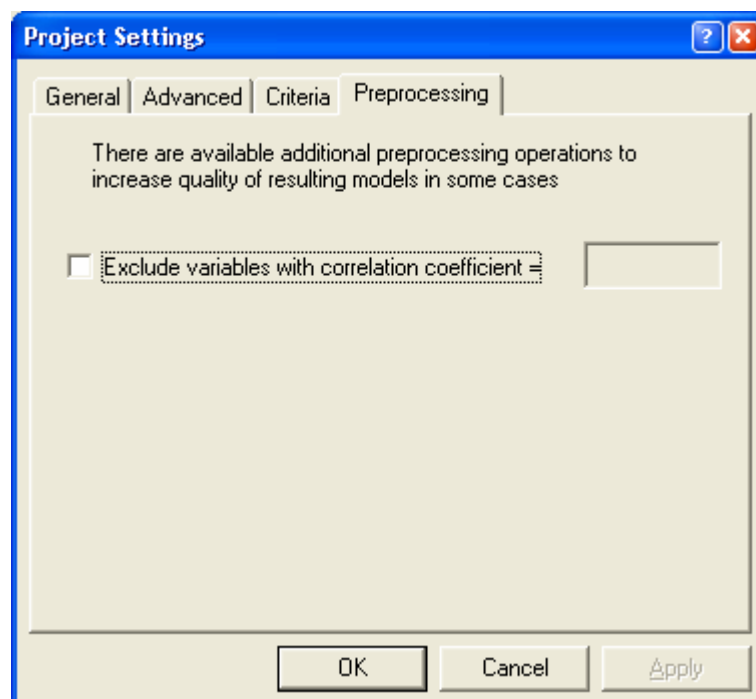
Fixed parameter value – can take values from the interval 1.0 to 1.5 (Default value is 1.0).



To include additional external criteria estimation for resulting regressions, set criteria parameters. Two additional criteria included in the PNN Discovery Client:

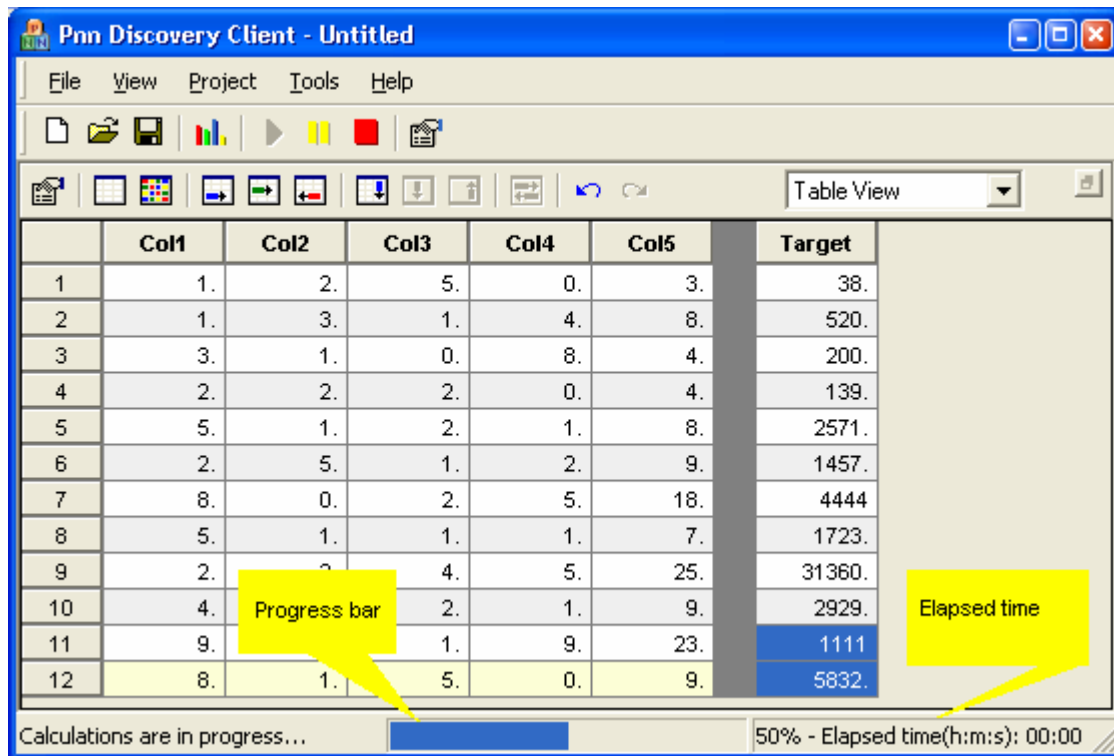
Correlation coefficient – well known as r2 criteria

Cross-validated correlation coefficient - like SIBYL q2 estimation






To include additional data preprocessing and increase quality of resulting regressions, set preprocessing parameters. Now, the only data preprocessing operation, included to the PNN Discovery Client is:

Exclude variables with correlation coefficient – variables with given value of correlation coefficient excluded from searching process.



The screenshot shows the 'Pnn Discovery Client - Untitled' window. It features a menu bar (File, View, Project, Tools, Help), a toolbar with various icons, and a 'Table View' dropdown. The main area contains a table with 12 rows and 7 columns. The columns are labeled 'Col1', 'Col2', 'Col3', 'Col4', 'Col5', and 'Target'. The 'Target' column is highlighted in blue. A progress bar at the bottom indicates 'Calculations are in progress...' at 50% completion, with an elapsed time of 00:00. Two yellow callout boxes point to the 'Progress bar' and 'Elapsed time' labels.

	Col1	Col2	Col3	Col4	Col5	Target
1	1.	2.	5.	0.	3.	38.
2	1.	3.	1.	4.	8.	520.
3	3.	1.	0.	8.	4.	200.
4	2.	2.	2.	0.	4.	139.
5	5.	1.	2.	1.	8.	2571.
6	2.	5.	1.	2.	9.	1457.
7	8.	0.	2.	5.	18.	4444.
8	5.	1.	1.	1.	7.	1723.
9	2.	2.	4.	5.	25.	31360.
10	4.	2.	2.	1.	9.	2929.
11	9.	1.	1.	9.	23.	1111.
12	8.	1.	5.	0.	9.	5832.

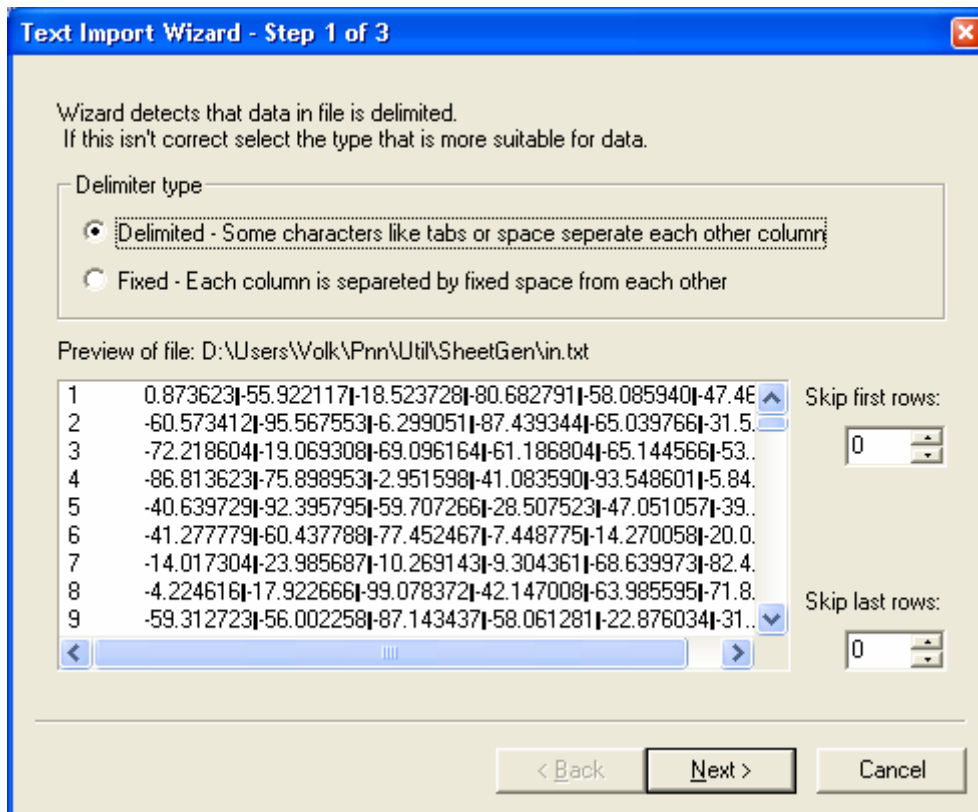
Click on  button in the PNN Discovery Client framework toolbar to start searching the best models with PNN algorithm. Click on  button to break searching best models any time processing takes place. Click on  button to suspend execution temporarily, so you can resume it later by clicking again. If you change data in the spreadsheet or parameters of PNN algorithm while calculation is going on, changes will be effective next time you start calculations. Calculation progress is shown in status bar of PNN Discovery Client framework (see figure above).

3.7 Import and export

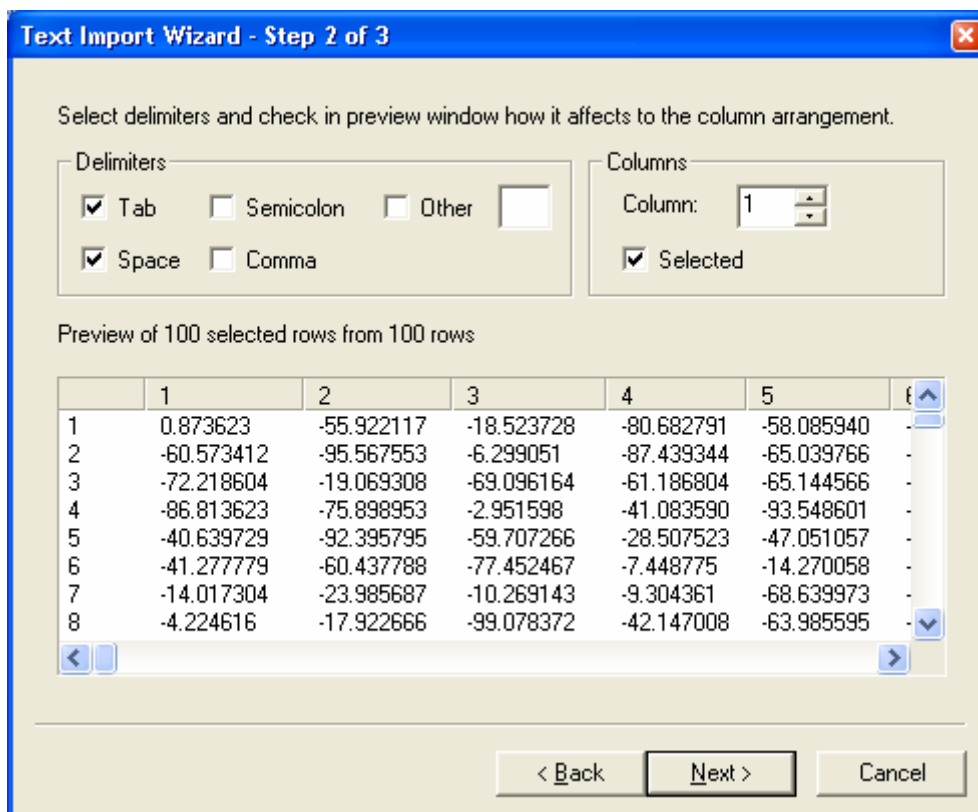
The PNN Discovery Client supports import of data presented in the following formats: plain text (.txt), XML (.xml), Microsoft Excel workbooks (.xls).

3.7.1 Text Import Wizard

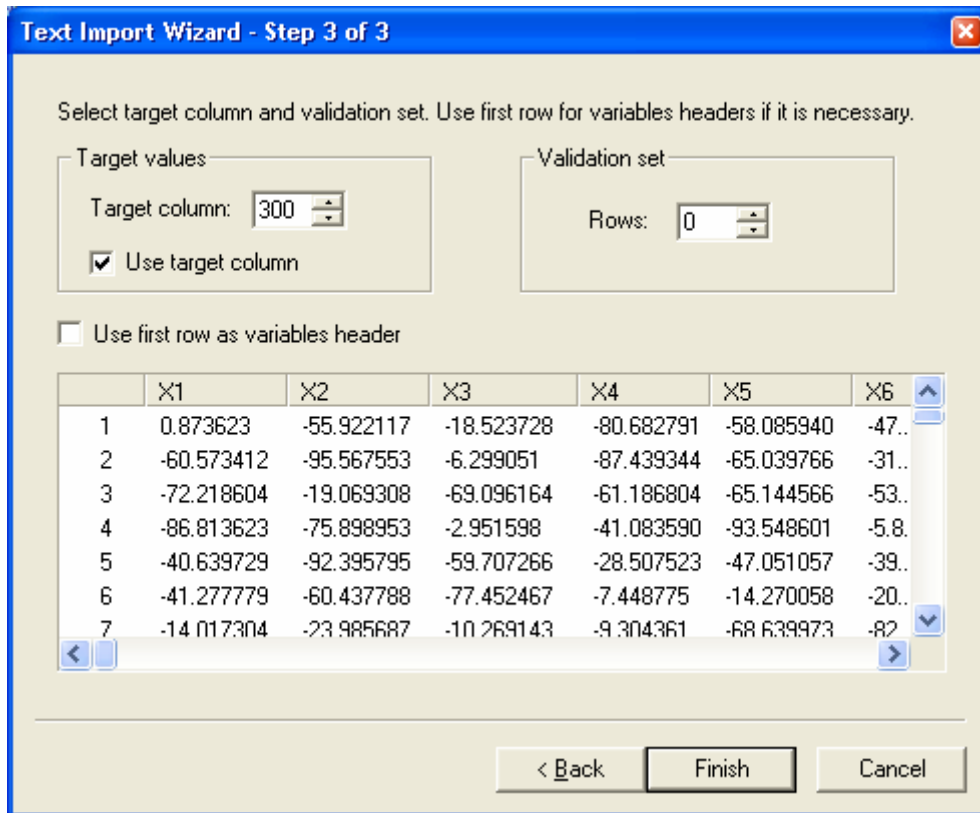
On the first page you can select data format: delimited values or fixed size columns.



Also, you can specify the range of rows you are going to import.
 For **delimited** data on the second page you can specify symbol(s) treated as delimiter and exclude columns you are not going to import.



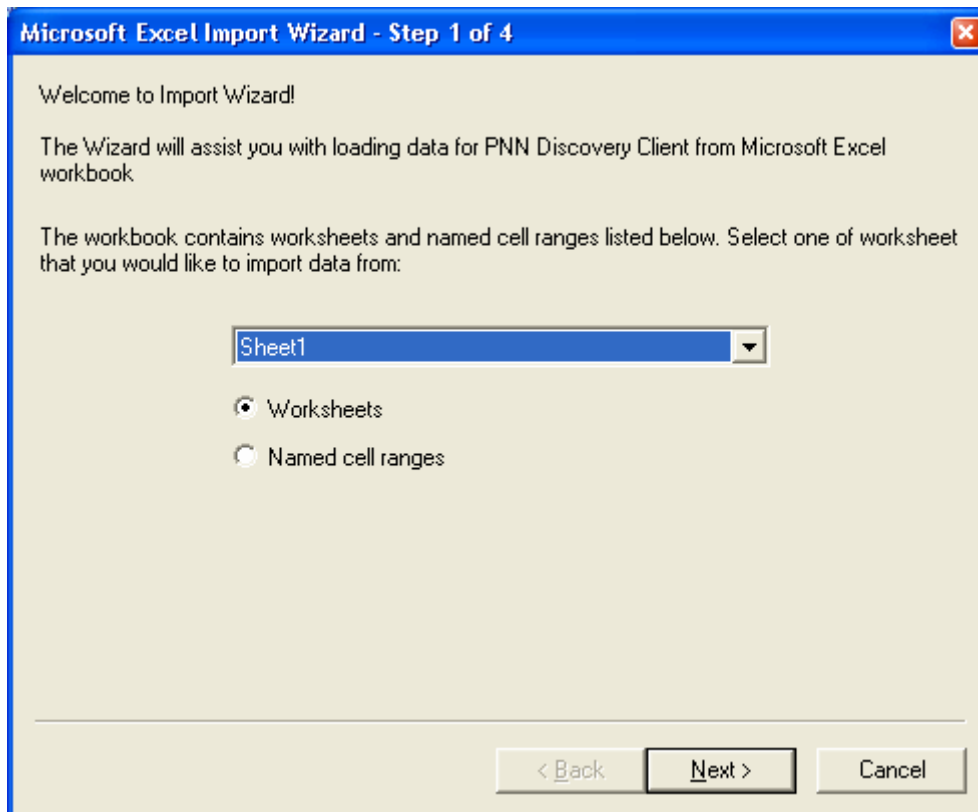
For **fixed** data on the second page you can specify column width for each column and exclude columns you are not going to import. On each step, you can monitor and verify correctness of export settings. And on the last page you can select “**target**” column and rows that will be treated as “**validation**” set.



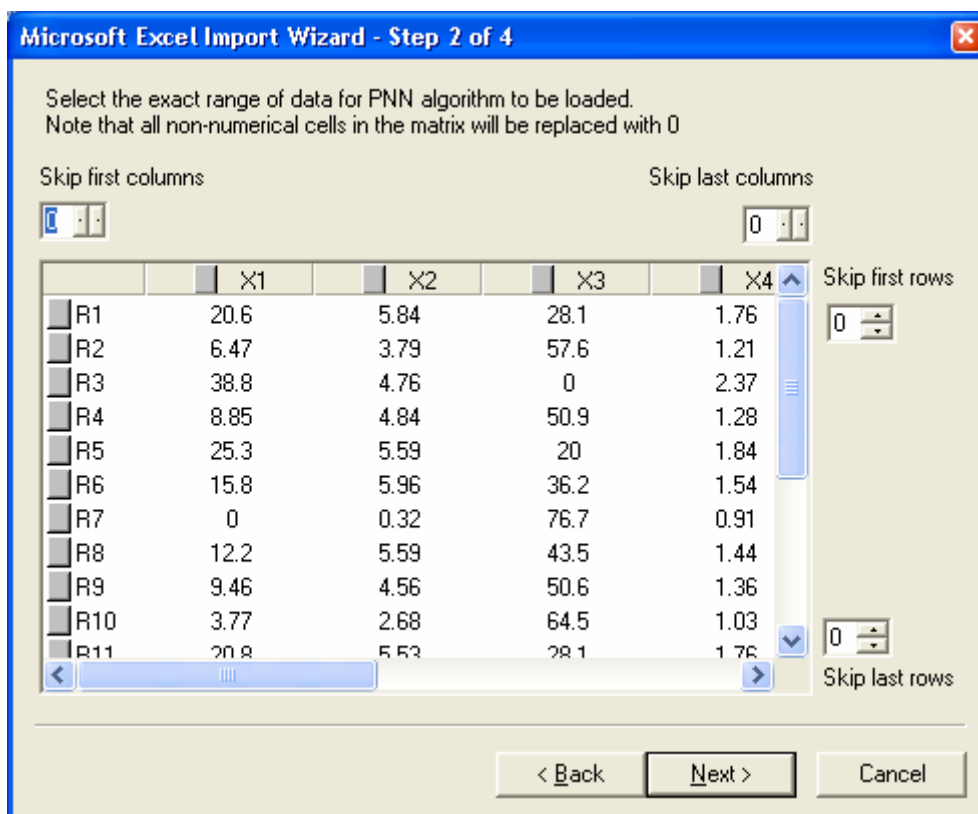
You can select to treat first row as columns (variables) names. Click “**Finish**” button to start import of data. Imported data is displayed in the current Data panel View. For huge data sheets, import may take some time.

3.7.2 Microsoft Excel Import Wizard

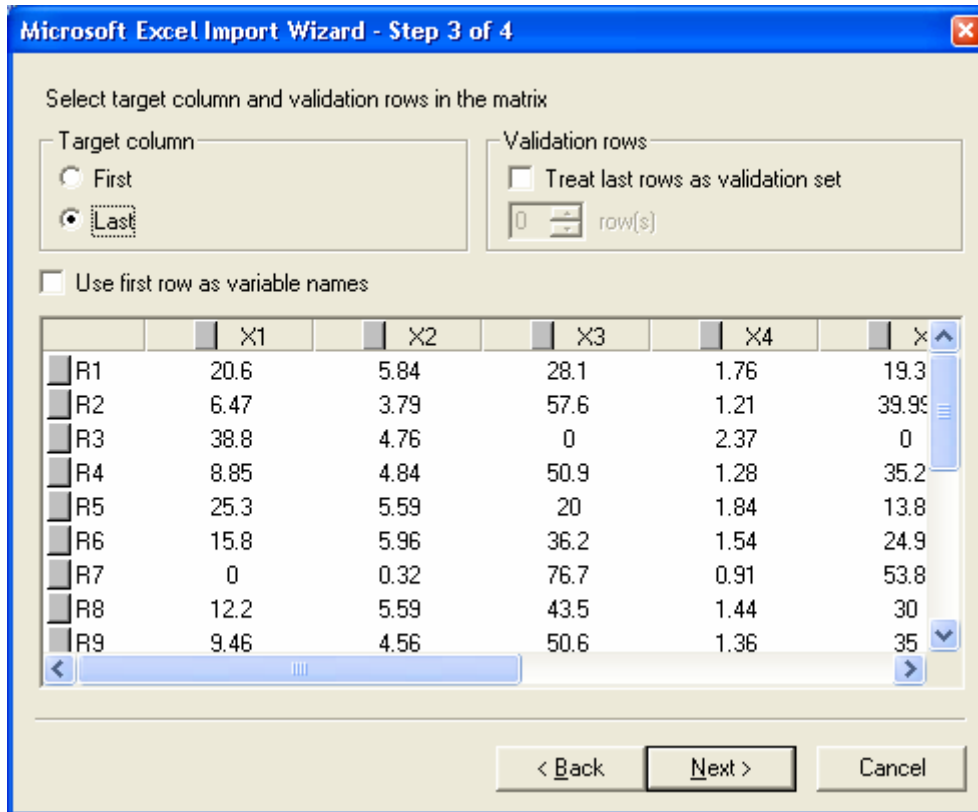
Microsoft Excel Import wizard allows importing data from Microsoft Excel workbook. On the first page you can select worksheet or named cell range from which data is to be imported.



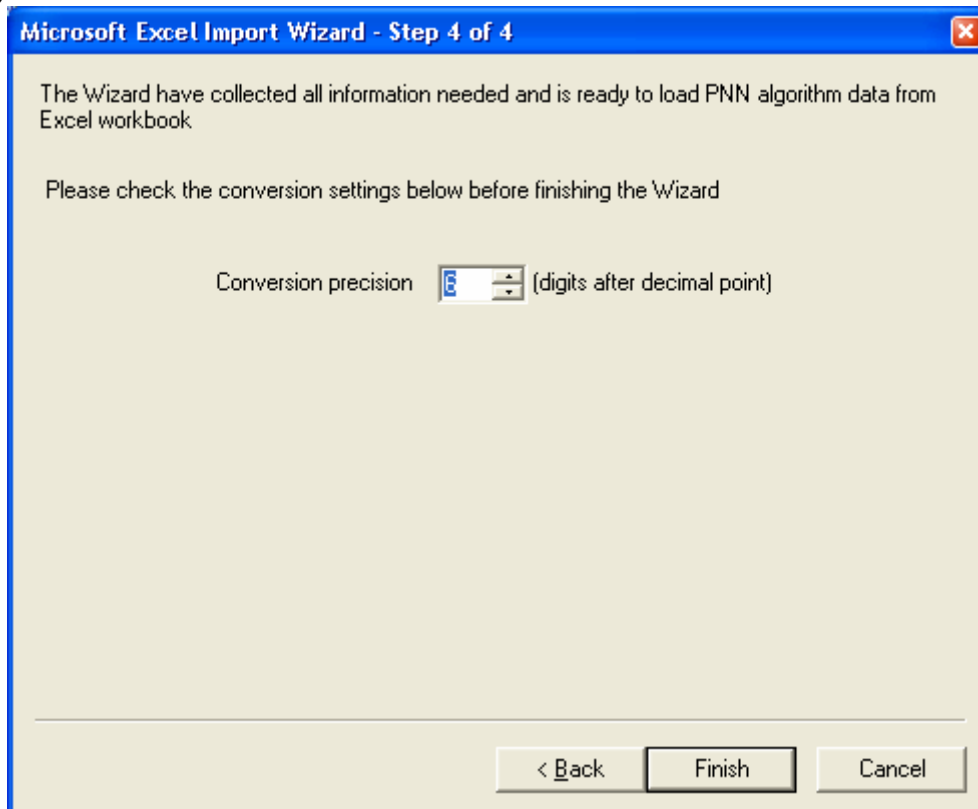
On the second page you can specify columns and rows you wish to skip during import operation.



On the third page you can specify “**target**” column and rows that will be treated as “**validation**” set. Also, by selecting first row as columns headers, you can define names for variables.



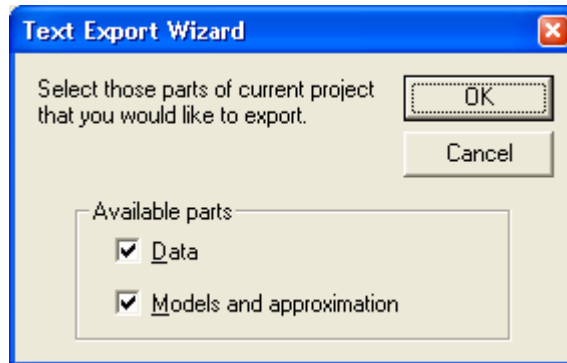
And on the last page you can specify conversion precision for real values (number of decimal digits after decimal point).



Click “**Finish**” button to start import of data. Imported data is displayed in the current Data panel View. For huge data sheets, import may take some time.

3.7.3 Export Wizard

Select menu command “Tools”->“Export To” in the PNN Discovery Client framework menu to export input data sheet and best models found to following formats: tab delimited plain text (*.txt), comma delimited plain text (*.csv), XML (.xml) and Microsoft Excel workbook (.xls).



Specify information, you wish to export - input data sheet and/or models found. Export Wizard saves it in the selected format.

Note. Number of columns in Microsoft Excel spreadsheet is limited to 256 columns. Therefore, we recommend using our Export Wizard to export large data sheets to Microsoft Excel format. Our Export Wizard guarantees correctness of data layout on the Microsoft Excel spreadsheet.