

MODEL

With AAA Model, build a machine learning analysis that allows to import, prepare & cleanse and consolidate data for better deployment and deliver an actionable business insights



FEATURES

- 1. Descriptive, Predictive, Anomaly, and Time series algorithms.



BENEFITS

- 1. Analyze data using multiple algorithms through ensemble.
- 2. Quickly and easily create models.
- 3. Discover structure and easily predict categories.
- 4. Capability to empower even a novice user.
- 5. Various modeling techniques are selected and applied and their parameters are calibrated to optimal values

PREREQUISITE

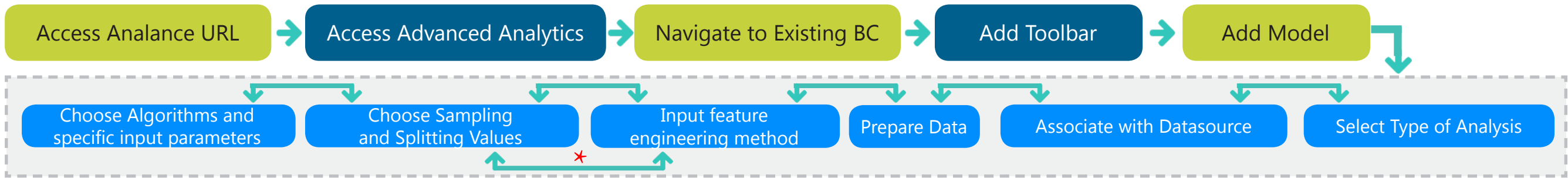
- 1. Data should be available in any one of different data connector format supported in Analance.

RELATIONAL DATABASES	BIG DATA	API	FILE	SOCIAL MEDIA

MODEL - WORK FLOW

WORK FLOW TO ADD MODEL IS SHOWN BELOW:

MENU PATH: >> Add > Click Model icon



- 1. Log into Analance using Analance URL
- 2. Select Advanced Analytics
- 3. Navigate to existing BC.
- 4. Select Add Toolbar- Navigate to the Model path to create Model. Refer to the Path section for more details.
- 5. Add a Model

★ There are two models: Supervised and Unsupervised. Both the steps, feature engineering method and sampling and splitting values step is not required for unsupervised models.

PROCEDURE TO ADD PREDICTION MODEL

Create Model

Name *

Description

UPLOAD YOUR IMAGE

Drag and drop files here or click to upload

Create

1

Create Model :
Associate a model with a name, description and if required identify it with an image to view as a thumb nail icon.

Choose Analysis :
Select Prediction analysis from the list of options.

2

New Model

Choose analysis

Select data

Prepare data

Feature engineering

Sample and split

Choose algorithm(s)

Please choose a type of analysis that will better suit model requirement

Prediction

Classification

Clustering

Survival analysis

Anomaly dimensionality reduction

Forecasting

Text classification

Text clustering

Association rules

Cancel

Next

Choose existing datasource

Available datasources

Search...

Test DS

DS sample

EmpData

asd

asd

Test DS-copy

dfghj

ttt

new

Data

sd

TestDB

qw

CustomerData

sdf

sdd

xfghj

ccx

Preview - DS sample

MachineOS	DecommissionDate	LastModifiedDate	LoadDate	MachineDescription	Machine
Linux2.6	2010-06-25T00:00:00	2009-07-10T00:00:00	2009-07-10T00:00:00	Kit contains one convenie...	15" LCD
Linux2.6	2011-06-17T00:00:00	2009-07-06T00:00:00	2009-07-07T00:00:00	Kit contains one convenie...	15" LCD
Linux2.6	2015-01-01T00:00:00	2009-07-06T00:00:00	2009-07-06T00:00:00	Kit contains one convenie...	15" LCD
Linux2.6	2014-06-16T00:00:00	2009-07-06T00:00:00	2009-07-06T00:00:00	Kit contains one convenie...	15" LCD
Linux2.6	2010-06-16T00:00:00	2009-07-06T00:00:00	2009-07-06T00:00:00	Kit contains one convenie...	15" LCD
Linux2.6	2011-06-17T00:00:00	2009-07-06T00:00:00	2009-07-07T00:00:00	Kit contains one convenie...	15" LCD
Linux2.6	2015-01-01T00:00:00	2009-07-06T00:00:00	2009-07-06T00:00:00	Kit contains one convenie...	15" LCD
Linux2.6	2014-06-16T00:00:00	2009-07-10T00:00:00	2009-07-10T00:00:00	Kit contains one convenie...	15" LCD
Linux2.6	2010-06-16T00:00:00	2009-07-10T00:00:00	2009-07-10T00:00:00	Kit contains one convenie...	15" LCD
Linux2.6	2011-06-17T00:00:00	2009-07-10T00:00:00	2009-07-10T00:00:00	Kit contains one convenie...	15" LCD
Linux2.6	2015-01-01T00:00:00	2009-07-10T00:00:00	2009-07-10T00:00:00	Kit contains one convenie...	15" LCD
Linux2.6	2014-06-16T00:00:00	2009-07-10T00:00:00	2009-07-10T00:00:00	Kit contains one convenie...	15" LCD
Linux2.6	2010-06-16T00:00:00	2009-07-10T00:00:00	2009-07-10T00:00:00	Kit contains one convenie...	15" LCD
Linux2.6	2011-06-17T00:00:00	2009-07-10T00:00:00	2009-07-10T00:00:00	Kit contains one convenie...	15" LCD

In memory

Add

3

Choose existing Datasource :
Select the database to add data and the preview will be shown. Option to add a new datasource is also available.

New Model

Choose analysis — Select data — **Prepare data** — Feature engineering — Sample and split — Choose algorithm(s)

Please prepare data with required column for model preparation

Skip cleansing ☐ Skip normalization ☐

Column name	Data type	Is factor	Unique id	Target	Predictor	Imputation	Lower outlier	Upper outlier	Normalization
<input type="checkbox"/> MachineOS	nvarchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> DecommissionDate	datetime	<input type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>		Select	Select	Select
<input type="checkbox"/> LastModifiedDate	datetime	<input type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>		Select	Select	Select
<input type="checkbox"/> LoadDate	datetime	<input type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>		Select	Select	Select
<input type="checkbox"/> MachineDescription	nvarchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> MachineHardware	nvarchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> UpdateDate	datetime	<input type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>		Select	Select	Select
<input type="checkbox"/> MachineLabel	nvarchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> MachineName	nvarchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> Status	nvarchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> MachineSoftware	nvarchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> MachineSource	nvarchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> MachineType	number	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select

Total columns: 18, Selected columns: 0

4

Clear all

Cancel Next

Prepare data : Prepare the data with the required columns.

Is factor: Select if the variable is to be treated as categorical.

Unique id: Select the unique variable if the datasource has a unique value.

Target: This variable value is mandatory. Select a variable that you will like to predict that is dependent for the analysis.

Predictor: This column value is mandatory and independent. Select columns to influence the target variable.

Imputation: Select substitute value in case there are missing values for each variable. The options will vary based on the Data Type. If it is a continuous / numeric value then the options are: Mean, Median, Mode and User Input (Number). If the variable is factor / categorical then the Options: Most frequent, Least frequent, User Input (Text).

Lower outlier: Replace the lower outlier values with any one of the options: Percentile 2nd, Percentile 5th, User input.

Upper outlier: Replace the upper outlier values with any one of the options: Percentile 95th, Percentile 98th, User input.

Normalization: Select from the options Z-score, Min Max, Logistic, Log normal, Tanh to normalize data.

Skip Cleansing and Skip Normalization: if available data is cleansed and normalized or user wants to view raw data to be analyzed.

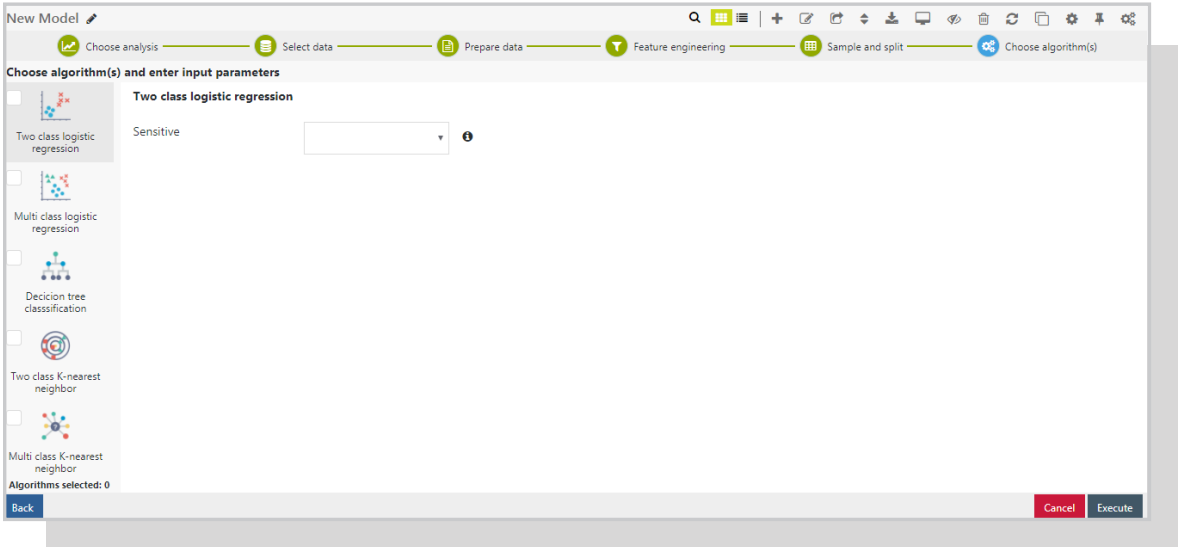
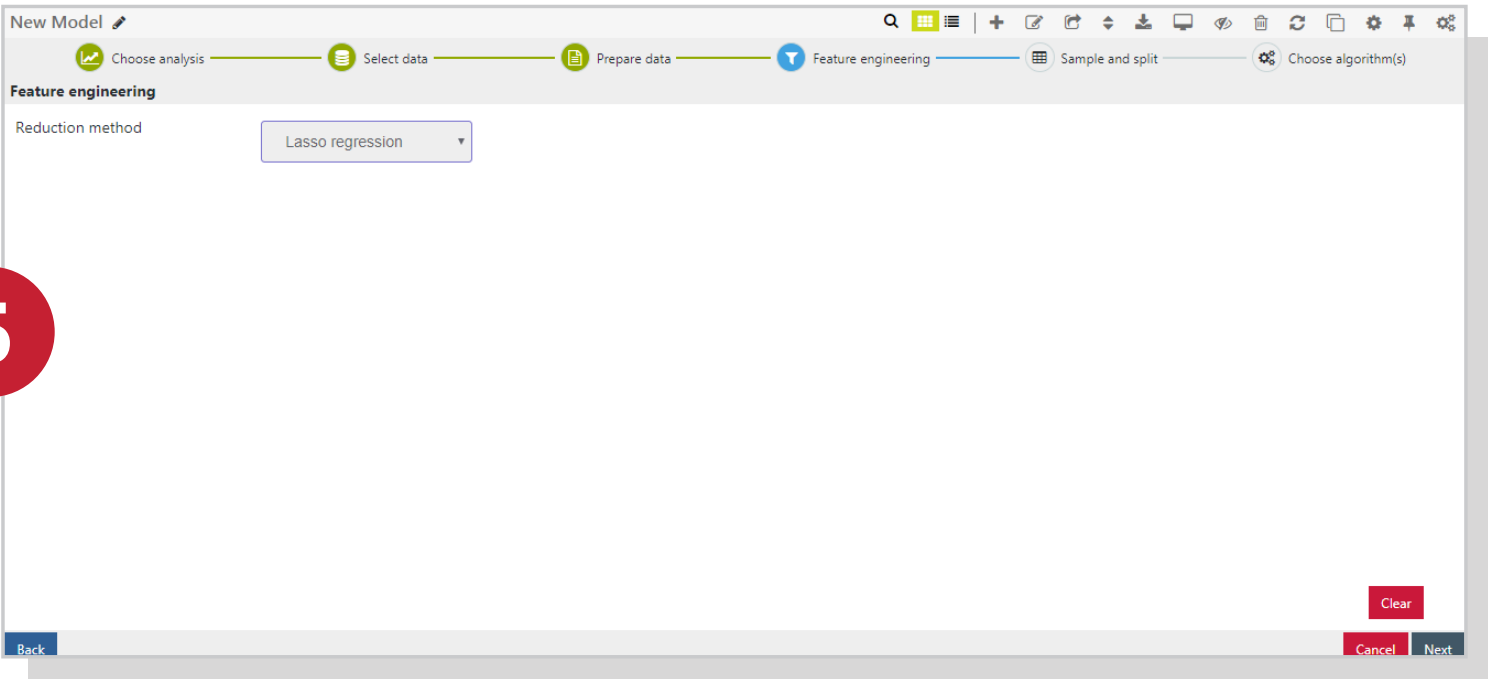
Feature engineering :
Select the reduction method. Tree based, Principle component analysis, Lasso Regression.

Tree based :
Specify Number of Trees and Depth of Trees.
Principle component: Select Matrix generation based on: Correlation / Covariance.

Proportion of variance to :
Use the slider to increase or reduce the value as required.

Normalization :
Select Z-score to standardize the values.

5



6

Choose algorithms :
Choose the type of algorithm(s) from the available algorithms for calculating the values for the model and specify the input parameters.

ALGORITHMS	INPUT PATRAMETERS
Multiple linear regression	No input parameter
Decision Tree regression	Minimum split, Minimum bucket, Maximum Depth
Bayesian linear regression	No input parameter
Neural network regression	Sensitive value
Multi class k-nearest neighbor	Hidden layer size, Maximum iterations

PROCEDURE TO ADD CLASSIFICATION MODEL

Create Model :

Associate a model with a name, description and if required identify it with an image to view as a thumb nail icon.

1

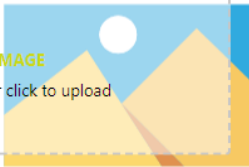
Create Model

Name *

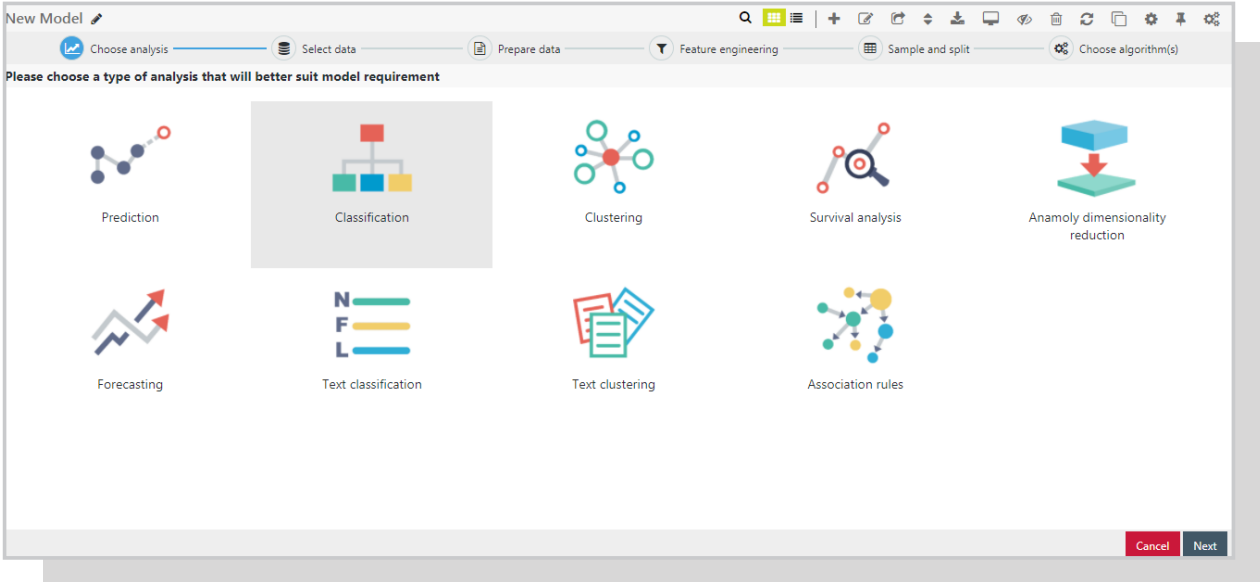
Description

UPLOAD YOUR IMAGE

Drag and drop files here or click to upload



Create



2

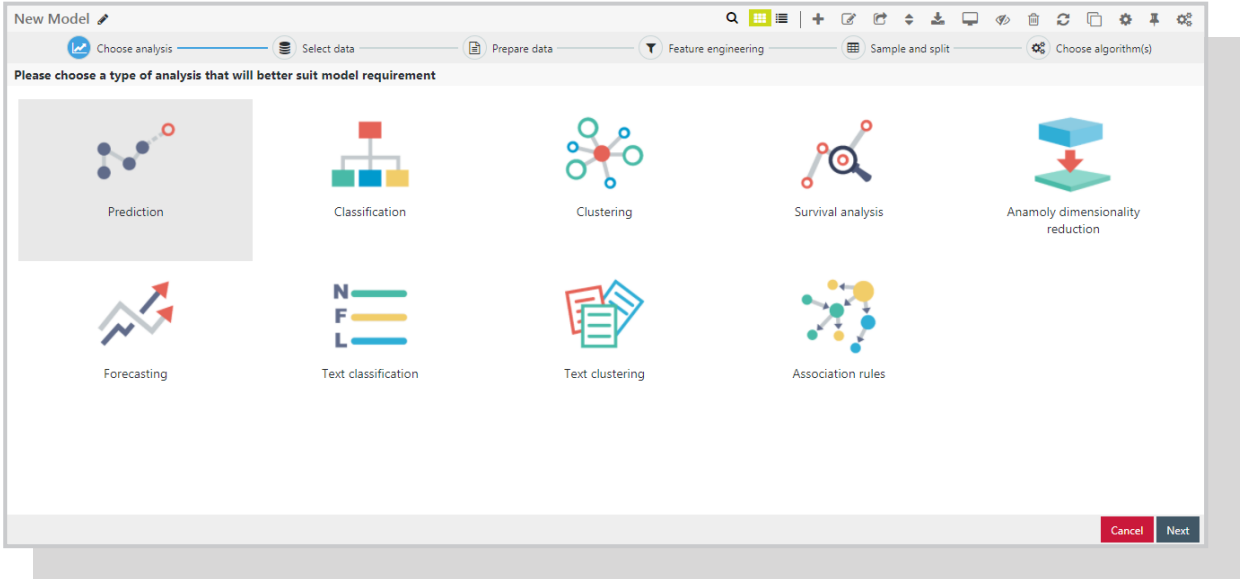
Choose Analysis :

Select Classification analysis from the list of options.

Choose existing datasource :

Select the database to add data and the preview will be shown. Option to add a new datasource is also available.

3



New Model

Choose analysis | Select data | Prepare data | Feature engineering | Sample and split | Choose algorithm(s)

Please prepare data with required column for model preparation

Column name	Data type	Is factor	Unique id	Target	Predictor	Imputation	Lower outlier	Upper outlier	Normalization
<input type="checkbox"/> MachineOS	nvarchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> DecommissionDate	datetime	<input type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>		Select	Select	Select
<input type="checkbox"/> LastModifiedDate	datetime	<input type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>		Select	Select	Select
<input type="checkbox"/> LoadDate	datetime	<input type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>		Select	Select	Select
<input type="checkbox"/> MachineDescription	nvarchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> MachineHardware	nvarchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> UpdateDate	datetime	<input type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>		Select	Select	Select
<input type="checkbox"/> MachineLabel	nvarchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> MachineName	nvarchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> Status	nvarchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> MachineSoftware	nvarchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> MachineSource	nvarchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> MachineType	nvarchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select

Total columns: 18, Selected columns: 0

Clear all

Back Cancel Next

4

- Prepare Data :** Prepare the data with the required columns.
- Is factor:** Select if the variable is to be treated as categorical.
- Unique id:** Select the unique variable if the datasource has an unique value.
- Target:** This variable value is mandatory. Select a variable that you will like to predict that is dependent for the analysis.
- Predictor:** This column value is mandatory and independent. Select columns to influence the target variable.
- Imputation:** Select substitute value in case there are missing values for each variable. Options: Most frequent, Least frequent, User Input.
- Lower outlier:** Select the lower value that is less than the variable value, the options are: Percentile 2nd, Percentile 5th, User input.
- Upper outlier:** Select the upper value that is greater than the column value, the options are: Percentile 95th , Percentile 98th , User input.
- Normalization:** Select from the options Z-score, Min Max, Logistic, Log normal, Tanh to normalize data.
- Skip Cleansing and Skip Normalization** if available data is cleansed and normalized or user wants to view raw data to be analyzed.

- Feature engineering:** Select the reduction method. No reduction, PCA based dimensionally reduction.
- PCA based dimensionally reduction:**
Select Matrix generation based on: Correlation / Covariance.
- Proportion of variance to:**
Use the slider to increase or reduce the value as required.
- Normalization:**
Select Z-score to standardize the values.

5

New Model

Choose analysis | Select data | Prepare data | Feature engineering | Sample and split | Choose algorithm(s)

Choose algorithm(s) and enter input parameters

☐ Two class logistic regression

☐ Multi class logistic regression

☐ Decision tree classification

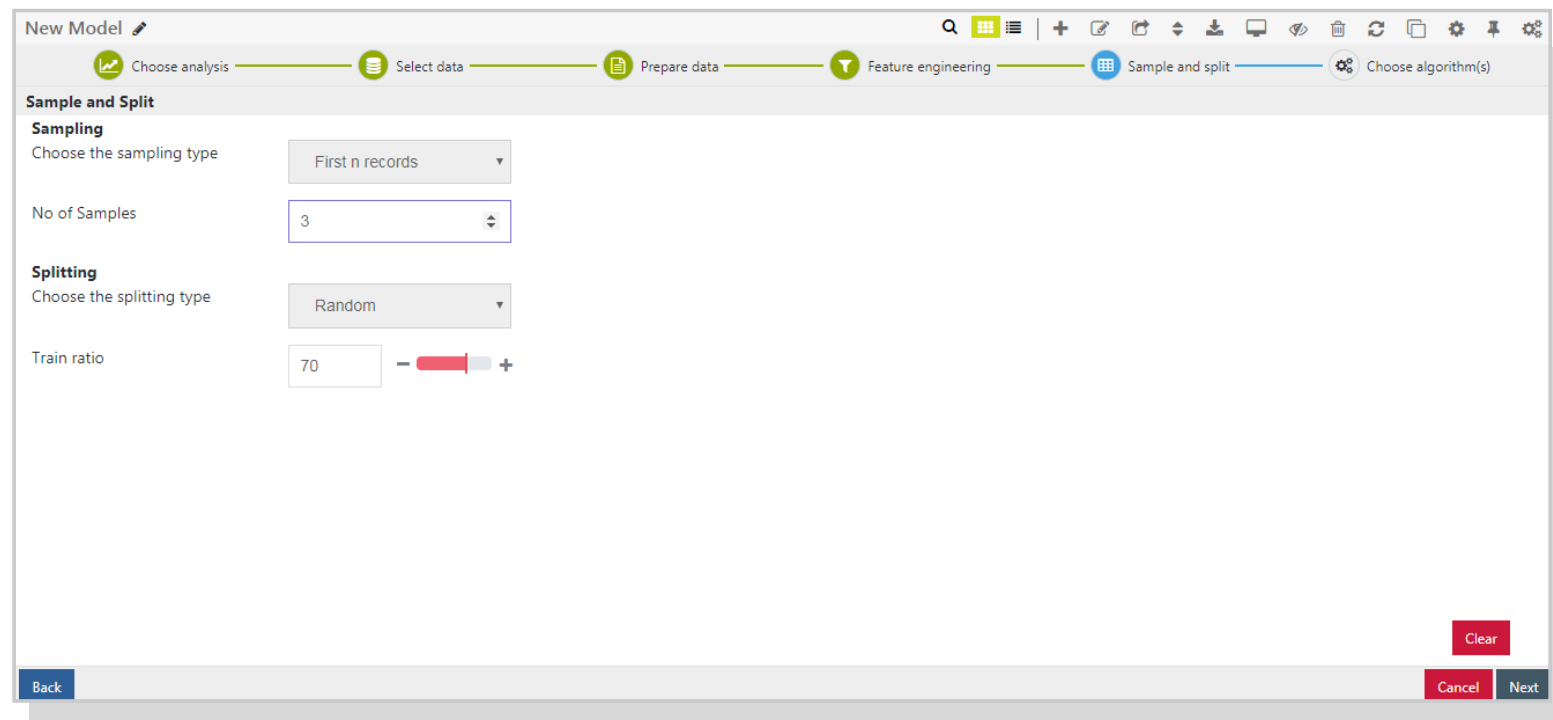
☐ Two class K-nearest neighbor

☐ Multi class K-nearest neighbor

Algorithms selected: 0

Back Cancel Execute

Quick Start Reference Card - MODEL



New Model

Choose analysis — Select data — Prepare data — Feature engineering — Sample and split — Choose algorithm(s)

Sample and Split

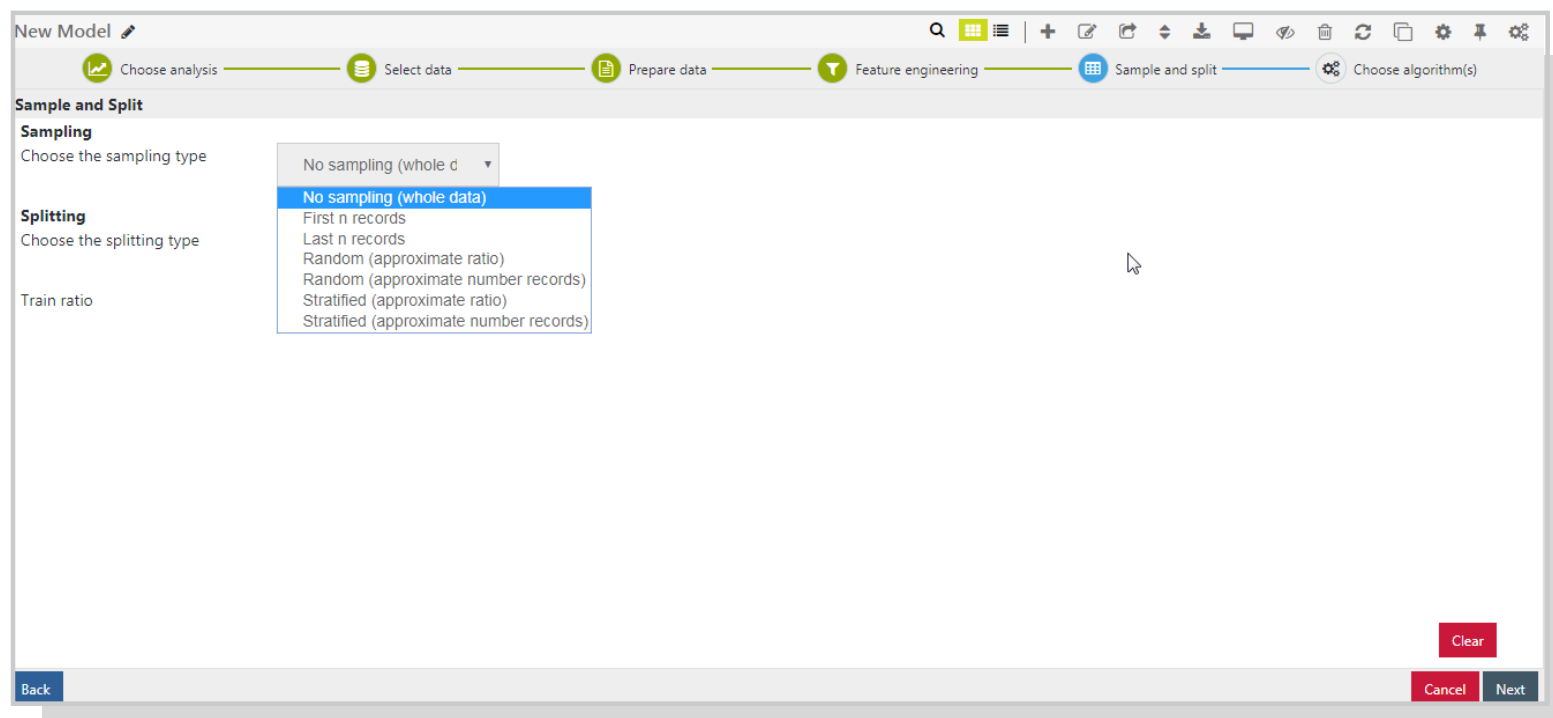
Sampling
Choose the sampling type: First n records

No of Samples: 3

Splitting
Choose the splitting type: Random

Train ratio: 70

Back Clear Cancel Next



New Model

Choose analysis — Select data — Prepare data — Feature engineering — Sample and split — Choose algorithm(s)

Sample and Split

Sampling
Choose the sampling type: No sampling (whole data)

Splitting
Choose the splitting type: Random

Train ratio

Back Clear Cancel Next

Sampling and Split : Choose the sampling and splitting the data method from the options.

Sampling: Options for selecting the Sampling method:

- First n records
- Last n records
- Random (approximate ratios)
- Random (approximate no.of.records)
- Stratified (approximate ratio): Specify - No.of. samples, Column Information: select from the drop down list.
- Stratified (approximate number of records): Specify - No.of.samples, Column Information: select from the drop down list.

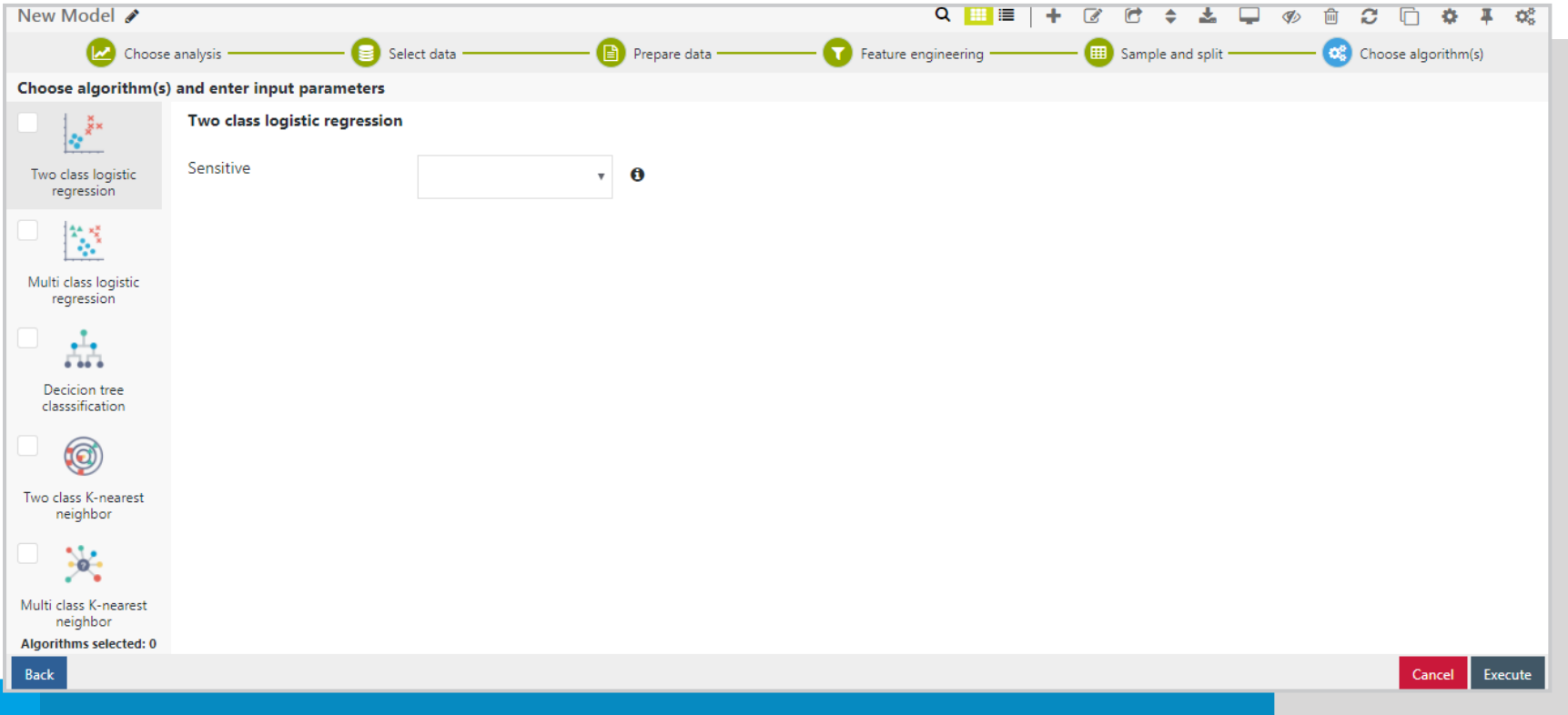
Splitting: Choose the Splitting type. Random

- Train ratio: The defaulted value is specified, the ratio can be increased or decreased using the slider.

6

Choose algorithms :
Choose the type of algorithm(s) from the available algorithms for calculating the values for the model and specify the input parameters.

7



ALGORITHMS	INPUT PATRAMETERS
Two Class logistic regression	Sensitive value
Multi class logistic regression	No input parameter
Decision tree classification	Sensitive value
Two class K-nearest neighbor	Sensitive value
Multi class k-nearest neighbor	No input parameter
Two class random forest	Maximum depth, number of tress, sensitive, minimum split, minimum bucket, maximum leaf node
Two class neural network	Sensitive value
Two class support vector machine	Sensitive, Cost function C, Gamma value, Kernel
Gradient boosting machine	Minimum split, Maximum depth, Sensitive, Bagging size, Estimators, Feature strategy, Learning rate, Loss function
Ada boosting	Minimum split, Maximum depth, Sensitive, Criterion, Estimators, Feature strategy, Learning rate, Type
Multi class random forest	Minimum split, Minimum bucket, Maximum depth, Number of trees, Maximum leaf node
Ordinal regression	No input parameter

PROCEDURE TO ADD CLUSTERING MODEL

Create Model

Name *

Description

UPLOAD YOUR IMAGE

Drag and drop files here or click to upload

Create

1

Create Model :

Associate a model with a name, description and if required identify it with an image to view as a thumb nail icon.

Choose Analysis :

Select Clustering analysis from the list of options.

2

Choose existing datasource

Available datasources

Search...

HR_Data

e

WineData

MBA

WineData

SampleComments

Cox

HR

WorkAccident (varchar)

Left (varchar)

Department (varchar)

Promotion (varchar)

SalaryType (varchar)

EmployeeId (int)

LastEvaluation (float)

AverageMonthlyHours (int)

SatisfactoryLevel (float)

TimeSpent (int)

Preview - HR

WorkAccident	Left	Department	Promotion	SalaryType	Employee
0	1	sales	0	low	1000001
0	1	sales	0	medium	1000002
0	1	sales	0	medium	1000003
0	1	sales	0	low	1000004
0	1	sales	0	low	1000005
0	1	sales	0	low	1000006
0	1	sales	0	low	1000007
0	1	sales	0	low	1000008
0	1	sales	0	low	1000009
0	1	sales	0	low	1000010
0	1	sales	0	low	1000011
0	1	sales	0	low	1000012
0	1	sales	0	low	1000013
0	1	sales	0	low	1000014
0	1	sales	0	low	1000015
0	1	sales	0	low	1000016
0	1	sales	0	low	1000017
0	1	sales	0	low	1000018

+ New Datasource

Add

3

Choose existing datasource :

Select the database to add data and the preview will be shown. Option to add a new datasource is also available.

New Model

Choose analysis

Select data

Prepare data

Feature engineering

Choose algorithm(s)

Please choose a type of analysis that will better suit model requirement

Prediction

Classification

Clustering

Survival analysis

Anomaly dimensionality reduction

Forecasting

Text classification

Text clustering

Association rules

Cancel

Next

- Prepare Data:** Prepare the data with the required columns.
- Is factor:** Select if the variable is to be treated as categorical.
- Unique id:** Select the unique variable if the datasource has a unique value.
- Cluster Variable:** This variable value is mandatory. Select a variable that you will like to cluster for the analysis.
- Imputation:** Select substitute value in case there are missing values for each variable. Options: Most frequent, Least frequent, User Input.
- Lower outlier:** Select the lower value that is less than the variable value, the options are: Percentile 2nd, Percentile 5th, User input.
- Rescaling:** Select from options to apply rescaling
- Apply rescaling for the entire matrix:** Enable to apply rescaling and select from the options listed.
- Skip cleansing:** Enable to skip cleansing. On selecting Imputation, Lower outlier, Upper outlier columns will be deactivated.

4

New Model

Choose analysis | Select data | Prepare data | Feature engineering | Choose algorithm(s)

Please prepare data with required column for model preparation

Apply rescaling for the entire matrix ☐ Select Skip cleansing ☐

Column name	Data type	Is factor	Unique id	Cluster variable	Imputation	Lower outlier	Upper outlier	Rescaling
<input type="checkbox"/> WorkAccident	varchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> Left	varchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> Department	varchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> Promotion	varchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> SalaryType	varchar	<input checked="" type="checkbox"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> EmployeeId	int	<input type="checkbox"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> LastEvaluation	float	<input type="checkbox"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> AverageMonthlyHours	int	<input type="checkbox"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> SatisfactoryLevel	float	<input type="checkbox"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> TimeSpent	int	<input type="checkbox"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select
<input type="checkbox"/> NumberProject	int	<input type="checkbox"/>	<input type="radio"/>	<input type="checkbox"/>	Select	Select	Select	Select

Total columns: 11, Selected columns: 0

Preview limited to 50 rows

Back Clear all Cancel Next

New Model

Choose analysis | Select data | Prepare data | Feature engineering | Choose algorithm(s)

Feature engineering

Reduction method

No feature reduction enabled.

No reduction

No reduction

PCA based dimensionality reduction

Clear

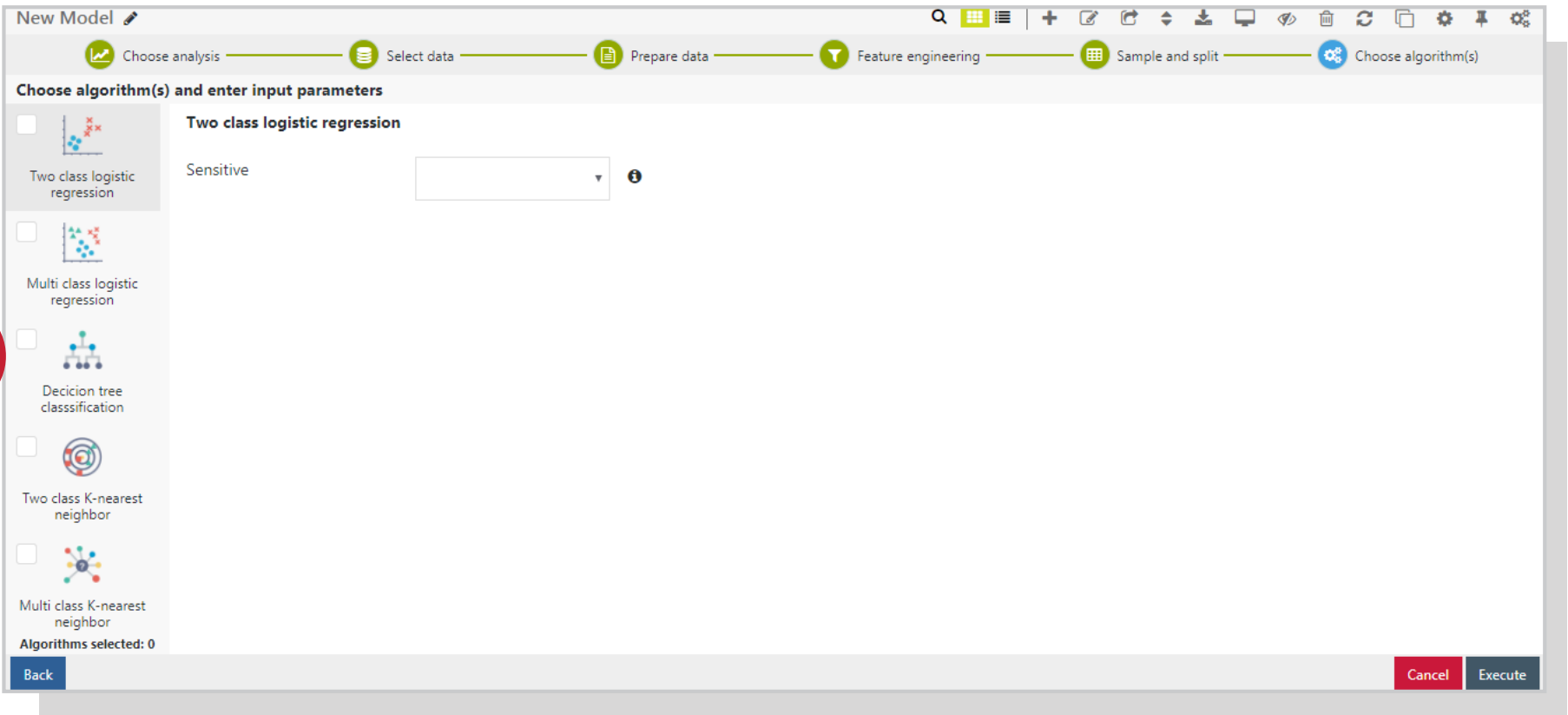
Back Cancel Next

5

- Feature engineering:** Select the reduction method. Tree based, Principle component analysis, Lasso Regression.
- Tree based:** Specify No.ofTrees and Depth of Trees.
- Principle component:** Select Matrix generation based on: Coorelation / Covariance.
- Proportion of variance to:** Use the slider to increase or reduce the value as required.
- Normalization:** Select Z-score to standardize the values.

Choose algorithms :
Choose the type of algorithm(s) from the available algorithms for calculating the values for the model and specify the input parameters.

6

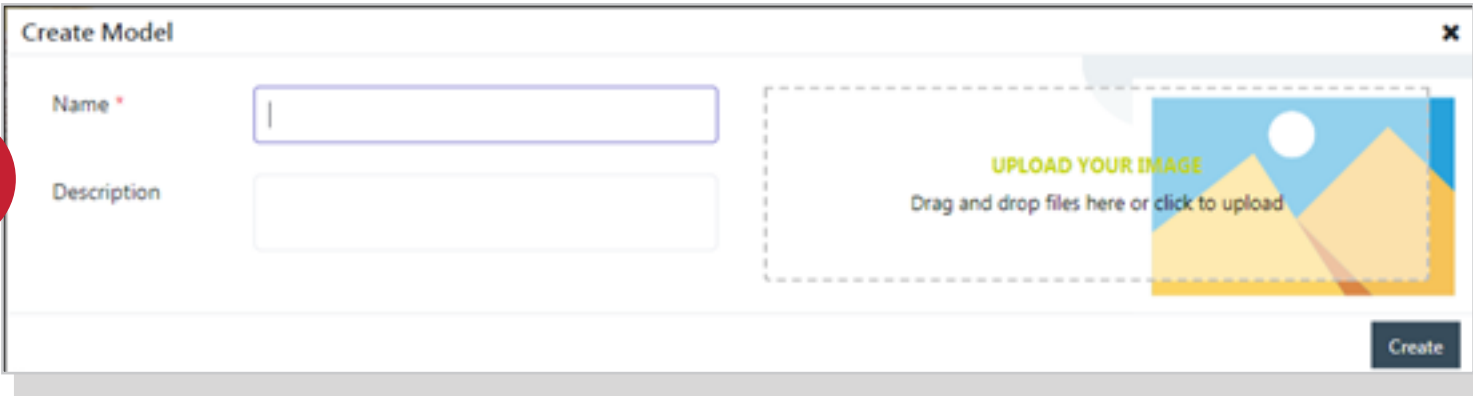


ALGORITHMS	INPUT PATRAMETERS
K-means clustering	Exact K
Mini batch K – means clustering	Exact K, Batch size, Maximum iterations, Number of initialization, Initialization fraction, Initialization
Gaussian mixture model	Minimum iterations, Number of components
Partition around medioids	Exact K, Distance metric
Ward hierarchical clustering	Exact K, Distance metric, Method
DB Scan	Optimal epsilon(eps), Minimum points
Adaptive K-means clustering	Maximum iterations, Distance metric, Minimum K, Maximum K, Threshold value, N Start

PROCEDURE TO ADD SURVIVAL ANALYSIS MODEL

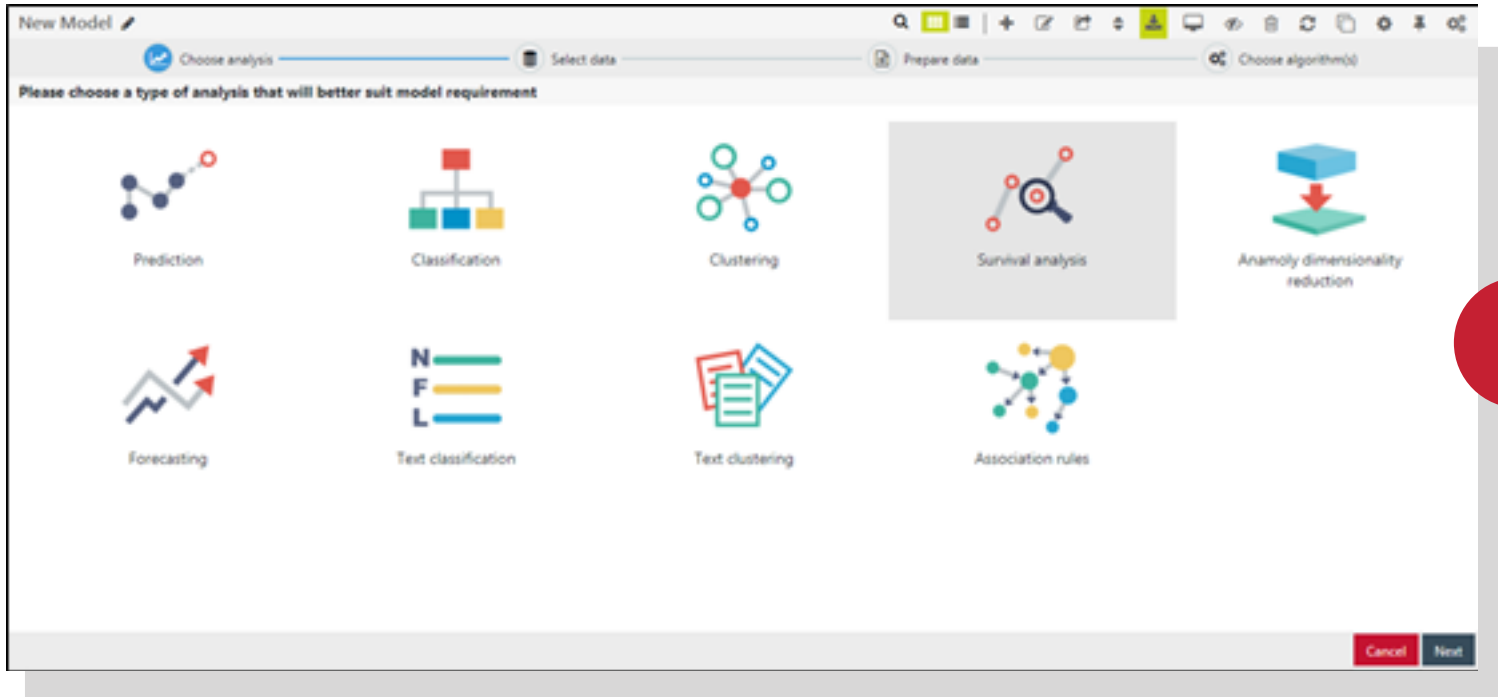
Create Model:
Associate a model with a name, description and if required identify it with an image to view as a thumb nail icon.

1



The 'Create Model' dialog box contains the following elements:

- Name:** A text input field with a red asterisk indicating it is required.
- Description:** A larger text input field.
- Image Upload:** A dashed rectangular box containing the text 'UPLOAD YOUR IMAGE' and 'Drag and drop files here or click to upload'. To the right of this box is a thumbnail image of a landscape with mountains and a sun.
- Create Button:** A dark blue button with the text 'Create' in white, located at the bottom right of the dialog.



The 'New Model' screen features a top navigation bar with the following tabs: 'Choose analysis' (active), 'Select data', 'Prepare data', and 'Choose algorithm(s)'. Below the navigation bar, a message reads: 'Please choose a type of analysis that will better suit model requirement'. The main area displays a grid of analysis options, each with an icon and a label:

- Prediction (Icon: three connected dots)
- Classification (Icon: a tree diagram)
- Clustering (Icon: a network of nodes)
- Survival analysis** (Icon: a magnifying glass over a line graph, highlighted with a grey background)
- Anomaly dimensionality reduction (Icon: a 3D cube with a red arrow pointing down)
- Forecasting (Icon: a line graph with an upward arrow)
- Text classification (Icon: the letters N, F, and L stacked vertically)
- Text clustering (Icon: three overlapping document icons)
- Association rules (Icon: a network of nodes with colored circles)

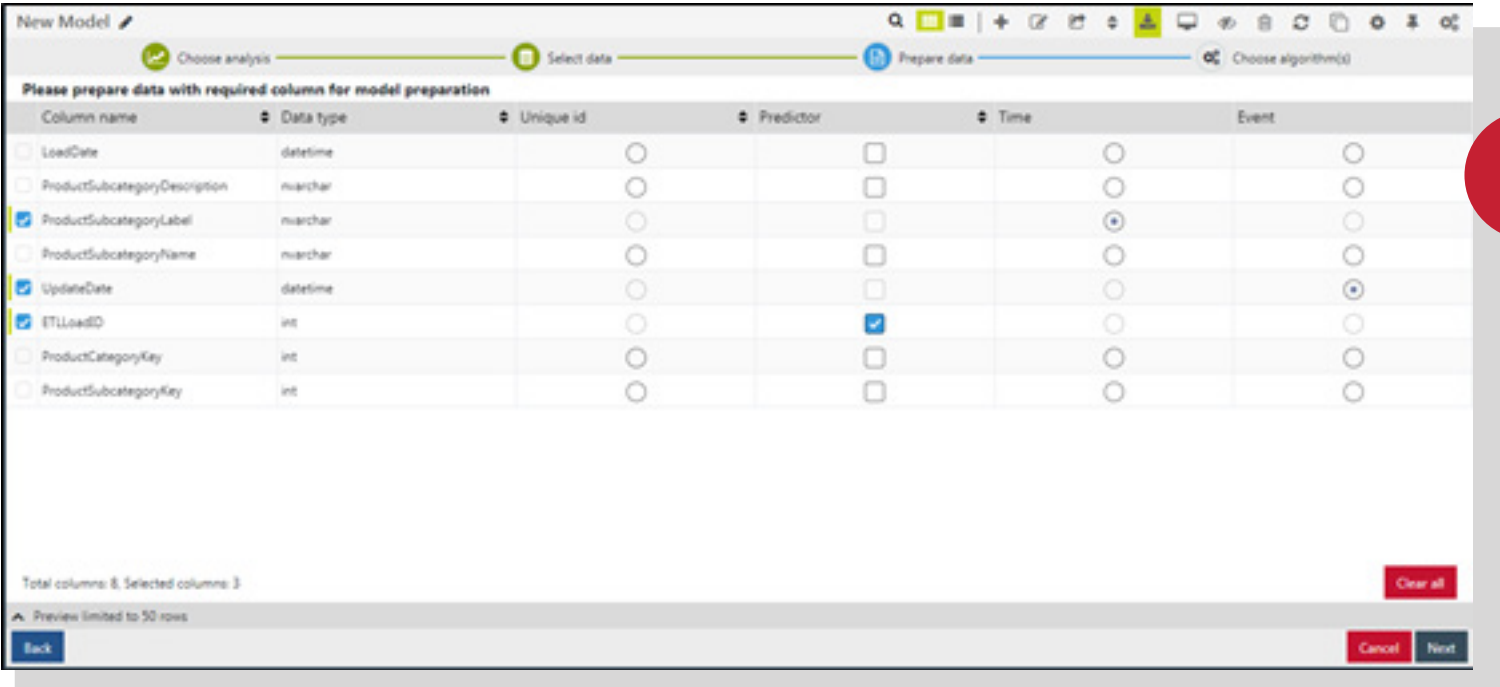
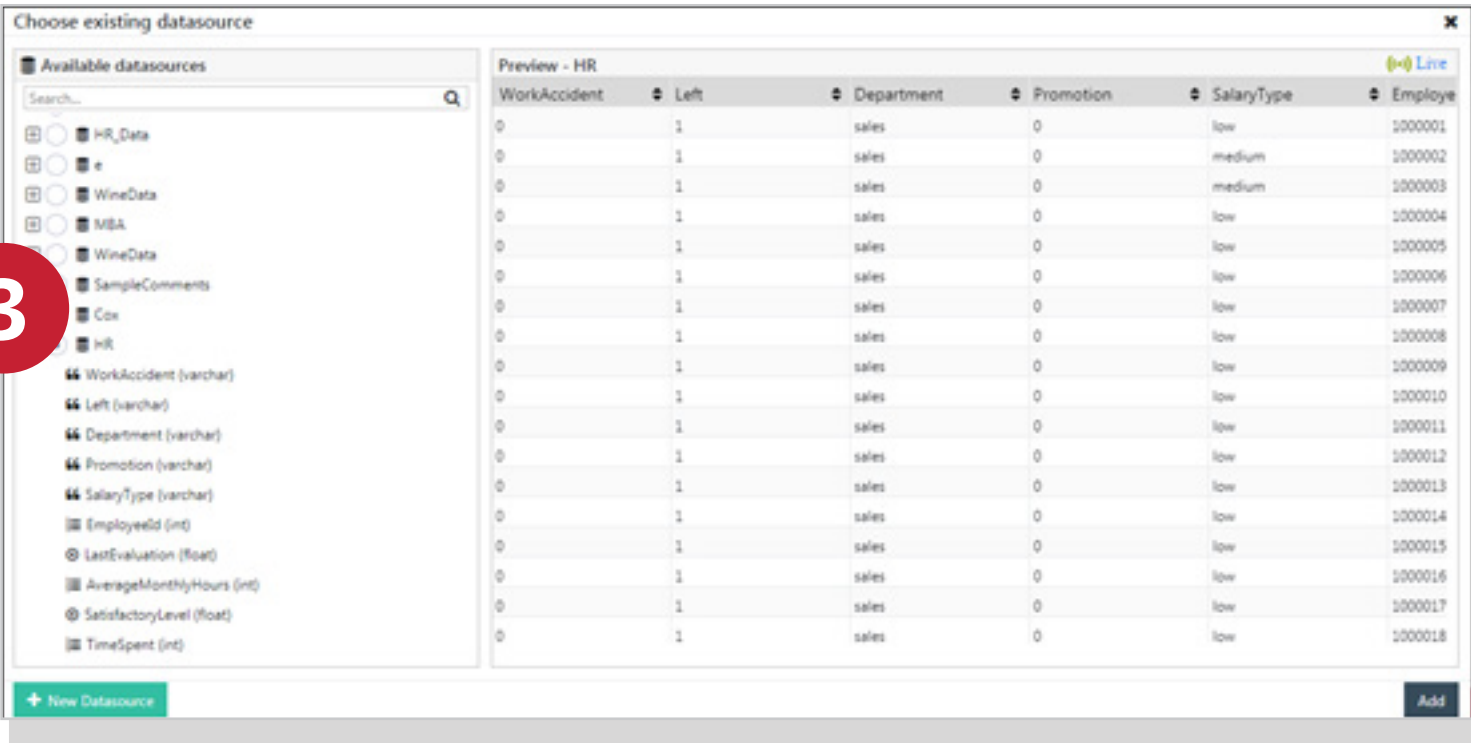
At the bottom right of the screen are two buttons: 'Cancel' (red) and 'Next' (dark blue).

2

Choose Analysis:
Select Survival analysis from the list of options.

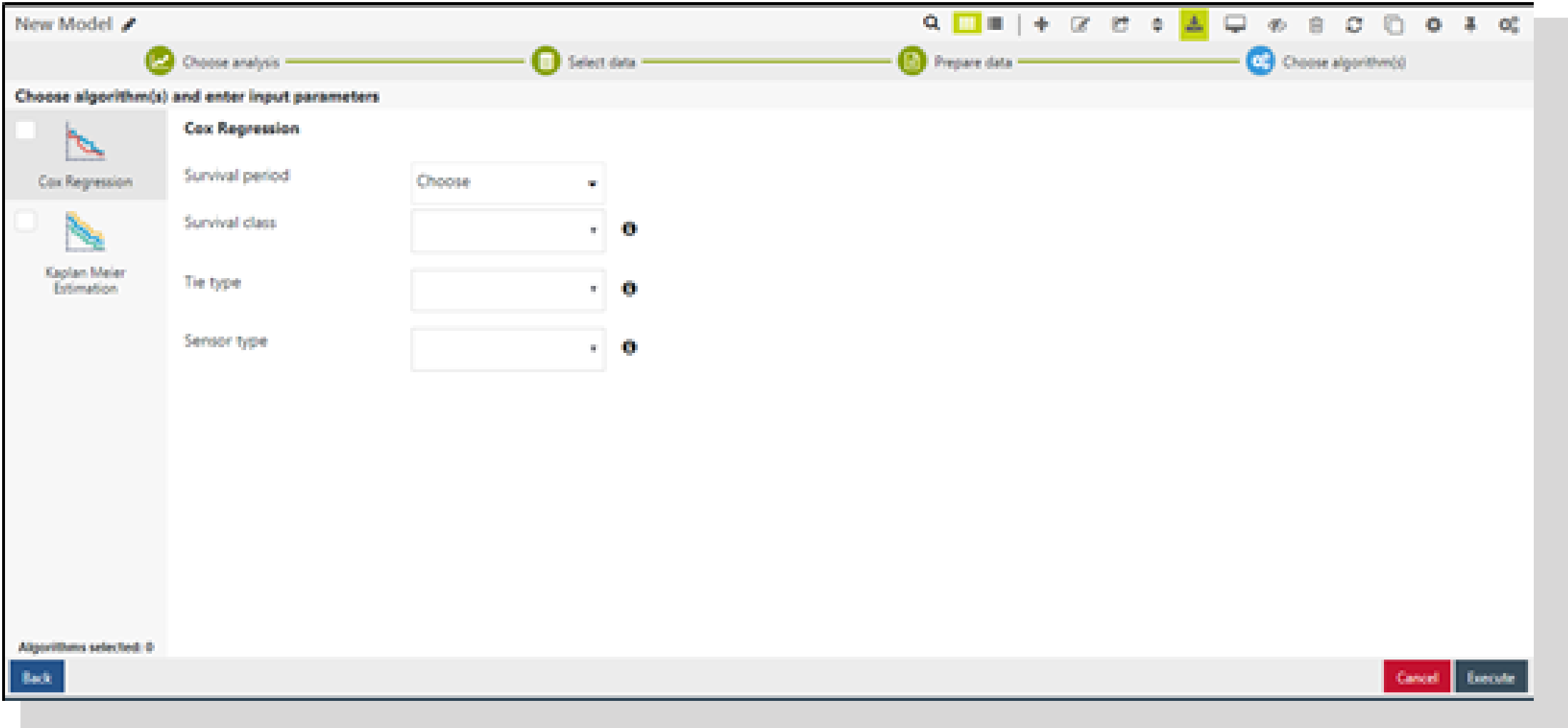
Choose existing datasource:
Select the database to add data and the preview will be shown. Option to add a new datasource is also available.

3



4

Prepare Data: Prepare the data with the required columns.
Data type: List of data type from the selected table.
Unique id: Select the unique variable if the datasource has a unique value.
Predictor: This column value is mandatory and independent. Select columns to influence the target variable. On selecting Predictor for a particular variable all the other parameter are disabled for the particular column.
Time: This column value is mandatory. On selecting Time for a particular variable all the other parameter are disabled for the particular column.
Event: This column is mandatory. On selecting an event for a particular variable all other factors for the variable such as Predictor, Time are disabled.



5

Choose Algorithm(s):
Select the algorithm(s) to analyze data:

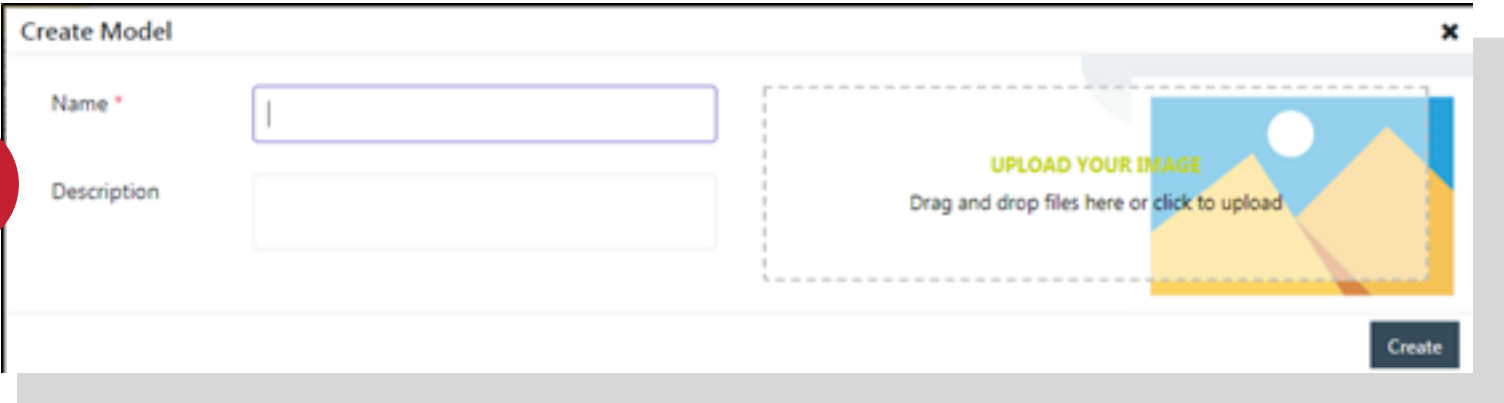
ALGORITHMS	INPUT PATRAMETERS
Cox Regression	Survival period, Survival class, Tie Type, Sensor Type
Kaplar Meier Estimation	Censor Type, Variable Type, Confidence Interval.

PROCEDURE TO ADD ANOMALY / DIMENSIONALITY REDUCTION MODEL

Create Model:

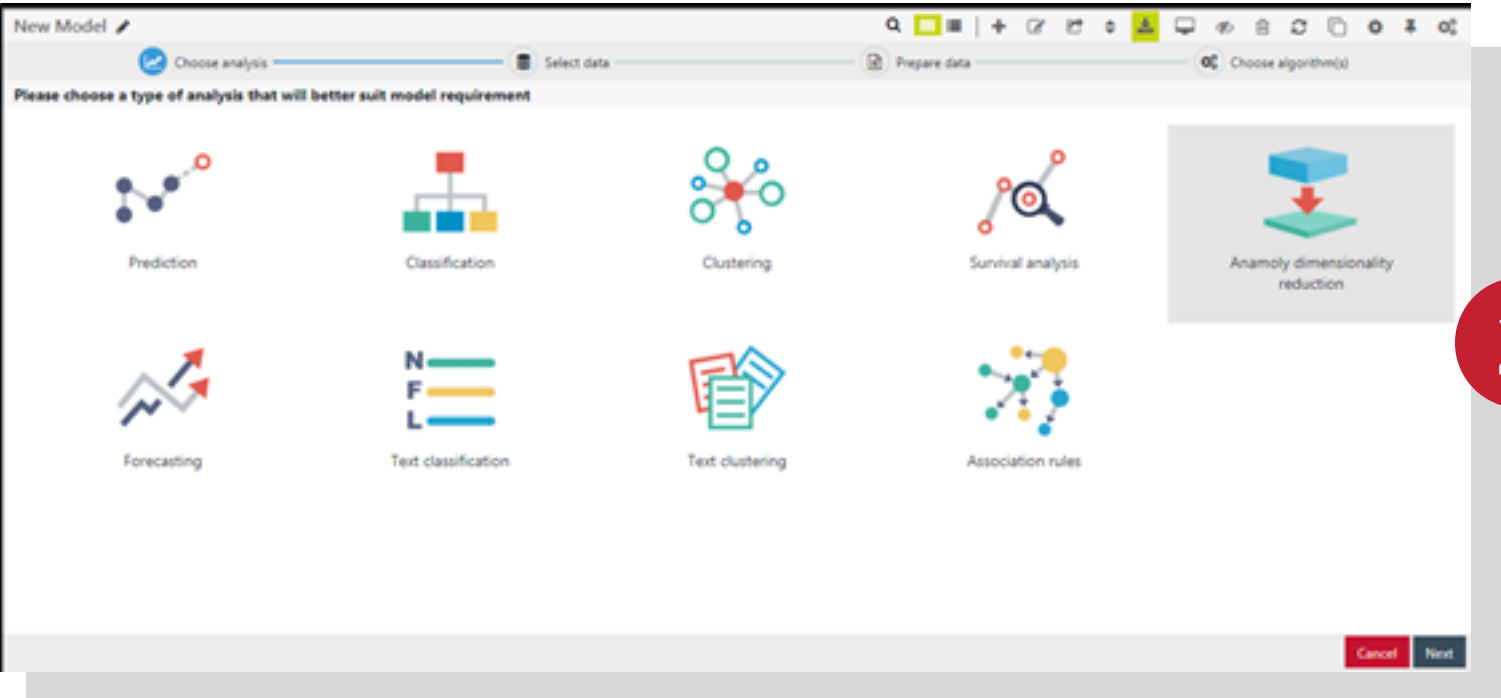
Associate a model with a name, description and if required identify it with an image to view as a thumb nail icon.

1



The 'Create Model' dialog box contains the following elements:

- Name ***: A text input field.
- Description**: A text input field.
- Image Upload Area**: A dashed box containing the text 'UPLOAD YOUR IMAGE' and 'Drag and drop files here or click to upload', with a thumbnail image of a landscape.
- Create**: A button at the bottom right.



The 'New Model' dialog box features a progress bar at the top with four steps: 'Choose analysis' (active), 'Select data', 'Prepare data', and 'Choose algorithm(s)'. Below the progress bar, it says 'Please choose a type of analysis that will better suit model requirement'. The main area displays eight analysis options with icons and labels:

- Prediction
- Classification
- Clustering
- Survival analysis
- Forecasting
- Text classification
- Text clustering
- Association rules

The 'Anomaly dimensionality reduction' option is highlighted with a grey background. At the bottom right, there are 'Cancel' and 'Next' buttons.

2

Choose Analysis:

Select Anomaly / dimensionality reduction from the list of options.

Choose existing datasource:
Select the database to add data and the preview will be shown. Option to add a new datasource is also available.

3

Choose existing datasource

Available datasources

Search...

HelloWorld_DS

MBA

Demo_datasource

sdflgh

MBA-copy

HR_data

lala

ra

HR_Data

e

WineData

MBA

WineData

SampleComments

Cox

HR

House Data

testtt

+ New Datasource

Preview - HelloWorld_DS

Live

LoadDate	ProductSubcategory	ProductSubcategory	ProductSubcategory	UpdateDate	ETLLoadID
2009-07-07T00:00:00	MP4&MP3-testing	0001	MP4&MP3	2009-07-07T00:00:00	1
2015-08-20T00:00:00	Recorder	0002	Recorder-tst	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Radio	0003	Radio	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Recording Pen	0004	Recording Pen	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Headphones	0005	Headphones	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Bluetooth Headphones	0006	Bluetooth Headphones	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Speakers	0007	Speakers	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Televisions	0001	Televisions	2009-07-07T00:00:00	1
2009-07-07T00:00:00	VCD & DVD	0002	VCD & DVD	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Home Theater System	0003	Home Theater System	2009-07-07T00:00:00	2
2009-07-07T00:00:00	Car Video	0005	Car Video	2009-07-07T00:00:00	1
2009-07-07T00:00:00	TV & Video Accessories	0006	TV & Video Accessories	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Laptops	0001	Laptops	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Netbooks	0002	Netbooks	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Desktops	0003	Desktops	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Monitors	0004	Monitors	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Projectors & Screens	0005	Projectors & Screens	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Printers, Scanners & Fax	0006	Printers, Scanners & Fax	2009-07-07T00:00:00	1

Add

New Model

Choose analysis

Select data

Prepare data

Choose algorithm(s)

Please prepare data with required column for model preparation

Column name	Data type	Unique id	Target
<input checked="" type="checkbox"/> LoadDate	datetime	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> ProductSubcategoryDescription	nvarchar	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> ProductSubcategoryLabel	nvarchar	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> ProductSubcategoryName	nvarchar	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> UpdateDate	datetime	<input type="radio"/>	<input type="radio"/>
<input checked="" type="checkbox"/> ETLLoadID	int	<input type="radio"/>	<input checked="" type="radio"/>
<input type="checkbox"/> ProductCategoryKey	int	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> ProductSubcategoryKey	int	<input type="radio"/>	<input type="radio"/>

Total columns: 8, Selected columns: 2

Preview limited to 50 rows

Back

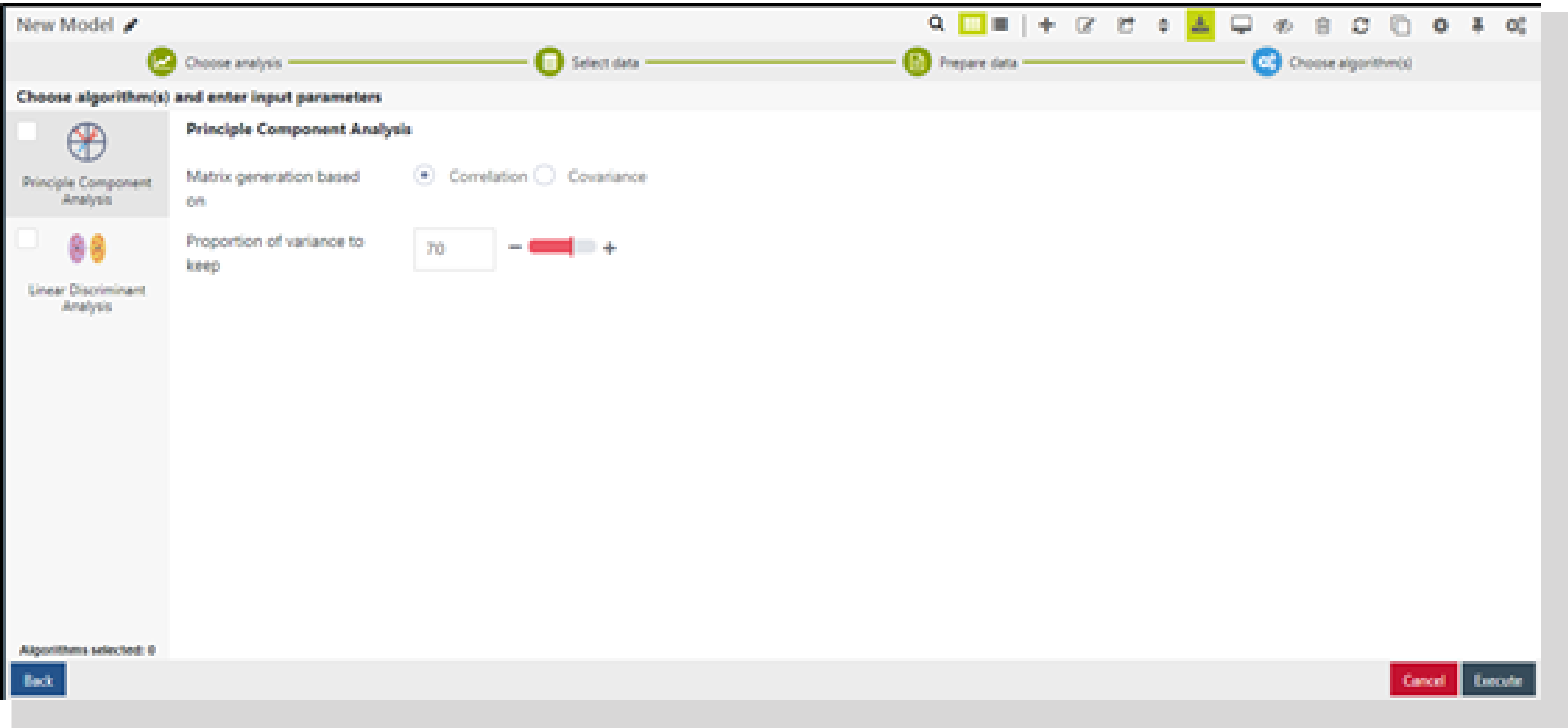
Clear all

Cancel

Next

4

Prepare Data: Prepare the data with the required columns.
Data type: List of data type from the selected table.
Unique id: Select the unique variable if the datasource has a unique value.
Target: This variable value is mandatory. Select a variable that you will like to predict that is dependent for the analysis.
On selecting either of Unique id or Target the variable will be disabled.



5

Choose Algorithm(s):
Choose the type of algorithm(s) from the available algorithms for calculating the values for the model and specify the input parameters.


ALGORITHMS	INPUT PATRAMETERS
Principle Component Analysis	Matrix generated based on: Correlation / Covariance Proportion of variance to keep: Increase or decrease the slider
Linear Discriminant Analysis	No input parameter

PROCEDURE TO ADD FORECASTING MODEL

Create Model:

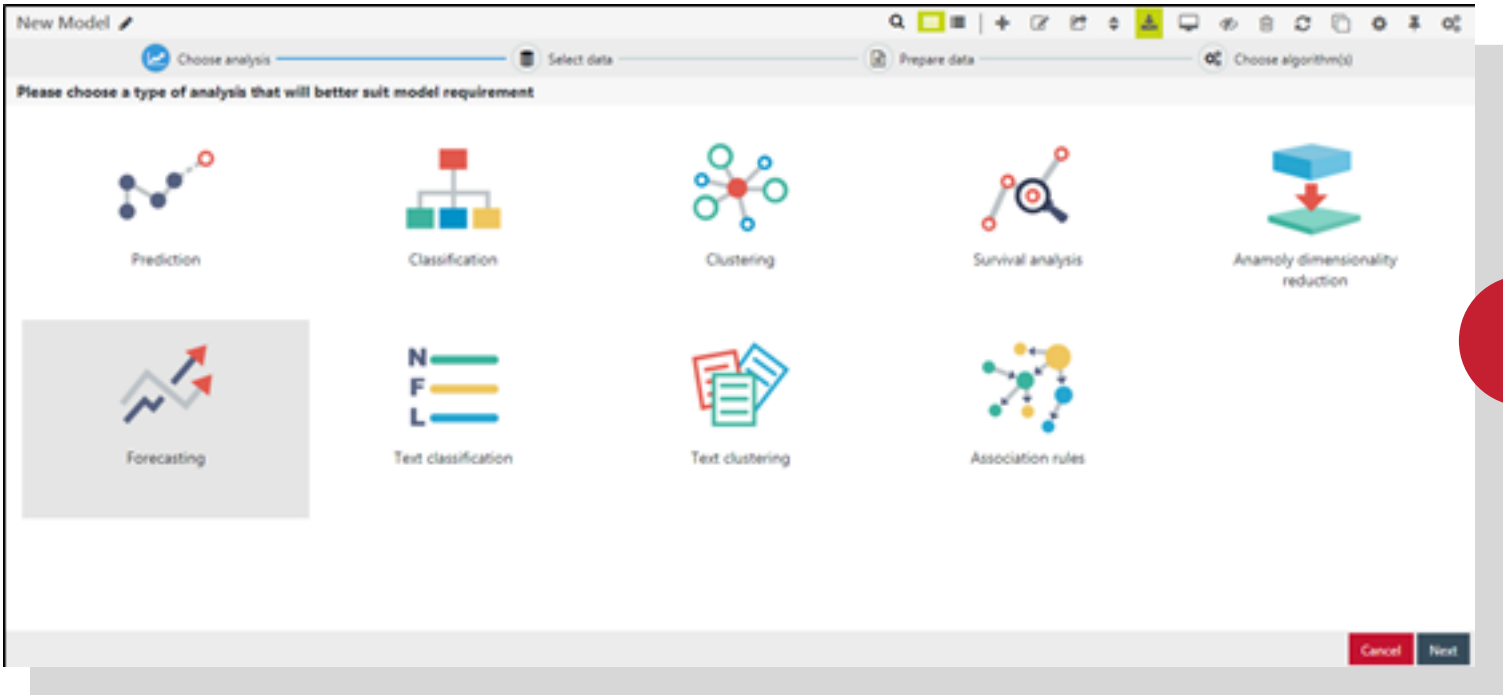
Associate a model with a name, description and if required identify it with an image to view as a thumb nail icon.

1



The 'Create Model' dialog box contains the following elements:

- Name ***: A text input field.
- Description**: A text input field.
- Image Upload**: A dashed box containing the text 'UPLOAD YOUR IMAGE' and 'Drag and drop files here or click to upload', with a thumbnail image of a landscape.
- Create**: A button at the bottom right.



The 'New Model' dialog box features a top navigation bar with tabs: 'Choose analysis' (selected), 'Select data', 'Prepare data', and 'Choose algorithm(s)'. Below the tabs, it says 'Please choose a type of analysis that will better suit model requirement'. The main area displays ten analysis options with icons and labels:

- Prediction
- Classification
- Clustering
- Survival analysis
- Anomaly dimensionality reduction
- Forecasting (highlighted with a grey background)
- Text classification
- Text clustering
- Association rules

At the bottom right are 'Cancel' and 'Next' buttons.

2

Choose Analysis:

Select Forecasting from the list of options.

Choose existing datasource:
Select the database to add data and the preview will be shown. Option to add a new datasource is also available.

3

Choose existing datasource

Available datasources

Search...

HelloWorld_DS

MBA

Demo_datasource

sdflgh

MBA-copy

HR_data

lala

ra

HR_Data

e

WineData

MBA

WineData

SampleComments

Cox

HR

House Data

testtt

+ New Datasource

Preview - HelloWorld_DS

LoadDate

ProductSubcategory

ProductSubcategory

ProductSubcategory

UpdateDate

ETLLoadID

2009-07-07T00:00:00	MP4&MP3-testing	0001	MP4&MP3	2009-07-07T00:00:00	1
2015-08-20T00:00:00	Recorder	0002	Recorder-tst	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Radio	0003	Radio	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Recording Pen	0004	Recording Pen	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Headphones	0005	Headphones	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Bluetooth Headphones	0006	Bluetooth Headphones	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Speakers	0007	Speakers	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Televisions	0001	Televisions	2009-07-07T00:00:00	1
2009-07-07T00:00:00	VCD & DVD	0002	VCD & DVD	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Home Theater System	0003	Home Theater System	2009-07-07T00:00:00	2
2009-07-07T00:00:00	Car Video	0005	Car Video	2009-07-07T00:00:00	1
2009-07-07T00:00:00	TV & Video Accessories	0006	TV & Video Accessories	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Laptops	0001	Laptops	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Netbooks	0002	Netbooks	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Desktops	0003	Desktops	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Monitors	0004	Monitors	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Projectors & Screens	0005	Projectors & Screens	2009-07-07T00:00:00	1
2009-07-07T00:00:00	Printers, Scanners & Fax	0006	Printers, Scanners & Fax	2009-07-07T00:00:00	1

Add

New Model

Choose analysis

Select data

Prepare data

Choose algorithm(s)

Please prepare data with required column for model preparation

Column name	Data type	Unique id	Target
<input checked="" type="checkbox"/> LoadDate	datetime	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> ProductSubcategoryDescription	nvarchar	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> ProductSubcategoryLabel	nvarchar	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> ProductSubcategoryName	nvarchar	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> UpdateDate	datetime	<input type="radio"/>	<input type="radio"/>
<input checked="" type="checkbox"/> ETLLoadID	int	<input type="radio"/>	<input checked="" type="radio"/>
<input type="checkbox"/> ProductCategoryKey	int	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> ProductSubcategoryKey	int	<input type="radio"/>	<input type="radio"/>

Total columns: 8, Selected columns: 2

Clear all

Preview limited to 50 rows

Back

Cancel

Next

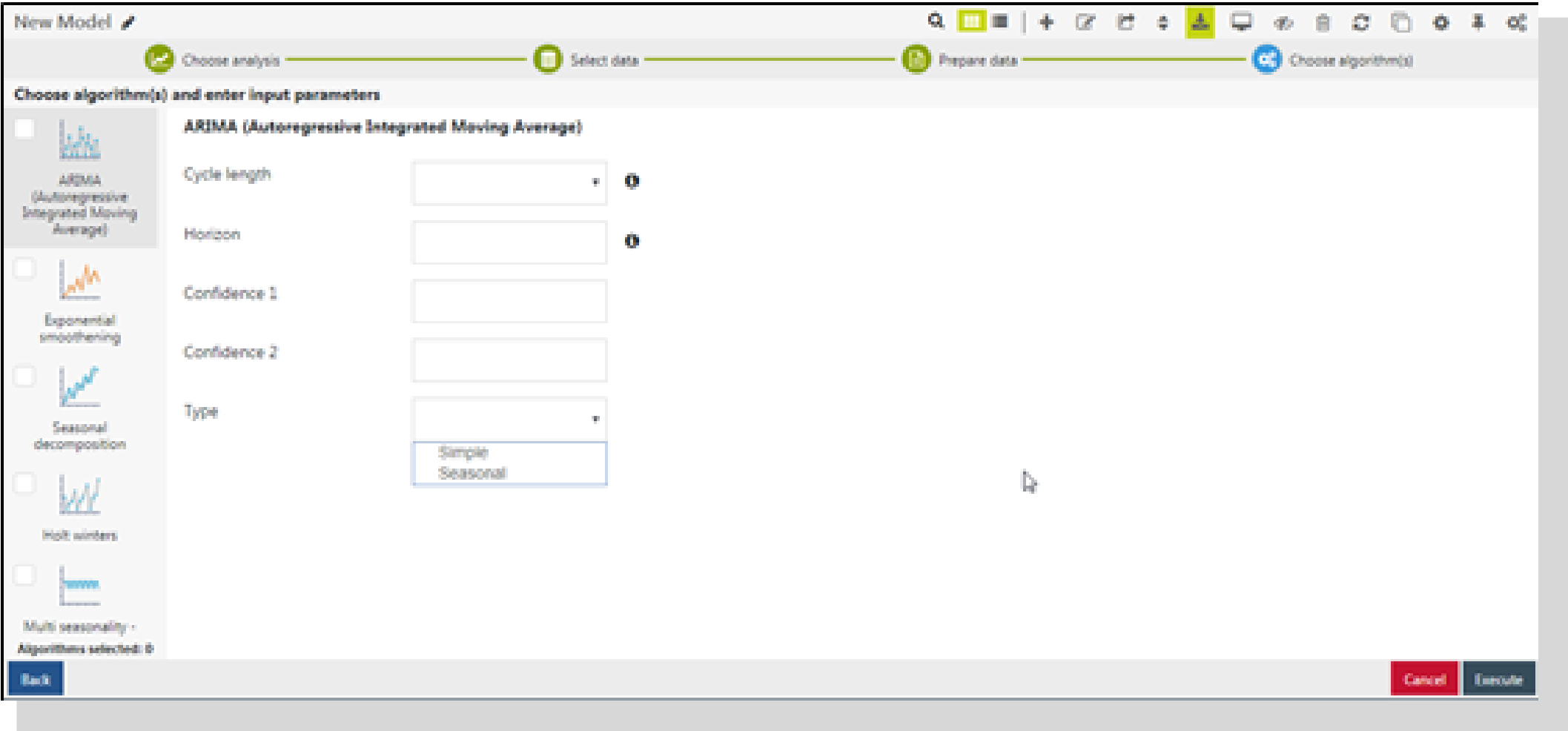
4

Prepare Data: Prepare the data with the required columns.

Data type: List of data type from the selected table.

Date column: This is mandatory column. Select the date variable that will be the base for forecasting the value. On selecting the Date column for a variable the Value column is disabled for the variable.

Value column: This is also a mandatory column. Select the desired forecasted value. On selecting the Value column for a variable Date column is disabled for the variable.



5

Choose Algorithm(s):

Choose the type of algorithm(s) from the available algorithms for calculating the values for the model and specify the input parameters.

ALGORITHMS	INPUT PATRAMETERS
ARIMA (Auto regressive integrated moving average)	Cycle length, Horizon, Confidence 1, Confidence 2, Type
Exponential smoothening	Cycle length, Horizon, Confidence 1, Confidence 2
Seasonal decomposition	Cycle length, Horizon, Confidence 1, Confidence 2
Holt winters	Cycle length, Horizon, Confidence 1, Confidence 2, Type
Multi seasonality (TBATS)	Cycle length, Horizon, Confidence 1, Confidence 2

PROCEDURE TO ADD TEXT CLASSIFICATION MODEL

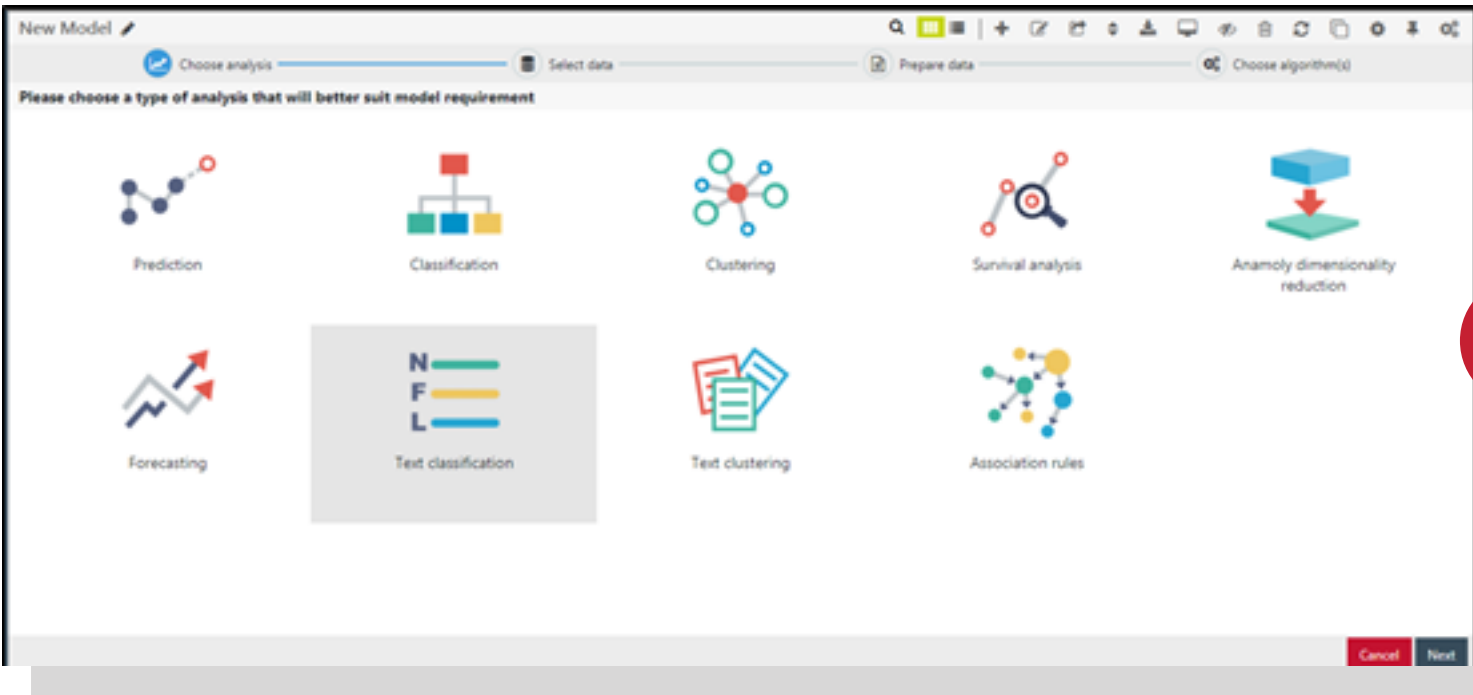
Create Model:
Associate a model with a name, description and if required identify it with an image to view as a thumb nail icon.

1



The 'Create Model' dialog box contains the following elements:

- Name:** A text input field with a red asterisk indicating it is required.
- Description:** A larger text input field.
- Image Upload:** A dashed box containing the text 'UPLOAD YOUR IMAGE' and 'Drag and drop files here or click to upload', with a thumbnail image of a landscape.
- Create:** A button at the bottom right.



The 'New Model' dialog box features a progress bar at the top with four steps: 'Choose analysis' (active), 'Select data', 'Prepare data', and 'Choose algorithm(s)'. Below the progress bar, it says 'Please choose a type of analysis that will better suit model requirement'. The main area displays ten analysis options with icons:

- Prediction
- Classification
- Clustering
- Survival analysis
- Anomaly dimensionality reduction
- Forecasting
- Text classification** (highlighted with a grey background)
- Text clustering
- Association rules

At the bottom right, there are 'Cancel' and 'Next' buttons.

2

Choose Analysis:
Select Text Classification analysis from the list of options.

Choose existing datasource:
Select the database to add data and the preview will be shown. Option to add a new datasource is also available.

3

Choose existing datasource

Available datasources

Search...

HelloWorld_DS

MBA12

Demo_datasource

HR_Data

MBA

Column 23 (varchar)

Column 0 (varchar)

Column 10 (varchar)

Column 11 (varchar)

Column 12 (varchar)

Column 13 (varchar)

Column 14 (varchar)

Column 15 (varchar)

Column 16 (varchar)

Column 17 (varchar)

Column 18 (varchar)

Column 19 (varchar)

Column 2 (varchar)

+ New Datasource

Preview - Demo_datasource

Live

ProductCategoryName	UpdateDate	ProductName	Manufacturer	ColorName	ClassNar
Bluetooth Headphones	2000-01-01T00:00:00	Contoso Battery charger ...	Contoso, Ltd	Black	Economy
Cameras and camcorders	2000-01-01T00:00:00	Fabrikam Budget Mouiem...	Fabrikam, Inc.	White	Regular
Cameras and camcorders	2000-01-01T00:00:00	Fabrikam SLR Camera 35'...	Fabrikam, Inc.	Blue	Regular
TV and Video	2000-01-01T00:00:00	Adventure Works 32" LC...	Adventure Works	Black	Regular
Games and Toys	2000-01-01T00:00:00	MOS King& Myths: The A...	Tailspin Toys	Black	Regular
Bluetooth Headphones	2000-01-01T00:00:00	Adventure Works LCD17 ...	Adventure Works	Black	Economy
Bluetooth Headphones	2000-01-01T00:00:00	Proseware Projector 480p...	Proseware, Inc.	Black	Regular
Cameras and camcorders	2000-01-01T00:00:00	Contoso Carrying Case E3...	Contoso, Ltd	Blue	Economy
Games and Toys	2000-01-01T00:00:00	MOS Age of Mythology: ...	Tailspin Toys	Silver	Economy
Home Appliances	2000-01-01T00:00:00	Contoso Microwave 1.0C...	Contoso, Ltd	Black	Regular
Music, Movies and Audio ...	2000-01-01T00:00:00	Contoso DVD 15-Inch Pla...	Contoso, Ltd	Silver	Deluxe
Games and Toys	2000-01-01T00:00:00	MOS Fable: The Lost Cha...	Tailspin Toys	Silver	Economy
Bluetooth Headphones	2000-01-01T00:00:00	WWI LCD22 M2002 Black	Wide World Importers	Black	Regular
Audiotest	2000-01-01T00:00:00	WWI 1GBPulse Smart pen...	Wide World Importers	Black	Economy
Audiotest	2000-01-01T00:00:00	WWI 2GB Pulse Smart pe...	Wide World Importers	White	Regular
Audiotest	2000-01-01T00:00:00	SV 16xDVD M360 Silver	Southridge Video	Silver	Regular
Bluetooth Headphones	2000-01-01T00:00:00	Proseware Laser Jet Color...	Proseware, Inc.	White	Regular
Bluetooth Headphones	2000-01-01T00:00:00	Proseware Screen 100in ...	Proseware, Inc.	Silver	Economy

Add

New Model

Choose analysis

Select data

Prepare data

Choose algorithm(s)

Please prepare data with required column for model preparation

Column name	Data type	Unique id	Comment	Category
<input type="checkbox"/> ProductCategoryName	nvarchar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> UpdateDate	datetime	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input checked="" type="checkbox"/> ProductName	nvarchar	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> Manufacturer	nvarchar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> ColorName	nvarchar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input checked="" type="checkbox"/> ClassName	nvarchar	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
<input type="checkbox"/> Size	nvarchar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> AvailableForSaleDate	datetime	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input checked="" type="checkbox"/> ProductSubcategoryName	nvarchar	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
<input type="checkbox"/> UnitPrice	money	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> SalesQuantity	int	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> ReturnQuantity	int	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Total columns: 17, Selected columns: 3

Clear all

Preview limited to 50 rows

Back

Cancel

Next

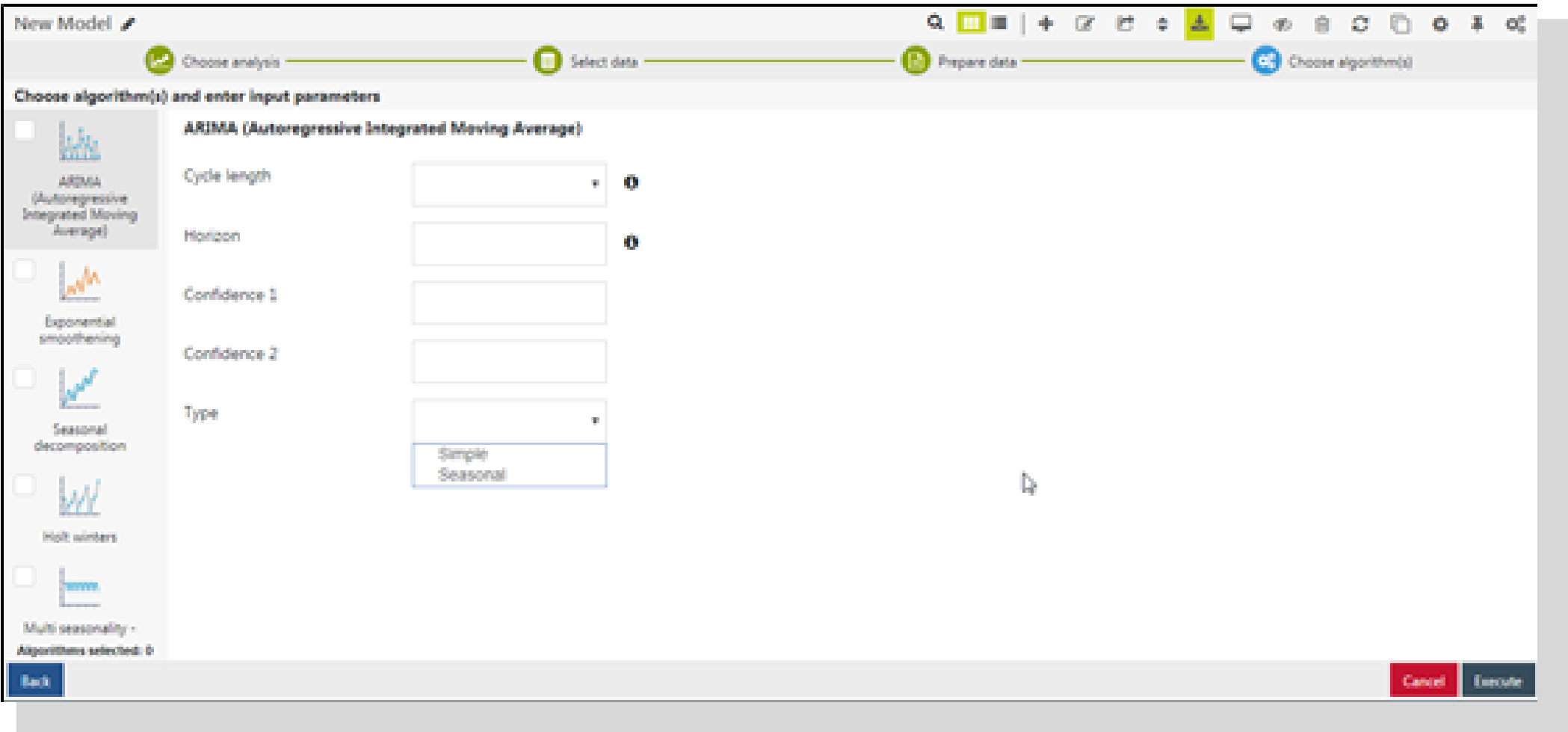
4

Prepare Data: Prepare the data with the required columns.

Unique id: This column is mandatory. Select the unique label to identify it as unique comment.

Comment: This column is mandatory. Select the column that has text comments. If 'Comment' is selected for a column then 'Unique id' and 'Category' for that column are disabled. User will not be able to define more than one criteria for a column.

Category: This column is mandatory. Select a column to define it as Category.



5

Choose Algorithm(s):

Choose the type of algorithm(s) from the available algorithms for calculating the values for the model and specify the input parameters.

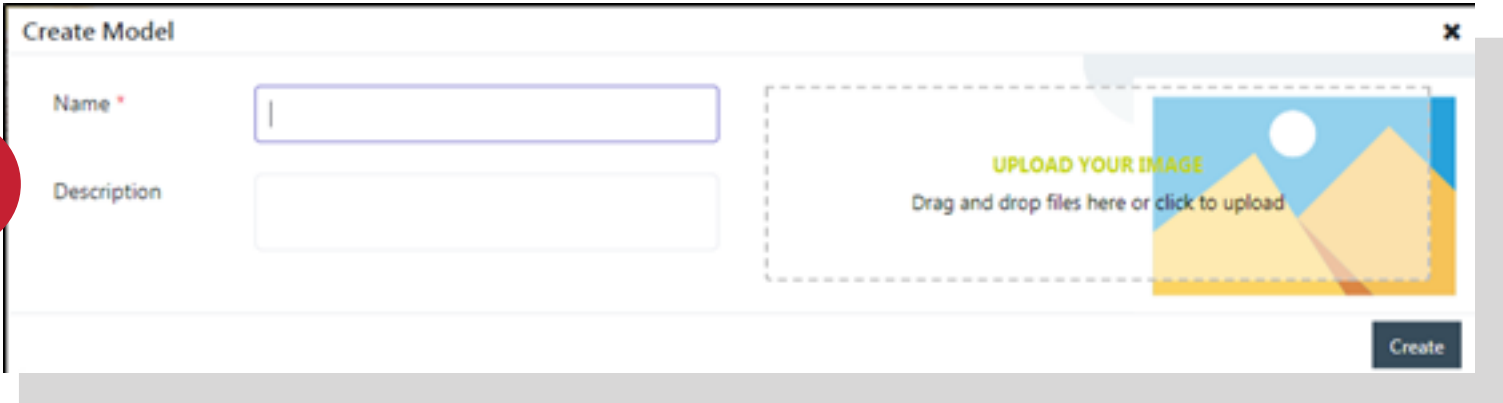
ALGORITHMS	INPUT PATRAMETERS
Two class naïve bayes	Sensitive class
Multi class naïve bayes	Alpha (Laplace Correlation)
Two class K-nearest neighbor	Sensitive class
Multi class support vector machine	Cost function 'C', Gamma value, Kernel

PROCEDURE TO ADD TEXT CLUSTERING MODEL

Create Model:

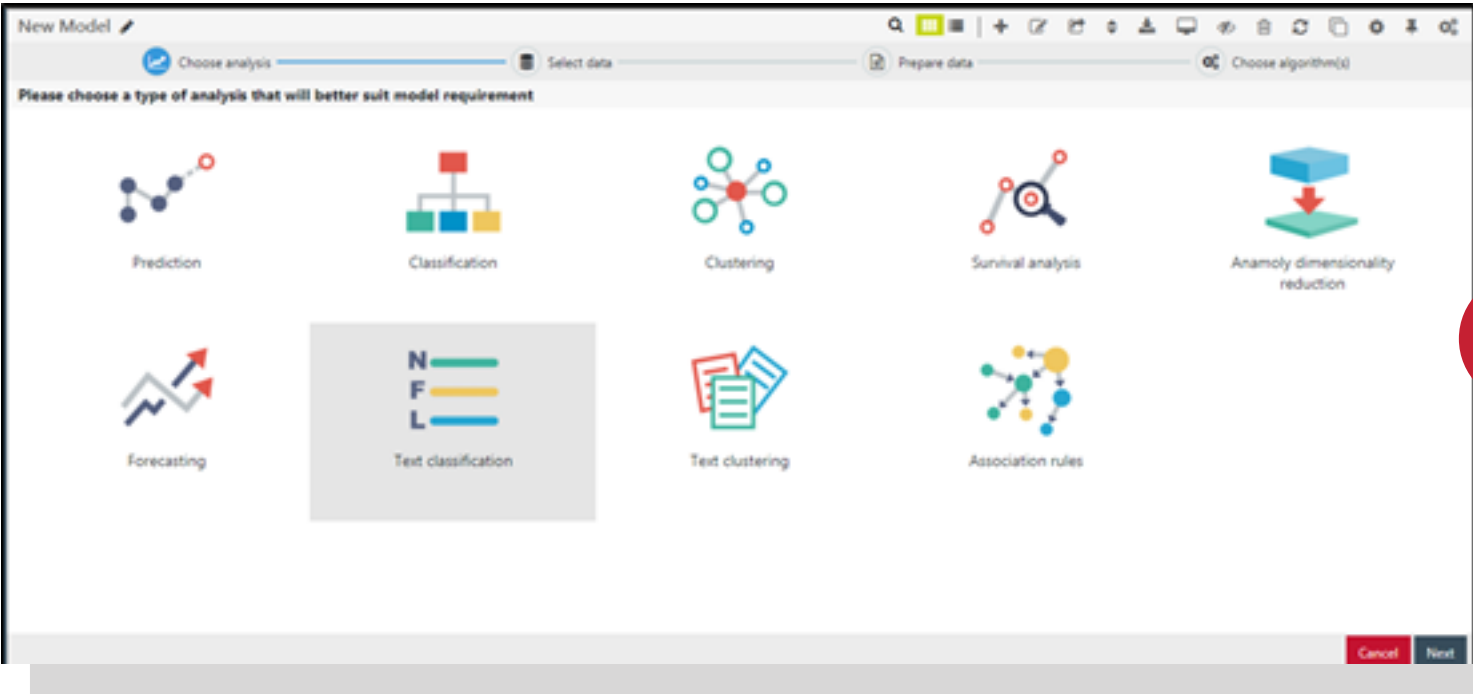
Associate a model with a name, description and if required identify it with an image to view as a thumb nail icon.

1



The 'Create Model' form contains the following elements:

- Name ***: A text input field.
- Description**: A text input field.
- Image Upload**: A dashed box with the text 'UPLOAD YOUR IMAGE' and 'Drag and drop files here or click to upload'. A thumbnail image of a mountain is shown.
- Create**: A button at the bottom right.



The 'New Model' screen shows a progress bar with four steps: 'Choose analysis' (active), 'Select data', 'Prepare data', and 'Choose algorithm(s)'. Below the progress bar, it says 'Please choose a type of analysis that will better suit model requirement'. There are ten analysis options, each with an icon and a label:

- Prediction
- Classification
- Clustering
- Survival analysis
- Anomaly dimensionality reduction
- Forecasting
- Text classification
- Text clustering
- Association rules

At the bottom right, there are 'Cancel' and 'Next' buttons.

2

Choose Analysis:

Select Text Clustering analysis from the list of options.

Choose existing datasource:
Select the database to add data and the preview will be shown. Option to add a new datasource is also available.

3

Choose existing datasource

Available datasources

Search...

e-copy

new

nr

test

Inmemory

HRData

adbfudf

HelloWorld_DS-SQL

cfhsafds

sss

fr

MBA-copy

MBA-copy(2)

ghf

sf

bgh

rtrey

Infymontly

+ New Datasource

Preview - HRData

WorkAccident

Left

Department

Promotion

SalaryType

Employee

0	1	sales	0	low	1000001
0	1	sales	0	medium	1000002
0	1	sales	0	medium	1000003
0	1	sales	0	low	1000004
0	1	sales	0	low	1000005
0	1	sales	0	low	1000006
0	1	sales	0	low	1000007
0	1	sales	0	low	1000008
0	1	sales	0	low	1000009
0	1	sales	0	low	1000010
0	1	sales	0	low	1000011
0	1	sales	0	low	1000012
0	1	sales	0	low	1000013
0	1	sales	0	low	1000014
0	1	sales	0	low	1000015
0	1	sales	0	low	1000016
0	1	sales	0	low	1000017
0	1	sales	0	low	1000018

Add

New Model

Choose analysis

Select data

Prepare data

Choose algorithm(s)

Please prepare data with required column for model preparation

Column name	Data type	Unique id	Comment
<input checked="" type="checkbox"/> WorkAccident	varchar	<input type="radio"/>	<input checked="" type="radio"/>
<input type="checkbox"/> Left	varchar	<input type="radio"/>	<input type="radio"/>
<input checked="" type="checkbox"/> Department	varchar	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> Promotion	varchar	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> SalaryType	varchar	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> EmployeeId	int	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> LastEvaluation	float	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> AverageMonthlyHours	int	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> SatisfactoryLevel	float	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> TimeSpent	int	<input type="radio"/>	<input type="radio"/>
<input type="checkbox"/> NumberProject	int	<input type="radio"/>	<input type="radio"/>

Total columns: 11, Selected columns: 2

Clear all

Preview limited to 50 rows

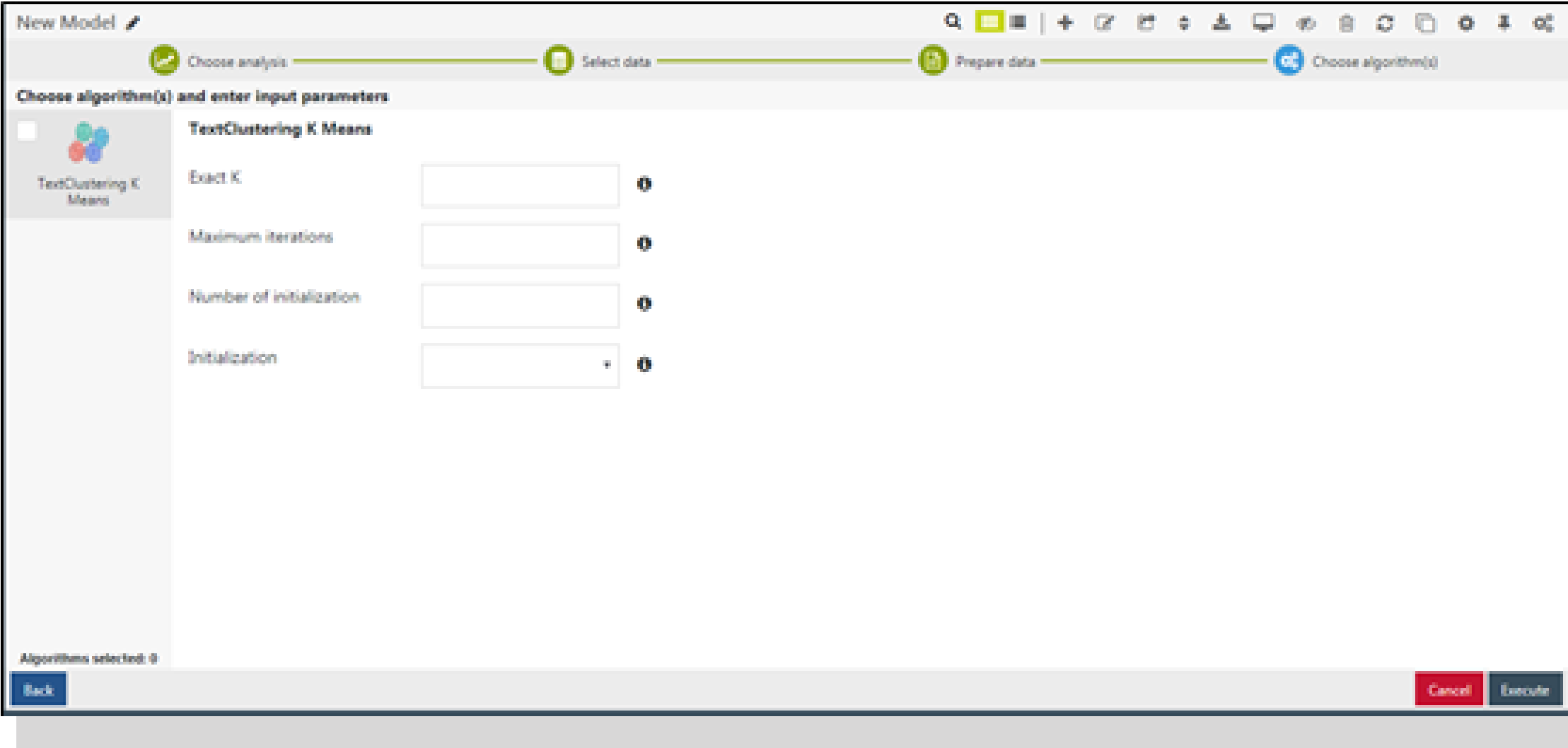
Back

Cancel

Next

4

Prepare Data: Prepare the data with the required columns.
Unique id: This column is mandatory. Select the unique label to identify it as unique comment.
Comment: This column is mandatory. Select the column that has text comments.



5

Choose Algorithm(s):

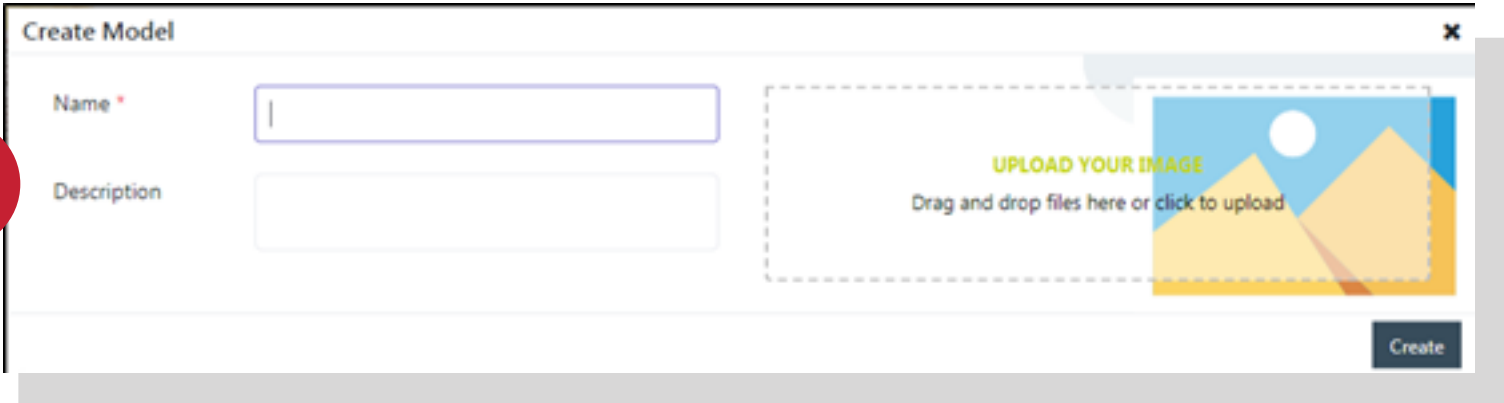
Choose the type of algorithm(s) from the available algorithms for calculating the values for the model and specify the input parameters.

ALGORITHMS	INPUT PATRAMETERS
TextClustering K Means	Exact K, Maximum number of iterations, Number of initialization, Initialization

PROCEDURE TO ADD ASSOCIATION RULES MODEL

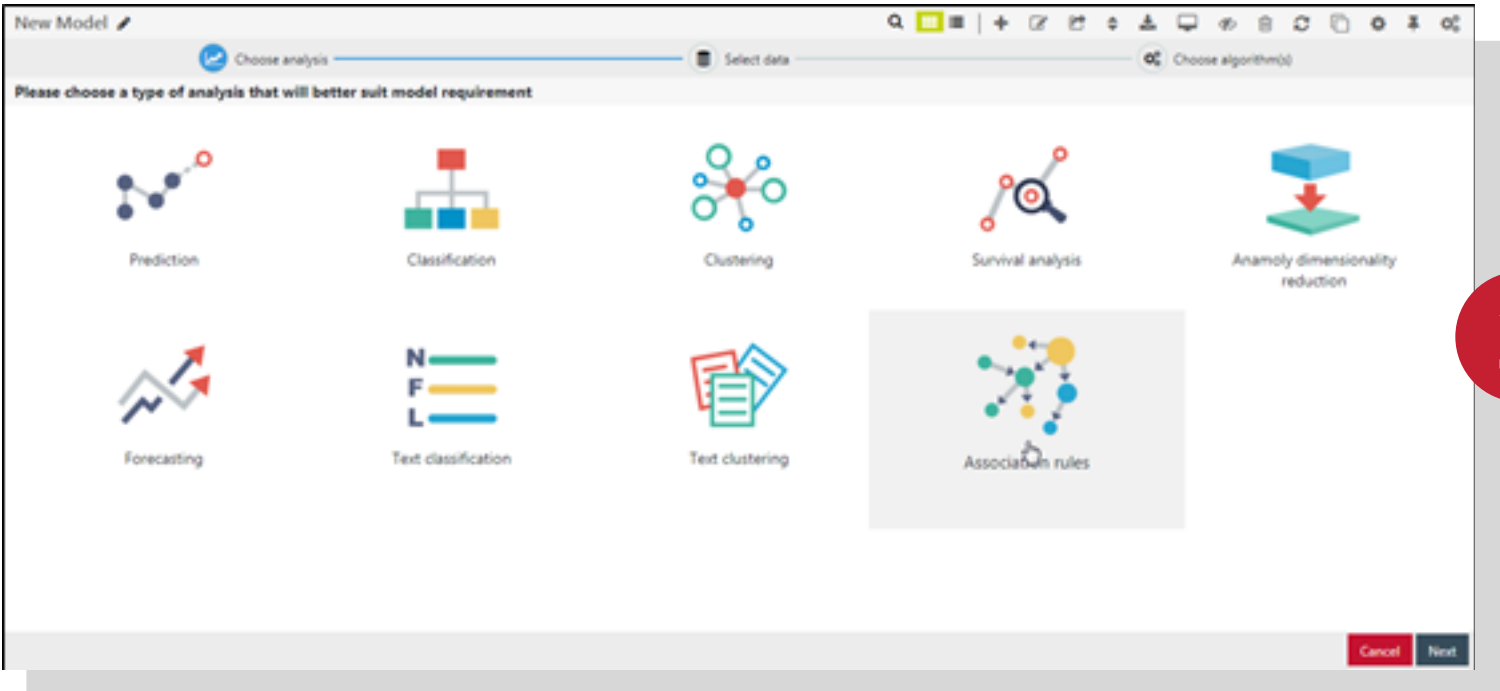
Create Model:
Associate a model with a name, description and if required identify it with an image to view as a thumb nail icon.

1



The 'Create Model' dialog box contains the following elements:

- Name ***: A text input field.
- Description**: A text input field.
- Image Upload Area**: A dashed box containing the text 'UPLOAD YOUR IMAGE' and 'Drag and drop files here or click to upload', with a placeholder image of a landscape.
- Create**: A button at the bottom right.



The 'New Model' dialog box features a toolbar with icons for 'Choose analysis', 'Select data', and 'Choose algorithm(s)'. Below the toolbar, it says 'Please choose a type of analysis that will better suit model requirement'. The main area displays several analysis options with icons:

- Prediction
- Classification
- Clustering
- Survival analysis
- Anomaly dimensionality reduction
- Forecasting
- Text classification
- Text clustering
- Association rules** (highlighted with a grey background)

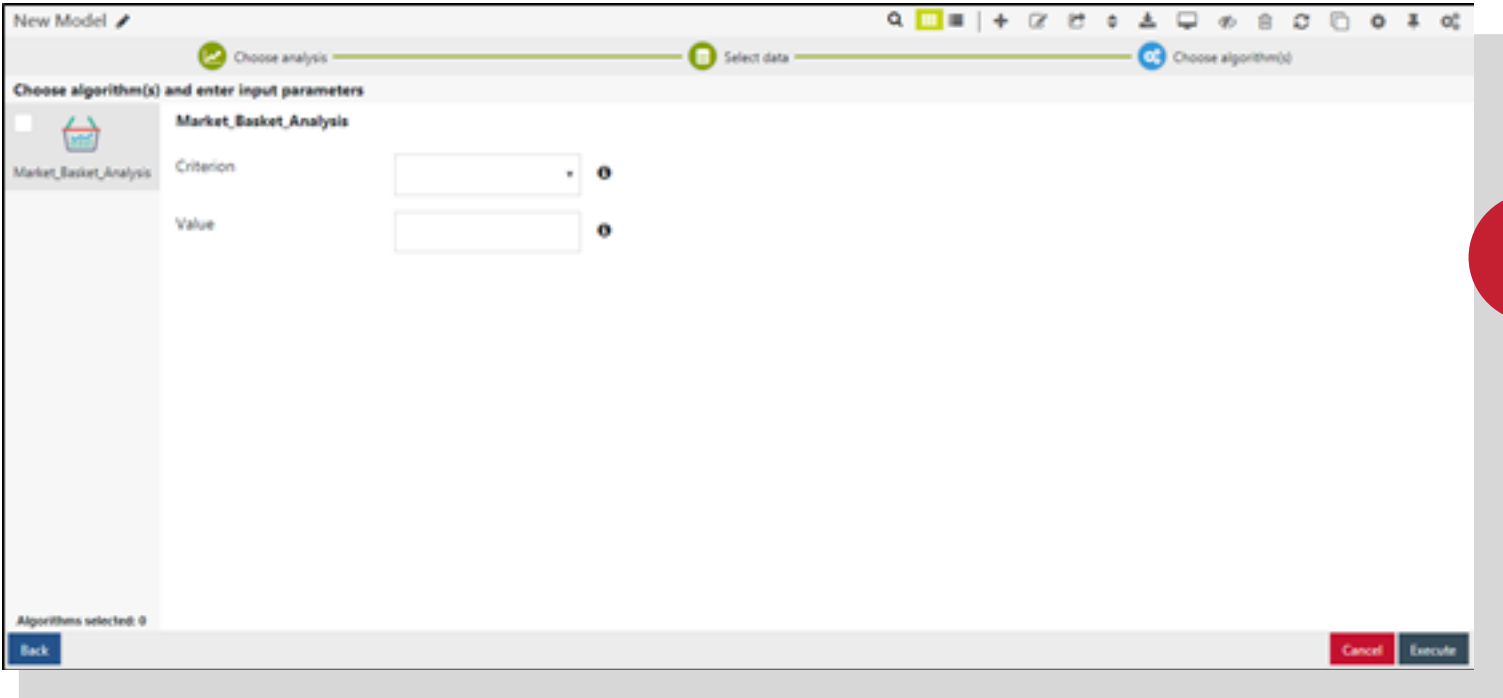
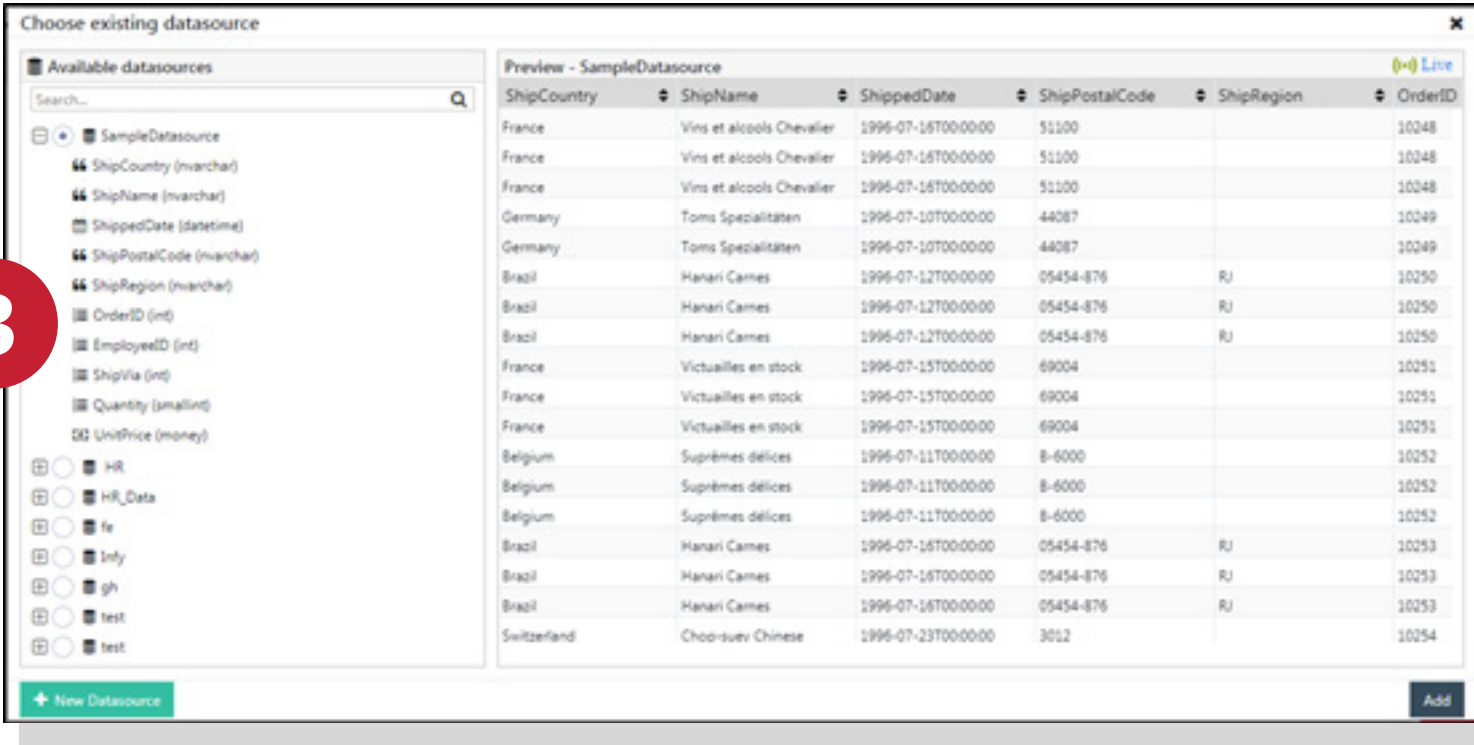
At the bottom right, there are 'Cancel' and 'Next' buttons.

2

Choose Analysis:
Select Association rules analysis from the list of options.

Choose existing datasource:
Select the database to add data and the preview will be shown. Option to add a new datasource is also available.

3



4

Choose algorithms:
Choose the type of algorithm(s) from the available algorithms for calculating the values for the model and specify the input parameters.

ALGORITHMS	INPUT PATRAMETERS
Market_Basket_Analysis	Criterion, Value