Lift Chart

SQL Server 2005

You can view two types of charts in the **Lift Chart** tab of the **Mining Accuracy Chart** tab of Data Mining Designer: a lift chart and a profit chart. A lift chart compares the accuracy of the predictions of each model, while a profit chart displays the theoretical increase in profit that is associated with using each model.

Use the **Chart Type** list to select the type of chart you want. When you select **Profit Chart** from the list, the **Profit Chart Setting** dialog box automatically opens. (This dialog box also opens when you click **Settings**.) You can use this dialog box to set the parameters that define the profit chart.

Only mining models that contain discrete predictable attributes can be compared in a lift chart. The **Mining Accuracy Chart** tab cannot be used with time series models or with models that have continuous predictable attributes.

For more information: Mining Accuracy Chart Tab How-to Topics[1], Column Mappings (Lift Chart)[2], Validating Data Mining Models[3]

Chart Types

The following sections describe each chart type in more detail:

- Lift Chart Type
- Profit Chart Type

Lift Chart Type

The **Lift Chart** tab displays a graphical representation of the change in *lift* that a mining model causes. For example, consider the case in which the Marketing department at Adventure Works Cycles wants to create a targeted mailing campaign. From past campaigns, they know that a 10 percent response rate is typical. They have a list of 10,000 potential customers stored in a table in the database. Based on the typical response rate, they can expect 1,000 customers to respond.

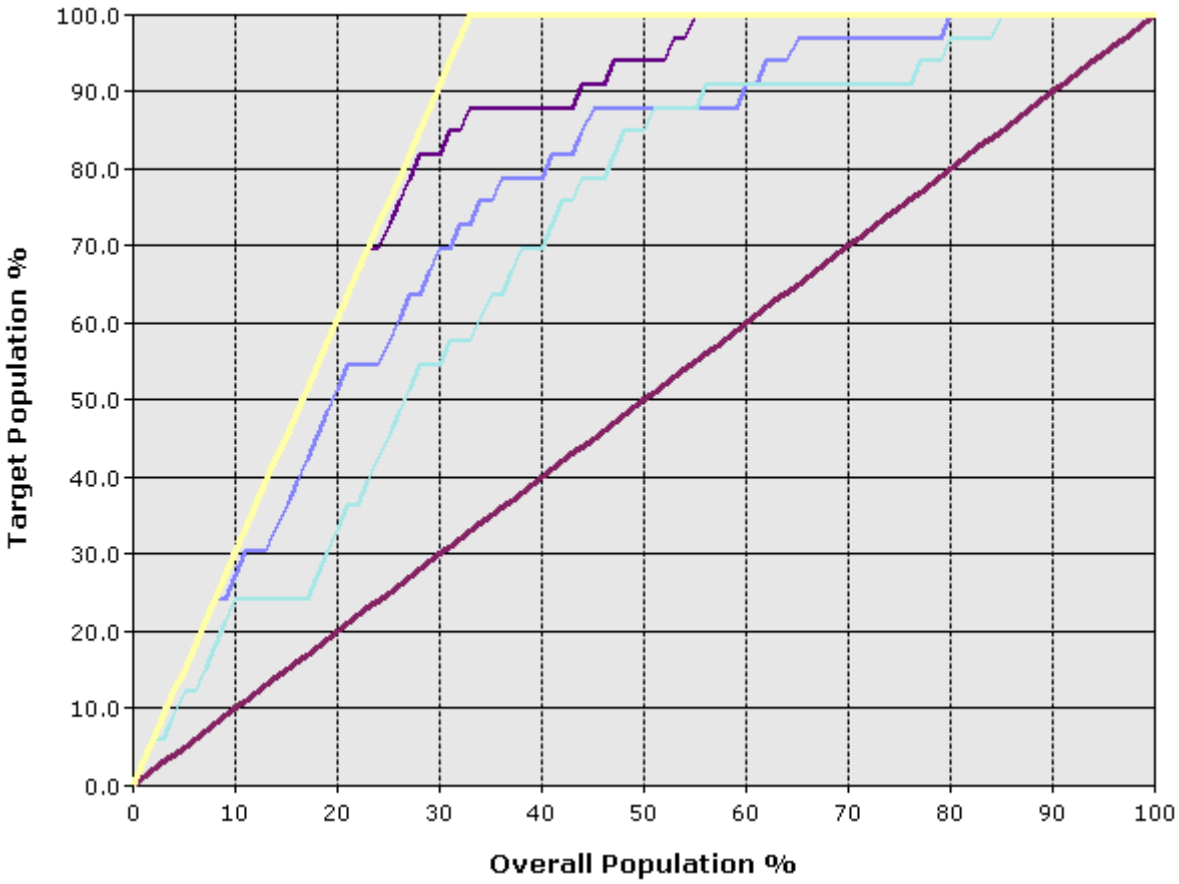
Consider also that the budget for the project is less than the amount of money that is required to reach all 10,000 customers in the database. Based on the budget, they can afford to mail an advertisement to only 5,000 customers. The Marketing department has two choices:

- Randomly select 5,000 customers to target
- Use a mining model to target the 5,000 customers who are most likely to respond

If the company randomly selects 5,000 customers, they can expect to reach only 500 of the estimated 1,000 positive responses, because only 10 percent of those who are targeted typically respond. This scenario is what the random line in the lift chart represents. However, if the Marketing department uses a mining model to target their mailing, they can expect a greater response rate because they can target those customers who are most likely to respond. If the model was perfect, meaning that it could create predictions that are never wrong, the company could expect to reach all 1,000 of the estimated responses by mailing to 1,000 potential customers who the model recommends. This scenario is represented by the ideal line in the lift chart. Reality is that the mining model most likely falls between these two extremes, between a random guess and an ideal or perfect prediction. Any improvement in response from the random guess is considered to be lift.

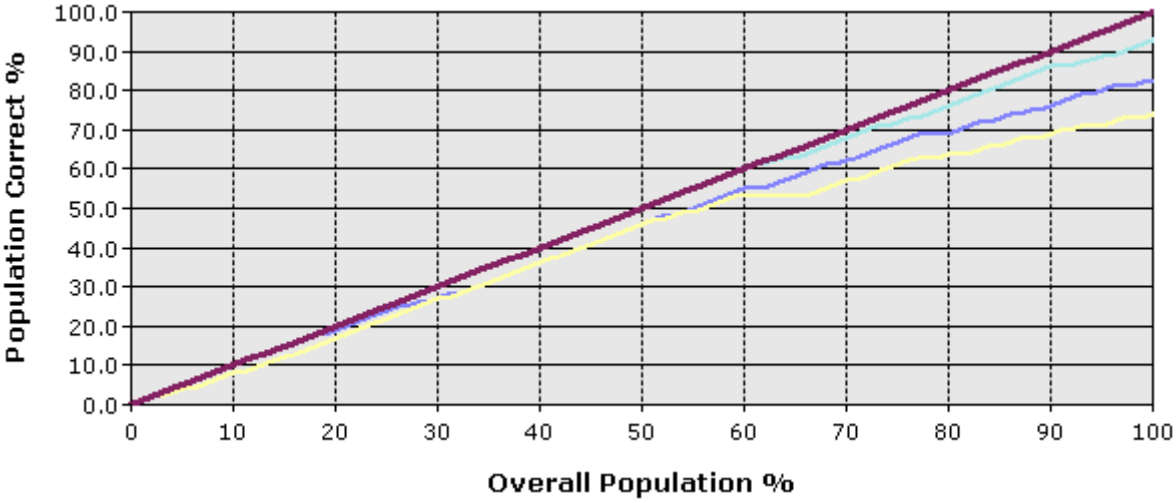
You can create two types of charts, one in which you specify a state of the predictable column, and one in which you do not specify the state.

If you specify the state of the predictable column, you create the type of chart shown in the following diagram.



The X-axis of the chart represents the percentage of the test dataset that is used to compare the predictions. The Y-axis of the chart represents the percentage of values that are predicted to be the specified state. In the chart, the red line represents the random line, and the yellow line represents the ideal model.

If you do not specify the state of the predictable column, you create the type of chart shown in the following diagram.



The X-axis is the same as in the chart with the predictable column specified, but the Y-axis now represents the percentage of predictions that are correct. In this chart, the red line represents the ideal model.

When you switch between the **Column Mapping** tab and the **Lift Chart** tab, the chart is updated to reflect any changes made in the column mappings.

Back to Top

Profit Chart Type

A profit chart displays the estimated profit increase that is associated with using a mining model to determine which customers a company should contact in a business scenario. The Y-axis of the chart represents the profit, while the X-axis represents the percentage of the population that the company contacted. A typical profit chart shows an increase in profits up to a point, after which profits decrease as more of the population is contacted.

Use the **Chart Type** list to display the profit chart. After you select **Profit Chart**, the **Profit Chart Settings** dialog box opens. You can use this dialog box to set the parameters that define the profit chart. The following list describes the parameters that you can set.

Population

The number of cases in the dataset that is being used to create the lift chart. For example, the number of potential customers.

Fixed Cost

The fixed cost that is associated with the business problem. If this were for a targeted mailing solution, the cost would not depend on variables such as the number of telephone calls made or the number of promotional mailings sent.

Individual Cost

Costs that are in addition to the fixed cost, that can be associated with each customer contact. For example, promotional mailings or telephone calls.

Revenue Per Individual

The amount of revenue that is associated with each successful sale.

You can also open the **Profit Chart Settings** dialog box by clicking **Settings** in the **Lift Chart** tab.

The profit chart contains a gray vertical line, which you can move by clicking a location in the chart. The **Mining Legend** displays a score, the population correct and the predict probability that are associated with the location of the gray line on the chart. If you select the maximum point of profits in the chart by using the gray line, you can use the predict probability value to determine a probability threshold for contacting a customer.

For example, if the peak of the profit curve is at 55 percent of the population and the associated predict probability is 20 percent, this indicates that to achieve maximum profits you should only contact those customers whose response is predicted with a 20 percent or greater chance.

Back to Top

See Also

Concepts

Column Mappings (Lift Chart)[2]
Data Mining Concepts[4]
Working with Data Mining[5]
Validating Data Mining Models[3]

Other Resources

Mining Accuracy Chart Tab How-to Topics[1]

Help and Information

Getting SQL Server 2005 Assistance [6]

Links Table

[1] *http://technet.microsoft.com/en-us/library/ms174767(v=SQL.90).aspx*

[2] *http://technet.microsoft.com/en-us/library/ms174947(v=SQL.90).aspx*

[3] *http://technet.microsoft.com/en-us/library/ms174493(v=SQL.90).aspx*

[4] *http://technet.microsoft.com/en-us/library/ms174949(v=SQL.90).aspx*

[5] *http://technet.microsoft.com/en-us/library/ms174861(v=SQL.90).aspx*

[6] *http://technet.microsoft.com/en-us/library/ms166016(v=SQL.90).aspx*

Community Content