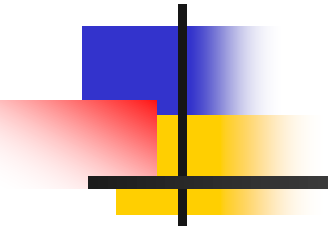

Demo LR Advanced Topics

Dr. Goutam Chakraborty

Outline of Demo

- Adding Transformation
  - Remove the connection between the Data Partition node and the Impute node
  - Drag the **Transform Variables** node into the diagram workspace.
  - Connect the **Data Partition** node to the **Transform Variables** node.
  - Connect the **Transform Variables** node to the **Impute** node.
  - Select **Variables** ⇨ ellipsis button
  - Select all inputs with **Gift** in the name
  - Select **Explore**. Look at distributions
  - Deselect the two inputs with **GiftTime** in their names
  - Select **Method** ⇨ **Log** for one of the remaining selected inputs with **Gift** in the name
  - Run the **Transform** node and view results

Outline of Demo (Continued)

- Regression (optimal) with transformed inputs
  - Run the Regression (optimal) node and view results
  - Select **View** ⇨ **Model** ⇨ **Iteration Plot**.

Performance of Regression Models

- No Selection:
  - Number of Parameters = 86, ASE (V) = 0.2438, Misclassification (V) = 0.4319
- Stepwise:
  - Number of Parameters = 13, ASE (V) = 0.2423, Misclassification (V) = 0.4245
- Optimal:
  - Number of Parameters = 15, ASE (V) = 0.2417, Misclassification (V) = 0.4319
- Optimal with Transformation
  - Number of Parameters = 5, ASE (V) = **0.2407**, Misclassification (V) = **0.4212**

Outline of Demo (Continued)

- Using **Replacement** node to facilitate combining input levels of a categorical input.
  - Remove the connection between the Transform Variables node and the Impute node.
  - Click the **Modify** tab.
  - Drag a **Replacement** node into the diagram workspace.
  - Connect the **Transform Variables** node to the **Replacement** node.
  - Connect the **Replacement** node to the **Impute** node
  - In the Interval Variables property group, select **Default Limits Method** ⇨ **None**
  - In the Class Variables property group, select **Replacement Editor** ⇨ ellipsis button

Outline of Demo (Continued)

- The levels of StatusCat96NK are consolidated as follows:
- Levels A and S (active and star donors) indicate consistent donors and are grouped into a single level, A.
  - Enter A as the Replacement level for StatusCat96NK levels A and S.
- Levels F and N (first-time and new donors) indicate new donors and are grouped into a single level, N.
  - Enter N as the Replacement level for StatusCat96NK levels F and N.
- Levels E and L (inactive and lapsing donors) indicate lapsing donors and are grouped into a single level L.
  - Enter L as the Replacement level for StatusCat96NK levels L and E.
- Run the Replacement node and view the results.
- Select **View ⇨ Model ⇨ Iteration Plot**

Performance of Regression Models

- No Selection:
  - Number of Parameters = 86, ASE (V) = 0.2438, Misclassification (V) = 0.4319
- Stepwise:
  - Number of Parameters = 13, ASE (V) = 0.2423, Misclassification (V) = 0.4245
- Optimal:
  - Number of Parameters = 15, ASE (V) = 0.2417, Misclassification (V) = 0.4319
- Optimal with Transformation
  - Number of Parameters = 5, ASE (V) = **0.2407**, Misclassification (V) = **0.4212**
- Optimal with Transformation and Consolidation of Categorical Input
  - Number of Parameters = 5, ASE (V) = **0.2407**, Misclassification (V) = **0.4212**

Polynomial Regression Demo

- Right-click the **Regression (optimal)** node and select **Copy** from the menu.
- Right-click the diagram workspace and select **Paste** from the menu. Rename the new **as Polynomial Regression**.
- Connect the **Polynomial Regression** node to the **Impute** node
- Follow these steps to explore a full two-factor stepwise selection process:
  - Select **Two-Factor Interaction** ⇨ **Yes** in the Polynomial Regression Properties panel.
  - Select **Polynomial Terms** ⇨ **Yes** in the Polynomial Regression Properties panel.
  - Run and view results (*warning*: it will take some time to churn through all quadratic terms and all 2 way interactions)
- Do self study