Models/Results

**Standard Scalar, Train-Test Split Logistic Regression on Numerical Columns**

A picture containing application

Description automatically generatedTable

Description automatically generated

**Standard Scalar, Train-Test Split Logistic Regression on WoE Encoded Columns**

Chart

Description automatically generatedTable

Description automatically generated

**DataFrame mapping technique used in 2nd model:**

**Table

Description automatically generated**

Note: WoE labels for each age (not grouped)

Target Salary Groupings used for WoE estimates (grouped):

["15k-26k","26k-40k","40k-55k","55k-67k","67k-81k","81k-100k", "100k-120k","120k-135k","135k-150k"]

**Logistic Regression on one-hot encoded variables. K-fold split**

**A picture containing text

Description automatically generated**

**Logistic Regression on one-hot encoded variables, Train-Test split**

**Table

Description automatically generated**

**Dataframe used for both models:**

**A sheet of music

Description automatically generated with medium confidence**

Note: One-Hot Encoding for age (grouped):

["18-25","26-35","36-45","46-55","55+"]

Target Salary Groupings used for one-hot encoding (grouped):

["15k-26k","26k-40k","40k-55k","55k-67k","67k-81k","81k-100k", "100k-120k","120k-135k","135k-150k"]

**Dataframe Mapping technique used for IV/WoE mapping in following models:**

Table

Description automatically generated

Table

Description automatically generated

Table

Description automatically generated

Final Map before reduction:

Table

Description automatically generated

**Results of WoE encoding log regression *without* standard scaling , K-fold split**

Note: columns

A picture containing table

Description automatically generated

**Results of WoE encoding log regression *with* standard scaling, K-fold split**

Text

Description automatically generated

**DataFrame used on both WoE models:**

Table

Description automatically generated

**Results of IV Encoding *without* standard scaling**

Table

Description automatically generated with low confidence

**Results of IV encoding *with* standard scaling**

Text

Description automatically generated with medium confidence

**Reduced DataFrame used in both IV models:**

Table

Description automatically generated