BANK MARKETING ANALYSIS

NOORPINDER KAUR

15th Dec, 2023

Big Data and Predictive Analysis

Resmi Ann Thomas

Summary

In this research, our primary objectives were to explore the factors influencing term deposit subscriptions, gain insights into customer behavior, and develop predictive models for targeted marketing strategies using the Bank Marketing dataset spanning May 2008 to November 2010. The dataset comprises 21 variables encompassing demographic, economic, and campaign-related attributes, with 41,188 rows.

Our methodology involved thoughtful steps such as employing SMOTE for sampling to address class imbalances, utilizing Variable Clustering to understand relationships and eliminate multicollinearity, rejecting irrelevant variables, and employing Data Partitioning for model evaluation.

We navigated potential skewness and outliers using Cap and Floor with standard deviations from the mean as default limits and log transformation. The data underwent Dummy Re-coding for categorical variables. Decision trees, including Maximal Tree, Misclassification Tree, and Average Squared Error Tree, were employed to identify key variables influencing term deposit subscriptions. Regression nodes provided insights into the relationship between variables.

Our Neural Network analysis revealed that the Data Partitioning Neural Network with 50 iterations and 3 hidden units emerged as the best model. The most influential variables identified through the Stepwise Regression model were "LOG_REP_duration", "LOG_REP_campaign", "REP_education", "REP_job", "REP_month", "contact", "default", "marital", and "poutcome".

In conclusion, our research not only identifies critical variables affecting subscription decisions but also emphasizes the effectiveness of the Data Partitioning Neural Network and Stepwise Regression in predicting term deposit subscriptions. These findings equip marketers with actionable insights for crafting targeted strategies to enhance subscription rates in banking campaigns.

TABLE OF CONTENT

TABLE OF CONTENT **3**
INTRODUCTION **6**
OVERVIEW OF THE DATASET 6
IMPORTANCE 6
PURPOSE OF THE RESEARCH 7
SCOPE OF THE ANALYSIS 7
BACKGROUND CONTEXT 9
DATA EXPLORATION **10**
FILE IMPORT 10
DATA DESCRIPTION 10
DATA VALIDATION 13
Summary statistics 13
Sampling 15
Variables Clustering (checking on correlation) 17
Rejected Variables 19
DATA PARTITIONING 20
DUMMY RE-CODING 21
MODELING AND PREDICTIVE ANALYSIS **24**
DECISION TREES 24
Maximal Decision Tree Model 25
The Story 28
Misclassification Decision Tree Models 28
Average Square Error Trees 31
Decision Trees Comparison 33
The Best Misclassification Decision Tree 34
The Story 35
The Best ASE Decision Tree 35
The Story 36
OUTLIERS AND SKEWNESS 37
Capping and Flooring 37
Logarithmic Transformation 40
REGRESSION 43
Full regression 44
Type 3 Analysis of Effects. 44
Average Squared Error (ASE) 46
Misclassification Rate 46
Odds Ratio Estimates 47
Findings 49
Stepwise and Forward Regression 49
Type 3 Analysis of Effects: 52
Average Squared Error (ASE) 53
Misclassification Error 54
Odds Ratio Estimates 55
Findings 56

Backward Regression 57
Average Squared Error (ASE) 58
Misclassification Rate 58
Type 3 Analysis of Effects 59
Odds Ratio Estimates 60
Findings 61
NEURAL NETWORK 62
Data Partition Neural Networks 65
Data Partition Neural Network Average Squared Error, 50 iterations, 3 hidden layers 66
Data Partition Neural Network Average Squared Error, 50 iterations, 4 hidden layers 68
Data Partition Neural Network Average Squared Error, 50 iterations, 5 hidden layers 69
Data Partition Neural Network Average Squared Error, 50 iterations, 6 hidden layers 71
Data Partition Neural Network Average Squared Error, 100 iterations, 3 hidden layers 72
Data Partition Neural Network Average Squared Error, 100 iterations, 4 hidden layers 74
Data Partition Neural Network Average Squared Error, 100 iterations, 5 hidden layers 75
Data Partition Neural Network Average Squared Error, 100 iterations, 6 hidden layers 76
Cap & Floor NN 78
Cap and Floor Neural Network Average Squared Error 50 iterations, 3 hidden units 79
Cap and Floor Neural Network Average Squared Error 50 iterations, 4 hidden units 81
Cap and Floor Neural Network Average Squared Error 50 iterations, 5 hidden units 82
Cap and Floor Neural Network Average Squared Error 50 iterations, 6 hidden units 84
Cap and Floor Neural Network Average Squared Error 100 iterations, 3 hidden units 85
Cap and Floor Neural Network Average Squared Error 100 iterations, 4 hidden units 86
Cap and Floor Neural Network Average Squared Error 100 iterations, 5 hidden units 88
Cap and Floor Neural Network Average Squared Error 100 iterations, 6 hidden units 89
Transform Variables Neural Network 91
Transform Variables Neural Network 50 iterations, 3 hidden units 93
Transform Variables Neural Network 50 iterations, 4 hidden units 94
Transform Variables Neural Network 50 iterations, 5 hidden units 96
Transform Variables Neural Network 50 iterations, 6 hidden units 97
Transform Variables Neural Network 100 iterations, 3 hidden units 99
Transform Variables Neural Network 100 iterations, 4 hidden units 100
Transform Variables Neural Network 100 iterations, 5 hidden units 102
Transform Variables Neural Network 100 iterations, 6 hidden units 104
Stepwise Regression Neural Network 105
Stepwise Regression Neural Network 50 iterations, 3 hidden units 107
Stepwise Regression Neural Network 50 iterations, 4 hidden units 108
Stepwise Regression Neural Network 50 iterations, 5 hidden units 110
Stepwise Regression Neural Network 50 iterations, 6 hidden units 111
Stepwise Regression Neural Network 100 iterations, 3 hidden units 113
Stepwise Regression Neural Network 100 iterations, 4 hidden units 114
Stepwise Regression Neural Network 100 iterations, 5 hidden units 116
Stepwise Regression Neural Network 100 iterations, 6 hidden units 118
MODEL COMPARISON 120
Data Partition Neural Network Results, 50 iterations, 3 hidden units 125
CONCLUSION 130
KEY FEATURES OF THE RESEARCH 130

The Maximal Tree 130
The Best Misclassification Tree 130
The Best Regression Model: Stepwise Regression 131
The Data Partitioning Neural Network Average Squared Error, 50 iterations, 3 hidden units 132
The Transform Variables Neural Network 132
Cap and Floor Neural Network 132
Stepwise Regression Neural Network 132
Model Comparison Node 133
ACTIONABLE INSIGHTS 133
Decision Trees 133
Stepwise Regression (Best Model) 134
Data Partitioning Neural Network : 50 Iterations 3 Hidden Units 134
RECOMMENDATIONS 136
CITATION AND PAST USAGE 137
REFERENCES 138
APPENDIX 139

INTRODUCTION

OVERVIEW OF THE DATASET

The Bank Marketing dataset, taken from the UCI Machine Learning Repository, provides a deep insight into the complex world of marketing strategies implemented by a Portuguese banking institution. Carefully gathered and including a vast number of customer and campaign-specific attributes, this dataset is a trove for understanding the dynamics of telemarketing initiatives to promote term deposit subscriptions. The significance of the dataset lies in shedding light on the multiple factors that affect customer behavior and subscription decision making in the banking sector..

This dataset consists of 45,211 rows and 18 columns, organized carefully in chronological order from May 2008 to November 2010. The research associated with this dataset, conducted by S. Moreau, et al; 2014, focuses on a data-driven approach to predict the success of bank telemarketing. This study provides a framework for exploring the nuances of the data set and insights into the performance of telephone marketing campaigns.

The dataset, which includes a diverse set of variables, covers a variety of demographic attributes, economic indicators and communication characteristics. Variables such as age, marital status, education level, occupation as well as financial indicators such as average annual balance and loans in force provide a comprehensive view of the customer base. Moreover, characteristics such as number of contacts, call duration, and results of previous marketing promotions significantly help in understanding the effectiveness of telephone marketing strategies in attracting term deposit subscribers. Thus, this dataset provides an invaluable opportunity to decipher the relationship between these drivers and the final decision of customers to subscribe or not to a term deposit offered by a bank

IMPORTANCE

Predicting Term Deposit Subscriptions. The main objective is to develop a model to predict whether customers will subscribe to term deposits. Such forecasting is extremely important for banks, as the subscription to these deposits largely generates their income.

Understanding Telemarketing Impact. Analyze the Bank Marketing dataset to find out the effect of telephone marketing on the bank's success in attracting customers to sign up for term deposits.

Optimizing Marketing Strategies. Identifying the factors that drive customer decisions allows banks to optimize their marketing efforts. This optimization enables them to target potential subscribers more effectively, saving resources and increasing efficiency.

Enhancing Customer Relations. In addition to the financial benefits, understanding customer behavior contributes to the development of more customer-centric strategies. This understanding allows banks to build stronger relationships with customers, resulting in increased customer loyalty and satisfaction.

PURPOSE OF THE RESEARCH

The primary objectives of the research endeavor:

- To explore the key factors influencing term deposit subscriptions based on the provided dataset.
- To develop insights into customer behavior and characteristics affecting subscription decisions.
- To build predictive models for identifying potential subscribers, aiding in targeted marketing strategies.

SCOPE OF THE ANALYSIS

Temporal Range. The analysis focuses on the Bank Marketing dataset covering the period from May 2008 to November 2010. Such a time interval allows us to assess banks' marketing campaigns and customers' reactions over a certain period.

Features Considered. There are 21 variables in the dataset which are related to different groups:

1. Bank client data:
    1. Age
    2. Occupation
    3. Marital status
    4. Educational level
    5. Has credit in default?
    6. Has a housing loan?
    7. Has a personal loan?
2. Related with the last contact of the current campaign:
    1. Contact communication type
    2. Last contact day of the week

   3. Last contact duration, in seconds
   4. Last contact month of year
3. Social and economic context attributes:
   1. Euro Interbank Offered Rate (3 months)
   2. Employment variation rate (represents the percentage change in the number of employed individuals over a quarter)
   3. Consumer price index (measures the average change in prices of a basket of goods and services consumed by households over time, monthly)
   4. Consumer confidence index (reflects the sentiment or optimism levels of consumers regarding the economy's future prospects, monthly)
   5. Number of employees (the total count of individuals currently employed within an economy, quarterly)
4. Other attributes:
   1. Number of contacts performed during this campaign and for this client (includes last contact)
   2. Number of days that passed by after the client was last contacted from a previous campaign (-1 means client was not previously contacted)
   3. Outcome of the previous marketing campaign:
   4. Number of contacts performed before this campaign and for this client
5. Output variable (desired target):
   1. Has the client subscribed to a term deposit?

The first four groups of client-related attributes were considered for their influence on term deposit subscriptions. The main focus of the analysis is on the last target variable “Has the client subscribed to a term deposit?”, which means whether the customer subscribed to the term deposit or not. Understanding the factors that lead to a positive or negative result on this variable is important for predicting the customer's subscription behavior based on the given attributes.

BACKGROUND CONTEXT

Term deposits are critical to banks as they provide a stable source of funding by investing customers' fixed amounts at interest rates. These deposits significantly increase the liquidity and lending capacity of the bank.

However, telephone marketing campaigns, while successful, are financially challenging due to the cost of managing call centers and extensive awareness-raising. In the absence of targeted approaches, such campaigns often result in high operational costs and lower conversion rates. To address these challenges, banks need accurate customer identification to optimize their performance and ensure that marketing messages are effectively delivered to receptive customers, therefore cutting costs and leveraging campaign efficiency.

DATA EXPLORATION

FILE IMPORT

The dataset utilized in this analysis is provided in a CSV (Comma-Separated Values) format, which facilitates seamless data import and manipulation across various analytical platforms. The CSV file was directly integrated into the SAS (Statistical Analysis System) environment for comprehensive examination and analysis: Add Note - Sample - File Export - Import file in Properties window.

DATA DESCRIPTION

The dataset was carefully explored using the exploratory function (Edit Variables - Explore), which allowed the variables to assess their relative roles and levels in the dataset. The study evaluated the significance and contribution of each variable to the analysis, factors such as distribution, statistical performance and relevance to the study objectives. As a result of the comprehensive evaluation, it was decided not to reject any variable from the analysis at this stage of study. Each attribute including demographic, economic and campaign specific characteristics were considered essential to a complete understanding of banks' marketing campaigns and their impact on term deposit subscriptions. Thus, all variables were retained and deemed necessary for analysis, allowing for a comprehensive examination of their interaction and influence on predicting subscription decisions.

Table 1. Data Dictionary

#	Variable Name	Role	Level	Demographic	Description		Units
	Bank client data						
1	age	Input	Interval	Age			years
2	default	Input	Nominal		Has credit in default? “yes”, “no”, “unknown”		
3	education	Input	Nominal	Education Level	“basic.4y” “basic.6y” “basic.9y” “high.school”	“illiterate” “professional.course” “university.degree” “unknown”	
4	housing	Input	Nominal		Has a housing loan? “yes”, “no”, “unknown”		
5	job	Input	Nominal	Occupation	Type of job:		

					“admin” “blue-collar” “entrepreneur” “housemaid” “management” “retired”	“self-employed” “services” “student” “technician” “unemployed” “unknown”	
6	loan	Input	Nominal		Has a personal loan? “yes”, “no”, “unknown”		
7	marital	Input	Nominal	Marital Status	Marital status:		
					“divorced” “married”	“single” “unknown”	
					Note: “divorced” means divorced or widowed)		
	Related with the last contact of the current campaign						
8	contact	Input	Nominal		Contact communication type:		
					“cellular”	“telephone”	
9	day_of_week	Input	Nominal		Last contact day of the week:		
					"mon" "tue" “wed”	“thu” “fri”	
10	duration	Input	Interval		Last contact duration, in seconds. Important note: this attribute highly affects the output target (e.g., if duration=0 then y='no'). Yet, the duration is not known before a call is performed. Also, after the end of the call y is obviously known. Thus, this input should only be included for benchmark purposes and should be discarded if the intention is to have a realistic predictive model.		Seconds
11	month	Input	Nominal		Last contact month of year:		Month
					“Jan” “feb” “mar”	... “nov” “dec”	
	Social and economic context attributes						
12	euribor3m	Input	Interval		euribor 3 month rate - daily indicator		Euros
13	emp.var.rate	Input	Interval		Employment variation rate - quarterly indicator (numeric)		
14	cons.price.idx	Input	Interval		Consumer price index - monthly indicator (numeric)		
15	cons.conf.idx	Input	Interval		Consumer confidence index - monthly indicator (numeric)		
16	nr.employed	Input	Interval		Number of employees - quarterly indicator (numeric)		

	Other attributes						
17	campaign	Input	Interval		Number of contacts performed during this campaign and for this client (numeric, includes last contact)		
18	pdays	Input	Nominal		Number of days that passed by after the client was last contacted from a previous campaign (numeric)	999 means client was not previously contacted. Because of this number we assume this variable as nominal	
19	poutcome	Input	Nominal		Outcome of the previous marketing campaign: “failure”	“nonexistent” “success”	
20	previous	Input	Interval		Number of contacts performed before this campaign and for this client		
	Output variable (desired target)						
21	y	Target	Binary		Has the client subscribed to a term deposit? “yes”, “no”		

Data structure. The dataset consists of 41,188 rows and 21 columns.

Missing Attribute Values. Several categorical attributes were found to contain missing values labeled "unknown", and it presents an inherent problem of data completeness, requiring careful consideration and possible processing strategies in subsequent stages of analysis to ensure reliable and valid inferences drawn from the dataset.

Variable Categories. Variables in the dataset are categorized into three types:

- Interval variables: 10 attributes
- Nominal variables: 10 attributes
- Target variable: 1 binary attribute denoting the subscription status of clients to a term deposit.

The presence of interval variables representing numeric attributes with meaningful intervals, nominal variables for categorical attributes, and a critical binary target variable reflecting subscription status is a fundamental component that influences subsequent analyses.

DATA VALIDATION

Summary statistics

Table 2. Summary statistics, interval variables (from Bank Marketing Node)

Variable	Minimum	Maximum	Mean	Skewness	Standard Deviation	Median	Kurtosis
campaign	1	56	2.567593	4.762507	2.770014	2	36.9798
previous	0	7	0.172963	3.832042	0.494901	0	20.10882
duration	0	4918	258.285	3.263141	259.2792	180	20.24794
age	17	98	40.02406	0.784697	10.42125	38	0.791312
cons_conf_idx	-50.8	-26.9	-40.5026	0.30318	4.628198	-41.8	-0.35856
cons_price_idx	92.201	94.767	93.57566	-0.23088	0.57884	93.749	-0.83154
euribor3m	0.634	5.045	3.621291	-0.70919	1.734447	4.857	-1.4068
emp_var_rate	-3.4	1.4	0.081886	-0.7241	1.57096	1.1	-1.06263
nr_employed	4964	5228	5167.019	-1.04163	72.17807	5191	-0.01556

Age Distribution. The average age of individuals in the dataset is approximately 40 years, the range from 17 to 98 years, displaying a slightly positively skewed distribution (0.78).

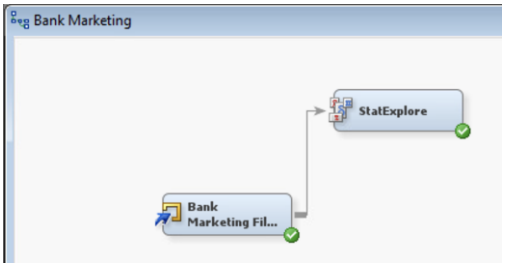
Economic Variables. The range of employment variation rate (emp_var_rate) is from -3.4 to 1.4, and it has a mean equal to 0.08 that indicates a slightly negative skewed distribution (-0.72). The range of Euribor 3-month rate (euribor3m) is from 0.634 to 5.045, and it has mean equals to 3.62 that indicates a slightly negative skewed distribution (-0.7). Also the number of employees (nr_employed) showcase range from 4964 to 5228 and mean equals to 5167.019 displaying a highly negative skewness (-1.04).

Consumer Confidence and Price Indices. The consumer confidence index (cons_conf_idx) and consumer price index (cons_price_idx) have averages around -40.50 (a significantly lower level of consumer confidence) and 93.58, respectively.

Campaign Contacts. On average, clients were contacted approximately 3 times during the marketing campaign, with a maximum of 56 contacts, indicating potential outliers and highly skewed distribution.

Pdays and Previous Contacts. The variable pdays, representing the number of days since the client was last contacted. We assume that this variable is nominal because it has a value 999 that means

Picture 1. Import File and StatExplore



that this particular client has not been contacted before. The average number of contacts performed before this campaign (previous) is approximately 0 with a range from 0 to 7 with a high positive skewness (3.82).

Duration of Last Contact. The average duration of the last contact (duration) with clients was about 258 seconds. The duration varies widely, ranging from 0 to 4918 seconds, displaying a highly positively skewed distribution (3.26).

Table 3. Summary statistics, interval variables (from StatExplore Node)

Variable	Target Level	Minimum	Maximum	Mean	Skewness	Standard Deviation	Median	Kurtosis
campaign	no	1	56	2.633085	4.679777	2.873438	2	35.21286
campaign	yes	1	23	2.051724	3.387153	1.666245	2	19.35519
previous	no	0	7	0.132374	4.047547	0.409199	0	23.81719
previous	yes	0	6	0.492672	2.132789	0.860344	0	5.267404
duration	no	0	4918	220.8448	3.668784	207.0963	163	31.21465
duration	yes	37	4199	553.1912	1.837301	401.1719	449	6.318574
age	no	17	95	39.91119	0.652875	9.898132	38	0.363927
age	yes	17	98	40.91315	0.999425	13.83748	37	0.671076
cons_conf_idx	no	-50.8	-26.9	-40.5931	0.250342	4.391155	-41.8	-0.47939
cons_conf_idx	yes	-50.8	-26.9	-39.7898	0.259931	6.139668	-40.4	-0.78703
cons_price_idx	no	92.201	94.767	93.60376	-0.238	0.558993	93.918	-0.8243
cons_price_idx	yes	92.201	94.767	93.35439	0.119767	0.67664	93.2	-1.05428
euribor3m	no	0.634	5.045	3.811491	-0.94332	1.638187	4.857	-1.00919
euribor3m	yes	0.634	5.045	2.123135	0.890782	1.742598	1.266	-1.08895
emp_var_rate	no	-3.4	1.4	0.248875	-0.9086	1.482932	1.1	-0.75163
emp_var_rate	yes	-3.4	1.4	-1.23345	0.556589	1.623626	-1.8	-0.99048
nr_employed	no	4964	5228	5176.137	-1.20599	64.52434	5196	0.528275
nr_employed	yes	4964	5228	5095.201	0.282066	87.40568	5099	-1.19989

StatExplore shows that we have the following skewed variables: “campaign”, “previous” and “duration” with target levels “yes” and “no” and “nr.employed” with target level “no”.

Sampling

Our dataset had a significant imbalance in our target binary variable “Has the client subscribed to a term deposit?”, with 36,548 instances labeled “no” and only 4,640 labeled “yes”. To fix this, we used the Synthetic Minority Over-sampling Technique (SMOTE), a technique that balances such imbalances.

Picture 2. Summary statistics for Class Targets

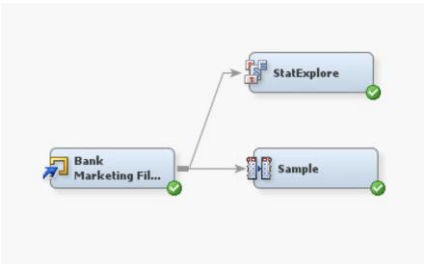
Summary Statistics for Class Targets

Data=DATA

Variable	Numeric Value	Formatted Value	Frequency Count	Percent	Label
y	.	no	36548	88.7346	y
y	.	yes	4640	11.2654	y

We added a “Sample” node after importing the data. This step allowed us to create a more balanced dataset for analysis. With SMOTE, we made sure to have an equal number of both “yes” (4,640) and ”'no” instances, ensuring a fair representation of both outcomes.

Picture 3.Adding Sample Node into the Model



By using SMOTE to balance the dataset, we aimed to improve the accuracy of our models. This approach ensured that our models were trained on a more even set of data, providing better insights into both scenarios of clients subscribing (“yes”) and not subscribing (“no”) to term deposits.

For sampling we used the following settings:

- Sample Method - Stratify
- Size type - Percentage
- Percentage - 100
- Stratified Criterion - Level Based
- Ignore Small Strata - No
- Minimum Strata Size - 5
- Level Selection - Event
- Level Proportion - 100
- Sample Proportion - 50
- Adjust Frequency - No

Sample Results:

- Target Level “no” - 4640
- Target level “yes” - 4640

Table 4. Summary statistics, interval variables (from StatExplore Node after Sampling)

Variable	Target Level	Median	Minimum	Maximum	Mean	Standard Deviation	Skewness	Kurtosis
emp_var_rate	no	1.1	-3.4	1.4	0.270129	1.473832	-0.94326	-0.67785
emp_var_rate	yes	-1.8	-3.4	1.4	-1.23345	1.623626	0.556589	-0.99048
previous	no	0	0	5	0.140302	0.422391	3.799768	19.06447
previous	yes	0	0	6	0.492672	0.860344	2.132789	5.267404
duration	no	166	0	4918	222.505	222.8422	5.390659	69.63531
duration	yes	449	37	4199	553.1912	401.1719	1.837301	6.318574
euribor3m	no	4.858	0.635	5.045	3.835632	1.6265	-0.97815	-0.94089
euribor3m	yes	1.266	0.634	5.045	2.123135	1.742598	0.890782	-1.08895
campaign	no	2	1	33	2.678233	2.834693	3.968208	23.28107
campaign	yes	2	1	23	2.051724	1.666245	3.387153	19.35519
age	no	38	17	88	39.98793	9.823154	0.611795	0.22527
age	yes	37	17	98	40.91315	13.83748	0.999425	0.671076
cons_conf_idx	no	-41.8	-50.8	-26.9	-40.5479	4.392402	0.276459	-0.39105
cons_conf_idx	yes	-40.4	-50.8	-26.9	-39.7898	6.139668	0.259931	-0.78703
nr_employed	no	5196	4964	5228	5177.074	63.86761	-1.23794	0.641159
nr_employed	yes	5099	4964	5228	5095.201	87.40568	0.282066	-1.19989
cons_price_idx	no	93.918	92.201	94.767	93.60721	0.558607	-0.25992	-0.80801
cons_price_idx	yes	93.2	92.201	94.767	93.35439	0.676644	0.119767	-1.05428

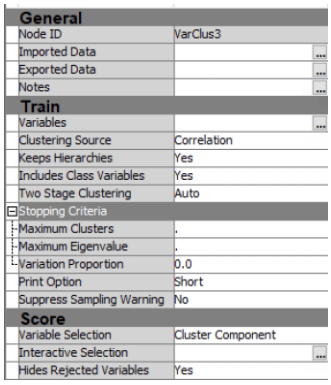
StatExplore shows that we have the same skewness and outliers as we had before sampling: “campaign”, “previous” and “duration” with target levels “yes” and “no” and “nr.employed” with target level “no”.

Variables Clustering (checking on correlation)

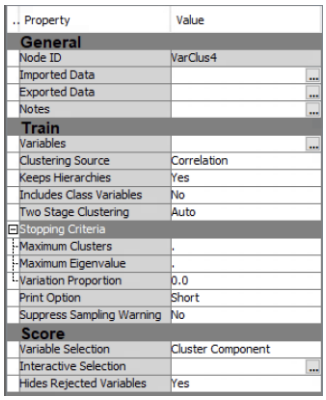
To understand the relationships between our variables and exclude multicollinearity we ran Variables Clustering node. Because multicollinearity leads to high correlation between variables and makes it hard to interpret our model. Also highly correlated variables may contain redundant information.Including redundant variables in the model is of little benefit and can lead to problems such as overfitting.

We added 2 Variable Clustering Nodes into our model with the following settings:

Picture 5. Variable Clustering Nodes Settings



General	
Node ID	VarClus3
Imported Data	
Exported Data	
Notes	
Train	
Variables	
Clustering Source	Correlation
Keeps Hierarchies	Yes
Includes Class Variables	Yes
Two Stage Clustering	Auto
Stopping Criteria	
Maximum Clusters	.
Maximum Eigenvalue	.
Variation Proportion	0.0
Print Option	Short
Suppress Sampling Warning	No
Score	
Variable Selection	Cluster Component
Interactive Selection	
Hides Rejected Variables	Yes



.. Property	Value
General	
Node ID	VarClus4
Imported Data	
Exported Data	
Notes	
Train	
Variables	
Clustering Source	Correlation
Keeps Hierarchies	Yes
Includes Class Variables	No
Two Stage Clustering	Auto
Stopping Criteria	
Maximum Clusters	.
Maximum Eigenvalue	.
Variation Proportion	0.0
Print Option	Short
Suppress Sampling Warning	No
Score	
Variable Selection	Cluster Component
Interactive Selection	
Hides Rejected Variables	Yes

Variable Clustering Node #1 (with Class Variables):

- Clustering Source - Correlation
- Keeps Hearachies - Yes
- Includes Class Variables - Yes
- Two Stage Clustering - Auto
- Variable Selection - Cluster Component
- Hides Rejected Variables - Yes

Variable Clustering Node #2 (without Class Variables):

- Clustering Source - Correlation
- Keeps Hearachies - Yes
- Includes Class Variables - No
- Two Stage Clustering - Auto
- Variable Selection - Cluster Component
- Hides Rejected Variables - Yes

After adding these nodes we explored Variable Correlation tables and found that there are few highly correlated pairs of variable:

- emp_var_rate - euribor3m - 0.958834
- euribor3m - nr_employed - 0.941528
- emp_var_rate - nr_employed - 0.869678
- cons_price_idx - emp_var_rate - 0.72169

Picture 6. Adding Variable Clustering Nodes into Model

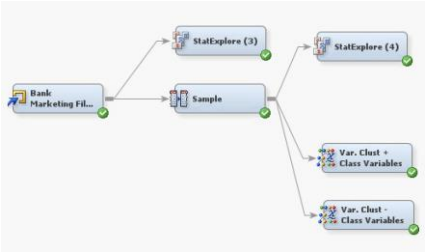


Table 5. Correlation between variables (Variable Clustering Node with Class Variables)

Variablle	Variable	Correlation
emp_var_rate	euribor3m	0.958834
euribor3m	emp_var_rate	0.958834
euribor3m	nr_employed	0.941528
nr_employed	euribor3m	0.941528
emp_var_rate	nr_employed	0.869678
nr_employed	emp_var_rate	0.869678
cons_price_idx	emp_var_rate	0.72169
emp_var_rate	cons_price_idx	0.72169

pdays=999	previous	-0.7084
previous	pdays=999	-0.7084
marital=married	marital=single	-0.79139
marital=single	marital=married	-0.79139
poutcome=nonexistent	previous	-0.84914
previous	poutcome=nonexistent	-0.84914
loan=no	loan=yes	-0.91345
loan=yes	loan=no	-0.91345
housing=no	housing=yes	-0.95127
housing=yes	housing=no	-0.95127
pdays=999	poutcome=success	-0.95566
poutcome=success	pdays=999	-0.95566

Table 6. Correlation between variables (Variable Clustering Node without Class Variables)

Variablle	Variable	Correlation
emp_var_rate	euribor3m	0.958834178
euribor3m	emp_var_rate	0.958834178
euribor3m	nr_employed	0.94152801
nr_employed	euribor3m	0.94152801
emp_var_rate	nr_employed	0.86967763
nr_employed	emp_var_rate	0.86967763
cons_price_idx	emp_var_rate	0.721690152
emp_var_rate	cons_price_idx	0.721690152

Rejected Variables

Picture 7. Rejected Variables

Name	Role	Level
age	Input	Interval
campaign	Input	Interval
cons_conf_idx	Rejected	Interval
cons_price_idx	Rejected	Interval
contact	Input	Nominal
day_of_week	Input	Nominal
default	Input	Nominal
duration	Input	Interval
education	Input	Nominal
emp_var_rate	Rejected	Interval
euribor3m	Rejected	Interval
housing	Input	Nominal
job	Input	Nominal
loan	Input	Nominal
marital	Input	Nominal
month	Input	Nominal
nr_employed	Rejected	Interval
pdays	Input	Nominal
poutcome	Input	Nominal
previous	Rejected	Interval
y	Target	Binary

To exclude this correlation and avoid multicollinearity we decided to reject all social and economic context attributes because they are irrelevant: “euribor3m”, “emp.var.rate”, “cons.price.idx”, “cons.conf.idx”, and “nr.employed”. Also we detected the correlation between “previous” = 0 and “pday” = 999. Therefore, we decided that the variable “previous” is redundant to a certain content and we rejected it too.

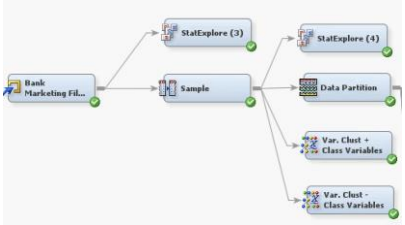
We corrected the roles for these variables in the Edit Variable Setting in our dataset node.

Table 7. Sample Statistics after rejecting variables

#	Variable	Role	Level	Minimum	Maximum	Mean	Number of Levels
1	contact	Input	Nominal	.	.	.	2
2	day_of_week	Input	Nominal	.	.	.	5
3	default	Input	Nominal	.	.	.	2
4	education	Input	Nominal	.	.	.	8
5	housing	Input	Nominal	.	.	.	3
6	job	Input	Nominal	.	.	.	12
7	loan	Input	Nominal	.	.	.	3
8	marital	Input	Nominal	.	.	.	4
9	month	Input	Nominal	.	.	.	10
10	poutcome	Input	Nominal	.	.	.	3
11	y	Input	Nominal	.	.	.	2
12	age	Input	Interval	17	98	41.52533333	.
13	campaign	Input	Interval	1	16	2.057833333	.
14	duration	Input	Interval	1	3785	272.6285	.
15	pdays	Input	Nominal	0	999	792.4865	.
16	y	Target	Binary	.	.	.	2

DATA PARTITIONING

Picture 8. Adding Data Partition Node into the Model



For the predictive analysis, the dataset was split into two different subsets: a training subset consisting of 50% of the data and a separate test subset also consisting of 50% of the data (Add Node - Sample - Data Partition).

1. **Training Data.** Used for model training and parameter estimation. There is 50% from the whole dataset in our Model.

2. **Validation Data.**

- Used for evaluating model performance.
- Independent dataset not used during model training.
- Helps assess how well the model performs on unseen data.
- There is 50% from the whole dataset in our Model.

This partitioning strategy ensures a balanced representation of instances in both subsets, preserving the original distribution of data characteristics during the training and testing phases.

Picture 9. Data Partition results

```
Partition Summary

                                          Number of
Type              Data Set              Observations

DATA        EMWS5.Smpl2_DATA                 9280
TRAIN       EMWS5.Part2_TRAIN                4638
VALIDATE    EMWS5.Part2_VALIDATE             4642
```

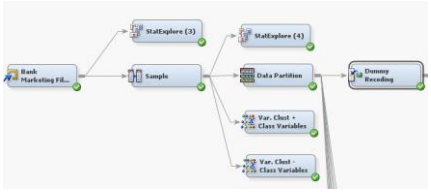
In our results, a total of 9280 observations were recorded. Following the partitioning of our dataset, an equitable distribution of 50% was allocated to both the Training Data and the Validation Data sets. Specifically, the software assigned 4638 observations to the Training Data subset and 4642 observations to the Validation Data subset. This systematic division ensures a balanced representation of data for training and validating our model for our research on Bank Marketing .

DUMMY RE-CODING

After exploring our class variables and their levels we decided to collapse them into smaller categories. For this reason it was decided to use Replacement Node and Dummy re-coding technique. Dummy re-coding and collapsing level recording are techniques used to transform categorical variables in a dataset also for the following reasons:

1. **Simplify and Group Categories.** By collapsing we reduce the number of categories, making the data more manageable and easier to interpret.
2. **Reducing Dimensionality.** This approach is beneficial in cases where too many levels might introduce noise or unnecessary complexity into models.**Enhancing Interpretability.** This technique can make the interpretation of the data and the resulting analyses more straightforward and meaningful.

Picture 10. Adding Replacement Node into the Model



3. **Handling Rare Categories:** This avoids problems with overfitting due to insufficient data for these less frequent categories.

The setting of the Replacement Node (that was renamed to Dummy recoding):

- Default Limits Method - None
- Unknown Levels - Ignore
- Replacement Values - Computed
- Score Hide - No

In the context of our study of the bank's marketing campaign, a strategic decision was made to group days of the week, education, work, and months together for the purposes outlined above.

1. **Days of the week.** Monday through Thursday days were grouped together to reflect typical working days, while Friday was considered separately due to its unique characteristic as a shorter working day in Portugal. To represent this grouping in the dataset, we used a concise labeling system in which "M/T/W/TH" values indicate the combined Monday through Thursday workdays. In contrast, the "F" value refers to Friday only.
2. **Education**. We simplified the variable "education" as follows: categories such as "basic.4y", "basic.6y" and "basic.9y" were combined into one group called "basic". The categories “University degree” and “Professional Courses” were combined into “Un/Pr”. In addition, we combined the categories "unknown"

and "illiterate" into a new category called "UN/Il" to account for cases where education status is either missing or categorized as illiterate.

3. **Occupation**. Within the "job" variable, occupations were categorized strategically to increase interpretability and decrease noise. Categories such as "students" and "service", "administration" were grouped into "S/S/A", representing sectors characterized by service-oriented roles. Similarly, the occupations of "blue-collar," "technician," and "housekeeper" were combined and renamed "B/T/H," representing occupations with common characteristics involving manual labor. Finally, the categories "entrepreneur", "self-employed", and "management" were combined into "E/S/M", reflecting occupations with entrepreneurial or managerial characteristics.
4. **Months**. For a more organized time-based analysis, the variable "months" was transformed into quarters. For example, March was assigned to Quarter 1 (Q1). This restructuring allows for a more consistent examination of campaign trends and results over specific time intervals. The only thing to add is that the data in our dataset was collected for 1 year from March to December. Therefore, in Q1 we have only one month March.
5. **Number of days after last contact (pdays)**. The number of days was defined by how many weeks had passed since the person was last contacted. In this case, if the pdays were 14, group it into "2", which means it has been 2 weeks since the person was last contacted. Also there is a category when a person has never been contacted with values "0" and "999", it was collapsed into one group with the name "0".

Table 8. Re-coding variables

Variables	Formatted Value	New Value	Description
day_of _week	mon, tue, wed, thu	M/T/W/Th	Monday, Tuesday, Wednesday, Thursday
day_of_week	fri	Fr	Friday
education	university.degree, professional.course	Un/Pr	University Degree, Professional Course
education	high.school	HS	High School
education	basic.9y	Basic	Basic education
education	basic.4y	Basic	Basic education
education	basic.6y	Basic	Basic education

education	unknown, illiterate	Un/Il	Unknown, illiterate
job	student, services, admin	S/S/A	Student, services, admin
job	blue-collar, technician, housemaid	B/T/H	Blue-collar, technician, housemaid
job	entrepreneur, self-employed, management	E/S/M	Entrepreneur, self-employed, management
job	unemployed	Unempl	Unemployed
job	unknown	Unknown	Unknown
month	mar	Q1	Quarters of the year
month	apr, may, jun	Q2	Quarters of the year
month	jul, aug, sep	Q3	Quarters of the year
month	oct, nov, dec	Q4	Quarters of the year
pdays	0, 999	0	Never contacted
pdays	1, 2, 3, 4, 6	1	1 Week since last contacted
pdays	8,10,11,12,13,14	2	2 Weeks since last contacted
pdays	15,17,18,16,18,21	3	3 Weeks since last contacted

MODELING AND PREDICTIVE ANALYSIS

DECISION TREES

SAS Enterprise Miner that we usd for our analysis has 3 main groups of tools: primary, specialty, and multiple models. The primary tools interface with the three most commonly used predictive modeling methods: decision tree, regression, and neural network.

Decision Tree is a tree-like structure where each internal node represents a decision based on a particular feature, each branch represents the outcome of the decision, and each leaf node represents a class label or a decision. Cases are evaluated using *prediction rules*. A *split search* algorithm facilitates the selection of input data. Model complexity is addressed by *pruning*.

The main features of Decision Trees.

1. **Interpretability**. Decision trees are easy to understand and interpret. They mimic human decision-making processes, making them intuitive and explainable. Each decision and partitioning in the tree is based on specific criteria, allowing analysts to interpret the logic behind the predictions.
2. **Handling Missing Values and Outliers**. Decision trees can handle missing values and outliers without needing to preprocess the data. They can efficiently handle mixed data types (categorical and numeric).
3. **Handling Non-linear Relationships**. Decision trees can model complex relationships between features and the target variable without assuming a linear relationship. They are able to capture non-linear patterns in the data, making them universally applicable for solving different types of problems.

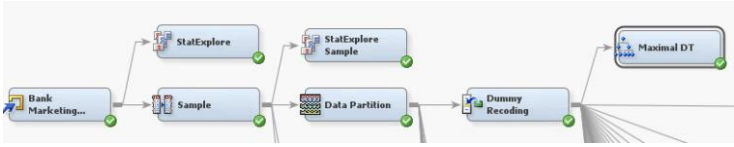
In our analysis we ran three different types of Decision Trees:

- Maximal Decision Tree with default settings.
- Misclassification Decision Tree (assessment measure - Misclassification).
- Average Square Error Decision Tree (assessment measure - Average Square Error).

We started from 2 branches for all of them except Maximal Tree and continuously increased branches to get the optimal model without repeating variables.

Maximal Decision Tree Model

Picture 11. Adding Maximal Tree Node



The model employed for this analysis is a Decision Tree of the maximal type. It was trained interactively via the Properties window - Train function. The Assessment method was applied with Decision Assessment Measure.

Model Parameters:

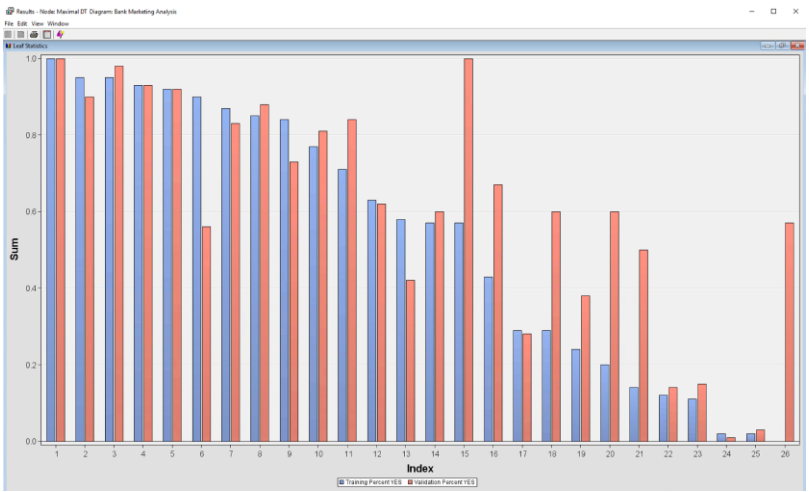
- Number of Leaves: 26
- Average Square Error for Validation Data: 0.133443
- Misclassification Rate: 0.179664
- Number of Branches: 2

Picture 12. Summary Statistics for Maximal Tree

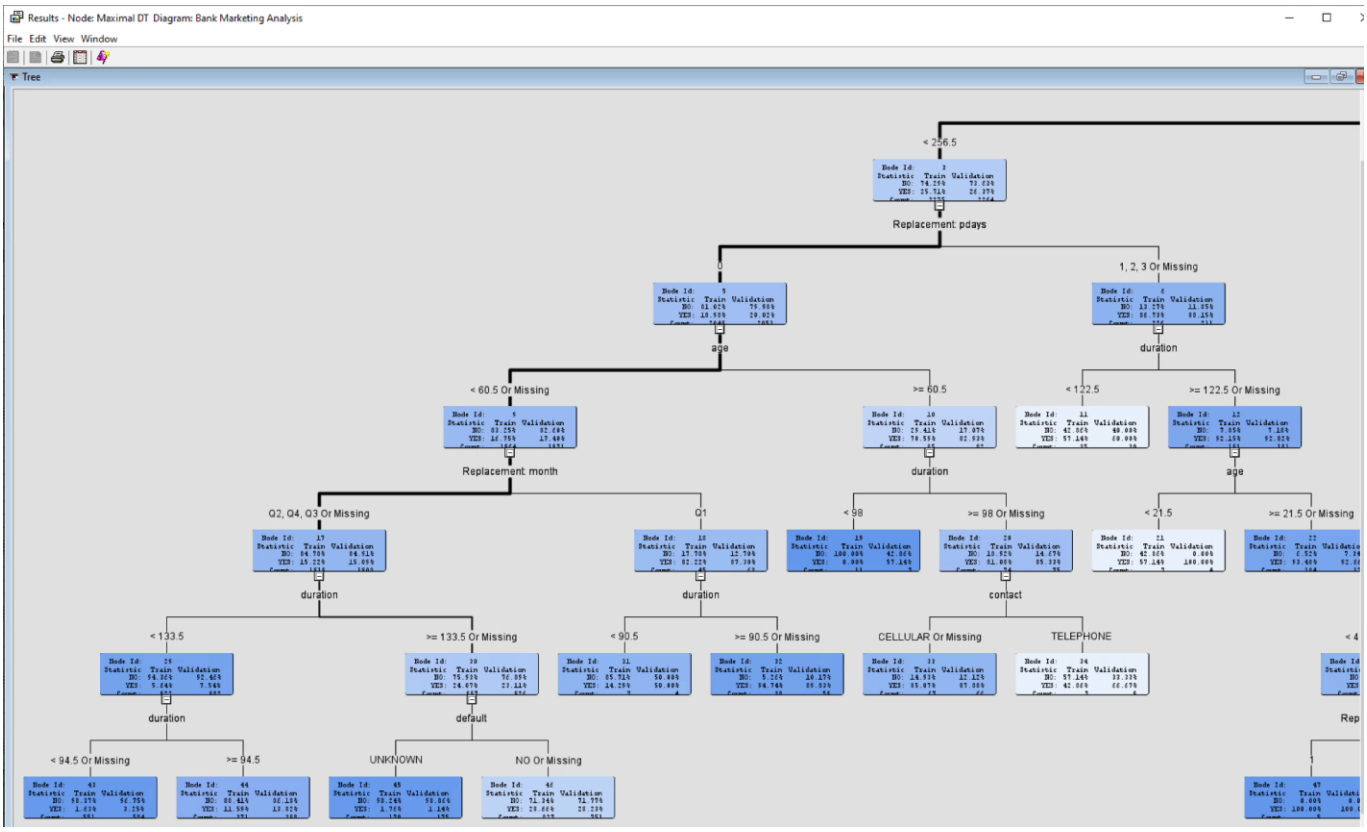
Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification R...	0.167098	0.179664	.
y	y	_MAX_	Maximum Absolute...	0.983666	1	.
y	y	_SSE_	Sum of Squared Er...	1149.932	1238.881	.
y	y	_ASE_	Average Squared E...	0.123969	0.133443	.
y	y	_RASE_	Root Average Squa...	0.352092	0.365298	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Fr...	4638	.	.

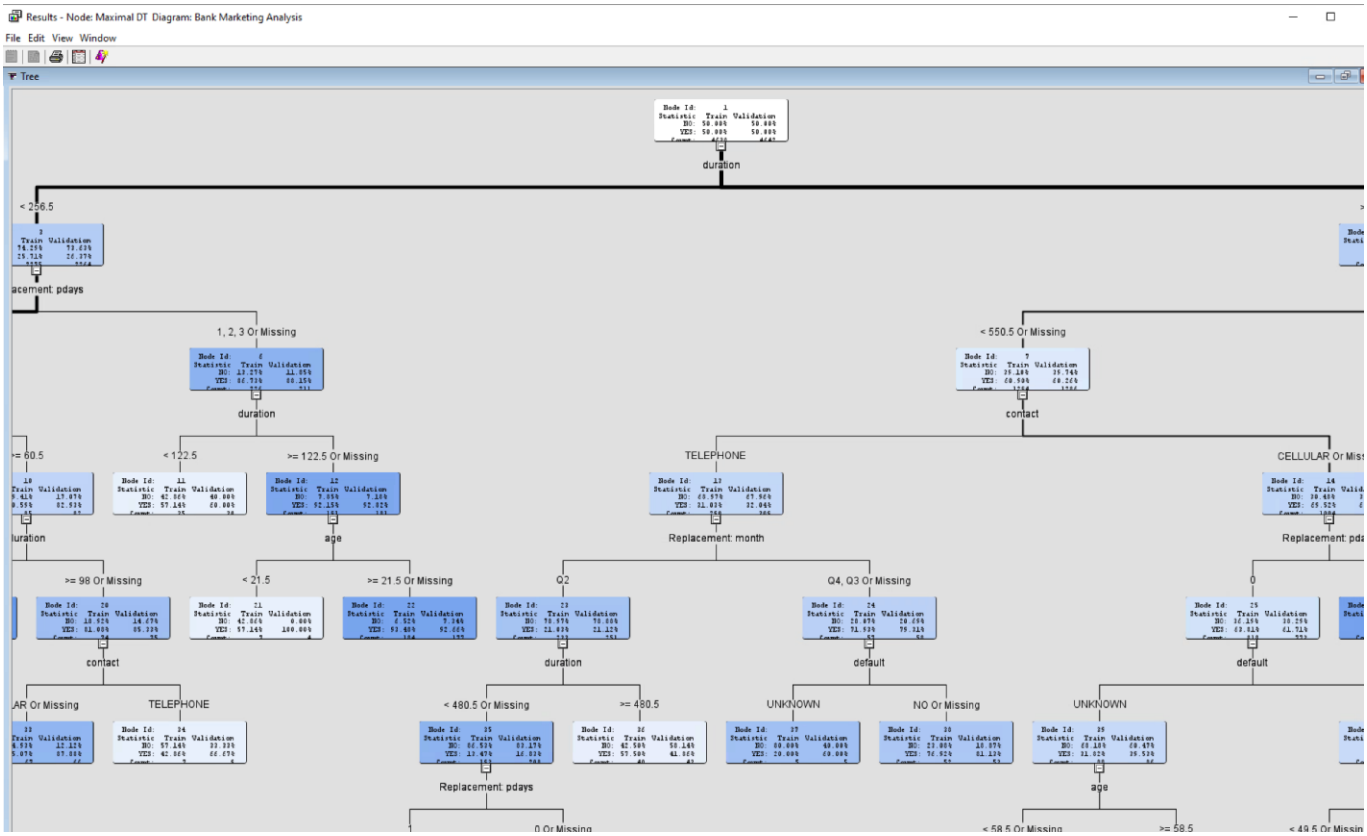
Picture 13. Maximal Tree, number of leafs



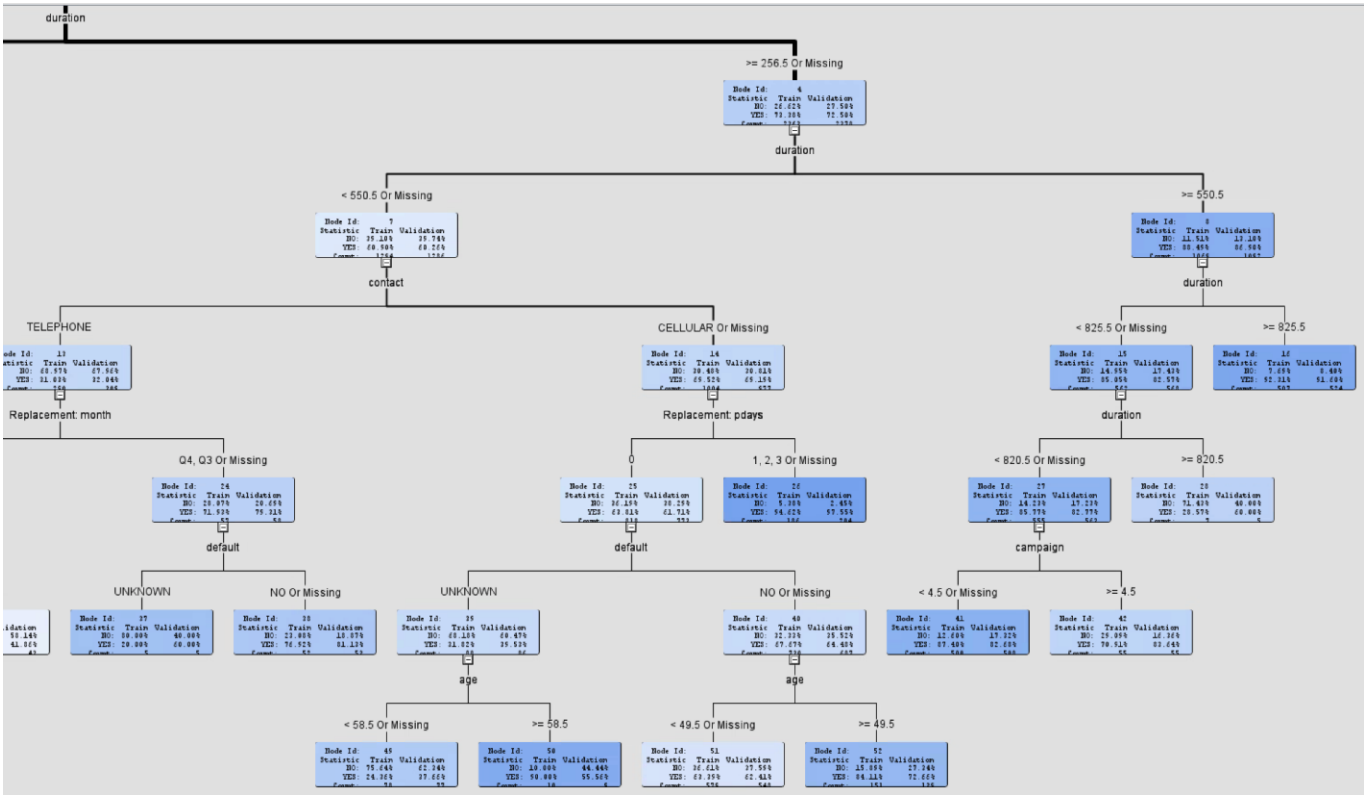
Picture 14. Maximal Tree, the model, part 1 (left)



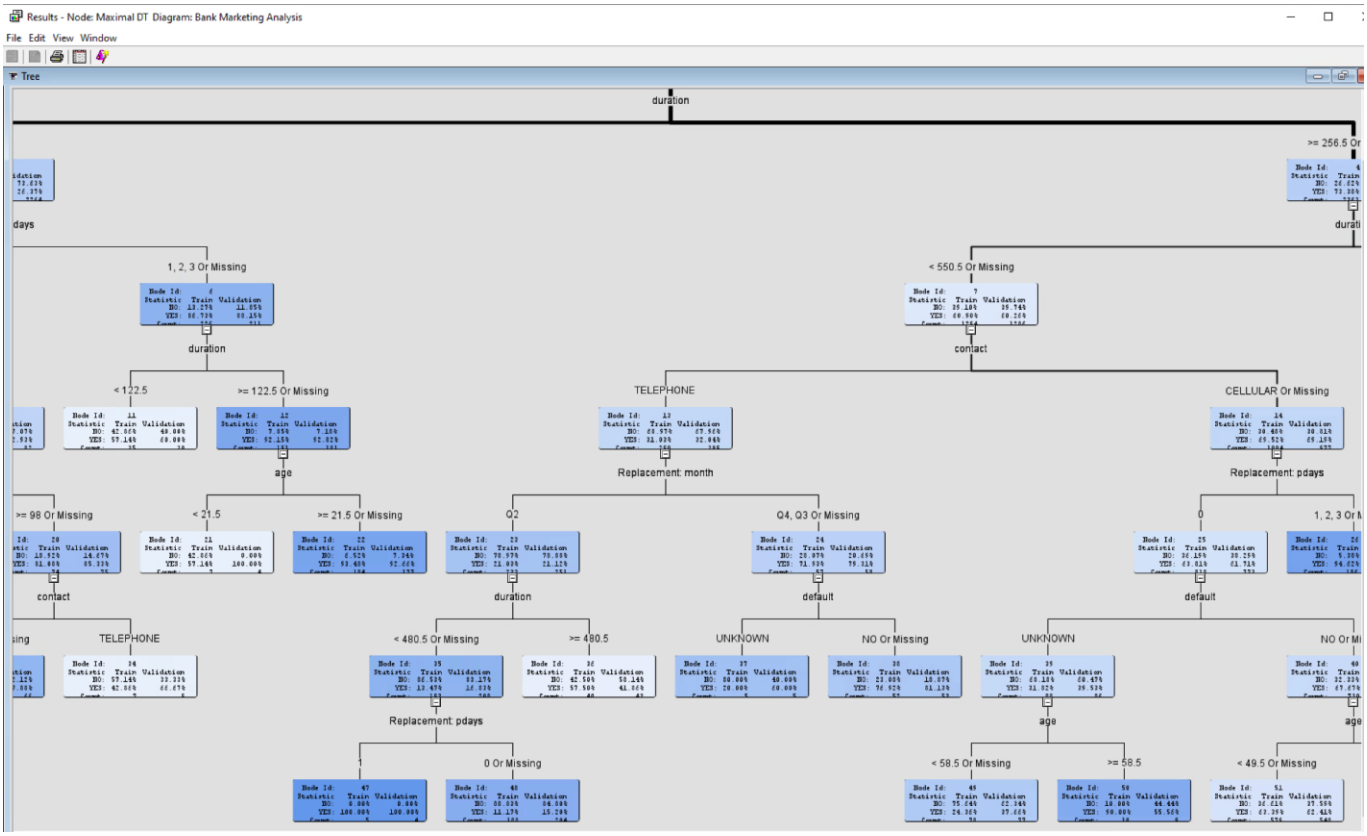
Picture 15. Maximal Tree, the model, part 2 (middle)



Picture 16. Maximal Tree, the model, part 3 (right side)



Picture 17. Maximal Tree, the model, part 4 (middle bottom)



Picture 18. Maximal Tree, Variable Importance

Results - Node: Maximal DT Diagram: Bank Marketing Analysis

File Edit View Window

Variable Importance

Variable Name	Label	Number of Splitting Rules	Importance	Validation Importance	Ratio of Validation to Training Importance
duration	duration	10	1.0000	1.0000	1.0000
REP_pdays	Replacement: pdays	3	0.5646	0.6020	1.0662
contact	contact	2	0.3137	0.3244	1.0342
age	age	4	0.3091	0.3258	1.0539
REP_month	Replacement: month	2	0.3004	0.3922	1.3057
default	default	3	0.2493	0.2160	0.8663
campaign	campaign	1	0.0620	0.0000	0.0000

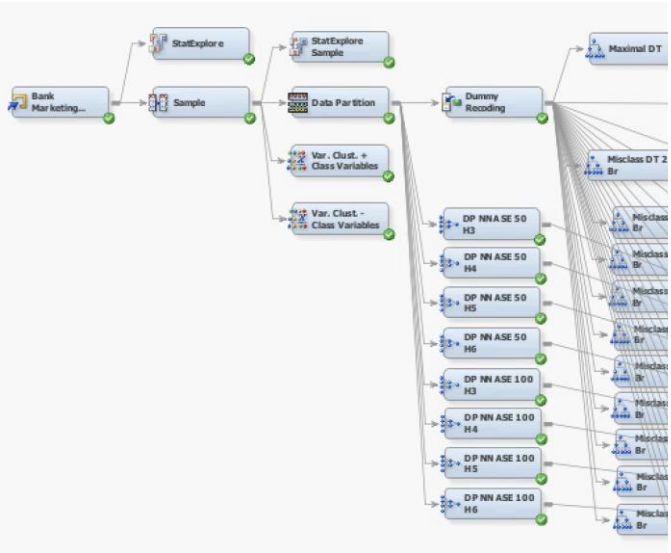
The most important variables for the Maximal Tree model are: “duration”, “REP_pdays”, “contact”, “age”, “REP_month”, “default”, and “campaign”.

The Story

One of the stories that we can tell by analyzing our Maximal tree is: in 100% of cases a person will subscribe if a representative of a bank calls by telephone from April to June in the next 7 days after the previous call and this conversation should be at least 256 sec but no more than 480 sec.

Misclassification Decision Tree Models

Picture 19. Adding Misclassification Trees into the Model



Picture 20. Misclassification Tree, settings

Subtree	
Method	Assessment
Number of Leaves	1
Assessment Measure	Misclassification
Assessment Fraction	0.25
Cross Validation	
Perform Cross Validation	No
Number of Subsets	10
Number of Repeats	1

We added a Decision Tree node into our Model changing the Assessment Measure in Subtree Setting to “Misclassification”.

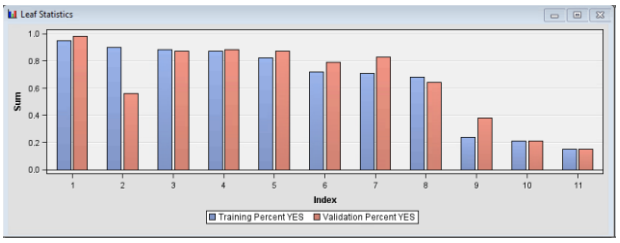
The initial evaluation was performed using an Misclassification tree with 2 branches, yielding 11 leaves, an ASE of 0.132973 and a misclassification rate of 0.177725.

Picture 21. Misclassification Tree 2 branches, fit statistics

Fit Statistics

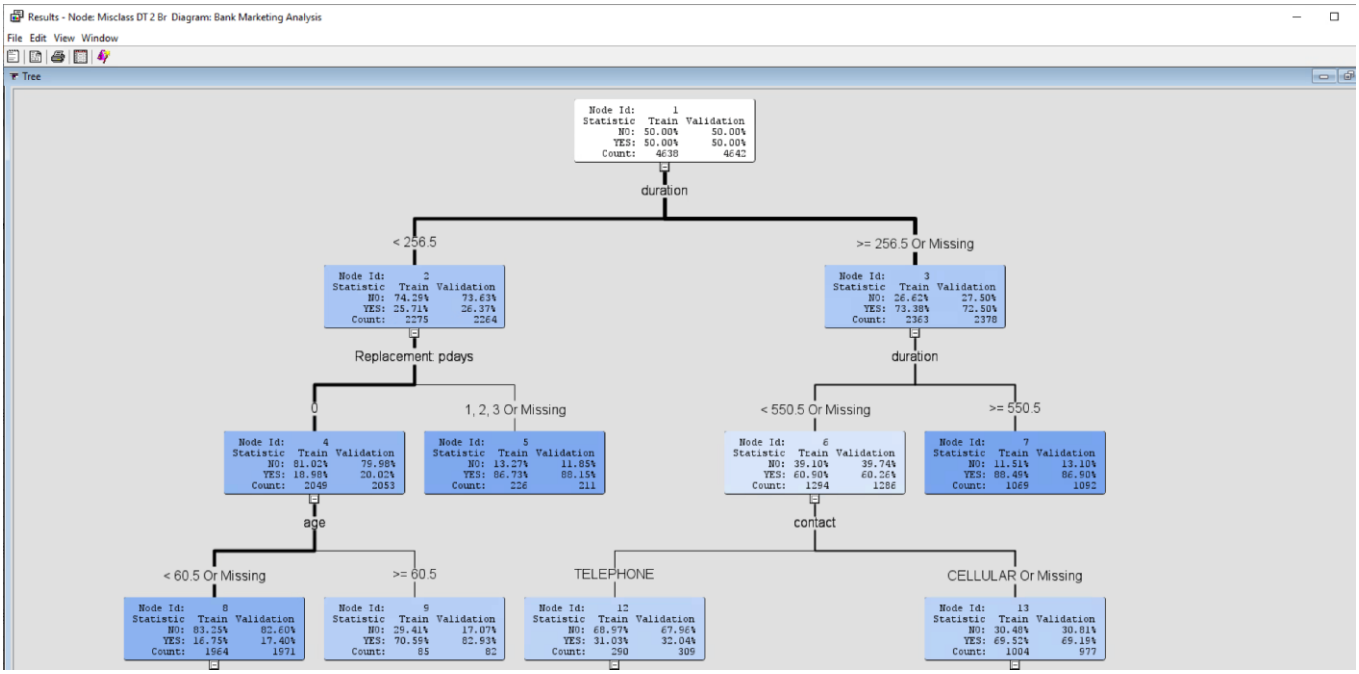
Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification R...	0.174429	0.177725	.
y	y	_MAX_	Maximum Absolute...	0.946237	0.946237	.
y	y	_SSE_	Sum of Squared Er...	1282.602	1297.257	.
y	y	_ASE_	Average Squared E...	0.138271	0.13973	.
y	y	_RASE_	Root Average Squa...	0.371848	0.373805	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Fr...	4638	.	.

Picture 22. Misclassification Tree, 2 branches, number of leafs

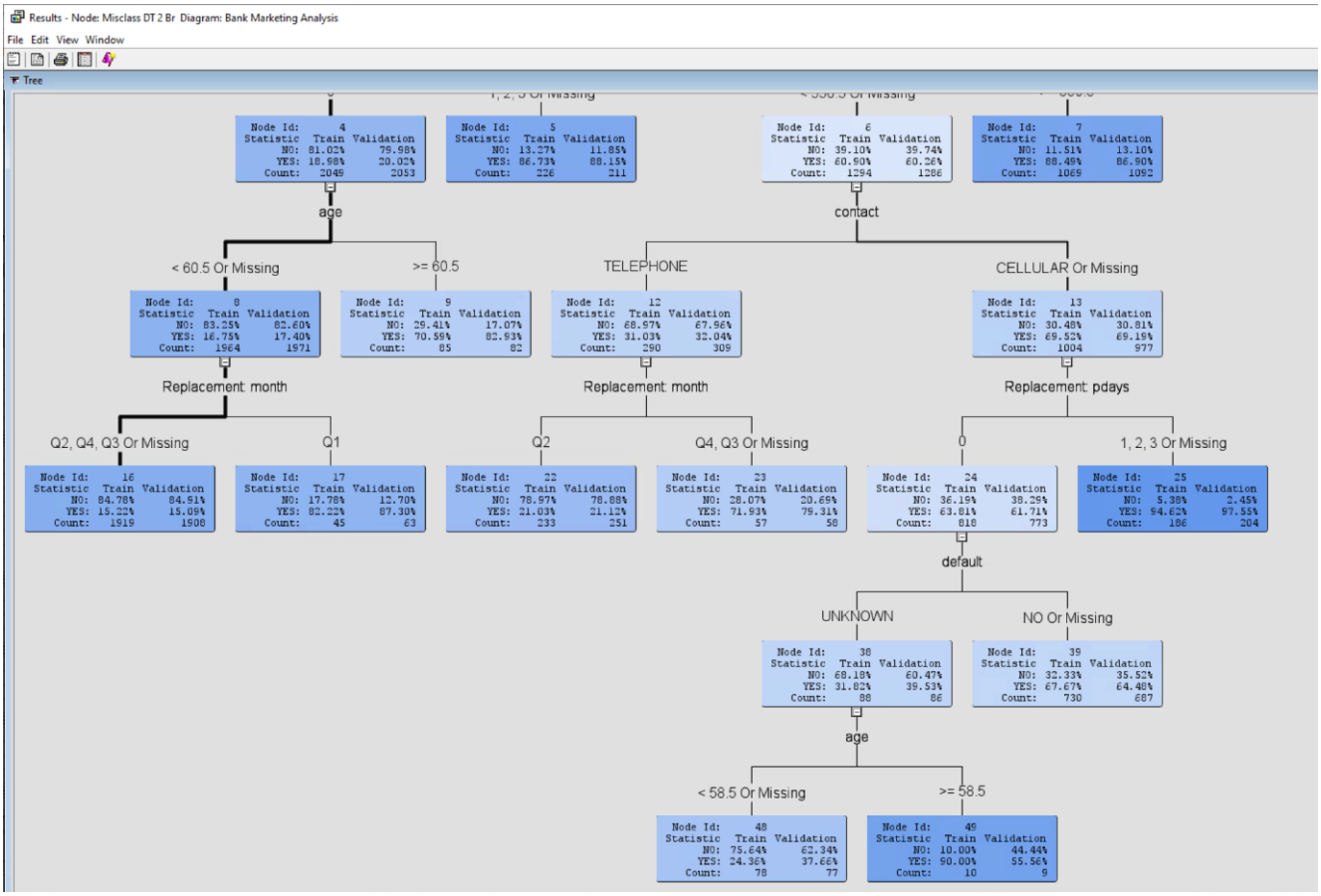


As we can see on the Model there are few repeating variables such as “duration”. That means that we should increase the number of branches for the next tree. We repeated to increase these numbers until we did not get the Model without repeating variables. During this process we compared ASE and Misclassification rates to define which Model performs better.

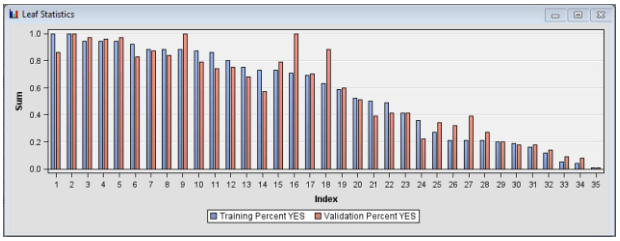
Picture 23. Misclassification Tree 2 branches, the model, part 1 (upper part)



Picture 24. Misclassification Tree 2 branches, the model, part 2 (bottom part)



Picture 25. Misclassification Tree, 11 branches, number of leafs



<u>A model with 11 branches</u> proved promising, yielding 35 leaves, an ASE of 0.134435, and a misclassification rate of 0.181172. Notably, this model contained no repeated variables in its structure, which may indicate an improvement in discrimination ability without the redundant variables observed in previous iterations.

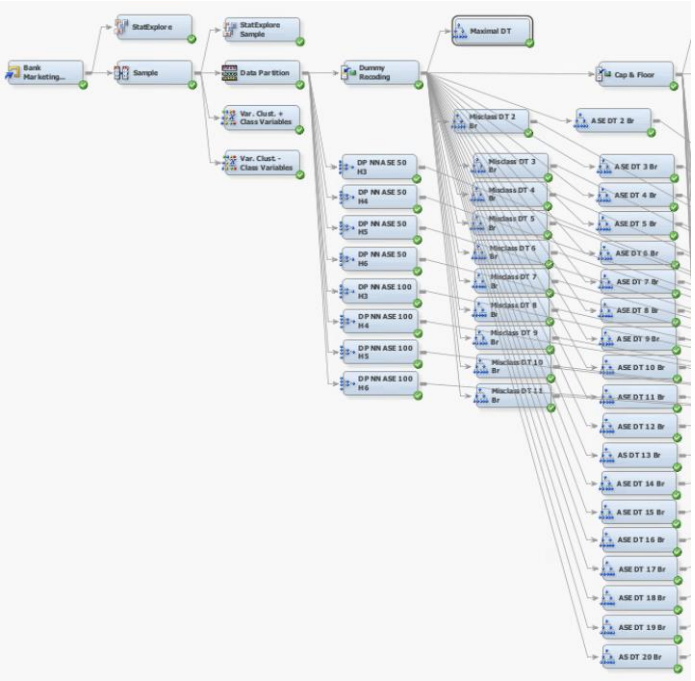
Picture 26. Misclassification Tree 11 branches, fit statistics

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification R...	0.172057	0.181172	.
y	y	_MAX_	Maximum Absolute...	0.986264	1	.
y	y	_SSE_	Sum of Squared Er...	1187.912	1248.09	.
y	y	_ASE_	Average Squared E...	0.128063	0.134435	.
y	y	_RASE_	Root Average Squa...	0.357859	0.366653	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Fr...	4638	.	.

Average Square Error Trees

Picture 27. Adding Average Square Error Trees into the Model



We added a Decision Tree Node into our Model changing the Assessment Measure to Average Square Error.

The initial evaluation was performed using an Average Square Error (ASE) tree with 2 branches, yielding 18 leaves, an ASE of 0.132279 and a misclassification rate of 0.178587.

Picture 28. Average Square Error Tree 2 branches, settings

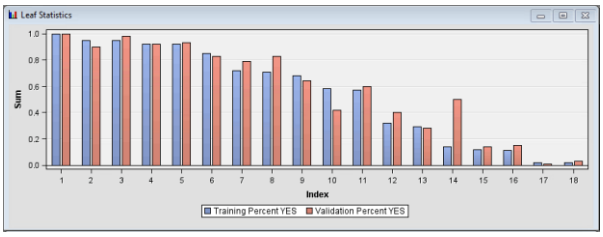
Subtree	
Method	Assessment
Number of Leaves	1
Assessment Measure	Average Square Error
Assessment Fraction	0.25
Cross Validation	

Picture 29. Average Square Error Tree 2 branches, fit statistics

Fit Statistics

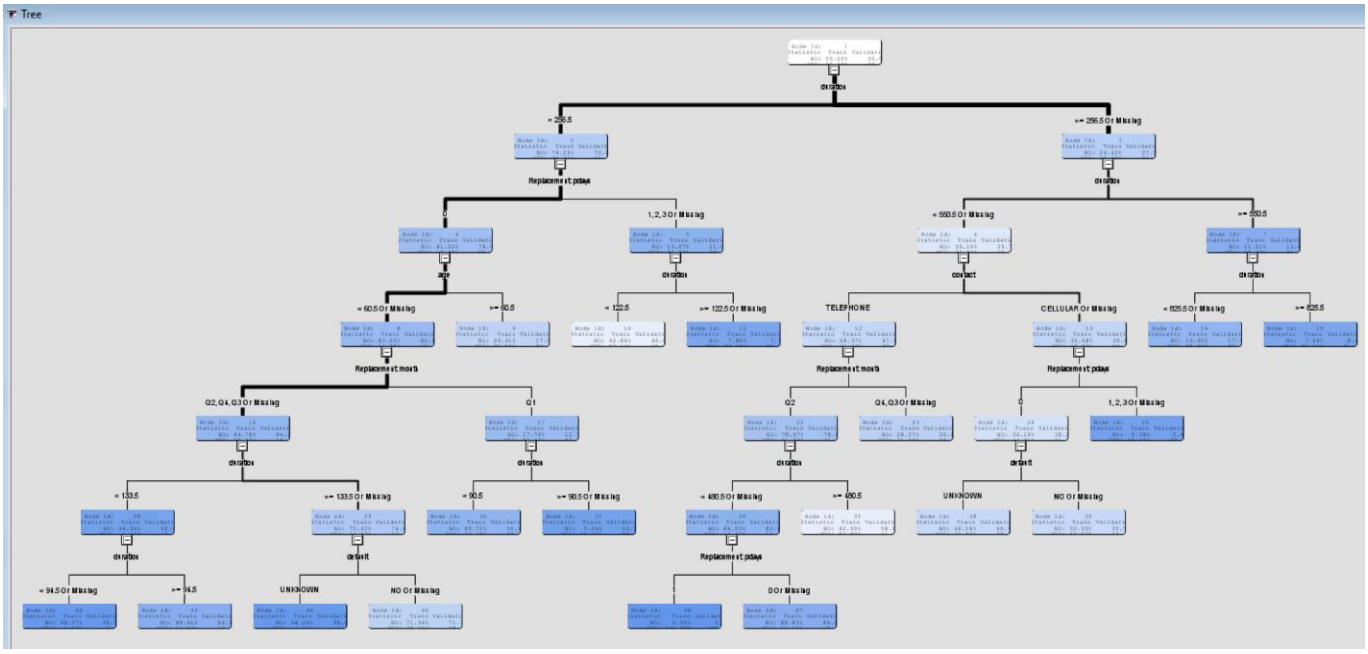
Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification R...	0.172704	0.178587	.
y	y	_MAX_	Maximum Absolute...	0.983666	0.983666	.
y	y	_SSE_	Sum of Squared Er...	1194.66	1228.079	.
y	y	_ASE_	Average Squared E...	0.12879	0.132279	.
y	y	_RASE_	Root Average Squa...	0.358874	0.363702	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Fr...	4638	.	.

Picture 30. Average Square Error Tree 2 branches, number of leafs

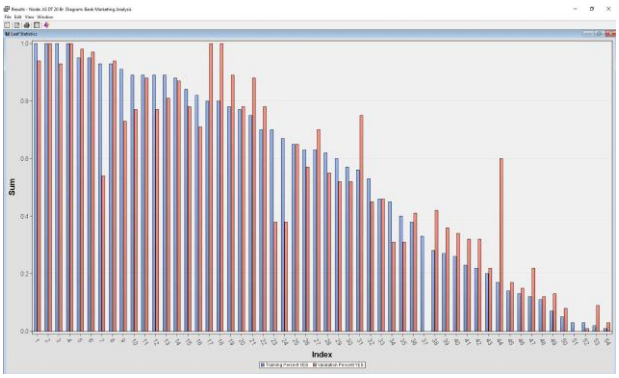


As we can see on the Model there are few repeating variables such as “duration”. That means that we should increase the number of branches for the next tree. We repeated to increase these numbers until we did not get the Model without repeating variables. During this process we compared ASE and Misclassification rates to define which Model performs better.

Picture 31. Average Square Error Tree 2 branches, the model



Picture 32. Average Square Error Tree 20 branches, number of leafs (54)



<u>A model with 20 branches</u> proved promising, yielding 50 leaves, an ASE of 0.129366, and a misclassification rate of 0.179449. Notably, this model contained no repeated variables in its structure, which may indicate an improvement in discrimination ability without the redundant variables observed in previous iterations.

Picture 33. Average Square Error Tree 20 branches, fit statistics

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification R...	0.159552	0.179449	.
y	y	_MAX_	Maximum Absolute...	0.991379	1	.
y	y	_SSE_	Sum of Squared Er...	1076.875	1201.038	.
y	y	_ASE_	Average Squared E...	0.116093	0.129366	.
y	y	_RASE_	Root Average Squa...	0.340724	0.359675	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Fr...	4638	.	.

Decision Trees Comparison

When evaluating a model that decides between “yes” or “no”, we find the misclassification rate to be more suitable than Average Square Error (ASE). The misclassification rate directly tells us how often the model gets the answer wrong. This is important because, in our case, we want to see how well the model can correctly say “yes” or “no”. It's like checking how many times it makes mistakes. Unlike ASE, which is better for different types of problems, the misclassification rate is perfect for tasks like ours where we are trying to predict between two options. So, we focus on reducing the misclassification rate to ensure our model is good at getting the right answers for our binary (“yes” or “no”) question.

Table 9. Results of Average Square Error and Misclassification Trees

Decision Tree	Number of Branches	Number of Leaves	Misclassification	ASE
ASE DT 2BR	2	18	0.178587	0.132279
ASE DT 3BR	3	33	0.177941	0.130065
ASE DT 4BR	4	38	0.179554	0.126260
ASE DT 5BR	5	39	0.188927	0.128365
ASE DT 6BR	6	39	0.188927	0.128365
ASE DT 7BR	7	39	0.188927	0.128365
ASE DT 8BR	8	39	0.188927	0.128366
ASE DT 9BR	9	39	0.188927	0.128365
ASE DT 10BR	10	39	0.188927	0.128365
ASE DT 11BR	11	45	0.182034	0.127342
ASE DT 12BR	12	49	0.177941	0.125615
ASE DT 13BR	13	50	0.184519	0.128195
ASE DT 14BR	14	45	0.181603	0.127719
ASE DT 15BR	15	40	0.182680	0.126428
ASE DT 16BR	16	44	0.183757	0.127171

ASE DT 17BR	17	44	0.187635	0.127508
ASE DT 18BR	18	49	0.186558	0.127707
ASE DT 19BR	19	45	0.180095	0.126936
ASE DT 20BR	20	54	0.179449	0.129366
MIS DT 2BR	2	11	0.177725	0.139730
MIS DT 3BR	3	28	0.177294	0.133616
MIS DT 4BR	4	31	0.178156	0.128879
MIS DT 5BR	5	33	0.185696	0.133840
MIS DT 6BR	6	33	0.185696	0.133840
MIS DT 7BR	7	33	0.185696	0.133840
MIS DT 8BR	8	33	0.185696	0.133840
MIS DT 9BR	9	33	0.185696	0.133840
MIS DT 10BR	10	33	0.185696	0.133840
MIS DT 11BR	11	35	0.181172	0.134435

The Best Misclassification Decision Tree

We found that there are two the best Misclassification Decision Trees:

- Misclassification Decision Tree with 3 branches: ASE = 0.133616, Misclassification Rate = 0.177294, 28 leafs.
- Misclassification Decision Tree with 4 branches: ASE = 0.128879, Misclassification Rate = 0.178156, 31 leafs.

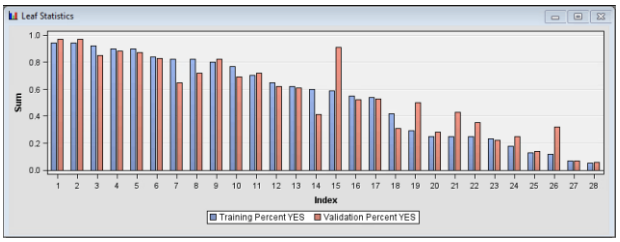
As we decided to base the Misclassification Rate for Decision Trees, we assume that the Misclassification Decision Tree with 3 branches is the best.

Picture 34. Misclassification Tree 3 branches, fit statistics

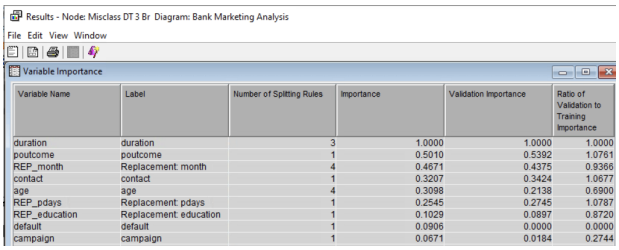
Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification R...	0.162139	0.177294	.
y	y	_MAX_	Maximum Absolute...	0.954831	0.954831	.
y	y	_SSE_	Sum of Squared Er...	1149.27	1240.489	.
y	y	_ASE_	Average Squared E...	0.123897	0.133616	.
y	y	_RASE_	Root Average Squa...	0.35199	0.365535	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Fr...	4638	.	.

Picture 35. Misclassification Tree, 3 branches, number of leafs (28)



Picture 36. Misclassification Tree, 3 branches, Variable Importance



Results - Node: Misclass DT 3 Br Diagram: Bank Marketing Analysis

File Edit View Window

Variable Importance

Variable Name	Label	Number of Splitting Rules	Importance	Validation Importance	Ratio of Validation to Training Importance
duration	duration	3	1.0000	1.0000	1.0000
poutcome	poutcome	1	0.5010	0.5392	1.0761
REP_month	Replacement: month	4	0.4671	0.4375	0.9366
contact	contact	1	0.3207	0.3424	1.0677
age	age	4	0.3098	0.2138	0.6900
REP_pdays	Replacement: pdays	1	0.2545	0.2745	1.0787
REP_education	Replacement: education	1	0.1029	0.0897	0.8720
default	default	1	0.0906	0.0000	0.0000
campaign	campaign	1	0.0671	0.0184	0.2744

The most important variables for the ASE Decision Tree Model with 12 branches are: “duration”, “poutcome”, “REP_month”, “contact”, “age”, “REP_pdays”, “Rep_education”, “default”, and “campaign”.

The Story

There is we can see the same story that the Maximal Tree showed us, also we can tell the following story: in 96.88% of cases a person will subscribe if a representative of a bank will call by cellular and the duration of this conversation will be at least 157 sec but no more than 441, it will be in March, and outcome of the previous marketing campaign will be nonexisten.

The Best ASE Decision Tree

Based on Misclassification Rate we can conclude that there are two the best ASE Decision Trees:

- ASE Decision Tree 3 Branches: ASE = 0.130065, Misclassification Rate = 0.177941, and 33 leafs.
- ASE Decision Tree 12 Branches: ASE = 0.125615, Misclassification rate = 0.177941, and 49 leafs.

But we decided to take into account that ASE Decision Tree with 12 branches has the lowest Misclassification rate and ASE as well. So, we assumed that this is the best model from ASE Decision Trees.

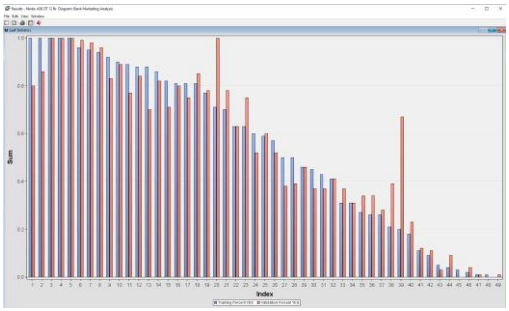
Picture 37. Average Square Error Tree 12 branches, fit statistics

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification R...	0.166667	0.177941	.
y	y	_MAX_	Maximum Absolute...	0.991304	1	.
y	y	_SSE_	Sum of Squared Er...	1098.849	1166.211	.
y	y	_ASE_	Average Squared E...	0.118461	0.125615	.
y	y	_RASE_	Root Average Squa...	0.344182	0.354422	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Fr...	4638	.	.

Picture 38. Average Square Error Tree 12 branches, number of leafs (49)

Picture 39. Average Square Error Tree 12 branches, Variable Importance



Results - Node: ASE DT 12 Br Diagram: Bank Marketing Analysis

File Edit View Window

Variable Importance

Variable Name	Label	Number of Splitting Rules	Importance ▼	Validation Importance	Ratio of Validation to Training Importance
duration	duration	2	1.0000	1.0000	1.0000
REP_month	Replacement: month	9	0.4094	0.3981	0.9723
poutcome	poutcome	3	0.4002	0.3986	0.9959
contact	contact	4	0.3554	0.3751	1.0554
age	age	3	0.3246	0.3147	0.9696
REP_pdays	Replacement: pdays	4	0.2831	0.3127	1.1043
default	default	1	0.1860	0.1466	0.7879

The most important variables for the ASE Decision Tree Model with 12 branches are: “duration”, “REP_month”, “poutcome”, “contact”, “age”, “REP_pdays”, and “default”.

The Story

There is we can see the same story that the Maximal Tree showed us, also we can tell the following story: in 100% of cases a person will subscribe if the duration of a call from representative of a bank will be at least 162 sec but no more than 256, it will be in March, the age of this person will be 26-60 years old, and outcome of the previous marketing campaign will be nonexistent.

OUTLIERS AND SKEWNESS

Picture 40. Summary Statistics for Data Partition Node (StatExplore)

Results - Node: StatExplore Sample Diagram: Bank Marketing Analysis

File Edit View Window

Interval Variables

Data Role	Target	Target Level	Variable	Median	Missing	Non Missing	Minimum	Maximum	Mean	Standard Deviation	Skewness	Kurtosis	Role	Label	Scaled Mean Deviation	Maximum Deviation	Level Id
TRAIN	y	no	duration	169	0	2319	4	4918	223.9418	234.4896	6.775871	98.35723	INPUT	duration	-0.42207	0.422067	1
TRAIN	y	yes	duration	453	0	2319	37	4199	551.0332	392.273	1.726539	5.715318	INPUT	duration	0.422067	0.422067	2
TRAIN	y	no	campaign	2	0	2319	1	30	2.561449	2.639629	4.031323	24.64788	INPUT	campaign	0.116436	0.116436	1
TRAIN	y	yes	campaign	1	0	2319	1	17	2.027167	1.623715	3.191397	16.00937	INPUT	campaign	-0.11644	0.116436	2
TRAIN	y	no	age	38	0	2319	17	86	39.89823	9.863214	0.615843	0.168726	INPUT	age	-0.0098	0.009803	1
TRAIN	y	yes	age	37	0	2319	17	92	40.68823	13.62146	1.047476	0.778568	INPUT	age	0.009803	0.009803	2

During this analysis, we identified that certain variables exhibit <u>significant skewness:</u>

1. **“campaign”**: 4.031323 for “no” and 3.19397 for “yes”
2. **“duration”**: 6.775871 for “no” and 1.726539 for “yes”
3. **“age”**: 1.047476 for “yes”

Skewness is a measure of the asymmetry in a distribution, and extreme values can contribute to a skewed distribution. It could potentially impact regression analysis and subsequent modeling due to their non-normal distributions and the associated challenges in meeting regression assumptions and interpretation.

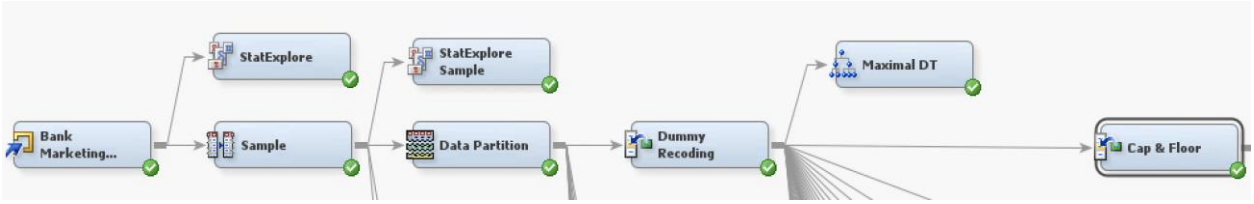
1. **Bias in Model Outputs.** Skewed variables can introduce bias into regression models. Extreme skewness can lead to non-normality in the data, violating the assumptions of the regression analysis. This can affect the accuracy and reliability of coefficient estimates and model predictions.
2. **Outlier Sensitivity.** Skewed distributions often indicate the presence of outliers or extreme values. Outliers can have an excessive influence on a regression model, leading to over- or underestimates, which negatively affects the predictive performance of the model.

In our regression analysis for this dataset, we chose to deal with extreme values by using transformation techniques, such as "cap and floor" and logarithmic transformation, for our input variables that have skewness. This purposeful approach aims to reduce the impact of extreme values, which increases the stability of our model and removes the problems associated with overfitting and underfitting.

Capping and Flooring

Capping-flooring is a technique used to handle extreme values (outliers) in a dataset by replacing them with a specified upper or lower limit. It is effective in reducing skewness by limiting the influence of extreme values. By replacing extreme values with capped or floored values, the distribution becomes more symmetric.

Picture 41. Adding Cap & Floor Node into the Model (Replacement Node)



Picture 42. Settings Cap & Floor Node

Property	Value
General	
Node ID	Repl2
Imported Data	
Exported Data	
Notes	
Train	
Interval Variables	
Replacement Editor	
Default Limits Method	Standard Deviations from the Mean
Cutoff Values	
Class Variables	
Replacement Editor	
Unknown Levels	Ignore
Score	
Replacement Values	Computed
Hide	No
Report	
Replacement Report	Yes

In our analysis we decided to use Standard Deviations from the Meant as Default Limits Method. This method involves setting upper and lower bounds for data values, usually defining thresholds in terms of the standard deviation from the mean.

1. **Calculate Mean and Standard Deviation.** Begin by calculating the mean and standard deviation of the dataset for the interval variables.
2. **Define Thresholds.** There are 3 standard deviations that will be used as the threshold for identifying extreme values.
3. **Identify Outliers.** Identify data points that fall outside the defined thresholds. Values that lie above or below these thresholds are considered outliers or extreme values.
4. **Cap and Floor Values**. Once the outliers are identified, cap (set an upper limit) or floor (set a lower limit) these extreme values to the threshold limits by replacing them with the highest or lowest threshold value.
5. **Preserving data integrity.** Capping and flooring preserve the essential information contained in extreme values, minimizing their impact on statistical analyses or predictive models.

Overall, this method provides a systematic approach to handling outliers in a dataset, ensuring that extreme values do not unduly influence subsequent analyses or models, while preserving the overall structure and information of the data.

Picture 43. Total Replacements Count

Total Replacement Counts

Variable	Label	Role	Train	Validation
age	age	INPUT	71	71
campaign	campaign	INPUT	105	128
duration	duration	INPUT	75	78

Picture 44. Thresholds with highest and lowest boundaries for interval variables

Interval Variables

Variable	Replace Variable	Lower limit	Upper Limit	Label	Limits Method	Replacement Method	Lower Replacement Value	Upper Replacement Value
age	REP_age	4.602157	75.9843	age	STDDEV	COMPUTED	4.602157	75.9843
campaign	REP_campaign	-4.32774	8.916355	campaign	STDDEV	COMPUTED	-4.32774	8.916355
duration	REP_duration	-699.002	1473.977	duration	STDDEV	COMPUTED	-699.002	1473.977

Limits and Replacement Values for Interval Variables:

1. Age (REP_age)
   - Replacement Value for values below the lower limit: 4.602
   - Replacement Value for values above the upper limit: 75.98
2. Campaign (REP_campaign)
   - Replacement Value for values below the lower limit: -4.328
   - Replacement Value for values above the upper limit: 8.92
3. Duration (REP_duration)
   - Replacement Value for values below the lower limit: -699.002
   - Replacement Value for values above the upper limit: 1473.98,

Picture 45. Summary Statistics for Cap & Floor Node (StatExplore)

Results - Node: StatExplore C&F Diagram: Bank Marketing Analysis

File Edit View Window

Interval Variables

Data Role	Target	Target Level	Variable	Median	Missing	Non Missing	Minimum	Maximum	Mean	Standard Deviation	Skewness	Kurtosis	Role	Label	Scaled Mean Deviation	Maximum Deviation	Level Id
TRAIN	y	no	REP_duration	169	0	2319	4	1473.977	220.7533	198.4418	2.705579	10.89055	INPUT	Replacement...	-0.41987	0.419868	1
TRAIN	y	yes	REP_duration	453	0	2319	37	1473.977	540.2913	353.1727	0.948574	0.156221	INPUT	Replacement...	0.419868	0.419868	2
TRAIN	y	no	REP_campaign	2	0	2319	1	8.916355	2.42237	1.947267	1.822747	2.941256	INPUT	Replacement...	0.094895	0.094895	1
TRAIN	y	yes	REP_campaign	1	0	2319	1	8.916355	2.002476	1.476382	2.237525	6.040691	INPUT	Replacement...	-0.09489	0.094895	2
TRAIN	y	no	REP_age	38	0	2319	17	75.9843	39.89043	9.831878	0.5854	0.001012	INPUT	Replacement...	-0.00825	0.00825	1
TRAIN	y	yes	REP_age	37	0	2319	17	75.9843	40.55411	13.22717	0.902445	0.188287	INPUT	Replacement...	0.00825	0.00825	2

After applying capping and flooring techniques to mitigate extreme values and reduce skewness in the dataset, we observed that some degree of asymmetry persists.

4. **“campaign” - “REP_campaign”**:
    a. “no”: from 4.031323 to 1.822747
    b. “yes”: from 3.19397 to 2.237525
5. **“duration” - “REP_duration”**:
    a. “no”: from 6.775871 to 2.705579
    b. “yes”: from 1.726539 to 0.948574
6. **“age” - “REP_age”**:
    a. “yes”: from 1.047476 to 0.902445

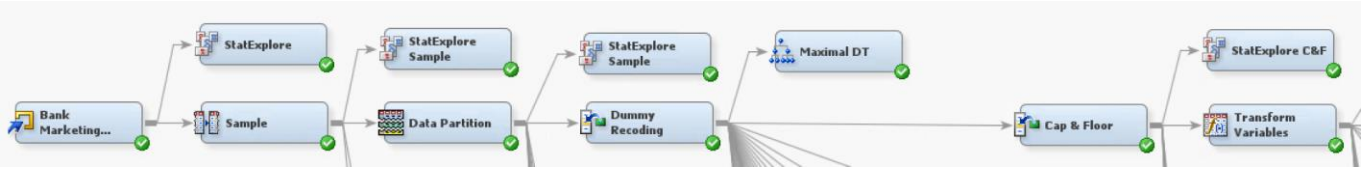
To further enhance the symmetry of the data and meet the assumptions of certain statistical models, a logarithmic transformation will be applied.

Logarithmic Transformation

For our analysis we used the natural logarithm “log” in the settings of SAS Enterprise Miner.

1. **Reducing Extreme Influence.** Transforming input variables using techniques such as logarithmic transformations aims to reduce their disproportionate impact on the model and get a more balanced impact on the overall model.
2. **Resolving Overfitting and Underfitting.** Logarithmic transformations aid in better capturing relationships between variables accurately. This helps strike a balance: the model can detect essential patterns while being less affected by extreme values. By doing so, this transformation helps prevent overfitting and underfitting as well.
3. **Improved Model Interpretability**. Such transformations facilitate a clearer understanding of the associations between variables, increasing the interpretability of the model and contributing to more accurate predictions and conclusions.

Picture 46. Adding Transform Variables Node into the Model



By adding the Transform Variables Node into the Model, we utilized the “Log” function within the “Edit Variables” option specifically for the “REP_campaign” and “REP_duration” variables. This step became necessary as, even after applying the “Cap and Floor” technique, these variables exhibited skewness, prompting the application of the “Log” transformation. By applying a logarithmic function, we achieved a significant reduction in skewness.

Results of utilization the log transformation function:

1. **“REP_campaign” - “LOG REP_campaign”**:
    a. “no”: from 1.822747 to 0.930683
    b. “yes”: from 2.237525 to 1.115448
2. **“REP_duration” - “LOG REP_duration”**:
    a. “no”: from 2.705579 to -0.53269
    b. “yes”: from 0.948574 to -0.1862

Picture 47. Applying log transformation function

Name	Method	Number of Bins	Role	Level
REP_age	**Default**	4	Input	Interval
REP_campaign	**Log**	4	Input	Interval
REP_day_of_week	**Default**	4	Input	Nominal
REP_duration	**Log**	4	Input	Interval
REP_education	**Default**	4	Input	Nominal
REP_job	**Default**	4	Input	Nominal
REP_month	**Default**	4	Input	Nominal
REP_pdays	**Default**	4	Input	Nominal
age	**Default**	4	Rejected	Interval
campaign	**Default**	4	Rejected	Interval
cons_conf_idx	**Default**	4	Rejected	Interval
cons_price_idx	**Default**	4	Rejected	Interval
contact	**Default**	4	Input	Nominal
day_of_week	**Default**	4	Rejected	Nominal
default	**Default**	4	Input	Nominal
duration	**Default**	4	Rejected	Interval
education	**Default**	4	Rejected	Nominal
emp_var_rate	**Default**	4	Rejected	Interval
euribor3m	**Default**	4	Rejected	Interval
housing	**Default**	4	Input	Nominal
job	**Default**	4	Rejected	Nominal
loan	**Default**	4	Input	Nominal
marital	**Default**	4	Input	Nominal
month	**Default**	4	Rejected	Nominal
nr_employed	**Default**	4	Rejected	Interval
pdays	**Default**	4	Rejected	Nominal
poutcome	**Default**	4	Input	Nominal
previous	**Default**	4	Rejected	Interval
y	**Default**	4	Target	Binary

Picture 48. Summary Statistics for Transform Variables Node (StatExplore)

Data Role	Target	Target Level	Variable	Skewness	Median	Missing	Non Missing	Minimum	Maximum	Mean	Standard Deviation	Kurtosis
TRAIN	y	no	LOG_REP_duration	-0.53269	5.135798	0	2319	1.609438	7.296398	5.071693	0.857093	1.041649
TRAIN	y	yes	LOG_REP_duration	-0.1862	6.118097	0	2319	3.637586	7.296398	6.072186	0.689736	-0.72828
TRAIN	y	no	LOG_REP_campaign	0.930683	1.098612	0	2319	0.693147	2.294185	1.110067	0.460273	-0.01415
TRAIN	y	yes	LOG_REP_campaign	1.115448	0.693147	0	2319	0.693147	2.294185	1.012838	0.388218	0.662705
TRAIN	y	no	REP_age	0.5854	38	0	2319	17	75.9843	39.89043	9.831878	0.001012
TRAIN	y	yes	REP_age	0.902445	37	0	2319	17	75.9843	40.55411	13.22717	0.188287

After utilizing this approach, it was noticed that positive skewness for “REP_durations” transformed into negative. The transformation of a positively skewed variable (skewed to the right) using a logarithmic function can result in a negatively skewed distribution.

1. **Skewness Alteration**. The logarithmic transformation compresses larger values more strongly than smaller values. When applied to a positively skewed distribution with a tail pointing toward larger values, this transformation compresses larger values more tightly than smaller values. As a result, this can lead to a redistribution of data toward the left side of the distribution, resulting in a change from positive to negative skewness.
2. **Impact on Distribution**. In a positively skewed distribution, most values tend to be smaller, while a few larger values lengthen the tail to the right. The logarithmic transformation reduces the difference between these larger values, which "pulls" the data toward smaller values, affecting the shape of the distribution and potentially causing reverse skewness.
3. **Skewness value**. A strongly skewed distribution is likely to have a more pronounced change in skewness after transformation

It should be noted separately that after applying the logarithmic transformation there may be difficulties with interpretation of values. Larger original values are compressed more than smaller ones. To reverse this transformation and obtain the original value, we should compute the antilog using the exponential function. While log-transformed values might make certain relationships or patterns more apparent, especially in linear relationships, we should remember to interpret the findings in the context of the log-transformed scale.

After the preparations of our data, we could come to the regression analysis.

REGRESSION

Regression is a statistical modeling technique used to investigate the relationship between a dependent variable (also known as the response or outcome variable) and one or more independent variables (predictors or features). The primary goal of regression analysis is to understand the nature of this relationship and use it for making predictions or inferences.

1. **Understanding Relationships**. It helps to understand how the value of the dependent variable changes when one or more independent variables change. It assists in determining which independent variables are statistically significant in predicting the dependent variable.
2. **Prediction**. This forecasting capability is particularly useful when you have historical data and want to predict future results.
3. **Feature Importance**. It quantifies the influence or impact of each feature on the outcome.

The features used in regression analysis are the independent variables that are considered to predict the dependent variable. Ideally, these features should be relevant, meaningful and have a logical or theoretical justification for their inclusion in the model. Also these variables should be prepared before applying regression analysis: no missing values, no skewness or outliers. The selection of attributes should be based on knowledge of the subject area, exploratory data analysis, and statistical methods that assess the significance and relevance of the variables in relation to the target variable.

For utilizing our analysis, we used the following methods that are used for variable selection in SAS Enterprise Miner:

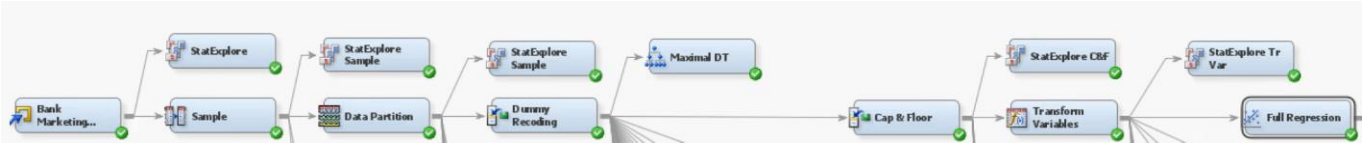
1. Full Regression.
2. Backward Regression.
3. Forward Regression.
4. Stepwise Regression.

Full regression

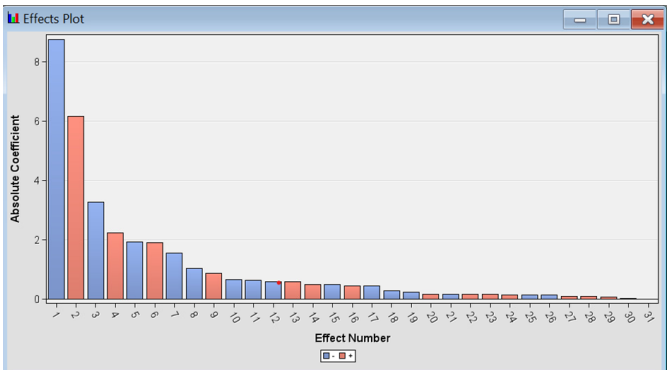
In full regression, all available independent variables are included in the regression model at the same time. This method considers all predictor variables without any selection or exclusion. Although it ensures that no variable is omitted, it may lead to multicollinearity problems or include irrelevant variables that do not contribute significantly to the model.

We added the Regression Node into our Model without changing any changes in its setting (Logistic Regression).

Picture 49. Adding the Full RegressionNode into the Model



Picture 50. The Full Regression Model, the distribution of effects of independent variables



Type 3 Analysis of Effects.

Picture 51. Type 3 Analysis of Effects

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
LOG_REP_campaign	1	35.8633	<.0001
LOG_REP_duration	1	1014.2911	<.0001
REP_age	1	5.2032	0.0225
REP_day_of_week	1	1.8955	0.1686
REP_education	3	25.1788	<.0001
REP_job	5	55.4509	<.0001
REP_month	3	108.0107	<.0001
REP_pdays	4	10.7542	0.0295
contact	1	67.5166	<.0001
default	1	59.5829	<.0001
housing	2	2.5648	0.2774
loan	1	2.6153	0.1058
marital	3	21.5749	<.0001
poutcome	2	8.8356	0.0121

Significant Effects. Several variables have a significant impact on the target variable ("y") based on the Wald Chi-Square test. Notably, "LOG_REP_duration" and "LOG_REP_campaign" have extremely low p-values (<0.0001), indicating strong significance.

Picture 52. Analysis of Maximum Likelihood Estimates

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Standardized Estimate	Exp(Est)
Intercept		1	-8.7650	7.8203	1.26	0.2624		0.000
LOG_REP_campaign		1	-0.6198	0.1035	35.86	<.0001	-0.1464	0.538
LOG_REP_duration		1	2.2349	0.0702	1014.29	<.0001	1.1396	9.346
REP_age		1	0.0112	0.00489	5.20	0.0225	0.0717	1.011
REP_day_of_week	Fr	1	0.0750	0.0545	1.90	0.1686		1.078
REP_education	Basic	1	-0.2756	0.0881	9.78	0.0018		0.759
REP_education	HS	1	-0.2210	0.0889	6.17	0.0130		0.802
REP_education	Un/Pr	1	0.1613	0.0763	4.47	0.0344		1.175
REP_job	B/T/H	1	-0.5791	0.1095	27.98	<.0001		0.560
REP_job	E/S/M	1	-0.4403	0.1298	11.50	0.0007		0.644
REP_job	Ret	1	0.8524	0.1918	19.75	<.0001		2.345
REP_job	S/S/A	1	-0.1562	0.1112	1.97	0.1602		0.855
REP_job	Unemployed	1	0.1425	0.2134	0.45	0.5044		1.153
REP_month	Q1	1	1.8902	0.2353	64.52	<.0001		6.621
REP_month	Q2	1	-0.6519	0.0991	43.24	<.0001		0.521
REP_month	Q3	1	-1.0349	0.1014	104.23	<.0001		0.355
REP_pdays	0	1	-3.2572	7.8103	0.17	0.6766		0.038
REP_pdays	1	1	-1.5359	7.8068	0.04	0.8440		0.215
REP_pdays	2	1	-1.9148	7.8106	0.06	0.8063		0.147
REP_pdays	3	1	6.1510	31.1641	0.04	0.8435		469.186
contact	cellular	1	0.4449	0.0541	67.52	<.0001		1.560
default	no	1	0.4923	0.0638	59.58	<.0001		1.636
housing	no	1	-0.1324	0.1021	1.68	0.1946		0.876
housing	unknown	1	0.1462	0.1843	0.63	0.4276		1.157
loan	no	1	0.0934	0.0577	2.62	0.1058		1.098
loan	unknown	0	0	.	.	.	.	.
marital	divorced	1	0.0614	0.2327	0.07	0.7920		1.063
marital	married	1	0.1353	0.2160	0.39	0.5311		1.145
marital	single	1	0.5782	0.2200	6.91	0.0086		1.783
poutcome	failure	1	-0.1296	0.1938	0.45	0.5037		0.878
poutcome	nonexistent	1	-0.4729	0.2038	5.39	0.0203		0.623

1. <u>Significant Predictors:</u>

- "duration": Highly significant (p-value < 0.0001) with a positive coefficient. This suggests that longer call durations are associated with a higher likelihood of success.
- "LOG_REP_campaign": More campaign contacts decrease the odds of subscription.
- "REP_job": Occupation significantly affects subscription odds.
- "REP_month": Quarters of the year impact subscription odds.
- "contact", "default", "marital", "poutcome": Communication type, credit status, marital status, and previous campaign outcome significantly influence subscription odds.

2. <u>Less Significant Predictors</u>:

- REP_age: (p = 0.0225) Older clients have slightly higher odds of subscription.
- REP_education: (Significance: Varies) Education level impacts subscription odds.
- REP_pdays: (p = 0.0295) Days since the last contact moderately impact subscription odds.
- "housing", "loan", "REP_day_of_week": (Moderate to low) Housing loan, personal loan, and the day of the week of the last contact have a moderate to low impact on subscription odds.
- "pdays" categories (0, 1, 2, 3): Significance: Varies - The re coded categories for the number of weeks since the last contact have mixed impacts.

Picture 53. Full Regression Node, fit statistics

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation
V	V	AIC	Akaike's Information Criterion	3732.271	
V	V	ASE	Average Squared Error	0.124802	0.125833
V	V	AVERR	Average Error Function	0.39589	0.396983
V	V	DFE	Degrees of Freedom for Error	4608	.
V	V	DFM	Model Degrees of Freedom	30	.
V	V	DFT	Total Degrees of Freedom	4638	
V	V	DIV	Divisor for ASE	9276	9284
V	V	ERR	Error Function	3672.271	3685.59
V	V	FPE	Final Prediction Error	0.126427	.
V	V	MAX	Maximum Absolute Error	0.999467	0.999978
V	V	MSE	Mean Square Error	0.125614	0.125833
V	V	NOBS	Sum of Frequencies	4638	4642
V	V	NW	Number of Estimate Weights	30	.
V	V	RASE	Root Average Sum of Squares	0.353273	0.35473
V	V	RFPE	Root Final Prediction Error	0.355565	.
V	V	RMSE	Root Mean Squared Error	0.354421	0.35473
V	V	SBC	Schwarz's Bayesian Criterion	3925.532	.
V	V	SSE	Sum of Squared Errors	1157.661	1168.235
V	V	SUMW	Sum of Case Weights Times Freq	9276	9284
V	V	MISC	Misclassification Rate	0.174213	0.182034

To evaluate the goodness of fit and performance of the model we explored Average Square Error and Misclassification Rate for validation data.

Average Squared Error (ASE)

It measures the squared difference between predicted probabilities and actual outcomes. Lower ASE values indicate better model performance.

ASE (Train): 0.124802

ASE (Validation): 0.125833

The slight increase from training to validation ASE suggests that the model may slightly overfit the training data.

Misclassification Rate

Misclassification Rate is the proportion of incorrectly predicted instances. Lower misclassification rates indicate better model accuracy.

Misclassification Rate (Training): 0.174213

Misclassification Rate (Validation): 0.182034

Odds Ratio Estimates

Picture 54. Full Regression Node, Odds Ratio Estimates

Odds Ratio Estimates

Effect		Point Estimate
LOG_REP_campaign		0.538
LOG_REP_duration		9.346
REP_age		1.011
REP_day_of_week	Fr vs M/T/W/Th	1.162
REP_education	Basic vs Unw/Il	0.543
REP_education	HS vs Unw/Il	0.573
REP_education	Un/Pr vs Unw/Il	0.840
REP_job	B/T/H vs Unknown	0.468
REP_job	E/S/M vs Unknown	0.537
REP_job	Ret vs Unknown	1.958
REP_job	S/S/A vs Unknown	0.714
REP_job	Unemployed vs Unknown	0.963
REP_month	Q1 vs Q4	8.114
REP_month	Q2 vs Q4	0.639
REP_month	Q3 vs Q4	0.435
REP_pdays	0 vs 4	0.022
REP_pdays	1 vs 4	0.123
REP_pdays	2 vs 4	0.084
REP_pdays	3 vs 4	268.815
contact	cellular vs telephone	2.435
default	no vs unknown	2.676
housing	no vs yes	0.888
housing	unknown vs yes	1.173
loan	no vs yes	1.205
loan	unknown vs yes	.
marital	divorced vs unknown	2.308
marital	married vs unknown	2.485
marital	single vs unknown	3.869
poutcome	failure vs success	0.481
poutcome	nonexistent vs success	0.341

Table 10. Odds Results of Full Regression explanation

Variable	Recoded Variable	Odds Ratio	Percentage Change Interpretation
LOG_REP_duration	-	9.346	For a one-unit increase in log-transformed duration, the odds of subscription increase by approximately 834.6%. The high odds ratio for 'LOG_REP_duration' underscores its importance. Longer contact durations significantly increase the odds of subscription.
LOG_REP_campaign	-	0.538	For a one-unit increase in log-transformed number of contacts during the campaign, the odds of subscription decrease by approximately 46.2%.
contact: cellular	Cellular vs Telephone	2.435	Being contacted via cellular communication increases the odds of subscription by approximately 143.5% higher compared to being contacted by telephone.
REP_age	-	1.011	A 1% increase in age is associated with a 1.1% increase in the odds of subscription.
REP_day_of_week	M/T/W/Th	1.162	Odds of subscription are 16.2% higher on Fri as compared to Mon, Tue, Wed.

REP_education	Basic	0.543	Odds of subscription are 45.7% lower for Basic education compared to Un/Il.
REP_education	HS	0.573	Having a high school education lower the odds of subscription by approximately 42.7% compared to being unknown or illiterate.
REP_education	Un/Pr	0.84	Odds of subscription are 16.0% higher for University/Professional education compared to Un/Il.
REP_job	B/T/H	0.468	Odds of subscription are 53.2% lower for Blue-collar/Technician/Housemaid jobs compared to Unknown.
REP_job	E/S/M	0.537	Odds of subscription are 46.3% lower for Entrepreneur/Self-employed/Management jobs compared to Unknown.
REP_job	Ret	1.958	Being retired increases the odds of subscription by approximately 95.8% higher compared to being unknown.
REP_job	S/S/A	0.714	Odds of subscription are 28.6% lower for Student/Services/Admin jobs compared to Unknown.
REP_job	Unempl	0.963	Odds of subscription are 3.7% lower for Unemployed individuals compared to Unknown.
REP_month	Q1	8.114	Being contacted in the first quarter increases the odds of subscription by approximately 711.4% higher compared to the fourth quarter
REP_month	Q2	0.639	Odds of subscription are 36.1% lower in Q2 compared to Q4.
REP_month	Q3	0.435	Odds of subscription are 56.5% lower in Q3 compared to Q4.
REP_pdays	0	0.022	Odds of subscription are 97.8% lower for clients never contacted compared to 3 weeks since last contact.
REP_pdays	1	0.123	Odds of subscription are 87.7% lower for clients contacted 1 week ago compared to 3 weeks since last contact.
REP_pdays	2	0.084	Odds of subscription are 91.6% lower for clients contacted 2 weeks ago compared to 3 weeks since last contact.
REP_pdays	3	26881.5	Odds of subscription are 26881.5% higher for clients contacted 3 weeks ago compared to 3 weeks since last contact.
contact	Cellular	2.435	Odds of subscription are 143.5% higher for cellular communication compared to telephone.
default	No	2.676	Not having credit in default increases the odds of subscription by approximately 167.6% than having credit in default.
housing	No	0.888	Odds of subscription are 11.2% lower for clients without a housing loan compared to with.
housing	Unknown	1.173	Odds of subscription are 17.3% higher for clients with unknown housing status compared to with.
loan	No	1.205	Not having a personal loan increases the odds of subscription by approximately 20.5% compared to having a personal loan.
marital	Divorced	1.063	Odds of subscription are 6.3% higher for divorced clients compared to unknown.
marital	Married	1.145	Odds of subscription are 14.5% higher for married clients compared to unknown.
marital	Single	3.869	Being single increases the odds of subscription by 286.9% than for the married.

poutcome	Failure	0.481	Odds of subscription are 51.9% lower for clients with a previous marketing failure compared to success.
poutcome	Nonexistent	0.341	Odds of subscription are 65.9% lower for clients with no previous marketing outcome compared to success.

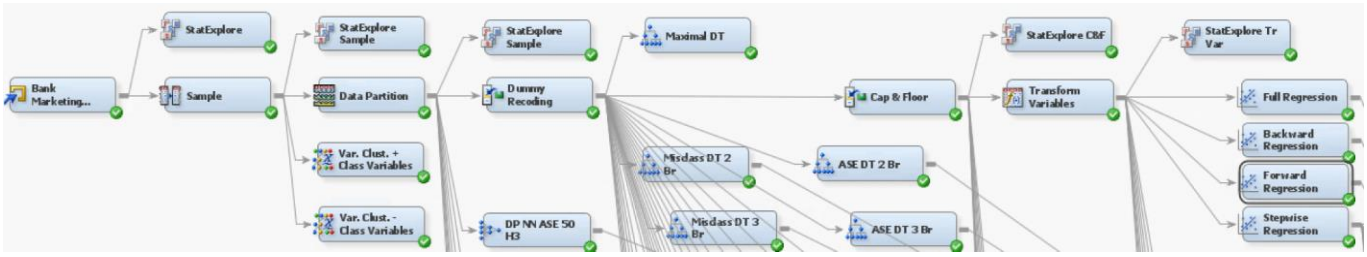
Findings

Based on the results of the full regression model, we suggest client to longer and engaging interactions ("LOG_REP_duration") as it stands out as a crucial factor positively influencing the likelihood of subscription. Additionally, the choice of communication channel ("contact") plays a significant role, with the cellular channel the effectiveness of telemarketing campaigns targeting term deposit subscriptions increases, so efforts should be intensified through this medium. Furthermore, special focus on clients with no credit in default ("defaut_no") and single marital status "'marital_status") could also prove beneficial, suggesting tailored marketing approaches to unique preferences and behaviours of this demographic.

Stepwise and Forward Regression

Forward regression is a variable selection technique used in statistical modeling to build a predictive model. The process involves gradually adding variables to the model one at a time, starting with the most statistically significant or relevant variable whereas Stepwise regression is an extension of forward regression that involves both adding and removing variables from the model.

Picture 55. Adding Forward and Stepwise Regression Nodes into the Model



Picture 56. Forward Regression Node Settings

Model Selection	
Selection Model	Forward
Selection Criterion	Validation Error
Use Selection Defaults	Yes
Selection Options	

We added the Forward Regression and Stepwise Nodes into our model and utilized the Selection Model "Forward" and "Stepwise" accordingly and Selection Criterion "Validation Error" for both of them.

We added the Forward Regression and Stepwise Nodes into our model and utilized the Selection Model “Forward” and “Stepwise” accordingly and Selection Criterion “Validation Error” for both of them. The goal of using “Validation Error” is to select the subset of variables that minimizes the error in predicting outcomes on data not used for model training (validation data). Using Validation Error as a selection criterion helps prevent overfitting by evaluating the model's performance on independent data, ensuring that the selected variables are truly predictive and do not simply reflect noise or random associations present in the training data. It aims to find a balance between model complexity and predictive accuracy.

Picture 57. Stepwise Regression Node Settings

Model Selection	
Selection Model	Stepwise
Selection Criterion	Validation Error
Use Selection Defaults	Yes
Selection Options	...

The process includes fitting the model to a portion of the dataset (training set) and then evaluating its performance on a separate dataset (validation set).

The variables that have the lowest validation error or the best predictive performance are selected to build the final model. By selecting variables based on their ability to minimize validation error, the forward selection method seeks to create a simple model that generalizes well to new, unknown data, thereby improving the model's applicability and predictive ability.

Picture 58. Summary of Forward Selection

Summary of Forward Selection

Step	Effect Entered	DF	Number In	Score Chi-Square	Pr > ChiSq	Validation Error Rate
1	LOG_REP_duration	1	1	1357.2272	<.0001	4822.4
2	poutcome	2	2	462.7507	<.0001	4252.2
3	REP_month	3	3	186.5347	<.0001	4029.8
4	contact	1	4	112.4899	<.0001	3874.5
5	default	1	5	81.2877	<.0001	3816.6
6	REP_job	5	6	79.3614	<.0001	3705.3
7	LOG_REP_campaign	1	7	35.3271	<.0001	3693.5
8	REP_education	3	8	30.2826	<.0001	3680.4
9	marital	3	9	17.4568	0.0006	3665.4
10	REP_pdays	4	10	14.6568	0.0055	3679.7
11	REP_age	1	11	5.0411	0.0248	3684.6

The selected model, based on the error rate for the validation data, is the model trained in Step 9. It consists of the following effects:

Intercept LOG_REP_campaign LOG_REP_duration REP_education REP_job REP_month contact default marital poutcome

Picture 59. Summary of Stepwise Selection

Summary of Stepwise Selection

Step	Effect Entered	DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq	Validation Error Rate
1	LOG_REP_duration	1	1	1357.2272		<.0001	4822.4
2	poutcome	2	2	462.7507		<.0001	4252.2
3	REP_month	3	3	186.5347		<.0001	4029.8
4	contact	1	4	112.4899		<.0001	3874.5
5	default	1	5	81.2877		<.0001	3816.6
6	REP_job	5	6	79.3614		<.0001	3705.3
7	LOG_REP_campaign	1	7	35.3271		<.0001	3693.5
8	REP_education	3	8	30.2826		<.0001	3680.4
9	marital	3	9	17.4568		0.0006	3665.4
10	REP_pdays	4	10	14.6568		0.0055	3679.7
11	REP_age	1	11	5.0411		0.0248	3684.6

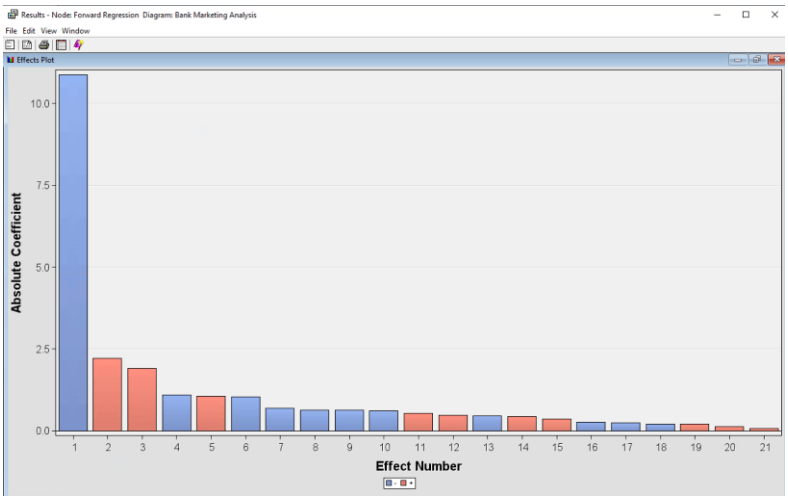
The selected model, based on the error rate for the validation data, is the model trained in Step 9. It consists of the following effects:

Intercept LOG_REP_campaign LOG_REP_duration REP_education REP_job REP_month contact default marital poutcome

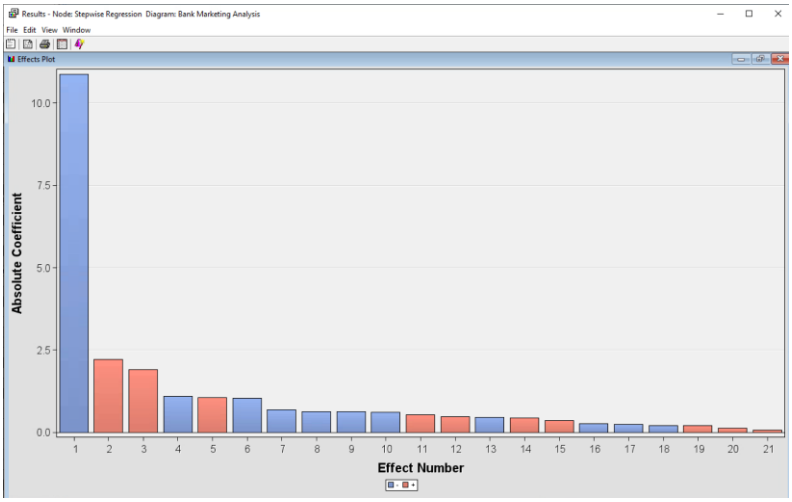
The results of Forward and Stepwise selection are identical; it indicates that both procedures have independently arrived at the same optimal set of predictor variables for the model. In this scenario, the selection of the stepwise model is justified, and we proceed with the stepwise model.

The Stepwise Regression Model underwent a selection process resulting in a model with 11 significant effects. Starting with “LOG_REP_duration” in Step 1, the model sequentially included variables such as “poutcome”, “REP_month”, “contact”, “default”, “REP_job”, ‘LOG_REP_campaign”, “REP_education”, “marital”, and “REP_pdays”.

Picture 60. The Forward Regression Model, the distribution of effects of independent variables



Picture 61. The Stepwise Regression Model, the distribution of effects of independent variables



Type 3 Analysis of Effects:

Picture 62. Type 3 Analysis of Effects

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
LOG_REP_campaign	1	36.3262	<.0001
LOG_REP_duration	1	1016.8019	<.0001
REP_education	3	25.1380	<.0001
REP_job	5	83.7686	<.0001
REP_month	3	110.7591	<.0001
contact	1	65.6277	<.0001
default	1	56.7052	<.0001
marital	3	17.2934	0.0006
poutcome	2	203.2882	<.0001

Significant Predictors:

“LOG_REP_campaign”: Chi-Square = 36.3262, $p < 0.0001$

“LOG_REP_duration”: Chi-Square = 1016.8019, $p < 0.0001$

Other significant predictors related to “education”, “job”, “REP_month”, “contact”, “default”, “marital”, and “poutcome”.

Picture 63. Analysis of Maximum Likelihood Estimates

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Standardized Estimate	Exp(Est)
Intercept		1	-10.8879	0.4559	570.35	<.0001		0.000
LOG_REP_campaign		1	-0.6213	0.1031	36.33	<.0001	-0.1468	0.537
LOG_REP_duration		1	2.2206	0.0696	1016.80	<.0001	1.1323	9.212
REP_education	Basic	1	-0.2699	0.0878	9.45	0.0021		0.763
REP_education	HS	1	-0.2412	0.0882	7.47	0.0063		0.786
REP_education	Un/Pr	1	0.1438	0.0757	3.61	0.0576		1.155
REP_job	B/T/H	1	-0.6246	0.1062	34.57	<.0001		0.535
REP_job	E/S/M	1	-0.4617	0.1281	12.98	0.0003		0.630
REP_job	Ret	1	1.0540	0.1709	38.04	<.0001		2.869
REP_job	S/S/A	1	-0.2125	0.1070	3.95	0.0470		0.809
REP_job	Unemployed	1	0.0657	0.2112	0.10	0.7558		1.068
REP_month	Q1	1	1.9100	0.2339	66.70	<.0001		6.753
REP_month	Q2	1	-0.6872	0.0984	48.81	<.0001		0.503
REP_month	Q3	1	-1.0305	0.1006	104.86	<.0001		0.357
contact	cellular	1	0.4360	0.0538	65.63	<.0001		1.547
default	no	1	0.4738	0.0629	56.71	<.0001		1.606
marital	divorced	1	0.1218	0.2328	0.27	0.6007		1.130
marital	married	1	0.1934	0.2175	0.79	0.3739		1.213
marital	single	1	0.5409	0.2209	5.99	0.0143		1.718
poutcome	failure	1	-0.6141	0.1008	37.15	<.0001		0.541
poutcome	nonexistent	1	-1.0858	0.0822	174.45	<.0001		0.338

Picture 64. Fit Statistics for both Forward and Stepwise Regression Models

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation
V	V	AIC	Akaike's Information Criterion	3742.779	.
V	V	ASE	Average Squared Error	0.125979	0.126304
V	V	AVERR	Average Error Function	0.398963	0.39481
V	V	DFE	Degrees of Freedom for Error	4617	.
V	V	DFM	Model Degrees of Freedom	21	.
V	V	DFT	Total Degrees of Freedom	4638	.
V	V	DIV	Divisor for ASE	9276	9284
V	V	ERR	Error Function	3700.779	3665.413
V	V	FPE	Final Prediction Error	0.127125	.
V	V	MAX	Maximum Absolute Error	0.999412	0.998773
V	V	MSE	Mean Square Error	0.126552	0.126304
V	V	NOBS	Sum of Frequencies	4638	4642
V	V	NW	Number of Estimate Weights	21	.
V	V	RASE	Root Average Sum of Squares	0.354935	0.355393
V	V	RFPE	Root Final Prediction Error	0.356545	.
V	V	RMSE	Root Mean Squared Error	0.355741	0.355393
V	V	SBC	Schwarz's Bayesian Criterion	3878.062	.
V	V	SSE	Sum of Squared Errors	1168.578	1172.606
V	V	SUMW	Sum of Case Weights Times Freq	9276	9284
V	V	MISC	Misclassification Rate	0.176154	0.182464

Average Squared Error (ASE)

Training ASE: 0.125979

The training ASE is a measure of the model's performance on the data it was trained on.

Validation ASE: 0.126304

The validation ASE is a measure of how well the model generalizes to new, unseen data.

ASE is a measure of the model's predictive accuracy. A lower ASE indicates better predictive performance. The ASE values for both training and validation data are relatively close, so, it suggests that the model is not overfitting or underfitting.

Misclassification Error

Training Misclassification Error: 0.176154

The proportion of misclassified instances in the training set. It represents the percentage of cases where the model predictions do not match the actual values observed in the training data.

Validation Misclassification Error: 0.182464

The proportion of misclassified instances in the validation set. It shows the percentage of cases where the model predictions do not match the true observed values in the validation data that the model did not see during training.

A lower Misclassification Error Rate is preferable, indicating better predictive performance. In this case, the training Misclassification Error (0.176154) is slightly lower than the validation Misclassification Error (0.182464). The results show that the model's performance in the validation set is slightly worse than when dealing with training data. The relatively close Misclassification Error Rate between the training and validation sets suggests that the model generalizes reasonably well, but may show a slight decrease in performance when dealing with new, unseen data.

Odds Ratio Estimates

Picture 65. Odds Ratio Estimates for both Forward and Stepwise Regression Models

Odds Ratio Estimates

Effect		Point Estimate
LOG_REP_campaign		0.537
LOG_REP_duration		9.212
REP_education	Basic vs Unw/Il	0.529
REP_education	HS vs Unw/Il	0.544
REP_education	Un/Pr vs Unw/Il	0.800
REP_job	B/T/H vs Unknown	0.448
REP_job	E/S/M vs Unknown	0.527
REP_job	Ret vs Unknown	2.399
REP_job	S/S/A vs Unknown	0.676
REP_job	Unemployed vs Unknown	0.893
REP_month	Q1 vs Q4	8.185
REP_month	Q2 vs Q4	0.610
REP_month	Q3 vs Q4	0.432
contact	cellular vs telephone	2.392
default	no vs unknown	2.579
marital	divorced vs unknown	2.659
marital	married vs unknown	2.857
marital	single vs unknown	4.043
poutcome	failure vs success	0.099
poutcome	nonexistent vs success	0.062

Table 11. Odds Results of Stepwise Regression explanation

Variable	Odds Ratio	Interpretation
LOG_REP_campaign	0.537	An additional contact reduces the odds of subscription by 46.3%, suggesting diminishing returns with excessive contacts.
LOG_REP_duration	9.212	Each additional second of contact significantly increases the odds of subscription by 821.2%, highlighting the positive influence of longer interactions.
REP_education (Basic vs Unw/Il)	0.529	Clients with basic and high school education have 52.9% and 54.4% lower odds of subscribing compared to those with unknown education.

REP_education (HS vs Unw/Il)	0.544	
REP_education (Un/Pr vs Unw/Il)	0.8	University/professional education shows 20% lower odds.
REP_job (B/T/H vs Unknown)	0.448	Specific occupations impact subscription odds differently compared to unknown occupations, providing nuanced insights into the influence of job types.
		55.2% lower odds for specific jobs compared to unknown
REP_job (E/S/M vs Unknown)	0.527	47.3% lower odds for specific jobs compared to unknown
REP_job (Ret vs Unknown)	2.399	139.9% higher odds for retired clients compared to unknown
REP_job (S/S/A vs Unknown)	0.676	32.4% lower odds for specific jobs compared to unknown
REP_job (Unemployed vs Unknown)	0.893	10.7% lower odds for unemployed clients compared to unknown
REP_month (Q1 vs Q4)	8.185	Odds are 718.5% higher for Q1 compared to Q4. Odds are significantly higher for Q1 and lower for Q2 and Q3 compared to Q4, indicating a potential seasonality effect.
REP_month (Q2 vs Q4)	0.61	39.0% lower odds for Q2 compared to Q4
REP_month (Q3 vs Q4)	0.432	56.8% lower odds for Q3 compared to Q4
contact (cellular vs telephone)	2.392	The odds of subscription are 139.2% higher when the communication type is cellular compared to telephone.
default (no vs unknown)	2.579	Clients with no credit default have 157.9% higher odds of subscribing compared to those with an unknown credit default status.
marital (divorced vs unknown)	2.659	Divorced, married, and single clients have higher odds of subscription compared to those with an unknown marital status. 165.9% higher odds for divorced clients compared to unknown
marital (married vs unknown)	2.857	185.7% higher odds for married clients compared to unknown
marital (single vs unknown)	4.043	304.3% higher odds for single clients compared to unknown
poutcome (failure vs success)	0.099	Clients with a previous campaign failure or no previous campaign have significantly lower odds of subscribing compared to those with a successful outcome.
		90.1% lower odds for previous campaign failure compared to success.
poutcome (nonexistent vs success)	0.062	93.8% lower odds for no previous campaign compared to success.

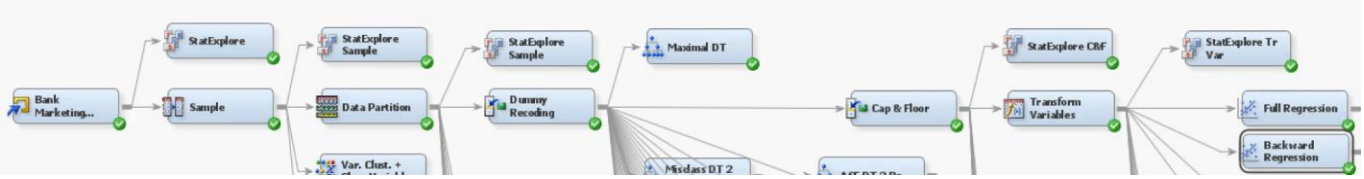
Findings

In summary, our analysis of the Stepwise/Forward Regression Models highlights key strategies for improving the bank's telemarketing success. We recommend focusing on longer calls, using cell phones for communication, and tailoring campaigns to customers with good credit. Targeting single individuals and learning from past campaign outcomes, especially failures, will fine-tune future strategies. Regularly monitoring and adapting predictive models ensures they stay effective amid changing customer behaviors. By following these suggestions, the bank can not only boost subscription rates but also streamline and improve the efficiency of their telemarketing efforts for overall success.

Backward Regression

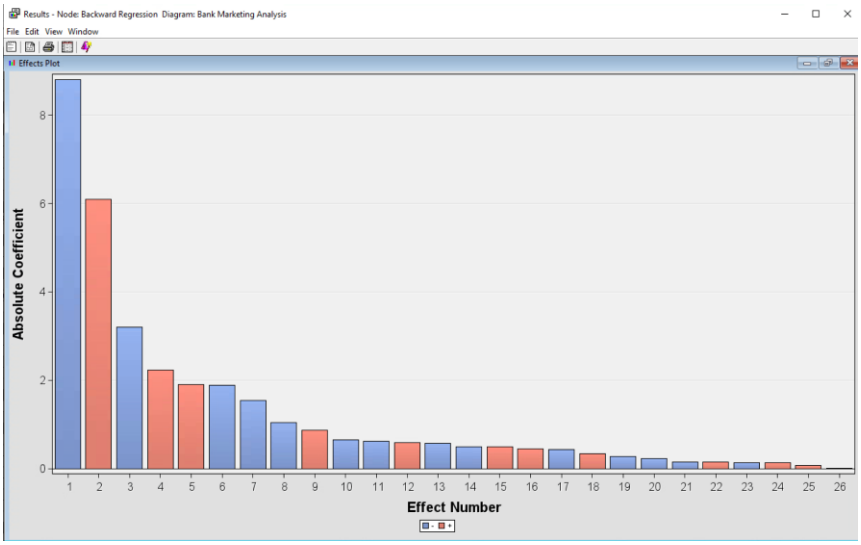
Backward Regression is a stepwise method used in statistical modeling, especially in linear regression. Its main purpose is to simplify the model by removing less important variables.

Picture 66. Adding Backward Regression Nodes into the Model



Variables with the highest p-values (indicating less significance) are removed.

Picture 67. The Backward Regression Model, the distribution of effects of independent variables



Picture 68. Summary of Backward Selection

Summary of Backward Elimination

Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq	Validation Error Rate
1	housing	2	13	2.5648	0.2774	3686.4
2	loan	2	12	2.7221	0.2564	3686.0
3	REP_day_of_week	1	11	1.8709	0.1714	3684.6

The selected model, based on the error rate for the validation data, is the model trained in Step 3. It consists of the following effects:

Intercept LOG_REP_campaign LOG_REP_duration REP_age REP_education REP_job REP_month REP_pdays contact default marital poutcome

The backward elimination process resulted in a model that includes variables considered significant based on p-values.

- “housing” (Chi-Square: 2.5648, p-value: 0.2774)
- “loan” (Chi-Square: 2.7221, p-value: 0.2564)
- “REP_day_of_week” (Chi-Square: 1.8709, p-value: 0.1714)

Selected Model contains variables:

“LOG_REP_campaign”, “LOG_REP_duration”, “REP_age”, “REP_education”, “REP_job”, “REP_month”, “REP_pdays”, “contact”, “default”, “marital”, “poutcome”.

Picture 69. Fit Statistics for Backward Regression Model

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation
y	y	AIC	Akaike's Information Criterion	3730.848	.
y	y	ASE	Average Squared Error	0.125131	0.125881
y	y	AVERR	Average Error Function	0.396599	0.396876
y	y	DFE	Degrees of Freedom for Error	4612	.
y	y	DFM	Model Degrees of Freedom	26	.
y	y	DFT	Total Degrees of Freedom	4638	.
y	y	DIV	Divisor for ASE	9276	9284
y	y	ERR	Error Function	3678.848	3684.595
y	y	FPE	Final Prediction Error	0.126541	.
y	y	MAX	Maximum Absolute Error	0.999431	0.999975
y	y	MSE	Mean Square Error	0.125836	0.125881
y	y	NOBS	Sum of Frequencies	4638	4642
y	y	NW	Number of Estimate Weights	26	.
y	y	RASE	Root Average Sum of Squares	0.353738	0.354797
y	y	RFPE	Root Final Prediction Error	0.355727	.
y	y	RMSE	Root Mean Squared Error	0.354734	0.354797
y	y	SBC	Schwarz's Bayesian Criterion	3898.341	.
y	y	SSE	Sum of Squared Errors	1160.711	1168.676
y	y	SUMW	Sum of Case Weights Times Fr...	9276	9284
y	y	MISC	Misclassification Rate	0.173351	0.181172

Average Squared Error (ASE)

Train ASE: 0.125131

It measures how well the model fits the training data.

Validation ASE: 0.125881

It assesses how well the model generalizes to new, unseen data that it hasn't encountered during training.

Lower ASE values indicate better prediction performance because they indicate that the model predictions are closer to the actual observed values. In this case, the ASE for training (0.125131) and ASE

for validation (0.125881) are relatively low and quite close to each other. The similarity between the ASE values in training and validation means that the model generalizes reasonably well to unseen data. This suggests that the model's performance remains consistent between the dataset on which it was trained and the new data, indicating a good balance in its ability to fit the training data while performing well on the new observations.

Misclassification Rate

Train Misclassification Rate: 0.173351

It indicates the percentage of cases where the model's predictions do not match the actual observed values in the training data.

Validation Misclassification Rate: 0.181172

It indicates the percentage of cases where the model's predictions do not align with the true observed values in the validation data, which the model hasn't seen during training.

A lower misclassification rate is preferable as it suggests better predictive performance. In this scenario, the training misclassification rate (0.173351) is slightly lower than the validation misclassification rate (0.181172). These results mean that the performance of the model in the validation set is slightly lower than when predicting on training data. However, both figures are relatively close, suggesting that the model generalizes reasonably well to new data, but may have slightly lower performance on unseen observations compared to training data.

Type 3 Analysis of Effects

Picture 70. Type 3 Analysis of Effects

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
LOG_REP_campaign	1	35.9786	<.0001
LOG_REP_duration	1	1014.9034	<.0001
REP_age	1	5.0322	0.0249
REP_education	3	24.8433	<.0001
REP_job	5	55.7113	<.0001
REP_month	3	109.4703	<.0001
REP_pdays	4	10.1933	0.0373
contact	1	67.7516	<.0001
default	1	59.0462	<.0001
marital	3	21.8523	<.0001
poutcome	2	9.5294	0.0085

Picture 71. Analysis of Maximum Likelihood Estimates

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Standardized Estimate	Exp(Est)
Intercept		1	-8.8037	7.8300	1.26	0.2609		0.000
LOG_REP_campaign		1	-0.6187	0.1032	35.98	<.0001	-0.1462	0.539
LOG_REP_duration		1	2.2287	0.0700	1014.90	<.0001	1.1364	9.288
REP_age		1	0.0109	0.00488	5.03	0.0249	0.0703	1.011
REP_education	Basic	1	-0.2777	0.0881	9.94	0.0016		0.758
REP_education	HS	1	-0.2198	0.0889	6.12	0.0134		0.803
REP_education	Un/Pr	1	0.1547	0.0761	4.13	0.0421		1.167
REP_job	B/T/H	1	-0.5761	0.1093	27.78	<.0001		0.562
REP_job	E/S/M	1	-0.4339	0.1297	11.20	0.0008		0.648
REP_job	Ret	1	0.8665	0.1912	20.53	<.0001		2.379
REP_job	S/S/A	1	-0.1563	0.1110	1.98	0.1592		0.855
REP_job	Unemployed	1	0.1348	0.2130	0.40	0.5270		1.144
REP_month	Q1	1	1.9020	0.2354	65.30	<.0001		6.700
REP_month	Q2	1	-0.6560	0.0991	43.85	<.0001		0.519
REP_month	Q3	1	-1.0406	0.1013	105.62	<.0001		0.353
REP_pdays	0	1	-3.2039	7.8206	0.17	0.6820		0.041
REP_pdays	1	1	-1.5372	7.8166	0.04	0.8441		0.215
REP_pdays	2	1	-1.8894	7.8205	0.06	0.8091		0.151
REP_pdays	3	1	6.0983	31.2061	0.04	0.8451		445.083
contact	cellular	1	0.4443	0.0540	67.75	<.0001		1.559
default	no	1	0.4884	0.0636	59.05	<.0001		1.630
marital	divorced	1	0.0672	0.2348	0.08	0.7748		1.069
marital	married	1	0.1422	0.2183	0.42	0.5148		1.153
marital	single	1	0.5868	0.2223	6.96	0.0083		1.798
poutcome	failure	1	-0.1404	0.1938	0.52	0.4689		0.869
poutcome	nonexistent	1	-0.4942	0.2039	5.88	0.0153		0.610

Significant Variables (p < 0.05):

"LOG_REP_campaign", "LOG_REP_duration", "REP_age", "REP_education", "REP_job", "REP_month", "REP_pdays", "contact", "default", "marital", "poutcome".

Odds Ratio Estimates

Picture 72. Odds Ratio Estimates for Backward Regression Models

Odds Ratio Estimates

Effect		Point Estimate
LOG_REP_campaign		0.539
LOG_REP_duration		9.288
REP_age		1.011
REP_education	Basic vs Unw/Il	0.538
REP_education	HS vs Unw/Il	0.570
REP_education	Un/Pr vs Unw/Il	0.828
REP_job	B/T/H vs Unknown	0.477
REP_job	E/S/M vs Unknown	0.549
REP_job	Ret vs Unknown	2.017
REP_job	S/S/A vs Unknown	0.725
REP_job	Unemployed vs Unknown	0.970
REP_month	Q1 vs Q4	8.227
REP_month	Q2 vs Q4	0.637
REP_month	Q3 vs Q4	0.434
REP_pdays	0 vs 4	0.024
REP_pdays	1 vs 4	0.126
REP_pdays	2 vs 4	0.089
REP_pdays	3 vs 4	261.394
contact	cellular vs telephone	2.432
default	no vs unknown	2.656
marital	divorced vs unknown	2.371
marital	married vs unknown	2.556
marital	single vs unknown	3.987
poutcome	failure vs success	0.461
poutcome	nonexistent vs success	0.323

Table 12. Odds Results of Backward Regression explanation

Variable	Category	Odds Ratio	Interpretation
LOG_REP_campaign	-	0.539	For each unit increase in contact in the current campaign, the odds of subscribing decrease by 46.1%.
LOG_REP_duration	-	9.288	For each additional second in the last contact duration, the odds of subscribing increase by an impressive 828.8%.
REP_age	-	1.011	With each additional year of age, the odds of subscribing increase by 1.1%.
REP_education	Basic	0.538	Clients with basic education have 46.2% lower odds of subscribing compared to those with unknown or illiterate education.
REP_education	HS	0.57	Clients with high school education have 43.0% lower odds of subscribing compared to those with unknown or illiterate education.
REP_education	Un/Pr	0.828	Clients with university or professional education have 17.2% lower odds of subscribing compared to those with unknown or illiterate education.
REP_job	B/T/H	0.477	Compared to individuals with an unknown job, retirees have 101.7% higher odds of subscribing, while those in blue-collar, entrepreneur/self-employed, and services/sales/administration jobs have lower odds by 52.3%, 45.1%, and 27.5%, respectively.
REP_job	E/S/M	0.549	
REP_job	Ret	2.017	
REP_job	S/S/A	0.725	
REP_job	Unemployed	0.97	
REP_month	Q1	8.227	Compared to the fourth quarter, subscribing odds increase by 722.7% in Q1 but decrease by 36.3% and 56.6% in Q2 and Q3, respectively.
REP_month	Q2	0.637	
REP_month	Q3	0.434	
REP_pdays	0	0.024	Odds of subscribing decrease by 97.6%, 87.4%, 91.1%, and increase dramatically by 26039.4% for clients contacted on the 1st, 2nd, and 3rd week compared to those not contacted in the previous campaign.
REP_pdays	1	0.126	
REP_pdays	2	0.089	
REP_pdays	3	261.394	
contact	Cellular	2.432	Clients contacted via cellular communication have 143.2% higher odds of subscribing compared to telephone communication.
default	No	2.656	Clients with no credit default have 165.6% higher odds of subscribing compared to those with unknown credit status.
marital	Divorced	2.371	Compared to clients with an unknown marital status, divorced, married, and single clients have higher odds of subscribing by 137.1%, 155.6%, and 298.7%, respectively.
marital	Married	2.556	
marital	Single	3.987	

poutcome	Failure	0.461	Clients with a previous marketing campaign failure have 53.9% lower odds of subscribing, and clients with no previous campaign have 67.7% lower odds compared to those with a successful previous campaign.
poutcome	Nonexistent	0.323	

Findings

In summary, the backward regression process includes significant variables, “LOG_REP_campaign”, “LOG_REP_duration”, “REP_age”, “REP_education”, “REP_job”, ‘REP_month”, “REP_pdays”, “contact”, “default”, “marital”, and “poutcome”. Noteworthy findings include the substantial impact of call duration (“LOG_REP_duration”) on subscription odds, the influence of education and job categories, and the importance of marital status and previous campaign outcomes.

NEURAL NETWORK

Neural networks are powerful tools for solving real-life problems by understanding complex connections between inputs and outputs. In most cases it is a natural extension of a regression model. They're great at figuring out hidden patterns, making broad predictions, and handling intricate relationships. This makes them highly valuable for our work in understanding and predicting outcomes in various situations.

In our study on bank marketing, we smoothly combined Neural Network analysis with Regression and Decision Trees to get a better grasp of our findings. We carefully tested different setups using 50 and 100 iterations and varying hidden units (3, 4, 5, and 6). By blending Neural Network analysis with tools like Data Partition, Cap and Floor, Transform Variables, and Stepwise Regression, we fine-tuned our results, aiming for the best outcomes with the least errors. Despite similar results with Stepwise and Forward Regression, we chose Stepwise for its simplicity. This thorough approach ensured a deep dive into our data, making our conclusions more accurate and reliable in the dynamic realm of bank marketing analysis. Also, we applied “Average Error” as a Model Selection Criterion for each Neural Network Node.

Picture 73. Network Settings for Neural Network Nodes

Network

Property	Value
Architecture	Multilayer Perceptron
Direct Connection	No
Number of Hidden Units	3
Randomization Distribution	Normal
Randomization Center	0.0
Randomization Scale	0.1
Input Standardization	Standard Deviation
Hidden Layer Combination Function	Default
Hidden Layer Activation Function	Default
Hidden Bias	Yes
Target Layer Combination Function	Default
Target Layer Activation Function	Default
Target Layer Error Function	Default
Target Bias	Yes
Weight Decay	0.0

Picture 74. Optimisation Settings for Neural Network Nodes

Optimization

Property	Value
Training Technique	Default
Maximum Iterations	50
Maximum Time	4 Hours
Nonlinear Options	
Use Defaults	Yes

Optimization

Property	Value
Training Technique	Default
Maximum Iterations	100
Maximum Time	4 Hours
Nonlinear Options	
Use Defaults	Yes

Picture 75. Enabling Preliminary Training

Preliminary Training	
Enable	No
Number of Runs	5
Maximum Iterations	10
Maximum Time	1 Hour

Using the average root mean square error (ASE) as a model selection criterion for neural network nodes is a common practice in model training. ASE is a measure that quantifies the mean squared difference between predicted and actually observed values in a data set.

1. **Accuracy Measure**. ASE is a measure of forecast accuracy. Minimizing ASE means that the model predictions are closer to the real values on average. A lower ASE value means better prediction performance.
2. **Model Optimization**. During the training process, the neural network seeks to minimize the ASE by adjusting its weights and biases. Therefore, selecting models with lower ASE helps in selecting models that have been better optimized and exhibit higher predictive ability.
3. **Overfitting Control**. By tracking the ASE on the training and validation datasets, it is possible to determine if the model is overfitting. If the difference between ASE for training and validation is significant, this may indicate overfitting, suggesting that the model performs well on training data but less well on new, unknown data.
4. **Comparative Evaluation**. ASE allows the comparison of different models or different configurations of the same model. Selecting models with lower ASE on validation dataset helps to select the most efficient models among all the options considered.

By changing number of Maximum Number of Iterations we took into account the main features of this setting in Neural Networks:

1. **Convergence**. The model might not have converged within the default or earlier specified number of iterations. By increasing the maximum iterations, the model is given more opportunities to converge, potentially improving its performance.
2. **Avoiding Premature Stopping.** Sometimes, the model may stop training before reaching an optimal solution due to reaching the iteration limit prematurely. Increasing the iterations helps prevent premature stopping, allowing the model to potentially achieve better performance.

Also, it should be noticed that changing the number of Hidden Units in a neural network can significantly impact its performance and representational capacity.

1. **Model Complexity**. Increasing the number of hidden layers increases the complexity of the network, allowing it to learn more complex patterns and representations from the data. Increased complexity can help the model capture more subtle relationships in complex datasets.
2. **Overfitting/Underfitting**. Adjusting the number of hidden layers can help prevent overfitting or underfitting. Too many layers can lead to overfitting, and too few layers can lead to the model failing to capture significant patterns in the data (underfitting).

By analysing our models we also explored different other settings.

To describe classification results there uses the terms "true positive", "true negative", "false positive" and "false negative" in the context of binary classification models the terms "true positive", "true negative", "false positive" and "false negative".

1. **True positive**. The model correctly predicts a positive class (e.g., "yes") as positive.
2. **True Negative**. The model correctly predicts a negative class (e.g., "no") as negative.
3. **False positive**. The model incorrectly predicts a negative class as positive. In other words, it incorrectly identifies a case as positive when it is actually negative.
4. **False Negative**. The model incorrectly predicts a positive class as negative. It fails to identify the positive case.

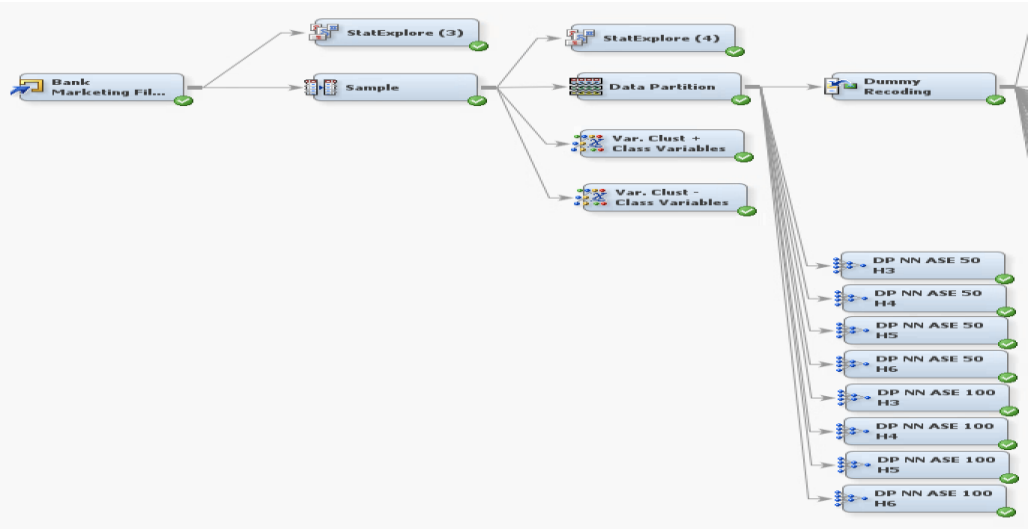
Cumulative Lift and Cumulative % Response.

In predictive modeling and binary classification tasks, examining score ratings and overlaying them with plots of cumulative lift and cumulative % response provides insight into how well the model identifies positive cases (out target “yes” - subscription).

1. **Cumulative Lift**. This metric compares the performance of the model to random chance. It shows how much better the model is at identifying positive cases compared to random sampling. A higher lift value in the upper segments of the ranking indicates that the model is better at identifying positive cases compared to random sampling.
2. **Cumulative % Response**. This chart shows the percentage of actual positive cases identified by the model as it moves down the rankings. The sharper the curve, the more effective the model is at identifying positive cases early in the ranking process.

Data Partition Neural Networks

Picture 76. Adding Data Partition Neural Network Nodes into the Model



Firstly, we added Neural Network to the data partition node, we denied preliminary training in optimisation. Then we increased Hidden layers in network settings and we carried 6 Neural Networks with 50 Iterations and the next 5 with 100 Iterations. We looked for the lowest ASE and Misclassification rate in our results for each Neural Network to find the best optimisation model. We also noted the Parameter Estimation which determined the complexity of our model. We noted Parameter Estimation because in some cases the ASE or Misclassification rate was the same and it was harder for us to decide which could be the best model. With

the help of Parameter Estimation, we could decide which was a less complex structure and we selected that. For our best model, DP NN 50 Iterations and 3 Hidden Units and DP NN 100 Iterations and 3 Hidden Units, the best results were given after 23 & 27 iterations and it showed a “QUANEW” warning message because may be struggling to converge to a solution within the given number of iterations.

Table 13. Data Partition Neural Networks Results

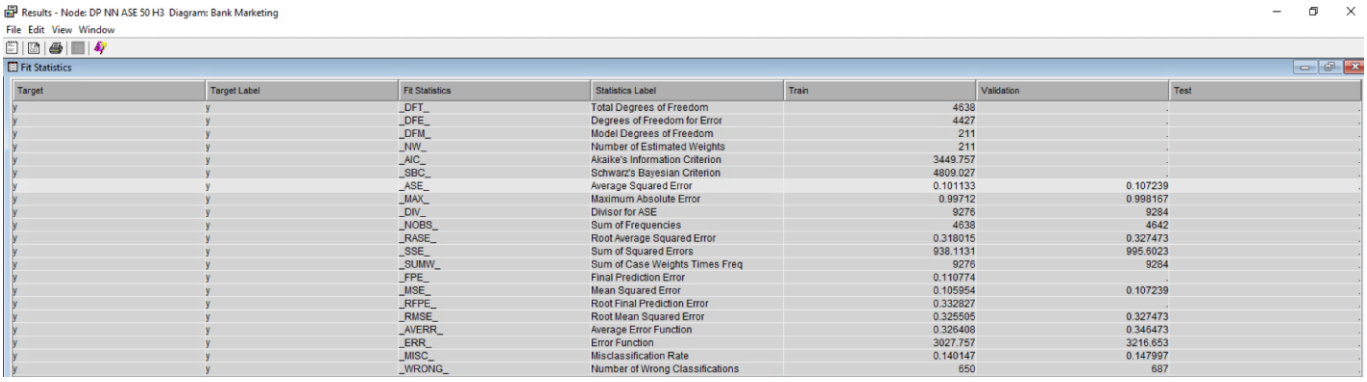
Neural Network	Number of Iterations	Number of Hidden Layers	Misclassification	ASE	Parameter Estimate
DP NN ASE 50	50	3	0.140147	0.107239	211
DP NN ASE 50	50	4	0.155967	0.111473	281
DP NN ASE 50	50	5	0.150582	0.109376	351
DP NN ASE 50	50	6	0.153813	0.112011	421
DP NN ASE 100	100	3	0.147997	0.107239	211
DP NN ASE 100	100	4	0.155967	0.111473	281
DP NN ASE 100	100	5	0.150582	0.109376	351
DP NN ASE 100	100	6	0.153813	0.112011	421

The lowest ASE was for DP NN ASE with 50 Iterations and 3 Hidden Units and for DP NN ASE with 100 Iterations with 3 Hidden Units, these were our best Neural Network models for Data Partitioning

Overall the ASE remained the same after we increased the number of Iterations.

Data Partition Neural Network Average Squared Error, 50 iterations, 3 hidden layers

Picture 77. Results for Data Partition Neural Network 50 iterations 3 hidden units



Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Freedom	4638	.	.
y	y	_DFE_	Degrees of Freedom for Error	4427	.	.
y	y	_DFM_	Model Degrees of Freedom	211	.	.
y	y	_NW_	Number of Estimated Weights	211	.	.
y	y	_AIC_	Akaike's Information Criterion	3449.757	.	.
y	y	_SBC_	Schwarz's Bayesian Criterion	4809.027	.	.
y	y	_ASE_	Average Squared Error	0.101133	0.107239	.
y	y	_MAX_	Maximum Absolute Error	0.99712	0.998167	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squared Error	0.318015	0.327473	.
y	y	_SSE_	Sum of Squared Errors	938.1131	995.6023	.
y	y	_SUMW_	Sum of Case Weights Times Freq	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.110774	.	.
y	y	_MSE_	Mean Squared Error	0.105954	0.107239	.
y	y	_RFPE_	Root Final Prediction Error	0.332827	.	.
y	y	_RMSE_	Root Mean Squared Error	0.325505	0.327473	.
y	y	_AVERR_	Average Error Function	0.326408	0.346473	.
y	y	_ERR_	Error Function	3027.757	3216.653	.
y	y	_MISC_	Misclassification Rate	0.140147	0.147997	.
y	y	_WRONG_	Number of Wrong Classifications	650	687	.

Picture 78. Optimization Start, Parameter Estimates for Data Partition Neural Network 50 iterations 3 hidden units

```
The NEURAL Procedure

                        Optimization Start
                        Parameter Estimates
                                                        Gradient
                                                       Objective
  N Parameter                              Estimate       Function

  1 age_H11                               -0.005212              0
  2 campaign_H11                          -0.012125              0
  3 duration_H11                          -0.029528              0
  4 age_H12                               -0.026953              0
  5 campaign_H12                           0.008917              0
  6 duration_H12                           0.060557              0
  7 age_H13                               -0.184471              0
  8 campaign_H13                           0.096020              0
  9 duration_H13                           0.069390              0
 10 contactcellular_H11                    0.021308              0
```

…

```
200 pdays22_H13                            0.016455              0
201 pdays25_H13                           -0.031278              0
202 pdays26_H13                            0.013142              0
203 poutcomefailure_H13                   -0.114525              0
204 poutcomenonexistent_H13               -0.063467              0
205 BIAS_H11                              -0.794813              0
206 BIAS_H12                              -0.794908              0
207 BIAS_H13                               0.201300              0
208 H11_yyes                                      0       0.037241
209 H12_yyes                                      0      -0.017805
210 H13_yyes                                      0      -0.034998
211 BIAS_yyes                                     0              0

Value of Objective Function = 0.6931471806
```

Picture 79. Optimization Results for Data Partition Neural Network 50 iterations 3 hidden units

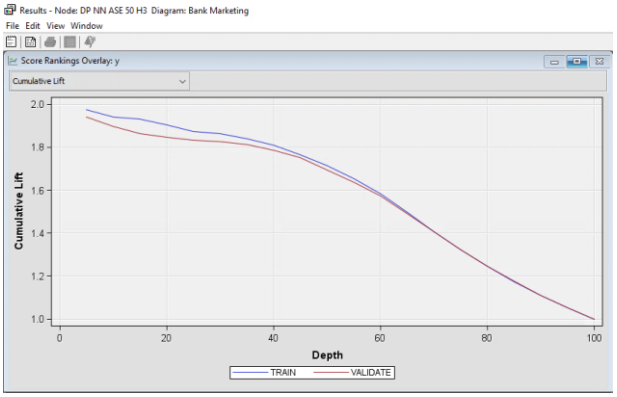
```
                                        Optimization Results

Iterations                                        50  Function Calls                          126
Gradient Calls                                    56  Active Constraints                        0
Objective Function                     0.3131501632  Max Abs Gradient Element        0.0036785644
Slope of Search Direction              -0.001064499

QUANEW needs more than 50 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Picture 80. Classification Results for Data Partition Neural Network 50 iterations 3 hidden units

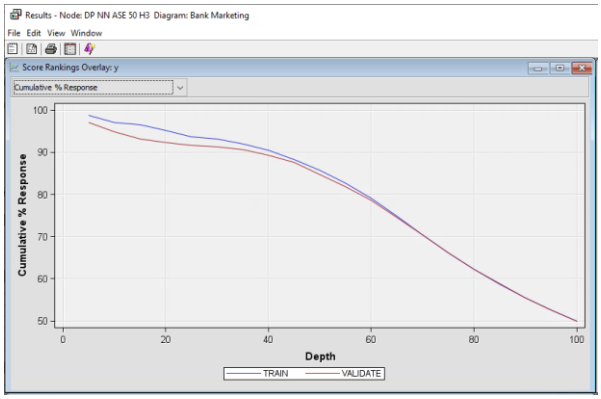
```

Data Role=VALIDATE Target=y Target Label=y

   False       True        False       True
 Negative    Negative    Positive    Positive

     270        1904         417        2051
```

There was a warning message.The optimization algorithm (QUANEW) may be struggling to converge to a solution within the given number of iterations The maximum number of iterations, by default, is 50. Apparently, this is not enough for the network training process to converge.

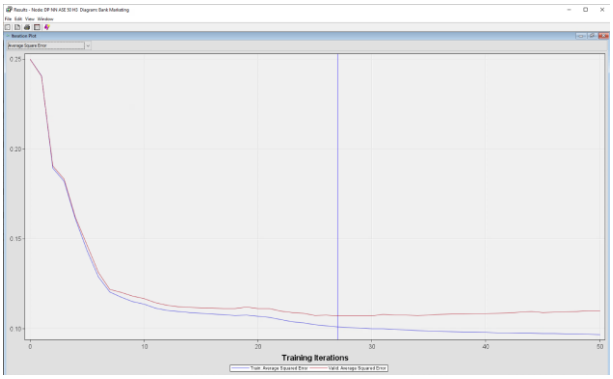
Picture 81. Cumulative Lift Plot for Data Partition Neural Network 50 iterations 3 hidden units



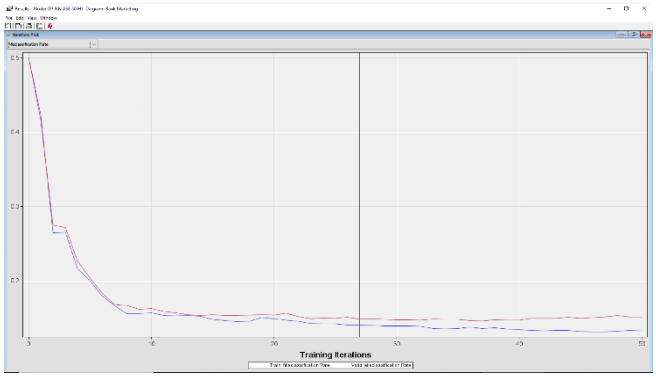
Picture 82. Cumulative % Response for Data Partition Neural Network 50 iterations 3 hidden units



Picture 83. Iteration Plot (ASE) for Data Partition Neural Network 50 iterations 3 hidden units



Picture 84. Iteration Plot (Misclassification) for Data Partition Neural Network 50 iterations 3 hidden units



ASE. As seen in the graph, approximately 27 iterations are enough after that it gives us the same results. **Misclassification**. As seen in the graph, approximately 27 iterations are enough after that it gives us the same results. The iteration plot shows optimal validation average squared error occurring on this iteration.

Data Partition Neural Network Average Squared Error, 50 iterations, 4 hidden layers

Picture 85. Results for Data Partition Neural Network 50 iterations 4 hidden units

Results - Node: DP NN ASE 50 H4 Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Freedom	4638	.	.
y	y	_DFE_	Degrees of Freedom for Error	4357	.	.
y	y	_DFM_	Model Degrees of Freedom	281	.	.
y	y	_NW_	Number of Estimated Weights	281	.	.
y	y	_AIC_	Akaike's Information Criterion	3762.203	.	.
y	y	_SBC_	Schwarz's Bayesian Criterion	5572.416	.	.
y	y	_ASE_	Average Squared Error	0.106642	0.111473	.
y	y	_MAX_	Maximum Absolute Error	0.994798	0.995159	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squared Error	0.32656	0.333875	.
y	y	_SSE_	Sum of Squared Errors	989.2071	1034.911	.
y	y	_SUMW_	Sum of Case Weights Times Freq	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.120397	.	.
y	y	_MSE_	Mean Squared Error	0.113519	0.111473	.
y	y	_RFPE_	Root Final Prediction Error	0.346983	.	.
y	y	_RMSE_	Root Mean Squared Error	0.336926	0.333875	.
y	y	_AVERR_	Average Error Function	0.344998	0.359575	.
y	y	_ERR_	Error Function	3200.203	3338.297	.
y	y	_MISC_	Misclassification Rate	0.145968	0.155967	.
y	y	_WRONG_	Number of Wrong Classifications	677	724	.

Picture 86. Optimization Start, Parameter Estimates for Data Partition Neural Network 50 iterations 4 hidden units

```
The NEURAL Procedure

                    Optimization Start
                    Parameter Estimates
                                                  Gradient
                                                 Objective
 N Parameter                          Estimate    Function

 1 age_H11                           -0.005212           0
 2 campaign_H11                      -0.012125           0
 3 duration_H11                      -0.029528           0
 4 age_H12                           -0.026953           0
 5 campaign_H12                       0.008917           0
 6 duration_H12                       0.060557           0
 7 age_H13                           -0.184471           0
 8 campaign_H13                       0.096020           0
 9 duration_H13                       0.069390           0
10 age_H14                            0.021308           0
```

…

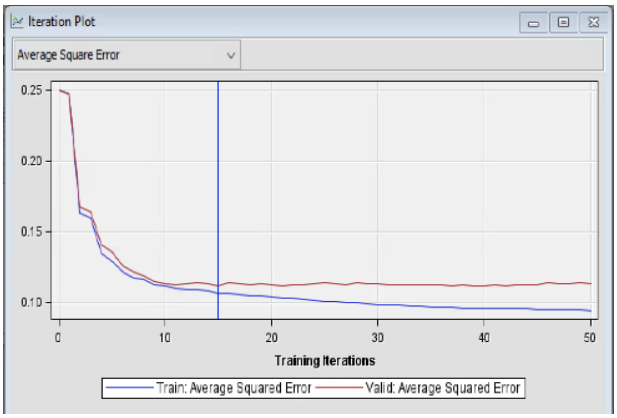
```
270 pdays26_H14                     -0.015754           0
271 poutcomefailure_H14              0.109466           0
272 poutcomenonexistent_H14         -0.069959           0
273 BIAS_H11                         0.205731           0
274 BIAS_H12                        -0.495543           0
275 BIAS_H13                        -0.327556           0
276 BIAS_H14                         0.392078           0
277 H11_yyes                                0      -0.015445
278 H12_yyes                                0      -0.019336
279 H13_yyes                                0      -0.018180
280 H14_yyes                                0      -0.014110
281 BIAS_yyes                               0               0

Value of Objective Function = 0.6931471806
```

Picture 87. Optimization Results for Data Partition Neural Network 50 iterations 4 hidden units

```
                                    Optimization Results

Iterations                                    50  Function Calls                          128
Gradient Calls                                59  Active Constraints                        0
Objective Function                 0.3066348786  Max Abs Gradient Element       0.0017077794
Slope of Search Direction         -0.000824453

QUANEW needs more than 50 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

There was a warning message. The optimization algorithm (QUANEW) may be struggling to converge to a solution within the given number of iterations The maximum number of iterations, by default, is 50. Apparently, this is not enough for the network training process to converge.

Picture 88. Iteration Plot (ASE) for Data Partition Neural Network 50 iterations 4 hidden units



As seen in the graph, approximately 15 iterations are enough after that it gives us the same results. The iteration plot shows optimal validation average squared error occurring on this iteration.

Data Partition Neural Network Average Squared Error, 50 iterations, 5 hidden layers

Picture 89. Results for Data Partition Neural Network 50 iterations 5 hidden units

Results - Node: DP NN ASE 50 H5 Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Freedom	4638	.	.
y	y	_DFE_	Degrees of Freedom for Error	4287	.	.
y	y	_DFM_	Model Degrees of Freedom	351	.	.
y	y	_NW_	Number of Estimated Weights	351	.	.
y	y	_AIC_	Akaike's Information Criterion	3593.218	.	.
y	y	_SBC_	Schwarz's Bayesian Criterion	5854.373	.	.
y	y	_ASE_	Average Squared Error	0.095866	0.109376	.
y	y	_MAX_	Maximum Absolute Error	0.992132	0.999983	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squared Error	0.309623	0.330721	.
y	y	_SSE_	Sum of Squared Errors	889.2573	1015.451	.
y	y	_SUMW_	Sum of Case Weights Times Freq	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.111565	.	.
y	y	_MSE_	Mean Squared Error	0.103716	0.109376	.
y	y	_RFPE_	Root Final Prediction Error	0.334013	.	.
y	y	_RMSE_	Root Mean Squared Error	0.322049	0.330721	.
y	y	_AVERR_	Average Error Function	0.311688	0.357807	.
y	y	_ERR_	Error Function	2891.218	3321.883	.
y	y	_MISC_	Misclassification Rate	0.135619	0.150582	.
y	y	_WRONG_	Number of Wrong Classifications	629	699	.

Picture 90. Optimization Start, Parameter Estimates for Data Partition Neural Network 50 iterations 5 hidden units

```
The NEURAL Procedure

                        Optimization Start
                        Parameter Estimates
                                                    Gradient
                                                   Objective
  N Parameter                          Estimate     Function

  1 age_H11                           -0.005212            0
  2 campaign_H11                      -0.012125            0
  3 duration_H11                      -0.029528            0
  4 age_H12                           -0.026953            0
  5 campaign_H12                       0.008917            0
  6 duration_H12                       0.060557            0
  7 age_H13                           -0.184471            0
  8 campaign_H13                       0.096020            0
  9 duration_H13                       0.069390            0
 10 age_H14                            0.021308            0
```

...

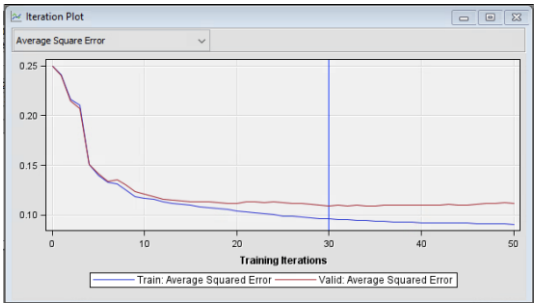
```
340 poutcomenonexistent_H15          -0.112449            0
341 BIAS_H11                          0.432637            0
342 BIAS_H12                          1.523502            0
343 BIAS_H13                          0.189475            0
344 BIAS_H14                         -0.484969            0
345 BIAS_H15                          1.114734            0
346 H11_yyes                                 0    -0.011558
347 H12_yyes                                 0     0.002410
348 H13_yyes                                 0    -0.007753
349 H14_yyes                                 0    -0.039543
350 H15_yyes                                 0    -0.013742
351 BIAS_yyes                                0            0

Value of Objective Function = 0.6931471806
```

Picture 91. Optimization Results for Data Partition Neural Network 50 iterations 5 hidden units

```
                                Optimization Results

Iterations                                   50  Function Calls                    137
Gradient Calls                               62  Active Constraints                  0
Objective Function                 0.2956237455  Max Abs Gradient Element   0.0065305392
Slope of Search Direction         -0.001320879

QUANEW needs more than 50 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Notice the warning message. The optimization algorithm (QUANEW) may be struggling to converge to a solution within the given number of iterations The maximum number of iterations, by default, is 50. Apparently, this is not enough for the network training process to converge.

Picture 92. Iteration Plot (ASE) for Data Partition Neural Network 50 iterations 5 hidden units



As seen in the graph, approximately 30 iterations are enough after that it gives us the same results.The iteration plot shows optimal validation average squared error occurring on this iteration.

Data Partition Neural Network Average Squared Error, 50 iterations, 6 hidden layers

Picture 93. Results for Data Partition Neural Network 50 iterations 6 hidden units

Results - Node: DP NN ASE 50 H6 Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Freedom	4638	.	.
y	y	_DFE_	Degrees of Freedom for Error	4217	.	.
y	y	_DFM_	Model Degrees of Freedom	421	.	.
y	y	_NW_	Number of Estimated Weights	421	.	.
y	y	_AIC_	Akaike's Information Criterion	3934.804	.	.
y	y	_SBC_	Schwarz's Bayesian Criterion	6646.903	.	.
y	y	_ASE_	Average Squared Error	0.103463	0.112011	.
y	y	_MAX_	Maximum Absolute Error	0.987117	0.99918	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squared Error	0.321657	0.33468	.
y	y	_SSE_	Sum of Squared Errors	959.7246	1039.906	.
y	y	_SUMW_	Sum of Case Weights Times Freq	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.124121	.	.
y	y	_MSE_	Mean Squared Error	0.113792	0.112011	.
y	y	_RFPE_	Root Final Prediction Error	0.352309	.	.
y	y	_RMSE_	Root Mean Squared Error	0.337331	0.33468	.
y	y	_AVERR_	Average Error Function	0.33342	0.358623	.
y	y	_ERR_	Error Function	3092.804	3329.46	.
y	y	_MISC_	Misclassification Rate	0.145106	0.153813	.
y	y	_WRONG_	Number of Wrong Classifications	673	714	.

Picture 94. Optimization Start, Parameter Estimates for Data Partition Neural Network 50 iterations 6 hidden units

```
The NEURAL Procedure

                        Optimization Start
                        Parameter Estimates
                                                   Gradient
                                                  Objective
 N Parameter                            Estimate   Function

 1 age_H11                             -0.005212          0
 2 campaign_H11                        -0.012125          0
 3 duration_H11                        -0.029528          0
 4 age_H12                             -0.026953          0
 5 campaign_H12                         0.008917          0
 6 duration_H12                         0.060557          0
 7 age_H13                             -0.184471          0
 8 campaign_H13                         0.096020          0
 9 duration_H13                         0.069390          0
10 age_H14                              0.021308          0
```

…

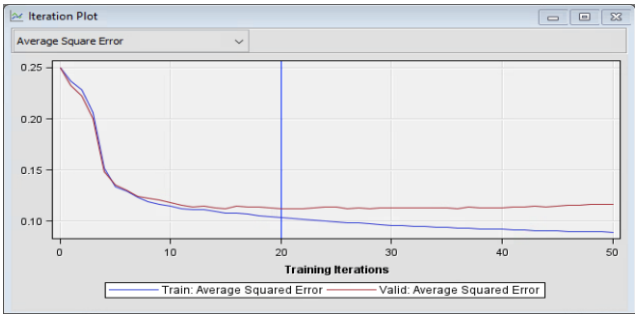
```
410 BIAS_H12                           -1.186628          0
411 BIAS_H13                           -0.276756          0
412 BIAS_H14                            1.053376          0
413 BIAS_H15                            0.826377          0
414 BIAS_H16                            1.110835          0
415 H11_yyes                                   0   -0.027700
416 H12_yyes                                   0   -0.007840
417 H13_yyes                                   0    0.031970
418 H14_yyes                                   0   -0.029007
419 H15_yyes                                   0    0.008398
420 H16_yyes                                   0   -0.018442
421 BIAS_yyes                                  0          0

Value of Objective Function = 0.6931471806
```

Picture 95. Optimization Results for Data Partition Neural Network 50 iterations 6 hidden units

```
                                        Optimization Results

Iterations                                        50  Function Calls                                  122
Gradient Calls                                    56  Active Constraints                                0
Objective Function                     0.2910763391  Max Abs Gradient Element               0.0034117618
Slope of Search Direction              -0.000982871

QUANEW needs more than 50 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Notice the warning message.The optimization algorithm (QUANEW) may be struggling to converge to a solution within the given number of iterations The maximum number of iterations, by default, is 50. Apparently, this is not enough for the network training process to converge.

Picture 97. Iteration Plot (ASE) for Data Partition Neural Network 50 iterations 6 hidden units



As seen in the graph, approximately 20 iterations are enough after that it gives us the same results. The iteration plot shows optimal validation average squared error occurring on this iteration.

Data Partition Neural Network Average Squared Error, 100 iterations, 3 hidden layers

Picture 98. Results for Data Partition Neural Network 100 iterations 3 hidden units

Results - Node: DP NN ASE 100 H3 Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Freedom	4638	.	.
y	y	_DFE_	Degrees of Freedom for Error	4427	.	.
y	y	_DFM_	Model Degrees of Freedom	211	.	.
y	y	_NW_	Number of Estimated Weights	211	.	.
y	y	_AIC_	Akaike's Information Criterion	3449.757	.	.
y	y	_SBC_	Schwarz's Bayesian Criterion	4809.027	.	.
y	y	_ASE_	Average Squared Error	0.101133	0.107239	.
y	y	_MAX_	Maximum Absolute Error	0.99712	0.998167	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4542	.
y	y	_RASE_	Root Average Squared Error	0.318015	0.327473	.
y	y	_SSE_	Sum of Squared Errors	938.1131	995.6023	.
y	y	_SUMW_	Sum of Case Weights Times Freq	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.110774	.	.
y	y	_MSE_	Mean Squared Error	0.105954	0.107239	.
y	y	_RFPE_	Root Final Prediction Error	0.332827	.	.
y	y	_RMSE_	Root Mean Squared Error	0.325505	0.327473	.
y	y	_AVERR_	Average Error Function	0.326408	0.346473	.
y	y	_ERR_	Error Function	3027.757	3216.653	.
y	y	_MISC_	Misclassification Rate	0.140147	0.147997	.
y	y	_WRONG_	Number of Wrong Classifications	650	687	.

Picture 99. Optimization Start, Parameter Estimates for Data Partition Neural Network 100 iterations 3 hidden units

```
The NEURAL Procedure

                        Optimization Start
                        Parameter Estimates
                                                        Gradient
                                                       Objective
  N Parameter                              Estimate       Function

  1 age_H11                                -0.005212              0
  2 campaign_H11                           -0.012125              0
  3 duration_H11                           -0.029528              0
  4 age_H12                                -0.026953              0
  5 campaign_H12                            0.008917              0
  6 duration_H12                            0.060557              0
  7 age_H13                                -0.184471              0
  8 campaign_H13                            0.096020              0
  9 duration_H13                            0.069390              0
 10 contactcellular_H11                     0.021308              0
```

…

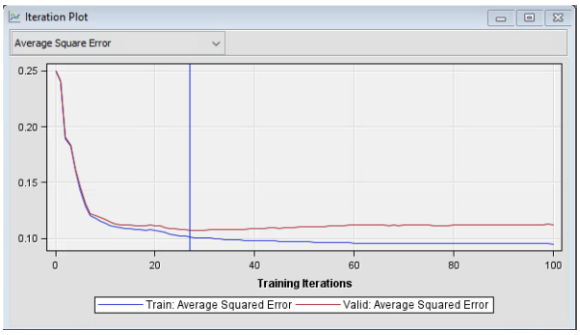
```
200 pdays22_H13                             0.016455              0
201 pdays25_H13                            -0.031278              0
202 pdays26_H13                             0.013142              0
203 poutcomefailure_H13                    -0.114525              0
204 poutcomenonexistent_H13                -0.063467              0
205 BIAS_H11                               -0.794813              0
206 BIAS_H12                               -0.794908              0
207 BIAS_H13                                0.201300              0
208 H11_yyes                                       0       0.037241
209 H12_yyes                                       0      -0.017805
210 H13_yyes                                       0      -0.034998
211 BIAS_yyes                                      0              0

Value of Objective Function = 0.6931471806
```

Picture 99. Optimization Results for Data PartitionNeural Network 100 iterations 3 hidden units

```
                                  Optimization Results

Iterations                                  100  Function Calls                          252
Gradient Calls                              114  Active Constraints                        0
Objective Function                 0.3076287909  Max Abs Gradient Element       0.0026450927
Slope of Search Direction          -0.000054088

QUANEW needs more than 100 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Notice the warning message.The optimization algorithm (QUANEW) may be struggling to converge to a solution within the given number of iterations The maximum number of iterations, by default, is 100. Apparently, this is not enough for the network training process to converge.

Picture 100. Iteration Plot (ASE) for Data Partition Neural Network 100 iterations 3 hidden units



As seen in the graph, approximately 27 iterations are enough after that it gives us the same results.The iteration plot shows optimal validation average squared error occurring on this iteration.

Data Partition Neural Network Average Squared Error, 100 iterations, 4 hidden layers

Picture 101. Results for Data Partition Neural Network 100 iterations 4 hidden units

Results - Node: DP NN ASE 100 H4 Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Freedom	4638	.	.
y	y	_DFE_	Degrees of Freedom for Error	4357	.	.
y	y	_DFM_	Model Degrees of Freedom	281	.	.
y	y	_NW_	Number of Estimated Weights	281	.	.
y	y	_AIC_	Akaike's Information Criterion	3762.203	.	.
y	y	_SBC_	Schwarz's Bayesian Criterion	5572.416	.	.
y	y	_ASE_	Average Squared Error	0.106642	0.111473	.
y	y	_MAX_	Maximum Absolute Error	0.994798	0.995159	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squared Error	0.32656	0.333875	.
y	y	_SSE_	Sum of Squared Errors	989.2071	1034.911	.
y	y	_SUMW_	Sum of Case Weights Times Freq	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.120397	.	.
y	y	_MSE_	Mean Squared Error	0.113519	0.111473	.
y	y	_RFPE_	Root Final Prediction Error	0.346983	.	.
y	y	_RMSE_	Root Mean Squared Error	0.336926	0.333875	.
y	y	_AVERR_	Average Error Function	0.344998	0.359575	.
y	y	_ERR_	Error Function	3200.203	3338.297	.
y	y	_MISC_	Misclassification Rate	0.145968	0.155967	.
y	y	_WRONG_	Number of Wrong Classifications	677	724	.

Picture 102. Optimization Start, Parameter Estimates for Data Partition Neural Network 100 iterations 4 hidden units

```
The NEURAL Procedure

                        Optimization Start
                        Parameter Estimates
                                                        Gradient
                                                       Objective
  N Parameter                            Estimate        Function

  1 age_H11                             -0.005212               0
  2 campaign_H11                        -0.012125               0
  3 duration_H11                        -0.029528               0
  4 age_H12                             -0.026953               0
  5 campaign_H12                         0.008917               0
  6 duration_H12                         0.060557               0
  7 age_H13                             -0.184471               0
  8 campaign_H13                         0.096020               0
  9 duration_H13                         0.069390               0
 10 age_H14                              0.021308               0
```

…

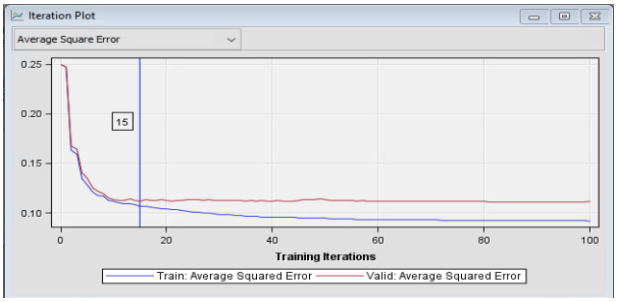
```
270 pdays26_H14                         -0.015754               0
271 poutcomefailure_H14                  0.109466               0
272 poutcomenonexistent_H14             -0.069959               0
273 BIAS_H11                             0.205731               0
274 BIAS_H12                            -0.495543               0
275 BIAS_H13                            -0.327556               0
276 BIAS_H14                             0.392078               0
277 H11_yyes                                    0       -0.015445
278 H12_yyes                                    0       -0.019336
279 H13_yyes                                    0       -0.018180
280 H14_yyes                                    0       -0.014110
281 BIAS_yyes                                   0               0

Value of Objective Function = 0.6931471806
```

Picture 103. Optimization Results for Data Partition Neural Network 100 iterations 4 hidden units

```
                                Optimization Results

Iterations                                100  Function Calls                          276
Gradient Calls                            125  Active Constraints                        0
Objective Function                0.2987912338  Max Abs Gradient Element       0.0013418758
Slope of Search Direction        -0.000082917

QUANEW needs more than 100 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Notice the warning message.The optimization algorithm (QUANEW) may be struggling to converge to a solution within the given number of iterations The maximum number of iterations, by default, is 100. Apparently, this is not enough for the network training process to converge.

Picture 104. Iteration Plot (ASE) for Data Partition Neural Network 100 iterations 4 hidden units



As seen in the graph, approximately 16 iterations are enough after that it gives us the same results.The iteration plot shows optimal validation average squared error occurring on this iteration.

Data Partition Neural Network Average Squared Error, 100 iterations, 5 hidden layers

Picture 105. Results for Data Partition Neural Network 100 iterations 5 hidden units

Results - Node: DP NN ASE 100 H5 Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Freedom	4638	.	.
y	y	_DFE_	Degrees of Freedom for Error	4287	.	.
y	y	_DFM_	Model Degrees of Freedom	351	.	.
y	y	_NW_	Number of Estimated Weights	351	.	.
y	y	_AIC_	Akaike's Information Criterion	3593.218	.	.
y	y	_SBC_	Schwarz's Bayesian Criterion	5854.373	.	.
y	y	_ASE_	Average Squared Error	0.095866	0.109376	.
y	y	_MAX_	Maximum Absolute Error	0.992132	0.999983	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squared Error	0.309623	0.330721	.
y	y	_SSE_	Sum of Squared Errors	889.2573	1015.451	.
y	y	_SUMW_	Sum of Case Weights Times Freq	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.111565	.	.
y	y	_MSE_	Mean Squared Error	0.103716	0.109376	.
y	y	_RFPE_	Root Final Prediction Error	0.334013	.	.
y	y	_RMSE_	Root Mean Squared Error	0.322049	0.330721	.
y	y	_AVERR_	Average Error Function	0.311688	0.357807	.
y	y	_ERR_	Error Function	2891.218	3321.883	.
y	y	_MISC_	Misclassification Rate	0.135619	0.150582	.
y	y	_WRONG_	Number of Wrong Classifications	629	699	.

Picture 106. Optimization Start, Parameter Estimates for Data Partition Neural Network 100 iterations 5 hidden units

```
The NEURAL Procedure

                        Optimization Start
                        Parameter Estimates
                                                   Gradient
                                                  Objective
  N Parameter                          Estimate     Function

  1 age_H11                           -0.005212            0
  2 campaign_H11                      -0.012125            0
  3 duration_H11                      -0.029528            0
  4 age_H12                           -0.026953            0
  5 campaign_H12                       0.008917            0
  6 duration_H12                       0.060557            0
  7 age_H13                           -0.184471            0
  8 campaign_H13                       0.096020            0
  9 duration_H13                       0.069390            0
 10 age_H14                            0.021308            0
```

…

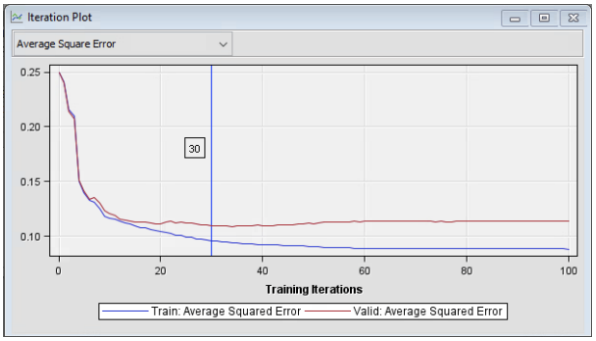
```
340 poutcomenonexistent_H15          -0.112449            0
341 BIAS_H11                          0.432637            0
342 BIAS_H12                          1.523502            0
343 BIAS_H13                          0.189475            0
344 BIAS_H14                         -0.484969            0
345 BIAS_H15                          1.114734            0
346 H11_yyes                                 0    -0.011558
347 H12_yyes                                 0     0.002410
348 H13_yyes                                 0    -0.007753
349 H14_yyes                                 0    -0.039543
350 H15_yyes                                 0    -0.013742
351 BIAS_yyes                                0            0

Value of Objective Function = 0.6931471806
```

Picture 107. Optimization Results for Data Partition Neural Network 100 iterations 5 hidden units

```
                                        Optimization Results

Iterations                                     100  Function Calls                                271
Gradient Calls                                 123  Active Constraints                              0
Objective Function                0.2886351057  Max Abs Gradient Element                0.001046381
Slope of Search Direction         -7.868493E-6

QUANEW needs more than 100 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Notice the warning message.The optimization algorithm (QUANEW) may be struggling to converge to a solution within the given number of iterations The maximum number of iterations, by default, is 100. Apparently, this is not enough for the network training process to converge.

Picture 108. Iteration Plot (ASE) for Data Partition Neural Network 100 iterations 5 hidden units



As seen in the graph, approximately 30 iterations are enough after that it gives us the same results.The iteration plot shows optimal validation average squared error occurring on this iteration.

Data Partition Neural Network Average Squared Error, 100 iterations, 6 hidden layers

Picture 109. Results for Data Partition Neural Network 100 iterations 6 hidden units

Results - Node: DP NN ASE 100 H6 Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Freedom	4638	.	.
y	y	_DFE_	Degrees of Freedom for Error	4217	.	.
y	y	_DFM_	Model Degrees of Freedom	421	.	.
y	y	_NW_	Number of Estimated Weights	421	.	.
y	y	_AIC_	Akaike's Information Criterion	3934.804	.	.
y	y	_SBC_	Schwarz's Bayesian Criterion	6646.903	.	.
y	y	_ASE_	Average Squared Error	0.103463	0.112011	.
y	y	_MAX_	Maximum Absolute Error	0.987117	0.99918	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squared Error	0.321657	0.33468	.
y	y	_SSE_	Sum of Squared Errors	959.7246	1039.906	.
y	y	_SUMW_	Sum of Case Weights Times Freq	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.124121	.	.
y	y	_MSE_	Mean Squared Error	0.113792	0.112011	.
y	y	_RFPE_	Root Final Prediction Error	0.352309	.	.
y	y	_RMSE_	Root Mean Squared Error	0.337331	0.33468	.
y	y	_AVERR_	Average Error Function	0.33342	0.358623	.
y	y	_ERR_	Error Function	3092.804	3329.46	.
y	y	_MISC_	Misclassification Rate	0.145106	0.153813	.
y	y	_WRONG_	Number of Wrong Classifications	673	714	.

Picture 110. Optimization Start, Parameter Estimates for Data Partition Neural Network 100 iterations 6 hidden units

```
The NEURAL Procedure

                        Optimization Start
                        Parameter Estimates
                                                    Gradient
                                                   Objective
  N Parameter                          Estimate     Function

  1 age_H11                           -0.005212            0
  2 campaign_H11                      -0.012125            0
  3 duration_H11                      -0.029528            0
  4 age_H12                           -0.026953            0
  5 campaign_H12                       0.008917            0
  6 duration_H12                       0.060557            0
  7 age_H13                           -0.184471            0
  8 campaign_H13                       0.096020            0
  9 duration_H13                       0.069390            0
 10 age_H14                            0.021308            0
```

...

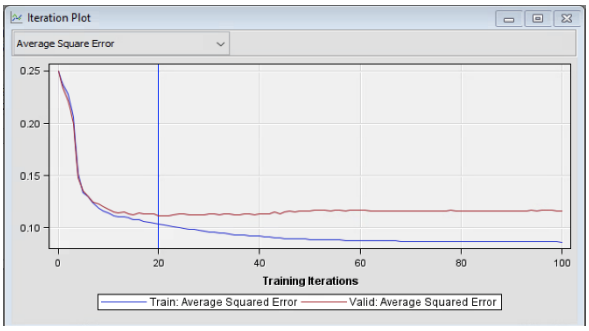
```
409 BIAS_H11                          -0.016807            0
410 BIAS_H12                          -1.186628            0
411 BIAS_H13                          -0.276756            0
412 BIAS_H14                           1.053376            0
413 BIAS_H15                           0.826377            0
414 BIAS_H16                           1.110835            0
415 H11_yyes                                  0     -0.027700
416 H12_yyes                                  0     -0.007840
417 H13_yyes                                  0      0.031970
418 H14_yyes                                  0     -0.029007
419 H15_yyes                                  0      0.008398
420 H16_yyes                                  0     -0.018442
421 BIAS_yyes                                 0             0

Value of Objective Function = 0.6931471806
```

Picture 111. Optimization Results for Data Partition Neural Network 100 iterations 6 hidden units

```
                                    Optimization Results

Iterations                                   100  Function Calls                          251
Gradient Calls                               116  Active Constraints                        0
Objective Function                 0.2823268085  Max Abs Gradient Element     0.0017011644
Slope of Search Direction         -0.000085168

QUANEW needs more than 100 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Notice the warning message.The optimization algorithm (QUANEW) may be struggling to converge to a solution within the given number of iterations The maximum number of iterations, by default, is 100. Apparently, this is not enough for the network training process to converge.

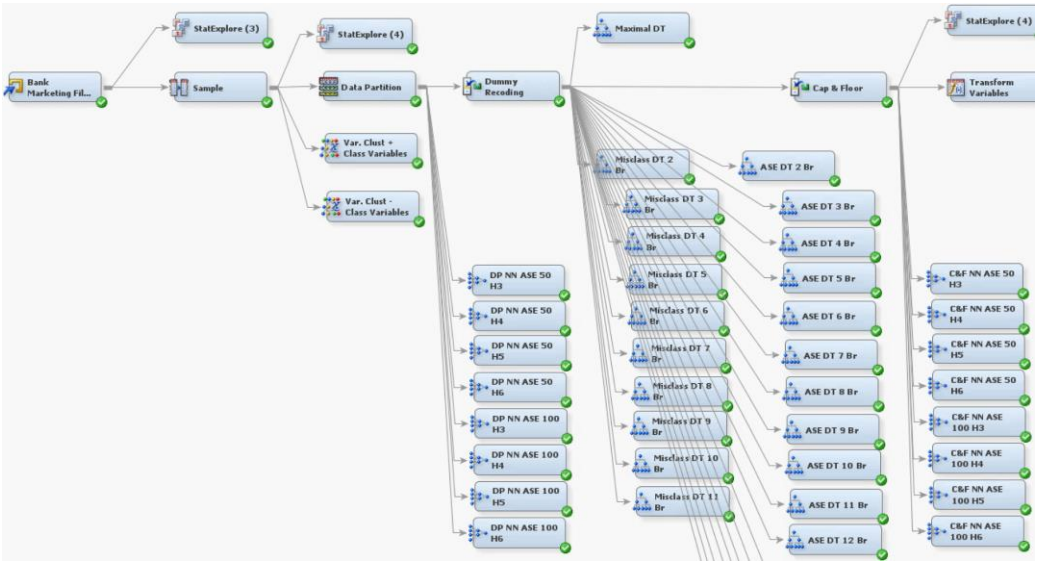
Picture 112. Iteration Plot (ASE) for Data Partition Neural Network 100 iterations 6 hidden units



As seen in the graph, approximately 20 iterations are enough after that it gives us the same results.The iteration plot shows optimal validation average squared error occurring on this iteration.

Cap & Floor NN

Picture 113. Adding Cap and Floor Neural Networks into the Model



Then we added Neural Network to the Cap N Floor node, we denied preliminary training in optimisation. Then we increased Hidden layers in network settings and we carried 6 Neural Networks with 50 Iterations and the next 5 with 100 Iterations. We looked for the lowest ASE and Misclassification rate in our results for each Neural Network to find the best optimisation model. We also noted the Parameter Estimation which determined the complexity of our model. We noted Parameter Estimation because in some cases the ASE or Misclassification rate was the same and it was harder for us to decide which could be the best model. With the help of Parameter Estimation, we could decide which was a less complex structure and we selected that. For our best model, CF NN 50 Iterations and 6 Hidden Units and CF NN 100 Iterations and 6 Hidden Units , the best results were given after 37 & 38 iterations and it showed a “QUANEW” warning message because may be struggling to converge to a solution within the given number of iterations.

Table 14. Cap and Floor Neural Networks Results

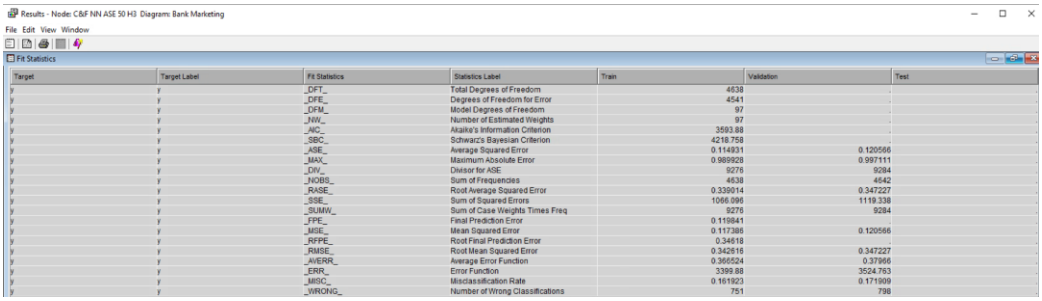
Neural Network	Number of Iterations	Number of Hidden Layers	Misclassification	ASE	Parameter Estimate
CF NN ASE 50	50	3	0.171909	0.120566	97
CF NN ASE 50	50	4	0.162861	0.117521	129
CF NN ASE 50	50	5	0.167169	0.118519	161
CF NN ASE 50	50	6	0.160060	0.116799	193
CF NN ASE 100	100	3	0.171909	0.120566	97
CF NN ASE 100	100	4	0.162215	0.116841	129

CF NN ASE 100	100	5	0.163292	0.116917	161
CF NN ASE 100	100	6	0.160060	0.116799	193

The lowest ASE was for CF NN ASE with 50 Iterations and 6 Hidden Units and for CF NN ASE with 100 Iterations with 6 Hidden Units, these were our best Neural Network models for Cap N Floor. Overall the ASE Reduced after we increased the number of Iterations.

Cap and Floor Neural Network Average Squared Error 50 iterations, 3 hidden units

Picture 114. Results for Cap and Floor Neural Network 50 iterations 3 hidden units



Results - Node: C&F NN ASE 50 H3 Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Freedom	4538		
y	y	_DFE_	Degrees of Freedom for Error	4541		
y	y	_DFM_	Model Degrees of Freedom	97		
y	y	_NW_	Number of Estimated Weights	97		
y	y	_AIC_	Akaike's Information Criterion	3593.88		
y	y	_SBC_	Schwarz's Bayesian Criterion	4218.758		
y	y	_ASE_	Average Squared Error	0.114931	0.120566	
y	y	_MAX_	Maximum Absolute Error	0.989928	0.997111	
y	y	_DIV_	Divisor for ASE	9276	9284	
y	y	_NOBS_	Sum of Frequencies	4538	4642	
y	y	_RASE_	Root Average Squared Error	0.339014	0.347227	
y	y	_SSE_	Sum of Squared Errors	1066.096	1119.338	
y	y	_SUMW_	Sum of Case Weights Times Freq	9276	9284	
y	y	_FPE_	Final Prediction Error	0.119941		
y	y	_MSE_	Mean Squared Error	0.117386	0.120566	
y	y	_RFPE_	Root Final Prediction Error	0.34618		
y	y	_RMSE_	Root Mean Squared Error	0.342616	0.347227	
y	y	_AVERR_	Average Error Function	0.366524	0.37966	
y	y	_ERR_	Error Function	3399.88	3524.763	
y	y	_MISC_	Misclassification Rate	0.161923	0.171909	
y	y	_WRONG_	Number of Wrong Classifications	751	798	

Picture 115. Optimization Start, Parameter Estimates for Cap and Floor Neural Network 50 iterations 3 hidden units

```
The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                Gradient
                                               Objective
 N Parameter                     Estimate       Function

 1 REP_age_H11                  -0.007847              0
 2 REP_campaign_H11             -0.018255              0
 3 REP_duration_H11             -0.044456              0
 4 REP_age_H12                  -0.040579              0
 5 REP_campaign_H12              0.013425              0
 6 REP_duration_H12              0.091171              0
 7 REP_age_H13                  -0.277730              0
 8 REP_campaign_H13              0.144563              0
 9 REP_duration_H13              0.104470              0
10 REP_day_of_weekFr_H11         0.032080              0
```

…

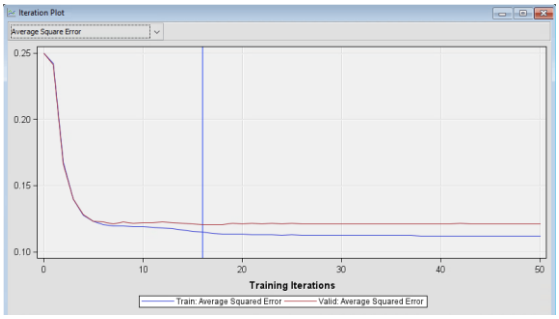
```
90 poutcomenonexistent_H13       0.194201              0
91 BIAS_H11                      0.628974              0
92 BIAS_H12                      0.202838              0
93 BIAS_H13                     -1.278167              0
94 H11_yyes                             0      -0.007612
95 H12_yyes                             0       0.040558
96 H13_yyes                             0      -0.000478
97 BIAS_yyes                            0              0

Value of Objective Function = 0.6931471806
```

Picture 116. Optimization Results for Cap and Floor Neural Network 50 iterations 3 hidden units

```
                                    Optimization Results

Iterations                                   50  Function Calls                          61
Jacobian Calls                               52  Active Constraints                       0
Objective Function              0.3583420739  Max Abs Gradient Element        0.0019225887
Lambda                          0.0073345809  Actual Over Pred Change         0.3801621714
Radius                          0.0617364045

LEVMAR needs more than 50 iterations or 2147483647 function calls.

WARNING: LEVMAR Optimization cannot be completed.
```

Notice the warning message. It can be interpreted to mean that the model-fitting process did not converge. The maximum number of iterations, by default, is 50. Apparently, this is not enough for the network training process to converge.

Picture 117. Iteration Plot (ASE) forCap and Floor Neural Network 50 iterations 3 hidden units



As seen in the graph, approximately 16 iterations are enough after that it gives us the same results.The iteration plot shows optimal validation average squared error occurring on this iteration.

Cap and Floor Neural Network Average Squared Error 50 iterations, 4 hidden units

Picture 118. Results for Cap and Floor Neural Network 50 iterations 4 hidden units

Results - Node: C&F NN ASE 50 H4 Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Freedom	4638	.	.
y	y	_DFE_	Degrees of Freedom for Error	4509	.	.
y	y	_DFM_	Model Degrees of Freedom	129	.	.
y	y	_NW_	Number of Estimated Weights	129	.	.
y	y	_AIC_	Akaike's Information Criterion	3510.777	.	.
y	y	_SBC_	Schwarz's Bayesian Criterion	4341.8	.	.
y	y	_ASE_	Average Squared Error	0.109084	0.117521	.
y	y	_MAX_	Maximum Absolute Error	0.990032	0.997885	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squared Error	0.330279	0.342814	.
y	y	_SSE_	Sum of Squared Errors	1011.866	1091.067	.
y	y	_SUMW_	Sum of Case Weights Times Freq	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.115326	.	.
y	y	_MSE_	Mean Squared Error	0.112205	0.117521	.
y	y	_RFPE_	Root Final Prediction Error	0.339597	.	.
y	y	_RMSE_	Root Mean Squared Error	0.33497	0.342814	.
y	y	_AVERR_	Average Error Function	0.350666	0.372843	.
y	y	_ERR_	Error Function	3252.777	3461.479	.
y	y	_MISC_	Misclassification Rate	0.151358	0.162861	.
y	y	_WRONG_	Number of Wrong Classifications	702	756	.

Picture 119. Optimization Start, Parameter Estimates for Cap and Floor Neural Network 50 iterations 4 hidden units

```
The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                Gradient
                                               Objective
 N Parameter                     Estimate       Function

 1 REP_age_H11                  -0.007847              0
 2 REP_campaign_H11             -0.018255              0
 3 REP_duration_H11             -0.044456              0
 4 REP_age_H12                  -0.040579              0
 5 REP_campaign_H12              0.013425              0
 6 REP_duration_H12              0.091171              0
 7 REP_age_H13                  -0.277730              0
 8 REP_campaign_H13              0.144563              0
 9 REP_duration_H13              0.104470              0
10 REP_age_H14                   0.032080              0
```

…

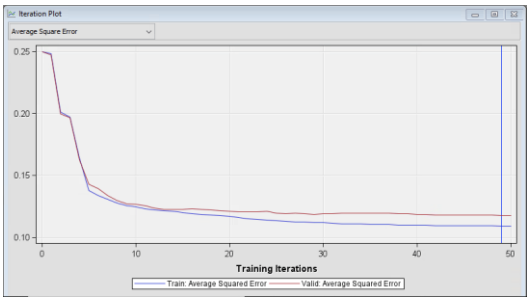
```
120 poutcomenonexistent_H14      0.060935              0
121 BIAS_H11                    -0.283225              0
122 BIAS_H12                     1.216321              0
123 BIAS_H13                     0.941225              0
124 BIAS_H14                     0.044303              0
125 H11_yyes                            0       0.035687
126 H12_yyes                            0      -0.000659
127 H13_yyes                            0      -0.026665
128 H14_yyes                            0      -0.021828
129 BIAS_yyes                           0              0

Value of Objective Function = 0.6931471806
```

Picture 120. Optimization Results for Cap and Floor Neural Network 50 iterations 4 hidden units

```
                                   Optimization Results

Iterations                                   50  Function Calls                          120
Gradient Calls                               55  Active Constraints                        0
Objective Function                  0.350463508  Max Abs Gradient Element       0.0027223901
Slope of Search Direction            -0.0001503

QUANEW needs more than 50 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Notice the warning message.The optimization algorithm (QUANEW) may be struggling to converge to a solution within the given number of iterations The maximum number of iterations, by default, is 50. Apparently, this is not enough for the network training process to converge.

Picture 121. Iteration Plot (ASE) for Cap and Floor Neural Network 50 iterations 4 hidden units



As seen in the graph, approximately 48 iterations are enough after that it gives us the same results.The iteration plot shows optimal validation average squared error occurring on this iteration.

Cap and Floor Neural Network Average Squared Error 50 iterations, 5 hidden units

Picture 122. Results for Cap and Floor Neural Network 50 iterations 5 hidden units

Results - Node: C&F NN ASE 50 H5 Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Freedom	4638	.	.
y	y	_DFE_	Degrees of Freedom for Error	4477	.	.
y	y	_DFM_	Model Degrees of Freedom	161	.	.
y	y	_NW_	Number of Estimated Weights	161	.	.
y	y	_AIC_	Akaike's Information Criterion	3699.029	.	.
y	y	_SBC_	Schwarz's Bayesian Criterion	4736.197	.	.
y	y	_ASE_	Average Squared Error	0.113272	0.118519	.
y	y	_MAX_	Maximum Absolute Error	0.996966	0.996907	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squared Error	0.336559	0.344265	.
y	y	_SSE_	Sum of Squared Errors	1050.71	1100.327	.
y	y	_SUMW_	Sum of Case Weights Times Freq	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.121419	.	.
y	y	_MSE_	Mean Squared Error	0.117345	0.118519	.
y	y	_RFPE_	Root Final Prediction Error	0.348452	.	.
y	y	_RMSE_	Root Mean Squared Error	0.342557	0.344265	.
y	y	_AVERR_	Average Error Function	0.364061	0.375393	.
y	y	_ERR_	Error Function	3377.029	3485.15	.
y	y	_MISC_	Misclassification Rate	0.158258	0.167169	.
y	y	_WRONG_	Number of Wrong Classifications	734	776	.

Picture 123. Optimization Start, Parameter Estimates for Cap and Floor Neural Network 50 iterations 5 hidden units

```

The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                Gradient
                                               Objective
  N Parameter                    Estimate       Function

  1 REP_age_H11                  -0.007847              0
  2 REP_campaign_H11             -0.018255              0
  3 REP_duration_H11             -0.044456              0
  4 REP_age_H12                  -0.040579              0
  5 REP_campaign_H12              0.013425              0
  6 REP_duration_H12              0.091171              0
  7 REP_age_H13                  -0.277730              0
  8 REP_campaign_H13              0.144563              0
  9 REP_duration_H13              0.104470              0
 10 REP_age_H14                   0.032080              0
```

…

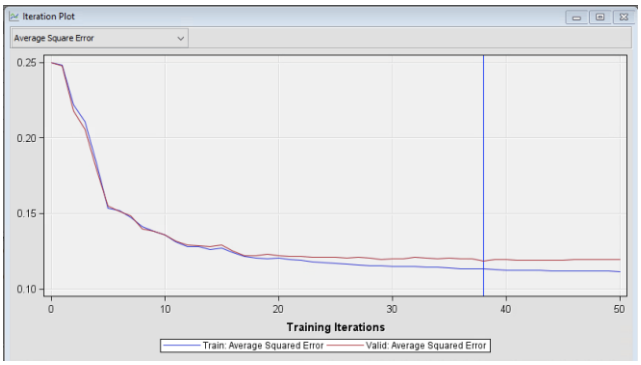
```
150 poutcomenonexistent_H15      -0.228058              0
151 BIAS_H11                     -1.837805              0
152 BIAS_H12                     -2.928428              0
153 BIAS_H13                     -1.950622              0
154 BIAS_H14                      0.712494              0
155 BIAS_H15                     -0.711007              0
156 H11_yyes                             0       0.003523
157 H12_yyes                             0      -0.001103
158 H13_yyes                             0       0.000742
159 H14_yyes                             0       0.013720
160 H15_yyes                             0      -0.022811
161 BIAS_yyes                            0              0

Value of Objective Function = 0.6931471806
```

Picture 124. Optimization Results for Cap and Floor Neural Network 50 iterations 5 hidden units

```

                                   Optimization Results

Iterations                                   50  Function Calls                          131
Gradient Calls                               59  Active Constraints                        0
Objective Function                 0.3575819538  Max Abs Gradient Element       0.0025687692
Slope of Search Direction          -0.000446624

QUANEW needs more than 50 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Notice the warning message.The optimization algorithm (QUANEW) may be struggling to converge to a solution within the given number of iterations The maximum number of iterations, by default, is 50. Apparently, this is not enough for the network training process to converge.

Picture 125. Iteration Plot (ASE) for Cap and Floor Neural Network 50 iterations 5 hidden units



As seen in the graph, approximately 38 iterations are enough after that it gives us the same results.The iteration plot shows optimal validation average squared error occurring on this iteration.

Cap and Floor Neural Network Average Squared Error 50 iterations, 6 hidden units

Picture 126. Results for Cap and Floor Neural Network 50 iterations 6 hidden units

Results - Node: C&F NN ASE 50 H6 Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Freedom	4638	.	.
y	y	_DFE_	Degrees of Freedom for Error	4445	.	.
y	y	_DFM_	Model Degrees of Freedom	193	.	.
y	y	_NW_	Number of Estimated Weights	193	.	.
y	y	_AIC_	Akaike's Information Criterion	3596.786	.	.
y	y	_SBC_	Schwarz's Bayesian Criterion	4840.1	.	.
y	y	_ASE_	Average Squared Error	0.10795	0.116799	.
y	y	_MAX_	Maximum Absolute Error	0.990368	0.999578	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squared Error	0.328558	0.341758	.
y	y	_SSE_	Sum of Squared Errors	1001.345	1084.36	.
y	y	_SUMW_	Sum of Case Weights Times Freq	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.117324	.	.
y	y	_MSE_	Mean Squared Error	0.112637	0.116799	.
y	y	_RFPE_	Root Final Prediction Error	0.342526	.	.
y	y	_RMSE_	Root Mean Squared Error	0.335615	0.341758	.
y	y	_AVERR_	Average Error Function	0.346139	0.373235	.
y	y	_ERR_	Error Function	3210.786	3465.112	.
y	y	_MISC_	Misclassification Rate	0.148124	0.16006	.
y	y	_WRONG_	Number of Wrong Classifications	687	743	.

Picture 127. Optimization Start, Parameter Estimates for Cap and Floor Neural Network 50 iterations 6 hidden units

```
The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                Gradient
                                               Objective
 N Parameter                     Estimate       Function

 1 REP_age_H11                  -0.007847              0
 2 REP_campaign_H11             -0.018255              0
 3 REP_duration_H11             -0.044456              0
 4 REP_age_H12                  -0.040579              0
 5 REP_campaign_H12              0.013425              0
 6 REP_duration_H12              0.091171              0
 7 REP_age_H13                  -0.277730              0
 8 REP_campaign_H13              0.144563              0
 9 REP_duration_H13              0.104470              0
10 REP_age_H14                   0.032080              0
```

…

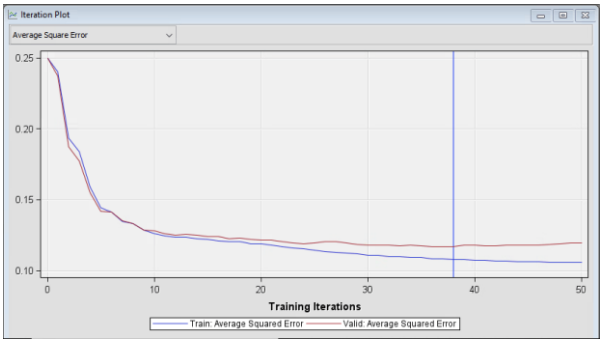
```
190 H14_yyes                    0       -0.029468
191 H15_yyes                    0       -0.016548
192 H16_yyes                    0        0.014591
193 BIAS_yyes                   0               0

Value of Objective Function = 0.6931471806
```

Picture 128. Optimization Results for Cap and Floor Neural Network 50 iterations 6 hidden units

```
                                    Optimization Results

Iterations                                   50  Function Calls                          129
Gradient Calls                               58  Active Constraints                        0
Objective Function              0.3397404558  Max Abs Gradient Element        0.0051809117
Slope of Search Direction       -0.000675616

QUANEW needs more than 50 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Notice the warning message.The optimization algorithm (QUANEW) may be struggling to converge to a solution within the given number of iterations The maximum number of iterations, by default, is 50. Apparently, this is not enough for the network training process to converge.

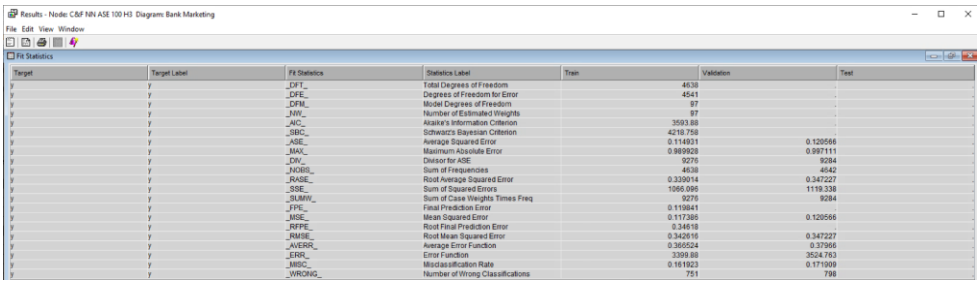
Picture 129. Iteration Plot (ASE) for Cap and Floor Neural Network 50 iterations 6 hidden units



As seen in the graph, approximately 37 iterations are enough after that it gives us the same results.The iteration plot shows optimal validation average squared error occurring on this iteration.

Cap and Floor Neural Network Average Squared Error 100 iterations, 3 hidden units

Picture 130. Results for Cap and Floor Neural Network 100 iterations 3 hidden units



Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Freedom	4638		
y	y	_DFE_	Degrees of Freedom for Error	4541		
y	y	_DFM_	Model Degrees of Freedom	97		
y	y	_NW_	Number of Estimated Weights	97		
y	y	_AIC_	Akaike's Information Criterion	3593.88		
y	y	_SBC_	Schwarz's Bayesian Criterion	4218.758		
y	y	_ASE_	Average Squared Error	0.114931	0.120566	
y	y	_MAX_	Maximum Absolute Error	0.989928	0.997111	
y	y	_DIV_	Divisor for ASE	9276	9284	
y	y	_NOBS_	Sum of Frequencies	4638	4642	
y	y	_RASE_	Root Average Squared Error	0.339014	0.347227	
y	y	_SSE_	Sum of Squared Errors	1066.096	1119.338	
y	y	_SUMW_	Sum of Case Weights Times Freq	9276	9284	
y	y	_FPE_	Final Prediction Error	0.119841		
y	y	_MSE_	Mean Squared Error	0.117386	0.120566	
y	y	_RFPE_	Root Final Prediction Error	0.346118		
y	y	_RMSE_	Root Mean Squared Error	0.342616	0.347227	
y	y	_AVERR_	Average Error Function	0.366524	0.37966	
y	y	_ERR_	Error Function	3399.88	3524.763	
y	y	_MISC_	Misclassification Rate	0.161923	0.171909	
y	y	_WRONG_	Number of Wrong Classifications	751	798	

Picture 131. Optimization Start, Parameter Estimates for Cap and Floor Neural Network 100 iterations 3 hidden units

```
The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                   Gradient
                                                  Objective
 N Parameter                      Estimate        Function

 1 REP_age_H11                    -0.007847               0
 2 REP_campaign_H11               -0.018255               0
 3 REP_duration_H11               -0.044456               0
 4 REP_age_H12                    -0.040579               0
 5 REP_campaign_H12                0.013425               0
 6 REP_duration_H12                0.091171               0
 7 REP_age_H13                    -0.277730               0
 8 REP_campaign_H13                0.144563               0
 9 REP_duration_H13                0.104470               0
10 REP_day_of_weekFr_H11           0.032080               0
```

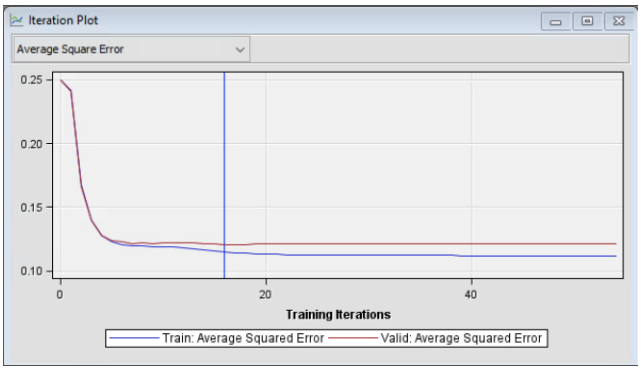
…

```
90 poutcomenonexistent_H13         0.194201               0
91 BIAS_H11                        0.628974               0
92 BIAS_H12                        0.202838               0
93 BIAS_H13                       -1.278167               0
94 H11_yyes                               0       -0.007612
95 H12_yyes                               0        0.040558
96 H13_yyes                               0       -0.000478
97 BIAS_yyes                              0               0

Value of Objective Function = 0.6931471806
```

Picture 132. Optimization Results for Cap and Floor Neural Network 100 iterations 3 hidden units

```
                                   Optimization Results

Iterations                                   54  Function Calls                          65
Jacobian Calls                               56  Active Constraints                       0
Objective Function                 0.3582663441  Max Abs Gradient Element        0.0018944748
Lambda                             0.0066402856  Actual Over Pred Change         0.3504635716
Radius                             0.0617364045

Convergence criterion (FCONV=0.0001) satisfied.

NOTE: At least one element of the gradient is greater than 1e-3.
```

Picture 133. Iteration Plot (ASE) for Cap and Floor Neural Network 100 iterations 3 hidden units



As seen in the graph, approximately 16 iterations are enough after that it gives us the same results.The iteration plot shows optimal validation average squared error occurring on this iteration.

Cap and Floor Neural Network Average Squared Error 100 iterations, 4 hidden units

Picture 134. Results for Cap and Floor Neural Network 100 iterations 4 hidden units

Results - Node: C&F NN ASE 100 H4 Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Freedom	4638	.	.
y	y	_DFE_	Degrees of Freedom for Error	4509	.	.
y	y	_DFM_	Model Degrees of Freedom	129	.	.
y	y	_NW_	Number of Estimated Weights	129	.	.
y	y	_AIC_	Akaike's Information Criterion	3503.59	.	.
y	y	_SBC_	Schwarz's Bayesian Criterion	4334.613	.	.
y	y	_ASE_	Average Squared Error	0.108947	0.116841	.
y	y	_MAX_	Maximum Absolute Error	0.986376	0.997281	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squared Error	0.330071	0.34182	.
y	y	_SSE_	Sum of Squared Errors	1010.591	1084.749	.
y	y	_SUMW_	Sum of Case Weights Times ...	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.115181	.	.
y	y	_MSE_	Mean Squared Error	0.112064	0.116841	.
y	y	_RFPE_	Root Final Prediction Error	0.339383	.	.
y	y	_RMSE_	Root Mean Squared Error	0.334759	0.34182	.
y	y	_AVERR_	Average Error Function	0.349891	0.370541	.
y	y	_ERR_	Error Function	3245.59	3440.105	.
y	y	_MISC_	Misclassification Rate	0.153083	0.162215	.
y	y	_WRONG_	Number of Wrong Classificati...	710	753	.

Picture 135. Optimization Start, Parameter Estimates for Cap and Floor Neural Network 100 iterations 4 hidden units

```
The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                Gradient
                                               Objective
 N Parameter                     Estimate       Function

 1 REP_age_H11                  -0.007847              0
 2 REP_campaign_H11             -0.018255              0
 3 REP_duration_H11             -0.044456              0
 4 REP_age_H12                  -0.040579              0
 5 REP_campaign_H12              0.013425              0
 6 REP_duration_H12              0.091171              0
 7 REP_age_H13                  -0.277730              0
 8 REP_campaign_H13              0.144563              0
 9 REP_duration_H13              0.104470              0
10 REP_age_H14                   0.032080              0
```

...

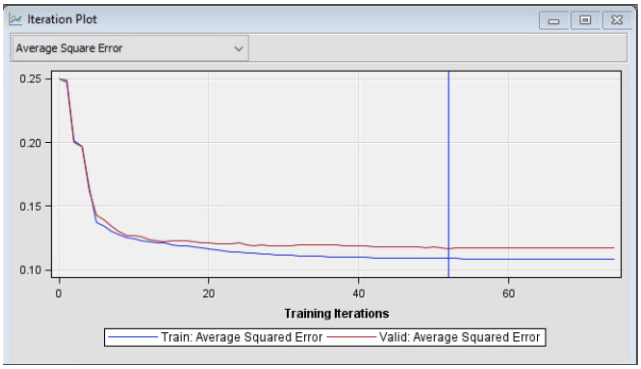
```
120 poutcomenonexistent_H14      0.060935              0
121 BIAS_H11                    -0.283225              0
122 BIAS_H12                     1.216321              0
123 BIAS_H13                     0.941225              0
124 BIAS_H14                     0.044303              0
125 H11_yyes                            0       0.035687
126 H12_yyes                            0      -0.000659
127 H13_yyes                            0      -0.026665
128 H14_yyes                            0      -0.021828
129 BIAS_yyes                           0              0

Value of Objective Function = 0.6931471806
```

Picture 136. Optimization Results for Cap and Floor Neural Network 100 iterations 4 hidden units

```
                                Optimization Results

Iterations                                   74  Function Calls                          177
Gradient Calls                               82  Active Constraints                        0
Objective Function                  0.348508437  Max Abs Gradient Element       0.0003470355
Slope of Search Direction          -0.000030357

Convergence criterion (FCONV=0.0001) satisfied.
```

Convergence criterion satisfied, and it was observed after approximately 74 iterations.

Picture 137. Iteration Plot (ASE) for Cap and Floor Neural Network 100 iterations 4 hidden units



As seen in the graph, approximately 50 iterations are enough after that it gives us the same results. The iteration plot shows optimal validation average squared error occurring on this iteration.

Cap and Floor Neural Network Average Squared Error 100 iterations, 5 hidden units

Picture 138. Results for Cap and Floor Neural Network 100 iterations 5 hidden units

Results - Node: C&F NN ASE 100 H5 Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Freedom	4638	.	.
y	y	_DFE_	Degrees of Freedom for Error	4477	.	.
y	y	_DFM_	Model Degrees of Freedom	161	.	.
y	y	_NW_	Number of Estimated Weights	161	.	.
y	y	_AIC_	Akaike's Information Criterion	3569.426	.	.
y	y	_SBC_	Schwarz's Bayesian Criterion	4606.595	.	.
y	y	_ASE_	Average Squared Error	0.108644	0.116917	.
y	y	_MAX_	Maximum Absolute Error	0.996257	0.99883	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squared Error	0.329612	0.341931	.
y	y	_SSE_	Sum of Squared Errors	1007.784	1085.458	.
y	y	_SUMW_	Sum of Case Weights Times ...	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.116458	.	.
y	y	_MSE_	Mean Squared Error	0.112551	0.116917	.
y	y	_RFPE_	Root Final Prediction Error	0.34126	.	.
y	y	_RMSE_	Root Mean Squared Error	0.335487	0.341931	.
y	y	_AVERR_	Average Error Function	0.350089	0.373267	.
y	y	_ERR_	Error Function	3247.426	3465.411	.
y	y	_MISC_	Misclassification Rate	0.148555	0.163292	.
y	y	_WRONG_	Number of Wrong Classificati...	689	758	.

Picture 139. Optimization Start, Parameter Estimates for Cap and Floor Neural Network 100 iterations 5 hidden units

```
The NEURAL Procedure

                    Optimization Start
                  Parameter Estimates
                                                Gradient
                                               Objective
 N Parameter                     Estimate       Function

 1 REP_age_H11                  -0.007847              0
 2 REP_campaign_H11             -0.018255              0
 3 REP_duration_H11             -0.044456              0
 4 REP_age_H12                  -0.040579              0
 5 REP_campaign_H12              0.013425              0
 6 REP_duration_H12              0.091171              0
 7 REP_age_H13                  -0.277730              0
 8 REP_campaign_H13              0.144563              0
 9 REP_duration_H13              0.104470              0
10 REP_age_H14                   0.032080              0
```

…

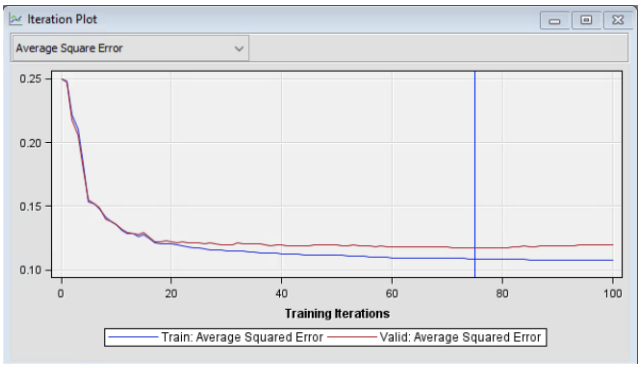
```
150 poutcomenonexistent_H15      -0.228058            0
151 BIAS_H11                     -1.837805            0
152 BIAS_H12                     -2.928428            0
153 BIAS_H13                     -1.950622            0
154 BIAS_H14                      0.712494            0
155 BIAS_H15                     -0.711007            0
156 H11_yyes                             0     0.003523
157 H12_yyes                             0    -0.001103
158 H13_yyes                             0     0.000742
159 H14_yyes                             0     0.013720
160 H15_yyes                             0    -0.022811
161 BIAS_yyes                            0            0

Value of Objective Function = 0.6931471806
```

Picture 140. Optimization Results for Cap and Floor Neural Network 100 iterations 5 hidden units

```
                                    Optimization Results

Iterations                                  100  Function Calls                          247
Gradient Calls                              114  Active Constraints                        0
Objective Function                 0.3461518038  Max Abs Gradient Element     0.0033957458
Slope of Search Direction         -0.000080289

QUANEW needs more than 100 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Notice the warning message. It can be interpreted to mean that the model-fitting process did not converge. The maximum number of iterations, by default, is 100. Apparently, this is not enough for the network training process to converge.

Picture 141. Iteration Plot (ASE) for Cap and Floor Neural Network 100 iterations 5 hidden units



As seen in the graph, approximately 73 iterations are enough after that it gives us the same results.The iteration plot shows optimal validation average squared error occurring on this iteration.

Cap and Floor Neural Network Average Squared Error 100 iterations, 6 hidden units

Picture 142. Results for Cap and Floor Neural Network 100 iterations 6 hidden units

Results - Node: C&F NN ASE 100 H6 Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Freedom	4638	.	.
y	y	_DFE_	Degrees of Freedom for Error	4445	.	.
y	y	_DFM_	Model Degrees of Freedom	193	.	.
y	y	_NW_	Number of Estimated Weights	193	.	.
y	y	_AIC_	Akaike's Information Criterion	3596.786	.	.
y	y	_SBC_	Schwarz's Bayesian Criterion	4840.1	.	.
y	y	_ASE_	Average Squared Error	0.10795	0.116799	.
y	y	_MAX_	Maximum Absolute Error	0.990368	0.999578	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squared Error	0.328558	0.341758	.
y	y	_SSE_	Sum of Squared Errors	1001.345	1084.36	.
y	y	_SUMW_	Sum of Case Weights Times ...	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.117324	.	.
y	y	_MSE_	Mean Squared Error	0.112637	0.116799	.
y	y	_RFPE_	Root Final Prediction Error	0.342526	.	.
y	y	_RMSE_	Root Mean Squared Error	0.335615	0.341758	.
y	y	_AVERR_	Average Error Function	0.346139	0.373235	.
y	y	_ERR_	Error Function	3210.786	3465.112	.
y	y	_MISC_	Misclassification Rate	0.148124	0.16006	.
y	y	_WRONG_	Number of Wrong Classificati...	687	743	.

Picture 143. Optimization Start, Parameter Estimates for Cap and Floor Neural Network 100 iterations 6 hidden units

```
The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                  Gradient
                                                 Objective
  N Parameter                      Estimate       Function

  1 REP_age_H11                   -0.007847              0
  2 REP_campaign_H11              -0.018255              0
  3 REP_duration_H11              -0.044456              0
  4 REP_age_H12                   -0.040579              0
  5 REP_campaign_H12               0.013425              0
  6 REP_duration_H12               0.091171              0
  7 REP_age_H13                   -0.277730              0
  8 REP_campaign_H13               0.144563              0
  9 REP_duration_H13               0.104470              0
 10 REP_age_H14                    0.032080              0
```

...

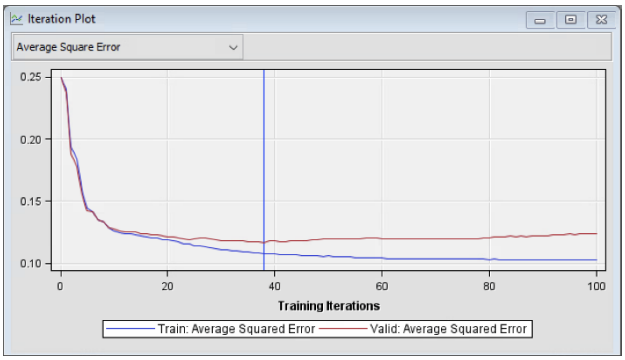
```
190 H14_yyes                          0        -0.029468
191 H15_yyes                          0        -0.016548
192 H16_yyes                          0         0.014591
193 BIAS_yyes                         0                0

Value of Objective Function = 0.6931471806
```

Picture 144. Optimization Results for Cap and Floor Neural Network 100 iterations 6 hidden units

```
                                    Optimization Results

Iterations                                  100  Function Calls                          241
Gradient Calls                              111  Active Constraints                        0
Objective Function                 0.3292857134  Max Abs Gradient Element     0.0008820123
Slope of Search Direction         -0.000110294

QUANEW needs more than 100 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Notice the warning message. It can be interpreted to mean that the model-fitting process did not converge. The maximum number of iterations, by default, is 100. Apparently, this is not enough for the network training process to converge.

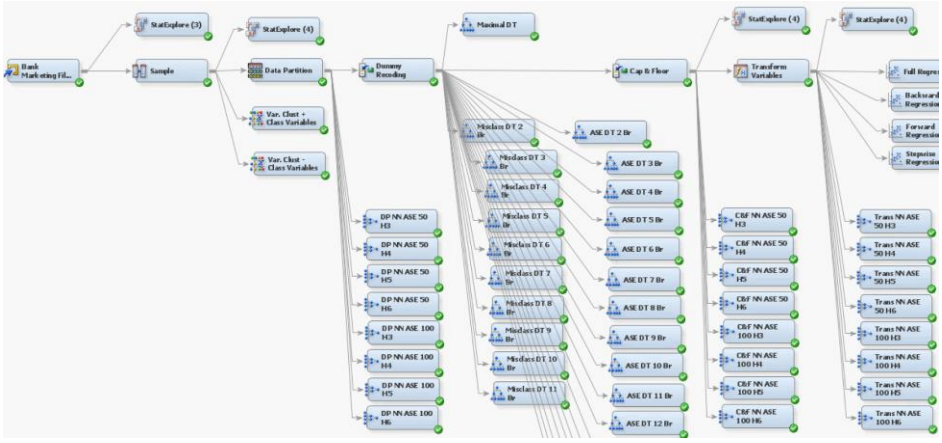
Picture 145. Iteration Plot (ASE) for Cap and Floor Neural Network 100 iterations 6 hidden units



As seen in the graph, approximately 38 iterations are enough after that it gives us the same results.The iteration plot shows optimal validation average squared error occurring on this iteration.

Transform Variables Neural Network

Picture 146. Adding Transform Variables Neural Networks into the Model



We added Neural Network to the Transform Variables node, we denied preliminary training in optimisation. Then we increased Hidden layers in network settings and we carried 6 Neural Networks with 50 Iterations and the next 5 with 100 Iterations. We looked for the lowest ASE and Misclassification rate in our results for each Neural Network to find the best optimisation model. We also noted the Parameter Estimation which determined the complexity of our model. We noted Parameter Estimation because in some cases the ASE or Misclassification rate was the same and it was harder for us to decide which could be the best model. With the help of Parameter Estimation, we could decide which was a less complex structure and we selected that. For our best model, TV NN 50 Iterations and 4 Hidden Units and TV NN 100 Iterations

and 4 Hidden Units , the best results were given after 39 iterations and it showed a “QUANEW” warning message because may be struggling to converge to a solution within the given number of iterations. Convergence was observed for TV NN 100 Iterations and 4 Hidden Units.

Table 15. Transform Variables Neural Networks Results

Neural Network	Number of Iterations	Number of Hidden Layers	Misclassification	ASE	Parameter Estimate
TV NN ASE 50	50	3	0.172340	0.119533	97
TV NN ASE 50	50	4	0.163723	0.116953	129
TV NN ASE 50	50	5	0.168677	0.119392	161
TV NN ASE 50	50	6	0.169970	0.119281	193
TV NN ASE 100	100	3	0.172340	0.119533	97
TV NN ASE 100	100	4	0.163723	0.116953	129
TV NN ASE 100	100	5	0.168677	0.119392	161
TV NN ASE 100	100	6	0.169970	0.119281	193

The lowest ASE was for TV NN ASE with 50 Iterations and 4 Hidden Units and for TV NN ASE with 100 Iterations with 4 Hidden Units, these were our best Neural Network models for Transform Variables. Overall the ASE almost remained the same after we increased the number of Iterations.

Transform Variables Neural Network 50 iterations, 3 hidden units

Picture 147. Results for Transform Variables Neural Network 50 iterations 3 hidden units

Results - Node: Trans NN ASE 50 H3 Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Freedom	4638	.	.
y	y	_DFE_	Degrees of Freedom for Error	4541	.	.
y	y	_DFM_	Model Degrees of Freedom	97	.	.
y	y	_NW_	Number of Estimated Weights	97	.	.
y	y	_AIC_	Akaike's Information Criterion	3517.403	.	.
y	y	_SBC_	Schwarz's Bayesian Criterion	4142.281	.	.
y	y	_ASE_	Average Squared Error	0.112166	0.119533	.
y	y	_MAX_	Maximum Absolute Error	0.99409	0.998701	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squared Error	0.334912	0.345736	.
y	y	_SSE_	Sum of Squared Errors	1040.45	1109.746	.
y	y	_SUMW_	Sum of Case Weights Times ...	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.116958	.	.
y	y	_MSE_	Mean Squared Error	0.114562	0.119533	.
y	y	_RFPE_	Root Final Prediction Error	0.341991	.	.
y	y	_RMSE_	Root Mean Squared Error	0.33847	0.345736	.
y	y	_AVERR_	Average Error Function	0.35828	0.376465	.
y	y	_ERR_	Error Function	3323.403	3495.099	.
y	y	_MISC_	Misclassification Rate	0.148555	0.17234	.
y	y	_WRONG_	Number of Wrong Classificati...	689	800	.

Picture 148. Optimization Start, Parameter Estimates for Transform Variables Neural Network 50 iterations 3 hidden units

```
The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                Gradient
                                               Objective
 N Parameter                     Estimate       Function

 1 LOG_REP_campaign_H11          -0.007847              0
 2 LOG_REP_duration_H11          -0.018255              0
 3 REP_age_H11                   -0.044456              0
 4 LOG_REP_campaign_H12          -0.040579              0
 5 LOG_REP_duration_H12           0.013425              0
 6 REP_age_H12                    0.091171              0
 7 LOG_REP_campaign_H13          -0.277730              0
 8 LOG_REP_duration_H13           0.144563              0
 9 REP_age_H13                    0.104470              0
10 REP_day_of_weekFr_H11          0.032080              0
```

...

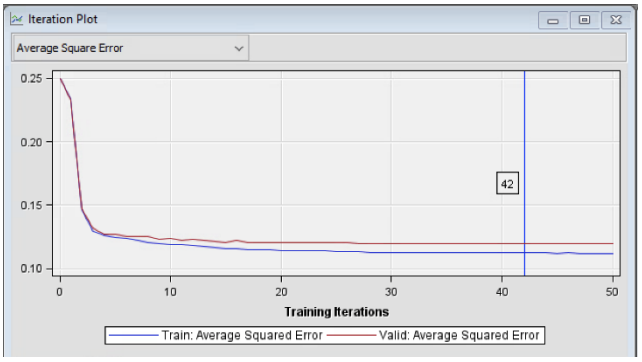
```
90 poutcomenonexistent_H13        0.194201              0
91 BIAS_H11                       0.628974              0
92 BIAS_H12                       0.202838              0
93 BIAS_H13                      -1.278167              0
94 H11_yyes                              0      -0.010081
95 H12_yyes                              0       0.048651
96 H13_yyes                              0      -0.005972
97 BIAS_yyes                             0              0

Value of Objective Function = 0.6931471806
```

Picture 149. Optimization Results for Transform Variables Neural Network 50 iterations 3 hidden units

```
                                    Optimization Results

Iterations                                   50  Function Calls                          59
Jacobian Calls                               52  Active Constraints                       0
Objective Function                0.3574654928  Max Abs Gradient Element        0.0028893823
Lambda                            0.0076461945  Actual Over Pred Change         0.1311245772
Radius                            0.1815636367

LEVMAR needs more than 50 iterations or 2147483647 function calls.

WARNING: LEVMAR Optimization cannot be completed.
```

Notice the warning message. It can be interpreted to mean that the model-fitting process did not converge. The maximum number of iterations, by default, is 50. Apparently, this is not enough for the network training process to converge.

Picture 150. Iteration Plot (ASE) for Transform Variables Neural Network 50 iterations 3 hidden units



As seen in the graph, approximately 42 iterations are enough after that it gives us the same results.The iteration plot shows optimal validation average squared error occurring on this iteration.

Transform Variables Neural Network 50 iterations, 4 hidden units

Picture 151. Results for Transform Variables Neural Network 50 iterations 4 hidden units

Results - Node: Trans NN ASE 50 H4 Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of...	4638	.	.
y	y	_DFE_	Degrees of Free...	4509	.	.
y	y	_DFM_	Model Degrees ...	129	.	.
y	y	_NW_	Number of Esti...	129	.	.
y	y	_AIC_	Akaike's Informa...	3508.224	.	.
y	y	_SBC_	Schwarz's Baye...	4339.247	.	.
y	y	_ASE_	Average Square...	0.108624	0.116953	.
y	y	_MAX_	Maximum Absol...	0.998356	0.997845	.
y	y	_DIV_	Divisor for ASE ...	9276	9284	.
y	y	_NOBS_	Sum of Frequen...	4638	4642	.
y	y	_RASE_	Root Average Sq...	0.329582	0.341984	.
y	y	_SSE_	Sum of Squared...	1007.601	1085.793	.
y	y	_SUMW_	Sum of Case W...	9276	9284	.
y	y	_FPE_	Final Prediction ...	0.11484	.	.
y	y	_MSE_	Mean Squared E...	0.111732	0.116953	.
y	y	_RFPE_	Root Final Predi...	0.33888	.	.
y	y	_RMSE_	Root Mean Squa...	0.334264	0.341984	.
y	y	_AVERR_	Average Error Fu...	0.350391	0.3735	.
y	y	_ERR_	Error Function	3250.224	3467.574	.
y	y	_MISC_	Misclassification...	0.151574	0.163723	.
y	y	_WRONG_	Number of Wron...	703	760	.

Picture 152. Optimization Start, Parameter Estimates for Transform Variables Neural Network 50 iterations 4 hidden units

```
The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                Gradient
                                               Objective
 N Parameter                     Estimate       Function

 1 LOG_REP_campaign_H11          -0.007847              0
 2 LOG_REP_duration_H11          -0.018255              0
 3 REP_age_H11                   -0.044456              0
 4 LOG_REP_campaign_H12          -0.040579              0
 5 LOG_REP_duration_H12           0.013425              0
 6 REP_age_H12                    0.091171              0
 7 LOG_REP_campaign_H13          -0.277730              0
 8 LOG_REP_duration_H13           0.144563              0
 9 REP_age_H13                    0.104470              0
10 LOG_REP_campaign_H14           0.032080              0
```

…

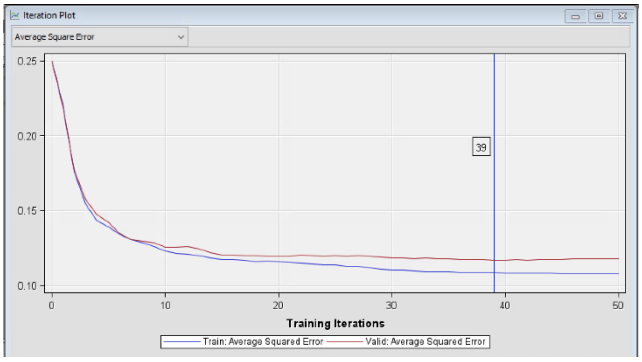
```
120 poutcomenonexistent_H14        0.060935              0
121 BIAS_H11                      -0.283225              0
122 BIAS_H12                       1.216321              0
123 BIAS_H13                       0.941225              0
124 BIAS_H14                       0.044303              0
125 H11_yyes                              0       0.032275
126 H12_yyes                              0       0.000795
127 H13_yyes                              0      -0.051967
128 H14_yyes                              0       0.048515
129 BIAS_yyes                             0              0

Value of Objective Function = 0.6931471806
```

Picture 152. Optimization Results for Transform Variables Neural Network 50 iterations 4 hidden units

```
                                   Optimization Results

Iterations                                   50  Function Calls                          131
Gradient Calls                               60  Active Constraints                        0
Objective Function                 0.3478759036  Max Abs Gradient Element       0.0032004082
Slope of Search Direction           -0.00041494

QUANEW needs more than 50 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Notice the warning message.The optimization algorithm (QUANEW) may be struggling to converge to a solution within the given number of iterations The maximum number of iterations, by default, is 50. Apparently, this is not enough for the network training process to converge.

Picture 153. Iteration Plot (ASE) for Transform Variables Neural Network 50 iterations 4 hidden units



As seen in the graph, approximately 39 iterations are enough after that it gives us the same results. The iteration plot shows optimal validation average squared error occurring on this iteration.

Transform Variables Neural Network 50 iterations, 5 hidden units

Picture 154. Results for Transform Variables Neural Network 50 iterations 5 hidden units

Results - Node: Trans NN ASE 50 H5 Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Freedom	4638	.	.
y	y	_DFE_	Degrees of Freedom for Error	4477	.	.
y	y	_DFM_	Model Degrees of Freedom	161	.	.
y	y	_NW_	Number of Estimated Weights	161	.	.
y	y	_AIC_	Akaike's Information Criterion	3652.937	.	.
y	y	_SBC_	Schwarz's Bayesian Criterion	4690.105	.	.
y	y	_ASE_	Average Squared Error	0.111466	0.119392	.
y	y	_MAX_	Maximum Absolute Error	0.998581	0.998874	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squared Error	0.333865	0.345532	.
y	y	_SSE_	Sum of Squared Errors	1033.956	1108.439	.
y	y	_SUMW_	Sum of Case Weights Times Freq	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.119483	.	.
y	y	_MSE_	Mean Squared Error	0.115474	0.119392	.
y	y	_RFPE_	Root Final Prediction Error	0.345663	.	.
y	y	_RMSE_	Root Mean Squared Error	0.339815	0.345532	.
y	y	_AVERR_	Average Error Function	0.359092	0.375687	.
y	y	_ERR_	Error Function	3330.937	3487.875	.
y	y	_MISC_	Misclassification Rate	0.15028	0.168677	.
y	y	_WRONG_	Number of Wrong Classifications	697	783	.

Picture 155. Optimization Start, Parameter Estimates for Transform Variables Neural Network 50 iterations 5 hidden units

```
The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                Gradient
                                               Objective
  N Parameter                     Estimate       Function

  1 LOG_REP_campaign_H11          -0.007847              0
  2 LOG_REP_duration_H11          -0.018255              0
  3 REP_age_H11                   -0.044456              0
  4 LOG_REP_campaign_H12          -0.040579              0
  5 LOG_REP_duration_H12           0.013425              0
  6 REP_age_H12                    0.091171              0
  7 LOG_REP_campaign_H13          -0.277730              0
  8 LOG_REP_duration_H13           0.144563              0
  9 REP_age_H13                    0.104470              0
 10 LOG_REP_campaign_H14           0.032080              0
```

…

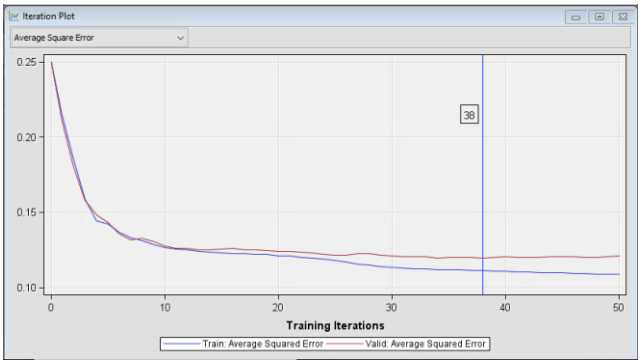
```
150 poutcomenonexistent_H15      -0.228058          0
151 BIAS_H11                     -1.837805          0
152 BIAS_H12                     -2.928428          0
153 BIAS_H13                     -1.950622          0
154 BIAS_H14                      0.712494          0
155 BIAS_H15                     -0.711007          0
156 H11_yyes                             0   0.003179
157 H12_yyes                             0  -0.000493
158 H13_yyes                             0  -0.000237
159 H14_yyes                             0   0.088569
160 H15_yyes                             0  -0.093114
161 BIAS_yyes                            0          0

Value of Objective Function = 0.6931471806
```

Picture 156. Optimization Results for Transform Variables Neural Network 50 iterations 5 hidden units

```
                                        Optimization Results

Iterations                                     50  Function Calls                          127
Gradient Calls                                 56  Active Constraints                        0
Objective Function                 0.3510896762  Max Abs Gradient Element         0.0073540875
Slope of Search Direction          -0.000799645

QUANEW needs more than 50 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Notice the warning message.The optimization algorithm (QUANEW) may be struggling to converge to a solution within the given number of iterations The maximum number of iterations, by default, is 50. Apparently, this is not enough for the network training process to converge.

Picture 157. Iteration Plot (ASE) for Transform Variables Neural Network 50 iterations 5 hidden units



As seen in the graph, approximately 38 iterations are enough after that it gives us the same results.The iteration plot shows optimal validation average squared error occurring on this iteration.

Transform Variables Neural Network 50 iterations, 6 hidden units

Picture 158. Results for Transform Variables Neural Network 50 iterations 6 hidden units

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Fr...	4638	.	.
y	y	_DFE_	Degrees of Freedo...	4445	.	.
y	y	_DFM_	Model Degrees of ...	193	.	.
y	y	_NW_	Number of Estimat...	193	.	.
y	y	_AIC_	Akaike's Informatio...	3701.75	.	.
y	y	_SBC_	Schwarz's Bayesia...	4945.063	.	.
y	y	_ASE_	Average Squared E...	0.112462	0.119281	.
y	y	_MAX_	Maximum Absolute...	0.987289	0.993379	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squa...	0.335353	0.345371	.
y	y	_SSE_	Sum of Squared Er...	1043.194	1107.406	.
y	y	_SUMW_	Sum of Case Weig...	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.122228	.	.
y	y	_MSE_	Mean Squared Error	0.117345	0.119281	.
y	y	_RFPE_	Root Final Predicti...	0.349611	.	.
y	y	_RMSE_	Root Mean Square...	0.342556	0.345371	.
y	y	_AVERR_	Average Error Func...	0.357455	0.380522	.
y	y	_ERR_	Error Function	3315.75	3532.766	.
y	y	_MISC_	Misclassification R...	0.156964	0.16997	.
y	y	_WRONG_	Number of Wrong ...	728	789	.

Picture 159. Optimization Start, Parameter Estimates for Transform Variables Neural Network 50 iterations 6 hidden units

```
The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                  Gradient
                                                 Objective
  N Parameter                      Estimate       Function

  1 LOG_REP_campaign_H11          -0.007847              0
  2 LOG_REP_duration_H11          -0.018255              0
  3 REP_age_H11                   -0.044456              0
  4 LOG_REP_campaign_H12          -0.040579              0
  5 LOG_REP_duration_H12           0.013425              0
  6 REP_age_H12                    0.091171              0
  7 LOG_REP_campaign_H13          -0.277730              0
  8 LOG_REP_duration_H13           0.144563              0
  9 REP_age_H13                    0.104470              0
 10 LOG_REP_campaign_H14           0.032080              0
```

…

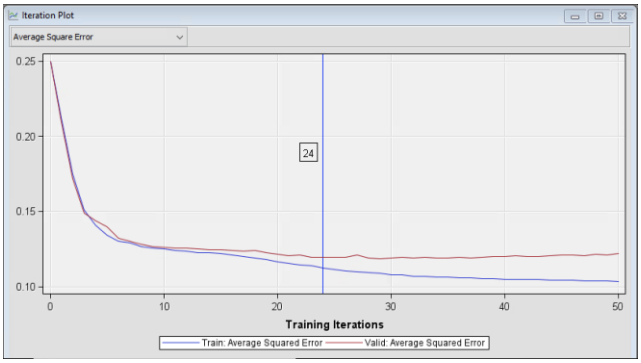
```
180 poutcomenonexistent_H16       -0.197132              0
181 BIAS_H11                       1.035066              0
182 BIAS_H12                       1.107832              0
183 BIAS_H13                       0.269358              0
184 BIAS_H14                       0.764637              0
185 BIAS_H15                       0.017364              0
186 BIAS_H16                       0.366147              0
187 H11_yyes                              0       0.023115
188 H12_yyes                              0      -0.016268
189 H13_yyes                              0      -0.031829
190 H14_yyes                              0       0.033332
191 H15_yyes                              0      -0.099538
192 H16_yyes                              0       0.005369
193 BIAS_yyes                             0              0

Value of Objective Function = 0.6931471806
```

Picture 160. Optimization Results for Transform Variables Neural Network 50 iterations 6 hidden units

```
                                    Optimization Results

Iterations                                    50  Function Calls                          126
Gradient Calls                                57  Active Constraints                        0
Objective Function                 0.3317327248  Max Abs Gradient Element       0.0122922377
Slope of Search Direction          -0.000631674

QUANEW needs more than 50 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Notice the warning message. The optimization algorithm (QUANEW) may be struggling to converge to a solution within the given number of iterations The maximum number of iterations, by default, is 50. Apparently, this is not enough for the network training process to converge.

Picture 161. Iteration Plot (ASE) for Transform Variables Neural Network 50 iterations 6 hidden units



As seen in the graph, 24 iterations are enough after that it gives us the same results. The iteration plot shows optimal validation average squared error occurring on this iteration.

Transform Variables Neural Network 100 iterations, 3 hidden units

Picture 162. Results for Transform Variables Neural Network 100 iterations 3 hidden units

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Fr...	4638	.	.
y	y	_DFE_	Degrees of Freedo...	4541	.	.
y	y	_DFM_	Model Degrees of ...	97	.	.
y	y	_NW_	Number of Estimat...	97	.	.
y	y	_AIC_	Akaike's Informatio...	3517.403	.	.
y	y	_SBC_	Schwarz's Bayesia...	4142.281	.	.
y	y	_ASE_	Average Squared E...	0.112166	0.119533	.
y	y	_MAX_	Maximum Absolute...	0.99409	0.998701	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squa...	0.334912	0.345736	.
y	y	_SSE_	Sum of Squared Er...	1040.45	1109.746	.
y	y	_SUMW_	Sum of Case Weig...	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.116958	.	.
y	y	_MSE_	Mean Squared Error	0.114562	0.119533	.
y	y	_RFPE_	Root Final Predicti...	0.341991	.	.
y	y	_RMSE_	Root Mean Square...	0.33847	0.345736	.
y	y	_AVERR_	Average Error Func...	0.35828	0.376465	.
y	y	_ERR_	Error Function	3323.403	3495.099	.
y	y	_MISC_	Misclassification R...	0.148555	0.17234	.
y	y	_WRONG_	Number of Wrong ...	689	800	.

Picture 163. Optimization Start, Parameter Estimates for Transform Variables Neural Network 100 iterations 3 hidden units

```
The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                   Gradient
                                                  Objective
 N Parameter                      Estimate        Function

 1 LOG_REP_campaign_H11          -0.007847               0
 2 LOG_REP_duration_H11          -0.018255               0
 3 REP_age_H11                   -0.044456               0
 4 LOG_REP_campaign_H12          -0.040579               0
 5 LOG_REP_duration_H12           0.013425               0
 6 REP_age_H12                    0.091171               0
 7 LOG_REP_campaign_H13          -0.277730               0
 8 LOG_REP_duration_H13           0.144563               0
 9 REP_age_H13                    0.104470               0
10 REP_day_of_weekFr_H11          0.032080               0
```

…

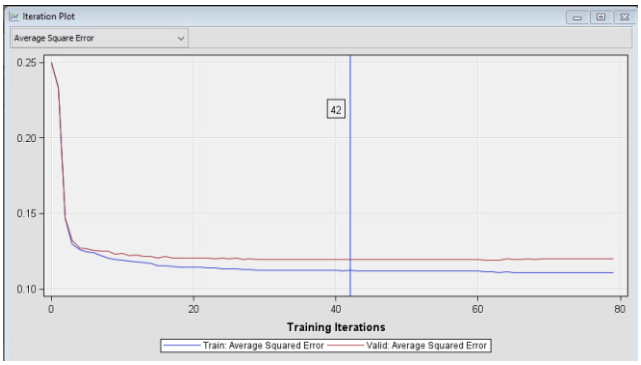
```
90 poutcomenonexistent_H13        0.194201               0
91 BIAS_H11                       0.628974               0
92 BIAS_H12                       0.202838               0
93 BIAS_H13                      -1.278167               0
94 H11_yyes                              0       -0.010081
95 H12_yyes                              0        0.048651
96 H13_yyes                              0       -0.005972
97 BIAS_yyes                             0               0

Value of Objective Function = 0.6931471806
```

Picture 164. Optimization Results for Transform Variables Neural Network 100 iterations 3 hidden units

```
                                   Optimization Results

Iterations                                   79  Function Calls                          93
Jacobian Calls                               82  Active Constraints                       0
Objective Function                 0.3529449948  Max Abs Gradient Element       0.0008841324
Lambda                             0.0047229692  Actual Over Pred Change        0.1727549062
Radius                             0.0986901203

Convergence criterion (FCONV=0.0001) satisfied.
```

Convergence was observed after 79 iterations

Picture 165. Iteration Plot (ASE) for Transform Variables Neural Network 100 iterations 3 hidden units



As seen in the graph, 42 iterations are enough after that it gives us the same results. The iteration plot shows optimal validation average squared error occurring on this iteration.

Transform Variables Neural Network 100 iterations, 4 hidden units

Picture 166. Results for Transform Variables Neural Network 100 iterations 4 hidden units

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Fr...	4638	.	.
y	y	_DFE_	Degrees of Freedo...	4509	.	.
y	y	_DFM_	Model Degrees of ...	129	.	.
y	y	_NW_	Number of Estimat...	129	.	.
y	y	_AIC_	Akaike's Informatio...	3508.224	.	.
y	y	_SBC_	Schwarz's Bayesia...	4339.247	.	.
y	y	_ASE_	Average Squared E...	0.108624	0.116953	.
y	y	_MAX_	Maximum Absolute...	0.998356	0.997845	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squa...	0.329582	0.341984	.
y	y	_SSE_	Sum of Squared Er...	1007.601	1085.793	.
y	y	_SUMW_	Sum of Case Weig...	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.11484	.	.
y	y	_MSE_	Mean Squared Error	0.111732	0.116953	.
y	y	_RFPE_	Root Final Predicti...	0.33888	.	.
y	y	_RMSE_	Root Mean Square...	0.334264	0.341984	.
y	y	_AVERR_	Average Error Func...	0.350391	0.3735	.
y	y	_ERR_	Error Function	3250.224	3467.574	.
y	y	_MISC_	Misclassification R...	0.151574	0.163723	.
y	y	_WRONG_	Number of Wrong ...	703	760	.

Picture 167. Optimization Start, Parameter Estimates for Transform Variables Neural Network 100 iterations 4 hidden units

```
The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                  Gradient
                                                 Objective
  N Parameter                      Estimate       Function

  1 LOG_REP_campaign_H11          -0.007847              0
  2 LOG_REP_duration_H11          -0.018255              0
  3 REP_age_H11                   -0.044456              0
  4 LOG_REP_campaign_H12          -0.040579              0
  5 LOG_REP_duration_H12           0.013425              0
  6 REP_age_H12                    0.091171              0
  7 LOG_REP_campaign_H13          -0.277730              0
  8 LOG_REP_duration_H13           0.144563              0
  9 REP_age_H13                    0.104470              0
 10 LOG_REP_campaign_H14           0.032080              0
```

…

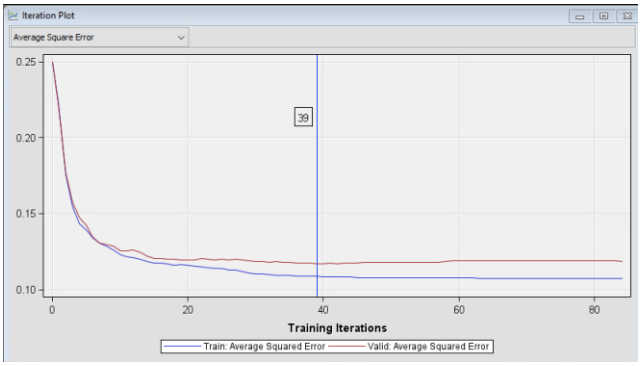
```
120 poutcomenonexistent_H14        0.060935              0
121 BIAS_H11                      -0.283225              0
122 BIAS_H12                       1.216321              0
123 BIAS_H13                       0.941225              0
124 BIAS_H14                       0.044303              0
125 H11_yyes                              0       0.032275
126 H12_yyes                              0       0.000795
127 H13_yyes                              0      -0.051967
128 H14_yyes                              0       0.048515
129 BIAS_yyes                             0              0

Value of Objective Function = 0.6931471806
```

Picture 168. Optimization Results for Transform Variables Neural Network 100 iterations 4 hidden units

```
                                    Optimization Results

Iterations                                   84  Function Calls                          211
Gradient Calls                               98  Active Constraints                        0
Objective Function              0.3451448251  Max Abs Gradient Element          0.001210644
Slope of Search Direction       -0.000020442

Convergence criterion (FCONV=0.0001) satisfied.

NOTE: At least one element of the gradient is greater than 1e-3.
```

Conversion was observed after 84 iterations. In this case, SAS Miner wants to make sure that these tweaks are not too big (greater than -3) during the optimization process

Picture 169. Iteration Plot (ASE) for Transform Variables Neural Network 100 iterations 4 hidden units



As seen in the graph, 39 iterations are enough after that it gives us the same results. The iteration plot shows optimal validation average squared error occurring on this iteration.

Transform Variables Neural Network 100 iterations, 5 hidden units

Picture 170. Results for Transform Variables Neural Network 100 iterations 5 hidden units

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Fr...	4638	.	.
y	y	_DFE_	Degrees of Freedo...	4477	.	.
y	y	_DFM_	Model Degrees of ...	161	.	.
y	y	_NW_	Number of Estimat...	161	.	.
y	y	_AIC_	Akaike's Informatio...	3652.937	.	.
y	y	_SBC_	Schwarz's Bayesia...	4690.105	.	.
y	y	_ASE_	Average Squared E...	0.111466	0.119392	.
y	y	_MAX_	Maximum Absolute...	0.998581	0.998874	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squa...	0.333865	0.345532	.
y	y	_SSE_	Sum of Squared Er...	1033.956	1108.439	.
y	y	_SUMW_	Sum of Case Weig...	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.119483	.	.
y	y	_MSE_	Mean Squared Error	0.115474	0.119392	.
y	y	_RFPE_	Root Final Predicti...	0.345663	.	.
y	y	_RMSE_	Root Mean Square...	0.339815	0.345532	.
y	y	_AVERR_	Average Error Func...	0.359092	0.375687	.
y	y	_ERR_	Error Function	3330.937	3487.875	.
y	y	_MISC_	Misclassification R...	0.15028	0.168677	.
y	y	_WRONG_	Number of Wrong ...	697	783	.

Picture 171. Optimization Start, Parameter Estimates for Transform Variables Neural Network 100 iterations 5 hidden units

```
The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                Gradient
                                               Objective
  N Parameter                     Estimate      Function

  1 LOG_REP_campaign_H11          -0.007847             0
  2 LOG_REP_duration_H11          -0.018255             0
  3 REP_age_H11                   -0.044456             0
  4 LOG_REP_campaign_H12          -0.040579             0
  5 LOG_REP_duration_H12           0.013425             0
  6 REP_age_H12                    0.091171             0
  7 LOG_REP_campaign_H13          -0.277730             0
  8 LOG_REP_duration_H13           0.144563             0
  9 REP_age_H13                    0.104470             0
 10 LOG_REP_campaign_H14           0.032080             0
```

…

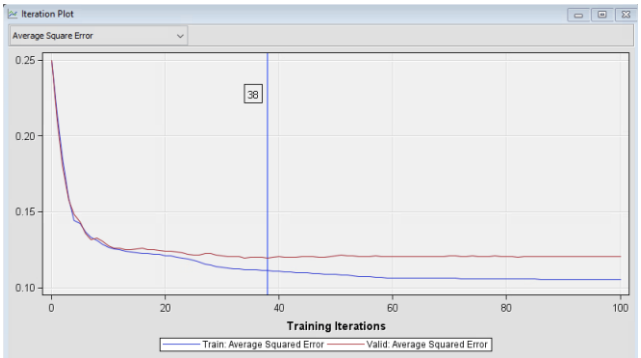
```
150 poutcomenonexistent_H15       -0.228058             0
151 BIAS_H11                      -1.837805             0
152 BIAS_H12                      -2.928428             0
153 BIAS_H13                      -1.950622             0
154 BIAS_H14                       0.712494             0
155 BIAS_H15                      -0.711007             0
156 H11_yyes                              0      0.003179
157 H12_yyes                              0     -0.000493
158 H13_yyes                              0     -0.000237
159 H14_yyes                              0      0.088569
160 H15_yyes                              0     -0.093114
161 BIAS_yyes                             0             0

Value of Objective Function = 0.6931471806
```

Picture 172. Optimization Results for Transform Variables Neural Network 100 iterations 5 hidden units

```
                                   Optimization Results

Iterations                                    100  Function Calls                          241
Gradient Calls                                110  Active Constraints                        0
Objective Function                   0.3388366523  Max Abs Gradient Element       0.0028737785
Slope of Search Direction           -0.000151916

QUANEW needs more than 100 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Notice the warning message. The optimization algorithm (QUANEW) may be struggling to converge to a solution within the given number of iterations The maximum number of iterations, by default, is 100. Apparently, this is not enough for the network training process to converge

Picture 173. Iteration Plot (ASE) for Transform Variables Neural Network 100 iterations 5 hidden units



As seen in the graph, 38 iterations are enough after that it gives us the same results. The iteration plot shows optimal validation average squared error occurring on this iteration.

Transform Variables Neural Network 100 iterations, 6 hidden units

Picture 174. Results for Transform Variables Neural Network 100 iterations 6 hidden units

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Fr...	4638	.	.
y	y	_DFE_	Degrees of Freedo...	4445	.	.
y	y	_DFM_	Model Degrees of F...	193	.	.
y	y	_NW_	Number of Estimat...	193	.	.
y	y	_AIC_	Akaike's Informatio...	3701.75	.	.
y	y	_SBC_	Schwarz's Bayesia...	4945.063	.	.
y	y	_ASE_	Average Squared E...	0.112462	0.119281	.
y	y	_MAX_	Maximum Absolute ...	0.987289	0.993379	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squa...	0.335353	0.345371	.
y	y	_SSE_	Sum of Squared Er...	1043.194	1107.406	.
y	y	_SUMW_	Sum of Case Weig...	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.122228	.	.
y	y	_MSE_	Mean Squared Error	0.117345	0.119281	.
y	y	_RFPE_	Root Final Predictio...	0.349611	.	.
y	y	_RMSE_	Root Mean Square...	0.342556	0.345371	.
y	y	_AVERR_	Average Error Funct...	0.357455	0.380522	.
y	y	_ERR_	Error Function	3315.75	3532.766	.
y	y	_MISC_	Misclassification R...	0.156964	0.16997	.
y	y	_WRONG_	Number of Wrong ...	728	789	.

Picture 175. Optimization Start, Parameter Estimates for Transform Variables Neural Network 100 iterations 6 hidden units

```
The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                Gradient
                                               Objective
  N Parameter                     Estimate      Function

  1 LOG_REP_campaign_H11          -0.007847             0
  2 LOG_REP_duration_H11          -0.018255             0
  3 REP_age_H11                   -0.044456             0
  4 LOG_REP_campaign_H12          -0.040579             0
  5 LOG_REP_duration_H12           0.013425             0
  6 REP_age_H12                    0.091171             0
  7 LOG_REP_campaign_H13          -0.277730             0
  8 LOG_REP_duration_H13           0.144563             0
  9 REP_age_H13                    0.104470             0
 10 LOG_REP_campaign_H14           0.032080             0
```

…

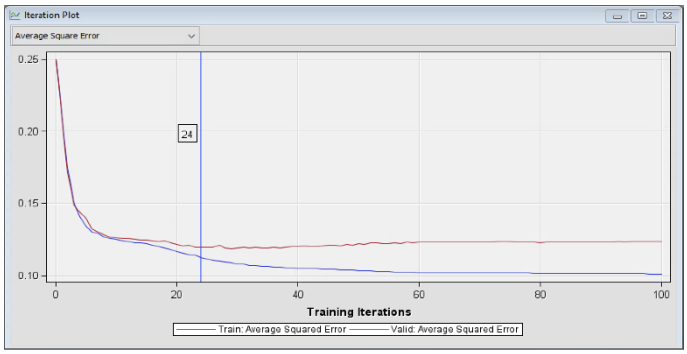
```
180 poutcomenonexistent_H16       -0.197132             0
181 BIAS_H11                       1.035066             0
182 BIAS_H12                       1.107832             0
183 BIAS_H13                       0.269358             0
184 BIAS_H14                       0.764637             0
185 BIAS_H15                       0.017364             0
186 BIAS_H16                       0.366147             0
187 H11_yyes                              0      0.023115
188 H12_yyes                              0     -0.016268
189 H13_yyes                              0     -0.031829
190 H14_yyes                              0      0.033332
191 H15_yyes                              0     -0.099538
192 H16_yyes                              0      0.005369
193 BIAS_yyes                             0             0

Value of Objective Function = 0.6931471806
```

Picture 176. Optimization Results for Transform Variables Neural Network 100 iterations 6 hidden units

```
                                    Optimization Results

Iterations                                    100  Function Calls                          244
Gradient Calls                                113  Active Constraints                        0
Objective Function                   0.3257504087  Max Abs Gradient Element     0.0015480911
Slope of Search Direction           -0.000028781

QUANEW needs more than 100 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Notice the warning message. The optimization algorithm (QUANEW) may be struggling to converge to a solution within the given number of iterations The maximum number of iterations, by default, is 50. Apparently, this is not enough for the network training process to converge.

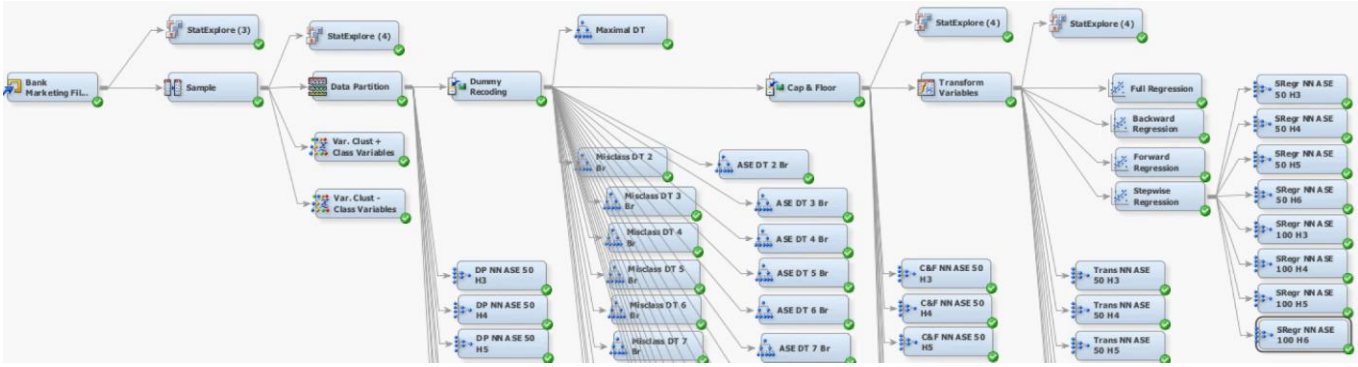
Picture 177. Iteration Plot (ASE) for Transform Variables Neural Network 100 iterations 6 hidden units



As seen in the graph, 24 iterations are enough after that it gives us the same results. The iteration plot shows optimal validation average squared error occurring on this iteration.

Stepwise Regression Neural Network

Picture 178. Adding Stepwise Regression Neural Networks into the Model



We added a Neural Network to the Stepwise node because it is our best Regression Node, and we denied preliminary training in optimization. Then we increased Hidden layers in network settings and we carried 6 Neural Networks with 50 Iterations and the next 5 with 100 Iterations. We looked for the lowest ASE and Misclassification rate in our results for each Neural Network to find the best optimisation model. We also noted the Parameter Estimation which determined the complexity of our model. We noted Parameter Estimation because in some cases the ASE or Misclassification rate was the same and it was harder for us to decide which could be the best model. With the help of Parameter Estimation, we could decide which was

a less complex structure and we selected that. For our best model, SR NN 100 Iterations and 5 Hidden Units, the best results were given after 95 iterations and it showed a “QUANEW” warning message because may be struggling to converge to a solution within the given number of iterations.

Table 16. Stepwise Regression Neural Network Results

Neural Network	Number of Iterations	Number of Hidden Layers	Misclassification	ASE	Parameter Estimate
SR NN ASE 50	50	3	0.172277	0.118650	67
SR NN ASE 50	50	4	0.173632	0.120201	89
SR NN ASE 50	50	5	0.171909	0.118800	111
SR NN ASE 50	50	6	0.170616	0.118298	133
SR NN ASE 100	100	3	0.172770	0.118650	67
SR NN ASE 100	100	4	0.173632	0.120201	89
SR NN ASE 100	100	5	0.167385	0.117893	111
SR NN ASE 100	100	6	0.170616	0.118298	133

The lowest ASE was for SR NN ASE with 100 Iterations and 5 Hidden Units, these were our best Neural Network models for Stepwise Regression. Overall the ASE reduced after we increased the number of Iterations.

Stepwise Regression Neural Network 50 iterations, 3 hidden units

Picture 179. Results for Stepwise Regression Neural Network 50 iterations 3 hidden units

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Fr...	4638	.	.
y	y	_DFE_	Degrees of Freedo...	4571	.	.
y	y	_DFM_	Model Degrees of ...	67	.	.
y	y	_NW_	Number of Estimat...	67	.	.
y	y	_AIC_	Akaike's Informatio...	3533.057	.	.
y	y	_SBC_	Schwarz's Bayesia...	3964.674	.	.
y	y	_ASE_	Average Squared E...	0.114849	0.11865	.
y	y	_MAX_	Maximum Absolute...	0.996118	0.994912	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squa...	0.338893	0.344457	.
y	y	_SSE_	Sum of Squared Er...	1065.336	1101.549	.
y	y	_SUMW_	Sum of Case Weig...	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.118215	.	.
y	y	_MSE_	Mean Squared Error	0.116532	0.11865	.
y	y	_RFPE_	Root Final Predicti...	0.343825	.	.
y	y	_RMSE_	Root Mean Square...	0.341368	0.344457	.
y	y	_AVERR_	Average Error Func...	0.366436	0.370807	.
y	y	_ERR_	Error Function	3399.057	3442.568	.
y	y	_MISC_	Misclassification R...	0.160414	0.17277	.
y	y	_WRONG_	Number of Wrong ...	744	802	.

Picture 180. Optimization Start, Parameter Estimates for Stepwise Regression Neural Network 50 iterations 3 hidden units

```
The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                    Gradient
                                                   Objective
 N Parameter                      Estimate          Function

 1 LOG_REP_campaign_H11          -0.009611                 0
 2 LOG_REP_duration_H11          -0.022358                 0
 3 LOG_REP_campaign_H12          -0.054447                 0
 4 LOG_REP_duration_H12          -0.049699                 0
 5 LOG_REP_campaign_H13           0.016443                 0
 6 LOG_REP_duration_H13           0.111661                 0
 7 REP_educationBasic_H11        -0.340148                 0
 8 REP_educationHS_H11            0.177052                 0
 9 REP_educationUn_Il_H11         0.127949                 0
10 REP_jobB_T_H_H11               0.039290                 0
```

…

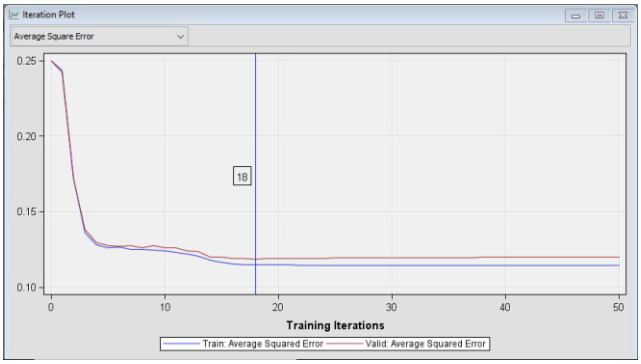
```
60 poutcomenonexistent_H13        0.317382                 0
61 BIAS_H11                      -0.220168                 0
62 BIAS_H12                       0.397161                 0
63 BIAS_H13                       1.303395                 0
64 H11_yyes                              0         -0.007759
65 H12_yyes                              0          0.040399
66 H13_yyes                              0          0.013491
67 BIAS_yyes                             0                 0

Value of Objective Function = 0.6931471806
```

Picture 181. Optimization Results for Stepwise Regression Neural Network 50 iterations 3 hidden units

```
                                   Optimization Results

Iterations                                   50  Function Calls                          58
Jacobian Calls                               52  Active Constraints                       0
Objective Function                 0.3643938736  Max Abs Gradient Element        0.0006617334
Lambda                             0.0303313853  Actual Over Pred Change         0.5128680514
Radius                             0.0329975122

LEVMAR needs more than 50 iterations or 2147483647 function calls.

WARNING: LEVMAR Optimization cannot be completed.
```

Notice the warning message. It can be interpreted to mean that the model-fitting process did not converge. The maximum number of iterations, by default, is 50. Apparently, this is not enough for the network training process to converge.

Picture 182. Iteration Plot (ASE) for Stepwise Regression Neural Network 50 iterations 3 hidden units



As seen in the graph, 18 iterations are enough after that it gives us the same results. The iteration plot shows optimal validation average squared error occurring on this iteration.

Stepwise Regression Neural Network 50 iterations, 4 hidden units

Picture 183. Results for Stepwise Regression Neural Network 50 iterations 4 hidden units

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Fr...	4638	.	.
y	y	_DFE_	Degrees of Freedo...	4549	.	.
y	y	_DFM_	Model Degrees of ...	89	.	.
y	y	_NW_	Number of Estimat...	89	.	.
y	y	_AIC_	Akaike's Informatio...	3575.591	.	.
y	y	_SBC_	Schwarz's Bayesia...	4148.933	.	.
y	y	_ASE_	Average Squared E...	0.114821	0.120201	.
y	y	_MAX_	Maximum Absolute...	0.997356	0.997316	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squa...	0.338853	0.3467	.
y	y	_SSE_	Sum of Squared Er...	1065.081	1115.948	.
y	y	_SUMW_	Sum of Case Weig...	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.119314	.	.
y	y	_MSE_	Mean Squared Error	0.117068	0.120201	.
y	y	_RFPE_	Root Final Predicti...	0.345419	.	.
y	y	_RMSE_	Root Mean Square...	0.342151	0.3467	.
y	y	_AVERR_	Average Error Func...	0.366278	0.376259	.
y	y	_ERR_	Error Function	3397.591	3493.192	.
y	y	_MISC_	Misclassification R...	0.158042	0.173632	.
y	y	_WRONG_	Number of Wrong ...	733	806	.

Picture 184. Optimization Start, Parameter Estimates for Stepwise Regression Neural Network 50 iterations 4 hidden units

```
The NEURAL Procedure

                    Optimization Start
                  Parameter Estimates
                                                Gradient
                                               Objective
 N Parameter                     Estimate       Function

 1 LOG_REP_campaign_H11          -0.009611              0
 2 LOG_REP_duration_H11          -0.022358              0
 3 LOG_REP_campaign_H12          -0.054447              0
 4 LOG_REP_duration_H12          -0.049699              0
 5 LOG_REP_campaign_H13           0.016443              0
 6 LOG_REP_duration_H13           0.111661              0
 7 LOG_REP_campaign_H14          -0.340148              0
 8 LOG_REP_duration_H14           0.177052              0
 9 REP_educationBasic_H11         0.127949              0
10 REP_educationHS_H11            0.039290              0
```

...

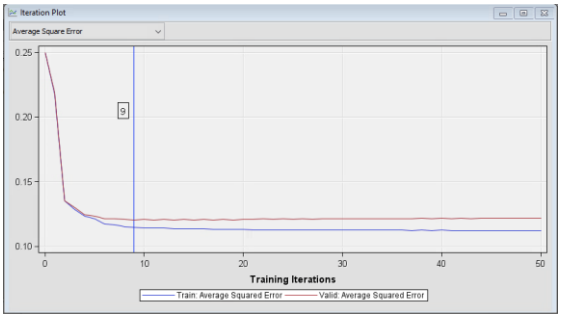
```
80 poutcomenonexistent_H14        0.015195              0
81 BIAS_H11                      -0.400147              0
82 BIAS_H12                      -0.642126              0
83 BIAS_H13                       0.585875              0
84 BIAS_H14                      -0.050069              0
85 H11_yyes                              0       0.025681
86 H12_yyes                              0       0.005571
87 H13_yyes                              0      -0.017007
88 H14_yyes                              0      -0.066564
89 BIAS_yyes                             0              0

Value of Objective Function = 0.6931471806
```

Picture 185. Optimization Results for Stepwise Regression Neural Network 50 iterations 4 hidden units

```
                                Optimization Results

Iterations                                   50  Function Calls                          57
Jacobian Calls                               52  Active Constraints                       0
Objective Function                 0.3571743603  Max Abs Gradient Element       0.000958374
Lambda                             0.0012742233  Actual Over Pred Change        0.567911366
Radius                              0.261773714

LEVMAR needs more than 50 iterations or 2147483647 function calls.

WARNING: LEVMAR Optimization cannot be completed.
```

Notice the warning message. It can be interpreted to mean that the model-fitting process did not converge. The maximum number of iterations, by default, is 50. Apparently, this is not enough for the network training process to converge.

Picture 186. Iteration Plot (ASE) for Stepwise Regression Neural Network 50 iterations 4 hidden units



As seen in the graph, 9 iterations are enough after that it gives us the same results.The iteration plot shows optimal validation average squared error occurring on this iteration.

Stepwise Regression Neural Network 50 iterations, 5 hidden units

Picture 187. Results for Stepwise Regression Neural Network 50 iterations 5 hidden units

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Fr...	4638	.	.
y	y	_DFE_	Degrees of Freedo...	4527	.	.
y	y	_DFM_	Model Degrees of ...	111	.	.
y	y	_NW_	Number of Estimat...	111	.	.
y	y	_AIC_	Akaike's Informatio...	3509.569	.	.
y	y	_SBC_	Schwarz's Bayesia...	4224.635	.	.
y	y	_ASE_	Average Squared E...	0.110543	0.1188	.
y	y	_MAX_	Maximum Absolute...	0.99358	0.994929	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squa...	0.33248	0.344673	.
y	y	_SSE_	Sum of Squared Er...	1025.397	1102.935	.
y	y	_SUMW_	Sum of Case Weig...	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.115964	.	.
y	y	_MSE_	Mean Squared Error	0.113254	0.1188	.
y	y	_RFPE_	Root Final Predicti...	0.340535	.	.
y	y	_RMSE_	Root Mean Square...	0.336532	0.344673	.
y	y	_AVERR_	Average Error Func...	0.354417	0.376068	.
y	y	_ERR_	Error Function	3287.569	3491.419	.
y	y	_MISC_	Misclassification R...	0.152221	0.171909	.
y	y	_WRONG_	Number of Wrong ...	706	798	.

Picture 188. Optimization Start, Parameter Estimates for Stepwise Regression Neural Network 50 iterations 5 hidden units

```
The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                Gradient
                                               Objective
  N Parameter                     Estimate      Function

  1 LOG_REP_campaign_H11          -0.009611             0
  2 LOG_REP_duration_H11          -0.022358             0
  3 LOG_REP_campaign_H12          -0.054447             0
  4 LOG_REP_duration_H12          -0.049699             0
  5 LOG_REP_campaign_H13           0.016443             0
  6 LOG_REP_duration_H13           0.111661             0
  7 LOG_REP_campaign_H14          -0.340148             0
  8 LOG_REP_duration_H14           0.177052             0
  9 LOG_REP_campaign_H15           0.127949             0
 10 LOG_REP_duration_H15           0.039290             0
```

…

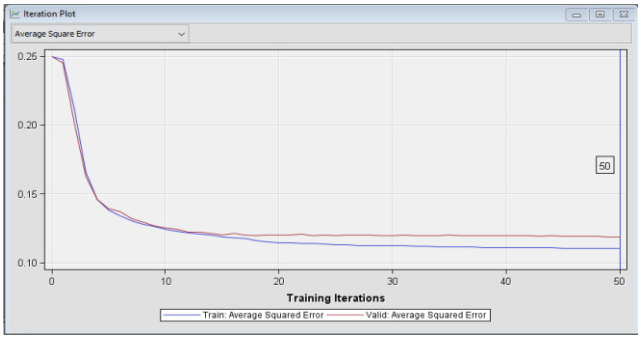
```
100 poutcomenonexistent_H15        0.323542             0
101 BIAS_H11                       0.505428             0
102 BIAS_H12                      -0.315904             0
103 BIAS_H13                      -0.594813             0
104 BIAS_H14                      -0.461791             0
105 BIAS_H15                       0.645601             0
106 H11_yyes                              0     -0.011265
107 H12_yyes                              0     -0.017772
108 H13_yyes                              0      0.006127
109 H14_yyes                              0     -0.033832
110 H15_yyes                              0      0.024929
111 BIAS_yyes                             0             0

Value of Objective Function = 0.6931471806
```

Picture 189. Optimization Results for Stepwise Regression Neural Network 50 iterations 5 hidden units

```
                                    Optimization Results

Iterations                                     50  Function Calls                          121
Gradient Calls                                 56  Active Constraints                        0
Objective Function                 0.3544166298  Max Abs Gradient Element       0.0042087233
Slope of Search Direction           -0.00016487

QUANEW needs more than 50 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Notice the warning message.The optimization algorithm (QUANEW) may be struggling to converge to a solution within the given number of iterations The maximum number of iterations, by default, is 50. Apparently, this is not enough for the network training process to converge.

Picture 190. Iteration Plot (ASE) for Stepwise Regression Neural Network 50 iterations 5 hidden units



As seen in the graph, 50 iterations are enough after that it gives us the same results. The iteration plot shows optimal validation average squared error occurring on this iteration.

Stepwise Regression Neural Network 50 iterations, 6 hidden units

Picture 191. Results for Stepwise Regression Neural Network 50 iterations 6 hidden units

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Fr...	4638	.	.
y	y	_DFE_	Degrees of Freedo...	4505	.	.
y	y	_DFM_	Model Degrees of ...	133	.	.
y	y	_NW_	Number of Estimat...	133	.	.
y	y	_AIC_	Akaike's Informatio...	3583.641	.	.
y	y	_SBC_	Schwarz's Bayesia...	4440.433	.	.
y	y	_ASE_	Average Squared E...	0.111796	0.118298	.
y	y	_MAX_	Maximum Absolute...	0.9943	0.998474	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squa...	0.334359	0.343945	.
y	y	_SSE_	Sum of Squared Er...	1037.02	1098.279	.
y	y	_SUMW_	Sum of Case Weig...	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.118397	.	.
y	y	_MSE_	Mean Squared Error	0.115097	0.118298	.
y	y	_RFPE_	Root Final Predicti...	0.344089	.	.
y	y	_RMSE_	Root Mean Square...	0.339259	0.343945	.
y	y	_AVERR_	Average Error Func...	0.357659	0.371199	.
y	y	_ERR_	Error Function	3317.641	3446.213	.
y	y	_MISC_	Misclassification R...	0.15718	0.170616	.
y	y	_WRONG_	Number of Wrong ...	729	792	.

Picture 192. Optimization Start, Parameter Estimates for Stepwise Regression Neural Network 50 iterations 6 hidden units

```
The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                Gradient
                                               Objective
  N Parameter                    Estimate       Function

  1 LOG_REP_campaign_H11         -0.009611              0
  2 LOG_REP_duration_H11         -0.022358              0
  3 LOG_REP_campaign_H12         -0.054447              0
  4 LOG_REP_duration_H12         -0.049699              0
  5 LOG_REP_campaign_H13          0.016443              0
  6 LOG_REP_duration_H13          0.111661              0
  7 LOG_REP_campaign_H14         -0.340148              0
  8 LOG_REP_duration_H14          0.177052              0
  9 LOG_REP_campaign_H15          0.127949              0
 10 LOG_REP_duration_H15          0.039290              0
```

…

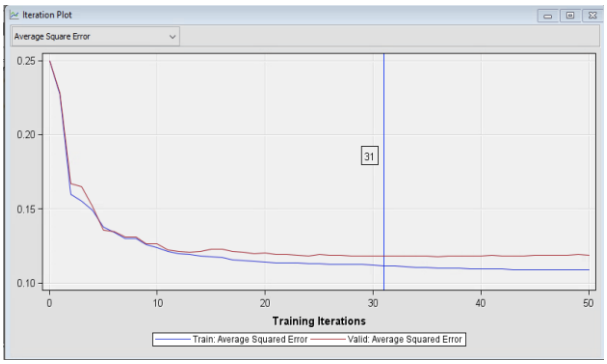
```
120 poutcomenonexistent_H16       0.074630              0
121 BIAS_H11                     -0.283225              0
122 BIAS_H12                      1.216321              0
123 BIAS_H13                      0.941225              0
124 BIAS_H14                      0.044303              0
125 BIAS_H15                      0.680919              0
126 BIAS_H16                      0.652068              0
127 H11_yyes                             0      -0.040848
128 H12_yyes                             0       0.019871
129 H13_yyes                             0      -0.022027
130 H14_yyes                             0      -0.030000
131 H15_yyes                             0      -0.010755
132 H16_yyes                             0      -0.009282
133 BIAS_yyes                            0              0

Value of Objective Function = 0.6931471806
```

Picture 192. Optimization Results for Stepwise Regression Neural Network 50 iterations 6 hidden units

```
                                  Optimization Results

Iterations                                   50  Function Calls                          129
Gradient Calls                               58  Active Constraints                        0
Objective Function                 0.348734798  Max Abs Gradient Element        0.0048825414
Slope of Search Direction          -0.00039496

QUANEW needs more than 50 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Notice the warning message. The optimization algorithm (QUANEW) may be struggling to converge to a solution within the given number of iterations The maximum number of iterations, by default, is 50. Apparently, this is not enough for the network training process to converge.

Picture 193. Iteration Plot (ASE) for Stepwise Regression Neural Network 50 iterations 6 hidden units



As seen in the graph, 31 iterations are enough after that it gives us the same results. The iteration plot shows optimal validation average squared error occurring on this iteration.

Stepwise Regression Neural Network 100 iterations, 3 hidden units

Picture 194. Results for Stepwise Regression Neural Network 100 iterations 3 hidden units

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Fr...	4638	.	.
y	y	_DFE_	Degrees of Freedo...	4571	.	.
y	y	_DFM_	Model Degrees of ...	67	.	.
y	y	_NW_	Number of Estimat...	67	.	.
y	y	_AIC_	Akaike's Informatio...	3533.057	.	.
y	y	_SBC_	Schwarz's Bayesia...	3964.674	.	.
y	y	_ASE_	Average Squared E...	0.114849	0.11865	.
y	y	_MAX_	Maximum Absolute...	0.996118	0.994912	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squa...	0.338893	0.344457	.
y	y	_SSE_	Sum of Squared Er...	1065.336	1101.549	.
y	y	_SUMW_	Sum of Case Weig...	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.118215	.	.
y	y	_MSE_	Mean Squared Error	0.116532	0.11865	.
y	y	_RFPE_	Root Final Predicti...	0.343825	.	.
y	y	_RMSE_	Root Mean Square...	0.341368	0.344457	.
y	y	_AVERR_	Average Error Func...	0.366436	0.370807	.
y	y	_ERR_	Error Function	3399.057	3442.568	.
y	y	_MISC_	Misclassification R...	0.160414	0.17277	.
y	y	_WRONG_	Number of Wrong ...	744	802	.

Picture 195. Optimization Start, Parameter Estimates for Stepwise Regression Neural Network 100 iterations 3 hidden units

```
The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                   Gradient
                                                  Objective
 N Parameter                      Estimate        Function

 1 LOG_REP_campaign_H11          -0.009611               0
 2 LOG_REP_duration_H11          -0.022358               0
 3 LOG_REP_campaign_H12          -0.054447               0
 4 LOG_REP_duration_H12          -0.049699               0
 5 LOG_REP_campaign_H13           0.016443               0
 6 LOG_REP_duration_H13           0.111661               0
 7 REP_educationBasic_H11        -0.340148               0
 8 REP_educationHS_H11            0.177052               0
 9 REP_educationUn_Il_H11         0.127949               0
10 REP_jobB_T_H_H11               0.039290               0
```

…

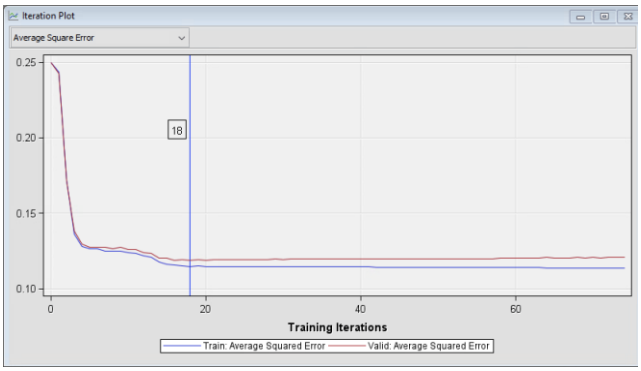
```
60 poutcomenonexistent_H13        0.317382               0
61 BIAS_H11                      -0.220168               0
62 BIAS_H12                       0.397161               0
63 BIAS_H13                       1.303395               0
64 H11_yyes                              0       -0.007759
65 H12_yyes                              0        0.040399
66 H13_yyes                              0        0.013491
67 BIAS_yyes                             0               0

Value of Objective Function = 0.6931471806
```

Picture 196. Optimization Results for Stepwise Regression Neural Network 100 iterations 3 hidden units

```
                                   Optimization Results

Iterations                                   74  Function Calls                          89
Jacobian Calls                               77  Active Constraints                       0
Objective Function                 0.3635469215  Max Abs Gradient Element       0.0002119915
Lambda                             0.6815779562  Actual Over Pred Change        0.2838455847
Radius                             0.0070341774

Convergence criterion (FCONV=0.0001) satisfied.
```

Convergence was observed after 74 iterations.

Picture 197. Iteration Plot (ASE) for Stepwise Regression Neural Network 100 iterations 3 hidden units



As seen in the graph, 18 iterations are enough after that it gives us the same results. The iteration plot shows optimal validation average squared error occurring on this iteration.

Stepwise Regression Neural Network 100 iterations, 4 hidden units

Picture 200. Results for Stepwise Regression Neural Network 100 iterations 4 hidden units

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Fr...	4638	.	.
y	y	_DFE_	Degrees of Freedo...	4549	.	.
y	y	_DFM_	Model Degrees of ...	89	.	.
y	y	_NW_	Number of Estimat...	89	.	.
y	y	_AIC_	Akaike's Informatio...	3575.591	.	.
y	y	_SBC_	Schwarz's Bayesia...	4148.933	.	.
y	y	_ASE_	Average Squared E...	0.114821	0.120201	.
y	y	_MAX_	Maximum Absolute...	0.997356	0.997316	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squa...	0.338853	0.3467	.
y	y	_SSE_	Sum of Squared Er...	1065.081	1115.948	.
y	y	_SUMW_	Sum of Case Weig...	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.119314	.	.
y	y	_MSE_	Mean Squared Error	0.117068	0.120201	.
y	y	_RFPE_	Root Final Predicti...	0.345419	.	.
y	y	_RMSE_	Root Mean Square...	0.342151	0.3467	.
y	y	_AVERR_	Average Error Func...	0.366278	0.376259	.
y	y	_ERR_	Error Function	3397.591	3493.192	.
y	y	_MISC_	Misclassification R...	0.158042	0.173632	.
y	y	_WRONG_	Number of Wrong ...	733	806	.

Picture 201. Optimization Start, Parameter Estimates for Stepwise Regression Neural Network 100 iterations 4 hidden units

```
The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                Gradient
                                               Objective
  N Parameter                     Estimate       Function

  1 LOG_REP_campaign_H11          -0.009611              0
  2 LOG_REP_duration_H11          -0.022358              0
  3 LOG_REP_campaign_H12          -0.054447              0
  4 LOG_REP_duration_H12          -0.049699              0
  5 LOG_REP_campaign_H13           0.016443              0
  6 LOG_REP_duration_H13           0.111661              0
  7 LOG_REP_campaign_H14          -0.340148              0
  8 LOG_REP_duration_H14           0.177052              0
  9 REP_educationBasic_H11         0.127949              0
 10 REP_educationHS_H11            0.039290              0
```

…

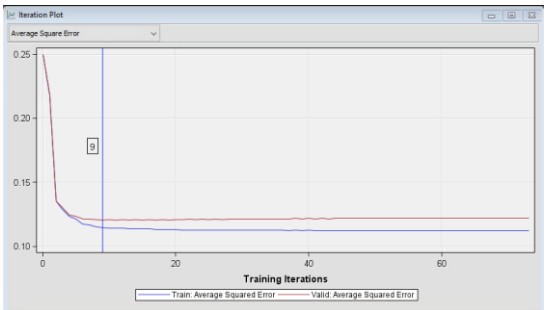
```
 80 poutcomenonexistent_H14        0.015195              0
 81 BIAS_H11                      -0.400147              0
 82 BIAS_H12                      -0.642126              0
 83 BIAS_H13                       0.585875              0
 84 BIAS_H14                      -0.050069              0
 85 H11_yyes                              0       0.025681
 86 H12_yyes                              0       0.005571
 87 H13_yyes                              0      -0.017007
 88 H14_yyes                              0      -0.066564
 89 BIAS_yyes                             0              0

Value of Objective Function = 0.6931471806
```

Picture 202. Optimization Results for Stepwise Regression Neural Network 100 iterations 4 hidden units

```
                                    Optimization Results

Iterations                                   73  Function Calls                    85
Jacobian Calls                               76  Active Constraints                 0
Objective Function                 0.3561439224  Max Abs Gradient Element  0.0001886796
Lambda                              1.303462794  Actual Over Pred Change    0.975904179
Radius                             0.0043864051

Convergence criterion (FCONV=0.0001) satisfied.
```

Convergence criterion satisfied after 73 iterations.

Picture 203. Iteration Plot (ASE) for Stepwise Regression Neural Network 100 iterations 4 hidden units



As seen in the graph, 9 iterations are enough after that it gives us the same results. The iteration plot shows optimal validation average squared error occurring on this iteration.

Stepwise Regression Neural Network 100 iterations, 5 hidden units

Picture 204. Results for Stepwise Regression Neural Network 100 iterations 5 hidden units

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Fr...	4638	.	.
y	y	_DFE_	Degrees of Freedo...	4527	.	.
y	y	_DFM_	Model Degrees of ...	111	.	.
y	y	_NW_	Number of Estimat...	111	.	.
y	y	_AIC_	Akaike's Informatio...	3461.485	.	.
y	y	_SBC_	Schwarz's Bayesia...	4176.552	.	.
y	y	_ASE_	Average Squared E...	0.109104	0.117893	.
y	y	_MAX_	Maximum Absolute...	0.997743	0.994864	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squa...	0.330309	0.343356	.
y	y	_SSE_	Sum of Squared Er...	1012.049	1094.52	.
y	y	_SUMW_	Sum of Case Weig...	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.114454	.	.
y	y	_MSE_	Mean Squared Error	0.111779	0.117893	.
y	y	_RFPE_	Root Final Predicti...	0.338311	.	.
y	y	_RMSE_	Root Mean Square...	0.334334	0.343356	.
y	y	_AVERR_	Average Error Func...	0.349233	0.372941	.
y	y	_ERR_	Error Function	3239.485	3462.382	.
y	y	_MISC_	Misclassification R...	0.155024	0.167385	.
y	y	_WRONG_	Number of Wrong ...	719	777	.

Picture 205. Optimization Start, Parameter Estimates for Stepwise Regression Neural Network 100 iterations 5 hidden units

```
The NEURAL Procedure

                    Optimization Start
                  Parameter Estimates
                                                  Gradient
                                                 Objective
  N Parameter                     Estimate        Function

  1 LOG_REP_campaign_H11          -0.009611               0
  2 LOG_REP_duration_H11          -0.022358               0
  3 LOG_REP_campaign_H12          -0.054447               0
  4 LOG_REP_duration_H12          -0.049699               0
  5 LOG_REP_campaign_H13           0.016443               0
  6 LOG_REP_duration_H13           0.111661               0
  7 LOG_REP_campaign_H14          -0.340148               0
  8 LOG_REP_duration_H14           0.177052               0
  9 LOG_REP_campaign_H15           0.127949               0
 10 LOG_REP_duration_H15           0.039290               0
```

…

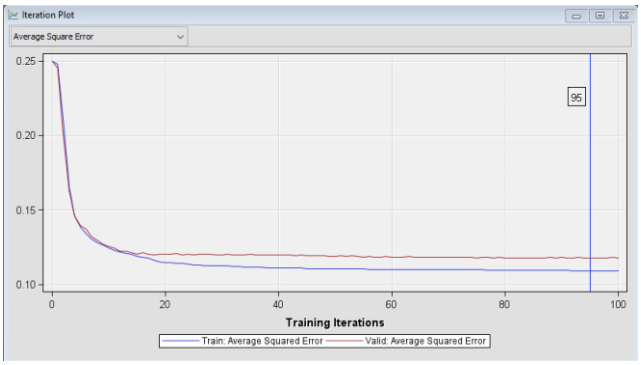
```
100 poutcomenonexistent_H15        0.323542               0
101 BIAS_H11                       0.505428               0
102 BIAS_H12                      -0.315904               0
103 BIAS_H13                      -0.594813               0
104 BIAS_H14                      -0.461791               0
105 BIAS_H15                       0.645601               0
106 H11_yyes                              0       -0.011265
107 H12_yyes                              0       -0.017772
108 H13_yyes                              0        0.006127
109 H14_yyes                              0       -0.033832
110 H15_yyes                              0        0.024929
111 BIAS_yyes                             0               0

Value of Objective Function = 0.6931471806
```

Picture 206. Optimization Results for Stepwise Regression Neural Network 100 iterations 5 hidden units

```
                                   Optimization Results

Iterations                                   100  Function Calls                          240
Gradient Calls                               113  Active Constraints                        0
Objective Function                 0.3486352097  Max Abs Gradient Element          0.002141942
Slope of Search Direction         -0.000182875

QUANEW needs more than 100 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Notice the warning message. The optimization algorithm (QUANEW) may be struggling to converge to a solution within the given number of iterations The maximum number of iterations, by default, is 100. Apparently, this is not enough for the network training process to converge.

Picture 207. Iteration Plot (ASE) for Stepwise Regression Neural Network 100 iterations 5 hidden units



As seen in the graph, 95 iterations are enough after that it gives us the same results. The iteration plot shows optimal validation average squared error occurring on this iteration.

Stepwise Regression Neural Network 100 iterations, 6 hidden units

Picture 208. Results for Stepwise Regression Neural Network 100 iterations 6 hidden units

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Fr...	4638	.	.
y	y	_DFE_	Degrees of Freedo...	4505	.	.
y	y	_DFM_	Model Degrees of ...	133	.	.
y	y	_NW_	Number of Estimat...	133	.	.
y	y	_AIC_	Akaike's Informatio...	3583.641	.	.
y	y	_SBC_	Schwarz's Bayesia...	4440.433	.	.
y	y	_ASE_	Average Squared E...	0.111796	0.118298	.
y	y	_MAX_	Maximum Absolute...	0.9943	0.998474	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squa...	0.334359	0.343945	.
y	y	_SSE_	Sum of Squared Er...	1037.02	1098.279	.
y	y	_SUMW_	Sum of Case Weig...	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.118397	.	.
y	y	_MSE_	Mean Squared Error	0.115097	0.118298	.
y	y	_RFPE_	Root Final Predicti...	0.344089	.	.
y	y	_RMSE_	Root Mean Square...	0.339259	0.343945	.
y	y	_AVERR_	Average Error Func...	0.357659	0.371199	.
y	y	_ERR_	Error Function	3317.641	3446.213	.
y	y	_MISC_	Misclassification R...	0.15718	0.170616	.
y	y	_WRONG_	Number of Wrong ...	729	792	.

Picture 209. Optimization Start, Parameter Estimates for Stepwise Regression Neural Network 100 iterations 6 hidden units

```
The NEURAL Procedure

                    Optimization Start
                   Parameter Estimates
                                                Gradient
                                               Objective
  N Parameter                     Estimate      Function

  1 LOG_REP_campaign_H11          -0.009611             0
  2 LOG_REP_duration_H11          -0.022358             0
  3 LOG_REP_campaign_H12          -0.054447             0
  4 LOG_REP_duration_H12          -0.049699             0
  5 LOG_REP_campaign_H13           0.016443             0
  6 LOG_REP_duration_H13           0.111661             0
  7 LOG_REP_campaign_H14          -0.340148             0
  8 LOG_REP_duration_H14           0.177052             0
  9 LOG_REP_campaign_H15           0.127949             0
 10 LOG_REP_duration_H15           0.039290             0
```

…

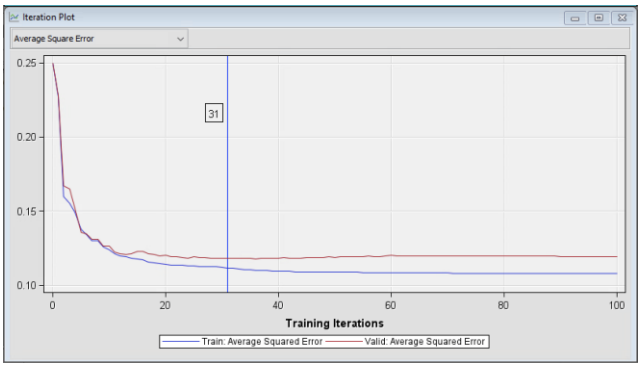
```
120 poutcomenonexistent_H16        0.074630             0
121 BIAS_H11                      -0.283225             0
122 BIAS_H12                       1.216321             0
123 BIAS_H13                       0.941225             0
124 BIAS_H14                       0.044303             0
125 BIAS_H15                       0.680919             0
126 BIAS_H16                       0.652068             0
127 H11_yyes                              0     -0.040848
128 H12_yyes                              0      0.019871
129 H13_yyes                              0     -0.022027
130 H14_yyes                              0     -0.030000
131 H15_yyes                              0     -0.010755
132 H16_yyes                              0     -0.009282
133 BIAS_yyes                             0             0

Value of Objective Function = 0.6931471806
```

Picture 210. Optimization Results for Stepwise Regression Neural Network 100 iterations 6 hidden units

```
                                        Optimization Results

Iterations                                     100  Function Calls                                249
Gradient Calls                                 113  Active Constraints                              0
Objective Function                    0.3452335961  Max Abs Gradient Element               0.0007384107
Slope of Search Direction             -0.000020476

QUANEW needs more than 100 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Notice the warning message. The optimization algorithm (QUANEW) may be struggling to converge to a solution within the given number of iterations The maximum number of iterations, by default, is 100. Apparently, this is not enough for the network training process to converge.

Picture 211. Iteration Plot (ASE) for Stepwise Regression Neural Network 100 iterations 6 hidden units



As seen in the graph, 31 iterations are enough after that it gives us the same results. The iteration plot shows optimal validation average squared error occurring on this iteration.

MODEL COMPARISON

The Model Comparison node is like a helper that helps us figure out which predictive model works the best. It checks how good different models, such as decision trees or neural networks, are at making predictions. This way, we can pick the one that does the job well for our specific needs. It's a handy tool for choosing the right model and making sure it performs well in what we want it to do.

We added a Model Comparison Node to our model and connected it to our ASE Decision Tree with all our 13 Branches, we connected it to Misclassification Decision Tree with all our 11 branches, we connected it with our Full Regression, Stepwise Regression, Forward and Backward Regression, we finally connected it with all Data Partition, Cap n Floor, Transform Variables, and Stepwise Regression Neural Networks with 50 Iterations with 3,4,5,6 Hidden Units and with 100 Iterations with 3,4,5,6 Hidden Units for optimal results. It helped us in comparing the performance of various models by considering metrics like misclassification rate, average squared error, ROC (Receiver Operating Characteristic) curves. The most optimal Model chosen by our Model Comparison Node was Data Partition Neural Network with 50 Iterations and 3 Hidden Units.

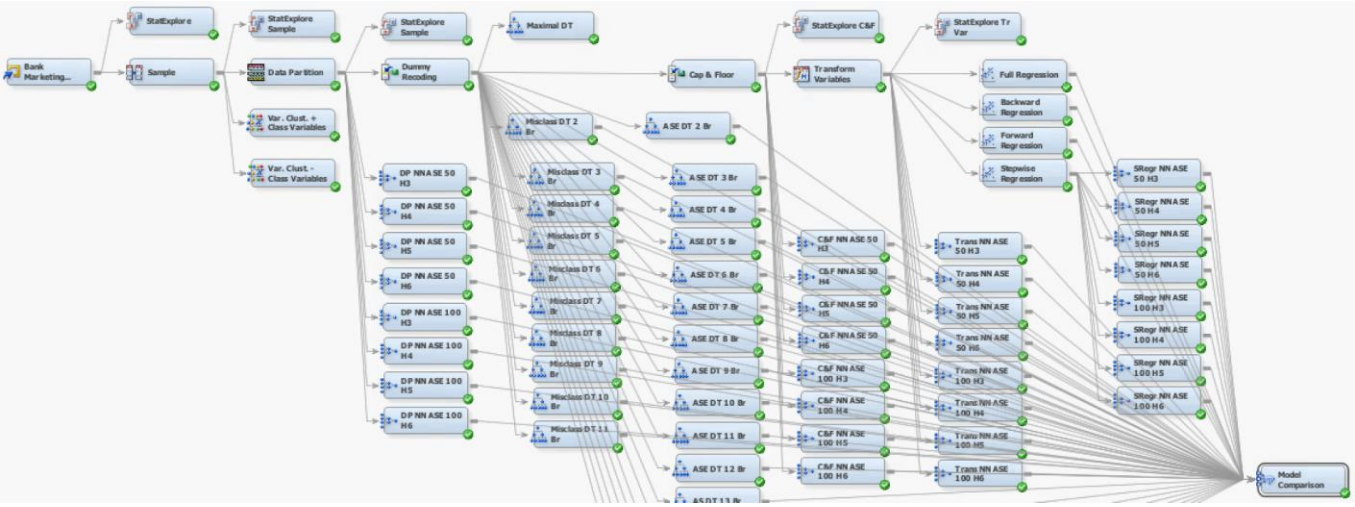
Picture 212. Model Comparison, settings

Model Selection	
Selection Data	Default
Selection Statistic	Average Squared Error
HP Selection Statistic	Default
SAS Viya Selection Statistic	...
Selection Table	Validation
Selection Depth	10

We added a Model Comparison into our Model and changed Selection Statistics to Average Squared Error and Selection Table to Validation in the Settings. It was connected to:

- Misclassification Decision Tree Nodes
- ASE Decision Tree Nodes
- Data Partition Neural Network Nodes
- Cap & Floor Neural Network Nodes
- Transform Variables Neural Network Nodes
- StepWise Regression Neural Network Nodes
- Full Regression Node
- Backward Regression Node
- Forward Regression Node
- Stepwise Regression Node

Picture 213. Adding Model Comparison Node into the Model

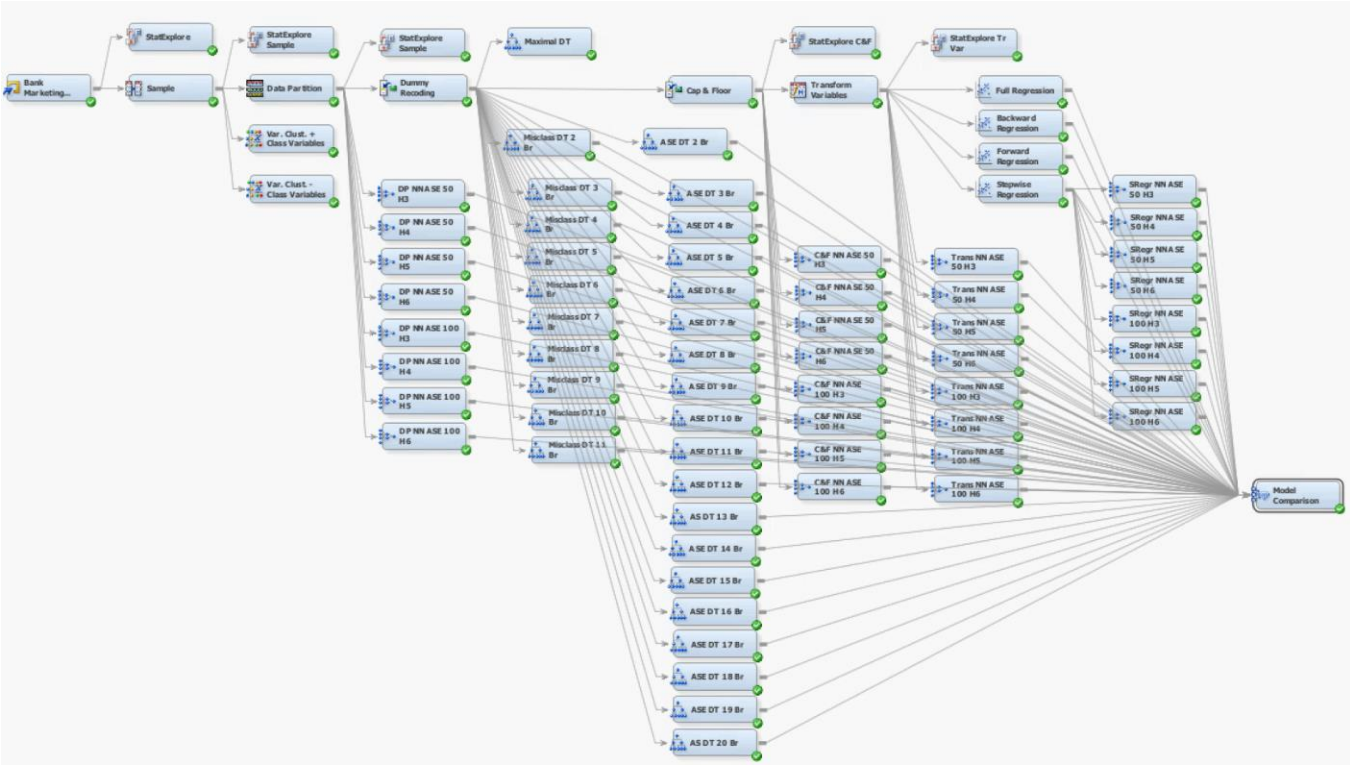


In our study, we explored different models to find the best one for predicting outcomes in our bank marketing analysis. We considered Decision Trees, Neural Networks with various settings, and Regression analyses. After careful evaluation, the Data Partition Neural Network with 50 Iterations and 3 Hidden Units stood out as the top performer.

This particular model achieved an impressive ROC Index of 0.922, showcasing its excellent ability to distinguish between different outcomes. With an Average Squared Error of 0.107239 and a misclassification rate of 0.147997, it demonstrated high accuracy and reliability.

While a similar performance was observed with the Data Partition Neural Network with 100 Iterations and 3 Hidden Units, the model comparison node opted for the 50 Iterations version. This decision aligns with the idea of choosing the simplest model that performs well. Our chosen configuration not only showed outstanding predictive capabilities but also represents an efficient solution for our bank marketing analysis.

Picture 214. The whole Model



Picture 215. Model Comparison, fit statistics

Results - Node: Model Comparison Diagram: Bank Marketing Analysis

File Edit View Window

Fit Statistics

Selected Model	Predecessor Node	Model Node	Model Description	Target Variable	Target Label	Selection Criterion: Valid: Average Squared Error ▲	Valid: Average Squared Error	Valid: Misclassification Rate	Valid: Roc Index	Valid: Lift	Valid: Cumulative Lift	Valid: Percent Response	Valid: Cumulative Percent Response	Valid: Percent Captured Response	Valid: Cumulative Percent Captured Response
Y	Neural	Neural	DP NN ASE 50 H3	y	y	0.107239	0.107239	0.147997	0.922	1.853448	1.896774	92.67241	94.83871	9.263249	19.00043
	Neural5	Neural5	DP NN ASE 100 H3	y	y	0.107239	0.107239	0.147997	0.922	1.853448	1.896774	92.67241	94.83871	9.263249	19.00043
	Neural3	Neural3	DP NN ASE 50 H5	y	y	0.109376	0.109376	0.150582	0.919	1.862069	1.87957	93.10345	93.97849	9.306333	18.82809
	Neural7	Neural7	DP NN ASE 100 H5	y	y	0.109376	0.109376	0.150582	0.919	1.862069	1.87957	93.10345	93.97849	9.306333	18.82809
	Neural2	Neural2	DP NN ASE 50 H4	y	y	0.111473	0.111473	0.155967	0.918	1.827586	1.883871	91.37931	94.19355	9.133994	18.87118
	Neural6	Neural6	DP NN ASE 100 H4	y	y	0.111473	0.111473	0.155967	0.918	1.827586	1.883871	91.37931	94.19355	9.133994	18.87118
	Neural4	Neural4	DP NN ASE 50 H6	y	y	0.112011	0.112011	0.153813	0.917	1.87069	1.913978	93.53448	95.69892	9.349418	19.17277
	Neural8	Neural8	DP NN ASE 100 H6	y	y	0.112011	0.112011	0.153813	0.917	1.87069	1.913978	93.53448	95.69892	9.349418	19.17277
	Neural12	Neural12	C&F NN ASE 50 H6	y	y	0.116799	0.116799	0.16006	0.913	1.965517	1.922581	98.27586	96.12903	9.823352	19.25894
	Neural16	Neural16	C&F NN ASE 100 H6	y	y	0.116799	0.116799	0.16006	0.913	1.965517	1.922581	98.27586	96.12903	9.823352	19.25894
	Neural14	Neural14	C&F NN ASE 100 H4	y	y	0.116841	0.116841	0.162215	0.913	1.853448	1.883871	92.67241	94.19355	9.263249	18.87118
	Neural15	Neural15	C&F NN ASE 100 H5	y	y	0.116917	0.116917	0.163292	0.913	1.896552	1.883871	94.82759	94.19355	9.478673	18.87118
	Neural17	Neural17	Trans NN ASE 50 H4	y	y	0.116953	0.116953	0.163723	0.911	1.896552	1.87957	94.82759	93.97849	9.478673	18.82809
	Neural21	Neural21	Trans NN ASE 100 H4	y	y	0.116953	0.116953	0.163723	0.911	1.896552	1.87957	94.82759	93.97849	9.478673	18.82809
	Neural10	Neural10	C&F NN ASE 50 H4	y	y	0.117521	0.117521	0.162861	0.912	1.862069	1.888172	93.10345	94.4086	9.306333	18.91426
	Neural30	Neural30	SRegr NN ASE 100 H5	y	y	0.117893	0.117893	0.167385	0.911	1.827586	1.883871	91.37931	94.19355	9.133994	18.87118
	Neural27	Neural27	SRegr NN ASE 50 H6	y	y	0.118298	0.118298	0.170616	0.913	1.905172	1.931183	95.25862	96.55914	9.521758	19.34511

Picture 215. Model Comparison, fit statistics (continue)

Results - Node: Model Comparison Diagram: Bank Marketing Analysis

File Edit View Window

Fit Statistics

Selected Model	Predecessor Node	Model Node	Model Description	Target Variable	Target Label	Selection Criterion: Valid: Average Squared Error ▲	Valid: Average Squared Error	Valid: Misclassification Rate	Valid: Roc Index	Valid: Lift	Valid: Cumulative Lift	Valid: Percent Response	Valid: Cumulative Percent Response	Valid: Percent Captured Response	Valid: Cumulative Percent Captured Response
	Neural31	Neural31	SRegr NN ASE 100 H6	y	y	0.118298	0.118298	0.170616	0.913	1.905172	1.931183	95.25862	96.55914	9.521758	19.34511
	Neural11	Neural11	C&F NN ASE 50 H5	y	y	0.118519	0.118519	0.167169	0.911	1.853448	1.888172	92.67241	94.4086	9.263249	18.91426
	Neural28	Neural28	SRegr NN ASE 100 H3	y	y	0.11865	0.11865	0.17277	0.912	1.956897	1.944086	97.84483	97.2043	9.780267	19.47436
	Neural32	Neural32	SRegr NN ASE 50 H3	y	y	0.11865	0.11865	0.17277	0.912	1.956897	1.944086	97.84483	97.2043	9.780267	19.47436
	Neural26	Neural26	SRegr NN ASE 50 H5	y	y	0.1188	0.1188	0.171909	0.91	1.827586	1.858065	91.37931	92.90323	9.133994	18.61267
	Neural19	Neural19	Trans NN ASE 50 H6	y	y	0.119281	0.119281	0.16997	0.908	1.896552	1.875269	94.82759	93.76344	9.478673	18.78501
	Neural23	Neural23	Trans NN ASE 100 H6	y	y	0.119281	0.119281	0.16997	0.908	1.896552	1.875269	94.82759	93.76344	9.478673	18.78501
	Neural18	Neural18	Trans NN ASE 50 H5	y	y	0.119392	0.119392	0.168677	0.91	1.913793	1.944086	95.68966	97.2043	9.564843	19.47436
	Neural22	Neural22	Trans NN ASE 100 H5	y	y	0.119392	0.119392	0.168677	0.91	1.913793	1.944086	95.68966	97.2043	9.564843	19.47436
	Neural20	Neural20	Trans NN ASE 100 H3	y	y	0.119533	0.119533	0.17234	0.91	1.87069	1.901075	93.53448	95.05376	9.349418	19.04352
	Neural24	Neural24	Trans NN ASE 50 H3	y	y	0.119533	0.119533	0.17234	0.91	1.87069	1.901075	93.53448	95.05376	9.349418	19.04352
	Neural25	Neural25	SRegr NN ASE 50 H4	y	y	0.120201	0.120201	0.173632	0.91	1.87931	1.913978	93.96552	95.69892	9.392503	19.17277
	Neural29	Neural29	SRegr NN ASE 100 H4	y	y	0.120201	0.120201	0.173632	0.91	1.87931	1.913978	93.96552	95.69892	9.392503	19.17277
	Neural13	Neural13	C&F NN ASE 100 H3	y	y	0.120566	0.120566	0.171909	0.908	1.827586	1.836559	91.37931	91.82796	9.133994	18.39724
	Neural9	Neural9	C&F NN ASE 50 H3	y	y	0.120566	0.120566	0.171909	0.908	1.827586	1.836559	91.37931	91.82796	9.133994	18.39724
	Tree16	Tree16	ASE DT 12 Br	y	y	0.125615	0.125615	0.177941	0.902	1.886731	1.91828	94.33653	95.91398	9.429589	19.21586
	Tree26	Tree26	ASE DT 19 Br	y	y	0.125615	0.125615	0.177941	0.902	1.886731	1.91828	94.33653	95.91398	9.429589	19.21586
	Reg	Reg	Full Regression	y	y	0.125833	0.125833	0.182034	0.902	1.853448	1.892473	92.67241	94.62366	9.263249	18.95735
	Reg2	Reg2	Backward Regression	y	y	0.125881	0.125881	0.181172	0.902	1.862069	1.896774	93.10345	94.83871	9.306333	19.00043
	Tree19	Tree19	ASE DT 4 Br	y	y	0.12626	0.12626	0.179664	0.895	1.764003	1.852639	88.20015	92.63196	8.816215	18.55832
	Reg3	Reg3	Forward Regression	y	y	0.126304	0.126304	0.182464	0.902	1.853448	1.901075	92.67241	95.05376	9.263249	19.04352
	Reg4	Reg4	Stepwise Regression	y	y	0.126304	0.126304	0.182464	0.902	1.853448	1.901075	92.67241	95.05376	9.263249	19.04352
	Tree15	Tree15	ASE DT 11 Br	y	y	0.127342	0.127342	0.182034	0.897	1.841139	1.894175	92.05693	94.70873	9.201727	18.97439
	Tree25	Tree25	ASE DT 18 Br	y	y	0.127342	0.127342	0.182034	0.897	1.841139	1.894175	92.05693	94.70873	9.201727	18.97439
	Tree17	Tree17	AS DT 13 Br	y	y	0.128195	0.128195	0.184619	0.898	1.876966	1.921193	93.8483	96.05966	9.380786	19.24504
	Tree27	Tree27	AS DT 20 Br	y	y	0.128195	0.128195	0.184619	0.898	1.876966	1.921193	93.8483	96.05966	9.380786	19.24504
	Tree20	Tree20	ASE DT 5 Br	y	y	0.128365	0.128365	0.188927	0.894	1.849781	1.895167	92.48903	94.75836	9.244919	18.98433
	Tree21	Tree21	ASE DT 6 Br	y	y	0.128365	0.128365	0.188927	0.894	1.849781	1.895167	92.48903	94.75836	9.244919	18.98433
	Tree22	Tree22	ASE DT 7 Br	y	y	0.128365	0.128365	0.188927	0.894	1.849781	1.895167	92.48903	94.75836	9.244919	18.98433
	Tree23	Tree23	ASE DT 8 Br	y	y	0.128365	0.128365	0.188927	0.894	1.849781	1.895167	92.48903	94.75836	9.244919	18.98433
	Tree24	Tree24	ASE DT 9 Br	y	y	0.128365	0.128365	0.188927	0.894	1.849781	1.895167	92.48903	94.75836	9.244919	18.98433
	Tree28	Tree28	ASE DT 14 Br	y	y	0.128365	0.128365	0.188927	0.894	1.849781	1.895167	92.48903	94.75836	9.244919	18.98433
	Tree29	Tree29	ASE DT 15 Br	y	y	0.128365	0.128365	0.188927	0.894	1.849781	1.895167	92.48903	94.75836	9.244919	18.98433
	Tree30	Tree30	ASE DT 16 Br	y	y	0.128365	0.128365	0.188927	0.894	1.849781	1.895167	92.48903	94.75836	9.244919	18.98433
	Tree4	Tree4	Misclass DT 4 Br	y	y	0.128879	0.128879	0.178156	0.892	1.837586	1.88886	91.87931	94.44301	9.183972	18.92115
	Tree18	Tree18	ASE DT 3 Br	y	y	0.130065	0.130065	0.177941	0.894	1.892916	1.913716	94.64581	95.68579	9.460503	19.17014
	Tree12	Tree12	ASE DT 2 Br	y	y	0.132279	0.132279	0.178587	0.888	1.849489	1.881179	92.47445	94.05893	9.24346	18.84421
	Tree13	Tree13	ASE DT 17 Br	y	y	0.132279	0.132279	0.178587	0.888	1.849489	1.881179	92.47445	94.05893	9.24346	18.84421
	Tree14	Tree14	ASE DT 10 Br	y	y	0.132279	0.132279	0.178587	0.888	1.849489	1.881179	92.47445	94.05893	9.24346	18.84421
	Tree3	Tree3	Misclass DT 3 Br	y	y	0.133616	0.133616	0.177294	0.881	1.849432	1.89202	92.47159	94.60102	9.243175	18.95281
	Tree5	Tree5	Misclass DT 5 Br	y	y	0.13384	0.13384	0.185696	0.883	1.869438	1.904506	93.47188	95.22529	9.343161	19.07788
	Tree6	Tree6	Misclass DT 6 Br	y	y	0.13384	0.13384	0.185696	0.883	1.869438	1.904506	93.47188	95.22529	9.343161	19.07788
	Tree7	Tree7	Misclass DT 7 Br	y	y	0.13384	0.13384	0.185696	0.883	1.869438	1.904506	93.47188	95.22529	9.343161	19.07788
	Tree8	Tree8	Misclass DT 8 Br	y	y	0.13384	0.13384	0.185696	0.883	1.869438	1.904506	93.47188	95.22529	9.343161	19.07788
	Tree9	Tree9	Misclass DT 9 Br	y	y	0.13384	0.13384	0.185696	0.883	1.869438	1.904506	93.47188	95.22529	9.343161	19.07788
	Tree11	Tree11	Misclass DT 11 Br	y	y	0.134435	0.134435	0.181172	0.883	1.851876	1.898773	92.59378	94.93863	9.255389	19.02045
	Tree10	Tree10	Misclass DT 10 Br	y	y	0.13973	0.13973	0.177725	0.857	1.738095	1.819355	86.90476	90.96774	8.686732	18.2249
	Tree2	Tree2	Misclass DT 2 Br	y	y	0.13973	0.13973	0.177725	0.857	1.738095	1.819355	86.90476	90.96774	8.686732	18.2249

Picture 7. Model Comparison, fit statistics based on ASE

Fit Statistics

Model Selection based on Valid: Average Squared Error (_VASE_)

Selected Model	Model Node	Model Description	Valid: Average Squared Error	Train: Average Squared Error	Train: Misclassification Rate	Valid: Misclassification Rate
Y	Neural	DP NN ASE 50 H3	0.10724	0.10113	0.14015	0.14800
	Neural5	DP NN ASE 100 H3	0.10724	0.10113	0.14015	0.14800
	Neural3	DP NN ASE 50 H5	0.10938	0.09587	0.13562	0.15058
	Neural7	DP NN ASE 100 H5	0.10938	0.09587	0.13562	0.15058
	Neural2	DP NN ASE 50 H4	0.11147	0.10664	0.14597	0.15597
	Neural6	DP NN ASE 100 H4	0.11147	0.10664	0.14597	0.15597
	Neural4	DP NN ASE 50 H6	0.11201	0.10346	0.14511	0.15381
	Neural8	DP NN ASE 100 H6	0.11201	0.10346	0.14511	0.15381

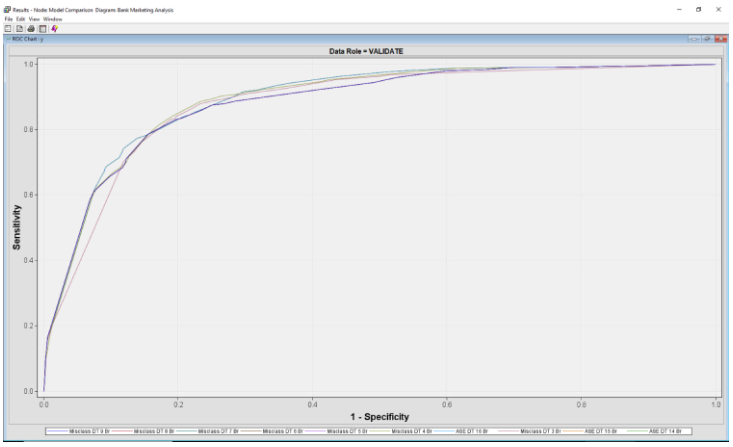
The results for Decision Partition Neural Network Average Squared Error with 50 Iterations and 3 Hidden Units and Decision Partition Neural Network Average Squared Error with 100 Iterations and 3 Hidden Units are the same, but the model preferred to choose the one with 50 Iterations and 3 Hidden Units with an ASE of 0.10724 for our validation data.

Picture 7. The main statistical features of chosen final Model

Data Role=Valid

Statistics	Neural	Neural5	Neural3	Neural7	Neural2
Valid: Kolmogorov-Smirnov Statistic	0.71	0.71	0.71	0.71	0.69
Valid: Average Squared Error	0.11	0.11	0.11	0.11	0.11
Valid: Roc Index	0.92	0.92	0.92	0.92	0.92
Valid: Average Error Function	0.35	0.35	0.36	0.36	0.36
Valid: Bin-Based Two-Way Kolmogorov-Smirnov Probability Cutoff	0.52	0.52	0.52	0.52	0.60
Valid: Cumulative Percent Captured Response	19.00	19.00	18.83	18.83	18.87
Valid: Percent Captured Response	9.26	9.26	9.31	9.31	9.13
Valid: Divisor for VASE	9284.00	9284.00	9284.00	9284.00	9284.00
Valid: Error Function	3216.65	3216.65	3321.88	3321.88	3338.30
Valid: Gain	89.68	89.68	87.96	87.96	88.39
Valid: Gini Coefficient	0.84	0.84	0.84	0.84	0.84
Valid: Bin-Based Two-Way Kolmogorov-Smirnov Statistic	0.70	0.70	0.70	0.70	0.69
Valid: Kolmogorov-Smirnov Probability Cutoff	0.53	0.53	0.46	0.46	0.46
Valid: Cumulative Lift	1.90	1.90	1.88	1.88	1.88
Valid: Lift	1.85	1.85	1.86	1.86	1.83
Valid: Maximum Absolute Error	1.00	1.00	1.00	1.00	1.00
Valid: Misclassification Rate	0.15	0.15	0.15	0.15	0.16
Valid: Mean Square Error	0.11	0.11	0.11	0.11	0.11
Valid: Sum of Frequencies	4642.00	4642.00	4642.00	4642.00	4642.00
Valid: Root Average Squared Error	0.33	0.33	0.33	0.33	0.33
Valid: Cumulative Percent Response	94.84	94.84	93.98	93.98	94.19
Valid: Percent Response	92.67	92.67	93.10	93.10	91.38
Valid: Root Mean Square Error	0.33	0.33	0.33	0.33	0.33
Valid: Sum of Square Errors	995.60	995.60	1015.45	1015.45	1034.91
Valid: Sum of Case Weights Times Freq	9284.00	9284.00	9284.00	9284.00	9284.00
Valid: Number of Wrong Classifications	687.00	687.00	699.00	699.00	724.00

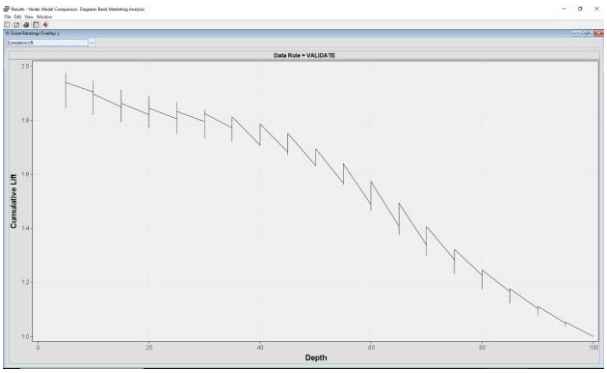
Picture 216. Model Comparison, ROC chart



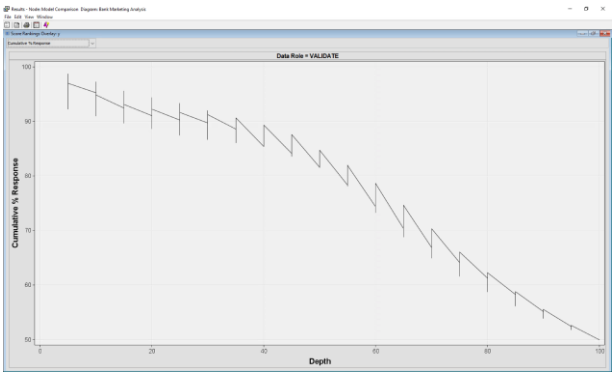
The ROC chart represents the union of similar calculations for all selection fractions. Models that capture mostly primary outcome cases without capturing secondary outcome cases are strong models.

The results for most of the presented models are the same and it represents strong models.

Picture 217. Model Comparison, Cumulative Lift



Picture 218. Model Comparison, Cumulative % Response



Our Model was too large to be interpreted in this graph and hence we can see this graph in this way.

Data Partition Neural Network Results, 50 iterations, 3 hidden units

Picture 219. Data Partition Neural Network, 50 iterations, 3 HU, Fit Statistics

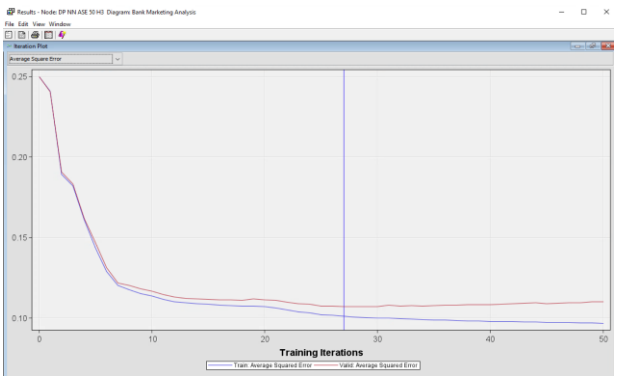
Results - Node: DP NN ASE 50 H3 Diagram: Bank Marketing Analysis

File Edit View Window

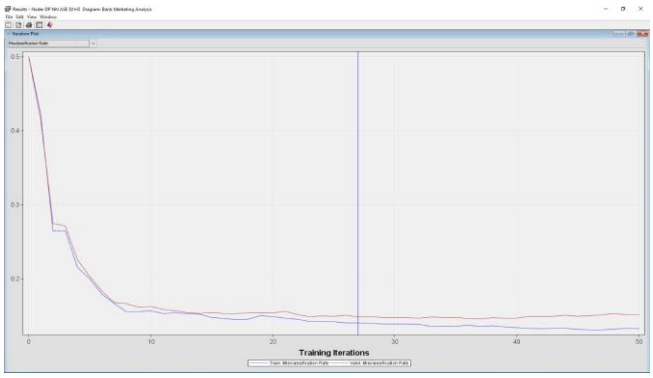
Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_DFT_	Total Degrees of Freedom	4638	.	.
y	y	_DFE_	Degrees of Freedom for Error	4427	.	.
y	y	_DFM_	Model Degrees of Freedom	211	.	.
y	y	_NW_	Number of Estimated Weights	211	.	.
y	y	_AIC_	Akaike's Information Criterion	3449.757	.	.
y	y	_SBC_	Schwarz's Bayesian Criterion	4809.027	.	.
y	y	_ASE_	Average Squared Error	0.101133	0.107239	.
y	y	_MAX_	Maximum Absolute Error	0.99712	0.998167	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_RASE_	Root Average Squared Error	0.318015	0.327473	.
y	y	_SSE_	Sum of Squared Errors	938.1131	995.6023	.
y	y	_SUMW_	Sum of Case Weights Times ...	9276	9284	.
y	y	_FPE_	Final Prediction Error	0.110774	.	.
y	y	_MSE_	Mean Squared Error	0.105954	0.107239	.
y	y	_RFPE_	Root Final Prediction Error	0.332827	.	.
y	y	_RMSE_	Root Mean Squared Error	0.325505	0.327473	.
y	y	_AVERR_	Average Error Function	0.326408	0.346473	.
y	y	_ERR_	Error Function	3027.757	3216.653	.
y	y	_MISC_	Misclassification Rate	0.140147	0.147997	.
y	y	_WRONG_	Number of Wrong Classificati...	650	687	.

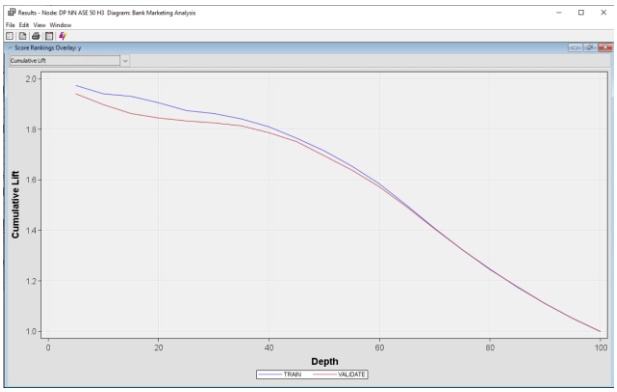
Picture 220. Data Partition Neural Network, 50 iterations, 3 HU, ASE Plot



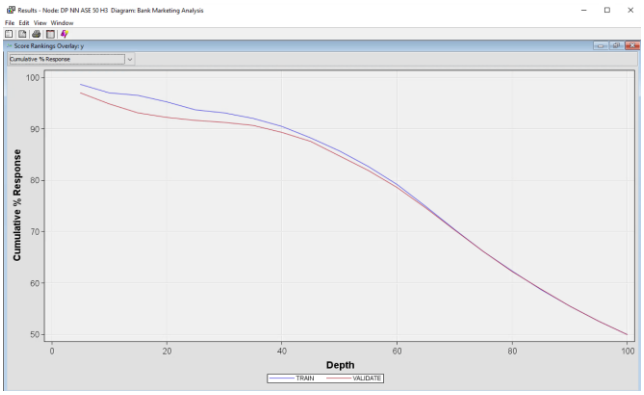
Picture 221. Data Partition Neural Network, 50 iterations, 3 HU, Misclassification Rate



Picture 222. Data Partition Neural Network, 50 iterations, 3 HU, Cumulative Lift



Picture 223. Data Partition Neural Network, 50 iterations, 3 HU, Cumulative % Response



Picture 224. Data Partition Neural Network, 50 iterations, 3 HU, Optimization Start

```
The NEURAL Procedure

                        Optimization Start
                        Parameter Estimates
                                                    Gradient
                                                   Objective
 N Parameter                             Estimate   Function

 1 age_H11                              -0.005212          0
 2 campaign_H11                         -0.012125          0
 3 duration_H11                         -0.029528          0
 4 age_H12                              -0.026953          0
 5 campaign_H12                          0.008917          0
 6 duration_H12                          0.060557          0
 7 age_H13                              -0.184471          0
 8 campaign_H13                          0.096020          0
 9 duration_H13                          0.069390          0
10 contactcellular_H11                   0.021308          0
```

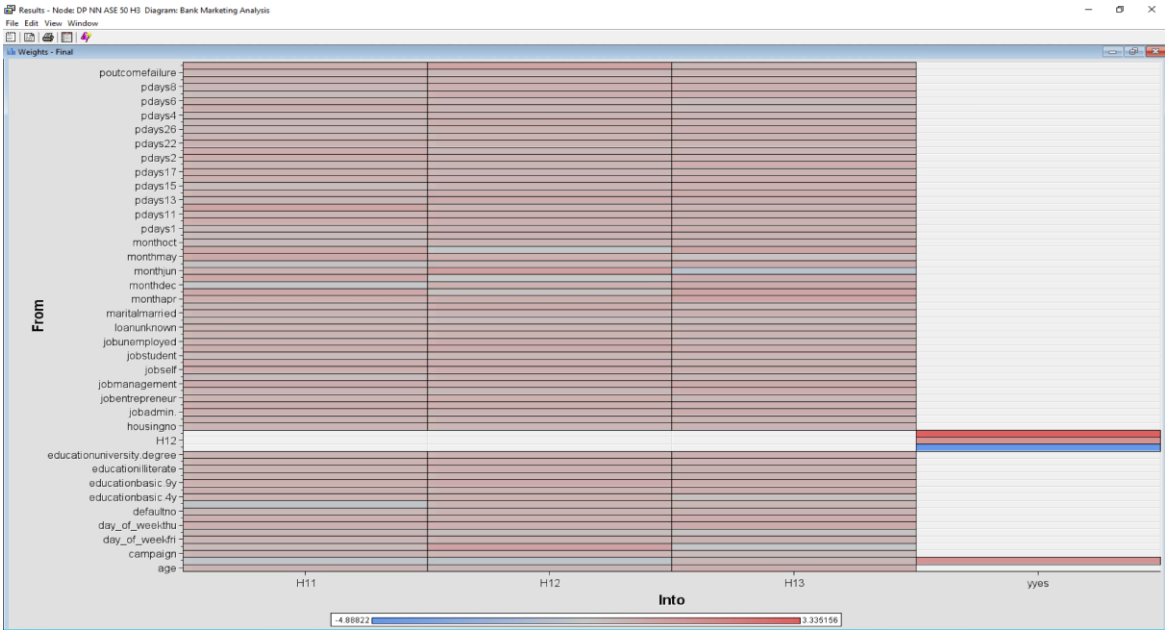
...

```
200 pdays22_H13                          0.016455          0
201 pdays25_H13                         -0.031278          0
202 pdays26_H13                          0.013142          0
203 poutcomefailure_H13                 -0.114525          0
204 poutcomenonexistent_H13             -0.063467          0
205 BIAS_H11                            -0.794813          0
206 BIAS_H12                            -0.794908          0
207 BIAS_H13                             0.201300          0
208 H11_yyes                                    0   0.037241
209 H12_yyes                                    0  -0.017805
210 H13_yyes                                    0  -0.034998
211 BIAS_yyes                                   0          0

Value of Objective Function = 0.6931471806
```

Picture 225. Data Partition Neural Network, 50 iterations, 3 HU, Optimization Results

```
                                    Optimization Results

Iterations                                   50  Function Calls                          126
Gradient Calls                               56  Active Constraints                        0
Objective Function                 0.3131501632  Max Abs Gradient Element       0.0036785644
Slope of Search Direction         -0.001064499

QUANEW needs more than 50 iterations or 2147483647 function calls.

WARNING: QUANEW Optimization cannot be completed.
```

Picture 226. Data Partition Neural Network, 50 iterations, 3 HU, Weights - Final (plot)



Picture 227. Data Partition Neural Network, 50 iterations, 3 HU, the most significant Weights - Final (table)

Table: Weights - Final

Label	From	Into	Weight ▼
H13 -> yyes	H13	yyes	3.335156
H12 -> yyes	H12	yyes	1.344029
BIAS -> yyes	BIAS	yyes	1.284198
monthjun -> H12	monthjun	H12	0.736374
contactcellular -> H12	contactcellular	H12	0.628391
poutcomenonexistent -> H12	poutcomenonexistent	H12	0.591432
monthapr -> H13	monthapr	H13	0.545164
pdays12 -> H11	pdays12	H11	0.53996
monthaug -> H13	monthaug	H13	0.492789
monthnov -> H13	monthnov	H13	0.474442
monthmay -> H11	monthmay	H11	0.467981
monthjul -> H11	monthjul	H11	0.419719
monthaug -> H11	monthaug	H11	0.372776
day_of_weekthu -> H13	day_of_weekthu	H13	0.33663
educationbasic.9y -> H12	educationbasic.9y	H12	0.328647
monthnov -> H11	monthnov	H11	0.304783
monthdec -> H13	monthdec	H13	0.292613
pdays13 -> H12	pdays13	H12	0.277751
jobadmin. -> H13	jobadmin.	H13	0.261596
jobtechnician -> H12	jobtechnician	H12	0.25788
jobmanagement -> H13	jobmanagement	H13	0.257631
monthmar -> H13	monthmar	H13	0.246173
pdays14 -> H13	pdays14	H13	0.241895
jobmanagement -> H11	jobmanagement	H11	0.232757
monthjul -> H13	monthjul	H13	0.227243
pdays10 -> H11	pdays10	H11	0.225541
pdays7 -> H12	pdays7	H12	0.225191
pdays5 -> H11	pdays5	H11	0.218971
jobunemployed -> H12	jobunemployed	H12	0.215323
pdays18 -> H13	pdays18	H13	0.2046
day_of_weektue -> H13	day_of_weektue	H13	0.203432
pdays21 -> H11	pdays21	H11	0.202412
jobhousemaid -> H13	jobhousemaid	H13	0.20217
jobself-employed -> H12	jobself	H12	0.200029

Picture 228. Data Partition Neural Network, 50 iterations, 3 HU, top 45 clients to contact with highest probability

EMWS7.Neural_VALIDATE

	Observation Number	age	job	marital	education	Predicted: y=yes	default	housing	loan	contact	month	day_of_week	duration	campaign	pdays	previous	poutcome	euribor3m	y	emp_var_rate	cons_price_idx	cons_conf_idx	nr_employed
1	2128.0	26	admin.	single	university.degree	0.9998146035962 95	no	yes	no	telephone	sep	thu	471	1	7	1	success	0.809	yes	-3.4	92.379	-29.8	5018.0
2	1855.0	37	technician	married	university.degree	0.9997783292751363	no	yes	no	telephone	sep	tue	773	1	3	2	success	0.755	yes	-3.4	92.379	-29.8	5018.0
3	43.0	33	admin.	single	university.degree	0.9993674680459284	no	yes	no	cellular	apr	thu	280	1	7	2	success	0.635	yes	-1.8	93.749	-34.6	5009.0
4	1222.0	48	admin.	married	high.school	0.999233115691333	unknown	no	yes	cellular	dec	thu	274	1	8	1	success	0.72	yes	-3.0	92.713	-33.0	5024.0
5	2310.0	83	retired	married	university.degree	0.9991700364139884	no	no	no	telephone	jul	tue	617	1	12	1	success	0.835	yes	-1.7	94.215	-40.3	4992.0
6	323.0	34	admin.	single	university.degree	0.9991073672798859	no	no	no	telephone	mar	tue	225	1	3	1	success	0.652	yes	-1.8	93.369	-34.8	5009.0
7	279.0	23	student	single	high.school	0.9989555946879942	no	no	no	telephone	mar	fri	290	1	3	1	success	0.649	yes	-1.8	93.369	-34.8	5009.0
8	984.0	53	housemaid	married	basic.4y	0.9988990803366825	no	no	yes	telephone	nov	wed	513	1	3	1	success	0.715	yes	-3.4	92.649	-30.1	5018.0
9	13061.0	34	blue-collar	married	basic.9y	0.9988483216395472	unknown	yes	no	telephone	apr	mon	1353	2	5	1	success	1.466	yes	-1.8	93.075	-47.1	5099.0
10	10793.0	60	admin.	married	high.school	0.9987620710487681	no	no	no	cellular	apr	thu	447	1	2	1	success	1.365	yes	-1.8	93.075	-47.1	5099.0
11	904.0	55	admin.	married	professional.course	0.9987390374861974	no	yes	no	cellular	dec	thu	798	1	9	2	success	0.714	yes	-3.0	92.713	-33.0	5024.0
12	4119.0	60	admin.	married	unknown	0.9987121618087009	no	no	no	cellular	nov	tue	787	1	7	2	success	1.046	yes	-1.1	94.767	-50.8	4964.0
13	3633.0	61	retired	married	university.degree	0.9986995341929 09	no	no	no	telephone	aug	wed	250	2	7	1	success	0.9	yes	-1.7	94.027	-38.3	4992.0
14	338.0	38	blue-collar	single	high.school	0.9986813264895508	no	no	no	telephone	mar	tue	450	3	6	2	success	0.652	yes	-1.8	93.369	-34.8	5009.0
15	10781.0	53	admin.	single	university.degree	0.9986130984134689	no	no	no	cellular	apr	thu	494	1	2	1	success	1.365	yes	-1.8	93.075	-47.1	5099.0
16	3703.0	26	admin.	single	university.degree	0.9985833595656134	no	no	no	telephone	aug	mon	1087	1	3	1	success	0.904	yes	-1.7	94.027	-38.3	4992.0
17	10942.0	28	admin.	single	university.degree	0.9985798053308366	no	no	yes	telephone	apr	thu	917	1	999	0	nonexistent	1.365	yes	-1.8	93.075	-47.1	5099.0
18	2915.0	45	unemployed	married	professional.course	0.9983558266508629	unknown	no	no	telephone	sep	thu	1405	1	6	2	failure	0.879	yes	-1.1	94.199	-37.5	4964.0
19	2797.0	30	technician	single	professional.course	0.9982841501404139	no	unknown	unkn...	telephone	sep	tue	808	1	12	2	failure	0.877	yes	-1.1	94.199	-37.5	4964.0
20	4005.0	62	retired	married	university.degree	0.9982528065805953	no	yes	no	cellular	nov	thu	208	1	1	6	success	1.031	yes	-1.1	94.767	-50.8	4964.0
21	11027.0	34	technician	married	university.degree	0.9981673427601755	no	yes	no	cellular	apr	thu	193	2	8	1	success	1.365	no	-1.8	93.075	-47.1	5099.0
22	2505.0	61	retired	married	university.degree	0.9980168922958731	no	no	no	telephone	aug	tue	249	2	3	1	success	0.859	yes	-2.9	92.201	-31.4	5076.0
23	3265.0	42	admin.	married	university.degree	0.9979918042762179	no	yes	no	cellular	aug	tue	576	1	4	1	success	0.884	yes	-2.9	92.201	-31.4	5076.0
24	4160.0	18	student	single	basic.4y	0.9979771772023125	no	yes	no	telephone	nov	tue	394	1	13	2	success	1.049	yes	-1.1	94.767	-50.8	4964.0
25	420.0	37	technician	married	professional.course	0.9979445451622337	no	no	no	cellular	apr	thu	277	1	18	1	success	0.659	yes	-1.8	93.749	-34.6	5009.0
26	10964.0	40	technician	married	professional.course	0.9979053517129436	no	yes	no	cellular	apr	thu	580	2	2	1	success	1.365	yes	-1.8	93.075	-47.1	5099.0
27	12359.0	32	blue-collar	married	basic.9y	0.9978318025327674	no	no	no	cellular	apr	thu	583	1	5	1	success	1.41	yes	-1.8	93.075	-47.1	5099.0
28	2617.0	42	management	married	university.degree	0.9977607375535942	no	yes	no	cellular	aug	fri	436	1	13	1	success	0.869	yes	-2.9	92.201	-31.4	5076.0
29	10915.0	56	entrepreneur	married	university.degree	0.9977385391644285	no	no	no	cellular	apr	thu	457	1	2	1	success	1.365	yes	-1.8	93.075	-47.1	5099.0
30	344.0	28	unemployed	married	high.school	0.9975880504052944	no	no	no	telephone	mar	fri	341	4	6	1	success	0.653	yes	-1.8	93.369	-34.8	5009.0
31	1028.0	59	retired	married	professional.course	0.9975538296518304	no	no	no	telephone	dec	tue	476	1	999	1	failure	0.715	yes	-3.0	92.713	-33.0	5024.0
32	11040.0	60	unemployed	married	high.school	0.997115623712492	no	yes	no	cellular	apr	thu	1260	2	2	1	success	1.365	yes	-1.8	93.075	-47.1	5099.0
33	416.0	38	admin.	single	high.school	0.9971083641381625	no	yes	no	cellular	apr	thu	882	1	6	1	success	0.659	yes	-1.8	93.749	-34.6	5009.0
34	192.0	50	technician	single	university.degree	0.9969721379481112	no	yes	no	telephone	apr	fri	634	1	999	0	nonexistent	0.644	yes	-1.8	93.749	-34.6	5009.0
35	130.0	45	management	married	unknown	0.9969667591499325	no	yes	no	cellular	apr	tue	379	3	7	1	success	0.642	yes	-1.8	93.749	-34.6	5009.0
36	140.0	53	blue-collar	single	basic.9y	0.9969473848291792	no	unknown	unkn...	cellular	mar	thu	362	1	9	2	success	0.643	yes	-1.8	93.369	-34.8	5009.0
37	193.0	58	unemployed	married	basic.9y	0.9969036430755241	no	yes	no	cellular	apr	thu	229	1	6	1	success	0.644	yes	-1.8	93.749	-34.6	5009.0
38	2848.0	49	admin.	married	high.school	0.9968372987686809	no	yes	no	cellular	aug	wed	668	2	15	1	success	0.879	yes	-2.9	92.201	-31.4	5076.0
39	1157.0	44	admin.	divorced	high.school	0.9967692974293261	no	yes	no	telephone	nov	mon	1628	2	6	1	success	0.719	yes	-3.4	92.649	-30.1	5018.0
40	3997.0	32	admin.	married	university.degree	0.9966192361913417	no	no	no	telephone	nov	wed	651	1	999	1	failure	1.03	yes	-1.1	94.767	-50.8	4964.0
41	1008.0	86	retired	married	professional.course	0.9965251512779091	no	no	no	telephone	dec	wed	343	2	999	0	nonexistent	0.715	yes	-3.0	92.713	-33.0	5024.0
42	13168.0	28	blue-collar	single	high.school	0.9965242574805433	no	yes	no	telephone	apr	thu	771	1	999	0	nonexistent	1.483	yes	-1.8	93.075	-47.1	5099.0
43	1057.0	54	admin.	married	high.school	0.9964424673987246	no	yes	no	cellular	nov	fri	402	1	3	3	success	0.716	yes	-3.4	92.649	-30.1	5018.0
44	3618.0	76	retired	married	university.degree	0.9964053984127083	no	yes	no	cellular	aug	thu	504	2	6	3	success	0.899	yes	-1.7	94.027	-38.3	4992.0
45	2200.0	38	technician	married	professional.course	0.9963938162905009	no	unknown	unkn...	telephone	aug	mon	473	2	9	1	success	0.821	yes	-2.9	92.201	-31.4	5076.0

In summary, our study has pinpointed the Data Partition Neural Network with 50 Iterations and 3 Hidden Units as the standout model. This model is quite exciting as it generates a list of top clients, each with a remarkable probability of over 99% to subscribe to the bank's term deposits.

The detailed client information, covering age, job, marital status, education, and other important factors, gives us a full picture of what influences their decisions. This powerful model allows the bank to prioritize and reach out to clients most likely to deposit subscriptions.

Our research not only highlights the model's predictive success but also arms the bank with practical insights for smarter marketing decisions. The identification of these highly probable top clients emphasizes the real-world impact of our findings, offering the bank a more targeted and effective approach to engaging potential subscribers.

CONCLUSION

KEY FEATURES OF THE RESEARCH

The Maximal Tree

Picture 229. The Maximal Tree, Variable Importance

Results - Node: Maximal DT Diagram: Bank Marketing Analysis

File Edit View Window

Variable Importance

Variable Name	Label	Number of Splitting Rules	Importance	Validation Importance	Ratio of Validation to Training Importance
duration	duration	10	1.0000	1.0000	1.0000
REP_pdays	Replacement: pdays	3	0.5646	0.6020	1.0662
contact	contact	2	0.3137	0.3244	1.0342
age	age	4	0.3091	0.3258	1.0539
REP_month	Replacement: month	2	0.3004	0.3922	1.3057
default	default	3	0.2493	0.2160	0.8663
campaign	campaign	1	0.0620	0.0000	0.0000

The most Important Variables : “duration”, “REP_pdays”, “contact”, “age”, “REP_month”, “default”, and “campaign”.

The Best Misclassification Tree

Picture 230. The Misclassification Tree, Variable Importance

Results - Node: Misclass DT 3 Br Diagram: Bank Marketing Analysis

File Edit View Window

Variable Importance

Variable Name	Label	Number of Splitting Rules	Importance	Validation Importance	Ratio of Validation to Training Importance
duration	duration	3	1.0000	1.0000	1.0000
poutcome	poutcome	1	0.5010	0.5392	1.0761
REP_month	Replacement: month	4	0.4671	0.4375	0.9366
contact	contact	1	0.3207	0.3424	1.0677
age	age	4	0.3098	0.2138	0.6900
REP_pdays	Replacement: pdays	1	0.2545	0.2745	1.0787
REP_education	Replacement: education	1	0.1029	0.0897	0.8720
default	default	1	0.0906	0.0000	0.0000
campaign	campaign	1	0.0671	0.0184	0.2744

The most Important Variables: “duration”, “poutcome”, “REP_month”, “contact”, “age”, “REP_pdays”, “Rep_education”, “default”, and “campaign”.

The Best Average Squared Error Tree

Picture 231. The ASE Tree, Variable Importance

Results - Node: ASE DT 12 Br Diagram: Bank Marketing Analysis

File Edit View Window

Variable Importance

Variable Name	Label	Number of Splitting Rules	Importance ▼	Validation Importance	Ratio of Validation to Training Importance
duration	duration	2	1.0000	1.0000	1.0000
REP_month	Replacement: month	9	0.4094	0.3981	0.9723
poutcome	poutcome	3	0.4002	0.3986	0.9959
contact	contact	4	0.3554	0.3751	1.0554
age	age	3	0.3246	0.3147	0.9696
REP_pdays	Replacement: pdays	4	0.2831	0.3127	1.1043
default	default	1	0.1860	0.1466	0.7879

The most important variables for the ASE Decision Tree Model with 12 branches are: “duration”, “REP_month”, “poutcome”, “contact”, “age”, “REP_pdays”, and “default”.

The Best Regression Model: Stepwise Regression

Picture 232. The Regression Model, Fit Statistics

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation
y	y	AIC	Akaike's Information Criterion	3742.779	.
y	y	ASE	Average Squared Error	0.125979	0.126304
y	y	AVERR	Average Error Function	0.398963	0.39481
y	y	DFE	Degrees of Freedom for Error	4617	.
y	y	DFM	Model Degrees of Freedom	21	.
y	y	DFT	Total Degrees of Freedom	4638	.
y	y	DIV	Divisor for ASE	9276	9284
y	y	ERR	Error Function	3700.779	3665.413
y	y	FPE	Final Prediction Error	0.127125	.
y	y	MAX	Maximum Absolute Error	0.999412	0.998773
y	y	MSE	Mean Square Error	0.126552	0.126304
y	y	NOBS	Sum of Frequencies	4638	4642
y	y	NW	Number of Estimate Weights	21	.
y	y	RASE	Root Average Sum of Squares	0.354935	0.355393
y	y	RFPE	Root Final Prediction Error	0.356545	.
y	y	RMSE	Root Mean Squared Error	0.355741	0.355393
y	y	SBC	Schwarz's Bayesian Criterion	3878.062	.
y	y	SSE	Sum of Squared Errors	1168.578	1172.606
y	y	SUMW	Sum of Case Weights Times Fr...	9276	9284
y	y	MISC	Misclassification Rate	0.176154	0.182464

ASE: The average squared error for validation datasets is 0.126304.

Misclassification rate: 0.182464, the misclassification rate is around 18%, meaning the model accurately predicts the term deposit subscription status about 82% of the time. While not perfect, this rate provides a reasonable level of accuracy in classification.

The most significant variables are “LOG_REP_duration”, “LOG_REP_campaign”, “REP_education”, “REP_job”, “REP_month”, “contact”, “default”, “marital”, and “poutcome”.

The Data Partitioning Neural Network Average Squared Error, 50 iterations, 3 hidden units

Table 17. Data Partition Neural Network Results

Neural Network	Number of Iterations	Number of Hidden Layers	Misclassification	ASE	Parameter Estimate
DP NN ASE 50	50	3	0.140147	0.107239	211

The Transform Variables Neural Network

Best Model: Transform Variables Neural Network : 50 Iterations 6 Hidden Units

Table 18. The Transform Variables Neural Network Results

Neural Network	Number of Iterations	Number of Hidden Layers	Misclassification	ASE	Parameter Estimate
TV NN ASE 50	50	4	0.163723	0.116953	129

Cap and Floor Neural Network

Best Model: Cap n Floor Neural Network: 50 Iterations 4 Hidden Units.

Table 19. Cap and Floor Neural Network Results

Neural Network	Number of Iterations	Number of Hidden Layers	Misclassification	ASE	Parameter Estimate
CF NN ASE 50	50	6	0.160060	0.116799	193

Stepwise Regression Neural Network

Best Model : Stepwise Regression Neural Network : 100 Iterations 5 Hidden Units

Table 20. Stepwise Regression Neural Networks Results

Neural Network	Number of Iterations	Number of Hidden Layers	Misclassification	ASE	Parameter Estimate
SR NN ASE 100	100	5	0.167385	0.117893	111

Model Comparison Node

Picture 233. Model Comparison, Fit Statistics

Results - Node: Model Comparison Diagram: Bank Marketing Analysis

File Edit View Window

Fit Statistics

Selected Model	Predecessor Node	Model Node	Model Description	Target Variable	Target Label	Selection Criterion: Valid: Average Squared Error ▲	Valid: Average Squared Error	Valid: Misclassification Rate	Valid: Roc Index	Valid: Lift	Valid: Cumulative Lift	Valid: Percent Response	Valid: Cumulative Percent Response	Valid: Percent Captured Response	Valid: Cumulative Percent Captured Response
Y	Neural	Neural	DP NN ASE 50 H3	y	y	0.107239	0.107239	0.147997	0.922	1.853448	1.896774	92.67241	94.83871	9.263249	19.00043
	Neural5	Neural5	DP NN ASE 100 H3	y	y	0.107239	0.107239	0.147997	0.922	1.853448	1.896774	92.67241	94.83871	9.263249	19.00043
	Neural3	Neural3	DP NN ASE 50 H5	y	y	0.109376	0.109376	0.150582	0.919	1.862069	1.87957	93.10345	93.97849	9.306333	18.82809
	Neural7	Neural7	DP NN ASE 100 H5	y	y	0.109376	0.109376	0.150582	0.919	1.862069	1.87957	93.10345	93.97849	9.306333	18.82809
	Neural2	Neural2	DP NN ASE 50 H4	y	y	0.111473	0.111473	0.155967	0.918	1.827586	1.883871	91.37931	94.19355	9.133994	18.87118
	Neural6	Neural6	DP NN ASE 100 H4	y	y	0.111473	0.111473	0.155967	0.918	1.827586	1.883871	91.37931	94.19355	9.133994	18.87118
	Neural4	Neural4	DP NN ASE 50 H6	y	y	0.112011	0.112011	0.153813	0.917	1.87069	1.913978	93.53448	95.69892	9.349418	19.17277
	Neural8	Neural8	DP NN ASE 100 H6	y	y	0.112011	0.112011	0.153813	0.917	1.87069	1.913978	93.53448	95.69892	9.349418	19.17277
	Neural12	Neural12	C&F NN ASE 50 H6	y	y	0.116799	0.116799	0.16006	0.913	1.965517	1.922581	98.27586	96.12903	9.823352	19.25894
	Neural16	Neural16	C&F NN ASE 100 H6	y	y	0.116799	0.116799	0.16006	0.913	1.965517	1.922581	98.27586	96.12903	9.823352	19.25894
	Neural14	Neural14	C&F NN ASE 100 H4	y	y	0.116841	0.116841	0.162215	0.913	1.853448	1.883871	92.67241	94.19355	9.263249	18.87118
	Neural15	Neural15	C&F NN ASE 100 H5	y	y	0.116917	0.116917	0.163292	0.913	1.896552	1.883871	94.82759	94.19355	9.478673	18.87118
	Neural17	Neural17	Trans NN ASE 50 H4	y	y	0.116953	0.116953	0.163723	0.911	1.896552	1.87957	94.82759	93.97849	9.478673	18.82809
	Neural21	Neural21	Trans NN ASE 100 H4	y	y	0.116953	0.116953	0.163723	0.911	1.896552	1.87957	94.82759	93.97849	9.478673	18.82809
	Neural10	Neural10	C&F NN ASE 50 H4	y	y	0.117521	0.117521	0.162861	0.912	1.862069	1.888172	93.10345	94.4086	9.306333	18.91426
	Neural30	Neural30	SRegr NN ASE 100 H5	y	y	0.117893	0.117893	0.167385	0.911	1.827586	1.883871	91.37931	94.19355	9.133994	18.87118
	Neural27	Neural27	SRegr NN ASE 50 H6	y	y	0.118298	0.118298	0.170616	0.913	1.905172	1.931183	95.25862	96.55914	9.521758	19.34511

The most optimal effective model: Data Partition Neural Network Average Squared Error 50 Iterations 3 Hidden Units.

- ASE: 0.107239
- Misclassification Rate: 0.147997
- ROC Index : 0.922

ACTIONABLE INSIGHTS

Decision Trees

Optimal Time for Calls. By analyzing our Maximal tree, we have identified a crucial insight. In 100% of cases, individuals are more likely to subscribe if a bank representative calls them from April to June, within the next 7 days after the previous call. Additionally, the conversation duration should be between 256 and 480 seconds.

Effective Cellular Communication. Our analysis reveals a consistent pattern echoed in the Maximal Tree. In 96.88% of cases, successful subscriptions occur when a bank representative makes a call via cellular communication. The optimal conversation duration ranges from 157 to 441 seconds, and the call takes place in March. Moreover, a positive outcome from the previous marketing campaign contributes to the likelihood of subscription.

Tailored Approach for Success. Building on the Maximal Tree's insights, we find that in 100% of cases, a person is inclined to subscribe when the call duration falls between 162 and 256 seconds. This is particularly effective in March, targeting individuals aged 26-60 years with no positive outcome from the

previous marketing campaign. Tailoring calls based on these parameters can significantly increase subscription rates.

Stepwise Regression (Best Model)

Optimize Call Duration for Higher Subscriptions. Encourage representatives to engage in longer interactions with clients during telemarketing calls. Considering that each additional second of contact significantly increases the odds of subscription by 821.2%, extending the duration of conversations may positively impact subscription rates.

Target Retired Clients. Focus marketing efforts on retired clients, as they have 139.9% higher odds of subscribing compared to clients with an unknown status. Tailoring campaigns and messaging to resonate with the needs and preferences of retired individuals may yield more successful outcomes

Seasonal Campaign Adjustments. Consider adjusting marketing strategies based on the seasonality effect observed. The odds of subscription are 718.5% higher in Q1 compared to Q4. This suggests potential variations in client responsiveness during different quarters, and adapting campaigns to align with seasonal trends could enhance effectiveness.

Leverage Cellular Communication. Emphasize the use of cellular communication methods in marketing campaigns, as the odds of subscription are 139.2% higher when the communication type is cellular compared to telephone. Prioritizing mobile channels may better resonate with clients and increase subscription rates.

Target Specific Marital Status Groups. Tailor marketing approaches for specific marital status groups. Divorced, married, and single clients collectively have higher odds of subscription compared to those with an unknown marital status. Specifically, divorced clients exhibit 165.9% higher odds, indicating potential success in targeting this demographic with customized campaigns.

Data Partitioning Neural Network: 50 Iterations 3 Hidden Units

This was our best for the entire model, the Model Comparison node also picked the same Neural Network as the best fit. This particular model achieved an impressive ROC Index of 0.922, showcasing its excellent ability to distinguish between different outcomes. With an Average Squared Error of 0.107239 and a misclassification rate of 0.147997, it demonstrated high accuracy and reliability.

Here is a list of Top 45 people who have the highest chance to subscribe with the bank. The predicted probability for them to sign up is more than 99% .

Picture 234. Data Partition Neural Network, 50 iterations, 3 HU, top 45 clients to contact with highest probability

EMWS7.Neural_VALIDATE

	Observation Number	age	job	marital	education	Predicted: y=yes	default	housing	loan	contact	month	day_of_week	duration	campaign	pdays	previous	poutcome	euribor3m	y	emp_var_rate	cons_price_idx	cons_conf_idx	nr_employed
1	2128.0	26	admin.	single	university.degree	0.999814603596295	no	yes	no	telephone	sep	thu	471	1	7	1	success	0.809	yes	-3.4	92.379	-29.8	5018.0
2	1855.0	37	technician	married	university.degree	0.999778329275136 3	no	yes	no	telephone	sep	tue	773	1	3	2	success	0.755	yes	-3.4	92.379	-29.8	5018.0
3	43.0	33	admin.	single	university.degree	0.999367468045928 4	no	yes	no	cellular	apr	thu	280	1	7	2	success	0.635	yes	-1.8	93.749	-34.6	5009.0
4	1222.0	48	admin.	married	high.school	0.999233115691333	unknown	no	yes	cellular	dec	thu	274	1	8	1	success	0.72	yes	-3.0	92.713	-33.0	5024.0
5	2310.0	83	retired	married	university.degree	0.999170036413988 4	no	no	no	telephone	jul	tue	617	1	12	1	success	0.835	yes	-1.7	94.215	-40.3	4992.0
6	323.0	34	admin.	single	university.degree	0.999107362779885 9	no	no	no	telephone	mar	tue	225	1	3	1	success	0.652	yes	-1.8	93.369	-34.8	5009.0
7	279.0	23	student	single	high.school	0.998955594687994 2	no	no	no	telephone	mar	fri	290	1	3	1	success	0.649	yes	-1.8	93.369	-34.8	5009.0
8	984.0	53	housemaid	married	basic.4y	0.998899080336682 5	no	no	yes	telephone	nov	wed	513	1	3	1	success	0.715	yes	-3.4	92.649	-30.1	5018.0
9	13061.0	34	blue-collar	married	basic.9y	0.998848321639547 2	unknown	yes	no	telephone	apr	mon	1353	2	5	1	success	1.466	yes	-1.8	93.075	-47.1	5099.0
10	10793.0	60	admin.	married	high.school	0.998762071048768 1	no	no	no	cellular	apr	thu	447	1	2	1	success	1.365	yes	-1.8	93.075	-47.1	5099.0
11	904.0	55	admin.	married	professional.course	0.998739037486197 4	no	yes	no	cellular	dec	thu	798	1	9	2	success	0.714	yes	-3.0	92.713	-33.0	5024.0
12	4119.0	60	admin.	married	unknown	0.998712161808700 9	no	no	no	cellular	nov	tue	787	1	7	2	success	1.046	yes	-1.1	94.767	-50.8	4964.0
13	3633.0	61	retired	married	university.degree	0.998699534192909	no	no	no	telephone	aug	wed	250	2	7	1	success	0.9	yes	-1.7	94.027	-38.3	4992.0
14	338.0	38	blue-collar	single	high.school	0.998681326489950 8	no	no	no	telephone	mar	tue	450	3	6	2	success	0.652	yes	-1.8	93.369	-34.8	5009.0
15	10781.0	53	admin.	single	university.degree	0.998613098413468 9	no	no	no	cellular	apr	thu	494	1	2	1	success	1.365	yes	-1.8	93.075	-47.1	5099.0
16	3703.0	26	admin.	single	university.degree	0.998583359565613 4	no	no	no	telephone	aug	mon	1087	1	3	1	success	0.904	yes	-1.7	94.027	-38.3	4992.0
17	10942.0	28	admin.	single	university.degree	0.998579805330836 6	no	no	yes	telephone	apr	thu	917	1	999	0	nonexistent	1.365	yes	-1.8	93.075	-47.1	5099.0
18	2915.0	45	unemployed	married	professional.course	0.998355826650862 9	unknown	no	no	telephone	sep	thu	1405	1	6	2	failure	0.879	yes	-1.1	94.199	-37.5	4964.0
19	2797.0	30	technician	single	professional.course	0.998284150140413 9	no	unknown	unkn...	telephone	sep	tue	808	1	12	2	failure	0.877	yes	-1.1	94.199	-37.5	4964.0
20	4005.0	62	retired	married	university.degree	0.998252806580595 3	no	yes	no	cellular	nov	thu	208	1	1	6	success	1.031	yes	-1.1	94.767	-50.8	4964.0
21	11027.0	34	technician	married	university.degree	0.998167342760175 5	no	yes	no	cellular	apr	thu	193	2	8	1	success	1.365	no	-1.8	93.075	-47.1	5099.0
22	2505.0	61	retired	married	university.degree	0.998016892295873 1	no	no	no	telephone	aug	tue	249	2	3	1	success	0.859	yes	-2.9	92.201	-31.4	5076.0
23	3265.0	42	admin.	married	university.degree	0.997991804276217 9	no	yes	no	cellular	aug	tue	576	1	4	1	success	0.884	yes	-2.9	92.201	-31.4	5076.0
24	4160.0	18	student	single	basic.4y	0.997971772023125	no	yes	no	telephone	nov	tue	394	1	13	2	success	1.049	yes	-1.1	94.767	-50.8	4964.0
25	420.0	37	technician	married	professional.course	0.997944545162233 7	no	no	no	cellular	apr	thu	277	1	18	1	success	0.659	yes	-1.8	93.749	-34.6	5009.0
26	10964.0	40	technician	married	professional.course	0.997905351712943 6	no	yes	no	cellular	apr	thu	580	2	2	1	success	1.365	yes	-1.8	93.075	-47.1	5099.0
27	12359.0	32	blue-collar	married	basic.9y	0.997831802532767 4	no	no	no	cellular	apr	thu	583	1	5	1	success	1.41	yes	-1.8	93.075	-47.1	5099.0
28	2617.0	42	management	married	university.degree	0.997760737553594 2	no	yes	no	cellular	aug	fri	436	1	13	1	success	0.869	yes	-2.9	92.201	-31.4	5076.0
29	10915.0	56	entrepreneur	married	university.degree	0.997738539164428 5	no	no	no	cellular	apr	thu	457	1	2	1	success	1.365	yes	-1.8	93.075	-47.1	5099.0
30	344.0	28	unemployed	married	high.school	0.997588850485294 4	no	no	no	telephone	mar	fri	341	4	6	1	success	0.653	yes	-1.8	93.369	-34.8	5009.0
31	1028.0	59	retired	married	professional.course	0.997553829651830 4	no	no	no	telephone	dec	tue	476	1	999	1	failure	0.715	yes	-3.0	92.713	-33.0	5024.0
32	11040.0	60	unemployed	married	high.school	0.997115623712492	no	yes	no	cellular	apr	thu	1260	2	2	1	success	1.365	yes	-1.8	93.075	-47.1	5099.0
33	416.0	38	admin.	single	high.school	0.997108364138162 5	no	yes	no	cellular	apr	thu	882	1	6	1	success	0.659	yes	-1.8	93.749	-34.6	5009.0
34	192.0	50	technician	single	university.degree	0.996972137948111 2	no	yes	no	telephone	apr	fri	634	1	999	0	nonexistent	0.644	yes	-1.8	93.749	-34.6	5009.0
35	130.0	45	management	married	unknown	0.996966759149932 5	no	yes	no	cellular	apr	tue	379	3	7	1	success	0.642	yes	-1.8	93.749	-34.6	5009.0
36	140.0	53	blue-collar	single	basic.9y	0.996947384829179 2	no	unknown	unkn...	cellular	mar	thu	362	1	9	2	success	0.643	yes	-1.8	93.369	-34.8	5009.0
37	193.0	58	unemployed	married	basic.9y	0.996903643075524 1	no	yes	no	cellular	apr	thu	229	1	6	1	success	0.644	yes	-1.8	93.749	-34.6	5009.0
38	2848.0	49	admin.	married	high.school	0.996837298768680 9	no	yes	no	cellular	aug	wed	668	2	15	1	success	0.879	yes	-2.9	92.201	-31.4	5076.0
39	1157.0	44	admin.	divorced	high.school	0.996769297429326 1	no	yes	no	telephone	nov	mon	1628	2	6	1	success	0.719	yes	-3.4	92.649	-30.1	5018.0
40	3997.0	32	admin.	married	university.degree	0.996619236191341 7	no	no	no	telephone	nov	wed	651	1	999	1	failure	1.03	yes	-1.1	94.767	-50.8	4964.0
41	1008.0	86	retired	married	professional.course	0.996525151277909 1	no	no	no	telephone	dec	wed	343	2	999	0	nonexistent	0.715	yes	-3.0	92.713	-33.0	5024.0
42	13168.0	28	blue-collar	single	high.school	0.996524257480543 3	no	yes	no	telephone	apr	thu	771	1	999	0	nonexistent	1.483	yes	-1.8	93.075	-47.1	5099.0
43	1057.0	54	admin.	married	high.school	0.996442467398724 6	no	yes	no	cellular	nov	fri	402	1	3	3	success	0.716	yes	-3.4	92.649	-30.1	5018.0
44	3618.0	76	retired	married	university.degree	0.996405398412708 3	no	yes	no	cellular	aug	thu	504	2	6	3	success	0.899	yes	-1.7	94.027	-38.3	4992.0
45	2200.0	38	technician	married	professional.course	0.996393816290500 9	no	unknown	unkn...	telephone	aug	mon	473	2	9	1	success	0.821	yes	-2.9	92.201	-31.4	5076.0

RECOMMENDATIONS

Based on these actionable insights, here are recommendations we would like to present.

Optimize Call Timing and Duration

- Schedule calls between April and June, within 7 days after the previous call.
- Aim for conversations lasting between 256 and 480 seconds, as indicated by the Maximal Tree's success pattern.

Leverage Cellular Communication

- Prioritize cellular communication methods, as individuals are 139.2% more likely to subscribe through this channel.

Tailor Calls for Seasonal Trends

- Adjust marketing strategies based on seasonality, with a focus on Q1, where odds of subscription are 718.5% higher.

Engage with Specific Demographics

- Target retired individuals, as they exhibit 139.9% higher odds of subscribing compared to those with an unknown status.
- Customize campaigns for divorced, married, and single clients, who collectively show higher odds of subscription.

Encourage Longer Interactions

- Encourage representatives to engage in longer conversations, as each additional second of contact significantly increases the odds of subscription by 821.2%.

CITATION AND PAST USAGE

The dataset used in this analysis has been the subject of previous studies related to bank telemarketing, indicating its relevance and importance in understanding marketing dynamics. In particular, S. Moro, P. Cortes, and P. Rita analyzed this dataset in detail in their paper published in Decision Support Systems entitled "A Data-Driven Approach to Predict the Success of Bank Telemarketing" (June 2014). Their study emphasized the key role of the dataset in predicting the results of bank telemarketing campaigns and laid the groundwork for subsequent research in this area.

In addition, the dataset has been cited and used in studies on banks' marketing strategies. Moro, Laureano, and Cortez (Moro, Laureano, and Cortez, 2011) reviewed its application in their paper "Using Data Mining for Bank Direct Marketing: An Application of the CRISP-DM Methodology" presented at the European Modeling and Simulation Conference (ESM'2011). This research reveals the use of the dataset and its contribution to understanding customer behavior and campaign effectiveness in bank telemarketing.

REFERENCES

Moro, S., Cortes, P., & Rita, P. (2014). *A Data-Driven Approach to Predict the Success of Bank Telemarketing*. Decision Support Systems, Volume 62, 2014, Pages 22-31, ISSN 0167-9236, https://doi.org/10.1016/j.dss.2014.03.001

APPENDIX

Appendix. Picture 1. StatExplore Node after BankMarketing Node

Results - Node: StatExplore (3) Diagram: Bank Marketing

File Edit View Window

Interval Variables

Data Role	Target	Target Level	Variable	Median	Missing	Non Missing	Minimum	Maximum	Mean	Standard Deviation	Skewness	Kurtosis	Role	Label	Scaled Mean Deviation	Maximum Deviation	Level Id
TRAIN	y	no	emp_var_rate	1.1	0	36548	-3.4	1.4	0.248875	1.482932	-0.9086	-0.75163	INPUT	emp.var.rate	2.03931	16.06309	1
TRAIN	y	yes	emp_var_rate	-1.8	0	4640	-3.4	1.4	-1.23345	1.623626	0.556589	-0.99048	INPUT	emp.var.rate	-16.0631	16.06309	2
TRAIN	y	no	previous	0	0	36548	0	7	0.132374	0.409199	4.047547	23.81719	INPUT	previous	-0.23467	1.848427	1
TRAIN	y	yes	previous	0	0	4640	0	6	0.492672	0.860344	2.132789	5.267404	INPUT	previous	1.848427	1.848427	2
TRAIN	y	no	duration	163	0	36548	0	4918	220.8448	207.0963	3.668784	31.21465	INPUT	duration	-0.14496	1.141786	1
TRAIN	y	yes	duration	449	0	4640	37	4199	553.1912	401.1719	1.837301	6.318574	INPUT	duration	1.141786	1.141786	2
TRAIN	y	no	euribor3m	4.857	0	36548	0.634	5.045	3.811491	1.638187	-0.94332	-1.00919	INPUT	euribor3m	0.052523	0.413708	1
TRAIN	y	yes	euribor3m	1.266	0	4640	0.634	5.045	2.123135	1.742598	0.890782	-1.08895	INPUT	euribor3m	-0.41371	0.413708	2
TRAIN	y	no	campaign	2	0	36548	1	56	2.633085	2.873438	4.679777	35.21286	INPUT	campaign	0.025507	0.200915	1
TRAIN	y	yes	campaign	2	0	4640	1	23	2.051724	1.666245	3.387153	19.35519	INPUT	campaign	-0.20092	0.200915	2
TRAIN	y	no	age	38	0	36548	17	95	39.91119	9.898132	0.652875	0.363927	INPUT	age	-0.00282	0.022214	1
TRAIN	y	yes	age	37	0	4640	17	98	40.91315	13.83748	0.999425	0.671076	INPUT	age	0.022214	0.022214	2
TRAIN	y	no	cons_conf_idx	-41.8	0	36548	-50.8	-26.9	-40.5931	4.391155	0.250342	-0.47939	INPUT	cons.conf.idx	0.002234	0.017599	1
TRAIN	y	yes	cons_conf_idx	-40.4	0	4640	-50.8	-26.9	-39.7898	6.139668	0.259931	-0.78703	INPUT	cons.conf.idx	-0.0175	0.017599	2
TRAIN	y	no	nr_employed	5196	0	36548	4964	5228	5176.137	64.52434	-1.20599	0.528275	INPUT	nr.employed	0.001765	0.013899	1
TRAIN	y	yes	nr_employed	5099	0	4640	4964	5228	5095.201	87.40568	0.282066	-1.19989	INPUT	nr.employed	-0.0139	0.013899	2
TRAIN	y	no	cons_price_idx	93.918	0	36548	92.201	94.767	93.60376	0.558993	-0.238	-0.8243	INPUT	cons.price.idx	.0003002	0.002365	1
TRAIN	y	yes	cons_price_idx	93.2	0	4640	92.201	94.767	93.35439	0.676644	0.119767	-1.05428	INPUT	cons.price.idx	-0.00236	0.002365	2

Appendix. Picture 2. Sample Node Settings

Property	Value
General	
Node ID	Smpl2
Imported Data	
Exported Data	
Notes	
Train	
Variables	
Output Type	Data
Sample Method	Stratify
Random Seed	12345
Size	
Type	Percentage
Observations	.
Percentage	100.0
Alpha	0.01
PValue	0.01
Cluster Method	Random
Stratified	
Criterion	Level Based
Ignore Small Strata	No
Minimum Strata Size	5
Level Based Options	
Level Selection	Event
Level Proportion	100.0
Sample Proportion	50.0
Oversampling	
Adjust Frequency	No
Based on Count	No
Exclude Missing Levels	No
Report	
Interval Targets	Yes
Class Targets	Yes
Status	
Create Time	06/12/23 3:33 PM
Run ID	0cf138dc-25a3-4859-89c5-06
Last Error	
Last Status	Complete
Last Run Time	11/12/23 5:51 PM
Run Duration	0 Hr. 0 Min. 9.01 Sec.
Grid Host	
User-Added Node	No

Appendix. Picture 3. Sample Results

```
Summary Statistics for Class Targets
(maximum 500 observations printed)

Data=DATA

                Numeric     Formatted     Frequency
Variable         Value        Value         Count       Percent     Label

   y               .            no           36548       88.7346       y
   y               .            yes           4640       11.2654       y


Data=SAMPLE

                Numeric     Formatted     Frequency
Variable         Value        Value         Count       Percent     Label

   y               .            no            4640          50         y
   y               .            yes           4640          50         y
```

Appendix. Picture 4. StatExplore after Sampling

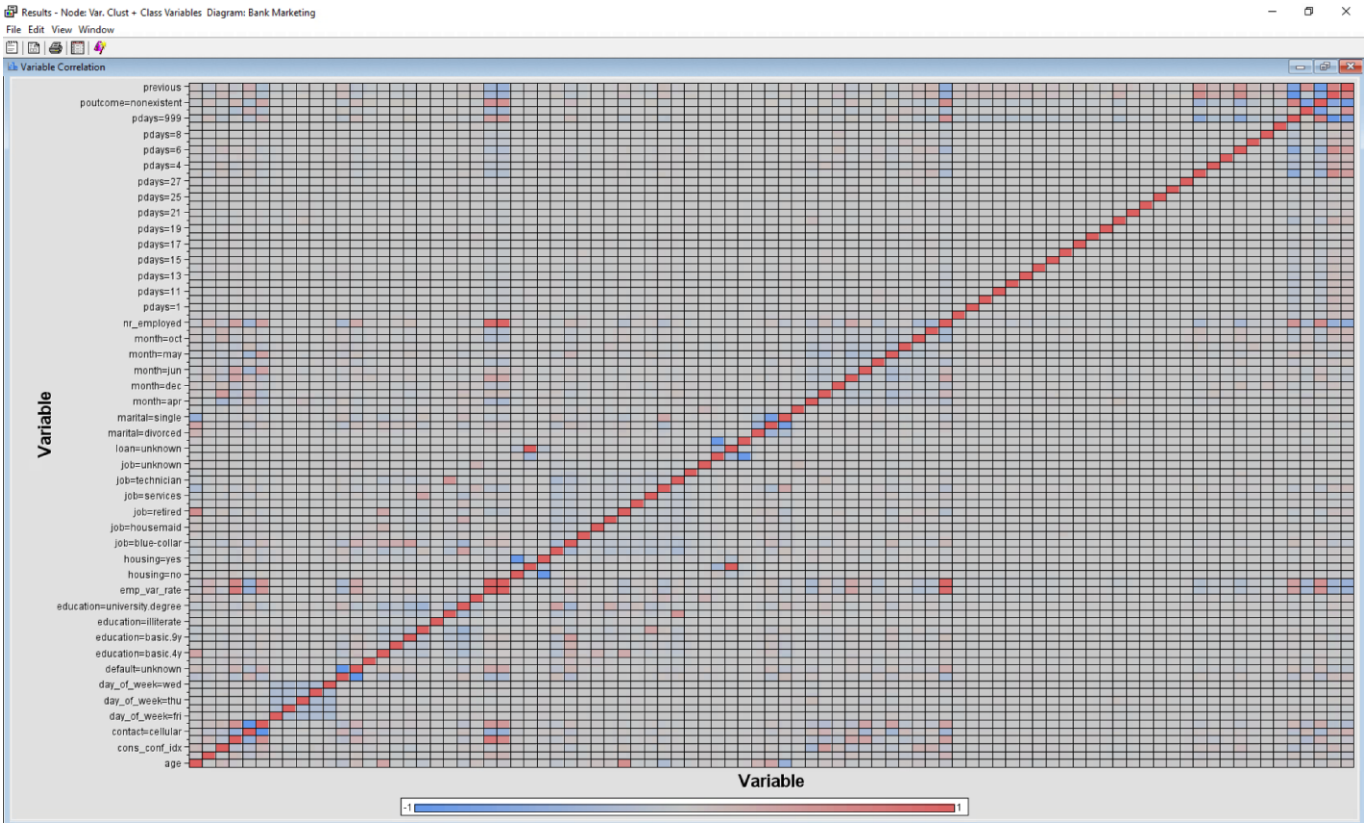
Results - Node: StatExplore (4) Diagram: Bank Marketing

File Edit View Window

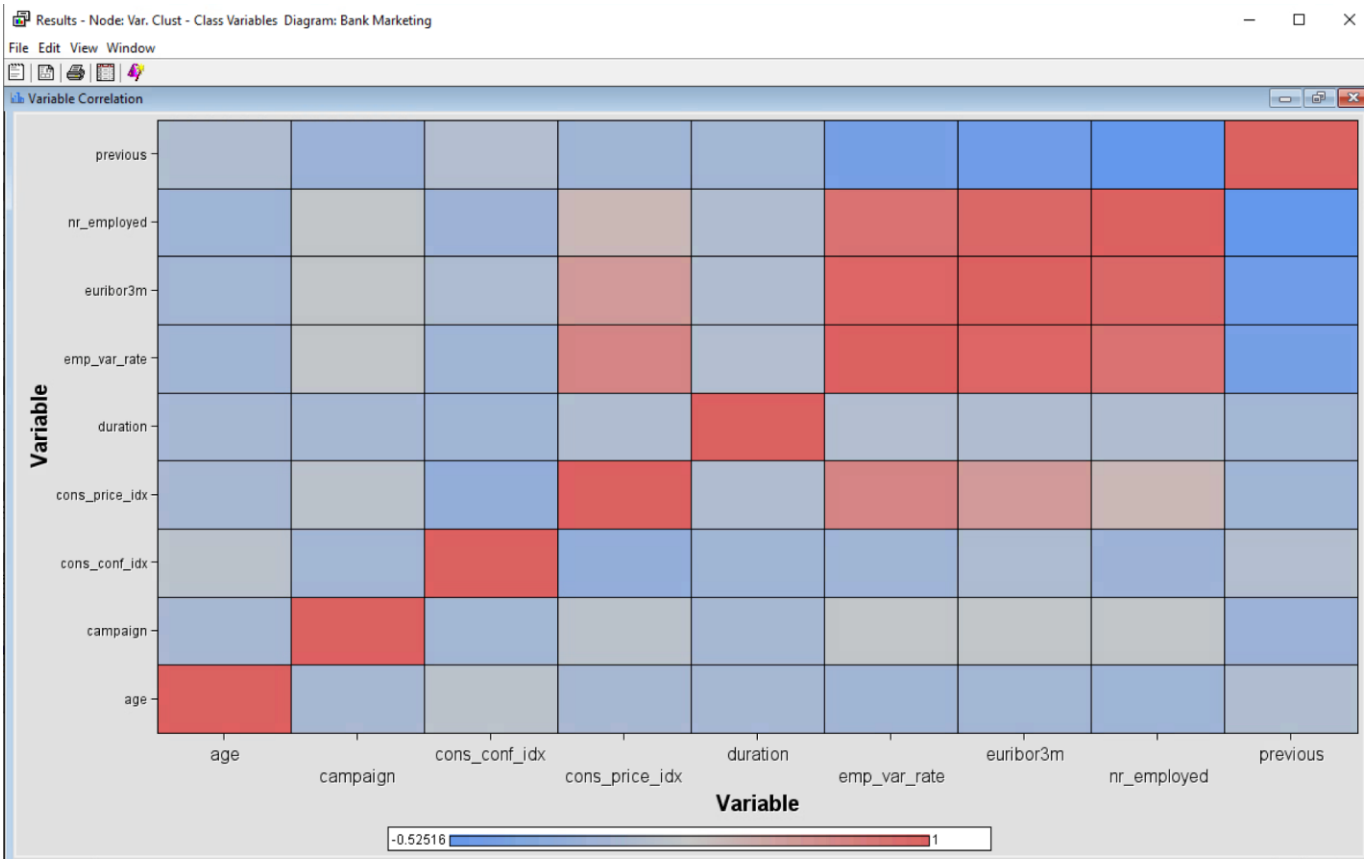
Interval Variables

Data Role	Target	Target Level	Variable	Median	Missing	Non Missing	Minimum	Maximum	Mean	Standard Deviation	Skewness	Kurtosis	Role	Label	Scaled Mean Deviation	Maximum Deviation	Level Id
TRAIN	y	no	emp_var_r...	1.1	0	4640	-3.4	1.4	0.270129	1.473832	-0.94326	-0.67785	INPUT	emp.var.rate	-1.56083	1.56083	1
TRAIN	y	yes	emp_var_r...	-1.8	0	4640	-3.4	1.4	-1.23345	1.623626	0.556589	-0.99048	INPUT	emp.var.rate	1.56083	1.56083	2
TRAIN	y	no	previous	0	0	4640	0	5	0.140302	0.422391	3.799768	19.06447	INPUT	previous	-0.55669	0.556691	1
TRAIN	y	yes	previous	0	0	4640	0	6	0.492672	0.860344	2.132789	5.267404	INPUT	previous	0.556691	0.556691	2
TRAIN	y	no	duration	166	0	4640	0	4918	222.505	222.8422	5.390659	69.63531	INPUT	duration	-0.42631	0.426309	1
TRAIN	y	yes	duration	449	0	4640	37	4199	553.1912	401.1719	1.837301	6.318574	INPUT	duration	0.426309	0.426309	2
TRAIN	y	no	euribor3m	4.858	0	4640	0.635	5.045	3.835632	1.6265	-0.97815	-0.94089	INPUT	euribor3m	0.287391	0.287391	1
TRAIN	y	yes	euribor3m	1.266	0	4640	0.634	5.045	2.123135	1.742598	0.890782	-1.08895	INPUT	euribor3m	-0.28739	0.287391	2
TRAIN	y	no	campaign	2	0	4640	1	33	2.678233	2.834693	3.968208	23.28107	INPUT	campaign	0.132455	0.132455	1
TRAIN	y	yes	campaign	2	0	4640	1	23	2.051724	1.666245	3.387153	19.35519	INPUT	campaign	-0.13246	0.132455	2
TRAIN	y	no	age	38	0	4640	17	88	39.98793	9.823154	0.611795	0.22527	INPUT	age	-0.01144	0.011436	1
TRAIN	y	yes	age	37	0	4640	17	98	40.91315	13.83748	0.999425	0.671076	INPUT	age	0.011436	0.011436	2
TRAIN	y	no	cons_conf_...	-41.8	0	4640	-50.8	-26.9	-40.5479	4.392402	0.276459	-0.39105	INPUT	cons.conf.idx	0.009437	0.009437	1
TRAIN	y	yes	cons_conf_...	-40.4	0	4640	-50.8	-26.9	-39.7898	6.139668	0.259931	-0.78703	INPUT	cons.conf.idx	-0.00944	0.009437	2
TRAIN	y	no	nr_employed	5196	0	4640	4964	5228	5177.074	63.86761	-1.23794	0.641159	INPUT	nr.employed	0.00797	0.00797	1
TRAIN	y	yes	nr_employed	5099	0	4640	4964	5228	5095.201	87.40568	0.282066	-1.19989	INPUT	nr.employed	-0.00797	0.00797	2
TRAIN	y	no	cons_price...	93.918	0	4640	92.201	94.767	93.60721	0.558607	-0.25992	-0.80801	INPUT	cons.price.i...	0.001352	0.001352	1
TRAIN	y	yes	cons_price...	93.2	0	4640	92.201	94.767	93.35439	0.676644	0.119767	-1.05428	INPUT	cons.price.i...	-0.00135	0.001352	2

Appendix. Picture 5. Variable Clustering with Class variables results



Appendix. Picture 6. Variable Clustering without Class variables results



Appendix. Picture 7. Sample Statistics after rejecting variables

Explore - EMWS5.FIMPORT2_train

File View Actions Window

Sample Statistics

Obs #	Variable ...	Label	Type	Percent ...	Minimum	Maximum	Mean	Number o...	Mode Per...	Mode
1	contact	contact	CLASS	0	.	.	.	2	87.28333	CELLULAR
2	day_of_week	day_of_week	CLASS	0	.	.	.	5	24.11667	FRI
3	default	default	CLASS	0	.	.	.	2	95.51667	NO
4	education	education	CLASS	0	.	.	.	8	37.5	UNIVERSIT...
5	housing	housing	CLASS	0	.	.	.	3	55.55	YES
6	job	job	CLASS	0	.	.	.	12	30.83333	ADMIN.
7	loan	loan	CLASS	0	.	.	.	3	82.73333	NO
8	marital	marital	CLASS	0	.	.	.	4	51.13333	MARRIED
9	month	month	CLASS	0	.	.	.	10	24.58333	MAY
10	poutcome	poutcome	CLASS	0	.	.	.	3	56.15	NONEXIST...
11	y	y	CLASS	0	.	.	.	2	62.21667	NO
12	age	age	VAR	0	17	98	41.52533	.	.	
13	campaign	campaign	VAR	0	1	16	2.057833	.	.	
14	duration	duration	VAR	0	1	3785	272.6285	.	.	
15	pdays	pdays	VAR	0	0	999	792.4865	.	.	

Appendix. Picture 8. Data Partition Settings

.. Property	Value
General	
Node ID	Part
Imported Data	...
Exported Data	...
Notes	...
Train	
Variables	...
Output Type	Data
Partitioning Method	Default
Random Seed	12345
⊟ Data Set Allocations	
Training	50.0
Validation	50.0
Test	0.0
Report	
Interval Targets	Yes
Class Targets	Yes
Status	
Create Time	14/12/23 6:43 AM
Run ID	aa8bfefd-2586-4779-b448-e
Last Error	
Last Status	Complete
Last Run Time	14/12/23 6:50 AM
Run Duration	0 Hr. 0 Min. 4.12 Sec.
Grid Host	
User-Added Node	No

Appendix. Picture 9. Fit statistics, ASE 3 branches, 33 leafs

Results - Node: ASE DT 3 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.164511	0.177941	.
y	y	_MAX_	Maximum Absolute Error	0.990446	0.990446	.
y	y	_SSE_	Sum of Squared Errors	1118.58	1207.526	.
y	y	_ASE_	Average Squared Error	0.120589	0.130065	.
y	y	_RASE_	Root Average Squared Error	0.347259	0.360646	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 10. Fit statistics, ASE 4 branches,38 leafs

Results - Node: ASE DT 4 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.163648	0.179664	.
y	y	_MAX_	Maximum Absolute Error	0.988095	1	.
y	y	_SSE_	Sum of Squared Errors	1077.667	1172.197	.
y	y	_ASE_	Average Squared Error	0.116178	0.12626	.
y	y	_RASE_	Root Average Squared Error	0.340849	0.355331	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 11. Fit statistics,ASE 5 branches,39 leafs

Results - Node: ASE DT 5 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.173351	0.188927	.
y	y	_MAX_	Maximum Absolute Error	0.983696	0.983696	.
y	y	_SSE_	Sum of Squared Errors	1104.469	1191.741	.
y	y	_ASE_	Average Squared Error	0.119067	0.128365	.
y	y	_RASE_	Root Average Squared Error	0.345061	0.358281	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 12. Fit statistics, ASE 6 branches,39 leafs

Results - Node: ASE DT 6 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.173351	0.188927	.
y	y	_MAX_	Maximum Absolute Error	0.983696	0.983696	.
y	y	_SSE_	Sum of Squared Errors	1104.469	1191.741	.
y	y	_ASE_	Average Squared Error	0.119067	0.128365	.
y	y	_RASE_	Root Average Squared Error	0.345061	0.358281	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 13. Fit statistics, ASE 7 branches,39 leaves

Results - Node: ASE DT 7 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.173351	0.188927	.
y	y	_MAX_	Maximum Absolute Error	0.983696	0.983696	.
y	y	_SSE_	Sum of Squared Errors	1104.469	1191.741	.
y	y	_ASE_	Average Squared Error	0.119067	0.128365	.
y	y	_RASE_	Root Average Squared Error	0.345061	0.358281	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 14. Fit statistics, ASE 8 branches 39 leafs

Results - Node: ASE DT 8 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.173351	0.188927	.
y	y	_MAX_	Maximum Absolute Error	0.983696	0.983696	.
y	y	_SSE_	Sum of Squared Errors	1104.469	1191.741	.
y	y	_ASE_	Average Squared Error	0.119067	0.128365	.
y	y	_RASE_	Root Average Squared Error	0.345061	0.358281	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 15. Fit statistics, ASE 9 branches 39 leafs

Results - Node: ASE DT 9 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.173351	0.188927	.
y	y	_MAX_	Maximum Absolute Error	0.983696	0.983696	.
y	y	_SSE_	Sum of Squared Errors	1104.469	1191.741	.
y	y	_ASE_	Average Squared Error	0.119067	0.128365	.
y	y	_RASE_	Root Average Squared Error	0.345061	0.358281	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 16. Fit statistics, ASE 10 branches 39 leafs

Results - Node: ASE DT 10 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.173351	0.188927	.
y	y	_MAX_	Maximum Absolute Error	0.983696	0.983696	.
y	y	_SSE_	Sum of Squared Errors	1104.469	1191.741	.
y	y	_ASE_	Average Squared Error	0.119067	0.128365	.
y	y	_RASE_	Root Average Squared Error	0.345061	0.358281	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 17. Fit statistics, ASE 11 branches 45 leafs

Results - Node: ASE DT 11 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.16947	0.182034	.
y	y	_MAX_	Maximum Absolute Error	0.986264	1	.
y	y	_SSE_	Sum of Squared Errors	1107.841	1182.239	.
y	y	_ASE_	Average Squared Error	0.119431	0.127342	.
y	y	_RASE_	Root Average Squared Error	0.345588	0.35685	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 18. Fit statistics, ASE 12 branches 49 leafs

Results - Node: ASE DT 12 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.166667	0.177941	.
y	y	_MAX_	Maximum Absolute Error	0.991304	1	.
y	y	_SSE_	Sum of Squared Errors	1098.849	1166.211	.
y	y	_ASE_	Average Squared Error	0.118461	0.125615	.
y	y	_RASE_	Root Average Squared Error	0.344182	0.354422	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 19. Fit statistics, ASE 13 branches 50 leafs

Results - Node: ASE DT 13 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.16602	0.184619	.
y	y	_MAX_	Maximum Absolute Error	0.989474	1	.
y	y	_SSE_	Sum of Squared Errors	1092.832	1190.159	.
y	y	_ASE_	Average Squared Error	0.117813	0.128195	.
y	y	_RASE_	Root Average Squared Error	0.343239	0.358043	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 20. Fit statistics, ASE 14 branches 45 leafs

Results - Node: ASE DT 14 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.170548	0.181603	.
y	y	_MAX_	Maximum Absolute Error	0.979695	1	.
y	y	_SSE_	Sum of Squared Errors	1114.328	1185.746	.
y	y	_ASE_	Average Squared Error	0.12013	0.127719	.
y	y	_RASE_	Root Average Squared Error	0.346598	0.357378	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 21. Fit statistics, ASE 15 branches 40 leafs

Results - Node: ASE DT 15 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.165589	0.18268	.
y	y	_MAX_	Maximum Absolute Error	0.992958	0.992958	.
y	y	_SSE_	Sum of Squared Errors	1088.786	1173.755	.
y	y	_ASE_	Average Squared Error	0.117377	0.126428	.
y	y	_RASE_	Root Average Squared Error	0.342603	0.355567	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 22. Fit statistics, ASE 16 branches 44 leafs

Results - Node: ASE DT 16 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.166882	0.183757	.
y	y	_MAX_	Maximum Absolute Error	0.987879	1	.
y	y	_SSE_	Sum of Squared Errors	1097.423	1180.655	.
y	y	_ASE_	Average Squared Error	0.118308	0.127171	.
y	y	_RASE_	Root Average Squared Error	0.343959	0.35661	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 23. Fit statistics, ASE 17 branches 44 leafs

Results - Node: ASE DT 17 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.167745	0.187635	.
y	y	_MAX_	Maximum Absolute Error	0.979487	1	.
y	y	_SSE_	Sum of Squared Errors	1094.307	1183.788	.
y	y	_ASE_	Average Squared Error	0.117972	0.127508	.
y	y	_RASE_	Root Average Squared Error	0.34347	0.357083	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 24. Fit statistics, ASE 18 branches 49 leafs

Results - Node: ASE DT 18 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.163217	0.186558	.
y	y	_MAX_	Maximum Absolute Error	0.980769	1	.
y	y	_SSE_	Sum of Squared Errors	1073.367	1185.627	.
y	y	_ASE_	Average Squared Error	0.115714	0.127707	.
y	y	_RASE_	Root Average Squared Error	0.340168	0.35736	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 25. Fit statistics, ASE 19 branches 45 leafs

Results - Node: ASE DT 19 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.161492	0.180095	.
y	y	_MAX_	Maximum Absolute Error	0.986755	1	.
y	y	_SSE_	Sum of Squared Errors	1077.075	1178.473	.
y	y	_ASE_	Average Squared Error	0.116114	0.126936	.
y	y	_RASE_	Root Average Squared Error	0.340755	0.356281	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 26. Fit statistics, ASE 20 branches 54 leafs

Results - Node: ASE DT 20 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.159552	0.179449	.
y	y	_MAX_	Maximum Absolute Error	0.991379	1	.
y	y	_SSE_	Sum of Squared Errors	1076.875	1201.038	.
y	y	_ASE_	Average Squared Error	0.116093	0.129366	.
y	y	_RASE_	Root Average Squared Error	0.340724	0.359675	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 27. Fit statistics, Misclassification 3 br 28 leafs

Results - Node: Misclass DT 3 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.162139	0.177294	.
y	y	_MAX_	Maximum Absolute Error	0.954831	0.954831	.
y	y	_SSE_	Sum of Squared Errors	1149.27	1240.489	.
y	y	_ASE_	Average Squared Error	0.123897	0.133616	.
y	y	_RASE_	Root Average Squared Error	0.35199	0.365535	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 28. Fit statistics, Misclassification 4 br 31 leafs

Results - Node: Misclass DT 4 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.167745	0.178156	.
y	y	_MAX_	Maximum Absolute Error	0.983696	1	.
y	y	_SSE_	Sum of Squared Errors	1132.356	1196.517	.
y	y	_ASE_	Average Squared Error	0.122074	0.128879	.
y	y	_RASE_	Root Average Squared Error	0.34939	0.358998	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 29. Fit statistics, Misclassification 5 br 33 leafs

Results - Node: Misclass DT 5 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.1768	0.185696	.
y	y	_MAX_	Maximum Absolute Error	0.983696	0.983696	.
y	y	_SSE_	Sum of Squared Errors	1189.558	1242.574	.
y	y	_ASE_	Average Squared Error	0.12824	0.13384	.
y	y	_RASE_	Root Average Squared Error	0.358107	0.365842	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 30. Fit statistics, Misclassification 6 br 33 leafs

Results - Node: Misclass DT 6 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.1768	0.185696	.
y	y	_MAX_	Maximum Absolute Error	0.983696	0.983696	.
y	y	_SSE_	Sum of Squared Errors	1189.558	1242.574	.
y	y	_ASE_	Average Squared Error	0.12824	0.13384	.
y	y	_RASE_	Root Average Squared Error	0.358107	0.365842	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 31. Fit statistics, Misclassification 7 br 33 leafs

Results - Node: Misclass DT 7 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.1768	0.185696	.
y	y	_MAX_	Maximum Absolute Error	0.983696	0.983696	.
y	y	_SSE_	Sum of Squared Errors	1189.558	1242.574	.
y	y	_ASE_	Average Squared Error	0.12824	0.13384	.
y	y	_RASE_	Root Average Squared Error	0.358107	0.365842	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 32. Fit statistics, Misclassification 8 br 33 leafs

Results - Node: Misclass DT 8 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.1768	0.185696	.
y	y	_MAX_	Maximum Absolute Error	0.983696	0.983696	.
y	y	_SSE_	Sum of Squared Errors	1189.558	1242.574	.
y	y	_ASE_	Average Squared Error	0.12824	0.13384	.
y	y	_RASE_	Root Average Squared Error	0.358107	0.365842	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 33. Fit statistics, Misclassification 9 br 33 leafs

Results - Node: Misclass DT 9 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.1768	0.185696	.
y	y	_MAX_	Maximum Absolute Error	0.983696	0.983696	.
y	y	_SSE_	Sum of Squared Errors	1189.558	1242.574	.
y	y	_ASE_	Average Squared Error	0.12824	0.13384	.
y	y	_RASE_	Root Average Squared Error	0.358107	0.365842	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 34. Fit statistics, Misclassification 10 br 33 leafs

Results - Node: Misclass DT 10 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.1768	0.185696	.
y	y	_MAX_	Maximum Absolute Error	0.983696	0.983696	.
y	y	_SSE_	Sum of Squared Errors	1189.558	1242.574	.
y	y	_ASE_	Average Squared Error	0.12824	0.13384	.
y	y	_RASE_	Root Average Squared Error	0.358107	0.365842	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 35. Fit statistics, Misclassification 11 br 35 leafs

Results - Node: Misclass DT 11 Br Diagram: Bank Marketing

File Edit View Window

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation	Test
y	y	_NOBS_	Sum of Frequencies	4638	4642	.
y	y	_MISC_	Misclassification Rate	0.172057	0.181172	.
y	y	_MAX_	Maximum Absolute Error	0.986264	1	.
y	y	_SSE_	Sum of Squared Errors	1187.912	1248.09	.
y	y	_ASE_	Average Squared Error	0.128063	0.134435	.
y	y	_RASE_	Root Average Squared Error	0.357859	0.366653	.
y	y	_DIV_	Divisor for ASE	9276	9284	.
y	y	_DFT_	Total Degrees of Freedom	4638	.	.

Appendix. Picture 36. Dummy Re-coding (replacement) settings

Property	Value
General	
Node ID	Repl
Imported Data	
Exported Data	
Notes	
Train	
Interval Variables	
Replacement Editor	
Default Limits Method	None
Cutoff Values	
Class Variables	
Replacement Editor	
Unknown Levels	Ignore
Score	
Replacement Values	Computed
Hide	No
Report	
Replacement Report	Yes
Status	
Create Time	14/12/23 6:45 AM
Run ID	1d69d9f7-13d6-417b-b552-2
Last Error	
Last Status	Complete
Last Run Time	14/12/23 7:10 AM
Run Duration	0 Hr. 0 Min. 6.51 Sec.
Grid Host	
User-Added Node	No

Appendix. Picture 37. Dummy re-coding, replacement editor

Replacement Editor-WORK.OUTCLASS

Variable	Formatted Value	Replacement Value
contact	cellular	
contact	telephone	
contact	_UNKNOWN_	**_DEFAULT_**
day_of_week	thu	**M/T/W/Th**
day_of_week	wed	**M/T/W/Th**
day_of_week	tue	**M/T/W/Th**
day_of_week	mon	**M/T/W/Th**
day_of_week	fri	**Fr**

Replacement Editor-WORK.OUTCLASS

Variable	Formatted Value	Replacement Value
education	university.degree	**Un/Pr**
education	high.school	**HS**
education	professional.course	**Un/Pr**
education	basic.9y	**Basic**
education	basic.4y	**Basic**
education	basic.6y	**Basic**
education	unknown	**Un/Il**
education	illiterate	**Un/Il**

Appendix. Picture 37. Dummy re-coding, replacement editor (continue)

Replacement Editor-WORK.OUTCLASS

Variable	Formatted Value	Replacement Value
job	admin.	**S/S/A**
job	blue-collar	**B/T/H**
job	technician	**B/T/H**
job	services	**S/S/A**
job	management	**E/S/M**
job	retired	**Ret**
job	student	**S/S/A**
job	self-employed	**E/S/M**
job	unemployed	**Unempl**
job	entrepreneur	**E/S/M**
job	housemaid	**B/T/H**
job	unknown	**Unknown**
job	_UNKNOWN_	**_DEFAULT_**

Replacement Editor-WORK.OUTCLASS

Variable	Formatted Value	Replacement Value
month	may	**Q2**
month	jul	**Q3**
month	aug	**Q3**
month	jun	**Q2**
month	nov	**Q4**
month	apr	**Q2**
month	oct	**Q4**
month	sep	**Q3**
month	mar	**Q1**
month	dec	**Q4**

pdays	999	0
pdays	3	1
pdays	6	1
pdays	4	1
pdays	9	2
pdays	2	1
pdays	10	2
pdays	13	2
pdays	7	1
pdays	12	2
pdays	5	1
pdays	15	3
pdays	11	2
pdays	1	1
pdays	8	2
pdays	0	0
pdays	14	2
pdays	17	3
pdays	22	4
pdays	16	3
pdays	18	3
pdays	21	3
pdays	25	4
pdays	26	4
pdays	_UNKNOWN_	_DEFAULT_

Appendix. Picture 38. Cap and Floor Settings

Property	Value
General	
Node ID	Repl2
Imported Data	
Exported Data	
Notes	
Train	
Interval Variables	
Replacement Editor	
Default Limits Method	Standard Deviations from the
Cutoff Values	
Class Variables	
Replacement Editor	
Unknown Levels	Ignore
Score	
Replacement Values	Computed
Hide	No
Report	
Replacement Report	Yes
Status	
Create Time	14/12/23 7:27 AM
Run ID	d0f54a43-1f2a-4780-be62-9
Last Error	
Last Status	Complete
Last Run Time	14/12/23 7:29 AM
Run Duration	0 Hr. 0 Min. 7.98 Sec.
Grid Host	
User-Added Node	No

Appendix. Picture 39. Transform Variables Settings

Property	Value
General	
Node ID	Trans
Imported Data	...
Exported Data	...
Notes	...
Train	
Variables	...
Formulas	...
Interactions	...
SAS Code	...
Default Methods	
Interval Inputs	None
Interval Targets	None
Class Inputs	None
Class Targets	None
Treat Missing as Level	No
Sample Properties	
Method	First N
Size	Default
Random Seed	12345
Optimal Binning	
Number of Bins	4
Missing Values	Use in Search
Grouping Method	
Cutoff Value	0.1
Group Missing	No
Number of Bins	Variables
Add Minimum Value to Offse	Yes
Offset Value	1
Score	
Use Meta Transformation	Yes
Hide	Yes
Reject	Yes
Report	
Summary Statistics	Yes

Appendix. Picture 40. Full Regression Settings

.. Property	Value
General	
Node ID	Reg
Imported Data	...
Exported Data	...
Notes	...
Train	
Variables	...
Equation	
Main Effects	Yes
Two-Factor Interactions	No
Polynomial Terms	No
Polynomial Degree	2
User Terms	No
Term Editor	...
Class Targets	
Regression Type	Logistic Regression
Link Function	Logit
Model Options	
Suppress Intercept	No
Input Coding	Deviation
Model Selection	
Selection Model	None
Selection Criterion	Default
Use Selection Defaults	Yes
Selection Options	...
Optimization Options	
Technique	Default
Default Optimization	Yes
Max Iterations	0
Max Function Calls	0
Maximum Time	1 Hour
Convergence Criteria	
Uses Defaults	Yes
Options	...
Output Options	
Confidence Limits	No
Save Covariance	No
Covariance	No
Correlation	No
Statistics	No
Suppress Output	No
Details	No
Design Matrix	No
Score	
Excluded Variables	Reject
Status	
Create Time	14/12/23 7:42 AM
Run ID	a080dcbb-ef17-4738-b296-

Appendix. Picture 41. Forward Regression Settings

Property	Value
General	
Node ID	Reg3
Imported Data	...
Exported Data	...
Notes	...
Train	
Variables	...
Equation	
Main Effects	Yes
Two-Factor Interactions	No
Polynomial Terms	No
Polynomial Degree	2
User Terms	No
Term Editor	...
Class Targets	
Regression Type	Logistic Regression
Link Function	Logit
Model Options	
Suppress Intercept	No
Input Coding	Deviation
Model Selection	
Selection Model	Forward
Selection Criterion	Validation Error
Use Selection Defaults	Yes
Selection Options	...
Optimization Options	
Technique	Default
Default Optimization	Yes
Max Iterations	0
Max Function Calls	0
Maximum Time	1 Hour
Convergence Criteria	
Uses Defaults	Yes
Options	...
Output Options	
Confidence Limits	No
Save Covariance	No
Covariance	No
Correlation	No
Statistics	No
Suppress Output	No
Details	No
Design Matrix	No
Score	
Excluded Variables	Reject
Status	
Create Time	14/12/23 7:44 AM
Run ID	34fe687a-4480-4c3d-995c-

Appendix. Picture 42. Stepwise Regression Node Settings

Property	Value
General	
Node ID	Reg4
Imported Data	
Exported Data	
Notes	
Train	
Variables	
Equation	
Main Effects	Yes
Two-Factor Interactions	No
Polynomial Terms	No
Polynomial Degree	2
User Terms	No
Term Editor	
Class Targets	
Regression Type	Logistic Regression
Link Function	Logit
Model Options	
Suppress Intercept	No
Input Coding	Deviation
Model Selection	
Selection Model	Stepwise
Selection Criterion	Validation Error
Use Selection Defaults	Yes
Selection Options	
Optimization Options	
Technique	Default
Default Optimization	Yes
Max Iterations	0
Max Function Calls	0
Maximum Time	1 Hour
Convergence Criteria	
Uses Defaults	Yes
Options	
Output Options	
Confidence Limits	No
Save Covariance	No
Covariance	No
Correlation	No
Statistics	No
Suppress Output	No
Details	No
Design Matrix	No
Score	
Excluded Variables	Reject
Status	
Create Time	14/12/23 7:45 AM
Run ID	c7999085-86de-485a-85ad

Appendix. Picture 43. Backward Regression Node Settings

Property	Value
General	
Node ID	Reg2
Imported Data	
Exported Data	
Notes	
Train	
Variables	
Equation	
Main Effects	Yes
Two-Factor Interactions	No
Polynomial Terms	No
Polynomial Degree	2
User Terms	No
Term Editor	
Class Targets	
Regression Type	Logistic Regression
Link Function	Logit
Model Options	
Suppress Intercept	No
Input Coding	Deviation
Model Selection	
Selection Model	Backward
Selection Criterion	Validation Error
Use Selection Defaults	Yes
Selection Options	
Optimization Options	
Technique	Default
Default Optimization	Yes
Max Iterations	0
Max Function Calls	0
Maximum Time	1 Hour
Convergence Criteria	
Uses Defaults	Yes
Options	
Output Options	
Confidence Limits	No
Save Covariance	No
Covariance	No
Correlation	No
Statistics	No
Suppress Output	No
Details	No
Design Matrix	No
Score	
Excluded Variables	Reject
Status	
Create Time	14/12/23 7:43 AM

Appendix. Picture 44. Neural Network General Settings

Property	Value
General	
Node ID	Neural5
Imported Data	
Exported Data	
Notes	
Train	
Variables	
Continue Training	No
Network	
Optimization	
Initialization Seed	12345
Model Selection Criterion	Average Error
Suppress Output	No
Score	
Hidden Units	No
Residuals	Yes
Standardization	No
Status	
Create Time	07/12/23 9:11 AM
Run ID	fa02b608-d79a-4b71-9b05-d
Last Error	
Last Status	Complete
Last Run Time	13/12/23 2:06 PM
Run Duration	0 Hr. 0 Min. 20.73 Sec.
Grid Host	
User-Added Node	No

Appendix. Picture 45. Model comparison settings

Property	Value
General	
Node ID	MdlComp
Imported Data	
Exported Data	
Notes	
Train	
Variables	
Assessment Reports	
Number of Bins	20
ROC Chart	Yes
Recompute	No
Model Selection	
Selection Data	Default
Selection Statistic	Average Squared Error
HP Selection Statistic	Default
SAS Viya Selection Statistic	
Selection Table	Validation
Selection Depth	10
Score	
Selection Editor	
Report	
Selected Model	
Target	y
Model Node	Neural
Model Description	DP NN ASE 50 H3
Selection Criteria	Valid: Average Squared Error
Status	
Create Time	14/12/23 7:48 AM
Run ID	594c98fa-4c75-4a09-b84f-52
Last Error	
Last Status	Complete
Last Run Time	14/12/23 7:53 PM
Run Duration	0 Hr. 4 Min. 36.07 Sec.
Grid Host	
User-Added Node	No