

```
>Warning # 849 in column 23.  Text: es_41
>The LOCALE subcommand of the SET command has an invalid parameter.  It could
>not be mapped to a valid backend locale.
```

```
GET DATA  /TYPE=TXT
  /FILE="G:\Mi unidad\UNL\UNL 9A Desktop\Inteligencia Artificial\DATASETS\train.csv"
  /ENCODING='Locale'
  /DELCASE=LINE
  /DELIMITERS=","
  /ARRANGEMENT=DELIMITED
  /FIRSTCASE=2
  /IMPORTCASE=ALL
  /VARIABLES=
V1 F5.0
obj_ID A22
alpha F17.0
delta A18
u F8.0
g F8.0
r F8.0
i F8.0
z F8.0
run_ID F4.0
rerun_ID F3.0
cam_col F1.0
field_ID F3.0
spec_obj_ID A22
class A6
redshift F13.0
plate F5.0
MJD F5.0
fiber_ID F3.0.
CACHE.
EXECUTE.
DATASET NAME Conjunto_de_datos1 WINDOW=FRONT.
* Árbol de decisiones.
TREE class [n] BY V1 [s] alpha [s] delta [n] u [s] g [s] r [s] i [s] z [s] redshift [s] MJD [s]
  /TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES SCALE=AUTO
  /DEPCATEGORIES USEVALUES=[VALID]
  /PRINT MODELSUMMARY CLASSIFICATION RISK
  /METHOD TYPE=CHAID
  /GROWTHLIMIT MAXDEPTH=AUTO MINPARENTSIZE=100 MINCHILDSIZE=50
  /VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES
  /CHAID ALPHASPLIT=0.05 ALPHAMERGE=0.05 SPLITMERGED=NO CHISQUARE=PEARSON CONVERGE=0.001 MAXITE
  /COSTS EQUAL
  /MISSING NOMINALMISSING=MISSING.
```

## Árbol de clasificación

### Notas

Resultados creados		14-FEB-2022 18:45:22
Comentarios		
Entrada	Datos	G:\Mi unidad\UNL\UNL 9A Desktop\Inteligencia Artificial\DATASETS\train.csv
	Conjunto de datos activo	Conjunto_de_datos1
	Filtro	<ninguno>
	Peso	<ninguno>
	Dividir archivo	<ninguno>
	Núm. de filas del archivo de trabajo	80000
Tratamiento de los datos perdidos	Definición de perdidos	El tratamiento de los valores perdidos definidos por el usuario de las variables independientes nominales depende del método de crecimiento.
	Casos utilizados	Solamente se utilizarán los casos con valores válidos para la variable dependiente y todas o alguna variable independiente en el cálculo de los estadísticos.
Sintaxis		TREE class [n] BY V1 [s] alpha [s] delta [n] u [s] g [s] r [s] i [s] z [s] redshift [s] MJD [s] /TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES SCALE=AUTO /DEPCATEGORIES USEVALUES= [VALID] /PRINT MODELSUMMARY CLASSIFICATION RISK /METHOD TYPE=CHAID /GROWTHLIMIT MAXDEPTH=AUTO MINPARENTSIZE=100 MINCHILDSIZE=50 /VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES /CHAID ALPHASPLIT=0.05 ALPHAMERGE=0.05 SPLITMERGED=NO CHISQUARE=PEARSON CONVERGE=0.001 MAXITERATIONS=100 ADJUST=BONFERRONI INTERVALS=10 /COSTS EQUAL /MISSING...

### Notas

Recursos	Tiempo de procesador	00:00:10.16
	Tiempo transcurrido	00:00:11.45
Archivos guardados	Archivo de reglas	

[ Conjunto\_de\_datos1 ]

### Advertencia

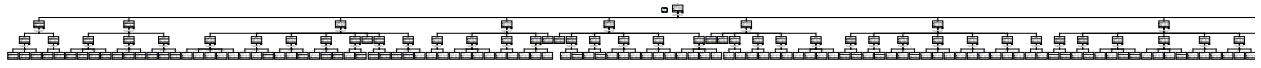
Se han excluido una o más variables independientes del proceso de crecimiento del árbol en uno o más nodos ya que el número de categorías supera el número máximo que permite el método de crecimiento.

No se muestran las tablas de resúmenes de ganancias ya que no se han definido los beneficios.

No se muestran las tablas de ganancias de las categorías objetivo ya que no se han definido las categorías objetivo.

### Resumen del modelo

Especificaciones	Método de crecimiento	CHAID	
	Variable dependiente	class	
	Variables independientes	V1, alpha, delta, u, g, r, i, z, redshift, MJD	
	Validación	Ninguna	
	Máxima profundidad de árbol		3
	Mínimo de casos en un nodo filial		100
	Mínimo de casos en un nodo parental		50
Resultados	Variables independientes incluidas	redshift, i, u, alpha, MJD, z, r, g, V1	
	Número de nodos		195
	Número de nodos terminales		136
	Profundidad		3





### Riesgo

Estimación	Típ. Error
,039	,001

Métodos de crecimiento:

CHAID

Variable dependiente:

class

### Clasificación

Observado	Pronosticado			
	GALAXY	QSO	STAR	Porcentaje correcto
GALAXY	46284	674	616	97,3%
QSO	1423	13782	19	90,5%
STAR	414	0	16788	97,6%
Porcentaje global	60,2%	18,1%	21,8%	96,1%

Métodos de crecimiento: CHAID

Variable dependiente: class

\* Árbol de decisiones.

```
TREE class [n] BY V1 [s] alpha [s] delta [n] u [s] g [s] r [s] i [s] z [s] redshift [s] MJD [s]
/TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES SCALE=AUTO
/DEPCATEGORIES USEVALUES=[VALID]
/PRINT MODELSUMMARY CLASSIFICATION RISK
/METHOD TYPE=CRT MAXSURROGATES=AUTO PRUNE=NONE
/GROWTHLIMIT MAXDEPTH=AUTO MINPARENTSIZE=100 MINCHILDSIZE=50
/VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES
/CRT IMPURITY=GINI MINIMPROVEMENT=0.0001
/COSTS EQUAL
/PRIORS FROMDATA ADJUST=NO
/MISSING NOMINALMISSING=MISSING.
```

## Árbol de clasificación

## Notas

Resultados creados		14-FEB-2022 18:46:31
Comentarios		
Entrada	Datos	G:\Mi unidad\UNL\UNL 9A Desktop\Inteligencia Artificial\DATASETS\train.csv
	Conjunto de datos activo	Conjunto_de_datos1
	Filtro	<ninguno>
	Peso	<ninguno>
	Dividir archivo	<ninguno>
	Núm. de filas del archivo de trabajo	80000
Tratamiento de los datos perdidos	Definición de perdidos	El tratamiento de los valores perdidos definidos por el usuario de las variables independientes nominales depende del método de crecimiento.
	Casos utilizados	Solamente se utilizarán los casos con valores válidos para la variable dependiente y todas o alguna variable independiente en el cálculo de los estadísticos.
Sintaxis		TREE class [n] BY V1 [s] alpha [s] delta [n] u [s] g [s] r [s] i [s] z [s] redshift [s] MJD [s] /TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES SCALE=AUTO /DEPCATEGORIES USEVALUES=[VALID] /PRINT MODELSUMMARY CLASSIFICATION RISK /METHOD TYPE=CRT MAXSURROGATES=AUTO PRUNE=NONE /GROWTHLIMIT MAXDEPTH=AUTO MINPARENTSIZE=100 MINCHILDSIZE=50 /VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES /CRT IMPURITY=GINI MINIMPROVEMENT=0.0001 /COSTS EQUAL /PRIORS FROMDATA ADJUST=NO...
Recursos	Tiempo de procesador	00:00:48.81
	Tiempo transcurrido	00:00:48.91
Archivos guardados	Archivo de reglas	

[Conjunto\_de\_datos1]

### Advertencia

Se han excluido una o más variables independientes del proceso de crecimiento del árbol en uno o más nodos ya que el número de categorías supera el número máximo que permite el método de crecimiento.

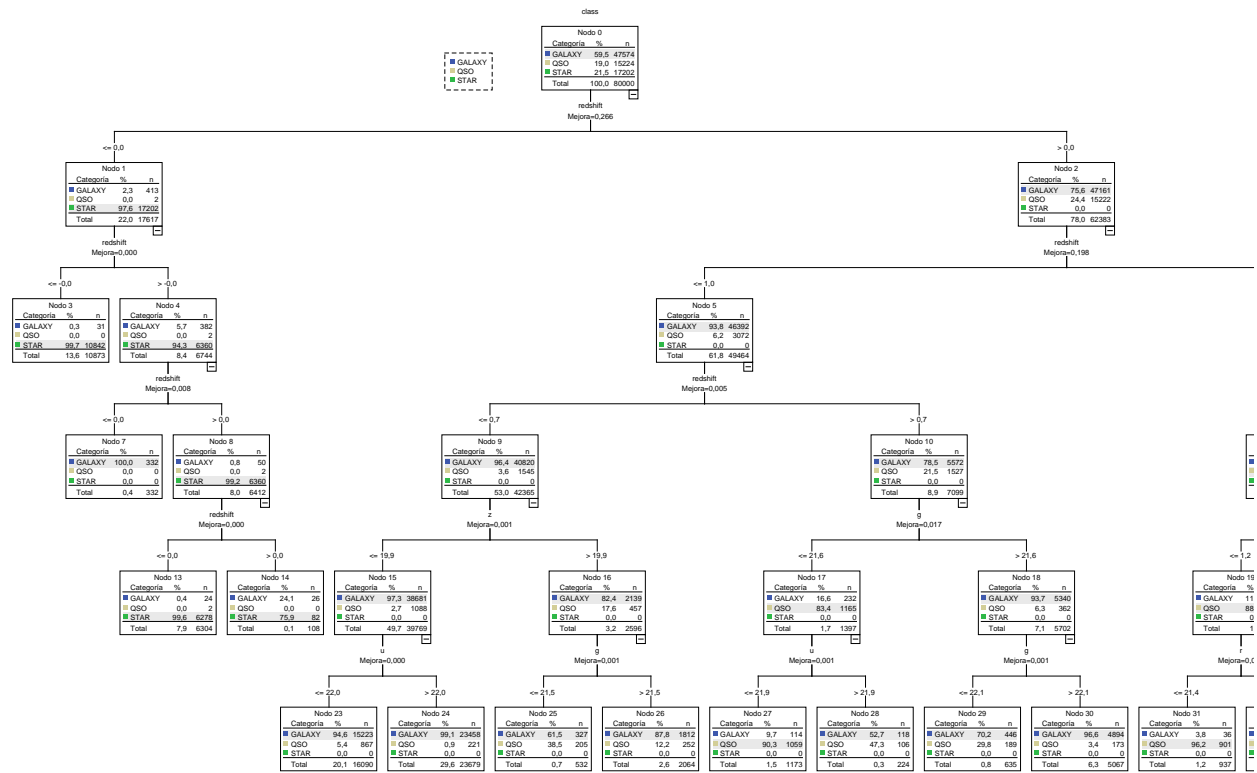
No se muestran las tablas de resúmenes de ganancias ya que no se han definido los beneficios.

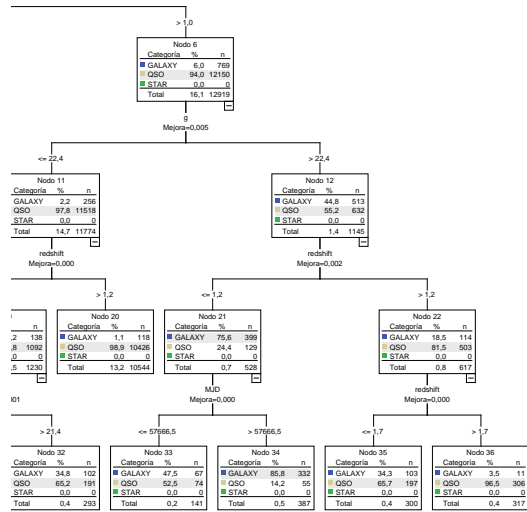
No se muestran las tablas de ganancias de las categorías objetivo ya que no se han definido las categorías objetivo.

### Resumen del modelo

Especificaciones	Método de crecimiento	CRT	
	Variable dependiente	class	
	Variables independientes	V1, alpha, delta, u, g, r, i, z, redshift, MJD	
	Validación	Ninguna	
	Máxima profundidad de árbol		5
	Mínimo de casos en un nodo filial		100
	Mínimo de casos en un nodo parental		50
Resultados	Variables independientes incluidas	redshift, u, g, r, i, z, alpha, MJD, V1	
	Número de nodos		37
	Número de nodos terminales		19
	Profundidad		5







### Riesgo

Estimación	Típ. Error
,034	,001

Métodos de crecimiento:

CRT

Variable dependiente:

class

### Clasificación

Observado	Pronosticado			
	GALAXY	QSO	STAR	Porcentaje correcto
GALAXY	46942	551	81	98,7%
QSO	2068	13154	2	86,4%
STAR	0	0	17202	100,0%
Porcentaje global	61,3%	17,1%	21,6%	96,6%

Métodos de crecimiento: CRT

Variable dependiente: class

```
GET DATA /TYPE=TXT
/FILE="G:\Mi unidad\UNL\UNL 9A Desktop\Inteligencia Artificial\DATASETS\test.csv"
/ENCODING='Locale'
/DELCASE=LINE
/DELIMITERS=","
/ARRANGEMENT=DELIMITED
/FIRSTCASE=2
/IMPORTCASE=ALL
/VARIABLES=
V1 F5.0
obj_ID A22
alpha F17.0
delta F18.0
u F8.0
g F8.0
r F8.0
i F8.0
z F8.0
run_ID F4.0
rerun_ID F3.0
cam_col F1.0
field_ID F3.0
spec_obj_ID A22
class A6
redshift F13.0
```

```

plate F5.0
MJD F5.0
fiber_ID F3.0.
CACHE.
EXECUTE.
DATASET NAME Conjunto_de_datos2 WINDOW=FRONT.
* Árbol de decisiones.
TREE class [n] BY V1 [s] alpha [s] delta [s] u [s] g [s] r [s] i [s] z [s] redshift [s] MJD [s]
/TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES SCALE=AUTO
/DEPCATEGORIES USEVALUES=[VALID]
/PRINT MODELSUMMARY CLASSIFICATION RISK
/METHOD TYPE=CHAID
/GROWTHLIMIT MAXDEPTH=AUTO MINPARENTSIZE=100 MINCHILDSIZE=50
/VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES
/CHAID ALPHASPLIT=0.05 ALPHAMERGE=0.05 SPLITMERGED=NO CHISQUARE=PEARSON CONVERGE=0.001 MAXITE

```

## Árbol de clasificación

### Notas

Resultados creados		14-FEB-2022 18:49:30
Comentarios		
Entrada	Datos	G:\Mi unidad\UNL\UNL 9A Desktop\Inteligencia Artificial\DATASETS\test.csv
	Conjunto de datos activo	Conjunto_de_datos2
	Filtro	<ninguno>
	Peso	<ninguno>
	Dividir archivo	<ninguno>
	Núm. de filas del archivo de trabajo	20000
Tratamiento de los datos perdidos	Definición de perdidos	El tratamiento de los valores perdidos definidos por el usuario de las variables independientes nominales depende del método de crecimiento.
	Casos utilizados	Solamente se utilizarán los casos con valores válidos para la variable dependiente y todas o alguna variable independiente en el cálculo de los estadísticos.

### Notas

<p>Sintaxis</p>	<pre> TREE class [n] BY V1 [s] alpha [s] delta [s] u [s] g [s] r [s] i [s] z [s] redshift [s] MJD [s]   /TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES SCALE=AUTO   /DEPCATEGORIES USEVALUES= [VALID]   /PRINT MODELSUMMARY CLASSIFICATION RISK   /METHOD TYPE=CHAID   /GROWTHLIMIT MAXDEPTH=AUTO MINPARENTSIZE=100 MINCHILDSIZE=50   /VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES   /CHAID ALPHASPLIT=0.05 ALPHAMERGE=0.05 SPLITMERGED=NO CHISQUARE=PEARSON CONVERGE=0.001 MAXITERATIONS=100 ADJUST=BONFERRONI ... </pre>
<p>Recursos</p>	<p>Tiempo de procesador 00:00:01.36</p> <p>Tiempo transcurrido 00:00:01.38</p>
<p>Archivos guardados</p>	<p>Archivo de reglas</p>

[Conjunto\_de\_datos2]

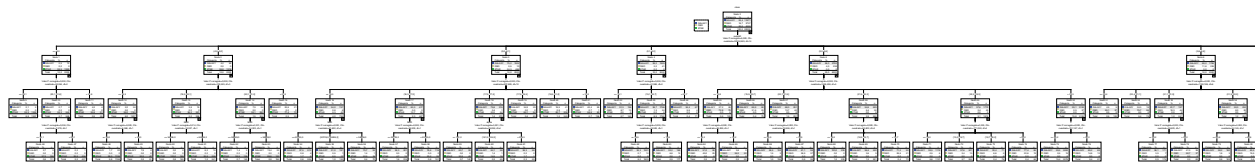
### Advertencia

No se muestran las tablas de resúmenes de ganancias ya que no se han definido los beneficios.

No se muestran las tablas de ganancias de las categorías objetivo ya que no se han definido las categorías objetivo.

### Resumen del modelo

Especificaciones	Método de crecimiento	CHAID	
	Variable dependiente	class	
	Variables independientes	V1, alpha, delta, u, g, r, i, z, redshift, MJD	
	Validación	Ninguna	
	Máxima profundidad de árbol		3
	Mínimo de casos en un nodo filial		100
Resultados	Mínimo de casos en un nodo parental		50
	Variables independientes incluidas	redshift, r, i, u, MJD, z, alpha, g	
	Número de nodos		91
	Número de nodos terminales		62
	Profundidad		3







### Riesgo

Estimación	Típ. Error
,047	,001

Métodos de crecimiento:

CHAID

Variable dependiente:

class

### Clasificación

Observado	Pronosticado			
	GALAXY	QSO	STAR	Porcentaje correcto
GALAXY	11479	173	219	96,7%
QSO	399	3335	3	89,2%
STAR	146	0	4246	96,7%
Porcentaje global	60,1%	17,5%	22,3%	95,3%

Métodos de crecimiento: CHAID

Variable dependiente: class

\* Árbol de decisiones.

```
TREE class [n] BY V1 [s] alpha [s] delta [s] u [s] g [s] r [s] i [s] z [s] redshift [s] MJD [s]
/TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES SCALE=AUTO
/DEPCATEGORIES USEVALUES=[VALID]
/PRINT MODELSUMMARY CLASSIFICATION RISK
/METHOD TYPE=CRT MAXSURROGATES=AUTO PRUNE=NONE
/GROWTHLIMIT MAXDEPTH=AUTO MINPARENTSIZE=100 MINCHILDSIZE=50
/VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES
/CRT IMPURITY=GINI MINIMPROVEMENT=0.0001
/COSTS EQUAL
/PRIORS FROMDATA ADJUST=NO.
```

## Árbol de clasificación

## Notas

Resultados creados		14-FEB-2022 18:50:01
Comentarios		
Entrada	Datos	G:\Mi unidad\UNL\UNL 9A Desktop\Inteligencia Artificial\DATASETS\test.csv
	Conjunto de datos activo	Conjunto_de_datos2
	Filtro	<ninguno>
	Peso	<ninguno>
	Dividir archivo	<ninguno>
	Núm. de filas del archivo de trabajo	20000
Tratamiento de los datos perdidos	Definición de perdidos	El tratamiento de los valores perdidos definidos por el usuario de las variables independientes nominales depende del método de crecimiento.
	Casos utilizados	Solamente se utilizarán los casos con valores válidos para la variable dependiente y todas o alguna variable independiente en el cálculo de los estadísticos.
Sintaxis		TREE class [n] BY V1 [s] alpha [s] delta [s] u [s] g [s] r [s] i [s] z [s] redshift [s] MJD [s] /TREE DISPLAY=TOPDOWN NODES=STATISTICS BRANCHSTATISTICS=YES NODEDEFS=YES SCALE=AUTO /DEPCATEGORIES USEVALUES=[VALID] /PRINT MODELSUMMARY CLASSIFICATION RISK /METHOD TYPE=CRT MAXSURROGATES=AUTO PRUNE=NONE /GROWTHLIMIT MAXDEPTH=AUTO MINPARENTSIZE=100 MINCHILDSIZE=50 /VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES /CRT IMPURITY=GINI MINIMPROVEMENT=0.0001 /COSTS EQUAL /PRIORS FROMDATA...
Recursos	Tiempo de procesador	00:00:01.64
	Tiempo transcurrido	00:00:01.87
Archivos guardados	Archivo de reglas	

[Conjunto\_de\_datos2]

### Advertencia

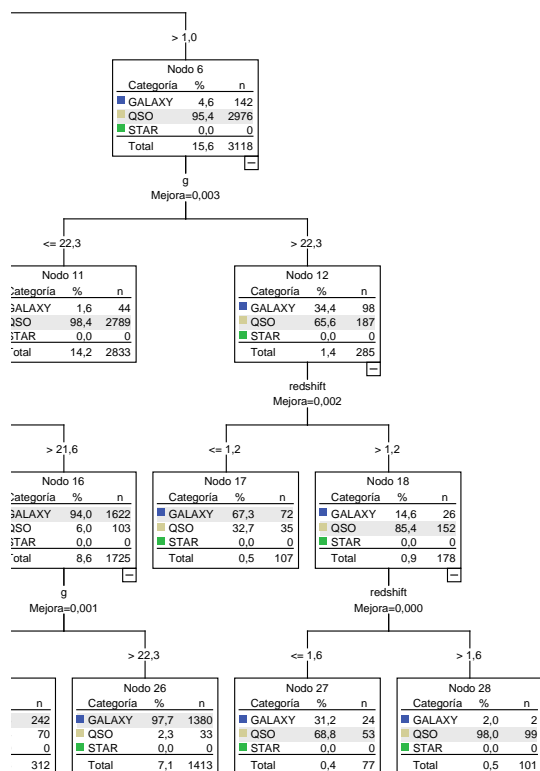
No se muestran las tablas de resúmenes de ganancias ya que no se han definido los beneficios.

No se muestran las tablas de ganancias de las categorías objetivo ya que no se han definido las categorías objetivo.

### Resumen del modelo

Especificaciones	Método de crecimiento	CRT	
	Variable dependiente	class	
	Variables independientes	V1, alpha, delta, u, g, r, i, z, redshift, MJD	
	Validación	Ninguna	
	Máxima profundidad de árbol		5
Resultados	Mínimo de casos en un nodo filial		100
	Mínimo de casos en un nodo parental		50
	Variables independientes incluidas	redshift, u, g, delta, r, i, alpha, MJD, V1, z	
	Número de nodos		29
	Número de nodos terminales		15
	Profundidad		5





### Riesgo

Estimación	Típ. Error
,032	,001

Métodos de crecimiento:

CRT

Variable dependiente:

class

### Clasificación

Observado	Pronosticado			
	GALAXY	QSO	STAR	Porcentaje correcto
GALAXY	11720	132	19	98,7%
QSO	483	3253	1	87,0%
STAR	0	0	4392	100,0%
Porcentaje global	61,0%	16,9%	22,1%	96,8%

Métodos de crecimiento: CRT

Variable dependiente: class