DATAVISION

Healthcare Dataset

Benítez, Figueroa, Gavenda e Ibañez

ÍNDICE

Problema

- Tiempos de espera prolongados para obtener resultados
- Demoras en el diagnóstico y tratamiento debido a la alta demanda
- Recursos hospitalarios (laboratorios, médicos, equipamiento) limitados.
- Dificultad para priorizar pacientes según nivel de urgencia

Necesidad

Surge la necesidad de contar con herramientas que permitan anticipar el resultado probable de un test médico, incluso antes de que esté disponible, utilizando variables clínicas y administrativas recogidas durante la internación.

Modelo predictivo capaz de anticipar el resultado de un test médico realizado a un paciente en su ingreso al hospital

Beneficios del modelo

- Anticipación del diagnóstico
- Priorización inteligente
- Eficiencia operativa.
- Reducción de costos

EDA

Variable Objetivo

Test Results

Describe el resultado del test realizado en el ingreso del paciente.

Los resultados pueden ser:

Normal

Abnormal

Inconcluso

Variables Independientes

Medical condition

specifies the diagnosis associated with the patient (Diabetes, Hypertention, etc

Insurance Provider

Patient's insurance provider including "Aetna," "Blue Cross," "Cigna," "UnitedHealthcare," and "Medicare."

Blood Type

The patient's blood type. ("A+", "O−", etc.)

Admission Type

Specifies the type of admission, can be "Emergency," "Elective," or "Urgent".

Gender

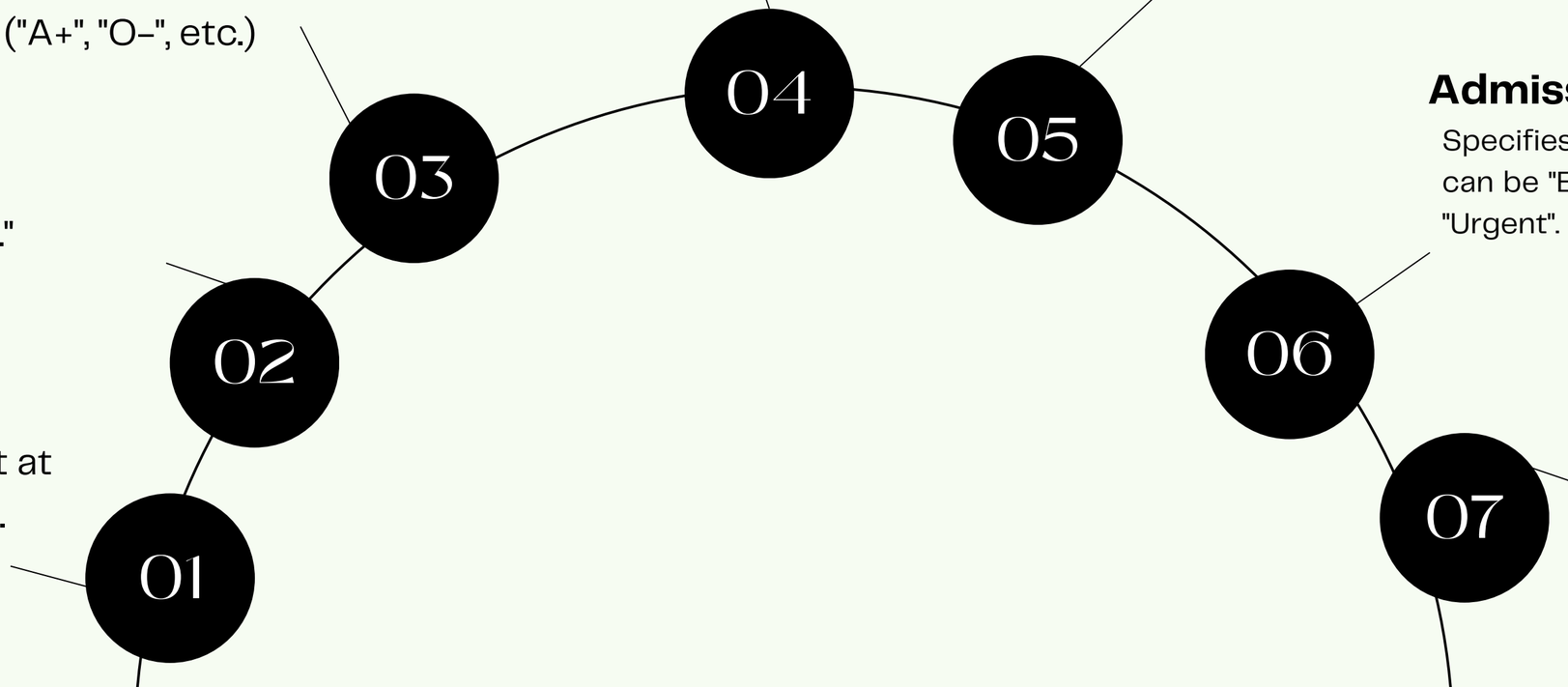
"Male" or "Female."

Medication

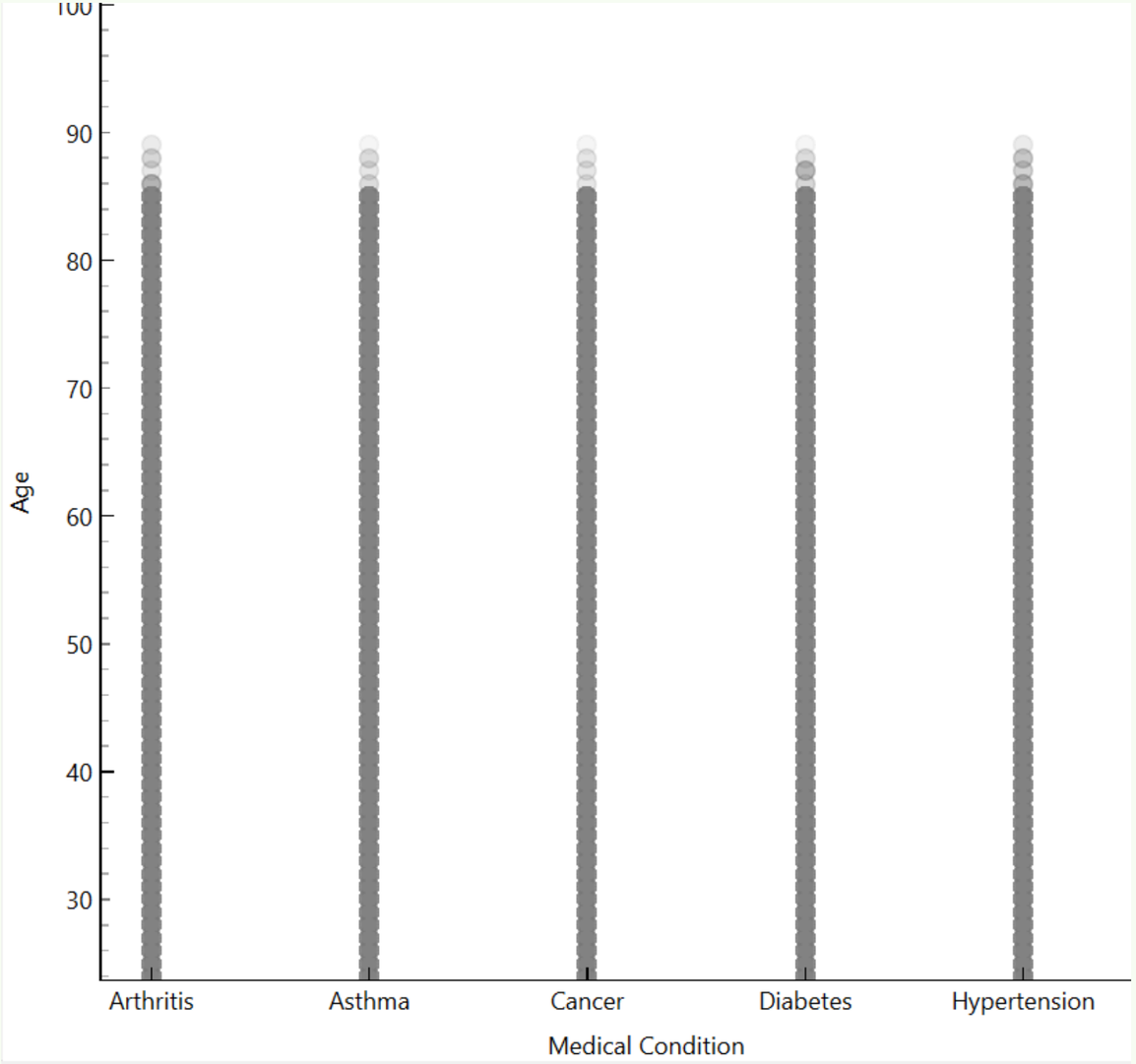
a medication prescribed or administered to the patient "Aspirin," "Ibuprofen," "Penicillin"

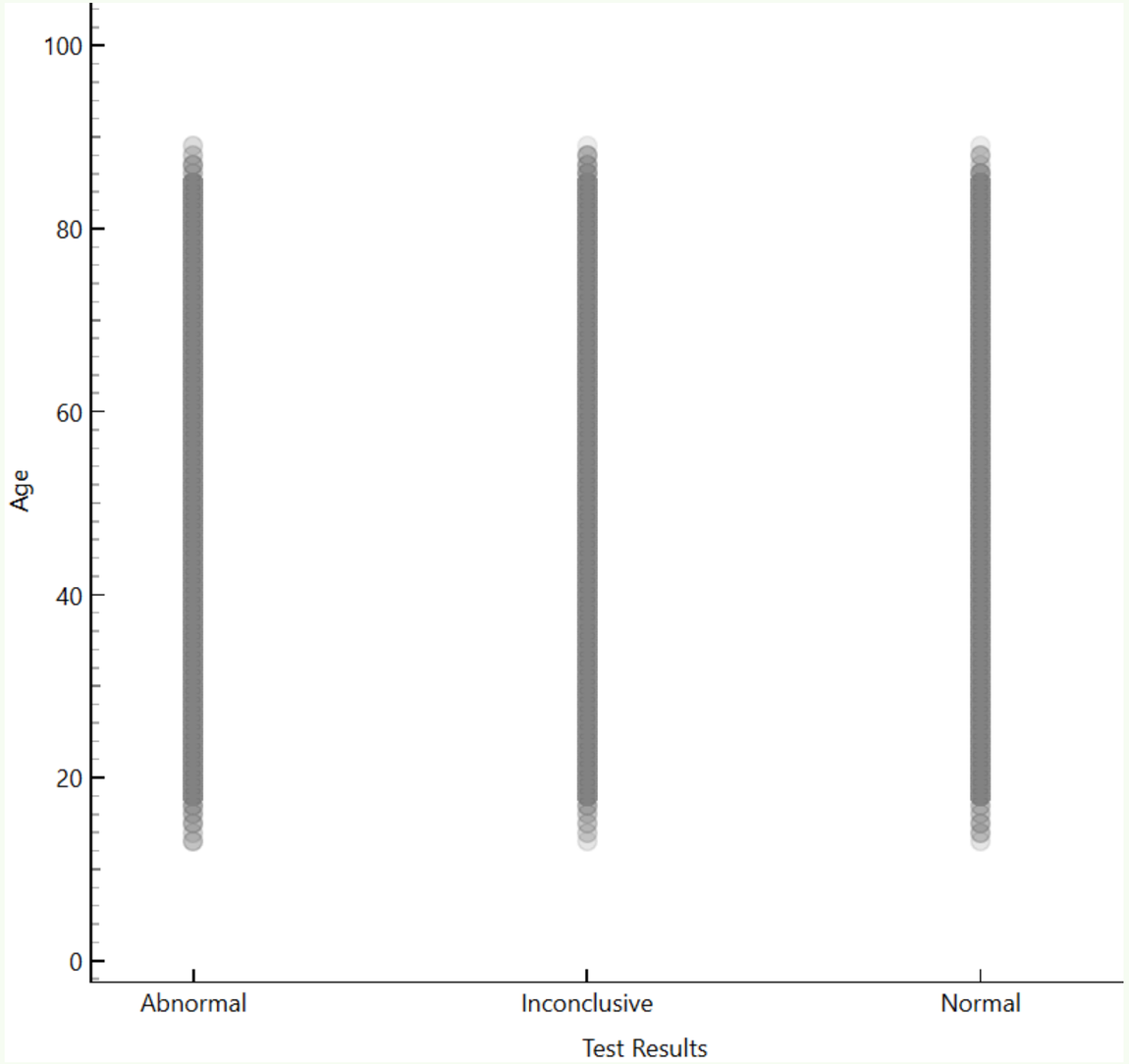
Age

The age of the patient at the time of admission.



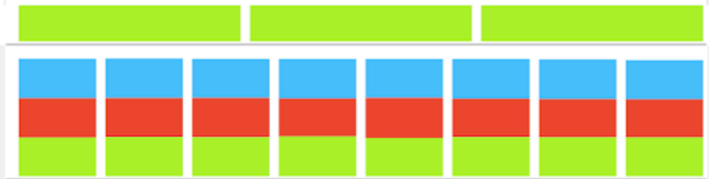
Scatter Plott





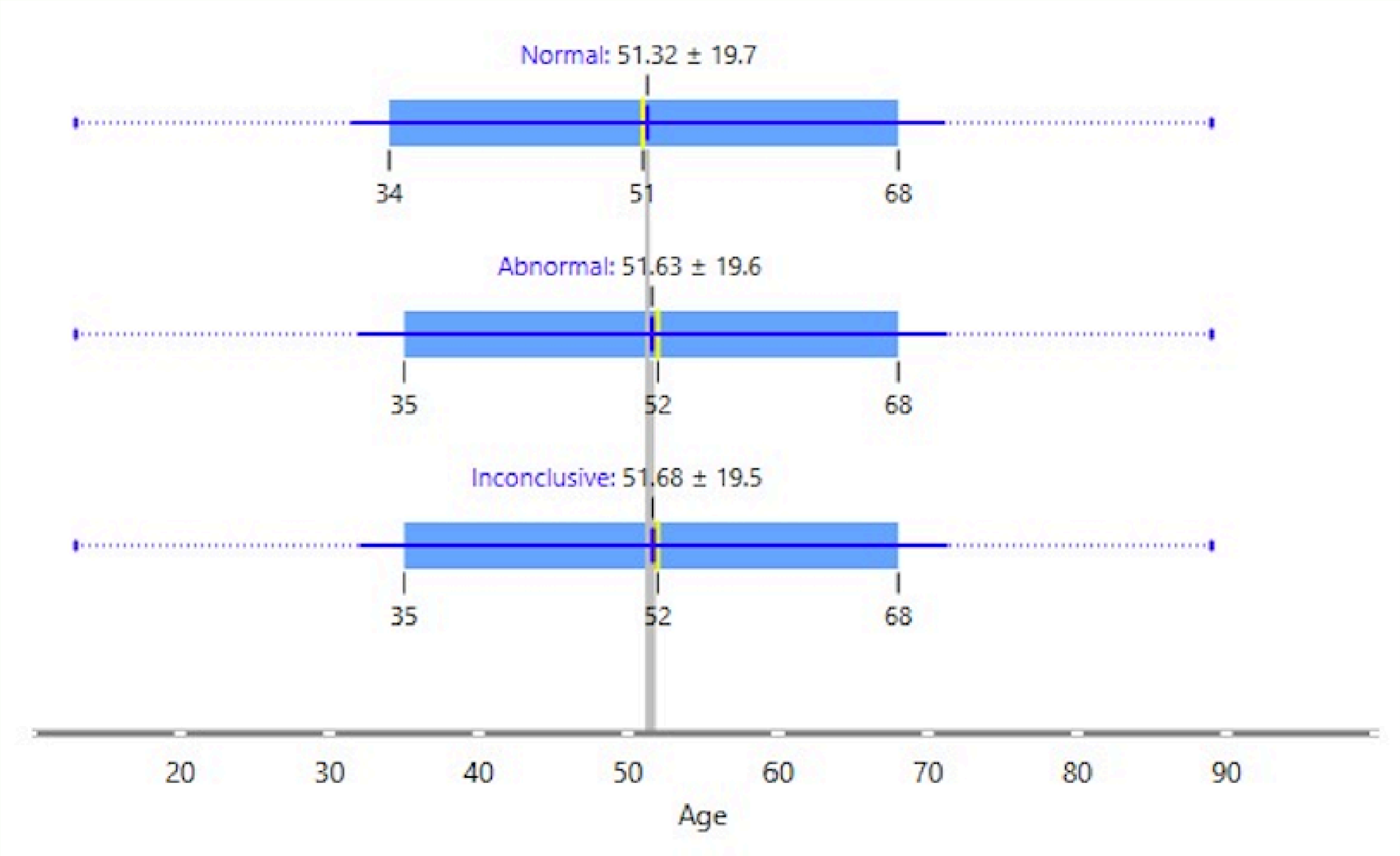
Feature Statistics

	Name	Distribution	Mean	Mode	Median	Dispersion	Min.	Max.	Missing
N	Age		51.54	38	52	0.38	13	89	0 (0 %)

	Name	Distribution	Mean	Mode	Median	Dispersion	Min.	Max.	Missing
C	Blood Type			A-		2.08			0 (0 %)
C	Gender			Male		0.693			0 (0 %)
C	Insurance Provider			Cigna		1.61			0 (0 %)
C	Medical Condition			Arthritis		1.79			0 (0 %)
C	Medication			Lipitor		1.61			0 (0 %)
C	Test Results			Abnormal		1.1			0 (0 %)

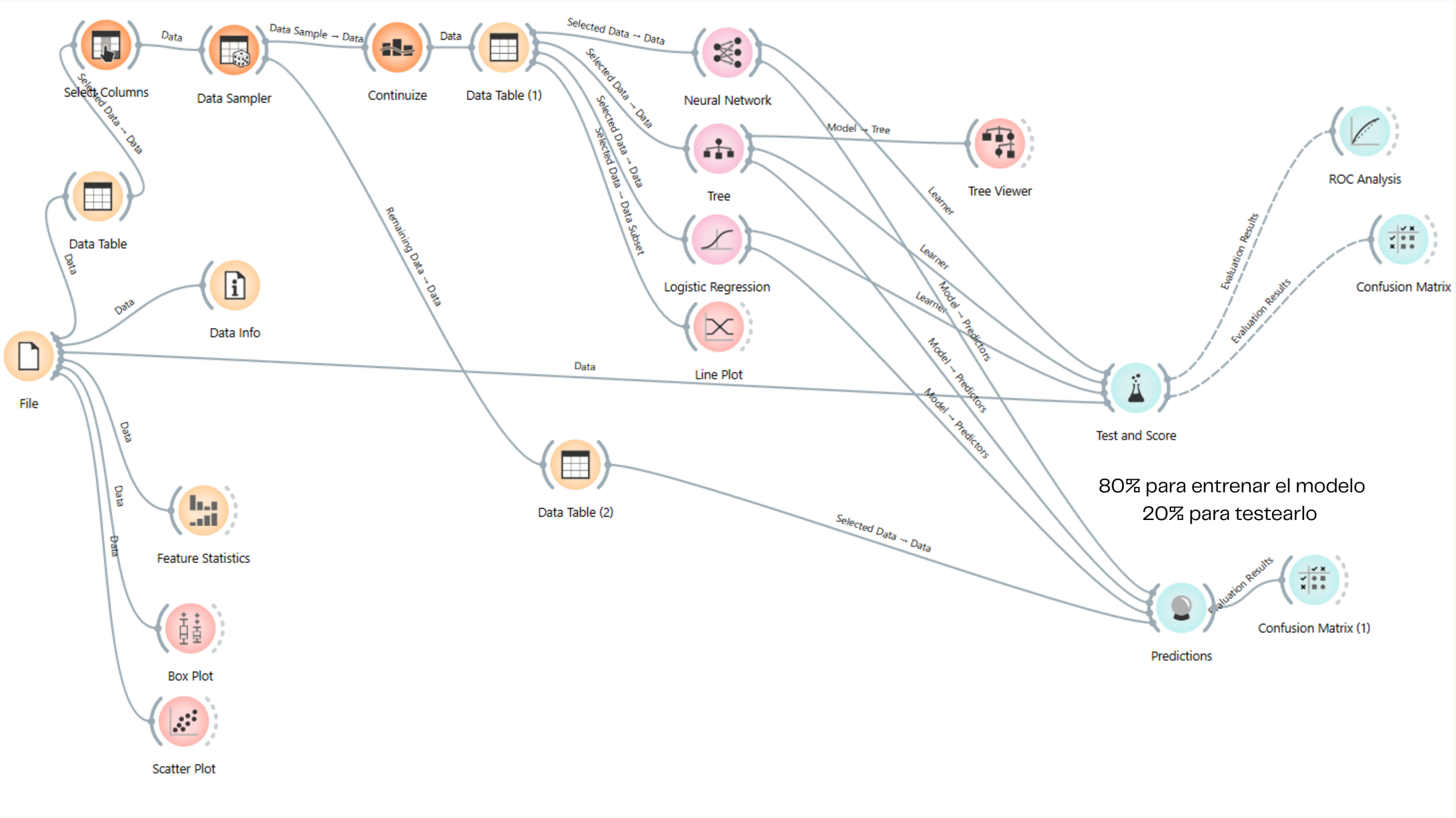
Abnormal Inconclusive Normal

Box Plot



La edad no influye en el resultado del test

PIPELINE


Select Columns
Data Sampler
Continuize
Data Table (1)
Neural Network
Tree
Tree Viewer
Logistic Regression
Line Plot
Data Table
Data Info
File
Data Table (2)
Feature Statistics
Box Plot
Scatter Plot
Test and Score
ROC Analysis
Confusion Matrix
Predictions
Confusion Matrix (1)
Data
Data Sample → Data
Selected Data → Data
Remaining Data → Data
Selected Data → Data Subset
Model → Tree
Learner
Model → Predictors
Evaluation Results
80% para entrenar el modelo
20% para testearlo

Análisis de Métricas

F1

Representa el balance entre la precisión y el recall

RECALL

Capacidad del modelo para detectar correctamente los casos positivos reales (en este caso, pacientes con resultados “Abnormal”).

PRECISIÓN

De las veces que el modelo predijo positiva (“Abnormal”) cuantas veces acertó

Análisis de Modelos

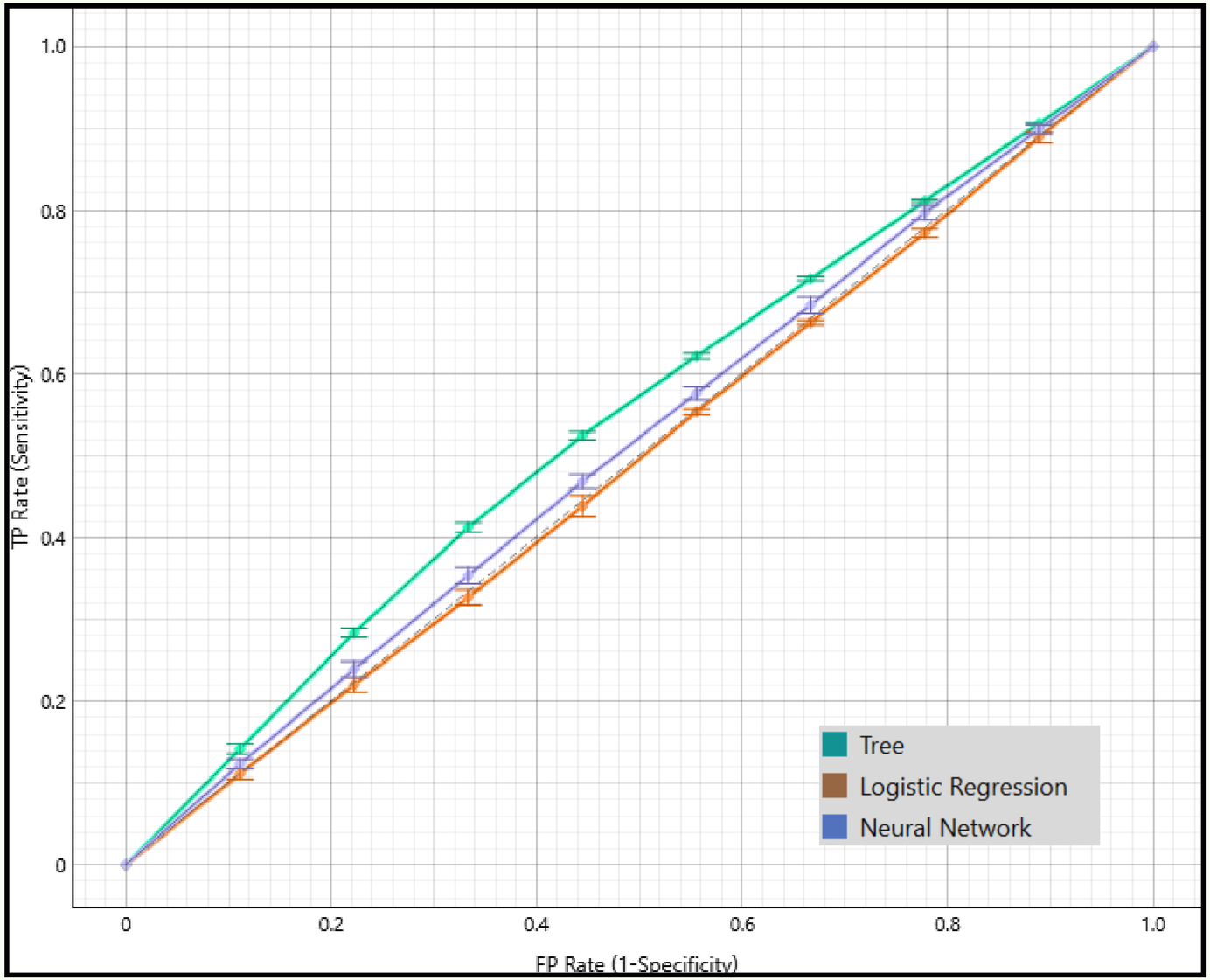
Model	AUC	CA	F1	Prec	Recall	MCC
Tree	0.548	0.386	0.386	0.386	0.386	0.080
Neural Network	0.514	0.346	0.345	0.345	0.346	0.018
Logistic Regression	0.499	0.334	0.330	0.334	0.334	0.000

→ Mejor rendimiento

→ Leve mejora del rendimiento

→ Bajo rendimiento

Curvas ROC



Ninguno de los modelos logra una discriminación clara entre los casos positivos y negativos

MATRICES DE CONFUSIÓN

Tree

Predicted

Actual	Abnormal	Inconclusive	Normal	Σ
Abnormal	1550	1126	1073	3749
Inconclusive	1286	1292	1065	3643
Normal	1250	1160	1298	3708
Σ	4086	3578	3436	11100

Regresión Logística

Predicted

Actual	Abnormal	Inconclusive	Normal	Σ
Abnormal	1442	1025	1282	3749
Inconclusive	1415	1004	1224	3643
Normal	1518	992	1198	3708
Σ	4375	3021	3704	11100

Los modelos cometen una alta proporción de errores en la clasificación de los resultados del test médico

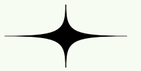
Red Neuronal

Predicted

Actual	Abnormal	Inconclusive	Normal	Σ
Abnormal	1014	1497	1238	3749
Inconclusive	916	1504	1223	3643
Normal	910	1482	1316	3708
Σ	2840	4483	3777	11100

	Tree	error	Logistic Regression	error	Neural Network	error	Test Results
1	1.00 : 0.00 : 0.00 → Abnormal	1.000	0.33 : 0.33 : 0.34 → Normal	0.657	0.30 : 0.54 : 0.16 → Inconclusive	0.835	Normal
2	1.00 : 0.00 : 0.00 → Abnormal	0.000	0.34 : 0.32 : 0.34 → Abnormal	0.660	0.40 : 0.30 : 0.30 → Abnormal	0.596	Abnormal
3	1.00 : 0.00 : 0.00 → Abnormal	1.000	0.32 : 0.34 : 0.34 → Inconclusive	0.665	0.30 : 0.32 : 0.39 → Normal	0.614	Normal
4	1.00 : 0.00 : 0.00 → Abnormal	1.000	0.34 : 0.32 : 0.34 → Normal	0.682	0.34 : 0.28 : 0.39 → Normal	0.723	Inconclusive
5	1.00 : 0.00 : 0.00 → Abnormal	0.000	0.35 : 0.31 : 0.34 → Abnormal	0.654	0.48 : 0.21 : 0.31 → Abnormal	0.519	Abnormal
6	1.00 : 0.00 : 0.00 → Abnormal	1.000	0.36 : 0.32 : 0.33 → Abnormal	0.684	0.24 : 0.25 : 0.50 → Normal	0.747	Inconclusive
7	1.00 : 0.00 : 0.00 → Abnormal	1.000	0.34 : 0.33 : 0.33 → Abnormal	0.668	0.19 : 0.55 : 0.26 → Inconclusive	0.738	Normal
8	1.00 : 0.00 : 0.00 → Abnormal	1.000	0.33 : 0.34 : 0.33 → Inconclusive	0.661	0.42 : 0.30 : 0.28 → Abnormal	0.698	Inconclusive
9	1.00 : 0.00 : 0.00 → Abnormal	0.000	0.35 : 0.32 : 0.33 → Abnormal	0.654	0.27 : 0.40 : 0.33 → Inconclusive	0.727	Abnormal
10	1.00 : 0.00 : 0.00 → Abnormal	1.000	0.34 : 0.34 : 0.32 → Abnormal	0.677	0.23 : 0.44 : 0.33 → Inconclusive	0.668	Normal
11	1.00 : 0.00 : 0.00 → Abnormal	1.000	0.34 : 0.33 : 0.34 → Normal	0.663	0.36 : 0.37 : 0.27 → Inconclusive	0.733	Normal
12	1.00 : 0.00 : 0.00 → Abnormal	0.000	0.33 : 0.32 : 0.35 → Normal	0.667	0.43 : 0.18 : 0.40 → Abnormal	0.573	Abnormal
13	1.00 : 0.00 : 0.00 → Abnormal	1.000	0.32 : 0.33 : 0.35 → Normal	0.667	0.24 : 0.26 : 0.50 → Normal	0.739	Inconclusive
14	1.00 : 0.00 : 0.00 → Abnormal	1.000	0.36 : 0.33 : 0.31 → Abnormal	0.686	0.34 : 0.24 : 0.42 → Normal	0.580	Normal
15	1.00 : 0.00 : 0.00 → Abnormal	1.000	0.33 : 0.33 : 0.34 → Normal	0.668	0.19 : 0.33 : 0.48 → Normal	0.675	Inconclusive
16	1.00 : 0.00 : 0.00 → Abnormal	1.000	0.34 : 0.32 : 0.34 → Normal	0.683	0.59 : 0.13 : 0.28 → Abnormal	0.865	Inconclusive
17	1.00 : 0.00 : 0.00 → Abnormal	0.000	0.35 : 0.33 : 0.32 → Abnormal	0.654	0.23 : 0.43 : 0.35 → Inconclusive	0.774	Abnormal

	Tree	error	Logistic Regression	error	Neural Network	error	Test Results
1136	0.00 : 0.00 : 1.00 → Normal	0.000	0.35 : 0.32 : 0.33 → Abnormal	0.669	0.22 : 0.43 : 0.35 → Inconclusive	0.646	Normal
1137	0.00 : 0.00 : 1.00 → Normal	0.000	0.35 : 0.33 : 0.32 → Abnormal	0.677	0.30 : 0.45 : 0.25 → Inconclusive	0.751	Normal
1138	0.00 : 0.00 : 1.00 → Normal	0.000	0.33 : 0.33 : 0.33 → Abnormal	0.667	0.13 : 0.36 : 0.51 → Normal	0.488	Normal
1139	0.00 : 0.00 : 1.00 → Normal	1.000	0.33 : 0.32 : 0.35 → Normal	0.680	0.30 : 0.43 : 0.28 → Inconclusive	0.573	Inconclusive
1140	0.00 : 0.00 : 1.00 → Normal	0.000	0.34 : 0.32 : 0.34 → Normal	0.662	0.33 : 0.49 : 0.18 → Inconclusive	0.822	Normal
1141	0.00 : 0.00 : 1.00 → Normal	0.000	0.32 : 0.33 : 0.35 → Normal	0.653	0.26 : 0.41 : 0.33 → Inconclusive	0.666	Normal
1142	0.00 : 0.00 : 1.00 → Normal	0.000	0.33 : 0.34 : 0.33 → Inconclusive	0.666	0.21 : 0.41 : 0.39 → Inconclusive	0.613	Normal
1143	0.00 : 0.00 : 1.00 → Normal	0.000	0.32 : 0.34 : 0.34 → Inconclusive	0.663	0.27 : 0.52 : 0.21 → Inconclusive	0.792	Normal
1144	0.00 : 0.00 : 1.00 → Normal	1.000	0.32 : 0.34 : 0.34 → Normal	0.677	0.28 : 0.34 : 0.38 → Normal	0.723	Abnormal
1145	0.00 : 0.00 : 1.00 → Normal	1.000	0.33 : 0.32 : 0.34 → Normal	0.677	0.16 : 0.16 : 0.68 → Normal	0.836	Inconclusive
1146	0.00 : 0.00 : 1.00 → Normal	1.000	0.34 : 0.33 : 0.33 → Abnormal	0.670	0.34 : 0.25 : 0.41 → Normal	0.755	Inconclusive
1147	0.00 : 0.00 : 1.00 → Normal	1.000	0.34 : 0.33 : 0.33 → Abnormal	0.660	0.33 : 0.55 : 0.13 → Inconclusive	0.673	Abnormal
1148	0.00 : 0.00 : 1.00 → Normal	0.000	0.33 : 0.34 : 0.34 → Inconclusive	0.664	0.36 : 0.51 : 0.12 → Inconclusive	0.876	Normal
1149	0.00 : 0.00 : 1.00 → Normal	0.000	0.36 : 0.33 : 0.31 → Abnormal	0.692	0.35 : 0.25 : 0.40 → Normal	0.600	Normal
1150	0.00 : 0.00 : 1.00 → Normal	0.000	0.34 : 0.33 : 0.33 → Abnormal	0.667	0.18 : 0.36 : 0.46 → Normal	0.536	Normal
1151	0.00 : 0.00 : 1.00 → Normal	1.000	0.34 : 0.33 : 0.33 → Abnormal	0.663	0.33 : 0.27 : 0.40 → Normal	0.671	Abnormal
1152	0.00 : 0.00 : 1.00 → Normal	0.000	0.33 : 0.33 : 0.34 → Normal	0.662	0.40 : 0.25 : 0.35 → Abnormal	0.647	Normal



¡Muchas Gracias!