

Modelos para la gestión de pacientes con problemas derivados de la diabetes


DATA ANALYTICS

Cristina Martín Bris

Natalia Medina Rivas



Problema

- La diabetes es una enfermedad crónica que aparece cuando el páncreas no produce insulina suficiente o cuando el organismo no utiliza eficazmente la insulina que produce.
 - Existen tres tipos de diabetes:
 - Diabetes tipo 1.
 - Diabetes tipo 2
 - Diabetes gestacional
- 

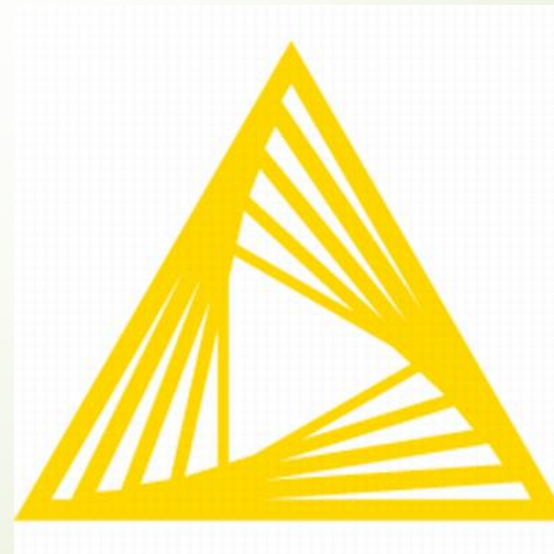


Objetivo de negocio

Objetivo	Atributos
Predecir la recaída de pacientes enfermos de diabetes	Datos aportados por parte del cliente

Análisis inicial y pre-procesado de los datos

➤ Herramientas utilizadas:





Limpieza datos

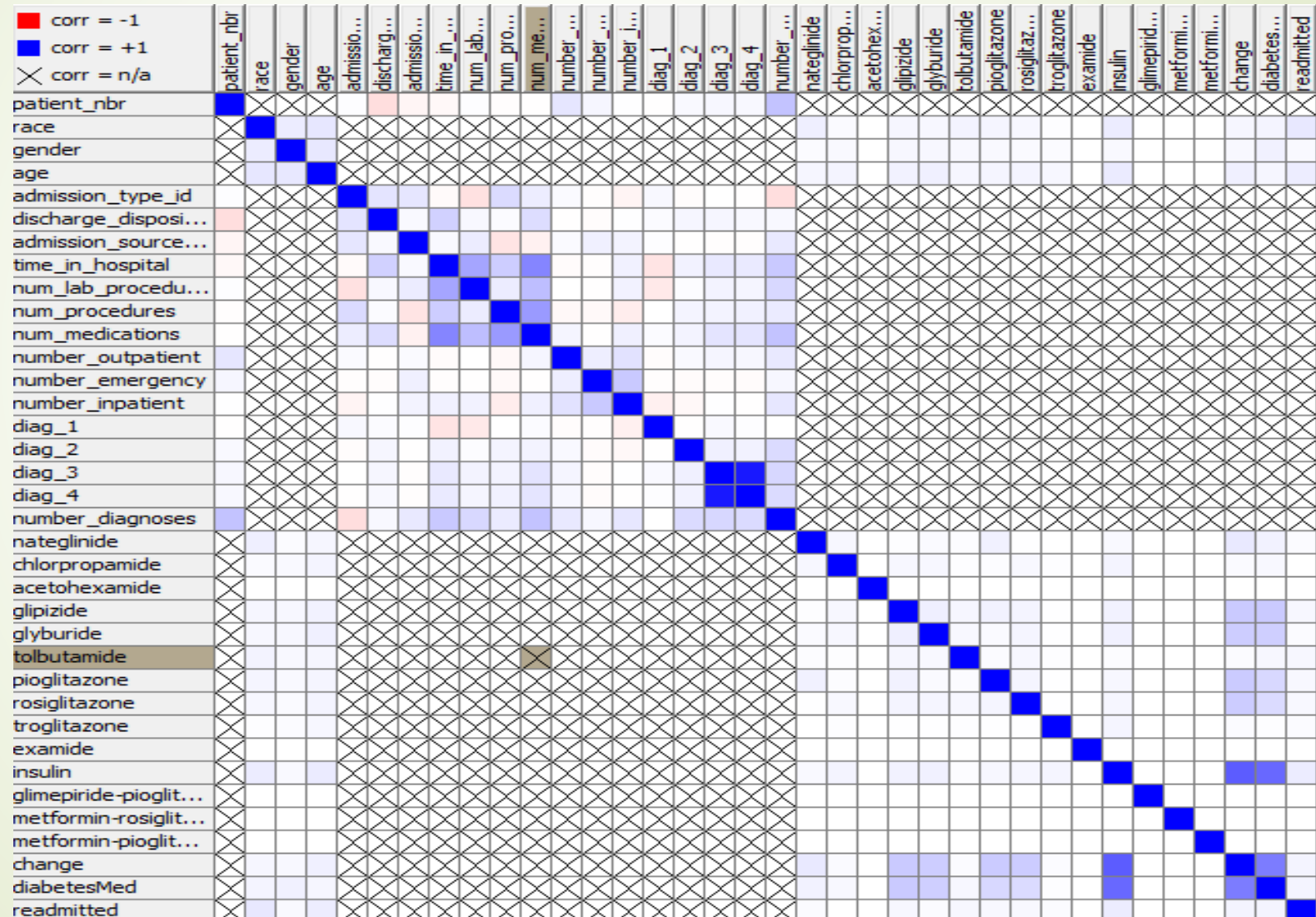
- Búsqueda de duplicados
- Búsqueda de nulos
- Selección de features más significativas

Limpieza datos

- No existen duplicados
- Existen 6 variables con valores nulos.

Variable	Número de nulos
Race	134
Gender	1
Admission_type_id	283
Discharge_disposition_id	176
Admission_source_id	335
Tobultamide	4012

Correlation test





Variables descartadas

- X1
- Patient_nbr
- Tobultamide
- Acetoexamide
- Examide
- Glimepiride-pioglitazone
- Metformin-rosiglitazone
- Metformin-pioglitazone
- Diag_3



Modelos entrenados

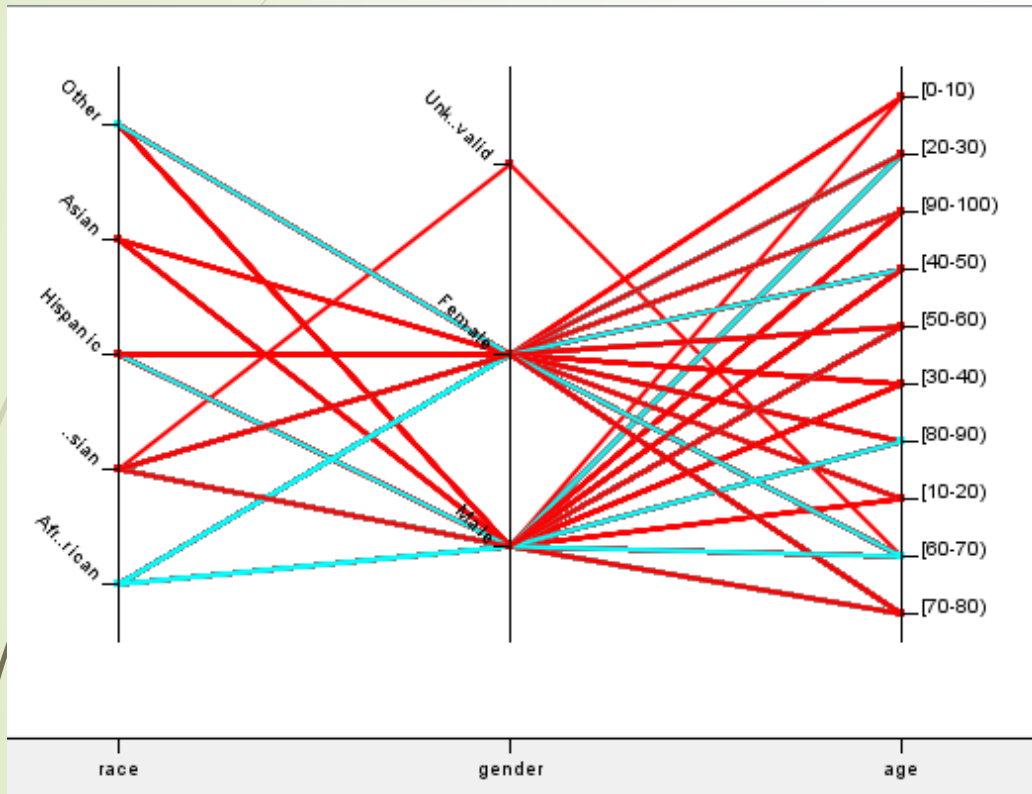
Modelo	Exactitud	Falsos positivos
Bayes	61,22%	842
Árbol de decisión	56,76%	1060
KNN	54%	126
Red neuronal	51,8%	1097
Random Forest	62,56%	665



Modelos entrenados

Modelo	Exactitud	Falsos positivos
Bayes	61,22%	842
Árbol de decisión	56,76%	1060
KNN	54%	126
Red neuronal	51,8%	1097
Random Forest	62,56%	665

Búsqueda de patrones



GENDER	RACE	AGE
Female	Other, African American	[40, 50) [60, 70)
Male	Hispanic, African American	[20, 30) [60, 70) [80, 90)

Búsqueda de patrones

GENDER	RACE	AGE
Female	Other, Asian, African American	[60, 70)
Male	Asian, Hispanic, African American	[70, 80) [80, 90) [90, 100)