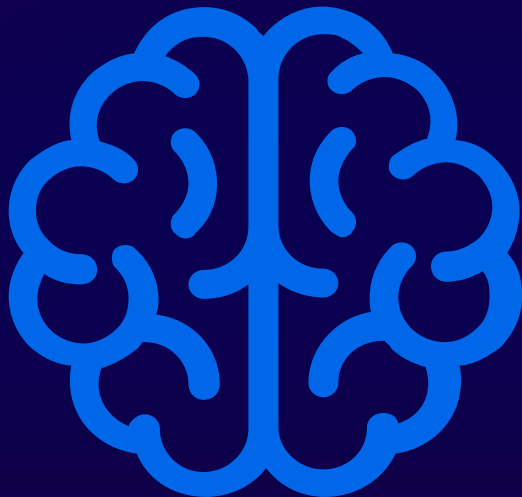




Prevención de ACV

Proyecto Final
Coderhouse

Tabla de contenidos



01

Caso de negocio

Situación actual y
objetivo del modelo

02

Dataset

Descripción y
procesamiento de los datos

03

EDA

Análisis univariado,
bivariado y multivariado

04

Algoritmo

Elección del algoritmo y
performance

05

Conclusiones

Propuestas y siguientes
pasos

Equipo



Bautista Pazos

Estudiante de ingeniería
industrial. Data Analyst.

INTRODUCCIÓN

Un accidente cerebrovascular (ACV) ocurre cuando algo bloquea el suministro de sangre a una parte del cerebro o cuando un vaso sanguíneo en el cerebro estalla.

En cualquier caso, partes del cerebro se dañan o mueren. Un derrame cerebral puede causar daño cerebral duradero, discapacidad a largo plazo o incluso la muerte.

INTRODUCCIÓN

En la Argentina:

- Se produce un accidente cerebrovascular (ACV) cada nueve minutos.
- 126 mil casos de ACV por año, de los cuales 18 mil terminan en muerte.

En Estados Unidos:

- USD\$ 46 mil millones fueron los costos relacionados a ACVs en 2015. Esto incluye atención médica, medicamentos y días de trabajo perdidos.

Según la OMS, **el 80% de los casos son prevenibles**

OBJETIVO

Construir un modelo de Machine Learning utilizando algoritmos de clasificación para que médicos, clínicas, aseguradoras y empresas den un diagnóstico más certero a la hora de prevenir un accidente cerebrovascular.

Se utilizó un dataset de kaggle.com
cuya fuente es anónima.

Cuenta con 5110 observaciones y 12
variables categóricas y continuas.

DATASET



VARIABLES CATEGÓRICAS



VARIABLES CONTINUAS

ID



Índice de masa corporal
(BMI)

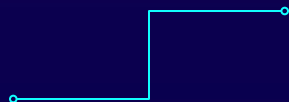


Promedio de glucosa en
sangre



Desafío

Realizar un modelo predictivo en base a los datos obtenidos.



Resultados

La métrica que definirá que tan bueno es el modelo será la de Recall.



Solución

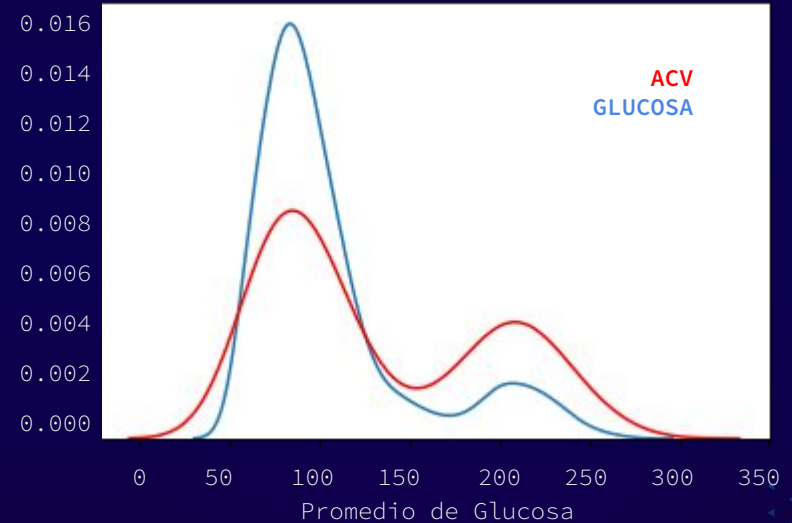
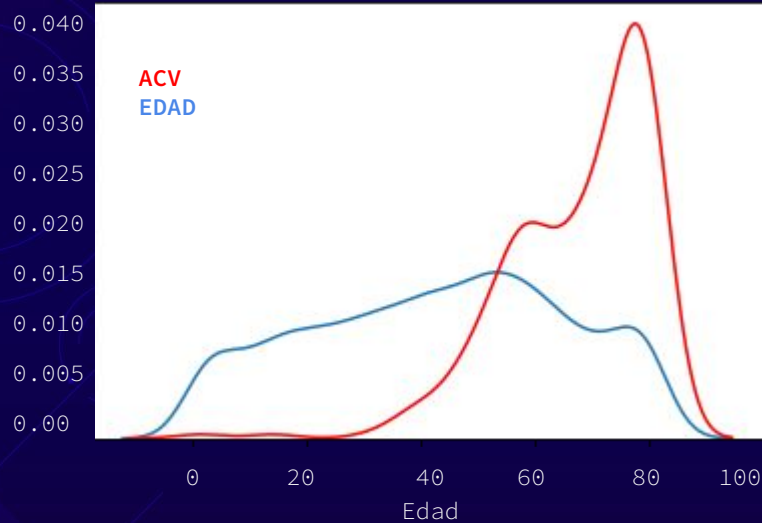
Dar a los clientes una herramienta que les permita dar diagnósticos más certeros

Análisis univariado, bivariado y
multivariado.

EXPLORATORY DATA ANALYSIS (EDA)

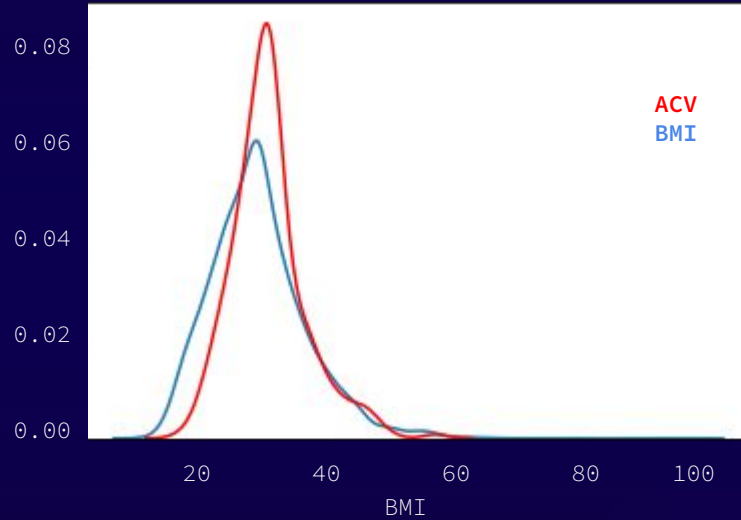


EDA



- Una mayor edad parece ser un factor determinante respecto al riesgo de sufrir un acv.
- Un promedio alto de glucosa puede aumentar el riesgo de sufrir acv.

EDA



- Los casos de ACV en su mayoría contaban con un BMI de 30

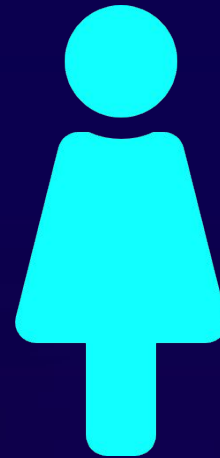
EDA Género

41%
Hombres



5% TUVO UN ACV

61%
Mujeres



4% TUVO UN ACV

EDA Hipertensión

10%
CON HIPERTENSIÓN



14% TUVO UN ACV

90%
SIN HIPERTENSIÓN



4% TUVO UN ACV

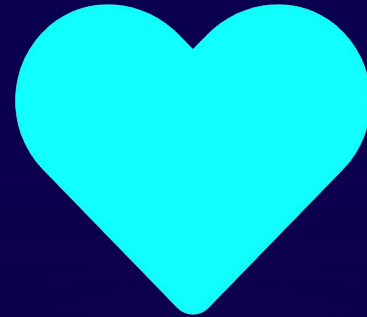
EDA Enfermedad cardíaca

5%
CON ENF. CARDÍACA



17% TUVO UN ACV

95%
SIN ENF. CARDÍACA



4% TUVO UN ACV

EDA

Estado civil

66%
CASADOS



6% TUVO UN ACV

34%
SOLTEROS



4% TUVO UN ACV

EDA

Tipo de trabajo

54%
SECTOR PRIVADO



5% TUVO UN ACV

15%
INDEPENDIENTE



8% TUVO UN ACV

12%
GOBIERNO



5% TUVO UN ACV

13%
HIJOS



1% TUVO UN ACV

1%
NUNCA TRABAJO



SIN ACVs

EDA

Tipo de residencia

51%
URBANA



5% TUVO UN ACV

49%
RURAL



4% TUVO UN ACV

EDA

Estado de fumador

15%
FUMA



5% TUVO UN ACV

17%
SOLÍA FUMAR



8% TUVO UN ACV

37%
NUNCA FUMÓ



5% TUVO UN ACV

30%
NO SE SABE



3% TUVO UN ACV

EDA

Matriz de Correlaciones

age	age	bmi	hypertension	heart_disease	avg_glucose_level	stroke
age	1	0.28	0.26	0.24	0.33	0.25
bmi	0.28	1	0.11	0.17	0.16	0.13
hypertension	0.26	0.11	1	0.16	0.04	0.13
heart_disease	0.24	0.17	0.16	1	0.17	0.13
avg_glucose_level	0.33	0.16	0.04	0.17	1	0.046
stroke	0.25	0.13	0.13	0.13	0.046	1

- La edad y bmi son las variables con mayor correlación.
- La variable que más se relaciona con la posibilidad de tener un acv es la edad.

PRIORIDAD DE MÉTRICAS



1º RECALL

Porcentaje de clasificación correcta de los verdaderos positivos.

2º PRECISION

Porcentaje de clasificación correcta de los falsos negativos.

3º ACCURACY

Porcentaje de aciertos sobre el total.

ALGORITMO UTILIZADO:

RANDOM FOREST



MÉTRICAS OBTENIDAS

100%

RECALL

99%

PRECISION

99%

ACCURACY

Valores reales	Predicción	
	No	Si
Si	964	8
No	0	972

SIGUIENTES PASOS



Inversión

Buscar inversores interesados en la investigación y desarrollo del proyecto.



APP

Desarrollar un front-end para poder interactuar con el modelo.



Promoción

Llevar la aplicación a los potenciales clientes.

MUCHAS GRACIAS!



bautistapazos97@gmail.com



[linkedin.com/bautista-pazos](https://www.linkedin.com/bautista-pazos)