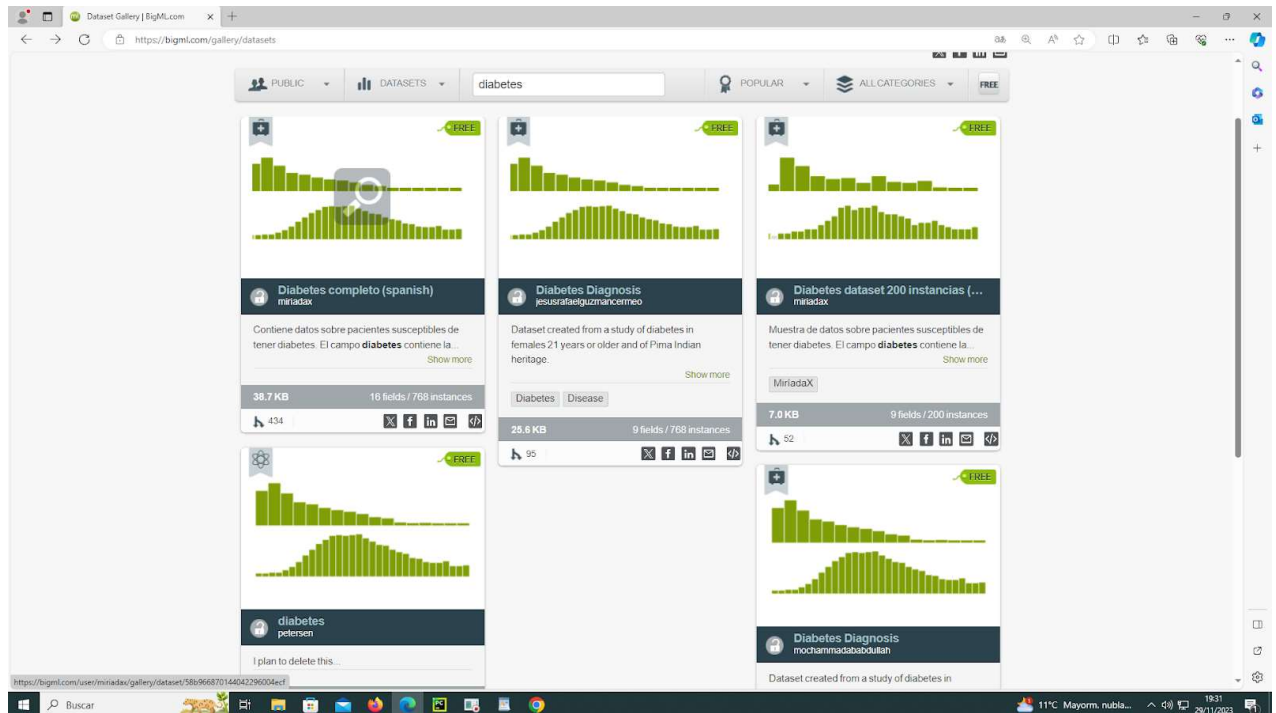
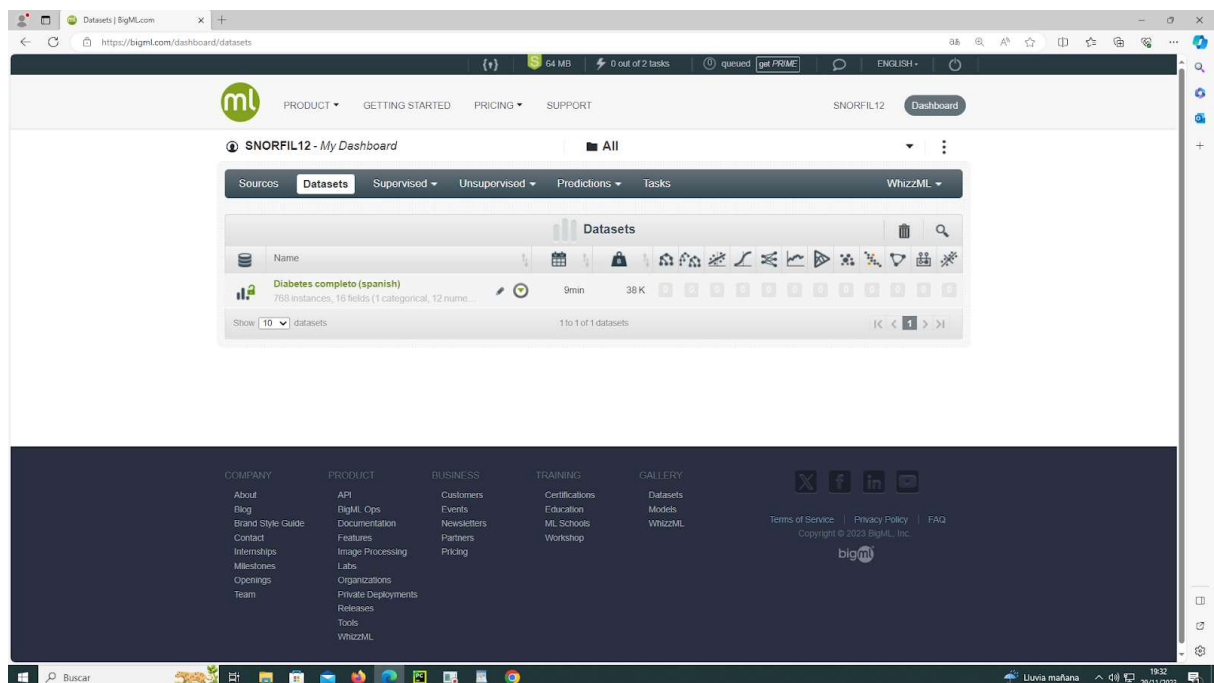


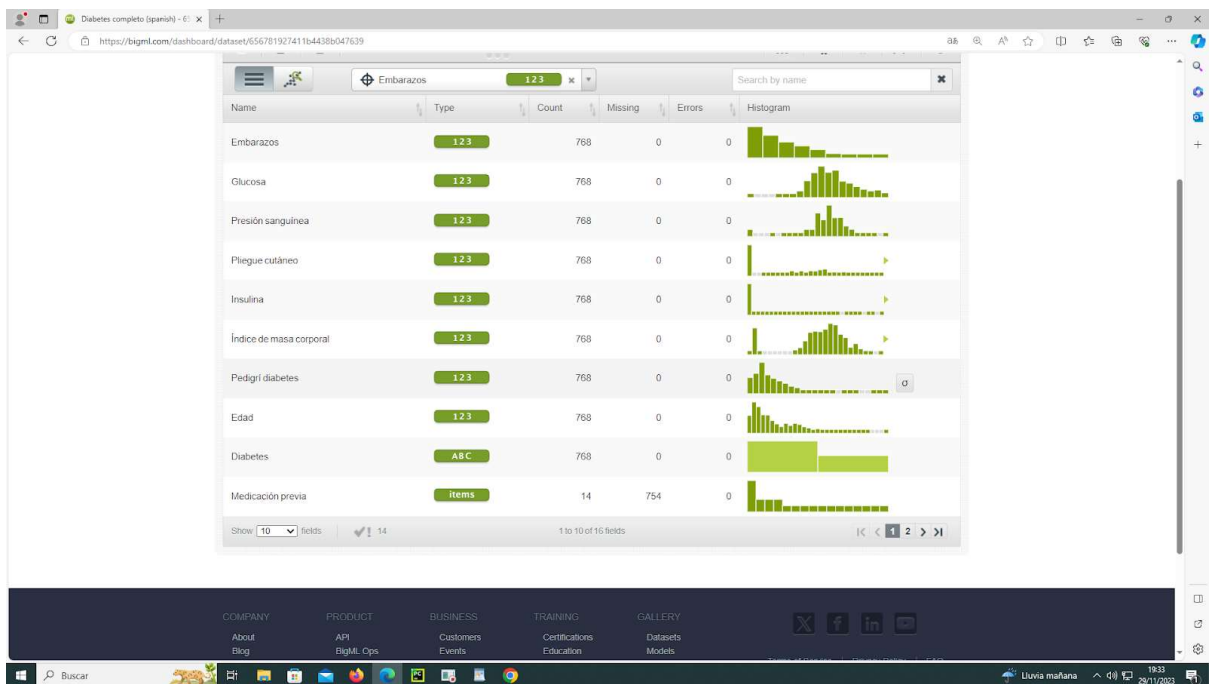
APARTADO 1



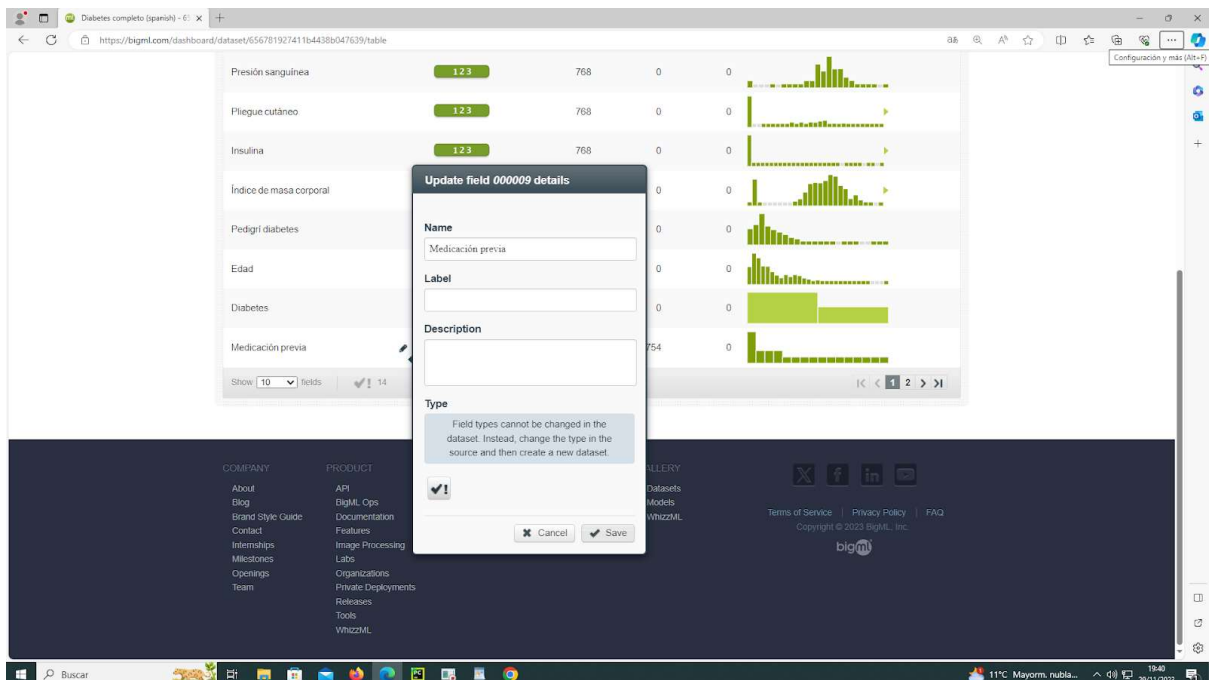
en el buscador de dataset, situado en la parte de abajo de la pagina principal, en el footer esta el link.

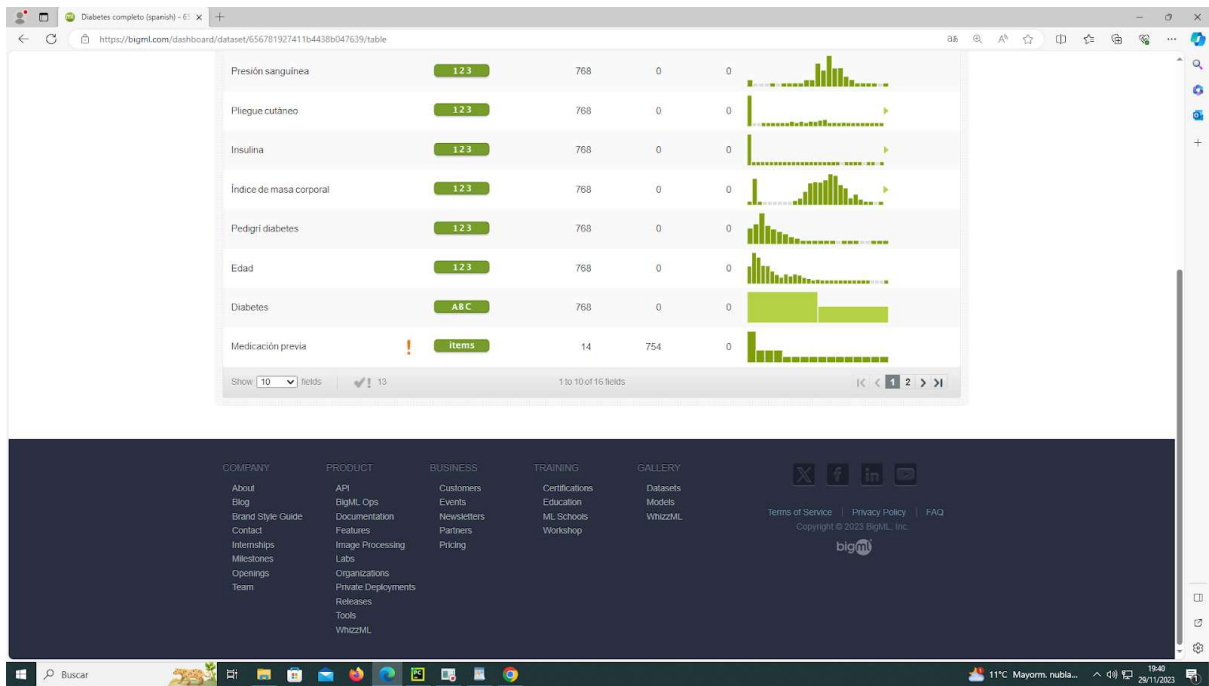


APARTADO 2



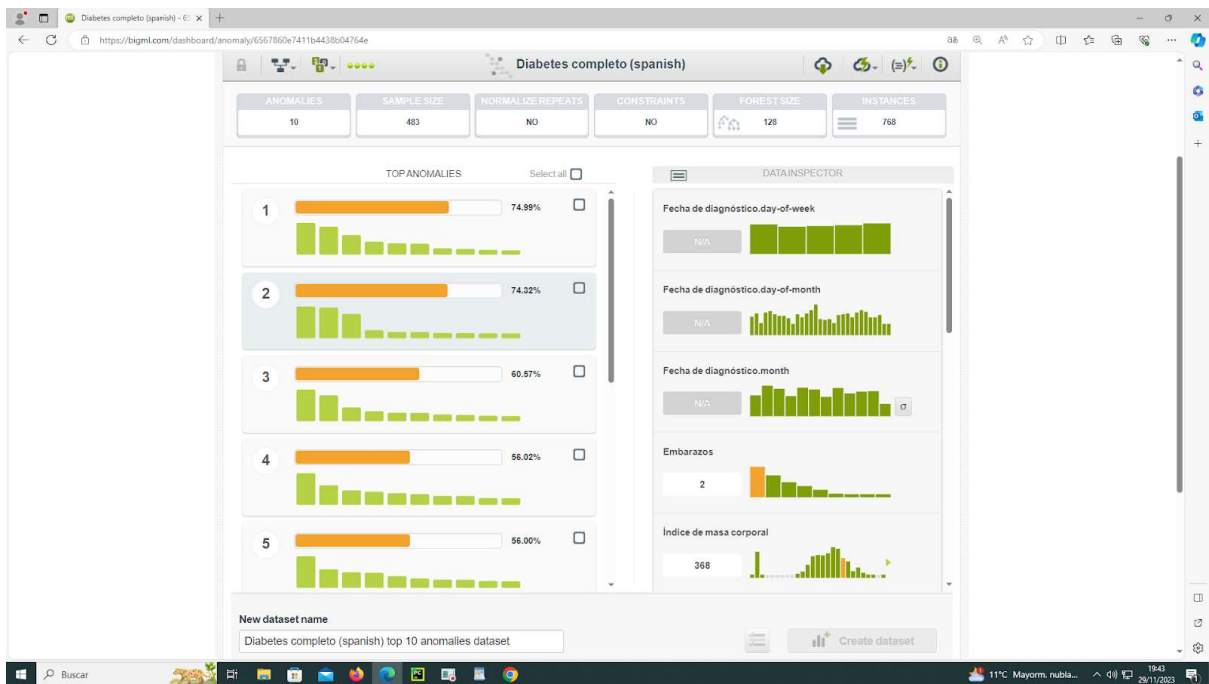
La imagen muestra datos con 768 instancias cada uno hay una categoría de diabeticos o no esa es una categoría y como anomalía podrían ser medicación previa, por que sus porcentajes no varían por lo tanto puede no ser condicionante para detectar la diabetes.





La exclamación excluye el dato concreto para su posterior análisis.

Anomalías detectadas en porcentajes según el análisis



Apartado 3

Diabetes completo (spanish)

SPLIT DATASET CONFIGURATION

Training: 80% | Test: 20% | Seed: | Linear split: ?

Training dataset name: Diabetes completo (spanish) | Training | Test dataset name: Diabetes completo (spanish) | Test (20%)

Create Training | Test

Name	Type	Count	Missing	Errors	Histogram
Embarazos	1,2,3	768	0	0	
Glucosa	1,2,3	768	0	0	
Presión sanguínea	1,2,3	768	0	0	
Plegue cutáneo	1,2,3	768	0	0	
Insulina	1,2,3	768	0	0	

Diabetes completo (spanish)

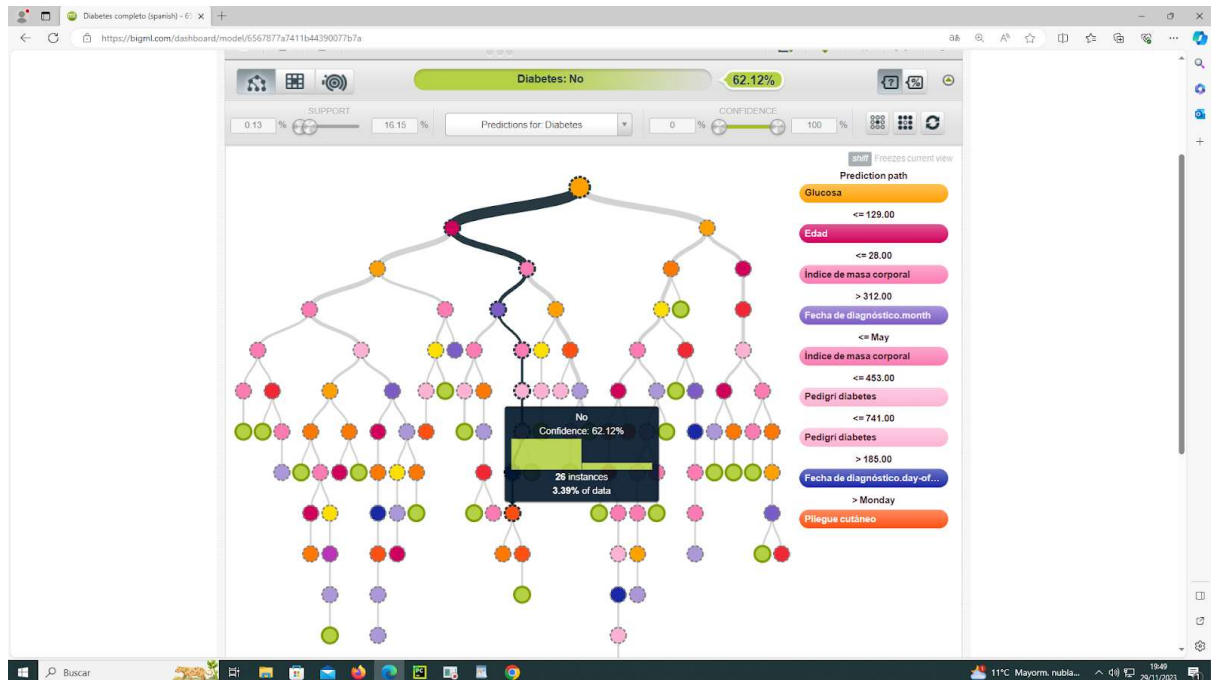
Datasets

Name	Instances	Fields	Time	Size
Diabetes completo (spanish) Training (80%)	614 instances	16 fields (1 categorical, 12 numeric)	0min	31 K
Diabetes completo (spanish) Test (20%)	154 instances	16 fields (1 categorical, 12 numeric)	0min	7 K
Diabetes completo (spanish)	768 instances	16 fields (1 categorical, 12 numeric)	23min	38 K

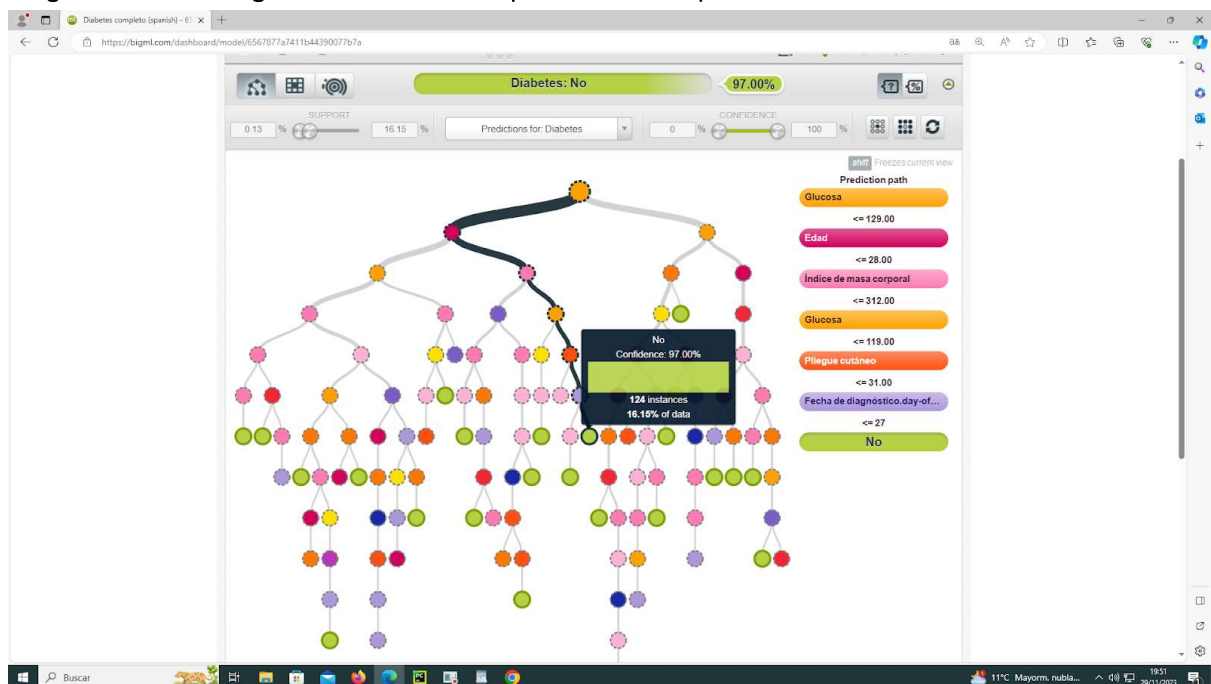
Show 10 datasets | 1 to 3 of 3 datasets

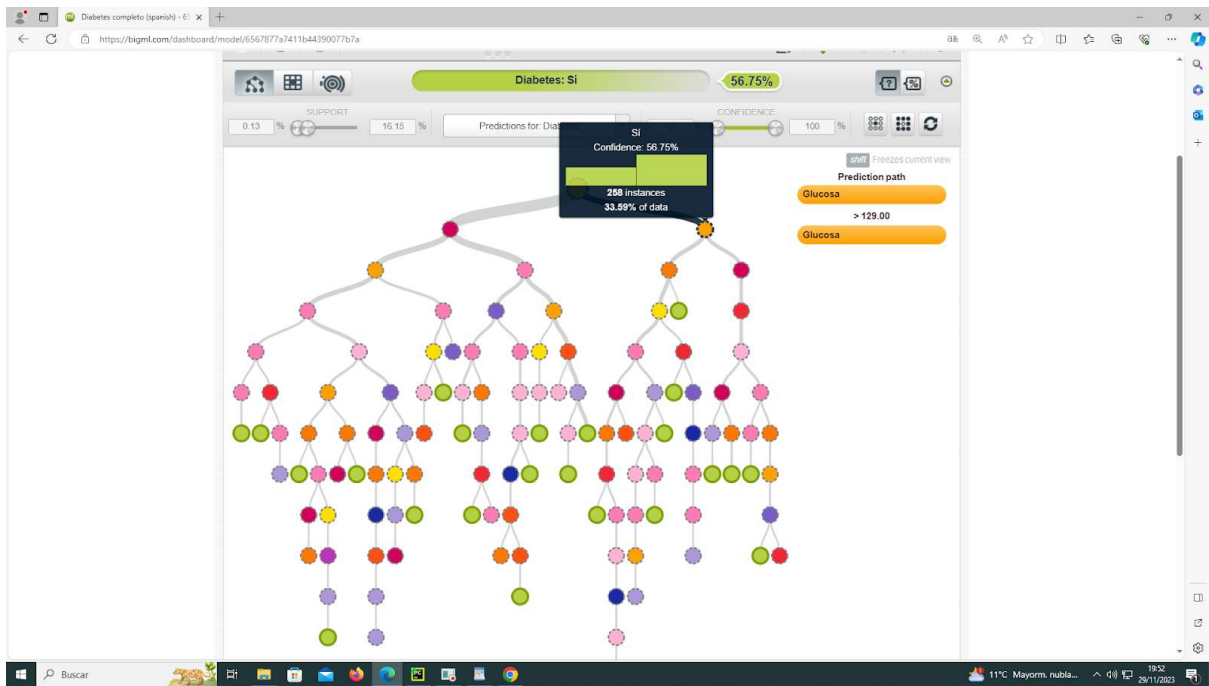
Resultado de la creacion de los datos

APARTADO 4

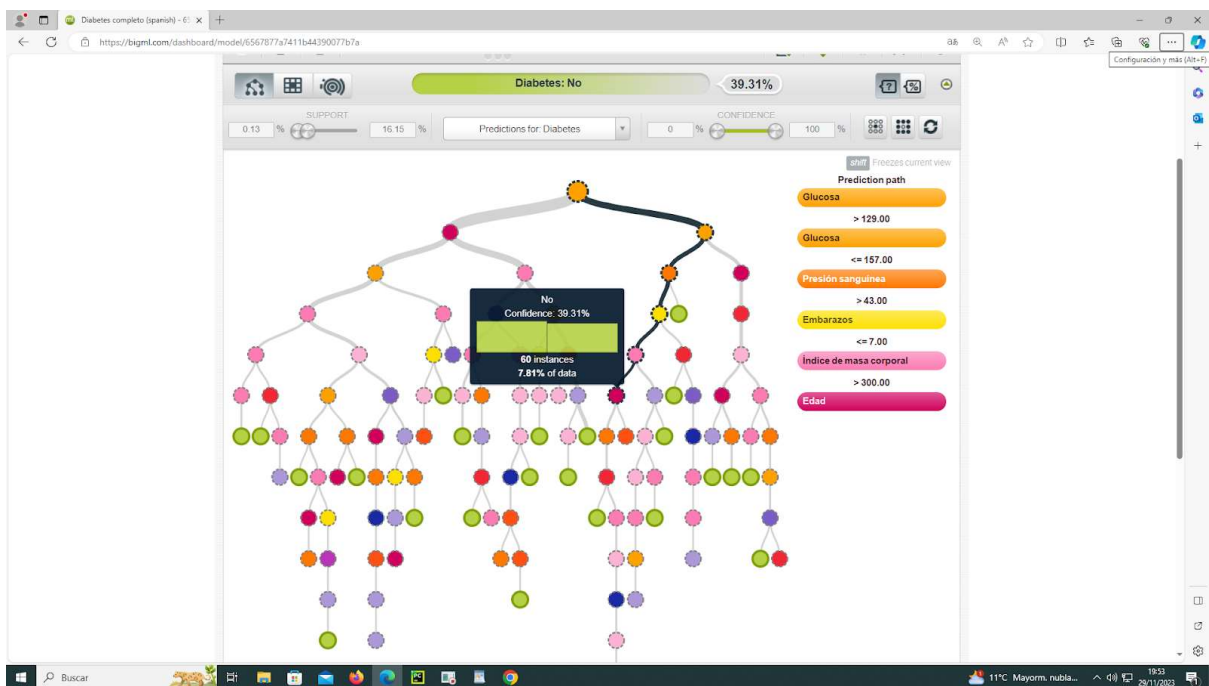


Como se recalca en la pantalla nos podemos fiar en un 75,48 del diagnostico de que no tengas diabetes segun las variables implicadas en el proceso de 194 instancias

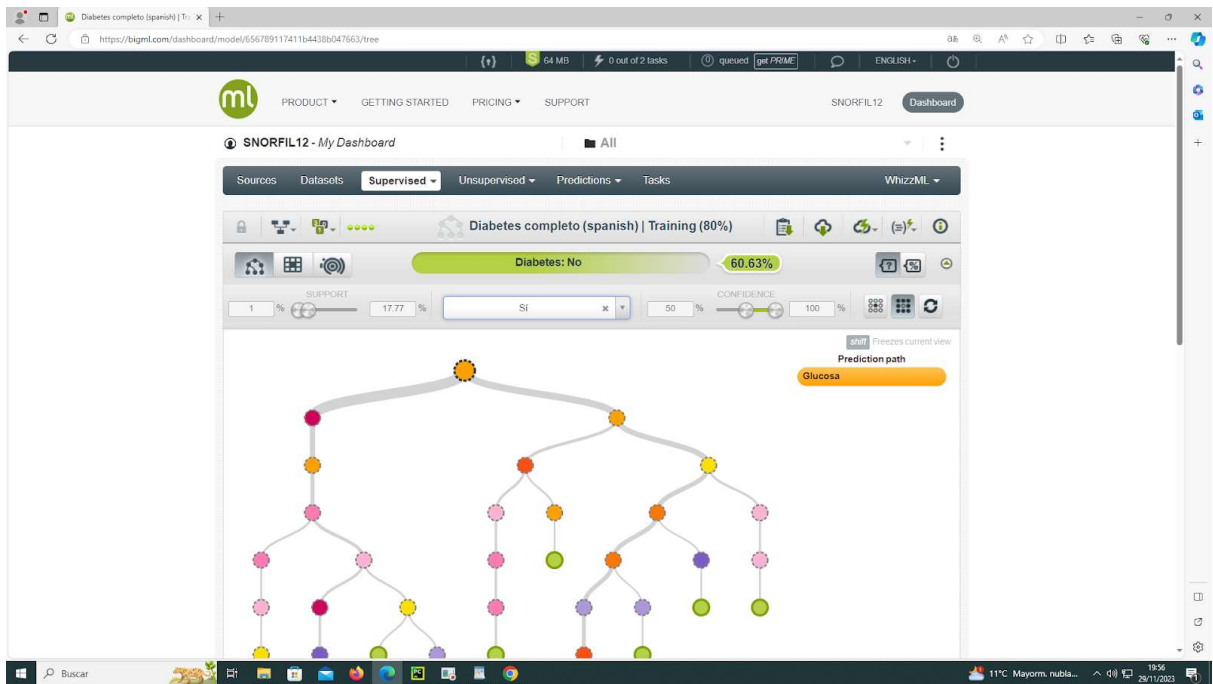
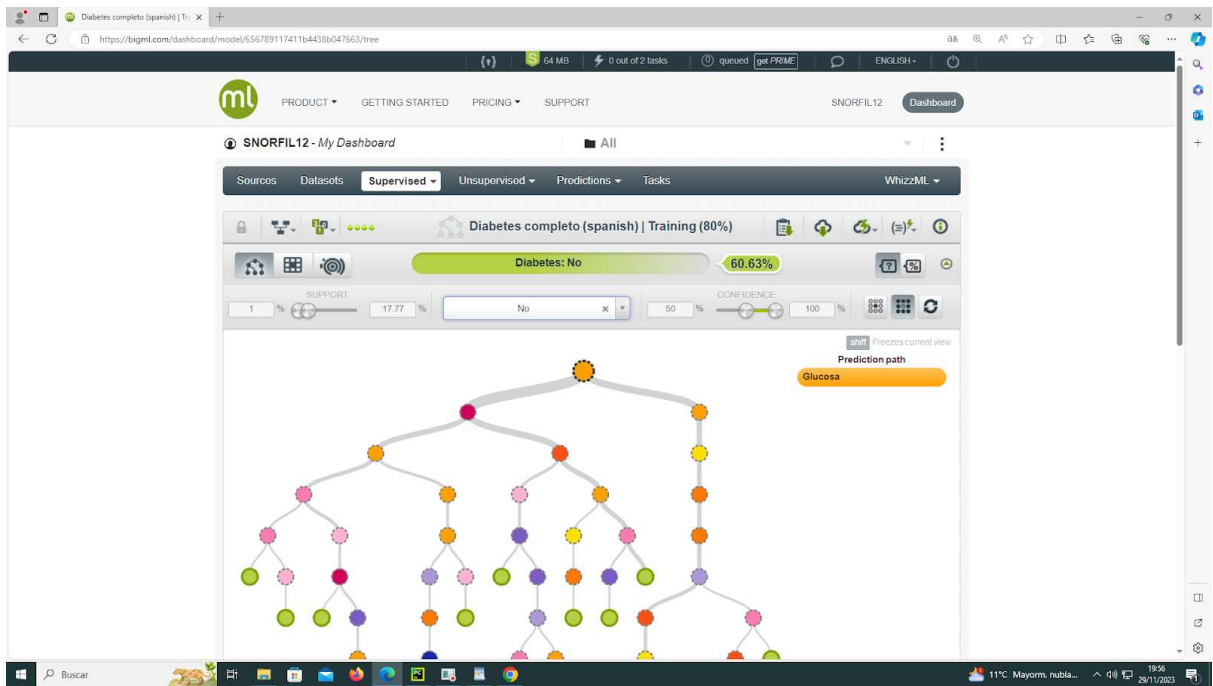


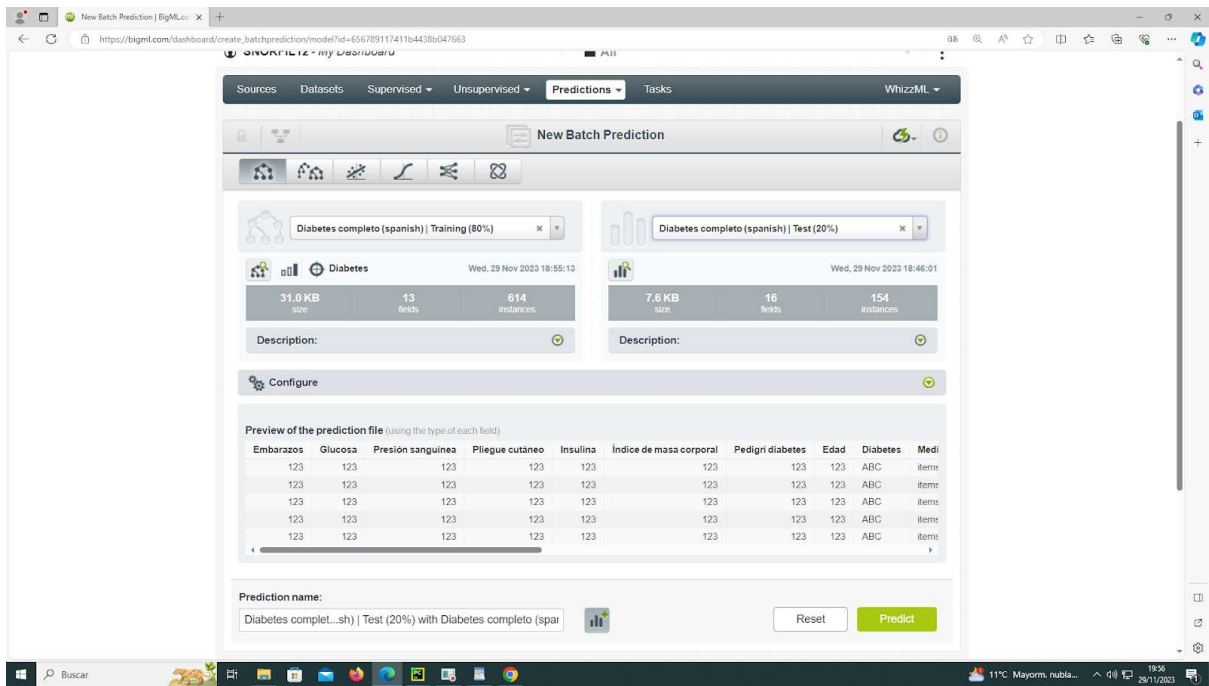


Sin embargo no tenemos una fiabilidad del 52 % de detectar la diabetes.



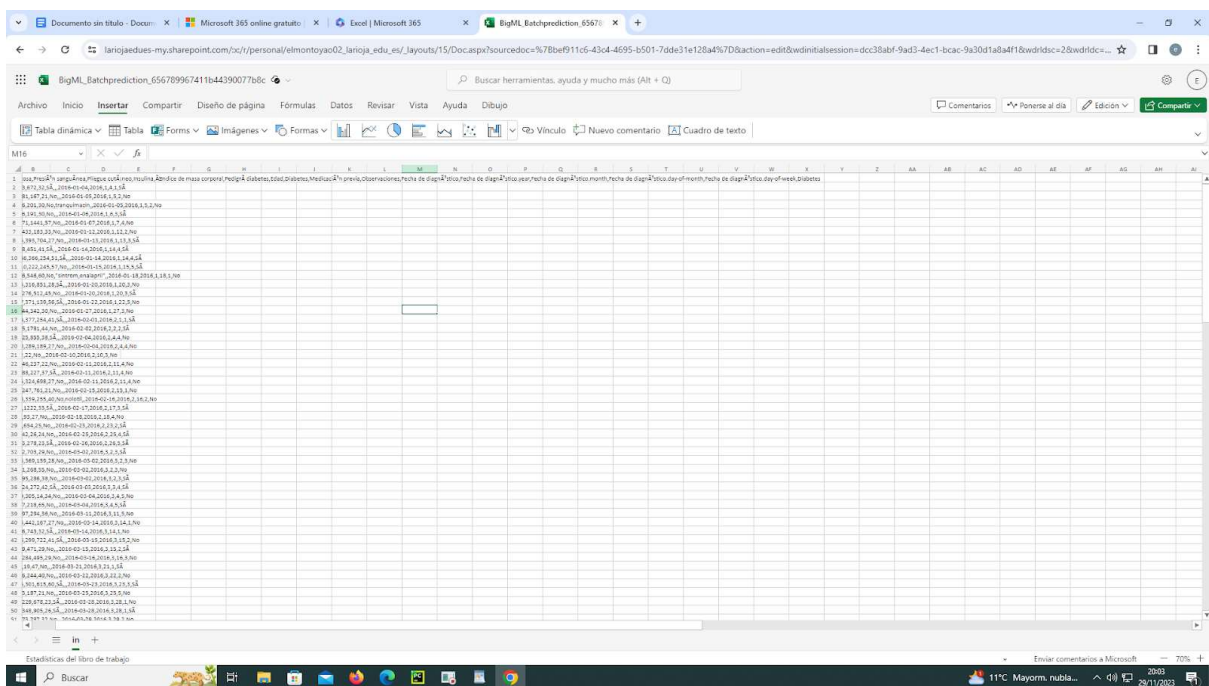
y es poco fiable por el porcentaje y sobre todo por la cantidad de instancias tambien que pueden salir como diabeticos.
no diria que es del todo fiable para detectar si es o no diabetico



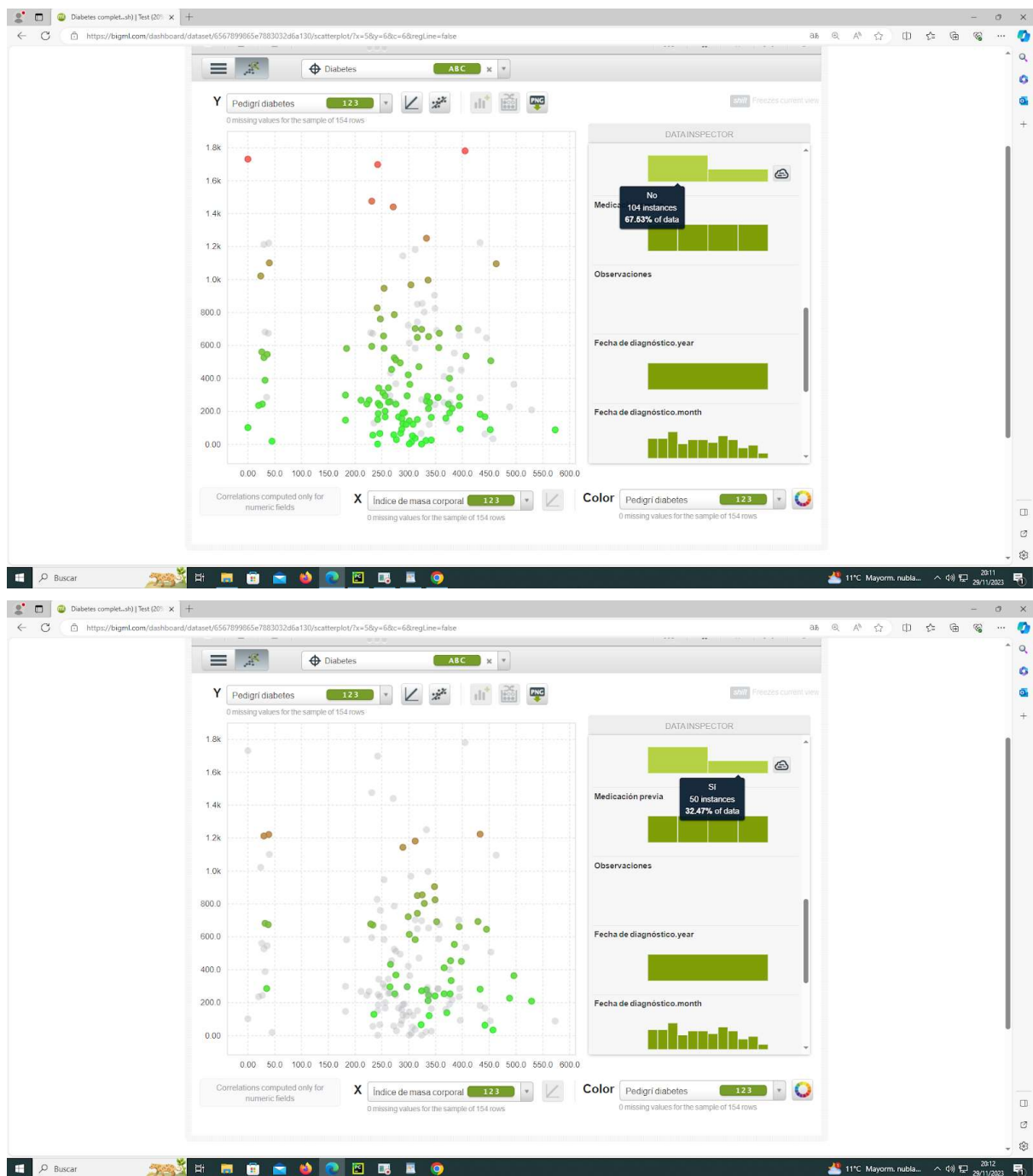


Hacemos una predicción con el 20% de los datos restantes.

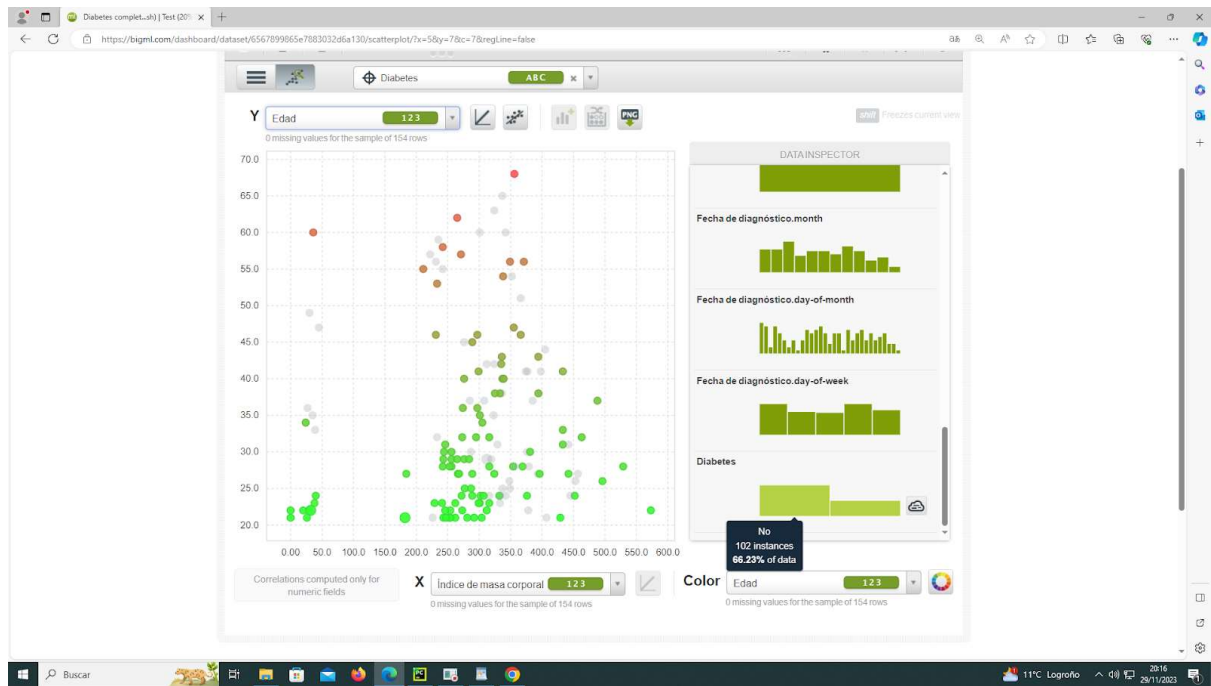
APARTADO 5



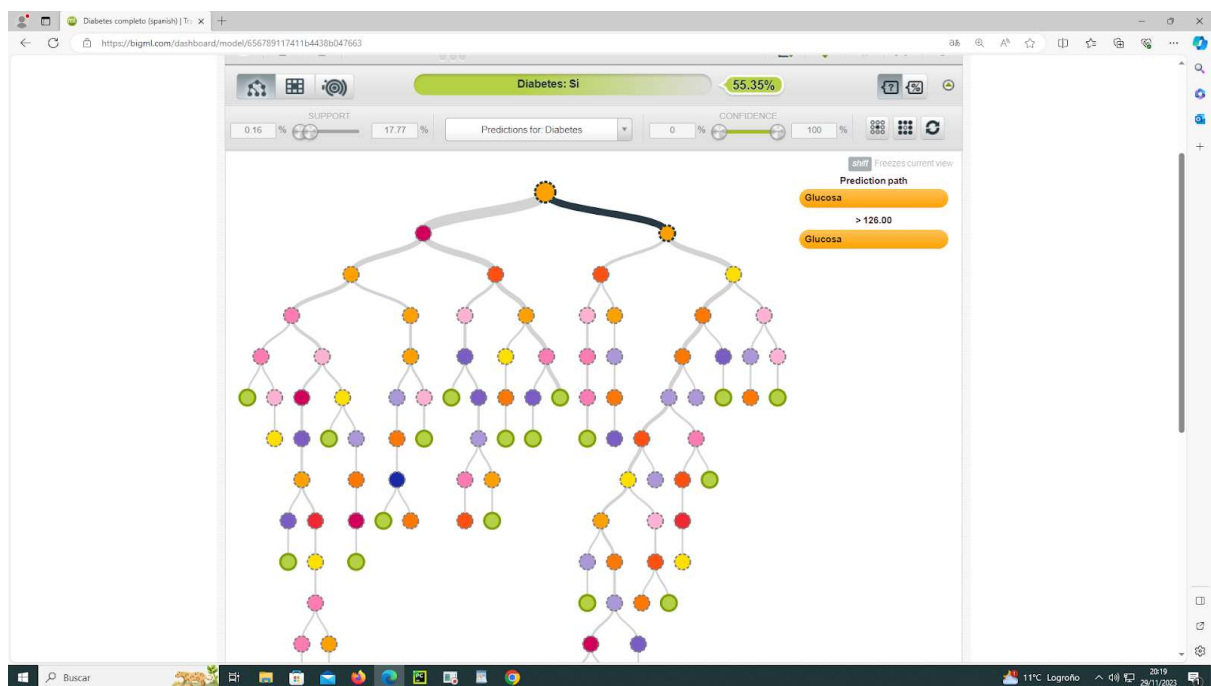
exportamos los datos a xml.



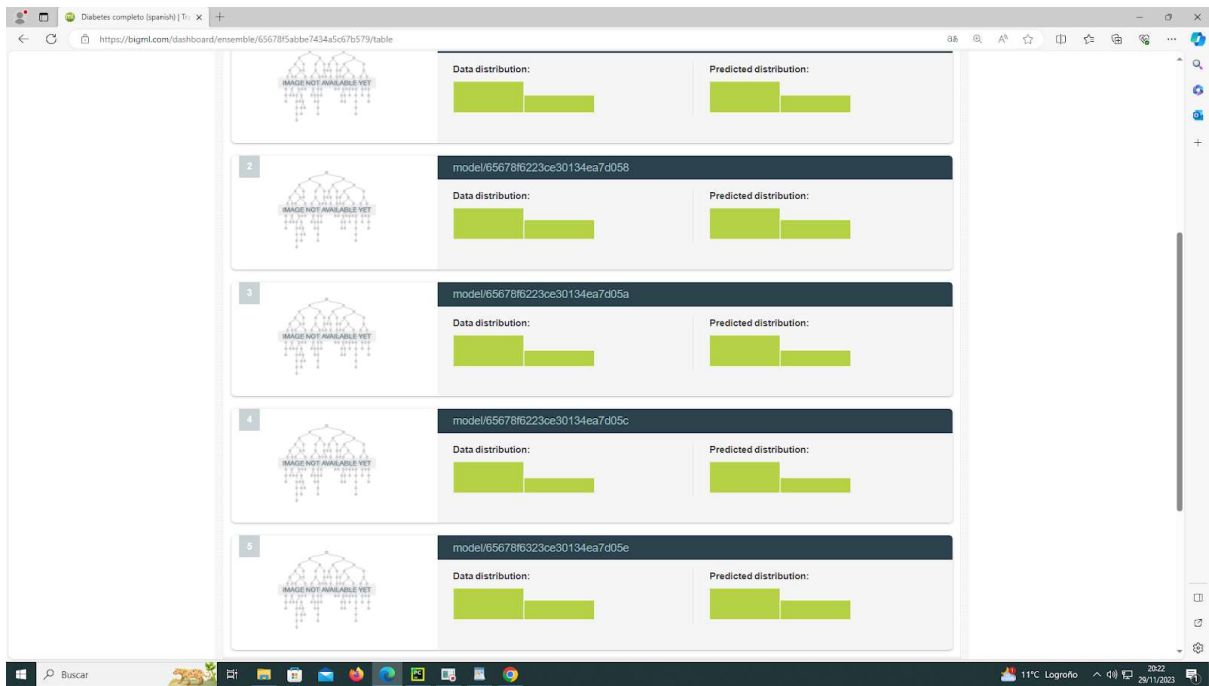
muchos de los datos dependen del indice corporal y se si tienen pedigri con diabetes, se podria decir que el modelo si que se adapto a los parametros de lo que es o no diabetes segun las variables de entrenamiento, y pienso que con los de entrenamiento no lo hizo tan dispar aunque , no es muy fiable un por que sus valores para determinar no es de un 70% y para decir que si a veces es del 20% o del 30% por que lo que no te puedes fiar mucho .



en este otro grafico comparo las edades con su indice de masa corporal con el fin de encontrar alguna relacion visualizando los que no tienen respecto los que si tienen.



Tambien he visto que la glucosa es un factor predomienante para detectar un si en la diabetes. con un 53,35 por ciento de fiabilidad.



con los diferentes arboles generados, vemos que mayormente suele ser no con las diferentes comparaciones entre los datos.

