Saludos,

Con el objetivo de gestionar eficientemente el tiempo y la cantidad de información disponible se implementará un modelo de aprendizaje automático que pronostique el nivel de riesgo de la enfermedad Bipolaridad de los pacientes, considerando los registros de los familiares del estudio.

Dado que contamos con una cantidad limitada de información en este momento, se desarrollará un flujo de trabajo en Python que permitirá importar un archivo Excel con los registros de pacientes con una estructura predefinida para calcular el nivel de riesgo de la enfermedad bipolaridad. Esta implementación nos brindará la flexibilidad necesaria para reutilizar y actualizar el modelo a medida que se obtengan más registros de pacientes.

Para resolver el problema de vacíos en la base de datos por familiar, se propone el tomar un numero predefinido de familiares de cada paciente con el objetivo de mitigar los valores faltantes:

Criterios de selección:

Usar los familiares directos de los pacientes:

* Mama
* Papa

En caso de que no exista el registro usar los siguientes:

* ⁠⁠Hermanos
* Abuelos

En caso de no existir se optará por

* Cualquier otro parentesco existente en la bd

Así mismo con el objetivo de aprovechar la mayor cantidad de información se propone adicionar un registro auxiliar para el modelo, que se obtendrá con el registro de los demás familiares.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Bipolar\_DIGS\_Yes\_Familiar1** | **Eating\_Disorders\_DIGS\_Yes\_Familiar1** | **Alcohol\_Abuse/Dependence\_DIGS\_Yes\_Familiar2** | **Drugs\_Abuse/Dependence\_DIGS\_Yes\_Familiar2** | **Bipolar\_DIGS\_Yes\_Familiar\_adicional** | **Eating\_Disorders\_DIGS\_Yes\_ Familiar\_adicional** | **Alcohol\_Abuse/Dependence\_DIGS\_Yes\_ Familiar\_adicional** |
| 1 | 0 | 0 | 1 | 1 | 1 | 0 |
| 1 | 1 | 0 | 0 | 0 | 0 | 1 |
| 0 | 0 | 1 | 0 | 0 | 1 | 1 |

Lista propuesta de columnas con datos extra de familiares:

1. Psychiatric\_Diagnosis\_DIGS\_Familiar\_adicional
2. Psychosis\_DIGS\_Familiar\_adicional
3. D+AH1epression\_DIGS\_Familiar\_adicional
4. Anxiety\_DIGS Bipolar\_DIGS\_Familiar\_adicional
5. Eating\_Disorders\_DIGS\_Familiar\_adicional
6. Alcohol\_Abuse/Dependence\_DIGS\_Familiar\_adicional
7. Drugs\_Abuse/Dependence\_DIGS\_Familiar\_adicional
8. Behaviour\_Disorders\_Familiar\_adicional
9. Personality\_Disorders\_Familiar\_adicional

Donde el criterio para sumarizar los datos de los familiares sería el considerar las siguientes condiciones:

* En caso de tener más de un familiar adicional a la clasificación propuesta, se revisará el contenido de la columna y en caso de tener un pariente positivo la columna adicional se tratará como un caso positivo, en caso de no tener un pariente con enfermedad registrada se tratará como caso negativo.
* En caso de no existir más familiares la columna se tratará como un negativo