

# TFM

## PREPARACION

### FASE I: Clasificación entre: ALIVE, DEATH, NO-DATA y MD/GR

- En esta fase, se ha realizado una clasificación de los datos dados según 4 resultados finales:
- Fallecidos o con discapacidades severas (SD-D)
- Con discapacidad moderada o buena recuperación (MR-GR)
- Vivos (pero sin resultados finales)
- Sin datos
- Para este procesamiento se han tenido en cuenta principalmente las siguientes variables:
- EO\_Outcome
- EO\_Symptoms
- TH\_Outcome
- TH\_Symptoms
- GOS5
- GOS8

#### Cuando las variables de GOS5 y GOS8 tienen datos

- Si las filas ya contenían datos en las columnas de GOS5 y GOS8, directamente se han clasificado -según estas variables-. De lo contrario, se ha tenido que analizar las otras variables.

```
head(datos.modelo[,c(17,18,27,28,29,30)])
```

```
##   EO_Outcome EO_Symptoms TH_Outcome TH_Symptoms GOS5 GOS8
## 1           4           1         NA           NA <NA> <NA>
## 2           4           3         NA           NA <NA> MD+
## 3           4           2         NA           NA SD* <NA>
## 4           4           2         NA           NA <NA> GR+
## 5           4           1         NA           NA <NA> <NA>
## 6           4           2         NA           NA <NA> SD-
```

#### Cuando las variables de GOS5 y GOS8 no tienen datos

- Si las variables de Outcome contenían el valor de 1 (fallecimiento) o las variables de Symptoms contenían el valor de 6, directamente esas filas del dataset pasaban a clasificarse como fallecidos.

```
##   EO_Outcome EO_Symptoms TH_Outcome TH_Symptoms GOS5 GOS8 TH_Cause
## 22           1           6         NA           NA <NA> <NA>      NA
## 38           1           6         NA           NA <NA> <NA>      NA
## 50           1           6         NA           NA <NA> <NA>      NA
## 55           1           6         NA           NA <NA> <NA>      NA
## 61           1           6         NA           NA <NA> <NA>      NA
## 85           1           6         NA           NA <NA> <NA>      NA
```

- Si las variables de Outcome contenian el valor de 4 (alta) y las de Symptoms el valor de 1, entonces se han clasificado como “Vivos (pero sin resultados finales)”.

```
##      EO_Outcome EO_Symptoms TH_Outcome TH_Symptoms GOS5 GOS8 TH_Cause
## 1           4           1           NA           NA <NA> <NA>           NA
## 5           4           1           NA           NA <NA> <NA>           NA
## 18          4           1           NA           NA <NA> <NA>           NA
## 20          4           1           NA           NA <NA> <NA>           NA
## 36          4           1           NA           NA <NA> <NA>           NA
## 51          4           1           NA           NA <NA> <NA>           NA
```

- Se clasificaran como “Sin datos” todas aquellas filas que no contengan valores ni en las columnas de Symptoms. Se tienen en cuenta los transferidos a otros hospitales.

```
##      EO_Outcome EO_Symptoms TH_Outcome TH_Symptoms GOS5 GOS8 TH_Cause
## 384           NA           NA           NA           NA <NA> <NA>           NA
## 417           NA           NA           NA           NA <NA> <NA>           NA
## 985           NA           NA           NA           NA <NA> <NA>           NA
## 997           NA           NA           NA           NA <NA> <NA>           NA
## 2270          NA           NA           NA           NA <NA> <NA>           NA
## 2292          NA           NA           NA           NA <NA> <NA>           NA
```

- Si los Symptoms son de 4 o de 5 (Discapacidad Severa), entonces se clasificaran como “Fallecidos o con discapacidades severas”

```
##      EO_Outcome EO_Symptoms TH_Outcome TH_Symptoms GOS5 GOS8 TH_Cause
## 160           5           5           NA           NA <NA> <NA>           NA
## 241           5           5           NA           NA <NA> <NA>           NA
## 317           5           5           NA           NA <NA> <NA>           NA
## 336           5           5           NA           NA <NA> <NA>           NA
## 357           5           5           NA           NA <NA> <NA>           NA
## 573           5           5           NA           NA <NA> <NA>           NA
```

- Asi mismo, si las variables de Outcome contenian el valor de 4 y las de Symptoms el valor de 9, significa que el paciente ha sido dado de alta, pero no se tiene ningun dato sobre el estado final, por lo tanto, se han incluido en la clasificacion de “Vivos (pero sin resultados finales)”.

```
##      EO_Outcome EO_Symptoms TH_Outcome TH_Symptoms GOS5 GOS8 TH_Cause
## 409           4           9           NA           NA <NA> <NA>           NA
## 1000          4           9           NA           NA <NA> <NA>           NA
## 4859          4           9           NA           NA <NA> <NA>           NA
```

- Se han visto 3 elementos de NODATA, cuyos pacientes obtienen un estado de symptoma 4, por lo que se envia a estado de fallecido, son datos anomalos.

```
##      EO_Outcome EO_Symptoms TH_Outcome TH_Symptoms GOS5 GOS8 TH_Cause
## 52            2           4           NA           4 <NA> <NA>           3
## 699           2           4           NA           4 <NA> <NA>           NA
## 3025          2           4           NA           4 <NA> <NA>           1
```

## DATOS FINALES:

- Fallecidos o con discapacidades severas

```
## [1] 3559
```

- Con discapacidad moderada o buena recuperacion

```
## [1] 5997
```

- Vivos (pero sin resultados finales)

## [1] 127

- Sin datos

## [1] 86

- Con NA

## [1] 239

## FASE II: Clasificación entre: ESCANEADOS y NO ESCANEADOS

- En primer lugar, se han encontrado ciertos datos anómalos, en los que aparecen datos escaneados (1) y no tienen los datos del escaner, entonces deberíamos ponerlo como no escaneado (2).

```
##      EO_Head.CT.scan EO_1.or.more.PH EO_Subarachnoid.bleed
## 201                1                NA                NA
## 314                1                NA                NA
## 1277               1                NA                NA
## 3234               1                NA                NA
## 3687               1                NA                NA
## 4256               1                NA                NA
##      EO_Obliteration.3rdVorBC EO_Midline.shift..5mm EO_Non.evac.haem
## 201                NA                NA                NA
## 314                NA                NA                NA
## 1277               NA                NA                NA
## 3234               NA                NA                NA
## 3687               NA                NA                NA
## 4256               NA                NA                NA
##      EO_Evac.haem
## 201                NA
## 314                NA
## 1277               NA
## 3234               NA
## 3687               NA
## 4256               NA
```

- A continuación se van a clasificar los datos como:
- Escaneados

```
##      EO_Head.CT.scan EO_1.or.more.PH TH_Head.CT.scan TH_1.or.more.PH outcome
## 2                1                2          <NA>          NA      MDGR
## 7                1                2          <NA>          NA      MDGR
## 8                1                1          <NA>          NA      MDGR
## 14               1                2                1          2          D
## 17               1                1          <NA>          NA      MDGR
## 22               1                2          <NA>          NA          D
```

- No escaneados

```
##      EO_Head.CT.scan EO_1.or.more.PH TH_Head.CT.scan TH_1.or.more.PH outcome
## 3                2                NA          <NA>          NA          D
## 4                2                NA          <NA>          NA      MDGR
## 6                2                NA          <NA>          NA          D
## 9                2                NA          <NA>          NA      MDGR
## 10               2                NA          <NA>          NA      MDGR
```

## 11	2	NA	<NA>	NA	MDGR
-------	---	----	------	----	------

- En analisis

##	EO_Head.CT.scan	EO_1.or.more.PH	TH_Head.CT.scan	TH_1.or.more.PH
## 192	<NA>	NA	<NA>	NA
## 227	<NA>	NA	<NA>	NA
## 240	<NA>	NA	<NA>	NA
## 267	<NA>	NA	1	2
## 329	<NA>	NA	1	2
## 354	<NA>	NA	1	1
##	outcome			
## 192	D			
## 227	D			
## 240	D			
## 267	D			
## 329	D			
## 354	D			

- Si el Outcome es 2 (el paciente se ha transferido a otro hospital), se ha escaneado en dicho hospital (TH\_SCAN) y no se tiene ninguna informaci3n en los esc3neres, se clasificaran como “En an3lisis”.

##	EO_Outcome	TH_Head.CT.scan	TH_1.or.more.PH	TH_Subarachnoid.bleed
## 52	2	<NA>	NA	NA
## 128	2	<NA>	NA	NA
## 135	2	<NA>	NA	NA
## 188	2	<NA>	NA	NA
## 193	2	<NA>	NA	NA
## 207	2	<NA>	NA	NA

- Sobre el dataset NOSCANEADO: Si el Outcome es 2 (el paciente se ha transferido a otro hospital) y no se ha realizado ningun escaner pero si contiene datos en el escaner, entonces se clasificara como “Escaneado”.

##	EO_Outcome	TH_Head.CT.scan	TH_1.or.more.PH	TH_Subarachnoid.bleed
## 201	2	1	2	2
## 217	2	1	2	1
## 257	2	1	1	2
## 314	2	1	1	2
## 318	2	1	2	2
## 1184	2	1	2	2

- Sobre el dataset NOSCANEADO: Nos hemos dado cuenta que existen datos anomalos, que contienen varios escaners, pero sin embargo, no se indica como escaneado, son los registros: 2628,3276,3279,8469,8655, etc. (En total son 12)

##	EO_Head.CT.scan	EO_1.or.more.PH	EO_Subarachnoid.bleed
## 2628	2	2	2
## 3276	2	2	2
## 3279	2	2	2
## 3720	2	2	2
## 7286	2	2	2
## 8469	2	2	2

- Sobre el dataset ENANALISIS: Nos hemos dado cuenta de que existen datos anomalos. Para las variables de los pacientes que se han transferido a otro hospital (TH), existen variables de escaner (TH\_Head.CT.scan) que se encuentran vacias, junto con el resto de variables del escaner en particular. Por lo tanto, se ha asignado el valor de 2 a la variable de escaner (TH\_Head.CT.scan) y se han incluido

en los escaneados, puesto que en todos ellos, en la variable EO\_Head.CT.scan si que existe un valor de 1 (escaneados) y no se han encontrado mas anomalias en dichos datos.

```
##      EO_Head.CT.scan EO_1.or.more.PH EO_Outcome TH_Head.CT.scan
## 681                1                2          2          <NA>
## 1639               1                2          2          <NA>
## 5743               1                2          2          <NA>
## 8434               1                2          2          <NA>
## 8972               1                2          2          <NA>
##      TH_1.or.more.PH TH_Subarachnoid.bleed
## 681                NA                  NA
## 1639               NA                  NA
## 5743               NA                  NA
## 8434               NA                  NA
## 8972               NA                  NA
```

- Sobre el dataset SCANEADO: Se van a eliminar todas las filas que no tengan informacion en el TH\_Major.EC.injury y en el EO\_Major.EC.injury

```
##      EO_Outcome TH_Major.EC.injury
## 76             2                  NA
## 90             2                  NA
## 315            2                  NA
## 361            2                  NA
## 510            2                  NA
## 565            2                  NA
```

- Sobre el dataset SCANEADO: Comprobamos que las variables: EO\_Cause y EO\_Symptoms, no contengan valores nulos.

```
##      EO_Cause EO_Major.EC.injury
## 177          3                  2
## 211          NA                  1
## 242          NA                  2
## 255          2                  2
## 293          NA                  2
## 321          NA                  1
```

## DATOS FINALES:

- Vivos y escaneados

```
## [1] 4158
```

- Vivos y no escaneados

```
## [1] 1535
```

- Vivos en análisis

```
## [1] 304
```

- Fallecidos y escaneados

```
## [1] 2830
```

- Fallecidos y no escaneados

```
## [1] 439
```

- Fallecidos en análisis

## [1] 290

### FASE III: Eliminacion y centralizacion de variables

- Se va a centralizar las variables de PUPIL\_REACT\_LEFT y PUPIL\_REACT\_RIGHT

##	PUPIL_REACT_LEFT	PUPIL_REACT_RIGHT	ESTADOESCANER
## 1	1	1	SCANEADO
## 2	1	1	SCANEADO
## 3	1	1	SCANEADO
## 4	1	1	SCANEADO
## 5	1	1	SCANEADO
## 6	1	1	SCANEADO

- Both reactive

## [1] 5664

- No response unilateral

## [1] 497

- No response

## [1] 634

- Unable to assess

## [1] 193

- Ahora vamos a ver si podemos prescindir o aunar las variables de EO\_Cause y TH\_Cause. Para ello veremos en que caso, ambas variables difieren:

##	EO_Cause	TH_Cause
## 2308	1	3
## 2814	3	1
## 3286	2	3
## 4022	3	1

- Como se puede observar, podriamos prescindir de la variable TH\_Cause, puesto que recoge la misma información que EO\_Cause.