



AN APPROACH TO THE USE OF MACHINE LEARNING TOOLS FOR THE PREDICTION OF ADVERSE EVENTS IN CANCER **PATIENTS ON IMMUNOTHERAPY**



A. Paradela Carreiro¹, L. Otero Millan¹, M. Alfonsín Lara¹, E.Y. Romero Ventosa1, C. Veiga García², M.J. Villanueva Silva³, A.M. Regueira Arcay¹, I. Agra Blanco¹, A. Romero Rendón⁴, N. Martínez López De Castro¹.

2022

INTRODUCTION

The FDA Adverse Reporting System (FAERS) is a tool to voluntary report adverse events (AE), These data can be downloaded and used to apply "Machine learning" (ML) techniques. The bibliography is very limited, although it has already been the subject of a systematic review (Kim et al, 2022). FAERS data set could be useful to elaborate potential predictive modeling.

1 Vigo University Hospital Group, Pharmacy Department, i-FARMA Vigo Research Group (IISGS). 2 Galicia Sur Health Research Institute, Senior Researcher, Vigo, Spain. **3 Vigo University Hospital Group, Oncology Department, Vigo, Spain.** 4 Galicia Sur Health Research Institute, Research Fellow, Vigo, Spain.



PURPOSE

To test a tool of ML to develop a potential predictive model of AE caused by immune checkpoint Inhibitors (ICI), using FAERS data set.

To contrast and explain the ML results with a reference model through conventional (RM), obtained processing data (spreadsheet).

MATERIAL AND METHODS

All FAERS records from 2022 were downloaded, selecting those of the group ICIs notified as "main suspected drug" (inclusion

				1000
22			Therapy Primaryid caseid	Primaryid caseid pt
October - December 2022 posted on 30-Jan-2023	ASCII (ZIP + 69MB)	XML (ZIP+144MB)	dsg_drug_seq start_dt end_dt dur dur_cod V primaryid V primaryid	nic yid
July - Sepember 2022 posted on 4 Nov 2022	ASCII (ZIP+63.2MB)	XML (ZIP - 132MB)	Indication caseid caseid Indication role_cod i_f_cod Image: seq caseid drug_seq caseid Image: seq caseid role_cod i_f_cod Image: seq caseid grimaryid grimaryid caseid prod_ai mfr_dt	rsion de dt
April - June 2022 posted on 2 Aug-2022	ASCII (ZIP+63MB)	XML (ZIP+140MB)	indi_drug_seq indi_pt val_vbm dose_vbm cum_dose_chr cum_dose_unit indi_pt indi_pt indi_pt indi_pt indi_pt indi_pt indi_pt indi_pt indi_drug indi_pt indi_pt indi_pt indi_drug indi_pt indi_drug indi_pt indi_drug indi_pt indi_pt indi_drug indi_pt indi_drug indi_pt indi_drug indi_pt indi_drug indi_pt indi_drug indi_pt indi_drug indi_drug indi_drug indi_drug indi_pt indi_drug indi indi_drug indi indi indi indi indi indi indi ind	a_dt
January - March 2022 posted on 2-May-2022	ASCII (ZIP - 64.7MB)	XML (ZIP - 136MB)	dechal mfr_sn rechal lit_ref lot_num age exp_dt age_co nda_num age_gr dose_amt sex dose_unit e_sub	dr pd rp
			dose_freq wt cor	4
Asistente para importación de texto ¿Qué delimitador separa los campos? Seleccione que aparece abajo. — Elija el delimitador que separa los campos:	el delimitador apropiado y compr	uebe cómo se ve afectado el texto	ista previa	d t cod er_country pountry
Asistente para importación de texto ¿Qué delimitador separa los campos? Seleccione que aparece abajo. Elija el delimitador que separa los campos: O <u>T</u> abulación O P <u>u</u> nto y coma	el delimitador apropiado y compr O Co <u>m</u> a O E <u>s</u> pac	uebe cómo se ve afectado el texto o O tros: \$	ista previa	eka GUI C – D X
Asistente para importación de texto ¿Qué delimitador separa los campos? Seleccione que aparece abajo. Elija el delimitador que separa los campos: <u>T</u> abulación <u>Pu</u> nto y coma <u>P</u> rimera fila contiene nombres de campos	el delimitador apropiado y compr O Co <u>m</u> a O E <u>s</u> pac <u>C</u> ualificac	uebe cómo se ve afectado el texto o Otros: \$ or de texto: {ninguno}	x ista previa X X X X X X X X X	eka GUI C – 🗆 X Applications
Asistente para importación de texto ¿Qué delimitador separa los campos? Seleccione que aparece abajo. Elija el delimitador que separa los campos: <u>Tabulación</u> <u>Punto y coma</u> <u>Primera fila contiene nombres de campos</u> <u>imaryid</u> <u>caseid</u> <u>drug seq roj</u>	el delimitador apropiado y compr Co <u>m</u> a Espac <u>C</u> ualificac <u>e cod drugname</u>	uebe cómo se ve afectado el texto o Otros: \$ or de texto: {ninguno}	ista previa	eka GUI C – 🗆 X Applications Explorer
Asistente para importación de texto ¿Qué delimitador separa los campos? Seleccione que aparece abajo. Elija el delimitador que separa los campos: <u>Tabulación</u> Punto y coma Primera fila contiene nombres de campos <u>imaryid</u> caseid drug seq rol 2058834 10005883 1 PS 2058834 10005883 2 SS 2058834 10005883 3 SS	el delimitador apropiado y compr Coma Espac Cualificac <u>Cualificac</u> <u>e cod drugname</u> ACETAMINOPHEN TRAMADOL BROMAZEPAM ZOL DI DEM	uebe cómo se ve afectado el texto o Otros: \$ or de texto: {ninguno}	x ista previa	eka GUI C – 🗆 X Applications Explorer Experimenter
Asistente para importación de texto ¿Qué delimitador separa los campos? Seleccione que aparece abajo. Elija el delimitador que separa los campos: <u>Tabulación</u> Punto y coma <u>Primera fila contiene nombres de campos</u> <u>imaryid</u> <u>caseid</u> <u>drug seq</u> rol 2058834 10005883 1 PS 2058834 10005883 2 SS 2058834 10005883 3 SS 2058834 10005883 4 SS 2058834 10005883 5 SS 2058834 10005883 5 SS 20597211 10005972 1 PS 20597211 10005972 2 C	el delimitador apropiado y compr Coma Espac Cualificac e cod drugname ACETAMINOPHEN TRAMADOL BROMAZEPAM ZOLPIDEM CETIRIZINE NILOTINIB HYDROCHLOROTHI	uebe cómo se ve afectado el texto o O Otros: \$ or de texto: {ninguno}	Acetraminophen TRAMADOL BROMAZEPAM ZOLPIDEM TARTRAT CETIRIZINE HYDRO NILOTINIB HYDROCHLOROTHIAZ	eka GUI C – 🗆 × Applications Explorer Experimenter KnowledgeFlow
Asistente para importación de texto ¿Qué delimitador separa los campos? Seleccione que aparece abajo. Elija el delimitador que separa los campos: 	el delimitador apropiado y compr Coma Espac Cualificac Cualificac Cualificac Cualificac ACETAMINOPHEN TRAMADOL BROMAZEPAM ZOLPIDEM CETIRIZINE NILOTINIB HYDROCHLOROTHI LEVOTHYROXINE TORSEMIDE GLUCOPHAGE	uebe cómo se ve afectado el texto o O Otros: \$ or de texto: {ninguno} AZIDE\IRBESARTAN SODIUM\POTASSIUM IOI	ACETAMINOPHEN TRAMADOL BROMAZEPAM ZOLPIDEM TARTRAT CETIRLIZINE HYDRO NILOTINIB HYDROCHLOROTHIAZ LEVOTHYROXINE SO TORSEMIDE METFORMIN HYDRO	eka GUI C – 🗆 × Applications Explorer Experimenter KnowledgeFlow Workbench

Figure 3: FAERS data import tool to the local database in Access.



criteria). Collected variables from FAERS data set were: AE, age, drug and sex. The ML decision tree classification algorithm J48 implemented in the Weka application (version 3.8.6) was used to elaborate the ML model. The RM was built using a spreadsheet to tabulate and analyze the data (pivot tables and descriptive statistics).

RESULTS

1,702,222 notifications were downloaded and 86,053 records were selected according to inclusion criteria. The J48 algorithm applied to a subset including "adverse effect", "sex" and "drug", allowed us to estimate, for each AE the most likely responsible ICI drug. The metrics of the ML model obtained were satisfactory and compatible with the RM analysis. The J48 algorithm produced a complex tree (to be expected given the large number of AE). The application of J48 on another subset that includes "adverse effect", "age" and "drug", had a lower predictive capacity, due to the lower consistency of the data (age is only recorded as younger or older than 65 years) and that there is a higher proportion of missing values. The RM allows the results obtained with ML to be easily explained and understood.

	Attributes		No.	Label Co	unt Weight			
	All None	Invert Pat	ttern 1 Abdom	inal discomfort 59	59			
		N	2 Abdom	inal distension 68	68			
	No.	Name	3 Abdom	inal pain 206	206			
	1 Efecto		4 Abdom	inal pain upper 123	123			
	2 sex		5 Acute k	Idney Injury 415	415			
	3 Clase		6 Adrena	disorder 72	72			
			7 Adrena	Insufficiency 394	394			
			8 Adreno	corticotropic no 136	130			
	Figure 5: Exam	ple of analyzin	10 Alapadi	aminotransferas 225	225			
				a 101				
	a subset of da	ta with the	Class: clase (Nom)					
	<i>"explorer" too</i>	Ι.						
🕽 Weka Explorer	Re	move						
Preprocess Classify Cluster Ass	cia' Status							
lassifier	OK				Log x0			
Choose J48 -C 0.25 -M 2								
est options	Classifier output							
Use training set	Correctly Classified Instances	17470 42	.984 %					
Supplied test set Set	Incorrectly Classified Instances	23173 57	.016 %					
Cross-validation Folds 10	Kappa statistic	0.1873						
	Mean absolute error	0.1963						
Percentage split % 66	Root mean squared error	0.3151						
More options	Relative absolute error	92.44/1 %						
	Total Number of Instances	40643						
Nom) clase 🗸 🗸	Total Number of Instances	10015						
Start Stop	=== Detailed Accuracy By Class ===	:						
esult list (right-click for options)	TP Rate FP Rate	Precision Recall F-Me	asure MCC ROC Area	PRC Area Class				
1:44:53 - trees.J48	0,634 0,439	0,422 0,634 0,50	0,184 0,644	0,468 NIVOLUMA				
	0,093 0,025	0,329 0,093 0,14	6 0,123 0,689	0,230 IPILIMUM				
	0,507 0,250	0,469 0,507 0,48	0,251 0,672	0,485 PEMBROLI				
	0,259 0,100	0,382 0,259 0,30	9 0,187 0,672	0,334 ATEZOLIZ				
	0,000 0,000	? 0,000 ?	? 0,659	0,035 AVELUMAE				
	0,056 0,002	0,555 0,056 0,10	0,169 0,697	0,131 DURVALUM				
	0,000 0,000	2 0,000 2	2 0,579	0,006 CEMIPLIM				
	Weighted Avg. 0,430 0,240	. 0,150 .	. 0,005	0,100				
	=== Confusion Matrix ===							
			Weka Explorer	eka Explorer				
	a b c d e f	g < classified as	Preprocess Classify Cluster	Associate Select attributes	Visualize			
	06/1 40/ 33/1 1198 0 22 2520 426 1006 500 0 16	0 a = NIVOLUMAB	Classifier	, associate select attributes				
	4763 224 6257 1024 0 0							
	3438 225 2099 2025 0 19	$d = \Delta \pi R Z O T Z I M A B$	Choose J48 -C 0.25 -M 2					
	351 8 159 62 0 0	$0 \mid e = \Delta V FLIMAB$	Test options	Classifier output				
	679 19 342 321 0 81	0 f = DIRVALUMAR	O Use training set	=== Run informatio	on ===			
	94 7 31 14 0 0	0 g = CEMIPLIMAB	Supplied test set Set					
		, g olititutero	Cross-validation Folds 10	Scheme: weka	.classifiers.trees.J48 -C 0.25 -M 2			
			Dercentage split 0/ 55	Relation: mode	elo02_hasta80			
			Percentage spin 70 60	Instances: 4064	3			
			More options	Attributes: 3	to			

CONCLUSIONS

The results of the J48 algorithm were useful for the association between AE, sex and drug. Despite the inherent limitations of voluntary AE reporting, this study will serve as a starting point for applying ML techniques in any other group, using FAERS data.



