



## **Algorry\_Eje3**

# **Determinación de la capacidad predictiva de técnicas de aprendizaje automático aplicadas al comportamiento de alumnos en evaluaciones sistematizadas**

**Aldo M. ALGORRY, Laura C. DIAZ, Carlos A. BARTO**  
**Facultad de Ciencias Exactas, Físicas y Naturales - UNC**  
[aalgorry@gmail.com](mailto:aalgorry@gmail.com), [lcd\\_ic@yahoo.com.ar](mailto:lcd_ic@yahoo.com.ar), [cbarto@gmail.com](mailto:cbarto@gmail.com)

### **Eje y sub-eje seleccionados**

Eje 3: Procesos de enseñanza y aprendizaje con TIC

Sub eje: Incorporación de las TIC en la enseñanza disciplinar (Sociales, Naturales, Exactas y Artes).

### **Resumen**

Este trabajo es la continuación del desarrollado por el grupo de Sistemas Tutores Inteligentes del Departamento de Computación de la FCEFyN de la UNC aplicados a la enseñanza del lenguaje de programación Python en la materia de Informática que se dicta para las carreras de Ingeniería.

En la intención de avanzar en la investigación de los Tutores inteligentes y contando con datos sistematizados del comportamiento de los alumnos por dos años este trabajo avanza en la posibilidad de realizar estimaciones a priori sobre el comportamiento futuro de un alumno ante una pregunta concreta constituyendo un input para un hipotético tutor inteligente.

El presente es continuación del trabajo “Descubrimiento de Patrones cognitivos en Evaluaciones de Informática basado en Explotación de Información” donde se determina mediante el uso de árboles de decisión que existe una correlación que

merece ser estudiada y susceptible de ser predicha con suficiente exactitud a los fines prácticos. Lo que nos lleva a estudiar con mas detalle diferentes técnicas de IA y valorar su performace como predictores de conductas cognitivas.

Lo que se pretende mostrar con este trabajo que nivel de capacidad podría tener un futuro tutor inteligente para determinar cuales opciones de respuesta elegirá un alumno para las preguntas de una prueba basado en las opciones que eligió como respuesta en pruebas anteriores. En definitiva la determinación de patrones cognitivos enlazados que no solo permitan predecir si responderá la opción correcta o una incorrecta sino intentar identificar cuales opciones tienen mayor posibilidad de ser seleccionadas para direccionar el aprendizaje asistido. Como conclusión, se muestran los diferentes resultados obtenidos utilizando distintos clasificadores, con distintas configuraciones y estrategias de preprocesamiento con el propósito de seleccionar una configuración recomendable para este dominio del problema.

## **Palabras Clave**

Aprendizaje Automático, Educación Superior, Tutores Inteligentes

## **Introducción**

Para alcanzar los resultados se utilizaron los datos recolectados mediante un curso virtual implementado en la plataforma Moodle. Este curso corresponde a la materia Informática y se encuentra disponible para los alumnos de primer año de las carreras de Ingeniería de la Facultad de Ciencias Exactas Físicas y Naturales.

El estudio se realizó sobre los datos obtenidos en los años 2013 y 2014, trabajándose sobre la actuación en la materia.

La forma de evaluación de la materia contempla tomar 7 exámenes, llamados Evaluaciones Conceptuales (EC), de 9 preguntas y con 5 opciones de respuesta cada una.

Las opciones incorrectas estan pensadas como representativas de un modelo mental incorrecto que pueda tener un alumno.

Para el presente trabajo se tomó la EC 5 pregunta 5, y se intentó predecir para cada una de las 4 opciones incorrectas si un alumno que no contestó correctamente

seleccionaría esta opción o no. La predicción se basa en las respuestas que el alumno brindó en las 4 EC previas.

## Los datos y su preparación

De cada alumno se tomaron 180 atributos de entrenamiento, correspondientes a los datos de 4 evaluaciones de 9 preguntas y 5 opciones, además de un atributo de clasificación a ser predicho correspondiente a la opción incorrecta y pregunta en estudio.

Los atributos de entrenamiento pueden tener 3 valores posibles: 1 que indica que esa opción se seleccionó, 0 que no se seleccionó y -1 que indica que no respondió la pregunta.

Por cada opción incorrecta de la pregunta 5 de la evaluación 5 se entrenaron los modelos, en total 4 modelos.

Los datos fueron preprocesados de modo de eliminar las instancias de aquellos alumnos que no seleccionaron ninguna opción para la pregunta en estudio (pregunta 5 de la EC 5), con lo que quedaron 1372 instancias de alumnos que respondieron la pregunta, Estas instancias se dividieron en dos grupos de manera aleatoria, uno de entrenamiento con 904 instancias correspondiente al 66% y el otro de test con 468 equivalente al 33%.

La opción correcta para la pregunta es la A, esta opción no se analizó porque el objetivo es el de determinar si un alumno podría marcar una de las opciones incorrectas a fin de que un tutor inteligente detecte el modelo mental incorrecto.

La distribución de las selecciones o no de las distintas opciones incorrectas para los conjuntos de entrenamiento y de test se muestran en el siguiente cuadro:



	Entrenamiento			Test		
	Selecciona	No selecciona	% No Sel/Sel	Selecciona	No selecciona	% No Sel/Sel
Opción B	741	163	22,00%	387	81	20,93%
Opción C	841	63	7,49%	435	33	7,59%
Opción D	692	212	30,64%	332	136	40,96%
Opción E	813	91	11,19%	418	50	11,96%

Cuadro 1: Distribución de las selecciones de las opciones en los conjuntos de entrenamiento y de test.

Como la proporción de las instancias para cada opción en estudio que fueron seleccionadas contra las que no lo fueron es baja, como se puede apreciar en el cuadro 1, el conjunto de entrenamiento fue redistribuido de modo de lograr una cantidad similar de las dos clases posibles (opción seleccionada y opción no seleccionada). Esto se logra mediante la repetición aleatoria de instancias de menor proporción. Dado que los modelos predictivos tienden a premiar las clases de mayor ocurrencia, se realiza este preprocesamiento para evitar ese sesgo.

## Los modelos de aprendizaje automático

Para el presente trabajo se utilizó la herramienta Weka, que permite aplicar diferentes técnicas o modelos de aprendizaje automático.

Se evaluaron los siguientes clasificadores:

- Redes neuronales perceptrón multicapa con una capa de entrada de 180 neuronas, una capa intermedia de 90 neuronas y una capa de salida de una neurona binaria.
- Naive Bayes
- AdaBoostM1. con 10 iteraciones y un umbral de 100.
- Random Forest: con 100 árboles

## Resultados

Primero mostraremos los resultados de los distintos modelos y su capacidad predictiva para la opción B:

		TP Rate	FP Rate	Precision
Red Neuronal	Selecciona	0,099	0,083	0,2
	NO Selecciona	0,917	0,901	0,829
Naive Bayes	Selecciona	<b>0,531</b>	0,408	0,214
	NO Selecciona	0,592	0,469	0,858
ID3	Selecciona	0,425	0,283	0,243
	NO Selecciona	0,717	0,575	0,854
Random Forest	Selecciona	0,123	0,0088	0,227
	NO Selecciona	0,912	0,877	0,833
AdaBoostM1	Selecciona	<b>0,642</b>	0,625	0,177
	NO Selecciona	0,375	0,358	0,592

Cuadro 2: Indicadores de la capacidad predictiva de cada modelo para la Opción B.

Donde:

TP Rate o Recall =  $Tp / (Tp + Fn)$

FP Rate =  $Fp / (Fp + Fn)$  - Mejor menor valor

Precision =  $Tp / (Tp + Fp)$

con:

Tp : Positivos verdaderos

Fp: Falsos positivos

Fn: Falsos negativos

Las matrices de confusión resultaron:

Red Neuronal:

a	b	<-- clasificado como
355	32	a = No selecciona
73	8	b = Selecciona

Naive Bayes:

a	b	<-- clasificado como
229	158	a = No selecciona
38	43	b = Selecciona

ID3:

a	b	<-- clasificado como
269	106	a = No selecciona
46	34	b = Selecciona

Random Forest:

```
a b <-- clasificado como  
353 34 | a = No selecciona  
71 10 | b = Selecciona
```

AdaBOostM1:

```
a b <-- clasificado como  
145 242 | a = No selecciona  
29 52 | b = Selecciona
```

A los objetos del presente trabajo lo que nos interesa es el valor de la TP Rate o Recall para la clase Selecciona, ya que este indicador nos dice que mientras mayor sea su valor (entre 0 y 1) mayor será la capacidad del clasificador de no dejar de señalar como positivos instancias que sí lo son, aún cuando haya mucho falsos positivos.

Decimos que este es el indicador que nos interesa porque lo que queremos es que el predictor instruya al tutor para que trabaje en la modificación de un modelo mental erróneo de los alumnos que van a optar por esta opción aún cuando cuando se seleccionen alumnos que no lo necesitan.

Por ejemplo, un indicador del recall de 0,6 nos dice que está detectando un 60% de los positivos.

Por esto es que para las restantes opciones mostraremos los resultados obtenidos por los modelos Naive Bayes y AdaBoostM1 que lograron los mejores valores de recall.

A continuación se muestra un cuadro con los resultados obtenidos para las otras 3 opciones analizadas.



			TP Rate	FP Rate	Precision
Opción C	Naive Bayes	Selecciona	0,636	0,667	0,068
		NO Selecciona	0,333	0,667	0,068
	AdaBoostM1	Selecciona	0,364	0,324	0,078
		NO Selecciona	0,676	0,636	0,933
Opción D	Naive Bayes	Selecciona	0,618	0,667	0,068
		NO Selecciona	0,333	0,364	0,924
	AdaBoostM1	Selecciona	0,669	0,392	0,669
		NO Selecciona	0,608	0,331	0,818
Opción E	Naive Bayes	Selecciona	0,52	0,385	0,139
		NO Selecciona	0,615	0,48	0,915
	AdaBoostM1	Selecciona	0,66	0,612	0,114
		NO Selecciona	0,388	0,34	0,905

Cuadro 3: Indicadores de la capacidad predictiva de cada modelo para las Opciones C,D y E.

## Análisis y conclusiones

Si bien los resultados obtenidos no son los óptimos permiten pensar que es posible el uso de clasificadores automáticos para que instruyan a tutores inteligentes sobre que temas reforzar de una manera personalizada.

Si miramos los valores obtenidos vemos que hemos desarrollado una capacidad de detección del orden del 60% de los alumnos que tienen un modelo mental equivocado. Es de destacar que el tipo de problema que se aborda y que se materializa a través de los datos, es de naturaleza muy ruidosa, dado que en la elección de una opción como respuesta correcta influyen muchos factores mas allá de los exámenes anteriores y las selecciones previas de los alumnos.

Se ve claramente que mientras mayor sea la proporción de alumnos que elijen una opción mayor es la capacidad relativa de predicción de los modelos utilizados. Esto se debe a que el ruido baja.

Un tema que queda para futuros estudios es la determinación de la correlación del contenido de la pregunta bajo análisis con los temas anteriores y de como eso influye en la capacidad de predicción de los modelos. Es de esperar que la calidad de los modelos sea mayor cuanto mayor sea la correlación temática.

## Bibliografía



*Díaz, L., Algorry, A., Eschoyez, M., Barto, C., Marangunic, R. (2013). "Actions towards the application of intelligent systems in computer education. IEEE Latin America Transactions", 11(1): 591-595. ISSN 1548-0992.*

*Laluf A., Saavedra L., Díaz L., Algorry A. Bartó C. (2015): "Descubrimiento de Patrones cognitivos en Evaluaciones de Informática basado en Explotación de Información". TE&ET 2015. Corrientes.*