



Consideraciones sobre la portabilidad de modelos predictivos de rendimiento académico

Susana Carabias López¹, Jose Luis Arroyo-Barrigüete¹,
Adolfo Hernández Estrada², Tomás Curto Gonzalez¹

¹ Universidad Pontificia Comillas

² Universidad Complutense de Madrid

Marco Teórico del estudio

Área de investigación

Modelos de predicción de rendimiento académico

- Learning Analytics
- Educational Data Mining
- Academic Analytics

Nivel

Implementación en grupo, curso, departamento. (LA, EDM)

Finalidad

Predicción temprana con el objetivo apoyar a los docentes en el diseño de intervenciones (LA)

Necesidad de adaptación a la metodología docente.

- Integra enseñanza directa y elementos de teoría constructiva.
- Presencial con apoyo de plataformas de aprendizaje.

Marco Teórico del estudio

Factores considerados

Evaluaciones previas en la asignatura y en cursos anteriores.

- Es sencillo acceder a la información, por lo que facilita la implementación de los modelos
- Las intervenciones basadas en los modelos pueden apoyarse en los principios de evaluación formativa.

Análisis de portabilidad

- El modelo sólo puede estimarse en grupos en los que la asignatura ha finalizado, en los cuales no se pueden tomar medidas para reducir el fracaso académico.
- La aplicación de los modelos para predecir siempre tendrá lugar en un grupo diferente a aquél en el que se estimó, lo que hace imprescindibles los estudios para ver si los modelos son transferibles entre grupos o estudios de portabilidad.
- La portabilidad se ha estudiado ampliamente cuando los factores son datos de plataformas de aprendizaje, pero no cuando son evaluaciones previas.

Descripción de los datos

Muestra

170 alumnos del grado en ADE, evaluados en dos asignaturas diferentes: Estadística Empresarial y Matemáticas Financieras.

Control de variables de confusión

- Rendimiento previo: misma nota de acceso a la universidad en todos los grupos (ANOVA)
- Nivel de exigencia: mismo examen final para todos los grupos

Elección de variables para evaluar portabilidad

	Grupo	Asignatura	Profesor	Idioma
GB (Estadística)				Grupo de referencia
GB (Mat. Financieras)	Sí			Sí
GR (Estadística)		Sí		Sí
GBil (Estadística)		Sí	Sí	
GR (Mat. Financieras)				Sí
Gbil (Mat. Financieras)				

Alejamiento del GB

Metodología

Modelo de predicción para el GB

En el Grupo Base (GB) ajuste de modelos tipo logit (aprobar/suspender) con VC de 0,5 y 0,9

Selección de los cuatro mejores modelos en función de:

- Precisión. Al menos regla de la mayoría
- Sensibilidad (aprobados correctamente predichos)
- Especificidad (suspensos correctamente predichos)

De especial importancia para la finalidad de intervención temprana en los casos de posible fracaso.

Predicción realizada al mes de comenzar las clases.

Análisis de portabilidad

Evaluación del rendimiento de los cuatro modelos elegidos en el resto de grupos

Resultados

Modelo de predicción para el GB

Se seleccionan cuatro modelos cuya precisión (porcentaje de casos bien clasificados) es superior al 83% tanto para VC =0,5, como para VC=0,9

- Valores entre 85% y 93% para VC=0,5
- Valores entre 83% y 90% para VC=0,9

Especificidad

Porcentaje de suspensos bien identificados

Variable más significativa dado que el objetivo es el de diseñar intervenciones para prevención del fracaso

En los cuatro modelos seleccionados, la especificidad es del 86%

Análisis de portabilidad

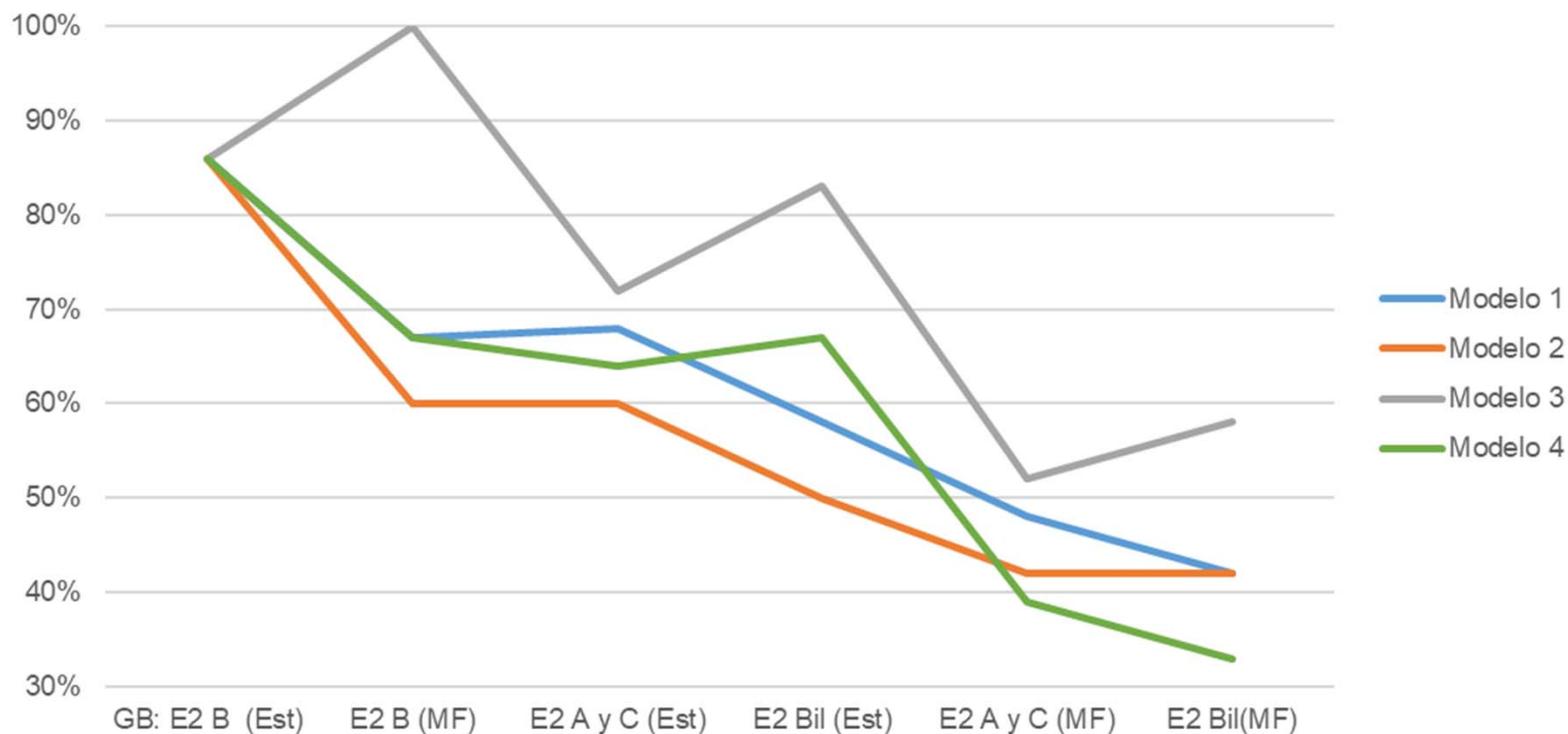
Se da prioridad a evitar la pérdida de especificidad, por lo que se seleccionan los resultados para VC=0,9

Especificidad

Empeoramiento consistente con lo esperado

Resultados

Especificidad de los cuatro modelos, para un VC de 0.9, en cada uno de los grupos



Empeoramiento consistente con lo esperado: al alejarse del GB (menos características en común), la especificidad tiende a disminuir en los cuatro modelos

Conclusiones

Modelo de predicción para el GB

Las evaluaciones previas son buenos predictores del fracaso académico en el grupo en el que se estima el modelo.

Análisis de portabilidad

No parece posible encontrar modelos de predicción temprana universales basados en evaluaciones previas.

Sensibilidad

La identificación correcta de los alumnos que tienen éxito es alta al aplicar el modelo en diferentes grupos.

Especificidad

La identificación correcta de los alumnos que no superan la asignatura, necesaria para intervenciones de prevención del fracaso, disminuye al aplicar el modelo en diferentes grupos.

Resultados consistentes

La portabilidad del modelo como predictor temprano de fracaso académico empeora cuando aumenta el número de características diferentes respecto del GB.