

# MACHINE LEARNING REGRESSION MODEL DEL ÍNDICE ÍNDICE ECOGRÁFICO DE ENTESIS MASEI EN PACIENTES CON ESPONDILOARTRITIS

David Castro Corredor <sup>1</sup>, Luis Ángel Calvo Pascual <sup>2</sup>, M.A. Ramírez Huaranga <sup>1</sup>, M. Paulino Huertas <sup>1</sup>.

<sup>1</sup> Servicio de Reumatología. Hospital General Universitario de Ciudad Real.

<sup>2</sup> Departamento de Métodos Cuantitativos. ICADE. Universidad Pontificia de Comillas.

Email: d.castrocorredor@gmail.com

## INTRODUCCIÓN

Los modelos de regresión de aprendizaje automático se construyen minimizando el error de generalización en la predicción y evitando al mismo tiempo el sobreajuste.

Estas técnicas mejoran los modelos y pronósticos de investigaciones médicas anteriores al lograr una mayor precisión que las técnicas econométricas tradicionales.

## MATERIAL Y MÉTODOS

Se realizaron investigaciones transversales, descriptivas y observacionales. Recopilamos datos de 24 pacientes con espondiloartritis que recibieron tratamiento en nuestras consultas durante 5 meses y se sometieron a una ecografía musculoesquelética utilizando el índice MASEI. Utilizando pruebas F-test y mutual information, reducimos las variables a las más importantes (figura 1). Finalmente, utilizamos el aprendizaje automático (machine learning) para estimar algunos modelos de regresión.

## OBJETIVOS

Predecir el Índice de Entesis Ecográfico de Madrid (MASEI) en pacientes con espondiloartritis utilizando las mejores variables predictoras, incluida la actividad de la enfermedad y otros factores.

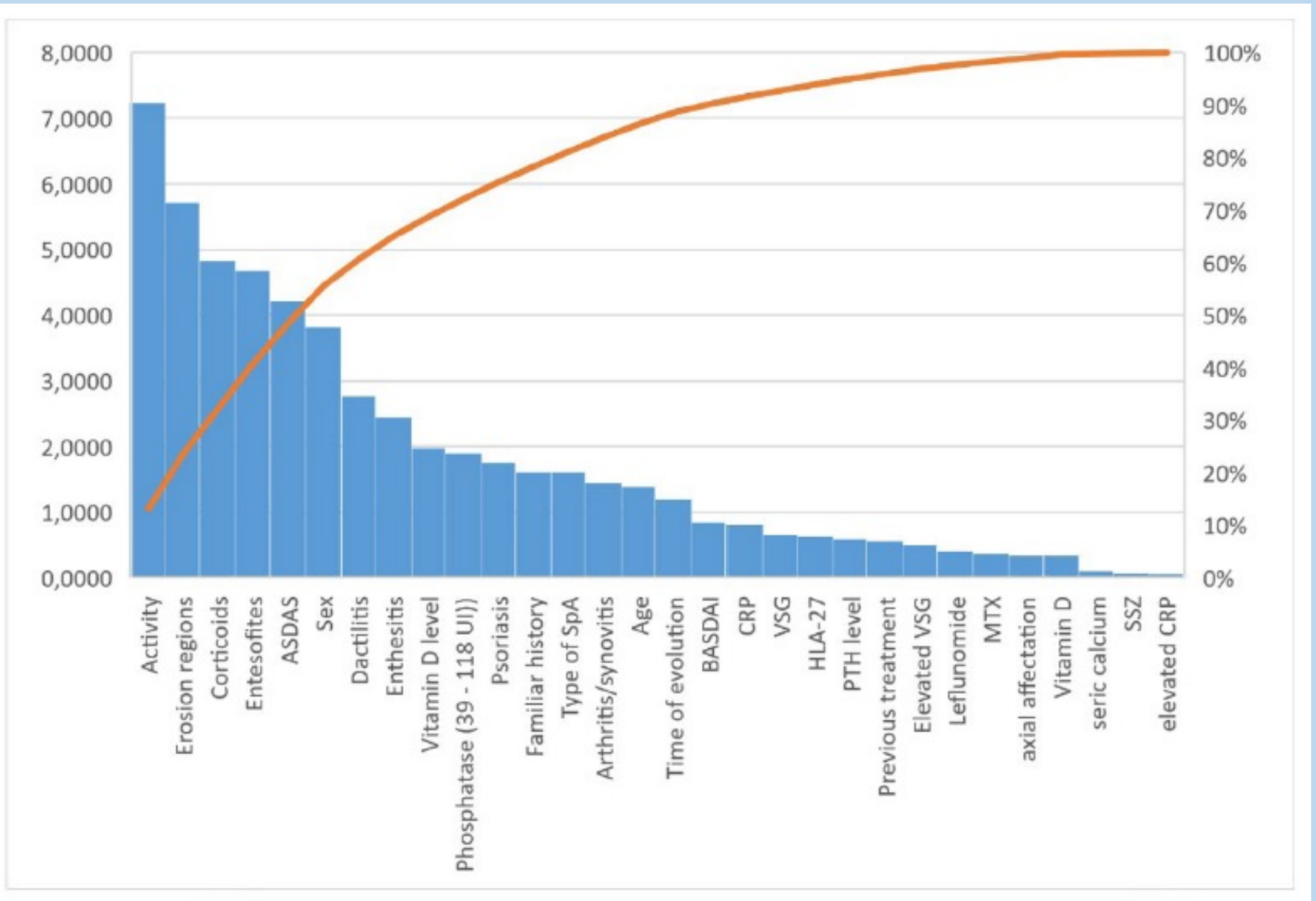


Figura 1. Diagrama de Pareto de las pruebas F del Índice MASEI con el Resto de las Variables

## RESULTADOS

Las variables predictoras con valores más altos en las pruebas de selección que explican el índice MASEI son la actividad, el puntaje de actividad de la enfermedad de espondilitis anquilosante (ASDAS), los corticosteroides, los entesofitos y el sexo masculino (figura 2).

El modelo de regresión más preciso fue el modelo de Support Vector Machine (R cuadrado = 0,81 en la validación). Dado que este modelo es una caja negra, también calculamos una regresión lineal clásica (R-Cuadrado=0,60 en la validación) porque proporciona un modelo explícito dado por:  $MASEI = -4,02 + 2,75 \text{Entesofitos} + 4,838 \text{ ASDAS}$  (figura 3).

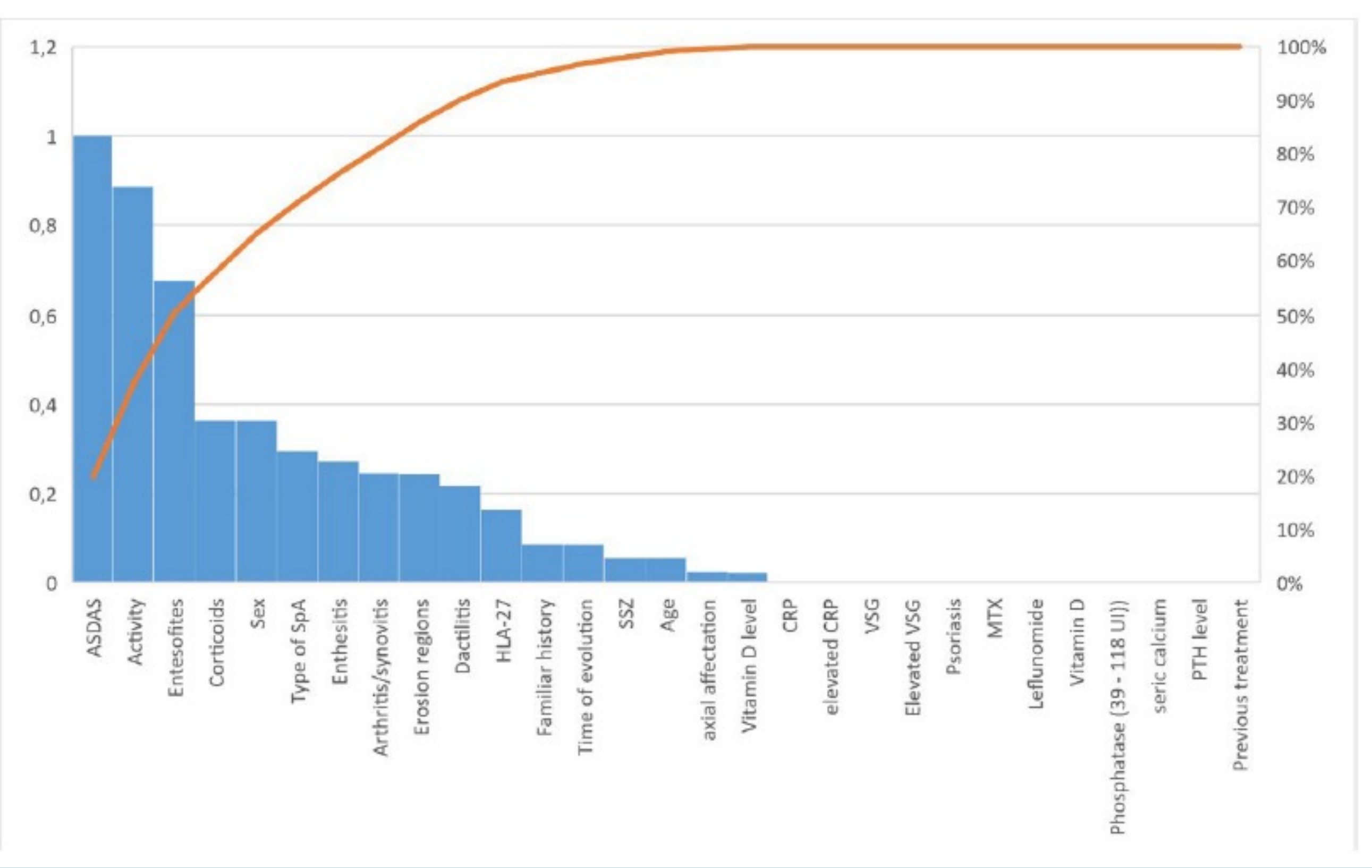


Figura 2. Diagrama de Pareto de la prueba de Información Mutua del Índice MASEI con el Resto de las Variables

## CONCLUSIONES

- Hasta la fecha se conocía la correlación entre el índice MASEI y la actividad de la enfermedad en pacientes con espondiloartritis.
- Para examinar más a fondo esta relación, predijimos 2 modelos regresivos para pronosticar el índice MASEI:
  - Un modelo de regresión lineal clásico
  - Y un modelo support vector machine con el error cuadrático medio estimado más bajo entre 32 modelos alternativos de aprendizaje automático.

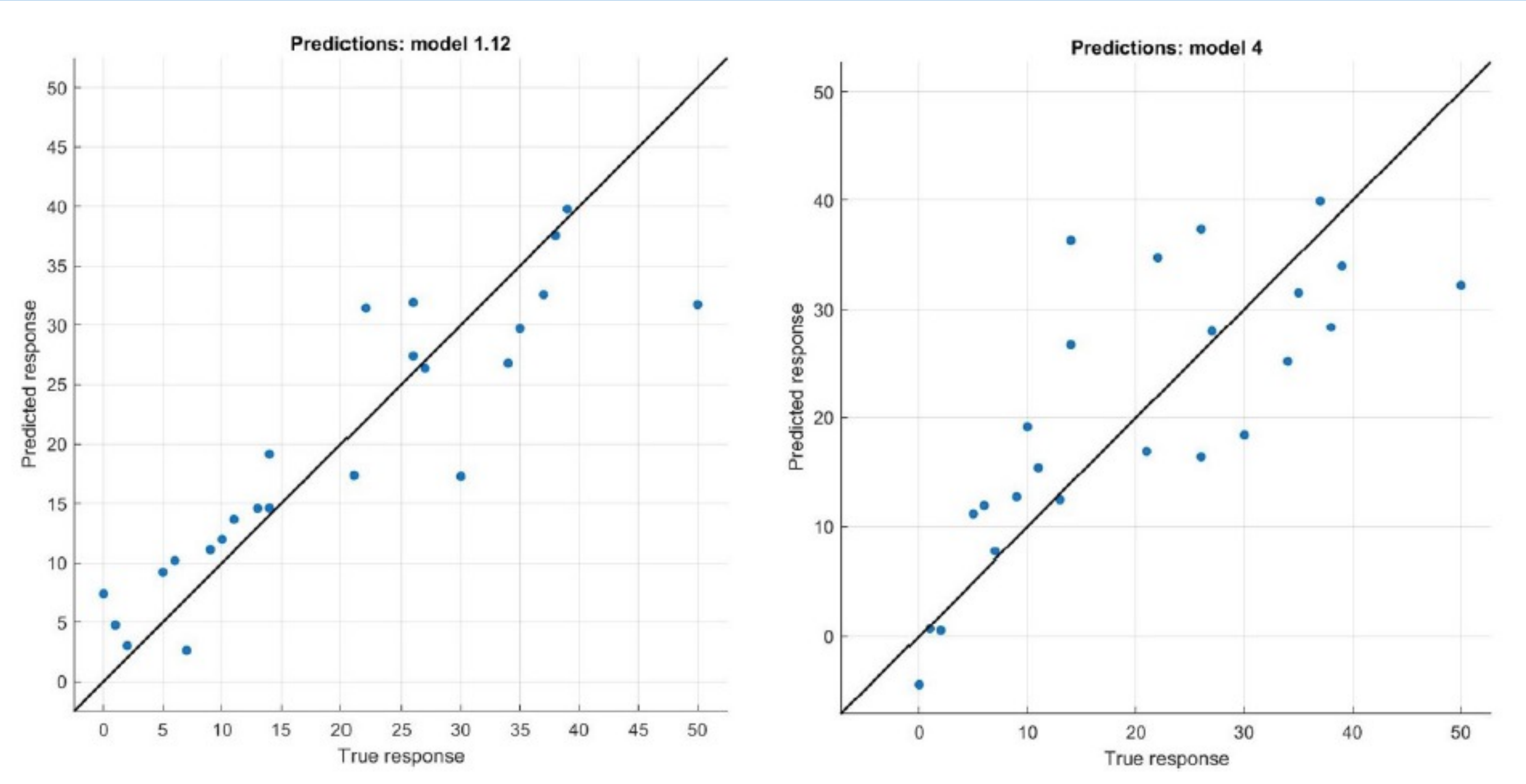


Figura 3. Respuesta verdadera versus prevista del modelo SVM y el modelo de regresión lineal

