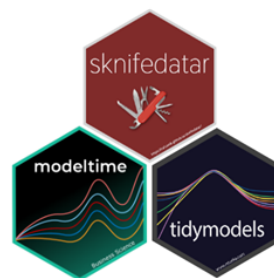


{sknifedatar}: Un paquete para el modelado y visualización de múltiples series de tiempo

Rafael Zambrano , Karina Bartolomé

Abstract Desarrollar análisis estadísticos o descriptivos sobre series temporales, requiere un flujo de trabajo ordenado, consistente y compacto. Esto es fundamental cuando se consideran múltiples series de tiempo, donde surge la necesidad de visualizar dichas series y ajustar múltiples modelos predictivos para cada una. Con este espíritu surge la creación del paquete {sknifedatar}, sirviendo como una extensión del ecosistema {tidymodels-modeltime} para el modelado de series de tiempo, factor diferenciador con respecto a otros paquetes similares de modelado, permitiendo extender su aplicación sobre otros sub-paquetes de tidymodels como {workflowsets} o funcionalidades de visualización sobre {rmarkdown}.. Desde el mes de junio de 2021, el paquete se encuentra disponible en The Comprehensive R Archive Network (CRAN).

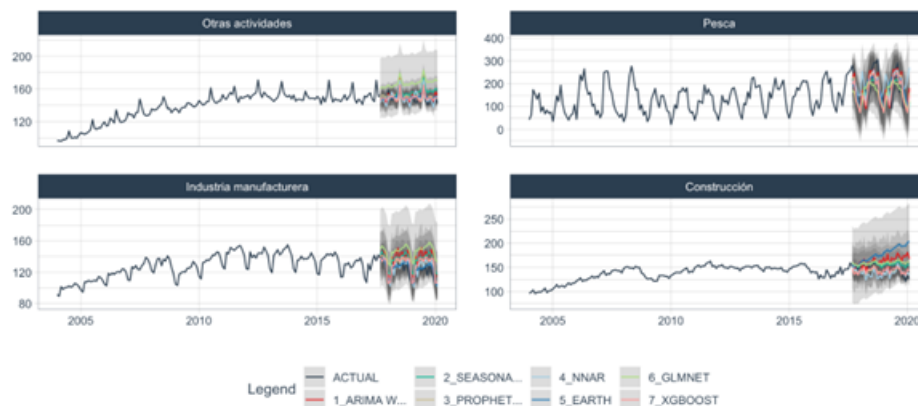
Palabras clave: series de tiempo - modeltime - rmarkdown - tidymodels - workflows - tabs



Estructura general

■ Múltiples modelos sobre múltiples series de tiempo (sin considerar datos de panel):

La principal contribución del paquete al ecosistema {tidymodels}¹ reside en ampliar las funcionalidades de modelado sobre series temporales, permitiendo ajustar múltiples modelos sobre múltiples series de tiempo. Para ello, se parte de la base de {modeltime}², un paquete destinado al modelado de series de tiempo con un enfoque *tidy*. Adicionalmente facilita la evaluación, reentrenamiento, diagnóstico y comparación de las múltiples combinaciones de modelos y series.

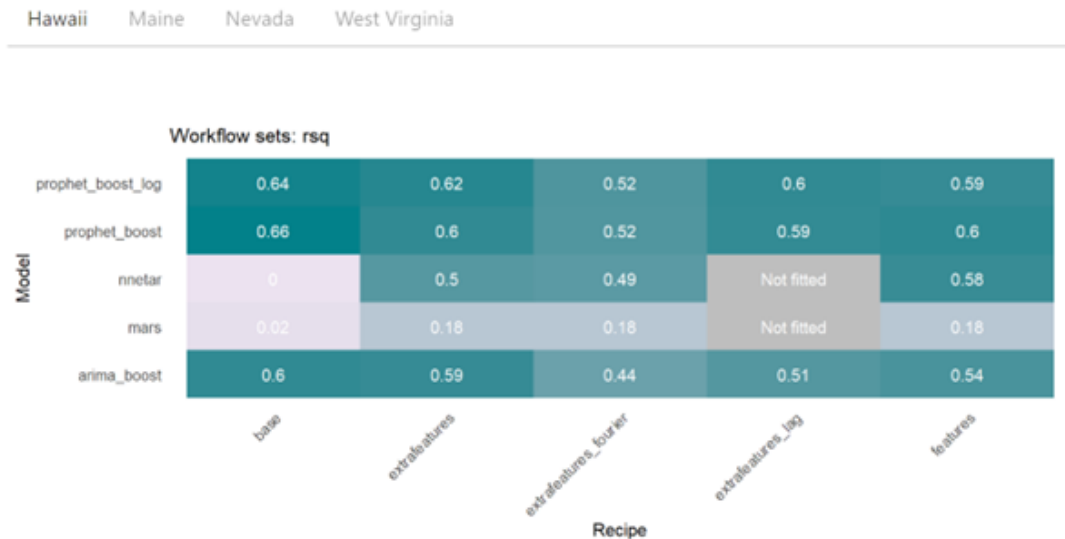


■ {workflowsets} extensión:

¹(Kuhn and Wickham 2020)

²(Dancho 2021)

{workflowsets}³ es un paquete que permite combinar múltiples procesos de ingeniería de características con sobre múltiples modelos dentro del ecosistema {tidymodels}, sin embargo, su aplicación directa no está diseñada para modelar series temporales. {sknifedatar}⁴ permite extender la aplicación de {workflowsets} sobre múltiples series de tiempo, mediante un flujo de trabajo ordenado para verificar las mejores combinaciones de modelos e ingenierías de características para cada serie.



■ Automagic Tabs:

La visualización de datos juega un papel central en el tratamiento de series temporales, dentro del paquete se incluyen funciones para generar automáticamente tabs (pestañas) en documentos “.Rmd” renderizados como html. La generación manual de tabs es posible, pero demanda el desarrollo de largas extensiones de código y es susceptible a errores. En este sentido, mediante metaprogramación automática se generan chunks de código que permiten visualizar gráficos, tablas o texto en múltiples tabs.

En esta presentación se exponen las funcionalidades del paquete {sknifedatar} sobre un caso de uso con datos reales, facilitando el uso y divulgación del ecosistema {tidymodels} en referencia al modelado de series temporales.

Referencias

- 10 Dancho, Matt. 2021. *Modeltime: The Tidymodels Extension for Time Series Modeling*. <https://github.com/business-science/modeltime>.
- Kuhn, Max. 2022. *Workflowsets: Create a Collection of Tidymodels Workflows*. <https://CRAN.R-project.org/package=workflowsets>.
- Kuhn, Max, and Hadley Wickham. 2020. *Tidymodels: A Collection of Packages for Modeling and Machine Learning Using Tidyverse Principles*. <https://www.tidymodels.org>.
- Zambrano, Rafael, and Karina Bartolome. 2021. *Sknifedatar: Swiss Knife of Data*. <https://github.com/rafzamb/sknifedatar>.

³(Kuhn 2022)

⁴(Zambrano and Bartolome 2021)