

Autor: Javier
Amado Bouza

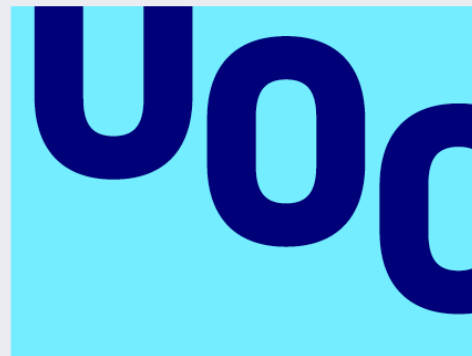
10/01/2021

Profesora: Núria
Pérez Álvarez

Responsable: Marc
Maceira Duch

Máster
Universitario de
Bioinformática y
Bioestadística

PREDICCIÓN DE TIEMPO DE FRACASO DE UN TRATAMIENTO ANTIRRETROVIRAL MEDIANTE ALGORITMOS DE APRENDIZAJE AUTOMÁTICO DE SUPERVIVENCIA



Universitat Oberta
de Catalunya

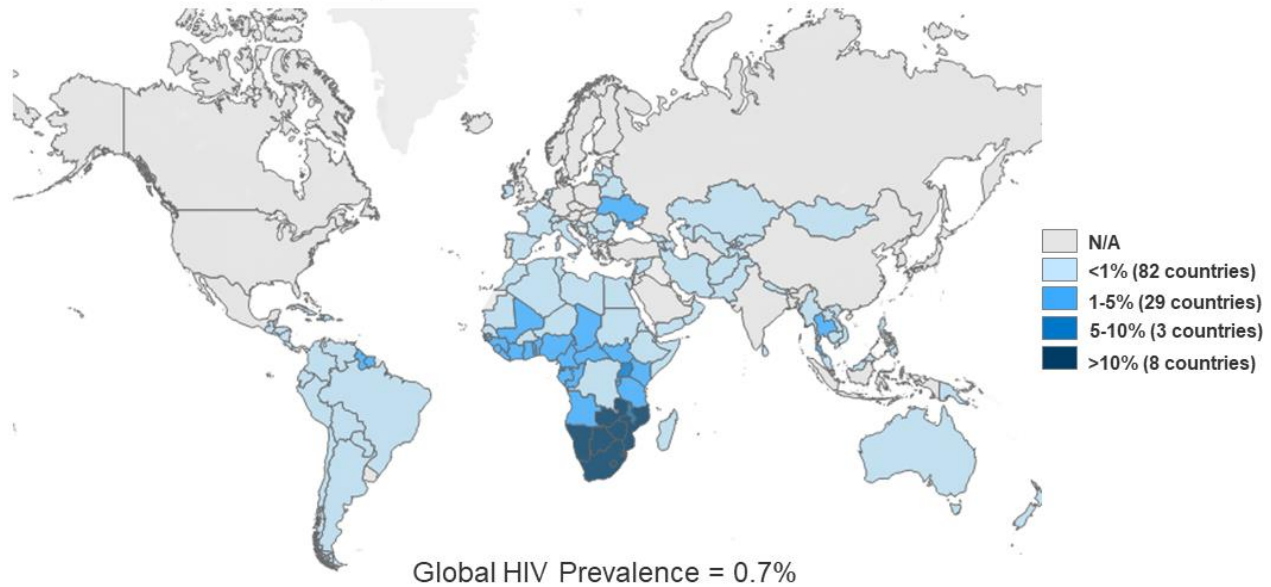
ÍNDICE:

- El VIH
- El ensayo Lake
- Gestión de datos
- La imputación
- Aprendizaje automático
- Desempeño de los algoritmos
- Comparación de tratamientos
- De cara al futuro
- Conclusiones

EL VIH

Figure 1

Adult HIV Prevalence, 2019



NOTES: Data are estimates. Prevalence includes adults ages 15-49.
SOURCE: KFF, based on UNAIDS, AIDSinfo, Accessed July 2020.



0,7%

De los adultos

38 millones

De seropositivos

1,7 millones

De nuevos casos

690.000

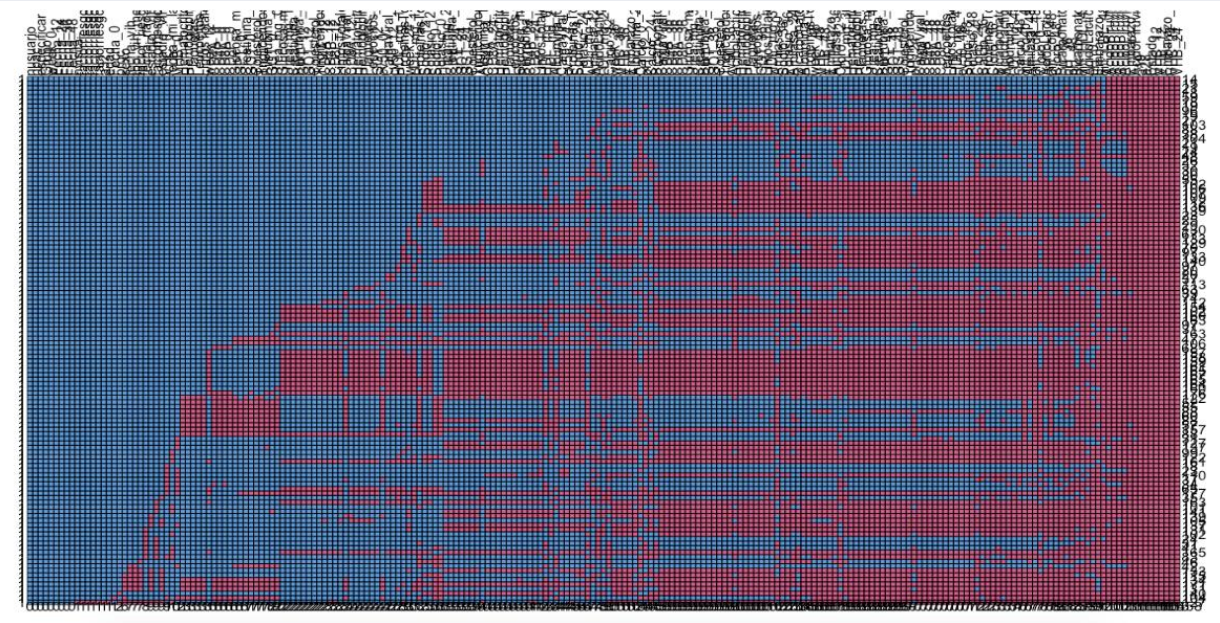
Muertes anuales

EL ENSAYO LAKE

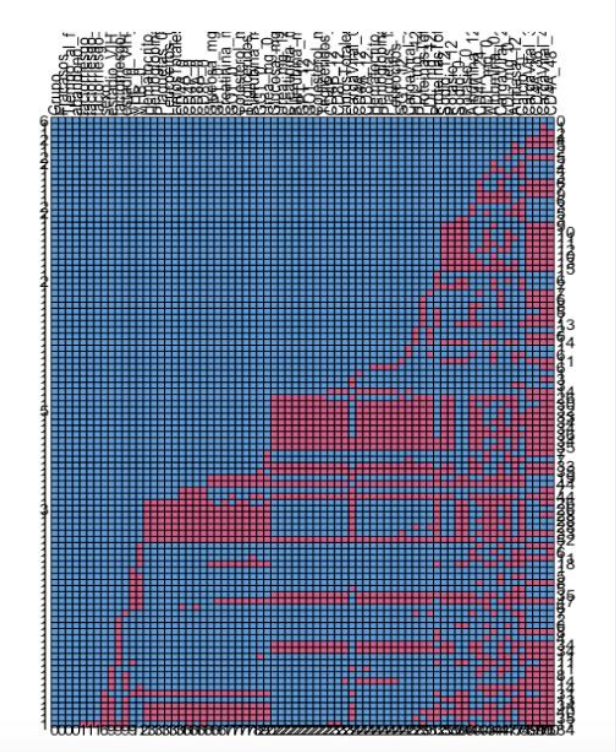


GESTIÓN DE DATOS

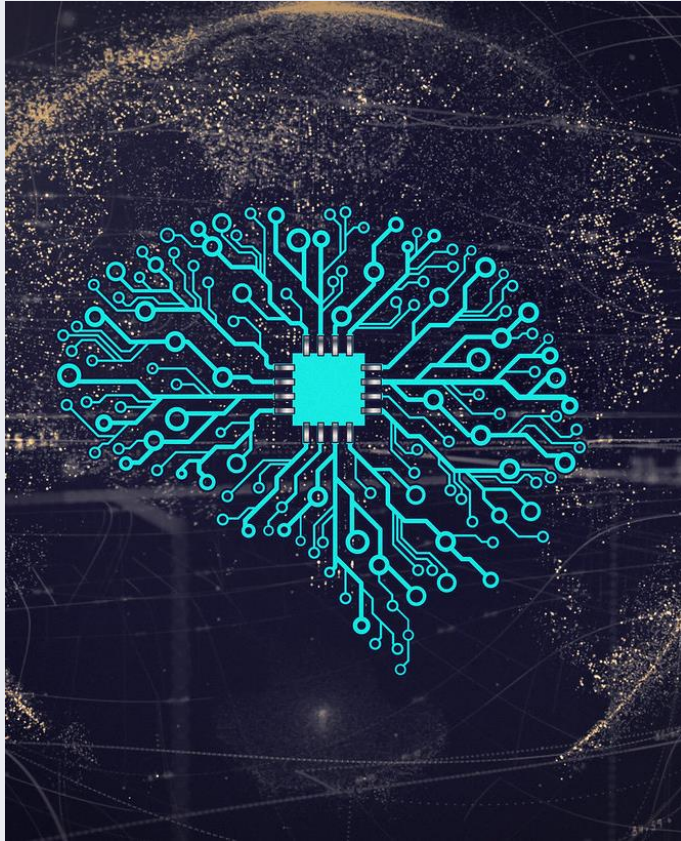
Datos Crudos



Datos procesados



APRENDIZAJE AUTOMÁTICO



- **Bosque aleatorio de supervivencia**
- **Máquina de soporte vectorial de supervivencia**
- **Boosting**

DESEMPEÑO DE LOS ALGORITMOS

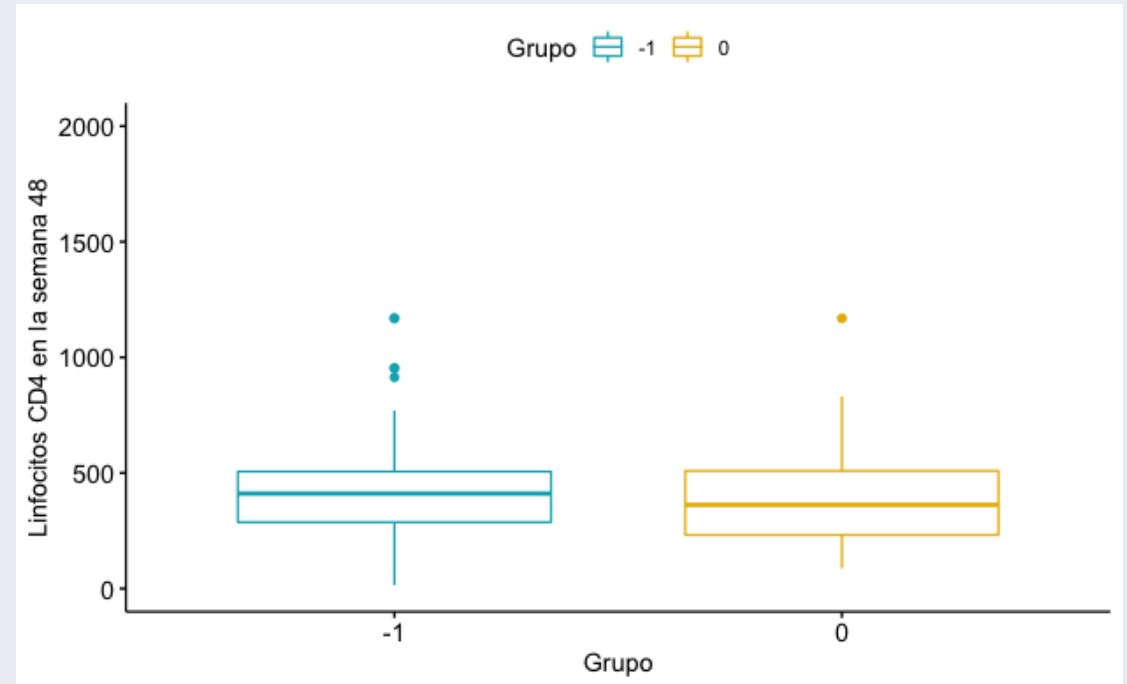
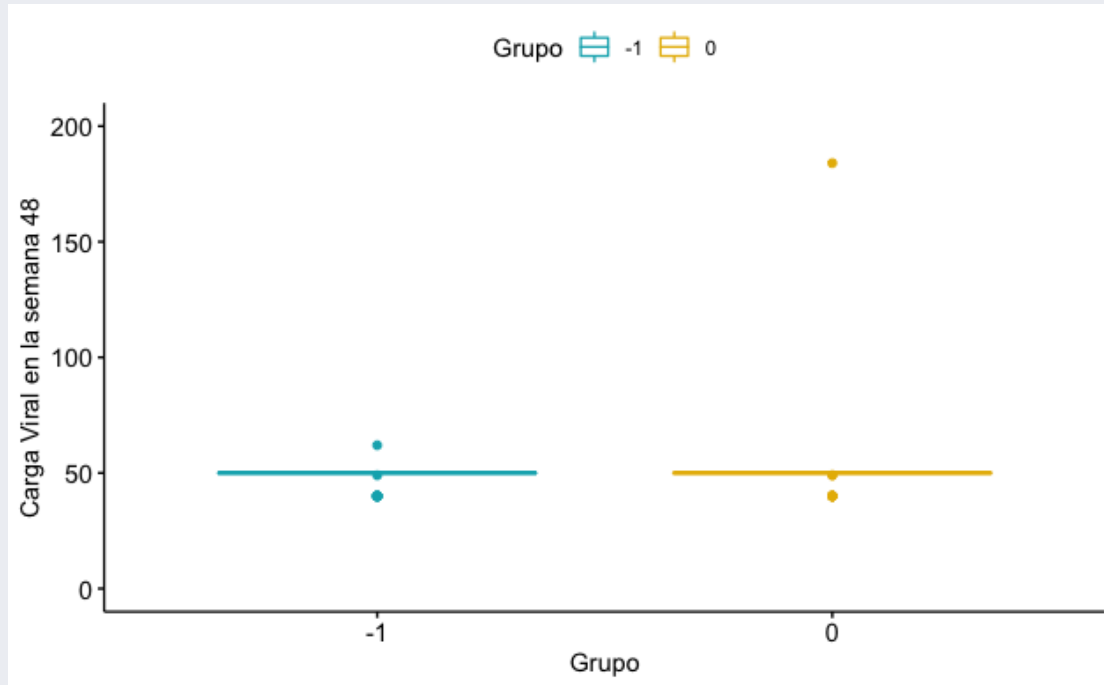
Obtención del índice C de Harrell

Índice	Bosque supervivencia	Máquina soporte vectorial	Boosting
C de Harrell	0.7960526	0.7333333	0.8

COMPARACIÓN DE TRATAMIENTOS

Carga viral final tratamiento

Linfocitos CD4 final tratamiento



DE CARA AL FUTURO

- Ensayos clínicos con menor cantidad de pacientes
- Predicción mediante aprendizaje automático
- Mayor calidad de vida en pacientes VIH+

CONCLUSIONES

- La imputación permite sacar más provecho a las BBDD.
- Tratamiento A es el mejor tratamiento
- Boosting es el mejor algoritmo
- Con pocos pacientes se obtiene buena predicción

BIBLIOGRAFÍA

- 1. Vinzamuri B, Li Y, Reddy CK. Active learning based survival regression for censored data. CIKM2014 - Proceedings of the 2014 ACM International Conference on Information and Knowledge Management. 2014;241–50.
- 2. Wang P, Li Y, Reddy CK. Machine learning for survival analysis: A survey. ACM Computing Surveys[Internet]. 2019;51(6):1–39. Available from: <http://arxiv.org/abs/1708.04649>
- 3. Taylor JMG. Random survival forests. Journal of Thoracic Oncology [Internet]. 2011;6(12):1974–5. Available from: <http://dx.doi.org/10.1097/JTO.0b013e318233d835>
- 4. Roy J, Lin X. Missing covariates in longitudinal data with informative dropouts: Bias analysis and inference. Biometrics. 2005;61(3):837–46.
- 5. Dziura JD, Post LA, Zhao Q, Fu Z, Peduzzi P. Strategies for dealing with missing data in clinical trials: From design to analysis. Yale Journal of Biology and Medicine. 2013;86(3):343–58.
- 6. Little RJA, Rubin DB. Statistical Analysis with Missing Data. 2002.
- 7. Calafati RO. Estrategias para el tratamiento de datos faltantes ("missing data") en estudios con datos longitudinales [Internet] [PhD thesis]. 2017. p. 82. Available from: <http://openaccess.uoc.edu/webapps/o2/bitstream/10609/64085/6/romancalafatiTFG0617memoria.pdf>
- 8. Buuren S van. Flexible Imputation of Missing Data, Second Edition [Internet]. Second edition. | Boca Raton, Florida : CRC Press, [2019] | : Chapman; Hall/CRC; 2018. Available from: <https://www.taylorfrancis.com/books/9780429492259>
- 9. Echeverría P, Negredo E, Carosi G, Gálvez J, Gómez JL, Ocampo A, et al. Similar antiviral efficacy and tolerability between efavirenz and lopinavir/ritonavir, administered with abacavir/lamivudine (Kivexa), in antiretroviral-naïve patients: A 48-week, multicentre, randomized study (Lake Study). Antiviral Research. 2010;85(2):403–8.
- 10. Medicine AB of I. ABIM Laboratory Test Reference Ranges - January 2020 [Internet]. 2020. pp. 1–12. Available from: <https://www.abim.org/%7B~%7D/media/ABIM%20Public/Files/pdf/exam/laboratoryreference-ranges.pdf>

IMÁGENES UTILIZADAS

- Imagen de la diapositiva 7 utilizada bajo licencia CC BY 2.0. Siendo el autor [Mike Mackenzie](#)