Nombre: Modelos Estadísticos Interpretables LCDA210062 **(LicDatos/Mat)**

Tópicos de Modelos Interpretables DOC8801300 **(Doc)**

Profesor: Dra. María Eugenia Szretter (nueva IC2024)

Puntaje: 4 puntos. Arancelado(DOC)

Correlatividades: Introducción a la estadística y Ciencia de datos **Carácter**: Optativo

Carga horaria: 2**4 hs totales** 8 semanas **Carreras:** LCD /MATE/ DOC

WEB: https://modelosinterpretables.wordpress.com/2024/02/16/modelos-estadisticos-interpretables/

Alumnos externos: La materia que ofreceremos es para el doctorado en Cs matemáticas y Cs de Datos, requiere como correlativa la asignatura obligatoria Introducción a la Estadística y Cs. de Datos o Estadística (matemáticos) y asume conocimientos solidos de álgebra lineal y de probabilidades (a nivel de la licenciatura en matemáticas o Cs. de datos).

PROGRAMA:

Motivación

En la actualidad existe un auge de algoritmos altamente eficientes desde el punto de vista predictivo.
Sin embargo, estos carecen en su mayoría de la capacidad de identificar los mecanismos inherentes que producen esas buenas predicciones. En contraposición los métodos estadísticos tradicionales explicitan esos mecanismos en forma clara, pero no poseen la capacidad predictiva de los algoritmos del machine learning. Este curso procura brindar herramientas estadísticas para el modelado supervisado, tanto en clasificación como en regresión, que privilegian la capacidad explicativa e inferencial en el trabajo con datos. Las técnicas que veremos son útiles como metodologías en sí mismas y a la vez como herramientas complementarias al uso de técnicas de modelado complejo (tipo caja negra) para aumentar la capacidad interpretativa.

Programa resumido: Modelo lineal múltiple. Interpretación de distintos predictores: cuantitativos, cualitativos, interacción entre ellos y transformaciones polinomiales de las predictoras. ANOVA un factor. Mínimos cuadrados pesados. Estimación robusta. Técnicas de regularización. Selección del parámetro de suavizado usando validación cruzada. Selección de modelos. Compromiso sesgo varianza. Árboles de clasificación y regresión. Métricas involucradas. Criterio de parada. Poda y penalización por complejidad. Modelos predictivos usando el paquete CARET. Bagging, Random Forest y Boosting. Introducción a modelos lineales generalizados. Modelos lineales de efectos mixtos.
Interpretación de modelos complejos: metodologías SHAP (Shapley Additive exPlanations) y LIME (Local Interpretable Model-gnostic Example).

Programa Analítico

1. Modelo lineal múltiple. Revisión de los conceptos básicos de Modelo Lineal. Análisis de casos especiales: interpretación del modelo y ajuste bajo casos particulares de covariables, predictores cualitativos con dos categorías y con más de dos categorías, interacción entre distintos tipos de variables, transformaciones polinomiales de las covariables, test para comparar dos medias como caso particular y el modelo de análisis de la varianza (ANOVA) de un factor como otro caso particular.
2. Diagnóstico de problemas potenciales del ajuste. Gráficos y otras herramientas de diagnóstico para detectar problemas de ajuste. Soluciones para la falta de linealidad y/o heteroscedasticidad. Mínimos cuadrados pesados. Estimación robusta. Comparación con mínimos cuadrados ordinarios. Colinealidad.

3. Selección de Modelos. Criterios para elegir el modelo óptimo. Muestra de entrenamiento y validación. Medidas de ajuste: coeficientes de determinación R^2-ajustado, estadístico de allows Cp, AIC y BIC. Selección de modelos por remuestreo: validación cruzada y bootstrap. Compromiso sesgo-varianza.
4. Técnicas de regularización. Métodos de Penalización. Comparación de distintas penalizaciones entre sí y con mínimos cuadrados ordinarios. Selección del parámetro de suavizado usando validación cruzada. Interpretación de la propiedad oráculo.
5. Árboles de clasificación y regresión. Métricas involucradas. Criterio de parada. Poda y penalización por complejidad. Arboles vs. Modelo lineal. Modelos predictivos usando el paquete CARET (Classification And REgression Training): conjunto de funciones para ajustar modelos predictivos. El paquete contiene herramientas para el pre-procesamiento de los datos, ajuste mediante remuestreo y la estimación de la importancia de cada variable.
Métodos de agregación: Bagging, Random Forests y Boosting.
6. Generalización del modelo lineal. Introducción a modelos lineales generalizados. Modelos lineales de efectos mixtos. Efectos aleatorios. Ajuste del modelo de efectos aleatorios.
Evaluación del modelo ajustado. Cómo evaluar los términos de efectos fijos. Modelos de efectos mixtos para mediciones repetidas. Predicciones de la respuesta y de los efectos aleatorios.
7. Interpretación. ¿Cómo y cuánto aportan las variables incluidas en un modelo a su interpretación? Metodología SHAP (Shapley Additive exPlanations) para intentar dar una explicación a la predicción que hacen modelos complejos (de "caja negra"). Metodología LIME (Local Interpretable Model-gnostic Example). Introducción de la función de pérdida local y medidas de complejidad del modelo explicativo.

Bibliografía

- Draper, N. R., y Smith, H. (1998). Applied Regression Analysis, Third Edition. Wiley
- series in probability and statistics.
- Seber, G. A. F., Lee A. J. (2003). Linear Regression Analysis, 2nd Edition, Wiley Series
- in Probability and Mathematical Statistics.
- Kutner, M. H., Nachtsheim, C., Neter, J., y Li, W. (2005). Applied linear statistical
- models. McGraw-Hill Irwin.
- James, G., Witten, D., Hastie, T., y Tibshirani, R. (2013). An introduction to statistical
- learning (Vol. 112). Springer.
- Weisberg, S. (2005). Applied linear regression (3ra. edición). John Wiley & Sons.
- Hansen, Bruce E. (2022) Econometrics. Princeton University Press.
- Pinheiro J.C., Bates D. M. (2000). Mixed-Effects Models in S and S-PLUS. Springer,
- New York.
- Stasinopoulos, M. D., Rigby, R. A., Heller, G. Z., Voudouris, V., De Bastiani, F. (2017).
- Flexible regression and smoothing: using GAMLSS in R. CRC Press.
- Kuhn, M., Johnson, K. (2018). Applied predictive modeling. Springer.
- Molnar, C. (2022). Interpretable Machine Learning: A Guide for Making Black Box
- Models Explainable (2nd ed.). christophm.github.io/interpretable-ml-book/
- Lundberg, S. M., Su-In L. A unified approach to interpreting model predictions.
- Advances in Neural Information Processing Systems (2017).
- Ribeiro, M. T., Singh, S., Guestrin, C. Why should I trust you?: Explaining the
- predictions of any classifier. Proceedings of the 22nd ACM SIGKDD international
- conference on knowledge discovery and data mining. ACM (2016).

2do. Bim. 2024