



Modelos de clasificación de perfiles topográficos para la industria forestal

Camila Escobar Troncoso
Profesor Guía: Nora Serdyukova
Profesor Colaborador: Leonardo Padilla
Profesor Consejero: Martín Pincheira

Universidad de Concepción
Facultad de Ciencias Físicas y Matemáticas
Departamento de Estadística
Ingeniería Estadística



Índice de contenidos

1 Contexto Forestal

2 Problema

3 Objetivos

4 Metodología

- Regresión logística
- Machine Learning

5 Datos

- Flujo de trabajo

6 Plan de trabajo



Actividad forestal

Se trata de una práctica que implica la transformación de una zona donde no hay árboles para crear nuevos bosques. Con estas acciones se consiguen componentes de la madera que funcionan como materia prima para fabricar ciertos productos.



Figura: Tractor en el bosque cargado con troncos de madera.



Ciclo forestal

El ciclo forestal corresponde a todas las etapas consideradas por la actividad forestal para llevar a cabo el proceso de un producto maderero.

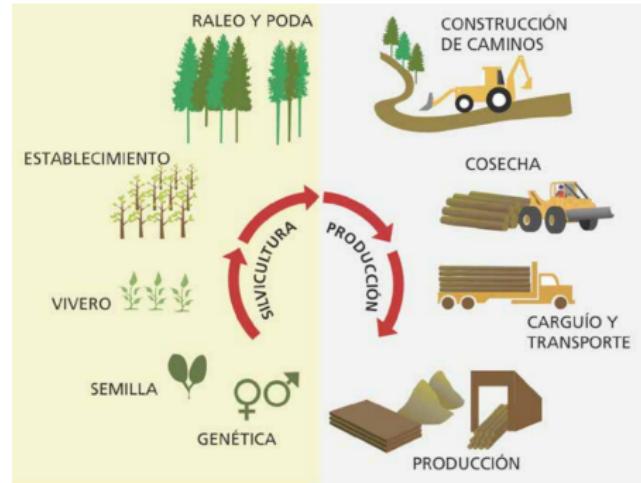


Figura: Ciclo forestal de una plantación forestal.



Cosecha

Una de los etapas más importantes de la actividad forestal es la cosecha. En esta se ve reflejado el esfuerzo y resultado de todo el periodo productivo y de las estrategias utilizadas. Es por esta razón que la planificación de la cosecha es primordial, ya que determinará los resultados del negocio.



Cosecha manual

Históricamente, la cosecha manual ha sido la principal forma de recolección en la actividad forestal. Un ejemplo clásico es el uso de bueyes para realizar la tala a ras del suelo. Esta técnica tradicional refleja la dependencia de herramientas básicas y la fuerza animal o humana para llevar a cabo la tarea.



Figura: Extracción de trozas largas en cosecha a tala rasa



Cosecha Mecanizada

Con la llegada de la tecnología, la cosecha mecanizada ha transformado significativamente la industria forestal. Esta técnica utiliza maquinaria avanzada para realizar la cosecha de manera más rápida y eficiente.



Figura: Tractor shovel volteando un árbol



Cosecha Mecanizada: Costo v/s Beneficio

La cosecha mecanizada requiere una inversión considerable en maquinaria y herramientas avanzadas. Aunque el costo inicial puede ser alto, los beneficios a largo plazo suelen superar esta inversión. La mecanización aumenta la eficiencia y mejora la seguridad en el trabajo. En muchas situaciones, dependiendo de la extensión de la superficie, esta inversión se amortiza rápidamente. Implementar tecnologías mecanizadas puede resultar en una mayor cantidad de materia prima recolectada y una cosecha más rápida y efectiva.



Forestal Arauco

Zonas Forestales: distribuidas entre las regiones del Libertador Bernardo O'Higgins (VI), Maule (VII), Ñuble (XVI), Biobío (VIII), Araucanía (IX), Los Ríos(XIV) y Los Lagos (X).

Zona Constitución



Zona Chillán



Zona Arauco



Zona Valdivia



Figura: Patrimonio forestal de Arauco S.A.



Terreno de cosecha



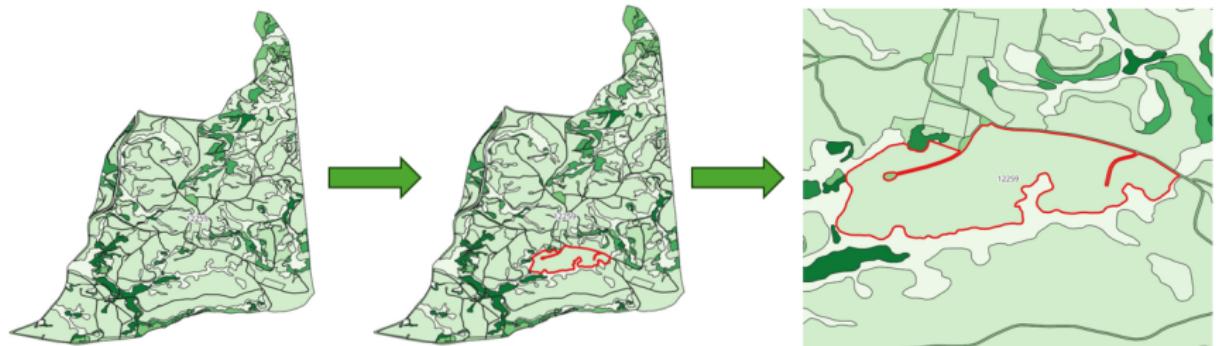
Figura: Predio Forestal

Muchas de las plantaciones de forestales de ARAUCO se encuentran ubicadas en sectores donde la topografía del terreno supera el 15 % de pendiente y el acceso para máquinas de uso terrestre pueden llegar a ser una alternativa muy costosa y en algunos casos su utilización, desde un punto de vista técnico, es imposible.



Escenarios de cosecha

Los predios forestales son conformados por rodales, los cuales son unidades homogéneas en cuanto al tipo de especie plantada y año de plantación. En base a esto es como se planifican las plantaciones y posteriores cosechas. Finalmente, un escenario de cosecha corresponde al área del rodal que se cosechará.





Elementos de un escenario





Elementos de un escenario



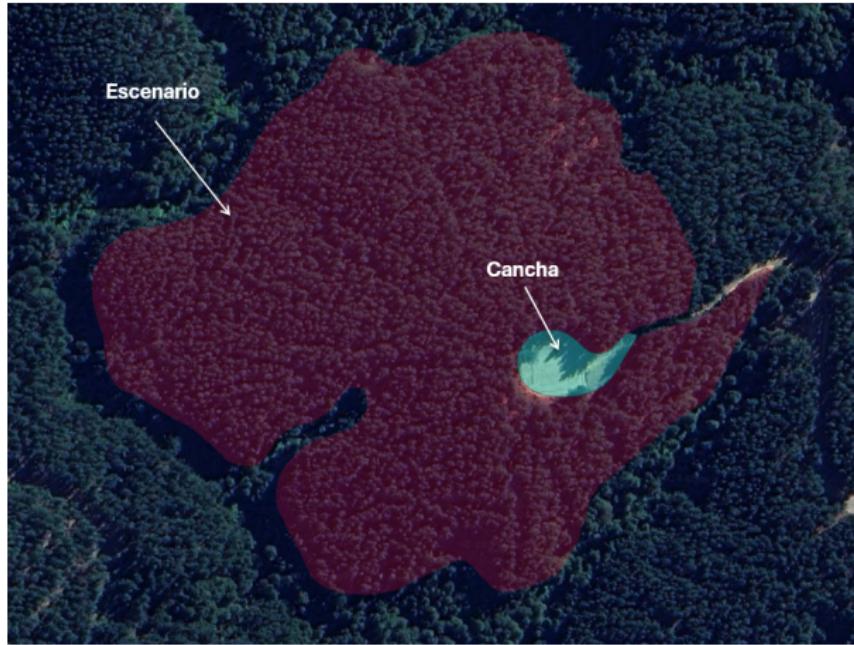


Elementos de un escenario



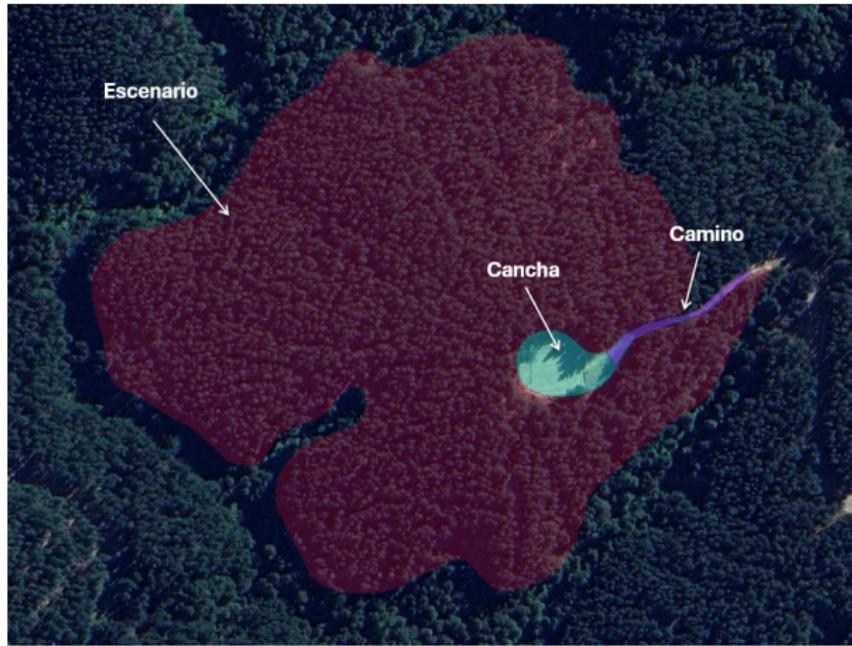


Elementos de un escenario





Elementos de un escenario





Perfil del escenario

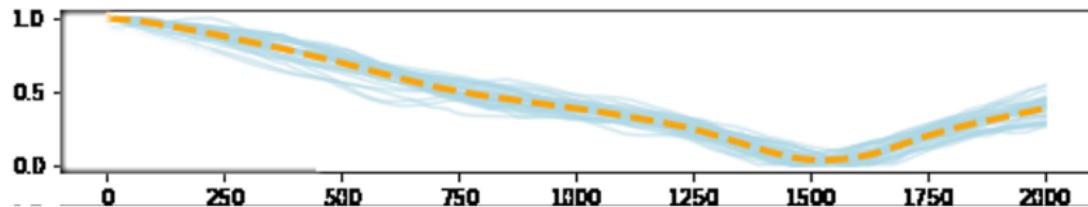


Figura: Perfil topográfico del escenario



Perfiles topográficos

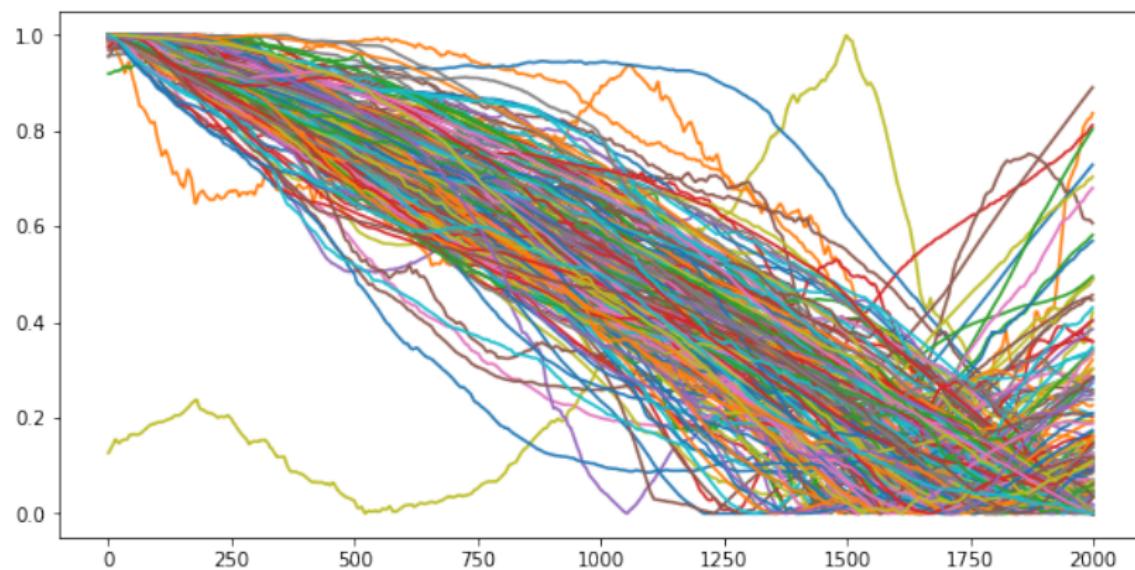


Figura: Tipos de perfiles topográficos



Madereo con torre



En estas condiciones de relieve, donde la pendiente supera el 15 % de pendiente la torre de madereo es la única alternativa técnica y económicamente viable para realizar las faenas de cosecha.

Figura: Torre de madereo aéreo



Cosecha Línea Directa

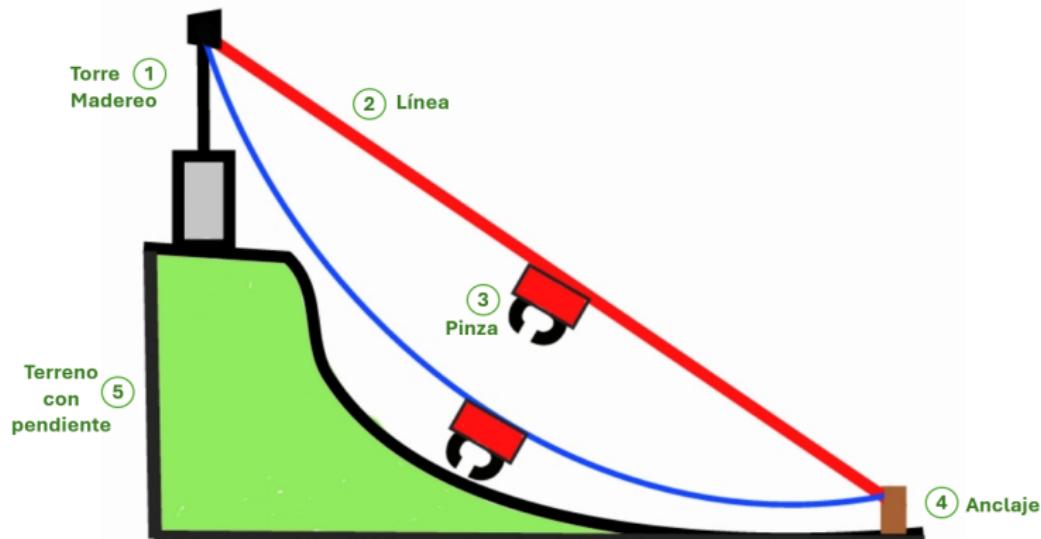


Figura: Línea de cosecha directa



Cosecha Soporte Intermedio



Figura: Línea de cosecha con soporte intermedio



Versus

Entre estos dos tipos de cosecha área, el más óptimo es el de línea directa, puesto que:

- ✓ No utiliza soporte intermedio
- ✓ Tiempos de cosecha más rápidos
- ✓ Más eficiente



Clasificación de escenarios

Bajo este contexto se busca clasificar los escenarios en los cuales es factible utilizar la cosecha con torre y aquellos en los que no, con el objetivo de optimizar los recursos y maximizar la eficiencia en la cosecha forestal.



Clasificador de líneas - ARAUCO

En la actualidad existe una metodología local que permite la selección de línea de cosecha bajo un modelo determinístico, sin embargo el proyecto no ha sido puesto en producción por diversas razones.

Al día de hoy se está retomando la iniciativa con la perspectiva de aplicar un enfoque estadístico.



Objetivo General

- Desarrollar e implementar un modelo de clasificación de escenarios de cosecha de acuerdo a la factibilidad de trazado de línea directa.



Objetivos Específicos

- Evaluar y comparar el rendimiento del modelo de clasificación de perfiles con modelos estadísticos.
- Desarrollar el modelo haciendo uso del lenguaje de programación Python.



Modelos de Clasificación

Ante esta problemática, se plantea la utilización de modelos de clasificación como una herramienta fundamental para identificar qué escenarios son factibles para utilizar torre de madereo.

Se propone emplear:

- Modelo Clásico: Regresión logística.
- Machine Learning: Random Forest, XGBoost



Modelo de regresión logística

Se busca ajustar un modelo de regresión logístico

$$\text{Logit} \{ E[Y | \tilde{X}] \} = \beta^T \tilde{X} \quad (1)$$

donde se deben identificar los predictores clave que influyen en el resultado. En este caso, se define una serie de variables que caracterizan los perfiles:

$$\tilde{X} = \begin{pmatrix} \text{Distancia Torre - Anclaje} \\ \text{Tipo superficie anclaje} \\ \text{Pendiente anclaje} \\ \text{Pendiente perfil} \\ h_1 \\ h_2 \\ \vdots \\ h_{500} \end{pmatrix}$$



Random Forest

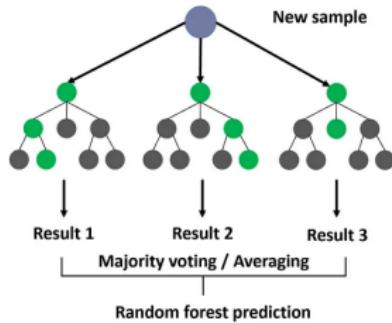
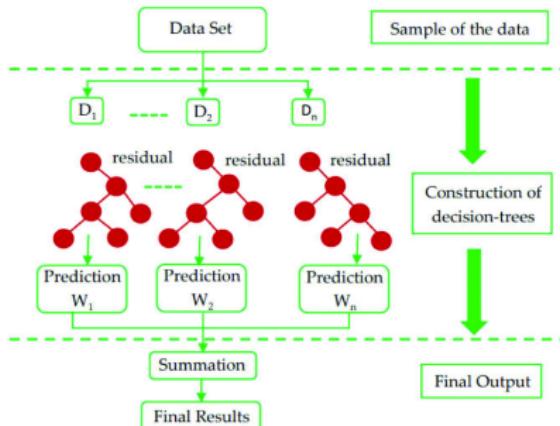


Figura: Estructura algoritmo
Random Forest

Cada árbol en el bosque evalúa un subconjunto aleatorio de las características de las muestras de entrenamiento y tiene una regla en cada nivel del árbol que clasifica según estas características aleatorias.



XGBoost



XGBoosting es una técnica que se basa en la construcción de múltiples árboles de decisión de forma secuencial para mejorar la precisión de las predicciones.

Figura: Estructura algoritmo XGBoost



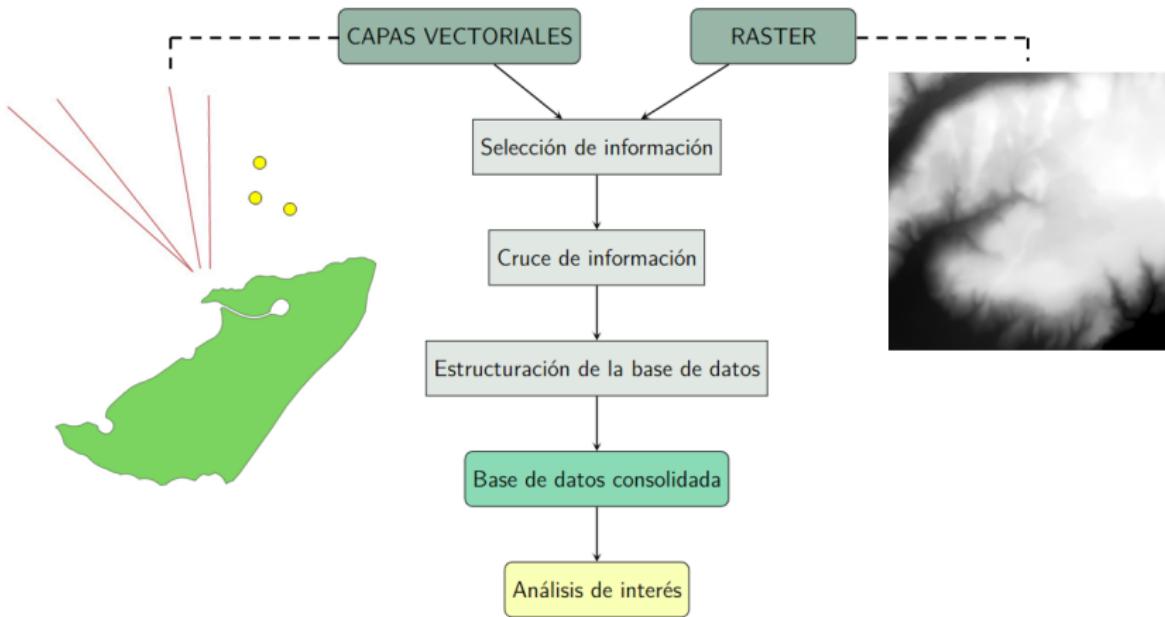
Datos

Para este proyecto, se utilizarán capas de información georreferenciada, las cuales contiene información del patrimonio forestal de FORESTAL ARAUCO S.A.

Estas capas se procesarán y analizarán utilizando el software QGIS. Para la consolidación, procesamiento y modelamiento de la base de datos se utilizará el lenguaje de programación Python.



Datos





Carta Gantt

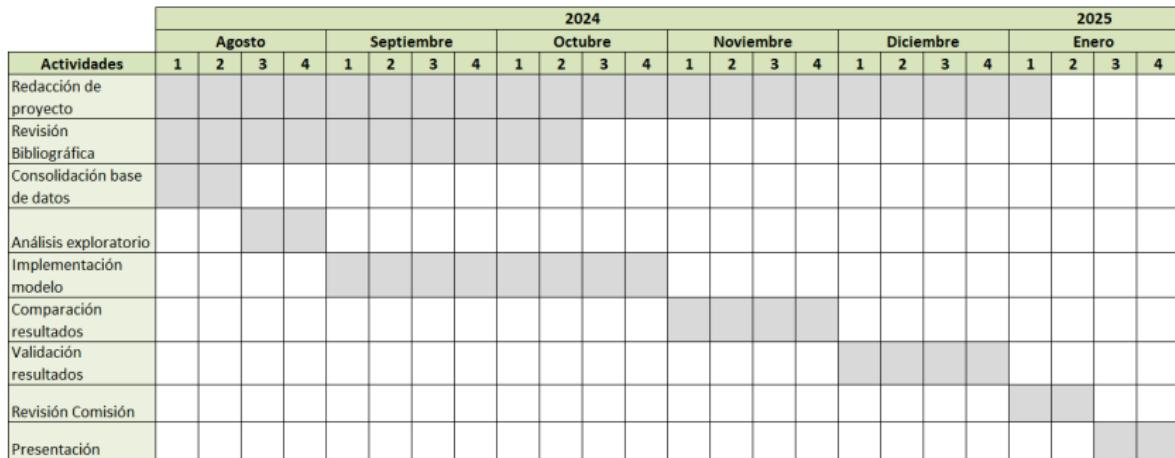


Figura: Programación del avance mediante una carta Gantt.



Referencias I

Agresti, A. (2018). *Foundations of linear and generalized linear models*. John Wiley & Sons.

Beyer, J. (1991). Modelo de apoyo a la toma de decisiones respecto a la localización de canchas de madereo en conjunto con el trazado de caminos forestales. In *Actas III Taller de Producción Forestal*, page 9, Concepción, Chile. Fundación Chile-Grupo de Producción Forestal.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1984). *Classification and regression trees*.

Carrasco, S. (1984). Análisis del madereo mediante una torre koller-300 (live sky line sistem).



Referencias II

- Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 785–794.
- Conway, S. (1982). *Logging Practices: Principles of Timber Harvesting Systems*. San Francisco, USA.
- Cox, D. R. (1958). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 20(2):187–220.
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5):1189–1232.



Referencias III

Gayoso, J. and Acuña, M. (1999). *Mejores Prácticas de Manejo Forestal. Guía de Campo.* Facultad de Ciencias Forestales, Universidad Austral de Chile, Valdivia, Chile.

Genuer, R. and Poggi, J.-M. (2020). *Random Forests with R.* Springer Cham.

Giordano, G. (1959). *Logging Cableways.* United Nation Publications, Génova.

Halevy, A., Norvig, P., and Pereira, F. (2009). The unreasonable effectiveness of data. *IEEE Intelligent Systems*, 24(2):9–12.

Harrell, Frank E., J. (2015). *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis.* Springer.



Referencias IV

- Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2nd edition.
- Hosmer Jr, D. W., Lemeshow, S., and Sturdivant, R. X. (2004). *Applied logistic regression*. John Wiley & Sons.
- Jordahl, K. (2016). Geopandas documentation. URL: <http://sethc23.github.io/wiki/Python/geopandas.pdf-Download vom, 26:2022>.
- Largo, S. (1985). Principios y técnicas de extracción forestal con cables. Boletín de Extensión 9, Facultad de Ciencias Agrarias y Forestales, Universidad de Concepción, Concepción, Chile.



Referencias V

- Larraín, O., Becker, J., and Garrido, G. (1989). Evaluación del madereo con torres. In *II Taller de Producción Forestal*, page 13, Concepción, Chile. Fundación Chile-Grupo de Producción Forestal.
- Lou, H., Lian, B., and Hageman, M. J. (2021). Applications of machine learning in solid oral dosage form development. *Journal of Pharmaceutical Sciences*, 110(9):3150–3165.
- McCulloch, W. S. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5(4):115–133.
- Natekin, A. and Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in neurorobotics*, 7:21.



Referencias VI

- Ruckstuhl, A. (2010). *Introduction to nonlinear regression.* IDP Institut fur Datenanalyse und Prozessdesign, Zurich University of Applied Sciences, Winterthur, Switzerland.
- Sahin, E. K. (2020). Assessing the predictive capability of ensemble tree methods for landslide susceptibility mapping using xgboost, gradient boosting machine, and random forest. *SN Applied Sciences*, 2(7):1308.
- Sapunar, P., Mansilla, G., and Fuentealba, G. (1999). Sistema cartográfico de apoyo a la planificación y habilitación de la cosecha. *XIII Silvotecnia*.
- Troncoso, J. (1996). Pladec: Un modelo para la planificación de sistemas de cosecha en pinus radiata (d. don).



Referencias VII

- Urra, G. (1999). Estudio de tiempo, rendimiento y costo para los equipos de madereo mecanizado: logger's dream y skidder.
- Valdebenito, G. and Neuenschwander, R. (1995). Evaluación y selección de número de canchas en función de costos mínimos de madereo. *Actas Seminario Internacional. Sistema de Producción Forestal: Decisiones y Técnicas. Universidad de Talca. Talca-Chile.* pp, pages 85–93.
- Vergara, M. (2000). Evaluación del rendimiento de torres de madereo en faenas a tala rasa.
- Wackerman, A., Hagenstein, W., and Michell, A. (1966). *Harvesting timber crops.* McGraw-Hill, New York, USA.