

PROYECTO FINAL *PRICING*

**CASOS DE ESTUDIO APLICADOS AL SECTOR
SEGUROS Y FINANCIEROS EN COLOMBIA**

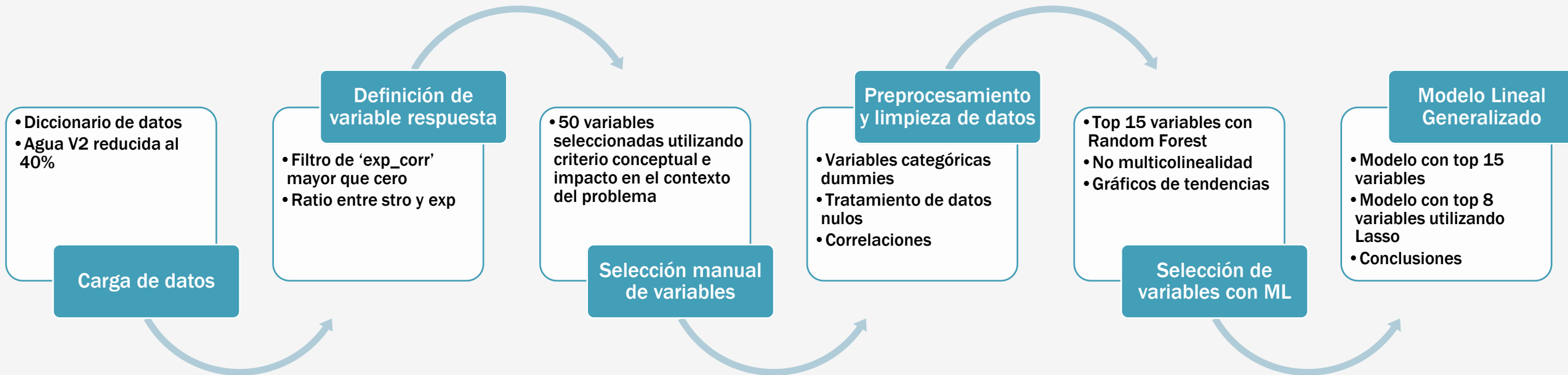
Elizabeth Cano Mejia



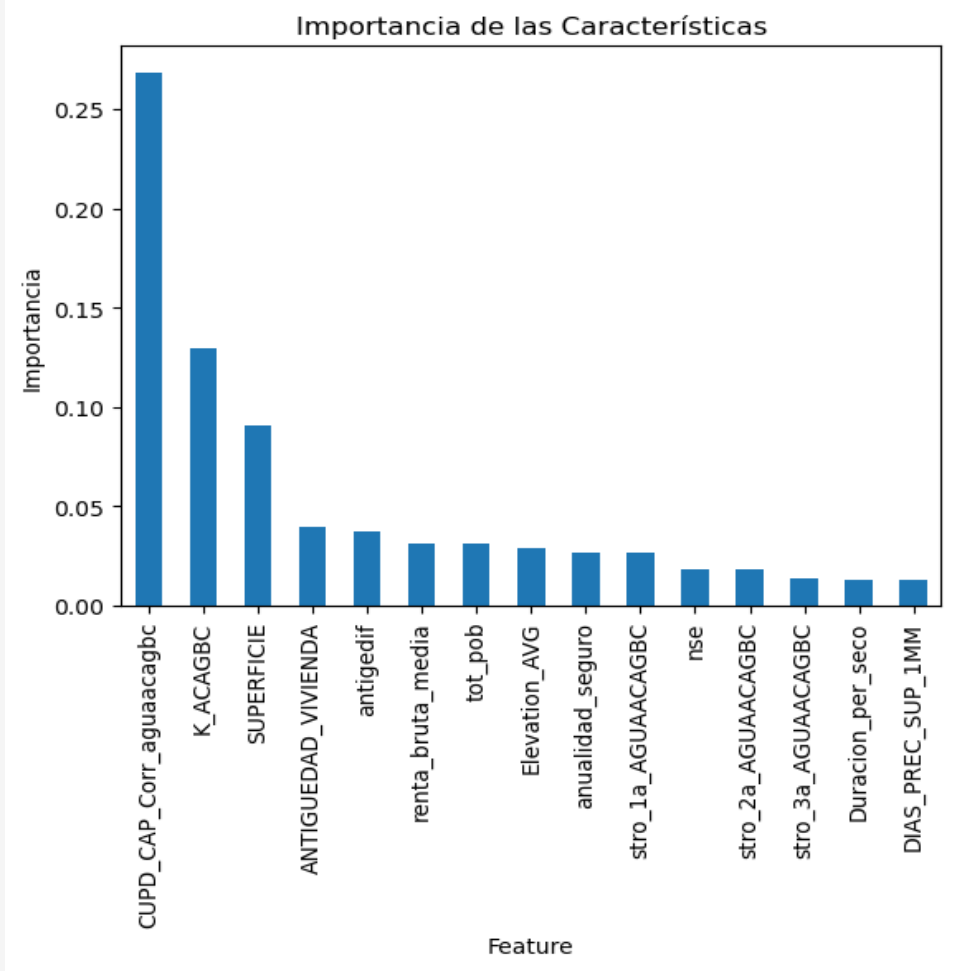
PRIME
BUSINESS SCHOOL
UNIVERSIDAD SERGIO ARBOLEDA

Metodología de trabajo

2



Selección de variables utilizando *Random Forest*



	Variable	VIF
0	const	98.27
1	CUPD_CAP_Corr_aguaacagbc	1.03
2	K_ACAGBC	1.47
3	SUPERFICIE	1.04
4	ANTIGUEDAD_VIVIENDA	1.59
5	antigedif	1.74
6	renta_bruta_media	3.14
7	tot_pob	1.42
8	Elevation_AVG	1.21
9	anualidad_seguro	1.16
10	stro_1a_AGUAACAGBC	2.03
11	nse	2.48
12	stro_2a_AGUAACAGBC	4.77
13	stro_3a_AGUAACAGBC	4.01
14	Duracion_per_seco	2.35
15	DIAS_PREC_SUP_1MM	2.31

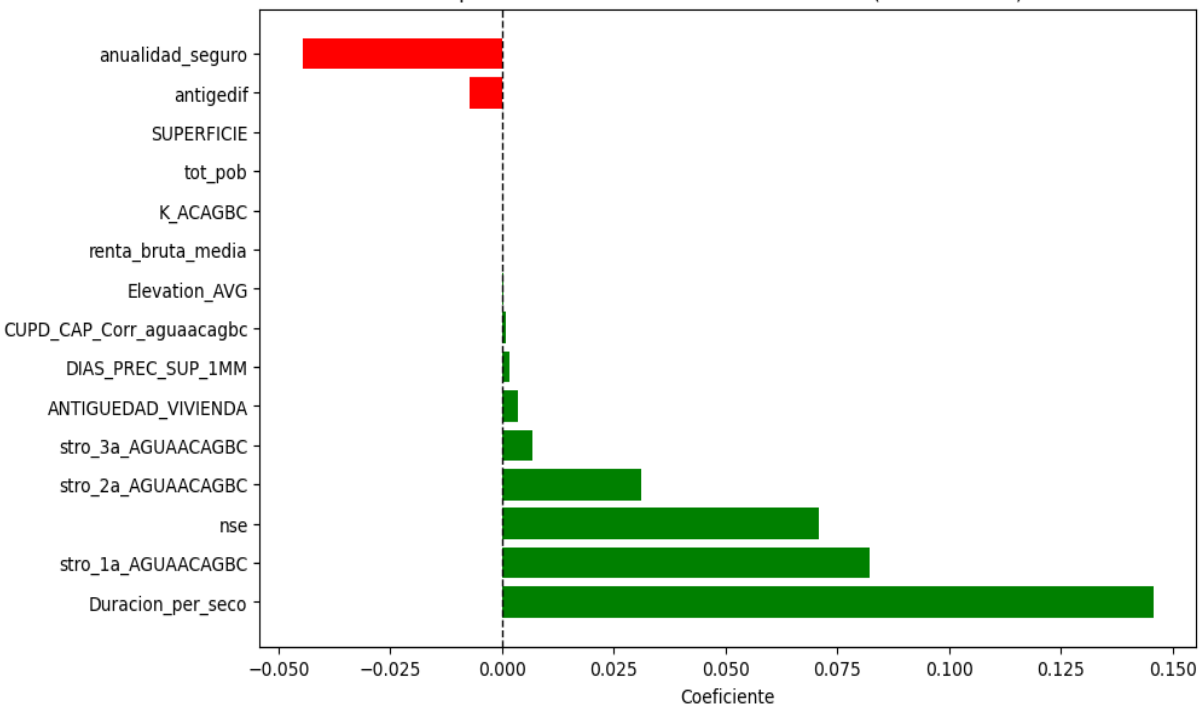
El VIF es una métrica utilizada para evaluar la multicolinealidad, mide cuánto aumenta la varianza de un coeficiente de regresión debido a la correlación con otras variables independientes del modelo.

Los valores de VIF obtenidos se encuentran en el rango de 1 a 5 para todas las variables, lo que indica que no existe un problema de multicolinealidad en este caso.

Variable	Impacto y sentido de negocio
CUPD_CAP_Corr_aguaacagbc	A mayor costo de reparación, más frecuencia de siniestros. Importante para ajustar primas y reservas.
K_ACAGBC	Mayor riesgo de agua incrementa la frecuencia de siniestros. Útil para clasificar propiedades y fijar precios de seguros.
SUPERFICIE	Propiedades más grandes tienen más puntos vulnerables. Considerar para cálculos de primas.
ANTIGUEDAD_VIVIENDA	Las viviendas antiguas tienen mayor riesgo de siniestros por desgaste. Usar para evaluar riesgo de siniestros.
antigedif	Edificios antiguos pueden tener menor riesgo debido a mantenimiento. Ajustar evaluación de riesgo.
renta_bruta_media	Zonas con más ingresos tienen mayor frecuencia de siniestros. Refleja capacidad de reporte y manejo de incidentes.
tot_pob	Más población implica mayor exposición a siniestros. Importante para evaluación de riesgos en áreas densas.
Elevation_AVG	A mayor altitud, mayor frecuencia de siniestros. Importante en zonas propensas a lluvias intensas.
anualidad_seguro	Mayor anualidad está asociada con menor frecuencia de siniestros, indicando una mejor gestión de riesgos.
stro_1a_AGUAACAGBC	Historial de daños pasados incrementa la probabilidad de siniestros futuros. Clave para la evaluación de riesgos.
nse	Mayor nivel socioeconómico está asociado con mayor frecuencia de siniestros. Considerar para segmentación de riesgos.
Duracion_per_seco	Períodos secos más largos pueden afectar la integridad de infraestructuras. Importante para evaluación de exposición.
DIAS_PREC_SUP_1MM	Incluso lluvias moderadas pueden aumentar la frecuencia de siniestros. Útil para modelar riesgos climáticos.

Modelo con top 15 variables

Impacto de las Variables en el Modelo GLM (sin constante)



HALLAZGOS

5

•**R² y Pseudo R²:** El Pseudo R² (CS) de 1.000 indica un ajuste perfecto del modelo, aunque en modelos de Poisson puede reflejar sobreajuste.

•**Significancia de las Variables:** Todas las variables son estadísticamente significativas ($p < 0.0001$), confirmando su impacto en la predicción de siniestros.

•**Valor del BIC:** El BIC es 17,681,493.01. Un valor que permite comparar la calidad del ajuste y la complejidad del modelo con otros modelos candidatos.

•**Impacto de Variables Clave:** Variables como costo de reparación, antigüedad de la vivienda y historial de siniestros son cruciales para comprender la frecuencia de los daños.

•**Implicaciones para la Gestión de Riesgos:** El análisis permite diseñar estrategias de prevención, ajustar pólizas y programas de mantenimiento para reducir siniestros.

•**Conclusión General:** El modelo es robusto y esencial para la gestión de riesgos y toma de decisiones en el sector asegurador.

Modelo con top 8 variables utilizando Lasso

HALLAZGOS

- Pseudo R²:** Valor de 0.9998, mostrando un ajuste casi perfecto a los datos, con un modelo simplificado que conserva gran capacidad explicativa.

- Significancia de las Variables:** Todas las variables seleccionadas son estadísticamente significativas ($P < 0.05$), validando su relevancia en la predicción de siniestros.

stro_1a_AGUAACAGBC: Historial de siniestros inicial aumenta el riesgo futuro.

Duracion_per_seco: Períodos prolongados incrementan siniestros.

anualidad_seguro: Prima más alta reduce frecuencia de siniestros.

ANTIGUEDAD_VIVIENDA: Viviendas más antiguas tienen más siniestros.

Elevation_AVG: Mayor altitud aumenta riesgos.

- BIC:** Valor de 18,224,691, competitivo pese a ser mayor al modelo completo, destacando la eficiencia del modelo reducido.

- Modelo Simplificado y Eficaz:** Selección de variables con Lasso genera un modelo más sencillo y aún robusto, facilitando su interpretación y aplicación.

- Relevancia para la Industria de Seguros:** Modelo útil para evaluar riesgos, ajustar pólizas, y enfocar esfuerzos en variables clave como historial de siniestros, antigüedad, y duración de períodos secos.

- Conclusión General:** Modelo robusto y eficiente, con aplicaciones directas para la gestión de riesgos y estrategias de personalización en el sector asegurador.

Impacto de las Variables en el Modelo GLM (sin constante)

