

TABLA DE CONTENIDO

1	Introducción	1
1.1	Antecedentes generales.....	1
1.2	Motivación	1
1.3	Objetivos	3
1.3.1	Objetivo General.....	3
1.3.2	Objetivos Específicos	3
1.4	Alcances y Limitaciones	4
1.5	Resultados Esperados	4
2	Marco Conceptual.....	4
2.1	Metodología KDD	4
2.2	Propensity Score	6
2.3	Adopción e intención de compra online.....	7
2.4	Problema de Clasificación	8
2.5	Técnicas de muestreo para set de datos desbalanceados	8
2.6	Algoritmos de Clasificación	9
2.6.1	Árboles de Decisión.....	10
2.6.2	Boosted Logit.....	10
2.6.3	Extreme Gradient Boosting.....	10
2.7	Indicadores de Rendimiento.....	11
2.7.1	Matriz de confusión.....	11
2.7.2	Curva ROC	12
2.8	Experimentos	13
2.9	Cálculo de población de grupos de control y de tratamiento.....	15
2.10	Indicadores de rendimiento campaña de primera compra web	15
2.11	Weight of Evidence e Information Value	16
3	Desarrollo Metodológico	17
3.1	Creación de un set de datos objetivo	17
3.2	Preprocesamiento de datos	18
3.2.1	Transformación de datos	20
3.3	Análisis Exploratorio.....	21
3.4	Propensión de primera compra online.....	28
3.4.1	Entrenamiento de modelos.....	29
3.4.2	Resultados de modelos	31
3.5	Variables más relevantes según modelo escogido	34
3.6	Caracterización propensión primera compra web	34

3.7	Diseño experimental.....	39
3.8	Hipótesis por testear	39
3.9	Variable experimental y tiempo de medición.....	40
3.10	Muestra total de clientes	40
3.11	Segmentación de clientes según propensión.....	41
3.12	Tamaño muestral	42
3.13	Comprobación de balance de grupos experimentales	43
3.14	Resultados experimento.....	44
3.15	Test de hipótesis	46
3.15.1	Hipótesis 1	46
3.15.2	Hipótesis 2.....	46
3.15.3	Hipótesis 3.....	47
3.16	Análisis de resultados experimento.....	47
4	Conclusiones	49
5	Trabajo futuro	51
6	Bibliografía.....	52
7	Anexos.....	54