

TABLA DE CONTENIDO

Resumen	I
Dedicatoria.....	II
Agradecimientos.....	III
Indice de tablas	VII
Indice de figuras	VIII
1 Introducción	1
1.1 Caracterización de la Industria.....	1
1.1.1 Rubro	1
1.1.2 Industria Postal en Chile.....	1
1.1.3 Productos y/o Servicios	1
1.1.4 Clientes	2
1.1.5 Dimensionamiento de actividad realizada.....	3
1.2 Mercado y/o Marco Institucional.....	4
1.2.1 Actores.....	4
1.2.2 Marco Normativo	5
1.2.3 Tendencias del Mercado.....	5
2 Descripción del proyecto.....	7
2.1 Información general del rubro postal.....	7
2.1.1 Proceso Postal.....	7
2.1.2 Productos y servicios	8
2.1.3 Cliente.....	8
2.2 Identificación del problema y oportunidad	9
2.2.1 Descripción del Problema.....	9
2.2.2 Declaración del Problema.....	9
2.2.3 Causas	9
2.2.4 Consecuencias	10
2.3 Objetivos del Proyecto.....	11
2.3.1 Objetivo general	11
2.3.2 Objetivos específicos.....	11
2.4 Alcances.....	11
2.4.1 Normalización de direcciones	11
2.4.2 Datos a utilizar.....	12
2.4.3 Aplicabilidad del modelo.....	12
3 Marco conceptual	13

3.1	Estandarización de direcciones postales	13
3.1.1	Clasificación de Texto	13
3.1.2	Importancia de la clasificación de texto	14
3.1.3	Coincidencia de dirección	14
3.1.4	Código Postal.....	15
3.2	Distancia de Levenshtein	15
3.2.1	Definición de la distancia de Levenshtein (o distancia mínima de edición)	16
3.2.2	Distancia con lógica difusa – FuzzyWuzzy.....	16
3.3	Clasificación de texto.....	17
3.3.1	Clasificador en base a reglas lógicas	17
3.3.2	Clasificador Probabilísticos.....	17
4	Metodología	19
4.1	Entendimiento del Negocio.....	19
4.1.1	Estandarización de direcciones.....	19
4.1.2	Mercado y Clientes.....	19
4.2	Entendimiento de los datos	20
4.3	Preparación de los Datos.....	21
4.4	Modelamiento	22
4.4.1	Clasificador de Texto.....	22
4.4.2	Coincidencia de direcciones	22
5	Modelo Clasificador reglas lógicas	23
6	Modelamiento – Asignación Código Postal.....	25
6.1	Coincidencia de textos	25
6.2	Modelamiento – Criterios de asignación	27
6.3	Modelamiento – Enfoques a utilizar	28
6.4	Modelamiento – Resultados.....	29
7	Resultados	30
7.1	Coincidencia directa.....	30
7.2	Clasificador en base a reglas lógicas	30
7.3	Asignación código postal.....	31
7.3.1	Rendimiento general de enfoques de coincidencia.....	32
7.3.2	Generación de categorías de confiabilidad de asignaciones.....	35
7.4	Resultados generales.....	39
7.5	Costo computacional.....	40
8	Conclusiones y Trabajo Futuro	42

9 Bibliografía.....	45
Anexo A.....	47
Anexo B.....	48