

TABLA DE CONTENIDO

1	INTRODUCCIÓN.....	1
1.1	Consideraciones Preliminares.....	1
1.2	Justificación.....	3
2	OBJETIVOS Y METODOLOGÍA.....	6
2.1	Objetivos.....	6
2.1.1	Objetivo General.....	6
2.1.2	Objetivos Específicos.....	6
2.1.3	Alcances.....	6
2.2	Metodología.....	7
3	INDUSTRIA DEL NITRATO.....	9
3.1	Contexto Industrial del Nitrato.....	9
3.2	Producción de Nitratos.....	9
3.3	Empresas.....	11
3.4	Nitrato de Potasio.....	12
4	DESCRIPCIÓN DEL PROCESO.....	15
4.1	Plantas.....	15
4.2	Contaminantes.....	17
4.3	Descripción de la Base de Datos.....	19
4.4	Variables Químicas.....	19
4.5	Variables de Control.....	20
5	MARCO TEÓRICO.....	22
5.1	Minería De Datos.....	22
5.1.1	Metodología KDD.....	22
5.1.2	Métodos Supervisados.....	24
5.1.3	Regresión Logística.....	25
5.1.4	Support Vector Machine.....	26
5.1.5	Random Forest.....	29
5.1.6	Valores Perdidos.....	30
5.2	Series De Tiempo.....	33
5.2.1	Definición.....	33
5.2.2	Metodología De Box-Jenkins.....	34

5.2.3	Procesos Estocásticos	35
5.2.4	Autocorrelación	36
5.2.5	Correlogramas	36
6	ESTUDIO DEL PROCESO Y VARIABLES IMPLICADAS	40
6.1	Determinación de la Unidad de Tiempo	40
6.2	Reducción de Variables por Completitud de los Datos	42
6.3	Tratamiento de los Valores Perdidos	43
6.4	Selección de Variables	45
6.5	Relación Temporal de Variables.....	45
7	MODELAMIENTO DEL PROCESO.....	49
7.1	Medidas de Rendimiento	49
7.2	Modelos de Clasificación.....	50
7.2.1	Regresión Logística	52
7.2.2	Support Vector Machine.....	52
7.2.3	Random Forest.....	53
7.2.4	Contraste De Clasificación y Selección de Algoritmo	53
7.3	Modelos de Regresión	54
7.3.1	Regresión Lineal.....	55
7.3.2	Support Vector Regression.....	55
7.3.3	Random Forest Regression.....	56
7.3.4	Contraste De Clasificación y Selección de Algoritmo	56
7.4	Elección del Modelo	57
8	DISEÑO DE PRUEBAS Y VALIDACIÓN DEL MODELO	59
8.1	Descripción del Modelamiento	59
8.1.1	Descripción Con Modelos De Clasificación	59
8.1.2	Descripción Con Modelos De Regresión	61
8.2	Importancia de variables	63
8.2.1	Ranking de Variables en Modelos de Clasificación.....	63
8.2.2	Ranking de Variables en Modelos de Regresión.....	68
8.3	Contribución Marginal de las Variables	73
9	ANÁLISIS DE SENSIBILIDAD DE RESULTADOS.....	76
9.1	Introducción	76
9.2	Primera Política de Rangos.....	79

9.3	Segunda Política de Rangos.....	81
9.4	Tercera Política de Rangos	84
9.5	Mejoras en la producción.....	86
10	CONCLUSIONES.....	90
	BIBLIOGRAFÍA	94
	ANEXOS	96
	ANEXO A.	96
	ANEXO B.	97
	ANEXO C.	98
	ANEXO D.	105
	ANEXO E.....	106
	ANEXO F.....	107

INDICE DE TABLAS

<i>Tabla 1.2.1: Contribución al Margen Bruto por segmento.....</i>	<i>5</i>
<i>Tabla 3.1.1: Industrias relevantes para el uso del KNO3</i>	<i>9</i>
<i>Tabla 4.2.1: Concentraciones de compuestos que describen calidad</i>	<i>17</i>
<i>Tabla 4.2.2: Cantidad de datos de los contaminantes por corriente.....</i>	<i>17</i>
<i>Tabla 4.2.3: Corriente L entre los años 2006-2012.....</i>	<i>18</i>
<i>Tabla 4.2.4: Corriente M entre los años 2006-2012</i>	<i>18</i>
<i>Tabla 4.4.1: Frecuencia de mediciones Planta de Cristalización</i>	<i>20</i>
<i>Tabla 4.5.1: Ejemplo de variables de control.....</i>	<i>20</i>
<i>Tabla 4.5.2: Frecuencia de mediciones de control en Planta de Cristalización</i>	<i>21</i>
<i>Tabla 5.1.1: Kernel más comunes.....</i>	<i>28</i>
<i>Tabla 5.1.2: Técnicas de tratamiento de valores perdidos más comunes</i>	<i>33</i>
<i>Tabla 6.1.1: Número de variables por unidad de frecuencia</i>	<i>41</i>
<i>Tabla 6.3.1: Tratamiento de datos perdidos</i>	<i>43</i>
<i>Tabla 6.3.2: Subconjunto de variables analizadas para el paso 2.....</i>	<i>44</i>
<i>Tabla 6.3.3: Distribución de pares correlacionados</i>	<i>44</i>
<i>Tabla 6.5.1: Referencias de rezagos</i>	<i>47</i>
<i>Tabla 7.1.1: Matriz de Confusión</i>	<i>49</i>