

Tabla de Contenido

1. Contextualización del trabajo de título	1
1.1. Introducción	1
1.2. Antecedentes generales	2
1.2.1. Antecedentes de los datos a utilizar	2
1.2.2. Antecedentes Machine Learning	3
1.3. Objetivos	4
1.3.1. Objetivo General	4
1.3.2. Objetivos Específicos	4
1.4. Alcances	4
2. Antecedentes	5
2.1. Antecedentes del dataset	5
2.2. Conceptos generales machine learning	8
2.2.1. Aprendizaje supervisado y aprendizaje no supervisado	8
2.2.2. Underfitting y overfitting	9
2.2.3. Métricas de desempeño	10
2.2.3.1. Clasificación binaria	10
2.2.3.2. Clasificación no binaria	11
2.2.4. Escalar los datos	11
2.2.5. Ventanas de tiempo y parámetros de tiempo	12
2.3. Dimensionalidad del dataset	14
2.3.1. La maldición de la dimensionalidad	14
2.3.2. Métodos de reducción de dimensionalidad	14
2.3.2.1. Linear Discriminant Analysis	15
2.3.3. Selección hacia adelante	15
2.4. Modelos a usar	16
2.4.1. Random Forests	16
2.4.2. K-Nearest Neighbors	18
2.4.3. Support Vector Machines	18
2.4.4. Redes neuronales	21
2.4.4.1. Hiperparámetros	23
3. Metodología	25
3.1. Recopilación de antecedentes y exploración del dataset	25
3.2. Preprocesamiento de los datos	27
3.2.1. Separación en conjuntos de entrenamiento y testeo	28
3.2.2. Obtención de parámetros de tiempo	28

3.2.3.	Escalamiento de los datos	29
3.2.4.	Cambio en las dimensiones de las etiquetas	29
3.3.	Selección de features y de hiperparámetros	29
3.3.1.	Selección de features	29
3.3.2.	Selección de parámetros	31
3.4.	Análisis de resultados	32
4.	Resultados	33
4.1.	Gráficos de los datos sin procesar	33
4.1.1.	Gráficos de la temperatura	33
4.1.2.	Gráficos de Flujo	37
4.1.3.	Gráficos de Presión	40
4.1.4.	Gráficos de otras variables	46
4.2.	Gráficos de los parámetros de tiempo	52
4.2.1.	Diferencias entre los parámetros de tiempo	52
4.2.2.	Cambios según las diferentes clasificaciones	56
4.2.3.	Cambios diferentes ventanas de tiempo	59
4.3.	Resultados a partir de los modelos LDA	62
4.4.	Resultados a partir de los modelos RF	66
4.5.	Resultados a partir de los modelos KNN	72
4.6.	Resultados a partir de los modelos SVM	77
4.6.1.	Accuracy en función del largo de ventana para diferentes kernels	77
4.6.2.	Heatmaps con la variación de los parámetros C y gamma	79
4.7.	Resultados a partir de los modelos NN	84
4.7.1.	Accuracy en función del largo de ventana para diferentes funciones de activación	84
4.7.2.	Heatmaps con la variación del learning rate y el parámetro alpha	89
5.	Discusión	95
5.1.	Efecto del largo de las ventanas de tiempo	95
5.2.	Diferencias entre parámetros de tiempo	98
5.3.	Efecto del cambio de parámetros de los modelos	100
5.3.1.	Random Forest	100
5.3.2.	K Nearest Neighbors	100
5.3.3.	Linear Discriminant Analysis	100
5.3.4.	Support Vector Machines	100
5.3.5.	Multi layer perceptron (MLP)	101
5.4.	Diferencias entre sensores	102
5.5.	Resultados Generales	103
6.	Conclusiones	104
	Bibliografía	105
	Anexo A. Resultados ampliados	108
A.1.	Gráficos de los datos sin procesar	108
A.1.1.	Gráficos de temperatura	108
A.1.1.1.	Gráficos del enfriador	108

A.1.1.2.	Gráficos de la válvula	111
A.1.1.3.	Gráficos de la bomba	113
A.1.1.4.	Gráficos del acumulador	116
A.1.1.5.	Gráficos de la estabilidad del sistema	118
A.1.2.	Gráficos de flujo	121
A.1.2.1.	Gráficos del enfriador	121
A.1.2.2.	Gráficos de la válvula	123
A.1.2.3.	Gráficos de la bomba	126
A.1.2.4.	Gráficos del acumulador	128
A.1.2.5.	Gráficos de la estabilidad del sistema	131
A.1.3.	Gráficos de presión	134
A.1.3.1.	Gráficos del enfriador	134
A.1.3.2.	Gráficos de la válvula	139
A.1.3.3.	Gráficos de la bomba	144
A.1.3.4.	Gráficos del acumulador	149
A.1.3.5.	Gráficos de la estabilidad del sistema	154
A.1.4.	Gráficos de otras variables	159
A.1.4.1.	Gráficos del enfriador	159
A.1.4.2.	Gráficos de la válvula	164
A.1.4.3.	Gráficos de la bomba	169
A.1.4.4.	Gráficos del acumulador	174
A.1.4.5.	Gráficos de la estabilidad del sistema	179
A.2.	Gráficos de los parámetros de tiempo	184
A.2.1.	Gráficos del enfriador	184
A.2.1.1.	Gráficos usando 1 dato por ciclo	184
A.2.1.2.	Gráficos usando 2 datos por ciclo	188
A.2.1.3.	Gráficos usando 6 datos por ciclo	192
A.2.2.	Gráficos de la válvula	195
A.2.2.1.	Gráficos usando 1 dato por ciclo	196
A.2.2.2.	Gráficos usando 2 datos por ciclo	200
A.2.2.3.	Gráficos usando 6 datos por ciclo	204
A.2.3.	Gráficos de la bomba	207
A.2.3.1.	Gráficos usando 1 dato por ciclo	208
A.2.3.2.	Gráficos usando 2 datos por ciclo	212
A.2.3.3.	Gráficos usando 6 datos por ciclo	216
A.2.4.	Gráficos del acumulador	219
A.2.4.1.	Gráficos usando 1 dato por ciclo	220
A.2.4.2.	Gráficos usando 2 datos por ciclo	224
A.2.4.3.	Gráficos usando 6 datos por ciclo	228
A.2.5.	Gráficos de la estabilidad del sistema	231
A.2.5.1.	Gráficos usando 1 dato por ciclo	232
A.2.5.2.	Gráficos usando 2 datos por ciclo	236
A.2.5.3.	Gráficos usando 6 datos por ciclo	240
A.3.	Resultados a partir de los modelos RF	244
A.4.	Resultados a partir de los modelos KNN	249
A.5.	Resultados a partir de los modelos SVM	254

A.5.1. Accuracy en función de la cantidad de ventanas de entrenamiento para distintos kernel	254
A.5.2. Accuracy al variar los parámetros “C” y “gamma”	259