

Tabla de Contenido

1. Introducción	1
1.1. Motivación	1
1.2. Objetivos	2
1.2.1. Objetivo general	2
1.2.2. Objetivos específicos	2
1.3. Alcances	2
1.4. Estructura de la tesis	3
2. Marco teórico	4
2.1. El sistema de transporte	4
2.1.1. Sistemas de Transporte Inteligente (ITS)	5
2.2. Teoría y modelamiento de tráfico	5
2.2.1. Teoría de flujo de tráfico	5
2.2.2. Diagramas fundamentales de tráfico	7
2.2.3. Congestión de tráfico	7
2.3. Predicción mediante aprendizaje automático	9
2.3.1. Predicción de variables de tráfico mediante aprendizaje automático	10
3. Caracterización del caso de estudio	13
3.1. Descripción de los datos	14
3.2. Selección de la muestra	15
4. Metodología	17
4.1. Identificación del estado de congestión de tráfico	18
4.2. Modelos de aprendizaje automático	20
4.2.1. Predicción del estado de congestión futuro	20
4.2.1.1. Naive-Bayes	21
4.2.1.2. Árbol de decisión (clasificación)	22
4.2.1.3. <i>Random forest</i> (clasificación)	24
4.2.1.4. Red neuronal MLP (clasificación)	25
4.2.2. Predicción de velocidad media futura	30
4.2.2.1. Árbol de decisión (regresión)	30
4.2.2.2. <i>Random forest</i> (regresión)	31
4.2.2.3. Red neuronal MLP (regresión)	31
4.3. Medidas de error	31
4.4. Entrenamiento y elección de hiperparámetros	33
4.4.1. Entrenamiento-validación-testeo	33

4.4.2. Sobreajuste	34
4.5. Tratamiento de clases desbalanceadas	35
5. Análisis y Resultados	37
5.1. Preprocesamiento de datos	37
5.1.1. Imputación de datos faltantes	37
5.1.2. Generación de atributos	37
5.2. Exploración de los datos	38
5.3. Identificación del estado de congestión	43
5.3.1. Ejemplo gráfico del etiquetado del estado de tráfico	46
5.4. Comparación y evaluación de modelos de predicción	47
5.4.1. Clasificación de congestión futura	47
5.4.1.1. Importancia de los atributos en la predicción de congestión .	49
5.4.1.2. Ejemplo gráfico para la predicción de congestión	50
5.4.2. Estimación de velocidad media futura	51
5.4.2.1. Importancia de los atributos en la predicción de velocidad .	54
5.4.2.2. Ejemplo gráfico para la predicción de velocidad	55
5.4.2.3. Regla de velocidad fija para identificar congestión	57
6. Conclusiones	58
6.1. Trabajo futuro	59
Bibliografía	60
Anexos	63
Anexo A. Preprocesamiento de los datos	63
A.1. Datos faltantes en la muestra	63
Anexo B. Especificaciones de los modelos de predicción de congestión	65
B.1. Configuración de los modelos y elección de hiperparámetros	65
B.2. Selección de hiperparámetros mediante validación cruzada	66
B.3. Resultados para distintos usos de información pasada mediante rezagos . . .	68
Anexo C. Importancias de los atributos para el modelo de <i>random forest</i>	69
Anexo D. Especificaciones de los modelos de predicción de velocidad	71
D.1. Configuración de los modelos y elección de hiperparámetros	71
D.2. Selección de hiperparámetros mediante validación cruzada	73
D.3. Resultados para distintos usos de información pasada mediante rezagos . . .	78