

Tabla de Contenido

Índice de Tablas	vii
Índice de Ilustraciones	viii
1. Introducción	1
1.1. Hipótesis	4
1.2. Objetivos	4
1.3. Estructura del documento	4
2. Antecedentes	6
2.1. Redes neuronales artificiales	6
2.1.1. Introducción a las Redes Neuronales Convolucionales	11
2.1.2. Estructura de las redes convolucionales	13
2.2. Teoría de la Información	20
2.2.1. Conceptos básicos de teoría de la información	20
2.2.2. Desigualdad de procesamiento de información y redes neuronales	25
2.2.3. Propiedades de representaciones de datos	27
2.2.4. Método de <i>information bottleneck</i> y plano de información	29
2.2.5. Estimación de medidas de información a partir de datos	31
2.3. Astronomía	36
2.3.1. Eventos transitorios en Astronomía	36
2.3.2. The High cadence Transient Survey (HiTS)	37
3. Flujo de información en transformaciones lineales y no lineales	43
3.1. Flujo de información en transformaciones lineales	43
3.2. Flujo de información en funciones de activación	47
3.3. Flujo de información en capas con sigmoides saturadas	50
4. Medición de propiedades de las representaciones intermedias	53
4.1. Estimación de suficiencia vía información mutua	53
4.1.1. Aproximación de Monte Carlo	55
4.1.2. Kernel Density Estimation	55
4.2. Robustez de las representaciones intermedias	57
4.2.1. Estimación de la robustez	59
4.3. Mediciones en redes neuronales convolucionales	60
5. Aprendizaje de representaciones utilizando medidas de información	63

5.1.	Maximización de suficiencia y robustez	63
5.2.	Descripción del algoritmo	65
5.2.1.	Aprendizaje de representaciones	65
5.3.	Entrenamiento de redes neuronales	68
6.	Deep-HiTS: Detección de eventos astronómicos transitorios con redes convolucionales	72
6.1.	Descripción del problema	72
6.1.1.	Base de datos	73
6.2.	Modelo propuesto	74
7.	Experimentos de monitoreo de propiedades	77
7.1.	Monitoreo de MLP en base de datos Espiral	77
7.1.1.	Descripción del experimento	77
7.1.2.	Resultados y discusión	78
7.2.	Monitoreo de MLP en base de datos MNIST	86
7.2.1.	Descripción del experimento	86
7.2.2.	Resultados y discusión	87
7.3.	Monitoreo de ConvNet en base de datos MNIST	90
7.3.1.	Descripción del experimento	90
7.3.2.	Resultados y discusión	92
7.4.	Discusión general	95
8.	Experimentos de entrenamiento optimizando medidas de información	97
8.1.	Relación entre robustez y separabilidad lineal	97
8.1.1.	Descripción del experimento	97
8.1.2.	Resultados y discusión	98
8.2.	Visualización de la maximización de suficiencia y robustez	100
8.2.1.	Descripción del experimento	100
8.2.2.	Resultados y discusión	101
8.3.	Maximización de robustez en MLP con base de datos Espiral	103
8.3.1.	Descripción del experimento	103
8.3.2.	Resultados y discusión	104
8.4.	Maximización de robustez en MLP con base de datos MNIST	110
8.4.1.	Descripción del experimento	110
8.4.2.	Resultados y discusión	111
8.5.	Maximización de robustez en red convolucional con base de datos CIFAR-10	117
8.5.1.	Descripción del experimento	117
8.5.2.	Resultados y discusión	118
8.6.	Discusión general	122
9.	Experimentos en Deep-HiTS	126
9.1.	Comparación entre Deep-HiTS y Random Forests	126
9.2.	Monitoreo de propiedades en Deep-HiTS	128
9.3.	Deep-HiTS: arquitectura simplificada	133
9.3.1.	Monitoreo de propiedades en Deep-HiTS simplificado con entrenamiento tradicional	133

9.3.2. Entrenamiento de Deep-HiTS simplificado con maximización de robustez 136

10. Conclusión	142
10.1. Trabajo futuro	146
11. Bibliografía	149
Anexos	159
A. Antecedentes adicionales de redes neuronales	160
A.1. Algoritmo backpropagation	160
A.2. Algoritmos de optimización de primer orden	163
A.3. Regularización	165
A.4. Función de activación	166
A.5. Cómo diseñar una red convolucional	168
A.6. Visualizar y comprender las redes convolucionales	169
B. Antecedentes adicionales de teoría de la información	171
B.1. Entropía en el caso continuo	171
B.1.1. Problema de convergencia de la entropía en caso continuo	171
B.1.2. Entropía diferencial de variable uniforme	172
B.2. Estadístico suficiente	172
C. Medición del rango en transformaciones lineales aleatorias	174