

CONTENIDO DETALLADO

PREÁMBULO	xvii
AGRADECIMIENTOS	xix
INTRODUCCIÓN	xxi
¿Para quién es este libro?	xxi
¿Qué conseguirá con este libro?	xxii
Cómo leer este libro	xxii
Recursos en línea	xxvii

PARTE I

REDES NEURONALES Y APRENDIZAJE PROFUNDO

1. INCRUSTACIONES, ESPACIO LATENTE Y REPRESENTACIONES	3
Incrustaciones	3
Espacio latente	5
Representación	6
Ejercicios	6
Referencias	7
2. APRENDIZAJE AUTOSUPERVISADO	9
El aprendizaje autosupervisado frente al aprendizaje por transferencia	9
Aprovechamiento de datos sin etiqueta	11
Autopredicción y aprendizaje autosupervisado contrastivo	12
Ejercicios	15
Referencias	15
3. APRENDIZAJE CON POCOS GOLPES	17
Conjuntos de datos y terminología	17
Ejercicios	20
4. LA HIPÓTESIS DEL BOLETO DE LOTERÍA	21
El procedimiento de formación con boleto de lotería	22
Implicaciones prácticas y limitaciones	23
Ejercicios	23
Referencias	24
5. REDUCCIÓN DEL SOBREAJUSTE CON DATOS	25
Métodos comunes	26
<i>Recopilación de más datos</i>	26
<i>Aumento de datos</i>	26
<i>Formación previa</i>	27

Otros métodos	28
Ejercicios	28
Referencias	29
6. REDUCCIÓN DEL SOBREAJUSTE CON MODIFICACIONES DEL MODELO	31
Métodos comunes	32
<i>Regularización</i>	32
<i>Modelos más pequeños</i>	33
<i>Precauciones con los modelos más pequeños</i>	34
<i>Métodos de ensamble</i>	35
Otros métodos	37
Elección de una técnica de regularización	38
Ejercicios	38
Referencias	38
7. PARADIGMAS DE LA FORMACIÓN MULTI-GPU	41
Los paradigmas de la formación	42
<i>Paralelismo del modelo</i>	42
<i>Paralelismo de los datos</i>	42
<i>Paralelismo tensorial</i>	43
<i>Paralelismo de las tuberías</i>	44
<i>Paralelismo de secuencias</i>	45
Recomendaciones	46
Ejercicios	46
Referencias	47
8. EL ÉXITO DE LOS TRANSFORMADORES	49
El mecanismo de atención	49
Formación previa con el aprendizaje autosupervisado	51
Número de parámetros grande	52
Paralelización sencilla	52
Ejercicios	53
Referencias	53
9. MODELOS DE IA GENERATIVA	55
Modelado generativo frente a modelado discriminativo	55
Tipos de modelos generativos profundos	56
<i>Modelos basados en energía</i>	56
<i>Autocodificadores variacionales</i>	57
<i>Redes adversarias generativas</i>	59
<i>Modelos basados en el flujo</i>	60
<i>Modelos autorregresivos</i>	61
<i>Modelos de difusión</i>	63
<i>Modelos de consistencia</i>	64

Recomendaciones.....	65
Ejercicios	65
Referencias	65
10. FUENTES DE ALEATORIEDAD.....	67
Inicialización del peso del modelo.....	67
Muestreo y distribución de conjuntos de datos	68
Algoritmos no determinísticos	69
Algoritmos de tiempo de ejecución	70
Hardware y controladores	71
Aleatoriedad e IA generativa	71
Ejercicios	73
Referencias	73

PARTE II VISIÓN INFORMÁTICA

11. CÁLCULO DEL NÚMERO DE PARÁMETROS.....	77
Cómo contar los parámetros	77
<i>Capas convolucionales</i>	78
<i>Capas totalmente conectadas</i>	80
Usos prácticos	81
Ejercicios	81
12. CAPAS CONVOLUCIONALES Y TOTALMENTE CONECTADAS	83
Cuando los tamaños del núcleo y de la entrada son iguales.....	84
Cuando el tamaño del núcleo es 1	85
Recomendaciones.....	86
Ejercicios	86
13. CONJUNTOS DE FORMACIÓN GRANDES PARA TRANSFORMADORES DE VISIÓN.....	87
Los sesgos inductivos en las CNN.....	88
Los ViT pueden superar a las CNN.....	91
Sesgos inductivos en los ViT	92
Recomendaciones.....	93
Ejercicios	94
Referencias	94

PARTE III PROCESAMIENTO DEL LENGUAJE NATURAL

14. LA HIPÓTESIS DISTRIBUCIONAL	97
Word2vec, BERT y GPT	98
¿Se cumple la hipótesis?	100

Ejercicios	100
Referencias	101
15. AUMENTO DE DATOS PARA TEXTO	103
Reemplazo de sinónimo	103
Eliminación de palabras	104
Intercambio de posición de palabras	104
Barajado de oraciones	105
Inyección de ruido	106
Traducción inversa	106
Datos sintéticos	107
Recomendaciones	107
Ejercicios	108
Referencias	108
16. AUTOATENCIÓN	109
Atención en las RNN	110
El mecanismo de autoatención	112
Ejercicios	113
Referencias	113
17. TRANSFORMADORES TIPO CODIFICADOR Y DECODIFICADOR.....	115
El transformador original	116
<i>Codificadores</i>	117
<i>Decodificadores</i>	119
Híbridos codificadores-decodificadores	120
Terminología.....	121
Modelos de transformadores contemporáneos.....	121
Ejercicios	122
Referencias	123
18. USO Y AFINACIÓN DE TRANSFORMADORES CON FORMACIÓN PREVIA.....	125
Uso de transformadores para tareas de clasificación	126
Aprendizaje en contexto, indexación y afinación de indicaciones	128
Ajuste fino con parámetros eficientes	132
El aprendizaje por refuerzo con retroalimentación humana	137
Adaptación de modelos de lenguaje preformados.....	138
Ejercicios	138
Referencias	139
19. EVALUACIÓN DE MODELOS DE LENGUAJE GRANDES GENERATIVOS	141
Métricas de evaluación para los LLM.....	141
<i>Perplejidad</i>	142
<i>Puntuación BLEU</i>	143

<i>Puntuación ROUGE</i>	145
<i>Puntuación BERT (BERTScore)</i>	146
Métricas de sustitución	148
Ejercicios	149
Referencias	149

PARTE IV PRODUCCIÓN E INSTALACIÓN

20. FORMACIÓN SIN ESTADO Y CON ESTADO	153
(Re)formación sin estado	153
Formación con estado	154
Ejercicios	155
21. IA CENTRADA EN LOS DATOS	157
IA centrada en los datos frente a IA centrada en los modelos	157
Recomendaciones	159
Ejercicios	160
Referencias	160
22. ACELERACIÓN DE LA INFERENCIA	161
Paralelización	161
Vectorización	162
Mosaico de bucles	163
Fusión de operadores	164
Cuantificación	165
Ejercicios	166
Referencias	166
23. CAMBIOS EN LA DISTRIBUCIÓN DE LOS DATOS	167
Cambio de covariable	167
Cambio de etiqueta	169
Deriva del concepto	169
Cambio de dominio	170
Tipos de cambios en la distribución de los datos	170
Ejercicios	172
Referencias	172

PARTE V DESEMPEÑO PREDICTIVO Y EVALUACIÓN DE LOS MODELOS

24. REGRESIÓN DE POISSON Y ORDINAL	175
Ejercicios	176

25. INTERVALOS DE CONFIANZA.....	177
Definición de intervalos de confianza.....	178
Los métodos.....	180
<i>Método 1: Intervalos de aproximación normal.....</i>	<i>180</i>
<i>Método 2: Conjuntos de formación con pruebas de arranque.....</i>	<i>181</i>
<i>Método 3: Predicciones con conjunto de pruebas de arranque.....</i>	<i>184</i>
<i>Método 4: Reformación de modelos con semillas aleatorias diferentes.....</i>	<i>184</i>
Recomendaciones.....	185
Ejercicios.....	186
Referencias.....	186
26. INTERVALOS DE CONFIANZA Y PREDICCIONES CONFORMES.....	187
Intervalos de confianza e intervalos de predicción.....	188
Intervalos de predicción y predicciones conformes.....	188
Regiones, intervalos y conjuntos de predicción.....	189
Cálculo de predicciones conformes.....	189
Un ejemplo de predicción conforme.....	191
Los beneficios de las predicciones conformes.....	192
Recomendaciones.....	193
Ejercicios.....	193
Referencias.....	194
27. MÉTRICAS ADECUADAS.....	195
Los criterios.....	195
El error cuadrático medio.....	196
La pérdida de entropía cruzada.....	198
Ejercicios.....	199
28. LA K EN LA VALIDACIÓN CRUZADA DE K ITERACIONES.....	201
Compensaciones en la selección de valores para k	202
Determinación de los valores apropiados para k	204
Ejercicios.....	205
Referencias.....	205
29. DISCORDANCIA ENTRE LOS CONJUNTOS DE FORMACIÓN Y LOS DE PRUEBA.....	207
Ejercicios.....	209
30. DATOS ETIQUETADOS LIMITADOS.....	211
Mejora del rendimiento del modelo con datos etiquetados limitados.....	211
<i>Etiquetado de más datos.....</i>	<i>212</i>
<i>Arranque de los datos.....</i>	<i>212</i>
<i>Aprendizaje por transferencia.....</i>	<i>212</i>
<i>Aprendizaje autosupervisado.....</i>	<i>212</i>
<i>Aprendizaje activo.....</i>	<i>213</i>

<i>Aprendizaje con pocos golpes</i>	214
<i>Metaprendizaje</i>	214
<i>Aprendizaje con supervisión débil</i>	215
<i>Aprendizaje semisupervisado</i>	216
<i>Autoformación</i>	217
<i>Aprendizaje multitarea</i>	218
<i>Aprendizaje multimodal</i>	219
<i>Sesgos inductivos</i>	220
Recomendaciones.....	220
Ejercicios	222
Referencias	222
EPÍLOGO	225
RESPUESTAS A LOS EJERCICIOS	227