

Tabla de contenido

Índice de figuras	XIX
Índice de tablas	XXII
CAPÍTULO 1	1
1.1. Inteligencia artificial generativa.....	1
1.1.1. Funcionamiento de la IA generativa.....	3
1.1.2. Focos de la IA generativa.....	6
1.1.3. Aplicaciones.....	7
1.2. Modelos de lenguaje generativos.....	8
1.3. Conclusiones	14
CAPÍTULO 2	15
2.1. Introducción.....	15
2.2. Modelos de lenguaje autorregresivos	19
2.3. Modelos de lenguaje estadísticos.....	21
2.4. Modelos de lenguaje neuronales.....	22
2.4.1. Modelos de lenguaje preentrenados	25
2.5. Grandes modelos de lenguaje	26
2.6. Modelos de <i>embeddings</i> de palabras	27
2.7. Redes neuronales recurrentes.....	35
2.7.1. Redes neuronales recurrentes simples	35
2.7.2. Redes de memoria a corto-largo plazo.....	40
2.8. Autoencoders.....	44
2.8.1. Cuello de botella de la información.....	46
2.8.2. Variables latentes	47
2.8.3. Arquitectura de un Autoencoder	49
2.8.4. Tipos de Autoencoders.....	50
2.9. Redes adversarias generativas.....	56

- 2.10. Modelos de atención 59
 - 2.10.1. Problema del *encoder-decoder* 61
 - 2.10.2. Atención en modelos de secuencia 63
- 2.11. Transformers..... 80
 - 2.11.1. Capa del *encoder* 84
 - 2.11.2. Codificación posicional 85
 - 2.11.3. Conexiones residuales 89
 - 2.11.4. Capa del *decoder* 90
 - 2.11.5. Capa lineal y SoftMax 93
 - 2.11.6. Entrenamiento 94
 - 2.11.7. Inferencia..... 96
 - 2.11.8. Función de pérdida..... 98
- 2.12. Conclusiones 100
- CAPÍTULO 3** 101
 - 3.1. Introducción..... 101
 - 3.1.1. Habilidades emergentes..... 102
 - 3.1.2. Técnicas de mejoramiento de capacidades..... 104
 - 3.1.3. Corpus comunes..... 105
 - 3.1.4. Tipos de entrenamiento 106
 - 3.1.5. Tipos de aprendizaje..... 107
 - 3.1.6. Tipos de tokenización 109
 - 3.2. BERT 110
 - 3.2.1. Funcionamiento..... 112
 - 3.2.2. Arquitectura 115
 - 3.2.3. Entrada del modelo 115
 - 3.2.4. Salida del modelo 116
 - 3.2.5. Modelos preentrenados basados en BERT..... 118
 - 3.3. GPT..... 119
 - 3.3.1. El modelo GPT y GPT-2 121
 - 3.3.2. El modelo GPT-3 131
 - 3.3.3. El modelo GPT-4 134
 - 3.3.4. Reinforcement Learning from Human Feedback 135

3.4.	PaLM	140
3.4.1.	Vocabulario	143
3.4.2.	Entrenamiento	144
3.4.3.	PaLM-2	145
3.5.	LLaMA	148
3.5.1.	Datos de preentrenamiento	149
3.5.2.	Arquitectura	150
3.6.	LaMDA.....	151
3.6.1.	Objetivos y métricas	153
3.6.2.	Preentrenamiento de LaMDA.....	154
3.7.	MEGATRON.....	156
3.7.1.	Datos de entrenamiento	159
3.8.	Otros LLM.....	160
3.9.	Conclusiones	162
CAPÍTULO 4	165
4.1.	Introducción.....	165
4.2.	Tareas de evaluación	166
4.2.1.	Tareas básicas de evaluación	167
4.2.2.	Tareas avanzadas de evaluación	171
4.2.3.	Tareas de cumplimiento de regulaciones.....	172
4.3.	Métricas y puntos de referencia	176
4.4.	<i>Datasets</i> de Benchmark	178
4.4.1.	SQuAD (Stanford Question Answering Dataset)	178
4.4.2.	GLUE (General Language Understanding Evaluation)	179
4.4.3.	SNLI (Stanford Natural Language Inference)	180
4.4.4.	ARC (Abstraction and Reasoning Corpus).....	180
4.5.	Evaluación de LLM	181
4.6.	Conclusiones	186

CAPÍTULO 5	189
5.1. Introducción.....	189
5.2. Clasificación de sentimientos.....	190
5.3. Búsqueda semántica en textos	197
5.4. Razonamiento con agentes de lenguaje	198
5.5. Inferencia causal	201
5.6. Acceso a bases de datos en lenguaje natural	203
5.7. Cargando y preguntando por datos propios	206
5.8. Realizando ajuste fino de un modelo con datos propios	209
5.9. Diseño y optimización de <i>prompts</i>	214
5.10. Sistema conversacional ChatGPT	221
5.11. Sistema conversacional BARD.....	229
5.12. Conclusiones	231
CAPÍTULO 6	233
6.1. Introducción.....	233
6.2. Habilidades emergentes	234
6.3. LLM en producción	236
6.4. Alineación entre humanos y LLM.....	238
6.5. Ética	240
6.6. Aspectos regulatorios	242
6.7. Complejidad	243
6.8. Riesgos	244
6.9. Limitaciones	245
6.10. Conclusiones	247
Índice onomástico	249
Bibliografía	253