

Vorwort	11
1 Grundlagen von GPT-4 und ChatGPT	15
Einführung in Large Language Models	15
Die Grundlagen von Sprachmodellen und NLP	16
Die Transformer-Architektur und ihre Rolle in LLMs	19
Die Tokenisierungs- und Vorhersageschritte in GPT-Modellen entmystifizieren	23
Einbinden von Vision in ein LLM	25
Eine kurze Geschichte des GPT: von GPT-1 bis GPT-4	26
GPT-1	26
GPT-2	28
GPT-3	28
Von GPT-3 zu InstructGPT	29
GPT-3.5, ChatGPT und Codex	32
GPT-4	33
Die Entwicklung der KI hin zur Multimodalität	38
Anwendungsfälle für LLMs und Beispielprodukte	39
Be My Eyes	39
Morgan Stanley	40
Khan Academy	41
Duolingo	41
Yabble	42
Waymark	43
Inworld AI	43

Vorsicht, KI-Halluzinationen: Einschränkungen und Überlegungen 44

Das Potenzial von GPT durch fortgeschrittene Features ausreizen . . . 47

Zusammenfassung 50

2 Die APIs von OpenAI 53

Grundlegende Konzepte 54

In der OpenAI-API verfügbare Modelle 55

 GPT Base 56

 InstructGPT (Legacy) 56

 GPT-3.5 56

 GPT-4 57

GPT-Modelle mit dem OpenAI Playground ausprobieren 58

Einstieg in die OpenAI-Python-Bibliothek 63

 OpenAI-Zugriff und API-Schlüssel 63

 »Hallo Welt«-Beispiel 66

Chat-Completion-Modelle einsetzen 68

 Eingabeoptionen für den Chat-Completion-Endpoint 69

 Mit der Temperatur und top_p experimentieren 74

 Ausgabeformat für den Chat-Completion-Endpoint 77

 Vision 78

 Eine JSON-Ausgabe erzwingen 82

Andere Modelle zur Textvervollständigung verwenden 86

 Eingabeoptionen für den Textvervollständigungs-Endpoint 87

 Ausgabeformat für den Textvervollständigungs-Endpoint 89

Überlegungen zu Kosten und Datenschutz 89

 Preise und Token-Beschränkungen 89

 Sicherheit und Privatsphäre: Achtung! 90

Andere OpenAI-APIs und Funktionen 91

 Embeddings 91

 Moderation-Modell 94

 Text-to-Speech 98

 Speech-to-Text 100

 Images-API 104

Zusammenfassung (und Spickzettel) 115

3	Apps mit GPT-4 und ChatGPT bauen	117
	Überblick über die Anwendungsentwicklung	117
	API-Schlüsselmanagement	118
	Sicherheit und Datenschutz	120
	Prinzipien zum Design der Softwarearchitektur	121
	LLM-Fähigkeiten in Ihre Projekte integrieren	122
	Gesprächsfähigkeiten	122
	Sprachverarbeitungsfähigkeiten	123
	Mensch-Maschine-Interaktion	125
	Fähigkeiten kombinieren	126
	Beispielprojekte	127
	Projekt 1: Einen News-Generator bauen – Sprachverarbeitung	127
	Projekt 2: YouTube-Videos zusammenfassen – Sprachverarbeitung	130
	Projekt 3: Einen Experten für Zelda BOTW erschaffen – Sprachverarbeitung und Unterhaltung	135
	Projekt 4: Persönlicher Assistent – Mensch-Maschine- Schnittstelle	142
	Projekt 5: Dokumente organisieren – Sprachverarbeitung	150
	Projekt 6: Sentimentanalyse – Sprachverarbeitung	152
	Kostenmanagement	160
	Angriffspunkte in LLM-gestützten Apps	162
	Eingaben und Ausgaben analysieren	164
	Die Unvermeidbarkeit der Prompt Injection	165
	Mit einer externen API arbeiten	165
	Umgang mit Fehlern und unerwarteten Verzögerungen	165
	Rate Limits	167
	Responsivität und User Experience verbessern	168
	Zusammenfassung	172
4	Fortgeschrittene Integrationsstrategien mit OpenAI	173
	Prompt Engineering	173
	Effektive Prompts mit Rollen, Kontexten und Aufgaben designen	175
	Schritt für Schritt denken	182