

मशीन लर्निंग, डीप लर्निंग, और जीपीटी

1. मशीन लर्निंग (ML) कंप्यूटर साइंस का एक क्षेत्र है जो प्रणालियों को डेटा से सीखने और बिना स्पष्ट प्रोग्रामिंग के अपने प्रदर्शन को सुधारने की अनुमति देता है।
2. डीप लर्निंग (DL) एक ML का उपक्षेत्र है जो बहु-स्तरीय न्यूरल नेटवर्क का उपयोग करके डेटा में जटिल पैटर्न को मॉडल करने के लिए करता है।
3. न्यूरल नेटवर्क मानव दिमाग से प्रेरित गणनात्मक मॉडल हैं, जो इंटरकनेक्टेड नोड (न्यूरॉन) से बने होते हैं जो सूचना को स्तरों में प्रोसेस करते हैं।
4. ट्रेनिंग डेटा लेबल्ड या अनलेबल्ड डेटासेट है जो मशीन लर्निंग मॉडल को एक टास्क कैसे करने के लिए सिखाता है।
5. सुपरवाइज्ड लर्निंग लेबल्ड डेटा पर एक मॉडल को ट्रेन करने में शामिल होता है, जहां प्रत्येक उदाहरण में एक इनपुट और एक संबद्ध सही आउटपुट होता है।
6. अनसुपरवाइज्ड लर्निंग अनलेबल्ड डेटा का उपयोग करता है, जिससे मॉडल को बिना स्पष्ट निर्देश के छुपे पैटर्न या समूहों को खोजने की अनुमति होती है।
7. रिनफोर्समेंट लर्निंग (RL) एजेंट्स को रिवार्ड करने और अनिच्छा करने के लिए अनिच्छा करने के लिए निर्णय लेने के लिए ट्रेन करता है।
8. जनरेटिव मॉडल अपने ट्रेनिंग उदाहरणों के समान नया डेटा पैदा करने में सीखते हैं (उदाहरण के लिए, टेक्स्ट, छवियां)।
9. डिस्क्रिमिनेटिव मॉडल इनपुट को श्रेणियों में वर्गीकृत करने या विशेष परिणामों को भविष्यवाणी करने पर ध्यान केंद्रित करते हैं।
10. ट्रांसफर लर्निंग एक मॉडल को एक टास्क पर ट्रेन करने के बाद एक संबंधित टास्क पर पुनः उपयोग या फाइन-ट्यून करने की अनुमति देता है।
11. GPT (जनरेटिव प्री-ट्रेन्ड ट्रांसफॉर्मर) एक परिवार है जो OpenAI द्वारा विकसित किया गया है जो मानव-प्रेरणा टेक्स्ट पैदा कर सकता है।
12. चैट GPT एक इंटरैक्टिव वेरिएंट है GPT, जो संवाद और निर्देश-फॉलोइंग टास्क्स के लिए फाइन-ट्यून किया गया है।
13. ट्रांसफॉर्मर आर्किटेक्चर “एटेंशन इज़ ऑल यू निड” पेपर में पेश किया गया था, जो नेचरल लैंग्वेज प्रोसेसिंग को क्रांतिकारी बना दिया, जो एटेंशन मैकेनिज्म पर निर्भर करता है।
14. सेल्फ-एटेंशन मैकेनिज्म मॉडल को इनपुट सीक्वेंस के विभिन्न हिस्सों को वजन देने की अनुमति देते हैं जब एक आउटपुट प्रतिनिधित्व बनाया जाता है।
15. ट्रांसफॉर्मर्स में पोजिशनल एन्कोडिंग मॉडल को एक सीक्वेंस में टोकन का क्रम पहचानने में मदद करता है।
16. प्री-ट्रेनिंग प्रारंभिक चरण है जहां एक मॉडल बड़े पैमाने पर डेटा से सामान्य विशेषताएं सीखता है, फिर विशेष टास्क्स पर फाइन-ट्यून किया जाता है।
17. फाइन-ट्यूनिंग एक प्री-ट्रेन्ड मॉडल को लेना और इसे एक छोटे, टास्क-स्पेसिफिक डेटासेट का उपयोग करके एक नैरो टास्क में अनुकूलित करना है।
18. लैंग्वेज मॉडलिंग एक टास्क है जिसमें एक सीक्वेंस में अगला टोकन (शब्द या सबवर्ड) भविष्यवाणी करना है, जो GPT-प्रकार के मॉडलों के लिए आधारभूत है।
19. जीरो-शॉट लर्निंग एक मॉडल को बिना स्पष्ट ट्रेनिंग उदाहरणों के टास्क्स को संभालने की अनुमति देता है, सीखी हुई सामान्य जानकारी पर निर्भर करता है।
20. फ्यू-शॉट लर्निंग सीमित संख्या में टास्क-स्पेसिफिक उदाहरणों का उपयोग करके मॉडल भविष्यवाणियों या व्यवहार को मार्गदर्शन करने के लिए करता है।
21. GPT (रिनफोर्समेंट लर्निंग फ्रॉम ह्यूमन फीडबैक) मॉडल आउटपुट्स को मानव पसंद और मान्यताओं के साथ संरेखित करने के लिए उपयोग किया जाता है।

22. मानव फीडबैक रैंकिंग या लेबल्स शामिल हो सकते हैं जो मॉडल के जनरेशन को अधिक इच्छित प्रतिक्रियाओं की ओर मार्गदर्शित करते हैं।
23. प्रॉम्प्ट इंजीनियरिंग बड़े लैंग्वेज मॉडलों को प्रभावी रूप से मार्गदर्शन करने के लिए इनपुट क्वेरी या निर्देशों को तैयार करने का कला है।
24. कॉन्टेक्ट विंडो मॉडल एक बार में प्रोसेस कर सकता है; $\square\square\square$ मॉडल एक सीमित कॉन्टेक्ट लंबाई रखते हैं।
25. इनफेरेन्स एक चरण है जहां एक ट्रेन्ड मॉडल नए इनपुट्स के लिए भविष्यवाणियां बनाता है या आउटपुट पैदा करता है।
26. पैरामीटर काउंट मॉडल क्षमता का एक प्रमुख कारक है; बड़े मॉडल अधिक जटिल पैटर्न को कैच कर सकते हैं, लेकिन अधिक गणना की आवश्यकता होती है।
27. मॉडल कमप्रेशन तकनीकें (उदाहरण के लिए, प्रूनिंग, क्वांटाइजेशन) एक मॉडल की आकार को कम कर सकती हैं और कम से कम दक्षता हानि के साथ इनफेरेन्स को तेज कर सकती हैं।
28. ट्रांसफॉर्मर्स में एटेंशन हेड्स इनपुट के विभिन्न पहलुओं को समानांतर में प्रोसेस करते हैं, प्रतिनिधित्व शक्ति को बेहतर बनाते हैं।
29. मास्कड लैंग्वेज मॉडलिंग (उदाहरण के लिए, $\square\square\square\square$ में) एक वाक्य में गायब टोकन भविष्यवाणी करने में शामिल होता है, जिससे मॉडल को सांस्कृतिक सीखने में मदद मिलती है।
30. कॉसल लैंग्वेज मॉडलिंग (उदाहरण के लिए, $\square\square\square$ में) सभी पूर्व टोकन के आधार पर अगला टोकन भविष्यवाणी करने में शामिल होता है।
31. एन्कोडर-डिकोडर आर्किटेक्चर (उदाहरण के लिए, $\square 5$) एक नेटवर्क को इनपुट को एन्कोड करने और दूसरे को डिकोड करने के लिए उपयोग करता है।
32. कॉन्वोल्यूशनल न्यूरल नेटवर्क ($\square\square\square\square$) कॉन्वोल्यूशनल लेयर के माध्यम से ग्रिड-प्रकार के डेटा (उदाहरण के लिए, छवियां) को प्रोसेस करने में उत्कृष्ट होते हैं।
33. रिकरेंट न्यूरल नेटवर्क ($\square\square\square\square$) सीक्वेंशियल डेटा को प्रोसेस करते हैं, समय चरणों के साथ छुपे हुए स्टेट्स को पास करते हैं, हालांकि वे लंबे समय के निर्भरताओं से संघर्ष कर सकते हैं।
34. लॉन्ग शॉर्ट-टर्म मेमोरी ($\square\square\square\square$) और $\square\square\square$ $\square\square\square$ वेरिएंट हैं जो लंबे दूरी के निर्भरताओं को बेहतर ढंग से कैच करने के लिए डिजाइन किए गए हैं।
35. बैच नॉर्मलाइजेशन ट्रेनिंग को स्थिर बनाने में मदद करता है, मध्यवर्ती लेयर आउटपुट्स को नॉर्मलाइज करने के द्वारा।
36. ड्रॉपआउट एक नियमितकरण तकनीक है जो ट्रेनिंग के दौरान न्यूरॉन्स को “ड्रॉप” करने के लिए अकस्मात् करता है, ओवरफिटिंग से बचने के लिए।
37. ऑप्टिमाइजर एल्गोरिदम्स जैसे स्टोकास्टिक ग्रेडिएंट डिसेंट ($\square\square\square$), $\square\square\square\square$, और $\square\square\square\square\square\square\square\square$ ग्रेडिएंट्स के आधार पर मॉडल पैरामीटर्स को अपडेट करते हैं।
38. लर्निंग रेट एक हाइपरपैरामीटर है जो ट्रेनिंग के दौरान वजन अपडेट की तीव्रता को निर्धारित करता है।
39. हाइपरपैरामीटर (उदाहरण के लिए, बैच साइज, लेयर की संख्या) ट्रेनिंग से पहले चुने गए कनफिगरेशन सेटिंग्स हैं जो सीखने को कैसे विकसित होता है, को नियंत्रित करने के लिए।
40. मॉडल ओवरफिटिंग तब होता है जब एक मॉडल ट्रेनिंग डेटा को बहुत अच्छी तरह से सीखता है, नए डेटा पर सामान्यीकरण करने में विफल रहता है।
41. नियमितकरण तकनीकें (उदाहरण के लिए, $\square 2$ वेट डेके, ड्रॉपआउट) ओवरफिटिंग को कम करने और सामान्यीकरण को बेहतर बनाने में मदद करती हैं।
42. वैलिडेशन सेट हाइपरपैरामीटर को ट्यून करने के लिए उपयोग किया जाता है, जबकि टेस्ट सेट मॉडल के अंतिम प्रदर्शन को मूल्यांकन करता है।
43. क्रॉस-वैलिडेशन डेटा को कई उपसेट में विभाजित करता है, प्रणालीतः ट्रेनिंग और वैलिडेशन करके एक अधिक स्थिर प्रदर्शन अनुमान प्राप्त करता है।
44. ग्रेडिएंट एक्सप्लोडिंग और वैनिशिंग समस्याएं गहन नेटवर्कों में होती हैं, जो ट्रेनिंग को अस्थिर या असमर्थ बनाती हैं।

45. रिजिडुअल कनेक्शंस (स्किप कनेक्शंस) नेटवर्क जैसे $\square\square\square\square\square$ में वैनिशिंग ग्रेडिएंट्स को कम करने में मदद करते हैं, डेटा पथों को शॉर्टकट करके।
46. स्केलिंग लॉज सुझाव देते हैं कि मॉडल आकार और डेटा को बढ़ाने से आम तौर पर बेहतर प्रदर्शन होता है।
47. कंप्यूट एफिशेंसी महत्वपूर्ण है; बड़े मॉडलों को ट्रेन करने के लिए ऑप्टिमाइज्ड हार्डवेयर ($\square\square\square\square$, $\square\square\square\square$) और एल्गोरिदम्स की आवश्यकता होती है।
48. नैतिक विचारों में वेश्यावृत्ति, न्याय और संभावित नुकसान शामिल हैं— $\square\square$ मॉडलों को सावधानी से टेस्ट और मॉनिटर करना चाहिए।
49. डेटा ऑगमेंटेशन ट्रेनिंग डेटासेट को कृत्रिम रूप से बढ़ाता है, विशेष रूप से छवि और स्पीच टास्क में मॉडल की दृढ़ता को बेहतर बनाने के लिए।
50. डेटा प्रीप्रोसेसिंग (उदाहरण के लिए, टोकनाइजेशन, नॉर्मलाइजेशन) प्रभावी मॉडल ट्रेनिंग के लिए आवश्यक है।
51. टोकनाइजेशन टेक्स्ट को टोकन (शब्द या सबवर्ड) में विभाजित करता है, जो लैंग्वेज मॉडलों द्वारा प्रोसेस किए जाने वाले मूलभूत इकाइयां हैं।
52. वेक्टर एम्बेडिंग टोकन या अवधारणाओं को संख्यात्मक वेक्टर के रूप में प्रतिनिधित्व करते हैं, सांस्कृतिक संबंधों को संरक्षित करते हैं।
53. पोजिशनल एम्बेडिंग्स प्रत्येक टोकन के पोजिशन के बारे में जानकारी जोड़ते हैं, जिससे ट्रांसफॉर्मर सीक्वेंस क्रम को समझने में मदद मिलती है।
54. एटेंशन वेट्स एक मॉडल को इनपुट के विभिन्न हिस्सों पर ध्यान केंद्रित करने की प्रक्रिया को दिखाते हैं।
55. बीम सर्च एक डिकोडिंग स्ट्रेटेजी है जो प्रत्येक चरण पर कई उम्मीदवारों को रखता है, सबसे अच्छा कुल सीक्वेंस खोजने के लिए।
56. ग्रीडी सर्च प्रत्येक चरण पर सबसे प्रबल टोकन को चुनता है, लेकिन अंतिम आउटपुट को कम से कम कर सकता है।
57. नमूने में तापमान लैंग्वेज जनरेशन की रचनात्मकता को समायोजित करता है: अधिक तापमान = अधिक अकस्मात्।
58. टॉप-क और टॉप-प (न्यूक्लियस) नमूने तकनीकें उम्मीदवारों को \square सबसे प्रबल या एक संयुक्त प्रायिकता \square तक सीमित करते हैं, विविधता और सहजता को संतुलित करते हैं।
59. पर्लक्सिटी एक प्रायिकता मॉडल द्वारा एक नमूने को भविष्यवाणी करने की गुणवत्ता को मापता है; कम पर्लक्सिटी बेहतर भविष्यवाणी प्रदर्शन को इंगित करता है।
60. प्रिसिजन और रिकॉल वर्गीकरण टास्क के लिए मेट्रिक्स हैं, जो सही और पूर्णता पर ध्यान केंद्रित करते हैं, क्रमशः।
61. $\square 1$ स्कोर प्रिसिजन और रिकॉल का हार्मोनिक माध्य है, दोनों मेट्रिक्स को एक एकल मान में संतुलित करता है।
62. दक्षता सही भविष्यवाणियों का अनुपात है, लेकिन असंतुलित डेटासेट में भ्रामक हो सकता है।
63. $\square\square\square$ कर्व के नीचे का क्षेत्र ($\square\square\square$) एक वर्गीकरण के विभिन्न थ्रेशोल्ड पर प्रदर्शन को मापता है।
64. गलतफहमी मेट्रिक्स सत्य सकारात्मक, गलत सकारात्मक, गलत नकारात्मक और सत्य नकारात्मक की गिनती दिखाता है।
65. अनिश्चितता अनुमान तकनीकें (उदाहरण के लिए, मॉन्टे कार्लो ड्रॉपआउट) एक मॉडल को अपने भविष्यवाणियों में कितना विश्वास रखता है, यह बताती हैं।
66. सक्रिय लर्निंग नए डेटा उदाहरणों को पूछता है जो मॉडल सबसे कम विश्वास रखता है, डेटा दक्षता को बेहतर बनाता है।
67. ऑनलाइन लर्निंग नए डेटा के आने पर मॉडल को अग्रिम रूप से अपडेट करता है, फिर से ट्रेनिंग से बचता है।
68. विकासवादी एल्गोरिथम और जेनेटिक एल्गोरिथम्स मॉडलों या हाइपरपैरामीटर को जैविक प्रेरित म्यूटेशन और चयन का उपयोग करके ऑप्टिमाइज करते हैं।
69. बेसियन विधियां पूर्व जानकारी को शामिल करती हैं और आगंतुक डेटा के साथ विश्वासों को अपडेट करती हैं, अनिश्चितता मापन के लिए उपयोगी हैं।

70. एन्सेम्बल विधियां (उदाहरण के लिए, रैंडम फॉरेस्ट, ग्रेडिएंट बूस्टिंग) कई मॉडलों को संयोजित करके प्रदर्शन और स्थिरता को बेहतर बनाती हैं।
71. बैगिंग (बूटस्ट्रैप एग्रीगेशन) अलग-अलग डेटा उपसेट पर कई मॉडलों को ट्रेन करता है, फिर उनके भविष्यवाणियों को औसत करता है।
72. बूस्टिंग नए मॉडलों को पूर्व ट्रेन किए गए मॉडलों द्वारा किए गए त्रुटियों को सुधारने के लिए अग्रिम रूप से ट्रेन करता है।
73. ग्रेडिएंट बूस्टेड डिसीजन ट्रीज (GBDT) संरचित डेटा के लिए शक्तिशाली होते हैं, अक्सर सरल न्यूरल नेटवर्कों को पार कर जाते हैं।
74. ऑटोरेग्रेसिव मॉडल एक सीक्वेंस में पूर्व आउटपुट्स के आधार पर अगला मान (या टोकन) भविष्यवाणी करते हैं।
75. ऑटोएन्कोडर एक न्यूरल नेटवर्क है जो डेटा को एक लैटेंट प्रतिनिधित्व में एन्कोड करता है और फिर इसे वापस डिकोड करता है, संपीड़ित डेटा प्रतिनिधित्व सीखता है।
76. वैरिएशनल ऑटोएन्कोडर (VAE) नए डेटा पैदा करने के लिए एक प्रायिकता ट्विस्ट जोड़ता है जो ट्रेनिंग सेट के समान होता है।
77. जनरेटिव एडवर्सरियल नेटवर्क (GAN) एक जनरेटर और एक डिस्क्रिमिनेटर को एक दूसरे के खिलाफ लड़ा देता है, वास्तविक छवियां, टेक्स्ट, या अन्य डेटा पैदा करता है।
78. सेल्फ-सुपरवाइज्ड लर्निंग बड़े मात्रा में अनलेबल्ड डेटा का उपयोग करता है, कृत्रिम ट्रेनिंग टास्क्स (उदाहरण के लिए, गायब हिस्सों को भविष्यवाणी करना) बनाकर।
79. फाउंडेशन मॉडल बड़े प्री-ट्रेन्ड मॉडल हैं जो कई डाउनस्ट्रीम टास्क्स में अनुकूलित किए जा सकते हैं।
80. मल्टीमोडल लर्निंग कई स्रोतों (उदाहरण के लिए, टेक्स्ट, छवियां, ऑडियो) से डेटा को एकत्रित करता है, अधिक समृद्ध प्रतिनिधित्व बनाता है।
81. डेटा लेबलिंग अक्सर $\frac{1}{n}$ का सबसे समय लेने वाला हिस्सा होता है, सटीकता के लिए सावधानी से एनोटेशन की आवश्यकता होती है।
82. एज कंप्यूटिंग $\frac{1}{n}$ इनफेरेन्स को डेटा स्रोत के करीब लाता है, लेटेंसी और बैंडविड्थ उपयोग को कम करता है।
83. फेडरेटेड लर्निंग मॉडलों को डेटा सैंपल्स को रखने वाले डिसेंट्रलाइज्ड डिवाइसों या सर्वरों पर ट्रेन करता है, बिना उन्हें आदान-प्रदान किया।
84. प्राइवसी-प्रेसर्विंग $\frac{1}{n}$ तकनीकें जैसे डिफरेंशियल प्राइवसी और होमोमोर्फिक एन्क्रिप्शन संवेदनशील डेटा को संरक्षित करने के लिए शामिल हैं।
85. एक्सप्लेनेबल $\frac{1}{n}$ (XAI) जटिल मॉडलों के निर्णयों को मानवों के लिए अधिक समझने योग्य बनाने का लक्ष्य रखता है।
86. वेश्यावृत्ति और न्याय $\frac{1}{n}$ में सावधानी से निगरानी की आवश्यकता होती है, क्योंकि मॉडल सामाजिक वेश्यावृत्ति को सीख सकते हैं और बढ़ा सकते हैं।
87. अवधारणा ड्रिफ्ट तब होता है जब लक्ष्य चर के सांख्यिकीय गुण समय के साथ बदल जाते हैं, मॉडल प्रदर्शन को प्रभावित करते हैं।
88. ए/बी टेस्टिंग दो या अधिक मॉडल वरजनों को तुलना करता है, ताकि पता चल सके कि वास्तविक वातावरण में कौन बेहतर प्रदर्शन करता है।
89. $\frac{1}{n}$ एक्सीलरेशन ग्राफिक्स कार्डों पर समानांतर गणना का उपयोग करके $\frac{1}{n}$ ट्रेनिंग को तेजी से बढ़ाता है।
90. $\frac{1}{n}$ (टेंसर प्रोसेसिंग यूनिट्स) $\frac{1}{n}$ द्वारा विकसित विशेष हार्डवेयर एक्सीलरेटर्स हैं, जो गहन लर्निंग वर्कलोड्स के लिए अधिक कुशल हैं।
91. ओपन-सोर्स फ्रेमवर्क (उदाहरण के लिए, टेंसरफ्लो, $\frac{1}{n}$) $\frac{1}{n}$ मॉडल विकास के लिए ब्लॉक और उपकरण प्रदान करते हैं।
92. मॉडल सर्विंग एक अभ्यास है जहां ट्रेन्ड मॉडल को डिप्लॉय किया जाता है ताकि वे रियल-टाइम या बैच भविष्यवाणियों को संभाल सकें।
93. स्केलेबिलिटी बड़े डेटासेट या भारी ट्रैफिक को संभालने के लिए आवश्यक है, वितरित ट्रेनिंग और इनफेरेन्स स्ट्रेटेजी की आवश्यकता होती है।
94. $\frac{1}{n}$ $\frac{1}{n}$ विकास को ऑपरेशन प्रैक्टिस के साथ संयोजित करता है, पुनरुत्पादन, टेस्टिंग और निरंतर एकीकरण पर ध्यान केंद्रित करता है।
95. डेटा और मॉडलों के लिए वर्सन कंट्रोल एकसंगत प्रयोग ट्रैकिंग और सहयोग को सुनिश्चित करता है।

96. डिप्लॉयमेंट स्ट्रेटेजी (उदाहरण के लिए, कंटेनर्स, माइक्रोसर्विसेज) मॉडलों को कैसे पैकेज और स्केल पर सेवा दी जाती है, को संगठित करते हैं।
97. मॉनिटरिंग डिप्लॉयमेंट के बाद मॉडल प्रदर्शन को ट्रैक करता है, डिग्रेडेशन या अनोमलियों की निगरानी करता है।
98. रिट्रेनिंग और मॉडल अपडेट्स नए डेटा और बदलते परिस्थितियों के साथ मॉडलों को अद्यतन रखता है।
99. टाइम कॉम्प्लेक्सिटी (O -नोटेशन) एक एल्गोरिथम के रनटाइम को कैसे इनपुट आकार के साथ स्केल होता है, $O(1)$ स्थिर समय को इंगित करता है।
100. ML का भविष्य अधिक जटिल और सामान्य मॉडलों की ओर बढ़ता है, लेकिन नैतिक, सामाजिक और पर्यावरणीय विचारों को संबोधित करना होगा।