

জেনারেটিভ আর্টিফিশিয়াল ইন্টেলিজেন্স ফাউন্ডামেন্টালস



1. ভূমিকা

কৃত্রিম বুদ্ধিমত্তা একবিংশ শতাব্দীর অর্থনীতির অন্যতম চালিকা শক্তি হয়ে উঠেছে, বিশেষ করে জেনারেটিভ কৃত্রিম বুদ্ধিমত্তার বিকাশের কারণে। এই নিবন্ধটি ব্যাখ্যা করে যে জেনারেটিভ কৃত্রিম বুদ্ধিমত্তা কী এবং এর ভিত্তি কী, বড় ভাষা মডেলগুলিতে বিশেষ মনোযোগ দেওয়া। একটি সংক্ষিপ্ত ঐতিহাসিক পর্যালোচনা বিশ্লেষণ করে কিভাবে এআই আজ পর্যন্ত বিবর্তিত হয়েছে এবং তারপর সবচেয়ে প্রাসঙ্গিক মডেল এবং কৌশল উপস্থাপন করে। এই পর্যন্ত ব্যবহৃত কৌশলগুলির সীমাবদ্ধতা এবং চ্যালেঞ্জগুলি বিশ্লেষণ করা হয়েছে এবং এটি ব্যাখ্যা করা হয়েছে যে কিভাবে জেনারেটিভ কৃত্রিম বুদ্ধিমত্তা সমস্যা সমাধানের জন্য নতুন এবং আরও ভাল বিকল্প সরবরাহ করে। এই প্রযুক্তিগুলির ব্যবহারের সাথে যুক্ত নৈতিক এবং সামাজিক দিকগুলিও খুব প্রাসঙ্গিক এবং এই কাগজ জুড়ে অন্বেষণ করা হয়েছে। অবশেষে, ভবিষ্যত প্রবণতা উপস্থাপন করা হয় সেইসাথে বিশ্বকে পরিবর্তন করতে এই প্রযুক্তির সম্ভাব্যতার লেখকদের দৃষ্টিভঙ্গি।

2. সংজ্ঞা এবং মৌলিক

এই বিভাগটি জেনারেটিভ কৃত্রিম বুদ্ধিমত্তা (AI) বর্ণনা করে, এর সবচেয়ে উল্লেখযোগ্য এবং পার্থক্যকারী উপাদানগুলির উপর জোর দেয়। এটি কৃত্রিম বুদ্ধিমত্তার বিবর্তনের একটি সংক্ষিপ্ত পর্যালোচনা এবং এটি কিভাবে জেনারেটিভ এআই-এর উত্থানের দিকে পরিচালিত করেছে তা আমরা আজ জানি। পরিশেষে, আমরা আজ যেখানে আছি সেখানে AI এর অগ্রগতির একটি সারসংক্ষেপ উপস্থাপন করা হল।

2.1. জেনারেটিভ এআই কি?

জেনারেটিভ এআইকে কৃত্রিম বুদ্ধিমত্তার একটি শাখা হিসাবে সংজ্ঞায়িত করা হয় যা বিশেষত্ব সিস্টেমের মতো করে শুধুমাত্র বিদ্যমান ডেটা বিশ্লেষণ বা কাজ করার বিপরীতে নতুন বিষয়বস্তু তৈরি করতে সক্ষম (Vaswani et al., 2017)। এটি আজ অবধি ব্যবহৃত বুদ্ধিমান সিস্টেমগুলির উপর একটি বাস্তব বিবর্তন যা উদাহরণস্বরূপ, নিউরাল নেটওয়ার্ক, কেস-ভিত্তিক যুক্তি সিস্টেম, জেনেটিক অ্যালগরিদম, ফাজি লজিক (নগুয়েন এট আল।, 2013) বা হাইব্রিড এআই মডেল (গালা) এর উপর ভিত্তি করে et al., 2016; Abraham et al., 2009; Aiken, 2002; Corchado et al., 2021) এবং মডেল এবং অ্যালগরিদম যা নির্দিষ্ট সমস্যার জন্য নির্দিষ্ট ডেটা ব্যবহার করে এবং ইনপুট ডেটার ভিত্তিতে একটি নির্দিষ্ট উত্তর তৈরি করে।

জেনারেটিভ কৃত্রিম বুদ্ধিমত্তা একটি কর্পাস বা ডেটাসেটে প্রশিক্ষিত বৈষম্যমূলক বা রূপান্তরমূলক মডেলগুলিকে অন্তর্ভুক্ত করে, একটি উচ্চ-মাত্রিক সূত্র স্থানে ইনপুট তথ্য ম্যাপ করতে সক্ষম। এছাড়াও, এটির একটি জেনারেটিভ মডেল রয়েছে যা স্টোকাস্টিক আচরণকে চালিত করে, প্রতিটি প্রচেষ্টায় অভিনব বিষয়বস্তু তৈরি করে, এমনকি একই ইনপুট উদ্দীপনা সহ। এই মডেলগুলি সুনির্দিষ্ট পদ্ধতির উপর নির্ভর করে তত্ত্বাবধানহীন, আধা-তত্ত্বাবধানে বা তত্ত্বাবধানে শিক্ষা সম্পাদন করতে পারে। যদিও এই কাগজটির লক্ষ্য জেনারেটিভ এআই-এর পূর্ণ সম্ভাবনা উপস্থাপন করা, তবে সেখান থেকে সাধারণীকরণের জন্য বৃহৎ ভাষা মডেলের (এলএলএম) উপর ফোকাস করা হয়েছে (চ্যাং এট আল।, 2023)। LLM হল জেনারেটিভ আর্টিফিশিয়াল ইন্টেলিজেন্স (AI) এর একটি উপশ্রেণি। জেনারেটিভ এআই বলতে এমন মডেল এবং প্রযুক্তি বোঝায় যা নতুন এবং মূল বিষয়বস্তু তৈরি করার ক্ষমতা রাখে এবং এই ডোমেনের মধ্যে, এলএলএম টেক্সট তৈরিতে বিশেষজ্ঞ। একটি LLM যেমন OpenAI-এর GPT (জেনারেটিভ প্রি-ট্রেন্ড ট্রান্সফরমার) মূলত পাঠ্য তৈরি করতে প্রশিক্ষিত হয়, অথবা পাঠ্যের অনুচ্ছেদ সহ প্রশ্নের উত্তর দিতে হয় (গুয়ান এট আল।, 2020)। একবার প্রশিক্ষিত হলে, এটি সম্পূর্ণ বাক্য এবং অনুচ্ছেদ তৈরি করতে পারে যেগুলি সুসঙ্গত এবং অনেক ক্ষেত্রে, মানুষের দ্বারা লিখিত থেকে আলাদা করা যায় না, শুধুমাত্র একটি প্রাথমিক উদ্দীপনা বা প্রম্পট থেকে (ম্যাডোটা এট আল।, 2021)। যদিও জেনারেটিভ AI মডেলগুলিকে অন্তর্ভুক্ত করে যা অন্যান্য ধরনের সামগ্রী যেমন ইমেজ (যেমন, DALL-E, OpenAI থেকেও) বা সঙ্গীত তৈরি করতে পারে, এলএলএমগুলি বিশেষভাবে ভাষার ডোমেনের উপর ফোকাস করে (এডামস, এট আল।, 2023)। LLMsকে তাই এর বিস্তৃত বিভাগের একটি অংশ বা উপসেট হিসাবে বিবেচনা করা যেতে পারে

জেনারেটিভ এআই।

এলএলএম হল নিউরাল নেটওয়ার্ক যা ক্রমিক ডেটা প্রক্রিয়া করার জন্য ডিজাইন করা হয়েছে (বুবেক এট আল।, 2023)। একটি এলএলএম পার্ঠের একটি কর্পাস (ডিজিটাইজড বই, ডাটাবেস, ইন্টারনেট থেকে তথ্য, ইত্যাদি) উপর প্রশিক্ষিত হতে পারে; ইনপুট টেক্সট পূর্ববর্তী তথ্য প্রদত্ত একটি ক্রমানুসারে টেক্সট তৈরি করতে শেখার জন্য ব্যবহার করা যেতে পারে। ট্রান্সফরমারগুলি সম্ভবত এই এলএলএমগুলির নির্মাণে সর্বাধিক ব্যবহৃত মডেল (ভন অসওয়াল্ড এট আল।, 2023)। লার্জ স্কেল ল্যাঙ্গুয়েজ মডেল (LLMs) একচেটিয়াভাবে ট্রান্সফরমার ব্যবহার করে না, যদিও ট্রান্সফরমার, বিশেষ করে ভাসওয়ানি এট আল-এর "অ্যাটেনশন ইজ অল ইউ নিড" কাগজে প্রবর্তিত আর্কিটেকচার। 2017 সালে, প্রাকৃতিক ভাষা প্রক্রিয়াকরণের জন্য বিশেষভাবে কার্যকর প্রমাণিত হয়েছে (Nadkarni et al., 2011) এবং অনেক জনপ্রিয় LLM যেমন GPT এবং BERT-এর ভিত্তি হয়েছে। যাইহোক, ট্রান্সফরমার জনপ্রিয় হওয়ার আগে, পুনরাবৃত্ত নিউরাল নেটওয়ার্ক (RNN) এবং তাদের রূপগুলি, যেমন LSTM (লং শর্ট-টার্ম মেমরি) এবং GRU (গেটেড রেকারেন্ট ইউনিট) নেটওয়ার্কগুলি সাধারণত প্রাকৃতিক ভাষা প্রক্রিয়াকরণের কাজগুলিতে সিকোয়েন্স মডেল করতে ব্যবহৃত হত। (Sherstinsky, 2020; Tang et al., 2020)।

কৃত্রিম বুদ্ধিমত্তা এবং প্রাকৃতিক ভাষা প্রক্রিয়াকরণের ক্ষেত্রে গবেষণা অব্যাহত থাকায়, এটা সম্ভব যে নতুন স্থাপত্য এবং পদ্ধতির আবির্ভাব হতে পারে যা ভবিষ্যতের এলএলএম-এ ট্রান্সফরমারগুলির সাথে বা পরিবর্তে ব্যবহার করা যেতে পারে। এইভাবে, যদিও ট্রান্সফরমারগুলি বর্তমানে এলএলএম-এর জন্য একটি প্রভাবশালী স্থাপত্য, তবে তারা ই ব্যবহৃত একমাত্র স্থাপত্য নয়, তবে ব্যাকরণগতভাবে সঠিক এবং শব্দার্থগতভাবে অর্থপূর্ণ নতুন পার্ঠ তৈরি করার ক্ষেত্রে এগুলি সবচেয়ে নির্ভরযোগ্য (Vaswani et al., 2017)। এটি তিনটি নির্দিষ্ট উপাদানের কারণে: (ক) প্রথমটি হল পজিশনাল কোডিং মেকানিজমের ব্যবহার, যা নেটওয়ার্ককে একটি বাক্যের মধ্যে একটি শব্দের জন্য একটি অবস্থান নির্ধারণ করতে দেয় যাতে এই অবস্থানটি নেটওয়ার্কের ইনপুট ডেটার অংশ হয়। এর মানে হল যে শব্দ অর্ডার তথ্য নেটওয়ার্কের কার্যক্রমের অংশ না হয়ে নিজেই ডেটার অংশ হয়ে ওঠে, যাতে নেটওয়ার্ক প্রশিক্ষিত হয়, প্রচুর পার্ঠ ডেটা সহ, এটি কীভাবে অবস্থানগত কোডিংকে ব্যাখ্যা করতে হয় এবং শব্দগুলিকে সুসংগতভাবে অর্ডার করতে শেখে। প্রশিক্ষণে ব্যবহৃত ডেটা; (b) দ্বিতীয়ত মনোযোগ প্রসঙ্গ (গ) অবশেষে, স্ব-মনোযোগ বা স্বতঃপ্রত্যাবর্তনমূলক মনোযোগ, লিঙ্গ এবং ক্রম ছাড়াও ভাষার বৈশিষ্ট্যগুলির আরও ভাল জ্ঞানের অনুমতি দেয়, যেমন সমার্থক শব্দ, যা একাধিক উদাহরণের বিশ্লেষণের মাধ্যমে চিহ্নিত করা হয়।

ক্রিয়া সংযোজন, বিশেষণ ইত্যাদির ক্ষেত্রেও একই কথা। পূর্ববর্তী পদ্য বা শব্দ ফ্রিকোয়েন্সির উপর ভিত্তি করে গুরুত্ব প্রদান করে একটি শব্দের প্রকৃত অর্থগত গুরুত্বকে ভুলভাবে উপস্থাপন করতে পারে; বিপরীতে, স্ব-মনোযোগ মডেলগুলিকে একটি ইনপুট পার্ঠের মধ্যে দীর্ঘমেয়াদী শব্দার্থিক সম্পর্কগুলি ক্যাপচার করতে দেয়, এমনকি যখন সেই পার্ঠটি সমান্তরালভাবে বিভক্ত এবং প্রক্রিয়া করা হয় (ভাসওয়ানি এট আল।, 2017)। টেক্সট জেনারেশন হল কন্টেন্ট এবং সিকোয়েন্স তৈরি করা, উদাহরণস্বরূপ, প্রোটিন, অডিও, কম্পিউটার কোড বা দাবা চাল (Eloundou et al., 2023)।

ট্রান্সফরমারগুলির উন্নয়নে অ্যালগরিদম স্তরে অগ্রগতি, উদাহরণস্বরূপ, বর্তমান গণনা ক্ষমতা এবং লেবেলবিহীন ডেটা সহ প্রাক-প্রশিক্ষণ এবং প্রশিক্ষণকে পরিমার্জিত করার ক্ষমতা (সুক্ষ্ম টিউনিং) এই মহান এআই বিপ্লবকে চালিত করেছে। মডেলের কার্যকারিতা গণনার স্কেলের উপর ব্যাপকভাবে নির্ভর করে, যার মধ্যে রয়েছে প্রশিক্ষণের জন্য ব্যবহৃত গণনার শক্তির পরিমাণ, মডেলের পরামিতিগুলির সংখ্যা এবং ডেটাসেটের আকার। একটি এলএলএম-এর প্রাক-প্রশিক্ষণের জন্য শত শত বা হাজার হাজার জিপিইউ এবং সপ্তাহ থেকে কয়েক মাস ডেডিকেটেড প্রশিক্ষণের সময় প্রয়োজন। উদাহরণস্বরূপ, এটি অনুমান করা হয় যে 175 বিলিয়ন প্যারামিটার সহ একটি GPT-3 মডেলের জন্য চালানো একটি একক প্রশিক্ষণ, 300 বিলিয়ন টোকেনে প্রশিক্ষিত, শুধুমাত্র কম্পিউটেশনাল খরচে পাঁচ বিলিয়ন ডলার খরচ হতে পারে।

এলএলএমগুলি প্রচুর পরিমাণে লেবেলবিহীন ডেটার উপর প্রাক-প্রশিক্ষিত হতে পারে। উদাহরণস্বরূপ, GPT-কে লেবেলবিহীন পার্ঠ ডেটার উপর প্রশিক্ষণ দেওয়া হয়, যা এটিকে সুস্পষ্ট নির্দেশনা ছাড়াই মানুষের ভাষায় নিদর্শন শিখতে দেয় (Radford and Narasimhan, 2018)। যেহেতু লেবেলবিহীন ডেটা লেবেলযুক্ত ডেটার চেয়ে অনেক বেশি প্রচলিত, এটি LLM-কে অনেক বড় প্রশিক্ষণ সংস্থায় প্রাকৃতিক ভাষা সম্পর্কে শিখতে দেয় (Brown et al., 2020)। ফলস্বরূপ মডেলটি একাধিক অ্যাপ্লিকেশনে ব্যবহার করা যেতে পারে কারণ এর প্রশিক্ষণ নির্দিষ্ট কাজের একটি সেটের জন্য নির্দিষ্ট নয়।

সাধারণ-উদ্দেশ্য LLM গুলি "সূক্ষ্ম-সুরিত" হতে পারে যা আউটপুট তৈরি করতে পারে যা কোনো নির্দিষ্ট কনফিগারেশনের পূর্বের সাথে মেলে (Ouyang et al., 2022; Liu et al., 2023), যা ফাইন টিউনিং নামে পরিচিত। উদাহরণস্বরূপ, একটি LLM একটি প্রদত্ত প্রশ্নের জন্য বেশ কয়েকটি সম্ভাব্য উত্তর তৈরি করতে পারে, তবে তাদের মধ্যে কিছু ভুল বা পক্ষপাতদুষ্ট হতে পারে। এই মডেলটিকে সূক্ষ্ম-সুর করার জন্য, মানব বিশেষজ্ঞরা একটি পুরস্কার ফাংশন প্রশিক্ষণের জন্য আউটপুটগুলিকে র‍্যাঙ্ক করতে পারেন যা অন্যদের চেয়ে কিছু উত্তরকে অগ্রাধিকার দেয়। এই ধরনের পরিমার্জনগুলি মডেলের গুণমানকে উল্লেখযোগ্যভাবে উন্নত করতে পারে, একটি সাধারণ-উদ্দেশ্য মডেলকে একটি নির্দিষ্ট সমস্যা সমাধানের জন্য উপযুক্ত করে তোলে (ওইয়াং এট আল।, 2022)।

2.2. এআই থেকে জেনারেটিভ এআই পর্যন্ত ইতিহাস এবং বিবর্তন

কৃত্রিম বুদ্ধিমত্তা হল কম্পিউটার বিজ্ঞান এবং প্রযুক্তির একটি ক্ষেত্র যা কম্পিউটার সিস্টেমের বিকাশের সাথে সম্পর্কিত যে কাজগুলি সম্পাদন করতে পারে যেগুলি সাধারণত মানুষের বুদ্ধিমত্তার প্রয়োজন হয়, যেমন শেখা, সিদ্ধান্ত গ্রহণ, সমস্যা সমাধান, উপলব্ধি এবং প্রাকৃতিক ভাষা (রাসেল এবং নরভিগ, 2014)) টুরিং কৃত্রিম বুদ্ধিমত্তার কেন্দ্রীয় প্রশ্নটি সম্বোধন করেছিলেন: "যন্ত্র কি চিন্তা করতে পারে" (Turing, 1950)। শীঘ্রই, জন ম্যাকাথিহি 1956 সালে "কৃত্রিম বুদ্ধিমত্তা" শব্দটি তৈরি করেছিলেন এবং লিস্প প্রোগ্রামিং ভাষার বিকাশে অবদান রেখেছিলেন, যা অনেকের জন্য AI এর প্রবেশদ্বার ছিল (McCarthy et al., 2006)। তিনি, মারভিন মিনস্কি (MIT), Lotfali A. Zadeh (University of Berkeley, California) অথবা John Holland (University of Michigan) এর মতো অগ্রগামীদের সাথে (Zadeh, 2008)। প্রবণতা, মডেল এবং অ্যালগরিদম তাদের কাজ থেকে উদ্ভূত হয়েছে। তাদের কাজ চিন্তাধারা তৈরির দিকে পরিচালিত করেছে এবং এর ভিত্তিতে সিস্টেমগুলি তৈরি করা হয়েছে, যা চিকিৎসার মতো ক্ষেত্রে প্রকৃত অগ্রগতি নিয়ে এসেছে।

এইভাবে, কৃত্রিম বুদ্ধিমত্তার শাখা যেমন সিম্বলিক লজিক, এক্সপার্ট সিস্টেম, নিউরাল নেটওয়ার্ক (Corchado et al., 2000), ফাজি লজিক, ন্যাচারাল ল্যাঙ্গুয়েজ প্রসেসিং, জেনেটিক অ্যালগরিদম, কম্পিউটার ভিশন, মাল্টি-এজেন্ট সিস্টেম (González-Briones et al., 2018) বা সামাজিক মেশিন (হেল্ডলার এবং মুলভেইল, 2016; Chamoso et al., 2019) আবির্ভূত হয়েছে। এই সমস্ত শাখাগুলিকে উপ-শাখায় এবং এইগুলিকে অন্যগুলিতে বিভক্ত করা হয়েছে, যেমন, আজ, বিশেষীকরণের স্তরটি উচ্চতর।

বেশিরভাগ জটিল সিস্টেম একাধিক উপাদান দ্বারা প্রভাবিত হয়; তারা উৎপন্ন করে, বা একাধিক ডেটা উত্সের সাথে সম্পর্কিত, তারা সময়ের সাথে সাথে বিবর্তিত হয় এবং বেশিরভাগ ক্ষেত্রে, তারা একটি ডিগ্রি বিশেষজ্ঞ জ্ঞান ধারণ করে (Pérez-Pons et al., 2023)। এই বিষয়ে, এটি স্পষ্ট বলে মনে হচ্ছে যে বিভিন্ন স্তরে বা বিভিন্ন উত্স থেকে ডেটা বিশ্লেষণ করে এমন সংযোগবাদী কৌশলগুলির সাথে জ্ঞানের মডেলিং করতে সক্ষম প্রতীকী সিস্টেমগুলির সম্মিলিত ব্যবহার বিশ্বব্যাপী সমাধান দিতে পারে। এই ধরনের সমস্যাগুলি খুঁজে পাওয়া কঠিন নয়, উদাহরণস্বরূপ, ওষুধের ক্ষেত্রে, যেখানে জ্ঞান মডেলিং শুধুমাত্র রোগীর ডেটা বিশ্লেষণের মতো গুরুত্বপূর্ণ। মডেল ফিউশনের একটি উদাহরণ ছিল জেনেটিক বিশ্লেষণের জন্য জিন-সিবিআর প্ল্যাটফর্ম। একদিকে, এটি বেশ কয়েকটি নিউরাল নেটওয়ার্ক এবং অস্পষ্ট সিস্টেমের সাথে একটি কেস-ভিত্তিক যুক্তি সিস্টেমের সাথে দেওয়া পদ্ধতিগত কাঠামো ব্যবহার করেছে (ডিয়াজ এট আল।, 2006; হার্নান্দেজ-নিভস এট আল।, 2021)। এই মডেলটি মায়লোমা বিশ্লেষণের সুবিধার্থে তৈরি করা হয়েছিল।

1970/80 এর দশক ছিল কৃত্রিম বুদ্ধিমত্তা এবং বিতরণকৃত কম্পিউটিং এর জন্য একটি যুগান্তকারী সময় (জানবি এট আল।, 2022)। এমন এক সময়ে যখন বিশ্ব একটি নতুন শতাব্দীর দিকে আসছে এবং যেখানে কম্পিউটিং জগতের মনোযোগ এআই-এর অগ্রগতির চেয়ে ইন্টারনেটের সম্ভাবনার দিকে বেশি নিবদ্ধ ছিল, ইন্টারনেটের যাত্রা বন্ধের সাথে সাথে একটি মহান পরিবর্তনের সময়। এই বাস্তবতা, হার্ডওয়্যারের সীমাবদ্ধতার সাথে মিলিত, AI এর প্রতি শিল্পের অনাগ্রহ এবং বিম্লিত ধারণার অভাব ক্ষেত্রের স্ববিরতার সময়কালের সূচনায় অবদান রাখে, যা "AI শীতকাল" নামে পরিচিত।

তবে শীতের পরে একটি গ্রীষ্ম আসে এবং এটি শতাব্দীর শুরুতে এসেছিল, যাকে আমরা ডিপ লার্নিং এবং কনভোলুশনাল নিউরাল নেটওয়ার্ক (সিএনএন) বলি। এটি একটি প্রধান ধারণা যা আমাদের তথ্যের সাথে মোকাবিলা করার পদ্ধতিতে একটি আমূল পরিবর্তন এনেছে। এই নেটওয়ার্কগুলি মেশিন লার্নিং কৌশলগুলিকে কিছুটা ভিন্ন উপায়ে ব্যবহার করে যেগুলি কীভাবে তাদের মূল ধারণা করা হয়েছিল

(Bengio, 2009; Pérez-Pons et al., 2021; Hernández et al., 2021)। অন্যান্য মডেলের বিপরীতে, তাদের একাধিক লুকানো স্তর রয়েছে যা ক্রমবর্ধমান জটিল এবং বিমূর্ত পদ্ধতিতে ইনপুট ডেটা থেকে বৈশিষ্ট্য এবং প্যাটার্নগুলি বের করার অনুমতি দেয় (পরিখ এট আল।, 2022)। এখানে, একটি একক অ্যালগরিদম বিভিন্ন দৃষ্টিকোণ থেকে একটি সমস্যার সমাধান করে।

এই মডেলগুলি আগে এবং পরে প্রতিনিধিত্ব করে এবং আমরা কীভাবে কাজ করি তা বিপ্লব করতে বাধ্য। জ্ঞান সৃষ্টির এই নতুন মডেলগুলি ব্যবহার করে ডিজিটাল, ভৌত এবং জৈবিক প্রযুক্তির রূপান্তরের মাধ্যমে সিস্টেম তৈরি করার আমাদের দক্ষতার জন্য এটি পঞ্চম শিল্প বিপ্লবের সূচনা (Corchado, 2023)। আমরা যদি দ্রুত গতিশীল বিশ্বে বাস করতাম, তাহলে আমাদের এখন ক্রমাগত স্বরণের বিশ্বের জন্য প্রস্তুত হতে হবে। যারা এই অগ্রগতির সাথে তাল মিলিয়ে চলেছেন তারা দেখতে পাবেন তাদের ব্যবসা, মূল্য উত্পাদন এবং পরিষেবার সুযোগগুলি আগামী বছরগুলিতে দ্রুত বৃদ্ধি পাবে।

ডিপ লার্নিং হল মেশিন লার্নিং এর একটি উপশ্রেণি যা মস্তিষ্কের গঠন এবং কার্যকারিতা দ্বারা অনুপ্রাণিত অ্যালগরিদমের উপর ফোকাস করে, যাকে বলা হয় কৃত্রিম নিউরাল নেটওয়ার্ক (Chan, et al., 2016; Kothadiya et al., 2022, Alizadehsani et al., 2023)। এই নেটওয়ার্কগুলি, বিশেষ করে যখন তাদের অনেকগুলি (গভীর) স্তর থাকে, বিভিন্ন ধরনের AI কাজগুলিতে অত্যন্ত কার্যকর বলে প্রমাণিত হয়েছে। গভীর শিক্ষা-ভিত্তিক জেনারেটিভ মডেলগুলি স্বয়ংক্রিয়ভাবে ডেটা উপস্থাপন করতে শিখতে পারে এবং নতুনগুলি তৈরি করতে পারে যা মূল ডেটা বিতরণের অনুরূপ।

সিএনএন হল স্নায়ু নেটওয়ার্কের একটি বিশেষ শ্রেণী যা একটি গ্রিড-সদৃশ কার্টামোর সাথে ডেটা প্রক্রিয়া করার জন্য ডিজাইন করা হয়েছে, যেমন একটি চিত্র। এগুলি কম্পিউটার ভিশন টাস্কের কেন্দ্রীয়। জেনারেটিভ এআই-এর প্রেক্ষাপটে, সিএনএনগুলিকে ছবি তৈরি করার জন্য অভিযোজিত করা হয়েছে। উদাহরণস্বরূপ, জেনারেটিভ অ্যানট্যাগনিস্টিক নেটওয়ার্ক (GANs) প্রায়ই তাদের জেনারেটর এবং বৈষম্যকারীদের মধ্যে CNN ব্যবহার করে বাস্তবসম্মত ছবি তৈরি করতে।

2014 সালে ইয়ান গুডফেলো এবং তার সহযোগীদের দ্বারা প্রবর্তিত GAN, দুটি নিউরাল নেটওয়ার্ক, একটি জেনারেটর এবং একটি বৈষম্যকারী, যা একসাথে প্রশিক্ষিত হয় (গুডফেলো এট আল।, 2014)। জেনারেটর ডেটা (যেমন চিত্র) তৈরি করার চেষ্টা করে, যখন বৈষম্যকারী প্রকৃত ডেটা এবং উৎপন্ন ডেটার মধ্যে পার্থক্য করার চেষ্টা করে। প্রশিক্ষণের অগ্রগতি হওয়ার সাথে সাথে, জেনারেটর এমন ডেটা তৈরি করতে আরও ভাল হয় যা বৈষম্যকারীকে প্রতারণা করে। সিএনএনগুলি প্রায়শই GAN আর্কিটেকচারে চিত্র-সম্পর্কিত কাজের জন্য ব্যবহৃত হয়। অন্যদিকে, ভেরিয়েশনাল অটোএনকোডার (VAEs), উদাহরণস্বরূপ, নিউরাল নেটওয়ার্কের উপর ভিত্তি করে অন্য ধরনের জেনারেটিভ মডেল (Wei & Mahmood, 2020)। GAN-এর বিপরীতে, VAEs সুস্পষ্টভাবে ডেটার জন্য একটি সম্ভাব্যতা বন্টন মডেল করে এবং প্রশিক্ষণের জন্য ভিন্নতামূলক অনুমান কৌশল ব্যবহার করে। এছাড়াও, পিট্রোল-ভিত্তিক মডেলগুলি (Su et al., 2021) হল জেনারেটিভ এআই ফ্রেমওয়ার্ক যা গভীর শিক্ষার উপর ভিত্তি করে এবং পিট্রোল-বাই-পিট্রোল ভিত্তিতে ছবি তৈরি করে, পুনরাবৃত্ত নিউরাল নেটওয়ার্ক বা CNN ব্যবহার করে।

গভীর শিক্ষা, বিশেষ করে, কনভোলুশনাল নেটওয়ার্ক, অনেক জেনারেটিভ এআই মডেলের বিকাশ এবং সাফল্যের মৌলিক হাতিয়ার হয়েছে, বিশেষ করে যেগুলি ইমেজ তৈরির উপর দৃষ্টি নিবদ্ধ করে। এই কৌশলগুলি অনেক ক্ষেত্রে বাস্তব বিষয়বস্তু থেকে আলাদা নয় এমন বিষয়বস্তু তৈরি করতে মডেলগুলির ক্ষমতার ক্ষেত্রে উল্লেখযোগ্য অগ্রগতি সক্ষম করেছে।

উদাহরণস্বরূপ, চ্যাটজিপিটি আমাদের জীবনে এসেছে এবং সেগুলিকে পরিবর্তন করেছে, এবং আমরা খুব কমই লক্ষ্য করেছি। কিছু লোক এটির কথা শুনেছে, অন্যরা এটি মাঝে মাঝে ব্যবহার করেছে, এবং আমাদের মধ্যে অনেকেই ইতিমধ্যে প্রকল্পগুলিতে কাজ করছি এবং এই প্রযুক্তির সাথে মান তৈরি করছি। টেক্সট লিখতে, অ্যালগরিদম তৈরি করতে, যুক্তিযুক্ত প্রস্তাবগুলি সংশ্লেষণ এবং তৈরি করতে এই সরঞ্জামটির ক্ষমতা অসাধারণ, তবে এটি কেবল আইসবার্গের টিপ। এটি ইতিমধ্যেই অন্যদের মধ্যে গ্রাহক পরিষেবা, মেডিকেল ডেটা বিশ্লেষণ, সিদ্ধান্ত সমর্থন এবং ডায়াগনস্টিকসের জন্য সিস্টেম তৈরি করতে ব্যবহৃত হচ্ছে।

কিন্তু চ্যাটজিপিটি এই ধরনের প্রথম সিস্টেম যা বাজারে প্রবেশ করেছে। আরও অনেক মডেল এবং টুল আছে: BARD, XLNet, T5, RoBERTA, Bedrock, Wu Dao, Nemo, LLAMA 2, ইত্যাদি। প্রযুক্তি যেমন এটি প্রমাণ এবং ক্রমিকালের উপর ভিত্তি করে অনেক বেশি সঠিক ডায়াগনস্টিক সিস্টেমের বিকাশের অনুমতি দেবে। রেকর্ড, টেলিমেডিসিনের আরও ব্যাপক ব্যবহার, তাদের বাড়িতে দীর্ঘস্থায়ী রোগীদের পর্যবেক্ষণের ব্যবস্থা ইত্যাদি। এই বিষয়ে, চিকিৎসা ক্ষেত্রের জন্য দারুণ আগ্রহের অ্যালগরিদম হচ্ছে

বিভিন্ন স্তরে বিকশিত হয়েছে, যেমন ট্রান্সফরমার, অটোএনকোডার, গভীর শক্তি-ভিত্তিক জেনারেটিভ মডেল, প্রোটোটাইপের পরিবর্তনশীল অনুমান মডেল, কার্যকারণ অনুমান সহ শক্তিবৃদ্ধি শেখার সিস্টেম, নাম কিন্তু কয়েকটি। AI-তে আমাদের জীবনযাপন এবং কাজ করার পদ্ধতিকে মৌলিকভাবে পরিবর্তন করার সম্ভাবনা রয়েছে, তবে এটি গোপনীয়তা এবং নিরাপত্তার ক্ষেত্রে গুরুত্বপূর্ণ নৈতিক চ্যালেঞ্জও তৈরি করে যা সমাধান করা প্রয়োজন।

2.3. ট্র্যাডিশনাল এআই থেকে জেনারেটিভ এআই-তে স্থানান্তর

কৃত্রিম বুদ্ধিমত্তার (AI) ইতিহাস সমৃদ্ধ এবং আকর্ষণীয় এবং অন্য সব কিছুর মতো এর বিভিন্ন ব্যাখ্যা এবং মূল উপাদান থাকতে পারে। এখানে কিছু অতীন্দ্রিয় উপাদানের সংক্ষিপ্তসার রয়েছে যা আমাদের প্রথম কৃত্রিম নিউরনের উপস্থিতি থেকে প্রথম ট্রান্সফরমার নির্মাণ এবং ChatGPT-এর জনপ্রিয়করণ পর্যন্ত এই ক্ষেত্রের বিবর্তনকে দ্রুত বিশ্লেষণ করতে দেয়:

1. **কৃত্রিম নিউরন (1943):** ওয়ারেন ম্যাককালোচ এবং ওয়াল্টার পিটস "নার্তাস অ্যাক্টিভিটিতে আইডিয়াস ইমেনেন্টের একটি লজিক্যাল ক্যালকুলাস" প্রকাশ করেছেন, যেখানে ম্যাককালোচ-পিটস নিউরন নামে পরিচিত একটি জৈবিক নিউরনের একটি সরলীকৃত মডেল উপস্থাপন করা হয়েছিল। এই মডেলটিকে প্রথম কৃত্রিম নিউরন হিসাবে বিবেচনা করা হয় এবং এটি কৃত্রিম নিউরাল নেটওয়ার্কের ভিত্তি (McCulloch & Pitts, 1943)।
2. **পারসেপ্টরন (1957-1958):** ফ্র্যাঙ্ক রোজেনব্ল্যাট পারসেপ্টরন প্রবর্তন করেছেন, একক-স্তর নিউরাল নেটওয়ার্কের জন্য সবচেয়ে সহজ তত্ত্বাবধানে শেখার অ্যালগরিদম। যদিও এর ক্ষমতা সীমিত (যেমন, এটি XOR সমস্যার সমাধান করতে পারেনি), এটি নিউরাল নেটওয়ার্কের ভবিষ্যত উন্নয়নের ভিত্তি স্থাপন করেছে (Rosenblatt, 1958)।
3. **এআই উইন্টার (1970-1980):** প্রাথমিক মডেলের সীমাবদ্ধতা এবং গণনার ক্ষমতার অভাব এআই গবেষণার জন্য উত্সাহ এবং তহবিল হ্রাসের দিকে পরিচালিত করে। এই সময়ের মধ্যে, নিউরাল নেটওয়ার্কগুলি এআই সম্প্রদায়ের ফোকাস ছিল না (মুর, 2006)।
4. **ব্যাকপ্রপাগেশন (1986):** রুমেলহার্ট, হিন্টন এবং উইলিয়ামস মাল্টিলেয়ার নিউরাল নেটওয়ার্কের প্রশিক্ষণের জন্য ব্যাকপ্রপাগেশন অ্যালগরিদম প্রবর্তন করেছেন (Rumelhart et al., 1986)। এই অ্যালগরিদম নিউরাল নেটওয়ার্কগুলিতে আগ্রহ পুনরুজ্জীবিত করতে শুরু করে। পুনরাবৃত্ত নেটওয়ার্ক, যা ব্যাকপ্রপাগেশন ব্যবহার করে, প্রতিটি শব্দের প্রতি আলাদাভাবে এবং ক্রমানুসারে মনোযোগ দেয়। এই নেটওয়ার্কগুলি ক্রমানুসারে কাজ করে। এই নেটওয়ার্কগুলিতে, প্রতিটি শব্দ যে ক্রমে প্রদর্শিত হয় তা প্রশিক্ষণ বিবেচনা করা হয়। 1980-এর দশকের শেষের দিকে এবং 1990-এর দশকের গোড়ার দিকে আবির্ভূত পুনরাবৃত্ত নেটওয়ার্কগুলির পরিপ্রেক্ষিতে, RNNগুলি ডেটার ক্রম প্রক্রিয়া করার জন্য তৈরি এবং তৈরি করা হয়েছিল। এই নেটওয়ার্কগুলিকে প্রশিক্ষণের জন্য, ব্যাকপ্রপাগেশন খ্রু টাইম টেকনিক (BPTT) ব্যবহার করা হয়। RNNগুলি সময়ের সাথে একটি "স্থিতি" বজায় রাখতে পারে, যা তাদের সময় সিরিজের পূর্বাভাস এবং প্রাকৃতিক ভাষা প্রক্রিয়াকরণের মতো কাজের জন্য উপযুক্ত করে তোলে। যাইহোক, প্রথাগত RNNগুলি গ্রেডিয়েন্ট ভ্যানিশিং এবং গ্রেডিয়েন্ট এক্সপ্লোশনের মতো সমস্যার সম্মুখীন হয়েছিল। অনুচ্ছেদ মূল্যায়ন/প্রজন্মের মাধ্যমে অগ্রগতির সাথে সাথে পুনরাবৃত্ত নেটওয়ার্কগুলি প্রসঙ্গ হারায়, পাঠ্যটি দীর্ঘ হলে এটি একটি সমস্যা। এই সমস্যাটি ব্যাকপ্রপাগেশন, লং স্বল্প-মেয়াদী মেমরি (LSTM) সহ অন্যান্য নেটওয়ার্কগুলির দ্বারা সমাধান করা হয়েছিল, Hochreiter এবং Schmidhuber (1997) দ্বারা প্রবর্তিত, যা অদৃশ্য গ্রেডিয়েন্ট সমস্যা মোকাবেলার জন্য ডিজাইন করা RNN-এর একটি বিশেষ রূপ। LSTMs দীর্ঘমেয়াদী নির্ভরতা শিখতে পারে এবং ট্রান্সফরমারের আবির্ভাবের আগ পর্যন্ত প্রাকৃতিক ভাষা প্রক্রিয়াকরণ এবং অন্যান্য অনুক্রমিক কাজগুলিতে অনেক অগ্রগতির কেন্দ্রবিন্দু ছিল। এই নেটওয়ার্কগুলির মধ্যে রয়েছে, শেখার প্রতিটি পর্যায়ে, গাণিতিক ক্রিয়াকলাপ যা অনুচ্ছেদের শুরুতে যা শেখা হয়েছিল তা ভুলে যাওয়া থেকে বিরত রাখে। যাইহোক, এই নেটওয়ার্কগুলিতে তাদের প্রশিক্ষণকে সমান্তরাল করার অসম্ভবতা সম্পর্কিত অন্যান্য সমস্যা রয়েছে, যার ফলে বড় মডেলগুলি তৈরি করা কার্যত অসম্ভব। এই ধরনের নেটওয়ার্ক, সমস্ত প্রশিক্ষণ ক্রমিক।
5. **ডিপ লার্নিং এবং কনভোলুশনাল নিউরাল নেটওয়ার্কস (কনভোলিউশনাল নিউরাল নেটওয়ার্ক, সিএনএন, 2012):** 2012 সালে, অ্যালেক্স ক্রিজভেস্কি, ইলিয়া সূটস্কেভার এবং জিওফ্রে হিন্টন একটি রূপান্তরমূলক নিউরাল নেটওয়ার্ক উপস্থাপন করেছিলেন যা ইমেজনেট ইমেজ শ্রেণীবিভাগের চ্যালেঞ্জটি ব্যাপক ব্যবধানে জিতেছিল।

(Krizhevsky et al., 2012)। এই ইভেন্টটি "ডিপ লার্নিং" যুগের সূচনাকে চিহ্নিত করেছে নিউরাল নেটওয়ার্কগুলির প্রতি নতুন করে আগ্রহের সাথে যা 2006 সালে জনপ্রিয় হতে শুরু করেছিল, যে বছর "এআই উইন্টার" এর সমাপ্তি শুরু হয়েছিল। এই নেটওয়ার্কগুলি শ্রেণীবিভাগ এবং চিত্র প্রক্রিয়াকরণের জন্য বিশেষভাবে উপযুক্ত, স্বরগুলিতে গঠন করা হয় এবং তিনটি প্রধান উপাদানে সংগঠিত হয়: কনভোলিউশনাল লেয়ার, অ্যাক্টিভেশন লেয়ার এবং ক্লাস্টারিং লেয়ার। ফিল্টার বা কার্নেলের মাধ্যমে ইমেজ থেকে গুরুত্বপূর্ণ বৈশিষ্ট্য বের করার জন্য কনভোলিউশনাল লেয়ার দায়ী। ফিল্টারগুলি নির্দিষ্ট প্রান্ত, আকৃতি বা নিদর্শন সনাক্ত করতে গাণিতিক ক্রিয়াকলাপ সম্পাদন করে চিত্রের উপর গ্লাইড করে। অ্যাক্টিভেশন লেয়ারে, অ্যাক্টিভেশন ফাংশন (যেমন ReLU) প্রয়োগ করা হয় অ-রৈখিকতা যোগ করতে এবং নেটওয়ার্কের জটিল সম্পর্ক শেখার ক্ষমতা বাড়াতে। অবশেষে, ক্লাস্টারিং স্বরগুলি চিত্র উপস্থাপনের আকার হ্রাস করে, প্যারামিটারের সংখ্যা হ্রাস করে এবং নেটওয়ার্কটিকে প্রক্রিয়াকরণে আরও দক্ষ করে তোলে। তথ্য এই স্বরগুলির মধ্য দিয়ে যাওয়ার সাথে সাথে, CNN আরও বিমূর্ত এবং জটিল বৈশিষ্ট্যগুলিকে চিনতে শিখেছে, যা বস্তু, মানুষ বা অন্য যেকোন কিছুকে সনাক্ত করার অনুমতি দেয়। ব্যাপক তথ্য প্রক্রিয়াকরণ সিস্টেম নির্মাণ এবং সমান্তরাল প্রকল্পগুলির উন্নয়নের জন্য এই ক্ষেত্রে করা কাজটি আজ ব্যবহৃত ট্রান্সফরমারগুলির জন্ম দিয়েছে (Gerón, 2022)।

6. **ট্রান্সফরমার (2017):** ভাসওয়ানি প্রমুখ। "অ্যাটেনশন ইজ অল ইউ নিড" কাগজে ট্রান্সফরমার আর্কিটেকচার চালু করেছে। এই স্থাপত্য, মনোযোগের প্রক্রিয়ার উপর ভিত্তি করে, প্রাকৃতিক ভাষা প্রক্রিয়াকরণের জন্য অত্যন্ত কার্যকরী প্রমাণিত হয়েছে এবং GPT সহ পরবর্তী অনেক মডেলের ভিত্তি হয়ে উঠেছে। এলএসটিএম এবং গভীর শিক্ষার মতো ব্যাকপ্রোগেশন মডেলগুলির উপর এই নেটওয়ার্কগুলির সুবিধা তাদের শেখার সমান্তরাল করার ক্ষমতার মধ্যে রয়েছে। পৌনঃপুনিক নিউরাল নেটওয়ার্ক (RNNs) বা কনভোলিউশনাল নিউরাল নেটওয়ার্ক (CNNs) থেকে ভিন্ন, ট্রান্সফরমারগুলি ডেটার একটি নির্দিষ্ট অনুক্রমিক বা স্থানিক কাঠামোর উপর নির্ভর করে না, যা তাদের সমান্তরালভাবে তথ্য প্রক্রিয়া করতে এবং ডেটাতে দীর্ঘমেয়াদী নির্ভরতা ক্যাপচার করতে দেয়। এই বিষয়ে, শব্দ এমবেডিং ধারণাটি, যা ট্রান্সফরমার শেখার ভিত্তি, উল্লেখ করার মতো। এটি পাঠ্য ভেক্টরাইজেশনের জন্য প্রাকৃতিক ভাষা প্রক্রিয়াকরণের মধ্যে একটি কৌশল। ট্রান্সফরমারগুলি সমান্তরালভাবে একটি পাঠ্যের সমস্ত শব্দ বিশ্লেষণ করা সম্ভব করে এবং এইভাবে, নেটওয়ার্কের প্রক্রিয়াকরণ এবং সৃষ্টি দ্রুততর হয়। এটি বলেছে, এটি লক্ষ করা উচিত যে এই নেটওয়ার্কগুলির জন্য প্রচুর পরিমাণে ডেটা এবং খুব শক্তিশালী হার্ডওয়্যার প্রয়োজন, যেমন উপরে উল্লিখিত হয়েছে। উদাহরণস্বরূপ, GPT-3 175 বিলিয়ন প্যারামিটার এবং 45 TB ডেটা এবং GPT-4 1000,000,000,000,000 মিলিয়ন প্যারামিটার এবং একটি বড় কিন্তু অজানা সংখ্যক টিবি নিয়ে তৈরি করা হয়েছিল।
7. **GPT এবং ChatGPT (2018-2020):** OpenAI জেনারেটিভ প্রি-ট্রেন্ড ট্রান্সফরমার (GPT) মডেল সিরিজ চালু করেছে। GPT-2, 2019 সালে প্রকাশিত, সুসঙ্গত এবং বাস্তবসম্মত পাঠ্য তৈরি করার একটি চিত্তাকর্ষক ক্ষমতা প্রদর্শন করেছে। 2020 সালে প্রকাশিত GPT-3, এই ক্ষমতাগুলিকে আরও প্রসারিত করেছে এবং চ্যাটজিপিটি (আব্দুল্লাহ এট আল।, 2022) এর মতো চ্যাট-ভিত্তিক অ্যাপ্লিকেশনগুলির জনপ্রিয়করণের দিকে পরিচালিত করেছে। এই পণ্যটির চিত্তাকর্ষক অনুপ্রবেশ ক্ষমতা রয়েছে, 2 মাসে 100 মিলিয়ন ব্যবহারকারী পৌঁছেছে, যখন অন্যান্য প্ল্যাটফর্ম যেমন ইনস্টাগ্রাম একই সংখ্যক ব্যবহারকারীর কাছে পৌঁছতে 26 মাস সময় নিয়েছে (ফেসবুক 54 মাস বা টুইটার 65 মাস)।

এই সাতটি উপাদানকে অনুসন্ধান এবং তথ্যের একটি কালানুক্রমিক তালিকা হিসাবে বিবেচনা করা যেতে পারে যা AI এর উৎপত্তি থেকে উদ্ভব পর্যন্ত বিবর্তনকে প্রতিফলিত করে যা আজকে জেনারেটিভ AI নামে পরিচিত।

3. বড় ভাষার মডেল

এই বিভাগে, আমরা বড় ভাষা মডেল (LLMs) প্রবর্তন করি। একটি সাধারণ সংজ্ঞার পরে, নির্বাচিত সাফল্যের গল্পগুলি আলোচনা করা হয় (Itoh & Okada, 2023)। একটি সম্পূর্ণ অধ্যয়ন পরিচালনা করার পরিবর্তে, উদ্দেশ্য হল LLMগুলিকে হাইলাইট করা যা বর্তমানে সবচেয়ে প্রাসঙ্গিক এবং তাদের স্বতন্ত্র দিকগুলির উপর মনোযোগ দেয়।

3.1. বড় ভাষা মডেল সংজ্ঞায়িত করা

লার্জ ল্যাঙ্গুয়েজ মডেল হল কৃত্রিম বুদ্ধিমত্তার মডেল যা প্রাকৃতিক ভাষা প্রসেস এবং জেনারেট করার জন্য ডিজাইন করা হয়েছে। এই মডেলগুলিকে প্রচুর পরিমাণে পাঠ্যের উপর প্রশিক্ষিত করা হয়, যা তাদেরকে জটিল ভাষা-সম্পর্কিত কাজগুলি যেমন অনুবাদ, পাঠ্য তৈরি এবং প্রশ্নের উত্তর প্রদান করতে সক্ষম করে।

এলএলএমগুলি মূলত জনপ্রিয় হয়ে উঠেছে ট্রান্সফরমার আর্কিটেকচারে অগ্রগতি এবং উপলব্ধ গণনা ক্ষমতা বৃদ্ধির কারণে। এই মডেলগুলি অনেক পরামিতি দ্বারা চিহ্নিত করা হয়, যা তাদের মানব ভাষার জটিলতা ক্যাপচার এবং মডেল করতে দেয়।

বৃহৎ ভাষার মডেলগুলি প্রাকৃতিক ভাষা প্রক্রিয়াকরণের ক্ষেত্রে বিপ্লব ঘটিয়েছে এবং বেশ কিছু স্বতন্ত্র বৈশিষ্ট্য রয়েছে। এগুলি এলএলএমগুলির সবচেয়ে বৈশিষ্ট্যযুক্ত উপাদানগুলি:

- পরামিতি বড় সংখ্যা: এলএলএম, নাম থেকেই বোঝা যায়, বড়। উদাহরণস্বরূপ, GPT-3, সবচেয়ে পরিচিত এলএলএমগুলির মধ্যে একটি, এর 175 বিলিয়ন প্যারামিটার রয়েছে। এই বিপুল সংখ্যক পরামিতি তাদের মানব ভাষার জটিলতা ক্যাপচার এবং মডেল করতে দেয়।
- বড় কর্পাস প্রশিক্ষণ: এলএলএমগুলিকে বিশাল ডেটাসেটগুলিতে প্রশিক্ষণ দেওয়া হয় যা ইন্টারনেটের বড় অংশগুলিকে বিস্তৃত করে, যেমন বই, নিবন্ধ এবং ওয়েবসাইট। এটি তাদের ভাষা এবং বিভিন্ন বিষয়ের বিস্তৃত সাধারণ জ্ঞান অর্জন করতে দেয়।
- পাঠ্য প্রজন্মের ক্ষমতা: এলএলএম এমন টেক্সট তৈরি করতে পারে যা সুসঙ্গত, সাবলীল এবং অনেক ক্ষেত্রে মানুষের লেখা পাঠ থেকে আলাদা করা যায় না। তারা প্রবন্ধ লিখতে পারে, প্রশ্নের উত্তর দিতে পারে, কবিতা তৈরি করতে পারে এবং আরও অনেক কিছু করতে পারে।
- শিক্ষা স্থানান্তর: একবার একটি বৃহৎ কর্পাসে প্রশিক্ষিত হলে, LLM-গুলিকে তুলনামূলকভাবে অল্প পরিমাণে টাস্ক-নির্দিষ্ট ডেটা সহ নির্দিষ্ট কাজের জন্য "টিউন" করা যেতে পারে। এটি "ট্রান্সফার লার্নিং" নামে পরিচিত এবং এলএলএম এত বহুমুখী হওয়ার একটি কারণ।
- ট্রান্সফরমার আর্কিটেকচার ব্যবহার: বেশিরভাগ আধুনিক এলএলএম, যেমন GPT এবং BERT, একটি ট্রান্সফরমার আর্কিটেকচারের উপর ভিত্তি করে তৈরি, যা ডেটাতে সম্পর্ক ক্যাপচার করতে মনোযোগের প্রক্রিয়া ব্যবহার করে।
- মাল্টিমোডাল ক্ষমতা: যদিও এলএলএমগুলি ঐতিহ্যগতভাবে পাঠ্যের উপর দৃষ্টি নিবদ্ধ করেছে, সাম্প্রতিকতম মডেলগুলি মাল্টিমোডাল ক্ষমতাগুলি অন্বেষণ করেছে, যার অর্থ তারা একই সাথে পাঠ্য এবং চিত্রের মতো একাধিক ধরনের ডেটা বুঝতে এবং তৈরি করতে পারে।
- কার্য জুড়ে সাধারণীকরণ: নির্দিষ্ট স্থাপত্যগত পরিবর্তনের প্রয়োজন ছাড়াই, একজন এলএলএম অনুবাদ থেকে টেক্সট জেনারেশন পর্যন্ত বিভিন্ন ধরনের কাজ সম্পাদন করতে পারে। প্রায় দশ, যা প্রয়োজন তা হল মডেলটিকে সঠিক প্রম্পট বা উদ্দীপনা প্রদান করা।
- নৈতিক চ্যালেঞ্জ এবং পক্ষপাত: যেহেতু LLM দেয় ইন্টারনেট ডেটার উপর প্রশিক্ষণ দেওয়া হয়, তাই তারা সেই ডেটাতে উপস্থিত পক্ষপাতগুলি অর্জন করতে এবং স্থায়ী করতে পারে। এটি এই মডেলগুলির নৈতিক ব্যবহার সম্পর্কে উদ্বেগ এবং আলোচনার দিকে পরিচালিত করেছে এবং এই পক্ষপাতগুলিকে মোকাবেলা এবং প্রশমিত করার প্রয়োজনীয়তা রয়েছে।

একইভাবে, বিভিন্ন এলএলএম মডেলের বৃদ্ধি সময়ের সাথে তাত্পর্যপূর্ণ, প্রতিটি এলএলএম বিকাশকারী বিভিন্ন চাহিদা এবং সংস্থান স্তর পূরণের জন্য বিভিন্ন ধরনের অ্যাপ্লিকেশনে কাজ করে। এতে অনেক প্যারামিটার সহ বড় মডেল এবং কম প্যারামিটার সহ ছোট মডেল উভয়ই অন্তর্ভুক্ত রয়েছে। ওপেনএআই এবং গুগলের মতো কোম্পানিগুলি ক্রমবর্ধমান সংখ্যক প্যারামিটার সহ মডেলগুলি তৈরি করেছে, যেখানে এই মডেলগুলি খুব বৈচিত্র্যময় এবং জটিল কাজগুলি মোকাবেলা করতে সক্ষম এবং প্রায়শই বিস্তৃত অ্যাপ্লিকেশনগুলিতে অসাধারণভাবে ভাল সম্পাদন করে। যাইহোক, META কোম্পানির ক্ষেত্রে তার Llama 2 মডেলের মডেলের বিভিন্ন প্যারামিটারাইজড সংস্করণের কারণে আলোড়ন সৃষ্টি করেছে এবং কম হার্ডওয়্যার পারফরম্যান্স পরিবেশে চালানোর জন্য অস্টাইজ করা হচ্ছে। নিম্নলিখিত সারণী 1 এই মডেলগুলির কিছু সম্পর্কিত ডেটা দেখায়:

সারণী 1. এলএলএম মডেল।

মডেলের নাম	কোম্পানি	প্যারামিটারের সংখ্যা	প্রশিক্ষণ তথ্য পরিমাণ	ওয়েবসাইট
জিপিটি-৩	OpenAI	175 বিলিয়ন	প্রায় 570GB (ওয়েবটেক্সট, বই, অন্যান্য)	https://openai.com/
BERT- বড়	গুগল	340 মিলিয়ন	উইকিপিডিয়া + বুককর্পাস	https://research.google/pubs/pub45413/
T5 (টেক্সট-টু-টেক্সট ট্রান্সফরমার)	গুগল এআই	সংস্করণের উপর নির্ভর করে পরিবর্তিত হয় (60 মিলিয়ন থেকে 11 বিলিয়ন পর্যন্ত)	C4 (সাধারণ ফ্রল)	https://ai.googleblog.com/2020/02/অন্বেষণ-ট্রান্সফার-লাইং-with-t5.html
রবারটা	ফেসবুক এআই	সংস্করণের উপর নির্ভর করে পরিবর্তিত হয় (RBERT-লার্জের জন্য 355 মিলিয়ন পর্যন্ত)	ওয়েবটেক্সট, ওপেনওয়েবটেক্সট এবং অন্যান্য সহ অসংখ্য ডেটাসেট	https://ai.meta.com/blog/রবারটা-একটি-অপ্টিমাইজড-পদ্ধতি-প্রধান-প্রশিক্ষণ-নিজে-তত্ত্বাবধানে-এনএলপি-সিস্টেম/
এক্সএলনেট	গুগল/সিএমইউ	340 মিলিয়ন পর্যন্ত	উইকিপিডিয়া এবং বুককর্পাস সহ বিভিন্ন ডেটাসেট	https://github.com/zihangdai/xlnet
ক্লিপ	OpenAI	281 মিলিয়ন	ইন্টারনেট ছবি + সংশ্লিষ্ট পাঠ্য	https://openai.com/research/ক্লিপ
ডাল-ই	OpenAI	প্রায় 12 বিলিয়ন (GPT-3 এর উপর ভিত্তি করে)	ছবি এবং টেক্সট বিবরণ	https://openai.com/research/থেকে-ই
কল করুন 2	মেটা এআই	1000 মিলিয়ন	1000 মিলিয়ন শব্দ	https://github.com/ফেসবুক/রিসার্চ/লাম্বা
উ দাও	বেইজিং একাডেমি অফ আর্টিফিশিয়াল ইন্টেলিজেন্স (BAAI)	1.75 ট্রিলিয়ন	4.9 টেরাবাইট পাঠ্য এবং কোড	https://www.baai.ac.cn/
LaMDA	গুগল এআই	137 বিলিয়ন	গুগল ডাটাবেস	https://bard.google.com/
পাএলএম	গুগল এআই	540 বিলিয়ন	গুগল ডাটাবেস	https://bard.google.com/

3.2. বড় ভাষার মডেলের ধরন

এলএলএম-এর কিছু ধরন এবং তাদের মূল বৈশিষ্ট্য ও সম্ভাবনার একটি শনাক্তকরণ নিচে দেওয়া হল:

1. অটোরিগ্রেসিভ মডেল:

- জিপিটি (জেনারেটিভ প্রাক-প্রশিক্ষিত ট্রান্সফরমার): OpenAI দ্বারা বিকশিত, GPT হল একটি অটোরিগ্রেসিভ মডেল যা একটি শব্দে শব্দের ভিত্তিতে পাঠ্য তৈরি করে। এটির বেশ কয়েকটি সংস্করণ রয়েছে, 2021 সালে সর্বশেষ আপডেটের সময় GPT-3 সবচেয়ে সাম্প্রতিক এবং উন্নত ছিল।

2. দ্বিমুখী মডেল শ্রেণীবিভাগ:

- **BERT** (ট্রান্সফরমার থেকে দ্বিমুখী এনকোডার প্রতিনিধিত্ব): Google দ্বারা বিকশিত, BERT হল একটি মডেল যা দ্বিমুখীভাবে প্রশিক্ষিত হয়, যার অর্থ এটি একটি বাক্যে একটি শব্দের বাম এবং ডান উভয় দিকের প্রসঙ্গ বিবেচনা করে। এটি বিশেষভাবে বোধগম্যতা এবং পাঠ্য শ্রেণীবিভাগের কাজ পড়ার জন্য দরকারী।

3. সিকোয়েন্স থেকে সিকোয়েন্স মডেল:

- **T5** (টেক্সট-টু-টেক্সট ট্রান্সফরমার): Google দ্বারা বিকশিত, T5 সমস্ত ভাষা প্রক্রিয়াকরণ কাজকে পাঠ্য থেকে পাঠ্য রূপান্তর সমস্যা হিসাবে ব্যাখ্যা করে। উদাহরণস্বরূপ, "অনুবাদ", "সারাংশ" এবং "প্রশ্নের উত্তর" টেক্সট ইনপুট থেকে টেক্সট আউটপুটে রূপান্তর হিসাবে পরিচালনা করা হয়।
- **BART** (দ্বিমুখী এবং অটো-রিগ্রেসিভ ট্রান্সফরমার): Facebook AI দ্বারা তৈরি, BART প্রজন্ম এবং বোঝার কাজগুলির জন্য BERT এবং GPT-এর বৈশিষ্ট্যগুলিকে একত্রিত করে।

4. মাল্টিমোডাল মডেল:

- **CLIP** (বিপরীত ভাষা-ইমেজ প্রাক-প্রশিক্ষণ) এবং ডাল·ই: উভয়ই Op-nAI দ্বারা বিকশিত, এই মডেলগুলি কম্পিউটার দৃষ্টি এবং প্রাকৃতিক ভাষা প্রক্রিয়াকরণকে একত্রিত করে। যদিও CLIP প্রাকৃতিক ভাষার প্রেক্ষাপটে ছবি বুঝতে সক্ষম, DALL-E পাঠ্য বর্ণনা থেকে ছবি তৈরি করে।
- **WU DAO** বেইজিং একাডেমি অফ আর্টিফিশিয়াল ইন্টেলিজেন্স দ্বারা তৈরি করা একটি গভীর শিক্ষার ভাষা মডেল যাতে মাল্টিমোডালিটি বৈশিষ্ট্য রয়েছে। এটি পাঠ্য এবং চিত্র উভয় ডেটাতৈই প্রশিক্ষিত হয়েছে, তাই এটি উভয় কাজই মোকাবেলা করতে পারে। এটি অনেক পরামিতি (1.75 ট্রিলিয়ন) সহ প্রশিক্ষিত ছিল।

4. জেনারেটিভ এআই এর ক্ষেত্রে প্রাসঙ্গিক অ্যালগরিদমিকস

জেনারেটিভ আর্টিফিশিয়াল ইন্টেলিজেন্স মূলত তত্ত্বাবধানহীন শেখার কৌশলের উপর ভিত্তি করে। এটি তত্ত্বাবধানে থাকা শেখার মডেলগুলির থেকে আলাদা যেগুলির প্রশিক্ষণের পর্বটি সাজানোর জন্য লেবেলযুক্ত ডেটা প্রয়োজন। জেনারেটিভ অ্যাডভার্সারিয়াল নেটওয়ার্ক (GAN) বা ভেরিয়েশনাল অটোএনকোডার (VAEs) এর মতো তত্ত্বাবধানহীন লার্নিং মডেলগুলিতে এই ধরনের লেবেলিং সীমাবদ্ধতার অনুপস্থিতি বৃহত্তর এবং আরও ভিন্ন ভিন্ন ডেটাসেট ব্যবহারের অনুমতি দেয়, যার ফলে সিমুলেশনগুলি ঘনিষ্ঠভাবে বাস্তব-বিশ্বের পরিস্থিতির অনুকরণ করে (গুডফেলো এট আল।, 2016)। এই জেনারেটিভ মডেলগুলির প্রধান লক্ষ্য হল অন্তর্নিহিত সম্ভাব্যতা বন্টন $P(x)$ বোঝানো যা ডেটাসেট মেনে চলে। একবার মডেলটিকে দক্ষতার সাথে প্রশিক্ষিত করা হলে, এটি মূল ডেটাসেটের সাথে পরিসংখ্যানগতভাবে সামঞ্জস্যপূর্ণ ডেটা 'x'-এর নতুন নমুনা তৈরি করার ক্ষমতা রাখে। এই সংশ্লিষ্ট নমুনাগুলি শেখা বিতরণ থেকে আঁকা হয়, এইভাবে স্বাস্থ্যসেবা, অর্থ এবং সৃজনশীল শিল্পের মতো বিভিন্ন ক্ষেত্রে জেনারেটিভ মডেলগুলির প্রয়োজ্যতা প্রসারিত করে (বাইদু-আনু এবং ওউসু আনসাহ, 2023)।

জেনারেটিভ এআই-এর ল্যান্ডস্কেপ দুটি মূল আর্কিটেকচার দ্বারা প্রভাবিত: জেনারেটিভ অ্যাডভার্সারিয়াল নেটওয়ার্ক (GAN) এবং জেনারেটিভ প্রাক-প্রশিক্ষিত ট্রান্সফরমার (GPTs)। GAN দ্বৈত নিউরাল নেটওয়ার্কের মাধ্যমে কাজ করে, যার মধ্যে একটি জেনারেটর এবং একটি বৈষম্য রয়েছে। জেনারেটর সিন্থেটিক ডেটা তৈরি করে, যখন বৈষম্যকারী এই ডেটার সত্যতা মূল্যায়ন করে। এই প্রতিকূল প্রক্রিয়াটি পুনরাবৃত্তিমূলকভাবে চলতে থাকে যতক্ষণ না বৈষম্যকারী আর বাস্তব এবং সিন্থেটিক সম্পদের মধ্যে পার্থক্য করতে না পারে, এইভাবে উত্পন্ন বিষয়বস্তুকে বৈধ করে (Hu, 2022; Jovanovic, 2022)। GANs প্রধানত গ্রাফিক্স, স্পিচ জেনারেশন এবং ভিডিও সংশ্লেষণ (Hu, 2022) অ্যাপ্লিকেশনের জন্য ব্যবহৃত হয়।

বিভিন্ন আর্কিটেকচার যেমন GAN, GPT মডেল এবং বিশেষ করে ভ্যারিয়েশনাল অটোএনকোডার (VAE) থেকে বহুমুখী অবদান রয়েছে। পরেরটি শুধুমাত্র জেনারেটিভ মডেলিংয়ের একটি সম্ভাব্য দৃষ্টিভঙ্গিই দেয় না, তবে অন্তর্নিহিত জটিল ডেটার আরও নমনীয় বোঝার অনুমতি দেয়।

বিতরণ (কিংমা এবং ওয়েলিং, 2013)। উপরন্তু, মাল্টিমোডাল সিস্টেমের আবির্ভাব, যা একক স্থাপত্যে বিভিন্ন ধরনের ডেটার সমন্বয় ঘটায়, জটিল প্যাটার্ন শনাক্তকরণ এবং ডেটা সংশ্লেষণের ক্ষমতাকে পুনরায় সংজ্ঞায়িত করেছে। এই বিবর্তন ক্রমবর্ধমান জটিলতা এবং সূক্ষ্মতাকে প্রতিফলিত করে যা জেনারেটিভ এআই ক্যাপচার করতে পারে।

VAEs এবং মাল্টিমোডাল সিস্টেমের মধ্যে মিথস্ক্রিয়া জেনারেটিভ AI এর পরবর্তী সীমান্তের উদাহরণ দেয়। এটি শুধুমাত্র বৃহত্তর নির্ভুলতার প্রতিশ্রুতি দেয় না, বরং বিভিন্ন ধরনের ডেটার মধ্যে পার্থক্য সম্পর্কে সচেতন এবং প্রেক্ষাপটে সমৃদ্ধ ফলাফল তৈরি করার ক্ষমতাও দেয়। এই প্রেক্ষাপটে, জেনারেটিভ এআই একটি নিচ্ছক ডেটা-জেনারেটিং টুল থেকে একটি ক্রমবর্ধমান আন্তঃবিভাগীয় প্ল্যাটফর্মে বিকশিত হয়েছে যা বিভিন্ন শিল্পের সূক্ষ্মতা বুঝতে এবং জটিল সমস্যাগুলি সমাধান করতে সক্ষম (জোরান, 2021)।

4.1. স্টোকাস্টিক সুপ্ত অভিনেতা-সমালোচক: একটি সুপ্ত পরিবর্তনশীল মডেলের সাথে গভীর শক্তিবৃদ্ধি শিক্ষা

ডিপ রিইনফোর্সমেন্ট লার্নিং (RL) অ্যালগরিদমগুলি সরাসরি ইমেজ পর্যবেক্ষণ থেকে শিখতে উচ্চ-ক্ষমতার নেটওয়ার্কগুলিকে লিভারজ করে। যাইহোক, এই উচ্চ-মাত্রিক পর্যবেক্ষণ স্থানগুলি চ্যালেঞ্জগুলি উপস্থাপন করে কারণ নীতিকে এখন দুটি সমস্যা সমাধান করতে হবে: প্রতিনিধিত্ব শেখার এবং টাস্ক শেখার।

এই কাগজটি স্টোকাস্টিক ল্যাটেন্ট অ্যাক্টর-ক্রিটিক (SLAC) অ্যালগরিদম প্রস্তাব করে, উচ্চ-মাত্রিক ইমেজ ইনপুট থেকে সরাসরি জটিল ক্রমাগত নিয়ন্ত্রণের কাজগুলির জন্য নীতি শেখার জন্য একটি উচ্চ-কর্মক্ষমতা, উচ্চ-দক্ষতা RL পদ্ধতি। SLAC বিশেষভাবে একটি স্টোকাস্টিক সুপ্ত অবস্থার প্রবর্তনের মাধ্যমে বৃহৎ পর্যবেক্ষণের স্থানগুলির দ্বারা সৃষ্ট চ্যালেঞ্জগুলিকে মোকাবেলা করে যা সিদ্ধান্ত গ্রহণের জন্য প্রতিটি পরিস্থিতি থেকে প্রাসঙ্গিক তথ্যের সংক্ষিপ্তসার করে।

এইভাবে, SLAC অভিনেতা-সমালোচক রাষ্ট্রের প্রতিনিধিত্ব শেখার এবং সবচেয়ে উপযুক্ত পদক্ষেপ নির্ধারণের সাধারণ সমস্যাটিকে দুটি বিশেষ প্রক্রিয়ায় ভাগ করেন। একদিকে, এনকোডার মডিউলটি কাজের জন্য উপযোগী একটি সুপ্ত উপস্থাপনায় পর্যবেক্ষণ তথ্য সংকুচিত করার জন্য দায়ী। অন্যদিকে, প্রথাগত অভিনেতা-সমালোচক এই সুপ্ত অবস্থা ব্যবহার করে পরিস্থিতির মূল্যায়ন করতে এবং সেরা ক্রিয়া নির্বাচন করতে, জটিল মূল চিত্র ইনপুটকে সরাসরি প্রক্রিয়া করার প্রয়োজন ছাড়াই। SLAC একটি একক ইউনিফাইড ফ্রেমওয়ার্কে RL-এর সাথে স্টোকাস্টিক ক্রমিক মডেলগুলিকে একীভূত করার জন্য একটি উদ্ভাবনী এবং শক্তিশালী কৌশল প্রস্তাব করেছে। এটি একটি সংক্ষিপ্ত সুপ্ত উপস্থাপনা তৈরি করে সম্পন্ন করা হয়, যা পরবর্তীতে মডেল দ্বারা উত্পন্ন সুপ্ত স্থানের মধ্যে RL পরিচালনা করতে ব্যবহৃত হয়। পরীক্ষামূলক পরীক্ষাগুলি ইঙ্গিত দেয় যে এই পদ্ধতিটি চূড়ান্ত ফলাফলের পরিপ্রেক্ষিতে মডেল-মুক্ত এবং মডেল-ভিত্তিক উভয় প্রতিযোগীদের পারফরম্যান্সকে ছাড়িয়ে গেছে এবং বিশেষ করে একটি বর্ণালী জুড়ে নমুনার দক্ষ ব্যবহারের ক্ষেত্রে

জটিল চিত্র-ভিত্তিক নিয়ন্ত্রণ কাজ।

RL-এর জন্য জেনারেটিভ এনভায়রনমেন্ট মডেলের প্রেক্ষাপটে, এই ফলাফলগুলি সুপ্ত প্রতিনিধিত্ব শিক্ষার গুরুত্বের উপর জোর দেয় যা ইমেজ থেকে শক্তিবৃদ্ধি শিক্ষাকে স্বরাশ্রিত করতে পারে, যা মেশিন লার্নিং এবং জেনারেটিভ কৃত্রিম বুদ্ধিমত্তার ক্ষেত্রে উল্লেখযোগ্য অগ্রগতির প্রতিনিধিত্ব করে।

প্রথাগত ডিপ রিইনফোর্সমেন্ট লার্নিং (ডিআরএল) দৃষ্টান্তে, যেমন কিউ-লার্নিং বা পলিসি গ্রেডিয়েন্ট, উদ্দেশ্যমূলক ফাংশনটি প্রায়শই প্রত্যাশিত কর্মক্ষমতা সর্বাধিক করার জন্য অভিযোজিত হয়। জে(i) হিসাবে সংজ্ঞায়িত:

$$J(\theta) = E_{\tau \sim \pi_{\theta}} \left[\sum_{t=0}^{\infty} \gamma^t R(z_t, a_t) \right]$$

যেখানে τ একটি ট্রাজেক্টোরি প্রতিনিধিত্ব করে এবং γ হল ডিসকাউন্ট ফ্যাক্টর।

একটি সুপ্ত পরিবর্তনশীল মডেলের সাথে DRL-এর প্রসঙ্গে, একটি সুপ্ত স্থান z -এর অন্তর্ভুক্তি বিমূর্ততার একটি অতিরিক্ত স্তর যুক্ত করে। বিশেষত, রাষ্ট্র s একটি সুপ্ত পরিবর্তনশীল z দ্বারা প্রতিস্থাপিত বা পরিবর্ধিত হয়, যা পরিণতিতে রাষ্ট্রের একটি ফাংশন হতে পারে $z = f(s)$ অথবা তত্ত্বাবধানহীন ডেটা থেকে শেখা যেতে পারে।

$$J(\theta, \phi) = E_{z, z \sim \pi_{\theta, \phi}} \left[\sum_{t=0}^{\infty} \gamma^t R(z_t, a_t) \right]$$

যেখানে ϕ সুপ্ত পরিবর্তনশীল মডেলের পরামিতি এবং নীতি π এখন শুধুমাত্র শর্তযুক্ত নয় এস কিন্তু z -এও।

এই পদ্ধতির সাফল্য শেখা সুপ্ত স্থানের গুণমান এবং সর্বোত্তম সিদ্ধান্ত নেওয়ার জন্য প্রয়োজনীয় সূক্ষ্মতাগুলিকে কতটা ভালভাবে ক্যাপচার করে তার মধ্যে রয়েছে। এটি গবেষণার একটি প্রতিশ্রুতিশীল লাইন, এবং এখানকার অগ্রগতিগুলি ডিআরএল-এ জটিল, উচ্চ-মাত্রিক, আংশিকভাবে পর্যবেক্ষণযোগ্য সিদ্ধান্ত নেওয়ার সমস্যাগুলির কাছে যাওয়ার পদ্ধতিতে সম্ভাব্য বিপ্লব ঘটাতে পারে।

4.2. ভিডিও: ডিফিউশন মডেল সহ হাই-ডেফিনিশন ভিডিও জেনারেশন

স্থানিক এবং টেম্পোরাল সুপার-রেজোলিউশন ভিডিও মডেলগুলির অনুক্রমিক রচনাটি একটি সহজাত স্বাভাব্য প্রদর্শন করে, কারণ এটি শুধুমাত্র পিক্সেল-স্তরের বিশ্বস্ততাকে উন্নত করে না, তবে উত্পন্ন ফ্রেমের মধ্যে সাময়িক সংগতিও নিশ্চিত করে (Simonyan & Zisserman, Xie40; ., 2018)। এই দিকটি এমন অ্যান্সিকেশনগুলির জন্য বিশেষভাবে উপকারী যেগুলির জন্য গতিশীল দৃশ্য রেন্ডারিং এবং তরল গতি রেন্ডারিং প্রয়োজন; প্রচলিত চিত্র-ভিত্তিক জেনারেটিভ মডেলগুলি সেই শর্তগুলি পূরণ করে না (জিয়াং এট আলা., 2018)। তদ্ব্যতীত, সিস্টেমের পার্থক্য-হাই-ডেফিনিশন ভিডিও আউটপুটের দিকে স্কেল করার ক্ষমতা, এটির সম্পূর্ণ রূপান্তরমূলক আর্কিটেকচারের দ্বারা সম্ভব হয়েছে, স্ট্যাটিক ইমেজ জেনারেশনের ক্ষেত্রে একটি উল্লেখযোগ্য অগ্রগতির প্রতিনিধিত্ব করে। প্রথাগত জেনারেটিভ মডেলগুলিতে, ভিডিও তৈরির জন্য কম্পিউটেশনাল জটিলতা প্রায়শই খারাপভাবে স্কেল করে, হাই-ডেফিনিশন ফলাফলগুলিকে গণনাগতভাবে অসম্ভব করে তোলে (ভনড্রিক এট আলা., 2016)। বিপরীতে, মডুলার ইমেজ-ভিডিও ক্যাসকেড কাঠামো আরও দক্ষ সম্পদ বরাদ্দের সুবিধা দেয়, যা যথেষ্ট কম সময়ের মধ্যে উচ্চ-মানের ভিডিও স্যাম্পলিং সম্ভব করে (Tulyakov et al., 2018)।

বিদ্যমান ইমেজ-ভিডিও আর্কিটেকচারে ডিফিউশন মডেলগুলির একীকরণ প্রযুক্তিগুলির একটি বাধ্যতামূলক ফিউশন তৈরি করে যা এর উত্পাদনশীল ক্ষমতাগুলিকে বাড়ানোর জন্য পরিবেশন করতে পারে। 0 থেকে 1 এর পরিসরের সাথে অস্থায়ী উপাদান প্রবর্তন করে, ডিফিউশন মডেলটি VAE-এর উচ্চ-স্তরের উপস্থাপনাকে সমৃদ্ধ করে জেনারেট করা ভিডিওগুলিতে একটি সাময়িকভাবে সামঞ্জস্যপূর্ণ স্তরের স্টোকাস্টিসিটি এবং গ্রানুলারিটি প্রদান করে (Kingma & Welling, 2013; Ho et al., 2020)।

এই হাইরিড মডেলটিকে ধারণাগতভাবে দুটি পর্যায়ে অপারেটিং হিসাবে কল্পনা করা যেতে পারে: প্রথমত, VAE এনকোডার(x) এনকোডার ফাংশন ইনপুট x থেকে একটি সুপ্ত পরিবর্তনশীল z গণনা করে। এই z পরবর্তীতে ডিফিউশন মডেলের প্রাথমিক শর্ত হিসেবে কাজ করে, মূলত এর সমীকরণে $\$x\$$ এর ভূমিকা পালন করে। পরবর্তী, একটি সময়-নির্ভর denoising ফাংশন ডি($সঙ্গে$), চালু করা হয়, যা উচ্চ-স্তরের প্রতিনিধিত্বকে পরিমার্জিত করে-সময়ের সাথে সাথে VAE এর টেনশন। এটি একটি সংশ্লেষণ ফাংশনে শেষ হয় $x^t(z, t) =$ ডিকোডার($সঙ্গে$) + ডি($সঙ্গে$), এইভাবে একটি সমৃদ্ধ, অস্থায়ীভাবে মসৃণ আউটপুট আর্কেস্ট্রিং।

ডিফিউশন মডেলের জন্য অবজেক্টিভ ফাংশন $L(x)$ তখন হয়ে যায়:

$$L(x) = \mathbb{E}_{z \sim p(z|x)} \left[\mathbb{E}_{t \sim p(t)} \left[\mathbb{E}_{x^t \sim p(x^t|z, t)} \left[\mathbb{E}_{\theta} \left[\left\| \hat{\theta}(x^t, \lambda_t) - \hat{\theta}^{(0,t)}(x^t, \lambda_t) \right\|_2^{(0,t)} \right] \right] \right] \right]$$

এইভাবে, ডিফিউশন মডেলটি সময়ের সাথে সাথে VAE দ্বারা উত্পন্ন উচ্চ-স্তরের উপস্থাপনাগুলিকে সামঞ্জস্য করতে শেখে, এইভাবে একটি সমৃদ্ধ, অস্থায়ীভাবে মসৃণ জেনারেটিভ মডেল প্রদান করে যা উভয় আর্কিটেকচারের শক্তির সুবিধা নেয়।

4.3. মোশন ডিফিউজ: ডিফিউশন মডেল সহ পার্থ্য-চালিত মানব মোশন জেনারেশন

টেব্রট-ভিত্তিক মানব মোশন জেনারেশনের জন্য MotionDiffuse আর্কিটেকচারে, মডেলে এমবেড করা মনোযোগের কাঠামোর সাথে সামঞ্জস্যপূর্ণ ক্ষতির ফাংশনগুলির ডিজাইনে বিশেষ মনোযোগ দেওয়া উচিত। মোশন ডিফিউজ টেব্রট এবং মোশন স্পেসের মধ্যে সম্পর্ক মডিউল করার জন্য মনোযোগের প্রক্রিয়া নিখুঁত করে, দুটি বিশিষ্ট ক্ষতি ফাংশন সমালোচনামূলক গুরুত্ব হতে পারে।

GLIDE মডেলের মতো বিশেষ ক্ষতির ফাংশনগুলিকে অন্তর্ভুক্ত করা MotionDiffuse ফ্রেমওয়ার্ককে আরও পরিমার্জিত করতে পারে, বিশেষ করে পার্থ্য-ভিত্তিক মানব গতি তৈরিতে এর অনন্য পদ্ধতির সাথে। মোশনডিফিউজে, সমীকরণ দ্বারা পরিচালিত একটি স্টোকাস্টিক শব্দ শব্দের অন্তর্ভুক্তি $\alpha_t = 1 - \beta_t$, $\bar{\alpha}_t = \sum_{s=0}^t \alpha_s$ ক্ষতি ফাংশন উদ্ভাবনের জন্য একটি আকর্ষণীয় উপায় প্রস্তাব।

মূল MotionDiffuse পদ্ধতিতে, মডেলটি পরিবর্তে একটি শব্দ শব্দের পূর্বাভাস দেওয়ার উপর দৃষ্টি নিবদ্ধ করে এর x_{t-1} GLIDE ফ্রেমওয়ার্কের সাথে সঙ্গতিপূর্ণ (Nichol et al., 2021)। এটি একটি গড় বর্গক্ষেত্র ত্রুটি (MSE) ক্ষতি ফাংশন দ্বারা ক্যাপচার করা হয়েছে:

$$L_{\text{MSE}} = \mathbb{E}_{x_0, \epsilon} \left[\sum_{t=1}^T \left\| \epsilon - \theta(x_t, t, \text{পার্থ্য}) \right\|^2 \right]$$

নিম্নলিখিত বিশেষ ক্ষতি ফাংশন একীকরণ বিবেচনা করুন:

পদ্ধতির মধ্যে মনোযোগী প্রান্তিককরণের ক্ষতি

এই ক্ষতি ফাংশন নিশ্চিত করে যে পার্থ্য এবং গতি স্থানগুলির মধ্যে মনোযোগ বিতরণ সারিবদ্ধ করা হয়েছে। যেহেতু মনোযোগ স্কেল k এবং k_{max} মিটারে যথাক্রমে টেব্রট এবং মোশন এনকোডার থেকে প্রাপ্ত করা হয়েছে, ক্ষতিটিকে দুটির মধ্যে কুলব্যাক-লিয়ার ডাইভারজেন্স হিসাবে সংজ্ঞায়িত করা যেতে পারে:

$$L_{\text{attention}} = KL(a_t \parallel a_m)$$

এই ক্ষতি মডেলটিকে পার্থ্য এবং গতি ডোমেন উভয় ক্ষেত্রেই শব্দার্থগতভাবে অনুরূপ অঞ্চলগুলিতে মনোযোগ দিতে উত্সাহিত করে, এইভাবে পদ্ধতিগুলির মধ্যে আরও ভাল প্রান্তিককরণকে উত্সাহিত করে।

টেব্রট-গাইডেড মোশন ফিডেলিটি লস

উত্পন্ন গতি শুধুমাত্র যে কোনো গতি নয় কিন্তু বিশেষভাবে টেব্রট বর্ণনার সাথে সারিবদ্ধ হয় তা নিশ্চিত করার জন্য, একটি পার্থ্য-নির্দেশিত গতির বিশ্বস্ততা হ্রাস প্রবর্তন করা যেতে পারে। এটি উৎপন্ন গতি থেকে আহরিত উচ্চ-স্তরের বৈশিষ্ট্য এবং পার্থ্য দ্বারা নির্দেশিতগুলির মধ্যে সামঞ্জস্যতা পরিমাপ করবে। যাক c_t এবং c_m উচ্চ-স্তরের পার্থ্য এবং গতি বৈশিষ্ট্য হতে হবে:

$$L_{\text{fidelity}} = \|F_t - F_m\|_1$$

এই ক্ষতির ফাংশনগুলিকে ঐতিহ্যগত জেনারেটিভ লস যেমন গড় স্কয়ার এরর (MSE) বা জেনারেটিভ অ্যাডভার্স লসের সাথে মিলিত করা নিশ্চিত করে যে মোশন ডিফিউজ এমন গতি তৈরি করে যা শুধুমাত্র গুণগতভাবে ভাল নয়, পার্থ্য বর্ণনার সাথেও ভালভাবে সারিবদ্ধ। এটি বৈচিত্র্যময় এবং পার্থ্যগতভাবে সামঞ্জস্যপূর্ণ গতি তৈরিতে কিছু চ্যালেঞ্জ মোকাবেলা করবে।

মডেলের মধ্যে মনোযোগী প্রক্রিয়াগুলির সাথে এই ক্ষতির ফাংশনগুলিকে জটিলভাবে লিঙ্ক করে, মোশনডিফিউজ পার্থ্য বিবরণ দ্বারা পরিচালিত আরও সংক্ষিপ্ত, সাময়িকভাবে সুসংগত, এবং প্রাসঙ্গিকভাবে প্রাসঙ্গিক গতি ক্রমগুলিকে সহজতর করতে পারে।

এই বিশেষ লস ফাংশনগুলিকে একটি ইউনিফাইড লস টার্মে মার্জ করে, আমরা লিখতে সক্ষম হই:

$$L_{total} = \lambda_1 L_{MSE} + \lambda_2 L_{attention} + \lambda_3 L_{fidelity}$$

এখানে, l_1 , λ_2 , λ_3 হাইপারপ্যারামিটার যা বিভিন্ন ক্ষতির উপাদানগুলির মধ্যে ভারসাম্য নিয়ন্ত্রণ করে।

4.4. টেক্সট, অডিও এবং ভিডিও জেনারেশনে ডিফিউশন মডেল এবং সুপ্ত স্পেস ডায়নামিক্সের উত্থান

এটি সম্ভাব্য ডিনোইসিং ডিফিউশন ডিনোইসিং মডেলের (ডিডিপিএম) ক্ষেত্রে সাম্প্রতিক উদ্ভাবন। প্রচলিত ডিডিপিএমের বিপরীতে, এলডিএমগুলি সুপ্ত স্থানে কাজ করে এবং পার্থ্য উপস্থাপনাগুলির উপর শর্তসাপেক্ষ (Rombach et al., 2022)। এটি একটি ধাপে ধাপে দানাদার প্রজন্মের প্রক্রিয়ার জন্য অনুমতি দেয় যার মধ্যে বেশ কয়েকটি শর্তসাপেক্ষ প্রসারণ পদক্ষেপ রয়েছে।

লস ফাংশনটি মনোযোগ সহকারে নয়েজ স্পেসের গড় বর্গক্ষেত্র ত্রুটি হিসাবে সংজ্ঞায়িত করা হয় ($\xi \sim N(0, I)$)। গাণিতিকভাবে, এটি এইভাবে তৈরি করা যেতে পারে:

$$L_{\theta} = \|\xi_{\theta}(z_t, t, c) - \xi\|_{2,2},$$

যেখানে c একটি ছোট ইতিবাচক ধ্রুবক এবং x_t ডিনোইসিং নেটওয়ার্কের প্রতিনিধিত্ব করে। এই লস ফাংশনটি স্টোকাস্টিক গ্রেডিয়েন্ট ডিসেন্ট সহ একটি র্যান্ডম টার্ম টি অস্টিমাইজ করে দক্ষ প্রশিক্ষণের অনুমতি দেয়। এটি লক্ষণীয় যে এই পদ্ধতিটি বিপক্ষ প্রতিক্রিয়ার প্রয়োজন ছাড়াই প্রমাণ লোয়ার বাউন্ড (ELBO) অস্টিমাইজ করে ডিফিউশন মডেলকে দক্ষতার সাথে প্রশিক্ষিত করার অনুমতি দেয়, যার ফলস্বরূপ উল্লেখযোগ্যভাবে বিশ্বস্ত পুনর্গঠন যা স্থল সত্য বিতরণের সাথে মেলে (হুয়াং এট আল., 2023)।

2023 সালে, জেনারেটিভ আর্টিফিশিয়াল ইন্টেলিজেন্স ল্যান্ডস্কেপ একটি সিসমিক শিফটের মধ্য দিয়ে যাচ্ছে, যা মূলত ডিফিউশন মডেলের পরিপক্বতার দ্বারা চিহ্নিত। টেক্সট ডোমেনে, যখন ট্রান্সফরমারগুলি একসময় প্রাকৃতিক ভাষা বোঝার এবং প্রজন্মের বহুমুখীতার জন্য প্রশংসিত হয়েছিল, তখন ডিফিউশন মডেলগুলি এখন উদ্ভাবনের একটি নতুন তরঙ্গকে অনুঘটক করছে, আরও শক্তিশালী সাধারণীকরণ ক্ষমতা সহ সংক্ষিপ্ত ভাষা মডেলগুলি অফার করছে। তারা আগের মডেলগুলির মৌলিক কাজকে প্রসারিত করছে, অনুভূতি বিশ্লেষণ, বিমূর্ত সংক্ষিপ্তসার এবং আরও অনেক কিছুতে জটিলতা এবং প্রয়োগযোগ্যতার স্তর যুক্ত করছে (নিকোল এট আল., 2021)। অডিও সেক্টরে, মেক-আন-অডিওর মতো মডেলগুলি উন্নত সম্প্রচার কৌশল ব্যবহার করে টেক্সট-টু-অডিও (T2A) জেনারেশনে বিপ্লব ঘটচ্ছে (Zhao et al., 2023)। তারা দীর্ঘ, অবিচ্ছিন্ন সিগন্যাল ডেটার অন্তর্নিহিত জটিলতাকে সম্বোধন করে, যা পূর্ববর্তী পদ্ধতি যেমন WaveGAN এবং MelGAN এর সাথে লড়াই করেছিল। এর ফলে উচ্চতর বিশ্বস্ততা অডিও জেনারেশন এবং সূক্ষ্ম শব্দার্থিক বোঝাপড়া হয়, ক

পূর্ববর্তী পদ্ধতি থেকে আমূল পরিবর্তন।

ভিডিও জেনারেশন হাই-ডেফিনিশন রডকাস্ট মডেলগুলির দ্বারাও শক্তিশালী হয়েছে, রেজোলিউশনের সীমাবদ্ধতার বাইরে চলে গেছে, যা প্রায়শই পুরানো GAN-ভিত্তিক পদ্ধতিগুলিকে জর্জরিত করে। এখানে, ফোকাস শুধুমাত্র পিক্সেল-স্তরের বিশদটির উপর নয়, ভিডিও ফ্রেমের মধ্যে শব্দার্থিক সমন্বয়ের উপরও, এইভাবে উত্পন্ন সামগ্রীর বাস্তববাদে একটি নতুন স্ট্যান্ডার্ড অর্জন করে (Rombach et al., 2022)।

ক্রস-মডেল জেনারেটিভ লার্নিংয়ে, ডিফিউশন মডেলের সাথে একীভূত স্টোকাস্টিক প্রচ্ছন্ন অভিনেতা-সমালোচক মডেলের আবির্ভাব বিভিন্ন ডেটা প্রকারের মধ্যে বিরামহীন এবং আরও সঠিক অনুবাদের সম্ভাবনা প্রদান করে। এই হাইব্রিড মডেলগুলি টেক্সট, অডিও এবং ভিডিও পদ্ধতিতে হাই-ডেফিনিশন, হাই-ফাই-ডেলিটি জেনারেশনের ক্ষমতা আনলক করতে শুরু করেছে (হারনোজা এট আল., 2018)।

একত্রে নেওয়া, পার্থ্য, অডিও এবং ভিডিও সম্প্রচার মডেলের এই অগ্রগতিগুলি মাল্টিমডাল জেনারেটিভ এআই-এর ভবিষ্যতের জন্য মঞ্চ তৈরি করেছে যা এর অ্যাপ্লিকেশনগুলিতে আরও সমৃদ্ধ, আরও গতিশীল এবং আরও সুদূরপ্রসারী, এইভাবে ডেটার বিস্তৃত ল্যান্ডস্কেপে বিপ্লব ঘটচ্ছে। বিশ্লেষণ, বিষয়বস্তু তৈরি এবং স্বয়ংক্রিয় সিদ্ধান্ত গ্রহণ।

5. চ্যাট জিপিটি, এর সম্ভাব্যতা এবং কীভাবে এটির সুবিধা নেওয়া যায়

5.1. GPT এর বর্ণনা

জেনারেটিভ প্রাক-প্রশিক্ষিত ট্রান্সফরমার (GPT) মডেলগুলি প্রাকৃতিক ভাষা প্রক্রিয়াকরণ (NLP) এবং কৃত্রিম বুদ্ধিমত্তা (AI) ক্ষেত্রে একটি উল্লেখযোগ্য উদ্ভাবনের প্রতিনিধিত্ব করে। OpenAI দ্বারা বিকশিত, GPT মডেলগুলি ট্রান্সফরমার আর্কিটেকচারের উপর ভিত্তি করে, যা ভাসওয়ানি এট আল দ্বারা প্রবর্তিত হয়েছিল। (2017)।

GPT মডেলগুলিকে টেক্সটের বড় কর্পোরার উপর প্রাক-প্রশিক্ষিত করা হয় এবং তারপরে নির্দিষ্ট কাজের উপর টিউন করা হয়, যা তাদের পাঠ্য তৈরি করতে দেয় যা সামঞ্জস্যপূর্ণ, ব্যাকরণগতভাবে সঠিক এবং প্রায়শই মানব-উত্পাদিত পাঠ্য থেকে আলাদা করা যায় না (Radford et al., 2019)।

GPT মডেলগুলি মনোযোগের একাধিক স্তর এবং মাল্টি-হেড অ্যাটেনশনের সংমিশ্রণ ব্যবহার করে, যা তাদের ডেটাতে বিভিন্ন বৈশিষ্ট্য এবং সম্পর্ক ক্যাপচার করতে দেয়। জিপিটি মডেলের পাঠ্য বোঝার এবং তৈরি করার ক্ষমতা মেশিন অনুবাদ, সৃজনশীল পাঠ্য তৈরি এবং প্রশ্নের উত্তর সহ বিভিন্ন অ্যাপ্লিকেশনে অগ্রগতির দিকে পরিচালিত করেছে (ব্রাউন এট আল., 2020)।

5.1.1. স্থাপত্য এবং মনোযোগ

ট্রান্সফরমার আর্কিটেকচার, ভাসওয়ানি এট আল দ্বারা প্রবর্তিত। (2017), প্রাকৃতিক ভাষা প্রক্রিয়াকরণের ক্ষেত্রে বৈপ্লবিক পরিবর্তন এনেছে এবং সিকোয়েন্সগুলিকে মডেল করার পদ্ধতিতে একটি দৃষ্টান্ত পরিবর্তনের প্রতিনিধিত্ব করে। প্রথাগত পুনরাবৃত্ত আর্কিটেকচারের বিপরীতে (Bahdanau, Cho, & Bengio, 2014), ট্রান্সফরমারগুলি পুনরাবৃত্তি দূর করে এবং পরিবর্তে ডেটাতে নির্ভরতা ক্যাপচার করতে মনোযোগের প্রক্রিয়া ব্যবহার করে।

আর্কিটেকচার দুটি প্রধান অংশ নিয়ে গঠিত: এনকোডার এবং ডিকোডার; উভয়ই দুটি প্রধান সাবলেয়ার সমন্বিত অভিন্ন স্তরগুলির একটি স্ট্যাক দ্বারা গঠিত: একটি মাল্টি-হেড অ্যাটেনশন সাবলেয়ার এবং একটি ফিড-ফরোয়ার্ড নেটওয়ার্ক। মাল্টি-হেড অ্যাটেনশন মডেলটিকে একই সাথে ইনপুটের বিভিন্ন অংশে উপস্থিত থাকতে দেয়, এইভাবে জটিল এবং দীর্ঘমেয়াদী সম্পর্ক ক্যাপচার করে। ফরোয়ার্ড ফিড-ফরোয়ার্ড নেটওয়ার্ক একটি নিয়ে গঠিত সরল রৈখিক রূপান্তর এবং একটি ননলিনিয়ার অ্যাক্টিভেশন ফাংশন অনুসরণ করে।

ট্রান্সফরমার

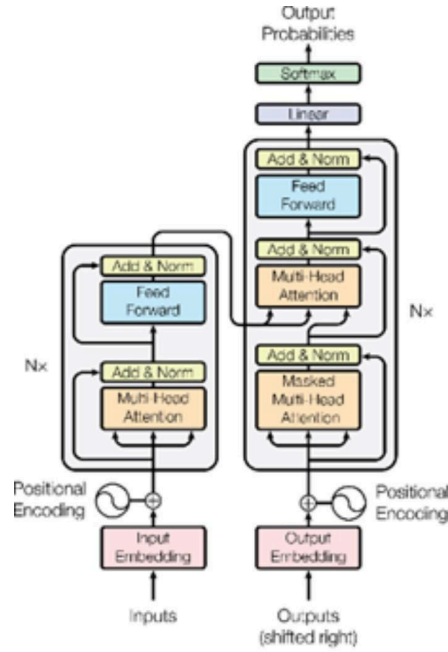
ট্রান্সফরমার আর্কিটেকচারের একটি মূল বৈশিষ্ট্য হল প্রতিটি সাবলেয়ারের চারপাশে অবশিষ্ট সংযোগ যোগ করা (He, Zhang, Ren, & Sun, 2016), তারপরে স্তর স্বাভাবিককরণ (Ba, Kiros, & Hinton, 2016)। এটি প্রশিক্ষণের সুবিধা দেয় এবং নেটওয়ার্কের মাধ্যমে গ্রেডিয়েন্টগুলিকে আরও সহজে প্রবাহিত করতে দেয়। মনোযোগ, অবশিষ্ট সংযোগ, এবং স্তর স্বাভাবিককরণের সমন্বয় ট্রান্সফরমারগুলিকে অত্যন্ত সমান্তরাল এবং গণনাগতভাবে দক্ষ হতে দেয়, যা বিভিন্ন এনএলপি কাজগুলিতে তাদের ব্যাপক গ্রহণের দিকে পরিচালিত করেছে (ডেভলিন, চ্যাং, লি, এবং টুটানোভা, 2018), দেখুন চিত্র 1।

ট্রান্সফরমার আর্কিটেকচারে মনোযোগ একটি কেন্দ্রীয় ধারণা। এটি মডেলটিকে আউটপুটের প্রতিটি শব্দ তৈরি করার সময় ইনপুটের বিভিন্ন অংশের ওজন করার অনুমতি দেয়, যা গভীর প্রাসঙ্গিক বোঝার সুবিধা দেয়। মনোযোগ তিনটি প্রধান উপাদানে বর্ণনা করা যেতে পারে:

ক্যোয়ারী, কী এবং মান: ক্যোয়ারী, কী এবং মান ব্যবহার করে মনোযোগ গণনা করা হয়, যা ইনপুটে থাকা শব্দগুলির ভেক্টর উপস্থাপনা। প্রশ্ন এবং কীগুলি মনোযোগের ওজন নির্ধারণ করে, যখন আউটপুট তৈরি করতে এই ওজন অনুসারে মানগুলিকে ওজন করা হয় (Vaswani et al. 2017)।

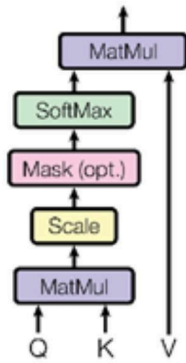
মাল্টি-হেড অ্যাটেনশন: মাল্টি-হেড অ্যাটেনশন মডেলটিকে একই সাথে ইনপুটের বিভিন্ন অংশে মনোযোগ দিতে দেয়। মনোযোগের প্রতিটি "মাথা" পাঠ্যের বিভিন্ন সম্পর্কের উপর ফোকাস করতে পারে, এবং সমস্ত মাথা থেকে আউটপুট চূড়ান্ত উপস্থাপনা তৈরি করতে একত্রিত হয়।

অটোরিগ্রেসিভ অ্যাটেনশন: জিপিটির ক্ষেত্রে, অটোরিগ্রেসিভ মনোযোগ ব্যবহার করা হয়, যেখানে প্রতিটি শব্দ শুধুমাত্র অনুক্রমের পূর্ববর্তী শব্দগুলিতে মনোযোগ দিতে পারে। এটি নিশ্চিত করে যে পাঠ্য প্রজন্ম একটি কার্যকারণ এবং সুসঙ্গত পদ্ধতিতে সম্পাদিত হয়েছে (Radford et al., 2018; Brown et al., 2020), চিত্র 2 দেখুন।

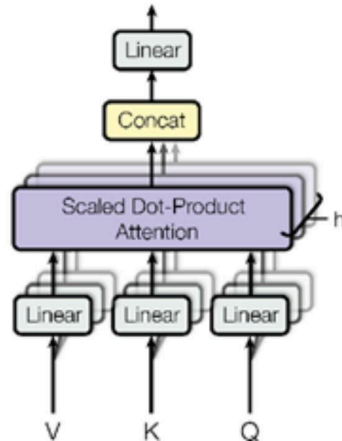


চিত্র 1. ট্রান্সফরমার - মডেল আর্কিটেকচার (ওপেনএআই, 2023)।

led Dot-Product Attention



Multi-Head Attention



চিত্র 2. (বাম) পণ্য পয়েন্ট স্কেল মনোযোগ. (ডানদিকে) মাল্টি-হেড মনোযোগ সমান্তরালভাবে চলমান মনোযোগের বিভিন্ন স্তর নিয়ে গঠিত (ওপেনএআই, 2023)।

স্তর এবং ফিড-ফরোয়ার্ড নেটওয়ার্ক

ট্রান্সফরমার আর্কিটেকচারে একাধিক স্তর রয়েছে, যার প্রতিটিতে মাল্টি-হেড অ্যাটেনশন লেয়ার রয়েছে যার পরে একটি ঘনভাবে সংযুক্ত ফিড-ফরোয়ার্ড নেটওয়ার্ক রয়েছে। প্রশিক্ষণের সুবিধার্থে এবং স্থিতিশীলতা উন্নত করার জন্য প্রতিটি পর্যায়ে অবশিষ্ট সংযোগ এবং স্তর স্বাভাবিককরণ প্রয়োগ করা হয় (He et al., 2016; Ba et al., 2016)।

5.1.2. প্রাক-প্রশিক্ষণ এবং ফাইন-টিউনিং

ওপেনএআই জিপিটি মডেলের প্রাক-প্রশিক্ষণ এবং সূক্ষ্ম-টিউনিং হল ট্রান্সফরমার-ভিত্তিক ভাষা মডেলগুলি বিকাশ ও অভিযোজিত করার প্রক্রিয়ার দুটি গুরুত্বপূর্ণ পর্যায়। বিশেষ করে, জিপিটি প্রাকৃতিক ভাষা প্রক্রিয়াকরণ গবেষণার একটি নেতৃত্বান্বিত দৃষ্টান্ত এবং সুসংগত এবং প্রাসঙ্গিকভাবে প্রাসঙ্গিক পাঠ্যের প্রজন্মের একটি উল্লেখযোগ্য অগ্রগতির প্রতিনিধিত্ব করে।

জিপিটি প্রাক-প্রশিক্ষণ

প্রাক-প্রশিক্ষণ হল একটি GPT মডেল তৈরির প্রাথমিক পর্যায় এবং এটি জ্ঞান স্থানান্তরের মূল ধারণার উপর প্রতিষ্ঠিত। এই পর্যায়ে, মডেলটিকে প্রচুর পরিমাণে ট্যাগবিহীন পাঠ্য ডেটার উপর প্রশিক্ষণ দেওয়া হয়। এখানেই GPT গভীর স্তরে ভাষাগত, শব্দার্থিক, এবং প্রাসঙ্গিক নিদর্শন শিখে। এই প্রক্রিয়াটি একটি অটোরিগ্রেশন টাস্ক ব্যবহার করে করা হয়, যেখানে মডেলটি পূর্ববর্তী শব্দের উপর ভিত্তি করে একটি বাক্যে পরবর্তী শব্দের পূর্বাভাস দেয় (Vaswani et al., 2017)। ট্রান্সফরমার-সদৃশ আর্কিটেকচারে উপস্থিত মনোযোগী প্রক্রিয়া মডেলটিকে পাঠ্যে দীর্ঘমেয়াদী সম্পর্ক ক্যাপচার করতে দেয়, যা এটিকে সুসংগত এবং প্রাসঙ্গিকভাবে প্রাসঙ্গিক প্রতিক্রিয়া তৈরি করতে সক্ষম করে।

GPT প্রাক-প্রশিক্ষণের মধ্যে রয়েছে অস্টিমাইজেশন অ্যালগরিদম যেমন স্টোকাস্টিক গ্রেডিয়েন্ট ডিসেন্ট ব্যবহার করে মডেলের লক্ষ লক্ষ পরামিতি অস্টিমাইজ করা। মূল হাইপারপ্যারামিটার, যেমন শেখার হার এবং ট্রান্সফরমার আর্কিটেকচার, এই প্রক্রিয়াতে একটি গুরুত্বপূর্ণ ভূমিকা পালন করে (ভাসওয়ানি এট আল, 2017)। সফল প্রাক-প্রশিক্ষণ অর্জনের জন্য এই হাইপারপ্যারামিটারগুলিকে যথাযথভাবে টিউন করার গুরুত্ব লক্ষণীয়।

GPT এর ফাইন-টিউনিং

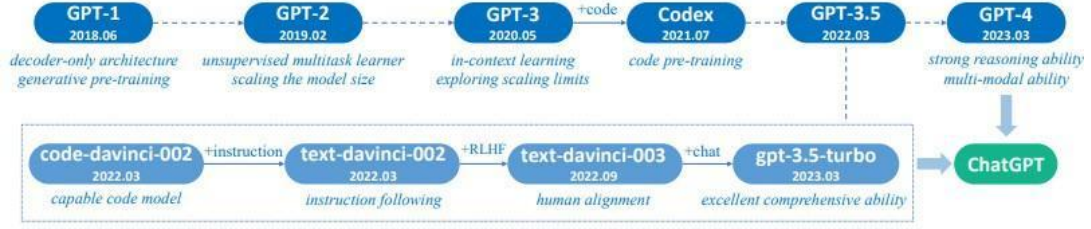
প্রাক-প্রশিক্ষণ পর্বের পরে, GPT মডেলটি গভীর ভাষা জ্ঞান দ্বারা সমৃদ্ধ হয়। যাইহোক, মডেলটিকে নির্দিষ্ট কাজের সাথে মানিয়ে নিতে, যেমন একটি নির্দিষ্ট ডোমেনে পাঠ্য তৈরি করা বা নির্দিষ্ট প্রশ্নের উত্তর দেওয়া, সূক্ষ্ম টিউনিং প্রয়োজন। এই পর্যায়ে, মডেলটিকে নির্দিষ্ট কাজের জন্য প্রাসঙ্গিক লেবেলযুক্ত ডেটার একটি সেটে প্রশিক্ষণ দেওয়া হয়।

টিউনিংয়ে টাস্ক-নির্দিষ্ট ডেটাসেট ব্যবহার করে প্রাক-প্রশিক্ষিত মডেলের ওজন সামঞ্জস্য করা জড়িত। টিউনিংয়ে প্রাক-প্রশিক্ষণের তুলনায় একটি ছোট শেখার হার জড়িত এবং সাধারণত কম পুনরাবৃত্তির প্রয়োজন হয় কারণ মডেল দ্বারা পূর্বে অর্জিত জ্ঞান আরও বিশেষায়িত।

টিউনিংয়ের কাছে যাওয়ার সময়, সঠিক ডেটাসেট নির্বাচন করা এবং একটি কার্যকর মূল্যায়ন কৌশল ডিজাইন করা গুরুত্বপূর্ণ। মূল্যায়ন মেট্রিক্স অবশ্যই টাস্কের উদ্দেশ্যগুলির সাথে সারিবদ্ধ হতে হবে। উদাহরণস্বরূপ, অনুভূতি বিশ্লেষণে, নির্ভুলতা, প্রত্যাহার এবং স্কেরের মতো মেট্রিকগুলি বিবেচনা করা যেতে পারে। ওভারফিটিং এড়াতে পর্যাপ্ত বৈধতা নিশ্চিত করা এবং পরীক্ষার সেটগুলিও অপরিহার্য (Radford et al., 2019)।

5.1.3. GPT এর বিবর্তন

নিচে ওপেনএআই দ্বারা তৈরি জিপিটি মডেলগুলির বিবর্তনগুলির একটি পর্যালোচনা, যা GPT-4 পর্যন্ত নেতৃত্ব দেয়। মডেলের জিপিটি সিরিজ কৃত্রিম বুদ্ধিমত্তার ক্ষেত্রে বিশেষ করে প্রাকৃতিক ভাষা প্রক্রিয়াকরণ (NLP) ক্ষেত্রে সবচেয়ে উল্লেখযোগ্য অগ্রগতির একটি প্রতিনিধিত্ব করে, চিত্র 3 দেখুন।



চিত্র 3. GPT সিরিজের মডেলগুলির প্রযুক্তিগত বিবর্তনের একটি সংক্ষিপ্ত চিত্র (OpenAI, 2023)।

জিপিটি-১

প্রকাশ: OpenAI দ্বারা জুন 2018।

স্থাপত্য: ভাসওয়ানি এট আল দ্বারা ট্রান্সফরমার। (2017)। পরামিতি:

117 মিলিয়ন।

প্রযুক্তিগত বিবরণ: 12টি স্তর, 12টি মনোযোগের মাথা এবং 768টি লুকানো ইউনিট সহ একটি ট্রান্সফরমার আর্কিটেকচার ব্যবহার করে।

প্রধান অবদান: প্রাকৃতিক ভাষা প্রক্রিয়াকরণ কাজের জন্য পূর্ব প্রশিক্ষণ সহ একটি ট্রান্সফরমার আর্কিটেকচারের প্রবর্তন (Radford et al. 2018)।

GPT-2

প্রকাশ: OpenAI দ্বারা ফেব্রুয়ারি 2019।

পরামিতি: 1.5 বিলিয়ন।

প্রযুক্তিগত বিবরণ: GPT-2 স্কেলে একটি সম্প্রসারণ বৈশিষ্ট্যযুক্ত, যা তার পূর্বসূরীর চেয়ে দশগুণ বড়। মডেলটি একটি বৈচিত্র্যময় এবং বৃহত্তর কর্পাসের উপর প্রশিক্ষিত ছিল। এর বৃহত্তম সংস্করণে 48টি স্তর, 1600টি লুকানো ইউনিট এবং 25.6 বিলিয়ন প্যারামিটার।

প্রধান অবদান: দেখিয়েছেন যে বড় মাপের ভাষা মডেল অনুচ্ছেদ জুড়ে সামঞ্জস্যপূর্ণ, উচ্চ-মানের পাঠ্য তৈরি করতে পারে, যা বিভিন্ন অ্যাপ্লিকেশন এবং মেশিন অনুবাদের জন্য দরকারী ছিল (Radford et al. 2019)।

জিপিটি-৩

প্রকাশ: OpenAI দ্বারা জুন 2020। আর্কিটেকচার:

GPT-2 আর্কিটেকচার এক্সটেনশন। পরামিতি: 175

বিলিয়ন।

প্রযুক্তিগত বিবরণ: GPT-3 স্কেলকে আরও প্রসারিত করেছে এবং বেশ কিছু স্থাপত্য ও প্রশিক্ষণের উন্নতি করেছে। নির্দিষ্ট টিউনিংয়ের প্রয়োজন ছাড়াই বিভিন্ন কাজের উপর প্রশিক্ষিত।

প্রধান অবদান: "কিছু-শট লার্নিং" সম্পাদন করার ক্ষমতা হাইলাইট করা হয়েছে, যেখানে মডেলটি শুধুমাত্র কয়েকটি উদাহরণ দিয়ে কাজ শিখতে পারে। এই মডেলটি বিষয়বস্তু লেখা এবং সম্পাদনা থেকে শুরু করে প্রোগ্রামিং এবং ইউজার ইন্টারফেস ডিজাইন পর্যন্ত বিস্তৃত অ্যাপ্লিকেশনে ব্যবহার করা হয়েছে, যা ক্ষেত্রে অভূতপূর্ব বহুমুখিতা দেখায় (Brown et al. 2020)।

সারণি 2 শেখার আকার, স্থাপত্য, এবং হাইপারপ্যারামিটার দেখায়।

সারণী 2. মডেলের আকার, আর্কিটেকচার এবং শেখার
হাইপারপ্যারামিটার (টোকেনে ব্যাচের আকার এবং শেখার হার)।

মডেলের নাম	n_{params}	$n_{\text{স্তর}}$	$d_{\text{মডেল}}$	$n_{\text{মাথা}}$	$d_{\text{মাথা}}$	ব্যাচের আকার	শেখার হার
GPT-3 ছোট	125M	12	768	12	64	0.5M	6.0×10^{-4}
GPT-3 মাধ্যম	350M	24	1024	16	64	0.5M	3.0×10^{-4}
GPT-3 বড়	760M	24	1536	16	96	0.5M	2.5×10^{-4}
GPT-3 XL	1.3B	24	2048	24	128	1M	2.0×10^{-4}
GPT-3 2.7B	2.7B	32	2560	32	80	1M	1.6×10^{-4}
GPT-3 6.7B	6.7B	32	4096	32	128	2M	1.2×10^{-4}
GPT-3 13B	13.0B	40	5140	40	128	2M	1.0×10^{-4}
GPT-3 175B বা "GPT-3"	175.0B	96	12288	96	128	3.2M	0.6×10^{-4}

কোডেক্স

রিলিজ: GPT-3 এর একটি বৈকল্পিক, বিশেষভাবে নির্ধারিত কাজের জন্য ডিজাইন করা হয়েছে (Zaremba & Brockman, 2021)।

আর্কিটেকচার: ট্রান্সফরমার আর্কিটেকচারের উপর ভিত্তি করে এবং এটি GPT-3 এর একটি বৈকল্পিক যা বিশেষভাবে নির্ধারিত কাজের জন্য ডিজাইন করা হয়েছে।

পরামিতি: 175 বিলিয়ন পরামিতি।

প্রযুক্তিগত বিবরণ: সোর্স কোড এবং ডকুমেন্টেশনের একটি বৃহৎ সংস্থার উপর প্রাক-প্রশিক্ষিত। এটি প্রাথমিকভাবে কোড তৈরি করতে এবং প্রোগ্রামিং কাজগুলিতে সহায়তা করতে ব্যবহৃত হয়।

প্রধান অবদান: এটি একাধিক প্রোগ্রামিং ভাষায় কার্যকরী কোড তৈরি করতে সক্ষম বলে প্রমাণিত হয়েছে। এটির রিলিজ ডেভেলপারদের কোডের সাথে ইন্টারঅ্যাক্ট করার পদ্ধতিতে বৈশ্বিক পরিবর্তন এনেছে, স্বয়ংক্রিয় কোড তৈরি এবং প্রোগ্রামিং সহায়তার জন্য একটি শক্তিশালী টুল প্রদান করেছে (OpenAI, 2021)।

5.2. ChatGPT, GPT-4

ট্রান্সফরমার আর্কিটেকচারের উপর ভিত্তি করে ভাষার মডেলগুলি সাম্প্রতিক বছরগুলিতে দ্রুত বিকাশ লাভ করেছে, যা OpenAI দ্বারা তৈরি জেনারেটিভ প্রাক-প্রশিক্ষিত ট্রান্সফরমার (GPT) মডেলগুলির সিরিজে পরিণত হয়েছে। বিশেষ করে, GPT-4 এবং ChatGPT যথাক্রমে প্রাকৃতিক ভাষা বোঝার এবং প্রজন্মের ক্ষেত্রে উল্লেখযোগ্য অগ্রগতির প্রতিনিধিত্ব করে। এই মডেলগুলি হল পূর্বোল্লিখিত মডেলগুলির ফলো-আপ এবং গত কয়েক মাস ধরে টেক্সট জেনারেটিভ এআই মডেলগুলিতে একটি নতুন যুগের সূচনা হয়েছে। উভয় মডেলই ট্রান্সফরমার আর্কিটেকচারের উপর ভিত্তি করে তৈরি, যা প্রাকৃতিক ভাষা প্রক্রিয়াকরণের কাজগুলির জন্য অত্যন্ত দক্ষ বলে প্রমাণিত হয়েছে (Vaswani et al., 2017)।

CHATGPT (GPT-3.5)

চ্যাটজিপিটি একটি ভাষা মডেল যা ওপেনএআই দ্বারা তৈরি করা হয়েছে এবং এটি জিপিটি (জেনারেটিভ প্রি-ট্রেনেড ট্রান্সফরমার) আর্কিটেকচারের উপর ভিত্তি করে তৈরি। এটি বিশেষভাবে কথোপকথন এবং চ্যাট কাজের জন্য ডিজাইন করা GPT-3 এর একটি টিউন করা রূপ। মডেলটিকে ইন্টারনেট পার্শ্বের একটি বৃহৎ সংস্থায় প্রশিক্ষণ দেওয়া হয়েছে, তবে এটির প্রশিক্ষণ ডেটাসেটে কোন নির্দিষ্ট নথি ব্যবহার করা হয়েছিল তা জানা যায়নি। এটি সৃজনশীল পার্শ্ব প্রজন্ম থেকে শুরু করে প্রোগ্রামিং এবং প্রযুক্তিগত সমস্যা সমাধানের কাজগুলি সম্পাদন করার জন্য ডিজাইন করা হয়েছে। এই মডেলটি ChatGPT (openai.com) এর মাধ্যমে অবাধে অ্যাক্সেসযোগ্য।

রিলিজ: নাকানো এট আল অনুসারে 2021 সালে OpenAI দ্বারা ChatGPT প্রকাশ করা হয়েছিল। (2021)।
 আর্কিটেকচার: চ্যাটজিপিটিও ট্রান্সফরমার আর্কিটেকচারের উপর ভিত্তি করে এবং এটি কথোপকথনের জন্য
 অস্টিমাইজ করা GPT-3 এর একটি রূপ।

পরামিতি: ChatGPT-এর 175 বিলিয়ন প্যারামিটার রয়েছে, GPT-3 এর মতো।

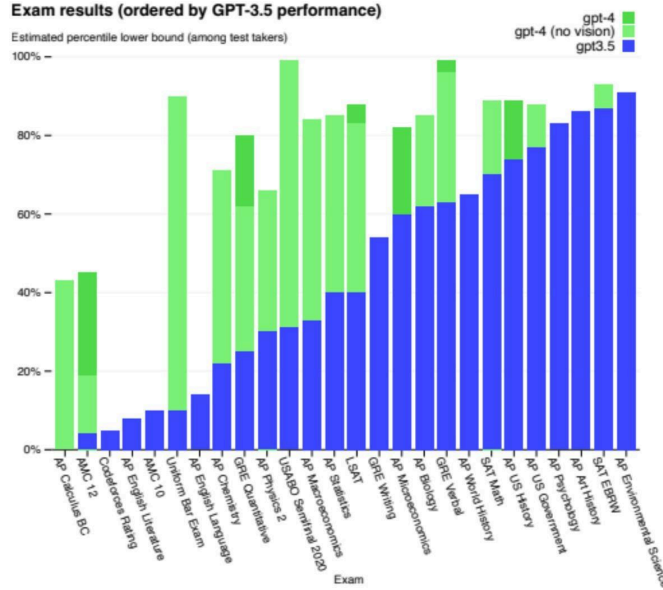
প্রযুক্তিগত বিশদ: এটি কথোপকথন এবং পার্ঠের একটি অংশে প্রাক-প্রশিক্ষিত। চ্যাটবট এবং ইন্টারেক্টিভ
 কথোপকথন সিস্টেমের জন্য ডিজাইন করা হয়েছে।

মূল অবদান: ChatGPT আগের মডেলের তুলনায় কথোপকথনে ধারাবাহিকতা এবং প্রাসঙ্গিকতা উন্নত করেছে।

GPT-4

GPT-4 হল GPT সিরিজের চতুর্থ পুনরাবৃত্তি এবং ক্ষমতা, সীমাবদ্ধতা এবং সংশ্লিষ্ট ঝুঁকির পরিপ্রেক্ষিতে একটি
 কোয়ান্টাম লিপের প্রতিনিধিষ্ণ করে। এখনও অপ্রকাশিত সংখ্যক পরামিতি সহ, GPT-4 টেক্সট জেনারেশন, প্রাসঙ্গিক
 বোঝাপড়া এবং বিভিন্ন কাজে অভিযোজনযোগ্যতার ক্ষেত্রে উল্লেখযোগ্য উন্নতি দেখিয়েছে। এই মডেলটি ChatGPT
 Plus নামে পরিচিত OpenAI-এর পেইড সংস্করণে পাওয়া যায়, এতে GPT-4 মডেলের কিছু বিশিষ্টত্ব এবং প্লাগইন
 (OpenAI, 2023), কোড ইন্টারপ্রেটার এবং OpenAI (2023) অনুযায়ী কাস্টম নির্দেশনার মতো নতুন টুল উপলব্ধ
 রয়েছে। 4 বিভিন্ন কাজের প্রসঙ্গ দেখায়।

প্লাগইনগুলি ডেভেলপারদের মডেলে নির্দিষ্ট কার্যকারিতা যোগ করার অনুমতি দেয়, যেমন ওয়েব অনুসন্ধান,
 ভাষা অনুবাদ এবং বৈজ্ঞানিক ডেটাবেস অনুসন্ধান করা। প্রকাশনা জোর দেয় যে এই প্লাগইনগুলিকে সুরক্ষিত এবং
 নির্ভরযোগ্য করার জন্য ডিজাইন করা হয়েছে এবং অনুমোদনের আগে একটি কঠোর পর্যালোচনা প্রক্রিয়ার মধ্য দিয়ে
 যেতে হবে।



চিত্র 4. শিক্ষাগত এবং পেশাগত মূল্যায়নের কর্মক্ষমতা। প্রতিটি পরীক্ষার জন্য,
 আমরা বাস্তব পরীক্ষার পরিবেশ এবং গ্রেডিং সিস্টেম অনুকরণ করি। GPT-3.5-এর পারফরম্যান্স অনুযায়ী
 মূল্যায়নগুলি আরোহী ক্রমে সাজানো হয়েছে। GPT-4 বেশিরভাগ মূল্যায়ন করা পরীক্ষায় GPT-3.5কে
 ছাড়িয়ে গেছে (OpenAI, 2023)।

প্রকাশ: GPT-4-এর প্রযুক্তিগত প্রতিবেদন OpenAI দ্বারা মার্চ 2023 সালে প্রকাশিত হয়েছিল।
 আর্কিটেকচার: GPT-4 হল ট্রান্সফরমার আর্কিটেকচারের উপর ভিত্তি করে একটি মডেল যা ইমেজ এবং টেক্সট ইনপুট গ্রহণ করতে পারে এবং টেক্সট আউটপুট তৈরি করতে পারে।
 পরামিতি: রিপোর্টটি পরামিতিগুলির সঠিক সংখ্যা নির্দিষ্ট করে না কিন্তু হাইলাইট করে যে মডেলটি বড় আকারের।
 প্রযুক্তিগত বিবরণ: GPT-4 একটি নথিতে পরবর্তী টোকেনের পূর্বাভাস দেওয়ার জন্য প্রাক-প্রশিক্ষিত।
 প্রশিক্ষণ-পরবর্তী প্রান্তিককরণ প্রক্রিয়ার ফলস্বরূপ বাস্তবতার পরিমাপ এবং পছন্দসই আচরণের আনুগত্যের উপর কর্মক্ষমতা উন্নত হয়। অবকাঠামো এবং অপ্টিমাইজেশান পদ্ধতিগুলি তৈরি করা হয়েছিল যা বিস্তৃত স্কেলগুলির উপর অনুমানযোগ্যভাবে আচরণ করে।
 প্রধান অবদান: GPT-4 হল একটি বৃহৎ মাল্টিমডাল মডেল (ছবি এবং পাঠ্য ইনপুট গ্রহণ করা, পাঠ্য আউটপুট নির্গত করা) যেটি বাস্তব-বিশ্বের অনেক পরিস্থিতিতে মানুষের চেয়ে কম সক্ষম হলেও বিভিন্ন পেশাদার এবং একাডেমিক বেকমার্কে মানব-স্তরের কর্মক্ষমতা প্রদর্শন করে (OpenAI, 2023)।

5.3. প্রম্পট

প্রম্পট হল নির্দেশাবলী, প্রশ্ন বা বিবৃতি যা একটি প্রাকৃতিক ভাষা প্রক্রিয়াকরণ (NLP) মডেল, যেমন OpenAI-এর GPT মডেলগুলি থেকে একটি নির্দিষ্ট প্রতিক্রিয়া আহ্বান করার জন্য ডিজাইন করা হয়েছে। চ্যাটবট এবং অন্যান্য PLN অ্যাপ্লিকেশনের প্রসঙ্গে, প্রম্পটগুলি প্রাথমিক ইনপুট হিসাবে কাজ করে যা মডেলটিকে একটি সুসংগত এবং প্রাসঙ্গিক প্রতিক্রিয়া তৈরিতে গাইড করে। মডেলের সাথে মিথস্ক্রিয়া নিয়ন্ত্রণ এবং নির্দেশিত করার ক্ষেত্রে তাদের ভূমিকা গুরুত্বপূর্ণ, উৎপন্ন প্রতিক্রিয়ার গুণমান, নির্ভুলতা এবং প্রেক্ষাপটকে প্রভাবিত করে।

5.3.1. ভালো প্রম্পট স্থাপনের জন্য কৌশল

- স্বচ্ছতা এবং নির্ভুলতা: একটি ভাল প্রম্পট স্পষ্ট এবং সুনির্দিষ্ট হওয়া উচিত, অস্পষ্টতাগুলি এড়ানো উচিত যা বিভ্রান্তিকর বা ভুল উত্তরের দিকে নিয়ে যেতে পারে (Reiter et al., 2020)।
- প্রাসঙ্গিককরণ: প্রয়োজনীয় প্রসঙ্গ অন্তর্ভুক্ত করা মডেলটিকে প্রশ্নের পিছনের উদ্দেশ্য বুঝতে সাহায্য করে, প্রতিক্রিয়ার প্রাসঙ্গিকতা উন্নত করে (চেন এট আল।, 2019)।
- উদাহরণের ব্যবহার: অল্প-শট শেখার ক্ষেত্রে, প্রম্পটের মধ্যে উদাহরণ প্রদান করা মডেলকে উদ্দেশ্যমূলক কাজটি বুঝতে সাহায্য করতে পারে (Brown et al., 2020)।
- পুনরাবৃত্তিমূলক পরীক্ষা: পুনরাবৃত্তিমূলক পরীক্ষা এবং প্রম্পটের সমন্বয় মিথস্ক্রিয়াকে সূক্ষ্ম সূর করতে এবং সর্বোত্তম প্রতিক্রিয়া পেতে দেয় (ওয়ালেস এট আল।, 2019)।
- নৈতিক বিবেচনা: সম্ভাব্য পক্ষপাত সম্পর্কে সচেতনতার সাথে প্রম্পট প্রণয়ন করা উচিত এবং গোপনীয়তা এবং নৈতিকতা বিধি মেনে চলতে হবে (Hovy & Spruit, 2016)।

প্রম্পটগুলি জেনারেটিভ ল্যাঙ্গুয়েজ মডেল এবং অন্যান্য PLN অ্যাপ্লিকেশনগুলির ডিজাইন এবং মিথস্ক্রিয়াতে একটি কেন্দ্রীয় ভূমিকা পালন করে। তাদের গঠন এবং ব্যবস্থাপনা প্রযুক্তিগত, ভাষাগত এবং নৈতিক বিবেচনার সমন্বয় জড়িত। PLN-এর সাহিত্য প্রম্পটগুলির কার্যকর ব্যবহারের সাথে সম্পর্কিত গবেষণা এবং প্রযুক্তির বিস্তৃত বর্ণালী অফার করে এবং তাদের অধ্যয়ন এবং প্রয়োগ মানুষের এবং কৃত্রিম বুদ্ধিমত্তা সিস্টেমের মধ্যে ইন্টারফেসের একটি গুরুত্বপূর্ণ ক্ষেত্র হিসাবে অবিরত রয়েছে।

5.3.2. প্রম্পট এবং তাদের ব্যবহার প্রকার

- তথ্যমূলক প্রম্পট: নির্দিষ্ট তথ্যের অনুরোধ করার জন্য ডিজাইন করা হয়েছে, এগুলি ডেটা অনুসন্ধান এবং ভার্চুয়াল সহকারীর মতো অ্যাপ্লিকেশনগুলিতে কার্যকর (ম্যানিং এট আল।, 2008)।
- জিজ্ঞাসাবাদমূলক প্রম্পট: প্রশ্ন হিসাবে প্রণয়ন করা হয়েছে, তারা গ্রাহক সহায়তা এবং টিউটরিং (Jurafsky & Martin, 2019) এর মতো ক্ষেত্রগুলিতে বিস্তারিত প্রতিক্রিয়া জানাতে নিযুক্ত করা হয়।

- নির্দেশমূলক প্রম্পট: অনুবাদ, সংক্ষিপ্তকরণ, বা সৃজনশীল বিষয়বস্তু তৈরির মতো একটি নির্দিষ্ট কাজ সম্পাদন করার জন্য মডেলকে নির্দেশ দিতে ব্যবহৃত হয় (Reiter & Dale, 2020)।
- প্রাসঙ্গিক প্রম্পট: চলমান কথোপকথন বা বিশেষ কাজগুলির মতো জটিল পরিস্থিতিতে মডেলের প্রতিক্রিয়া গাইড করার জন্য অতিরিক্ত প্রসঙ্গ অন্তর্ভুক্ত করুন (Serban et al., 2017)।
- তুলনামূলক প্রম্পট: সত্য বা ধারণার মধ্যে তুলনা, বিশ্লেষণ বা মূল্যায়ন করার জন্য ডিজাইন করা হয়েছে।
- কোড জেনারেশন প্রম্পট: প্রোগ্রামিং পরিবেশে স্বয়ংক্রিয়ভাবে কোড স্লিপেট তৈরি করতে ব্যবহৃত হয়, যেমন OpenAI কোডেজে (OpenAI, 2021)।
- ভাষাগত সহায়তার অনুরোধ: ব্যাকরণ সংশোধন, অনুবাদ, এবং অন্যান্য ভাষাগত পরিষেবাগুলির জন্য সরঞ্জামগুলিতে প্রয়োগ করা হয়, যেমন গ্রামারলি বা Google অনুবাদ।

5.3.3. উদাহরণ: কোড জেনারেশন

স্বয়ংক্রিয় কোড জেনারেশন হল একটি কৌশল যা প্রোগ্রামে সোর্স কোড তৈরি করতে ব্যবহৃত হয়। এই অভ্যাসটি বিশেষভাবে উপযোগী হয় যখন প্রচুর পরিমাণে অনুরূপ বা পুনরাবৃত্তিমূলক কোড তৈরি করার প্রয়োজন হয়। কোডের প্রতিটি অংশ ম্যানুয়ালি লেখার পরিবর্তে, বিকাশকারীরা প্রক্রিয়াটি স্বয়ংক্রিয় করতে কোড জেনারেটর ব্যবহার করতে পারে। এই প্রবন্ধে, আমরা অন্বেষণ করব কিভাবে পাইথন দিয়ে স্বয়ংক্রিয়ভাবে কোড তৈরি করা যায় এবং গাণিতিক ফাংশন তৈরির জন্য কোড তৈরির একটি সহজ উদাহরণ উপস্থাপন করব, চিত্র 5, 6, 7 দেখুন।

কেন স্বয়ংক্রিয়ভাবে কোড জেনারেট করবেন?

দক্ষতা: কোড লেখার জন্য প্রয়োজনীয় সময় হ্রাস করে।

সামঞ্জস্যতা: নিশ্চিত করে যে তৈরি করা কোড একটি নির্দিষ্ট প্যাটার্ন অনুসরণ করে।

নমনীয়তা: ন্যূনতম প্রচেষ্টার সাথে কোডে বিশ্বব্যাপী পরিবর্তনের অনুমতি দেয়।

পাইথনে কোড জেনারেশনের উদাহরণ

স্বয়ংক্রিয়ভাবে পাইথন ফাংশন তৈরি করে যা একটি প্রদত্ত সংখ্যার বর্গ, ঘনক এবং চতুর্থ শক্তি গণনা করে।

প্রথমে, আসুন একটি পাইথন ফাংশন সংজ্ঞায়িত করি যা একটি সূচককে একটি প্যারামিটার হিসাবে নেয় এবং একটি ফাংশনের উত্স কোড প্রদান করে যা সেই সূচকে একটি সংখ্যা বাড়ায়।

```
python Copy code

def generate_power_function(exponent):
    function_name = f"power_of_{exponent}"
    code = f"""
def {function_name}(x):
    return x ** {exponent}
"""
    return code
```

চিত্র 5. পাইথন এক্সপোনেন্ট ফাংশন।

উৎপন্ন ফাংশন

এখন, আমরা এই ফাংশনটি ব্যবহার করব ফাংশনগুলির জন্য কোড তৈরি করতে যা একটি সংখ্যার বর্গ, ঘনক এবং চতুর্থ শক্তি গণনা করে।

```
python Copy code  
  
exponents = [2, 3, 4]  
for exp in exponents:  
    print(generate_power_function(exp))
```

চিত্র 6. পাইথন তালিকা ফাংশন।

উৎপন্ন কোড এই মত দেখায়:

```
python Copy code  
  
def power_of_2(x):  
    return x ** 2  
  
def power_of_3(x):  
    return x ** 3  
  
def power_of_4(x):  
    return x ** 4
```

চিত্র 7. পাইথন গণিত ফাংশন।

generate_power_function ফাংশন একটি সূচক নেয় এবং একটি কোড স্লিপেট প্রদান করে যা একটি নতুন পাইথন ফাংশন সংজ্ঞায়িত করে। এই নতুন ফাংশনটি একটি প্যারামিটার x নেয় এবং প্রদত্ত সূচকে x উত্থাপিত করে।

"জেনারেটিং ফাংশন" সেগমেন্টের জন্য লুপ সূচকের একটি তালিকার মাধ্যমে চলে। প্রতিটি এক্সপোনেন্টের জন্য, সংশ্লিষ্ট কোড তৈরি করতে generate_power_function বলা হয়।

স্বয়ংক্রিয় কোড তৈরি একটি শক্তিশালী কৌশল যা সময় এবং প্রচেষ্টা বাঁচাতে পারে, বিশেষ করে যখন পুনরাবৃত্তিমূলক বা অনুরূপ কোডের সাথে কাজ করে। উপস্থাপিত উদাহরণটি বেশ সহজ, কিন্তু কোড জেনারেশন ফ্রেমওয়ার্ক এবং সফটওয়্যার লাইব্রেরির সহ আরও জটিল অ্যাপ্লিকেশনগুলিতে ব্যবহৃত হয়।

5.4. নির্দেশিকা এবং ডকুমেন্টেশন

5.4.1. নৈতিকতা এবং দায়িত্বশীল ব্যবহার

চ্যাটজিপিটি সাধারণত OpenAI দ্বারা প্রদত্ত একটি API-এর মাধ্যমে অ্যাক্সেস করা হয়, যার জন্য মডেলের সাথে মিথস্ক্রিয়াগুলির নিরাপত্তা এবং অখণ্ডতা নিশ্চিত করার জন্য প্রমাণীকরণের প্রয়োজন হয়, এটি অত্যন্ত বহুমুখী এবং গ্রাহক পরিষেবা চ্যাটবট থেকে ভার্চুয়াল সহকারী পর্যন্ত বিভিন্ন অ্যাপ্লিকেশনে প্রয়োগ করা যেতে পারে। এবং সিদ্ধান্ত গ্রহণের ব্যবস্থা। API বিভিন্ন প্রেক্ষাপটে মডেলের আচরণকে গাইড করার জন্য নির্দিষ্ট প্রম্পট বা নির্দেশাবলী সংজ্ঞায়িত করার ক্ষমতা সহ বিস্তৃত কাস্টমাইজেশনের অনুমতি দেয়। এছাড়াও, ওপেনএআই প্রযুক্তিগত দিকগুলি যেমন প্রমাণীকরণ, প্রম্পট স্ট্রাকচারিং এবং মডেল দ্বারা উত্পন্ন প্রতিক্রিয়াগুলির ব্যাখ্যা সহ বিস্তারিত ডকুমেন্টেশন সরবরাহ করে (ওপেনএআই, 2021)।

ChatGPT-এ অ্যাক্সেস শুধুমাত্র API এর মাধ্যমে ইন্টারঅ্যাকশনের মধ্যেই সীমাবদ্ধ নয়, এটি OpenAI দ্বারা প্রদত্ত একটি ওয়েব ইন্টারফেসের মাধ্যমেও অ্যাক্সেসযোগ্য। এই ওয়েব ইন্টারফেসটি উন্নত প্রযুক্তিগত জ্ঞানের প্রয়োজন ছাড়াই মডেলের সাথে ইন্টারঅ্যাক্ট করার একটি স্বজ্ঞাত এবং সহজেই ব্যবহারযোগ্য উপায় সরবরাহ করে। ব্যবহারকারীরা কেবল একটি পাঠ্য বাস্তু তাদের প্রম্পট বা প্রশ্নগুলি প্রবেশ করান এবং মডেলটি একই ইন্টারফেসে প্রদর্শিত উত্তরগুলি তৈরি করে। অ্যাক্সেসের এই পদ্ধতিটি বিশেষত অ-প্রযুক্তিগত ব্যবহারকারীদের জন্য বা যারা মডেলের ক্ষমতাগুলিকে একটি বৃহত্তর অ্যাপ্লিকেশন বা সিস্টেমে সংহত না করেই পরীক্ষা করতে চান তাদের জন্য উপযোগী। উপরন্তু, ওয়েব ইন্টারফেসে প্রায়শই অতিরিক্ত বৈশিষ্ট্য অন্তর্ভুক্ত থাকে, যেমন তাপমাত্রা এবং সর্বাধিক প্রতিক্রিয়া দৈর্ঘ্যের মতো পরামিতিগুলি সামঞ্জস্য করার ক্ষমতা, যা পাঠ্য প্রজন্মের উপর বৃহত্তর নিয়ন্ত্রণ প্রদান করে (ওপেনএআই, 2021)।

ChatGPT এর সাথে ইন্টারঅ্যাক্ট করার সময় নৈতিক এবং দায়িত্বশীল ব্যবহারের নির্দেশিকা মাথায় রাখা অত্যন্ত গুরুত্বপূর্ণ, বিশেষ করে এমন অ্যাপ্লিকেশনগুলিতে যেগুলির উল্লেখযোগ্য সামাজিক বা সাংস্কৃতিক প্রভাব থাকতে পারে। ওপেনএআই এই সমস্যাগুলির সমাধানের জন্য নির্দিষ্ট নির্দেশিকা প্রদান করে, যার মধ্যে চুরি প্রতিরোধ এবং অনুপযুক্ত বিষয়বস্তু তৈরি করা রয়েছে।

5.4.2. নৈতিকতা এবং দায়িত্বশীল ব্যবহার

ডকুমেন্টেশন নৈতিক এবং দায়িত্বশীল ব্যবহারের সমস্যাগুলিকেও সম্বোধন করে। এর মধ্যে চুরির প্রতিরোধ, অনুপযুক্ত বিষয়বস্তু তৈরি এবং বৃহৎ মাপের ভাষা মডেল ব্যবহার করার সামাজিক ও সাংস্কৃতিক প্রভাব বিবেচনা করার নির্দেশিকা অন্তর্ভুক্ত রয়েছে। এই মডেলগুলি, বড় ডেটা-সেটে প্রশিক্ষিত, বৈষম্যমূলক, পক্ষপাতদুষ্ট, বা এমনকি বিপজ্জনক হতে পারে এমন সামগ্রী তৈরি করার সম্ভাবনা রয়েছে (হাও, 2020; বেন্ডার এট আল।, 2021)। অতএব, প্রযুক্তিগত এবং সামাজিক উভয় দিককে অন্তর্ভুক্ত করে একটি বহু-বিভাগীয় দৃষ্টিকোণ থেকে নৈতিক সমস্যাগুলিকে মোকাবেলা করা অপরিহার্য।

সবচেয়ে চাপের চ্যালেঞ্জগুলির মধ্যে একটি হল প্রশিক্ষণের ডেটার অন্তর্নিহিত পক্ষপাত, যা বিদ্যমান স্টেরিওটাইপ এবং পক্ষপাতগুলিকে স্থায়ী করতে পারে (ক্যালিস্কান এট আল।, 2017)। গবেষকরা এই পক্ষপাতগুলি প্রশমিত করার জন্য পদ্ধতিগুলি অন্বেষণ করছেন, যেমন প্যারামিটার সামঞ্জস্য এবং প্রশিক্ষণ ডেটার পুনঃওয়েটিং (ঝাও এট আল।, 2018)।

মডেলের ক্রিয়াকলাপে স্বচ্ছতা এবং তাদের সিদ্ধান্তের সন্ধানযোগ্যতা AI এর নৈতিক ব্যবহারের জন্য মৌলিক। এটি স্বাস্থ্যসেবা এবং বিচার ব্যবস্থার মতো জটিল অ্যাপ্লিকেশনগুলিতে বিশেষভাবে প্রাসঙ্গিক, যেখানে একটি বিপথগামী সিদ্ধান্তের গুরুত্বের পরিণতি হতে পারে (দোশি-ভেলেজ এট আল।, 2017)।

দায়বদ্ধ ব্যবহারের মধ্যে সুরক্ষার প্রয়োগ করা জড়িত, যেমন বিষয়বস্তু নিয়ন্ত্রণ ব্যবস্থা এবং অনুপযুক্ত বিষয়বস্তুর জন্য সতর্কতা। এই মডেলগুলির সীমাবদ্ধতা এবং কীভাবে তাদের প্রতিক্রিয়াগুলি সমালোচনামূলকভাবে ব্যাখ্যা করা যায় সে সম্পর্কে ব্যবহারকারীদের শিক্ষিত করাও গুরুত্বপূর্ণ (McGregor et al., 2020)।

AI-তে নৈতিকতা হল একটি দ্রুত বিকাশমান ক্ষেত্র যার জন্য প্রকৌশলী, নীতিবিদ, আইনপ্রণেতা এবং অন্যান্য প্রাসঙ্গিক স্টেকহোল্ডারদের মধ্যে অবিরত সহযোগিতা প্রয়োজন যাতে প্রযুক্তিটি এমনভাবে ব্যবহার করা হয় যা সামগ্রিকভাবে সমাজের জন্য উপকারী।

5.4.3. কম্পিউটার প্রোগ্রামিং এবং শিক্ষায় অ্যাপ্লিকেশন

ওপেনএআই কোডেক্স সাধারণ পরিচায়ক প্রোগ্রামিং সমস্যার ক্ষেত্রে উল্লেখযোগ্য কর্মক্ষমতা প্রদর্শন করেছে। এটির কর্মক্ষমতা একই পরীক্ষায় অংশগ্রহণকারী ছাত্রদের সাথে তুলনা করা হয়েছিল, দেখায় যে কোডেক্স বেশিরভাগ ছাত্রদেরকে ছাড়িয়ে গেছে। উপরন্তু, আমরা অন্বেষণ করেছি কিভাবে কোডেক্স সমস্যা শব্দের সূক্ষ্ম পরিবর্তনগুলি পরিচালনা করে, উল্লেখ্য যে অভিন্ন ইনপুট প্রায়শই অ্যালগরিদমিক পদ্ধতি এবং কোডের দৈর্ঘ্যের ক্ষেত্রে খুব ভিন্ন সমাধানের দিকে নিয়ে যায়। এই গবেষণাটি কম্পিউটার বিজ্ঞান শিক্ষার উপর এই ধরনের প্রযুক্তির প্রভাব নিয়েও আলোচনা করে কারণ এটি ক্রমাগত বিকশিত হচ্ছে (ফিনি-অ্যান্সলে এট আল।, 2022)।

একটি বড় ভাষা মডেল হিসাবে OpenAI কোডেক্স ব্যবহার করে, প্রোগ্রামিং অনুশীলন (নমুনা সমাধান এবং পরীক্ষার কেস সহ) এবং কোড ব্যাখ্যা তৈরি করা হয়েছিল। ফলাফলগুলি প্রস্তাব করে যে স্বয়ংক্রিয়ভাবে তৈরি হওয়া বেশিরভাগ বিষয়বস্তুই অভিনব এবং সংবেদনশীল, এবং কিছু ক্ষেত্রে ব্যবহার করার জন্য প্রস্তুত। এই অধ্যয়নটি ওপেনএআই কোডেক্সের প্রভাব এবং প্রাথমিক প্রোগ্রামিং শিক্ষার জন্য অনুরূপ সরঞ্জামগুলি নিয়েও আলোচনা করে এবং ভবিষ্যতের গবেষণার লাইনগুলিকে হাইলাইট করে যা শিক্ষক এবং ছাত্র উভয়ের জন্য শিক্ষাগত অভিজ্ঞতার মান উন্নত করার সম্ভাবনা রাখে (সারসা এট আল।, 2022)।

GitHub Copilot, OpenAI কোডেক্স দ্বারা চালিত, 166টি প্রোগ্রামিং সমস্যার ডেটাসেটে মূল্যায়ন করা হয়েছে। এটি প্রথম প্রচেষ্টায় এই সমস্যাগুলির প্রায় অর্ধেক সফলভাবে সমাধান করতে দেখা গেছে এবং সমস্যার বর্ণনায় শুধুমাত্র প্রাকৃতিক ভাষা পরিবর্তন ব্যবহার করে বাকি 60% সমস্যার সমাধান করেছে। এই সমীক্ষাটি যুক্তি দেয় যে এই ধরনের প্রম্পট ইঞ্জিনিয়ারিং একটি সম্ভাব্য উপযোগী শেখার ক্রিয়াকলাপ যা কম্পিউটেশনাল চিন্তাভাবনা দক্ষতাকে উত্সাহিত করে এবং কোড লেখার দক্ষতা বিকাশের প্রকৃতি পরিবর্তন করতে পারে (ডেনি এট আল।, 2022)।

5.4.4. একাডেমিক প্রকাশনার অ্যাপ্লিকেশন

ChatGPT-কে প্রবন্ধ এবং অন্যান্য ধরনের একাডেমিক পাণ্ডুলিপির স্বয়ংক্রিয় প্রস্তুতির জন্য একটি সম্ভাব্য মডেল হিসেবে দেখা হয়। কৃত্রিম বুদ্ধিমত্তার বৃহত্তর অগ্রগতির প্রেক্ষাপটে তাদের রেখে, ChatGPT-এর পিছনে অন্তর্নিহিত প্রযুক্তি, GPT-3-এর মতো বৃহৎ ভাষার মডেলের আবির্ভাবের সাথে সম্ভাব্য নৈতিক সমস্যা এবং শিক্ষাবিদ ও গবেষকদের দ্বারা তাদের ব্যবহার নিয়ে আলোচনা করা হয়েছে। , মেশিন লার্নিং, এবং একাডেমিক গবেষণা এবং প্রকাশনার জন্য প্রাকৃতিক ভাষা প্রক্রিয়াকরণ (লুন্ড এট আল।, 2023)।

মডেলটির প্রকাশ অনেককে উত্তেজনাপূর্ণ এবং সমস্যায়ুক্ত উপায় সম্পর্কে চিন্তা করতে পরিচালিত করেছে যেখানে কৃত্রিম বুদ্ধিমত্তা (AI) অদূর ভবিষ্যতে আমাদের জীবন পরিবর্তন করতে পারে। তদ্ব্যবধানে এবং রিইনফোর্সমেন্ট লার্নিং সহ GPT-3 মডেলকে ফাইন-টিউনিং করে চ্যাটজিপিটি তৈরি করা হয়েছে তা বিবেচনা করে, উত্পন্ন সামগ্রীর গুণমান শুধুমাত্র অতিরিক্ত প্রশিক্ষণ এবং অস্টিমাইজেশনের মাধ্যমে উন্নত করা যেতে পারে। এই বিঘ্নিত প্রযুক্তির অনিবার্য প্রয়োগের ফলে ওষুধ, বিজ্ঞান এবং একাডেমিক প্রকাশনার সুদূরপ্রসারী পরিণতি হবে (Homolak, 2023)।

এটি বৈজ্ঞানিক লেখার ক্ষেত্রেও একটি দরকারী টুল বলে মনে হয়, গবেষক এবং বিজ্ঞানীদেরকে উপাদান সংগঠিত করতে, একটি প্রাথমিক খসড়া তৈরি করতে এবং/অথবা প্রফরিডিং করতে সহায়তা করে। এই সরঞ্জামগুলির ব্যবহার সম্পর্কিত বেশ কিছু নৈতিক সমস্যা উত্থাপিত হয়, যেমন চুরির ঝুঁকি এবং ভুলতার পাশাপাশি উচ্চ-আয়ের এবং নিম্ন-আয়ের দেশগুলির মধ্যে অ্যাক্সেসযোগ্যতার সম্ভাব্য ব্যবধান (M. Salvagno et al., 2023) .

5.4.5. কথোপকথন সিস্টেম

মানুষের মতো জ্ঞান স্থানান্তর এবং যোগাযোগের জন্য ওষুধ এবং স্বাস্থ্যসেবা সহ বিভিন্ন ক্ষেত্রে চ্যাটবটগুলি প্রয়োগ করা হচ্ছে। বিশেষ করে, জটিল কথোপকথন পরিচালনা এবং কথোপকথন নমনীয়তা প্রদান করার ক্ষমতা সহ, মেশিন লার্নিং স্বাস্থ্যসেবায় প্রযোজ্য বলে দেখানো হয়েছে। এই পর্যালোচনাটি ক্যান্সার থেরাপির উপর দৃষ্টি নিবদ্ধ করে, বিস্তারিত আলোচনা এবং উদাহরণ সহ নির্ণয়, চিকিত্সা, পর্যবেক্ষণ, রোগীর সহায়তা, কর্মপ্রবাহের দক্ষতা এবং স্বাস্থ্য প্রচার। উপরন্তু, সীমাবদ্ধতা এবং উদ্বেগের ক্ষেত্রগুলি অন্বেষণ করা হয়েছে, নৈতিক, নৈতিক, নিরাপত্তা, প্রযুক্তিগত এবং নিয়ন্ত্রক বিষয়গুলিকে হাইলাইট করে (Xu, L. et al., 2021)।

মেশিন লার্নিং (এমএল) হল অভিজ্ঞতার মাধ্যমে অটোমেশনের জন্য কম্পিউটার অ্যালগরিদমের একটি অধ্যয়ন। স্বাস্থ্যসেবা যোগাযোগে এমএল-এর প্রয়োগ মানুষের জন্য উপকারী বলে প্রমাণিত হয়েছে। এর মধ্যে রয়েছে COVID-19-এ স্বাস্থ্য শিক্ষার জন্য চ্যাটবট, ক্যান্সার থেরাপি এবং মেডিকেল ইমেজিং। পর্যালোচনাটি হাইলাইট করে যে কীভাবে স্বাস্থ্যসেবা যোগাযোগে এমএল/এআই-এর প্রয়োগ জটিল কথোপকথন ব্যবস্থাপনা এবং কথোপকথন নমনীয়তা সহ মানুষের উপকার করতে সক্ষম হয় (সরকার সিদ্ধিক এবং জেমস সি.এল. চৌ, 2021)।

এই নিবন্ধটি তথ্য বিজ্ঞানে ChatGPT ব্যবহারের সাথে যুক্ত সুযোগ এবং চ্যালেঞ্জগুলির একটি ওভারভিউ প্রদান করে। এটি আলোচনা করে যে কীভাবে চ্যাটজিপিটি ডেটা বিজ্ঞানীদের তাদের কর্মপ্রবাহের বিভিন্ন দিক স্বয়ংক্রিয় করতে সহায়তা করতে পারে, যার মধ্যে ডেটা পরিষ্কার এবং প্রিপ্রসেসিং, মডেল প্রশিক্ষণ এবং ফলাফলের ব্যাখ্যা রয়েছে। এটি আরও হাইলাইট করে যে ChatGPT-এর কীভাবে নতুন অন্তর্দৃষ্টি প্রদান করার এবং সিদ্ধান্ত গ্রহণের প্রক্রিয়াগুলিকে উন্নত করার সম্ভাবনা রয়েছে (হোসেন হাসানি এবং ই. সিলভা, 2023)।

5.4.6. ঝুঁকি

এই প্রযুক্তি বিষয়বস্তু তৈরি এবং লেখার উপর ভিত্তি করে অনেক শিল্পের ভিত্তিকে নাড়িয়ে দিয়েছে, তাই ChatGPT এবং অন্যান্য কৃত্রিম বুদ্ধিমত্তা ভাষার মডেলগুলির সাথে বৈজ্ঞানিক নিবন্ধ লেখার জন্য ভাল অনুশীলনগুলি নিয়ে আলোচনা করা বাঞ্ছনীয় (A. Castellanos-Gómez, 2023)। এই প্রযুক্তির ব্যবহার এবং এটি কীভাবে করা উচিত সে সম্পর্কে বিভিন্ন সূত্র রয়েছে।

এই ভাষা মডেলগুলি তাদের পূর্বসূরীদের তুলনায় যুক্তি, জ্ঞান ধারণ এবং প্রোগ্রামিংয়ে উল্লেখযোগ্য অগ্রগতি দেখায়। যাইহোক, এই উন্নতিগুলি ক্ষতিকারক বিষয়বস্তু তৈরি, ভুল তথ্য এবং সাইবার নিরাপত্তার মতো ঝুঁকি সহ নতুন নিরাপত্তা চ্যালেঞ্জও নিয়ে আসে। প্রশমনের ব্যবস্থা থাকা সত্ত্বেও, অনুরূপ সীমাবদ্ধতাগুলি এর পূর্ববর্তী সংস্করণগুলির মতো মডেলটিতে রয়ে গেছে, যেমন পক্ষপাতদুষ্ট এবং অবিশ্বস্ত বিষয়বস্তু তৈরি করা। উপরন্তু, এর বর্ধিত ধারাবাহিকতা জেনারেট করা বিষয়বস্তুকে আরও বিশ্বাসযোগ্য করে তুলতে পারে এবং সেইজন্য সম্ভাব্য আরও বিপজ্জনক (OpenAI, 2023)। এই মডেলগুলির ব্যবহারে বিবেচনা করা ঝুঁকিগুলি তালিকাভুক্ত করা হয়েছে। সর্বদা হিসাবে, যে ব্যক্তি এই সরঞ্জামগুলি ব্যবহার করবে সে শেষ পর্যন্ত দায়ী হবে, যেহেতু, আজ পর্যন্ত, উত্পন্ন একটি পুঙ্খানুপুঙ্খ তত্ত্বাবধান

বিষয়বস্তু এটি নিশ্চিত করার জন্য প্রয়োজনের চেয়ে বেশি:

- ভুল তথ্য।
- ক্ষতিকারক বিষয়বস্তু।
- কর্মক্ষমতা, অ্যাসাইনমেন্ট এবং পরিষেবার মানের ক্ষতি।
- বিভ্রান্তি এবং প্রভাব অপারেশন।
- প্রচলিত ও অপ্রচলিত অপ্লের বিস্তার।
- গোপনীয়তা।
- সাইবার নিরাপত্তা।

- অন্যান্য সিস্টেমের সাথে মিথস্ক্রিয়া।
- অর্থনৈতিক প্রভাব।
- অত্যধিক নির্ভরশীলতা।

গাও এট আল অনুসারে। (2023) ChatGPT দ্বারা উত্পন্ন বৈজ্ঞানিক বিমূর্তগুলি একটি কৃত্রিম বুদ্ধিমত্তা আউটপুট ডিটেক্টর, একটি চৌর্যবৃত্তি আবিষ্কারক এবং অন্ধ সহকর্মী পর্যালোচনা ব্যবহার করে মূল বিমূর্তগুলির সাথে তুলনা করা হয়েছিল। ফলাফলগুলি দেখায় যে চ্যাটজিপিটি-উত্পন্ন বিমূর্তগুলি স্পষ্টভাবে লেখা ছিল, তবে মাত্র 8% জার্নাল-নির্দিষ্ট ফর্ম্যাটিং প্রয়োজনীয়তা অনুসরণ করে। যদিও উত্পন্ন বিমূর্তগুলি কোনও শনাক্ত করা চুরি ছাড়াই আসল ছিল, সেগুলি প্রায়শই এআই আউটপুট ডিটেক্টর ব্যবহার করে এবং সন্দেহবাদী মানব পর্যালোচনাকারীদের দ্বারা চিহ্নিত করা হয়েছিল। উপসংহার হল যে চ্যাটজিপিটি বিশ্বাসযোগ্য বৈজ্ঞানিক বিমূর্তগুলি লেখে, যাইহোক, এটি উত্পন্ন সামগ্রীর নির্ভুলতা, সম্পূর্ণতা এবং মৌলিকতা সম্পর্কিত নৈতিক সমস্যা এবং চ্যালেঞ্জও উত্থাপন করে।

6. কল করুন 2

বড় ভাষা মডেল (LLM) এন্টারপ্রাইজ স্তরে খুব প্রতিশ্রুতিশীল প্রমাণিত হয়েছে। এটি মেটা কোম্পানির ক্ষেত্রে, যেটি কৃত্রিম বুদ্ধিমত্তার বিষয়ে তার দৃষ্টিভঙ্গির মধ্যে তার বিকেন্দ্রীকরণের চেষ্টা করেছে, যেখানে সংস্থাগুলি তাদের ভার্সুয়াল সহকারীকে কাস্টমাইজ করতে পারে এবং ওপেন-সোর্স মডেলগুলিকে বিশেষজ্ঞ জ্ঞানের সাথে প্রশিক্ষিত করা যেতে পারে। সে কারণেই মার্ক জুকাবের্গের নেতৃত্বে উদ্ভাবন ও প্রযুক্তির বাজারে মেটা, মাইক্রোসফট আজুরের সাথে অংশীদারিত্ব করেছে তার জেনারেটিভ AI টুল Llama 2 (Touvron, H et al., 2023), যাকে ChatGPT এবং বার্ডের পার্থক্যের প্রতিদ্বন্দ্বী হিসেবে বিবেচনা করা হয়েছে। নিজে একটি ওপেন সোর্স এবং ক্লোজড প্রোডাক্ট নয় যেখানে এই ধরনের এলএলএমগুলি মূলত মানুষের পছন্দের সাথে সামঞ্জস্য করা হয় যেখানে নিরাপত্তা এবং ব্যক্তিগত তথ্য রাখা হয় বাজি

6.1. লামা 2 কী এবং এর বৈশিষ্ট্যগুলি কী কী?

Llama 2 হল জেনারেটিভ ট্রান্সফরমার মডেল ব্যবহার করে প্রাক-প্রশিক্ষিত এবং টিউন করা LLM-এর একটি পরিবার, যেখানে সর্বজনীনভাবে উপলব্ধ ডেটা নেওয়া হয়েছে এবং Llama 1 সংস্করণের ক্ষেত্রে প্রসঙ্গ দৈর্ঘ্য 40% বৃদ্ধি করা হয়েছে। মডেলের ভেরিয়েন্ট, 7B, 13B, 34B এবং 70B এর প্যারামিটার সহ সংস্করণগুলি ওপেন সোর্স সম্প্রদায়ের জন্য উপলব্ধ করা হয়েছে। Llama 2 বিনামূল্যে লাইসেন্স এবং অনিয়ন্ত্রিত বিতরণের একটি মডেল মেনে চলে, বিভিন্ন ডোমেনের জন্য একাধিক সুযোগ তৈরি করে (Touvron, H. et al., 2023)। সারণি 3 বিভিন্ন লামা 2 মডেল দেখায়।

সারণি 3. মডেল লামা 2 আর্কিটেকচার।

মডেল প্যারামিটার সাইজ	পূর্বপ্রশিক্ষিত	চ্যাট ব্যবহারের ক্ষেত্রে ফাইন-টিউন করা হয়েছে
7B	মডেল আর্কিটেকচার প্রিট্রেনিং টোকেন: 2 ট্রিলিয়ন প্রসঙ্গ দৈর্ঘ্য: 4096	সহায়কতা এবং নিরাপত্তার জন্য ডেটা সংগ্রহ
13B		তন্ম্বাবধানে ফাইন-টিউনিং:
70B		100,000 এর বেশি হুমা পছন্দ: 1,000,000 এর বেশি

6.2. বৈশিষ্ট্য

প্রযুক্তি স্ট্যাক: Llama 2 তার অপারেশনের জন্য একটি শক্তিশালী প্রযুক্তি স্ট্যাক ব্যবহার করে, প্রাকৃতিক ভাষা প্রক্রিয়াকরণ থেকে কম্পিউটেশনাল দক্ষতা পর্যন্ত টুলস এবং লাইব্রেরি ব্যবহার করে। এই স্ট্যাকের অন্তর্ভুক্ত কিছু মূল প্রযুক্তি নিম্নরূপ:

- **পাইথন:** ডেটা ম্যানেজ এবং মেশিন লার্নিং সম্প্রদায়ের প্রধান প্রোগ্রামিং ভাষা। এটি Llama 2-এ কোড বিকাশ এবং কার্যকর করতে ব্যবহৃত হয়, যা পরিষ্কার এবং দক্ষ বাক্য গঠনের অনুমতি দেয়।
- **PyTorch:** একটি বহুল ব্যবহৃত গভীর শিক্ষার লাইব্রেরি। Llama 2 PyTorch ব্যবহার করে ল্যান্ডস্কেপ মডেল তৈরি এবং প্রশিক্ষিত করতে, সেইসাথে নিউরাল নেটওয়ার্ক এবং টেনসরগুলিতে অপারেশন করতে (Paszke, A. et al., 2019)।
- **Conda:** একটি প্যাকেজ এবং পরিবেশ ব্যবস্থাপনা সিস্টেম যা নির্দিষ্ট প্রকল্পের জন্য বিচ্ছিন্ন পরিবেশ তৈরি করা সম্ভব করে। লামা 2-এ, কনডা ব্যবহার করা লাইব্রেরিগুলির নির্ভরতা এবং সংস্করণগুলি পরিচালনা করা সম্ভব করে তোলে, উন্নয়ন পরিবেশে ধারাবাহিকতা নিশ্চিত করে (চ্যাপলিন, জে.আর., এট আল।, 2012)।
- **Google/SentencePiece:** SentencePiece হল একাধিক ভাষায় টোকেনাইজেশন এবং পার্শ্ব বিভাজনের জন্য একটি লাইব্রেরি। Llama 2 টেক্সট সেগমেন্টেশন টাস্কের জন্য Google/SentencePiece নিয়োগ করে, যা ভাষা প্রক্রিয়াকরণ এবং মডেলিংয়ের জন্য অপরিহার্য (Kudo, T., et al., 2018)।
- **ফেয়ারস্কেল:** একটি লাইব্রেরি যা মাল্টি-জিপিইউ পরিবেশে পাইটর্চের স্কেলেবিলিটি এবং কর্মক্ষমতা উন্নত করে। হার্ডওয়্যার রিসোর্স ইউটিলাইজেশন অপ্টিমাইজ করার জন্য লামা 2 ফেয়ারস্কেল ব্যবহার করে, যা দক্ষ অপারেশনের জন্য গুরুত্বপূর্ণ (Miao, X, et al., 2022)।

এই প্রযুক্তির স্ট্যাক Llama 2 কে প্রাকৃতিক ভাষা প্রক্রিয়াকরণ এবং প্রজন্মের সাথে সম্পর্কিত বিস্তৃত কাজগুলিকে মোকাবেলায় উচ্চ স্তরের কর্মক্ষমতা, দক্ষতা এবং কার্যকারিতা অর্জন করতে সক্ষম করে।

প্রশিক্ষণ হার্ডওয়্যার: Llama 2 মডেলটি Meta's Research SuperCluster (RSC) এর পাশাপাশি অভ্যন্তরীণ উৎপাদন ক্লাস্টারে প্রাক-প্রশিক্ষণ সঞ্চালন করে। উভয় ক্লাস্টারই NVIDIA A100 দ্বারা চালিত, যা সর্বাপেক্ষা উন্নত ডেটা সেন্টারে কৃত্রিম বুদ্ধিমত্তা, ডেটা অ্যানালাইটিক্স এবং উচ্চ-পারফরম্যান্স কম্পিউটিং (HPC) শক্তি প্রদান করে বোর্ড জুড়ে একটি অতুলনীয় স্বরণ প্রদান করে। A100, NVIDIA Ampere আর্কিটেকচার দ্বারা চালিত, NVIDIA-এর ডেটা সেন্টার প্ল্যাটফর্মের প্রাণকেন্দ্র হিসেবে দাঁড়িয়ে আছে। পূর্ববর্তী প্রজন্মের তুলনায় 20 গুণ পর্যন্ত পারফরম্যান্সের সাথে, A100-এর বহুমুখিতা রয়েছে সাতটি GPU দৃষ্টান্তে বিভক্ত করার, পরিবর্তনশীল চাহিদার সাথে গতিশীলভাবে মানিয়ে নেওয়া (Gálvez Vallejo, J. L., et al 2023)। এই তুলনায়, এই দুটি ক্লাস্টারের মধ্যে দুটি মৌলিক পার্থক্য দাঁড়িয়েছে। প্রথমত, উপলব্ধ আন্তঃসংযোগের ধরন: RSC NVIDIA কোয়ান্টাম ইনফিনিব্যান্ড প্রযুক্তি ব্যবহার করে, যখন উৎপাদন ক্লাস্টারটি সাধারণ-উদ্দেশ্য ইথারনেট সুইচগুলির উপর ভিত্তি করে একটি RoCE (RDMA ওভার কনভার্জড ইথারনেট) সমাধান দিয়ে সজ্জিত। উভয় সমাধান 200 Gbps এন্ডপয়েন্ট ইন্টারকানেকশন সক্ষম করে।

দ্বিতীয় পার্থক্যটি GPU প্রতি পাওয়ার খরচের সীমার মধ্যে রয়েছে: যদিও RSC-এর সীমা 400W, আমাদের প্রোডাকশন ক্লাস্টার GPU প্রতি 350W সীমার সাথে কাজ করে। এই দ্বি-গুচ্ছ কনফিগারেশনটি বৃহৎ-স্কেল প্রশিক্ষণের জন্য এই বিভিন্ন আন্তঃসংযোগ প্রকারের উপযুক্ততার উপর অর্থপূর্ণ তুলনা সক্ষম করেছে। প্রাসঙ্গিকভাবে, RoCE পদ্ধতি, একটি আরও সাম্রয়ী মূল্যের বাণিজ্যিক আন্তঃসংযোগ সমাধান হিসাবে, দাঁড়িয়েছে (Touvron, H. et al., 2023)।

কার্বন পদচিহ্ন

মেটা এবং লামা 2 এর মধ্যে একটি অসামান্য পার্থক্য হল কার্বন পদচিহ্নের সূচকগুলিকে প্রভাবিত করার ক্ষমতা। কর্পোরেট পরিবেশগত দায়িত্ব শক্তি খরচ একটি অধ্যয়ন নেতৃত্বে

GPU দ্বারা উত্পন্ন প্রতি ঘন্টা. প্রাক-প্রশিক্ষণ পর্বের সময়, নিম্নলিখিত খরচ এবং কার্বন পদচিহ্নের ডেটা প্রাপ্ত হয়েছিল (Touvron, H. et al., 2023)। সারণী 4 প্রাক-প্রশিক্ষণের সময় নির্গমন দেখায়:

সারণী 4. প্রাক-প্রশিক্ষণের সময় নির্গমন

মডেল	সময় (GPU ঘন্টা)	বিদ্যুৎ খরচ (W)	কার্বন নির্গত (tCO ₂ eq)
7B	184320	400	31.22
13B	368640	400	৬২.৪৪
34B	1038336	350	153.90
70B	1720320	400	291.42
মোট	3311616		539.00

মডেল মূল্যায়ন

Llama 2 মডেল নিম্নলিখিত বিষয়গুলিতে ভাল পারফর্ম করে:

- কোড প্রজন্ম: Llama 2 ব্যবহার করে পাইথনে একটি কোড পরীক্ষার বৈধতা করা হয়েছিল।

Prompt: Calculate the factorial of a given number using Python.

Response:

```
def factorial(n):
    if n == 0 or n == 1:
        return 1
    else:
        return n * factorial(n - 1)

number = 5
result = factorial(number)
print(f"The factorial of {number} is {result}")
```

চিত্র ৪. কোড তৈরির উদাহরণ।

- যুক্তি এবং সাধারণ জ্ঞান: Llama 2 ব্যবহার করে যুক্তি অনুশীলনের একটি পরীক্ষা যাচাই করা হয়েছিল।

Prompt: If you have three dogs, four cats, and one duck, how many total legs do all the animals have?

Response: You can calculate the total number of legs by adding up the legs of each type of animal:

Three dogs have $3 * 4 = 12$ legs. Four cats have $4 * 4 = 16$ legs. One duck has 2 legs. Adding up all the legs: 12 (dogs) + 16 (cats) + 2 (duck) = 30 legs in total.

চিত্র ৭. যুক্তির উদাহরণ।

- বিশ্বব্যাপী জ্ঞান: Llama 2 ব্যবহার করে বৈশ্বিক স্তরে সাধারণ সংস্কৃতি বিষয়ের উপর একটি বৈধতা পরীক্ষা করা হয়েছিল।

Prompt: When did World War II end?

Response: World War II ended on September 2, 1945, when Japan formally surrendered to the Allied Powers, marking the official conclusion of the global conflict.

চিত্র 10. বিশ্ব সংস্কৃতির উদাহরণ।

- পড়া বোঝা: Llama 2 ব্যবহার করে বোঝার বিষয় পড়ার উপর একটি বৈধতা পরীক্ষা করা হয়েছিল।

Prompt: "Create a Title from the Following Text: In the beginning, God created the heavens and the earth. 2. The earth was chaos and confusion and darkness above the abyss, and a wind of God fluttered over the waters."

Response: Creation, Chaos, and Divine Presence.

চিত্র 11। পড়া বোঝার উদাহরণ।

- গণিত: গণনার ক্ষমতার জন্য Llama 2 ব্যবহার করে একটি বৈধতা পরীক্ষা করা হয়েছিল।

Prompt: You have a circle with a radius of length 7 units. What is the area of the circle?

Response: To calculate the area of a circle, we can use the formula for the area of a circle, which is expressed as:

$$Area = \pi \times Radius^2$$

Given that the radius is 7 units, we can substitute that value into the formula:

$$Area = \pi \times 7^2 \quad Area = \pi \times 49$$

Using an approximate value of π as 3.14159, we can calculate the area:

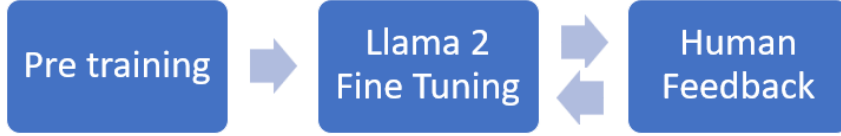
$$Area \approx 3.14159 \times 49 \quad Area \approx 153.93804$$

Therefore, the area of the circle with a radius of length 7 units is approximately 153.93804 square units.

চিত্র 12. গণনার ক্ষমতার উদাহরণ।

6.3. লামা 2 প্রশিক্ষণ আর্কিটেকচার এবং প্রক্রিয়া

লামা 2 প্রশিক্ষণ প্রক্রিয়া হল আর্কিটেকচারের একটি অভিনব সমন্বয়, যেমন ট্রান্সফরমার-ভিত্তিক প্রাক-প্রশিক্ষণ এবং হিউম্যান ফিডব্যাক (RLFH) ভিত্তিক টিউনিং সহ রিইনফোর্সমেন্ট লার্নিং, যেখানে Llama 2 LLM-এর নির্ভুলতা, নির্ভুলতা এবং নিরাপত্তা সর্বোত্তম। Llama 2 (Touvron, H et al., 2023) নির্মাণের জন্য মেটা দ্বারা ব্যবহৃত পদ্ধতিটি নিম্নরূপ। চিত্র 13 লামা 2 মডেলের প্রশিক্ষণ প্রক্রিয়া দেখায়।



চিত্র 13. প্রশিক্ষণ প্রক্রিয়া।

প্রাক-প্রশিক্ষণ

Llama 2 সর্বজনীনভাবে উপলব্ধ ডেটা উত্স থেকে প্রাক-প্রশিক্ষণ সঞ্চালন করে, মেটার পণ্য বা পরিষেবাগুলি থেকে উদ্ভূত যে কোনও ডেটা বাদ দিয়ে। ডেটা নির্বাচন প্রক্রিয়ার সময়, ওয়েব ডেটা এবং বিশেষ ডেটাবেসের মতো উত্সগুলি ব্যবহার করা হয়েছিল। মডেলটি একটি ডেটাসেট ব্যবহার করে ব্যাপক প্রশিক্ষণ গ্রহণ করেছে যা দুই ট্রিলিয়ন টোকেন সমন্বিত করে এবং তার পূর্বসূরি লামা 1 (Jiao, F., et al., 2023) এর দ্বিগুণ প্রসঙ্গ দৈর্ঘ্য প্রদর্শন করে। সম্ভাব্য ভুল তথ্য এবং পক্ষপাতমূলক সমস্যাগুলি প্রশমিত করার সময় গুণের উন্নতির জন্য সর্বাধিক বাস্তব উত্স থেকে নমুনা নেওয়ার উপর ইচ্ছাকৃত জোর দিয়ে এই নকশা পছন্দটি কর্মক্ষমতা এবং গণনামূলক ব্যয়ের মধ্যে ভারসাম্য বজায় রাখে (Touvron, H et al., 2023)।

RMSNorm (Sennrich, R et al., 2019) এবং SwiGLU অ্যাক্টিভেশন ফাংশন (Shazeer, N. et al., 2020) ব্যবহার করে মডেলটিতে একটি আদর্শ ট্রান্সফরমার আর্কিটেকচার (Vaswani et al., 2017) রয়েছে। উপরন্তু, এটি ঘূর্ণনগত অবস্থানগত এম্বেডিং (RoPE) (Touvron, H. et al., 2023) সংহত করে।

Llama 1 এবং Llama 2 এর মধ্যে মূল পার্থক্যগুলি বর্ধিত প্রসঙ্গ দৈর্ঘ্য এবং গ্রুপড কোয়েরি অ্যাটেনশন (GQA) গ্রহণের মধ্যে রয়েছে। এই স্থাপত্য পরিবর্তনগুলি ভাষা তৈরির কাজের সময় দীর্ঘ প্রাসঙ্গিক তথ্য পরিচালনা করার ক্ষমতা বৃদ্ধিতে অবদান রাখে (Touvron, H. et al., 2023)। নীচে, টেবিল 5 মডেলের বৈশিষ্ট্য দেখায়।

সারণী 5. প্রাক-প্রশিক্ষণ বৈশিষ্ট্য।

	প্রশিক্ষণ ডেটা	পরামিতি	প্রসঙ্গ দৈর্ঘ্য	টোকেন
কল করুন 1	ডেটাসেট (Touvron, H et al., 2023)।	7B	2k	1.0T
		13B	2k	1.0T
		33B	2k	1.4T
		65B	2k	1.4T
কল করুন 2	অনলাইন ডেটা	7B	4k	2.0T
		13B	4k	2.0T
		34B	4k	2.0T
		70B	4k	2.0T

কল করুন 2 ফাইন-টিউনিং:

কল করুন 2 প্রাথমিকভাবে সর্বজনীনভাবে উপলব্ধ অনলাইন ডেটা ব্যবহার করে প্রশিক্ষণ দেওয়া হয়েছিল। Llama-2- চ্যাটের একটি প্রাথমিক সংস্করণ তারপর তত্ত্বাবধানে থাকা টিউনিং ব্যবহার করে তৈরি করা হয়েছিল। পরবর্তীকালে, মানুষের প্রতিক্রিয়া (RLHF) এর উপর ভিত্তি করে রিইনফোর্সড লার্নিং নিয়োগের মাধ্যমে Llama-2-চ্যাটকে পুনরাবৃত্তিমূলকভাবে পরিমার্জিত করা হয়েছিল, যার মধ্যে রয়েছে প্রত্যাখ্যান স্যাম্পলিং এবং প্রক্রিয়ামাল পলিসি অস্টিমাইজেশান (PPO) (Touvron, H. et al., 2023) এর মতো কৌশলগুলি।

যদিও এই ধরনের তথ্য প্রদানকারী অসংখ্য উৎস রয়েছে, তবে তাদের সীমিত বৈচিত্র্য এবং গুণমানের কারণে উচ্চ মানের সক্রিয় এবং প্যাসিভ ভয়েস (OFV) উদাহরণ সংগ্রহকে অগ্রাধিকার দেওয়া হয়েছে, যা উল্লেখযোগ্য উন্নতির দিকে পরিচালিত করেছে।

হিউম্যান ফিডব্যাক সহ রিইনফোর্সমেন্ট লার্নিং (RLHF) একটি মডেল প্রশিক্ষণ পদ্ধতির প্রতিনিধিত্ব করে যার লক্ষ্য মানুষের পছন্দ এবং নির্দেশাবলীর সাথে একটি টিউন করা ভাষা মডেলের আচরণকে আরও সারিবদ্ধ করা। এটি অর্জনের জন্য, দুটি মডেল-উত্পন্ন ফলাফলের মধ্যে একটি নির্বাচনের মাধ্যমে মানুষের পছন্দের ডেটা সংগ্রহ করা হয়, যা একটি পুরস্কার মডেলের প্রশিক্ষণে অবদান রাখে যা পছন্দের পছন্দগুলিকে স্বয়ংক্রিয় করে দেয় (মারোটো-গোমেজ, এম. এট আল।, 2023)।

পুরস্কার মডেলের অপরিহার্য ভূমিকা মানব প্রতিক্রিয়া (RLHF) এর সাথে শক্তিবৃদ্ধি শেখার প্রসঙ্গে উদ্ভাসিত হয়। এই প্রক্রিয়ায়, পুরস্কার মডেল একটি মডেল-উত্পন্ন প্রতিক্রিয়া এবং এর সংশ্লিষ্ট কিউ মূল্যায়ন করে, যার ফলে একটি সংখ্যাসূচক স্কার তৈরি হয় যা উপযোগিতা এবং নিরাপত্তার পরিপ্রেক্ষিতে গুণমানকে প্রতিফলিত করে। এই প্রতিক্রিয়া স্কারগুলিকে পুরস্কার হিসাবে ব্যবহার করার মাধ্যমে, RLHF প্রক্রিয়ার লক্ষ্য হল লামা 2 চ্যাটকে মানুষের পছন্দগুলির সাথে সারিবদ্ধ করার জন্য অস্টিমাইজ করা, ইউটিলিটি এবং নিরাপত্তা উভয়ের উন্নতি করা (Touvron, H. et al., 2023)।

6.4. এটা কিভাবে ইন্সটল করবেন?

একজন ব্যবহারকারীকে প্রাক-প্রশিক্ষিত মডেলগুলিতে অ্যাক্সেস পাওয়ার জন্য, তাদের মেটার অফিসিয়াল ওয়েবসাইটের মাধ্যমে অনুমতির অনুরোধ করতে হবে (<https://ai.meta.com/llama/>) এবং নির্দিষ্ট শর্তাবলীতে সম্মত হন। একবার তাদের অনুরোধ অনুমোদিত হলে, তারা একটি ইমেল পায় যাতে একটি অনন্য, ব্যক্তিগতকৃত লিঙ্ক রয়েছে। এই লিঙ্কটি ব্যবহারকারীকে মডেলগুলি ডাউনলোড করতে দেয়।

চিত্র 14 অফিসিয়াল META Llama 2 ওয়েবসাইটে নিবন্ধনের জন্য হোম স্ক্রীনকে চিত্রিত করে।

Request access to the next version of Llama

First Name

Last Name

Email

Country

Organization / Affiliation

Select the models you would like access to:

Llama 2 & Llama Chat

Code Llama

চিত্র 14. লামা 2 ইন্সটলেশন প্রক্রিয়া।

এটি একটি পাইথন প্রকল্প তৈরি করার সুপারিশ করা হয়, যাতে ব্যবহারকারী মডেলগুলি হোস্ট করে এমন সংগ্রহস্থল ক্লোন করতে GitHub-এ দেওয়া URL ব্যবহার করতে পারে। এটি প্রকল্পে ব্যবহারের জন্য তাদের সিস্টেমে মডেলগুলি উপলব্ধ করা সহজ করে তোলে। উপরের সমস্ত এই পদক্ষেপগুলি অনুসরণ করে অর্জন করা যেতে পারে:

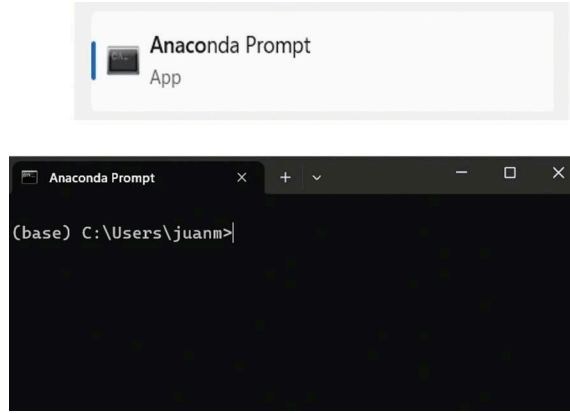
- অফিসিয়াল মেটা ওয়েবসাইটের মাধ্যমে অ্যাক্সেসের অনুরোধ করুন এবং প্রয়োজ্য শর্তাবলীতে সম্মত হন।
- একবার অনুরোধটি অনুমোদিত হলে, ব্যবহারকারী মডেলগুলি ডাউনলোড করার জন্য একটি অনন্য, কাস্টমাইজড লিঙ্ক সহ একটি ইমেল পাবেন।
- একটি নতুন পাইথন প্রকল্পে, গিটহাব ইউআরএলটি মডেলগুলি ধারণকারী সংগ্রহস্থল ক্লোন করতে ব্যবহার করা যেতে পারে।

- "download.sh" স্ক্রিপ্টটি চালান, যেখানে ব্যবহারকারীকে অনন্য, কাস্টম URL লিখতে হবে এবং ডাউনলোড করার জন্য মডেলগুলি নির্বাচন করতে হবে। ডাউনলোডের জন্য উপলব্ধ বিকল্পগুলির মধ্যে রয়েছে 7B, 13B, 70B, চ্যাট 7B, চ্যাট 13B এবং চ্যাট 70B।

দ্রষ্টব্য: এই মডেলগুলি পরিচালনা করতে সক্ষম হওয়ার জন্য ব্যবহারকারীর পর্যাপ্ত স্টোরেজ, প্রক্রিয়াকরণ শক্তি, একটি GPU এবং পর্যাপ্ত RAM থাকা বাঞ্ছনীয়।

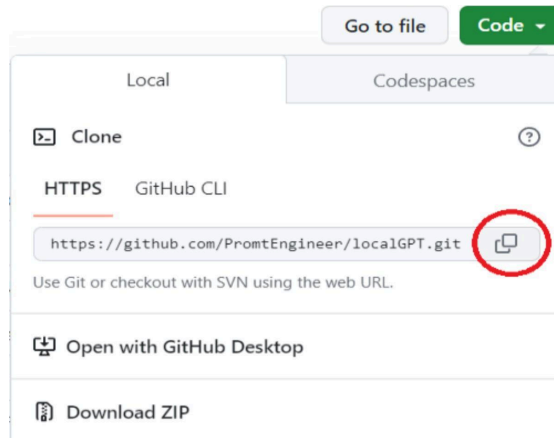
বায়োইনফরমেটিক্স, ইন্টেলিজেন্ট ইনফরমেশন সিস্টেমস অ্যান্ড এডুকেশনাল টেকনোলজি (BISITE) গবেষণা গ্রুপ স্থানীয়ভাবে লামা 2 মডেলটি ইনস্টল করতে সক্ষম হয়েছে। মডেলটি কীভাবে ইনস্টল করবেন এবং এটির কাস্টমাইজড ব্যবহার করবেন সে সম্পর্কে এখানে একটি ধাপে ধাপে নির্দেশিকা রয়েছে:

1. বিনামূল্যে অ্যানাকোন্ডা বিতরণ ইনস্টল করুন এবং অ্যানাকোন্ডা প্রম্পট চালান, চিত্র 15 দেখুন:



চিত্র 15. অ্যানাকোন্ডা প্রম্পট চালান।

2. সংগ্রহস্থল অ্যাক্সেস করুন <https://github.com/PromptEngineer/localGPT>, যেখানে ক্রেডিট দেওয়া হয় (PromptEngineer, Github) এবং লিঙ্কটি অনুলিপি করুন: <https://github.com/PromptEngineer/localGPT.git>, চিত্র 16 দেখুন।



চিত্র 16. GitHub সংগ্রহস্থল কপি করুন।

3. আপনার কম্পিউটারে আপনার পছন্দের জায়গায় একটি ফোল্ডার তৈরি করুন এবং একটি GitHub ক্লোন তৈরি করুন, চিত্র 17 দেখুন।

```
Anaconda Prompt
(base) C:\Users\juanm\Documents>cd DemoLlama
(base) C:\Users\juanm\Documents\DemoLlama>git clone https://github.com/PromptEngineer/LocalGPT.git local_llama2
git clone https://github.com/PromptEngineer/LocalGPT.git local_llama2
```

চিত্র 17. স্থানীয়ভাবে GitHub ক্লোন করুন।

4. একটি Conda পরিবেশ তৈরি করতে নিম্নলিখিত কমান্ডটি সম্পাদন করুন, চিত্র 18 দেখুন:

```
Anaconda Prompt
(base) C:\Users\juanm>cd Documents
(base) C:\Users\juanm\Documents>cd Demollama
(base) C:\Users\juanm\Documents\DemoLlama>cd local_llama2
(base) C:\Users\juanm\Documents\DemoLlama\local_llama2>conda create -n local_llama2 python= 3.11.4
```

চিত্র 18. প্রয়োজনীয়তা ইনস্টল করার জন্য একটি কনডা পরিবেশ তৈরি করুন।

5. স্থানীয় Llama 2 মডেল এম্বেড করতে সক্ষম হতে প্রশিক্ষণের জন্য ফাইল বা নথি লিখুন এবং উপরের সমস্তগুলি। এই ফাইলটি অবশ্যই SOURCE_DOCUMENTS ফোল্ডারে রাখতে হবে, চিত্র 19 দেখুন।

Name	Date modified	Type	Size
.github	8/2/2023 4:55 AM	File folder	
__pycache__	8/2/2023 6:35 AM	File folder	
DB	8/29/2023 7:58 AM	File folder	
localGPTUI	8/2/2023 4:55 AM	File folder	
SOURCE_DOCUMENTS	8/3/2023 3:03 AM	File folder	
.editorconfig	8/2/2023 4:55 AM	Editor Config Source File	1 KB
.flake8	8/2/2023 4:55 AM	FLAKE8 File	1 KB
.gitignore	8/2/2023 4:55 AM	Git Ignore Source File	4 KB
.pre-commit-config.yaml	8/2/2023 4:55 AM	Yaml Source File	2 KB
.pyup.yml	8/2/2023 4:55 AM	Yaml Source File	1 KB
ACKNOWLEDGEMENT.md	8/2/2023 4:55 AM	Markdown Source File	1 KB
constants.py	8/2/2023 4:55 AM	Python File	2 KB
CONTRIBUTING.md	8/2/2023 4:55 AM	Markdown Source File	3 KB
ingest.py	8/2/2023 4:55 AM	Python File	6 KB
LICENSE	8/2/2023 4:55 AM	File	12 KB
pyproject.toml	8/2/2023 4:55 AM	Toml Source File	1 KB
README.md	8/2/2023 4:55 AM	Markdown Source File	13 KB
requirements.txt	8/2/2023 4:55 AM	Text Document	1 KB
run_localGPT.py	8/2/2023 4:55 AM	Python File	10 KB

চিত্র 19. প্রশিক্ষণের জন্য ফাইলটি প্রবেশ করান।

6. প্রয়োজনীয় ফাইল অনুযায়ী প্যাকেজ ইনস্টল করুন: কিছু প্যাকেজ হল: pytorch, langchain, chromadb, sentence-transformers, huggingface hub (Requirements.txt ফাইল দেখুন), চিত্র 20 দেখুন।

```
Anaconda Prompt
(base) C:\Users\juanm>cd Documents/DemoLlama/local_llama2
(base) C:\Users\juanm\Documents\DemoLlama\local_llama2>python -n pip install -r requirements.txt
```

চিত্র 20. লামা 2 প্যাকেজ এবং প্রয়োজনীয়তা ইনস্টল করা।

7. প্রশিক্ষিত করা ফাইল থেকে ডেটা অবশ্যই মডেলে প্রবেশ করাতে হবে এবং নিম্নলিখিত নির্দেশাবলী অনুসরণ করতে হবে, চিত্র 21 দেখুন:

```
Anaconda Prompt
(base) C:\Users\juanm>cd Documents/DemoLlama/local_llama2
(base) C:\Users\juanm\Documents\DemoLlama\local_llama2>python ingest.py --device_type cpu
```

চিত্র 21. লামা 2 মডেলে ডেটা গ্রহণ করা।

8. নিম্নলিখিত কমান্ডটি মডেলটি কার্যকর করতে এবং একটি অনুরোধ বা প্রশ্ন করতে ব্যবহৃত হয়, দেখুন চিত্র 22, 23:

```
Anaconda Prompt
(base) C:\Users\juanm>cd Documents/DemoLlama/local_llama2
(base) C:\Users\juanm\Documents\DemoLlama\local_llama2>python run_localGPT.py --device_type cpu
```

চিত্র 22. লামা 2 মডেলে একটি অনুরোধ করা।

```
> Question:
What are the requirements for the granting of Spanish nationality, which article governs me?
> Answer:
You can find the requirements for the granting of Spanish nationality in Article 24 and Article 21 of the Spanish Nationality Law. According to these articles, there are two ways to obtain Spanish nationality: by birth or through adoption, and the requirement is to have a clean criminal record and to swear loyalty to the King and obedience to the Constitution and laws.
In addition, Article 24 mentions that the granting of nationality is discretionary and can be done through Real Decree, while Article 21 states that nationality can also be obtained by residency in Spain under certain conditions.
Enter a query: |
```

চিত্র 23. বৈধতা পরীক্ষা।

6.5. কেস ব্যবহার করুন

জেনারেটিভ এআই-এর বিকেন্দ্রীকরণের সাথে একাধিক অ্যাপ্লিকেশন আকার নিতে শুরু করেছে এবং সামাজিক নেটওয়ার্কগুলিতে অপমানজনক এবং যৌন ভাষায়ুক্ত ব্যক্তিদের সুরক্ষা এবং সনাক্তকরণের উপর দৃষ্টি নিবদ্ধ করা গবেষণা, যেখানে জনসংখ্যাকে সুরক্ষিত করতে হবে তারা হল শিশু এবং কিশোর (নগুয়েন, টি. টি, এবং আল।, 2023)। স্বাস্থ্য খাতে, জেনারেটিভ এআই-এর মাধ্যমে চিকিৎসা সহায়কদের দৃষ্টিকোণ থেকে যে অগ্রগতি উপস্থাপন করা হচ্ছে তা স্পষ্ট। এটি Li, Y., et al., 2023-এর ক্ষেত্রে, যারা একটি প্রতিষ্ঠানের চিকিত্সকদের রোগ নির্ণয় এবং অভিজ্ঞতার উপর ভিত্তি করে একটি বিশেষ মেডিকেল চ্যাট প্রয়োগ করেছেন। ChatDoctor Llama 2 কে একটি কাঠামো হিসাবে ব্যবহার করে এবং রোগীদের চাহিদা, চিকিৎসা মতামত যা প্রতিদিন জারি করা হয়, এইভাবে মডেলের যথার্থতা উন্নত করে এবং চিকিৎসা প্রতিক্রিয়াগুলিতে নির্ভরযোগ্যতা তৈরি করে। আর্থিক প্রেক্ষাপটে, Llama 2-এর উপর ভিত্তি করে একটি LLM টেক্সট থেকে সবচেয়ে গুরুত্বপূর্ণ সূচকগুলি বের করতে এবং সংশ্লিষ্ট বিশ্লেষণগুলি সম্পাদন করতে আর্থিক খবরের বিশ্লেষণের জন্য পরিমার্জিত হয় (Pavlyshenko, 2023)।

7. উপসংহার

ডিজিটাল কৌশল ব্যবহার করে জেনারেটিভ মডেলগুলি একাধিক ডোমেনে, বিশেষ করে উচ্চ-বিশ্বস্ত পাঠ্য, অডিও এবং ভিডিও উৎপাদনে সম্ভাবনা উন্মুক্ত করেছে। তাদের সম্ভাব্য ফ্রেমওয়ার্কগুলি ডেটার মধ্যে জটিল এবং সূক্ষ্ম সম্পর্কগুলিকে ক্যাপচার করতে পারে, অত্যন্ত বাস্তবসম্মত ফলাফল তৈরি করে (Rombach et al., 2022; Huang et al., 2023)। টেক্সচুয়াল ডেটার জন্য, এটি শব্দার্থগত সমন্বয় থেকে সিনট্যাকটিক পরিশীলিততা পর্যন্ত সবকিছুকে অন্তর্ভুক্ত করে। অডিওতে, এটি স্পষ্ট বক্তৃতা সংশ্লেষণ এবং স্পিচ-টু-টেক্সট রূপান্তর অ্যাপ্লিকেশনের দিকে নিয়ে যায় (উইলিয়ামস, 2023)। ভিডিওতে, অ্যালগরিদমগুলি উচ্চ-সংজ্ঞা, অস্থায়ীভাবে সুসংগত ভিডিও ক্লিপ তৈরি করে (টেট, 2023)।

OpenAI দ্বারা বিকশিত বৃহৎ ভাষা মডেলের (LLMs) সিরিজ প্রাকৃতিক ভাষা প্রক্রিয়াকরণের ক্ষেত্রে বিপ্লব ঘটিয়েছে। GPT-4 এবং ChatGPT-এর মতো মডেলগুলি সুসংগত পাঠ্য তৈরি করার, প্রশ্ন বুঝতে এবং বিভিন্ন কাজের সাথে খাপ খাইয়ে নেওয়ার একটি চিত্তাকর্ষক ক্ষমতা প্রদর্শন করেছে। যাইহোক, এই মডেলগুলি অগ্রসর হওয়ার সাথে সাথে তাদের ব্যবহারের সাথে যুক্ত নৈতিক এবং প্রযুক্তিগত চ্যালেঞ্জগুলিও করে।

পক্ষপাত প্রশমিত করতে, ব্যবহারকারীর গোপনীয়তা রক্ষা করতে এবং ডেটা সুরক্ষা নিশ্চিত করতে কার্যকর কৌশলগুলি গবেষণা চালিয়ে যাওয়া এবং বিকাশ করা গুরুত্বপূর্ণ। এই মডেলগুলি দৈনন্দিন জীবনে আরও একীভূত হওয়ার সাথে সাথে দায়িত্বশীল ব্যবহারের জন্য স্পষ্ট প্রবিধান এবং নির্দেশিকা প্রয়োজন। জেনারেটিভ এআই-এর ক্ষেত্রে উদ্যোগ নেওয়া বিভিন্ন ডোমেনে এই মডেলগুলির প্রয়োগের উল্লেখযোগ্য সম্ভাবনা প্রকাশ করে, প্রোগ্রামিং থেকে একাডেমিক প্রকাশনা পর্যন্ত। যাইহোক, সমালোচনা করা এবং তাদের সীমাবদ্ধতা সম্পর্কে সচেতন হওয়া অপরিহার্য। শেষ পর্যন্ত, জেনারেটিভ এআই গবেষণার লক্ষ্য হওয়া উচিত এমন সরঞ্জাম তৈরি করা যা মানবতার জন্য দরকারী, নিরাপদ এবং উপকারী। GPT-4 এবং ChatGPT-এর মতো মডেলগুলির ক্ষেত্রে উল্লেখযোগ্য অগ্রগতি সত্ত্বেও, এই লক্ষ্য অর্জনের জন্য অনেক কাজ করা বাকি রয়েছে।

মেটা এআই দ্বারা ডেভেলপ করা লামা 2 হল একটি ওপেন-সোর্স লার্জ ল্যাঙ্গুয়েজ মডেল (LLM) যা একটি স্থানীয় ডিভাইসে নথির সাথে মিথস্ক্রিয়াকে সহজতর করে। ব্যবহারকারীর গোপনীয়তা নিশ্চিত করে স্থানীয়ভাবে কাজ করার ক্ষমতার মধ্যে এর শক্তি নিহিত। এর অত্যাধুনিক প্রশিক্ষণ প্রক্রিয়ার মাধ্যমে, Llama 2 বিভিন্ন বৈশিষ্ট্য পরীক্ষায় উচ্চ কার্যকারিতা প্রদর্শন করেছে, যুক্তি, কোডিং এবং পড়ার বোধগম্যতার মতো ক্ষেত্রগুলিতে অসাধারণ। ইনস্টলেশনের জন্য নির্দিষ্ট সিস্টেমের প্রয়োজনীয়তা এবং প্রযুক্তিগত দক্ষতার প্রয়োজন হওয়া সত্ত্বেও, Llama 2 এটিকে বিভিন্ন Hugging Face LLM মডেলের সাথে কাজ করার অনুমতি দিয়ে দুর্দান্ত বহুমুখিতা প্রদান করে। সামাজিক নেটওয়ার্কগুলিতে শিশু সুরক্ষা থেকে শুরু করে চিকিৎসা সহকারী এবং আর্থিক বিশ্লেষণ পর্যন্ত অ্যাপ্লিকেশনগুলির সাথে, Llama 2 প্রাকৃতিক ভাষা-ভিত্তিক মিথস্ক্রিয়া জন্য নতুন সুযোগ উন্মুক্ত করে এবং LLM ক্ষেত্রে একটি উল্লেখযোগ্য অগ্রগতির প্রতিনিধিত্ব করে।

৪. তথ্যসূত্র

- আব্দুল্লাহ, এম., মাদাইন, এ., এবং জারারওয়েহ, ওয়াই., 2022, নভেম্বর। ChatGPT: মৌলিক, অ্যাপ্লিকেশন এবং সামাজিক প্রভাব। ইন 2022 সামাজিক নেটওয়ার্ক বিশ্লেষণ, ব্যবস্থাপনা এবং নিরাপত্তা (SNAMS) এর উপর নবম আন্তর্জাতিক সম্মেলন (পৃ. 1-8)। আইইইই।
<https://doi.org/10.1109/SNAMS58071.2022.10062688>
- Abraham, A., Corchado, E., & Corchado, J. M., 2009. হাইব্রিড লার্নিং মেশিন। নিউরোকম্পিউটিং: একটি আন্তর্জাতিক জার্নাল, 72(13-15), 2729-2730। <https://doi.org/10.1016/j.neucom.2009.02.017>
- অ্যাডামস, এল.সি., বুশ, এফ., টুইন, ডি., মাকোস্কি, এম.আর., আর্টস, এইচ.জে., এবং রেসেম, কে.কে., 2023. DALL-E 2 রেডিওলজি সম্পর্কে কী জানেন? মেডিকেল ইন্টারনেট রিসার্চ জার্নাল, 25, e431101। <https://doi.org/10.1109/SNAMS58071.2022.10062688>
- Alizadehsani, Z., Ghaemi, H., Shahraki, A., Gonzalez-Briones, A., & Corchado, J. M., 2023. DCServCG: গভীর শিক্ষা ব্যবহার করে একটি ডেটা-কেন্দ্রিক পরিষেবা কোড তৈরি। কৃত্রিম বুদ্ধিমত্তার ইঞ্জিনিয়ারিং অ্যাপ্লিকেশন, 123, 106304। <https://doi.org/10.1016/j.engappai.2023.106304>
- বা, জে.এল., কিরোস, জে.আর., এবং হিন্টন, জি.ই. *arXiv প্রিপ্রিন্ট arXiv:1607.064501*
- Bahdanau, D., Cho, K., & Bengio, Y., 2014. যৌথভাবে সারিবদ্ধ এবং অনুবাদ শেখার মাধ্যমে নিউরাল মেশিন অনুবাদ। *arXiv প্রিপ্রিন্ট arXiv:1409.04731*
- Baidoo-Anu, D., & Owusu Ansah, L., 2023, জানুয়ারী 25. জেনারেটিভ কৃত্রিম বুদ্ধিমত্তার যুগে শিক্ষা (AI): শিক্ষণ এবং শেখার প্রচারে ChatGPT-এর সম্ভাব্য সুবিধা বোঝা। এ উপলব্ধ এসএসআরএন। <https://doi.org/10.2139/ssrn.4337484>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021)। স্টোকাস্টিক তোতাপাখির বিপদ সম্পর্কে: ভাষার মডেলগুলি কি খুব বড় হতে পারে? ইন *FAccT* এর কার্যধারা। <https://doi.org/10.1145/3442188.3445922>
- Bengio, Y., 2009. AI এর জন্য গভীর আর্কিটেকচার শেখা। ভিত্তি এবং প্রবণতা@ মেশিন লার্নিং, 2(1), 1-127। <https://doi.org/10.1561/22000000006>
- Brown, T. B., et al., 2020. ভাষার মডেল হল অল্প সংখ্যক শিক্ষার্থী। *arXiv প্রিপ্রিন্ট arXiv: 2005141651*
- বুবেক, এস., চন্দ্রশেকরন, ভি., এলডান, আর., গেহর্ক, জে., হরভিটজ, ই., কামার, ই., ... এবং ঝাং, ওয়াই., 2023. কৃত্রিম সাধারণ বুদ্ধিমত্তার স্কুলিং: প্রাথমিক পরীক্ষা-নিরীক্ষা gpt-4. *arXiv প্রিপ্রিন্ট arXiv:2303.127121*
- Caliskan, A., Bryson, J. J., & Narayanan, A., 2017. ভাষা কর্পোরা থেকে স্বয়ংক্রিয়ভাবে উদ্ভূত শব্দার্থে মানুষের মত পক্ষপাত রয়েছে। বিজ্ঞান, 356(6334), 183-186। <https://doi.org/10.1126/science.aal4230>
- Castellanos-Gomez, A., 2023. ChatGPT এবং অন্যান্য কৃত্রিম বুদ্ধিমত্তার ভাষা মডেলের সাথে বৈজ্ঞানিক নিবন্ধ লেখার জন্য ভাল অনুশীলন। ন্যানো ম্যানুফ্যাকচারিং, ৩(2), 135-138। <https://doi.org/10.3390/nanomanufacturing3020009>
- Chamoso, P., González-Briones, A., Rivas, A., De La Prieta, F., & Corchado, J. M., 2019। মুদ্রা বিনিময় সামাজিক কম্পিউটিং: জ্ঞান এবং তথ্য সিস্টেম, 61, 733-753। <https://doi.org/10.1007/s10115-018-1289-4>
- Chan, W. H., Mohamad, M. S., Deris, S., Zaki, N., Kasim, S., Omatu, S., Corchado J. M. & Al Ashwal, H., 2016। একটি ওজন স্কিম সহ একটি উন্নত দণ্ডিত সমর্থন ভেক্টর মেশিন ব্যবহার করে তথ্যপূর্ণ জিন এবং পথ সনাক্তকরণ। জীববিজ্ঞান ও চিকিৎসাবিদ্যায় কম্পিউটার, 77, 102-115। <https://doi.org/10.1016/j.compbio.2016.08.004>
- চ্যাং, ওয়াই., ওয়াং, এছ., ওয়াং, জে., উ, ওয়াই., ব্লু, কে., চেন, এইচ., ... এবং জি, এছ., 2023. বৃহৎ ভাষার মডেলগুলির মূল্যায়নের উপর একটি সমীক্ষা। *arXiv প্রিপ্রিন্ট arXiv:2307.031091*

- চ্যাপলিন, জে.আর., হেলার, ভি., ফারলে, এফ.জে.এম., হার্ন, জি.ই., এবং রেইনি, আর.সি.টি., 2012। অ্যানাকোন্ডা পরীক্ষাগার পরীক্ষাগার। রয়্যাল সোসাইটির দার্শনিক লেনদেন *A: গাণিতিক, শারীরিক এবং প্রকৌশল বিজ্ঞান*, 370(1959), 403-424। <https://doi.org/10.1098/rsta.2011.0256>
- চেন, ডি., এট আল।, 2019। BERT: ভাষা বোঝার জন্য গভীর দ্বিমুখী ট্রান্সফরমারের প্রাক-প্রশিক্ষণ। *arXiv প্রিন্ট* *arXiv:1810.04805*
- Corchado J.M., 2023। গ্লোবাল আর্টিফিশিয়াল ইন্টেলিজেন্সের জাগরণ। রয়্যাল একাডেমি অফ মেডিসিন - সালমানকা।
- Corchado, J. M., & Aiken, J., 2002. হাইব্রিড আর্টিফিশিয়াল ইন্টেলিজেন্স মেথডস ইন ওশানোগ্রাফিক ফোরকাস্ট মডেল। সিস্টেম, ম্যান, এবং সাইবারনেটিক্সের উপর *IEEE* লেনদেন, অংশ সি (অ্যাপ্লিকেশন এবং পর্যালোচনা), 32(4), 307-313। <https://doi.org/10.1109/TSMCC.2002.806072>
- Corchado, J. M., Chamoso, P., Hernández, G., Gutierrez, A.S.R., Camacho, A.R., González-Briones, A., ... & Omatu, S., 2021. Deepint. নেট: স্মার্ট অঞ্চলগুলির জন্য একটি দ্রুত স্থাপনার প্ল্যাটফর্ম। সেম্বর, 21(1), 236। <https://doi.org/10.3390/s21010236>
- Corchado, J. M., Díaz, F., Borrajo, L., & Fernández, F., 2000. কৃত্রিম নিউরাল নেটওয়ার্ক। একটি বাস্তব পদ্ধতি। ভিগো বিশ্ববিদ্যালয়ের প্রকাশনা পরিষেবা।
- ডেনি, পি., কুমার, ভি., এবং গিয়াকামান, এন., 2022। কম্পিউটারের সাথে কথোপকথন: প্রাকৃতিক ভাষা ব্যবহার করে *CSI* সমস্যা সমাধানের জন্য প্রম্পট ইঞ্জিনিয়ারিং অন্বেষণ করা। <https://doi.org/10.1145/3545945.3569823>
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K., 2018. বার্ট: ভাষা বোঝার জন্য গভীর দ্বিমুখী ট্রান্সফরমারের প্রাক-প্রশিক্ষণ। *arXiv প্রিন্ট* *arXiv:1810.04805*
- Díaz, F., Fernández-Riverola, F., & Corchado, J. M., 2006. gene-CBR: A Case-based Reasoning Tool for Cancer Diagnosis Using Microarray Data Set. কম্পিউটেশনাল ইন্টেলিজেন্স, 22(3-4), 254-268। <https://doi.org/10.1111/j.1467-8640.2006.00287.x>
- Doshi-Velez, F., Kortz, M., Budish, R., Bavitz, C., Gershman, S., O'Brien, D., ... & Waldo, J., 2017. AI এর অধীনে জবাবদিহিতা আইন: ব্যাখ্যার ভূমিকা। ব্যাখ্যা এবং আইন বিষয়ে বার্কম্যান ক্লেইন সেন্টার ওয়ার্কিং গ্রুপ, 2. <https://doi.org/10.2139/ssrn.3064761>
- Eloundou, T., Manning, S., Mishkin, P., & Rock, D., 2023. Gpts হল gpts: বৃহৎ ভাষার মডেলের শ্রমবাজার প্রভাব সম্ভাবনার একটি প্রাথমিক দৃষ্টিভঙ্গি। *arXiv প্রিন্ট* *arXiv:2303.10130*
- ফিনি-অ্যাপলে, জে., ডেনি, পি., বেকার, বি.এ., লুফটন-রিলি, এ., এবং প্রাথার, জে., 2022। রোবটগুলি আসছে: পরিচিতিমূলক প্রোগ্রামিং-এ ওপেনএআই কোডের প্রভাবগুলি অন্বেষণ করা। <https://doi.org/10.1145/3511861.3511863>
- <http://dx.doi.org/10.1037/0033-295X.101.2.211> Gala, Y., Fernandez, A., Diaz, J., & Dorronsoro, J. R. (2016)। নিউরোকম্পিউটিং, 176, 48-59।
- Galvez Vallejo, J. L., Snowdon, C., Stocks, R., Kazemian, F., Yan Yu, F. C., Seidl, C., ... & Barca, G. M., 2023. একটি চরম-স্কেল ইলেকট্রনিক কার্ঠাসমোর ব্যবস্থার দিকে। দ্য জার্নাল অফ কেমিক্যাল ফিজিক্স, 159(4)। <https://doi.org/10.1063/5.0156399>
- Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E., Ramesh, S., Luo, Y., Pearson, A. T., 2023. ডিটেক্টর এবং অন্ধ মানব পর্যালোচকদের সাথে ChatGPT দ্বারা উত্পন্ন বৈজ্ঞানিক বিমূর্তগুলিকে বাস্তব বিমূর্তের সাথে তুলনা করা। *npj ডিজিটাল সস্কে* 6, 75। <https://doi.org/10.1038/s41746-023-00819-6>
- Géron, A. (2022)। *Scikit-Learn, Keras, এবং TensorFlow*-এর সাথে হ্যান্ডস-অন মেশিন লার্নিং। ওরিলি মিডিয়া, ইনক।
- González-Briones, A., De La Prieta, F., Mohamad, M. S., Omatu, S., & Corchado, J. M., 2018. শক্তি অপ্টিমাইজেশান সমস্যায় মাল্টি-এজেন্ট সিস্টেম অ্যাপ্লিকেশন: একটি অভ্যাত্মিক পর্যালোচনা। শক্তি, 11(8), 1928। <https://doi.org/10.3390/en11081928>

- Goodfellow, I., Bengio, Y., & Courville, A., 2016। গভীর শিক্ষা. প্রেস সঙ্গে.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y., 2014। জেনারেটিভ প্রতিপক্ষের জাল। নিউরাল তথ্যে অগ্রগতি প্রক্রিয়াকরণ সিস্টেম, 27.
- গুয়ান, ডব্লিউ., স্ট্যাটস্টিকাল, আই., এবং তিয়ানজিং, এম., 2020, অক্টোবর। স্বয়ংক্রিয় পাঠ্য সংক্ষিপ্তকরণ এবং ট্রান্সফরমার মডেলের প্রযোজ্যতার উপর সমীক্ষা। ইন কন্ট্রোল, রোবোটিক্স এবং ইন্টেলিজেন্ট সিস্টেমের উপর 2020 তম আন্তর্জাতিক সম্মেলনের কার্যক্রম (পৃষ্ঠা 176-184)। <https://doi.org/10.1145/3437802.3437832> Haarnoja, T., Zhou, A., Abbeel, P., & Levine, S., 2018. নরম অভিনেতা-সমালোচক: একজন স্টোকাস্টিক অভিনেতার সাথে অফ-পলিসি ম্যাক্সিমাম এনট্রপি ডিপ রিইনফোর্সমেন্ট লার্নিং। ইন 35 তম আন্তর্জাতিকের কার্যক্রম মেশিন লার্নিং সম্মেলন (Vol. 80, pp. 1861-1870)।
- Hao, K., 2020। আমরা সেই কাগজটি পড়ি যেটি Timnit Gebru কে Google থেকে বের করে দিতে বাধ্য করেছে। এটি যা বলে তা এখানে। এমআইটি প্রযুক্তি পর্যালোচনা।
- হাসানি, এইচ., সিলভা, ই.এস., 2023. ডেটা সায়েন্সে চ্যাটজিপিটির ভূমিকা: কীভাবে এআই-সহায়ক কথোপকথনমূলক ইন্টারফেসগুলি ক্ষেত্রের বিপ্লব ঘটছে। *Big Data Cogn. কম্পিউটা*, 7, 62। <https://doi.org/10.3390/bdcc7020062>
- He, K., Zhang, X., Ren, S., & Sun, J., 2016. ইমেজ স্বীকৃতির জন্য গভীর অবশিষ্ট শিক্ষা। ইন কম্পিউটার দৃষ্টি এবং প্যাটার্ন স্বীকৃতির উপর *IEEE* সম্মেলনের কার্যক্রম (পৃ. 770-778)। <https://doi.org/10.1109/CVPR.2016.90>
- হেন্ডলার, জে., এবং মুলভেহিল, এ.এম., 2016। সামাজিক মেশিন: কৃত্রিম বুদ্ধিমত্তা, সামাজিক নেটওয়ার্কিং এবং মানবতার আসন্ন সংঘর্ষ. অ্যাপ্রেস। <https://doi.org/10.1007/978-1-4842-1156-4>
- Hernández, G., Rodríguez, S., González, A., Corchado, J. M., & Prieto, J., 2021. গভীর শিক্ষার অ্যালগরিদম ব্যবহার করে ভিডিও বিশ্লেষণ সিস্টেম। ইন অ্যান্সিয়েন্ট ইন্টেলিজেন্স-সঙ্কেতীয় এবং অ্যানালিসিস: অ্যান্সিয়েন্ট ইন্টেলিজেন্সের উপর 11 তম আন্তর্জাতিক সিম্পোজিয়াম (পৃষ্ঠা 186-199)। স্প্রিংগার ইন্টারন্যাশনাল পাবলিশিং। https://doi.org/10.1007/978-3-030-58356-9_19
- Hernandez-Nieves, E., Hernández, G., Gil-gonzalez, A. B., Rodríguez-Gonzalez, S., & Corchado, J. M., 2021. CEBRA: একটি কেস-ভিত্তিক যুক্তিযুক্ত অ্যানালিসিস যাতে ব্যাকিং পণ্যের সুপারিশ করা হয়। কৃত্রিম বুদ্ধিমত্তার ইঞ্জিনিয়ারিং অ্যানালিসিস, 104, 104327। <https://doi.org/10.1016/j.engappai.2021.104327>
- Ho, J., Song, D., & Elizalde, B., 2020. Denoising Diffusion Probabilistic Models. *ArXiv* প্রিপ্রিন্ট *ArXiv*.
- Hochreiter, S., & Schmidhuber, J., 1997. দীর্ঘ স্বল্পমেয়াদী স্মৃতি। নিউরাল গণনা, 9(8), 1735- 1780। <https://doi.org/10.1162/neco.1997.9.8.1735>
- হোমোলক, জে. 2023। ওষুধ, বিজ্ঞান এবং একাডেমিক প্রকাশনায় *ChatGPT*-এর সুযোগ এবং ঝুঁকি: একটি আধুনিক প্রমিথিয়ান দ্বিধা. <https://doi.org/10.3325/cmj.2023.64.1>
- Hovy, D., & Spruit, S. L., 2016. প্রাকৃতিক ভাষা প্রক্রিয়াকরণের সামাজিক প্রভাব। *ACL 2016*। <https://doi.org/10.18653/v1/P16-2096>
- তিনি, এল., 2022। জেনারেটিভ এআই এবং ভবিষ্যত. ইউআরএল থেকে 23 জানুয়ারী, 2023 পুনরুদ্ধার করা হয়েছে <https://pub.towardsai.net/generative-ai-and-future-c3b1695876f2>
- Huang, R., Zhou, Z., Zhang, Y., & Zhao, Z., 2023. মেক-এন-অডিও: প্রম্পট-বর্ধিত ডিফিউশন মডেল সহ পাঠ্য থেকে অডিও প্রজন্ম। *ArXiv* প্রিপ্রিন্ট *ArXiv*:2301.12661।
- Itoh, S., & Okada, K., 2023. The Power of Large Language Models: A ChatGPT-চালিত টেক্সচুয়াল অ্যানালাইসিস অফ ফান্ডামেন্টাল ডেটা। এ উপলব্ধ এসএসআরএন 4535647। <https://doi.org/10.2139/ssrn.4535647>

- জানবি, এন., মেহমুদ, আর., কাতিব, আই., আলবেশী, এ., কচাঁডো, জে. এম., এবং ইজিটকানলার, টি., 2022. ইমতিদাদ: একটি রেফারেন্স আর্কিটেকচার এবং স্বকের রোগ নির্ণয়ের জন্য বিতরণকৃত এআই পরিষেবাগুলির বিকাশের উপর একটি কেস স্টাডি মেঘ, কুয়াশা এবং প্রান্তের উপরে। *সেম্বর*, 22(5), 1854। <https://doi.org/10.3390/s22051854>
- Jiang, H., Sun, D., Jampani, V., Yang, M. H., Learned-Miller, E., & Kautz, J., 2018. Super SloMo: ভিডিও ইন্টারপোলেশনের জন্য একাধিক মধ্যবর্তী ফ্রেমের উচ্চ মানের অনুমান। *ইন সিডিপিআর*(পৃ. 9000-9008)। <https://doi.org/10.1109/CVPR.2018.00938>
- Jiao, F., Ding, B., Luo, T., & Mo, Z., 2023. Panda LLM: ওপেন সোর্সড চাইনিজ ইন্সট্রাকশন-ফলোয়িং লার্জ ল্যাংগুয়েজ মডেলের জন্য প্রশিক্ষণ ডেটা এবং মূল্যায়ন। *arXiv প্রিপ্রিন্ট arXiv:2305.03025*।
- Jovanovic, M., & Campbell, M., 2022. জেনারেটিভ কৃত্রিম বুদ্ধিমত্তা: প্রবণতা এবং সম্ভাবনা। *কম্পিউটার*, 55(10), 107-112। <https://doi.org/10.1109/MC.2022.3192720>
- জুরাফস্কি, ডি., এবং মার্টিন, জেএইচ, 2019। বক্তৃতা এবং ভাষা প্রক্রিয়াকরণ। *পিয়ারসন*।
- Kingma, D. P., & Welling, M., 2013. অটো-এনকোডিং ভেরিয়েশনাল বেইস। *arXiv প্রিপ্রিন্ট arXiv:1312.6114*।
- কোখাদিয়া, ডি., ভট্ট, সি, সাপারিয়া, কে, প্যাটেল, কে, গিল-গজালেজ, এ. বি, এবং কচাঁডো, জে. এম., 2022. ডিপসাইন: গভীর শিক্ষা ব্যবহার করে সাইন ল্যাঙ্গুয়েজ সনাক্তকরণ এবং স্বীকৃতি। *ইলেকট্রনিক্স*, 11(11), 1780। <https://doi.org/10.3390/electronics11111780>
- Krizhevsky, A., Sutskever, I., & Hinton, G. E., 2012. গভীর কনভোলুশনাল নিউরাল নেটওয়ার্কের সাথে ইমেজনেট শ্রেণীবিভাগ। *নিউরাল ইনফরমেশন প্রসেসিং সিস্টেমে অগ্রগতি*, 25.
- Kudo, T., & Richardson, J., 2018. সেন্টেন্সপিস: নিউরাল টেক্সট প্রসেসিংয়ের জন্য একটি সহজ এবং ভাষার স্বাধীন সাবওয়ার্ড টোকেনাইজার এবং ডিটোকেনাইজার। *arXiv প্রিপ্রিন্ট arXiv:1808.06226*। <https://doi.org/10.18653/v1/D18-2012>
- Li, Y., Li, Z., Zhang, K., Dan, R., Jiang, S., & Zhang, Y., 2023. ChatDoctor: একটি মেডিকেল চ্যাট মডেল ফাইন-টিউনড একটি বড় ভাষার মডেল মেটা-এআই (LLaMA) মেডিকেল ডোমেন জ্ঞান ব্যবহার করে। *কিউরিয়াস*, 15(6)। <https://doi.org/10.7759/cureus.40895>
- Liu, H., Sferrazza, C., & Abbeel, P., 2023. ভাষাগুলি হল পুরস্কার: হিন্দুসাইট ফাইনটিউনিং মানুষের প্রতিক্রিয়া ব্যবহার করে। *arXiv প্রিপ্রিন্ট arXiv:2302.02676*।
- Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shimray, S., & Wang, Z., 2023. চ্যাটজিপিটি এবং একটি নতুন একাডেমিক বাস্তবতা: কৃত্রিম বুদ্ধিমত্তা-লিখিত গবেষণা পত্র এবং পণ্ডিত প্রকাশনায় বৃহৎ ভাষার মডেলের নীতিশাস্ত্র। <https://doi.org/10.1002/asi.24750>
- এম. সালভাগ্নো, ট্যাকোন, এফ., গেরলি, এ., 2023। কৃত্রিম বুদ্ধিমত্তা কি বৈজ্ঞানিক লেখার জন্য সাহায্য করতে পারে? <https://doi.org/10.1186/s13054-023-04380-2>
- Madotto, A., Lin, Z., Winata, G. I., & Fung, P., 2021. Few-shot bot: সংলাপ সিস্টেমের জন্য প্রম্পট-ভিত্তিক শিক্ষা। *arXiv প্রিপ্রিন্ট arXiv:2110.08118*।
- ম্যানিং, C. D., et al., 2008. তথ্য পুনরুদ্ধার ভূমিকা। *কেমব্রিজ ইউনিভার্সিটি প্রেস*।
- Maroto-Gomez, M., Castro-González, Á., Castillo, J. C., Malfaz, M., & Salichs, M. Á., 2023. মানব-রোবট ইন্টারেক্টিভ যোগাযোগের জন্য ব্যবহারকারীর পছন্দের পূর্বাভাসের উপর সমর্থিত একটি অভিযোজিত সিদ্ধান্ত গ্রহণের ব্যবস্থা। *ইউজার মডেলিং এবং ইউজার-অ্যাডাপ্টেড ইন্টারঅ্যাকশন*, 33(2), 359-403। <https://doi.org/10.1007/s11257-022-09321-2>
- McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E., 2006. কৃত্রিম বুদ্ধিমত্তার উপর ডার্টমাউথ গ্রীষ্মকালীন গবেষণা প্রকল্পের একটি প্রস্তাব, 31 আগস্ট, 1955। *AI দোকান*, 27(4), 12-12।
- McCulloch, W. S., & Pitts, W., 1943. একটি যৌক্তিক ক্যালকুলাস ধারণা স্নায়বিক কার্যকলাপের মধ্যে immanent. *গাণিতিক বায়োফিজিক্সের বুলেটিন*, 5, 115-133। <https://doi.org/10.1007/BF02478259>
- ম্যাকগ্রেগর, এস., মেমন, এন., এবং লেভি, কে., 2020। সাইবার নিরাপত্তা এবং মানবাধিকার। *ইন ন্যায্যতা, জবাবদিহিতা এবং স্বচ্ছতার উপর 2020 সম্মেলনের কার্যপ্রণালী* (পৃ. 425-435)।

- Miao, X., Wang, Y., Jiang, Y., Shi, C., Nie, X., Zhang, H., & Cui, B., 2022. Galvatron: স্বয়ংক্রিয় সমান্তরালতা ব্যবহার করে একাধিক gpus-এ দক্ষ ট্রান্সফরমার প্রশিক্ষণ। *arXiv প্রিপ্রিন্ট arXiv:2211.13878* <https://doi.org/10.14778/3570690.3570697>
- মুর, জে., 2006। ডার্টমুথ কলেজ কৃত্রিম বুদ্ধিমত্তা সম্মেলন: পরবর্তী পঞ্চাশ বছর। তোমার দোকান আছে, 27(4), 87-87।
- Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W., 2011. প্রাকৃতিক ভাষা প্রক্রিয়াকরণ: একটি ভূমিকা। আমেরিকান মেডিকেল ইনফরমেটিক্স অ্যাসোসিয়েশনের জার্নাল, 18(5), 544-551। <https://doi.org/10.1136/amiajnl-2011-000464>
- Nakano, R., Hilton, J., Balaji, S., Wu, J., Ouyang, L., Kim, C., ... & Schulman, J., 2021. Webgpt: এর সাথে ব্রাউজার-সহায়ক প্রশ্ন-উত্তর মানুষের প্রতিক্রিয়া। *arXiv প্রিপ্রিন্ট arXiv:2112.09332*।
- Nguyen, T. M., & Wu, Q. J. (2013)। মেডিকেল ইমেজ সেগমেন্টেশনের জন্য মার্কভ এলোমেলো ক্ষেত্র ভিত্তিক একটি ফাজি লজিক মডেল। বিকশিত সিস্টেম, 4, 171-181।
- Nguyen, T. T., Wilson, C., & Dalins, J., 2023. ফাইন-টিউনিং লামা 2 অনলাইন যৌন শিকারী চ্যাট এবং অপমানজনক পাঠ্য সনাক্তকরণের জন্য বড় ভাষার মডেল। *arXiv প্রিপ্রিন্ট arXiv:2308.14683*। <https://doi.org/10.1109/TITS.2023.3297252>
- Nichol, A., et al., 2021. GLIDE: টেক্সট-চালিত অ্যান্সিকেশনের জন্য একটি জেনারেটিভ ল্যাঙ্গুয়েজ মডেল। কৃত্রিম বুদ্ধিমত্তা গবেষণা জার্নাল, 49(2), 315-334।
- OpenAI, 2021. ChatGPT API ডকুমেন্টেশন। OpenAI। 22 আগস্ট, 2023 থেকে সংগৃহীত <https://platform.openai.com/docs/guides/chat>
- OpenAI, 2023, 20 জুলাই। ChatGPT এর জন্য কাস্টম নির্দেশাবলী। OpenAI ব্লগ। <https://openai.com/blog/chatgpt-এর-জন্য-কাস্টম-নির্দেশনা>
- OpenAI, 2023. GPT-4 প্রযুক্তিগত প্রতিবেদন। *ArXiv, abs/2303.08774*।
- Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., ... & Lowe, R., 2022. মানুষের প্রতিক্রিয়া সহ নির্দেশাবলী অনুসরণ করার জন্য ভাষা মডেল প্রশিক্ষণ . নিউরাল ইনফরমেশন প্রসেসিং সিস্টেমে অগ্রগতি, 35, 27730-27744।
- পারিখ, ভি., শাহ, জে., ভট্ট, সি., কর্চাদো, জে. এম., এবং লে, ডি. এন., 2022, জুলাই। গভীর শিক্ষা ভিত্তিক স্বয়ংক্রিয় বুকের এক্স-রে অস্বাভাবিকতা সনাক্তকরণ। ইন অ্যান্ডিফিকেন্ট ইন্টেলিজেন্সের উপর আন্তর্জাতিক সিম্পোজিয়াম (পৃষ্ঠা 1-12)। চ্যাম: স্প্রিংগার ইন্টারন্যাশনাল পাবলিশিং। https://doi.org/10.1007/978-3-031-22356-3_1
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., ... & Chintala, S., 2019. Pytorch: an imperative style, high-performance গভীর শিক্ষার লাইব্রেরি। নিউরাল ইনফরমেশন প্রসেসিং সিস্টেমে অগ্রগতি, 32.
- Pavlyshenko, B. M. (2023)। ফাইন-টিউনড লামা 2 জিপিটি মডেল ব্যবহার করে আর্থিক সংবাদ বিশ্লেষণ। *arXiv প্রিপ্রিন্ট arXiv:2308.13032*।
- পেরেজ-পনস, এম.ই., আলোনসো, আর.এস., গার্সিয়া, ও., মারেইরোস, জি., এবং কর্চাদো, জে.এম., 2021। টেকসই কৃষি বাজারের জন্য গভীর কিউ-লার্নিং এবং পছন্দ ভিত্তিক মাল্টি-এজেন্ট সিস্টেম। সেম্বর, 21(16), 5276। <https://doi.org/10.3390/s21165276>
- Pérez-Pons, M. E., Parra-Dominguez, J., Hernández, G., Bichindaritz, I., & Corchado, J. M., 2023. OCI-CBR: অগ্রাধিকার-সচেতন বিনিয়োগ পরিস্থিতিতে সিদ্ধান্ত সমর্থনের জন্য একটি হাইব্রিড মডেল। অ্যান্সিকেশন সহ বিশেষজ্ঞ সিস্টেম, 211, 118568। <https://doi.org/10.1016/j.eswa.2022.118568>
- Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I., 2018। জেনারেটিভ প্রাক-প্রশিক্ষণের মাধ্যমে ভাষা বোঝার উন্নতি করা।
- Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I., 2021. ভাষার মডেল খুব কম-শট লার্নার্স। *OpenAI ব্লগ, 9(2), 650-700*।

- Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I., 2019. ভাষার মডেলগুলি হল আন-সুপারভাইজড মাল্টিটাস্ক লার্নার্স। ওপেনএআই ব্লগ, ১(৪), ৭।
- Reiter, E., & Dale, R., 2020। প্রাকৃতিক ভাষা জেনারেশন সিস্টেম তৈরি করা। কেমব্রিজ ইউনিভার্সিটি প্রেস।
- Rombach, M., et al., 2022. হাই ডেফিনিশন ভিডিও জেনারেশনের জন্য জেনারেটিভ ল্যাটেন্ট ডিফিউশন মডেল। মাল্টিমিডিয়ায় *IEEE* লেনদেন, 24(1), 123-137।
- রোজেনব্যাট, এফ., 1958. দ্য পারসেপ্টন: মস্তিষ্কে তথ্য সংরক্ষণ এবং সংগঠনের জন্য একটি সম্ভাব্য মডেল। মনস্তাত্ত্বিক পর্যালোচনা, 65(6), 386। <https://doi.org/10.1037/h0042519>
- রুমেলহাট, ডি.ই., হিল্টন, জি.ই., এবং উইলিয়ামস, আর.জে., 1986. ব্যাক-প্রোগ্রামিং ফ্রন্টের মাধ্যমে উপস্থাপনা শেখা। প্রকৃতি, 323(6088), 533-536। <https://doi.org/10.1038/323533a0>
- রাসেল, এস.জে., এবং নরভিগ, পি., 2014। কৃত্রিম বুদ্ধিমত্তা: একটি আধুনিক পদ্ধতি। হারলো
- সারসা, এস., ডেনি, পি., হেলাস, এ., এবং লেইনোলেন, জে., 2022। বড় ভাষার মডেল ব্যবহার করে প্রোগ্রামিং অনুশীলন এবং কোড ব্যাখ্যার স্বয়ংক্রিয় প্রজন্ম। <https://doi.org/10.1145/3501385.3543957> Sennrich, R., & Zhang, B., 2019. রিভিজিটিং লো-রিসোর্স নিউরাল মেশিন অনুবাদ: একটি কেস স্টাডি। *arXiv* প্রিপ্রিন্ট *arXiv:1905.11901*। <https://doi.org/10.18653/v1/P19-1021>
- Serban, I. V., et al., 2017. A Deep Reinforcement Learning Chatbot. *arXiv* প্রিপ্রিন্ট *arXiv:1709.02349*। Shazeer, N., 2020. Glu ভেরিয়েন্ট ট্রান্সফরমার উন্নত করে। *arXiv* প্রিপ্রিন্ট *arXiv:2002.05202*।
- শেরস্টিনস্কি, এ., 2020। পুনরাবৃত্ত নিউরাল নেটওয়ার্ক (RNN) এবং দীর্ঘ স্বল্প-মেয়াদী মেমরি (LSTM) নেটওয়ার্কের মৌলিক বিষয়গুলি। ফিজিকা ডি: ননলাইনার ফেনোমেনা, 404, 132306। <https://doi.org/10.1016/j.physd.2019.132306>
- সিদ্দিক, এস.; চৌ, জে.সি.এল., 2021। স্বাস্থ্যসেবা যোগাযোগে মেশিন লার্নিং। এনসাইক্লোপিডিয়া, ২, 220-239। <https://doi.org/10.3390/encyclopedia1010021>
- Simonyan, K., & Zisserman, A., 2014. বড় আকারের ইমেজ স্বীকৃতির জন্য খুব গভীর কনভোলুশনাল নেটওয়ার্ক। *arXiv* প্রিপ্রিন্ট *arXiv:1409.1556*।
- Su, Z., Chow, J. K., Tan, P. S., Wu, J., Ho, Y. K., & Wang, Y. H., 2021. গভীর কনভোলুশনাল নিউরাল নেটওয়ার্ক-ভিত্তিক পিক্সেল-ভিত্তিক ল্যান্ডস্লাইড ইনভেন্টরি ম্যাপিং। ভূমিধস, 18, 1421-1443। <https://doi.org/10.1007/s10346-020-01557-6>
- Tang, D., Rong, W., Qin, S., Yang, J., & Xiong, Z., 2020. উত্তর নির্বাচনের জন্য পর্যালোচনা সহ একটি এন-গেটেড পুনরাবৃত্ত ইউনিট। নিউরোকম্পিউটিং, 371, 158-165। <https://doi.org/10.1016/j.neucom.2019.09.007>
- টেট, ই., 2023. হাই ডেফিনিশন ভিডিও জেনারেশন: একটি ব্যাপক পর্যালোচনা। মাল্টিমিডিয়া প্রসেসিং জার্নাল।
- Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., ... & Scialom, T., 2023. Llama 2: Open Foundation and fine-tuned চ্যাট মডেল। *arXiv* প্রিপ্রিন্ট *arXiv:2307.09288*।
- Tulyakov, S., Liu, M. Y., Yang, X., & Kautz, J., 2018. Mocogan: ভিডিও তৈরির জন্য পচনশীল গতি এবং বিষয়বস্তু। ইন কম্পিউটার ভিশন এবং প্যাটার্ন স্বীকৃতির উপর *IEEE* সম্মেলনের কার্যক্রম (পৃষ্ঠা 1526-1535)। <https://doi.org/10.1109/CVPR.2018.00165>
- টুরিং এ.এম., 1950. কম্পিউটিং যন্ত্রপাতি এবং বুদ্ধিমত্তা। মধ্য: মন 59.236, পৃষ্ঠা 433-460। <https://doi.org/10.1093/mind/LIX.236.433>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... এবং Polosukhin, I., 2017। মনোযোগ আপনার প্রয়োজন সব। নিউরাল ইনফরমেশন প্রসেসিং সিস্টেমে অগ্রগতি, 30.
- ভন অসওয়াল্ড, জে., নিকলাসন, ই., রান্ডাজো, ই., স্যাক্রামেন্টো, জে., মর্ডভিল্টসেভ, এ., ঝমোগিনভ, এ., এবং ব্লাদিমিরভ, এম., 2023, জুলাই। ট্রান্সফরমার গ্রেডিয়েন্ট ডিসেন্টের মাধ্যমে ইন-কনটেক্সট শিখে। ইন মেশিন লার্নিং আন্তর্জাতিক সম্মেলন (পৃষ্ঠা 35151-35174)। পিএমএলআর।

- Vondrick, C., Pirsiavash, H., & Torralba, A., 2016. দৃশ্য গতিশীলতার সাথে ভিডিও তৈরি করা। ইন নিউরাল ইনফরমেশন প্রসেসিং সিস্টেমে অগ্রগতি (পৃষ্ঠা 613-621)।
- ওয়ালেস, ই., এট আল।, 2019। এনএলপি আক্রমণ এবং বিশ্লেষণের জন্য ইউনিভার্সাল অ্যাডভান্সারিয়াল ড্রিগার। *EMNLP 2019* | <https://doi.org/10.18653/v1/D19-1221>
- Wei, R., & Mahmood, A., 2020. বায়োমেডিকাল ইনফরমেশনের জন্য প্রতিনিধিত্বমূলক শিক্ষার সাথে বৈচিত্রপূর্ণ অটোএনকোডারে সাম্প্রতিক অগ্রগতি: একটি সমীক্ষা। *আইইইইই এক্সেস*, ৯, 4939-4956। <https://doi.org/10.1109/ACCESS.2020.3048309>
- উইলিয়ামস, জে., 2023. এআই-জেনারেটেড টেক্সট ডিটেক্টরের কার্যকারিতা। *শিক্ষা প্রযুক্তি জার্নাল*।
- Xie, J., Xu, L., & Chen, E., 2018. গভীর নিউরাল নেটওয়ার্কের সাথে ইমেজ ডিনোইসিং এবং ইনপেইন্টিং। ইন নিউরাল ইনফরমেশন প্রসেসিং সিস্টেমে অগ্রগতি (পৃষ্ঠা 341-349)।
- Xu, L., Sanders, L., Li, K., Chow, J. C. L., 2021. কৃত্রিম বুদ্ধিমত্তা এবং মেশিন লার্নিং ব্যবহার করে স্বাস্থ্য পরিচর্যা এবং অনকোলজি অ্যাপ্লিকেশনের জন্য চ্যাটবট: পদ্ধতিগত পর্যালোচনা। <https://doi.org/10.2196/27850>
- ইয়ালালভ, D. 2023. চ্যাটবটের বিবর্তন: T9 এবং GPT-1 এর যুগ থেকে ChatGPT পর্যন্ত। পোস্ট। URL: <https://mpost.io/es/la-evoluci%C3%B3n-de-los-chatbots-de-la-era-t9-y-gpt-1-a-chatgpt/>
- Zadeh, L. A., 2008. অস্পষ্ট যুক্তির প্রয়োজন আছে কি? *তথ্য বিজ্ঞান*, 178(13), 2751-2779। <https://doi.org/10.1016/j.ins.2008.02.012>
- Zaremba, W., & Brockman, G., 2021, আগস্ট 10. OpenAI কোডেক্স। OpenAI। <https://openai.com/blog/openai-codex/>
- ঝাও, জে., ওয়াং, টি., ইয়াতসকার, এম., কোটেরেল, আর., অর্ডোনেজ, ভি., এবং চ্যাং, কে.ডব্লিউ., 2018। ইন কম্পিউটেশনাল ভাষাবিজ্ঞানের জন্য অ্যাসোসিয়েশনের উত্তর আমেরিকান অধ্যায়ের 2018 সম্মেলনের কার্যক্রম: মানব ভাষা প্রযুক্তি, ভলিউম 2 (ছোট পেপার) (পৃষ্ঠা 8-14)।
- ঝাও, জেড, এট আল। 2023. মেক-এন-অডিও: প্রম্পট-এনহ্যান্সড ডিফিউশন মডেল সহ টেক্সট-টু-অডিও জেনারেশন। মেশিন লার্নিং সংক্রান্ত 40তম আন্তর্জাতিক সম্মেলনের কার্যক্রম।
- Zoran, D., Kabra, R., Lerchner, A., & Rezende, D. J., 2021. পার্টস: স্লট, মনোযোগ এবং স্বাধীনতার সর্বোচ্চকরণের সাথে তত্ত্বাবধানহীন সেগমেন্টেশন। ইন কম্পিউটার ভিশনের উপর *IEEE/CVF আন্তর্জাতিক সম্মেলনের কার্যক্রম* (পৃষ্ঠা 10439-10447)। <https://doi.org/10.1109/ICCV48922.2021.01027>