شماره مجوز:
۱۹۷۵/دم آ

CSSC
computer student's scientific community
انجمن علمی دانشجویی کامپیوتر ارشاد دماوند

تکنیک

آذر ماه ۱۴۰۳
شماره ۳۶

- انواع حملات سایبری
- بررسی نسخه های CHAT-GPT
- کاربرد علم آمار در کامپیوتر

@CSSC_EDU

@CSSCEDU

سخنی با دانشجویان

با عرض سلام و تبریک سال جدید تحصیلی خدمت شما دانشجویان گرامی به ویژه نو دانشجویان رشته کامپیوتر و اساتید محترم.

مگیت نشریه ای از طرف انجمن علمی کامپیوتر است که در ابتدا به صورت هفتگی و سپس به صورت ماهانه منتشر می‌شد. این مجله شامل انتشار اخبار روز تکنولوژی، بررسی مفاهیم مرتبط با کامپیوتر، معرفی کتاب یا فیلم های کامپیوتری و..... بود. ما در انتشار این نشریه وقفه ای داشتیم اما بعد از تلاش و پیگیری مجدد و همچنین همکاری دانشجویان رشته کامپیوتر توانستیم دوباره فعالیت مگیت را زیر نظر استاد مشاور انجمن علمی کامپیوتر، خانم مهندس مریم تیموری از سر بگیریم.

امیدواریم که از مطالب مگیت بهره لازم را ببرید. همچنین ما از همکاری شما در مگیت استقبال می‌کنیم. در صورتی که تمایل به همکاری در نشریه مگیت را دارید، به آی دی sysreamr@ یا EVE2800@ پیام دهید.

نازنین فاطمه ترکمن دانشجوی مهندسی کامپیوتر ترم ۵

انواع حملات سایبری

حمله سایبری چیست؟

حمله سایبری نوعی حمله به شبکه ها و سیستم ها است که سبب دسترسی غیر مجاز مهاجمین (بازیگران تهدید، مجرمین) به آنها می‌شود که اهدافی غیر اخلاقی و جنایتکارانه در پس آنها قرار دارد.

این اهداف انواع گوناگونی دارند که برخی از آن‌ها اهداف مالی، ایجاد آشوب و دلهره، انتقام و خصومت شخصی، جنگی و بین‌المللی و... هستند.

انواع حملات سایبری

حملات سایبری به شیوه‌های مختلف اجرا می‌شوند از جمله Malware, Phishing, SQL Injection, DNS Spoofing, Backdoors, Formjacking, Insider Threat, Zero-Day Exploit, Drive-by Download, Eavesdropping Attack. اگرچه این روش‌ها تنها برخی از حملات سایبری هستند اما ممکن است نامشان به گوشتان خورده باشد. در ادامه به معرفی برخی از این روش‌ها می‌پردازیم.

Malware

این کلمه مخفف عبارت Malicious Software به معنی بدافزار می‌باشد. در این روش، مهاجم از طریق بدافزارها یا نرم‌افزارهای مشکوک و لینک های مخرب، کنترل سیستم را به دست گرفته و می‌تواند به آن آسیب برساند یا اطلاعات موجود در سیستم را بدزدد.

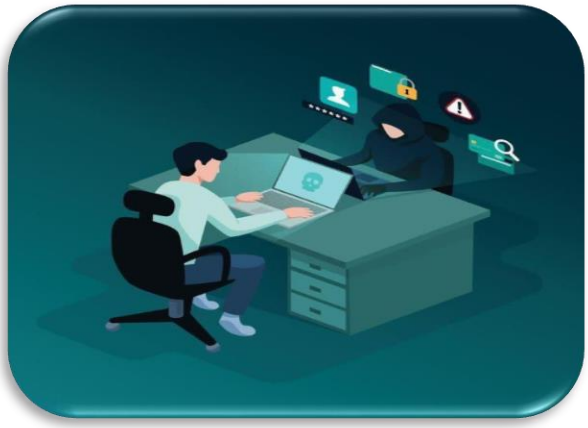
Phishing

"فیشینگ" نوعی حمله سایبری است که در آن یک یا مجموعه ای از ایمیل های پنهان، پیام‌های مستقیم، تبلیغات جعلی یا حتی اخبار غیر موثق، برای بدست آوردن اطلاعات حساس از جمله اطلاعات شخصی، رمز عبور و کارت اعتباری و... استفاده می‌شود. از رایج ترین موضوعات مربوط به "فیشینگ" می‌توان به مشکلات اعتباری_پرداختی، تنظیم مجدد رمز عبور و ادعای برنده شدن جوایز، اشاره کرد.

Formjacking

"فرم جکینگ" روش دیگری از حمله سایبری است که در آن مهاجم کد مخرب و ناسالم جاوا اسکریپت را برای هک کردن یک وبسایت و به دست گرفتن کنترل عملکرد صفحه سایت جهت جمع‌آوری اطلاعات حساس کاربر، بکار می‌برد. این روش بگونه ای کار می کند که کاربر وبسایت موردنظر، اطلاعات و داده های کارت اعتباری خود را وارد و سپس روی گزینه پرداخت، کلیک می‌کند. آن‌گاه کد مخرب جاوا اسکریپت، داده های وارده را جمع آوری کرده و به سرقت می‌برد.

انواع حملات سایبری



اما چیزی که واضح است این است که جاوا اسکریپت تنها زبانی نیست که می‌توان از کدهای مخربش جهت حمله سایبری استفاده کرد. پنج زبان برنامه‌نویسی موردعلاقه هکرها در سرتاسر جهان که بطور گسترده از آنها استفاده می‌شود، به ترتیب بدین شرح می‌باشند:

- Python
- Java Script
- PHP
- SQL
- C Programming

تنها در سال ۲۰۲۳ تا به امروز، بیش از ۲/۳ میلیون گزارش حمله سایبری در بریتانیا دریافت شده است که میانگین هزینه هر حمله حدود ۳۲۳۰ پوند، برآورد شده است. البته جالب است بدانید که ایالات متحده در لیست کشورهایی با بیشترین دریافت حمله سایبری، رتبه اول را دارا می‌باشد. البته که حمله سایبری چیز جدیدی نیست. از زمان تولد اینترنت، شرکت‌ها و سازمان های بزرگ و کوچک در شکل ها و ابعاد متفاوتی، مورد حمله مهاجمین سایبری قرار گرفته اند.

منابع:

- https://www.ibm.com
- https://em360tech.com
- https://www.simplilearn.com

بررسی نسخه های Chat GPT

در سال های اخیر پیشرفت هوش مصنوعی به ویژه در حوزه چت بات ها به طور چشمگیری افزایش یافته است. ChatGPT به عنوان یک مدل زبانی مبتنی بر هوش مصنوعی یکی از چت بات های پرکاربرد است که قطعاً اسم آن را زیاد شنیده اید. طبیعتاً هیچ نوآوری و ابداعی بدون چالش نبوده و همیشه نقص هایی داشته است که به مرور کم تر شده یا به صورت کامل از بین رفته است. در مواقعی نیز چالش ها همچنان پابرجاست و رفع آنها یک دغدغه اساسی است. ChatGPT هم از این قاعده استثنا نبوده و به مرور پیشرفت کرده که البته همچنان با چالش ها و مشکلاتی رو به رو است. از همین رو می‌خواهیم به بررسی انواع مدل های ChatGPT از ابتدا تا کنون بپردازیم.

ChatGPT بر پایه پردازش زبان طبیعی (NLP) بنیان گذاری شده است که قابلیت فهم زبان انسان و تعامل با وی را ایجاد می‌کند. همچنین استفاده از معماری ترانسفورمر در ChatGPT ها باعث می‌شود تا نسبت به مدل های قبلی بهتر عمل کنند. یکی از قابلیت های مهم این معماری این است که به جای توجه به کلمات پشت سر هم، به کل کلمات در یک جمله دقت می‌شود تا درک بهتری صورت گیرد. قبل از ایجاد ChatGPT روی ساخت و گسترش GPT ها کار شد. این مدل ها بر روی داده های زیادی آموزش می‌دیدند و ساختار و الگوهای زبان را درک می‌کردند تا بتوانند متن های جدیدی تولید کنند. با بهبود مدل ها سرانجام ChatGPT نیز بر پایه مدل GPT-3.5 ساخته شد. در حقیقت ChatGPT ها بر اساس برخی از نسخه های GPT ها ساخته شده اند. در ادامه ابتدا به مدل های GPT و سپس معرفی نسخه های ChatGPT می‌پردازیم.

انواع GPT

GPT-1

این نسخه اولین مدل GPT است که در سال ۲۰۱۸ بر پایه معماری ترانسفورمر ایجاد شد که وظیفه پردازش زبان طبیعی مانند ترجمه ماشینی و مدل سازی زبان را بر عهده داشت. این مدل با استفاده از حجم عظیمی از داده های متنی مثل کتاب، مقالات و صفحات آموزش داده شد. با این روش GPT-1 توانست کلمات بعدی را پیش بینی کند. از مهم ترین کاربردهای این نسخه ترجمه زبان، طبقه بندی متون و تحلیل احساسات بود. این نسخه از ۱۱۷ میلیون پارامتر تشکیل شده بود که با این وجود در برابر نسخه های بعدی مقدار ناچیزی محسوب می‌شد. به طور خلاصه باید گفت که تعداد بیش تر پارامترها به مدل زبانی کمک می‌کند تا الگوهای پیچیده تری را یاد بگیرند.

GPT-2

نسخه GPT-2 با ۱/۵ میلیارد پارامتر ساخته شد که نسبت به نسخه قبلی پیشرفت بسیار بزرگی محسوب می‌شد. نحوه آموزش و پیش آماده سازی این نسخه نیز مانند GPT-1 بود؛ با این تفاوت که در تولید متن های طولانی و منسجم موفق تر از نسخه قبلی خود بود. یکی از موفقیت های مهم GPT-2 تولید متن هایی بود که شباهت بسیار زیادی به نوشته های انسان ها داشت و تشخیص تفاوت را بسیار سخت می‌کرد. سوء استفاده

بررسی نسخه های Chat GPT

از این مدل برای اخبار جعلی نگران کننده بود! از همین رو OpenAI به جای انتشار نسخه کامل، یک نسخه با قابلیت های کم تر را منتشر کرد.

GPT3

این مدل یکی قدرتمندترین و بزرگترین مدل های ایجاد شده محسوب می‌شود که دارای ۱۷۵ میلیارد پارامتر است. GPT3 با صفحات وب، کتاب ها، مقالات و سایر مطالب نوشتاری پیش آماده سازی شد و همچنان مانند GPT2 کلمات بعدی را پیش بینی می‌کرد اما در تولید متن پیشرفت بسیار زیادی داشت؛ طبقه بندی متون، پرسش و پاسخ و تحلیل و تجزیه احساسات و عواطف با دقت بیش تری انجام می‌شد.

یادگیری چند وظیفه ای یکی از ویژگی های منحصر به فرد این مدل است که می‌تواند چند کار را به صورت همزمان انجام دهد و با چند مثال آموزش ببیند. همین ویژگی باعث شد تا در بسیاری از زمینه ها مانند چت بات ها، تولید محتوا، ترجمه و حتی کد نویسی مورد استفاده قرار گیرد. این نسخه تغییر بزرگی در جامعه هوش مصنوعی ایجاد کرد.

instructGPT

این مدل بر پایه GPT3 توسط OpenAI به وجود آمد تا بتواند تعامل بهتری با انسان برقرار کند. مدل یادگیری بر اساس بازخورد انسانی بود تا نحوه پاسخ دهی به سوالات نسبت به قبل بهبود یابد. روند آموزش به این گونه بود که خروجی GPT را با پاسخ ایده آل مقایسه می‌کردند. سپس با برچسب گذاری روی پاسخ ها مشخص می‌شد که کدام پاسخ ها به حالت مطلوب نزدیک تر است. در نهایت نیز برچسب ها به GPT داده می‌شود تا در پاسخ های بعدی پیشرفت کند.

ProtGPT2

این مدل تنها برای طراحی و مهندسی پروتئین طراحی شده که قادر به تولید توالی های بعدی پروتئین با حفظ ویژگی های طبیعی آن است. این مدل با استفاده از معماری GPT2 پیش می‌رود و روی یک پایگاه داده بزرگ از توالی پروتئین، از پیش آموزش دیده است.

BioGPT

این مدل به طور اختصاصی برای تولید متن های زیست پزشکی ایجاد شده است و بر اساس معماری ترانسفورمر کار می‌کند. BioGPT از خلاصه های PubMed آموزش دیده است.

ChatGPT

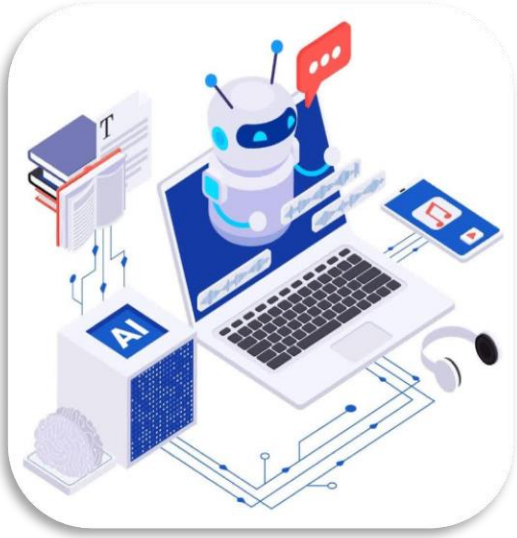
این مدل با کشف الگو و روابط بین کلمات و عبارات، پاسخ های منسجم و واقع گرایانه را در هنگام گفت و گو ارائه می‌دهد. آموزش به ChatGPT نیز با استفاده از داده های بزرگ متنی صورت گرفته است. این مدل بر اساس ساختار GPT-۳.۵ ساخته شده است که برای نخستین بار در سال ۲۰۲۲ با حدود ۶/۷ میلیارد پارامتر عرضه شد که به صورت خاص برای محاوره به کار می‌رفت. در حقیقت GPT-۳.۵ نسخه کوچک تری از GPT۳ محسوب می‌شود که بهبود یافته است.

بررسی نسخه های Chat GPT

GPT-4

این نسخه از پیشرفته ترین مدل های زبانی است که می‌تواند علاوه بر متن تصویر را نیز پردازش کند و خروجی متنی تحویل دهد. در بسیاری از معیارها، عملکردی مشابه انسان را از خود نشان داده است.

پس از بررسی نسخه های GPT، نسخه های عرضه شده ChatGPT را بیان می‌کنیم.



بررسی انواع نسخه های ChatGPT

ChatGPT-3.5 turbo and ChatGPT-3.5

قادر به درک و تولید زبان و کد است. بهینه سازی گفتگو با استفاده از API تکمیل گفت و گو (یک ابزار برنامه نویسی) از ویژگی های این سری است. نسخه Turbu برای افزایش سرعت و انجسام بیش تری در تولید متن ساخته شد. این نسخه ها همچنان در درک سوالات پیچیده دچار ابهام شوند.

ChatGPT-4 Turbo and ChatGPT-4

این مدل که در مارچ ۲۰۲۳ منتشر شد، در وظایف پایه‌ای تفاوت چندانی با ChatGPT-3.5 ندارد اما در استدلال های پیچیده تر موفق تر می‌کند و نه تنها در زبان انگلیسی، بلکه در سایر زبان ها عملکرد خوبی از خود نشان داده است. نسخه Turbo نیز در سال اواسط سال ۲۰۲۳ عرضه شد.

o1-preview و o1-mini

این سری ها با هدف انجام استدلال های پیچیده تر و با استفاده از آموزش تقویتی ایجاد شدند. این نسخه ها قبل از پاسخ دهی زمان بیش تری را فکر و تحلیل می‌کند تا با دقت بیش تری جواب دهند.

نسخه مینی از سرعت بیش تری برخوردار بوده و برای استدلال علوم و ریاضیات گزینه مناسبی است.

ChatGPT-4o Audio and Realtime

- Audio: این سری نیز در تکمیل گفت و گو قادر به تولید محتوای صوتی است.
- Realtime: این مدل با استفاده از رابط WebSocket به ورودی های متنی و صوتی پاسخ می‌دهد.

ChatGPT-4o mini

این مدل نیز به صورت چند وجهی عمل می‌کند و نسبت به مدل های قبلی در تولید و درک زبان باهوش تر است. همچنین در پردازش تصاویر و ارائه اطلاعات به صورت دقیق تر عمل می‌کند.

ChatGPT-4o

این نسخه سرعتش از ChatGPT-4 Turbo حدود ۲ برابر بیش تر است. همچنین خلاقیت این نسخه برای تولید محتوا و پردازش تصاویر نسب به قبل افزایش قابل توجهی داشته است و برای زبان های دیگر به غیر از انگلیسی از همه نسخه ها موفق تر بوده است.

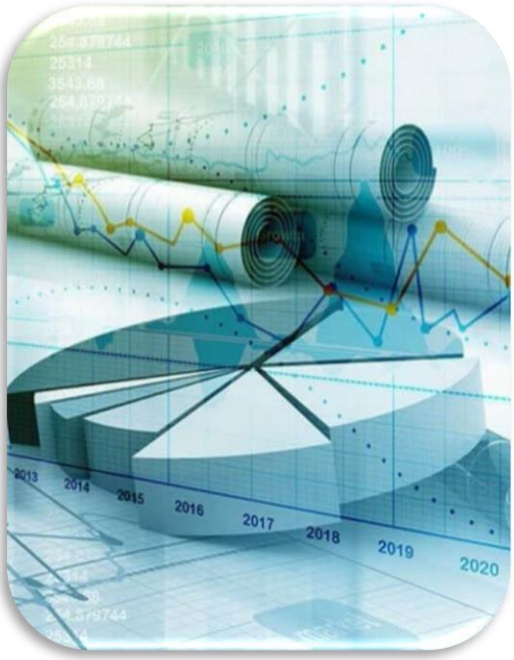
منابع:

- https://www.sciencedirect.com/science/article/pii/S266734522300024X?__cf_chl_tk=_.yMnPOsDo4LF_J.46K2z51lE9HfI79BRAurlVH_meo-1729260684-1.0.1.1-nX1BMosQRprW6joDh6fbO.GyA_6hCWCp_7agLs.Doio#bib6
- https://platform.openai.com/docs/models

کاربرد علم آمار در کامپیوتر

آمار شامل شاخه‌ای از ریاضیات است که به گردآوری، تحلیل، و ارائه داده ها می‌پردازد. آمار برای تحلیل آنچه در جهان اطراف ما اتفاق می‌افتد مورد استفاده قرار می‌گیرد. یکی از کاربردهای آمار این است که اتفاقات گذشته را بررسی نمود و در جهت پیش بینی آینده به کار گرفت.

آمار مدرن جهت انجام بعضی از محاسبات پیچیده به وسیله رایانه ها مورد استفاده قرار می‌گیرد. کل شاخه‌های آمار به وسیله محاسبات کامپیوتری امکان پذیر شده‌اند. انقلاب کامپیوتری یک توجه ویژه ای به آمار «آزمایشی» و «شناختیک» داشته که رویکردهایی برای آینده آمار به حساب می‌آید.



یکی از مهم‌ترین کاربردهای آمار و احتمال با استفاده از کامپیوتر شبیه سازی است . شبیه سازی نسخه‌ای از بعضی ابزار حقیقی یا موقعیت‌های کاری است. شبیه سازی سعی می‌کند تا بعضی جنبه‌های رفتاری یک سیستم فیزیکی یا انتزاعی را به وسیله رفتار سیستم دیگری نمایش دهد.

برخی از کاربردهای دیگر علم آمار در کامپیوتر عبارتند از:

۱. طراحی آزمایش‌ها و تحلیل داده ها: آمار به عنوان یک ابزار کلیدی در جمع‌آوری داد ها، طراحی آزمایش ها و تحلیل آنها در زمینه های مختلف از جمله علوم کامپیوتر و مهندسی نرم افزار به کار می‌آید.
۲. یادگیری ماشین و هوش مصنوعی: آمار به عنوان یک ابزار اساسی در تحلیل داده ها، پیش‌بینی الگوها و آموزش مدل های یادگیری ماشین مورد استفاده قرار می‌گیرد. در حقیقت الگوریتم های یادگیری ماشین و هوش مصنوعی از روش هایی مانند طبقه بندی، خوشه بندی و تحلیل رگرسیون برای تحلیل داده ها بهره می‌گیرند. مبانی آماری به الگوریتم ها کمک می‌کند تا از داده ها آموزش ببینند و به مرور عملکرد رو به رشدی داشته باشند.
۳. امنیت اطلاعات: تحلیل الگوهای نفوذ و آشکار سازی تهدیدات امنیتی از طریق علم آمار امکان پذیر می‌شود. به بیان دیگر در بحث امنیت سایبری با رصد الگوهای ترافیک شبکه و رفتار کاربران علاوه بر شناسایی حملات سایبری می‌توان از وقوع آن نیز پیشگیری کرد.
٤. پردازش تصویر: در تشخیص الگوها، شناسایی اشیاء، ترجمه تصاویر و تحلیل تصاویر مربوط به پردازش تصویر، آمار نقش بسیار مهمی را ایفا

می‌کند. موارد متعددی نظیر کوواریانس، مد، میانه، چولگی، کشیدگی، میانگین هارمونیک، انحراف استاندارد و.... برای پردازش تصویر استفاده می‌شوند. پردازش تصویر می‌تواند با یک یا چند متد آمار صورت گیرد. به عنوان مثال میانه، پیکسل های با شدت بالاتر را از پیکسل های با شدت پایین تر جدا می‌کند. به نوعی فرآیند مرتب سازی را انجام می‌دهد.

۵. شبکه‌های اجتماعی و تحلیل داده های اجتماعی: بررسی الگوهای رفتاری، پردازش داده ها از شبکه های اجتماعی و تحلیل داده‌های آنها از کاربردهای دیگر آمار در زمینه کامپیوتر است. مثلاً از روش خوشه بندی برای شناسایی گروه های مرتبط به هم استفاده می‌شود. تحلیل توصیفی، مدل سازی گراف، مدل های رگرسیون از جمله روش های آماری دیگری هستند که برای تحلیل شبکه های اجتماعی کاربرد دارند.

منابع:

- https://www.tpbin.com/jarticle/20220811102542511106f3c9b
- https://www.apu.apus.edu/area-of-study/information-technology/resources/computer-science-and-statistics-exploring-intersections/#:~:text=In%20computer%20science%2C%20statistics%20are,the%20advancement%20of%20computer%20science
- https://www.researchgate.net/publication/۲۸۷۱۰۹۹۲۲_Significant_Role_of_Statistics_in_Computational_Sciences

گاهنامه انجمن علمی دانشجویی کامپیوتر زیر نظر معاونت فرهنگی دانشجویی موسسه آموزش عالی ارشاد دماوند

استاد مشاور:

خانم مهندس مریم تیموری، دانشجوی دکترای هوش مصنوعی، استاد گروه آموزشی کامپیوتر موسسه آ.ع ارشاد دماوند

اعضای انجمن:

ساغر سیدپور (دبیر انجمن)

نگار ابوالقاسمی

مریم عبدالهاشمی

مبینا اشکبار

نازنین فاطمه ترکمن

نویسندگان این مگیت:

مریم عبدالهاشمی

مبینا اشکبار

نازنین فاطمه ترکمن

سردبیر نشریه:

نازنین فاطمه ترکمن