Znaczenie inżynierii natychmiastowej w generatywnych modelach AI

Streszczenie

Dla efektywności modelu w generatywnej sztucznej inteligencji, inżynieria natychmiastowa i projektowanie odgrywają bardzo ważną rolę. W kontekście wpływu modelu i inżynierii natychmiastowej na zachowanie, poddziedzina uczenia maszynowego i przetwarzania języka naturalnego (NLP) odgrywa istotną rolę w określaniu wyników modelu. W tym rozdziale omówiono solidność, znaczenie interpretowalności wydajności oraz sposoby jej udoskonalenia. Pierwsza część książki dotyczy omawianych technik, zasad i idei. Trafność, pomysłowość i spójność to niezbędne parametry wejściowe dla danej funkcji. Aby sprostać zadaniom i celom, omówiono związek między projektowaniem natychmiastowym a możliwościami i złożonością relacji. W rozdziale określono również różne techniki szybkiego tworzenia i ograniczania metod lingwistycznych z wykorzystaniem szablonów i wskazówek dotyczących konkretnej dziedziny. W tym rozdziale pokazano, jak badana jest interoperacyjność modelu i ograniczanie stronniczości w inżynierii natychmiastowej. W tym rozdziale omówiono również przejrzystość i rozpoznawanie stronniczości w systemie AI. W tym rozdziale omówiono zagadnienia związane z generowaniem tekstu, syntezą obrazu oraz agentami konwersacyjnymi w czasie rzeczywistym i studiami przypadków. Omówiono wyzwania i przyszłe kierunki rozwoju inżynierii natychmiastowej.

Słowa kluczowe: podpowiedzi oparte na szablonach, podpowiedzi oparte na ograniczeniach, uczenie się przez wzmacnianie, względy etyczne, ochrona prywatności i minimalizacja uprzedzeń, inżynieria natychmiastowa, GenAI

Wprowadzenie

Definiowanie podpowiedzi w generatywnej sztucznej inteligencji (AI)

Inżynieria podpowiedzi to podzbiór generatywnej sztucznej inteligencji (AI), który ma na celu poznanie kontekstu, zrozumienie danych wejściowych modelu i stworzenie odpowiedniego wyniku. Zazwyczaj zaczyna się od podpowiedzi, które mają na celu podanie instrukcji lub wskazówek, które mają zbudować model. Podpowiedzi te są niezwykle ważne i koncentrują się na tym, na czym należy się skupić, a na czym nie, aby zapewnić spójność podpowiedzi. Dane wejściowe dla inżynierii podpowiedzi występują w różnych formatach, takich jak obrazy, tekst i wideo . W dziedzinie inżynierii podpowiedzi generowanie tekstu odgrywa ważną rolę. Jeśli chcesz napisać wiersz o naturze, informacja podana w podpowiedzi brzmi: „Wygeneruj wiersz oddający majestat natury”. Sztuczna inteligencja wygeneruje wiersz o naturze w bardziej realistycznym stylu. Synteza obrazu to kolejny przykład. Podpowiedzi określają rodzaj obrazu danych, który chcesz wygenerować. Załóżmy na przykład, że chcemy stworzyć nowy świeży obraz z zachodem słońca nad jeziorem, wtedy podpowiedź będzie taka: „Stwórz obraz uchwycający spokój zachodu słońca nad spokojnym jeziorem”. Wtedy AI użyje tego opisu, aby namalować obraz dokładnie tego, co miałeś na myśli . Zamiast używać pisania lub generowania obrazów, używa się generowania tekstu. Słowa kluczowe są podane w podpowiedzi do generowania treści. Możemy nawet podzielić się naszymi zainteresowaniami i hobby z systemem. Następnie AI generuje pytanie, pytając o twoją historię i rozpoczyna rozmowę od tego momentu . Generatywna AI wyprodukuje treść i odpowiednie informacje na podstawie generatywnej AI . Podpowiedzi są ważne, ponieważ bez ich znajomości nie możemy rozmawiać dalej.Tabela opisuje porównanie różnych inżynierii podpowiedzi.

Prompt type	Description	Examples	Advantages	Limitations
Open-ended	Allows for unrestricted user input	"Write a story about a mysterious island"	Encourages creativity and originality	May result in varied or divergent responses
Constrained	Imposes specific guidelines or restrictions	"Write a poem with the theme of love"	Provides structure and focus	Limits creative freedom
Image-based	Uses visual stimuli as prompts	Provides an image as inspiration	Sparks visual imagination	Relies on the interpretation of visuals
Interactive	Engages users in a dialogue or conversation	"Tell me about your favorite childhood memory"	Facilitates dynamic interaction	Requires natural language understanding

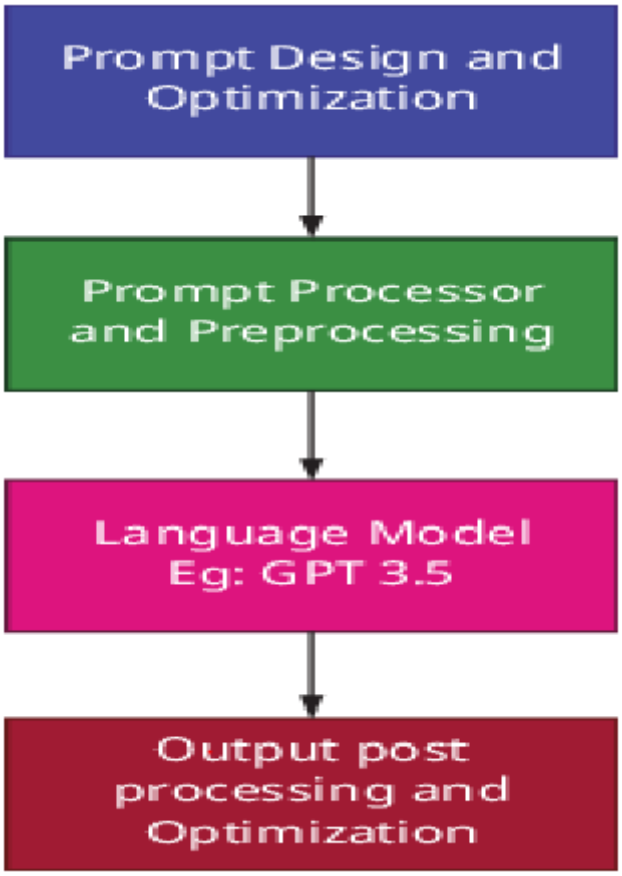
Rozwój inżynierii natychmiastowej

Dzięki postępowi w badaniach, interakcji człowiek-maszyna i inteligencji obliczeniowej w dziedzinie generatywnej sztucznej inteligencji (AI), nastąpił rozwój inżynierii natychmiastowej . Jednym z rozwijających się obszarów jest inżynieria natychmiastowa . Wczesne etapy: W przypadku inżynierii natychmiastowej, modelowanie języka i inne metody przetwarzania języka naturalnego (NLP) wzbogacają obszar inżynierii natychmiastowej. Na wczesnym etapie system oparty na AI otrzymuje proste dane tekstowe i różne instrukcje. Stanowi to podstawę inżynierii natychmiastowej do rozwiązywania złożonych danych . Pojawienie się dużych modeli językowych: Wielkoskalowe modele językowe odgrywają drugą ważną rolę w inżynierii natychmiastowej. Głównymi przykładami są LLM firmy Google, takie jak BERT (Bidirectional Encoder Representations from Transformers) oraz seria GPT (Generative Pretrained Transformer) firmy OpenAI . Modele te są trenowane z wykorzystaniem ogromnych ilości danych, aby tworzyć treść w języku naturalnym i w sposób swobodny poprawiać wydajność modelu. Ulepszanie metod projektowania natychmiastowego: W przypadku formułowania natychmiastowego badacze projektu aktywnie uczestniczą w opracowywaniu modelu. Tworzenie zachowania modelu jest prostym i zaawansowanym algorytmem wykorzystującym podejścia oparte na szablonach, które zawierają predefiniowane szablony i struktury]. Zgodnie z ustalonymi standardami i w celu uzyskania wyników, praktycy stosują techniki oparte na ograniczeniach . Integracja iteracyjnego udoskonalania i ludzkiej informacji zwrotnej: Kluczowym osiągnięciem w inżynierii natychmiastowej jest ulepszenie i połączenie informacji zwrotnej od człowieka i metody iteracyjnej. Badacze odkrywają trudności w gromadzeniu informacji zwrotnych i iteracyjnym ulepszaniu formułowania inżynierii natychmiastowej w oparciu o dane użytkownika i rzeczywiste wykorzystanie . Dzięki znaczeniu inżynierii natychmiastowej, jest ona bardziej przyjazna dla użytkownika i kontekstualna, co zwiększa jego zadowolenie i wydajność. Dywersyfikacja zastosowań: Inżynieria błyskawiczna jest wykorzystywana w różnych dziedzinach, poza generowaniem tekstu. Jest ona wykorzystywana w różnych dziedzinach sztuki, a jej implikacje obejmują przykładowe komponowanie muzyki w czasie rzeczywistym, syntezę obrazu, generowanie kodu i tworzenie dialogów. Zwiększa to rozwój technologiczny i udoskonalanie tych technik. Przyszłe kierunki: Naukowcy odkrywają potencjalne zastosowanie inżynierii błyskawicznej oraz ograniczenia, wyzwania i luki do rozwiązania w przyszłości. Zidentyfikowane luki polegają na dostrajaniu komend do rozwiązywania złożonych danych, zrozumieniu modelu i zapewnieniu kompleksowego wyglądu komendy w oparciu o sytuację kontekstową i preferencje użytkownika . Inne obszary badań to interpretowalność danych, ograniczanie stronniczości danych oraz etyczne aspekty ewoluujących danych.

Teoretyczne podstawy inżynierii promptów

Projektowanie promptów i teoria lingwistyczna

Rozwój promptów opiera się na teorii teorii lingwistycznej oraz jej złożoności językowej i komunikacji. Inżynieria promptów może wykorzystywać zachowanie systemu i sensowne interakcje z wykorzystaniem predefiniowanej struktury języka, semantyki danych i pragmatyki danych z teorii lingwistycznej. Ta część rozdziału ukazuje powiązania między teorią lingwistyczną a projektowaniem promptów, wyjaśniając fundamentalne idee i koncepcje, które stanowią podstawę teoretycznych ram inżynierii promptów . Rysunek przedstawia udoskonalone wyniki modelu językowego uzyskane dzięki szybkiej inżynierii i przetwarzaniu końcowemu.



Składnia i semantyka: Podstawą teorii lingwistyki jest kontrola języka i zrozumienie jego znaczenia. Główną metodą jest analiza semantyczna; zajmuje się ona sposobem uporządkowania błędów gramatycznych i fraz. Do stworzenia jednoznacznej i precyzyjnej inżynierii komunikatów niezbędna jest składnia i semantyka. Proces generowania i wyprowadzania to spójny językowo sposób tworzenia kontekstu i zadań. Spójność dyskursu i pragmatyka: Pragmatyka to kolejny ważny aspekt analizy języka, a cele komunikacji są istotne wraz z wypowiedziami językowymi. Inżynieria komunikatów czerpie dane wejściowe z pragmatyki. Modele generyczne są ważne dla kontekstowego generowania odpowiednich danych i spójnego udzielania odpowiedzi z wykorzystaniem czynników kontekstowych, implikacji konwersacyjnych i intencji mówiącego. Struktura dyskursu i relacje spójności to projektowanie komunikatów w celu generowania spójnych danych i logicznego wytwarzania komunikatów . Struktura informacji i analiza semantyczna: Analiza semantyczna jest podzbiorem lingwistyki obliczeniowej, który służy do przedstawiania znaczenia w sposób sformatowany i ustrukturyzowany oraz do ekstrakcji znaczenia z tekstu w języku naturalnym. Ta technika analizy semantycznej jest wykorzystywana w inżynierii komunikatów do trafnej ekstrakcji danych i poznania ich znaczenia, celów zadania, kontekstu

danych oraz inkluzywności ograniczeń danych. Interpretowalność i skuteczność osiąga się poprzez integrację Gen AI z projektowaniem komunikatów i analizą semantyczną. Komunikaty są prezentowane i organizowane w lepszy sposób, co sprzyja efektywnej komunikacji i rządzi zasadami, koncentrując się na temacie i dyskursie wyeksponowania [→26]. Rozważania międzyjęzykowe i międzykulturowe: Przed zaprojektowaniem komunikatu znane są różne kultury, a języki przez nie używane są analizowane i rozpoznawane przez projektanta „komunikatu". Dzięki teorii lingwistycznej i znajomości języków w różnorodności jesteśmy w stanie ocenić uniwersalną dostępność danych, a zmienność danych jest znana . Aby zapewnić łatwą dostępność i inkluzywność danych, uwzględniono związek między różnorodnością językową a parametrami międzykulturowymi.

Kognitywistyka

Projektując podpowiedź, powinniśmy pamiętać o jej użyteczności dla danej osoby. Możemy zastanowić się, jak taki projekt poprawia rezultaty i jak dokładnie znany jest sposób wyszukiwania danych przed jego zaprojektowaniem. Omówimy kilka ważnych parametrów w kognitywistyce .

1. Obciążenie poznawcze i pamięć robocza: Aby przezwyciężyć ograniczenia mocy ludzkiego mózgu, pamięci i zapamiętywania danych, projektowanie podpowiedzi koncentruje się na tych obszarach. Projektując podpowiedzi w sposób prostszy i bardziej zwięzły, podpowiedź jest użyteczna dla użytkownika i pozwala na generowanie dokładnych danych, a także umożliwia ich dokładną ekstrakcję, a odpowiedzi są szybkie i łatwe do odczytania .

2. Modele mentalne i mapowanie konceptualne: Aby zrozumieć model, mózg tworzy modele mentalne. Za pomocą tego modelu mentalnego sztuczna inteligencja może szybko i łatwo przewidzieć i zrozumieć, co próbujemy powiedzieć. Model ten może łatwo dopasować naszą dokładną intencję do potrzebnego zapytania.

3. Błędy poznawcze i heurystyki: Czasami to, co myślimy w umyśle, generuje wyodrębnione informacje w czasie rzeczywistym. Dlatego projektując model, możemy przezwyciężyć to skrzywienie, podając dokładne dane. Na przykład, gdy ktoś zna argumenty jednej strony, ten model przewidzi dane z innej perspektywy. Dzięki temu dane będą tworzone w sposób bardziej zrównoważony i wiarygodny.

4. Metapoznanie i samoregulacja: Aby wyciągnąć wnioski i rozwiązać problem, musimy zastanowić się, jak będziemy myśleć i jak możemy na niego zareagować. Aby uzyskać strategie i przemyślenia w inżynierii podpowiedzi, możemy zbudować podpowiedź, wykorzystując model. Aby uzyskać lepszy wynik, prosimy system o pobranie danych wejściowych i łatwe pozyskanie danych z systemu.

5. Zasady interakcji człowiek-komputer (HCI): Projektując nową podpowiedź, powinniśmy pamiętać, że wszystkie tworzone dane są idealne, łatwe w użyciu i spójne. Model HCI jest implementowany podczas tworzenia projektu. Informacje zwrotne w czasie rzeczywistym i adaptacyjne są wykorzystywane w procesie tworzenia danych, co sprawia, że system jest bardziej satysfakcjonujący i komfortowy dla użytkownika.

Podejścia lingwistyki obliczeniowej w tworzeniu podpowiedzi dla modeli AI

Tekst w języku naturalnym odgrywa ważną rolę w tworzeniu danych, które służą głównie do zrozumienia przez podpowiedź, jakie dane wejściowe i jakie dane wyjściowe ma ona dostarczyć, co jest dokładnie znane projektantowi. Model ten sprawia, że projekt działa sprawnie, dostarczając dokładne wyniki. W lingwistyce obliczeniowej istnieje kilka metod ulepszania podpowiedzi.

1. Techniki przetwarzania języka naturalnego (NLP): NLP jest główną koncepcją tworzenia podpowiedzi, ułatwiającą zrozumienie tekstu. Tokenizacja to jedna z głównych koncepcji NLP, która

dzieli dane na krótkie teksty i przekształca je w odpowiednie znaki interpunkcyjne i słowa. Pomaga to projektantowi i użytkownikowi podpowiedzi w zaprojektowaniu jej w odpowiedni sposób, z łatwą strukturą podpowiedzi i danymi.

2. Analiza i reprezentacja semantyczna: Aby wydobyć znaczenie w sposób ustrukturyzowany, stosuje się koncepcję zwaną grafem semantycznym, która pozwala na wyszukiwanie lub odczytywanie danych w ustrukturyzowany i znany sposób, a także na poznanie znaczenia i powiązań między tekstem. Koncepcja ta służy do zbudowania modelu w bardziej ustrukturyzowany sposób.

3. Analiza sentymentu i wykrywanie emocji: Projektując podpowiedź, powinniśmy znać stan emocjonalny użytkownika i to, w jakim stanie korzysta z podpowiedzi. Na przykład: jeśli użytkownik korzystający z podpowiedzi jest w stanie szczęścia lub gniewu, jego zdolności emocjonalne i umysłowe są rozumiane, a dane są odpowiednio przekazywane.

4. Rozpoznawanie jednostek nazwanych (NER) i ekstrakcja informacji: NER służy do identyfikacji obiektów, a także do identyfikacji obiektów takich jak miejsca, tekst i organizacja na podstawie podpowiedzi. Poznanie tych elementów i identyfikacja frazy w podpowiedzi inżynieryjnej w odniesieniu do danych.

5. Analiza dyskursu i modelowanie spójności: Techniki te są przydatne w identyfikacji różnych części tekstu oraz w określaniu ich powiązań i dopasowania do danych. Dane te służą do identyfikacji danych i łatwego śledzenia ich. Na przykład, wykorzystanie znaczników dyskursu i struktury retorycznej do zapewnienia spójności danych w ustrukturyzowanej odpowiedzi.

6. Tłumaczenie maszynowe i przetwarzanie wielojęzyczne: Ma to na celu stworzenie modelu w różnych językach i w różnych formach. Jest to pomocne w projektowaniu modelu w różnych językach i w różnych formatach. Na przykład, jeśli ktoś zna język hiszpański, łatwo będzie zrozumieć ten model.

Metodologie w inżynierii promptów

Prompt oparty na szablonach

Głównym modelem wykorzystania inżynierii promptów jest prompt oparty na szablonach i jest to model podstawowy. Jest on bardzo popularny ze względu na swoją elastyczność, prostotę i spójność danych oraz szeroki zakres zastosowań. Za pomocą szablonów możemy tworzyć rusztowania (scaffold). Wiemy, dlaczego budujemy model, a struktura jest obsługiwana, co jest bardzo popularne w inżynierii promptów. Aby stworzyć prompt bardziej spersonalizowany, tworzymy dane i wypełniamy nimi szablony. Pierwszym krokiem jest inscenizacja szablonu, która pozwala na utworzenie szablonu i wygenerowanie danych w bardziej adekwatny sposób. Przykładem danych jest generowanie tekstu i promptu dialogowego. Projektując prompt, należy zawsze brać pod uwagę konkretną aplikację, dla której projektujemy, cel, jaki chce osiągnąć użytkownik, język, którego będziemy używać, a także używane słowa kluczowe i tematy. Możemy również wykorzystać zmienne zastępcze, dane i inną zawartość danych. Stosujemy różne techniki, takie jak parametryzacja i abstrakcja, służące do przenoszenia danych do innego wiersza poleceń. Wizualizacja oparta na szablonach jest wykorzystywana w różnych procesach generowania tekstu i tworzenia treści, takich jak pisanie historii i podsumowywanie danych. Dzięki temu możemy tworzyć pliki audio, wideo i obrazy, a także pomaga to w bardziej naturalnym wdrażaniu modelu i budowaniu danych. Ma to zarówno zalety, jak i wady. Jest bardzo prosta, spójna i elastyczna, a wyzwania polegają na tym, że czasami pokazuje tylko ograniczoną jakość szablonu, a generowane dane oznaczają, że nie wygenerują dokładnych danych. To przyszłościowa technika ulepszania modelu. Możemy używać modelu jako szablonu zagnieżdżonego i szablonu hierarchicznie.

Projektowanie wiersza poleceń opartego na ograniczeniach

Do określania reguł i zasad używa się wiersza poleceń opartego na ograniczeniach. Służy on do tworzenia celów i reguł danych. Kiedy tworzymy obraz, samo podpowiedź dotyczy rodzaju koloru, jaki należy zastosować, oraz struktury danych, jakie mają zostać użyte. W tym modelu ograniczenia i metody są uwzględnione w modelu opartym na ograniczeniach i pomagają w tworzeniu reguł i przepisów, a także służą do tworzenia danych o wysokiej trafności i ich tworzenia. Jeśli podajemy dokładne dane i ograniczenie, zapewnia to dokładną jakość i trafność. Istnieją różne sposoby formułowania modelu i projektowania go w oparciu o ograniczenie. Przed rozpoczęciem pracy nad modelem musimy zdecydować, dla jakich konkretnych aplikacji pracujemy. Jest on wykorzystywany do generowania tekstów, tworzenia filmów i chatbotów. W generowaniu tekstów teksty są generowane w bardziej ograniczony sposób, aby generować dane w sposób prawidłowy. Podsumowując, ważne punkty są zapamiętywane, a reszta danych jest zapominana, a zapamiętywanie danych i chatbotów jest łatwe lub trudne. Dzięki wykorzystaniu danych tworzone treści są bardziej wiarygodne i rzetelne, a model opiera się na ograniczeniach. Przy większych ograniczeniach sztuczna inteligencja nie może uzyskać dużego wyniku; wygeneruje błąd. Zatem dokładne określenie ograniczenia daje prawidłowe wyniki. Interaktywne i adaptowalne dane są bardziej niezawodne w tworzeniu i generowaniu danych. Dzięki temu system jest bardziej interaktywny i generuje dane w sposób bardziej wiarygodny. Uczenie ze wzmocnieniem (RL) jest wykorzystywane do dokładniejszego tworzenia modelu; na przykład, jeśli instruujesz dane, aby posprzątały pokój i prosisz je o posprzątanie pokoju, rozumieją one dane dokładniej. Z pomocą RL jest ono wykorzystywane do karania modelu lub do przyznawania nagród w bardziej precyzyjny sposób.

Techniki uczenia się przez wzmacnianie (RL)

Uczenie się przez wzmacnianie (RL) służy do tworzenia danych i generowania tekstu w bardziej efektywny sposób. Aby tworzyć dane i RL w bardziej angażujący i interesujący sposób, można użyć RL, aby podpowiedzi były bardziej angażujące i interesujące dla osób czytających historie. Istnieją pewne zasady i metodologie korzystania z RL w inżynierii podpowiedzi, ale tutaj postawimy na prostotę. Główną ideą jest wykorzystanie agentów RL, takich jak generatory podpowiedzi lub ewaluatory, do interakcji z modelem AI i uczenia się z niego. Środowiskiem jest sam model, a agent stara się znaleźć sposób na dostarczenie mu podpowiedzi, które sprawią, że zrobi to, czego od niego oczekujesz, na przykład wygeneruje spójny tekst lub przeprowadzi wartościową konwersację. Istnieją różne sposoby wykorzystania RL w inżynierii podpowiedzi, takie jak metody gradientu polityki, Q-learning i architektury aktor-krytyk. Wszystkie działają na różne sposoby, ale wszystkie pomagają agentowi nauczyć się, jak dawać najlepsze podpowiedzi modelowi AI. RL może być używane do najróżniejszych rzeczy, takich jak generowanie tekstu, systemy dialogowe, a nawet tworzenie treści multimedialnych, takich jak obrazy, filmy lub muzyka. Jest dość wszechstronne! Ale, jak w przypadku każdego narzędzia, istnieją pewne wyzwania i ograniczenia w używaniu RL do szybkiego inżynierowania. Czasami trudno jest ustalić najlepszy sposób nagrodzenia modelu i mogą wystąpić problemy z agentem utkniętym w lokalnych optimach lub nie badającym wystarczającej liczby opcji . W przyszłości jednak uważamy, że RL może być jeszcze lepszy w szybkim inżynierowaniu! Być może zobaczymy nowe algorytmy, które mogą uczyć się szybciej lub eksplorować wydajniej, a może dowiemy się, jak je łączyć.

Studia empiryczne i przykłady przypadków

Generowanie tekstu: Podpowiedzi do kreatywnego pisania

W tej sekcji omówiono studia przypadków inżynieryjnych i przykłady podpowiedzi.

Studia empiryczne

Podpowiedzi i kreatywność: Badacze prowadzą badania, aby napisać historię w formie narracji i stworzyć wiersz. Są one wykorzystywane do tworzenia treści, tworzenia danych i ich przetwarzania. Preferencje użytkowników i opinie: Podpowiedzi do tworzenia danych i ich jakości [→38].

Przykłady przypadków

Warsztaty pisarskie: Podczas warsztatów wykorzystuje się je do tworzenia danych. Polegają one na podpowiadaniu pytań, takich jak „kiedy czujesz się szczęśliwy", i rytmicznym pisaniu wiersza. Są one wykorzystywane do tworzenia informacji zwrotnej i wyzwań, a także do lepszego pisania historii i danych.

Kluczowe wnioski

Podpowiedzi są ważne dla dokładnego zapisu danych w celu tworzenia treści. Podpowiedzi są wykorzystywane w różnych rodzajach tekstów. Do tworzenia jasnych i interesujących danych stosuje się podpowiedzi w sposób interesujący. Ta metoda służy lepszemu poznaniu danych.

Przyszłe badania

Prowadzi się wiele badań nad tym, jak wprowadzać dane do podpowiedzi i jak wpływa to na tekst, a także jak zadawać pytania w podpowiedzi. Kwestie te nie są jeszcze znane. Prowadzony jest szeroki obszar badań nad dostosowywaniem podpowiedzi. Kolejnym przyszłym obszarem badań jest wprowadzanie danych za pomocą obrazów i dźwięków w inżynierii podpowiedzi.

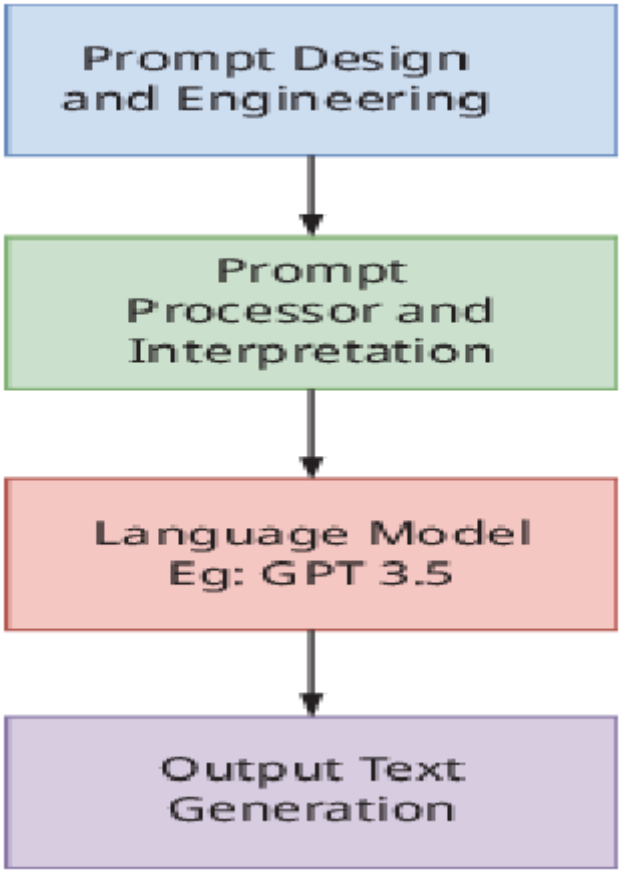
Badania empiryczne i ocena technik inżynierii natychmiastowej

Skuteczność metodologii i technik odgrywa kluczową rolę w badaniach empirycznych i rozwoju dziedziny inżynierii natychmiastowej. W tej sekcji omówiono oceny i efekty, przyszłe badania, wyzwania i ograniczenia. Na podstawie badania omówiono słabe i mocne strony techniki natychmiastowej, a także scenariusz rzeczywisty. Aby wygenerować trafne, spójne i wysokiej jakości podejścia, badacze testują je na wysokiej jakości zbiorze danych. Badania empiryczne potwierdzają założenia dotyczące zachowań użytkowników, wydajności modelu danych i skuteczności projektowania technik natychmiastowych. Niektóre z wyzwań, z jakimi borykają się badania empiryczne, to:

1. Metryki i benchmark w subiektywnej technice oceny inżynierii natychmiastowej.

2. Standaryzacja i benchmark są trudne w przypadku ocen inżynierii natychmiastowej. W zależności od różnorodności i jakości zbioru danych, wydajność sztucznej inteligencji generacji może się różnić.

3. Badania użytkowników wymagają wydajności zadań i wpływu sztucznej inteligencji generacji, co powoduje zużycie czasu i zasobów. Rysunek 2 przedstawia znaczenie inżynierii natychmiastowej w modelu językowym.



Aby pokonać ograniczenia i wyzwania, niezbędna jest skalowalność, interoperacyjność i ulepszanie zestawów danych testowych. Przetwarzanie języka naturalnego to nadchodząca metoda, która koncentruje się na zaawansowanych narzędziach i technikach uczenia maszynowego. Współpraca interdyscyplinarna jest niezbędna w kontekście badaczy zajmujących się inżynierią natychmiastową, psychologią oraz interakcją człowiek-komputer w efektywnym generowaniu tekstu i podejmowaniu decyzji.

Destylacja wiedzy w inżynierii szybkiego dostępu

Destylacja wiedzy to nowa metoda i technika w uczeniu maszynowym. Opiera się na koncepcji pozyskiwania wiedzy z dużego, skomplikowanego modelu (nauczyciela) i przekazywania jej mniejszemu, bardziej wydajnemu modelowi (uczniowi). Efektywność i szybkość działania modelu są osiągane poprzez model. Podstawowa idea destylacji wiedzy jest prosta: model ucznia jest szybszy niż model nauczyciela. Dane wyjściowe modelu nauczyciela są wykorzystywane do przekazywania danych wejściowych do modelu ucznia. Do podejmowania decyzji wykorzystywany jest model nauczyciela.

1. Architektura nauczyciel–uczeń: Model nauczyciela to bardzo duży model, który jest trenowany z dużą ilością danych, podczas gdy model ucznia to tylko niewielka ilość danych i pracuje z modelem nauczyciela w zespole.

2. Funkcje strat destylacji: Model ucznia uczy się od modelu nauczyciela, gdzie jest trenowany z danymi wyjściowymi obu modeli, a uczeń zdobywa wiedzę o modelu nauczyciela.

3. Skalowanie temperatury: Dane wyjściowe modelu nauczyciela są podawane bardziej szczegółowo i czasami generują określone dane wyjściowe. W tym celu stosujemy technikę skalowania zwaną

„zmiękczaniem” z wykorzystaniem parametru „temperatura”. Model nauczyciela działa jak przewodnik, w którym model ucznia podąża za danymi wejściowymi i zasadami przewodnika.

Destylacja wiedzy w inżynierii komunikatów ma kilka ciekawych zastosowań

1. Kompresja modelu: Moc obliczeniowa i wydajność modelu ucznia są efektywnie trenowane poprzez trenowanie danych z wykorzystaniem większej ilości danych modelu nauczyciela. Pomaga to we wdrażaniu modelu przy użyciu niewielkich zasobów, takich jak drony i smartfony.

2. Transfer uczenia: To jak zdobywanie wiedzy i przenoszenie jej do innego modelu. W tym modelu służy to do wydajniejszego tworzenia modelu. W destylacji wiedzy jest to jak pozyskiwanie wiedzy o modelu szkoleniowym od nauczyciela i przekazywanie jej modelowi ucznia.

3. Personalizacja modelu: W konkretnym przypadku użycia tworzone dane są bardziej efektywne i wstępnie trenowane, podczas gdy model ucznia jest trenowany w oparciu o model nauczyciela, dla określonych celów i użytkownika.

Wyzwania

1. Równoważenie złożoności modelu i wydajności: Musimy uzyskać model w zrównoważony sposób. Gdy model jest prosty, jest to bardzo łatwe, a gdy jest skomplikowany, jest to dość trudne, ale jest to bardzo wydajne. 2. Generalizacja i odporność: Opracowując model ucznia, powinniśmy mieć świadomość, że działa on we wszystkich scenariuszach, zarówno w świecie rzeczywistym, jak i w laboratorium. Musimy opracować model, który wygeneruje dokładne dane.

3. Interpretowalny transfer wiedzy: W tym kontekście znany jest sposób, w jaki dane są przesyłane z modelu ucznia do modelu nauczyciela, a także sposób, w jaki model jest trenowany i jak działa w oparciu o ten transfer wiedzy.

Badanie wpływu podpowiedzi

Poglądy multidyscyplinarne

Perspektywy poznawcze dotyczące interakcji podpowiedzi z modelem

Błędy poznawcze i ramowanie: Błąd zakotwiczenia lub potwierdzenia to błąd generowany przez nieistotne pytania, który wpływa również na mentalność użytkownika. Modele mentalne i oczekiwania: Tworzone są możliwości i model, a także znany jest sposób ich interakcji. Motywacja i zaangażowanie: Motywacja użytkownika i projekt są znane, a powód ich powstania jest w znacznym stopniu znany podczas korzystania z inżynierii podpowiedzi.

Znaczenie projektowania podpowiedzi w socjologii

Utrwalanie uprzedzeń – Rasa, płeć i uprzedzenia społeczne są znane, a dane i ich przyczyny są znane podczas reakcji podpowiedzi. Normy społeczne i wartości – Sprawdzany jest wpływ danych wyjściowych na użytkownika, sposób ich włączenia i sposób reakcji. Narracje kulturowe i tożsamość – Sposób, w jaki tożsamości i inicjatywy kulturowe są znane, a także sposób, w jaki dane są przekazywane, jest identyfikowany i znany.

Interakcja człowiek-komputer

Dla łatwości użytkowania, informacji zwrotnej i przejrzystości, zawsze, gdy korzystamy z projektu podpowiedzi, powinniśmy pamiętać o następujących kwestiach:

* Rozpoznawanie celów i potrzeb użytkownika: Aby spełnić cele i wymagania, musimy zmodyfikować zawartość danych.

* Projektowanie komunikatów z uwzględnieniem interakcji z człowiekiem: Aby ulepszyć projekt komunikatu, powinniśmy postępować zgodnie ze strategiami projektowymi.

Implikacje etyczne i społeczne

W tej sekcji omówiono kwestie etyczne i wpływ społeczeństwa:

Strategie ograniczania stronniczości w inżynierii poleceń

Wykrywanie stronniczości w danych szkoleniowych: Aby ograniczyć stronniczość i zidentyfikować dane, należy pamiętać o ryzyku i błędach, które mogą wystąpić w czasie rzeczywistym. Tworzenie uczciwych i inkluzywnych poleceń: Możemy stworzyć model generujący dane inkluzywnie, bez zatrzymywania danych strategie projektowania danych i stereotypy danych. Techniki usuwania stronniczości: Możemy wykorzystać różne przykłady i błędy, aby uznać je za mniej stronnicze w inżynierii poleceń.

Zagadnienia dotyczące prywatności i ochrony danych

Polecenia jako dane osobowe: Podczas projektowania polecenia brane są pod uwagę informacje wrażliwe i preferencje. Bezpieczeństwo i własność danych: Podczas projektowania polecenia brane jest pod uwagę bezpieczeństwo danych. Brane jest pod uwagę prawo własności polecenia i to, kto odpowiada za opracowanie danych. Przejrzystość w zakresie wykorzystania danych: sposób przechowywania monitów i danych jest wyraźnie pokazywany, a przejrzystość jest zachowana.

Wytyczne etyczne dotyczące odpowiedzialnego projektowania promptów

Tworzenie ram etycznych: Należy przestrzegać wytycznych dotyczących generowania modelu i etycznego wykorzystywania danych. Odpowiedzialność za wyniki generowane przez prompty: Tworzony jest test zaprojektowany przez użytkownika, który powinien uwzględniać osobę odpowiedzialną za prompt. Promowanie sprawiedliwości algorytmicznej: Możemy bezpieczniej wykorzystywać dane algorytmiczne i etycznie je generować. Tabela 2 przedstawia zagadnienia etyczne związane z inżynierią promptów.

Ethical consideration	Description
Algorithmic bias	The potential for prompts to introduce or perpetuate biases present in training data or models.
Privacy protection	Ensuring that prompts do not inadvertently reveal sensitive information or compromise user privacy.
User consent	Obtaining informed consent from users regarding the collection and use of their data for prompts.
Fairness and equity	Ensuring that prompts are fair and equitable, considering the diverse backgrounds and perspectives of users.
Transparency and accountability	Providing transparency into how prompts are generated and how user data is used in prompt-guided generation systems.

Badania empiryczne i ilustrujące studia przypadków

W tej sekcji omówiono praktyczne zastosowania danych:

Międzydomenowe zastosowania inżynierii promptów

Sygnalizacja dla różnych zadań: Szeroki zakres zastosowań obejmuje różne dziedziny, w tym uzupełnianie kodu danych, kreatywne generowanie tekstu danych i tworzenie podpisów do obrazów danych.

Zmiana promptów dla poszczególnych pól: Analiza promptów i modyfikacja poszczególnych branż, takich jak opieka zdrowotna i finanse, oraz dostosowanie ich do specyfiki danej dziedziny i terminologii.

Dokładne dostrajanie na podstawie monitów jest możliwe dzięki trenowaniu danych przy użyciu techniki uczenia transferowego, służącej trenowaniu danych i uczeniu się na podstawie małych ilości danych.

Ocena wpływu podpowiedzi na wydajność modelu

Skuteczność podpowiedzi: Spójność danych, dokładność danych i płynność danych są opisane i skwantyfikowane. Testowanie A/B różnych iteracji podpowiedzi: Korzyści i zastosowania są utrzymywane i znane dzięki testowaniu A/B.

Ocena generowanych wyników przez ludzi: Stosowalność i jakość danych są utrzymywane w ocenie danych i generacji przez ludzi.

Badania nad doświadczeniem użytkownika i analiza informacji zwrotnych

Określanie potrzeb i preferencji użytkowników: Zachowania użytkowników są badane za pomocą projektu podpowiedzi w celu zidentyfikowania dokładnych funkcji i ważnych narzędzi.

Analiza danych wprowadzanych przez użytkownika w komunikatach: Znajdź najlepszy model tworzenia danych i twórz dane w inżynierii komunikatów. Aby projekt był wydajny i przyjazny dla użytkownika, dane są tworzone w sposób bardziej znaczący i wydajny.

Interdyscyplinarne perspektywy inżynierii podpowiedzi

Spostrzeżenia psychologiczne

1. Teoria obciążenia poznawczego: Zbadaj dane i to, jak ta teoria pomaga w projektowaniu podpowiedzi, analizuj zdolności poznawcze i inne dane użytkownika.

2. Modele mentalne: Zbadaj wiedzę użytkownika i jej wpływ na rozumienie danych oraz interakcję z podpowiedzią.

3. Motywacja wewnętrzna: Powinniśmy znaleźć powód użycia podpowiedzi. Analizujemy, czy sprawia ona użytkownikowi przyjemność, czy go zasmuca.

4. Informacja zwrotna i wzmocnienie: Badacze powinni wiedzieć, jak przetwarzane są dane oraz jak analizowana jest adaptacja systemu i zachowanie.

Perspektywy socjologiczne

1. Wrażliwość kulturowa: Zidentyfikuj inkluzywność i punktualność danych, a także dane kulturowe.

2. Dynamika społeczna: gdzie dane, ich analiza i interakcje kształtów są brane pod uwagę, aby zwiększyć ich wpływ, a także zgodność danych i ich znaczenie .

3. Zagadnienia etyczne: niektóre z zagadnień dotyczących danych to prywatność danych, zgoda na dane i stronniczość danych, które mają wpływ społeczny.

Perspektywy HCI

1. Projektowanie zorientowane na użytkownika: kładą nacisk na dostępność danych, ich wykorzystanie i reagowanie na potrzeby użytkownika oraz preferencje dotyczące danych, zgodnie z zasadami projektowania zorientowanego na użytkownika i inżynierii użyteczności.

2. Techniki interaktywnego podpowiadania: niektóre z podpowiedzi, takie jak adaptacyjność podpowiedzi, sprzężenie zwrotne w czasie rzeczywistym i konwersacja agentów, sprawiają, że interakcja danych z generatywnymi modelami sztucznej inteligencji jest bardziej naturalna.

3. Interakcja multimodalna: integracja różnych rodzajów modalności, takich jak ruchy ciała, mowa i dotykowe sprzężenie zwrotne, zapewnia immersyjne i bogatsze dane.

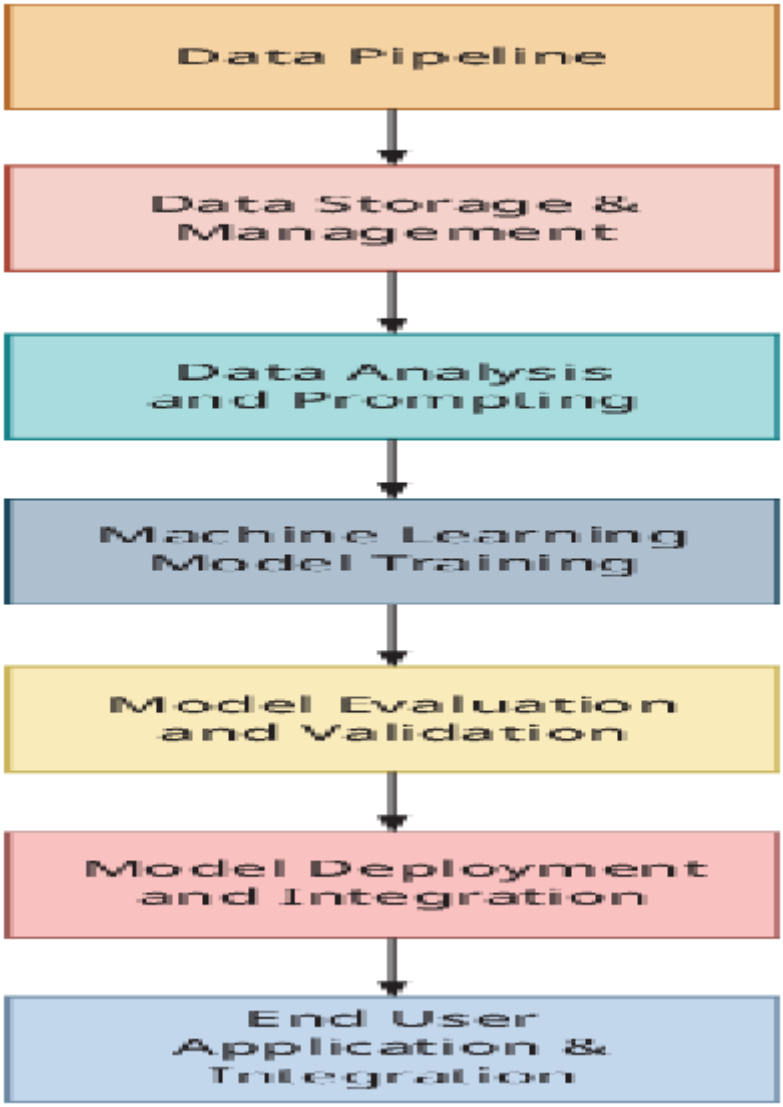
Te multidyscyplinarne perspektywy przyczyniają się do projektowania danych, ich wdrażania i oceny podpowiedzi w tworzeniu systemów AI, uwzględniających kwestie społecznej ewolucji danych.

Przyszłe kierunki i wyzwania

Podpowiedzi multimodalne: Aby zwiększyć interaktywność, dodawanie danych polega na dodawaniu do nich obrazów, danych audio lub innych elementów. Dzięki temu dane stają się ciekawsze i bardziej interaktywne, a samo tworzenie danych staje się bardziej interesujące. Podpowiedzi interaktywne: W przypadku treści dynamicznych i adaptacyjnych podpowiedź informuje, że dane są bardziej interaktywne i generuje dokładne dane. Dzięki temu dane stają się ciekawsze i ciekawsze, a same reagują na interaktywną treść .

Podpowiedzi semantyczne: Przetwarzanie języka naturalnego generuje dane i wyniki na podstawie danych, a podpowiedź podaje odpowiedź z analizą semantyczną i dostarcza dokładne wyniki.

Podpowiedzi spersonalizowane: Tworząc bardziej spersonalizowane dane, powinniśmy poznać je dokładniej i generować dane w sposób bardziej odpowiedzialny i precyzyjny. Możemy określić swoje preferencje i preferencje. Rysunek 3 przedstawia kompleksowy proces pracy w nauce o danych i uczeniu maszynowym



Nowe trendy w inżynierii błyskawicznej

Modele oparte na transformatorach: GPT wykorzystuje uczenie oparte na transformatorach, które jest trenowane na milionach i bilionach danych, a następnie generowane na ich podstawie ogromne ilości, co zwiększa dokładność systemu i precyzję interakcji.

Metauczenie i uczenie się na małej liczbie prób: Działa to, gdy dostępna jest niewielka liczba danych, a do trenowania wykorzystywane są tylko ograniczone dane, które nie są znane przed zastosowaniem tego rodzaju algorytmu.

Uczenie federacyjne: Wiele osób współpracuje ze sobą, nie udostępniając swoich prywatnych danych, aby trenować model z wykorzystaniem wielu różnych danych, bez dokładnego udostępniania danych, co zapewnia bezpieczeństwo danych.

Wytłumaczalna sztuczna inteligencja: Wprowadź techniki wyjaśnialnej sztucznej inteligencji do systemów generowania kierowanych błyskawicznie, aby pomóc użytkownikom zrozumieć, jak sztuczna inteligencja podejmuje decyzje i generuje wyniki. Może to zwiększyć zaufanie i zrozumienie między użytkownikami a sztuczną inteligencją, czyniąc cały proces bardziej przejrzystym i przyjaznym dla użytkownika .

Rozwiązywanie ograniczeń i rozważania etyczne

Ograniczanie błędów: Trwają badania nad projektowaniem komunikatów i starannym ograniczaniem danych. Pozwoli to na bezpieczniejsze i dokładniejsze generowanie danych. Techniki ochrony prywatności: Aby zapewnić prywatność i bezpieczeństwo danych, w uczeniu federacyjnym stosuje się techniki takie jak uczenie federacyjne i bezpieczne obliczenia wielostronne. Służą one do bezpieczniejszego i dokładniejszego generowania danych.

Wytyczne etyczne: Podczas korzystania z komunikatu zachowane są względy etyczne, takie jak przejrzystość, poufność, prywatność i spójność danych, a projekt uwzględnia odpowiednie zasady i regulacje dotyczące tworzenia danych.

Przejrzystość algorytmiczna: Modele sztucznej inteligencji (AI) zapewniają bardziej przejrzyste i dokładniejsze generowanie danych. Aby zapewnić kontrolę i wygodę pracy, dane są wykorzystywane i generowane na podstawie danych użytkownika.

Możliwości badań i współpracy interdyscyplinarnej

Nauki humanistyczne i społeczne: Aby pracować z danymi kulturowymi i społecznymi, nauka i nauki humanistyczne są badane, a szerszy obszar badań będzie kontynuowany w przyszłości. Pozwala nam to lepiej zrozumieć dane i uzyskiwać trafne wyniki.

Kognitywistyka i psychologia: Współpraca z psychologiem i kognitywistą zwiększa efektywność danych, ich interaktywność, a system jest bezpieczniejszy i bardziej poufny.

Interakcja człowiek-komputer (HCI): W przypadku responsywnego tworzenia danych i danych generowanych w dużej ilości, system staje się bardziej otwarty, atrakcyjny i responsywny, co sprzyja szybkiemu tworzeniu projektów. To sprawia, że system jest bardziej atrakcyjny dla tworzenia danych.

Ekspertyza w danej dziedzinie: Aby generować wysokiej jakości dane, możemy gromadzić dane z różnych dziedzin, takich jak sztuka, rysunek i rzeźba, co zwiększa efektywność systemu. 4.8 Perspektywy i trudności

Inżynieria natychmiastowa to stale rozwijająca się i rozwijająca się dziedzina, w której powinniśmy skupić się na tym, co i jak robić, a potencjalne ograniczenia powinny być znane przed rozpoczęciem korzystania z systemu.

Nowe osiągnięcia w dziedzinie szybkiej inżynierii

Uczenie się od skryptu do polecenia: Powinniśmy prowadzić badania, znając gromadzone dane, bez uczenia dużej ilości danych. Wiadomo, że pozwala to na tworzenie bardziej interaktywnych i elastycznych danych.

Optymalizacja metody „bez danych”: W celu tworzenia wysokiej jakości danych i technik, takich jak tworzenie poleceń z małą liczbą i zerową liczbą poleceń, z wykorzystaniem technik projektowania słabych danych .

Interaktywne pętle sprzężenia zwrotnego i polecenia: Gromadzone są dane inżynierii natychmiastowej, a następnie generowane są pętle sprzężenia zwrotnego.

Uwzględnianie kwestii etycznych i ograniczeń

Interpretowalność i wyjaśnialność: W przypadku poleceń i wyników wykorzystywane jest tworzenie danych, a także zachowywana jest ich wyjaśnialność i interpretowalność.

Wyzwanie związane ze zmieniającymi się uprzedzeniami: Uprzedzenia to nieoczekiwane przestrzenie i interaktywność, a uprzedzenia są brane pod uwagę.

Projektowanie dla bezpieczeństwa i ochrony. Nieprawidłowe dane: Podpowiedź projektowa będzie generować dane w sposób zwodniczy.

Możliwości badań i współpracy multidyscyplinarnej

Tworzenie obrazów odbywa się poprzez współpracę w oparciu o dane, interakcję człowiek-komputer i psychologię człowieka:

Sztuczna inteligencja i lingwistyka: Zrozumienie, jak subtelności języka w podpowiedziach wpływają na wyniki modelu, wymaga wiedzy lingwistycznej. Pełny potencjał projektowania podpowiedzi w zakresie manipulacji językiem można wykorzystać poprzez badania multidyscyplinarne.

Etyka i prawo: Solidne ramy etyczne i jasne wytyczne prawne są niezbędne do rozwoju inżynierii podpowiedzi. Współpracując, etycy, prawnicy i badacze sztucznej inteligencji mogą zagwarantować etyczne i rzetelne wykorzystanie tej technologii .

Wnioski

Inżynieria natychmiastowa jest pionierem innowacji w generatywnej sztucznej inteligencji (AI), wykorzystując ustrukturyzowane dane wejściowe do sterowania modelami AI na wiele sposobów. W trakcie naszych badań przekonaliśmy się, jak wszechstronna i potężna może być inżynieria natychmiastowa: przekraczając granice dyscyplin takich jak informatyka, psychologia, socjologia i etyka. Od podstaw teoretycznych po zastosowania praktyczne, inżynieria natychmiastowa pokazała, jak może zmienić sposób, w jaki korzystamy ze sztucznej inteligencji i wchodzimy z nią w interakcje. Inżynieria natychmiastowa to nie tylko dziedzina techniczna, ale także źródło wiedzy z różnych obszarów tej dziedziny. Dzięki temu system jest bardziej przyjazny dla użytkownika i generuje dokładne wyniki. To sprawia, że system jest bardziej efektywny i przyjazny dla użytkownika. Niektóre z działań ograniczających błędy, terminów błędów i nakład pracy są znaczące wraz z rozwojem tego systemu. Badacze muszą bardziej skoncentrować się na tej dziedzinie, aby uzyskać dokładne wyniki. W przyszłości pojawią się różne luki, takie jak multimodalne podpowiedzi danych, interaktywne podpowiedzi oparte na danych oraz spersonalizowane dane dotyczące ludzi i sztucznej inteligencji, które będą współdziałać w ramach multidyscyplinarnej współpracy, zwiększając bezpieczeństwo systemu.