dr Marcin Trepczyński
Instytut Badań Edukacyjnych – Państwowy Instytut Badawczy

Predyspozycje czy zawód – jak wspierać wybory edukacyjne i zawodowe uczniów z wykorzystaniem AI. Doświadczenia projektu ZRK

Abstrakt

Aby sztuczna inteligencja skutecznie wspierała wybory edukacyjne i zawodowe uczniów, konieczne jest rozsądne zaprojektowanie sposobów rozpoznawania ich mocnych stron, predyspozycji, zainteresowań, a także celów związanych z przyszłością zawodową. Modele językowe są tu niezwykle przydatne, jednak na obecnym etapie rozwoju AI jest to wciąż zadanie przede wszystkim dla człowieka, będące zarazem poważnym wyzwaniem. W artykule zaprezentowano wybrane rozwiązania pozwalające na wykorzystanie modeli językowych do rozpoznawanie preferencji i potrzeb użytkowników. Są to doświadczenia zebrane z prac prowadzonych w Instytucie Badań Edukacyjnych – Państwowym Instytucie Badawczym w ramach projektu polegającego na rozwijaniu i modernizowaniu Zintegrowanego Rejestru Kwalifikacji (ZRK), uwzględniających testowanie wielu możliwych podejść w zakresie diagnozowania predyspozycji, potrzeb i celów użytkowników, w tym uczniów, oraz proponowania kwalifikacji, które odpowiadają ich profilom. W artykule przedstawiono przede wszystkim rozwiązania testowane w powstającej obecnie nowej wersji chatbotki Zeterka w zakresie stosowania testu preferencji, identyfikowania zainteresowań, umiejętności, doświadczenia zawodowego, a także rozpoznawania intencji użytkownika w ramach prowadzonej rozmowy.

1. Wprowadzenie

Czy sztuczna inteligencja (dalej: SI) może skutecznie wspierać użytkowników, w tym uczniów, gdy wybierają ścieżki uczenia się lub zawód? Część osób zapewne wierzy, że tak jest. Okazuje się jednak, że wiele zależy od szczegółów.

Różnego rodzaju modele językowe, stanowiące część technologii nazywanych SI, stwarzają w tym zakresie ogromne możliwości. Proste modele, typu Word2Vec, lub nieco bardziej zaawansowane, typu *sentence vector*, umożliwiają np. przeszukiwanie dużych zasobów danych na podstawie bliskoznaczności (semantycznie). Mogą więc uchwycić sens wypowiedzi użytkownika i dopasować najbliższe temu zasoby danych. W najprostszym wypadku wpisanie słowa "rumak" zaowocuje wskazaniem zasobów, w których pojawia się słowo "koń". Natomiast językowe modele generatywne, zwłaszcza duże modele językowe (Large Language Models – LLMs), takie jak słynne modele GPT-4o, Gemini 2.5 czy polskie modele jak Bielik lub PLLuM, potrafią nie tylko swobodnie konwersować z użytkownikiem, lecz także wykonywać szereg zadań nakierowanych na dobranie informacji przydatnych użytkownikowi. Przykłady takich zadań zostaną podane w sekcjach w dalszej części.

Jednak gdy modele wykorzystywane są w tych celach, często otrzymywane odpowiedzi nie są odpowiedziami modelu wprost na zapytanie (ang. *prompt*) użytkownika. Zapytanie to jest zazwyczaj odpowiednio obudowane i wzbogacone dzięki procesom poprzedzającym wysłanie zapytania do modelu oraz uzupełnione o dodatkowy prompt niewidoczny dla użytkownika (tzw. *meta prompt*). Dlaczego tak jest?

Model jedynie zwraca możliwie najlepszą odpowiedź na otrzymane zapytanie – najlepszą zgodnie z tym, jak został wytrenowany. Jeśli LLM został wytrenowany na danych, o które pytamy (np. w zbiorze treningowym znajdowały się opracowania o koniach i pytamy o różne ich rasy), z dużym prawdopodobieństwem poda prawidłową odpowiedź. Gdy jednak poprosimy model o informację, której sam nie posiada (nie uczył się o koniach), będzie on halucynował, czyli zmyślał. Jeśli zaś chcemy, aby wykorzystał informację z zewnętrznego zasobu wiedzy (np. baza informacji o koniach), musimy mu tę informację dostarczyć, dodając ją do zapytania użytkownika. Bywa też, że zależy nam, by w ramach konwersacji model zachowywał się w różny sposób, o czym będzie mowa niżej.

Dlatego też, aby model językowy odpowiednio reagował w ramach konwersacji z użytkownikiem oraz korzystał z dodatkowych zasobów danych (np. danych o kierunkach studiów), w programie służącym do konwersacji, czyli chatbocie, poza modelem trzeba umieścić również szereg instrukcji, zaprojektować ścieżki konwersacji, a także ustalić, które dotychczas uzyskane od użytkownika informacje model ma ponownie uwzględnić, co z nimi zrobić i co z nich zachować, a wreszcie: jak interpretować intencje rozmówcy i jak reagować na jego wypowiedzi w różnych sytuacjach. Ta cała obudowa, składająca się na aplikację, jaką jest chatbot, to nadal zadanie dla ludzi.

Tego typu wyzwania stoją przed twórcami opartych na modelach językowych rozwiązań służących wspieraniu wyborów edukacyjnych i zawodowych. Stworzenie takiej aplikacji, w tym chatbota, wymaga rozsądnego zaprojektowanie sposobów rozpoznawania przez LLM mocnych stron użytkowników, ich predyspozycji, preferencji, zainteresowań, a także celów związanych z przyszłością zawodową. Wszystko po to, by zaproponować im inspirujące ścieżki rozwoju edukacyjnego i zawodowego, m.in. w postaci kwalifikacji, które mogą zdobyć.

Poniżej zostaną w skrócie przedstawione tego typu rozwiązania, testowane obecnie w ramach prac nad nową wersją chatbotki Zeterka. Jest to narzędzie do automatycznego wspierania doradztwa edukacyjnego i zawodowego, będące częścią prowadzonego przez Instytut Badań Edukacyjnych – Państwowego Instytutu Badawczego (dalej: IBE PIB) Zintegrowanego Rejestru Kwalifikacji (dalej: ZRK), dostępnego pod adresem: https://kwalifikacje.gov.pl/rejestr. Obecnie udostępniona wersja tej chatbotki korzysta z prostych modeli językowych. Równolegle w ramach realizowanego aktualnie projektu polegającego na modernizacji ZRK (w skrócie: projekt ZRK3) powstaje nowa wersja chatbotki, która w dużej mierze wykorzystuje model językowy Bielik. Rozwiązania te są owocem zbiorowych prac zespołu działającego w projekcie ZRK3 (tzw. zespół zadania 4.), a wcześniej w poprzedzającym go projekcie ZRK2, a także pracowników współpracującego z IBE PIB w tym zakresie wykonawcy, czyli Emplocity S.A.

Celem niniejszej prezentacji jest pokazanie, w jaki sposób tego typu narzędzie może pomóc we wspieraniu wyborów edukacyjnych i zawodowych poprzez rozpoznawanie mocnych stron i preferencji użytkowników. Wydaje się, że w czasach, kiedy coraz bardziej różnorodnie i elastycznie patrzy się na ścieżki karier (zob. Sullivan i Arris, 2021), w których nie jest niczym nowym mówienie o wspieraniu wyborów edukacyjnych czy zawodowych przez SI, o czym świadczą publikacje jeszcze z roku 2020 (de Laet et al., 2020; Guerra et al., 2020; Gutierrez et al., 2020) i stale bada się wykorzystanie SI zarówno w analizie danych na temat edukacji i rynku pracy (Gedrimiene et al., 2024), w tym w kontekście tzw. wyjaśnialnej SI (Guleria i Sood, 2023), jak i wprost w edukacji (Grassini, 2023) i w decyzjach edukacyjnych, np. w kontekście zapewniania równości (Viberg et al., 2024), dyskutowanie i dzięki temu rozwijanie takich rozwiązań jest zadaniem zdecydowanie pożądanym.

Dodatkowym celem niniejszego artykułu jest zilustrowanie, jak można projektować tego typu rozwiązania, a także wskazanie, jakie obecnie istnieją ograniczenia w wykorzystaniu SI w tym zakresie. Poniżej przedstawione zostaną kolejno rozwiązania związane z: wykorzystaniem tzw. testu preferencji; interpretowaniem podanych przez użytkownika informacji o jego zainteresowaniach, umiejętnościach, interesujących go zawodach; rozpoznawaniem potrzeb użytkownika.

2. Test predyspozycji i preferencji

Diagnozowanie predyspozycji lub preferencji zawodowych ma długą tradycję. W potocznej świadomości funkcjonuje przede wszystkim tzw. test Hollanda, oparty na teorii Johna Hollanda wyodrębniającej sześć typów osobowości (realistyczny, badawczy, artystyczny, społeczny, przedsiębiorczy, konwencjonalny) (Holland, 2017), będącej ważnym krokiem w ponadstuletniej historii stosowania kryteriów psychologicznych do prognozowania przydatności zawodowej (Syroka, 2003, s. 82–83) i wciąż stosowanej w praktyce badawczej (red. Paszkowska-Rogacz, 2013). Dużym zainteresowaniem cieszy się też od ponad 40 lat obrazkowy test skłonności zawodowych Martina Achtnicha (Achtnich, 1979; Jarosiewicz, 2014). Różne warianty tych testów lub testy nimi inspirowane doczekały się swoich wersji w postaci narzędzi elektronicznych, pomagając zwłaszcza uczniom poznawać samych siebie i podejmować wybory edukacyjne lub zawodowe zgodne z tym, co im bliskie.

W podobnym duchu stworzono test będący pomocniczym modułem w czatbotce Zeterka. Jest to ankieta, którą chatbotka proponuje do wypełnienia użytkownikowi. Jeśli się on na to zdecyduje, zostanie poproszony o ustosunkowanie się do 30 stwierdzeń (wylosowanych z puli o liczbie 60). Jego zadaniem jest określenie, w skali 1–5, na ile zgadza się (gdzie: 1 to "zupełnie się nie zgadzam", 5 – "w pełni się zgadzam") z takimi stwierdzeniami jak np. "Mam bujną wyobraźnię" czy "Przekonują mnie badania, logika, argumenty i fakty, a nie intuicje czy wrażenia". Na podstawie udzielonych odpowiedzi użytkownik otrzymuje propozycję kwalifikacji, które do nich najlepiej pasują. Jeśli kwalifikacji jest dużo, Zeterka proponuje kontynuowanie testu i ustosunkowanie się do pozostałych 30 z puli 60 stwierdzeń. Przed wyświetleniem ankiety pyta

jeszcze o poziom wykształcenia użytkownika, by ewentualnie wyeliminować te kwalifikacje, które nie znajdują się w polu zainteresowań użytkownika na danym poziomie wykształcenia.

Co ważne, Zeterka nie podaje tu wyniku w postaci profilu danego ucznia wyrażonego np. za pomocą typów Hollanda. Wynik, jaki otrzymuje jej rozmówca, to zestaw proponowanych kwalifikacji. Nie jest bowiem celem ani ambicją twórców tego testu określenie profilu rozmówcy Zeterki, a jedynie sformułowanie sensownej propozycji w zakresie kwalifikacji. Co ciekawe, zazwyczaj do pewnego stopnia wyniki te są dla osób, które go rozwiązują, dość oczywiste, co potwierdza względną skuteczność i przydatność testu (dzięki temu użytkownik dociera do interesujących go kwalifikacji), w pozostałym zakresie zaś bywają one zaskakujące, lecz – jak przyznają użytkownicy – nadal trafne. Ten ostatni efekt jest w istocie najbardziej oczekiwany i godny podkreślenia. Pozwala on bowiem poszerzyć użytkownikowi spektrum interesujących go kwalifikacji o te, o których dotąd nie myślał. Jest to ważna funkcja polegająca na rozwijaniu wyobraźni w myśleniu o kwalifikacjach.

Choć – jak zaznaczono – Zeterka nie podaje wyniku wyrażonego w terminach teorii Hollanda, w obecnej wersji wykorzystuje samą zaproponowaną przez tego uczonego typologię. Służy ona do powiązania wskazań rozmówcy Zeterki z kwalifikacjami. Każde ze stwierdzeń pojawiających się w teście jest przypisane do jednego z typów Hollanda. Gdy użytkownik określa, na ile zgadza się z danym stwierdzeniem przypisanym np. do typu artystycznego, odpowiednia liczba punktów zostaje doliczona do tego właśnie typu. Na koniec algorytm podlicza, ile punktów uzyskała każda z kategorii, i wybiera te, które otrzymały największą liczbę punktów w kolejności od najlepszego, podobnie jak w teście Hollanda. Zarazem w bazie danych kwalifikacji każda kwalifikacja ma przypisane trzy najbardziej pasujące do niej typy Hollanda, w kolejności od najbardziej pasującej. Algorytm porównuje wynik użytkownika z tymi informacjami i w przypadku pełnej zgodności – przedstawia użytkownikowi odfiltrowane w ten sposób kwalifikacje.

Ta stosunkowo nieskomplikowana metoda pozwoliła uzyskać dość zadowalające rezultaty. Uniknięto zarazem ewentualnych wątpliwości metodologicznych w zakresie psychometrii. Test nie aspiruje bowiem do określania typu osobowości rozmówcy Zeterki. Rezultat w postaci wyniku składającego się z typów Hollanda nie zostaje ujawniony, a jedynie wykorzystany do odfiltrowania bliskich użytkownikowi kwalifikacji. Zestaw tych kwalifikacji jest zaś jedynie miękką propozycją, mającą pomóc użytkownikowi zorientować się w świecie kwalifikacji, z którą może się on bardziej lub mniej zgodzić, a w tym drugim przypadku – kontynuować z Zeterką swoje poszukiwania.

Dotychczasowe rozwiązanie wiąże się jednak z kłopotami. Po pierwsze, rozwiązanie oparte na filtrowaniu za pomocą typów Hollanda narażone jest na pewną arbitralność i przez to niedokładność. Dzieje się tak, gdy trudno rozstrzygnąć, który typ w przypadku danej kwalifikacji powinien zostać do niej przypisany raczej na pierwszym miejscu niż na drugim, i analogicznie: raczej na drugim niż na trzecim, wreszcie: który powinien być na owym trzecim miejscu, a który nie powinien być uwzględniony. Przykładowo, dla

danej kwalifikacji można mieć wątpliwość, czy przypisać jej trójkę RBP (R – realistyczny, B – badawczy, P – przedsiębiorczy) czy raczej BRP; intuicja może też wskazywać, że na równi pasuje tu przedsiębiorczy, co społeczny – który wówczas wybrać? Każda taka decyzja ma wpływ na to, czy dana kwalifikacja zostanie lub nie zostanie zaproponowana użytkownikowi w przypadku danego wyniku testu. Po drugie, istnieje problem praktyczny w postaci konieczności przypisywania typów Hollanda do nowych kwalifikacji pojawiających się w Zintegrowanym Rejestrze Kwalifikacji, co do tej pory wykonywali eksperci wykonawcy chatbotki (Emplocity S.A.) lub eksperci IBE PIB.

W związku z tym drugim problemem w IBE PIB na przełomie 2024 i 2025 r. w zespole Data Science działającym w ramach projektu ZRK3 podjęto prace rozwojowe nad trenowaniem modeli służących do automatycznego przypisywania typów Hollanda do kwalifikacji. Prace te jeszcze trwają, jednak pierwsze rezultaty w postaci modeli wytrenowanych na różnych zestawach danych okazały się obiecujące. Równolegle podejmowane są testy związane z wykonywaniem takich przypisań przez duże modele językowe, m.in. polski model Bielik-11B-v2. Wydaje się, że w tym zakresie SI może zostać skutecznie wykorzystana do wykonywania tego zadania i rozwiązać wspomniany problem praktyczny.

Z kolei pierwszy problem mógłby zostać rozwiązany przez zastąpienie trójek przypisywaniem wag do każdej z kwalifikacji. A zatem zamiast przypisanej danej kwalifikacji trójki RBP można podawać np. procentową wagę: realistyczny [75], badawczy [67], artystyczny [44], społeczny [45], przedsiębiorczy [52], konwencjonalny [44]. Z takimi przypisaniami można porównywać pełny wynik testu wykonanego przez użytkownika. Pełny, czyli nie w postaci trójki typów Hollanda, które otrzymały najwięcej punktów, lecz w postaci matrycy prezentującej liczbę punktów zliczonych dla każdego z typów Hollanda. Rozwiązanie to jest jednym ze scenariuszy rozwoju tego modułu chatbotki Zeterka. Aby został on wdrożony, konieczne byłoby sprawdzenie, czy wytrenowany w tym celu model lub też któryś z dużych modeli językowych w zadowalający sposób przypisywałby do każdej kwalifikacji wagi dla poszczególnych typów Hollanda, tak jak to zilustrowano powyżej.

Obecnie jednak w ramach prac nad nową wersją Zeterki testowane jest zupełnie inne rozwiązanie, pomijające podejście, które możemy określić jako „ilościowe". Wykorzystuje ono potencjał dużego modelu językowego. Aktualnie testowany jest zainstalowany w infrastrukturze IBE PIB model Bielik-11B-v2.

Rozwiązanie to zakłada możliwość trafnego zinterpretowania przez LLM kluczowych preferencji czy predyspozycji rozmówcy Zeterki na podstawie wyników testu. Po wykonaniu przez użytkownika testu algorytm sprawdza, z którymi stwierdzeniami rozmówca Zeterki w pełni się zgodził, z którymi raczej się zgodził, a także – z którymi zupełnie się nie zgodził i z którymi raczej się nie zgodził.

Algorytm przygotowuje więc listy stwierdzeń, np. w postaci: "Użytkownik w pełni zgadza się ze stwierdzeniami: 1. Lubię książki/filmy zmuszające do myślenia. 2. Lubię uczyć się nowych rzeczy. 3. Chętnie podejmuję wyzwania. 3. Jestem osobą ambitną, stawiam sobie cele wysoko.", następnie "Użytkownik zupełnie nie zgadza się ze stwierdzeniami: 1. Bez trudu naprawię zepsuty

sprzęt. 2. Inni uważają, że mam zmysł praktyczny i dbam o porządek. 3. Lubię wyrażać uczucia przez muzykę, taniec, obraz, teatr lub język. 4. W czasie wolnym rysuję, lepię z gliny, komponuję, zmieniam wystrój, po prostu tworzę". Podobnie jest dla sytuacji typu "raczej". Informacje te dostarczone są LLM-owi, który na tej podstawie formułuje listę kluczowych preferencji i listę głównych antypreferencji użytkownika. W ten sposób otrzymujemy interpretację wyniku testu przez LLM.

Kolejny krok to dobranie kwalifikacji pasujących do preferencji i niepasujących do antypreferencji. Aby jak najlepiej się to udało, LLM został wcześniej wykorzystany do przypisania każdej z kwalifikacji związanych z nimi kluczowych preferencji/predyspozycji. Porównanie interpretacji LLM-u z danymi o kwalifikacjach zgromadzonymi w bazie danych następuje z wykorzystaniem modelu językowego typu *sentence vector*, zapewniającego dobranie kwalifikacji najlepiej pasujących do interpretacji LLM-a pod względem zgromadzonych informacji o kwalifikacjach, w tym pod względem predyspozycji i antypredyspozycji.

Rozwiązanie to jest obecnie dopracowywane. Wstępne wyniki są jednak obiecujące. W razie powodzenia kolejnych testów zostanie ono wdrożone w nowej wersji Zeterki. Będzie to kolejne oryginalne rozwiązanie wykorzystujące nowy potencjał, który pojawił się wraz z przełomem w rozwijaniu dużych modeli językowych, jaki nastąpił w roku 2022, wraz z uruchomieniem Chata GPT firmy OpenAI opartego na modelu GPT 3, a następnie GPT 3.5.

Tego typu wykorzystanie LLM-ów to nowe podejście, które może bardzo wzmocnić rozpoznawanie preferencji i predyspozycji uczniów.

3. Określanie zainteresowań, umiejętności itp.

Do kwalifikacji można jednak dotrzeć nie tylko przez predyspozycje. Są też inne drogi, uwzględniające zainteresowanie, umiejętności, a ostatecznie również interesujący użytkownika, w tym ucznia, zawód.

Obecnie udostępniona Zeterka posiada moduł do poszukiwania kwalifikacji w kontekście dalszego rozwoju. W ramach tych poszukiwań możliwe jest filtrowanie kwalifikacji przez wskazywanie wymaganego poziomu wykształcenia, wskazywania interesujących rozmówcę branż, wybieranie z listy posiadanych umiejętności czy preferencji, a także interesujących użytkownika zawodów. Zebrane informacje wykorzystane są do przeszukania bazy informacji o kwalifikacjach i dobrania najlepiej pasujących kwalifikacji. Wyszukiwanie to ma charakter semantyczny, a zatem dokonuje się na podstawie sensu użytych wyrażeń, a nie ich dokładnego brzmienia. W obecnej wersji wykorzystywany jest do tego model typu *sentence vector*.

Ponadto w module tym użytkownik może własnymi słowami opisać posiadane umiejętności czy interesujące go zadania. Dzięki wykorzystaniu modelu typu *sentence vector* w ramach wyszukiwania semantycznego Zeterka dobiera na tej podstawie kwalifikacje, które najlepiej pasują do opisu podanego przez użytkownika.

W Zeterce istnieje również możliwość wybierania z oddzielnej listy hobby czy zainteresowań. W obecnej wersji wybór ten porównywany jest z zainteresowaniami przypisanymi do każdej z kwalifikacji. Wiąże się to z ograniczeniami analogicznymi jak w przypadku testu preferencji.

W ramach prac nad nową wersją Zeterki przyjęto nowe rozwiązanie oparte na wykorzystaniu LLM na różnych polach. Po pierwsze, LLM został użyty do samodzielnego przypisania do każdej kwalifikacji listy pasujących do niej zainteresowań. Zakończone prace ewaluacyjne wykazały konieczność udoskonalenia promptu i powtórzenia tej procedury, jednak wyniki te są obiecujące i prace w tym kierunku są kontynuowane.

Po drugie, powstała nowa możliwość istotnie wykraczająca poza wybieranie zainteresowań. Zgodnie z nowym rozwiązaniem LLM sam interpretuje wypowiedzi użytkownika na temat tego, czym się on interesuje, co lubi robić itp. Na podstawie takiej wypowiedzi LLM formułuje listę zainteresowań użytkownika. Następnie lista ta zostaje porównana – znów dzięki wykorzystaniu modelu typu *sentence vector* – z listami zainteresowań przypisanych wcześniej do kwalifikacji. W ten sposób dobrane zostaną kwalifikacje najlepiej pasujące do tego, co napisał o sobie użytkownik.

Wreszcie, nowa wersja Zeterki nie ogranicza się do interpretowania tego typu wypowiedzi w zakresie samych zainteresowań. LLM otrzymuje zadanie rozpoznawania i kategoryzowania kluczowych informacji, takich jak: dotychczasowe wykształcenie, posiadane umiejętności, preferencje, predyspozycje, wspomniane już zainteresowania, a także ewentualnie doświadczenie zawodowe. Na tej podstawie tworzy profil użytkownika. Dzięki temu rozmówca Zeterki może zdecydować, by do wyszukania pasujących do niego kwalifikacji wykorzystać cały ten profil.

Ważna z perspektywy ucznia operacja wykonywana jest również wtedy, gdy pyta on o konkretny zawód. Zeterka nie przeszukuje wówczas bazy na podstawie wprost podanej wypowiedzi. W wypowiedzi takiej można przecież podać zawód wprost (np. "chcę być architektem"), lecz także można go opisać (np. "chciałbym projektować budynki"). Zeterka najpierw – znów – wykorzystuje LLM do interpretacji tej wypowiedzi. LLM interpretuje ją, formułując samodzielnie nazwę zawodu i ewentualnie dodając kluczowe związane z nim umiejętności. W ten sposób istotnie wzmacnia skuteczność wyszukiwania pasujących kwalifikacji.

Są to więc kolejne korzyści z zastosowania dużego modelu językowego. Rozwiązania te pozwalają na bardziej swobodną konwersację z chatbotem, z której może on uzyskać wystarczająco dużo informacji, by sformułować wartościową propozycję.

Warto zarazem zauważyć, że dzięki interpretowaniu przez LLM wypowiedzi rozmówcy i wydobywanie istotnych elementów, z których następnie buduje się jego profil, Zeterka umożliwia użytkownikowi lepsze poznanie samego siebie, a zarazem jest ogromną pomocą przy rozmowie z doradcą. W przypadku ucznia jego swobodna opowieść o sobie będzie podstawą do powstania profilu, w którym wyodrębnione zostaną jego umiejętności, predyspozycje,

preferencje, zainteresowania, a w rezultacie – bliskie mu kwalifikacje, obejmujące zarówno kierunki studiów, jak i kwalifikacje ze szkolnictwa branżowego oraz kwalifikacje wolnorynkowe i sektorowe.

Warto tu jeszcze poruszyć problem wspomniany we wprowadzeniu. Jak widać na podanych przykładach, na poszczególnych etapach działania chatbotki dwa różne modele wykonują różne zadania. Raz LLM ma za zadanie wydobyć z wypowiedzi dane kategorie informacji i je zinterpretować. Innym razem ma zinterpretować wynik testu. Co więcej, o czym jeszcze nie było mowy, na koniec ma on jeszcze sformułować swobodną wypowiedź, w której zaprezentuje wybrane z bazy danych kwalifikacje. Poza tym wykorzystuje się model typu *sentence vector*, który umożliwia przeszukiwanie semantyczne bazy danych, a wcześniej LLM wykorzystany jest do przypisania nowych informacji, typu zainteresowania czy predyspozycje/preferencje, do każdej z kwalifikacji. Można zadać sobie pytanie: czy musi tak to wyglądać? Czy nie wystarczy, że LLM sam będzie rozmawiał z użytkownikiem, mając do dyspozycji bazę danych o kwalifikacjach? Niestety, w praktyce nigdy tak to nie wygląda. Sam model nie będzie wiedział, co zrobić. Nie będzie zdolny do tego, by zdecydować, że należy skorzystać z takich, a nie innych danych z bazy lub w taki, a nie inny sposób wykorzystać wyniki testu predyspozycji. Na każdym etapie musi otrzymać w prompcie odpowiednie instrukcje oraz dostać pakiet informacji, na których ma bazować. W przeciwnym wypadku będzie się opierał wyłącznie na zapytaniu rozmówcy oraz własnej wiedzy nabytej w ramach treningu modelu.

Jak zaś odpowiada model, który nie korzysta z zasobów zewnętrznych, takich jak np. aktualna baza informacji o kwalifikacjach? Na pierwszy rzut oka wydaje się, że rozmawia sensownie, jednak szybko okazuje się, że często są to informacje zmyślone, a praktycznie zawsze są one niepełne.

Modelowi Bielik-11B-v2 można zadać następujące pytanie: "Lubię pracować na świeżym powietrzu, lubię przestrzeń i przyrodę. Chcę mieć spokojną pracę. Jakie kwalifikacje powinienem zdobyć, by wykonywać związany z tym zawód?". Model odpowiada bardzo sensownie i na pewno pomoże nam w poszukiwaniach. Ogranicza się jednak do ogólnych nazw kierunków studiów oraz ogólnych określeń kursów. Nie znajdziemy jednak w jego odpowiedzi konkretnych kwalifikacji, w tym konkretnych kierunków na konkretnych uczelniach, a tym bardziej takich kwalifikacji jak rynkowe czy sektorowe. W innej sytuacji są ChatGPT lub Gemini, które korzystają z zasobów internetu i są w stanie znaleźć aktualne kwalifikacje z ZRK. Można poprosić ChatGPT o podobną rzecz: "Podaj mi, proszę, nazwy kilku kwalifikacji wolnorynkowych lub sektorowych dla osób, które lubią być na dworze, lubią przyrodę i spokojną pracę, podlinkowane do opisów tych kwalifikacji na stronie Zintegrowanego Rejestru Kwalifikacji". Jako wynik otrzymamy nazwy istniejących kwalifikacji, ale bez bezpośrednich linków. Co więcej, można odnieść wrażenie, że pewne istotne kwalifikacje zostały pominięte. W przypadku zaś Bielika odpowiedź na ten ostatni prompt jest już pełna zmyślonych informacji, łącznie ze zmyślonymi niedziałającymi linkami.

Na obecnym etapie rozwoju SI jest więc nieodzowne odpowiednie instruowanie modeli i wskazywanie konkretnych zadań wykonywanych w odpowiedniej kolejności, uzupełnione o wykorzystanie modelu pozwalającego na

wyszukiwanie semantyczne. Wymaga to więc projektowania tzw. *flow*, czyli schematu działania chatbota w zależności od sytuacji, oraz zaprojektowania sekwencji wykonywanych zadań. Przede wszystkim zaś – sensownego zaprojektowania sposobów wykorzystania LLM w celu zdobycia od użytkownika informacji, które pozwolą na trafne zaproponowanie mu kwalifikacji.

4. Rozpoznawanie potrzeb (intencje)

Użytkownik, który rozmawia z chatbotem, zazwyczaj potrzebuje, by go poprowadzić. Jest gotowy, by swobodnie rozmawiać. Nie zawsze ma jasno określony cel. Zeterka pomaga w tym, zachęcając użytkownika do opowiedzenia o sobie i do rozmowy o kwalifikacjach. Zarazem jednak musi sensownie reagować na różne jego wypowiedzi.

W obecnie udostępnionej wersji to Zeterka nadaje ton. Mówi użytkownikowi, co ma on zrobić. Daje różne opcje do wyboru. Tam, gdzie użytkownik ma pole do wyboru, Zeterka za tym wyborem podąża. Jeśli jednak Zeterka o coś prosi, a użytkownik robi co innego, do niczego sensownego nie dochodzi. Jeśli np. użytkownik ma wpisać swoje umiejętności, a on zada pytanie typu "ale po co?" lub zacznie opowiadać o swoim życiu, Zeterka odpowie, że nie potrafi pomóc. Nastawiona jest bowiem na szukanie kwalifikacji, a podane przez użytkownika treści zupełnie do tego nie pasują.

Wykorzystanie LLM-a w nowej wersji chatbotki pozwala na udzielenie sensownej odpowiedzi na każdą wypowiedź użytkownika. Zarazem jednak LLM musi być poinstruowany:

1. kiedy może sobie pozwolić na swobodną odpowiedź, tak jak sam potrafi, w tym kiedy ma uciąć temat lub zaoponować;
2. kiedy ma nie odpowiadać swobodnie, tylko wykonać zadanie (np. rozpoznać, że użytkownik podał dane do wyszukania mu kwalifikacji, więc należy je zinterpretować, podając np. listę zainteresowań czy umiejętności potrzebną do przeszukania bazy, by wyszukać pasujące kwalifikacje), żeby dopiero w kolejnym kroku sformułować odpowiedź na podstawie dostarczonych danych (np. listy otrzymanych kwalifikacji).

W tym celu LLM ma dodatkowe zadanie. Zanim odpowie użytkownikowi, musi rozpoznać intencję, czyli potrzebę, jaka stoi za wypowiedzią użytkownika. Musi np. ustalić, czy użytkownik:

1. pyta o rejestr kwalifikacji lub system kwalifikacji;
2. szuka kwalifikacji i mówi o sobie coś, na podstawie czego można zaproponować kwalifikacje;
3. wprost mówi, że chciałbym wykonać test preferencji;
4. mówi coś nie na temat (luźna rozmowa).

Jeśli LLM rozpozna przypadek (1), otrzyma instrukcję, by odpowiedzieć na pytanie użytkownika, ale dodatkowo otrzyma związaną z tematem zapytania informację z bazy wiedzy o rejestrze kwalifikacji i systemie kwalifikacji, dzięki czemu jego odpowiedź będzie wiarygodna. Jeśli LLM rozpozna (2), wypowiedź użytkownika zostanie w kolejnym kroku zinterpretowana przez LLM, by wydobyć z niej dane o umiejętnościach, zainteresowaniach czy preferencjach.

Następnie otrzyma wynik w postaci listy kwalifikacji, by sformułować na tej podstawie rekomendację. Jeśli rozpozna sytuację (3), powinien udostępnić użytkownikowi test, upewniwszy się, czy na pewno chce on go wypełnić. Jeśli rozpozna (4), po prostu odpowiada, ewentualnie korzystając z dodatkowych zasobów wzorcowych pytań i odpowiedzi, jednak dość swobodnie. Zarazem jednak jego meta prompty zawierają instrukcję, by w przypadku wypowiedzi wulgarnych, prowokacyjnych itp. zdecydowanie zaoponować i uciąć dany wątek rozmowy.

Każdorazowo więc LLM, który otrzymuje zapytanie użytkownika ma za zadanie rozpoznanie intencji wypowiedzi, a wśród tych intencji – jego potrzeb. Odpowiada za to meta prompt zawierający listę reguł postępowania typu: „Jeśli użytkownik robi/pisze to a to, odpowiedz jedną literą X". Umożliwia to odpowiednie kierowanie działaniami chatbota, tak by jak najbardziej odpowiadały one potrzebom użytkownika, o ile nie wykraczają poza to, co moralnie i prawnie dopuszczalne, lecz zgodne są z celem, jakim jest wsparcie użytkownika.

Wreszcie, użycie w czatbotce LLM-a pozwala na to, by maksymalnie dopasować się do potrzeby użytkownika, o której on sam wyraźnie napisze. Choć w końcowej fazie LLM otrzymuje z bazy listę kwalifikacji, jego odpowiedź nie będzie taka sama, lecz dostosuje się do tego, o co prosił użytkownik. Przykładowo: jeśli poprosi on wprost o kwalifikacje, np. „Podaj mi kwalifikacje, które pasują do takich a takich umiejętności", otrzyma odpowiedź skoncentrowaną na kwalifikacjach. Jeśli jednak poprosi o kierunki studiów i związane z nimi zawody, LLM poda kwalifikacje ze szkolnictwa wyższego, lecz zarazem – zgodnie z potrzebą użytkownika – uzupełni tę listę o informacje o zawodach (oczywiście na podstawie swojej wiedzy nabytej w ramach trenowania modelu). Podobnie zachowa się w innych analogicznych sytuacjach, gdy na przykład zapytamy go też o kluczowe umiejętności, jakie się z tym wiążą.

Wykorzystanie LLM-a pozwala więc chatbotce nie tylko sensownie odpowiedzieć na każdą wypowiedź rozmówcy, lecz także rozpoznać, o co chodzi mu w danej wypowiedzi, i odpowiednio zareagować. W szczególności pozwala to rozpoznać, kiedy szuka on kwalifikacji i opowiada o sobie, podaje informacje, które umożliwiają ich znalezienie.

W przypadku ucznia rozmawiającego z Zeterką ma to niezwykle ważne znaczenie. Pozwala na swobodną rozmowę, gdy uczeń tego potrzebuje, a następnie odpowiednie nakierowanie go na właściwe tory, by – gdy LLM rozpozna istotne informacje – podać sensowną rekomendację w postaci ścieżki edukacyjnej i zawodowej, wzbogaconej o konkretne kwalifikacje, lecz także dostosowaną do jego pytania.

5. Podsumowanie

Powyżej zaprezentowano, jak narzędzia cyfrowe oparte na SI mogą wspierać użytkowników, w tym uczniów, w ich wyborach edukacyjnych. Pokazano, że dzięki wykorzystaniu modeli językowych mogą one nie tylko skutecznie przeszukiwać bazy danych, w prezentowanym przypadku – bazę z informacjami o kwalifikacjach, lecz także skuteczniej rozpoznawać preferencje, predyspozycje, mocne strony, w tym umiejętności oraz zainteresowania uczniów. Wraz ze

związanymi z nimi wyzwaniami przedstawiono rozwiązania zastosowane w celu rozpoznawania wspomnianych tu elementów. Wśród nich, w jaki sposób można wykorzystać SI do interpretacji testów predyspozycji lub preferencji. Co ważne, w nowym rozwiązaniu uwzględniono aspekt negatywny w postaci antypreferencji/antypredyspozycji, dzięki czemu bierze się pod uwagę również okoliczności, których poszukujący swojej drogi edukacyjnej uczeń chciałby uniknąć.

Wreszcie, ukazano istotny (zazwyczaj niedoceniany) aspekt, jakim jest interpretacja intencji wypowiedzi użytkownika, która w kontekście celu, jakim jest wsparcie w wyborach edukacyjnych i zawodowych, nabiera szczególnego znaczenia. Stworzony w takim celu chatbot musi bowiem łączyć dwie funkcje: prowadzenie swobodnej rozmowy, dzięki której rozmówca będzie rozmawiał w sposób choć częściowo naturalny, co stwarza okazję do wydobycia od niego potrzebnych informacji, a zarazem udzielania konkretnych informacji i rekomendacji. Rozpoznawanie intencji i dopasowywanie się do bieżących potrzeb jest tu jednym z ważniejszych wyzwań, stwarzających ogromne szanse.

Podsumowując, przykłady te pokazują, że chatbot wykorzystujący LLM może istotnie wesprzeć uczniów w rozpoznawaniu tego, co istotne, aby zaproponować im dostosowane do ich potrzeb kwalifikacje. Na przykładach tych pokazano ograniczenia SI na obecnym etapie rozwoju. Nawet bardzo zaawansowane chatboty, korzystające z zasobów internetu, nie dostarczają jeszcze wystarczająco wiarygodnych, a na pewno kompletnych danych. Wskazano też, w jakim zakresie niezbędny jest wysiłek osób, które projektują tego typu narzędzia, jeśli mają one udzielać wiarygodnych informacji i realnie wspierać uczniów w ich edukacyjnych wyborach.

Warto na koniec dodać jeszcze jeden wniosek. Z tym ostatnim faktem wiąże się także jedna istotna sprawa. Choć rozwiązania te wymagają wiele wysiłku ze strony projektujących je osób, ma to również swoje zalety w postaci kontroli nad procesami prowadzącymi do udzielenia rozmówcy odpowiedzi. Dzięki tej kontroli mamy świadomość, w jaki sposób i na podstawie jakich danych przekazywana mu informacja została wytworzona i w jakim zakresie jest wiarygodna. W dużej mierze spełnia to więc postulat wyjaśnialnej SI (ang. *explainable AI*). Przede wszystkim daje zaś większą pewność co do jakości udostępnianego produktu.

Bibliografia

Achtnich, M. (1979). *Der Berufsbilder-Test. Projektives Verfahren zur Abklarung der Berufsneigung*, Verlag Hans Huber Bern, Stuttgart-Wien.

De Laet, T., Millecamp, M., Ortiz-Rojas, M., Jimenez, A., Maya, R., & Verbert, K. (2020). *Adoption and impact of a learning analytics dashboard supporting the advisor-student dialogue in a higher education institute in Latin America.* British Journal of Educational Technology, 51(4), 1002–1018. https://doi.org/10.1111/bjet.12962

Gedrimiene, E., Celik, I., Kaasila, A. et al. (2024). A*rtificial Intelligence (AI)-enhanced learning analytics (LA) for supporting Career decisions: advantages and challenges from user perspective.* Educ Inf Technol, 29, 297–322. https://doi.org/10.1007/s10639-023-12277-4

Grassini, S. (2023). *Shaping the Future of Education: Exploring the Potential and Consequences of AI and ChatGPT in Educational Settings*. Education Sciences, 13(7), 692. https://doi.org/10.3390/educsci13070692

Guerra, J., Ortiz-Rojas, M., Zúñiga-Prieto, M. A., Scheihing, E., Jiménez, A., Broos, T., De Laet, T., & Verbert, K. (2020). *Adaptation and evaluation of a learning analytics dashboard to improve academic support at three Latin American universities*. British Journal of Educational Technology, 51(4), 973–1001. https://doi.org/10.1111/bjet.12950

Guleria, P., Sood, M. (2023). *Explainable AI and machine learning: performance evaluation and explainability of classifiers on educational data mining inspired career counseling*. Education and Information Technologies, 28, 1081–1116. https://doi.org/10.1007/s10639-022-11221-2

Gutierrez, F., Seipp, K., Ochoa, X., Chiluiza, K., De Laet, T., & Verbert, K. (2020). *LADA: A learning analytics dashboard for academic advising*. Computers in Human Behavior, 107, Article 105826. https://doi.org/10.1016/j.chb.2018.12.004

Holland, J.L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (wyd. 3). Odessa, Florida: Psychological Assessment Resources, Inc.

Jarosiewicz, H. (2014). *Pomiar skłonności zawodowych–zastosowanie Obrazkowego Testu Zawodów*. Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, 350, 166-178.

Paszkowska-Rogacz, A., red. (2013). *Wieloaspektowa Ocena Preferencji Zawodowych (WOPZ). Podręcznik testowy*. Częstochowa. APUS Robert Meller.

Syroka, J. (2003). *Typologiczna teoria karier – założenia konceptualne i prezentacja narzędzia*. Annales Universitatis Mariae Curie-Skłodowska: Sectio J, 16: 81-90.

Sullivan, E. S., & Al Ariss, A. (2021). *Making sense of different perspectives on career transitions: A review and agenda for future research*. Human Resource Management Review, 31(1). https://doi.org/10.1016/j.hrmr.2019.100727

Viberg, O., Kizilcec, R.F., Wise, A.F., Jivet, I. and Nixon, N. (2024). *Advancing equity and inclusion in educational practices with AI-powered educational decision support systems (AI-EDSS)*. Br J Educ Technol, 55: 1974-1981. https://doi.org/10.1111/bjet.13507