

Nagy nyelvi modellek tanulói alkalmazásának tapasztalatai és lehetőségei a köznevelésben

Vaskó András, Mentor: Fodor Richárd

MATHIAS CORVINUS COLLEGIUM KÖZÉPISKOLÁS ÖSZTÖNDÍJPROGRAM

Tartalomjegyzék

1.Bevezetés	2
1.2. Kutatási probléma és kutatói kérdések.....	3
1.3. Hipotézisek.....	4
2. Szakirodalmi áttekintés.....	5
3. Kutatásmódszertan	9
4. Nyelvi modellek összehasonlítása:.....	10
4.1. A kutatás menete	10
4.2. Harmadik féltől származó alkalmazások	11
4.3. Bing Chat/ Microsoft Copilot.....	11
3.4. Google Gemini	12
4.5. Baidu Ernie.....	13
4.5. Meta Llama 2	14
4.6. Claude 3	14
4.7. ChatGPT 3/3.5, 4	15
4.8. Gyakorlati teljesítményvizsgálat.....	16
4.9. Összefoglaló	18
5.Kísérlet	21
5.1.A kísérlet menete	21
5.2. A kísérlet eredményei	23
5.2.1.Irodalom teszt	23
5.2.2.Történelem teszt.....	26
5.3.A kísérlet összegzése	29
6. Konklúzió	31
Mellékletek	38
1. számú melléklet: Fogalomtár	38
2. számú melléklet: Modellek összehasonlítása (táblázat)	39
3. számú melléklet: kiadott feladatok.....	40
4. számú melléklet: Diákok összehasonlítása - Táblázat	41
5. Számú melléklet – kérdőív és interjúkérdések	42

1.Bevezetés

2023-ban egy hatalmas változás történt a világon. Ebben az évben, a Chat GPT megjelenésének hatására, gyorsan kezdett elterjedni a mesterséges intelligencia. Sok program jelent meg különböző területeken. A nagy nyelvi modellek (Large Language Model, LLM) lényege, hogy a korábbiaktól eltérően komplexen képesek kérdésekre válaszolni, így jelentősen kibővült a felhasználási területük. Képesek egy adatbázis alapján különböző témákban összetett gondolatmeneteket alkotni és értelmesen reflektálni az emberi interakcióra. Az embereket váratlanul érte ez a technológiai áttörés. Korábban is léteztek olyan programok, amik képesek voltak kommunikálni az emberrel, de ezek felhasználhatósága korlátozott volt.

Az oktatás hazánkban, ahogy az egész világon, lassan követi le a technológiai újításokat. Példának okáért 2000-ben vezették be a szélessávú internetszolgáltatást Magyarországon. Ezt követően 2016-ban jelentették be, a DOS-t, tehát Magyarország Digitális Oktatási Stratégiáját. Ebben megfogalmazták, hogy 2018-2019-es tanév kezdetére az összes iskolába bevezetik a szélessávú internetkapcsolatot (Magyarország Digitális Oktatási Stratégiája, 2016). Ebből arra lehet következtetni, hogy 2016-ban még nem volt minden iskolában szélessávú internetkapcsolat, pedig akkoriban már ez alapvető volt a háztartásokban. A felsőoktatásban sokkal gyorsabban történtek digitalizációs lépések. 2005-ben több egyetemen bevezették a NEPTUN.net szolgáltatást, mely rendszer digitalizálta az adminisztráció jelentős részét, beleértve a pénzügyi és tanulmányi területeket. A közoktatásban viszont csak a 2017/18-as tanévtől kezdve indult el a KRÉTA digitális adminisztrációs rendszer, hasonló céllal (Oktatási Hivatal, 2024). Ezen példák alapján kitűnik, hogy a köznevelés digitális fejlesztése egy rendkívül időigényes feladat.

Kutatásomban a nagy nyelvi modellek jelentőségét vizsgálom a köznevelésben, különös hangsúlyt helyezve a diákok szemszögére. Ez azért fontos, mert az emberek nagyon hamar felismerték, hogy hatalmas lehetőség lakozik ebben a technológiában, így a diákok is használni kezdték. A többség nem ismeri a technológia működési elvét, így olyan feladatokra próbálják használni, amire nem feltétlenül alkalmas. Bár a kezelése elsőre egyszerűnek tűnik, ugyanis csak szövegesen kell betáplálni az utasításokat, mégis rengeteg buktatója van a megfelelő bemenet megválasztásának. A nyelvünk egy roppant bonyolult rendszer, így a megfelelő szöveges parancs megfogalmazása is összetett.

Kutatásommal helyi és országos szinten is szeretném felhívni a figyelmet a témára, és eredményeimmel szeretném elősegíteni a mihamarabbi integrációt.

A beszámolómban sok fogalom témaspecifikus. Ezeket a későbbiekben magyarázom. A magyarázat rövidített, összesített változatát az 1. számú mellékletben lehet megtalálni.

1.2. Kutatási probléma és kutatói kérdések

Jelenleg nem sok kutatás van, amely a diákok szemszögéből közelíti meg a nyelvi modellek diákok általi felhasználásának témáját. Hiányzik egy komplex összehasonlítás a jelenleg fellelhető nagy nyelvi modellek tanulói felhasználhatóságáról a köznevelésben. A diákok és a tanárok csak a saját belátásuk és az interneten fellelhető, olykor megtévesztő információk alapján használják. Ez sokszor ahhoz vezet, hogy impulzívan ítélik meg a technológia felett, és többet vagy épp mást gondolnak róla, mint amire optimálisan alkalmazható a működési elve alapján. Ez rengeteg félrevezetést eredményez. Azok az emberek, akik nem megfelelően vagy egyáltalán nem tudják használni a technológiát, sokkal rosszabb eséllyel indulnak a munkaerőpiacon, ugyanis egy feladatot lassabban tudnak csak megoldani.

Mindenképp fontosnak tartom, hogy a jövőben bevezessék az oktatás részeként ezen új technológia alapszintű alkalmazását. Jelentős mértékben felgyorsítja a munkavégzést, így fontos lenne készség szinten ismerni ezt a segítséget mindenkinek. Ennek elősegítésére kutatásomban kitérek a nagy nyelvi modellek technikai működésére.

Az elmúlt 1,5 évben rengeteg mesterséges intelligencia alapú csevegőszoftver jelent meg a piacon. A helyzet egyik indikátora az OpenAI nevű vállalat volt, ami bemutatta a ChatGPT nevű, mesterséges intelligencián alapuló csevegőszoftverét. A csevegés, mint ige arra hagy következtetni, hogy informális beszélgetésekre lett kitalálva, de ez koránt sincs így. Lehetséges ez is vele, viszont korlátoltsága és személytelensége miatt nem ajánlatos erre alkalmazni. Sokkal inkább használható különböző feladatok elvégzésére, információk gyűjtésére, szövegek átformálására. Kutatásomban vizsgálom a különböző nyelvi platformokat, hogy megtudjam:

K1: Melyik mesterséges intelligencián alapuló nyelvi platform a legalkalmasabb a magyar köznevelésben való alkalmazásra a diákok számára?

Természetesen a nagy nyelvi modellek és a diákok kapcsolatának két oldala van. Az első kérdésem kapcsán a modellek oldalával második kérdés kapcsán pedig a diákok oldalával foglalkozok.

K2: Mely iskolai tantárgyak és milyen típusú feladatok megoldásában nyújthatnak segítséget a mesterséges intelligencia felhasználásával működő szoftverek?

Lényegesnek tartom a felhasználás módját is. Nagyon fontos, hogy miként dolgozunk a nagy nyelvi modellekkel. Mivel csak egy módon lehet utasítást adni, szöveges formában, ezért meghatározó, hogy milyen szöveget adunk bemeneti utasításként. Ehhez kapcsolódik a harmadik kutatási kérdésem.

K3: Hogyan érdemes használni diákként a nagy nyelvi modelleket?

1.3. Hipotézisek

Hipotéziseimet az imént felsorolt kérdésekhez kapcsolódva fogalmaztam meg. Téma szerint a kérdések sorrendjében osztom fel, így az első kérdéshez az első hipotézis tartozik.

H1: Feltételezésem szerint a diákok számára bármelyik modell megfelelhet. Valójában a felhasználás módja, a modellek teljesítményében van különbség és ez mérvadó lehet. A jobb teljesítményű modell jobb válaszokat ad, így könnyebb, gyorsabb és pontosabb vele a munka.

H2: Feltételezésem szerint, azon tantárgyak esetében, melyek nagyban ragaszkodnak a tankönyvben megfogalmazott diszciplináris tudáshoz, és kevés teret engednek a kreativitás kibontakoztatásához, jelentős segítséget tud nyújtani a használatuk. Ennek ellenére, szerintem, minden területen képes megrövidíteni a feladatvégzést. Használatuk kis eséllyel vezet minőségromláshoz, megfelelő alkalmazás esetén csak hozzátesz a munka értékéhez.

H3: Azt gondolom, hogy a használat kulcsa a megfelelő prompt megírása. Fontos a világos utasítás megfogalmazása, hogy könnyen meg lehessen érteni. Persze ehhez pontosan tudni kell a felhasználás célját és hogy milyen eredményt várunk. Valószínűleg nem tudja kiváltani az emberi munkát, viszont ötleteket adhat, összefoglalókat készíthet, repetitív feladatokat el tud látni.

2. Szakirodalmi áttekintés

Ahhoz, hogy vizsgálni tudjuk a közoktatásbeli felhasználását, ismerkedjünk meg a mesterséges intelligencia fogalmával, illetve az egyéb, manapság népszerű, viszont sok esetben helytelenül használt hozzá kapcsolódó kifejezésekkel.

Az alábbi fejezetben bemutatom a mesterséges intelligencia rövid fogalomtörténetét, interdiszciplináris kapcsolódási pontjait más tudományterületekhez, legfontosabb kulcsfogalmait és kategóriarendszereit.

A „mesterséges intelligencia” kifejezést John McCarthy alkotta meg 1955-ben. McCarthy a mesterséges intelligenciát úgy definiálta, mint "olyan gépek viselkedését, amelyet intelligensnek neveznénk, ha egy ember így viselkedne" (McCarthy *et al.*, 2006). Viszont a kérdés nem ennyire egyszerű. A XX. század második felében és a XXI. században a technológia rohamos ütemben fejlődni kezdett és a számítógép szerepe a mindennapokban jelentősen megnőtt, tehát ezek a programok is többet tudnak, mint az ember. „A mesterséges intelligencia definiálásának kihívása tehát a mesterséges jellemzők és az emberi intelligenciától eltérő szempontok meghatározása.” (Cope, Kalantzis and Searsmith, 2021) Bizonyos szemszögből nézve egy egyszerű számológép is okosabb, mint az ember, mivel jelentősen gyorsabb számítási kapacitással bír. Ellenben a legtöbb esetben jócskán elmarad az embertől. A legtöbb számítógépes program nagyon specifikus feladatot lát el, ezért nem lehet intelligensnek nevezni őket. Az emberi intelligencia meghatározása és mérése kihívást jelent, bár számos kísérlet történt ennek kvantifikálására, egy olyan elvont és szubjektív fogalom objektív definiálása, mint az intelligencia, a gyakorlatban megvalósíthatatlan. (Gardner, 2011; Kaplan, 2016)

Az emberi felhasználás miatt, kénytelenek vagyunk besorolni és kategorizálni ezeket a programokat, illetve szabályok közé szorítani. Ennek eredményeként a tanulmányokban, kutatási, oktatáspolitikai vagy piaci jelentésekben található legtöbb definíció homályos. Hiányzik egy közös meghatározás, ellenben számos közös vonást figyelhetünk meg az elemzett definíciókban. Például a környezet észlelése, beleértve a valóság bonyolultságának figyelembevételét, információfeldolgozás, a bevitt információ gyűjtése és értelmezése (adatformában), döntéshozatal (beleértve a következtetést és tanulást). Cselekvések végrehajtása, feladatok teljesítése (beleértve az alkalmazkodást, a környezeti változásokra adott reakciót) bizonyos fokú önállósággal. Konkrét célok elérése: ezt tekintik az MI rendszerek végső céljának. (European Commission. Joint Research Centre, 2020) Ezekből is látszik, hogy a MI képes minden esetben az emberhez hasonló módon reagálni racionálisan. Felismeri,

értelmezi és megoldja az előtte álló problémákat a legjobb tudása szerint. Mindenféle feladatot meg tud oldani, illetve képes tanulni a saját hibáiból, tudja fejleszteni önmagát a visszajelzések alapján. Ezzel szorosan összefügg a gépi tanulás fogalma, amely egy informatikai módszerre utal különféle programok létrehozásához. A gépi tanulás (Machine Learning) statisztikai technikákat alkalmaz az észlelt minták alapján történő előrejelzésekhez. (Zhai, C. et al., 2016). Egyszerűbben kifejezve, ha nagy mennyiségű macskaképet adunk neki, és megjelöljük, hogy ezek a képek macskákat ábrázolnak, különféle mintázatokat fog azonosítani a képeken. A mélytanulás (Deep Learning) és a neurális hálózatok olyan többrétegű statisztikai modellek, amelyek mintákat azonosítanak az adatokban (Krizhevsky et al., 2017; Rumelhart et al. 1986). Ezen technológiák működéséhez hatalmas mennyiségű adat és jelentős számítási kapacitás szükséges. A hálózatok többrétegű elemzése olyan eredményeket hoz létre, amelyek kevésbé intuitívak, mint az egyszerűbb gépi tanulási modellek - a mélytanulás során a gép az egyes rétegekben újraszámolva saját eredményeit, folyamatosan tanul és fejleszti magát (Hossenfelder, 2018). A gyakorlatban a mélytanulás felerősített gépi tanulásként működik. A nagyobb adatállomány, nagyobb számítási kapacitás és bonyolultabb algoritmusok révén ezek a rendszerek sokkal pontosabb eredményeket képesek produkálni.

A "korpusz" (corpus) eredeti latin jelentése "test", a modern nyelvben pedig az írott és beszélt szövegek összességét jelenti (Oxford English Dictionary, 2024). Minél nagyobb a korpusz, annál sokrétűbb és mélyebb témákat tartalmaz, ami változatosabb szóhasználatot tesz lehetővé. A mélytanulás nem csak szövegek esetében működik, hanem képeken, hangokon és más típusú adatokon is alkalmazható.

A természetes nyelvfeldolgozás (Natural Language Processing - NLP) lehetővé teszi, hogy a számítógépek szöveget és beszédet értsenek meg emberi módon. Az NLP a számítógépes nyelvészet, a statisztikai modellek, a gépi tanulás és a mélytanulás kombinációja (Holdsworth, 2024). Ez a technológia alapvető fontosságú a chatbotok (más néven társalgási algoritmusok) működéséhez, amelyek képesek interakcióra lépni a felhasználókkal szöveges vagy hangos formában, és különböző feladatokat végrehajtani.

Az NLP és a mélytanulás alkalmazásai széleskörűek. Az ügyfélszolgálatoktól kezdve az egészségügyi diagnosztikán át az autonóm járművekig számos területen hasznosítják ezeket a technológiákat. Az NLP például lehetővé teszi, hogy a chatbotok pontosabban értsék és válaszolják meg a felhasználói kérdéseket, míg a mélytanulási modellek képesek komplex mintázatokat felismerni az adatokban, ami javítja a predikciók és a döntéshozatal pontosságát.

Az ilyen rendszerek folyamatosan fejlődnek, és a jövőben még inkább integrálódnak mindennapi életünkbe, új lehetőségeket teremtve a technológiai innovációk számára.

Az érzelelemzés (Sentiment Analysis) egyre fontosabbá válik, mivel gyorsan és hatékonyan tanulmányozza a közösségi médiában és más online platformokon megjelenő véleményeket. Célja az adatok feldolgozása és azok polaritásának (pozitív, negatív vagy semleges) meghatározása (Aqlan, et al. 2019). A vélemények kategorizálása elengedhetetlen a gépi tanulási algoritmusok számára. Az érzelelemzés előnyös a marketingben, politikában, pénzügyi szektorban és ügyfélszolgálatban, mivel lehetővé teszi a közvélemény követését és az ügyfélélegedtség javítását. A nagy mennyiségű szöveges adat automatizált feldolgozása jelentősen csökkenti az elemzéshez szükséges időt és erőforrásokat, azonban fejlett algoritmusokra van szükség. A chatbotok, amelyeket mesterséges intelligencia és gépi tanulás alkalmazásával fejlesztenek, a természetes nyelvfeldolgozás és érzelelemzés koncepcióit használják a felhasználókkal való kommunikációra, mint például az HDFC bank EVA botja és az Amazon Chat ügyfélszolgálat (Singh, et al. 2023). Az elmúlt évtizedben a chatbotok leggyakrabban cégek ügyfélszolgálatánál jelentek meg, ahol az egyszerű online ügyintézészt gyorsították meg. Ezek a chatbotok specifikus feladatokat látnak el, például az adott oldalon fellelhető információk gyorsabb megtalálását segítik. Az utóbbi időben azonban a többfunkciós chatbotok terjedtek el szélesebb körben. Ezek a chatbotok nemcsak céges környezetben, hanem általános beszélgetések szimulálására és személyi asszisztensként is használhatók. A chatbotok fejlődése a mesterséges intelligencia és a gépi tanulás terén elért előre lépéseknek köszönhetően egyre több feladatot tudnak hatékonyan ellátni. A jövőben várhatóan tovább nő a szerepük az ügyfélszolgálati, oktatási, egészségügyi és számos egyéb területen, ahol az emberekkel való kommunikáció és a feladatok automatizálása kiemelt fontosságú.

A legismertebb chatbot, a ChatGPT, nevének utolsó három nagybetűje (GPT) jelentős szerepet játszik a platform képességeinek megjelölésében. A Generative Pre-trained Transformer (GPT) egy önmagát tanító modell, amely a természetes nyelvfeldolgozás feladataihoz a Transformer architektúrát alkalmazza. A nagy nyelvi modell (LLM) egy olyan mesterséges intelligencia modell, amelyet kiterjedt szöveges adathalmazokon képeznek, hogy felismerje és megértse az emberi nyelv mintázatait, szerkezetét és finom árnyalatait. Az LLM-ek képesek a szöveggörnyezet alapján prediktív módon meghatározni a következő szót egy mondatban vagy szószerkezetben (Wickman, 2023). Egyszerűbben megfogalmazva, egy nagy méretű alaphalmazból (korpusz) kiválasztja a legvalószínűbb következő szót, ami a legjobban tükrözi az emberi viselkedést. A ChatGPT fejlesztéséhez az előzetes kiképzés során egy nagy

menyiségű szövegtörzset alkalmazták a modell tanításához (Lee et al., 2023). Rengeteg irodalmi, tudományos és közéleti szöveget „olvastattak” a modellel, hogy elemezze azokat és figyeljen meg mintázatokat. A ChatGPT a világ leggyorsabban növekvő internetes platformjává vált, mindössze két hónap alatt 100 millió havi felhasználót elérve, ami jóval gyorsabb növekedést mutatott, mint a Facebook, a Twitter és az Instagram esetében (Adamopoulou et al., 2020; Lee et al., 2023).

A promptok a felhasználó és a ChatGPT közötti kommunikáció elsődleges eszközeként szolgálnak. Ezek a rövid szöveges utasítások vagy kérdések irányítják a modellt, segítve a releváns és kontextusba illeszkedő válaszok generálását. A promptok révén a felhasználók konkrét információkat kérhetnek, feladatokat adhatnak, vagy irányt szabhatnak a beszélgetés menetére. Az így megadott útmutatások alapján a ChatGPT képes azonosulni a felhasználó szándékával, és a lehető legpontosabban reagálni a kérésre. A promptok minősége és részletessége nagyban befolyásolja a ChatGPT által adott válaszok pontosságát és relevanciáját (Negin Yazdani Motlagh et al., 2023). A token a MI modellek, különösen a nagy nyelvi modellek (LLM-ek) által feldolgozott szöveg legkisebb egysége. Ezek a tokenek a szöveg építőelemei, amelyeket a gépi tanulási modellek be- és kimenetként használnak. A tokenek lehetnek teljes szavak, szórészek, egyedi karakterek vagy akár nagyobb szövegegységek is. A MI rendszerek a tokenek segítségével értik meg és dolgozzák fel az emberi nyelvet, így képesek kommunikálni és válaszokat generálni. A tokenizáció az a folyamat, amely során a szöveget kisebb, kezelhető részekre, tokenekre bontják. Ez lehetővé teszi a MI rendszerek számára, hogy elemezzék és megértsék a nyelvi mintákat és összefüggéseket. Például a „It’s raining outside” mondat tokenizálása során három token keletkezik: „It’s”, „raining”, „outside”. A tokenizáció segít az AI-nak felismerni a szavak és kifejezések közötti kapcsolatokat, amelyek alapján előre tudják jelezni a következő szavakat vagy kifejezéseket a válaszok generálása során (Agomuh, 2024, Zhao Q, et. al., 2024).

A nagy nyelvi modellek (LLM-ek), mint például a GPT, a mesterséges intelligencia fejlett eszközei, melyeket hatalmas mennyiségű szöveges adaton képeznek ki. Ezek a modellek a természetes nyelvfeldolgozás (NLP) feladatait látják el, és képesek prediktív módon meghatározni a következő szót egy mondatban a szövegtörzset alapján. A modellek tanítása során irodalmi, tudományos és közéleti szövegeket dolgoznak fel, hogy felismerjék és megértsék az emberi nyelv mintázatait. A mélytanulás (Deep Learning) és a több rétegű neurális hálózatok lehetővé teszik a modellek számára, hogy bonyolultabb és pontosabb eredményeket érjenek el. A tokenizáció során a szöveget kisebb egységekre, tokenekre bontják, amelyek

segítik a modelleket a szavak és kifejezések közötti kapcsolatok felismerésében és az összefüggések megértésében. A felhasználói promptok irányítják a modell válaszait, biztosítva a specifikus és releváns információkat. Az LLM-ek képesek folyamatosan tanulni és fejlődni a visszajelzések alapján, ami növeli pontosságukat és alkalmazhatóságukat különféle területeken, például ügyfélszolgálatokban, oktatásban és egészségügyben. Az ilyen modellek jelentősen hozzájárulnak a technológiai innovációhoz, lehetővé téve az automatizált és hatékony kommunikációt és problémamegoldást.

3. Kutatásmódszertan

A kutatásom kvalitatív megközelítésű deduktív empirikus kutatás. A feltáró szakirodalmi áttekintést és elemzéseket követően két fő vonalon vizsgáltam a mesterséges intelligenciák felhasználását a köznevelésben. Az egyik egy elméleti elemzés, a másik egy gyakorlati megközelítésű kísérlet. Az elméleti vizsgálat során a népszerűbb nagy nyelvi modelleket hasonlítottam össze, technikai, illetve gyakorlati alapon. A technikai vizsgálat a specifikációkra fókuszál, a gyakorlati vizsgálat során egy történelemérettségi esszéfeladatot adtam mindegyik modellnek, és az általános érettségi szempontok alapján vizsgálom a válaszokat.

A kutatás másik fő vonala egy tanteremi kísérlet volt, mely során középiskolás diákok készítettek beadandó dolgozatokat történelem és magyar irodalom tantárgyból. Két spontán választott diákcsoport alakult ki a feladatmegoldás formája alapján. Az első csoportban a mesterséges intelligenciát igénybe vevő diákok, a másodikban az MI segítségével dolgozó diákoké. A feladatok alapvetően a középszintű érettségi teljesítéséhez szükséges tudást vártak el, korábban tanult témában. A kísérlet célja volt, hogy a két csoport közötti különbségeket tanulmányozzam. A vizsgálati szempontok között szerepelt a dolgozatok minősége, a munka hatékonysága, illetve a diákok háttere, tanulmányi eredménye, a nagy nyelvi modellek használatában való jártasság. A kísérlet nagyon korlátozottan, 35 fő részvételével készült, így nem nevezhető reprezentatívnak. 11. osztályos országos viszonylatban átlagosnak mondható gimnázium nyelvi tagozaton tanulóinak munkái jó alapot adhatnak más kísérletekhez. Fontos és szembetűnő megállapításokat lehet ugyanis megtenni ilyen szűk mintából is. A feladatok mellett a diákok háttérét is szerettem volna megismerni. Erre egy kérdőív, illetve rövid interjúk formájában

4. Nyelvi modellek összehasonlítása:

4.1. A kutatás menete

A ChatGPT megjelenése óta rengeteg különféle cég dobott piacra hasonló programot. A következőkben vizsgálom ezeket a szoftvereket. A vizsgálat célja, hogy megállapítsam, milyen segítséget tudnak nyújtani az egyes szoftverek a diákoknak. Az elemzésben nagy cégek programjaira fókuszálok, ahol a legtöbb esetben a modell is egyedi fejlesztés. Ez alól egy kivétel van, mégpedig a Microsoft cég Copilot nevű szoftvere, ami az OpenAI által fejlesztett GPT-4 alatt fut. Ez a szűkítés nagyon fontos a kutatás szempontjából. Az összehasonlításomat elsősorban technikai adatokra alapozom, de természetesen a felhasználási módokat, lehetőségeket, a korlátokat és a gyártó ajánlásait is figyelembe veszem. Ezek alapján a szempontrendszerem az alábbiak szerint alakul: A program pontos neve (amennyiben rövidítés, annak kifejtése), megjelenési dátuma, fejlesztője, a modell, ami a program motorja, illetve az információbázisa. Ezen legfontosabb jellemzők mellett, fontosnak tekinthető még az információk visszakereshetősége, felhasznált források feltüntetése, illetve korlátossága, a megadható utasítás maximális hosszúsága. Ezen szempontokon felül fontos még az ár és az érte nyújtott szolgáltatás. Ez az integráció szempontjából is jelentős meghatározó erő. A kutatásomban elsősorban a bárki számára ingyenesen elérhető verziókat veszem alapul, ugyanis jelenleg ezt használják a legtöbben. Amennyiben az egyes modellek új, egyedi szempontot hoznak be, azt ez egyéb jellemző kategóriában fogom feltüntetni.

Az összehasonlítás részeként egy vizsgálatot is készítek, melyben letesztetem a vizsgált eszközöket. A vizsgálat célja, hogy megtudjam mennyire képesek jól teljesíteni ezek a modellek egy önálló szövegalkotási feladatban. Mivel a kutatásom második felében diákoknak is el kellett végezni egy hasonló feladatot, így a későbbiekben ez jó összehasonlítási alapot jelent. A pontos feladat a 2023 októberi középszintű történelemérettségi 1848 előtti rövid esszéfeladata. A feladat során az iszlám vallással és annak 5 alappillérevel foglalkozik. A kísérlet során az érettségifeladat leírását adtam meg prompt formájában az egyes modelleknek. Így az abban szereplő forrás szolgált a válaszok fő kiindulópontjaként. Sajnos az Ernie és a Llama nem elérhető jelenleg Magyarországon. A Llama esetében ezt sikerült orvosolnom egy harmadik féltől származó alkalmazáson keresztül. A segítségem a Sider nevű program volt, amely egy pdf file alapján dolgozik, így az általa adott eredmény nem pontos és a feladat szempontjából nagyon korlátozott.

4.2. Harmadik féltől származó alkalmazások

Nagyon sok olyan weboldal van, ami különböző AI eszközöket kínál. A legtöbb esetben ezek nem teljesen saját fejlesztésűek, és csak kiegészítésként szolgálnak. Előnyük, hogy optimalizált kezelőfelülettel rendelkeznek egy-egy feladatra specializálódva, ezzel megkönnyítve a felhasználhatóságot. Rengeteg területen hoztak létre ilyeneket. Csak említés szintjén hozok most pár példát, a legfőbb jellemzőikkel. Az információkat a szoftverek saját hivatalos weboldalukról származnak.

A Quizgecko AI szövegből készít tanulókártyákat, kvizeket. (e-mail). Három verziója van, ingyenes, diák, ami 64\$ évente és oktató, ami 89\$ évente. Csak a fizetős verziókban van AI segédlet, az ingyenes verzióval csak manuálisan lehet kvizeket készíteni.

A Sider egy adatszűkítő program. Lényege, hogy egy személyen feltöltött PDF anyagából dolgozik, kimondottan onnan válaszol kérdésekre. A Saider többféle motorral is dolgozik, GPT-3.5 Turbo; Claude 3 Haiku; Claude 3 Sonnet; Gemini 1,5 Flash; Llama 3. A felhasználó személyesen választhat ezek közül. Van ingyenes verziója, illetve 3 különböző fizetős csomagja 80\$/150\$/200\$-os árazással. Ezek a havi lekérdezések mennyiségében, illetve a feltölthető pdf-ek méretében térnek el.

A tanároknak is nyújtanak különböző segédleteket. Az eduaide tanórai segédleteket, kvizeket, feladatokat, óravázlatokat készít AI segítségével. Különböző motorok alatt fut, melyek között a feladat függvényében vált. (e-mail)v . Az ingyenes verzió nagyon korlátozott, nagyjából demo verziónak nevezhető, csak tesztelésre alkalmas. A fizetős, korlátlan verzió évi 50\$-ba kerül.

A felsorolt platformok kiegészítő szolgáltatások, ugyanis egy hagyományosnak nevezhető LLM is képes kiváltani a munkájukat. Sokszor viszont ez nagyon sok munkát és odafigyelést jelentene (különösen az eduaide esetében), így hasznos AI eszköznek tekinthetők. Egy másik előnyük, hogy sokszor teljes elérést biztosítanak az általuk alkalmazott modellhez is. Az áruk pedig a kiegészítő szolgáltatásokkal együtt sem magasabbak. A kutatás szempontjából nem tartom relevánsnak az imént említett, illetve felsorolt programok részletes elemzését.

Legfőképpen azért, mert nagyon sok ilyen program van, hasonló tulajdonságokkal, nincs jelentős különbség köztük technikailag. A legtöbb esetben a személyes preferencia dönt, hogy ki melyiket használja. Természetesen nem vitatom el a relevanciájukat. Sok esetben jobb a felhasználhatóságuk, mint a nyers modelleknek.

4.3. Bing Chat/ Microsoft Copilot

A Microsoft 2023 február 7-én mutatta be Bing keresőmotorjának átdolgozott verzióját. (Negin Yazdani Motlagh *et al.*, 2023) A legnagyobb újítása a mesterséges intelligencia magasfokú

integrációja. Az új Bing MI alapú keresőmotort kapott, ami biztosítja a pontosabb keresési eredményeket. E mellett egy nagy nyelvi modell alapú chatbotot is implementáltak. (Microsoft, 2023) Az OpenAI által fejlesztett szoftver növeli a válaszok pontosságát és gyorsaságát, a mesterséges intelligencia által. A Bing Orchestratorban összekapcsolja a keresést a GPT mesterséges intelligenciájával. A Bing GPT-4 modellen fut, amely internet-hozzáféréssel rendelkezik. A rendszer tartalmaz lábjegyzeteket és tud akadémiai hivatkozásokat is szolgáltatni (Negin Yazdani Motlagh *et al.*, 2023) A megújult keresés, illetve a beépített nyelvi modell nagyban megnöveli a felhasználói élményt és a Google keresőmotorjának, korábban kevés népszerűségnek örvendő Binget, egy komoly versenytársává teszi. Jelenleg a Bing ingyenesen nyújtja ezt a szolgáltatást, minden magán Microsoft fiókhoz. (Microsoft, 2024)

A Microsoft Copilot különböző beszélgetési stílusokat is kínál (Kreatív, Kiegyensúlyozott és Precíz), amelyek lehetővé teszik a válaszok hangnemének és részletességének testre szabását. Összességében a Microsoft Copilot sokoldalú és dinamikus MI eszköz, amely interaktív és testre szabható élményt nyújt. (Ciornei, S. 2024)

A Copilot előnyeikhez legfőképpen a GPT-4-hez való ingyenes hozzáférés, a vizuális funkciók, illetve az internethozzáférés tartozik. Hátránya pedig a korlátozott (4000 karakteres) prompt méret.

3.4. Google Gemini

Korábban Google Bardnak hívták, de egy nagyfokú frissítés után a fejlesztő átnevezte Geminira. Ez az frissítés az írás ideje alatt történt. A Gemini 2023. december 6-án indult el, mint a Google DeepMind MI modellje (Hsiao. 2024).

Az eszköz két verzióval rendelkezik: normál, és advanced. A normál verzió ingyenes és az 1.0 verzió alatt fut. Az Advanced verzióra való előfizetés 20 dollár havonta. Ez a verzió 1.5-ös modell alatt fut. Míg a Gemini ingyenesen elérhető és ideális az egyéni felhasználók, diákok vagy kisebb vállalkozások számára, addig a Gemini Advanced előfizetéshez kötött, és elsősorban fejlesztőknek, adatelemzőknek és szakembereknek szól. A két verzió közötti technikai különbségek az MI modellben és a funkciókban mutatkoznak meg. A Gemini Advanced használatakor a Gemini Ultra, a legfejlettebb MI modell kerül alkalmazásra, míg a Gemini csak a Gemini Pro-t használja. Ez a különbség jelentős előnyt jelenthet azoknak, akiknek komplex feladatokat kell megoldaniuk vagy fejlett MI funkciókat szükséges használniuk. Felhasználói szempontból a Gemini Advanced sokkal több lehetőséget kínál az integrációra a Google Workspace alkalmazásaival, ami különösen hasznos lehet fejlesztők és

vállalati felhasználók számára. A Gemini ingyenes verziója ideális, akik alapvető MI funkciókat keresnek, míg a Gemini Advanced azoknak ajánlott, akiknek komplexebb feladatokra van szükségük és akik számára fontos a Google alkalmazásokkal való szoros integráció. (Hsiao 2024)

A Gemini sokoldalúsága lehetővé teszi a különféle adatok, például szöveg, kép, hang, PDF és videó feldolgozását, ami hasznos különféle feladatok és alkalmazások esetén. A Gemini segítség lehet különféle feladatokban, mint például szövegelemzés, programozás, logikai feladatok, szövegértés, matematikai problémák megoldása és kódgenerálás. A kommunikáció terén is sokoldalú, képes különféle kommunikációs feladatok és stílusok kezelésére, interaktív szimulációk és tanulási környezetek létrehozására. (Imran, Almusharraf, 2024)

A kutatás szempontjából viszont a program azért nem releváns, mivel csak 18 év felett lehet használni hivatalosan és a fejlesztő ellenőrzi a felhasználói fiókokat a használat előtt. Mivel a Google bevezette ezt a korhatárt, nem beszélhetünk a Bard vagy (most már Gemini) köznevelésbeli felhasználásáról. Ezen szabály miatt legkorábban csak az egyetemeken lehet oktatni a program felhasználását.

4.5. Baidu Ernie

A kínai tech vállalatok, hasonlóan nyugati társaikhoz, beléptek a generatív AI versenybe. A Baidu által kifejlesztett Ernie chatbot nyolc hónap alatt már 200 millió felhasználót szerzett. Míg az amerikai fogyasztók már 2022 novemberétől hozzáférhetnek generatív MI-hoz, például a ChatGPT-hoz, Kínában ezek az eszközök nem érhetők el, ezért a hazai tech vállalatok, mint a Baidu, igyekeznek kihasználni az igényt. Az Ernie-t a Baidu 2023. március 16-án indította el, és teljes neve Enhanced Representation through Knowledge Integration, amely nagyjából „szövegfejlesztés tudásintegráció révén” jelentéssel bír (Negin Yazdani Motlagh et al., 2023). Az Ernie egy nagy nyelvi modell és generatív MI eszköz, amely a felhasználók szöveges kommunikációjára reagál kínai nyelven. Képes szöveges és vizuális tartalom létrehozására, és kulturális ismeretei szempontjából felülmúlja más kínai technológiai versenytársait. Kifejez különböző kínai dialektusokban, beleértve a szecsuaní, kantoni és hokkien nyelveket (Negin Yazdani Motlagh et al., 2023). Az Ernie 4.0 2023 októberében jelent meg, amely komplex feladatokat és szöveg-, kép- és videógenerálást jobban megért (Shankland, S. 2024). A modellt 10 milliárd paraméterrel betanították egy 4TB méretű korpuszon, amely szövegeket és egy nagy méretű tudásgrafikont tartalmaz. Az empirikus eredmények szerint az ERNIE 3.0 minden eddigi modellt felülmúlt 54 kínai NLP feladatban, és az első helyet érte el a SuperGLUE

benchmarkon, meghaladva az emberi teljesítményt. (Sun et al., 2021). A modell nagy teljesítménye és a kínai technológia fejlettsége mindenképp említésre méltóvá teszi, annak ellenére, hogy hazánkban sajnos várhatóan sosem lesz elérhető.

4.5. Meta Llama 2

A Llama 2, a Meta terméke, a nyílt forráskódú nagy nyelvi modellek (NNM-ek) legújabb fejlesztését képviseli. Egy hatalmas, 2 billió szavas adathalmazon lett kiképezve, ami jelentős növekedés a Llama 1-hez képest (Masalkhi et al., 2024). A nyílt forráskód itt kulcsfogalomnak tekinthető, ugyanis a többi program nem rendelkezik ezzel a tulajdonsággal. Mivel nyílt forráskódú projekt, a Llama 2 átláthatóságot és követhetőséget kínál a fejlesztési folyamatában. A felhasználók hozzáférhetnek a modell forráskódjához, amely lehetővé teszi számukra, hogy megértsék és ellenőrizzék annak funkcióit. (Roumeliotis, et al. 2023) Ez az engedmény a professzionális felhasználókat érintheti nagyon pozitívan. Képesek lehetnek saját kiegészítéseket tenni a programhoz. Ez lehetőséget nyújt egy különleges verzió kialakítására, ami a diákok igényeire lenne szabva. Természetesen ennek megvalósulásához rengeteg tőke szükséges, így nem látok valós esélyt rá, hogy a közeljövőben ez létrejönne. A Llama 2 modell egy erős LLM, amelyet a természetes nyelvi feldolgozási képességei miatt ismert, különböző területeken alkalmazható. A Llama 2 fejlett nyelvi feldolgozási képességekkel rendelkezik, amelyek lehetővé teszik számára, hogy hatékonyan végezzen olyan feladatokat, mint a dokumentumalapú kérdések megválaszolása. (Roumeliotis, et al. 2023)(Masalkhi et al., 2024) A program nem nevezhető kiemelkedőnek a vetélytársakkal szemben, több esetben alulmarad tudásában. Egyetlen pozitívuma a nyílt forráskódjából fakad, amivel a megfelelő tudású felhasználók tudnak élni. Összességében nem tartom alkalmasnak a közoktatásban történő használatát. A Llama 2 megengedő licenclési modellt követ, azt jelenti, hogy a felhasználók jelentős szabadságot élveznek a modell különböző célokra történő alkalmazásában minimális korlátozásokkal (Roumeliotis, et al. 2023).

4.6. Claude 3

Claude 3 egy nagy teljesítményű nagy nyelvi modell, amely kiválóan teljesít a kontrollmérnöki problémák megoldásában. A programot az Allen Institute for AI fejlesztette, és az egyik legfejlettebb LLM a piacon. Claude 3 Opus megjelenése viszonylag friss, és a legújabb generációs nagy nyelvi modellek közé tartozik. A program hatalmas információbázissal rendelkezik, amely lehetővé teszi számára, hogy széles körű problémákat megoldjon. (Kevian

et al., 2024) A fejlesztő, az Anthropic, egy olyan MI startup, amely többek között az Amazon által támogatott jelentős, 4 milliárd dolláros, befektetést kapott. A Claude 3 három különböző modellt tartalmaz: a prémium kategóriájú Claude 3 Opus-t, a második szintű Claude 3 Sonnet-et és a még nem megjelent Claude 3 Haiku-t. A Claude 3 Opus az Anthropic legfejlettebb modellje, amely csak a Claude Pro felhasználók számára érhető el. Az Anthropic mérései szerint az Opus modell posztgraduális szintű érvelési képességei 50,4%-ot értek el, szemben a GPT-4 35,7%-os eredményével. A Claude 3 Sonnet az ingyenes verzió, amely bár kevésbé fejlett, mint az Opus, még mindig magas színvonalú teljesítményt nyújt. Érvelési képességei 40,4%-os eredményt mutattak ugyanazon teszten. A Claude 3 Haiku a legkevésbé fejlett modell, amelyet elsősorban vállalati felhasználásra terveztek. Ez a modell a gyorsaságáról ismert, és olyan feladatokhoz ideális, mint a weboldalak ügyfélszolgálati chatbotjainak működtetése. A Claude 3 különböző modelljeinek ára változó. A Claude 3 Opus havi 20 dollárba kerül a fogyasztók számára, míg az üzleti felhasználók token alapú rendszert használnak: 15 dollár/ millió input-token és 75 dollár/ millió output-tokenen. A Claude 3 Sonnet ingyenesen használható, de az üzleti felhasználóknak 3 dollár/millió input-token és 15 dollár/ millió output-token. A Haiku modell ára szintén token alapú: 0,25 dollár/ millió input-tokenen és 1,25 dollár/millió output-token. Claude 3-t a GPT-4-gyel és a Google Gemini Ultra-val hasonlítva, az Anthropic saját tesztjei szerint a Claude 3 család kimagasló teljesítményt nyújt. Azonban a végleges értékeléshez további független tesztek szükségesek. Az Anthropic bemutatott tesztjei azt sugallják, hogy a Claude 3 jelentős képességekkel rendelkezik, de az objektív összehasonlítás érdekében szükség van külső, független értékelésre is. (Laird, J. 2024)

4.7. ChatGPT 3/3.5, 4

A ChatGPT fejlesztője a OpenAI, egy cég, amely a természetes nyelvfeldolgozásra és a mesterséges intelligenciára specializálódott. A ChatGPT-ben szereplő három utolsó nagybetű nagyon fontos megjelölése a platform képességeinek. A Generative Pre-trained Transformer (GPT) egy autoregresszív, ön-tanuló modell, amely a Természetes Nyelvfeldolgozás feladataihoz alkalmazza a Transformer architektúrát.(Wickman, 2023) A ChatGPT 2022 novemberében jelent meg. A GPT-3.5 modell egy változatát használja, amely egy generatív előre betanított transformer típusú modell. A ChatGPT nem biztosít közvetlen forráskövetést, de képes szövegrészleteket vagy teljes szövegeket is generálni a felhasználói bemenet alapján. (Somoye, F. 2024.) A ChatGPT tudásának határideje 2022, így nem biztos, hogy rendelkezik információval a nagyon friss eseményekről vagy fejleményekről. A ChatGPT nem mindig érti

az emberi nyelv árnyalatait vagy egy beszélgetés kontextusát. Ezért használatakor nagyon fontos az egyértelmű megfogalmazás, a logikailag nem komplex megfogalmazási mód, illetve a helyes nyelvhasználat. A ChatGPT tükrözheti azokat az elfogultságokat, amelyek az általa betanított adatokban jelen vannak. A ChatGPT-3.5 modellje 175 millió tokenből áll, amelyek segítik a modellt az információk feldolgozásában és a szövegek generálásában. Bár a ChatGPT képes szövegeket generálni, nem mindig képes új vagy kreatív ötletekkel előállni. (Negin Yazdani Motlagh *et al.*, 2023; Levkovich, 2023)

A ChatGPT-4 a ChatGPT sorozat legújabb verziója, amely továbbfejlesztett funkciókkal és képességekkel rendelkezik. Az OpenAI által fejlesztett ChatGPT-4 a mesterséges intelligencia területén nyújt új lehetőségeket és fejlesztéseket. A ChatGPT-4 motorja a GPT-4 modellre épül, amely magasabb szintű nyelvi modellezést és szöveggenerálást tesz lehetővé. A ChatGPT-4 nagy adatbázisokon alapul, amelyek segítik a modellt a széles körű információk feldolgozásában és a pontos válaszok generálásában. A ChatGPT-4 esetében hangsúlyosabb az információkeresés, mivel a modell képes hatékonyabban keresni és feldolgozni az információkat a szövegekben. Képes forrásokkal szolgálni és visszakereshetővé tenni azokat, ami segíti a felhasználókat az információk hitelességének ellenőrzésében. 200 millió tokenből áll, ami lehetővé teszi a modell számára a nagyobb információfeldolgozási kapacitást és a komplex összefüggések kezelését. Esetében a maximális prompt méret bővült, ami lehetővé teszi a felhasználók számára a nagyobb és részletesebb bemeneti promptok használatát a válaszok generálásához. (Levkovich, 2023)

4.8. Gyakorlati teljesítményvizsgálat

A technikai elemzés mellett fontosnak tartom, hogy gyakorlati úton is megvizsgáljam az egyes szoftvereket. Ennek formájára, a közoktatás témája miatt, egy érettségi feladatot választottam. Pontosabban a 2023. októberi középszintű történelem érettségit adtam meg. A prompt minden esetben a feladatkiírás volt, a hosszúságot tettem hozzá, ami 150-200 szóban határoztam meg, ami nagyjából megfelel az elvártaknak. Sajnos, hazánkban nem mindegyik szoftver elérhető. Ezért az Ernie és a Llama esetében nem tudtam a gyártó oldalát használni. A Llama esetében sikerült egy köztes megoldást találnom egy harmadik féltől származó alkalmazással. Ebben az esetben egy pdf alapján dolgozott, így ez csökkenthette a modell szabadságát. A pdf tartalma a feladatban foglalt idézet volt. A feladat értékeléséhez szintén a mesterséges intelligencia segítségét kértem, lényegében önértékelést végeztem. Természetesen ezt felülvizsgáltattam és kikértem szaktanár véleményét. A következőkben ennek értékeléséről lesz szó.

Minden megoldás szaknyelvet használó, összefüggő mondatokból állt. A Copilot és a Gemini vázlatpontokban adta meg a választ, míg a másik három platform folyószöveget írt. A legtöbb válasz megfelelt a 100-200 szavas keretnek, kivéve a Gemini választ, amely 258 szavas volt.

A Chat GPT és a Copilot megoldásaiban szerepelt a legtöbb iszlám valláshoz kapcsolódó idegen kifejezés (pl. szalát, sehede), amelyeket minden esetben magyarázattal láttak el. A legnagyobb különbség a hivatkozott források feltüntetésében volt; csak a Copilot sorolta fel az öt internetes forrást, amelyeket használt. Ezek az érettségi tételekkel foglalkozó weboldalakból, valamint Farkas Judit közismert történelemtanár írásaiból és egy névtelen tételből származtak, amelyet a Magyarországi Muszlimok Egyháza lektorált.

A történelemdidaktikai értékelés során kulcsfontosságú volt a forráshasználat jelenléte és minősége. A Copilot kivételével minden platform utalt a forrásra, míg a Copilot válaszában ez nem volt jelen. A Copilot csak az iszlám öt pillérét sorolta fel, amelyek közvetetten érintik a forrásban szereplő tiltásokat, de nem alakított ki kapcsolatot a kettő között. A többi platform belefoglalmazta a válaszbba a forrás fogalmait és további szempontokat is.

A platformok legnagyobb hiányossága az időben és térben való elhelyezés volt. Egyik sem említ semmilyen iszlámmal kapcsolatos eseményt vagy helyszínt. Ennek okai valószínűleg a feladatban keresendőek, ugyanis sem idézet, sem a feladat kiírása nem említi, hogy erre is ki kell térni. Itt a megfelelő eredményhez fontos lenne tudni a kontextust is.

Ezt követően a platformoknak az érettségi szempontok alapján értékelniük kellett saját válaszaikat. Az önértékelések nyolc és tizenöt pont között mozogtak a tizenhetes skálán. A Copilot önértékelése felismerte a forráshasználat hiányosságát. Az érettségi értékelési útmutató szigorú értelmezése alapján a Copilot és a Gemini válasza nem felelt meg a feladatmegoldásnak a vázlatpontos szerkezet miatt. A Chat GPT és a Claude 3 tizennégy és tizenöt pontos önértékelése reálisnak tekinthető, a Chat GPT megoldása alaposabbnak tűnik.

Összességében az öt megoldás mind megfelelhet a középszintű érettségi gondolkodási szintjének. A Copilot a források hivatkozásával, a Gemini a Koránra utaló pontos utalással, a Chat GPT pedig a koherens megfogalmazással emelkedik ki. Fontos megjegyezni, hogy mindegyik válasz csak alapot szolgáltathat egy megfelelő és elfogadható esszéhez, amelyhez szükség van az értékelési szempontokban jártas, gyakorlott tanulói történelmi gondolkodásra is. További érdekes próba lehetne a megoldás ismétlése az ismert értékelési szempontok mentén.

4.9. Összefoglaló

Az eddigiekben bemutatam röviden a mesterséges intelligenciák működését. Ezt követően sorra vettem a számomra relevánsnak tekintett nyelvi modelleket és ezeket elemeztem technikai alapon. Majd egy tesztet is végeztem velük, ahol egy érettségi szituációban vizsgáltam őket önértékelés segítségével, illetve szaktanár bevonásával. Ezeket kívánom összefoglalni egy rövid útmutató formájában. Ennek alapjául a feljebb megfogalmazottakat, egy szakirodalmi és két internetes forrás szolgált. Tehát akkor adódik a kérdés, hogyan használjuk a nagy nyelvi modelleket? Mi a jó prompt titka?

A prompt megírása előtt egy nulladik lépésként kiemelem a modell megértését. Ez az alapja a nagy nyelvi modellekkel történő munkának. Itt fontos nem csak maguknak a nagy nyelvi modelleknek a működését ismerni, hanem magát az adott modellt is. A nyelvi modellek működését korábban ismertettem. Röviden összegezve, olyan gépi tanulást alkalmazó algoritmusok, melyek a szövegekkel dolgoznak. Egy adott utasítást követően egy adatbázisban kikeresi a témához releváns információkat. Azokból egy szöveget alkot olyan módon, hogy statisztikai és valószínűség-számítási alapon megpróbálja kikövetkeztetni a megfelelő következő szót. Az egyes modellek működésében számottevő eltérések lehetnek. A korábbi elemzéseim alapján két fő csoportra lehet tagolni ezeket. Az egyik a szöveg alapú modellek, ezek sokkal jobban dolgoznak a szövegekkel a többihez képest. Általában hosszabb prompt is megadható nekik, illetve hosszabb választ is képesek adni. Ellenben könnyebben tévednek és dolgoznak hamis információkkal. Sajnos a legtöbb esetben forrással nem tudnak szolgálni. Ezen hátrányukat kompenzálja, hogy hosszabb prompt is megadható nekik. Ez lehetőséget nyújt teljes szövegek megadására is. Ezen tulajdonságuk miatt sokkal könnyebben használhatóak szövegek értelmezésére, átdolgozására, kivonatolására. Így segítség lehet pl.: teszt, jegyzet, összefoglaló készítésében, átfogalmazásban, fordításban. A másik nagy csoport az információ alapú chatbotok, melyek általában sokkal rövidebb bemenettel dolgoznak, kapcsolódnak az internethez és segítenek a kérdéseinkre, témánkra a lehető legpontosabb információkat megtalálni. Mivel ezek a modellek hozzá vannak kapcsolódva az internethez, így mindig naprakész információkkal tudnak szolgálni. Sajnos hátrányuk, hogy nem tudnak csak a megadott szövegre koncentrálni, mindig megpróbálják kiegészíteni, illetve rövidebb a bemenet és a válasz. A két csoport közti különbségek megjelentek a modellek tesztjei során is. Megfogalmazás terén jobban teljesítettek a szöveg alapúak, a téma általánossága és kevésbé friss mivolta miatt tartalom terén sem voltak rosszak. Feltételezem, hogy egy specifikusabb és frissebb témában nagyobb diverzitás lenne megfigyelhető a válaszokban.

Miután megismertük a modellek működését, tudjuk melyiket mire lehet és érdemes használni. Elkezdhetjük megírni a promptot, ennek az alapja a feladat pontos meghatározása. Nélkülözhetetlen, hogy tudjuk mit szeretnénk, mert ez alapján kell a modellt kiválasztani. Fontos ismerni a technológia korlátait és lehetőségeit. Lényeges megadni a feladat háttérét, milyen előzetes információk birtokában lehet nekiállni a feladatnak, ugyanis a program semmit nem tud a dologról. Például, ha egy teszt készítése a cél, akkor érdemes tudatni a célcsoportot, vagy ha a válasznak meghatározott formátumúnak vagy struktúrájúnak kell lennie (pl. lista, bekezdések, pontok formájában), akkor ezt is szükséges jelezni a promptban. Meg kell határozni a kommunikáció jellegét és hangnemét, legyen az hivatalos, barátságos, tájékoztató vagy más. Ez befolyásolja, hogyan kommunikál a modell. Érdemes példát mutatni a kívánt válasz típusára, ez segít a modellnek jobban megérteni az elvárásokat, pontosabb válaszhoz vezet. A modell megfigyelés alapján tanul. Fontos figyelni a stilisztikára is. Bár ezek a modellek képesek javítani és megérteni a rossz nyelvtant és helyesírást, még a legkisebb hibák is befolyásolhatják a prompt eredményét.

A nagy nyelvi modellek hatékony használatának kulcsa a megfelelő prompt kialakítása. Ahogy a korábban bemutatott példa is mutatja, egy jól megfogalmazott prompt egyértelművé teszi az elvárásokat és segít a modellnek pontos, releváns válaszokat adni. Ennek összefoglalása látható az alábbi ábrán is.

1. ábra: A prompt írásának szempontjai, saját szerkesztés



A modell működésének és sajátosságainak ismerete, a feladat pontos meghatározása és a stilisztikai részletekre való odafigyelés mind hozzájárulnak ahhoz, hogy a mesterséges intelligencia a lehető legjobban teljesítsen. Ahogyan a jó kommunikáció alapja a világos és részletes instrukciók adása, úgy a mesterséges intelligenciával való munka sikere is a precíz és átgondolt promptokon múlik. Ha figyelembe vesszük ezeket a szempontokat, akkor a nagy nyelvi modellek hatékony eszközzé válnak, amelyekkel kiváló eredményeket érhetünk el.

5.Kísérlet

A kutatásom második szekciójában gyakorlati úton vizsgálom a témát. Egy kíséreltet végzek középiskolás diákokkal. A kísérlet lényege, hogy megtudjam, hatékonyabban dolgoznak-e a diákok a nagy nyelvi modellek segítségével. Hogyan használják a mesterséges intelligenciát, illetve kik használják szívesebben. A kísérlet korlátozott, nincs kellően nagy merítése, hogy teljes képet adjon a hazai diákok képességeiről. A kísérlet célja tehát nem az, hogy globálisan vagy országosan érvényes megállapításokat tegyen a diákságra, inkább jellemző tendenciák felvázolása vezérli, melyek későbbi kutatások alapjául szolgálhatnak.

5.1.A kísérlet menete

A kísérlet lebonyolítására megkaptam az iskolám és az osztályfőnököm támogatását, így a győri Péterfy Sándor Evangélikus Gimnázium 11./VII osztályának tanulói voltak a résztvevők (n=35). A feladatok önkéntesek voltak, kiskorúak esetében szülő beleegyezésével vettek rajta részt a diákok, amelyről írásos nyilatkozatot töltöttek ki. A kísérlet kétfordulós volt, egy magyar irodalom és egy történelem tematikájú feladattal. Mindkét esetben a szaktanárok támogatására és közreműködésére hagyatkoztam a feladatkiadáskor és az értékelés folyamán is. A feladatok elvégzésére az első forduló esetében két hét, a második fordulónál egy hónap állt rendelkezésre a diákoknak. Ennek oka a feladatok időzítése volt. Az első feladat a tavaszi szünet előtt került kiadásra, és a szünet folyamán értékelte a szaktanár. A második feladat esetében az érettségik miatt a szaktanár csak május végén tudta megkezdeni a kiértékelést, viszont a feladatot korábban kiadtuk a nagyobb részvevőszám reményében.

Mindkét esetben két lehetőséget adtam a diákoknak. Az egyik, hogy használhatnak mesterséges intelligenciát, a másik, hogy nem használhatnak mesterséges intelligenciát a feladat elvégzéséhez. A feladatok mindkét esetben érettségi típusúak voltak, kiegészítve, hogy használni kell külső forrásokat. Ezt azért tartottam fontosnak, hogy akik éltek a mesterséges intelligencia nyújtotta lehetőségekkel, kritikusan gondolkodjanak az általuk adott információkkal kapcsolatban és ne hagyatkozzanak rá kizárólagosan. Az irodalmi feladat esetében egy magyar vers elemzését kellett elvégeznie a diákoknak. A feladat illeszkedett a tananyaghoz, így Ady Endre: A csókok átka című versét választottuk a szaktanárommal közösen. A dolgozat formátuma az általános normákat követte, 1-1,5 gépelt oldal terjedelmű dokumentumot kellett készíteniük, amelyet Times New Roman betűtípussal, 12-es betűmérettel és 1,5-es sortávval, sorkizárt formában nyújtottak be digitális formában. Az elemzés arányos bekezdésekre tagolt és címmel ellátott. Értelmezni kellett a verset, bemutatva a lírai én

gondolatait a teremtett világ múltjáról, jelenéről és jövőjéről. Emellett ki kellett térniük a szövegösszefüggést és a hangvételt megteremtő stílus eszközök és alakzatok szerepére, valamint a vers téma szerinti besorolhatóságának kérdéseire. A történelmi feladat során a feltételek hasonlóak voltak. Itt a téma a nagy gazdasági világválság kibontakozása és következményei Németországban. A két feladat témájában a lényegi különbség a hozzá kapcsolódó információk mértéke az interneten. Míg az irodalmi téma nagyon szűk és a magyar kapcsolódás miatt kimondottan kevés segítséget lehet találni hozzá, a történelmi téma kimondottan tág, nagyon sok forrás segíti a munkát.

Az elkészült dolgozatok értékeléséhez szaktanárokat kértem meg. Korábban kiadtam egy feladatot, amelyben arra kértem, hogy a dolgozatokat ne egyenként értékeljék, hanem egy visszajelző dokumentumban összesítsék az eredményeket. Ennek a feladatnak a célja az volt, hogy átfogó képet kapjak a dolgozatok általános színvonaláról és az egyes csoportok teljesítményéről. A visszajelző dokumentumnak az alábbi pontokat kértem, hogy jellemezzék. Először is emeljék ki a leggyengébb dolgozatot és a legjobb dolgozatot. Itt is fontos volt megvizsgálni a mesterséges intelligencia használatát, az átlagtól való eltérés mértékét és az érettségi követelményeknek való megfelelést. Ezen kívül általánosan értékelték a dolgozatok minőségét. Ez magában foglalta annak vizsgálatát, hogy az összes dolgozat milyen színvonalat képvisel, és hogy az egyes csoportok dolgozatai milyen minőségűnek tekinthetők. A visszajelző dokumentum célja az volt, hogy átfogó és részletes értékelést nyújtson a dolgozatok teljesítményéről, megkönnyítve ezzel a hiányosságok és a kiemelkedő teljesítmények azonosítását. Emellett egy általános véleményre is kíváncsi voltam, hogy a dolgozatokat látva mit gondolnak a témában.

A kísérlet gyakorlati része mellett felmértem a diákok háttérét is. Az első forduló után egy kérdőíven keresztül, a második esetben rövid írásos interjúk formájában. A kérdőívet az elején két részre bontottam. Annak függvényében haladhattak tovább a válaszadók az elágazódó kérdőívben, hogy milyen módszerrel készítették el a feladatot. Ezt követően kérdéseket fogalmaztam meg a résztvevő neméről, valamint a szülők legmagasabb iskolai végzettségéről. Továbbá, rákérdeztem az elmúlt félévi tanulmányi átlagra és magyar nyelvből elért tanulmányi átlagra is. A következő szekció arra irányult, hogy a résztvevők hallottak-e korábban a nagy nyelvi modellekről, illetve, hogy melyeket használták korábban és mire. Ha egyéb programokat ismertek, ezeket is fel kellett tüntetniük, valamint azt, hogy mire használták ezeket a programokat. A következő szakasz csak azokra vonatkozott, akik a feladat elvégzéséhez használtak mesterséges intelligenciát. Az alkalmazott programokat és azok használati céljait

vizsgálta. Kérdéseket tartalmazott arról, hogy mennyiben hagyatkoztak az elkészített anyagra, mennyire voltak elégedettek a végeredménnyel, mennyi időt töltöttek a feladat elvégzésével, és mennyivel érezték gyorsabbnak a munkát a mesterséges intelligencia használatával. A kérdőív utolsó szekciója Ady Endre "A csókok átka" című versével kapcsolatos kérdéseket tartalmazott. Ezek a kérdések a vers egyes motívumainak és szimbólumainak értelmezésére irányultak, mint például, hogy mikor történik az éjszaka, amikor az élet megfog a versben, milyen következményei vannak a csóknak, milyen képeken keresztül jelenik meg a csókok ereje, mit jelent a kétféle csók Ady szerint, hogyan jellemzi a csókok által született embereket, milyen irodalmi stílusokat képvisel a versben, és milyen jelentősége van az Olimpia kifejezésnek. Itt a válaszok helyességét vizsgáltam. Feltételezésem szerint jobban teljesítenek azok a diákok, akik nem használnak mesterséges intelligenciát, a feladattal töltött többletidő miatt. A kérdőív végén lehetőséget biztosítottam az egyéb kérdések és észrevételek megadására is, hogy teljes képet kapjunk a résztvevők tapasztalatairól és véleményéről.

5.2. A kísérlet eredményei

Elsőként az irodalmi és utána a történelmi feladatot fogom értékelni. A feladatokon belül pedig a szaktanári vélemény és utána a háttérkérdőív lesz a sorrend. A 20 résztvevő diák által írt esszéket egy javítótanár értékelte, aki előzetesen nem tudta, melyik diák mely módszert alkalmazta. Az értékelés célja az volt, hogy megállapítsa, melyik csoport tudta jobb minőségben elvégezni a feladatot. A javítótanárok feladata volt, hogy megpróbálják kitalálni, hogy mely dolgozatok készültek mesterséges intelligencia segítségével és melyek nem. Az első forduló találati aránya 60%. Ez a szám jól mutatja, hogy nem minden esetben könnyű kitalálni, hogy teljesen önállóan dolgozott-e az adott diák vagy nem, viszont sok esetben könnyen leszűrhető.

5.2.1. Irodalom teszt

5.2.1.1. Szaktanári értékelés

A leggyengébb esszékről azt gondolta, hogy mesterséges intelligencia segítségével készültek, de utólag kiderült, hogy ez nem volt igaz. Ezekben a dolgozatokban magyartalanságok, stílusterések és logikátlanságok voltak, amelyek a középszintű elvárásokat sem érték el. Ez arra utal, hogy a gyenge minőség nem feltétlenül a mesterséges intelligencia használatának eredménye, hanem a diákok saját kifejezőképességének és szerkesztési készségeinek hiányosságaira vezethető vissza. Azok a diákok, akik javaslatait ügyesen szerkesztették és logikailag felülbírálták, megfelelő stílusban tudtak írni. Az ő esetükben segítség lehetett a mesterséges intelligencia, de a végső tartalom minőségét elsősorban saját ötleteik határozták meg.

Azoknál a dolgozatoknál, ahol a javítótanár helyesen találta ki, hogy a diákok Chat GPT-t használtak, jellemző volt, hogy nem jelöltek meg más írásos forrást. Ez arra utal, hogy a diákok a mesterséges intelligenciát kizárólagos forrásként használták, ami a dolgozatok sablonos és felszínes jellegéhez vezetett. Azokban a dolgozatokban, ahol a diákok nem használtak Chat GPT-t, több önálló gondolat és egyéniség jelent meg. Ezekben az esetekben a diákok aktívan gondolkodtak a témáról és kifejezték véleményüket, nem csak a feladat teljesítésére törekedtek. Ez azt sugallja, hogy a mesterséges intelligencia nélküli munkák gyakran gazdagabbak voltak tartalom szempontjából, mivel a diákok saját véleményüket és értelmezésüket is belefoglalták az esszéikbe.

A dolgozatok összességét tekintve egyik sem érte el a jeles szintet tartalmi-szerkezeti logikátlanságok, formai hibák és stílusterések miatt. A diákok gyakran sablonos kifejezéseket és értelmetlen frázisokat halmoztak egymás mellé, anélkül, hogy részletezték volna állításaikat. A Chat GPT használata csak azoknál a diákoknál jelentett segítséget, akiknek eleve volt valamilyen alapgondolatuk a témáról, és ezeket ki is akarták fejteni. Számukra a mesterséges intelligencia használata nem hozott döntő minőségjavulást, de lerövidítette a megoldási időt. Ezek a diákok képesek voltak a mesterséges intelligencia javaslatait kritikusan és kreatívan felhasználni, és ezáltal javítani munkájuk minőségén.

Azok a diákok, akik kritikátlanul átvették a mesterséges intelligencia javaslatait, megfosztották magukat az önálló gondolkodás lehetőségétől. A Chat GPT használata így gátolhatja a kreativitást, ha a diákok nem csak kiindulási alapként tekintenek rá. A mesterséges intelligencia önállóan nem képes kreatív és eredeti gondolatokat generálni, amelyek egy irodalmi esszéhez szükségesek. Az önálló gondolkodás és a kritikai szemlélet elengedhetetlenek a minőségi munkához.

5.2.1.2. Az irodalom teszt háttérkérdőívének eredményei

Az alábbiakban részletesen elemzem a diákok által kitöltött kérdőívet a kérdések sorrendjében. Különös hangsúlyt fektetve a két csoport közti különbségekre. Sajnos nem minden diák töltötte ki a kérdőívet, az esszé elkészítői közül, így itt még kevésbé reprezentatívak az eredmények. Mindkét csoportban vegyesen szerepelnek férfiak és nők. Az AI-t használók között nagyobb arányban vannak férfiak, de ez nem jelentős különbség. Mindkét csoportban a szülők többsége felsőfokú végzettséggel rendelkezik. Az AI-t használók átlagosan 4.87-es tanulmányi átlagot értek el, míg az AI-t nem használók átlagosan 4.82-es eredményt. Magyar nyelvből az MI-t

használók átlaga 4.75, míg az MI-t nem használók 4.7-es átlagot értek el. Mindegyik érték kimondottan magas és nincs számottevő különbség, ezért valószínűleg a tanulmányi átlag nem befolyásolja a hajlandóságot, az önkéntesség háttérében más tényezők állhatnak.

A feladathoz MI-t használók többsége hallott már korábban különböző programokról, mint például ChatGPT, Bard, Gemini és Copilot. Mégpedig, különböző célokra alkalmazták korábban a mesterséges intelligenciát, például prezentációs anyagok készítésére, beadandók írására és összefoglalók készítésére. A feladat elvégzése során is a legtöbben szinte az egész munkát a ChatGPT-re bízta. A ChatGPT megnevezést itt tudatosan használtam, ugyanis az összes válaszadó ezt a programot vallotta egyedüli segítségének a nyelvi modellek oldaláról, sőt korábban is csak ezt használták a feladataikhoz. Azok közül, akik nem éltek ezzel a segítséggel a feladat elvégzéséhez, kevesebben ismerték vagy használták a nyelvi modelleket.

A mesterséges intelligenciát használók közül sokan sokkal gyorsabbnak találták a feladatvégzést, mintha nem használták volna. Ők kevesebb mint 2, de többségük kevesebb mint 1 óra alatt végzett. A másik csoport ezzel szemben átlagosan 2-3 órát töltött a feladat elvégzésével. Összességében mindkét csoport elégedett volt a munkájával, viszont akik igénybe vették a nyelvi modell segítségét, többségben, könnyűnek találta a feladatot, míg a másik csoport számára a feladat, saját bevallásuk szerint, általában nagyobb kihívást jelentett. Mindkét csoport résztvevői hasonló válaszokat adtak a versértelmezési kérdésekre, ami arra utal, hogy a nyelvi modellek használata nem befolyásolta jelentősen a vers értelmezésének minőségét.

5.2.1.3. Összegzés

A mesterséges intelligencia használata önmagában nem garantálja a magas minőséget, ami a szaktanári értékelésből is világosan kiderül. Azok a diákok, akik kritikátlanul vették át az MI javaslatait, gyakran szembesültek azzal, hogy dolgozataik sablonossá és felszínessé váltak. Ez arra utal, hogy az MI-t használó diákoknak is szükségük van a saját kritikai gondolkodásukra és kreativitásukra ahhoz, hogy igazán jó minőségű munkát hozzanak létre. A szaktanári értékelés szerint a gyenge minőség nem feltétlenül az MI használatának eredménye, hanem inkább a diákok kifejezőképességének és szerkesztési készségeinek hiányosságaira vezethető vissza. Ez azt mutatja, hogy az MI nem helyettesítheti a diákok saját gondolatait és ötleteit, amelyek elengedhetetlenek a minőségi munkához.

Bár az MI nem helyettesítheti az önálló gondolkodást, hasznos eszközként szolgálhat a diákok számára, különösen azok esetében, akik már rendelkeznek megfelelő alapgondolatokkal. A

mesterséges intelligencia megfelelő használata esetén valóban támogathatja a diákok munkáját, de nem helyettesítheti a saját gondolkodási folyamatukat. Az önálló gondolatok és egyéniség különösen fontos szerepet játszanak a minőségi irodalmi esszék elkészítésében. A szaktanári értékelés alapján azok a dolgozatok, amelyekben a diákok nem használtak MI-t, gyakran gazdagabbak voltak tartalom tekintetében. A háttérkérdőíves elemzésből kiderült, hogy az MI-t használó diákok gyorsabbnak találták a feladatvégzést. A MI-t használók többsége kevesebb mint egy óra alatt végzett a feladattal, míg a nem használók átlagosan 2-3 órát töltöttek a feladat elvégzésével. Az időmegtakarítás azonban nem feltétlenül jár együtt a minőség javulásával.

Mindkét csoport, a MI-t használóknak és a nem használóknak is tetszett a saját munkájuk, tehát a MI használata nem befolyásolta jelentősen a diákok elégedettségét. Bár, míg a MI-t használók többsége könnyűnek találta a feladatot, a másik csoport számára nagyobb kihívást jelentett. Azonban fontos megjegyezni, hogy az elégedettség nem mindig korrelál a munka minőségével. A háttérkérdőíves elemzés azt mutatta, hogy a versértelmezési kérdésekre adott válaszok hasonlóak voltak mindkét csoportban, ami arra utal, hogy az MI használata nem befolyásolta jelentősen a vers értelmezésének minőségét. Az irodalmi elemzés és értelmezés képességei továbbra is nagymértékben függenek a diákok saját intellektuális erőfeszítéseitől és kritikai gondolkodásától.

Összefoglalva, a mesterséges intelligencia használata vegyes eredményeket hozott az irodalmi esszék esetében. Az MI hasznos eszköz lehet, ha a diákok kritikusan és kreatívan használják, és ha már rendelkeznek megfelelő alapgondolatokkal. Azonban az MI önmagában nem pótolhatja az önálló gondolkodást és kreativitást, amelyek elengedhetetlenek a minőségi irodalmi munkák elkészítéséhez. Az önálló gondolatok és egyéniség fontossága kiemelkedő, és a diákok saját szellemi erőfeszítései továbbra is döntő szerepet játszanak a dolgozatok minőségében.

5.2.2.Történelem teszt

Sajnálatos módon a tanév végére maradt ennek feladatnak a kiadása, így nagyon kevesen készítették el. Összesen hét dolgozat született. Viszont még ilyen csekély részvételi arány mellett is szembetűnő különbségek vannak ezen és a másik feladat eredményei között.

5.2.2.1.Szaktanári értékelés

Az elkészült hét esszé alapján megállapítható, hogy a diákok mindegyike elfogadható minőségű munkát nyújtott. A szaktanár véleménye szerint két diák, mesterséges intelligenciát

használhatott a feladat megoldása során. Bár a dolgozatok között nem voltak jelentős különbségek a tartalom és terjedelem tekintetében,

A leggyengébb valamelyest kilógott a sorból. Kevesebb konkrét tény és adatot tartalmazott, inkább általánosságokra épült. Ezért is feltételezi, hogy itt segítség volt a mesterséges intelligencia. Ennek ellenére, ha érettségi esszéről lenne szó, ez a megközelítés elegendő lenne a feladat jó megoldásához, bár a forráshasználat kompetenciája nem érvényesülne. Az utóbbi azonban most nem szempont, mivel előzetesen nem adott forrásokat a feladathoz. A korábban említett két dolgozat utolsó bekezdése hasonló stílusú, ami didaktikussá tette a megoldásukat. Ez a rész inkább elvett a dolgozatok értékéből, mint hozzáadott.

A legkidolgozottabb munka azonban meghaladná az érettségi követelményeinek terjedelmi kereteit, és a beillesztett táblázatok sem szükségesek egy érettségi dolgozatban. Valószínűleg az általa felhasznált szakirodalom volt túlságosan részletes, bár forrást nem tüntetett fel, a dolgozat alapos munka eredménye. Valószínűleg nem használt mesterséges intelligenciát.

Összességében elmondható, hogy mindegyik dolgozat megfelel a középszintű érettségi elvárásainak, és nem volt nagyon kiemelkedő dolgozat egyik irányban sem. Talán, amelyik a leghosszabb terjedelmű lett, hossz tekintetében, túlteljesítette a célt.

5.2.2.2. Rövid interjúk eredményei

A résztvevők száma miatt a visszajelzéseiket félig strukturált interjú formájában rögzítettem. A diákok válaszai alapján egyértelműen megállapítható, hogy a mesterséges intelligencia (MI) használata jelentős hatással volt a feladatmegoldás folyamatára és eredményére. Az alábbiakban részletes következtetéseket vonok le az alapján, hogy a diákok használtak-e MI-t vagy sem.

Az MI-t használó diákok jelentős időmegtakarítást értek el, ami átlagosan 10-30 percet vett igénybe. Ez a hatékonyság azonnal szembeötlő, különösen a több órát eltöltő diákokkal szemben, akik nem használtak MI-t. Az MI gyorsan képes volt összefoglalni az információkat és megfogalmazni az esszét, így a diákoknak kevesebb időt kellett a feladatra fordítaniuk. Az MI jelentős mértékben növelheti a feladatvégzés hatékonyságát, különösen időszűkében lévő diákok esetében. Azonban a gyorsaság gyakran a mélyreható elemzés és a részletes ellenőrzés rovására mehet.

Az MI-t használó diákok közül többen is megemlítették, hogy az MI által generált szöveget nem ellenőrizték vagy javították ki alaposan, mert megbíztak a programban vagy nem akartak több időt szánni rá. Ez bizonyos mértékű bizalmatlanságot is tükrözött az MI iránt, attól tartva,

hogy az esetlegesen hibás vagy félrevezető információkat tartalmazhat. Ezzel szemben az MI-t nem használó diákok sokkal több időt fordítottak a kutatásra, forrásgyűjtésre és a szöveg alapos megírására, ezért kimondható, hogy a MI használata nem feltétlenül vezet magasabb minőségű munkához. A diákoknak továbbra is fontos, hogy alaposan ellenőrizzék és javítsák az MI által generált szöveget, hogy biztosak legyenek a tartalom helyességében és relevanciájában.

A MI-t használó diákok legnagyobb nehézségei közé tartozott a programmal szembeni bizalmatlanság és a magyar nyelvű fordítások pontatlansága. A MI-t nem használó diákok számára a legnagyobb kihívást a hiteles források megtalálása és a nagy mennyiségű információ feldolgozása jelentette. Ezek a nehézségek rámutatnak arra, hogy bár az MI gyorsan generál tartalmat, a források hitelessége és a részletes információk feldolgozása továbbra is komoly kihívás marad a diákok számára. A MI használata csökkentheti a kutatási időt, de a diákoknak továbbra is fontos figyelmet kell fordítaniuk a források hitelességére és az információk pontos feldolgozására. Az MI-t nem használó diákoknak pedig meg kell tanulniuk hatékonyabb forráskutatási technikákat.

A diákok egybehangzóan állították, hogy éles helyzetben több időt és energiát fektettek volna a feladatba. Ez különösen igaz azokra, akik MI-t használtak, mivel ők gyakran gyorsan, kevés átnézéssel végezték el a feladatot. Az MI-t nem használó diákok már a próbahelyzetben is alaposabb munkát végeztek, de éles helyzetben ők is több időt fordítottak volna az ellenőrzésre és az átfogalmazásra. Az éles helyzet nagyobb motivációt jelent a diákok számára a részletesebb és alaposabb munkavégzéshez. Az MI-t használó diákoknak különösen fontos lenne, hogy éles helyzetben is szánjanak időt az ellenőrzésre és a javításra.

5.2.2.3.Összegzés

Az összehasonlítás alapján látható, hogy az MI-t használó diákok jelentős időmegtakarítást értek el a feladatok elkészítésében, míg azok, akik hagyományos módon dolgoztak, jelentősen több időt fordítottak ugyanerre. Például az MI-t használók általában 10-30 perc közötti időtartamban végeztek, míg a hagyományos módszerekkel dolgozók 1-4 óra közötti időtartamot említettek. Ez az időbeli különbség jól tükrözi az MI hatékonyságát az időgazdálkodás szempontjából.

Az MI-t használó diákok gyakran említették, hogy bár az MI jelentős segítséget nyújtott a feladat gyorsabb megoldásában, bizalmatlanságuk volt a program pontosságával és a magyar nyelvű fordításokkal kapcsolatban. Ez a pontatlanság a szaktanári értékelésekben is

kiemelkedett, különösen akkor, ha a dolgozatok túlzottan általános összegzéseket vagy didaktikus megoldásokat tartalmaztak.

A hagyományos módszerekkel dolgozó diákoknak, a fő nehézséget a hiteles források megtalálása és a szükséges információk összegyűjtése okozta. Ez a szempont a szaktanári véleményekben is megjelent, hangsúlyozva az alapos kutatómunka fontosságát és annak hatását a dolgozat minőségére.

Az értékelések szerint a legjobban kidolgozott dolgozatok gyakran túlteljesítették az érettségi követelményeit, ami részben annak köszönhető, hogy ezek a diákok mélyebben ásták bele magukat a témába és alaposabb munkát végeztek, függetlenül attól, hogy használtak-e MI-t.

Összefoglalva, bár az MI használata gyorsabb megoldást eredményezhet, nem minden esetben járul hozzá a dolgozatok minőségének javulásához. A hagyományos módszerekkel dolgozók gyakran több időt és energiát fektetnek a feladatokba, ami alaposabb és részletesebb dolgozatokhoz vezethet. A szaktanárok értékelése alapján mindkét megközelítés elfogadható lehet az érettségi elvárások teljesítésére, de az alapos kutatómunka és a mélyebb tematikai elemzés előnyösebb lehet.

5.3.A kísérlet összegzése

A két teszt, az irodalmi és történelmi esszék összehasonlítására szolgált, valamint a mesterséges intelligencia (MI) használatával járó hatások vizsgálata volt a diákok munkájában. Az elemzés alapján számos fontos következtetés rajzolható ki.

Az MI használata időmegtakarítást eredményezett a dolgozatok elkészítésében, de ez nem mindig járt kézzelfogható minőségi javulással. Az MI segítségével gyorsabban meg lehetett írni az esszéket, de a tartalom gyakran sablonos és felszínes maradt, különösen, ha a diákok nem végeztek alapos javítást és szerkesztést az MI által generált szövegen.

Az MI-t használó diákoknak komoly kihívást jelentett az önálló gondolkodás és kreativitás megőrzése. Sokan nem kritikusan vizsgálták az MI által javasolt információkat, ami a sablonos és kevésbé egyedi dolgozatokhoz vezetett. Ez alapján látható, hogy az MI önmagában nem képes pótolni az emberi kreativitást és kritikai gondolkodást.

Bár az MI használói elégedettek voltak a munkájukkal, és gyorsabbnak találták a feladatmegoldást, ez nem mindig korrelált a munka minőségével. A hagyományos

módszerekkel dolgozó diákok általában több időt és energiát fordítottak a kutatásra és a szöveg kidolgozására, ami alaposabb és részletesebb esszéket eredményezett.

A történelmi esszék esetében is látszódtak az MI használatának hatásai, például gyorsabb elkészülés, de nem minden esetben járt minőségi előnnyel. Az irodalmi esszék esetében pedig az MI használata gyakran sablonossá és felszínessé tette a dolgozatokat, hiányoztak belőlük az egyéni gondolatok és mélyebb elemzések.

A tanárok értékelése szerint a gyenge minőségű dolgozatok nem feltétlenül az MI használatának, hanem inkább a diákok saját kifejezőképességének és szerkesztési készségeinek hiányosságaira vezethetők vissza. Ez alapján az MI támogathatja a munkafolyamatot, de nem helyettesítheti a diákok saját kritikai gondolkodását és kreativitását.

Összességében az MI használata potenciálisan segíthet a diákoknak a feladatok gyorsabb elkészítésében, de a minőségi munkához továbbra is szükség van az önálló gondolkodásra, kreativitásra és alapos szerkesztésre. A tanulási folyamat részeként fontos, hogy a diákok kritikusan kezeljék az MI által generált tartalmat, és ne tekintsék kizárólagos forrásként a feladatok megoldásához.

6. Konklúzió

Arra a kérdésre, hogy tud-e segíteni a diákoknak a mesterséges intelligencia, sajnos nem lehet egyértelmű választ adni. Megannyi aspektust kell vizsgálni, melyek közül csak párra volt lehetőségem, azonban azok esetében is csak korlátozott vizsgálatot tudtam végezni. Ennek ellenére mindenképpen le lehet vonni következtetéseket. Remélem, hogy eredményeim elősegítik a mesterséges intelligencia integrációját a köznevelésbe. Összefoglalójuk, rövid formában az alábbi táblázatban olvasható.

2. ábra: A kutatás összegző eredményei a kutatói kérdések alapján, saját szerkesztés.

K1: Melyik mesterséges intelligencián alapuló nyelvi platform a legalkalmasabb a magyar köznevelésben való alkalmazásra a diákok számára?	E1: Nem lehet kizárólagosságot adni egyetlen platformnak sem, de a kutatás több szempontból is kiemelt szoftvereket.: a MS Copilot feltünteti a forrásait, a Google Gemini és a Copilot korpuszuk mellett az interneten is keresnek, a két platform integrálható a fejlesztők további alkalmazásaival. A cél, hogy a diákok otthonosan használják ezt a technológiát, ne pedig egy platformot ismerjenek csak.
K2: Mely iskolai tantárgyak és milyen típusú feladatok megoldásában nyújthatnak segítséget a mesterséges intelligencia felhasználásával működő szoftverek?	E2: Alapvetően a sok lexikai tudást igénylő tantárgyakban hasznosabb lehet. A kreatív területeken az önálló munkát gátolhatja. Tehát fontos, hogy viszonylag kevés szellemi értéket hordozó feladatokban érdemes használni, melyek adott esetekben repetitívek is.
K3: Hogyan érdemes használni diákként a nagy nyelvi modelleket	E3: Csak a technológia működési alapjait, lehetőségeit és határait ismerve lehet felelősségteljesen használni. Nem szabad, hogy az egyén teljesítményének gátat szabjon, csak eszközként kell tekinteni rá.

Az első kérdésem volt, hogy mely nyelvi modell a legalkalmasabb a köznevelésbeli felhasználásra. A nagy nyelvi modellek elemzése alapján kitűnt, hogy nem lehet egyértelmű és egzakt válasz találni a kérdésre. A modell kiválasztása mindenképp összeforr a feladattal. Mindenképp fontos, hogy nem egy abszolút megoldás létezik, hanem feladat és szituáció függvényében kell megtalálni a legjobb eszközt. Alapvetően létezik két irány melyet legjobban a ChatGPT és a Copilot szentesít meg. Az előbbi a szövegekkel való munkában, utóbbi az információkeresésben nevezhető jobbnak. Az oktatásban tehát nem egy kimondott modellt kell kínálni a diákok számára, hanem bátorítani kell őket a kísérletezésre. Ehhez szükséges bemutatni az egyes opciókat. Ezt próbáltam meg felvázolni a diákoknak a feladatok kiadásakor, ahol egy rövid prezentáció keretében bemutattam az egyes általam vizsgált modelleket. Sajnos, ezzel nem sikerült elérnem a kívánt hatást, ugyanis a diákok megmaradtak az általuk megismert és megszokott eszközöknél.

A mesterséges intelligencia technológiai alapjai szerint a platformok egy adott szövegtörzsből képesek szöveget generálni, korábbi információk alapján tudnak csak összefüggő szöveget alkotni. Ezek alapján mindenképpen olyan feladatok elvégzésében tartom hatékonynak, ami leszűkíti a merítést, és kisebb eséllyel tud csak tévútra futni. Folytatva a gondolatmenetet, mely szerint csak korábbi információkból képes dolgozni, nem tud egyéni gondolatokat megfogalmazni, egzaktabb tudományterületeken hatékonyabb a használata. Ezt erősíti meg, hogy a történelmi feladatban sokkal jobban szerepeltek átlagosan a diákok, mint az irodalmi esetében, ahol egy verset kellett elemezni a saját tudásuk és kreativitásukat felhasználva. A szaktanárok által készített visszajelző dokumentumok alapján megállapítható, hogy a diákok teljesítménye változó volt, és a mesterséges intelligencia használata egyes esetben előnyös volt, míg más esetben nem hozott jelentős javulást. Ez a megállapítás rávilágít arra, hogy a mesterséges intelligencia alkalmazása nem helyettesítheti a diákok kreatív gondolkodását és önálló munkáját, hanem inkább kiegészíti azt.

Diákként rengeteg dologtól függ a nagy nyelvi modellek felhasználása. Természetesen a leglényegesebb az egyén oldala, a felelősségteljes használat. Meg kell ismerni a technológia korlátait és lehetőségeit. A megfelelő használathoz rengeteg tapasztalat szükséges, mind az egyén, mind a társadalom oldaláról és ezen tapasztalatok alapján kell dolgozni. Nem szabad megragadni egy szinten, mert jelenleg ez a tudományterület a leggyorsabban fejlődő. Napról napra nagyobb tudással rendelkeznek ezek a modellek, mivel visszacsatoláson alapul a működésük. Ez egy kimondottan nehéz feladat, ugyanis a résztvevők válaszai igazolják, hogy

a diákok a nagyon konzervatívan állnak az új módszerekhez. Figyelniük kellene arra, hogy a MI nem képes kiváltani a saját teljesítményüket. Ez egy nehéz dolog, ugyanis, a feladat során, a legtöbb diák nehezen rugaszkodott el a ChatGPT által kiadott szövegtől. A legfontosabb dolog a tanulók számára, hogy informálódjanak a technológiáról és ismerjék azt. Kellőképpen hagyatkozzanak rá, de semmiképpen ne legyen gátja a kreativitásuknak. Fontos tudni, hogy melyik modell mire való, mire tudják használni. A legfontosabb pedig, hogy fejlődjenek együtt a technológiával.

Összességében a kutatás megerősítette, hogy a mesterséges intelligencia hasznos eszköz lehet a köznevelésben, különösen a strukturált és diszciplináris tantárgyak esetében. Ugyanakkor hangsúlyozni kell, hogy a mesterséges intelligencia nem helyettesítheti az emberi kreativitást és a kritikai gondolkodást, amelyek elengedhetetlenek a tanulási folyamatban. A jövőbeni kutatásoknak érdemes lenne további mélységben vizsgálniuk a mesterséges intelligencia és a diákok közötti interakciót, valamint a kreatív feladatokban való alkalmazásának lehetőségeit. Bár a kutatásom legfőképpen a diákok oldaláról közelítette meg a témát, mindenképpen érdemes kiemelni, hogy a tanári segítség elengedhetetlen. A tanulók önmaguktól nehezen rugaszkodnak el a megszokottól. A mesterséges intelligencia integrálása a tanulási folyamatba tehát nem csupán technológiai kérdés, hanem pedagógiai kihívás is, amely új megközelítéseket és módszereket igényel a tanárok és diákok részéről egyaránt.

Idézett források jegyzéke

Adamopoulou, E. Moussiades, L. (2020) 'Chatbots: History, technology, and applications. International Hellenic University, Agios Loukas: <https://doi.org/10.1016/j.mlwa.2020.100006>, (2024.07.27.)

Agomuoh, F. (2024) What is an AI token?, Digitaltrends, <https://www.digitaltrends.com/computing/what-is-an-ai-token/#dt-heading-what-is-an-ai-token> (2024.06.10)

Aqlan, A, Manjula, B. Lakshman, R. (2019) 'A Study of Sentiment Analysis: Concepts, Techniques, and Challenges', Singapore: Springer Singapore (Lecture Notes on Data Engineering and Communications Technologies), pp. 147–162. https://doi.org/10.1007/978-981-13-6459-4_16, (2024.06.14.)

Ciornei, S.(2024), AI Showdown: Microsoft Copilot vs. Google Gemini vs. ChatGPT 3.5 vs. Mistral vs Claude 3, Medium, <https://medium.com/thereach-ai/ai-showdown-microsoft-copilot-vs-google-gemini-vs-chatgpt-3-5-vs-mistral-vs-claude-3-f9ab1e3c314c>

Cope, B., Kalantzis, M. and Sears, D. (2021) 'Artificial intelligence for education: Knowledge and its assessment in AI-enabled learning ecologies', *Educational Philosophy and Theory*, 53(12), pp. 1229–1245. <https://doi.org/10.1080/00131857.2020.1728732>, (2024.07.23.)

Gardner, H.E. (2011) *Frames of Mind: The Theory of Multiple Intelligences*. Basic Books. https://books.google.hu/books/about/Frames_of_Mind.html?id=2IEfFSYouKUC&redir_esc=y, (2024.07.27.)

Holdsworth, J. (2024), What is NLP (natural language processing)?, IBM Topics, <https://www.ibm.com/topics/natural-language-processing> (2024.06.10.)

Hsiao, S. (2024) Bard becomes Gemini: Try Ultra 1.0 and a new mobile app today, Google, The Keyword <https://blog.google/products/gemini/bard-gemini-advanced-app/> (2024.06.10.)

Imran, M. and Almusharraf, N. (2024) 'Google Gemini as a next generation AI educational tool: a review of emerging educational technology', *Smart Learning Environments*, 11(1), p. 22, <https://doi.org/10.1186/s40561-024-00310-z>. (2024.07.27.)

Kevian, D. Syed, U. Guo, X. Havens, A. Dullerud, G. Seiler, A. Qin, L. Hu, L. (2024) 'Capabilities of Large Language Models in Control Engineering: A Benchmark Study on GPT-4, Claude 3 Opus, and Gemini 1.0 Ultra'. arXiv. <http://arxiv.org/abs/2404.03647> (2024.06.12.).

Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2017) 'ImageNet classification with deep convolutional neural networks', *Communications of the ACM*, 60(6), pp. 84–90. <https://doi.org/10.1145/3065386>, (2024.07.27.)

Laird, J. 2024 What Is Claude 3? Anthropic's New AI Models Explained Forrás: <<https://tech.co/news/what-is-claude-3-anthropic>> (2024.06.27.)

Lee, M. (2023) 'A Mathematical Interpretation of Autoregressive Generative Pre-Trained Transformer and Self-Supervised Learning', *Mathematics*, 11(11), p. 2451. <https://doi.org/10.3390/math11112451>. (2024.06.10.)

Lee, U. Jung, H. Jeon, H. Sohn, Y. Hwang, W. Moon, J. Kim, H. Authors Info & Claims (2023) 'Few-shot is enough: exploring ChatGPT prompt engineering method for automatic question generation in english education', *Education and Information Technologies*, <https://doi.org/10.1007/s10639-023-12249-8>. (2024.07.10.)

Magyarország Digitális Oktatási Stratégiája, <https://digitalisjoletprogram.hu/hu/tartalom/dos-magyarorszag-digitalis-oktatasi-strategiaja> (2024.06.11.)

Masalkhi, M. Ong, J. Waisberg, E. Zaman, N. Sarker, P. Lee, A. Tavakkoli, A (2024) 'A side-by-side evaluation of Llama 2 by meta with ChatGPT and its application in ophthalmology', *Eye* [Preprint], <https://doi.org/10.1038/s41433-024-02972-y>. (2024.07.10.)

McCarthy, J. Minsky, M. Rochester, N. Shannon, C. (2006) 'A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955', *AI Magazine*, 27(4), pp. 12–12. <https://doi.org/10.1609/aimag.v27i4.1904>. (2024.06.10.)

Zhao, Q. Xu, M. Gupta, K. Asthana, A. Zheng, L. Gould, S (2024) The First to Know: How Token Distributions Reveal Hidden Knowledge in Large Vision-Language Models? <https://doi.org/10.48550/arXiv.2403.09037>

Negin Yazdani Motlagh *et al.* (2023) ‘War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education’, *Journal of Applied Learning & Teaching*, 6(1), <https://doi.org/10.37074/jalt.2023.6.1.23>. (2023.07.10.)

Roumeliotis, K.I., Tselikas, N.D. and Nasiopoulos, D.K. (2023) ‘Llama 2: Early Adopters’ Utilization of Meta’s New Open-Source Pretrained Model’, <https://doi.org/10.20944/preprints202307.2142.v2>. (2024.06.27.)

Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986) ‘Learning representations by back-propagating errors’, *Nature*, 323(6088), pp. 533–536., <https://doi.org/10.1038/323533a0>. (2024.06.23.)

Oktatási Hivatal https://www.oktatas.hu/koznevelas/ajanlas_szemelyes_talalkozas_nelkuli_oktatas_nevelas_modszereire/tavoktatast_tamogato_eszkozok_kreta_rendszer; [Jogszábályi háttér - KRÉTA Elektronikus napló - KRÉTA Tudásbázis \(ekreta.hu\)](https://www.oktatas.hu/koznevelas/ajanlas_szemelyes_talalkozas_nelkuli_oktatas_nevelas_modszereire/tavoktatast_tamogato_eszkozok_kreta_rendszer) (2024.06.10.)

Oxford English Dictionary <https://www.oxfordlearnersdictionaries.com/definition/english/corpus?q=corpus>, (2024.05.27.)

Hossenfelder, S *Lost in Math: How Beauty Leads Physics Astray*, Basic Books, 2018, <https://link.springer.com/article/10.1007/s00016-019-00233-0> (2024.01.24.).

Shankland, S. (2024.), OpenAI Dall-E 3 Review: Generative AI for Fanciful, Fun Illustrations, cnet, <https://www.cnet.com/tech/services-and-software/openai-dall-e-3-review-generative-ai-for-fanciful-fun-illustrations/> (2024.07.14.)

Singh, S.K., Kumar, S. and Mehra, P.S. (2023) ‘Chat GPT & Google Bard AI: A Review’, in *2023 International Conference on IoT, Communication and Automation Technology (ICICAT)*., Gorakhpur, India: IEEE, pp. 1–6. <https://doi.org/10.1109/ICICAT57735.2023.10263706>. (2024.06.13.)

Somoye, F. 2024. ChatGPT’s word limit is 3000, but here’s how to get around it, pcguide, <https://www.pcguides.com/apps/chat-gpt-word-limit/> (2024.06.25.)

Sun, Y. Wang, S. Feng, S. Ding, S. Pang, C. Shang, J. Liu, J. Chen, X. Zhao, Y. Lu, Y. Liu, W. Wu, Z. Gong, W. Liang, J. Shang, Z. Sun, P. Liu, W. Ouyang, W. Yu, D. Tian, H. Wu, H. Wang, H. (2021) ‘ERNIE 3.0: Large-scale Knowledge Enhanced Pre-training for Language Understanding and Generation’. arXiv., <http://arxiv.org/abs/2107.02137> (2024.06.05.).

Wickman, S. (2023), Transforming Education into Chatbot Chats, KHT Royal Institute of Technology, <https://www.diva-portal.org/smash/get/diva2:1814507/FULLTEXT01.pdf>
(2024.07.27.)

Mellékletek

1. számú melléklet: Fogalomtár

Magyar kulcsfogalom	Angol kulcsfogalom	Meghatározás
Bemenet	Prompt	Rövid szöveges utasítás vagy kérdés, amely irányítja a nyelvi modellt a válasz generálásában. (Negin Yazdani Motlagh <i>et al.</i> , 2023)
Gépi tanulás	Machine Learning	Informatikai módszer, amely statisztikai technikákat alkalmaz az előrejelzésekhez az észlelt minták alapján.(Zhai, C., Massung, S. 2016)
Mélytanulás	Deep Learning	Többretegű neurális hálózatok, amelyek nagy mennyiségű adat alapján tanulnak és fejlesztik magukat (Krizhevsky et al. 2017; Rumelhart et al. 1986).
Szövegtörzs	Corpus	Írott és beszélt szövegek összessége, amelyet nyelvi modellek tanítására használnak (Oxford English Dictionary, 2024).
Érzelelemzés	Sentiment Analysis	A szöveges adatok érzelmi polaritásának (pozitív, negatív, semleges) meghatározása (Aqlan, Manjula and Lakshman Naik, 2019).
Természetes nyelvfeldolgozás	Natural Language Processing (NLP)	Számítógépes módszer, amely lehetővé teszi a szöveg és beszéd emberi módon történő megértését. Az NLP ötvözi a számítógépes nyelvészetet, a statisztikai, gépi tanulás és mély tanulás modellekkel (IBM, 2024)
Chatbot	Chatbot	Mesterséges intelligencia által működtetett program, amely szöveges vagy hangos formában kommunikál a felhasználókkal (Singh, et al. 2023).
GPT	Generative Pre-trained Transformer	Előre betanított nyelvi modell, amely a Transformer architektúrát alkalmazza a természetes nyelvfeldolgozáshoz (Lee, 2023)
LLM	Large Language Model	Nagy nyelvi modell, amely kiterjedt szöveges adathalmazokon tanul, hogy felismerje és megértse az emberi nyelv mintázatait (Wickman, 2023).

2. számú melléklet: Modellek összehasonlítása (táblázat)

Szempont	Bing Chat / Microsoft Copilot	Google Gemini	Baidu Ernie	Meta Llama 2	Claude 3	ChatGPT 3.5 / 4
Fejlesztő	Microsoft	Google	Baidu	Meta	Anthropic	OpenAI
Nyelvi modell / technológiai alap	GPT-4	Gemini 1.0 / 1.5	Ernie 4.0	Llama 2	Claude 3 Opus / Sonnet / Haiku	GPT-3.5 / GPT-4
Információbázis	internethozzáférés	internethozzáférés	nincs pontos adat	Nyílt forráskódú	nincs pontos adat	Internetes adatok, 2022-ig
Ár	Ingyenes minden magán Microsoft fiókhoz.	Sima verzió: Ingyenes Advanced verzió: 20 USD havonta	Ingyenes	Ingyenes	Van ingyenes és fizetős verzió is, a fizetős verziók token alapú díjazást alkalmaznak.	A ChatGPT 3.5 ingyenes, a ChatGPT 4 fizetős, 20\$havonta.
Funkcionalitás	elsősorban keresés, információfeldolgozás	elsősorban keresés, információfeldolgozás	Szöveg, vizuális tartalom generálás	Szövegfeldolgozás, nyílt fejlesztés	Különböző modellverziók, érvelési képességek	Szövegenerálás, információkeresés
Feladatkör	Keresés, válaszadás, beszélgetési stílusok	Szövegelemzés, programozás, logikai feladatok	Szöveges tartalom generálás, kínai nyelvi feldolgozás	Dokumentumalapú kérdések megválaszolása	Érvelési feladatok, ügyfélszolgálati chatbot	Szövegalkotás, információkeresés
Korlátok	4000 karakteres prompt méret	Csak 18 év felett használható	Kínai nyelvű	Kiseb tudásbázis mint mások	Verziók közötti különbségek, költségek	Tudás 2022-ig, prompt komplexitás
Hátrányok	Korlátozott prompt méret	Korlátozott felhasználói csoportok	Kormányzati cenzúra, elfogultság	Nem kiemelkedő más modellekhez képest	Független tesztek hiánya	Elfogultság, pontatlanság néha
Specifikus tulajdonság	Ingyenes hozzáférés, vizuális funkciók	Integráció Google Workspace-szel	Különböző kínai dialektusok kezelése	Nyílt forráskódú, testreszabhatóság	Érvelési képességek, különböző modellverziók	Képes szolgálni forrásokkal (GPT-4)
Egyéb	Különböző beszélgetési stílusok	Ingyenes és fizetős verziók	Jelentős felhasználói bázis Kínában	Átláthatóság, követhetőség	Amazon támogatás, nagy befektetés	Nagy adatbázis, különböző verziók

3. számú melléklet: kiadott feladatok

Történelem:

Teams feladatként kiadott pontos leírás:

„Szorgalmi feladat

Határidő: 2024. május 31. 23:59

Utasítások

Vaskó András kérésének megfelelően a következő témát kellene kifejteni: A nagy gazdasági világválság kibontakozása és következményei Németországban. (gazdasági, társadalmi és politikai téren)”

Irodalom:

Teams feladatként kiadott pontos leírás:

„MCC-kutatás, szövegalkotási feladat

Határidő: 2024. március 11. 23:59•

Lezárás: 2024. március 11. 23:59

Utasítások

- 1-1,5 gépelt oldal terjedelemben
- Times New Roma betűtípus
- 12-es betűméret
- 1,5-es sortáv
- sorkizárt
- van címe
- arányos bekezdésekre tagolt

Értelmezze **Ady Endre: A csókok átk**a c. versét! Mutassa be a lírai én gondolatait a teremtett világ múltjáról, jelenéről és jövőjéről! Térjen ki a szövegösszefüggést és a hangvételt megteremtő stíluseszközök és alakzatok szerepére, a vers téma szerinti besorolhatóságainak kérdéseire!”

4. számú melléklet: Diákok összehasonlítása - Táblázat

Szempont	MI-t használó diákok	MI-t nem használó diákok
Feladatvégzés ideje	Gyors (10-30 perc)	Lassabb (1-4 óra)
Tartalom minősége	Gyakran sablonos és felszínes	Általában alaposabb és részletesebb
Kreativitás és egyedi gondolatok	Korlátozott, gyakran hiányzik	Több egyedi gondolat és értelmezés
Forráshasználat	Főként az MI által generált tartalom	Változatos források, alaposabb kutatómunka
Szerkesztési készségek	Gyakran minimális utólagos szerkesztés	Alaposabb szerkesztés és átdolgozás
Elégedettség a munkával	Általában elégedettek	Elégedettek, de nagyobb kihívásként élték meg
Nehézségek	MI pontosságával kapcsolatos bizonytalanság, fordítási problémák	Hiteles források megtalálása, nagy mennyiségű információ feldolgozása
Tanári értékelés	Elfogadható, de gyakran sablonos	Változó, de gyakran részletesebb és egyedibb
Versértelmezés (irodalom)	Hasonló színvonalú a nem MI-t használókéhoz	Hasonló színvonalú az MI-t használókéhoz
Történelmi esszék minősége	Elfogadható, de néha túl általános	Változó, de gyakran részletesebb
Önálló gondolkodás mértéke	Alacsonyabb, gyakran az MI javaslataira támaszkodnak	Magasabb, saját gondolatok és értelmezések
Időgazdálkodás hatékonysága	Magas	Alacsonyabb
Tanulási folyamat mélysége	Felszínesebb, kevesebb saját kutatás	Mélyebb, több önálló kutatás és gondolkodás
Éles helyzetben várható teljesítmény	Valószínűleg alaposabb munkát végeznének	Már a próbahelyzetben is alaposabb munkát végeztek
Fejlesztendő területek	Kritikai gondolkodás, önálló kutatás, szerkesztési készségek	Időgazdálkodás, hatékony információfeldolgozás

5. Számú melléklet – kérdőív és interjúkérdések

Interjúkérdések

Használtál mesterséges intelligenciát a feladat elvégzéséhez?

Amennyiben igen

Milyen programot/programokat használtál és miért?

Miben tudott neked segíteni?

Mennyi időt töltöttél a feladattal?

Mennyire volt a mesterséges intelligencia munkája?

Mi volt a legnagyobb nehézség, amivel találkoztl?

Több időt töltöttél volna vele, ha élesben ment volna és számított volna az eredmény?

Mit változtatnál, ha máskor találkoznál hasonló feladattal?

Kérdőív

Melyik csoportot választottad?

Előzetes információk

Nemed?

Édesanyád legmagasabb iskolai végzettsége?

Édesapád legmagasabb iskolai végzettsége?

Elmúlt félév tanulmányi átlagod egy tizedesjegyre kerekítve?

Elmúlt félév tanulmányi átlagod magyar nyelvből tizedesjegyre kerekítve?

Az előadás előtt hallottál-e már az alábbi programokról?

Ha az egyébre nyomtál, miről hallottál még?

Melyiket használtad már korábban?

Ha az egyébre nyomtál, mit használtál még?

Ha használtál már korábban ilyen programot, mire használtad?

Ha az egyébre nyomtál, mire használtad?

Feladattal kapcsolatos információk (Az AI-os kérdések csak annak a csoportnak voltak)

Mennyire találtad nehéznek a feladatot?

Melyik programot használtad?

Ha az egyébre nyomtál, mit használtál?

Mire használtad?

Mennyiben hagyatkoztál az általa létrehozott anyagra?

Mennyire voltál megelégedve a végeredménnyel?

Mennyi időt töltöttél a feladat elvégzésével?

Mennyivel érezted gyorsabbnak a munkát a megszokott (MI nélküli, önálló) feladatvégzéshez képest?

Jobbnak érzed a munkádat, mintha nem használtad volna?

Verssel kapcsolatos tesztkérdések:

Mikor történik az éjszaka, amikor az élet megfog a versben?

Milyen következményei vannak annak, ha az élet csókba téved minden kis mozgás során?

Milyen képeken keresztül jelenik meg a csókok ereje a versben?

Mit jelent számára Ady Endre szerint a két csókféle, a forró és a lanya?

Hogyan jellemzi Ady a csókok által született embereket?

Milyen irodalmi stílusokat képvisel Ady Endre a versben?

Milyen jelentősége van a versben szereplő Olimpia kifejezésnek?

7.2 Mesterséges intelligencia a pályamunka kutatási témájához kapcsolódik

NYILATKOZAT

Kijelentem, hogy a pályamunka elkészítése során, a kutatási téma kapcsolódása miatt, mesterséges intelligencia alkalmazást/programot (pl. képalkotó, szövegalkotó) használtam.

Jelen nyilatkozat aláírásával tudomásul veszem, hogy amennyiben bizonyítható, hogy a dolgozatot nem magam készítettem, a dolgozattal kapcsolatban szerzői jogi jogsértés ténye merül fel, a Szekció megtagadja a dolgozat befogadását.

Dátum: 2024.09.11.



Aláírás
VASKÓ ANDRÁS

Név (olvashatóan)