



Mesterséges intelligencia alapú beszéd felismerés
a mindennapi munkavégzésben

EGYSZERŰ DIKTÁLÁS A FÁRADÁSÁGOS GÉPELÉS HELYETT



Horváth Kata



A mesterséges intelligencia alapú alkalmazások fontossága napjainkban már nem vitatott, azok tényleges felhasználási területeit és üzleti környezetben való alkalmazását viszont sokan nem ismerik. Cikkünkben összefoglaljuk, hogy mely felhasználási eseteken keresztül lehet a napi munkavégzésünket a beszéd felismerés által könnyen felgyorsítani.

Habár a jelenlegi digitális világunkban mindennapjaink jelentős részében használjuk mobiltelefonunkat és számítógép vagy tablet eszközeinket, a gépelés feladata még mindig egy mostoha feladatnak minősül. Sokan a többéves rutin dacára is lassan gépelnek számítógépen, mások a mobilkészülékek apró billentyűzetmezőivel küzdenek, vagy esetleg számos elgépelést vétenek. És akkor még nem ejtettünk szót arról, hogy mennyivel időigényesebb

tevékenység a gépelés a szóbeli beszédnél. Ezekre a helyzetekre nyújtanak hatékony megoldást az úgynevezett deep learning algoritmusokra épülő beszéd felismerő szoftveralkalmazások az elhangzott beszéd írásbeli rögzítése által.

A LEGGYAKORIBB FELHASZNÁLÁSI ESETEK

Számos iparágban és munkakörben jellemző napi feladatnak minősül a megbeszéléseken és tárgyalásokon

elhangzott információk rögzítése. A beszéd felismerő megoldások által a teljes megbeszélés szövege egyszerűen leiratozható, akár a beszélők megkülönböztetésének jelölésével. Az AI (Artificial Intelligence azaz Mesterséges intelligencia, MI – a szerk.) által támogatott diktálás emellett olyan rutinfeladatokban is segíthet, mint az e-mailek megírása vagy saját dokumentumok elkészítése. Bizonyos munkakörökben a munka jellege vagy a helyszíni adottságok

miatt nem lehetséges az azonnali gépelés, gondoljunk például egy megrendelővel történő helyszíni egyeztetésre. Ilyen esetekben elég elindítani a telefonunk hangrögzítő alkalmazásának segítségével egy felvételt, melyet utólag valamely beszédfelismerő alkalmazásba feltöltve könnyedén áttekinthető szöveggé alakíthatjuk az elhangzott információkat. Ezáltal elkerülhető a hibásan megjegyzett információk alapján történő munkavégzés.

Az AI alapú megoldások segítségével ezek a hanganyagok könnyen kereshetővé válhatnak, így nem szükséges egyetlen részletért órákon át különböző hang- és videóanyagokban keresgélni. A keresett kifejezés fájlban való előfordulását másodpercre pontosan meg tudják határozni a rendszerek, így könnyen meghallgatható a szükséges részlet. A beszédfelismerésre épülő alkalmazások további fontos felhasználási lehetősége, hogy képesek azonosítani az elhangzottak mil-

liszekundumra pontos idejét egy fájlban, így általa egyszerűen készíthető pontosan illeszkedő felirat videókhoz, mindennemű technikai képzettség nélkül. Mivel kutatások alapján a mobiltelefonon nézett videós tartalmak jelentős részét hang nélkül nézik a felhasználók, a megfelelő videófeliratozás már elengedhetetlen.

BESZÉDFELISMERÉST KANÁLÓ ALKALMAZÁSOK

Angol nyelven már számos megoldás létezik a beszéd leiratozására, viszont a magyar nyelvre optimalizált megoldásokat kínáló szereplők száma ennél jelentősen alacsonyabb. Cikkünkben a teljesség igénye nélkül összegyűjtöttünk néhány megoldást, amelyet ön is kipróbálhat és a mindennapi tevékenységeit, munkáját kényelmesebbé és gyorsabbá teheti általa.

Az Alrite beszédfelismerő szoftvermegoldás a digitális transzformáció területén szakértő Régens Zrt. saját

fejlesztésű megoldása. Az Alrite alkalmazás 95%-os beszédfelismerési pontosságával, letisztult felületével és széles körű funkcionalitásával kiemelten népszerű a felhasználók körében. Az alkalmazásban elérhetőek a korábban rögzített hanganyagok, lehetőség van mind közvetlen diktálásra, mind korábban rögzített hangfájl feltöltésére és leiratozására. Videókhöz villámgyorsan készíthető általa pontosan illeszkedő felirattá. A megoldás nagy előnye más elérhető alkalmazásokhoz képest, hogy megfelelően kezeli a központozást, vagyis az írásjelek elhelyezését a szövegben, jelentősen lerövidítve az utólagos szerkesztés idejét. Az alkalmazás ingyenesen kipróbálható a www.regens.com/alrite oldalon regisztrálva.

A Google széles körben népszerű fejlesztései mellett a beszédfelismerésre is hangsúlyt fektet, amely szintén 95%-os pontossággal képes felismerni a magyar nyelvű beszédet. A megoldást számos hozzájuk kapcsolódó termékben alkalmazzák, ilyen például a Google Docs-ban elérhető diktálás vagy az *Azonnali átírás* (Live Transcribe) nevű applikáció, amely a Google Play áruházból letölthető.

Az iOS eszközöket használók számára releváns megoldás az Apple által fejlesztett beszédfelismerő algoritmus, amely az iPhone/iPad és Mac készülékeken könnyíti meg az adatbevitelt. Üzenetek írása-
kor vagy jegyzetek felvételekor a billentyűzeten elérhető mikrofon ikonnal egyszerűen átválthat a hang alapú adatbevitelre. ■

A szerző a Régens Zrt. marketing igazgatója

Képek:

<https://www.istockphoto.com>

A diktálás mellett olyan rutinfeladatokban is segíthet a szövegfelismerő alkalmazás, mint például az e-mailek megírása