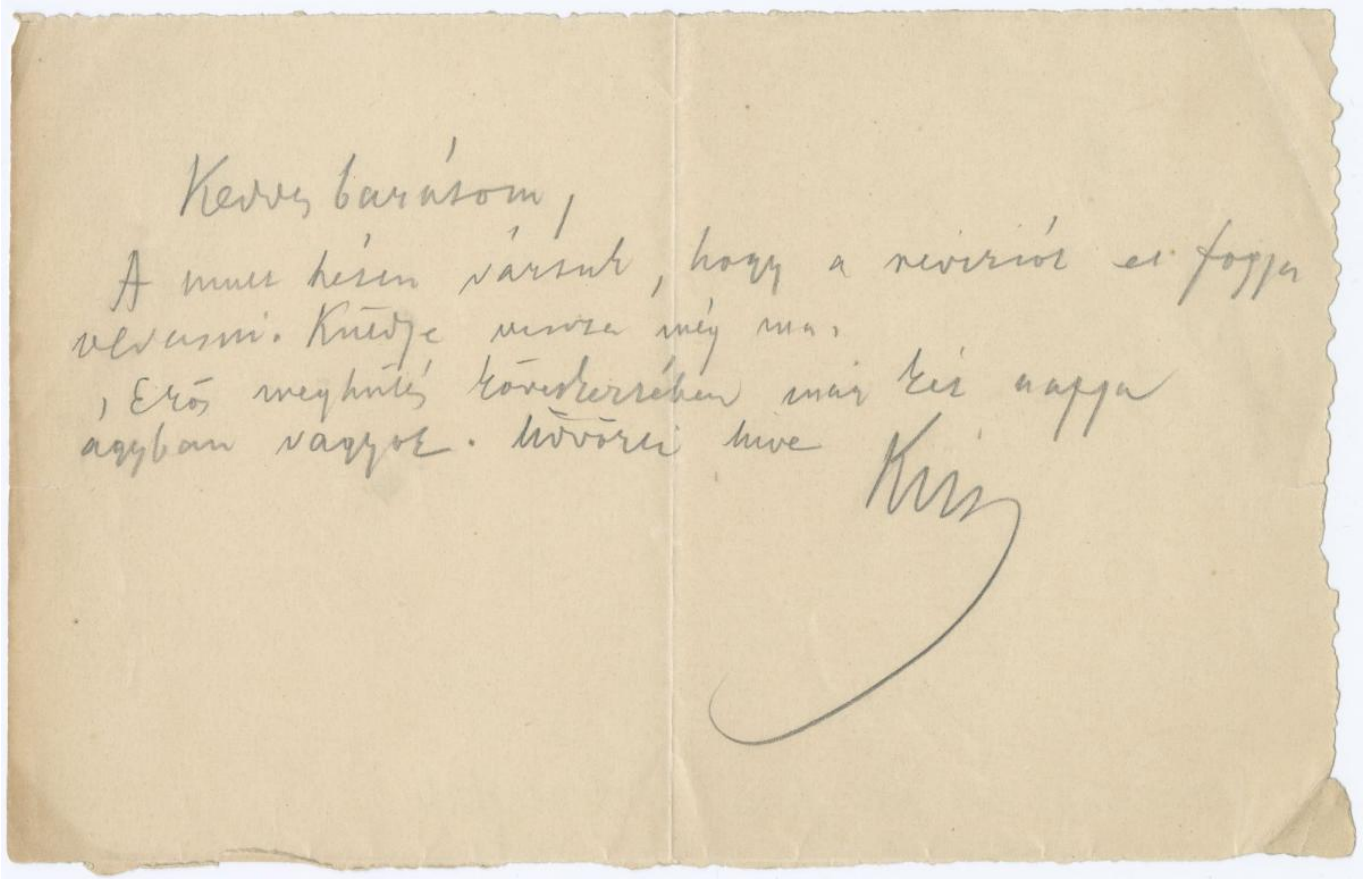
Gépi kézírás-felismerő modell készült az Országos Széchényi Könyvtárban

2022/07/07

1-1 # Kedves barátom,
2-1 # A mult héten ~~vartuk,~~ vártu
2-2 # ~~oevemi.~~ olvasni. Küldje vi
3-1 # ~~kös meghüles~~ Erős meghü
3-2 # ágyban vagyok.
4-1 # ~~Üdvözei~~ Üdvözli híve
5-1 # ~~Kissi~~ Kiss

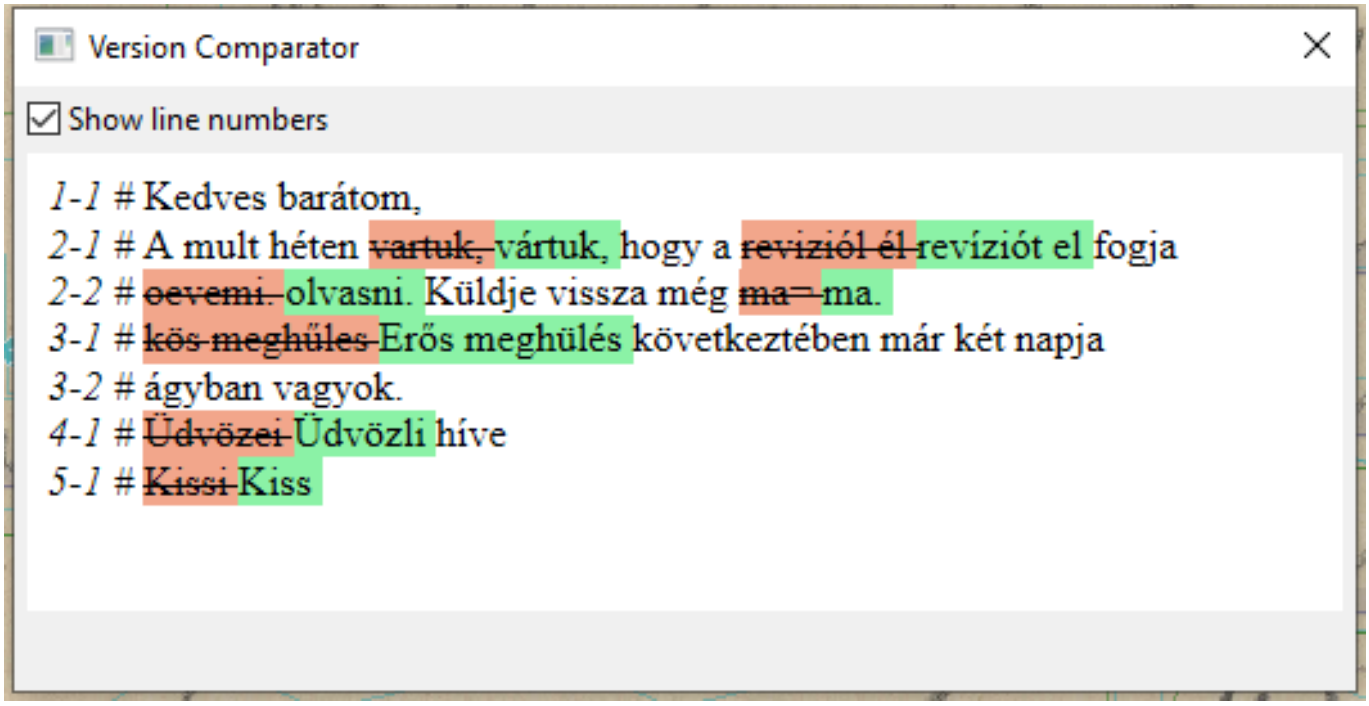
Megjelent az első publikus, magyar nyelvű számítógépes kézírás-felismerő modell. A modellt a Digitális Bölcsészeti Központ hozta létre.



[1]

Megjelent az első publikus, magyar nyelvű számítógépes kézírás-felismerő modell. A HTR- (Handwritten Text Recognition-) programot az Országos Széchényi Könyvtár Digitális Bölcsészeti Központja hozta létre, és mindenki által elérhetővé tette a Transkribus nevű szoftver komponenseként. Alapját Kiss Józsefnek, a 19–20. század fordulóján élő magyar költőnek, a Nyugat előfutáraként számontartott A Hét című hetilap szerkesztőjének a szakmai és személyes levelezése

képezi.



[2]A kézírás-felismerést mesterséges intelligencián alapuló algoritmus végzi. Ezt először be kell tanítani, hogy a különféle kézírásokat felismerje, majd a megadott minták alapján felépít egy modellt, amellyel képessé válik arra, hogy egy ismeretlen, általa még nem látott kézírásképet értelmezzen. Minél többféle anyagon tanítják, annál jobban teljesít az egyes szövegeken. A modell először Kiss József, majd a levelezőpartnerek vegyes kézírásán tanult. Az algoritmus csaknem kilencvenszázalékos pontossággal meg tudja állapítani, milyen karakterekből áll a szöveg.

Az eddig felhasznált kéziratok a Petőfi Irodalmi Múzeumban találhatók, a tanulásra felhasznált szavak száma körülbelül 75 ezer. A dokumentumok között szerepelnek borítékok, képeslapok, hagyományos és fejléces levelek, névjegykártyák is. A levélírók Kiss József és családja, illetve a századforduló írói, újságírói és művészei, mint Ady Endre, Móricz Zsigmond vagy Tömörkény István. Ez összesen 300 darab változó hosszúságú és minőségű levelet jelent, amelyeket a DBK folyamatosan közzétesz [3] az olvasók számára. A levélváltások további kéziratai jelenleg is feldolgozás alatt állnak a Petőfi Irodalmi Múzeumban, illetve az Országos Széchényi Könyvtárban.

Az új magyar kézírás-felismerő modell megtalálható a Transkribus [4] asztali alkalmazásban és a webes felületen is.
2022/07/07 - 10:54

Source URL: https://oszk.hu/en/node/5680

Links:
[1] https://oszk.hu/sites/default/files/Kiss-Jozsef-levele-Fenyo-Miksahoz_0.jpg
[2] https://oszk.hu/sites/default/files/A%20modell%20%C3%A1ltal%20ejtett%20hib%C3%A1k%20a%20138.%20oldalon.PNG
[3] http://dhupla.hu/collection/kiss-jozsef-levelezes
[4] https://readcoop.eu/model/hungarian/
[5] https://oszk.hu/en/category/foszotar-es-pozicionalo/hirek
[6] https://oszk.hu/en/category/foszotar-es-pozicionalo/hirek-cimlapon