|  |
| --- |
| **NASLOV: Prepoznavanje i generiranje govora za početnike u Scratchu** |

|  |  |  |  |
| --- | --- | --- | --- |
| **SCENARIJ UČENJA** | | | |
| ***Škola:*** | | ***Trajanje (minute):*** | 90 |
| ***Učitelj:*** |  | ***Dob učenika:*** | 13-14 |

|  |  |
| --- | --- |
| ***Motivacijsko pitanje:*** | Što je prepoznavanje govora i generacija? |

|  |
| --- |
| ***Teme:*** |
| * prepoznavanje i generiranje govora za početnike u Scratchu |
| ***Cilj:*** |
| * upoznavanje prepoznavanja i generiranja govora za početnike u Scratchu |
| ***Ishodi:*** |
| * razumijevanje prepoznavanja i generiranja govora uz pomoć jednostavnog primjera u Scratchu. |
| ***Oblici rada:***   * *rad u parovima, grupni rad* |
| ***Metode:***   * *prezentacija, razgovor, rasprava, interaktivna vježba* |

|  |
| --- |
| **ARTIKULACIJA** |
| **Tijek akcije (trajanje u minutama)** |
| **UVOD**  *Zatražite od učenika da pokušaju definirati prepoznavanje i generiranje govora.*  *Možete ih zamoliti da objasne razliku između prepoznavanja i generiranja govora.*  *Upitajte učnika poznaju li bilo koji sustav prepoznavanja govora. (Odgovor: Google NOW, SIRI, ...)*  **Uvod u prepoznavanje govora**  **Prepoznavanje govora** je sposobnost prevođenja diktata ili izgovorene riječi u tekst. Poznat je i kao govor-u-tekst (*Speech-To-Text STT*) i prepoznavanje glasa (*Voice Recognition*).  Prepoznavanje se postiže slijedom određenih radnj, a program koji vrši prepoznavanje govora poznat je kao "Sustav prepoznavanja govora" (*Speech Recognition system*). Ovi sustavi obično se implementiraju u obliku programa za diktiranje i inteligentnih asistenata u osobnim računalima, pametnim telefonima, web preglednicima i mnogim drugim uređajima.  **Uvod u generiranje govora**  **Generiranje** ili sinteza govora (skraćeno i kao TTS, što je kratica od engleskih riječi [Text-to-Speech](https://vivoka.com/text-to-speech/)), za razliku od prepoznavanja govora, nije tehnologija koja iskorištava glas, ona ga proizvodi. Sintetizirani govori općenito su završna faza globalnih procesa koji postaju sve demokratskiji zbog važnosti cjelokupnog iskustva vezanog uz pojam "glasa".  Sinteza govora (TTS) definira se kao umjetna proizvodnja ljudskih glasova. Glavna uporaba (i ono što je potaknulo stvaranje te sinteze) je sposobnost automatskog prevođenja teksta u izgovorenu riječ.  **Definiranje cilja lekcije:**  Uvod u prepoznavanje i generiranje govora za početnike kroz jednostavan primjer programa u Makeblock aplikaciji. |
|  |
| **GLAVNI DIO**  **Aplikacija**  **Scratch (MIT)** - <https://scratch.mit.edu/projects/editor/>  Dostupno je samo proširenje "*Text-to-Speech*" (3 bloka)  Snimka zaslona računala  Opis automatski generiran  **Scratch (ML4KIDS)** - <https://machinelearningforkids.co.uk/scratch3/>  Dostupna su proširenja pretvaranja teksta u govor (*Text-to-Speech,* 3 bloka) i pretvaranje govora u tekst (*Speech-to-Text*, 3 bloka)  Snimka zaslona računala  Opis automatski generiran  **Makeblock** - <https://ide.mblock.cc/>  Dostupna su proširenja za pretvaranje teksta u govor (*Text-to-Speech,* 3 bloka), i kognitivne usluge (*Cognitive Services,* 3 bloka za prepoznavanje govora)    *\* DODATAK* Makeblock u *Cognitive Services* proširenju ima i blokove za prepoznavanje jezika i analizu raspoloženja u tekstu    **Analiza** raspoloženja je proces otkrivanja pozitivnog ili negativnog raspoloženja u tekstu. Tvrtke ga često koriste za otkrivanje osjećaja u društvenim podacima, procjenu ugleda robnih marki i razumijevanje kupaca. Modeli analize sentimenta fokusiraju se na polaritet (*pozitivan, negativan, neutralan*), ali i na osjećaje i emocije (ljuti, sretni, tužni itd.), hitnost (hitni, manje hitni), pa čak i namjere (zainteresirani ili nisu zainteresirani).  Saznajte više o analizi sentimenta na poveznici: <https://monkeylearn.com/sentiment-analysis/>  Da bismo to ilustrirali, uzmimo primjer s Makeblockom.  Korak 1: Otvorite Makeblock sučelje na poveznici: <https://ide.mblock.cc/>  Korak 2: Dodajte proširenja: *Cognitive services* i *Text-to-Speech*  Korak 3: Označite za prikaz sljedeće blokove izvještajnog tipa (*reporter type block*)  **✓** *speech recognition result* **✓** *language recognition result* **✓** *sentiment analysis result*  Korak 4: Koristimo blok *speech recognition result* unutar *recognize language* i *analyse text sentiment* blokova  Korak 5: Programsko rješenje koje ćemo koristiti nalazi se na slici dolje:    Korak 6: Uključite mikrofon i zvučnike, pokrenite skriptu i recite "Sretan sam" (vidjet ćete skočni prozor koji snima vaš glas 2 sekunde - ako vam je potrebno više vremena, povećajte vrijednost u bloku recognize speech in English for 3 secs (ili više sekundi). Program će zatim izreći "Ovo je dobro" (*This is good*) ili "Ovo nije dobro" (*This is not good*) ovisno o rezultatu analize raspoloženja.      Korak 7: Ponovno pokrenite ovu skriptu, recite nešto drugo i pričekajte rezultat analize sentimenta. |
| **Osnovna načela prepoznavanja govora**  Najmanja jedinica govornog jezika poznata je kao **Fonem**. Engleski jezik sadrži oko 44 fonema koji predstavljaju sve samoglasnike i suglasnike koje koristimo za govor. Možemo uzeti primjer tipične riječi kao što je Mjesec (*Moon*) koja se može podijeliti na tri fonema: *m, ue, n*.  Da bismo protumačili govor, moramo imati način identificiranja komponenti izgovorenih riječi i identificiranja markera unutar govora. Algoritam se mora koristiti za daljnje tumačenje govora. Za tu svrhunajčešće se koristi matematički model pod nazivom *Hidden Markov* *Model*. Da biste stvorili modul za prepoznavanje govora, potrebno je stvoriti veliku bazu podataka s modelima koji odgovaraju svakom fonemu.  Saznajte više o ovome na poveznici: [*https://www.ibm.com/cloud/learn/speech-recognition*](https://www.ibm.com/cloud/learn/speech-recognition)  **Osnovna načela generiranja govora**  Za razliku od sustava za prepoznavanje govora koji koriste foneme (najmanje jedinice zvuka) za izdvajanje iz riječi i rečenica, generiranje govora (*Text-To-Speech*) će se temeljiti na onome što je poznato kao grafemi: slova i skupine slova koja se pripisuju odgovarajućim fonemima. To znači da ovdje osnovni resurs nije zvuk, već tekst. Taj proces se obično odvija u dva koraka.  Prvi će „izrezati“ tekst na rečenice, a zatim na riječi (naši poznati grafemi) i dodijeliti im pripadajuće foneme, odnosno izgovor za svaki grafem. Nakon što se identificiraju sve tekstualne, odnosno glasovne skupine, drugi korak sastoji se od pretvaranja tih jezičnih prikaza u zvuk. Drugim riječima, morate spojiti tonske zapise svih fonema kako biste dobili glas koji čita tekst.  Isprobajte generiranje govora iz teksta (TTS) na poveznici: [*https://www.readspeaker.com/*](https://www.readspeaker.com/)  Prepoznavanje govora je sposobnost prevođenja diktata ili izgovorene riječi u tekst. Poznat je i kao Tekst-u-Govor i Prepoznavanje govora (*Speech-to-Text* i *Voice recognition*). To se postiže pomoću određenih koraka, a sustav koji to radi poznat je i kao "Sustav prepoznavanja govora". Sustavi prepoznavanja govora obično se primjenjuju u obliku aplikacija za diktiranje i inteligentnih asistenata u osobnim računalima, pametnim telefonima, web preglednicima i mnogim drugim uređajima. |
| **ZAKLJUČAK** |
| **Prepoznavanje govora** je sposobnost prevođenja diktata ili izgovorene riječi u tekst. **Generiranje** govora je tehnologija koja stvara glas. |

|  |  |
| --- | --- |
| ***Metode*** | ***Oblici rada*** |
| ***prezentacija***  ***interaktivna vježba / simulacija na računalu*** | ***rad u parovima***  ***grupni rad*** |

|  |
| --- |
| ***Materijali:*** |
| * <https://scratch.mit.edu/projects/editor/> * <https://machinelearningforkids.co.uk/scratch3/> * <https://ide.mblock.cc/> * <https://monkeylearn.com/sentiment-analysis/> * <https://www.ibm.com/cloud/learn/speech-recognition> * <https://www.readspeaker.com/> |

|  |
| --- |
| ***Literatura:*** |

|  |
| --- |
| **OSOBNA ZAPAŽANJA, KOMENTARI I BILJEŠKE** |
|  |