|  |
| --- |
| **TEMAT:** Rozpoznawanie i generowanie mowy w Scratchu dla początkujących |

|  |  |  |  |
| --- | --- | --- | --- |
| **SCENARIUSZ LEKCJI** | | | |
| ***Szkoła:*** | | ***Czas (minuty):*** | 90 |
| ***Nauczyciel:*** |  | ***Wiek uczniów:*** | 13-14 |

|  |  |
| --- | --- |
| ***Zagadnienie główne:*** | Czym jest rozpoznawanie i generowanie mowy? |

|  |
| --- |
| ***Tematyka:*** |
| * Rozpoznawanie i generowanie mowy dla początkujących w Scratchu |
| ***Cele:*** |
| * Nauka o rozpoznawaniu i generowaniu mowy dla początkujących w Scratchu |
| ***Oczekiwane efekty:*** |
| * Zrozumienie rozpoznawania i generowania mowy za pomocą prostego przykładu w Scratchu. |
| ***Formy pracy:***   * praca w parach, praca w grupach   ***Metody:*** |
| * prezentacja, dyskusja, ćwiczenia interaktywne |

|  |
| --- |
| **TOK LEKCJI** |
| **Przebieg zajęć** |
| **WPROWADZENIE**  *Poproś uczniów, aby spróbowali zdefiniować rozpoznawanie i generowanie mowy.*  *Możesz poprosić ich o wyjaśnienie różnicy między rozpoznawaniem mowy a generowaniem mowy.*  *Zapytaj uczniów, czy znają jakieś systemy rozpoznawania mowy. (Odpowiedź: Google NOW, SIRI, …)*  **Wprowadzenie do Rozpoznawania Mowy**  **Rozpoznawanie Mowy** to umiejętność przekształcania wypowiedzi mówionej na tekst. Jest to również znane jako konwersja mowy na tekst (STT) oraz rozpoznawanie głosu.  Osiąga się to poprzez realizację określonych kroków, a oprogramowanie odpowiedzialne za to nazywane jest "Systemem Rozpoznawania Mowy" (SR). Systemy SR są zazwyczaj wdrażane w postaci oprogramowania do dyktowania i inteligentnych asystentów w komputerach osobistych, smartfonach, przeglądarkach internetowych i wielu innych urządzeniach.  **Wprowadzenie do Generowania Mowy**  **Generowanie mowy** lub syntezowanie mowy (również skrótowo TTS, Tekast-do-Mowy) w przeciwieństwie do rozpoznawania mowy, to nie technologia, która wykorzystuje głos, ale go produkuje. Sztuczne głosy są zazwyczaj ostatnim etapem procesu i stają się coraz bardziej powszechne, ponieważ są ważne dla ogólnego doświadczenia "głosu".  Synteza mowy (TTS) jest definiowana jako sztuczne generowanie ludzkich głosów. Głównym zastosowaniem (i tym, co skłoniło do jego stworzenia), jest umożliwienie automatycznego przekształcania tekstu na mowę.  **Przedstawienie celu głównego lekcji:**  Wprowadzenie do rozpoznawania i generowania mowy dla początkujących na przykładzie jednego prostego programu w Makeblock. |
|  |
| **CZĘŚĆ GŁÓWNA**  **Aplikacje**  **Scratch (MIT)** - <https://scratch.mit.edu/projects/editor/>  Dostępne jest tylko rozszerzenie Text-to-Speech (3 bloki)  A screenshot of a computer  Description automatically generated  **Scratch (ML4KIDS)** - <https://machinelearningforkids.co.uk/scratch3/>  Dostępne są rozszerzenia Text-to-Speech (3 bloki) i Speech-to-Text (3 bloki)  A screenshot of a computer  Description automatically generated  **Makeblock** - <https://ide.mblock.cc/>  Text-to-Speech (3 bloki), Cognitive Services (3 bloki związane z Speech-to-Text)    *\* BONUS*  Makeblock zawiera również bloki do rozpoznawania języka i analizy sentymentu tekstu.    **Analiza sentymentu** jest procesem wykrywania pozytywnego lub negatywnego nastroju w tekście. Jest często wykorzystywana przez firmy do wykrywania sentymentu w danych społecznościowych, oceny reputacji marki i zrozumienia potrzeb klientów. Modele analizy sentymentu skupiają się na polaryzacji (pozytywna, negatywna, neutralna), ale także na uczuciach i emocjach (złość, radość, smutek itp.), pilności (pilne, niepilne) i nawet intencjach (zainteresowany vs niezainteresowany).  Dowiedz się więcej o analizie sentymentu: <https://monkeylearn.com/sentiment-analysis/>  Aby to zobrazować, weźmy przykład z Makeblock.  Krok 1: Otwórz stronę Makeblock: https://ide.mblock.cc/  Krok 2: Dodaj rozszerzenia: Usługi poznawcze (Cognitive services) i Tekst do mowy (Text to Speech)  Krok 3: Sprawdź, czy wyświetlane są następujące bloki typu reporter:  ✓ Wynik rozpoznawania mowy (speech recognition result)  ✓ Wynik rozpoznawania języka (language recognition result)  ✓ Wynik analizy sentymentu (sentiment analysis result)  Krok 4: Wykorzystaj wynik rozpoznawania mowy do rozpoznawania języka i analizowania bloków sentymentu tekstu  Krok 5: Użyj tej sekwencji bloków    Krok 6: Włącz mikrofon i głośniki, uruchom kod i powiedz "Jestem szczęśliwy/a"  (pojawi się okno dialogowe nagrywające Twoją mowę przez 2 sekundy - jeśli potrzebujesz więcej, zwiększ wartość w bloku rozpoznawania mowy w języku angielskim na 3 sekundy (lub dłużej). Usłyszysz odpowiedź "To jest dobre" lub "To nie jest dobre", w zależności od wyników analizy sentymentu.    Krok 7: Uruchom kod ponownie, powiedz coś innego i poczekaj na wyniki analizy sentymentu.  **Podstawowe zasady rozpoznawania mowy**  Najmniejszą jednostką mówionego języka jest fonem. Język angielski zawiera około 44 fonemy, które reprezentują samogłoski i spółgłoski używane w mowie. Możemy przyjąć jako przykład typowe słowo takie jak "moon", które można podzielić na trzy fonemy: m, ue, n.  Aby interpretować mowę, musimy mieć sposób na identyfikację składników wypowiedzianych słów i fonemów, które pełnią rolę znaczników identyfikacyjnych w mowie. Do dalszej interpretacji mowy konieczne jest użycie odpowiedniego algorytmu. Ukryty Model Markowa (Hidden Markov Model) to powszechnie stosowany matematyczny model, który to umożliwia. Aby stworzyć silnik rozpoznawania mowy, tworzy się dużą bazę modeli, które odpowiadają każdemu fonemowi.  Dowiedz się więcej: <https://www.ibm.com/cloud/learn/speech-recognition>  **Podstawowe zasady generowania mowy**  W przeciwieństwie do systemów rozpoznawania mowy, które w pierwszej kolejności korzystają z fonemów (najmniejszych jednostek dźwiękowych) do wyodrębnienia zdań, TTS (tekst do mowy) będzie oparte na tym, co nazywamy grafemami: literami i grupami liter, które transkrybują fonem. Oznacza to, że podstawowym źródłem nie jest dźwięk, lecz tekst. Proces ten zazwyczaj odbywa się w dwóch etapach.  Pierwszy etap polega na podzieleniu tekstu na zdania i słowa (nasze znane grafemy) oraz przypisaniu transkrypcji fonetycznej, czyli wymowy, do tych grup. Gdy różne grupy tekstowe/fonetyczne zostaną zidentyfikowane, drugi etap polega na przekształceniu tych reprezentacji lingwistycznych w dźwięk. Innymi słowy, odczytanie tych wskazówek w celu wygenerowania głosu, który przeczyta informacje.  Wypróbuj TTS online: <https://www.readspeaker.com/>  Rozpoznawanie mowy to zdolność przekształcania mówionych słów na tekst. Jest również znane jako przekształcanie mowy na tekst (Speech-to-Text) oraz rozpoznawanie głosu. Osiąga się to poprzez przestrzeganie określonych kroków, a oprogramowanie odpowiedzialne za to nazywane jest "Systemem Rozpoznawania Mowy". Systemy rozpoznawania mowy są zazwyczaj wdrażane w postaci oprogramowania do dyktowania oraz inteligentnych asystentów w komputerach osobistych, smartfonach, przeglądarkach internetowych i wielu innych urządzeniach.  Opracuj tabelę K.W.L. (Know, Want, Learned) ze swoimi uczniami.   |  |  |  | | --- | --- | --- | | Co wiem | Co chcę wiedzieć | Czego się nauczyłem/nauczyłam | |  |  |  | |
|  |
| **PODSUMOWANIE**  Rozpoznawanie mowy to zdolność przekładania wypowiedzi ustnych na tekst. Generowanie mowy to technologia, która tworzy głos. |
|  |

|  |  |
| --- | --- |
| ***Metody*** | ***Formy pracy*** |
| ***prezentacja***  ***ćwiczenie interaktywne/symulacja komputerowa*** | ***praca w parach***  ***praca grupowa*** |

|  |
| --- |
| ***Materiały:*** |
| * <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*** |

|  |
| --- |
| **OBSERWACJE, UWAGI, NOTATKI** |
|  |