| **TÍTULO:** Speech recognition and generation for beginners in Scratch |
| --- |

| **CENÁRIO DE APRENDIZAGEM** | | | |
| --- | --- | --- | --- |
| ***Escola:*** | | ***Duração (minutos):*** | 90 |
| ***Professor:*** |  | ***Alunos***  ***idade:*** | 13-14 |

| ***Questão Essencial*** | What is Speech recognition and generation? |
| --- | --- |

| ***Tópicos:*** |
| --- |
| * Reconhecimento e geração de voz para principiantes em Scratch |
| ***objetivos:*** |
| * Reconhecer a geração de voz para principiantes no Scratch |
| ***Outcomes:*** |
| * Compreender o reconhecimento e a geração de fala com a ajuda de um exemplo simples no Scratch. |
| ***Formas de trabalho:***   * trabalho individual, trabalho a pares, trabalho de grupo   ***Métodos:*** |
| * apresentação, palestra, debate, exercício interativo |

| **ARTICULAÇÃO** |
| --- |
| O curso (duração, minutos) |
| **INTRODUÇÃO**  *Peça aos alunos para tentarem definir reconhecimento e geração de fala.*  *Pode pedir-lhes que expliquem a diferença entre reconhecimento de fala e geração de fala.*  *Pergunte aos alunos se conhecem algum sistema de reconhecimento de fala. (Resposta: Google NOW, SIRI, ...)*  *Introdução ao reconhecimento de fala*  *O reconhecimento da fala é a capacidade de traduzir um ditado ou uma palavra falada em texto. É também conhecido por Speech-to-Text (STT) e Voice Recognition (reconhecimento de voz).*  *É conseguido através de determinados passos e o software responsável por ele é conhecido como "Sistema de Reconhecimento de Fala". Os sistemas de SR são geralmente implementados sob a forma de software de ditado e assistentes inteligentes em computadores pessoais, smartphones, navegadores Web e muitos outros dispositivos.*  *Introdução à geração de fala*  *A geração ou síntese de fala (também abreviada como TTS, Text-to-Speech), ao contrário do reconhecimento de fala, não é uma tecnologia que explora a voz, mas sim que a produz. As vozes sintéticas são geralmente a fase final do processo e estão a tornar-se cada vez mais democráticas porque são importantes na experiência global da "voz".*  *A síntese da fala (TTS) é definida como a produção artificial de vozes humanas. A principal utilização (e o que induziu a sua criação) é a capacidade de traduzir automaticamente um texto em discurso falado.*  *Definir o objetivo da aula:*  *Introdução ao reconhecimento e geração de fala para principiantes, usando um exemplo de um programa simples em Makeblock.* |
|  |
| **PARTE PRINCIPAL**  **Apliações**  **Scratch (MIT)** - <https://scratch.mit.edu/projects/editor/>  Apenas a extensão Text-to-Speech está disponível (3 blocos)  A screenshot of a computer  Description automatically generated  **Scratch (ML4KIDS)** - <https://machinelearningforkids.co.uk/scratch3/>  Estão disponíveis extensões de conversão de texto em voz (3 blocos) e de conversão de texto em voz (3 blocos)  A screenshot of a computer  Description automatically generated  **Makeblock** - <https://ide.mblock.cc/>  Conversão de texto em fala (3 blocos), serviços cognitivos (3 blocos relacionados com a conversão de texto em fala)    *\* BONUS* O Makeblock também inclui blocos para reconhecimento de linguagem e análise de sentimentos de texto  A análise de sentimentos é o processo de deteção de sentimentos positivos ou negativos no texto. É frequentemente utilizada pelas empresas para detetar sentimentos em dados sociais, avaliar a reputação da marca e compreender os clientes. Os modelos de análise de sentimentos centram-se na polaridade (positivo, negativo, neutro), mas também nos sentimentos e emoções (zangado, feliz, triste, etc.), na urgência (urgente, não urgente) e até nas intenções (interessado vs. não interessado).  Saiba mais sobre a análise de sentimentos: https://monkeylearn.com/sentiment-analysis/  Para ilustrar isto, vamos dar um exemplo com o Makeblock.  Passo 1: Abrir a página do Makeblock: https://ide.mblock.cc/  Passo 2: Adicionar extensões: Serviços cognitivos e Texto para discurso  Passo 3: Marque para apresentar os seguintes blocos de tipo de repórter:  ✓ resultado do reconhecimento de fala  ✓ resultado do reconhecimento da língua  ✓ resultado da análise de sentimentos  Step 4: Use speech recognition result in recognize language and analyse text sentiment blocks  Step 5: Use this sequence of blocks    Passo 6: Ligue o microfone e os altifalantes, inicie o código e diga "I am happy"  (verá a janela pop-up a gravar a sua voz durante 2 segundos - se precisar de mais, aumente o valor em reconhecer discurso em inglês para 3 segundos (ou mais). Irá ouvir "Isto é bom" ou "Isto não é bom", dependendo dos resultados da análise de sentimentos.    Passo 7: Recomeçar o código, dizer outra coisa e esperar pelos resultados da análise de sentimentos . |
|  |
| **CONCLUSÃO**  Princípios básicos do reconhecimento de fala  A unidade mais pequena da linguagem falada é conhecida como fonema. A língua inglesa contém aproximadamente 44 fonemas que representam todas as vogais e consoantes que utilizamos na fala. Podemos dar o exemplo de uma palavra típica como moon, que pode ser dividida em três fonemas: m, ue, n.  Para interpretar o discurso, temos de ter uma forma de identificar os componentes das palavras faladas e os fonemas que actuam como marcadores de identificação dentro de um discurso. Deve ser utilizado um algoritmo para interpretar o discurso. O modelo oculto de Markov é um modelo matemático utilizado para o fazer. Para criar um motor de reconhecimento de fala, é criada uma grande base de dados de modelos para corresponder a cada fonema.  Saiba mais: https://www.ibm.com/cloud/learn/speech-recognition  Princípios básicos da geração de fala  Ao contrário dos sistemas de reconhecimento de voz que utilizam fonemas (as unidades mais pequenas de som) para cortar frases, o TTS baseia-se nos chamados grafemas: as letras e grupos de letras que transcrevem um fonema. Isto significa que o recurso básico não é o som, mas o texto. Isto é normalmente feito em duas etapas.  Na primeira, o texto é dividido em frases e palavras (os nossos famosos grafemas) e são atribuídas transcrições fonéticas, a pronúncia, a todos estes grupos. Uma vez identificados os diferentes grupos de texto/fonética, a segunda etapa consiste em converter estas representações linguísticas em som. Por outras palavras, ler estas indicações para produzir uma voz que leia a informação.  Experimente o TTS em linha: https://www.readspeaker.com/  O reconhecimento da fala é a capacidade de traduzir um ditado ou uma palavra falada em texto. Também é conhecido como Speech-to-Text e reconhecimento de voz. É conseguido seguindo certos passos e o software responsável por ele é conhecido como "Sistema de reconhecimento de fala". Os sistemas de reconhecimento de voz são normalmente implementados sob a forma de software de ditado e assistentes inteligentes em computadores pessoais, smartphones, navegadores Web e muitos outros dispositivos.  Faça a tabela K.W.L. (Know, Want, Learned) com os seus alunos.   | O que eu sei | O que eu quero saber | O que aprendi | | --- | --- | --- | |  |  |  | |
|  |

| ***Métodos*** | ***Formas de trabalho*** |
| --- | --- |
| ***apresentação***  ***exercício interativo/simulação no computador*** | ***trabalho individual***  ***trabalho em pares***  ***trabalho de grupo*** |

| ***Material:*** |
| --- |
| * <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*** |
| --- |

| * OBSERVAÇÕES PESSOAIS, COMENTÁRIOS E NOTAS |
| --- |
|  |