# SAINT Curriculum

## UNIT 3: Application of AI in Speech & vision

Deliverable: WP2/2.2

SAINT

HANDS ON INTRODUCTION TO ARTIFICIAL INTELLIGENCE IN PRIMARY EDUCATION USING MINECRAFT

APRIL 2023

EKVASIS
Authored by: EKVASIS Team
Project Number: 2022-1-FR01-KA220-SCH-000087794

Co-funded by the European Union

# REVISION HISTORY

| Version | Date | Author | Description | Action | Pages |
|---|---|---|---|---|---|
| 1.0 | 17/01/2023 | TARAN | Creation | C | TBS |
| | | | | | |
| | | | | | |

(*) Action: C = Creation, I = Insert, U = Update, R = Replace, D = Delete

# REFERENCED DOCUMENTS

| ID | Reference | | Title |
|---|---|---|---|
| 1 | 2022-1-FR01-KA220-SCH-000087794 | | SAINT Proposal |
| 2 | | | |

# APPLICABLE DOCUMENTS

| ID | Reference | | Title |
|---|---|---|---|
| 1 | | | |
| 2 | | | |

# Contents

1 Introduction of the project ..... 4
1.1 The scope of the project ..... 4
1.2 The target groups ..... 4
1.3 The purpose of this document ..... 5
2 Glossary of the Unit ..... 5
3 Introduction of the Unit ..... 6
3.1 Description ..... 6
3.2 Learning objectives & outcomes ..... 6
3.3 Estimated seat time ..... 7
4 Course content of the Unit ..... 7
4.1 Introduction ..... 7
4.2 Idea 1: Perception ..... 7
4.3 Idea 2: Representation & reasoning ..... 9
4.4 Idea 3: Learning ..... 9
4.5 Idea 4: Natural interaction ..... 11
4.6 Idea 5: Societal impact ..... 12
4.7 Case Studies & Success Stories ..... 13
5 Additional materials and resources ..... 14
6 Wrap-up ..... 16
7 Quiz ..... 16
8 References ..... 18

# 1 Introduction of the project

## 1.1 The scope of the project

Working as an ideal digital learning environment to teach children about the practical applications of AI based on theAI4K12 project guidelines, the motivation for this project comprises the following goals:

- Introduce pupils, teachers and educators to AI concepts, its impacts on our society and related practical implementations,
- Address the growing need to develop remote learning solutions facilitating student engagement, creativity, problem-solving and decision-making skills,
- Upskill the teachers and educators with new sets of skills (PBL, AI, gamification etc) developed through innovative ways of teaching,
- Improve engagement rates in children through the use of an innovative way of teaching, helping children develop creativity,
- Reduce the gap between need and availability for AI related skills.

AI Adventures in Minecraft teaches AI related skills to children aged from 9-12 years old, using a Minecraft World. With this, we create a fun, interactive and creative learning environment through specific activities and challenges aligned with the AI4K12 guidelines (ai4ka12.org) and the 5 big ideas of AI: 1) Perception, 2) Representation & Reasoning, 3) Learning, 4) Natural interactions, 5) Societal impact.

To that end, the project develops and promotes the following tangible results:

- This curriculum: a complete learning course for introducing AI in school teaching based on the 5 big ideas of the AI4K12 framework. The course disseminates knowledge about AI4K12's AI education guidelines and the 5 big ideas, explore the impact of AI in our society and enhance understanding of relevant concepts.
- A tailored Minecraft world (AI Adventures World) delivering educational challenges based on the learning course. It makes use of the escape room concept and offer Problem Based Learning activities. One challenge for each unit or lesson.
- The foundry virtual space supporting a growing community of adopters of SAINT and guiding the corrective/perfective and evolutive maintenance of the training package.

## 1.2 The target groups

The project sees the direct involvement of teachers, mainly teachers of children aged from 9-12 years old or Higher Education staff involved in the teaching of educators. These teachers are either teachers of STEM subjects or have some knowledge and interest in AI and/or Minecraft.

Concerning the Indirect target groups identified, the following can be involved:

- STEM centres looking to develop their catalogue of innovative teaching technologies or their catalogue of product enhancing AI knowledge,
- Higher education institutions collaborating with companies / public authorities engaged in the creation of educational material,
- Organisation, associations or networks looking to provide parents and or educators with educational material on AI: such as coding clubs, adult learning centres, entrepreneurial coaching services, continuing education centres, etc.

## 1.3 The purpose of this document

The work package n°2 - AI4K12 Educational Programme focuses on producing a complete course on AI with a set of 5 challenges in the related Minecraft World to illustrate the practical implementation of the technology.

This AI Curriculum is composed of a total of 5 units of pedagogical material based on the AI4K12 education guidelines and the learning objectives put into light following national surveys:
1. Application of AI in Machine learning,
2. Application of AI in Work and Entrepreneurship,
3. Application of AI in Speech & vision,
4. Application of AI in Games & puzzles,
5. Application of AI in everyday life.

Additionally, a glossary is created in each Unit in order to ease the adoption of the SAINT package by the teachers and schools.

# 2 Glossary of the Unit

| Words | Definition |
|---|---|
| **Machine Learning** | A subset of AI that uses algorithms and statistical models to enable computer systems to improve their performance on a specific task based on data input, without being explicitly programmed. |
| **Speech Recognition** | The ability of a machine or program to identify and understand human speech, often used in applications such as virtual assistants, automated customer service, and transcription. |
| **Computer Vision** | A field of study focused on enabling computers to interpret and understand visual data from the world around them, often used in applications such as object recognition, facial recognition, and autonomous vehicles. |
| **Deep Learning** | A subset of machine learning that uses artificial neural networks to model complex patterns in data, often used in applications such as image and speech recognition |
| **Natural Language Processing (NLP)** | A subset of AI that focuses on enabling computers to understand, interpret, and generate human language, often used in applications |

| | |
|---|---|
| | such as chatbots, language translation, and sentiment analysis. |
| **Reinforcement Learning** | A type of machine learning in which an agent learns to make decisions by trial and error, receiving rewards or punishments based on its actions. |
| **Augmented Reality** | An interactive experience in which digital information is overlaid on the physical world, often used in applications such as gaming, education, and marketing. |
| **Virtual Reality** | A computer-generated environment that simulates a user's physical presence and allows for interaction, often used in applications such as gaming, training, and therapy. |

# 3 Introduction of the Unit

## 3.1 Description

This Unit will introduce learners to the exciting field of AI and its application in speech and vision. It will explore the different types of machine learning, including supervised, unsupervised, and reinforcementlearning, and how they are used in speech recognition, image classification, and other real-world applications. The Unit will also delve into natural interaction, such as how AI interacts with humans andhow virtual and augmented reality are changing the way we interact with technology. Additionally, the Unit will address the societal impact of AI, including ethical considerations and the pros and cons of AItechnology. By the end of the Unit, learners should have a fundamental understanding of AI in speechand vision, as well as its potential for solving real-world problems.

## 3.2 Learning objectives & outcomes

In this unit, learners will become acquainted with the application of AI in speech and vision. They will learn about the fundamentals of machine learning, including supervised, unsupervised, and reinforcement learning. The unit will also cover natural interaction and its relevance in AI, as well as thesocietal impact and ethical considerations associated with the use of AI.

On successful completion of this Unit, learners:

- should be able to identify the different types of machine learning and their applications in speechand vision.
- will be able to explain the importance of natural interaction and its role in the development of AIsystems.
- will have a solid understanding of the societal impact of AI and the ethical considerations thatmust be taken into account in the development and deployment of these systems.
- will be able to evaluate real-world applications of AI in speech and vision and recognize theirpotential for solving complex problems.

## 3.3 Estimated seat time

The completion of the module along with the implementation of the knowledge provided will last around5-6 hours, based on the amount of content and quizzes provided in the module. However, the actual time for learners to complete the module and apply the acquired knowledge may vary, depending on their learning pace and familiarity with the subject.

# 4 Course content of the Unit

## 4.1 Introduction

Welcome to the unit on the application of AI in Speech and Vision. In this unit, we will explore how AI is being used to transform the fields of speech recognition and computer vision. We will begin by discussing the basics of machine learning and the various techniques used in Speech and Vision applications, including supervised, unsupervised, and reinforcement learning. Next, we will delve into natural interaction and how AI is changing the way we interact with machines, including speech and vision interaction. We will also examine the societal impact of AI in Speech and Vision applications, including its pros and cons and ethical considerations. Finally, we will highlight some of the most exciting case studies and success stories in the field, showcasing how AI is being used to solve real-world problems in healthcare, education, and entertainment. By the end of this unit, you will have a solid understanding of how AI is transforming the fields of Speech and Vision and the potential impact it can have on society as a whole.

## 4.2 Idea 1: Perception

Artificial Intelligence (AI) is a rapidly evolving field that is transforming our lives in many ways. From speech recognition to image analysis, AI has become an integral part of our everyday lives. However, teaching AI concepts to children can be challenging, especially given the complex nature of the technology. To address this challenge, the AI4K12 framework has been developed as a set of guidelines for teaching AI concepts and practical applications to K-12 students. The framework is based on five big ideas, the first of which is Perception. In this unit, we will explore the first big idea of the AI4K12 framework, Perception, which involves understanding how AI perceives the world around us. The importance of Perception in AI cannot be overstated, as it is what allows machines to interpret and respond to the world around them in a way that is similar to how humans do. To perceive the world around them, AI systems rely on a range of sensors, each designed to collect data from different sources. These sensors play a critical role in enabling machines to understand theirsurroundings and make decisions based on that information. In this section, we will introduce differenttypes of sensors used in AI and how they work.

**Examples**:

**Microphones:** One of the most common types of sensors used in AI is the microphone, which is used

for speech recognition. Microphones collect sound waves and convert them into digital signals that can be processed by AI algorithms. Voice assistants like Amazon's Alexa and Apple's Siri rely on microphonesto understand and respond to voice commands.

**Cameras:** Another critical type of sensor used in AI is the camera, which is used for computer vision. Cameras capture images or video frames and use algorithms to extract information from them. This information can include object detection, facial recognition, and motion tracking. Cameras are used in many real- world applications, including surveillance systems, autonomous vehicles, and security cameras.

**Infrared Sensors:** Infrared sensors detect infrared radiation, which is emitted by all objects that have a temperature aboveabsolute zero. These sensors are used in applications like home security systems and night vision cameras. They detect the heat signatures of objects, allowing AI systems to detect and track movement in the dark.

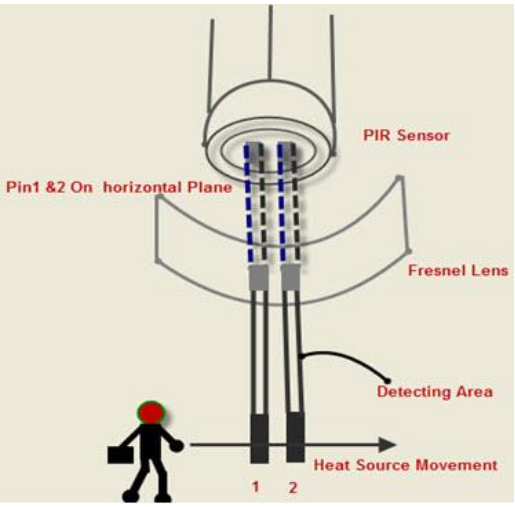


*Figure 1: Infrared Sensor*

**Touch Sensors:** Touch sensors are used in applications like touchscreens, robotic grippers, and prosthetics. Touch sensors can detect pressure, temperature, and other physical properties, allowing AI systems to interact with the physical world.

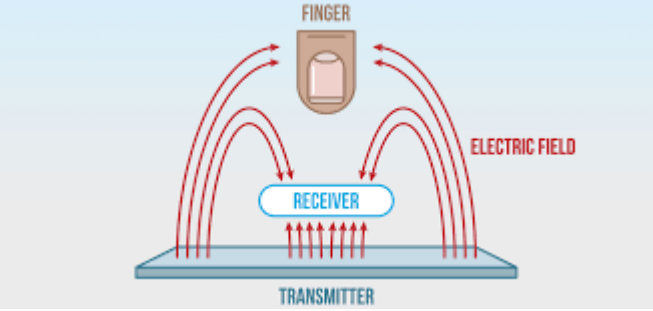


*Figure 2: Touch Sensor*

## 4.3 Idea 2: Representation & reasoning

Representation refers to how data is stored and processed by AI systems. This representation allows the AI system to analyze and extract featuresof the sound, such as phonemes or words. In AI, data is represented in a way that machines can understand and process efficiently. The most common way data is represented in AI is through symbols or numbers. For example, a picture of a catcan be represented as a matrix of numbers that correspond to the color of each pixel in the image.
Reasoning is the process of drawing conclusions from the data that has been collected and represented. AI systems use different reasoning techniques to process and analyze data, such as deductive reasoning, inductive reasoning, and abductive reasoning. These techniques help machines to make predictions, recognize patterns, and solve problems.

**Examples:**

**Speech recognition:** AI systems represent spoken language as a sequence of symbols or numbers, and then use reasoning techniques to interpret and understand the meaning of the speech. AI system might use reasoning to match the audio data to a database of known words or phrases, or to identify the intent ofa spoken command.

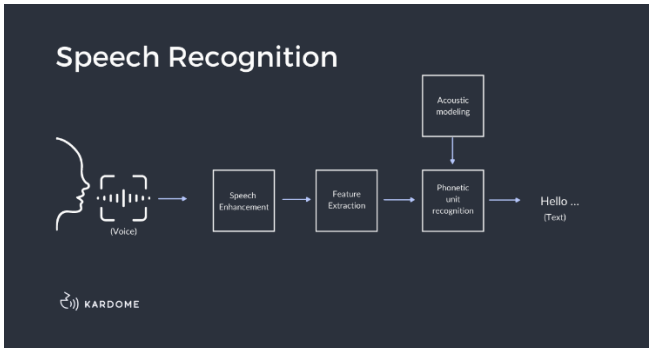


*Figure 3: Speech Recognition*

**Computer vision:** AI systems represent images as matrices of numbers, and then use reasoning techniques to recognizeobjects and patterns in the images. reasoning might involve using image recognition algorithms to identify objects in a scene or to track the movement of objects over time.

**Medical field**: AI can analyze large amounts of patient data to identify patterns and predict potential health issues. This can help doctors make more informed decisions about patient care.

**Finances:** AI can analyze market trends and data to identify patterns and predict future market movements. This can help investors make more informed decisions about investments and manage risk.

## 4.4 Idea 3: Learning

Learning is a critical aspect of AI and is what sets it apart from traditional computer programming. Machine learning algorithms enable AI systems to learn and improve from experience without being

explicitly programmed. In the context of Speech and Vision applications, learning plays a crucial role in enabling machines to recognize patterns and make predictions based on data. The importance of Learning in Speech and Vision applications cannot be overstated. By learning from large amounts of data, machines can recognize complex patterns and make more accurate predictions, leading to improved performance in real-world applications such as speech recognition, object detection, and autonomous driving.

The three main types of machine learning include:

**Supervised learning:** Supervised learning is used for tasks like speech and image recognition. Supervised learning is a type of machine learning in which the algorithm is trained on a labelled dataset. The labelled dataset consists of input data (e.g., images or audio files) along with corresponding output labels (e.g., text or image classification). The goal of the algorithm is to learn a mapping function from the input data to the output labels. In speech recognition, supervised learning is used to train the algorithm to recognize spoken words and convert them into text. The input data consists of audio files, and the output labels are the corresponding transcriptions. The algorithm is trained on a large dataset of labelled audio files and transcriptions, and it learns to recognize the patterns and relationships between the audio data and the corresponding text. Some common supervised learning algorithms used in speech and vision applications are convolutional neural networks (CNNs) and recurrent neural networks (RNNs). CNNs are particularly well-suited for image recognition tasks, as they can automatically learn hierarchical representations of the visual features. RNNs are particularly well-suited for speech recognition tasks, as they can model the temporal dependencies between the audio data and the corresponding text.

**Unsupervised learning:** Unsupervised learning is a type of machine learning that involves learning patterns and relationships within data without the need for explicit labels or supervision. Unsupervised learning is particularly useful in computer vision applications where large amounts of unstructured visual data need to be analyzed to identify meaningful patterns and relationships. For example, clustering algorithms can be used to group together similar images or objects in a dataset based on their visual features, without the need for explicit labelling. Another example of an unsupervised learning algorithm is the autoencoder. An autoencoder is a neural network that is trained to reconstruct input data from a lower- dimensional representation of that data. The encoder part of the network reduces the input data to a lower-dimensional latent space, and the decoder part of the network attempts to reconstruct the original input from this latent representation. Autoencoders can be used for tasks such as image denoising, dimensionality reduction, and feature learning. Another example of an unsupervised learning algorithm is the generative adversarial network (GAN).

**Reinforcement learning**: In Speech and Vision applications, reinforcement learning is used in areas such as robotics, autonomous driving, and game playing. For example, reinforcement learning can be used to teach a robot to navigate a cluttered environment or to control the movements of a self-driving car. In game playing, reinforcement learning can be used to develop agents that can play games like chess or Go at a superhuman level. One of the most popular reinforcement learning algorithms is Q-learning, which uses a Q-table to store the expected reward for each action in each state. The agent updates the Q-table based on the rewards it receives and uses the updated table to choose actions in the future.

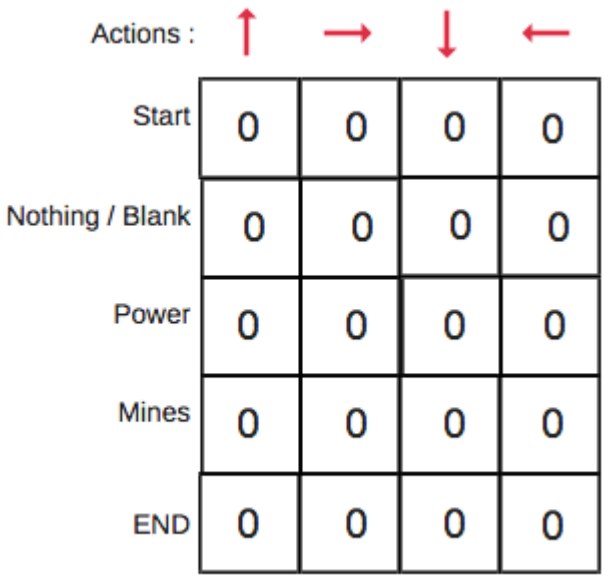


*Figure 4: Q table*

## 4.5 Idea 4: Natural interaction

Natural interaction in AI and robots is the ability to understand and respond to human natural language and actions, such as voice and gestures. This is particularly relevant in speech and vision applications, where the goal is to create interfaces that can understand human language and visual cues and respond in a way that is understandable, intuitive, and useful. Speech and Vision Interaction technologies, such as speech recognition, gesture recognition, and facialrecognition, are transforming the way we interact with machines and opening new possibilities for applications in gaming, robotics, security, and more.

**Speech Interaction:** Speech interaction is the use of natural language to communicate between humans and AI systems, and it includes speech recognition and speech synthesis. Speech recognition converts spoken language into text using algorithms, while speech synthesis generates artificial speech from text usingTTS systems. One of the primary applications of speech interactionin AI is virtual assistants, such as Apple's Siri, Amazon's Alexa, and Google Assistant. These virtual assistants use speech recognition and NLP to understand user requests and respond with relevant information or actions. Speech interaction is also used in call centers, where automated voice systemscan route calls and answer basic questions, freeing up human agents for more complex tasks.

**Vision Interaction:** Vision interaction refers to the ways in which humans can interact with AI systems using visual cues, gestures, and other non-verbal communication methods. This includes technologies such as facial recognition, gaze tracking, and motion capture. These technologies have many practical applications, such as controlling deviceswith hand gestures, tracking eye movements, and recognizing facial expressions and emotions.

**Virtual and Augmented Reality:** Virtual and Augmented Reality (VR and AR) are technologies used in AI to create immersive interactionsbetween humans and machines. VR creates a fully virtual environment using headsets and controllerswhile AR overlays virtual objects onto the user's view of the real world. These technologies have limitations, such as requiring specialized hardware and the potential for

motion sickness, but they are evolving and have numerous applications in fields such as education, healthcare, and entertainment.

*Figure 5:Virtual Reality*


*Figure 6:Augmented Reality*

## 4.6 Idea 5: Societal impact

The societal impact of AI in speech and vision applications has become a crucial topic of discussion in recent years. While these technologies have the potential to bring about significant advancements in various fields, they also come with ethical, social, and economic implications that need to be addressed.

**Pros of AI in Speech and Vision Applications**: AI in speech and vision has pros and cons for society. On the positive side, it has improved healthcare,education, and public safety. For example, speech recognition helps doctors transcribe medical recordsaccurately and assistive technologies aid people with disabilities. Vision applications contribute to facialrecognition and self-driving vehicles. However, ethical considerations must be considered for responsible development and implementation of AI technology.

**Cons of AI in Speech and Vision Applications**: AI in speech and vision has potential drawbacks. One major concern is the potential for bias and discrimination, as AI algorithms can replicate societal prejudices. There is also a risk of job displacement and economic disparities due to the high cost of AI systems. Privacy and surveillance are ethical concerns, with AI collecting personal data. Additionally, there is a broader concern about the long-term impact of AI, as it becomes more advanced and autonomous. It's essential to address these concerns and create more transparent, equitable, and accountable systems that benefit society.

**Ethical Considerations**: While AI technologies have the potential to bring about significant benefits, there are also several ethicalconsiderations that must be considered. Some of the key ethical issues associated with AI in Speech and Vision applications include privacy, transparency, bias, and accountability.

- **Privacy**: The use of AI in Speech and Vision applications raises concerns about privacy. With the increasing amount of data being collected, there is a risk that personal information could be misused ormishandled. As such, it is important for developers and organizations to ensure that appropriate privacyprotections are in place to safeguard personal data.

- **Transparency**: Another ethical concern related to AI in Speech and Vision applications is the issue of transparency. It can be difficult to understand how an AI system arrives at its decisions, which can makeit challenging to assess the fairness and accuracy of its outputs. To address this concern, it is importantfor developers to prioritize transparency in their AI systems and make sure that they can be audited andexplained.

- **Bias**: Bias is another ethical issue that can arise in AI in Speech and Vision applications. If the data used to train an AI system is biased, then the system's outputs may also be biased. This can have significant implications in areas such as hiring, lending, and criminal justice. To address this issue, developers must take steps to ensure that their data sets are representative and unbiased.

- **Accountability**: Finally, the use of AI in Speech and Vision applications raises questions aboutaccountability. Who is responsible if an AI system makes a mistake or causes harm? As AI becomes more integrated into our lives, it is important to ensure that there are clear lines of accountability and responsibility in place.

## 4.7 Case Studies & Success Stories

Case studies showcase practical AI applications in Speech and Vision. Google's DeepMind AI for Lip-Reading improves speech recognition for people with hearing difficulties, while IBM's Watson for Oncology assists oncologists in creating personalized cancer treatment plans. Amazon's Alexa for Speech Recognition has revolutionized the way people interact with technology. Other notable AI applications in Speech and Vision include Apple's Face ID and Tesla's Autopilot. These examples demonstrate the transformative potential of AI in enhancing our lives.

### 4.7.1 Example 1: Google’s DeepMind AI for Lip-Reading

Google's DeepMind created an AI system trained on 5,000 hours of TV footage that achieved an accuracy rate of 46.8% in lip-reading, surpassing human experts' 12.4%. It has potential applications in speech recognition, security, and aiding individuals with hearing impairments. Lip-reading is challenging, and even the best human lip-readers can only understand about 30% of spoken words accurately. DeepMind's AI has the potential to revolutionize communication for people with hearing impairments. There are privacy concerns related to the collection and analysis of large amounts of video footage, and the potential for misuse in surveillance and law enforcement contexts. Overall, this technology demonstrates how AI can improve communication for individuals with hearing impairments, and highlights the need for ethical considerations in AI development and use.

### 4.7.2 Example 2: IBM's Watson for Oncology

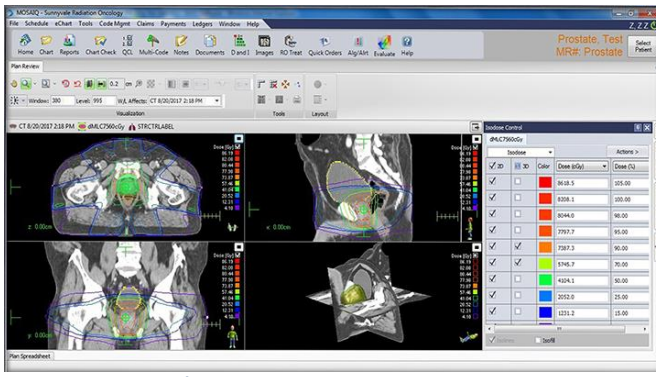

*Figure 7: IBM Watson*

IBM's Watson for Oncology is an AI system that assists oncologists in making personalized treatment decisions for cancer patients. The system uses natural language processing and machine learning to analyze vast amounts of medical data and provide treatment recommendations based on a patient's unique circumstances. Although there have been some challenges and criticisms of the system, it has been used successfully in healthcare settings and represents a promising example of how AI can improve patient outcomes.

### 4.7.3 Example 3: Amazon's Alexa for Speech Recognition

Amazon's Alexa is a popular virtual assistant that uses speech recognition technology to enable hands-free operation of various devices. It recognizes and responds to natural language commands and uses machine learning to provide personalized recommendations. Alexa has also helped people with disabilities by enabling them to perform everyday tasks more easily. However, privacy and security concerns must be addressed. The success of Alexa demonstrates the potential of speech recognition technology to enhance our daily lives, but ethical considerations and the risks associated with AI must be considered.

# 5 Additional materials and resources

| Type of resource | Title | Topic | Link |
|---|---|---|---|
| Article | "Speech Recognition: The Future is Now" | Speech Recognition | https://www.wired.com/story/speech-recognition-future-is-now/ |

| | | | |
|---|---|---|---|
| Article | "Computer Vision: The Future of Artificial Intelligence" | Computer Vision | https://www.forbes.com/sites/forbestechcouncil/2021/01/05/computer-vision-the-future-of-artificial-intelligence/?sh=39c773b742f4 |
| Article | "How Deep Learning is Changing the Future of Visual Content" | Deep Learning | https://blog.hubspot.com/marketing/deep-learning-visual-content |
| Article | "The Pros and Cons of AI in Healthcare" | AI in Healthcare | https://www.healthtechmagazine.net/article/2020/03/pros-and-cons-ai-healthcare |
| Video | "TED Talk: How AI is making it easier to diagnose disease" | Pros of Vision and Speech in Healthcare | https://www.ted.com/talks/pratik_shah_how_ai_is_making_it_easier_to_diagnose_disease |
| Video | "How AI is changing the way we see the world" | The world-changing effects of AI development and use | https://www.youtube.com/watch?v=BDBTJOGvCv4 |
| Link | TensorFlow | TensorFlow: an open-source machine learning platform for building deep learning models | https://www.tensorflow.org/ |

| | | | |
|---|---|---|---|
| Link | OpenCV | OpenCV: an open-source computer vision library for building real-time computer vision applications | https://opencv.org/ |
| Link | PyTorch | PyTorch: an open-source machine learning library for building deep learning models | https://pytorch.org/ |
| Link | NVIDIA | NVIDIA Deep Learning Institute | https://www.nvidia.com/en-us/deep-learning-ai/education/ |

# 6 Wrap-up

In conclusion, the unit on the application of AI in speech and vision has covered various topics, including natural language processing, computer vision, machine learning, and societal impact. We haveexplored case studies such as Google's DeepMind AI for lip-reading, IBM's Watson for Oncology, and Amazon's Alexa for speech recognition, which illustrate the potential of AI in solving real-world problems. However, ethical considerations and the potential for bias and limitations of AI must also becarefully examined. With this knowledge, learners can continue to explore the vast opportunities and challenges in the field of AI.

# 7 Quiz

Question 1: True/False
Deep learning is a subset of machine learning. (**True**/False)
Bias and overfitting are common challenges in machine learning. (True/**False**)

Question 2: Match the following machine learning types with their descriptions:

| | |
|---|---|
| a. Supervised learning | i. Learning with labeled data |
| b. Unsupervised learning | ii. Learning with no labeled data |
| c. Reinforcement learning | iii. Learning with a reward system |

**Answers: a. i, b. ii, c. iii**

Question 3: Match the following machine learning algorithms with their applications:

| | |
|---|---|
| a. Convolutional neural networks (CNNs) | i. Speech recognition |
| b. Autoencoders | ii. Image classification |
| c. Q-learning | iii. Image classification |

**Answers: a. ii, b. i, c. iii**

Question 4: What is the main difference between supervised and unsupervised learning?

**a) Supervised learning requires labelled data while unsupervised learning does not.**
b) Unsupervised learning requires labelled data while supervised learning does not.
c) Both types of learning require labelled data.
d) There is no difference between supervised and unsupervised learning.

Question 5: Which type of learning is used for robotics and game playing?

a) Supervised learning
b) Unsupervised learning
**c) Reinforcement learning**
d) None of the above

Question 6: What is the main limitation of machine learning in Speech and Vision applications?

a) Lack of computational power
b) Lack of labelled data
**c) Bias and overfitting**
d) Lack of interpretability

Question 7: What is the main advantage of using Deep Learning?

**a) Better accuracy on complex tasks**
b) Lower computational requirements
c) Less prone to overfitting
d) Easier to interpret

Question 8: What is the main ethical consideration in developing AI for Societal Impact?
a) Ensuring the safety and security of users
**b) Ensuring fairness and avoiding bias**
c) Ensuring profitability for the company
d) Ensuring compatibility with existing systems

Question 9: What is the main advantage of using Augmented Reality in combination with AI?
**a) Enhancing the user experience**
b) Reducing computational requirements
c) Improving accuracy on complex tasks
d) Making AI more interpretable

# 8 References


- Kelleher, J. D., & Tierney, B. (2018). Data science: An introduction. CRC Press.Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
- Jurafsky, D., & Martin, J. H. (2020). Speech and language processing (3rd ed.). Pearson.
- Zhang, L., Tan, T., & Li, Y. (2016). Deep learning-based face recognition: A survey. International Journalof Automation and Computing, 13(4), 261-279.
- Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., ... & Dieleman, S. (2016). Mastering the game of Go with deep neural networks and tree search. Nature, 529(7587), 484- 489.
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Petersen, S. (2015).Human-level control through deep reinforcement learning. Nature, 518(7540), 529-533.
- Weng, L., Zhang, Y., Xue, X., & Chen, K. (2021). The applications of artificial intelligence in computer vision. In 2021 International Conference on Computer Science and Artificial Intelligence (CSAI) (pp. 139-143). IEEE.
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 580-587).
- Dhillon, G., & Moores, D. (2018). Natural language processing in AI: past, present and future. In Proceedings of the 2nd International Conference on Natural Language Processing and Information Retrieval (pp. 1-10).
- Canny, J. (2021). Ethics and artificial intelligence. Communications of the ACM, 64(1), 18-20.
- Li, X., Liang, J., & Li, Y. (2021). Research on the impact of augmented reality technology on the experience of cultural tourism. Journal of Physics: Conference Series, 1838(1), 012064.
- Yaqoob, I., Ahmed, M. M., Gani, A., Imran, M., Guizani, M., & Hitha, H. (2019). Virtual reality for education:A survey. Journal of Network and Com