

5.DA.3 Data and Analysis

The student will explain the significance of training data in machine learning. (a) Compare how training data is utilized in supervised, unsupervised and reinforcement learning. (b) Explain how training data is used to make classification predictions. (c) Discuss the need and significance of diverse, inclusive, and large datasets.



Integration Opportunities

Understanding the Standard

Machine learning, while rooted in the study of artificial intelligence (AI), is different from the generative AI popular in contemporary media. Machine learning is more useful in finding practical approaches to analyzing massive, complex data sets that may be too cumbersome to process with an explicit, human-written algorithm. The data used as the foundation of these predictions is called "training data". Companies making machine learning technologies can collect this data in all sorts of ways, from scraping data available on the public Internet to collecting data from individual users on social media. The training data has a big impact on the predictions the machine learning technology can make. Training data that is limited in scope or quantity will result in flawed, inconsistent, or even harmful predictions.

Term	Definition
Machine learning (ML)	A branch of AI (artificial intelligence) focused on creating computer systems that accomplish tasks without explicit instructions. (from ComputerHope)
Supervised machine learning	A type of ML in which a model is trained on a labeled dataset.
Unsupervised machine learning	A type of ML in which training data is "unlabeled," meaning its desired output is not provided to the algorithm. (from ComputerHope)
Reinforcement learning	A type of ML in which the system focuses on how to increase the chances of accumulating a quantifiable reward. (from ComputerHope)
Training data	The set or collection of data used to train an AI model to recognize patterns or make decisions.

Science 5.4a Create a machine learning model to identify conductors vs. insulators and explain how training data is used to make predictions for future materials.

Math 5.MG.3d Use Teachable Machine to classify different types of triangles based on angle measurements and use training data to help the model make accurate predictions, while discussing how training data improves classification accuracy.

Physical Education 5.5 a,b Explore the Recommended Dietary Allowance and the data used to create these guidelines. Discuss the importance of a diverse dataset and have students explain why there are different RDAs for various age groups.

Music 5.3a Create a machine learning model to group a variety of instruments based on how their sounds are produced.

Prerequisite Knowledge

Students should have a basic understanding of "AI", and "machine learning" in broad terms before engaging with this standard.

Summary of a Lesson

Have students list technologies they think might be "artificial intelligence". Then, introduce the main characteristics of AI: collecting data, analyzing data, and making predictions. Facilitate a discussion where students identify what data the AI technologies use to make predictions. Then, brainstorm ways that the attributes of that data might result in bad predictions. You can address this standard with a hands-on activity in which students create simple machine learning models using [Teachable Machine](#) or another similar tool.