APRIL 17

The Role of Data in Artificial Intelligence Literacy in School Education

Ismaila Sanusi, PhD

School of computing, UEF

ProDaBi
Projekt Data Science und Big Data in der Schule

Overview

State of teaching and learning AI in K-12 Education
– Early and ongoing effort

Data science in AI for school education

How data can be further integrated in AI lessons to support students learning.

State of teaching & learning AI in K-12

GatesNotes THE BLOG OF BILL GATES

A NEW ERA

The Age of AI has begun

Artificial intelligence is as revolutionary as mobile phones and the Internet.

By **Bill Gates** | March 21, 2023 • 14 minute read

Insight

29 February 2024

> "The rapid growth and adoption of AI technology in recent years is just the beginning

The future of AI: What to expect in the next 5 years

AI's impact in the next five years? Human life will speed up, behaviors will change and industries will be transformed -- and that's what can be predicted with certainty.

By **Michael Bennett,** Northeastern University

Published: **25 Jan 2024**

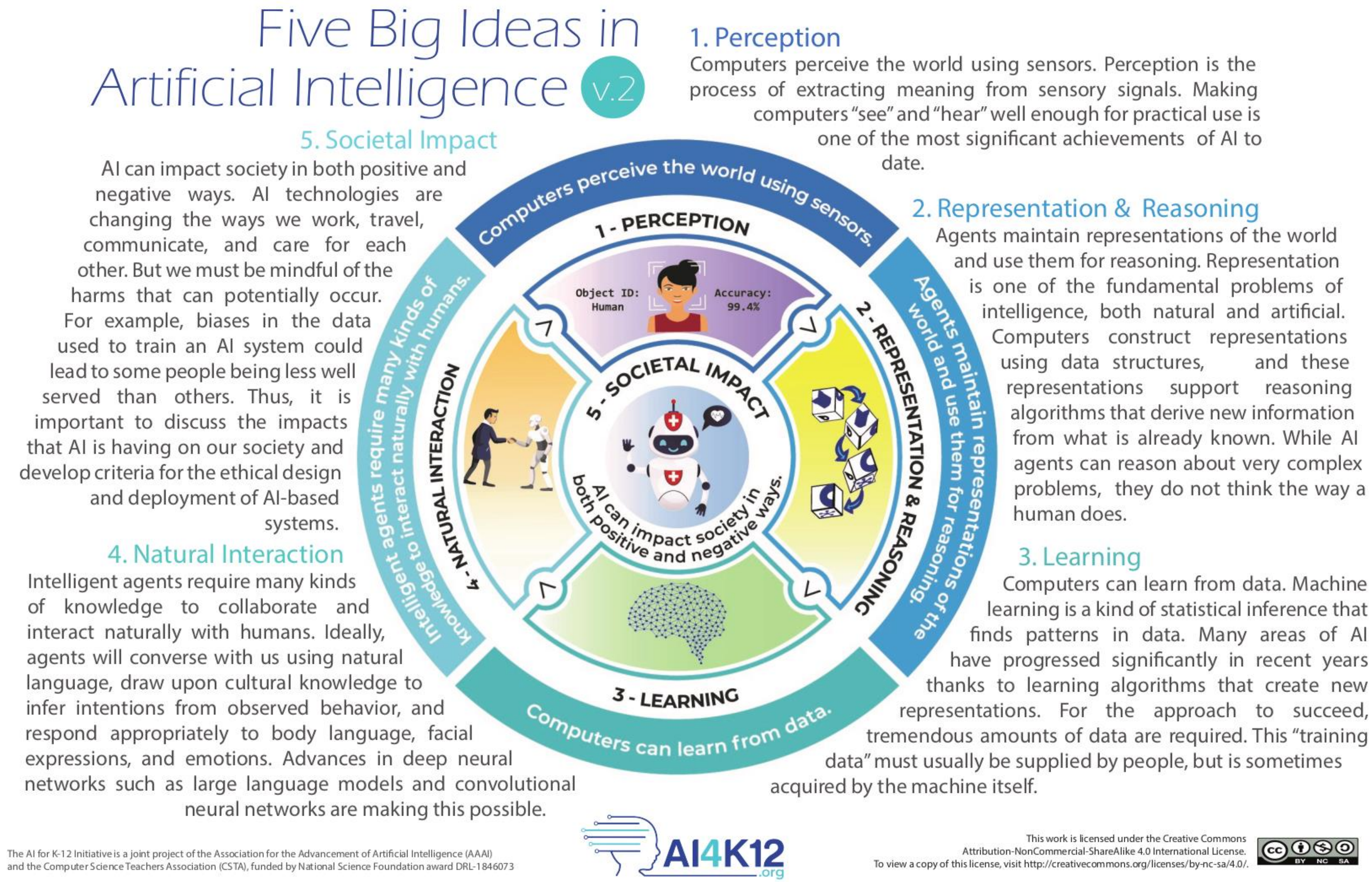
Why teach AI at K-12 levels

- educational: future learning; support learning of other subjects

- personal : Career readiness, critical thinking

- broader societal benefits: social good; citizen empowerment

Teaching AI in K-12



AI4K12.org

Integrating AI and Machine Learning in Software Engineering Course for High School Students

Sperling, Ahuva
Leo Baeck Education Center
Haifa, Israel
97248300500
asperling@leobaeck.net

Lickerman, Dorit
Leo Baeck Education Center
Haifa, Israel
97248300500
dlickerman@gmail.com

2012

ABSTRACT

This paper describes a unique software engineering curriculum for high-school students that includes subjects in artificial intelligence and machine learning. The students in the course deal with the implementation of solutions to riddles and games (complex algorithmic problems), use the DrRacket functional programming language as a tool that supports their comprehension and thorough understanding of blind search algorithms, informed search algorithms, search games trees and machine learning algorithms. During their studies, the students engage in self-learning,

2. LEO BAECK EDUCATION CENTER

The Leo Baeck education center in Haifa was founded in 1938 for children who had survived the Holocaust. Its pluralistic approach is committed to democracy, egalitarianism and human rights, as well as to the teaching of the living values of progressive Judaism which inspires social change and imroving the world. The education center is one of Israel's finest institutions for academic excellence and it also promotes community outreach and social action.

AN INTRODUCTION TO MACHINE LEARNING FOR STUDENTS IN SECONDARY EDUCATION

2011

Steven D. Essinger, Gail L. Rosen

Drexel University
Department of Electrical & Computer Engineering
3141 Chestnut Street
Philadelphia, PA 19104

ABSTRACT

We have developed a platform for exposing high school students to machine learning techniques for signal processing problems, making use of relatively simple mathematics and engineering concepts. Along with this platform we have created two example scenarios which give motivation to the stu-

the brain [3]. While this class of methods undoubtedly employs coarse approximations of actual neuron function, they have shown tremendous success in several ML applications [4].

A few examples of ML applications include speech recognition aka natural language processing, image processing such as face detection, DNA sequence classification, financial

An Action Research Report from a Multi–Year Approach to Teaching Artificial Intelligence at the K–6 Level

2016

Dr. Clint Heinze
Air Operations Division
Defence Science & Technology Organisation
Fishermans Bend, Victoria, 3207, Australia
clinton.heinze@dsto.defence.gov.au

Janet Haase and **Helen Higgins**
Manchester Primary School
Department of Education & Early Childhood Development
Mooroolbark, Victoria, 3138, Australia
haase.janet.e@edumail.vic.gov.au

Abstract

In Australia, the *Scientists–in–Schools* program partners professional scientists with teachers from K-12 schools to improve early engagement and educational outcomes in the sciences and mathematics. An overview of the developing syllabus of a K-6 course resulting from the pairing of a senior AI researcher with teachers from a K-6 (primary) school is presented. Now entering its third year, the course introduces

monwealth Scientific and Industrial Research Organisation (CSIRO) and it:

"... *promotes science education in primary and secondary schools, helps to engage and motivate students in their learning of science, and broadens awareness of the types and variety of exciting careers available in the sciences.*"— *Dr. Jim Peacock, Australia's Chief Scientist 2006–2008*

What students should know and should be able to do.

Big Idea 1: Perception

- **Sensing**
  Living things, Computer Sensors, Digital Encoding
- **Processing**
  Sensing vs. Perception, Feature extraction, Abstraction pipeline: Language, Abstraction pipeline: Vision
- **Domain Knowledge**
  Types of domain knowledge, Inclusivity

Big Idea 2: Representation and Reasoning

- **Representation**
  Abstract, symbolic representation, data structure, feature vectors
- **Search**
  State spaces and operations, combinatorial search
- **Reasoning**
  Type of reasoning problems, reasoning algorithms

Big Idea 3: Learning

- **Nature of learning**
  Human vs machine, finding patterns in data, training models, constructing vs using reasoners, adjusting internal representation, learning from experience
- **Neural networks**
  Structure of neural network, weight adjustment
- **Datasets**
  Feature set, large dataset, bias

Big Idea 4: Natural Interaction

- **Natural language**
  Structure of language, ambiguity of language, reasoning about text, applications
- **Common sense reasoning**
- **Understanding emotion**
- **Philosophy of mind**

Big Idea 5: Societal Impact

- **Ethical AI**
  Diversity of interests and Disparate Impacts, Ethical Design Criteria, Practicing Ethical Design
- **AI and Culture**
  AI in daily life, Trust and Responsibility
- **AI and the Economy**
  Impact of AI on Sectors of the society, Effects on Unemployment
- **AI for social good**
  Democratization of AI technology, Using AI to solve societal problems

AI LITERACY
COMPETENCIES & DESIGN CONSIDERATIONS

WHAT IS AI?

COMPETENCIES
1. Recognizing AI
2. Understanding Intelligence
3. Interdisciplinarity
4. General vs. Narrow

WHAT CAN AI DO?

COMPETENCIES
5. AI's Strengths & Weaknesses
6. Imagine Future AI

HOW DOES AI WORK?

COMPETENCIES
7. Representations
8. Decision-Making
9. Explainability
10. ML Steps
11. Data Literacy
12. Learning from Data
13. Critically Interpreting Data
14. Action & Reaction
15. Sensors

DESIGN CONSIDERATIONS
1. Explainability
2. Embodied Interactions
3. Contextualizing Data

WHAT SHOULD AI DO?

COMPETENCIES
16. Ethics

HOW DO PEOPLE PERCEIVE AI?

COMPETENCIES
17. Programmability

DESIGN CONSIDERATIONS
4. Promote Transparency
5. Unveil Gradually
6. Opportunities to Program
7. Milestones
8. Critical Thinking
9. Culture
10. Support for Parents
11. Social Interaction
12. Leverage Learners' Interests
13. Acknowledge Preconceptions
14. New Perspectives
15. Low Barrier to Entry

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1-16).

Machine Learning Education Framework

For Transforming ML Consumers to be ML Contributors

Knowledge

- **General ML Knowledge:** What ML is how it is created
- **Knowledge of ML Methods**
- **Bias in ML Systems:** ML encodes and amplifies bias
- **Societal Implications of ML**

Skills

- **ML Problem Scoping:** When can you vs. should you use ML
- **ML Project Planning**
- **Creating ML Artifacts**
- **Analysis of ML Design Intentions & Results**
- **ML Advocacy –** Critical, and education community engagement with ML policies, products
- **Independent Out-of-Class Learning**

Attitudes

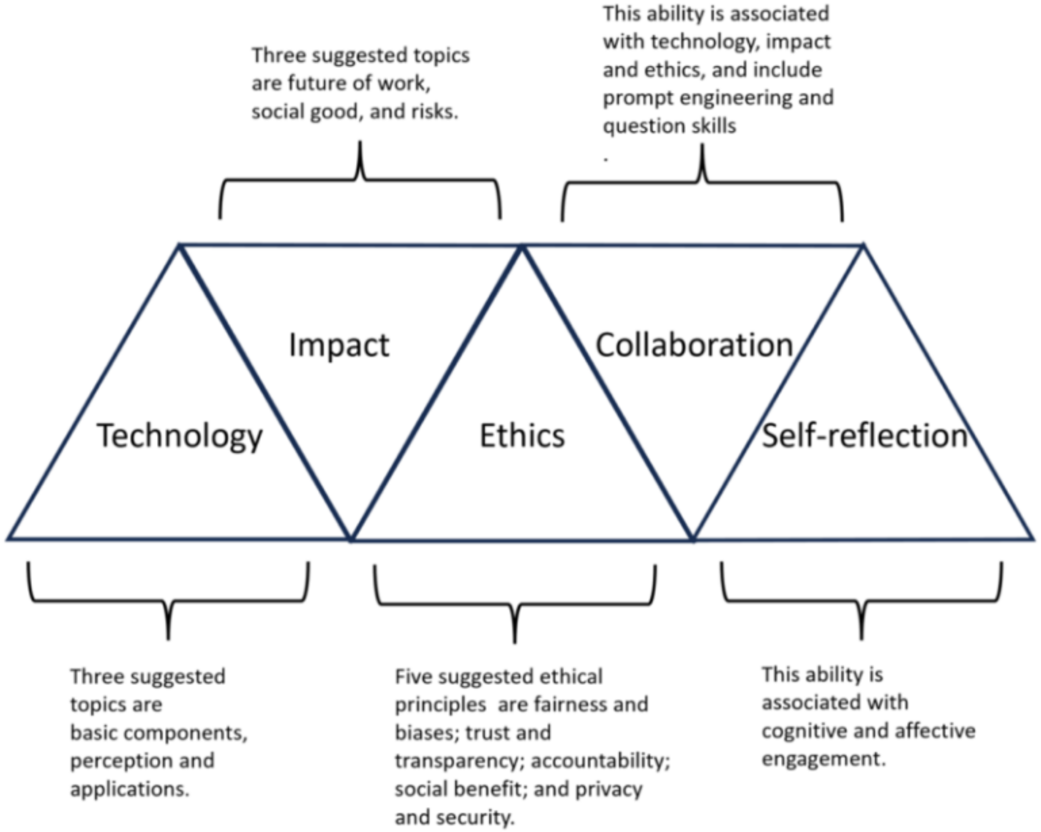
- **Interest**
- **Identity & Community**
- **Self-Efficacy**
- **Persistence**

Applicable for educating ML-engaged citizens and ML tinkerers. Extendable to technical education in ML.

- Minimum requirements for ML-engaged citizen → tinkerer → researcher/engineer

Lao, N. (2020). *Reorienting machine learning education towards tinkerers and ML-engaged citizens* (Doctoral dissertation, Massachusetts Institute of Technology).

AI Literacy and Competency



Chiu, K.F., Ahmad, Z., Ismailov, M & Sanusi, I.T. (2024). What are Artificial Intelligence Literacy and Competency? A Comprehensive Framework to Support Them. *Computers and Education Open*, 100171.

Putting the frameworks together

- Emphasized the knowledge of AI
- The need to engage effectively with AI applications
- attitude including perception which highlights the perspectives of stakeholders
- Ethical and societal implication of AI

My Works

Curriculum activities

Developing AI/ML techniques understandings (e.g., kNN, decision trees)

Ethics

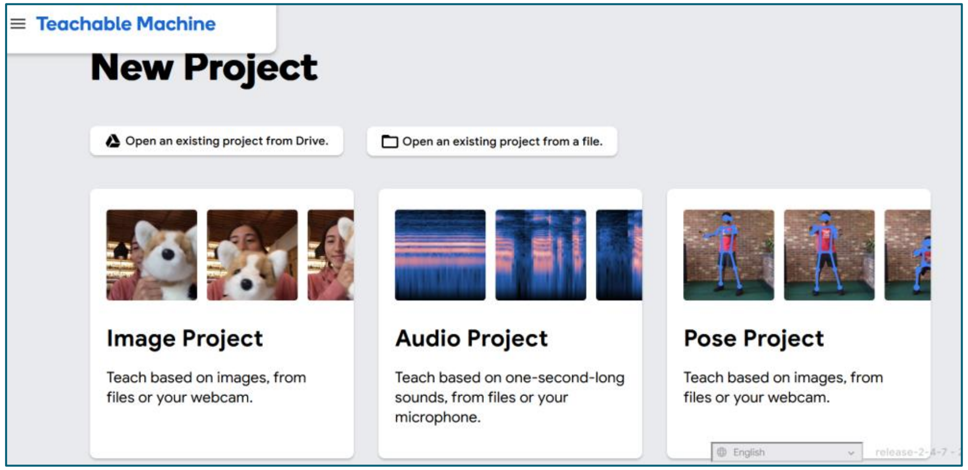
Tools and the use of diverse pedagogies for developing an understanding about AI

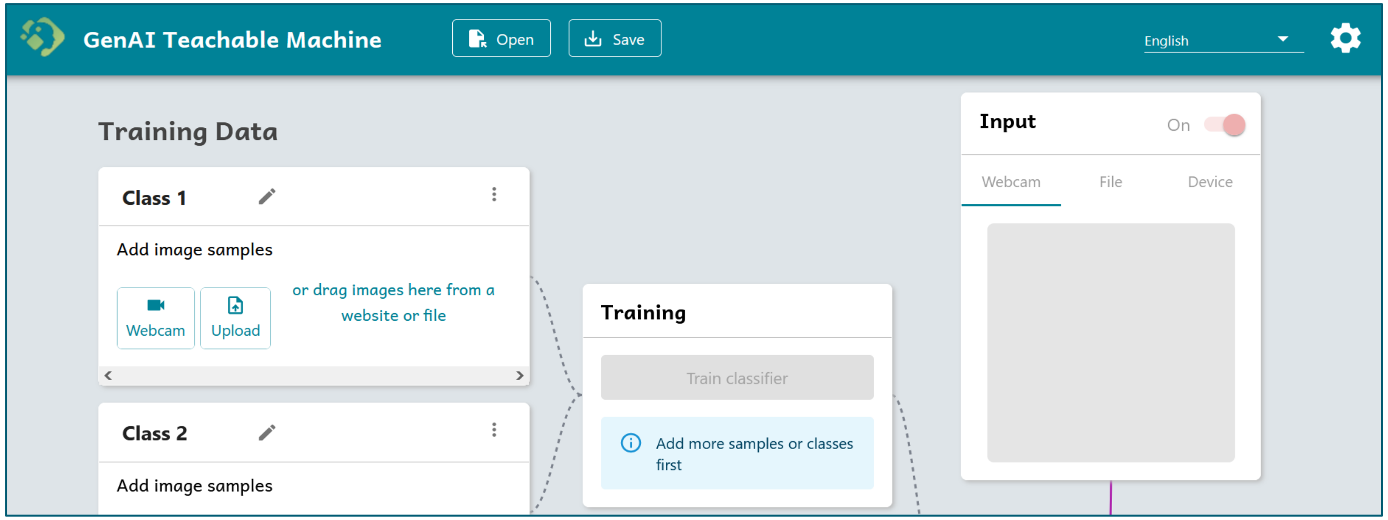
Broadening participation

Integration of AI learning into other subjects

Co-design with teachers and teacher preparation

Tools

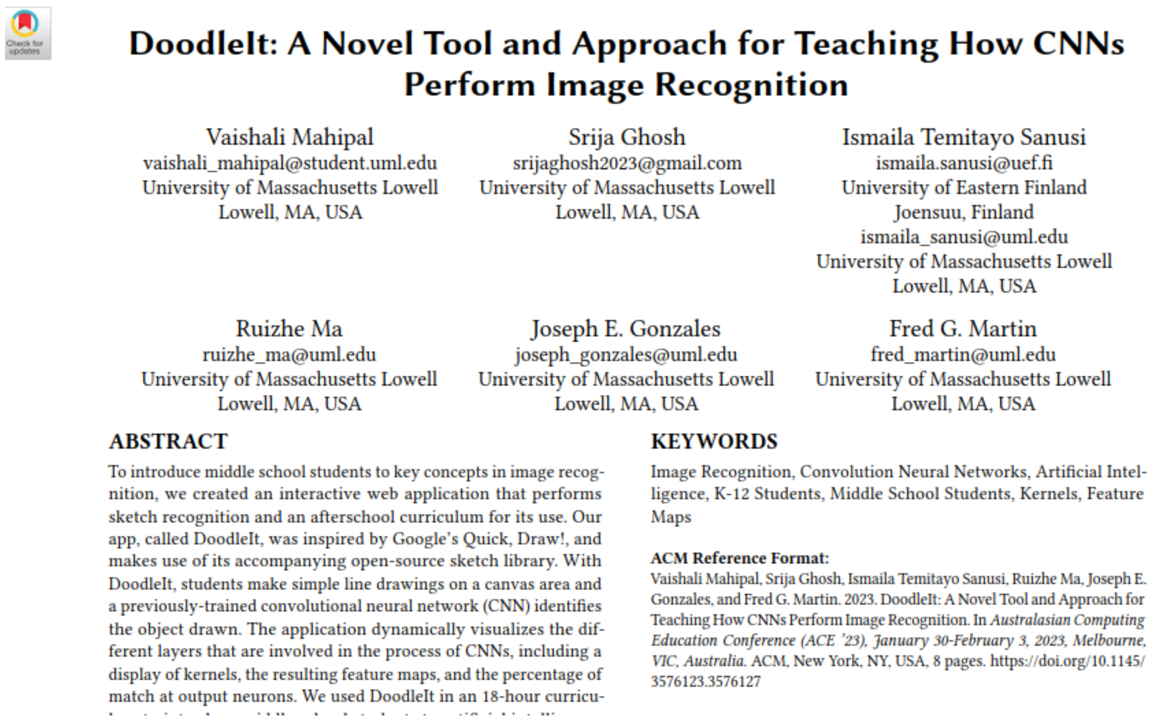




https://teachablemachine.withgoogle.com/

https://tm.gen-ai.fi/

Tools



DoodleIt: A Novel Tool and Approach for Teaching How CNNs Perform Image Recognition

Vaishali Mahipal
vaishali_mahipal@student.uml.edu
University of Massachusetts Lowell
Lowell, MA, USA

Srija Ghosh
srijaghosh2023@gmail.com
University of Massachusetts Lowell
Lowell, MA, USA

Ismaila Temitayo Sanusi
ismaila.sanusi@uef.fi
University of Eastern Finland
Joensuu, Finland
ismaila_sanusi@uml.edu
University of Massachusetts Lowell
Lowell, MA, USA

Ruizhe Ma
ruizhe_ma@uml.edu
University of Massachusetts Lowell
Lowell, MA, USA

Joseph E. Gonzales
joseph_gonzales@uml.edu
University of Massachusetts Lowell
Lowell, MA, USA

Fred G. Martin
fred_martin@uml.edu
University of Massachusetts Lowell
Lowell, MA, USA

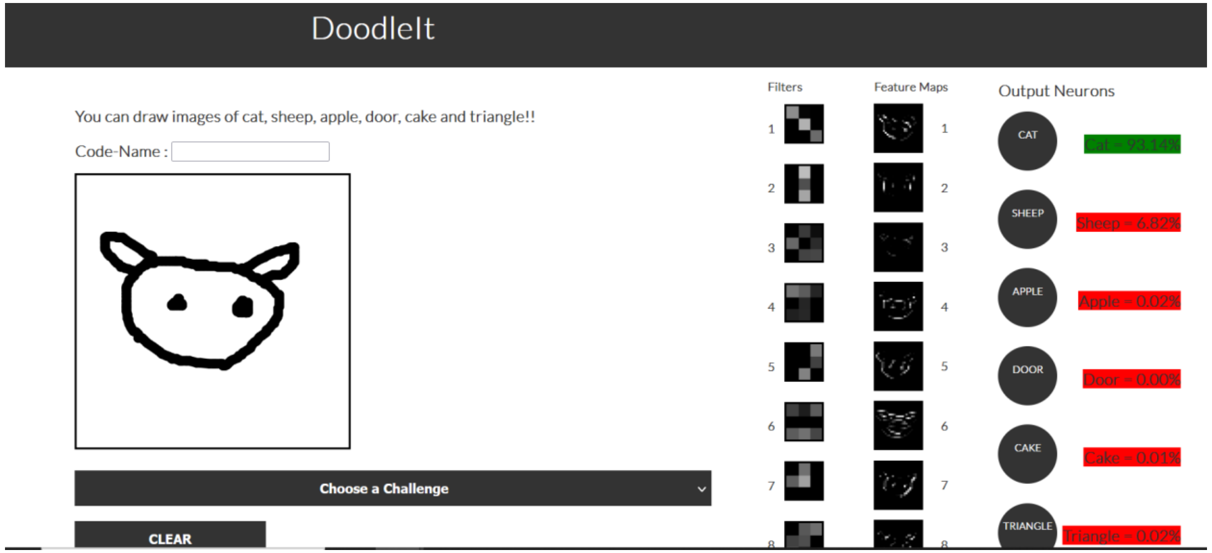
ABSTRACT

To introduce middle school students to key concepts in image recognition, we created an interactive web application that performs sketch recognition and an afterschool curriculum for its use. Our app, called DoodleIt, was inspired by Google's Quick, Draw!, and makes use of its accompanying open-source sketch library. With DoodleIt, students make simple line drawings on a canvas area and a previously-trained convolutional neural network (CNN) identifies the object drawn. The application dynamically visualizes the different layers that are involved in the process of CNNs, including a display of kernels, the resulting feature maps, and the percentage of match at output neurons. We used DoodleIt in an 18-hour curricu-

KEYWORDS

Image Recognition, Convolution Neural Networks, Artificial Intelligence, K-12 Students, Middle School Students, Kernels, Feature Maps

ACM Reference Format:
Vaishali Mahipal, Srija Ghosh, Ismaila Temitayo Sanusi, Ruizhe Ma, Joseph E. Gonzales, and Fred G. Martin. 2023. DoodleIt: A Novel Tool and Approach for Teaching How CNNs Perform Image Recognition. In *Australasian Computing Education Conference (ACE '23), January 30-February 3, 2023, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3576123.3576127



https://www.cs.uml.edu/~fredm/DoodleItUpdated/LatestFiles:%20v2

In *Proceedings of the 25th Australasian computing education conference*

Tools

ChemAIstry: A Novel Software Tool for Teaching Model Training in K-8 Education

Fred Martin
fred.martin@utsa.edu
University of Texas at San Antonio
San Antonio, Texas, USA

Vaishali Mahipal
vaishali_mahipal@uml.edu
University of Massachusetts Lowell
Lowell, Massachusetts, USA

Garima Jain
garima_jain@student.uml.edu
University of Massachusetts Lowell
Lowell, Massachusetts, USA

Srija Ghosh
srijaghosh2015@gmail.com
University of Massachusetts Lowell
Lowell, Massachusetts, USA

Ismaila Temitayo Sanusi
ismaila.sanusi@uef.fi
University of Eastern Finland
Joensuu, Finland

ABSTRACT

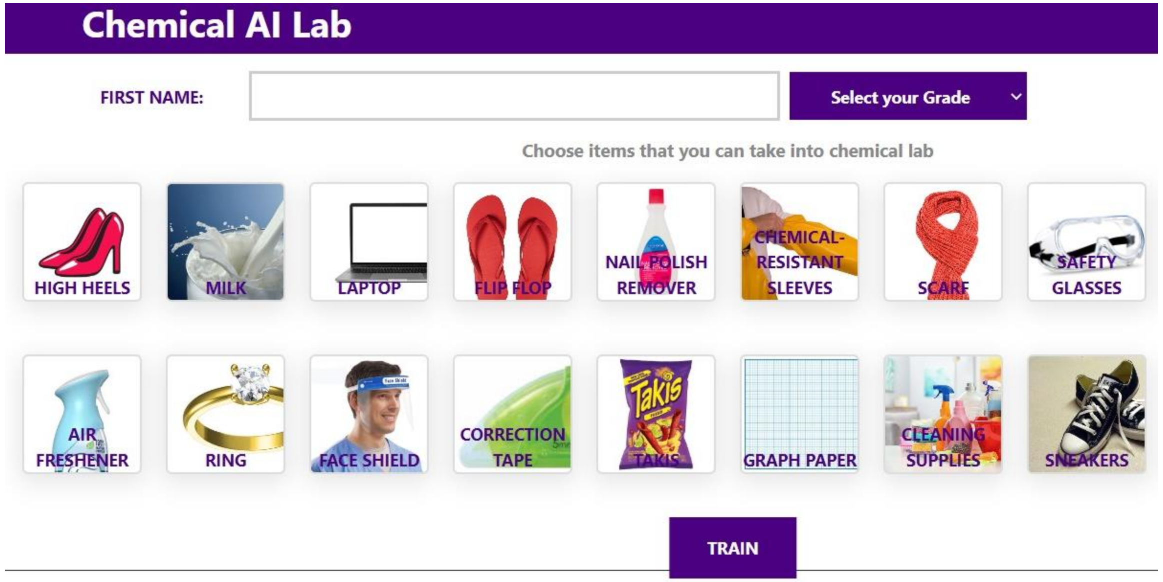
Machine learning (ML) systems are increasingly in use in society. For young learners to be informed citizens and have full career potential it is important for them to understand these concepts. To support this learning, we created "ChemAIstry," an interactive software tool for children which demonstrates training and classification in machine learning. Students select which everyday items are safe to bring into a chemistry lab (e.g., a lab coat is safe; pizza is not). These selections serve as training input for a decision tree classifier. After training, students see how the trained model performs in classifying new objects. ChemAIstry was tested with 40 students aged 7 to 14 years at a public K–8 school. The software captured student selections during training. We analyzed these interactions to yield a "Correspondence Score," a measure of stu-

ACM Reference Format:
Fred Martin, Vaishali Mahipal, Garima Jain, Srija Ghosh, and Ismaila Temitayo Sanusi. 2024. ChemAIstry: A Novel Software Tool for Teaching Model Training in K-8 Education. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1 (SIGCSE 2024), March 20–23, 2024, Portland, OR, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3626252.3630804

1 INTRODUCTION AND MOTIVATION

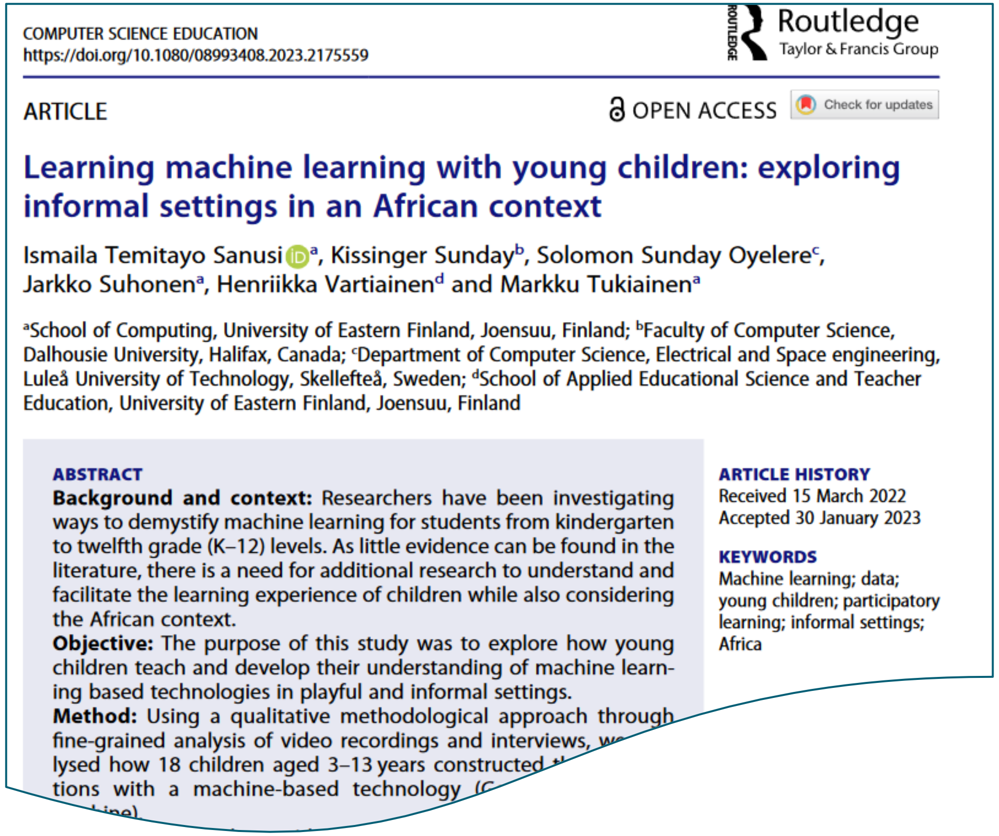
Artificial Intelligence (AI) is having an unprecedented impact on society as the amount of data and processing capacity are expanding quickly. The widespread deployment of AI in many different disciplines and industries emphasizes the need to develop a workforce with strong computing abilities and the capacity to work with

In Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1



https://engaging-computing.github.io/ChemAIstry/

Pedagogy


COMPUTER SCIENCE EDUCATION
https://doi.org/10.1080/08993408.2023.2175559

Routledge
Taylor & Francis Group

ARTICLE

OPEN ACCESS Check for updates

Learning machine learning with young children: exploring informal settings in an African context

Ismaila Temitayo Sanusi[a], Kissinger Sunday[b], Solomon Sunday Oyelere[c], Jarkko Suhonen[a], Henriikka Vartiainen[d] and Markku Tukiainen[a]

[a]School of Computing, University of Eastern Finland, Joensuu, Finland; [b]Faculty of Computer Science, Dalhousie University, Halifax, Canada; [c]Department of Computer Science, Electrical and Space engineering, Luleå University of Technology, Skellefteå, Sweden; [d]School of Applied Educational Science and Teacher Education, University of Eastern Finland, Joensuu, Finland

ABSTRACT

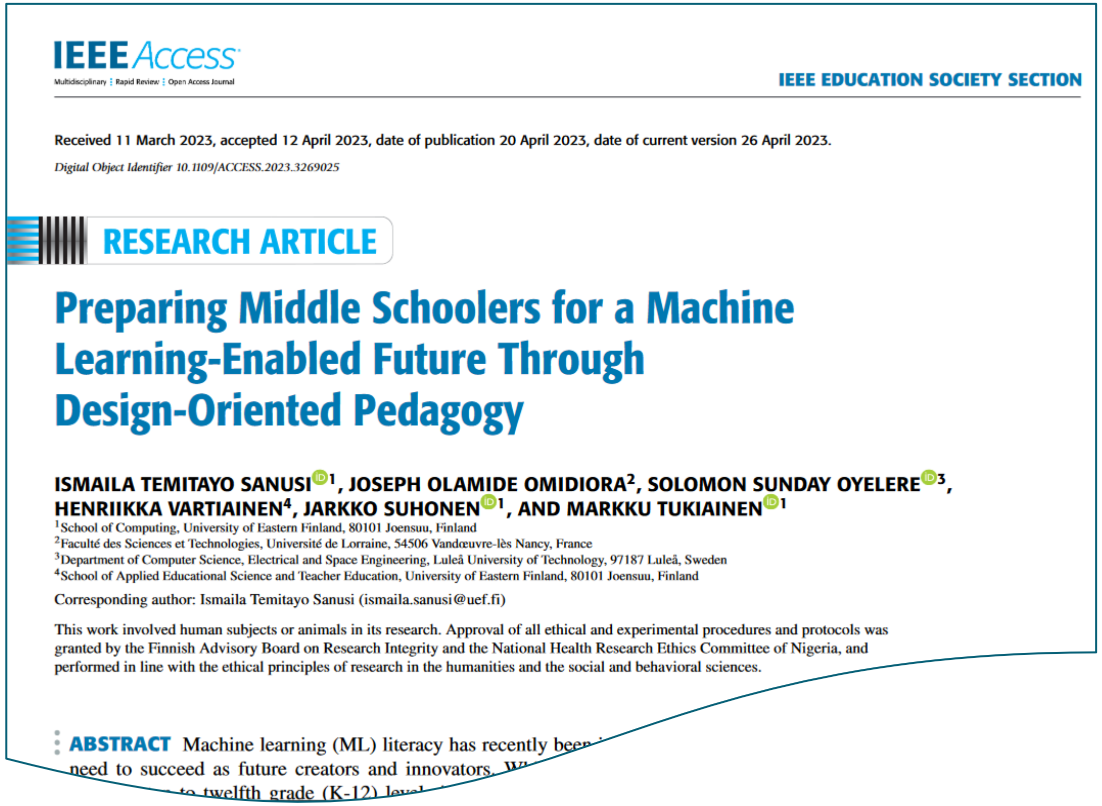
Background and context: Researchers have been investigating ways to demystify machine learning for students from kindergarten to twelfth grade (K–12) levels. As little evidence can be found in the literature, there is a need for additional research to understand and facilitate the learning experience of children while also considering the African context.

Objective: The purpose of this study was to explore how young children teach and develop their understanding of machine learning based technologies in playful and informal settings.

Method: Using a qualitative methodological approach through fine-grained analysis of video recordings and interviews, we ... lysed how 18 children aged 3–13 years constructed ... tions with a machine-based technology (...

ARTICLE HISTORY
Received 15 March 2022
Accepted 30 January 2023

KEYWORDS
Machine learning; data; young children; participatory learning; informal settings; Africa


IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

IEEE EDUCATION SOCIETY SECTION

Received 11 March 2023, accepted 12 April 2023, date of publication 20 April 2023, date of current version 26 April 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3269025

RESEARCH ARTICLE

Preparing Middle Schoolers for a Machine Learning-Enabled Future Through Design-Oriented Pedagogy

ISMAILA TEMITAYO SANUSI[1], JOSEPH OLAMIDE OMIDIORA[2], SOLOMON SUNDAY OYELERE[3], HENRIIKKA VARTIAINEN[4], JARKKO SUHONEN[1], AND MARKKU TUKIAINEN[1]

[1]School of Computing, University of Eastern Finland, 80101 Joensuu, Finland
[2]Faculté des Sciences et Technologies, Université de Lorraine, 54506 Vandœuvre-lès Nancy, France
[3]Department of Computer Science, Electrical and Space Engineering, Luleå University of Technology, 97187 Luleå, Sweden
[4]School of Applied Educational Science and Teacher Education, University of Eastern Finland, 80101 Joensuu, Finland

Corresponding author: Ismaila Temitayo Sanusi (ismaila.sanusi@uef.fi)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Finnish Advisory Board on Research Integrity and the National Health Research Ethics Committee of Nigeria, and performed in line with the ethical principles of research in the humanities and the social and behavioral sciences.

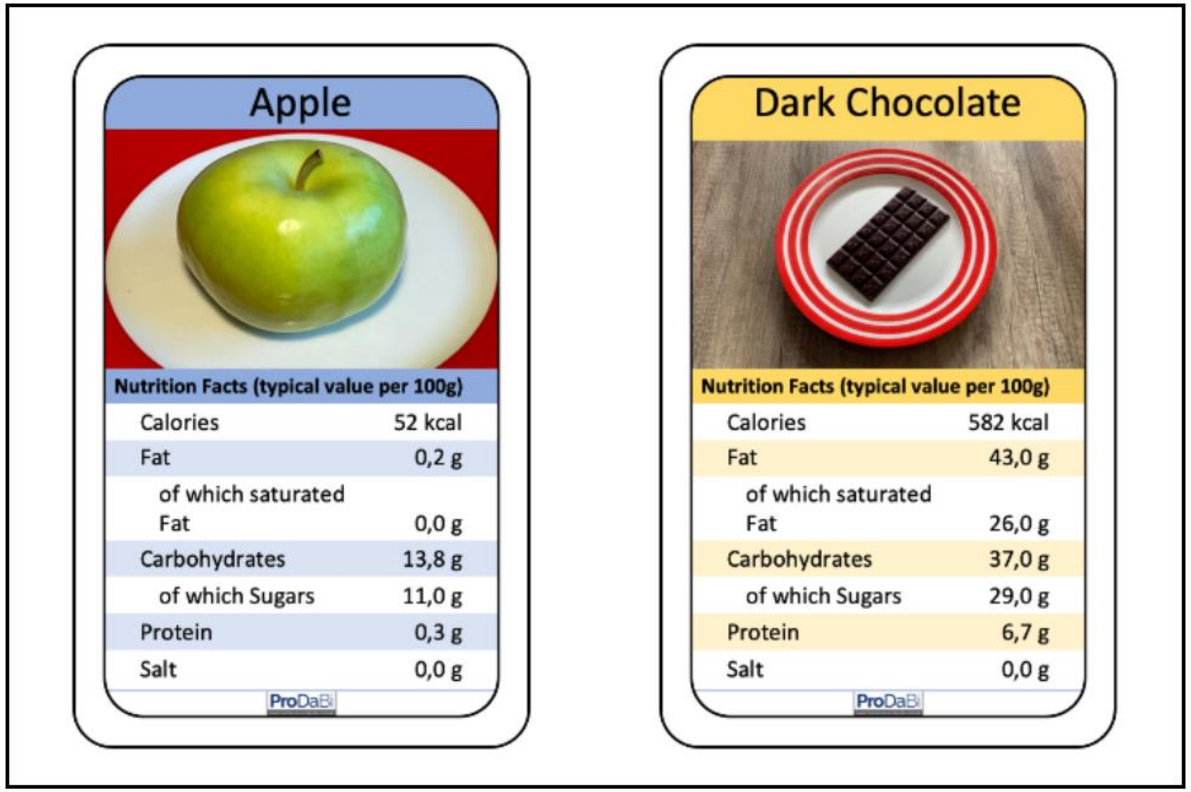
ABSTRACT Machine learning (ML) literacy has recently been ... need to succeed as future creators and innovators. ... to twelfth grade (K-12) level ...

What have been the practices regarding explicitly teaching about data in the context of AI education?

Data is an integral element of AI

Using data cards for teaching decision trees

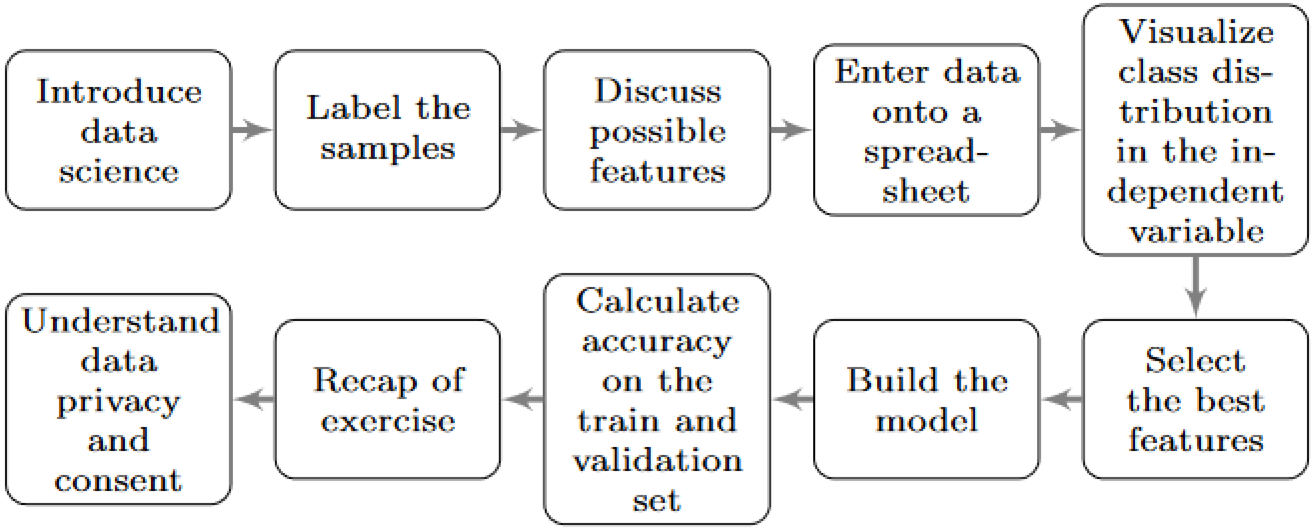
- Data modelling in ML
- Data labelling
- Test data
- Training data



Podworny, S., Fleischer, Y., Hüsing, S., Biehler, R., Frischemeier, D., Höper, L., & Schulte, C. (2021). Using data cards for teaching data based decision trees in middle school. In *Proceedings of the 21st Koli Calling International Conference on Computing Education Research* (pp. 1-3).

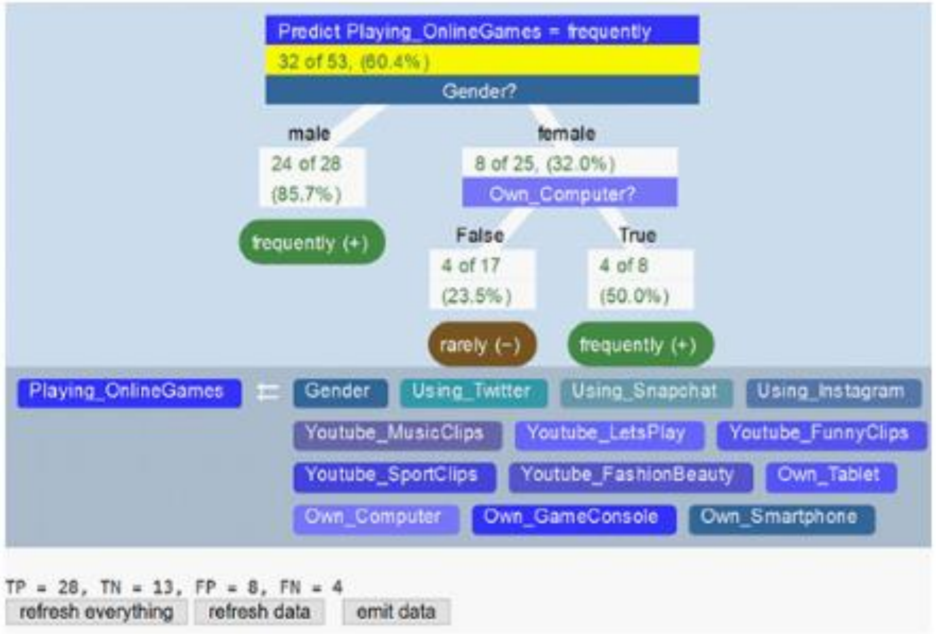
Introducing students to the full cycle of a typical supervised learning approach

- supervised learning approach
  - data collection
  - data entry,
  - data visualization,
  - feature engineering,
  - model building,
  - model testing and data permissions

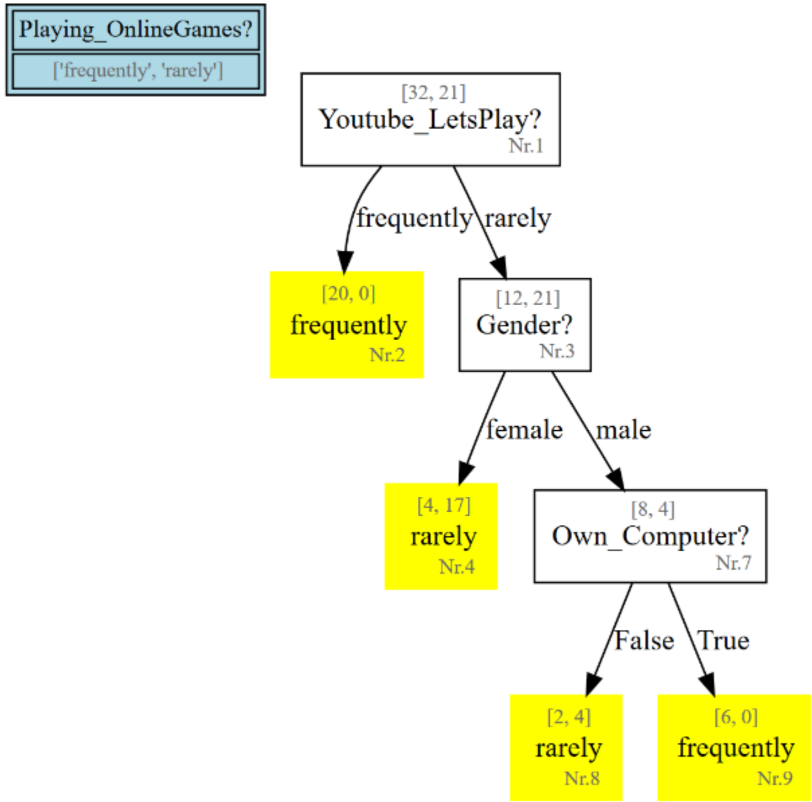


Srikant, S., & Aggarwal, V. (2017). Introducing data science to school kids. In *Proceedings of the 2017 ACM SIGCSE technical symposium on computer science education* (pp. 561-566).

Introducing students to ML with decision trees using CODAP and Jupyter Notebooks

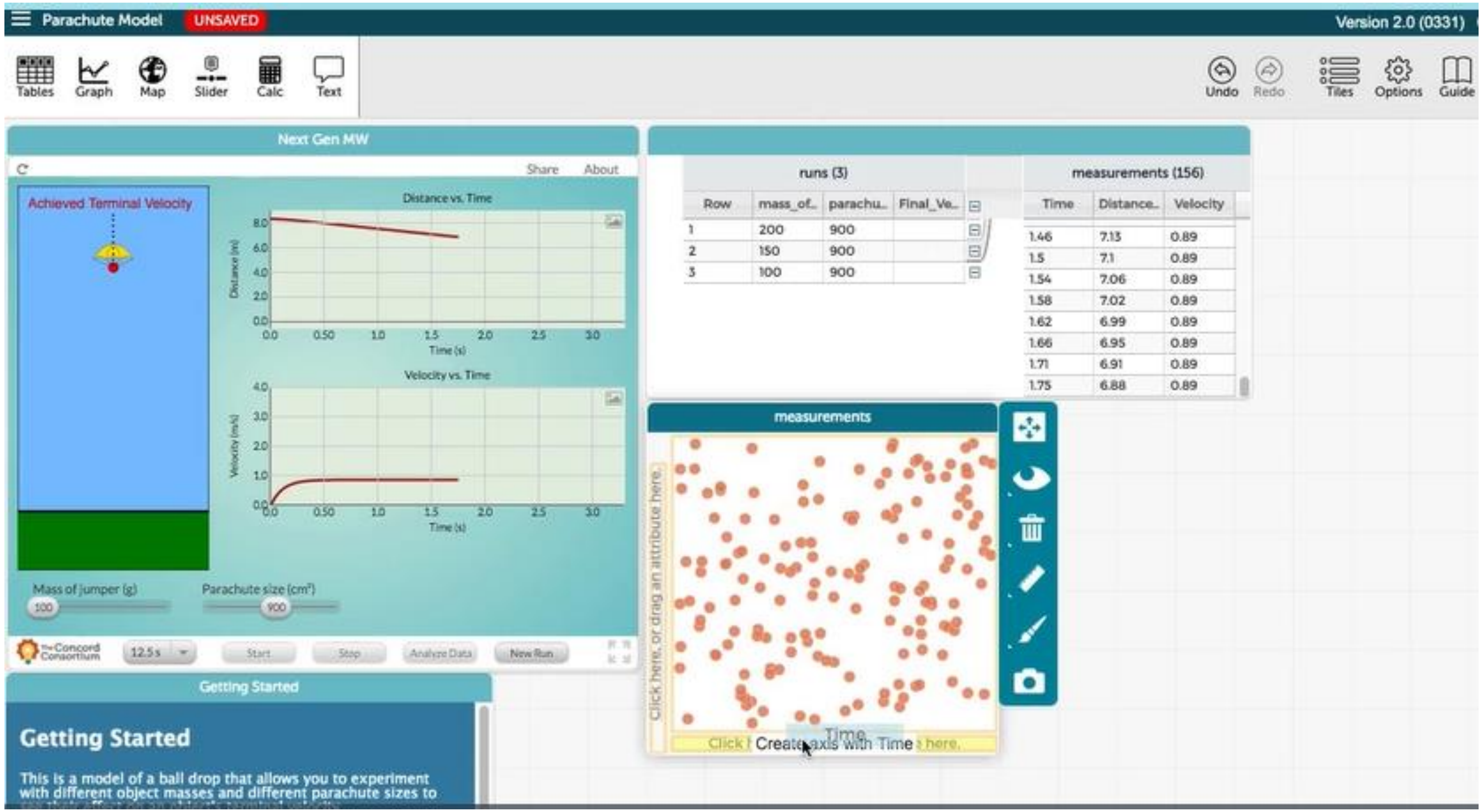


Using CODAP with the Decision Tree Plug-In



Decision Trees created with the ProDaBi Decision Tree Jupyter Notebook

Biehler, R., & Fleischer, Y. (2021). Introducing students to machine learning with decision trees using CODAP and Jupyter Notebooks. *Teaching Statistics*, *43*, S133-S142.



Common Online Data Analysis Platform (CODAP)

https://codap.concord.org/

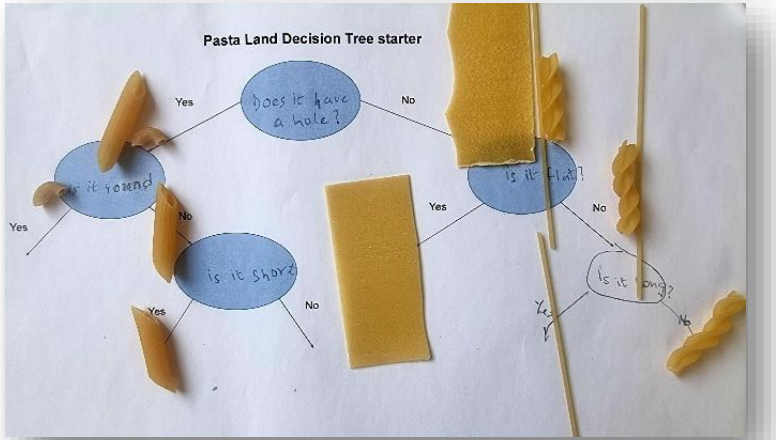
AI MYDATA

1. Introduction to AI and Image Recognition.
2. Machine Learning Techniques.
3. Data Creation, Exploration and Visualization.
4. Ethical Implication of AI.

Sanusi et al. (2024). AI MyData: Fostering middle school students' engagement with machine learning through an ethics-infused AI curriculum.

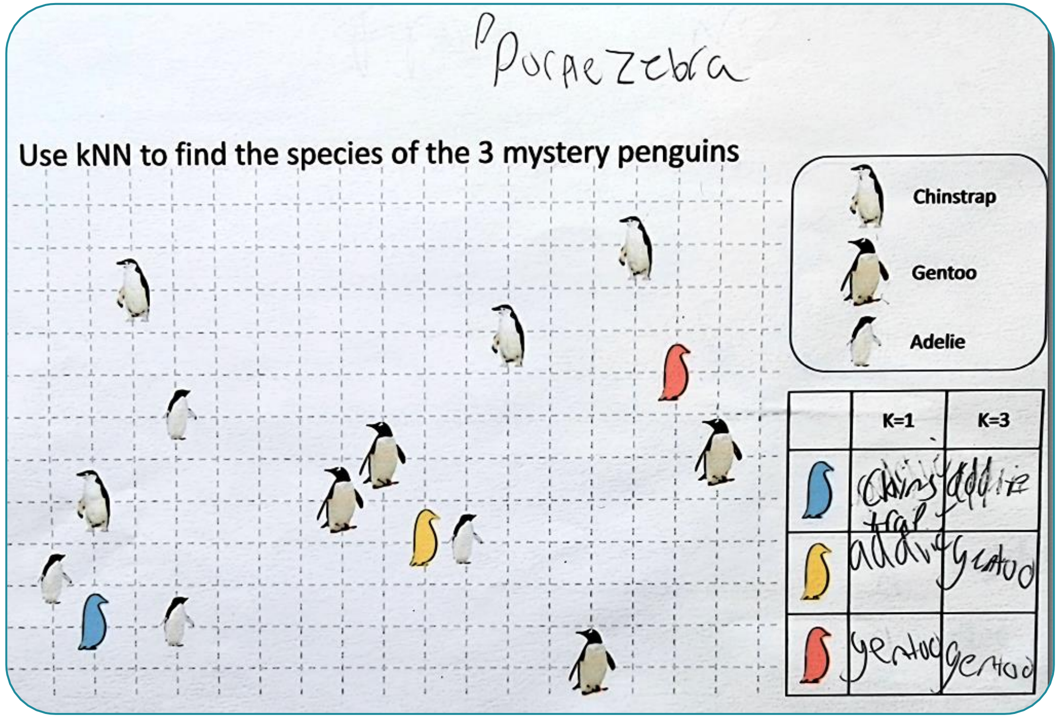
Pasta Land – Decision tree

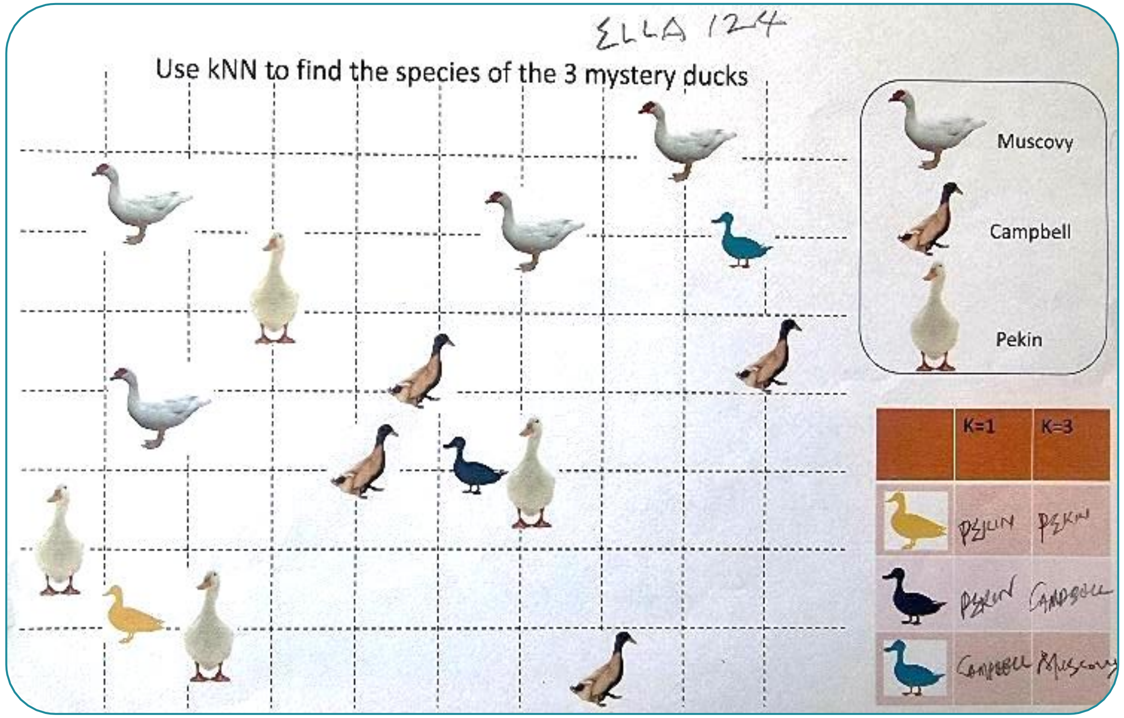
- Introduces the idea of a decision tree classifier.
- “Unplugged” – no computers needed.
- Co-developed Fred Martin & Irene Lee, MIT and UTSA.



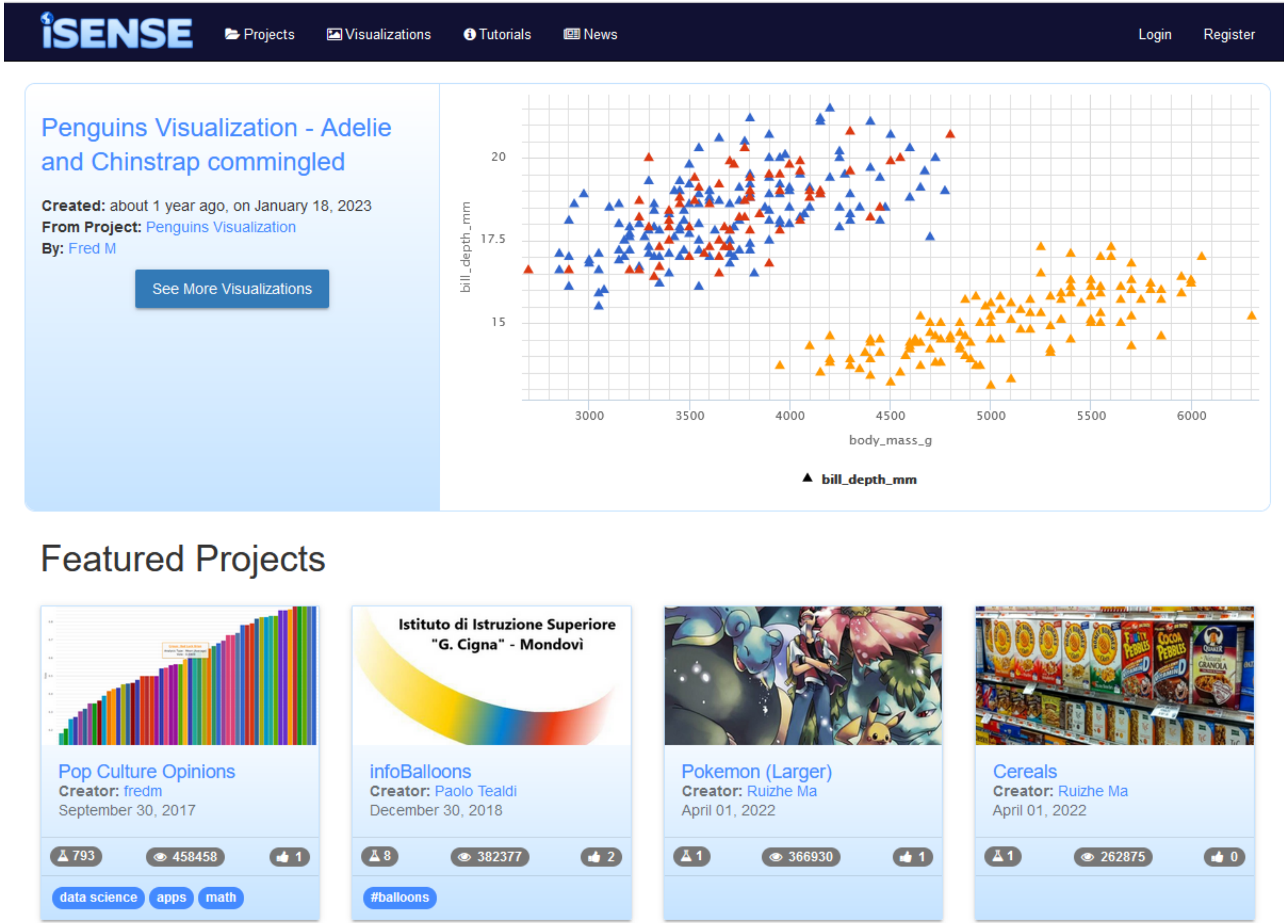
Lots of pasta

k-NN

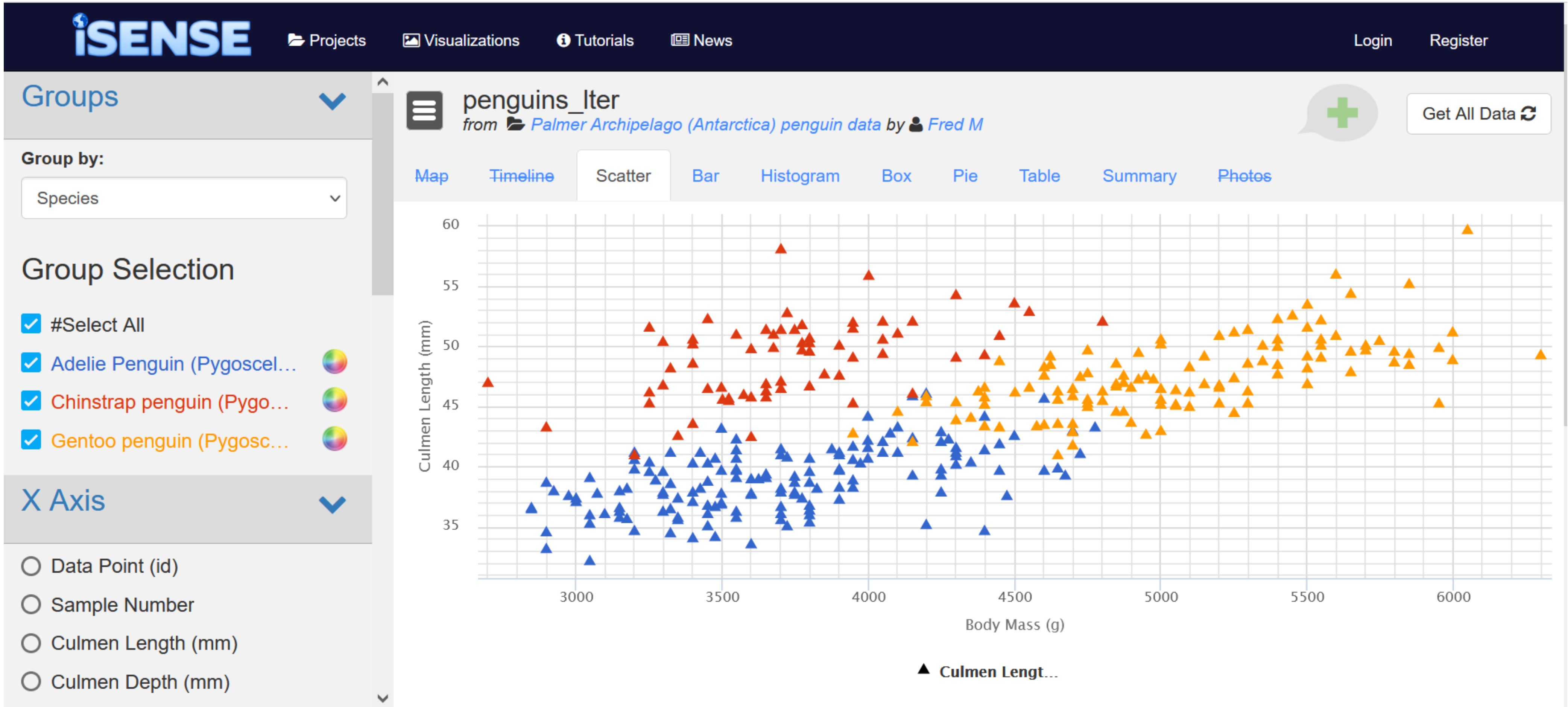




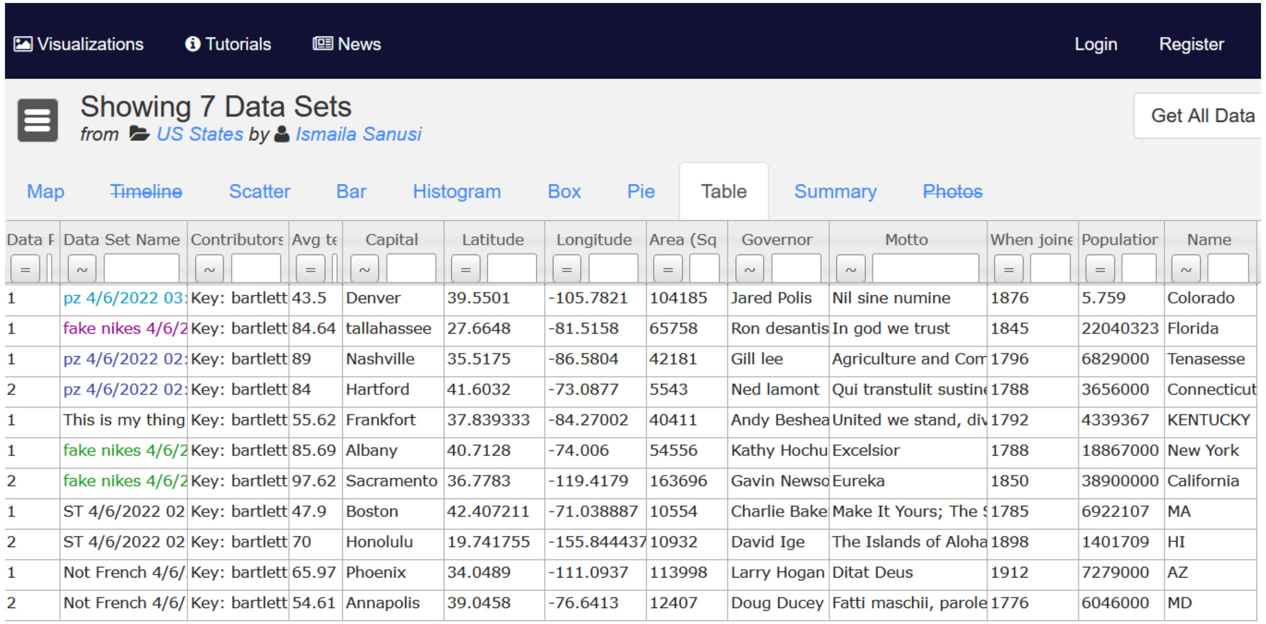
Worksheet completed by the students to understand k-NN concepts



https://isenseproject.org/



iSENSE visualization of penguin culmen length as a function of body mass for three different species



Visualizations Tutorials News Login Register

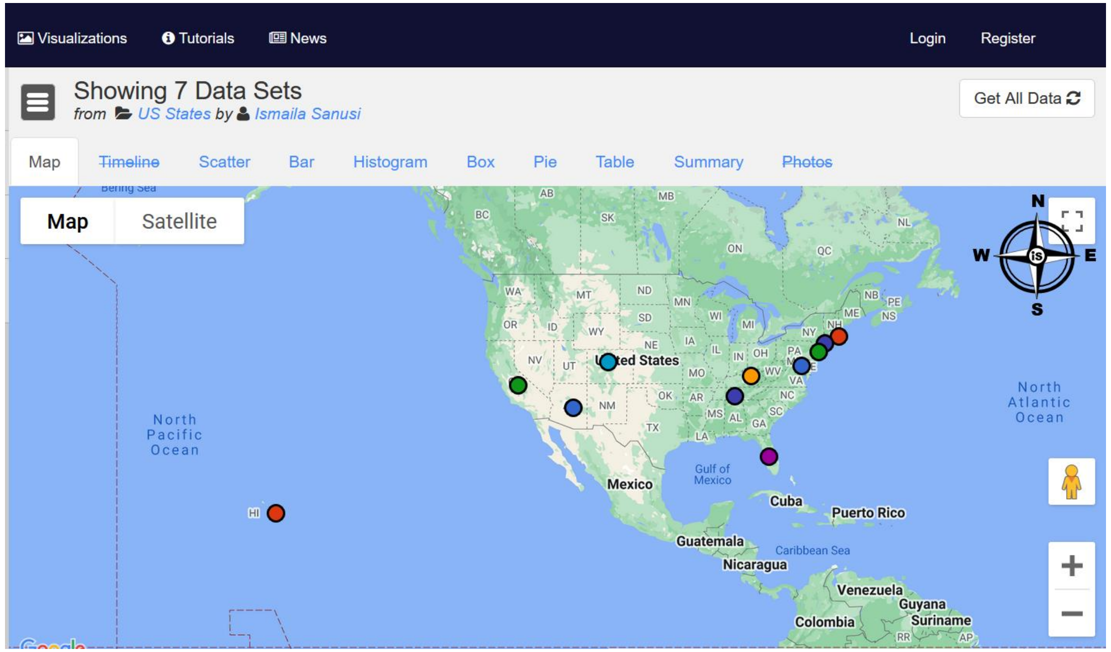
Showing 7 Data Sets
from US States by Ismaila Sanusi

Get All Data

Map Timeline Scatter Bar Histogram Box Pie Table Summary Photos

Data F	Data Set Name	Contributors	Avg te	Capital	Latitude	Longitude	Area (Sq	Governor	Motto	When joine	Populatior	Name
1	pz 4/6/2022 03:	Key: bartlett	43.5	Denver	39.5501	-105.7821	104185	Jared Polis	Nil sine numine	1876	5.759	Colorado
1	fake nikes 4/6/2	Key: bartlett	84.64	tallahassee	27.6648	-81.5158	65758	Ron desantis	In god we trust	1845	22040323	Florida
1	pz 4/6/2022 02:	Key: bartlett	89	Nashville	35.5175	-86.5804	42181	Gill lee	Agriculture and Com	1796	6829000	Tenasesse
2	pz 4/6/2022 02:	Key: bartlett	84	Hartford	41.6032	-73.0877	5543	Ned lamont	Qui transtulit sustine	1788	3656000	Connecticut
1	This is my thing	Key: bartlett	55.62	Frankfort	37.839333	-84.27002	40411	Andy Beshea	United we stand, div	1792	4339367	KENTUCKY
1	fake nikes 4/6/2	Key: bartlett	85.69	Albany	40.7128	-74.006	54556	Kathy Hochu	Excelsior	1788	18867000	New York
2	fake nikes 4/6/2	Key: bartlett	97.62	Sacramento	36.7783	-119.4179	163696	Gavin Newso	Eureka	1850	38900000	California
1	ST 4/6/2022 02	Key: bartlett	47.9	Boston	42.407211	-71.038887	10554	Charlie Bake	Make It Yours; The !	1785	6922107	MA
2	ST 4/6/2022 02	Key: bartlett	70	Honolulu	19.741755	-155.844437	10932	David Ige	The Islands of Aloha	1898	1401709	HI
1	Not French 4/6/	Key: bartlett	65.97	Phoenix	34.0489	-111.0937	113998	Larry Hogan	Ditat Deus	1912	7279000	AZ
2	Not French 4/6/	Key: bartlett	54.61	Annapolis	39.0458	-76.6413	12407	Doug Ducey	Fatti maschii, parole	1776	6046000	MD

U.S. State dataset created by students



Map visualization created based on the U.S. State dataset

How working with data can further support children's understanding of AI.

Since the development of AI application entails the knowledge of how data is prepared and handed, mere reading and inferring from data may not be sufficient to understand and create an AI system.

“The best model is only as good as the data it learns from.”

- Curtis Northcutt

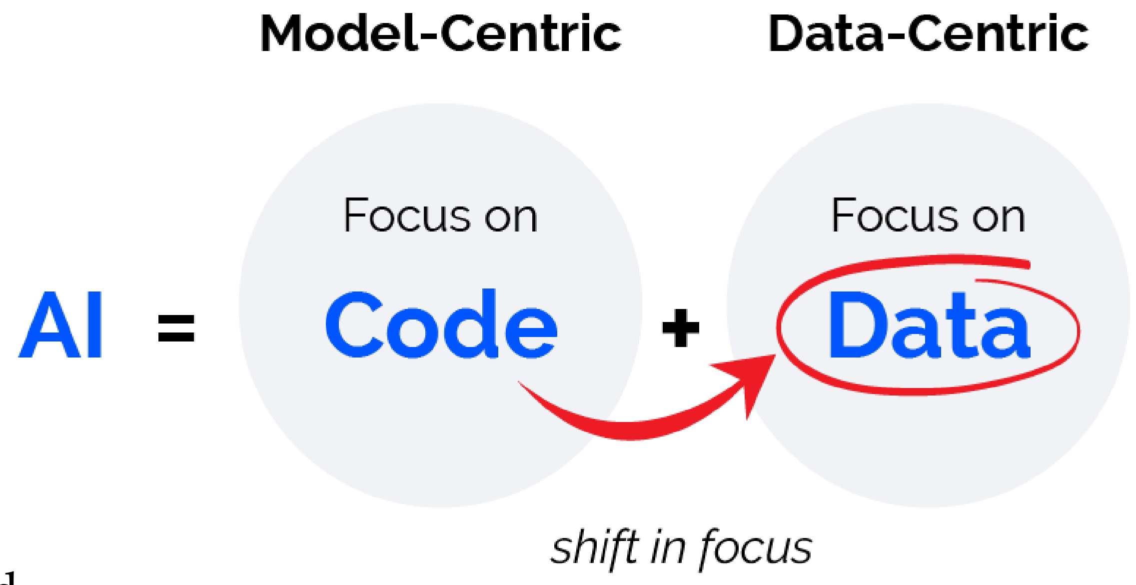
Cleanlab

Data-Centric vs Model-Centric

Data-Centric AI is the discipline of systematically engineering the data used to build an AI system.

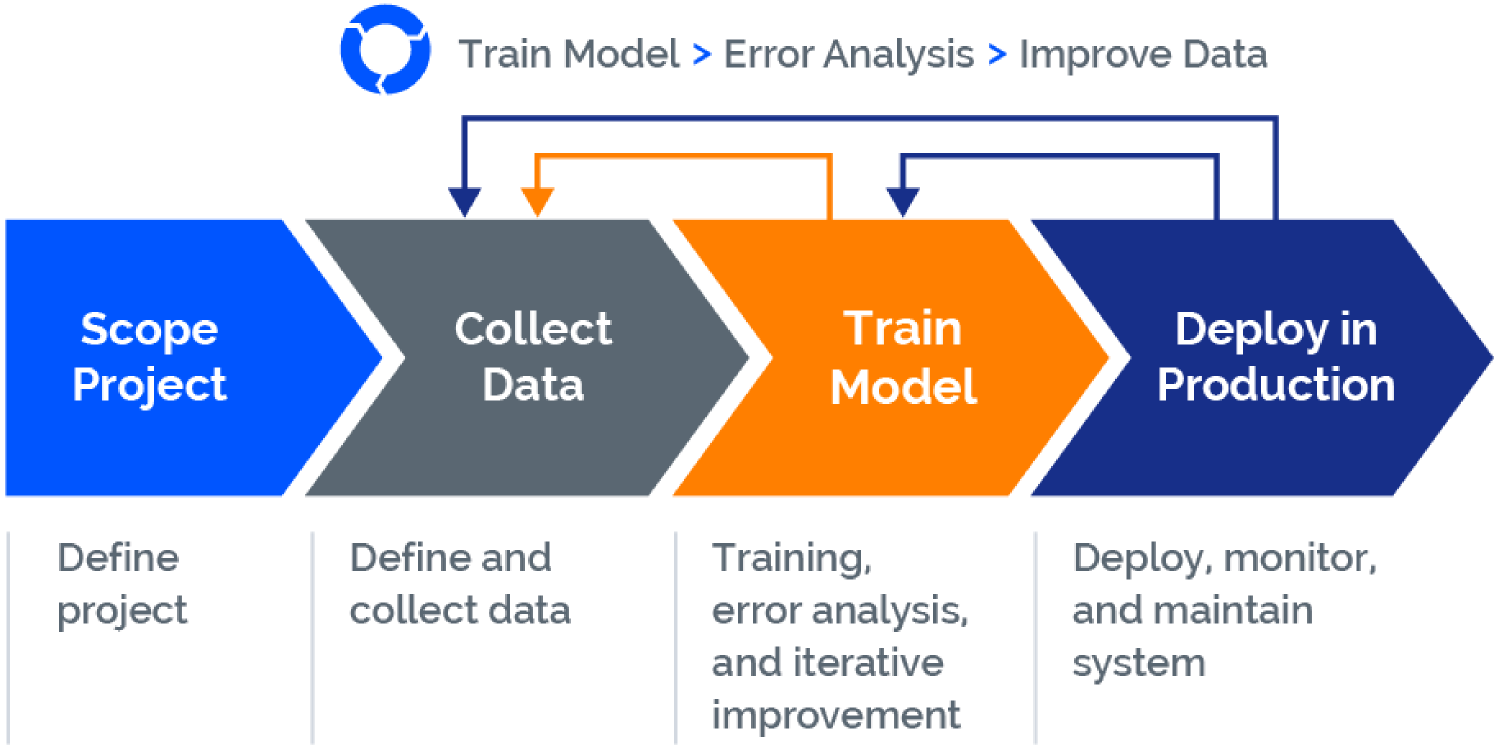
OpenAI has ‘open’ly stated that one of the biggest issues with Dall-E and GPT-3 is errors in the data and labels used during training.

It’s not the model, it’s the data! - Curtis Northcutt



https://landing.ai/data-centric-ai/

Data-Centric AI



https://landing.ai/data-centric-ai/

Some observations

Tasks involved - Classification

- Classification

  - audio or text classifier,
  - facial expression recognition,
  - face detection, object or speech recognition etc.

➢ Discussion centers on algorithms

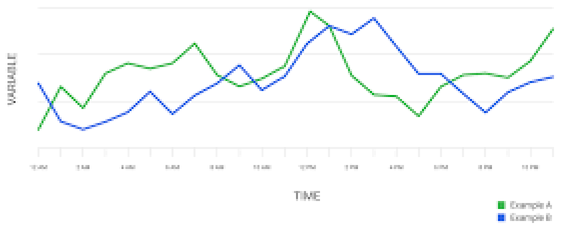
➢ How do we prepare and process the data?

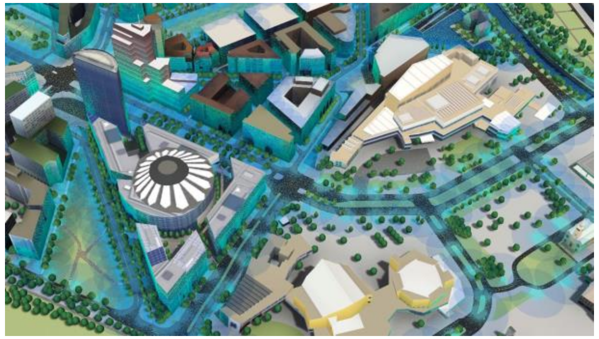
Modality of data

Modality of the data

- (*audio*, *text*, video, image, time series, Geodata, etc.)

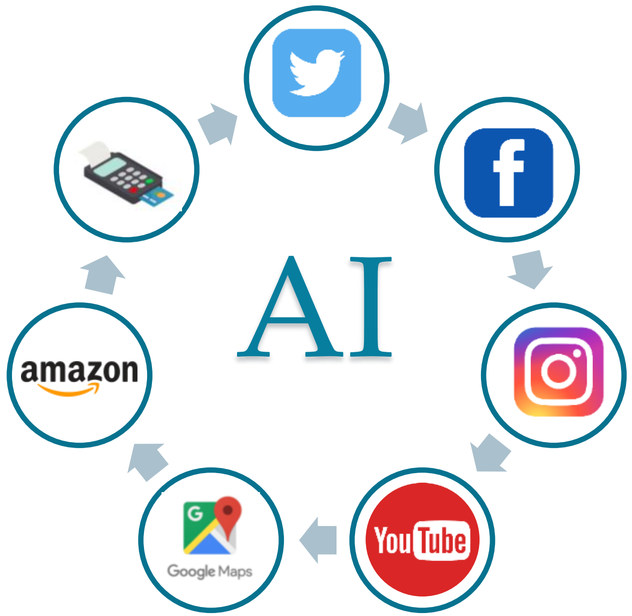
- Heavy dependence on image, audio & text





Data Collection

- Digital family photographs.
- Surveillance videos, tweets.
- Legislative documents.
- Event logs from computer systems.
- Sensor readings over time.
- Any other information in digital form.



What are the steps involved in data collection for an AI system? This may be important to support students understanding of how AI operates.

Data Modeling in AI

Data labeling

Label quality

Labeling strategy

Cleaning of data
(e.g. image or text)

- In many real-world ML applications, the dataset is not fixed!
- Errorneous data
- “Data and label issues plague the most-used AI tech. e.g. Dall-E, ChatGPT” - Curtis Northcutt.

MIT Technology Review

Featured Topics Newsletters Events Podcasts SIGN IN SUBSCRIBE

POLICY

When algorithms mess up, the nearest human gets the blame

A look at historical case studies shows us how we handle the liability of automated systems.

By Karen Hao May 28, 2019

An image showing the aftermath of a self-driving car accident, with an uber vehicle on its side

https://www.technologyreview.com/2019/05/28/65748/ai-algorithms-liability-human-blame/

Training, validation, and test data sets

- The need to further stress the value of training and test datasets
  - data modalities
- Understanding how to engineer data to build better AI systems

Conclusion

Some ongoing works

- Exploring empirical literature on how students understand data and the kind of tasks involved which influence students' understanding.

- Investigating the implications of data comprehension research for students' understanding of machine learning and for teaching machine learning concepts.

Some ongoing works

Integrating ML and data science in school subjects

What we are doing:

- Exploring a few machine learning software tools & approaches
- Brainstorming connections between machine learning to curriculum topics
- Prototyping idea for bringing machine learning to students
- Teachers share experiences with colleagues and program researchers

We believe integration of AI across the curriculum is the key and that data fluency should be more incorporated into AI lessons.

UNIVERSITY OF
EASTERN FINLAND

Thank You

Ismaila Temitayo Sanusi

ismaila.sanusi@uef.fi