Instructional Lesson Plan

Machine Learning Applications in the High School Biology Classroom

Steven Clemens, Michael Stanley, Dr. Emmanuel Salifu, Dr. Andreas Spanias

LESSON DETAILS

Subject Area(s): Biology, Microbiology, Coding, Computer Science, Mathematics

Focus Grade Level: 9-10th grade

Grade Level Range: 9-12th grade

RESEARCH BACKGROUND

Machine learning is not exclusively taught at all high schools. According to the code.org (2021) state of computer science education and availability report, a significant gap exists between the demand for computer science education and the availability of programs in US high schools. In the 2019-2020 school year, only 45% of high schools in the US offered computer science courses. This represents a significant increase from just five years earlier, when only 25% of high schools offered such courses (code.org). This lesson will seek to connect a general education high school biology classroom with basic machine learning applications that can be used as a supplemental lesson with microbiology and plating lessons, or as a standalone lesson.

LESSON SUMMARY

Teachers have two options to complete this lesson, option 1 includes completing the microbiology lab collection and plating. Option 2 would be for teachers who do not have access to the proper equipment and cannot complete the plating in a sterile environment. Option 2 will have the students use the photographs taken in the laboratory at Arizona State of the plated yeast and use them to run enumerations through the OpenCFU program. For either option the OpenCFU program will be used to give colony counts. Classification algorithms will be used (pre-coded) to predict the strain of yeast used when plated as well as the dilution of that stain. The methods used in this lesson are not overtly difficult and are simpler ML tasks but complex enough to get students curious about other ML applications that are out there. The overall goal would be to spark an interest in finding out more about machine learning and the growing field of data scientists.

MATERIALS AND EQUIPMENT

Option 1 and Option 2: Chromebooks or laptop computers for each student or at least 1 per every 2 students with access to the OpenCFU program https://opencfu.sourceforge.net/

Option 2: Access to the high resolution photos of plated yeast from Arizona State

Option 1 Only: Microbiology and Plating

Equipment:

Autoclave or pressure cooker for sterilization

Bunsen burner or alcohol lamp for sterilization

Incubator for growing yeast

Microscope for observing yeast morphology and structure.

Pipettes for transferring liquids.

Petri dishes for plating yeast

Refrigerator for storing yeast cultures.

Sterile loop or wire for streaking yeast cultures

Materials:

Agar plates for yeast growth

Distilled water for making media and dilutions.

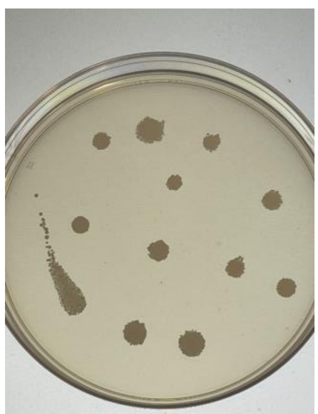
Ethanol or other disinfectant for surface sterilization

Nutrient media for yeast growth

Yeast strains for culturing and plating

ATTACHMENTS

Sample image of Petri Dish



EDUCATIONAL STANDARDS

K-12 TEACHERS

NGSS:

HS-ETS1-3: Evaluate a solution to a complex real-world problem based on prioritized criteria and trade-offs that account for a range of constraints, including cost, safety, reliability, and aesthetics as well as possible social, cultural, and environmental impacts.

HS-LS2-2: Use mathematical representations to support and revise explanations based on evidence about factors affecting biodiversity and populations in ecosystems of different scales.

HS-ETS1-2: Design a solution to a complex real-world problem by breaking it down into smaller, more manageable problems that can be solved through engineering.

ITEEA Standard 7 - Students will develop an understanding of the influence of technology on history.

LEARNING OBJECTIVES

Option 1: Each student will be able to create 3 sterilized agar plates and successfully stain 3 plates with yeast at differing dilutions.

Option 1 and 2: Each student will be able to run their images of petri dishes through the OpenCFU program and determine the correct colony forming units from each dish

Each student will be able to transpose the data from the OpenCFU program into the Jupyter notebook for classification.

Each student will be able to write a 500 word research paper on the benefits of using machine learning algorithms for use in biological and microbiology applications or other everyday applications.

VOCABULARY

Open source- refers to software or hardware that is made available to the public for use, modification, and distribution, with the source code being accessible and editable by anyone.

Classification- the act of categorizing or organizing items or data into groups or classes based on shared characteristics or attributes.

Enumeration- the process of counting or listing items or data, often used to create a comprehensive inventory.

Colony forming units (CFUs)- a term used in microbiology to quantify the number of viable microorganisms in a sample, often through the counting of colonies that arise from single cells.

Microbiology- the branch of biology that deals with the study of microorganisms, including bacteria, viruses, fungi, and protozoa.

Plating- a laboratory technique used to isolate and grow colonies of microorganisms on solid nutrient media.

Agar- a gelatinous substance derived from algae that is commonly used as a nutrient medium for culturing microorganisms in the laboratory.

Autoclave- a device used to sterilize laboratory equipment and materials using high pressure and temperature steam.

Sterilization- the process of eliminating all forms of life, including microorganisms and viruses, from a given environment or material.

Decision tree- a graphical representation of a series of decisions and their potential consequences, often used in decision-making and problem-solving.

Anaconda- a distribution of the Python programming language that includes many commonly used packages and tools for scientific computing and data analysis.

Jupyter Notebooks- an open-source web application that allows users to create and share documents that contain live code, equations, visualizations, and narrative text. It is commonly used in data analysis and scientific research.

Lesson Procedure

For option 1:

1. Sterilize all equipment and materials that will come into contact with yeast, including the autoclave, Bunsen burner, petri dishes, pipettes, and sterile loop or wire.
2. Prepare the nutrient media by mixing the appropriate amount of agar and other ingredients with distilled water and autoclaving.
3. Allow the media to cool to around 45-50°C, and then add the yeast strains to the media using sterile technique.
4. Pour the media into sterilized petri dishes and let them solidify.
5. Use a sterile loop or wire to streak the yeast on the surface of the agar in a zigzag pattern.
6. Incubate the plates at the appropriate temperature for the yeast strain being cultured.
7. Once colonies have grown (after 24 hours), use a digital camera and place to petri dishes on a solid background for pictures. Take photos at the 24, 48, and 72 intervals. DO NOT OPEN THE PETRI DISHES or they will become contaminated.
8. Properly discard cultured media after the 72-hour data collection period
9. Remember to follow proper safety procedures when working with microorganisms, including wearing gloves, lab coat, and eye protection, and disposing of all biohazardous waste appropriately.

For Option 1 and 2 (after the data collection or the use of the photographs as part of this lesson)

1. Provide access to the photos taken during the data collection or to the photos that are given as part of this lesson.
2. Upload the photos to the program OpenCFU. Allow students to manipulate some of the parameters such as color identification and surface area. Give specific parameters to students to create controls for the data collection from OpenCFU

3. Once data is available from the OpenCFU program, download data as a CSV file.
4. Ensure that the CSV filename matches the filename in the Jupyter notebook file (attached to this lesson and preprogrammed)
5. Students will run Jupyter notebook and data will be classified. Students should be able to note that the algorithm was able to classify the yeast by the parameters that were being tested (length of incubation, amount of colonies present, type of yeast).
6. Students will create a research paper on the benefits of using the machine learning algorithms in applications such as the one tested in this lab activity and other applications in the real world. Students

INTRODUCTION/MOTIVATION

Teacher will provide this picture to students and ask the students, how many colonies do you see?

Students will provide their best guess to the teacher and the class. (teachers could also use an actual petri dish with colonies pre-grown)

The correct answer is 462… now what if you had to do this all day everyday? What if your job was to sit there and hand count every single colony you see? How many do you think you could get through in a day? Would it even be accurate? (get students discussing this scenario).

Teacher will then go over that there are programs that can help microbiologists and other people who work with lab specimens count accurately the amount of colonies present in a sample or to help them classify by specific parameters viruses, bacteria, and other microscopic organisms using algorithms and machine learning. \

Teacher will inform students that they are going to be data scientists for the next two class periods. They will run there own lab in which the data they calculate will be used to classify a new strand of yeast that is taking over the lab.

LEARNING ACTIVITIES/STRATEGIES

1. Active participation: One of the most important learning strategies for hands-on lab activities is active participation. This means engaging with the lab activity, asking questions, and taking notes as you go along. Actively participating in the activity helps students stay focused and engaged, which can enhance their understanding of the concepts being explored.
2. Reflection: Reflection is an important strategy for learning from hands-on lab activities. After students complete the lab activities (either option 1 or 2), they will reflect on what was learned, what went well, and what they could do differently next time. Reflection helps solidify their learning. Students are also writing a research paper at the end, which will take the concepts they learning and put them into a reflective piece.
3. Collaboration: Working with others can help students to gain different perspectives and approaches to the lab. Students can also identify areas where they may need additional help or support.
4. Visual aids: Using visual aids, such as diagrams or flowcharts, can be a helpful strategy for learning from hands-on lab activities. Visual aids can help to simplify complex concepts, making them easier to understand and remember.
5. Practice and repetition: Practicing and repeating hands-on lab activities can help to reinforce your learning and improve your skills.

CLOSURE

Teacher will lead a discussion with students on what it was like to be a data scientist for the last two class periods . Teacher will solicit feedback from the students on what they think the future will look like as computers become more powerful. Students will then use their data, classifications, and skills learned throughout the lab and the data collection process to write a 500-word research paper that describes why using tools like the OpenCFU program can help improve society and functioning within the scientific community or in real world applications. Students will be learning and applying multiple skills throughout these lessons. Advanced computer skills and critical thinking skills will be evident throughout the lesson.

ASSESSMENT

FORMATIVE ASSESSMENT

Formative assessments will be conducted throughout the lesson

For option 1:

- Successful agar plate
- Growth of yeast
- Lab safety check
- Lab skills throughout the lab

For option 2:

- OpenCFU pictures upload
- Download of the data file
- Upload of the data to the correct file path
- Classification data from the Jupyter notebook

SUMMATIVE ASSESSMENT

Completed research paper based on the use of programs such as OpenCFU that describes why using tools like the OpenCFU program can help improve society and functioning within the scientific community or in real world applications. Students will be graded on the following rubric:

5 - Excellent

- The research paper is clear, concise, and well-organized.
- The introduction provides a clear thesis statement and context for the research.
- The body of the paper is well-developed and effectively uses evidence to support the argument.
- The conclusion effectively summarizes the key findings and draws a clear conclusion.
- The paper demonstrates a thorough understanding of the topic and provides original insights or ideas.
- The paper is well-written, with few errors in grammar, spelling, and punctuation.

4 - Good

- The research paper is clear and well-organized but may lack some depth or complexity.
- The introduction provides a clear thesis statement but may not provide as much context for the research.
- The body of the paper is developed and uses evidence to support the argument but may not be as well-structured or cohesive as a 5-point paper.
- The conclusion effectively summarizes the key findings but may not be as clear or concise as a 5-point paper.
- The paper demonstrates a good understanding of the topic and provides some original insights or ideas.
- The paper is well-written, with some errors in grammar, spelling, and punctuation.

3 - Satisfactory

- The research paper is generally clear but may lack organization or coherence.
- The introduction provides a thesis statement but may not provide much context or background for the research.
- The body of the paper provides some evidence to support the argument but may not be well-developed or use evidence effectively.
- The conclusion summarizes the key findings but may not be as clear or well-constructed as a 4- or 5-point paper.
- The paper demonstrates a satisfactory understanding of the topic but may not provide original insights or ideas.
- The paper has some errors in grammar, spelling, and punctuation.

2 - Needs Improvement

- The research paper is unclear or poorly organized.
- The introduction may not provide a clear thesis statement or context for the research.
- The body of the paper may not provide enough evidence or support for the argument or may be poorly organized.
- The conclusion may not effectively summarize the key findings.
- The paper demonstrates a limited understanding of the topic and may not provide original insights or ideas.
- The paper has several errors in grammar, spelling, and punctuation.

1 - Unsatisfactory

- The research paper is unclear, disorganized, or does not meet the minimum length requirement.
- The introduction does not provide a clear thesis statement or context for the research.
- The body of the paper does not provide adequate evidence or support for the argument or may not be well-organized or developed.
- The conclusion does not effectively summarize the key findings.
- The paper demonstrates a poor understanding of the topic and does not provide original insights or ideas.
- The paper has numerous errors in grammar, spelling, and punctuation.

CONTRIBUTORS

INDIVIDUALS

Steven Clemens – Biological Sciences Teacher; Dysart Unified School District

Michael Stanley- Industry partner and mentor for summer RET session

Dr. Emmanual Salifu- Arizona State University

Dr. Andreas Spanias – Arizona State University

Huda Clemens – K-4 STEAM teacher; Dysart Unified School District

REFERENCES

Code.org. "2021 State of Computer Science Education: Policy and Implementation," 2021. [Online]. Available: https://advocacy.code.org/stateofcs. [Accessed: Aug. 20, 2022].

Q. Geissmann, "OpenCFU, a New Free and Open-Source Software to Count Cell Colonies and Other Circular Objects," in PLoS ONE, vol. 8, no. 2, pp. e54072, 2013. doi: 10.1371/journal.pone.0054072.

SUPPORTING PROGRAM

RET Site: Sensor, Signal and Information Processing Algorithms and Software
Sensor, Signal and Information Processing Center (SenSIP), in partnership with Arizona State University and the National Science Foundation.

FUNDING ACKNOWLEDGEMENTS

This project is funded by the National Science Foundation (NSF) Award 1953745. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect those of the NSF.