Neural Networks and Computer Vision with MNIST

Total Number of Lessons: [number of lessons]
Total Time for Module: [number of minutes]

SECTION 1: MODULE OVERVIEW AND CURRICULAR CONNECTIONS

- **Module overview**: [High level description of the module, approximately one paragraph.]
  - Data is all around us, and it can be powerfully harnessed to answer interesting questions about the world that we live in. But how can we get computers to help us with this? How can computers make inferences from thousands of bits of information, in order to answer interesting questions? In this module, students will be introduced to artificial neural networks, and train one to answer an interesting image classification question.

- **Module goal**: [At the end of this module, SWBAT…; try to make it one big goal]
  - SWBAT train an ANN using the MNIST data set and use it to classify handwritten numbers provided by students.
  - SWBAT gather their own dataset and train another ANN to classify images from their dataset.

- **Module scope and rationale**: [Will this be a multi-day activity or a full module? Why should someone want to do this module?]
  - This module is a multi-day activity, which draws on knowledge of Python programming. It makes a connection between big data, computer vision, and machine learning, and introduces students to the world of machine learning and data science.

- **Connection to standards**: [List standards that will be addressed by the module. Should include Indiana standards; NGSS and CCSS are strongly recommended; CSTA could also be useful for the Computing RET.]
  - CSTA
    - 3A-DA-12 Create computational models that represent the relationships among different elements of data collected from a phenomenon or process.
    - 3B-DA-05 Use data analysis tools and techniques to identify patterns in data representing complex systems.
    - 3B-DA-06 Select data collection tools and techniques to generate data sets that support a claim or communicate information.
  - Indiana
    - CSI-3.2: Assess the use of algorithms to provide a solution.
    - CSI-6 Domain - documentation
    - CSI-7.2: Explain the privacy concerns related to the collection and generation of data through implicit and explicit processes.

- CSI-8.4: Investigate post-secondary training opportunities and industry certifications that are available.

- **RET materials/ideas to be leveraged**: [Describe what elements of your summer experience you will use as part of the module for students.]
  - This module will leverage the experience we gained in python programming and using the MNIST dataset. It harnesses an understanding of how ANNs work, including building the layers, training, and testing ANNs.

- **Prior knowledge needed for module**: [Describe what would be helpful for students to know before beginning the module.]
  - Before attempting this module, students need a working knowledge of python programming. They need to have some idea about the difficulty of using classical programming techniques to classify images or make inferences about big data.

SECTION 2: OVERVIEW OF MODULE FRAMEWORK

- **Real-world context**: [Describe in 1-2 paragraphs]

  - **Possible Lesson Ideas**: [What are some ideas for a lesson or two focused on the real-world context and framing the project for students?]
  - We can start with a broad discussion of facial recognition in everyday life. Ask students about where they have seen facial recognition used. Discuss how you might write an algorithm to compare two faces from what they know about programming so far. This will be difficult for students. Ask an easier question, how would they write a program to distinguish between images of cats and dogs. (use Dr. Niemier's slides here), this question will still be difficult, so ask an easier question, how could they write a program that will distinguish between 10 handwritten digit images.
  - Brainstorm with students as many applications of image recognition as they can. This can also lead to an ethical discussion about privacy, security, data gathering, and big data.

- **Background STEM content**: [Describe in 1-2 paragraphs]

  - **Possible Lesson Ideas**: [What are some ideas for a lesson or two focused on the background STEM content?]
  - Introductory lesson: Introduce the idea of an ANN, explain on a high level how they work. Discuss matrices, matrix dot notation, and how an ANN works (use Dr. Niemier's slides). Discuss the differences between conventional programming

and using Neural Networks. Discuss the use of CNNs in facial recognition and other applications.
- Second lesson: Discuss privacy concerns of widespread use of facial recognition. Students brainstorm uses of image recognition in daily life, and try to break down how they might work.
- Later lesson: Discuss implications of bias within facial recognition algorithms - bring in research question from this summer (why are COTS Facial Recognition software better at recognizing men than women?)

- **Final Project**: [Describe in 1-2 paragraphs]

  - **Possible Project Ideas**: [What is the goal of the project? What is the timeline for completing the project? At what point would students work in groups or individually?]
  - After introducing the topics and having some discussions about ANNs and how they work, I will demonstrate the power of ML using this website, or this one. Then I will guide students in small groups(2-3 students) to build and train an ANN on the MNIST dataset. They will use jupyter notebooks to install tensorflow and keras, and then train and run the ANN on MNIST data set to understand how it works.
    - Extension - students can try different amounts/types of layers in the network to see if the predictions get better or worse.
    - Extension - students can use a set number of epochs for training, and then tweak other variables to try to get the best accuracy.
  - For the final project, students will get into larger groups (4-5 students), and brainstorm an interesting binary classification question. Then they will gather data in the form of images, to build their own data set to answer a binary classification question. Once the dataset is built, they will use it to train an ANN and then test on part of the dataset.
    - Students will need significant guidance and support for this part of the final project, especially as it relates to gathering, formatting, sizing, and assembling images.
    - An alternative would be for students to search out another dataset that already exists and train an ANN on this dataset using the example of the MNIST dataset ANN.

  - **Possible Project Deliverables**: [What would students produce at the end of the project? Would it be shared in some way? How?]
    - Students would end up with a trained network, and a python script that trains and runs an ANN to perform some classification problem. They will also have a dataset of images.

        - Extension: If students want to take this further, we could host their trained network on a website. They could have users upload photos and the algorithm classify them.

Questions:

- How can I diversify this so that students can modify it/do their own thing?
    - Could have students decide on their own problem, then gather data and design their own network to classify something - could tie it to nature/biology/another class.
        - Could do this as a class, or as two big groups?
        - Just for fun, Friday project
    - Train a network to differentiate between different types of music?!
    - Once its functional - can tweak different parameters
    - Given a set number of epochs, and a set amount of training data, can you tweak and adjust for better training? Different layers?
    - Train a set on own data - two choice classification
        - Can I get students to write their own data to test?
- What is the best way to break this down to their level?
- Time of year?
- What level of students should I do this with?
    - Probably CSP
- Is it okay if this doesn't explicitly tie to any standards? But I see it as hugely valuable.
    - Yeah - this is really high level stuff.

SECTION 3: MODULE SEQUENCING AND ASSESSMENT

- **Description of sequenced learning objectives**: [How do learning objectives build from lesson to lesson to meet the module goal?]
    - Learning objectives build from theoretical understanding of ANNs to understandings of the applications of ANNs in everyday life, to practical application in their own lives and the ability to build and train an ANN.

- **Description of formative assessment approaches**: [When/what techniques will you use to assess student progress towards objectives during, and at the end of, each class period?]
    - I will assess student progress by 5 finger poll questions throughout the lesson asking students to rate their understanding on a scale of 1-5.
    - I will look at student work as I move through the room to assess how well they are understanding.

- **Description of summative assessment approaches**: [How will you assess the final project and/or performance assessment? If you are using a rubric, what criteria will you examine?]
    - I will assess the success of students' final projects based on:

- the success of their ANNs to correctly identify new data
- their written explanation of their work
- their verbal explanations of how ANNs work

SECTION 4: FINAL COMMENTS AND ATTACHED FILES

- **Recommendations for implementation**: [Describe any "safety tips" or advice you have for other educators who might implement this module.]
  - Plan ahead of time for how to set up and run the appropriate python environment with the installations you will need.
  - Plan ahead of time how you will collect and assess student work
  - Carefully group students, as this may be beyond some students current ability level/interest.

- **List of attached files**: [should include any lesson plans and handouts relevant to the Module, if you have them ready]

  - Overview video of NNs
  - 06-17-19 - ML Overview: Slides by Dr. Niemier to draw from for discussion of how to classify images (slides 20 - 50), and the MNIST dataset (slides 51 - 59).
  - Quick Engaging AI Websites
  - MNIST Recognition Tutorial
  - Teachable Machine
  - Jupyter Tutorial
  - Second Jupyter Python Tutorial