



2020 weSTEM Conference Workshop

Smart ZooKeeper

Objective

The objective of the workshop is to introduce students to the concepts and ideas that build the foundation of machine learning. Students will build a 'Smart ZooKeeper' classification model using a Zoo Animal Classification Dataset, consisting of 16 features. They will be programming alongside the instructors on a pre-coded Jupyter Notebook which will illustrate the basics of coding and machine learning.

At a Glance

Instructor's Name: Prajna Soni and Alia Albastaki

E-mail: prajna.soni@nyu.edu and alia.albastaki@nyu.edu

Time needed: 1h

Preparation time: 20 minutes

Number of Participants: N/A

Materials List

- A fully charged laptop per participant
- A white board to illustrate/draw concepts out for students
- A box of blue and red balls

Preparation and Setup

1. Ensure each laptop has working internet connection and is fully charged (or plugged in)
2. Open Google Chrome
3. Start up the Kaggle Kernel and upload the [zoo.csv](#) and [class.csv](#) file to each hub (separate instruction file will be provided for people setting up)
4. Upload [Intro to ML.ipynb](#) to the hubs
5. Set up AV for powerpoint and connect Presenter's laptop to projector

Workshop Runthrough

1. Group Introductions (5 minutes)
 - a. Everyone says their name and their favourite book
 - b. Facilitators introduce themselves and why they are interested in STEM
2. Introduction to Machine Learning and Artificial Intelligence (15 minutes)
 - a. Ask students what their understanding of ML and AI are
 - b. How do you differentiate between a cat and a dog as a human being? How do you teach a machine how to do that?
 - c. Define ML and AI and provide some examples and basic use cases
 - d. Introduction into classifiers
3. Explore the zoo dataset (10 minutes)
 - a. Work through the notebook to open the zoo.csv dataset
 - b. Facilitators define features, instance, training set and testing set
 - c. Split the dataset into test and train
4. Train the model on the training dataset (10 minutes)
 - a. Students follow the notebook alongside facilitators to run the train model on the training data
 - b. Facilitators explain what classification means and draw a classification tree for the zoo keeping problem
5. Run the model on the test data and explain results (10 minutes)
 - a. Students follow the notebook alongside facilitators to run the test model on the trained model
 - b. Facilitators debrief what the accuracy means
6. Real Life Applications (5 minutes)

- a. Facilitators provide examples of where this is used in real life
7. Questions? (5 minutes)

Conclusion

Come up with 2-3 discussion questions that reflect on the workshop process?

What does machine learning now mean to you?

Do you feel more confident in your ability to learn other machine learning algorithms?

Can you think of a few other examples where you would use as classifier?

Be Prepared!

What will you do in the event that you run out of time?

If we run out of time we will skip over some parts of the code and show the main results, so that there is enough time for discussion and explanation of the most important parts.

What if you finish with too much time remaining?

We might show different ways of trying to improve model accuracy (changing model parameters, etc), or introduce the idea of multi-classifiers.

There will be a weSTEM Conference Team member at every workshop to help in the event of AV problems or any other things that come up.

Thanks for your time in creating and leading a workshop. We look forward to seeing you at the conference on March 13, 2020.