Kickstart Workshop

1 hour 30 minutes

The kickstart program is an early career development program for women and gender-diverse individuals. As part of the program we are developing a workshop to act as a preview to the world of AI and ML which will showcase some popular tools and machine learning applications, built for both technical and non-technical folks. The goal of this workshop is to gauge interest of the participants in AI/ML and to encourage them to gain further knowledge and dive deeper into the ML industry.

Link: https://www.amii.ca/training/kickstart/

Workshop Outline & Times:

1) Describe what is machine learning in detail (5 mins)
   a) Provides ability to predict outcomes based on historical data
   b) Advanced statistical models amplified with computer science
   c) Types of machine learning problems
      i) Supervised
         (1) Classification
         (2) Regression
      ii) Unsupervised
   d) Some ML tools
      i) Deep learning
         (1) Provide some examples
            (a) Image
            (b) NLP
   e) Discuss the most basic form of machine learning is a linear regression which everyone has done at middle/high school

2) Talk about data (5 mins)
   a) Types of data (unstructured vs structured)
   b) Sources of data
   c) Discuss the importance of good quality data

3) Python (5 mins)
   a) Most popular language for machine learning
   b) Mention ease of learning and usefulness of the language
   c) Quick overview of Python
      i) Snippets of Python code
         (1) 'hello world'
         (2) Loading libraries

(3) Simple math
- d) Talk about libraries
  - i) pandas
  - ii) NumPy
  - iii) scikit-learn

4) Jupyter lab setup (10-15 mins)
  - a) Install anaconda
  - b) pip install pandas, numpy
  - c) Quick start guide for Jupyter notebooks
  - d) Some students will have bugs and errors (help them fix it and install properly)

– BREAK (5 minutes) –

5) Showcase machine learning applications (50 mins)
  - a) Linear Regression (10 mins)
    - i) Quickly describe how linear regression works
    - ii) Install libraries
    - iii) Discuss data and its X and Y variables
    - iv) Quick overview of code
    - v) Run code and show results (accuracy is simply metric for the purposes of this workshop)

  - b) k-means clustering (10 mins)
    - i) Quickly describe this is a unsupervised model with no intended y variable
    - ii) Discuss why this is useful sometimes
    - iii) Install any libraries
    - iv) Load data
    - v) Quick overview of the code
    - vi) Display Visualization of the clusters formed with centroid being visible(this is important)
  - c) ML application with regular neural network (15 mins)
    - i) Maybe recommender systems
    - ii) Talk about how deep learning is a different type of machine learning
    - iii) Talk about how neural nets are inspired by the brain (neurons and neural network)
    - iv)
  - d) Image Classification (15 mins)
    - i) Explain how images are a 2D matrix of pixel values
    - ii) Address the need for a CNN rather than regular deep net
    - iii) Install libraries
    - iv) Mnist dataset to classify numbers
    - v) Feed data into network and output results

6) Ending the Workshop (5 Mins)
   a) Mention these are basics of machine learning
   b) Field is growing very fast and newer technologies are coming out in the industry
   c) Encourage participants to dive deeper into the topic
   d) Mention Amii and ML tech

Notes:

- Module 1, 2 and 3 will be presented with a slide deck
- Module 4 and 5 will be done on a Jupyter notebook
- Module 6 on slide deck
- Break after module 5.a
- Jupyter lab sets them up better for further learning and feels more more engaging if setup during a workshop, as compared to using just collab