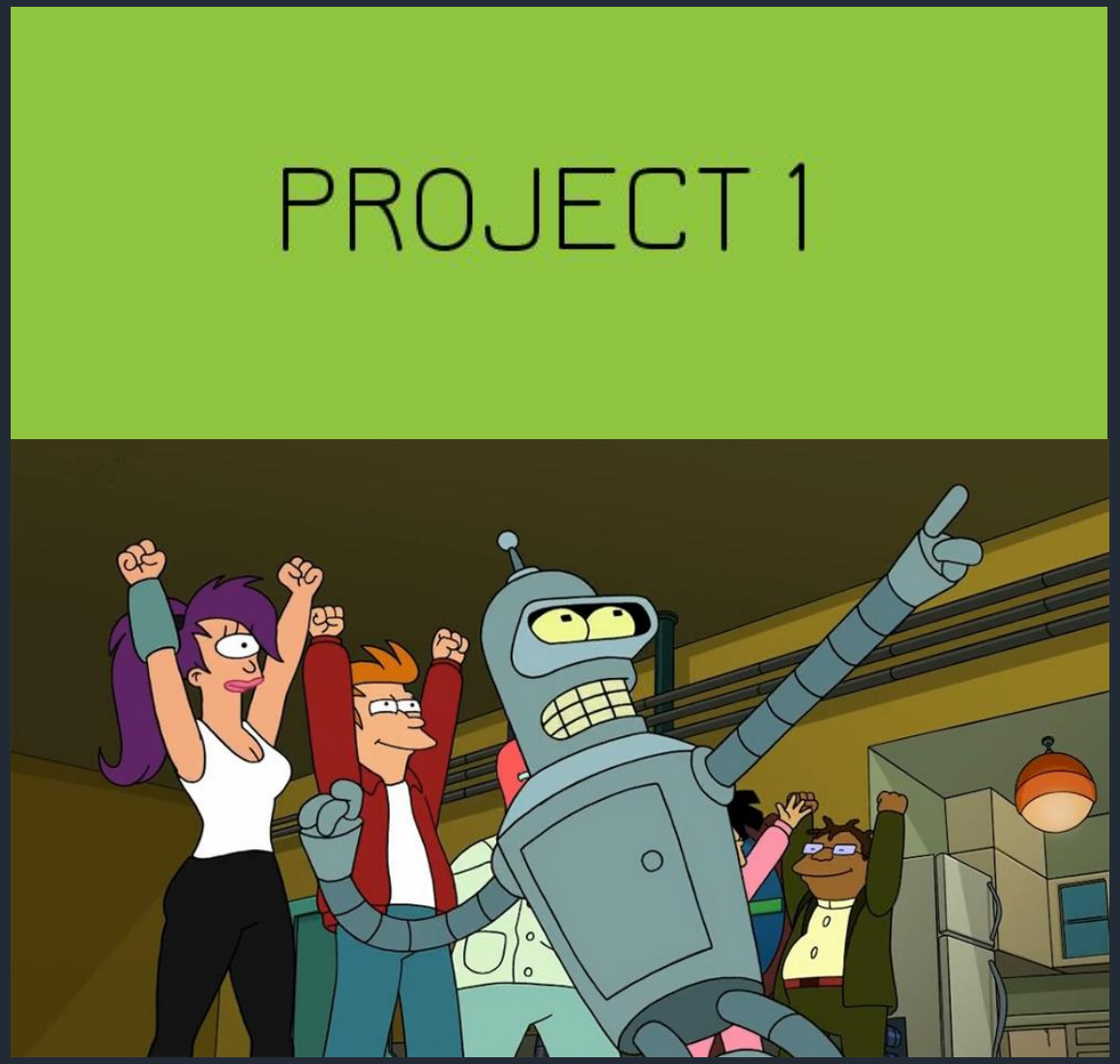
MechMath Autumn 2019:

Introduction to Machine Learning using Python

Project1: It has arrived!



Project1: How to start

- Download Project1_Guideline.pdf from our Slack-channel
- Download Project1 distribution code “resources_sentiment_analysis.tar.gz” from our Slack-channel
- Untar (unzip) the code
- Download Project1_Tasks.pdf from our Slack-channel
- Open main.py, test.py, utils.py, project1.py
- Try to run test.py!
- Try to understand what is going on and how are these files connected

Project1: First step with the Hinge Loss function

main.py | test.py | utils.py | project1.py

Match Case | Words | Regex

```
#pragma: coderesponse template
def hinge_loss_single(feature_vector, label, theta, theta_0):
    """
    Finds the hinge loss on a single data point given specific classification
    parameters.

    Args:
        feature_vector - A numpy array describing the given data point.
        label - A real valued number, the correct classification of the data
            point.
        theta - A numpy array describing the linear classifier.
        theta_0 - A real valued number representing the offset parameter.


    Returns: A real number representing the hinge loss associated with the
    given data point and parameters.
    """
```

What would you like to get

```
C:\Users\Dell\PycharmProjects\untitled\venv\Scripts\python.exe
PASS Import projectl
PASS Get order
PASS Hinge loss single
PASS Hinge loss full
PASS Perceptron single update
PASS Perceptron
PASS Average perceptron
PASS Pegasos single update
PASS Pegasos
PASS Classify
PASS Classifier accuracy
PASS Bag of words
PASS Bag of words  stopwords removed
PASS Extract bow feature vectors
PASS Extract bow feature vectors : correct non binary features

Process finished with exit code 0
```

"correct code"
(test.py)

+

```
C:\Users\Dell\PycharmProjects\untitled\venv\Scripts\python.exe "C:/U
theta for Perceptron is 3.9173999999999918, 4.164000000000001
theta_0 for Perceptron is -8.0
theta for Average Perceptron is 3.47826049999999, 3.611060999999974
theta_0 for Average Perceptron is -6.373
theta for Pegasos is 0.7346463119064065, 0.6300224592973831
theta_0 for Pegasos is -1.2195071848898564
```

"correct value"
(main.py)

Project1: Keep in mind

- **Deadline:** 14:00, 04.12.2019
- Be sure that you untared (“unzipped”) all the files to the same directory!
- Other than PyCharm IDE’s at your risk.
- It is not obligatory to understand EVERYTHING in the distribution code. But at least try to.
- Correct solutions could be the first pieces of your github portfolio.
- Try to check the “**correct value**” with your classmates. Feel free to discuss Project1 with the colleagues…
- …but don’t steal their code!

How to avoid fear of programming ML in Python

JUST DO IT.

Thank you for attention!