**USING NEURAL NETWORK TO PREDICT VIEWER’S FEELING TOWARDS A YOUTUBER’S STYLE THROUGH FOOD DESCRIPTION**

**Answer to required questions (detailed report below):**

* The model: neural network.
* The architecture: The model has 1 hidden layer, and both activation functions for the hidden layer and the output layer is sigmoid.
* Reason to choose:
  + Based on theory: all 3 models: linear regression, logistic regression and neural network is usable on the problem since the output is discrete (classification) and ranges from 1 to 5 (regression).
  + Based on the data: the covariance matrix informs a non-linear relation between the inputs and the output, which eliminate the use of non-polynomial feature linear regression.
  + Based on practice: Neural network is more flexible than multiclass logistic regression due to having to only build a single model.
* Result:
  + Started out with a seemingly underfit model. After increasing the complexity, the accuracy doesn’t change (only 45%).

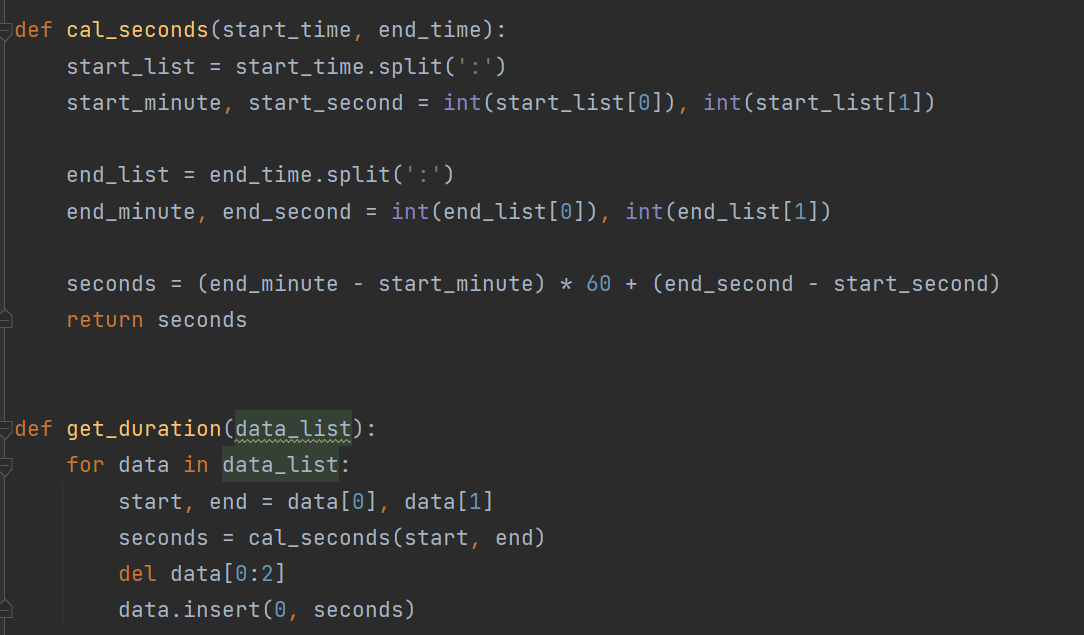
🡪 The problem itself is hard to solve by algorithm.

* + After testing with a linear regression model, additional conclusion is added: a neural network might not be needed for this problem. A simple linear regression models works as well.

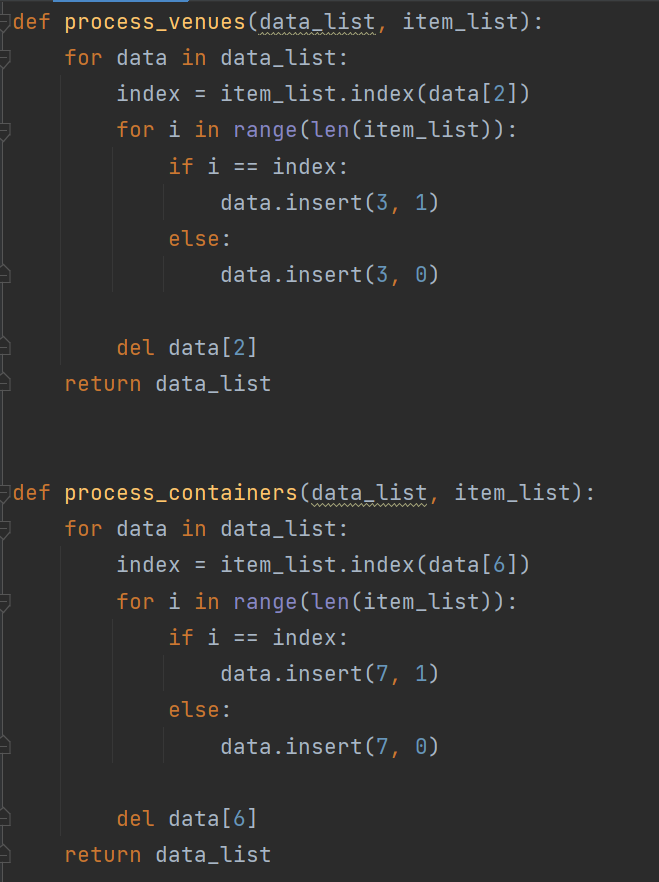
**Detailed report**

**Data:**

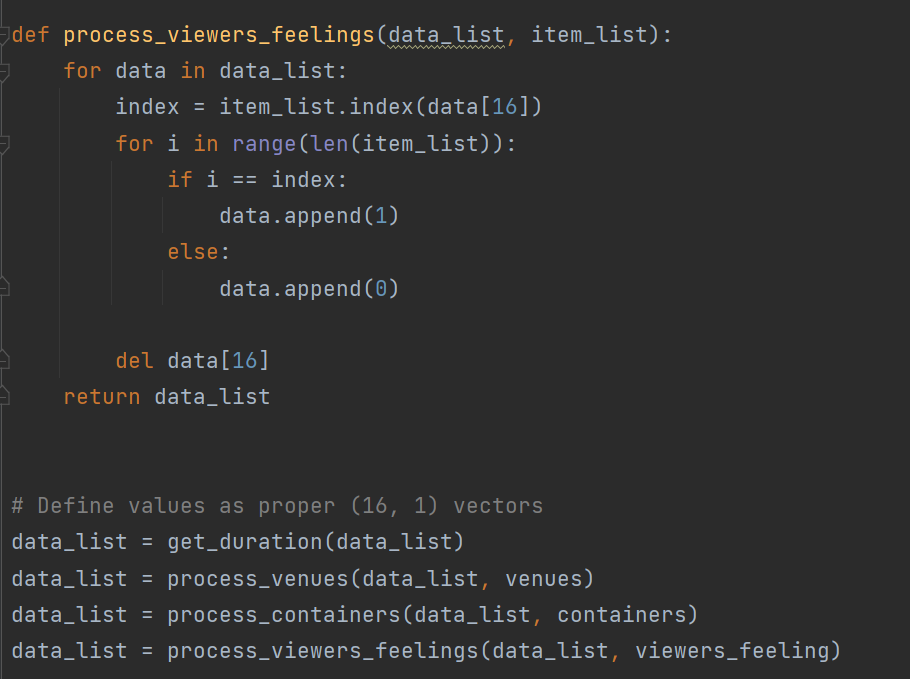
* Description:
  + There are 5 inputs and 1 output:
    - Inputs: start time, end time, number of ingredients, venue, container.
    - Output: viewer’s feeling of youtuber’s style.
  + Venue and container only take certain values mention later in data cleaning.
* Cleaning: takes place on Excel using a filter.
  + Start time: only takes in a single time value, minute and second is separated by a colon.
  + End time: only takes in a single time value, minute and second is separated by a colon.
  + Number of ingredients: takes in a non-negative integer.
  + Venue: only takes in 'fine restaurant', 'casual restaurant', 'street restaurant' and 'other'.
  + Container: only takes in 'bag', 'bottle', 'bowl', 'box', 'glass', 'hands-on', 'plate', 'pot', 'tray' and 'other'.
  + After cleaning, 3622 examples remain.
* Processing:
  + Duration is calculated by subtracting start time and end time.



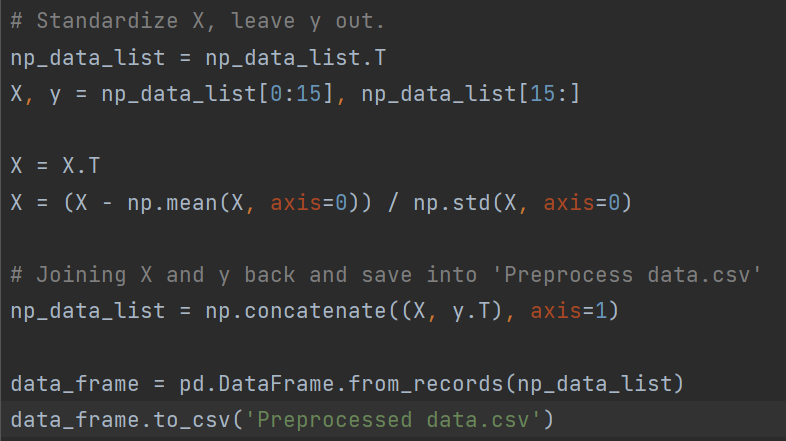
* + Venues and containers are processed into vectors.



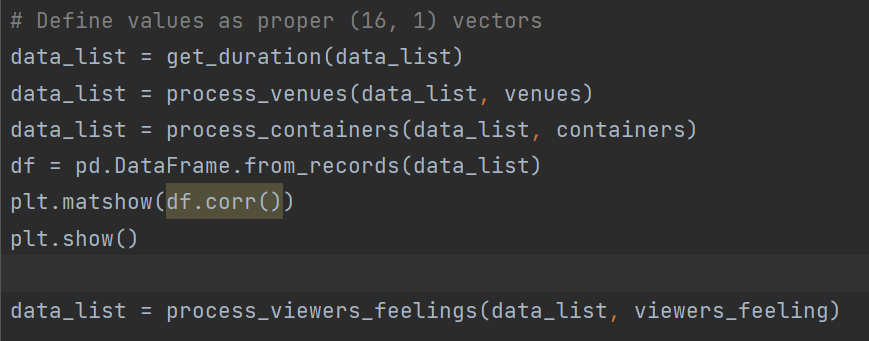
* + The same goes for viewer’s feeling of youtuber’s style.

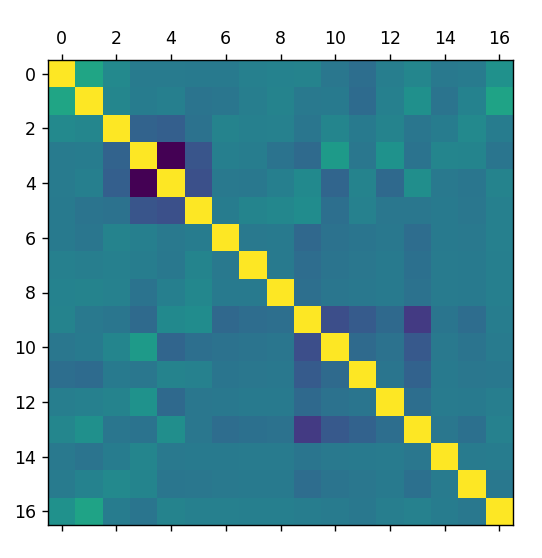


* + After that, there are 15 inputs and 5 outputs.
  + The data is now, normalized for the inputs and saved.



* Visualization:
  + Before standardizing the inputs and vectorization of the output, a covariance matrix of the data is calculated and plotted.

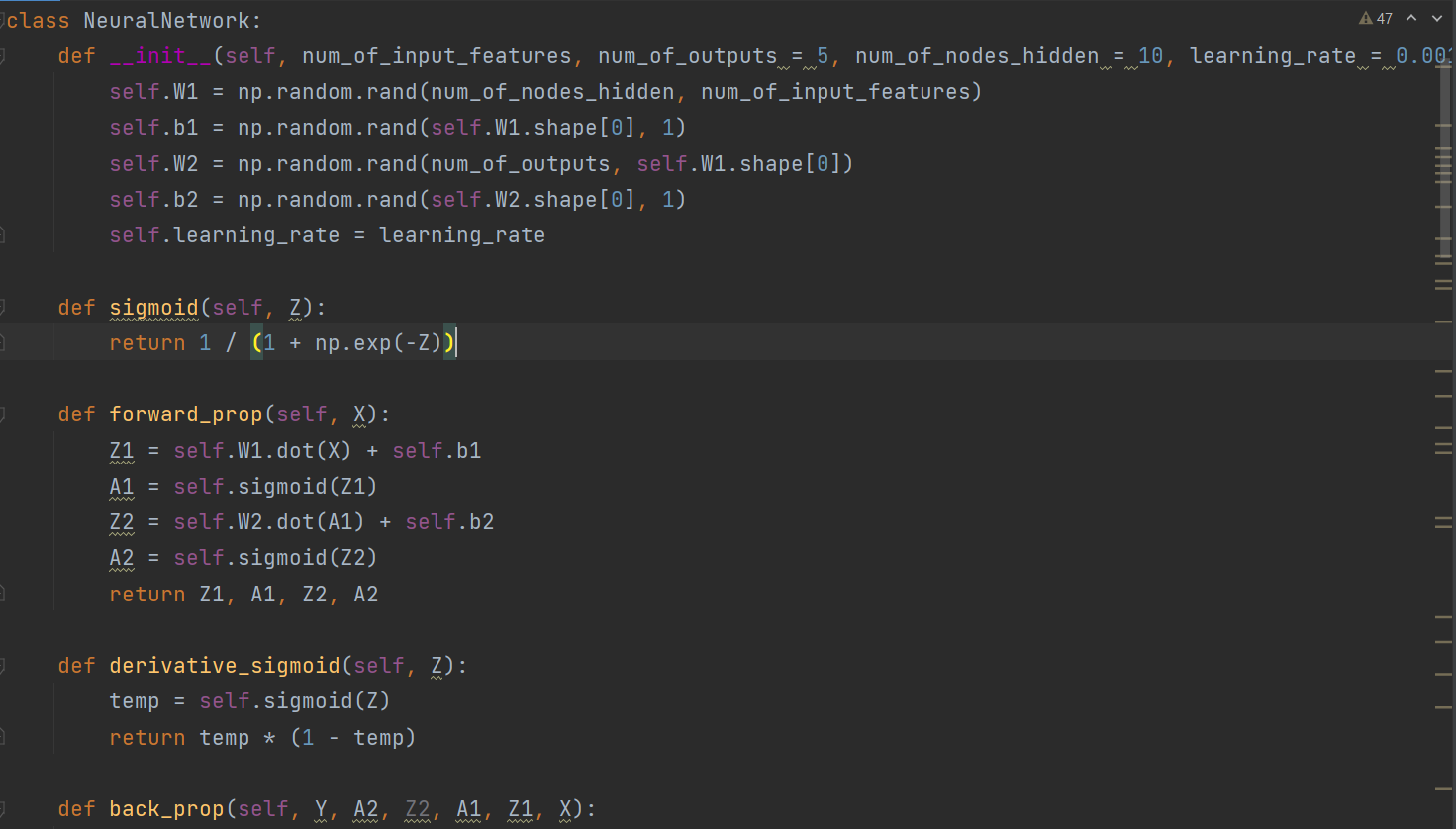




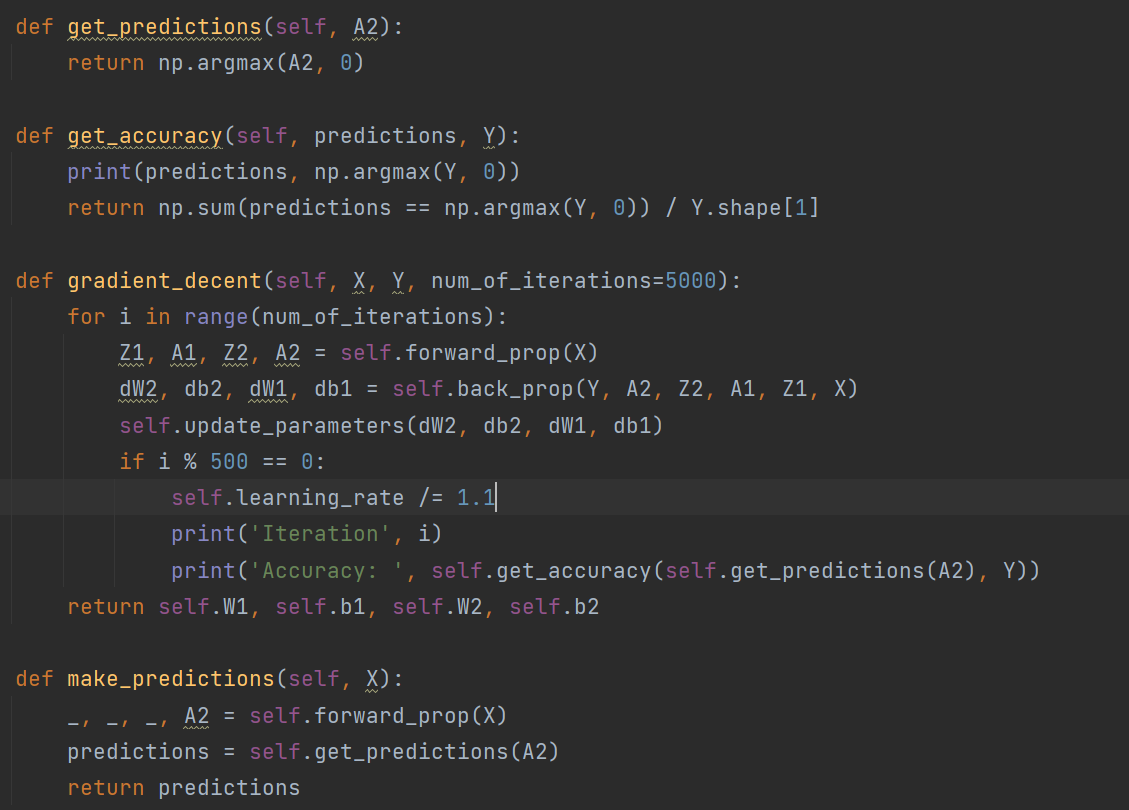
* + There are little to no correlation between the inputs and the output. This model’s accuracy is expected to be low, if not overfits.

**Model:** Neural Network.

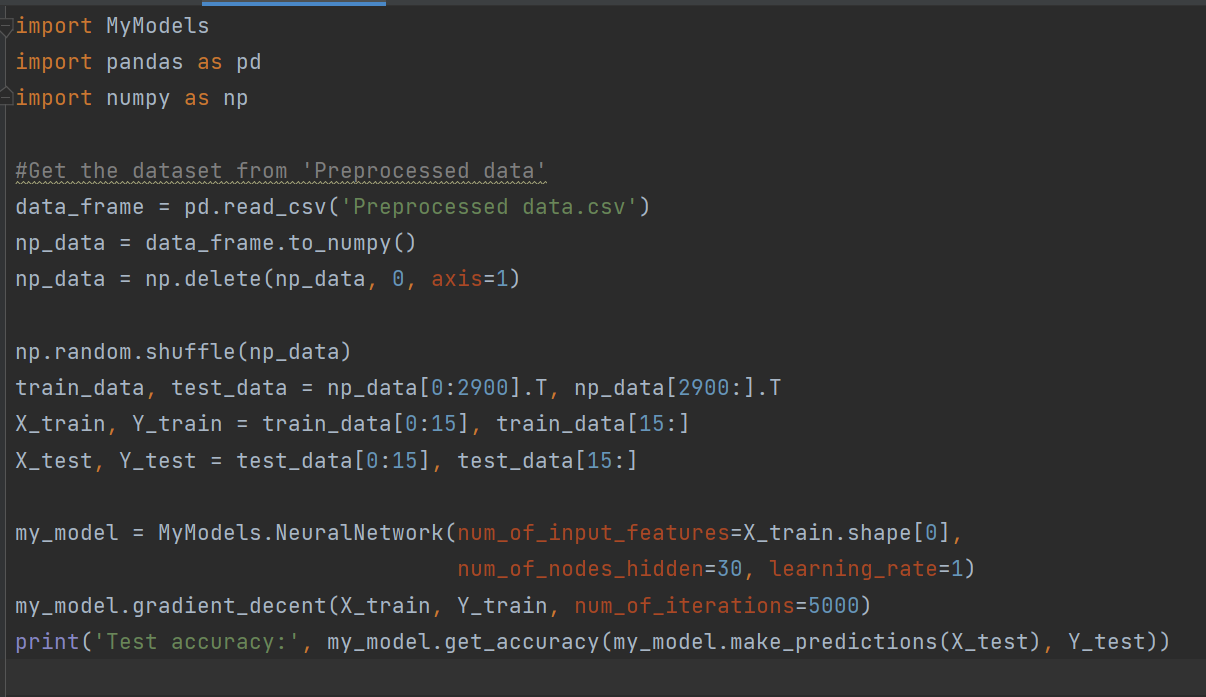
* Reason to choose: the covariance matrix informs a non-linear relation between the inputs and the output. That leaves multiclass logistic regression and neural network. Neural network is more flexible than multiclass logistic regression due to having to only build a single model.
* Architecture:
  + The model has 1 hidden layer, and both activation functions for the hidden layer and the output layer is sigmoid.



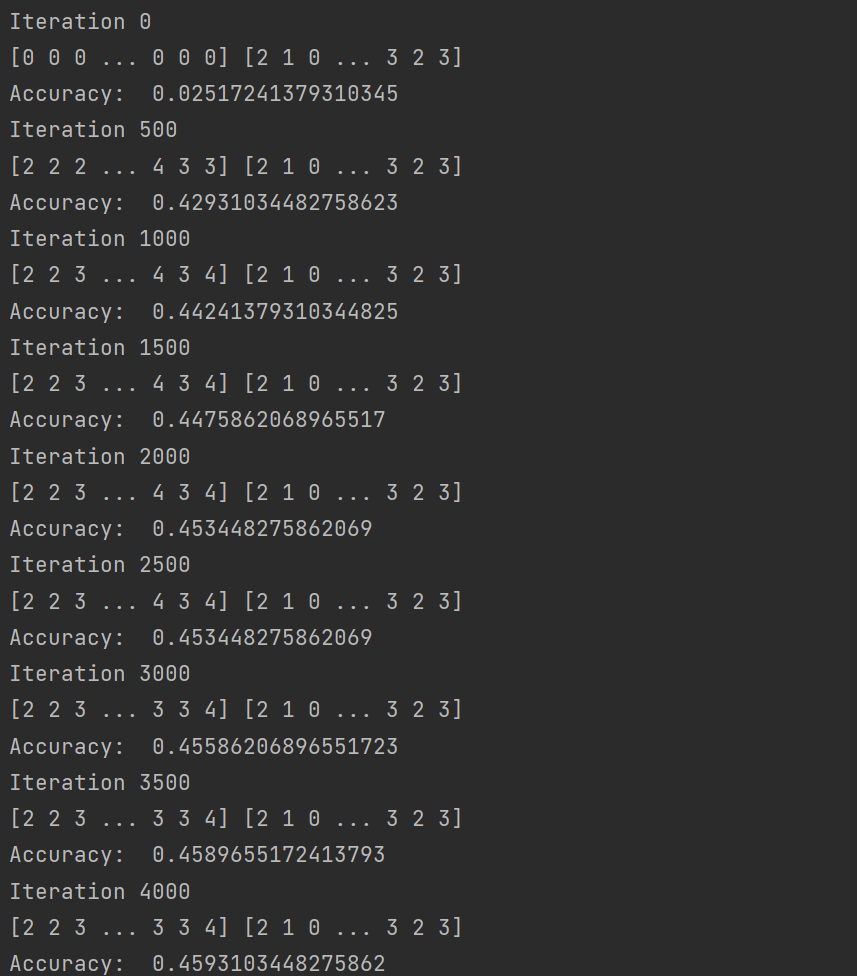


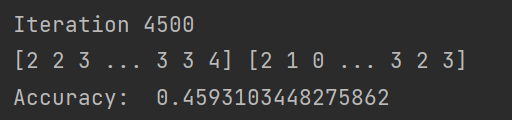


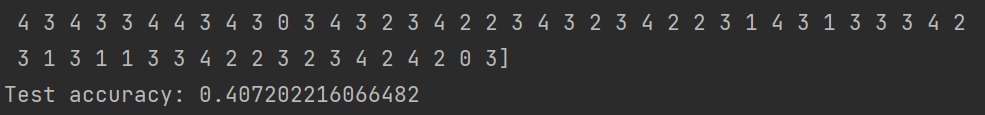
* Training:
  + The data is splitted into train set and test set. (2900/722 ~ 80/20).
  + The model will be trained for 5000 iterations, with 30 nodes in the hidden layer and the learning rate is set to 1, decreasing by 1.1 every 500 iterations.



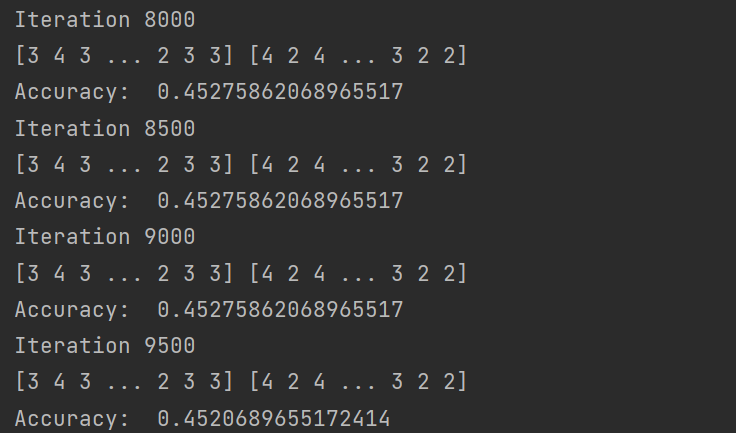
* Result:
  + Not very promising, the accuracy stucks at around 45% and the test accuracy is 40.7%.

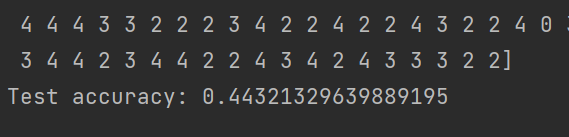






* + Based on the result, it can be that the model is underfitting. The solution is to raise the complexity of the model: 300 hidden nodes, 10000 iterations.





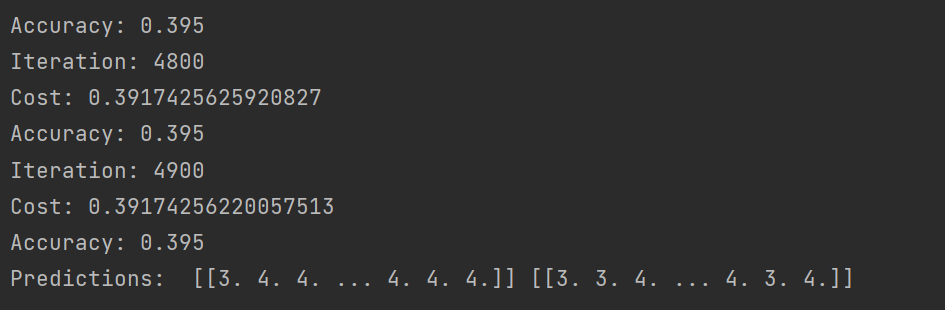
* + Even after increasing the complexity of the model, the result doesn’t change much. The model has done its best. Lying at 45% accuracy.

**Conclusion:**

* Predicting a viewer’s feeling of youtuber’s style through food description is after all, depends on human’s emotions. It is difficult to offer a model to predict precisely the rating for a feeling.

**Additionals:**

* Since the output of this problem is ranged for 1 to 5, this can be defined as a regression problem.
* A simple linear regression model has been built and with 5000 iterations and learning rate of 0.01, it yields almost identical result.



* In additional conclusion, a neural network might not be needed for this problem. A simple linear regression models works as well.