Tutorial 1: Building an NLC Application Part 1

Tutorial and walkthrough

The Application

- Working through one of the COMM190 Demo Applications;
- **Idea:** Use the NLC service to classify Amazon reviews on a particular product
  - Decided to use a Keurig [K55] as the Amazon.com product has 3,645 product reviews
    https://www.amazon.com/Keurig-K55-Single-Programmable-Coffee/dp/B018UQ5AMS/
  - We will train the classifier on custom classes, and then see how well it generalizes to reviews which it has not seen before.

Plan to Build

Part 1:

1. Generate Credentials on Bluemix;
2. Collecting a Sample;
3. Determine the Base-Classes for Classifier;
4. Collect the Initial Training Set;
5. Try the web interface for Training;
6. Test the Classifier Online;

Part 2:

7. Run the API Call in Python;
8. Generate a CF Starter Pack Flask Application;
9. Build Out MVP Flask Application on the Above API Calls;
10. Push to Bluemix Hosting

Step 1

Generate Bluemix Credentials

Generate Bluemix Credentials

- Visit https://bluemix.net/ and sign-in

- Click "Catalog" from the top menu bar

- In the search box search for "Natural Language Classifier" and click on the resulting AI Service Tile.

- Fill in a "Service name"
  - **Comm190-nlc**

- Press "Create" and you will have your Bluemix service registered
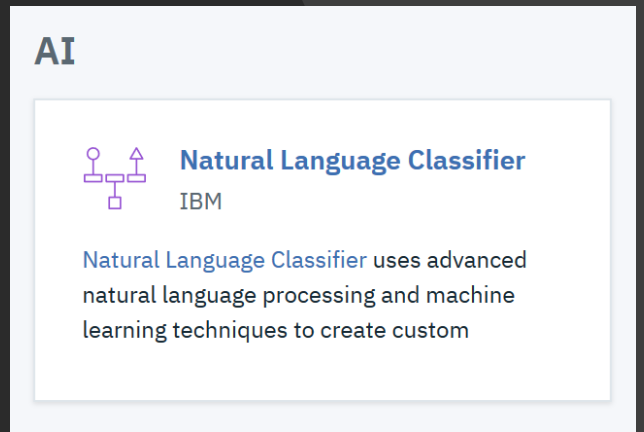
Step 2

Collecting a Sample

The Extracted Reviews

- Using a scraping tool, we extracted reviews from Amazon for the Keurig K55. We now have a CSV file with the extracted reviews for the product (we will post this for you to use on the course website).

- It is helpful to take a look at the content of a few of these reviews to get a sense of what we are dealing with.

Comment
loved it
works fine, very happy with it.
Thank you very muchPackage arrived a day earlyMuch appreciated
Great for just me! I love it!
This was my second defective unit. I would like to return it
reservoir is not big enough compared to the cusinart
Container came broken ......
Keurigs suck
AV djcfik

The machine started to have problems after only a couple of weeks. I thought this was a one time occurrence, but it continues to fail. It works for a couple of days, then it just doesn't make any coffee. The coffee just drips out slowly, as if no pressure is used to "push" the water through the k cup. The k cup is heavy and full of water after removing it. It's not clogged or anything, I rinse it regularly. A total disappointment, I will ever buy any Keurig product ever again. I have this machine for less than 2 months and can't return anymore.

Updates are located at the bottom of this review.***Original Review***Up until recently, I had never been much of a coffee drinker except for the occasional seasonal flavored latte drinks but I was having trouble keeping up with my very active twin toddlers. We already had a standard coffee maker but I didn't want to brew a whole pot of coffee at once and single-serve coffee makers seem to be very popular so I figured it couldn't hurt to give it a try. I had seen several different Keurigs from the basic ones to very fancy ones but wanted to purchase one that both fulfilled my needs to brew a single cup of coffee and wasn't too expensive so I opted to buy the Keurig K55 Coffee Maker in black.When the Keurig arrived, I eagerly opened up the box to get it set up and start using it! I found it fairly simple to set up since there is a Use and Care Guide included and it's easy to follow the step-by-step instructions. I'd say the setup only takes a few minutes and then you are ready to brew your first cup of coffee. There are lights on the brewer that illuminate and walk you through the brewing process but once you are used to it, it becomes second nature and now, I don't even have to think about it, which is good in the early morning when I'm not thinking straight and need my coffee!Brewing: To brew a cup of coffee is very simple. You'll need to turn on the brewer by pushing the power button. The power button and heating lights will illuminate and the Keurig will make some noises while it heats up. When the three cup sizes illuminate in blue, you know it's time to brew your coffee. Set your coffee mug on top of the drip tray and lift the handle for the K-cup pod holder. Place a K-cup pod in the holder and lower the handle completely until the three cup sizes are blinking. The three sizes are: small cup (5.25 ounces), small mug (7.25 ounces) and large mug (9.25 ounces). Select your size and the coffee will be ready in about a minute. If the water reservoir is low on water, the "add water" light will illuminate and the process won't continue until you've filled the reservoir. Once it's stopped dripping, you may remove your mug and enjoy your coffee! If someone else wants to brew a cup of coffee right after yours is done, all they have to do is remove your K-cup pod since they are single-use pods, add theirs and repeat the process.Water Reservoir: The reservoir holds 48 ounces of water so you don't have to fill it up every time you want to brew a new cup of coffee, which is one of the reasons I selected this model. I make about two cups of coffee every day so I end up refilling the reservoir every few days. There are filters to use with the water but I purchase bottled water so I haven't used the filters at this point. When the reservoir needs to be filled, I remove the reservoir, take off the lid, fill it up with water, place the reservoir back in the Keurig and replace the top.Cleaning: I feel this Keurig is fairly simple to clean. The drip tray is removable and I usually wipe it down with a damp paper towel a couple times a week and wash it in water, soapy water occasionally. I also like to wipe down the exterior since the front part of the Keurig gets splatters of coffee on it. You can also clean the pod holder, funnel, needles and water reservoir and lid. Be careful as the two needles used to puncture holes in the K-cup pods are very sharp.Coffee Pods (the best part!): The Keurig K55 Coffee Maker uses the K-cup pods, which you can purchase nearly everywhere and there are many brands and flavors to choose from. You don't need to buy any specific brand with this coffee maker. As long as they are K-cups, they should work and I haven't come across any pods that don't work. I've been having a lot of fun purchasing different flavors and brands since I have several friends and family members who also have Keurigs so it's a lot of fun to swap coffees! I prefer flavored coffees and there are so many out there to try but my favorites are the Green Mountain flavors in Caramel Vanilla Creme and Wild Mountain Blueberry. There are K-cup pods in different price ranges and it's nice to have that option since everyone has a different budget. The nice thing with this brewer is that the flavors don't tend to blend from one cup to the next. For instance, my mom and I have brewed cups of coffee one right after the other (different flavors) and neither of us have noticed the other flavor in our respective cups. I've only brewed one cup of apple cider with this and haven't noticed the previous coffee flavor in the apple cider or the apple cider in my next cup of coffee. In this brewer, you can use all kinds of coffee flavors from flavored to medium and dark roasts to hot chocolate to apple cider. I typically choose the small mug (7.25 ounces) setting since I like the flavor. It isn't too strong or too watered-down for me but others may prefer other settings.Programmable Auto Off Timer: You can set the brewer to turn off two hours after the last cup of coffee has brewed (included in the instructions). I just turn the Keurig on and off as I need it but it's nice to have that option.Noise: Some people have stated that this brewer is noisy. I've found it's the noisiest when it's taking water from the reservoir and it sounds like a loud hum. It's not too loud when it's brewing the cup of coffee. I occasionally get up earlier than my husband and kids and brew a cup of coffee and no one has woken up (I'd say there is about 30 feet between the kitchen and bedrooms with the doors closed). When we have guests and they brew coffee before I wake up, I can't hear it at all.This Keurig is made in China and mine had a packaging date of August 2016, a month before I purchased it. It measures about 12 inches deep (from front to back), nine inches wide (side to side) and 13 1/2 inches tall. I've had several family members use it and my parents and in-laws have found it fairly simple to use after I show them how to brew the coffee since we all have different models. Ever since I started drinking coffee, I've had better success with staying more alert and awake to take care of my twins and buying this Keurig has been well-worth every penny we spent on it! I have no regrets buying it and have been having a lot of fun finding different flavors to try. As with all of my reviews and especially any appliances and electronics reviews, I will periodically update this review, add in any new information I discover and adjust my thoughts and star rating, if needed. If you, or someone you know, is interested in a Keurig, I highly recommend this one. It's a great model that is nice on your wallet and will brew a great cup of coffee as often as you'd like!***Update 04/25/2017 - I've had this Keurig now for almost eight months and it's still working great for us! I typically brew four cups of coffea a day now to keep up with my kids and I've had no trouble with using this Keurig. When we have guests and are switching pods a lot, I never notice a strange taste at all (since we all like different flavors). I clean out the drip tray and wipe the whole machine down at least once a week to keep it looking nice. We only use bottled water in it so I've never had to use the filter. When we go away for a few days, I always wipe everything down and empty the water reservoir so it can air out and dry. I still highly recommend this Keurig if you want a good single-cup brewer that isn't incredibly expensive!

Step 3

Determine the Base Classes

Determine the Base Classes

- The NLC works by classifying snippets of text into defined categories. The categories are defined by you when you implement the classifier.

- Ideally, the categories should comprehensively cover the breadth of possibilities. In business-speak, the classes should be MECE (mutually exclusive, collectively exhaustive).

Determine the Base Classes

Looking at the data and our business objectives, we hypothesize the following classes:

1. **Complaint re: Product**

2. **Complaint re: Process**

3. **Ambivalent**

4. **Satisfied**

5. **Overjoyed / Praise**

6. **Gibberish**

Whether or not these classes will be comprehensive enough to hold all of the reviews, while also being differentiated enough to be clear is currently unknown.

If more classes are needed once training begins, that can be easily accommodated.

Step 4

Collect the Initial Training Set

Collect the Initial Training Set

https://www.linkedin.com/pulse/big-data-conundrum-garbage-out-other-challenges-business-platform/

- When training we must be mindful of quantity, quality, and representativeness of the data.
  - Too little data will not provide useful results
  - Too much data will overfit to the training
  - "Bad" data will produce "bad" results
  - Unrepresentative/Skewed Data with, you guessed it, will provide Unrepresentative/Skewed Results
- Guidelines: https://console.bluemix.net/docs/services/natural-language-classifier/using-your-data.html#size-limitations

Collect the Initial Training Set

Goal:

- Select 100 comments at random, which are at most 1024 characters long.

Hope:

- These become a representative sample of all the comments, and that each class has at least 5 to 10 comments in it.

Method:

- Use excel to limit the results to those with fewer than 1024 characters.

- Using excel, count the number of records you have with fewer than 1024 characters.

- Of these records, randomly select ~100 records for the training set. Note – 100 is a very small sample, but as you will find, classifying these records is time consuming! So for the purposes of this tutorial, we will only use 100 records

Classify the Training Set

- Create a new CSV file and copy into it the records you identified for training in the last step (copy into the first column)

- In the second column (and more if you are assigning more than one class to a comment), type the class label that characterizes the comment best (i.e., one of the base classes identified earlier, or a new class if you think it is warranted).

- You should have at least 5 records in each class. If you do not, you will need to add records that represent those classes. One way to do this is to extract them from the full set of comments. But for this exercise, I would advise just creating a comment(s).

Step 5

Use the Web Interface for Training

Use the Web Interface for Training

- Sign into your Bluemix account, and on the main (Dashboard) page find the service you registered for the NLC (under Services).

- Click on this service and then click "Launch Tool".
  - This launches a web application that allows you create and train your NLC model

- Click on Create Model

- Under "Define project details", give your project a name.

- Then under "Natural Language Classifier service instance" select the service you just created.

- Then click the Create button in the lower right hand of the screen.
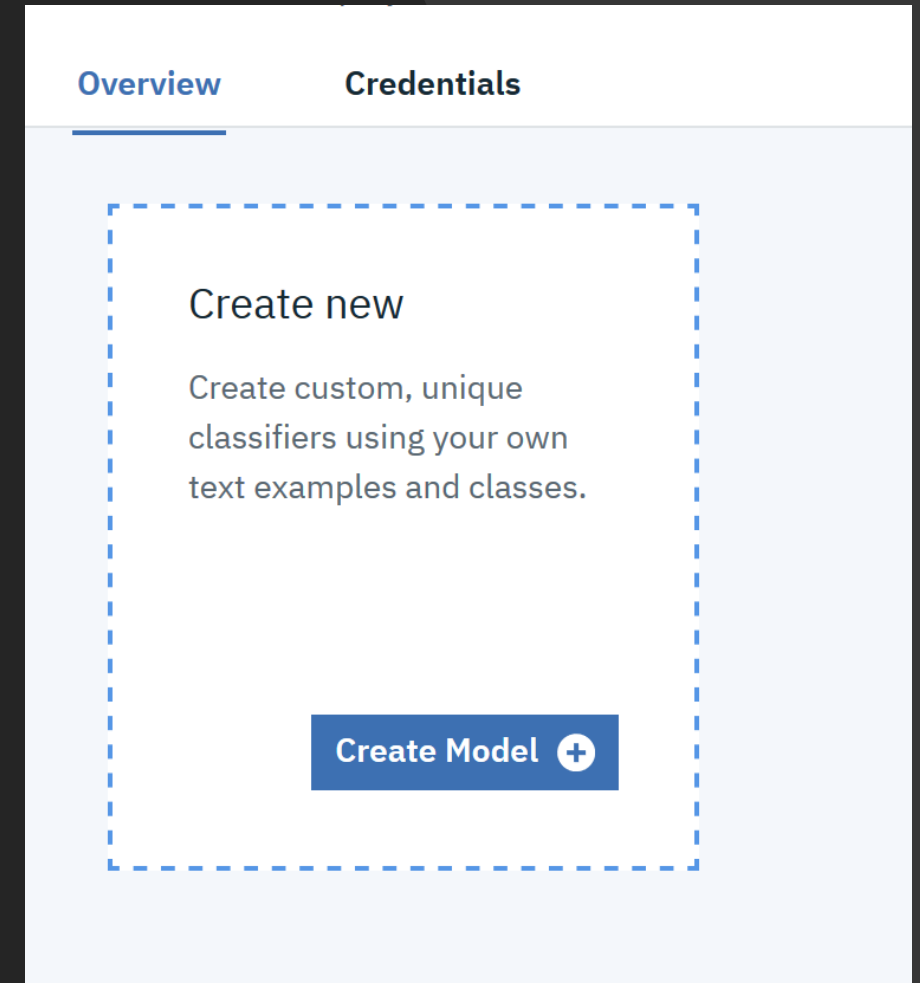
Use the Web Interface for Training

- Click Browse under "Upload to Project"

- Select the CSV training file you created earlier. This will import your trained data.

- Now select the check box beside the file you just uploaded (see box below). And then click Add to Model.

1. Upload to project

To add files to your project, drop a valid .csv or .txt file here.

Browse

2. Add from project

Drag .csv or .txt files from your project to the training area to add them to your model.

0 selected

There are no .csv or .txt files in your project.

2. Add from project

Drag .csv or .txt files from your project to the training area to add them to your model.

1 selected	Add to model

training_sample_clear_with

classes Nov 28.csv

21 Dec 2018, 5:58:29 pm
47.99 KB

Use the Web Interface for Training (2)

- You will now see the number of comments assigned to each class

- Click the "Train Model" button and enter a name for it (e.g., Keurig). Now wait for it to train!
- This may take a while…

- When you see the "Training Successful" message, click on the link to test it.

Step 6

Test the Classifier Online
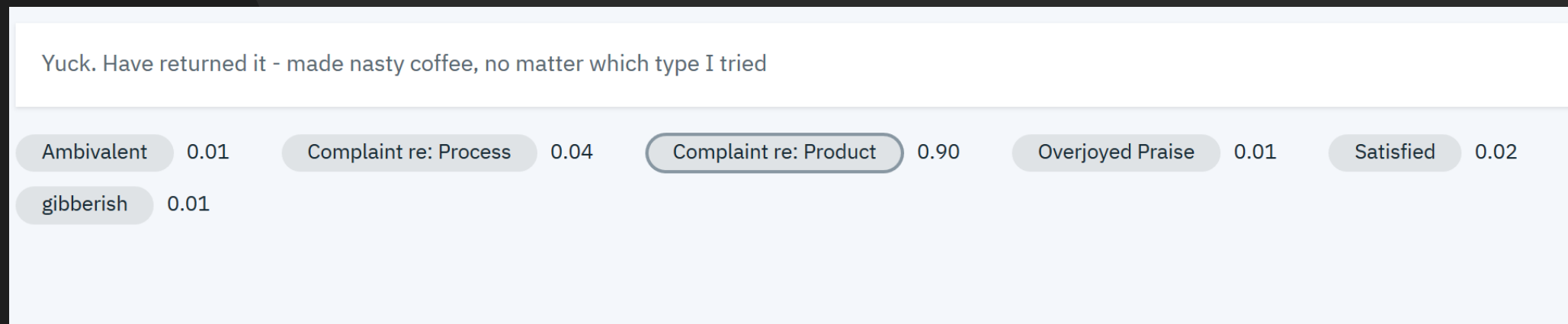
Test the Classifier Online

- After the training runs, we can use the web interface to test out how well the classification works.

Try:

"Works great and love the coffee."

"Yuck. Have returned it - made nasty coffee, no matter which type I tried."

"replaced one that quit after several years. Great system, could have a larger water tank."

Yuck. Have returned it - made nasty coffee, no matter which type I tried

| Ambivalent | 0.01 | Complaint re: Process | 0.04 | Complaint re: Product | 0.90 | Overjoyed Praise | 0.01 | Satisfied | 0.02 |
| gibberish | 0.01 | | | | | | | | |

Test the Classifier Online (2)

- To add the comment to your training set, click the + button beside the comment. This screen will allow you to add the new comment to a given class.
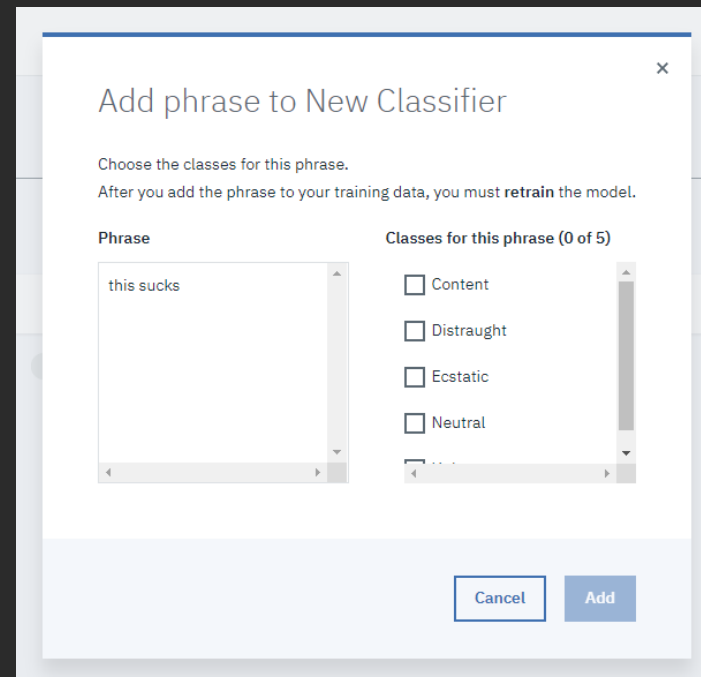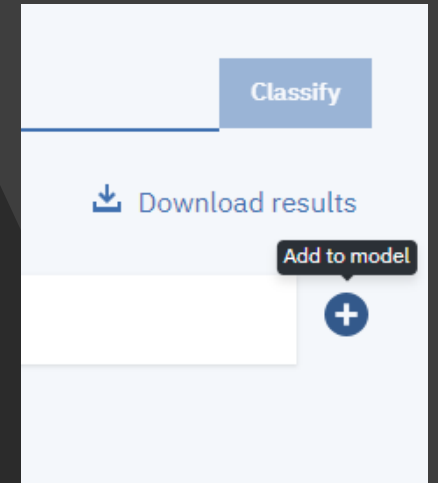
Deliverables

This concludes the walkthrough. Here are your deliverables:

Activity: Test your classifier using a variety of comments, attempting to trigger each one of the classes. These comments should not be from training set.

Submission: In a word document, submit a paragraph explaining one of the most significant misclassifications, and a screenshot of the test results for that misclassification. In your explanation, discuss why you think the comment was misclassified and how you would recommend fixing this issue in training.