CSC401: Natural Language Processing

Tutorial 2: Watson

Jan 15, 2015
TA: Stefania Raimondo
Today’s Tutorial

• IBM Watson Intro
• IBM BlueMix Intro
• Walk-through of BlueMix Sign-up
• Walk-through of creating and using a Service
• Natural Language Classifier Service Overview (Assignment 1)
IBM Watson
IBM Watson

2011: Watson beats humans at Jeopardy!

2005: “Impossible problem.” Only 35% correct responses and took minutes to answer.

- Natural language input
- Wide range of topics
- Clues aren’t straight forward questions
- Question understanding

- 4 Terabytes of structured and unstructured content (including all of Wikipedia)

- In 2011, Watson was just a QA platform...
IBM Watson Platform

• “IBM Watson is a technology platform that uses natural language processing and machine learning to reveal insights from large amounts of unstructured data” - http://www.ibm.com/smarterplanet/us/en/ibmwatson/what-is-watson.html
IBM Watson Services
Ex. IBM Watson Trend

Find and understand consumer trends.
Find the next hot thing, why it’s hot, and how long it’ll last.

• How?
  1. Search the internet for conversations about purchases
  2. Analyze sentiment and context
  3. Group conversations into topics and monitor topic development
  4. Forecast trend development
  5. Provide a summary of the trend

Check it out here: https://ibmwatson.trend.com/?cm_mcid=81589843967914351576502&cm_mcid=50200000-1452527711#

Other Examples:
IBM Bluemix

• “Bluemix is a place for **developers** to go to **quickly** create, deploy, and manage **applications in the cloud** without dealing with any of the underlying **infrastructure.”** – David Barnes, IBM

• Designed for **developers** — still need to know how to code

• Connect all the components of your application — databases, virtual machines, cloudfoundry, git, docker. Supports any programming language and development environment

• Provide components as “Services”

• Includes “Watson Services”

See it here:
http://www.ibm.com/cloud-computing/bluemix/
Setting up Bluemix!

• Part 1: Set up a Bluemix Account
• Part 2: Register Your Promo Code
• Part 3: Adding your teammate
Part 1: Setup Your Bluemix Account
Step 0: Start Registration

Go to the Bluemix Cloud-computing Student Developer Website:

https://developer.ibm.com/students/

Click Register (circled)
Step 1: Fill out the Signup info

Fill in all your info...

Make sure you have “YES! Create a Bluemix trial account for me” selected
Step 2: Sign in to IBM

Now that you’re registered, sign in.

Use the email you registered with for the IBM id

Use the password you just registered

Click Sign in
Part 2: Redeem your promo code
Step 1: Log into BlueMix if necessary

You might have to login to BlueMix

You can get to this page directly:
https://console.ng.bluemix.net/

Click Login
Alternate Login

Back at the Student Developer Website...
https://developer.ibm.com/students/

Click to redeem your promo code (circled)
Alternate Login continued:

To get to this page directly:

https://developer.ibm.com/students/learn/redeem-promo/

Log into your Bluemix account if necessary

Follow these steps to redeem your unique promo code. Remember that your promo code can be used only once; do not share it.

1. Register for an IBM Bluemix trial account if you don’t already have one. You will be asked to validate your email address, so please check your inbox and click the validation link.

2. Log in to your IBM Bluemix dashboard, click the number of days remaining on your account in the upper right corner, and click promo code.

3. Cut and paste the promo code that your educator provided into the “Enter promo code” field, and click APPLY. Refresh the page to see your extension reflected in the calendar.

Now you are ready to get started by accessing the IBM Bluemix tutorials and resources to help you on your way. Check out some of the cool projects other students have done. Or look for a nearby hackathon. Take advantage of the IBM Bluemix community!
Step 2: Validate Your Account

When you try to login you might see this screen.

Have patience!

You may have to wait 24 hrs for your Bluemix account to be activated and to receive the confirmation email.

But go check anyway.
Validate Your Account

Click “Validate Email Address” in the confirmation email you received
Step 3: Add the promo code

Now go back and login to Bluemix.

You might see either of these pages...

Either way, click the calendar sheet with a number in the top-left corner (circled).

This number shows you how many days you have left in your trial account.
Add the promo code

In the drop-down box, click the “promo code” link.
Add the promo code

In this window, add your promo code.

Your promo code will be provided to you via email.
Add the promo code

You should a “redeemed successfully” confirmation and your promo code in the dropdown.
Part 3: Adding your teammate

• 2 options
  • Method 1: Have one person register first and then invite their teammate to join Bluemix
    • Teammate will be added as a “member”
  • Method 2: Have both people register individually, then have one teammate invite the other to join their organization
    • Teammate will be added as a “collaborator”

• It doesn’t matter which you do for this course
• The following slides demonstrate Method 1
Step 1: Go to Manage your organization

Click the user icon in the top left of the screen (circled) to show the dropdown menu of user settings.
Go to Manage your organization

Click “Manage Organizations”.
Step 2: Invite a User

Click “Invite a User”
Step 3: Add your teammate’s info

Add your teammates name and email address.

Check of all of the circled boxes.
I.e. give them access to your organization and give them access to “All spaces” in your organization.
Step 4: Finish Inviting your Teammate

You will see this box if you were using Method 2 and your teammate already has a Bluemix account.

Just click ok.
Step 4: Finish Inviting your Teammate

You will see this box if your teammate doesn’t yet have a Bluemix account.

Select the “Invite” radio button, and click “OK”.

![Image of Bluemix invite process]
Step 5: Have your teammate Accept the Invitation

Have your teammate check their email and click the “Accept Invitation” link
Step 6: Have your teammate register for Bluemix

Have your teammate fill in their credentials and click “Register”
All Done!
All Done!
Quick Tour of the Bluemix console
The console
Dropdown menu

Use this to access the Dashboard and the Catalog
The Catalog

Here are Services that you can add to your application.

Scroll down to see the Watson Services
Creating and Adding a Service

• Add the Natural Language Classifier Service that you will use for Assignment 1
Create a Natural Language Classifier Service

From the Services page, click on the Natural Language Classifier Service icon.

You should see this.

Rename the service, if you like.

Ensure that you have “Leave unbound selected”.

Click Create.
Voila!

Now you can see your service as in icon on your Dashboard.
Natural Language Classifier

• Classifies short texts (< 1000 characters) into categories
  • Ex. Classify SMS as personal, work, promotional
  • Ex. Classify tweets as events, news, opinions
  • Ex. Classify tweets as positive, negative (Assignment 1)

• Blackbox - Uses deep learning
Natural Language Classifier

Using the Natural Language Classifier Service

1. Prepare training data
   - Identify class labels
   - Collect representative texts
   - Match classes to texts

2. Create and train the classifier
   - Use the API to upload the training data
   - Training begins immediately

3. Query the trained classifier
   - Use the API to send text to the classifier
   - The service returns the top matching class and other possible matches

4. Evaluate results and update training data
   - Update your training data based on the classification results
   - Create and train a classifier using updated training data
How to use it?

• Find your credentials
• Make the api calls
Finding your Credentials

Go back to the Dashboard and click on the Natural Language Classifier Service that you created earlier.
Finding your credentials

Click on the little arrow (circled) to get the dropdown menu on the left.

Click “Service Credentials” to see this screen.

You’ll need the username and password, as well as the url (which is the same for everyone).

Note that this username and password are NOT the same as the ones for your IBM Bluemix Account.
Using the API

This page outlines the API calls and is shown on the right:


You may also find set of pages helpful:

Using the API

- 5 Calls:
  - Create Classifier
  - List classifiers
  - Get info
  - Delete Classifier
  - Classify

- Each Call has a “Request” and a “Response” (as shown to the right for “Create classifier”)
Details of the Create Classifier API

- Request components:
  - training_data – a csv file containing your training instances
  - training_metadata – JSON formatted
Details of the Create Classifier API

• Request components:
  • training_data – a csv file containing your training instances
    • 2 column comma-separated format:
      • Text – an example of the text you want to classify
      • Class – the correct class for that example
  • training_metadata – JSON formatted
Details of the Create Classifier API

• Request components:
  • training_data – a csv file containing your training instances
  • training_metadata – JSON formatted
    • 2 components:
      • Language: “en” for English
      • Name: the name of the classifier

```json
{
  "language": "en",
  "name": "My Classifier"
}
```
Details of the Create Classifier API

• Making the Request
  • Examples on the website use Curl, Node.js, and Java, but you can use whatever you want
    • For example, you can also use python’s request library...
  • Components:
    • Username, password
    • training_data file name
    • training_metadata JSON
    • url for the NLClassifier service

```
curl -u "{username}”:"{password}" -F training_data=@train.csv -F training_metadata="{"language":"en","name":"My Classifier"}" "https://gateway.watsonplatform.net/natural-language-classifier/api/v1/classifiers"
```
Details of the Create Classifier API

• Getting the response
  • The actual format of the response (HTML, JSON) depends on whether you are using curl, a python json library etc.
  • But here are the components:

```json
{
  "classifier_id": "10D41B-nlc-1",
  "name": "My Classifier",
  "language": "en",
  "created": "2015-05-28T18:01:57.393Z",
  "url": "https://gateway.watsonplatform.net/natural-language-classifier/api/v1/classifiers/10D41B-nlc-1",
  "status": "Training",
  "status_description": "The classifier instance is in its training phase, not yet ready to accept classify requests"
}
```
Have fun!