Natural Language Understanding

Requirements:

1. Bluemix account; sign up for a trial account
2. Knowledge studio account; sign up for a trial account
3. Python 2.7 or 3
4. Get the Watson Developer Cloud Python SDK

Step-by-step instructions:

Let's go through all the steps to build our project!

Create Project

It is assumed that you have an IBM Bluemix account. Sign in to your account and select **Catalog** and search **Natural Language Understanding** and create new instance of service.

View Credentials

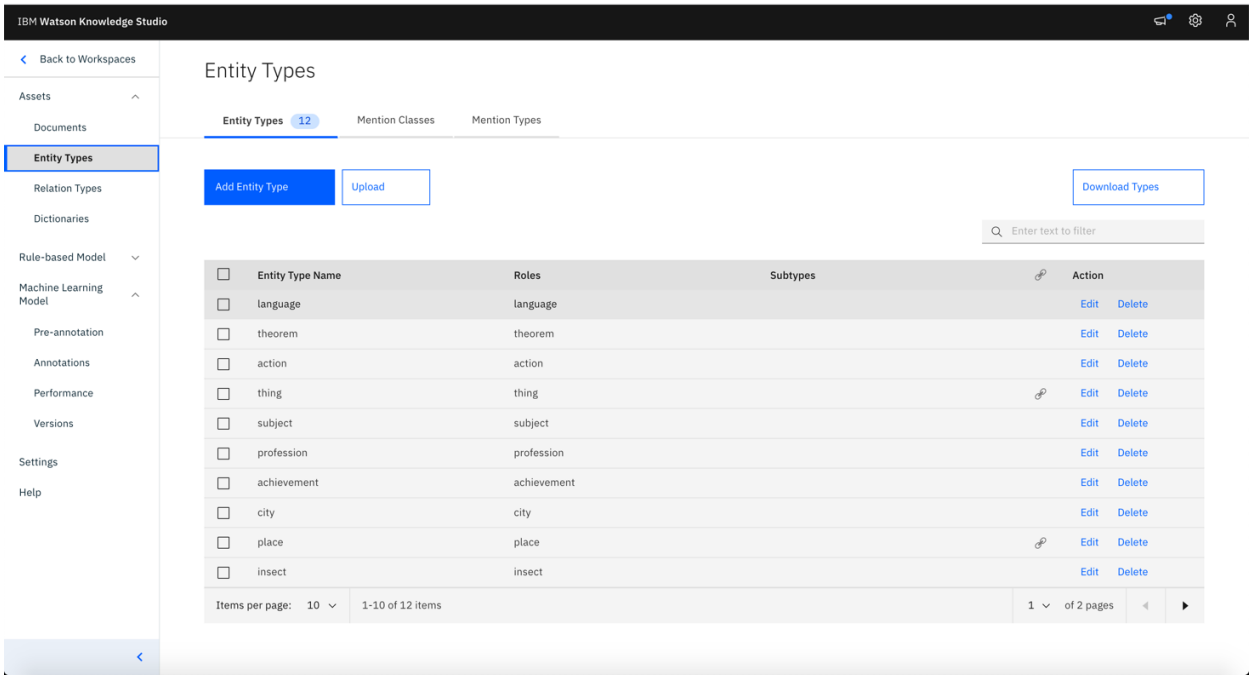
You will need to click on the **Service Credentials** and **View Credentials** link to get the details that we need to populate the Jupyter notebook.

Configure Machine Learning Annotator in IBM Knowledge Studio

As defined in the Requirements section, it is assumed that you have a Knowledge Studio account. Sign in to your account and launch Knowledge Studio.

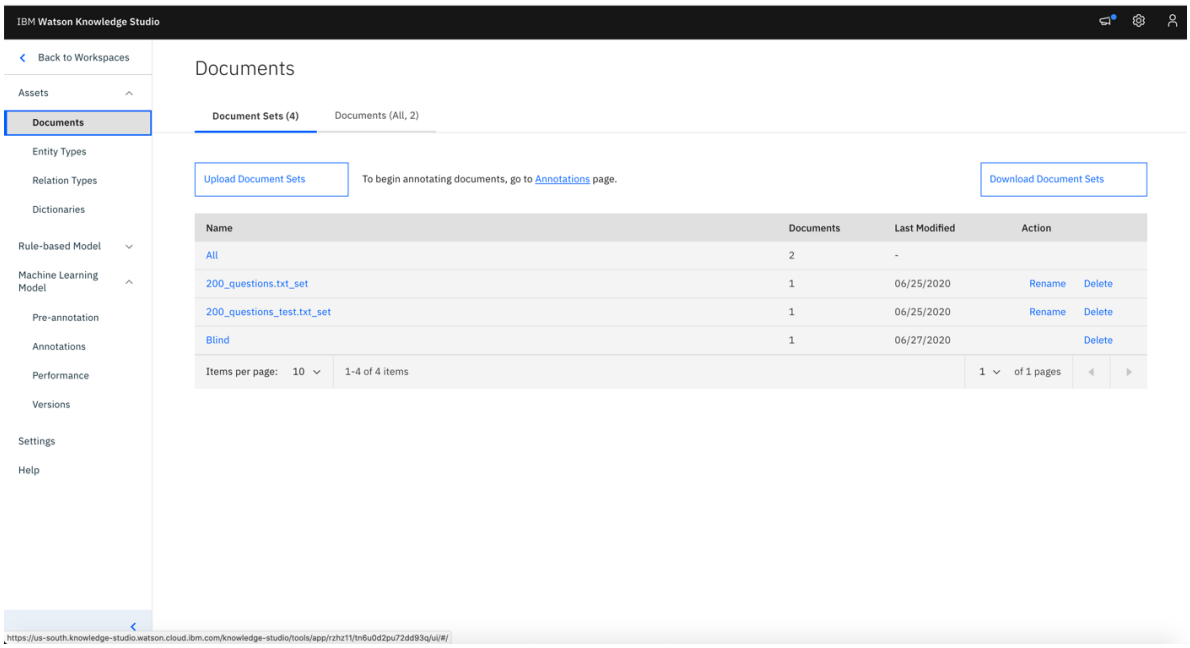
Add Entities Type

The first step is to click **Assets > Entity types.** We defined ten entity types.



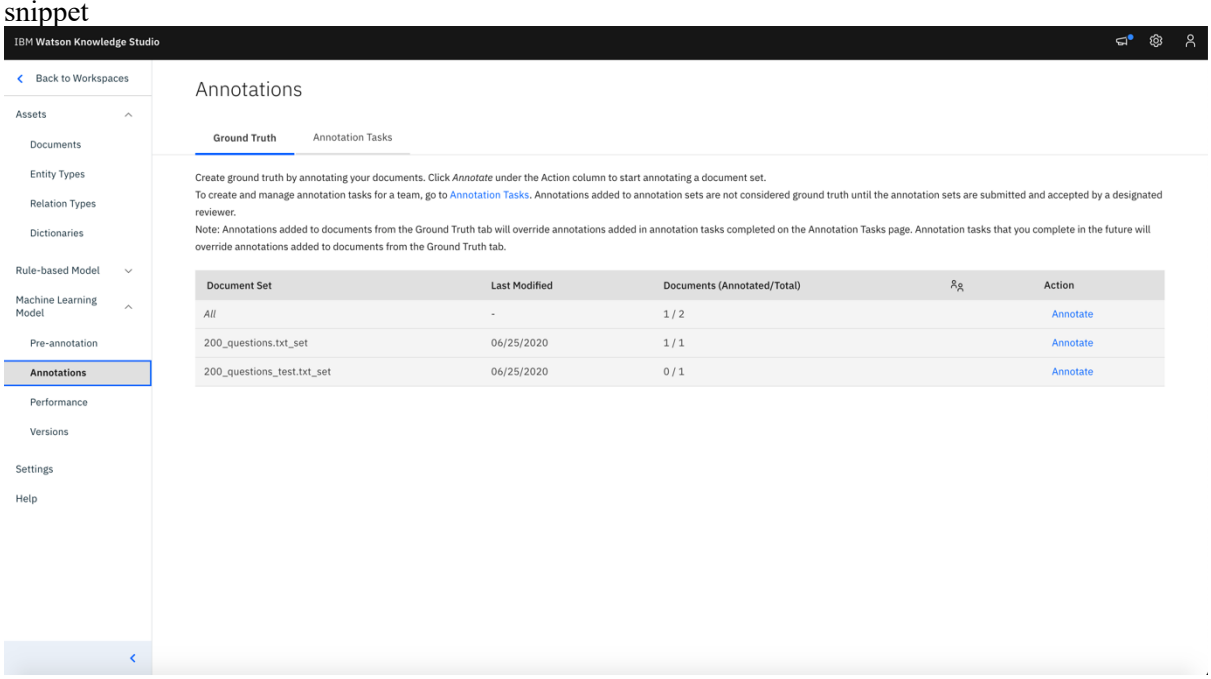
Upload Documents for Annotation

The next step is to click **Documents** > **Upload Document Set**. These documents contain questions data.

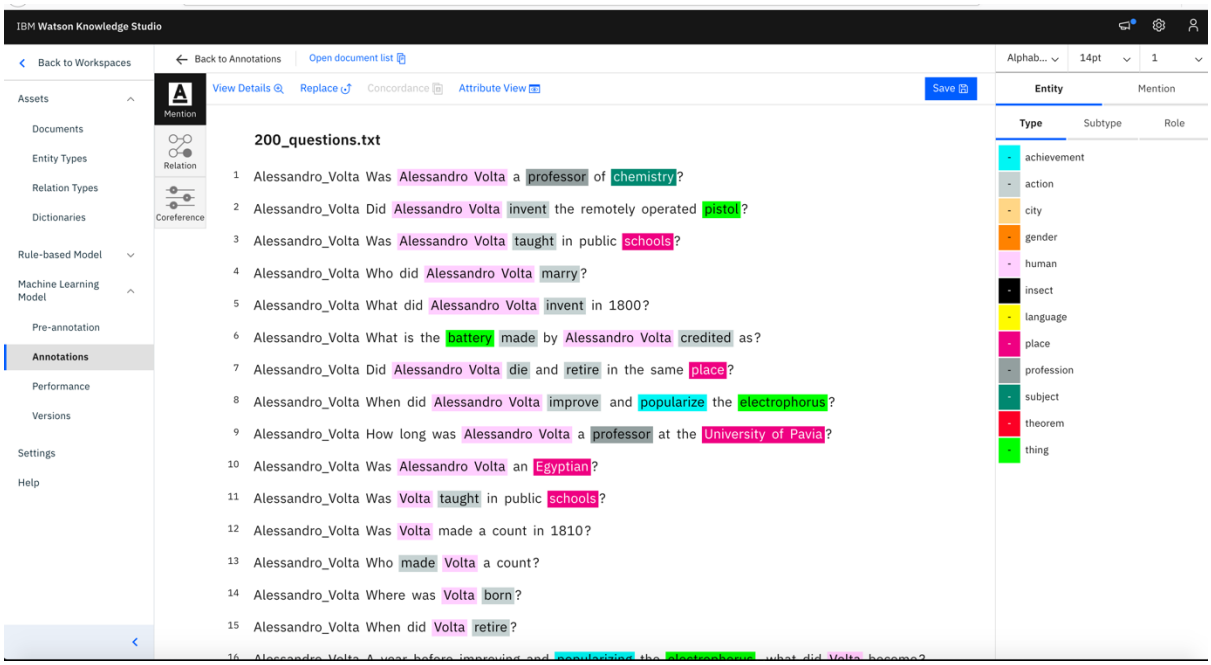


Create Annotation Set

The system needs a set of human annotators for identifying all entities for each document containing questions. Under **Machine Learning Model > Annotations** you can see two documents in below snippet

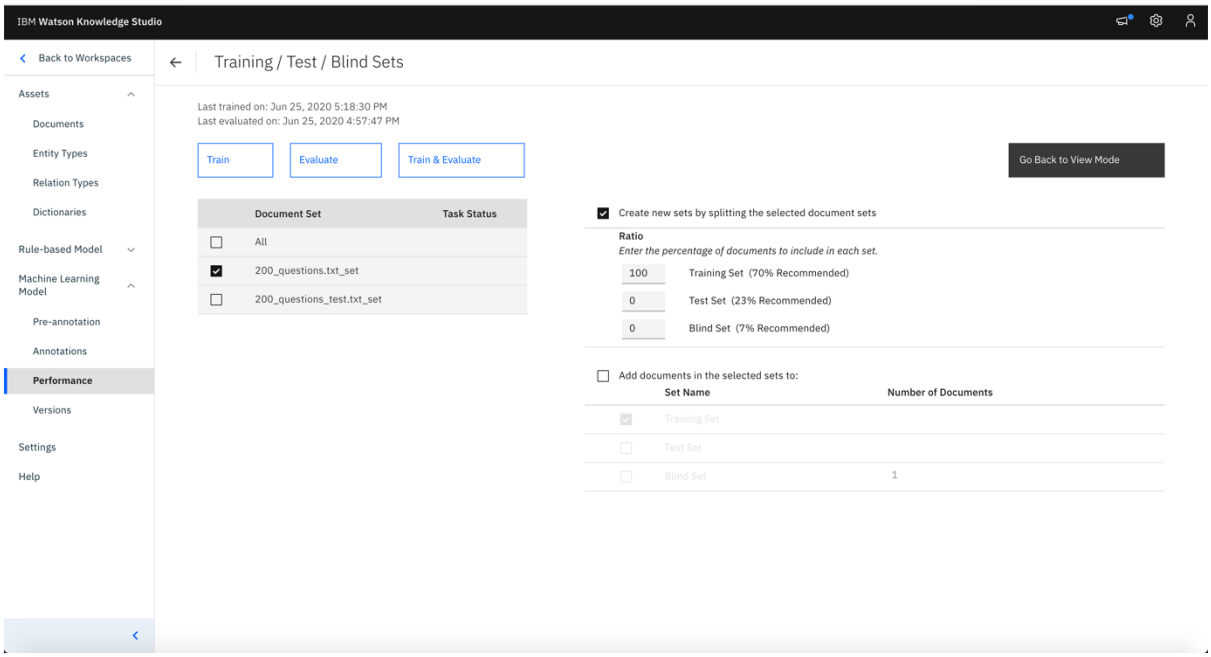


Click on **Annotate,** we have annotated different entities which found relevant according to the data



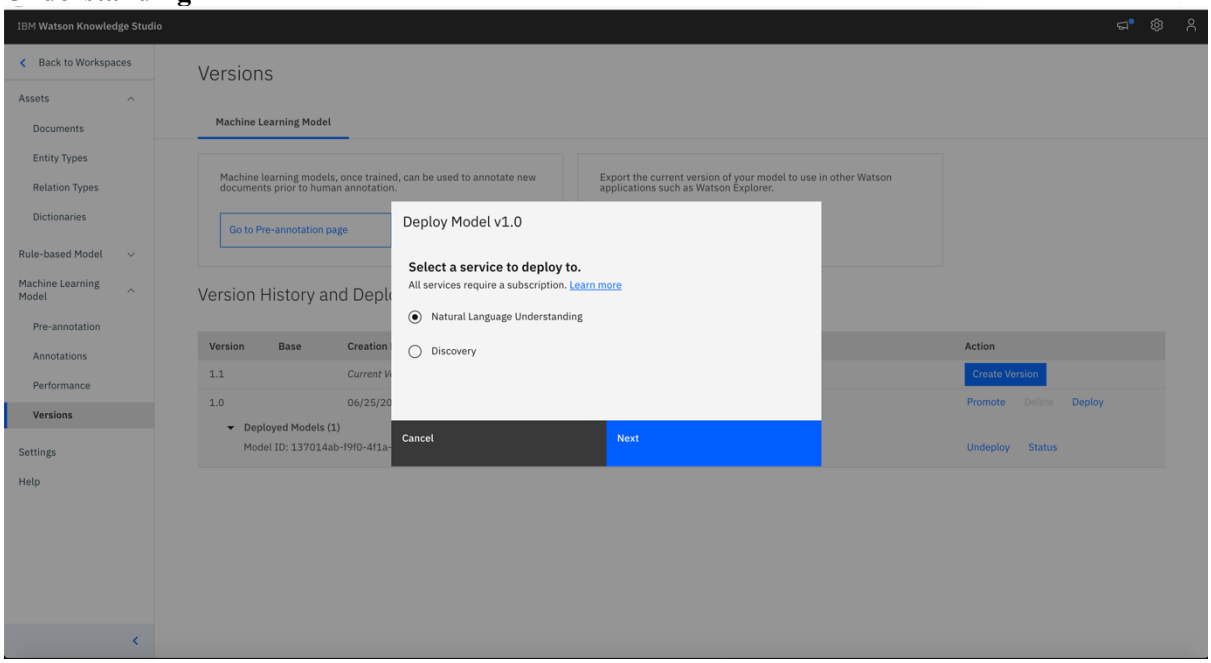
Create a Machine Learning Annotator Model

We trained and evaluated with separate training and evaluation files.



Deploy the Machine Learning Annotator Model

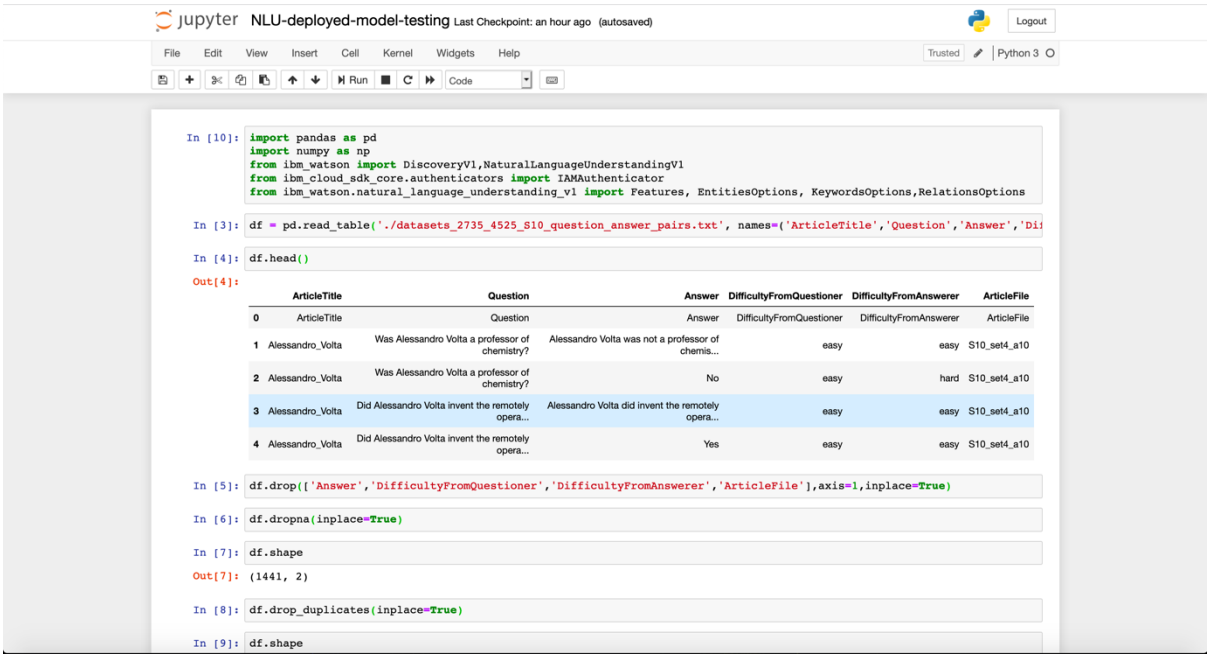
Go to **Machine Learning Model > Versions > Create Version > Deploy > Natural Language Understanding**



Copy model id of deployed model

Create Jupyter notebook Back-End on Python:

1. **Load testing data**



2. Copy API key of NLU service and model id of deployed model. Populate IMAuthenticator with NLU API key, and model id in model dictionary.

```
In [11]: authenticator = IAMAuthenticator("IPaKtGaZ8aVqTL4P1lJzoO88dwQXTGgs0tZqaSHMbSlI")

In [12]: url = "https://gateway.watsonplatform.net/natural-language-understanding/api"
         service = NaturalLanguageUnderstandingV1(version='2018-03-16',authenticator=authenticator)

In [13]: service.set_service_url(url)

In [21]: response = service.analyze(
             text=df['Question'].iloc[1],    features={

                 'relations': {
                     'model': '137014ab-f9f0-4f1a-8db0-371697970218'
                 },
                 'entities':{ 'model': '137014ab-f9f0-4f1a-8db0-371697970218'}
                 }
             ).get_result()

In [22]: response

Out[22]: {'usage': {'text_units': 1, 'text_characters': 46, 'features': 2},
          'relations': [],
          'language': 'en',
          'entities': [{'type': 'human',
            'text': 'Alessandro Volta',
            'disambiguation': {'subtype': ['NONE']},
            'count': 1,
            'confidence': 0.999277},
           {'type': 'profession',
            'text': 'professor',
            'disambiguation': {'subtype': ['NONE']},
            'count': 1,
            'confidence': 0.998283},
           {'type': 'subject',
            'text': 'chemistry',
            'disambiguation': {'subtype': ['NONE']},
            'count': 1,
            'confidence': 0.982997}]}
```

In response model will return tagged entities.