

Resolving some issues with pytorch-nlp
(due to updates to PyTorch & spaCy)...

Issue: "int_classes"

```
~/opt/anaconda3/lib/python3.8/site-packages/torch/nlp/_third_party/weighted_random_sampler.py in <module>
  2
  3 from torch.utils.data.sampler import Sampler
----> 4 from torch._six import int_classes as _int_classes
  5
  6

ImportError: cannot import name 'int_classes' from 'torch._six' (/Users/georgehc/opt/anaconda3/lib/python3.8/site-packages/torch/_six.py)
```

Edit this file (weighted_random_sampler.py) so that it looks like the following (comment out one line and define int_classes to be int):

```
1 import torch
2
3 from torch.utils.data.sampler import Sampler
4 # from torch._six import int_classes as _int_classes
5 int_classes = int
6
7
8 class WeightedRandomSampler(Sampler):
9
10     def __init__(self, weights, num_samples, replacement=True):
11         # NOTE: Adapted 'WeightedRandomSampler' to accept 'num_samples'
12         if not isinstance(num_samples, int) or isinstance(num_samples, bool) or num_samples < 0:
13             raise ValueError("num_samples should be a positive integer value")
```

I found this fix here: <https://discuss.pytorch.org/t/cannot-import-name-int-classes-from-torch-six/126611>

Issue: deprecated spacy "en" model...

OSError: [E941] Can't find model 'en'. It looks like you're trying to load a model from a as of spaCy v3.0. To load the model, use its full name instead:

```
nlp = spacy.load("en_core_web_sm")
```

For more details on the available models, see the models directory: <https://spacy.io/model>
blank model, use spacy.blank: nlp = spacy.blank("en")

During handling of the above exception, another exception occurred:

```
ValueError                                Traceback (most recent call last)
<ipython-input-9-6f68fb0d3ff0> in <module>
----> 1 encoder = SpacyEncoder([data['text'] for data in proper_train_dataset])

~/opt/anaconda3/lib/python3.8/site-packages/torch/nlp/encoders/text/spacy_encoder.py in __i
s)
    54         self.spacy = spacy.load(language, disable=['parser', 'tagger', 'ne
    55     except OSError:
----> 56         raise ValueError(("Language '{0}' not found. Install using "
    57                             "spaCy: `python -m spacy download {0}`").format(
    58             else:
```

ValueError: Language 'en' not found. Install using spaCy: `python -m spacy download en`

```
44     # Use English as default when no language was specified
45     # language = kwargs.get('language', 'en')
46     language = kwargs.get('language', 'en_core_web_sm')
47
48     # All languages supported by spaCy can be found here:
49     # https://spacy.io/models/#available-models
50     # supported_languages = ['en', 'de', 'es', 'pt', 'fr', 'it', 'nl', 'xx']
51     supported_languages = ['en_core_web_sm', 'de', 'es', 'pt', 'fr', 'it', 'nl', 'xx']
52
53     if language in supported_languages:
54         # Load the spaCy language model if it has been installed
55         try:
56             # self.spacy = spacy.load(language, disable=['parser', 'tagger', 'ner'])
57             self.spacy = spacy.load(language, disable=['parser', 'tagger', 'ner', 'lemmatizer'])
58         except OSError:
59             raise ValueError(("Language '{0}' not found. Install using "
60                             "spaCy: `python -m spacy download {0}`").format(language))
61
```

Edit
spacy_encoder.py
and modify 3 lines