16_all_sentences

March 21, 2017

2/24/17 - smiel

1 Collect all sentence data into a single csv

```python
In [1]: %matplotlib inline
        # run me when first starting this notebook
        import os

        import numpy as np
        import pandas as pd

        path = '/research/ella/rivendell'

        datasets = {
            'ceeaus': 'CEEAUS',
            'fce': 'FCE',
            'icnale': 'ICNALE',
            'nog': 'NOG',
            'ra': 'RA-2016',
            'teccl': 'TECCL',
            'asap': 'ASAP',
            'chungdahm': 'Chungdahm',
            'gachon': 'Gachon',
            'moecs': 'MOECS',
            'nucle': 'NUCLE',
            'sbac': 'SBAC-Field',
            'toefl': 'TOEFL-11'
        }

        dfs = []
        for folder, dataset in datasets.items():
            print(dataset)
            data_path = os.path.join(path, folder)

            # load sentences
            df_in = pd.read_csv(os.path.join(data_path, '{}_sentences.csv'.format(dataset)), enc
            dfs.append(df_in)
```

```python
        print('Creating one big data frame')
        df_out = pd.concat(dfs, ignore_index=True)

        print('Saving')
        df_out.to_csv(os.path.join(path, 'all_sentences.csv'), encoding='utf8', index=False)
        print('Done!')
```

```
NUCLE
ASAP
NOG
CEEAUS
FCE
SBAC-Field
TECCL
ICNALE


/home/smiel/.venvs/rivendell/local/lib/python2.7/site-packages/IPython/core/interactiveshell.py:
    interactivity=interactivity, compiler=compiler, result=result)


TOEFL-11
Chungdahm
RA-2016
Gachon
MOECS
Creating one big data frame
Saving
Done!
```