
Neural Fingerprint (nfp)

Peter St. John

Jun 14, 2022

CONTENTS

1	Tutorials	3
1.1	Tokenizer usage examples	3
2	Neural Fingerprint (nfp), tf.keras layers for molecular graphs	5

NFP can be installed with pip into a suitable python environment with

```
pip install nfp
```

The only dependency that is not pip-installable is rdkit, which can be [installed via conda-forge](#).

1.1 Tokenizer usage examples

The Tokenizer class allows transforming arbitrary inputs into integer classes

```
[1]: from nfp.preprocessing import Tokenizer
```

```
[2]: tokenizer = Tokenizer()  
tokenizer.train = True
```

The 0 and 1 classes are reserved for the <MASK> and missing labels, respectively.

```
[3]: [tokenizer(item) for item in ['A', 'B', 'C', 'A']]
```

```
[3]: [2, 3, 4, 2]
```

When train is set to False, unknown items are assigned the missing label

```
[4]: tokenizer.train = False  
[tokenizer(item) for item in ['A', 'D']]
```

```
[4]: [2, 1]
```

The total number of seen classes is available from the num_classes property, useful to initializing embedding layer weights.

```
[5]: tokenizer.num_classes
```

```
[5]: 4
```


NEURAL FINGERPRINT (NFP), TF.KERAS LAYERS FOR MOLECULAR GRAPHS

A collection of tensorflow.keras layers, losses, and preprocessors for building graph neural networks on molecules.