

TFDS now supports the [Croissant](https://mlcommons.org/croissant)  format (<https://mlcommons.org/croissant>)! Read the [documentation](https://www.tensorflow.org/datasets/format_specific_dataset_builders#croissantbuilder) (https://www.tensorflow.org/datasets/format_specific_dataset_builders#croissantbuilder) to know more.

mtnt

- **Description:**

MTNT: Machine Translation of Noisy Text

- **Homepage:** <https://pmichel31415.github.io/mtnt/index.html>
(<https://pmichel31415.github.io/mtnt/index.html>)
- **Source code:** `tlds.translate.mtnt.Mtnt`
(https://github.com/tensorflow/datasets/tree/master/tensorflow_datasets/translate/mtnt/mtnt.py)
- **Versions:**
 - **1.0.0** (default): Initial release.
- **Download size:** 35.08 MiB
- **Auto-cached** ([documentation](https://www.tensorflow.org/datasets/performances#auto-caching) (<https://www.tensorflow.org/datasets/performances#auto-caching>)): Yes
- **Feature structure:**

```
FeaturesDict({  
    'dst': Text(shape=(), dtype=string),  
    'src': Text(shape=(), dtype=string),  
})
```

- **Feature documentation:**

Feature	Class	Shape	Dtype	Description
	FeaturesDict			
dst	Text		string	

Feature	Class	Shape	Dtype	Description
src	Text		string	

- **Supervised keys** (See [as_supervised doc](#) (https://www.tensorflow.org/datasets/api_docs/python/tfds/load#args)): ('src', 'dst')
- **Figure** ([tfds.show_examples](#) (https://www.tensorflow.org/datasets/api_docs/python/tfds/visualization/show_examples)): Not supported.
- **Citation:**

```
@InProceedings{michel2018mtnt,
  author    = {Michel, Paul and Neubig, Graham},
  title     = {MTNT: A Testbed for Machine Translation of Noisy Text},
  booktitle = {Proceedings of the 2018 Conference on Empirical Methods in Natural Lan
}
```

mtnt/en-fr (default config)

- **Dataset size:** 11.33 MiB
- **Splits:**

Split	Examples
'test'	1,020
'train'	35,692
'valid'	811

- **Examples** ([tfds.as_dataframe](#) (https://www.tensorflow.org/datasets/api_docs/python/tfds/as_dataframe)):

[Display examples...](#)

mtnt/en-ja

- **Dataset size:** 4.41 MiB
- **Splits:**

Split	Examples
'test'	946
'train'	5,746
'valid'	892

- **Examples** (`tfds.as_dataframe` (https://www.tensorflow.org/datasets/api_docs/python/tfds/as_dataframe)):

[Display examples...](#)

mtnt/fr-en

- **Dataset size:** 8.28 MiB
- **Splits:**

Split	Examples
'test'	1,022
'train'	18,942

Split	Examples
'valid'	876

- **Examples** ([tfds.as_dataframe](https://www.tensorflow.org/datasets/api_docs/python/tfds/as_dataframe) (https://www.tensorflow.org/datasets/api_docs/python/tfds/as_dataframe)):

[Display examples...](#)

Except as otherwise noted, the content of this page is licensed under the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) (<https://creativecommons.org/licenses/by/4.0/>), and code samples are licensed under the [Apache 2.0 License](https://www.apache.org/licenses/LICENSE-2.0) (<https://www.apache.org/licenses/LICENSE-2.0>). For details, see the [Google Developers Site Policies](https://developers.google.com/site-policies) (<https://developers.google.com/site-policies>). Java is a registered trademark of Oracle and/or its affiliates.

Last updated 2022-12-06 UTC.