

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.

xtreme_xnli

- **Description:**

This dataset contains machine translations of MNLI into each of the XNLI languages. The translation data is provided by XTREME. Note that this is different from the machine translated data provided by the original XNLI paper.

- **Homepage:** <https://www.nyu.edu/projects/bowman/xnli/>
(<https://www.nyu.edu/projects/bowman/xnli/>)
- **Source code:** `tfds.text.xtreme_xnli.XtremeXnli`
(https://github.com/tensorflow/datasets/tree/master/tensorflow_datasets/text/xtreme_xnli/xtreme_xnli.py)
- **Versions:**
 - `1.1.0` (default): No release notes.
- **Download size:** 2.31 GiB
- **Dataset size:** 1.59 GiB
- **Auto-cached** ([documentation](https://www.tensorflow.org/datasets/performances#auto-caching) (<https://www.tensorflow.org/datasets/performances#auto-caching>)): No
- **Splits:**

Split	Examples
'train'	392,570

- **Feature structure:**

```
FeaturesDict({  
  'hypothesis': TranslationVariableLanguages({  
    'language': Text(shape=(), dtype=string),
```

```

        'translation': Text(shape=(), dtype=string),
    )),
    'label': ClassLabel(shape=(), dtype=int64, num_classes=3),
    'premise': Translation({
        'ar': Text(shape=(), dtype=string),
        'bg': Text(shape=(), dtype=string),
        'de': Text(shape=(), dtype=string),
        'el': Text(shape=(), dtype=string),
        'en': Text(shape=(), dtype=string),
        'es': Text(shape=(), dtype=string),
        'fr': Text(shape=(), dtype=string),
        'hi': Text(shape=(), dtype=string),
        'ru': Text(shape=(), dtype=string),
        'sw': Text(shape=(), dtype=string),
        'th': Text(shape=(), dtype=string),
        'tr': Text(shape=(), dtype=string),
        'ur': Text(shape=(), dtype=string),
        'vi': Text(shape=(), dtype=string),
        'zh': Text(shape=(), dtype=string),
    })),
))

```

- **Feature documentation:**

Feature	Class	Shape	Dtype	Description
	FeaturesDict			
hypothesis	TranslationVariableLanguages			
hypothesis/language	Text		string	
hypothesis/translation	Text		string	
label	ClassLabel		int64	
premise	Translation			
premise/ar	Text		string	
premise/bg	Text		string	
premise/de	Text		string	
premise/el	Text		string	
premise/en	Text		string	

Feature	Class	Shape	Dtype	Description
premise/es	Text		string	
premise/fr	Text		string	
premise/hi	Text		string	
premise/ru	Text		string	
premise/sw	Text		string	
premise/th	Text		string	
premise/tr	Text		string	
premise/ur	Text		string	
premise/vi	Text		string	
premise/zh	Text		string	

- **Supervised keys** (See [as_supervised doc](https://www.tensorflow.org/datasets/api_docs/python/tfds/load#args) (https://www.tensorflow.org/datasets/api_docs/python/tfds/load#args)): **None**
- **Figure** ([tfds.show_examples](https://www.tensorflow.org/datasets/api_docs/python/tfds/visualization/show_examples) (https://www.tensorflow.org/datasets/api_docs/python/tfds/visualization/show_examples)): Not supported.
- **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...](#)

- **Citation:**

```
@article{hu2020xtreme,
  author    = {Junjie Hu and Sebastian Ruder and Aditya Siddhant and Graham Neubig},
  title     = {XTREME: A Massively Multilingual Multi-task Benchmark for Evaluating Machine Translation Models},
  journal   = {CoRR},
  volume    = {abs/2003.11080},
  year      = {2020},
}
```

```
    archivePrefix = {arXiv},  
    eprint       = {2003.11080}  
}
```

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.