# tf.string_split

```
string_split(
    source,
    delimiter=' ',
    skip_empty=True
)
```

Defined in `tensorflow/python/ops/string_ops.py`.

See the guide: Strings > Splitting

Split elements of `source` based on `delimiter` into a `SparseTensor`.

Let N be the size of source (typically N will be the batch size). Split each element of `source` based on `delimiter` and return a `SparseTensor` containing the split tokens. Empty tokens are ignored.

If `delimiter` is an empty string, each element of the `source` is split into individual strings, each containing one byte. (This includes splitting multibyte sequences of UTF-8.) If delimiter contains multiple bytes, it is treated as a set of delimiters with each considered a potential split point.

For example: N = 2, source[0] is 'hello world' and source[1] is 'a b c', then the output will be

st.indices = [0, 0; 0, 1; 1, 0; 1, 1; 1, 2] st.shape = [2, 3] st.values = ['hello', 'world', 'a', 'b', 'c']

## Args:

- `source`: `1-D` string `Tensor`, the strings to split.
- `delimiter`: `0-D` string `Tensor`, the delimiter character, the string should be length 0 or 1.
- `skip_empty`: A `bool`. If `True`, skip the empty strings from the result.

## Raises:

- `ValueError`: If delimiter is not a string.

## Returns:

A `SparseTensor` of rank `2`, the strings split according to the delimiter. The first column of the indices corresponds to the row in `source` and the second column corresponds to the index of the split component in this row.



*Last updated November 2, 2017.*

GitHub

Twitter

**Support**

Issue Tracker

Release Notes

Stack Overflow

English

**Terms** | **Privacy**