

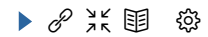
nlp_stopwords by ars0107



```
%pyspark
# import stopwords library
from pyspark.ml.feature import StopWordsRemover
```

```
%pyspark
# Create DataFrame
sentenceData = spark.createDataFrame([
    (0, ["Big", "data", "is", "super", "powerful"]),
    (1, ["This", "is", "going", "to", "be", "epic"])
], ["id", "raw"])

sentenceData.show()
```



```
+---+-----+
| id|          raw|
+---+-----+
|  0|[Big, data, is, s...|
|  1|[This, is, going,...|
+---+-----+
```

Interpreter: spark.pyspark. **FINISHED** Took 4 sec 184 millisec. Updated by ars0107 on February 04 2019, 8:48:04 AM (CST)



Run

Started

Juno ▾



```
%pyspark (/U4G66226D/spaces)
# Instantiate Remover
remover = StopWordsRemover(inputCol="raw", outputCol="filtered")
```

```
%pyspark
# Transform and show data
remover.transform(sentenceData).show(truncate=False)
```

```
+---+-----+-----+
|id|raw                                     |filtered                               |
+---+-----+-----+
|0 | [Big, data, is, super, powerful] | [Big, data, super, powerful] |
|1 | [This, is, going, to, be, epic]  | [going, epic]                |
+---+-----+-----+
```

Run

Started

Juno ▾

