

Overview

Dataset Statistics

Number of Variables	5
Number of Rows	1087
Missing Cells	0
Missing Cells (%)	0.0%
Duplicate Rows	0
Duplicate Rows (%)	0.0%
Total Size in Memory	1.3 MB
Average Row Size in Memory	1.3 KB
Variable Types	Categorical: 4 Numerical: 1

Dataset Insights

input_seq_length

 is skewed

Skewed

input

 has a high cardinality: 1087 distinct values

High Cardinality

membrane

 has constant length 1

Constant Length

labels

 has constant length 1

Constant Length

input

 has all distinct values

Unique

Variables

input

categorical

Show Details

Approximate Distinct Count	1087
Approximate Unique (%)	100.0%
Missing	0
Missing (%)	0.0%
Memory Size	1.2 MB

input

Top 10 of 1087 input

Approximate Distinct Count

10

<div>loc</div> <div>categorical</div> <div>Show Details</div>	Approximate Unique (%)		0.9%	
	Missing		0	
	Missing (%)		0.0%	
	Memory Size		81.9 KB	
<div>membrane</div> <div>categorical</div> <div>Show Details</div>	Approximate Distinct Count		2	
	Approximate Unique (%)		0.2%	
	Missing		0	
	Missing (%)		0.0%	
	Memory Size		70.1 KB	
<div>labels</div> <div>categorical</div> <div>Show Details</div>	Approximate Distinct Count		2	
	Approximate Unique (%)		0.2%	
	Missing		0	
	Missing (%)		0.0%	
	Memory Size		70.1 KB	
<div>input_seq_length</div> <div>numerical</div> <div>Show Details</div>	Approximate Distinct Count	712	Mean	546.5658
			Minimum	40
	Approximate Unique (%)	65.5%	Maximum	5654
	Missing	0	Zeros	0
	Missing (%)	0.0%	Zeros (%)	0.0%
	Infinite	0	Negatives	0
	Infinite (%)	0.0%	Negatives (%)	0.0%
	Memory Size	17.0 KB		

Interactions

Correlations

Pearson

Spearman

KendallTau

Missing Values

Bar Chart

Spectrum

Heat Map

Dendogram

Report generated with [DataPrep](#)