

Overview

Dataset Statistics

Number of Variables	5
Number of Rows	6622
Missing Cells	0
Missing Cells (%)	0.0%
Duplicate Rows	0
Duplicate Rows (%)	0.0%
Total Size in Memory	8.2 MB
Average Row Size in Memory	1.3 KB
Variable Types	Categorical: 4 Numerical: 1

Dataset Insights

<code>input_seq_length</code> is skewed	Skewed
<code>input</code> has a high cardinality: 6622 distinct values	High Cardinality
<code>membrane</code> has constant length 1	Constant Length
<code>labels</code> has constant length 1	Constant Length
<code>input</code> has all distinct values	Unique

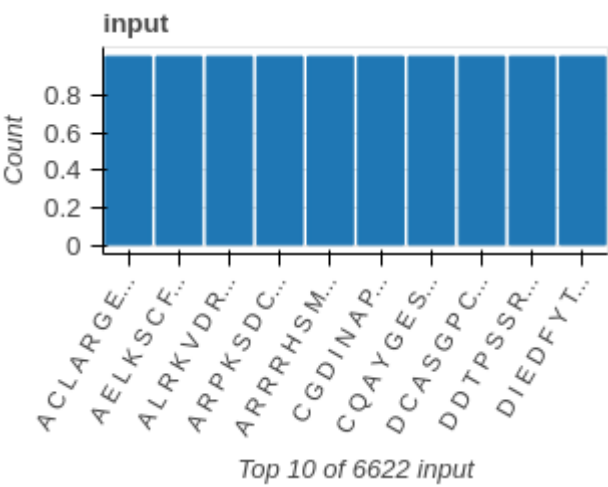
Variables

input

categorical

Show Details

Approximate Distinct Count	6622
Approximate Unique (%)	100.0%
Missing	0
Missing (%)	0.0%
Memory Size	7.4 MB



<div>loc</div> <div>categorical</div> <div>Show Details</div>	Approximate Unique (%)		0.2%	
	Missing		0	
	Missing (%)		0.0%	
	Memory Size		487.2 KB	
<div>membrane</div> <div>categorical</div> <div>Show Details</div>	Approximate Distinct Count		3	
	Approximate Unique (%)		0.0%	
	Missing		0	
	Missing (%)		0.0%	
	Memory Size		426.8 KB	
<div>labels</div> <div>categorical</div> <div>Show Details</div>	Approximate Distinct Count		2	
	Approximate Unique (%)		0.0%	
	Missing		0	
	Missing (%)		0.0%	
	Memory Size		426.8 KB	
<div>input_seq_length</div> <div>numerical</div> <div>Show Details</div>	Approximate Distinct Count	1491	Mean	557.0979
			Minimum	40
	Approximate Unique (%)	22.5%	Maximum	13100
			Zeros	0
	Missing	0	Zeros (%)	0.0%
	Missing (%)	0.0%	Negatives	0
	Infinite	0	Negatives (%)	0.0%
	Infinite (%)	0.0%		
Memory Size		103.5 KB		

Interactions

Correlations

Pearson

Spearman

KendallTau

Missing Values

Bar Chart

Spectrum

Heat Map

Dendogram

Report generated with [DataPrep](#)