

Package ‘transformer’

May 8, 2026

Title Implementation of Transformer Deep Neural Network with Vignettes

Version 0.2.0

Description Transformer is a Deep Neural Network Architecture based i.a. on the Attention mechanism (Vaswani et al. (2017) <[doi:10.48550/arXiv.1706.03762](https://doi.org/10.48550/arXiv.1706.03762)>).

License MIT + file LICENSE

Encoding UTF-8

RoxygenNote 7.2.3

Imports attention (>= 0.4.0)

Suggests covr, testthat (>= 3.0.0)

Config/testthat/edition 3

NeedsCompilation no

Author Bastiaan Quast [aut, cre] (ORCID:
<<https://orcid.org/0000-0002-2951-3577>>)

Maintainer Bastiaan Quast <bquast@gmail.com>

Repository CRAN

Date/Publication 2023-11-10 12:30:02 UTC

Contents

| | |
|------------------------|---|
| feed_forward | 2 |
| layer_norm | 2 |
| multi_head | 3 |
| row_means | 3 |
| row_vars | 4 |
| transformer | 4 |

| | |
|--------------|----------|
| Index | 6 |
|--------------|----------|

| | |
|--------------|---------------------------|
| feed_forward | <i>Feed Forward Layer</i> |
|--------------|---------------------------|

Description

Feed Forward Layer

Usage

```
feed_forward(x, dff, d_model)
```

Arguments

| | |
|---------|----------------------------------|
| x | inputs |
| dff | dimensions of feed-forward model |
| d_model | dimensions of the model |

Value

output of the feed-forward layer

| | |
|------------|----------------------------|
| layer_norm | <i>Layer Normalization</i> |
|------------|----------------------------|

Description

Layer Normalization

Usage

```
layer_norm(x, epsilon = 1e-06)
```

Arguments

| | |
|---------|--------|
| x | inputs |
| epsilon | scale |

Value

outputs of layer normalization

| | |
|------------|-------------------------------|
| multi_head | <i>Multi-Headed Attention</i> |
|------------|-------------------------------|

Description

Multi-Headed Attention

Usage

```
multi_head(Q, K, V, d_model, num_heads, mask = NULL)
```

Arguments

| | |
|-----------|-------------------------|
| Q | queries |
| K | keys |
| V | values |
| d_model | dimensions of the model |
| num_heads | number of heads |
| mask | optional mask |

Value

multi-headed attention outputs

| | |
|-----------|------------------|
| row_means | <i>Row Means</i> |
|-----------|------------------|

Description

Row Means

Usage

```
row_means(x)
```

Arguments

| | |
|---|--------|
| x | matrix |
|---|--------|

Value

vector with the mean of each of row of the input matrix

Examples

```
row_means(t(matrix(1:5)))
```

| | |
|----------|----------------------|
| row_vars | <i>Row Variances</i> |
|----------|----------------------|

Description

Row Variances

Usage

```
row_vars(x)
```

Arguments

| | |
|---|--------|
| x | matrix |
|---|--------|

Value

vector with the variance of each of row of the input matrix

Examples

```
row_vars(t(matrix(1:5)))
```

| | |
|-------------|--------------------|
| transformer | <i>Transformer</i> |
|-------------|--------------------|

Description

Transformer

Usage

```
transformer(x, d_model, num_heads, dff, mask = NULL)
```

Arguments

| | |
|-----------|----------------------------------|
| x | inputs |
| d_model | dimensions of the model |
| num_heads | number of heads |
| dff | dimensions of feed-forward model |
| mask | optional mask |

Value

output of the transformer layer

Examples

```
x <- matrix(rnorm(50 * 512), 50, 512)
d_model <- 512
num_heads <- 8
dff <- 2048

output <- transformer(x, d_model, num_heads, dff)
```

Index

feed_forward, 2

layer_norm, 2

multi_head, 3

row_means, 3

row_vars, 4

transformer, 4