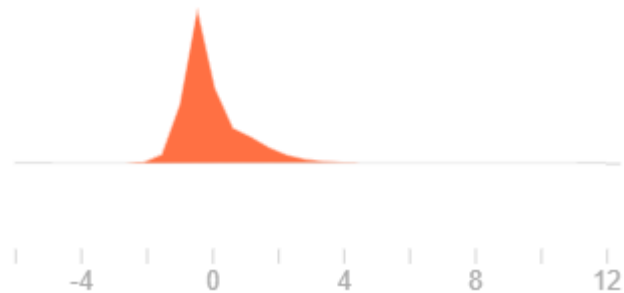


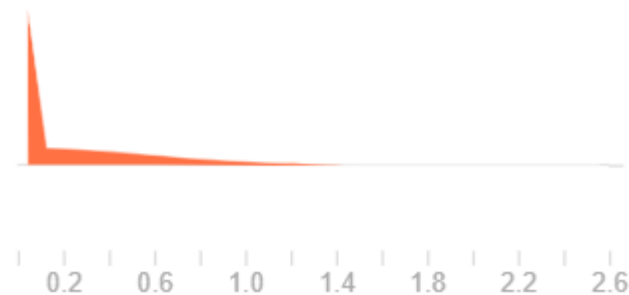
after\_resblock\_1  
tag: after\_resblock\_1

my\_model



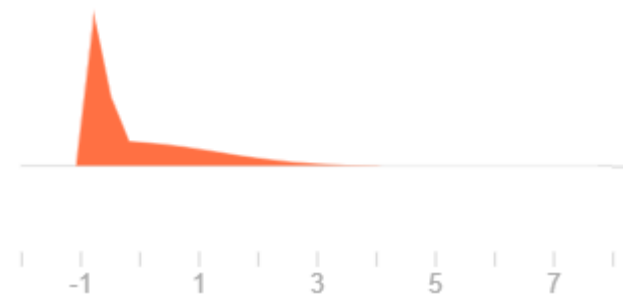
before\_norm msa\_1  
tag: before\_norm msa\_1

my\_model



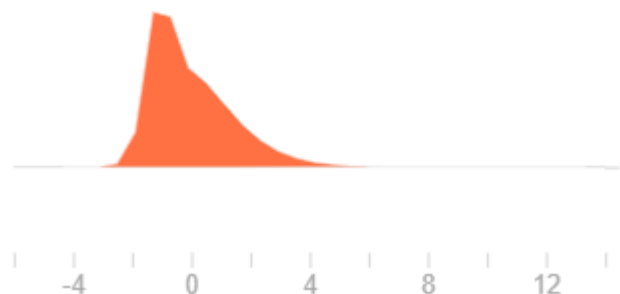
after\_norm msa\_1  
tag: after\_norm msa\_1

my\_model



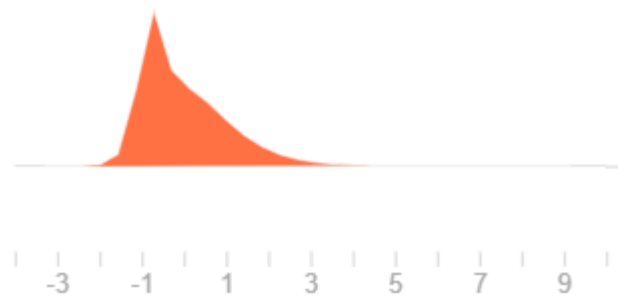
after\_msa+\_x\_1  
tag: after\_msa+\_x\_1

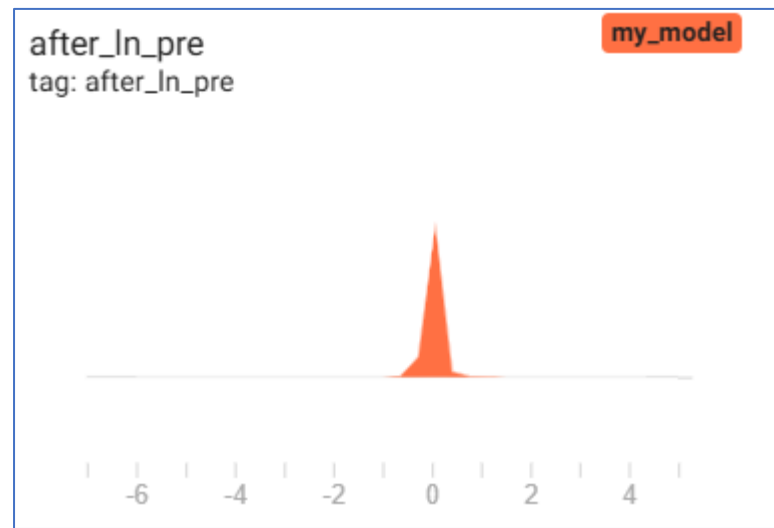
my\_model



after\_layer\_norm\_msa+\_x\_1  
tag: after\_layer\_norm\_msa+\_x\_1

my\_model





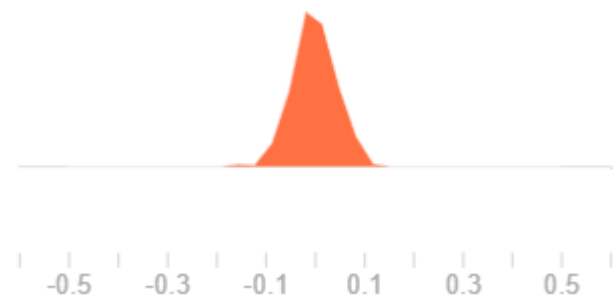
before\_norm\_token\_prefix  
tag: before\_norm\_token\_prefix

my\_model



before\_norm\_token\_suffix  
tag: before\_norm\_token\_suffix

my\_model



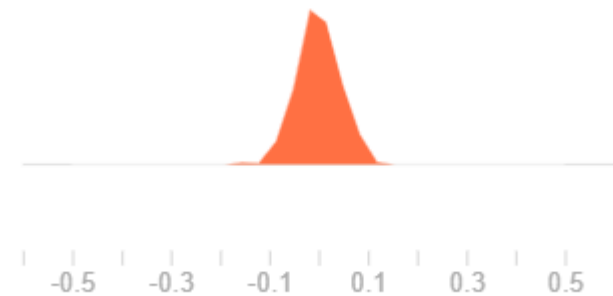
after\_norm\_token\_prefix  
tag: after\_norm\_token\_prefix

my\_model



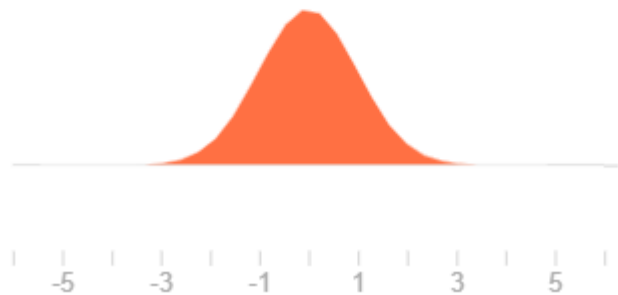
after\_norm\_token\_suffix  
tag: after\_norm\_token\_suffix

my\_model



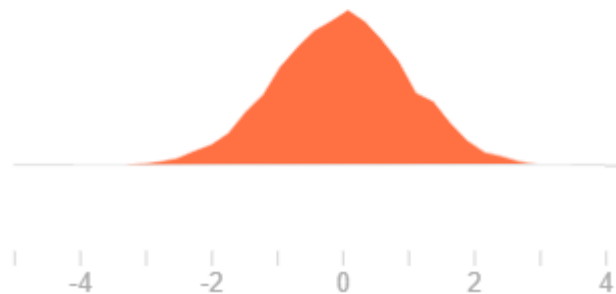
image\_features\_standardized  
tag: image\_features\_standardized

my\_model



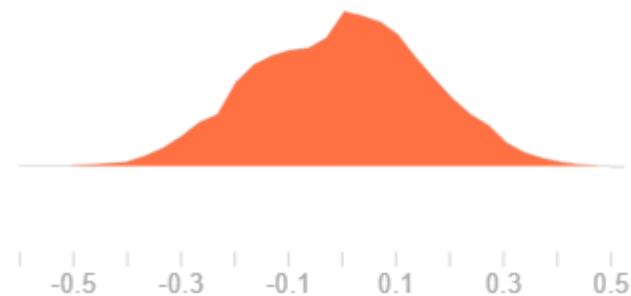
ctx\_features\_standardized  
tag: ctx\_features\_standardized

my\_model

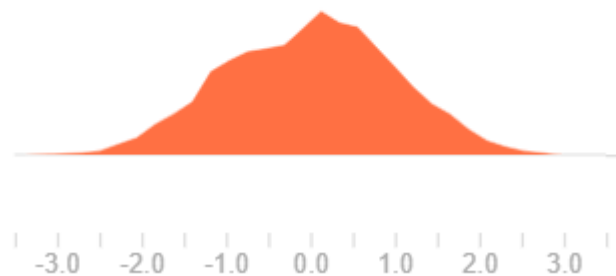


ctx\_features\_attn\_output  
tag: ctx\_features\_attn\_output

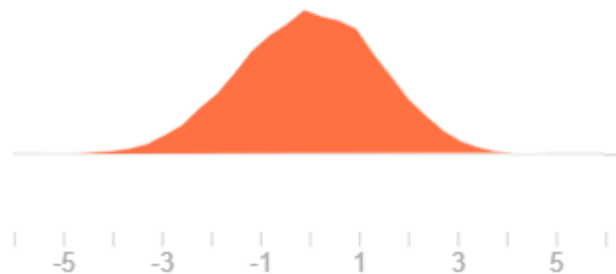
my\_model



ctx\_features\_attn\_output\_standardized  
my\_model  
tag: ctx\_features\_attn\_output\_standardized

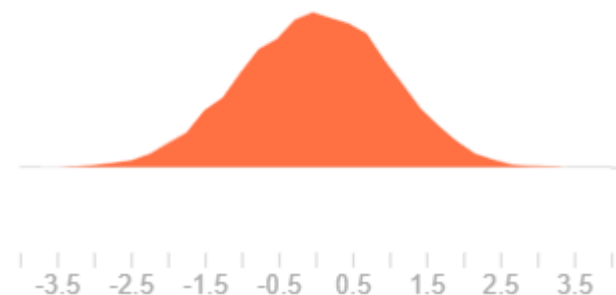


ctx\_features\_attn\_output\_standardized +  
ctx\_features\_standardized  
my\_model  
tag: ctx\_features\_attn\_output\_standardized +  
ctx\_features\_standardized



ctx\_features\_attn\_output\_final  
tag: ctx\_features\_attn\_output\_final

my\_model



final\_ctx\_features\_attn\_output  
tag: final\_ctx\_features\_attn\_output

my\_model

