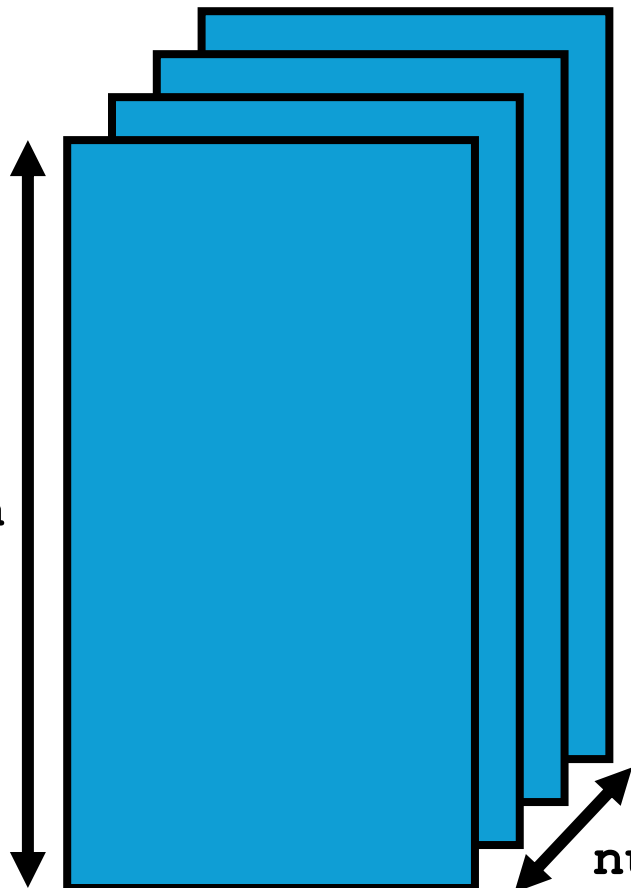


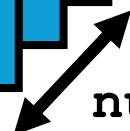
b = sentences

num\_token

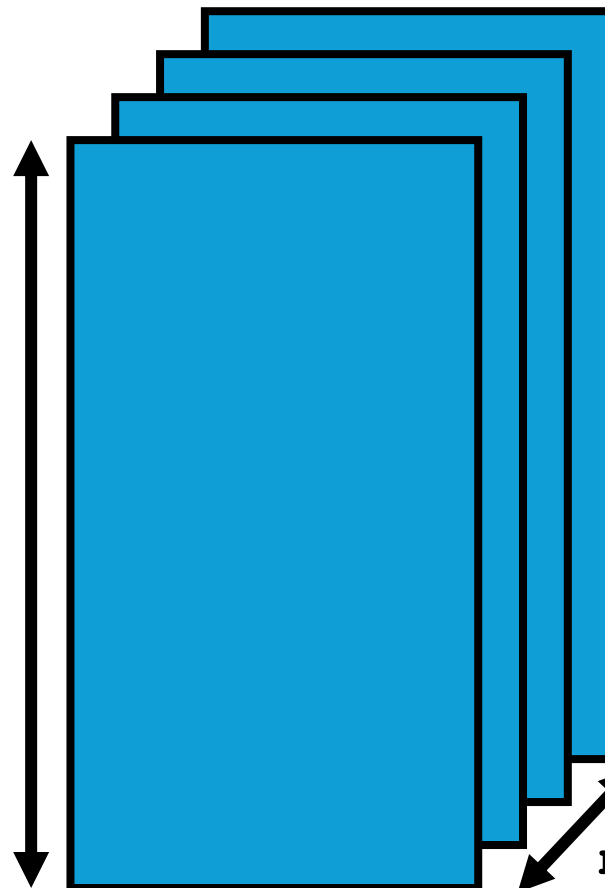


head\_dim

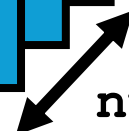
num\_token  
=  
words



num\_head



head\_dim



num\_head

=

Different Rules

=

Rule to Learn