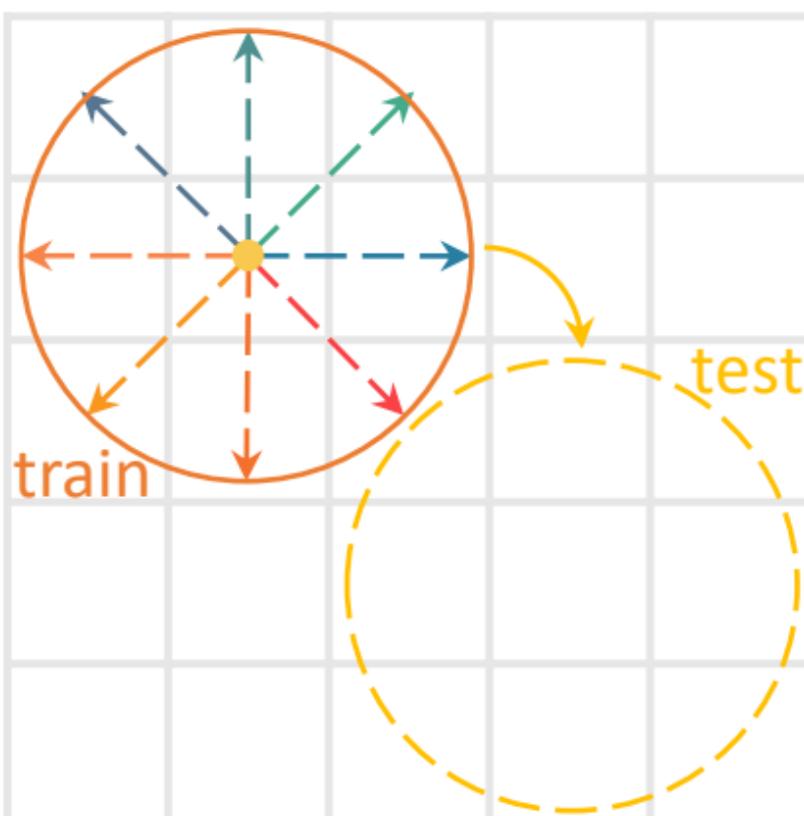


$\{W_{-1,-1}, W_{-1,0}, W_{-1,1}, W_{0,-1}, W_{0,0}, W_{0,1}, W_{1,-1}, W_{1,0}, W_{1,1}\}$

$W_{-1,-1}$	$W_{-1,0}$	$W_{-1,1}$
$W_{0,-1}$	$W_{0,0}$	$W_{0,1}$
$W_{1,-1}$	$W_{1,0}$	$W_{1,1}$



(a) Convolution

$W \in \{W^q, W^k, W^v\}$

$W$	$W$	$W$	$W$	$W$
$W$	$W$	$W$	$W$	$W$
$W$	$W$	$W$	$W$	$W$
$W$	$W$	$W$	$W$	$W$



(b) Self-attention

with positional embedding