

object (that queries)



query
transformation

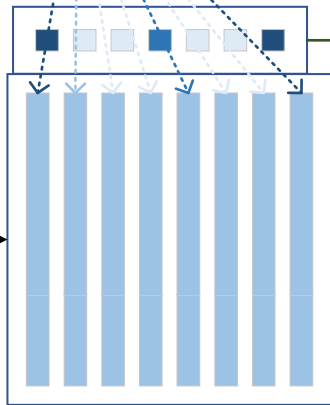
transformed query



compute
similarity
score

set (to be queried)

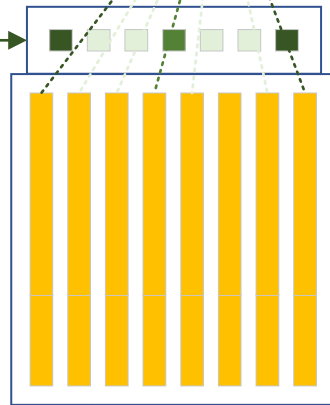
key transformation



key set

softmax

mix
according to
attention
weights



value set

value transformation

output

