

multi heads attention 在各个模态之间share weights

作者: @ry664663 | 来自: 知犀思维导图