

RL Problem

=

Decision Process Model

*e.g. MDP, POMDP,
multi-armed bandit*

+

Learning Objective

*e.g. discounted return with
specific discount factor*