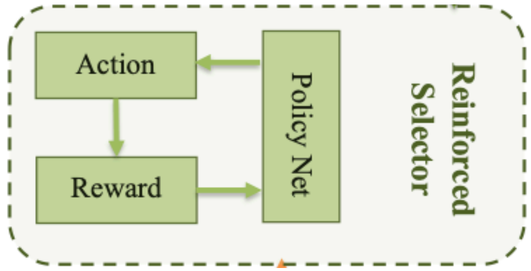
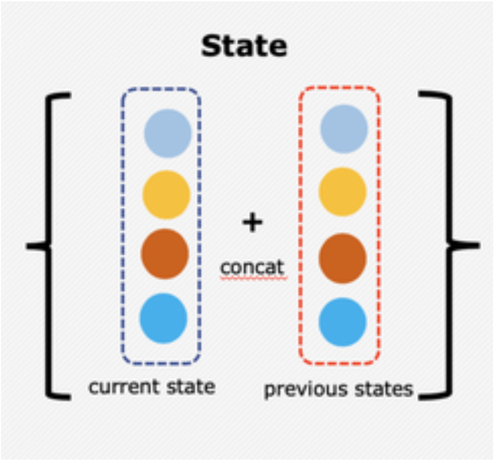
Methodology

Data Selection via Reinforcement Learning - *State*

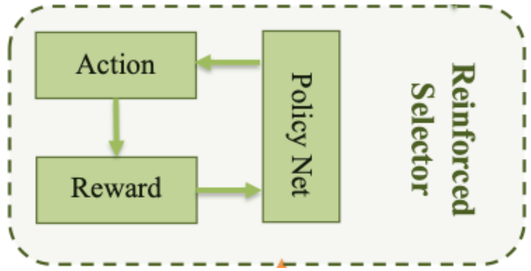


- $s_i^{(k)}$: state vector of the sample $x_i^{(k)}$
- Every action is made based on the current sample and the chosen sample, the state vector mainly consists of two components:
  - Representation of the current sample (related to data quality and diversity)
  - Average representation of the chosen samples
- The concatenation of the current state vector and the average of previous state vectors is considered as the final state vector $s_i^{(k)}$



Methodology

Data Selection via Reinforcement Learning - *State*



- The current state vector contains four elements:

  1) output probability from the annotator (quality) ▪
  2) output probability from fake news detector (quality) ▪
  3) maximum of cosine similarity between the current sample and the chosen samples (diversity) ▪
  4) weak label of the current sample (balance the distribution of classes) ▪