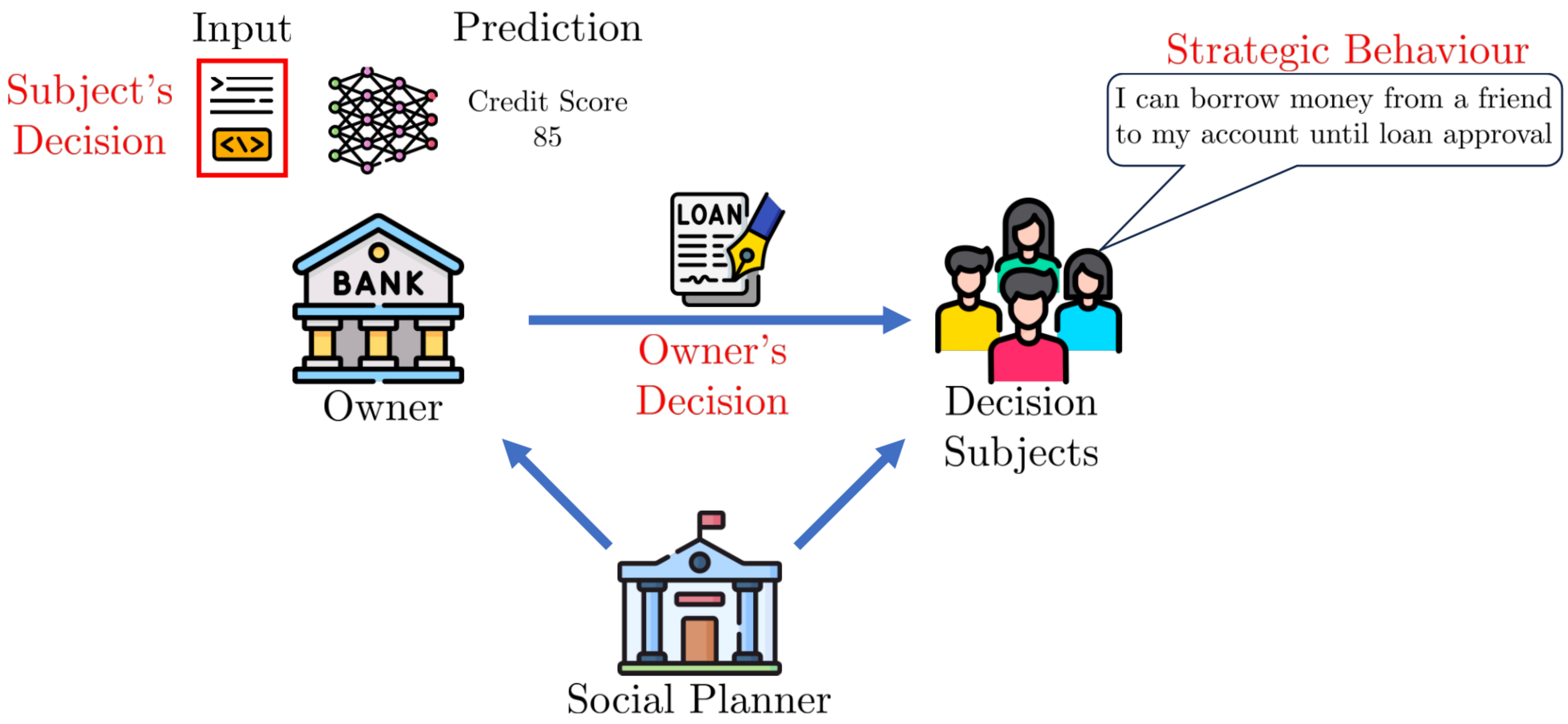
Decision Making

Stackelberg game & Bayesian persuasion in Machine Learning

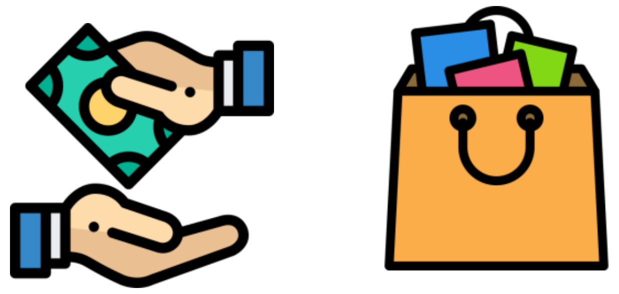
Recap

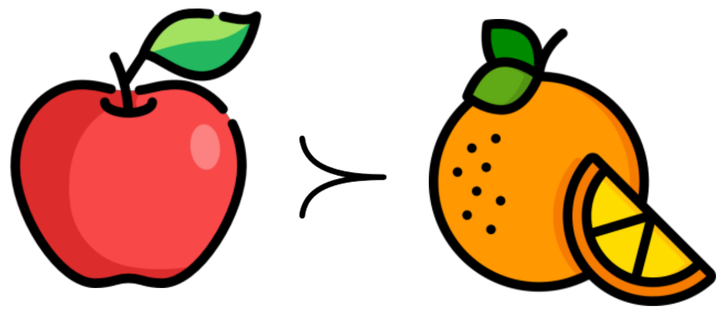

Input
Prediction
Subject's Decision
Credit Score
85
Strategic Behaviour
I can borrow money from a friend to my account until loan approval
LOAN
BANK
Owner
Owner's Decision
Decision Subjects
Social Planner

Recap

Utility and Preferences

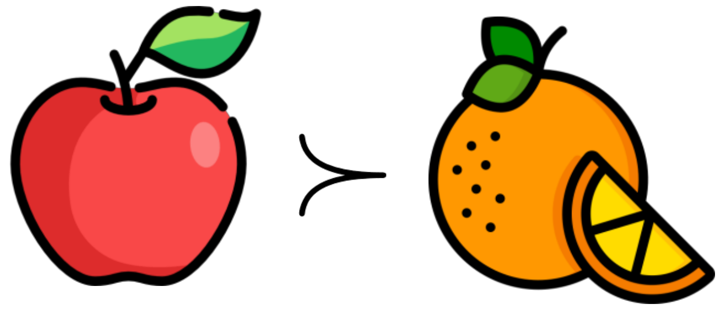
- Utility tries to assign numbers to a subjective idea of "*value*".
- Preferences are more objective since they represent ordering over items.
- Utility is intrinsic and hard to measure. Therefore it is inferred with *revealed preferences*.



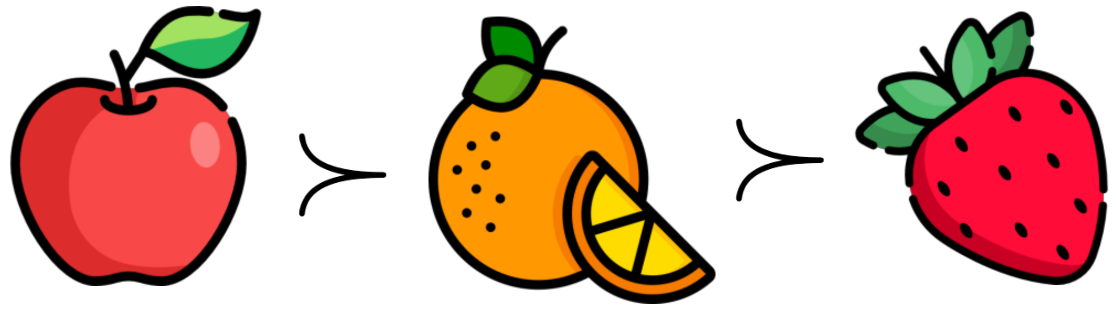


Rationality Axioms on Agent's Preferences

- Completeness: Agent must be able to compare any two items in the set i.e. $\forall a, b \in A$ either $a \succeq b$ or $b \succeq a$ or both.

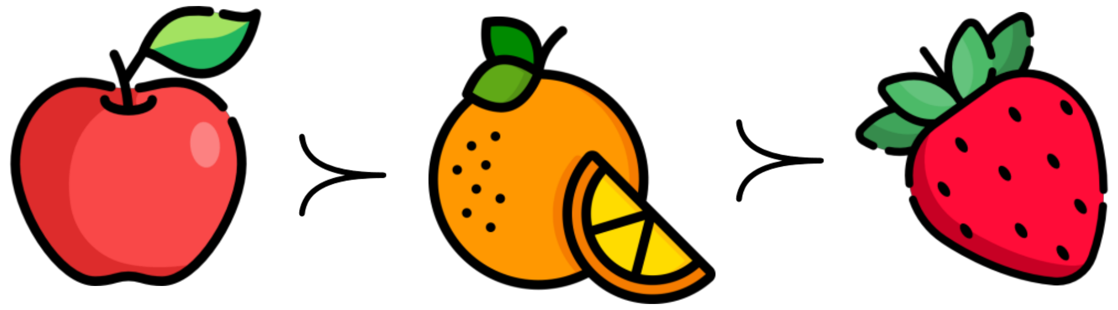


- Transitivity: $\forall a, b, c \in A$, $a \succeq b$ and $b \succeq c$ must imply $a \succeq c$.



Recap

Utility construction with rational preferences



$\succ \quad \succ$

u(apple)=1 $\quad \cdots \quad$ u(strawberry)=0

How do we construct utility in uncertainty?

Utility construction for uncertainty

Assuming a preference on distributions $\Delta(\lambda)$

- **(Completeness):** Either $q_1 \succeq q_2$ or $q_2 \succeq q_1$
- **(Trasitivity):** If $q_1 \succ q_2$ and $q_2 \succ q_1$ then $q_1 \succ q_3$
- **(Archimedean Property):** If $q_1 \succ q_2 \succ q_3$ then $\exists\ \epsilon \in (0,1)$ such that

$$(1-\epsilon)q_1 + \epsilon q_3 \succ q_2 \succ \epsilon q_1 + (1-\epsilon)q_3$$

- **(Independence of Irrelevant Alternatives)** For any q_3 and $\epsilon \in (0,1]$,

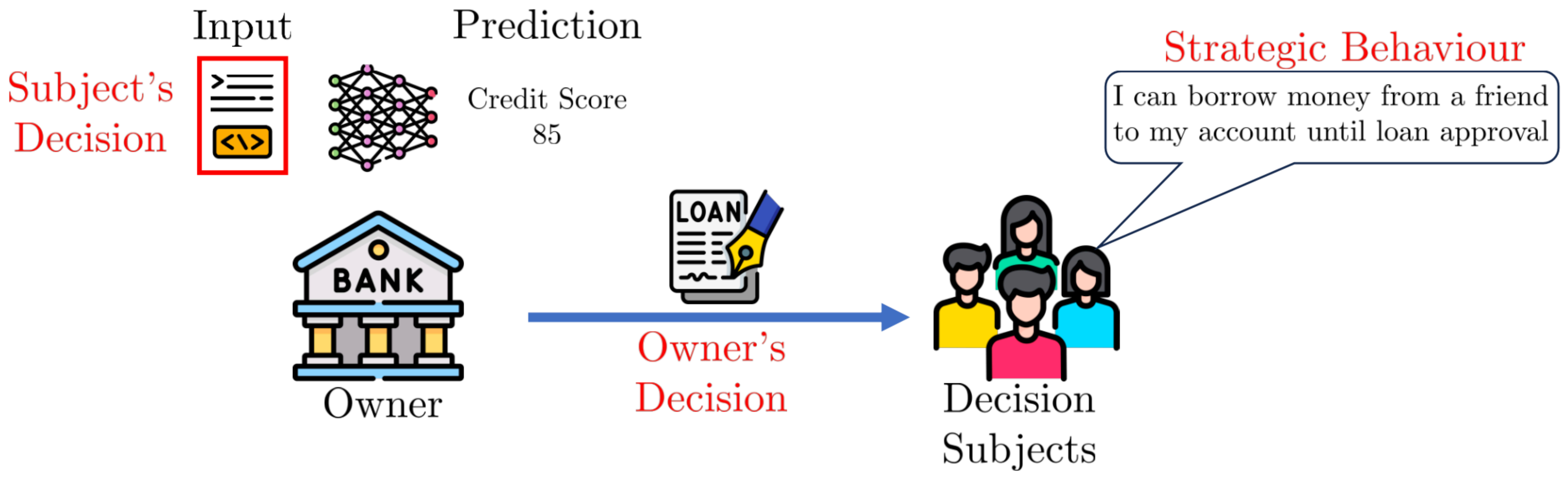
$$q_1 \succ q_2 \quad \text{iff} \quad \epsilon q_1 + (1-\epsilon)q_3 \succ \epsilon q_2 + (1-\epsilon)q_3$$

Utility

- $U : \mathscr{R} \to \mathbb{R}$, such that
  - $r_1 \succ r_2 \iff U(r_1) > U(r_2) \quad \forall r_1, r_2 \in \mathscr{R}$
  - $P_1 \succ P_2 \iff \mathbb{E}_{r \sim P_1}\left[U(r)\right] > \mathbb{E}_{r \sim P_2}\left[U(r)\right] \quad \forall P_1, P_2 \in \Delta(\mathscr{R})$

Stackelberg game

in Strategic Classification

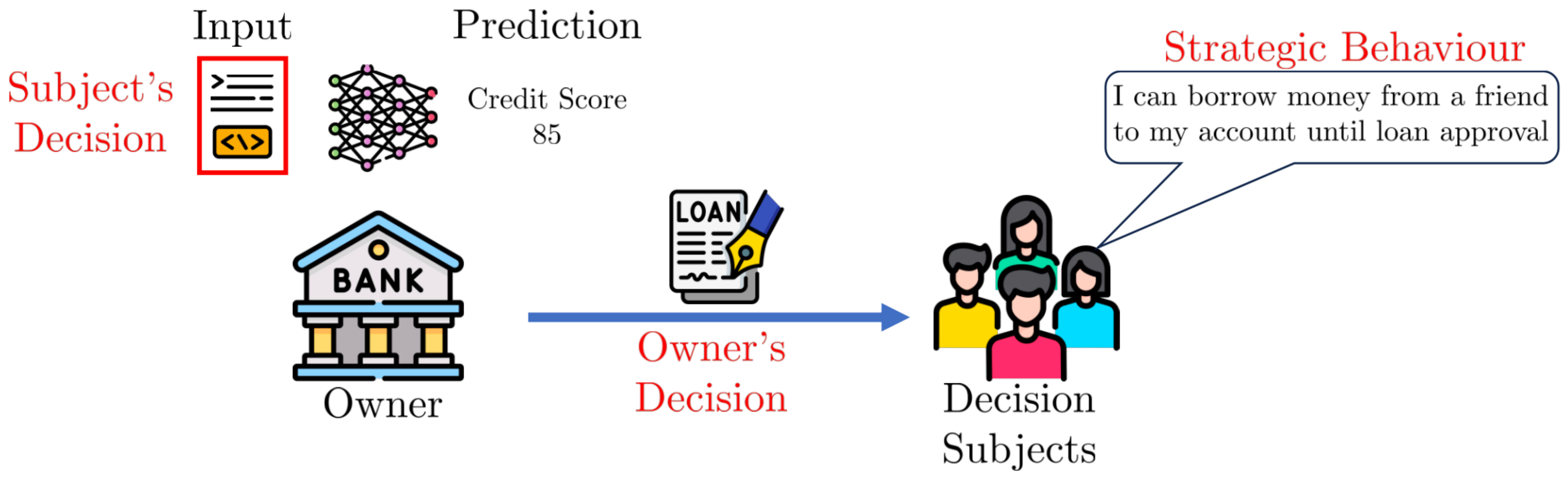


- Given a distribution D over a population $\mathscr{X}$, a cost function $c : \mathscr{X} \times \mathscr{X} \to \mathbb{R}^+$, and a target classifier $h : \mathscr{X} \to \{-1,1\}$:

  1. Decision maker (DM) publishes a classifier $f : \mathscr{X} \to \{-1,1\}$.

  2. Decision subject (or agent) observes their initial value $x_0 \sim D$ and produces a new value $x' = \psi(x_0)$, for some strategy $\psi : \mathscr{X} \to \mathscr{X}$.

Stackelberg game
in Strategic Classification



- Given a distribution D over a population $\mathscr{X}$, a cost function $c : \mathscr{X} \times \mathscr{X} \rightarrow \mathbb{R}^+$, and a target classifier $h : \mathscr{X} \rightarrow \{-1,1\}$:

    1. Decision maker (DM) publishes a classifier $f : \mathscr{X} \rightarrow \{-1,1\}$.

    2. Decision subject (or agent) observes their initial value $x_0 \sim D$ and produces a new value $x' = \psi(x_0)$, for some strategy $\psi : \mathscr{X} \rightarrow \mathscr{X}$.

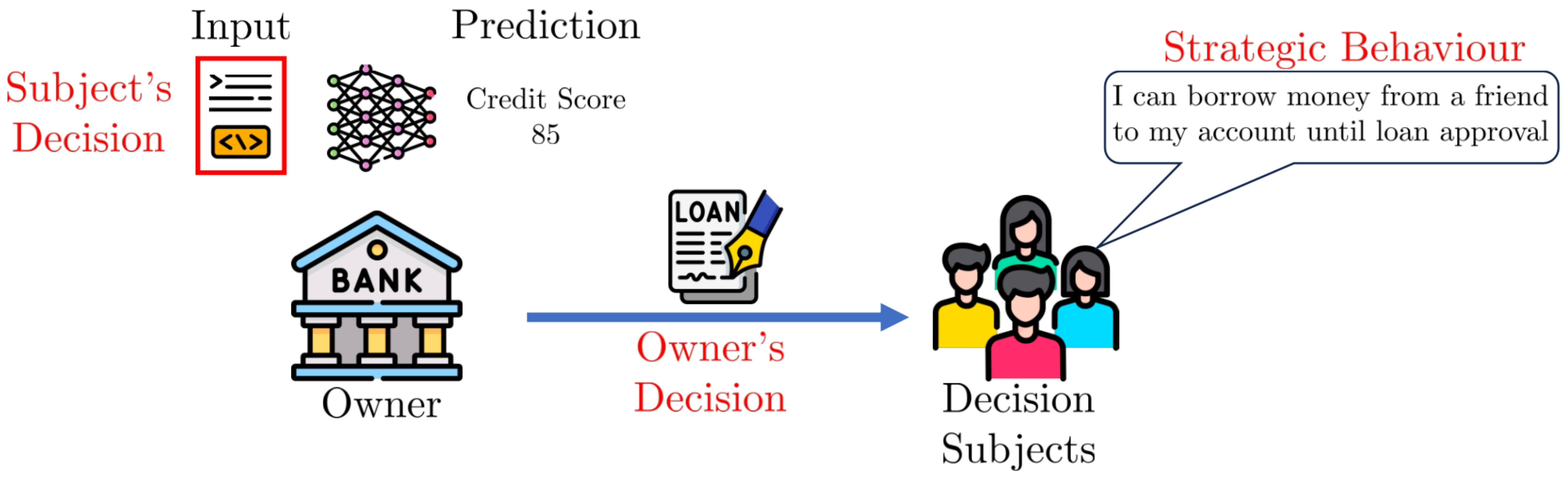
- DM's payoff: $r_{DM}(f, x_0) = 1\left\{h\left(x'\right) = f\left(x'\right)\right\}$

- Agent's payoff: $r_{Ag}(x_0, \psi) = f\left(x'\right) - c\left(x_0, x'\right)$

- DM's expected utility: $\mathbb{E}_{x_0 \sim D}\left[r_{DM}(f, x_0)\right] = \mathbb{E}_{x_0 \sim D}\left[1\left\{h\left(x'\right) = f\left(x'\right)\right\}\right]$

Stackelberg game

in Strategic Classification



- Given a distribution D over a population $\mathcal{X}$, a cost function $c : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+$, and a target classifier $h : \mathcal{X} \to \{-1,1\}$:

  1. Decision maker (DM) publishes a classifier $f : \mathcal{X} \to \{-1,1\}$.

  2. Decision subject (or agent) observes their initial value $x_0 \sim D$ and produces a new value $x' = \psi(x_0)$, for some strategy $\psi : \mathcal{X} \to \mathcal{X}$.

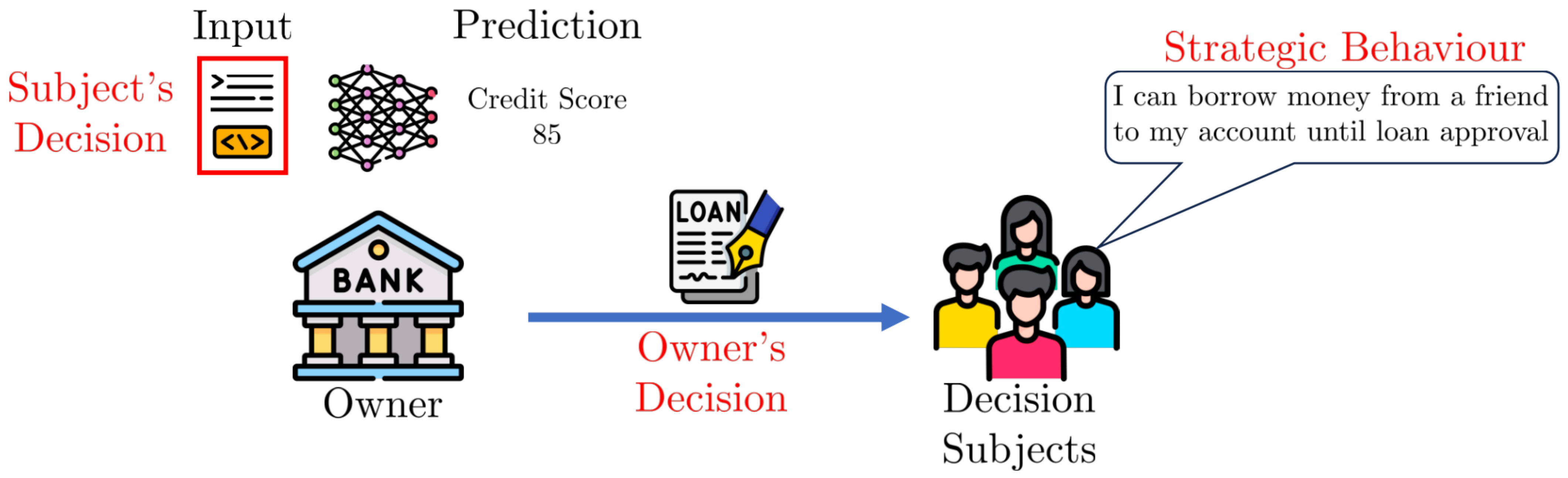
- DM's payoff: $r_{DM}(f, x_0) = 1\left\{h(x') = f(x')\right\}$

- Agent's payoff: $r_{Ag}(x_0, \psi) = f(x') - c\left(x_0, x'\right)$

- DM's expected utility: $\mathbb{E}_{x_0 \sim D}\left[r_{DM}(f, x_0)\right] = \mathbb{E}_{x_0 \sim D}\left[1\left\{h(x') = f(x')\right\}\right]$

No uncertainty in r_{Ag}!

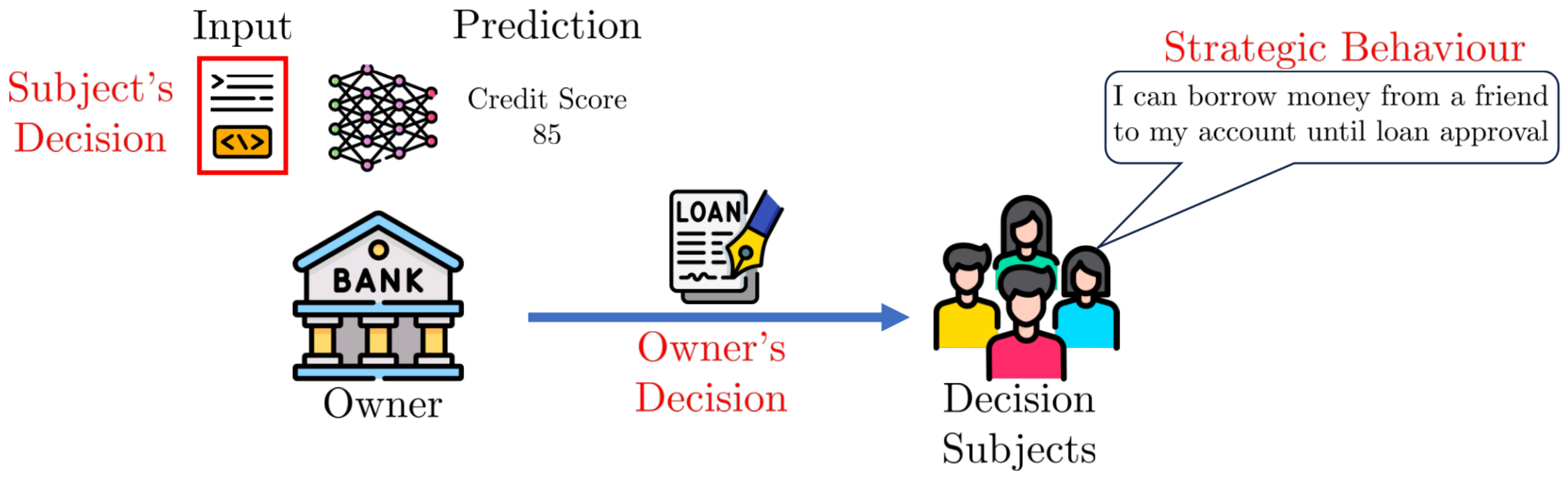
Stackelberg game
Equilibrium



- DM's expected utility: $\mathbb{E}_{x_0 \sim D}\left[r_{DM}(f, x_0)\right] = \mathbb{E}_{x_0 \sim D}\left[1\left\{h\left(x'\right) = f\left(x'\right)\right\}\right]$
- Agent's payoff: $r_{Ag}(x_0, \psi) = f\left(x'\right) - c\left(x_0, x'\right)$

- Stackelberg equilibrium (subgame perfect Nash equilibrium):
  - Agent's best response: $\psi(x_0) = \arg\max_{x \in \mathcal{X}} f(x) - c(x_0, x)$
  - DM's optimal strategy: $f^* = \arg\max_{f:\mathcal{X} \to \{-1,1\}} \mathbb{E}_{x_0 \sim D}\left[1\left\{h\left(\psi(x_0)\right) = f\left(\psi(x_0)\right)\right\}\right]$

Stackelberg game



- Stackelberg equilibrium:
  - Agent's <u>best response</u>: $\psi(x_0) = \arg\max_{x \in \mathcal{X}} f(x) - c(x_0, x)$
  - DM's optimal strategy: $f^* = \arg\max_{f:\mathcal{X} \to \{-1,1\}} \mathbb{E}_{x_0 \sim D}\left[1\left\{h\left(\psi(x_0)\right) = f\left(\psi(x_0)\right)\right\}\right]$

$\Rightarrow$ DM requires knowledge of ψ. $\Rightarrow$ Intersection between Decision Making and Machine Learning!

Bayesian Persuasion

1. The sender observes the realised state of the world $\theta \sim \Pi$, and produces a signal $\sigma \sim S(\theta)$.

2. The receiver observes σ, the signalling policy $p(\sigma \,|\, \theta)$, and the prior Π.

- For any utility function $u_{receiver}(a, \theta)$,
  - The receiver's ***subjective*** expected utility: $\mathbb{E}_{\tilde{\theta} \sim \Pi'}\left[u_{receiver}(a, \tilde{\theta}) \;\middle|\; \sigma\right]$,
  - The receiver's posterior belief: $\Pi'(\tilde{\theta} \mid \sigma) \propto p(\sigma \,|\, \theta)\Pi(\theta)$

Bayesian Persuasion

- For any utility function $u_{receiver}(a, \theta)$,
  - The receiver's *subjective* expected utility: $\mathbb{E}_{\tilde{\theta} \sim \Pi'} \left[u_{receiver}(a, \tilde{\theta}) \;\middle|\; \sigma \right]$,
  - The receiver's posterior belief: $\Pi'(\tilde{\theta} \mid \sigma) \propto p(\sigma | \theta)\Pi(\theta)$

- The receiver's optimal action: $a^* := \arg\max_{a \in \mathscr{A}} \mathbb{E}_{\tilde{\theta} \sim \Pi'} \left[u_{receiver}(a, \tilde{\theta}) \;\middle|\; \sigma \right]$
- A straightforward signalling policy $S(\theta)$ is such that:
  - $a' := \arg\max_{a \in \mathscr{A}} \mathbb{E}_{\tilde{\theta} \sim \Pi'} \left[u_{receiver}(a, \tilde{\theta}) \;\middle|\; \sigma = a' \right] \quad \forall a' : p(\sigma = a' | \theta) > 0$

Bayesian Persuasion

- For any utility function $u_{receiver}(a, \theta)$,
  - The receiver's ***subjective*** expected utility: $\mathbb{E}_{\tilde{\theta} \sim \Pi'}\left[u_{receiver}(a, \tilde{\theta}) \,\middle|\, \sigma\right]$,
  - The receiver's posterior belief: $\Pi'(\tilde{\theta} \mid \sigma) \propto p(\sigma | \theta)\Pi(\theta)$

- The receiver's optimal action: $a^* := \arg\max_{a \in \mathscr{A}} \mathbb{E}_{\tilde{\theta} \sim \Pi'}\left[u_{receiver}(a, \tilde{\theta}) \,\middle|\, \sigma\right]$
- A straightforward signalling policy $S(\theta)$ is such that:
  - $a' := \arg\max_{a \in \mathscr{A}} \mathbb{E}_{\tilde{\theta} \sim \Pi'}\left[u_{receiver}(a, \tilde{\theta}) \,\middle|\, \sigma = a'\right] \quad \forall a' : p(\sigma = a' | \theta) > 0$

⇒ The class of straightforward signalling policy $S(\theta)$ is sufficient to rationalise any receiver's behaviour.

Bayesian Persuasion for Algorithmic Recourse

- Given an agent with the initial value $x_0 \in \mathcal{X}$, the cost function $c : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+$, and the target classifier $h : \mathcal{X} \to \{-1,1\}$.
- Given a decision maker (DM) with a classifier $f_\theta :=$ sign$(x^\top \theta)$ and a (stochastic) signalling policy $S : \Theta \to \mathcal{X}$:
    1. Agent reports x_0 to the DM.
    2. DM publishes $x_r \sim S(\theta)$.
    3. Agent produces a new value $x' = \psi(x_0, a_r)$, for some strategy $\psi : \mathcal{X} \times \mathcal{X} \to \mathcal{X}$.

Bayesian Persuasion for Algorithmic Recourse

- Given an agent with the initial value $x_0 \in \mathscr{X}$, the cost function $c : \mathscr{X} \times \mathscr{X} \to \mathbb{R}^+$, and the target classifier $h : \mathscr{X} \to \{-1,1\}$.
- Given a decision maker (DM) with a classifier $f_\theta :=$ sign$(x^\top \theta)$ and a (stochastic) signalling policy $S : \Theta \to \mathscr{X}$:
    1. Agent reports x_0 to the DM.
    2. DM publishes $x_r \sim S(\theta)$.
    3. Agent produces a new value $x' = \psi(x_0, x_r)$, for some strategy $\psi : \mathscr{X} \times \mathscr{X} \to \mathscr{X}$.

- DM's payoff: $r_{DM}(x')$, e.g., $r_{DM}(x') = 1\left\{h(x') = f_\theta(x')\right\}$
- Agent's payoff: $r_{Ag}(x')$, e.g., $r_{Ag}(x') = f_\theta(x') - c\left(x_0, x'\right)$

- DM's expected utility: $\mathbb{E}_{x_r \sim S(\theta),\ \theta \sim \Pi}\left[r_{DM}(x')\right]$
- Agent's expected utility: $\mathbb{E}_{\theta \sim \Pi'}\left[r_{Ag}(x') \;\middle|\; x_r\right]$

Bayesian Persuasion for Algorithmic Recourse

- DM's expected utility: $\mathbb{E}_{a_r \sim S(\theta),\ \theta \sim \Pi}\left[r_{DM}(x')\right]$
- Agent's expected utility: $\mathbb{E}_{\theta \sim \Pi'}\left[r_{Ag}(x') \mid x_r\right]$

- Bayesian incentive-compatibility (BIC):
  - $\mathbb{E}_{\theta \sim \Pi'}\left[r_{Ag}(x' = x_r) \mid x_r\right] \geq \mathbb{E}_{\theta \sim \Pi'}\left[r_{Ag}(x' = x^\bullet) \mid x_r\right] \quad \forall x_r, x^\bullet \in \mathcal{X}.$
  - $S(\theta)$ is BIC.

Bayesian Persuasion for Algorithmic Recourse

- DM's expected utility: $\mathbb{E}_{a_r \sim S(\theta),\ \theta \sim \Pi}\left[r_{DM}(x')\right]$.
- BIC constraint: $\mathbb{E}_{\theta \sim \Pi'}\left[r_{Ag}(x' = x_r) \;\middle|\; x_r\right] \geq \mathbb{E}_{\theta \sim \Pi'}\left[r_{Ag}(x' = x^\bullet) \;\middle|\; x_r\right] \quad \forall x_r, x^\bullet \in \mathscr{X}$.

- DM's optimal strategy:

$$\max_{S} \mathbb{E}_{a_r \sim S(\theta),\ \theta \sim \Pi}\left[r_{DM}(x')\right]$$

s.t. S is BIC

Bayesian Persuasion for Algorithmic Recourse

Definition 4.1 (Equivalence Region)**.** *Two assessments* $\boldsymbol{\theta}, \boldsymbol{\theta}'$ *are* equivalent *(w.r.t.* u_{ds}*) if* $u_{ds}(a, \boldsymbol{\theta}) - u_{ds}(a', \boldsymbol{\theta}) = u_{ds}(a, \boldsymbol{\theta}') - u_{ds}(a', \boldsymbol{\theta}')$, $\forall a, a' \in \mathcal{A}$*. An* equivalence region R *is a subset of* $\boldsymbol{\Theta}$ *such that for any* $\theta \in R$*, all* θ' *equivalent to* θ *are also in* R*. We denote the set of all equivalence regions by* $\mathcal{R}$*.*

Theorem 4.2 (Optimal signaling policy)**.** *The decision maker's optimal signaling policy can be characterized by the following linear program OPT-LP:*

$$\begin{aligned}
\max_{p(\sigma=a|R), \forall a \in \mathcal{A}, R \in \mathcal{R}} \quad & \sum_{a \in \mathcal{A}} \sum_{R \in \mathcal{R}} p(R) p(\sigma = a|R) u_{dm}(a) \\
\textit{s.t.} \quad & \sum_{R \in \mathcal{R}} p(\sigma = a|R) p(R)(u_{ds}(a, R) - u_{ds}(a', R)) \geq 0, \ \forall a, a' \in \mathcal{A} \\
& \sum_{a \in \mathcal{A}} p(\sigma = a|R) = 1, \ \forall R, \quad p(\sigma = a|R) \geq 0, \ \forall R \in \mathcal{R}, a \in \mathcal{A},
\end{aligned} \tag{OPT-LP}$$