Proceedings of the Thirty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-26)
Survey Track

A Survey on Active Feature Acquisition Strategies

Linus Aronsson[1,2] and **Arman Rahbar**[1,2] and **Morteza Haghir Chehreghani**[1,2]
[1]Chalmers University of Technology
[2]University of Gothenburg
{linaro, armanr, morteza.chehreghani}@chalmers.se

Abstract

Active feature acquisition (AFA) studies how a predictive system can sequentially choose which feature values to obtain for each instance to balance predictive accuracy against feature acquisition cost. This survey provides the first unified treatment of modern AFA through a partially observable Markov decision process (POMDP) formulation, showing that most existing methods can be understood as different approximations of the same underlying sequential decision problem. The survey proposes an up-to-date taxonomy organizing AFA into four categories: (i) embedded cost-aware predictors (notably cost-sensitive decision trees and ensembles), (ii) model-based methods that plan using learned probabilistic components, (iii) model-free methods that learn acquisition policies from simulated episodes, and (iv) hybrid methods that combine the strengths of model-based and model-free approaches. We argue that this POMDP-centric view clarifies connections among existing methods and motivates more principled algorithm design. Since much prior work is heuristic and lacks formal guarantees, we also outline routes to guarantees by connecting AFA to adaptive stochastic optimization. We conclude by highlighting open challenges and promising directions for future research.

1 Introduction

Many predictive systems rely on features that are costly, slow, invasive, or privacy sensitive to obtain. In medical decision support, ordering all tests for every patient can be expensive and burdensome, and it may delay treatment [Turney, 1995]. In personalized marketing and recommender systems, collecting all behavioral signals can be costly and raises privacy concerns. In interactive troubleshooting, diagnostic queries (for example, checks, measurements, or questions) consume time and user effort. Similar constraints arise in sensor systems and robotics, where measurements require power, time, or motion [Golovin and Krause, 2011; Lauri *et al.*, 2023]. These settings motivate *active feature acquisition* (AFA), where an agent sequentially decides which feature values to acquire for each instance, and when to stop and predict, trading accuracy against feature acquisition cost.

AFA differs fundamentally from standard (static) feature selection, which commits to a single fixed subset used for all instances [Guyon and Elisseeff, 2003]. The advantage of AFA is most apparent when the informative features vary across instances. For example, suppose there is one feature that indicates which of many otherwise independent features is the only one relevant for predicting the label on the current instance. In order to be accurate on all instances, a static subset rule must include the indicator feature and all other features. In contrast, AFA learns an *adaptive* acquisition policy: it can first acquire the indicator feature and then, based on its realized value, choose the next feature accordingly, achieving the same accuracy while using far fewer features per instance [Valancius *et al.*, 2024].

At a conceptual level, AFA is a form of *optimal information acquisition*. Optimal information acquisition problems are typically instances of adaptive stochastic optimization and are often computationally intractable, which explains the prevalence of approximations such as myopic value of information rules [Golovin and Krause, 2011]. AFA connects to a broad set of related fields that study cost aware information gathering, including sequential experimental design, active learning, bandits, decision analysis, and information gathering in robotics and perception [Howard, 1966; Settles, 2009; Lauri *et al.*, 2023].

As described in Section 3, AFA can be formulated as a *partially observable Markov decision process* (POMDP) [Åström, 1965]. Existing approaches largely differ in how they approximate the optimal action-value function in (10). Methods for decision-making in the broader POMDP literature are commonly organized by when computation is performed (offline versus online planning) and by whether the environment model is assumed known or must be learned (planning with a model versus reinforcement learning approaches, including model-based and model-free variants) [Ross *et al.*, 2008]. While many AFA papers do not explicitly adopt POMDP terminology, their algorithmic choices map naturally onto these paradigms. The goal of this survey is to structure the AFA literature around the underlying POMDP formulation and the approximation decisions it entails. Accordingly, we group AFA methods into four families (Table 1): (i) *embedded cost-aware predictors*, such as cost-sensitive decision trees and en-

sembles; (ii) *model-based methods* that learn or approximate the probabilistic components of the acquisition process and use them for model-based POMDP planning; (iii) *model-free methods* that learn a policy or action-value function directly from simulated acquisition episodes; and (iv) *hybrid methods* that combine complementary elements of model-based and model-free approaches.

Historical precursors of AFA include decision theoretic medical expert systems, with the Pathfinder system as a canonical example [Heckerman *et al.*, 1992]. In parallel, cost sensitive decision trees embedded instance-wise feature selection into a tree policy and introduced cost adjusted split criteria [Núñez, 1991; Norton, 1989]. Modern AFA generalizes these ideas to arbitrary predictors and richer policy classes, and it makes the sequential decision structure explicit, enabling generic planning and reinforcement learning (RL) methods.

Finally, [Attenberg *et al.*, 2011] surveys selective data acquisition and briefly mentions AFA, and [Zois and Chelmis, 2024] surveys feature selection and acquisition across multiple settings (including AFA to some extent). To the best of our knowledge, this is the first survey focused specifically on AFA that (i) formulates AFA in detail as a POMDP, (ii) uses this framework to unify the diverse literature under a common lens, and (iii) provides a up-to-date taxonomy of AFA methods.

2 Problem Formulation

In this section, we define the AFA problem. For each instance, an agent sequentially acquires feature values (incurring costs) and decides when to stop and predict.

Notation. Let $(x, y) \sim p(\mathbf{x}, \mathbf{y})$ be the data distribution, where $x = (x_1, \ldots, x_d) \in \mathcal{X}$ and $y \in \mathcal{Y}$. In this paper, we represent random variables using bold symbols (e.g. $\mathbf{x}, \mathbf{y}$), while their possible values are shown in regular font. Acquiring feature $a \in [d] \triangleq \{1, \ldots, d\}$ reveals x_a at cost $c_a \in \mathbb{R}_+$. For $S \subseteq [d]$, let $x_S = \{x_a : a \in S\} \in \mathcal{X}_S$, and $c(S) = \sum_{a \in S} c_a$.

Predictor Under Partial Observability. A predictor must handle arbitrary subsets of observed features:

$$f : \{(S, x_S) \mid S \subseteq [d],\, x_S \in \mathcal{X}_S\} \to \mathcal{Y}. \tag{1}$$

For brevity, we write $f(x_S)$ as shorthand for $f(S, x_S)$ whenever S is clear from context. In practice, f is often *probabilistic*, meaning it outputs a distribution $p(\mathbf{y} \mid x_S)$, from which a point prediction can be obtained (for example, $\hat{y} = \arg\max_y p(y \mid x_S)$). To apply f to partially observed inputs, one typically maps (S, x_S) to a fixed-dimensional representation. A common approach is *masking*: represent the current information by a binary mask $m \in \{0, 1\}^d$ together with a full-length feature vector $\tilde{x} \in \mathbb{R}^d$, where $m_i = 1$ indicates that feature i has been acquired and $m_i = 0$ otherwise. The input to f is then a concatenation such as $(m, \tilde{x})$. The unobserved entries of $\tilde{x}$ can be filled with (i) a fixed placeholder value (for example, 0) or (ii) *imputed* values produced by an imputation rule or model, for example $\tilde{x}_j \sim p(\mathbf{x}_j \mid x_S)$.

Acquisition Process. The features to be acquired for an instance $x \in \mathcal{X}$ are determined *adaptively* by a (stochastic) policy π. Let (S, x_S) be the features acquired so far, $U = [d] \setminus S$ the remaining unobserved features, and $B \geq 0$ the acquisition budget per instance. The policy maps the current partial observation to an action $\pi(S, x_S) \in \mathcal{A}_B(S)$, where

$$\mathcal{A}_B(S) \triangleq \{a \in U :\ c(S) + c_a \leq B\} \cup \{\texttt{STOP}\} \tag{2}$$

is the set of available actions, and $\mathcal{F}_B(S) \triangleq \mathcal{A}_B(S) \setminus \{\texttt{STOP}\}$. In practice, the policy often outputs a probability distribution over actions. If π chooses $a \in \mathcal{F}_B(S)$, we observe x_a and update $S = S \cup \{a\}$ and $U = U \setminus \{a\}$. If π chooses STOP, we output $\hat{y} = f(x_S)$ and terminate. We write $\pi[x]$ for the final acquired set when the policy stops on instance x. In practice, the policy π is usually implemented via masking, similar to the predictor f.

Optimization Objective. A standard objective trades off prediction loss and acquisition cost under a budget constraint:

$$\begin{aligned} \min_{f,\pi}\ & \mathbb{E}_{\mathbf{x},\mathbf{y}}\, \mathbb{E}_{\pi}\left[\ell(f(\mathbf{x}_{\pi[\mathbf{x}]}), \mathbf{y}) + \alpha\, c(\pi[\mathbf{x}])\right] \\ \text{s.t.}\quad & c(\pi[x]) \leq B \ \text{ for all } x \in \mathcal{X}. \end{aligned} \tag{3}$$

where $\alpha \geq 0$ controls the cost–accuracy trade-off and ℓ is a loss function. The term $\alpha\, c(\pi[\mathbf{x}])$ lets acquisition cost vary across instances, so harder instances can use more resources. Thus, α induces a *soft* budget, whereas B imposes a *hard* budget. Two common special cases are $\alpha = 0$ (only a hard budget) and $B = c([d])$ (all features are acquirable, so only the soft-budget term matters). The extended version of this survey will cover other AFA formulations, but the formulation in (3) is by far the most common.

Training. Most AFA papers assume access to i.i.d. training samples $(x, y) \sim p(\mathbf{x}, \mathbf{y})$ with fully observed features and labels (costs are only incurred at test time). The predictor f and policy π are learned on this fully observed data (either jointly or separately) and then deployed on test instances from the same distribution, where features are initially unobserved and must be acquired sequentially at a cost. Only a small fraction of work studies an *online* variant in which the learner acquires features during training as well and updates f and π from streaming experience, so learning is cost-sensitive from the start rather than relying on fully observed training data (see e.g., [Kachuee *et al.*, 2019; Rahbar *et al.*, 2023; Rahbar *et al.*, 2025]).

3 AFA-POMDP Formulation

The objective in (3) can be formulated as reward maximization in an episodic, finite-horizon POMDP. We provide a formal definition of the AFA-POMDP in the extended version of this survey.

State and Action Space. Let $S \subseteq [d]$, $U = [d] \setminus S$, and $\mathcal{S} = \{(S, x, y) : S \in 2^{[d]}, x \in \mathcal{X}, y \in \mathcal{Y}\}$. Since $x = (x_S, x_U)$, we write states as $(S, x_S, x_U, y) \in \mathcal{S}$, with observed part (S, x_S) and latent part (x_U, y). Given budget B, the available actions are $\mathcal{A}_B(S)$ from (2).

Transition and Observation Model. Given state (S, x_S, x_U, y), both the transition and observation model are deterministic. If $a \in \mathcal{F}_B(S)$ is chosen, x_a is observed from the latent component x_U and the next state is $(S \cup \{a\}, x_{S \cup \{a\}}, x_{U \setminus \{a\}}, y)$. If STOP is chosen, the episode terminates and the predictor outputs a label.

Belief State and Observation Probability. The belief state corresponds to our current belief over the latent part (x_U, y) given the observed information (S, x_S). We write

$$b(x_S) \triangleq p(\mathbf{x}_U, \mathbf{y} \mid x_S). \tag{4}$$

The probability of observing $o \in \mathcal{X}_a$ after acquiring $a \in \mathcal{F}_B(S)$ in the current belief state is

$$p(\mathbf{x}_a = o \mid x_S) = \mathbb{E}_{\mathbf{x}_{U\setminus\{a\}}, \mathbf{y}|x_S}[p(\mathbf{x}_a = o \mid x_S, \mathbf{x}_{U\setminus\{a\}}, \mathbf{y})]. \tag{5}$$

Belief Update. After acquiring $a \in \mathcal{F}_B(S)$ and observing $o \in \mathcal{X}_a$ in current belief state $b(x_S)$, let

$$b(x_S, a, o) \triangleq p(\mathbf{x}_{U\setminus\{a\}}, \mathbf{y} \mid x_S, \mathbf{x}_a = o) \tag{6}$$

denote the updated belief given the new evidence. This belief can be obtained via Bayesian updating. The updated label belief is

$$p(\mathbf{y} \mid x_S, \mathbf{x}_a = o) = \frac{p(\mathbf{x}_a = o \mid x_S, \mathbf{y}) p(\mathbf{y} \mid x_S)}{p(\mathbf{x}_a = o \mid x_S)}. \tag{7}$$

(7) is the standard belief update rule in POMDPs [Åström, 1965]. The Bayesian update of the full belief in (6) will be given in the extended version of this survey.

Belief-MDP and Sufficient Statistic. Any POMDP can be rewritten as a fully observable *belief-MDP*. For the exact AFA-POMDP, when the full belief is explicitly represented, the pair $(S, b(x_S))$ is a sufficient statistic for the current observation history (S, x_S) and defines the belief-MDP state: Bayesian updating folds observed values into $b(x_S)$, so the value function need not separately store x_S. The set S determines feasible actions and acquired costs, while $b(x_S)$ determines observation probabilities, belief updates, and stopping rewards.

Reward. For a state $s = (S, x_S, x_U, y)$, define the immediate reward $R(s, a) = -\alpha c_a$ for each acquisition action $a \in \mathcal{F}_B(S)$, and $R(s, \texttt{STOP}) = -\ell(f(x_S), y)$ for stopping. The expected stopping reward in belief state $(S, b(x_S))$ is

$$\begin{aligned} r(S, b(x_S), \texttt{STOP}) &= \mathbb{E}_{\mathbf{x}_U, \mathbf{y}|x_S}[R((S, x_S, \mathbf{x}_U, \mathbf{y}), \texttt{STOP})] \\ &= \mathbb{E}_{\mathbf{y}|x_S}[-\ell(f(x_S), \mathbf{y})]. \end{aligned} \tag{8}$$

Similarly, for any acquisition action $a \in \mathcal{F}_B(S)$, $r(S, b(x_S), a) = \mathbb{E}_{\mathbf{x}_U, \mathbf{y}|x_S}[R((S, x_S, \mathbf{x}_U, \mathbf{y}), a)] = -\alpha c_a$. Here $f(x_S)$ represents the predictive distribution $p(\mathbf{y} \mid x_S)$ in the Bayes-optimal case. In the exact POMDP, this distribution is the label marginal of $b(x_S)$ and can be maintained by Bayesian updating as in (7). In approximate implementations that query a predictive model directly, the computational state must still retain x_S (or an equivalent encoding) to evaluate $f(x_S)$. Under log loss this gives $r(S, b(x_S), \texttt{STOP}) = -H(\mathbf{y} \mid x_S)$, while under 0–1 loss it gives $r(S, b(x_S), \texttt{STOP}) = \max_y p(y \mid x_S)$.

Optimal Value Function of Belief-MDP. We show the k-step truncated value function V_k over the belief-MDP state $(S, b(x_S))$, where STOP is forced after k selections:

$$\begin{aligned} V_k(S, b(x_S)) = \max \Big\{ \underbrace{V_0(S, b(x_S))}_{a=\texttt{STOP}}, \max_{a \in \mathcal{F}_B(S)} \Big\{ -\alpha c_a + \\ \mathbb{E}_{\mathbf{x}_a|x_S}[V_{k-1}(S \cup \{a\}, b(x_S, a, \mathbf{x}_a))] \Big\} \Big\}. \end{aligned} \tag{9}$$

Here $V_0(S, b(x_S)) = r(S, b(x_S), \texttt{STOP})$, $p(\mathbf{x}_a \mid x_S)$ is given by (5), and $b(x_S, a, \mathbf{x}_a)$ is the updated belief in (6).

Equivalence to the AFA Objective. Consider an acquisition episode for $(x, y) \sim p(\mathbf{x}, \mathbf{y})$, and let T be the first step where STOP is chosen. The undiscounted return is $G_\pi(x, y) = \sum_{t=1}^{T-1}(-\alpha c_{a_t}) + R(s_T, \texttt{STOP})$, where $s_t = (S_t, x_{S_t}, x_{U_t}, y)$ is the state at step t. Hence $-\mathbb{E}_{\mathbf{x},\mathbf{y}}\mathbb{E}_\pi[G_\pi(\mathbf{x}, \mathbf{y})]$ is exactly the objective in (3); maximizing expected return is therefore equivalent to minimizing the AFA objective. This connection was first made in [Dulac-Arnold *et al.*, 2012].

Model-Free vs. Model-Based Methods. Every AFA method needs a predictor for $\mathbf{y}$ to evaluate the stopping reward, typically through $p(\mathbf{y} \mid x_S)$. We use *model-based* to denote methods that also explicitly model the unobserved features, e.g., $p(\mathbf{x}_U \mid x_S)$ or the marginals $p(\mathbf{x}_a \mid x_S)$ that determine the transition dynamics in (9). Exact POMDP planning is expressed in terms of the full belief $b(x_S)$ over $(\mathbf{x}_U, \mathbf{y})$, which the Bellman recursion uses through the next-observation distribution $p(\mathbf{x}_a \mid x_S)$ and the predictive distribution $p(\mathbf{y} \mid x_S)$ needed for the stopping reward. These quantities can be obtained in two ways: (i) By marginalizing the full joint belief $p(\mathbf{x}_U, \mathbf{y} \mid x_S)$, which is maintained via sequential Bayesian updating (see (7)). (ii) By using predictive models that take x_S (or an equivalent history encoding) as input to directly predict the marginals $p(\mathbf{x}_a \mid x_S)$ and $p(\mathbf{y} \mid x_S)$. *Model-free* methods do not explicitly model the feature dynamics, but instead learn a policy directly from acquisition rollouts.

AFA as an MDP. As in any POMDP, the full action-observation history induces a belief state and can be used to define a fully observable history-MDP. In AFA, this history has the compact form (S, x_S): it records which features have been acquired and their values. Thus the corresponding MDP has states $(S, x_S) \in \{(O, x_O) : O \subseteq [d], x_O \in \mathcal{X}_O\}$, and each state induces a unique belief $b(x_S)$. Its value function is exactly as in (9), with the belief terms understood as functions of the current MDP state (S, x_S). Much prior AFA work effectively uses this MDP view, often without making the underlying POMDP explicit; in this survey, we make the POMDP formulation explicit because it clarifies the role of beliefs and the probabilistic quantities required for planning.

Reduced Sufficient Statistics under Conditional Independence. Under a Naive Bayes assumption, $p(\mathbf{x}_a \mid x_S, y) = p(\mathbf{x}_a \mid y)$, so x_S affects future decisions only through the posterior $p(\mathbf{y} \mid x_S)$. The information state can then be reduced from $(S, b(x_S))$ to $(S, p(\mathbf{y} \mid x_S))$, which greatly improves tractability and explains why conditional independence assumptions are common in AFA [Liyanage and Zois, 2021].

Motivation for POMDP View of AFA. The MDP view over partial observations is conceptually simple, but highly intractable: the number of states (S, x_S) grows combinatorially with d and becomes effectively infinite for continuous features. The POMDP perspective helps mitigate this by making the belief dynamics of sequential acquisition explicit and by clarifying which probabilistic quantities, such as $p(\mathbf{x}_a \mid x_S)$ and $p(\mathbf{y} \mid x_S)$, must be estimated. It also exposes distributional conditions on $p(\mathbf{x}, \mathbf{y})$ under which reduced sufficient statistics can be used, as in the preceding paragraph, thereby

reducing the search space for planning. It also emphasizes that planning can focus on reachable belief states along observed trajectories, rather than every possible subset-value pair. For discretized features, the AFA-POMDP admits the standard piecewise-linear structure of value functions over beliefs [Liyanage and Zois, 2021], which connects AFA to point-based and online POMDP solvers [Ross *et al.*, 2008].

Properties of the AFA-POMDP. The AFA-POMDP has *mixed observability*: the agent observes (S, x_S), while $(\mathbf{x}_U, \mathbf{y})$ remains latent. Thus, AFA belongs to the class of mixed-observability MDPs (MOMDPs), where the belief only needs to track the latent components. It is also an information-gathering POMDP in a passive environment, since actions reveal features but do not alter the underlying instance. A distinctive property of the AFA-POMDP is that different belief marginals play different roles: the belief over unobserved features determines the observation distribution, and hence the transition dynamics of the belief-MDP in (9), while the label marginal determines the stopping reward in (8). Thus the reward is belief-dependent, and maximizing terminal reward often corresponds to reducing uncertainty in $p(\mathbf{y} \mid x_S)$. This interpretation is not automatic in general POMDPs, where rewards and transitions need not decompose around information about the current belief over states. Finally, the explicit `STOP` action gives action-based termination, so the undiscounted objective can be viewed as a stochastic shortest path problem.

4 Myopic vs. Non-Myopic AFA

A policy can be derived from the value function V_k in (9) by defining the corresponding action-value function over $(S, b(x_S))$. For $a = \texttt{STOP}$, $Q_k(S, b(x_S), \texttt{STOP}) = V_0(S, b(x_S))$ for all k. Then, for any $a \in \mathcal{F}_B(S)$,

$$\begin{aligned} Q_k(S, b(x_S), a) &= -\alpha c_a \\ &\quad + \mathbb{E}_{\mathbf{x}_a \mid x_S}\big[V_{k-1}(S \cup \{a\}, b(x_S, a, \mathbf{x}_a))\big]. \end{aligned} \tag{10}$$

An optimal (deterministic) k-step truncated policy is then $\pi_k(S, b(x_S)) \in \arg\max_{a \in \mathcal{A}_B(S)} Q_k(S, b(x_S), a)$. All AFA methods effectively amount to approximating the action-value function $Q \approx Q_d$ (or directly the policy π_d for policy-gradient methods) by making different modeling and computational choices. A natural approximation is $Q_k \approx Q_d$ for $k < d$.

Myopic AFA. The myopic approximation corresponds to the 1-step truncated action-value function Q_1, replacing the continuation value $V_{k-1}(S \cup \{a\}, b(x_S, a, \mathbf{x}_a))$ in (10) with the value of stopping immediately after acquiring a. Concretely, $Q_1(S, b(x_S), \texttt{STOP}) = r(S, b(x_S), \texttt{STOP})$, and for any $a \in \mathcal{F}_B(S)$,

$$\begin{aligned} Q_1(S, b(x_S), a) &= -\alpha c_a \\ &\quad + \mathbb{E}_{\mathbf{x}_a \mid x_S}\big[r(S \cup \{a\}, b(x_S, a, \mathbf{x}_a), \texttt{STOP})\big]. \end{aligned} \tag{11}$$

Computing Q_1 requires no planning beyond the marginalization over $\mathbf{x}_a$. Assuming the log-loss stopping reward $r(S, b(x_S), \texttt{STOP}) = -H(\mathbf{y} \mid x_S)$, the benefit of acquiring feature $a \in \mathcal{F}_B(S)$ over stopping is

$$\begin{aligned} &Q_1(S, b(x_S), a) - Q_1(S, b(x_S), \texttt{STOP}) \\ &= H(\mathbf{y} \mid x_S) - \mathbb{E}_{\mathbf{x}_a \mid x_S}[H(\mathbf{y} \mid \mathbf{x}_a, x_S)] - \alpha c_a \\ &= I(\mathbf{y}; \mathbf{x}_a \mid x_S) - \alpha c_a. \end{aligned} \tag{12}$$

$I(\mathbf{y}; \mathbf{x}_a \mid x_S)$ is the conditional mutual information (CMI) between $\mathbf{x}_a$ and $\mathbf{y}$. Hence, the myopic CMI policy stops ($a = \texttt{STOP}$) if $I(\mathbf{y}; \mathbf{x}_a \mid x_S) - \alpha c_a \le 0$ for all $a \in \mathcal{F}_B(S)$, and selects the feasible feature with largest information gain (relative to its cost) about the label $\mathbf{y}$ otherwise. The myopic CMI policy is one of the most common heuristics in the AFA literature (often in combination with the Naive Bayes assumption), and often performs well. Other variants of this heuristic exist (e.g., $I(\mathbf{y}; \mathbf{x}_a \mid x_S)/c_a$ [Gadgil *et al.*, 2024; Golovin and Krause, 2011]).

Non-Myopic AFA. The optimal action-value function in (10) and the myopic heuristic in (11) are at the two extreme ends of policies for AFA. However, it is often of interest to find a policy between these extreme ends that finds a balance between (i) computation time, and (ii) performance. To understand what makes a policy non-myopic (i.e., planning beyond a one-step lookahead), we consider the two-step truncated action-value function under log loss. For $a \in \mathcal{F}_B(S)$, we have

$$\begin{aligned} Q_2(S, b(x_S), a) = -\alpha c_a + \mathbb{E}_{\mathbf{x}_a \mid x_S}[\max\{-H(\mathbf{y} \mid x_S, \mathbf{x}_a), \\ \max_{j \in \mathcal{F}_B(S \cup \{a\})} (-\alpha c_j + \mathbb{E}_{\mathbf{x}_j \mid x_S, \mathbf{x}_a}[-H(\mathbf{y} \mid x_S, \mathbf{x}_a, \mathbf{x}_j)])\}]. \end{aligned} \tag{13}$$

(13) compares stopping after observing $\mathbf{x}_a$, with reward $-H(\mathbf{y} \mid x_S, \mathbf{x}_a)$, against acquiring a second feature j and then stopping, with expected reward $\mathbb{E}_{\mathbf{x}_j \mid x_S, \mathbf{x}_a}[-H(\mathbf{y} \mid x_S, \mathbf{x}_a, \mathbf{x}_j)]$ minus the acquisition cost αc_j. Equivalently, the incremental value of the second acquisition relative to stopping is $I(\mathbf{y}; \mathbf{x}_j \mid x_S, \mathbf{x}_a) - \alpha c_j$. The nested entropy term $\mathbb{E}_{\mathbf{x}_a \mid x_S} \mathbb{E}_{\mathbf{x}_j \mid x_S, \mathbf{x}_a}[H(\mathbf{y} \mid x_S, \mathbf{x}_a, \mathbf{x}_j)]$ makes explicit that non-myopic planning must account for interactions between a candidate feature $a \in U$ and remaining unobserved features in $U \setminus \{a\}$. For example, two features $a \neq j$ can each be individually uninformative (small $I(\mathbf{y}; \mathbf{x}_a \mid x_S)$ and $I(\mathbf{y}; \mathbf{x}_j \mid x_S)$) yet be jointly decisive, meaning $H(\mathbf{y} \mid x_S, \mathbf{x}_a, \mathbf{x}_j)$ can drop close to zero, an effect the myopic one-step heuristic in (11) cannot capture. This motivates intermediate heuristics that condition on a small lookahead group O containing a and score a by

$$\begin{aligned} &\mathbb{E}_{\mathbf{x}_{O \setminus \{a\}} \mid x_S}[I(\mathbf{y}; \mathbf{x}_a \mid x_S, \mathbf{x}_{O \setminus \{a\}})] \\ &\quad = I(\mathbf{y}; \mathbf{x}_O \mid x_S) - I(\mathbf{y}; \mathbf{x}_{O \setminus \{a\}} \mid x_S). \end{aligned} \tag{14}$$

where $a \in O \subseteq U$. This score reasons about jointly informative feature groups by averaging the marginal utility of acquiring a over sampled partial completions $\mathbf{x}_{O \setminus \{a\}}$ of other unobserved features. It can also be viewed as a non-adaptive approximation to the value function, where a lookahead group is fixed or sampled rather than chosen adaptively after each observation. This will be discussed in detail in the extended version of this survey. The main difficulty is that selecting the subsets O to consider (and how to sample them) is problem-dependent and can strongly affect performance.

5 AFA Methods

As discussed in Section 3, AFA is naturally a POMDP, and existing methods mainly differ in how they approximate the

Category	Papers
Embedded Cost-Aware Predictors	
Cost-sensitive decision trees	[Norton, 1989; Tan, 1990; Núñez, 1991; Turney, 1995; Ling *et al.*, 2004; Yang *et al.*, 2006; Sheng and Ling, 2006; Ling *et al.*, 2006]
Cost-sensitive ensembles	[Xu *et al.*, 2012; Nan *et al.*, 2015; Peter *et al.*, 2017]
Model-Based Methods	
Myopic generative heuristics	[Geman and Jedynak, 1996; Chai *et al.*, 2004; Ji and Carin, 2007; Ma *et al.*, 2019; Rangrej and Clark, 2021; Chattopadhyay *et al.*, 2023b]
Non-myopic generative heuristics	[Bilgic and Getoor, 2007; Valancius *et al.*, 2024; Rahbar *et al.*, 2023; Rahbar *et al.*, 2025; Huang *et al.*, 2025; Norcliffe *et al.*, 2025]
Dynamic programming	[Maliah and Shani, 2018; Liyanage and Zois, 2021]
Search	[Bayer-Zubek and Dietterich, 2005]
Model-Free and Hybrid Methods	
Myopic discriminative heuristics	[He *et al.*, 2012; Covert *et al.*, 2023; Ghosh and Lan, 2023; Chattopadhyay *et al.*, 2023a; Gadgil *et al.*, 2024]
Non-myopic discriminative heuristics	[He *et al.*, 2016; Guney *et al.*, 2025]
Model-free RL	[Dulac-Arnold *et al.*, 2012; Rückstieß *et al.*, 2013; Shim *et al.*, 2018; Kachuee *et al.*, 2019; Janisch *et al.*, 2020; Aronsson and Chehreghani, 2026]
Model-based RL (hybrid)	[Zannone *et al.*, 2019; Li and Oliva, 2021]

Table 1: Taxonomy of AFA methods (Sections 5.1–5.3).

optimal action-value in (10). We group existing AFA methods into four categories: (i) *embedded cost-aware predictors*, where the model structure implicitly defines an acquisition trajectory; (ii) *model-based methods*, which estimate the probabilistic quantities of the POMDP (i.e., $p(\mathbf{x}_U \mid x_S)$) and then find a policy via model-based planning; (iii) *model-free methods*, which learn a policy directly from acquisition rollouts; and (iv) *hybrid methods*, which combine acquisition rollouts with a learned model to improve stability. See Table 1 for a taxonomy of all papers. This taxonomy also parallels standard POMDP solution paradigms [Ross *et al.*, 2008; Lauri *et al.*, 2023] (often not stated in AFA papers). Some approaches are primarily *offline* (they learn or precompute a policy from training data or a learned model), while others are *online* (they plan at test time from the current state). In addition, some are myopic while others are non-myopic by planning over a longer future horizon.

5.1 Embedded Cost-Aware Predictors

Any deterministic AFA policy π can be represented as a decision tree [Quinlan, 1993]. At test time, a decision tree evaluates only a subset of features, thereby implementing adaptive, instance-wise feature acquisition within the model itself. Consequently, decision-tree learners that explicitly incorporate feature costs yield valid policies for AFA. This section covers such cost-aware decision tree methods. All methods in this section also fall under one of the broader (POMDP-based) categories introduced later, since they can be interpreted as approximations to the underlying POMDP for AFA (as any AFA policy can). We nevertheless present them as a separate category to emphasize that they are specific to decision tree learning. Moreover, in contrast to other AFA methods introduced later, they inherit familiar limitations of standard decision trees (e.g., overfitting and discretizing continuous features). Notably, they are rarely included in modern AFA benchmarks.

Cost-Sensitive Decision Trees and Ensembles. A standard decision tree is typically constructed myopically. For example, it decides which feature to split based on the feature with largest information gain $I(\mathbf{y}; \mathbf{x}_a \mid x_S)$ [Quinlan, 1993]. An early example of a cost-sensitive decision tree is [Norton, 1989], where the split criterion now takes the cost of the feature into account as $I(\mathbf{y}; \mathbf{x}_a \mid x_S)/c_a$ (i.e., connected to the myopic CMI heuristic in (12)). This requires access to $p(\mathbf{x}_a \mid x_S)$ and $p(\mathbf{y} \mid x_S)$, which are estimated from the empirical frequencies of the training examples that reach the current node (i.e., consistent with x_S). This is a form of generative estimation of these probabilities. These methods hence also belong to the category *myopic generative heuristics* in Section 5.2. Other methods are similar in spirit, but consider other myopic cost-sensitive split criteria [Núñez, 1991; Tan, 1990; Turney, 1995; Ling *et al.*, 2004; Sheng and Ling, 2006; Ling *et al.*, 2006; Yang *et al.*, 2006]. Some of these methods also consider non-myopic variants via lookahead during splits [Norton, 1989] or post-processing of the myopic tree by optimizing some global objective (e.g., (3)) [Turney, 1995]. This aligns them with one of the *non-myopic heuristic* categories discussed later. Notably, [Ling *et al.*, 2006] rebuilds a cost-sensitive decision tree at test time after each feature is acquired, with cost of already acquired features set to zero. This corresponds to a kind of online planning (a common approach to solve complex POMDPs [Ross *et al.*, 2008]).

This idea has also been extended to ensemble methods, e.g., random forests [Xu *et al.*, 2012; Nan *et al.*, 2015] and boosting methods [Peter *et al.*, 2017]. The general idea is to build multiple cost-sensitive decision trees and combine them similar to standard ensemble methods. A feature is only acquired the first time it is used by a tree in the ensemble.

5.2 Model-Based Methods

Model-based methods estimate the distribution over unobserved features given the observed features, $p(\mathbf{x}_U \mid x_S)$, to perform model-based planning in the AFA-POMDP (see Section 3). These estimates are needed to evaluate observation probabilities, belief updates, and lookahead values. The main difference across methods is how they approximate the action-value in (10), ranging from one-step proxies (myopic) to objectives that explicitly account for interactions with other unobserved features (non-myopic), and to explicit planning via dynamic programming or search (non-myopic).

Myopic Generative Heuristics. These methods use the myopic approximation in (11) (with different choices of the terminal reward $r(S, b(x_S), \texttt{STOP})$ in (8)). Early methods include [Geman and Jedynak, 1996; Chai *et al.*, 2004], both of which assume a Naive Bayes setting. [Ji and Carin, 2007] goes beyond the Naive Bayes assumption to richer class-conditional generative models (e.g., IOHMM/HMM beliefs to compute both $p(\mathbf{y} \mid x_S)$ and $p(\mathbf{x}_a \mid x_S)$). More recently, deep models have been used to estimate these probabilities without making any simplifying assumptions. This is implemented using arbitrary-conditional deep generative models (e.g., Partial VAE) to approximate the required conditionals efficiently [Ma *et al.*, 2019; Rangrej and Clark, 2021; Chattopadhyay *et al.*, 2023b]. For these methods, the deep model learns to directly predict relevant probabilities for arbitrary subset x_S, instead of sequential Bayesian updating. Consequently, they are very hard to train and often have difficulties converging.

Non-Myopic Generative Heuristics. The methods in this category are non-myopic because they consider the unobserved features beyond a one-step lookahead. However, they do not use standard POMDP planning to achieve this. Instead, they are non-myopic by exploiting the specific structure of the AFA problem (i.e., by estimating the joint informativeness of unobserved features, see discussion in Section 4).

[Bilgic and Getoor, 2007] estimate the joint informativeness of feature groups and then select a single feature from the most informative group. To make this tractable, they assume conditional independence among features. Subsequently, the EC^2 objective was introduced in [Golovin and Krause, 2011] and used for AFA in [Rahbar *et al.*, 2023; Rahbar *et al.*, 2025]. EC^2 prioritizes features that reduce uncertainty in both $p(\mathbf{y} \mid x_S)$ (aligning with the myopic CMI rule in (12)) and $p(\mathbf{x} \mid x_S)$ (rapidly eliminating feature vectors $x \in \mathcal{X}$ that are inconsistent with the observations). EC^2 is also *adaptive submodular* [Golovin and Krause, 2011], which yields a near-optimality guarantee.

Recently, [Valancius *et al.*, 2024] proposed a method that randomly samples subsets of unobserved feature indices from U. For each subset, it evaluates the features' joint informativeness about the label, then selects the most informative feature (under the myopic CMI rule) within the most informative subset. When log loss is used, this procedure is related to the non-myopic CMI heuristic in (14). To compute the required probabilities, they rely on neighborhood-based probability estimation: they identify instances close to x_S and evaluate the objective using this local neighborhood. This approach is reminiscent of probability estimation in decision trees and avoids training a deep generative model for arbitrary S. [Huang *et al.*, 2025] extend this framework by learning to identify subsets that outperform random sampling via an offline set-optimization problem. Similarly, [Norcliffe *et al.*, 2025] use gradient-based local attributions and a deep generative model to estimate expectations over unobserved features in a latent space, which is related to (14) with gradient-based feature attributions replacing mutual information. Because these methods compute their scores at test time, they can be viewed as a form of online planning, as is common in expensive POMDP settings where full offline planning is intractable [Ross *et al.*, 2008].

Dynamic Programming and Search. [Liyanage and Zois, 2021] proposes to solve the POMDP of AFA via dynamic programming. To make it tractable, they utilize (i) the piecewise-linear structure of the value function of the AFA-POMDP, and (ii) the reduced sufficient statistic via conditional independence as described in Section 3. An alternative is to discretize the information state via learned abstractions, for example using the leaves of decision trees as surrogate states [Maliah and Shani, 2018]. Moreover, the state space of the AFA-POMDP corresponds to an AND-OR graph. [Bayer-Zubek and Dietterich, 2005] proposes to search for a policy in this graph using the AO^* search algorithm. They propose an admissible heuristic and statistical pruning to make it tractable.

5.3 Model-Free and Hybrid Methods

Model-free methods sidestep explicit probabilistic modeling of $p(\mathbf{x}_U \mid x_S)$, and instead learn acquisition behavior directly from data instances $(x, y) \sim p(\mathbf{x}, \mathbf{y})$. Rather than estimating probabilities and plugging them into an acquisition objective, they learn to predict the value of actions in a given state (e.g., via an action-value function), typically using a neural network.

Myopic Discriminative Heuristics (Oracle Guided). Oracle guided methods can be viewed as imitation learning (often simple behavioral cloning) for AFA: using fully observed training instances, one simulates states (S, x_S) along rollouts (see the MDP view of AFA in Section 3), computes an oracle action for each visited state, and trains a parametric policy to imitate this oracle. The oracle action is typically computed by utilizing the true label y and unobserved features x_U of the instance (as both are assumed to be known in the offline setting). This perspective was made explicit early on by [He *et al.*, 2012]. They can also be seen as simplified variants of the RL methods discussed later: rather than learning from delayed returns, they replace credit assignment with supervised oracle labels, which makes training easier but typically biases the learned policy toward the oracle's (often more myopic) structure.

[Covert *et al.*, 2023] learns a policy network π_θ (represented by a neural net). It simulates episodes on training instances by following π_θ. At any state (S, x_S), an action is drawn $a \sim \pi_\theta(S, x_S)$. Subsequently, the loss $\ell(f(x_{S\cup\{a\}}), y)$ is computed in the resulting state. The policy parameters are then updated based on this loss. In essence, they allow gradients to backpropagate from the loss to the policy network via the reparameterization trick. The predictor f is also trained jointly. This is a unique aspect of the POMDP of AFA: we can

optimize policies via gradient methods without having to rely on high-variance methods such as the REINFORCE algorithm [Ghosh and Lan, 2023]. This is because, via a continuous relaxation of the AFA problem, the terminal reward can be made differentiable w.r.t. the action. Moreover, [Covert *et al.*, 2023] shows that this policy is equivalent to the selections made by the myopic CMI criterion in (12) (assuming log loss). Note that the method in [He *et al.*, 2012] is very similar to this. Similar methods were concurrently proposed by [Ghosh and Lan, 2023] and [Chattopadhyay *et al.*, 2023a]. [Gadgil *et al.*, 2024] is closely related too, but it learns to predict the one-step action-value function instead of the policy directly.

Non-Myopic Discriminative Heuristics (Oracle Guided). The methods covered here are similar to those above, but they use a non-myopic oracle signal instead of a myopic one, yielding a non-myopic policy. The work in [He *et al.*, 2016] simulates episodes on training instances $(x, y) \sim p(\mathbf{x}, \mathbf{y})$. At any state (S, x_S) during an episode, it executes each candidate action and rolls it out to termination to obtain a terminal loss. This yields a non-myopic signal for each action. Explainability-driven rankings provide another oracle signal: [Guney *et al.*, 2025] compute local SHAP/LIME feature attributions on fully observed instances and train a policy to follow the highest-ranked unacquired feature. Because some explanation methods (notably SHAP) can reflect feature interactions and joint informativeness, the resulting signal may be non-myopic to a limited extent.

Model-Free RL. [Dulac-Arnold *et al.*, 2012] were the first to show the connection between (3) and maximizing long-term reward in the corresponding MDP of AFA. Motivated by this connection, they use RL to learn a policy in a model-free way by simulating trajectories offline on fully observed data. They learn the predictor and policy jointly via approximate policy iteration. [Rückstieß *et al.*, 2013] also uses model-free RL to solve the problem, but uses a slightly different state representation and uses fitted Q-iteration. More recently, it has also been extended to deep RL by, e.g., [Shim *et al.*, 2018; Kachuee *et al.*, 2019; Janisch *et al.*, 2020] using deep RL algorithms such as DQN and PPO. [Aronsson and Chehreghani, 2026] also considers a continuous relaxation of the AFA problem, similar to [Covert *et al.*, 2023; Ghosh and Lan, 2023]. However, it is non-myopic because gradients flow through the entire acquisition trajectory rather than a single acquisition step. It also focuses on the soft-budget setting, whereas these myopic relaxation-based methods are limited to hard budgets with uniform feature costs.

Model-Based RL (Hybrid). Some RL-based methods train an auxiliary surrogate model (for example, the PVAE in [Ma *et al.*, 2019]) to approximate the conditional distribution $p(\mathbf{x}_a \mid x_S)$. This surrogate can improve training stability in otherwise model-free approaches. For instance, [Zannone *et al.*, 2019] uses it to generate synthetic trajectories by sampling from $p(\mathbf{x}_a \mid x_S)$, though the empirical benefit of this augmentation appears limited [Schütz *et al.*, 2026]. By contrast, [Li and Oliva, 2021] uses the surrogate more directly to support learning, enriching the policy input (for example, with imputations and uncertainty summaries) and providing denser reward signals than those obtained from purely model-free rollouts. Although these methods learn a dynamics model (as model-based methods covered in Section 5.2), they remain primarily model-free RL, using it mainly to stabilize training.

6 Conclusion and Future Directions

This survey organizes AFA around its underlying POMDP structure, showing that most methods can be understood as different approximations to the same sequential decision problem. We introduced a unified taxonomy that separates four categories: (i) embedded cost-aware predictors, (ii) model-based methods, (iii) model-free methods, and (iv) hybrid methods. We hope this POMDP-centric view both clarifies existing work and motivates new AFA methods that more directly build on the mature literature on POMDP planning and approximation.

Despite substantial progress, several research directions remain open. First, recent myopic discriminative heuristics [Covert *et al.*, 2023; Gadgil *et al.*, 2024] often perform remarkably well on common AFA datasets, in many cases outperforming non-myopic approaches, including RL-based methods [Schütz *et al.*, 2026]. At the same time, it has been shown that when datasets are synthetically constructed so that optimal performance requires non-myopic selection, non-myopic methods expectedly outperform myopic ones [Schütz *et al.*, 2026; Aronsson and Chehreghani, 2026]. This motivates a more principled understanding of when non-myopic policies are necessary for a given data distribution $p(\mathbf{x}, \mathbf{y})$, and whether such insights can be used to decide when the additional complexity of non-myopic planning is justified.

Second, most AFA papers assume an offline setting with fully observed feature vectors, split into training and test sets for learning and evaluation. This can conflict with the premise that features are costly, and real datasets often have missing values (e.g., medical records include only ordered tests). Consequently, if a policy acquires a feature that is systematically missing in the logged data, it cannot be reliably evaluated. This motivates online and semi-offline settings and evaluation protocols that account for realistic logging and missingness. Notably, [von Kleist *et al.*, 2025] studies this issue, and further work is needed to develop realistic AFA systems.

Third, most AFA methods come without performance guarantees. A notable exception constitutes the methods based on the EC^2 objective [Rahbar *et al.*, 2023; Rahbar *et al.*, 2025], which admit near-optimality guarantees, but the objective is typically expensive to compute.

Due to space constraints, several specialized and less common AFA settings were not covered in depth here. In addition, many methods in Section 5 were covered only briefly, and a number of additional AFA works were omitted from this shorter version. The extended version of this survey will treat these works and settings in more detail and provides a more extensive discussion of future directions.

Acknowledgments

This work was partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation. We would like to thank Daphney-Stavroula Zois for pointing us to a number of relevant works.

References

[Aronsson and Chehreghani, 2026] Linus Aronsson and Morteza Haghir Chehreghani. Non-myopic active feature acquisition via pathwise policy gradients. *arXiv preprint arXiv:2605.05511*, 2026.

[Attenberg *et al.*, 2011] Josh Attenberg, Prem Melville, Foster Provost, and Maytal Saar-Tsechansky. Selective data acquisition for machine learning. *Cost-sensitive machine learning*, 2011.

[Bayer-Zubek and Dietterich, 2005] Valentina Bayer-Zubek and Thomas G. Dietterich. Integrating learning from examples into the search for diagnostic policies. *J. Artif. Int. Res.*, 24:263–303, 2005.

[Bilgic and Getoor, 2007] Mustafa Bilgic and Lise Getoor. Voila: efficient feature-value acquisition for classification. In *Proceedings of the 22nd National Conference on Artificial Intelligence*, 2007.

[Chai *et al.*, 2004] Xiaoyong Chai, Lin Deng, Qiang Yang, and Charles X Ling. Test-cost sensitive naive bayes classification. In *IEEE International Conference on Data Mining*, 2004.

[Chattopadhyay *et al.*, 2023a] Aditya Chattopadhyay, Kwan Ho Ryan Chan, Benjamin David Haeffele, Donald Geman, and Rene Vidal. Variational information pursuit for interpretable predictions. In *The Eleventh International Conference on Learning Representations*, 2023.

[Chattopadhyay *et al.*, 2023b] Aditya Chattopadhyay, Stewart Slocum, Benjamin D Haeffele, Rene Vidal, and Donald Geman. Interpretable by design: Learning predictors by composing interpretable queries. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(6):7430–7443, 2023.

[Covert *et al.*, 2023] Ian Covert, Wei Qiu, Mingyu Lu, Nayoon Kim, Nathan White, and Su-In Lee. Learning to maximize mutual information for dynamic feature selection. In *Proceedings of the 40th International Conference on Machine Learning*, 2023.

[Dulac-Arnold *et al.*, 2012] Gabriel Dulac-Arnold, Ludovic Denoyer, Philippe Preux, and Patrick Gallinari. Sequential approaches for learning datum-wise sparse representations. *Mach. Learn.*, 89:87–122, 2012.

[Gadgil *et al.*, 2024] Soham Gadgil, Ian Connick Covert, and Su-In Lee. Estimating conditional mutual information for dynamic feature selection. In *The Twelfth International Conference on Learning Representations*, 2024.

[Geman and Jedynak, 1996] Donald Geman and Bruno Jedynak. An active testing model for tracking roads in satellite images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(1):1–14, 1996.

[Ghosh and Lan, 2023] Aritra Ghosh and Andrew Lan. Difa: Differentiable feature acquisition. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2023.

[Golovin and Krause, 2011] Daniel Golovin and Andreas Krause. Adaptive submodularity: Theory and applications in active learning and stochastic optimization. *Journal of Artificial Intelligence Research*, 42:427–486, 2011.

[Guney *et al.*, 2025] Osman Berke Guney, Ketan Suhaas Saichandran, Karim Elzokm, Ziming Zhang, and Vijaya B Kolachalama. Active feature acquisition via explainability-driven ranking. In *Forty-second International Conference on Machine Learning*, 2025.

[Guyon and Elisseeff, 2003] Isabelle Guyon and André Elisseeff. An introduction to variable and feature selection. *J. Mach. Learn. Res.*, 3:1157–1182, 2003.

[He *et al.*, 2012] He He, Jason Eisner, and Hal Daume. Imitation learning by coaching. In *Advances in Neural Information Processing Systems*, 2012.

[He *et al.*, 2016] He He, Paul Mineiro, and Nikos Karampatziakis. Active information acquisition. *arXiv preprint arXiv:1602.02181*, 2016.

[Heckerman *et al.*, 1992] David E. Heckerman, Eric J. Horvitz, and Bharat N. Nathwani. Toward normative expert systems: Part i. the pathfinder project. *Methods of Information in Medicine*, 31(2):90–105, 1992.

[Howard, 1966] Ronald A. Howard. Information value theory. *IEEE Transactions on Systems Science and Cybernetics*, 2(1):22–26, 1966.

[Huang *et al.*, 2025] Hung-Tien Huang, Dzung Dinh, and Junier B. Oliva. Information templates: A new paradigm for intelligent active feature acquisition. *arXiv preprint arXiv:2508.18380*, 2025.

[Janisch *et al.*, 2020] Jaromír Janisch, Tomás Pevný, and Viliam Lisý. Classification with costly features as a sequential decision-making problem. *Mach. Learn.*, 109(8):1587–1615, 2020.

[Ji and Carin, 2007] Shihao Ji and Lawrence Carin. Cost-sensitive feature acquisition and classification. *Pattern Recogn.*, 40:1474–1485, 2007.

[Kachuee *et al.*, 2019] Mohammad Kachuee, Orpaz Goldstein, Kimmo Kärkkäinen, and Majid Sarrafzadeh. Opportunistic learning: Budgeted cost-sensitive learning from data streams. In *International Conference on Learning Representations*, 2019.

[Lauri *et al.*, 2023] Mikko Lauri, David Hsu, and Joni Pajarinen. Partially observable markov decision processes in robotics: A survey. *IEEE Transactions on Robotics*, 39(1):21–40, 2023.

[Li and Oliva, 2021] Yang Li and Junier Oliva. Active feature acquisition with generative surrogate models. In *Proceedings of the 38th International Conference on Machine Learning*, 2021.

[Ling *et al.*, 2004] Charles X Ling, Qiang Yang, Jianning Wang, and Shichao Zhang. Decision trees with minimal costs. In *Proceedings of the twenty-first international conference on Machine learning*, 2004.

[Ling *et al.*, 2006] C.X. Ling, V.S. Sheng, and Q. Yang. Test strategies for cost-sensitive decision trees. *IEEE Transactions on Knowledge and Data Engineering*, 18(8):1055–1067, 2006.

[Liyanage and Zois, 2021] Yasitha Warahena Liyanage and Daphney–Stavroula Zois. Optimum feature ordering for dynamic instance–wise joint feature selection and classification. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2021.

[Ma *et al.*, 2019] Chao Ma, Sebastian Tschiatschek, Konstantina Palla, Jose Miguel Hernandez-Lobato, Sebastian Nowozin, and Cheng Zhang. EDDI: Efficient dynamic discovery of high-value information with partial VAE. In *Proceedings of the 36th International Conference on Machine Learning*, 2019.

[Maliah and Shani, 2018] Shlomi Maliah and Guy Shani. Mdp-based cost sensitive classification using decision trees. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1), 2018.

[Nan *et al.*, 2015] Feng Nan, Joseph Wang, and Venkatesh Saligrama. Feature-budgeted random forest. In *Proceedings of the 32nd International Conference on Machine Learning*, 2015.

[Norcliffe *et al.*, 2025] Alexander Luke Ian Norcliffe, Changhee Lee, Fergus Imrie, Mihaela van der Schaar, and Pietro Lio. Stochastic encodings for active feature acquisition. In *Forty-second International Conference on Machine Learning*, 2025.

[Norton, 1989] Steven W. Norton. Generating better decision trees. In *Proceedings of the 11th International Joint Conference on Artificial Intelligence*, 1989.

[Núñez, 1991] Marlon Núñez. The use of background knowledge in decision tree induction. *Machine Learning*, 6(3):231–250, 1991.

[Peter *et al.*, 2017] Sven Peter, Ferran Diego, Fred A Hamprecht, and Boaz Nadler. Cost efficient gradient boosting. In *Advances in Neural Information Processing Systems*, 2017.

[Quinlan, 1993] J Ross Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers, 1993.

[Rahbar *et al.*, 2023] Arman Rahbar, Ziyu Ye, Yuxin Chen, and Morteza Haghir Chehreghani. Efficient online decision tree learning with active feature acquisition. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, 2023.

[Rahbar *et al.*, 2025] Arman Rahbar, Niklas Åkerblom, and Morteza Haghir Chehreghani. Cost-efficient online decision making: A combinatorial multi-armed bandit approach. *Transactions on Machine Learning Research*, 2025.

[Rangrej and Clark, 2021] Samrudhdhi B. Rangrej and James J. Clark. A probabilistic hard attention model for sequentially observed scenes. In *32nd British Machine Vision Conference*, 2021.

[Ross *et al.*, 2008] Stéphane Ross, Joelle Pineau, Sébastien Paquet, and Brahim Chaib-draa. Online planning algorithms for pomdps. *J. Artif. Int. Res.*, 32:663–704, 2008.

[Rückstieß *et al.*, 2013] T. Rückstieß, C. Osendorfer, and P. van der Smagt. Minimizing data consumption with sequential online feature selection. *International Journal of Machine Learning and Cybernetics*, 4:235–243, 2013.

[Schütz *et al.*, 2026] Valter Schütz, Han Wu, Reza Rezvan, Linus Aronsson, and Morteza Haghir Chehreghani. Afabench: A generic framework for benchmarking active feature acquisition. In *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2026.

[Settles, 2009] Burr Settles. Active learning literature survey. Technical report, University of Wisconsin–Madison, 2009. Computer Sciences Technical Report.

[Sheng and Ling, 2006] Victor S. Sheng and Charles X. Ling. Feature value acquisition in testing: a sequential batch test algorithm. In *Proceedings of the 23rd International Conference on Machine Learning*, 2006.

[Shim *et al.*, 2018] Hajin Shim, Sung Ju Hwang, and Eunho Yang. Joint active feature acquisition and classification with variable-size set encoding. In *Advances in Neural Information Processing Systems*, 2018.

[Tan, 1990] Ming Tan. Csl: A cost-sensitive learning system for sensing and grasping objects. In *Proceedings of the IEEE International Conference on Robotics and Automation*, 1990.

[Turney, 1995] Peter D. Turney. Cost-sensitive classification: empirical evaluation of a hybrid genetic decision tree induction algorithm. *J. Artif. Int. Res.*, 2:369–409, 1995.

[Valancius *et al.*, 2024] Michael Valancius, Max Lennon, and Junier B. Oliva. Acquisition conditioned oracle for nongreedy active feature acquisition. In *Proceedings of the 41st International Conference on Machine Learning*, 2024.

[von Kleist *et al.*, 2025] Henrik von Kleist, Alireza Zamanian, Ilya Shpitser, and Narges Ahmidi. Evaluation of active feature acquisition methods for time-varying feature settings. *Journal of Machine Learning Research*, 26(60):1–84, 2025.

[Xu *et al.*, 2012] Zhixiang Xu, Kilian Q. Weinberger, and Olivier Chapelle. The greedy miser: learning under test-time budgets. In *Proceedings of the 29th International Conference on Machine Learning*, 2012.

[Yang *et al.*, 2006] Qiang Yang, C. Ling, X. Chai, and Rong Pan. Test-cost sensitive classification on data with missing values. *IEEE Transactions on Knowledge and Data Engineering*, 18(5):626–638, 2006.

[Zannone *et al.*, 2019] Sara Zannone, Jose Miguel Hernandez Lobato, Cheng Zhang, and Konstantina Palla. Odin: Optimal discovery of high-value information using model-based deep reinforcement learning. In *Real-world Sequential Decision Making Workshop, ICML*, 2019.

[Zois and Chelmis, 2024] Daphney-Stavroula Zois and Charalampos Chelmis. From feature selection to instance-wise feature acquisition. Minitutorial at SIAM International Conference on Data Mining, 2024.

[Åström, 1965] K.J Åström. Optimal control of markov processes with incomplete state information. *Journal of Mathematical Analysis and Applications*, 10(1):174–205, 1965.