Active Feature Acquisition via Human Interaction in Relational domains

Nandini Ramanan
nandhiniramanan5@gmail.com
University of Texas at Dallas
USA

Phillip Odom
Georgia Institute of Technology
USA
Phillip.Odom@gtri.gatech.edu

Kristian Kersting
TU Darmstadt
Germany
kersting@cs.tu-darmstadt.de

Sriraam Natarajan
UT Dallas
USA
Sriraam.Natarajan@utdallas.edu

ABSTRACT

We consider the problem of *interactive and explainable active feature elicitation* in relational domains in which a small subset of data is fully observed while the rest of the data is minimally observed. The goal is to identify the most informative set of entities for whom acquiring additional relations would yield a more robust model. We assume the presence of a human expert who can interactively provide the relations. Thus there is a need for an explainable model. Consequently, we employ an relational tree-based distance metric to identify the most diverse set of relational examples (entities) to obtain more relational feature information on. The model that is learned iteratively is an interpretable and explainable model that is presented to the human expert for eliciting additional features. Our empirical evaluation demonstrates both the efficiency and the interpretability of the proposed approach.

KEYWORDS

Relational Learning, Interactive ML, Active learning

ACM Reference Format:
Nandini Ramanan, Phillip Odom, Kristian Kersting, and Sriraam Natarajan. 2023. Active Feature Acquisition via Human Interaction in Relational domains. In *6th Joint International Conference on Data Science & Management of Data (10th ACM IKDD CODS and 28th COMAD) (CODS-COMAD 2023), January 4–7, 2023, Mumbai, India.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3570991.3571001

1 INTRODUCTION

Feature-value acquisition is a necessary step for deployment of AI and ML systems. It has long been pursued in standard domains (i.e, the ones that can be easily described by a flat vector representation) from the perspective of active learning [11, 36]. Such approaches, while successful, cannot be directly extended to relational domains such as networks or hyper-graphs [3]. For instance, simply acquiring a single feature value, say, a paper for a specific author, say *Mitchell* would provide minimal value in learning. On the other hand, acquiring all the values of the features, say all the papers that *Mitchell has co-authored* would involve cumbersome data gathering. Instead, soliciting a specific query such as *list of papers that Mitchell co-authored with Cohen* could provide significant information value when learning a robust classifier. Note that this requires reasoning at different levels of abstraction – individual feature values, subgroups of entities/relations and at the level of the set of all objects. We consider this type of learning under the paradigm of Statistical Relational Learning (SRL) [9, 12, 27] that naturally combines the power of relational representations such as first-order logic with the ability of statistical/probabilistic models to handle uncertainty. Specifically, we consider actively acquiring features in relational data.

While previous research in relational data for active learning have mainly considered acquiring the labels in the context of link prediction or relation extraction [3, 15, 18], we consider a different yet related task of identifying the best entities (examples) to acquire more feature information on. This is quite natural in settings such as networks where a subset of the network is fully observed while in the other part of the network, only the entities are known with partial relations. For instance, in a clinical study where participants live in a smart home, one could observe all the social interactions. However, their outside interactions are unobserved and it is essential to obtain some of these information when building a relational classifier (predicting friendships). Our goal is to identify the *most informative set of* individuals to build a robust model.

This problems poses several challenges in relational settings that we directly address – First, a need for *active selection strategy* of the elicitable (relational) examples on whom acquiring additional features will improve the classifier performance. Our hypothesis in this work is that acquiring diverse set of examples will help improve generalization. This brings out the second challenge – that of computing distances between relational examples. While in Natarajan et al. [22], it was easily computed using a divergence metric such as KL-divergence [14], in relational models, the distance calculation is not straightforward. Thus we employ the notion of first-order tree-based distance due to Khot et al. [13] for computing the most diverse set of examples. Third challenge is that of *feature-subspace selection*

CODS-COMAD 2023, January 4–7, 2023, Mumbai, India

ACM ISBN 978-1-4503-9797-1/23/01…$15.00
https://doi.org/10.1145/3570991.3571001

for each example, to avoid uninformative/redundant features. Feature sub-space selection in relational domains is intractable due to the number of groundings. To mitigate this, inspired by the ideas of subjective understanding in explanation-based learning literature [33, 35], we perform inference on the current model based on the chosen examples from the tree-based distances. This inference step **explains the most important aspects of the current model to the human learner**. Our base learner is a *relational gradient-boosted tree* that is combined to a single tree at each explanation step. The algorithm picks the most uncertain part of this tree for the current example and elicits the additional features from the human. The final challenge in relational domains is identifying the appropriate level of abstraction – should the query be over an instance (an individual co-author), a sub-group (area of research) or the population (over all the journals)? We note that this challenge is automatically handled by our **explainable, interpretable and elicitable** base learner which is a (combined) relational decision-tree. Identifying the appropriate paths will automatically present the set of features to the human expert. The presence of human expert makes our work distinctly different from the previous active learning work in relational data. It reinforces the need for **explainable** models and our approach precisely addresses this issue.

To summarize we make the following contributions: (1) We consider and address the problem of training-time feature acquisition in the presence of a human-expert. (2) We develop the first **F**eature **A**cquisition via **I**nteraction in **R**elational domains (*FAIR*) algorithm that actively solicits features on the most informative examples. (3) We adapt a relational tree-based distance metric for identifying the most diverse set of examples. (4) We employ an **explainable and interpretable** subjective feature subset elicitation strategy for relational domains to present the most informative set of features to the human expert. (5) Our empirical evaluation across two standard relational domains demonstrate the efficiency, effectiveness of the *FAIR* and most importantly, the explainability and interpretability of the feature selection strategy.

As far as we are aware, this is the first work on SRL models that can seamlessly incorporate human-in-the-loop while performing the computationally intensive task of full model-learning with partially observed examples.

The rest of the paper is organized as follows: we introduce the necessary background on active learning, relational learning and explanations before outlining the framework and then discuss the algorithm. We then present the empirical evaluations before concluding by outlining areas of future research.

2 BACKGROUND

2.1 Active Learning

Active Learning [30] is a paradigm which deals with sample-efficient learning where the assumption is that labels are expensive to obtain. Thus, the goal is to select the most informative instances for whom labels need to be acquired in order to provide maximum information to the underlying training model. Typically, techniques such as uncertainty sampling [17], query by committee [32] etc. are used to select informative samples. Active learning has been successfully used with traditional machine learning models ranging from logistic regression [17], support vector machines [39] to Bayesian networks [37]. It has also been used for more expressive modes of communication like seeking advice actively for discriminative models such as probabilistic logic models [25] to decision making tasks like inverse reinforcement learning [24]. There is a broad set of active learning algorithms that have been applied to resource-constrained domains [31, 38], however in propositional setting.

2.2 Active Learning and Active Inference in Relational Domains

Recent work in active learning in the context of relational or structured domains have just begun to explore the idea of actively acquiring labels for both learning and collective inference. The fundamental difference between active learning and active inference is when these labels are collected. Related work by Bilgic et al. and Macskassy [2, 18] focused on acquiring labels for examples that will improve the accuracy of collective inference by considering properties of the network structure. These "reflect and correct" approaches only query for labels while inference. Bilgic et al. proposed ALFNET [3], an active learning method for relational data that builds on uncertainty sampling, committee-based sampling, and clustering. ALFNET employs the disagreement score between three classifiers a). content-only classifier (i.e., trained using attributes of the example in isolation), b). a collective classifier (i.e., trained using attributes and network structure) and c) clustering, and chooses the instance with highest score for labeling.

Kuwadekar et al. proposed a semi-supervised method called RAL that learns a relational dependency network and relies on an ensemble of models to select the most informative instances to query [15]. Their intuition behind this approach was to select these instances to acquire the label, whose predictions are potentially most certain, which is opposite of what had been seen previously in the literature.

Shi et al. proposed a batch mode active learning (BMAL) using graph-based metrics to define the informativeness of instances [34]. Ji et al. select examples that minimize the total variance of the distribution of the unlabeled samples along with the the total generalization error [10]. It is worth noticing that all the above techniques combine the relations (links) information with examples-specific features to train a classifier and then employ various query strategies for instance selection [3, 15, 34].

2.3 Active Feature Elicitation

Inspired by the success of active learning methods for structured domains, we consider the novel problem of *active feature elicitation* for relational domains. Active feature elicitation is similar to active learning in the sense that both the settings have budgetary constraint and learns from most useful examples. However, they are different in their assumptions and problem setting because active learning assumes that labels are expensive whereas AFE makes the assumption that feature subsets are expensive, the labels being fully observed for all the examples. Our goal is to select the best set of examples on whom the missing features (links/attributes) can be queried on to best improve the classifier performance. The problem of *active feature elicitation* for propositional domains has previously been studied in literature and has been solved using various imputation techniques [19, 20, 40] to fill in the missing feature set and then using traditional active learning.

Kanani et al. [11] proposed a test time elicitation framework that employed uncertainty sampling on the observed feature sets in a propositional setting. We propose a training time elicitation framework that employs a tree based relational distance metric to compute the distance between the two sets of points to acquire the most useful relational example and its attributes. Later, Thahir et al. [36] built on Kanani et al. [11] by adding an extra term to capture the utility of adding an example to the training set and was extended to a more generalized setting by Saar-Tsechansky et al. [29] where even class labels can be missing and acquired in addition to the feature subsets. A similar task with various costs for information acquisition and misclassification was addressed by Bilgic and Getoor [1] by using Probabilistic Graphical models to model the feature dependencies.

Our work is heavily inspired by the work of Natarajan et al. [22] on active feature elicitation for propositional domains, whereas we formulate the first relational active feature elicitation problem. Also in the Natarajan et al. framework, the elicitable feature set is fixed apriori and are acquired fully for an example, once deemed important by the diversity metric. They assumed that the human input is restricted to only obtaining the full set of features. Consequently, human/domain expertise was not fully exploited. However, we propose an interactive explanations framework which can facilitate a more informed active subspace elicitation for the queryable relational examples.

2.4 Explanation Based Learning

We propose a robust explanation framework which allows for explaining why we acquire a feature subset from the elicitable feature set for the group of queryable examples. As a research question this is not new and explanation based learning (EBL) systems are developed and actively pursued in literature [33, 35]. In this type of learning, EBL computes a generalisation of the training example into a form that can be applied to solve conceptually similar problems. The generalization is driven by the explanation of why the solution worked. Motivated by this, we are the first to explore EBL frameworks in the context of actively learning with relational domains. Standard EBL methods often produce overly specific rules that are shown to impact the generalization performance [21]. Cohen and Leckie et al. show that by learning simple and approximate control rules, one can improve the utility of acquired knowledge [6, 16]. To this effect, we propose to generate explanations for active feature subset elicitation for queryable examples using an approximate relational model.

3 FEATURE ACQUISITION VIA INTERACTION IN RELATIONAL DOMAINS

Traditional ML methods typically assume that all features are observed for each example. However, not all features require the same amount of effort to obtain. For example, medical applications depend on multiple modalities of data such as imagery, text, family history, and demographic or epigenetic information. While family history may be collected from patients, diagnostic procedures, which produce different data modalities, vary in terms of their cost and invasiveness. Therefore, selecting the correct procedure, and consequently the most informative feature subspace can be critical in improving decision-making while minimizing the cost of acquiring features.

3.1 Problem Formulation

Identifying the best subset of features to elicit from the experts is intractable, especially in relational domains. Instead, we present an approximate method that first selects a set of examples about whom to query and then identifies relevant features for those examples. Formally, we assume that a given dataset, $\mathbf{D} = \mathbf{D}^b \cup \mathbf{D}^o$, is composed of examples with only baseline features ($\mathbf{D}^b$) and examples which have additional observed features that have been previously queried ($\mathbf{D}^o$). For each previously elicited example, $\langle(\mathbf{x}_i^b, \mathbf{x}_i^o, y_i)\rangle \in \mathbf{D}^o$, $\mathbf{x}_i^b$ represents the base features while $\mathbf{x}_i^{o_i} \subseteq \mathbf{x}_i^e$ represents any features from the elicitable set that have been previously acquired for example i, whose label is y_i. We denote the set of observed features for all examples as $\mathbf{x^o}$. Note that each example may have a different set of observed features. Interactive feature elicitation iteratively expands $\mathbf{x^o}$ in order to improve the model. Thus, the learning problem in our setting is as follows:

Given: A relational data base $\mathbf{D}$, Relational database schema $\mathcal{R}$, Query budget B and The expert, E.

To Do: Identify the most useful set of examples $\mathbf{y}_b \in \mathbf{D}^b$ for which to obtain more attributes or relations $\mathbf{x}_i^{o_i} \subseteq \mathbf{x}_i^e$ in order to improve classifier performance.

Our proposed approach (*FAIR*), shown in Figure 1, interatively identifies a representative subset of query-able examples (1), learns an explanation model to identify subset of features from elicitable set (2) and elicits the features from experts (3). Intuitively, *FAIR* aims to acquire additional features for a diverse set of examples which could not be classified using the available features. By prioritizing examples/features, *FAIR* efficiently utilizes the resources for feature elicitation. While previous work focused on propositional distance measures [22], we faithfully capture the underlying relational data via a tree-based distance measure to capture the semantic similarity which uses the path similarity in relational decision trees [13].

The two key components of *FAIR* are: 1) **Example Subset Selection**, which relies on a relational distance measure over the full data, $\mathbf{D}$, to pick the most diverse and informative examples and 2) **Feature Sub-Space Selection**, which leverages a relational base model on the observed data $\mathbf{D}^o$, to generate robust explanation which can be used to do informed feature subspace elicitation for the queryable examples. The advantage of using a relational representation is that it succinctly captures probabilistic (noisy) dependencies among the attributes of different objects, leading to a compact representation of learned models. We chose SRL for this task for two key reasons: their ability to handle *generalized, noisy* knowledge and data and their capability to produce *explainable and interpretable* hypotheses. We will now describe the components of *FAIR* in more detail.

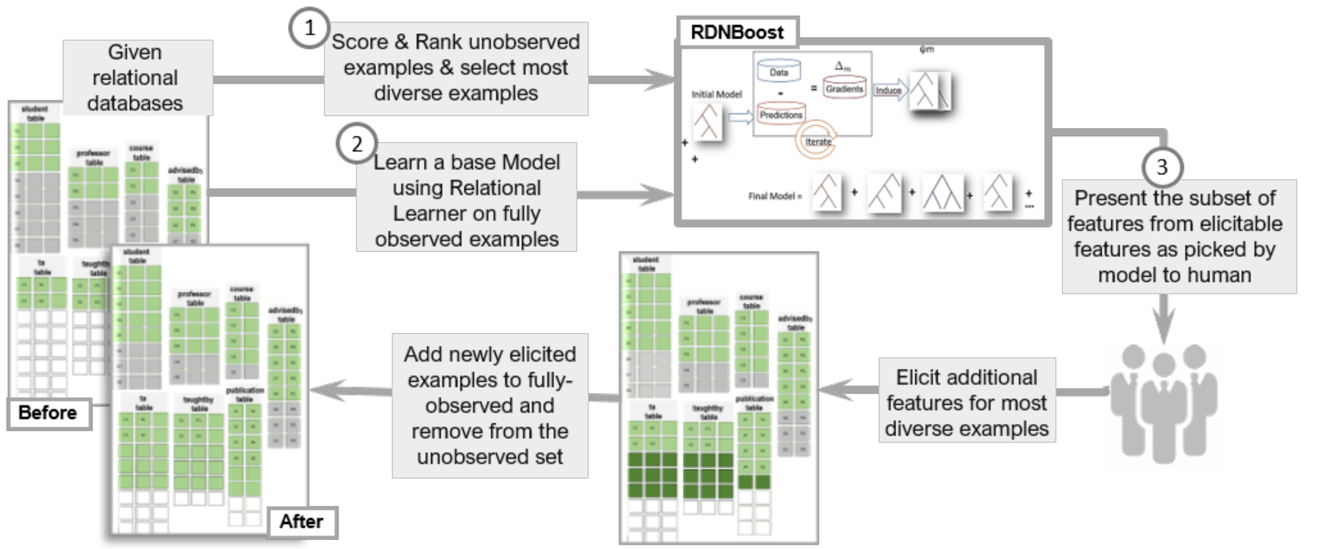

Figure 1: We introduce *F̲eature A̲cquisition via I̲nteraction in R̲elational domains*, where the feature space is characterized by base features ($\mathbf{x}^b$) and elicitable features ($\mathbf{x}^e$), which are typically more expensive to obtain. *FAIR* iteratively seeks to identify examples in $\mathbf{D}^b$ that are difficult to classify using the available features alone, and queries the human-in-the-loop to the most relevant features from the elicitable set for each example to acquire. By identifying the most important examples in $\mathbf{D}^b$ for whom to elicit features, *FAIR* systematically grows the set of fully-observed examples $\mathbf{D}^o$.

Algorithm 1 F̲eature A̲cquisition via I̲nteraction in R̲elational domains

function FAIR($\mathbf{D}^o$, $\mathbf{D}^b$, $\mathcal{R}$, B)
 where $\langle(\mathbf{x}_i^b, y_i)\rangle \in \mathbf{D}^b$, $\langle(\mathbf{x}_i^b, \mathbf{x}_i^o, y_i)\rangle \in \mathbf{D}^o$, R := Relational schema, B := Query budget
 count := 0 ▷ Initilization
 repeat
  $d(u,v)_{\{u\in\mathbf{D^o}\},\{v\in\mathbf{D^b}\}}$:= RelDistanceMetric($\mathbf{D}^b$, $\mathbf{D}^o$)
  $\zeta(u_j, v) := \sum_i d_i(u_j, v)$
  Compute Score := $\sum_j \zeta(u_j, v)$
  (Example Subset Selection.)
  $\mathbf{x}_b, \mathbf{y}_b$:= GetTopN(Score, R) , s.t. $\mathbf{y}_b \subseteq \mathbf{D}^b$
  $\mathcal{F}_e$:= RelationalFGB($\mathbf{D}^o$)
  $\mathbf{x}^o$ = FeatureSubspaceSelection($\mathcal{F}_e, \mathbf{x}_b, \mathbf{y}_b$)
  $\mathbf{D}^o = \mathbf{D}^o \cup \mathbf{x}^b \cup \mathbf{x}^o$, $\mathbf{D}^b = \mathbf{D}^b - (\mathbf{x}^b \cup \mathbf{x}^o)$
  (Feature Sub-Space Selection.)
  count = *count* + $|\mathbf{x}^o|$
 until *count* >= B ▷ i.e., budget exhausted
 return LearnFinalModel($\mathbf{D}^o$)
end function

3.2 Example Subset Selection via Relational Distance:

FAIR aims to select a representative set of examples about which to elicit features. Traditional methods leverage distance metrics (e.g., Euclidean) to quantify diversity. However, these metrics do not capture the relational structure inherent in the domain. Therefore, we leverage a tree-based relational distance measure [13], which compares path similarities across the examples, to select a representative set of examples.

We learn a series of relational regression trees on the entire data $\mathbf{D}$, which includes all the baseline and observed examples ($\mathrm{P}(y \mid \mathbf{x}^b, \mathbf{x}^o)$). Note that observed features are a subset of all possible elicitable features ($\mathbf{x}_i^o \subseteq \mathbf{x}_i^e$). We apply the standard closed world assumption for all elicitable features which have not been queried. Given the learned trees, we compute the distances between examples in $\mathbf{D}^o$ and $\mathbf{D}^b$. We use the distance measure based on paths taken by the relational examples, such that the distance between a pair of relational examples u and v is given as,

$$d(u,v) = \begin{cases} 0, & \mathrm{LCA}(u,\ v) \text{ is leaf;} \\ \exp^{-\lambda\cdot\mathsf{depth}(\mathsf{LCA}(u,v))} & \text{otherwise} \end{cases}$$

where LCA refers to the *least common ancestor* of the examples u and v and parameter $\lambda = 0.5$ ensures the distance decreases as depth increases. The relational distance between two examples is inversely proportional to the lowest common ancestor of the two examples in the learned tree. Intuitively, examples which share more common nodes in the tree are more similar. Figure 2 shows a sample first-order decision tree with examples $u \equiv$ `advisedBy("S1", "P1")` from the data pool $\mathbf{D}^o$ and $v \equiv$ `advisedBy("S4", "P1")` from data pool $\mathbf{D}^b$. They both have a distance of 1, as both examples differ at the root node with no common relational attribute.

To calculate the mean distance of an example v in $\mathbf{D^b}$ from every labeled examples in $\mathbf{D^o}$, we define a model to combine distances from multiple trees and then distances from multiple examples.

The distances from multiple incrementally learned trees are aggregated, such that $\zeta(\mathsf{u},\mathsf{v})_{\{u\in\mathbf{D^o}\},\{v\in\mathbf{D^b}\}} = \mathsf{f}(\mathsf{d}_1,\mathsf{d}_2,\mathsf{d}_3..\mathsf{d}_t)$ where $\mathsf{d_i}(\mathsf{u},\mathsf{v})$ is distance between u and v based on $i-th$ tree. $\zeta(u,v)$ combines the tree distances to return the overall distance between u and v. Then, we apply instance level combination by averaging the distances of each example which has no observed features with all the examples which have any observed features. We finally compute the overall metric (Algorithm 1 line[7]) for v using the distances from all the examples $\{u_1,u_2,..,u_n\}$ as, $\mathsf{Score}(v) = \mathsf{g}(\zeta(\mathsf{u}_1,\mathsf{v}),\zeta(\mathsf{u}_2,\mathsf{v}),..,\zeta(\mathsf{u}_n,\mathsf{v}))$ for $u_i \in \mathbf{D^o}$. The combination function f and g that we use in both levels is mean:

$$\zeta(u_j,v) = \frac{1}{t}\sum_{i=1}^{t} d_i(u_j,v)$$

$$\mathsf{Score}(v) = \frac{1}{n}\sum_{j=1}^{n} \zeta(u_j,v)$$

Using the Score metric, we identify a set of examples to acquire features by selecting N examples that are furthest from examples with observed features.

3.3 Feature Sub-Space Selection using Explanations:

Given the selected examples, we now address how to identify the subset of features to elicit. A key challenge for feature sub-space selection in the relational setting is there are an intractable number of potential groundings to consider. For example, in medical tasks the same lab test could be repeated numerous times, which may not yield useful information. Such tests can be expensive so it is important that we select informative features within the budgetary constraints. The key idea of our approach is to identify the features which would have the largest impact on the selected examples. We also present an explanation of the current model for the queried examples so that a human expert can understand and potentially guide the feature sub-space selection.

Once we have selected the diverse set of queryable examples using a relational distance measure, we need to select the features from elicitable set to be acquired per example ensuring budgetary constraints. In order to learn an interpretable representation for elicitation, *FAIR* leverages **RDN-B** [23], a relational dependency learner based on relational functional gradient boosting [8], that has previously leveraged to solicit knowledge from human experts [26]. It represents each conditional distribution as a set of relational regression trees (RRT) [4] with first-order logic in the nodes of the tree and regression values on the leaves. *FAIR* learns an explanation model $\mathcal{F}_e = \mathsf{P}(y \mid \mathbf{x}^b,\ \mathbf{x}^o)$, for examples in $\mathbf{D}^o$ (Algorithm 1 line[8]). Using the empirical approach suggested by Craven and Shavlik [7], we convert $\mathcal{F}_e$, consisting of a series of trees, into a single relational regression tree $\hat{\mathcal{F}_e}$. The resulting tree, although approximate, acts as good surrogate for generating accurate explanations of the underlying model.

Explanations are generated for each examples obtained during example subset selection. By capturing the paths that these examples take in $\mathcal{F}_e$, we highlight the elicitable features $\mathbf{x}_i^e$ that are potentially reachable for each example. For $i \in \mathbf{b}$, if i reaches a node v in the tree branch it takes, such that $v \in \mathbf{x}^e$ then that elicitable feature and every other elicitable feature in the lower subtree are reachable as shown in Figure. 3. We group the examples based on the paths and acquire the relevant features for each cluster of examples (Algorithm 1 lines[9-11]). For examples $i = \{1,2,3\} \subset \mathbf{b}$ that are closer together and acquire {`Taught_By, TA`} $\subset \mathbf{x}^e$. Similarly for examples $i = \{4,5\} \subset \mathbf{b}$, we elicit {`Publication`}. At this point, we present the elicitable feature set for the query-able examples to the human-expert as shown in the Figure 4. However, it must be noted that incorporating human advice to refine $\mathbf{x}_i^o$ for $i \in \mathbf{b}$ is natural in our framework.

Finally, the selected examples and their elicited features are added to $\mathbf{D}_o$. We update the budget B at this point based on the number of ground instances of features acquired. As in traditional active learning, *FAIR* continues until the query budget is exhausted.

4 EMPIRICAL EVALUATIONS

For our relational distance learner, we used the *Relocc* algorithm [13]. We learnt 10 first order trees with $max\_depth = 3$ in all our domains. For the fully observed relational model, we used *RDN-B* algorithm [23]. We learnt 20 relational regression trees with $max\_depth = 4$ in all our domains. At each active learning iteration, we solicit 10-20 new relational examples until the budget is exhausted, depending on the domain. In all our domains the observed example set is randomly picked. We compare four different evaluation metrics: *recall*, *F1-score*, *AUC-PR* and *AUC-ROC*, that provide a reasonably robust evaluation in the presence of high class imbalance, that is prevalent in relational domains.

4.1 Datasets

(1) **University of Washington Department of Computer Science and Engineering** data set :
The **UW-CSE** data set [28] was created from the University of Washington's Computer Science & Engineering department's student database and consists of details about professors, students and courses from 5 different subareas of computer science (AI, programming languages, theory, system and graphics). The task is to predict the `AdvisedBy` relation between a student and a professor. We have fixed train and test set. This data has $|\mathbf{D}_o| = 47$ and $|\mathbf{D}_b| = 187$. Results are averaged over 5-fold.

(2) **Never-Ending Language Learner: Sport** data set:
We consider the relational **NELL:Sport** data set [5], consisting of relations generated by the Never Ending Language Learner (NELL). The NELL consisting of information

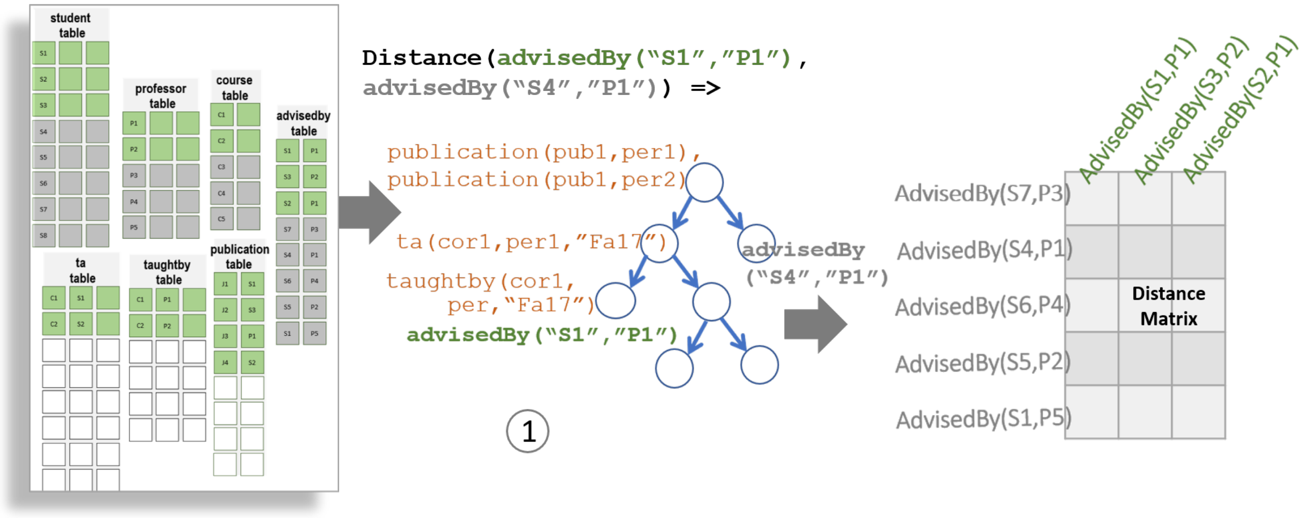

Figure 2: Distance Module in *FAIR* Framework.

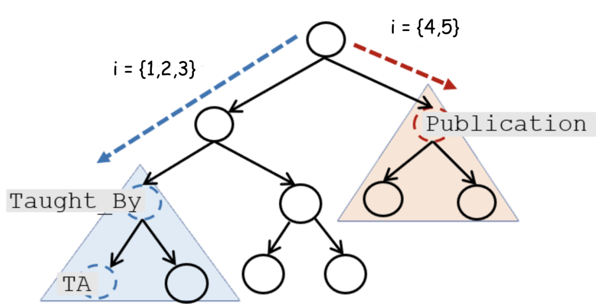

Figure 3: Explanations in *FAIR* Framework.

about players and teams. The task is to predict the relation TeamPlaysSport i.e., whether a team plays a particular sport. The data has fixed train and test set. This data has $|\mathbf{D}_o| = 119$ and $|\mathbf{D}_b| = 480$. We average the results over 3-folds and report them.

4.2 Baselines

In addition to the proposed *FAIR* approach, we considered three other baselines:

(1) Randomly choosing points to query and acquiring the full set of elicitable features such that $\mathbf{x}_o = \mathbf{x}_e$. This method is denoted as RND-ALL;
(2) Randomly choosing points to query and acquiring the features for the queryable examples using our proposed active feature subset elicitation. This method is denoted as RND;
(3) Choosing points using our proposed way of relational distance metric and acquiring the full set of elicitable features such that $\mathbf{x}_o = \mathbf{x}_e$. This is denoted as FAIR-ALL.

Comparing with these baselines, helps us do an ablation study which validates the following two important contribution in this

Data Sets	Types	Predicates	neg:pos Ratio
UW-CSE	AdvisedBy	12	539.629
NELL Sports	TeamPlaysSport	5	2.702

Table 1: Details of relational domains used in our experiments. These data sets have high ratios of negative to positive examples, which is a key characteristic of relational data sets.

algorithm: 1) the hypothesis to acquire diverse set of examples in order to improve generalization. We answer this question by accessing the gain of FAIR-ALL over RND-ALL 2) Need explainable feature-subset elicitation for the set of examples than using all the features for the queryable examples. This is answered by accessing the gain of FAIR over FAIR-ALL.

4.3 Results

As can be seen from the two domains in Figures 5 and 6, *FAIR* outperforms RND and RND-ALL on all domains across all metrics, specifically in recall where the effect of choosing the most informative set of examples can have the maximal impact on the classifier performance. We observe that the variance in recall due to random selection of examples is high as expected in both the domains for RND and RND-ALL. However, this effect is minimal for *FAIR* that chooses good training examples compared to the baselines. These relational domains under consideration, are highly imbalanced as shown in Table 1. However, it can be seen clearly that the proposed *FAIR* achieves a recall of over 0.95 in UWCS and 0.8 in NELL Sports after just a few early iterations, without significantly sacrificing f1 score making it an ideal choice for imbalanced data sets.

Similar results can be observed when comparing *FAIR* to *FAIR*-All in that *FAIR* consistently does at par or sometimes better than the strongest baseline *FAIR*-All for a lower budget, where we acquire all of the features for queryable examples. Between the two

```
>> Full elicitable feature set = {TA, Taught_By, Publication}
>> For examples i in {1, 2, 3}: acquire {TA,Taught_By}
>> For examples i in {4, 5}: acquire {Publication}
>> If ok:
>>       Press "Y" to continue.
>> Else:
>>       Select examples Id in {1-5} to change it's elicitable features.
```

Figure 4: Commands to interact with explanations in FAIR Framework.

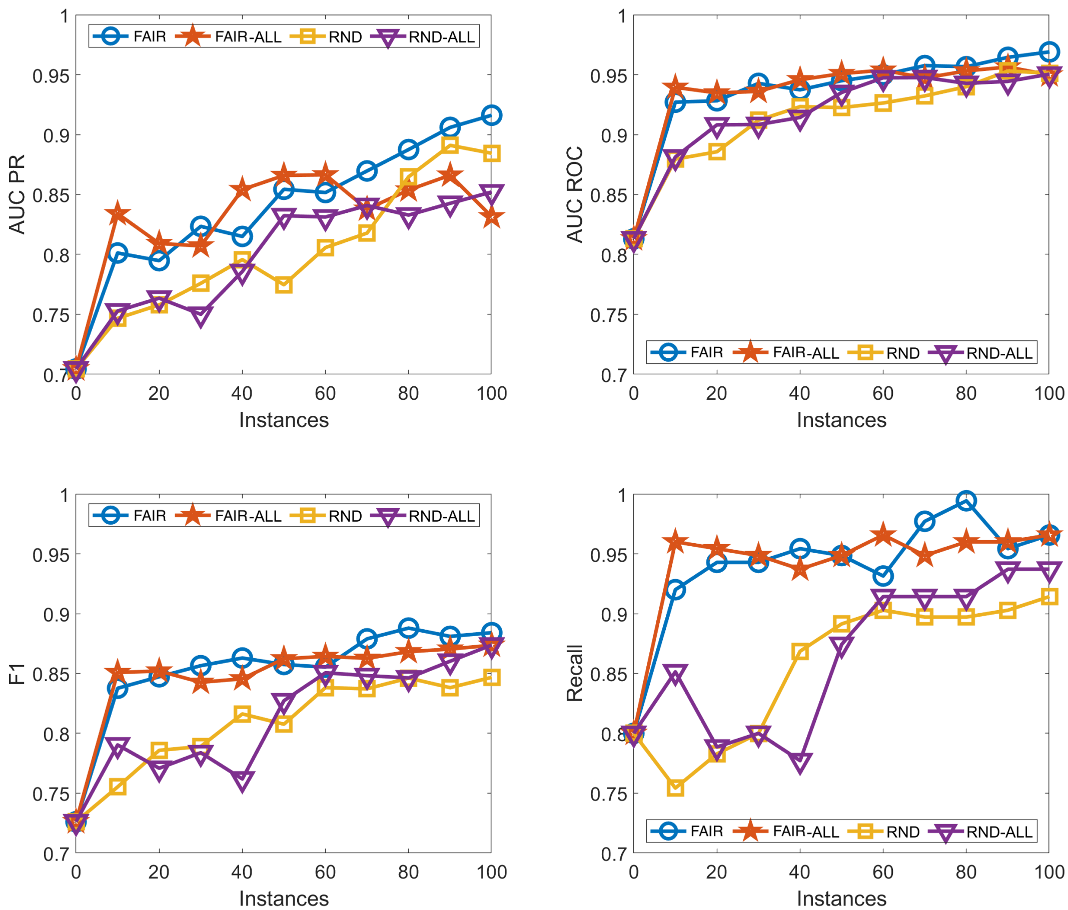

Figure 5: UWCS Results showing AUCPR, AUCROC, F1 and Recall.

domains the gain in performance across all the metrics due explainable feature-subset elicitation is significantly more visible in UWCS which has more features to query from as compared to NELL Sports. In general, the use of explainable feature-subset elicitation for the most informative set of examples still appears better and cost efficient than using all the features for the queryable examples. This helps us validate our two key components of example selection and explainable feature-subset elicitation in our empirical evaluations.

5 CONCLUSION

We present a novel approach for actively soliciting features in relational domain through interaction with human expert. At a high level, our algorithm *FAIR* identifies instances that are the farthest from the current set of fully observed relational instances. Our distance metric uses tree-based distances that are interpretable as well. Consequently, we employ the successful *explanation based learning* methods to select only the relevant feature-subspace for the informative examples. These are then presented to the expert with suitable explanations and their feature-values are obtained

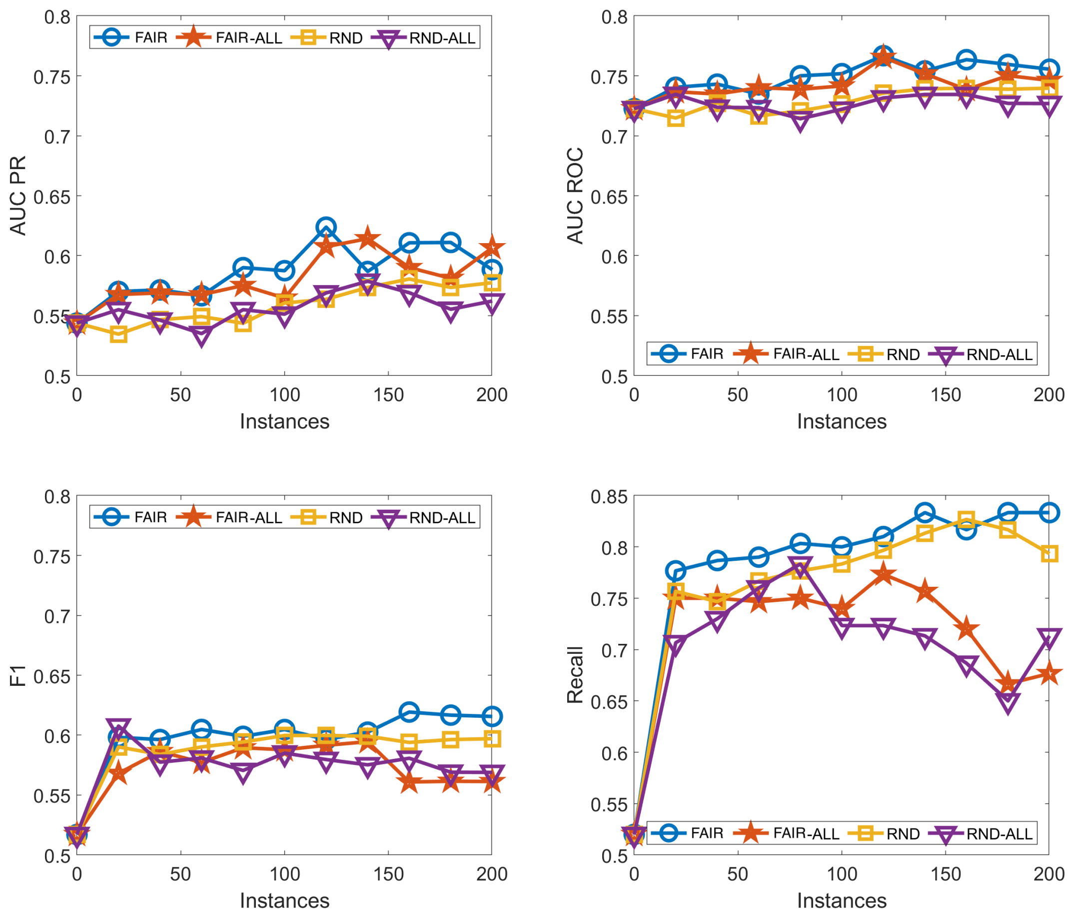

Figure 6: Nell Sports Results showing AUCPR, AUCROC, F1 and Recall.

and added to the current training set. Finally, the model is actively updated and the process is repeated until our budget is exhausted. Our empirical evaluations on standard relational domains demonstrate the efficacy of the proposed approach when compared to baseline techniques.

More rigorous evaluations on larger domains is the immediate next step. Extending the algorithm to dynamic domains where the feature sets can change consistently can result in interesting insights. Finally, performing large scale human evaluation for explainability and interpretability remains interesting direction of future research.

ACKNOWLEDGMENTS

The authors thank the anonymous reviewers. SN and NR gratefully acknowledge the support of AFOSR award FA9550-18-1-0462. Any opinions, findings, and conclusion or recommendations expressed in this material are those of the authors and do not necessarily reflect the view of the AFOSR,or the US government.

REFERENCES

[1] M. Bilgic and L. Getoor. 2007. VOILA: efficient feature-value acquisition for classification *(NCAI)*.
[2] Mustafa Bilgic and Lise Getoor. 2008. Effective label acquisition for collective classification. In *KDD*.
[3] Mustafa Bilgic, Lilyana Mihalkova, and Lise Getoor. 2010. Active learning for networked data. In *ICML*.
[4] H. Blockeel. 1999. Top-down induction of first order logical decision trees. *AIC* (1999).
[5] A. Carlson, J. Betteridge, B. Kisiel, B. Settles, E. Hruschka Jr, and T. Mitchell. 2010. Toward an Architecture for Never-Ending Language Learning.. In *AAAI*.
[6] William W Cohen. 1990. Learning approximate control rules of high utility. In *Machine Learning Proceedings 1990*. Elsevier.
[7] Mark Craven and Jude W Shavlik. 1996. Extracting tree-structured representations of trained networks. In *NeurIPS*.
[8] J.H. Friedman. 2001. Greedy function approximation: A gradient boosting machine. *Annals of Statistics* (2001).
[9] L. Getoor and B. Taskar. 2007. *Introduction to Statistical Relational Learning*. MIT Press.
[10] Ming Ji and Jiawei Han. 2012. A variance minimization criterion to active learning on graphs. In *Artificial Intelligence and Statistics*.
[11] P. Kanani and P. Melville. 2008. Prediction-time active feature-value acquisition for cost-effective customer targeting. *Workshop on Cost Sensitive Learning at NIPS* (2008).
[12] K. Kersting and L. De Raedt. 2007. Bayesian Logic Programming: Theory and Tool. In *An Introduction to Statistical Relational Learning*.

[13] Tushar Khot, Sriraam Natarajan, and Jude W Shavlik. 2014. Relational One-Class Classification: A Non-Parametric Approach.. In *AAAI*.
[14] Solomon Kullback and Richard A Leibler. 1951. On information and sufficiency. *The annals of mathematical statistics* (1951).
[15] Ankit Kuwadekar and Jennifer Neville. 2011. Relational active learning for joint collective classification models. In *ICML*.
[16] Christopher Leckie and Ingrid Zukerman. 1993. An inductive approach to learning search control rules for planning. In *IJCAI*.
[17] D. D. Lewis and J. Catlett. 1994. Heterogeneous uncertainty sampling for supervised learning *(ICML)*.
[18] Sofus A Macskassy. 2009. Using graph-based metrics with empirical risk minimization to speed up active learning on networked data. In *KDD*.
[19] P. Melville, M. Saar-Tsechansky, et al. 2004. Active feature-value acquisition for classifier induction *(ICDM)*.
[20] P. Melville, M. Saar-Tsechansky, et al. 2005. An expected utility approach to active feature-value acquisition *(ICDM)*.
[21] S MINTON. 1988. Quantitative Results Concerning the Utility of Explanation-Based Learning. In *AAAI-88*.
[22] Sriraam Natarajan, Srijita Das, Nandini Ramanan, Gautam Kunapuli, and Predrag Radivojac. 2018. On Whom Should I Perform this Lab Test Next? An Active Feature Elicitation Approach.. In *IJCAI*.
[23] S. Natarajan, T. Khot, K. Kersting, B. Gutmann, and J. Shavlik. 2012. Gradient-based boosting for statistical relational learning: The relational dependency network case. *MLJ* (2012).
[24] P. Odom and S. Natarajan. 2016. Active advice seeking for inverse reinforcement learning *(AAMAS)*.
[25] Phillip Odom and Sriraam Natarajan. 2016. Actively Interacting with Experts: A Probabilistic Logic Approach. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer.
[26] Phillip Odom and Sriraam Natarajan. 2018. Human-Guided Learning for Probabilistic Logic Models. *Frontiers in Robotics and AI* 5 (2018).
[27] L.D. Raedt, K. Kersting, S. Natarajan, and D. Poole. 2016. *Statistical relational artificial intelligence: Logic, probability, and computation*. Morgan Claypool.
[28] M. Richardson and P. Domingos. 2006. Markov logic networks. *MLJ* (2006).
[29] M. Saar-Tsechansky, P. Melville, and F. Provost. 2009. Active feature-value acquisition. *Manag Sci* (2009).
[30] B. Settles. 2012. Active learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning* (2012).
[31] Burr Settles, Mark Craven, and Soumya Ray. 2008. Multiple-instance active learning. In *NIPS*.
[32] H Sebastian Seung, Manfred Opper, and Haim Sompolinsky. 1992. Query by committee. In *Proceedings of the fifth annual workshop on Computational learning theory*. ACM.
[33] Jude W Shavlik and Geoffrey G Towell. 1989. Combining explanation-based learning and artificial neural networks. In *Proceedings of the sixth international workshop on Machine learning*. Elsevier.
[34] Lixin Shi, Yuhang Zhao, and Jie Tang. 2012. Batch mode active learning for networked data. *ACM Transactions on Intelligent Systems and Technology (TIST)* (2012).
[35] Prasad Tadepalli. 1991. A Formalization of Explanation-Based Macro-operator Learning.. In *IJCAI*.
[36] M. Thahir, T. Sharma, and M. K. Ganapathiraju. 2012. An efficient heuristic method for active feature acquisition and its application to protein-protein interaction prediction. *BMC Proc* (2012).
[37] S. Tong and D. Koller. 2000. Active learning for parameter estimation in Bayesian networks *(NIPS)*.
[38] S. Tong and D. Koller. 2001. Active learning for structure in Bayesian networks *(IJCAI)*.
[39] S. Tong and D. Koller. 2001. Support vector machine active learning with applications to text classification. *J Mach Learn Res* (2001).
[40] Z. Zhang, J. Kwok, and D. Yeung. 2004. Surrogate Maximization/Minimization Algorithms for AdaBoost and the Logistic Regression Model. In *ICML*.