Hitting Sets for Cluster Explanations and Practical Applications

Contact Information:
Bob Downey (University of Virginia, Charlottesville, VA) – kuh4bd@virginia.edu
Professor S. S. Ravi - ssravi@virginia.edu

Background

Goals of Explainable AI (XAI)

- Enable human users to understand, appropriately trust, and effectively manage the generation of outcomes by software tools based on AI.
- To produce more explainable models, while maintaining a high level of learning performance (prediction accuracy).

Unsupervised Learning and Clustering

- Unsupervised Learning: Analyzes and finds patterns in unlabeled datasets.
- Clustering: A ML application that divides a set of data points into groups such that the data points in any given group are more like other data points in the same group than those in other groups.

Integer Linear Programming (ILP)

- A method to achieve an optimal outcome (such as maximum profit or lowest cost) in a mathematical model whose requirements are represented by linear constraints and objectives.
- Solutions obtained using the Gurobi solver
- Variables are constrained to have values from suitable subsets of integers (e.g., {0, 1}).

Project Description

Motivation

- Obtaining an optimal explanation is computationally intractable (NP-hard) in general.
- Instead, we use ILP and heuristic algorithms to achieve near-optimal explanations (i.e. descriptors).

Task

- Given a cluster produced by an algorithm, we want to understand why a clustering algorithm placed those data items in the same cluster.
- Explore algorithmic techniques for developing explanations of outputs produced by clustering algorithms.

General Approach

- Explaining outputs of clustering algorithms using auxiliary information (tags).
- This approach was developed in [Davidson, Gourru & Ravi, 2018].

Explaining Clusters Using Tags

Tag Descriptors

- $S = \{s_1, \dots, s_n\}$ are items in the dataset
- $T = \{t_1, \dots, t_N\}$ are tags that the items can have (that describe the items in some way)
- Each $s_i \in S$ is associated with a subset $T_i \subseteq T$ of tags
- A tag descriptor D is represented as a collection of tags $D = \{t_1, \dots, t_K\}$ where $D_k \subseteq T$ that solves the **Minimum Hitting Set Problem** for cluster C_k.

Hitting Sets

- A set of tags for which every item in the dataset is guaranteed to have at least one tag from the hitting set.
- Disjunctive Explanation: {t4, t11, t3}
- Conjunctive Normal Form (CNF): {t4, t6, t3} and {t8, t2, t7}

Overall Goal

- Develop both exact and heuristic approximation algorithms for the **Minimum Hitting Set** problem.
- Experiments are being done on both synthetic and real-world data sets.

Current Work

Tools/software/data

- Developed a simple heuristic approximation method to solve the Minimum Hitting Set Problem which provides "small enough" descriptors for a given cluster.
- Extended this heuristic with the Conjunctive Normal Form (CNF) explanation.
- Compared these algorithms to an exact solver which uses Gurobi optimization software to find a minimum hitting set using Integer Linear Programming Methods (ILPs).

```
[8, 5]
[1, 1, 1, 1, 1, 0]
[2, 0, 1, 0, 1, 0]
[3, 1, 0, 0, 1, 1]
[4, 1, 0, 1, 1, 1]
[5, 1, 1, 1, 0, 0]
[6, 1, 1, 1, 0, 1]
[7, 1, 0, 1, 0, 1]
[8, 0, 0, 1, 1, 1]

['t1', 't3', 't4']

['t3', 't4']

(['t1', 't3', 't4'], ['t5', 't2'])
```

Figure 1: Tag Descriptors for a Synthetic Dataset

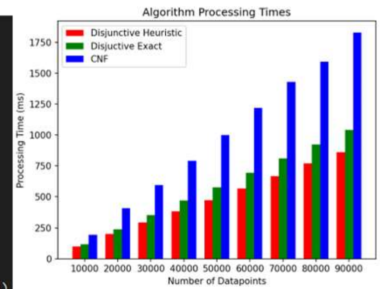

Figure 2: Hitting Set Solver Algorithms Processing Times

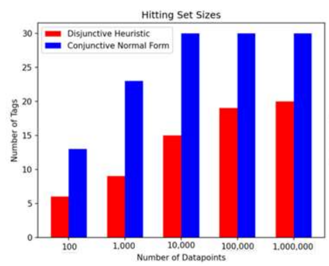

Figure 3: Hitting Set Sizes for Heuristic Algorithms

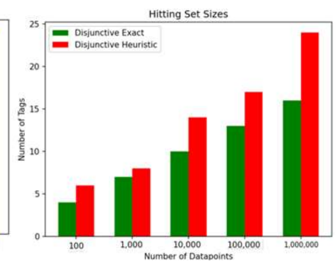

Figure 4: Hitting Set Sizes for Approximation and Exact Algorithms

Real World Application

- 1994 Adults Census Data: group census data into two clusters and develop tags and apply them to this data to determine what individuals in the clusters have in common.

```
Cluster 1
Cluster Size: 5201
Approximate Hitting Set: ['t18']
CNF Descriptor: (['t18'], ['t20', 't3', 't11', 't4', 't16', 't7'])
Minimum Hitting Set: ['t18']

Cluster 2
Cluster Size: 25518
Approximate Hitting Set: ['t19', 't3', 't11', 't20', 't17', 't4', 't7']
CNF Descriptor: (['t19', 't3', 't11', 't20', 't17', 't4', 't7'], ['t6', 't10', 't5', 't9', 't16', 't8'])
Minimum Hitting Set: ['t20', 't21']
```

Figure 5: Hitting Set Results for Adults Census Dataset

C1 Tag Percentages: {'t1': 0.173, 't2': 8.981, 't3': 90.846, 't4': 79.135, 't5': 20.865, 't6': 49.308, 't7': 50.692, 't8': 28.788, 't9': 32.538, 't10': 38.673, 't11': 84.808, 't12': 3.096, 't13': 0.923, 't14': 10.308, 't15': 0.865, 't16': 52.288, 't17': 47.712, 't18': 100.0, 't19': 0.0, 't20': 93.308, 't21': 6.692}

C2 Tag Percentages: {'t1': 0.035, 't2': 14.265, 't3': 85.7, 't4': 65.498, 't5': 34.502, 't6': 44.602, 't7': 55.398, 't8': 28.883, 't9': 32.069, 't10': 39.048, 't11': 85.621, 't12': 3.155, 't13': 0.964, 't14': 9.445, 't15': 0.815, 't16': 29.173, 't17': 70.827, 't18': 8.277, 't19': 91.723, 't20': 72.516, 't21': 27.484}

Figure 6: Cluster Tag Percentages for Adults Census Dataset

```
(t1, t2, t3) No income, govt. employee, private/other
(t4, t5) Does not have/has at least bachelors degree
(t6, t7) Number of years of education less/greater than 10 years
(t8, t9, t10) General Service, Corporate Support, Specialist/Managerial position
(t11, t12, t13, t14, t15) White, Asian-Pac-Islander, Amer-Indian-Eskimo, Black, Other
(t16, t17) Female, male
(t18, t19) Hours per week below or above 40 hours per week
(t20, t21) Yearly income below or above $50K
```

Figure 7: Tag Definitions for Adults Census Dataset

Future Work

- Automate a new process for efficiently generating optimal tags for both continuous and categorical data.
- See how the number of tags effects the quality of the hitting set descriptor.
- Assist Dr. Ravi in writing a paper based on this work.

Acknowledgement: This work is supported by NSF Grant IIS-1908530.

Computing for Global Challenges

University of Virginia
BIOCOMPLEXITY INSTITUTE