***Implementation of anytime and contract decision tree classifiers algorithms based on Esmeir & Markovitch papers.***

Amir Avivi: 305183873, [saavivi@campus.technion.ac.il](mailto:saavivi@campus.technion.ac.il)

Eliad Ben Yishay: 312193634, ebenyishay@campus.technion.ac.il

## Introduction

Nowadays success and popularity of Deep learning algorithms has risen the demand for high performance learning algorithms, even at a cost of large computational resources.   
DNN’s have several drawbacks, we will aim to address a few among them:  
 1. “black box”/”black magic” - uninterpretable decisions, doesn’t suite for medicine or military. For example: need to explain why drone killed a suspect.  
2. DNN’s need an exponentially big dataset in the number of features - not always present. For example: rare disease - only tens/hundreds of documented cases.   
Esmeir & Markovitch work on anytime and contract decision tree classifiers was based on the fact that tree based algorithms are probably the most interpretable family of learning algorithms known to us today, but most of these algorithms להוסיף על האדישות למשאבי חישוב גדולי יותר  
In our project we will implement some of their algorithms described in the paper *“Anytime Learning of Decision Trees” - 2007*. We will conduct sanity level benchmarks and analyze the results.

## Main algorithmic ideas summary

Heuristics: The main heuristics that these algorithms follow is based on Occam's razor principle – the simpler solution is the better one.  
For decision tree classifiers the projection of this philosophical view is – whenever possible, prefer a smaller decision tree (not always true).  
Thus, the algorithms basic idea is to start with a well-defined, easy calculated decision tree classifier (ID3 in most cases), and then improving the sub-trees that derive from that tree to reduce its size and enhance its performance.  
Anytime: The anytime version of these algorithms will start with a known, easily computed decision tree classifier, and try to improve its sub-trees as long as you allow them to run. The only constraint for such algorithms, is that if it is given any reasonable time to run (in order to build the basic decision tree) it should return some classifier.  
Contract: The contract version for the algorithms will have a user given argument which indicates the contract between the user and the algorithm, meaning, the bigger the parameter, then the algorithm can use larger computational resources to calculate the classifier.

## Project goals

1. Implement algorithms based on the paper *“Anytime Learning of Decision Trees – 2007 Saher Esmeir & Shaul Markovitch”*.
2. Conduct sanity level benchmarks.
3. Analyze the results.

## Tests and experiments plan

We will test our implementation against some of the datasets mentioned in *“Anytime Learning of Decision Trees – 2007 Saher Esmeir & Shaul Markovitch”*.

* Balance scale
* Breast cancer
* Glass
* Iris
* Monk 1-3
* Mushroom
* Solar flare
* Tic-tac-toe
* Wine
* Zoo

Some datasets are irretrievable, so we will build new datasets, representing the same concept. These datasets are:

* Numeric XOR 3D
* Numeric XOR 4D
* Multiplexer 20
* XOR 5
* XOR 5, with 10% noise
* XOR 10

We will compare our implementation against several known IDT algorithms using the following metrics:

* Average size and standard deviation of the induced trees
* T-test significance using
* The generalization accuracy’s Average and standard deviation

## Technical

**Programming language:** Python 3.7

**Version control:** GitHub

**Instructor:** Guy Kushilevitz