ITEA automatic report

ITEA_summarizer

Wednesday 24$^{\text{th}}$ November, 2021, 20:12

Automatic report created by *ITEA_summarizer* package. This report makes usage of several methods to automatically inspect and explain the final expression found in the evolutionary process performed by the ITEA algorithm.

Descriptive statistics of the data

Reporting descriptive statistics for 5 (from a total of 8) features contained on the training data. The features were selected based on the absolute final importance.

	AveBedrms	MedInc	AveOccup	AveRooms	Latitude
count	13828.000000	13828.000000	13828.000000	13828.000000	13828.000000
mean	1.097533	3.876745	3.128660	5.436556	35.651238
std	0.445688	1.903102	12.646130	2.449446	2.134064
min	0.333333	0.499900	0.692308	0.888889	32.550000
25%	1.006623	2.568575	2.432189	4.459802	33.940000
50%	1.049552	3.538750	2.819702	5.232422	34.270000
75%	1.100283	4.756600	3.282093	6.058566	37.720000
max	25.636364	15.000100	1243.333333	141.909091	41.950000

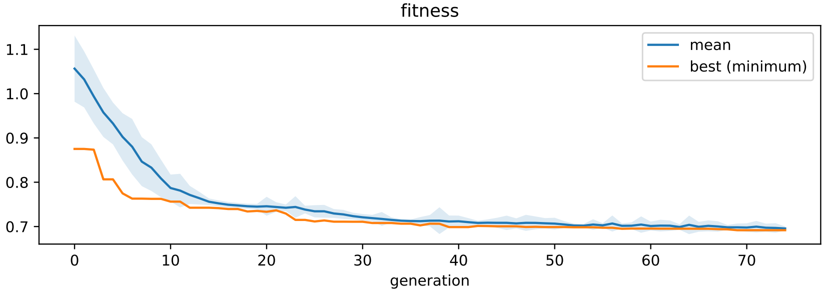
Algorithm Hyper-parameters

The following hyperparameters were used to execute the algorithm. If the random_state parameter was set to an integer value (or a numpy randomState instance was given), then it is possible to repeat the exact execution by using the same training data and the parameters listed below.

```
expolim : (-1, 1)
gens : 75
max_terms : 5
popsize : 75
random_state : 42
simplify_method : None
verbose : 10
tfuncs : [log, sqrt.abs, id, sin, cos, exp]
```

Evolution convergence

The algorithm took 80.499 seconds to completely run. Below are the plots for the average fitness of the population and the best individual fitness for each generation.



Best expression

The best expression corresponds to the expression with the best fitness on the last generation before the evolution ends. Not necessarily it will be the simpliest or the global optimum expression of the evoution. The final expression is a regressor with a fitness of 0.69172, and the number of IT terms is 5. Below is an LaTeX representation of the expression:

$$\begin{aligned} ITExpr = &\underbrace{\beta_0 \cdot sqrt.abs(\frac{HouseAge \cdot Longitude}{AveRooms \cdot AveBedrms \cdot Latitude})}_{\text{term 0}} + \underbrace{\beta_1 \cdot sqrt.abs(MedInc)}_{\text{term 1}} \\ &+ \underbrace{\beta_2 \cdot sqrt.abs(\frac{MedInc \cdot HouseAge \cdot AveBedrms}{AveRooms \cdot AveOccup \cdot Latitude})}_{\text{term 2}} \\ &+ \underbrace{\beta_3 \cdot sqrt.abs(\frac{Longitude}{MedInc \cdot AveRooms \cdot Latitude})}_{\text{term 3}} \\ &+ \underbrace{\beta_4 \cdot sqrt.abs(\frac{MedInc \cdot HouseAge \cdot Population}{AveRooms \cdot Latitude})}_{\text{term 4}} + I_0 \end{aligned} \tag{1}$$

Best expression metrics

On the next page is reported a table containing the coefficients for the previous expression, as well as some metrics calculated for each term individually:

- **coef:** coefficient of each term (or coefficients, if the itexpr is an instance of ITExpr_classifier);
- **coef stderr:** the standard error of the coefficients;
- **disentang.:** mean pairwise disentanglement between each term when compared with the others;
- **M.I.:** mean continuous mutual information between each term when compared with the others;
- **pred. var.:** variance of the predicted outcomes for each term when predicting the training data.

	coef	func	coef stderr	disentang.	M.I.	pred. var.
term 0	-0.317	sqrt.abs	0.009	0.213	0.301	0.174
term 1	2.029	sqrt.abs	0.028	0.253	0.506	0.840
term 2	4.258	sqrt.abs	0.068	0.171	0.280	0.417
term 3	3.673	sqrt.abs	0.083	0.256	0.470	0.341
term 4	0.008	sqrt.abs	0.001	0.107	0.205	0.007
term 5	-4.299	intercept	0.079	0.000	0.000	0.000

Partial derivatives

$$\begin{aligned}
&\frac{\partial}{\partial MedInc} ITExpr \\
&\quad = \underbrace{1\beta_1 \cdot sqrt.abs'(MedInc)()}_{\text{term 1}} \\
&\qquad + \underbrace{1\beta_2 \cdot sqrt.abs'(\frac{MedInc \cdot HouseAge \cdot AveBedrms}{AveRooms \cdot AveOccup \cdot Latitude})(\frac{HouseAge \cdot AveBedrms}{AveRooms \cdot AveOccup \cdot Latitude})}_{\text{term 2}} \\
&\qquad + \underbrace{-1\beta_3 \cdot sqrt.abs'(\frac{Longitude}{MedInc \cdot AveRooms \cdot Latitude})(\frac{Longitude}{MedInc^2 \cdot AveRooms \cdot Latitude})}_{\text{term 3}} \\
&\qquad + \underbrace{1\beta_4 \cdot sqrt.abs'(\frac{MedInc \cdot HouseAge \cdot Population}{AveRooms \cdot Latitude})(\frac{HouseAge \cdot Population}{AveRooms \cdot Latitude})}_{\text{term 4}}
\end{aligned} \tag{2}$$

$$
\begin{aligned}
&\frac{\partial}{\partial HouseAge} ITExpr \\
&\quad = \underbrace{1\beta_0 \cdot sqrt.abs'(\frac{HouseAge \cdot Longitude}{AveRooms \cdot AveBedrms \cdot Latitude})(\frac{Longitude}{AveRooms \cdot AveBedrms \cdot Latitude})}_{\text{term 0}} \\
&\qquad + \underbrace{1\beta_2 \cdot sqrt.abs'(\frac{MedInc \cdot HouseAge \cdot AveBedrms}{AveRooms \cdot AveOccup \cdot Latitude})(\frac{MedInc \cdot AveBedrms}{AveRooms \cdot AveOccup \cdot Latitude})}_{\text{term 2}} \\
&\qquad + \underbrace{1\beta_4 \cdot sqrt.abs'(\frac{MedInc \cdot HouseAge \cdot Population}{AveRooms \cdot Latitude})(\frac{MedInc \cdot Population}{AveRooms \cdot Latitude})}_{\text{term 4}}
\end{aligned} \tag{3}
$$

$$
\begin{aligned}
&\frac{\partial}{\partial AveRooms} ITExpr \\
&\quad = \underbrace{-1\beta_0 \cdot sqrt.abs'(\frac{HouseAge \cdot Longitude}{AveRooms \cdot AveBedrms \cdot Latitude})(\frac{HouseAge \cdot Longitude}{AveRooms^2 \cdot AveBedrms \cdot Latitude})}_{\text{term 0}} \\
&\qquad + \underbrace{-1\beta_2 \cdot sqrt.abs'(\frac{MedInc \cdot HouseAge \cdot AveBedrms}{AveRooms \cdot AveOccup \cdot Latitude})(\frac{MedInc \cdot HouseAge \cdot AveBedrms}{AveRooms^2 \cdot AveOccup \cdot Latitude})}_{\text{term 2}} \\
&\qquad + \underbrace{-1\beta_3 \cdot sqrt.abs'(\frac{Longitude}{MedInc \cdot AveRooms \cdot Latitude})(\frac{Longitude}{MedInc \cdot AveRooms^2 \cdot Latitude})}_{\text{term 3}} \\
&\qquad + \underbrace{-1\beta_4 \cdot sqrt.abs'(\frac{MedInc \cdot HouseAge \cdot Population}{AveRooms \cdot Latitude})(\frac{MedInc \cdot HouseAge \cdot Population}{AveRooms^2 \cdot Latitude})}_{\text{term 4}}
\end{aligned} \tag{4}
$$

$$\begin{aligned}\frac{\partial}{\partial AveBedrms} ITExpr \\ = \underbrace{-1\beta_0 \cdot sqrt.abs'(\frac{HouseAge \cdot Longitude}{AveRooms \cdot AveBedrms \cdot Latitude})(\frac{HouseAge \cdot Longitude}{AveRooms \cdot AveBedrms^2 \cdot Latitude})}_{\text{term 0}} \\ + \underbrace{1\beta_2 \cdot sqrt.abs'(\frac{MedInc \cdot HouseAge \cdot AveBedrms}{AveRooms \cdot AveOccup \cdot Latitude})(\frac{MedInc \cdot HouseAge}{AveRooms \cdot AveOccup \cdot Latitude})}_{\text{term 2}}\end{aligned} \tag{5}$$

$$\frac{\partial}{\partial Population} ITExpr = \underbrace{1\beta_4 \cdot sqrt.abs'(\frac{MedInc \cdot HouseAge \cdot Population}{AveRooms \cdot Latitude})(\frac{MedInc \cdot HouseAge}{AveRooms \cdot Latitude})}_{\text{term 4}} \tag{6}$$

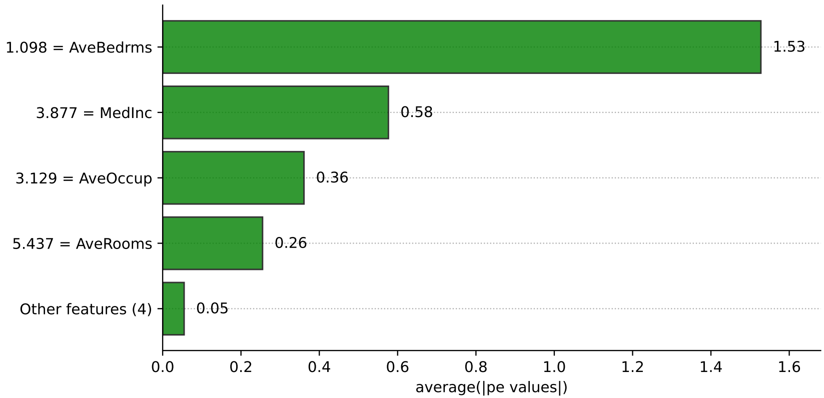
$$\begin{aligned}\frac{\partial}{\partial AveOccup} ITExpr \\ = \underbrace{-1\beta_2 \cdot sqrt.abs'(\frac{MedInc \cdot HouseAge \cdot AveBedrms}{AveRooms \cdot AveOccup \cdot Latitude})(\frac{MedInc \cdot HouseAge \cdot AveBedrms}{AveRooms \cdot AveOccup^2 \cdot Latitude})}_{\text{term 2}}\end{aligned} \tag{7}$$

$$
\begin{aligned}
&\frac{\partial}{\partial Latitude} ITExpr \\
&\quad = \underbrace{-1\beta_0 \cdot sqrt.abs'(\frac{HouseAge \cdot Longitude}{AveRooms \cdot AveBedrms \cdot Latitude})(\frac{HouseAge \cdot Longitude}{AveRooms \cdot AveBedrms \cdot Latitude^2})}_{\text{term 0}} \\
&\quad\quad + \underbrace{-1\beta_2 \cdot sqrt.abs'(\frac{MedInc \cdot HouseAge \cdot AveBedrms}{AveRooms \cdot AveOccup \cdot Latitude})(\frac{MedInc \cdot HouseAge \cdot AveBedrms}{AveRooms \cdot AveOccup \cdot Latitude^2})}_{\text{term 2}} \\
&\quad\quad + \underbrace{-1\beta_3 \cdot sqrt.abs'(\frac{Longitude}{MedInc \cdot AveRooms \cdot Latitude})(\frac{Longitude}{MedInc \cdot AveRooms \cdot Latitude^2})}_{\text{term 3}} \\
&\quad\quad + \underbrace{-1\beta_4 \cdot sqrt.abs'(\frac{MedInc \cdot HouseAge \cdot Population}{AveRooms \cdot Latitude})(\frac{MedInc \cdot HouseAge \cdot Population}{AveRooms \cdot Latitude^2})}_{\text{term 4}}
\end{aligned} \tag{8}
$$

$$
\begin{aligned}
&\frac{\partial}{\partial Longitude} ITExpr \\
&\quad = \underbrace{1\beta_0 \cdot sqrt.abs'(\frac{HouseAge \cdot Longitude}{AveRooms \cdot AveBedrms \cdot Latitude})(\frac{HouseAge}{AveRooms \cdot AveBedrms \cdot Latitude})}_{\text{term 0}} \\
&\qquad + \underbrace{1\beta_3 \cdot sqrt.abs'(\frac{Longitude}{MedInc \cdot AveRooms \cdot Latitude})(\frac{1}{MedInc \cdot AveRooms \cdot Latitude})}_{\text{term 3}}
\end{aligned} \tag{9}
$$

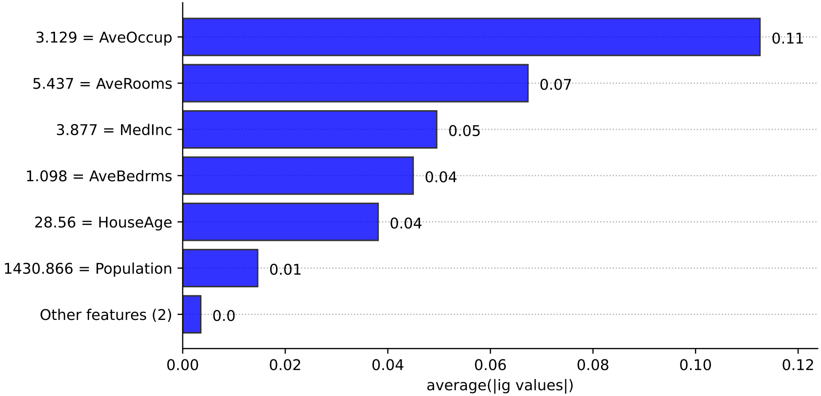
Global importances with *Average partial Effects*

Feature importances with Average Partial Effects. This method attributes the importance to the i-th variable by calculating the average of the partial derivative w.r.t. i, evaluated for all data in the training set.



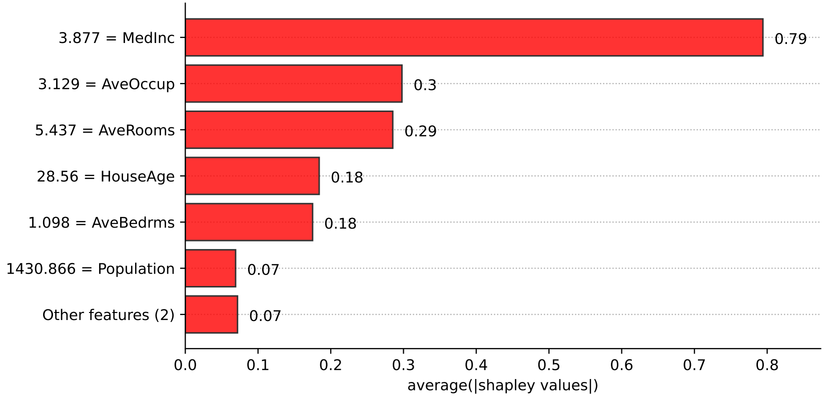
Global importances with *Integrated Gradients*

Feature importance using the Integrated Gradients method. The idea of the integrated gradient is to calculate a local importance score for a feature i by evaluating — given a baseline $\mathbf{x}'$ and a specific point $\mathbf{x}$ — the integral of the models' gradients $\frac{\partial f}{\partial x_i}$ along a straight line between the baseline and the specific point. Since gradients describe how minimal local changes interfere with the model's predictions, the calculated importance represents the accumulation of the gradient of each variable to go from the baseline to the specific point.


3.129 = AveOccup
5.437 = AveRooms
3.877 = MedInc
1.098 = AveBedrms
28.56 = HouseAge
1430.866 = Population
Other features (2)
0.11
0.07
0.05
0.04
0.04
0.01
0.0
0.00
0.02
0.04
0.06
0.08
0.10
0.12
average(|ig values|)

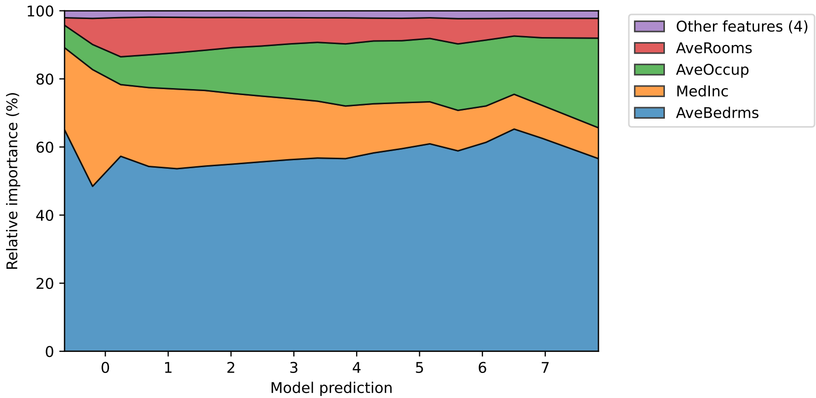
Global importances with *Shapley Values*

Feature importance estimation through approximation of the Shapley values with the gradient information. The shapley values are based on coalition game theory, where players can contribute differently to a result of a game. The total contribution of each player is the Shapley value, and represents the overall contribution that the player presents, based on all possible distribution of teams.


3.877 = MedInc
3.129 = AveOccup
5.437 = AveRooms
28.56 = HouseAge
1.098 = AveBedrms
1430.866 = Population
Other features (2)
0.79
0.3
0.29
0.18
0.18
0.07
0.07
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
average(|shapley values|)

Normalized partial Effects

Feature importances with Normalized Partial Effects. To create this plot, first, the output interval is discretized. Then, for each interval, the partial effect of all samples in the training set that results in a prediction within the interval are calculated. Finally, they are normalized in order to make the total contribution by 100%.



Partial Effects at the Means

Partial Effects plots created by fixing the co-variables at the means and evaluating the model's output when only one variable changes. For simplicity, at most 5 variables are selected to create the plot (the 5 most important variables considering their Average Partial Effects).