DOCTORAL PROGRAM IN ENGINEERING SCIENCES AT ITESO

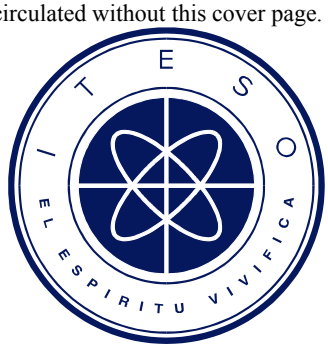
ADAPTATIVE DISCOVERING ALGORITHM BASED ON NEURAL NETWORKS

Edgar-D. Ramirez-de-L.
Maria-del-P. Pozos-P.
Ivan Villalon-Turrubiates

Algorithm 3 - chooseRandomFeatures.doc

May 15, 2018
Tlaquepaque, Mexico 45604

Doctoral Program in Engineering Sciences
ITESO (*Instituto Tecnológico y de Estudios Superiores de Occidente*)



ADAPTATIVE DISCOVERING ALGORITHM BASED ON NEURAL NETWORKS

Edgar-D. Ramirez-de-L.
Maria-del-P. Pozos-P.
Ivan Villalon-Turrubiates

May 15, 2018

Doctoral Program in Engineering Sciences
ITESO (*Instituto Tecnológico y de Estudios Superiores de Occidente*)

Tlaquepaque, Mexico 45604
Tel +52 33 3669 3598
E-mail: dci@iteso.mx

Keywords *neural networks, deep learning, machine learning, classifiers, python*

Abstract

We present the Algorithm 3 (`chooseRandomFeatures`) which is part of the Adaptative Discovering Algorithm based on Neural networks (ADAN algorithm).

Algorithm 3 chooseRandomFeatures

Require: $args \neq \varnothing \wedge sourceDataFrame \neq \varnothing$
Ensure: $returnedDataFrame \neq \varnothing \wedge size(returnedDataFrame.keys) = ssfe$

$tsfe \leftarrow size(sourceDataFrame.keys)$
$spfe \leftarrow readArg('spfe', args)$
$ssfe \leftarrow int(tsfe * spfe)$
$series \leftarrow \varnothing$
$processed \leftarrow \varnothing$
$returnedDataFrame \leftarrow DataFrame()$
$x \leftarrow randomIntBetween(0, ssfe - 1)$
while $size(returnedDataFrame.keys) < ssfe$ **do**
  while $x \in processed$ **do**
    $x \leftarrow randomIntBetween(0, ssfe - 1)$
  end while
  $processed \leftarrow append(processed, x)$
  $serie \leftarrow getColumn(sourceDataFrame, x)$
  $returnedDataFrame \leftarrow concat(returnedDataFrame, serie, 'columns')$
  $x \leftarrow randomIntBetween(0, ssfe - 1)$
  return $returnedDataFrame$
end while

Some considerations about the Algorithm 3:

1. The function `int(arg)`, in line 3, will return the int value for the `arg` specified. It's a casting.

2. In line 6, `DataFrame()` returns a new empty data frame, and this will depend from the programming language, tools, libraries and so on. In our case, for Python, we use Pandas, as we said at the beginning of this document.

3. The function `randomIntBetween(from, to)` (lines 7, 10 and 15) also will depend from the programming language. We think that is pretty obvious its behaviour, but is better make it clearer: will pick an integer random value, starting `from` and including the `to` value. In our Python implementation we use `random.randint(from, to)`.

4. The function `getColumn(dataFrame, columnNumber)` will return the column in `columnNumber` for the specified `dataFrame`. In our implementation, using Pandas, we have: `serie = source_df[source_df.columns[x]]`