Decision Tree Classifier

sales@nextgeninvent.com

Contents

1. Introduction

NextGen Invent corporation is a leading global technology services provider. Our multi-disciplinary teams, Data Scientists, Change Agents, Strategists and Process Analysis framework are delivering Robotic Process Automation solution to its clients.
NGI in partnership with Automation Anywhere is delivering Stats/ML/AI/NLP/OR/Deep Learning models as ready to use Cloud Bots that are providing efficiencies across business processes, improved customer conversion rate and reduced churn rate.

1.1. Overview

A business user could use this bot, which is based on the principles of 'Decision Tree Classifier' to expedite business decisions. The Bot takes in the input data in CSV file and performs the Decision Tree based Classification and provides output in the MS excel.

1.2. Use cases

- Identifying key decision factors
- Decision trees help organizations to view alternatives decisions
- Decision tree methodology is also commonly used data mining method for establishing classification systems
- This bot could predict the binary and multiclass classification problems

2. Requirements & Prerequisites

2.1.System Requirements

For the PC or server where the bot needs to run:

Device	Processor	RAM	Storage	Network
Machine	Intel Core i5 or higher	8GB or higher	16GB	1GbE
Bot Creator and Bot Runner	No additions to the machine requirements	No additions to the machine requirements	No additions to the machine requirements	No additions to the machine requirements

2.2.Prerequisites

Software's required -

Automation Anywhere Plugin version	Excel Requirement Version
11.x	2013 version or higher

Data Requirement-

Data should be pre-processed before feeding it to the bot and it should contain at least 4,500 rows for training with around 8-9 minimum features every time to obtain desired accuracy.

3. Getting Started

3.1. Skill Matrix

- A potential companion to analytics for classification problem
- Each internal node of the decision tree denotes an attribute and each leaf node denotes a class label outcome
- The Bot takes in the CSV file and performs the decision tree-based classification and output in the MS Excel

3.2. Installation Hierarchy

- Download the bot from Bot Store.
- Double click on the .msi file.
- On Welcome to Installation wizard, click Next to continue.
- Click I agree to the terms in the license agreement radio button to accept the agreement.
- Get/Copy the License key from Bot Store Downloads into License Key, click Next to continue.
- Click Install to begin the installation.
- Click Finish to complete the installation.
- To view the installation, go to 'My Tasks' folder on AAE Client to see bot files.

3.3. Configuration

Inputs:

Parameter	Type	Direction	Additional Info
Path	String	Input	Path of the CSV file For example: C:\~\Employee_data.csv
ColumnCount	Numeric	Input	Count of columns available in CSV file being provided with input data
InputColumnNames	String	Input	Name of columns being provided in CSV file. For example: 'Age','Work Class','Education','Marital_Status','Occupation','Relationship','Race','Gender','Hours_Per_Week','Native_Country','Income'
InputColumnDataType	String	Input	Data Type of columns being provided in CSV file in the same order as column names. For example, 'Numeric','String','String','String','String','String','String','String','Numeric','String','String'

MissingNumberValueReplacemsent	String	Input	A value that can be used by Model in case CSV file has blank field/s or missing numerical value/s.
IgnoreColumns	String	Input	Names of the column that model should ignore while calculating correlation. For example: 'Race'
PredictionColumnName	String	Input	Header of the Column on which you want to predict Ex: 'Income'

Input Sample:

Ex: This input tells us about the different salaries of the person based on various features like Education, occupation, etc.

Age	Work Class	Education	Marital_ Status	Occupation	Relationship	Race	Gender	Hours_Per_ Week	Native_ Country	Income
39	Private	HS-grad	Married-civ-spouse	Prof-specialty	Husband	White	Male	45		>50K
25	Local-gov	Bachelors	Never-married	Adm-clerical	Own-child	Black	Female	40	United-States	<=50K
20	Private	HS-grad	Never-married	Machine-op-inspct	Own-child	White	Male	40	United-States	<=50K
46	Private	Masters	Married-civ-spouse	Prof-specialty	Wife	White	Female	22	United-States	>50K
40	Private	Assoc-voc	Married-civ-spouse	Craft-repair	Husband	Black	Male	45	United-States	<=50K
66	Federal-gov	10th	Married-civ-spouse	Craft-repair	Husband	White	Male	40	United-States	<=50K
30	Private	Assoc-voc	Divorced	Adm-clerical	Unmarried	White	Female	24	United-States	<=50K
36	Private	7th-8th	Married-civ-spouse	Craft-repair	Husband	White	Male	40	United-States	<=50K
57	Local-gov	HS-grad	Married-civ-spouse	Craft-repair	Husband	White	Male	40	United-States	>50K
46	Private	Some-college	Married-civ-spouse	Exec-managerial	Husband	White	Male	48	United-States	>50K
27	Private	HS-grad	Never-married	Other-service	Unmarried	White	Female	64	United-States	<=50K
33	Private	HS-grad	Married-civ-spouse	Craft-repair	Husband	Black	Male	40	United-States	<=50K
58	Private	11th	Married-civ-spouse	Sales	Husband	White	Male	40	United-States	<=50K
30	Private	HS-grad	Divorced	Transport-moving	Not-in-family	White	Male	55	United-States	
26	Private	Assoc-acdm	Never-married	Machine-op-inspct	Other-relative	White	Female	15	United-States	
81		Assoc-voc	Divorced		Unmarried	White	Female	1		
32	Private	10th	Married-civ-spouse	Transport-moving	Husband	White	Male	40	United-States	
22	Private	Some-college	Never-married	Adm-clerical	Own-child	White	Male	40	United-States	
31	Private	HS-grad	Married-civ-spouse	Machine-op-inspct	Wife	White	Female	40	United-States	
29	Private	HS-grad	Separated	Sales	Unmarried	White	Female	35	United-States	
35		Bachelors	Married-civ-spouse		Wife	White	Female	55	United-States	
30		Bachelors	Never-married		Not-in-family	Asian-Pac-Islander	Female	99	United-States	
34	Private	Doctorate	Married-civ-spouse	Prof-specialty	Husband	White	Male	60	United-States	
54	Private	Bachelors	Married-civ-spouse	Exec-managerial	Husband	Asian-Pac-Islander	Male	50	Japan	
37	Private	Some-college	Divorced	Adm-clerical	Unmarried	White	Female	39	United-States	
22	Private	12th	Never-married	Protective-serv	Own-child	Black	Male	35	United-States	
34	Private	Bachelors	Never-married	Exec-managerial	Not-in-family	White	Female	55	United-States	
30	Private	HS-grad	Never-married	Craft-repair	Not-in-family	Black	Male	46	United-States	
38	Private	Bachelors	Divorced	Prof-specialty	Unmarried	Black	Female	45	United-States	
71		Doctorate	Married-civ-spouse		Husband	White	Male	10	United-States	

Outputs:

- Output: the result in the form of the MS Excel
- Predictions in the prediction class

  Ex: The below result output shows the two classes:
  Gold_Class: This column is taken as the predicted target

  Predicted_Class: This tells us the predicted type of the person salary i.e. it is above 50k or below it

Age	Work Class	Education	Marital_ Status	Occupation	Relationship	Race	Gender	Hours_Per_ Week	Native_ Country	Gold_ Class	Predicted_ Class
39	Private	HS-grad	Married-civ-spouse	Prof-specialty	Husband	White	Male	45		>50K	>50K
25	Local-gov	Bachelors	Never-married	Adm-clerical	Own-child	Black	Female	40	United-States	<=50K	<=50K
20	Private	HS-grad	Never-married	Machine-op-inspct	Own-child	White	Male	40	United-States	<=50K	<=50K
46	Private	Masters	Married-civ-spouse	Prof-specialty	Wife	White	Female	22	United-States	>50K	>50K
40	Private	Assoc-voc	Married-civ-spouse	Craft-repair	Husband	Black	Male	45	United-States	<=50K	<=50K
66	Federal-gov	10th	Married-civ-spouse	Craft-repair	Husband	White	Male	40	United-States	<=50K	<=50K
30	Private	Assoc-voc	Divorced	Adm-clerical	Unmarried	White	Female	24	United-States	<=50K	<=50K
36	Private	7th-8th	Married-civ-spouse	Craft-repair	Husband	White	Male	40	United-States	<=50K	<=50K
57	Local-gov	HS-grad	Married-civ-spouse	Craft-repair	Husband	White	Male	40	United-States	>50K	>50K
46	Private	Some-college	Married-civ-spouse	Exec-managerial	Husband	White	Male	48	United-States	>50K	>50K
27	Private	HS-grad	Never-married	Other-service	Unmarried	White	Female	64	United-States	<=50K	<=50K
33	Private	HS-grad	Married-civ-spouse	Craft-repair	Husband	Black	Male	40	United-States	<=50K	<=50K
58	Private	11th	Married-civ-spouse	Sales	Husband	White	Male	40	United-States	<=50K	<=50K
30	Private	HS-grad	Divorced	Transport-moving	Not-in-family	White	Male	55	United-States		<=50K
26	Private	Assoc-acdm	Never-married	Machine-op-inspct	Other-relative	White	Female	15	United-States		<=50K
81		Assoc-voc	Divorced		Unmarried	White	Female	1			<=50K
32	Private	10th	Married-civ-spouse	Transport-moving	Husband	White	Male	40	United-States		<=50K
22	Private	Some-college	Never-married	Adm-clerical	Own-child	White	Male	40	United-States		<=50K
31	Private	HS-grad	Married-civ-spouse	Machine-op-inspct	Wife	White	Female	40	United-States		>50K
29	Private	HS-grad	Separated	Sales	Unmarried	White	Female	35	United-States		<=50K
35		Bachelors	Married-civ-spouse		Wife	White	Female	55	United-States		>50K
30		Bachelors	Never-married		Not-in-family	Asian-Pac-Islander	Female	99	United-States		<=50K
34	Private	Doctorate	Married-civ-spouse	Prof-specialty	Husband	White	Male	60	United-States		>50K
54	Private	Bachelors	Married-civ-spouse	Exec-managerial	Husband	Asian-Pac-Islander	Male	50	Japan		>50K
37	Private	Some-college	Divorced	Adm-clerical	Unmarried	White	Female	39	United-States		<=50K
22	Private	12th	Never-married	Protective-serv	Own-child	Black	Male	35	United-States		<=50K
34	Private	Bachelors	Never-married	Exec-managerial	Not-in-family	White	Female	55	United-States		<=50K
30	Private	HS-grad	Never-married	Craft-repair	Not-in-family	Black	Male	46	United-States		<=50K
38	Private	Bachelors	Divorced	Prof-specialty	Unmarried	Black	Female	45	United-States		<=50K
71		Doctorate	Married-civ-spouse		Husband	White	Male	10	United-States		>50K

4. Logs

Error Handling In case of Errors, Error Logs & Screenshots are generated within Error Folder.

Each Bot folder contains the below hierarchy.

1 Error Folder

1.1 Logs

1.1.1 ErrorLogsMonth-Day-Year.txt: In case of any error, this file logs error message along with timestamp

1.2 Snapshots

1.2.1 Error SnapMonth-Day-Year-Hour-Minute-SecondMillisecond.png: In case of any error, this file captures screenshot of error.

Task Status of bot is set to failed in case of error.

5. Troubleshooting & Support

5.1 Support

Please Email bot.support@nextgeninvent.com for any support related questions.

Appendix A - Record of Changes

Version #	Date of Change	Author	Notes
V1.0	June 26th 2020	Nextgen Invent Corporation	

Appendix B - Acronyms

#	Acronym	Description
1.	AA	Automation Anywhere
2.	API	Application Programming Interface
3.	RAM	Random Access Memory
4.	OS	Operating System

Appendix C - References

#	Topic	Reference Link
1.	Nextgen Invent Corporation	https://www.nextgeninvent.com/