



NextGen
INVENT CORP
think. invent. solve.

Gaussian Mixture Model



Contents

1. INTRODUCTION	3
1.1. OVERVIEW	3
1.2. USE CASES	3
2. REQUIREMENTS & PREREQUISITES	4
2.1. SYSTEM REQUIREMENTS	4
2.2. PREREQUISITES	4
3. GETTING STARTED	5
3.1. SKILL MATRIX	5
3.2. INSTALLATION HIERARCHY	5
3.3. CONFIGURATION	5
4. LOGS	7
5. TROUBLESHOOTING & SUPPORT	8
5.1 SUPPORT	8
APPENDIX A - RECORD OF CHANGES	9
APPENDIX B - ACRONYMS	9
APPENDIX C - REFERENCES	9



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

The Bot takes in the **input** CSV file and prepare the clusters using Gaussian mixture model where sub-populations exists within normally distributed whole population, provide outputs in the **MS Excel**

1.2.Use cases

- Customer Segmentation to target most profitable customers sub-groups
- Customer Segmentation analysis has been used in marketing for various purposes such customer group at risk
- Helping Marketing managers in evaluating performance of various market tests to define future marketing strategy



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



3. Getting Started

3.1.Skill Matrix

- Rapid segmentation of homogeneous sub-groups
- This bot model could be used for cluster and sub-groups creations for a customer analytics project
- Prepare CSV file to pass it to the model and processing MS Excel output.

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:\~\ Wholesale customers.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: 'Channel','Region','Fresh','Milk','Grocery','Frozen','Detergents_Paper','Delicassen'
InputColumnDataType	String	Input	Data Type of columns being provided in CSV file in the same order as column names. For example: 'Numeric', 'Numeric', 'Numeric', 'Numeric', 'Numeric', 'Numeric', 'Numeric'



MissingNumberValueReplacement	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: 'Detergents_Paper'

Input Sample:

Input Ex:-

The customer segments data is included as a selection of 440 data points collected on data found from clients of a wholesale distributor in Lisbon, Portugal.

Channel	Region	Fresh	Milk	Grocery	Frozen	Detergents_Paper	Delicassen
2	3	12669	9656	7561	214	2674	1338
2	3	7057	9810	9568	1762	3293	1776
2	3	6353	8808	7684	2405	3516	7844
1	3	13265	1196	4221	6404	507	1788
2	3	22615	5410	7198	3915	1777	5185
2	3	9413	8259	5126	666	1795	1451
2	3	12126	3199	6975	480	3140	545
2	3	7579	4956	9426	1669	3321	2566
1	3	5963	3648	6192	425	1716	750
2	3	6006	11093	18881	1159	7425	2098
2	3	3366	5403	12974	4400	5977	1744
2	3	13146	1124	4523	1420	549	497
2	3	31714	12319	11757	287	3881	2931
2	3	21217	6208	14982	3095	6707	602
2	3	24653	9465	12091	294	5058	2168
1	3	10253	1114	3821	397	964	412
2	3	1020	8816	12121	134	4508	1080
1	3	5876	6157	2933	839	370	4478
2	3	18601	6327	10099	2205	2767	3181
1	3	7780	2495	9464	669	2518	501

Outputs:

1. Output: the result in the form of the MS Excel

Ex: The data has been divided into clusters using gaussian mixture model for the customer segmentation



Output Sample:

Channel	Region	Fresh	Milk	Grocery	Frozen	Delicassen	Cluster
1	1	514	8323	6869	529	1040	0
1	1	583	685	2216	469	18	0
1	1	964	4984	3316	937	7	0
1	1	1537	3748	5838	1859	806	0
1	1	1838	6380	2824	1218	295	0
1	1	1869	577	572	950	203	0
1	1	1956	891	5226	1383	1328	0
1	1	2083	5007	1563	1120	1550	0
1	1	2153	1115	6684	4324	411	0
1	1	2344	10678	3828	1439	490	0
1	1	2806	10765	15538	1374	2388	0
1	1	3191	1993	1799	1730	710	0
1	1	3317	6602	6861	1329	1215	0
1	1	4042	2204	1563	2286	689	0
1	1	4155	367	1390	2306	130	0
1	1	4720	1032	975	5500	56	2
1	1	5041	1115	2856	7496	375	2
1	1	5396	7503	10646	91	239	0
1	1	5909	23527	13699	10155	3636	1
1	1	6202	7775	10817	1183	1970	0
1	1	6373	780	950	878	285	0

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 15 th 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/

