NON-DISCLOSURE

The purpose of this document is the provision of information by OpenJaw Technologies Ltd. ("OpenJaw") to you ("Customer").

This document contains confidential information of OpenJaw. The Customer shall not disclose or make subsidiary disclosures of the confidential information contained in this document to any third party.

The Customer shall exercise all reasonable care to ensure that its employees, agents and associates maintain the confidential information in this document in strict confidence.

DISCLAIMER

While every effort has been made to ensure that this documentation is accurate and reliable, OpenJaw cannot assume liability for any damage caused by inaccuracies in this documentation or the failure of the software to function in a particular manner. The information in this document does not represent a commitment on the part of OpenJaw.

No part of this document may, without the prior written permission of OpenJaw, be reproduced or transmitted in any form or by any means, for any purpose other than the customer’s private use.

© 2021 OpenJaw. All rights reserved.

TRADEMARK ACKNOWLEDGEMENT

All product and service names used in this document are trademarks or registered trademarks of their respective companies.

The OpenJaw logo and xDistributor® are registered trademarks of OpenJaw Technologies.

AMAZON SAGEMAKER is a trademark of Amazon Technologies, Inc.

ABOUT THIS DOCUMENT

This document provides the guidelines for setting up the Customer Life Score process using Amazon Sagemaker. It explains how to run training job, create models and running batch inference jobs. It also explains how to prepare the training/inference datasets and understand the outputs.
# Contents

1. INTRODUCTION TO CUSTOMER LIFE SCORE ........................................................................................................... 1

2. USING CUSTOMER LIFE SCORE .......................................................................................................................... 2

2.1. CREATING A TRAINING JOB ........................................................................................................................... 2

2.2. CREATE A MODEL PACKAGE ............................................................................................................................. 4

2.3. CREATING BATCH INFEERENCE JOBS ................................................................................................................ 6

3. DATA PREPARATION AND INTERPRETATION .................................................................................................. 10

3.1. DATA PREPARATION ........................................................................................................................................ 10

3.2. STEPS FOR EXECUTING SAMPLE NOTEBOOK IN SAGEMAKER .............................................................. 11

3.3. OUTPUTS INTERPRETATION ............................................................................................................................. 11
1. Introduction to Customer Life Score

The Customer Life Score is a personalised, future-spend estimator, machine-learning driven solution, designed to mine your customer shopping data and produce easy access insights on your customers' future spend.

It can help marketers to run more effective campaigns through the identification of the right customers and right timing, resulting in reduced customer frictions, improved conversions and increased revenue.

Enterprise loyalty programs can also benefit from this solution in more effective membership management; from retaining and growing the customer-base to preventing potential churns. The solution returns the forecasted future-spend for each individual, as well as the analytics of their historical spend. The combination of these provides deep insights on an individual's spending pattern, relative to themselves and to other customers.

For details about Amazon Sagemaker, see https://aws.amazon.com/pm/sagemaker/.
2. Using Customer Life Score

There are three steps involved in the Customer Life Score process using Sagemaker. The steps are as follows:

1. Create a Training Job.
2. Creation of Sagemaker model package and model.
3. Creating Batch Inference Jobs.

2.1. CREATING A TRAINING JOB

To create a training job, do the following:

2. Click Algorithms.
3. Open AWS Marketplace Subscriptions and choose the algorithm for which you want to create training job.
4. Click Actions> Create training job.
5. Enter the Job name and select the Instance type. Prior to choosing the Instance type, refer the instance configuration and pricing in [https://aws.amazon.com/sagemaker/pricing/](https://aws.amazon.com/sagemaker/pricing/).
6. Open Create training job and choose the VPC to launch the EC2 instances needed to train the algorithm.
7. Specify the Hyperparameters settings. There are four parameters necessary to train the Customer Lifetime Score Algorithm:
   - **date_col**: Transaction timestamp of the format (yyyy-mm-dd HH:MM:SS).
   - **monetary_col**: Price of the item
   - **id_col**: Identifier column, for instance Customer ID.
   - **fwd_time_window**: The number of days to predict customer lifetime score. The number of days in fwd_time_window should be less than the number of days between overall first and last purchase of the training dataset. For example, training sample containing 2 years of purchase history can have a fwd_time_window of 180 days (approximately 6 months). Therefore, customer life score process will predict the purchase probability and customer life score for next 6 months.
8. Enter the S3 location of the input transactional data into the **Input data configuration** section. The input must be in the format `<id_col, date_col, monetary_col>`.

9. Enter the S3 output path in the **Output data configuration**.
10. Click **Create training job**. Upon successful creation of training job, the training job will be visible in the **Training jobs** list.

On completion of the job, you can check whether the model artifact is present in the output s3 location.

### 2.2. CREATE A MODEL PACKAGE

To create a model package, do the following:

1. Select the training job for which you want to create model package and click **Actions»Create model package**.
2. Enter the model package name and click **Next**. The **Validation specifications** page will open.

3. Select **No** option for both **Publish this model package on AWS Marketplace** and **Validate this resource** and then click **Create Model Package**.

4. Create a model using the new model package by selecting the model package you wanted to create a model and select **Actions> Create model**.
5. The model will appear on the **Models** list.

### 2.3. CREATING BATCH INFERENCE JOBS

To create a batch inference job, use the following steps:

1. Select the model for which you want to create an Inference job. The Model page opens.
2. Click **Batch transform**.
3. Enter the **Job name**.
4. Set the **Max payload size** to zero in **Additional configuration**.
5. Set the Input data content type to be **text/csv**.

6. Set the S3 location for Input data configuration.

7. Set the S3 location for Output configuration.
8. Click Create job. The job will appear in the Batch Transform page under Inferences. Once the job is complete, you will find the output in the output s3 location. The output csv will contain the <identifier column, customer lifetime value>.
Figure 11 Batch transform job list

<table>
<thead>
<tr>
<th>Name</th>
<th>Status</th>
<th>Duration</th>
<th>Creation Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>customer-lifetimevalue-demos-1-1</td>
<td>InProgress</td>
<td>a few seconds</td>
<td>Aug 24, 2021 15:31 UTC</td>
</tr>
<tr>
<td>customer-lifetimevalue-demos-1</td>
<td>Stopped</td>
<td>2 minutes</td>
<td>Aug 24, 2021 15:29 UTC</td>
</tr>
<tr>
<td>AlgorithmValidation-3e3df85b-2-12-6773-a332-23005f50-00b-1</td>
<td>Completed</td>
<td>5 minutes</td>
<td>Aug 24, 2021 11:07 UTC</td>
</tr>
<tr>
<td>AlgorithmValidation-80d20797-9050-4354-925c-101f9fe4e310e-1</td>
<td>Failed</td>
<td>5 minutes</td>
<td>Aug 24, 2021 11:11 UTC</td>
</tr>
<tr>
<td>AlgorithmValidation-24938035-50f6-4b3b-8b7c-9b28a24a827-1</td>
<td>Failed</td>
<td>5 minutes</td>
<td>Aug 24, 2021 11:15 UTC</td>
</tr>
<tr>
<td>AlgorithmValidation-76a6958a-7b4d-4f31-9ed7-624a627550c-1</td>
<td>Completed</td>
<td>5 minutes</td>
<td>Aug 24, 2021 10:06 UTC</td>
</tr>
<tr>
<td>cli-bkpapth-50</td>
<td>Completed</td>
<td>4 minutes</td>
<td>Aug 23, 2021 22:54 UTC</td>
</tr>
<tr>
<td>cli-bkpapth-29</td>
<td>Stopped</td>
<td>a minute</td>
<td>Aug 23, 2021 22:21 UTC</td>
</tr>
<tr>
<td>cli-bkpapth-27-test-1</td>
<td>Stopped</td>
<td>7 minutes</td>
<td>Aug 25, 2021 16:51 UTC</td>
</tr>
<tr>
<td>cli-bkpapth-27</td>
<td>Stopped</td>
<td>2 minutes</td>
<td>Aug 23, 2021 16:36 UTC</td>
</tr>
</tbody>
</table>
3. Data Preparation and Interpretation

3.1. DATA PREPARATION

Training data and inference data must be in comma separated values (csv) format. The column names and order of the columns should be consistent across both the training and inference data.

Sample notebooks for executing the training and inference is available in the following link: https://customer-life-score.s3.eu-west-1.amazonaws.com/sample-notebooks/CustomerLifeScore-Demo.ipynb

Sample training data is available in the following link: https://customer-life-score.s3.eu-west-1.amazonaws.com/sample-data/training.csv

Sample inference data is available in the following link: https://customer-life-score.s3.eu-west-1.amazonaws.com/sample-data/testing.csv

The input data should contain the three fields:

- **id_col**: Identifier column. For example, customer ID
- **date_col**: Date column. For instance, order_date, booking_date
- **monetary_col**: Price column. For instance, flight booking price, meal price

![Figure 12 Example input CSV data](image)

An important factor to consider while preparing training data is the **fwd_time_window**. The **fwd_time_window** represents the number of days to predict customer lifetime score. The number of days in **fwd_time_window** should be less than the number of days between overall first and last purchase of the training dataset. For instance, training sample containing 2 years of purchase history can have a **fwd_time_window** of 180 days (approx. 6 months). The training job will split the
data into calibration period (for example, 1.5 years) and observation period (for example, last 6 months). Therefore, customer life score process will predict the purchase probability and customer life score for next 6 months.

3.2. STEPS FOR EXECUTING SAMPLE NOTEBOOK IN SAGEMAKER

1. Open Sagemaker console and click on Notebook Instances under Notebook section
2. Create a new notebook instance or start an existing notebook instance
3. Open Jupyter and open a terminal session. In the terminal, type the following commands:
   `wget https://customer-life-score.s3.eu-west-1.amazonaws.com/sample-notebooks/CustomerLifeScore-Demo.ipynb`
4. Once the notebook is downloaded, open the sagemaker notebook from the jupyter file tab

3.3. OUTPUTS INTERPRETATION

The output of the process will contain customer life score and the probability of purchase for each customer. The customer life score is represented by the column: `estimated_spend_amount_in_forecasting_period`. The probability of purchase is represented by the column: `estimated_probability_of_purchase_in_forecasting_period`.

Besides the customer life score and purchase probability, additional columns are produced by the process to evaluate the use cases and to help the customer perform some advanced analytics.

The columns listed in the output are:

- customer_id
- number_of_purchases
- maximum_spend_per_purchase
- rank_maximum_spend_per_purchase
- average_spend_per_purchase
- total_spend
- rank_total_spend
- date_of_first_purchase
- date_of_last_purchase
- duration_between_first_and_last_purchase
- days_elapsed_since_last_purchase
- estimated_probability_of_purchase_in_forecasting_period
- estimated_spend_amount_in_forecasting_period
- rank_estimated_spend_amount_in_forecasting_period