Continuous Assurance Specialist Case Study 2

This case study aims to assess your ability to understand some basic concepts how to use the available technologies for machine learning model building and simple UI automation using Robotics Process Automation (RPA).

Pre-requisites

To be able to perform this case study, the following tools need to be installed on your laptop:

1. Machine learning development tool – either R or Python only
2. UiPath Community Edition

*Obtain installer from here:* [*https://www.uipath.com/start-trial*](https://www.uipath.com/start-trial)

Case Overview

In order to proactively managing the customer churn issues, the Sales Dept in a Telco has requested the BI department to develop a simple predictive analytics automation engine using either R or Python. In order to automate the process on continuous basis, it is preferred that the developed machine learning algorithms are executed automatically on monthly basis using the Robotics Process Automation tool UiPath.

Using the available data, the Sales Dept would like to build a model that is able to predict whether a customer is likely to churn or not.

The Raw Data

The raw data for the analysis is extracted by the IT Dept periodically and dumped into the following location (the current data contains the actual Churn column for model development purposes):

<https://maxis365.sharepoint.com/:f:/r/sites/InternalAudit/Public%20Documents/churn_data?csf=1&e=TopNtW>

(\*\*\* Please provide your email address to us in order to give you the access to the data)

Below are the column details of the raw data:

|  |  |  |
| --- | --- | --- |
| Column | Type | Description |
| customerID | Text | Customer ID |
| gender | Text | Whether the customer is a male or a female |
| SeniorCitizen | Number | Whether the customer is a senior citizen or not (1, 0) |
| Partner | Text | Whether the customer has a partner or not (Yes, No) |
| Dependents | Text | Whether the customer has dependents or not (Yes, No) |
| tenure | Number | Number of months the customer has stayed with the company |
| PhoneService | Text | Whether the customer has a phone service or not (Yes, No) |
| MultipleLines | Text | Whether the customer has multiple lines or not (Yes, No, No phone service) |
| InternetService | Text | Customer's internet service provider (DSL, Fiber optic, No) |
| OnlineSecurity | Text | Whether the customer has online security or not (Yes, No, No internet service) |
| OnlineBackup | Text | Whether the customer has online backup or not (Yes, No, No internet service) |
| DeviceProtection | Text | Whether the customer has device protection or not (Yes, No, No internet service) |
| TechSupport | Text | Whether the customer has tech support or not (Yes, No, No internet service) |
| StreamingTV | Text | Whether the customer has streaming TV or not (Yes, No, No internet service) |
| StreamingMovies | Text | Whether the customer has streaming movies or not (Yes, No, No internet service) |
| Contract | Text | The contract term of the customer (Month-to-month, One year, Two year) |
| PaperlessBilling | Text | Whether the customer has paperless billing or not (Yes, No) |
| PaymentMethod | Text | The customer's payment method (Electronic check, Mailed check, Bank transfer (automatic), Credit card (automatic)) |
| MonthlyCharges | Number | The amount charged to the customer monthly |
| TotalCharges | Number | The total amount charged to the customer |
| Churn | Text | Whether the customer churned or not (Yes or No) |

The Tasks

The automation for the analytics specified in the Case Overview is going to be performed using either R or Python combined with UiPath software for UI task automation. The following is the step-by-step guide to help you with the automation development for this case study:

1. **Step 1: Install the required tools**

* Ensure that R or Python is installed on your machine with the corresponding machine learning development libraries
* Ensure that UiPath (the free Community Edition) is installed and you are able to perform some automations using the tool

1. **Step 2: Develop the predictive analytics model using the tools**

* Using either R or Python, explore and develop a few applicable predictive analytics models and select one that could give the best performance
* For multiple algorithms/techniques explored, note and discuss the performance of each in terms of the applicable model evaluation metrics
* Select the best model evaluation metrics that you think is best in representing the model in this use case context

1. **Step 3: Automate the process using UiPath**

* Downloading the data, feeding it to the scripts and analysing the result can be tedious if need to be repeated many times on periodic basis. This can be automated using RPA.
* Using UiPath, develop an RPA automation flow that does the following:

1. Download the churn data from the URL
2. Run R/Python scripts to predict customer churn and export the results into a CSV file
3. Email the results to a recipient (you can use your own email for demo purpose)

Deliverables

Following are the deliverables expected for this case study:

1. The R Markdown or Jupyter Notebook file containing the steps of the various machine learning model development in step 2, which include the performance results discussion for each model
2. The UiPath xml file containing the flow for the automation

Please send to us the above files prior to the interview session.

You should be able to demonstrate the work done above during the interview session using the tools. For this, you need to bring your own laptop to the session.

See you soon!